

950/2019г.



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ



ОПЕРАТИВНА ПРОГРАМА
НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-С 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

Рег. №...../..... 2.

ВЪЗЛОЖИТЕЛ: УНИВЕРСИТЕТ ЗА НАЦИОНАЛНО И СВЕТОВНО СТОПАНСТВО
ИЗПЪЛНИТЕЛ: „Консорциум Биг Дейта Системс“ ДЗЗД

ДОГОВОР

№ 2011-79 / 2019 г.

Днес, 26.9.2019 г., в гр. София, между:

УНИВЕРСИТЕТ ЗА НАЦИОНАЛНО И СВЕТОВНО СТОПАНСТВО, с адрес: гр. София, п.к. 1700, Студентски град “Христо Ботев”, ул. “Осми декември”, ЕИК: 000670602, ИН по ДДС: BG000670602, представляван от проф. д.ик.н. Стати Статев - ректор на УНСС и Светослава Филчева-Иванова, директор на дирекция „Финанси“, наричан за краткост **ВЪЗЛОЖИТЕЛ**, от една страна,

и

„Консорциум Биг Дейта Системс“ ДЗЗД, със седалище и адрес на управление: гр.София - 1592, Хай Тех Бизнес Център, ж.к.“Дружба - 1“, ул.“Мюнхен“ № 8, ЕИК: 177357133, ИН по ДДС: BG177357133, представлявано от Деян Желев Йовков, в качеството на управител на основание Договор за консорциум от 25.04.2019 г., сключен между **„Секюрити Солюшънс Инститют“ ЕООД**, със седалище и адрес на управление: гр.София-1528, район Искър, ул.“Мюнхен“ № 8, ЕИК: 200149195, представлявано от Деян Желев Йовков, в качеството на управител и **„Перси“ООД**, със седалище и адрес на управление: гр.София- 1618, район Овча купел, бул.“Цар Борис III“ № 136б, представлявано от Евгений Георгиев Кърпачев, наричан за краткост **ИЗПЪЛНИТЕЛ**, от друга страна,

ВЪЗЛОЖИТЕЛЯТ и **ИЗПЪЛНИТЕЛЯТ**, наричани заедно „Страните“, а всеки от тях поотделно „Страна“,

след проведена открита процедура за избор на изпълнител на обществена поръчка по ЗОП с предмет: „Доставка, инсталиране, конфигуриране, тестване и поддържане на Специализирана централизирана система под ключ HADOOP за работа с големи данни по проект „Дигитализация на икономиката в среда на Големи данни“ (ДИГД), № BG05M2OP001-1.002-0002“, на основание Решение на РС № 21/08.04.2019 г. и Решение № 54/29.07.2019 г. на **ВЪЗЛОЖИТЕЛЯ** за определяне на **ИЗПЪЛНИТЕЛ**, и чл. 112, ал. 1 от ЗОП, се сключи настоящият договор за следното:

1



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ



ОПЕРАТИВНА ПРОГРАМА
НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-С 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

I. ПРЕДМЕТ НА ДОГОВОРА

Чл. 1. ВЪЗЛОЖИТЕЛЯТ възлага, а ИЗПЪЛНИТЕЛЯТ приема срещу възнаграждение и при условията на този Договор, да достави Специализирана централизирана система под ключ HADOOP (или еквивалент) за работа с големи данни по проект „Дигитализация на икономиката в среда на Големи данни“ (ДИГД), № BG05M2OP001-1.002-0002, като я инсталира, конфигурира, тества и поддържа, наричано за краткост „доставката“.

Чл. 2. ИЗПЪЛНИТЕЛЯТ се задължава да извърши доставката в съответствие с Техническата спецификация за доставки на ВЪЗЛОЖИТЕЛЯ, Техническото предложение на ИЗПЪЛНИТЕЛЯ и Ценовото предложение на ИЗПЪЛНИТЕЛЯ, съставляващи съответно Приложения №№ 1, 2 и 3 към този Договор /“Приложенията“/ и представляващи неразделна част от него. Доставката, инсталирането, конфигурирането и тестването се извършва от ИЗПЪЛНИТЕЛЯ чрез негов персонал и/или ръководен състав, съгласно Приложение № 4.

Чл. 3. В срок до 3 /три/ дни от датата на сключване на Договора, но най-късно преди започване на неговото изпълнение, ИЗПЪЛНИТЕЛЯТ уведомява ВЪЗЛОЖИТЕЛЯ за името, данните за контакт и представителите на подизпълнителите, посочени в офертата на ИЗПЪЛНИТЕЛЯ. ИЗПЪЛНИТЕЛЯТ уведомява ВЪЗЛОЖИТЕЛЯ за всякакви промени в предоставената информация в хода на изпълнението на Договора в срок до 3 /три/ дни от настъпване на съответното обстоятелство.

II. СРОК НА ДОГОВОРА. СРОК И МЯСТО НА ИЗПЪЛНЕНИЕ

Чл. 4. Договорът влиза в сила от датата на подписването му.

Чл. 5. (1) Срокът за изпълнение на доставката по чл. 2 е **3 (три) месеца**, считано от датата на подписване на договора.

(2) Гаранционното обслужване на доставката - Специализирана централизирана система под ключ HADOOP (или еквивалент) (хардуер, софтуер и други компоненти), е **5 (пет) години**, считано от дата на подписване на окончателния приемо-предавателен протокол за доставката, инсталирането, конфигурирането и тестването.

(3) Стойността на гаранционното обслужване, за сметка на ИЗПЪЛНИТЕЛЯ, на софтуера на Специализираната централизирана система под ключ HADOOP (или еквивалент), за срока на гаранцията ще е **0.00 лв. (нула лева и нула стотинки)** без ДДС. Стойността на гаранционното обслужване не се дължи от ВЪЗЛОЖИТЕЛЯ на ИЗПЪЛНИТЕЛЯ.

Чл. 6. Мястото на изпълнение на Договора е: Специализирано компютърно помещение в сградата на Университет за национално и световно стопанство, гр. София.

III. ЦЕНИ, РЕД И СРОКОВЕ НА ПЛАЩАНЕ

Чл. 7. (1) За доставката, ВЪЗЛОЖИТЕЛЯТ се задължава да плати на ИЗПЪЛНИТЕЛЯ обща цена в размер на **3 187 007.00 лв. /три милиона сто осемдесет и седем хиляди и седем лева/ без ДДС или 3 824 408.40 лв. /три милиона осемстотин двадесет и четири хиляди четиристотин и осем лева и четиридесет стотинки/ с ДДС, наричана по-нататък „Цената“ или „Стойността на Договора“, съгласно Ценовото предложение на ИЗПЪЛНИТЕЛЯ, съставляващо Приложение № 3.**

M .

2

шч

1 - / - - - - - /

-



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ



ОПЕРАТИВНА ПРОГРАМА
НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

(2) В Цената по ал. 1 са включени всички разходи на ИЗПЪЛНИТЕЛЯ за изпълнение на доставката, включително и разходите за персонала, който ще изпълнява поръчката и/или на членовете на ръководния състав, които ще отговарят за изпълнението и за неговите подизпълнители, като ВЪЗЛОЖИТЕЛЯТ не дължи заплащането на каквито и да е други разноси, направени от ИЗПЪЛНИТЕЛЯ.

(3) Цената, посочена в ал. 1 е фиксирана и крайна за времето на изпълнение на Договора и не подлежи на промяна, освен в случаите, изрично уговорени в този Договор и в съответствие с разпоредбите на ЗОП.

Чл. 8. (1) ВЪЗЛОЖИТЕЛЯТ плаща на ИЗПЪЛНИТЕЛЯ Цената по този Договор, както следва:

а) аванс в размер на 40 % (четиридесет на сто) от сумата по чл. 7, ал. 1, в срок до 10 дни след представяне гаранция за аванса и на фактура-оригинал от ИЗПЪЛНИТЕЛЯ, след датата на сключване на договора.

б) 30 % (тридесет на сто) от сумата по чл. 7, ал. 1, в срок до 30 дни след представяне на фактура-оригинал от ИЗПЪЛНИТЕЛЯ, при изпълнена доставка на всичките хардуерни, мрежови и системни софтуерни компоненти на системата и представяне на „Концептуална архитектура за разширение на системата за хиляди РВ“ и приемо-предавателен протокол за това.

в) останалите 30% (тридесет на сто) от сумата по чл. 7, ал. 1, в срок до 30 /тридесет/ дни след инсталиране, конфигуриране, тестване, изграждане на информационната сигурност, интегриране с MS SQL Server, и пускане в експлоатация на цялостната система под ключ и представяне на Окончателен приемо-предавателен протокол и фактура –оригинал.

(2) Във фактурите трябва да има текст „Разходът е по Договор за безвъзмездна помощ № BG05M2OP001-1.002-0002-C01, по оперативна програма „Наука и образование за интелигентен растеж“ 2014-2020.“

Чл. 9. (1) Плащането по чл. 8, ал. 1, б. „в“ от този Договор, се извършва въз основа на следните документи:

1. Окончателен приемо-предавателен протокол за приемане на доставката, инсталирането, конфигурирането и тестването на доставката по чл. 1 от Договора, подписан от ВЪЗЛОЖИТЕЛЯ и ИЗПЪЛНИТЕЛЯ, при съответно спазване на разпоредбите на Раздел VI /Предаване и приемане на изпълнението/ от Договора;

2. Фактура–оригинал, издадена от ИЗПЪЛНИТЕЛЯ и представена на ВЪЗЛОЖИТЕЛЯ.

Чл.10. (1) Всички плащания по този Договор се извършват в лева чрез банков превод по следната банкова сметка на ИЗПЪЛНИТЕЛЯ:

(2) ИЗПЪЛНИТЕЛЯТ е длъжен да уведомява писмено ВЪЗЛОЖИТЕЛЯ за всички последващи промени по ал. 1 в срок от 3 /три/ дни, считано от момента на промяната. В случай, че ИЗПЪЛНИТЕЛЯТ не уведоми Възложителя в този срок, счита се, че плащанията са надлежно извършени.

Чл. 11. Когато ИЗПЪЛНИТЕЛЯТ е сключил договор/и за подизпълнение, ВЪЗЛОЖИТЕЛЯТ извършва окончателно плащане към него, след като бъде представено копие на фактура–оригинал, че ИЗПЪЛНИТЕЛЯТ е заплатил на подизпълнителя за изпълнените от него работи.

IV. ГАРАНЦИЯ ЗА ИЗПЪЛНЕНИЕ



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ



ОПЕРАТИВНА ПРОГРАМА
НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

ГАРАНЦИЯ ЗА ИЗПЪЛНЕНИЕ

Чл. 12. (1) При подписването на този Договор, ИЗПЪЛНИТЕЛЯТ представя на ВЪЗЛОЖИТЕЛЯ гаранция за изпълнение в размер на 5 % (пет на сто) от Стойността на Договора без ДДС, а именно 159 350.35 лв. (сто петдесет и девет хиляди триста и петдесет лева и тридесет и пет стотинки) „Гаранцията за изпълнение“, която служи за обезпечаване на изпълнението на Договора. Страните се съгласяват, че Сума в размер на 79 675,18 (седемдесет и девет хиляди шестстотин седемдесет и пет лв. и 18 ст.), представляваща 2,5 % (две цяло и пет на сто) от Гаранцията за изпълнение, е предназначена по-конкретно за обезпечаване на гаранционното поддържане, предвидено в Договора.

Чл. 13. (1) В случай на изменение на Договора, извършено в съответствие с този Договор и приложимото право, ИЗПЪЛНИТЕЛЯТ се задължава да предприеме необходимите действия за привеждане на Гаранцията за изпълнение в съответствие с изменените условия на Договора, в срок до 10 (десет) дни от подписването на допълнително споразумение за изменението.

(2) Действията за привеждане на Гаранцията за изпълнение в съответствие с изменените условия на Договора могат да включват, по избор на ИЗПЪЛНИТЕЛЯ:

1. внасяне на допълнителна парична сума по банковата сметка на ВЪЗЛОЖИТЕЛЯ, при спазване на изискванията на чл. 15 от Договора; и/или;

2. предоставяне на документ за изменение на първоначалната банкова гаранция или нова банкова гаранция, при спазване на изискванията на чл. 16 от Договора; и/или

3. предоставяне на документ за изменение на първоначалната застраховка или нова застраховка, при спазване на изискванията на чл. 17 от Договора.

Чл. 14. Когато като Гаранция за изпълнение се представя парична сума, сумата се внася по следната банкова сметка на ВЪЗЛОЖИТЕЛЯ:

Банка: БНБ

BIC: BNBGBGSD

IBAN: BG03 BNBG 9661 3100 1746 01.

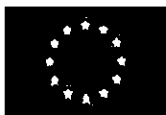
Чл. 15. (1) Когато като гаранция за изпълнение се представя банкова гаранция, ИЗПЪЛНИТЕЛЯТ предава на ВЪЗЛОЖИТЕЛЯ оригинален екземпляр на банкова гаранция, издадена в полза на ВЪЗЛОЖИТЕЛЯ, която трябва да отговаря на следните изисквания:

1. да бъде безусловна и неотменяема банкова гаранция, да съдържа задължение на банката - гарант да извърши плащане при първо писмено искане от ВЪЗЛОЖИТЕЛЯ, деклариращ, че е налице неизпълнение на задължение на ИЗПЪЛНИТЕЛЯ или друго основание за задържане на Гаранцията за изпълнение по този Договор;

2. да бъде със срок на валидност за целия срок на действие на Договора по чл. 4, плюс 30 (тридесет) дни след прекратяването на Договора, като при необходимост срокът на валидност на банковата гаранция се удължава или се издава нова.

(2) Банковите разходи по откриването и поддържането на Гаранцията за изпълнение във формата на банкова гаранция, както и по усвояването на средства от страна на ВЪЗЛОЖИТЕЛЯ, при наличието на основание за това, са за сметка на ИЗПЪЛНИТЕЛЯ.

Чл. 16. (1) Когато като Гаранция за изпълнение се представя застраховка, ИЗПЪЛНИТЕЛЯТ предава на ВЪЗЛОЖИТЕЛЯ оригинален екземпляр на застрахователна полица, издадена в полза на ВЪЗЛОЖИТЕЛЯ / в която ВЪЗЛОЖИТЕЛЯТ е посочен като трето ползващо се лице (бенефициер), която трябва да отговаря на следните изисквания:



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ



ОПЕРАТИВНА ПРОГРАМА
НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

1. да обезпечава изпълнението на този Договор чрез покритие на отговорността на ИЗПЪЛНИТЕЛЯ;

2. да бъде със срок на валидност за целия срок на действие на Договора по чл. 4, плюс 30 (тридесет) дни след прекратяването на Договора.

(2) Разходите по сключването на застрахователния договор и поддържането на валидността на застраховката за изисквания срок, както и по всяко изплащане на застрахователно обезщетение в полза на ВЪЗЛОЖИТЕЛЯ, при наличието на основание за това, са за сметка на ИЗПЪЛНИТЕЛЯ.

Чл. 17. (1) Възложителят освобождава гаранцията за изпълнение на договора, както следва:

а) 2,50 % (две цяло и пет на сто) - в срок от 30 (тридесет) календарни дни, след подписване на окончателният приемо-предавателен протокол и подадена молба от Изпълнителя, в случай, че не е налице основание за задържане съгласно този договор.

б) 2,50 % (две цяло и пет на сто) - в срок от 30 (тридесет) дни, след изтичане на срокът за гаранционно обслужване.

(2) Ако Изпълнителят е представил банкова гаранция за изпълнение на договора, преди частичното ѝ освобождаване следва да представи гаранция за изпълнение в остатъчния изискуем по договора размер на гаранцията след приспадане на сумата по ал. 1, буква „а“.

(3) Ако Изпълнителят е внесъл гаранцията за изпълнение на договора по банков път, Възложителят освобождава съответната част от нея в срока и при условията на ал. 1.

(4) Възложителят не дължи лихви върху сумите по предоставените гаранции, независимо от формата, под която са предоставени.

(5) Гаранцията или съответната част от нея не се освобождава от ВЪЗЛОЖИТЕЛЯ, ако в процеса на изпълнение на Договора е възникнал спор между Страните относно неизпълнение на задълженията на ИЗПЪЛНИТЕЛЯ и въпросът е отнесен за решаване пред съд. При решаване на спора в полза на ВЪЗЛОЖИТЕЛЯ той може да пристъпи към усвояване на гаранциите.

Чл. 18. ВЪЗЛОЖИТЕЛЯТ има право да задържи съответна част и да се удовлетвори от Гаранцията за изпълнение, когато ИЗПЪЛНИТЕЛЯТ не изпълни някое от неговите задължения по Договора, както и в случаите на лошо, частично и забавено изпълнение на което и да е задължение на ИЗПЪЛНИТЕЛЯ, като усвои такава част от Гаранцията за изпълнение, която съответства на уговорената в Договора неустойка за съответния случай на неизпълнение.

Чл. 19. ВЪЗЛОЖИТЕЛЯТ има право да задържи Гаранцията за изпълнение в пълен размер, в следните случаи:

1. ако ИЗПЪЛНИТЕЛЯТ не започне работа по изпълнение на Договора в срок до 15 (петнадесет) дни след Датата на получаване на аванса и ВЪЗЛОЖИТЕЛЯТ развали Договора на това основание;

2. при пълно неизпълнение, в т.ч. когато доставката не отговаря на изискванията на ВЪЗЛОЖИТЕЛЯ, и разваляне на Договора от страна на ВЪЗЛОЖИТЕЛЯ на това основание;

3. при прекратяване на дейността на ИЗПЪЛНИТЕЛЯ или при обявяването му в несъстоятелност.

Чл. 20. (1) Когато ВЪЗЛОЖИТЕЛЯТ се е удовлетворил от Гаранцията за изпълнение и Договорът продължава да е в сила, ИЗПЪЛНИТЕЛЯТ се задължава в срок до 10 (десет) дни да допълни Гаранцията за изпълнение, като внесе усвоената от ВЪЗЛОЖИТЕЛЯ сума по сметката на ВЪЗЛОЖИТЕЛЯ или предостави документ за изменение на първоначалната банкова гаранция или нова банкова гаранция, съответно застраховка, така че във всеки момент от действието на Договора размерът на Гаранцията за изпълнение да бъде в съответствие с чл. 12 от Договора.

(2) В всеки случай на задържане на Гаранцията за изпълнение, ВЪЗЛОЖИТЕЛЯТ уведомява



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ



ОПЕРАТИВНА ПРОГРАМА
НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

ИЗПЪЛНИТЕЛЯ за задържането и неговото основание. Задържането на Гаранцията за изпълнение изцяло или частично не изчерпва правата на ВЪЗЛОЖИТЕЛЯ да търси обезщетение в по-голям размер.

ГАРАНЦИЯ ЗА АВАНСОВО ПРЕДОСТАВЕНИ СРЕДСТВА

Чл. 21. (1) Гаранцията за авансово предоставените средства е в размер на тяхната стойност и се предоставя по избор на ИЗПЪЛНИТЕЛЯ в една от формите, посочени в ЗОП при спазване на изискванията на чл. 12-20 от настоящия договор и чл. 111 от ЗОП. Гаранцията се предоставя от ИЗПЪЛНИТЕЛЯ на ВЪЗЛОЖИТЕЛЯ при подписването на този договор.

(2) Гаранцията за авансово предоставените средства се освобождава в срок от 3 (три) календарни дни след:

1) представяне на ВЪЗЛОЖИТЕЛЯТ от ИЗПЪЛНИТЕЛЯТ на отчетни документи – (копия на фактури и други), че е усвоил аванса, чрез закупуване на хардуерни, мрежови и системни софтуерни компоненти на системата и приемо-предавателен протокол за доставката им на ВЪЗЛОЖИТЕЛЯ и предаването на тези документи на ВЪЗЛОЖИТЕЛЯ, подписан от представители на двете страни, или

2) възстановяване на аванса от ИЗПЪЛНИТЕЛЯ на ВЪЗЛОЖИТЕЛЯ.

V. ПРАВА И ЗАДЪЛЖЕНИЯ НА СТРАНИТЕ

Чл. 22. Изброяването на конкретни права и задължения на Страните в този раздел от Договора е неизчерпателно и не засяга действието на други клаузи от Договора или от приложимото право, предвиждащи права и/или задължения на която и да е от Страните.

ОБЩИ ПРАВА И ЗАДЪЛЖЕНИЯ НА ИЗПЪЛНИТЕЛЯ

Чл. 23. (1) ИЗПЪЛНИТЕЛЯТ има право:

1. да получи аванса и възнаграждение в размера, сроковете и при условията по чл.7-чл.11 от договора;

2. да иска и да получава от ВЪЗЛОЖИТЕЛЯ необходимото съдействие за изпълнение на задълженията по този Договор, както и всички необходими документи, информация и данни, пряко свързани или необходими за изпълнение на Договора.

(2) ИЗПЪЛНИТЕЛЯТ не носи отговорност в следните случаи:

- за повреди и отклонения от качеството на стоката, възникнали вследствие на нейната неправилна употреба, експлоатация, неспазване указанията в съпровождащата документация или неправилно съхранение от страна на Възложителя;

- поражения при стихийни бедствия: пожар, земетресение, наводнение, големи колебания в електрическата мрежа, посегателства и др. подобни.

Чл. 24. ИЗПЪЛНИТЕЛЯТ се задължава:

1. да извърши доставките, инсталирането, конфигурирането и тестването на хардуерните системи и да изпълнява задълженията си по този Договор в уговорения срок и качествено, в съответствие с договора и Приложенията;

2. да представи на ВЪЗЛОЖИТЕЛЯ разработките, съгласно Приложение № 1, и да извърши преработване и/или допълване в указания от ВЪЗЛОЖИТЕЛЯ срок, когато ВЪЗЛОЖИТЕЛЯТ е



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ



НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

поискал това;

3. да информира своевременно ВЪЗЛОЖИТЕЛЯ за всички пречки, възникващи в хода на изпълнението на работа, да предложи начин за отстраняването им, като може да поиска от ВЪЗЛОЖИТЕЛЯ указания и/или съдействие за отстраняването им;

4. да изпълнява всички законосъобразни указания и изисквания на ВЪЗЛОЖИТЕЛЯ;

5. да пази поверителна Конфиденциалната информация, в съответствие с уговореното в чл. 40 от Договора.

6. да осигури гаранционно обслужване и сервиз:

- гаранционното сервизно обслужване се извършва на мястото на ползване на доставката;
- ако е неизпълнимо горното условие – в сервиза, посочен от ИЗПЪЛНИТЕЛЯ: Сервиз на „Перси“ ООД в гр.София, ул.“Златна Добруджа“ № 18.

7. да извършва съответните гаранционни ремонти при следните условия:

а) Срокът за реакция при регистрация/уведомление от страна на Възложителя за хардуерна повредата да е: при подадена регистрация/уведомление в работно време - до 2 часа, и при подадена регистрация/уведомление в извънработно време – до началото на следващия работен ден, като реакцията е свързана с посещение на съответния специалист на място в УНСС;

б) Срокът за реакция при регистрация/уведомление от страна на Възложителя за повредата в инсталацията и/или конфигурацията на управляващия софтуер на Cloudera (или еквивалентен) на Nadoor (или еквивалентен) кълстера да е: при подадена регистрация/уведомление в работно време - до 2 часа, и при подадена регистрация/уведомление в извънработно време – до началото на следващия работен ден, като реакцията е свързана с посещение на сертифициран специалист или такъв с преминато обучение/положен изпит при производител/сертифициран център за обучение по „Cloudera Administrator“ (или еквивалентен), на място в УНСС.

в) Срокът за реакция при регистрация/уведомление от страна на Възложителя за повредата в информационната сигурност да е: при подадена регистрация/уведомление в работно време - до 2 часа, и при подадена регистрация/уведомление в извънработно време – до началото на следващия работен ден, като реакцията е свързана с посещение на съответния специалист, на място в УНСС;

г) Срокът за отстраняване на повредата е, както следва:

- За повреди в сървъри, дискови масиви и комуникационни устройства - до 2 работни дни;

- За повреди във всички останали хардуерни компоненти – до 14 работни дни;

- За повредата в инсталацията и/или конфигурацията на управляващия софтуер на Cloudera (или еквивалентен) – до 2 работни дни, когато повредата е в рамките на текущата версия на управляващия софтуер и до 14 дни, когато повредата е свързана с нова версия на управляващия софтуер;

- За повреда в информационната сигурност - до 2 работни дни, свързани със софтуерни и конфигурационни проблеми на информационната сигурност, и до 14 дни при проблеми свързани с хардуерни компоненти на информационната сигурност.

8. По време на първите 6 месеца от гаранционният срок, при установяване на еднороден дефект на компоненти и това е установено на не по-малко от 20% от техниката от един вид да подмени дефектния компонент и на другата техника от същият вид, които не са проявили дефекта;

9. при невъзможност за отстраняване на повреда в дефектиралото устройство в рамките на уговорените срокове, същото се заменя с работоспособно от същият или по-висок клас;

10. да не възлага работата или части от нея на подизпълнители, извън посочените в офертата на ИЗПЪЛНИТЕЛЯ, освен в случаите и при условията, предвидени в ЗОП;

∩

∩ - 7

С

∩ - 1



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ



ОПЕРАТИВНА ПРОГРАМА
НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

11. да участва във всички работни срещи, свързани с изпълнението на този Договор;
12. преди да предаде Специализираната централизирана система под ключ като готова система за използване, тя трябва да се тества за частична и пълна функционалност. За целта трябва да се проведе тестване на Специализирана централизирана система под ключ, което да се удостовери за всеки тестван компонент с приемо-предавателен протокол и съпътстващи документи съгласно Приложение № 1.
13. да изпълни договора с посоченият от него персонал притежаващ съответната професионална компетентност за изпълнението му.
14. да проведе следните видове курсове на обучение на колективи от заявителя на територията на УНСС:
 - Инсталиране и конфигуриране на Hadoop система (или еквивалентна);
 - Основи на програмирането в среда MapReduce (или еквивалентна);
 - Създаване на тестови данни за провеждане на изследвания.Всеки курс да бъде не по-кратък от 5 (пет) дни.
Броят участници от заявителя в един курс да бъде до 15 човека.
15. ИЗПЪЛНИТЕЛЯТ се задължава да сключи договор/договори за подизпълнение с посочените в офертата му подизпълнители в срок от 3 /три/ дни от сключване на настоящия Договор. В срок до 3 /три/ дни от сключването на договор за подизпълнение или на допълнително споразумение за замяна на посочен в офертата подизпълнител, ИЗПЪЛНИТЕЛЯТ изпраща копие на договора или на допълнителното споразумение на ВЪЗЛОЖИТЕЛЯ, заедно с доказателства, че са изпълнени условията по чл. 66, ал. 2 и 11 от ЗОП.
16. да представи, при подписването на Окончателният приемо-предавателен протокол, на ВЪЗЛОЖИТЕЛЯ Декларация, че техническите параметри на доставеното оборудване отговарят на офериранияте от него в техническата спецификация.

ОБЩИ ПРАВА И ЗАДЪЛЖЕНИЯ НА ВЪЗЛОЖИТЕЛЯ

Чл. 25. ВЪЗЛОЖИТЕЛЯТ има право:

1. да изисква и да получи доставките в уговорения срок, количество и качество, съгласно договора и Приложенията;
2. да контролира изпълнението на поетите от ИЗПЪЛНИТЕЛЯ задължения, в т.ч. да иска и да получава информация от ИЗПЪЛНИТЕЛЯ през целия Срок на Договора, или да извършва проверки, при необходимост и на мястото на изпълнение на Договора, но без с това да пречи на изпълнението;
3. да изисква, изпълнението на гаранционното обслужване и ремонти да се изпълняват в срок и в съответствие с условията на този договор;
4. да изисква от ИЗПЪЛНИТЕЛЯ поддържането на гаранциите, в съответствие с уговореното в Договора;
5. да не приеме някои от доставките, в съответствие с уговореното в чл. 28, ал. 1, т. 2 и т. 3 от Договора.
6. В случай, че ИЗПЪЛНИТЕЛЯТ не предприеме действия по отстраняването на повредите или замяната на дефектиралата техника в срок, ВЪЗЛОЖИТЕЛЯТ може да отстрани повредата за сметка на ИЗПЪЛНИТЕЛЯ. Всички разходи по поправката или замяната са за сметка на ИЗПЪЛНИТЕЛЯ и се удържат от гаранцията за добро изпълнение, ако размерът ѝ е недостатъчен, остатъкът – по съдебен ред, или цялата сума по съдебен ред, по преценка на ВЪЗЛОЖИТЕЛЯ.

11

6

11/11/11

✓



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ



ОПЕРАТИВНА ПРОГРАМА
НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

Чл. 26. ВЪЗЛОЖИТЕЛЯТ се задължава:

1. да приеме изпълнението на доставката, когато отговаря на договореното, по реда и при условията на този Договор;
2. да заплати на ИЗПЪЛНИТЕЛЯ Цената в размера, по реда и при условията, предвидени в този Договор;
3. да предостави и осигури достъп на ИЗПЪЛНИТЕЛЯ до информацията, необходима за извършването на доставката, предмет на Договора, при спазване на относимите изисквания или ограничения, съгласно приложимото право;
4. да пази поверителна Конфиденциалната информация, в съответствие с уговореното в чл. 40 от Договора;
5. да оказва съдействие на ИЗПЪЛНИТЕЛЯ във връзка с изпълнението на този Договор, включително и за отстраняване на възникнали пречки пред изпълнението на Договора, когато ИЗПЪЛНИТЕЛЯТ поиска това.

VI. ПРЕДАВАНЕ И ПРИЕМАНЕ НА ИЗПЪЛНЕНИЕТО

Чл. 27. (1) Предаването на изпълнението на различни части от доставката по чл. 1 се документира с протоколи за приемане и предаване, заедно с приложени доказателства, съгласно Приложенията, които се подписват от представители на ВЪЗЛОЖИТЕЛЯ и ИЗПЪЛНИТЕЛЯ в два оригинални екземпляра – по един за всяка от Страните /“Приемо-предавателен протокол“/. Приемо-предавателните протоколи се съставят и съдържат документите, съобразно изискванията на Приложение № 1.

(2) Доставката по чл. 1 се изпълнява от ИЗПЪЛНИТЕЛЯ през целия срок на изпълнение на договора.

(3) Приемането на работи от ВЪЗЛОЖИТЕЛЯ, за които е сключен договор за подизпълнение, се извършва в присъствието на ИЗПЪЛНИТЕЛЯ и подизпълнителя.

(4) При наличие на подизпълнител, Приемо-предавателният протокол се подписва от ВЪЗЛОЖИТЕЛЯ, ИЗПЪЛНИТЕЛЯ и от подизпълнителя.

(5) Независимо от ползването на подизпълнители, отговорността за изпълнение на договора е на ИЗПЪЛНИТЕЛЯ.

Чл. 28. (1) ВЪЗЛОЖИТЕЛЯТ има право:

1. да приеме изпълнението, когато отговаря на договореното;
2. когато бъдат установени несъответствия на изпълненото с уговореното или бъдат констатирани недостатъци, ВЪЗЛОЖИТЕЛЯТ може да откаже приемане на изпълнението до отстраняване на недостатъците, като даде подходящ срок за отстраняването им за сметка на ИЗПЪЛНИТЕЛЯ;
3. да откаже да приеме изпълнението при съществени отклонения от договореното и/или в случай, че констатирани недостатъци са от такова естество, че не могат да бъдат отстранени в рамките на срока за изпълнение на Договора или резултатът от изпълнението става безполезен за ВЪЗЛОЖИТЕЛЯ;

(2) Окончателното приемане на изпълнението на доставката по този Договор се извършва с подписване на Окончателен Приемо-предавателен протокол, с включени в него приемо-предавателните протоколи за изпълнението заедно с приложени доказателства, съгласно Приложенията, подписан от страните в срок до 3 /три/ дни след изтичането на срока на изпълнение

11 .

11/11/1



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ



ОПЕРАТИВНА ПРОГРАМА
НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

по чл. 5, ал. 1 от Договора. В случай, че към този момент бъдат констатирани недостатъци в изпълнението, те се описват в окончателния Приемо-предавателен протокол и се определя подходящ срок за отстраняването им или налагането на санкция, съгласно чл. 29 - чл. 32 от Договора.

(3) При подписването на Окончателният приемо-предавателен протокол ИЗПЪЛНИТЕЛЯТ представя на ВЪЗЛОЖИТЕЛЯ Декларация, че техническите параметри на доставеното оборудване отговарят на офериранията от него в техническото предложение.

VII. САНКЦИИ ПРИ НЕИЗПЪЛНЕНИЕ

Чл. 29. При просрочване изпълнението на задълженията по този Договор, неизправната Страна дължи на изправната, неустойка в размер на 0,5% (нула цяло и пет на сто) от Цената за всеки ден забава, но не повече от 50 % (петдесет на сто) от стойността на неизпълнението.

Чл. 30. При констатирано лошо или друго неточно или частично изпълнение на отделна дейност или при отклонение от изискванията на ВЪЗЛОЖИТЕЛЯ, посочени в Техническата спецификация, ВЪЗЛОЖИТЕЛЯТ има право да поиска от ИЗПЪЛНИТЕЛЯ да изпълни изцяло и качествено съответната дейност, без да дължи допълнително възнаграждение за това. В случай, че и повторното изпълнение на доставката или част от нея е некачествено, ВЪЗЛОЖИТЕЛЯТ има право да прекрати договора.

чл. 31. при разваляне на договора поради неизпълнение на някоя от страните, неизправната страна дължи на изправната неустойка в размер на 50% (петдесет на сто) от стойността на неизпълнението

Чл. 32. Плащането на неустойките, уговорени в този Договор, не ограничава правото на изправната Страна да търси реално изпълнение и/или обезщетение за понесени вреди и пропуснати ползи в по-голям размер, съгласно приложимото право.

VIII. ПРЕКРАТЯВАНЕ НА ДОГОВОРА

Чл. 33. (1) Този Договор се прекратява:

1. с изтичане на Срока на Договора по чл. 5 от Договора;
2. с изпълнението на всички задължения на Страните по него;
3. при настъпване на пълна обективна невъзможност за изпълнение, за което обстоятелство засегнатата Страна е длъжна да уведоми другата Страна в срок до 3 /три/ дни от настъпване на невъзможността и да представи доказателства;
4. при прекратяване на юридическо лице – Страна по Договора без правопримство, по смисъла на законодателството на държавата, в която съответното лице е установено;
5. при условията по чл. 5, ал. 1, т. 3 от ЗИФОДРЮПДРКЛТДС;

(2) Договорът може да бъде прекратен:

1. по взаимно съгласие на Страните, изразено в писмена форма;
2. когато за ИЗПЪЛНИТЕЛЯ бъде открито производство по несъстоятелност или ликвидация – по искане на ВЪЗЛОЖИТЕЛЯ;

Чл. 34. (1) Всяка от Страните може да развали Договора при виновно неизпълнение на съществено задължение на другата страна по Договора, при условията и с последиците, съгласно чл. 87 и сл. от Закона за задълженията и договорите, чрез отправяне на писмено предупреждение от изправната Страна до неизправната и определяне на подходящ срок за изпълнение. Разваляне на Договора не се



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ



ОПЕРАТИВНА ПРОГРАМА
НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

допуска, когато неизпълнената част от задължението е незначителна с оглед на интереса на изправната Страна.

(2) За целите на този Договор, Страните ще считат за виновно неизпълнение на съществено задължение на ИЗПЪЛНИТЕЛЯ всеки от следните случаи:

1. когато ИЗПЪЛНИТЕЛЯТ не е започнал изпълнението на доставката в срок до 15 /петнадесет/ дни, считано от Датата на получаване на заявката за доставката;
2. ИЗПЪЛНИТЕЛЯТ е прекратил изпълнението на доставката за повече от 3 /три/ дни;
3. ИЗПЪЛНИТЕЛЯТ е допуснал съществено отклонение от Техническата спецификация и Техническото предложение;

(3) ВЪЗЛОЖИТЕЛЯТ може да развали Договора само с писмено уведомление до ИЗПЪЛНИТЕЛЯ и без да му даде допълнителен срок за изпълнение, ако поради забава на ИЗПЪЛНИТЕЛЯ, то е станало безполезно или ако задължението е трябвало да се изпълни непременно в уговореното време.

Чл. 35. ВЪЗЛОЖИТЕЛЯТ прекратява Договора в случаите по чл.118, ал.1 от ЗОП, без да дължи обезщетение на ИЗПЪЛНИТЕЛЯ за претърпени от прекратяването на Договора вреди, освен ако прекратяването е на основание чл. 118, ал. 1, т. 1 от ЗОП. В последния случай, размерът на обезщетението се определя в протокол или споразумение, подписано от Страните, а при непостигане на съгласие – по реда на клаузата за разрешаване на спорове по този Договор.

Чл. 36. Във всички случаи на прекратяване на Договора, освен при прекратяване на юридическо лице – Страна по Договора без правоприемство:

1. ВЪЗЛОЖИТЕЛЯТ и ИЗПЪЛНИТЕЛЯТ съставят констативен протокол за извършената към момента на прекратяване работа и размера на евентуално дължимите плащания;

2. ИЗПЪЛНИТЕЛЯТ се задължава:

а/ да преустанови предоставянето на доставката, с изключение на такива дейности, каквито може да бъдат необходими и поискани от ВЪЗЛОЖИТЕЛЯ;

б/ да предаде на ВЪЗЛОЖИТЕЛЯ всички отчети/разработки/материали, изготвени от него в изпълнение на Договора до датата на прекратяването; и

в/ да върне на ВЪЗЛОЖИТЕЛЯ всички документи и материали, които са собственост на ВЪЗЛОЖИТЕЛЯ и са били предоставени на ИЗПЪЛНИТЕЛЯ във връзка с предмета на Договора.

Чл. 37. При предсрочно прекратяване на Договора, ВЪЗЛОЖИТЕЛЯТ е длъжен да заплати на ИЗПЪЛНИТЕЛЯ реално изпълнението и приета по установения ред доставка.

IX. ОБЩИ РАЗПОРЕДБИ

Дефинирани понятия и тълкуване

Чл. 38. (1) Освен ако са дефинирани изрично по друг начин в този Договор, използваните в него понятия имат значението, дадено им в ЗОП, съответно в легалните дефиниции в Допълнителните разпоредби на ЗОП или, ако няма такива за някои понятия – според значението, което им се придава в основните разпоредби на ЗОП.

(2) При противоречие между различни разпоредби или условия, съдържащи се в Договора и Приложенията, се прилагат следните правила:

1. специалните разпоредби имат предимство пред общите разпоредби;
2. разпоредбите на Приложенията имат предимство пред разпоредбите на Договора;

Л

С

11/11/11

11/11/11

11/11/11

11/11/11



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ



ОПЕРАТИВНА ПРОГРАМА
НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

Спазване на приложими норми

Чл. 39. При изпълнението на Договора, ИЗПЪЛНИТЕЛЯТ и неговите подизпълнители са длъжни да спазват всички приложими нормативни актове, разпоредби, стандарти и други изисквания, свързани с предмета на Договора, и в частност, всички приложими правила и изисквания, свързани с опазване на околната среда, социалното и трудовото право, приложими колективни споразумения и/или разпоредби на международното екологично, социално и трудово право, съгласно Приложение № 10 към чл. 115 от ЗОП.

Конфиденциалност

Чл. 40. (1) Всяка от Страните по този Договор се задължава да пази в поверителност и да не разкрива или разпространява информация за другата Страна, станала ѝ известна при или по повод изпълнението на Договора („Конфиденциална информация“). Конфиденциална информация включва, без да се ограничава до: обстоятелства, свързани с търговската дейност, техническите процеси, проекти или финанси на Страните, както и ноу-хау, изобретения, полезни модели или други права от подобен характер, свързани с изпълнението на Договора.

Конфиденциална информация включва, без да се ограничава до: всякаква финансова, търговска, техническа или друга информация, анализи, съставени материали, изследвания, документи или други материали, свързани с бизнеса, управлението или дейността на другата Страна, от каквото и да е естество или в каквато и да е форма, включително, финансови и оперативни резултати, пазари, настоящи или потенциални клиенти, собственост, методи на работа, персонал, договори, ангажименти, правни въпроси или стратегии, продукти, процеси, свързани с документация, чертежи, спецификации, диаграми, планове, уведомления, данни, образци, модели, мостри, софтуер, софтуерни приложения, компютърни устройства или други материали или записи или друга информация, независимо дали в писмен или устен вид, или съдържаща се на компютърен диск или друго устройство.

(2) С изключение на случаите, посочени в ал. 3 на този член, Конфиденциална информация може да бъде разкривана само след предварително писмено одобрение от другата Страна, като това съгласие не може да бъде отказано безпричинно.

(3) Не се счита за нарушение на задълженията за неразкриване на Конфиденциална информация, когато:

1. информацията е станала или става публично достъпна, без нарушаване на този Договор от която и да е от Страните;
2. информацията се изисква по силата на закон, приложим спрямо която и да е от Страните;
- или
3. предоставянето на информацията се изисква от регулаторен или друг компетентен орган и съответната Страна е длъжна да изпълни такова изискване;

В случаите по точки 2 или 3, Страната, която следва да предостави информацията, уведомява незабавно другата Страна по Договора.

(4) Задълженията по тази клауза се отнасят до съответната Страна, всички нейни поделения, контролирани от нея фирми и организации, всички нейни служители и наети от нея физически или юридически лица, като съответната Страна отговаря за изпълнението на тези задължения от страна на такива лица.

M .



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ



ОПЕРАТИВНА ПРОГРАМА
НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

Задълженията, свързани с неразкриване на Конфиденциалната информация остават в сила и след прекратяване на Договора на каквото и да е основание.

Публични изявления

Чл. 41. ИЗПЪЛНИТЕЛЯТ няма право да дава публични изявления и съобщения, да разкрива или разгласява каквато и да е информация, която е получил във връзка с извършване на Услугите, предмет на този Договор, независимо дали е въз основа на данни и материали на ВЪЗЛОЖИТЕЛЯ или на резултати от работата на ИЗПЪЛНИТЕЛЯ, без предварителното писмено съгласие на ВЪЗЛОЖИТЕЛЯ, което съгласие няма да бъде безпричинно отказано или забавено.

Авторски права

Чл. 42. (1) Страните се съгласяват, на основание чл. 42, ал. 1 от Закона за авторското право и сродните му права, че авторските права върху всички документи и материали, и всякакви други елементи или компоненти, създадени в резултат на или във връзка с изпълнението на Договора, принадлежат изцяло на ВЪЗЛОЖИТЕЛЯ в същия обем, в който биха принадлежали на автора. ИЗПЪЛНИТЕЛЯТ декларира и гарантира, че трети лица не притежават права върху изготвените документи и други резултати от изпълнението на Договора, които могат да бъдат обект на авторско право.

(2) В случай че бъде установено с влязло в сила съдебно решение или в случай че ВЪЗЛОЖИТЕЛЯТ и/или ИЗПЪЛНИТЕЛЯТ установят, че с изготвянето, въвеждането и използването на документи или други материали, съставени при изпълнението на този Договор, е нарушено авторско право на трето лице, ИЗПЪЛНИТЕЛЯТ се задължава да направи възможно за ВЪЗЛОЖИТЕЛЯ използването им:

1. чрез промяна на съответния документ или материал; или
2. чрез замяната на елемент от него със защитени авторски права с друг елемент със същата функция, който не нарушава авторските права на трети лица; или
3. като получи за своя сметка разрешение за ползване на продукта от третото лице, чиито права са нарушени.

(3) ВЪЗЛОЖИТЕЛЯТ уведомява ИЗПЪЛНИТЕЛЯ за претенциите за нарушени авторски права от страна на трети лица в срок до 7 /седем/ дни от узнаването им. В случай, че трети лица предявят основателни претенции, ИЗПЪЛНИТЕЛЯТ носи пълната отговорност и понася всички щети, произтичащи от това. ВЪЗЛОЖИТЕЛЯТ привлича ИЗПЪЛНИТЕЛЯ в евентуален спор за нарушено авторско право във връзка с изпълнението по Договора.

(4) ИЗПЪЛНИТЕЛЯТ заплаща на ВЪЗЛОЖИТЕЛЯ обезщетение за претърпените вреди и пропуснатите ползи вследствие на окончателно признато нарушение на авторски права на трети лица.

Прехвърляне на права и задължения

Чл. 43. Никоя от Страните няма право да прехвърля никое от правата и задълженията, произтичащи от този Договор, без съгласието на другата Страна. Паричните вземания по Договора и по договорите за подизпълнение могат да бъдат прехвърляни или залагани, съгласно приложимото право.

11



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ



ОПЕРАТИВНА ПРОГРАМА
НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

Изменения

Чл. 44. Този Договор може да бъде изменен само с допълнителни споразумения, изготвени в писмена форма и подписани от двете Страни, в съответствие с изискванията и ограниченията на ЗОП.

Непреодолима сила

Чл. 45. (1) Страните не отговарят за неизпълнение на задължение по този Договор, когато невъзможността за изпълнение се дължи на непреодолима сила.

(2) За целите на този Договор, „непреодолима сила“ има значението на това понятие по смисъла на чл.306, ал.2 от Търговския закон.

(3) Страната, засегната от непреодолима сила, е длъжна да предприеме всички разумни усилия и мерки, за да намали до минимум понесените вреди и загуби, както и да уведоми писмено другата Страна в срок до 3 /три/ дни от настъпване на непреодолимата сила. Към уведомлението се прилагат всички релевантни и/или нормативно установени доказателства за настъпването и естеството на непреодолимата сила, причинната връзка между това обстоятелство и невъзможността за изпълнение, и очакваното времетраене на неизпълнението.

(4) Докато трае непреодолимата сила, изпълнението на задължението се спира. Засегнатата Страна е длъжна, след съгласуване с насрещната Страна, да продължи да изпълнява тази част от задълженията си, които не са възпрепятствани от непреодолимата сила.

(5) Не може да се позовава на непреодолима сила Страна:

1. която е била в забава или друго неизпълнение преди настъпването на непреодолима сила;
2. която не е информирала другата Страна за настъпването на непреодолима сила; или
3. чиято небрежност или умишлени действия или бездействия са довели до невъзможност за изпълнение на Договора;

(6) Липсата на парични средства не представлява непреодолима сила.

Нищожност на отделни клаузи

Чл. 46. В случай, че някоя от клаузите на този Договор е недействителна или неприложима, това не засяга останалите клаузи. Недействителната или неприложима клауза се замества от повелителна правна норма, ако има такава.

Уведомления

Чл. 47. (1) Всички уведомления между Страните във връзка с този Договор се извършват в писмена форма и могат да се предават лично или чрез препоръчано писмо, по куриер, по факс, електронна поща.

(2) За целите на този Договор данните и лицата за контакт на Страните са, както следва:

1. За ВЪЗЛОЖИТЕЛЯ:

Адрес за кореспонденция: УНСС, гр. София, Студентски град “Христо Ботев”, ул. “Осми декември”

M

S

— 1000

1



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ



ОПЕРАТИВНА ПРОГРАМА
НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

Тел.: (02) 8195 506

Кабинет: ИТ-04

e-mail: mimi@unwe.bg

Лице за контакт: Мария Иванова – ръководител отдел „Техническа поддръжка“, ДИТ

2. За ИЗПЪЛНИТЕЛЯ:

Адрес за кореспонденция: гр.София - 1592, Хай Тех Бизнес Център, ж.к.“Дружба - 1“, ул.“Мюнхен“ № 8

Тел.: 0887 776 797, (+359 2) 421 01 04; (+359 2) 421 01 00

Факс: 02/979 8008

e-mail: deyan.yovkov@ssi-bg.com

Лице за контакт: Деян Йовков - управител

(3) За дата на уведомлението се счита:

1. датата на предаването – при лично предаване на уведомлението;
2. датата на пощенското клеймо на обратната разписка – при изпращане по пощата;
3. датата на доставка, отбелязана върху куриерската разписка – при изпращане по куриер;
3. датата на приемането – при изпращане по факс;
4. датата на получаване – при изпращане по електронна поща;

(4) Всяка кореспонденция между Страните ще се счита за валидна, ако е изпратена на посочените по-горе адреси (в т.ч. електронни), чрез посочените по-горе средства за комуникация и на посочените лица за контакт. При промяна на посочените адреси, телефони и други данни за контакт, съответната Страна е длъжна да уведоми другата в писмен вид в срок до 3 /три/ дни от настъпване на промяната. При неизпълнение на това задължение всяко уведомление ще се счита за валидно връчено, ако е изпратено на посочените по-горе адреси, чрез описаните средства за комуникация и на посочените лица за контакт.

(5) При преобразуване без прекратяване, промяна на наименованието, правноорганизационната форма, седалището, адреса на управление, предмета на дейност, срока на съществуване, органите на управление и представителство на ИЗПЪЛНИТЕЛЯ, същият се задължава да уведоми ВЪЗЛОЖИТЕЛЯ за промяната в срок до 3 /три/ дни от вписването ѝ в съответния регистър.

Приложимо право

Чл. 48. Този Договор, в т.ч. Приложенията към него, както и всички произтичащи или свързани с него споразумения, и всички свързани с тях права и задължения, ще бъдат подчинени на и ще се тълкуват съгласно българското право.

Чл. 49. За неуредените в този Договор въпроси се прилагат разпоредбите на действащото българско законодателство.

Разрешаване на спорове

Чл. 50. Всички спорове, породени от този Договор или отнасящи се до него, включително споровете, породени или отнасящи се до неговото тълкуване, недействителност, изпълнение или прекратяване, както и споровете за попълване на празноти в Договора или приспособяването му към

11 .

15

C

Иванова

Йовков



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ



ОПЕРАТИВНА ПРОГРАМА
НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

нововъзникнали обстоятелства, ще се уреждат между Страните чрез преговори, а при непостигане на съгласие – спорът ще се отнася за решаване от компетентния български съд.

Екземпляри

Чл. 51. Този Договор се състои от 16 /шестнадесет/ страници и е изготвен и подписан в 3 /три/ еднообразни екземпляра – от които два за ВЪЗЛОЖИТЕЛЯ и един за ИЗПЪЛНИТЕЛЯ.

Приложения:

Чл. 52. Към този Договор се прилагат и са неразделна част от него следните приложения:

1. Приложение № 1 – Техническа спецификация за доставка на ВЪЗЛОЖИТЕЛЯ;
2. Приложение № 2 – Техническо предложение на ИЗПЪЛНИТЕЛЯ;
3. Приложение № 3 – Ценово предложение на ИЗПЪЛНИТЕЛЯ;
4. Приложение № 4 – Списък на персонала и/или ръководният състав, изпълняващ поръчката

ВЪЗЛОЖИТЕЛ:
УНСС

РЕКТОР:

ПРОФ. ДИКН. СТАТИСТАРЕВ

ДИРЕКТОР ДИРЕКЦИЯ „ФИНАНСИ“

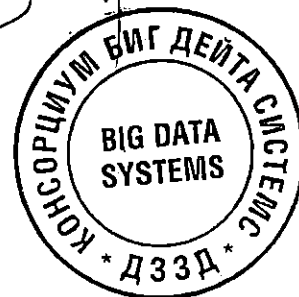
СВЕТОСЛАВА ФИЛЧЕВА-ИВАНОВА

ИЗПЪЛНИТЕЛ:

„Консорциум Биг Дейта Системс“ ДЗЗД

УПРАВИТЕЛ:

ДЕЯН ЙОВКОВ





ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ



НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-С 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

РАЗДЕЛ II. ТЕХНИЧЕСКА СПЕЦИФИКАЦИЯ

1. ВЪВЕДЕНИЕ

Доставка, инсталиране, конфигуриране, тестване и поддържане на СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ HADOOP за работа с Големи данни по проект „Дигитализация на икономиката в среда на Големи данни“ (ДИГД), № BG05M2OP001-1.002-0002.

2. ОПИСАНИЕ НА ПРЕДМЕТА НА ОБОСОБЕНАТА ПОЗИЦИЯ:

Специализираната централизирана система под ключ (turnkey system) Hadoop е предназначена да осигурява възможност за работа с Големи данни – структурирани, полу-структурирани и неструктурирани данни. От функционална гледна точка, Hadoop е колекция от програми с отворен код изпълнявани върху клъстерна съвкупност от сървъри / компютри с общо предназначение. Той се използва главно за обработка на различни форми на структурирани, полу-структурирани и неструктурирани данни, който дава голяма гъвкавост при събирането, обработката и анализа на данни, отколкото релационните бази данни и складовете за данни. От хардуерна гледна точка, Hadoop системата е изградена като клъстер, съставен от стандартни сървъри / компютри. От софтуерна гледна точка Hadoop системата се състои от следните основни софтуерни модули: Hadoop Common - съдържа библиотеки и помощни програми, необходими на други модули на Hadoop; Hadoop Distributed File System (HDFS) - разпределена файлова система, която съхранява данни върху стандартни компютри; Hadoop YARN - отговорна за управлението на изчислителните ресурси в клъстера и използването им за планиране и управление изпълнението на приложенията на потребителите; Hadoop MapReduce - прилагаща програмния модел MapReduce за мащабна обработка на данни.

Специализираната централизирана система под ключ Hadoop трябва да се достави, да се инсталира, да се извърши на нея съответното конфигуриране и интегриране, и да се тества за доказване както на цялостната работа на системата (изисква се предлаганата система да е „система по ключ“), така и на отделни нейни компоненти. Специализираната централизирана система под ключ Hadoop е предназначена да бъде изградена като едно цялостно съоръжение от хардуерни компоненти (интегрирани ракове, сървъри, устройства за информационна сигурност, мрежови устройства, окабеляване и пр.) и софтуерни компоненти (инсталирани операционни системи и

системни софтуерни елементи), работещи като един комплектен продукт, която система да допуска развитие и разширение в бъдеще. Тази система под ключ има за предназначение да обслужва оперирането на големи данни, прилагайки архитектурата с отворен код „Hadoop“. Специализираната централизирана система под ключ Hadoop трябва да бъде управлявана чрез софтуерната система Cloudera минимална версия 6.0, работеща със софтуерната рамка Hadoop минимална версия 3.0, за да се оформи логически като едно цялостно изделие готово за използване след предаването му.

Специализираната система под ключ трябва да бъде конфигурирана и проведено с нея съответното администриране, като тя се предостави на заявителя във вид на цялостно изградена и готова за опериране и непосредствено използване система. Имайки предвид, че Специализираната система под ключ ще предоставя достъп на много потребители и потребителски организации, то информационната сигурност на системата е от съществено значение. Всички посочени компоненти на системата, заедно с тяхното LAN и токово окабеляване, трябва да бъдат изградени физически в отделни шкафове (Ракове) с размери 42U. Връзката на отделните ракове следва да бъде обезпечена с минимум 2 броя 20 метрови LAN оптически кабели със скорост за обмен минимум по 40Gbps, а връзката между компонентите, монтирани във всеки рак трябва да е обезпечена с минимум 2 броя LAN кабели със скорост на обмен минимум по 10Gbps за обмен на данни и 1Gbps за управление. При използване, всички LAN портове със скорост за обмен 10Gbps и 40Gbps на устройствата (изчислителни и комуникационни модули) трябва да бъдат със съответно SFP+ / QSFP+ конектори, като за съответните свързвания да се доставят и съответните видове кабели. Специализираната централизирана система под ключ трябва да бъде тествана и предадена на заявителя за опериране като една цялостна и готова за работа система.

Предложената система под ключ трябва да е с минимум 5 години гаранция.

Специализираната централизирана система под ключ Hadoop трябва да бъде изградена в специализирано компютърно помещение, посочено от заявителя, намиращо се в сградата на УНСС София.

Специализираната централизирана система под ключ Hadoop трябва да има следните 7 елемента:

- i. Специализирани хардуерни системи;
- ii. Специализирано системно софтуерно осигуряване;
- iii. Специализирана информационна сигурност;
- iv. Интегриране с Майкрософт SQL Server минимум версия 2017;
- v. Прототипна под-система Hadoop за управление тестването на прототипни решения за големи данни;
- vi. Система за тестване пълна функционалност на Специализирана централизирана система под ключ;
- vii. Концептуална архитектура за разширение на системата Hadoop за хиляди PetaBytes.

Елементите i-iv включително създават т.н. „Централизирана Hadoop система“ в първи Hadoop клъстер, представляващ ядрото на Специализираната централизирана система под ключ Hadoop. „Прототипната под-система Hadoop за управление тестването на прототипни решения за големи данни“, представляваща елемент v, ще се създава върху втори Hadoop клъстер, която система трябва да представлява от една страна отделна Hadoop клъстерна конфигурация и от друга - да бъде част от Специализираната Централизирана система под ключ Hadoop. По този начин, Специализираната централизирана система под ключ Hadoop трябва да има в състава си „Централизирана Hadoop система“ и „Прототипна под-система Hadoop за управление тестването на прототипни решения за големи данни“. Тази Прототипна под-система Hadoop има за цел да служи за специализирани прототипни изследвания (тестване и изследвания със силно ограничен набор от данни и силно ограничено използване на изчислителна мощност), като да може да работи автономно от другите компоненти на Специализираната централизирана система под ключ Hadoop. По този начин, в един момент от функционирането на Специализираната централизирана система под ключ Hadoop, може да работи само „Прототипната под-система Hadoop за управление тестването на прототипни решения за големи данни“, във втори момент – да работи Специализираната централизирана система под ключ Hadoop, без да се използва Прототипната под-система Hadoop (т.е. да работи само „Централизирана Hadoop система“), а в трети момент - да работи цялостно Специализираната централизирана система под ключ Hadoop, като съвместно работи „Централизирана Hadoop система“ и „Прототипната под-система Hadoop“.

Целта на „Концептуалната архитектура за разширение на системата Hadoop за хиляди PetaBytes“ на Специализираната централизирана система под ключ Hadoop е да докаже, че предложените хардуерни, програмни, конфигурационни и интеграционни решения, заедно с решенията за информационна сигурност, са създадени като основа без ограничения за потенциално бъдещо интензивно разширение на системата до капацитет стотици и хиляди PetaBytes (без да се отчита участието на Прототипната под-система Hadoop за управление тестването на прототипни решения за големи данни).

2. Специализирани хардуерни системи

Специализираната централизирана система под ключ Hadoop трябва да се състои от 4 типа базови компютърни възела (node) – Възел за данни (DataNode), Възел за имена (NameNode), Управляващ възел (Management Node) и Граничен възел (EdgeNode). Всеки от тези възела може да се състои от един или няколко изчислителни модула, свързани чрез мрежови модули – комутатори (switches). Всичките тези модули трябва да се разположат в 4 броя шкафове (Ракове) – Рак №10, Рак №11, Рак №12 и Рак №13 (при доставка на по-големи модули за DataNode, броят на раковете може да се увеличи с още един - Рак №13-А). Във всеки рак ще се разположи отделна специализирана хардуерна система. По този начин, от хардуерна гледна точка, Специализираната централизирана система под ключ ще се състои от 5 броя специализирани хардуерни системи, всяка

изградена в един от посочените 4 рака. За всяка специализирана хардуерна система, състояща се от изчислителни и/или мрежови модули, се изисква:

- да бъде изградена чрез съответната форма на структурно окабеляване съответстващо на изискваната скорост за обмен;
- да бъдат извършени съответните конфигурационни, интеграционни и административни действия за осъществяване на нейната работоспособност;
- да бъде тествана за функциониране чрез демонстриране на обмен на данни между всяко устройство и останалите устройства в рамката на хардуерната система в рака, така също и обмен на данни между устройство от един рак и устройства от другите ракове, като тестването се представи на заявителя чрез приемо-предавателен протокол.

Тези 5 броя специализирани хардуерни системи трябва да бъдат:

- i. Базова Специализирана хардуерна система за обработка на данни;
- ii. Разширяема Специализирана хардуерна система за обработка на данни;
- iii. Управляваща Специализирана хардуерна система за управление обработка на данни;
- iv. Комуникационна Специализирана хардуерна система за обмен на данни;
- v. Прототипна хардуерна под-система Hadoop за управление тестването на прототипни решения за големи данни

2.1. Общи принципи заложи в Специализираната централизирана система под ключ Hadoop

Специализираната централизирана система под ключ Hadoop, всеки DataNode трябва да се разположи в един рак като цялостно изделие, поради което ще бъде наричан Раково ориентиран възел за данни (RackDataNode). Раково ориентираният възел за данни трябва да се състои минимум от 6 изчислителни модула за данни (МодулDataNode) от гледна точка на ускорен обмен и репликиране на големи файлове, на организация за опростено техническо обслужване и на по-висока надеждност на механично движещите се части. Всеки Модул за данни (МодулDataNode) трябва да има минимум 12 позиции за включване на 3,5" дискове със сменяемост по време на опериране (hot-swap 3,5" drive bays). Дисковете, които се поставят в позициите за включване на дискове в Модула за данни (МодулDataNode) трябва да имат минимален капацитет 10 TB, с възможност в бъдеще да се включват дискове с няколко десетки TB. Изхождайки от общите изисквания на Cloudera за обем дисково пространство на DataNode и принципите на Hadoop 3.0, всеки Модул за данни (МодулDataNode) трябва да има мрежова LAN връзка с обобщен капацитет за обмен (изграден чрез една или няколко физически LAN връзки) минимум 20Gbps (при използване на няколко LAN физически връзки за този обобщен капацитет, е необходимо те да се конфигурират като една обединена логическа LAN връзка, като всяка единична физическа връзка да се състои не по-малко от 10Gbps). От гледна точка потенциалното развитие в бъдеще на Специализираната система за използване на много големи обеми от данни (стотици и повече PB), то Модулите за данни (МодулDataNode) физически не трябва да бъдат по-големи от 2U, за предпочитане 1U (единицата "U" е за

фронтална височина на устройство разположено в рак, наречена форм фактор – 1U е около 1,75 инча или около 44.45 милиметра). Модулите за данни (МодулDataNode) трябва да не бъдат с функционалност като на устройства за съхранение на данни или като на устройства само за изчислителни операции. От гледна точка създаване на Балансирана Специализираната централизирана система под ключ Hadoop, обслужваща множество различни по характер приложения, и с възможност за разширение с ниска по стойност и използване на общо разпространени компютри като модули за данни, то дисковете на изчислителния модул на възела за данни (МодулDataNode) не трябва да са оформени физически в отделни самостоятелни запаметяващи устройства, например SAN устройства, а да бъдат свързани в изчислителния модул като локални дискове. Тези изчислителни модули трябва да бъдат с достатъчно изявени изчислителни възможности, използващи минимум 2 броя процесори от тип Intel Xeon Scalable Processors, всеки от тях с минимум 8 ядра, работещи с честота минимум 2GHz и използващи оперативна памет минимум 128 GB. По този начин изчислителният модулът за възела за данни (МодулDataNode) е изчислителен и запаметяващ модул едновременно. Този подход създава възможност за еднотипност в бъдещо разширение на Специализираната централизирана система под ключ Hadoop с произволна скалабилност, опериращи със сървъри с общо предназначение. Специализираната централизирана система под ключ Hadoop трябва се изгради от 6 броя DataNode, т.е. от 6 броя Раково ориентирани възела за данни (PakDataNode), което кореспондира на съответен брой Модули за данни (МодулDataNode). Възелът за имена (NameNode) поради своята функционалност трябва да се състои от 1 брой изчислителен модул - Модул за имена (МодулNameNode), т.е. NameNode и МодулNameNode представляват едно и също нещо. Всеки МодулNameNode трябва да има мрежова LAN връзка с обобщен капацитет за предаване на данни (изграден чрез една или няколко LAN връзки) минимум 20Gbps (при използване на няколко LAN връзки за този обобщен капацитет, е необходимо те да се конфигурират като една обединена логическа LAN връзка, всяка състояща се от поне 10Gbps). От гледна точка спецификите на Hadoop версия 3.0 и следващи версии, изчислителните възможности на МодулNameNode трябва да са увеличени, поради което модулът МодулNameNode трябва да има минимум 2 броя процесори от тип Intel Xeon Scalable Processors, всеки от тях с минимум 8 ядра, работещи с честота минимум 1,7GHz и използващи оперативна памет минимум 128 GB. Модулът за имена (МодулNameNode) трябва да има минимум 8 позиции за включване на 2,5" или 3,5" дискове със сменяемост по време на опериране (hot-swap drive bays). Дисковете, които се поставят в позициите за включване на дискове в Модула за имена (МодулNameNode) трябва да имат минимален капацитет 2 TB. С оглед осигуряване на необходимата надеждност, Специализираната централизирана система под ключ Hadoop трябва се изгради от 3 броя NameNode, т.е. от 3 броя модули МодулNameNode. Граничният възел - EdgeNode обикновено действа като портал за свързване с крайните потребители, за да могат крайните потребители по-организирано да достигат до възлите за данни на Hadoop системата. Неговата цел е да действа като интерфейс (gateway) за Hadoop клъстера и външния свят. Той може да изпълнява и ролята на компютър за потребителски програми. Граничният възел (EdgeNode) поради своята функционалност трябва да се състои от 1 брой изчислителен модул – Граничен модул (МодулEdgeNode), т.е. EdgeNode и МодулEdgeNode представляват едно и също нещо.

Всеки МодулEdgeNode трябва да има минимум 2 броя процесори от тип Intel Xeon Scalable Processors, всеки от тях с минимум 8 ядра, работещи с честота минимум 1,7GHz и използващи оперативна памет минимум 128GB. Граничният модул (МодулEdgeNode) трябва да има минимум 8 позиции за включване на 2,5" или 3,5" дискове със сменяемост по време на опериране (hot-swap drive bays). Дисковете, които се поставят в позициите за включване на дискове в Граничния модул (МодулEdgeNode) трябва да имат минимален капацитет 2 TB.

За управление работата на цялостната Hadoop клъстерна система, е необходимо Специализираната централизирана система под ключ Hadoop да има възел за управление - ManagementNode. Възелът за управление (ManagementNode) поради своята функционалност трябва да се състои от 1 брой изчислителен модул - Модул за управление (МодулManagementNode), т.е. ManagementNode и МодулManagementNode представляват едно и също нещо. Всеки МодулManagementNode трябва да има мрежова LAN връзка с обобщен капацитет за предаване на данни (изграден чрез една или няколко LAN връзки) минимум 20Gbps (при използване на няколко LAN връзки за този обобщен капацитет, е необходимо те да се конфигурират като една обединена логическа LAN връзка, всяка състояща се от поне 10Gbps). Модулът за управление МодулManagementNode трябва да има минимум 2 броя процесори от тип Intel Xeon Scalable Processors, всеки от тях с минимум 8 ядра, работещи с честота минимум 1,7GHz и използващи оперативна памет минимум 128GB. Модулът за управление (МодулManagementNode) трябва да има минимум 8 позиции за включване на 2,5" или 3,5" дискове със сменяемост по време на опериране (hot-swap drive bays). Дисковете, които се поставят в позициите за включване на дискове в Модула за управление (МодулManagementNode) трябва да имат минимален капацитет 2 TB. Специализираната централизирана система под ключ Hadoop трябва се изгради от 1 брой ManagementNode, т.е. от 1 брой модул МодулManagementNode.

За Прототипната под-система Hadoop за управление тестването на прототипни решения за големи данни, Възелст за данни (DataNode) трябва да се състои от отделен компютър (ComputerDataNode), като се използват 3 броя такива компютъра, а Възелът за имена (NameNode) трябва да се състои от отделен компютър (ComputerNameMgmtNode). В този компютър трябва да се зареди и софтуерът за Възела за управление (ManagementNode). Тези 4 компютъра, заедно с тяхното мрежово свързване, от хардуерна гледна точка, ще се предоставят от Заявителя. Заявителят трябва да достави за Прототипната под-система от хардуерна гледна точка, само 1 брой LAN мрежова платка с 10Gbps скорост за обмен на данни за модула ComputerNameMgmtNode (трите броя ComputerDataNode съдържат в себе си LAN мрежова платка с 10Gbps скорост за обмен на данни). Доставчикът трябва да изгради от тези 4 броя модули – 3 броя ComputerDataNode и 1 брой ComputerNameMgmtNode, Прототипната под-система Hadoop за управление тестването на прототипни решения за големи данни. Цялата Прототипна под-система Hadoop ще трябва да се свърже към Комуникационната Специализирана хардуерна система за обмен на данни на Специализираната централизирана система под ключ Hadoop (Пак №10), чрез 10Gps връзки. Ангажимент на

Доставчикът е да инсталира и конфигурира Прототипната под-система Hadoop като автономно работеща Hadoop система.

2.2. Базова Специализирана хардуерна система за обработка на данни

Тази система трябва да бъде изградена в отделен rack – Rack №12.

2.2.1. Изчислителни модули на възел за данни на Базовата Специализирана хардуерна система за обработка на данни

Базовата Специализирана хардуерна система за обработка на данни трябва да се състои от 2 броя DataNode – RackDataNode, всеки от които да се състои от минимум 6 броя Модули за данни (МодулDataNode), т.е. Базовата Специализирана хардуерна система за обработка на данни трябва да се състои общо от минимум 12 броя Модули за данни (МодулDataNode). Всеки Модул за данни (МодулDataNode) трябва да се състои от:

- минимум 2 броя процесори от тип Intel Xeon Bronze Scalable Processors, всеки от тях с минимум 8 ядра, работещи с честота минимум 2GHz;
- оперативна памет тип ECC с минимум 128GB с честота на работа минимум 2600 MHz;
- минимум 12 позиции за включване на 3,5" дискове за потребителски данни със сменяемост по време на опериране (hot-swap 3,5" drive bays), с включени в тях 3,5" дискови устройства с минимален капацитет 10TB работещи със стандартизираната дискова архитектура SAS 3.0 12.0Gb/s, която архитектура се поддържа от изчислителния модул, организирани в архитектура JBOD;
- минимум 2 броя дискови устройства за системния софтуер на изчислителния модул, всеки от тях с минимален обем 0,5TB, организирани за работа в RAID1, осигурявана от съответен контролер за RAID1 организация на дискове в изчислителния модул;
- минимум 2 порта за LAN Ethernet със скорост за обмен 1Gbps и интерфейс RJ45;
- минимум 2 порта за LAN Ethernet със скорост за обмен 10Gbps и интерфейс SFP+;
- софтуер за управление на хардуера на модула;
- минимум 2 броя интерфейси PCIe 3.0 x8 или PCIe 3.0 x16;
- минимум удвоено захранване;
- Физическа организация – за инсталиране в rack, с размери не по-големи от 2U, за предпочитане 1U.

2.2.2. Комуникационни модули – комутатори (switches) на Базовата Специализирана хардуерна система за обработка на данни

А) Основен комуникационен модул на Базовата Специализирана хардуерна система за обработка на данни – 1 брой. Той да бъде управляем комутатор на минимално ниво 2 от OSI Референтната архитектурата, със следното минимално количество портове:

- 48 порта за Ethernet 10Gbps скорост за обмен с интерфейс SFP+;

- 2 порта uplink връзки за Ethernet 40Gbps скорост за обмен с интерфейс QSFP+;
- 2 порта за Ethernet 1Gbps скорост за обмен с интерфейс RJ45;

Б) Допълнителен комуникационен модул на Базовата Специализирана хардуерна система за обработка на данни – 1 брой. Той да бъде управляем комутатор на минимално ниво 2 от OSI Референтната архитектурата, със следното минимално количество портове:

- 24 порта за Ethernet 1Gbps скорост за обмен с интерфейс RJ45;
- 1 порт за връзка с по-висока йерархия комуникационни модули за Ethernet 1-10Gbps скорост за обмен с интерфейс SFP+.

2.2.3. Физическа организация на Базова Специализирана хардуерна система за обработка на данни

Базова Специализирана хардуерна система за обработка на данни трябва да бъде монтирана и инсталирана за работа в рак с размери 42U – рак №12, който да се достави. От този рак да се осигурят минимум 2 броя кабели за скорост на обмен 40Gbps, с дължина минимум 20 метра, чрез които да се осигури връзка с Комутатор с портове от по 40Gbps, разположен в Рак №10. Рак №12 трябва да бъде със съответното управление – KVM switch за управление на минимум 12 сървъра с дисплей, клавиатура и мишка и кабели за управление на 12 компютъра. Тези дисплей, клавиатура и мишка да са в размерите сумарно на 1U.

Всички изчислителни модули и комуникационни модули на Базова Специализирана хардуерна система за обработка на данни да бъдат монтирани в посочения рак №12. Изчислителните модули да бъдат свързани с Основния комуникационен модул с необходимите LAN кабели за 10Gbps скорост за обмен с интерфейс SFP+. Допълнителният комуникационен модул да бъде свързан към Основния комуникационен модул. При инсталирането в рака, комуникационните модули да се поставят отгоре в рака, след които да се поставят изчислителните модули. Да се доставят минимум 12 броя LAN кабели за 1Gbps скорост за обмен с интерфейс RJ45 за връзки в рамките на рака. Всички LAN кабели да се наименоуват за това с кои устройства се прави връзка, и се подредят и организират в рака с оглед улеснено техническо обслужване. Да се доставят 12 разделителни носещи плочи за поставяне в рак, върху които да се поставят евентуални tower-based сървъри и компютри.

Ракът да бъде доставен с UPS система. UPS да бъде с размери до 4U, като доставчикът да определи неговите технически параметри за да се осигурява токоподаване на всички устройства в рака минимум 3 минути след спиране на електричеството.

2.3. Разширяема Специализирана хардуерна система за обработка на данни

Тази система трябва да бъде изградена в отделен рак – Рак №13, а по изключение (при Модул за данни (МодулDataNode) по-голям от 1U) – и в Рак №13-А.

Разширяемата Специализирана хардуерна система за обработка на данни ще служи като основа за бъдещо разширение на Специализираната хардуерна система за обработка на данни – в рамките на един рак и/или в рамките на множество такива ракове.

2.3.1. Изчислителни модули на възел за данни на Разширяемата Специализирана хардуерна система за обработка на данни

Разширяемата Специализирана хардуерна система за обработка на данни трябва да се състои от 4 броя DataNode – РакDataNode, всеки от които да се състои от минимум 6 броя Модули за данни (МодулDataNode), т.е. Разширяемата Специализирана хардуерна система за обработка на данни трябва да се състои минимум общо от 24 броя Модули за данни (МодулDataNode). Всеки Модул за данни (МодулDataNode) трябва да се състои от:

- минимум 2 броя процесори от тип Intel Xeon Bronze Scalable Processors, всеки от тях с минимум 8 ядра, работещи с честота минимум 2GHz;
- оперативна памет минимум 128GB с честота на работа минимум 2600 MHz;
- минимум 12 позиции за включване на 3,5" дискове за потребителски данни със сменяемост по време на опериране (hot-swap 3,5" drive bays), с включени в тях 3,5" дискови устройства с минимален капацитет 10TB работещи със стандартизираната дискова архитектура SAS 3.0 12.0Gb/s, която архитектура се поддържа от изчислителния модул, организирани в архитектура JBOD;
- минимум 2 броя дискови устройства с технология SSD за системния софтуер на изчислителния модул, всеки от тях с минимален обем 0,5TB, организирани за работа в RAID1, осигурявана от съответен контролер за RAID1 организация на дискове в изчислителния модул;
- минимум 2 порта за LAN Ethernet със скорост за обмен 1Gbps и интерфейс RJ45;
- минимум 2 порта за LAN Ethernet със скорост за обмен 10Gbps и интерфейс SFP+;
- софтуер за управление на хардуера на модула;
- минимум 2 броя интерфейси PCIe 3.0 x8 или PCIe 3.0 x16;
- минимум удвоено захранване;
- Физическа организация – за инсталиране в рак, с размери не по-големи от 2U, за предпочитане 1U. При избор на Модула за данни (МодулDataNode) който е по-голям от 1U, то да се достави допълнителен рак със съответното управление, който да се обозначи като Рак №13-А.

2.3.2. Комуникационни модули - комутатори (switches) на Разширяемата Специализирана хардуерна система за обработка на данни

А) Основен комуникационен модул на Разширяемата Специализирана хардуерна система за обработка на данни – 3 броя. Те да бъдат управляеми комутатори на минимално ниво 2 от OSI Референтната архитектурата, със следното минимално количество портове:

- 48 порта за Ethernet 10Gbps скорост за обмен с интерфейс SFP+;
- 2 порта uplink връзки за Ethernet 40Gbps скорост за обмен с интерфейс QSFP+;
- 2 порта за Ethernet 1Gbps скорост за обмен с интерфейс RJ45;

При използване на Модула за данни (МодулDataNode) който е по-голям от 1U, то в Рак №13 да се поставят два от тези комуникационни модула, а в Рак №13-А – третият комуникационен модул.

Б) Допълнителен комуникационен модул на Разширяемата Специализирана хардуерна система за обработка на данни – 1 брой. При използване на Рак №13-А този допълнителен комуникационен модул се поставя там. Този комуникационен модул да бъде управляем комутатор на минимално ниво 2 от OSI Референтната архитектурата, със следното минимално количество портове:

- 24 порта за Ethernet 1Gbps скорост за обмен с интерфейс RJ45;
- 1 порт за връзка с по-висока йерархия комуникационни модули за Ethernet 1-10Gbps скорост за обмен с интерфейс SFP+.

2.3.3. Физическа организация на Разширяемата Специализирана хардуерна система за обработка на данни

Разширяемата Специализирана хардуерна система за обработка на данни трябва да бъде монтирана и инсталирана за работа в рак с размери 42U – Рак №13 (или съответно и Рак №13-А), който (които) да се достави. При използване само на Рак №13 (използва се Модула за данни (МодулDataNode) който е 1U), то от този рак да се осигурят 6 броя кабели за скорост на обмен 40Gbps, с дължина минимум 20 метра, чрез които да се осигури връзка с Комутатор с портове от по 40Gbps, разположен в Рак №10. При използване на Рак №13 и Рак №13-А (използва се Модула за данни (МодулDataNode) който е по-голям от 1U), то от Рак №13 да се осигурят 4 броя кабели за скорост на обмен 40Gbps, с дължина минимум 20 метра, чрез които да се осигури връзка с Комутатор с портове от по 40Gbps, разположен в Рак №10, а от Рак №13-А да се осигурят 2 броя кабели за скорост на обмен 40Gbps, с дължина минимум 20 метра, чрез които да се осигури връзка с Комутатор с портове от по 40Gbps, разположен в Рак №10. Рак №13-А трябва да бъде със съответното управление – KVM switch за управление на 12 сървъра с дисплей, клавиатура и мишка и кабели за управление на 12 компютъра. Тези дисплей, клавиатура и мишка да са в размерите сумарно на 1U.

Всички изчислителни модули и комуникационни модули на Разширяемата Специализирана хардуерна система за обработка на данни да бъдат монтирани в посочения рак (ракове). Изчислителните модули да бъдат свързани с Основните комуникационни модули с необходимите LAN кабели за 10Gbps скорост за обмен с интерфейс SFP+. Допълнителният комуникационен модул да бъде свързан към един от Основните комуникационни модули. При инсталирането в рака, комуникационните модули да се поставят отгоре в рака, след които да се поставят изчислителните модули.

Да се доставят минимално 24 броя LAN кабели за 1Gbps скорост за обмен с интерфейс RJ45. Всички LAN кабели да се наименоват за това с кои устройства се прави връзка, и се подредят и организират в рака с оглед улеснено техническо обслужване. За всеки рак да се доставят 12 разделителни носещи плочи за поставяне в рак, върху които да се поставят евентуални tower-based сървъри и компютри.

При прилагане на Модула за данни (МодулDataNode) който е по-голям от 1U, т.е. използва се РАК №13 и Рак №13-А, то в Рак №13 се поставят 18 броя Модула за данни (МодулDataNode), а в Рак №13-А се поставят 6 броя Модула за данни (МодулDataNode).

Ракът (раковете при използване на Рак №13-А) да бъде доставен с UPS система. UPS да бъде с размери до 4U, като доставчикът да определи неговите технически параметри за да се осигурява токоподаване на всички устройства в рака минимум 3 минути след спиране на електричеството.

2.4. Управляваща Специализирана хардуерна система за управление обработка на данни

Тази система трябва да бъде изградена в отделен рак – Рак №11.

Предназначението на Управляващата Специализирана хардуерна система е да извършва управление на цялостната Специализирана система за обработка на данни, както и да управлява основното функциониране на Информационната сигурност.

Управляващата Специализирана хардуерна система трябва да се състои от следните 6 броя изчислителни модули: 3 броя Модули за имена (МодулNameNode), 1 брой Модул за управление (МодулManagementNode), 1 брой Сървър за УАЛС (сървър за Управление и Анализ на Логове и Събития - за управление и анализ на всички логове и настъпващи събития в Специализираната система под ключ) и 1 брой Сървър за МУД (сървър за Мрежово Управление и Достъп – за управление на мрежовите комутатори и изпълняваните от тях управления на достъпа в Специализираната система под ключ). Изчислителните модули 3 броя Модули за имена (МодулNameNode) и 1 брой Модул за управление (МодулManagementNode) са свързани с управление работата на Hadoop средата, докато изчислителните модули Сървър за УАЛС и Сървър за МУД са свързани с управлението на Информационната сигурност на Специализираната система под ключ. Управляващата Специализирана хардуерна система трябва да се състои също и от комуникационни модули (switches): Основен комуникационен модул и Допълнителен комуникационен модул.

2.4.1. Изчислителни модули на Управляващата Специализирана хардуерна система

А) Изчислителни модули свързани с управление работата на Hadoop средата

Всеки от 4-те изчислителните модули - 3 броя Модули за имена (МодулNameNode) и 1 брой Модул за управление (МодулManagementNode), трябва да се състоят от:

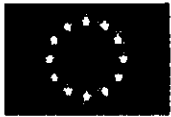
- минимум 2 броя процесори от тип Intel Xeon Bronze Scalable Processors, всеки от тях с минимум 8 ядра, работещи с честота минимум 1,7GHz;
- оперативна памет минимум 128GB с честота на работа минимум 2600 MHz;
- минимум 8 позиции за включване на 2,5" или 3,5" дискове със сменяемост по време на опериране (hot-swap drive bays) с инсталирани дискове с минимум 2.0TB с архитектура SAS 3.0 12.0Gb/s със съответен SAS-RAID контролер, работещи в архитектура RAID 1/5/6, които да служат за системен софтуер и за потребителски данни;
- оптично устройство DVD-RW / 24;
- минимум 2 порта за LAN Ethernet със скорост за обмен 1Gbps и интерфейс RJ45;
- минимум 2 порта за LAN Ethernet със скорост за обмен 10Gbps и интерфейс SFP+;
- елемент за backup на захранване при RAID контролера;
- минимум удвоено захранване;
- софтуер за управление на хардуера на модула;
- Физическа организация – за инсталиране в рак, с размери не по-големи от 2U, за предпочитане 1U.

Б) Изчислителен модул „Сървър за УАЛС“ (Сървър за управление и анализ на логове и събития)

Изчислителният модул Сървър за УАЛС трябва да има следните минимални технически параметри:

- минимум 2 броя процесори от тип Intel Xeon Bronze Scalable Processors, всеки от тях с минимум 8 ядра, работещи с честота минимум 1,7GHz;
- оперативна памет тип ECC с минимум 128GB с честота на работа минимум 2600 MHz;
- минимум 12 позиции за включване на 3,5" дискове за потребителски данни със сменяемост по време на опериране (hot-swap 3,5" drive bays), с включени в 4 от тях 3,5" дискови устройства с минимален капацитет 10TB работещи със стандартизираната дискова архитектура SAS 3.0 12.0Gb/s, която архитектура се поддържа от изчислителния модул, организирани в архитектура RAID10;
- минимум 2 броя дискови устройства за системния софтуер на изчислителния модул, всеки от тях с минимален обем 480GB, организирани за работа в RAID1, осигурявана от съответен контролер за RAID1 организация на дискове в изчислителния модул;
- минимум 2 порта за LAN Ethernet със скорост за обмен 1Gbps и интерфейс RJ45;
- минимум 2 порта за LAN Ethernet със скорост за обмен 10Gbps и интерфейс SFP+;
- софтуер за управление на хардуера на модула;
- минимум 2 броя интерфейси PCIe 3.0 x16;
- Физическа организация – за инсталиране в рак, с размери не по-големи от 2U, за предпочитане 1U.

В този изчислителен модул трябва да се зареди системен софтуер, който да извършва функциониране по управление на LOG файловете на отделните устройства (изчислителни и комуникационни модули) на Специализираната система под ключ Hadoop и възникналите събития свързани с Информационната система в реално време.



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН СЪЮЗ
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ



НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

Образец №2

ТЕХНИЧЕСКО ПРЕДЛОЖЕНИЕ

До:

УНИВЕРСИТЕТ ЗА НАЦИОНАЛНО И СВЕТОВНО СТОПАНСТВО

Техническо предложение за изпълнение на обществена поръчка с предмет: Доставка, инсталиране, конфигуриране, тестване и поддържане на СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ NADOOP за работа с Големи данни по проект „Дигитализация на икономиката в среда на Големи данни“ (ДИГД), № BG05M2OP001-1.002-0002

От:

Участник: „Консорциум Биг Дейта Системс“ ДЗЗД

ЕИК/БУЛСТАТ: ЕИК: 177357133

Седалище и адрес на управление: ж.к. Дружба 1, ул. Мюнхен 8

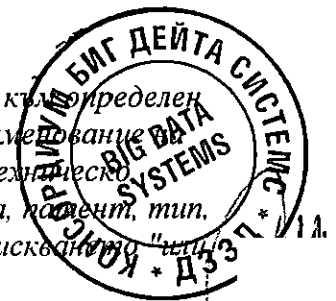
представявано от: Деян Йовков, в качеството му на Управител

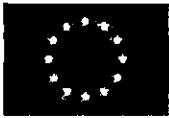
УВАЖАЕМИ ГОСПОЖИ И ГОСПОДА,

С настоящото Ви представяме нашето предложение за изпълнение на обявената обществената поръчка като се задължаваме да спазваме всички условия на възложителя, посочени в документацията за участие, които се отнасят до изпълнението на поръчката.

В случай че същата ни бъде възложена след като подробно се запознахме с изискванията на възложителя за участие в обществената поръчка с предмет:

В настоящата документацията, не се съдържат изисквания, насочващи към определен производител или доставчик. В случай, че наименованието или част от наименованието съвпада с конкретен стандарт, спецификация, техническа оценка, техническо одобрение, технически еталон и модел, източник, процес, търговска марка, материал, тип, произход или производство, да се приема, че възложителят е поставил изискванията еквивалентно/и". 000022





ЕВРОПЕЙСКИ СЪЮ
 ЕВРОПЕЙСКИ
 СОЦИАЛЕН ФОНД
 ЕВРОПЕЙСКИ ФОНД
 ЗА РЕГИОНАЛНО РАЗВИТИЕ



НАУКА И ОБРАЗОВАНИЕ ЗА
 ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

Доставка, инсталиране, конфигуриране, тестване и поддържане на СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ HADOOP за работа с Големи данни по проект „Дигитализация на икономиката в среда на Големи данни“ (ДИГД), № BG05M2OP001-1.002-0002

предлагаме:

При изпълнението на обществената поръчка няма да ползваме/ще ползваме (вярното се подчертава) следните подизпълнители:

1.
2.

(наименование на подизпълнителя и дела от поръчка който ще изпълни);

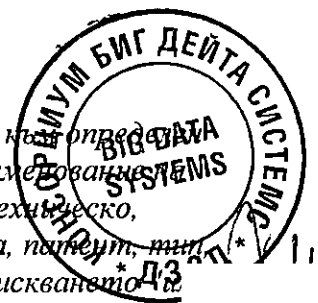
Съгласно чл. 66, ал. 1 от ЗОП представяме **декларация (ПО ОБРАЗЕЦ)** от всеки подизпълнител, че е поел задължение да изпълни посочения по-горе дял от поръчката и (други документи, подписани от подизпълнителите ако е приложимо);

Ако бъдем определени за потенциален изпълнител, ще сключим договора в законоустановения срок.

Подаването на настоящата оферта и предложение удостоверява безусловното приемане на всички изисквания и задължения, поставени от Възложителя в провежданата процедура.

Поемаме ангажимент да изпълним дейностите по предмета на обществената поръчка в съответствие с изискванията Ви, посочени в Техническата спецификация на настоящата поръчка, изискванията на Възложителя в документацията за участие и всички приложими нормативни актове, както и в съответствие с предложената от нас оферта.

В настоящата документацията, не се съдържат изисквания, насочващи към определена производител или доставчик. В случай, че наименование или част от наименованието съвпада с конкретен стандарт, спецификация, техническа оценка, техническо, одобрение, технически еталон и модел, източник, процес, търговска марка, патент, търговски произход или производство, да се приема, че възложителят е поставил изискването еквивалентно/и".



0000-3



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ



НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

Ангажираме се да изпълним предмета на обществената поръчка съгласно сроковете, както следва: за срок от 3 месеца, считано от датата на подписване на договора.

ДЕКЛАРИРАМ, ЧЕ: Приемам всички клаузи на приложения проект на договор за изпълнение на настоящата обществена поръчка

ДЕКЛАРИРАМ, ЧЕ: при изготвяне на офертата са спазени задълженията, свързани с данъци и осигуровки, опазване на околната среда, закрила на заетостта и условията на труд

ДЕКЛАРИРАМ, ЧЕ: предлаганата техника следва е нова, неупотребявана, нерициклирана и в производствената листа на производителя към момента на подаване на офертата, което е видно от линк:

<https://www.supermicro.com/products/system/1U/6019/SSG-6019P-ACR12L.cfm>

<https://www.supermicro.com/products/accessories/Networking/SSE-X3348S.cfm>

https://www.socomec.com/range-ups-three-phase_en.html?product=/ups-modulys-rm-gp-green-power_en.html

<https://www.supermicro.com/products/system/1U/1029/SYS-1029P-WTR.cfm>

<https://gtec-power.eu/en/ups/singlephase-online/ap160n-on-line-double-conversion-ups/>

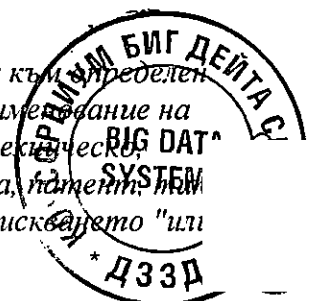
<https://www.supermicro.com/products/accessories/Networking/SSE-C3632S.cfm>

https://www.supermicro.com/wftp/Networking_Drivers/CDR-NIC_1.52_for_Add-on_NIC_Cards/MANUALS/AOC-STGN-i1S-i2S.pdf

<https://www.alliedtelesis.com/products/switches/x510l-28gt>

<https://www.watchguard.com/wgrd-products/rack-mount/firebox-m4600-m5600>, чрез който може да се провери истинността на горепосоченото.

В настоящата документацията, не се съдържат изисквания, насочващи към определен производител или доставчик. В случай, че наименование или част от наименованието на съвпада с конкретен стандарт, спецификация, техническа оценка, техническо одобрение, технически еталон и модел, източник, процес, търговска марка, предмет, произход или производство, да се приема, че възложителят е поставил изискването "или еквивалентно/и".



000024



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ



НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

ДЕКЛАРИРАМ, ЧЕ: Предлаганата техника да е съвместима с операционните системи Windows и Linux и това е доказано с линк:

[https://www.supermicro.com/support/resources/OS/OS_Certification_Intel.cfm?MProduct_Name=SSG-6019P-ACR12L;](https://www.supermicro.com/support/resources/OS/OS_Certification_Intel.cfm?MProduct_Name=SSG-6019P-ACR12L)

https://www.supermicro.com/support/resources/OS/OS_Certification_Intel.cfm?MProduct_Name=SYS-1029P-WTR към сайта

ДЕКЛАРИРАМ, ЧЕ: са налице резултати от тест, чрез тестова програма или чрез персонализирани тестове на компонентите в клъстерната хардуерна системата и това е доказано, чрез линк достъпен чрез FTP клиент: - <ftp://data.persy.com> Потребител: tests Парола: Tests123!@# и документ за тест на хардуерната система (приложение към настоящото техническо предложение)

ДЕКЛАРИРАМ, ЧЕ: Всички хардуерни системи да са снабдени с технически средства за наблюдение и управление

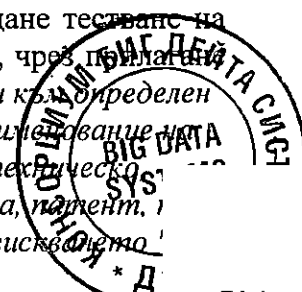
ДЕКЛАРИРАМ, ЧЕ: сме партньор от най - високо ниво на производителя на техниката

ДЕКЛАРИРАМ, ЧЕ: сме оторизирани от производителя на предлаганото оборудване

Прилагаме:

1. Документ - Техническо предложение съдържащо функционални и нефункционални изисквания на СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ, включващо следния набор от документи:
 - 1.1. „Интегриране на Специализираната централизирана система под ключ с Майкрософт SQL Server минимум версия 2017“, съгл. Стр.43, Т.6 от документацията.
 - 1.2. Документ „Ръководство на администратора за Провеждане тестване на работоспособността на отделните МодулDataNode“, чрез прилагане на „Host Heartbeats“, съгл. Стр. 44, Т.7.1 от документацията.
 - 1.3. Документ „Ръководство на администратора за провеждане тестване на изпълнението на MapReduce задание в Hadoop система, чрез прилагане

В настоящата документацията, не се съдържат изисквания, насочващи към определен производител или доставчик. В случай, че наименование или част от наименованието съвпада с конкретен стандарт, спецификация, техническа оценка, техническо одобрение, технически еталон и модел, източник, процес, търговска марка, патент, произход или производство, да се приема, че възложителят е поставил изискването еквивалентно/и“.



000025



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ СОЛИАЛЕН СЪЮЗ
ЕВРОПЕЙСКИ СЪЮЗ
ЗА РЕГИОНАЛНО РАЗВИТИЕ

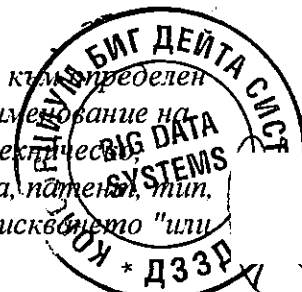


НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

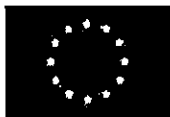
Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

- на „Pi Estimator“, на WorldCount приложение или на друго“, съгл. Стр. 45, Т.7.2 от документацията.
- 1.4. Документ „Ръководство на администратора за провеждане тестване на комплексната работа на Специализираната централизирана система под ключ Hadoop чрез прилагане на „Ние приложение“ или чрез прилагане на „Clean test“, съгл. Стр. 45, Т.7.3 от документацията.
 - 1.5. Концептуална архитектура за разширение на системата Hadoop за хиляди PetaBytes.
 - 1.6. Обучение - описание на КУРС 1 – Инсталиране и конфигуриране на Hadoop система.
 - 1.7. Обучение - описание на КУРС 2 – Основи на MapReduce програмирането
 - 1.8. Обучение - описание на КУРС 3 – Създаване на тестови данни за провеждане на изследвания.
 - 1.9. Документ „Изисквания към Специализираната централизирана система под ключ и нейното поддържане“, съгл. Стр. 48, Т.11
 - 1.10. Документ с приложен списък на ръководствата, във връзка с изискване от страна на възложителя, съгл. Стр.51, както следва:
 - 1.10.1. Документ „Ръководство на администратора за Създаване на централизирано автентикиране на потребителите на Hadoop система“
 - 1.10.2. Документ „Ръководство на администратора за Създаване на права за достъп за сървъри и сегменти от данни в Hadoop система, на вече създадени потребители“
 - 1.10.3. Документ „Ръководство на администратора за Създаване контрол на достъпа до HDFS директории и файлове на Hadoop система“
 - 1.10.4. Документ „Ръководство на администратора за Създаване централизиран Лист за управление на достъпа до Hadoop система“
 - 1.10.5. Документ „Ръководство на администратора за Създаване на механизъм за пълен одит на Hadoop система“
 - 1.10.6. Документ „Ръководство на администратора за Осигуряване на защита и криптиране на данните в Hadoop система за данни върху диск и данни в движение“

В настоящата документацията, не се съдържат изисквания, насочващи към определен производител или доставчик. В случай, че наименованието или част от наименованието на съпада с конкретен стандарт, спецификация, техническа оценка, техническо одобрение, технически еталон и модел, източник, процес, търговска марка, патент, тип, произход или производство, да се приема, че възложителят е поставил изискването "или еквивалентно/и".



000026



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ

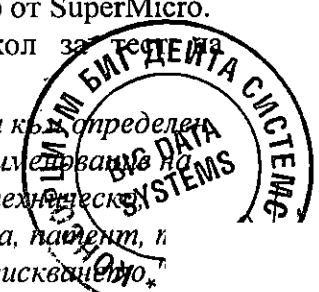


НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

- 1.10.7. Документ „Ръководство на администратора за Изграждане управление на ключовете за криптиране в Hadoop система“
- 1.10.8. Документ: „Примерно функциониране на Адаптивна Архитектура за Информационна Сигурност“
- 1.11. Документ „Съдържащ информация за интегриране на двете системи MS SQL Server 2017 или следваща версия и система Cloudera/Hadoop“
- 1.12. Документ представящ наличие на програмен продукт разработен от доставчика за интегриране на двете системи MS SQL Server 2017 или следваща версия и система Cloudera/Hadoop, който програмен продукт е с достатъчно ниво на зрялост - с пазарна реализация минимум 6 месеца, съгл. Стр. 54, Т.1.3.
 - 1.12.1. Копие на фактурите за продажба
 - 1.12.2. Банково извлечение като доказателство за плащането
 - 1.12.3. Протоколи от предаването на лицензите
- 1.13. Документ „Ръководство за прилагане на технологията за машинно обучение (Machine Learning – ML) на MS SQL Server 2017 или следваща версия, с използване на данни от система Cloudera / Hadoop и от MS SQL Server“, съгл. Т2 и Т.2.1, стр. 55
- 1.14. Документ представящ наличие на програмен продукт разработен от доставчика, за прилагане на технологията за машинно обучение (Machine Learning – ML) на MS SQL Server 2017 или следваща версия, с използване на данни от система Cloudera / Hadoop и от MS SQL Server, който програмен продукт е с достатъчно ниво на зрялост - с пазарна реализация минимум 6 месеца, съгл. Стр.55, Т.2.3.
 - 1.14.1. Копие на фактурите за продажба
 - 1.14.2. Банково извлечение като доказателство за плащането
 - 1.14.3. Протоколи от предаването на лицензите
- 1.15. Документ относно изискване от страна на Възложителя за предоставяне на писмено предложена детайлно разписана процедура за провеждане на отделните тестове, базирана на добрите практики и препоръки от страна на производителя на предлаганото оборудване в 4 приложения:
 - 1.15.1. Приложение 1 – Процедура за провеждане на тестове и изпитания свързани с предложения хардуер от SuperMicro.
 - 1.15.2. Приложение 2 – Производствен протокол за тестове на хардуерната система SuperMicro.

В настоящата документацията, не се съдържат изисквания, насочващи към определени производител или доставчик. В случай, че наименованието или част от наименованието съвпада с конкретен стандарт, спецификация, техническа оценка, техническо одобрение, технически еталон и модел, източник, процес, търговска марка, патент, и произход или производство, да се приема, че възложителят е поставил изискването еквивалентно/и“.



000027



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ



НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

1.15.3. Приложение 3 – Добри практики при конфигурацията на WatchGuard Firebox устройство.

1.15.4. Приложение 4 – Хардуерни тестове на IPS устройство на WatchGuard

1.16. Документ, във връзка с изискване от страна на Възложителя за Предоставяне на разпечатка на екрани от провеждане на подобно конфигуриране и тестове на специализирани системи, свързани с: Конфигуриране на VLAN на комутаторите (switches); Обмен на данни между два изчислителни модула с използване на конфигурираната VLAN;

1.17. Описание на съдържанието на външния хард диск (HDD)

Участник: Деян Йовков (име и фамилия)

Длъжност: Управител

Дата: 15.05.2019

Подпис и печат:



В настоящата документацията, не се съдържат изисквания, насочващи към определен производител или доставчик. В случай, че наименования или част от наименования на съвпада с конкретен стандарт, спецификация, техническа оценка, техническо, одобрение, технически еталон и модел, източник, процес, търговска марка, патент, тип, произход или производство, да се приема, че възложителят е поставил изискването "или еквивалентно/и".

000028



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ



ОПЕРАТИВНА ПРОГРАМА
НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-С 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

Образец №6

ДЕКЛАРАЦИЯ ЗА КОНФИДЕНЦИАЛНОСТ

До:

"КОНСОРЦИУМ БИГ ДЕЙТА СИСТЕМС" ДЗЗД

Долуподписаният/-ната/ Деян Йовков, в качеството ми на Управител (*посочва се длъжността и качеството, в което лицето има право да представлява и управлява - напр. изпълнителен директор, управител или др.*) на "КОНСОРЦИУМ БИГ ДЕЙТА СИСТЕМС" ДЗЗД, ЕИК/ БУЛСТАТ/ друга идентифицираща информация (В съответствие със законодателството на държавата, в която участникът е установен) ЕИК 177357133, със седалище и адрес на управление ж.к. ДРУЖБА ул. МЮНХЕН №8, София 1528 – участник в процедура за възлагане на обществена поръчка с предмет:

„Доставка, инсталиране, конфигуриране, тестване и поддържане на СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ NADOOP за работа с Големи данни по проект „Дигитализация на икономиката в среда на Големи данни“ (ДИГД), № BG05M2OP001-1.002-0002“

(моля, посочете номера и наименованието на обособената позиция)

ДЕКЛАРИРАМ, че:

В представената от мен оферта, **нямаме конфиденциална информацията.**

**Участниците не могат да се възползват на конфиденциалност по отношение на предложенията от офертите, които подлежат на оценка.*

дата 15.05.2019

Декларатор: Деян Йовков



ПОДПИС И ПЕЧАТ

В настоящата документацията, не се съдържат изисквания, насочващи към определен производител или доставчик. В случай, че наименование или част от наименование на съвпада с конкретен стандарт, спецификация, техническа оценка, техническо, одобрение, технически еталон и модел, източник, процес, търговска марка, патент, тип, произход или производство, да се приема, че възложителят е поставил изискването "или еквивалентно/и".

000029

ТЕХНИЧЕСКО ПРЕДЛОЖЕНИЕ

Функционални и нефункционални изисквания на СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД
HADOOP.

Раздел II. Техническа спецификация от Документацията за участие в Открита процедура за възлагане на обществена поръчка
с предмет:

„Доставка, инсталиране, конфигуриране, тестване и поддържане на СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА
ПОД КЛЮЧ HADOOP за работа с Големи данни по проект „Дигитализация на икономиката в среда на Големи данни“
(ДИГД), № BG05M2OP001-1.002-0002“

Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависимости
T.2. Специализирани хардуерни системи			
T. 2.2. Базова Специализирана хардуерна система за обработка на данни: Да бъде изградена в отделен рак – Рак	ДА	Предлагаме Базовата Специализирана хардуерна система за обработка на данни да бъде изградена с хардуерно оборудване на SuperMicro. <ul style="list-style-type: none"> • Базовата Специализирана хардуерна система за обработка на данни ще бъде изградена в отделен рак – Рак с номер 12. 	



Изискване	Изпълнение (Да/Не)	Начин на реализация	Зави
<p>Т. 2.2.1 Изчислителни модули на възел за данни на Базовата Специализирана хардуерна система за обработка на данни.</p> <ul style="list-style-type: none"> Базовата Специализирана хардуерна система за обработка на данни трябва да се състои от 2 броя DataNode – PakDataNode, всеки от които да се състои от минимум 6 броя Модули за данни (МодулDataNode), т.е. Базовата Специализирана хардуерна система за обработка на данни трябва да се състои общо от минимум 12 броя Модули за данни (МодулDataNode). Всеки Модул за данни (МодулDataNode) трябва да се състои от: <ul style="list-style-type: none"> минимум 2 броя процесори от тип Intel Xeon Bronze Scalable Processors, всеки от тях с минимум 8 ядра, работещи с честота 2GHz; 	ДА	<ul style="list-style-type: none"> Хардуерната система на SuperMicro модел SSG-6019P-ACR12L, като Базовата Специализирана система за обработка на данни ще се състои от 2 броя DataNode – PakDataNode, всеки от които ще се състои от минимум 6 броя Модули за данни (МодулDataNode). Това прави общо 12 броя Модули за данни (МодулDataNode). Всеки модул за данни ще съдържа следните характеристики: <ul style="list-style-type: none"> Всеки МодулDataNode на SuperMicro модел SSG-6019P-ACR12L, като част от Базовата Специализирана хардуерна система за обработка на данни включва 2 броя процесори тип 	

000031



000032



Изискване	Изпълнение (Да/Не)	Начин на реализация	
<p>- оперативна памет тип ECC с минимум 128GB с честота на работа минимум 2600 MHz;</p> <p>м 12 позиции за включване на 3,5" за потребителски данни със</p>		<p>Intel® Xeon® Silver 4110 Processor 8-core 2.10GHz 11.00MB Cache</p> <p>https://www.supermicro.com/products/system/1U/6019/SSG-6019P-ACR12L.cfm;</p> <p>https://ark.intel.com/content/www/us/en/ark/products/123547/intel-xeon-silver-4110-processor-11m-cache-2-10-ghz.html</p> <p>- Оперативната памет във всеки МодулDataNode на SuperMicro модел SSG-6019P-ACR12L, като част от Базовата Специализирана хардуерна система за обработка на данни е тип ECC. Всеки МодулDataNode ще съдържа 8 памети x 16GB PC4-21300 2666MHz DDR4 ECC. Това прави 128 GB.</p> <p>https://store.supermicro.com/16gb-ddr4-2666-mem-dr416l-sl04-er26.html</p> <p>- Всеки МодулDataNode на SuperMicro модел SSG-6019P-ACR12L се състои</p>	



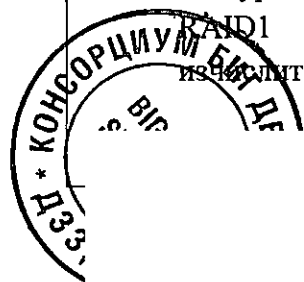
✓

✓

21

✓

000633



Изискване	Изпълнение (Да/Не)	Начин на реализация	
<p>сменяемост по време на опериране (hot-swap 3,5" drive bays), с включени в тях 3,5" дискови устройства с минимален капацитет 10TB работещи със стандартизираната дискова архитектура SAS 3.0 12.0Gb/s, която архитектура се поддържа от изчислителния модул, организирани в архитектура JBOD;</p> <p>- минимум 2 броя дискови устройства за системния софтуер на изчислителния модул, всеки от тях с минимален обем 0,5TB, организирани за работа в RAID1, осигурявана от съответен контролер за RAID1 организация на дискове в изчислителния модул;</p>		<p>от 12 позиции за включване на 3,5" дискове за потребителски данни със сменяемост по време на опериране (hot-swap 3,5" drive bays). Всеки диск е с минимален капацитет 10.0TB SAS 3.0 12.0Gb/s 7200RPM - 3.5" - Seagate Exos X10 Series (Helium) (512e), която архитектура се поддържа от изчислителния модул, организирани в архитектура JBOD.</p> <p>https://www.supermicro.com/products/system/1U/6019/SSG-6019P-ACR12L.cfm</p> <p>- Всеки МодулDataNode на SuperMicro модел SSG-6019P-ACR12L включва 2 броя дискови устройства за системния софтуер - Samsung SSD PM871b 512GB, 2.5 Inch, организирани за работа в RAID 1, осигурявана от съответен контролер за RAID1 организация на дискове в изчислителния модул.</p>	

Изискване	Изпълнение (Да/Не)	Начин на реализация	Забелжки
<ul style="list-style-type: none"> - минимум 2 порта за LAN Ethernet със скорост за обмен 1Gbps и интерфейс RJ45; - минимум 2 порта за LAN Ethernet със скорост за обмен 10Gbps и интерфейс SFP+; 		<ul style="list-style-type: none"> - Всеки Модул DataNode на SuperMicro модел SSG-6019P-ACR12L включва два порта за LAN Ethernet със скорост за обмен 1Gbps и 10Gbps вградени на дънната платка 10GBase-T ports via Intel C622 с включени 2x RJ-45. https://www.supermicro.com/products/system/1U/6019/SSG-6019P-ACR12L.cfm?parts=SHOW#jump - Всеки Модул DataNode на SuperMicro модел SSG-6019P-ACR12L включва два порта за LAN Ethernet със скорост за обмен 10Gbps както следва Supermicro 10-Gigabit Ethernet AOC-STGN-i2S LAN 2-port 10G SFP+ https://www.supermicro.com/wftp/Networking_Drivers/CDR-NIC_1.52_for_Add-on_NIC_Cards/MANUALS/AOC-STGN-i1S-i2S.pdf 	

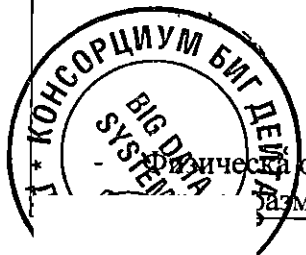
000034



Изискване	Изпълнение (Да/Не)	Начин на реализация	
<ul style="list-style-type: none"> - софтуер за управление на хардуера на модула; - минимум 2 броя интерфейси PCIe 3.0 x8 или PCIe 3.0 x16; - минимум удвоено захранване; 		<ul style="list-style-type: none"> - Софтуерът за управление на Super Micro е Supermicro Server Manager (SSM) (DataCenter Management Package). https://www.supermicro.com/en/solutions/management-software/supermicro-server-manager - Всеки МодулDataNode на SuperMicro модел SSG-6019P-ACR12L има 2 броя PCI-E 3.0 x16 слота. https://www.supermicro.com/products/system/1U/6019/SSG-6019P-ACR12L.cfm - Всеки МодулDataNode на SuperMicro модел SSG-6019P-ACR12L има двойно захранване (PSU) https://www.supermicro.com/products/system/1U/6019/SSG-6019P-ACR12L.cfm - Всеки МодулDataNode на SuperMicro модел SSG-6019P-ACR12L са с 	

ЮСТИ

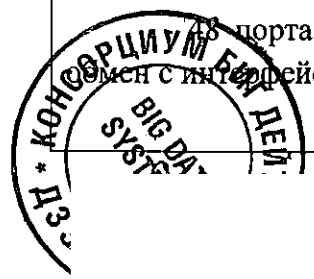
000635



...ическа организация – за инсталиране в
...размери не по-големи от 2U, за

Изискване	Изпълнение (Да/Не)	Начин на реализация	
предпочитане 1U.		<p>размери 1U (1U Rackmount)</p> <p>https://www.supermicro.com/products/system/1U/6019/SSG-6019P-ACR12L.cfm</p>	
<p>1.2.2. Комуникационни модули – комутатори (switches) на Базовата Специализирана хардуерна система за обработка на данни:</p> <p>А) Основен комуникационен модул на Базовата Специализирана хардуерна система за обработка на данни – 1 брой. Той да бъде управляем комутатор на минимално ниво 2 от OSI Референтната архитектурата, със следното минимално количество портове:</p> <p>1 порта за Ethernet 10Gbps скорост за обмен с интерфейс SFP+;</p>	ДА	<p>Предлагаме Комуникационните модули – комутатори (switches) на Базовата Специализирана хардуерна система за данни да бъдат SuperMicro.</p> <p>А) За 1 брой основен комуникационен модул на Базовата Специализирана хардуерна система за обработка на данни предлагаме Super Micro модел SSE-X3348S, Layer 2 / 3 / 4. Този модел е управляем комутатор със следните портове:</p> <ul style="list-style-type: none"> - Комутатор Super Micro модел SSE-X3348S има 48 порта за Ethernet 10Gbps скорост за обмен с интерфейс SFP+ 	

000030



Изискване	Изпълнение (Да/Не)	Начин на реализация	З:
<p>- 2 порта uplink връзки за Ethernet 40Gbps скорост за обмен с интерфейс QSFP+;</p> <p>- 2 порта за Ethernet 1Gbps скорост за обмен с интерфейс RJ45;</p> <p>Допълнителен комуникационен модул на Базовата Специализирана хардуерна система данни – 1 брой. Той да бъде</p>		<p>https://www.supermicro.com/products/accessories/Networking/SSE-X3348S.cfm</p> <p>- Комутатор Super Micro модел SSE-X3348S има uplink минимум 2x40Gbps QSFP+</p> <p>https://www.supermicro.com/products/accessories/Networking/SSE-X3348S.cfm</p> <p>- Комутатор Super Micro модел SSE-X3348S има 2 порта с по 1Gbps скорост за обмен с интерфейс RJ45.</p> <p>https://www.supermicro.com/products/accessories/Networking/SSE-X3348S.cfm</p> <p>Б) За 1 брой допълнителен комуникационен модул на Базовата Специализирана хардуерна система за обработка на данни предлагаме Allied Telesis модел AT-x510L-</p>	

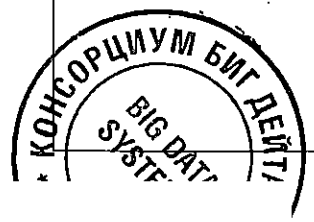
000037



1
>

Изискване	Изпълнение (Да/Не)	Начин на реализация	3
<p>управляем комутатор на минимално ниво 2 от OSI Референтната архитектурата, със следното минимално количество портове:</p> <ul style="list-style-type: none"> - 24 порта за Ethernet 1Gbps скорост за обмен с интерфейс RJ45; - 1 порт за връзка с по-висока йерархия комуникационни модули за Ethernet 1-10Gbps скорост за обмен с интерфейс SFP+. 		<p>28GT. Layer 3. Този модел е управляем комутатор, с включени следните портове.</p> <ul style="list-style-type: none"> - Комутатор Allied Telesis модел AT-x510L-28GT притежава 24 порта за Ethernet 1Gbps скорост за обмен с интерфейс RJ45. Layer 3 от OSI. <p>https://www.alliedtelesis.com/products/switches/x510l-28gt</p> <ul style="list-style-type: none"> - Комутатор Allied Telesis модел AT-x510L-28GT има 4 порт за връзка с по-висока йерархия комуникационни модули за Ethernet 1-10Gbps скорост за обмен с интерфейс SFP+. <p>https://www.alliedtelesis.com/products/switches/x510l-28gt</p>	<p>1</p> <p>5</p> <p>></p>

000033



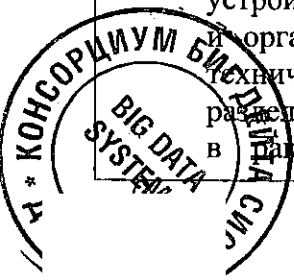
Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависимости
<p>T.2.2.3. Физическа организация на Базова Специализирана хардуерна система за обработка на данни.</p> <p>Базова Специализирана хардуерна система за обработка на данни трябва да бъде монтирана и инсталирана за работа в рак с размери 42U – рак №12, който да се достави. От този рак да се осигурят минимум 2 броя кабели за скорост на обмен 40Gbps, с дължина минимум 20 метра, чрез които да се осигури връзка с Комутатор с портове от по 40Gbps, разположен в Рак №10.</p> <p>- Рак №12 трябва да бъде със съответното управление – KVM switch за управление на минимум 12 сървъра с дисплей, клавиатура и мишка и кабели за управление на 12 компютъра. Тези дисплей, клавиатура и мишка да са в размерите сумарно на 1U.</p>	<p>ДА</p>	<p>Физическата организация на Базовата Специализирана хардуерна система за обработка на данни ще бъде направена така както е описано в изискването на Възложителя, както следва:</p> <ul style="list-style-type: none"> - Базова Специализирана хардуерна система за обработка на данни ще бъде доставена и инсталирана за работа в Rack 12, с размери 42U. От този рак ще бъдат осигурени 2 броя кабели за скорост на обмен 40Gbps, с дължина 20 метра, чрез които ще бъде осигурена връзката с Комутатор с портове от по 40Gbps, който е разположен в Рак 10. - Предлагаме в Рак №12 да бъде разположен KVM switch за управление на 16 сървъра с дисплей, клавиатура и мишка и кабели за управление на 16 компютъра. Като предложените дисплей, клавиатура и мишка ще са с 	<p>- За да бъде монтирана и инсталирана -Базова Специализирана хардуерна система, да се осигури от Възложителя:</p> <ul style="list-style-type: none"> - Достъп до помещението на специалистите ангажирани с монтажа и инсталцията; - Необходимото стабилно и достатъчно по мощност захранване на помещението; - Необходимата и достатъчна по мощност климатизация

000039



Изискване	Изпълнение (Да/Не)	Начин на реализация	Забелюшки
<p>- Всички изчислителни модули и комуникационни модули на Базова Специализирана хардуерна система за обработка на данни да бъдат монтирани в посочения рак №12. Изчислителните модули да бъдат свързани с Основния комуникационен модул с необходимите LAN кабели за 10Gbps скорост за обмен с интерфейс SFP+. Допълнителният комуникационен модул да бъде свързан към Основния комуникационен модул. При инсталирането в рака, комуникационните модули да се поставят отгоре в рака, след които да се поставят изчислителните модули. Да се доставят минимално 12 броя LAN кабели за 1Gbps скорост за обмен с интерфейс RJ45 за връзки в рамките на рака. Всички LAN кабели да се наименоват за това с кои устройства се прави връзка, и се подредят и организират в рака с оглед улеснено техническо обслужване. Да се доставят 12 разделителни носещи плочи за поставяне в рака, върху които да се поставят</p>		<p>размера на 1U.</p> <p>- Всички изчислителни модули и комуникационни модули на Базова Специализирана хардуерна система за обработка на данни ще бъдат монтирани в рак №12, както е посочено в настоящата документация. Изчислителните модули ще бъдат свързани с Основния комуникационен модул с необходимите за тази връзка LAN кабели за 10Gbps скорост за обмен с интерфейс SFP+. Допълнителният комуникационен модул ще бъде свързан към Основния комуникационен модул. При инсталиране на техниката в рака, комуникационните модули ще бъдат поставени най - отгоре в рака, след което ще бъдат поставени изчислителните модули. Ще бъдат доставени заедно с МодулитеDataNote 12 броя LAN кабели за осъществяване на 1Gbps скорост за обмен с интерфейс RJ45, което е необходимо за връзки в рамките на рака. Всички LAN кабели ще бъдат наименовани за</p>	

000040



Изискване	Изпълнение (Да/Не)	Начин на реализация	
<p>евентуални tower-based сървъри и компютри.</p> <p>Ракът да бъде доставен с UPS система. UPS да бъде с размери до 4U, като доставчикът да определи неговите технически параметри за да се осигурява токоподаване на всички устройства в рака минимум 3 минути след спиране на електричеството.</p>		<p>това с кои устройства се прави връзка, и ще бъдат старателно подредени и организирани в рака за улеснение на техническо обслужване. Ще бъдат доставени 12 разделителни носещи плочи необходими за поставяне в съответния рак, върху който могат се поставят евентуално и tower-based сървъри и компютри.</p> <p>Ракът ще бъде доставен с UPS система модел Gtec AP160N с мощност 10kVA/9kW ще бъде с размери до 5U, като моделът осигурява токоподаване на всички устройства в рака за 3 минути след спиране на електричеството.</p>	

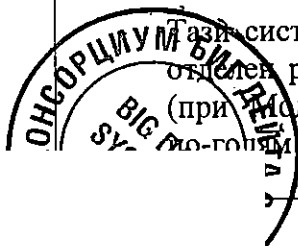
Базовата специализирана хардуерна система за обработка на данни ще съдържа следните компоненти:

1. Модул DataNote модел SuperMicro SSG-6019P-ACR12L – 12 броя
2. Основен комуникационен модул – Super Micro модел SSE-X3348S – 1 брой
3. Допълнителен комуникационен модул Allied Telesis модел AT-x510L-28GT –



Изискване	Изпълнение (Да/Не)	Начин на реализация	ТИ
		<p>1 брой</p> <p>4. KVM switch – 1 брой</p> <p>5. UPS система модел Gtec API60N – 1 брой</p> <p>6. Кабели за скорост на обмен 40Gbps с дължина 20 метра – 2 броя</p> <p>7. LAN кабели за осъществяване на 1Gbps скорост за обмен с интерфейс RJ45 – 12 броя</p> <p>8. Разделителни носещи плочи – 12 броя</p> <p>9. Физическата организация за разполагането и инсталирането на Базовата Специализирана система за обработка на данни.</p>	
<p>Т.2.3.Разширяема Специализирана хардуерна система за обработка на данни.</p> <p>Тази система трябва да бъде изградена в отделен рак – Рак №13, а по изключение (при необходимост) – и в Рак №13-А.</p>	<p>ДА</p>	<p>Предлагаме Разширяемата Специализирана хардуерна система за обработка на данни да бъде изградена с хардуерно оборудване на SuperMicro.</p> <p>- Предложеното оборудване ще бъде инсталирано в отделен рак – Рак №13, тъй като предложеното от нас оборудване за тази система е с размер 1U.</p>	

000013



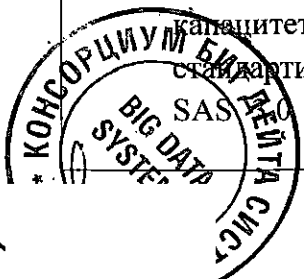
Изискване	Изпълнение (Да/Не)	Начин на реализация
<p>T.2.3.1. Изчислителни модули на възел за данни на Разширяемата Специализирана хардуерна система за обработка на данни:</p> <ul style="list-style-type: none"> Разширяемата Специализирана хардуерна система за обработка на данни трябва да се състои от 4 броя DataNode – RackDataNode, всеки от които да се състои от минимум 6 броя Модули за данни (МодулDataNode), т.е. Разширяемата Специализирана хардуерна система за обработка на данни трябва да се състои минимум общо от 24 броя Модули за данни (МодулDataNode). Всеки Модул за данни (МодулDataNode) трябва да се състои от: <ul style="list-style-type: none"> минимум 2 броя процесори от тип Intel Xeon Bronze Scalable Processors, всеки от тях с минимум 8 ядра, работещи с честота минимум 2GHz; 	<p>ДА</p>	<ul style="list-style-type: none"> Хардуерната система МодулDataNode на SuperMicro модел SSG-6019P-ACR12L, като Разширяема Специализирана система за обработка на данни ще се състои от 4 броя DataNode – RackDataNode, всеки от които се състои от минимум 6 броя Модули за данни (МодулDataNode). Това прави общо 24 броя Модули за данни (МодулDataNode). Всеки модул за данни ще съдържа: <ul style="list-style-type: none"> Във всеки МодулDataNode на SuperMicro модел SSG-6019P-ACR12L, като част от Разширяемата Специализирана хардуерна система за обработка на данни включва 2 броя процесори тип Intel® Xeon® Silver 4110 Processor 8-core 2.10GHz 11.00MB Cache <p>https://www.supermicro.com/products/system/1U/6019/SSG-6019P-ACR12L.cfm; https://ark.intel.com/content/www/us/en/ark/prod</p>

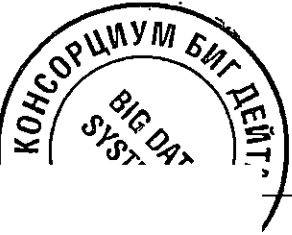
000013



Изискване	Изпълнение (Да/Не)	Начин на реализация	И
<p>- оперативна памет минимум 128GB с честота на работа минимум 2600 MHz;</p> <p>- минимум 12 позиции за включване на 3,5" дискове за потребителски данни със сменяемост по време на опериране (hot-swap 3,5" drive bays), с включени в тях 3,5" дискови устройства с минимален капацитет 10TB работещи със стандартизираната дискова архитектура SAS 12.0Gb/s, която архитектура се</p>		<p>ucts/123547/intel-xeon-silver-4110-processor-11m-cache-2-10-ghz.html</p> <p>- Оперативната памет във всеки МодулDataNode на SuperMicro модел SSG-6019P-ACR12L, като част от Разширяемата Специализирана хардуерна система за обработка на данни е тип ECC. Всеки МодулDataNode ще съдържа 8 памети x 16GB PC4-21300 2666MHz DDR4 ECC. Това прави 128 GB.</p> <p>https://store.supermicro.com/16gb-ddr4-2666-mem-dr416l-sl04-er26.html</p> <p>- Всеки МодулDataNode на SuperMicro модел SSG-6019P-ACR12L се състои от 12 позиции за включване на 3,5" дискове за потребителски данни със сменяемост по време на опериране (hot-swap 3,5" drive bays). Всеки диск е с минимален капацитет 10.0TB SAS 3.0 12.0Gb/s 7200RPM - 3.5" - Seagate Exos X10 Series (Helium) (512e), която архитектура се поддържа от изчислителния модул, организирани в архитектура JBOD.</p>	

000047





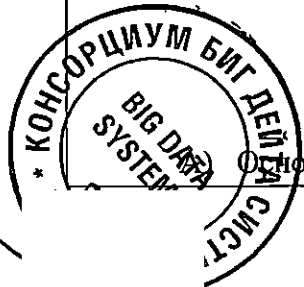
Изискване	Изпълнение (Да/Не)	Начин на реализация	
<p>поддържа от изчислителния модул, организирани в архитектура JBOD;</p> <p>минимум 2 броя дискови устройства с технология SSD за системния софтуер на изчислителния модул, всеки от тях с минимален обем 0,5TB, организирани за работа в RAID1, осигурявана от съответен контролер за RAID1 организация на дискове в изчислителния модул;</p> <p>- минимум 2 порта за LAN Ethernet със скорост за обмен 1Gbps и интерфейс RJ45;</p> 		<p>https://www.supermicro.com/products/system/1U/6019/SSG-6019P-ACR12L.cfm</p> <ul style="list-style-type: none"> - Всеки МодулDataNode на SuperMicro модел SSG-6019P-ACR12L включва 2 броя дискови устройства за системния софтуер - Samsung SSD PM871b 512GB, 2.5 Inch, организирани за работа в RAID 1, осигурявана от съответен контролер за RAID1 организация на дискове в изчислителния модул. - Всеки МодулDataNode на SuperMicro модел SSG-6019P-ACR12L включва два порта за LAN Ethernet със скорост за обмен 1Gbps и 10Gbps вградени на дънната платка- 10GBase-T ports via Intel C622 с включени 2x RJ-45. <p>https://www.supermicro.com/products/system/1U/6019/SSG-6019P-ACR12L.cfm?parts=SHOW#jump</p>	

Изискване	Изпълнение (Да/Не)	Начин на реализация
<p>- минимум 2 порта за LAN Ethernet със скорост за обмен 10Gbps и интерфейс +;</p> <p>- софтуер за управление на хардуера на модула;</p> <p>Минимум 2 броя интерфейси PCIe 3.0 x8 или PCIe 3.0 x16;</p>		<p>- Всеки МодулDataNode на SuperMicro модел SSG-6019P-ACR12L включва два порта за LAN Ethernet със скорост за обмен 10Gbps както следва Supermicro 10-Gigabit Ethernet AOC-STGN-i2S LAN 2-port 10G SFP+</p> <p>https://www.supermicro.com/wftp/Networking_Drivers/CDR-NIC_1.52_for_Add-on_NIC_Cards/MANUALS/AOC-STGN-i1S-i2S.pdf</p> <p>- Софтуерът за управление на Super Micro е Supermicro Server Manager (SSM) (DataCenter Management Package).</p> <p>https://www.supermicro.com/en/solutions/management-software/supermicro-server-manager</p> <p>- Всеки МодулDataNode на SuperMicro модел SSG-6019P-ACR12L има 2 броя PCI-E 3.0 x16 слота.</p> <p>https://www.supermicro.com/products/system/1</p>

01:00:00



Изискване	Изпълнение (Да/Не)	Начин на реализация	Зав.
<p>- минимум удвоено захранване;</p> <p>- Физическа организация – за инсталиране в рак, с размери не по-големи от 2U, за предпочитане 1U. При избор на Модула за данни (МодулDataNode) който е по-голям от 1U, то да се достави допълнителен рак със съответното управление, който да се обозначи като Рак №13-А.</p>		<p>U/6019/SSG-6019P-ACR12L.cfm</p> <p>- Всеки МодулDataNode на SuperMicro модел SSG-6019P-ACR12L има двойно захранване.</p> <p>https://www.supermicro.com/products/system/1U/6019/SSG-6019P-ACR12L.cfm</p> <p>- МодулDataNode на SuperMicro модел SSG-6019P-ACR12L са с размери 1U (1U Rackmount), те ще бъдат разположени в рак №13.</p> <p>https://www.supermicro.com/products/system/1U/6019/SSG-6019P-ACR12L.cfm</p>	
<p>Т 2.3.2 Комуникационни модули - комутатори (switches) на Разширяемата Специализирана хардуерна система за обработка на данни.</p> <p>Основен комуникационен модул на</p>	ДА	<p>Предлагаме Комуникационните модули – комутатори (switches) на Разширяемата Специализирана хардуерна система за данни да бъдат на SuperMicro.</p> <p>А) Всеки брой основен комуникационен</p>	



Изискване	Изпълнение (Да/Не)	Начин на реализация	
<p>Разширяемата Специализирана хардуерна система за обработка на данни – 3 броя. Те да бъдат управляеми комутатори на минимално ниво 2 от OSI Референтната архитектурата, със едно минимално количество портове:</p> <ul style="list-style-type: none"> - 48 порта за Ethernet 10Gbps скорост за мен с интерфейс SFP+; - 2 порта uplink връзки за Ethernet 40Gbps скорост за обмен с интерфейс QSFP+; - 2 порта за Ethernet 1Gbps скорост за обмен с интерфейс RJ45; 		<p>модул на Разширяемата Специализирана хардуерна система за обработка на данни предлагаме Super Micro модел SSE-X3348S. Layer 2 / 3 / 4. Този модел е управляем комутатор със следните портове:</p> <ul style="list-style-type: none"> - Комутатор Super Micro модел SSE-X3348S има 48 порта за Ethernet 10Gbps скорост за обмен с интерфейс SFP+ <p>https://www.supermicro.com/products/accessories/Networking/SSE-X3348S.cfm</p> <ul style="list-style-type: none"> - Комутатор Super Micro модел SSE-X3348S има uplink минимум 2x40Gbps QSFP+ <p>https://www.supermicro.com/products/accessories/Networking/SSE-X3348S.cfm</p> <ul style="list-style-type: none"> - Комутатор Super Micro модел SSE-X3348S има 2 порта с по 1Gbps скорост за обмен с интерфейс RJ45. 	11



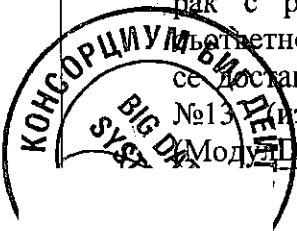
Изискване	Изпълнение (Да/Не)	Начин на реализация	И
<p>- При използване на Модула за данни (МодулDataNode) който е по-голям от 1U, то в Рак №13 да се поставят два от тези комуникационни модула, а в Рак №13-А – третият комуникационен модул.</p> <p>Б) Допълнителен комуникационен модул на Разширяемата Специализирана хардуерна система за обработка на данни – 1 брой. При използване на Рак №13-А този допълнителен комуникационен модул се поставя там. Този комуникационен модул да бъде управляем комутатор на минимално ниво 2 от OSI Референтната архитектурата, със следното минимално количество портове:</p> <p>24 порта за Ethernet 1Gbps скорост за обмен с интерфейс RJ45;</p>		<p>https://www.supermicro.com/products/accessories/Networking/SSE-X3348S.cfm</p> <p>- Модула за данни (МодулDataNode) на SuperMicro модел SSG-6019P-ACR12L са с размери 1U (1U Rackmount), ще бъдат разположени в рак №13.</p> <p>Б) За 1 брой допълнителен комуникационен модул на Разширяемата Специализирана хардуерна система за обработка на данни предлагаме Allied Telesis модел AT-x510L-28GT Layer 3 . Този модел е управляем комутатор, с включени следните портове.</p> <p>- Комутатор Allied Telesis модел AT-x510L-28GT притежава 24 порта за Ethernet 1Gbps скорост за обмен с интерфейс RJ45.</p>	<p>А А</p>

0010100



Изискване	Изпълнение (Да/Не)	Начин на реализация	Забелжки
<p>- 1 порт за връзка с по-висока йерархия комуникационни модули за Ethernet 1-10Gbps скорост за обмен с интерфейс SFP+.</p>		<p>https://www.alliedtelesis.com/products/switches/x510l-28gt</p> <p>- Комутатор Allied Telesis модел AT-x510L-28GT има 4 порт за връзка с по-висока йерархия комуникационни модули за Ethernet 1-10Gbps скорост за обмен с интерфейс SFP+.</p> <p>https://www.alliedtelesis.com/products/switches/x510l-28gt</p>	
<p>2.3.3. Физическа организация на Разширяемата специализирана хардуерна система за обработка на данни.</p> <p>- Разширяемата Специализирана хардуерна система за обработка на данни трябва да бъде монтирана и инсталирана за работа в Рак с размери 42U – Рак №13 (или съответно и Рак №13-А), който (които) да се достави. При използване само на Рак №13 използва се Модула за данни (DataNode) който е 1U), то от този</p>	ДА	<p>Физическата организация на Разширяемата Специализирана хардуерна система за обработка на данни ще бъде направена така както е описано в изискването на Възложителя, както следва:</p> <p>- Разширяемата Специализирана хардуерна система за обработка на данни ще бъде доставена и инсталирана за работа в Rack 13, с размери 42U. От този рак ще бъдат осигурени 6 броя кабели за скорост на обмен 40Gbps, с дължина 20 метра, чрез които ще бъде осигурена връзката</p>	<p>- За да бъде монтирана и инсталирана - Разширяемата Специализирана хардуерна система, да се осигури от Възложителя:</p> <p>- Достъп до помещението на</p>

000050



1000051

Изискване	Изпълнение (Да/Не)	Начин на реализация	3
<p>рак да се осигурят 6 броя кабели за скорост на обмен 40Gbps, с дължина минимум 20 метра, чрез които да се осигури връзка с Комутатор с портове от по 40Gbps, разположен в Рак №10. При използване на Рак №13 и Рак №13-А (използва се Модула за данни (МодулDataNode) който е по-голям от 1U), то от Рак №13 да се осигурят 4 броя кабели за скорост на обмен 40Gbps, с дължина минимум 20 метра, чрез които да се осигури връзка с Комутатор с портове от по 40Gbps, разположен в Рак №10, а от Рак №13-А да се осигурят 2 броя кабели за скорост на обмен 40Gbps, с дължина минимум 20 метра, чрез които да се осигури връзка с Комутатор с портове от по 40Gbps, разположен в Рак №10.</p> <p>- Рак №13 трябва да бъде със съответното управление – KVM switch за управление на 12 сървъра с дисплей, клавиатура и мишка и кабели за управление на 12 компютъра. Тези дисплей, клавиатура и мишка да са в размерите сумарно на 1U.</p>		<p>с Комутатор с портове от по 40Gbps, който е разположен в Рак 10. От Рак 13 ще бъдат осигурени 4 броя кабели за скорост на обмен 40Gbps, с дължина 20 метра, чрез които да се осигури връзка с Комутатора с портове от по 40Gbps, разположен в Рак №10.</p> <p>- Предлагаме в Рак №13 да бъде разположен KVM switch за управление на 16 сървъра с дисплей, клавиатура и мишка и кабели за управление на 16 компютъра. Като предложените дисплей, клавиатура и мишка да са с размера на 1U.</p>	<p>специалистите ангажирани с монтажа и инсталцията;</p> <ul style="list-style-type: none"> - Необходимото стабилно и достатъчно по мощност захранване на помещението; - Необходимата и достатъчна по мощност климатизация

A

Изискване	Изпълнение (Да/Не)	Начин на реализация	3
<p>- Всички изчислителни модули и комуникационни модули на Разширяемата Специализирана хардуерна система за обработка на данни да бъдат монтирани в тосочения рак (ракове). Изчислителните модули да бъдат свързани с Основните комуникационни модули с необходимите LAN кабели за 10Gbps скорост за обмен с интерфейс SFP+. Допълнителният комуникационен модул да бъде свързан към един от Основните комуникационни модули. При инсталирането в рака, комуникационните модули да се поставят отгоре в рака, след които да се поставят изчислителните модули. Да се доставят минимално 24 броя LAN кабели за 1Gbps скорост за обмен с интерфейс RJ45. Всички LAN кабели да се наименоуват за това с кои устройства се прави връзка, и се подредят и организират в рака с оглед улеснено техническо обслужване. За всеки рак да се доставят 12 разделителни носещи плочи за поставяне в рак, върху които да се поставят евентуални tower-based сървъри и компютри.</p>		<p>- Всички изчислителни модули и комуникационни модули на Разширяемата Специализирана хардуерна система за обработка на данни ще бъдат монтирани в рак №13, както е посочено и в изискването за 1U. Изчислителните модули ще бъдат свързани с Основния комуникационен модул с необходимите за тази връзка LAN кабели за 10Gbps скорост за обмен с интерфейс SFP+. Допълнителният комуникационен модул ще бъде свързан към Основния комуникационен модул. При инсталиране на техниката в рака, комуникационните модули ще бъдат поставени най-отгоре в рака, след което ще бъдат поставени изчислителните модули. Ще бъдат доставени 24 броя LAN кабели за осъществяване на 1Gbps скорост за обмен с интерфейс RJ45, което е необходимо за връзки в рамките на рака. Всички LAN кабели ще бъдат наименоувани за това с кои устройства се прави връзка, и ще бъдат старателно подредени и организирани в рака за улеснение на техническо обслужване.</p>	<p>✓</p>

000052



Изискване	Изпълнение (Да/Не)	Начин на реализация
<p>- При прилагане на Модула за данни (МодулDataNode) който е по-голям от 1U, т.е. използва се РАК №13 и Рак №13-А, то в Рак №13 се поставят 18 броя Модула за данни (МодулDataNode), а в Рак №13-А се поставят 6 броя Модула за данни (МодулDataNode).</p> <p>- Ракът (раковете при използване на Рак №13-А) да бъде доставен с UPS система. UPS да бъде с размери до 4U, като доставчикът да определи неговите технически параметри за да се осигурява токоподаване на всички устройства в рака минимум 3 минути след спиране на електричеството.</p>		<p>Ще бъдат доставени 12 разделителни носещи плочи необходими за поставяне в съответния рак, върху който могат се поставят евентуално и tower-based сървъри и компютри.</p> <p>- Модула за данни (МодулDataNode) на SuperMicro ще бъде с размер 1U, затова в Рак 13 ще бъдат поставени всичките 24 модула за данни.</p> <p>- Ракът ще бъде доставен с UPS система SOCOMEC MODULYS GP RM с мощност 25kVA/25kW ще бъде с размери до 9U, като моделът осигурява токоподаване на всички устройства в рака за минимум 3 минути след спиране на електричеството.</p> <p>https://www.socomec.com/range-ups-three-phase_en.html?product=/ups-modulys-rm-gp-green-power_en.html</p>

000053



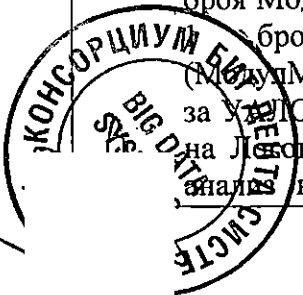
Изискване	Изпълнение (Да/Не)	Начин на реализация	СИ
<p>Разширяемата специализирана хардуерна система за обработка на данни ще съдържа следните компоненти:</p>		<ol style="list-style-type: none"> 1. Модул DataNote модел SuperMicro SSG-6019P-ACR12L – 24 броя 2. Основен комуникационен модул – Super Micro модел SSE-X3348S – 3 броя 3. Допълнителен комуникационен модул - Allied Telesis модел AT-x510L-28GT – 1 брой 4. KVM switch – 1 брой 5. UPS система модел SOCOMEC MODULYS GP RM – 1 брой 6. Кабели за скорост на обмен 40Gbps с дължина над 20 метра – 10 броя 7. LAN кабели за осъществяване на 1Gbps скорост за обмен с интерфейс RJ45 – 24 броя 8. Разделителни носещи плочи – 12 броя 9. Физическата организация за разполагането и инсталирането на Разширяемата Специализирана система за обработка на данни. 	


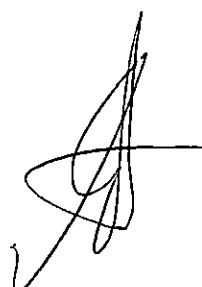

000000



Изискване	Изпълнение (Да/Не)	Начин на реализация	ТИ
<p>Г.2.4. Управляваща Специализирана хардуерна система за управление обработка на данни.</p> <ul style="list-style-type: none"> - Тази система трябва да бъде изградена в отделен рак – Рак №11. - Предназначението на Управляващата Специализирана хардуерна система е да извършва управление на цялостната Специализирана система за обработка на данни, както и да управлява основното функциониране на Информационната сигурност. - Управляващата Специализирана хардуерна система трябва да се състои от следните 6 броя изчислителни модули: 3 броя Модули за имена (МодулNameNode), 1 брой Модул за управление (МодулManagementNode), 1 брой Сървър за УАЛС (сървър за Управление и Анализ на Логове и Събития - за управление и анализ на всички логове и настъпващи 	<p>ДА</p>	<p>Предлагаме Управляващата Специализирана хардуерна система за обработка на данни да бъде изградена с хардуерно оборудване на SuperMicro.</p> <ul style="list-style-type: none"> - Тази система ще бъде разположена и изградена в Рак №11. - В тази система Управляващата Специализирана хардуерна система за обработка на данни ще бъде разположени и решенията за Информационна сигурност, обект на тази поръчка. - Предлагаме Управляващата хардуерна система за данни да се състои от 3 броя Модули за имена (МодулNameNode) – Super Micro модел SYS-1029P-WTR, 1 брой Модул за управление (МодулManagementNode) – SuperMicro модел SYS-1029P-WTR, 1 брой Сървър за УАЛС (сървър за 	<p>7</p>

00050

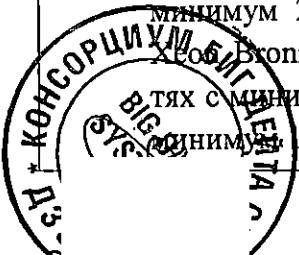


Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависимости
<p>събития в Специализираната система под ключ) и 1 брой Сървър за МУД (сървър за Мрежово Управление и Достъп – за управление на мрежовите комутатори и изпълняваните от тях управления на достъпа в Специализираната система под ключ). Изчислителните модули 3 броя Модули за имена (МодулNameNode) и 1 брой Модул за управление (МодулManagementNode) са свързани с управление работата на Hadoop средата, докато изчислителните модули Сървър за УАЛС и Сървър за МУД са свързани с управлението на Информационната сигурност на Специализираната система под ключ. Управляващата Специализирана хардуерна система трябва да се състои също и от комуникационни модули (switches): Основен комуникационен модул и Допълнителен комуникационен модул.</p>		<p>Управление и Анализ на Логове и Събития - за управление и анализ на всички логове и настъпващи събития в Специализираната система под ключ) – SuperMicro модел SSG-6019P-ACR12L и 1 брой Сървър за МУД (сървър за Мрежово Управление и Достъп – Super Micro модел SYS-1029P-WTR. Като Изчислителните модули 3 броя Модули за имена (МодулNameNode) и 1 брой Модул за управление (МодулManagementNode) ще бъдат свързани с управление работата на Hadoop средата, докато изчислителните модули Сървър за УАЛС и Сървър за МУД ще бъдат свързани с управлението на Информационната сигурност на Специализираната система под ключ. Управляващата Специализирана хардуерна система ще се състои също и от комуникационни модули (switches): Основен комуникационен модул и Допълнителен комуникационен модул.</p> <p>https://www.supermicro.com/products/system/1U/1029/SYS-1029P-WTR.cfm</p>	  



Изискване	Изпълнение (Да/Не)	Начин на реализация	
<p>T.2.4.1. Изчислителни модули на Управляващата Специализирана хардуерна система</p> <p>А) Изчислителни модули свързани с управление на Hadoop средата</p> <p>- Всеки от 4-те изчислителните модули - 3 броя Модули за имена (МодулNameNode) и 1 брой Модул за управление (МодулManagementNode), трябва да се състоят от:</p> <p>минимум 2 броя процесори от тип Intel Xeon Bronze Scalable Processors, всеки от тях с минимум 8 ядра, работещи с честота минимум 2,7GHz;</p>	<p>ДА</p>	<p>https://www.supermicro.com/products/system/1U/6019/SSG-6019P-ACR12L.cfm</p> <p>Предлагаме изчислителните модули на Управляващата Специализирана хардуерна система да бъдат SuperMicro.</p> <p>А) Изчислителни модули свързани с управление работата на Hadoop средата да бъдат изградени със Super Micro.</p> <ul style="list-style-type: none"> - Предлагаме 3 те броя Модули за имена (МодулNameNode) да бъдат изградени със SuperMicro модел SYS-1029P-WTR, и 1 брой Модул за управление (МодулManagementNode) да бъде изграден със SuperMicro модел SYS-1029P-WTR. Всеки от изброените изчислителни модула ще съдържа: - Във всеки МодулNameNode на SuperMicro модел SYS-1029P-WTR, като част от Управляващата Специализирана хардуерна система за обработка на данни включва 2 броя 	

750057



Изискване	Изпълнение (Да/Не)	Начин на реализация	За
<p>- оперативна памет минимум 128GB с честота на работа минимум 2600 MHz;</p> <p>минимум 8 позиции за включване на 2,5"</p>		<p>процесори тип 2 Intel® Xeon® Bronze 3106 Processor 8-core 1.70GHz 11.00MB Cache</p> <p>https://www.supermicro.com/products/system/1U/1029/SYS-1029P-WTR.cfm; https://ark.intel.com/content/www/us/en/ark/products/123540/intel-xeon-bronze-3106-processor-11m-cache-1-70-ghz.html</p> <p>- Оперативната памет във всеки МодулNameNode на SuperMicro модел SYS-1029P-WTR, като част от Управляемата Специализирана хардуерна система за обработка на данни е тип ECC. Всеки сървър ще съдържа 8 памети x 16GB PC4-21300 2666MHz DDR4 ECC. Това прави 128 GB.</p> <p>https://www.supermicro.com/products/system/1U/1029/SYS-1029P-WTR.cfm</p> <p>- Всеки МодулNameNode на SuperMicro модел SYS-1029P-WTR се състои от 8 позиции за включване на 2,5"</p>	<p>3а</p>

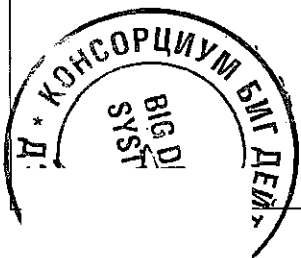


(Handwritten mark)

(Handwritten mark)

Изискване	Изпълнение (Да/Не)	Начин на реализация	Завършено
<p>или 3,5" дискове със сменяемост по време на опериране (hot-swap drive bays) с инсталирани дискове с минимум 2.0TB с архитектура SAS 3.0 12.0Gb/s със съответен SAS-RAID контролер, работещи в архитектура RAID 1/5/6, които да служат за системен софтуер и за потребителски данни;</p> <ul style="list-style-type: none"> - оптично устройство DVD-RW / 24; - минимум 2 порта за LAN Ethernet със скорост за обмен 1Gbps и интерфейс RJ45; 		<p>дискове за потребителски данни със сменяемост по време на опериране (hot-swap 2,5" drive bays). Всеки диск е с минимален капацитет 2.0TB SAS 3.0 12.0Gb/s 7200RPM - 2.5" - Seagate Exos 7E2000 Series (512e), работещи в архитектура RAID 1/5/6.</p> <ul style="list-style-type: none"> - В МодулNameNode на Super Micro модел SYS-1029P-WTR се включва следното оптично устройство LG Slim 8x DVD-RW / 24x CDR Combo (SATA) <p>https://www.supermicro.com/products/system/1U/1029/SYS-1029P-WTR.cfm</p> <ul style="list-style-type: none"> - Всеки МодулNameNode на SuperMicro модел SYS-1029P-WTR включва два порта за LAN Ethernet със скорост за обмен 1Gbps и 10Gbps вградени на дънната платка- 10GBase-T ports via Intel C622 с включени 2x RJ-45. <p>https://www.supermicro.com/products/system/1U/6019/SSG-6019P-ACR12L.cfm?parts=SHOW#jump</p>	<p>✓</p> <p>✓</p> <p>✓</p>

0000530



Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависимост
<ul style="list-style-type: none"> - минимум 2 порта за LAN Ethernet със скорост за обмен 10Gbps и интерфейс SFP+; - елемент за backup на захранване при RAID контролера; - минимум удвоено захранване; - Софтуерът за управление на хардуера на 		<ul style="list-style-type: none"> - Всеки МодулNameNode на SuperMicro модел SYS-1029P-WTR включва два порта за LAN Ethernet със скорост за обмен 10Gbps както следва Supermicro Supermicro 10-Gigabit Ethernet AOC-STGN-i2S <p>https://www.supermicro.com/wftp/Networking_Drivers/CDR-NIC_1.52_for_Add-on_NIC_Cards/MANUALS/AOC-STGN-i1S-i2S.pdf</p> <ul style="list-style-type: none"> - Всеки МодулNameNode на SuperMicro модел SYS-1029P-WTR притежава елемент за backup на захранване при RAID контролера. - Всеки МодулNameNode на SuperMicro модел SYS-1029P-WTR има двойно захранване. <p>https://www.supermicro.com/products/system/1U/1029/SYS-1029P-WTR.cfm</p> <ul style="list-style-type: none"> - Софтуерът за управление на Super Micro e Supermicro Server Manager 	

000060



Изискване	Изпълнение (Да/Не)	Начин на реализация	
<p>модула;</p> <p>- Физическа организация – за инсталиране в рак, с размери не по-големи от 2U, за предпочитане 1U.</p> <p>Б) Изчислителен модул „Сървър за УАЛС“ (Сървър за управление и анализ на логове и събития)</p> <p>Изчислителният модул Сървър за УАЛС трябва да има следните минимални технически параметри:</p> <p>минимум 2 броя процесори от тип Intel Xeon Scalable Processors, всеки от</p>		<p>(SSM) (DataCenter Management Package).</p> <p>https://www.supermicro.com/en/solutions/management-software/supermicro-server-manager</p> <p>- Сървърите SuperMicro SSG- SYS-1029P-WTR са с размери 1U (1U Rackmount).</p> <p>https://www.supermicro.com/products/system/1U/1029/SYS-1029P-WTR.cfm</p> <p>Б) Предлагаме Изчислителният модул за „Сървър за УАЛС“ (Сървър за управление и анализ на логове и събития) да бъде SuperMicro модел SSG-6019P-ACR12L.</p> <p>- Във всеки изчислителен модул SuperMicro модел SSG-6019P-</p>	<p>7</p> <p>2</p> <p>L</p> <p>U</p>

000061



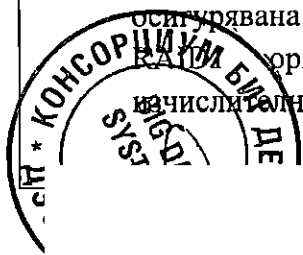
Изискване	Изпълнение (Да/Не)	Начин на реализация	Заб.
<p>тях с минимум 8 ядра, работещи с честота минимум 1,7GHz;</p> <p>- Оперативна памет тип ECC с минимум 128GB с честота на работа минимум 2600 MHz;</p>		<p>ACR12L, като част от Управляващата Специализирана хардуерна система за обработка на данни включва 2 броя процесори тип 2 Intel® Xeon® Bronze 3106 Processor 8-core 1.70GHz 11.00MB Cache</p> <p>https://www.supermicro.com/products/system/1U/6019/SSG-6019P-ACR12L.cfm; https://ark.intel.com/content/www/us/en/ark/products/123540/intel-xeon-bronze-3106-processor-11m-cache-1-70-ghz.html</p> <p>- Оперативната памет във всеки изчислителен модул на SuperMicro модел SSG-6019P-ACR12L, като част от Управляемата Специализирана хардуерна система за обработка на данни е тип ECC. Всеки изчислителен модул ще съдържа 8 памети x 16GB PC4-21300 2666MHz DDR4 ECC. Това прави 128 GB.</p> <p>https://store.supermicro.com/16gb-ddr4-2666-mem-dr416l-sl04-er26.html</p>	<p>1</p> <p>✓</p> <p>1</p>

000000



Изискване	Изпълнение (Да/Не)	Начин на реализация	Зави
<p>- минимум 12 позиции за включване на 3,5" дискове за потребителски данни със сменяемост по време на опериране (hot-swap 3,5" drive bays), с включени в 4 от тях 3,5" дискови устройства с минимален капацитет 10TB работещи със стандартизираната дискова архитектура SAS 3.0 12.0Gb/s, която архитектура се поддържа от изчислителния модул, организирани в архитектура RAID10;</p> <p>- минимум 2 броя дискови устройства за системния софтуер на изчислителния модул, всеки от тях с минимален обем 480GB, организирани за работа в RAID1, осигурявана от съответен контролер за RAID организация на дискове в изчислителния модул;</p>		<p>- Всеки изчислителен модул на SuperMicro модел SSG-6019P-ACR12L се състои от 12 позиции за включване на 3,5" дискове за потребителски данни със сменяемост по време на опериране (hot-swap 3,5" drive bays) с включени в 4 от тях 3,5" дискови устройства с минимален капацитет 10.0TB SAS 3.0 12.0Gb/s 7200RPM - 3.5" - Seagate Exos X10 Series (Helium) (512e), работещи в архитектура RAID 10.</p> <p>https://www.supermicro.com/products/system/1U/6019/SSG-6019P-ACR12L.cfm</p> <p>- Във всеки изчислителен модул на SuperMicro модел SSG-6019P-ACR12L има 2 броя дискови устройства за системния софтуер на изчислителния модул, всеки от тях с 512 GB организирани в RAID 1 както следва Samsung SSD PM871b 512GB, 2.5 Inch</p>	<p>А</p>

000000



Изискване	Изпълнение (Да/Не)	Начин на реализация	Завис
<ul style="list-style-type: none"> - минимум 2 порта за LAN Ethernet със скорост за обмен 1Gbps и интерфейс RJ45; - минимум 2 порта за LAN Ethernet със скорост за обмен 10Gbps и интерфейс SFP+; 		<ul style="list-style-type: none"> - Всеки изчислителен модул на SuperMicro модел SSG-6019P-ACR12L включва два порта за LAN Ethernet със скорост за обмен 1Gbps и 10Gbps вградени на дънната платка- 10GBase-T ports via Intel C622 с включени 2x RJ-45. https://www.supermicro.com/products/system/1U/6019/SSG-6019P-ACR12L.cfm?parts=SHOW#jump - Всеки изчислителен модул на SuperMicro модел SSG-6019P-ACR12L включва два порта за LAN Ethernet със скорост за обмен 10Gbps както следва Supermicro Supermicro 10-Gigabit Ethernet AOC-STGN-i2S LAN 2-port 10G SFP+ https://www.supermicro.com/wftp/Networking_Drivers/CDR-NIC_1.52_for_Addon_NIC_Cards/MANUALS/AOC-STGN-i1S-i2S.pdf - Софтуерът за управление на Super 	

Handwritten mark

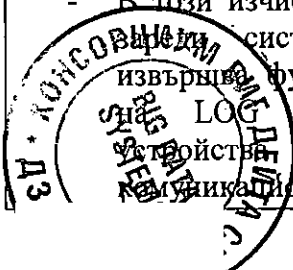
Handwritten mark



000064

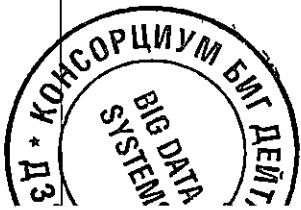
Изискване	Изпълнение (Да/Не)	Начин на реализация	За
<ul style="list-style-type: none"> - софтуер за управление на хардуера на модула; - минимум 2 броя интерфейси PCIe 3.0 x16; - Физическа организация – за инсталиране в рак, с размери не по-големи от 2U, за предпочитане 1U. - В този изчислителен модул трябва да се извършва системно функциониране по управление на LOG файловете на отделните устройства (изчислителни и комуникационни модули) на 		<p>Micro e Supermicro Server Manager (SSM) (DataCenter Management Package).</p> <p>https://www.supermicro.com/en/solutions/management-software/supermicro-server-manager</p> <ul style="list-style-type: none"> - Всеки изчислителен модул SuperMicro SSG-6019P-ACR12L включва 2 PCI-E 3.0 x16 слота. - Изчислителните модули на SuperMicro модел SSG- SSG-6019P-ACR12L са с размери 1U (1U Rackmount). <p>https://www.supermicro.com/products/system/1U/6019/SSG-6019P-ACR12L.cfm</p> <ul style="list-style-type: none"> - В този изчислителен модул ще бъде зареден системен софтуер, който ще извършва управление на LOG файловете на отделните устройства (изчислителни и комуникационни модули) на Специализираната система 	

0000055



Изискване	Изпълнение (Да/Не)	Начин на реализация	Зави
<p>Специализираната система под ключ Nadoor и възникналите събития свързани с Информационната система в реално време.</p> <p>) Изчислителен модул „Сървър за МУД“ (Сървър за Мрежово управление и достъп)</p> <p>Изчислителният модул „Сървър за МУД“ трябва да има следните минимални технически параметри:</p> <ul style="list-style-type: none"> - минимум 2 броя процесори от тип Intel Xeon Bronze Scalable Processors, всеки от тях с минимум 8 ядра, работещи с честота минимум 1,7GHz; 		<p>под ключ Nadoor и възникналите събития свързани с Информационната система в реално време.</p> <p>В) Предлагаме Изчислителният модул за „Сървър за МУД“ (Сървър за Мрежово управление и достъп) да бъде SuperMicro модел SYS-1029P-WTR.</p> <ul style="list-style-type: none"> - Във всеки изчислителен модул SuperMicro модел SYS-1029P-WTR, като част от Управляващата Специализирана хардуерна система за обработка на данни включва 2 броя процесори тип 2 Intel® Xeon® Bronze 3106 Processor 8-core 1.70GHz 11.00MB Cache <p>https://www.supermicro.com/products/system/1U/1029/SYS-1029P-WTR.cfm; https://ark.intel.com/content/www/us/en/ark/products/123540/intel-xeon-bronze-3106-processor-11m-cache-1-70-ghz.html</p> <ul style="list-style-type: none"> - Оперативната памет във всеки 	

000000



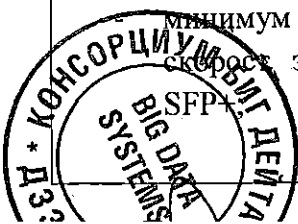
Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависимости
<p>- оперативна памет минимум 128GB с честота на работа минимум 2600 MHz;</p> <p>- минимум 8 позиции за включване на 2,5" или 3,5" дискове със сменяемост по време на опериране (hot-swap drive bays) с инсталирани дискове с минимум 2.0TB с архитектура SAS 3.0 12.0Gb/s със съответен SAS-RAID контролер, работещи в архитектура RAID 1/5/6, които да служат за системни софтуер и за потребителски данни</p>		<p>изчислителен модул на SuperMicro модел SYS-1029P-WTR, като част от Управляемата Специализирана хардуерна система за обработка на данни е тип ECC. Всеки изчислителен модул ще съдържа 8 памети x 16GB PC4-21300 2666MHz DDR4 ECC. Това прави 128 GB.</p> <p>https://store.supermicro.com/16gb-ddr4-2666-mem-dr416l-sl04-er26.html</p> <p>- Всеки изчислителен модул на SuperMicro модел SYS-1029P-WTR се състои от 8 позиции за включване на 2,5" дискове със сменяемост по време на опериране (hot-swap drive bays) с инсталирани дискове 2.0TB SAS 3.0 12.0Gb/s 7200RPM - 2.5" - Seagate Exos 7E2000 Series (512e), работещи в архитектура RAID 1/5/6.</p> <p>https://www.supermicro.com/products/system/1U/1029/SYS-1029P-WTR.cfm</p>	

000000



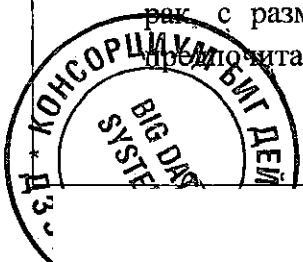
Изискване	Изпълнение (Да/Не)	Начин на реализация	За
<ul style="list-style-type: none"> - оптическо устройство DVD-RW / 24; - минимум 2 порта за LAN Ethernet със скорост за обмен 1Gbps и интерфейс RJ45; - минимум 2 порта за LAN Ethernet със скорост за обмен 10Gbps и интерфейс 		<ul style="list-style-type: none"> - Всеки изчислителен модул на SuperMicro модел SYS-1029P-WTR има оптическо устройство LG Slim 8x DVD-RW / 24x CDR Combo (SATA) https://www.supermicro.com/products/system/1U/1029/SYS-1029P-WTR.cfm - Всеки изчислителен модул на SuperMicro модел SYS-1029P-WTR включва два порта за LAN Ethernet със скорост за обмен 1Gbps и 10Gbps вградени на дънната платка- 10GBase-T ports via Intel C622 с включени 2x RJ-45 https://www.supermicro.com/products/system/1U/6019/SSG-6019P-ACR12L.cfm?parts=SHOW#jump - Всеки изчислителен модул на SuperMicro модел SYS-1029P-WTR включва два порта за LAN Ethernet със скорост за обмен 10Gbps както следва Supermicro Supermicro 10-Gigabit Ethernet AOC-STGN-i2S 10G 2 	

000068

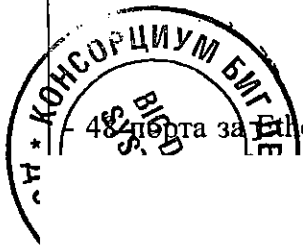


Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависимост
<p>- елемент за backup на хранване при RAID контролера;</p> <p>- софтуер за управление на хардуера на модула;</p> <p>- Физическа организация – за инсталиране в рак с размери не по-големи от 2U, за предпочитане 1U.</p>		<p>ports SFP+</p> <p>https://www.supermicro.com/wftp/Networking_Drivers/CDR-NIC_1.52_for_Add-on_NIC_Cards/MANUALS/AOC-STGN-i1S-i2S.pdf</p> <p>- Всеки изчислителен модул на SuperMicro модел SYS-1029P-WTR има елемент за backup на хранване при RAID контролера.</p> <p>- Софтуерът за управление на Super Micro е Supermicro Server Manager (SSM) (DataCenter Management Package).</p> <p>https://www.supermicro.com/en/solutions/management-software/supermicro-server-manager</p> <p>- Изчислителните модули на SuperMicro модел SYS-1029P-WTR са с размери 1U (1U Rackmount).</p> <p>https://www.supermicro.com/products/system/1U/1029/SYS-1029P-WTR.cfm</p>	

690000



Изискване	Изпълнение (Да/Не)	Начин на реализация	Завис
<p>- В този изчислителен модул трябва да се зареди системен софтуер, който да извършва функциониране по управление на мрежовите портове в Специализираната система под ключ, с оглед осигуряване контрол на достъпа.</p>		<p>- В този изчислителен модул ще бъде зареден системен софтуер, който ще извършва функциониране по управление на мрежовите портове в Специализираната система под ключ, с оглед осигуряване контрол на достъпа.</p>	<p>✓</p>
<p>Т.2.4.2 Комуникационни модули – комутатори (switches) на Управляващата Специализирана хардуерна система за обработка на данни.</p> <p>А) Основен комуникационен модул на Управляващата Специализирана хардуерна система за обработка на данни – 1 брой. Той да бъде управляем комутатор на минимално ниво 2 от OSI Референтната архитектурата, със следното минимално количество портове:</p> <p>48 порта за Ethernet 10Gbps скорост за обмен с</p>	<p>ДА</p>	<p>Предлагаме Комуникационните модули – комутатори (switches) на Управляващата Специализирана хардуерна система за обработка на данни да бъдат на SuperMicro.</p> <p>А) За 1 брой основен комуникационен модул на Управляващата Специализирана хардуерна система за обработка на данни предлагаме Super Micro модел SSE-X3348S. Layer 2 / 3 / 4. Този модел е управляем комутатор със следните портове:</p> <p>- Комутатор Super Micro модел SSE-X3348S има 48 порта за Ethernet</p>	<p>✓</p>



000070

✓

✓

Изискване	Изпълнение (Да/Не)	Начин на реализация	За
<p>интерфейс SFP+;</p> <p>- 2 порта uplink връзки за Ethernet 40Gbps скорост за обмен с интерфейс QSFP+;</p> <p>- 2 порта за Ethernet 1Gbps скорост за обмен и интерфейс RJ45;</p> <p>Допълнителен комуникационен модул на Управляващата Специализирана хардуерна система за обработка на данни – 1 брой. Той да бъде управляем комутатор на минимално ниво 2 от OSI Референтната архитектура, със следното минимално количество портове:</p>		<p>10Gbps скорост за обмен с интерфейс SFP+</p> <p>https://www.supermicro.com/products/accessories/Networking/SSE-X3348S.cfm</p> <p>- Комутатор Super Micro модел SSE-X3348S има uplink минимум 2x40Gbps QSFP+</p> <p>https://www.supermicro.com/products/accessories/Networking/SSE-X3348S.cfm</p> <p>- Комутатор Super Micro модел SSE-X3348S има 2 порта с по 1Gbps скорост за обмен с интерфейс RJ45.</p> <p>https://www.supermicro.com/products/accessories/Networking/SSE-X3348S.cfm</p> <p>Б) За 1 брой допълнителен комуникационен модул на Управляващата Специализирана хардуерна система за обработка на данни предлагаме предлагаме Allied Telesis модел AT-x510L-28GT Layer 3. Този модел е управляем комутатор, с включени следните</p>	<p>За</p>

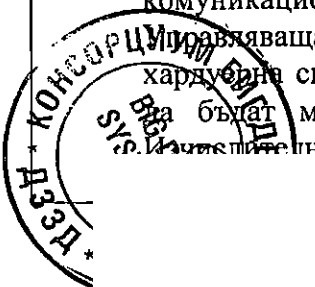
Изискване	Изпълнение (Да/Не)	Начин на реализация	Заб.
<p>- 24 порта за Ethernet 1Gbps скорост за обмен с ъс RJ45;</p> <p>- порт за връзка с по-висока йерархия комуникационни модули за Ethernet 1-10Gbps скорост за обмен с интерфейс SFP+.</p>		<p>портове.</p> <p>- Комутатор Allied Telesis модел AT-x510L-28GT притежава 24 порта за Ethernet 1Gbps скорост за обмен с интерфейс RJ45.</p> <p>https://www.alliedtelesis.com/products/switches/x510l-28gt</p> <p>- Комутатор Allied Telesis модел AT-x510L-28GT има 4 порт за връзка с по-висока йерархия комуникационни модули за Ethernet 1-10Gbps скорост за обмен с интерфейс SFP+.</p> <p>https://www.alliedtelesis.com/products/switches/x510l-28gt</p>	
<p>Т.2.4.3. Физическа организация на Управляващата Специализирана хардуерна система за обработка на данни</p>	ДА	<p>Физическата организация на Управляващата Специализирана хардуерна система за обработка на данни ще бъде направена така както е описано в изискването на Възложителя, както следва:</p>	

270000



Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависим
<ul style="list-style-type: none"> - Управляващата Специализирана хардуерна система за обработка на данни трябва да бъде монтирана и инсталирана за работа в rack с размери 42U – rack №11. От този rack да се осигурят 2 броя кабели за скорост на обмен 40Gbps, с дължина минимум 20 метра, чрез които да се осигури връзка с Комутатор с портове от по 40Gbps, разположен в Rack №10. - Rack 11 трябва да бъде със съответното управление – KVM switch за управление на 12 сървъра с дисплей, клавиатура и мишка и кабели за управление на 12 компютъра. Тези дисплей, клавиатура и мишка да са в размерите сумарно на 1U. - Всички изчислителни модули и комуникационни модули на Управляващата Специализирана хардуерна система за обработка на данни трябва бъдат монтирани в посочения rack. Изчислителните модули да бъдат свързани муникационен модул с 		<ul style="list-style-type: none"> - Управляваща Специализирана хардуерна система за обработка на данни ще бъде доставена и инсталирана за работа в Rack 11, с размери 42U. От този rack ще бъдат осигурени 2 броя кабели за скорост на обмен 40Gbps, с дължина 20 метра, чрез които ще бъде осигурена връзката с Комутатор с портове от по 40Gbps, който е разположен в Rack 10. - Предлагаме в Rack №11 да бъде разположен KVM switch за управление на 16 сървъра с дисплей, клавиатура и мишка и кабели за управление на 16 компютъра. Като предложените дисплей, клавиатура и мишка да са с размера на 1U. - Всички изчислителни модули и комуникационни модули на Управляващата Специализирана хардуерна система за обработка на данни ще бъдат монтирани в rack №11, както е посочено и в изискването за 1U. Изчислителните модули ще бъдат 	<ul style="list-style-type: none"> - За да бъде монтирана и инсталирана - Управляващата Специализирана хардуерна система обработка на данни осигури от Възлож - Достъп до помещение специалист ангажирани монтажа и инсталцията; - Необходимото стабилно и достатъчно по мощност захранване на помещението; - Необходимата и достатъчна по мощност климатизация

000073



Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависим
<p>необходимите LAN кабели за 10Gbps скорост за обмен с интерфейс SFP+. Допълнителният комуникационен модул да бъде свързан към Основния комуникационен модул. При инсталирането в рака, комуникационните модули да се поставят отгоре в рака, след които да се поставят изчислителните модули. Да се доставят минимално 12 броя LAN кабели за 1Gbps скорост за обмен с интерфейс RJ45 за връзки в рамките на рака. Всички LAN кабели да се наименоуват за това с кои устройства се прави връзка, и се подредят и организират в рака с оглед улеснено техническо обслужване. Да се доставят 12 разделителни носещи плочи за поставяне в рак, върху които да се поставят евентуални tower-based сървъри и компютри.</p>		<p>свързани с Основния комуникационен модул с необходимите за тази връзка LAN кабели за 10Gbps скорост за обмен с интерфейс SFP+. Допълнителният комуникационен модул ще бъде свързан към Основния комуникационен модул. При инсталиране на техниката в рака, комуникационните модули ще бъдат поставени най - отгоре в рака, след което ще бъдат поставени изчислителните модули. Ще бъдат доставени 12 броя LAN кабели за осъществяване на 1Gbps скорост за обмен с интерфейс RJ45, което е необходимо за връзки в рамките на рака. Всички LAN кабели ще бъдат наименовани за това с кои устройства се прави връзка, и ще бъдат старателно подредени и организирани в рака за улеснение на техническо обслужване. Ще бъдат доставени 12 разделителни носещи плочи необходими за поставяне в съответния рак, върху който могат се поставят евентуално и tower-based сървъри и компютри</p>	

000074



111

Изискване	Изпълнение (Да/Не)	Начин на реализация	Зави
<p>- Ракът да бъде доставен с UPS система. UPS да бъде с размери до 4U, като доставчикът да определи неговите технически параметри за да се осигурява токоподаване на всички устройства в рака минимум 3 минути след спиране на електричеството.</p>		<p>- Ракът ще бъде доставен с UPS система модел GTEC AP160N 1. С мощност 6kVA/5,4kW ще бъде с размери до 3U, като моделът осигурява токоподаване на всички устройства в рака за минимум 3 минути след спиране на електричеството.</p> <p>https://gtec.pl/produkt/ap160n/ap160n-6kva/</p>	<p>11</p>
<p>Управляващата специализирана хардуерна система за обработка на данни ще съдържа следните компоненти:</p>		<ol style="list-style-type: none"> 1. МодулNameNode модел SuperMicro модел SYS-1029P-WTR – 3 броя 2. МодулManagementNode модел SuperMicro SYS-1029P-WTR – 1 брой 3. Сървър за УАЛС модел SuperMicro SSG-6019P-ACR12L – 1 брой 4. Сървър за МУД модел SYS-1029P-WTR – 1 брой 5. Основен комуникационен модул – Super Micro модел SSE-X3348S – 1 брой 6. Допълнителен комуникационен модул - Allied Telesis модел AT-x510L-28GT – 1 брой 7. KVM switch – 1 брой 8. UPS система модел GTEC AP160N 1 – 1 брой 	<p>U</p>

000075



Изискване	Изпълнение (Да/Не)	Начин на реализация	Завис
		9. Кабели за скорост на обмен 40Gbps с дължина над 20 метра – 2 броя 10. LAN кабели за осъществяване на 1Gbps скорост за обмен с интерфейс RJ45 – 12 броя 11. Разделителни носещи плочи – 12 броя 12. Физическата организация за разполагането и инсталирането на Управляващата Специализирана система за обработка на данни.	11 —
<p>Комуникационна Специализирана система за обмен на данни.</p> <p>- Тази система трябва да бъде изградена в отделен рак – Рак №10.</p> <p>- Комуникационната Специализирана хардуерна система за обмен на данни трябва да се състои от 2 броя комуникационни модула (switches) – Централен комуникационен модул, Допълнителен комуникационен модул, 1 ичнен изчислителен модул</p>	<p>ДА</p>	<p>Предлагаме Комуникационна Специализирана хардуерна система за обмен на данни да бъде изградена с оборудване на Super Micro.</p> <p>- Описаната Комуникационна система ще бъде изградена в отделен рак №10.</p> <p>- Предлагаме Комуникационната Специализирана хардуерна система за обмен на данни трябва да се състои от 2 броя комуникационни модула (switches) – Централен комуникационен модул SuperMicro модел SSE-C3632S, Допълнителен</p>	<p>✓</p> <p>11</p>

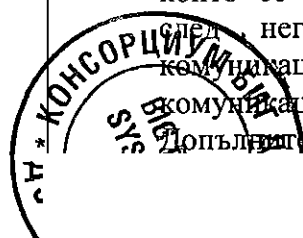
0078



Handwritten signature or mark.

Изискване	Изпълнение (Да/Не)	Начин на реализация	3
<p>Модул EdgeNode, 1 брой изчислителен сървър за LDAP и 1 брой IPS устройство, които трябва да се доставят от участника. Комуникационната Специализирана хардуерна система за обмен на данни трябва да се изгради и с 1 брой Комуникационен модул на Центъра за компетентност, предоставен от заявителя.</p> <p>- Йерархически от гледна точка на комуникационни връзки, Комуникационната Специализирана хардуерна система за обмен на данни трябва да представлява най-отгоре Комуникационния модул на Центъра за компетентност (в който се подава входния трафик към всички елементи на Специализираната система под ключ), към който се включва IPS устройство, като след него се включва Централния комуникационен модул, след Централния комуникационен модул се включва Допълнителния комуникационен модул и / се включва изчислителен</p>		<p>комуникационен модул Allied Telesis модел AT-x510L-28GT, 1 брой Граничен изчислителен модул Модул EdgeNode SuperMicro модел SYS-1029P-WTR, 1 брой изчислителен сървър за LDAP Super Micro SYS-1029P-WTR и 1 брой IPS устройство WatchGuard M4600. Специализирана хардуерна система за обмен на данни ще се изгради и с 1 брой Комуникационен модул на Центъра за компетентност, предоставен от заявителя.</p> <p>- В Комуникационната Специализирана хардуерна система за обмен на данни най-отгоре ще бъде поставен Комуникационния модул на Центъра за компетентност (в който ще се подава входния трафик към всички елементи на Специализираната система под ключ), към който се включва IPS устройство WatchGuard M4600, като след него се включва Централния комуникационен модул SuperMicro модел SSE-C3632S, след Централния комуникационен модул се включва Допълнителния</p>	

000077



Изискване	Изпълнение (Да/Не)	Начин на реализация	Зави
<p>сървър за LDAP. Граничният модул МодулEdgeNode трябва да се свърже към Комуникационния модул на Центъра за компетентност. По принцип IPS устройството (Intrusion Prevention System гема за предпазване от нежелано влизане) има за цел да анализира входящия трафик в Специализираната централизирана система под ключ Nadoor и когато трафикът не отговаря на конфигурираните условия от гледна точка на информационна сигурност, да издаде сигнал към Сървъра за МУД (разположена в Рак № 11), за да затвори той определен порт на даден комуникационен възел. Комуникационният модул на Центъра за компетентност се доставя от заявителя.</p>		<p>комуникационен модул Allied Telesis модел AT-x510L-28GT и най-отдолу ще се включи изчислителен сървър за LDAP Super Micro SYS-1029P-WTR. Граничният модул МодулEdgeNode SuperMicro модел SYS-1029P-WTR ще се свърже към Комуникационния модул на Центъра за компетентност. По принцип IPS устройството WatchGuard M4600 (Intrusion Prevention System – Система за предпазване от нежелано проникване) има за цел да анализира входящия трафик в Специализираната централизирана система под ключ Nadoor и когато трафикът не отговаря на конфигурираните условия от гледна точка на информационна сигурност, ще издаде сигнал към Сървъра за МУД (разположена в Рак № 11), за да затвори той определен порт на даден комуникационен възел. Комуникационният модул на Центъра за компетентност ще бъде доставен от заявителя.</p>	
<p>Т.2.5.1. Централен комуникационен модул</p>	<p>ДА</p>		

000078



7

11

✓

✓

✓

Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависимост
<p>- Централният комуникационен модул – комутатор (switch) трябва да бъде управляем комутатор на минимално ниво 2 от OSI Референтната архитектурата, със следното минимално количество портове:</p> <p>- 24 порта за Ethernet 40Gbps скорост за обмен с интерфейс QSFP+;</p> <p>- 2 порта за Ethernet 1Gbps скорост за обмен с интерфейс RJ45;</p>		<p>- За 1 брой Централен комуникационен модул на Специализирана хардуерна система за обработка на данни предлагаме SuperMicro модел SSE-C3632S. Layer 2 / 3/4. Този модел е управляем комутатор, с включени следните портове.</p> <p>https://www.supermicro.com/products/accessories/Networking/SSE-C3632S.cfm</p> <p>- Комутатор SuperMicro модел SSE-C3632S притежава 32 порта за 40Gbps скорост за обмен с интерфейс QSFP+.</p> <p>https://www.supermicro.com/products/accessories/Networking/SSE-C3632S.cfm</p> <p>- Комутатор SuperMicro модел SSE-C3632S има 2 порта за Ethernet 1Gbps скорост за обмен и интерфейс RJ45 с включена опция за целта.</p>	

000079



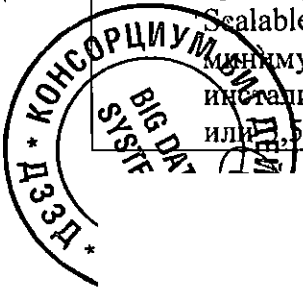
Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависимости
		https://www.supermicro.com/products/accessories/Networking/SSE-C3632S.cfm	
<p>5.2. Допълнителен комуникационен модул</p> <p>Допълнителният комуникационен модул – мутатор (switch) трябва да бъде управляем мутатор на минимално ниво 2 от OSI ферентната архитектурата, със следното минимално количество портове:</p> <p>- 24 порта за Ethernet 1Gbps скорост за обмен;</p>	ДА	<p>- За 1 брой Допълнителен комуникационен модул на Специализирана хардуерна система за обработка на данни предлагаме Allied Telesis модел AT-x510L-28GT. Този модел е управляем комутатор, с включени следните портове.</p> <p>https://www.alliedtelesis.com/products/switches/x510l-28gt</p> <p>- Комутатор Allied Telesis модел AT-x510L-28GT притежава 24 порта за Ethernet 1Gbps скорост за обмен;</p> <p>https://www.alliedtelesis.com/products/switches/x510l-28gt</p> <p>- Комутатор Allied Telesis модел AT-</p>	<p>AA</p>

000030



Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависим
<ul style="list-style-type: none"> - 1 порт за връзка с по-висока йерархия комуникационни модули за Ethernet 1-10Gbps скорост за обмен с интерфейс SFP+. - Допълнителният комуникационен възел трябва да бъде свързан към Централния комуникационен възел, като при необходимост от преобразователи на интерфейсите връзки, да се доставят необходимите такива. 		<p>x510L-28GT да притежава 4 порт за връзка с по-висока йерархия комуникационни модули за Ethernet 1-10Gbps скорост за обмен с интерфейс SFP+.</p> <p>https://www.alliedtelesis.com/products/switches/x510l-28gt</p> <ul style="list-style-type: none"> - Допълнителният комуникационен възел трябва ще бъде свързан към Централния комуникационен възел, като при необходимост от преобразователи на интерфейсите връзки, ще бъдат доставени необходимите такива. 	<p>1/1</p> <p>✓</p>
<p>2.5.3. Изчислителен сървър за LDAP</p> <p>Изчислителният сървър за LDAP трябва да има:</p> <ul style="list-style-type: none"> - Техническа конфигурация - минимум 2 брой процесор от тип Intel Xeon Bronze Scalable Processors, оперативна памет минимум 128GB, и минимум 5 инсталирани позиции за включване на 2,5" или 3,5" дискове със сменяемост по време 	ДА	<p>Преплагаме Изчислителният сървър за LDAP да бъде изграден от SuperMicro модел SYS-1029P-WTR.</p> <ul style="list-style-type: none"> - В изчислителния сървър на SuperMicro модел SYS-1029P-WTR включва 2 брой процесори тип 2 Intel® Xeon® Bronze 3106 Processor 8-core 1.70GHz 11.00MB Cache, оперативната памет е тип ECC. 	<p>✓</p>

180000



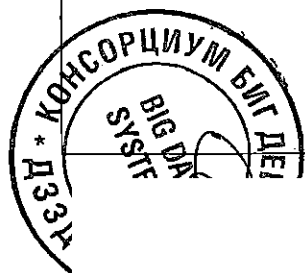
Изискване	Изпълнение (Да/Не)	Начин на реализация	Зави
<p>на опериране (hot-swap drive bays) в RAID1/5 с обем минимум по 0,5 TB;</p> <p>- 2 порта за Ethernet 1Gbps скорост за обмен.</p>		<p>Сървърът ще съдържа 8 памети x 16GB PC4-21300 2666MHz DDR4 ECC. Това прави 128 GB, ще бъдат инсталирани 5 позиции за включване на 2,5" дискове със сменяемост по време на опериране (hot-swap drive bays) с обем Samsung Samsung SSD PM871b 512GB, 2.5 Inch, работещи в архитектура RAID 1/5.</p> <p>https://www.supermicro.com/products/system/1U/1029/SYS-1029P-WTR.cfm ; https://ark.intel.com/content/www/us/en/ark/products/123540/intel-xeon-bronze-3106-processor-11m-cache-1-70-ghz.html ; https://store.supermicro.com/16gb-ddr4-2666-mem-dr416l-sl04-er26.html</p> <p>- Изчислителния сървър на SuperMicro модел SYS-1029P-WTR включва два порта за LAN Ethernet със скорост за обмен 1Gbps и 10Gbps вградени на дънната платка- 10GBase-T ports via Intel C622 включително 2x RJ-45.</p> <p>https://www.supermicro.com/products/system/1U/6019/SSG-6019P-ACR12L.cfm?parts=SHOW#jump</p>	

000082



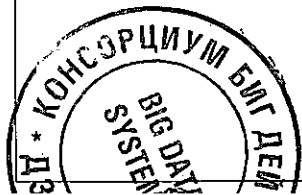
SM

Изискване	Изпълнение (Да/Не)	Начин на реализация	3
<ul style="list-style-type: none"> - Изчислителният сървър за LDAP трябва да бъде свързан към Допълнителния комуникационен възел. - В Изчислителният сървър за LDAP трябва да се зареди системен софтуер, който основно да изпълнява ролята на LDAP сървър. 		<ul style="list-style-type: none"> - Изчислителният сървър за LDAP ще бъде свързан към Допълнителния комуникационен възел. - В Изчислителния сървър за LDAP ще бъде зареден системен софтуер, който основно ще изпълнява ролята на LDAP сървър. 	
<p>2.5.4. Граничен изчислителен модул МодулEdgeNode</p> <p>Граничният модул МодулEdgeNode трябва да се състои от:</p> <ul style="list-style-type: none"> - минимум 2 броя процесори от тип Intel Xeon Bronze Scalable Processors, всеки от тях с минимум 8 ядра, работещи с честота минимум 1,7GHz; 	ДА	<p>Преплагаме Граничен изчислителен модул МодулEdgeNode да бъде изграден от SuperMicro модел SYS-1029P-WTR.</p> <ul style="list-style-type: none"> - В Граничен изчислителен модул МодулEdgeNode на SuperMicro модел SYS-1029P-WTR включва 2 броя процесори тип 2 Intel® Xeon® Bronze 3106 Processor 8-core 1.70GHz 11.00MB Cache. <p>https://ark.intel.com/content/www/us/en/ark/products/123540/intel-xeon-bronze-3106-processor-11m-cache-1-70-ghz.html</p>	✓



Изискване	Изпълнение (Да/Не)	Начин на реализация	З
<p>- оперативна памет минимум 128GB с честота на работа минимум 2600 MHz;</p> <p>- минимум 8 позиции за включване на 2,5" или 3,5" дискове със сменяемост по време на опериране (hot-swap drive bays) с инсталирани дискове с минимум 2.0TB с архитектура SAS 3.0 12.0Gb/s със съответен SAS-RAID контролер, работещи в архитектура RAID 1/5/6, които да служат за системен софтуер и за потребителски данни;</p> <p>- оптическо устройство DVD-RW / 24;</p>		<p>- оперативната памет е тип ECC. Сървърът ще съдържа 8 памети x 16GB PC4-21300 2666MHz DDR4 ECC. Това прави 128 GB.</p> <p>https://store.supermicro.com/16gb-ddr4-2666-mem-dr416l-sl04-er26.html</p> <p>- ще бъдат инсталирани 8 позиции за включване на 2,5" дискове със сменяемост по време на опериране (hot-swap drive bays) с обем 2.0TB SAS 3.0 12.0Gb/s 7200RPM - 2.5" - Seagate Exos 7E2000 Series (512e), работещи в архитектура RAID 1/5/6.</p> <p>https://www.supermicro.com/products/system/1U/1029/SYS-1029P-WTR.cfm</p> <p>- В Граничен изчислителен модул МодулEdgeNode на SuperMicro SYS-1029P-WTR се включва оптическо устройство - LG Slim 8x DVD-RW / 24x CDR Combo (SATA)</p> <p>https://www.supermicro.com/products/system/1U/1029/SYS-1029P-WTR.cfm</p>	3

000084



Изискване	Изпълнение (Да/Не)	Начин на реализация	
<p>- минимум 2 порта за LAN Ethernet със скорост за обмен 1Gbps и интерфейс RJ45;</p> <p>- минимум 1 порт за LAN Ethernet със скорост за обмен 40Gbps и интерфейс QSFP+;</p>		<p>- В Граничен изчислителен модул на SuperMicro модел SYS-1029P-WTR включва два порта за LAN Ethernet със скорост за обмен 1Gbps и 10Gbps вградени на дънната платка- 10GBase-T ports via Intel C622 включително 2x RJ-45.</p> <p>https://www.supermicro.com/products/system/1U/6019/SSG-6019P-ACR12L.cfm?parts=SHOW#jump</p> <p>- В Граничен изчислителен модул SuperMicro модел SYS-1029P-WTR има включен Supermicro AOC-S40G-i2Q,2-port 40GbE controller, based on Intel Fortville XL710</p> <p>ftp://supermicro.com/Networking_Drivers/CDR-NIC_1.53_for_Add-on_NIC_Cards/MANUALS/datasheet-AOC-S40G-i1Q_i2Q.pdf</p> <p>- В Граничен изчислителен модул SuperMicro SYS-1029P-WTR има включено backup на захранване при</p>	

000085



Изискване	Изпълнение (Да/Не)	Начин на реализация
<ul style="list-style-type: none"> - елемент за backup на захранване при RAID контролера; - софтуер за управление на хардуера на модула; - Физическа организация – за инсталиране в рак, с размери не по-големи от 2U, за предпочитане 1U. - Граничният модул МодулEdgeNode трябва да бъде свързан към Комуникационния модул на Центъра за компетентност. 		<p>RAID контролера;</p> <ul style="list-style-type: none"> - Софтуерът за управление на Super Micro е Supermicro Server Manager (SSM) (DataCenter Management Package). <p>https://www.supermicro.com/en/solutions/management-software/supermicro-server-manager</p> <ul style="list-style-type: none"> - Граничен изчислителен модул SuperMicro SYS-1029P-WTR е с размер 1U. <p>https://www.supermicro.com/products/system/1U/1029/SYS-1029P-WTR.cfm</p> <ul style="list-style-type: none"> - Граничният модул МодулEdgeNode ще бъде свързан към Комуникационния модул на Центъра за компетентност.
<p>Т. 2.5.5 IPS Устройство</p> <p>IPS Устройство трябва да има следните</p>	<p>ДА</p>	<p>Устройството на WatchGuard M4600 покрива всички описани в документацията функционалности както следва:</p>

000086



Изискване	Изпълнение (Да/Не)	Начин на реализация	За
<p>минимални функции:</p> <ul style="list-style-type: none"> самостоятелно физическо устройство с два броя интерфейси. <p>пропускателна способност при включена защита Intrusion Prevention System (IPS) да бъде поне 13 Gbps;</p> <ul style="list-style-type: none"> защита от атаки тип „SQLInjections” 		<ul style="list-style-type: none"> Устройството WatchGuard M4600 е физическо самостоятелно устройство с два броя интерфейси. https://www.rdggroup.co.za/images/wg/wg_firebox_m4600-m5600_ds.pdf Устройството WatchGuard M4600 има пропускателна способност при включена защита Intrusion Prevention System (IPS) е 13 Gbps https://www.rdggroup.co.za/images/wg/wg_firebox_m4600-m5600_ds.pdf - стр.2 Устройството WatchGuard M4600 със функционалността си за Intrusion Prevention System защитава от атаки тип „SQLInjections”. https://www.watchguard.com/wgrd-products/security-services/intrusion-prevention-service Устройството WatchGuard M4600 притежава функционалност за “Gateway AntiVirus” и „Spam Preventions”, с 	<p>За</p> <p>✓</p>

000087



Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависимост
<ul style="list-style-type: none"> защита от вируси и спам; защита от атаки тип „Cross Site Scripting“; защита от атаки тип DoS (Denial of Services); защита от атаки използващи гласов 		<p>които осъществява защитата от вируси и спам.</p> <p>https://www.watchguard.com/wgrd-products/security-services/gateway-antivirus https://www.watchguard.com/wgrd-products/security-services/spam-prevention</p> <ul style="list-style-type: none"> Устройството WatchGuard M4600 със функционалността си за Intrusion Prevention System защитава от атаки тип „Cross Site Scripting“. <p>https://www.watchguard.com/wgrd-products/security-services/intrusion-prevention-service</p> <ul style="list-style-type: none"> Устройството WatchGuard M4600 осъществява защита срещу атаки тип DoS (Denial of Services). <p>https://www.rdggroup.co.za/images/wg/wg_fir_ebox_m4600-m5600_ds.pdf</p> <ul style="list-style-type: none"> Устройството WatchGuard M4600 осъществява защита срещу атаки използващи гласов трафик - H.323, SIP, call setup and session security 	

880000



Изискване	Изпълнение (Да/Не)	Начин на реализация	Заб.
<p>трафик – с протоколи H.323, SIP;</p> <ul style="list-style-type: none"> да идентифицира и блокира злонамерен трафик; да притежава методи за криптиране например като DES, 3DES, AES 128-, 192-, 256-bit; да разпознава и управлява минимум 300 броя WLAN 		<p>https://www.rdggroup.co.za/images/wg/wg_firebox_m4600-m5600_ds.pdf</p> <ul style="list-style-type: none"> Устройството WatchGuard M4600 може да идентифицира и блокира злонамерен трафик на база уеб репутация (Reputation-Based Threat Prevention). <p>https://www.watchguard.com/wgrd-products/security-services/reputation-based-threat-prevention</p> <ul style="list-style-type: none"> Устройството WatchGuard M4600 притежава следните алгоритми за криптиране: AES 128, 192, 256 bit, DES, 3 DES <p>https://www.rdggroup.co.za/images/wg/wg_firebox_m4600-m5600_ds.pdf</p> <ul style="list-style-type: none"> Устройството WatchGuard M4600 разпознава и управлява 1000 броя VLAN. <p>https://www.rdggroup.co.za/images/wg/wg_firebox_m4600-m5600_ds.pdf</p>	<p>1 -</p> <p>✓</p>

680000



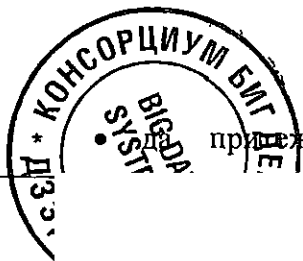
Изискване	Изпълнение (Да/Не)	Начин на реализация	Забелески
<ul style="list-style-type: none"> да притежава функционалност за откриване, наблюдение и контрол на трафика на приложения (Application Control); да може да се интегрира с Активна директория/LDAP; да може да бъдат извлечени статистически доклади/отчети за събития и активности; 		<ul style="list-style-type: none"> Устройството WatchGuard M4600 притежава функционалността „Application Control“, чрез която открива, наблюдава и контролира трафика към приложенията. https://www.watchguard.com/wgrd-products/security-services Устройството WatchGuard M4600 може да се интегрира успешно с Активна директория/LDAP. https://www.rdggroup.co.za/images/wg/wg_firebox_m4600-m5600_ds.pdf Устройството WatchGuard притежава конзола за управление, от където могат да бъдат извлечени най – различни статистически доклади за събития и активности. https://www.rdggroup.co.za/images/wg/wg_firebox_m4600-m5600_ds.pdf Устройството WatchGuard M4600 притежава функционалността да 	<p>1</p> <p>✓</p>

000000



Изискване	Изпълнение (Да/Не)	Начин на реализация	Заб.
<ul style="list-style-type: none"> • да може да изпраща уведомление за критични събития по е-мейл; • да притежава централизирана конзола за управление с уеб-интерфейс; • да притежава сертификати за информационна сигурност, например като ICSA Firewall, ICSA IPSec VPN, CC EAL4+, FIPS 140-2 		<p>изпраща уведомление за критични събития по е-мейл;</p> <p>https://www.rdggroup.co.za/images/wg/wg_firebox_m4600-m5600_ds.pdf</p> <ul style="list-style-type: none"> • Устройството WatchGuard M4600 притежава централизирана конзола за управление с уеб интерфейс (Web UI). <p>https://www.rdggroup.co.za/images/wg/wg_firebox_m4600-m5600_ds.pdf</p> <ul style="list-style-type: none"> • Устройството WatchGuard M4600 притежава сертификати за информационна сигурност както следва: ICSA IPSec VPN, FIPS 140-2, CC EAL4+ , ICSA Firewall. <p>https://www.rdggroup.co.za/images/wg/wg_firebox_m4600-m5600_ds.pdf</p> <ul style="list-style-type: none"> • Устройството WatchGuard M4600 притежава функционалността „WebBlocker URL Filtering” за филтриране на уеб съдържанието, като се прави проверка с над 130 URL 	

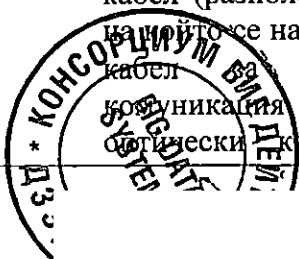
000091



притежава функционалност за

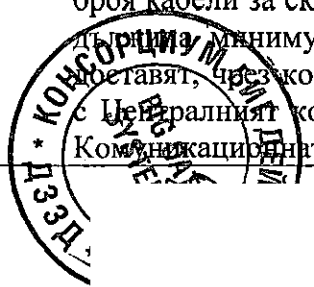
Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависим
<p>филтриране на уеб съдържанието.</p> <ul style="list-style-type: none"> да притежава минимум 2 порта по 40Gb с интерфейс QSFP+; 		<p>категории.</p> <p>https://www.watchguard.com/wgrd-products/security-services/webblocker-url-filtering</p> <ul style="list-style-type: none"> Устройството WatchGuard M 4600 притежава 2 Port 40Gb QSFP+ Fiber Module <p>https://www.rdggroup.co.za/images/wg/wg_firebox_m4600-m5600_ds.pdf</p>	
<p>2.5.6. Физическа организация на Комуникационната Специализирана хардуерна система за обмен на данни.</p> <ul style="list-style-type: none"> Комуникационният модул на Центъра за компетентност (предоставен от заявителя) трябва да бъде свързан с 2 броя оптически кабели към Панела за входен оптически кабел (разположен на стената на залата), на който се намира терминиран оптически кабел за връзка с Централната Комуникационна на УНСС. Тези 2 броя оптически кабели, наричани „външни 	ДА	<p>Предлагаме Физическата организация на Комуникационната Специализирана хардуерна система за обмен на данни да бъде осъществена както следва:</p> <ul style="list-style-type: none"> Комуникационният модул на Центъра за компетентност ще бъде предоставен от заявителя и ще бъде свързан с 2 броя оптически кабели към Панела за входен оптически кабел (разположен на стената на залата по данни от заявителя), на който се намира терминиран оптически кабел за връзка 	

0000092



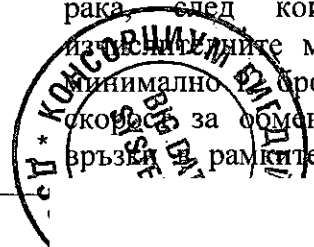
Изискване	Изпълнение (Да/Не)	Начин на реализация
<p>оптически кабели“, които трябва да се доставят, от едната страна трябва да имат интерфейс QSFP+ за връзка към Комуникационният модул на Центъра за компетентност, а от другата страна към Панела за входен оптически кабел и да бъдат с дължина минимум 30 метра. Комуникационният модул на Центъра за компетентност трябва да се инсталира в същия рак – Рак №10, където се инсталира Комуникационната Специализирана хардуерна система.</p> <p>- Комуникационната Специализирана хардуерна система за обмен на данни трябва да бъде монтирана и инсталирана за работа в рак с размери 42U – Рак №10. Комуникационният модул на Центъра за компетентност трябва да се свърже чрез 2 броя кабели за скорост на обмен 40Gbps, с дължина минимум 3 метра, които да се доставят, чрез които да се осигури връзка с Централният комуникационен модул на Комуникационната Специализирана</p>		<p>с Централната комуникация на УНСС (по данни на заявителя). Тези 2 броя оптически кабели, ще се наричат „външни оптически кабели“, които ще бъдат доставени, от едната страна ще имат интерфейс QSFP+ за връзка към Комуникационният модул на Центъра за компетентност, а от другата страна към Панела за входен оптически кабел ще бъдат с дължина 30 метра. Комуникационният модул на Центъра за компетентност ще бъде инсталиран в Рак №10, където ще бъде инсталирана и Комуникационната Специализирана хардуерна система.</p> <p>- Комуникационната Специализирана хардуерна система за обмен на данни ще бъде монтирана и инсталирана за работа в рак с размери 42U – Рак №10. Комуникационният модул на Центъра за компетентност ще бъде свързан чрез 2 броя кабели за скорост на обмен 40Gbps, с дължина 3 метра, които да се доставят, чрез които ще се осигури връзка с Централният комуникационен модул на Комуникационната</p>

000099



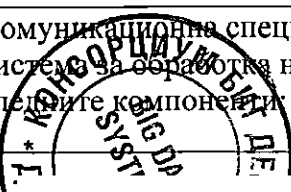
Изискване	Изпълнение (Да/Не)	Начин на реализация
<p>хардуерна система, в рамките на Рак№10.</p> <p>- Рак №10 трябва да бъде със съответното управление – KVM switch за управление на 12 сървъра с дисплей, клавиатура и мишка и кабели за управление на 12 компютъра. Тези дисплей, клавиатура и мишка да са в размерите сумарно на 1U.</p> <p>Всички изчислителни модули и комуникационни модули на Комуникационната Специализирана хардуерна система за обмен на данни да бъдат монтирани в посочения рак. При инсталирането в рака, Комуникационният модул на Центъра за компетентност и Централният комуникационен модул на Комуникационната Специализирана хардуерна система да се поставят отгоре в рака, след които да се поставят изчислителните модули. Да се доставят минимално броя LAN кабели за 1Gbps скорост за обмен с интерфейс RJ45 за връзка в рамките на рака. Всички LAN</p>		<p>Специализирана хардуерна система, в рамките на Рак№10.</p> <p>- В Рак №10 ще бъде разположено и съответното управление – KVM switch за управление на 16 сървъра с дисплей, клавиатура и мишка и кабели за управление на 16 компютъра. Тези дисплей, клавиатура и мишка ще бъдат с размерите на 1U.</p> <p>- Всички изчислителни модули и комуникационни модули на Комуникационната Специализирана хардуерна система за обмен на данни ще бъдат монтирани в посочения рак. При инсталирането в рака, Комуникационният модул на Центъра за компетентност и Централният комуникационен модул на Комуникационната Специализирана хардуерна система ще бъдат поставени отгоре в рака, след което ще бъдат поставени изчислителните модули. Ще бъдат доставени 12 броя LAN кабели за 1Gbps скорост за обмен</p>

000091



Изискване	Изпълнение (Да/Не)	Начин на реализация
<p>кабели да се наименоват за това с кои устройства се прави връзка, и се подредят и организират в рака с оглед улеснено техническо обслужване. Да се доставят 12 разделителни носещи плочи за поставяне в рак, върху които да се поставят евентуални tower-based сървъри и компютри.</p> <ul style="list-style-type: none"> - За този рак да бъде доставена UPS система. UPS системата да бъде с размери до 4U, като доставчикът да определи неговите технически параметри за да се осигурява токоподаване на всички устройства в рака минимум 3 минути след спиране на електричеството. - Самият Рак №10 като отделно устройство ще бъде предоставено от заявителя. 		<p>с интерфейс RJ45 за връзки в рамките на рака. Всички LAN кабели ще се наименоват за това с кои устройства се прави връзка, и ще се подредят и организират в рака за улеснено техническо обслужване. Ще бъдат доставени и 12 разделителни носещи плочи за поставяне в рак, върху които да се поставят евентуални tower-based сървъри и компютри.</p> <ul style="list-style-type: none"> - За този рак ще бъде доставена UPS система модел GTEC AP160N C мощност 6kVA/5,4kW. UPS системата ще бъде с размери до 3U, и ще предоставя възможност за осигуряване на токоподаване на всички устройства в рака за минимум 3 минути след спиране на електричеството. - Самият Рак №10 като отделно устройство ще бъде предоставено от заявителя.
<p>Комуникационна специализирана хардуерна система за обработка на данни ще съдържа следните компоненти:</p>	ДА	<p>1. Централен комуникационен модел – Super Micro модел SSE-C3632S – 1 брой</p>

000095

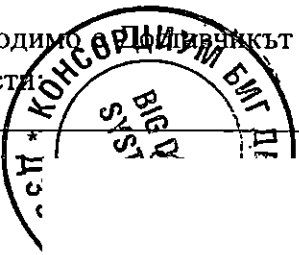


Изискване	Изпълнение (Да/Не)	Начин на реализация	Зам.
		<ol style="list-style-type: none"> 2. Комуникационен модул – предоставен от заявителя. 3. Допълнителен комуникационен модул - Allied Telesis модел AT-x510L-28GT – 1 брой 4. Граничен изчислителен модул Edge Node – SuperMicro модел SYS-1029P-WTR – 1 брой 5. Изчислителен сървър за LDAP – SuperMicro модел SYS-1029P-WTR 6. IPS устройство WatchGuard M4600 – 1 брой 7. KVM switch – 1 брой 8. UPS система модел GTEC AP160N – 1 брой 9. Оптически кабели с интерфейс QSFP+ дължина 30 метра – 2 броя 10. Кабели за скорост на обмен 40Gbps с дължина над 3 метра – 2 броя 11. LAN кабели за осъществяване на 1Gbps скорост за обмен с интерфейс RJ45 – 12 броя 	C

000096



Изискване	Изпълнение (Да/Не)	Начин на реализация	
		<p>12. Разделителни носещи плочи – 12 броя</p> <p>13. Физическата организация за разполагането и инсталирането на Комуникационната Специализирана система за обработка на данни.</p>	
<p>3. Хардуерна система на Прототипната под-система Hadoop за управление тестването на прототипни решения за големи данни</p> <p>- Прототипната под-система Hadoop за управление тестването на прототипни решения за големи данни трябва да се състои от 4 броя модули - 3 броя модули ComputerDataNode и 1 брой модул ComputerNameMgmtNode. Тези 4 модула от хардуерна гледна точка ще се предоставят от Заявителя. За 3-те броя модули ComputerDataNode Заявителят ще е предоставил по 1 брой LAN мрежова платка с 10Gbps скорост за обмен на данни.</p> <p>Необходимо е Заявителят да извърши следните дейности:</p>	ДА	<p>- Прототипната под-система Hadoop за управление тестването на прототипни решения за големи данни ще бъде доставена от Заявителя и ще се състои от 4 броя модули - 3 броя модули ComputerDataNode и 1 брой модул ComputerNameMgmtNode. Тези 4 модула от хардуерна гледна точка ще се предоставят от Заявителя. За 3-те броя модули ComputerDataNode Заявителят ще предостави по 1 брой LAN мрежова платка с 10Gbps скорост за обмен на данни.</p> <p>Съгласни сме да извършим следните дейности:</p>	



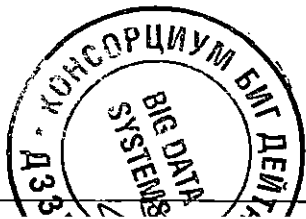
Изискване	Изпълнение (Да/Не)	Начин на реализация	Завършено
<ul style="list-style-type: none"> - достави 1 брой LAN мрежова платка с 10Gbps скорост за обмен на данни модул ComputerNameMgmtNode; инсталира в 4-те модула на Прототипната под-система, 4 броя LAN мрежова платка с 10Gbps скорост за обмен на данни – в трите модула ComputerDataNode и в модула ComputerNameMgmtNode; - Да достави 4 броя LAN кабели за скорост за обмен 10Gbps за свързване на 4-те модула на Прототипната под-система, с интерфейс SFP+, с дължина поне 3 метра; - Да свърже 4-те модула на Прототипната под-система към Специализираната централизирана система под ключ Hadoop (Пак №10) към Централния комуникационен модул; 		<ul style="list-style-type: none"> - Да доставим 1 брой LAN мрежова платка PCIe с 10Gbps скорост за обмен на данни за модул ComputerNameMgmtNode – предоставен от заявителя. - Да инсталираме 4-те модула на Прототипната под-система, в това число 4 броя LAN мрежова платка с 10Gbps скорост за обмен на данни – в трите модула ComputerDataNode и в модула ComputerNameMgmtNode - Да доставим 4 броя LAN кабели за скорост за обмен 10Gbps за свързване на 4-те модула на Прототипната под-система, с интерфейс SFP+, с дължина 3 метра. - Да свържем 4-те модула на Прототипната под-система към Специализираната централизирана система под ключ Hadoop (Пак №10) към Централния комуникационен 	

860000



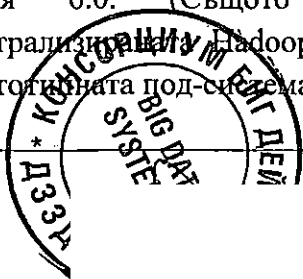
Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависен
<p>- Да инсталира и конфигурира Прототипната под-система Hadoop като автономно работеща Hadoop система – да работи като самостоятелна Cloudera/Hadoop система.</p>		<p>модул;</p> <p>- Да инсталираме и конфигурираме Прототипната под-система Hadoop като автономно работеща Hadoop система – така че да работи като самостоятелна Cloudera/Hadoop система.</p>	
<p>Хардуерна система на Прототипна под – система Hadoop за управление тестването на прототипни решения за големи данни ще съдържа следните компоненти:</p>		<ol style="list-style-type: none"> 1. ComputerDataNode – 3 броя – предоставени от заявителя 2. ComputerNameMgmtNode – 1 брой – предоставен от заявителя 3. LAN мрежова платка с 10Gbps – 1 брой – предоставен от заявителя 4. LAN мрежова платка PCIe с 10 Gbps – 1 брой 5. Инсталация на 4 те модула на Прототипната подсистема. 6. LAN мрежова платка с 10Gbps – 4 бр. 7. LAN кабели за осъществяване на 10 Gbps скорост за обмен с интерфейс SFP+ – 4 броя 	<p>6</p>

000099



Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависимост
		8. Свързване на 4 те модула на Прототипната подсистема. 9. Инсталиране и конфигуриране на Прототипната подсистема.	
4 Специализирано системно софтуерно осигуряване			
Т.4.1 Всички изчислителни модули на цялостната Специализирана централизирана система под ключ Hadoop (включително на „Централизираната Hadoop система“ и на „Прототипната под-система Hadoop“) трябва да са заредени със и да работят с операционна система от типа Linux, по-специално тези за подчертано корпоративни приложения SUSE Linux Enterprise или Red Hat Enterprise Linux, минимални версии според това за кои операционни системи са сертифицирани изчислителните модули и за кои работи управляващият софтуер Cloudera с минимална версия 6.0. (Същото важи както за „Централизираната Hadoop система“, така и за „Прототипната под-система Hadoop“).	ДА	Предлагаме всички изчислителни модули на цялостната Специализирана централизирана система под ключ Hadoop (включително на „Централизираната Hadoop система“ и на „Прототипната под-система Hadoop“) да са заредени със и да работят с операционна система от типа Linux, по-специално тези за подчертано корпоративни приложения SUSE Linux Enterprise или RedHat, сертифицирани за работа с изчислителните модули както и за управляващият софтуер Cloudera с минимална версия 6.0. Същото важи както за „Централизираната Hadoop система“, така и за „Прототипната под-система Hadoop“.	

007000



Изискване	Изпълнение (Да/Не)	Начин на реализация	3
<p>Специализирано системно софтуерно осигуряване ще съдържа следните компоненти:</p>	<p>ДА</p>	<p>1. За всеки сървър от основната система на Hadoop, описана в настоящето Техническо предложение предлагаме RedHat Enterprise Linux Academic Server, Self support (16 sockets) (up to 1 guest) with Smart Management.</p> <p>1. За всеки сървър описан в документацията за Прототипната система, предлагаме RED HAT Enterprise Linux Server Entry Level, Self Support (2 sockets).</p>	
<p>Г.4.2 Специализираната централизирана система под ключ (turnkey system) Hadoop да бъде инсталирана със софтуерната управляваща система Cloudera минимална версия 6.0 работеща със софтуерната рамка Hadoop с минимална версия 3.0, за да се оформи логически като едно цялостно изделие готово за използване след предаването му. (Същото важи и за прототипната под-система).</p>	<p>ДА</p>	<p>Преплагаме Специализираната централизирана система под ключ (turnkey system) Hadoop да бъде инсталирана със софтуерната управляваща система Cloudera версия 6.0 работеща със софтуерната рамка Hadoop с минимална версия 3.0. Същото ще бъде направено и за прототипната под-система.</p> <p>- Компонентите на Cloudera</p>	

000101



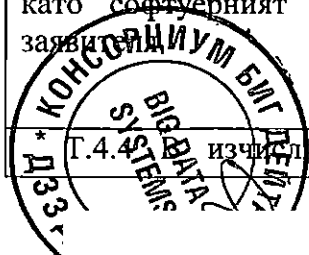
Изискване	Изпълнение (Да/Не)	Начин на реализация	Зави
<ul style="list-style-type: none"> - Компонентите на Cloudera да покриват следните корпоративни (enterprise) минимални функционалности: да допуска изграждане и управлява Hadoop клъстер с до 100 броя изчислителни модули МодулDataNode; да осигурява платформа за данни, включваща Apache Hadoop 3.0; - да осъществява автоматизирано създаване и управление на Hadoop клъстер; - да може да изгражда и управлява няколко Hadoop клъстера; - да осигурява конфигуриране и управление за висока надеждност (High Availability); - да предоставя инструменти за диагностика, за предупреждение и за опериране с използване на API; - притежава гъвкавост при интегриране в облачни среди; - има гъвкавост при създаване, 		<p>покриват следните корпоративни (enterprise) функционалности:</p> <ul style="list-style-type: none"> - допуска изграждане и управлява Hadoop клъстер с до 100 броя изчислителни модули МодулDataNode; - осигурява платформа за данни, включваща Apache Hadoop 3.0; - осъществява автоматизирано създаване и управление на Hadoop клъстер; - може да изгражда и управлява няколко Hadoop клъстера; - осигурява конфигуриране и управление за висока надеждност (High Availability); - предоставя инструменти за диагностика, за предупреждение и за опериране с използване на API; - притежава гъвкавост при интегриране в облачни среди; - има гъвкавост при създаване, 	

000102



Изискване	Изпълнение (Да/Не)	Начин на реализация	За
<p>премахване на модули, промяна на размера на Hadoop клъстера;</p> <ul style="list-style-type: none"> - да осигурява наличие на Rollback функционалност. 		<p>премахване на модули, промяна на размера на Hadoop клъстера;</p> <ul style="list-style-type: none"> - осигурява наличие на Rollback функционалност. <p>https://www.cloudera.com/fm/emea/kits/rs-components---big-data-fundamentals/cloudera-enterprise--the-modern-platform-for-machine-learning-an.html</p>	
<p>4.3. Изчислителният сървър за LDAP трябва да бъде минимум със следното системно програмно осигуряване:</p> <ul style="list-style-type: none"> - Системно осигуряване базирано на операционна система Windows Server минимална версия 2016, предоставена от заявителя; - Инсталиран Microsoft SQL Server минимум версия 2017 с активирана компонента PolyBase, като софтуерният продукт е предоставен от заявителя; 	ДА	<ul style="list-style-type: none"> - Системно осигуряване базирано на операционна система Windows Server минимална версия 2016, ще бъде предоставено от заявителя; - Ще бъде Инсталиран Microsoft SQL Server минимум версия 2017 с активирана компонента PolyBase, като софтуерният продукт е предоставен от заявителя. 	✓
Г.4.4. Изчислителният сървър за УАЛС	ДА	Предлагаме в изчислителният сървър за	

00000



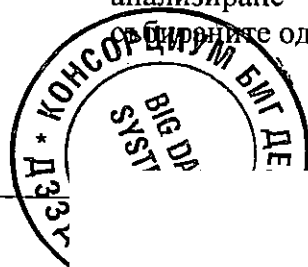
Изискване	Изпълнение (Да/Не)	Начин на реализация	З
<p>трябва да се зареди системно програмно осигуряване, което да изпълнява следните минимални функции:</p> <ul style="list-style-type: none"> да може да събира данни от регистрите (LOG файлове) за събития – наречени „одитни данни“ от всички изчислителни и комуникационни модули на Специализираната система под ключ, включително но не изчерпателно: информация за използване на изчислителните и комуникационни модули, уязвимости, цифрови идентичности и др. <p>Предоставя възможност за събиране на</p>		<p>УАЛС да бъде заредено системно програмно осигуряване (софтуерно решение) – Sentinel Enterprise на MicroFocus.</p> <ul style="list-style-type: none"> Решението “Sentinel Enterprise“ на Micro Focus притежава функционалността да събира данни за събития като информация за използване на изчислителните и комуникационни модули, включително за активи, уязвимости, цифрови идентичности и др. Притежава функционалността за „одитни данни“ от всички изчислителни и комуникационни модули на Специализираната система под ключ. <p>https://www.netiq.com/documentation/sentinel-82/pdfdoc/install/install.pdf</p> <ul style="list-style-type: none"> Решението на “Sentinel Enterprise“ на Micro Focus може да събира данни от 	3

000104



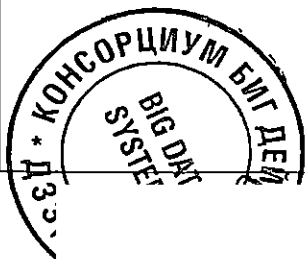
Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависимост
<p>данни чрез протокола SNMP;</p> <ul style="list-style-type: none"> да предоставя възможност за конфигуриране на източниците на събития за събирането на данни през интерфейса за управление; да притежава функционалност за анализиране и нормализация на събираните одитни данни; 		<p>SNMP.</p> <p>https://www.netiq.com/documentation/sentinel-82/pdffdoc/admin/admin.pdf</p> <ul style="list-style-type: none"> Решението на "Sentinel Enterprise" на Micro Focus притежава функционалността за конфигуриране на източниците на събития за събирането на данни през интерфейса за управление. <p>https://www.netiq.com/documentation/sentinel-82/pdffdoc/admin/admin.pdf</p> <ul style="list-style-type: none"> Решението на "Sentinel Enterprise" на Micro Focus включва функционалността за анализиране и нормализация на събираните одитни данни. <p>https://www.netiq.com/documentation/sentinel-82/pdffdoc/admin/admin.pdf</p>	

000105



Изискване	Изпълнение (Да/Не)	Начин на реализация	
<ul style="list-style-type: none"> да притежава функционалност за таксономична класификация на одитните данни за събития; да притежава функционалност за контекстно съпоставяне (contextual mapping) на данните в процеса на събиране; да притежава функционалност за търсене на регистрирани и запазени събития; 		<p>82/pdfdoc/admin/admin.pdf</p> <ul style="list-style-type: none"> Решението "Sentinel Enterprise" на Micro Focus може да прави таксономична класификация на одитните данни за събития. <p>https://www.netiq.com/documentation/sentinel-82/pdfdoc/install/install.pdf</p> <ul style="list-style-type: none"> Решението „Sentinel Enterprise“ на Micro Focus може да прави контекстно съпоставяне на данните в процеса на събиране на самите данни, так. нар. "contextual mapping". <p>https://www.netiq.com/documentation/sentinel-82/pdfdoc/install/install.pdf</p> <ul style="list-style-type: none"> Решението „Sentinel Enterprise“ на Micro Focus може да осъществява търсене в регистрирани и запазени събития (event search). <p>https://www.netiq.com/documentation/sentinel-82/pdfdoc/install/install.pdf</p>	

000106



Изискване	Изпълнение (Да/Не)	Начин на реализация	
<ul style="list-style-type: none"> да притежава функционалност за ръчно създаване на събития; да притежава функционалност за създаване на политики за съхранение на събития; да притежава функционалност за филтриране на събития; да притежава функционалност за докладване/отчети за събития; 		<ul style="list-style-type: none"> Решението „Sentinel Enterprise” на Micro Focus притежава функционалността за ръчно създаване на събития. https://www.netiq.com/documentation/sentinel-82/pdfdoc/install/install.pdf Решението „Sentinel Enterprise” на Micro Focus може да създава политики за съхранение на събития. https://www.netiq.com/documentation/sentinel-82/pdfdoc/install/install.pdf Решението „Sentinel Enterprise” на Micro Focus притежава функционалността за филтриране на събития (event filtering). https://www.netiq.com/documentation/sentinel-82/pdfdoc/install/install.pdf Решението „Sentinel Enterprise” на 	<p style="text-align: right;">✓</p>

000107



Изискване	Изпълнение (Да/Не)	Начин на реализация	
<p>• да има възможност за доклади/отчети за съответствие с нормативни и законови актове, вътрешни политики и разпоредби , които да бъдат използвани при одити и разследвания;</p> <p>• да има възможност за създаване на доклади/отчети дефинирани от потребителя;</p>		<p>Micro Focus може да прави различни доклади/отчети за събития (Event reporting)</p> <p>https://www.netiq.com/documentation/sentinel-82/pdfdoc/install/install.pdf</p> <ul style="list-style-type: none"> Решението „Sentinel Enterprise” на Micro Focus предоставя възможност за изготвяне на доклади/отчети за съответствие с нормативни и законови актове, вътрешни политики и разпоредби , които да бъдат използвани при одити и разследвания. <p>https://www.netiq.com/documentation/sentinel-82/s821-user/data/bookinfo.html</p> <ul style="list-style-type: none"> Решението „Sentinel Enterprise” на Micro Focus предоставя възможност за изготвяне на доклади/отчети дефинирани от потребителите на системата. <p>https://www.netiq.com/documentation/sentinel-82/s821-admin/data/bookinfo.html</p>	

807:000



Изискване	Изпълнение (Да/Не)	Начин на реализация
<ul style="list-style-type: none"> да има потребителски интерфейс поддържащ разпространени браузери и да <ul style="list-style-type: none"> да предоставя графично анализиране на аномалии; да може да поддържа анализ на данните от източниците на събития; да може да дефинира последователност от действия във формата на работен процес, които да бъдат изпълнявани при възникване на събитие или събития; 		<ul style="list-style-type: none"> Решението „Sentinel Enterprise” има потребителски интерфейс, който поддържа различни, често най - разпространените браузери, както при събирането и съхранението на данни от одитни записи, така и при изготвяне на доклади и извършване на търсения и разследвания и графично анализиране на аномалии. <p>https://www.netiq.com/documentation/sentinel-82/pdfdoc/install/install.pdf</p> <ul style="list-style-type: none"> Решението „Sentinel Enterprise” може да прави анализ на събраните данни от източниците на събития. <p>https://www.netiq.com/documentation/sentinel-82/pdfdoc/admin/admin.pdf</p> <ul style="list-style-type: none"> Решението „Sentinel Enterprise” може да дефинира последователност от действия под формата на работен процес, които могат да бъдат

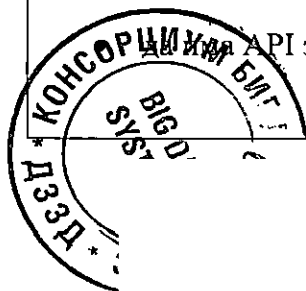
607000



Handwritten marks and signatures on the right side of the page, including a large checkmark, a signature, and several smaller checkmarks.

Изискване	Изпълнение (Да/Не)	Начин на реализация	
<ul style="list-style-type: none"> да притежава механизъм за „ранно предупреждение“; да може да прави корелации и да създава правила за корелации; да има API за достъп до данните му; 		<p>изпълнявани при възникване на събитие или множество събития.</p> <p>https://www.netiq.com/documentation/sentinel-82/pdfdoc/install/install.pdf</p> <ul style="list-style-type: none"> Решението „Sentinel Enterprise“ притежава механизъм за „ранно предупреждение“, чрез функционалността си за „Anomaly detection“. <p>https://www.microfocus.com/en-us/products/netiq-sentinel/features</p> <ul style="list-style-type: none"> Решението „Sentinel Enterprise“ може да прави корелации и да създава правила за корелации. <p>https://www.microfocus.com/en-us/products/netiq-sentinel/features</p> <ul style="list-style-type: none"> Решението „Sentinel Enterprise“ притежава възможност за създаване на API за достъп до данните му. 	

000110



Handwritten signature or mark.

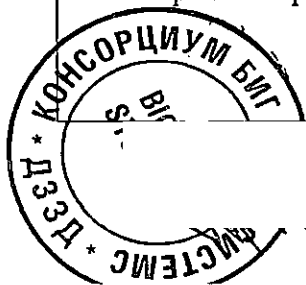
Изискване	Изпълнение (Да/Не)	Начин на реализация	Забелжки
<ul style="list-style-type: none"> да се предложат лицензи за минимум 100 броя събития за секунда (Events per second (EPS)). 		<p>https://www.novell.com/developer/pluginsdk/ref/restapi/7.0/#API_documentation</p> <ul style="list-style-type: none"> Решението „Sentinel Enterprise” се лицензира на събития за секунда (Events per second (EPS)). В ценовото предложение, ще бъдат предложени лицензи за 100 броя събития за секунда. <p>https://www.netiq.com/documentation/sentinel-82/pdfdoc/s821-install/s821-install.pdf</p>	
<p>Системно програмно осигуряване ще бъде заредено в изчислителният сървър за УАЛС и ще съдържа следните компоненти:</p>	ДА	<p>1. Sentinel Enterprise 100 EPS / 50 Devices License с включена 5 годишна поддръжка</p>	✓
<p>T.4.5 В изчислителният сървър за МУД трябва да се зареди системно програмно осигуряване, което да изпълнява следните минимални функции:</p>	ДА	<p>Предлагаме в изчислителният сървър за МУД да се зареди системно програмно осигуряване, което представлява софтуерно решение за прилагане на политики и управление на достъпа до мрежата, чрез модула Network bundle NAC (Network Access Control) на масмон.</p>	

000211



Изискване	Изпълнение (Да/Не)	Начин на реализация	Завис
<ul style="list-style-type: none"> да приема указания за работа от IPS устройството чрез API; да приема указания за работа от Сървъра за УАЛС чрез API; да може да предостави цялостен преглед на всички управляеми устройства (лаптопи, компютри, десктоп машини, сървъри, комуникационни модули, принтери и др.) включени към мрежата на Специализираната система под ключ, в реално време; 		<ul style="list-style-type: none"> Решението "NAC на macmon" може да приема указания за работа от IPS устройството посредством API. https://www.macmon.eu/en/wissen/media/ Решението на macmon за "NAC", чрез модула Network bundle NAC може да приема указания за работа от Сървъра за УАЛС чрез API. https://www.macmon.eu/en/solutions/switch-viewer/ Решението на macmon за "NAC", чрез Network bundle NAC осъществява преглед на достъпа до мрежата в реално време на всички управляеми устройства (таблети, смартфони, лаптопи, компютри, десктоп машини, медицински устройства, принтери, сървъри и комуникационни устройства), независимо дали са 	<p>✓</p> <p>✓</p>

000112



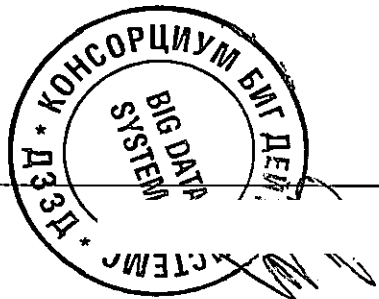
Изискване	Изпълнение (Да/Не)	Начин на реализация	Заб.
<ul style="list-style-type: none"> да може да извършва филтриране като минимум по IP адрес, име, VLAN; да може да управлява достъпа на всички устройства включени в мрежата на Специализираната система под ключ; 		<p>познати или непознати за организацията.</p> <p>https://www.macmon.eu/fileadmin/users_all/content/Produktinfos/ENG_Datasheets/macmon_NAC_Datasheet_module-bundles_6-pages.pdf</p> <ul style="list-style-type: none"> Решението на macmon за "NAC", чрез модула Graphical Topology съдържащ се в Network bundle NAC може да филтрира по IP адрес, име, VLAN. <p>https://www.macmon.eu/fileadmin/users_all/content/Produktinfos/ENG_Datasheets/macmon_N</p> <ul style="list-style-type: none"> Решението на macmon за "NAC", чрез Network bundle NAC може, както да осъществява цялостен преглед на всички крайни точки включени в мрежата, така и да ги управлява. <p>https://www.macmon.eu/fileadmin/users_all/content/Produktinfos/ENG_Datasheets/macmon_NA</p>	<p>1 /</p> <p>✓ ✓</p>

000113



Изискване	Изпълнение (Да/Не)	Начин на реализация
<ul style="list-style-type: none"> да може да осъществява мрежов контрол на достъпа с комуникационни модули както по SNMP/Telnet, така и с 802.1x; да може да прилага на конкретен порт правила на комуникационните модули за разрешаване или отказване на достъп, както и да се получават уведомления за аномалии с помощта на SNMP команди; 		<p><u>C Datasheet module-bundles 6-pages.pdf</u></p> <ul style="list-style-type: none"> Решението на macmon за "NAC", чрез Network bundle NAC може да осъществява връзка с всички комутатори и маршрутизатори в организацията, както по SNMP, така и с 802,1x. <p><u>https://www.macmon.eu/fileadmin/users_all/content/Produktinfos/ENG_Datasheets/macmon_N</u></p> <ul style="list-style-type: none"> Решението на macmon за "NAC", чрез Network bundle NAC, може да прави различни правила с портовете на комутаторите. Например да предоставя уведомления за аномалии с помощта на SNMP команди, свързани с портовете, да разрешава или отказва достъп до ресурси в мрежата на организацията. <p><u>https://www.macmon.eu/fileadmin/users_all/content/Produktinfos/ENG_Datasheets/macmon_N</u></p>

000114



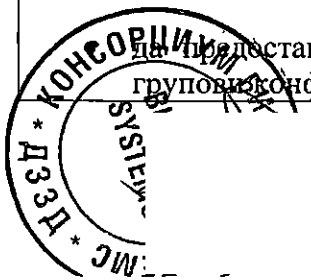
Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависимост
<ul style="list-style-type: none"> да може да наблюдава мрежата на Специализираната система под ключ и да реагира при нестандартно поведение, като: сканиране на портовете, сканиране на източниците на ARP данни, неочаквани и/или несъвместими отговори от източниците на ARP данни, порои от пакети и пр. Реакцията да може да се простира от подаване на тревога до деактивиране на порта на мрежовия порт на устройството; да предоставя възможност за анализ на системните данни като: име на хост, IP адрес, операционна система и пр.; 		<p>ent/Produktinfos/ENG_Datasheets/macmon_N</p> <ul style="list-style-type: none"> Решението на macmon за "NAC", чрез модулът Advanced Security, включен в Network bundle NAC може да наблюдава мрежата на Специализираната система под ключ, в това число сканиране на портове сканиране на източниците на ARP данни, неочаквани и/или несъвместими отговори от източниците на ARP данни, порои от пакети и пр. Реакцията се простира от подаване на тревога до деактивиране на порта на мрежовия порт на устройството; <p>https://www.macmon.eu/fileadmin/users_all/content/Produktinfos/ENG_Datasheets/macmon_N</p> <ul style="list-style-type: none"> Решението на macmon за "NAC", чрез модулът Advanced Security, включен в Network bundle NAC анализира системните данни име на хост, IP адрес, операционна система, отворени или затворени IP портове, като по този 	<p>✓ V Y</p>

000115



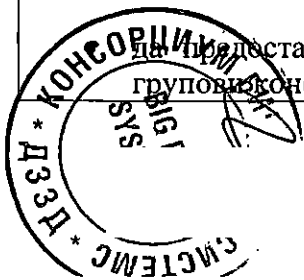
Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависим
<ul style="list-style-type: none"> да може да прехвърля опасните крайни точки в карантинен VLAN докато статуса им не бъде напълно изяснен; да може да се интегрира с LDAP на Microsoft - Active Directory; да предоставя възможност да се правят групови конфигурации и политики; 		<p>начин предоставя по - голяма сигурност за организацията.</p> <p>https://www.macmon.eu/fileadmin/users_all/content/Produktinfos/ENG_Datasheets/macmon_N</p> <ul style="list-style-type: none"> Решението на macmon за "NAC", чрез модулт VLAN Manager, включен в Network bundle NAC може да прави карантинен VLAN и да прехвърля там всички опасни крайни точки до изясняване на техния статус. <p>https://www.macmon.eu/fileadmin/users_all/content/Produktinfos/ENG_Datasheets/macmon_N</p> <ul style="list-style-type: none"> Решението на macmon за "NAC" има възможност за интеграция с LDAP на Microsoft - Active Directory <p>https://www.rdggroup.co.za/images/wg/Datasheet_M200_M370.pdf</p> <ul style="list-style-type: none"> Решението на macmon за "NAC", чрез 	<p>✓</p>


000116



Изискване	Изпълнение (Да/Не)	Начин на реализация	Зави
<ul style="list-style-type: none"> да може да прехвърля опасните крайни точки в карантинен VLAN докато статуса им не бъде напълно изяснен; да може да се интегрира с LDAP на Microsoft - Active Directory; 		<p>начин предоставя по - голяма сигурност за организацията.</p> <p>https://www.macmon.eu/fileadmin/users_all/content/Produktinfos/ENG_Datasheets/macmon_N</p> <ul style="list-style-type: none"> Решението на macmon за "NAC", чрез модулт VLAN Manager, включен в Network bundle NAC може да прави карантинен VLAN и да прехвърля там всички опасни крайни точки до изясняване на техния статус. <p>https://www.macmon.eu/fileadmin/users_all/content/Produktinfos/ENG_Datasheets/macmon_N</p> <ul style="list-style-type: none"> Решението на macmon за "NAC" има възможност за интеграция с LDAP на Microsoft - Active Directory <p>https://www.rdggroup.co.za/images/wg/Datasheet_M200_M370.pdf</p> <ul style="list-style-type: none"> Решението на macmon за "NAC", чрез 	

000116



Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависим
<p>000117</p> <ul style="list-style-type: none"> да притежава графичен потребителски интерфейс за управление; да прехвърля опасни крайни точки в конфигуриран LAN, докато статуса им не бъде изяснен; 		<p>модулът 802.1x , включен в Network bundle NAC много лесно могат да бъдат направени различни групови конфигурации и политики и така да бъде автоматизиран процеса в организацията.</p> <p>https://www.macmon.eu/fileadmin/users_all/content/Produktinfos/ENG_Datasheets/macmon_N</p> <ul style="list-style-type: none"> Решението на macmon за "NAC", чрез Network bundle NAC притежава графичен потребителски интерфейс. Данните извлечени от потребителския интерфейс могат да бъдат в графична форма и могат да бъдат сортирани и прилагани в различни филтри. Модулът може да предостави редица доклади, включваща възможността за експорт в CSV формат. <p>https://www.macmon.eu/en/knowledge/media/</p> <ul style="list-style-type: none"> Решението на macmon за "NAC", чрез модулът VLAN Manager, включен в Network bundle NAC може да прави 	

Изискване	Изпълнение (Да/Не)	Начин на реализация	Зав.
<ul style="list-style-type: none"> предложените лицензи да обхващат минимум 150 MAC адреса 		<p>карантинен VLAN и да прехвърля там всички опасни крайни точки до изясняване на техния статус.</p> <p>https://www.macmon.eu/fileadmin/users_all/content/Produktinfos/ENG_Datasheets/macmon_N</p> <ul style="list-style-type: none"> Лицензирането на решението на macmon за "NAC", чрез Network bundle NAC се лицензира пер MAC адрес. В ценово предложение ще предоставим цена за 150 MAC адреса. <p>https://www.macmon.eu/en/knowledge/faq/</p>	✓
<p>Системно програмно осигуряване ще бъде заредено в изчислителният сървър за МУД и ще съдържа следните компоненти:</p>	ДА	<p>1. Macmon NAC network bundle за 150 MAC адреса с включена 5 годишна поддръжка</p>	
<p>Т 5. Информационна сигурност на Специализираната централизирана система под ключ Nadoor.</p> <p>Доставчикът трябва да изгради следните компоненти на информационна сигурност за Специализираната централизирана</p>	ДА	<p>- Предлагаме да изградим изброените по – долу компоненти на информационна сигурност за Специализираната централизирана</p>	



000118

Изискване	Изпълнение (Да/Не)	Начин на реализация	За
<p>система под ключ (turnkey system) Hadoop, отчитайки потенциалното голямо количество крайни потребители и съобразявайки се с основни изисквания на Европейския регламент за защита на личните данни (GDPR).</p>		<p>система под ключ (turnkey system) Hadoop, като ще вземем предвид голямото количество крайни потребители и ще се съобразим с основните изисквания на Европейския регламент за защита на личните данни (GDPR). В следващите точки са изброени компонентите.</p>	
<p>Т. 5.1. Централизирано автентикиране на потребителите на изградената Специализирана централизирана система под ключ Hadoop система с използване на LDAP сървър (например като Майкрософт АД или подобни), чрез опериране с „Име“ и „Парола“, което да се интегрира с Cloudera Manager. Да се създадат минимум 2 групи потребители с по минимум трима потребителя в група. Да се дефинира управление на поне 1 брой Hadoop услуга;</p> <p>Компонентата се предава чрез Приемо-предавателен протокол с разпечатка от създаденото централизирано автентикиране на потребителите на изградената Специализирана централизирана система под ключ Hadoop.</p>	ДА	<ul style="list-style-type: none"> - На етап “Предаване на системата“ компонента ще бъде предаден с Приемо-предавателен протокол с разпечатка от създаденото централизирано автентикиране на потребителите на изградената Специализирана централизирана система под ключ Hadoop. - На етап оферта предоставяме Документ „Ръководство на администратора за Създаване на централизирано автентикиране на потребителите на Hadoop система“, в което е описано: <p>а)подход за създаване съответното изискване</p>	<p>7</p>

000119



Изискване	Изпълнение (Да/Не)	Начин на реализация	Завис
		<p>с посочено техническо и програмно-конфигурационно решение;</p> <p>б) един или няколко примера за прилагане на изискването.</p> <p>Прилагаме описания документ веднага след Т.11, стр. 117 от настоящето Техническо предложение</p>	
<p>Т. 5.2. Създаване права за достъп за сървърите и сегментиране на данните, като се обхвалят NameNode, DataNode, JobTracker, ApplicationMaster, TaskTracker, NodeManager, за предварително създадени потребители на Hadoop система;</p> <p>Компонентата се предава чрез Приемо-предавателен протокол с разпечатка от създадените права за достъп за сървърите и сегментирането на данните на изградената Специализираната централизирана система под ключ Hadoop.</p>	ДА	<ul style="list-style-type: none"> - На етап "Предаване на системата" компонента ще бъде предаден с Приемо-предавателен протокол с разпечатка от създадените права за достъп за сървърите и сегментирането на данните на изградената Специализираната централизирана система под ключ Hadoop. - На етап оферта предоставяме Документ „Ръководство на администратора за Създаване на права за достъп за сървъри и сегменти от данни в Hadoop система, на вече създадени потребители“, в което е описано: <p>а) подход за създаване съответното изискване с посочено техническо и програмно-</p>	

000120

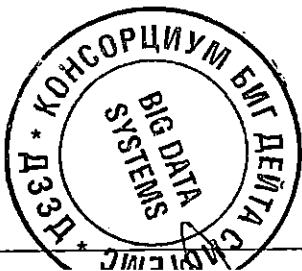


Изискване	Изпълнение (Да/Не)	Начин на реализация	Заб.
		<p>конфигурационно решение;</p> <p>б)един или няколко примера за прилагане на изискването.</p> <p>Прилагаме описания документ веднага след Т.11, стр. 117 от настоящето Техническо предложение</p>	✓
<p>Т. 5.3. Създаване контрол на достъпа до HDFS директории и файлове на изградената Специализирана централизирана система под ключ Hadoop система от страна на потребителски приложения (такива като работа с CLI шел или програми написани с използване на Hadoop API), като за поне 2 от създадени предварително потребителя да се изгради контрол на достъпа до поне 2 директории с по минимум 2 файла.</p> <p>Компонентата се предава чрез Приемо-предавателен протокол с разпечатка от създаден контрол на достъпа до HDFS директории и файлове на данните на изградената Специализирана централизирана система под ключ Hadoop.</p>	ДА	<ul style="list-style-type: none"> - На етап "Предаване на системата" компонента ще бъде предаден с Приемо-предавателен протокол с разпечатка от създаден контрол на достъпа до HDFS директории и файлове на данните на изградената Специализирана централизирана система под ключ Hadoop. - На етап оферта предоставяме Документ „Ръководство на администратора за Създаване контрол на достъпа до HDFS директории и файлове на Hadoop система“, в което се описва: <p>а)подход за създаване съответното изискване с посочено техническо и програмно-конфигурационно решение;</p>	✓

000121



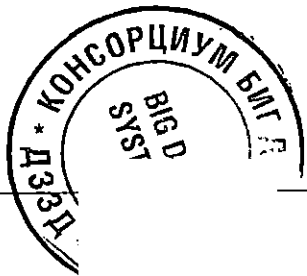
Изискване	Изпълнение (Да/Не)	Начин на реализация	
		<p>б)един или няколко примера за прилагане на изискването.</p> <p>Прилагаме описания документ веднага след Т.11, стр. 117 от настоящето Техническо предложение</p>	
<p>Т. 5.4. Създаване централизиран Лист за управление на достъпа до изградената Специализирана централизирана система под ключ Hadoop система, като за минимум 2 от предварително създадени потребителя да се изгради контрол на достъпа.</p> <p>Компонентата се предава чрез Приемо-предавателен протокол с разпечатка от създаден централизиран Лист за управление на достъпа на изградената Специализирана централизирана система под ключ Hadoop.</p>	ДА	<ul style="list-style-type: none"> - На етап "Предаване на системата" компонента ще бъде предаден с Приемо-предавателен протокол с разпечатка от създаден централизиран Лист за управление на достъпа на изградената Специализирана централизирана система под ключ Hadoop. - На етап оферта предоставяме Документ „Ръководство на администратора за Създаване централизиран Лист за управление на достъпа до Hadoop система“, в което се описва: <ul style="list-style-type: none"> а)подход за създаване съответното изискване с посочено техническо и програмно-конфигурационно решение; б)пример за прилагане на изискването 	<p>Л</p> <p>С</p> <p>С</p>



000122

Изискване	Изпълнение (Да/Не)	Начин на реализация	
<p>Т. 5.5. Създаване на механизъм за пълен одит на изградената Специализирана централизирана система под ключ Nadoop система с параметри: а) произход на данните – от къде са дошли и за какво са били използвани; б) осъществен достъп до данни; в) идентификация на потребител осъществил достъп до Nadoop система с IP адрес на потребителя.</p> <p>Компонентата се предава чрез Приемо-предавателен протокол с разпечатка от създаден механизъм за пълен одит на изградената Специализирана централизирана система под ключ Nadoop.</p>	ДА	<p>Прилагаме описания документ веднага след Т.11, стр. 117 от настоящето Техническо предложение</p> <ul style="list-style-type: none"> - На етап "Предаване на системата" компонента ще бъде предаден с Приемо-предавателен протокол с разпечатка от създаден механизъм за пълен одит на изградената Специализирана централизирана система под ключ Nadoop. - На етап оферта предоставяме Документ „Ръководство на администратора за Създаване на механизъм за пълен одит на Nadoop система“, в което се описва: <ul style="list-style-type: none"> а) подход за създаване съответната одитна система с 3 параметъра – произход на данни, достъп до данни и идентификация на потребител, с посочено техническо и програмно-конфигурационно решение; б) разпечатка от работата на създадена одитна система с посочените 3-те броя параметри. <p>Прилагаме описания документ веднага след Т.11, стр. 117 от настоящето</p>	

000123



Изискване	Изпълнение (Да/Не)	Начин на реализация
<p>Т. 5.6. Осигуряване на защита и криптиране на данните в изградената Специализирана централизирана система под ключ Hadoop система – криптиране на данни намиращи се върху диск “data-at-rest” и криптиране на данни в движение в Hadoop система – “data-in-transit”</p> <p>Компонентата се предава чрез Приемо-предавателен протокол с разпечатка от създадена защита и криптиране на данните на изградената Специализирана централизирана система под ключ Hadoop.</p>	ДА	<p>Техническо предложение</p> <ul style="list-style-type: none"> - На етап “Предаване на системата” компонента ще бъде предаден с Приемо-предавателен протокол с разпечатка от създадена защита и криптиране на данните на изградената Специализирана централизирана система под ключ Hadoop. - На етап оферта предоставяме Документ „Ръководство на администратора за Осигуряване на защита и криптиране на данните в Hadoop система за данни върху диск и данни в движение“, в което се описва: <ul style="list-style-type: none"> а) подход за създаване съответното изискване с посочено техническо и програмно-конфигурационно решение; б) по един пример за двата вида криптиране, като се използва един от предварително създадени потребители. <p>Прилагаме описания документ веднага след Т.11, стр. 117 от настоящето Техническо предложение</p>
управление на ключовете за	ДА	- На етап “Предаване на системата”

060124



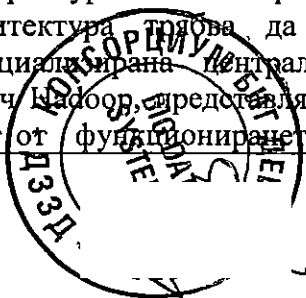
000125

Изискване	Изпълнение (Да/Не)	Начин на реализация	
<p>криптиране в изградената Специализирана централизирана система под ключ Nadoop.</p> <p>Компонентата се предава чрез Приемо-предавателен протокол с разпечатка от създадено управление на ключовете за криптиране на изградената Специализирана централизирана система под ключ Nadoop .</p>		<p>компонента ще бъде предаден с Приемо-предавателен протокол с разпечатка от създадено управление на ключовете за криптиране на изградената Специализирана централизирана система под ключ Nadoop .</p> <p>- На етап оферта предоставяме Документ „Ръководство на администратора за Изграждане управление на ключовете за криптиране в Nadoop система“, в което се описва:</p> <p>а)подход за създаване съответното изискване с посочено техническо и програмно-конфигурационно решение;</p> <p>б)пример как са създадени 2 броя потребители се създават ключове и се представи как може те да се управляват.</p> <p>Прилагаме описания документ веднага след Т.11, стр. 117 от настоящето Техническо предложение</p>	<p>V</p> <p>2</p>
<p>Т.11. Адаптивната Информационна сигурност е подход за киберсигурността, който анализира</p>	ДА	Специализирана централизирана система под	За да може да



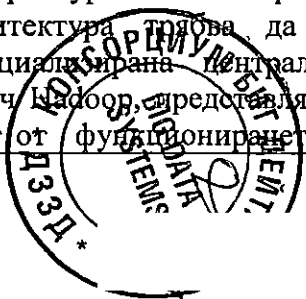
Изискване	Изпълнение (Да/Не)	Начин на реализация	За и
<p>поведението и събитията, за да ги предпази и да се адаптира към заплахите, преди те да се случат. С Адаптивна Архитектура за Информационна Сигурност една организация може непрекъснато да оценява риска и автоматично да осигурява пропорционално правоприлагане. Поради постоянното развитие на заплахите за сигурността, една Специализирана под ключ Nadoor система не може достатъчно просто да използва механизми за блокиране или да прилага процедури след края на събитие за уязвимост. Тя трябва да използва по-усъвършенствани платформи за сигурност, които да могат да се адаптират към най-новите заплахи и да използват динамични механизми за защита и реакция.</p>		<p>ключ Nadoor ще бъде реализирана на база на Адаптивна Архитектура за Информационна сигурност, която обхваща три компонента: Откриване (Predict), Защита (Prevent) и Отговор (Respond). Тези компоненти ще отговарят за идентифицирането, анализирането, блокирането и адаптирането към заплахите като ще се използват динамични механизми за защита и реакции. Тези механизми ще проучват поведението на заплахите и ще придобиват умения на базата, на които ще се адаптират и ще отговарят адекватно на бъдещи заплахи.</p>	<p>осъществи внедряването на системата е необходимо да бъде осигурен физически достъп до средата където ще бъде ситуирана и където ще функционира системата и до съответните крайни точки. Внедряването е нужно да бъде съобразено с изисквания описани в ръководствата за инсталация и конфигурация на съответните продукти.</p>
<p>Адаптивната Архитектура за Информационна Сигурност на Специализирана централизирана система под ключ Nadoor трябва да бъде архитектурно решение, при което наблюдението на заплахите да става непрекъснато и да се подобрява, тъй като рисковете в киберсигурността се променят с времето. Тази архитектура трябва да бъде интегрирана в Специализирана централизирана система под ключ Nadoor, представляваща нейна неразделна част от функционирането. Тази архитектура на</p>	ДА	<p>Трите компоненти от архитектурата Откриване (Predict), Защита (Prevent) и Отговор (Respond) включват респективно Модул за управление и анализ на събития (Модул за УАЛС), IPS Модул, и Модул за Мрежово управление и достъп (Модул за МУД). Продуктите, които ще бъдат използвани, които покриват изискванията на тези модули са респективно MicroFocus Sentinel SIEM, WatchGuard UTM и Macmon. Тези модули ще си взаимодействат със</p>	<p>✓</p>


000126



Изискване	Изпълнение (Да/Не)	Начин на реализация	
<p>поведението и събитията, за да ги предпази и да се адаптира към заплахите, преди те да се случат. С Адаптивна Архитектура за Информационна Сигурност една организация може непрекъснато да оценява риска и автоматично да осигурява пропорционално правоприлагане. Поради постоянното развитие на заплахите за сигурността, една Специализирана под ключ Nadoor система не може достатъчно просто да използва механизми за блокиране или да прилага процедури след края на събитие за уязвимост. Тя трябва да използва по-усъвършенствани платформи за сигурност, които да могат да се адаптират към най-новите заплахи и да използват динамични механизми за защита и реакция.</p>		<p>ключ Nadoor ще бъде реализирана на база на Адаптивна Архитектура за Информационна сигурност, която обхваща три компонента: Откриване (Predict), Защита (Prevent) и Отговор (Respond). Тези компоненти ще отговарят за идентифицирането, анализирането, блокирането и адаптирането към заплахите като ще се използват динамични механизми за защита и реакции. Тези механизми ще проучват поведението на заплахите и ще придобиват умения на базата, на които ще се адаптират и ще отговарят адекватно на бъдещи заплахи.</p>	<p>осъществи внедряването на системата е необходимо да бъде осигурен физически достъп до средата където ще бъде ситуирана и където ще функционира системата и до съответните крайни точки. Внедряването е нужно да бъде съобразено с изисквания описани в ръководствата за инсталация и конфигурация на съответните продукти.</p>
<p>Адаптивната Архитектура за Информационна Сигурност на Специализирана централизирана система под ключ Nadoor трябва да бъде архитектурно решение, при което наблюдението на заплахите да става непрекъснато и да се подобрява, тъй като рисковете в киберсигурността се променят с времето. Тази архитектура трябва да бъде интегрирана в Специализирана централизирана система под ключ Nadoor, представляваща нейна неразделна част от функционирането. Тази архитектура на</p>	ДА	<p>Трите компоненти от архитектурата Откриване (Predict), Защита (Prevent) и Отговор (Respond) включват респективно Модул за управление и анализ на събития (Модул за УАЛС), IPS Модул, и Модул за Мрежово управление и достъп (Модул за МУД). Продуктите, които ще бъдат използвани, които покриват изискванията на тези модули са респективно MicroFocus Sentinel SIEM, WatchGuard UTM и Macmon. Тези модули ще си взаимодействат със</p>	<p>✓</p>

000126



Изискване	Изпълнение (Да/Не)	Начин на реализация	Сти
б) „Защита (Prevent)” – чрез внедряване на IPS устройството; в) “Отговор (Respond)” – чрез внедряване на Модул за Мрежово управление и достъп – Модул за МУД.			
<p>Адаптивната Архитектура за Информационна Сигурност на Специализираната система под ключ Hadoop трябва да извършва следните 3 вида интеграции:</p> <ul style="list-style-type: none"> - Между Модула за УАЛС и LOG файловете на модулите МодулDataNode, МодулNameNode, МодулManagementNode и МодулEdgeNode, както и тези на операционните системи на тези модули; - Между Модула за УАЛС и LOG файла на Cloudera Manager и на неговата операционна система; - Между IPS устройството и Модула за МУД; - Между Модул за УАЛС и Модул за 	ДА	<p>Във фаза изграждане ще бъдат извършени 4 вида интеграции. На този етап ще предоставим Документ „Примерно функциониране на Адаптивна Архитектура за Информационна Сигурност“, в който са представени разпечатки от примерно функциониране и конфигуриране в компютърна среда на доставчика, на следните компоненти: Модула за УАЛС, IPS устройството и Модула за МУД. В този документ да се представят разпечатки от следните 4 броя примерно интегриране:</p> <p>Между Модул за УАЛС и LOG файловете на един модул МодулDataNode, и на един модул МодулNameNode, както и на техните операционни системи;</p> <p>Резултат от корелация между съдържание на LOG файл на един модул МодулDataNode и</p>	

000128



Изискване	Изпълнение (Да/Не)	Начин на реализация	ТИ
МУД.		<p>на LOG файл на един модул МодулNameNode;</p> <p>Между IPS устройството и Модул за прилагане на политики– Модула за МУД;</p> <p>Между Модул за УАЛС и Модул за МУД.</p> <p>Прилагаме описания документ веднага след Т.11, стр. 117 от настоящето Техническо предложение</p>	Т

000129



„Консорциум Биг Дейта Системс“ ДЗЗД

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

Всички документи по **Т.5** (включително **Т.5.1, Т.5.2, Т.5.3, Т.5.4, Т.5.5, Т.5.6, Т.5.7, Т.5.8**) свързани с Информационна сигурност на Специализираната централизирана система под ключ Nadoor, ще бъдат представени след Т.11 от настоящото техническо предложение.



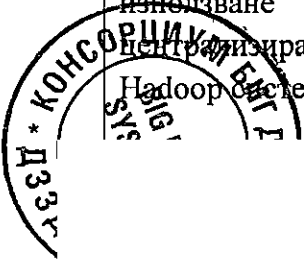
Изискване	Изпълнение (Да/Не)	Начин на реализация	Зав.
<p>T.6 Интегриране на Специализираната централизирана система под ключ с Майкрософт SQL Server минимум версия 2017</p> <p>Доставчикът трябва да интегрира Майкрософт SQL Server минимум версия 2017, инсталиран на Изчислителният сървър за LDAP, със Специализирана централизирана система под ключ Hadoop.</p>	ДА	<p>Приемаме да интегрираме Майкрософт SQL Server минимум версия 2017, инсталиран на Изчислителният сървър за LDAP, със Специализирана централизирана система под ключ Hadoop.</p>	<p style="text-align: right;">/</p>
<p>T.6.1 Доставчикът трябва да предостави софтуер в първичен код (source code), в изпълняващ код (binary code) и конфигурационни файлове за интегриране между данните на двете системи: MS SQL Server минимална версия 2017 (във вид на релационни таблици) и Специализирана централизирана система под ключ Cloudera/Hadoop (във вид на HDFS файлове), прилагайки PolyBase технологията на MS SQL Server, и който софтуер да има следната функционалност:</p> <p>а) да експортира и импортира данни от едната система (MS SQL Server) в другата (Специализираната централизирана система под ключ) и обратно;</p> <p>б) да четене, модифицира и изтрива данни от</p>	ДА	<ul style="list-style-type: none"> - В документ „Интегриране на Специализираната централизирана система под ключ с Майкрософт SQL Server минимум версия 2017“ предоставяме първичен код (source code), на последна страница в описания документ, а конфигурационните файлове + изпълняващ код (binary code) ще бъдат качени на външен хард диск. - експортиране и импортиране на данни от едната система (MS SQL Server) в другата (Специализираната централизирана система под ключ) и обратно. - четене, модифициране и изтриване на данни 	<p style="text-align: right;">/</p>

000131



Изискване	Изпълнение (Да/Не)	Начин на реализация	
<p>едната система (MS SQL Server) в другата (Специализираната централизирана система под ключ) и обратно;</p> <p>в) Данните в системата MS SQL Server да бъдат в таблици създадени за работа в паметта (memory-optimized tables in the main memory);</p> <p>г) Да се конфигурират Аналитичните услуги (Analytical Services) на MS SQL Server 2017 да работят с интегрирани данни от Cloudera/Hadoop (HDFS файлове).</p>		<p>от едната система (MS SQL Server) в другата (Специализираната централизирана система под ключ) и обратно;</p> <p>- Информация за данните в системата MS SQL Server, създадени за работа в паметта (memory-optimized tables in the main memory)</p> <p>- конфигуриране на Аналитичните услуги (Analytical Services) на MS SQL Server 2017 да работят с интегрирани данни от Cloudera/Hadoop (HDFS файлове).</p>	<p>11/11</p> <p>6</p>
<p>Т.6.2 Доставчикът трябва да предостави софтуер в първичен код (source code), в изпълняващ код (binary code) и конфигурационни файлове за прилагане на технологията за машинно обучение (Machine Learning – ML) на MS SQL Server 2017 или следваща версия, с използване на данни от Специализирана централизирана система под ключ Cloudera / Hadoop система и от MS SQL Server.</p>	ДА	<p>В магнитен носител предоставяме софтуер в първичен код (source code), в изпълняващ код (binary code) и конфигурационни файлове за прилагане на технологията за машинно обучение (Machine Learning – ML) на MS SQL Server 2017 или следваща версия, с използване на данни от Специализирана централизирана система под ключ Cloudera / Hadoop система и от MS SQL Server</p>	<p>11/11</p>

000132



ДОКУМЕНТ:

„Интегриране на Специализираната централизирана система под ключ с Майкрософт SQL Server минимум версия 2017“

Съгласно страница 43, Т.6 от документацията

Съгласно Т.6, стр.43 от документацията - прилагаме изискванията по тази точка на външен хард диск, както следва: **софтуер в първичен код (source code), в изпълняващ код (binary code) и конфигурационни файлове за интегриране между данните на двете системи: MS SQL Server минимална версия 2017 (във вид на релационни таблици) и Специализирана централизирана система под ключ Cloudera/Hadoop (във вид на HDFS файлове), прилагайки PolyBase технологията на MS SQL Server, и който софтуер да има следната функционалност:**

Т.6.1 а) експортиране и импортиране на данни от едната система (MS SQL Server) в другата (Специализираната централизирана система под ключ) и обратно в другата, както следва:

Този код използва JDO за връзка към MSSQL сървър 2017 за създаване на таблица и индекс към нея, вмъкване на ред в таблица, получаване на ред, получаване на стойност на колона, изпълнение на заявка и извършване на някои допълнителни операции.

```
package com.apache.hadoop.hbase.client.jdo.examples;  
  
import java.io.File;  
import java.io.FileInputStream;  
import java.io.InputStream;  
import java.util.Hashtable;  
import org.apache.hadoop.fs.Path;  
import org.apache.hadoop.hbase.client.tableindexed.IndexedTable;  
import com.apache.hadoop.hbase.client.jdo.AbstractHBaseDBO;  
import com.apache.hadoop.hbase.client.jdo.HBaseBigFile;
```



„Консорциум Биг Дейта Системс“ ДЗ:

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

```
import com.apache.hadoop.hbase.client.jdo.HBaseDBOImpl;
import com.apache.hadoop.hbase.client.jdo.query.DeleteQuery;
import com.apache.hadoop.hbase.client.jdo.query.HBaseOrder;
import com.apache.hadoop.hbase.client.jdo.query.HBaseParam;
import com.apache.hadoop.hbase.client.jdo.query.InsertQuery;
import com.apache.hadoop.hbase.client.jdo.query.QSearch;
import com.apache.hadoop.hbase.client.jdo.query.SelectQuery;
import com.apache.hadoop.hbase.client.jdo.query.UpdateQuery;

/**
 * Hbase JDO.
 *
 * dependency library.
 * - commons-beanutils.jar
 * - commons-pool-1.5.5.jar
 * - hbase0.90.0-transactionl.jar
 *
 * you can expand Delete,Select,Update,Insert Query classes.
 *
 */

public class HBaseExample {
    public static void main(String[] args) throws Exception {
        AbstractHBaseDBO dbo = new HBaseDBOImpl();

        /**drop if table is already exist.*
        if(dbo.isTableExist("user")){
            dbo.deleteTable("user");
```



000134

„Консорциум Биг Дейта Системс“ |

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

}

/*create table*

dbo.createTableIfNotExist("user",HBaseOrder.DESC,"account");

//dbo.createTableIfNotExist("user",HBaseOrder.ASC,"account");

//create index.

String[] cols={"id","name"};

dbo.addIndexExistingTable("user","account",cols);

//insert

InsertQuery insert = dbo.createInsertQuery("user");

UserBean bean = new UserBean();

bean.setFamily("account");

bean.setAge(20);

bean.setEmail("example@gmail.com");

bean.setId("ncanis");

bean.setName("ncanis");

bean.setPassword("1111");

insert.insert(bean);

//select 1 row

SelectQuery select = dbo.createSelectQuery("user");

UserBean resultBean = (UserBean)select.select(bean.getRow(),UserBean.class);

// select column value.

String value = (String)select.selectColumn(bean.getRow(),"account","id",String.class);

// search with option (QSearch has EQUAL, NOT_EQUAL, LIKE)

// select id,password,name,email from account where id='ncanis' limit startRow

HBaseParam param = new HBaseParam();



000_35

„Консорциум Биг Дейта Системс“ Д

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

```
param.setPage(bean.getRow(),20);
param.addColumn("id","password","name","email");
param.addSearchOption("id","ncanis",QSearch.EQUAL);
select.search("account", param, UserBean.class);
// search column value is existing.
boolean isExist = select.existColumnValue("account","id","ncanis".getBytes());
// update password.
UpdateQuery update = dbo.createUpdateQuery("user");
Hashtable<String, byte[]> colsTable = new Hashtable<String, byte[]>();
colsTable.put("password","2222".getBytes());
update.update(bean.getRow(),"account",colsTable);
//delete
DeleteQuery delete = dbo.createDeleteQuery("user");
delete.deleteRow(resultBean.getRow());
////////////////////////////////////
// etc
// HTable pool with apache commons pool
// borrow and release. HBasePoolManager(maxActive, minIdle etc..)
IndexedTable table = dbo.getPool().borrow("user");
dbo.getPool().release(table);
// upload bigFile by hadoop directly.
HBaseBigFile bigFile = new HBaseBigFile();
File file = new File("doc/movie.avi");
FileInputStream fis = new FileInputStream(file);
Path rootPath = new Path("/files/");
String filename = "movie.avi";
```



000136

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

```
bigFile.uploadFile(rootPath,filename,fis,true);
```

```
// receive file stream from hadoop.
```

```
Path p = new Path(rootPath,filename);
```

```
InputStream is = bigFile.path2Stream(p,4096);
```

```
}
```

```
}
```

Т.6.1 б) да чете, модифицира и изтрива данни от едната система (MS SQL Server) в другата (Специализираната централизирана система под ключ) и обратно, както следва:

Създаване, модифициране и изтриване на Таблица, ползвайки Java :

```
package com.example.hbase.admin;
```

```
import java.io.IOException;
```

```
import org.apache.hadoop.conf.Configuration;
```

```
import org.apache.hadoop.fs.Path;
```

```
import org.apache.hadoop.hbase.HBaseConfiguration;
```

```
import org.apache.hadoop.hbase.HColumnDescriptor;
```

```
import org.apache.hadoop.hbase.HConstants;
```

```
import org.apache.hadoop.hbase.HTableDescriptor;
```

```
import org.apache.hadoop.hbase.TableName;
```

```
import org.apache.hadoop.hbase.client.Admin;
```

```
import org.apache.hadoop.hbase.client.Connection;
```

```
import org.apache.hadoop.hbase.client.ConnectionFactory;
```

```
import org.apache.hadoop.hbase.io.compress.Compression.Algorithm;
```

```
public class Example {
```



000137

„Консорциум Биг Дейта Системс“ Д:

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

```
private static final String TABLE_NAME = "MY_TABLE_NAME_TOO";

private static final String CF_DEFAULT = "DEFAULT_COLUMN_FAMILY";

public static void createOrOverwrite(Admin admin, HTableDescriptor table) throws
IOException {

    if (admin.tableExists(table.getTableName())) {

        admin.disableTable(table.getTableName());

        admin.deleteTable(table.getTableName());

    }

    admin.createTable(table);

}

public static void createSchemaTables(Configuration config) throws IOException {

    try (Connection connection = ConnectionFactory.createConnection(config);

        Admin admin = connection.getAdmin()) {

        HTableDescriptor table = new HTableDescriptor(TableName.valueOf(TABLE_NAME));

        table.addFamily(new
HColumnDescriptor(CF_DEFAULT).setCompressionType(Algorithm.NONE));

        System.out.print("Creating table. ");

        createOrOverwrite(admin, table);

        System.out.println(" Done.");

    }

}

public static void modifySchema (Configuration config) throws IOException {

    try (Connection connection = ConnectionFactory.createConnection(config);

        Admin admin = connection.getAdmin()) {

        TableName tableName = TableName.valueOf(TABLE_NAME);

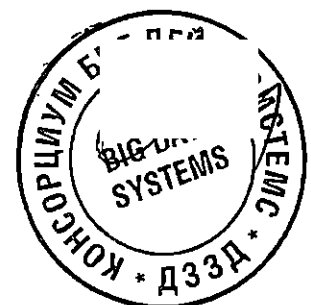
        if (!admin.tableExists(tableName)) {

            System.out.println("Table does not exist.");

        }

    }

}
```



„Консорциум Биг Дейта Системс“

Хай Тех Бизнес Ц,
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

```
System.exit(-1);
}
HTableDescriptor table = admin.getTableDescriptor(tableName);
// Update existing table
HColumnDescriptor newColumn = new HColumnDescriptor("NEWCF");
newColumn.setCompactionCompressionType(Algorithm.GZ);
newColumn.setMaxVersions(HConstants.ALL_VERSIONS);
admin.addColumn(tableName, newColumn);
// Update existing column family
HColumnDescriptor existingColumn = new HColumnDescriptor(CF_DEFAULT);
existingColumn.setCompactionCompressionType(Algorithm.GZ);
existingColumn.setMaxVersions(HConstants.ALL_VERSIONS);
table.modifyFamily(existingColumn);
admin.modifyTable(tableName, table);
// Disable an existing table
admin.disableTable(tableName);
// Delete an existing column family
admin.deleteColumn(tableName, CF_DEFAULT.getBytes("UTF-8"));
// Delete a table (Need to be disabled first)
admin.deleteTable(tableName);
}
}
public static void main(String... args) throws IOException {
    Configuration config = HBaseConfiguration.create();
    //Add any necessary configuration files (hbase-site.xml, core-site.xml)
```



„Консорциум Биг Дейта Системс“ ДЗС

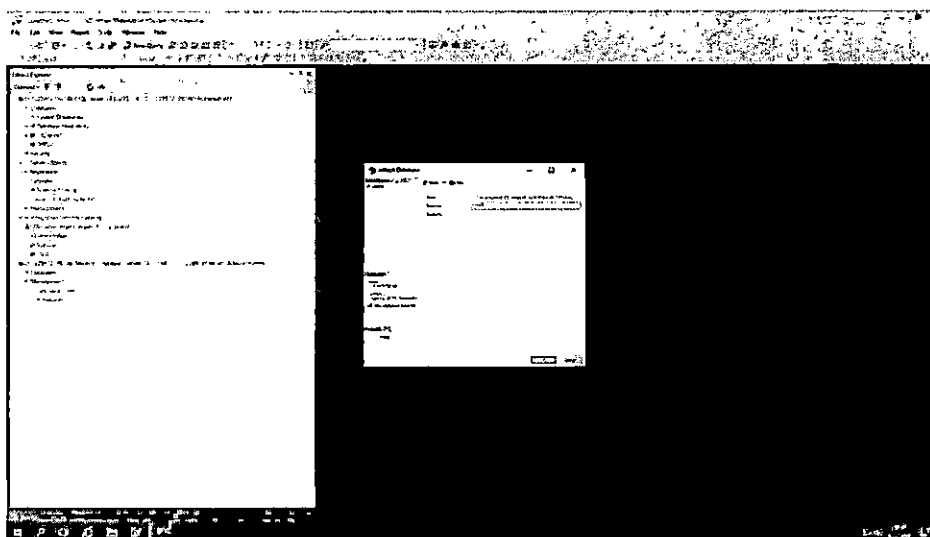
Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

```
config.addResource(new Path(System.getenv("HBASE_CONF_DIR"), "hbase-site.xml"));  
config.addResource(new Path(System.getenv("HADOOP_CONF_DIR"), "core-site.xml"));  
createSchemaTables(config);  
modifySchema(config);  
}  
}
```

Т.6.1 в) Данните в системата MS SQL Server да бъдат в таблици създадени за работа в паметта (memory-optimized tables in the main memory), както следва:

```
ALTER DATABASE temeo ADD FILEGROUP temeo_mod CONTAINS MEMORY_OPTIMIZED_DATA;  
ALTER DATABASE temeo ADD FILE (name='temeo1', filename='c:\DATA\temeo1') TO FILEGROUP temeo_mod;
```

Т.6.1 г) Конфигуриране на Аналитични услуги (Analytical Services) на MS SQL Server 2017 да работят с интегрирани данни от Cloudera/Hadoop (HDFS файлове)



„Консорциум Биг Дейта Системс“ ДЗЗД

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

Софтуер в първичен код (source code):

```
import os, fnmatch
import numpy as np
import matplotlib.patches as mpatches
import matplotlib.pyplot as plt
from operator import itemgetter
#Computing the average distance between lines
def average_distance(letters):
    average_distance = 0
    distances_n = 0
    for i in range(len(letters) - 1):
        letter = letters[i]
        next_letter = letters[i + 1]
        distance = abs(next_letter["y"] - letter["y"])
        if distance > 0:
            average_distance += distance
            distances_n += 1
    return average_distance / distances_n
#Average distance between pairs of letters on the same line
def average_word_distance(lines):
    average_distance = 0
    distances_n = 0
    for line in lines:
        for letter in line:
            distances_n = distances_n + 1
            average_distance = letter["distance_previous"] + average_distance
```



„Консорциум Биг Дейта Системс“ Д

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

```
return average_distance / distances_n

#Getting all PNG file locations in a folder
def get_png(folder):
    file_names = []
    files = os.listdir(folder)
    extension = ".png"
    for file_name in files:
        if fnmatch.fnmatch(file_name, extension):
            file_names.insert(len(file_names), file_name)
    return file_names

#Merge 2 Rectangle objects
def merge_rect(rect1, rect2):

    rect1_xLeft = rect1.get_x()
    rect1_xRight = rect1_xLeft + rect1.get_width()
    rect1_yBottom = rect1.get_y()
    rect1_yTop = rect1_yBottom + rect1.get_height()

    rect2_xLeft = rect2.get_x()
    rect2_xRight = rect2_xLeft + rect2.get_width()
    rect2_yBottom = rect2.get_y()
    rect2_yTop = rect2_yBottom + rect2.get_height()

    xLeft = min(rect1_xLeft, rect2_xLeft)
    xRight = max(rect1_xRight, rect2_xRight)
    yTop = max(rect1_yTop, rect2_yTop)
```



„Консорциум Биг Дейта Системс“ Д

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

```
yBottom = min(rect1_yBottom, rect2_yBottom)
```

```
width = xRight - xLeft
```

```
height = yTop - yBottom
```

```
final_rect = mpatches.Rectangle((xLeft, yBottom), width, height,  
                                fill=False, edgecolor='white', linewidth=2)
```

```
return final_rect
```

```
#Fixing letters
```

```
def fix_letters(current, patches, xFactor, yFactor, size):
```

```
    current_xLeft = current.get_x()
```

```
    current_xRight = current_xLeft + current.get_width()
```

```
    current_yBottom = current.get_y()
```

```
    current_yTop = current_yBottom + current.get_height()
```

```
    current_width = current_xRight - current_xLeft
```

```
    current_height = current_yTop - current_yBottom
```

```
    current_xMid = int((current_xRight - current_xLeft) / 2) + current_xLeft
```

```
    current_yMid = int((current_yTop - current_yBottom) / 2) + current_yBottom
```

```
for patch in patches:
```

```
    rect_xLeft = patch.get_x()
```

```
    rect_xRight = rect_xLeft + patch.get_width()
```



000143

„Консорциум Биг Дейта Системс“

Хай Тех Бизнес І,
ж.к. Дружба І, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

```
rect_yBottom = patch.get_y()
rect_yTop = rect_yBottom + patch.get_height()
rect_width = rect_xRight - rect_xLeft
rect_height = rect_yTop - rect_yBottom
rect_xMid = int((rect_xRight - rect_xLeft) / 2) + rect_xLeft
rect_yMid = int((rect_yTop - rect_yBottom) / 2) + rect_yBottom

# Check if they are part of the same letter
# Check for Overlaps
# Next to each other
if current_yBottom == rect_yBottom and current_yTop == rect_yTop:
    if (current_xRight > rect_xLeft and current_xLeft < rect_xRight) or (current_xLeft < rect_xRight and
current_xRight > rect_xRight):
        current = merge_rect(current, patch)
        patches.remove(patch)

    if rect_xMid in range(current_xLeft, current_xRight) or current_xMid in range(rect_xLeft,
rect_xRight):
        if rect_yMid in range(current_yBottom, current_yTop) or current_yMid in range(rect_yBottom,
rect_yTop):
            current = merge_rect(current, patch)
            patches.remove(patch)

    if rect_xMid in range(current_xMid - xFactor, current_xMid + xFactor) or current_xMid in
range(rect_xMid - xFactor, rect_xMid + xFactor):
        # Case 1 - on Top (dot)
        if rect_yBottom in range(current_yTop, current_yTop + yFactor) or rect_yTop in
range(current_yBottom - yFactor, current_yBottom):
            # Checking if one of the figures is small (dot)
            if (current_width < size and current_height < size) or (rect_width < size and rect_height < size):
```



„Консорциум Биг Дейта Системс

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

```
# Checking if one of the figures has a square-like shape
```

```
if current_width in range(current_height-5, current_height+5) or rect_width in  
range(rect_height-5, rect_height+5):
```

```
    current = merge_rect(current, patch)
```

```
    patch.edgcolor='white'
```

```
    patches.remove(patch)
```

```
# Check for dots/dashes above/below the rectangle
```

```
if rect_xMid in range(current_xMid - xFactor, current_xMid + xFactor) or current_xMid in  
range(rect_xMid - xFactor, rect_xMid + xFactor):
```

```
    # Case 2 - on Bottom (dash)
```

```
if rect_yTop in range(current_yBottom - yFactor, current_yBottom) or rect_yBottom in  
range(current_yTop, current_yTop + yFactor):
```

```
    # Checking if one of the figures has a rectangle-like shape
```

```
if current_width > current_height*3 or rect_width > rect_height*3:
```

```
    current = merge_rect(current, patch)
```

```
    patches.remove(patch)
```

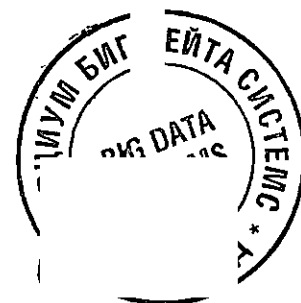
```
patches.append(current)
```

```
return patches
```

```
#Sorting letters in lines
```

```
def sort_letters(letters):
```

```
    sorted_lines = []
```



„Консорциум Биг Дейта Системс“

Хай Тех Бизнес
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

#Sorting all letters by y

```
letters = sorted(letters, key=itemgetter("y"))
```

#Putting letters in lines

```
lines = []
```

```
for i in range(len(letters)):
```

```
    letter = letters[i]
```

```
    if i == 0:
```

```
        k = 0
```

```
        lines.append([])
```

```
        lines[k].append(letter)
```

```
    else:
```

```
        last_letter = letters[i - 1]
```

```
        distance = abs(last_letter["y"] - letter["y"])
```

```
        if distance > average_distance(letters):
```

```
            k += 1
```

```
            lines.append([])
```

```
            lines[k].append(letter)
```

#Sorting all letters in each line by x

```
for line in lines:
```

```
    line = sorted(line, key=itemgetter("x"))
```

```
    sorted_lines.insert(len(sorted_lines), line)
```

#Adding distances

```
for i in range(len(sorted_lines)):
```

„Консорциум Биг Дейта Системс“ ДЗ

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

```
for k in range(len(sorted_lines[i])):
```

```
    if k > 0:
```

```
        current_letter = sorted_lines[i][k]
```

```
        previous_letter = sorted_lines[i][k - 1]
```

```
        current_letter["distance_previous"] = current_letter["x"] - previous_letter["x"]
```

```
        sorted_lines[i][k] = current_letter
```

```
return sorted_lines
```

```
def get_label(index, labels):
```

```
    return labels[index]
```



000147

Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависимо
<p>T.7 Система за тестване пълна функционалност на Специализирана централизирана система под ключ.</p> <p>T.7.1 Провеждане тестване работоспособността на отделните изчислителни модули на DataNode, т.е. на всичките МодулDataNode, от администратор на Специализираната система, чрез прилагане на „Host Heartbeats“;</p>	ДА	<p>Предоставяме следните документи , свързани със „Система за тестване пълна функционалност на Специализирана централизирана система под ключ“, както следва:</p> <ul style="list-style-type: none"> • Документ „Ръководство на администратора за Провеждане тестване на работоспособността на отделните МодулDataNode“, чрез прилагане на „Host Heartbeats“ 	
<p>T.7.2 Провеждане тестване изпълнението на MapReduce задание в Hadoop система, чрез прилагане на “Pi Estimator“, на WorldCount приложение или на друго.</p>	ДА	<ul style="list-style-type: none"> • Документ „Ръководство на администратора за провеждане тестване на изпълнението на MapReduce задание в Hadoop система, чрез прилагане на „Pi Estimator“, на WorldCount приложение или на друго“. 	
<p>T.7.3 Провеждане тестване на комплексната работа на хардуерните и софтуерни компоненти на Специализираната централизирана система под ключ Hadoop, чрез прилагане на „Hue приложение“ или чрез прилагане на „clean test“.</p>	ДА	<ul style="list-style-type: none"> • Документ „Ръководство на администратора за провеждане тестване на комплексната работа на Специализираната централизирана система под ключ Hadoop чрез прилагане на „Hue приложение“ или чрез прилагане на „Clean test“ 	

Ръководство на администратора за
Провеждане тестване
работоспособността на отделните
МодулDataNode



000149

Детайлно описание за провеждане на тест:

Тестването се извършва чрез използване на следната последователност от команди:

За извеждане на списък на състоянията на **DataNode** модулите на клъстерно ниво:

```
hdfs dfsadmin -D "fs.default.name=hdfs://10.10.6.20/" -report
```

Използвайте **fsck**, за да проверите състоянията на **DataNode** модули, съхраняващи данни на конкретен път (<path>).

```
hdfs fsck <path> // извеждане само на модули със статус „Decomission“
```

```
hdfs fsck <path> -maintenance // извеждане на гореспоменатите модули и на тези със статус „Maintenance“
```

Разпечатка от тестване на работоспособността на модул **DataNode**:

Изход:

```
Configured Capacity: 2532916322304 (2.30 TB)
Present Capacity: 2351330590629 (2.14 TB)
DFS Remaining: 2351325233152 (2.14 TB)
DFS Used: 5357477 (5.11 MB)
DFS Used%: 0.00%
Under replicated blocks: 0
Blocks with corrupt replicas: 0
Missing blocks: 0
Missing blocks (with replication factor 1): 0
```

Live datanodes (3):

```
Name: 10.10.6.14:30010 (wn2-ax.internal.cloudapp.net)
Hostname: wn2-.ax.internal.cloudapp.net
Decommission Status : Normal
Configured Capacity: 844305440768 (786.32 GB)
```



000150

DFS Used: 1785765 (1.70 MB)
Non DFS Used: 15721173083 (14.64 GB)
DFS Remaining: 785616138240 (731.66 GB)
DFS Used%: 0.00%
DFS Remaining%: 93.05%
Configured Cache Capacity: 0 (0 B)
Cache Used: 0 (0 B)
Cache Remaining: 0 (0 B)
Cache Used%: 100.00%
Cache Remaining%: 0.00%
Xceivers: 2
Last contact: Mon Apr 09 14:57:19 UTC 2018

Name: 10.10.6.6:30010 (wn1-ax.internal.cloudapp.net)
Hostname: wn1-ax.internal.cloudapp.net
Decommission Status : Normal
Configured Capacity: 844305440768 (786.32 GB)
DFS Used: 1789952 (1.71 MB)
Non DFS Used: 18485723136 (17.22 GB)
DFS Remaining: 782851584000 (729.09 GB)
DFS Used%: 0.00%
DFS Remaining%: 92.72%
Configured Cache Capacity: 0 (0 B)
Cache Used: 0 (0 B)
Cache Remaining: 0 (0 B)
Cache Used%: 100.00%
Cache Remaining%: 0.00%
Xceivers: 2
Last contact: Mon Apr 09 14:57:19 UTC 2018

Name: 10.10.6.5:30010 (wn0-ax.internal.cloudapp.net)
Hostname: wn0-ax.internal.cloudapp.net
Decommission Status : Normal
Configured Capacity: 844305440768 (786.32 GB)
DFS Used: 1781760 (1.70 MB)
Non DFS Used: 18479804416 (17.21 GB)
DFS Remaining: 782857510912 (729.09 GB)
DFS Used%: 0.00%
DFS Remaining%: 92.72%
Configured Cache Capacity: 0 (0 B)
Cache Used: 0 (0 B)
Cache Remaining: 0 (0 B)
Cache Used%: 100.00%



Cache Remaining%: 0.00%

Xceivers: 2

Last contact: Mon Apr 09 14:57:19 UTC 2018



0

Ръководство на администратора за провеждане
тестване на изпълнението на MapReduce
задание в Hadoop система

000253



/ /

Описание за провеждане на тестовете се извършва чрез следния пример:

Примерът `pi` изчислява цифрите на π , използвайки метода **quasi-Monte Carlo**. Ако не сте добавили потребители към **HDFS**, изпълнете тези тестове като потребител `hddfs`. За да изпълните примера с 16 `map` операции и 1 000 000 проби на `map`, въведете следната команда:

```
$ yarn jar $HADOOP_EXAMPLES/hadoop-mapreduce-examples.jar pi 16 1000000
```

Разпечатка от провеждане на тест на работоспособността на MapReduce:

Ако програмата работи правилно, трябва да видите изход, подобен на следния:

```
Number of Maps = 16
```

```
Samples per Map = 1000000
```

```
Wrote input for Map #0
```

```
Wrote input for Map #1
```

```
Wrote input for Map #2
```

```
Wrote input for Map #3
```

```
Wrote input for Map #4
```

```
Wrote input for Map #5
```

```
Wrote input for Map #6
```

```
Wrote input for Map #7
```

```
Wrote input for Map #8
```

000154



Wrote input for Map #9

Wrote input for Map #10

Wrote input for Map #11

Wrote input for Map #12

Wrote input for Map #13

Wrote input for Map #14

Wrote input for Map #15

Starting Job

...

15/05/13 20:10:30 INFO mapreduce.Job: map 0% reduce 0%

15/05/13 20:10:37 INFO mapreduce.Job: map 19% reduce 0%

15/05/13 20:10:39 INFO mapreduce.Job: map 50% reduce 0%

15/05/13 20:10:46 INFO mapreduce.Job: map 56% reduce 0%

15/05/13 20:10:47 INFO mapreduce.Job: map 94% reduce 0%

15/05/13 20:10:48 INFO mapreduce.Job: map 100% reduce 100%

15/05/13 20:10:48 INFO mapreduce.Job: Job job_1429912013449_0047 completed

successfully

15/05/13 20:10:48 INFO mapreduce.Job: Counters: 49

File System Counters

1

2



000155

FILE: Number of bytes read=358

FILE: Number of bytes written=1949395

FILE: Number of read operations=0

FILE: Number of large read operations=0

FILE: Number of write operations=0

HDFS: Number of bytes read=4198

HDFS: Number of bytes written=215

HDFS: Number of read operations=67

HDFS: Number of large read operations=0

HDFS: Number of write operations=3

Job Counters

Launched map tasks=16

Launched reduce tasks=1

Data-local map tasks=16

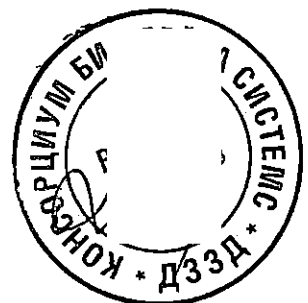
Total time spent by all maps in occupied slots (ms)=158378

Total time spent by all reduces in occupied slots (ms)=8462

Total time spent by all map tasks (ms)=158378

Total time spent by all reduce tasks (ms)=8462

Total vcore-seconds taken by all map tasks=158378



Total vcore-seconds taken by all reduce tasks=8462

Total megabyte-seconds taken by all map tasks=243268608

Total megabyte-seconds taken by all reduce tasks=12997632

Map-Reduce Framework

Map input records=16

Map output records=32

Map output bytes=288

Map output materialized bytes=448

Input split bytes=2310

Combine input records=0

Combine output records=0

Reduce input groups=2

Reduce shuffle bytes=448

Reduce input records=32

Reduce output records=0

Spilled Records=64

Shuffled Maps=16

Failed Shuffles=0

Merged Map outputs=16



GC time elapsed (ms)=1842

CPU time spent (ms)=11420

Physical memory (bytes) snapshot=13405769728

Virtual memory (bytes) snapshot=33911930880

Total committed heap usage (bytes)=17026777088

Shuffle Errors

BAD_ID=0

CONNECTION=0

IO_ERROR=0

WRONG_LENGTH=0

WRONG_MAP=0

WRONG_REDUCE=0

File Input Format Counters

Bytes Read=1888

File Output Format Counters

Bytes Written=97

Job Finished in 23.718 seconds

Estimated value of Pi is 3.14159125000000000000



000158

7

11

1

Ръководство на администратора за провеждане тестване
на комплексната работа на Специализираната
централизирана система под ключ Hadoop чрез прилагане
на Hue приложение



000159

Детайлно описание за провеждане на тест:

Един добър начин да тествате работоспособността на клъстер е като изпълните задача. Освен това можете да тествате клъстера, като изпълните едно от уеб приложенията на Hue. Hue е графичен потребителски интерфейс, който ви позволява да взаимодействате с вашите клъстери чрез стартиране на приложения, които ви позволяват да разглеждате HDFS, да управлявате HIVE metastore, да изпълнявате Hive, Impala и Search заявки, Pig скриптове и Oozie workflows.

1. В Cloudera Manager Admin Console, в раздел Home>Status, изберете Hue услугата.
2. Щракнете върху Hue Web UI връзката, която ще отвори Hue в нов прозорец.
3. Автентикарайте се с credenциите: потребителско име: hdfs, парола: hdfs.
4. Изберете приложение в лентата за навигация в горната част на прозореца на брауъра.

Разпечатка от провеждане на тест чрез прилагане на Hue приложение.

В този пример, тестът се състои в създаване на таблица employee в Hue, чрез HIVE. Следната таблица излага полетата и типовете на данни, които се съдържат в тях, в рамките на таблицата employee.

Sl.No	Име на полето	Тип на данните
1	Eid	int
2	Name	String
3	Salary	Float
4	Designation	String

Посочените по-долу данни са коментари, редове, форматирани полета като полёви терминатор, терминатор на линии и тип съхранен файл.

```
COMMENT 'Employee details'  
FIELDS TERMINATED BY '\t'  
LINES TERMINATED BY '\n'  
STORED IN TEXT FILE
```

Следната заявка създава таблица employee, използвайки данните по-горе.



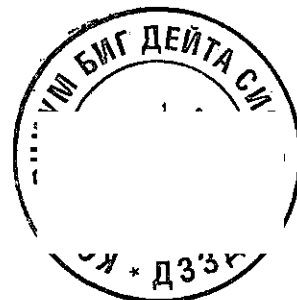
hive> CREATE TABLE IF NOT EXISTS employee (eid int, name String,
salary String, destination String)
COMMENT 'Employee details'
ROW FORMAT DELIMITED
FIELDS TERMINATED BY '\t'
LINES TERMINATED BY '\n'
STORED AS TEXTFILE;

Ако добавите опцията **IF NOT EXISTS**, Hive игнорира оператора в случай, че таблицата вече съществува.

При успешно създаване на таблица, получавате следния изход:

OK
Time taken: 5.905 seconds
hive>

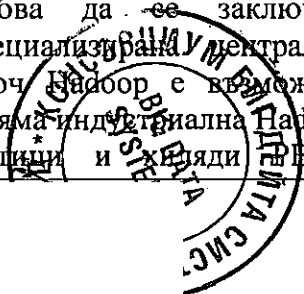
Успешното създаване на таблицата същевременно означава, че тестът чрез прилагане на Hue приложение е успешен.



000161

Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависимости
<p>T.8 Концептуална архитектура за разширение на системата Hadoop за хиляди PetaBytes.</p> <p>- Доставчикът трябва да представи материалът „Концептуалната архитектура за разширение на системата за хиляди PetaBytes (PB)“, отчитайки хардуерните и софтуерни елементи на предложената по-горе Специализирана централизирана система под ключ Hadoop, като се направи нейното теоретично разширение до капацитет хиляди PB. Целта на Концептуалната архитектура е да представи модел, според който предложената Специализирана централизирана система под ключ Hadoop да може да се разшири до стотици и хиляди PB капацитет, използвайки предложените компоненти на специализираната система под ключ. По този начин, Концептуалната архитектура ще докаже валидността на предложената Специализирана централизирана система под ключ Hadoop, за бъдещи разширения и големи индустриални приложения. От концептуалната архитектура трябва да се заключи, че предложената Специализирана централизирана система под ключ Hadoop е възможно да се разшири до голяма индустриална Hadoop система с капацитет стотици и хиляди PB, т.е. тя и нейната</p>	ДА	Ще предоставим обширен материал с наименование „Концептуалната архитектура за разширение на системата за хиляди PetaBytes (PB)“. В този материал са отчетени хардуерните и софтуерни елементи на описаната по-горе Специализирана централизирана система под ключ Hadoop, като се има предвид и нейното теоретично разширение до капацитет хиляди PB. От концептуалната архитектура става ясно, че предложената Специализирана централизирана система под ключ Hadoop е може да се разшири до голяма индустриална Hadoop система с капацитет стотици и хиляди PB.	

000162



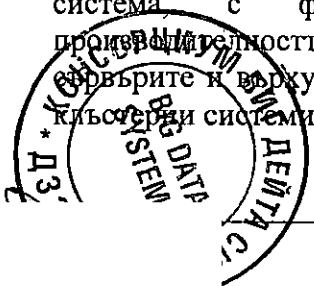
Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависимости
<p>архитектура може да се използват за бъдещо разширение в големи индустриални размери.</p> <ul style="list-style-type: none"> Концептуалната архитектура трябва да съдържа минимум следните части: <p>Същност на системата Hadoop и специфики на Hadoop 3, в която, изхождайки от архитектурните особености на Hadoop да се акцентира върху новите възможности на Hadoop версия 3 за изграждане на Hadoop клъстер;</p> <ul style="list-style-type: none"> Използване в Hadoop на не-RAID архитектури на дискови масиви за създаване на високо производителни DataNode сървърни системи; <p>Технологии влияещи на Hadoop клъстерните архитектури за изграждане на голям капацитет с стотици и хиляди PB, с</p>		<ul style="list-style-type: none"> Концептуалната архитектура отговаря на всички описани в заданието части както следва: <ul style="list-style-type: none"> Същност на системата Hadoop и специфики на Hadoop 3, в която може да се видят архитектурните особености на Hadoop и да се акцентира върху новите възможности на Hadoop версия 3 за изграждане на Hadoop клъстер; Използване в Hadoop на не-RAID архитектури на дискови масиви за създаване на високо производителни DataNode сървърни системи. Технологии влияещи на Hadoop клъстерните архитектури за изграждане на голям капацитет с обеми от стотици и хиляди PB, с акцент върху Същността на 	

000163



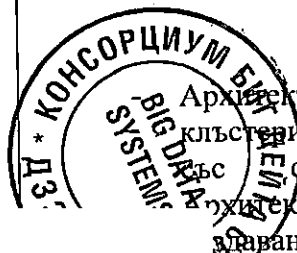
Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависимост
<p>акцент върху Същността на Кодирание срещу изтриване (Erasure coding) – КСИ, върху Кодирание срещу изтриване в Hadoop 3, върху възможностите за увеличава ефективното използвано дисково пространство, и върху Работа с малки файлове;</p> <ul style="list-style-type: none"> - Прилагане на твърдо дисково устройство (HDD) и Полупроводниково дисково устройство (SSD) в Hadoop в сървърни архитектури, заедно със съответните архитектури на дискови контролери; - Метод за повишаване производителността на Hadoop система с фокусиране върху производителността на DataNode сървърите и върху цялостните Hadoop кълъстерни системи; 		<p>Кодирание срещу изтриване (Erasure coding) – КСИ, върху Кодирание срещу изтриване в Hadoop 3, върху възможностите за увеличава ефективното използвано дисково пространство, и върху Работа с малки файлове</p> <ul style="list-style-type: none"> - Прилагане на твърдо дисково устройство (HDD) и Полупроводниково дисково устройство (SSD) в Hadoop в сървърни архитектури, заедно със съответните архитектури на дискови контролери. - Метод за повишаване производителността на Hadoop система, с фокусиране върху производителността на DataNode сървърите и върху цялостните Hadoop кълъстерни системи 	

000164

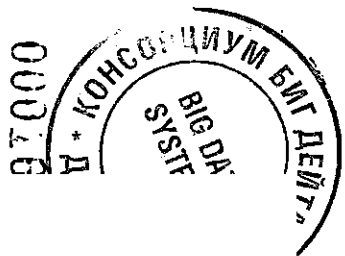


Изискване	Изпълнение (Да/Не)	Начин на реализация
<p>- Да се създаде Хардуерна Референтна архитектура на Hadoop с избор на управляваща система за Референтна архитектура на Hadoop, Типове сървъри използвани в Референтната архитектура, Инфраструктурна архитектура на Референтната архитектура, Прототипна Хардуерна Референтна архитектура на Hadoop, Референтна архитектура на Hadoop за разработка на приложения, Производствена Референтна архитектура на Hadoop, Архитектурни решения за преодоляване на практическата препоръка за 100 TB ограничение на DataNode при Cloudera-Hadoop, и определяне на Ключови индикатори за хардуерна оценка на Сървър за данни (DataNode);</p> <p>Архитектура на големи Hadoop клъстери с обеми стотици и хиляди PB, със специален фокус върху Архитектурен подход към създаването/конфигуриране на</p>		<p>- Създаване на Хардуерна Референтна архитектура на Hadoop с избор на управляваща система за Референтна архитектура на Hadoop, Типове сървъри използвани в Референтната архитектура, Инфраструктурна архитектура на Референтната архитектура, Прототипна Хардуерна Референтна архитектура на Hadoop, Референтна архитектура на Hadoop за разработка на приложения, Производствена Референтна архитектура на Hadoop, Архитектурни решения за преодоляване на практическата препоръка за 100 TB ограничение на DataNode при Cloudera-Hadoop, и определяне на Ключови индикатори за хардуерна оценка на Сървър за данни (DataNode).</p> <p>- Архитектура на големи Hadoop клъстери с обеми стотици и хиляди PB, със специален фокус върху Архитектурен подход към създаването/конфигуриране на</p>

000285



Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависимост
<p>DataNode сървър, Физическа архитектура на Hadoop клъстери с обеми стотици и хиляди PB, Използване на дискове с големи обеми в DataNode сървъри, и Hadoop Агрегирана система състояща се от няколко Hadoop клъстера;</p> <ul style="list-style-type: none"> Литературни източници ползвани при разработка на Концептуалната архитектура. 		<p>DataNode сървър, Физическа архитектура на Hadoop клъстери с обеми стотици и хиляди PB, Използване на дискове с големи обеми в DataNode сървъри, и Hadoop Агрегирана система състояща се от няколко Hadoop клъстера;</p> <ul style="list-style-type: none"> Използваните литературни източници при разработката на Концептуалната архитектура са описани в документа. <p>Прилагаме обширен материал Концептуалната архитектура в документацията.</p>	



„Консорциум Биг Дейта Системс“ ДЗ

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

Т.8 Концептуална архитектура за
разширение на системата Hadoop за хиляди
PetaBytes.



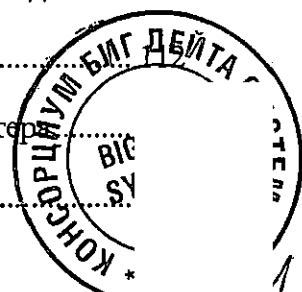
КОНЦЕПТУАЛНА АРХИТЕКТУРА НА СИСТЕМА “HADOOP” ЗА РАЗШИРЕНИЕ ДО ХИЛЯДИ ПЕТАБАЙ

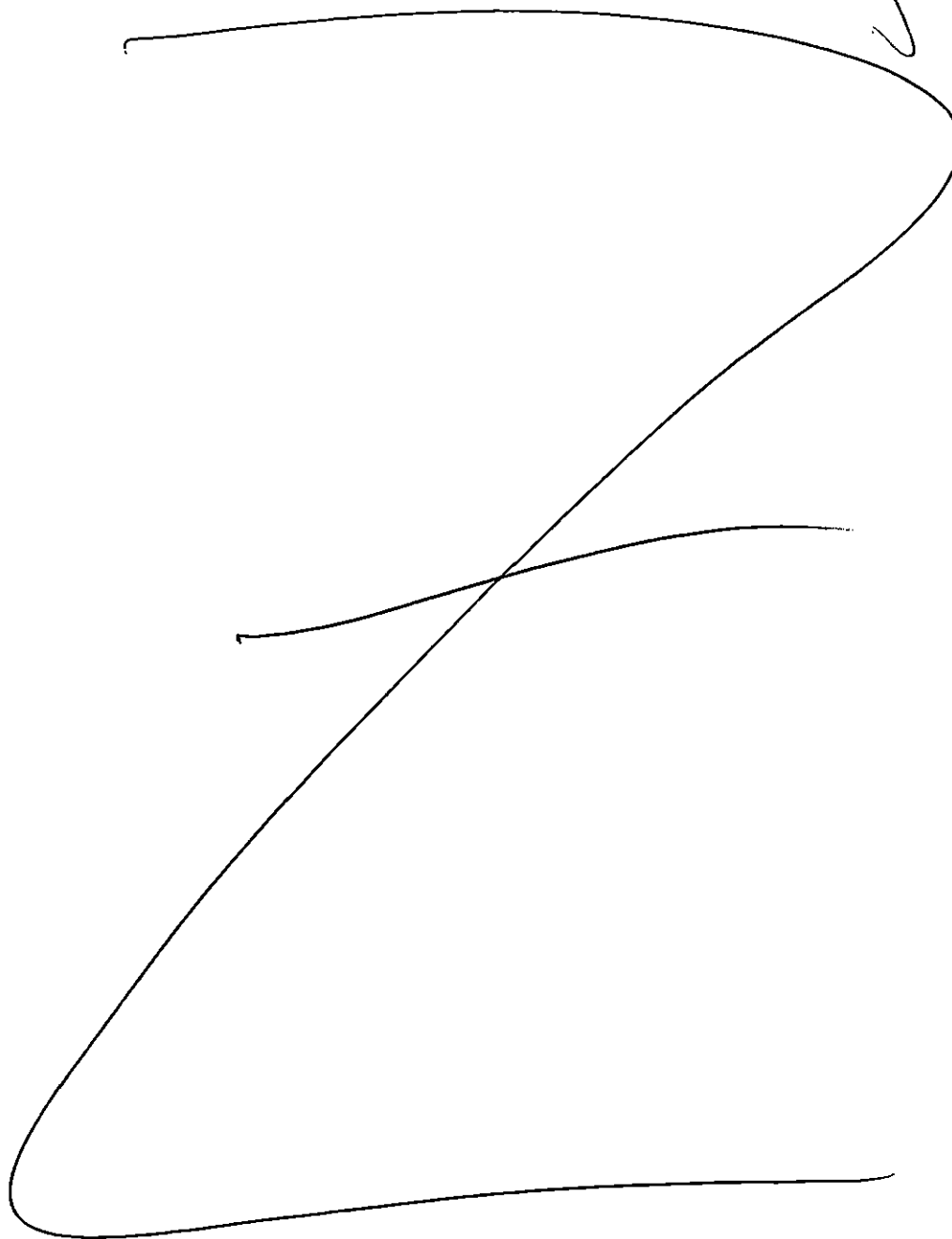
СЪДЪРЖАНИЕ

1	Същност на системата Hadoop и Hadoop 3.....	4
1.1	Големи данни.....	4
1.2	Основни архитектурни особености на система „Hadoop“	11
1.3	Архитектурни особености на Hadoop версия 3	29
2	Използване в Hadoop на не-RAID архитектури на дискови масиви	39
2.1	JBOD архитектура.....	40
2.2	SPAN/BIG архитектура.....	43
3	Технологии влияещи на Hadoop клъстерните архитектури за изграждане на голям капацитет	45
3.1	Същност на Кодирание срещу изтриване (КСИ) за увеличава ефективното използвано дисково пространство.....	45
3.1.1	Кодирание и Проектиране организацията на информацията при DataNode сървъри48	
3.1.2	Кодирание и Проектиране организацията на информацията при NameNode сървъри52	
3.1.3	Отчитане влиянието на кодиране-декодиране при КСИ.....	54
3.2	Кодирание срещу изтриване (КСИ) в Hadoop 3	55
3.3	Работа с малки файлове	60
4	Прилагане на твърдо дисково устройство (HDD) и Полупроводниково дисково устройство (SSD) в Hadoop	66
4.1	Същност на твърдия диск (HDD).....	
4.2	Същност на полупроводниково дисково устройство (SSD)	69
4.3	Основни параметри за сравнение между HDD и SSD	
4.4	Хибридни устройства и дуални системи	



4.5	SAS дискови устройства.....	
4.6	Защо HDD ще бъдат изместени от SSD	75
5	Използване на HDD и SSD устройства в Hadoop	77
5.1	Хетерогенна памет в HDFS – типове и полз	77
5.2	Анализ работата на Hadoop клъстер с използване на SSD устройства	81
5.3	Съвместно използване на KCSI и SSD устройства.....	87
6	Метод за повишаване производителността на Hadoop система.....	88
6.1	Повишаване производителност на HDFS системата	88
6.2	Повишаване производителността на Hadoop клъстера	90
7	Хардуерна Референтна архитектура на Hadoop.....	93
7.1	Избор на управляваща система за Референтна архитектура на Hadoop	94
7.2	Типове сървъри използвани в Референтната архитектура	96
7.3	Инфраструктурна архитектура на Референтната архитектура	97
7.4	Прототипна Хардуерна Референтна архитектура на Hadoop.....	102
7.5	Референтна архитектура на Hadoop за разработка на приложения.....	102
7.6	Производствена Референтна архитектура на Hadoop.....	103
7.7	Архитектурни решения за преодоляване на практическата препоръка за 100 TB ограничение на DataNode при Cloudera-Hadoop	105
7.7.1	Решение чрез използване на малки, 1U сървъри	106
7.7.2	Решение при прилагане на големи, 4U сървъри	106
7.8	Ключови индикатори за хардуерна оценка на Работен сървър	108
8	Архитектура на големи Hadoop клъстери с обеми стотици и хиляди PB	109
8.1	Архитектурен подход към създаването на DataNode сървър	109
8.2	Физическа архитектура на Hadoop клъстери с обеми стотици и хиляди PB... ..	109
8.3	Използване на дискове с големи обеми в DataNode сървъри	
8.4	Hadoop Агрегирана система състояща се от няколко Hadoop клъстери.....	
	Литература.....	





000170

1 Същност на системата Hadoop и H

1.1 Големи данни

„Големи данни“ е терминът за колекция от големи по обем и сложни по структура данни, което затруднява обработването им с помощта на инструменти за управление на релационни бази данни или традиционни приложения за обработка на данни. Трудно е да се приемат, пречистват, съхраняват, търсят, споделят, прехвърлят, анализират и визуализират големи данни. Това са съвкупност, структурирани, неструктурирани и полу-структурирани данни, чийто размер надхвърля възможностите на типичните база данни за съхранение, управление и анализ (особено за обработка на неструктурирани и полу-структурирани данни). Тази огромна информация се пази в хранилища, чийто обем надхвърля десетки, стотици и хиляди Ритабайта. Единиците Байт (B), Килобайт (KB), Мегабайт (MB), Гигабайт (GB), Терабайт (TB), Петабайт (PB), Ексабайт (EB) и Зетабайт (ZB) са в отношение 1:1024 една към друга (1KB=1024B, 1MB=1024KB, 1GB=1024MB, 1TB=1024GB, 1PB=1024TB, 1EB=1024PB, 1ZB=1024EB; примери – 1B е кодът на един символ, 1KB е половин страница текст, 1MB е снимка със средна резолюция, 1GB е 40 минутен филм но със занижени качества на резолюцията – като на не-HD телевизия, 1TB е съдържанието на книгите на библиотека на изследователски институт, 1PB е съдържанието на всички библиотеки на университетите на САЩ, 1EB са данните които целия свят създава в компютри за едно тримесечие на 2018 година, 1ZB е трафикът на целия световен интернет за половин година). Според статистиките, в 5 EB могат да поберат всички думи някога използвани от човечеството. Според Cisco все още не съществува толкова много информация, която да може да запълни 1ZB в една компютърна система.

Съществуват така наречените 5 основни „V“, които характеризират големите данни:

- **Volume** - Обем: Големите данни на първо място са "големи", а размерът се измерва като обем в единици байтове. Обемът се отнася до огромните количества данни, генерирани всяка секунда. Ако вземем например всички имейли, Twitter съобщения, снимки, видеоклипове, данни от които произвеждаме и споделяме всяка секунда. Тези данни са 1 ZB. Така при Facebook се изпращат 10 милиарда съобщения на ден, при които

4,5 милиарда пъти върху бутона "like" се качват 350 милиона нови снимки всеки ден. Данните генерирани в света през цялата 2008 г. сега се генерират за една минута! Това налага създаване на все по-големи хранилища от данни.

- **Velocity** – Скорост: Скоростта се отнася до количеството данни които се генерират за единица време, и количеството данни които се транспортират за единица време. Ако вземем например социалните медийни съобщения които се генерират н секунда, скоростта, с която се съхраняват и обработват транзакциите с кредитни, системите за онлайн и пр. търговия, за да анализират социалните медийни мрежи, за да вдигнат сигнали, които задействат решенията да купуват или продават акции. Съществуващата сега технология на обработка на големи данни позволява анализиране данните докато се генерират, без да записваме в хранилища.
- **Variety** – Разнообразието: Отнася се до различните видове данни, които трябва да се обработват сега. В миналото компютърната индустрия се фокусираше върху структурирани данни, които съответстват на таблици с редове и колони - във вида на релационни данни, съхранявани в релационни бази данни, като например финансови данни (например данна „продажби по продукти“ или данна „региони“). Всъщност 80% от данните в света вече не са структурирани, т.е. те са неструктурирани и поради това не могат лесно да бъдат включени в таблици (текстови документи, имейли, снимки, видео последователности или актуализации в социалните медии). С технологията за големи данни може да се използват различни видове данни (структурирани и неструктурирани), включително съобщения, социални медийни разговори, снимки, сензорни данни, видеозаписи или гласови записи, които да се обработят съвместно с по-традиционни структурирани данни.
- **Veracity** – Истинност: Истинността на данните се отнася до объркаността или достоверността на данните. Колкото повече са данните в едно хранилище, толкова качеството и точността са по-трудно контролируеми (например публикации в Twitter с хаш тагове, съкращения, печатни и разговори (срещи) както и тяхната надеждност и точност на съдържанието), като технологията за обработка на големи данни ни задължава да работим с този вид данни.
- **Value** - Стойност: С нарастването на обема на данните, все по-трудно става да се извлече обобщена стойност от тях, независимо, че това изглежда като най-

данни. Големите данни в индустрията са органично обвързани с индустриалния интернет / интранет. Те се характеризират с голям обем и разнородност, както и с изключително високи темпове на увеличаване обема на масивите. Ето защо за обработката им са необходими усъвършенствани технологии и инструменти. Обикновено Industrial Big Data масивите са по-структурирани и корелативни (от типа по-близки до полу-структурирани данни), по-подредени във времето и по-готови за анализ в сравнение с данните, генерирани в други сфери. Това се дължи на факта, че тези масиви се генерират от автоматизирано оборудване и процеси, при които средата и операциите са силно контролируеми, а човешката намеса е сведена до минимум. Дори да са генерирани от най-съвременните, мрежово свързани и високотехнологични автоматизирани системи, големите масиви индустриални данни все пак притежават характеристики, които ги превръщат в истинско предизвикателство за обработка. Традиционните Big Data аналитични инструменти (свързани с обработка на икономическа информация) са фокусирани върху изследването на връзките между явленията и събирането на информация за тях. IBD анализът обаче е насочен по-скоро към физическите причини за възникването на даден феномен и изискват по-задълбоченото му изследване.

Ето няколко интересни и показателни факти свързани с големите данни:

- Към 2010 г. корпоративните потребители по света са натрупали 7 EB данни ($1 \text{ EB} = 10^{18} \text{ B}$), а на персоналните компютри и ноутбуци на индивидуалните потребители са съхраняват 6 EB информация.
- Към 2010 г. в мрежите на мобилните оператори по света са обслужвани 5 млрд. телефони.
- Всеки месец в социалната мрежа Facebook със свободен достъп се публикуват 30 млрд. нови източници на информация.
- Всяка година обемът на съхраняваната информация нараства с 40%, докато глобалните разходи за ИТ растат с по-малко от 5%.
- Към април 2011 г. в библиотеката на Конгреса на САЩ са съхранявани 235 TB данни ($1 \text{ TB} = 10^{12} \text{ B}$).
- Американските компании от 15 сектора (от общо 17 сектора на икономиката) разполагат с обеми от данни, които са по-големи от тези на Библиотеката на Конгреса на САЩ.



- Датчиците монтирани на авиодвигател генерират около 10 ТВ данни за пс час. Приблизително същите потоци от данни се генерират от дат..... инсталирани на сондажни платформи и нефтопреработващи комплекси. Услугата за разпространение на кратки съобщения Twitter, въпреки ограниченията за дължина на посланието до 140 символа, генерира поток от 8 ТВ за денонощие. Ако всички подобни данни се натрупват за бъдеща обработка, техният общ обем щ и стотици петабайти (1 PB = 10¹⁵ B).

Данните в хранилищата за да предполагат условия за следните анализи, които да генерират резултати по възможност по-близко до реалното време:

- A/B testing. При този метод контролна група от елементи от данни се сравнява с други тестови групи, в които един или няколко показателя са били променени. Целта е да се изясни какви промени подобряват целевите показатели. С този метод може да се намери оптимална комбинация от показатели за достигане на определена цел – например най-добро възприемане на ново маркетингово предложение от страна на потребителите. “Големите данни” позволяват да бъдат проведени огромен брой итерации и да се получи статистически достоверен резултат.

- Association rule learning. Това е набор от методи за идентифициране на взаимовръзки, т.е. асоциативни правила за връзки между променливи величини в големи обеми от данни. Тези методи се ползват в решенията за data mining.

- Classification. Група методи, които позволяват да се предскаже поведението на потребителите в определен пазарен сегмент - например за вземане на решение за покупка, за обем на потребление, за отказ от използване на даден продукт и т.н.. Ползват се в решенията за data mining.

- Cluster analysis. Статистически метод за класификация на обекти по групи на базата на открити общи признаци, които не са били известни предварително. Използва се в решенията за data mining.

- Crowdsourcing. Метод за събиране на данни от голям брой източници



- Data fusion and data integration. Група методи, които позволяват да се анализират коментари на потребители на социалните мрежи и да се съпоставят с резултати от продажби в режим на реално време.

- Data mining. Група методи, които дават възможност за откриване на смислени корелации, зависимости, повтарящи се образци, тенденции и аномалии в масиви от данни. С инструменти за Data mining се реализират проекти за обработка на големи данни които имат за цел да се предскаже модел на поведение на потребители или например да се предскаже каква група потребители ще възприеме най-добре някакъв нов продукт. Счастлива са характерни за най-успешните служители и т.н.

- Ensemble learning. Разработван за целите на машинното обучение, този метод включва в действие множество предикативни модели, благодарение на което постига високо качество на изведените прогнози.

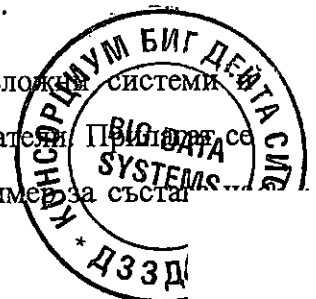
- Genetic algorithms. При този метод възможните решения се представят във вид на "хромозоми", които могат да се комбинират и да мутират. Както и в процеса на еволюцията в природата, оцеляват най-приспособените. Казано по-друг начин, това е евристичен алгоритъм за търсене, използван за решаване на задачи за оптимизация и моделиране чрез случаен подбор, комбиниране и вариации на целеви параметри и използване на механизми аналогични на естествения подбор.

- Machine learning. Направление от областта на изкуствения интелект, насочено към създаването на алгоритми за самообучение на базата на емпирични данни.

- Natural language processing (NLP). Методи за разпознаване и обработка на естествен език, взимствани от информатиката и лингвистиката.

- Network analysis. Група методи за анализ на връзки между възли в мрежи. Прилагат се към данни от социалните мрежи, като позволяват да се анализират връзки между отделни потребители, компании, общности и т.н.

- Optimization. Група от числени методи за редизайн на сложни системи и процеси с цел подобряване на един или няколко техни показателя. Прилагат се за подпомагане вземането на стратегически решения – например за съставяне на маршрутизиране на транспортни средства.



извежданата на пазара продуктова линия, за провеждане на инвестиционни анализи и т.н.

- Pattern recognition. Методи с елементи на самообучение, прилагани за предсказване на модели на поведение.

- Predictive modeling. Методи, които позволяват да бъде създаден математически модел за предварително зададен вероятен сценарий за развитието на дадени събития. Типичен пример за приложението на Predictive modeling е за анализ на данни от CRM система, с цел да се прогнозира възможните условия, при които част от абонатите на дадена компания ^п кажат от нейните услуги и ще започнат да използват тези на конкуренти

- Regression. Група статистически методи за намиране на закономерности между изменението на зависима променлива и една или няколко независими. Използва се за прогнозни анализи и при data mining. Широко се прилага методът “линейна регресия”, който определя влиянието на един числов параметър върху друг. Например, с каква средна стойност се променя обемът на продажбите при промяна на маркетинговия бюджет със 100 лв. Друг вариант е “логическата регресия” – т.е. когато зависимата променлива може да приема само две стойности (0 и 1) – това е един от най-разпространените методи за анализ на вероятността за настъпване на дадено събитие в зависимост от стойностите на някакви параметри.

- Sentiment analysis. Това са методи за оценка на настроеността на потребителите, базирани на технологии за разпознаване на естествен език. Те позволяват от общия информационен поток да бъдат извлечени съобщения, свързани с определена тема или предмет (например за даден продукт), а също така да се оцени дали изказваните по темата съждения са положителни или отрицателни, каква е степента на тяхната емоционалност и т.н.

- Signal processing. Група методи взимствани от радиотехниката, които имат за цел разпознаването на сигнал на фона на шум и анализ на този сигнал.

- Spatial analysis. Група методи взимствани от статистиката за анализ на пространствени данни – топология на местност, географски координати



геометрия на обекти. Източник на големи обеми от такива данни често са системите на крупни организации.

- Supervised learning. Методи базирани на технологиите за машинно обучение, които позволяват да бъдат открити функционални взаимовръзки в анализирания масив от данни.

- Simulation. Моделирането на поведението на сложни системи често се използва за прогнозиране и изпробване на различни сценарии за целите на планирането.

- Time series analysis. Група методи взаимствани от статистиката и от теорията за обработка на цифрови сигнали. Използват се за анализ на повтарящи се във времето поредици от данни. Типични приложения – проследяване на пазара на ценни книжа, проследяване на броя пациенти с някакво заболяване и т.н.

- Unsupervised learning. Методи базирани на технологиите за машинно обучение, които позволяват да бъдат идентифицирани скрити функционални взаимовръзки в анализирания масив от данни. Тези методи имат общи черти с Cluster Analysis.

- Visualization. Методи за графично представяне на резултатите от анализи на “големи данни” във вид на диаграми или анимирани изображения с цел опростяване на интерпретацията и по-лесно разбиране на получените резултати.

1.2 Основни архитектурни особености на система „Hadoop“

Hadoop е софтуерна Рамка с отворен код, използвана за разпределено съхранение и обработка на големи данни, прилагайки програмния модел MapReduce. Тази Рамка се инсталира върху компютърни клъстери, изградени от сървъри, организирани по ракове (Racks). По този начин компютърен клъстер с инсталирана върху него софтуерната Рамка Hadoop създават Система “Hadoop”. Всички модули в софтуерната Рамка Hadoop са проектирани с основното предположение, че хардуерните повреди са често срещани явления и независимо от тях Системата “Hadoop” трябва да продължава да обработва данните. В Системата “Hadoop” се зарежда за изпълнение потребителски приложения, които са насочени основно за модела MapReduce. Една приложна програма се разглежда като едно задание (job), което може да се разделя на по-малки наречени задачи (task). Отделните задачи може да се изпълняват в паралел на



рвъри. Данните в Hadoop се съхраняват във файлове, управлявани от файлова система HDFS (Hadoop Distributed File System).

Моделът MapReduce служи за разпределена обработка на големи обеми от данни, разположени върху разпределени компютри (отличителни компютри свързани чрез мрежа). Приложната програма представлява задание, съдържащо 3 компонента:

- Map метод е метод изпълняващ се в паралел от схема от сървъри за обработка на входни данни, и специално в онези сървъри в които се намират части от входните данни (предварително заредени в съответните сървъри). Този метод обработва като локални данни тези заредени в HDFS и тези които са част от входните данни, като и резултатът от обработката се записва във временни локални файлове – HDFS файлове. Файловете в Hadoop – HDFS файловете, се представят чрез записи, всеки във формат <key, value> (<ключ, стойност>). Map методът се изпълнява в сървър от системата “Hadoop” - сървър за обработка на данни - DataNode;
- Shuffle метод (переразпределяне), осигуряваща трансфериране на обработените от map данни към сървърите, които ще извършват обработката reduce, като данните от временните файлове на Map метода се поставят в един сървър (сървър за обработка на данни - DataNode) за обработка от Reduce метод;
- Reduce метод (редуциране), изпълняващ операциите по формиране на крайния резултат, като в един сървър (сървър за обработка на данни - DataNode) са обработват данни с еднакъв ключ и всичките такива обработки се изпълняват в паралел.

MapReduce е парадигма на програмиране заимствана от функционалното програмиране, позволяваща масивна скалируемост на стотици или хиляди сървъри в клъстер на Hadoop.

Обработката на MapReduce заданието обикновено разделя въведените данни на независими парчета, които се обработват като задачите за паралелно работещи сървъри. MapReduce сортира изходите от сървърите определени от схемата за разположение на данни, които след това се въвеждат в задачите за Reduce. Обикновено както входът, така и изходът на заданието се съхраняват в файловата система HDFS. Hadoop рамката се грижи за планирането на задачи, мониторинга им и повторен изпълнение на неуспешните задачи.



0000173

Ядрото на Hadoop се състои от *част за съхранение на данни*, известна като Hadoop Distributed File System (HDFS), и *обработваща част*, която е модел за програмиране и изпълнение на програмен код върху данните, наречен MapReduce. Hadoop разделя файловете на големи блокове и ги разпределя всеки в няколко сървъра в клъстера. След това той създава от едно задание една или няколко задачи, като всяка задача прехвърля за изпълнение в сървъри, изхождайки от принципа – къде (в кой сървър) са данните които трябва да се обработват. Това позволява обработването на набор от данни по-бързо и по-ефективно, отколкото при по-конвенционалната суперкомпютърна архитектура, която разчита на паралелна файлова система, в която изчисленията и данните се разпределят чрез високоскоростна мрежа.

За оперативност в настоящия документ, за *наименуване* на Рамката Hadoop ще се използва терминът просто "Hadoop", което е и *стандартното* разбиране в техническата литература.

Основата на Hadoop се състои от следните модули:

- i. Hadoop Common - съдържа библиотеки и помощни програми, необходими на други модули Hadoop;
- ii. Hadoop Distributed File System (HDFS) - разпределена файлова система, която съхранява данни на сървъри;
- iii. Hadoop YARN - платформа, отговорна за управлението на изчислителните ресурси в клъстерите и използването им за планиране на приложения на потребителите;
- iv. Hadoop MapReduce - прилагане на програмния модел MapReduce за мащабна пакетна обработка на данни.

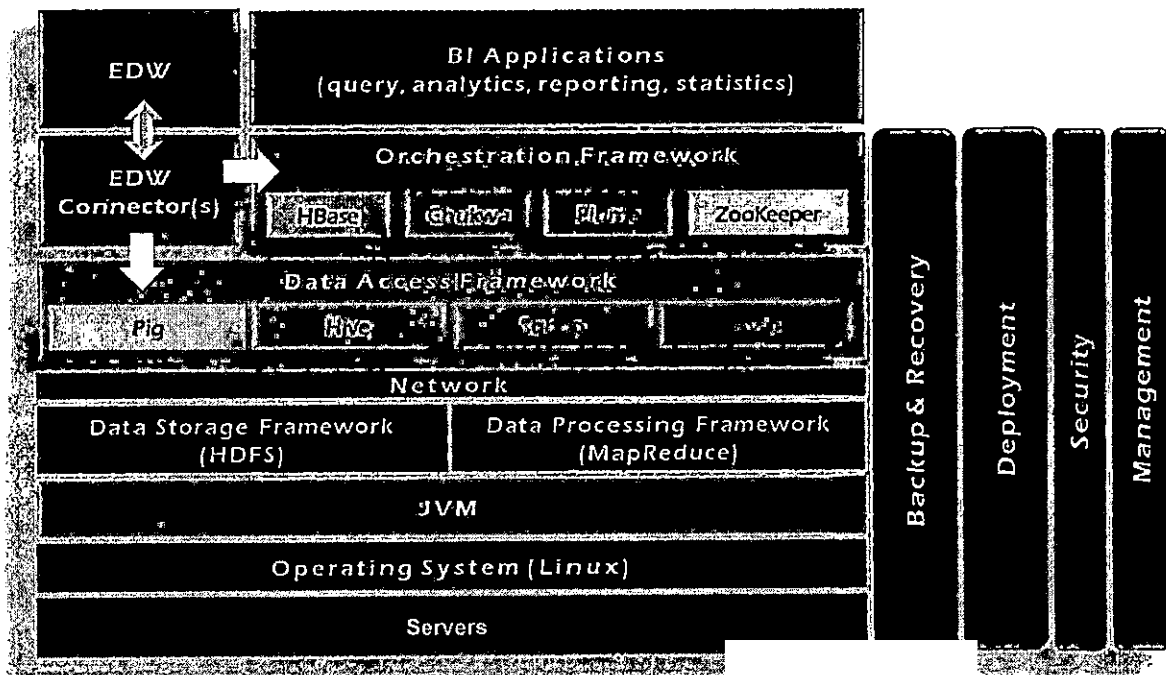
Терминът "Hadoop" се отнася не само до гореспоменатите базови модули и подмодули, но и до *екосистемата* или до включването на допълнителни софтуерни пакети, които могат да бъдат инсталирани на или надстроени над Hadoop, като Apache Pig, Apache Hive, Apache HBase, Apache Phoenix, Apache Spark, Apache ZooKeeper, Cloudera Impala, Apache Flume, Apache Sqoop, Apache Oozie, Apache Storm и пр. Hadoop е написан основно на езика Java, с някои елементи написани на езика C, както са използвани и определени помощни програми, написани като shell, написани също на езика Java. Независимо че за програмиране на MapReduce се използва основно езика Java.

000130



всеки език за програмиране може да се използва с "Hadoop Streaming" за изпълнение "map" и "reduce" части от потребителска програма.

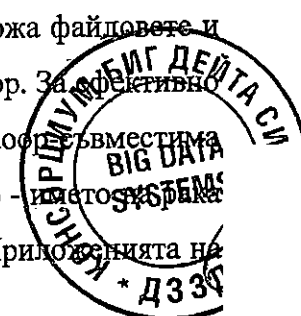
Концептуалната архитектура на Hadoop екосистемата е показана на фиг.1-1.



Фиг.1-1

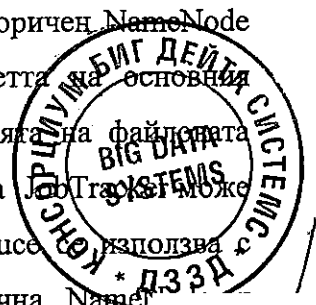
където EDW (Enterprise Data Wharehouse) представлява Склад от данни на предприятие, JVM е Java Виртуалната машина, която изпълнява Java програмен код, BI (Business Intelligence) е Бизнес интелигентност, Backup&Recovery е програмно средство за временно съхраняване и възстановяване при проблеми в оперирането, Deployment са програмните средствата за внедряване, Security са средствата за информационна сигурност, а Management са програмните средства за управление на цялостната екосистема.

Hadoop се състои от пакета Hadoop Common, който осигурява абстракции на ниво файлова система и операционна система, модулът за изпълнение на изчислителните ресурси MapReduce (MapReduce във версия 1 на Hadoop или YARN във версия 2) и Hadoop Distributed File System (HDFS). Пакетът Hadoop Common съдържа файловете и скриптовете на Java ARchive (JAR), необходими за стартиране на Hadoop. За ефективно планиране на изпълнението на приложната програма, всяка Hadoop е съвместима файлова система трябва да осигури осведоменост за местоположението - и местосъхраняване (или по-точно на мрежовия комутатор към който е свързан сървър). Приложенията на



Hadoop могат да използват тази информация, за да изпълнят по възможност програмния код в сървър, където са данните, а ако това не е възможно, да се използва сървър в същия рак / комутатор, за да се намали трафика на основната мрежа на Системата "Hadoop". HDFS използва този метод при репликиране на данни за съкращаване на данни в множество ракове. Този метод намалява въздействието от прекъсване на захранването на рака или на кому. възникне някоя от тези откази на хардуер, данните ще останат налични за обработка.

Минимална Система "Hadoop" включва един Главен възел (Master node) състоящ се от един сървър и множество работни възела (DataNodes) представляващи един или няколко сървъра. Тези сървъри физически са поставят в ракове, представляващи метални шкафове, в които като „по чекмеджета“ се поставят отделните устройства, в нашия случай сървъри и комутатори. В Главният възел се зареждат модулите JobTracker (за управление на заданията) и NameNode (за управление на имената), докато в Работният възел се зареждат модулите TaskTracker (за управление на задачи) и DataNode (за управление на данните). По принцип JobTracker е модул за проследяване изпълнението на задания, но за оперативност ще го наричаме „модул за управление на заданията“. По аналогичен начин, модулът TaskTracker е за проследяване изпълнението на задачи, но за оперативност ще го наричаме „модул за управление на задачи“. В сървъра DataNode се записват самите данни на файловете – данните на HDFS. Един блок данни се записва обикновено в 3 различни DataNode сървъра – два в един и същ рак, а третия в сървър намиращ се в друг рак. Първото копие на данните се поставя с първи сървър в определен рак, второто копие се зарежда във втори сървър в друг рак различен от този на първия сървър, а третото копие се зарежда в трети сървър в същия рак като на първия сървър. В сървъра NameNode се зареждат метаданни на файлове и тяхното физическо разположение – по подобие на директория на типична файлова система. В този сървър се създава и регистър – (log) за всички създадени и изтрени файлове. В по-голяма система "Hadoop" се използва специален сървър като основен NameNode за съхранение на индекса на файловете и евентуален вторичен NameNode сървър, съхраняващ моментни снимки на структурата на паметта на основния NameNode сървър, като по този начин се предотвратява корупцията на файловата система на Hadoop и загубата на данни. По същия начин модулът на JobTracker може да се разположи в самостоятелен сървър. Когато Hadoop MapReduce използва алтернативна файлова система, архитектурата NameNode, вторична NameNode



DataNode на HDFS се заменя от еквивалентите, специфични за файловата система. Хардуерно, DataNode може да бъде сървър или обикновен десктоп компютър.

Едно MapReduce приложение, представляващо задание, съдържа заявките за разположение на входно-изходните данни като предоставя видовете обработки. Тези и други параметри на заданието включват конфигурацията на заданието. След това клиентът за работа с Hadoop подава както заданието (jar / executable и т.н.), така и конфигурацията на JobTracker, който поема отговорността за разпространение на софтуера, планиране на задачите и наблюдение на тяхното изпълнение, като предоставя статус и диагностична информация на самия клиент (компютърът от който е стартирано клиентското приложение). Рамката MapReduce работи изключително върху двойки <key, value> (<ключ, стойност>), т.е. рамката приема входните данни към заданието като набор от двойки <ключ, стойност> и изпраща да набор от двойки <ключ, стойност> като изход на работа.

По време на изпълнение на MapReduce заданието, Hadoop разпределя задачите "Map" и "Reduce" до съответните сървъри в клъстера. Рамката Hadoop управлява всички подробности за преминаването на данни, като издаване на отделни задачи, проверка на завършването на задачите и копиране на данни в клъстера между отделните сървъри. Повечето изчисления се извършват в DataNodes с локални дискове, които намаляват трафика в мрежата. След приключване на дадените задачи клъстерът събира данните, за да формира подходящ резултат, и го изпраща обратно на сървър в Hadoop.

Друг начин да се погледне на MapReduce обработката е чрез 5-стъпковото паралелно-последователно разпределено изчисление:

- Подготовка на входа за метода Map - системата "MapReduce" определя сървъри за обработка на данни (DataNodes), присвоявайки стойността на ключ от входните данни, например K1 (според <ключ, стойност> на входните данни) които даден сървър ще обработва, и осигурява на този сървър всички входни данни, свързани с тази ключова стойност;
- Отделният сървър за обработка на данни изпълнява потребителския метод Map, който се изпълнява в дадения сървър, като се генерира резултат (също във формат <ключ, стойност>), като се създава нов резултатен ключ, например K2.
- Методът Shuffle разпределя данните от сървърите обработили Map метода към сървъри, които ще обработват Reduce метода, присвоявайки на всеки резултатен

000183



обработка на Reduce метода на ключ равен на такъв от резултатните ключ формиращи се в сървърите предназначени за изпълнение на Map метода;

- Отделният сървър за обработка на данни изпълнява потребителския метод Reduce, като всеки сървър работи с данни с ключ, определен от метода Shuffle, например тези с ключ K2;
- Формиране на крайния резултат - системата MapReduce събира всички изходи от сървърите, изпълнили метода Reduce и ги подрежда по ключ, например K2, за да произведе крайния резултат.

В много случаи входящите данни вече могат да бъдат намиращи се в различни сървъри в резултат от предишна обработка, като в този случай първата стъпка може да бъде много опростена - само присвояване на сървъри за Map обработка, които обработват наличните входни данни. По подобен начин, стъпка 3 може да бъде ускорена, като присвоите сървъри за обработка на Reduce метода да бъдат възможно най-близо до генерираните от Map данни. Потребителските методи Map и Reduce, обикновено написани на езика Java, са елементарни програми, изхождайки от обстоятелството, че обработката на големи данни (тези данни в Hadoop) често изискват не сложни алгоритмични обработки.

Файловата система HDFS съхранява големи файлове (обикновено в диапазона от гигабайта до терабайта) на няколко машини. Всеки файл се разбива на големи блокове (по подразбиране 64 MB за Hadoop версия 1, или 128 MB за Hadoop версия 2). Постига се надеждността на данните чрез репликиране на данните в множество компютри и оттам теоретично не изисква резервно пространство за съхранение на независими дискове (RAID) на компютри (но за да се увеличи ефективността на входно-изходните (I / O) конфигурации на RAID все още са полезни). При репликация на данните по подразбиране - 3 броя, данните се съхраняват на три възела: два в един и същ рак и една в различен рак. Отделните сървъри на данни (Data Nodes) могат да разговарят помежду си, за да възстановят равновесието на данните, да преместят копията на задачи наоколо (в рамките на същия рак) и да запазят желаната репликация на данните. HDFS не е напълно съвместима с POSIX, защото изискванията за POSIX файлова система се различават от целевите цели на приложението Hadoop. Компромисът за липса на пълна съвместима с POSIX файлова система е за сметка на поддържане на производителност свързана с придвижването на данните и поддръжка на операции, като например операция Append.

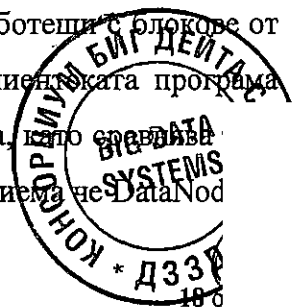


000184

HDFS добавя възможности за висока надеждност, което позволява сървърите метаданни (NameNode) да има резервно копие. Файловата система HDFS включва наречената вторичен сървър за метаданни (бекъп NameNode сървър), който се включва в опериране, когато основния NameNode сървър излезе офлайн. При нормална работа второстепенният NameNode сървър периодично се свързва с първичния NameNode сървър, като създава снимки на директорията на първичния NameNode, която снимка се копира като директория на вторичния NameNode. Тези контролирани снимки на директория могат да се използват за възстановяване и рестартиране на повреден и след това поправен първичен NameNode, без да се налага да се възпроизвежда цялото съдържание на първичния NameNode (да се редактира дневника, за да се създаде структурата на директорията и пр.). HDFS допуска да има множество пространства на имена, обслужвани от отделни NameNodes. Едно от предимствата на използването на HDFS е осведомеността за разположение на данни между инструмента за управление на задания и за проследяване изпълнението на задачи. Програмата за управление на задания създава карта за изпълнение на отделното задание, водещо до формиране на отделни задачи, свързани с местоположението на данните. Например: ако сървър А съдържа данни (x, y, z) и сървър В съдържа данни (a, b, c), инструментът за управление на задания подготвя сървър В за извършване на задача - Map или Reduce операции, с използване на данни (a, b, c), а на сървър А ще бъде планирана задача - да изпълни Map или Reduce операции за данни (x, y, z). Това разбиване на задание на задачи намалява размера на трафика, който пресича мрежата, и предотвратява нужното прехвърляне на данни от сървър в сървър. Когато Hadoop се използва в файлови системи, това предимство не винаги е налице. Това може да окаже значаен влияние върху времето за завършване на работа.

Мета данните които се съхраняват в NameNode се прехвърлят за обработка в Оперативната памет на сървъра, което увеличава бързодействието на цялата Система Hadoop. Метаданните съдържат например Списък на файловете, Списък на блоковете за всеки файл, Списък на информационните номера за всеки блок, Файлови атрибути, като време за създаване, репликационен фактор и пр. При работа с файлове, HDFS работи с т.н. „коригиращи суми“ – използвайки CRC32 метода, работещ с блокове от 512 байта, като коригиращите суми се записват в DataNode. Клиентската програма извършва създаване на коригираща сума за всяка част от 512 байта, като след това сумата се записва в DataNode. При несъвпадение се приема че DataNode

000285



работи коректно и обработката се пр
ъщите данни.

ode, съдържаща репли

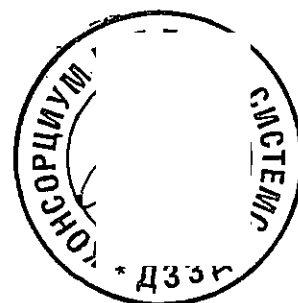
imeNode следи за работоспособността на DataNodes, като извършва избор на нови DataNodes за нови копия, Балансира използването на диск в DataNode, Балансира комуникационния трафик към DataNodes и пр.

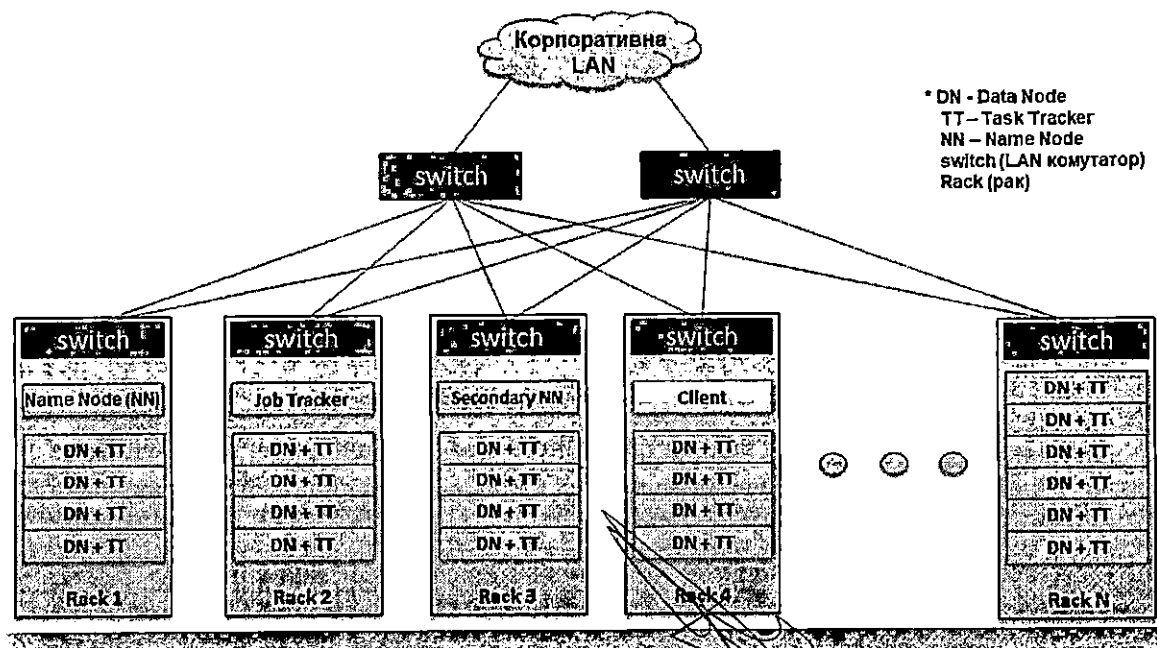
HDFS е разработена най-вече за преобладаващо не подлежащи на промяна файлове и може да не е подходяща ефективно за приложения, изискващи конкурентни операции по записване.

В обобщение, функциите на Hadoop се свеждат до:

- Съхранява и обработва големи и / или бързо развиващи се набори от данни
- Оперира със Структурирани и неструктурирани данни, като фокусът е върху неструктурираните данни
- Притежава Прости модели на програмиране
- Разрешава Висока мащабируемост и наличност (не спиране от опериране)
- Използва евтини (конвенционални) компютри с малко резервиране
- Осигурява толерантност на надеждността при хардуерни проблеми
- Фокусира се към Преместване на изчислителните процеси там където са данните, а не данните да се преместват там където са изчислителните мощности

Концептуалната Инфраструктурна архитектура на Hadoop с възможни разположения на софтуерните компоненти на Hadoop по ракове е представена на фиг.1-2

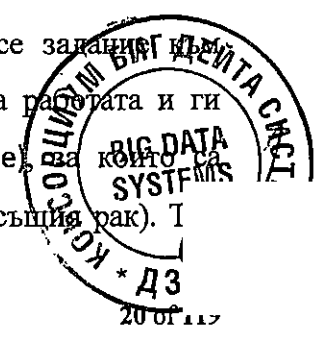




Фиг.1-2

Достъпът до файлове може да бъде постигнат чрез native Java API, Thrift API (генерира клиент на няколко езика, например C ++, Java, Python, PHP, Ruby, Erlang, Perl, Haskell, C #, Cocoa, Smalltalk и OCaml), интерфейсът за команден ред, уеб приложението HDFS-UI, чрез HTTP или чрез библиотеки на трети страни за мрежови клиенти. HDFS осигурява преносимост на различни хардуерни платформи и е съвместим с различни операционни системи. Обаче, HDFS въвежда определени ограничения на преносимостта, тъй като внедряването на Java не може да използва функции, които са изключителни за платформата, на която работи HDFS. Поради широкото си интегриране в инфраструктурата на ниво предприятие, наблюдението на производителността на HDFS по мащаб става все по-важен въпрос. Мониторингът на ефективността от край до край изисква да се проследят показателите от базите данни, имената и основната операционна система. Понастоящем има няколко платформи за мониторинг за проследяване на ефективността на HDFS, като например HortonWorks, Cloudera.

Логически погледнато, клиентското приложение изпраща MapReduce задания на JobTracker, JobTracker преобразува съответното задание в задачи за работата и ги изпраща на онези TaskTracker на сървъри (работещи като DataNode, които са възможно най-близо данните (данните са в самия сървър или поне в същия пак). I



000187

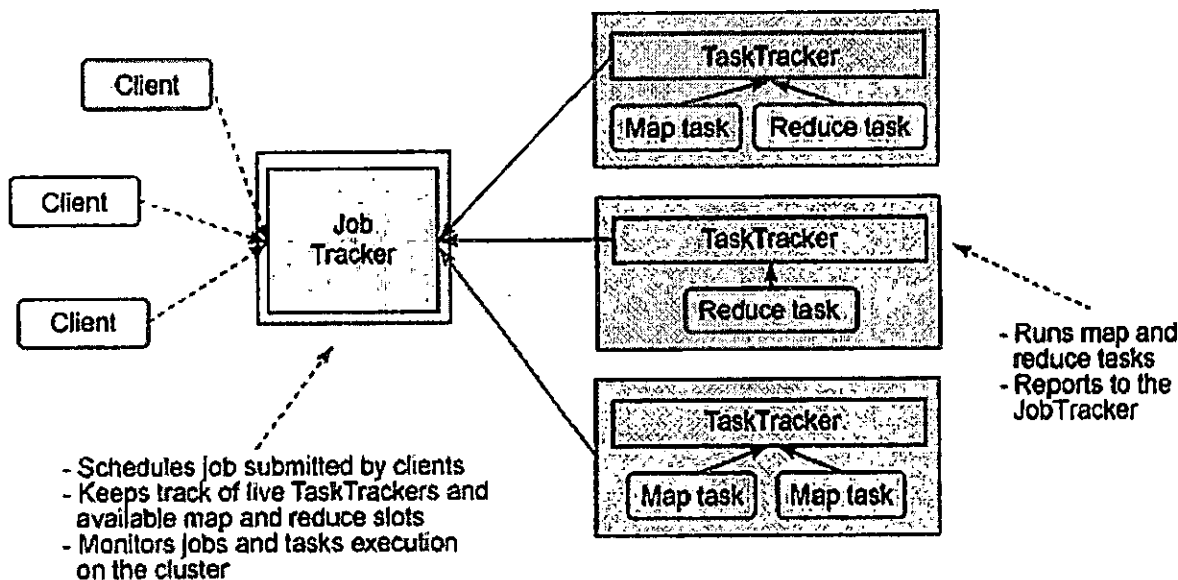
като HDFS файловата система е съобразена с разположението на данните по рако
JobTracker знае кой сървър съдържа данните и кои други сървъри (DataNodes) са в ра....
Ако работата не може да бъде хоствана на действителния сървър (DataNode), където се
намират данните, приоритет се дава на сървърите в същия рако. Това намалява
мрежовия трафик на мрежовата инфраструктура на Hadoop. Ако задачата се провали
или не може да се изпълни, то тя се пренасочва към друг DataNode. TaskTracker на
всеки сървър създава отделен процес на Java виртуална машина (JVM), за да
предотврати провала на самия TaskTracker, ако текущата задача срие JVM.
Информация за изпълнението на заданието се изпраща от TaskTracker до JobTracker на
всеки няколко минути, за да контролира състоянието му. Състоянието и информацията
за проследяване на заданията от JobTracker и на задачите от TaskTracker се показват от
Jet. Разпределението на работата на TaskTrackers е много проста. Всеки TaskTracker
има няколко свободни слота (например "4 слота"), като една активна задача (map или
reduce) заема един слот. При избор на сървър (DataNode) на който да се изпрати задача
за изпълнение, не се отчита текущото натоварване на сървъра, както и не се отчита
дали в момента този сървър е наличен (работещ) или не. Ако един TaskTracker е много
бавен, той може да забави цялото задание на MapReduce. По подразбиране Hadoop
използва FIFO график с алтернативни 5 приоритета за планиране, за да планира задачи
от работна опашка.

Действията на MapReduce се изпълняват от JobTracker и TaskTracker модулите.
Модулът JobTracker разделя заданието на по-малки задачи (Map) и ги изпраща за
изпълнение на отделните TaskTracker. Модулът TaskTracker отчита обратно на
JobTracker за извършената работа и изпраща получените резултатни данни за следваща
обработка. Приложната програма работи обикновено със съществуващи в Hadoop
файлове, т.е. приложните програми обикновено са ориентирани към използване на
съществуващи файлове. При записване на информация, клиентската програма получава
от NameNode списък от DataNodes в които да зареди блоковете от данни. Клиентската
програма записва данните в първия DataNode от списъка, като самия DataNode
осигурява записване на реликата от данни във втори DataNode, без участие на
приложната програма. И така до третия или всички DataNode, на които трябва да се
създаде реликата от тези данни (когато подразбиращия брой реплики на данните не е
3 а повече). След като и последния DataNode запише последната реплика



зания на приложната програма, че тези данни са записани, за не записването следващите данни. И това всичко се управлява от HDFS.

Типично изпълнение на MapReduce приложение е показано графично на фигура 1-3.



Фиг.1-3

При създаване на Hadoop 2 (втора версия), най-съществената разлика спрямо Hadoop 1 (първа версия) е технологията YARN. В първата версия на Hadoop, основните компоненти включват Hadoop Common, HDFS и MapReduce, но втората версия на Hadoop е с нова технология наречена YARN - Yet Another Resource Negotiator (YARN). Това е технология за управление на ресурсите. YARN се стреми да разпредели ресурсите за различни приложения ефективно. Той управлява два софтуерни демона (агенти), които се грижат за две различни задачи: проследяване на работата и наблюдение на напредъка. Тези два демона се наричат съответно *Мениджър на ресурсите* и *Магистър на приложението*. В понятията на YARN приложението представлява съвкупност от задачи които трябва да се изпълнят заедно. По този начин понятието „приложението“ в рамките на YARN отговаря на понятието „Задание“ в MapReduce. Мениджърът на ресурси разпределя ресурси за различни задачи, докато Магистърът на приложението следи изпълнението на отделната задача. YARN се състои от 3 компонента:

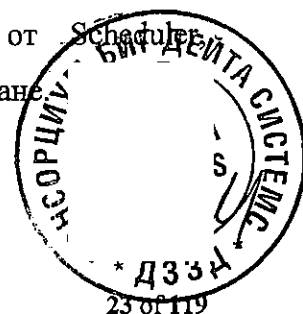
- ResourceManager (един на Hadoop клъстер) – Мениджър на ресурсите
- ApplicationMaster (по един на приложение) – Магистър на приложението
- NodeManagers (по един на DataNode) – Мениджър на сървъра (на DataNode).



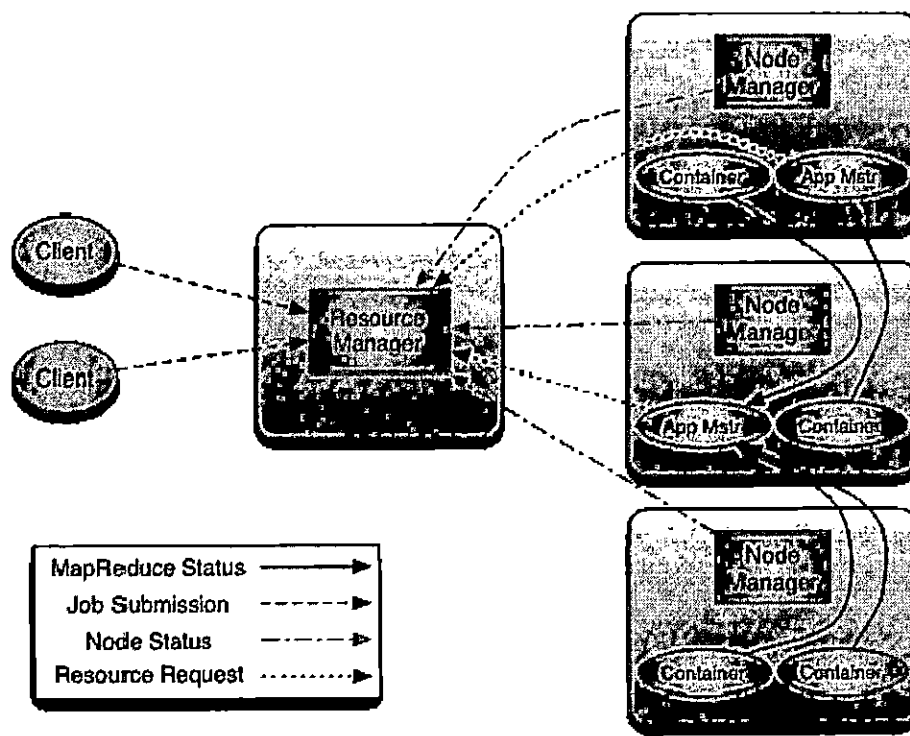
Основната идея на YARN е да раздели функционалностите на управлението на ресурсите от планирането / мониторинга на работата в отделни програмни модули - нодони. Идеята е да има глобален ResourceManager (RM) и ApplicationMaster (AM) за всяко приложение. ResourceManager и NodeManager образуват една рамката за изчисляване на данни. ResourceManager е програмнен модул, който решава назначаването на ресурсите за всички приложения в системата. ResourceManager има два основни компонента: Scheduler и ApplicationsManager. NodeManager е друг програмнен модул - агентът, разположен във всеки DataNode, който отговаря за контейнерите от ресурси (процесор, памет, диск, мрежа) и наблюдава тяхното използване, като докладва същото на ResourceManager - частта му Scheduler. Понятието „контейнер“ в Hadoop/YARN представлява контейнер от ресурси (Resource Container - RC), представляващ съвкупност от физически ресурси. Scheduler е отговорен за разпределянето на ресурси за различните изпълнявани приложения, в рамките на познати ограничения като капацитета, опашките и т.н. Scheduler е компонент за планиране в смисъл, че не извършва наблюдение или проследяване на състоянието на приложението. Също така, той не предлага гаранции за рестартиране на неуспешни задачи поради неработоспособност на приложението или хардуерни повреди. Scheduler изпълнява своята функция за планиране въз основа на ресурсните изисквания на приложенията; тя се основава на абстрактното понятие за ресурсен контейнер, който включва елементи като памет, процесор, диск, мрежа и т.н. Scheduler има програмно управляема (вграждаща се) политика, която е отговорна за разделянето на ресурсите на клъстерите между различните опашки, приложения и т.н. Текущите планиращи програми като CapacityScheduler и FairScheduler са примери за такива вграждащи се политики.

Компонентата ApplicationMaster е всъщност специфична библиотека и е натоварена с договаряне на ресурси от ResourceManager и работа с NodeManager (и), за да изпълнява и наблюдава задачите. ApplicationManager приема заявки за работа, договаря първия контейнер за изпълнение на ApplicationMaster и предоставя услугата за рестартиране на контейнера на ApplicationMaster при неуспех. Компонентата ApplicationMaster има отговорността да договаря подходящи контейнери с ресурси от Scheduler, проследявайки техния статус и наблюдение за напредъка на функционирането.

000190



На фигура 1-4 е показано графично взаимодействието на трите компоненти на YARN Resource Manager, Node Manager и Application Master, заедно със съответните контейнери на отделните сървъри DataNode, при изпълнение на клиентски програми

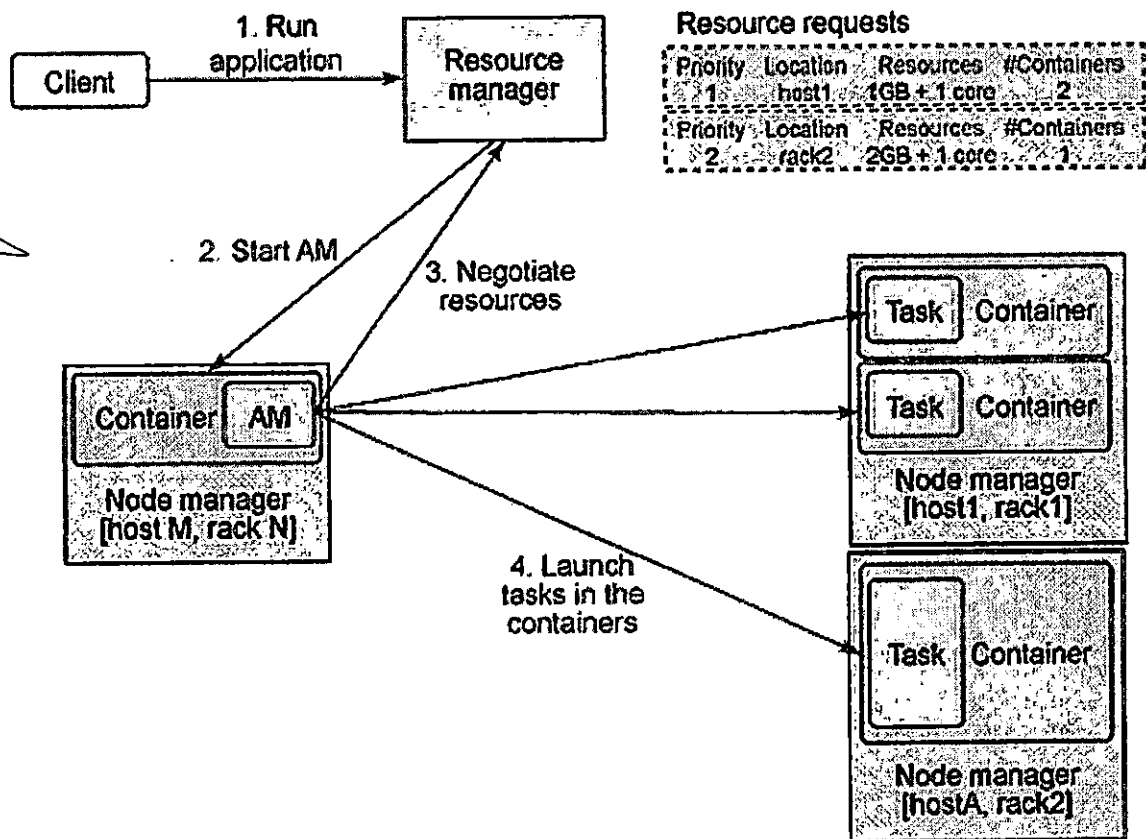


Фиг.1-4

Чрез компонентата YARN, Hadoop може да изпълняват интерактивни заявки едновременно с групови задания за пакетна работа (batch) на MapReduce. Разделянето на HDFS от MapReduce посредством YARN прави средата на Hadoop по-подходяща за операционни приложения, като аналитичности, които не могат да чакат до края на задачите за пакетна обработка.

Resource Manager, Application Master, Node Managers и контейнерите взаимодействат заедно, когато приложение е изпратено до клъстер на YARN. Пример за работа на приложение в среда YARN е показано на фигура 1-5.

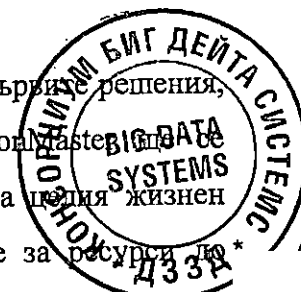




Фиг.1-5

Да предположим, че потребителите изпращат приложения към ResourceManager, като напишат командата `hadoop jar` по същия начин, както в MapReduce v1. ResourceManager поддържа списъка с приложения, които се изпълняват в клъстера, и списъка с наличните ресурси на всеки активен NodeManager. ResourceManager трябва да определи кое приложение да получи следващата част от ресурсите на клъстера. Решението е подложено на много ограничения, като капацитет на опашката, система за достъп до ресурси (ACL) и справедливост при разпределянето на ресурси. ResourceManager използва вграден Scheduler. Програмата Scheduler се фокусира само върху планирането; тя управлява кой получава клъстерните ресурси (под формата на контейнери) и кога, но не извършва никакъв мониторинг на задачите в рамките на дадено приложение, както и не се опитва да рестартира неуспешни задачи.

Когато ResourceManager приема нова заявка на приложение, едно от първите решения, които Scheduler прави, е избора на контейнер, в който ApplicationMaster ще се изпълнява. След като ApplicationMaster стартира, той е отговорен за целия жизнен цикъл на това приложение. Първо и най-вече той изпраща искане за ресурси до ResourceManager да поиска контейнери, необходими за изпълнение на задачите на



приложението. Искането за ресурс е просто заявка за няколко контейнера, които отговарят на определени изисквания за ресурсите, като например:

- Количество на ресурсите, изразена като мегабайти от паметта и части от CPU;
- Предпочитано местоположение, посочено по име на сървър
- Приоритет в рамките на даденото приложение, а не в рамките на цялото приложение.

Ако и когато е възможно, ResourceManager предоставя контейнер (изразен като ID на контейнера и име на сървър - хост), който удовлетворява изискванията, изисквани от ApplicationMaster в искането за ресурс. Контейнер позволява на приложението да използва определена група от ресурси на конкретен сървър. След като получи контейнер, ApplicationMaster ще поиска от NodeManager (който управлява хоста, на който е бил разпределен контейнера), да използва тези ресурси, за да стартира специфична за приложението задача. Тази задача може да бъде всеки процес, написан във всяка рамка (като задача от типа MapReduce или задача от типа Giraph). NodeManager не проследява задачи; той само следи използването на ресурсите в контейнерите и например спира контейнер, ако консумира по-голяма памет от първоначално разпределената. ApplicationMaster провежда преговори за установяване на контейнери, за да стартира изпълнението на всички задачи, необходими за изпълнение на приложението. Той също така следи напредъка на приложението и неговите задачи, рестартира неуспешните задачи в новите заявени контейнери и докладва напредъка на клиента, стартирал приложението. След като приложението завърши, ApplicationMaster се затваря и пуска свой собствен контейнер. Въпреки че ResourceManager не извършва никакъв мониторинг на задачите в дадено приложение, той проверява „здравословното състояние“ на ApplicationMasters. Ако ApplicationMaster пропадне, той може да бъде рестартиран от ResourceManager в нов контейнер. Може да се каже, че ResourceManager се грижи за ApplicationMasters, докато ApplicationMasters се грижи за задачите в приложението.

YARN предлага ясни предимства по отношение на мащабируемостта, ефективността и гъвкавостта в сравнение с класическата технология на MapReduce от първата версия на Hadoop. И двата - малки и големи кълстера Hadoop имат голяма полза от YARN. За крайния потребител (програмист, а не администратор) промените са почти неощутими.

000193



20 01 117

защото е възможно в среда на YARN да се изпълняват непроменени задания на MapReduce, като се използват същите API MapReduce и CLI.

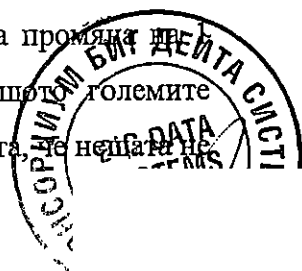
Най-екстремните приложения на Hadoop (Facebook, Amazon, Yahoo, Netflix, NYTimes и др.) работят със Системи Hadoop с обеми по-големи от 100 PB (петабайта), като тези обеми растат с около по половин PB на ден. Примерна голяма Система Hadoop се състои от 10 000-ди DataNodes, 100 милиона файла, в рамките 100 PB общ обем.

Когато клиентска програма иска да напише файл в HDFS се изпълняват следните стъпки:

- Клиентската програма трябва да раздели файла на блокове и ще трябва да изпрати заявка за запис до NameNode.
- За всеки блок NameNode ще предостави на клиентската програма списък, съдържащ IP адреса на DataNodes (в зависимост от фактора на репликация, по подразбиране 3), където блокът с данни трябва да бъде копиран в крайна сметка.
- Клиентската програма ще трябва да копира първия блок в първия DataNode и след това другите копия на блока ще бъдат копирани от самите DataNodes автоматично.

По принцип не могат да модифицират файловете, които вече са налице в HDFS, тъй като HDFS следва модела Запиши веднъж Чети много пъти (Write Once Read Multi). Но винаги могат да добавя данни към съществуващия файл на HDFS. За да се модифицира всяка част от файл, който вече е записан, трябва да се презапише целия файл и да се замените стария файл. Това важи даже и при необходимост от модифициране на 1 байт. Независимо, че предложения подход изглежда ужасно неефективно, но това е без значение, защото големите приложения за обработка на данни обикновено са изградени около идеята, че нещата не се променят на парчета. В такъв случай как хората променят файловете, съхранявани на HDFS? Краткият отговор на въпроса е, че за да се промени част от файл, който вече е записан, трябва да се презапише целия файл и да се смени стария файл с новосъздадения, включително и когато става въпрос за промяна на един байт. Това колкото и да изглежда неефективно няма значение, защото големите приложения за обработка на данни обикновено са изградени около идеята, че нещата не

000194



се променят често, на парче. За да се направи промяна на съществуващ файл, мож се приложат един от следните 2 варианта:

- *Вариант 1 – добавяне към съществуващ файл:*

```
echo "<Text to append>" | hdfs dfs -appendToFile - /user/hduser/myfile.txt
```

или

```
hdfs dfs -appendToFile - /user/hduser/myfile.txt
```

и след това се въвежда текста от клавиатурата, завършвайки 'Ctrl+D'

- *Вариант 2 – вземане на HDFS файла и прехвърляне в локалната файлова система, модифициране и след това записването му като нов HDFS файл със същото име:*

```
hdfs dfs -get /user/hduser/myfile.txt
```

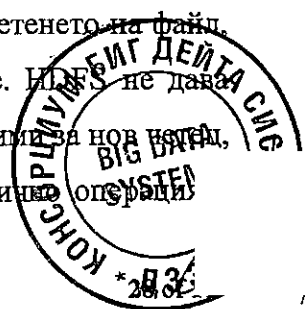
vi myfile.txt #или да се използва кой да е инструмент за модификация

```
hdfs dfs -put -f myfile.txt /user/hduser/myfile.txt
```

Няколко клиентски програми не могат да пишат едновременно в един и същ HDFS файл. HDFS следва модела с множествен четене и едно писане в един файл едновременно. Клиентската програма, която отваря файл за писане, получава лизингов договор от NameNode. Ако междувременно някоя друга клиентска програма иска да напише в този файл ще трябва да поиска от NameNode разрешението за писане. Първоначално NameNode ще провери дали лизинговият договор за записване в този специален файл е бил предоставен на някоя друга програма или не. В този случай тя ще отхвърли заявката за писане на другата клиентска програма, ако лизинговият договор е придобит от някоя друга програма, която в момента пише в самия файл.

Може да се прочете файла, който вече е отворен. Но проблемът при четенето на файл, който понастоящем съществува, е в съгласуваността на данните, т.е. HDFS не дава гаранция, че данните, които са били записани във файла, ще бъдат видими преди файлът да бъде затворен, За тази цел може да се извика изрично операцията

000195



hflush, която ще запише всички данни в буфера в тръбопровода за запис, а след тог операцията hflush ще изчака потвържденията от DataNodes. Следователно, когато изпълни тази команда, данните, които са били записани във файла операцията hflush, ще бъдат видими за читателите със сигурност.

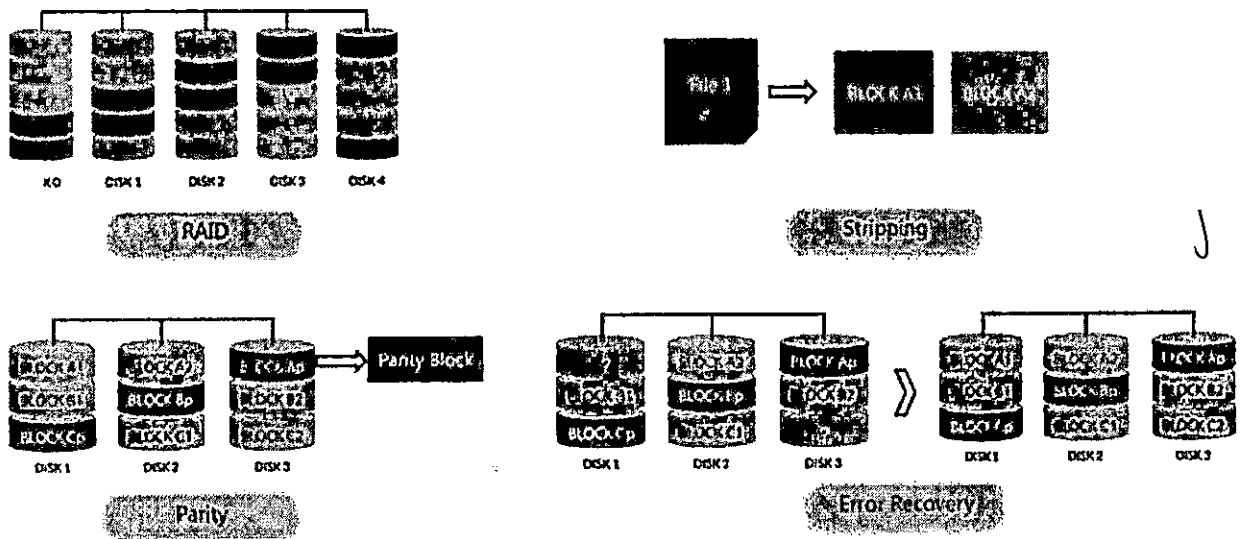
1.3 Архитектурни особености на Hadoop версия 3

Hadoop 3 (трета версия на Hadoop) е надстройка над Hadoop 2, като добавя следните допълнителни функции:

- i. Минималната версия на JVM е 8, с нейните допълнителни функционалности;
- ii. Кодиране срещу изтриване (Erasure encoding) - КСИ

Кодирането срещу изтриване (КСИ) може да се представи чрез следните последователни технологически стъпки, представени на фигура 1-6. По принцип, в системите за данни, Кодирането срещу изтриване (Erasure Coding) се използва като технологията в онези компютърни системи, съставени от евтини дискове. В съвременните сървъри тази КСИ технология се възприема все повече. Така както е в RAID технологията, Кодирането срещу изтриване работи чрез разделяне на един файл на последователност от отделни по-малки блокове (stripping) – КСИ блокове или наричани още „под-блокове“ и съхраняване на тези КСИ блокове на различни дискове. При това, за всеки от тези КСИ блокове се изчислява и записва един или няколко паритетни блокове, разположени на други дискове, не на тези на които са разположение самите оригинални КСИ данни. Грешка причинена от изтриване на данните в КСИ блоковете може да бъде възстановена чрез изчисляване, използващо паритетните блокове.





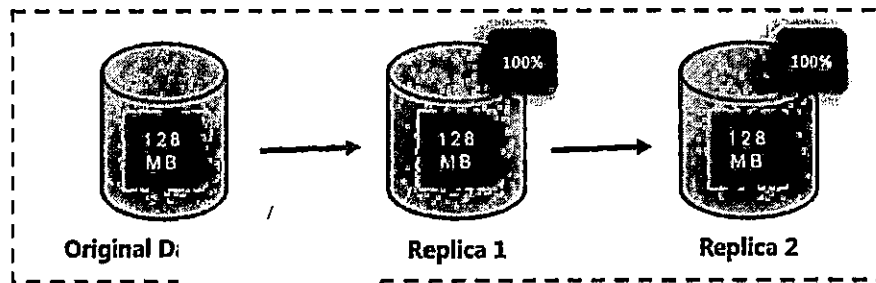
Фиг.1-6

При работа с RAID технологията, когато се повреди информацията в един диск, то цялата RAID дискова инфраструктура преминава към възстановяване, което е толкова по-продължително, колкото е обемът един от дисковете. По време на това възстановяване, не съществува възможност при последващо нарушаване на данните, те да се възстановят. И ако такова последващо нарушаване настъпи, се получава хардуерна грешка със спиране на системата. С навлизането в експлоатация на дискове с голям капацитет (8TB, 10TB), се оказва, че периодът на възстановяване на RAID данните става все по-продължителен. Този период може да бъде от минути до часове. Ако се вземе предвид и че през този период дисковете се използват за друго нормално опериране (изземващ от общия капацитет за обмен на данни на самите дискове), то възстановяването на такива RAID дискове с голям капацитет, може да достигне до много часове и дни. По тази причина, с въвеждането на дискове с капацитет десетки терабайта, светът търси алтернатива за RAID технологията.

В архитектурата на Hadoop се използва разделяне на един файл на блокове от данни (64MB за Hadoop версия 1, 128MB за Hadoop версия 2 и 3), като всеки блок се репликира (по подразбиране броят на репликираните копия е общо 3, представляващо един оригинален блок от данни и 2 броя реплики) - фиг.1-7.



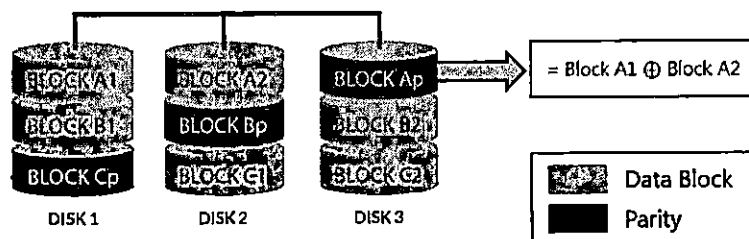
000197



Фиг.1-7

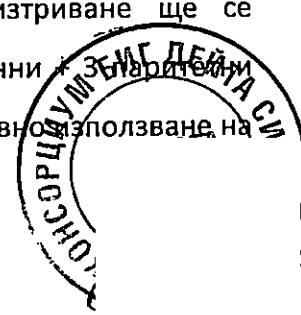
Понеже тези репликирани копия са 100% копие на оригиналния блок от данни, то се получава 200% неефективно изразходване на дисковото пространство (при общ брой на копия на блоковете от данни - 3).

Кодирането срещу изтриване (КСИ) осигурява същото ниво за толерантност към грешки като при репликиране в Hadoop (до Hadoop версия 2 включително), но с много по-малко неефективно изразходване на дисковото пространство – фигура 1-8.



Фиг.1-8

Този подход предлага само 50% неефективно използване на дисковото пространство, понеже Паритетния блок се прилага за 2 КСИ блока. Интегрирането на Кодиране срещу изтриване в сравнение с HDFS версия 2, може да поддържа същата толерантност към грешките на данните. Например, при файл с 6 под-блока, с прилагане стандартните принципи на Hadoop (до Hadoop версия 2 включително) на 3 броя репликирани копия, ще се консумира $6 * 3 = 18$ броя под-блока от дисково пространство, докато с въвеждането на Кодиране срещу изтриване ще се консумират само 9 броя под-блока (6 под-блока оригинални данни + 3 паритетни блока) на дисковото пространство. Това прави само 50% неефективно използване на дисковото пространство.



За да се поддържа Кодирането срещу изтриване в HDF зор версия 3
въвеждат следните допълнителни архитектурни решения:

a) *NameNode Разширение* - HDFS файловете са групирани чрез КСИ блокови групи, които имат определен паритетни блокове. За да се намали потреблението на паметта на NameNode сървъра за тези допълнителни паритетни блокове, беше въведен нов протокол за именуване на йерархични блокове. При това, всеки КСИ блок има Идентификационният номер в рамките на КСИ блокова група, което позволява управлението да бъде на нивото на КСИ блоковата група, а не на КСИ блок.

b) *Разширения на интерфейса за клиентски програми* - След въвеждане на Кодирането срещу изтриване в HDFS, NameNode работи на ниво КСИ блокова група, като достъпът за четене и запис на клиентската програма е усъвършенстван чрез възможност за работа с КСИ блокове и паралелна работа на няколко вътрешни КСИ блока в рамките на КСИ блокова група. В тази връзка:

- При работа със записване, DFSStripedOutputStream управлява набор от потоци на данни, по един за всеки DataNode, който съхранява КСИ блок от текущата блокова група. Координаторът поема отговорности за операциите по цялата блокова група, включително за прекратяване на текущата блокова група, разпределение на нова блокова група и др.;
- При работа с четене, DFSStripedInputStream осигурява КСИ блока от DataNode да се постави в съответната КСИ блокова група, като се допуска паралелна работа на отделните DataNodes.

c) *Разширения за DataNode* - DataNode изпълнява допълнителна задача ErasureCodingWorker (ECWorker) за възстановяване при неуспешно кодирани блокове за изтриване. Неуспешни блокове на КСИ се откриват от NameNode, който след това избира съответния DataNode, за да извърши възстановяването. Реконструкцията изпълнява три основни задачи:

- Прочитане на данните от изходните сървъри - само тези от КСИ блоковата група за възстановяване (заедно с паритетния блок



- Създава се ново кодиране срещу изтриване на прочетените блокове от данни;
- След като кодирането приключи, възстановените блокове от КСИ блоковата група се прехвърлят към съответните DataNodes за да запишат новите КСИ блокове.

d) *Политика на Кодиране срещу изтриване* – политика чрез която се приспособяват хетерогенните работни натоварвания, позволява файловете и директорите в HDFS клъстера да имат различна репликация на КСИ. За целта информацията за кодирането и декодирането на файлове се капсулира в клас ErasureCodingPolicy. Той съдържа 2 броя информация, т.е. Схема на КСИ (ECSchema) & размера на КСИ блока.

iii. Усъвършенстване на YARN Timeline услугата

YARN Timeline услугата се изпълнява от YARN Timeline сървъра, който в по-ранните версии се наричаше Generic Application History Server. По принцип, YARN Timeline сървъра изпълнява следните 2 функции:

- Обща информация за завършени приложения

Общата информация включва данни за ниво на изпълнение на приложение, като име на опашка, потребителска информация и т.н. в ApplicationSubmissionContext; списък на опитите за приложение, които се изпълняват за дадено приложение; информация за всеки опит за работа на приложението; списък на контейнерите, изпълнявани под всеки опит за работа на приложение; информация за всеки контейнер. Общите данни се съхраняват от ResourceManager в хранилище на история (по подразбиране в файловата система) и се използват от веб интерфейса за показване на информация за завършени приложения.

- Подробна информация за текущите и завършените приложения

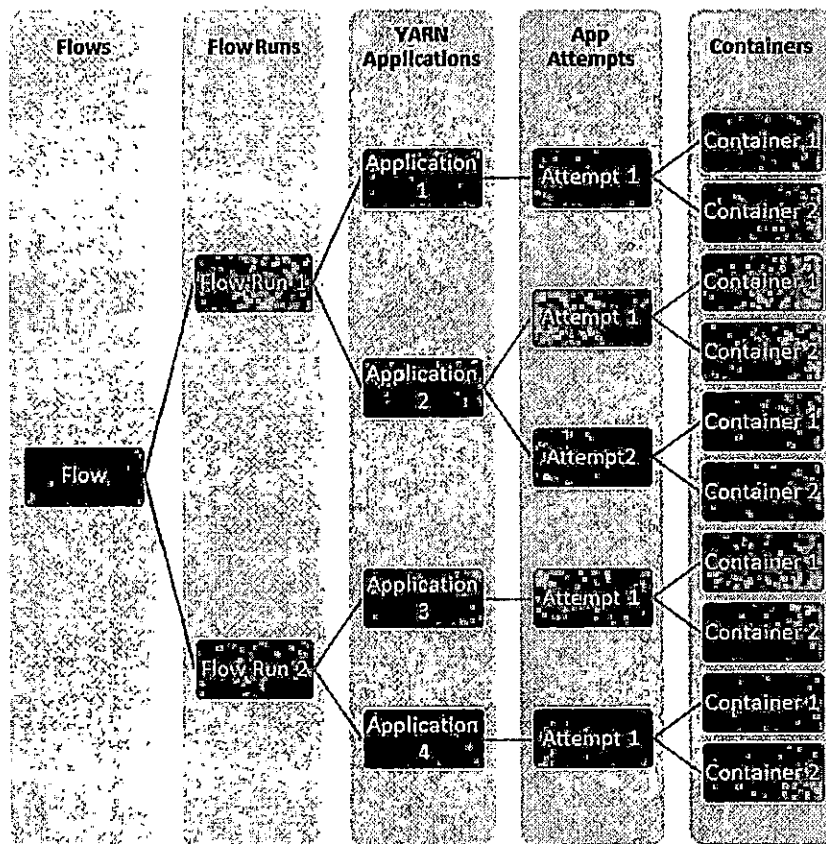
Подробната информацията е напълно специфична за приложение. Например, рамката на Hadoop MapReduce може да включва части от информация като брой задачи на Map, брой задачи на Reduce, броячи и т.н. Разработчиците на приложения могат да публикуват информация в Timeline сървъра чрез използване

000,00



TimelineClient от клиентската програма. След това тази информация може да б задавана чрез REST API за предоставяне на потребителски интерфейси.

Във версия 3 на Hadoop (версия 2.0 на YARN Timeline) се прави по-добро използване на тази услуга. Подобрието за използване е свързано с "потоци" или логически групи от приложения на YARN. Много по-често е да се стартира набор или серия от приложения на YARN за завършване на окрупнено бизнес приложение. Тази услуга поддържа изрично понятието „потоци“. Освен това тя поддържа обобщаване на показателите на ниво поток, както е показана на фигура 1-9 по-долу.



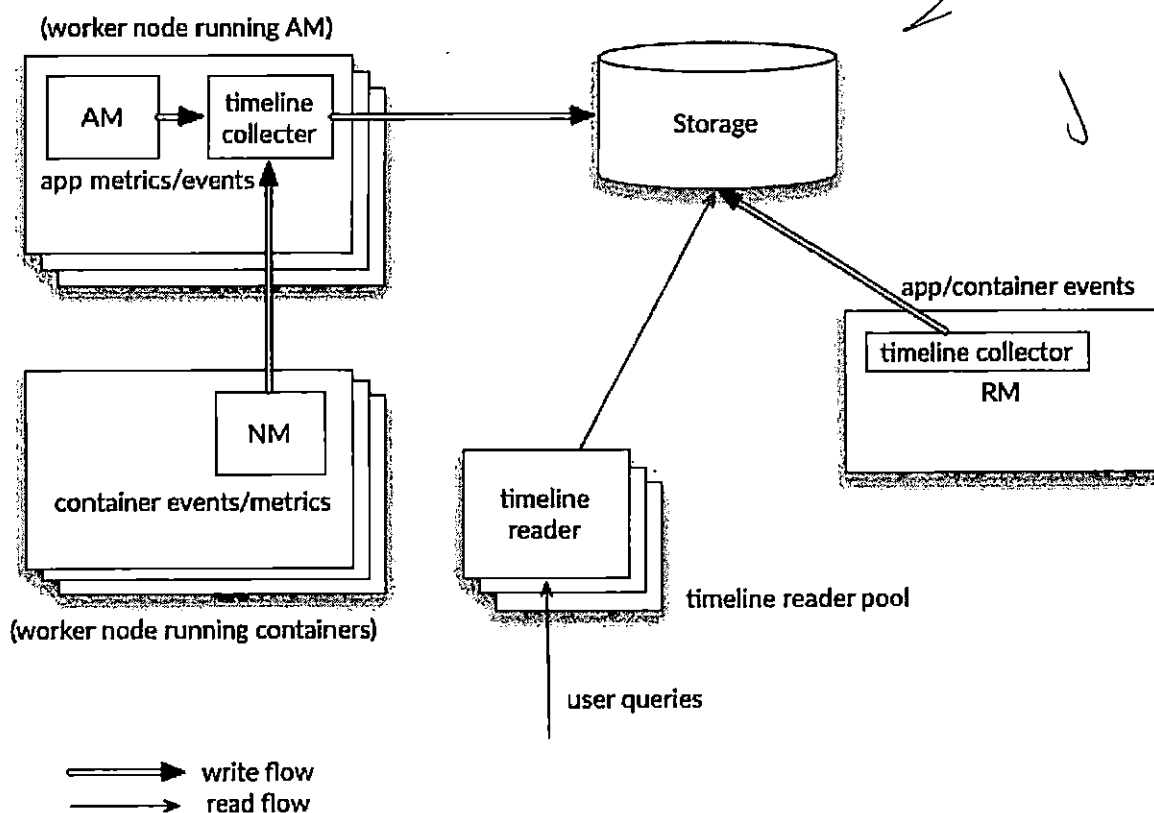
Фиг.1-9

YARN Timeline Service използва набор от колектори за записване на данни в базовото хранилище. Колекторите се разпределят и се намират съвместно с Application Masters (AM), към които са обвързани. Всички данни, които принадлежат на това приложение, се изпращат към колекторите, с изключение на колектора за мениджъра на ресурси. За дадено приложение Application Master записва данни за приложението в съвместно разположените колектори (което е поместено в услугата Node Manager - NM). Освен това мениджърите на възли от други възли, в които се



изпълняват контейнерите за приложението, също записват данни в колектора, изпълнява главното приложение. Resource Manager (RM) също поддържа свой собствен колектор. Той излъчва само YARN-генерични събития, за да запази обема си в ограничен обем. Компонентите за четене са отделни програмни демони, отделени от колекторите на приложенията и са посветени на сервизиране на заявки чрез REST API. Обобщено тази функционалност е показана на фигура 1-10.

фиг. 1-10



фиг. 1-10

iv. Пренаписване на кода за Shell

Скриптовете на Hadoop shell са пренаписани, за да поправят много бъгове, да разрешат проблеми със съвместимостта и да променят вече съществуващи инсталации. Също така са включени някои нови функции.

v. Поделени клиентски jar файлове



В Hadoop версия 3 има нови компоненти – `hadoop-client-api` и `hadoop-client-runtime` които премахват зависимостта от един `jar` файл. `hadoop-client-api` има действие в време на компилиране, докато `hadoop-client-runtime` има обхват по време на изпълнение на програма. Така че може да се свържат зависимостите в един `jar` файл предотвратявайки изтичането на зависимостите от `classpath` на приложението. Например, HBase може да разговаря с клъстер на Hadoop, без да вижда каквито и да било зависимости свързани с внедряването в конкретна инсталация.

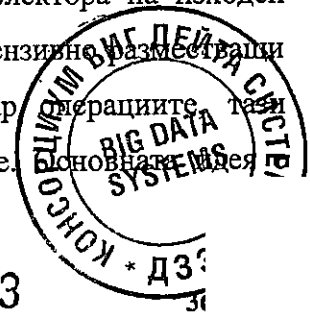
vi. Подкрепа за опортюнистични контейнери и разпределено планиране

Въведен е нов `ExecutionType`, т.е. Опортюнистични контейнери, които могат да бъдат изпратени за изпълнение в `NodeManager`, дори ако в момента на планиране няма налични ресурси. В такъв случай тези контейнери ще бъдат поставени на опашка в `NodeManager`, чакайки наличните ресурси за да започнат да се изпълняват. Опортюнистичните контейнери са с по-нисък приоритет в сравнение с Гарантираните контейнери (тези по подразбиране) и осигуряват приоритетна работа на гарантираните контейнери. Това води до да подобряване използването на клъстера.

Гарантираните контейнери съответстват на съществуващите контейнери на YARN. Те се разпределят от Диспечера и след като бъдат изпратени до сървър / възел, започва тяхното изпълнение, ако има налични ресурси за изпълнението им. След като се стартират, тези контейнери се изпълняват до завършване, стига да няма откази. Опортюнистичните контейнери се разпределят по подразбиране от централния `ResourceManager`, но е възможно те да се разпределят чрез разпределен планиращ план, който се изпълнява като интерфейс за `AMRMProtocol`.

vii. MapReduce оптимизация на ниво задача

В Hadoop 3, в MapReduce е добавена Java реализация за колектора на изходен резултат на Map, на базата на JNI интерфейса. За работа с интензивно разпространяващи се задания от гледна точка на изходни резултати от Map операцията тази възможност подобрява производителността с 30% или повече.



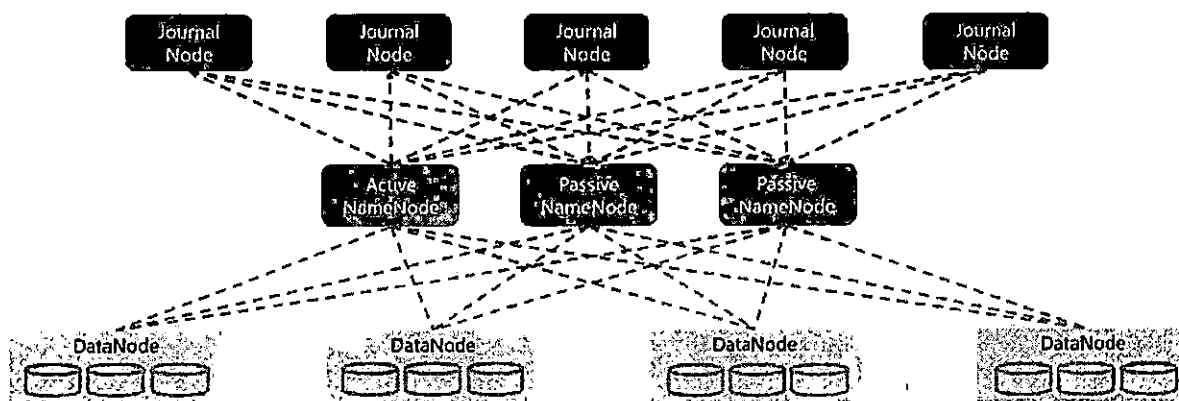
000.03

7

добавяне на NativeMapOutputCollector за обработка на двойката „ключ-стойност“ създадена от Map процеса.

viii. Поддържане на повече от 2 броя NameNode сървъра

В Hadoop 2, архитектурата за висока надеждност работи с един активен NameNode и един NameNode на изчакване (пасивен режим на работа). Чрез възможността за репликиране на три JournalNodes в Hadoop 3, тази архитектура е в състояние да толерира провала на който и да е от тези NameNode. В допълнение на това, в Hadoop 3 се разрешава да се използват няколко NameNode на изчакване, например 1 активен и 2 пасивни NameNode сървъра. По такъв начин чрез конфигуриране на три броя NameNodes сървъра (1 активен и 2 пасивни) и пет броя JournalNodes, Hadoop клъстерът може да понесе провала на два възела NameNode – фигура 1-11.

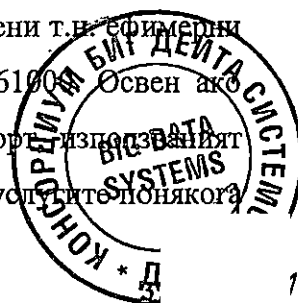


Фиг. 1-11

Посочените JournalNodes не е задължително да се разполагат на отделни сървъри (поради икономически съображения), като могат да поставят в отделни DataNodes сървъри.

ix. Промяна на портове по подразбиране за услуги

В по-ранните версии на Hadoop, на множество услуги бяха прикачени т.н. ефимерни портове – не постоянни назначени портове, в диапазона 32768-61440. Освен ако клиентската програма изрично не поиска конкретен номер на порт, използваният номер на порт е ефимерен номер на порт. Така че при стартиране услугите понякога



може да се стигне до конфликт поради използване на тези портове от друго приложение.

За да се предотврати този потенциален конфликт, в Hadoop 3 примерите обхват са преместени извън този диапазон.

х. Поддържане на конектор на файловата система

Hadoop вече поддържа интегрирането с Microsoft Azure Data Lake и Aliyun Object Storage System. Той може да се използва като алтернативна Hadoop-съвместима файлова система. Първоначално Microsoft Azure Data Lake бе добавено и след това и Aliyun Object Storage System.

xi. Балансьор между DataNodes

Един DataNode управлява няколко диска. По време на нормална операция по запис, данните се разделят равномерно по дисковете и по този начин дисковете се запълват равномерно. Но добавянето или замяната на дискове води до изкривяване на този баланс в рамките на DataNode. Тази ситуация се отнася до изкривяването на DataNode.

Hadoop 3 се справя с тази ситуация чрез новата функционалност за балансиране на данни в рамките на DataNode, която се извиква чрез CLI - diskbalancer на hdfs.



000205

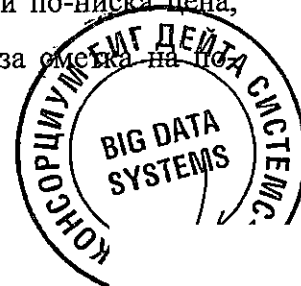
2 Използване в Hadoop на не-RAID архитектури на ди

ИВИ

Hadoop може да съхранява огромно количество данни. Той се изгражда с възможност за преодоляване на проблеми в дисковете. Но независимо от това, в тази масивна система за паралелна обработка най-малко надеждни компоненти са твърдите дискове. В типичен клъстер 100 – 1000 DataNodes, от дисковете работят в тандем, за да свършат работата. Според големите администратори на клъстерите на Hadoop 50% от проблемите на такива клъстерни системи са от дисковете. Затова е много важно да се конфигурират правилно дисковете в клъстера Hadoop по правилния начин, за да се получи максимална работа от тях.

Най-разпространеният стандарт за конфигуриране на множество дискови устройства в един компютър / сървър е RAID архитектурата (Redundant Array of Inexpensive/Independent Disks), която се предлага основно в сървърите. Освен това съществуват конфигурации на дискове с не-RAID архитектура, които се наричат по този начин, за да се различат от RAID архитектурата. Това са:

- JBOD (just a bunch of disks) архитектура, при която архитектура отделните дискови устройства се разглеждат и работят като отделни независими дискове;
- SPAN или BIG архитектура, при която свободно дисково пространство от няколко дискови устройства (с JBOD архитектура) се разглежда като едно логическо устройство. В BIG / SPAN архитектурата могат да се обединят дискови пространства от различни по тип и размер дискове.
- MAID архитектура (произлиза от Massive Array of Idle Drives) оперира със стотици до хиляди дискови устройства, наречени „устройства работещи в почти онлайн режим) - Nearline хранилища (Nearline произлиза от "near" и "online storage"), които работят в приложен режим "пиши веднъж, прочети от време на време" - Write Once, Read Occasionally - WORO). Nearline архитектурата е междинна между online и offline режим на работа на дискова архитектура. При тази Nearline архитектура са създава по-голям обем памет при по-ниска цена, постигнат чрез компромиса по-голямо закъснение на работа за сметка на по-ниска надеждност.



2.1 JBOD архитектура

Това е архитектура с използване на шкаф за поместване на множество дискови устройства. За разлика от RAID архитектурата, JBOD взема набор от дискове и ги превръща в един голям дял. Освен това JBOD не предлага излишък. В JBOD няма толерантност към грешки, няма подобрене в производителността, за разлика от RAID технологията. Освен това, при повреда на един от дисковете в JBOD архитектурата, целият дял, състоящ се от всичките JBOD дискове се оказва недостъпен за работа.

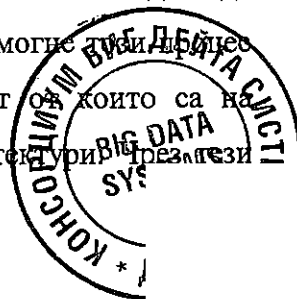
По принцип, целта на JBOD е преди всичко да увеличи капацитета на конфигурация.

Основните предимствата на JBOD над RAID архитектурата са:

- В JBOD устройството няма загуба на данни. Няма никаква загуба на капацитет, тъй като масивът JBOD позволява обединяване на отделните дискове в една голяма единица - дял. Например, при дискове – един 10GB диск и един 40GB диск ще образуват 50 GB обем на JBOD архитектурата, докато при RAID0, 30GB се губят. По този начин JBOD архитектурата е по-евтина пред RAID архитектурата;
- В JBOD устройството по време на повреда има по-лесно възстановяване. В RAID архитектурата, при повреда в един диск, данните на всеки от дисковете се унищожават и трябва да се възстановяват. По този начин възстановяването на JBOD архитектурата е по-бързо от възстановяването на RAID данни. Това е от особено значение при големи дискови устройства например 10 терабайта.

Основните недостатъци на JBOD над RAID архитектурата са:

- JBOD не предлага резервиране за разлика от RAID технологията.
- Няма подобрене в производителността на системата, използвайки JBOD масив за разлика от RAID, който подобрява производителността и предлага по-висока надеждност.
- JBOD няма толерантност към грешки.
- За разлика от RAID възстановяването, възстановяването на JBOD може да бъде малко по-трудно с чисто хардуерни решения. За да се подпомогне възстановяването, съществуват различни категории софтуерни програми, част от които са на разположение в интернет, които да възстановят JBOD архитектура през мрежи.



000207

програми, сравнително лесно и ефективно може да се възстановяват всички файлове, които са загубени.

Днес JBOD става все по-популярен, тъй като дисковете стават все по-големи по обем и съхранението става все по-сложно. Няколко дистрибуции на Hadoop, включително популярната CDH на Cloudera, препоръчват да се използват конфигурации на JBOD при изграждането на кълстери на Hadoop. За по-голяма точност, дисковете на NameNode е добре да се конфигурират в RAID 1, 5 или 10, докато дисковете на DataNode е добре да се конфигурират в RAID1 за операционната система, а за данните в JBOD.

Хардуерните производители вече предлагат дискове с капацитет 8 до 12TB, които могат да предложат стотици TB за съхранение в сървър. При избор на дискова конфигурация трябва да се има предвид, че количеството дискове в сървър (плътността на данните), съответно скоростта за обмен на данни в сървър е много скъп параметър и е много важен компонент за оразмеряване на Hadoop кълстер, като решенията са между сървъри за данни (DataNodes) с висока и с ниска скорост за обмен (плътност на данните в сървър).

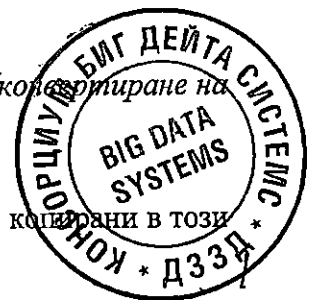
За яснота ще разгледаме практически пример от реална конфигурация:

- Да приемем, че шасито на сървъра позволява 12 дискови шпиндела (до 12 диска могат да се включат към този сървър);
- Да приемем, че за Операционната система се използват 2x1TB дискове в RAID1;
- Да приемем, че за данни се използват 10x2TB дискове в JBOD, общо 20TB данни;
- Да приемем, че скоростта на обмен на данни на един диск е 50 MB/s;

В случай на повреда на един сървър, можем да очакваме трафик от 4.8 Gb/s, т.е. Hadoop архитектурата трябва да допуска мрежов трафик към и от сървъра да бъде не по-малко от 4.8 Gb/s.

$$(10 \times 50 \text{ MB} / \text{s} \times 0.001 \text{ (конвертиране на MB в GB)}) = 0.5 \text{ GB/s} \times 8 \text{ (конвертиране на GByte в Gbit)} = 4.8 \text{ Gb/s}$$

Да приемем, че 16 TB данни трябва да бъдат записани или повторно кодирани в този сървър.



$$(16 \text{ TB} \times 1000 \text{ (конвертиране на TB в GB)}) = 16000 \text{ GB} \text{ (конвертиране на GB Gb)} = 96000 \text{ Gb}$$

Времето, необходимо за записване или за повторното копиране на загубени данни на тези 16TB е 5,55 часа.

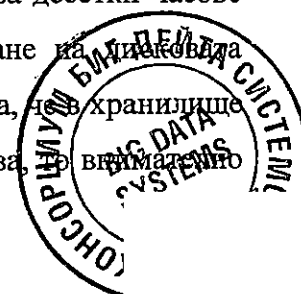
$$(96000 \text{ Gb} / 4,8 \text{ Gb/s} = 20000 \text{ секунди} / 60 \text{ секунди в минута} = 333,33 \text{ минути} = 5,55 \text{ часа.})$$

По същия начин, времето, необходимо за записване или за повторното копиране на загубени данни на 36TB е 12 часа.

Съвременните контролери на дискове допускат по-голяма скорост за трансфер на данни от изчислените 4.8 Gb/s, като съществуващите стандарти допускат 6 Gb/s или 12 Gb/s. Тогава, при 6 Gb/s скорост на трансфер и наличие поне на 16 диска, 36 TB биха се трансферирали за 9,6 часа; при 12 Gb/s скорост на трансфер и наличие поне на 32 диска, 36 TB биха се трансферирали за 4,8 часа.

Тези примерни изчисления показват зависимостта на времето за записване или на повторно копиране на загубени данни от скоростта на работа на дискове и от броя дискове в сървъра. Същевременно тези изчисления изглеждат теоретични и не практически насочени, защото при записване на 16TB или на 36 TB се използват едни и същи дискове, без да е възможно тези данни да се разпределят на 20-30-40 диска в сървъра. Тогава, ако за 16 TB при 12 диска се изисква време за трансфер 5,5 часа, за 36 TB биха се използвали 12 часа, стига капацитета на 12-диска да не е 2TB, а поне 4 TB. Към тези теоретични изчисления трябва да се включат и съображенията, че тази теоретична скорост за трансфер от 6 Gb/s или 12 Gb/s не са реални, защото в това време същият сървър може би изпълнява и други операции, както и върху тези дискове може да се изпълняват и други трансфери.

От тези теоретични изчисления може да се стигне до извода, че еднократна операция за обмен на десетки TB данни би отнела десетки часове. При оперативна обработка на данни не е често срещана ситуация на еднократна транзакция с такъв обем. Следователно, съображенията за обработка на десетки TB данни за десетки часове не е свързана с оперативна обработка на данни, а с поддържане на данни архитектура при повреда на дискове. Изхождайки от съображенията, че в хранилище за големи данни се използват стотици и хиляди дискови устройства,



трябва да се оценяват съображенията за поддържане на дисковата архитектура, които са сериозен фактор за цялостното опериране на такива хранилища.

Съвременните дискови устройства имащи капацитет 12TB/14TB/16TB имат скорост за обмен около 250MB/s и SATAинтерфейс от 6Gb/s. Това означава, че този интерфейс може да поддържа само няколко диска в паралел едновременно обменящи данни на пълен капацитет, т.е. при дискове в JBOD с голям брой устройства, SATA интерфейса е ограничаващ за обмена на данни от такава JBOD архитектура.

2.2 SPAN/BIG архитектура

Когато се изисква да се създаде дял (том), но нямат достатъчно неразпределено място на един диск, може да се създаде обем с желания размер, като се комбинират области от неразпределено пространство от няколко отделни диска. Зоните на неразпределеното пространство могат да бъдат с различен размер. Този вид диск се нарича „обхващащ“ диск (диск със SPAN/BIG архитектура на дял). Обхващащият диск позволява създаването на по-голям дисков обем от колкото е на физическите дискове. Чрез комбиниране на множество неразпределени пространства на физически дискове, обозначени с друго име на устройство (например J:) се създава възможност за работа с големи по обем файлове, по-големи от размер на физическо дисково устройство в компютъра.

При създаване на обхващащ диск, може да се използват части от свободно пространство на множество физически дискове. Но след като се създаде обхващащия диск, от него не може да се премахнат части от някои участващи физически дискове – може да се изтрие изцяло обхващащия диск, за да се използват части от някои участващи дискови пространства.

При UNIX-подобните операционни системи, терминът „устройство“ се отнася за физическо устройство за съхранение на данни, като например твърд диск, полупроводниково дисково устройство, подвижен USB флаш памет и т.н. В тези операционни системи устройствата са представени от специални обекти на файловата система, наречени device nodes, които се виждат в директория /dev. Устройствата за съхранение на данни обозначени в /dev са според типа на устройството, последвано от буква, което означава реда, в който те са филтрирани.

000210



от системата. В Linux преди ядрото 2.6.20 префикс hd означаваше IDE устр
така че например файловете на устройствата / dev / hda, / dev / hdb и / de
съответно съответстваха на първо, второ и трето IDE устройство. Префиксът sd
първоначално е бил използван за SCSI устройства, но сега се използва за всички
PATA и SATA устройства, включително устройства на IDE шина. Ако има повече от
26 такива устройства в системата, устройствата от 27-о нататък са означени с / dev /
sdAa, / dev / sdAb и т.н. Устройството за физическо съхранение на данни може да
бъде разделено на няколко логически единици, известни като дялове. Всеки дял ще
се покаже в / dev като отделен device node. Номер след буквата на устройството
означава номера на дяла, така че например / dev / sda1 и / dev / sda2 се отнасят до
първия и втория дял на първото устройство. Други Unix-подобни системи могат да
се обръщат към дискове и дялове от тях по други начини. Например FreeBSD
използва / dev / adaX (където X е една или повече цифри), за да се позове на дискове
и / dev / adaXpY (където X и Y са едновременно или повече цифри), за да се отнасят
до дялове на дискове. Терминът „дял (volume)” в Linux е свързан с Логическия дял
(LVM – Logical Volume Manager), който може да се използва за управление на
устройства за съхранение на данни. Физическият обем е устройство за съхранение
или дял. Физическият дял е цяло устройство или част от него, докато логическият
дял - LVM, е логическо устройство за съхранение, обхващащо множество физически
дялове, включително и обхващащ диск.



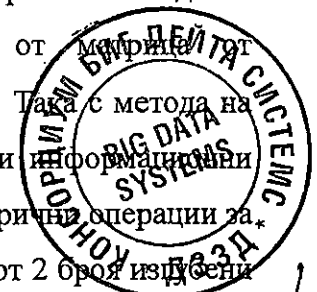
000111

3 Технологии влияещи на Hadoop клъстерните архитектури за изграждане на голям капацитет

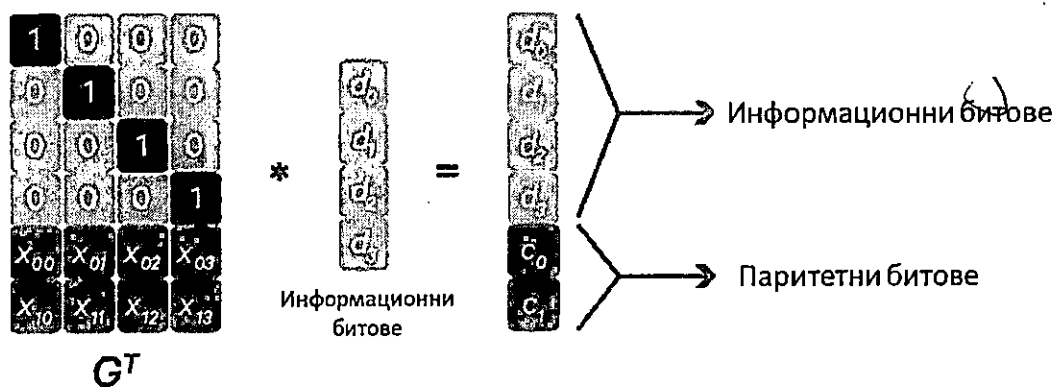
Представените хардуерни и софтуерни технологии по-долу имат за цел да предложат различни подходи за създаване на Hadoop клъстерни архитектури с голям капацитет от порядъка на стотици и хиляди петабайта – PB (1 екзабайт (EB) = 1024 петабайта (PB) = 1048576 терабайта (TB)).

3.1 Същност на Кодиране срещу изтриване (КСИ) за увеличаване ефективното използвано дисково пространство

Методът за Кодиране срещу изтриване (КСИ), наречен в оригинал Erasure Coding – EC или Erasure Encoding – EC, отдавна се използва в локалните системи за съхранение, особено под формата на RAID-5 и RAID-6. RAID-5 обикновено използва XOR кодиране, тъй като то трябва само да толерира единичен диск, докато RAID-6 използва Reed-Solomon с две паритетни диска, за да толерира до две грешки. Най-простата форма на КСИ е прилагането на функцията на „Изключващо ИЛИ – XOR“ между съвкупност от информационни битове, създавайки допълнителен бит за паритет, при което ако един от информационните битове се изгуби, то използвайки паритетния бит, съдържанието на изгубения бит може да се възстанови (регенерира). При загуба на повече от един бит, паритетния бит създаден на база на „Изключващо ИЛИ“ не може да се използва. За възстановяване на два изгубени бита може да се прилага метода на кодиране Reed-Solomon (RS), създавайки 2 паритетни бита от информацията от информационни битове – един по хоризонтала и един по вертикала. Така с метода на кодиране Reed-Solomon се възстановят (регенерират) до 2 изгубени информационни бита. Фактически КСИ-RS използва усъвършенствани линейни алгебрични операции за генериране на множество паритетни битове, осигурявайки по-голям от 2 броя изгубени

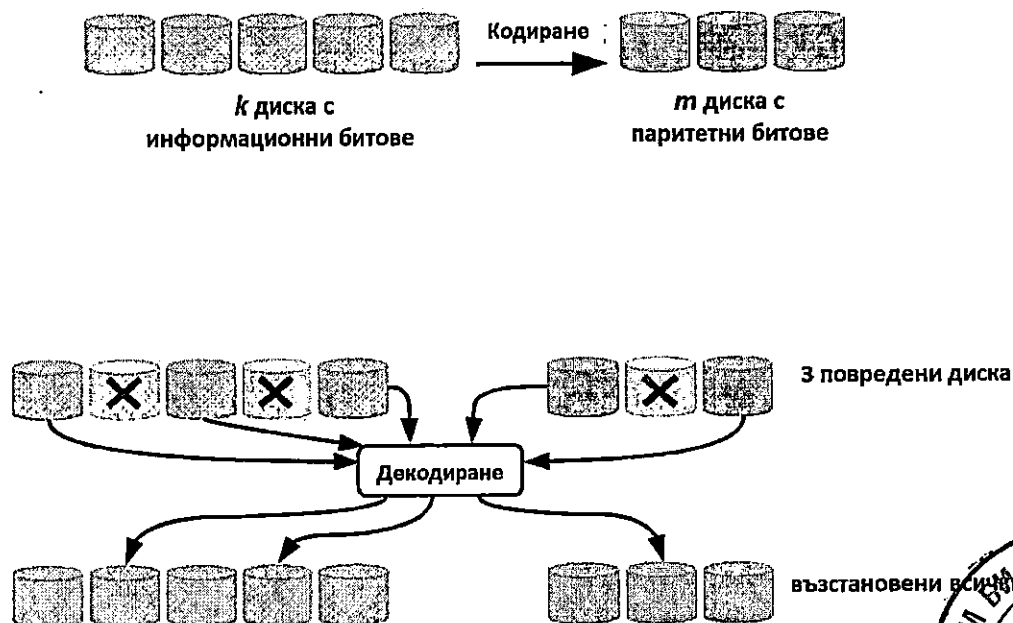


информационни битове. Това го прави по-всеобхватен метод за съхранение на данн при производствени системи. КСИ-RS е конфигурируем с два параметъра „ k “ и „ m “, като RS(k,m) работи чрез умножаване на вектор от информационните данни „ k “ с генераторна матрица (G^T) за генериране на разширен с „ k “ информационни бита данни и „ m “ паритетни бита – фигура 3-1.



Фиг.3-1

На фигура 3-2 е показана възможна реализация на RS(k,m) в дискова конфигурация, където процесът „кодиране“ създава дисковете с паритетна информация, а процесът „декодиране“ възстановява потенциално повредени дискове, без да се губи полезна информация.



Фиг. 3-2

000213

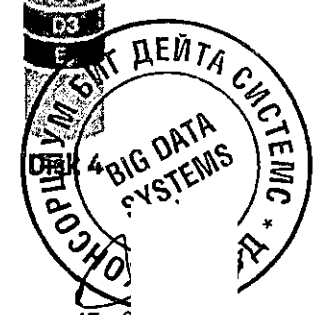
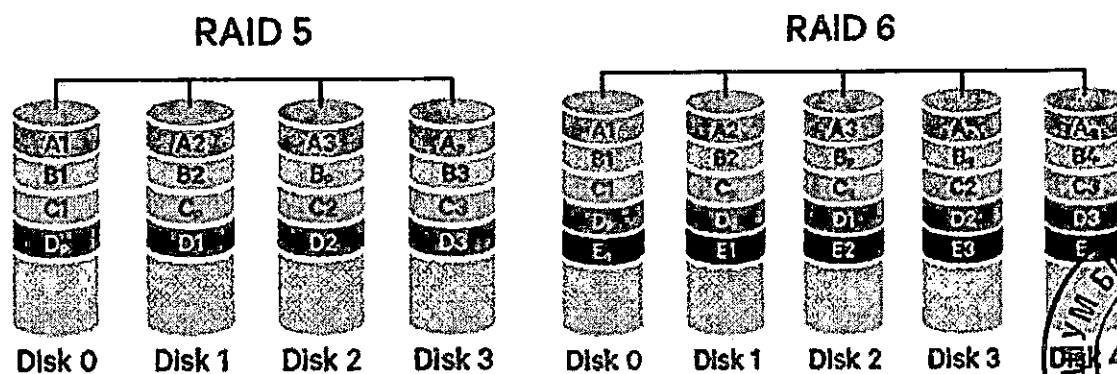


В таблица 3-1 са показани сравнения между различни форми на Репликиране и Кодирание срещу изтриване.

Таблица 3-1

Методи на кодиране	Стабилност срещу изтриване/повреда	Ефективност на използване на дисковото пространство
Без репликиране	Допустими 0 броя изтриване	100% използване
Тройно репликиране (основни данни + 2 реплики) – както в Hadoop 2	Допустими 2 броя изтриване	33% използване
КСИ - XOR	Допустими 1 брой изтриване	86% използване
КСИ – RS (6,3)	Допустими 2 броя изтриване	67% използване
КСИ – RS (10,4)	Допустими 3 броя изтриване	71% използване

КСИ се използва отдавна в компютърните системи при дисковите архитектури RAID5 и RAID6, като RAID5 работи по метода КСИ - XOR с допустима за корекция 1 грешка, а RAID6 работи по метода КСИ – RS с допустими за корекция 2 броя грешки – фигура 3-3.



000214

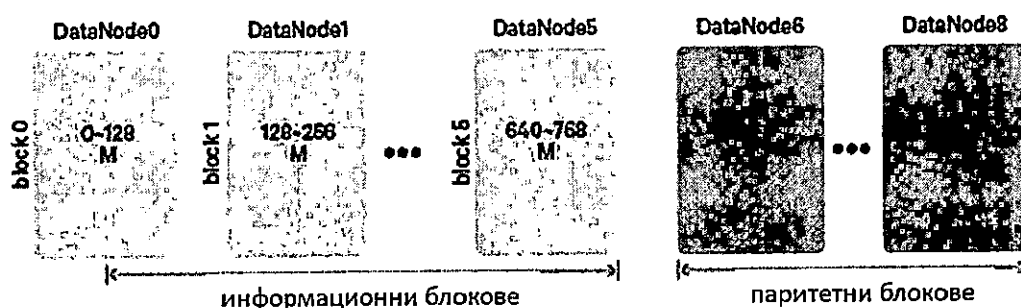
Фиг.3-3

За да управлява потенциално много големи файлове, каквито са в Hadoop клъстера, тези файлове обикновено се разделят файловете на логически последователности от байтове с фиксиран размер, наречени логически блокове. Тези логически блокове след това се преобразуват в блокове за дисково записване, които отразяват физическото оформление на данните в клъстера.

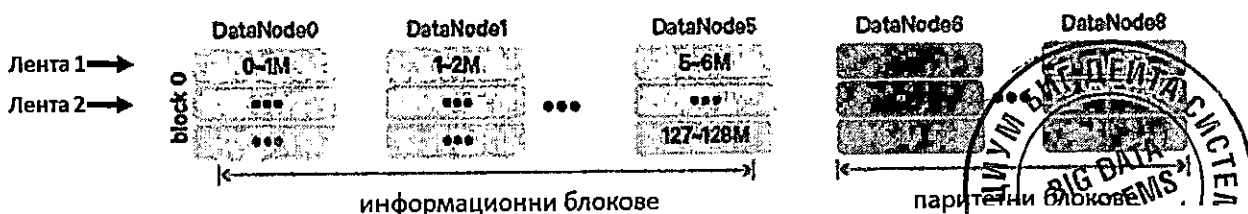
3.1.1 Кодирани и Проектирани организацията на информацията при DataNode сървъри

Информацията в логическите блокове може да се разполага на дисковете чрез 2 метода – чрез Съседно разполагане на блокове и чрез Разполагане на ленти – фигура 3-4. Най-простото съответствие между логическите блокове и дисковите блокове е едно към едно. Файл със съседно разполагане на блоковете съответства на линейно последователно разполагане на логическите блокове по дисковото пространство. За разлика от това, оформянето на блока с разполагане на ленти разделя логическия блок на много по-малки запаметяващи единици - ленти, като записва много ленти в един блок за съхранение. Четенето на файл организиран чрез разполагане на ленти изисква заявка за прочитане на набор от логически блокове, след което се четат ленти от тези блокове.

Съседно разполагане - contiguous



Разполагане на ленти - Stripping

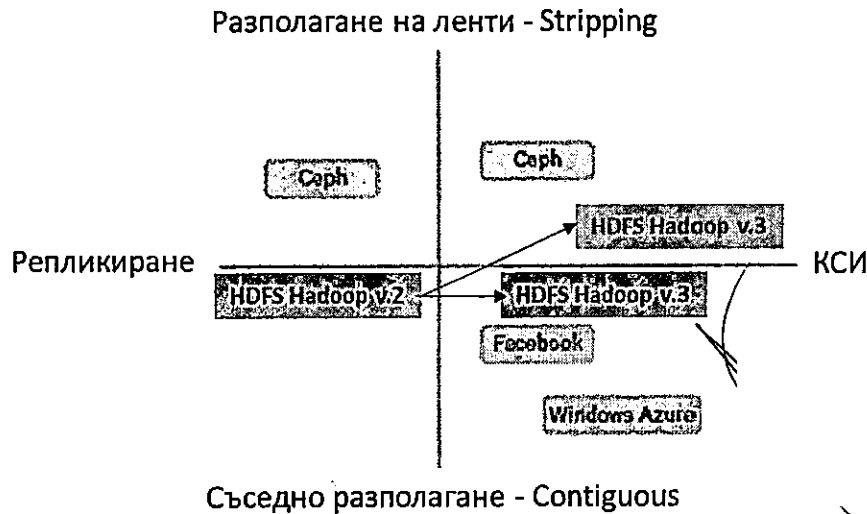


Фиг. 3-4

По принцип блоковото оформление (Съседно срещу Разполагане по ленти) и форма на съкращаване (Репликация срещу КСИ) са две ортогонални оси на четири квадранта възможни комбинации. Анализът на тези възможности за кодирани, показани на фигура 3-5, показват, че и четирите са приложими при съхранение на големи файлове. Някои системи, включително Serp и QFS, поддържат конфигуриране на излишеството на ниво директория или на ниво цялостен файл. Serp е безплатен софтуер за управление на данни в разпределен компютърен клъстер, като осигурява интерфейси за опериране на ниво обект, блок и файл. Serp поддържа автоматично изправяне при единична грешка и управление на големи обеми от данни на ниво exabyte. QFS (Quick File System) е файлова система на фирмата Oracle. Тя е тясно интегрирана с дисковата архитектура SAM (Storage и Archive Manager) и по този начин често се нарича SAM-QFS. SAM осигурява функционалността на йерархичен мениджър за съхранение. QFS поддържа някои възможности за управление на обема, което позволява много дискове да бъдат групирани заедно във файлова система. Метаданните на файловата система могат да се съхраняват на отделен набор от дискове, което е полезно за стрийминг на приложения, където дългото търсене на диск не може да бъде толерирано. Файловата система SAM-QFS може да има сравнително малък (от гигабайта в терабайсет) "дискон кеш", поддържан от петабайтово лентово устройство или оптични носители. При този вид файлова организация файловете се копират в архивното хранилище във фонов режим и прозрачно се възстановяват върху диск при достъп до данните. SAM-QFS поддържа до четири архивни копия, всеки от които може да бъде на диск, лента, оптичен носител или може да бъде съхраняван на отдалечен сайт, който изпълнява и SAM-QFS.



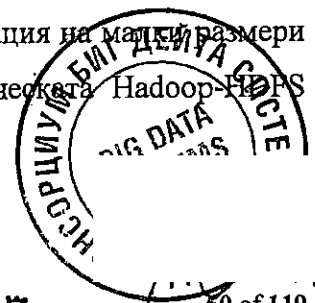
000216



фиг.3-5

Трябва отново да се отбележи, че Кодирането срещу изтриването (КСИ) има предимство пред репликацията по отношение на ефективността на използване на дисковото пространство, но пък изисква по-комплексно възстановяване при грешка.

Ако се изхожда от обстоятелството, че блокът в Hadoop е голям по размер (128 MB), то методът с Разполагане на ленти може да осигури по-добра входно-изходна пропускателна способност от метода за Съседното разполагане, тъй като може да използва паралелно няколко дискови шпиндела. При четене от няколко дискови устройства разположени в различни сървъри, методът за Разполагане на ленти поставя изискване за по-мощна мрежова среда – от порядъка на 10-ки Gbps. Същевременно методът на Разполагане на ленти на един блок в различни физически сървъра противоречи на традиционната парадигма на MapReduce за локализиране на данни от един блок в един сървър, като това може да бъде преодоляна чрез отдаване на специален фокус на приложението, където приложението трябва да чете и записва своите данни отчитайки размерите на отделните ленти. По този начин ще се осигури прилагането на философията на Hadoop и специално на MapReduce, с предимствата на метода на Разполагане на ленти. Представеният подход – прилагане на метода за създаване на данни върху диск - Разполагане на ленти, в съчетание с приложение, отчитащо размерите на отделните ленти, е средство за модификация на мащабни размери данни, което работи сравнително по-бързо от колкото класическата Hadoop HDFS версия 3 организация на работа.



За HDFS-КСИ основният въпрос е да се определи къде е най-подходящо да се разположи даден блок от данни. Методът за разполагане на блоковете от данни „Съседно разполагане“ в Hadoop версия 3 е лесно приложим, тъй като пътят за достъп до такива данни е много подобен на този при репликиране, съществуващ в Hadoop версия 2. Обаче, „Съседното разполагане“ е подходящо само когато файловете са достатъчно големи. Например, с RS (10,4) лента с един 128MB блок данни изисква създаване на четири 128MB паритетни блока, което е 400% използване на по-голямо дисково пространство и което е по-голямо разхищение отколкото тройното репликиране без КСИ. Съседното разполагане е подходящо предимно за офлайн или за пакетна работа, тъй като в противен случай клиентската програма ще трябва да буферира гигабайти информационни данни, за да изчисли паритетните данни.

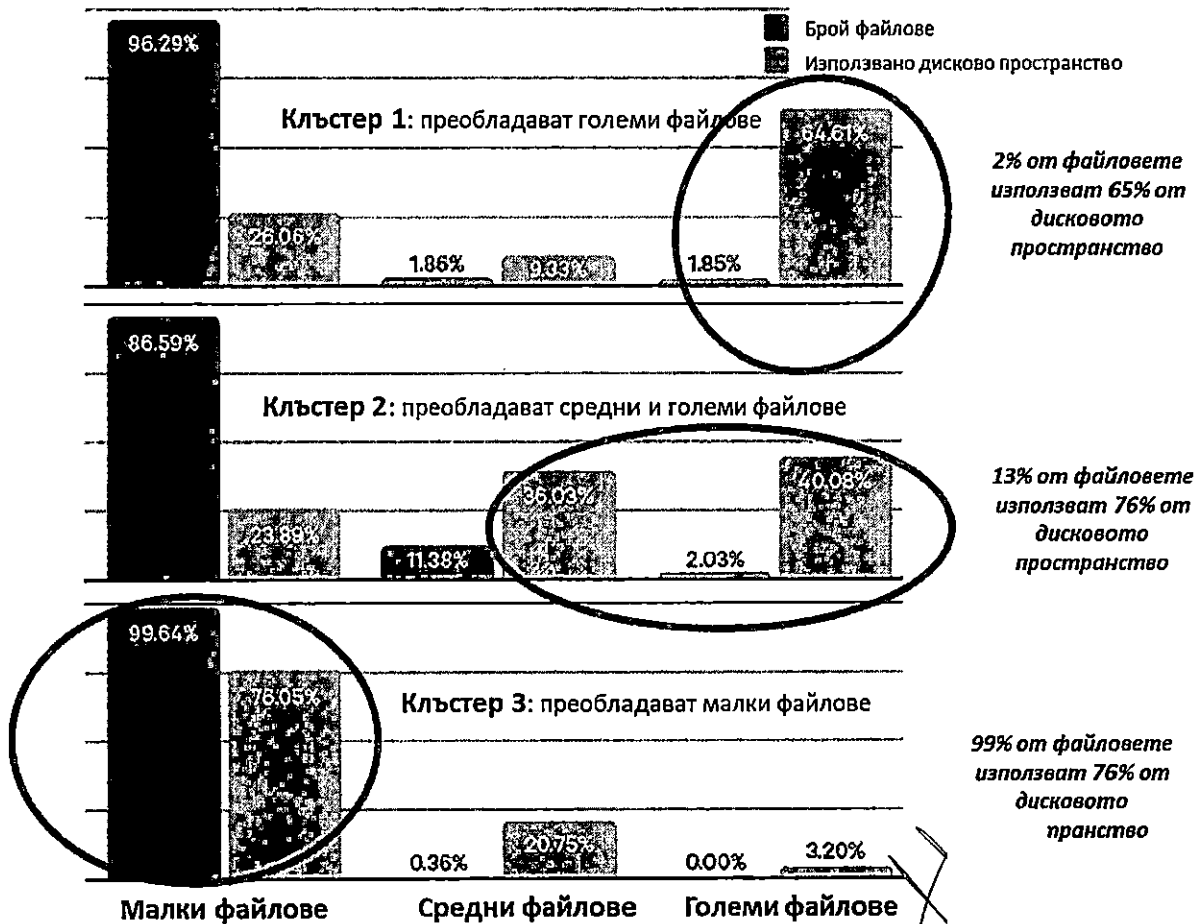
От друга страна, КСИ-RS с Разполагане на ленти може да реализира спестявания при съхранение както на малки, така и на големи файлове, защото размерът на лентите е много по-малък (обикновено 64KB или 1MB). Този общ размер на по-малките групи също дава възможност на онлайн опериране, където приложенията директно записват данни за изтриване, тъй като са необходими са габайта буфери за изчисляване на информацията за паритета.

Въз основа на този анализ, размерът на файла е ключовият определящ фактор за избор на метод на КСИ - „Съседното разполагане“ или „Разполагане на ленти“. Ако използването на клъстера е доминирано от големи файлове - това означава шест или повече блока от по 128 MB всеки, достатъчно за една пълна група от ЕК по схема RS (6,3) - тогава „Съседното разполагане“ е подходящо, тъй като можем да избегнем проблемите при внедряването на комбиниране на блокове от множество малки файлове в една група. Въпреки това, ако капацитетът на клъстерите е доминиран от малки файлове, тогава „Разполагане на ленти“ е по-естественият избор, за да се постигнат по-добро използване на дисковото пространство, са по-бързо опериране и по-ефективно администриране.

Емпирическо изследване проведено от Zhe Zhang, Andrew Wang, Kai Zheng, Uma Maheswara G., и Vinayakumar B - [8], проведено върху големи клиенти на Hadoop-Cloudera системи – фигура 3-6, показва, че между 24-76% от файловете в Hadoop системите са с малки файлове (в обхвата на един блок от 128 MB), което води до извода



за предпочитане е използването на кодиране по метода КСИ - „Разполаган ленти“.



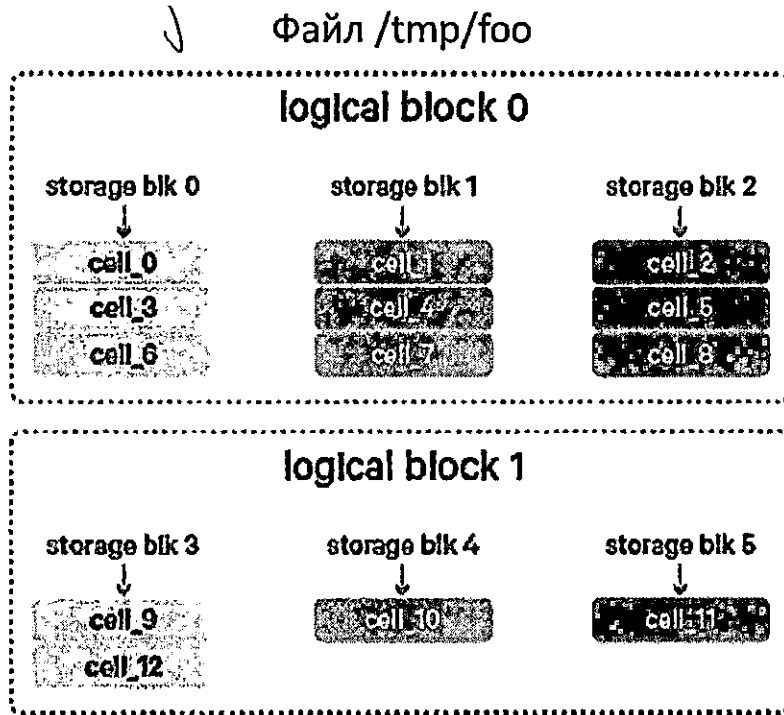
Фиг. 3-6

3.1.2 Кодиране и Проектиране организацията на информацията при NameNode сървъри

Предположението за „Съседното разполагане“ е широко и дълбоко вградено във вътрешната логика на HDFS. За да се поддържа „Разполагане на ленти“, концепцията за логически блок (logical block) трябваше да бъде отделена от тази на блока за съхранение (storage block). Първият представлява логически диапазон от байтове в един файл, а вторият представлява основната единица от данни, съхранявани в DataNode. Фигура 3-7 показва концепциите за логически блокове и блокове за съхранение. В този пример файлът / tmp / foo е логически разделен на 13 лентови клетки наречени „cell“ - („cell_0“ до „cell_12“). Логическият блок 0 обхваща от „cell_0“ до „cell_12“. Логическият блок 1 обхваща от „cell_9“ до „cell_12“. Клетките „cell_0“, „cell_3“, „cell_6“ образуват блок за съхранение 0 (storage blk 0), който ще съдържа данни като



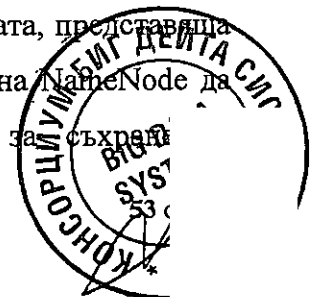
единица данни в NameNode. За краткост в тази фигура не се включват паритетни блокове/клетки



фиг. 3-7

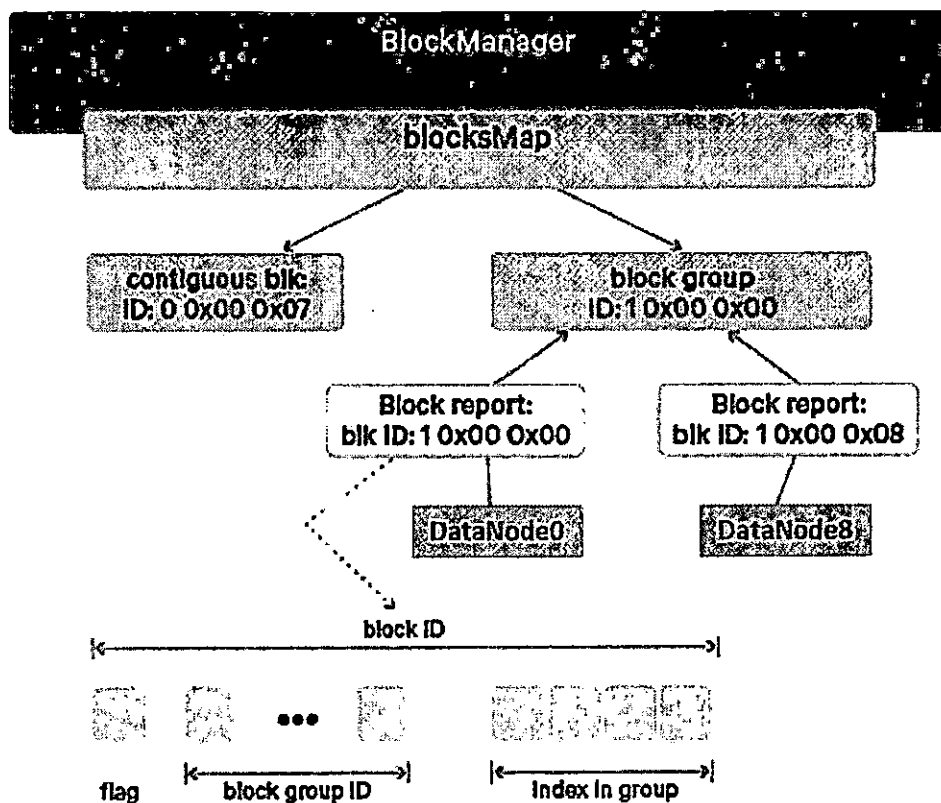
Естественният механизъм в HDFS NameNode е да се следи всеки блок за съхранение в неговата организация на блоковете (blocksMap), която организация обхваща съответствието от блоков идентификатор (blk ID) към съответния логически блок от една страна и от друга - от логически блок към блокове за съхранение. Това обаче означава, че малките файлове ще натрупат значителни разходи за памет в NameNode, тъй като рабиването на клетки води до много повече блокове за съхранение, отколкото при прилагане на стандартната за Hadoop репликация.

За да намалим тези разходи, се въведе нов протокол за именуване на блокове. Понастоящем HDFS разпределя блоковите идентификатори последователно въз основа на времето за създаване на блокове. Вместо това, в този нов протокол се разделя всеки идентификатор на блок на 2 - 3 секции, както е показано на фигура 3-8. Всеки блоков идентификатор започва с флаг, показващо разположението му (съседни = 0, ивица = 1). За ивичестите блокове останалата част от идентификационния номер се състои от две части: средната секция с идентификатор на логическия блок и опашката, представляваща индекса на блока за съхранение в логическия блок. Това позволява на NameNode да управлява логически блок като обобщение на блоковете му за съхранение



Идентификаторите на блок
 блок, като маскират индекс
 блок данни.

ение могат да бъдат прикачени към логическия
 зискава, когато NameNode обработва отчети за

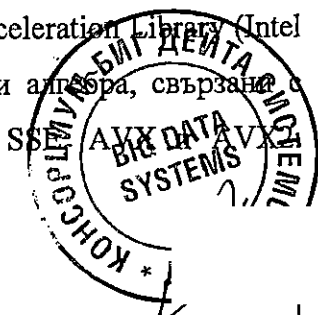


Фиг.3-8

Резултати от симулирането на работа на паметта на NameNode с прилагане на КСИ на данните представени от трите HDFS клъстера във Фигура 4-6, показват, че без прилагане на посочения по-горе нов протокол за йерархично именуване на блоковете, методът за „Разполагане на ленти“ в NameNode ще увеличи размера на блоковете с 250% - 440% - [8]. С прилагане на този протокол, увеличаването на дисковото пространство на имената ще бъде с 21% - 76% повече.

3.1.3 Отчитане влиянието на кодиране-декодиране при КСИ

Кодирането и декодирането на данни е много интензивно за процесора и може да бъде определяща функция на процесора на сървър (DataNode). За да се облекчи този процес в HDFS-EC може да се използва библиотеката Intel Intelligent Acceleration Library (Intel ISA-L) с отворен код, която ускорява изчисленията за линейни алгебра, свързани с КСИ, като използва съвременни хардуерни инструкции като SSE.



Библиотеката ISA-L поддържа всички
Unix подобните и Windows.

ционни системи, включително

Проведено бе изследване ([8]) за HDFS-КСИ използване на Reed-Solomon в две версии: едната на базата на ISA-L и другата в чиста Java програмна реализация. В допълнение за сравнение бе включена и реализация използвана от Facebook чрез хардуерно изпълнение на корекцията от грешка HDFS-RAID. Приложена е схемата на КСИ - Reed-Solomon (6,3), а за хардуерната реализация – съответстващата RAID. Резултатите от изпълнението на ISA-L надхвърля изпълнението на HDFS-EC Java с повече от 4 пъти и на HDFS-RAID с около 20 пъти. Въз основа на тези резултатите силно се препоръчва използване на ISA-L за всички DataNode производствени мощности.

Също така беше изследвана ефективността на входно-изходните операции в HDFS от край до край с посочените две програмни реализации и стандартното тройно репликиране в HDFS. Тестовите са извършени на клъстер с 11 възела (1 брой NameNode, 9 броя DataNodes, и 1 клиентски сървър), свързани с 10 GigE мрежа. Изследването е направено при запис от страна на клиент на 12GB файл в Hadoop, и при четене на 12GB файл от Hadoop. Това изследване показва, че, внедряването на ISA-L е много по-бързо от чисто Java програмната реализация, поради отличната ефективност на процесора.

От представените до момента изследвания показват, че КСИ може да намали разходите за съхранение на HDFS с приблизително 50% в сравнение с репликацията, като същевременно се запазят същите гаранции за гарантиране изправянето на грешка. Това води до значително намаляване на разходите за съхранение на хардуер.

3.2 Кодиране срещу изтриване (КСИ) в Hadoop 3

Фирмите Intel и Cloudera осигуриха възможност за прилагане на библиотеката Intel ISA-L за изпълнение на метода КСИ в Hadoop 3, без външни зависимости, което прави Hadoop 3 самостоятелен и независим от допълнителна поддръжка. Този осигурява работа на КСИ със съществуващите функции на HDFS като кеширане, моментна снимка, криптиране, висока достъпност и т.н. По този начин в Hadoop 3 се предоставят различни схеми за кодиране на EC, реализации също така и политики за различни сценарии за прилагане.



000222

Intel ISA-L (Intel Intelligent Storage Acceleration Library) е набор от оптимизирани функции на ниско ниво, използвани главно в съхранението на данни върху диск. Той включва метода КСИ с използване на процесорите с Intel архитектура - AVX и AVX2.

HDFS в Hadoop 3 осигурява натурална поддръжка на КСИ за по-ефективно съхраняване на данните. Всяка отделна директория може да бъде конфигурирана с политика на КСИ с команда "hdfs erasurecode -setPolicy". Когато се създаде файл, той ще наследи правилата на КСИ от най-близката директория на предците си, за да определи как се съхраняват блоковете му. В сравнение с тройната репликация от Hadoop 2, стандартната политика на КСИ в Hadoop 3 спестява 50% от дисковото пространство за съхранение, като в същото време толерира повече откази от съхранение.

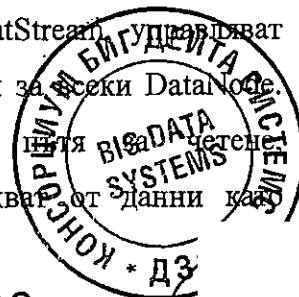
За да поддържат малки файлове, HDFS-КСИ на Hadoop 3 съхранява блокове в `_striped_` оформление (разполагане на ленти), където логическият файлов блок е разделен на малки единици (по подразбиране 64KB) и разпределен в съвкупността от DataNode съвъри. Това позволява паралелен вход/изход, но също така намалява локалната зависимост на данните. Поради това конфигурирането на Hadoop 3 клъстера и използването на входно/изходните операции трябва да се обмисли предварително.

В съвременните HDFS клъстери, малките файлове са преобладаващ брой от общото количество файлове в Hadoop. За по-добра поддръжка на малки файлове, първоначално в Hadoop 3 е приложено използването на КСИ с „Разполагане на ленти“, което осигурява ефективна работа на малки файлове, в следващите версии на Hadoop 3 ще се предложи по-ефективна работа на средни и големи файлове чрез КСИ „Съседно разполагане“.

За да се поддържа HDFS-КСИ в Hadoop 3 се въведени архитектурни разширения в трите типа компютърни архитектури – Клиентски компютър, DataNode и NameNode:

- *Разширение за Клиентски компютър* – Пътеките за четене и запис в клиентския компютър са подобрени, за да работят паралелно върху няколко вътрешни блока в блокова група. На пътя за запис, `DFSStripedOutputStream` предоставяват „Разполагане на ленти“ (чрез streamers на данни), по един за всеки DataNode. Тези streamers работят предимно асинхронно. На пътя за четене `DFSStripedInputStream` превежда заявения логически обхват от данни към диапазони във вътрешни блокове, съхранявани в DataNodes.

000223



5

- *Разширения за DataNode сървър* - DataNode изпълнява допълнителна задача ErasureCodingWorker (ECWorker) за възстановяване на фона на неуспешно кодирани КСИ блокове. Неуспешни блокове на КСИ се откриват от NameNode, който след това избира DataNode, за да извърши възстановяването.
- *Разширение за NameNode сървър* – Файловете работещи чрез „Разполагане на ленти“ са логически съставени от блокови групи, всяка от които съдържа определен брой вътрешни блокове. За да се намали потреблението на паметта на NameNode сървъра от тези допълнителни блокове, е въведен нов протокол за именуване на йерархични блокове. Идентификационният номер на блокова група може да бъде извлечен от идентификационния номер на всеки от нейните вътрешни блокове. Това позволява управлението да се извършва на нивото на блоковата група, а не на блока.

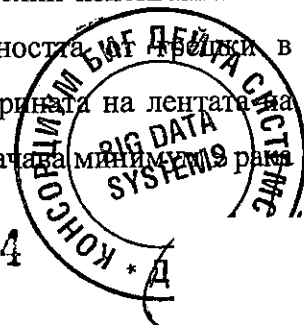
За да се осигури използване на разнородните варианти на КСИ, файловете и директорите в HDFS клъстера могат да имат различна политики на КСИ. Информация за това как да се кодира / декодира файл е капсулирана в клас ErasureCodingPolicy. Всяко правило се определя от следните две части:

- КСИ Schema: Това включва броя на блоковете за данни и паритет в група от КСИ (напр. 6 + 3), както и алгоритъм (напр. Reed-Solomon) – RS(6,3).
- Размерът на лентовите клетки: Това определя грануларността на лентите използвани при метода „Разполагане на ленти“, включително размера на клетките.

В момента се поддържат четири политики: RS-DEFAULT-3-2-64k, RS-DEFAULT-6-3-64k, RS-DEFAULT-10-4-64k и RS-LEGACY-6-3-64k. Всички са с размер на лентовата клетка по подразбиране 64KB. Правилата по подразбиране за системата са RS-DEFAULT-6-3-64k, които използват стандартната схема RS_6_3_SCHEMA с размер на клетката 64KB.

Файловете за КСИ се разпростират и в ракове за толерантност към грешка. Това означава, че при четене и писане с прилагане на метода „Разполагане на ленти“, се използват операции между няколко рака. Това поставя допълнителни изисквания към мрежовата инфраструктура на Hadoop клъстера. За толерантността на клъстера е важно да има най-малко толкова рака, колкото е ширината на лентата на КСИ. За политиката на КСИ по подразбиране на RS (6,3) това означава минимум четири рака.

000224



са необходими, а в идеалния случай 10 или 11, за да се справят с планираните и н планирани прекъсвания поради технически грешки. За групи с по-малко ракове отколкото ширината на лентата на КСИ, HDFS-КСИ не може да поддържа толерантност към грешки на ниво ракове, но ще поддържа толерантност към грешки на ниво сървър.

По подразбиране всички вградени правила за КСИ са деактивирани, с изключение на дефинираната в `dfs.namenode.ec.system.default.policy`, която е активирана по подразбиране. Администраторът на клъстерите може да активира набор от правила чрез командата `hdfs ec [-enablePolicy-policy <policyName>]`, базирана на размера на клъстера и желаните свойства за понасяне на грешки. Например, за клъстер с 9 рака политика като RS-10-4-1024k няма да запази толерантността към грешки на ниво ракове и RS-6-3-1024k или RS-3-2-1024k може да е по-подходяща. Ако администраторът се интересува само от толерантността към грешки на ниво сървър, RS-10-4-1024k би било подходящо, стига да има поне 14 DataNodes в клъстера

Политиката на КСИ по подразбиране за системата може да бъде указана чрез конфигурацията `"dfs.namenode.ec.system.default.policy"`. При тази конфигурация правилата на КСИ по подразбиране ще се използват, когато в командата `"-setPolicy"` не се предава име на правилата като аргумент. По подразбиране `"dfs.namenode.ec.system.default.policy"` е `"RS-6-3-1024k"`.

КСИ поставя допълнителни изисквания към Hadoop клъстера по отношение на процесора и мрежата, защото КСИ консумира допълнителен CPU както за HDFS клиентите, така и за DataNodes.

Прилаганите кодиращи / декодиращи програми за Reed-Solomon (RS) и XOR могат да бъдат конфигурирани със следните конфигурационни ключове за клиентски в DataNode сървъри: `io.erasurecode.codec.rs.rawcoders` за подразбиращите се RS програми, `io.erasurecode.codec.rs-legacy.rawcoders` за онаследени RS програми и `io.erasurecode.codec.xor.rawcoders` за XOR програми. Може да се създаде също така отделни кодиращи / декодиращи програми с конфигурационен ключ `io.erasurecode.codec.self-defined-codec.rawcoders`. Стойностите за тези ключове



000225

списъци с имена на кодиращи / декодиращи програми с механизъм за връщане назад. Тези кодиращи / декодиращи програми са заредени в реда, определен от конфигурационните стойности, докато кодиращи / декодиращи програми бъдат заредени успешно. Конфигурацията на RS и XOR кодиращи / декодиращи програми изискват реализации в чиста Java. За кодиращи се RS и XOR кодиращи / декодиращи програми може да се използва библиотеката Intel ISA-L за подобряване на производителността.

В Hadoop версия 3 е вградено по подразбиране използването на библиотеката Intel ISA-L за KCSI-RS. За да се активира използването на Intel ISA-L, се прилагат следните три стъпки:

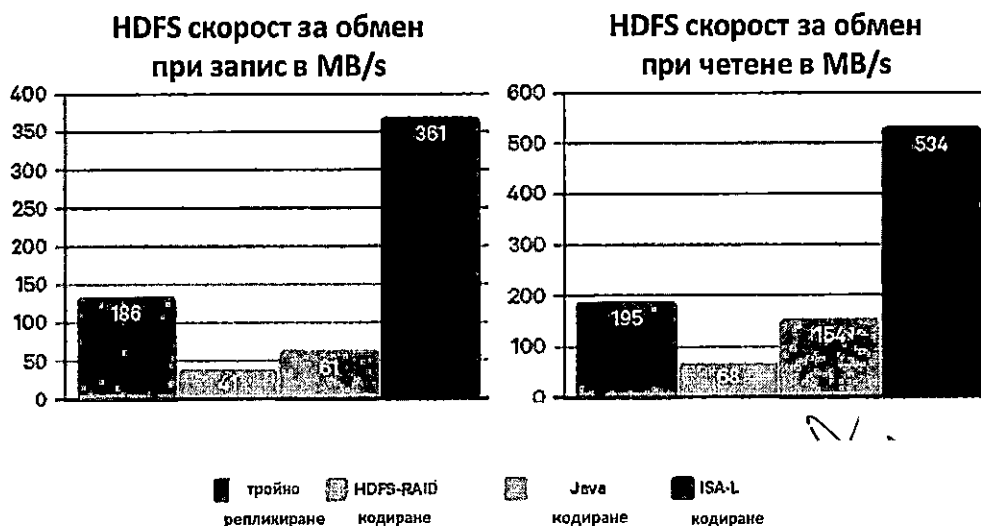
- Изграждане на ISA-L библиотека (за подробности - <https://github.com/01org/isa-l/>).
- Организиране Hadoop да използва ISA-L (за подробности - "Intel ISA-L build options" section в документа "Build instructions for Hadoop"(BUILDING.txt)). За целта се използва
 - Dbundle.isal за копиране съдържанието на директорията isal.lib в окончателния tar файл. Необходимо е библиотеката ISA-L да е достъпна както за HDFS клиентския компютър, така и за DataNodes.
 - Конфигурира се ключа io.erasurecode.codec.rs-default.rawcoder със стойност org.apache.hadoop.io.erasurecode.rawcoder.NativeRSRawErasureCoderFactory на HDFS клиентския компютър и DataNodes.

За да се провери дали състоянието на ISA-L библиотеката е активно може да се използва командата "Hadoop checknative". Той ще каже дали ISA-L библиотеката е активирана или не. Той също така изисква три стъпки, за да се даде възможност за естествено внедряване на XOR метода на KCSI. Първите две стъпки са същите като горната стъпка 1 и стъпка 2. В стъпка 3 трябва да се конфигурира ключа io.erasurecode.codec.xor.rawcoder с org.apache.hadoop.io.erasurecode.rawcoder.NativeXORRawErasureCoderFactory и на двата сървъра - HDFS клиентски компютър и DataNodes.

Използването на различните схеми за кодиране с оглед изправяне на грешки в Hadoop 3 е показано на фигура 3-9.



000126



Фиг.3-9

Изхождайки от резултатите показани на по-горната фигура, може да се стигне до извода, че ефективна схема на работа на Hadoop 3, при която се използва максимална схема за обмен на данни е чрез прилагане на КСИ RS(6,3) - използвайки Intel ISA-L библиотеката.

Използването на КСИ в Hadoop 3 има някои ограничения. Такива са операциите за запис на файлове като `hflush()`, `hsync()` и `append()`, поради някои технически предизвикателства.

3.3 Работа с малки файлове

Малки файлове са тези с обем по-малък от блока на HDFS (по подразбиране 64MB). Малките файлове създават проблем от това че в рамките на един Hadoop клъстер ще има множество по брой файлове. Всеки файл, директория и блок в HDFS е представен като обект в паметта на NameNode, всеки от които заема 150 байта, като правило. Така че 10 милиона файла, ще използва около 15 гигабайта памет. Същественото увеличаване на броя на файловете в Hadoop клъстера бе било проблем, заради хардуерните ограничения на обема на паметта на NameNode сървъра. Например един милиард файла не е възможно да се обработват в един Hadoop клъстер.



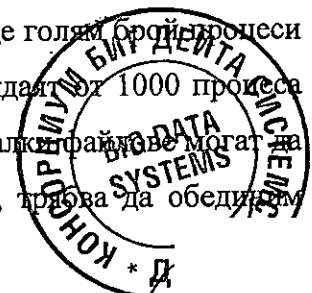
По принцип, HDFS файловата организация няма висока ефективност за открити

малките файлове и това налага специално насочено внимание. Чуден е
и целия комплекс Hadoop клъстер и приложни компютри.

Ето пример за работа с малки по големина файлове и малки-средни файлове е
обработка на OCR сканирани документи. Пример за малък файл е такъв с размер 1MB
файл, а за малък-среден файл е такъв с обем 128MB. Обработването на 128 броя 1MB
файлове ще изисква 128 пъти повече "административни" режимни, в сравнение с
обработката на един файл с дължина 128Mb. В обикновен текст без компресия от
сканирането, този 1MB файл може да съдържа 100 документа, а 128 MB файла може да
съдържа 12800 записа. В съществуващата компресия от OCR сканиране от 85-92% на
документи, се получава, че 1MB файл би съдържал над 1000 сканирани документа, а
128MB файл би съдържал над 1 милион сканирани документа. Всичко това изисква да
се използва лентова клетка с размер 64Mb, вместо 128Mb, използвана в предишни
версии на Hadoop.

Процесът "Map" обикновено обработва един входен блок в даден момент (използвайки
файла по подразбиране FileInputFormat). Ако файлът е много малък и има много от тях,
тогава всяка задача на Map обработва много малко входни данни и има много повече
задачи за разпределяне, всяка от които налага допълнителни компютърни ресурси. Ако
сравним 1GB файл, разбит на 16 64MB блока, с 10 000 броя файла с по 100kb блокове,
то тези 10 000 файла ще използват по един процес "Map" и работното време на тези
процеси може да бъде десетки или стотици пъти по-голямо от това на еквивалентния
процес "Map" работещ с един входен файл. В Hadoop има създадени няколко
функции, които спомагат за облекчаване на режима на това преговарване: стартирала
JVM за една задача може повторно да се използва за изпълнение на множество задачи
Map, като по този начин се избягва многократното стартиране на JVM
(mapred.job.reuse.jvm.num.tasks property), както и MultiFileInputSplit който може да
изпълнява повече от едно разделяне на Map. Обобщено, Hadoop се справя много добре
с големи файлове, но когато файловете са малки, то Hadoop предава всеки малък файл
на функция map(), което не е много ефективно, защото ще създаде голям брой процеси
Map. Например, файловете от 1000 (с по 2 до 3 MB) ще се нуждаят от 1000 процеса
Map, които са много неефективни. Следователно твърде много малки файлове могат да
бъдат проблематични в Hadoop. За да разрешим този проблем, трябва да обединим

000.28



много от тези малки файлове в едно и след това да ги обработим. Основната ц
решаването на проблема с малките файлове е да се ускори изпълнението на прог
) Hadoop, като се съчетаят малки файлове в по-големи файлове. Решаването на проблема
с малките файлове ще намали броя изпълнени функции на map() и по този начин ще
подобри цялостното изпълнение на задачата на Hadoop.

a) Подход на използване на стандартни програмни системи с отворен код

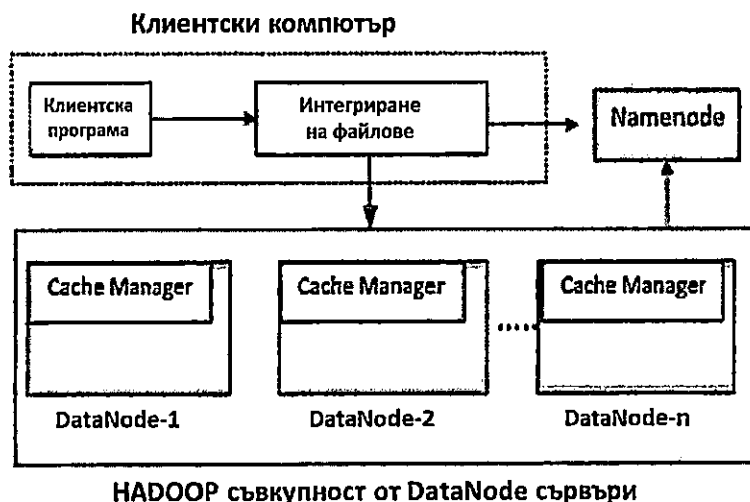
Понякога малки файлове не могат да бъдат избегнати. За да ги тяхното
влияние върху работата на Hadoop клъстера, може да се приложат следните
софтуерни системи за намаляване на броя на малките файлове:

- Apache Nifi – предлагаща обработка за консолидиране на данни и файлове, преди достигането им до Hadoop клъстер;
- Apache Flume - Използвайте архитектурна Flume за събиране, агрегиране на малките файлове и преместване в Hadoop клъстера;
- Apache Hive – Създаване на голям файл от малки файлове, който голям файл след това да се запише за последваща обработка в Hadoop. Възможно е също така и прочитане на създадения голям файл, неговото „разделяне“ до малки файлове, след което един от малките файлове да се подаде на фактическата обработка на малки файлове;
- Apache Sqoop – бидейки приложение с интерфейс използващ командния ред за прехвърляне на данни между релационни бази данни и Hadoop, да управлява броя на малките файлове за генериране на подходящи файлове с размери.

b) Подход на използване на клиентски софтуер за интегриране с Кеш механизъм в DataNode - Extended HDFS

Друг подход е използването на компонент разположен в клиентския компютър, който интегрира малки файлове в голям файл и използване на компонент на DataNode сървър, който удовлетворява кеша. Това е показано на фигура 3-10





Фиг.3-10

Същността на подхода, съответстваща на по-горе посочената фигура, е базиран на индекси. Въз основа на файловата зависимост, файловете, принадлежащи към същата директория, се интегрират в голям файл като се създава индекс за всеки малък файл, за да се намали загубата от малките файлове. Политиката за кеша е да се увеличи ефективността на четене на малки файлове. Мениджърът на кеш е разположен в DataNode, като при четене на данни от малки файлове всеки път, първо се търсят данните в кеша. Ако този ресурс не е в кеша, ресурсът може да бъде намерен от диска на DataNode. Интегрирането на файловете работи, като всеки голям файл има индексен файл, който съдържа отместването и дължината на всеки оригинален малък файл. Процесът на интегрирането се състои в:

- Сортиране на малките файлове в дадена директория въз основа на тяхното име и записване на тези малки файлове от директорията в голям файл - един по един.
- Определяне общата сума на малките файлове.
- Определяне размера на големия файл и сравнение на резултата с размера по подразбиране на HDFS блока

Този метод разширява HDFS и е наречен Extended HDFS (EHDFS). EHDFS се основава на индекси, когато броят на малките файлове е много голям в индексния файл, актуализирането на индекса може да стане по-трудно. Да се обогати прочетеното изпълнение на HDFS, се използва подход за предварително фиксиране на EHDFS осигурява подобрен механизъм за индексирание и предварително съхранение на индексна информация. EHDFS има четири техники, които играят важна роля за



000230

подобряване на ефективността, с която HDFS обработва малки файлове. Те са сливане на файлове, картографиране на файлове, предварително изтегляне и извличане на файлове. Този метод изисква създаване на съответната програмна рамка.

с) Подход използващ архивен файл

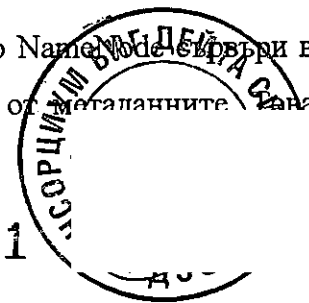
Архив на Hadoop е създаден, за да се справи с проблема с увеличаването на използването на паметта на NameNode за съхраняване на информацията за метаданните поради твърде много малки файлове. По принцип този архив позволява да се опаковат няколко малки HDFS файла в един архивен файл и следователно да се намали информацията за метаданните. Последният архивиран файл следва разширението .hag и може да се разглежда като слоеста файлова система над HDFS. Подходът на .hag файловете въвежда слой върху HDFS, който осигурява интерфейс за достъп до файловете. С помощта на командата Hadoop archive се създават .hag файлове, които изпълняват MapReduce задача, която опакова файловете, които се архивират в по-малък брой HDFS файлове. Четенето на файлове в .hag не е по-ефективно от четенето на файлове в HDFS, тъй като всеки достъп до .hag файлове изисква два индексни файла да се прочетат и след това файлът с данни да се чете.

д) Подход за използване на последователен файл

Друг подход за преодоляване на проблемите с малките файлове е изграждането на Последователен файл от малките файлове. При Последователния файл се използва името на малкия файл като ключ а неговото съдържание - като стойност. Ако имаме 10 000 файла от 100 KB, можем да напишем програма, която да ги постави в един последователен файл и след това да ги обработи използвайки подхода Разполагането на ленти.

е) Подход за използване на федерирани NameNode сървъри

Федерираните NameNode сървъри позволяват да има множество NameNode сървъри в Hadoop клъстера, всеки от които да съхранява подмножество от метаданните.



премахва необходимостта от съхраняване на всички метаданни на Hadoop клъстера в една машина, осигурявайки по-голям мащаб за използване на паметта. На повърхността, решаването на проблема с малката памет и техника е привлекателно, но този подход има определени ограничения.

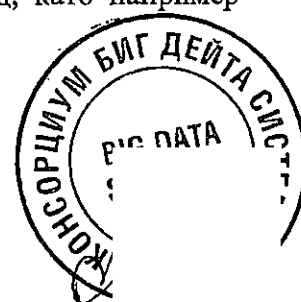
Федерираните NameNode сървъри изолират обектните метаданни - само един NameNode знае за всеки конкретен обект. Това означава да се използва файл, който трябва да знаете кой NameNode да използва. Ако Hadoop клъстерът има множество потребители и / или изолирани области на приложения, тогава обединените NameNode сървъри са естествено годни - може да се изолират обектните метаданни от потребителя или от приложението. Ако обаче се споделят данни във всички приложения в рамките на клъстера, този подход не е подходящ.

Тъй като федерацията всъщност не променя броя на обектите или блоковете в рамките на клъстера, тя не решава проблема с производителността на MapReduce. Обратно, федерацията добавя значителна и често ненужна сложност на инсталирането и администрирането на Hadoop клъстера.

f) Подход за решаване на проблема с ефективността на MapReduce

Както бе обсъдено по-горе, проблемът с ефективността на процедурата MapReduce се дължи на комбинацията от произволен достъп до диск и стартиране / управление на твърде много задачи на MapReduce. Решението изглежда очевидно - или използване на по-малко, по-големи файлове, или работа с по-малко задачи по MapReduce - но това често е по-лесно да се каже, отколкото да се направи. Някои от най-често срещаните решения включват:

- Да се промени интервала процеса на създаване на процедура Map;
- Консолидация на файловете за пакетна обработка;
- Използване на Последователни файлове (представени по-горе);
- Използване на стандартни програмни системи с отворен код, като например HBase (представени по-горе)
- Използване на CombineFileInputFormat;
- Конфигурационна настройка на Hive;
- Използване на способностите на Hadoop за добавяне.



4 Прилагане на твърдо дисково устройство (HDD) и Полупроводниково дисково устройство (SSD) в HDD

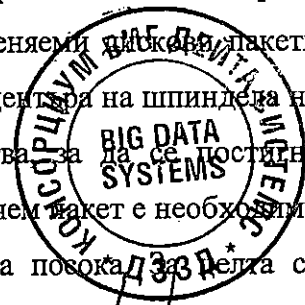
4.1 Същност на твърдия диск (HDD)

Твърдият диск (наричан още хард диск – HDD – Hard Disk Drive) е запомнящо устройство, енергонезависима компютърна памет. Информацията се съхранява по магнитен път върху покрити със специален слой плочи. Прилагателното „твърд“ подчертава разликата спрямо гъвкавия магнитен диск и напомня за исторически първото наименование запамятаващо устройство с твърд магнитен диск (ЗУТМД) използвано през 70-те и 80-те години на 20 век. Необходимостта от съхраняване на все повече информация налага устройства, използващи сменяеми дискови пакети с магнитни дискове, които разговорно също са наричани дискови пакети.

Запамятаващото устройство на твърд диск използва набор от една или повече дискови плочи (диска) около обща ос в т.нар. дисков пакет. Всяка плоча е покрита с магнитен слой, върху който информацията се записва и четете от магнитна глава. Данните се записват върху концентрични окръжности, които се наричат „пътечки“ (писти). Пистите се номерират за всяка плоча поотделно, като се започва от нулева (най-външната работна) до последната (с най-голям номер). Обикновено след нея има и няколко резервни (запасни).

Целият дисков пакет се върти с постоянна ъглова скорост около оста си, задвижван от електродвигател. При старите и големи устройства той е дори мощен трифазен електродвигател, свързан чрез ремъчна предавка с шпиндел, на който е фиксиран дисковия пакет. В съвременните устройства двигателят за пакета е най-често миниатюрен и плосък, куплиран директно към шпиндела, и управляван от специализиран контролер, стабилизиращ скоростта му на въртене.

За двете работни повърхности на всяка плоча има отделна глава за запис и четене. Блокът магнитни глави се задвижва чрез рамо, извършвайки операцията позициониране на главите чрез радиално преместване. В устройствата със сменяеми дискови пакети главите типично позиционират по права линия (радиално; към центъра на шпиндела на пакета), задвижвани от линеен двигател. При тези устройства за да се постигне съвместимост на запис и четене на информацията от всеки сменяем пакет е необходимо да се настроят магнитните глави в радиална и тангенциална посока. Делта са



използвани така наречени контролни пакети. При тези устройства магнитните глави се извеждат извън сменяемия дисков пакет и след това той може да се спре и смени.

Преди появата на персоналните компютри твърдите дискове се произвеждат в множество различни размери, като обикновено се поставят в самостоятелни или в стандартни 19-инчови шкафове. Тъй като първите персонални компютри имат вградени флопидискови устройства, твърдите дискове за този пазар започват да се конструират със същите размери – първоначално стандартните за флопидисковите устройства ширини 8", 5,25" и 3,5". По-късно се появяват и твърди дискове с по-малки размери. Към 2009 година най-малките размери (1,3", 1" и 0,85") са от производство, поради поевтиняването на флаш-паметта.

Основните технически параметри на твърдите дискове са:

- *Интерфейс* – съвкупността от връзки, сигнали, технически средства за поддръжка на връзките и правила на обмен. Модерните твърди дискове ползват интерфейси ATA (AT Attachment, или IDE – Integrated Drive Electronic, или Parallel ATA), (EIDE), Serial ATA, SCSI (Small Computer System Interface), SAS, FireWire, USB, SDIO и Fibre Channel;
- *Време за произволен (случаен) достъп* (на английски: random access time) – от 3 до 15 милисекунди;
- *Скорост на въртене на шпиндела* (на английски: spindle speed) – брой обороти на шпиндела за минута. От нея зависят времето за достъп и скоростта на предаване на данните. Към 2007 г. се произвеждат твърди дискове със скорости на въртене: 4200, 5400 и 7200 (ноутбуци), 7200 и 10 000 (персонални компютри), 10 000 и 15 000 об/мин. (сервъри и работни станции);
- *Надеждност* (на английски: reliability) – средно време за амортизация (до първия отказ) (Mean Time Between Failures, MTBF);
- *Количество входно-изходни операции за секунда* – съвременните дискове позволяват около 50 оп./сек при произволен достъп и около 100 оп./сек при последователен достъп;
- *Консумация на енергия* – важен фактор при мобилните устройства;
- *Ниво на шум* – шумът при работа на диска, измерен в децибел. За да се счита за тих, дискът трябва да има ниво на шума около 26 dB и по-малко.

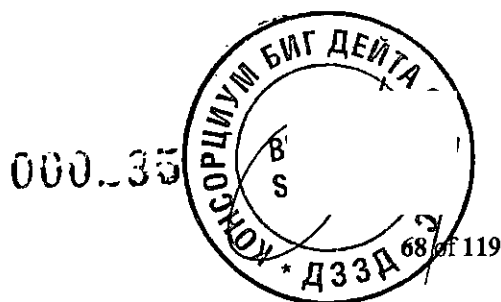


- *Устойчивост на удари* (на английски: G-shock rating) – допустимото натоварване се измерва в единици g;
- *Скорост на предаване на данни* (на английски: Transfer Rate)
 - Вътрешна зона на диска: от 44,2 до 74,5 MB/s
 - Външна зона на диска: от 74 до 260 MB/s

Дисковите контролери са управляващи устройства, които от една страна управляват работата на самото устройство, а от друга правят връзката с процесора на компютъра, чрез шините за достъп. Контролерите осигуряват доста по-бързи до данните в сравнение с това което могат да постигнат дисковете. Контролерът комуникира с електрониката на самото дисково устройство, като един от важните компоненти на електрониката на дисковото устройство е неговата кеш памет. Електрониката на дисковете имат кеш-памет и малки вградени микропроцесори, които заедно позволяват много бърз трансфер между компютъра и кеш-паметта, в сравнение с по-бавния трансфер между кеш-паметта и подложките на диска и обратно. Вградената кеш памет на дисковете варира от 32 KB до няколко MB. Кеш-паметта може да работи със скоростта на контролера и процесорът може да прочете пакети данни от диска, или пък да записва, докато работят програмите. Докато процесорът върши друга полезна работа, кеш-паметта работи с много по-бавния диск, за да завърши задачата. Употребата на кеш-памет при дисковете е по-различна, отколкото между процесора и паметта – дисковият кеш е посредник между устройства с голяма разлика в скоростите, докато кешът за паметта може да намали заявките към по-бавното устройство до нива, с които то може да се справи. Освобождаването на процесора за полезна работа, докато дисковата кеш памет буферира предаването на данни, е много важно. Поради тези взаимоотношения, увеличаването на скоростта на предаване до диска може значително да освободи процесора.

Съществуват няколко популярни стандарта за дискови контролери:

- AT;
- IDE (Integrated Drive Electronics);
- EIDE (Enhanced Integrated Drive Electronics);
- ESD (Enhanced Small Device Interface);
- SCSI (Small Computer Systems Interface);
- SATA (Serial ATA);



- SAS.

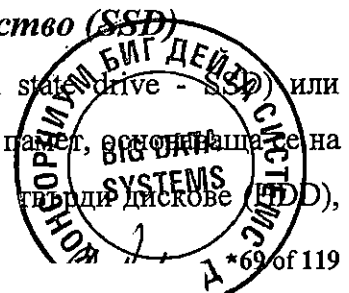
От гледна точка на Nadoop архитектура, особено внимание трябва да се обърне към SATA и SAS дисковете и контролерите. Използването на SATA контролери е възможно благодарение на факта, че и SATA контролерите и SAS контролерите имат сходни характеристики, или по-точно

- SATA устройства могат да бъдат включени в SAS контролери.
- SAS устройствата не могат да бъдат включени в SATA контролери.

Има разлика между SAS и SATA конекторите, като и двете включват мощност и данни. На SATA дисковете има разделение между захранването и данните, докато при SAS устройствата тя е унифицирана. От гледна точка на производителността, SATA устройство няма да може да съответства на производителността на SAS устройството. Ефективността на SAS дисковете надвишава тази на SATA устройства. Освен това, задвижванията SAS са предназначени за продължителна употреба. От друга страна, SATA дисковете обикновено имат по-голям капацитет за цената. Обикновено най-популярната скорост за въртене на шпиндела за SATA устройства са 7200 оборота за минута, докато SAS тя е два основни типа: 10000 и 15000 оборота за минута. SAS дисковете обикновено се използват за корпоративни изчисления, при които високата скорост и високата достъпност са от решаващо значение, като например банковите транзакции и електронната търговия. SATA дисковете се използват повече за настолни компютри, за крайни потребителите и за по-малко взискателни роли, като например съхранение на данни и архивиране. SAS устройствата са по-надеждни от SATA дисковете. Приетото средно време между неизправности (MTBF) за SAS устройства, е 1,2 милиона часа, докато за SATA устройства то е 700,000 часа. От физическа гледна точка кабелите също са значително различни. SATA кабелите са ограничени до 1 метър дължина, като частта за данни и за захранване са отделни, докато кабелите на SAS могат да бъдат с дължина до 10 метра с мощност и данни, предоставени чрез същия кабел.

4.2 Същност на полупроводниково дисково устройство (SSD)

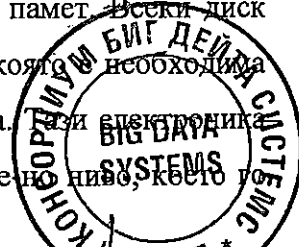
Полупроводниково дисково устройство (на английски: Solid state drive - SSD) или статично дисково устройство е енергонезависима компютърна памет, базирана на технологията на флаш-паметите. За разлика от традиционните твърди дискове (HDD),



статичните полупроводникови устройства не съдържат подвижни механични части това елиминира закъснението при четене и повишава значително скоростите на работа. Наименованието „solid state“ се дължи на историческото развитие в електронната техника и означава, че като градивен материал е използван полупроводник (силиций).

Флаш паметта е разработена от еднократно изтриваема програмируема памет за четене (EPROM) и програмируема памет с възможност за четене (EEPROM), която може да бъде изтрита с електромагнит. Флаш паметта е технически вариант на EEPROM, но индустрията запазва термина EEPROM за изтриваема памет за ниво байт и създава термина „флаш памет“ за ниво блок. Устройства, използващи флаш памет, изтриват данни на ниво блок и пренаписват данни на ниво байт - NOR flash - или многобайтово ниво на страница - NAND flash. NAND паметите са тези, които се използват в момента в компютърната индустрия. NAND флаш паметта е по-подходяща за съхранение на данни в настолни компютри и корпоративни сървъри и системи за съхранение поради по-ниската им цена на бит за съхранение на данни, по-голяма плътност и по-високи скорости на програмиране и изтриване (P/E). Устройствата с флаш памет имат ограничен брой P/E цикли, тъй като всеки цикъл причинява малко физическо увреждане на носителя. Това увреждане се натрупва с течение на времето, което в крайна сметка прави устройството неизползваемо. Броят на P/E цикли, които дадено устройство може да поддържа, преди проблемите да станат забранени, варира в зависимост от вида на технологията. Недостатък на SSD дисковете е ограниченият брой P/E цикли, преди да започнат да се амортизират и станат неизползваеми. Животът им не е никак кратък и не е по-малък от този на механичните хард дискове. Тестовите сочат продължителност на живота на тези устройства от около 10 години при сравнително високо натоварване. Освен това SSD дисковете може да загубят данните си ако електричеството прекъсне. Оставането на SSD без електричество повече от 8 дни може да доведе до загуба на данните.

Ключовите компоненти на всеки статичен диск са електрониката към него и паметта за съхраняване на данни. Основният компонент на паметта е NAND памет. Всеки диск включва електрониката, управляваща работата на устройството и която е необходима за свързване на NAND компонентите от паметта с хост компютъра. Тази електроника включва вграден микропроцесор, който изпълнява код на Firmware на ниво, което



000.37

прави един от най-важните фактори за бързодействието на тези дискове. Някои от

функциите на тази електроника са:

- Error-correcting code (ECC)
- Wear levelling
- Маркиране на лошите сектори
- Read scrubbing and read disturb management
- Кеш-памет
- Garbage collection
- Шифроване

Производителността на един статичен диск е пропорционална на количеството паралелни NAND чипове, използвани в устройството. Единичният NAND чип е относително бавен поради тесния (8/16 бита) асинхронен I/O интерфейс, както и с допълнителна висока латентност при основните I/O операции. Когато в един диск работят паралелно множество NAND устройства, честотата се мащабира и високата латентност може да се скрие, стига да има достатъчно чакащи операции и натоварването е разпределено по равно между различните устройства.

Най-често в практиката се прилагат дискове с 2,5-инчов и 3,5 инчов размер.

Предимствата на полупроводниковите дискови устройства пред устройствата с твърд диск са:

- значително по-бързи са – технологията им позволява да имат почти нулево време за търсене на „сектор“ с информация и много по-добри времена за запис и четене. Windows 10 стартира три до четири пъти по-бързо;
- липсата на движещи се механични части води до по-дълъг живот;
- четенето на данни не скъсява живота на устройството;
- имат много по-голяма плътност на единица площ (позволява създаването на малки по обем, но големи като капацитет дискове);
- по-издръжливи са на вибрации и механични смущения (удари, изпускания);
- използват доста по-малко енергия (поради липсата на механика);
- не се нагряват;
- напълно безшумни са;
- липсват вибрации при работа.



Недостатъците на полупроводниковите дискови устр

- все още по-високата цена;
- имат ограничен брой записи които да извършват, който обаче брой е съизмерим с издръжливостта на твърдите дискове.

Вследствие на постоянното намаляване на цената, полупроводниковите дискови устройства се доближават, но все още остават по-скъпи от традиционните: един статичен диск с капацитет 300 – 600 GB към декември 2012 г. е два пъти по-скъп от 2,5-инчов твърди диск с 500 GB за лаптоп[15]. В сегмента на бързите статични дискове цената на един бит към есента на 2014 г. практически е сравнима с тази при 2,5-инчовите твърди дискове от най-високо ниво със скорост на въртене на шпиндела 15 000 rpm (оборота в минута). Тези статични дискове ще станат по-евтини от споменатите твърди дискове още през 2019 г., а към 2025 г. се очаква и корпоративните статични да станат по-евтини от традиционните твърди дискове, прогнозира Toshiba.

Статичните дискови устройства обикновено използват същите файлови системи, както и твърдите дискове.

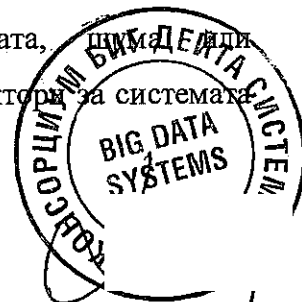
4.3 Основни параметри за сравнение между HDD и SSD

Както SSD, така и твърдите дискове изпълняват една и съща работа – съхраняват данни. Но всеки тип хранилище има свой уникален набор от функции. Те се различават по следните параметри:

- **Цена:** SSD са по-скъпи от твърдите дискове от гледна точка на долар на гигабайт. Това се изразява в 4 до 5 цента на гигабайт за твърдия диск и 20 цента на гигабайт за SSD. Тъй като твърдите дискове използват по-стари технологии, те ще останат по-евтини за близко бъдеще.
- **Максимален и общ капацитет:** Докато 500GB се счита за "базов" твърд диск през 2017 г., по ценообразуване това съответства на 128GB за по-евтини SSD дискове.
- **Скорост:** Това е параметър, където SSD дисковете имат определен приоритет. Един персонален компютър с SSD се зарежда само за секунди. Един твърд диск изисква време, за да ускори оперативните характеристики, и ще продължи да бъде по-бавен от SSD по време на нормална употреба.



- *Фрагментиране:* Поради ротационните си повърхности на запис, твърдите дискове работят най-добре с по-големи файлове, които се поставят в съседни блокове. По този начин главата на устройството може да започне и да приключи четенето си с едно непрекъснато движение. Когато твърдите дискове започват да се запълват, големи файлове могат да се разпръснат на различни части от диска, което кара устройството да страда от това, което се нарича фрагментация. Докато алгоритмите за четене / запис са се подобрили дотолкова, че ефектът е сведен до минимум, твърдите дискове все още могат да се фрагментират. SSD не може обаче, защото липсата на физическа четяща глава означава, че данните могат да се съхраняват навсякъде еквивалентно. Така SSD са по-бързи.
- *Устойчивост:* SSD няма движещи се части, така че е по-вероятно те да запазят данните си в безопасност, в случай физически удар, докато диска работи. Повечето твърди дискове паркират главите си за четене / запис, когато системата е изключена, но те летят над диска на разстояние няколко нанометра, когато са в експлоатация. Освен това дори спирачките за паркиране имат ограничения.
- *Наличност:* Твърдите дискове са по-разпространени в по-старите системи, докато SSD-тата стават все по-разпространени в ново пуснатите компютри. Видовете SSD модели се увеличават.
- *Шум:* Дори и най-тихият твърд диск ще излъчва малко шум, породен от въртенето на шпинделите на устройството или четящото рамо на главите, движещи се напред-назад. По-бързите твърди дискове ще направят повече шум от тези, които са по-бавни. Софтуерът за SSD няма почти никакъв шум, тъй като те не са механични и по този параметър са с огромно предимство.
- *Мощност:* Един SSD не трябва да изразходва електричество, когато не оперира, защото няма постоянно движещи се части. Следователно, нито една от консумираната от SSD енергия не се губи за триене или движение на главите, което ги прави по-ефективни.
- *Общо:* твърдите дискове печелят от цената, капацитета и наличността. SSD работи най-добре, ако скоростта, здравината, формата, фрагментацията (технически част от скоростта) са важни фактори за системата Hadoop.



4.4 Хибридни устройства и дуални системи

В средата на 2000-те, някои производители на твърди дискове, като Samsung и Seagate, теоретизираха, че ако добавите няколко гигабайта флаш чипове към въртящ се твърд диск, ще получите така нареченото "хибридно" устройство, комбиниращо голям твърд диск капацитет за съхранение с производителност на SSD, на цена, която е само малко по-висока от тази на типичния твърд диск. В тези хибридни устройства флаш паметта служи като буфер за често използвани файлове, така че системата има възможност за зареждане и стартиране на най-важните приложения по-бързо, въпреки че не може да се инсталира нищо директно в това пространство. На практика хибридните устройства работят, но те все още са по-скъпи и по-сложни от обикновените твърди дискове. Те работят най-добре за случаи, които се нуждаят едновременно от много голям капацитет и бързо време за зареждане. Тъй като те са продукт между продуктите, хибридните устройства не заместват непременно отделените твърди дискове или SSD устройства.

В система с дуални дискове, системата инсталира малък SSD основен диск (C :) за операционната система и приложенията и добави по-голям въртящ се твърд диск (D: или E :) за съхранение на файлове. Този тип приложение не е типичен за работа на Hadoop, поради което не представлява интерес в настоящата архитектура.

4.5 SAS дискови устройства

В изчислителната техника, Serial Attached SCSI (Serial Attached SCSI) - SAS е сериен протокол от точка до точка, който трансферира данни към и от устройства за съхранение на данни като твърди дискове и лентови устройства. SAS замества по-стария паралелен интерфейс SCSI (интерфейс за Small Computer System Interface), която за пръв път се появи в средата на 80-те години. SAS, подобно на своя предшественик, използва стандартния набор от команди за SCSI. SAS предлага опционална съвместимост със Serial ATA (SATA), версия 2 и по-нова версия. Това позволява свързването на SATA устройства към повечето SAS базови платформи или контролери. Обратното, свързващо SAS устройствата със SATA базови платформи, не е възможно.

Версиите на протокола SAS са:

- SAS-1: 3.0 Gbit/s, въведен през 2004;
- SAS-2: 6.0 Gbit/s, въведен през 2009;



- SAS-3: 12.0 Gbit/s, въведен през 2013;
- SAS-4: 22.5 Gbit/s наречен още "24G", въведен през 2017;

Дисковете, чиято електроника осигурява SAS интерфейс се наричат „SAS дискове“.

Технологията SAS-24Gb / s е нещо повече от ускорение за Serial Attached SCSI интерфейс от 3GB / s, но е основна промяна на технологията. Новото свързване на тези дискови устройства ще поддържа редица важни оптимизации за твърди дискове, ще използва изцяло новото 128b / 150b кодиране, 20-битовата корекция на грешки, както и редица нови функции. Докато SAS-24Gb / s ще изисква нови конектори, технологията ще поддържа обратната съвместимост с предишни поколения SAS интерфейси, което означава, че по-старите дискове ще бъдат съвместими с по-новите контролери. Докато SAS-4 е оптимизиран за SSD, той все пак поддържа различни SCSI-специфични команди, които ще повлияят на производителността. В резултат на това, ако подсистемата за съхранение трябва да е съвместима както със SSD, така и с твърди дискове, SAS-24Gb / s ще бъде естественият избор. Въпреки това, ако се използват само SSD, използването на SAS не е точно идеалното решение.

4.6 Защо HDD ще бъдат изместени от SSD

Хард дисковете (HDD) са неизменна част от компютрите в продължение на повече от 60 години. Съвременните устройства вече са оборудвани с още по-бързите SSD дискове. Според TrendForce, през 2017 година около 33% от компютрите на пазара имат SSD дискове, но се очаква тази стойност да се увеличи до 56% до 2018г., което води до силно намаляване на цената им. Само след две години повечето компютри на пазара ще разполагат с SSD дискове и така ще изместят класическите хард дискове. Развитието на технологиите ще позволи на производителите не само да ги направят по-бързи, но и да увеличат значително техния капацитет. Компаниите вече разработват ново поколение NAND флаш памет и е само въпрос на време те да станат част от устройствата. През 2016 година се появи SSD устройство с капацитет почти 16TB – [26]. В началото на 2018 година светът вече разполага със SSD диск с капацитет 30TB, което е невъзможно за достигане капацитет на HDD, като скоростта на четене е 2100MBps, а скоростта за запис е 1700 MBps – [27]. Независимо от сравнително високата им цена в момента на обявяването на тези устройства, то в рамките на няколко години, цената им ще спадне и ще започне да ги доближава до тази на HDD устройствата. В края на 2017

г. бе обявен най-големия за момента капацитет на SSD – 60 TB [28], първоначално като устройство с демонстративна технология и през 2018 г. да стане реален продукт на пазара.

Освен това тези устройства стават все по-близки до високо производителните устройства и вече имат над 5 години гарантирано опериране. Всичко това определя насоката на налагане на SSD устройствата над HDD.



000243

5 Използване на HDD и SSD устройства в Hadoop клъстер

Известно е, че полупроводниковите дискове (SSD) са бързи но скъпи от твърдите дискове (HDD). Но дали бързодействието им може да се компенсира от цената и при какви случаи в приложения на Hadoop клъстер? За да отговорим на този въпрос ще разгледаме два сценария: първият - когато се създава нов клъстер (дали SSD или HDD с еднаква обща скорост за обмен на данни са по-добри); и вторият - как да се конфигурират хибриден клъстер - от SSD и HDD устройства.

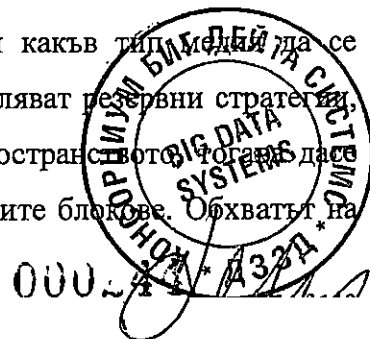
5.1 Хетерогенна памет в HDFS - типове и политики

В Hadoop 2 бе въведена възможността за хетерогенна медия (тип физическо устройство), чрез която да се изгради паметта на HDFS. Възможни видове медии (типове устройства) с техните обозначения са и основно използване:

- **DISK** - Съхранение в твърди дискове (HDD): Данните се съхраняват върху HDD;
- **SSD** - Съхранение в полупроводникови дискове (SSD или Flash): Ниво на съхранение, което се състои от полупроводникови дискове, имащи по-ниски нива на закъснение за четене / писане;
- **ARCHIVE** - Архивно съхранение, което може да бъде на различни физически типове устройства (прилагат се обикновено HDD, но по изключение - и в дискови системи от типа оптически дискове). Използва се за по-рядко достъпни масиви от данни;
- **RAM_DISK** - Единична реплика в паметта (Tech Preview): Разполага се едно копие от всичките копия за репликиране в RAM паметта, подкрепено от асинхронни записвания на диск за постоянно съхранение на данните.

Всеки DataNode в HDFS е конфигуриран за работа с дискове на които се запазват HDFS файловете. Администраторите могат да маркират всеки диск какъв вид медия представлява, например, монтираният диск може да бъде обозначен като архивно хранилище, а друг - като SSD.

HDFS използва политиките за съхранение, за да определя какъв тип медия да се използва за всяко репликирано копие. Тези политики позволяват резервни стратегии, при които ако желаният тип хранилище няма свободно пространство, тогава да се използва резервен тип хранилище за съхраняване на файловете блокове. Обхватът на



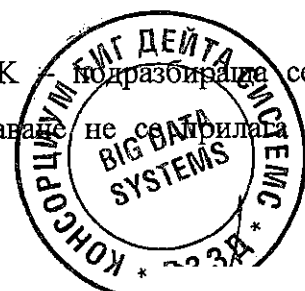
тези политики е на ниво директория - важи всички файлове в директорията. Политиките за съхранение могат да бъдат прилагани по време на създаването на файл, както и във всеки момент по време на живота на файла. За политиките за съхранение, които са се променили по време на живота на файла, HDFS въвежда нов инструмент, наречен Mover, който може да се изпълнява периодично, за да се мигрира всички файлове в клъстера, според политиките за съхранение.

Имаме следните 6 политики за съхранение, които се отнасят до предназначението на блоковете от данни (именувани са според "горещината на използване на данните"):

- Горещ (**Hot**) - както за съхранение, така и за място от което се вземат данните за изчисление. Данните, които текущо се обработват и все още се използват за обработка, трябва да бъдат в такава политика. Когато блокът е горещ, всички копия се съхраняват като DISK.
- Студен (**Cold**) - само за данни свързани с ограничени изчисления. Данните, които вече не се използват, или данните, които трябва да бъдат архивирани, се преместват от горещо съхранение в студено съхранение. Когато един блок е студен, всички копия се съхраняват като ARCHIVE.
- Топъл (**Warm**) - частично горещи и частично студени. Когато един блок е топъл, някои от неговите копия се съхраняват в DISK, а останалите копия се съхраняват в ARCHIVE.
- Всички копия в SSD (**All_SSD**) - за съхранение на всички копия в SSD.
- Едно копие на SSD (**One_SSD**) - за съхранение на едно от копията в SSD. Останалите копия се съхраняват в DISK SSD.
- В RAM (**Lazy_Persist**) - за писане на блокове с единична реплика в RAM паметта. Репликата първо се написана в RAM_DISK и след това се прехвърля в DISK.

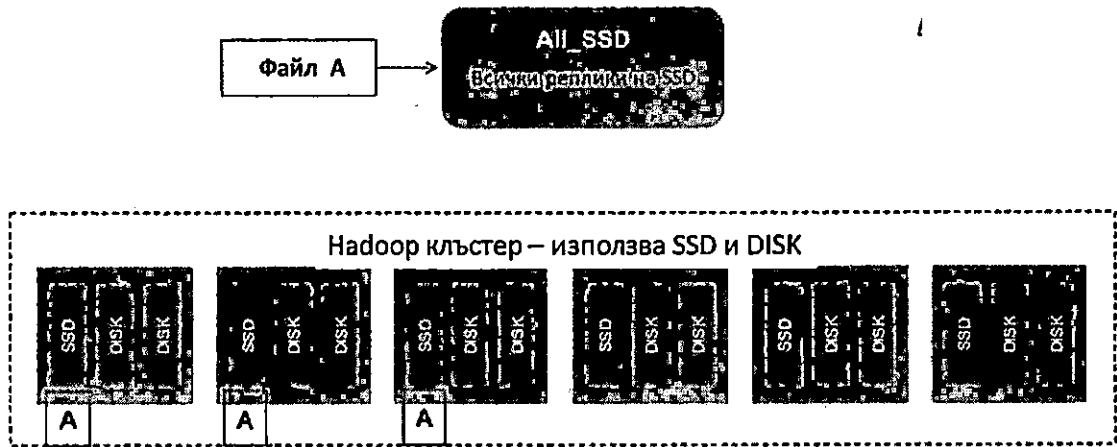
Типични сценарии на използване на различни устройства с различни политики са следните:

- Политика Hot – разполага всичките копия на DISK; Резервно съхранение по време на създаване на файл - ARCHIVE; Резервно съхранение по време на работа - ARCHIVE;



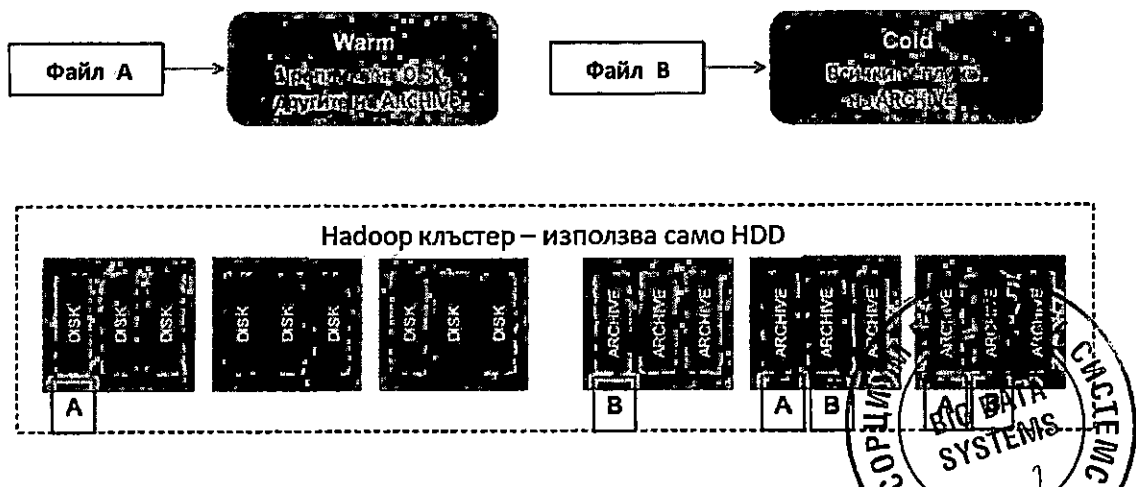
000245

- Политика One_SSD – разполага едно копие на SSD, а другите на DISK
Резервно съхранение по време на създаване – SSD, DISK / Резерв
съхранение по време на работа – SSD, DISK ;
- Политика All_SSD - разполага всички копия на SSD; Резервно съхранение по
време на създаване - DISK / Резервно съхранение по време на работа - DISK
– фигура 5-1;



Фиг.5-1

- Политика Warm – разполага първото копие на DISK а останалите на
ARCHIVE; Резервно съхранение по време на създаване – ARCHIVE, DISK /
Резервно съхранение по време на работа – ARCHIVE, DISK – фигура 5-2;
- Политика Cold – разполага всичките копия само на ARCHIVE – фигура 5-2;



Фиг.5-2

- Lazy_Persist – първото копие се разполага в RAM паметта, а остана DISK.

На тази фигура всички дисковете на архивните нива са маркирани с типа на съхранение "ARCHIVE". След това администраторите могат да прилагат "Студената" политика за съхранение на набори от данни, които трябва да бъдат съхранени в архивните възли за съхранение.

За приложения, които трябва да записват данни, които са временни или изискват кратко време за четене и запис, паметта (RAM) е подходяща среда за тяхното съхранение, която осигурява ниска латентност за четене и записване. Тъй като паметта е летлива среда за съхранение, данните, записани в паметта, ще бъдат асинхронно запазени на диска - DISK.

На таблица 5-1 е показано препоръчително прилагане на някои политики според използването на данните.

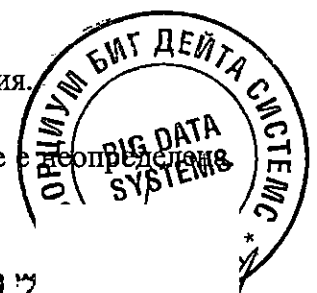
Таблица 5-1

Тип политика	Продължителност на съществуване на тези данни	Честота на опериране с тези данни
Hot	< 7 дни	20 пъти на ден
Warm	7 дни – 1 месец	5 пъти на седмица
Cold	1 месец – 3 месеца	5 пъти на месец

Формално, всяка политика за съхранение се състои от следните полета:

- Идентификационен номер на политиката;
- Име на политиката;
- Списък на типовете медии (технически устройства) за съхранение за всички блокове;
- Списък с резервни типове хранилище използвани по време на създаване на файлове;
- Списък на резервните типове хранилища за репликация.

Когато се създаде файл или директория, политиката му за съхранение е реопределена. Правилата за хранилището могат да се задават чрез командата



000-47

dfsadmin -setStoragePolicy.

Ефективната политика за съхранение на файл или директория се решава от следните правила.

- Ако файлът или директорията са конкретизирани с политика за съхранение, да се използва.
- За файл или директория с неопределена политика, ако това е основната директория, да се използва стандартната политика за клъстера. В противен случай да се използва предишно използваната политика.

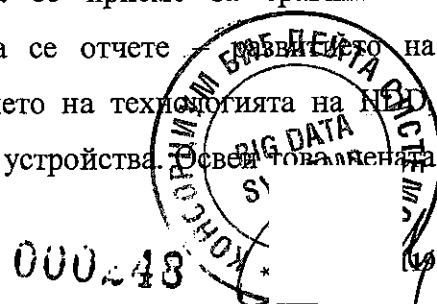
Ефективната политика за съхранение може да бъде извлечена чрез команда

dfsadmin -getStoragePolicy

5.2 Анализ работата на Hadoop клъстер с използване на HDD и SSD устройства

Едно от технологическите предимства на полупроводниковите памети пред тази на твърдите устройства – възможността за ефективна промяна на минимална единица информация (например един байт), не е от съществено значение за Hadoop клъстер, защото след промяна на един блок от данни, независимо дали той е 128 MB или е 64 MB, то се изисква преизчисляване на целия блок за корекция от грешки (при КСИ) или създаване на няколко негови реплики.

За провеждане на изследване свързано със сравнение използването на HDD и SSD, се използва конфигурацията за съхранение показана в таблица 5-2, където процесорите са Intel Xeon 2-socket, 8-ядрени, с 16 нива на прекъсване, с 48GB RAM и с 10Gbps Ethernet мрежово свързване, с HDD от 2 TB и SSD от 1,3 TB – [22]. Трябва да се има предвид, че посочените технически параметри за сравнение са стари – „много“ стари или „по-малко стари“, според гледната точка, но във всички случаи не са от текущия месец на създаване на настоящия документ. Но изхождайки от средните технологически различия, относителността на резултатите може да се приеме за сравнително достоверна. Има един факт, който също трябва да се отчете – развитието на технологията на SSD се развива по-бързо от развитието на технологията на HDD поради технологическата основа на двата вида дискови устройства.



на SSD спада по-бързо отколкото цената на HDD (при първоначално доста по-в цена на SSD), като обемите на двата типа устройства започват да се доближават. По-долу посочените резултати не са свързани с капацитета на отделните устройства и с техните цени, които са най-зависими от текущото състояние на технологиите. Тези резултати са свързани с времето за четене и запис, които са сравнително независими от текущото състояние на технологиите.

Таблица 5-2

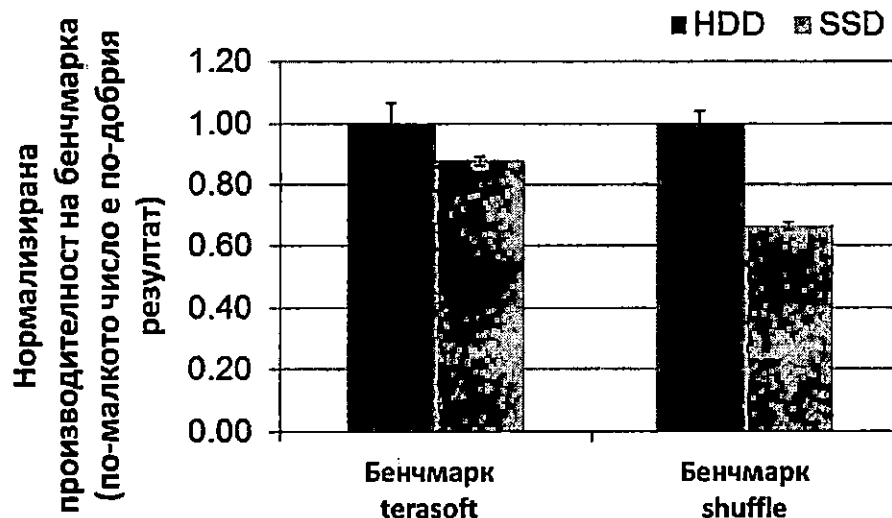
Устройство	Организация на паметта	Капацитет	Скорост за обмен на данни	Цена
HDD	11 броя HDD	22TB	1300 Mbps	\$4 400
SSD	1 брой SSD	1.3TB	1300 Mbps	\$14 000

Общото разбиране е, че MapReduce работи само с големи файлове, с последователна обработка на четения и запис. MapReduce показва добър резултат при последователно четене и записване от/към HDFS. По друг начин изглеждат нещата при четене и запис на по-малки файлове. Резултатът от едно задание Map се свежда до разделене на много задачи и всяка задача работи само със собствените си данни. Това води до обстоятелството, че всяка Reduce задача се работи с по-малки по обем данни – от по няколко MB.

В съответствие с физическия си дизайн, SSD не са ефективни за много търсения на малки данни в сравнение с твърдите дискове. SSD работят много сравнително по-добре с MapReduce задачи. За сравнение се използват 2 бенчмарка: "terasort" е общ бенчмарк със съотношение 1: 1: 1 между входни: междинни: изходни размери; "shuffle" е бенчмарк само за разбъркване, с който се набляга върху междинните резултати на MapReduce. Резултатите от тези 2 бенчмарка показват, че SSD предлага по-висока производителност – фигура 5-3.



000.49

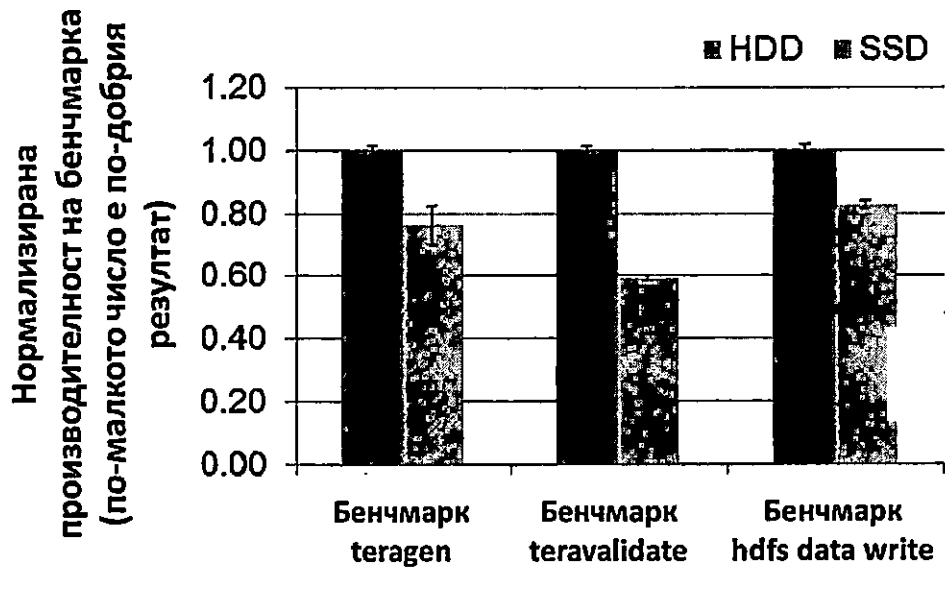


Фиг.5-3

Интересно е да се отбележи, че SSD показват по-добри резултати за задачите на MapReduce, които включват само четения и запис в HDFS, въпреки че твърдите дискове имат еднаква последователна скорост за предаване на данни според хардуерните спецификации. Трябва да се отбележи, че при бенчмарка "terasort", при който равномерно се чете, преместват и записват се данните по локалните дискове на DataNodes и резултатите от крайната Reduce обработка се записват в HDFS, то SSD дава предимство от около 15%-20%, което не е съществено предимство, особено като се има предвид значително по-високата цена на SSD.

За провеждане на следващото - второ изследване са използвани нови 3 бенчмарки: "teragen" за записване на данни в HDFS с трикратна репликация; "teravalidate" за четене резултат от работата на „terasort“ и проверява дали са в сортиран ред; и "hdfs write data" за записване на данни в HDFS с еднократна репликация. Резултатите от това сравнение на трите посочени бенчмарка е показано на фигура 5-4, при което се констатира отново предимството по производителност на SSD устройствата.

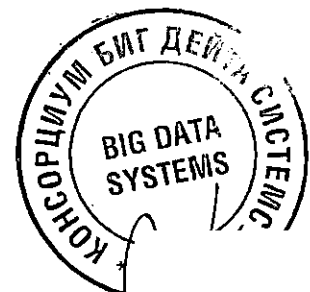


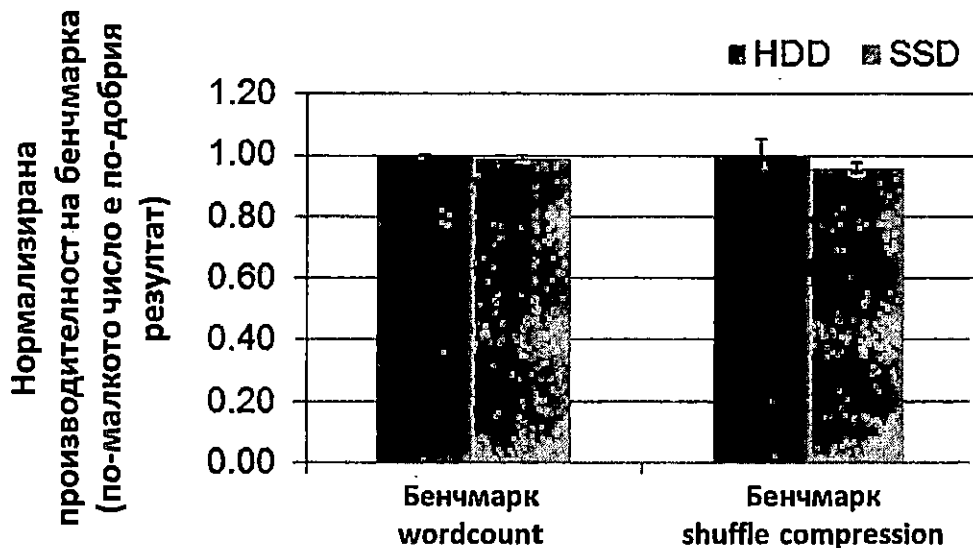


Фиг.5-4

Трябва да се отбележи, че значителен ефект в бързодействието от около 40% при използване на SSD се получава при четене на HDFS данни и не сложна изчислителна процедура като сортиране.

За провеждане на следващото – трето изследване, се използва друг вид бенчмарки при работа на MapReduce, доминирана от изчисления (процесорна обработка), а не от входно-изходни операции на дискове. Когато ограничаването на ресурсите на бенчмарка не е входно-изходната подсистема, изборът на дискове за съхранение не е от съществено значение. Тези 2 броя бенчмарки с които са проведени изследванията са: "wordcount" включващо преобладаваща работа на процесор за обработка на отчитане на честотите на думите; и "shuffle compressed" - за компресиране при разбъркване след Map и преди Reduce. Резултатите от това изследване са показани на фигура 5-5, от която се вижда, че предимството от SSD намалява значително.





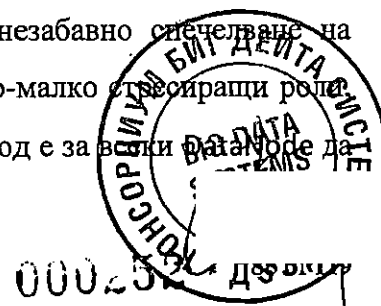
Фиг.5-5

Почти всички съществуващи клъстери MapReduce използват преобладаващо твърди дискове. При тях всеки DataNode има набор от твърди диска, създадени като JBODs.

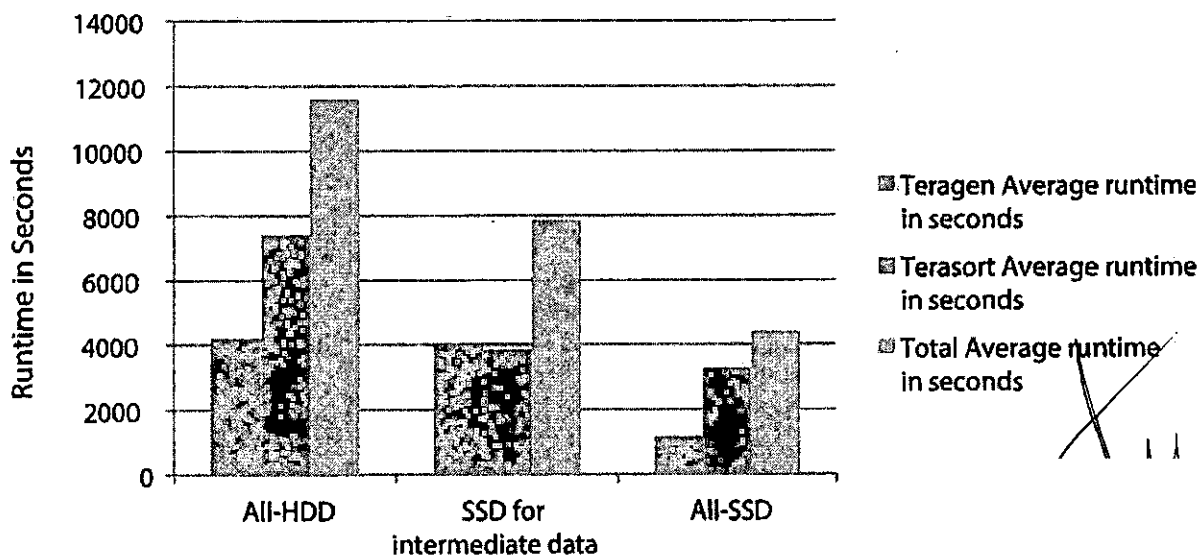
Има 5 начина за въвеждане на SSD:

- да се създаде нов клъстер със само SSD – SSD клъстер;
- да се добави нов рак (ракове) само със SSD – хибриден клъстер с SSD ракове;
- използване на SSD само за междинни съхранения на данни – за данни свързани с разбъркването след Map процедура, преди Reduce процедура;
- да се добави SSD към съществуваща съвкупност от HDD на съответните DataNodes – клъстер с комплексни SSD+HDD дискове за HDFS;
- да се добави SSD към съществуваща съвкупност от HDD на съответните DataNodes – клъстер с комплексни SSD+HDD дискове в DataNodes за HDFS и за съхранение при разбъркване след Map процедура, преди Reduce процедура.

Естествен подход изглежда че е да се заменят всички твърди дискове със SSD навсякъде в Hadoop клъстера. Въпреки това, внедряването на SSD в стотици или хиляди сървъри би могло да доведе до много скъпо решение. По-добрият подход е да се идентифицират критичните места, където SSD позволява незабавно спечелване на разходи за изпълнение и интегриране с HDD, използвани в по-малко стресиращи роли за да се спестят системни разходи. Честият практически подход е за всеки DataNode да



се използват HDD създадени като JBODs и един или няколко SSD. При случай на създаване на клъстер само със SSD дискове, те се конфигурират като JBOD. От изследване показано в [31], използващи посочените по-горе бенчмарки с 1TB-ови тестови файлове, се получават резултати показани на фигура 5-6.

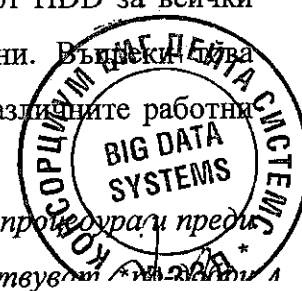


Фиг.5-6

Тези резултати показват, че използването на 1 брой SSD в DataNode, намаля в 2 пъти времето за използване с общо предназначение на Hadoop клъстера в сравнение с клъстер изграден само от HDD, докато при записване на данни в HDFS с трикратна репликация (бенчмарк "teragen"), то използването на 1 брой SSD не води до съществено намаляване времето за работа на дисковата система. Във всички случаи работата с DataNode сървъри, в които всички дискове са SSD намалява съществено времето за обработка при съществено увеличаване цената на целия Hadoop кръстер.

От представени всичките по-горе изследвания, обобщено може да се заключи:

- *Обща тенденция:* SSD е по-добър от съвкупност от HDD за всички задания, със и без междинна компресия на данни. Във всички случаи ползите от използването на SSD се различават в различните работни места и различните приложения;
- *SSD подобрява процеса на разбъркване между Map процедура и Reduce процедура, като подобрява съществено*



результати при големи размери на блоковете за разбъркване: SSD и полза при разбъркване. Интересното е, че ползите са слабо забележими когато междинните данни се компресират;

- SSD е от полза при четене и писане на HDFS: Важен резултат е, че SSD е от полза и за четене и писане на HDFS, като SSD е способен на приблизително 2 пъти повече последователни входно-изходни действие по сравнение на HDD;
- Задачите които са с по-активно използване на CPU не се влияят от типа на диска – SSD или HDD.

Съществуват решения, работещи не върху стандартен Hadoop, а върху модифициран Hadoop, при който се изгражда хибридна интегрирана архитектурна среда RAM-SSD с което се получава значително ускоряване на процесите на обработка в Hadoop.

За да с прави интегрирана оценка могат да се използват няколко параметъра:

- Скорост за предаване на данни за цена на устройство (в \$ или Euro);
- Обем на дисково устройство за цена на устройство (в \$ или Euro);
- Цена (US\$) за единица скорост за предаване на данни;
- Цена (US\$) за единица обем;

Трябва да се има предвид, че посочените параметри и архитектурното решение за SSD-HDD резолюция са валидни както за всеки отделен тип приложение, така и за всеки отделен софтуерен продукт, изпълняван в Hadoop клъстер.

5.3 Съвместно използване на КСИ и SSD устройства

Кодирането срещу изтриване (КСИ) се прилагат в Hadoop клъстери, включващи HDD и SSD, за да защитят масивите от дискове срещу откази. Прилагането на тези кодове в базирани на SSD системи води до допълнителен брой цикли Програмиране/Изтриване (P/E) на всеки диск, което може да ускори износването на SSD дискове. Това означава че докато КСИ подобряват надеждността и бързодействието на Hadoop системите, работата със SSD въвежда страничен ефект, който може да понижи надеждността на самите SSD устройства, тъй като броят на P/E циклите се увеличава.



6 Метод за повишаване производителността на Hadoop система

5.1 Повишаване производителност на HDFS системата

Един от стандартните подходи за повишаване на производителността на Hadoop система е избор на определени параметри, определящи производителността на работа на файловата система HDFS. За целта, предлаганият метод е насочен към определяне на следните параметри:

- а) Да се използват препоръчителните опции за всички HDFS дискове за данни

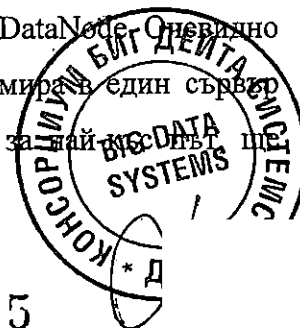
Има специфични опции за монтиране на файловата система, които се оказват ефективни за опериране на клъстерите на Hadoop. Тъй като опциите за монтиране се прилагат при монтиране на файловата система, например при системно зареждане или пренастройка, промените в / etc / fstab сами по себе си не са достатъчни, за да влязат в сила тези настройки. Настоящият метод препоръчва да се направи промяна на опцията за монтиране, като или ръчно да се подменят отделните файлови системи, или да се рестартира хоста, за да влязат в сила настройките.

- б) Да се определи чрез конфигуриране размера на блоковете на HDFS за системата

Оптималният размер на блоковете на HDFS увеличава производителността на NameNode, както и производителността на изпълнението на задачите. За целта е необходимо да се зададе размерът на блоковете ("dfs.blocksize" в hdfs-site.xml) да бъде в препоръчителния диапазон от 134217728 до 1073741824.

- в) Да се определи най-късият път при четене

В HDFS, прочитанията обикновено минават през DataNode. По този начин, когато клиентът поиска да чете файл, DataNode чете този файл от диска и изпраща данните до клиента през TCP socket. Така нареченият най-къс път на четене осигурява клиента да чете директно файла, прескачайки работата на самия DataNode. Очевидно е, че това е възможно само в случаите, когато клиентът се намира в един сървър заедно с данните. За да се конфигурират локалните четения за най-къс път, трябва да се активира libhadoop.so (dfs.domain.socket.path).



- д) Да се избягват размерите на файловете, които са по-малки от размера на блока

Средният размер на блока трябва да е по-голям от препоръчителния размер на блока от 67108864 MB. Средният размер под препоръчителния размер добавя повече натоварване към NameNode, освен това и причинява неефективно съхранение и обработка на данните.

- е) Да се настрои JVM на DataNode за оптимална производителност

DataNode е чувствителен към ефективността и поведението на JVM. Една типична конфигурация може да бъде:

Djava.net.preferIPv4Stack=true,

XX:ParallelGCThreads=8,

XX:+UseConcMarkSweepGC,

Xloggc:*,

verbose:gc,

XX:+PrintGCDetails,

XX:+PrintGCTimeStamps,

XX:+PrintGCDateStamps

- ф) Да се избягва четене или писане от остарели данни

DataNodes, които не са изпратили „heartbeat“ на NameNode за определен интервал, може да са под натоварване или да са умрели. Трябва да се избягва изпращане на заявки за четене / писане до такива "остарели" данни. За целта в hdfs-site.xml да се постави:

dfs.namenode.avoid.read.stale.datanode=true

dfs.namenode.avoid.write.stale.datanode=true

- г) Да се използва наблюдение на работата на системата, базирано на JMX

Надзор използва допълнително поставен интерфейс с множество възможни реализации за наблюдение на група потребител. Такова приложение, базирано на JMX, има по-добри



000.56

характеристики на ефективност в сравнение с други реализации. За целта в core-site трябва да се зададе:

```
hadoop.security.group.mapping=org.apache.hadoop.security.  
MappingWithFallback
```

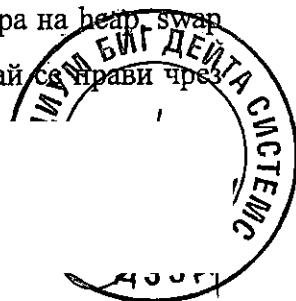
6.2 *Повишаване производителността на Hadoop кълъстера*

Един от важните критерии за създаване на Hadoop кълъстер за фирми е ефективната работа на самия кълъстер, обработващ огромни количества структурирани и неструктурирани данни. Един от параметрите за постигане на такава висока производителност е конфигурирането на хардуерните компоненти на кълъстера, водещо до най-пълно използване на изчислителните ресурси и капацитета за данни. Изхождайки от архитектурата на Hadoop, този вид кълъстер е хоризонтално мащабируем, т.е чрез добавяне на нови сървъри в кълъстера. Това обаче води до увеличаване на финансовите разходи и до допълнителни оперативни разходи за управление и поддържане.

В предлагания метод се повишава производителността на Hadoop кълъстера не чрез допълнителни финансови и хардуерни ресурси, а чрез допълнителна настройка и конфигуриране.

а) Настройка на паметта на сървърите за работа на задачи

Най-важната стъпка, за да се осигури максимална производителност на приложна задачата в Hadoop е да се настроят най-добрите конфигурационни параметри за паметта, като се следи използването на паметта на сървъра. В Hadoop кълъстера се използват различни опции за конфигуриране работата на паметта, диска, процесора и мрежата, които помагат да се оптимизира производителността. Всяка задача в Hadoop събира информация за различните записи на входните записи, броя на записите, които са настроени за по-нататъшно изпълнение, броя на рекордерите, размера на heap, swap на паметта и т.н. Повишаването на производителността в общия случай се прави чрез оптимизирайте използването на паметта.



5

Най-приложимото правило на предлагания метод при настройването на паметта гарантира, че задачите не се разменят по време на изпълнение на приложени задачи да бъде в паметта и коя да се прехвърли върху диска), поради недостатъчно място в оперативната памет. Такава размяна на задачи може да се наблюдава и управлява чрез софтуер като Ganglia, Nagios или Cloudera Manager. Всеки път, когато има прекомерно използване на такава размяна на задачи, използването на паметта трябва да се оптимизира чрез конфигуриране на собствеността `mapred.child.java.opts`, като се намали размерът на паметта, която е разпределена за всяка задача в `mapred.child.java.opts`. Паметта за задачата може да се задаване на `mapred.child.java.opts` на `-Xmx2048M` в файла `mapred-site.xml` с директива:

`mapred.child.java.opts`

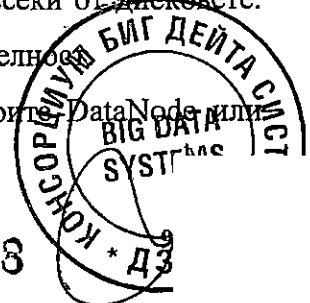
`-Xms1024M -Xmx2048M`

b) Подобряване на Входно-изходната производителност

Ето някои ключови моменти, които трябва да се следват, за да оптимизирате ефективността на приложение, която гарантира, че конфигурацията на клъстера Hadoop е настроена:

- Linux OS има контролен пункт за всеки файл, включително контролна сума, последно време за достъп, време за създаване, потребител, който е създал файла и т.н. За да постигнете по-добра производителност на Входно-изходната производителност, тази контролна точка трябва да бъде деактивирана в HDFS, понеже HDFS е организирана на принципа пиши веднъж – чети многократно.
- Монтирането на устройства за DataNode или директориите с данни трябва да бъдат конфигурирани с опцията `Noatime`, за да се гарантира, че метаданните не се обновяват от NameNode всеки път, когато се осъществява достъп до данните. Монтажът за памет MapReduce и DFS, когато е монтиран с опция `noatime`, деактивира проследяването на времето за достъп и това подобрява производителността на IO.
- Двата конфигурационни параметри `'mapreduce.local.dir'` и `'dfs.data.dir'` трябва да бъдат настроени така, че да сочат към една директория на всеки от дисковете. Това помага за подобряване на Входно-изходната производителност.
- Препоръчва се да не се използват LVM и RAID на сървърите DataNode или TaskTracker, тъй като това намаляват производителността.

000.58



- c) Намаляване на разточителното използване на диска чрез компресиране на изхода на Map частта от приложението

Входно-изходните операции на диска (IO) е един от най-големите ограничения в производителността на Hadoop кълстера и поради което в предлагания метод се предлагат два начина:

- Трябва разпределението за задачата MapReduce използва 70% от паметта за натрупване на буфера.
- Да се компресира изхода на Map частта от приложението

- d) Настройване на броя задачи за разпределение

Всяка Map или Reduce задача обикновено отнема 40 секунди, за да завърши изпълнението. Също така, когато има голяма работа, която да бъде изпълнена, тя не прави възможно най-доброто използване на всички сървъри в наличния кълстер. По този начин е изключително важно да се настрои броя на Map/Reduce задачите, като се прилагат следните техники:

- Ако задачата Map/Reduce има повече от 1 терабайт вход, тогава, за да е сигурно, че броят на задачите е по-малък, то размерът на блока на входния набор от данни трябва да бъде увеличен до 512MB или 256MB. Размерът на блока на съществуващите файлове може да бъде променен чрез конфигуриране на `dfs.block.size`. След като командата за промяна на размера на блока се изпълни, оригиналните данни могат да бъдат премахнати.
- Ако задачата Map/Reduce на хаотичен кълстер стартира няколко Map задачи, при които всяка задача завършва само за няколко секунди, то намаляването на броя на Map задачите, стартирани за това приложение, няма да се отрази на конфигурацията на Hadoop кълстера, а същевременно ще помогне за оптимизиране на производителността на кълстера.
- Установяването и насрочването на задачи изисква времеви разходи. Ако задачата отнема по-малко от 30 секунди за изпълнение, тогава е по-добре да се намали броят на задачите. Повторното използване на JVM е добра алтернатива на този проблем.



000259

7 Хардуерна Референтна архитектура на Hadoop

Референтната архитектура по принцип предлага типово решение (шаблон) за архитектура за конкретна област. Тя също така предоставя общ речник, с който да се обсъждат възможни реализации, често с цел да се подчертае общият характер. Референтната архитектура съдържа структурите на съответните елементи и взаимовръзки между тях. Референтните архитектури могат да бъдат дефинирани на различни нива на абстракция. Високото ниво на абстракция показва различни части на оборудването на комуникационна мрежа, всяка от които осигурява различни функции. По-ниско ниво на абстракция може да демонстрира взаимодействия в рамките на компютърна система. Референтната архитектура представлява обобщение на набор от решения.

Представената Хардуерна Референтна архитектура не е насочена към High Performance Computing (HPC) системи, а към система за обработка на големи данни с общо предназначение, т.е. „Hadoop система с общо предназначение“. Hadoop изглежда че прави същото нещо, което HPC прави – обработва много задачи в паралел. Но Hadoop прави това по принципно различен начин. В случая на HPC, съществува една задача, която се дава на много процесори да я изпълняват, докато при функциониране на Hadoop, първо задачата се разбива на под-задачи, всяка под-задача се дава на отделен сървър за изпълнение получавайки междинни резултати, и след това на други сървъри се изпълняват по-задачи-обединение на получените междинни резултати. От чисто хардуерна гледна точка, Hadoop системата е предназначена да оперира основно със сървъри с общо предназначение - работни сървъри (представени в Hadoop като DataNode), даже и със сървъри с не достатъчно голяма надеждност. Съществуват обаче и хардуерни системи за Hadoop, при която работните сървъри са разделени на два типа – такива за изчисления и такива за съхранение на данни. Последния тип хардуерни системи за Hadoop ги разглеждаме като изключение и не са фокус на настоящата разработка. Предложената Хардуерната Референтна архитектура ще използва работни сървъри с общо предназначение с x86 архитектура, в които се оперира както процесорите и техните ядра за изчисления, така и се осъществява управление



големи дискови масиви. Ето защо наричаме такава система „Hadoop система с предназначение“.

Референтната архитектура на Hadoop ще бъде представена на 3 функционални нива: Прототипна Референтна архитектура, Референтна архитектура за разработка на приложения и Производствена Референтна архитектура.

7.1 Избор на управляваща система за Референтна архитектура на Hadoop

Референтна архитектура на Hadoop се свързва основно с двете форми на индустриално управляващи системи – тази на Cloudera и тази на Hortonworks. Като по-малко популярни управляващи системи за Hadoop са тези на фирмите IBM Technologies, Amazon Web Services, Oracle.

Както Hortonworks, така и Cloudera са изградени върху едно и също ядро на Hadoop. Следователно и двете управляващи системи имат повече прилики, отколкото разлики, които се свеждат до:

- И двете предлагат дистрибуции на Hadoop готови за прилагане.
- Опакования софтуер осигурява сигурност и стабилност.
- Както Cloudera, така и Hortonworks предоставят платени услуги за обучение и поддръжка, за да запознаят специалистите, които са нови в домейна Big Data и Analytics.
- И двете са създали общности, които активно участват и помагат с проблемите, пред които са изправени, както и с необходимите демонстрации.
- И двете имат обща компютърна рамка.
- И двете разпределения имат архитектура master-slave.
- И двамата доставчици поддържат MapReduce и YARN.

Независимо от многото прилики и едно и също ядро, Cloudera и Hortonworks проявяват няколко разлики, които могат да се обобщят в:

- Cloudera и Hortonworks имат диаметрално противоположни стратегии. Cloudera продава търговски софтуер в допълнение към



разпространението си с отворен код Hadoop, докато Hortonworks предлага сертифициран софтуер на Apache Foundation.

- Стратегията за растеж на бизнеса на Hortonworks се фокусира върху вграждането на Hadoop в съществуващите платформи за данни, докато Cloudera използва подхода на традиционния доставчик на софтуер, който печели от продажбите на продукти и се конкурира с други доставчици на софтуер.
- Cloudera е по-насочена към Linux операционни системи, докато Hortonworks е по-насочена към Windows операционни системи.
- Hortonworks не идва с никакъв собствен софтуер, поради което използва за управление, Stinger за обработка на заявки и Apache Solr за търсене. Въпреки това, Cloudera притежава собствен софтуер за управление Cloudera Manager, Cloudera Search за достъп в реално време на продукти и Impala, интерфейс за обработка на SQL заявки.
- Cloudera притежава търговски лиценз, докато Hortonworks притежава лиценз с отворен код.
- Най-важното е, че Hortonworks е напълно безплатна и Cloudera осигурява платени услуги, въпреки че има и безплатни версии, като има безплатни производствени версии за университети и научно-изследователски институции.

От посоченото следва, че Cloudera и Hortonworks имат свои собствени силни и слаби страни. Поради това, при избора е необходимо да се избере основната добавена стойност за съответната цел на изграждане. Също така трябва да се прецени ефективността, мащабируемостта, управляемостта, надеждността и достъпа до данни за двете възможни управляващи системи, като се имат предвид краткосрочните и дългосрочните цели. Трябва да се акцентира и върху техническата поддръжка, разширената функционалност и надеждността на системата или нуждата от гъвкавост или бързо въздействие и цялостна рентабилност. Както Cloudera, така и Hortonworks са лидери в света на Hadoop. Въпреки че Cloudera е по-старият играч на пазара, Hortonworks бързо настига броя световни потребители.

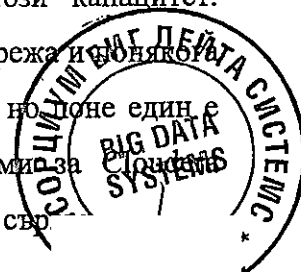
Изхождайки от по-горните съображения, то по-промишлената ориентирана управляваща система е тази на Cloudera. Такова е и мнението на световния анализатор Gartner Group – [41]. Ето защо предлаганата Референтна архитектура е насочена към използване на управляващата система на Cloudera. Системата Cloudera е с отворен код.



7.2 Типове сървъри използвани в Референтната архитектура

В предложената Референтна архитектура се използват следните типове сървъри:

- a) **Главен сървър** – това е сървърът в който се зареждат софтуерни компоненти на Hadoop за ръководене работата на Hadoop системата като цяло. Това са обикновено: NameNode, Standby NameNode, YARN Resource Manager, ZooKeeper, History Server, Journal Node, HBase Master daemon, Sentry server, Impala StateStore Server и Catalog Server, като NameNode, Standby NameNode, YARN Resource Manager, History Server са основни при предлаганата Референтна архитектура;
- b) **Работен сървър** – това е сървър в който се изпълняват основните дейности на Hadoop системата. В него се зареждат обикновено един или няколко от следните софтуерни компоненти: HDFS DataNode, YARN NodeManager, HBase RegionServer, Impala impalad, Search worker daemons and Kudu Tablet Servers, като HDFS DataNode е основният на предлаганата Референтна архитектура;
- c) **Управляващ сървър** – това е сървърът в който се зареждат софтуерните управляващи компоненти на Cloudera софтуера за управление на целия Hadoop клъстер. Това са обикновено : Cloudera Manager, Cloudera Management Services, MySQL (или друга база от данни), Hive, Sentry, като Cloudera Manager, Cloudera Management Services и MySQL са основни за предлаганата Референтна архитектура.
- d) **Клиентски сървър** – това е сървърът в който се зареждат за изпълнение потребителски програми, потребителски конфигурации и услуги, включително и gateway конфигурации за такива като например: HDFS, YARN, Impala, Hive, and HBase. Този сървър е препоръчително място също и за Hue, Oozie, HiveServer2, Impala HAProxy, gateway към външни приложения, такива Business Intelligence (BI) приложения.
- e) **Граничен сървър** - осигурява интерфейс между данните и капацитета за обработка, налични в Hadoop клъстер и потребител на този капацитет. Граничният сървър има допълнителна връзка към Граничната мрежа и понякога се нарича "gateway node." Граничният възел е незадължителен, но поне един е силно препоръчително. Оперативните бази данни, необходими за Cloudera Manager и допълнителни метасторите се разполагат в Граничния сър.



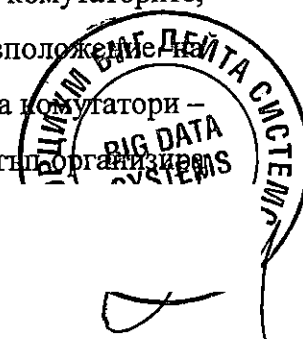
000-53

дно разположение на компонентите на Cloudera по отделните сървъри може

Тип сървър	Компоненти на
Граничен сървър	Cloudera Manager Operating System Provisic Yum Repositories Operational Databases (PostgreSQL) DMExpress Service (dmxd)
Активен Главен сървър	NameNode Resource Manager Zookeeper Quorum Journal Node HMaster Impala State Store and Catalog Daemons
Пасивен Главен сървър	Standby NameNode Standby Resource Manager Zookeeper Quorum Journal Node
Втори пасивен Главен сървър	Zookeeper Quorum Journal Node
Работен сървър	DataNode NodeManager RegionServer ImpalaDaemon DMX-h

7.3 Инфраструктурна архитектура на Референтната архитектура

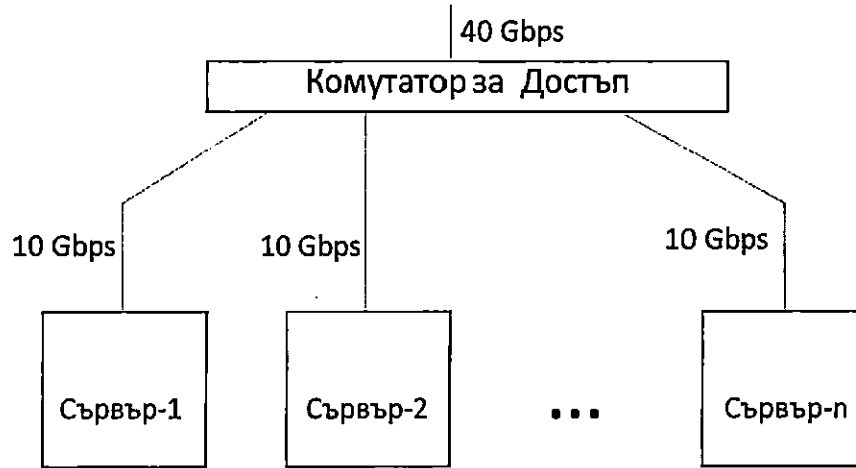
Инфраструктурната архитектура включва мрежовите компоненти – комутаторите, връзките между сървърите и комутаторите, както и физическото разположение на сървърите и комутаторите по Ракове (шкафове). В нея се използват 2 вида комутатори – Комутатор за достъп и Разпределителен комутатор. Комутаторът за достъп



000264

Handwritten mark resembling a stylized '2' or '7'.

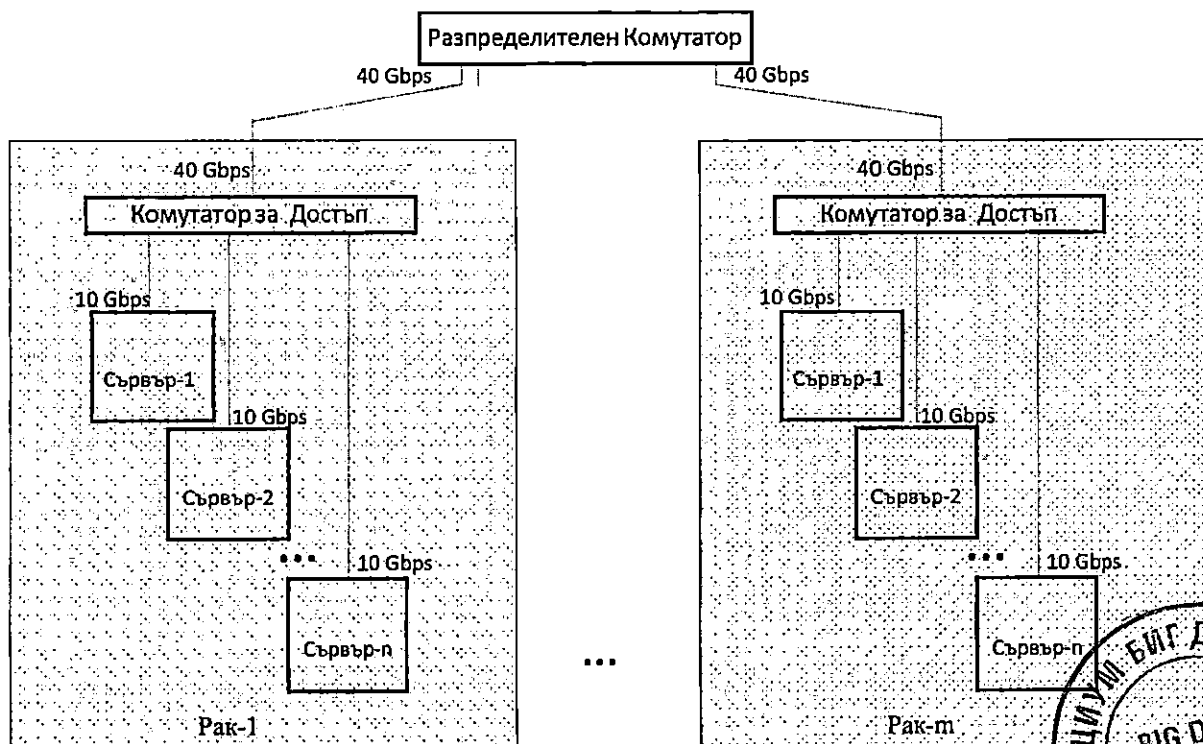
достъпа до отделния сървър на Хардуерната Референтна архитектура. Това е комутатор с n броя портове, всеки със скорост за обмен минимум 10 Gbps, и един или няколко връзки с минимум 40 Gbps – фигура 7-1.



Фиг. 7-1

В рамките на един Рак се предвижда 1 брой Комутатор за достъп.

Разпределителният комутатор има за цел да осигури мрежовата връзка между отделни ракове на Надоор системата. Той трябва да се състои от портове, всеки с минимум 40 Gbps – фигура 7-2.



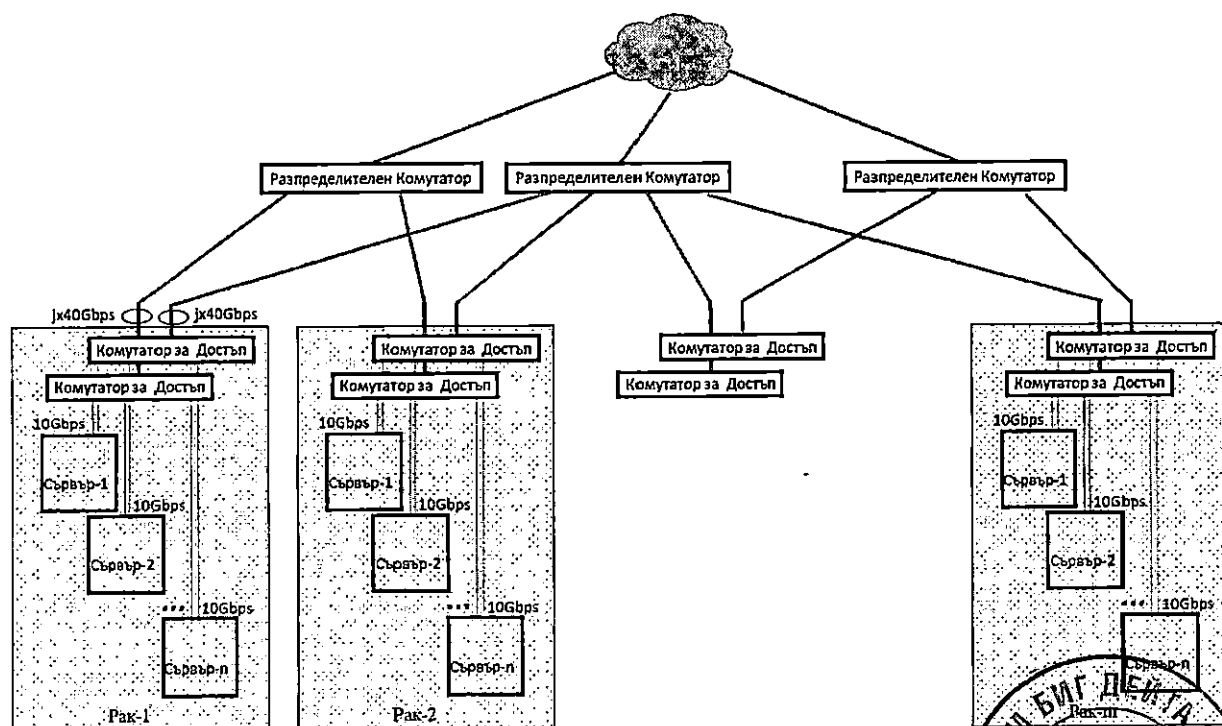
000265

Фиг. 7-2

Разпределителния комутатор се разполага в отделен Рак – Комуникационен рак, и някой от раковете, където са сървъри на Hadoop системата.

Връзката Комутатор за достъп – Разпределителен комутатор трябва да се състои от обединение на един или няколко изходящи 40 Gbps линии. Принципът е съотношението на сумарния капацитет на отделните портове към изходящия капацитет на комутатора не трябва да надвишава съотношението 4:1. Например при 24 порта на Комутатора за достъп, сумарният капацитет на отделните портове е $24 \times 10 \text{ Gbps} = 240 \text{ Gbps}$, което изисква 2 изходящи линии по 40 Gbps (общо 80 Gbps) да се обединят за връзка към Разпределителния комутатор ($240 \text{ Gbps} : 80 \text{ Gbps} \Rightarrow 3:1$). Подобен е примерът при Комутатор за достъп с 48 порта, при който сумарният капацитет на отделните портове е $48 \times 10 \text{ Gbps} = 480 \text{ Gbps}$, което изисква 4 изходящи линии по 40 Gbps (общо 160 Gbps) да се обединят за връзка към Разпределителния комутатор ($480 \text{ Gbps} : 160 \text{ Gbps} \Rightarrow 4:1$).

При изграждане на Hadoop система с капацитет стотици и хиляди Petabytes, се използват няколко рака, формиращи архитектура, показана на фигура 7-3.

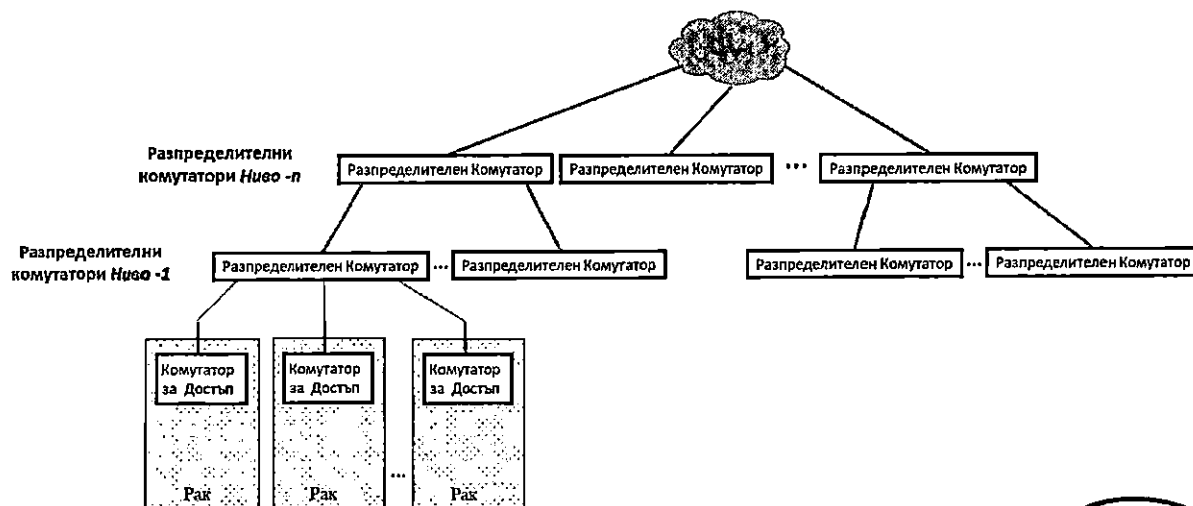


Фиг. 7-3

00016

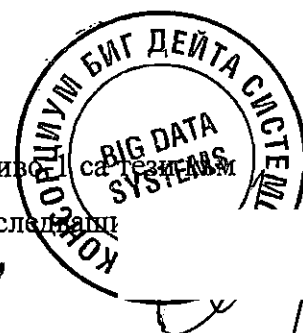
Тази архитектура е възможно да бъде изградена като Архитектура с излишество. При нея Комутаторите за достъп са по 2 броя в рак, като всеки сървър в рака е свързан към двата комутатора. От друга страна, Разпределителните комутатори са също в излишество и са няколко на брой, така че от един рак се прави връзка към 2 броя Разпределителни комутатори – единият Комутатор за достъп в рака е свързан към един Разпределителен комутатор, а вторият Комутатор за достъп в рака е свързан към друг Разпределителен комутатор. При Архитектурата с излишество захранването на отделните сървъри е двойно и оперативната памет е също с излишество. Архитектурата с излишество не използва излишество на дисковете, поради концептуалната характеристика на Hadoop един данни да имат няколко копия или да се разположени чрез т.н. схема „Кодиране срещу изтриване“ разгледани по-горе в този материал. По този начин се формира т.н. „високо ниво на излишество“. От това ниво може да се отива към по-ниско ниво на излишество, според изискванията на създаваната Hadoop система. Когато се изисква намаляване на нивото на излишество, то архитектурните решения се движат в следната посока: премахване излишеството на Комутаторите за достъп, премахване излишеството в сървърите, премахване излишеството на Разпределителните комутатори.

При изграждане на Hadoop система с капацитет стотици и хиляди PetaBytes, се използват десетки и стотици ракове, което налага създаването на Йерархия от Разпределителни комутатори – фигура 7-4.



Фиг. 7-4

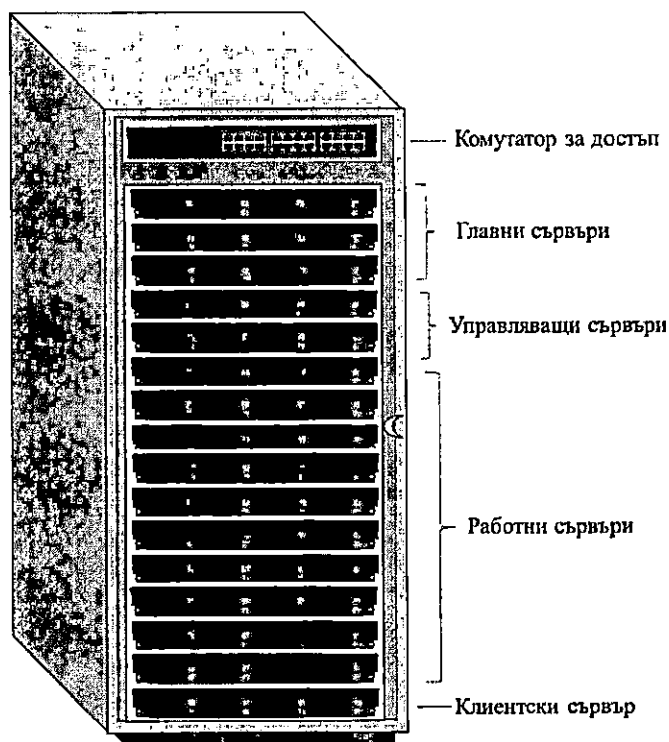
При това архитектурно решение, Разпределителните комутатори на Ниво n са свързани с интернет, а Разпределителните комутатори от следващи нива са свързани с раковете, а Разпределителните комутатори от следващи нива са свързани с раковете, а Разпределителните комутатори от следващи нива са свързани с раковете.



000267

осигуряват свързаност между Разпределителните протоколи. И при това архитектурно решение е възможно да се въведе Архитектура с излишество, която да бъде с различна на излишество, както е посочено по-горе. При организиране на Йерархията от изпределени комутатори трябва да се спазва посочения по-горе принцип за отношение до 4:1 на капацитета на отделните портове от един сървър към капацитета на изходящите линии на комутатори. При тази архитектурна йерархия може да се използват комутатори с портове с капацитет по 40 Gbps, но тогава фокусът на архитектурното решение ще се насочи към обединяването на броя връзки и капацитетите на отделните портове.

Примерно физическо разположението на инфраструктурните компоненти в един рак е показано на фигура 7-5.



Фиг. 7-5

Броят на сървърите в един рак зависи от типа Хардуерна Референтна архитектура, като при по-малките Nadoop системи е възможно ограничен брой Работни сървъри, Главните и Управляващи сървъри да са по един брой, както и съвместяване на Управляващ и Клиентски сървър в един хардуерен сървър.



000.68

7.4 Прототипна Хардуерна Референтна архитектура на Hadoop

Прототипната Хардуерна Референтна архитектура на Hadoop се състои от един сървър (Single Node Cluster). Не са необходими допълнителни устройства или специални мрежови системи, нито е необходимо да се прави конфигурация на Hadoop системата, освен да се зададе JAVA_HOME променливата. Предназначението на Прототипната Хардуерна Референтна архитектура е да може върху нея да се изгради напълно функционираща Hadoop система, с оглед запознаване с функциите на опериране на компонентите на Hadoop система. За всяка отделна компонента на Hadoop е необходимо да има съответните хардуерни ресурси за да се изпълнява отделна JVM машина. Производствената Референтна архитектура се състои от 10-ки, 100-ци или 1000-ди сървъри, докато при Прототипната Хардуерна Референтна архитектура отделните компоненти на Hadoop, такива като NameNode и DataNode използват части от един сървър, обикновено по едно ядро от многоядрения процесор(и) на този единичен сървър, в което се изпълнява по едно копие от NameNode, DataNode TaskTracker, JobTracker, ResourceManager и NodeManager.

Необходимите хардуерни ресурси за Прототипната Хардуерна Референтна архитектура на Hadoop са: 6-8 ядра в един или няколко процесора, 16 GB RAM и дискова памет в размера на 1 TB. Възможно е да се използва Прототипната Хардуерна Референтна архитектура с по-малко от посочените ресурси, но тогава прототипното функциониране ще работи доста ограничено.

Прототипната Референтна архитектура няма специфична Инфраструктурна архитектура, защото се състои от 1 сървър.

7.5 Референтна архитектура на Hadoop за разработка на приложения

Референтна архитектура на Hadoop за разработка на приложения служи за разработване и тестване на приложения за работа с Hadoop, като скоростта на работа – време за изпълнение на съответното приложение, е ограничена. Количеството сървъри за ManagementNode, NameNode, и DataNode са ограничени – няколко на брой (общо под 10), както и дисковото пространство на DataNode е също ограничено. Тази Референтна архитектура не може да се очаква развито ниво на надеждност, така както е при всички разработващи системи.



000.09

Инфраструктурната архитектура на Референтна архитектура на Nadoor за разработка на приложения е от типа на тази представена на фигура 7-1.

Диапазона на брой сървъри в Референтната архитектура на Nadoor за разработка на приложения варира според броя разработчици едновременно работещи, като тази Референтна архитектура може да се класифицира като „много малка“, използваща по-малко от 10 сървъра общо, от които:

- 1 брой Главен сървър,
- 1 брой сървър съчетаващ Управляващ сървър и Клиентски сървър,
- 2-8 броя Работни сървъри.

7.6 Производствена Референтна архитектура на Nadoor

Производствената Референтна архитектура на Nadoor се разглежда в 3 модификации – Малка архитектура, Средна архитектура и Голяма архитектура.

a) *Малката Производствената Референтна архитектура* на Nadoor се състои от поне 20 сървъра общо:

- 2 броя Главен сървър;
- 20-50 броя Работни сървъра;
- 1 брой сървър съчетаващ Управляващ сървър и Клиентски сървър.

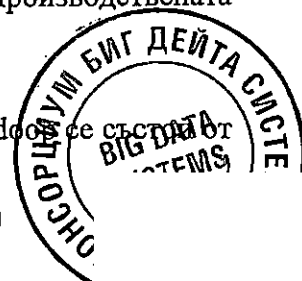
b) *Средната Производствената Референтна архитектура* на Nadoor се състои от поне 200-250 сървъра общо:

- 3 броя Главен сървър;
- 200-300 броя Работни сървъра;
- 2 броя Управляващ сървър;
- 1-2 броя Клиентски сървър.

Множеството Nadoor системи в света използват 10-50 Работни сървъра [46], което изисква фокусиране върху такъв тип Nadoor решения със Средната Производствената Референтна архитектура.

c) *Голямата Производствената Референтна архитектура* на Nadoor се състои от повече от 500 сървъра общо:

000.70



- 5 броя Главен сървър;
- Повече от 500 броя Работни сървъра;
- 2 броя Управляващ сървър
- 2 броя Клиентски сървър.

За всяка от 3-те модификации на Производствената Референтна архитектура на Hadoop е възможно да се използва Криптираща инфраструктура, за която се изискват:

- 2 броя сървъра за Управление на ключовете (Key Management server);
- 2 броя сървъри за Сървър на доверието (Trustee server).

Главният сървър е препоръчително да има следната дискова конфигурация:

- 2 х диска в RAID1 за операционната система и логовете;
- 4 х диска в RAID10 за данните. От гледна точка на бързодействие SSD дискове са за предпочитане, но от гледна точка на цена – HDD са за предпочитане.

Работният сървър е препоръчително да има следната дискова конфигурация:

- 2 х диска в RAID1 за операционната система и логовете;
- Дискове в режим JBOD или дискове в RAID0. От гледна точка на бързодействие SSD дискове са за предпочитане, но от гледна точка на цена – HDD са за предпочитане. Освен това, съществуват съвременни SSD дискове с много по-голям капацитет отколкото HDD дискове. Така например SSD с капацитет 60TB е едно добро техническо решение от гледна точка на бързодействие и обслужване поради по-малко устройства за наблюдение и управление, но от гледна точка на цена не - 60TB SSD диск има средна цена \$337/TB, докато 8TB HDD е \$37/TB.

Минималното количество процесорни ядра за Главния и Работния сървър са 4 броя, като за Работния сървър за активно работеща Hadoop система е добре да е значително повече.

За да се определи разпределението на хардуера и ресурсите за Hadoop системата трябва да се анализират видовете работни натоварвания, които се изискват и изпълняват в системата, и компонентите на CDH, които ще се използват, но така трябва да се има предвид размера на данните, които да се съхраняват и обр



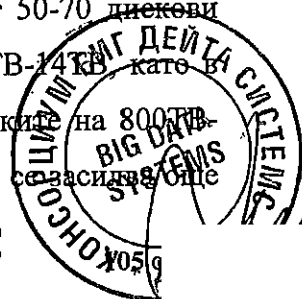
честотата на натоварванията, броя съпътстващи задачи, които трябва да се изпълняват, и скоростта, необходима за изпълнение на приложенията. Също така, при създаване на архитектурата на Hadoop системата, ще трябва да се разпределят ролите на Cloudera Manager и CDH според вида Референтна архитектура (малка, средна, голяма), за да увеличите максимално използването на ресурсите. Cloudera предоставя някои указания как да се определят ресурсите по роли в Hadoop системата.

Относно размера на оперативната памет на Работните сървъри, размер от 128 GB RAM е добър, защото обикновено ограничава забавянето което би се получило от разпределени услуги като YARN и Impala. Но препоръчителна стойност от 256 GB RAM удовлетворява различни приложения с необходимото задоволително бързодействие.

7.7 Архитектурни решения за преодоляване на практическата препоръка за 100 TB ограничение на DataNode при Cloudera-Hadoop

Съществува препоръка на Cloudera за Работен сървър при Cloudera-Hadoop система [46], според която Работният сървър не трябва да има повече от 100TB дискова памет, като този сървър е свързан в Hadoop системата чрез мрежов адаптер за предаване на данни със скорост 10Gbps. Това е свързано с практически ограничения като време за репликация (при стандартни приложения или при загуба на данни) и отчитане на натоварване (при операции с ниска латентност). Това е практическа препоръка свързано с времето за обмен на големи масиви от данни с Работен сървър, т.е. колко време е необходимо да се обработи файл с обем 1 TB от един Работен сървър при мрежова връзка на сървъра 10Gbps. По този начин е въведено ограничение, разпространено сред Cloudera обществото, че Работния сървър не трябва да има повече от 100TB дискова памет. По принцип, в програмния код на Hadoop няма такова ограничение за дисков капацитет в Работен сървър, но при репликиране на блоковете на един файл се наблюдава ролята на по-големия дисков капацитет върху намаляване производителността на процеса на репликиране. Това е проблем при големи Работни сървъри (например сървъри с размер 4U и много на брой дискове – повече от препоръчаните от Cloudera 24 броя), в които може да се разположат 50-70 дискови устройства, които при сегашни размери на твърди дискове от по 10TB-14TB като в този случай общия обем на дисков капацитет на сървъра става в рамките на 800TB-900TB – сериозно различен капацитет от този на 100TB. Този проблем се засилва още

000.72



повече с наличието на 60TB SSD дискове, където в един голям Работен сървър могло да се разположи сумарен дисков капацитет от няколко PB.

Посочената препоръка от Cloudera да не бъде дисковото пространство в Работен сървър по-голямо от 100TB се отнася основно за работа с репликиране (свойство на Hadoop 2), а не с Кодирането срещу изтриване (Erasure Coding). Hadoop е приела Кодирането срещу изтриване от версия 3.0, която Apache Software Foundation публикува през декември 2017 г. Това отбелязва значителна промяна за Hadoop Distributed File System (HDFS), която до този момент винаги е направила три копия на данни за защита загуба на данни. В сравнение с тази 3-кратно репликация на данни, Кодирането срещу изтриване води до фактор на репликация на данни от около 1,5X, което се превръща в нетен 50% увеличение на ефективността на съхранение. Докато Кодирането срещу изтриване увеличава използваемостта на дисковото пространство и елиминира репликирането, това натоварва повече мрежата, тъй като всички дискови прочитания се движат по мрежата. Съществуват и допълнителни разходи за CPU, свързан с Кодиране срещу изтриване. Това е голяма промяна от традиционната схема за тройна репликация (прилагана до версия 2 на Hadoop), при която важи концепцията за "местоположение на данните", за да се избегне необходимостта HDFS да извлича данни по мрежата.

7.7.1 Решение чрез използване на малки, 1U сървъри

При сървъри с размери от 1U броят на възможните работни дискове е обикновено 2-10 и тогава ограничението от 100TB на сървър се явява естествено практическо и физическо конфигуриране на Работен сървър.

7.7.2 Решение при прилагане на големи, 4U сървъри

Използването на Кодирането срещу изтриване (Erasure Coding) вместо тройно репликиране на данни е нов подход в прилагане на Hadoop, който практически стана възможен от началото на 2018 г. Кодирането срещу изтриване силно намалява изискванията посочени от Cloudera за максимален обем на памет на сървър от 100TB, но поставя допълнителни изисквания за увеличен мрежов капацитет между отделните сървъри, както и увеличаване на използваната изчислителна мощност на отделните сървъри. Това налага мрежовата пропускна възможност на такива големи сървъри да не е само 10Gbps, но няколко пъти по 10Gbps. Относно изискването за увеличаване

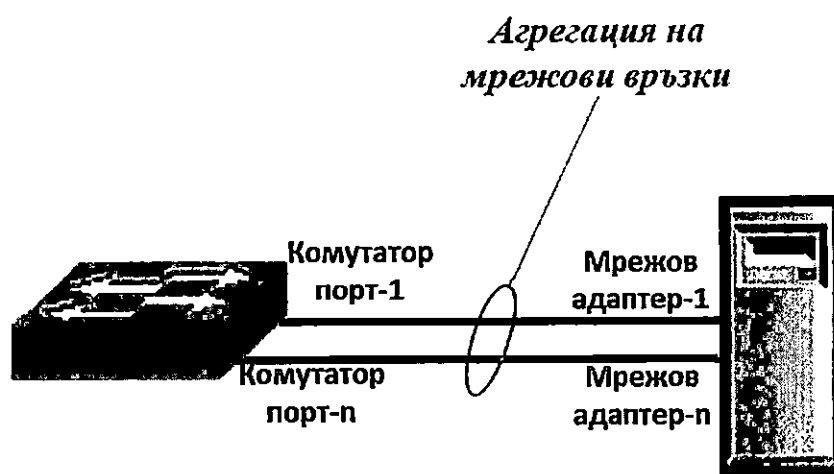


000.73

изчислителна мощност – се изисква да се използва в сървър поне 2 броя CPU с максимален брой ядра и максимална честота на работа.

Внедряването на подхода Кодирането срещу изтриване (Erasure Coding) вместо тройно пликване на данни е път за премахване на ограничението от 100TB дисково пространство на Работен сървър, като в замяна се изисква увеличаване на мрежовия капацитет на предаване на данни на Работния сървър и увеличаване броя и ядрата на CPU на Работния сървър. Съчетаването на изчислителна мощност и дисково пространство в един сървър е основа пък за по-голяма производителност на самата Hadoop система при нормална работа на системата.

При големи Работни сървъри (например сървъри с размер 4U и много дискове), в които може да се разположат 50-70 дискови устройства, трябва да се приложи т.н. подход на "агрегация на мрежови връзки" или "Network teaming". Терминът "агрегация на мрежови връзки" е термин, който описва комбинирането (или обединяването) на няколко мрежови връзки, създадени от няколко мрежови адаптера, за да се осигури логически мрежов слой от почти сумарен мрежов капацитет. Този подход първоначално бе въведен за сървъри с операционна система Linux, но сега се използва и за операционна система Windows. Графическо представяне на подхода „агрегация на мрежови връзки“ е представен на фигура 7-6.



Фиг. 7-6

За да се осъществи агрегацията на мрежови връзки е необходимо да се извърши определено конфигуриране както в сървъра, където са поставени n-брой мрежови



000.74

адаптери, така и конфигуриране на комутатора, към който са свързани съответните мрежови адаптери. Така се получава Агрегация на мрежовите връзки, чиито логически мрежов капацитет е близък до сумата от капацитетите на отделните мрежови адаптери. Така например, при 700ТВ общ дисков капацитет на сървър, то този сървър трябва да се свърже с 8 мрежови адаптера, всеки с по 10Gbps, за да се получи еквивалентна компютърна среда на 7 броя виртуални сървъра, всеки с по около 100ТВ виртуален капацитет, като всеки сървър е свързан с мрежов адаптер с виртуална скорост за предаване на данни около 10Gbps.

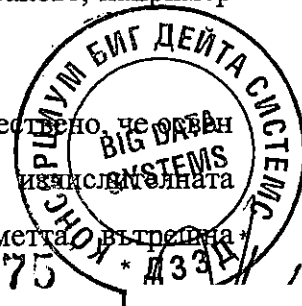
Използването на големи сървъри с няколко стотици ТВ дисково пространство, с прилагане на Кодирание срещу изтриване е удачен подход при създаване на не много големи Hadoop системи. Изследванията до момента не могат да дадат точни оценки, но въвеждането на Агрегация на мрежовите връзки при скорост на обмен на една връзка 40Gbps е основа за разширяване на предложения подход и към по-големи Hadoop системи.

7.8 Ключови индикатори за хардуерна оценка на Работен сървър

Чрез няколко ключови индикатори може да се специфицира и избере хардуера на Работните сървъри използвани в Hadoop система.

Съотношение на изчислителна мощност към дисков капацитет на сървър в еквивалентно 1U пространство в рак за хардуерна инфраструктура на компютърни сървъри. Ракът е стандартна физическа структура в която се разполагат инфраструктурни елементи като сървъри, комутатори, рутери и пр., според стандартната спецификация EIA-310, адаптирана от IEC 60297 Mechanical structures for electronic equipment. Размерът на рака се определя от брой „U“ единици (раков Unit) устройства които могат да се поберат в него, например максимално големият рак е 42U, т.е. за устройства чиято сумарна височина е до 42U. Физическите параметри на сървърите които са за разполагане в рак (rack mountable) се измерва в „U“, например, сървърите са обикновено 1U, 2U, 4U. Съществуват също и по-малки ракове, например 1.5U или 0.5U и пр.

Изчислителната мощност ще се измерва в брой ядра на процесори. Естествено, че броят на броя ядра на процесор и други технически параметри влияят на изчислителната мощност на процесора: честота на работа, размер на кеш паметта, вътрешна



архитектура, система команди, за какви типове изчислителни програми е преднази процесора, за какъв тип многоядрена работа е предназначен процесора и Изхождайки от обстоятелството, че предлаганата

8 Архитектура на големи Hadoop клъстери с обеми стотици и хиляди PB

8.1 Архитектурен подход към създаването на DataNode сървър

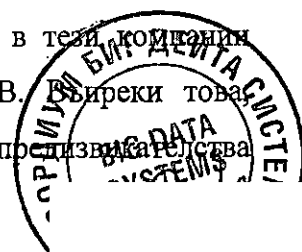
Съществуващите ограничения за физически параметри на DataNode сървър от Cloudera версия 6.0 – 100TB дисково пространство на сървър и 12 дискови устройства, водят до определяне на „балансиран“ DataNode да бъде 1U сървър с 12 дискови устройства, в които има максимално 10TB-12TB, което прави 120TB-144TB на сървър. Такъв сървър би трябвало да има минимум 2 броя процесори от тип Intel Xeon Scalable Processors, всеки от тях с минимум 8 ядра, работещи с честота минимум 1,7GHz-2GHz и използващи оперативна памет минимум 128 GB. Такива архитектурни параметри на сървър са допустими даже и в края на 2018 г. Тогава, за обем памет на Hadoop клъстера от около 125 PB ще са необходими около 1000 DataNode сървъра, а за 250PB клъстерен обем памет ще са необходими около 2000 DataNode сървъра.

8.2 Физическа архитектура на Hadoop клъстери с обеми стотици и хиляди PB

Използвайки 42U ракове, означава че в един рак ако разположим 40 DataNode сървъри от по 1U (другите 2U of 42U ги заделяме за комутатори), то обемът на един рак в дисков капацитет би бил 4,8PB-5,76PB. При тези условия, за да изградим 1000PB Hadoop клъстер са необходими 209 рака при дискове 10TB или 174 при дискове 12TB. Количество ракове 209 е еквивалентно на 8360 DataNode сървъри с 10TB дискове, а количество ракове 174 е еквивалентно на 6900 DataNode сървъри с 12TB дискове.

Очаква се през 2020 година [48] 75% от компаниите от Fortune 2000 да работят с 1000 DataNode сървъра в клъстер в производството. Това означава, че в тези компании трябва да има Hadoop клъстери с обеми около 125PB – 140PB. Въпреки това внедряването на Hadoop в производството все още е придружено от

000.76



за разгръщане и управление, като скалируемост, гъвкавост и ефективност на разход
При 1000 броя DataNode сървъри с капацитет 144ТВ на сървър, означава, че те
създават Hadoop клъстер с обем 125PB-140PB, т.е. в края на 2020 г. ще се изисква
ма търговски продукти за Hadoop клъстер за по 125PB-140PB с по 1000 DataN
сървъра.

Към края на 2016 г. [48] Yahoo разполага с 36 различни Hadoop клъстера, разпредел
в Apache HBase, Storm и YARN - общо 60 000 сървъра, създадени от 100 различни
хардуерни конфигурации, изградени в продължение на поколения. Според Yahoo,
фирмата управлява най-голямата инсталация за наемане на множество наематели в
света с широк набор от случаи на използване, Yahoo работи на 850 000 Hadoop работни
места дневно.

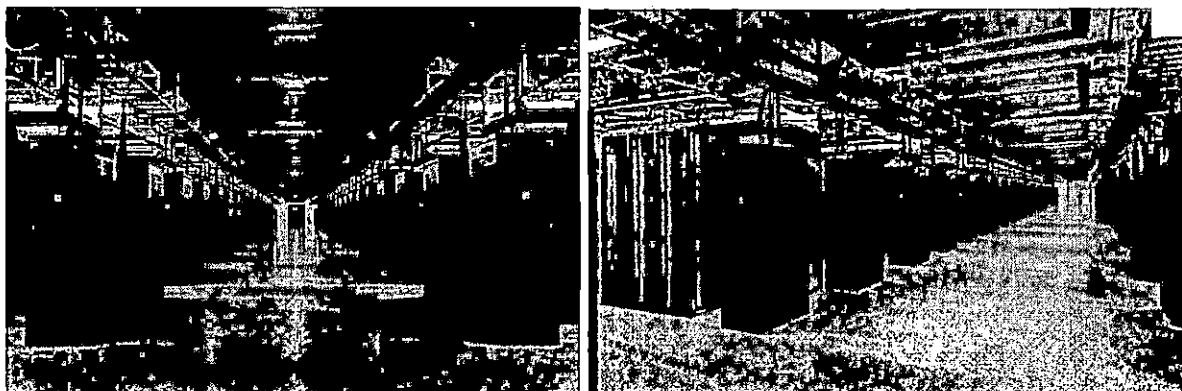
Към края на 2016 г. [48] Facebook (FB) работи с 1,59 милиарда потребители
(приблизително 1/5 от световното население), 30 милиона потребители на FB
актуализират състоянието си най-малко веднъж дневно, 10+ милиона видеоклипа се
качват всеки месец, 1+ милиард парчета съдържание се споделят всяка седмица и
повече от 1 милиард снимки са качени всеки месец, т.е. Facebook оперира с петбайтови
данни. Facebook поддържа най-големия Hadoop Cluster в света с повече от 4000
сървъра, които съхраняват стотици милиони гигабайта данни. По онова време Hadoop
клъстера във Facebook има около 2500 CPU ядра и 1 PB от дисково пространство, а
инженерите във Facebook зареждат повече от 250 GB компресирани данни (повече от 2
TB от некомпресирани данни) в HDFS дневно и има 100 задания работещи ежедневно в
тези набори от данни. FB използва 135 TB от компресирани данни, които са сканирани
ежедневно и 4 TB компресирани данни се добавят ежедневно. Сега FB използва
няколко Hadoop клъстера с по 2500 ядра. FB са прегърнали SQL като позната
парадигма за адресиране и работа с големи набори от данни. Повечето данни,
съхранявани в файловата система на Hadoop, се публикуват като таблици.
Разработчиците могат да изследват схемите и данните на тези таблици, както биха
направили с добра стара база данни. Когато искат да работят с тези набори от данни, те
могат да използват малка подгрупа от SQL, за да укажат необходимия набор от данни.
Аналитичният процес във Фейсбук започва от Склад за анализ на данни с размер на
300PB. За да се отговори на конкретна заявка, данните често се изтеглят от склада и се
поставят в таблица, така че да могат да бъдат проучени. Центърът за данни на FB е
около 27 800 квадратни метра от ракове.

000277



Този анализ показва, че най-големите Hadoop инсталации в света се състоят от десетки Hadoop клъстера, като всеки от тези клъстери се състои от няколко хиляди DataNode сървъра.

Частична визуална представа от Центъра за данни на FB може да се получи от фиг.8



Фиг.8.1

От фигурата и направените по-горе изчисления може да се направи извода, че инсталация на Hadoop клъстер в рамките на 100-200 рака е естествен подход при задача за 1000PB и повече, която задача е съизмерима през 2018 г. за големи Центрове за данни. В допълнение на тези разсъждения, ако за нормално обслужване на един рак отделим 4 квадратни метра в Център за данни, то 200 рака изискват 800 квадратни метра, което е 35 пъти по-малко от площта на центъра на FB.

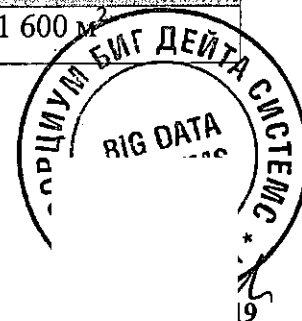
В резултат от направените анализи, може да се стигне до следните обобщени изводи, представени в таблица 8.1.

Таблица 8.1

Обем памет в Hadoop клъстер	Брой DataNode сървъри	Брой ракове	Необходима площ в Центъра за данни
125PB	1 000	25	100 м ²
250PB	2 000	50	200 м ²
1 000PB	8 000	200	800 м ²
2 000PB	16 000	400	1 600 м ²

Всичко това води до следните изводи:

000.78

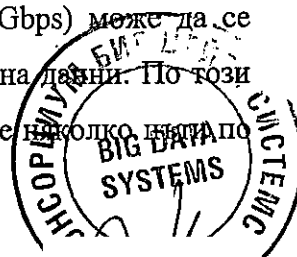


- Създаването на Hadoop клъстер с обем памет стотици и хиляди PB е възможно предложени архитектурен подход – както от брой сървъри, вид дискове в всеки сървър, брой ракове, така и необходима площ в Център за данни.
- Използването на DataNode сървър с посочените параметри е възможно да служи като основа на Hadoop клъстер със стотици и хиляди PB.

8.3 Използване на дискове с големи обеми в DataNode сървъри

При наличие на 12 дискови устройства и нови дискови устройства с обеми 14TB, 16TB, 20TB води до дисков капацитет на DataNode сървър от 200TB - 2 пъти повече от препоръчаното от Cloudera б.х. Същевременно са появили SSD устройства с 30TB и 60TB, което изглежда, че не могат да намерят приложение в DataNode и Hadoop клъстери. Това е така, когато се използват готови за използване сървъри с общо, без да се въвеждат в тях допълнителни архитектурни решения.

Съществуващите сървъри с общо предназначение имат следните решения от гледна точка на дисковите устройства: SAS-3 дисков контролер, работещ с 12Gbps скорост за обмен управлява всички дискови устройства, които имат средно скорост за обмен 150MBs-250MBs [49] (което прави $250\text{MBs} \times 8 = 2\text{Gbps}$), а сървърът има изходна мрежова скорост за обмен на данни 40Gbps. По този начин SAS-3 дисковият контролер (с 12Gbps) може да обслужва без забавяне на 6 дискови устройства ($6 \times 2\text{Gbps} = 12\text{Gbps}$). При наличие на 12 дискови устройства в DataNode сървър, то те биха се обслужили с максималната си скорост на обмен, ако сървърът разполага с няколко (поне 2 броя) SAS-3 дискови контролера. Същевременно, всеки SAS-3 дисков контролер обикновено се поставя в сървъра в съответен PCIe 3.0 или PCIe 4.0 интерфейсни куплунг. Скоростите на обмен на тези интерфейсни куплунги в съвременните сървъри достига до 7GBps (за PCIe 3.0x8), което прави 56Gbps. По този начин в DataNode сървър, работещ с 12 дискови устройства, могат да се използват 2 броя PCIe 3.0x8 интерфейсни куплунга, в които да се включат по един SAS-3 дискови контролери, като първите 6 диска се свържат към първия SAS-3 контролер, а вторите 6 диска се свържат към втория SAS-3 контролер. Скоростта на SAS-3 контролерите ($2 \times 12\text{Gbps}$) може да се поеме от 40Gbps за сървъра изходна мрежова скорост за предаване на данни. По този начин могат да се използват в сървъра дискове със сумарен капацитет няколко пъти по-голям от 100TB.



000.79

Това архитектурно решение отваря вратите за използване на SSD дискове с голем обем (30TB / 60TB), с използване на няколко PCIe 3.0x8, PCIe 3.0x16 или PCIe 4.0x8, в които да се включат SAS-3 дискови контролери. Това става все по-възможно с навлизането на SAS-4 контролерите и съответните дискови устройства.

Използването на дискове с големи обеми в DataNode сървъри ще намали броя ракове необходими за създаването на Hadoop клъстери със стотици и хиляди PB. Например с използване на 20TB дискове, посочените необходими брой DataNode сървъри, брой ракове и необходима площ в центъра за данни, посочени в Таблица 8.1. ще се намалят на половина.

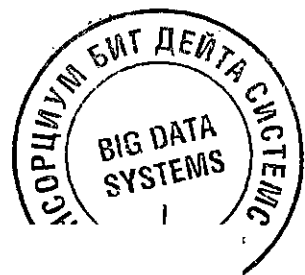
8.4 Hadoop Агрегирана система състояща се от няколко Hadoop клъстера

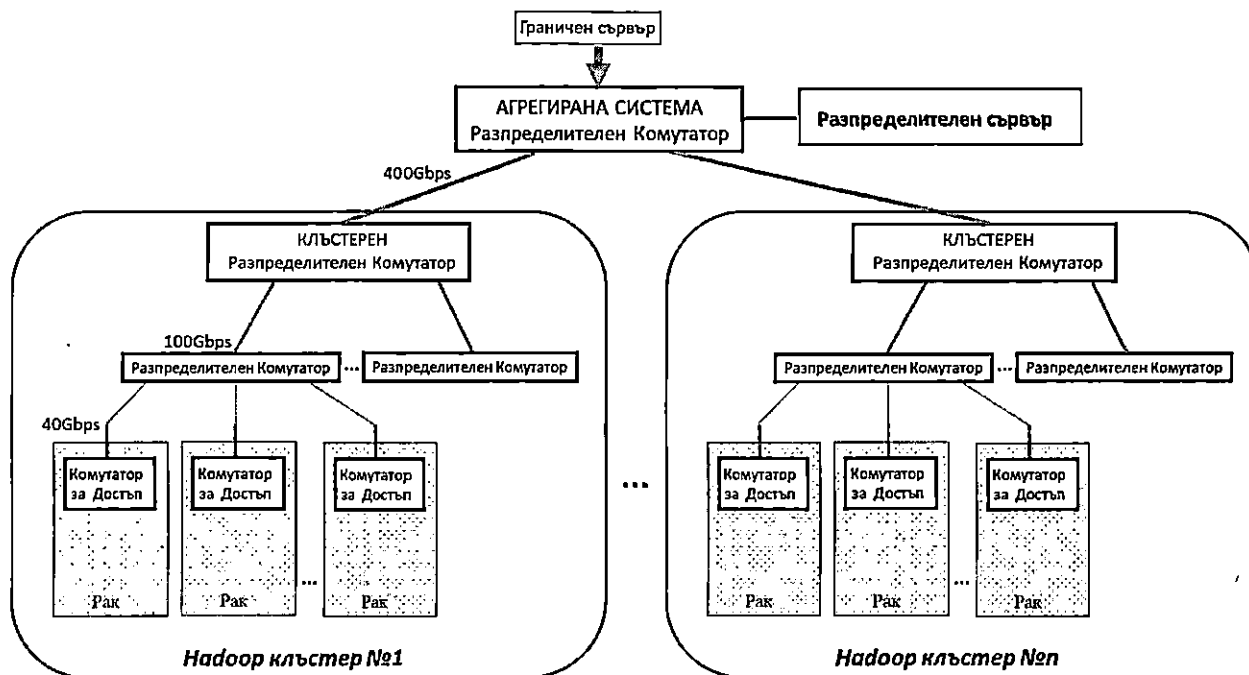
Създаването на Hadoop Агрегирана система се от няколко Hadoop клъстера не е теоретична постановка, а вече практически приложено решение, например във Facebook.

Предназначението на Hadoop Агрегираната система е да изгради едана система от няколко Hadoop клъстера от няколко събрагения:

- Да се използват Hadoop клъстери с по-малки обеми, например от по 150PB-200PB, които се обслужват технически по-лесно;
- Да се използват Hadoop клъстери, които има логическо разделение според типа използвани данни.

Архитектурата на Hadoop Агрегираната система е показана на фигура 8.2.





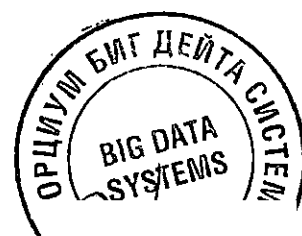
Фиг. 8.2.

Всеки Hadoop клъстер ще се състои от определен набор от ракове, в които са разположени както DataNode сървърите, така и NameNode и ManagementNode сървърите. От гледна точка на мрежово осигуряване, комуникацията от DataNode сървърите до Комутатора за достъп разположен в рака, може да бъде 10Gbps, а от Комутатора за достъп разположен в рака до Клъстерния разпределителен протокол да се използват връзки със 100Gbps. Между отделните Hadoop клъстери може да се създаде допълнителна мрежа чрез „Агрегирана система разпределителен комутатор“, който да има връзка с отделните Hadoop клъстери с връзки 200Gbps-400Gbps, в зависимост от натурата на разпределението между отделните Hadoop клъстери. Посочените 40Gbps, 100Gbps, 400Gbps са примерни, като тези - 40Gbps и 100Gbps са пряко свързани с работата на Hadoop клъстер и те не трябва да са по-малки от посочените, докато връзката между отделните Hadoop клъстери може да варира между 200Gbps и 400Gbps, в зависимост какво количество от данни предстои да се обменя между отделните Hadoop клъстери. При добра изолация на ниво данни между отделните Hadoop клъстери, то тези връзки може да бъдат даже и 100Gbps. Връзките от 200Gbps-400Gbps може да се изградят чрез няколко връзки от по 100Gbps, агрегиращи в една логическа връзка, както е показано по-горе в този материал.

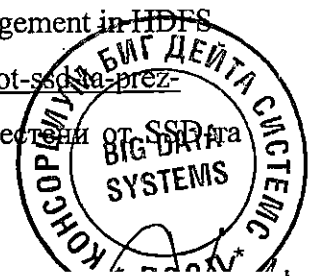


Граничният сървър има за задача да служи основно като потребителски сървър, в който се зареждат потребителските приложения и от който излизат заявките към Hadoop Агрегираната система.

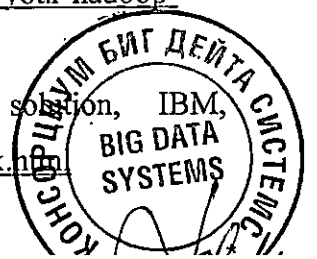
Специална функция в Hadoop Агрегирана система има Разпределителният сървър, който не е стандартен сървър в Hadoop философията, но в Hadoop Агрегирана система играе роля да насочи отделните заявки към отделните Hadoop клъстери, като при необходимост трябва да изпълнява интегриращи функции между отделните NameNode сървъри на отделните Hadoop кръстери.



- 0%BE %D1%83%D1%81%D1%82%D1%80%D0%BE%D0%B9%D1%81%
%D0%B2%D0%BE - Полупроводниково (статично) дисково устройство
13. https://en.wikipedia.org/wiki/Serial_Attached_SCSI - Serial Attached SCSI
 4. <https://www.slideshare.net/HadoopSummit/hdfs-analysis-for-small-files> - Analyzing small files in HDFS Cluster
 5. <https://hadoop.apache.org/docs/r3.0.0-alpha3/hadoop-project-dist/hadoop-hdfs/HDFSErasureCoding.html> - HDFS Erasure Coding
 16. James S . Plank, Erasure Codes for Storage Systems, https://www.usenix.org/system/files/login/articles/10_plank-online.pdf
 17. P. Vijay Kumar, Codes for Big Data: Erasure Coding for Distributed Storage, The 3rd Annual Storage Developer Conference, Bengaluru, May 25-26, 2017 - <https://www.snia.org/sites/default/files/SDCIndia/2017/Slides/Prof%20Kumar%20-%20IISc%20-%20Erasure%20codes.pdf>
 18. <https://www.edureka.co/blog/interview-questions/hadoop-interview-questions-hdfs-2/>
 19. <https://medium.com/arabamlabs/small-files-in-hadoop-88708e2f6a46> - Small files in Hadoop, Problem
 20. Sachin Bendea, Rajashree Shedge, Dealing with Small Files Problem in Hadoop Distributed File System, https://ac.els-cdn.com/S1877050916002581/1-s2.0-S1877050916002581-main.pdf?tid=spdf-d892eb68-05f7-4939-9575-2124d7cdbdb8&acdnat=1519734416_9d3e11a5164f82536cce0147530dad8d
 21. <https://community.hds.com/community/products-and-solutions/pentaho/blog/2017/11/07/working-with-small-files-in-hadoop-part-2> - Working with Small Files in Hadoop - Part 2
 22. Yanpei Chen, The truth about MapReduce performance on SSDs, <https://www.oreilly.com/ideas/the-truth-about-mapreduce-performance-on-ssds>
 23. <https://archive.cloudera.com/cdh5/cdh/5/hadoop/hadoop-project-dist/hadoop-hdfs/ArchivalStorage.html> - Archival Storage, SSD & Memory
 24. <http://hadoop.apache.org/docs/r2.8.3/hadoop-project-dist/hadoop-hdfs/CentralizedCacheManagement.html> - Centralized Cache Management in HDFS
 25. <https://www.digital.bg/zashto-hard-diskovete-shte-badat-izmesteni-ot-ssd-ata-prez-2018g-article523919.html> - Защо хард дисковете ще бъдат изместени от SSD ата през 2018г.



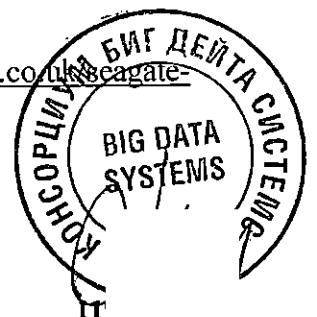
26. <https://www.digital.bg/nay-golemiqt-ssd-disk-v-sveta-e-s-kapacitet-ot-153tb-i-veche-mozhe-da-se-kupi-za-%E2%82%AC8990-article517228.html> - Най-големият SSD диск в света е с капацитет от 15.3TB и вече може да се купи за €8990
27. <https://www.digital.bg/samsung-predstaviha-parviq-ssd-disk-s-kapacitet-ot-30tb-article670501.html> - Samsung представиха първия SSD диск с капацитет от 30TB
28. <http://www.zdnet.com/article/worlds-largest-ssd-revealed-as-seagate-unveils-60tb-monster/> - 'World's largest' SSD revealed as Seagate unveils 60TB monster
29. <http://blog.cloudera.com/blog/2014/03/the-truth-about-mapreduce-performance-on-ssds/> - The Truth About MapReduce Performance on SSDs
30. <https://www.slideshare.net/SamsungBusinessUSA/big-data-ssd-architecture-digging-deep-to-discover-where-ssd-performance-pays-off> - Big Data SSD Architecture: Digging Deep to Discover Where SSD Performance Pays Off
31. <https://www.sandisk.com/business/datacenter/resources/white-papers/increasing-hadoop-performance-with-sandisk-ssds-whitepaper> - Increasing Hadoop Performance with SanDisk SSDs
32. <http://ieeexplore.ieee.org/document/7906530/> - Can erasure codes damage reliability in SSD-based Storage Systems?
33. Pradeep Subedi, Exploration of Erasure-Coded Storage Systems for High Performance, Reliability, and Inter-operability, <https://scholarscompass.vcu.edu/cgi/viewcontent.cgi?referer=https://www.google.bg/&httpsredir=1&article=5480&context=etd>
34. <http://i.dell.com/sites/doccontent/business/large-business/en/Documents/Dell-Cloudera-Apache-Hadoop-Soution-Reference-Architecture.pdf> - Dell | Cloudera Apache Hadoop Solution Reference Architecture Guide - Version 5.5.1
35. HDFS Settings for Better Hadoop Performance, <https://community.hortonworks.com/articles/103176/hdfs-settings-for-better-hadoop-performance.html>
36. How to ensure best performance for your Hadoop Cluster?, <https://www.dezyre.com/article/how-to-ensure-best-performance-for-your-hadoop-cluster/200>
37. Understanding the architectural layers of a big data solution, IBM, <https://www.ibm.com/developerworks/library/bd-archpatterns3/index.html>



000.35

11/01/2014

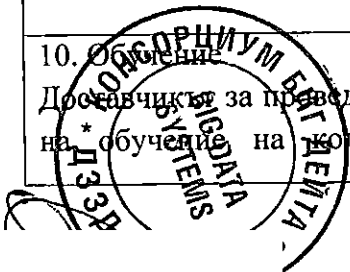
38. Cloudera Enterprise Reference Architecture for Bare Metal Deployment
http://www.cloudera.com/documentation/other/reference-architecture/PDF/cloudera_ref_arch_metal.pdf
39. Hardware Requirements ✓ Gui
https://www.cloudera.com/documentation/enterprise/release-notes/topics/hardware_requirements_guide.html
40. Hadoop Distribution : Cloudera vs Hortonworks – Which One is Better?,
<https://www.springpeople.com/blog/hortonworks-or-cloudera-which-one-is-better/>
41. Comparing Cloudera, Hortonworks, <https://www.gartner.com/reviews/market/data-warehouse-solutions/compare/cloudera-vs-hortonworks>
42. Seagate Introduces a 60TB SSD – Is a 3.6PB Storage Pod Next?,
<https://www.backblaze.com/blog/seagate-60tb-ssd-36pb-storage-pod-next/>
43. Generic Reference Architecture for Cloudera Enterprise Running in a Private Cloud,
https://www.cloudera.com/documentation/other/reference-architecture/topics/ra_private_cloud.html
44. Cloudera Enterprise Reference Architecture for Bare Metal Deployments,
https://www.cloudera.com/documentation/other/reference-architecture/topics/ra_bare_metal_deployment.html
45. Cloudera Enterprise Storage Device Acceptance Criteria Guide,
http://www.cloudera.com/documentation/other/reference-architecture/PDF/cloudera_ref_arch_stg_dev_accept_criteria.pdf
46. Cloudera Enterprise 5.x Release Notes, Hardware Requirements Guide,
https://www.cloudera.com/documentation/enterprise/release-notes/topics/hardware_requirements_guide.html
47. How Erasure Coding Changes Hadoop Storage Economics,
<https://www.datanami.com/2018/02/07/erasure-coding-changes-hadoop-storage-economics/>
48. Hadoop Architecture Explained-What it is and why it matters,
<https://www.dezyre.com/article/hadoop-architecture-explained-what-it-is-and-why-it-matters/317>
49. 10TB HDD Performance Testing, <https://www.tomshardware.com/uk/seagate-barracuda-pro-10tb-hdd,review-34183-2.html>
50. PCI Express, https://en.wikipedia.org/wiki/PCI_Express
- 51.



Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависим
-----------	-----------------------	---------------------	---------

<p>T.9 Архивиране на началната инсталация</p> <p>Създадената и подготвена за предаване в експлоатация Специализирана система по ключ Nadoor трябва да се архивира. Софтуерното съдържание на всеки изчислителен модул трябва да се архивира (backup) на отделен DVD диск. Конфигурационният файл на всеки комуникационен модул трябва да се архивира (backup) на отделен DVD диск. За два избрани от доставчика изчислителни модула и за един комуникационен модул трябва да се тества процеса на възстановяване (restore/rollback) на тяхното съдържание и последващото след това коректно функциониране на цялата Специализирана система по ключ Nadoor. Доставчикът да предостави набор от 300 DVD носителя за създаване на библиотека от архивирани копия на изчислителните и комуникационни модули.</p>	ДА	<p>На етап предаване на Специализирана система под ключ Nadoor ще бъде архивирано софтуерното съдържание на всеки изчислителен модул. Като всеки архивиран модул ще бъде на отделен диск. Конфигурационният файл на всеки комуникационен модул ще се архивира на отделен DVD диск. На етап предаване ще бъде тестван процеса на възстановяване (rollback/restore) на два изчислителни модула и на един комуникационен модул. На етап предаване на системата ще предоставим 300 DVD носителя за създаване на библиотека от архивирани копия на изчислителните и комуникационни модули.</p>	
<p>10. Обучение</p> <p>Доставчикът заведе следните видове курсове на обучение на колективи от заявителя на</p>	ДА	<p>На етап оферта ще предоставим писмен документ със съдържанието на обучението за трите основни курса, както следва:</p>	

000.87



Изискване	Изпълнение (Да/Не)	Начин на реализация	Завис
<p>територията на УНСС:</p> <ul style="list-style-type: none"> - Инсталиране и конфигуриране на Hadoop система; - Основи на програмирането в среда MapReduce; - Създаване на тестови данни за провеждане на изследвания. <p>еки курс да бъде няколко дневен.</p> <p>оят участници от заявителя в един курс да бъде до 15 човека.</p>		<ul style="list-style-type: none"> - Инсталиране и конфигуриране на Hadoop система; - Основи на програмирането в среда MapReduce; - Създаване на тестови данни за провеждане на изследвания. <p>Всеки курс ще бъде дву дневен. Брой участници – 15 човека.</p> <p>В документацията предоставяме документ със съдържанието на обучението по Т.10.</p>	

000-38



КУРС 1

Инсталиране и конфигуриране на Hadoop система

Ден 1:



Предпоставки за инсталацията (Prerequisites):

1. Linux OS;
2. Java 1.6+

1. Добавете системен Hadoop потребител със следните команди:

```
guru99@guru99-VirtualBox:~$ sudo adduser --ingroup hadoop_ hduser_
Adding user 'hduser_' ...
Adding new user 'hduser_' (1001) with group 'hadoop_' ...
Creating home directory '/home/hduser_' ...
Copying files from '/etc/skel' ...
Enter new UNIX password:
Retype new UNIX password:
passwd: password updated successfully
Adding the user information for hduser_
Enter the new value, or press ENTER for the default
Full Name []: Team
Room Number []: 1
Work Phone []: 1
Home Phone []: 1
Other []: 1
The information correct? [Y/n] y
guru99@guru99-VirtualBox:~$
```

Добавете
парола за
потребителя

```
sudo addgroup hadoop_
sudo adduser --ingroup hadoo
p_ hduser_
```

Изпълнете командата:

```
sudo adduser hduser_ sudo
Влезте като hduser_
```

2. Конфигурирайте SSH



Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

Първо сменете работния потребител със следната команда:

```
su - hduser_
```

Следната команда ще създаде нов ключ:

```
ssh-keygen -t rsa -P ""
```

```
guru99@guru99-VirtualBox:~$ su - hduser_
Password:
hduser_@guru99-VirtualBox:~$ ssh-keygen -t rsa -P ""
Generating public/private rsa key pair.
Enter file in which to save the key (/home/hduser_/.ssh/id_rsa):
Created directory '/home/hduser_/.ssh'.
Your identification has been saved in /home/hduser_/.ssh/id_rsa.
Your public key has been saved in /home/hduser_/.ssh/id_rsa.pub.
The key fingerprint is:
07:e2:3f:7d:7d:d1:0d:9d:12:0e:e7:27:ab:47:4a:22 hduser_@guru99-VirtualBox
The key's randomart image is:
+--[ RSA 2048 ]-----+
  |          .          |
  |         =...       |
  |        =.o.        |
  |       =.o          |
  |      .ES... o .o   |
  |     ..oo +.       |
  |    o .o...        |
  |     +. .          |
  |                    |
+-----+
hduser_@guru99-VirtualBox:~$
```



Активирайте SSH достъпа до локалната машина, като използвате този ключ.

```
cat $HOME/.ssh/id_rsa.pub >> $HOME/.ssh/authorized_keys
```

Сега тествайте SSH настройката, като се свържете към localhost като потребител 'hduser'.

```
ssh localhost
```



„Консорциум Биг Дейта Системс“

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

```
hduser_@guru99-VirtualBox:/$ ssh localhost
The authenticity of host 'localhost (127.0.0.1)' can't be established.
ECDSA key fingerprint is 4a:78:f7:93:32:0a:c1:b4:24:e2:a6:78:d7:cb:20:d6.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'localhost' (ECDSA) to the list of known hosts
Welcome to Ubuntu 12.04.1 LTS (GNU/Linux 3.2.0-29-generic-pae i686)

 * Documentation:  https://help.ubuntu.com/

The programs included with the Ubuntu system are free software;
the exact distribution terms for each program are described in the
individual files in /usr/share/doc/*/copyright.

Ubuntu comes with ABSOLUTELY NO WARRANTY, to the extent permitted by
applicable law.

hduser_@guru99-VirtualBox:~$
```

Инсталирайте ssh със следната команда:

```
sudo apt-get install openssh-server
```

Инсталация:

От <http://www.eu.apache.org/dist/hadoop/common/> изтеглете стабилна версия на Hadoop под формата на архивиран .tar.gz файл и я разархивирайте в локация по избор във файловата система.

```
sudo tar xzf hadoop-x.y.z.tar.gz
```

Преименувайте hadoop-x.y.z на hadoop.

```
sudo mv hadoop-2.2.0 hadoop
```

```
sudo chown -R hduser_:hadoop_ hadoop
```

Ден 2:

Конфигурация:

1. Модифицирайте ~/.bashrc файла

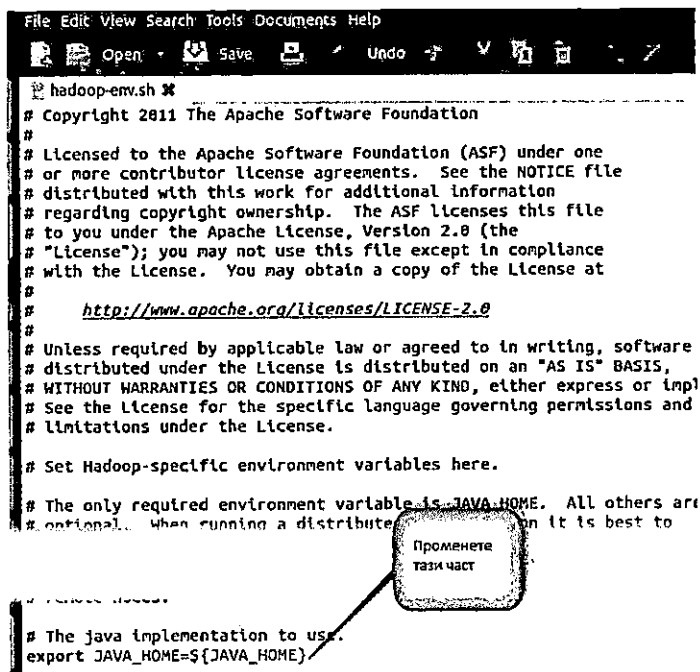


Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

Добавете следните редове на края на файла

```
#Set HADOOP_HOME
export HADOOP_HOME=<Инсталационна директория на Hadoop>
#Set JAVA_HOME
export JAVA_HOME=<Инсталационна директория на Java>
# Add bin/ directory of Hadoop to PATH
export PATH=$PATH:$HADOOP_HOME/bin
```

След това изпълнете командата `~/bashrc`



```
File Edit View Search Tools Documents Help
hadoop-env.sh
# Copyright 2011 The Apache Software Foundation
# Licensed to the Apache Software Foundation (ASF) under one
# or more contributor license agreements. See the NOTICE file
# distributed with this work for additional information
# regarding copyright ownership. The ASF licenses this file
# to you under the Apache License, Version 2.0 (the
# "License"); you may not use this file except in compliance
# with the License. You may obtain a copy of the License at
#
# http://www.apache.org/licenses/LICENSE-2.0
#
# Unless required by applicable law or agreed to in writing, software
# distributed under the License is distributed on an "AS IS" BASIS,
# WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
# See the License for the specific language governing permissions and
# limitations under the License.
#
# Set Hadoop-specific environment variables here.
#
# The only required environment variable is JAVA_HOME. All others are
# optional. When running a distributed cluster it is best to
#
# The java implementation to use.
export JAVA_HOME=${JAVA_HOME}
```

2. Конфигурации, свързани с HDFS
Настройте JAVA_HOME във файла \$HADOOP_HOME/etc/hadoop/hadoop-env.sh

Променете означения ред с:

```
export
JAVA_HOME=/home/guru99/Downloads/jdk.8
.0_05
```

където guru99 е името на root потребителя

Има 2 параметъра в \$HADOOP_HOME/etc/hadoop/core-site.xml, които трябва да бъдат настроени

- а) `hadoop.tmp.dir` – Използва се за насочване към директория, където Hadoop съхранява файлове с данни
- б) `fs.default.name` – Използва се за уточняване на файловата система по подразбиране

За да настроите тези параметри, отворете `core-site.xml`:

```
sudo gedit $HADOOP_HOME/etc/hadoop/core-site.xml
```



„Консорциум Биг Дейта Системс“ ДЗ

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

Копирайте следните редове между таговете <configuration></configuration>:

```
<property>
<name>hadoop.tmp.dir</name>
<value>/app/hadoop/tmp</value>
<description>Главна директория за временни поддиректории.</description>
</property>
<property>
<name>fs.defaultFS </name>
<value>hdfs://localhost:54310</value>
<description>Името на файловата система по подразбиране. </description>
</property>
```

Навигирайте до \$HADOOP_HOME/etc/Hadoop и създайте директорията, спомената в по-горната стъпка

```
sudo mkdir -p <Път на директорията в предишната стъпка>
```

Възложете права на директорията:

```
sudo chown -R hduser_:Hadoop_ < Път на директорията в предишната стъпка >
sudo chmod 750 < Път на директорията в предишната стъпка >
```

3. MapReduce конфигурация

Преди да започне тази конфигурация, трябва да се настрои HADOOP_HOME път.

```
sudo gedit /etc/profile.d/hadoop.sh и въведете export HADOOP_HOME=/home/guru99/Downlo
ads/Hadoop
```

След това въведете: `sudo chmod +x /etc/profile.d/hadoop.sh`

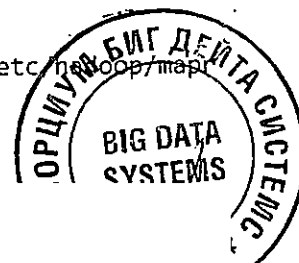
Излезте от терминала и рестартирайте. След рестартирането, въведете в терминала `echo $HADOOP_HOME`, за да потвърдите създадения път.

Сега копирайте файловете:

```
sudo cp $HADOOP_HOME/etc/hadoop/mapred-site.xml.template $HADOOP_HOME/etc/hadoop/mapred-site.xml
```

Отворете `mapred-site.xml` файла:

```
sudo gedit $HADOOP_HOME/etc/hadoop/mapred-site.xml
```



„Консорциум Биг Дейта Системс“ |

Хай Тех Бизнес Ц
ж.к. Дружба 1, ул. Мюл
1592 София, Р
Tel.: +359 2 42 10 10

Копирайте следните редове между таговете <configuration></configuration>:

```
<property>
<name>mapreduce.jobtracker.address</name>
<value>localhost:54311</value>
<description>MapReduce job tracker runs at this host and port.
</description>
</property>
```

Отворете \$HADOOP_HOME/etc/hadoop/hdfs-site.xml файла:

```
sudo gedit $HADOOP_HOME/etc/hadoop/hdfs-site.xml
```

Копирайте следните редове между таговете <configuration></configuration>:

```
<property>
<name>dfs.replication</name>
<value>1</value>
<description>Default block replication.</description>
</property>
<property>
<name>dfs.datanode.data.dir</name>
<value>/home/hduser_/hdfs</value>
</property>
```

Създайте директорията, описана в стъпките по-горе: `sudo mkdir -p /home/hduser_/hdfs`

```
sudo chown -R hduser_:hadoop_ /home/hduser_/hdfs
```

```
sudo chmod 750 /home/hduser_/hdfs
```

4. Преди да се задейства Hadoop за пръв път, форматирайте HDFS със следната команда:

```
$HADOOP_HOME/bin/hdfs namenode -format
```

5. Стартирайте Hadoop single node cluster със следната команда:

```
$HADOOP_HOME/sbin/start-dfs.sh
```

Примерен изход на командат



```
hduser@guru99-VirtualBox:~$ SHADOOP_HOME/sbin/start-dfs.sh
Starting namenodes on [localhost]
localhost: starting namenode, logging to /home/guru99/Downloads/hadoop/logs/hadoop-hduser_namenode-guru99-VirtualBox.out
localhost: starting datanode, logging to /home/guru99/Downloads/hadoop/logs/hadoop-hduser_datanode-guru99-VirtualBox.out
Starting secondary namenodes [0.0.0.0]
The authenticity of host '0.0.0.0 (0.0.0.0)' can't be established.
ECDSA key fingerprint is 4a:78:f7:93:32:0a:ci:b4:24:e2:a6:78:d7:ch:78:06.
Are you sure you want to continue connecting (yes/no)? yes
0.0.0.0: Warning: Permanently added '0.0.0.0' (ECDSA) to the list of known hosts.
0.0.0.0: starting secondarynamenode, logging to /home/guru99/Downloads/hadoop/logs/hadoop-hduser_secondarynamenode-guru99-VirtualBox.out
hduser@guru99-VirtualBox:~$
```

Въведете
Yes

\$HADOOP_HOME/sbin/start-yarn.sh

```
hduser@guru99-VirtualBox:~$ SHADOOP_HOME/sbin/start-yarn.sh
starting yarn daemons
starting resourcemanager, logging to /home/guru99/Downloads/hadoop/logs/yarn-hduser_resourceanager-guru99-VirtualBox.out
localhost: starting nodemanager, logging to /home/guru99/Downloads/hadoop/logs/yarn-hduser_nodemanager-guru99-VirtualBox.out
hduser@guru99-VirtualBox:~$
```

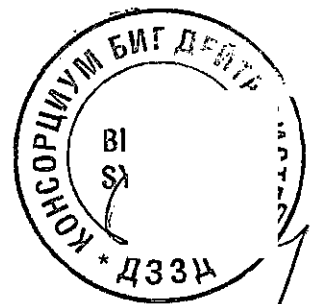
Използвайте `jps tool/command`, за да потвърдите, че всички процеси, свързани с Hadoop работят.
Успешен изход:

```
hduser@guru99-VirtualBox:~$ jps
3732 SecondaryNameNode
4326 Jps
3865 ResourceManager
3466 DataNode
4061 NodeManager
3279 NameNode
hduser@guru99-VirtualBox:~$
```

6. Спиране на Hadoop:

\$HADOOP_HOME/sbin/stop-dfs.sh
\$HADOOP_HOME/sbin/stop-yarn.sh

000_95



КУРС 2

Основи на MapReduce програмирането

Ден 1:

Двете най-големи предимства на MapReduce са:

1. Паралелна обработка

В MapReduce разделяме заданието между множество възли и всеки възел работи едновременно с част от задачата. Така че, MapReduce се основава на парадигмата "Разделяй и владей", която ни помага да обработваме данните, използвайки различни машини. Тъй като данните се обработват от множество машини вместо една машина паралелно, времето, необходимо за обработка на данните, намалява с огромно количество

2. Локалност на данните

Вместо да преместваме данни към обработващата единица, ние преместваме обработващата единица към данните в **MapReduce Framework**. В традиционната система, ние въвеждахме данни в обработващата единица и ги обработвахме. Но тъй като данните нарастваха и станаха много огромни, това огромно количество данни в обработващата единица постави следните въпроси:

- Преместването на огромни данни в обработка е скъпо и влошава производителността на мрежата.
- Обработката отнема време, тъй като данните се обработват от една единица, която става пречка.
- Главният възел може да бъде претоварен и може да се провал

Сега **MapReduce** ни позволява да преодоляваме горепосочените проблеми като пренасяме обработващата единица към данните. Данните се рапределят между множество възли, където всеки възел обраб



Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

от данните, които се намират в него. Този факт ни дава следните
преимущества.

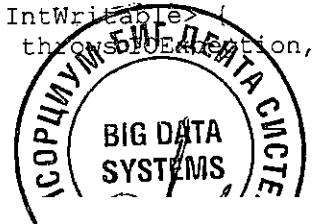
- Много е рентабилно да преместите процесора към данните.
- Времето за обработка се намалява, тъй като всички възли работят с тяхната част от данните паралелно.
- Всеки възел получава част от данните за обработване и следователно няма вероятност възел да бъде претоварен.

Примерна програма на MapReduce

Преди да прескочим в детайлите, нека се спрем на примерната програма на **MapReduce**, за да имаме основна идея за това как работят нещата в среда на **MapReduce** на практика. В този пример се наблюдава **wordcount** приложение, което следи броя пъти, в които всяка дума присъства в обем текст.

Source код:

```
1 package co.edureka.mapreduce;
2 import java.io.IOException;
3 import java.util.StringTokenizer;
4 import org.apache.hadoop.io.IntWritable;
5 import org.apache.hadoop.io.LongWritable;
6 import org.apache.hadoop.io.Text;
7 import org.apache.hadoop.mapreduce.Mapper;
8 import org.apache.hadoop.mapreduce.Reducer;
9 import org.apache.hadoop.conf.Configuration;
10 import org.apache.hadoop.mapreduce.Job;
11 import org.apache.hadoop.mapreduce.lib.input.TextInputFormat;
12 import org.apache.hadoop.mapreduce.lib.output.TextOutputFormat;
13 import org.apache.hadoop.mapreduce.lib.input.TextInputFormat;
14 import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;
15 import org.apache.hadoop.fs.Path;
16
17 public class WordCount
18 {
19     public static class Map extends Mapper<LongWritable, Text, Text, IntWritable> {
20         public void map(LongWritable key, Text value, Context context) throws IOException, Int
21         String line = value.toString();
22         StringTokenizer tokenizer = new StringTokenizer(line);
23         while (tokenizer.hasMoreTokens()) {
24             value.set(tokenizer.nextToken());
```



000.97

„Консорциум Биг Дейта Системс“ Д

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

```
21 context.write(value, new IntWritable(1));
22 }
23 }
24 }
25 public static class Reduce extends Reducer<Text, IntWritable, Text, IntWritable> {
26 public void reduce(Text key, Iterable<IntWritable> values, Context context) throws IC
27 int sum=0;
28 for(IntWritable x: values)
29 {
30 sum+=x.get();
31 }
32 context.write(key, new IntWritable(sum));
33 }
34 }
35
36 public static void main(String[] args) throws Exception {
37
38 Configuration conf= new Configuration();
39 Job job = new Job(conf, "My Word Count Program");
40 job.setJarByClass(WordCount.class);
41 job.setMapperClass(Map.class);
42 job.setReducerClass(Reduce.class);
43 job.setOutputKeyClass(Text.class);
44 job.setOutputValueClass(IntWritable.class);
45 job.setInputFormatClass(TextInputFormat.class);
46 job.setOutputFormatClass(TextOutputFormat.class);
47 Path outputPath = new Path(args[1]);
48 //Configuring the input/output path from the filesystem into the job
49 FileInputFormat.addInputPath(job, new Path(args[0]));
50 FileOutputFormat.setOutputPath(job, new Path(args[1]));
51 //deleting the output path automatically from hdfs so that we don't have to delete
52 outputPath.getFileSystem(conf).delete(outputPath);
53 //exiting the job only if the flag value becomes false
54 System.exit(job.waitForCompletion(true) ? 0 : 1);
55 }
56 }
```

Ден 2:

Обяснение на примерната програма

Цялата програма може да бъде фундаментално разделена на три части:

- Код за Mapper фаза
- Код за Reducer фаза
- Driver код

БМГ ДЕЙ.

„Консорциум Биг Дейта Системс“

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

Ще добием представа за всяка част в последователен ред.

Mapper код:

```
1 public static class Map extends Mapper<LongWritable, Text, Text, IntWritable> {
2
3 public void map(LongWritable key, Text value, Context context) throws IOException, In
4
5 String line = value.toString();
6 StringTokenizer tokenizer = new StringTokenizer(line);
7 while (tokenizer.hasMoreTokens()) {
8 value.set(tokenizer.nextToken());
9 context.write(value, new IntWritable(1));
10 }
```

- Създадохме клас Map, който разширява класа Mapper, който вече е дефиниран в MapReduce Framework
- Дефинираме типа на данните на входните и изходни key/value двойки след декларацията на класа чрез <>
- И входът, и изходът на Mapper са key/value двойка.

Вход:

- Key е нищо друго, освен отместването на всеки ред в текстовия файл: *LongWritable*
- Value е всеки отделен ред: *Text*

Изход:

- Key са думи под формата на токени: *Text*
- Имаме кодираното value, което в нашия случай е 1: *IntWritable*

Написали сме java код, където сме означили всяка дума и им приписали твърда стойност, равна на 1.

Reducer код:

```
1 public static class Reduce extends Reducer<Text, IntWritable, Text, IntWritable> {
2 public void reduce(Text key, Iterable<IntWritable> values, Context context) throws IOException, InterruptedException {
3
4 }
```

000.99



Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

```
3 throws IOException, InterruptedException {
4
5     int sum=0;
6     for(IntWritable x: values)
7     {
8         sum+=x.get();
9     }
10    context.write(key, new IntWritable(sum));
11 }
12
13
```

- Създадохме клас `Reduce`, който разширява класа `Reducer`, както този на `Mapper`.
- Дефинираме типа на данните на входните и изходни `key/value` двойки след декларацията на класа чрез `<>`
- И входът, и изходът на `Reducer` са `key/value` двойка.
- Вход:
 - **Key** е нищо друго, освен уникални думи, които са генерирани по време на `sorting` и `shuffling` фазите: *Text*
 - **Value** е списък с `integer` числа, отговарящи на всяка уникална дума: *IntWritable*
 - Пример – Bear, [1,1], и т.н.
- Изход:
 - **Key** са всички уникални думи, които присъстват във входния текстови файл: *Text*
 - **Value** е броят на пъти, в които всяка от тези уникални думи е присъствала в обема текст: *IntWritable*
 - Пример – Bear, 2; Car, 3 и т.н.
- Обобщихме стойностите, присъстващи във всеки от списъка, съответстващ на всеки ключ, и получихме окончателния отговор
- Като цяло за всяка от уникалните думи се създава единичен `reducer`, но можете да укажете броя на `reducer` елементите в `mapred-site.xml`

Driver код:

```
1 Configuration conf= new Configuration();
2 Job job = new Job(conf, "My Word Count Program");
3 job.setJarByClass(WordCount.class);
4 job.setMapperClass(Map.class);
5 job.setReducerClass(Reduce.class);
6 job.setOutputKeyClass(Text.class);
7
8 job.setOutputValueClass(IntWritable.class);
9 job.setInputFormatClass(TextInputFormat.class);
```

5 |

000300



„Консорциум Биг Дейта Системс

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

```
9  job.setOutputFormatClass(TextOutputFormat.class);
10 Path outputPath = new Path(args[1]);
11
12 //Configuring the input/output path from the filesystem into the job
13 FileInputFormat.addInputPath(job, new Path(args[0]));
14 FileOutputFormat.setOutputPath(job, new Path(args[1]));
15
```

- В **driver** класа, настройваме конфигурацията на нашата MapReduce задача да протече в Hadoop.
- Уточняваме името на задачата, типът на входните и изходните данни на mapper и reducer елементите.
- Също така уточняваме имената на mapper и reducer класовете.
- Пътят на входната и изходната папка са уточнени.
- Методът setInputFormatClass () се използва за уточняване на метода за четене на входни данни от mapper елемент, иначе казано обема на работа. В този случай сме избрали TextInputFormat ,за да може mapper елемента да чете от входния текстови файл линия по линия.
- Main () методът е входната точка за driver. В този метод създаваме нов обект за конфигурация на заданието.

Изпълнение на MapReduce код:

Командата а изпълнение на MapReduce код е:

```
hadoop jar hadoop-mapreduce-example.jar WordCount /sample/input /sample/output
```



00301

КУРС 3

Създаване на тестови данни за провеждане на изследвания.

Ден 1:

Когато използвате нов или актуализиран хардуер или софтуер, простите примери и бенчмаркове помагат за потвърждаване на правилната работа. Apache Hadoop включва много примери и бенчмаркове за подпомагане на тази задача. Този курс предоставя инструкции как да изпълнявате, наблюдавате и управлявате някои основни примери и бенчмаркове на MapReduce.

Изпълнение на примерни MapReduce задачи:

Всички издания на Hadoop идват с примерни приложения MapReduce. Стартирането на съществуващите примери за MapReduce е лесен за изпълнение процес след като файловете с примери са набавени. Например, ако сте инсталирали Hadoop версия 2.6.0 от Apache източници в /opt, примерите ще бъдат в следната директория:

```
/opt/hadoop-2.6.0/share/hadoop/mapreduce/
```

В други версии примерите могат да бъдат в /usr/lib/hadoop-mapreduce/ или друго място. Точното местоположение на примера на файла с jar може да бъде намерено чрез командата find:

```
$ find / -name "hadoop-mapreduce-examples*.jar" -print
```

В тази среда местоположението на примерите е /usr/hdp/2.2.4.2-2/hadoop-mapreduce. За целите на този пример, променлива на средата, наречена HADOOP_EXAMPLES, може да се дефинира както следва:



„Консорциум Биг Дейта Системс“ ДЗ:

Хай Тех Бизнес Цент
ж.к. Дружба 1, ул. Мюнхе
1592 София, РОВ 74
Tel.: +359 2 42 10 10

```
$ export HADOOP_EXAMPLES=/usr/hdp/2.2.4.2-2/hadoop-mapreduce
```

След като дефинирате пътя на примерите, можете да изпълните примерите на Hadoop с помощта на командите, разгледани в следващите раздели.

Изброяване на възможните примери.

Списък с наличните примери може да бъде намерен чрез изпълнение на следната команда. В някои случаи номерът на версията може да е част от jar файла (например в източниците на Apache версия 2.6, файлът се нарича `hadoop-mapreduce-examples-2.6.0.jar`).

```
$ yarn jar $HADOOP_EXAMPLES/hadoop-mapreduce-examples.jar
```

Възможните примери са следните:

Примерна програма трябва да бъде дадена като първия аргумент.

Валидни имена на програми са:

aggregatewordcount: Aggregate базирана map/reduce програма, която брои думите във входните файлове.

aggregatewordhist: An Aggregate базирана map/reduce програма, която изчислява хистограмата на думите във входните файлове.

bbp: map/reduce програма, която използва Bailey-Borwein-Plouffe, за да изчисли точност цифри на пи.

dbcoun: Примерна програма, която брои броя на преглежданите страници от база данни.



00003

„Консорциум Биг Дейта Системс“ ДЗЗ

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен
1592 София, РОВ 74
Tel.: +359 2 42 10 10

distbbp: map/reduce програма, която използва ВВР-тип формула, за да изчисли с точност стойността на пи.

grep: map/reduce програма, която брой съвпаденията на regex във входа.

join: Програма, която задейства обединението на сортирани, еднакво разделени масиви от данни.

multifilewc: Програма, която брой думи от няколко файла.

pentomino: map/reduce програма за намиране на решения на pentomino проблеми.

pi: map/reduce програма, която дава приблизителна стойност на пи, чрез quasi-Monte Carlo метод.

Carlo метод.

randomtextwriter: map/reduce програма, която записва 10GB произволни данни в текстови формат на всяко устройство.

randomwriter: map/reduce програма, която записва 10GB произволни данни на всяко устройство.

secondarysort: Пример, който дефинира вторично сортиране на редукцията.

sort: map/reduce програма, която сортира данни въведени от източник на произволни данни.

sudoku: Програма за решаване на sudoku.

teragen: Генериране на данни за terasort

terasort: Изпълнение на terasort

teravalidate: Проверка на резултати от terasort



000504

„Консорциум Биг Дейта Системс“ ДЗЗД

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

wordcount: map/reduce програма, която преброява думите във входни файлове.

wordmean: map/reduce програма, която изчислява средната дължина на думите във входни файлове.

wordmedian: map/reduce програма, която изчислява междинната дължина на думите във входни файлове.

wordstandarddeviation: map/reduce програма, която изчислява стандартната девиация в дължината на думите във входни файлове.

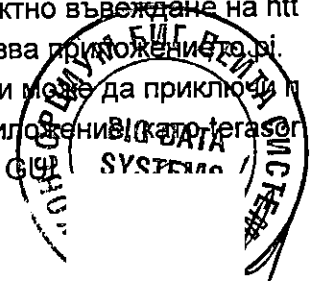
Изпълнение на примерна програма (в случая, избраната програма е pi)

```
$ yarn jar $HADOOP_EXAMPLES/hadoop-mapreduce-examples.jar pi 16 1000000
```

Ден 2:

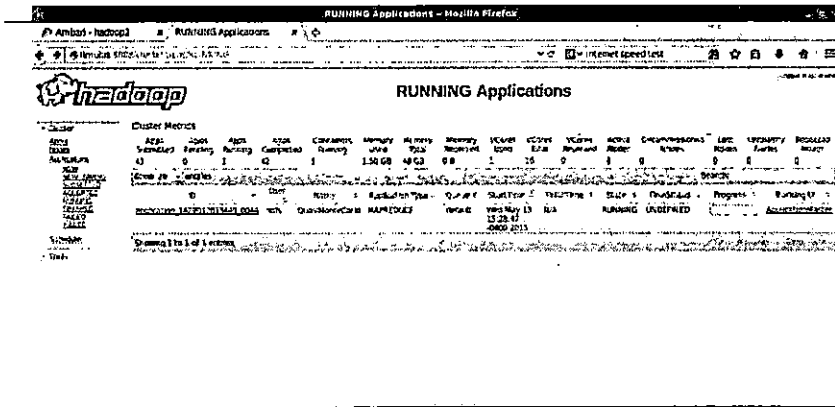
Използване на WebGUI за проверка на изпълнението на примерни задачи.

Този раздел предоставя илюстрация за използването на уеб GUI ResourceManager за наблюдение и намиране на информация за YARN задания. Версията Hadoop 2 YARN ResourceManager web GUI се различава значително от уеб графичния интерфейс на MapReduce, който се намира в Hadoop версия 1. Фигурата показва основния уеб интерфейс на YARN. Показателите на клъстера се показват в най-горния ред, докато изпълняваните приложения се показват в основната таблица. Меню отляво осигурява навигация към таблицата на възлите, различни категории задачи (напр. Ново, Приемано, Работещо, Завършено, Не успешно) и Планировчик на капацитета. Този интерфейс може да бъде отворен директно от менюто за бързи връзки на услугата Ambari YARN или чрез директно въвеждане на http://hostname:8088 в локален уеб браузър. За този пример се използва приложението pi. Обърнете внимание, че приложението може да работи бързо и може да приключи преди да сте напълно проучили GUI. По-продължителното приложение vi като terasort, може да бъде полезно при проучване на всички различни връзки в GUI.

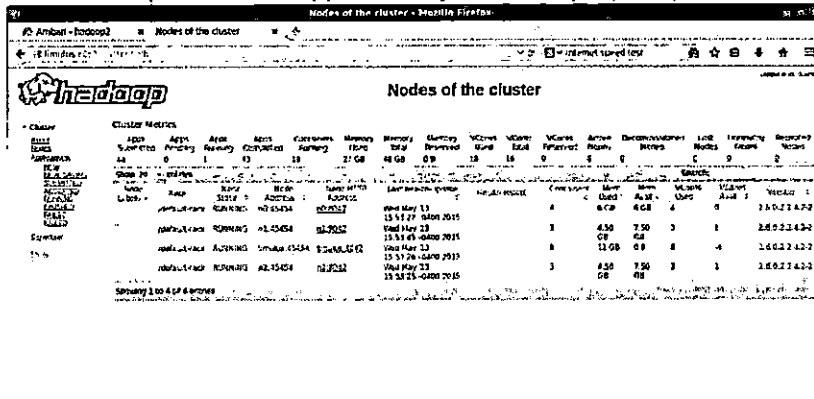


„Консорциум Биг Дейта Системс“ ДЗ:

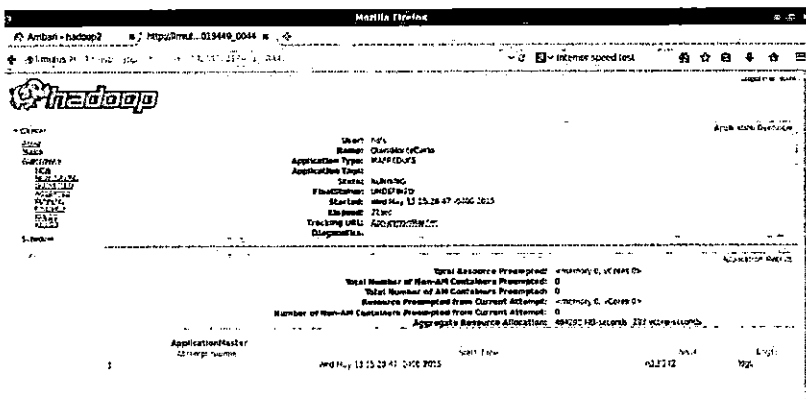
Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10



Фиг.1 Hadoop RUNNING Applications уеб GUI за pi примера.



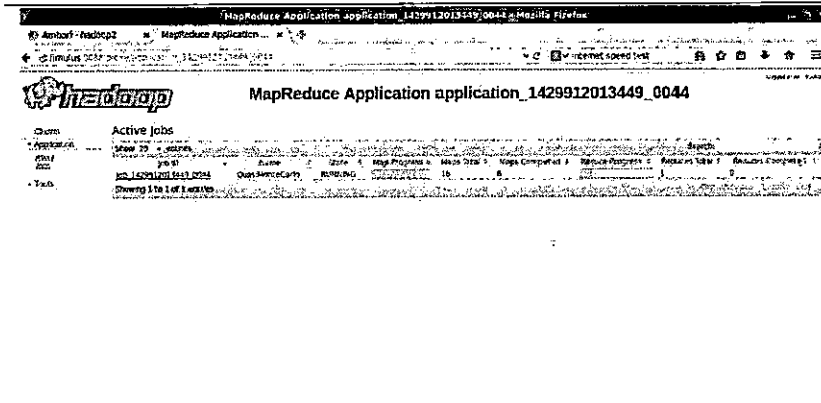
Фиг.2 Hadoop YARN ResourceManager статус на устройствата.



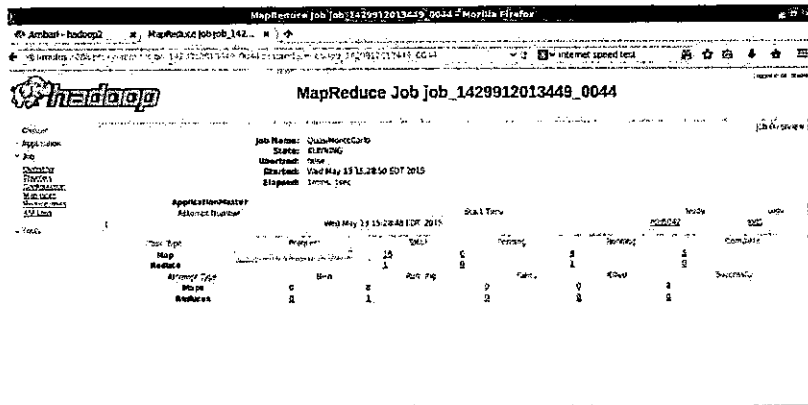
Фиг.3 Hadoop YARN application статус за pi примера.



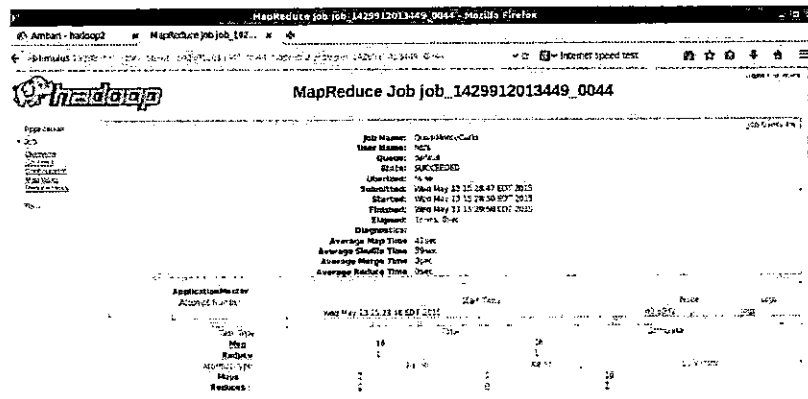
Хай Тех Бизнес Център
 ж.к. Дружба 1, ул. Мюнхен 8
 1592 София, РОВ 74
 Tel.: +359 2 42 10 10



Фиг.4 Hadoop YARN ApplicationMaster за MapReduce приложение.



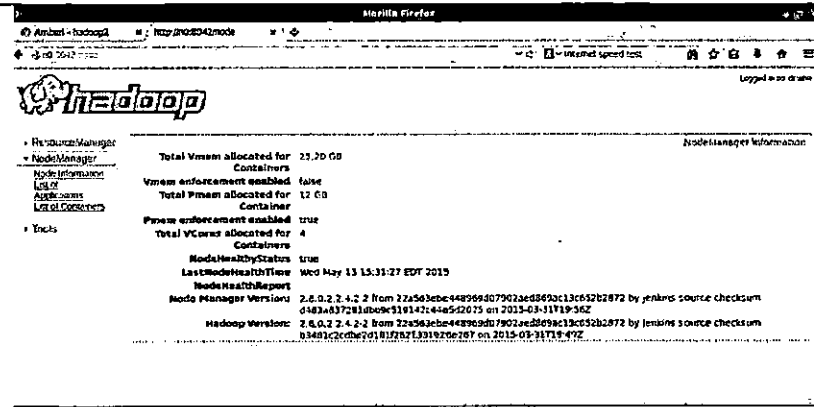
Фиг.5 Hadoop YARN MapReduce прогрес на задачата.



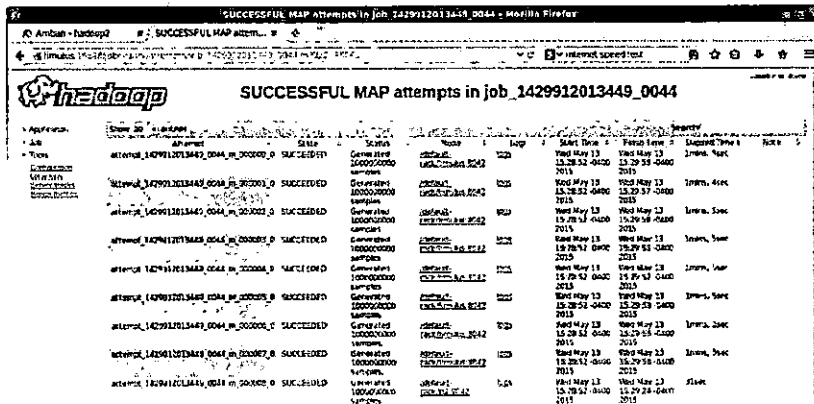
Фиг.6 Hadoop YARN обобщение на завършената MapReduce задача.



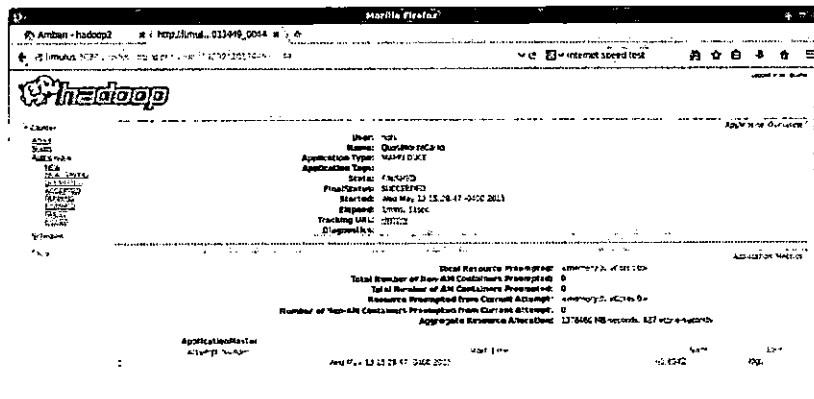
Хай Тех Бизне
 ж.к. Дружба 1, ул. Мюнхен 8
 1592 София, РОВ 74
 Tel.: +359 2 42 10 10



Фиг.7 Hadoop YARN NodeManager за п0 обобщение на изпълнението на задачата.



Фиг.8 Hadoop YARN MapReduce регистри, налични за преглед и анализ.



Фиг.9 Hadoop YARN application обобщителна страница.



Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависимост
-----------	-----------------------	---------------------	------------

<p>T.11 Изисквания към Специализираната централизирана система под ключ и нейното поддържане.</p>	<p>ДА</p>	<p>Предлагаме да изпълним изискванията към Специализираната централизирана система под ключ и нейното поддържане съгласно документацията за участие, като ще осигурим всички необходими условия, съгласно изискванията на Възложителя.</p> <p>Предоставяме документ за „Изпълнието изискванията към Специализираната централизирана система под ключ и нейното поддържане.“</p>	
---------------------------------------------------------------------------------------------------	-----------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--

000000



Предложение за изпълнение на поръчката в съответствие с техническата спецификация и изискванията на Възложителя по Т.11, стр. 48 от Документация за участие в Открита процедура за възлагане на обществена поръчка с предмет:

„Доставка, инсталиране, конфигуриране, тестване и поддържане на СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ НАDOOP за работа с Големи данни по проект „Дигитализация на икономиката в среда на Големи данни“ (ДИГД), № BG05M2OP001-1.002-0002“

Т.11 Изисквания към Специализираната централизирана система под ключ и нейното поддържане

1. Предлагаме да изпълним поръчката съгласно документацията за участие, като ще осигурим всички необходими условия, съгласно изискванията на Възложителя.
2. Приемаме:

- Общите изисквания по Т.11.1 както следва:

- a) Предлаганата техника ще бъде нова, неупотребявана, нерициклирана и в производствената листа на производителя към момента на подаване на офертата, за което посочаваме линк:
 - <https://www.supermicro.com/products/system/1U/6019/SSG-6019P-ACR12L.cfm>
 - <https://www.supermicro.com/products/accessories/Networking/SSE-X3348S.cfm>
 - https://www.socomec.com/range-ups-three-phase_en.html?product=/ups-modulys-rm-gp-green-power_en.html
 - <https://www.supermicro.com/products/system/1U/1029/SYS-1029P-WTR.cfm>
 - <https://gtec-power.eu/en/ups/singlephase-online/ap160n-on-line-double-conversion-ups/>
 - <https://www.supermicro.com/products/accessories/Networking/SSE-C3632S.cfm>
 - https://www.supermicro.com/wftp/Networking_Drivers/CDR-NIC_1.52_for_Add-on_NIC_Cards/MANUALS/AOC-STGN-i1S-i2S.pdf
 - <https://www.watchguard.com/wgrd-products/rack-mount/firebox-m4600-m5600>
 - <https://www.alliedtelesis.com/products/switches/x5101-28gt>за доказване на горепосоченото.
- b) Предлаганата техника е съвместима с операционните системи Windows и Linux, за което посочаваме линкове, за доказване на горепосоченото, както следва:

https://www.supermicro.com/support/resources/OS/OS_Certification.html?ProductID=SSG-6019P-ACR12L

000510



https://www.supermicro.com/support/resources/OS/OS_Certification_Intel.cfm?MPProduct_Name=SYS-1029P-WTR

-, Ще предоставим резултати от тест, чрез тестова програма или чрез персонализирана тестове на компонентите в кълстерната хардуерна системата, за което посочваме линк достъпен чрез FTP клиент:

- <ftp://data.persy.com> Потребител: tests Парола: Tests123!@#

d) Всички хардуерни системи са снабдени с технически средства за наблюдение и управление, за което посочваме линк:

<https://www.supermicro.com/en/solutions/management-software/supermicro-server-manager>

e) Цялата хардуерна кълстерна система в настоящата поръчка трябва да бъде напълно интегрирана от сертифициран от производителя специалист.

- Задължения на изпълнителя по Т.11.2 както следва:

Ще осигурим гаранционен срок на Цялата Специализирана централизирана система под ключ за 5 години, считано от датата на разписване на договора и приемаме необходимите гаранционни ремонти описани по – долу, както следва:

- Приемаме срокът за реакция при регистриране/уведомление от страна на Възложителя за хардуерна повреда да бъде следния: при подадена регистрация/уведомление в работно време - до 2 часа, и при подадена регистрация/уведомление в извънработно време – до началото на следващия работен ден, като реакцията е свързана с посещение на съответен специалист на място в УНСС;

- Приемаме срокът за реакция при регистрация/уведомление от страна на Възложителя за повреда в инсталацията и/или конфигурацията на управляващия софтуер на Cloudera на Hadoop кълстера да бъде следния: при подадена регистрация/уведомление в работно време - до 2 часа, и при подадена регистрация/уведомление в извънработно време – до началото на следващия работен ден, като реакцията е свързана с посещение на сертифициран специалист или такъв с преминато обучение/положен изпит при производител/сертифициран център за обучение по „Cloudera Administrator“, на място в УНСС.

- Приемаме срокът за реакция при регистрация/уведомление от страна на Възложителя за повреда в информационната сигурност да бъде следния: при подадена регистрация/уведомление в работно време - до 2 часа, и при подадена регистрация/уведомление в извънработно време – до началото на следващия работен ден, като реакцията е свързана с посещение на съответен специалист, на място в УНСС;

- Приемаме срокът за отстраняване на повреда да бъде, както следва:

- За повреди в сървъри, дискови масиви и комуникационни устройства - до 2 работни дни;

- За повреди във всички останали хардуерни компоненти – до 2 работни дни;

- За повреда в инсталацията и/или конфигурацията на управляващия софтуер на Cloudera – до 2 работни дни, когато повреда е в рамките на последната версия на управляващия софтуер и до 14 дни, когато повреда е свързана с друга версия на управляващия со



- За повреда в информационната сигурност - до 2 работни дни, свързани със софтуерни и конфигурационни проблеми на информационната сигурност, и до 14 дни при проблеми свързани с хардуерни компоненти на информационната сигурност.

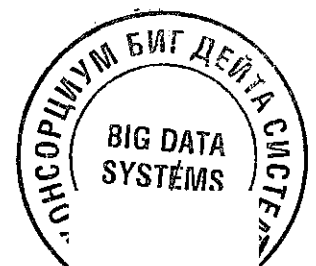
- Приемаме гаранцията за качеството на стоката и се задължаваме:

- да осигурим гаранционно обслужване и сервиз за срок от 5 години;
- да извършваме качествено сервизно техническо поддържане и ремонт на доставената компютърна техника в съответствие с действащите стандарти за срок от 5 години;
- да отстраним възникналите при нормалната експлоатация повреди в рамките на гаранционния срок от 5 години;
- да отстраняваме безвъзмездно всички повреди и отклонения от изискванията за качество, които са възникнали в рамките на посочения гаранционен срок;
- По време на първите 6 месеца от гаранционния срок, се задължаваме при установяване на еднороден дефект на компоненти, особено ако това е установено на не по - малко от 20% от техниката от един вид да подменим дефектния компонент и на другата техника от същия вид, които не са проявили дефекта.

- Приемаме да не носим отговорност в следните случаи:

- за повреди и отклонения от качеството на стоката, възникнали вследствие на нейната неправилна употреба, експлоатация, неспазване на указанията в съпровождащата документация или неправилно съхранение от страна на Възложителя;
- поражения при стихийни бедствия: пожар, земетресение, наводнение, големи колебания в електрическата мрежа, посегателства и др.

- Приемаме, че при невъзможност за отстраняване на повредата в дефектирало устройство в рамките на срока, същото се заменя с работоспособно от същия или по-висок клас.
- Приемаме, че ако не предприемем действия по отстраняване на повредите или замяна на дефектиралата техника в срок, Възложителят може да отстрани повредата за наша сметка. Всички разходи по поправката или замяната ще са за наша сметка и ще се удържат от гаранцията за добро изпълнение.
- Приемаме гаранционното сервизно обслужване да се извършва на мястото на използване на стоките – в УНСС, а ако това е невъзможно да се извършва в посочения сервиз: сервиз на ПЕРСИ ООД в гр. София ул. Златна Добруджа 18



000312

Изискване	Изпълнение (Да/Не)	Начин на реализация	
-----------	-----------------------	---------------------	--

<p>Участниците трябва да представят следните документи, свързани с „Информационна сигурност на Специализираната централизирана система под ключ“:</p> <p>- Документ „Ръководство на администратора за Създаване на централизирано автентикиране на потребителите на Hadoop система“, в което се описва:</p> <p>а) подход за създаване съответното изискване с посочено техническо и програмно-конфигурационно решение;</p> <p>б) един или няколко примера за прилагане на изискването.</p> <p>Документ „Ръководство на администратора за Създаване на права за достъп сървъри и сегменти от данни в Hadoop</p>	<p>ДА</p>	<p>Предоставяме описаните в документацията документи свързани с „Информационна сигурност на Специализираната централизирана система под ключ“, както следва:</p> <p>- Предоставяме: Документ „Ръководство на администратора за Създаване на централизирано автентикиране на потребителите на Hadoop система“, с описани като отделни раздели:</p> <p>а) подход за създаване съответното изискване с посочено техническо и програмно-конфигурационно решение;</p> <p>б) един или няколко примера за прилагане на изискването.</p> <p>- Предоставяме: Документ „Ръководство на администратора за Създаване на права за достъп за сървъри и сегменти от данни в Hadoop система, на вече създадени</p>	<p></p>
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------



Изискване	Изпълнение (Да/Не)	Начин на реализация	За
<p>система, на вече създадени потребители“, в което се описва:</p> <p>а) подход за създаване съответното изискване с посочено техническо и програмно-конфигурационно решение;</p> <p>б) един или няколко примера за прилагане на изискването.</p> <p>- Документ „Ръководство на администратора за Създаване контрол на достъпа до HDFS директории и файлове на Hadoop система“, в което се описва:</p> <p>а) подход за създаване съответното изискване с посочено техническо и програмно-конфигурационно решение;</p> <p>б) един или няколко примера за прилагане на изискването.</p> <p>- Документ „Ръководство на администратора за Създаване централизиран Лист за управление на достъпа до Hadoop</p>		<p>потребители“ с описани като отделни раздели:</p> <p>а) подход за създаване съответното изискване с посочено техническо и програмно-конфигурационно решение;</p> <p>б) един или няколко примера за прилагане на изискването.</p> <p>- Предоставяме: Документ „Ръководство на администратора за Създаване контрол на достъпа до HDFS директории и файлове на Hadoop система“, с описани като отделни раздели:</p> <p>а) подход за създаване съответното изискване с посочено техническо и програмно-конфигурационно решение;</p> <p>б) един или няколко примера за прилагане на изискването.</p> <p>- Предоставяме: Документ „Ръководство на администратора за Създаване централизиран Лист за управление на достъпа до Hadoop</p>	<p>✓</p> <p>✓</p> <p>✓</p>

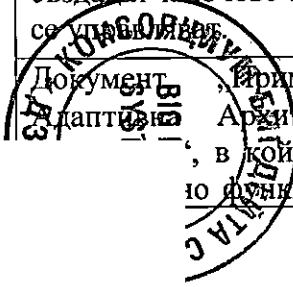
Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависим
<p>система“, в което се описва:</p> <p>а) подход за създаване съответното изискване с посочено техническо и програмно-конфигурационно решение;</p> <p>б) пример за прилагане на изискването.</p> <p>Документ „Ръководство на администратора за Създаване на механизъм за пълен одит на Nadoor система“, в което се описва:</p> <p>а) подход за създаване съответната одитна система с 3 параметъра - произход на данни, достъп до данни и идентификация на потребител, с посочено техническо и програмно-конфигурационно решение;</p> <p>б) разпечатка от работата на създадена одитна система с посочените 3-те броя параметри.</p> <p>Документ „Ръководство на администратора за Осигуряване на защита и</p>		<p>система“, с описани като отделни раздели:</p> <p>а) подход за създаване съответното изискване с посочено техническо и програмно-конфигурационно решение;</p> <p>б) пример за прилагане на изискването.</p> <p>- Предоставяме: Документ „Ръководство на администратора за Създаване на механизъм за пълен одит на Nadoor система“, с описани като отделни раздели:</p> <p>а) подход за създаване съответната одитна система с 3 параметъра - произход на данни, достъп до данни и идентификация на потребител, с посочено техническо и програмно-конфигурационно решение;</p> <p>б) разпечатка от работата на създадена одитна система с посочените 3-те броя параметри.</p> <p>- Предоставяме: Документ „Ръководство на администратора за Осигуряване на защита и криптиране на данните в Nadoor система за данни</p>	<p>Зависим</p> <p>С</p>

00015



Изискване	Изпълнение (Да/Не)	Начин на реализация	3
<p>криптиране на данните в Nadoor система за данни върху диск и данни в движение“, в което се описва:</p> <p>а) подход за създаване съответното изискване с посочено техническо и програмно-конфигурационно решение;</p> <p>б) по един пример за двата вида криптиране, като се използва един от предварително създадени потребители.</p> <p>- Документ „Ръководство на администратора за Изграждане управление на ключовете за криптиране в Nadoor система“, в което се описва:</p> <p>а) подход за създаване съответното изискване с посочено техническо и програмно-конфигурационно решение;</p> <p>б) пример как за създадени 2 броя потребители се създават ключове и се представи как може те да се управляват.</p>		<p>върху диск и данни в движение“, с описани като отделни раздели:</p> <p>а) подход за създаване съответното изискване с посочено техническо и програмно-конфигурационно решение;</p> <p>б) по един пример за двата вида криптиране, като се използва един от предварително създадени потребители.</p> <p>- Прилагаме: Документ „Ръководство на администратора за Изграждане управление на ключовете за криптиране в Nadoor система“, с описани като отделни раздели:</p> <p>а) подход за създаване съответното изискване с посочено техническо и програмно-конфигурационно решение;</p> <p>б) пример как за създадени 2 броя потребители се създават ключове и се представи как може те да се управляват.</p>	<p>3</p> <p>✓</p>
<p>Документ „Примерно функциониране на адаптивна Архитектура за Информационна Сигурност“, в който са представени разпечатки на функциониране и конфигуриране в</p>	ДА	<p>Предоставяме документ: „Примерно функциониране на Адаптивна Архитектура за Информационна Сигурност“, който съдържа</p>	

000010



Изискване	Изпълнение (Да/Не)	Начин на реализация
<p>компютърна среда на доставчика, на следните компоненти: Модула за УАЛС, IPS устройството и Модула за МУД. В този документ да се представят разпечатки от следните 4 броя примерно интегриране:</p> <p>Между Модул за УАЛС и LOG файловете на един модул МодулDataNode, и на един модул МодулNameNode, както и на техните операционни системи;</p> <p>Резултат от корелация между съдържание на LOG файл на един модул МодулDataNode и на LOG файл на един модул МодулNameNode;</p> <p>Между IPS устройството и Модул за прилагане на политики– Модула за МУД;</p> <p>Между Модул за УАЛС и Модул за МУД.</p>		<p>следните точки:</p> <ol style="list-style-type: none"> 1. Общо описание на реализацията 2. Архитектура на Адаптивна Архитектура за Информационна Сигурност – диаграми. 3. Предоставяме като приложения към документ „Примерно функциониране на Адаптивна Архитектура за Информационна Сигурност“ разпечатки от примерно функциониране и конфигуриране в компютърна среда на доставчика, на следните компоненти: Модула за УАЛС, IPS устройството и Модула за МУД както следва: <ul style="list-style-type: none"> - Между Модул за УАЛС и LOG файловете на един модул МодулDataNode, и на един модул МодулNameNode, както и на техните операционни системи; - Резултат от корелация между съдържание на LOG файл на един модул МодулDataNode и на LOG файл на един модул МодулNameNode; - Между IPS устройството и Модул за

000000



Изискване	Изпълнение (Да/Не)	Начин на реализация	
		<p>прилагане на политики– Модула за МУД; - Между Модул за УАЛС и Модул за МУД.</p> <p>Прилагаме документа за „Примерно функциониране на Адаптивна Архитектура за Информационна Сигурност“.</p>	<p>7</p> <p>8</p>

000018



ДОКУМЕНТ

Във връзка с изискване от страна на Възложителя,

„Участникът да представи документи, свързани с „Информационна сигурност на Специализираната централизирана система под ключ“

описани на стр. 51 от Документацията за участие в Открита процедура за възлагане на обществена поръчка с предмет:

Доставка, инсталиране, конфигуриране, тестване и поддържане на СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ NADOOR за работа с Големи данни по проект „Дигитализация на икономиката в среда на Големи данни“ (ДИГД), № BG05M2OP001-1.002-0002“

Прилагаме списък на документите:

1. Документ „Ръководство на администратора за Създаване на централизирано автентикиране на потребителите на Nadoor система“, в което се описва:

а) подход за създаване съответното изискване с посочено техническо и програмно-конфигурационно решение;

б) един или няколко примера за прилагане на изискването.

2. Документ „Ръководство на администратора за Създаване на права за достъп за сървъри и сегменти от данни в Nadoor система, на вече създадени потребители“, в което се описва:

а) подход за създаване съответното изискване с посочено техническо и програмно-конфигурационно решение;

б) един или няколко примера за прилагане на изискването.

3. Документ „Ръководство на администратора за Създаване контрол на достъпа до HDFS директории и файлове на Nadoor система“, в което се описва:

а) подход за създаване съответното изискване с посочено техническо конфигурационно решение;



000.19

„Консорциум Биг Дейта Системс“ ДЗ:

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

б) един или няколко примера за прилагане на изискването.

4. Документ „Ръководство на администратора за Създаване централизиран Лист за управление на достъпа до Hadoop система“, в което се описва:

а) подход за създаване съответното изискване с посочено техническо и програмно-конфигурационно решение;

б) пример за прилагане на изискването.

5. Документ „Ръководство на администратора за Създаване на механизъм за пълен одит на Hadoop система“, в което се описва:

а) подход за създаване съответната одитна система с 3 параметъра - произход на данни, достъп до данни и идентификация на потребител, с посочено техническо и програмно-конфигурационно решение;

б) разпечатка от работата на създадена одитна система с посочените 3-те броя параметри.

6. Документ „Ръководство на администратора за Осигуряване на защита и криптиране на данните в Hadoop система за данни върху диск и данни в движение“, в което се описва:

а) подход за създаване съответното изискване с посочено техническо и програмно-конфигурационно решение;

б) по един пример за двата вида криптиране, като се използва един от предварително създадени потребители.

7. Документ „Ръководство на администратора за Изграждане управление на ключовете за криптиране в Hadoop система“, в което се описва:

а) подход за създаване съответното изискване с посочено техническо и програмно-конфигурационно решение;

б) пример как за създадени 2 броя потребители се създадат ключове и се представи как може те да се управляват.

8. Документ „Примерно функциониране на Адаптивна Архитектура за Информационна Сигурност“, в който са представени разпечатки от примерно функциониране и конфигуриране в компютърна среда на доставчика, на следните компоненти: Модула за УАЛС, IPS устройството и Модула за МУД. В този документ да се представят разпечатки от следните 4 броя примерно интегриране:



„Консорциум Биг Дейта Системс“

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

Между Модул за УАЛС и LOG файловете на един модул МодулDataNode, и на един модул МодулNameNode, както и на техните операционни системи;

Резултат от корелация между съдържание на LOG файл на един модул МодулDataNode и на LOG файл на един модул МодулNameNode;

Между IPS устройството и Модул за прилагане на политики– Модула за МУД;

Между Модул за УАЛС и Модул за МУД.



11

Ръководство на администратора за Създаване на
централизирано автентикиране на потребителите
на Hadoop система



000522

11

Подход за създаване съответното изискване с посочено техническо и програмно-конфигурационно решение:

По подразбиране **Hadoop** и **YARN** действително автентикиране. потребител и услуга трябва използват защитен режим, при който не се изисква иране на **Hadoop** в защитен режим, всеки автентикирани от **Kerberos**, за да могат да се използват услугите **Hadoop**.

Когато се включи удостоверяване на ниво услуга, крайните потребители, които използват **Hadoop** в защитен режим, трябва да бъдат удостоверени от **Kerberos**. Най-простият начин да направите удостоверяване е използването на **kinit** командата на **Kerberos**.

Потребителски акаунти за Hadoop Daemons

Уверете се, че **HDFS** и **YARN** daemons се изпълняват като различни **Unix** потребители, напр. **hdfs** и **yarn**. Също така, уверете се, че **MapReduce JobHistory** сървърът работи като различен потребител, като например **mapred**. Препоръчва се те да споделят **Unix** група, напр. **hadoop**.

User:Group	Daemons
hdfs:hadoop	NameNode, Secondary NameNode, JournalNode, DataNode
yarn:hadoop	ResourceManager, NodeManager
mapred:hadoop	MapReduce JobHistory Server

Принципи на Kerberos за Hadoop Daemons и потребители

За да се изпълняват служебни **hadoop service daemons**, докато **Hadoop** е в защитен режим, се изискват принципи на **Kerberos (Kerberos principals)**. Всяка услуга чете информация за автентикиране, запаметена в **keytab** файла с подходящо разрешение. **HTTP** уеб-конзолите трябва да се обслужват от принцип, различен от този на **RPC**.

Пример за прилагане на изискването:

Подразделите по-долу показват примери за данни за автентикация за услугите на **Hadoop**.

HDFS

Файлът **keytab** на **NameNode** на същия хост трябва да изглежда по следния начин:

```

$ klist -e -k -t /etc/security/keytab/nn.service.keytab
Keytab name: FILE:/etc/security/keytab/nn.service.keytab
KVNO Timestamp Principal

```



4 07/18/11 21:08:09 nn/full.qualified.domain.name@REALM.TLD (AES-256 CTS mode with 96-bit SHA-1 HMAC)
4 07/18/11 21:08:09 nn/full.qualified.domain.name@REALM.TLD (AES-128 CTS mode with 96-bit SHA-1 HMAC)
4 07/18/11 21:08:09 nn/full.qualified.domain.name@REALM.TLD (ArcFour with HMAC/md5)
4 07/18/11 21:08:09 host/full.qualified.domain.name@REALM.TLD (AES-256 CTS mode with 96-bit SHA-1 HMAC)
4 07/18/11 21:08:09 host/full.qualified.domain.name@REALM.TLD (AES-128 CTS mode with 96-bit SHA-1 HMAC)
4 07/18/11 21:08:09 host/full.qualified.domain.name@REALM.TLD (ArcFour with HMAC/md5)

Файлт keytab на SecondaryNameNode на същия хост трябва да изглежда по следния начин:

```
$ klist -e -k -t /etc/security/keytab/sn.service.keytab
Keytab name: FILE:/etc/security/keytab/sn.service.keytab
KVNO Timestamp Principal
4 07/18/11 21:08:09 sn/full.qualified.domain.name@REALM.TLD (AES-256 CTS mode with 96-bit SHA-1 HMAC)
4 07/18/11 21:08:09 sn/full.qualified.domain.name@REALM.TLD (AES-128 CTS mode with 96-bit SHA-1 HMAC)
4 07/18/11 21:08:09 sn/full.qualified.domain.name@REALM.TLD (ArcFour with HMAC/md5)
4 07/18/11 21:08:09 host/full.qualified.domain.name@REALM.TLD (AES-256 CTS mode with 96-bit SHA-1 HMAC)
4 07/18/11 21:08:09 host/full.qualified.domain.name@REALM.TLD (AES-128 CTS mode with 96-bit SHA-1 HMAC)
4 07/18/11 21:08:09 host/full.qualified.domain.name@REALM.TLD (ArcFour with HMAC/md5)
```

Файлт keytab на DataNode на всеки хост трябва да изглежда по следния начин:

```
$ klist -e -k -t /etc/security/keytab/dn.service.keytab
Keytab name: FILE:/etc/security/keytab/dn.service.keytab
KVNO Timestamp Principal
4 07/18/11 21:08:09 dn/full.qualified.domain.name@REALM.TLD (AES-256 CTS mode with 96-bit SHA-1 HMAC)
4 07/18/11 21:08:09 dn/full.qualified.domain.name@REALM.TLD (AES-128 CTS mode with 96-bit SHA-1 HMAC)
4 07/18/11 21:08:09 dn/full.qualified.domain.name@REALM.TLD (ArcFour with HMAC/md5)
4 07/18/11 21:08:09 host/full.qualified.domain.name@REALM.TLD (AES-256 CTS mode with 96-bit SHA-1 HMAC)
4 07/18/11 21:08:09 host/full.qualified.domain.name@REALM.TLD (AES-128 CTS mode with 96-bit SHA-1 HMAC)
4 07/18/11 21:08:09 host/full.qualified.domain.name@REALM.TLD (ArcFour with HMAC/md5)
```

YARN

Файлт keytab на ResourceManager на същия хост трябва да изглежда по следния начин:

```
$ klist -e -k -t /etc/security/keytab/rm.service.keytab
Keytab name: FILE:/etc/security/keytab/rm.service.keytab
```



KVNO Timestamp Principal
4 07/18/11 21:08:09 rm/full.qualified.domain.name@REALM.TLD (AES-256 CTS mode with 96-bit SHA-1 HMAC)
4 07/18/11 21:08:09 rm/full.qualified.domain.name@REALM.TLD (AES-128 CTS mode with 96-bit SHA-1 HMAC)
4 07/18/11 21:08:09 rm/full.qualified.domain.name@REALM.TLD (ArcFour with HMAC/md5)
4 07/18/11 21:08:09 host/full.qualified.domain.name@REALM.TLD (AES-256 CTS mode with 96-bit SHA-1 HMAC)
4 07/18/11 21:08:09 host/full.qualified.domain.name@REALM.TLD (AES-128 CTS mode with 96-bit SHA-1 HMAC)
4 07/18/11 21:08:09 host/full.qualified.domain.name@REALM.TLD (ArcFour with HMAC/md5)

Файлт keytab на NodeManager на същия хост трябва да изглежда по следния начин:

```
$ klist -e -k -t /etc/security/keytab/nm.service.keytab
Keytab name: FILE:/etc/security/keytab/nm.service.keytab
KVNO Timestamp Principal
4 07/18/11 21:08:09 nm/full.qualified.domain.name@REALM.TLD (AES-256 CTS mode with 96-bit SHA-1 HMAC)
4 07/18/11 21:08:09 nm/full.qualified.domain.name@REALM.TLD (AES-128 CTS mode with 96-bit SHA-1 HMAC)
4 07/18/11 21:08:09 nm/full.qualified.domain.name@REALM.TLD (ArcFour with HMAC/md5)
4 07/18/11 21:08:09 host/full.qualified.domain.name@REALM.TLD (AES-256 CTS mode with 96-bit SHA-1 HMAC)
4 07/18/11 21:08:09 host/full.qualified.domain.name@REALM.TLD (AES-128 CTS mode with 96-bit SHA-1 HMAC)
4 07/18/11 21:08:09 host/full.qualified.domain.name@REALM.TLD (ArcFour with HMAC/md5)
```

MapReduce JobHistory Server

Файлт keytab на MapReduce JobHistory Server на същия хост трябва да изглежда по следния начин:

```
$ klist -e -k -t /etc/security/keytab/jhs.service.keytab
Keytab name: FILE:/etc/security/keytab/jhs.service.keytab
KVNO Timestamp Principal
4 07/18/11 21:08:09 jhs/full.qualified.domain.name@REALM.TLD (AES-256 CTS mode with 96-bit SHA-1 HMAC)
4 07/18/11 21:08:09 jhs/full.qualified.domain.name@REALM.TLD (AES-128 CTS mode with 96-bit SHA-1 HMAC)
4 07/18/11 21:08:09 jhs/full.qualified.domain.name@REALM.TLD (ArcFour with HMAC/md5)
4 07/18/11 21:08:09 host/full.qualified.domain.name@REALM.TLD (AES-256 CTS mode with 96-bit SHA-1 HMAC)
4 07/18/11 21:08:09 host/full.qualified.domain.name@REALM.TLD (AES-128 CTS mode with 96-bit SHA-1 HMAC)
4 07/18/11 21:08:09 host/full.qualified.domain.name@REALM.TLD (ArcFour with HMAC/md5)
```



Proxy User

Някои продукти като **Apache Oozie**, които имат достъп до услугите на **Hadoop** от името на крайните потребители, трябва да могат да се представят за крайни потребители. Можете да конфигурирате потребител на прокси, като използвате свойствата **hadoop.proxyuser.\$superuser.hosts** заедно с едното или и двете от **hadoop.proxyuser.\$superuser.groups** и **hadoop.proxyuser.\$superuser.users**. Например, както е посочено по-долу в **core-site.xml**, потребителят на име **oozie**, който получава достъп от всеки хост, може да се представя за всеки потребител, принадлежащ към някоя група.

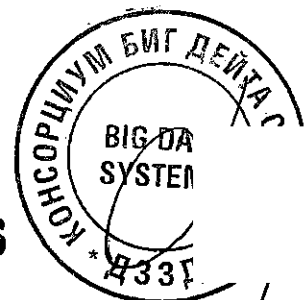
```
<property>
  <name>hadoop.proxyuser.oozie.hosts</name>
  <value>*</value>
</property>
<property>
  <name>hadoop.proxyuser.oozie.groups</name>
  <value>*</value>
</property>
```

Потребител на име **oozie**, който получава достъп от всеки хост, може да се представя за **user1** и **user2**, като посочва, както по-долу в **core-site.xml**

```
<property>
  <name>hadoop.proxyuser.oozie.hosts</name>
  <value>*</value>
</property>
<property>
  <name>hadoop.proxyuser.oozie.users</name>
  <value>user1,user2</value>
</property>
```

hadoop.proxyuser.\$superuser.hosts приема списък от IP адреси, обхвати в **CIDR** формат и/или имена на хостове. Например като посочва, както по-долу в **core-site.xml**, потребител на име **oozie** достъпва от хостове в обхват 10.222.0.0-15 и 10.113.221.221 може да се представи за всеки потребител, принадлежащ към някоя група.

```
<property>
  <name>hadoop.proxyuser.oozie.hosts</name>
  <value>10.222.0.0/16,10.113.221.221</value>
</property>
<property>
  <name>hadoop.proxyuser.oozie.groups</name>
  <value>*</value>
</property>
```



000.26

h

Ръководство на администратора за Създаване на
права за достъп за сървъри и сегменти от данни в
Nadoor система на вече създадени потребители



000527

5

Подробно описание на потребителски роли за Cloudera администратори.

Потребителските профили включват ролята на потребителя, която определя функциите **Cloudera Manager**, видими за потребителя, и действията, които потребителят може да изпълни. Всички задачи в документацията на **Cloudera Manager** указват коя роля е необходима за изпълнение на задачата. Потребителски акаунт в **Cloudera Manager** може да получи една от следните роли със съответните права:

1. Auditor

- Наблюдение на информация за конфигурация и мониторинг в **Cloudera Manager**.
- Наблюдение на одиторски събития.

2. Read-Only

- Наблюдение на информация за конфигурация и мониторинг в **Cloudera Manager**.
- Наблюдение на информация за услуги и мониторинг.
- Наблюдение на събития и логове.
- Наблюдение на задачи за репликация и snapshot политики.
- Наблюдение на **YARN** приложения и **Impala** заявки.

Read-Only ролята не позволява на потребителя:

- Да добавя услуги, или предприема действия, които засягат състоянието на клъстера.
- Да използва **HDFS** файлов браузър.
- Да използва **HBase** табличен браузър.
- Да използва **Solr Collection Statistics** браузър.

3. Dashboard

- Създаване, редактиране, или премахване на изгледи, принадлежащи на потребителя.
- Добавяне на съществуваща графа, или на нова графа към изгледа, принадлежащ на потребителя.
- Изпълнение на същите задачи като **Read-Only** ролята.

4. Limited Operator

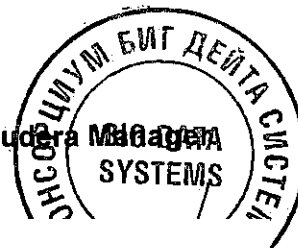
- Наблюдение на информация за конфигурация и мониторинг в **Cloudera Manager**.
- Наблюдение на информация за услуги и мониторинг.
- Извеждане на сървъри от експлоатация (освен сървъри изпълняващи **Cloudera Management Service** роли).
- Изпълнение на същите задачи като **Read-Only** ролята.
- **Limited Operator** ролята не позволява на потребителя да добавя услуги, или предприема действия, които засягат състоянието на клъстера.

5. Operator



000528

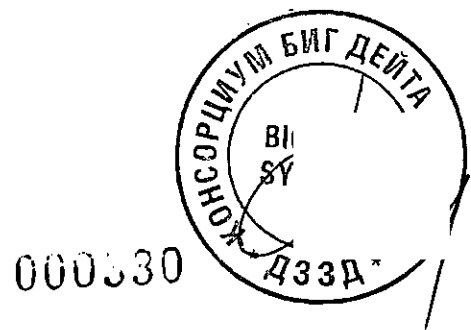
- ✓
- Наблюдение на информация за конфигурация и мониторинг на **Cloudera Manager**.
 - Наблюдение на информация за услуги и мониторинг.
 - Стартиране, прекратяване и рестартиране на клъстър, услуги (освен **Cloudera Management Service**) и роли.
 - Извеждане от експлоатация и преназначаване на сървъри (освен сървъри изпълняващи **Cloudera Management Service** роли).
 - Извеждане от експлоатация и преназначаване на роли (освен **Cloudera Management Service** роли).
 - Стартиране, прекратяване и рестартиране на **KMS**.
 - Изпълнение на същите задачи като **Read-Only** ролята.
 - **Operator** ролята не позволява на потребителя да добавя услуги, или предприема действия, които засягат състоянието на клъстера.
6. Configurator
- Наблюдение на информация за конфигурация и мониторинг на **Cloudera Manager**.
 - Изпълнение на всички операции на ролята **Operator**.
 - Настройка на услуги (освен на **Cloudera Management Service**).
 - Вход и изход от режим за поддръжка.
 - Управление на изгледи (включително **Cloudera Management Service** изгледи).
 - Стартиране, прекратяване и рестартиране на **KMS**
 - Изпълнение на същите задачи като **Read-Only** ролята.
7. Cluster Administrator — Преглед на всички данни и изпълнение на всички действия, **освен** следните:
- Администрация на **Cloudera Navigator**.
 - Наблюдение на графици за репликация и snapshot политики.
 - Наблюдение на одиторски събития.
 - Управление на потребителски акаунти и конфигурация на външна автентикация.
 - Управление на акаунти с **Full Administrator** роля.
 - Конфигурация на **HDFS** криптиране, администрация на Key Trustee сървър и управление на ключове за криптиране.
 - Използване на **HDFS** файлов браузър, **HBase** табличен браузър и **Solr Collection** браузър.
 - Изпълнение на същите задачи като **Read-Only** ролята.
8. BDR Administrator
- Наблюдение на информация за конфигурация и мониторинг на **Cloudera Manager**.
 - Наблюдение на информация за услуги и мониторинг.
 - Извършване на репликация и дефиниране на snapshot операции.



000329

- Използване на **HDFS** файлов браузър, **HBase** табличен браузър и **Solr Collection** браузър.
 - Изпълнение на същите задачи като **Read-Only** ролята.
9. Navigator Administrator
- Наблюдение на информация за конфигурация и мониторинг на **Cloudera Manager**.
 - Наблюдение на информация за услуги и мониторинг.
 - Администриране на **Cloudera Navigator**.
 - Наблюдение на одиторски събития.
 - Използване на **HDFS** файлов браузър, **HBase** табличен браузър и **Solr Collection** браузър.
 - Изпълнение на същите задачи като **Read-Only** ролята.
10. User Administrator
- Наблюдение на информация за конфигурация и мониторинг на **Cloudera Manager**.
 - Наблюдение на информация за услуги и мониторинг.
 - Управление на потребителски акаунти и конфигурация на външна автентикация.
 - Използване на **HDFS** файлов браузър, **HBase** табличен браузър и **Solr Collection** браузър.
 - Изпълнение на същите задачи като **Read-Only** ролята.
11. Key Administrator
- Наблюдение на информация за конфигурация и мониторинг на **Cloudera Manager**.
 - Конфигурация на **HDFS** криптиране, администрация на **Key Trustee** сървър и управление на ключове за криптиране.
 - Стартиране, прекратяване и рестартиране на **KMS**
 - Използване на **HDFS** файлов браузър, **HBase** табличен браузър и **Solr Collection** браузър.
 - Изпълнение на същите задачи като **Read-Only** ролята.
12. Full Administrator – Потребители с **Full Administrator** ролята имат права да наблюдават всички данни и да изпълняват всички действия, включително реконфигуриране и рестартиране на услуги и администриране на други потребители.

Подход за създаване съответното изискване с посочено техническо и програмно-конфигурационно решение:



Назначаване на потребителски роли

Според ролята на съответния потребител в организационната структура, може да му бъде назначена потребителска роля. Потребителските роли се назначават посредством следните конфигурационни стъпки:

1. Изберете **Administration > Users**.
2. Щракнете върху отметката до едно или повече потребителски имена.
3. Изберете **Actions for Selected > Assign User Roles**.
4. В падащото меню изберете роля.
5. Щракнете върху бутона **Assign Role**.

Пример за прилагане на изискването:

Ако длъжността на потребителя включва само наблюдение на информация без предприемане на действия в рамките на клъстера, следва да му се назначи **Read-Only** потребителска роля.

1. NameNode

NameNode инстанциите поддържат структурата на **HDFS** и правят връзката между файловете блокове и **DataNode** инстанциите, където се съхраняват данните. Един обикновен **HDFS** клъстер може да има само един основен **NameNode**, поддържан от вторичен **NameNode**, който периодично компресира регистрационния файл за редактиране на **NameNode**, който съдържа списък на модификации на метаданни на **HDFS**. Това намалява количеството дисково пространство, консумирано от регистрационния файл на **NameNode**, което също намалява времето за рестартиране на основния **NameNode**. **High Availability** клъстерът съдържа две инстанции на **NameNode**: активен и в режим на готовност.

2. DataNode

DataNodes инстанциите съхраняват данни в Hadoop клъстера и са отговорни за управлението им. Файловете данни се репликират на набор от инстанции на **DataNode** за надеждност и, за да може локализирано изчисление да бъде изпълнено близо до данните. Факторът за репликация по подразбиране в **HDFS** е три. Това означава, че три копия на даден обем от данни, се поддържат постоянно (Препоръчително е да не се конфигурира по-ниска стойност на фактора за репликация при наличие на три или повече инстанции на **DataNode**. По-ниска стойност може да доведе до значителна загуба на данни).

В рамките на клъстера, инстанциите на **DataNode** трябва да са еднородни. Ако не са еднородни, могат да възникнат проблеми. Например, **DataNode** инстанции с по-малко заделена памет я запълват по-бързо от инстанции на **DataNode** с повече памет, което може да доведе до неизпълнение на задачите.

3. Node Manager

NodeManager (NM) е агент на **YARN** за всяко устройство и се грижи за отделните изчислителни устройства в Hadoop клъстер. Това включва поддържане на актуална връзка с **ResourceManager (RM)**, наблюдение на управлението на жизнения цикъл на контейнерите; наблюдение на използването на ресурсите (памет, CPU) на отделни контейнери, проследяване на здравето, управление на логове и допълнителни услуги, които могат да се използват от различни **YARN** приложения.

4. JobTracker

JobTracker е услугата, която извежда задачите на **MapReduce** към специфични устройства в клъстера, в идеалния случай устройствата, които имат данните, или поне са в една и съща организационна единица. Клиентските приложения изпращат задания на **JobTracker**. **JobTracker** говори с **NameNode**, за да определи местоположението на данните.

JobTracker локализира **TaskTracker** възли с налични слотове на или близо до данните и предава работата на избраните **TaskTracker** възли. Възлите на **TaskTracker** се наблюдават. Ако те не подават достатъчно често сигнали за оперативно състояние, те се считат за неуспешни и работата е насрочена за друг **TaskTracker**. **TaskTracker** ще уведоми **JobTracker**, когато дадена задача не успее. **JobTracker** решава какво да прави след това: той може да изпрати отново работата на друго място, може да отбележи този конкретен случай като нещо, което да избегне, и може дори да го включи черния списък на **TaskTracker** като ненадежден. **JobTracker** е точка на неуспех за всички операции на **MapReduce**. Ако той спре да функционира, всички работещи задачи се спират.

YARN е проектиран да раздели функционалностите на **JobTracker** на отделни демони, които отделят управлението на ресурсите и планирането на задачите от изпълнението, мониторинга и осигуряването на поток за изпълнение при отказ. Отделните демони биват:

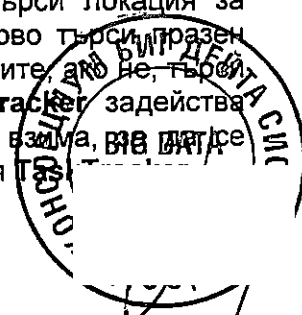
- **ResourceManager (RM)**: за управление и алокиране на ресурси и насрочване на приложения
- **ApplicationMaster (AM)**: за изпълнение на приложение и осигуряване на поток за изпълнение при отказ

Това разделяне на отговорностите позволява по-добро използване на ресурсите на клъстера и способността да се изпълняват приложения извън **MapReduce**.

Има един **RM** на ниво на клъстера, докато има един **AM** на приложение.

5. TaskTracker

TaskTracker е възел в клъстера, който приема задачи - **Map**, **Reduce** и **Shuffle** операции - от **JobTracker**. Всеки **TaskTracker** е конфигуриран с набор от слотове, които индикират броят на задачите, който може да бъде приет. Когато **JobTracker** търси локация за насрочване на приложение в рамките на **MapReduce** операциите, първо търси празен слот на същия сървър, на който е разположен **DataNode** съдържащ данните, а ако не, търси празен слот на машина в същата организационна структура. **TaskTracker** задейства отделни JVM процеси за изпълнението на задачата: тази мярка се взема, за да се подсигури, че евентуален неуспех на задачата няма да повлияе на самия **TaskTracker**.



Ръководство на администратора за Създаване
контрол на достъпа до HDFS директории и
файлове на Hadoop система

000533



D *A.* *S*

The Hadoop Distributed File System (HDFS) реализира модел за разрешения за файлове и директории, които споделят голяма част от POSIX модела. Всеки файл и директория е свързан със собственик и група. Файлът или директорията имат отделни разрешения за потребителя, който е собственик, за други потребители, които са членове на групата, както и за всички други потребители. За файловете се изисква разрешението *r* за четене на файла и се изисква разрешението *w* за записване или добавяне към файла. За директории се изисква разрешението *r*, за да се изброят съдържанието на директорията, за да се създадат или изтрият файлове или директории, и се изисква разрешение *x* за достъп до поддиректория.

За разлика от модела POSIX, за файловете няма **setuid** или **setgid** бита, тъй като няма вградена представа за изпълними файлове. За директории няма директория **setuid** или **setgid** бита за улеснение. **Sticky** бит може да бъде настроен на директории, като по този начин не се позволява на никой друг, освен собственика на суперпотребителя, собственика на директорията или собственика на файла, да изтрива или премества файловете в директорията. Задаването на **sticky** бит за файл няма ефект. Като цяло, разрешенията на даден файл или директория са неговият режим. Като цяло ще се използват **Unix Customs** за представяне и показване на режими, включително използването на осмични числа в това описание. Когато е създаден файл или директория, неговият собственик е потребителската идентичност на процеса на клиента, а неговата група е групата на родителската директория (правило **BSD**). Всеки клиентски процес, който осъществява достъп до **HDFS**, има идентичност от две части, съставена от потребителското име и списъка с групи. Всеки път, когато **HDFS** трябва да извърши проверка за разрешения за файл или директория, достъпни от клиентски процес,

- Ако името на потребителя съвпада със собственика на **foo**, тогава се тестват разрешенията на собственика;
- Иначе, ако групата от **foo** съвпада с някой от членовете на списъка с групи, тогава се тестват груповите разрешения;
- В противен случай се тестват останалите разрешения на **foo**.

Ако проверката на разрешения е неуспешна, операцията на клиента е неуспешна.

Идентичност на потребителя

От Hadoop 0.22, Hadoop поддържа два различни режима на работа, за да определи самоличността на потребителя, определена от свойството **hadoop.security.authentication**:

- **Simple**

В този режим на работа идентичността на клиентския процес се опр



операционната система на хоста. В Unix-подобни системи, потребителското име е еквивалентно на `'whoami'`.

- **kerberos**

В **Kerberized** операцията идентичността на клиентския процес се определя от нейните Kerberos идентификационни данни. Например, в една **Kerberized** среда, потребителят може да използва помощната програма **kinit**, за да получи **Kerberos** билет за издаване на билети (TGT) и да използва **klist**, за да определи текущия си принцип. Когато свързвате основния **Kerberos** към **HDFS** потребителско име, всички компоненти, с изключение на основния, се премахват. Например, главното `todd/foobar@CORP.COMPANY.COM` ще действа като просто потребителско име `todd` на **HDFS**.

Независимо от режима на работа, механизмът за идентифициране на потребителя е външен за самия **HDFS**. В **HDFS** няма разпоредба за създаване на потребителски идентичности, създаване на групи или обработка на потребителски идентификационни данни.

Проверки за разрешение

Всяка **HDFS** операция изисква потребителят да има специфични разрешения (някаква комбинация от **READ**, **WRITE** и **EXECUTE**), предоставена чрез собственост на файла, членство в група или други разрешения. Една операция може да извърши проверки за разрешение на множество компоненти на пътя, а не само на крайния компонент. Освен това, някои операции зависят от проверка на собственика на пътя.

Всички операции изискват достъп за пътя. Траверсалният достъп изисква разрешение **EXECUTE** за всички съществуващи компоненти на пътя, с изключение на крайния компонент на пътя. Например, за всяка операция с достъп до `/foo/bar/baz`, повикващият трябва да има разрешение **EXECUTE** за `/`, `/foo` и `/foo/bar`.

Следната таблица описва проверките за разрешения, извършени от **HDFS** за всеки компонент на пътя.

- **Ownership:** Дали да проверите дали повикващият е собственик на пътя. Обикновено операциите, които променят метаданните за собственост или разрешение, изискват обажданият се да е собственик.
- **Parent:** Родителската директория на заявения път. Например, за пътя `/foo/bar/baz`, родителят е `/foo/bar`.
- **Ancestor:** Последният съществуващ компонент на заявения път. Например, за пътя `/foo/bar/baz`, пътят на предшественика е `/foo/bar`, ако `/foo/bar` съществува. Пътят на предшественика е `/foo` ако `/foo` съществува, а пътят на предшественика е `/` ако `/` съществува.



- **Final:** Последният компонент на заявения път. Например, за пътя / foo / bar / baz, последният компонент на пътя е / foo / bar / baz.
- **Sub-tree:** За път, който е директория, самата директория и всичките му поддиректории, рекурсивно. Например, за пътя / foo / bar / baz, който има 2 поддиректории, наречени buz и boo, под-дървото е / foo / bar / baz, / foo / bar / baz / buz и / foo / bar / baz / boo.

[1] **WRITE** достъпът до крайния компонент на пътеката по време на създаването е необходим само ако повикването използва опцията за презапис и има съществуващ файл на пътя.

[2] Всяка операция, която проверява разрешението **WRITE** в родителската директория, също проверява собствеността, ако е зададен **sticky** бит.

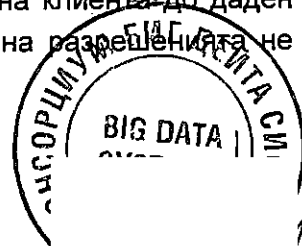
[3] Извикването на **setOwner** за промяна на потребителя, който притежава файл, изисква достъп до суперпотребител на **HDFS**. Суперпотребителският достъп на **HDFS** не се изисква за промяна на групата, но повикващият трябва да е собственик на файла и член на посочената група.

Подход за създаване съответното изискване с посочено техническо и програмно-конфигурационно решение:

Вникване в имплементацията

Разбирането на **ImplementEach** файла или операцията на директорията предава пълното име на пътя към **NameNode** и проверките за разрешения се прилагат по пътя за всяка операция. Клиентският **framework** ще асоциира имплицитно потребителската идентичност с връзката с **NameNode**, намалявайки необходимостта от промени в съществуващия клиентски **API**. Винаги е било така, че когато една операция върху даден файл успее, операцията може да се провали, когато се повтори, защото файлът или някоя директория на пътя вече не съществува. Например, когато клиентът първо започне да чете файл, той прави първа заявка към **NameNode**, за да открие местоположението на първите блокове на файла. Втората заявка за намиране на допълнителни блокове може да бъде неуспешна. От друга страна, изтриването на файл не отменя достъп от клиент, който вече знае блоковете на файла. С добавянето на разрешения, достъпът на клиента до даден файл може да бъде оттеглен между заявките. Отново, промяната на разрешенията не отменя достъпа на клиент, който вече знае **blockstation** на файла.

Промени на **API** на файловата система



000036

Всички методи, които използват параметър на път, ще активират `AccessControlException`, ако проверката на правата не успее

Нови методи:

- `public FSDataOutputStream create(Path f, FsPermission permission, boolean overwrite, int bufferSize, short replication, long blockSize, Progressable progress)` активира `IOException`;
- `public boolean mkdirs(Path f, FsPermission permission)` активира `IOException`;
- `public void setPermission(Path p, FsPermission permission)` активира `IOException`;
- `public void setOwner(Path p, String username, String groupname)` активира `IOException`;
- `public FileStatus getFileStatus(Path f)` активира `IOException`; допълнително връща потребителя, групата и режима, свързани с този път.

Режимът на нов файл или директория е ограничен от `umask` набор като конфигурационен параметър. Когато се използва съществуващия метод за създаване (път, ...) (без параметъра за разрешение), режимът на новия файл е `0666 & ^ umask`. Когато се използва новия метод на създаване (път, разрешение, ...) (с параметър на разрешение `P`), режимът на новия файл е `P & ^ umask & 0666`. Когато се създава нова директория с съществуващия метод на `mkdirs (path)` (без параметъра за разрешение), режимът на новата директория е `0777 & ^ umask`. Когато се използва новия метод `mkdirs (път, разрешение)` (с параметър на разрешение `P`), режимът на новата директория е `P & ^ umask & 0777`.

Промени на Application Shell

Нови операции:

- `chmod [-R] mode file ...`

Само собственикът на файл или суперпотребителят могат да променят режима на файл.

- `chgrp [-R] group file ...`

Потребителят, който извиква `chgrp`, трябва да принадлежи на посочената група и да е собственик на файла или да бъде суперпотребител.

- `chown [-R] [owner][:[group]] file ...`

Собственикът на даден файл може да бъде променян само от суперпотребител



- ls file ...
- lsr file ...

Изходът се преформатира, за да се покаже собственикът, групата и режимът.

Суперпотребителят

Суперпотребителят е потребител със същата идентичност като самия процес на **NameNode**. Иначе поставено, ако сте започнали **NameNode**, тогава сте суперпотребител. Суперпотребителят може да направи всичко, за което проверките на разрешения никога не се провалят за суперпотребителя. Няма постоянна представа кой е суперпотребителят; когато е стартиран **NameNode**, идентичността на процеса определя кой за сега е суперпотребителят. Суперпотребителят **HDFS** не трябва да бъде суперпотребител на **NameNode** хоста, нито е необходимо всички клъстери да имат един и същ суперпотребител. Също така, експериментатор, изпълняващ **HDFS** на лична работна станция, удобно става суперпотребител на тази инсталация без никаква конфигурация.

В допълнение, администраторът може да идентифицира разграничена група, като използва конфигурационен параметър. Ако са определени, членовете на тази група също са суперпотребители.

Уеб Сървърът

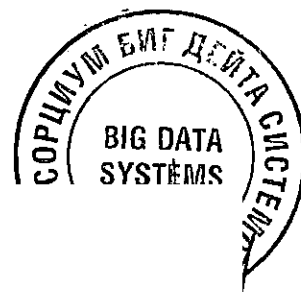
По подразбиране самоличността на уеб сървъра е конфигурационен параметър. Това означава, че **NameNode** няма представа за идентичността на реалния потребител, но уеб сървърът се държи така, сякаш има идентичност (потребител и групи) на потребител, избран от администратора. Освен ако избраната идентичност не съвпада със суперпотребителя, части от namespace могат да бъдат недостъпни за уеб сървъра.

Пример за прилагане на изискването:

1. Създайте OS акаунт на Linux система, от която искате да позволите на потребител да изпълнява Hadoop задания. Преди да създадете потребителя, може да се наложи да създадете група. (В този случай **analysts** е OS група, която сме създали за набор от потребители. Командата **passwd** ни позволя да настроим парола за потребителя.

```
$ group add analysts
```

```
$ useradd -g analysts alapati
```



000038

\$ passwd alapati

2. Уверете се, че сте настроили разрешенията в **temp** директорията на **Hadoop**, която сте настроили в **core-site.xml** файла така, че всички **Hadoop** потребители да могат да я достъпят.

<property>

<name>hadoop.tmp.dir</name>

<value>/tmp/hadoop-\$(user.name)</value>

</property>

3. Ако разрешенията за файлове в **temp** директорията не са **777**, ги направете така:

```
$ hdfs -dfs -chmod -R 777 //tmp/hadoop-alapati
```

4. За да „създадете“ нов **HDFS** потребител, трябва да създадете директория под **/user**. Тази директория ще служи като "home" директория на **HDFS** за потребителя.

```
$ hdfs dfs -mkdir /user/alapati
```

5. По подразбиране, когато създавате директория или файл, собственикът е потребителят, който създава директорията (или файла) и групата е групата на този потребител, както е показано тук.

```
# sudo -u hdfs
```

```
# hdfs dfs -ls /user
```

```
Found 135 items
```

```
drwxr-xr-x - hdfs supergroup  
/user/alapati
```



000039

-
6. Можете да проверите новата структура на директориите за потребителя със следната команда:

```
$ hdfs dfs -ls /user/alapati
```

Потребителят alapati вече може да съхранява изхода на своите MapReduce и други задачи под home директорията на този потребител в HDFS.

7. Обновете съпоставките (mappings) на потребителите и групите, за да може NameNode модула да усвои новия потребител.

```
$ hdfs dfsadmin -refreshUserToGroupMappings
```

8. Задайте пространствена квота за новата директория, която сте създали:

```
$ hdfs dfsadmin -setSpaceQuota 30g /user/alapati
```



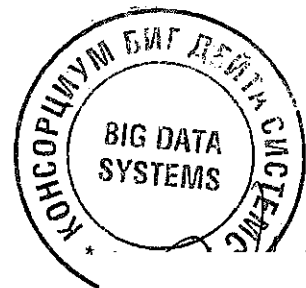
000040

✓

✓

✓

Ръководство на администратора за Създаване
централизиран Лист за управление на достъпа до
Hadoop система



000041

HDFS поддържа POSIX списъци за контрол на достъпа (ACLs), както и традиционния POSIX модел за права за достъп. ACL списъците контролират достъп до HDFS файлове като осигуряват начин да се разпределят различни права на определени потребители, или определени групи потребители. Те подобряват традиционния модел като позволят на специфични права да бъдат дефинирани за персонализируеми комбинации от потребители, или групи, вместо един определен потребител, или група.

Подход за създаване съответното изискване с посочено техническо и програмно-конфигурационно решение:

Активиране на HDFS списъци за контрол на достъпа (ACLs) през Cloudera Manager

1. Навигирайте до Cloudera Manager Admin конзолата и локализирайте HDFS услугата.
2. Щракнете върху раздела **Configuration**.
3. Изберете **Scope > Service_name (Service-Wide)**
4. Изберете **Category > Security**
5. Навигирайте до **Enable Access Control Lists** настройката и отбележете отметката, за да активирате HDFS ACLs.
6. Щракнете върху **Save Changes**, за да приложите промените.

Активиране на HDFS списъци за контрол на достъпа (ACLs) през командния ред

За да активирате списъците за контрол на достъпа през командния ред, променете `dfs.namenode.acls.enabled` настройката на `true` в конфигурационния файл на **NameNode** `hdfs-site.xml`.

```
<property>
<name>dfs.namenode.acls.enabled</name>
<value>true</value>
</property>
```

Команди

За да настроите и осигурите списъци за контрол на достъпа (ACLs), командите на файловата система `setfacl` и `getfacl`.



000342

getfacl

```
hdfs dfs -getfacl [-R] <path>
```

<!-- COMMAND OPTIONS

<path>: Път до файла, или директорията, за които да бъдат изброени ACLs.

-R: Използвайте тази опция, за да изброите рекурсивно ACLs за всички файлове и директории.

-->

Примери за прилагане на изискването:

<!-- За да изброите всички ACLs за файл, намиращ се в директория /user/hdfs/file -->

```
hdfs dfs -getfacl /user/hdfs/file
```

<!-- За да изброите рекурсивно всички ACLs за user/hdfs/file -->

```
hdfs dfs -getfacl -R user/hdfs/file
```

setfacl

```
hdfs dfs -setfacl [-R] [-b|-k|-m|-x <acl_spec> <path>][[--set <acl_spec> <path>]
```

<!-- COMMAND OPTIONS

<path>: Път до файла или директорията, за които трябва да бъдат зададени ACLs.

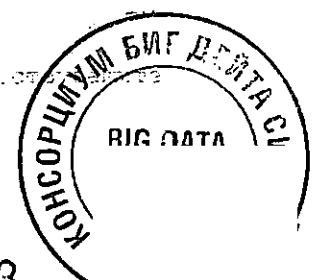
-R: Използвайте тази опция за рекурсивно изброяване на ACL за всички файлове и директории.

-b: Отменете всички разрешения, с изключение на базовите ACL за потребители, групи и други.

-k: Премахване на ACL от потребител.

-m: Добавяне на разрешения към ACL, тази опция. Не работи с опцията --set.

-x: Премахване на разрешения от ACL.



000043

MA

7

--set: Използвайте тази опция, за да замените напълно съществуващия ACL за посочения път.

Предишните записи на ACL вече няма да се прилагат.

<!-- За да назначите на потребител ben read & write права на /user/hdfs/file -->

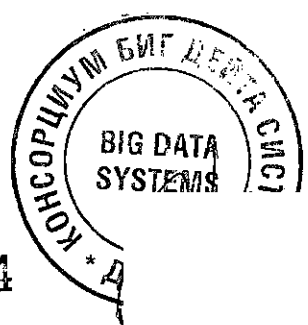
hdfs dfs -setfacl -m user:ben:rw- /user/hdfs/file

<!-- За да премахнете ACL записа на потребител alice за /user/hdfs/file -->

hdfs dfs -setfacl -x user:alice user/hdfs/file

<!-- За да назначите на потребител hadoop read & write права, а на group или others read-only права -->

hdfs dfs -setfacl --set user::rw-,user:hadoop:rw-,group::r--,other::r-- /user/hdfs/file



000.44

1

✓

2

75

Ръководство на администратора за Създаване на
механизъм за пълен одит на Hadoop система

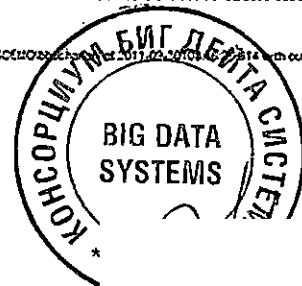


000545

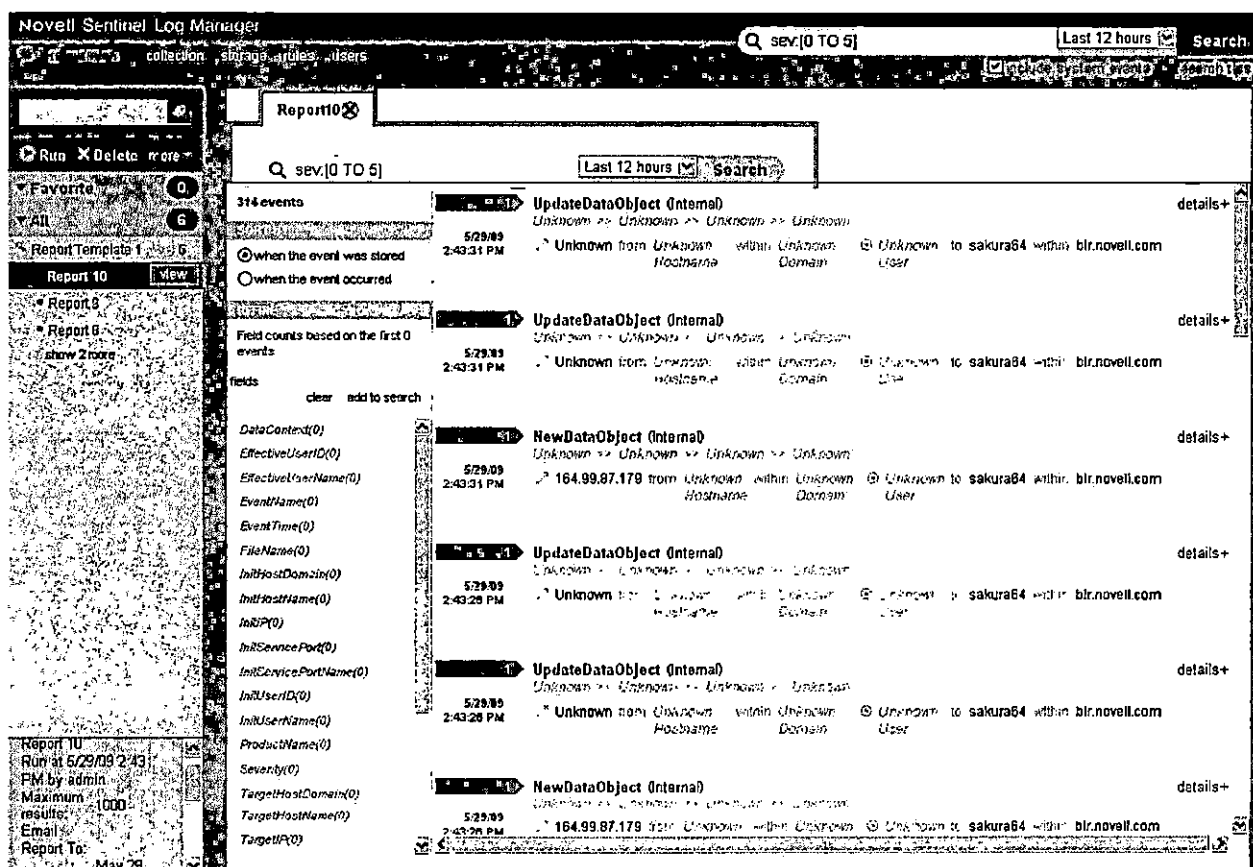
Подход за създаване съответната одитна система с 3 параметъра - произход на данни, достъп до данни и идентификация на потребител, с посочено техническо и програмно-конфигурационно решение;

Одитна система на Hadoop решението на NetIQ Sentinel. NetIQ® Sentinel™ осигурява на организациите визуализация в реално време в целия спектър на ИТ дейностите, за да се намалят заплахите за сигурността, да се подобрят операциите и автоматично да се приложат контролите на политиките във физическа, виртуална и облачна среда. Тя намалява сложността на традиционното управление на информацията и сигурността на събития (SIEM) и намалява бариерите пред приемането на SIEM, като прави информацията по сигурността достъпна за всички организации. NetIQ Sentinel също така предоставя на организациите по-ефективно решение SIEM чрез комбиниране на разузнавателна информация в реално време, откриване на аномалии и наблюдение на активността на потребителите, за да се осигури механизъм за ранно предупреждение и по-точна оценка на ИТ дейностите.

The screenshot displays the NetIQ Sentinel web interface. On the left, there is a navigation menu with options like 'Security Intelligence', 'Reports', and 'People'. The main area shows a search query: 'Q (sev:(3 TO 5)) AND (uname:"Brad Buchanan")'. Below the search bar, there is a list of events. The first event is 'File was read (Unmanaged) (File Integrity Monitoring: NetIQ Change Guardian)' with a severity of 81. The event details include the user 'houqdemo001', the file path 'Employee Salaries.tbl', and the outcome 'Success'. A second event is 'File was read (Unmanaged) (File Integrity Monitoring: NetIQ Change Guardian)' with a severity of 82. A third event is 'File was written (Unmanaged) (File Integrity Monitoring: NetIQ Change Guardian)' with a severity of 83. On the left side of the interface, there is a profile for Brad Buchanan, an Information Technology Manager, with contact information and a list of accounts belonging to him.



В случая имаме създаден колектор за данните от логовете на NameNode и DataNode на Hadoop, както и за логовете за достъп на операционната система RHEL. За всеки потребител в системата се създава профил, който позволява коректното одитиране и персонализиране на данните: Създаден е корелационен механизъм за известяване, базиран на следните параметри— произход на данните, достъп до данните като профил на потребителя, IP адрес, дата и час и типа на операция. Пример за визуализация:



Сентинел е заложен като основно решение в цялата екосистема на Hadoop / Cloudera като високата степен на сигурност се гарантира от интеграцията между МУД и УАЛС.



00004

7

✓

**Ръководство на администратора за Осигуряване на
защита и криптиране на данните в Hadoop система
– криптиране на данни намиращи се върху диск и
криптиране на данни в движение**



000548 ✓

HDFS първоначално поддържа криптиране на данни чрез механизъм, наречен зони за криптиране. Зона за криптиране представлява HDFS директория, която е свързана с ключ за криптиране. След като директорията е свързана с ключа за криптиране, всички файлове в директорията и поддиректорията ще бъдат автоматично кодирани.

Често срещано погрешно схващане за HDFS криптирането е убеждението, че данните са криптирани, когато се записват на диск в информационните възли като повечето решения за криптиране на дискове. Всъщност данните се криптират, преди да бъдат изпратени на възела за данни. Тази архитектура има два хубави странични ефекта: единият е, че данните са защитени при транзит, а другият е, че също така предотвратява излагането на ключовете на възлите на данните, където се съхраняват данните на риск.

Криптиране на данни в движение (Data-in-Transit)

Transport Layer Security (TLS) е индустриален стандартен набор от криптографски протоколи за осигуряване на комуникации по мрежа. TLS се развива от Secure Sockets Layer (SSL). Тъй като SSL терминология все още се използва широко, софтуерът и документацията на Cloudera се отнасят за TLS както за TLS / SSL, но реалният използван протокол е TLS. SSL не се използва в Cloudera софтуера.

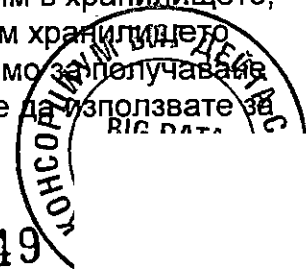
В допълнение към TLS / SSL криптиране, HDFS и HBase трансферират данни чрез remote procedure calls (RPC). За да осигурите това прехвърляне, трябва да активирате RPC криптирането.

TLS / SSL осигурява сигурност и защита на данните между приложенията, които комуникират по мрежата, като се криптират пакетите, предавани между крайните точки (портове на хост, например). Конфигурирането на TLS / SSL за всяка система обикновено включва създаване на частен ключ и публичен ключ за използване от процесите на сървъра и клиента за осъществяване на криптирана връзка по време на изпълнение. В допълнение, TLS / SSL може да използва сертификати за проверка на надеждността на ключовете, представени по време на осъществяването, за предотвратяване на спуфинг и смекчаване на други потенциални проблеми със сигурността.

Подход за създаване на съответното изискване с посочено програмно-конфигурационно решение:

Създаването на Cloudera кълъстери за използване на TLS / SSL изисква създаването на частен ключ, публичния ключ и съхраняването им в хранилището, наред с други задачи. Въпреки че добавянето на сертификат към хранилището може да бъде последната задача в процеса, времето, необходимо за получаване на сертификат, зависи от вида на сертификата, който планирате да използвате за кълъстера.

000-49



Бележка: TLS/SSL за основните услуги на Hadoop—HDFS, MapReduce, и YARN— трябва да бъде конфигурирано за група. TLS/SSL за други компоненти като HBase, Hue, и Oozie може да бъде конфигурирано отделно.

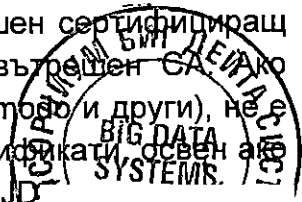
Сертификатът е цифрово подписан, обикновено от сертифициращ орган (CA), който косвено (чрез верига от доверие) проверява автентичността на публичния ключ, представен по време на преговорите. Сертификатите могат да бъдат подписани по един от трите различни начина, показани в таблицата:

Тип	Бележки за използване
Публични подписани сертификати	Препоръчително. Този тип сертификат се подписва от публичен сертифициращ орган (CA), като Symantec или Comodo. Обществените CA са надеждни трети страни, чиито сертификати могат да бъдат проверени чрез обществено достъпни вериги на доверие. Използването на този тип сертификати може да опрости внедряването, тъй като инфраструктурата за сигурност, като например root кодове, вече се съдържат в Java JDK по подразбиране.
Вътрешни подписани сертификати	Този тип сертификат се подписва от вътрешен CA на вашата организация. Организации, използващи OpenSSL Certificate Authority, Microsoft Active Directory сертификационна услуги, или друга вътрешна CA система могат да използват този тип сертификат.
Самоподписани сертификати	Не се препоръчва за производствени среди. Самоподписаните сертификати са приемливи за използване в непроизводствени инсталации, като например за настройка на доказателство за концепция (PoC).

По време на процеса на конфигуриране на TLS / SSL за клъстера, обикновено получавате сертификат за всеки хост в клъстера и повторно използвате сертификата, получен в даден формат (JKS, PEM), както е необходимо за различните услуги поддържани от host инстанцията.

Генериране на TLS сертификати

Следната процедура предполага, че се използва вътрешен сертифициращ орган (CA) и показва как да се установи доверие за този вътрешен CA, ако използвате доверен публичен CA (като Symantec, GeoTrust, Comodo и други), не е необходимо изрично да създавате доверие за издадените сертификати, освен ако не използвате по-стар JDK и по-нов обществен CA. По-старите JDK доверят на по-новите общес- то подразбиране.



Пример за прилагане на изискването:

На всеки хост в клъстера:

Извършете следната процедура за всеки хост в клъстера, включително и Cloudera Manager хоста.

1. Конфигурирайте работната си среда да зададе JAVA_HOME на JDK на Oracle.

```
export JAVA_HOME=/usr/java/jdk1.8.0_162
```

Ако излезете от хоста преди да завършите тази процедура, уверете се, че сте задали JAVA_HOME отново, когато влезете в системата, за да завършите стъпките.

2. Създайте директорията /opt/cloudera/security/pki

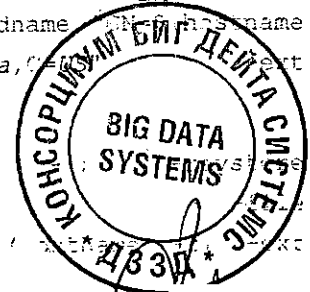
```
sudo mkdir -p /opt/cloudera/security/pki
```

3. Използвайте помощната програма keytool, за да генерирате заявка за хранилище на ключове на Java и заявка за подпис на сертификат (CSR). Заменете OU, O, L, ST и C записи със стойностите за вашата среда. Когато бъдете подканени, използвайте същата парола за хранилището на ключове и ключа. **Cloudera Manager** не поддържа различни пароли за ключа и хранилището на ключове.

```
sudo $JAVA_HOME/bin/keytool -certreq -alias $(hostname -f) -keystore /opt/cloudera/security/pki/$(hostname -f).jks -file /opt/cloudera/security/pki/$(hostname -f).csr
```

```
$JAVA_HOME/bin/keytool -genkeypair -alias $(hostname -f) -keyalg RSA -keystore /opt/cloudera/security/pki/$(hostname -f).jks -keysize 2048 -dname $(hostname -f),OU=Engineering,O=Cloudera,L=Palo Alto,ST=California,C=US
```

```
$JAVA_HOME/bin/keytool -certreq -alias $(hostname -f) -keystore /opt/cloudera/security/pki/$(hostname -f).jks -file /opt/cloudera/security/pki/$(hostname -f).csr -ext san=dns:$(hostname -f)
```



000351

4. Изпратете CSR файловете (например `cm01.example.com.csr`) на сертифициращия орган, за да получите сертификат за сървър. От съображения за сигурност много търговски СА пренебрегват исканите разширения в CSR. Уверете се, че информирате CO, че се нуждаете от сертификати с опции за удостоверяване на сървъра и клиента. Ако е възможно, получите сертификата в PEM (Base64 ASCII) формат. Файлт със сертификата е в PEM формат, ако изглежда подобен на този (някои редове са пропуснати):

-----BEGIN CERTIFICATE-----

```
MIIDAzCCAesCAQAwgY0xCzAJBgNVBAYTAlVTMRMwEQYDVQQTEwQwDlYwZm9ybm9hbnRl
NPTWEAYDVQQHEw1QYWxvIEFsdG9xETAPBgNVBAoTCENsb3VhZm9ybm9hbnRlYDVQQL
...
tudY0C32LjGjW0g5ALLiN9Oy1u2xRKGAVFapbzAZ2rcht1CZc7mtaT6BXgW8S+Db
0NhuObn1/STL4Ho9G+K1JB3Mwik2oEbOvQt0rBidMr9qaNX86m0i7pouXZelZ5c5
UnDEtrhW6A==
```

-----END CERTIFICATE-----

Ако издаденият ви сертификат е в бинарен формат (DER), конвертирайте го в PEM формат.

5. След като получите подписания сертификат, копирайте подписания сертификат на следното място:

```
opt cloudera/security/pki $(hostname -f).pem
```

6. Проверете подписания сертификат, за да проверите дали са налице опции за удостоверяване на сървъра и клиента, както и алтернативното име на обекта:

```
openssl x509 -in /opt/cloudera/security/pki/$(hostname -f).pem -noout
```

Потърсете изход, подобен на следния, за да проверите опциите за удостоверяване на сървъра и клиента:



000352

TLS Web Server Authentication, TLS Web Client Authentication

Потърсете изход, подобен на следния, за да потвърдите алтернативното име на обекта:

X509v3 Subject Alternative Name:

DNS:hostname.example.com

7. Копирайте кореновите и междинните сертификати на СА в `/opt/cloudera/security/pki/rootca.pem` и `/opt/cloudera/security/pki/intca.pem` на всеки хост. Ако имате свързващ файл, съдържащ основния СА и междинен сертификат на ЦА, разделете файла по границата `END CERTIFICATE / BEGIN CERTIFICATE` на отделни файлове. Ако има няколко междинни сертификата на СА, използвайте уникални имена на файлове като `intca-1.pem`, `intca-2.pem` и т.н.
8. Копирайте JDK файла `cacerts` в `jssecacerts` както следва:

```
sudo cp $JAVA_HOME/jre/lib/security/cacerts $JAVA_HOME/jre/lib/security/jssecacerts
```

Oracle JDK използва файла `jssecacerts` за своя доверителен файл по подразбиране, ако той съществува. В противен случай използва файла `cacerts`. Създаването на `jssecacerts` файл ви позволява да се доверите на вътрешен СА, без да променяте файла `cacerts`, който е включен в JDK.

9. Импортирайте root CA сертификата в JDK `truststore`.

```
sudo $JAVA_HOME/bin/keytool -importcert -alias rootca -keystore $JAVA_HOME/jre/lib/security/jssecacerts -file /opt/cloudera/security/pki/rootca.pem
```

10. Добавете междинния сертификат на СА към подписания сертификат за хост и след това го импортирайте в хранилището. Уверете се, че използвате оператора за добавяне (`>>`), а не оператора за презапис (`>`):

```
sudo cat /opt/cloudera/security/pki/intca.pem | keytool -importcert -alias intca -keystore $JAVA_HOME/jre/lib/security/jssecacerts -file /opt/cloudera/security/pki/intca.pem
```



11. Създаване на символни връзки (символна връзка) за файловете за сертификат и хранилище за ключове:

```
sudo ln -s /opt/cloudera/security/pki/${hostname} -f .pem  
/opt/cloudera/security/pki/agent.pem
```

Това ви позволява да използвате един и същ файл /etc/cloudera-scm-agent/config.ini на всички хостове на агенти, вместо да поддържате файл за всеки агент.

На Cloudera Manager сървърен хост:

На хоста на Cloudera Manager Server създайте допълнителна символна връзка за файла на хранилището:

```
sudo ln -s /opt/cloudera/security/pki/${hostname} -f .jks  
/opt/cloudera/security/pki/server.jks
```

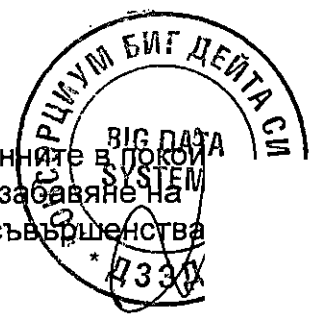
Криптиране на данни, намиращи се върху диск (Data-at-Rest)

HDFS криптирането прилага прозрачно, end-to-end криптиране на данни, които се четат и записват в HDFS, без да се налага промяна на кода на приложението. Тъй като криптирането е end-to-end, данните могат да бъдат криптирани и декриптирани само от клиента. HDFS не съхранява и няма достъп до некриптирани данни или ключове за шифроване. Това поддържа едновременно криптиране на данни в покой (данни за постоянни медии, като например диск) и криптиране на данни в движение (данни, пътуващи по мрежа).

Cloudera клъстерите могат да използват комбинация от механизми за криптиране, включително HDFS прозрачно криптиране и Cloudera Navigator Encrypt. И двата от тези механизми за криптиране могат да бъдат допълнени с управление на ключовете, използвайки Key Trustee Server на Cloudera Navigator и Key HSM на Cloudera Navigator.

Cloudera Navigator Encrypt

Cloudera Navigator Encrypt прозрачно криптира и обезопасява данните в покой без да изисква промени в приложенията и гарантира минимално забавяне на производителността в процеса на криптиране и декриптиране. Усъвършенства

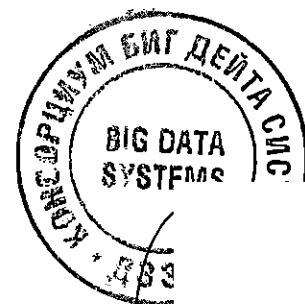


управление на ключове с Cloudera Navigator Key Trustee Server и процесно-базирани контроли за достъп в Navigator Encrypt дават възможност на организациите да спазват изискванията за съответствие и да гарантират, че неупълномощените страни и/или злонамерените такива няма да получат достъп до криптирани данни.

HDFS прозрачно криптиране

HDFS криптирането има тези възможности:

- Само HDFS клиенти могат да криптират и декриптират данни.
- Управлението на ключове е външно за HDFS. HDFS не може, да достъпи некриптирани данни или ключове за криптиране. Администрацията на HDFS и администрацията на ключовете са отделни длъжности, които се изпълняват от отделни потребителски роли (HDFS administrator, Key Administrator), като така се подсигурира, че няма единичен потребител с достъп до данните и до ключовете за криптиране.
- Операционната система и HDFS обменят само криптирани HDFS данни, като така отсяват заплахи на ОС и FileServer нива.
- HDFS използва Advanced Encryption Standard-Counter mode (AES-CTR) алгоритъма за криптиране. AES-CTR поддържа 128-битов ключ за криптиране (по подразбиране), или има възможност за поддръжка на 256-битов ключ за криптиране, когато Java Cryptography Extension (JCE) е инсталиран.
- HDFS криптирането е създадено, за да се възползва от набора от инструкции AES-NI - техниката за ускоряване на криптирането, базирана на хардуер, така че производителността на клъстера да не се влияе неблагоприятно от конфигурирането на криптирането. (Наборът инструкции AES-NI може да бъде с порядък по-бърз от софтуерните реализации на AES.) Въпреки това може да се наложи актуализация на криптографските библиотеки на HDFS и MapReduce клиентските хостове, за да се използва механизма за ускоряване.



000355

Ръководство на администратора за Изграждане
управление на ключовете за криптиране в Nadoor
система

000356



Взаимодействието с KMS и създаването на криптиращи зони изисква използването на две нови CLI команди: `hadoop key` и `hdfs crypto`.

Подход за създаване съответното изискване с посочено техническо и програмно-конфигурационно решение:

Проверка на Hadoop Key операции

Използвайте `hadoop key create`, за да създадете тестов ключ, след което `hadoop key list`, за да извлечете списъка с ключове:

```
$ sudo -u <key_admin> hadoop key create keytrustee_test  
$ hadoop key list
```

Създаване на зони за криптиране

След като е създаден KMS и клиентите на NameNode и HDFS са правилно конфигурирани, използвайте инструментите на командния ред `hadoop key` и `hdfs crypto` за създаване на ключове за криптиране и създаване на нови зони за криптиране.

- Създайте ключ за криптиране за вашата зона като `keyadmin` за потребителя / групата (независимо от приложението, което ще използва зоната за криптиране):

```
$ sudo -u hbase hadoop key create <key_name>
```

- Създайте нова празна директория и я направете зона за шифроване с помощта на създадения по-горе ключ.

```
$ sudo -u hdfs hadoop fs -mkdir /encryption_zone
```

```
$ sudo -u hdfs hdfs crypto -createZone -keyName <key_name> -path  
/encryption_zone
```

- Можете да проверите създаването на новата зона за криптиране като изпълните командата `-listZones`. Трябва да видите зоната за шифроване заедно с нейния ключ, както следва:

```
$ sudo -u hdfs hdfs crypto -listZones  
/encryption_zone <key_name>
```



000357

Предупреждение: Не изтривайте ключ за криптиране, докато той все още се използва за зона за криптиране. Това води до загуба на достъп до данни в тази зона. Също така, не изтривайте KMS услугата, тъй като криптираните HDFS данни ще бъдат недостъпни. За да предотвратите загуба на данни, първо копирайте криптираните HDFS данни в некриптирана зона, като използвате командата `distcp`.

Добавяне на файлове към зона за криптиране

Можете да добавяте файлове към зона за криптиране, като ги копирате в зоната за криптиране, като използвате `distcp`. Например:

```
sudo -u hdfs hadoop distcp /user/dir /encryption_zone
```

Съображения за DistCP

Често срещан случай на използване на DistCp е да се копират данни между клъстери за целите на архивирането и възстановяването при бедствия. Това обикновено се извършва от администратора на клъстера, който е суперпотребител на HDFS. За да се запази този работен процес, когато се използва HDFS криптиране, е въведен нов префикс за виртуален път, `/.reserved/raw/`, който дава директен достъп на суперпотребителите до основните блокови данни в файловата система. Това позволява на суперпотребителите да прилагат `distcp` на данни, без да изискват достъп до ключовете за криптиране, и избягва натоварването на дешифрирането и повторното кодиране на данни. Това също така означава, че данните за източника и дестинацията ще бъдат идентични по байт за байт, което не би било вярно, ако данните се прекодират с нов EDEK.

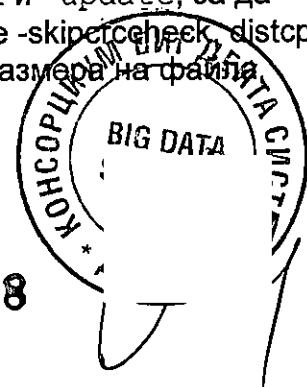
Копиране на данни от криптирани локации

По подразбиране `distcp` сравнява контролните суми, предоставени от файловата система, за да провери дали данните са копирани успешно в местоназначението. Когато копирате от криптирано място, контролните суми на файловата система няма да съвпадат, защото основните блокови данни са различни. Това е вярно, независимо дали местоположението на дестинацията е криптирано или некриптирано.

В този случай можете да укажете флаговете `-skipcrccheck` и `-update`, за да избегнете проверката на контролните суми. Когато използвате `-skipcrccheck`, `distcp` проверява целостта на файла, като извършва сравнение на размера на файла веднага след като копието завърши за всеки файл.

Изтриване на зони за криптиране

000358



За да изтриете зона за криптиране, изтрийте криптираната директория

```
$ sudo -u hdfs hadoop fs -rm -r -skipTrash /encryption_zone
```

Архивиране на ключовете за криптиране

Ако използвате Java KeyStore KMS, уверете се, че редовно архивирате Java KeyStore, който съхранява ключовете за шифроване.

Обновяване на ключовете за криптиране (Rolling)

Когато обновите EZ ключ, вие по същество създавате нова версия на ключа (ezKeyVersionName). Обновяването на EZ ключове редовно помагат на предприятията да намалят риска от пробив в сигурността на ключ. Ако нападателят получи EZ ключа и декриптира криптираните ключове за криптиране на данни (EDEK) в DEK, те биха могли да получат способността да декриптират HDFS файловете. Обновяването на EZ ключ гарантира, че всички DEK за новосъздадени файлове ще бъдат криптирани с новата версия на EZ ключа. По-старата версия на EZ ключа, която е получил, не може да декриптира тези EDEK. Може да искате да обновявате ключа за криптиране периодично, като част от вашата политика за сигурност, или когато бъде открит външна заплаха за сигурността.

За да обновите EZ ключ:

1. Влезте като **Key Administrator**. Тъй като ключовете могат да бъдат обновени, ключът може да има множество версии, където всяка ключова версия има свой собствен ключов материал (действителните тайни байтове, използвани по време на DEK криптиране и EDEK декриптиране). Можете да изтеглите ключ за криптиране от неговото име на ключа, връщане на последната версия на ключа или от конкретна версия на ключа.
Обновете ключа за криптиране. В този пример <key name> стойността е key1:

```
$ hbase o ke roll key1
```

Тази операция се свързва с KMS и обновява ключовете там. Имайте предвид, че това може да отнеме значително време, в зависимост от броя на ключовете, които се намират в KMS.



000359

Rolling key version from KeyProvider:

org.apache.hadoop.crypto.key.kms.LoadBalancingKMSClientProvider@5ea434c8

for keyName: key1

key1 has been successfully rolled.

org.apache.hadoop.crypto.key.kms.LoadBalancingKMSClientProvider@5ea434c8 has been updated.

2. Използвайте **KMS Rest API**, за да видите ключови метаданни и ключови версии. Елементите, които се появяват в скоби, трябва да бъдат заменени с действителните стойности. Така че в този случай, преди да преобърнете ключ, можете да видите основните метаданни и версии, както следва:

```
$ curl -k --negotiate -u: "https://<KMS_HOSTNAME>:16000/kms/v1/key/<key-name>_metadata"
```

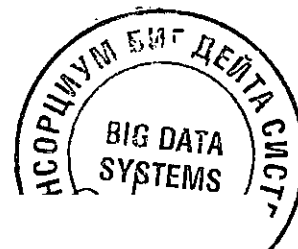
```
{  
  "name" : "<key-name>",  
  "cipher" : "<cipher>",  
  "length" : <length>,  
  "description" : "<description>",  
  "created" : <millis-epoch>,  
  "versions" : <versions> (For example, 1)  
}
```

```
$ curl -k --negotiate -u: "https://<KMS_HOSTNAME>:16000/kms/v1/key/<key-name>/_currentversion"
```

```
{  
  "material" : "<material>",  
  "name" : "<key-name>",  
  "versionName" : "<versionName>" (For example, version 1)  
}
```

3. Обновете ключа и сравнете резултатите:

```
curl -k --negotiate -u: "https://<KMS_HOSTNAME>:16000/kms/v1/key/<key-name>_metadata"
```



Rolling key version from KeyProvider:

KMSClientProvider[https://<KMS_HOSTNAME>:16000/kms/v1/]

For key name: <key-name>

key1 has been successfully rolled.

KMSClientProvider[https://<KMS_HOSTNAME>/kms/v1/] has been updated.

\$ curl -k --negotiate -u: "https://<KMS_HOSTNAME>:16000/kms/v1/key/<key-name>/_currentversion"

```
{
  "material" : "<material>", (New material)
  "name" : "<key-name>",
  "versionName" : "<versionName>" (New version name. For example, version 2)
}
```

\$ curl -k --negotiate -u: "https://<KMS_HOSTNAME>:16000/kms/v1/key/<key-name>/_metadata"

```
{
  "name" : "<key-name>",
  "cipher" : "<cipher>",
  "length" : <length>,
  "description" : "<description>",
  "created" : <millis-epoch>,
  "versions" : <versions> (For example, version 2)
}
```

Пример за прилагане на изискването за създаване на потребители и ключове, както и възможност да се управляват:

Конфигурация на KMS Proxyuser

Всеки proxyuser трябва да бъде конфигуриран в etc/hadoop/kms-site.xml със следните настройки:

<property>



```

<name>hadoop.kms.proxyuser.#USER#.users</name>
<value>*</value>
</property>

<property>
<name>hadoop.kms.proxyuser.#USER#.groups</name>
<value>*</value>
</property>

<property>
<name>hadoop.kms.proxyuser.#USER#.hosts</name>
<value>*</value>
</property>

```

USER # е потребителското име на **proxyuser** за конфигуриране.

Свойството **users** показва потребителите, които могат да бъдат представяни.

Свойството **groups** показва групите, на които потребителите се представят, трябва да принадлежат.

Трябва да се дефинира поне един от свойствата на потребителите или групите. Ако и двете са зададени, тогава конфигурираният **proxyuser** ще може да се представя и потребител в списъка с потребители и всеки потребител, принадлежащ към една от групите в списъка на групите.

Свойството **hosts** показва от кой хост прокси-сървърът може да прави заявки за самоличност.

Ако потребителите, групите или хостовете имат *, това означава, че няма ограничения за **proxyuser** по отношение на потребители, групи или хостове.

crypto command-line интерфейс:

Примери на възможни операции:

```

# Като нормалния потребител, създайте ключ за криптиране
hadoop key create mykey

```

```

# Като суперпотребителя, създайте нова празна директория и я делегирайте като
зона за криптиране
hadoop fs -mkdir /zone
hdfs crypto -createZone -keyName mykey -path /zone

```

```

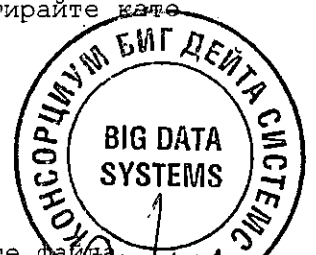
# chown към нормалния потребител
hadoop fs -chown myuser:myuser /zone

```

```

# Поставете файл в директорията, като нормален потребител, изчете файла
hadoop fs -put helloWorld /zone

```



```
hadoop fs -cat /zone/helloWorld
```

```
# Като нормалния потребител, изведете информация за криптиране от файла.  
hdfs crypto -getFileEncryptionInfo -path /zone/helloWorld  
# console output: {cipherSuite: {name: AES/CTR/NoPadding, algorithmBlockSize:  
16}, cryptoProtocolVersion: CryptoProtocolVersion{description='Encryption  
zones', version=1, unknownValue=null}, edek: 2010d301afbd43b58f10737ce4e93b39,  
iv: ade2293db2bab1a2e337f91361304cb3, keyName: mykey, ezKeyVersionName:  
mykey@0}
```

Проверка за операциите на ключовете на Hadoop

Предупреждение: Ако използвате или възнамерявате да използвате HSM на Cloudera Navigator Key съвместно с Cloudera Navigator Key Trustee Server, уверете се, че имената на ключовете започват с буквено-цифрови знаци и не използват специални символи, различни от тире (-), период (.), Или долна черта (_). Използването на други специални символи може да ви попречи да мигрирате ключовете си към HSM. Вижте Интегриране на ключови HSM с Key Trustee Server за повече информация.

Използвайте `hadoop key create`, за да създадете тестов ключ, след което използвайте списъка с ключове на `hadoop`, за да извлечете списъка с ключове:

```
$ sudo -u <key_admin> hadoop key create keytrustee_test
```

```
$ hadoop key list
```

Създаване на зони за шифроване

След като е създадена KMS и клиентите на NameNode и HDFS са правилно конфигурирани, използвайте инструментите на командния ред с ключ на `hadoop` и `hdfs crypto` за създаване на ключове за криптиране и създаване на нови зони за криптиране.

Създайте ключ за шифроване за вашата зона като ключ за управление на потребителя / групата (независимо от приложението, което ще използва зоната за шифроване):

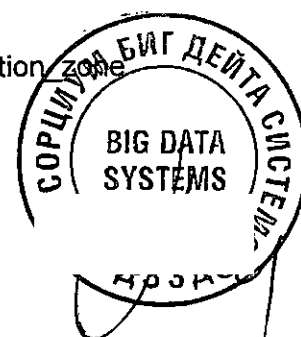
```
$ sudo -u hbase hadoop key create <key_name>
```

Създайте нова празна директория и я направете зона за шифроване с помощта на създадения по-горе ключ.

```
$ sudo -u hdfs hadoop fs -mkdir /encryption_zone
```

```
$ sudo -u hdfs hdfs crypto -createZone -keyName <key_name> -path /encryption_zone
```

000363



Можете да проверите създаването на новата зона за криптиране, като изпълните командата `-listZones`. Трябва да видите зоната за шифроване заедно с нейния ключ, както следва:

```
$ sudo -u hdfs hdfs crypto -listZones
```

```
/encryption_zone <key_name>
```

Предупреждение: Не изтривайте ключ за шифроване, докато той все още се използва за зона за криптиране. Това води до загуба на достъп до данни в тази зона.

Важно: Java Keystore KMS по подразбиране Truststore (например, `org.apache.hadoop.crypto.key.JavaKeyStoreProvider`) не поддържа главни имена на ключове.

За повече информация и препоръки за създаване на зони за криптиране за всеки CDH компонент вж. Конфигуриране на CDH услуги за HDFS криптиране.

Добавяне на файлове в зона за шифроване

Съществуващите данни могат да бъдат криптирани чрез копирането им в новите зони за криптиране с помощта на инструменти като `DistCp`. Можете да добавяте файлове към зона за криптиране, като ги копирате в зоната за шифроване, като използвате `distcp`.
Например:

```
$ sudo -u hdfs hadoop distcp /user/dir /encryption_zone
```

Изтриване на зони за шифроване

За да премахнете зона за шифроване, изтрийте шифрованата директория:

Предупреждение: Тази команда изтрива цялата директория и цялото ѝ съдържание. Уверете се, че данните не са необходими преди да изпълните тази команда.

```
$ sudo -u hdfs hadoop fs -rm -r -skipTrash /encryption_zone
```

Важно: Ключовото доверително управление KMS не изпълнява директно клавишното изтриване (например може да извърши меко изтриване, или да забави действителното изтриване, за да предотврати грешки). В тези случаи може да възникнат грешки при създаването или изтриването на ключ, използващ същото име, след като вече е бил изтрит.

Архивиране на ключовете за шифроване



000364

Предупреждение: Много е важно редовно да архивирате ключовете за шифроване. Ако не го направите, това може да доведе до безвъзвратна загуба на криптирани данни.

Ако използвате Java KeyStore KMS, уверете се, че редовно архивирате Java KeyStore, който съхранява ключовете за шифроване.

✓
✓

000365



Документ

„Примерно функциониране на Адаптивна Архитектура за Информационна Сигурност“

Предложение за реализация съгласно изискванията на проект:

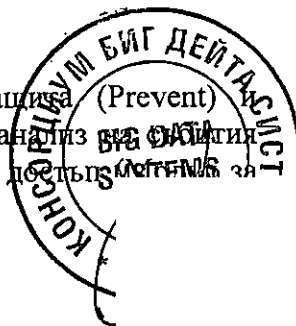
„Доставка, инсталиране, конфигуриране, тестване и поддържане на **СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ NADOOР** за работа с Големи данни по проект „Дигитализация на икономиката в среда на Големи данни“ (ДИГД), № BG05M2OP001-1.002-0002“, съгласно стр. 53

1. Общо описание на реализацията

В контекста на кибер сигурността традиционни методологии за откриване и предотвратяване на заплахи чрез внедряване на антивирусен софтуер и защитни стени стават все по-неефективни като организациите променят тяхната бизнес концепция интегрирайки облачно базирани системи и API системи с отворен код в техните дейности. Разширявайки границите на информационната си инфраструктура компаниите все по-трудно се справят с контрола върху тази инфраструктура и опазването на техните данни. Следователно реактивния подход на реакция към инцидент - „incident response“ е нужно да бъде заменен с проактивен динамичен такъв „continuous response“. Той ще се основава на това, че една организация може да бъде компроментирана във всеки един момент и продължителното мониториране на системите и анализ на поведението ще бъде единствения надежден начин за откриване и предотвратяване на заплахите.

Специализирана централизирана система под ключ NADOOР ще бъде реализирана на база на адаптивна архитектура за инфомационна сигурност, която обхваща три компонента: Откриване (Predict), Защита (Prevent) и Отговор (Respond). Тези компоненти ще отговарят за идентифицирането, анализирането, блокирането и адаптирането към заплахите като ще се използват динамични механизми за защита и реакции. Тези механизми ще проучват поведението на заплахите и ще придобиват умения на базата, на които ще се адаптират и ще отговарят адекватно на бъдещи заплахи.

Трите компоненти от архитектурата Откриване (Predict), Защита (Prevent) и Отговор (Respond) включват респективно Модул за управление и анализ на обектите (Модул за УАЛС), IPS Модул, и Модул за Мрежово управление и достъп до мрежата за



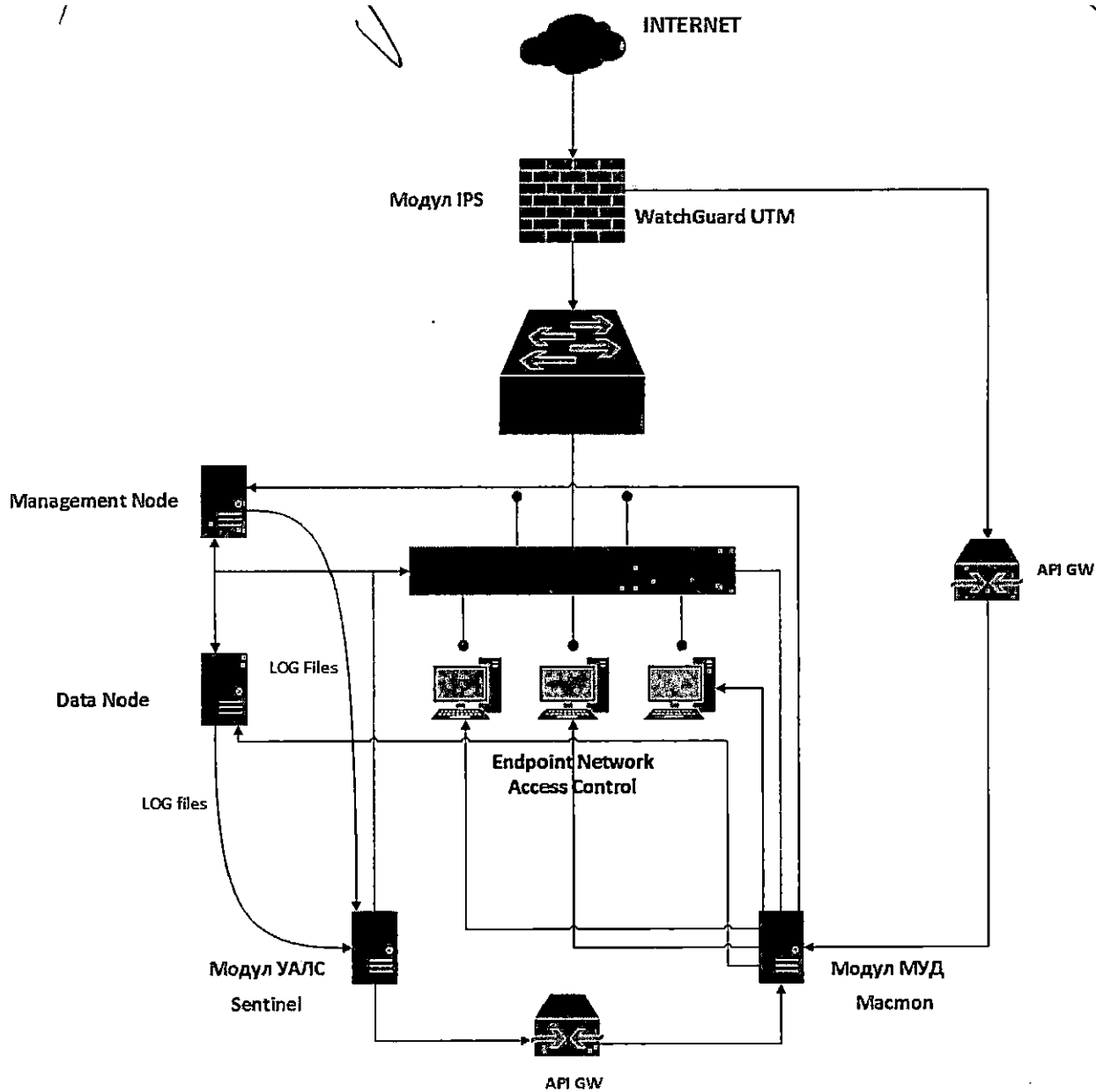
000366

МУД). Продуктите, които ще бъдат използвани, които покриват изискванията на тези модули са респективно MicroFocus Sentinel SIEM, WatchGuard UTM и Macmon. Тези модули ще си взаимодействат със Специализирана централизирана система под ключ Hadoop по следния начин: В случай че IPS модула открие, че някоя от крайните точки се опитва да прави нелегитимен трафик или че е заразена с вирус, то преустановява връзката на тази точка към Интернет и изпраща заявка посредством API Gateway към Модул за МУД с цел преместване на тази точка в отделна излорана мрежа за карантина, където няма да може да прави заявки към останалите хостове в локалната мрежа. Модулите DataNode, NameNode, ManagementNode и EdgeNode, част от системата Hadoop, както и техните Операционни системи ще изпращат техните LOG файлове на Модул за УАЛС, който ще ги анализира и в случай на нужда от някакво действие ще изпраща заявка съответно на Модул за МУД посредством API Gateway за изолиране на конкретния node.

2. Архитектура на Адаптивна Архитектура за Информационна Сигурност



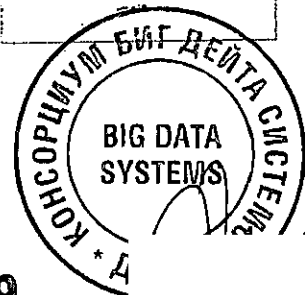
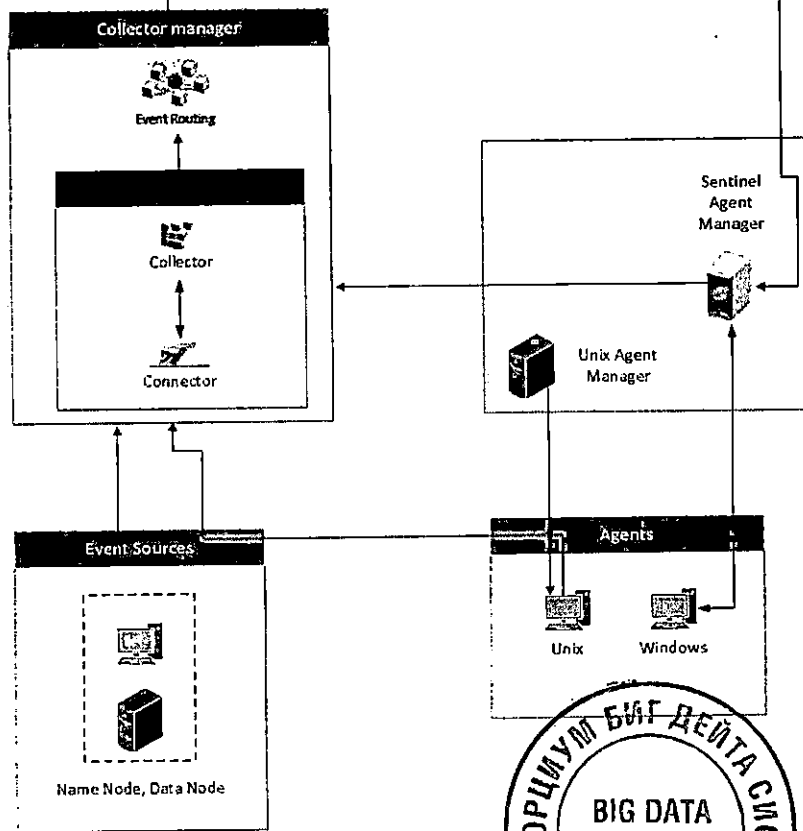
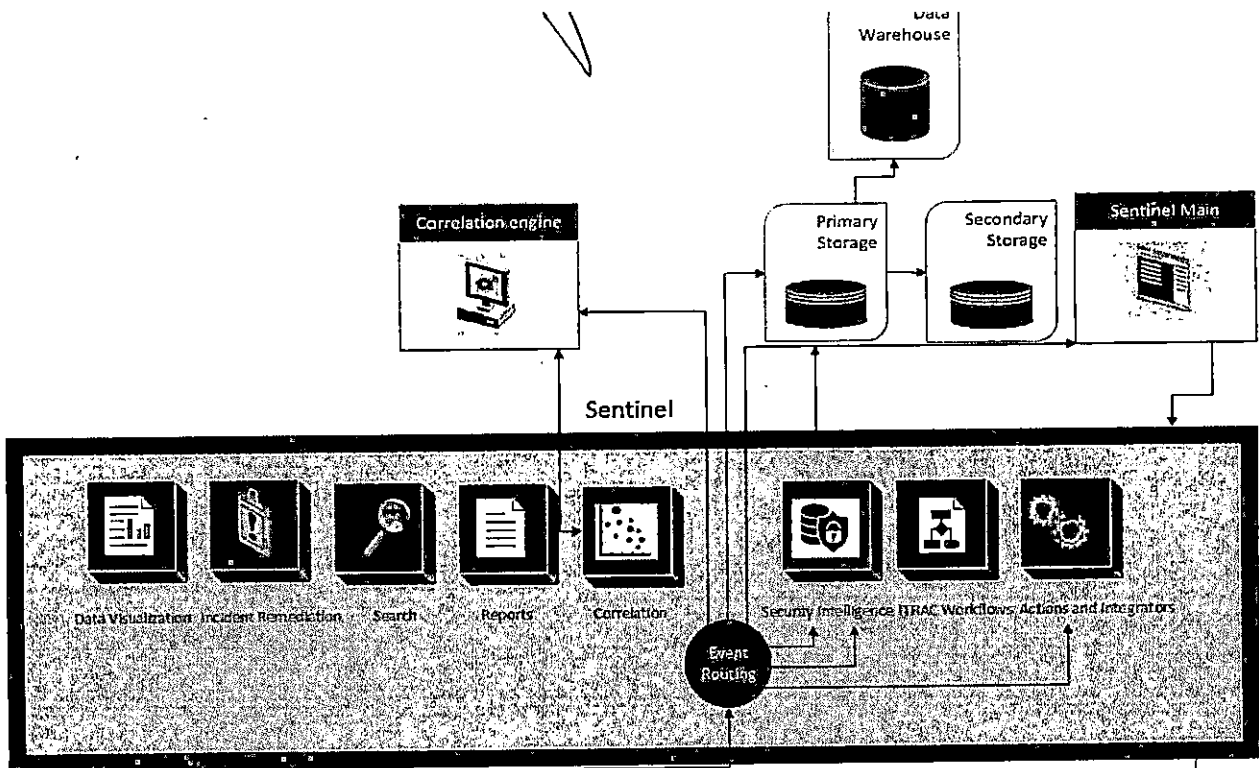
000367



3. Интеграцията между модул УАЛС и източници на данни

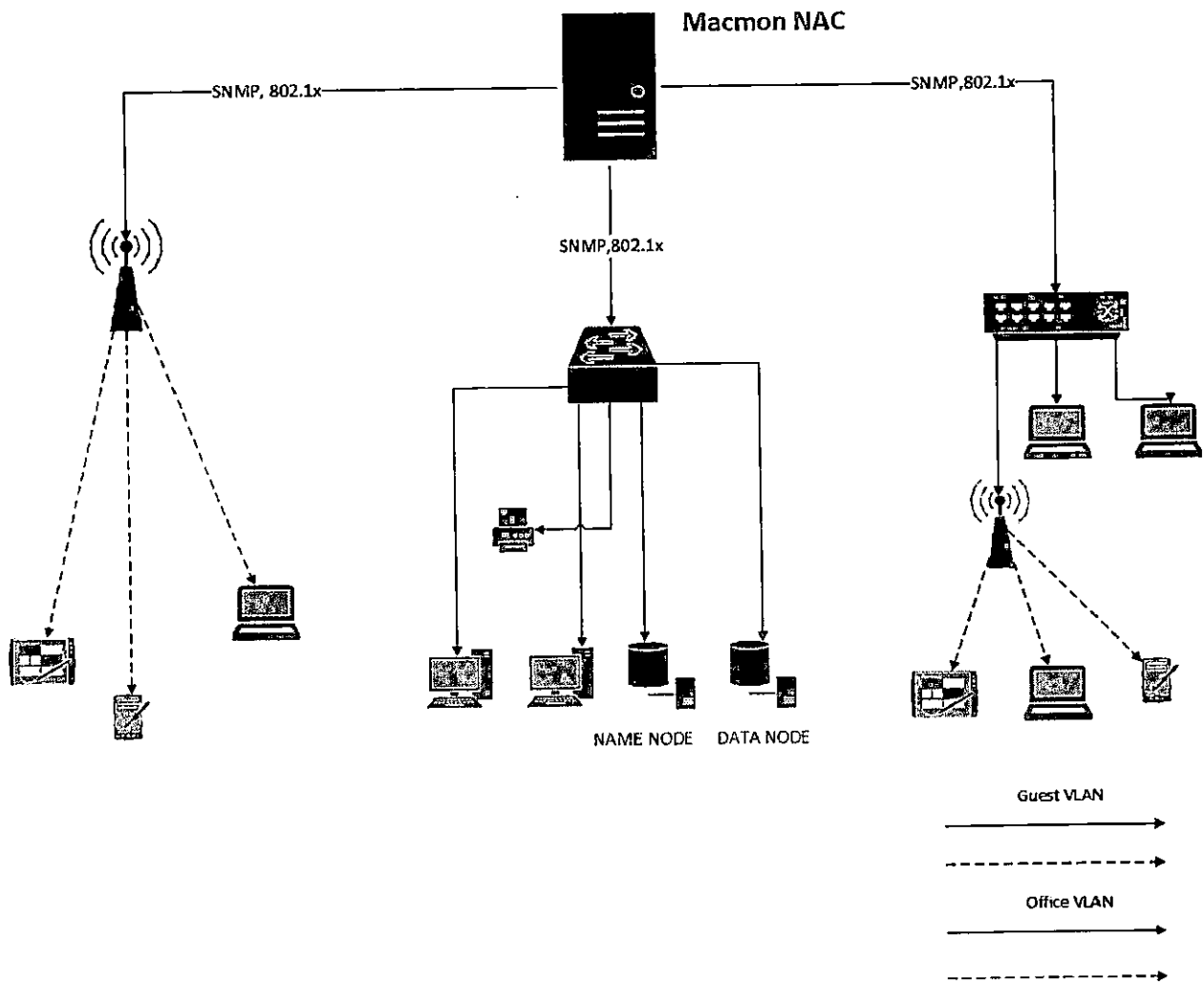


000368



000369

- Диаграма на модул МУД



Модул МУД реализиран чрез продукт Масмон позволява пълна видимост на всички крайни точки в локалната мрежа на базата на техните физически адреси и комуникация с голям набор комутатори на база на SNMP протокола. За работата му не са необходими структурни промени по мрежата както и инсталация на агенти или сензори на крайните точки. Продуктът позволява интеграция с Active Directory и аутентикация на база на 802.1x протокола. Могат да бъдат създадени правила въз основа на които при възникване на събитие крайни точки може да бъдат слагани в предварително конфигурирани виртуални локални мрежи. Така отделни точки могат да бъдат изолирани при необходимост.



000370

4. Разпечатки от примерно функциониране и конфигуриране в компютърна среда на доставчика, на следните компоненти: Модула за УАЛС, IPS устройството и Модула за МУД. В този документ са представени разпечатки от следните 4 броя примерно интегриране:

- a. Между Модул за УАЛС и LOG файловете на един модул МодулDataNode, и на един модул МодулNameNode, както и на техните операционни системи;
- b. Резултат от корелация между съдържание на LOG файл на един модул МодулDataNode и на LOG файл на един модул МодулNameNode;
- c. Между IPS устройството и Модул за прилагане на политики– Модула за МУД;
- d. Между Модул за УАЛС и Модул за МУД.



000371

Приложение 4.а.

Разпечатка от интеграцията между Модул за УАЛС и LOG файловете на един модул МодулDataNode, и на един модул МодулNameNode, както и на техните операционни системи.

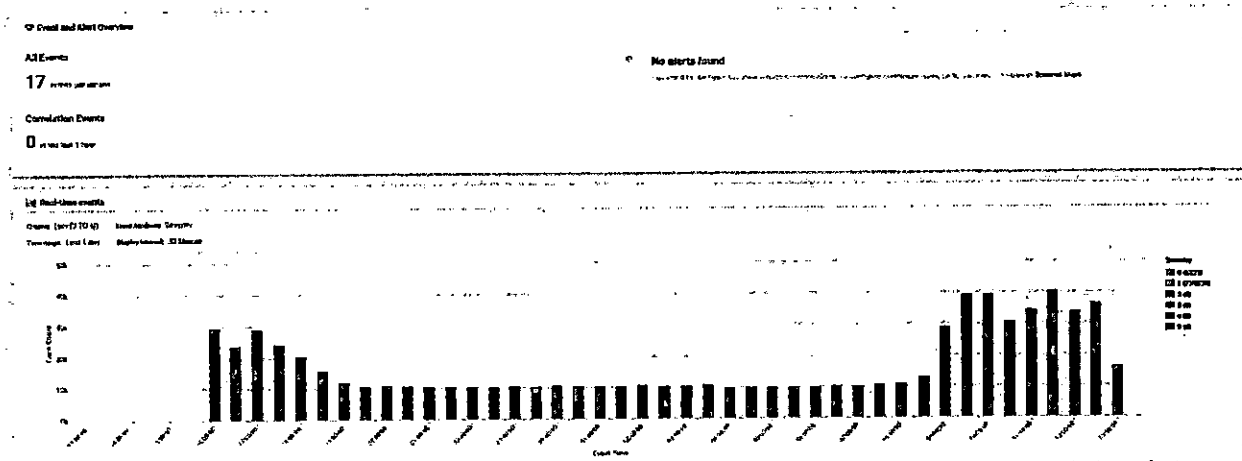
Task ID	Sort/Reduce Dur.	Host	Reduce input record	Red_InputWalks	Red_ItemsStored	Reduce output record	Red_InputEdge
0	13 mins, 38 sec	/default-rack/hb48	5140099	1377940	1377940	1126185586	3694577
16	6 mins, 26 sec	/default-rack/hb38	2455838	969936	969936	537624991	1485899
1	5 mins, 45 sec	/default-rack/hb35	2145360	875696	875696	45977429	1269657
4	5 mins, 7 sec	/default-rack/hb46	1914371	798103	798103	393474064	1116265
32	4 mins, 29 sec	/default-rack/hb30	1795805	755856	755856	361099409	1039947
2	3 mins, 39 sec	/default-rack/hb02	1599822	682502	682502	307798322	917313
8	2 mins, 18 sec	/default-rack/hb34	1182575	504203	504203	192394598	678357
17	2 mins, 14 sec	/default-rack/hb14	1018130	559563	559563	181444427	458563
20	1 mins, 49 sec	/default-rack/hb29	904356	506663	506663	153978269	397687
48	1 mins, 34 sec	/default-rack/hb41	849553	478388	478388	140742756	371262
5	1 mins, 29 sec	/default-rack/hb08	786533	450663	450663	125740548	335867
18	1 mins, 22 sec	/default-rack/hb03	755337	428253	428253	117784038	327079
33	1 mins, 20 sec	/default-rack/hb56	736614	425793	425793	114240201	310816
36	1 mins, 10 sec	/default-rack/hb09	653884	383764	383764	95971373	270116
26	1 mins, 7 sec	/default-rack/hb01	164681	119050	119050	9831727	45622
3	1 mins, 7 sec	/default-rack/hb24	655827	381362	381362	95269253	274459
6	59 sec	/default-rack/hb17	579106	341530	341530	79562804	237569
24	49 sec	/default-rack/hb50	554942	313570	313570	71584423	241358
34	49 sec	/default-rack/hb57	541764	322623	322623	71825128	219134
9	40 sec	/default-rack/hb40	481077	278603	278603	57210801	202463
40	35 sec	/default-rack/hb25	397581	234780	234780	41977549	162791
12	34 sec	/default-rack/hb28	424931	249541	249541	47167358	175382

New task table view of one job, showing reduce tasks in order of task duration. Columns can be rearranged, and rows sorted by any column header. Columns on the right show the values of job counters, which often correlate with job duration.



000372

Формат на дашборда в Sentinel – УАЛС модул



Приложение 4.б

Разпечатка на резултат от корелация между съдържание на LOG файл на един модул МодулDataNode и на LOG файл на един модул МодулNameNode

000373



Time taken by best performing Reduce task :
task 201111071101 1457 r 000063 : 1sec

Average time taken by Reduce tasks: 1mins, 1sec

Worse performing reduce tasks

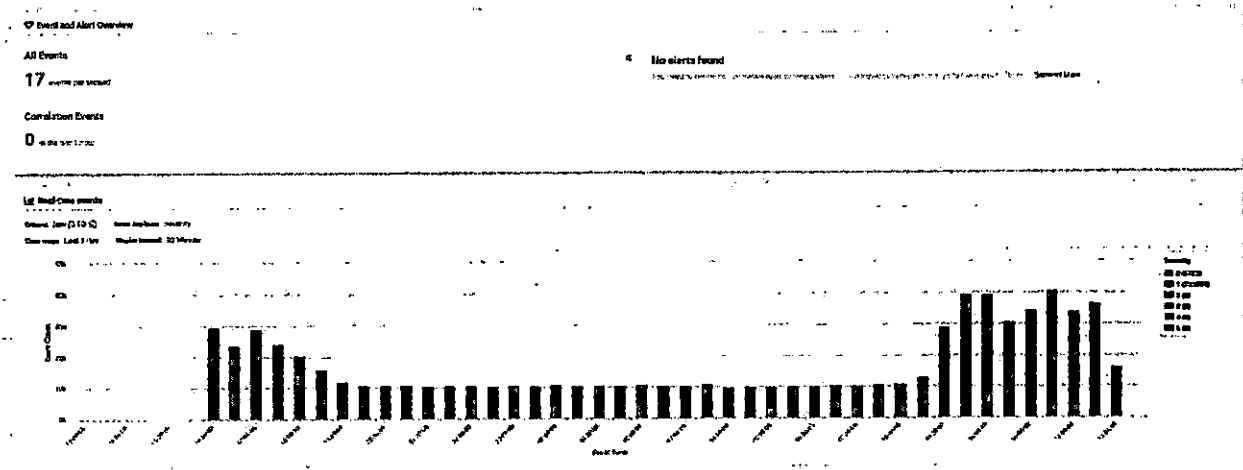
Task Id	Time taken
task 201111071101 1457 r 000000	13mins, 38sec
task 201111071101 1457 r 000016	6mins, 26sec
task 201111071101 1457 r 000001	5mins, 45sec
task 201111071101 1457 r 000004	5mins, 7sec
task 201111071101 1457 r 000032	4mins, 29sec
task 201111071101 1457 r 000002	3mins, 39sec
task 201111071101 1457 r 000008	2mins, 18sec
task 201111071101 1457 r 000017	2mins, 14sec



000374

U 7

Формат на дашборда в Сентинел – УАТС модул



Приложение 4.с

Разпечатка от интеграцията между IPS устройствата и Сентинел модул за прилагане на политики – Модула за МУД

Формат на събитията в IPS, един IP адрес и MAC адрес.

Source	Destination	Action	Severity	Status
192.168.1.78	192.168.1.1	IPsec Policy Violation	High	Blocked
192.168.1.78	192.168.1.1	IPsec Policy Violation	High	Blocked
192.168.1.78	192.168.1.1	IPsec Policy Violation	High	Blocked

На картинката е показано действие на IPS за вътрешно IP. В случая става дума за IP:192.168.1.78

в 11:58:08. Същото събитие се появява в Сентинел. От ARP таблица в Сентинел се идентифицира Mac



000375

адреса на устройството 18:A6:F7:67:FB:D7. Прави се забрана в MACMON за прехвърляне на устройството от стандартна в изолирана мрежа чрез SNMP протокол за съответният суич на което е закачено.

IP	MAC
192.168.1.41	90:1b:0e:45:c0:55 eth1
192.168.1.122	00:26:6c:46:33:78 eth1
192.168.1.139	d8:50:e6:4b:e9:6c eth1
169.254.13.2	44:85:00:21:a9:c6 eth1
192.168.1.109	f0:de:f1:af:28:d6 eth1
192.168.1.78	18:A6:F7:67:FB:D7 eth1
192.168.1.84	8c:73:6e:ba:e5:d4 eth1
192.168.1.54	c0:3f:d5:67:55:8c eth1
192.168.1.71	68:e7:c2:49:4d:89 eth1

API WatchGuard - IPS

За инсталация, можете да погледнете тук: [Install WatchGuard System Manager Software.](#)

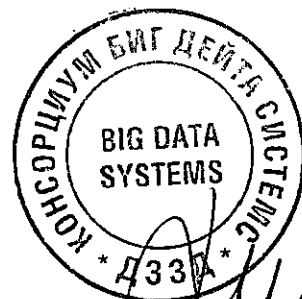
За да използвате Web Services API, вие трябва да имате API файлове и документация.

Web Services API файловете са инсталирани на тази локация:

C:\Program Files\WatchGuard\wsm11\wsserver\wsdl

Web Services API файловете включват:

- LogsService.wsdl — WSDL (Web Services Description Language) file
- LogsService.xsd — Web Services API schema



000376

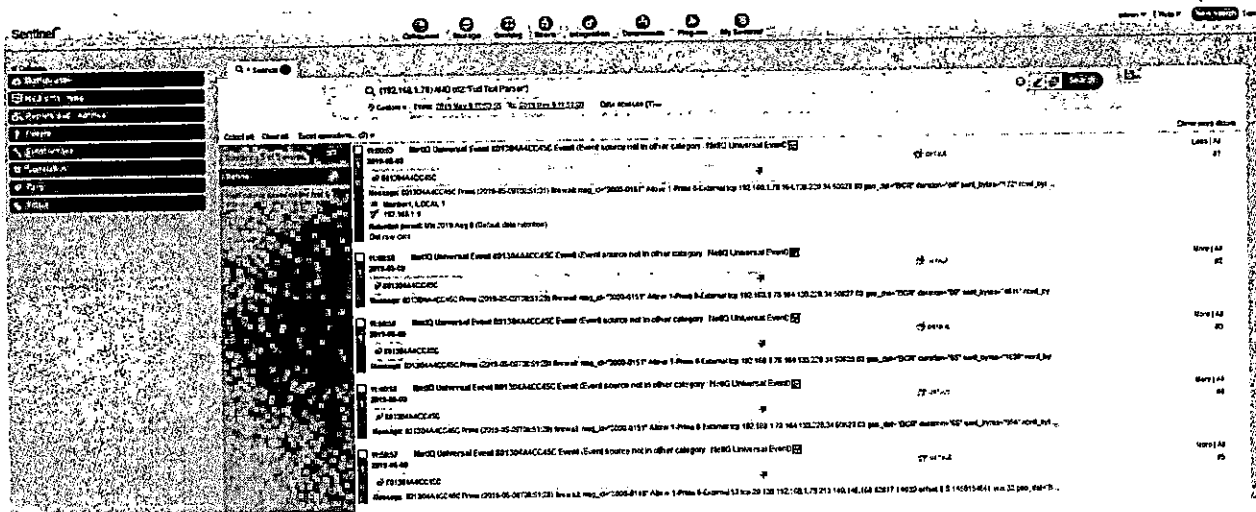
Web Services документацията се намира на тази локация:

C:\Program Files\WatchGuard\wsm11\wsserver\wsdl\docs

Web Services файловете включват:

- LogService.html — WSDL reference documentation
- LogService-UserGuide.pdf — A user guide for Web Services API for Reporting
- LogServiceEclipse.pdf — An interoperability example for Eclipse and Java

Справка от Sentinel за същото събитие за същия MAC адрес. Съответствието между IPS и МУД може да се види в изобразената диаграмата по – горе.



000377

Приложение 4.d

Разпечатка от интеграцията между модул за УАЛС и Модул за МУД

SENTINEL API - УАЛС:

GlobalDataSyncMetaData - GlobalDataSyncMetaData Retrieve, Update, and Delete Methods

URI

<https://164.99.19.131:8443/SentinelRESTServices/objects/datasyncmd/{ConfigId}>

Supported Methods

GET

PUT

DELETE

GET

Authentication

Authentication Types

- X-SAML

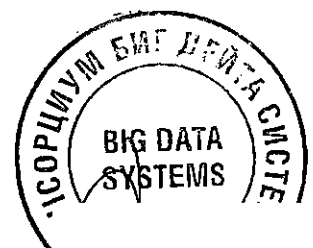
Sentinel Permissions Needed

- _adminRole_

Supported Formats

Response

- application/json



000378

URL Parameters

None.

Success Codes

- 200 OK
- 204 No Content

Fault Codes

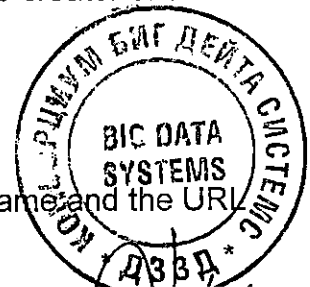
- 400 Bad Request
- 403 Forbidden
- 404 Not Found
- 500 Internal Server Error
- 503 Service Unavailable

Response Data

Object type: `datasync_md`

Needs description.

Field	Description
Application	Needs description.
createdate	The date and time when the object was created.
creator	The URL of the Sentinel User object that represents the creator of the object.
Data	Needs description. Definition
meta	The metadata for an object, including the object type name and the URI reference to the object. Definition



000379

- moddate The date and time when the object was last modified.
- modifier The URL of the Sentinel User object that represents of the
object.
- Status Needs description.
- Type Needs description.
- Unit Needs description.
- UserId Needs description.

Object type: meta

The metadata for an object, including the object type name and the URL reference to the object.

Field	Description
@href	The URL reference to the object.
type	The name of the object type

Object type: GlobalDataSyncConfig

Needs description.

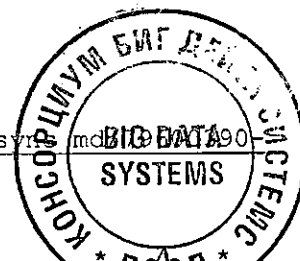
Field	Description
RetentionPeriod	Needs description.

Sample Request

GET
<https://164.99.19.131:8443/SentinelRESTServices/objects/datasync/mod/90-9B73-102E-A3E2-001676E4A757>

Sample Response for application/json

Status: 200



000380

```
{
  "meta":{
    "type":"datasync_md",
    "@href":"https://164.99.19.131:8443/SentinelRESTServices/objects/datasyn
c_md/79600390-9B73-102E-A3E2-001676E4A757"
  },
  "Data":{
    "RetentionPeriod":42
  },
  "Status":"Wildebeest",
  "moddate":"2012-04-25T13:33:44.441Z",
  "Type":"Wildebeest",
  "Application":"Wildebeest",
  "createdate":"2012-04-25T13:33:44.441Z",
  "UserId":42,
  "Unit":"Wildebeest"
}
```

PUT

Authentication

Authentication Types

- X-SAML

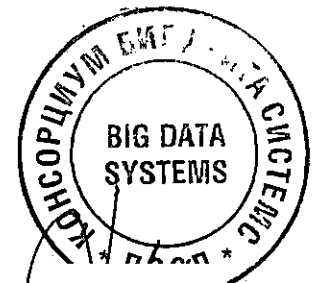
Sentinel Permissions Needed

- _adminRole_

Supported Formats

Request

- application/json



00381

URL Parameters

None.

Success Codes

- 204 No Content

Fault Codes

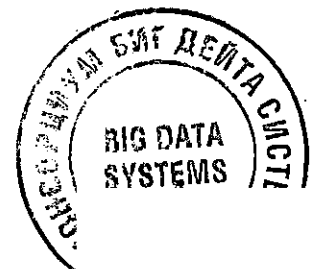
- 400 Bad Request
- 403 Forbidden
- 404 Not Found
- 500 Internal Server Error
- 503 Service Unavailable

Request Data

Object type: `datasync_md`

Needs description.

Field	Required	Description
Application	true	Needs description.
Data	false	Needs description. <u>Definition</u>
Status	false	Needs description.
Type	false	Needs description.
Unit	false	Needs description.
Userld	true	Needs description.



000382

Object type: meta

The metadata for an object, including the object type name and the URL reference to the object.

Field	Required	Description
@href	false	The URL reference to the object.
type	false	The name of the object type

Object type: GlobalDataSyncConfig

Needs description.

Field	Required	Description
RetentionPeriod	false	Needs description.

Sample Request

PUT
https://164.99.19.131:8443/SentinelRESTServices/objects/datasync_md/79600390-9B73-102E-A3E2-001676E4A757

```
{
  "Data":{
    "RetentionPeriod":42
  },
  "Status":"Wildebeest",
  "Type":"Wildebeest",
  "Application":"Wildebeest",
  "UserId":42,
  "Unit":"Wildebeest"
}
```



000383

Sample Response for application/json

Status: 200

DELETE

Authentication

Authentication Types

- X-SAML

Sentinel Permissions Needed

- _adminRole_

URL Parameters

None.

Success Codes

- 204 No Content

Fault Codes

- 400 Bad Request
- 403 Forbidden
- 404 Not Found
- 500 Internal Server Error
- 503 Service Unavailable

Sample Request

DELETE

<https://164.99.19.131:8443/SentinelRESTServices/objects/datasrcs/796003909B73-102E-A3E2-001676E4A757>



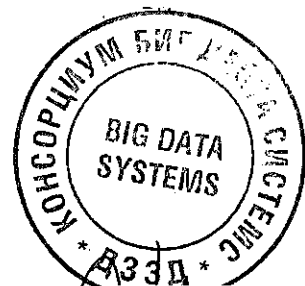
000384

API MACMON - МУД

30 .vscode/settings.json

@@ -0,0 +1,30 @@

```
{  
  "cSpell.enabledLanguageIds": [  
    "asciidoc",  
    "c",  
    "cpp",  
    "csharp",  
    "css",  
    "go",  
    "handlebars",  
    "html",  
    "jade",  
    "javascript",  
    "javascriptreact",  
    "json",  
    "latex",  
    "less",  
    "markdown",  
    "php",  
    "plaintext",  
    "pub",  
    "python",  
    "restructuredtext",
```



000385

"rust",
"scss",
"text",
"typescript",
"typescriptreact",
"yaml"

]
}

64 PSmacmon/Private/Get-MacmonFunctionString.ps1

@@ -0,0 +1,64 @@

function Get-MacmonFunctionString

{

<#

.EXAMPLE

\$Params = [ordered]@{

Filter = 'active==true and networkDeviceGroup.name == "Switch"'

Fields = 'id,userValues.*.value'

Limit = '10'

Offset = '2'

Sort = 'active,address'

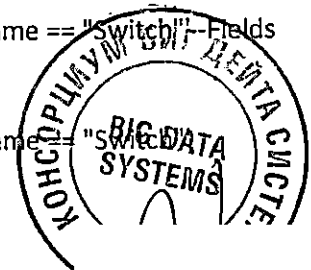
}

Get-MacmonFunctionString @Params

Get-MacmonFunctionString -Filter 'active==true and networkDeviceGroup.name == "Switch"' -Fields
'id,userValues.*.value' -Sort 'active,address'

Get-MacmonFunctionString -Filter 'active==true and networkDeviceGroup.name == "Switch"'

Get-MacmonFunctionString -Fields 'id,userValues.*.value' -Offset 1



00386

Get-MacmonFunctionString -Offset 0

Get-MacmonFunctionString -Fields " -Offset 0

#>

[CmdletBinding()]

param (

[string]

\$Fields,

[string]

\$Sort,

[int]

\$Limit,

[int]

\$Offset,

[string]

\$Filter

)

begin

{

}

process

{

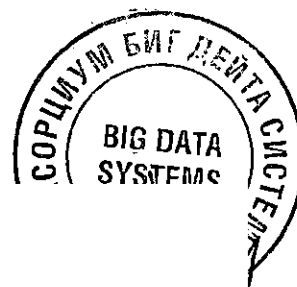
foreach (\$item in \$PSBoundParameters.GetEnumerator())

{

if (\$item.Value)

{

if (\$item.Value -is [int])

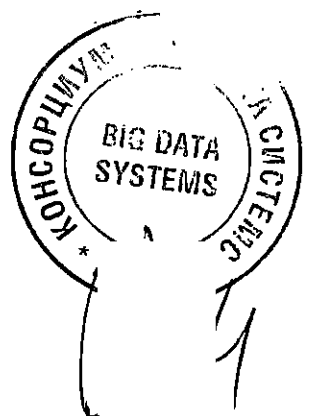


000387

```
{
    $Value = ($item.Value).toString()
}
else
{
    $Value = $item.Value
}
$FunctionString += ('&{0}={1}' -f $($item.Key).ToLower(), $Value)
}
}
$FunctionString -replace ('^&', '?')
}
end
{
}
}
```

2 PSmacmon/Private/Invoke-MacmonRestMethod.ps1

```
@@ -111,7 +111,7 @@ function Invoke-MacmonRestMethod
{
    Write-Warning -Message ('Internal server error. SessionURL: {0} Method: {1}' -f $SessionURL,
$Method)
}
901
911
{
```



000388

Handwritten signature or initials at the top of the page.

Handwritten mark or signature on the right side of the page.

Write-Warning -Message ('Application error. SessionURL: {0} Method: {1}' -f \$SessionURL, \$Method)

}

62 PSmacmon/Public/Get-MacmonEndpoint.ps1

@@ -27,12 +27,24 @@ function Get-MacmonEndpoint

Get-MacmonEndpoint -Hostname 'MACMONSERVER' -Credential \$Credential

#Ask for credential then get Endpoint from macmon NAC using provided credential

.EXAMPLE

\$Params = @{

Hostname = 'MACMONSERVER'

Fields = 'mac,endpointDeviceStatus.lastIp'

Sort = '-mac'

Limit = 1

Offset = 10

Filter = 'endpointGroupId==150'

}

Get-MacmonEndpoint @Params

Get mac and lastIP address from 11th endpoint from endpointgroup with ID 150 sorted by mac descending

.EXAMPLE

'00-00-FF-FF-FF-FF' | Get-MacmonEndpoint -Hostname 'MACMONSERVER'

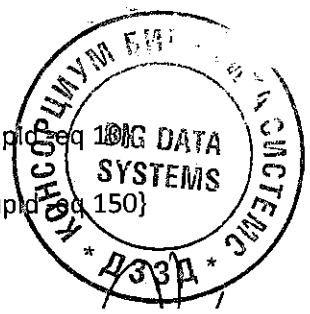
#Get Endpoint with MACAddress '00-00-FF-FF-FF-FF'

.EXAMPLE

(Get-MacmonEndpoint -Hostname 'MACMONSERVER').where{\$_.endpointGroupId -eq 150}

(Get-MacmonEndpoint -Hostname 'MACMONSERVER').where{\$_.endpointGroupId -eq 150}

#Get Endpoint with endpointGroupId 10



.LINK

@@ -43,7 +55,8 @@ function Get-MacmonEndpoint

#>

[CmdletBinding()]

#[CmdletBinding()]

[CmdletBinding(DefaultParameterSetName = 'All')]

param (

[Parameter(Mandatory)]

[string]

@@ -62,29 +75,56 @@ function Get-MacmonEndpoint

[System.Management.Automation.Credential()]

\$Credential = (Get-Credential -Message 'Enter your credentials'),

[Parameter(ValueFromPipeline)]

[Parameter(ValueFromPipeline, ParameterSetName = 'MAC')]

[ValidatePattern('([0-9A-Fa-f]{2}[:-]){5}[0-9A-Fa-f]{2}|([0-9A-Fa-f]{4}\.){2}[0-9A-Fa-f]{4}')]]

[string]

\$MACAddress

\$MACAddress,

[string]

\$Fields,

[Parameter(ParameterSetName = 'All')]

[string]

\$Sort,

[Parameter(ParameterSetName = 'All')]

[int]

\$Limit,

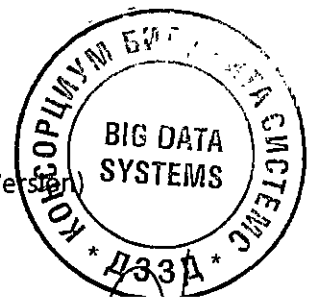


000390

```

[Parameter(ParameterSetName = 'All')]
[int]
$Offset,
[Parameter(ParameterSetName = 'All')]
[string]
$Filter
)
begin
{
    Invoke-MacmonTrustSelfSignedCertificate
    $BaseUrl = ('https://{0}:{1}/api/v{2}/endpoints' -f $HostName, $TCPPort, $ApiVersion)
    $FunctionStringParams = [ordered]@{
        Fields = $Fields
        Sort = $Sort
        Limit = $Limit
        Offset = $Offset
        Filter = $Filter
    }
    $FunctionString = Get-MacmonFunctionString @FunctionStringParams
}
process
{
    Invoke-MacmonTrustSelfSignedCertificate
    $BaseUrl = ('https://{0}:{1}/api/v{2}/endpoints' -f $HostName, $TCPPort, $ApiVersion)
    Switch ($MACAddress)
    Switch ($PsCmdlet.ParameterSetName)

```



00039

Handwritten mark resembling a stylized 'K' or '5'.

Handwritten mark resembling a 'd' or '0'.

```

{
  "
  'All'
  {
    $SessionURL = ('{0}' -f $BaseURL)
    $SessionURL = ('{0}{1}' -f $BaseURL, $FunctionString)
    (Invoke-MacmonRestMethod -Credential $Credential -SessionURL $SessionURL -Method
'Get').SyncRoot
  }
  default
  'MAC'
  {
    $SessionURL = ('{0}/{1}' -f $BaseURL, $MACAddress)
    $SessionURL = ('{0}/{1}{2}' -f $BaseURL, $MACAddress, $FunctionString)
    Invoke-MacmonRestMethod -Credential $Credential -SessionURL $SessionURL -Method 'Get'
  }
}

```

47 PSmacmon/Public/Get-MacmonNetworkDevice.ps1

@@ -43,7 +43,7 @@ function Get-MacmonNetworkDevice

#>

[CmdletBinding()]

[CmdletBinding(DefaultParameterSetName = 'All')]

param (

[Parameter(Mandatory)]

[string]

@@ -62,28 +62,55 @@ function Get-MacmonNetworkDevice



000392

Handwritten checkmark.

Handwritten vertical line.

[System.Management.Automation.Credential()]

\$Credential = (Get-Credential -Message 'Enter your credentials'),

[Parameter(ValueFromPipeline)]

[Parameter(ValueFromPipeline, ParameterSetName = 'ID')]

[int]

\$ID = -1

\$ID,

[string]

\$Fields,

[Parameter(ParameterSetName = 'All')]

[string]

\$Sort,

[Parameter(ParameterSetName = 'All')]

[int]

\$Limit,

[Parameter(ParameterSetName = 'All')]

[int]

\$Offset,

[Parameter(ParameterSetName = 'All')]

[string]

\$Filter

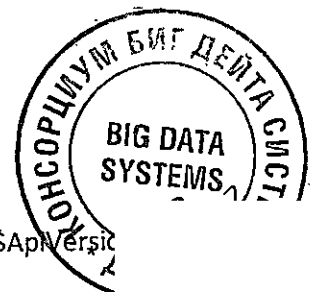
)

begin

{

Invoke-MacmonTrustSelfSignedCertificate

\$BaseUrl = ('https://{0}:{1}/api/v{2}/networkdevices' -f \$HostName, \$TCPPort, \$ApiVersion)



000393

\$FunctionStringParams = [ordered]@{

Fields = \$Fields

Sort = \$Sort

Limit = \$Limit

Offset = \$Offset

Filter = \$Filter

}

\$FunctionString = Get-MacmonFunctionString @FunctionStringParams

}

process

{

Invoke-MacmonTrustSelfSignedCertificate

\$BaseUrl = ('https://{0}:{1}/api/v{2}/networkdevices' -f \$HostName, \$TCPPort, \$ApiVersion)

Switch (\$ID)

Switch (\$PsCmdlet.ParameterSetName)

{

-1

'All'

{

\$SessionURL = ('{0}' -f \$BaseUrl)

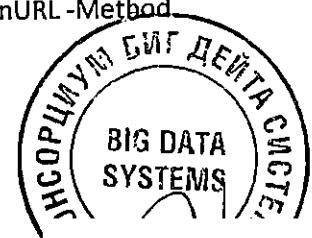
\$SessionURL = ('{0}{1}' -f \$BaseUrl, \$FunctionString)

(Invoke-MacmonRestMethod -Credential \$Credential -SessionURL \$SessionURL -Method
'Get').SyncRoot

}

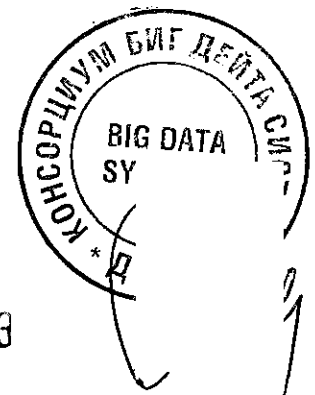
default

'ID'



000394

```
{
  $SessionURL = ('{0}/{1}' -f $BaseUrl, $ID)
  $SessionURL = ('{0}/{1}{2}' -f $BaseUrl, $ID, $FunctionString)
  Invoke-MacmonRestMethod -Credential $Credential -SessionURL $SessionURL -Method 'Get'
}
}
```

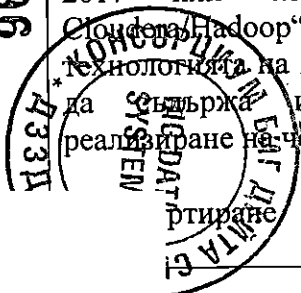


0003

Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависимости
-----------	-----------------------	---------------------	-------------

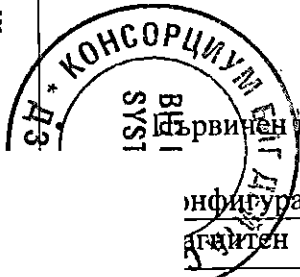
<p>Участниците следва да представят следните материали, свързани с „Интегриране на Специализираната централизирана система под ключ Hadoop работеща под управление на Cloudera, с Майкрософт SQL Server минимум версия 2017“:</p> <p>Материали свързани с интегриране на двете системи MS SQL Server 2017 или следваща версия и система Cloudera/Hadoop“, прилагайки PolyBase технологията на MS SQL Server:</p> <p>1.1. Документ „Ръководство за програмиста за интегриране на двете системи MS SQL Server 2017 или следваща версия и система Cloudera/Hadoop“, прилагайки PolyBase технологията на MS SQL Server, който документ да съдържа информация за потенциално реализиране на четирите типа интегриране: експортиране и импортиране на данни от</p>	<p>ДА</p>	<p>Предоставяме следните материали, относно „Интегриране на Специализираната централизирана система под ключ Hadoop работеща под управление на Cloudera, с Майкрософт SQL Server минимум версия 2017“, както и материали свързани с интегриране на двете системи MS SQL Server 2017 или следваща версия и система Cloudera/Hadoop“, прилагайки PolyBase технологията на MS SQL Server, както следва:</p> <p>Прилагаме: Документ „Ръководство за програмиста за интегриране на двете системи MS SQL Server 2017 или следваща версия и система Cloudera/Hadoop“, прилагайки PolyBase технологията на MS SQL Server, като документа съдържа информация за потенциално реализиране на четирите типа интегриране:</p> <p>а) експортиране и импортиране на данни от</p>	<p>AM</p> <p>U</p> <p>6</p> <p>7</p>
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------

000396



Изискване	Изпълнение (Да/Не)	Начин на реализация	Зав.
<p>едната система в другата;</p> <p>б)четене, модифициране и изтриване на данни от едната система в другата;</p> <p>в)данните в системата MS SQL Server 2017 да бъдат в таблици създадени за работа в паметта (memory-optimized tables in the main memory);</p> <p>г)конфигуриране Аналитичните услуги (Analytical Services) на MS SQL Server 2017 или следваща версия да работят с данни от система Cloudera/Hadoop данни (HDFS файлове), като се съдържа следната информация:</p> <ul style="list-style-type: none"> Аналитичните услуги (Analytical Services) на MS SQL Server 2017 да работят с данни от система Cloudera/Hadoop данни (HDFS файлове). Обяснения логиката на интегрирането; <p>Първичен код;</p> <p>Конфигурационни файлове.</p>		<p>едната система в другата;</p> <p>б)четене, модифициране и изтриване на данни от едната система в другата;</p> <p>в)данните в системата MS SQL Server 2017 да бъдат в таблици създадени за работа в паметта (memory-optimized tables in the main memory);</p> <p>г)конфигуриране Аналитичните услуги (Analytical Services) на MS SQL Server 2017 или следваща версия да работят с данни от система Cloudera/Hadoop данни (HDFS файлове), като се съдържа следната информация:</p> <ul style="list-style-type: none"> Аналитичните услуги (Analytical Services) на MS SQL Server 2017 да работят с данни от система Cloudera/Hadoop данни (HDFS файлове). Обяснения логиката на интегрирането; <p>Първичен код;</p> <p>Конфигурационни файлове.</p>	
<p>физически носител съдържащ изпълняващ</p>	<p>ДА</p>	<p>Прилагаме външен хард диск съдържащ</p>	

000397



Изискване	Изпълнение (Да/Не)	Начин на реализация	Зав
<p>код за Интегриране на двете системи: MS SQL Server 2017 или следваща версия и система Cloudera/Hadoop“, прилагайки PolyBase технологията на MS SQL Server, за четирите типа интегриране:</p> <p>а)експортиране и импортиране на данни от едната система в другата;</p> <p>б)четене, модифициране и изтриване на данни от едната система в другата;</p> <p>в)данните в системата MS SQL Server 2017 или следваща версия да бъдат в таблици създадени за работа в паметта (memory-optimized tables in the main memory);</p> <p>г)конфигуриране Аналитичните услуги (Analytical Services) на MS SQL Server 2017 или следваща версия да работят с данни от система Cloudera/Hadoop (HDFS файлове).</p>		<p>изпълняващ код за Интегриране на двете системи: MS SQL Server 2017 или следваща версия и система Cloudera/Hadoop“, прилагайки PolyBase технологията на MS SQL Server, за четирите типа интегриране:</p> <p>а)експортиране и импортиране на данни от едната система в другата;</p> <p>б)четене, модифициране и изтриване на данни от едната система в другата;</p> <p>в)данните в системата MS SQL Server 2017 или следваща версия да бъдат в таблици създадени за работа в паметта (memory-optimized tables in the main memory);</p> <p>г)конфигуриране Аналитичните услуги (Analytical Services) на MS SQL Server 2017 или следваща версия да работят с данни от система Cloudera/Hadoop (HDFS файлове).</p>	<p>✓</p> <p>✓</p>

000398



ДОКУМЕНТ

„Ръководство за програмиста съдържащо информация за Интегриране на двете системи: MS SQL Server 2017 или следваща версия и система Cloudera/Hadoop“, прилагайки PolyBase технологията на MS SQL Server, за четирите типа интегриране:

а) експортиране и импортиране на данни от едната система в другата;

б) четене, модифициране и изтриване на данни от едната система в другата;

в) данните в системата MS SQL Server 2017 или следваща версия да бъдат в таблици създадени за работа в паметта (memory-optimized tables in the main memory);

г) конфигуриране Аналитичните услуги (Analytical Services) на MS SQL Server 2017 или следваща версия да работят с данни от система Cloudera/Hadoop (HDFS файлове).

Съгласно страница 53, Т.1 от документацията

Т.1.1 а) експортиране и импортиране на данни от едната система в другата, както следва:

Този код използва JDO за връзка към MSSQL сървър 2017 за създаване на таблица и индекс към нея, вмъкване на ред в таблица, получаване на ред, получаване на стойност на колона, изпълнение на заявка и извършване на някои допълнителни операции.

```
package com.apache.hadoop.hbase.client.jdo.examples;  
  
import java.io.File;  
import java.io.FileInputStream;  
import java.io.InputStream;  
import java.util.Hashtable;  
import org.apache.hadoop.fs.Path;  
import org.apache.hadoop.hbase.client.tableindexed.IndexedTable;
```

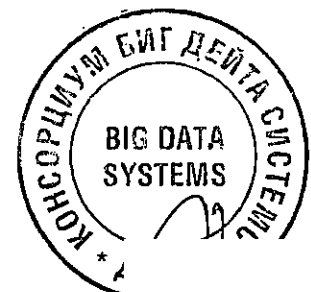


„Консорциум Биг Дейта Системс“

Хай Тех Бизнес
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

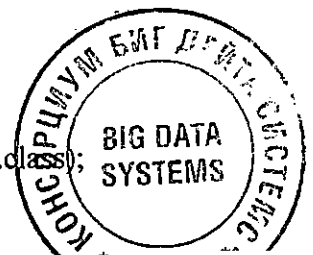
```
import com.apache.hadoop.hbase.client.jdo.AbstractHBaseDBO;
import com.apache.hadoop.hbase.client.jdo.HBaseBigFile;
import com.apache.hadoop.hbase.client.jdo.HBaseDBOImpl;
import com.apache.hadoop.hbase.client.jdo.query.DeleteQuery;
import com.apache.hadoop.hbase.client.jdo.query.HBaseOrder;
import com.apache.hadoop.hbase.client.jdo.query.HBaseParam;
import com.apache.hadoop.hbase.client.jdo.query.InsertQuery;
import com.apache.hadoop.hbase.client.jdo.query.QSearch;
import com.apache.hadoop.hbase.client.jdo.query.SelectQuery;
import com.apache.hadoop.hbase.client.jdo.query.UpdateQuery;
/**
 * Hbase JDO.
 *
 * dependency library.
 * - commons-beanutils.jar
 * - commons-pool-1.5.5.jar
 * - hbase0.90.0-transactionl.jar
 *
 * you can expand Delete,Select,Update,Insert Query classes.
 *
 */
public class HBaseExample {
    public static void main(String[] args) throws Exception {
        AbstractHBaseDBO dbo = new HBaseDBOImpl();

        /**drop if table is already exist.**
```




```
if(dbo.isTableExist("user")){
    dbo.deleteTable("user");
}

/*create table*
dbo.createTableIfNotExist("user",HBaseOrder.DESC,"account");
//dbo.createTableIfNotExist("user",HBaseOrder.ASC,"account");
//create index.
String[] cols={"id","name"};
dbo.addIndexExistingTable("user","account",cols);
//insert
InsertQuery insert = dbo.createInsertQuery("user");
UserBean bean = new UserBean();
bean.setFamily("account");
bean.setAge(20);
bean.setEmail("example@gmail.com");
bean.setId("ncanis");
bean.setName("ncanis");
bean.setPassword("1111");
insert.insert(bean);
//select 1 row
SelectQuery select = dbo.createSelectQuery("user");
UserBean resultBean = (UserBean)select.select(bean.getRow(),UserBean.class);
// select column value.
String value = (String)select.selectColumn(bean.getRow(),"account","id",String.class);
// search with option (QSearch has EQUAL, NOT_EQUAL, LIKE)
```



„Консорциум Биг Дейта Системс“ ДЗЗ

Хай Тех Бизнес Центъ
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

```
// select id,password,name,email from account where id='ncanis' limit
HBaseParam param = new HBaseParam();
param.setPage(bean.getRow(),20);
param.addColumn("id","password","name","email");
param.addSearchOption("id","ncanis",QSearch.EQUAL);
select.search("account", param, UserBean.class);

// search column value is existing.
boolean isExist = select.existColumnValue("account","id","ncanis".getBytes());

// update password.
UpdateQuery update = dbo.createUpdateQuery("user");
Hashtable<String, byte[]> colsTable = new Hashtable<String, byte[]>();
colsTable.put("password","2222".getBytes());
update.update(bean.getRow(),"account",colsTable);

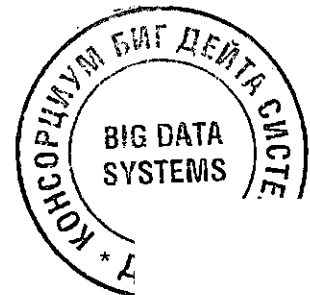
//delete
DeleteQuery delete = dbo.createDeleteQuery("user");
delete.deleteRow(resultBean.getRow());

////////////////////////////////////

// etc

// HTable pool with apache commons pool
// borrow and release. HBasePoolManager(maxActive, minIdle etc..)
IndexedTable table = dbo.getPool().borrow("user");
dbo.getPool().release(table);

// upload bigFile by hadoop directly.
HBaseBigFile bigFile = new HBaseBigFile();
File file = new File("doc/movie.avi");
FileInputStream fis = new FileInputStream(file);
```



```
Path rootPath = new Path("/files/");  
String filename = "movie.avi";  
bigFile.uploadFile(rootPath,filename,fis,true);  
  
// receive file stream from hadoop.  
Path p = new Path(rootPath,filename);  
InputStream is = bigFile.path2Stream(p,4096);  
}  
}
```

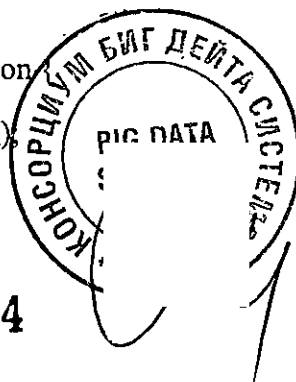
T.1.1 б) да чете, модифицира и изтрива данни от едната система (MS SQL Server) в другата (Специализираната централизирана система под ключ) и обратно, както следва:

Създаване, модифициране и изтриване на Таблица, ползвайки Java :

```
package com.example.hbase.admin;  
  
import java.io.IOException;  
import org.apache.hadoop.conf.Configuration;  
import org.apache.hadoop.fs.Path;  
import org.apache.hadoop.hbase.HBaseConfiguration;  
import org.apache.hadoop.hbase.HColumnDescriptor;  
import org.apache.hadoop.hbase.HConstants;  
import org.apache.hadoop.hbase.HTableDescriptor;  
import org.apache.hadoop.hbase.TableName;  
import org.apache.hadoop.hbase.client.Admin;  
import org.apache.hadoop.hbase.client.Connection;  
import org.apache.hadoop.hbase.client.ConnectionFactory;  
import org.apache.hadoop.hbase.io.compress.Compression.Algorithm;
```



```
public class Example {  
    private static final String TABLE_NAME = "MY_TABLE_NAME_TOO";  
    private static final String CF_DEFAULT = "DEFAULT_COLUMN_FAMILY";  
    public static void createOrOverwrite(Admin admin, HTableDescriptor table) throws  
    IOException {  
        if (admin.tableExists(table.getTableName())) {  
            admin.disableTable(table.getTableName());  
            admin.deleteTable(table.getTableName());  
        }  
        admin.createTable(table);  
    }  
    public static void createSchemaTables(Configuration config) throws IOException {  
        try (Connection connection = ConnectionFactory.createConnection(config);  
            Admin admin = connection.getAdmin()) {  
            HTableDescriptor table = new HTableDescriptor(TableName.valueOf(TABLE_NAME));  
            table.addFamily(new  
            HColumnDescriptor(CF_DEFAULT).setCompressionType(Algorithm.NONE));  
            System.out.print("Creating table. ");  
            createOrOverwrite(admin, table);  
            System.out.println(" Done.");  
        }  
    }  
    public static void modifySchema (Configuration config) throws IOException {  
        try (Connection connection = ConnectionFactory.createConnection(config);  
            Admin admin = connection.getAdmin()) {  
            TableName tableName = TableName.valueOf(TABLE_NAME);
```



```
if (!admin.tableExists(tableName)) {  
    System.out.println("Table does not exist.");  
    System.exit(-1);  
}  
  
HTableDescriptor table = admin.getTableDescriptor(tableName);  
  
// Update existing table  
HColumnDescriptor newColumn = new HColumnDescriptor("NEWCF");  
newColumn.setCompactionCompressionType(Algorithm.GZ);  
newColumn.setMaxVersions(HConstants.ALL_VERSIONS);  
admin.addColumn(tableName, newColumn);  
  
// Update existing column family  
HColumnDescriptor existingColumn = new HColumnDescriptor(CF_DEFAULT);  
existingColumn.setCompactionCompressionType(Algorithm.GZ);  
existingColumn.setMaxVersions(HConstants.ALL_VERSIONS);  
table.modifyFamily(existingColumn);  
admin.modifyTable(tableName, table);  
  
// Disable an existing table  
admin.disableTable(tableName);  
  
// Delete an existing column family  
admin.deleteColumn(tableName, CF_DEFAULT.getBytes("UTF-8"));  
  
// Delete a table (Need to be disabled first)  
admin.deleteTable(tableName);  
}  
}  
  
public static void main(String... args) throws IOException {  
    Configuration config = HBaseConfiguration.create();
```



Хай Тех Бизнес Ц
ж.к. Дружба 1, ул. Мюн
1592 София, РОВ 74
Tel.: +359 2 42 10 10

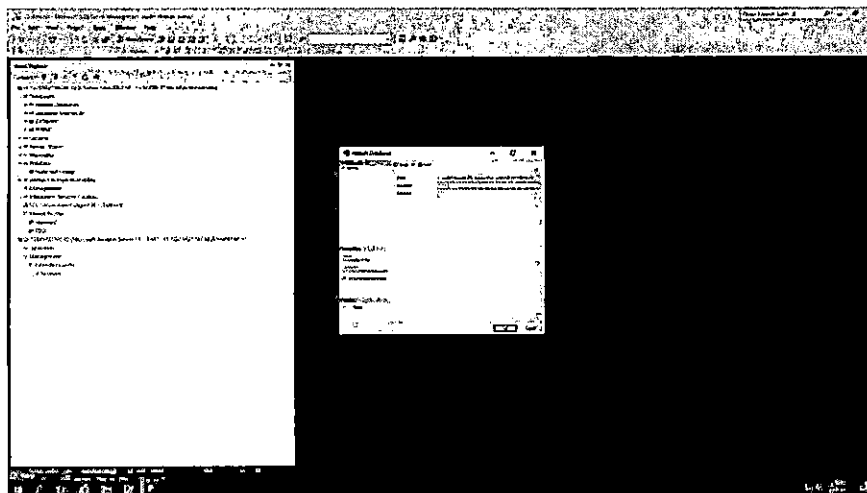
//Add any necessary configuration files (hbase-site.xml, core-site.xml)

```
config.addResource(new Path(System.getenv("HBASE_CONF_DIR"), "hbase-site.xml"));  
config.addResource(new Path(System.getenv("HADOOP_CONF_DIR"), "core-site.xml"));  
createSchemaTables(config);  
modifySchema(config);  
}  
}
```

Т.1.1 в) Данните в системата MS SQL Server да бъдат в таблици създадени за работа в паметта (memory-optimized tables in the main memory), както следва:

```
ALTER DATABASE temeo ADD FILEGROUP temeo_mod CONTAINS MEMORY_OPTIMIZED_DATA;  
ALTER DATABASE temeo ADD FILE (name='temeo1', filename='c:\DATA\temeo1') TO FILEGROUP temeo_mod;
```

Т.1.1 г) Конфигуриране на Аналитични услуги (Analytical Services) на MS SQL Server 2017 да работят с интегрирани данни от Cloudera/Hadoop (HDFS файлове)



Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависимости
-----------	-----------------------	---------------------	-------------

1.3. Документ представящ наличие на програмен продукт разработен от доставчика за интегриране на двете системи MS SQL Server 2017 или следваща версия и система Cloudera/Hadoop, който програмен продукт е с достатъчно ниво на зрялост - с пазарна реализация минимум 6 месец, като документът съдържа банково проследима пазарната реализация;

ДА

Прилагаме Документ представящ наличие на програмен продукт разработен от доставчика за интегриране на двете системи MS SQL Server 2017 или следваща версия и система Cloudera/Hadoop, който програмен продукт е с достатъчно ниво на зрялост - с пазарна реализация минимум 6 месеца, като документът съдържа :

1. Справка за исканите по т.1.3 продадени лицензи в 2018 г.
- Копие на фактурите за продажба
 - Банково извлечение като доказателство за плащането
 - Протоколи от предаването на лицензите



„Документ представящ наличие на програмен продукт разработен от доставчика, за интегриране на двете системи MS SQL Server 2017 или следваща версия и система Cloudera/Hadoop, който програмен продукт е с достатъчно ниво на зрялост - с пазарна реализация минимум 6 месец, като документът съдържа банково проследима пазарната реализация“ съгласно страница. 54 Т 1.3.

Във връзка с проект:

„Доставка, инсталиране, конфигуриране, тестване и поддържане на СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ за работа с Големи данни по проект „Дигитализация на икономиката в среда на Големи данни“ (ДИГД), № BG05M2OP001-1.002-0002“.

Док.№	Дата	Име	Име	Кол.
Ф-566	21.5.2018	Контракс АД	#YL3002610 Machine learning (ML) HADOOP система,	1
Ф-567	23.5.2018	Контракс АД	#YL6002610 Интегриране на MS SQL Server с Hadoop -	1
Ф-569	5.6.2018	ГРАВИС БЪЛГАРИЯ АД	Machine learning (ML) HADOOP система, базирана на PYTHON-	1
Ф-569	5.6.2018	ГРАВИС БЪЛГАРИЯ АД	Машинен анализ (алгоритми) на физиологични данни-	1
Ф-569	5.6.2018	ГРАВИС БЪЛГАРИЯ АД	Интегриране на MS SQL Server с Hadoop - годишен лиценз	1
Ф-572	6.7.2018	ITE Solutions sp. z o.o.	Machine learning (ML) HADOOP система, базирана на PYTHON-	1
Ф-572	6.7.2018	ITE Solutions sp. z o.o.	Комуникационен интерфейс TEMEO за предаване на данни,	1
Ф-572	6.7.2018	ITE Solutions sp. z o.o.	Машинен анализ (алгоритми) на физиологични данни-	1
Ф-572	6.7.2018	ITE Solutions sp. z o.o.	Интегриране на MS SQL Server с Hadoop - годишен лиценз	1
				9

Прилагаме следните документи като доказателство:

1. Копие на фактурите за продажба
2. Банков извлечение като доказателство за плащането
3. Копии от предаването на лицензите за Ф-566; Ф-567; Ф-569, Ф-572

000408

Фактура

за продадени продукция, стоки или услуги

Дол./пор. No и дата:		Дата дан.събитие/платяне:		No:	Дата:	КОПИЕ
		23.05.2018		0000000567	23.05.2018	
Получател, гр. (с.): Контракс АД, ул. Тинтява 13, 1113 София			Обект No:	Доставчик, гр. (с.): СЕКЮРИТИ СОЛЮШЪНС ИНСТИТЮТ ЕООД гр.София ,р-н Искър, ул.Мюнхен 8		Обект No: 10
			Място на сделката : София			
ДДС No: BG 175415627		ЕИК: 175415627		ДДС No: BG 200149195		вик: 200149195
ИВАН:			ИВАН:			
ВИС код:			ВИС код - Уникредит Булевард, ул. Младост 7			
Материалноотговорно лице: Йордан Йорданов			Материалноотговорно лице: Деян Йовков			
No	Код:	Наименование:				
1.	WEB поддр	#YL6002610 Интегриране на MS SQL Server с Nadoop -				
2.		годишен лиценз				
3.	ДДС20	Данък добавена стойност				

Вярно с оригинала

Стоката получена от:				Подпис:	
на дата:		Пълном. No дата:			
Пасп. данни:					
Съставил:		Деян Йовков		Съгласие за незаб. инкасо:	
Разрешил:	Предал:	За	Получател:	Сума	
				Форма на платяне:	
				Счетовод., подпис:	
				Платежно нареждане:	
				000409	



Security Solutions Institute

+359 242 10 100

+359 887 87 81 71

гр. София 1592

ж.к. Дружба I, ул. Мюнхен 8

Хай Тех Бизнес Център

Научно-производствена зона

ПРЕДАВАТЕЛНО - ПРИЕМАТЕЛЕН ПРОТОКОЛ

Ф(ПФ): 567/23.05.2018 г.

Договор No /

Днес 23.05.2018 г., представител на Секюрити Солюшънс Инститют ЕООД предаде на представител на Контракс АД следните стоки:

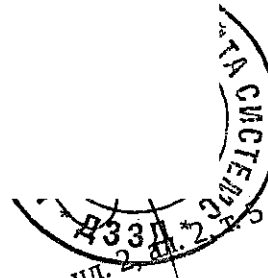
No	Описание	Брой
1	#YL6002610 Интегриране на MS SQL Server с Hadoop - годишен лиценз	1

Настоящият предавателно-приемателен протокол е съставен и подписан в два еднообразни екземпляра, по един за всяка от страните.

Предал:

(име, подпис)

Вярно е г



Заличени на основание чл. 2, ал. 2, т. 5 от ЗЗЛД

(име, подпис)

000410

Извлечение по Разплащателна сметка в BGN / Current account statement in BGN
№ 18/31.05.2018



Адрес: Пл. Света Неделя № 7, София 1000
Address: 7, Sveta Nedelya Sq., 1000 Sofia
Булстат/ЕИК: 831919536 / Bulstat: 831919536

Банков код/ SWIFT/ BIC: UNCRBGSF
Вашият филиал/ Your branch: София Интер Експо център
Вашият мениджър/ Relationship manager name:
Лиляна Цачева Стефанова
Вашият клиентски номер/ Your client's number:

Вашата сметка/ Your IBAN:
Разплащателна сметка в BGN / Current account in BGN

Получател.

СЕКУРИТИ СОЛЮШЪНС ИНСТИТУТ ЕООД

1528 СОФИЯ
Р-Н ИСКЪР
УЛ. МЮНХЕН 8

Уважаеми Госпожо / Господине,

По-долу Ви предоставяме обобщен резултат за всички движения по вашата сметка за периода 26.05.2018 - 30.05.2018

Dear Sir / Madam, Please find below an account statement about the period

Наличност към 26.05.2018

Общо дебит (-)

Общо кредит (+)

Вашето крайно салдо* към 30.05.2018

*Крайно салдо (Бъдеща наличност) отразява собствените средства по сметката с

***Информация за дебитен и кредитен оборот от началото на годината до края на

Детайлна информация за

Дата на транзакция/ Описание
Вальор

30.05.2018 / 30.05.2018 -Възвращаебанков превод FC Контрагент: КОНТРАКС АД
96601034361601 / / ф.566 и 567 По курс



Влоговете Ви в „УниКредит Булбанк“ АД са гарантирани в съгласие с
Your deposits in UniCredit Bulbank AD are guaranteed

е на влоговете в банките.
deposit Guarantee.

Благодарим Ви, че избрахте УниКредит Булбанк

За информация и съдействие позвънете на 0 700 1 84 34 или 1 84 34 от своя мобилен телефон!

For further information and assistance, please call us at 0700 1 84 34 or simply dial 1 84 34 from your mobile!

000411

Факт

за продадени продукция, стоки или услуги

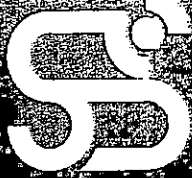
Дог./пор. No и дата:		Дата дан.събитие/платяне:		No:	Дата:	КОПИЕ	
		05.06.2018		0000000569	05.06.2018		
Получател, гр. (с.):			Обект No:	Доставчик, гр. (с.):			Обект No:
ГРАВИС БЪЛГАРИЯ АД, бул. Тодор Каблешков 55А ет.3 офис 5, 1618 София				СЕКЮРИТИ СОЛЮШЪНС ИНСТИТУТ ЕООД гр.София ,р-н Искър, ул.Мюнхен 8			10
			Място на сделката : София				
ДДС No: BG 121467570		ЕИК: 121467570		ДДС No: BG 200111111		ЕИК: 200111111	
IBAN:				IBAN:			
BIC код:				BIC код - Уникредит Булбанк АД, пл. Света Неделя 7			
Материалноотговорно лице:			Материалноотговорно лице:				
Георги Григоров			Деян Йовков				
No	Код:	Наименование:	Марка	Кол-во/%	Цена	Стойност	
1.	лицензи	Machine learning (ML) HADOOP система, базирана на PYTHON-	бр.				
2.		годишен лиценз #YL3002610					
3.	лицензи	Комуникационен интерфейс TEMEO за бр.					
4.		базирани на биосензори - годишен лице					
5.	лицензи	Машинен анализ (алгоритми) на физиол. бр.					
6.		ЕКГ, HR, BP, физ. активност - годишен л					
7.	лицензи	Интегриране на MS SQL Server с Hadoop бр.					
8.		#YL6002610					
9.	ДДС20	Данък добавена стойност					

Вярно с оригинала



000412

Стоката получена от:					
на дата:		Изпълн. No дата:			
Пасп. данни:					
Съставил:		Деян Йовков		Съгласие за изп. инкасо:	
Разрешил:	Предал:	За Получател:	За И:	Сума	
				Форма на плащане:	
Счетовод., подпис:				Платежно нареждане	



Security Solutions Institute

+359 2 42 10 100

+359 887 87 81 71

гр. София 1592

ж.к. Дружба 1, ул. Мюнхен 8

Хай Тех Бизнес Център

Научно-производствена зона

ПРЕДАВАТЕЛНО - ПРИЕМАТЕЛЕН ПРОТОКОЛ

Ф(ПФ): 569/05.06.2018 г.

Договор No /

Днес 05.06.2018 г., представител на Секюрити Солюшънс Инститют ЕООД предаде на представител на Гравис България АД следните стоки:

No	Описание	Брой
1	Machine learning (ML) HADOOP система, базирана на PYTHON- годишен лиценз #YL3002610	1
2	Комуникационен интерфейс TEMEO за предаване на данни базирани на биосензори - годишен лиценз	1
3	Машинен анализ (алгоритми) на физиологични данни- ЕКГ, HR, BP, физ. активност - годишен лиценз	1
4	Интегриране на MS SQL Server с Hadoop - #YL6002610	1

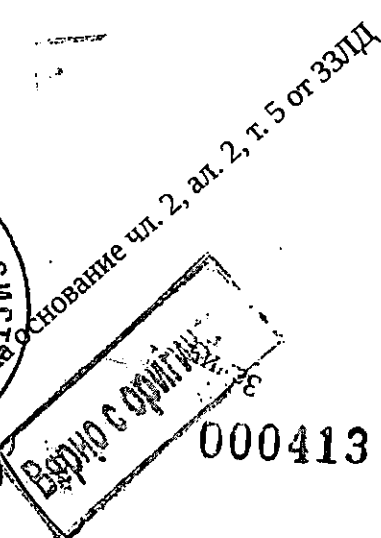
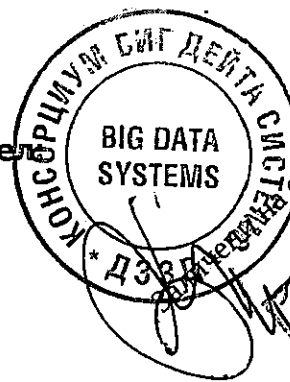
Настоящият предавателно-приемателен протокол е съставен и подписан екземпляра, по един за всяка от страните.

тазни

Предал:

(име, подпис)

Приел:



Извлечение по Разплащателна сметка в BGN / Current account statement in BGN
№ 20/13.06.2018



Адрес: Пл. Света Неделя № 7, София 1000
Address: 7, Sveta Nedelya Sq.; 1000 Sofia
Булстат/БИК: 831919536 / Bulstat: 831919536

Банков код/ SWIFT/ BIC: U
Вашият филиал/ Your branch: София Интер Експо център
Вашият мениджър/ Relationship manager name:
Лиляна Цачева Стефанова

Вашият клиентски номер/ Your client's number: 211000652

Вашата сметка/ Your IBAN:

Разплащателна сметка в BGN / Current account in BGN

Получател:

СЕКЮРИТИ СОЛЮШЪНС ИНСТИТУТ ЕООД

1528 СОФИЯ
Р-Н ИСКЪР
УЛ. МЮНХЕН 8

Уважаеми Госпожо / Господине,

По-долу Ви предоставяме обобщен резултат за всички движения по вашата сметка за периода 01.06.2018 - 12.06.2018

Dear Sir / Madam, Please find below an account statement about the period

Наличност към 01.06.2018

Общо дебит (-)

Общо кредит (+)

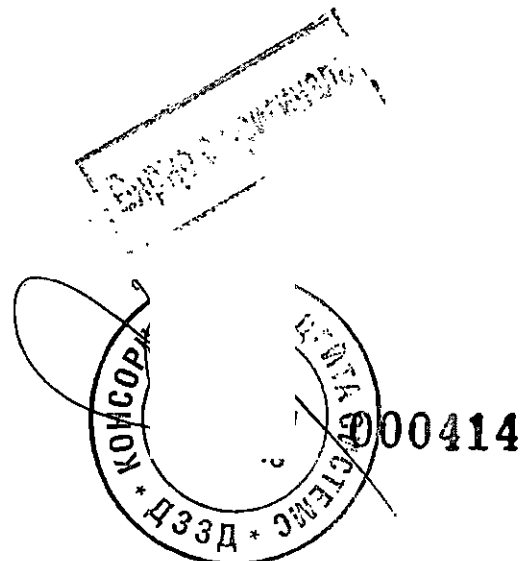
Вашето крайно салдо* към 12.06.2018

*Крайно салдо (бъдеща наличност) отразява собствените средства по сметката си

***Информация за дебитен и кредитен оборот от началото на годината до края на

Детайлна информация за

Дата на транзакция: Валюта	Описание
12.06.2018 / 12.06.2018	-Вътрешнобанков превод РС Контрагент: ПИМ ПРИМА АД 70001501509864 / / афз По курс:
12.06.2018 / 12.06.2018	-Такса за вътрешнобанков превод По курс:
12.06.2018 / 12.06.2018	-Дължима периодична такса Контрагент: 7291СА001 / / СССО ПОДД.И ОБСЛ.НА СКА ЮЛ за 31/05/2018 По курс:
12.06.2018 / 12.06.2018	-Получен междубанков превод 800FOIN1816334Q5 Контрагент: GRAVIS BULGARIA AD AD RZ999155 / BG66RZ9991551005732520 ФАКТУРА 569 05.06.2018г. / / УПо курс:
12.06.2018 / 12.06.2018	-Такса за междубанков превод По курс:



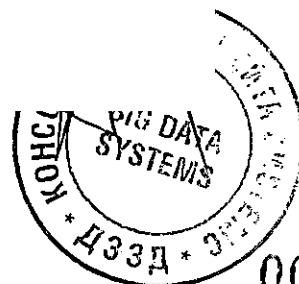
Фактура

за продадени продукция, стоки или услуги

Доз./пор. № и дата:		Дата дан. събитие/плащане:		№:		Дата:		КОПИЕ	
		21.05.2018		0000000566		21.05.2018			
Получател, гр. (с.): Контракс АД, ул. Тинтява 13, 1113 София			Обект №:		Доставчик, гр. (с.): СЕКЮРИТИ СОЛЮШЪНС ИНСТИТУТ ЕООД гр.София ,р-н Искър, ул.Мюнхен 8			Обект №: 10	
			Място на сделката :		София				
ДДС №: BG 175415627			ЕИК: 175415627		ДДС №: BG 200149195		ЕИК: 200149195		
IBAN:					IBAN:				
BIC код:					BIC код - Уникредит Булбанк АД, пл. Света Неделя 7				
Материалноотговорно лице:			Йордан Йорданов		Материалноотговорно лице: Деян Йовков				

No	Код:	Наименование:
1.	WEB поддр	#YL3002610 Machine learning (ML) HADOOP система, базирана на PYTHON - годишен лиценз
2.		
3.		#YL4002610 Комуникационен интерфейс TEMEO за предаване на данни, базирани на бисензори
4.		
5.		#YL5002610 Машинен анализ (алгоритми) на физиологични данни- ЕКТ, HR, BP, физ. активност
6.		
7.	ДДС20	Данък добавена стойност

Вярно с оригинала



000415

Стоката получена от:				Подпис:	
на дата:		Дължим. № дат.:			
Плщ. данни:					
Съставил: Деян Йовков				Съгласие за незаб. инкас:	
Разрешил: Предал:		За Получател:		Сума	
		За Изпълнител:		Форма на плащане:	
				Платежно нареждане	
				Извод., подпис:	

X

1 1



Security Solutions Institute

+359 887 87 81 71

гр. София 1592

ж.к. Дружба 1, ул. Мюнхен 8

Хай Тех Бизнес Център

Научно-производствена зона

ПРЕДАВАТЕЛНО - ПРИЕМАТЕЛЕН ПРОТОКОЛ

Ф(ПФ): 566/21.05.2018 г.

Договор No /

Днес 21.05.2018 г., представител на Секюрити Солюшънс Инститют ЕООД предаде на представител на Контракс АД следните стоки:

No	Описание	Брой
1	#YL3002610 Machine learning (ML) HADOOP система, базирана на PYTHON - годишен лиценз	1
2	Комуникационнен интерфейс TEMEO за предаване на данни базирани на биосензори	1
3	#YL5002610 Машинен анализ (алгоритми) на физиологични данни - ЕКГ, HR, BP, физ. активност	1

Настоящият предавателно-приемателен протокол е съставен и подписан в два еднообразни екземпляра, по един за всяка от страните.

Предал:

(име, подпис)

Приел:

Заличени на основание чл. 2, ал. 2, т. 5 от ЗЗЛД

000416

СЕКЮРИТИ СОЛЮШЪНС ИНСТИТУТ ЕООД

SYSTEMS

ДЗЗД

Извлечение по Разплащателна сметка в BGN / Current account statement in BGN
№ 18/31.05.2018



Адрес: Пл. Света Неделя № 7, София 1000
Address: 7, Sveta Nedelya Sq., 1000 Sofia
Булстат/ЕИК: 831919536 / Bulstat: 831919536

Банков код/ SWIFT/ BIC: UNICB330
Вашият филиал/ Your branch: София Интер Експо център
Вашият мениджър/ Relationship manager name:
Лиляна Цачева Стефанова
Вашият клиентски номер/ Your client's number: 211000652

Вашата сметка/ Your IBAN:
Разплащателна сметка в BGN / Current account in BGN

Получател:

СЕКЮРИТИ СОЛЮШЪНС ИНСТИТУТ ЕООД

1528 СОФИЯ
Р-Н ИСКЪР
УЛ. МЮНХЕН 8

Уважаеми Госпожо / Господине,

По-долу Ви предоставяме обобщен резултат за всички движения по вашата сметка за периода 26.05.2018 - 30.05.2018

Dear Sir / Madam, Please find below an account statement about the pe

Наличност към 26.05.2018

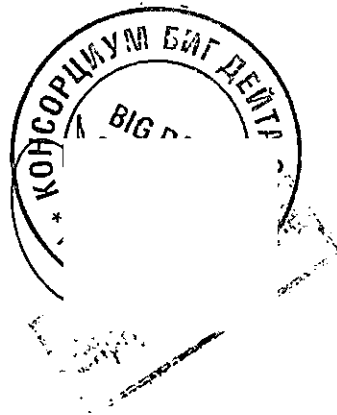
Общо дебит (-)

Общо кредит (+) 10

Вашето крайно салдо* към 30.05.2018 10

*Крайно салдо (бъдеща наличност) отразява собствените средства по сметката сле

***Информация за дебитен и кредитен оборот от началото на годината до края на 01



Детайлна информация за Вашите транзакции:

Дата на транзакция/ Описание
Валютар

30.05.2018 / 30.05.2018 -Вътрешнобанков превод FC Контрагент: КОНТРАКС АД
96601034361601 / ф.566 и 587 По курс:



000417

Влоговете Ви в „УниКредит Булбанк“ АД са гарантирани в съответств

Your deposits in UniCredit Bulbank AD are guaranteed in accordance with the Law on bank deposit guarantee.

Благодарим Ви, че избрахте УниКредит Булбанк!

За информация и съдействие позвънете на 0 700 1 34 84 или 1 34 84 от своя мобилен телефон!

For further information and assistance, please call us at 0700 1 34 34 or simply dial 1 34 34 from your mobile!

Фактура

за продадени продукция, стоки или услуги

Доз./пор. No и дата:		Дата дан.събитие/плащане:		No:	Дата:	КОПИЕ
		06.07.2018		0000000572	06.07.2018	
Получател, гр. (с.):			Обект No:	Доставчик, гр. (с.):		Обект No:
ITE Solutions sp. z o.o., Sztaszica 20/5, Lomianki, Polska, 05-092 Lomianki				СЕКЮРИТИ СОЛЮШЪНС ИНСТИТУТ ЕООД гр.София ,р-н Искър, ул.Мюнхен 8		10
			Място на сделката : София			
ДДС No: BG PL1182126781		ЕИК: PL11821267		ДДС No: BG 200149195		ЕИК: 200149195
IBAN:			IBAN:			
BIC код:			BIC код - УниКредит Булбанк ГД, ШТ. СЪСТА ПЕДЕЛЯ 7			
Материалноотговорно лице:			Материалноотговорно лице:			
ITE Solutions			Деян Йовков			

No	Код:	Наименование:	Марка
1.	лицензи	Machine learning (ML) HADOOP система, базирана на PYTHON- годишен лиценз #YL3002610	бр.
2.			
3.	лицензи	Комуникационен интерфейс TEMEO за	бр.
4.		базирани на биосензори - годишен лице	
5.	лицензи	Машинен анализ (алгоритми) на физиол	бр.
6.		.ЕКГ, HR, BP, физ. активност - годишен л	
7.	лицензи	.Интегриране на MS SQL Server с Hadoop	бр.
8.		#YL6002610	

Вярно с оригинала



000418

Стоката получена от:		ITE Solutions		Подпис:	
на дата:		Пълном. No дата:			
Пасп. данни:					
Съставил:	Деян Йовков			Съгласие за незаб. инкасо:	
Разрешил:	Предал:	За Получател:	За Изпълнител:	Сума	
					Форма на плащане:
				Счетовод., подпис:	Платежно нареждане



ул. Дружба 15, Мюнхен 8
Хай Тек Бизнес Център
Научно-производителна зона

ПРЕДАВАТЕЛНО - ПРИЕМ / ПРОТОКОЛ

Ф(ПФ): 572 / 06.07.2018 г.
Договор No /

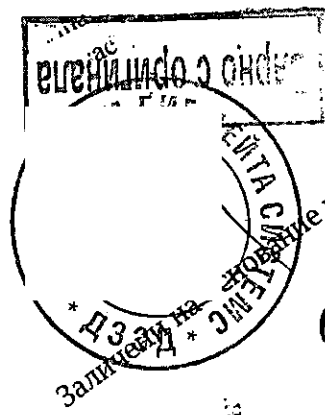
Днес 06.07.2018 г., представител на Секюрити Солюшънс Инститют ЕООД предаде на представител на ITE Solutions следните стоки:

No	Описание	Брой
1	Machine learning (ML) HADOOP система, базирана на PYTHON- годишен лиценз #YL3002610	1
2	Комуникационен интерфейс TEMEO за предаване на данни базирани на биосензори - годишен лиценз	1
3	Машинен анализ (алгоритми) на физиологични данни- ЕКГ, HR, BP, физ. активност - годишен лиценз	1
4	Интегриране на MS SQL Server с Hadoop - #YL6002610	1

Настоящият предавателно-приемателен протокол е съставен и подписан в два еднообразни екземпляра, по един за всяка от страните.

Преда:

Приел:



Закон чл. 2, ал. 2, т. 5 от ЗЗЛД
000419

Извлечение по Разплащателна сметка в EUR / Current account statement in EUR
 № 10/20.07.2018



Адрес: Пл. Света Неделя № 7, София 1000
 Address: 7, Sveta Nedelya Sq., 1000 Sofia
 Булстат/ЕИК: 831919536 / Bulstat: 831919536

Банков код/ SWIFT/ BIC: UNICB330
 Вашият филиал/ Your branch: Корпоративен филиал Света
 Вашият мениджър/ Relationship manager name:
 Лиляна Цачева Стефанова
 Вашият клиентски номер/ Your client's number:

Вашата сметка/ Your IBAN:
 Разплащателна сметка в EUR / Current account in EUR

Получ.
 СЕКИС
 ИНСТИТУТ ЕООД
 1528
 Р-Н И
 УЛ. МЮНХЕН 8

Уважаеми Госпожо / Господине,
 По-долу Ви предоставяме обобщен резултат за всички движения по вашата сметка за периода 01.06.2018 - 19.07.2018
 Dear Sir / Madam, Please find below an account statement about the period

Наличност към 01.06.2018

Общо дебит (-)

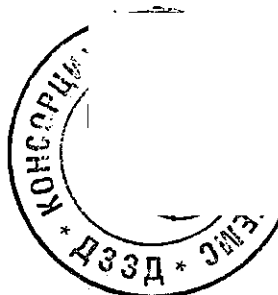
Общо кредит (+)

Вашето крайно салдо* към 19.07.2018

*Крайно салдо (бъдеща наличност) отразява собствените средства по сметката с
 ***Информация за дебитен и кредитен оборот от началото на годината до края на

Детайлна информация за

Дата на транзакция/ Валютор	Описание
19.07.2018 / 19.07.2018	-Получен превод във валута AZV-IT EXCELLENCE SOLUTIONS SPOLKA Z OG. Ord.Ref: NOTPROVIDED. Invoice , GPP Ref. 8200202415. . . ALBPPLPWXXX . PL3724901057000990294241440 Stanislaw Staszica 20 505 092 Lomianki . По курс: 1.9558
19.07.2018 / 19.07.2018	-Такси получени валутни преводи AZV-Commission 10440,00 EUR for GPP transaction Ref.: 8200202415По курс: 1.9559

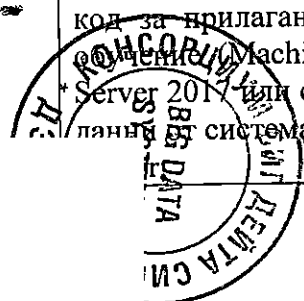


000420

Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависимости
-----------	-----------------------	---------------------	-------------

2. Материали свързани с прилагане на технологията за машинно обучение (Machine Learning – ML) на MS SQL Server 2017 или следваща версия, с използване на данни от система Cloudera / Hadoop и от MS SQL Server:	ДА	Прилагаме материали свързани с технологията за машинно обучение (Machine Learning – ML) на MS SQL Server 2017 или следваща версия, в това число използване на данни от система Cloudera / Hadoop и от MS SQL Server, както следва:	
<p>T.2.1. Документ „Ръководство за прилагане на технологията за машинно обучение (Machine Learning – ML) на MS SQL Server 2017 или следваща версия, с използване на данни от система Cloudera / Hadoop и от MS SQL Server“, в който документ да се съдържа следната информация:</p> <p><input type="checkbox"/> Първичен код;</p> <p><input type="checkbox"/> Конфигурационни файлове.</p>	ДА	<p>Прилагаме Документ „Ръководство за прилагане на технологията за машинно обучение (Machine Learning – ML) на MS SQL Server 2017 или следваща версия, с използване на данни от система Cloudera / Hadoop и от MS SQL Server“.</p> <p>Във външен хард диск се съдържа информация за Първичен код и Конфигурационни файлове.</p>	✓
T.2.2. Магнитен носител съдържащ изпълняващ код за прилагане на технологията за машинно обучение (Machine Learning – ML) на MS SQL Server 2017 или следваща версия, с използване на данни от система Cloudera / Hadoop и от MS SQL	ДА	Прилагаме външен хард диск съдържащ изпълняващ код за прилагане на технологията за машинно обучение (Machine Learning – ML) на MS SQL Server 2017 или следваща версия, с използване на данни от система Cloudera / Hadoop и от MS SQL Server	✓

000421



„Консорциум Биг Дейта Системс“ ДЗЗД

Хай Тех Бизнес Цент
ж.к. Дружба 1, ул. Мюнхен
1592 София, РОВ
Tel.: +359 2 42 10 10

ДОКУМЕНТ:

За предоставяне на материали свързани с прилагане на технологията за машинно обучение (Machine Learning – ML) на MS SQL Server 2017 или следваща версия, с използване на данни от система Cloudera / Hadoop и от MS SQL Server.

Съгласно страница 55, Т.2 и Т.2.1 от документацията

Съгласно Т.2 и Т.2.1, стр. 55 Предоставяме на хартия документ „Ръководство за прилагане на технологията за машинно обучение (Machine Learning – ML) на MS SQL Server 2017 или следваща версия, с използване на данни от система Cloudera / Hadoop и от MS SQL Server“, в който документ да се съдържа следната информация, включваща информация за Machine Learning – ML.

Предоставяме на външен хард диск:

- Първичен код;
- Конфигурационни файлове.



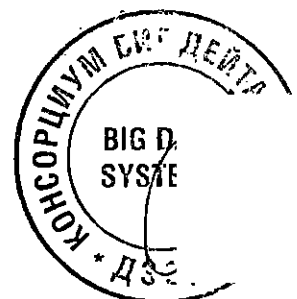
000422

Λ

✓

Ръководство за прилагане на технологията за машинно обучение (Machine Learning – ML) на MS SQL Server 2017 или следваща версия, с използване на данни от системата Cloudera/Hadoop и от MS SQL Server;

000423



Информация за Machine Learning

Machine Learning методи, които се използват при разработка на системи, разпознават отпечатан текст

Изисквания

За изпълнението на задачата са нужни следните средства

- Система с възможност за изпълнение на Python 3.7 код
- Jupyter
- Пакети
 - numpy, matplotlib, scipy, scipy-image, tensorflow, keras
- Файлове
 - scripts.py
 - ocr.ipynb
 - data/training1...16.png
 - tests/adobe.png

Ако има наличен графичен чип с CUDA, препоръчително е да използвате пакета tensorflow-gpu вместо нормалния пакет tensorflow. CUDA и cuDNN трябва да бъдат инсталирани преди изпълнението на който и да е от скриптовете.

Сегментация на изображението

За да може един модел ефективно да разпознае отпечатания текст, той трябва да бъде обучен в голям набор от данни. Въз основа на размера на обучителния комплект, моделът ще може да улавя фини тенденции за всяка буква, което ще е от полза за системата, когато става въпрос за класификация на букви, написани с неизвестен шрифт. Поради тази причина суровият набор от данни, използван за този проект, включваше всички букви от "a" до "z" (малки и главни), както и всички числа от "0" до "9" бяха написани с помощта на 144 различни шрифта. Това доведе до общо 8784 букви и цифри, които моделът би могъл да използва в процеса на обучение и валидиране. Всички букви и цифри бяха разделени на 16 PDF файла, които след това бяха преобразувани в PNG изображения, които системата би могла да анализира.

a b c d e f g h i j k l m n o p q r
s t u v w x y z 0 1 2 3 4 5 6 7 8 9
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
+ , - . : ; ' " & * ()

a b c d e f g h i j k l m n o p q r
s t u v w x y z A B C D E F G H I J K L M N
O P Q R S T U V W X Y Z 1 2 3 4 5 6 7

a b c d e f g h i j k l m n o p q r s
t u v w x y z A B C D E F G H I J K L M N
O P Q R S T U V W X Y Z 1 2 3 4 5 6 7

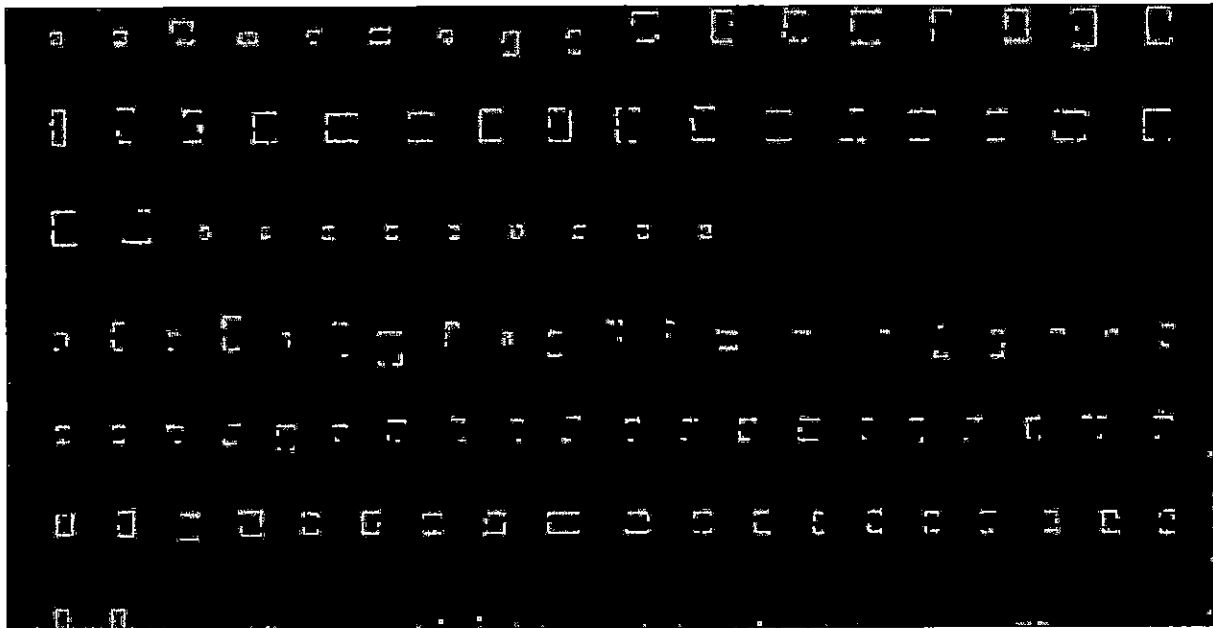


000424 1

Примерен сегмент на изображение, взето от набора от печатни букви. За да може системата да се обучава на тези изображения, те трябва да бъдат сегментирани на изображения с размери 20x20px, състоящи се от един символ. Тези по-малки изображения след това са обозначени със символа, който те представляват и използвани от модела на по-късен етап.

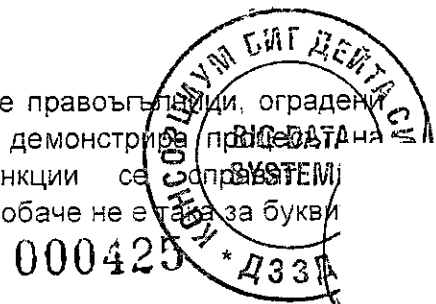
Сегментирането на буквите от суровия набор от данни може да се постигне чрез създаване на региони от групирани пиксели във всяко изображение. Това е възможно благодарение на библиотеката `scikit-image`. Тя предоставя функции за манипулиране на изображения като `clear_border`, която изчиства обекти, свързани с границата на изображението, функцията `label`, която връща списък на всички региони в изображението и функцията `region_props`, която предоставя информация за измерване за всеки регион (`label`) в изображение. Преди да приложите тези функции обаче, изображението трябва да има праг, приложен към всички негови пиксели, за да го направи двоичен.

Когато сегментите са създадени, те могат да бъдат визуализирани на изображението с помощта на 2D артисти, предоставени от библиотеката `matplotlib`. Правоъгълни обекти се създават за всеки регион на маркираното изображение, като се използват измерванията, осигурени от функцията `region_props`. След това артистите се добавят като маркери към графиката за съответното изображение и се показват. За да се избегнат пунктуации като точки, запетайи и тирета, се поставя праг върху размера на всеки регион, в резултат на което само региони по-големи от 50 пиксела получават правоъгълен маркер. Резултатът от горните стъпки е показан на изображението по-долу.

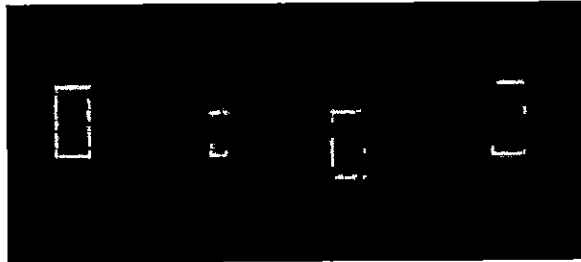


Фиг. 2

Примерен сегмент от маркираното изображение. Празните правоъгълници, оградени от сиво, са разположени върху всеки регион, за да се демонстрира процесът на етикетиране. Както е показано, предоставените функции се използват за сегментирането на изображението на отделни букви. Това обаче не е така за букви и j. Заради точките, които съдържат 2/отделни



региона се създават за точките и за частите от тях. Това не е основен проблем и обучението на модела за разпознаване на *j*, тъй като линията е уникална в сравнение с другите букви / цифри. Случаят е различен при *i*, тъй като линията, в зависимост от използвания шрифт, може да изглежда точно като *l*.

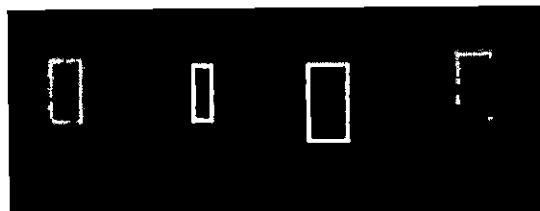


Фиг. 3

training15-1.png

Примерен сегмент на маркираното изображение, съдържащо *i* и *j*

Това би могло да обърка модела по време на фазите на обучение и валидиране, причинявайки спад в общата точност. За да се избегне това препятствие, ще бъде необходимо сливането на точките и линиите на тези букви. Функцията за сливане, предоставена от скриптовия файл, отнема 2 маркера и връща нов правоъгълен маркер, който покрива двата региона. `fix_letters` могат да бъдат използвани за проверка дали правоъгълният маркер е близо до друг правоъгълен маркер и да ги обедини заедно, връщайки актуализиран списък с всички маркери. Функцията позволява входните параметри `xFactor` и `yFactor`, които определят хоризонталните и вертикалните разстояния, които са допустими от линията до точката, и размера, който определя максималния размер на точка, за да се избегне сливането на 2 съседни символа.

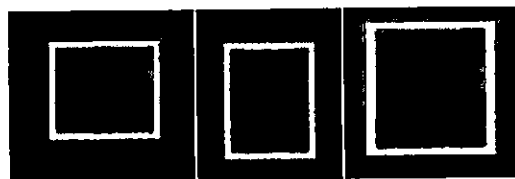


Фиг. 4

training15-1.png

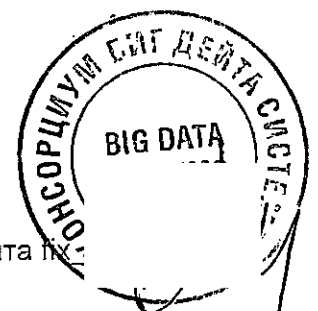
Примерният сегмент на маркираното изображение, съдържащо обединени правоъгълни маркери, предоставени от `fix_letters` функция

Функцията `fix_letters` също осигурява сливане за букви, които са несвързани, като например в примерите по-долу.

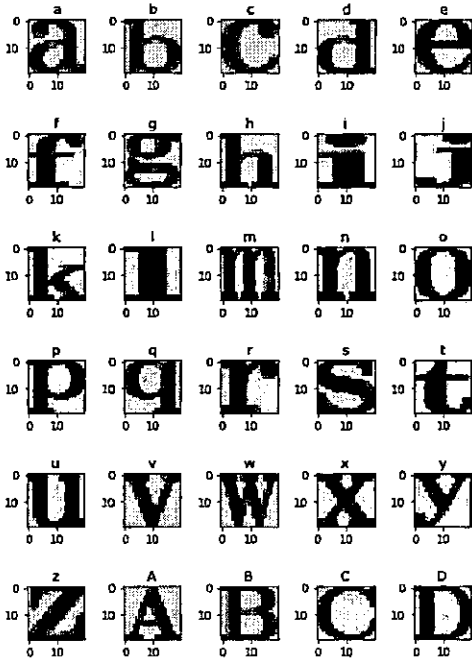


Фиг. 5

Несвързани букви, чиито маркери са правилно обединени от функцията `fix_letters`



След като проблемът с несвързаните букви е решен по полуавтоматичен начин (настройката на параметъра може да е необходима за някои от изображенията на набора от данни), следващата стъпка ще бъде да се разделят правоъгълните маркери в съответните им хоризонтални линии и след това да се подредят по x координатите им по възходящ ред.



Функцията `sort_letters` взема списък от структури на букви (речници, съдържащи 20x20 матрици на пикселите на символа за ключа на изображението, най-лявата x координата за x ключа и най-долната y координата за y ключа) и сортира структурите в редове. За да направи това, той подрежда всичките символи първо по техния y координат. След това се прелиства през вертикално подредените символи. За първия знак (най-нисък y) той създава списък, съдържащ знаците на първия ред и го вмъква. За следващите знаци, той сравнява вертикалното разстояние между предишния символ и текущия. Ако разстоянието е по-голямо от средното вертикално разстояние между всички букви, се създава нов ред и в него се вмъква знакът, в противен случай знакът се вмъква в последния ред. Резултатите от алгоритъма за сортиране са показани вляво, като изображенията на първите 30 елемента на сортираните линии се нанасят и са в правилния ред.

В предоставения Python файл, функцията `build_dataset` отнема папка и низ от маркери като параметри. Папката трябва да съдържа подредени изображения на необработени данни (подобно на изображението, показано на фиг. 1). Името на файла е решаващ фактор за поръчката - `training10-1.png` ще бъде последвано от `training10-2.png`. Всяко изображение трябва също да съдържа партиди от 61 символа (a-z, A-Z, 0-9), така че функцията ще може да маркира всяка област на изображението съответно. По време на изпълнението, броят на символите на всеки файл ще бъде регистриран в конзолата, за да се сигнализира за грешки по време на процеса на сегментиране. Функцията връща редиците X и Y, представляващи данни и маркери, които могат да се използват за обучение на модел на по-късен етап.

Обучение

и

валидация

След като разполагаме с всички налични данни, следващата стъпка ще бъде създаването, обучението и тестването на модел. За задачата за класификация беше избрана конволютивна невронна мрежа (CNN), използваща библиотеката `Keras`. Такива мрежи дават адекватни резултати в визуалната класификация и това е основната причина, поради която CNN е добър избор за задачата, вместо просто да използваме `Deep Neural Network (DNN)` или по-праволинеен алгоритъм като `K-Means`. Архитектурата се състои от входен слой, който приема 128 филтъра и прилага ядро от размер 3x3 и използва функцията за активация `ReLU` за изчисляване на състоянието на неврон. Следващите два слоя имат подобно поведение, с изключение на това, че всеки следващ отнема два пъти по-малко филтри в сравнение с предишния. Функция `Dropout` със скорост 0,5 се прилага на скритите слоеве, за да се оптимизира производителността на модела. Преди изходните данни от последния слой са сбити до броя на символите в низ етикетите. Функция `softmax` разходите се използва по време на фазата на обратното разпространение, за да се намали загубата на модела.

Генерираният по-рано набор от данни се разделя на две части, използвани за обучение и тестване. Препоръчва се по-голямата част да се използва за обучение, което съответства на 80% от данните и съответните маркери, използвани за тази цел. \ / то на данните, моделът е пригоден към данните за обучението и тестване \ / ът на партидата от 128 се използва през 10-те етапа на обучение.

```
Train on 7027 samples, validate on 1757 samples
Epoch 1/10
7027/7027 [-----] - 66s 9ms/step - loss: 1.8476 - acc: 0.5127 - val_loss: 0.5277 - val_acc: 0.7729
Epoch 2/10
7027/7027 [-----] - 65s 9ms/step - loss: 1.0438 - acc: 0.7057 - val_loss: 0.5277 - val_acc: 0.8239
Epoch 3/10
7027/7027 [-----] - 62s 9ms/step - loss: 0.7718 - acc: 0.7636 - val_loss: 0.4746 - val_acc: 0.8327
Epoch 4/10
7027/7027 [-----] - 65s 9ms/step - loss: 0.6260 - acc: 0.7965 - val_loss: 0.4741 - val_acc: 0.8270
Epoch 5/10
7027/7027 [-----] - 64s 9ms/step - loss: 0.5072 - acc: 0.8272 - val_loss: 0.4218 - val_acc: 0.8401
Epoch 6/10
7027/7027 [-----] - 69s 10ms/step - loss: 0.4480 - acc: 0.8455 - val_loss: 0.4297 - val_acc: 0.8469
Epoch 7/10
7027/7027 [-----] - 61s 9ms/step - loss: 0.3826 - acc: 0.8732 - val_loss: 0.4100 - val_acc: 0.8469
Epoch 8/10
7027/7027 [-----] - 70s 10ms/step - loss: 0.3487 - acc: 0.8773 - val_loss: 0.3838 - val_acc: 0.8486
Epoch 9/10
7027/7027 [-----] - 69s 10ms/step - loss: 0.2931 - acc: 0.8899 - val_loss: 0.4169 - val_acc: 0.8475
Epoch 10/10
7027/7027 [-----] - 60s 10ms/step - loss: 0.2675 - acc: 0.8958 - val_loss: 0.4065 - val_acc: 0.8418
```

Фиг. 6

Изход на конзолата по време на обучителния период

По време на процеса на настройка се наблюдава, че точността на обучението и валидирането е сравнително ниска (около 50%) през първия етап. Това се очаква, тъй като моделът е започнал обучението си и не е обработил голяма част от набора от данни. През последната част се получава точност от ~ 90%, базирана на набора от данни за обучението и 84% точност на базата на данните за валидиране. Това показва, че е постигнато значително подобрене по време обработката на набора от данни. По-добри резултати могат да бъдат постигнати, ако наборът от данни се разшири и размерът на етапиър се увеличи. Това обаче идва с цената на изчислителната мощност и времето, прекарано по време на процеса на монтиране.

Приложение

Моделът може да се приложи към сегментирано изображение на текст. Подобни методи на тези, описани в раздела Сегментиране на изображението, могат да бъдат приложени към изображение на отпечатан текст. Това се показва в Прилагането на моделната клетка на IPython файла. Като всеки символ като 20x20px изображение, той може да бъде даден на модела, който след това ще върне набор от прогнози и свързаните с тях сигурности. Предвиждането с най-висока сигурност трябва да бъде избрано. Методът predict_letter осигурява тази функционалност и връща маркера, който съответства на прогнозата с най-висока сигурност. Изпълнението на клетката ще доведе до подобен изход от този на изображението долу (фиг. 7).



000428

dobe.png
in a regulatory document filed with the SEC today
Adobe announced that chief technology officer Kevin
Lynch would be taking his leave as of this coming
Friday
on March 18 2013 Kevin Lynch resigned from his
position as executive vice president chief technology
officer of Adobe systems incorporated effective
March 22 2013 to pursue other opportunities the
filing reads Lynch who came to the company in 2005
during its acquisition of Macromedia led Adobe's
charge into some of the more cutting edge areas of
technology including multiscreen computing cloud
computing and social media
For ages Adobe had been rooted in the workflows
of the print design communityz Lynch was responsible
for the company's shift into web publishing starting
with Dreamweaver He also oversaw Adobe's research
and experience design teams and was as Adobe puts
it in charge of shaping Adobe's long term
technology vision and focusing innovation across the
company during a transformative time
Rumors around the web have pinpointed Apple as
Lynch's next destination and its not an entirely
nonsensical rum or Adobe's transition to web
technologies has been nothing if not profitable
Apple still a giant in consumer hardware could use
a helping hand when it comes to multiscreen
fluidity social media and web based software

Фиг. 7

Въз основа на резултата от класификацията може да се заключи, че моделът прави адекватна работа при класифицирането на знаците. Значителна грешка е грешната класификация на I с x в началото на текста. Това може да се дължи на недохранването на невронната мрежа, тъй като е обучена само на 8784 знака. MNIST е популярен набор от данни, който съдържа 70000 проби от ръкописни числа. Данните от такова количество биха били по-подходящи за обучението на модела и би трябвало да доведат до по-малко грешки като тези, наблюдавани на фиг. 7. Всички други грешки се дължат на неправилно разпознаване на букви като S и C, тъй като те имат много сходни главни букви. и малки букви. Като по-нататъшно усъвършенстване на системата, вместо да се преоразмерява всяка от буквите до 20×20 px, трябва да се вземе маркера с най-голям размер и всички останали маркери да се запълнят с пиксели с 0 интензивност, да съответстват на размера на най-големия маркер. По този начин моделът би могъл правилно да прави разлика между главните и малките букви.

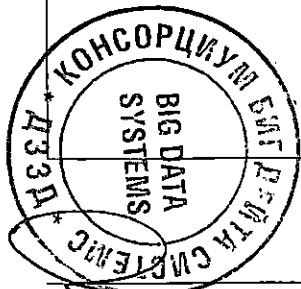


000429

Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависимости
-----------	-----------------------	---------------------	-------------

<p>2.3. Документ представящ наличие на програмен продукт разработен от доставчика, за прилагане на технологията за машинно обучение (Machine Learning – ML) на MS SQL Server 2017 или следваща версия, с използване на данни от система Cloudera / Hadoop и от MS SQL Server, който програмен продукт е с достатъчно ниво на зрялост - с пазарна реализация минимум 6 месец, като документът съдържа банково проследима пазарната реализация:</p>	<p>ДА</p>	<p>Прилагаме Документ представящ наличие на програмен продукт разработен от доставчика, за прилагане на технологията за машинно обучение (Machine Learning – ML) на MS SQL Server 2017 или следваща версия, с използване на данни от система Cloudera / Hadoop и от MS SQL Server, който програмен продукт е с достатъчно ниво на зрялост - с пазарна реализация минимум 6 месец, като документът съдържа:</p> <ol style="list-style-type: none"> 1. Справка за исканите в Т 2.3 продадени лицензи в 2018 г. - Копие на фактурите за продажба - Банково извлечение като доказателство за плащането - Протоколи от предаването на лицензите. 	<p>C</p>
---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------

000430



„Документ представящ наличие на програмен продукт разработен от доставчика, за прилагане на технологията за машинно обучение (Machine Learning – ML) на MS SQL Server 2017 или следваща версия, с използване на данни от система Cloudera / Hadoop и от MS SQL Server, който програмен продукт е с достатъчно ниво на зрялост - с пазарна реализация минимум 6 месец, като документът съдържа банково проследима пазарната реализация“ съгласно страница. 55 Т 2.3.

Във връзка с проект:

„Доставка, инсталиране, конфигуриране, тестване и поддържане на СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ HADOOP за работа с Големи данни по проект „Дигитализация на икономиката в среда на Големи данни“ (ДИГД), № BG05M2OP001-1.002-0002“.

Док.№	Дата	Име	Име	Кол.
Ф-566	21.5.2018	Контракс АД	#YL3002610 Machine learning (ML) HADOOP система,	1
Ф-567	23.5.2018	Контракс АД	#YL6002610 Интегриране на MS SQL Server с Hadoop -	1
Ф-569	5.6.2018	ГРАВИС БЪЛГАРИЯ АД	Machine learning (ML) HADOOP система, базирана на PYTHON-	1
Ф-569	5.6.2018	ГРАВИС БЪЛГАРИЯ АД	Машинен анализ (алгоритми) на физиологични данни-	1
Ф-569	5.6.2018	ГРАВИС БЪЛГАРИЯ АД	Интегриране на MS SQL Server с Hadoop - годишен лиценз	1
Ф-572	6.7.2018	ITE Solutions sp. z o.o.	Machine learning (ML) HADOOP система, базирана на PYTHON-	1
Ф-572	6.7.2018	ITE Solutions sp. z o.o.	Комуникационен интерфейс TEMEO за предаване на данни,	1
Ф-572	6.7.2018	ITE Solutions sp. z o.o.	Машинен анализ (алгоритми) на физиологични данни-	1
Ф-572	6.7.2018	ITE Solutions sp. z o.o.	Интегриране на MS SQL Server с Hadoop - годишен лиценз	1
				9

000431

Прилагам следните документи като доказателство:

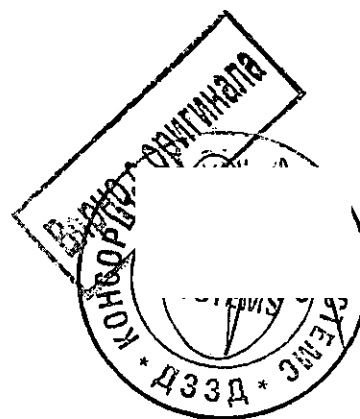
1. Копие на фактурите за продажба
2. Банково извлечение като доказателство за плащането и от предаването на лицензите за Ф-566; Ф-567; Ф-569, Ф-572

ФАКТ

за продадени продукция, стоки или услуги

Тог./пор. № и дата:		Дата дан. събитие/платяне:		№:		Дата:		КОПИЕ	
		05.06.2018		0000000569		05.06.2018			
Получател, гр. (с.):				Обект №:		Доставчик, гр. (с.):		Обект №:	
ГРАВИС БЪЛГАРИЯ АД, бул. Тодор Каблешков 55А ет.3 офис 5, 1618 София						СЕКЮРИТИ СОЛЮШЪНС ИНСТИТУТ ЕООД гр.София ,р-н Искър, ул.Мюнхен 8		10	
				Място на сделката:		София			
ДДС №: BG 121467570		ЕИК: 121467570		ДДС №: BG 200149195		ЕИК: 200149195			
ИВАН:				ИВАН:					
ВИС код:				ВИС код		- Уникредит Булбанк АД, пл. Света Неделя 7			
Материалноотговорно лице:				Материалноотговорно лице:		Деян Йовков			

No	Код:	Наименование:	Марка
1.	лицензи	Machine learning (ML) HADOOP система, базирана на PYTHON-	бр.
2.		годишен лиценз #YU3002610	
3.	лицензи	Комуникационен интерфейс TEMEO за бр.	
4.		базирани на биосензори - годишен лиценз	
5.	лицензи	Машинен анализ (алгоритми) на физиол	бр.
6.		ЕКГ, HR, BP, физ. активност - годишен п	
7.	лицензи	Интегриране на MS SQL Server с Hadoop	бр.
8.		#YU6002610	
9.	ДДС20	Данък добавена стойност	



Стоката получена от:		Пълном. № дат.:			
на дата:					
Пасп. данни:					
Съставил:		Деян Йовков		Съгласие за изп. инкасо:	
Разрешил:		Предал:		Сума	
За Получател:		За Изпълнител:			
				Форма на платяне:	
				Платежно нареждане	
				овод., подпис:	

000432

ПРЕДАВАТЕЛНО - ПРИЕМАТЕЛЕН ПРОТОС

Ф(ПФ): 569/05.0
 Договор No

Днес 05.06.2018 г., представител на Секюрити Солюшънс Инститют ЕООД предаде на представител на Гравис България АД следните стоки:

No	Описание	Брой
1	Machine learning (ML) HADOOP система, базирана на PYTHON- годишен лиценз #YL3002610	1
2	Комуникационен интерфейс TEMEO за предаване на данни базирани на биосензори - годишен лиценз	1
3	Машинен анализ (алгоритми) на физиологични данни- ЕКГ, HR, BP, физ. активност - годишен лиценз	1
4	Интегриране на MS SQL Server с Hadoop - #YL6002610	1

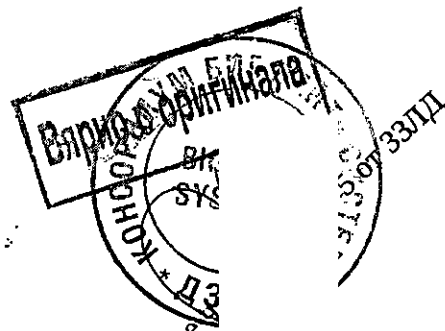
Настоящият предавателно-приемателен протокол е съставен и подписан в два еднообразни екземпляра, по един за всяка от страните.

Предал:

(име, подпис)

Прие

Заложени на основание



Адрес: Билег

000433

Извлечение по Разплащателна сметка в BGN / Current account statement in BGN
№ 20/13.06.2018



Адрес: Пл. Света Неделя № 7, София 1000
Address: 7, Sveta Nedelya Sq., 1000 Sofia
Булстат/ЕИК: 831919536 / Bulstat: 831919536

Банков код/ SWIFT/ BIC: UNCRBGSF
Вашият филиал/ Your branch: София Интер Експо център
Вашият мениджър/ Relationship manager name:
Лиляна Цачева Стефанова
Вашият клиентски номер/ Your client's number: 211000652
Вашата сметка/ Your IBAN:
Разплащателна сметка в BGN / Current account in BGN

Получател:

СЕКЮРИТИ СОЛЮШЪНС ИН

1528 СОФИЯ
Р-Н ИСКЪР
УЛ. МЮНХЕН 8

Уважаеми Госпожо / Господине,
По-долу Ви предоставяме обобщен резултат за всички движения по вашата сметка за периода 01.06.2018 - 12.06.2018
Dear Sir / Madam, Please find below an account statement about the period

Наличност към 01.06.2018

Общо дебит (-)

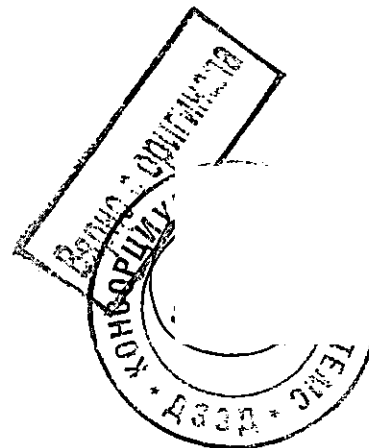
Общо кредит (+)

Вашето крайно салдо* към 12.06.2018

*Крайно салдо (бъдеща наличност) отразява собствените средства по сметката
***Информация за дебитен и кредитен оборот от началото на годината до края

Детайлна информация :

Дата на транзакция/ Валюта	Описание
12.06.2018 / 12.06.2018	-Вътрешнобанков превод FC Контрагент: ПИМ ПРИМА АД 70001501509864 / / ефз По курс:
12.06.2018 / 12.06.2018	-Такса за вътрешнобанков превод По курс:
12.06.2018 / 12.06.2018	-Дължима периодична такса Контрагент: 7291СА001 / / СССР ПОДД.И ОБСЛ.НА С/КА ЮЛ за 31/05/2018 По курс:
12.06.2018 / 12.06.2018	-Получен междубанков превод 800POIN181633405 Контрагент: GRAVIS BULGARIA AD AD RZBB9155 / BG66RZBB91551005732520 ФАКТУРА 569 05.06.2018Г. / УПо курс:
12.06.2018 / 12.06.2018	-Такса за междубанков превод По курс:

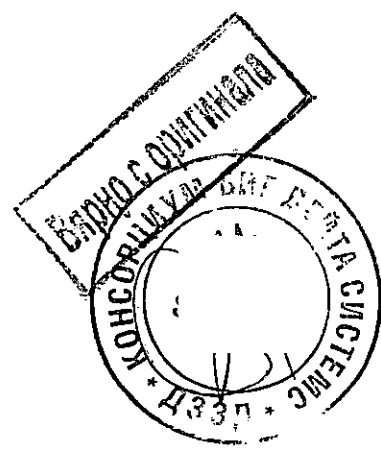


000434

Фактур

за продадени продукция, стоки или услуги

Дог./пор. № и дата:		Дата дан.събитие/плащане:		№:		Дата:		КОПИЕ		
		23.05.2018		0000000567		23.05.2018				
Получател, гр. (с.): Контракс АД, ул. Тинтява 13, 1113 София			Обект №:		Доставчик, гр. (с.): СЕКЮРИТИ СОЛЮШЪНС ИНСТИТЮТ ЕООД гр.София ,р-н Искър, ул.Мюнхен 8			Обект №: 10		
			Място на сделката : София							
ДДС №: BG 175415627		ЕИК: 175415627		ДДС №: BG 200149195		ЕИК: 200149195				
IBAN:			IBAN:							
BIC код:			BIC код - Уникредит Булбанк АД, пл. Света Неделя 7							
Материалноотговорно лице: Йордан Йорданов			Материалноотговорно лице: Деян Йовков							
№	Код:	Наименование:					Мярка			
1.	WEB поддр	#YL60026 10 Интегриране на MS SQL Server с Hadoop -					бр.			
2.		годишен лиценз								
3.	ДДС20	Данък добавена стойност								



000435

на дата:		Пълном. № дата:		Подпис:	
Пасп. данни:					
Съставил:		Деян Йовков		Съгласие за незаб. инкасo:	
Разрешил:		Предал:		За Изпълнител:	
				Сума	
				Счетовод., подпис:	
				Платежно нареждане	



Security Solutions Institute

+359 242 10 100

+359 887 87 81 71

гр. София 1592

ж.к. Дружба 1, ул. Мюнхен 8

Хай Тех Бизнес Център

Научно-производствена зона

ПРЕДАВАТЕЛНО - ПРИЕМАТЕЛЕН ПРОТОКОЛ

Ф(ПФ): 567/23.05.2018 г.

Договор No /

Днес 23.05.2018 г., представител на Секюрити Солюшънс Инститют ЕООД предаде на представител на Контракс АД следните стоки:

No	Описание	Брой
1	#YL6002610 Интегриране на MS SQL Server с Hadoop - годишен лиценз	1

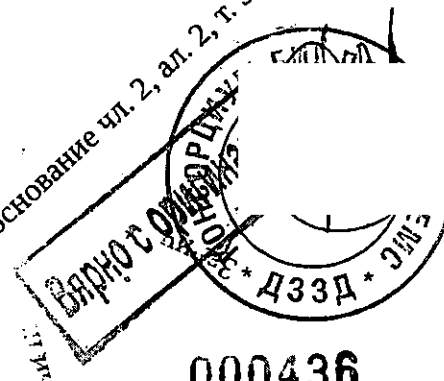
Настоящият предавателно-приемателен протокол е съставен и подписан в два еднообразни екземпляра, по един за всяка от страните.

Предал:

(име, подпис)

Приел:

Заличени на основание чл. 2, ал. 2, т. 5 от ЗЗЛД



000436

Извлечение по Разплащателна сметка в BGN / Current account statement in BGN
№ 18/31.05.2018

 **UniCredit Bulbank**

Адрес: Пл. Света Неделя № 7, София 1000
Address: 7, Sveta Nedelya Sq.; 1000 Sofia
Булстат/ЕИК: 831919536 / Bulstat: 831919536

Банков код/ SWIFT/ BIC: UNCRBGSF
Вашият филиал/ Your branch: София Интер Експо център
Вашият мениджър/ Relationship manager name:
Лиляна Цачева Стефанова

Вашият клиентски номер/ Your client's number: 211000652

Вашата сметка/ Your IBAN:

Разплащателна сметка в BGN / Current account in BGN

Получател:

СЕКЮРИТИ СОЛЮШЪНС ИНСТИТУТ ЕООД

1528 СОФИЯ
Р-Н ИСКЪР
УЛ. МЮНХЕН 8

Уважаеми Госпожо / Господине,

По-долу Ви предоставяме обобщен резултат за всички движения по вашата сметка за периода 26.05.2018 - 30.05.2018

Dear Sir / Madam, Please find below an account statement about the period

Наличност към 26.05.2018

Общо дебит (-)

Общо кредит (+)

Вашето крайно салдо* към 30.05.2018

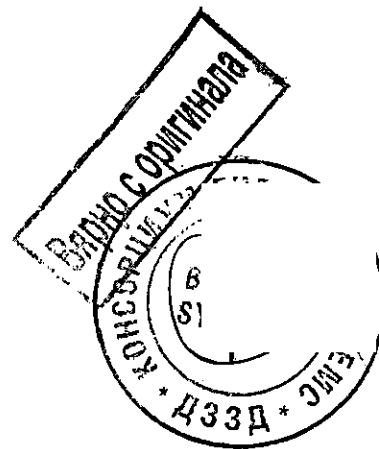
*Крайно салдо (бъдеща наличност) отразява собствениите средства по сметката

**Информация за дебитен и кредитен оборот от началото на годината до края

Детайлна информация

Дата на транзакция / Описание
Валюор

30.05.2018 / 30.05.2018 - Вътрешнобанков превод ФС Контрагент: КОНТРАКС АД
96601034381601 / / ф.566 и 567 По курс:



000437

Влоговете Ви в „УниКредит Булбанк“ АД са гарантирани в съответствие с
Your deposits in UniCredit Bulbank AD are guaranteed in accordance with

Благодарим Ви, че избрахте УниКредит Булбанк!

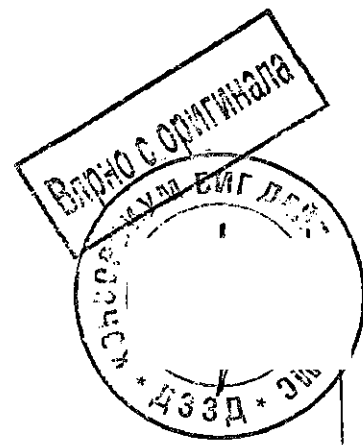
За информация и съдействие позвънете на 0 700 1 84 84 или 1 84 84 от своя мобилен телефон!

Фактура

за продадени продукция, стоки или услуги

Дог/пор. No и дата:		Дата на събитие/плащане:		No:	Дата:	КОПИЕ
		21.05.2018		0000000566	21.05.2018	
Получател, гр. (с.): Контракс АД, ул. Тинтява 13, 1113 София		Обект No:	Доставчик, гр. (с.): СЕКЮРИТИ СОЛЮШЪНС ИНСТИТУТ ЕООД гр.София ,р-н Искър, ул.Мюнхен 8			Обект No: 10
		Място на сделката : София				
ДДС No: BG 175415627		ЕИК: 175415627	ДДС No: BG 200149195		ЕИК: 200149195	
IBAN:		IBAN:				
BIC код:		BIC код - Уникредит Булбанк АД, пл. Света Неделя 7				
Материалноотговорно лице:		Материалноотговорно лице:				
Йордан Йорданов		Деян Йовков				

No	Код:	Наименование:	Мярка
1.	WEB поддр	#YL3002610 Machine learning (ML) HADOOP система,	бр.
2.		базирана на PYTHON - годишен лиценз	
3.		#YL4002610 Комуникационен интерфейс TEMEO за предаване	
4.		на данни, базирани на биосензори	
5.		#YL5002610 Машинен анализ (алгоритми) на физиологични данни-	
6.		ЕКГ, HR, BP, физ. активност	
7.	ДДС20	Данък добавена стойност	



000438

на дата:	Пълном. No дата:			
Пасп. данни:				
Съставил:	Деян Йовков			Съгласие за незаб. инкасо:
Разрешил:	Предал:	За Получател:	За Изпълнител:	Сума
				Форма на плащане:
				Счетовод., подпис: Платежно нареждане



Security Solutions Institute

+359 2 42 10 100

+359 887 87 81 71

гр. София 1592

ж.к. Дружба I, ул. Мюнхен 8

Хай Тех Бизнес Център

Научно-производствена зона

ПРЕДАВАТЕЛНО - ПРИЕМАТЕЛЕН ПРОТОКОЛ

Ф(ПФ): 566/21.05.2018 г.

Договор No 1

Днес 21.05.2018 г., представител на Секюрити Солюшънс Инститют ЕООД предаде на представител на Контракс АД следните стоки:

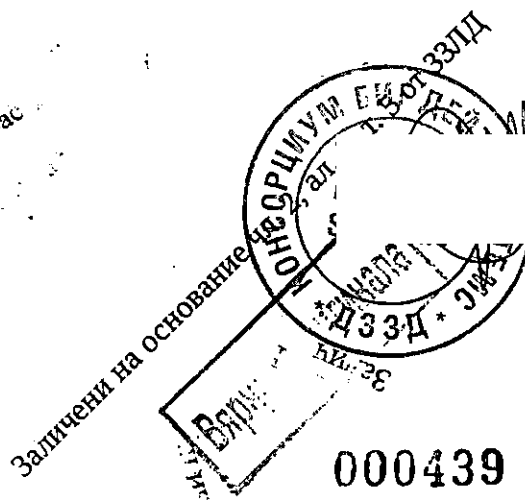
No	Описание	Брой
1	#YL3002610 Machine learning (ML) HADOOP система, базирана на PYTHON - годишен лиценз	1
2	Комуникационен интерфейс TEMEO за предаване на данни базирани на биосензори	1
3	#YL5002610 Машинен анализ (алгоритми) на физиологични данни - ЕКГ, HR, BP, физ. активност	1

Настоящият предавателно-приемателен протокол е съставен и подписан в два еднообразни екземпляра, по един за всяка от страните.

Предал:

(име, подпис)

Приел:



Извлечение по Разплащателна сметка в BGN / Current account statement in BGN
№ 18/31.05.2018

 UniCredit Bulbank

Адрес: Пл. Света Неделя № 7, София 1000
Address: 7, Sveta Nedelya Sq., 1000 Sofia
Булстат/ЕИК: 831919536 / Bulstat: 831919536

Банков код/ SWIFT/ BIC: UNCRBGSF
Вашият филиал/ Your branch: София Интер Експо център
Вашият мениджър/ Relationship manager name:
Лиляна Цачева Стефанова

Вашият клиентски номер/ Your client's number: 211000652

Вашата сметка/ Your IBAN:

Разплащателна сметка в BGN / Current account in BGN

Получател:

СЕКЮРИТИ СОЛЮШЪНС ИНСТИТУТ ЕО

1528 СОФИЯ
Р-Н ИСКЪР
УЛ. МЮНХЕН 8

Уважаеми Госпожо / Господине,

По-долу Ви предоставяме обобщен резултат за всички движения по вашата сметка за периода 26.05.2018 - 30.05.2018

Dear Sir / Madam, Please find below an account statement about the period

Наличност към 26.05.2018

Общо дебит (-)

Общо кредит (+)

Вашето крайно салдо* към 30.05.2018

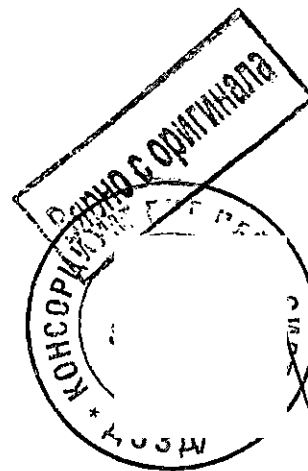
*Крайно салдо (бъдеща наличност) отразява собствените средства по сметката сг

***Информация за дебитен и кредитен оборот от началото на годината до края на

Детайлна информация за

Дата на транзакция/ Описание
Валюор

30.05.2018 / 30.05.2018 -Вътрешнобанков превод FC Контрагент: КОНТРАКС АД
96601034361601 / / ф.565 и 567 По курс:



000440

Влоговете Ви в „УниКредит Булбанк“ АД са гарантирани в съответствие със Закона за гарантиране на влоговете в банките.
Your deposits in UniCredit Bulbank AD are guaranteed in accordance with the Law on Bank Deposit Guarantee.
Благодарим Ви, че избрахте УниКредит Булбанк!

За информация и съдействие позвънете на 0 700 1 84 84 или 1 84 84 от своя мобилен телефон!

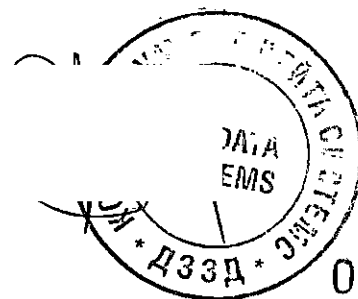
Фактура

за продадени продукция, стоки или услуги

Дог./пор. No и дата:		Дата дан.събитие/платяне:		No:	Дата:	КОПИЕ
		06.07.2018		0000000572	06.07.2018	
Получател, гр. (с.): ITE Solutions sp. z o.o., Sztaszica 20/5, Lomianki, Polska, 05-092 Lomianki			Обект No:	Доставчик, гр. (с.): СЕКЮРИТИ СОЛЮШЪНС ИНСТИТУТ ЕООД гр.София ,р-н Искър, ул.Мюнхен 8		Обект No: 10
			Място на сделката : София			
ДДС No: BG PL1182126781		ЕИК: PL11821267		ДДС No: BG 200149195		ЕИК: 200149195
IBAN:			IBAN:			
BIC код:			BIC код - Уникредит Булбанк АД, пл. Света Неделя 7			
Материалноотговорно лице:			Материалноотговорно лице:			
ITE Solutions			Деян Йовков			

No	Код:	Наименование:	Марка
1.	лицензи	Machine learning (ML) HADOOP система базирана на PYTHON-	бр.
2.		годишен лиценз #YL3002610	
3.	лицензи	Комуникационен интерфейс TEMEO за	бр.
4.		базирани на биосензори - годишен лице	
5.	лицензи	Машинен анализ (алгоритми) на физиол	бр.
6.		ЕКГ, HR, BP, физ. активност - годишен л	
7.	лицензи	Интегриране на MS SQL Server с Hadoop	бр.
8.		#YL6002610	

Варно с одитиране



000441

Стоката получена от:		ITE Solutions		Подпис:	
на дата:		Пълном. No дата:			
Пасл. данни:					
Съставил:	Деян Йовков			Съгласие за неизп. инкасо:	
Разрешил:	Предал:	За Получател:	За Изпълнител:	Сума	
				Счетовод., подпис:	Платежно нареждане

Извлечение по Разплащателна сметка в EUR / Current account statement in EUR
№ 10/20.07.2018



Адрес: Пл. Света Неделя № 7, София 1000
Address: 7, Sveta Nedelya Sq., 1000 Sofia
Булстат/ЕИК: 831919536 / Bulstat: 831919536

Банков код/ SWIFT/ BIC: UNCRBGSF
Вашият филиал/ Your branch: Корпоративен филиал Света
Вашият мениджър/ Relationship manager name:
Лиляна Цачева Стефанова
Вашият клиентски номер/ Your client's number: 211000652

Вашата сметка/ Your IBAN:
Разплащателна сметка в EUR / Current account in EUR

Получател:

СЕКЮРИТИ СОЛЮШЪНС

1528 СОФИЯ
Р-Н ИСКЪР
УЛ. МЮНХЕН 8

Уважаеми Госпожо / Господине,

По-долу Ви предоставяме обобщен резултат за всички движения по вашата сметка за периода 01.06.2018 - 19.07.2018

Dear Sir / Madam, Please find below an account statement about the period

Наличност към 01.06.2018

Общо дебит (-)

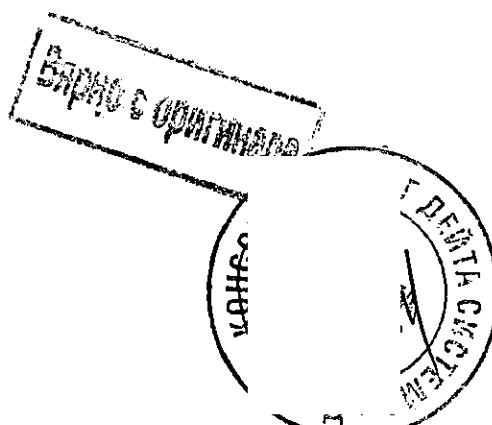
Общо кредит (+)

Вашето крайно салдо* към 19.07.2018

*Крайно салдо (бъдеща наличност) отразява собствените средства по сметка
**Информация за дебитен и кредитен оборот от началото на годината до края

Детайлна информация

Дата на транзакция/ Вальор	Описание
19.07.2018 / 19.07.2018	-Получен превод във валута AZV-IT EXCELLENCE SOLUTIONS SPOLKA Z OG, Ord.Ref: NOTPROVIDED... Invoice , GPP Ref.: 8200202415... ALBPPLPWXX, PL0724901057000990294241440 Stanislaw Staszica 20 505 092 Lomianki , По курс: 1.9558
19.07.2018 / 19.07.2018	-Такси получени валутни преводи AZV-Commission 10440.00 EUR for GPP transaction Ref.: 8200202415По курс: 1.9559



000442

Влоговете Ви в „УниКредит Булбанк“ АД са гарантирани в съответствие със Закона за гарантиране на влоговете в банките.
Your deposits in UniCredit Bulbank AD are guaranteed in accordance with the Law on Bank Deposit Guarantee.

За информация и съдействие позвънете на 0 700 1 84 84 или 1 84 84 от своя мобилен телефон!
For further information and assistance, please call us at 0700 1 84 84 or simply dial 1 84 84 from your mobile!



Security Solutions Institute

+359 2 42 10 100

+359 887 87 81 71

гр. София 1592

ж.к. Дружба 1, ул. Мюнхен 8

Хай Тех Бизнес Център

Научно-производствена зона

ПРЕДАВАТЕЛНО - ПРИЕМАТЕЛЕН ПРОТОКОЛ

Ф(ПФ): 572/06.07.2

Договор No

Днес 06.07.2018 г., представител на Секюрити Солюшънс Инститют ЕООД предаде на представител на ITE Solutions следните стоки:

No	Описание	Брой
1	Machine learning (ML) HADOOP система, базирана на PYTHON- годишен лиценз #YL3002610	1
2	Комуникационен интерфейс TEMEO за предаване на данни базирани на биосензори - годишен лиценз	1
3	Машинен анализ (алгоритми) на физиологични данни- ЕКГ, HR, BP, физ. активност - годишен лиценз	1
4	Интегриране на MS SQL Server с Hadoop - #YL6002610	1

Настоящият предавателно-приемателен протокол е съставен и подписан в два еднообразни екземпляра, по един за всяка от страните.

Преда:

Приел:

Сигнатура

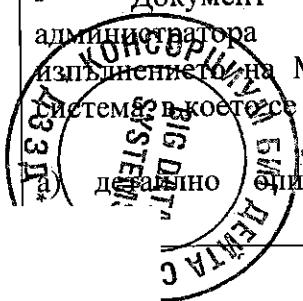
Заличени на основание чл. 2, ал. 2, т. 3 от ЗЗЛД

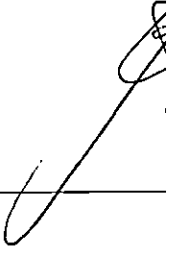
000443

Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависимости
-----------	-----------------------	---------------------	-------------

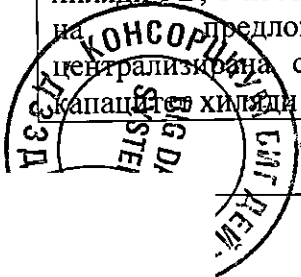
<p>Т.4 Участниците трябва да представят следните материали, свързани със „Система за тестване пълна функционалност на Специализирана централизирана система под ключ“:</p> <ul style="list-style-type: none"> - Документ „Ръководство на администратора за Провеждане тестване работоспособността на отделните МодулDataNode“, в което се представя: <ul style="list-style-type: none"> а) детайлно описание как да се проведе тестването според изискването; б) разпечатка от тестване на поне един МодулDataNode. - Документ „Ръководство на администратора за Провеждане тестване изпълнението на MapReduce задание в Hadoop система, в което се дава: <ul style="list-style-type: none"> детайлно описание как да се проведе 	ДА	<p>Прилагаме следните материали свързани със „Система за тестване пълна функционалност на Специализирана централизирана система под ключ“, както следва:</p> <ul style="list-style-type: none"> - Документ „Ръководство на администратора за Провеждане тестване работоспособността на отделните МодулDataNode“. Документа е предоставен в Т.7 от настоящата документация. - Документ „Ръководство на администратора за Провеждане тестване изпълнението на MapReduce задание в Hadoop система. Документа е предоставен в Т.7 от настоящата документация. 	
---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--

000444



Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависимости
<p>тестването според изискването;</p> <p>б) разпечатка от подобно тестване.</p> <p>- Документ „Ръководство на администратора за Провеждане тестване на комплексната работа на Специализираната централизирана система под ключ Hadoop чрез прилагане на „Hue приложение“ или чрез прилагане на „clean test“, в което се дава:</p> <p>а) детайлно описание как да се проведе тестването според изискването;</p> <p>б) разпечатка от подобно тестване.</p>		<p>- Документ „Ръководство на администратора за Провеждане тестване на комплексната работа на Специализираната централизирана система под ключ Hadoop чрез прилагане на „Hue приложение“ или чрез прилагане на „clean test“. Документа е предоставен в Т.7 от настоящата документация.</p>	
<p>Т. 5 5. Участниците трябва да представят материалът „Концептуалната архитектура за разширение на системата за хиляди РВ“, използвайки компонентите на предложената Специализирана централизирана система под ключ Hadoop, като се направи нейното теоретично разширение до капацитет стотици и хиляди РВ, с което да се докаже разширяемостта на предложената Специализирана централизирана система под ключ Hadoop до капацитет хиляди РВ.</p>	ДА	<p>Документа е представен при Т8 от настоящата документация.</p>	

000445



Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависимости
-----------	-----------------------	---------------------	-------------

Във връзка с провеждане на тестване на предлаганите хардуерни системи:

- i. Участникът да представи писмено предложена детайлно разписана процедура за провеждане на отделните тестове, базирана на добрите практики и препоръки от страна на производителя на предлаганото оборудване.

- Предоставяме писмено предложена детайлно разписана процедура за провеждане на отделните тестове, базирана на добрите практики и препоръки от страна на производителя на предлаганото оборудване в 4 приложения:

- Приложение 1 – Процедура за провеждане на тестове и изпитания свързани с предложения хардуер от SuperMicro.
- Приложение 2 – Производствен протокол за тест на хардуерната система SuperMicro.
- Приложение 3 – Добри практики при конфигурацията на WatchGuard Firebox устройство.
- Приложение 4 – Хардуерни тестове на IPS устройство на WatchGuard

000446



ДОКУМЕНТ

Във връзка с изискване от страна на Възложителя:

„Участникът да представи писмено предложена детайлно разписана процедура за провеждане на отделните тестове, базирана на добрите практики и препоръки от страна на производителя на предлаганото оборудване“, описано на стр. 57 от Документацията за участие в Открита процедура за възлагане на обществена поръчка с предмет:

Доставка, инсталиране, конфигуриране, тестване и поддържане на СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ HADOOP за работа с Големи данни по проект „Дигитализация на икономиката в среда на Големи данни“ (ДИГД), № BG05M2OP001-1.002-0002“

Предоставяме следните Приложения:

1. Приложение 1 - Процедура за провеждане на тестове и изпитания свързани с предложения хардуер от **SuperMicro**.
2. Приложение 2 - Производствен протокол за тест на хардуерната система **SuperMicro**
3. Приложение 3 - Добри практики при конфигурацията на **WatchGuard Firebox** устройство
4. Приложение 4 - Хардуерни тестове на **IPS** устройство на **WatchGuard**



ПРИЛОЖЕНИЕ 1

„Процедура за провеждана на тестове и изпитания свързани с предложени хардуер от SuperMicro“

Описани на стр. 57 от Документацията за участие в Открита процедура за възлагане на обществена поръчка с предмет:

Доставка, инсталиране, конфигуриране, тестване и поддържане на СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ HADOOP за работа с Големи данни по проект „Дигитализация на икономиката в среда на Големи данни“ (ДИГД), № BG05M2OP001-1.002-0002“

Всеки един от нодовете (сървърите) включени в състава на „СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ HADOOP“ се проверява за работоспособност, производителност, съвместимост и надеждност чрез тестова процедура включваща следните пет етапа:

1. Проверка за съвместимост на компонентите на BIOS ниво.
 - Дали се разпознават от системата.
 - Дали работят с честоти и размери указани в спецификациите.
2. Проверка за последни версии на фърмуеър.
 - BIOS
 - IPMI - модула за дистанционен достъп до сървъра
 - RAID контролер
3. Инсталация на операционна система
 - Инсталиране на последни драйвъри
4. Провеждане на тестове за доказване на работоспособност:
 - 4.1 Тестове под Windows server 2019 Standard със следните софтуери:
 - IO Meter - тестове за дисковата система (RAID, JBOD) (файловете .csv разп
 - SiSoft Sandra - тестове за производителност на централен процесор и мем
 - памет (Computer Report.txt)



„Консорциум Биг Дейта Системс“ ДЗ

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

- софтуер за репорт на компонентите на системата (Report.htm)

4.2 Тестове под CentOS 7.5 :

- тест за натоварване на цялата система. Папката съдържа логове от провежданите тестове и резултати от тях
- софтуер за репорт на компонентите на системата (hardinfo_report_linux.html)

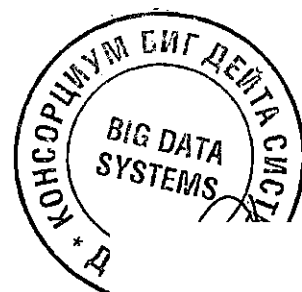
5. Проверка за надеждност на системата:

- Prime95 - максимално и продължително натоварване на централния процесор
- MemtestPRO - максимално и продължително натоварване на оперативната памет
- IO Meter - максимално и продължително натоварване на дисковата система
- SuperDoctor - софтуер за мониторинг на системата

* Снимки от различните етапи са качени в FTP папка Pictures, която е достъпна чрез FTP клиент: ftp://data.persy.com Потребител: tests Парола: Tests123!@#

** Резултати от тестовете са качени на същото FTP.

*** Общото време за изпълнение на тестовата процедура е между 48 и 72 часа за всеки един нод.



„Консорциум Биг Дейта Системс“ ДЗ

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

ПРИЛОЖЕНИЕ 2

“ПРОИЗВОДСТВЕН ПРОТОКОЛ ЗА ТЕСТ НА ХАРДУЕРНАТА СИСТЕМА SUPER MICRO”

Описани на стр. 57 от Документацията за участие в открита процедура за възлагане на обществена поръчка с предмет:

Доставка, инсталиране, конфигуриране, тестване и поддържане на
СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ НАДООР за работа с
Големи данни по проект „Дигитализация на икономиката в среда на Големи
данни“ (ДИГД), № BG05M2OP001-1.002-0002”



Производствен протокол PR79009 / 30.4.2019

Клиент: Светослав Кърпачев

Проформа фактура PR79009 30.4.2019 14:22:3

(Jordan 181377

Създава: svetoslav

Промяна на проформа фактура

Търговец:

Дата на предаване: 30.4.2019

Условия по изпълнението:

Може на място монтаж - NEW Hadoop Projec - SYS-1019P-WTR - Edge Node

1	Case_1065	Server Case Supermicro CSE-113AC2-R706WB2, 750W PSU, Redundant Power Supply 1U, 8 x 2.5" Hot-swap SAS/SATA Drive Bays, Slim DVD-ROM (optional), Rackmount rail kit	C1130I	16p
2	MB_4419	Supermicro X11DDW-L	NM191	16p
3	CPU_0618967	Intel Xeon Bronze 3106 (8 cores) 1.7Ghz, 11M, s3647	786V5	1
4	CPU_0618967	Intel Xeon Bronze 3106 (8 cores) 1.7Ghz, 11M, s3647	786V5	1
5	DDREG4_00001	16GB DDR4-2666 1Rx4 ECC REG DIMM Samsung	CN M393A2K40CB2	86p
6	HDD_037598	Seagate Constellation.2 2000GB ,7200 RPM , 2,5" SATA, 128MB - ST2000NX0403	S4610R9E	1
7	HDD_037598	Seagate Constellation.2 2000GB ,7200 RPM , 2,5" SATA, 128MB - ST2000NX0403	W46115YV	1
8	HDD_037598	Seagate Constellation.2 2000GB ,7200 RPM , 2,5" SATA, 128MB - ST2000NX0403	W46115LF	1
9	HDD_037598	Seagate Constellation.2 2000GB ,7200 RPM , 2,5" SATA, 128MB - ST2000NX0403	W460T7B2	1
10	HDD_037598	Seagate Constellation.2 2000GB ,7200 RPM , 2,5" SATA, 128MB - ST2000NX0403	S4610R8R	1
11	HDD_037598	Seagate Constellation.2 2000GB ,7200 RPM , 2,5" SATA, 128MB - ST2000NX0403	W4605P8T	1
12	HDD_037598	Seagate Constellation.2 2000GB ,7200 RPM , 2,5" SATA, 128MB - ST2000NX0403	W4601D2Z	1
13	HDD_037598	Seagate Constellation.2 2000GB ,7200 RPM , 2,5" SATA, 128MB - ST2000NX0403	W4601CSW	1
14	SCSI/SATA_002	AOC-S3108L-H8IR, 8-port (8 internal) 12Gb/s per port, Supports RAID 0, 1, 5, 6, 10, 50, 60, 2GB 1866MHz DDR3 on-card cache, PCI-E x8	VA183S025807	16p
15	DVD-RW_018	Slim DVD-RW/+RW	712H3QE000604	16p
16	Lan_00241	Supermicro AOC-SGP-I2, Dual port Server LAN Adapter - 10/100/1000Mbps, PCI-Express, Intel i350	VA191S016595	16p
17	Lan_00251	Supermicro AOC-S40G-I2Q, 2-port 40GbE controller, based on Intel Fortville XL710	*	16p
18	FAN-0128	FAN-0101L4 - Supermicro 1U, 40x56mm, (4-pin) PWM, SC819	FAN-0101L4	16p
19	Soft_0128	Supermicro SFT-DCMS-Single, Datacenter Management Package- consists of unified management of server Health monitoring, inventory management, power management, BIOS & firmware upgrades and provisioning, level 2 support	*	16p
20	Cab_269	CBL-SAST-Q556 INTERNAL MINI-SAS HD TO 4-SATA; 90/90/75/75CM	CBL-SAST-Q556	26p
21	battery_053	Supermicro BTR-CV3108-1U1, LSI 3108 CacheVault 1U dummy fan kit	PA188S000544	16p

0067 PSM - 2
531-1
552-1

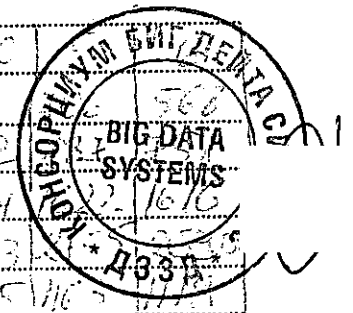
Тестове: Резултати - Desktop системи:

- Memtest
-
- 3D Mark 2003 / 2011
- ViewPerf 10.0
- IO Meter Write Test
- IO Meter Read Test
- Prime95 Time/Temp

Тестове: Резултати - сървърни системи:

- Event Log
- Memtest
-
- RAID level
- IO Meter Write Test
- IO Meter Read Test
- IO Meter Fileserver
- IO Meter Web Server
- IO Meter Database
- Prime 95 Time/Temp

1226/2133



000451

Инсталирана ОС: Windows 2019

BIOS Update BIOS

Report:

Sound Front Panel

Controller Firm ASUS

Controller BBU

IPMI IP address 1.49

IPMI Firm LSI

Проверил за съответствие,
асемблирал: Игорь АИ

Проверил за съответствие,
инсталирал, тествал всички
компоненти и като система: Игорь АИ

Проверил за комплектност,
механични дефекти и опаковал: _____

Забележки и проблеми при тестване:



0004

Производствен протокол PR79008 / 3.5.2019

Клиент: Светослав Кърпачев

Проформа фактура PR79008 2.5.2019 12:05:38

mihail
187405

Създадено: svetoslav

Дата на предаване: 30.4.2019

Промяна на проформа фактура

Търговец:

Дата/час:

Условия по изпълнението:

може на тестов монтаж - NEW 1U OK Hadoop Project - SSG-6019P-ACR12L - Data Node, RAM-a triabva da e - MEM-DR416L-SL04-ER26

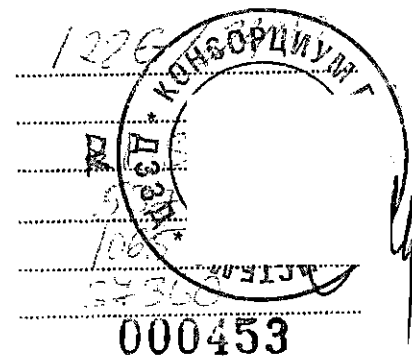
1	Case_20257	Server Case Supermicro CSE-826BE1C-R920LPB, Redundant Power 920W PSU, 2U, 12 x 3.5" SAS/SATA-3 Hot-Swap drive bays, 3 x 80mm cooling fans, Rackmount rail kit	C8260F110N90217	1 бр
2	MB_4098522	Supermicro X11DPI-N		1 бр
3	CPU_0618962	Intel Xeon Silver 4110 (8 cores) 2.1Ghz,11M, s3647		1
4	CPU_0618962	Intel Xeon Silver 4110 (8 cores) 2.1Ghz,11M, s3647		1
5	DDREG4_00000	16GB DDR4-2400 1Rx4 ECC REG DIMM Samsung	CN M393A2K40BB1-	8бр
6	HDD_0375687	Seagate,8TB, Constellation ES.3 , SAS 12Gb/s, 256MB, ST8000NM0075	ZA1EN7R0	1 бр
7	HDD_0375687	Seagate,8TB, Constellation ES.3 , SAS 12Gb/s, 256MB, ST8000NM0075	ZA1ET835	1 бр
8	HDD_0375687	Seagate,8TB, Constellation ES.3 , SAS 12Gb/s, 256MB, ST8000NM0075	ZA1EN7PZ	1 бр
9	HDD_0375687	Seagate,8TB, Constellation ES.3 , SAS 12Gb/s, 256MB, ST8000NM0075	ZA1EN8CY	1 бр
10	HDD_0375687	Seagate,8TB, Constellation ES.3 , SAS 12Gb/s, 256MB, ST8000NM0075	za1en9j5	1 бр
11	HDD_0375687	Seagate,8TB, Constellation ES.3 , SAS 12Gb/s, 256MB, ST8000NM0075	ZA1ENADG	1 бр
12	HDD_0375687	Seagate,8TB, Constellation ES.3 , SAS 12Gb/s, 256MB, ST8000NM0075	ZA1EN99Z	1 бр
13	HDD_0375687	Seagate,8TB, Constellation ES.3 , SAS 12Gb/s, 256MB, ST8000NM0075	Za1en87g	1 бр
14	HDD_0375687	Seagate,8TB, Constellation ES.3 , SAS 12Gb/s, 256MB, ST8000NM0075	ZA1EN8NB	1 бр
15	HDD_0375687	Seagate,8TB, Constellation ES.3 , SAS 12Gb/s, 256MB, ST8000NM0075	Za1en9ds	1 бр
16	HDD_0375687	Seagate,8TB, Constellation ES.3 , SAS 12Gb/s, 256MB, ST8000NM0075	ZA1EN9HG	1 бр
17	HDD_067009	Samsung SSD PM871b 512GB, 2.5 Inch, MZ7LN512HAJQ-00000	S3TVNF0KC03180	1
18	HDD_067009	Samsung SSD PM871b 512GB, 2.5 Inch, MZ7LN512HAJQ-00000	S3TVNF0KC03189	1
19	HDD_0375687	Seagate,8TB, Constellation ES.3 , SAS 12Gb/s, 256MB, ST8000NM0075	ZA1EN9H2	1 бр
20	Lan_00241	Supermicro AOC-SGP-I2, Dual prot Server LAN Adapter - 10/100/1000Mbps, PCI-Express, Intel i350	VA191S016494	1бр
21	Lan_1022	Supermicro AOC-STGN-I2S, Intel® 82599ES, PCI Express 2.0, Dual SFP + Connectors	VA18CS021606	1бр
22	FAN-011782	SNK-P0068PS, 2U Passive CPU HS for X11 Purley	SNK-P0068PS	2 бр
23	HBA-038	LSI LSI00344 9300-8i SGL, SAS 3, 8 port HBA, 12Gb/s, JBOD, PCI-Express 3.0 x8 LP	SP83306918	1бр
24	Cab_282	CBL-SAST-0532 - Internal MiniSAS HD SFF-8643 50cm Cable	CBL-SAST-0532	1 бр

Тестове: Резултати - Desktop системи:

- Memtest
- 3D Mark 2003 / 2011
- ViewPerf 10.0
- IO Meter Write Test
- IO Meter Read Test

Тестове: Резултати - сървърни системи:

- Event Log
- Memtest
- RAID level
- IO Meter Write Test
- IO Meter Read Test
- IO Meter Fileserver



Prime95 Time/Temp

IO Meter Web Server

IO Meter Database

Prime 95 Time/Temp

47952
35159
65h, 60/65°

Инсталирана ОС:

Win 2003 Std / Kubuntu

Controller Firm

16.00.00.00

BIOS Update

2.0c

Controller BBU

Report:

IPMI IP address

1.15

Sound

Front Panel

IPMI Firm

1.67

Проверил за съответствие,
асемблирал:

ИОРГАН

Проверил за съответствие,
инсталирал, тествал всички
компоненти и като система:

ИОРГАН

Проверил за комплектност,
механични дефекти и опаковал:

Забележки и проблеми при тестване:

MBD Results JBOB

№	1	2	3	4	5	6	7	8	9	10	11	12
W	247	241	258	245	239	250	239	239	236	240	238	239
R	242	242	241	242	242	252	241	242	241	241	241	240
FS	621	614	625	614	617	627	628	620	629	633	615	614
WS	641	650	635	644	656	660	661	659	642	662	650	651
OF	471	460	471	473	476	474	474	470	474	472	476	476



000454

ПРИЛОЖЕНИЕ 3

„Добри практики при конфигурацията на WatchGuard Firebox устройство“

Описани на **стр. 57** от Документацията за участие в Открита процедура за възлагане на обществена поръчка с предмет:

Доставка, инсталиране, конфигуриране, тестване и поддържане на СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ НАDOOP за работа с Големи данни по проект „Дигитализация на икономиката в среда на Големи данни“ (ДИГД), № BG05M2OP001-1.002-0002“

Препоръки от производителя WatchGuard

За да защитите вътрешните си мрежи, вашият Firebox отхвърля всички пакети, които не са специално разрешени от политиките на защитната стена. Всяка политика на защитната стена определя набор от правила, които казват на Firebox да разрешава или отхвърля трафик въз основа на фактори като източник и дестинация на пакета или TCP / IP порта или протокола, използван за пакета.

Препоръчителна конфигурация по подразбиране

Когато стартирате съветника за веб настройка или съветника за бърза настройка, за да настроите Firebox, съветникът за настройка автоматично конфигурира политиките за защитната стена по подразбиране и активира лицензирани услуги за сигурност с препоръчани настройки.

Ако новия ви Firebox е с версия на Fireware по-ниска от v12.0.1, се препоръчва да се направи ъпгрейд на Fireware, да се нулира Firebox и отново да се стартира съветника за настройка, за да се активират автоматично препоръчителните политики по подразбиране и услугите за сигурност.

Съветниците за настройка добавят следните политики по подразбиране:

- FTP-proxy, с Default-FTP-Client прокси действие(proxy action)

000455

„Консорциум Биг Дейта Системс“ ДЗЗД

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

- HTTP-proxy, с Default-HTTP-Client прокси действие(proxy action)
- HTTPs-proxy, с Default-HTTPs-Client прокси действие(proxy action)
- WatchGuard Certificate Portal (за версия на Firewall v17.3 или по-голяма)
- WatchGuard Web UI
- Ping
- DNS
- WatchGuard
- Outgoing

С тези политики по подразбиране, Firebox:

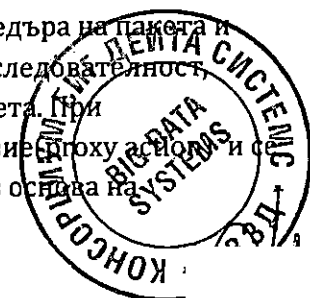
- Не позволява връзки от външната мрежа към trusted или optional мрежи или към Firebox
- Позволява мениджмънт връзки към Firebox само от trusted и optional мрежи
- Проверява изходящия FTP, HTTP и HTTPS трафик с препоръчани настройки за прокси действие(proxy action)
- Използва услугите за сигурност Application Control, WebBlocker, Gateway AntiVirus, Intrusion Prevention, Reputation Enabled Defense, Botnet Detection, Geolocation и APT Blocker APT Blocker за защита на trusted и optional мрежи
- Позволява изходящи FTP, Ping, DNS, TCP и UDP връзки от trusted и optional мрежи

Видове/Типове политики

Има два вида политики на защитната стена. Всеки разглежда връзките на различно ниво на детайлност.

Packet Filter Policy (Политика за филтриране на пакети) - проверява само хедъра на пакета. Пакетният филтър изследва IP и TCP / UDP хедъра на всеки пакет, но не инспектира съдържанието на пакета.

Proxy policy or application layer gateways (ALG) - проверява хедъра на пакета и съдържанието му. Прокси политиката или ALG отваря всеки пакет в последователност, премахва хедъра на мрежовия слой и инспектира съдържанието на пакета. При конфигуриране на прокси политика или ALG, се избира прокси действие(proxy action) и се конфигурират правила и действия, които трябва да се предприемат въз основа на характеристиките на съдържанието.



Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

Конфигурацията по подразбиране използва прокси политики, така че Firebox да може да проверява съдържанието и да използва лицензираните услуги за защита, за да защити вашата мрежа.

Свойства на политиките

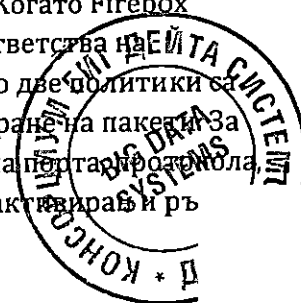
- Портове и протоколи - Политиката се определя от портовете и протоколите, на които слуша. Когато се добавя политика, избраният шаблон (темплейт) на съответната политика определя за кои портове и протоколи се прилага политиката. Портовете и протоколите са единствените настройки, които не можете да редактирате в една политика.
- Източник и дестинация – Политика се прилага само за връзки, които съответстват както на източника (From), така и на дестинацията (Do) в политиката. Източник или дестинация могат да включват псевдоними, потребители, групи, IP адреси и FQDN (Fully Qualified Domain Names).
- Диспозиция (Действие на защитната стена) - Диспозицията определя дали политиката позволява или отказва връзки, които съответстват на правилата в политиката. Можете да конфигурирате политика да разрешава или отказва връзки.

Приоритет на политиките

Приоритет е последователността, на база на която Firebox изследва мрежовия трафик и прилага правило на политика. По подразбиране политиките на Firebox се конфигурират в режим Auto-Order (автоматичен ред). В режим Auto-Order Firebox автоматично сортира политиките от най-частните към най-общите на базата на сравнение на тези свойства на политиките.

- Портове и протоколи
- Източник и дестинация
- Диспозиция
- Програма

Политиките, които са по-нагоре в списъка имат по-висок приоритет. Когато Firebox получи пакет, той прилага политиката с най-висок приоритет, която съответства на характеристиките на пакета. Когато режимът Auto-Order е активиран, ако две политики са еднакви, прокси политиката има предимство пред политиката за филтриране на пакети. За пакет се прилага политиката с най-висок приоритет, която съответства на порт, протокол, източника и дестинацията. Режимът Auto-Order можете също да бъде деактивиран и ръчно да се промени реда на политиките.



„Консорциум Биг Дейта Системс

Хай Тех Бизнес
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

Препоръка: Използвайте режима Auto-Order, докато не установите, че не работи за вашата конкретна ситуация. Този режим работи добре в повечето случаи.

Може да се добавят множество политики за един и същи порт / протокол с различни източници или дестинации, за да се разрешат различни нива на достъп до мрежови ресурси за различни потребители, групи или мрежи. Например, може да се конфигурира HTTP-проху политика за конкретен отдел, за да се разреши по-ограничен или по-широк достъп до ресурси, отколкото стандартната HTTP-проху политика с по-нисък приоритет.

Добри практики за конфигуриране на политики

Персонализиране на имена на политиките

Конфигурацията на Firebox по подразбиране използва стандартни имена за политики. Имената на политиките по подразбиране указват типа на трафика, който се обработва от политиката, но може да не е особено съдържателен в мрежовата ви среда, особено ако добавяте множество правила от един и същи тип.

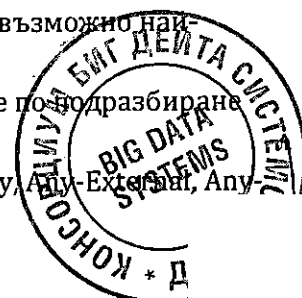
Препоръка: За да направите политиките си по-лесни за разбиране и поддържане, дайте на всяка политика смислено име, което показва целта на политиката, към кои потребители или мрежа се отнася, или всякакви други уникални характеристики, като например, когато политиката е активна.

Ограничение обхвата на политиките

По подразбиране политиките на защитната стена използват вградените псевдоними Any-External, Any-Trusted, Any-Optional и Any. Политиките, конфигурирани с тези псевдоними, могат да позволяват връзки от повече източници и дестинации, отколкото е необходимо. За по-голям контрол върху мрежовите връзки разгледайте източника и дестинацията във всяка политика, за да се уверите, че политиката не позволява връзки от повече източници или до повече дестинации, отколкото е необходимо.

Препоръки:

- За всяка политика конфигурирайте източника и дестинацията възможно най-подробно.
- Не използвайте псевдоним Any в политики (освен в политиките по подразбиране Ping, BOVPN-Allow.in и BOVPN-Allow.out)
- Прегледайте всички политики, които включват псевдоними Any, Any-External, Any-Optional, or Any-Trusted.



„Консорциум Биг Дейта Системс“

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, POB 74
Tel.: +359 2,42 10 10

- Уверете се, че псевдонимите във всяка политика са наистина необходими връзките, които искате да разрешите.
- Ако е възможно, заменете тези псевдоними с по-специфичен източник дестинация, за да стесните обхвата на политиката.

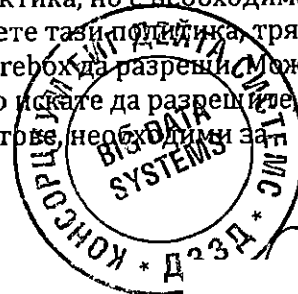
Ето няколко примера за това, кога да се ограничи обхватът на политиката на някои автоматично генерирани политики:

- a) Редактирайте дестинацията в стандартната DNS политика - политиката по подразбиране за DNS позволява връзки от Any-Trusted и Any-Optional към Any-External.
Препоръка: За да стесните обхвата на тази политика, можете да промените дестинацията, така че да включва само IP адресите или FQDN на външните DNS сървъри в настройките ви на DNS.
- b) Редактирайте източника в политиката за мениджмънт на Firebox - политиките по подразбиране WatchGuard и WatchGuard Web UI позволяват мениджмънт връзки от Any-Trusted и Any-Optional.
- c) Препоръка: Можете да премахнете Any-Optional от тези политики, за да предотвратите възможност за мениджмънт достъп от мрежите ви които са Optional. За да намалите още повече обхвата, премахнете Any-Trusted и добавете специфични подмрежи или IP адрес, от който искате Firebox да приема мениджмънт връзки.

Деактивиране на политика Outgoing

○ Конфигурацията на Firebox по подразбиране включва политиката за филтриране на пакети Outgoing. Политиката Outgoing позволява всички TCP и UDP връзки от всеки trusted или optional източник във вашата мрежа към всяка външна мрежа. Тъй като това е политика за филтриране на пакети, а не прокси политика, политиката Outgoing не филтрира съдържанието, когато изследва трафика през вашия Firebox. Политиката Outgoing съществува, за да се гарантира, че Firebox позволява изходящи TCP и UDP връзки, които не съвпадат с никоя друга политика.

Препоръка: За да се уверите, че Firebox позволява само връзки, които искате да разрешите, се препоръчва да забраните политиката Outgoing. Това е добра практика, но е необходимо внимателно планиране и последващи действия. Преди да премахнете тази политика, трябва да добавите политики за всички изходящи връзки, които искате Firebox да разреши. Можете да добавите отделна политика за всеки тип изходящ трафик, който искате да разрешите, или да създадете персонализиран пакетен филтър за конкретните портове, необходими за изходящите връзки, които искате да разрешите от вашата мрежа.



Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
ТГ-1 · 359 2 42 10 10

ПРИЛОЖЕНИЕ 4

„Процедура за провеждана на тестове и изпитания свързани с предложения хардуер от WatchGuard“

Описани на **стр. 57** от Документацията за участие в Открита процедура за възлагане
на обществена поръчка с предмет:

Доставка, инсталиране, конфигуриране, тестване и поддържане на
СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ НАDOOP за работа с
Големи данни по проект „Дигитализация на икономиката в среда на Големи
данни“ (ДИГД), № BG05M2OP001-1.002-0002“

1. Тест на системата

```
--Result of system running information.
--
CPU: 100%
MEM: 100%
DISK: 100%
NET: 100%
...

```



Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

2. Тест на оперативна памет

```
--Result of Memory test
--
memtester version 4.1.2 (64-bit)
Copyrights (C) 2009 Charles Cazabon.
Licensed under the GNU General Public license version 2 (only).

pagesize is 4096
pagesizemask is 0xffffffff000
want 75MB (78807040 bytes)
got 75MB (78807040 bytes), trying mlock ...locked.
Loop 1 1:
  Stack Address      : ok
  Random Value       : ok
  Compare XOR        : ok
  Compare SUB        : ok
  Compare MUL        : ok
  Compare DIV        : ok
  Compare OR         : ok
  Compare AND        : ok
  Sequential Increment: ok
  Solid Bits         : ok
  Block Sequential   : ok
  Checkerboard       : ok
  Bit Spread         : ok
  Bit Flip           : ok
  Walking Ones       : ok
  Walking Zeros      : ok

Done.
```



Хай Тех Бизнес Ц
ж.к. Дружба 1, ул. Мюн
1592 София, РОВ 74
Tel.: +359 2 42 10 10

3. Тест на флаш памет

```
W:\>msdskdiag hardware flash boot
It may take a long time to do the flash check,
Are you sure you want to do the flash check? yes or no? Yes
Doing flash check, please wait...
--
--Result of flash check
--
Checking blocks 0 to 1023
Checking for bad blocks read-only test done
Base completed, 0 bad blocks found.

W:\>msdskdiag hardware flash test
It may take a long time to do the flash check,
Are you sure you want to do the flash check? yes or no? Yes
Doing flash check, please wait...
--
--Result of flash check
--
Checking blocks 0 to 1023
Checking for bad blocks read-only test done
Base completed, 0 bad blocks found.

W:\>msdskdiag hardware flash test
It may take a long time to do the flash check,
Are you sure you want to do the flash check? yes or no? Yes
Doing flash check, please wait...
--
--Result of flash check
--
Checking blocks 0 to 1023
Checking for bad blocks read-only test done
Base completed, 0 bad blocks found.
```

4. Тест на мрежов интерфейс

```
--Result of ethernet interface diagnostic
--
Settings for eth0:
Supported port speeds: 1 TP
Supported link modes: 100baseT Half 100baseT Full
100baseT Half 100baseT Full
100baseT Full
Supported pause frame use: Symmetric
Suppress auto-negotiation: Yes
Advertised link modes: 100baseT Full
Advertised pause frame use: Symmetric
Advertised auto-negotiation: Yes
Speed: 100Mbps
Duplex: Full
Port: Twisted Pair
PHYAD: 1
Transceiver: internal
Auto-negotiation on
MII-VR: ok
Supports Wake-on: gmbp
Wake-on: g
Current message level: 0x00000007 (7)
drv probe link
Link detected: yes
```



Изискване	Изпълнение (Да/Не)	Начин на реализация	Зависимости
<p>ii. Участникът да представи разпечатка на екрани от провеждане на подобно конфигуриране и тестове на специализирани системи, свързани с:</p> <ul style="list-style-type: none"> - Конфигуриране на VLAN на комутаторите (switches); - Обмен на данни между два изчислителни модула с използване на конфигурираната VLAN; 		<p>Предоставяме разпечатка на екрани от провеждане на подобно конфигуриране и тестове на специализирани системи, свързани с:</p> <ul style="list-style-type: none"> - Конфигуриране на VLAN на комутаторите (switches); - Обмен на данни между два изчислителни модула с използване на конфигурираната VLAN; 	

000463



„Консорциум Биг Дейта Системс“ ДЗЗ

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

ДОКУМЕНТ

Във връзка с изискване от страна на Възложителя,

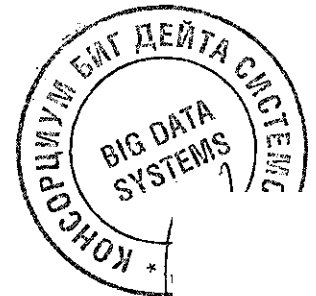
„Участникът да представи разпечатка на екрани от провеждане на подобно конфигуриране и тестове на специализирани системи, свързани с:

- Конфигуриране на VLAN на комутаторите (switches);
- Обмен на данни между два изчислителни модула с използване на конфигурираната VLAN;

описани на стр. 57 от Документацията за участие в Открита процедура за възлагане на обществена поръчка с предмет:

Доставка, инсталиране, конфигуриране, тестване и поддържане на СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ HADOOP за работа с Големи данни по проект „Дигитализация на икономиката в среда на Големи данни“ (ДИГД), № BG05M2OP001-1.002-0002“

1. Конфигуриране на VLAN на комутаторор (switches):



„Консорциум Биг Дейта Системс“ ДЗ:

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

2. Обмен на данни между два изчислителни модула с използване на конфигурираната VLAN;

No.	Time	Source	Destination	Protocol	Length	Info
297	8.166207	192.168.1.123	192.168.1.1	ICMP	74	Echo (ping) request 16=00001, seq=163/42240, ttl=64 (reply in 298)
298	8.166493	192.168.1.1	192.168.1.123	ICMP	74	Echo (ping) reply 16=00001, seq=163/42240, ttl=64 (request in 297)
299	8.166596	192.168.1.123	192.168.1.1	ICMP	74	Echo (ping) request 16=00001, seq=164/42240, ttl=64 (reply in 300)
300	8.166895	192.168.1.1	192.168.1.123	ICMP	74	Echo (ping) reply 16=00001, seq=164/42240, ttl=64 (request in 299)
467	18.371879	192.168.1.123	192.168.1.1	ICMP	74	Echo (ping) request 16=00001, seq=167/42720, ttl=64 (reply in 468)
468	18.372180	192.168.1.1	192.168.1.123	ICMP	74	Echo (ping) reply 16=00001, seq=167/42720, ttl=64 (request in 467)
413	11.374790	192.168.1.123	192.168.1.1	ICMP	74	Echo (ping) request 16=00001, seq=168/43200, ttl=64 (reply in 414)
414	11.375017	192.168.1.1	192.168.1.123	ICMP	74	Echo (ping) reply 16=00001, seq=168/43200, ttl=64 (request in 413)
1277	14.152968	192.168.1.123	192.168.1.1	ICMP	74	Echo (ping) request 16=00001, seq=169/43254, ttl=64 (reply in 1278)
1278	14.156280	192.168.1.1	192.168.1.123	ICMP	74	Echo (ping) reply 16=00001, seq=169/43254, ttl=64 (request in 1277)
1343	13.164128	192.168.1.123	192.168.1.1	ICMP	74	Echo (ping) request 16=00001, seq=170/43258, ttl=64 (reply in 1344)
1344	13.164391	192.168.1.1	192.168.1.123	ICMP	74	Echo (ping) reply 16=00001, seq=170/43258, ttl=64 (request in 1343)
1405	16.472944	192.168.1.123	192.168.1.1	ICMP	74	Echo (ping) request 16=00001, seq=171/43716, ttl=64 (reply in 1406)
1406	16.473204	192.168.1.1	192.168.1.123	ICMP	74	Echo (ping) reply 16=00001, seq=171/43716, ttl=64 (request in 1405)
1754	37.174224	192.168.1.123	192.168.1.1	ICMP	74	Echo (ping) request 16=00001, seq=172/44012, ttl=64 (reply in 1755)
1755	37.174586	192.168.1.1	192.168.1.123	ICMP	74	Echo (ping) reply 16=00001, seq=172/44012, ttl=64 (request in 1754)



000400

ДОКУМЕНТ

ОПИСАНИЕ НА СЪДЪРЖАНИЕТО НА ВЪНШНИЯ ХАРД ДИСК (F---

Като Външния хард диск се явява неделима част от настоящото техническо предложение

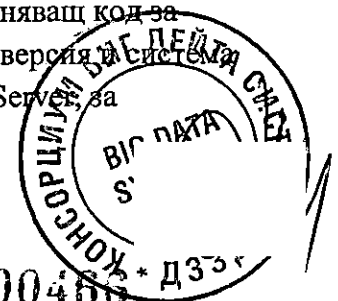
Съгласно изискванията от Възложителя относно открита процедура с предмет:

„Доставка, инсталиране, конфигуриране, тестване и поддържане на СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ HADOOP за работа с Големи данни по проект „Дигитализация на икономиката в среда на Големи данни“ (ДИГД), № BG05M2OP001-1.002-0002“

Съдържание на външния хард диск:

1. Във връзка с Т.6.1, на стр. 43 от Документацията, предоставяме изпълняващ код (binary code) и конфигурационни файлове за интегриране между данните на двете системи: MS SQL Server минимална версия 2017 (във вид на релационни таблици) и Специализирана централизирана система под ключ Cloudera/Hadoop (във вид на HDFS файлове).
2. Във връзка с Т.6.2, на стр. 44 от Документацията предоставяме софтуер в първичен код (source code), в изпълняващ код (binary code) и конфигурационни файлове за прилагане на технологията за машинно обучение (Machine Learning – ML) на MS SQL Server 2017 или следваща версия, с използване на данни от Специализирана централизирана система под ключ Cloudera / Hadoop система и от MS SQL Server
3. Във връзка с Т.1.2, на стр. 54 от Документацията, прилагаме изпълняващ код за Интегриране на двете системи: MS SQL Server 2017 или следваща версия на система Cloudera/Hadoop“, прилагайки PolyBase технологията на MS SQL Server, за четирите типа интегриране:

а) експортиране и импортиране на данни от едната система в другата;



„Консорциум Биг Дейта Системс“ ДЗЗД

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

б) четене, модифициране и изтриване на данни от едната система в другата;

в) данните в системата MS SQL Server 2017 или следваща версия да бъдат в таблици създадени за работа в паметта (memory-optimized tables in the main memory);

г) конфигуриране Аналитичните услуги (Analytical Services) на MS SQL Server 2017 или следваща версия да работят с данни от система Cloudera/Hadoop (HDFS файлове).

4. Във връзка с Т.1.1, на стр. 53 от Документацията, прилагаме изпълняващ код за прилагане на технологията за машинно обучение (Machine Learning – ML) на MS SQL Server 2017 или следваща версия, с използване на данни от система Cloudera / Hadoop и от MS SQL Server

5. Във връзка с 2.1, стр.55 от Документацията, предоставяме Първичен код и Конфигурационни файлове, а Документ „Ръководство за прилагане на технологията за машинно обучение (Machine Learning – ML) на MS SQL Server 2017 или следваща версия, с използване на данни от система Cloudera / Hadoop и от MS SQL Server“, е приложен в документацията.

6. Във връзка с 2.2 от Документацията, предоставяме изпълняващ код за прилагане на технологията за машинно обучение (Machine Learning – ML) на MS SQL Server 2017 или следваща версия, с използване на данни от система Cloudera / Hadoop и от MS SQL Server;



НАЦИОНАЛНА АГЕНЦИЯ ЗА ПРИХОДИТЕ
ЕИК по БУЛСТАТ 131063188

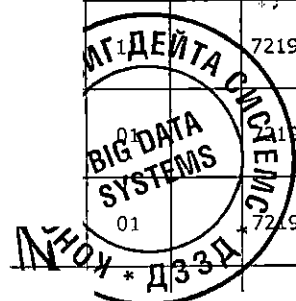
СПРАВКА ЗА АКТУАЛНО СЪСТОЯНИЕ НА ДЕЙСТВАЩИТЕ ТРУДОВИ ДОГОВОРИ
КЪМ ДАТА 30.04.2019 г.

ЕГН/ЛНЧ/Сл.номер/БУЛСТАТ: 200149195

Име: СЕКЮРИТИ СОЛЮШЪНС ИНСТИТУТ /ЕООД/
гр.СОФИЯ, община СТОЛИЧНА, област СОФИЯ, ул.
Мюнхен №8,

Дата и час на отпечатване: 30.04.2019 10:26:48

№	ЕГН/ЛНЧ/ Сл.номер	Име	Дата на сключване	Дата на посл. доп. спор./Промяна на работното място	Дата на прекр.	Осно- вание	Срок	Код КИД	Код НКПД	Длъжност наименование
1						.1		7219		ЕКСПЕРТ ИНФОРМАЦИОННИ ТЕХНОЛОГИИ
2						.1		7219		ЕКСПЕРТ ИНФОРМАЦИОННИ ТЕХНОЛОГИИ
3						.1		7219		СОФТУЕРЕН ИНЖЕНЕР
4						01		7219		МЕНИДЖЪР РАЗВОЙНА ДЕЙНОСТ
5)1		7219		РАЗРАБОТЧИК СОФТУЕР
6						01		7219		СПЕЦИАЛИСТ РЕНД МАЙКРО
7						.1		7219		СТАРШИ СИСТЕМЕН АНАЛИЗАТОР
8						.1		7219		СПЕЦИАЛИСТ ПРОДАЖБИ
9						.1		7219		СПЕЦИАЛИСТ ПРОДАЖБИ
10						.1		7219		Програмист софтуерни приложения
11						01		7219		СИСТЕМЕН АДМИНИСТРАТОР ИС
12						11		7219		ЕКСПЕРТ ИНФОРМАЦИОННИ ТЕХНОЛОГИИ
13						.1		7219		ЕКСПЕРТ ИНФОРМАЦИОННИ ТЕХНОЛОГИИ
14						11		7219		СПЕЦИАЛИСТ ПРОДАЖБИ
15						11		7219		ЕКСПЕРТ ИНФОРМАЦИОННИ ТЕХНОЛОГИИ
16						01		7219		МЛАДШИ ПРОГРАМИСТ
17						01		7219		МЛАДШИ ПРОГРАМИСТ



000468

18	7219	МЛАДШИ СПЕЦИАЛИСТ ИНФОРМАЦИОННА СИГУРНОСТ
<p>* - указва дата на сключване на договор по чл.123, ал.1 от КТ ** - служебно прекратяване при прекратена регистрация на работодателя или на работника/служителя *** - прекратяване на основание чл.327, ал.2 от КТ</p>		
<p>Легенда за поле "Основание": 01 - безсрочен трудов договор по чл. 67, ал. 1, т. 1 КТ 11 - допълнителен трудов договор по чл. 111 КТ</p>		

Вярно с оригинала



000469

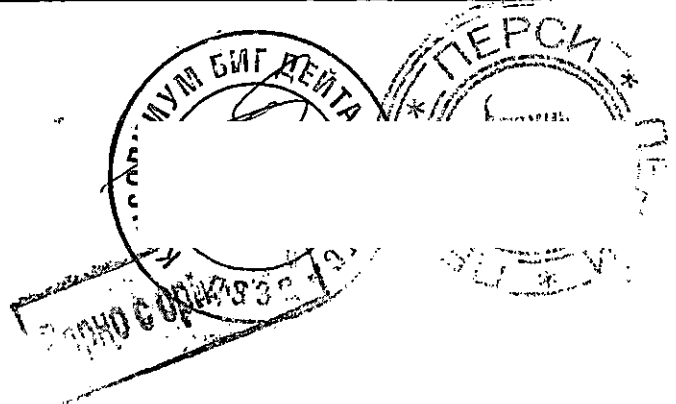
НАЦИОНАЛНА АГЕНЦИЯ ЗА ПРИХОДИТЕ БИК по БУЛСТАТ 131063183												
СПРАВКА ЗА АКТУАЛНО СЪСТОЯНИЕ НА ДЕЙСТВАЩИТЕ ТРУДОВИ ДОГОВОРИ КЪМ ДАТА 30.04.2019 г.												
ЕГН/ЛНЧ/Сл.номер/БУЛСТАТ: 030413948												
Име: ПЕРСИ /ООД/												
Адрес: гр.СОФИЯ, община СТОЛИЧНА, област СОФИЯ, бул. "Цар Борис III" №1366,												
Дата и час на отпечатване: 30.04.2019 12:53:54												
№	ЕГН/ЛНЧ/ Сл.номер	Име	Дата на сключване	Дата на посл. доп. спор./Промяна на работното място	Дата на прекр.	Осно- вание	Срок	Код КИД	Код НКПД	Длъжност наименование	Код по ЕКАТТЕ	Основна заплата
1								2620		СНАБДИТЕЛ ДОСТАВЧИК	68134	
2								2620		УПРАВИТЕЛ	68134	
3								2620		ИНЖЕНЕР ЕЛЕКТРОННО- ТЕХНИЧЕСКИ АРХИВ	68134	
4								7260		снабдители		
5								7260	4193	технически сътрудник		
6								2620		ОРГАНИЗАТОР ЕКСПЕДИЦИЯ ТОВАРО- РАЗТОВАРНА И СЛЕДИТОРСКА ДЕЙН	68134	
7								2620		РЪКОВОДИТЕЛ ОТДЕЛ В БИЗНЕС УСЛУГИТЕ	68134	
8								2620		ЕКСПЕРТ ФИНАНСОВИ И СТОПАНСКИ АНАЛИЗИ	68134	
9								262	13215020	ДИРЕКТОР ПРОИЗВОДСТВО	68134	
10								262	43212021	НАЧАЛНИК СКЛАД	68134	
11								262		АНАЛИЗАТОР ПРОУЧВАНЕ НА ПАЗАРИ	68134	
12								2620	13215021	ТЕХНИЧЕСКИ ДИРЕКТОР	68134	
13								2620	26296002	ЕКСПЕРТ ИНФОРМАЦИОННО ОСИГУРЯВАНЕ	68134	
14								2620	24216001	АНАЛИЗАТОР ПРОУЧВАНЕ НА ПАЗАРИ	68134	
15								2620	21526001	ИНЖЕНЕР КАРДЕР	68134	
16								2620	13216008	РЪКОВОДИТЕЛ ОТДЕЛ ПО ПРОДАЖБИТЕ	68134	
17								2408	28136001	ЕКСПЕРТ ПРОЕКТИРВАНЕ КОМПЮТЪРНИ СИСТЕМИ	68134	
18								2620	21526001	ИНЖЕНЕР КАРДЕР	68134	
19								2620	28296001	ИНЖЕНЕР КАРДЕР	68134	
20								2620	21526001	ИНЖЕНЕР КАРДЕР	68134	
21								2620	28136001	ЕКСПЕРТ ПРОЕКТИРВАНЕ КОМПЮТЪРНИ СИСТЕМИ	68134	

000470

		КОМП. СЪМ МРЕЖИ	
22	2620	ТЕХНИК КОМПЮТЪРНИ СИСТЕМИ	68134
23	2620	ТЕХНИК КОМПЮТЪРНИ СИСТЕМИ	68134

* - излиза дата на сключване на договор по чл.123, ал.1 от КТ
 ** - служебно прекратяване при поскратена регистрация на работодателя или на работника съгласно
 *** - прекратяване на основание чл.327, ал.2 от КТ

Легенда за поле "Основание"
 01 - безсрочен трудов договор по чл. 67, ал. 1, т. 1 КТ



000471

„Консорциум Биг Дейта Системс“ ДЗЗД

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

ДОКУМЕНТ:

**Доказателство за наличие на сервизна база за осъществяване на
сервизната услуга**

Съгласно стр. 68 Т. III.3. КРИТЕРИИ ЗА ПОДБОР

на открита процедура с предмет:

„Доставка, инсталиране, конфигуриране, тестване и поддържане на СПЕЦИАЛИЗИРАНА
ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ НАDOOP за работа с Големи данни по
проект „Дигитализация на икономиката в среда на Големи данни“ (ДИГД), №
BG05M2OP001-1.002-0002“

Във връзка с това изискване предоставяме следните документи като доказателство:

1. Удостоверение N:749/05.10.2018 г. за въвеждане в експлоатация на строеж, с възложител „Перси“ ООД, издадено от Столична община, направление „Архитектура и градоустройство“.
2. Снимков материал – 4 броя, доказващи реалното наличие на сервизна база.



000472



СТОЛИЧНА ОБЩИНА

Направление „Архитектура и градоустройство“

Главен архитект на Столична община

ул. „Сердика“ № 5, тел. 02/92 38 310, факс: 980 67 41

САГ18-ТК00-2063/2018 г.

На основание чл. 177, ал. 3 от ЗУТ, искане на възложителя вх. № САГ18-ТК00-2063- [2]/27.09.2018 г., придружено от окончателен доклад № 1824 от 29.08.2018 г., изготвен от „АРКА КОНСУЛТ“ ЕООД с ЕИК 131317505, представлявано от управителя Явор Милчев Стоянчев, упражняващо строителен надзор, съгласно Удостоверение № РК-0375/20.01.2015 г., издадено от ДНСК и Заповед № РА50-395/14.06.2017 г. на Главния архитект на Столична община, допълнена със Заповед №РА50-367/09.05.2018г., издавам:

УДОСТОВЕРЕНИЕ

№ 449 / 05.10.2018г. /

ЗА ВЪВЕЖДАНЕ В ЕКСПЛОАТАЦИЯ НА СТРОЕЖ:

„Преустройство и смяна на предназначението на магазин, аптека, кафе-клуб и гаражи №41, №52 и №54 и паркомясто №53 в Офис, кафе-клуб, гаражи №41, №52 и №54 и паркомясто №53“

находящ се в УПИ Х-351,352, кв. 2А, м. „НПЗ Средец“ по плана на гр. София, район „Красна поляна“-СО, с административен адрес: гр. София, ж.к. „Красна поляна“, ул. „Златна Добруджа“ №18, с възложител: „Перси“ ООД, ЕИК 030413948.

При спазване на следните ограничителни условия за ползване на строежа:

Без ограничения!

Строежът е изпълнен в съответствие с одобрените от район „Красна поляна“-СО на 04.03.2013 г. инвестиционни проекти към Разрешение за строеж № 05/11.03.2013 г., издадено от Главния архитект на район „Красна поляна“-СО, Заповед № РКП17-РД56-153/16.06.2017 г., издадена от Главния архитект на район „Красна поляна“-СО за допълване на Разрешението за строеж, във връзка с изменение на инвестиционния проект Заповед №РКП17-РД56-40/03.02.2017г. на Главния архитект на район „Красна поляна“-СО за поправка на допуснатата непълнота и очевидна фактическа грешка и изискванията към строежите, съгласно чл. 169 от ЗУТ.

Строежът е заснет и нанесен в кадастралната карта на град София, идентификатор № 68134.1107.219.1.60, съгласно Удостоверение № 25-100580/20.11.2017 г. на геодезия, картография и кадастр-София.

Гаранционните срокове за изпълнените строителни и монтажни работи са съгласно договора между възложителя и изпълнителя, но не по-малко от минималните срокове, определени с наредбата по чл. 160, ал. 3 от ЗУТ и започват да текат от деня на издаване на удостоверение за въвеждане в експлоатация на строежа.

При ползването на обекта да се спазват законовите разпоредби.

ГЛАВЕН АРХИТЕКТ
НА СТОЛИЧНА ОБЩИНА:

Вярно с оригинала



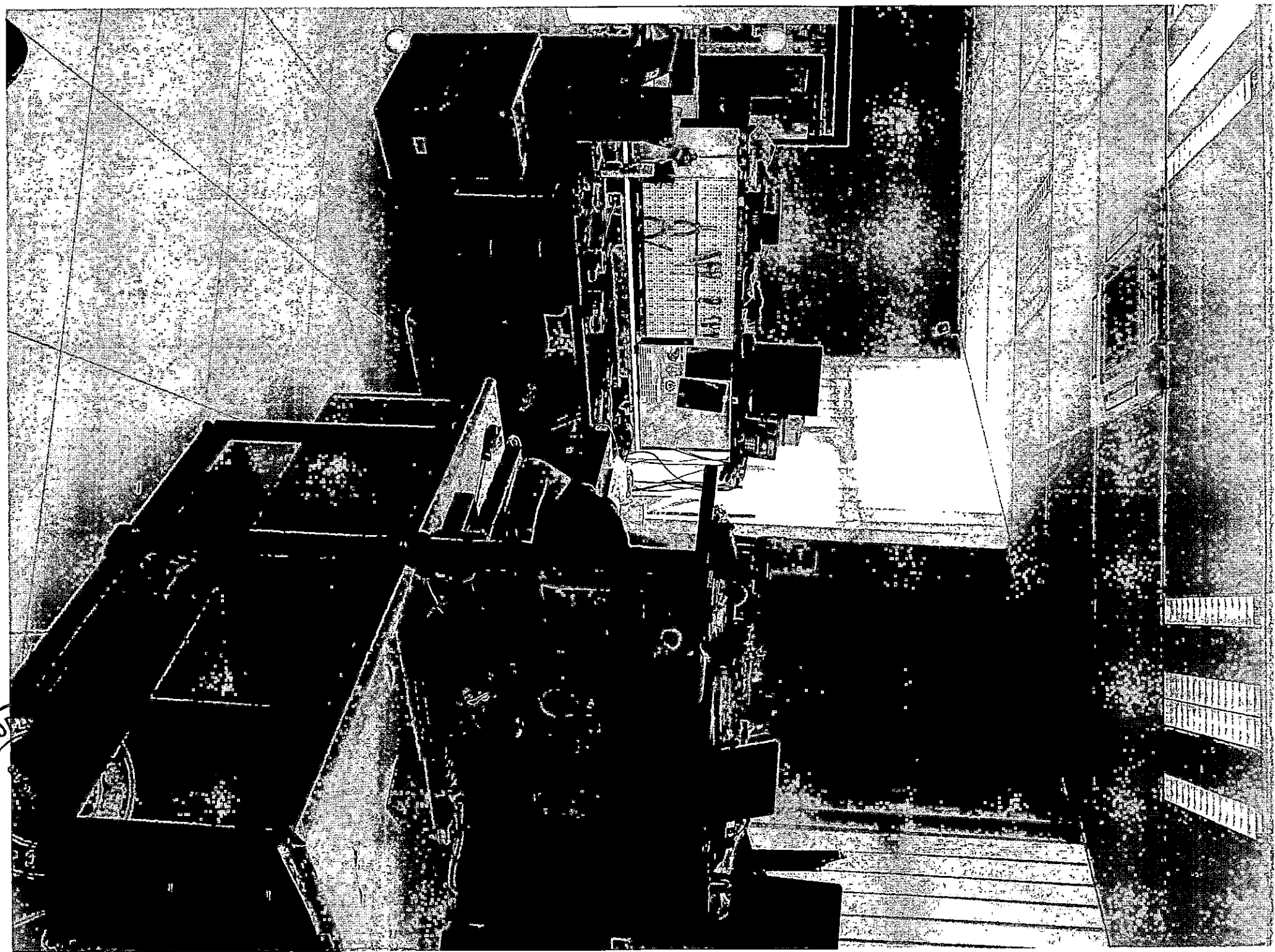
000472

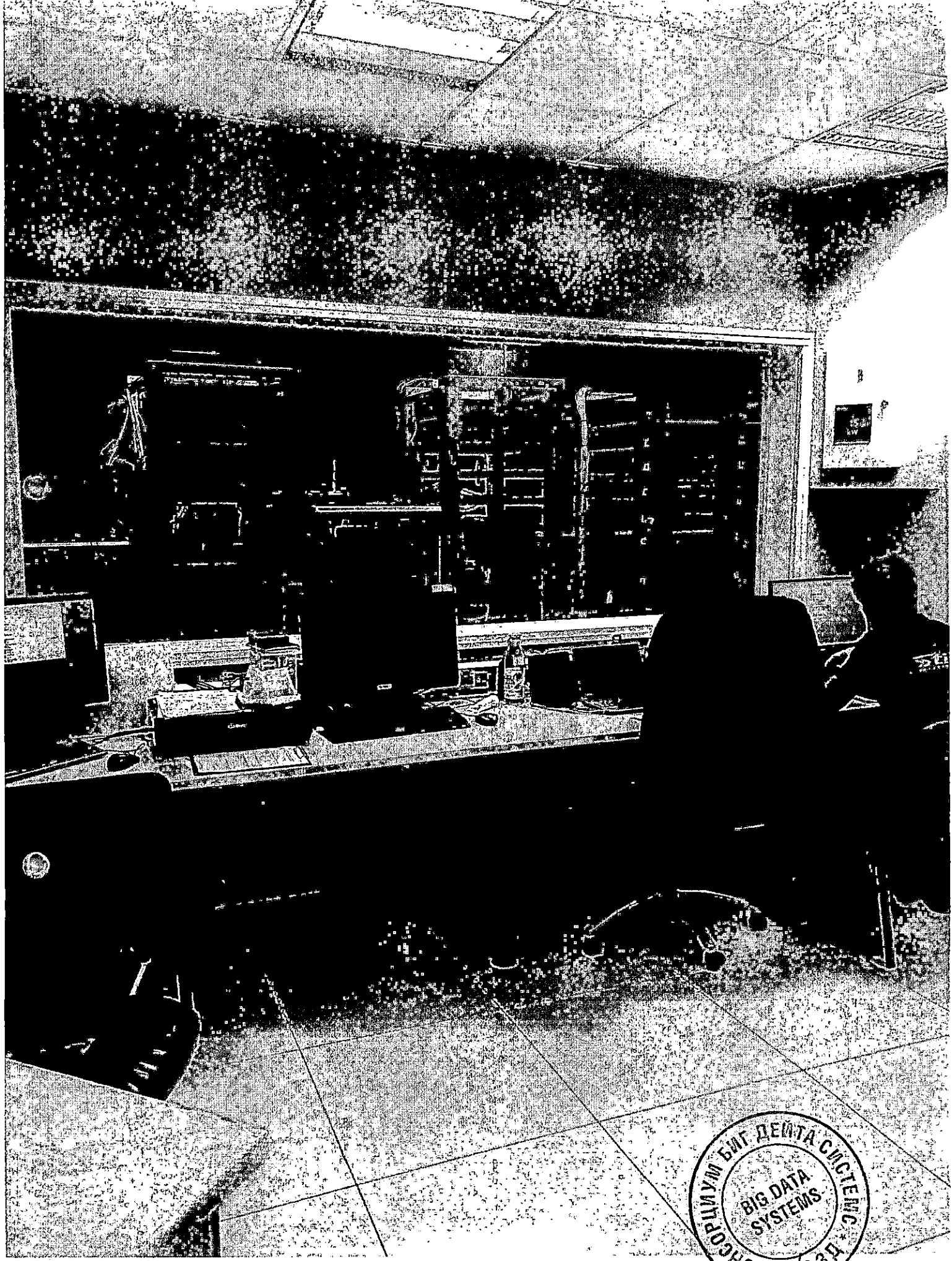
427000



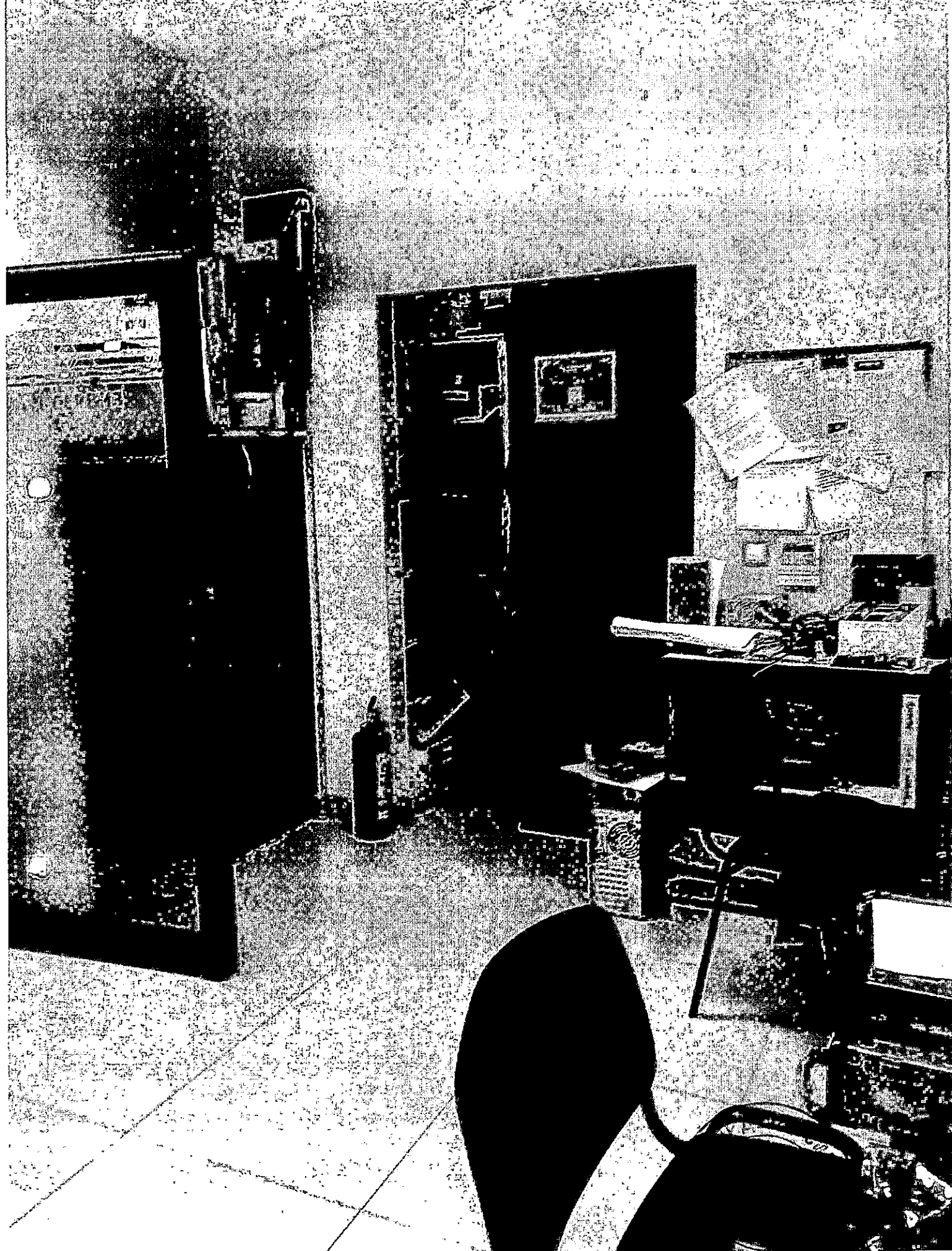
000475

COHCO





000476



1000 * 1230

0004 7

ИЗИСКВАНИЯ ЗА ПЕРСОНАЛ И/ИЛИ РЪКОВОДЕН СЪСТАВ С
ОПРЕДЕЛЕНА ПРОФЕСИОНАЛНА КОМПЕТЕНТНОСТ, т.П.3. Критерии
за подбор от Документацията за участие в Открита процедура за възлагане на
обществена поръчка с предмет:

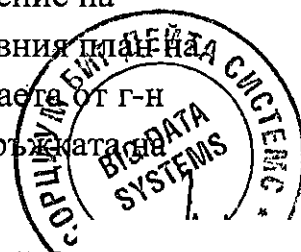
„Доставка, инсталиране, конфигуриране, тестване и поддържане на
Специализирана централизирана система под ключ HADOOP за работа с
големи данни по проект „Дигитализация на икономиката в среда на Големи
данни“ (ДИГД), № BG05M2OP001-1.002-0002“

Съгласно стр. 68 от Документацията

Оперативното ниво от организацията за изпълнение на поръчката при
Изпълнителя е в съответствие с изискванията описани в настоящата
документация и се състои от необходимите като брой експерти за успешното
изпълнение на поръчката, със следните роли и отговорности, специалисти:

1. Ръководител на проекта – 1 брой

Тази роля е от изключителна важност за успешното изпълнението на
поръчката, от страна на изпълнителя. Ръководителят на проекта ще отговаря
за цялостното изпълнение на поръчката, в сроковете, обема и качеството,
определени в Техническата спецификация, ще ръководи, координира и
контролира цялостното изпълнение на договора. Той ще взема необходимите
решения и мерки през целия жизнен цикъл на проекта по отношение на
всички аспекти от управлението на проекта при спазване на Главния план на
проекта и неговите приложения. Предлагаме тази роля да бъде заета от г-н
Александър Жеков, заради неговия опит в изпълнението и поддръжката на
ИТ проекти.



000478

„Консорциум Биг Дейта Системс“ ДЗЗД

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

- Специалист по продажби и управление на
проекти Образование

годишен опит.

2. Налични сертифицирани специалисти в следните области.

Ролята на сертифицираните специалисти е изключително важна, защото ще могат да предоставят на Възложителя много точна и компетентна помощ свързана с реализацията на целия проект.

- **управление и внедряване на хардуерни системи от производителя;**
- **Експерт информационни технологии хардуерна**

„Консорциум Биг Дейта Системс“ ДЗЗД

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

- **„Oracle Database SQL / Microsoft SQL Server” или еквивалент;**

– Старши системен анализатор Образование:

- **„ITIL Certificate in IT Service Management“ или еквивалент;**

– Експерт информационни технологии

– Експерт информационни технологии Образование:

- **„Microsoft Certified IT Professional” или еквивалент**

– Старши системен анализатор Образование:

„Консорциум Биг Дейта Системс“ ДЗЗД

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

- **„Microsoft Certified Solution Associate (MCSA)” или еквивалент;**

– Старши системен анализатор Образование:

- **управление на проекти – “PRINCE2” или еквивалент;**

– Експерт информационни технологии Образование:

- **мрежови технологии;**

- Специалист по мрежови технологии

„Консорциум Биг Дейта Системс“ ДЗЗД

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

- управление на достъпа до мрежата (Network Access Control или еквивалент) ;

----- Младши програмист Образование: Средно

- технология за защита от проникване - Intrusion Prevention System (IPS) или еквивалент;

----- Специалист по мрежови технологии

----- Мениджър на системна инфраструктура

„Консорциум Биг Дейта Системс“ ДЗЗД

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

- **Linux или еквивалент специалист .**

- Специалист Linux Essentials Образование: Магистър

3. **Минимум 1 брой наличен системен администратор (или на подобна длъжност специалист), който системен администратор е сертифициран или е с преминато обучение/положен изпит при производител/сертифициран център за обучение по „Cloudera Administrator“ или еквивалент, в която декларация се уверява, че чрез този системен администратор ще се осигурява реакция от страна на доставчика при регистрация/уведомление от страна на Възложителя за повредата в инсталацията и/или конфигурацията на управляващия софтуер на Cloudera на Hadoop клъстера (или еквивалент).**

– Младши програмист Образование



„Консорциум Биг Дейта Системс“ ДЗЗД

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

- Специалист развойна дейност Образование:

Допълнителен експерт:

– Експерт Информационни технологии



00

PEOPLECERT

AXELOS
GLOBAL BEST PRACTICE

Abid Ismail, CEO, AXELOS

Pandoraia Theleriti, Certification Qualifier, PEOPLECERT

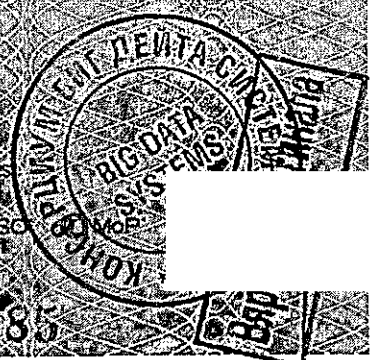
Printed on 24 October 2016

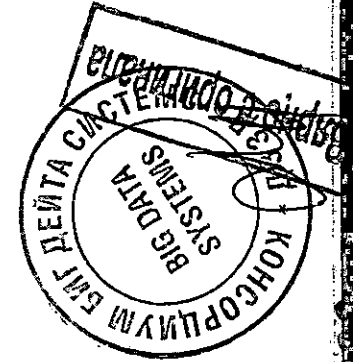
This certificate remains the property of the issuing Examination Institute and shall be returned immediately upon request.

ITIL RESILIA PRINCE2 PRINCE2 AGILE MSP M&P P3MS IP3C

AXELOS, the AXELOS logo, the AXELOS short logo, ITIL, PRINCE2, PRINCE2 AGILE, MSP, M&P, P3MS, IP3C, M&P and M&P are registered trademarks of AXELOS Limited. RESILIA is a trade mark of AXELOS Limited.

000485





CERTIFICATE *of* COMPLETION

01.02.2019

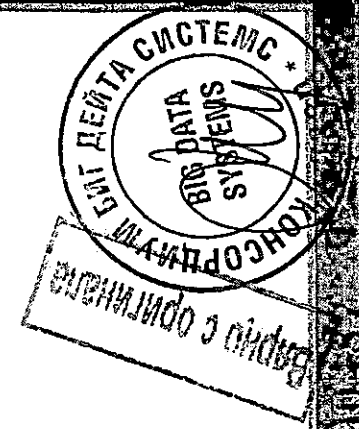
COMPLETION DATE

denov

СТОР



UPGRADING PEOPLE EVERY DAY



CERTIFICATE *of* COMPLETION

06.09.2018

COMPLETION DATE

В. П. П.

INSTRUCTOR

UPGRADING PEOPLE EVERY DAY

Certification Number
ECC2301957468



000488

CICISO

Chief Information Security Officer

Chief Information Security Officer

This is to acknowledge that

Issue Date: **24 January, 2019**

Expiry Date: **01 February, 2020**



#0702
ISO/IEC 17024
Personnel Certification Program

EC-Council

Sanjay Bavisi, President

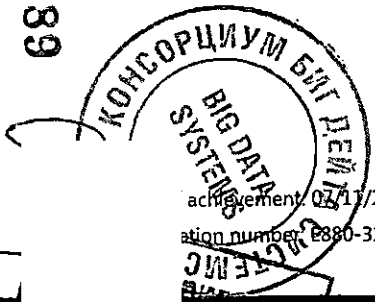


Microsoft Certified Solutions Associate

STEPHAN G GUEORGUIEV

Has successfully completed the requirements to be recognized as a Microsoft® Certified Solutions Associate: SQL Server 2012.

687000

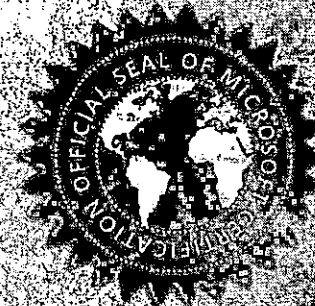


Microsoft Certificate of Excellence

Achievement Date: June 01, 2012

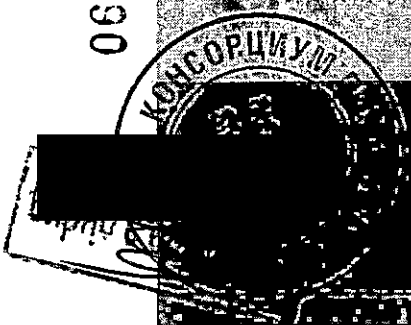
STEPHAN G GUEORGUIEV

Has successfully completed the requirements to be recognized as a Microsoft® Certified
IT Professional: Database Developer 2008



Steven A. Ballmer
Chief Executive Officer

000190



Peter Hepworth, CEO, AXELOS

Panorea Theletiti, Certification Qualifier
PEOPLECERT Group

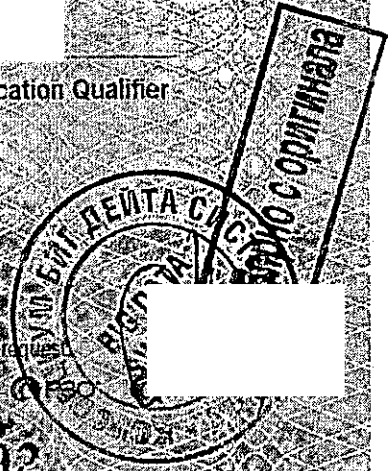
Printed on: 23 May 2018

This certificate remains the property of the issuing Examination Institute and shall be returned immediately upon request.

ITIL RESILIA PRINCE2 PRINCE2 AGILE MSP M&R P3M3 P3O M&P

AXELOS, the AXELOS logo, the AXELOS swirl logo, ITIL, PRINCE2, MSP, M&R, P3M3, P3O, M&P and May are registered trademarks of AXELOS Limited. PRINCE2 Agile and RESILIA are trademarks of AXELOS Limited.

000492



PeopleCert®

All talents, certified.

AXELOS

GLOBAL BEST PRACTICE

Peter Hepworth, CEO, AXELOS

Executive Director, Certification, PeopleCert

Printed on 19 November 2018

This certificate remains the property of the issuing Examination Institute and shall be returned immediately upon request.

ITIL® RESILIA® PRINCE2® PRINCE2 Agile® MSP® MGR® FRM® P3O®

AXELOS and AXELOS logo are trademarks of AXELOS LIMITED, a UK registered company. AXELOS and AXELOS logo are registered trademarks of AXELOS LIMITED in the UK and other countries. PeopleCert is a trademark of PeopleCert Limited.



000493

SvetlaTex

TRANSLATION CENTRE

ЕТ "СВЕЛТАТЕКС - СВЕЛЛАНА ПЕТРОВА - АЛХУСАМИ"

Ул. "Искър" №11, 1202 София, тел./факс 02 980 02 02, мобилен: 0877829091

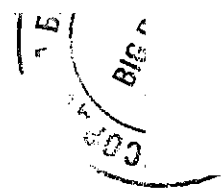
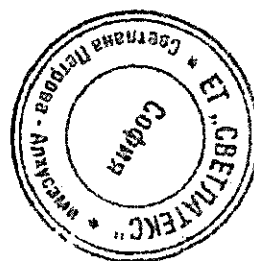
11 Iskar St., 1202 Sofia, Bulgaria, tel./fax: (+359) 2 980 02 02, mobile phone: (+359) 877829091

e-mail: svetlatex@yahoo.com

Превод от английски език



За верността на п.
Светлана Митева



000494

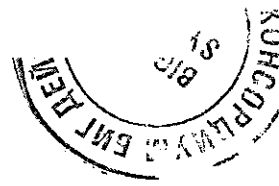
CompTIA



964000 <

WatchGuard Technical Certification

967000



Joanne Miller
Director of Training



WatchGuard Technical Certification



Joanne Miller
Director of Training





CERTIFICATE *of* COMPLETION

000498



13.01.2017

COMPLETION DATE



INSTRUCTOR

UPGRADING PEOPLE EVERY DAY

Achievement

MICROSOFT CERTIFICATE OF

This certificate accredits that

167000

New Horizons
Computer Learning Centers
CHOOSE. LEARN. SUCCEED.

Вярно с оригинала

Satya Nadella
Chief Executive Officer

Microsoft Certified Trainer

New Horizons Bulgaria
10.02.2017

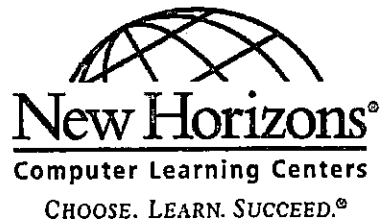
Certificate of Completion

000500

08.01.2016

Date

Vladimir Simeonov, Instructor



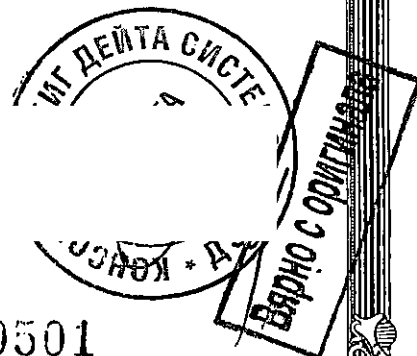
macmon
nac ■ intelligent einfach

ZERTIFIKAT

N
C

Christian Bucker
Geschäftsführung

macmon secure GmbH | Alte Jakobstraße 79-80 | 10179 Berlin
+49 30 2325 777-0 | nac@macmon.eu



000501

cloudera®

Certificate of Completion

[Signature]

Senior Director, Educational
Services

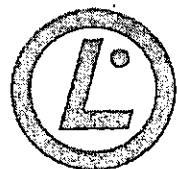
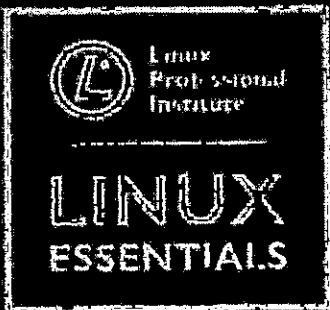
Cloudera, Inc.
www.cloudera.com

08/30/2018

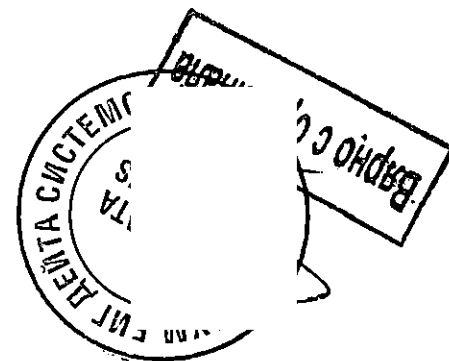
Date

Big Data Systems
1337
STEMC
Cloudera - OPTIMIZATION

000502



Linux
Professional
Institute



n

000503

Handwritten signature

Chairman of the Board

Executive Director

This certificate is not proof of certification, please visit:

ipi.org/v/LPI000413221/7eghzv3j9d

Linux Professional Institute and the circle-L logo are trademarks and registered trademarks (in Japan and the European Union) of Linux Professional Institute Inc. All Rights Reserved.



cloudera®

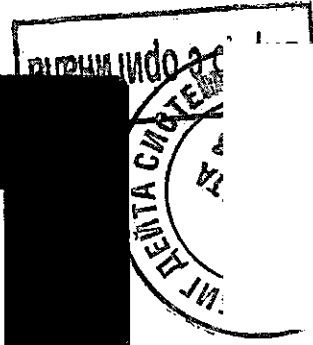
Certificate of Completion

Senior Director, Educational
Services

Cloudera, Inc.
www.cloudera.com

02/06/2019

Date



000504

SUPERMICR

WE KEEP IT GREEN.™

Supermicro - Certified Engineer

This certificate is awarded to

000305



certified

Intel® Technology Expert



The Intel® Technology Provider Program Recognizes

Maurits Tichelman
Vice President and General Manager,
Direct and Channel Sales

Date

Взяно с оригинала

0506

Your solutions. Our technology. Smarter together



ДОКУМЕНТ

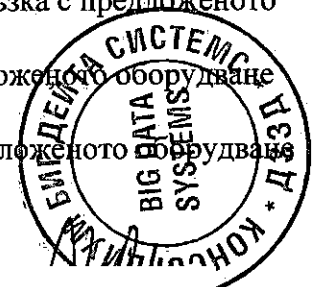
Описание на Оторизационните писма, предоставени от производителя на предложеното в Техническото предложение оборудване

Във връзка с Декларираното обстоятелство, че сме оторизирани от производителя на предложеното решение, описано на стр. 128 от Документацията за участие в открита процедура с предмет:

„Доставка, инсталиране, конфигуриране, тестване и поддръжане на СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ НАДООР за работа с Големи данни по проект „Дигитализация на икономиката в сѐ данни“ (ДИГД), № BG05M2OP001-1.002-0002“

Предоставяме следните оторизационни писма:

1. От MicroFocus към „Биг Дата Системс“ ДЗЗД във връзка с предложеното решение за управление и анализ на събития – **Sentinel Enterprise на MicroFocus.**
2. От Super Micro към „Биг Дата Системс“ ДЗЗД във връзка с предложените хардуерни решения на **SuperMicro** описани в Техническото предложение.
3. От WatchGuard към „Биг Дата Системс“ ДЗЗД във връзка с предложеното IPS устройство – **WatchGuard M4600**, описано в Техническото предложение.
4. От MACMON SECURE към „Биг Дата Системс“ ДЗЗД във връзка с Модул за Мрежово управление и достъп – **macmon Network bundle NAC (Network Access Control).**
5. От Allied Telesis International Ltd към „Биг Дата Системс“ ДЗЗД във връзка с предложеното оборудване за 24 портов комутатор.
6. От „G-TEC Europe“ S.r.l към „Биг Дата Системс“ ДЗЗД във връзка с предложеното оборудване за UPS.
7. От Socomes към „Биг Дата Системс“ ДЗЗД във връзка с предложеното оборудване за UPS системи.
8. От RED HAT към „Биг Дата Системс“ ДЗЗД във връзка с предложеното оборудване от типа Linux.



До: Университет за Национално и Световно Стопанство

Студентски град, София, България

ОТНОСНО: Процедура за обществена поръчка: „Доставка, инсталиране, конфигуриране, тестване и поддръжане на СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ NADDOOP за работа с Големи данни по проект „Дигитализация на икономиката в среда на Големи данни“ (ДИГД), № BG05M2OP001-1.002-0002“

Уважаеми дами, господа,

С настоящото даваме оторизация за участие в обществена поръчка: „Доставка, инсталиране, конфигуриране, тестване и поддръжане на СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ NADDOOP за работа с Големи данни по проект „Дигитализация на икономиката в среда на Големи данни“ (ДИГД), № BG05M2OP001-1.002-0002“ на Консорциум “Биг Дейта Системс“ ДЗЗД с адрес град София, ул. Мюнхен, № 8 и управител Деян Желев Йовков.

Участници в консорциума са оторизирани партньори на Micro Focus Intl. и притежават необходимите квалификации и опит да оферират, доставят, инсталират и поддръжат продуктите ни.

С уважение:

София, 18.04.2019 г.

Веселин Янков

Представител за България, Македония и Албания



ОТОРИЗАЦИОННО ПИСМО
ОТ
SUPER MICRO

До:
Университет за Национално и Световно Стопанство
Студентски град
София, 1700
България

25 Април 2019 г.

Открита процедура № BG05M2OP001-1.002-0002

Уважаеми г-н/г-жо,

С настоящото писмо потвърждаваме, че PERSY LTD, с офис на 136Б бул. Цар Борис 3, София, България са оторизиран дистрибутор на Super Micro Computer, Inc. ("Super Micro"). PERSY LTD имат право да предоставят оферти, да продават и да осъществяват дистрибуция на продуктите на SUPER MICRO, чрез „Консорциум Биг Дейта Системс“ ДЗЗД, с офис на ул. Мюнхен 8, София, 1592, България, както и да осъществяват след продажбени услуги и поддръжка (предоставени от екипа на SUPER MICRO GLOBAL Service) (заедно с "Продуктите"), на територията на България.

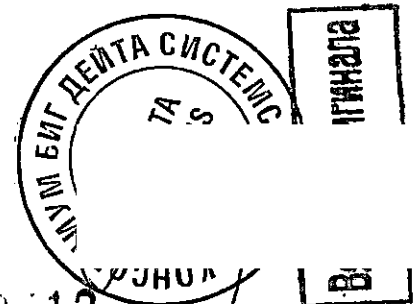
Следователно, „Консорциум Биг Дейта Системс“ са оторизирани да предоставят оферти и съответно да продават продуктите на SUPER MICRO за този проект в съответствие с посочените по-горе изисквания.

Това право е при условие, че то НЕ противоречи на листа на Американското Ембарго и може да бъде отменено и/или подновено от Super Micro по всяко едно време.

С Уважение,

Don Clegg

Старши Вицепрезидент в Световен мащаб



000512



University of National and World Economy (UNWE)
Student Town
Sofia, 1700
Bulgaria

April 25, 2019

Re: Tender No. BG05M2OP001-1.002-0002

Dear Sir,

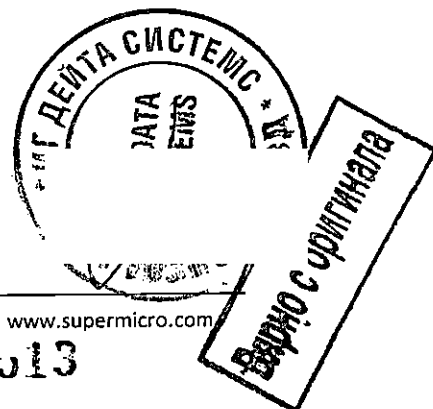
This letter confirms that PERSY LTD, having their office at 136B TZAAR BORIS III BLVD, SOFIA, 1618, Bulgaria is a Super Micro Computer, Inc. ("Super Micro") authorized distributor. PERSY LTD has the right to offer for sale, sell and distribute Super Micro products through Big Data Systems DZZD, having their office at 8 Munchen Str., Sofia, 1592, Bulgaria and related post-sales service and support (provided by Super Micro Global Service team) (together the "Products"), in the territory of Bulgaria.

Therefore, Big Data Systems DZZD is authorized to quote and sell Super Micro Products for this project in conformance with the aforementioned requirements.

This right is on the condition that it does NOT contravene the US Embargo list and may be revoked and/or renewed by Super Micro at any time.

Sincerely,

Don Clegg
Senior Vice President, Worldwide Sales



ДО: Университет за Национално и Световно Стопанство

Хамбург 24,04,2015

Оторизационно писмо от WATCHGUARD

Да се има предвид, че Watchguard Technologies, които са доказани и надеждни производители на WatchGuard Firebox и Access Points с настоящото оторизира:

„Консорциум Биг Дейта Системс“ ДЗЗД, София 1592, ул. Мюнхен 8

ДА участва в открита процедура с предмет:

„Доставка, инсталиране, конфигуриране, тестване и поддържане на СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ HADOOP за работа с Големи данни по проект „Дигитализация на икономиката в среда на Големи данни“ (ДИГД), № BG05M2OP001-1.002-0002“

Michael Haas
Търговски директор за Централна Европа
WatchGuard Technologies



100514





To: University of National and World Economy (UNWE)

Hamburg 24.04.2019

Manufacturer's Authorization Form

WHEREAS WatchGuard Technologies who are established and reputable manufacturers of WatchGuard Firebox and Access Points hereby authorizes:

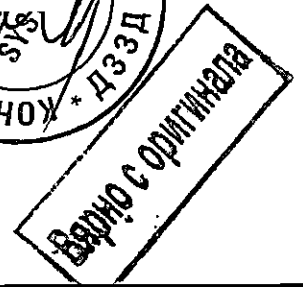
Big Data Systems" DZZD, 8 Munchen Str., Sofia 1592

TO:

Delivery, installation, configuration, testing and support of a "Specialized Centralized Turnkey System Hadoop" for Big Data Processing under the project "Digitalization of Economics in a Big Data Environment" (DEBD), № BG05M2OP001-1.002-0002.

Area Sales Director Central Europe
WatchGuard Technologies

Handwritten initials and marks: a large 'A' with a slash, a checkmark, and the letters 'W' and 'D'.



)

**Оторизационно писмо
ОТ
MACMON SECURE**

Във връзка с открита процедура:

„Доставка, инсталиране, конфигуриране, тестване и поддържане на СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ HADOOP за работа с Големи данни по проект „Дигитализация на икономиката в среда на Големи данни“ (ДИГД), № BG05M2OP001-1.002-0002“

ДО:

Университет за Национално и Световно Стопанство
София 1700, Студентски град

От:

„Консорциум Биг Дейта Системс“ ДЗЗД
София 1592, ул. Мюнхен 8

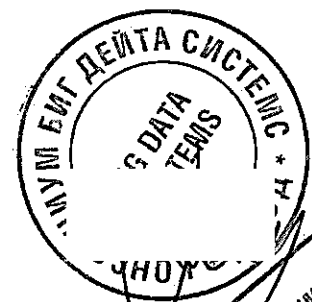
Да се има предвид, че MACMON SECURE GMBH със седалище в Европа ситуирано в **Aite Jakobstrabe 79-80, 10179 Берлин**, които са официални производители на продуктите на MACMON, оторизира: Big Data Systems DZZD, ситуирани на ул. Мюнхен 8, София 1592, България(по-долу, „Участник“) да предложи оферта и впоследствие да преговаря и подпише договор със Вас за продажба на продукти на MACMON SECURE.

Потвърждаваме, че в случай на приемане на офертата и съгласие по договора между Вас и Участникът, MACMON продуктите ще бъдат доставени с пълната си гаранция.

Име: Dragana Stanisic

Подпис и печат в оригинала: ДА

23.04.2019 г.



000516





Manufacturer's Authorization

Invitation for Bids

"Delivery, installation, configuration, testing and support of a "Specialized Centralized Turnkey System Hadoop" for Big Data Processing under the project "Digitalization of Economics in a Big Data Environment" (DEBD), № BG05M2OP001-1.002-0002"

Tender Reference:

University of National and World Economy (UNWE)
1700 Sofia, Student Town, UNWE

To:

Big Data Systems DZZD
8 Munchen Str., Sofia 1592, Bulgaria

WHEREAS MACMON SECURE GMBH with its European Headquarter based at **Alte Jakobstraße 79-80, 10179 Berlin** who are official producers of MACMON products do hereby authorize: Big Data Systems DZZD, located at 8 Munchen Str., Sofia 1592, Bulgaria (hereinafter, the "Bidder") to submit a bid and subsequently negotiate and sign a Contract with you for resale of MACMON SECURE product(s).

We confirm that, in case the bidding results in a Contract between you and the Bidder, MACMON products will come with our full ordered warranty.

Name: Dragana Stanisic

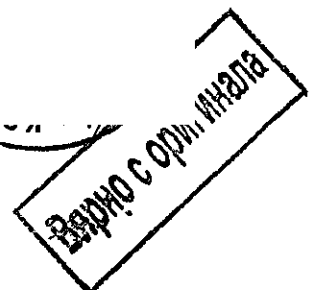


Signed and stamped_

nac@macmon.eu · www.macmon.eu

Duty authorized to sign the authorization for and on behalf of: **MACMON SECURE GMBH**

23.04.2019



000317

С
/

Оторизационно писмо

Allied Telesis International Ltd., които са установен и уважаван производител на хардуерни компоненти за мрежова среда и имат производствени съоразения в Allied Telesis Intl. Asia Pte Ltdq bldk 1026 Tai Seng Avenue, #05-3538q, Tai Seng Industrial Estate, Singapore 534412, оторизира, „Консорциум Биг Дейта Системс“ ДЗЗД, ул. Мюнхен № 8, София 1592, да предложи оферта и в последствие да подпише договор с вас във връзка с посочения търг.

„Консорциум Биг Дейта Системс“ ДЗЗД има сертифицирани служители, които могат да конфигурират, инсталират, поддържат и доставят решения базирани на нашето оборудване.

11.05.2019

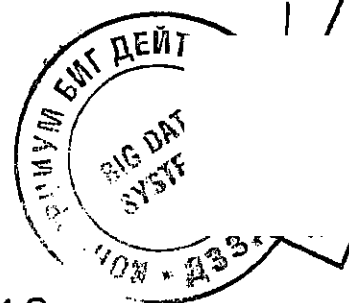
Dejan Malic

Country Manager

Balkan Region

For Allied Telesis International

Подпис и печат



000518

0101011011 001100101010100100001 010101



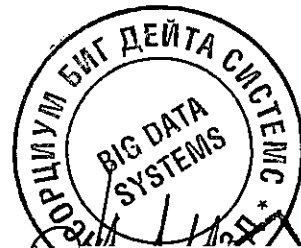
Manufacturer's Authorization Form

WHEREAS **Allied Telesis International Ltd.** who are established and reputable manufacturers or producers of **Hardware Components for Network Environments** having production facilities at **Allied Telesis Intl. Asia Pte Ltd, Bldk 1026 Tai Seng Avenue, #05-3538, Tai Seng Industrial Estate, Singapore 534412** do hereby authorize „Big Data Systems” DZZD. 8, Munich Str., Sofia 1592 to submit a bid, and subsequently negotiate and sign the Contract with you against the mentioned tender.

Big Data System d.o.o. has certified employees and capabilities to design, install, maintain and deliver solutions based on our equipment.

11.5.2019.

Dejan Malic
Country Manager
Balkan Region
For Allied Telesis International



Allied Telesis EMENA GmbH
Representative Office Belgrade * Alekse Nenadovića 15/3 | 11000 Belgrade, Serbia
Telephone +381-11-2416-410

000319

alliedtelesis.com

the solution : the network



ОТОРИЗАЦИОННО ПИСМО

Дата: 09.05.2019

До: УНИВЕРСИТЕТ ЗА НАЦИОНАЛНО И СВЕТОВНО СТОПАНСТВО

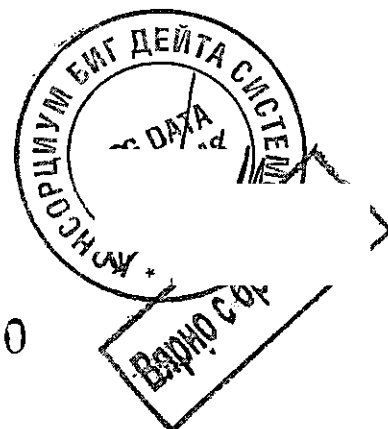
Бид: „Доставка, инсталиране, конфигуриране, тестване и поддържане на СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ HADOOP за работа с Големи данни по проект „Дигитализация на икономиката в среда на Големи данни“ (ДИГД), № BG05M2OP001-1.002-0002“.

Red Hat Ltd. („Red Hat“), официалният производител и дистрибутор на Red Hat Enterprise Linux и JBoss Enterprise Application Platform в EMEA, с регистрирани офиси на 6700 Cork Airport Business Park, Kinsale Road, Cork, Ирландия, с настоящото оторизира SSI Ltd. („Партньор“) да представи офертата и впоследствие да подпише договор с Клиента от свое име и от името си за препродажба на продукти и услуги на Red Hat, свързани с офертата, издадена от Клиента („Договор“).

В случай, че Офертата доведе до договор, продуктите на Red Hat ще бъдат поддържани в съответствие с правилата за поддръжка, описани на адрес <http://www.redhat.com/support> и сертифицираните хардуерни правила, описани в <http://www.redhat.com/hardware>.

Подпис и печат

Michael O'Neill





Red Hat Limited | 6700 Cork Airport Business Park | Kinsale Road | Cork Ireland
T: +353 21 2303400 | F: +353 21 2303499 | www.redhat.com

Manufacturer's Authorization Form

Date: 09/05/2019

To: University of National and World Economy (UNWE) ("Client")

Bid Description: "Delivery, installation, configuration, testing and support of a "Specialized Centralized Turnkey System Hadoop" for Big Data Processing under the project "Digitalization of Economics in a Big Data Environment" (DEBD) ("Bid")

On the date hereof, Red Hat Ltd. ("Red Hat"), the official producer and distributor* of Red Hat Enterprise Linux and JBoss Enterprise Application Platform in EMEA, with registered offices at 6700 Cork Airport Business Park, Kinsale Road, Cork, Ireland, hereby authorizes SSI Ltd. ("Partner") to submit the Bid and subsequently negotiate and sign a contract with Client in its own name and behalf, for the resale of Red Hat products and services associated with the Bid issued by Client ("Contract").

In case the Bid process results in a Contract, the Red Hat products will be supported in accordance with the support policies described at <http://www.redhat.com/support> and the certified hardware policies described at <http://www.redhat.com/hardware>.

RED HAT LTD
6700 Cork Airport Business Park,
Kinsale Road,
Cork

Red Hat Limited

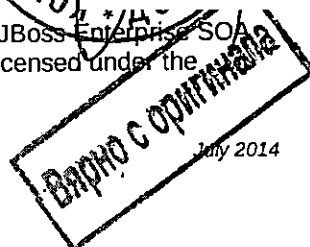
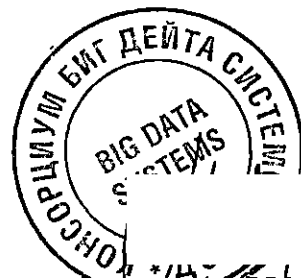
Signature
Michael O'Neill

Printed Name
EMEA Finance Senior Director

Title
May 9, 2019

Date

JJ



*Please note: JBoss Enterprise Application Platform, JBoss Enterprise Portal Platform, JBoss Enterprise SOA Platform and Red Hat Enterprise Linux are open source software compilations that are licensed under the

(Lesser) General Public License ((L)GPL), other open source licenses and Red Hat's End User License Agreement. The packages included in the software compilations itself are developed by independent engineers around the world who contribute source code (also known as the open source community). Therefore there is not a single developer or "producer" as in a traditional proprietary model. Each of the packages that are contained within the compilations are subject to their own copyright and Red Hat owns a copyright in the collective work.



Офис:

Corso Palladio, 147, 36100 Vicenza
Mailing Office/H Q/
Strada Marosticana 81/13
36031 Polovaro Di Dueville (VI)

ДО: УНИВЕРСИТЕТ ЗА НАЦИОНАЛНО И СВЕТОВНО СТОПАНСТВО

ОТОРИЗАЦИОННО ПИСМО

Референтен номер: AOP-00062-2019-0013

„Доставка, инсталиране, конфигуриране, тестване и поддържане на СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ NADDOOP за работа с Големи данни по проект „Дигитализация на икономиката в среда на Големи данни“ (ДИГД).

Имайки в предвид:

Ние, "G-TEC Europe" S.r.l., които сме официален производител на оборудване за непрекъсваема доставка на енергия, притежавайки фабрика в Страда Маростикана 81/15 – 36031 Поволаро (VI) – Италия, оторизираме консорциума "Big Data Systems DZZD" намиращи се в София, България да подаде оферта за търг, в която да достави UPS оборудване от тип "AP160N" произведено от нас и в последствие да договори и подпише договор. Ние удължаваме нашата пълна гаранция към продуктите предлагани от горепосочената компания и надлежно оторизираме участника в търга да действа от наше име, за да може да изпълни всички гаранционни задължения към изброените по-горе продукти оферирани за препродажба от участника в търга във връзка с тази покана за търг.

Ние също така потвърждаваме, че участника в търга има собствена организация и е квалифициран да достави услуги по поддръжка на територията на България, технически и хелп-деск съпорт, ъпгрейд към нови версии и/или други услуги свързани с описаните по-горе продукти.

Ние декларираме, че все тип продукт конфигуриран за тази тръжна процедура е оригинално оборудване директно от производителя.



Подписано от: инж. Лука Вивиан – Ериа мениджър на Източна Европа

000523

Вярно с оригинала



Registered office:

Corso Palladio,147 , 36100 Vicenza
Mailing Office/H Q/
Strada Marosticana 81/13
36031 Povolaro Di Dueville(VI)

Tel.+39-0444591420 or 592463 - Fax.+39-0444365191 ~ E-mail: <luca.vivian@gtec-power.eu>

TO: University of National and World Economy - Bulgaria

Manufacturer's Authorization

Reference Number: AOP – 00062-2019-0013

Subject: Supply, Installation, Configuring, Testing and Commissioning of a Specialized and Centralized System under key HADOOP for working with large data mass according the project "Digitalization of the economy in a large data environment"

Our sign:

Your sign:

Date: 09.05.2019

WHEREAS

We, "G-TEC Europe" S.r.l. who are official manufacturers of uninterruptible power supply equipment, having factories at Strada Marosticana 81/15 • 36031 Povolaro (VI) – Italy, do hereby authorize the Consortium "Big Data Systems DZZD" located in Sofia, Bulgaria, to submit a bid the purpose of which is to provide UPS equipment of the products range by the name of "API 60N" manufactured by us and to subsequently negotiate and sign a Contract. We hereby extend our full guarantee and warranty with respect to the Goods offered by the above company and duly authorize the Bidder to act on our behalf in fulfilling all warranty obligations with respect to the above-listed products, offered for resale by the Bidder in relation to this Invitation for Bids.

We also certify that the Bidder has his own organization and is qualified by us to provide the service maintenance at the territory of Bulgaria, technical or help desk support, new version upgrade and/or other services related to the above-listed Products.

We declare, that every type of equipment configured for this bidding procedure is from the same Original Equipment Manufacturer.

Signed: *Eng. Luca Vivian - East Europe Area Manager*

G-Tec Europe srl
Strada Marosticana, 81/13
36031 Povolaro di Dueville (VI)
Tel. 0444 361321 - Fax 0444 365191

Dated on 09th day of May, 2019

G-TEC Europe S.r.l - Strada Marosticana 81/15 • 36031 Povolaro (VI) ITALIA
Tel. +39 0444 591420 / +39 0444 592463 - Fax +39 0444 365191
<http://www.gtec-power.eu>
P. IVA 03185600248



000524



UPS - POWER SYSTEMS NAPAVALNI SISTEMI

Socomec UPS d.o.o.; Savlje 89
SI – 1000 Ljubljana
Tajništvo tel.: +386(0)15807860
Prodaja tel.: +386 (0)1 580 78 71
Fax: +386 (0) 1 561 11 73

TO: University of National and World Economy - Bulgaria

Your sign:

Our sign:

Date: 09.05.2019

Manufacturer's Authorization

FROM
SOCOMECS UPS d.o.o.
East Central Europe

Reference Number: AOP – 00062-2019-0013

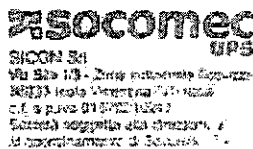
Subject: Supply, Installation, Configuring, Testing and Commissioning of a Specialized and Centralized System under key HADOOP for working with large data mass according to the project "Digitalization of the economy in a large data environment"

WHEREAS we, SOCOMECS UPS d.o.o., the subsidiary of the established official manufacturers of Uninterruptible Power Supply equipment SOCOMECS S.A., having factories located at 1/3, Via Sila – Zona Industriale Scovizzo, 36033 Isola Vicentina (VI), Italy, do hereby authorize the Consortium "Big Data Systems" DZZD to submit a bid in relation to the Invitation for Bids indicated above, the purpose of which is to provide the following Goods, manufactured by us - UPSs with brand Socomec, and to subsequently negotiate and sign the Contract.

We hereby extend our full guarantee and warranty with respect to the Goods offered by the above consortium in reply to this Invitation for Bids.

Name: Bostjan Flis
In the capacity of: Managing director

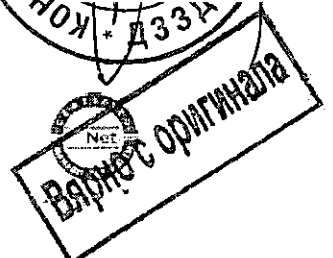
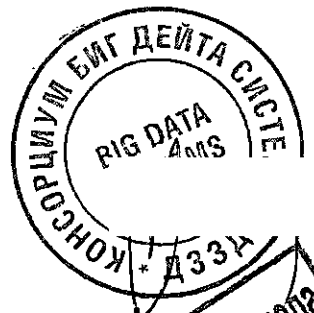
Signed:



Dated on 09th day of May, 2019

TR: 27000-000045380, MATICIMA ST. TR 5315.158, SIFRA DEJAVNOSTI: 31.10053151589T.
REG VLOZKA OKROZNEGA SODISCA V LJUBLJANI: 1/3852/00
OZNAKA IN ŠT. SKLEPA: SRG 2742/89. IDENTIFIKACIJSKA ŠT. ZA DDV: S185510947 OSNOVNI
KAPITAL DRUŽBE: 19.958.000.00 SIT

CERTIFIKAT
ŠT. 0315
000135
150901



Socomec UPS d.o.o; Savlje 89
SI-1000 Ljubljana

ДО: УНИВЕРСИТЕТ ЗА НАЦИОНАЛНО И СВЕТОВНО СТОПАНСТВО

ОТОРИЗАЦИОННО ПИСМО

ОТ: SOCOMEC UPS D.O.O

Референтен номер: АОП- 00062-2019-0013

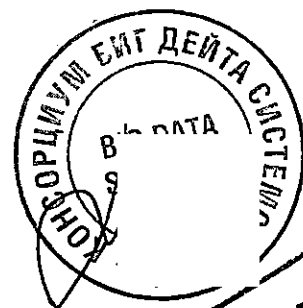
Относно: „Доставка, инсталиране, конфигуриране, тестване и поддържане на СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ HADOOP за работа с Големи данни по проект „Дигитализация на икономиката в среда на Големи данни“ (ДИГД).

Да се има предвид:

Ние от SOCOMEC UPS d.o.o. дружество на утвърдените официални производители на оборудване за UPS SOCOMEC S.A. притежаващи фабрики в 1/3 Виа Сила Индустриална зона Сковице, 36033 Изола Вицентина (VI), Италия оторизираме Консорциум “Big Data Systems” DZZD да подаде оферта във връзка с поканата за търг описана по-горе, целта на която е да достави следните продукти произведени от нас – UPSs с марката Socomec и в последствие да договори и подпише договора. Ние удължаваме нашата пълна гаранция към продуктите предлагани от горепосочения консорциум в отговор на поканата за този търг.

Име: Бостян Флис

В качеството си на: Managing director

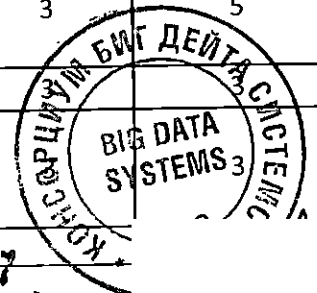


000526

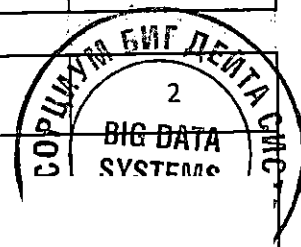
ПЛАН – ГРАФИК ЗА ИЗПЪЛНЕНИЕ НА ПРОЕКТ С ПРЕДМЕТ:

„Доставка, инсталиране, конфигуриране, тестване и поддръжане на СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ HADOOP за работа с Големи данни по проект „Дигитализация на икономиката в среда на Големи данни“ (ДИГД), № BG05M2OP001-1.0 0002“

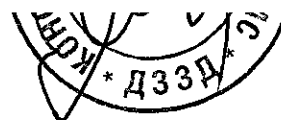
№	Планирана дейност	Начало	Завършване
ОБЩИ ПОКАЗАТЕЛИ			
1	Доставка и разполагане на хардуената техника - 44 физически сървъра и 1 хардуерно устройство за защита от прониквания.	2	3
2	Доставка на софтуера, необходим за изпълнението на поръчката.	1	1
3	Изготвяне съгласуване и предаване на отчетните материали за доставката на техниката и софтуера	1	1
ПРОЦЕСИ			
1	Определяне на цялостната организационна структура на Възложителя	1	1
2	Формиране на екипите	1	2
3	Провеждане на стартираща среща по проекта	1	2
4	Планиране на изпълнението на проекта	1	1
5	Изграждане на част от средите при изпълнителя	1	2
6	Изготвяне съгласуване и предаване на отчетните материали за фазата	1	1
Инсталиране на Cloudera / Hadoop			
1	Инсталиране на операционни системи върху физически хостове - 44 физически сървъра, 44 Linux	3	5
2	Вдигане на Apache HADOOP		
3	Инсталиране на Cloudera Express = Edge Node, Management Nodes, Data nodes, Name nodes.		



4	Конфигуриране на системата	3	3
5	Тестване на системата	1	3
Инсталиране на IPS и интегриране през API към NAC			
1	Инсталиране на WG FB M4600	1	2
2	Интегриране на WG M4600 през API Gateway към масов NAC	3	2
Инсталиране на масов NAC			
1	Инсталиране на виртуалния аплийнс за сървърната част, връзка със суичовете и сетъп на крайните точки, настройване на VLANs.	2	2
Инсталиране на Sentinel Enterprise			
1	Инсталиране на софтуер за виртуализация на всички виртуални машини на компонентите на Sentinel Enterprise	2	3
2	Свързване на системите които ще подават логове към Sentinel-a: WG IPS, масов NAC, Cloudera Manager	2	4
3	Настройване на дашбордове	1	2
4	Настройване на репорти	1	2
5	Интегриране на Sentinel Enterprise през API Gateway към масов NAC	3	3
6	Интеграция между между Модула за УАПС и LOG файловете на модулите МодулDataNode, МодулNameNode, МодулManagementNode и МодулEdgeNode, както и тези на операционните системи на тези модули	3	2
7	Интеграция между Модула за УАПС и LOG файла на Cloudera Manager и на неговата операционна система	3	2
8	Интеграция между IPS устройството и Модула за МУД	3	2
9	Интеграция между модул за УАПС и Модул за МУД	3	2
Тестване на цялата система, оживяване и довършителни работи			
1	Тестване на цялата система, оживяване и довършителни работи	1	4
2	Изготвяне съгласуване и предаване на отчетните материали за фазата	1	1
1	Предаване на цялостната система с реализирана пълна функционалност	1	2
2	Обучение на администраторите във връзка с настъпилите промени	6	



	<ul style="list-style-type: none"> - Инсталиране и конфигуриране на Hadoop система; - Основи на програмирането в среда MapReduce; - Създаване на тестови данни за провеждане на изследвания. 		
3	Изготвяне съгласуване и предаване на отчетните материали за фазата	1	1
	Общо работни дни за изпълнение на проекта	60	
	Общо месеци за изпълнение на проекта на база 20 работни дни в месеца	3	



000529



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН СЪЮЗ
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ



НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

Образец №3

ЦЕНОВО ПРЕДЛОЖЕНИЕ

До:

УНИВЕРСИТЕТ ЗА НАЦИОНАЛНО И СВЕТОВНО СТОПАНСТВО

Ценово предложение за обществена поръчка с предмет: „Доставка, инсталиране, конфигуриране, тестване и поддържане на СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ НАDOOP за работа с Големи данни по проект „Дигитализация на икономиката в среда на Големи данни“ (ДИГД), № BG05M2OP001-1.002-0002“

От: "КОНСОРЦИУМ БИГ ДЕЙТА СИСТЕМС" ДЗЗД

Участник: Деян Йовков

УВАЖАЕМИ ГОСПОЖИ И ГОСПОДА,

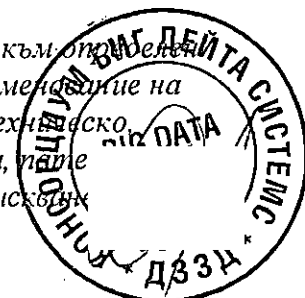
С настоящото Ви представяме нашето Ценово предложение за изпълнение на обявената от Вас обществена поръчка с предмет:

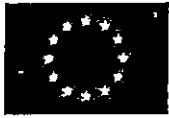
„Доставка, инсталиране, конфигуриране, тестване и поддържане на СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ НАDOOP за работа с Големи данни по проект „Дигитализация на икономиката в среда на Големи данни“ (ДИГД), № BG05M2OP001-1.002-0002“

За изпълнение на обществената поръчка, предлагаме следната обща цена **3 187 007 лева** без ДДС, (словом три милиона сто осемдесет и седем хиляди и седем лева) и **3 824 408.40 лева** с ДДС (три милиона осемстотин двадесет и четири хиляди и четиристотин и осем лева и четиридесет стотинки), от които за гаранционното обслужване за срока на гаранцията ще е **0,00 (нула) лв.** без ДДС и **0,00 лева** с ДДС (нула лева).

Забележка: Цените следва да бъдат представени в лева без ДДС и с ДДС. Отговорност за евентуално допуснати грешки или пропуски в изчисленията на предложените цени носи

В настоящата документацията, не се съдържа изисквания, насочващи към отговорен производител или доставчик. В случай, че наименование или част от наименованието на съпада с конкретен стандарт, спецификация, техническа оценка, техническо одобрение, технически еталон и модел, източник, процес, търговска марка, произход или производство, да се приема, че възложителят е поставил изискванията еквивалентно/и".





ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ



НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

единствено участникът в процедурата. При разлика в посочените цени, изписани с цифри и думи, възложителят ще приема за верни написаните с думи.

Приложение: Количествено - стойностна сметка (свободна редакция)

Участник: Деян Йовков (име и фамилия)

Длъжност: Управител

дата: 15.05.2019

подпис и печат:



В настоящата документацията, не се съдържа изисквания, насочващи към определен производител или доставчик. В случай, че наименование или част от наименование на съпада с конкретен стандарт, спецификация, техническа оценка, техническо, одобрение, технически еталон и модел, източник, процес, търговска марка, патент, тип, произход или производство, да се приема, че възложителят е поставил изискването "или еквивалентно/и".

„Консорциум Биг Дейта Системс“ ДЗЗД

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

ПРИЛОЖЕНИЕ

Количествено стойностна сметка

Във връзка с открита процедура:

„Доставка, инсталиране, конфигуриране, тестване и поддържане на СПЕЦИАЛИЗИРАНА
ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ НАDOOP за работа с Големи данни по
проект „Дигитализация на икономиката в среда на Големи данни“ (ДИГД), №
BG05M2OP001-1.002-0002“

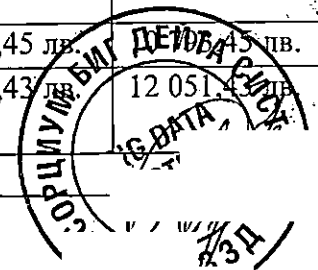
Базовата Специализирана хардуерна система					
Ред	Описание	Парт номер	Брой	Ед.стойност без ДДС	Обща стойност без ДДС
1	МодулDataNote модел SuperMicro с Cloudera с RedHat	SSG-6019P-ACR12L	12	35 328,77 лв.	423 945,24 лв. ✓
2	Основен комуникационен модул – Super Micro	SSE-X3348S	1	17 588,62 лв.	17 588,62 лв. ✓
3	Допълнителен комуникационен модул Allied Telesis	AT-x510L-28GT	1	12 288,27 лв.	12 288,27 лв. ✓
4	KVM switch	KVM	1	10 793,45 лв.	10 793,45 лв. ✓
5	UPS система модел Gtec AP160N	Gtec AP160N	1	12 051,43 лв.	12 051,43 лв. ✓
Разширяема Специализирана хардуерна система за обработка на данни.					
Ред	Описание	Парт номер	Брой	Ед.стойност без ДДС	Обща стойност без ДДС
1	МодулDataNote модел SuperMicro с Cloudera с RedHat	SSG-6019P-ACR12L	24	35 328,77 лв.	847 890,48 лв. ✓
2	Основен комуникационен модул – Super Micro	SSE-X3348S	3	17 588,62 лв.	52 765,86 лв. ✓
3	Допълнителен комуникационен модул - Allied Telesis	AT-x510L-28GT	1	12 288,27 лв.	12 288,27 лв. ✓
4	KVM switch	KVM	1	10 793,45 лв.	10 793,45 лв. ✓
5	UPS система модел SOCOMEC	MODULYS	1	35 088,50 лв.	35 088,50 лв. ✓



„Консорциум Биг Дейта Системс“ ДЗЗД

Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

Управляващата специализирана хардуерна система за обработка на данни					
Ред	Описание	Парт номер	Брой	Ед.стойност без ДДС	Обща стойност без ДДС
1	МодулNameNode модел SuperMicro с Cloudera с RedHat	SYS-1029P-WTR	3	24 633,18 лв.	73 899,54 лв. ✓
2	МодулManagementNode модел SuperMicro с RedHat	SYS-1029P-WTR	1	21 365,53 лв.	21 365,53 лв. ✓
3	Сървър за УАЛС модел SuperMicro с RedHat	SSG-6019P-ACR12L	1	54 036,65 лв.	54 036,65 лв. ✓
4	Сървър за МУД модел с RedHat	SYS-1029P-WTR	1	21 365,53 лв.	21 365,53 лв. ✓
5	Основен комуникационен модул – Super Micro модел	SSE-X3348S	1	17 588,62 лв.	17 588,62 лв. ✓
6	Допълнителен комуникационен модул - Allied Telesis	AT-x510L-28GT	1	12 288,27 лв.	12 288,27 лв. ✓
7	KVM switch	KVM	1	10 793,45 лв.	10 793,45 лв. ✓
8	UPS система модел GTEC AP160N	GTEC AP160N	1	12 051,43 лв.	12 051,43 лв. ✓
Комуникационна Специализирана хардуерна система за обмен на данни.					
Ред	Описание	Парт номер	Брой	Ед.стойност без ДДС	Обща стойност без ДДС
1	Централен комуникационен модел – Super Micro	SSE-C3632S	1	31 065,53 лв.	31 065,53 лв. ✓
2	Допълнителен комуникационен модул - Allied Telesis	AT-x510L-28GT	1	12 288,27 лв.	12 288,27 лв. ✓
3	Граничен изчислителен модул Edge Node – SuperMicro с RedHat	SYS-1029P-WTR	1	17 743,52 лв.	17 743,52 лв. ✓
4	Изчислителен сървър за LDAP – SuperMicro модел с RedHat	SYS-1029P-WTR	1	16 743,52 лв.	16 743,52 лв. ✓
5	KVM switch	KVM	1	10 793,45 лв.	10 793,45 лв. ✓
6	UPS система модел GTEC AP160N	GTEC AP160N	1	12 051,43 лв.	12 051,43 лв. ✓



„Консорциум Биг Дейта Системс“ ДЗЗД

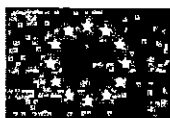
Хай Тех Бизнес Център
ж.к. Дружба 1, ул. Мюнхен 8
1592 София, РОВ 74
Tel.: +359 2 42 10 10

Ред	Описание	Парт номер	Брой	Ед.стойност без ДДС	Обща стойност без ДДС
1	Red Hat Enterprise Linux Server Entry Level, Self-support (2 sokets)	LS221	4	7 313,24 лв.	29 252,96 лв. ✓
2	Интегрирана система за информационна сигурност на Специализираната централизирана система под ключ Hadoop с компоненти: - Sentinel Enterprise 100 EPS / 50 Devices License - network bundle Network Access Control (NAC) - IPS устройство WatchGuard M4600 + WatchGuard Firebox M 2 Port 40Gb QSFP+ Fiber Module - Централизирано автентикиране на потребителите, Създаване права на достъп, Контрол на достъпа, Лист за управление, Пълен одит, Защита и криптиране на данните, Управление на ключовете.	Интеграция Конфигурация	1	725 258,55 лв.	725 258,55 лв. ✓
3	Интегриране на Специализираната централизирана система под ключ с Майкрософт SQL Server	Интеграция	1	400 410,00 лв.	400 410,00 лв. ✓
4	Концептуална архитектура (КА)за разширение на системата Hadoop за хиляди PetaBytes	КА	1	227 517,18 лв.	227 517,18 лв. ✓
5	Обучение	Обучение	1	45 000,00 лв.	45 000,00 лв. ✓
	Обща стойност без начислен ДДС:				3 187 007,00 лв.

*Задължаваме се да доставим всички изброени в техническото предложение като носещи плочи, както и да доставим и разположим цялата техника точно, така както е описано в техническото предложение.

**Цялата предложена техника, както и софтуерното осигуряване ще бъде доставено с гаранционна поддръжка за 5 години, както е описано в Документацията.





ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ



НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

Образец №8

СПИСЪК НА ПЕРСОНАЛА, КОЙТО ЩЕ ИЗПЪЛНЯВА ПОРЪЧКАТА, И/ИЛИ НА ЧЛЕНОВЕТЕ НА РЪКОВОДНИЯ СЪСТАВ, КОЙТО ЩЕ ОТГОВАРЯТ ЗА ИЗПЪЛНЕНИЕТО (ЧЛ. 64, АЛ. 1, Т. 6 ОТ ЗОП)

До:

УНИВЕРСИТЕТ ЗА НАЦИОНАЛНО И СВЕТОВНО СТОПАНСТВО

Долуподписаният/-ната/ Деян Йовков, в качеството ми на Управител (*посочва се длъжността и качеството, в което лицето има право да представлява и управлява - напр. изпълнителен директор, управител или др.*) на "КОНСОРЦИУМ БИГ ДЕЙТА СИСТЕМС" ДЗЗД, ЕИК/ БУЛСТАТ/ друга идентифицираща информация (В съответствие със законодателството на държавата, в която участникът е установен) ЕИК177357133, със седалище и адрес на управление ж.к. ДРУЖБА ул. МЮНХЕН №8, София 1528 – участник, избран за **ИЗПЪЛНИТЕЛ** в процедура за възлагане на обществена поръчка с предмет:

„Доставка, инсталиране, конфигуриране, тестване и поддържане на СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ НАDOOP за работа с Големи данни по проект „Дигитализация на икономиката в среда на Големи данни“ (ДИГД), № BG05M2OP001-1.002-0002“

заявяваме:

1. Експертите, с които предлагаме да изпълним обществената поръчка в съответствие с изискванията на възложителя, са:

Позиция, за която експертът отговаря спрямо изискванията на Възложителя	Експерт (трите имена)	образование/професионална компетентност (степен, специалност, номер на диплома)	Общо професионален опит и специфичен професионален опит по специалността

В настоящата документацията, не се съдържат изисквания, насочващи към определен производител или доставчик. В случай, че наименованието или част от наименованието на съпада с конкретен стандарт, спецификация, техническа оценка, техническо, одобрение, технически еталон и модел, източник, процес, търговска марка, патент, тип, производство, да се приема, че възложителят е поставил изискването "или еквивалентно/и".



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ

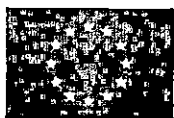


НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

Специалист по продажби и управление на проекти
Експерт информационни технологии хардуерна поддръжка

В настоящата документацията, не се съдържат изисквания, насочващи към определен производител или доставчик. В случай, че наименование или част от наименование на съпада с конкретен стандарт, спецификация, техническа оценка, техническо, одобрение, технически еталон и модел, източник, процес, търговска марка, патент, тип, про изводство, да се приема, че възложителят е поставил изискването "или еквивалентно/и".



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ



ОПЕРАТИВНА ПРОГРАМА
НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

Старши системен анализатор
Експерт информационни технологии
Експерт информационни технологии
Експерт информационни технологии и управление на проекти

В настоящата документацията, не се съдържат изисквания, насочващи към определен производител или доставчик. В случай, че наименование или част от наименование на съпада с конкретен стандарт, спецификация, техническа оценка, техническо одобрение, технически еталон и модел, източник, процес, търговска марка, патент, тип произход/ци производство, да се приема, че възложителят е поставил изискването "или еквивалентно/и".



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ

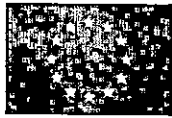


МИНИСТЕРСТВО НА
НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големите данни (ДИГД)“

Специалист по мрежови технологии
Мениджър на системна инфраструктура

В настоящата документацията, не се съдържа изисквания, насочващи към определен производител или доставчик. В случай, че наименования или част от наименования съвпада с конкретен стандарт, спецификация, техническа оценка, техническо, или технически еталон и модел, източник, процес, търговска марка, патент, тип, или производство, да се приема, че възложителят е поставил изискването "или еквивалентно/и".



ЕВРОПЕЙСКИ СЪЮЗ
 ЕВРОПЕЙСКИ
 СОБИРАТЕЛНИ ФОНД
 ЕВРОПЕЙСКИ ФОНД
 ЗА РЕГИОНАЛНО РАЗВИТИЕ



НАУКА И ОБРАЗОВАНИЕ ЗА
 ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

Специалист Linux Essentials
Специалист развойна дейност
Експерт Информационни технологии

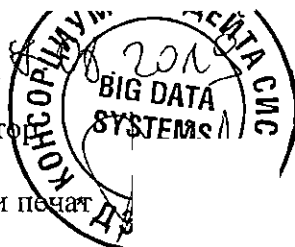
2. През целия период на изпълнение на обществената поръчка, ако същата ни бъде възложена, ще осигурим участие на посочените по-горе експерти.

Известна ми е и намерението в чл. 313 от Наказателния кодекс отговорност за вписване в настоящата декларация.

дата: 28

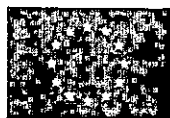
Декларатор

Подпис и печат



Забележка: ОБРАЗЕЦЪТ СЕ ПРЕДСТАВЯ САМО ОТ УЧАСТНИКА, ИЗБРАН ЗА ИЗПЪЛНИТЕЛ, НА ЕТАП СКЛЮЧВАНЕ НА ДОГОВОР!

В настоящата документацията, не се съдържа изисквания, насочващи към определен производител или доставчик. В случай, че наименование или част от наименование на съпада с конкретен стандарт, спецификация, техническа оценка, техническо, одобрение, технически еталон и модел, източник, процес, търговска марка, патент, тип, произход или производство, да се приема, че възложителят е поставил изискването "или еквивалентно/и".



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ



НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-С 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

Образец №8

СПИСЪК НА ПЕРСОНАЛА, КОЙТО ЩЕ ИЗПЪЛНЯВА ПОРЪЧКАТА, И/ИЛИ НА ЧЛЕНОВЕТЕ НА РЪКОВОДНИЯ СЪСТАВ, КОЙТО ЩЕ ОТГОВАРЯТ ЗА ИЗПЪЛНЕНИЕТО (ЧЛ. 64, АЛ. 1, Т. 6 ОТ ЗОП)

До:

УНИВЕРСИТЕТ ЗА НАЦИОНАЛНО И СВЕТОВНО СТОПАНСТВО

Долуподписаният/-ната/ Деян Йовков, в качеството ми на Управител (*посочва се длъжността и качеството, в което лицето има право да представлява и управлява - напр. изпълнителен директор, управител или др.*) на СЕКЮРИТИ СОЛЮШЪНС ИНСТИТУТ ЕООД, ЕИК 200149195, съдружник в "КОНСОРЦИУМ БИГ ДЕЙТА СИСТЕМС" ДЗЗД, ЕИК/ БУЛСТАТ/ друга идентифицираща информация (В съответствие със законодателството на държавата, в която участникът е установен) ЕИК177357133, със седалище и адрес на управление ж.к. ДРУЖБА ул. МЮНХЕН №8, София 1528 – участник, избран за **ИЗПЪЛНИТЕЛ** в процедура за възлагане на обществена поръчка с предмет:

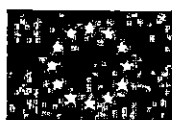
„Доставка, инсталиране, конфигуриране, тестване и поддържане на СПЕЦИАЛИЗИРАНА ЦЕНТРАЛИЗИРАНА СИСТЕМА ПОД КЛЮЧ НАDOOP за работа с Големи данни по проект „Дигитализация на икономиката в среда на Големи данни“ (ДИГД), № BG05M2OP001-1.002-0002“

заявяваме:

1. Експертите, с които предлагаме да изпълним обществената поръчка в съответствие с изискванията на възложителя, са:

Позиция, за която експертът отговаря спрямо изискванията на Възложителя	Експерт (трите имена)	образование/професионална компетентност (степен, специалност, номер на диплома)	Общо професионален опит и специфичен професионален опит по специалността

В настоящата документацията, не се съдържат изисквания, насочващи към определен производител или доставчик. В случай, че наименованието или част от наименованието на съпада с конкретен стандарт, спецификация, техническа оценка, техническо, одобрение, технически еталон и модел, източник, процес, търговска марка, патент, тип, ли производство, да се приема, че възложителят е поставил изискването "или еквивалентно/и".



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ



ОПЕРАТИВНА ПРОГРАМА
НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

Специалист по продажби и управление на проекти
Експерт информационни технологии хардуерна поддръжка

В настоящата документацията, не се съдържат изисквания, насочващи към определен производител или доставчик. В случай, че наименование или част от наименование на съпада с конкретен стандарт, спецификация, техническа оценка, техническо, одобрение, технически еталон и модел, източник, процес, търговска марка, патент, тип, или производство, да се приема, че възложителят е поставил изискването "или еквивалентно/и".



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ



НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

Старши системен анализатор		
Експерт информационни технологии		
Експерт информационни технологии		
Експерт информационни технологии и управление на проекти		

В настоящата документацията, не се съдържат изисквания, насочващи към определен производител или доставчик. В случай, че наименование или част от наименование на съпада с конкретен стандарт, спецификация, техническа оценка, техническо, одобрение, технически еталон и модел, източник, процес, търговска марка, патент, тип, произход или производство, да се приема, че възложителят е поставил изискването "или еквивалентно/и".



ЕВРОПЕЙСКИ СЪЮЗ
 ЕВРОПЕЙСКИ
 СОЦИАЛЕН ФОНД
 ЕВРОПЕЙСКИ ФОНД
 ЗА РЕГИОНАЛНО РАЗВИТИЕ

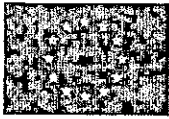


НАУКА И ОБРАЗОВАНИЕ ЗА
 ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

Специалист по мрежови технологии		
Мениджър на системна инфраструктура	1 1	

В настоящата документацията, не се съдържат изисквания, насочващи към определен производител или доставчик. В случай, че наименование или част от наименование на съпада с конкретен стандарт, спецификация, техническа оценка, техническо, одобрение, технически еталон и модел, източник, процес, търговска марка, патент, тип, произход или производство, да се приеме, че възложителят е поставил изискването "или еквивалентно/и".



ЕВРОПЕЙСКИ СЪЮЗ
ЕВРОПЕЙСКИ
СОЦИАЛЕН ФОНД
ЕВРОПЕЙСКИ ФОНД
ЗА РЕГИОНАЛНО РАЗВИТИЕ



НАУКА И ОБРАЗОВАНИЕ ЗА
ИНТЕЛИГЕНТЕН РАСТЕЖ

Процедура BG05M2OP001-1.002 „Изграждане и развитие на центрове за компетентност“, Оперативна програма „Наука и образование за интелигентен растеж“ 2014 – 2020. Проект BG05M2OP001-1.002-0002-C 01 „Дигитализация на икономиката в среда на Големи данни (ДИГД)“

Специалист Linux Essentials
Специалист развойна дейност
Експерт Информационни технологии

2. През целия период на изпълнение на обществената поръчка, ако същата ни бъде възложена, ще осигурим участие на посочените по-горе експерти.

Известна ми е предвидената в чл. 313 от Наказателния кодекс отговорност за вписване на неверни данни в настоящата декларация.

дата: 30.08.2018

Декларатор:

Подпис и печат



Забележка: ОБРАЗЕЦЪТ СЕ ПРЕДСТАВЯ САМО ОТ УЧАСТНИКА, ИЗБРАН ЗА ИЗПЪЛНИТЕЛ, НА ЕТАП СКЛЮЧВАНЕ НА ДОГОВОР!

В настоящата документацията, не се съдържа изисквания, насочващи към определен производител или доставчик. В случай, че наименование или част от наименование на съпада с конкретен стандарт, спецификация, техническа оценка, техническо, одобрение, технически еталон и модел, източник, процес, търговска марка, патент, тип, произход или производство, да се приема, че възложителят е поставил изискването "или еквивалентно/и".

ДИПЛОМА



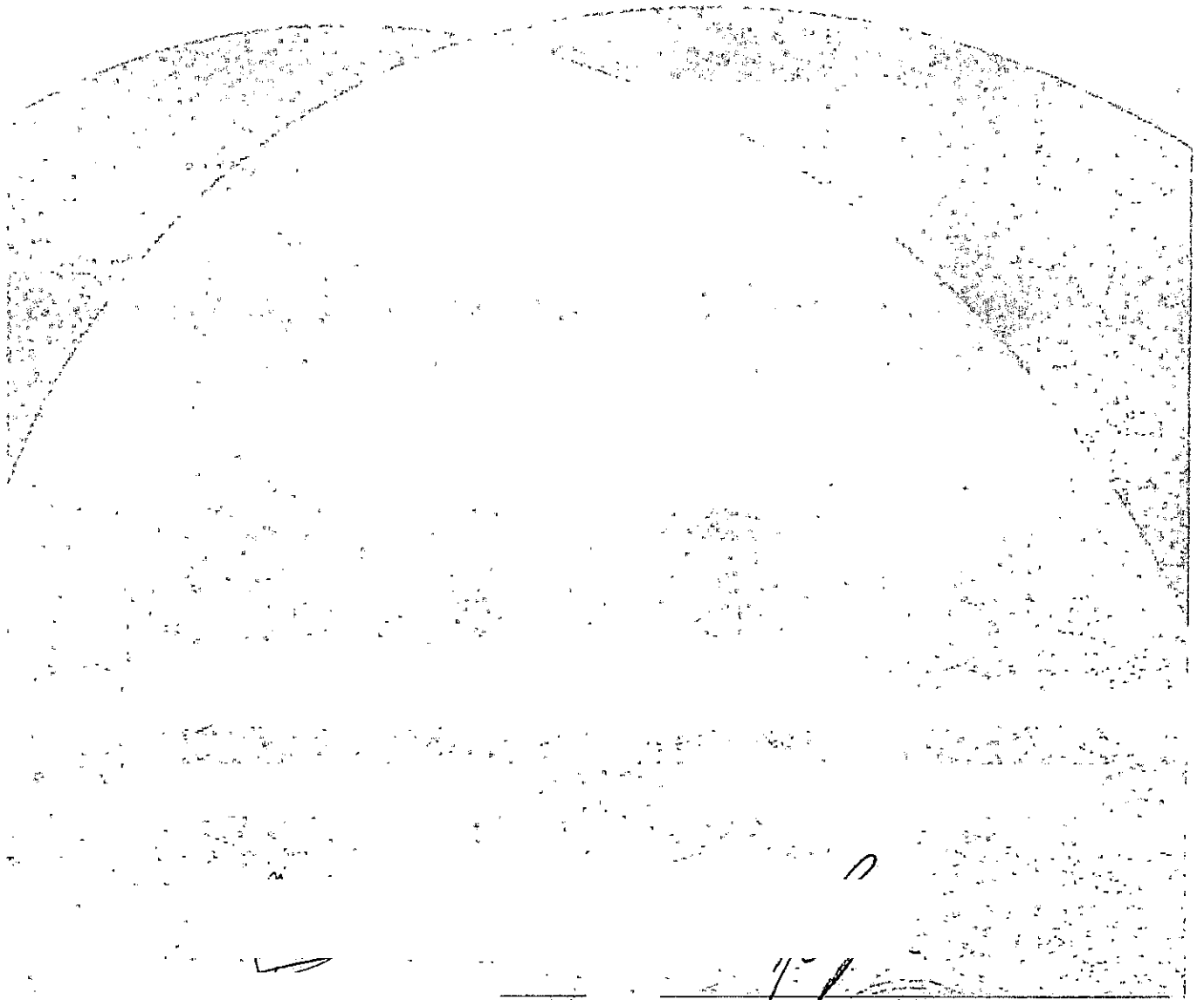
РЕПУБЛИКА
БЪГАРИЯ

ДИПЛОМА
— ЗА —
ВИСШЕ ОБРАЗОВАНИЕ

Ярн

PEOPLECERT

AXELOS
GLOBAL BEST PRACTICE



Вярно с оригинала

Printed on 24 October 2016

This certificate remains the property of the issuing Examination Institute and shall be returned immediately upon request.

QITIL' QRESILIA' QPRINCE2' QPRINCE2 AGILE QMSP' QMoR' QP3M3' QP30' QMoP' QMoV'

AXELOS, the AXELOS logo, the AXELOS swirl logo, ITIL, PRINCE2, PRINCE2 AGILE, MSP, MoR, P3M3, P30, MoP and MoV are registered trade marks of AXELOS Limited. RESILIA is a trade mark of AXELOS Limited.

Certification Number
ECC2301957468

C|CISO
Certified Chief Information Security Officer

Certified Chief Information Security Officer



#0732
ISO/IEC 17024
Personnel Certification Program

EC-Council

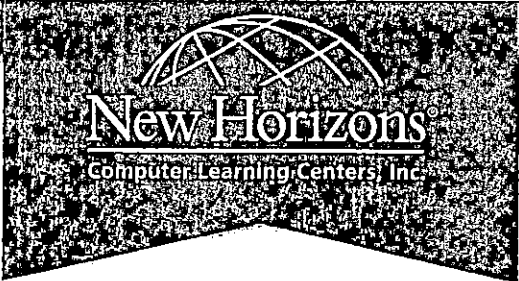
Sanjay Bavisj, President



CERTIFICATE *of* COMPLETION

1/10

UPGRADING PEOPLE EVERY DAY

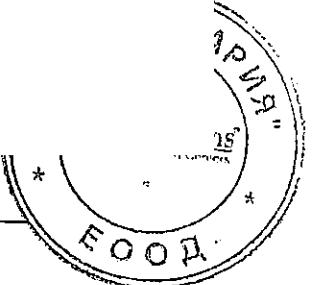


CERTIFICATE *of* COMPLETION

Вяр.

4

INSTRUCTOR



UPGRADING PEOPLE EVERY DAY



ДИПЛОМА
ЗА ВИСШЕ ОБРАЗОВАНИЕ

© Республиканское БУБ БГУ, сохранены дизайн и содержание. 2003

© Издательство «Классика», 2003



ДИПЛОМА
ЗА ВУСШЕ
ОБРАЗОВАНИЕ





WE KEEP IT GREEN.™

Supermicro - Certified Engineer

U.S. Patent # 7,700,000

Intel® Technology Expert



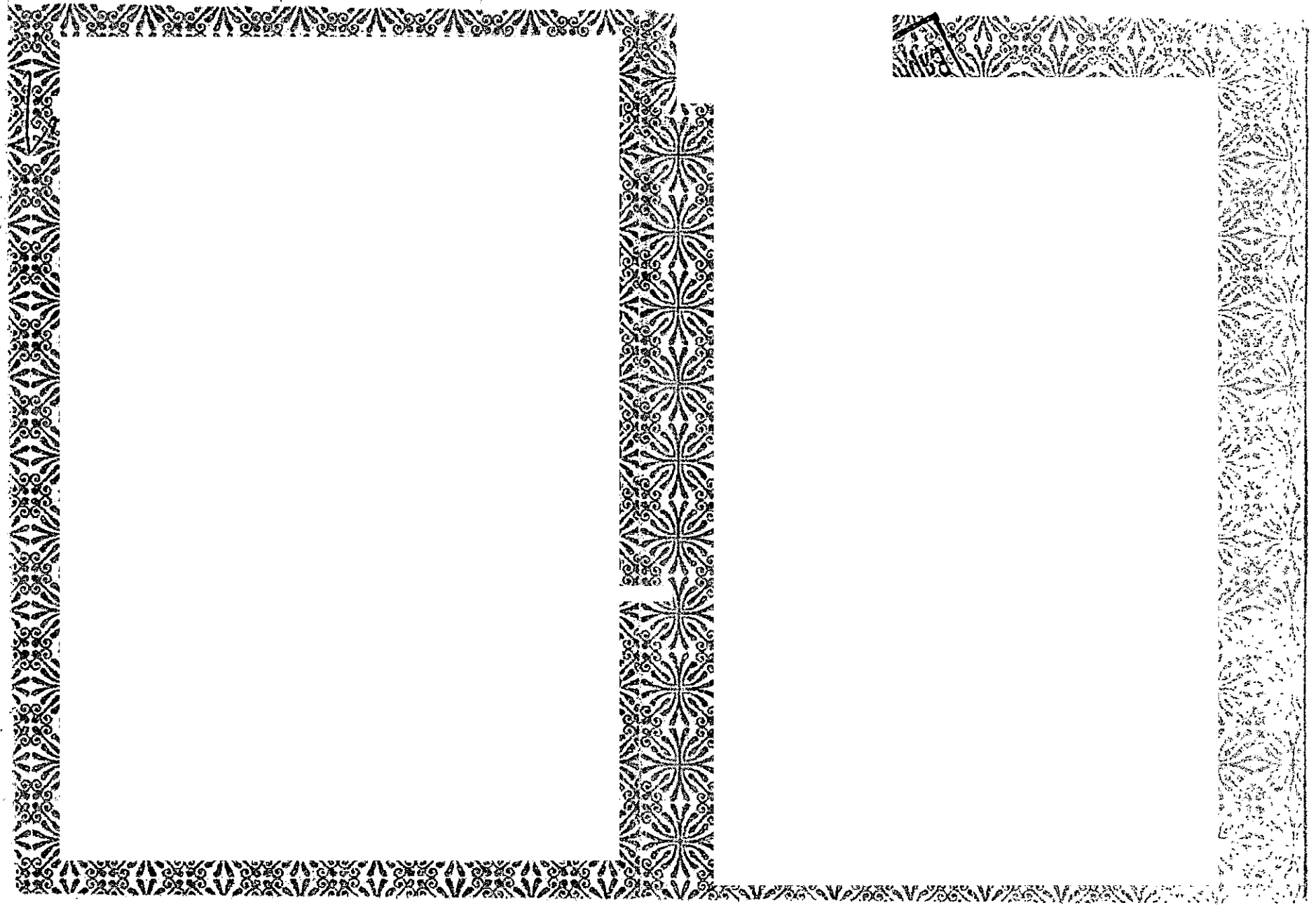
The Intel® Technology Provider Program Recognizes

Met
Vice President
Director

our solutions. Our technology. Smarter together.



Вярно с оригинала





Microsoft Certified Solutions Associate

Satya Nadella
Chief Executive Officer

Solutions Associate

Microsoft Certificate of Excellence

Microsoft
C E R T I F I E D
IT Professional

Database Developer, 2008

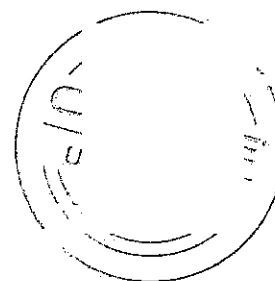
Microsoft
CERTIFIED

ORACLE®

Certified Expert

THIS CERTIFIES THAT

Senior Vice President, Oracle Corporation



AMITOMMA

ПРИЛОЖЕНИЕ
КЪМ ДИПЛОМА ЗА ВИСШЕ ОБРАЗОВАНИЕ

PEOPLECERT

 AXELOS
GLOBAL BEST PRACTICE

This is to certify that





Effe

Cer

Pete






Printed on 23 May 2010

This certificate remains the property of the issuing Examination Institute



returned immediately upon request.

ВІСНОВ

ДИПЛОМА
ЗА ВИСШЕ ОБРАЗОВАНИЕ

НО С О Р И

Δ

PeopleCert®

All talents, certified.

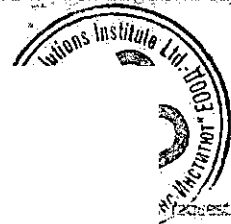
AXELOS

GLOBAL BEST PRACTICE

CEO
Peter Hepworth, CEO, AXELOS

11/11
Panorala Theeriti, Certification Qualifier, PeopleCert

ВЯНО С ОФИЦИАЛНА



Printed on 19 November 2018

This certificate remains the property of the issuing Examination Institute and sh

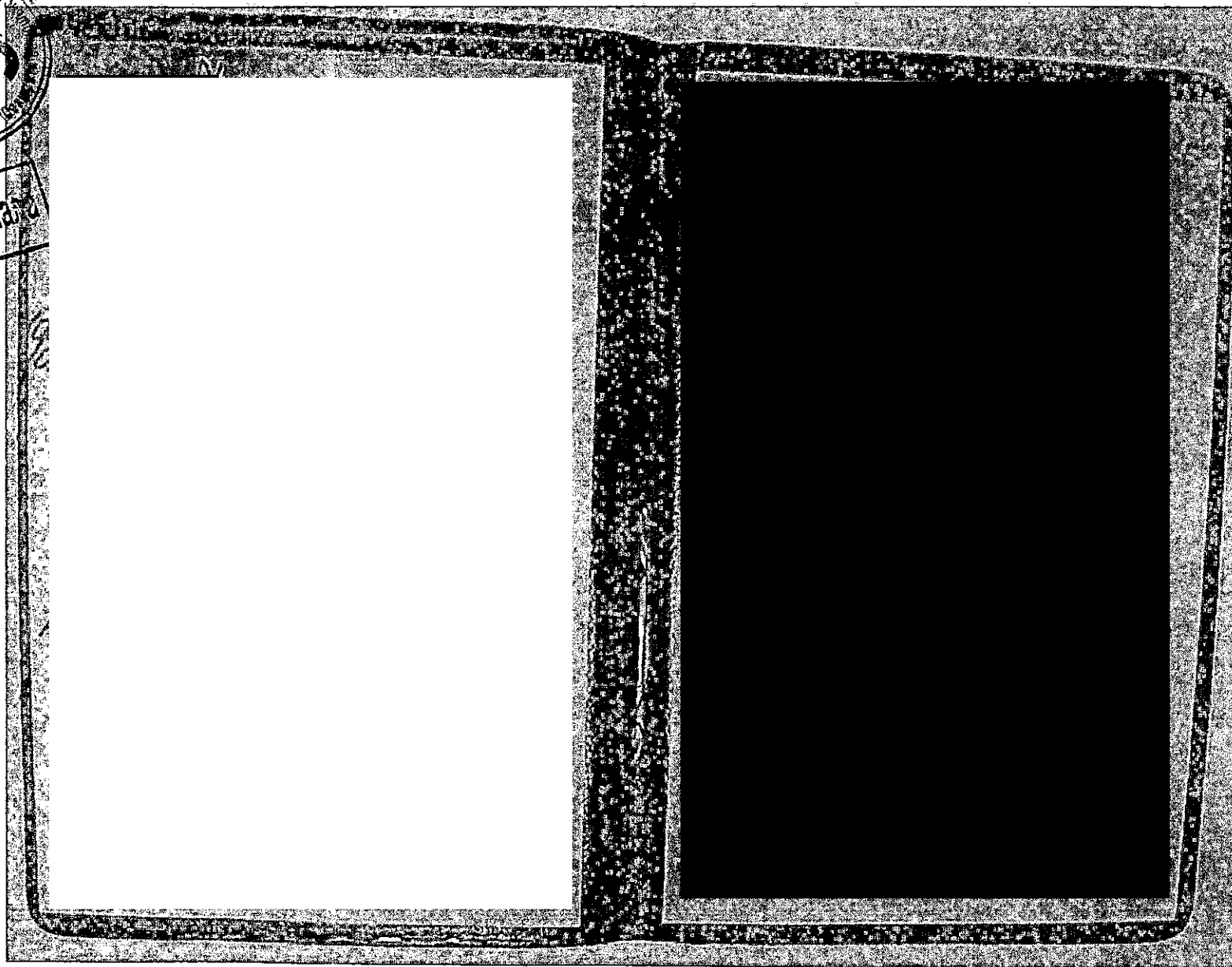
ITIL RESILIA PRINCE2 FRINCE2 MSF

PSO MoP MSV

AXELOS, the AXELOS logo, the AXELOS wordmark, the AXELOS logo and the AXELOS wordmark are trademarks of AXELOS Limited or AXELOS Group of Companies. All other trademarks are the property of their respective owners.



Вярно с оригинала



SvetlaTex

TRANSLATION CENTRE

ЕТ "СВЕЛАТЕКС – СВЕЛАНА ПЕТРОВА - АЛХУСАМИ"

Ул. "Искър" №11, 1202 София, тел./факс 02 980 02 02; мобилен: 0877829091

11 Iskar St., 1202 Sofia, Bulgaria, tel./fax: (+359)2 980 02 02, mobile phone: (+359) 877829091

e-mail: svetlatex@yahoo.com

Превод от английски език



ITIL
PRINCES2
MSP
M_o_R
P3m3
P3O
MoP
MoV



ДИПЛОМА
ЗА ВИСШЕ ОБРАЗОВАНИЕ



ДИПЛОМА
ЗА ВИСШЕ
ОБРАЗОВАНИЕ

подпис и печат на АУБ



CompTIA

VERIFIED

СЕРТИФИКАТ С ОРИГИНАЛОМ



Code: SHJL76WMFLE1QPPH
Verify at: <http://verify.CompTIA.org>

WatchGuard Technical Certification



macmon
nac ■ intelligent einfach

ZERTIFIKAT

Mit dem E
dass die ir
un

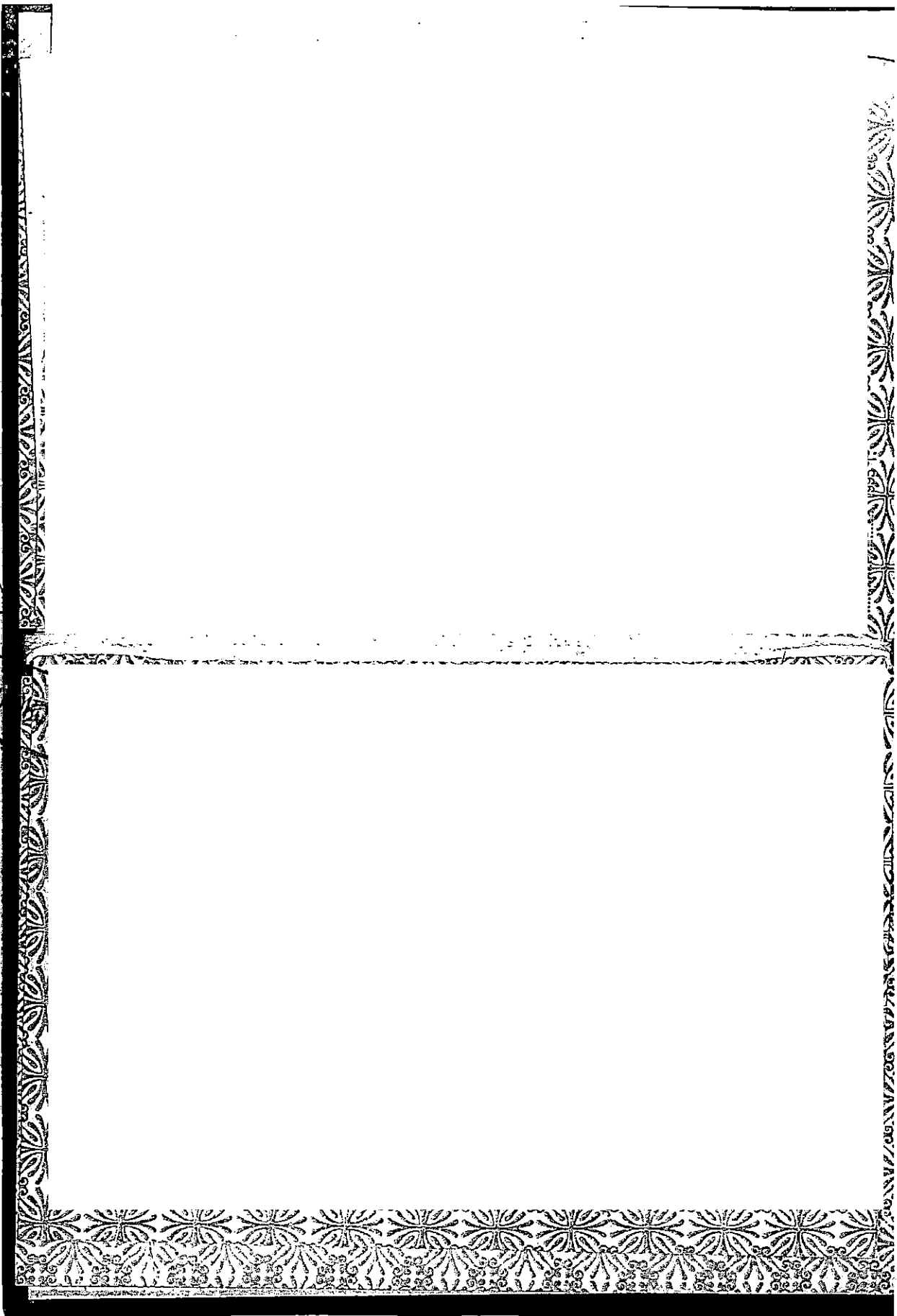
rden,
rden





[Handwritten signature]

ОПТИМАЛ



WatchGuard Technical Certification



CERTIFICATE *of* COMPLETION

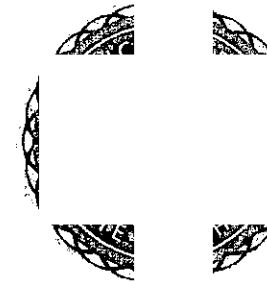


UPGRADING PEOPLE EVERY DAY

MICROSOFT CERTIFICATE OF

Achievement

This certificate accredits that



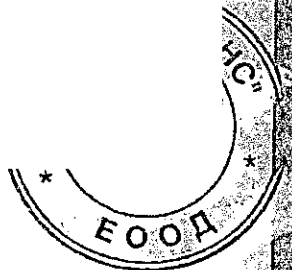
VC
Officer
Вярно с оригинала

Certificate of Completion

[Blank area for certificate details]

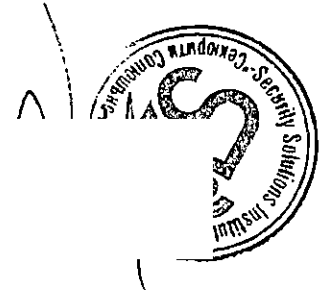
Bryan

Computer Learning Centers
CHOOSE. LEARN. SUCCEED.®



001111001
001111001
001111001

ДИПЛОМА
ЗА ЗАВЪРШЕНО ВИСШЕ ОБРАЗОВАНИЕ



РЕПУБЛИКА БЪЛГАРИЯ

ДИПЛОМА

ЗА ЗАВЪРШЕНО ВИСШЕ ОБРАЗОВАНИЕ



cloudera®

Certificate of Completion

Вярно с оригинала



НАРОДНА РЕПУБЛИКА БЪЛГАРИЯ

ДИПЛОМА
ЗА ВИСШЕ ОБРАЗОВАНИЕ

6

