

УДК 656.61:681.3

РОЗВИТОК ХМАРНИХ ТЕХНОЛОГІЙ В КОНТЕКСТІ УПРАВЛІННЯ ВЕЛИКИМИ ДАНИМИ В МАШИНОБУДУВАННІ

Гармаш О.М.¹, к.е.н., доц., Марчук В.Є.¹, к.т.н., доц., Безрека М.¹ асп.,
Градиський Ю.О.², к.т.н., доц.

¹Національний авіаційний університет, м. Київ,

²Харківський національний технічний університет сільського господарства імені
Петра Василенка)

Проаналізовано тенденції розвитку хмарних технологій у контексті управління великими даними для обробки та аналізу безперервного потоку неструктурованих даних з корисною інформацією в режимі реального часу. Показано перспективність використання даної технології в галузі машинобудування, логістики, освітянської діяльності та інших секторах економіки.

Ключові слова: хмарні технології, управління великими даними, інформаційні системи, інформаційні технології, бази даних.

Загальна постановка проблеми. З кожним роком з'являються більш досконалі методи збору інформації, як приклади джерел виникнення великих даних можна привести системи, з яких безперервно надходять дані з вимірювальних пристроїв: події від радіочастотних ідентифікаторів RFID, дані про місцезнаходження абонентів мереж стільникового зв'язку та ін. Відзначимо, що розвиток і початок широкого використання цих і подібних джерел ініціює проникнення технологій великих даних як в науково-дослідну діяльність, так і в машинобудуванні, комерційному секторі так і в сферах з державного управління, яка реалізується за допомогою хмарних технологій.

Огляд публікацій та аналіз невирішених проблем. Хмарні обчислення – це модель забезпечення зручного доступу до мережі по вимогам до загальних обчислень ресурсів, що конфігурують (наприклад, мережі передачі даних, серверам, пристроям зберігання даних, додаткам та сервісам), які можуть бути оперативно надані з мінімальними зусиллями зі сторони користувача або до постачальника послуг [1].

З часом область застосування хмарних технологій суттєво розширилася. Причиною цього став стрімкий розвиток комп'ютерних потужностей і ліній зв'язку. Хмарні технології почали застосовуватися в області біології [2], реалізації проектів автоматизації закладів охорони здоров'я [3], в навчальному процесі сучасної школи [4], самостійній роботі студентів [5], в процесі підготовки фахівців з фізичного виховання і спорту [6].

Еволюція технологій хмарних обчислень описується моделлю Cloud Computing Maturity Model (ССММ), яка включає в себе наступні рівні розвитку, представлені на рис. 1.

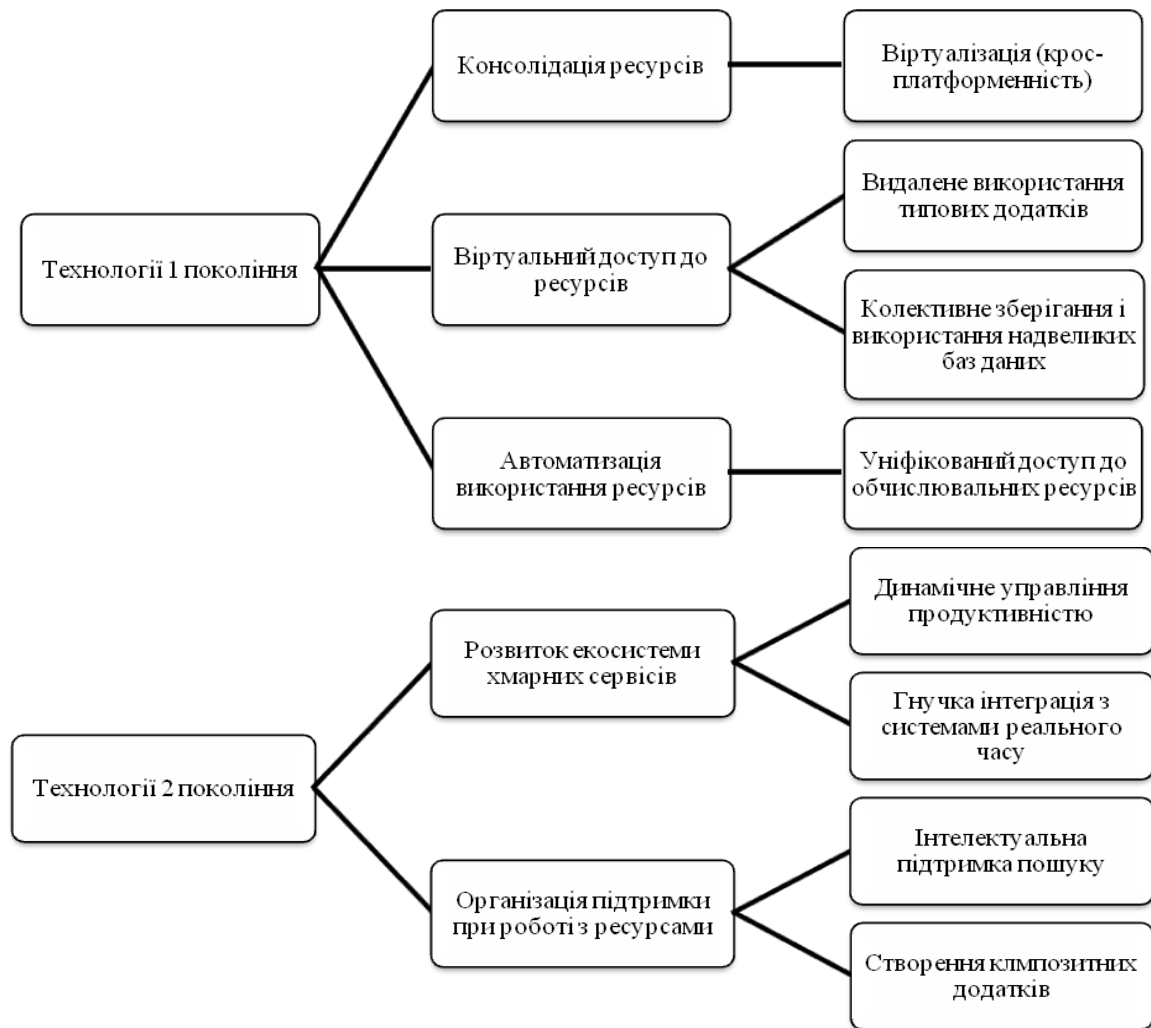


Рисунок 1. Еволюція технологій хмарних обчислень в моделі SSMM [7]

Технології 1 покоління повністю реалізовані і продовжують вдосконалюватися, а технології 2 покоління знаходяться на стадії розробок окремих пропозицій і платформ.

Як і в будь-якій технології, хмарні технології мають як переваги, так і недоліки, які представлені в табл. 1.

Оцінюючи економічну значимість різноманітних моделей розгортання хмарних обчислень для європейської економіки, дослідники СЕВР стверджують, що 39,3% економічних вигод приходить на гібридну модель, 35,1% - на приватні хмари і лише 25,6% на публічні хмарні сервіси [8].

Крім економічної вигоди хмарні обчислення ще і достатньо екологічні. Впровадження хмарних технологій не тільки дозволить зекономити на задіяному обладнанні, але і підвищить ефективність використання числа обчислювальних потужностей з розрахунку на кіловат-годину. В найближчому майбутньому основна частина отриманих даних буде збиратися і оброблятися кінцевими вбудованими інтелектуальними приладами (спеціалісти агентства

VDCResearchGroupInc називають їх ScalableEdgeNodes – SEN), об'єднаними в диференційну розумну мережу IP [8].

Таблиця 1. Переваги та недоліки хмарних технологій [7]

Хмарні технології	
Переваги	Недоліки
Доступ – в будь-якому місці в будь-який час, де є інтернет.	Постійне з'єднання з мережею – для роботи з «хмарою» необхідне постійне підключення до мережі.
Низька вартість – зниження витрат на обслуговування (використання технологій віртуалізації).	Програмне забезпечення – користувачу доступні тільки ті ПЗ, які є в «хмарі».
Гнучкість – необмеженість обчислювальних ресурсів.	Конфіденційність – в даний час немає технології, яка забезпечувала б конфіденційність на 100%.
Надійність – наявність додаткових джерел живлення, регулярне резервування даних, висока пропускну здатність інтернет-канала.	Надійність – втрата інформації в «хмарі» означає неможливість її відновлення.
Безпека – високий рівень безпеки при точній організації процесу.	Безпека – хоча «хмара» є достатньо надійною системою, але 100% захист від злону немає.
Великі обчислювальні потужності.	Вартісне обладнання – для створення своєї «хмари» необхідні значні матеріальні ресурси.

Компанії в будь-якій галузі докладають зусилля щоб приймати ефективні бізнес-рішення для власного просування в межах даного ринку. Існують різні дані, моніторинг яких допомагає підприємствам досягти успіху – очікувані обсяги продажів, переваги товарів серед аналогів на ринку, оптимізовані робочі графіки тощо. Пошук значимої інформації дуже складний через те, що наявне накопичення великих обсягів даних всередині компанії і в Інтернеті.

Як показує практика, відповідно до потреб сучасних умов праці зручнішим за локальне редагування документу є розміщення необхідного файлу в хмарне сховище, доступ до котрого може бути розмежований для ролей конкретних користувачів. Одні користувачі можуть змінювати файл, інші - тільки читати та рецензувати зміни. Загалом використання таких хмарних сервісів є простим у користуванні та не потребує особливих налаштувань.

Відзначимо, що за останні десятиліття змістовних робіт з питань обробки та аналізу неструктурованих даних з корисною інформацією не було, а те, що ми зазвичай називаємо теорією інформації Клода Шеннона, є статистичною теорією передачі сигналів, і до інформації, яка сприймається людиною, не має

відношення. Тому ігнорування ролі даних і інформації, як предметів дослідження, послужило виникненням потреби, коли обчислювальні можливості комп'ютерів виявилися набагато меншими, ніж швидкість надходження інформації та неструктурованість даних унеможливило проведення аналізу існуючими методами.

Ще у 2008 році, кількість доступних цифрових інформаційних частин, так званих, бітів, перевищила число зірок у Всесвіті, згідно «The Diverse and Exploding Digital Universe» [1]. Сьогодні «цифровий світ» розширюється зі швидкістю, яка подвоює обсяг даних кожні 2 роки (рис. 2) [9, 10].

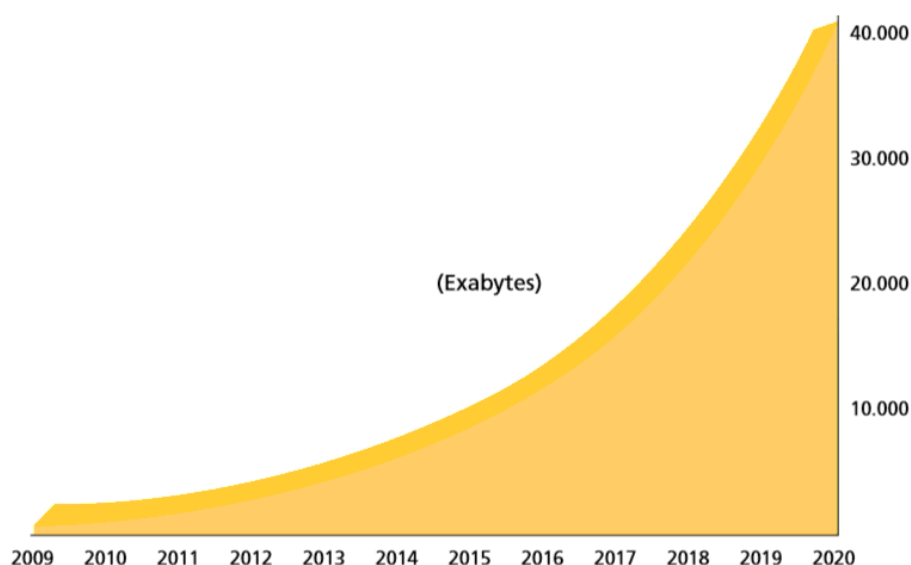


Рисунок 2. Експоненціальне зростання даних між 2010 і 2020 роками

В Україні більшість систем працює в основному зі статичними об'єктами і виконавці оперують категоріями станів. При використанні технології великих даних необхідно працювати з безперервним потоком даних. Технологія великих даних призначена для аналізу будь-яких обсягів інформації в режимі реального часу також з використанням існуючих сучасних хмарних сервісів. Відзначимо, що концепція великих даних – це сукупність підходів, які дозволяють працювати з великими обсягами даних, якими складно або неможливо управляти за допомогою звичайних засобів, тому що вони мають різну структуру і значну швидкість поповнення (табл. 2).

Раніше великі дані розглядали як завдання з причин відсутності місць для їх зберігання, мереж для передачі, потужностей для їх обробки та аналізу. Після появи даних технологій дані заповнили весь наданий їм об'єм, але як би не нарощували потужності з пропускну здатності та місткість засобів зберігання даних, завжди з'являються джерела, наприклад, експерименти з моделювання складних елементів машин і механізмів, які продукують інформації більше, ніж можуть обробити існуючі системи. Основні характеристики великих даних, а саме об'єм, швидкість, різноманіття, достовірність і цінність даних представлено на рис. 3. Згідно із законом Мура, продуктивність сучасних паралельних

обчислювальних систем стабільно зростає, зростають і швидкості мереж передачі даних.

Таблиця 2. Порівняльна таблиця параметрів традиційної бази і бази великих даних [11]

Характеристика	Традиційні бази даних	Великі дані
Об'єм інформації	1×10^9	1×10^{18}
Засоби зберігання	Централізований	Децентралізований
Структурованість даних	Структурована	Частково структурована та неструктурована
Модель зберігання та обробки даних	Вертикальна модель	Горизонтальна модель
Взаємозв'язок даних	Сильний	Слабкий



Рисунок 3. Основні характеристики великих даних [12]

У 2014 р. технології великих даних, стали одними з пріоритетних напрямків інвестування в сфері венчурної індустрії. На думку експертів, пов'язано це з тим, що розробки з даного напрямку почали приносити значні результати для їх користувачів. За минулий рік кількість компаній з реалізованими проектами в сфері управління великими даними збільшилася на 125%, обсяг ринку виріс на 45% в порівнянні з 2014 роком [10].

Таким чином, завдяки технології великих даних у наукового співтовариства і бізнесу з'являється безліч нових можливостей, наприклад, можливість проводити аналіз великих обсягів неструктурованих даних.

Також слід зазначити, що великі дані успішно реалізовані в багатьох проектах зарубіжних компаній, а саме: Nasdaq, Facebook, Google, IBM, VISA, Master Card, Bank of America, HSBC, AT&T, Coca Cola, Starbucks і Netflix. Сфери застосування обробленої інформації різноманітні і варіюються в залежності від галузі та завдань, які необхідно виконати, а саме [10]:

– HSBC використовує технології великих даних для протидії шахрайським операціям з пластиковими картами. За допомогою технологій великих даних компанія збільшила ефективність служби безпеки в 3 рази, розпізнавання шахрайських інцидентів – в 10 разів. Економічний ефект від впровадження даних технологій перевищив 10 млн дол. США.

– Procter&Gamble за допомогою великих даних проектує нові продукти і складають глобальні маркетингові кампанії. P&G створив спеціалізовані офіси Business Spheres, де можна переглядати інформацію в реальному часі. Таким чином, у менеджменту компанії з'явилася можливість миттєво перевіряти гіпотези і проводити експерименти.

– Міністерство праці Німеччини використовує технологію великих даних в роботі, пов'язаній з аналізом заявок на видачу допомоги безробітним. Так, проаналізувавши інформацію, стало зрозуміло, що 20% виплачувалося незаслужено. За допомогою великих даних міністерство праці скоротило витрати на 10 млрд. євро.

– Дитяча лікарня Торонто впровадила проект Project Artemis. Це інформаційна система, яка збирає і аналізує дані по немовлятам в реальному часі. Система щомиті відслідковує 1260 показників стану кожної дитини. Project Artemis дозволяє прогнозувати нестабільний стан дитини і почати профілактику захворювань у дітей.

Інструменти відстеження походження даних раніше були всього лише корисною функцією, тому що більшість даних у інформаційних системах надходило з перевірених сховищ даних. Але в сьогоденні технології великих даних дозволяють визначати походження даних та це стане обов'язковим для сучасних систем, так як користувачі мають справу як з корпоративними даними, так і зі сторонніми. Деякі дані можуть бути дуже високої якості, а інші набори даних можуть бути не ідеальними, але можуть бути придатними для прототипування та корисними для досягнення загального результату.

Також зазначимо, що розширення хмарних сервісів дозволить не тільки збирати дані, а й використовувати їх в аналізі та алгоритмах обробки «великих даних». Сучасні технології з забезпечення високого рівня безпеки хмарних сервісів IoT (Internet of Things) також допоможуть виробникам створювати нові продукти, які зможуть безпечно оперувати з аналізованими даними без втручання людини [13].

В роботах [11, 13] відмічається, що найбільшою популярністю користуються такі технології великих даних, як in-memory платформи компаній SAP, HANA, Oracle та ін. Результати опитування T-Systems показали, що їх обрали 30% опитаних компаній. Другими за популярністю стали NoSQL платформи (18% користувачів), також компанії використовували аналітичні платформи компаній Splunk і Dell, їх вибрало 15% компаній. Найменш корисними для компаній, за результатами опитування виявилися продукти Hadoop / MapReduce.

На думку експертів технології великих даних набули широкого поширення в багатьох галузях та секторах економіки. За даними опитування TechProResearch, проведеного у 2014 році, найбільшого поширення зазначені технології отримали в телекомунікаційній галузі, а також в інжинірингу, ІТ, в фінансових і державних підприємствах. Як видно за результатами даного опитування, вони менш популярні в галузі освіти та охорони здоров'я (рис. 4).

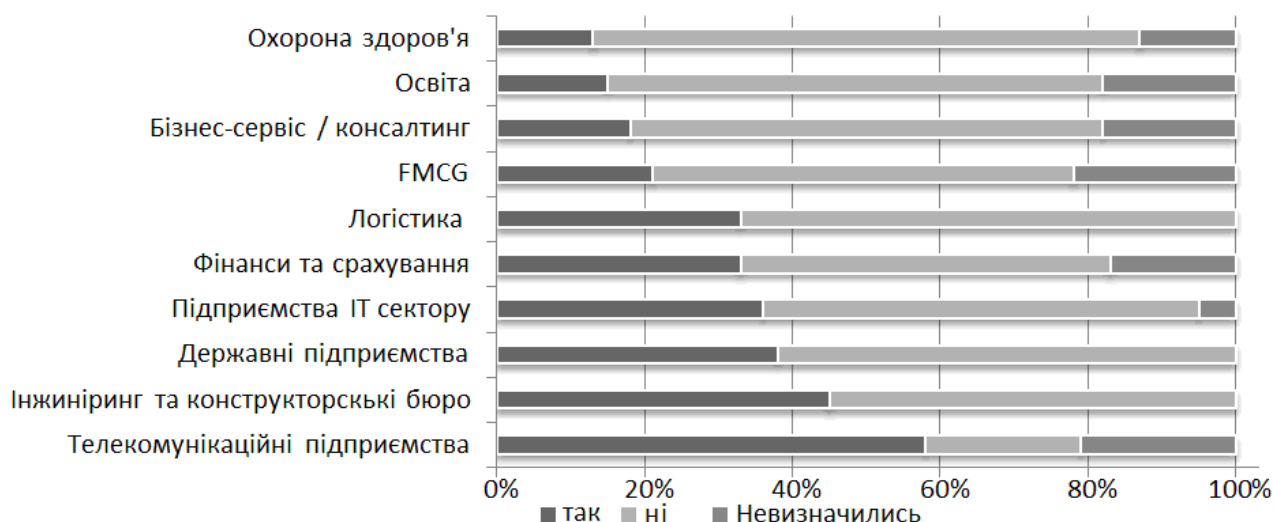


Рисунок 4. Аналіз застосування технологій великих даних за галузями [10]

Згідно з результатами дослідження CNewsAnalyticстаOracle, рівень зрілості ринку великих даних за останній рік підвищився. Респонденти, що представляють 108 великих підприємств з різних галузей, продемонстрували більш високий ступінь поінформованості про ці технології, а також склалося розуміння потенціалу подібних рішень для свого бізнесу.

Станом на 2014 рік, за даними IDC [10], в Україні накопичено приблизно 105 Ексабайт інформації, що становить всього лише 0,8% світових даних. Обсяг інформації до 2020 року може досягти 780 Ексабайт та забере 1,2%. Таким чином, середній темп зростання обсягу інформації прогнозується близько 28% в рік. Прогнозується збереження позитивної динаміки в цьому секторі українського ринку ІТ, навіть в умовах загальної стагнації економіки. Це пов'язано з тим, що бізнес як і раніше демонструє попит на рішення, що дозволяють підвищити ефективність роботи, а також оптимізацію витрат, поліпшення точності прогнозування і мінімізувати можливі ризики компанії.

Висновок. Застосування хмарних технологій з використанням великих даних для обробки та аналізу безперервного потоку неструктурованих даних з корисною інформацією в режимі реального часу є перспективною технологією впровадження в галузі машинобудування, в сфері державного управління, в науково-дослідній діяльності, комерційному секторі тощо. Це свідчить про необхідність проведення більш детальних наукових досліджень в сфері аналізу ринку використання технології великих даних, а також поширення даної технології в інших галузях промисловості, а саме охорони здоров'я, освітянської діяльності, логістики та інших секторах економіки.

Список літератури

1. Пять мифов о технологии облачных вычислений / IHS [Electronicresource]. – Modeofaccess: <http://www.cloudcomputing-news.net/news/2014/apr/08/why-cloud-services-spending-will-exceed-174b-in-2014>.
2. Оплачко Е.С. Облачные технологии и их применение в задачах вычислительной биологии / Е.С. Оплачко, Д.М. Устинин, М.Н. Устинин [Электронный ресурс] – Режим доступа: http://www.matbio.org/2013/Oplachko_8_449.pdf
3. Гусев А.В. Перспективы облачных вычислений и информатизация учреждений здравоохранения / А.В. Гусев // Врач и информационные технологии.–2011.– №2.– С.6-17. [Электронный ресурс]. – Режим доступа: <http://www.fesmu.ru/elib/Article.aspx?id=237446>.
4. Газейкина А.И. Применение облачных технологий в процессе обучения школьников / А.И. Газейкина, А.С. Кувина [Электронный ресурс]. – Режим доступа: <http://cyberleninka.ru/article/n/primenenie-oblachnyh-tehnologiy-v-protseesse-obucheniya-shkolnikov>.
5. Сажко Г.І. Щодо можливості використання хмарних технологій в організації самостійної роботи студентів / Г.І. Сажко, В.І. Шеховцова [Електронний ресурс] – Режим доступу: <http://repo.uipa.edu.ua/jspui/handle/123456789/4549>.
6. Денисова Л.В. Хмарні технології в освітньому процесі вищих навчальних закладів фізичної культури і спорту: стан питання та перспективи застосування / Л.В. Денисова [Електронний ресурс] – Режим доступу: [http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?I21DBN=LINK&P21DBN=UJRN&Z21ID=&S21REF=10&S21CNR=20&S21STN=1&S21FMT=ASP_meta&STR=VchdpuPN_2014_118\(2\)_10](http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?I21DBN=LINK&P21DBN=UJRN&Z21ID=&S21REF=10&S21CNR=20&S21STN=1&S21FMT=ASP_meta&STR=VchdpuPN_2014_118(2)_10).
7. Бурий А.С. Облачные технологии, как механизм распределенной переработки информации / А.С. Бурий [Электронный ресурс]. – Режим доступа: http://iea.gostinfo.ru/files/2014_01/2014_01_01.pdf.
8. Медведев А. Облачные технологи: тенденции развития, примеры исполнения / А. Медведев [Электронный ресурс]. – Режим доступа: <https://issuu.com/cta-mag/docs/20132006>.

9. Big Data in Logistics A DHL perspective on how to move beyond the hype / December 2013 Powered by Solutions & Innovation: Trend Research.
10. IDC's Digital Universe Study, sponsored by EMC, December 2012.
11. Mell P. The NIST Definition of Cloud Computing v15. / P. Mell [Electronic resource]. – <http://www.slideshare.net/crossgov/nist-definition-of-cloud-computing-v15>.
12. Романишина О.Я. Огляд інформаційних технологій та засобів їх реалізації у вищих навчальних закладах / О.Я. Романишина [Електронний ресурс]. – Режим доступу: <http://dspace.uzhnu.edu.ua:8080/jspui/handle/lib/417?mode=simple>.
13. “Trends and Strategies in Logistics and Supply Chain Management”, p. 51, BVL International, 2013.

Аннотация

РАЗВИТИЕ ОБЛАЧНЫХ ТЕХНОЛОГИЙ В КОНТЕКСТЕ УПРАВЛЕНИЯ БОЛЬШИМИ ДАННЫМИ В МАШИНОСТРОЕНИИ

Гармаш О.Н., Марчук В.Е., Безрека М., Градыский Ю.А.

Проанализированы тенденции развития облачных технологий в контексте управления большими данными для обработки и анализа непрерывного потока неструктурированных данных с полезной информацией в режиме реального времени. Показана перспективность использования данной технологии в области машиностроения, логистики, сферы образования и других секторах экономики.

Ключевые слова: *облачные технологии, управление большими данными, информационные системы, информационные технологии, базы данных.*

Abstract

DEVELOPMENT OF CLOUD TECHNOLOGY IN THE CONTEXT OF THESE GREAT IN MECHANICAL ENGINEERING

Garmash O., Marchuk V., Bezreka M., Gradiskij Y.

The tendencies of development of cloud computing in the context of large data processing and control for continuous flow analysis with unstructured data useful information in real time. The prospects of using this technology in the field of mechanical engineering, logistics, education and other sectors of the economy.

Keywords: *cloud computing, big data management, information systems, information technology, database.*