УДК 004.75

МАСЛЯНКО П.П.,
РЯБУШЕНКО А.В.

СИСТЕМНЕ КОНСТРУЮВАННЯ ТА МОДЕЛЬ РОЗГОРТАННЯ РОЗПОДІЛЕНОЇ СИСТЕМИ УПРАВЛІННЯ ІНВЕСТИЦІЙНИМ ПОРТФЕЛЕМ

В роботі представлена модель розгортання розподіленої системи управління інвестиційним портфелем цінних паперів з використанням розподіленої нереляційної бази данних Cassandra. Проаналізовані переваги використання розподілених баз у порівнянні з реляційними базами даних.

In this paper deployment model of distributed investment portfolio management system is presented with application of distributed non-relational database Cassandra. Benefits of distributed databases are analyzed in comparison with relational databases.

Вступ

Застосування прикладного системного аналізу і системної інженерії проектів інформатизації організаційних систем дозволяє суттєво вдосконалити процес створення систем фінансово-інвестиційної діяльності. Завдяки формалізації бізнес процесів та застосуванню компонентного процесу розробки вдається суттєво підвищити ефективність бізнес процесів управління інвестиційним портфелем, знизити ризики і зменшити повну вартість володіння системою [1].

З метою залучення закордонного капіталу в Україні з'явилась потреба створювати системи управління інвестиційним портфелем цінних паперів, які відповідають міжнародним стандартам: BASEL II, Sarbanes-Oxley та IFRS [2,3]. Створення подібних систем потребує строго формалізованого підходу до аналізу, проектування, конструювання та впровадження системи, який би відповідав би міжнародним вимогам. Компонентний процес розробки, що використовують автори, відповідає міжнародних нормам в області software engineering та system engineering [4].

В роботах [5-7] запропонована компонентна модель системи управління фінансово-інвестиційною діяльністю, що розроблена у відповідності міжнародним стандартам з використанням компонентного процесу розробки.

В цій роботі досліджується проблема системного впровадження системи управління інвестиційним портфелем [5-7] у корпоративній гетерогенній мережі банку UBS, що входить до п'ятірки найбільших банків світу. Пропонується метод системного впровадження у вигляді розподіленої системи.

Розподілені системи у фінансово-інвестиційній сфері дозволяють підвищити надійність та продуктивність, забезпечити масштабованість та розширюваність, саме тому впровадження розподілених систем є надзвичайно актуальною проблемою для підвищення ефективності функціонування систем управління фінансово-інвестиційної діяльності.

Постановка задачі

Мета дослідження полягає в застосуванні системної інженерії проектів інформатизації організаційних систем для реалізації фази системного конструювання розподіленої системи управління інвестиційним портфелем і пошуку ефективних методів системного впровадження.

Об'єкт дослідження – розподілені обчислення, як метод впровадження системи управління інвестиційним портфелем цінних паперів.

Предмет дослідження – системна інженерія фінансово-інвестиційної діяльності.

Системне конструювання розподіленої системи управління інвестиційним портфелем

Ідея застосування розподілених систем у фінансово-інвестиційній та банківській сфері досліджується вже доволі давно. Результати багаточисельних досліджень підсумовані у [8,9]. Розподілені системи останні 10 років вважаються перспективним напрямком досліджень у фінансовому секторі економіки,

але їх впровадження у корпоративних мережах просувається дуже повільно.

Розподілені системи мають ряд переваг: масштабованість, підвищена надійність, заміна апаратного забезпечення без зупинки надання сервісів, можливість використання обчислювальних потужностей вже закуплених серверів, які, як показує статистика задіяні менш ніж на 10%, тим самим зменшити витрати на апаратне забезпечення та багато інших.

Але розподілені системи мають також ряд недоліків. Не кожну задачу можна ефективно вирішувати за допомогою розподілених обчислень. Задача повинна піддаватись розпаралелюванню, тобто містити в собі окремі функціональні блоки, що можуть виконуватись незалежно на різних апаратних вузлах. Розробка ефективних розподілених алгоритмів це складна та трудомістка задача [10]. Існуюче програмне забезпечення має бути переписаним з використання розподілених технологій, що потребує значний коштів та часу, саме тому компанії не квапляться упроваджувати нові розподілені технології.

Математичні методи управління ризиками, що використовуються в UBS, детально описані у [11]. Проаналізувавши математичні методи моделювання можна зробити висновок, що вони є сильно зв'язаною задачею, що погано піддається розпаралелюванню. Тому впровадження розподілених обчислень у системі управління інвестиційним портфелем спіткає ряд серьозних труднощів. Тим не менш, навіть в таких умовах розподілені обчислення можуть привнести додаткові переваги.

Стандартний метод впровадження, що найбільш поширений серед фінансових компаній – це трьох рівнева архітектура (клієнт, сервер, база даних). В такій архітектурі доволі часта саме база данних стає вузьким місцем. Реляційні БД погано масштабуються, тому саме вони часто стають причиною низької швидкодії та точкою єдиного виходу з ладу.

Автори пронуються використати розподілену базу данних заміст реляційної БД з метою підвищення надійності, швидкодіїї та відходу від реляційної моделі пердставлення інформаціїї.

Компонента модель системи управління інвестиційним портфелем

Компонентна модель системи управління інвестиційним портфелем та її специфікація представлена у [5] та розширена в роботах [6,7]. Система складається з наступних компонентів (рис. 1): оптимізація інвестиційного портфеля, управління ризиками, прогнозування, оцінка вартості похідних фінансових інструментів, сек'юритизація, алгоритмічний трейдинг, арбітраж, статистичний арбітраж, фінансова інженерія, прийняття стратегічних рішень. Основне призначення системи – надати інвестиційним компаніям інтегровану систему управління, що дозволить проводити моделювання інвестиційного портфеля для сформування портфеля з оптимальними характеристиками, прискорення прийняття інвестиційного рішення, ідентифікації та оцінки ризиків.

На рис. 2 наведено суперклас системи, що є структурним мета-представленням сутностей розподіленої системи управління інвестиційним портфелем. Суперклас представлений у нотації UML [12] і визначає сутність проблеми управління інвестиційний портфелем, його інтеграцію з зовнішніми системами. Таке представлення системи управління дозволяє на етапі аналізу виділити задіяні зовнішні системи та інтерфейси їх взаємодії. Система управління інвестиційним портфелем спрямована на використання інституційними інвесторами для управління великими портфелями цінним паперів. Клієнтам, які бажають приймати участь у управлінні їхньої частини портфеля надається інтерфейс користувача. Система повинна оперувати як на біржовому ринку, так і на позабіржовому ринку (Over-the-counter), який по об'єму торгів значно перевищує біржовий. Також, до класу позабіржових систем входить важливий тип ринків "Чорні пули ліквідності" [13], який став грати важливу роль з зниженням ліквідності на світових інвестиційних ринках, що викликано фінансовою кризою.

Інформаційні агенції надають інформацію, яка необхідна для прийняття інвестиційного рішення. Інформаційні агенції можна умовно розділити на три групи: провайдери фінансових нових, такі як Reuters, Bloomberg; провайдери фінансових звітів емітентів, наприклад Bloomberg, FSA та про-

вайдери аналітичних звітів, наприклад JPMorgan Chase, Morgan Stanley.

Засоби реалізації

На рис. 3 наведено пакет засобів реалізації системи управління інвестиційним портфелем у профілі Еріксона-Пенкера [14]. Засоби реалізації розподілені на програмі та апаратні.

До апаратних відносять елементи корпоративної мережі: сервери додатків, сервери баз даних, маршрутизатори та мережеве обладнення. Згідно з внутрішніми корпоративними стандартами UBS, в якості серверів баз даних використовуються сервери SUN з процесорами UltraSparc. Для серверів додатків був зроблений виняток – використовуються сервери SUN з процесорами AMD Opteron.

Згідно з внутрьошнімі стандартми UBS до програмних засобів відносять: операційну систему – Solaris 10, файлову систему – ZFS, реляційну базу даних – Oracle 10g.

Взаємодія з зовнішніми системами угод відбувається за стандартним протоколом FIX (Financial Information eXchange). Для систем з підвищеними вимогами до продуктивності застосовується протокол FAST (FIX Adapted for STreaming). Для котирувань складних похідних фінансових інструментів використовується додаткові протоколи: MDDL (Market Data Description Language) та FpML (Financial Products Markup Language). З зовнішніми інформаційними системами взаємодія відбувається за стандартними протоколами, що базуються на XML: для обміну фінансовими новинам – NewsML (News Markup Language), фінансовою звітністю емітентів цінних паперів – XBRL (eXtensible Business Reporting Language), аналітичними звітами – RIXML (Research Information eXchange Markup Language) [15].

Системне розгортання розподіленої системи управління інвестиційним портфелем

На рис. 4 наведена модель розгортання системи управління інвестиційним портфелем. Система включає в себе: основну та резервну реляційні бази даних, що виконують роль центрального сховища для обробки аналітичних запитів, розподілювач навантаження з фільтрацією запитів. Для збереження критично важливої бізнес інформації забезпечується її архівація на магнітну стрічку. Загрузка даних із зовнішніх джерел реалізована окремими компонентами, які завантажують інформацію у реляційну БД.

Система складається з необхідної кількості типових обчислювальних вузлів з типовим програмним забезпеченням (рис. 4). Апаратна реалізація вузлів може бути різною.

В якості розподіленої БД використовується нереляційна БД − Cassandra. Кожен обчислювальних вузол має свій екземпляр розподіленої БД. База даних Cassandra поєднує масштабованість, надмірність та надійність такої повністю розподіленої бази данних, як Dynamo [16] та баз данних розділенних таблиць (partitioned table) BigTable і Hadoop [17]. Cassandra не підтримує мови запитів SQL, так як вона не є реляційною базою данних. Замість SQL використовується мова запитів, що схожа на MapReduce [18]. Сховище Cassandra орієнтовано на зберігання масивів (column-oriented storage) [19,20], що має кращу швидкодію у порівнянні зі зберіганням окремих значень (value-oriented storage), таких розподілених БД, як Memcache та Dynamo; та row-oriented storage реляційних БД.

База даних Cassandra забезпечує надійність завдяки автоматичній реплікації розподілених БД, а також підтримує прозоре введення нових обчислювальних вузлів у кластер без зупинки обслуговування. Нереляційна БД Cassandra, на відміну від реляційної БД, забезпечує значне підвищення продуктивності обчислень та дозволяє ефективно оперувати різними типами фінансових інструментів, які не можна ефективно представили у реляційній алгебрі. Погана структурованість фінансових інструментів та дуже великі їх відмінності один від одного не дозволяють спроектувати ефективну БД фінансових інструментів у реляційній алгебрі.

Обчислювальні задачі виконуються на тому вузлі, який має необхідну інформацію у власному екземплярі БД, що суттєво знижує навантаження на комунікаційну мережу. Результати обчислень спочатку зберігаються в нереляційній БД, а потім в асинхронному режимі передаються в реляційну БД.

Для прискорення обчислювально-складних алгоритмів в системі задіяні відеокарти.

Сучасні відеокарти мають від 500-1200 процесорів обробки чисел з плаваючою крапкою одинарної точності. Це величезна обчислювальна потужність, яку можна задіяти, якщо алгоритм піддається розпаралелюванню. Для цього застосовуються GPGPU бібліотеки [21], які дозволяють використовувати відеокарти в тому числі і для математичних обчислень.

Висновки

Застосування системної інженерії для реалізації фази системного конструювання систем управління інвестиційним портфелем цінних паперів забезпечує необхідні значення ключових властивостей ІКС, а саме: продуктивності, масштабованості, розширюваності, надійності, безпеки, прозорості, низької повної вартості володіння системою.

Необхідні значення ключових властивостей ІКС досягаються завдяки використанню розподіленої нереляційної БД Cassandra. Крім цього, відмова від реляційного принципу представлення даних, дозволило ефективно оперувати фінансовими інструментами максимально наблизити їх використання до реального масштабу часу.

Використання процесорів відео карт та GPGPU бібліотек [21] дозволило суттєво підвищити ефективність використання наявних обчислювальних ресурсів для складних алгоритмів, наприклад непараметричне Монте Карло [7], генерація багатовимірних випадкових чисел [22] та генетичного алгоритму оптимізації [23].

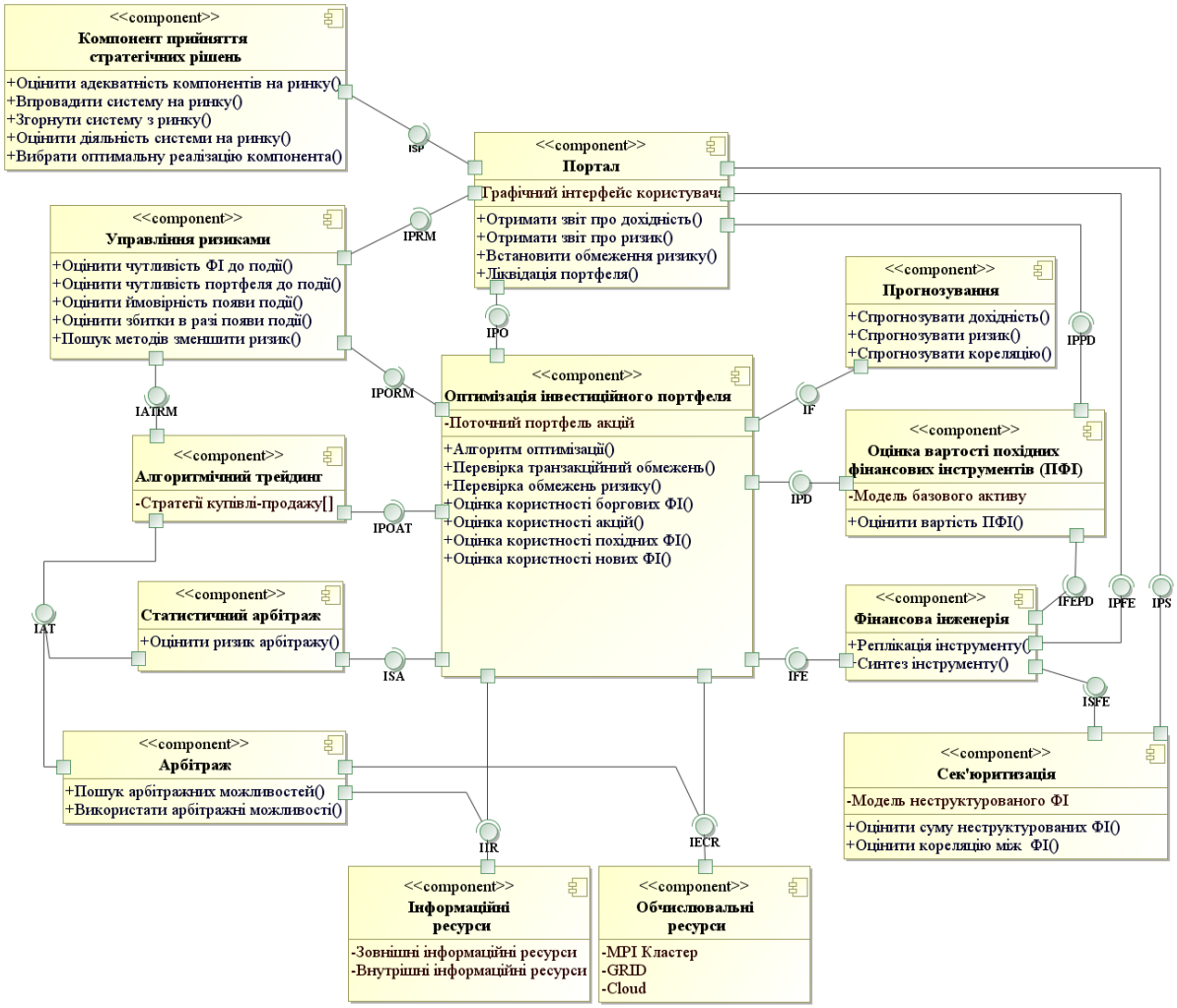

Рис. 4. Модель системи управління фінансово-інвестиційною діяльністю. Діаграма компонентів в нотації UML

Рис. 2. Суперклас системи управління інвестиційним портфелем. Діаграма класів у нотації UML

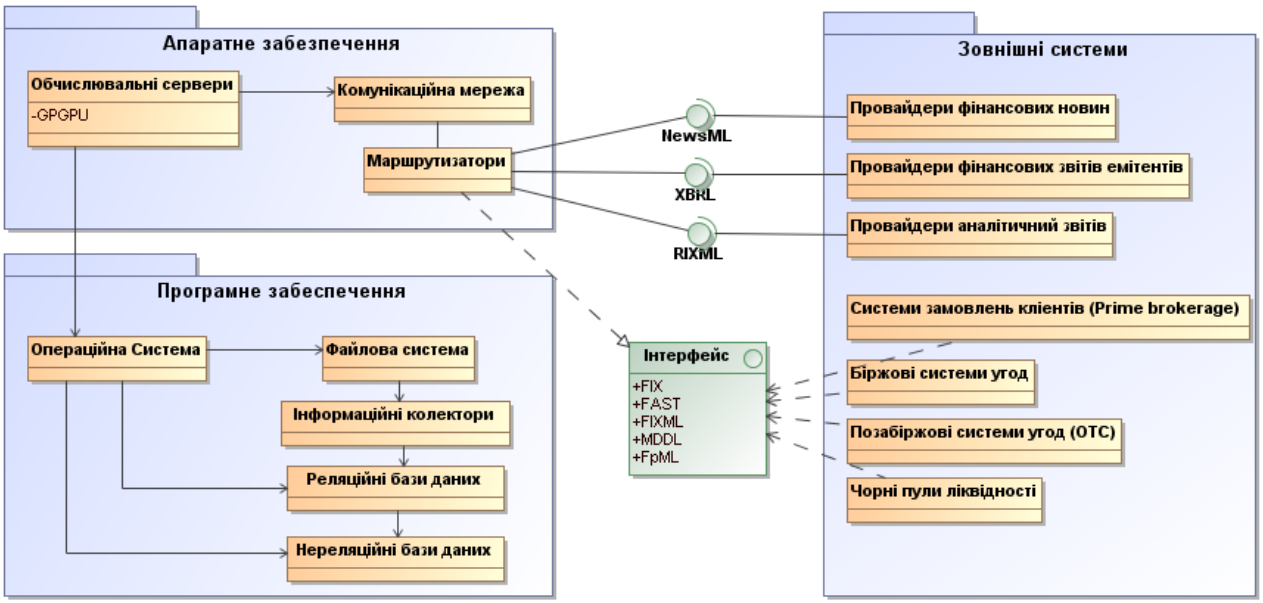

Рис. 3. Засоби реалізації системи управління інвестиційним портфелем. Пакет класів у профілі Еріксона-Пенкера

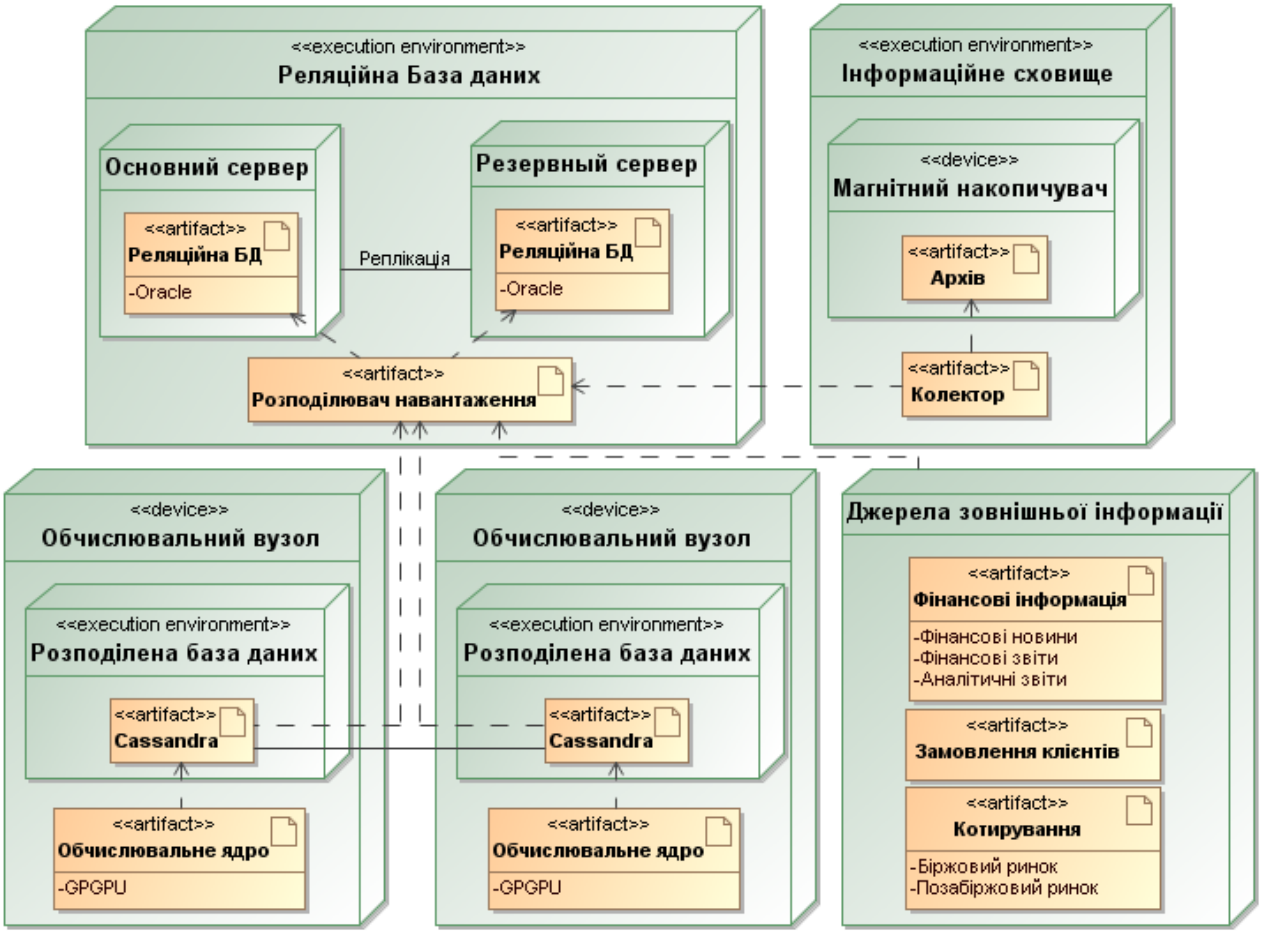

Рис. 4. Модель розгортання системи управління інвестиційним портфелем. Діаграма впровадження у нотації UML

Список посилань

1. Маслянко П.П. Концепція інформатизації корпоративних структур. Наукові вісті НТУУ “КПІ” 2001, № 3 - С.45-50.
2. Basel Committee on Banking Supervision. International Convergence of Capital Measurements and Capital Standards. A Revised Framework. 2005. Bank of international settlements - 284 pp.
3. Tarantino A. Manager's Guide to Compliance: Sarbanes-Oxley, COSO, ERM, COBIT, IFRS, BASEL II, OMB's A-123, ASX 10, OECD Principles, Turnbull Guidance, Best Practices, and Case Studies // Wiley, 2006. – 336pp.
4. Маслянко П.П. Основні положення методологій системного проектування інформаційно-комунікаційних систем. Наукові вісті НТУУ “КПІ” 2007, № 6 - С.54-60.
5. Маслянко П. П., Рябушенко А. В. Компонентна модель інформаційно-аналітичної системи та генетичний алгоритм формування оптимального портфеля акцій // Наукові вісті НТУУ “КПІ” 2009, № 1 - С.36-46.
6. Маслянко П. П., Рябушенко А. В. Створення компонента стратегічного планування системи управління фінансово-інвестиційною діяльністю // Наукові вісті НТУУ “КПІ” 2009, № 4 - С.53-65.
7. Маслянко П. П., Рябушенко А. В. Підсистема управління ризиками фінансово-інвестиційної діяльності // Вісник східноукраїнського національного університету ім. Володимира Даля. – 2009. – Т.2, №1. – С.370–378.
8. Tanenbaum A. S., Van Steen M. Distributed Systems: Principles and Paradigms. – Prentice Hall, 2006. – 2nd Edition. – 704 pp.
9. Dollimore J., Kindberg T., Coulouris G. Distributed Systems: Concepts and Design Addison. – Wesley, 2005. – 4th Edition. – 944 pp.
10. Ghosh S. Distributed Systems: An Algorithmic Approach. – Chapman & Hall/CRC Computer, 2006. – 424 pp

11. Cesari G., Aquilina J., Charpillon N. та ін. Modelling, Pricing and Hedging Counterparty Credit Exposure // Springer finance, 2009. – 246 pp.
12. Unified Modeling Language Specification, Version 2.0. Object Management Group, Framingham, Mass., 2004.
13. Banks E. Dark Pools: The Structure and Future of Off-Exchange Trading and Liquidity. – Palgrave Macmillan, 2010. – 240 pp.
14. Ericsson H.E., Penker M. Business Modeling with UML: Business Patterns at work // Wiley Computer Publishing, 2000. – 350 pp.
15. Martin S. MDDL and the quest for a market data. – Elsevier, 2007. – 320 p.
16. Hastorun D., Jampani M., Kakulapati G. та ін. Dynamo: amazon's highly available key-value store // In Proc. SOSP, 2007. – P.205-220.
17. Chang F., Dean J., Ghemawat S. Bigtable: A Distributed Storage System for Structured Data // OSDI'06: Seventh Symposium on Operating System Design and Implementation, 2006. – P. 205–218.
18. J Dean, S Ghemawat. MapReduce: Simplified data processing on large clusters // Communications of the ACM, 2008. – Vol. 51, No. 1. – P. 107-113.
19. Stonebraker M., Abadi D. J., Batkin Ad. та ін. C-store: a column-oriented DBMS // VLDB '05: Proceedings of the 31st international conference on Very large data bases. – VLDB Endowment, 2005. – p. 553-564
20. Abadi D., Madden S., Ferreira M. Integrating compression and execution in column-oriented database systems // SIGMOD '06: Proceedings of the 2006 ACM SIGMOD international conference on Management of data. – ACM, 2006. – p.671-682
21. Pharr M., Fernando R. GPU Gems 2: Programming Techniques for High-Performance Graphics and General-Purpose Computation // Addison-Wesley Professional, 2005. – 880 pp.
22. Маслянко П. П., Рябушенко А. В. Непараметричне Монте Карло: вдосконалений метод моделювання ризиків на фінансовому ринку // Всеукраїнський математичний конгрес. – К.: Інститут математики НАН України, 2009. http://www.imath.kiev.ua/~congress2009/Abstracts/Maslyanko_Ryabushenko.pdf
23. Маслянко П.П., Рябушенко А.В. Генетичний алгоритм оптимізації портфеля акцій на фондовому ринку // Системний аналіз та інформаційні технології. – К.: ННК "ІПСА" НТУУ "КПІ", 2009. – С.347.

Поступила в редакцию 3.12.2009