

GRID РЕСУРСИ ІНФОРМАЦІЙНО-ОБЧИСЛЮВАЛЬНОГО ЦЕНТРУ КИЇВСЬКОГО НАЦІОНАЛЬНОГО УНІВЕРСИТЕТУ ІМЕНІ ТАРАСА ШЕВЧЕНКА

О.О. Судаков, Ю.В. Бойко

Київський національний університет імені Тараса Шевченка
01033, Київ, вул. Володимирська, 64,
факс. 380 (44) 526-1214, тел. 380 (44) 259-0247,
E-mail: saa@grid.org.ua, boyko@univ.kiev.ua

GRID – це форма розподілених обчислень, що передбачає спільне використання та керування обчислювальними ресурсами, та ресурсами збереження даних в організаціях, які можуть бути географічно розподіленими та динамічно змінюватись. За допомогою GRID-систем з'являється можливість вирішення надскладних обчислювальних задач, які виникають у різних галузях людської діяльності. Створення GRID-систем вимагає об'єднання ресурсів фізичних організацій (обчислювальні засоби, системи збереження даних, програмні засоби, бази даних, мережеві ресурси) у віртуальну організацію для спільного використання. Інформаційно-обчислювальний центр (ІОЦ) Київського національного університету імені Тараса Шевченка вже понад сім років займається успішним розвитком і експлуатацією обчислювального кластера та інфраструктури комп'ютерних мереж, які застосовуються для розв'язання наукових і освітніх задач різними науковими та освітніми організаціями України. Наводиться огляд технічних характеристик обчислювального кластера і мережевих ресурсів ІОЦ університету та досвід їх застосування для GRID-обчислень.

GRID is a form of distributed computing that provides sharing and control of computing resources across dynamic and geographically distributed organizations. GRID computing provides the possibility of solving very computationally expensive problems that may arise in different branches of human activity. Creation of GRIDs requires connection of different computing resources of physical organizations (like CPU power, data storage, applications, databases, network resources) into virtual organization for common usage. Information and Computer Center of National Taras Shevchenko University of Kyiv more then seven years carries out successful development and usage of computing cluster and network infrastructure that found its application by different scientific and educational organizations of Ukraine. Description of technical characteristics and application experience of computing cluster and network resources of Kyiv University for GRID computing is provided.

Вступ

GRID (ГРІД) обчислення – це спеціальна форма розподілених обчислень, яка направлена на створення обчислювальних та інформаційних ресурсів, які спільно використовуються та контролюються географічно розподіленими та динамічно змінюваними віртуальними організаціями. Кожна віртуальна організація включає у себе фізичні обчислювальні ресурси (процесорну потужність, програмні засоби, засоби збереження даних, самі дані, ресурси комп'ютерних мереж) та декілька фізичних організацій, які надають частину своїх ресурсів у спільне використання та можуть користуватись цими спільними ресурсами. Термін GRID походить від англійського "мережа" і обраний у зв'язку з тим, що робота таких систем схожа на роботу єдиних енергосистем, коли потужність географічно розподілених електростанцій перерозподіляється між споживачами в залежності від потреб.

Сьогодні в світі до GRID-обчислень є надзвичайний інтерес, особливо в наукових та промислових колах [1], оскільки зазначені технології дають можливість на основі наявних обчислювальних ресурсів створювати потужні системи для розв'язання складних задач. У Європейській організації з ядерних досліджень (CERN) створюються GRID-системи для обробки експериментальних даних. В Україні інтерес до GRID-обчислень також зростає. У 2005 році вперше до системи ALIEN GRID [2] були підключені два українських обчислювальних кластера: кластер Інституту теоретичної фізики ім. М.М. Боголюбова НАН України та кластер Київського національного університету імені Тараса Шевченка, ведуться роботи з створення першого в Україні GRID-сегменту.

Для створення GRID-систем необхідна наявність обчислювальних ресурсів, до яких можна надати доступ зовнішнім користувачам. Мета даної роботи є опис принципів побудови обчислювальних ресурсів: кластера та комп'ютерної мережі інформаційно-обчислювального центру Київського національного університету імені Тараса Шевченка, які дали можливість успішно інтегрувати ці ресурси до GRID-системи.

Обчислювальні ресурси ІОЦ

Загальні характеристики GRID. Концепції GRID-систем з'явилися у кінці 90-х років минулого сторіччя, як спосіб промислового застосування метакомп'ютерних технологій – об'єднання великої кількості незалежних обчислювальних машин, що з'єднані через Інтернет з метою використання ресурсів обчислювальних систем, що в даний момент незадіяні. Потенційна процесорна потужність та обсяги оперативної та зовнішньої пам'яті

метакомп'ютерних систем можуть бути надзвичайно великими, проте такі системи характеризуються малою надійністю, низьким рівнем захисту, а також відсутністю єдиних стандартів. Відкриті стандарти для GRID-систем, які нині знаходяться у стадії розробки [3] мають на меті вирішення зазначених проблем.

Реалізація GRID-систем передбачає використання спеціального проміжного програмного забезпечення (middleware), яке реалізує всю складність взаємодії між комп'ютерами в розподіленій системі. Проміжне програмне забезпечення включає у себе бібліотеки, які застосовуються для розробки програмного забезпечення, GRID-служби – процеси, які забезпечують роботу системи (балансування навантаження, запуск задач та доступ до дискових ресурсів віддалених машин) та бази даних інформації про використання ресурсів розподіленої системи. Більшість існуючих реалізацій GRID-систем побудовано на основі WWW технологій [4]. Найпопулярнішою реалізацією проміжного програмного забезпечення, яке застосовується у багатьох GRID-проектах, таких як ALIEN GRID, LCG GRID, EDG, та інших є Globus Toolkit [5].

Для підключення ресурсів, таких як обчислювальний кластер до GRID-системи, необхідна наявність постійного підключення до Інтернет за допомогою швидкісних каналів зв'язку, а також сумісність програмного забезпечення обчислювального кластера з проміжним програмним забезпеченням, що використовується в GRID.

Комп'ютерна мережа. Комп'ютерна мережа Київського національного університету імені Тараса Шевченка є однією з найбільших серед наукових та освітніх закладів України. Інформаційно-обчислювальний центр (ІОЦ) університету постійно займається розвитком мережевої інфраструктури, її ресурсів та сервісів.

Мережа університету об'єднує факультети та інститути в різних частинах міста Києва (рис. 1). Вона має у своєму складі декілька точок присутності, а саме: Інститут міжнародних відносин, Червоний корпус університету, Національна бібліотека України імені В.І.Вернадського, Інститут фізіології, інформаційно-обчислювальний центр університету та АТС 526. Усі підрозділи та партнери університету підключені за допомогою різних технологій та технічних засобів. Зовнішні канали забезпечуються присутністю обладнання у провідних провайдерів міста Києва. Загальна пропускна спроможність зовнішніх Інтернет-каналів складає 10 МБіт за сек, швидкість українського Інтернет-сегмента з урахуванням паритетних каналів близько 400 МБіт за сек. Опорна мережа університету підтримує швидкість 1 ГБіт за сек.

Інформаційні ресурси. Київський національний університет імені Тараса Шевченка має розвинену інфраструктуру мережевих інформаційних ресурсів [6], які створені співробітниками інформаційно-обчислювального центру університету. До цих ресурсів належать www та ftp архіви наукової та технічної інформації, а також навчальні системи, служби мультимедійної Інтернет-трансляції. WWW та ftp ресурси містять технічну та наукову документацію з інформаційних технологій, природничих та гуманітарних наук, а також ресурси дистанційного навчання та електронні навчально-методичні матеріали [7]. Служба мультимедійних Інтернет-трансляцій застосовується для проведення аудіо- та відео-трансляцій лекцій та інших подій, що відбуваються в університеті [8]. Факультети, інститути та інші підрозділи університету також мають свої інформаційні та обчислювальні ресурси.

Обчислювальний кластер. Обчислювальний кластер Київського національного університету імені Тараса Шевченка [9] створений у 1999 році, як система для вирішення задач, що вимагають значних затрат обчислювальних ресурсів та оперують з великими обсягами даних. Система була створена, як відкрита з доступом у Інтернет для наукових та освітніх організацій України і складалася з 4 вузлів, мала 4 процесори, сумарний обсяг оперативної пам'яті 512 Мбайт і дисковий простір 160 Гбайт.

Обчислювальні ресурси кластера постійно розвиваються силами університету та інформаційно-обчислювального центру. До 2003 року це була найпотужніша подібна система в Україні, нині – один з найбільш використовуваних розрахункових ресурсів України [10]. Корпорація Intel, яка надавала підтримку при створенні кластера, неодноразово зазначала, що лабораторія кластерних обчислень ІОЦ Київського університету є важливим ресурсом для академічних та наукових установ України [11]. Нині на кластері працює більше сотні користувачів з близько двадцяти наукових та освітніх організацій України, серед яких Інститут молекулярної біології і генетики НАН України, Інститут фізики НАН України, Інститут фізики напівпровідників НАН України, Національний університет Києво-Могилянська академія та інші. Кластер застосовується для розрахунків задач квантової-хімії, молекулярної динаміки, теорії алгоритмів, теорії груп, фізики поверхні, розробки лікарських препаратів та ін.

На момент написання роботи кластер складається з 22 гетерогенних вузлів, має 44 процесори (Intel Pentium III-1ГГц, Intel Xeon-2.4 ГГц, Intel Xeon EM64T-3.2 ГГц). Вузли кластера об'єднано за допомогою технології Gigabit Ethernet з швидкістю передачі інформації 1 Гбіт за сек. Пікова продуктивність системи близько 178 GFlops.

Успішність застосування системи для розв'язання різних задач, у тому числі для GRID-обчислень, зумовлена вдалими особливостями побудови та конфігурації програмного та апаратного забезпечення системи.

Конфігурація апаратного забезпечення системи розроблялась з наступних міркувань: вимоги до обчислювальної потужності для існуючих розрахункових задач, можливість швидкого масштабування, надійність та економічна обґрунтованість. Це дозволило забезпечити оптимальне використання існуючих потужностей, гарантувати безперебійну роботу системи та раціонально використати економічні ресурси.

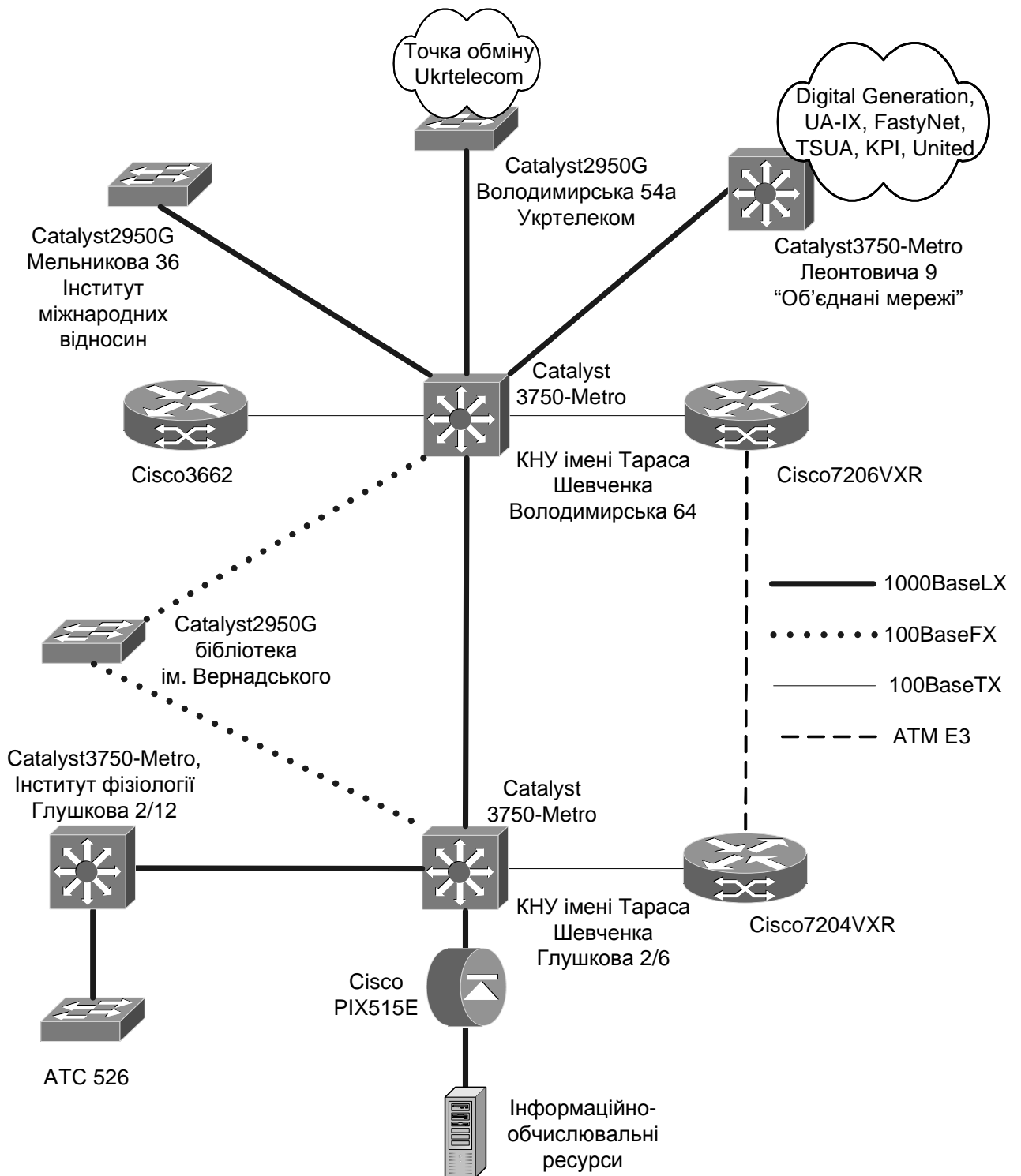


Рис. 1. Схема комп'ютерної мережі Київського національного університету імені Тараса Шевченка

Все програмне забезпечення кластера має відкритий вихідний код. Кластер належить до кластерів типу Beowulf, тобто має розподілений образ операційної системи. Структурна схема кластера показана на рис. 2.

Вузли кластера керуються операційною системою Linux (у даний час на основі Fedora Linux-1 та Fedora Linux-3 з розширеннями, що розроблені в лабораторії). Всі вузли кластера (наскільки це можливо) виконані однотипними. Кожен вузол містить жорсткий диск на якому зберігається копія операційної системи цього вузла, копія програмного забезпечення, простір для тимчасових даних, домашні каталоги користувачів та простір підкачки сторінок. Збереження копій прикладного програмного забезпечення та операційної системи на всіх вузлах хоч і призводить до надмірного використання дискового простору, однак це дозволяє тримати робочі резервні копії системи, швидко відновлювати вузли при збоях, та зменшити навантаження на мережу при запусках програм. Копії програмного забезпечення на різних вузлах дозволяє також оптимізувати програмне забезпечення для кожного вузла гетерогенної системи.

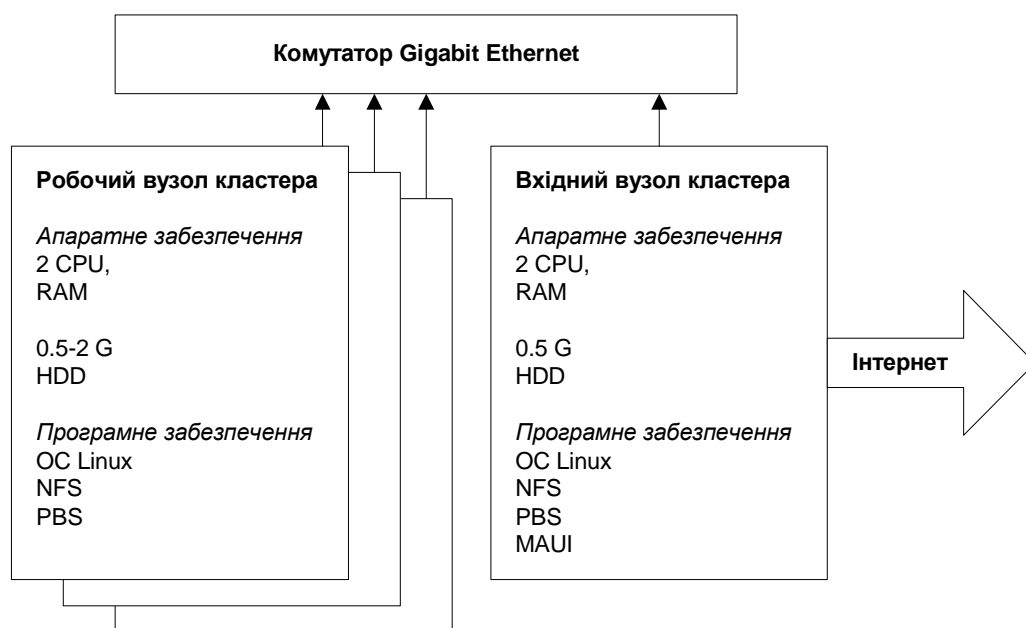


Рис. 2. Функціональна схема кластера

Запуск програм на інших вузлах виконується за допомогою протоколів rsh та ssh і доступ у межах кластера виконується без пароля. Це забезпечується за допомогою спільного простору системних імен (користувачі, імена машин та ін.) Розподілена система імен виконана на основі LDAP, що дає перевагу порівняно з іншими розподіленими системами імен (NIS, hesiod), оскільки вона добре масштабується, розрахована на велику кількість вузлів, роботу в Інтернет і добре захищена від несанкціонованого доступу. Недоліком LDAP є складність конфігурації та повільність. Другий недолік легко побороти за допомогою кешування. Вся авторитетна інформація знаходиться на вузлі керування, а кожен інший вузол містить копію авторитетної інформації, що дає можливість легко відновити роботу системи при збоях головного вузла. Вхідний вузол суміщений з вузлом керування, а також дозволяє запускати розрахункові задачі. При збоях вхідного вузла його функції можуть виконати інші вузли.

Дискові ресурси кластера розподілені між вузлами. Всі дискові ресурси всіх вузлів експортуються за допомогою протоколу NFS на всі інші вузли кластера. Спільна файлова система реалізована за допомогою автоматизування розподілених дискових ресурсів у каталоги з однаковим іменем. Таке рішення дає можливість реалізувати прозорість розташування ресурсів і виконувати непрозору міграцію даних шляхом зміни посилань на дані різних вузлів.

Розподіл та балансування ресурсів кластера виконується за допомогою системи torque [12] (клон PBS). Раніше для цих потреб використовувались системи динамічного балансування навантаження з міграцією процесів MOSIX та openMOSIX, проте при збільшенні кількості вузлів стабільність цих систем зменшувалась, а необхідність в динамічному балансуванні ставала все меншою. Система PBS забезпечує запуск задач у пакетному режимі. Розподіл ресурсів виконується на основі цільової функції, яка враховує відносну швидкість вузлів, кількість задач на вузлі, завантаженість вузла та використання пам'яті вузла (1). Задача запускається на вузлі з максимальним пріоритетом

$$p = S - A_{cpu} - L + 0.01A_{mem} - J \quad , \quad (1)$$

де p – пріоритет вузла, S – відносна швидкість вузла, A_{cpu} – кількість вільних процесорів на вузлі, L – середня завантаженість вузла, A_{mem} – обсяг доступної пам'яті на вузлі, J – кількість задач на вузлі. Планування задач виконується за допомогою планувальника MAUI [12].

Система пакетного режиму PBS є надзвичайно поширеною і підтримується багатьма реалізаціями програмного забезпечення для GRID-систем. Це однією причиною легкого інтегрування кластера з GRID-системами та легкості інсталяції нового програмного забезпечення є однаковість конфігурації програмного забезпечення різних вузлів кластера, а також відсутність надмірних заходів безпеки (таких як заборона доступу до вузлів в межах кластеру), які часто лише знижують ефективність та зручність застосування, а також зменшують надійність системи.

На кластері також інстальовані розширення операційної системи та прикладне програмне забезпечення, зокрема, розроблене в лабораторії програмне забезпечення для резервування та відновлення задач [13]. У 2005 році за аналогічними принципами побудована кластерна система в Інституті теоретичної фізики НАН України. Ця система також інтегрована в систему ALIEN GRID.

Висновки

Підсумовуючи сказане, можна зробити наступний висновок, що застосування програмного забезпечення з відкритим вихідним кодом та описаного підходу до побудови кластерних систем, дає можливість створити GRID-ресурси, які легко інтегруються з існуючими системами та можуть застосовуватись для розв'язання широкого кола задач. Цей висновок підтверджується багаторічним успішним застосуванням обчислювальних ресурсів ІОЦ Київського національного університету імені Тараса Шевченка, як співробітниками університету, так і партнерами – науковими та освітніми організаціями України, а також досвідом застосування описаного підходу в інших наукових організаціях України, зокрема, в Інституті теоретичної фізики НАН України.

1. *Chervenak A., Foster I., Kesselman C., Salisbury C., Tuecke S.* The Data Grid: Towards an Architecture for the Distributed Management and Analysis of Large Scientific Datasets. // J. of Network and Computer Applications, 23:187-200, 2001 (based on conference publication from Proceedings of NetStore Conference 1999).
2. Система моніторингу ресурсів Alice <http://pcalimonitor.cern.ch:8889/>
3. *Foster I., Kesselman C., Tuecke S.* The Anatomy of the Grid: Enabling Scalable Virtual Organizations // International J. Supercomputer Applications. – 2001. – Vol. 15. – P. 200–222.
4. *Humphrey M., Wasson G., Jackson K., Boverhof J.* State and Events for Web Services: A Comparison of Five WS-Resource Framework and WS-Notification Implementations. 4th IEEE International Symposium on High Performance Distributed Computing (HPDC-14), Research Triangle Park, NC, 24-27 July 2005.
5. *Foster I.* Globus Toolkit Version 4: Software for Service-Oriented Systems // IFIP International Conference on Network and Parallel Computing, Springer-Verlag LNCS 3779, 2005. – P. 2–13.
6. Ресурси інформаційно-обчислювального центру Київського національного університету імені Тараса Шевченка <http://icc.univ.kiev.ua/resources/>
7. Система дистанційного навчання Київського національного університету імені Тараса Шевченка <http://ilias.univ.kiev.ua/>
8. Сервер мультимедійних трансляцій Київського національного університету імені Тараса Шевченка <http://electure.kiev.ua/>
9. Обчислювальний кластер Київського національного університету імені Тараса Шевченка <http://www.cluster.kiev.ua/>
10. Система моніторингу ресурсів обчислювального кластера ІОЦ Київського національного університету імені Тараса Шевченка. <http://cluster.univ.kiev.ua/ganglia/>
11. Прес-реліз Intel, Київ 29 жовтня, 2003 р. <http://www.intel.com/ru/pressroom/2003/103003.htm>
12. Системи керування ресурсами обчислювальних кластерів <http://clusterresources.com/>
13. *Судаков О.О., Бойко Ю.В., Третяк О.В., Короткова Т.П., Мецераков Є.С.* Система резервування та відновлення процесів для Linux // Математичні машини і системи. – 2003. – № 2. – С. 146–153.