

УКРАЇНСЬКИЙ АКАДЕМІЧНИЙ ГРІД: ДОСВІД СТВОРЕННЯ І ПЕРШІ РЕЗУЛЬТАТИ ЕКСПЛУАТАЦІЇ

Abstract: Paper describes design features of Grid-infrastructure of The National Academy of Sciences of Ukraine and the first results of its application. History of creation of Grid-systems in Ukraine is depicted. The main requirements for computing resources, communication channels and middleware are listed. Beowulf clusters concept and Advanced resource connector (ARC) middleware were chosen and approved. Technical characteristics Grid-infrastructure clusters and results of its practical application for physics, biology and information technology problems are listed.

Key words: Grid, cluster, middleware.

Анотація: Описано особливості побудови та перші результати застосування Грід-інфраструктури Національної академії наук України. Наведено історичні відомості про побудову Грід-систем в Україні. Сформульовано основні вимоги до обчислювальних ресурсів, каналів передачі даних та проміжного програмного забезпечення. Обґрунтовано вибір концепції Beowulf-кластерів та проміжного програмного забезпечення Advanced resource connector (ARC). Наведено характеристики кластерних установок Грід-інфраструктури та результати практичного застосування системи для розв'язання задач фізики, біології, інформаційних технологій.

Ключові слова: Грід, кластер, проміжне програмне забезпечення.

Аннотация: Описаны особенности построения и первые результаты использования Грід-инфраструктуры Национальной академии наук Украины. Приведены исторические сведения о создании Грід-систем в Украине. Сформулированы основные требования к вычислительным ресурсам, каналам передачи информации и промежуточному программному обеспечению. Обосновано выбор концепции кластеров типа Beowulf и промежуточного программного обеспечения Advanced resource connector (ARC). Приведены характеристики кластерных установок Грід-инфраструктуры и результаты ее практического использования для решения задач из областей физики, биологии, информационных технологий.

Ключевые слова: Грід, кластер, промежуточное программное обеспечение.

1. Вступ

Метою проекту побудови Української академічної Грід-інфраструктури є об'єднання обчислювальних ресурсів наукових установ Національної академії наук України, навчальних закладів та інших організацій для ефективного розв'язання складних задач, що вимагають значних обчислювальних ресурсів. Побудова високошвидкісної оптоволоконної мережі, проектування і налаштування обчислювальних кластерів і систем зберігання даних, використання програмного забезпечення для об'єднання обчислювальних ресурсів дозволило створити повнофункціональну Грід-систему, що надає доступ користувачам до цих ресурсів.

У даній статті представлена концепція побудови Грід-інфраструктури НАН України, описано етапи реалізації проекту, обґрунтовані причини вибору проміжного програмного забезпечення, наведені перші результати експлуатації Грід-інфраструктури при розв'язанні наукових задач.

2. Коротка історична довідка

У наш час практично в кожній розвинутій країні реалізуються національні Грід-проекти, що мають на меті створення відповідної інфраструктури й розвиток технологій, які забезпечують віддалений доступ до різноманітних обчислювальних ресурсів незалежно від місця розташування споживача [1]. У Європі всі роботи в цій області координуються в рамках проекту EGEE (Enabling Grids for E-science in Europe).

В Україні впровадження Грід-технологій у наукових дослідженнях перебуває на початку свого життєвого циклу – фактично на стадії науково-дослідних розробок. Перший в Україні Грід-вузол був створений у Харківському фізико-технічному інституті НАН України в рамках спільних

проектів з Об'єднаним інститутом ядерних досліджень (м. Дубна, Росія) та Європейським центром ядерних досліджень (ЦЕРН) (м. Женева, Швейцарія). У 2005 році за ініціативою Інституту теоретичної фізики ім. М.М. Боголюбова НАН України (ІТФ) і Інформаційно-обчислювального центру Київського національного університету імені Тараса Шевченка (КНУ) у рамках системи AliEn-GRID для обробки даних експерименту ALICE у ЦЕРНі були створені ще два вузли [2, 3], однак національною Грід-системою їх назвати не можна, оскільки вони орієнтовані на розв'язання спеціалізованих завдань фізики високих енергій.

Розуміння необхідності розвитку Грід-технологій в Україні і набутий досвід експлуатації Грід-вузлів (потрібно відзначити, що в середньому на кластері ІТФ виконувалося до 40 завдань на добу), у 2006 році дозволили приступити до реалізації проекту створення Грід-інфраструктури НАН України. Ця Грід-мережа повинна була об'єднати провідні дослідницькі інститути й організації України, зацікавлені в розвитку та просуванні Грід-технології і застосувань Грід.

Дана ініціатива була підтримана Президією Національної академії наук України, що своїм рішенням затвердила концепцію й Програму впровадження Грід-технологій і створення кластерів у Національній академії наук України на 2006 рік. Цією Програмою були передбачені створення Грід-кластерів в Інституті молекулярної біології й генетики НАН України (ІМБГ), Інституті клітинної біології й генної інженерії НАНУ (ІКБГІ), Головній астрономічній обсерваторії (ГАО), а також модернізацію кластерів в Інституті теоретичної фізики ім. М.М. Боголюбова НАН України, Інституті фізики конденсованих систем (ІКС) у м. Львів, Харківському фізико-технічному інституті. Всі ці обчислювальні ресурси в підсумку повинні були створити основу Української Грід-інфраструктури. У грудні 2006 року Грід-інфраструктура була введена в тестову експлуатацію [4]. У березні 2007 року до Грід-мережі НАН України був підключений потужний обчислювальний кластер SKIT-3 Інституту кібернетики НАН України (ІК) і кластер Інституту космічних досліджень НАН – НКА України (ІКД).

Ця коротка історична довідка показує, що Грід-ініціатива в Україні тільки розвивається, і ми перебуваємо на початковій стадії, коли основним завданням є побудова Грід-інфраструктури, введення в експлуатацію обчислювальних кластерів, навчання Грід-користувачів. Тому в даній статті немає вичерпних експлуатаційних характеристик Грід-мережі, які можна одержати тільки через якийсь час після введення в експлуатацію.

3. Обчислювальні ресурси

Грід-інфраструктура [1] базується на трьох основних елементах:

- 1) обчислювальні ресурси (кластери, системи зберігання даних);
- 2) високошвидкісний і надійний доступ цих ресурсів до мережі Інтернет;
- 3) проміжне програмне забезпечення (middleware), що поєднує ці ресурси в єдиний обчислювальний комплекс.

На момент створення Грід-інфраструктури проміжне програмне забезпечення було доступним (про його вибір буде сказано нижче). Побудова ж кластерів і оптоволоконної мережі НАН України виконувалася практично одночасно зі створенням Грід-системи. Досвід Інформаційно-обчислювального центру КНУ по створенню кластерних систем, результати тестування й

експлуатації Грід-сайту системи AliEn-GRID [5] дозволили сформулювати основні шляхи розвитку Грід-інфраструктури в Національній академії наук України.

За концептуальну основу побудови обчислювальних кластерів у НАН України була прийнята концепція, відома як Beowulf [6], що базується на таких основних положеннях:

- 1) як сервери для виконання обчислень використовуються сервери з стандартною архітектурою PC, розподіленою оперативною пам'яттю і з стандартним мережевим обладнанням;
- 2) для об'єднання серверів як мережеве обладнання використовуються мережеві технології широкого вжитку, такі як Gigabit Ethernet;
- 3) як програмне забезпечення (операційна система й система керування завданнями) використовується вільно розповсюджене програмне забезпечення (Linux, Torque).

Враховуючи ці вимоги, всі обчислювальні кластери в НАН України і в КНУ побудовані на базі платформ x86 та x86_64, мають двох- і (чотирьох-) процесорні сервери з 1–4 ГБайтами оперативної пам'яті (RAM) і жорсткі диски (HDD) об'ємом 36-80 ГБайт. Для забезпечення міжсерверного обміну використана технологія Gigabit Ethernet з пропускною здатністю 1 Гбіт/с і тільки для двох кластерів використовується технологія InfiniBand з пропускною здатністю 2,5 Гбіт/с. Характеристики обчислювальних кластерів наведені в табл.1. Дисковий простір на жорстких дисках обчислювальних вузлів використовується для розміщення операційної системи (ОС) (завантаження виконується з локальних дисків), пакетів прикладних програм і тимчасових файлів, і він є не доступним користувачам для зберігання своїх файлів. Для зберігання програм і даних користувачів, а також даних загального користування, кожен кластер має дисковий масив (Network Attached Storage, NAS), характеристики якого наведені в табл.1. На всіх кластерах встановлена вільно розповсюджувана операційна система Linux, а для організації запуску завдань і розподілу навантаження на кластерах використовується система керування завданнями OpenPBS.

Інформація щодо технічних характеристик кластерів отримана з WWW-сайтів відповідних організацій. Наведені значення параметрів були доступні на момент написання статті. Поточні значення технічних параметрів, а також результати вимірювання продуктивності постійно змінюються і доступні на WWW-сайтах відповідних установ, а також частково на сайтах моніторингу <http://lcg.bitp.kiev.ua/loadmon.php> та <http://cluster.univ.kiev.ua/monitor/loadmon.php>.

Таблиця 1. Характеристики обчислювальних кластерів Грід-інфраструктури

	Кластер	Вузли	Процесори	RAM	HDD	Мережа	NAS	ОС
1	ІК (СКІТ-3)	75	2x Dual Core Intel Xeon 5160 CPU 3.0GHz	8 GB	80 GB	InfiniBand	20 TB	Linux CentOS -4.2
2	ІКС	36	17 nodes – 2 x Dual Core Intel Xeon 5130 2Ghz, 13 nodes – 2x AMD Athlon MP 2200+, CPU 1.8 GHz 6 nodes – 2 x AMD Opteron 246, CPU 1.8 GHz	4 GB	40 - 80 GB	InfiniBand Gigabit Ethernet	6 TB	Linux CentOS 4.0-4.4
3	КНУ	24	8 nodes-2xIntel Xeon CPU 2.40 GHz 7 nodes 2xIntel Xeon CPU 3.2 GHz 5 nodes 2xIntel Xeon 5130 2Ghz 4 nodes 2xIntel Pentium	512 Mb - 4 GB	36-72 GB	Gigabit Ethernet	6 TB	Fedora Core 1.x-3.x

			III CPU 1 GHz					
4	ІТФ	20	10 nodes – 2xIntel Xeon, CPU 2.80GHz 10 nodes – 2xIntel Xeon, CPU 3.2GHz	4 GB	36 GB	Gigabit Ethernet	12 TB	Scientific Linux 4 2.6.9-34.ELsmp
5	ГАО	8	2 x Dual Core Intel Xeon CPU 5130, CPU 2.00GHz	2 GB	80 GB	InfiniBand	3 TB	Debian GNU/Linux 4.0
6	ІКБГІ	8	2 x Intel Xeon, CPU 3.20GHz	4 GB	36 GB	Gigabit Ethernet	3 TB	Scientific Linux 4 2.6.9-34.ELsmp
7	ІМБГ	8	2 x Dual Core AMD Opteron, CPU 2.0 GHz	4 GB	80 GB	Gigabit Ethernet	1 TB	Scientific Linux 4 2.6.9-34.ELsmp

4. Інфраструктура каналів зв'язку

Наявність гарантованого, надійного й високошвидкісного каналу доступу до мережі Інтернет є однією з необхідних умов побудови Грід-середовища. Слід зазначити, що доступ до мережі Інтернет в інститутах Національної академії наук існував і раніше, але його швидкість була недостатня для обміну даними між обчислювальними кластерами. Це було основною перешкодою для побудови Грід-інфраструктури.

Слід зазначити, що на момент реалізації проекту Київський національний університет мав свою розгалужену мережу, що є однією з найбільших серед мереж навчальних закладів України. Локальна мережа КНУ поєднує будівлі в різних частинах міста Києва й побудована з використанням оптоволоконних каналів передачі даних із пропускною здатністю 1 Гбіт/с. Університет є учасником точки обміну Українським трафіком (UA-IX) і має зв'язок із нею на швидкості 1 Гбіт/с по оптоволоконному каналу. Університет підтримує паритетні Інтернет-канали з Інтернет-провайдерами м. Києва, побудовані на базі оптоволоконних з'єднань, зі швидкостями 1 Гбіт/с і 100 Мбіт/с. Закордонні інтернет-канали КНУ мають пропускну здатність 20 Мбіт/с.

На замовлення НАН України за останні два роки компанія UARNET (Ukrainian Academic and Research Network) побудувала інфраструктуру, що об'єднала академічні інститути оптоволоконними каналами зв'язку (пропускна здатність 100 Мбіт/с) з підключенням до магістрального каналу Львів – Київ – Харків (пропускна здатність 2,5 Гбіт/с) із виходом на Словаччину й Польщу (пропускна здатність 155 Мбіт/с).

Топологія мережі Національної академії наук є дворівневою. Перший рівень – це підключення оптоволоконними каналами зв'язку академічних і наукових інститутів академії в містах до точки обміну трафіком. Другий рівень – це підключення оптоволоконними каналами зв'язку точок обміну в містах до магістрального каналу Львів – Київ – Харків з наступним підключенням магістрального каналу до точки обміну трафіком закордонних партнерів у Польщі й Словаччині (рис.1).

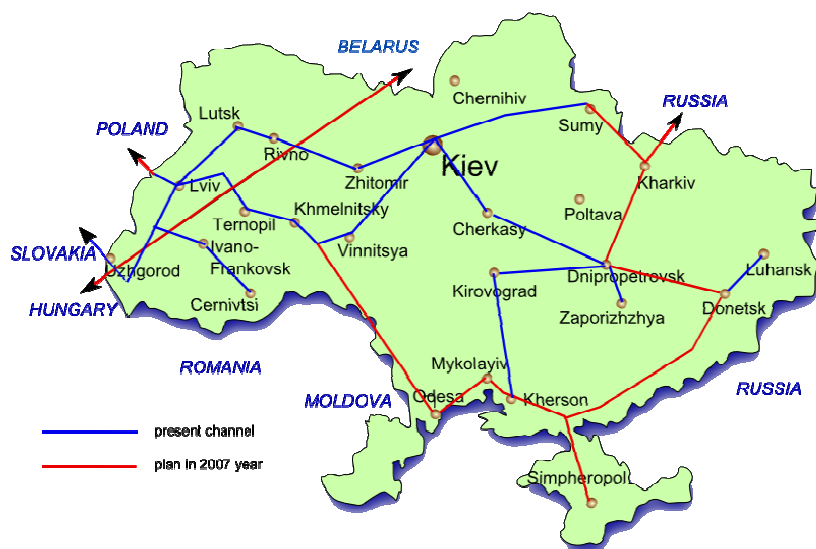


Рис. 1. Інфраструктура каналів зв'язку

Введення в експлуатацію оптоволоконного каналу, забезпечило для академічних інститутів надійний доступ до мережі Інтернет зі швидкістю від 4 Мбіт/с до 8 Мбіт/с з можливістю наступного збільшення пропускну здатності (зазначене обмеження визначається фінансовими можливостями інститутів по оплаті трафіка).

Для забезпечення надійного функціонування керуючих серверів Грід-системи Інститут теоретичної фізики має дві незалежні точки входу в мережу Інтернет: через академічну мережу UARnet і оптоволоконний паритетний канал із КНУ. У даний момент зовнішні канали для обміну між українськими учасниками Грід-системи не використовуються, тому що всі учасники сегмента, підключені до мережі UARnet, доступні через точку обміну UA-IX.

5. Проміжне програмне забезпечення

Третьою складовою Грід-інфраструктури є проміжне програмне забезпечення, що об'єднує обчислювальні ресурси в єдиний обчислювальний комплекс. Для побудови Грід-інфраструктури НАН України було використано програмне забезпечення проміжного рівня (middleware) ARC (Advanced Resource Connector), що також відоме під назвою проекту NorduGrid [7]. Вибір ARC як базового програмного забезпечення для побудови Грід-інфраструктури НАН України базується на таких положеннях:

- 1) ARC – це вільно розповсюджувана програмна система з відкритим кодом, що має вичерпну документацію по установці й підтримці;
- 2) ARC – це так званий «коробковий» варіант Грід-системи, який можна встановлювати без доробок і виправлень у програмному коді;
- 3) ARC не орієнтована на певну область прикладних обчислень (фізика, хімія), а являє собою загальносистемне програмне забезпечення для побудови Грід-систем;
- 4) ARC не вимагає спеціальних умов для установки (операційна система, система керування локальними завданнями й т.д.), що дозволяє встановити необхідні Грід-сервіси на вже працюючих обчислювальних ресурсах;

5) ARC дозволяє підключати нові обчислювальні ресурси без змін і порушень у роботі існуючого сегмента.

Крім того, ARC дозволяє без додаткових змін у майбутньому підключити Українську Грід-мережу до мережі в Данії, Естонії, Фінляндії, Норвегії й Швеції. Інші інструментальні засоби побудови Грід-систем, такі як gLite [8] та AliEn [5], мають засоби інтеграції з проміжним програмним забезпеченням ARC або досить легко можуть бути до цього адаптовані.

6. Архітектура ARC (Advanced Resource Connector)

Основні архітектурні рішення ARC слідують загальноприйнятим підходам побудови Грід. У ній використовується організація ресурсів, аналогічна тій, що застосовується у проекті EU DataGrid [8]. ARC забезпечує такі функції:

- 1) інформаційні (збір і надання інформації про ресурси Грід-системи);
- 2) динамічне включення ресурсів до Грід-системи і їхній моніторинг;
- 3) відправлення завдань на виконання в Грід-систему і керування завданнями;
- 4) розподіл завдань по ресурсах;
- 5) керування даними й ресурсами.

Всі функції реалізовані у вигляді служб, які спираються на відомі програмні засоби з відкритим кодом: OpenLDAP (Lightweight Directory Access Protocol), OpenSSL (Open Secure Socket Layer) і SASL (Simple Authentication Security Layer). Реалізація здійснена на основі бібліотек Globus Toolkit 2 (GT2), безпека досягається шляхом використання протоколів й інфраструктурних рішень GSI (Grid Security Infrastructure), основаних на надійній і поширеній інфраструктурі криптографії з відкритим ключем (Public Key Infrastructure - PKI).

Програмне забезпечення проміжного рівня ARC являє собою надбудову над інструментами Globus Toolkit 2 (GT2). Відмінною рисою NorduGrid ARC є те, що хоча ця платформа й опирається на протоколи GT2 і реалізована за допомогою API GT2, у ній запропонований власний набір служб, який замінює служби GT2. ARC не використовує інструменти зі складу Globus: GRAM (Globus Resource Allocation Manager) компоненти, відповідальні за створення / видалення процесів; утиліти керування завданнями; Gatekeeper і скрипти Job-manager, сервер Wuftp, схеми й інформаційні постачальники MDS (Monitoring and Discovery Service). Натомість ARC пропонує власні засоби: Grid Manager (сервіс керування завданнями користувачів); gridftp (сервіс – передачі файлів – ARC/NorduGrid GridFTP server); User Interface (інтерфейс користувача); Broker (планувальник завдань); систему моніторингу. Крім цього, впроваджено нову інформаційну схему, для неї розроблені постачальники даних, розширено мову опису ресурсів (xRSL). На рис. 2 представлені основні компоненти ARC і схематично показана взаємодія між ними. Нижче представлено короткий опис основних компонентів.

7. Інтерфейс користувача

Інтерфейс користувача (UI) є ключовим сервісом, що був розроблений у рамках проекту NorduGrid, він призначений для забезпечення запуску завдань користувачів у Грід-системі і керування виконуваними завданнями. Він являє собою набір команд для запуску, моніторингу й керування

завданнями, а також переміщення файлів і отримання інформації про стан ресурсів.

При розробці ARC переслідувалася мета створення програмного забезпечення виробничої якості, використовуючи як основний принцип максимально повну децентралізацію. Тому на кожному робочому місці користувача Грід-мережі встановлюється персональний брокер, функція якого – вибір найкращого ресурсу для виконання завдання користувача, яке необхідно виконати в Грід-мережі. Цей підхід відрізняє ARC від централізованої схеми EU DataGrid з єдиним брокером на всі робочі місця.

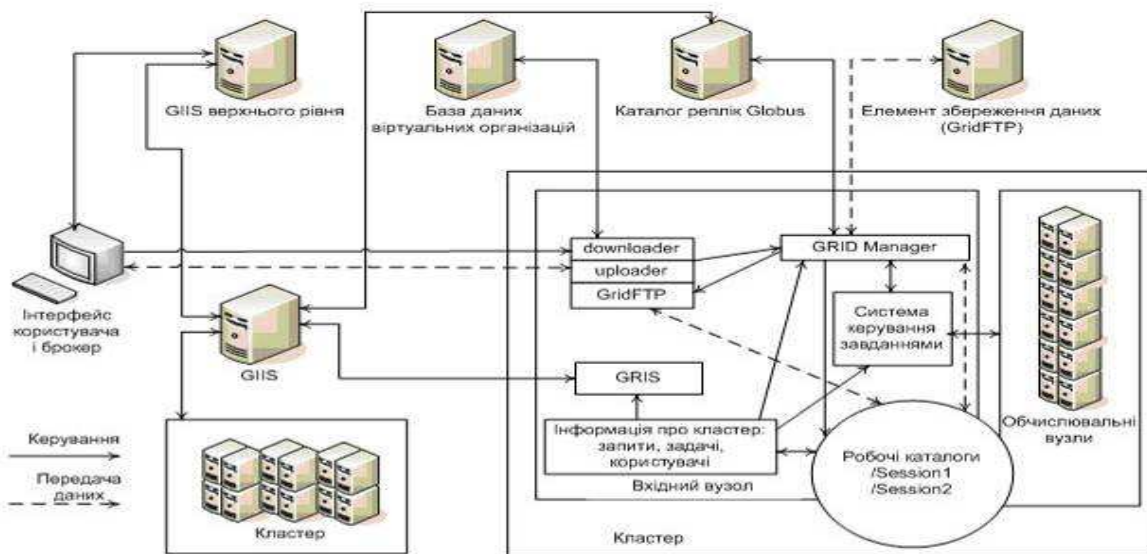


Рис. 2. Взаємодія компонентів ARC

8. Інформаційна система

Інформаційна система ARC є розподіленою динамічною системою, що надає інформацію про стан ресурсів Грід-середовища і використовує для роботи MDS (Monitoring and Discovery Service) – інформаційну систему Globus. Стандартний пакет MDS є відкритим інструментарієм для створення інформаційної системи для Грід і побудований на базі програмного забезпечення OpenLDAP. Інформація, що збирається інформаційною системою, може бути найрізноманітнішою і містити, наприклад, дані про конфігурацію або стан як всієї системи, так і окремих її ресурсів (тип ресурсу, доступний дисковий простір, кількість процесорів, об'єм пам'яті, продуктивність та інше). Вся інформація логічно організована у вигляді дерева, і доступ до неї здійснюється по стандартному протоколу LDAP – псевдорозподіленої бази даних.

Ієрархічна структура MDS представлена на рис. 3.

MDS складається із трьох основних компонентів:

1) IP (Information Provider) – є джерелом інформації про конкретний ресурс або частину ресурсу;

2) GRIS (Grid Resource Information Service) надає інформацію про вузол Грід-системи, що може бути як обчислювальним вузлом, так і яким-небудь іншим ресурсом. GRIS опитує окремі IP і поєднує отриману від них інформацію в рамках єдиної інформаційної схеми;

3) GIIS (Grid Index Information Service) поєднує інформацію з різних GRIS або інших GIIS. GIIS верхнього рівня містить всю інформацію про стан даної системи Грід.

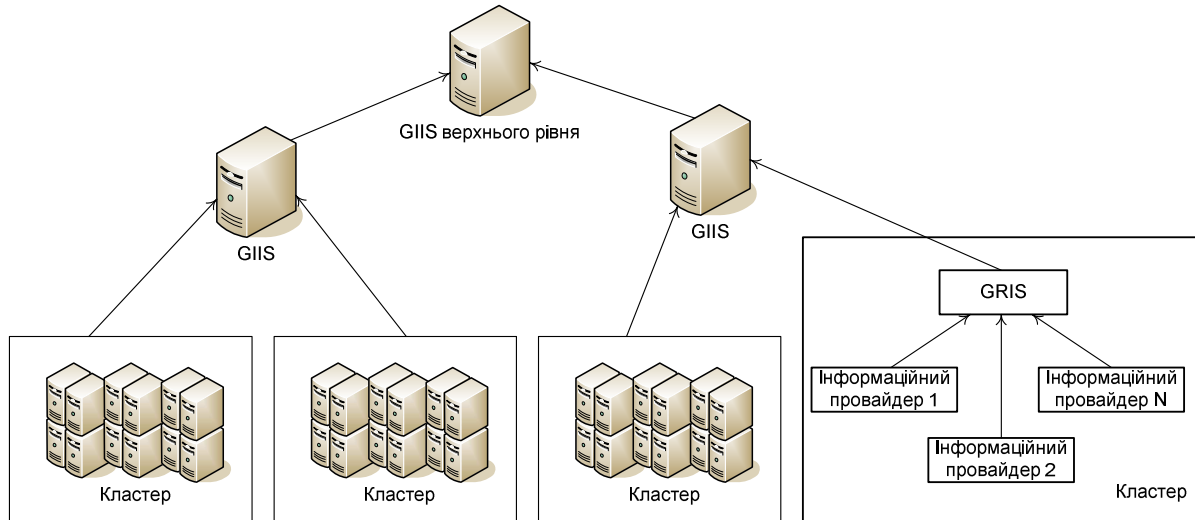


Рис. 3. Ієрархічна структура MDS

Інформація про стан ресурсів Грід-середовища (запису LDAP) генерується динамічно постачальниками інформації на ресурсі по запиту користувача або агента (pull-модель). Постачальники інформації – це спеціальні програми, які є інтерфейсом до ресурсів у Грід-середовищі і збирають про них інформацію. У складі ARC розроблений свій набір постачальників інформації для заповнення записів власної схеми. Далі згенерована інформація може кешуватися на різних рівнях розподіленої бази даних. У складі системи реалізовані такі елементи: локальні постачальники інформації, локальні бази даних (кеш першого рівня), індекси інформації (також з механізмом кешування – вищих рівнів), механізм м'якої реєстрації, а також інформаційна модель (схема) для представлення ресурсів у Грід. Інтерфейс користувача із вбудованим брокером і система моніторингу є головними споживачами динамічної інформації, що надається і зберігається в інформаційній системі. Локальні бази даних поєднані через індекси інформації GIIS, отримані з MDS. Використання динамічної реєстрації дає можливість будувати різні топології індексованих ресурсів. При побудові Грід-інфраструктури є доцільним використання декількох GIIS для забезпечення надійної та стійкої роботи. При такій реалізації система є більш відмовостійкою, і збій в роботі одного з вузлів не приведе до виходу з ладу всієї інфраструктури. На сьогодні функціонує два сервери індексації: в Інституті теоретичної фізики ім. Н.Н. Боголюбова (lsg.univ.kiev.ua) та в Інформаційно-обчислювальному центрі Київського національного університету імені Тараса Шевченка (alien.univ.kiev.ua).

На даний момент на виділеному сервері (lsg.bitp.kiev.ua) в ІТФ та на сервері Київського національного університету функціонують сервіси, дублюючи один одного:

GIIS – сервіс реєстрації обчислювальних ресурсів, що поєднує інформацію з GRIS, установлених на обчислювальних кластерах сегмента;

Certification Authority – сервіс керування сертифікатами користувачів і обчислювальних ресурсів;

Grid Monitor – система моніторингу Грід-інфраструктури, що періодично опитує розподілену інформаційну систему й представляє результати у вигляді взаємовкладених Web-сторінок.

Також на сервері Київського національного університету функціонує VOMS (Virtual Organization Membership Service), що надає можливість централізованого керування політиками доступу до ресурсів. На даний момент цей сервіс знаходиться на стадії тестування і не є впровадженим для всіх кластерів Грід-інфраструктури. Оновлення файлів відповідності з серверу VOMS налаштовано лише на кластері КНУ та кластері ІМБіГ НАН України. Найближчим часом планується впровадження його використання для всіх кластерів Грід-системи.

В ARC розроблена своя інформаційна модель, яка відображає обчислювальні ресурси, користувачів і черги у вигляді LDAP-дерева. Кластери описані об'єктивним класом LDAP `nordugrid-cluster`. Клас має набір атрибутів для опису властивостей обладнання, програмного забезпечення і проміжного програмного забезпечення кластера. Інформаційна модель розрізняє та підтримує кластери і черги як і ті, що виділені безпосередньо для Грід, так і змішаного використання. Кластер надає доступ до черг, виділених для завдань із Грід-мережі, які описуються класом `nordugrid-queue`. Це дерево містить піддерева `nordugrid-pauthuser` та `nordugrid-job` (поділ виконується класом `nordugrid-info-group`). Аналогічно, кожне `grid`-завдання, запущене в чергу, відображається відповідним записом. Запис `nordugrid-authuser` містить всю інформацію щодо авторизованого користувача Грід, навіть таку, як кількість процесорів, виділених для цього користувача, в даній черзі, обсяг доступного дискового простору і т.д. В ARC зроблений акцент на інформацію, що стосується користувача, тому що саме ця інформація становить більший інтерес для нього, ніж загальна інформація про ресурси кластера.

Записи `nordugrid-job` описують завдання, запущені для виконання на кластері. Докладна інформація про завдання включає: унікальний ідентифікатор у Грід-мережі, заголовок сертифіката власника завдання, статус завдання, обчислювальний вузол, на якому виконується завдання, і т.д. Записи про завдання зберігаються на кластері, на якому воно виконується. Додатково схема включає класи `nordugrid-se` та `nordugrid-rc` для опису сховищ даних і каталогів реплік відповідно. При розробці ARC було прийнято рішення щодо блокування функції кешування в GHS. У підсумку користувачі або агенти одержують результати запитів безпосередньо від ресурсів, а GHS виконує тільки функції індексації.

9. Обчислювальні кластери (Computing Clusters)

Ресурси в Грід-інфраструктурі НАН України під керуванням програмного забезпечення ARC формуються з окремих обчислювальних кластерів. Кластер складається із вхідного (інтерфейсного) вузла, що керує деякою кількістю робочих вузлів, використовуючи внутрішню локальну мережу. Програмне забезпечення містить у собі стандартну систему пакетної обробки і додаткові компоненти для взаємодії з Грід-мережею. Важливою властивістю архітектури ARC є той факт, що для робочих вузлів кластера не потрібен прямий доступ до Інтернет. Цей принцип реалізований відповідно до тієї вимоги, що ARC реалізовано як додатковий компонент, який впроваджує локальні

ресурси в Грід-мережі, і дозволяє завданням, що надійшли на кластер, працювати поряд з локальними завданнями.

Тому на вхідному вузлі кожного кластера, підключеного до мережі Інтернет, встановлені сервіси:

Grid Manager (Грід-менеджер) – сервіс керування завданнями користувачів. Grid Manager дозволяє користувачеві відправляти завдання в Грід-мережу і одержувати результати розрахунків;

GridFTP Server – сервіс, що забезпечує доступ до файлової системи. GridFTP Server забезпечує FTP-подібний доступ до локальних файлів і забезпечує інтерфейс для контролю над завданням. Крім того, GridFTP підтримує віртуальне дерево каталогів, що налаштовується для кожного користувача, і контроль доступу, що ґрунтується на імені в сертифікаті користувача;

Downloader and Uploader – сервіси передачі пакетних завдань і результатів розрахунків на комп'ютер користувача.

Грід-менеджер служить інтерфейсом для запуску завдань у Грід-мережу та, крім керування завданнями, керує переміщенням вхідних і вихідних даних з підтримкою використання каталогів метаданих. Дані обробляються тільки до і після завершення завдань. Передбачається, що користувач зазначає повну інформацію про вхідні дані, які необхідні для виконання завдання, і вихідні дані, які є результатами виконання завдання. Це найбільш істотне обмеження даного підходу. Грід-менеджер був розроблений в рамках проекту ARC. Він заснований на бібліотеках і сервісах Globus Toolkit і використовує такі елементи Globus: GridFTP-сервіс, що забезпечує швидкий і надійний доступ до даних у Грід з вбудованим механізмом безпеки; GASS-сервіс віддаленого доступу до додаткових елементів збереження по протоколах http й https; каталог реплік – сховище метаданих; сервіс розташування реплік – створений у співпраці Globus та EDG; мова специфікації ресурсів RSL (resource specification language). В ARC використовується xRSL (extendable resource specification language). Грід-менеджер реалізує децентралізовану модель доступу до ресурсів Грід, що означає відсутність центрального сервера, через який повинні проходити всі завдання в Грід-мережу.

Грід-менеджер реалізований в основному мовою програмування C/C++. Використання реалізації інтерфейсу за допомогою сценаріїв оболонки bash з локальною системою керування завданнями на кластері дозволяє адаптувати Грід-менеджер для роботи з будь-якою ЛСКЗ. Використання файлової системи для зберігання станів завдань, що обробляються, дозволяє Грід-менеджеру легко відновлювати поточний стан після системних помилок або після перезапуску. До його складу також включені утиліти користувача для здійснення передачі даних, реєстрації в каталогах метаданих (команда `ngtmove`) і видалення даних (команда `ngremove`).

На основі GridFTP був розроблений GridFTP-сервер (GFS), що підтримує такі можливості: віртуальне дерево каталогів, сконфігуроване для кожного користувача; контроль доступу за сертифікатами; доступ до локальної файлової системи побудований на основі двох бібліотек. Одна забезпечує локальний доступ до файлів, що реалізує звичайний доступ FTP-сервера; друга – забезпечує інтерфейс до функцій запуску і керування завданнями, що виконуються Грід-менеджером. GFS також використовується для створення серверів зберігання даних (storage element), заснованих на GridFTP, що відносно легко конфігуруються.

10. Виконання завдань в ARC

Для використання ресурсів Грід-мережі на комп'ютерах користувачів встановлюється спеціальне програмне забезпечення (User Interface – UI), що забезпечує взаємодію з Грід-середовищем. Користувач взаємодіє з ARC за допомогою утиліт командного рядка. ARC надає команди для надсилання завдань у Грід-мережу, запиту стану завдань і кластерів, отримання даних щодо завершених завдань, переривання роботи своїх завдань і т.д. Також існують засоби для копіювання та видалення файлів у сховищах даних і каталогах реплік. У табл. 2 наведений перелік основних команд та їх короткий опис.

Таблиця 2. Перелік основних команд ARC

Команда	Операція
ngsub	Запуск завдання
ngstat	Перегляд стану кластерів і статусу завдань
ngcat	Перегляд повідомлень та помилок завдань
ngget	Отримання вихідних файлів після завершення завдання
ngkill	Видалення завдання, що виконується
ngclean	Видалення вихідних даних завдання з кластера
ngmove	Копіювання файлів між сховищами даних, каталогами реплік
ngremove	Видалення файлів зі сховищ даних та каталогів реплік

Перед запуском завдання командою ngsub користувач повинен описати завдання мовою опису завдань xRSL (Extended Resource Specification Language). Опис має містити всю необхідну для запуску інформацію (виконавчий файл, параметри і т.д.), а також набір вимог, якому повинен відповідати кластер, щоб виконати завдання (вільний дисковий простір, наявне програмне забезпечення), і інформацію про те, де повинні зберігатися результати виконання.

Типовий xrsl-файл має вигляд:

```
&
(* this is comment *)
(executable=ex.sh)
(executables=example1)
(inputFiles=(example1 ""))
(arguments="100000000" "13" "0.324")
(stdout="out.txt")
(stderr="err.txt")
(outputFiles=("out.txt" "")("err.txt" "")("sol.ps" "")("err.ps" "")("data.txt" ""))
(gmlog="gridlog")
(jobname="Example")
(cputime=20)
(middleware>="nordugrid-arc-0.3.24")
```

У наведеному прикладі зазначено, що як вхідне завдання використовується файл сценарію ex.sh, а результати записуються у стандартний вихідний файл і файл реєстрації помилок, які потім будуть передані на комп'ютер користувача. Інтерфейс користувача проводить відповідність між вимогами, зазначеними в xRSL, і інформацією, що була отримана від кластерів, і обирає ресурс, який найкраще відповідає вимогам. Якщо кластер, що задовольняє вимогам xRSL, знайдений, завдання відправляються на кластер і стають у чергу для виконання.

ARC – версія мови (xRSL) побудована на основі мови специфікації ресурсів RSL, розробленої у рамках Globus. В архітектурі ARC мова специфікації ресурсів має подвійне призначення: окрім опису завдання користувачем вона описує комунікацію між інтерфейсом користувача і Грід-менеджером. Тому, через відмінності рівнів опису завдань, xRSL можна розділити на дві частини:

RSL користувача – набір атрибутів, зазначених користувачем у файлі опису завдання. Цей файл інтерпретується інтерфейсом користувача і після необхідних модифікацій передається до Грід-менеджера.

RSL Грід-менеджера – набір атрибутів попередньо оброблених інтерфейсом користувача, готових для інтерпретації Грід-менеджером. Досить зручний для опису завдань.

Такий функціональний розподіл, а також впровадження та перевизначення атрибутів, призвело до створення нової модифікації мови – Extended Resource Specification Language.

Одне з найбільш важливих завдань застосувань ARC – переміщення великої кількості файлів, найчастіше великого розміру, як до запуску, так і після завершення завдання. Для цього в xRSL були застосовані два атрибути: `inputFiles` та `outputFiles`, кожен з яких є списком пар «локальне ім'я файлу – віддалене ім'я файлу (або URL-адреса)». Локальні вхідні файли вузла запуску завдання завантажуються на виконавчий вузол інтерфейсом користувача, після чого контроль передається Грід-менеджеру.

Вихідні файли по завершенні завдання переміщуються в зазначене місце Грід-менеджером. Якщо місце не зазначене, то передбачається, що користувач сам виконає копіювання цих файлів за допомогою інтерфейсу користувача. Для будь-якого з вхідних файлів можна встановити дозвіл на виконання, якщо вказати його в списку атрибута `executables`. Як місце збереження вихідного файлу можна задати псевдо-URL каталогу реплік, у якому вихідний файл буде зареєстрований.

Використовуючи атрибут `cluster`, користувач може явно вказати або виключити список кластерів з переліку тих кластерів, серед яких буде проводитися пошук відповідності ресурсів. Використовуючи атрибут `stdlog`, користувач може переглянути повний лог завдання Грід-менеджера. Стандартний вивід завдання та звіт про помилки можуть бути записані в один файл, якщо зазначений атрибут `join`. Для того, щоб проводити пошук відповідності на основі конфігурації кластера, використовуються атрибути `middleware`, `architecture`, `runTimeEnvironment` і т.д.

11. Безпека

Інфраструктура безпеки Грід (Grid Security Infrastructure – GSI) забезпечує безпечну роботу в незахищених мережах загального доступу (Інтернет), надаючи такі сервіси, як аутентифікація, конфіденційність передачі інформації і єдиний вхід у Грід-систему. Під єдиним входом розуміють таке: користувачеві потрібно лише один раз пройти процедуру авторизації, а далі система сама подбає про те, щоб аутентифікувати його на всіх ресурсах, якими він збирається користуватись. GSI заснована на надійній і широко використовуваній інфраструктурі криптографії з відкритим ключем (Public Key Infrastructure – PKI).

Як ідентифікатори користувачів і ресурсів в GSI використовуються цифрові сертифікати X.509. У роботі із сертифікатами X.509 і у процедурі видачі/одержання сертифікатів учасниками є три сторони: Центр сертифікації (Certificate Authority – CA) – спеціальна організація, що володіє повноваженнями видавати (підписувати) цифрові сертифікати. Передплатник – це людина або ресурс, що користується сертифікаційними послугами CA. CA включає в сертифікат дані, надані передплатником (ім'я, організація й ін.) і ставить на ньому свій цифровий підпис. Користувач – це людина або ресурс, що покладається на інформацію, яка міститься в сертифікаті при одержанні

його від передплатника. Користувачі можуть приймати або відкидати сертифікати, підписані яким-небудь СА.

В Globus Toolkit використовуються два типи сертифікатів X.509:

Сертифікат користувача (User Certificate). Цей сертифікат повинен мати кожен користувач, що працює з Грід-системою. Сертифікат користувача містить ідентифікатор власника і відповідний відкритий ключ, термін дії, призначення сертифіката та інформацію про центр сертифікації, що надав даний сертифікат.

Сертифікат вузла (Host Certificate). Цей сертифікат повинен мати кожен вузол (ресурс) Грід-системи. Сертифікат вузла аналогічний до сертифіката користувача, але в ньому, замість імені користувача, вказується доменне ім'я конкретного обчислювального вузла. Технологія видачі сертифікатів містить неформальний момент, пов'язаний з ідентифікацією особистості власника і верифікацією його атрибутів, що припускає особистий контакт. Підписаний сертифікат разом з ключами користувач зберігає в себе. Користувачі в Грід-середовищі, як правило, не мають локальних облікових записів на обчислювальних ресурсах, які хочуть використати, але володіють сертифікатом, виданим Центром сертифікації. Авторизація (надання користувачеві прав на доступ до ресурсу) виконується відповідно до результатів аутентифікації і локальної політики авторизації, встановленої для даного ресурсу. Системні адміністратори кластерів зберігають повний контроль за вибором тих Грід-користувачів, кому надається доступ до ресурсу. Локальний процес авторизації виконується шляхом переліку у файлі (grid-mapfile) відповідностей між користувачами Грід і локальними обліковими записами.

Зараз Український Тестовий Центр сертифікації перебуває в Інституті теоретичної фізики НАН України та керує сертифікатами Грід НАН України. Також деякі користувачі системи мають сертифікати, підписані Центром сертифікації Київського національного університету. Слід зазначити, що всі ці сертифікати на даний момент не приймаються міжнародними Грід-системами і використовуються тільки в межах Грід-сегмента НАН України. Побудова Центру сертифікації європейської системи сертифікації є актуальним завданням, рішення якого планується на наступному етапі розвитку проекту.

12. Моніторинг

ARC надає зручний у використанні інструмент моніторингу, реалізований у вигляді web-інтерфейсу до інформаційної системи ARC. Грід-монітор дозволяє переглядати всю опубліковану інформацію про систему, забезпечуючи моніторинг у реальному часі.

Інформаційна система ARC забезпечує стійку динамічну модель для доступу як до квазістатичної інформації про ресурси та сервіси, так і до динамічної – про статус черг і завдань. Структура Грід-монітора багато в чому схожа на структуру інформаційної системи. Базові об'єкти визначаються такими класами схеми:

nordugrid-cluster: кластер;

nordugrid-queue: черга на кластері, доступна grid-користувачам;

nordugrid-job: Грід-завдання у відповідній черзі;

nordugrid-authuser: користувач, авторизований для запуску завдань у дану чергу.

Для кожного класу основна підмножина атрибутів (або всі атрибути) представлена в зручному зв'язному вигляді – за допомогою набору вікон, кожне з яких відповідає своєму модулю. У Грід-моніторі можна виділити сім основних модулів:

- загальний Грід-монітор;
- опис обчислювального ресурсу (кластера);
- інформація про стан черги на локальному ресурсі;
- інформація про завдання користувача;
- інформація про користувача;
- список атрибутів;
- список користувачів.

На рис. 4 наведено вигляд головного вікна Грід-монітора з web-сторінки сегмента Грід-інфраструктури НАН України (<http://lcg.bitp.kiev.ua/loadmon.php>). Також система моніторингу працює на кластері КНУ (<http://cluster.univ.kiev.ua/monitor/loadmon.php>). Більшість об'єктів головного вікна являють собою гіпер-посилання для доступу до інформації відповідних модулів системи моніторингу. Грід-монітор є зручним, надійним і необхідним інструментом для користувачів і системних адміністраторів. Web-сервер, що забезпечує роботу Грід-монітора, запущений на окремому, незалежному вузлі і, окрім частих LDAP-запитів, не збільшує навантаження системи ARC.

Особливо варто зупинитися на процедурі встановлення модулів ARC на обчислювальних ресурсах. Досвід інсталяції програмного забезпечення ARC показав, що воно відповідає концепції «коробкового варіанта» і являє собою легке неагресивне рішення, що підходить для будь-яких доступних ресурсів і потребує лише мінімального втручання з боку системних адміністраторів. Причому, інсталяція виконувалася як інсталяційним пакетом, так і повною збіркою з вихідних модулів. І це не викликало більших ускладнень у адміністраторів кластерів, не всі з яких є експертами в Грід-технологіях. Для зручності налаштування параметрів сервісів їх конфігурація в ARC була максимально спрощена, що в остаточному підсумку звелось до модифікації файлів `arc.conf` (налаштування сервісів ARC).

13. Розв'язування задач у Грід-середовищі НАН України

Побудова Грід-інфраструктури НАН України мала за мету створення потужного обчислювального територіально розподіленого комплексу, який би надавав вільні ресурси для розв'язування складних наукових завдань у різних прикладних областях. За своєю суттю це дуже різноманітні завдання. Відзначимо коротко лише основні напрямки наукових досліджень, які ведуться в інститутах та університеті.

Інститут теоретичної фізики НАН України. Обробка і аналіз експериментальних даних про андрон – андронні та ядерно – ядерні зіткнення на прискорювачі LHC (ALICE, ATLAS і TOTEM експерименти) і RHIC; порівняння теоретичних моделей виникнення кварк-глюонної плазми з майбутніми експериментальними даними.

Інститут молекулярної біології і генетики НАН України. Моделювання білків і розрахунки молекулярної динаміки ряду білків, які можуть бути мішенями для розробки нових лікарських

препаратів, у тому числі ВІЛ-протеази і її мутантів, резистентних до інгібіторів.

Інститут клітинної біології і генної інженерії НАНУ. Аналіз біологічних послідовностей, моделювання просторової структури та поведінки біологічних макромолекул, моделювання процесів специфічної взаємодії макромолекул між собою та з низькомолекулярними речовинами.

Головна астрономічна обсерваторія НАНУ. Дослідження спектрів зірок і суб-зіркових об'єктів, які перебувають на різних еволюційних стадіях; дослідження фізики активних процесів на поверхні кометних ядер; моделювання процесів формування та еволюції великомасштабних структур всесвіту, таких як галактики і ядра галактик (чорні дірки), скупчення галактик, процесів утворення і еволюції зірок.

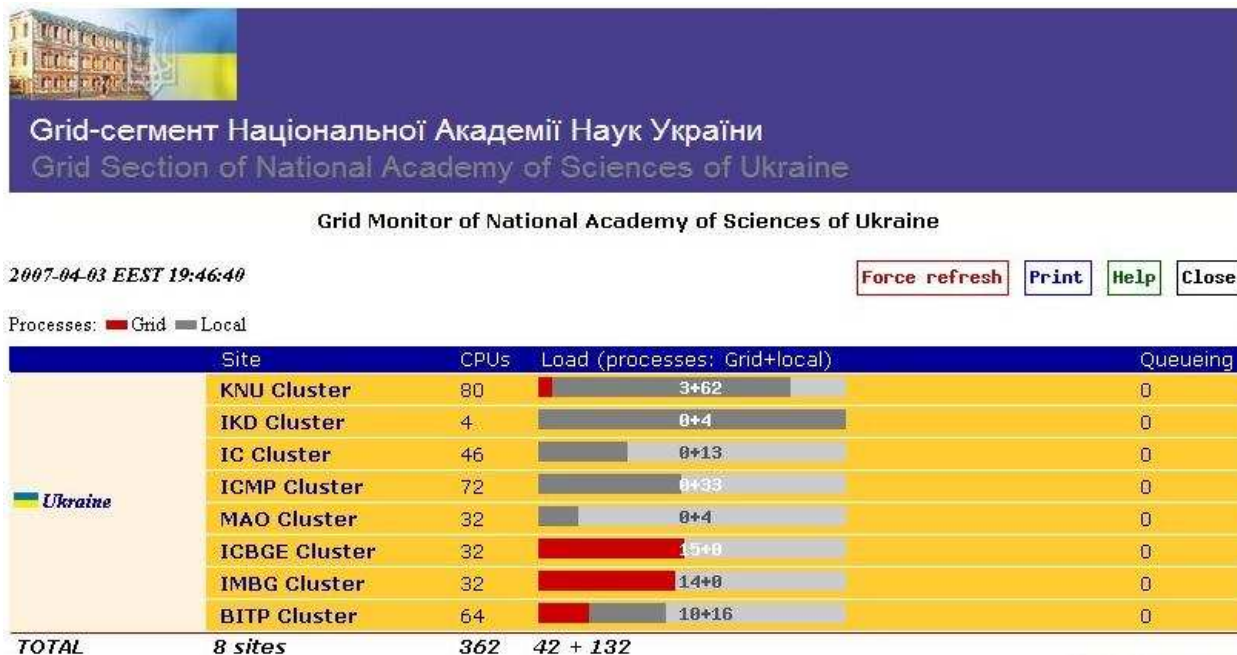


Рис. 4. Система моніторингу

Київський національний університет імені Тараса Шевченка. Квантово-механічні і напівкласичні розрахунки в області фізики поверхні, мікроелектроніки, нанотехнологій. Моделювання алгоритмів маршрутизації на графах. Дослідження з теорії груп. Дослідження з фізики твердого тіла, фізики плазми. Обробка сигналів.

Уже цей неповний перелік добре ілюструє той факт, що немає можливості використати єдине прикладне програмне забезпечення, яким можна було б скористатися для розв'язання цих завдань. Кожен науковий колектив використовує унікальне програмне забезпечення, що є загальноновизнаним у даній області, так як воно розроблено власними силами. Тому в Грід-інфраструктурі НАН України проміжне програмне забезпечення ARC використовується для забезпечення загальносистемних служб: відправлення завдань у Грід-мережу, передача файлів на обчислювальний кластер, відпрацювання запиту про стан завдання, моніторинг ресурсів, одержання даних від завершених завдань, аварійне завершення завдань. Прикладне програмне забезпечення встановлене на конкретних кластерах і не передається по мережі. Слід зазначити, що сервіс вилучення інсталяції програмного забезпечення в цій версії ARC відсутній.

На стадії тестування Грід-інфраструктури НАН України завдання користувача являє собою

повністю готовий до виконання модуль із необхідними даними, що зберігаються в окремому файлі. Даний модуль для виконання не вимагає спеціальних бібліотек, які потрібно динамічно інсталювати на обчислювальних кластерах. Природно такий підхід обмежує використання Грід-системи. Один із шляхів рішення цієї проблеми, що був випробуваний у результаті тестування, це попередня установка спеціалізованого програмного забезпечення на декількох кластерах. Інший варіант вирішення цієї проблеми – написання сценаріїв, які б попередньо інсталювали необхідні бібліотеки або збирали необхідні застосування з вихідних кодів, але він потребує від користувача знання в області програмування та написання сценаріїв.

ARC-версія мови (xRSL) – мови специфікації ресурсів RSL, виявилася дуже зручним засобом для реалізації такого підходу до використання обчислювальних ресурсів. Завдання атрибутів: `inputFiles` й `outputFiles`, кожний з яких є списком пар «локальне ім'я файлу – вилучене ім'я файлу(або URL-адреса)», дозволило використати передачу великої кількості локальних вхідних файлів, найчастіше великого розміру, до запуску завдання на виконавчий вузол. Вихідні файли по завершенні завдання переміщуються на вузол запуску завдання Грід-менеджером.

Як місце збереження вихідного файлу можна задати псевдо-URL каталогу реплік, у якому вихідний файл буде зареєстрований. Дуже корисним є атрибут `cluster`, що явно описує список кластерів, серед яких буде проводитися пошук необхідного ресурсу, із установленим прикладним програмним забезпеченням.

Саме використовуючи перший шлях, були проведені розрахунки по обробці даних рентгенівського випромінювання, одержані від міжнародного наукового супутника «Інтеграл» (проект Лозанського університету, Швейцарія) на базі пакетів XMM SAS (Scientific Analysis Software) і INTEGRAL OSA (Offline Science Analysis). Це дозволило використати пакетну обробку даних на декількох кластерах, які були менш завантажені локальними завданнями.

Іншим прикладом ефективного використання Грід-системи є розв'язування завдань молекулярного моделювання, що включають у себе комп'ютерну розробку лікарських препаратів, вивчення фізичних процесів у біологічних системах на молекулярному рівні і зв'язок цих процесів з біологічними функціями низько- та високомолекулярних з'єднань. Моделювання виконується методами квантової хімії й молекулярної механіки за допомогою програмного забезпечення GAMESS [9], GROMACS, Flo/qxp [10]. Зазначені програми не розраховані на роботу в Грід-середовищі. Тому для їхньої адаптації були розроблені спеціальні сценарії запуску завдань, які дозволяють звертатися до проміжного програмного забезпечення для доступу до файлових ресурсів і передавати необхідні дані на розрахункові кластерні системи. Необхідне прикладне програмне забезпечення попередньо інсталювалось у тимчасове сховище на обчислювальних кластерах. Тестування показало, що загальний час виконання завдання при такій схемі обчислень критично залежить від часу передачі даних між обчислювальними вузлами Грід при запуску і завершенні завдань.

Одним із прикладів ефективного використання Грід-системи НАН України є високопродуктивний молекулярний докінг за допомогою програмного забезпечення Flo/qxp. Це завдання ефективно розпаралелюється за схемою SIMD (Single Instruction Multiple Data), коли на кожному вузлі кластерів Грід-системи запускається та сама копія програмного забезпечення, що

працює зі своїми даними. Даними для завдання є частина бази з'єднань, які використовуються для докінгу на вказаний користувачем центр комунікації. Комунікаційні затримки при запуску такого завдання пов'язані тільки із завантаженням частини бази на вузол і зворотну передачу результатів. База даних з 1.000.000 з'єднань, що використовувалась для тестування, має обсяг 670 Мбайт. Для розрахунків у Грід-середовищі було підготовлено 480 незалежних завдань, кожне з яких обробляло свою частину бази даних. Всі завдання були запущені в Грід-систему одночасно. Загальна кількість процесорів, виділених у Грід для тестування, становила 120. Таким чином, на кожен процесор у середньому доводилося по 4 завдання для виконання оптимізації геометрії 2000 з'єднань. Повний час обчислення склав приблизно 73 години, що в середньому складає майже 2 хвилини на розрахунок одного з'єднання. Для порівняння подібний розрахунок на кластері КНУ вимагає порядку 2-х тижнів при повністю завантаженому кластері.

Таким чином, основну перевагу використання Грід-системи на даному етапі одержують користувачі, що обробляють великі масиви даних, завдання яких можуть бути ефективно розпаралелені за схемою SIMD. Прикладами таких додатків є завдання молекулярної динаміки (розрахунок великої кількості траєкторій для статистичного аналізу), алгоритми аналізу апаратних даних (обробка фотографічних зображень, отриманих від космічних станцій) і медичні діагностичні дані.

14. Висновки

Експериментальна Грід-система Національної академії наук України працює з січня 2007 року. Починаючи з двох обчислювальних кластерів ІТФ та КНУ, зараз вона поєднує 8 кластерів, розташованих у Києві та Львові. В 2007 році планується підключити обчислювальні кластери Радіоастрономічного інституту НАН України та Інституту сцинтиляційних матеріалів НАН України (м. Харків), Інституту геотехнічної механіки НАН України (м. Дніпропетровськ), Інституту фізики НАН України (м. Київ). Ріст обчислювальних ресурсів, збільшення кількості розв'язуваних задач поставлять нові завдання по забезпеченню стабільності роботи, надійності, моніторингу, навчанню користувачів, адаптації прикладного програмного забезпечення для Грід-розрахунків. Однак уже перші експерименти по використанню Грід НАН України показують, що українські наукові інститути мають достатній потенціал для розв'язування цих завдань.

Грід-система Національної академії наук України є добровільним об'єднанням обчислювальних ресурсів для спільного використання і відкрита для підключення нових кластерів. Кількість сайтів в Грід-системі може змінюватися з часом: будуть додаватися нові сайти, а окремі (тимчасово) припиняти свою роботу з різних причин (ремонтні роботи, профілактичні роботи та ін.). Однак здобутий досвід побудови розподіленої обчислювальної системи дозволить досить швидко підключити нові обчислювальні ресурси, які створені в Україні в рамках різних наукових проектів (наприклад, потужний обчислювальний кластер Національного технічного університету України "КПІ"), у разі, якщо нові учасники готові надати свої ресурси для використання.

Національна академія наук України знаходиться на самому початку шляху створення національного Грід-співтовариства і найближчою перспективою є включення України в офіційні члени Європейського об'єднання EGEE. Це відкриє для інститутів НАН України нові перспективи

міжнародного співробітництва в розвитку Грід-технологій і їх використання в нових наукових дослідженнях у різноманітних сферах.

СПИСОК ЛІТЕРАТУРИ

1. The Data Grid: Towards an Architecture for the Distributed Management and Analysis of Large Scientific Datasets / A. Chervenak, I. Foster, C. Kesselman et al // Journal of Network and Computer Applications (based on conference publication from Proceedings of NetStore Conference, 1999). – 2001. – Vol. 23. – P. 187–200.
2. Судаков О.О., Бойко Ю.В. GRID ресурси інформаційно-обчислювального центру Київського національного університету імені Тараса Шевченка // Проблеми програмування. Матеріали 5 міжнародної науково-практичної конференції з програмування УкрПРОГ. Київ, 23–25 травня 2006 р. – Київ, 2006. – С. 165–169.
3. Kyiv National Taras Shevchenko University High Performance Computing Cluster / Yu.V. Boyko, O.O. Vystoropsky, T.V. Nychyporuk et al // Proc. Third International Young Scientists Conference on Applied Physics. Kyiv, Ukraine, June 18–20 2003. – Kyiv, 2003. – P. 180–181. <http://www.cluster.kiev.ua>.
4. Ukrainian Grid Infrastructure: Practical Experience / M. Zynovyev, S. Svistunov, O. Sudakov et al // Proc. 4-th IEEE Workshop IDAACS 2007. Dortmund, Germany, September 6–8, 2007. – Dortmund, Germany, 2007. – P. 165–169.
5. Saiz P., P. and Others AliEn – ALICE environment on the GRID, Nucl. Instrum. Meth., A502. – 2003. – P. 437–440.
6. Gropp W., Lusk E., Sterling T. Beowulf Cluster Computing with Linux. – 2nd Edition. – MITPress. – 2003. – 504 p. <http://www.beowulf.org>.
7. Ellert M. et al Advanced Resource Connector middleware for lightweight computational Grids // Future Generation Computer Systems. – 2007. – Vol. 23. – P. 219–240. <http://www.nordugrid.org>.
8. Data Management in an International Data Grid Project / W. Hoschek, J. Jaen-Martinez, A. Samar et al // IEEE/ACM International Workshop on Grid Computing Grid'2000. – 2000. – 17–20 December. – P. 77–90.
9. General Atomic and Molecular Electronic Structure System / M.W. Schmidt, K.K. Baldridge, J.A. Boatz et al // J. Comput. Chem. – 1993. – Vol. 14. – P. 1347–1363.
10. McMartin C., Bohacek R.S.J. Powerful, Rapid Computer Algorithms for Structure-Based Drug Design // J. Comput.-Aided Mol. – 1997. – Vol. 11. – P. 333–344.

Стаття надійшла до редакції 15.11.2007