

**НАЦІОНАЛЬНА GRID-ІНФРАСТРУКТУРА ДЛЯ
ЗАБЕЗПЕЧЕННЯ НАУКОВИХ ДОСЛІДЖЕНЬ І ОСВІТИ**

А.І. ПЕТРЕНКО

Проведено аналіз концепції Grid, архітектури та ресурсів сучасних Grid-систем на прикладі побудови національної Grid-інфраструктури для наукових досліджень і освіти.

ВСТУП

Grid-технології і всесвітня Grid-мережа ідуть на зміну вже звичному Інтернету з його web-послугами як засіб сумісного використання обчислювальних потужностей та сховищ даних. Grid дозволяє вийти за рамки простого обміну даними між комп'ютерами і, зрештою, перетворити їхню глобальну мережу на свого роду гігантський віртуальний комп'ютер, доступний у режимі віддаленого доступу з будь-якої точки, незалежно від місця розташування користувача. Сфера застосування технологій Grid не обмежується лише складними науковими та інженерними задачами. Із розвитком Grid у промисловість проникає і бізнес, претендуючи на роль універсальної інфраструктури для обробки даних із безліччю служб (Grid Services), які не лише дозволяють розв'язувати конкретні прикладні задачі, а й пропонують послуги з пошуку необхідних ресурсів, збору інформації про стан ресурсів, зберігання і доставки даних.

Grid є технологією забезпечення гнучкого, безпечного і скоординованого загального доступу до ресурсів. При цьому слово «ресурс» розуміється у дуже широкому значенні, тобто ресурсом може бути апаратура — жорсткі диски, процесори, а також системне і прикладне програмне забезпечення (ПЗ) — бібліотеки, додатки.

Якщо перекладати дослівно, Grid означає «грати». Ближче всього за значенням, мабуть, power grid — мережа електроживлення, розподілений ресурс загального користування, коли кожний може легко підключитися через розетку і вжити електрики, скільки він потребує. Аналогічно користувачі комп'ютерних потужностей дістають можливість прямого підключення до віддаленої обчислювальної мережі, не цікавлячись, звідки саме приходять обчислювальні ресурси, необхідні для роботи, які для цього використовуються лінії передачі, паролі або протоколи і т.п. При цьому аналогом інфраструктури електричних мереж (ліній електропередач, підстанцій,

трансформаторів і т.ін.) виступає Grid програмне забезпечення (ПГЗ), або Middleware (рис. 1).

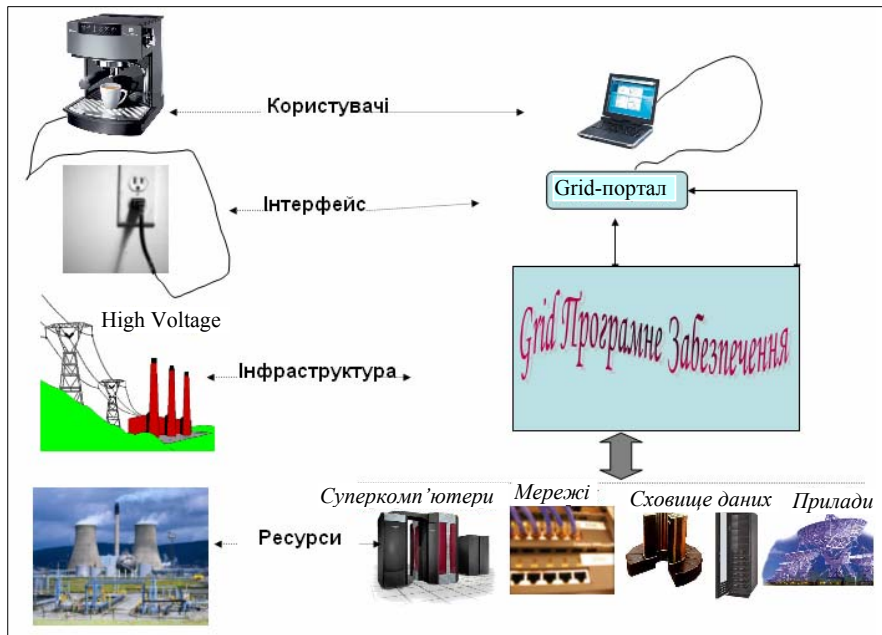


Рис. 1. Аналогія електро- і Grid-мереж

Розвиток ПГЗ починався від базових засобів, які підтримують дистанційний доступ до ресурсів, пройшов стадію окремих систем, їх пакетів і привів до створення **платформ** — взаємоузгоджених наборів засобів, здатних дати комплексне рішення задачі обслуговування Grid-інфраструктури виробничого призначення (рис. 2).

Створення багатьох сучасних Grid-інфраструктур супроводилося розробкою власного ПГЗ, тому зараз є платформи, які дещо розрізняються: ARC(NorduGrid), Alien, LCG, DataGrid (www.datagrid.org), Unicore (www.unicore.org), gLite (www.glite.org). Проте вони мають багато спільного, оскільки в тому або іншому ступені засновані на одній базі — системі Globus Toolkit (www.globus.org). З перерахованих вище платформ найбільший інтерес представляє gLite як найкрупніше і найбільш поширене в Європі ПГЗ, призначене для підтримки таких сервісів:

- аутентифікації і авторизації;
- брокера;
- моніторингу і діагностики;
- реплікації даних і програм;
- пошуку програмного забезпечення;
- каталогів даних і додатків;
- менеджменту ресурсів і завдань;
- розв'язання задачі;
- обліку і оплати отриманих послуг.



Рис. 2. Структура сервісів ПГЗ

Grid припускає колективний режим доступу до ресурсів і до пов'язаних з ними послуг (сервісів) у рамках глобально розподілених віртуальних організацій, які спільно використовують загальні ресурси. При цьому Grid визначає без допомоги користувача найбільш відповідне джерело даних і виконує їх аналіз, знайшовши найкраще місце для запуску відповідної програми на комп'ютерах, які простоюють. Якщо ж необхідно провести такий аналіз в інтерактивному режимі кількома користувачами із різних країн світу, то Grid зв'яже їхні комп'ютери так, що спільна робота не буде відрізнятися від роботи в локальній мережі. При цьому не треба хвилюватися про безліч паролів — Grid здатний зрозуміти, хто має право брати участь у спільній роботі. Для управління ресурсами і задачами Grid має інформаційну службу. Для управління ресурсами і розміщення задач використовується брокер ресурсів, і все це підтримується локальним диспетчером задач комп'ютерного кластера.

Серед ключових чинників, які сприяють упровадженню Grid, не тільки підвищення ефективності використання ресурсів і економія витрат, але й економія людських ресурсів, тому що обчислення Grid сприяють спільній роботі фахівців над проектами, дозволяючи їм застосовувати ту ж саму інфраструктуру, хоч стосовно України ще потрібно вирішити низку проблем щодо пропускнуєї спроможності мережі, управління даними і безпеки. Спочатку Grid-технології призначалися для вирішення складних наукових, виробничих та інженерних задач, які неможливо вирішити у прийнятні терміни на окремих обчислювальних установках. Проте тепер область застосування технологій Grid не обмежується тільки цими типами задач. У процесі свого розвитку Grid проникає у промисловість і бізнес, великі підприємства створюють Grid для вирішення власних виробничих задач. Та-

ким чином, Grid претендує на роль універсальної інфраструктури для обробки даних, в якій функціонують сервіси (Grid Services), що дозволяють розв'язувати не тільки конкретні прикладні задачі, але й пропонують додаткові послуги: пошук необхідних ресурсів, збір інформації про їх стан, зберігання і доставку даних.

Застосування Grid може дати нову якість розв'язанню таких класів задач:

- масова обробка потоків даних великого об'єму;
- багатопараметричний аналіз даних;
- моделювання на віддалених суперкомп'ютерах;
- реалістична візуалізація великих наборів даних;
- складні бізнес-додатки з великими об'ємами обчислень.

Grid-технології вже активно застосовуються у світі як державними організаціями управління, оборони, сфери комунальних послуг, так і приватними компаніями, наприклад, фінансовими і енергетичними. Область застосування Grid зараз обіймає ядерну фізику, захист навколишнього середовища, прогноз погоди і моделювання кліматичних змін, чисельне моделювання у машино- і авіабудуванні, біологічне моделювання, фармацевтику тощо.

GRID-СИСТЕМИ

Є два основні критерії, які виділяють Grid-системи (або інфраструктури) серед інших систем, що забезпечують роздільний доступ до ресурсів.

1. Grid-система координує розрізнені ресурси, які не мають загального центру управління, а Grid-система займається координацією їх використання, наприклад, балансування навантаження. Тому проста система управління ресурсами кластера не є системою Grid, оскільки здійснює централізоване управління всіма вузлами даного кластера, маючи до них повний доступ. Grid-системи мають лише обмежений доступ до ресурсів, залежний від політики того адміністративного домена (організації-власника), в якому цей ресурс знаходиться.

2. Grid-система будується на базі стандартних і відкритих протоколів, сервісів і інтерфейсів. Не маючи стандартних протоколів, неможливо легко і швидко підключати нові ресурси в Grid-систему, розробляти новий вигляд сервісів і т.ін.

Хоча технологія Grid не прив'язана до певних ресурсів, найчастіше реалізації Grid-систем забезпечують роботу з такими типами ресурсів:

- обчислювальними (окремі комп'ютери, кластери);
- ресурсами зберігання даних (диски і дискові масиви, стрічки, системи масового зберігання даних);
- мережевими;
- програмного забезпечення (спеціалізоване ПЗ).

Відзначимо різницю між технологією Grid і реалізаціями Grid-систем. Технологія Grid містить лише найбільш загальні і універсальні аспекти, однакові для будь-якої системи (архітектура, протоколи, інтерфейси, сервіси).

Використовуючи цю технологію і наповнюючи її конкретним змістом, можна реалізувати одну з Grid-систем, призначену для вирішення того або іншого класу прикладних завдань.

Слід розрізнити технологію Grid та технологією паралельних обчислень. У рамках конкретної Grid-системи, звичайно, можливо організувати паралельні обчислення з використанням існуючих технологій (PVM, MPI), оскільки Grid-систему можна розглядати як мета-комп'ютер, що має безліч обчислювальних вузлів. Проте технологія Grid не є технологією паралельних обчислень, в її завдання входить лише координація використання ресурсів.

Найбільшою в світі Grid-інфраструктурою є європейська EGEE (Enabling Grids for E-sciencE), ціль якої об'єднати національні, регіональні і тематичні Grid-розробки, що вже ведуться, в єдину цілісну Grid-інфраструктуру для підтримки наукових досліджень. EGEE надасть дослідникам як в академічних колах, так і в різних областях економіки, цілодобовий доступ до найбільш високопродуктивних обчислювальних ресурсів, незалежно від їх географічного положення. Користуються інфраструктурою географічно розподілені співтовариства дослідників, які потребують загальних обчислювальних можливостей Grid-систем, готові об'єднати свою власну обчислювальну інфраструктуру у відповідності до принципів загального доступу. Проект, впроваджений Європейським центром ядерних досліджень (ЦЕРН), підтримують, в основному, фінансуючі установи ЄС, але призначений він для роботи у всьому світі. Значні кошти надходять від США, Росії та інших учасників проекту, що не входять в ЄС.

Сьогодні в EGEE функціонують 240 вузлів у 45 країнах, в яких задіяні 41000 процесорів і 5 Пб (PetaBytes) пам'яті. Мережа обслуговує більше 10000 споживачів і 150 віртуальних організацій з продуктивністю більше 100000 обчислювальних завдань за день. У мережі експлуатуються додатки з археології, астрономії, астрофізики, захисту довкілля, комп'ютерної хімії, науки про Землю, фінансів, фізики плазми, геофізики, фізики високих енергій, науки про життя, мультимедіа, матеріалознавства тощо.

EGEE-II складається з мережних напрямів робіт (Networking Activities, NA), служб (Service Activities, SA) і напрямів сумісних досліджень (Joint Research Activities, JRA). Відповідно до сучасного, досконалішого стану Grid-технологій в EGEE-II частка фінансування SA і NA підвищиться, а JRA — знизиться. Це дозволить включити до інфраструктури проекту нові країни, додатки і сайти, а також у більшому масштабі поширювати інформацію, вести навчання і підтримувати додатки. Скоротиться розробка програмного забезпечення, оскільки його можна буде інтегрувати з інших проектів і джерел.

Мережні напрями. NA1 — управління проектом; NA2 — розповсюдження інформації і зв'язок; NA3 — навчання і включення в число користувачів; NA4 — визначення і підтримка Grid-додатків; NA5 — політика і міжнародна співпраця.

Служби. SA1 — підтримка, експлуатація і управління європейськими Grid; SA2 — забезпечення мережними ресурсами; SA3 — інтеграція, тестування і сертифікація; SA4 — новий напрям (в його рамках об'єднуюватимуться елементи програмного забезпечення з різних джерел для

отримання інтегрованих випусків програмного забезпечення, готових до розміщення в інфраструктурі проекту.

Напрями сумісних досліджень. JRA1 — перепроєктування проміжного програмного забезпечення. Продовжиться розробка і підтримка gLite. JRA2 — забезпечення якості робіт протягом всього проекту, зокрема, вестиметься загальна координація заходів безпеки.

З 2007 р. розпочалася розбудова Європейської Grid-інфраструктури (EGI), функціонування якої планується на постійній основі у вигляді скоординованої мережі національних Grid-систем, або NGI. ЦЕРН залишає за собою загальну координацію і відповідальність за модернізацію ПЗ і систему загальної безпеки.

Grid-система стає **національною** (NGI — National Grid Initiative), якщо вона задовольняє таким критеріям: має державну підтримку, наприклад, шляхом включення проекту створення Grid-інфраструктури у державну програму з гарантованим фінансуванням; представляє інтереси всіх шарів суспільства (вчених, вузівських працівників, промисловців, комерсантів і т.ін.); має розгалужену структуру з координуючих, регіональних і ресурсних центрів, які забезпечують функціонування базових Grid-сервісів, моніторинг і реагування на надзвичайні ситуації, ведення обліку ресурсів і виконаних робіт (accounting), управління і підтримку віртуальних організацій, сертифікацію Grid ПЗ; базується на дотриманні міжнародних стандартів і правил (наприклад, використовуванні протоколу SOAP для доступу до ресурсу, мови WSDL для опису Web-сервісів, каталогів UDDI як реєстрів описів і т.д.); підтримує безпеку інфраструктури, має право генерувати сертифікати користувачів CA з відома EUGridPMA (системи Європейської Grid-Аутентифікації); підключена до GEANT, Європейської науково-освітньої комп'ютерної мережі; має керівні органи Grid-інфраструктури у вигляді Ради з розвитку національної Grid-інфраструктури в напрямках NA, SA і JRA, наведених вище.

GRID-ПРОЕКТИ І ВІРТУАЛЬНІ ОРГАНІЗАЦІЇ

Grid дозволяє підвищити творчий потенціал людства, об'єднавши обчислювальні потужності для використання найважливіших додатків і дозволивши географічно розділеним користувачам мережі взаємодіяти і підтримувати спільну роботу. Існує п'ять основних груп Grid-додатків: розподілені суперобчислення, обчислення з високою пропускнуною спроможністю та за запитом, з інтенсивним обміном даними і сумісні додатки.

Треба розрізняти поняття Grid-системи і Grid-проекти, що здійснюються, наприклад, ЦЕРН у галузі фізики високих енергій на Європейській EGEE інфраструктурі, в яких сьогодні приймають участь два фізичні інститути — Харківський фізико-технічний інститут (ХФТІ) та Інститут теоретичної фізики ім. М.М. Боголюбова (ІТФ). При чому кластери ІТФ і Київського Національного Університету ім. Тараса Шевченка підключені до мережі AliEn-grid для обслуговування експерименту ALICE на прискорювачі ЛНС. Вузол ХФТІ підключений до мережі ЦЕРН через Російський RDIG, а ряд співробітників згаданих організацій отримали доступ до ресурсів EGEE також через Російський RDIG. Grid-проекти виконуються також в Інституті космічних

досліджень НАН України — НКАУ (обробка космічних знімків) і у Головній астрономічній обсерваторії НАН України (для задач астрофізики).

Для використання інфраструктури Grid при виконанні конкретного проекту користувачеві необхідно пройти процес реєстрації, після чого всі ресурси Grid стають доступними без будь-яких додаткових угод з окремими ресурсними центрами.

Процес реєстрації нового користувача здійснюється у два кроки:

1. Отримання персонального, призначеного для користувача, сертифіката.

2. Реєстрація у віртуальній організації.

Персональний, призначений для користувача, сертифікат (Personal Certificate), — це свого роду електронний документ, який підтверджує особу користувача при доступі до Grid-ресурсів. Сертифікати видаються центрами сертифікації (Certification Authority).

Фактично національна Grid-інфраструктура необхідна лише для того, щоб користувач міг запитати і отримати ті або інші ресурси від їх власників. Кожна віртуальна організація має свій власний центр реєстрації. Існує також можливість почати роботу в Grid, скориставшись учбовими порталами. У даному проєкті передбачається, що майбутні користувачі можуть спробувати попрацювати з невеликою інфраструктурою Grid, не будучи офіційними членами однієї з віртуальних організацій. Для більшості користувачів це найпростіший, а іноді і єдиний, спосіб навчитися працювати в Grid.

GRID-СИСТЕМИ УКРАЇНИ

Сьогодні важко знайти більш-менш розвинену країну, в якій не було б розгорнуто національні Grid-проєкти. Настала черга й України, хоч і з великим запізненням.

Національна Grid-інфраструктура створюється для обслуговування різних організацій з різноманітними характеристиками і вимогами до ресурсів. Для об'єднання різноманітних ресурсів єдина обчислювальна інфраструктура повинна мати єдину архітектуру. Можливість упорядкування початкового матеріалу забезпечується попереднім визначенням базових функцій Grid-інфраструктури.

Grid-сегмент НАН України (www.acadgrid.org.ua)

За ініціативою ІТФ у квітні 2006 р. в НАН України стартувала корпоративна програма «Впровадження Grid-технологій і створення кластерів в Національній академії наук України» (або Українська Академічна Grid Ініціатива — UAGI), що не враховує повною мірою загальнонаціональні потреби і наукові інтереси вчених і організацій, які не працюють у системі НАН України. Науковий керівник проєкту — А.Г. Загородній, директор ІТФ, академік НАН України. У квітні 2007 р. НАН України підписала Меморандум про взаєморозуміння з WLCG (Worldwide LHC Computing Grid), у ЦЕРН і обробці та аналізі даних — з LHC (Large Hadron Collider, великого адронного колайдери).

Завершується обладнання нових кластерів у Головній астрономічній обсерваторії, Інституті молекулярної біології і генетики, Інституті клітинної біології і генної інженерії, модернізація кластерів у ХФТІ, в Інституті фізики конденсованих станів у Львові. Загалом комп'ютерні ресурси цієї групи складають 330 процесорів і біля 53 Тбайтів дискового масиву. Всі ці кластери будуть об'єднані в Grid-сегмент НАН України (рис. 3). Принципи побудови цього сегменту, його конфігурації, методи керування ним вже практично відпрацьовано на прообразі цього сегменту — об'єднанні кластерів ІТФ, Київського Національного Університету ім. Тараса Шевченка та Інституту кібернетики НАН України. В ньому зосереджені найпотужніші суперкомп'ютери НАН України, серед яких є СКІТ-3 з такими параметрами: 280 процесорів (70 вузлів), Linpack продуктивністю 2,235 Тфлопс, система збереження даних — 10 Тбайт.

Передбачено також побудову кластерів в Інституті фізики (Київ), Інституті металофізики (Київ), Інституті математики (Київ), Радіоастрономічному інституті (Харків), Інституті сцинтиляційних матеріалів (Харків), Інституті геотехнічної механіки, Інституті транспортних систем і технологій, Інституті чорної металургії, Інституті проблем природокористування та екології (Дніпропетровськ).

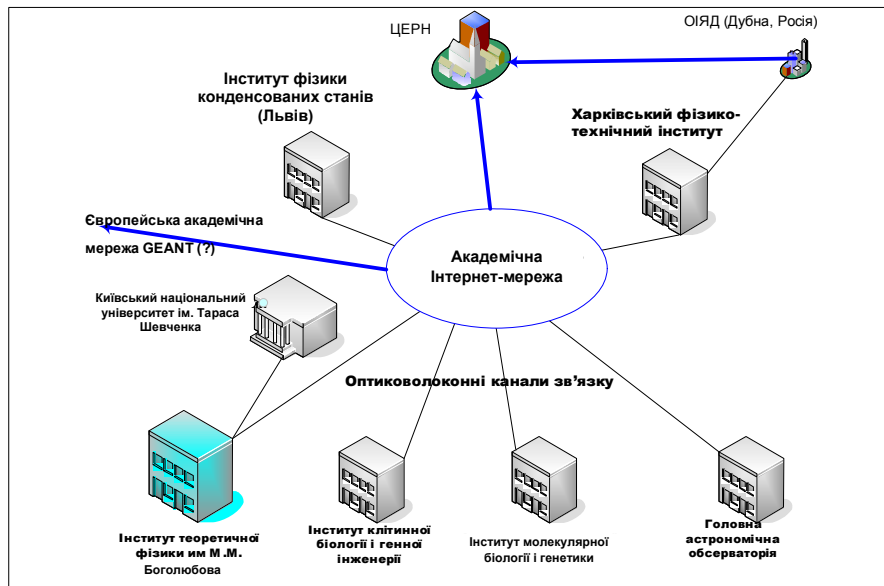


Рис. 3. Grid-сегмент НАН України

Слід зауважити, що реалізоване сьогодні в Grid-сегменті НАН України ПГЗ NorduGrid не забезпечує централізацію розподілення обчислювальних задач (Centralized workload management), загальної організації каталогу для всієї Grid-інфраструктури (Global names in file catalogue for whole Grid), глобальної організації послуг брокера і віртуалізації сховищ даних (Different kinds of backend for storage elements), і тому планується перехід до ПГЗ gLite.

Проект UGrid МОН України (www.grid.ntu-kpi.kiev.ua)

Згідно із Наказом Міністерства науки і освіти (МОН України) за № 758 від 22 серпня 2007 р. в Україні розпочата розробка Національної Grid-інфраструктури виробничої якості з високоякісними послугами (проект UGrid) на базі Національної науково-освітньої мережі УРАН для забезпечення загальнонаціональної віртуалізації розподілених обчислювальних ресурсів різних типів (процесорів, сховищ даних, мереж, унікального обладнання), здатних підтримувати життєдіяльність державних структур, наукових і освітніх організацій, промислових корпорацій. Проект виконує десять різних українських організацій (дві академічні, шість освітянських і дві промислові), які очолюють Інститут прикладного системного аналізу НАН України та Міносвіти і науки України (ІПСА). Науковий керівник проекту — М.З. Згуровський, ректор НТУУ «КПІ», академік НАН України.

Серед виконавців проекту Інститут проблем моделювання в енергетиці ім. Г.Є. Пухова НАН України, Харківський національний університет радіоелектроніки (ХНУРЕ), Львівський національний технічний університет «Львівська політехніка» (НУЛП), Запорізький національний технічний університет (ЗНТУ), Донецький національний технічний університет (ДонНТУ), Дніпропетровський національний гірничий університет (ДНГУ), державне підприємство «Львівський науково-дослідний радіотехнічний інститут» (ЛНДРІ) і підприємство ЮСТАР.

Проект передбачає створення високоякісної національної Grid-інфраструктури з відповідними сервісами для надання можливості вітчизняним науковцям плідно співпрацювати в Європейському науковому просторі (European Research Area, ERA) і сприяти створенню економіки інформаційного суспільства, заснованої на знаннях, шляхом впровадження наукових концепцій Grid і найбільш вагомих наукових додатків, які використовуються в Grid-середовищі. Ціль проекту — розв'язання таких задач:

- Забезпечення обслуговування Української філії Світового центру даних (УФ СЦД), надавши його клієнтам віддалений доступ до світових сховищ наукових даних, можливість ефективного сумісного використання комп'ютерів, унікальних експериментальних установок і приладів.
- Добудування і об'єднання у співдружності з НАН України існуючої наукової і освітньої обчислювальної і комунікаційної інфраструктури в національну Grid-інфраструктуру і підключення її до Європейської, що складається із сукупності національних Grid-інфраструктур, використання ПГЗ gLite.
- Розповсюдження в суспільстві знань про Grid-технології і навички з їх використання на рівні, притаманному європейським країнам, які мають більш тривалий досвід розробки та використання Grid.
- Розгортання, згідно із домовленістю з EUGridPMA, Сертифікаційного центру відкритих ключів в Україні, враховуючи наявність проекту Національної Grid-інфраструктури і визнання мережі УРАН в Європі та початок підготовки користувачів Grid.
- Організація обслуговування національної Grid-інфраструктури шляхом створення регіональних ресурсно-операційних центрів і постійно діючих груп з питань підтримки, експлуатації і управління Grid; розповсю-

дження інформації про проект і розширення кола користувачів; їх навчання і підготовка; визначення областей застосування і підтримки додатків; забезпечення мережних ресурсів тощо.

- Сприяння створенню спільних проектів в межах міжнародних віртуальних організацій українських науковців з іноземними колегами з питань сталого розвитку, геофізики і комп'ютеризації проектування, забезпечивши їх інформаційну підтримку.

Передбачається протягом 2007-2008 рр. створити Grid-інфраструктуру з шістьма ресурсно-операційними центрами (у Києві, Харкові, Донецьку, Дніпропетровську, Запоріжжі і Львові) та розпочати віддалене обслуговування майбутніх користувачів (науковців з університетів та наукових установ України). Звичайно, що ця початкова Grid-інфраструктура буде добуватися, а число її центрів і розробників з організацій МОН України і академічних інститутів НАН України буде зростати. Цьому сприятиме фактальність Grid-інфраструктури (система подібна сама до себе на кожній, різній за масштабом ділянці).

За минулий рік на шляху до підключення Українського Grid в якості повноцінної складової до Європейської Grid-інфраструктури зроблено кілька дуже важливих кроків. У червні 2007 р. підписана угода з європейською організацією DANTE про підключення Національної науково-освітньої мережі УРАН до Європейської мережі GEANT. У січні 2008 р. прийнято рішення іншої європейської організації EUGridPMA про відкриття в Україні Сертифікаційного центру відкритих ключів для надання українським користувачам доступу до європейських ресурсів і сховищ даних.

Передбачається, що вже у 2008 р. будуть створені і розпочнуть реальну співпрацю спільні наукові команди у вигляді віртуальної організації з українських і зарубіжних вчених. Згідно з проектом у кожному із шести ресурсно-операційних центрів національної Grid-інфраструктури створюються центри навчання і підтримки користувачів (рис. 4). Для цього передбачена розробка Grid-порталу, який буде задіяний в ресурсно-операційних центрах для навчання користувачів Grid-технологіям, і його будуть відвідувати чисельні гості з різних регіонів країни і зарубіжжя.

Крім того, в ПСА та ряді інших вузів з вересня 2007 р. розпочалася програма магістерської підготовки з напрямку «Grid-технології в науці і освіті». Всі організації учасниці проекту (крім обов'язків з організації ресурсно-операційних центрів, забезпечення їх функціонування, підготовки і підтримки користувачів) запланували проведення наукових досліджень у галузі сумісності (interoperability) ПГЗ, забезпечення наскрізної інформаційної безпеки при об'єднанні національних Grid-інфраструктур чи їх сегментів, розробки додаткових сервісів для розширення кола можливих користувачів (окрім науки) на представників різних прошарків суспільства (інженерії, бізнесу і соціальної сфери тощо). Наприклад, ЗНТУ планує створити Grid-додаток із задач молекулярної динаміки, ЛНДРІ — із задач комп'ютерної інженерії, НУЛП — із задач обробки сигнальної інформації та медицини, ДонНТУ — із методик і засобів паралельного програмування, ХНУРЕ — із задач бізнесу і економіки і т.д. Оскільки більшість вищих навчальних закладів, академічних інститутів і промислових підприємств поки не мають суперкомп'ютерів, то в проекті передбачено розробку заходів вір-

туалізації ресурсів звичайних персональних комп'ютерів за допомогою Grid-технологій. Довкола кожного провідного університету регіону групуються місцеві вищі навчальні заклади (ВНЗ), так що реальна кількість ВНЗ, які задіяні в Ugrid-проекті, перевищує 30.



Рис. 4. Розташування ресурсно-операційних вузлів UGrid

Обчислювальні ресурси UGrid складаються з суперкомп'ютера НТУУ «КПІ», найпотужнішого із комп'ютерів України (згідно із останнім рейтингу суперкомп'ютерів країн СНД — www.supercomputers.ru) з такими параметрами: 84 вузли, Linkcrack продуктивністю 3,13 Тфлопс, система збереження даних обсягом 12 Тбайт і архівна пам'ять 20 Тбайт, а також кластерів ХНУРЕ, НУЛП, ЗНТУ, ДонНТУ і підприємства ЮСТАР. Суперкомп'ютер НТУУ «КПІ» забезпечує функціонування українського філіалу Світового центру даних (СЦД) і віртуалізацію його різномірних систем зберігання даних у вигляді **єдиного сховища даних** (так званого віртуального пулу), управління яким здійснюється централізовано за допомогою ПЗ (IPStor фірми FalcomStor). Забезпечення користувачів сучасними даними (з фізики твердої Землі, сонячно-земної фізики, сталого розвитку суспільства та ін.) і обчислювальними ресурсами для їх обробки здійснювалося дистанційно через спеціалізований портал СЦД і Національну комп'ютерну мережу для науки і освіти УРАН.

Сумісний проект UNGI для EGI

Інтереси держави і життя вимагають спільних зусиль інститутів НАН України і організацій МОН України для об'єднання існуючих сегментів наукової і освітньої обчислювальної і комунікаційної інфраструктури в єдину Українську національну Grid-ініціативу (UNGI) та інтегрування UNGI в Європейську Grid-інфраструктуру EGEE – EGI. НАН України має авторитетне Від-

ділення інформатики, істотні обчислювальні ресурси, практичний досвід об'єднання високопродуктивних комп'ютерів у мережу, досвід проведення досліджень окремими вченими в реальних умовах європейських Grid-проектів, багаторічний досвід алгоритмізації різноманітних наукових задач і розробки для них відповідних додатків, починаючи від задач фізики та біології і до літературознавства.

Вузи МОН України і промисловість мають багаторічний досвід побудови і дослідження розподілених обчислювальних систем, зокрема, системи колективного мережного проектування виробів високих технологій на технологіях Grid, а також досвід партнерства з європейським проектом BalticGrid; порівняльне випробування ПГЗ Globus, NorduGrid і gLite. Існують офіційна угода з DANTE про підключення комп'ютерної мережі УРАН до європейської мережі GEANT; домовленість з EUGridPMA про заснування служби CA в Україні; величезний набір рішень науково-технічних і інженерних задач, починаючи від технічного і технологічного передбачення розвитку науки і техніки, вартості і наслідків рішень в цих галузях, пов'язаних із вибором пріоритетів і призначенням об'ємів фінансування на державному рівні або рівні підприємств, і закінчуючи моделюванням логічних схем з одноелектронними нанотранзисторами; швидко наростаючі обчислювальні ресурси і найсучасніший досвід кластеробудування (фірма ЮСТАР); практично невичерпний резерв талановитої творчої молоді.

У цілях забезпечення повномасштабної участі України в проекті глобальної європейської комп'ютерної інфраструктури типу Grid – EGI-EGEE ІТФ (провідна організація проекту UAGI) та ПІСА (провідна організація проекту UGrid) сформували в серпні 2007 р. спільний проект UNGI (Українська Національна Grid Ініціатива — Ukrainian National Grid Initiative), який вже включено в міжнародну сьому Рамочну програму. Проект UNGI, згідно із прийнятою в програмі EGEE структурою, входить в EGI-EGEE як складова регіональної федерації «Росія» («Russia») нарівні з ЦЕРН і іншими вісьмома національними і регіональними федераціями (Франція, Італія, Великобританія/Ірландія, Німеччина/Швейцарія, Скандинавські країни, консорціуми центральноєвропейських країн, Південно-Східної та Південно-Західної Європи). Тепер на карті Європи залишилося тільки дві білі плями (Білорусь і Молдова), не охоплені структурою EGI-EGEE.

Спільний проект UAGI увібрав найкраще від своїх попередників: UAGI — тісний зв'язок з ЦЕРН і участь в його Grid-проектах з фізики високих енергій, UGrid — потужну інформаційну систему у вигляді Української філії Світового центру даних, що входить в мережу з 52 центрів. Величезні обсяги доступної інформації стимулюють розробку методів її обробки, зберігання і оцінки (Data Mining), особливо в сучасних умовах, коли існуючі обсяги інформації почали перевершувати технічні можливості зберігання. Передбачено єдиний сертифікаційний центр відкритих ключів, єдине ПГЗ gLite, шлюзи між комп'ютерними мережами УРАН і УАРНЕТ, однакові Grid-портали у споживачів, спільні ресурсно-операційні вузли Grid-системи. Заявлені в Європі параметри UNGI зведені у таблицю.

Тепер черга за координацією чи об'єднанням робіт з проектів Grid, що проводяться організаціями НАН України і МОН України, в єдину дійсно національну Grid-інфраструктуру для наукових досліджень і освіти із залученням зацікавлених організацій і відомств.

Базові параметри UNGI у документах EGI

Параметр	Кількість
Як багато ресурсно-операційних центрів у Вашій Grid-системі?	8 зараз і ще 9 в найближчі півроку
Як багато процесорів у Вашій Grid-системі?	Більше 1000
Яка ємність існуючих сховищ даних?	Біля 100 TB
Чи фінансує Уряд Вашу розробку?	Так, Наказ МОНУ №758 від 22.08.07
Чи існує концепція впровадження Grid-технологій в країні?	Так
Коли розпочалися роботи з Grid?	Листопад 2004 (UAGI) Січень 2006 (UGrsd)
Чи є національні дослідницько-освітні мережі?	УРАН і УАРНЕТ
Чи є центри суперкомп'ютерних обчислень?	Так, в ІК НАН України і НТУУ «КПІ»
Чи займається Ваша NGI розробкою додатків?	Так
Скільки додатків планується створити?	20
До яких галузей відносяться ці додатки?	Наука, інженерія, бізнес, наука про Землю
Скільки людей працюють над Grid-розробками?	Біля 100

Примечание [P1]:

НА КОГО РОЗРАХОВАНА GRID-ІНФРАСТРУКТУРА

Потенціал технологій Grid вже зараз оцінюється дуже високо: він має стратегічний характер, і в близькій перспективі Grid повинен стати обчислювальним інструментарієм для розвитку високих технологій в різних сферах людської діяльності, подібно персональному комп'ютерові та Інтернет.

Україна розпочала «на рівних» співпрацювати з країнами Європейського Союзу по створенню і впровадженню технології 21-го століття — Grid, яка забезпечує сумісний доступ до комп'ютерних ресурсів (що змінюються не лише на рівні файлів і даних, але й комп'ютерів, сенсорів і мереж), реалізує різноманітні режими їх використання, забезпечує суворий контроль, управління і організацію системи безпеки, підтримує гетерогенність мережі, баланс навантаження на обчислювальних вузлах і т. ін.

Наведемо основні співтовариства, що сьогодні потребують застосування Grid-технологій.

- Урядові організації (службовці, експерти і науковці), які традиційно займаються питаннями національної безпеки, довгостроковими дослідженнями і плануванням.
- Організації охорони здоров'я, Grid-сегмент яких відрізняють відносно невеликі розміри, централізоване управління і складність корпоративної інфраструктури.

- Співтовариства вчених, яким необхідна віртуальна Grid-мережа, з універсальним доступом, відносно вузькою спрямованістю, динамічно змінним складом користувачів, децентралізованою управління, а також частим сумісним зверненням до існуючих ресурсів. Схожа модель може бути використана для міжгалузевих, міжвідомчих і міждисциплінарних дослідницьких груп і т.п. Прикладом такого співтовариства вчених може бути участь фахівців ІТФ у виконанні завдань проектів ALICE, CMS, що проводяться ЦЕРНОм у галузі фізики високих енергій.

- Співтовариство з великим числом учасників, яке охоплює весь існуючий ринок обчислень і якому властива, відсутність постійних схем і варіантів взаємодій.

У рамках спільного проекту UNGI головна мета України — включення UNGI у спільну Grid-інфраструктуру Європи і забезпечення постійного функціонування її як повноцінної операційної та функціональної складової цієї структури. Україна отримує можливість уже сьогодні співпрацювати з країнами Європейського Союзу над створенням і використанням Grid-технологій для забезпечення обміну науковими даними та організації їх колективного використання, а в найближчих кілька років подолати відставання від європейських країн і увійти в Європейський дослідницький простір (European Research Area, ERA) повноправним і кваліфікованим партнером.

ЛІТЕРАТУРА

1. *Zgurovsky M.Z.* Development of educational and research segment of information society in Ukraine // Proc. of world symposium on information society (WSIS). — Tunis, 2004. — P. 254–260.
2. *Zgurovsky M.Z., Petrenko A.I.* Grid-infrastructure for supporting education and research in Ukraine // CD-ROM with material of Baltic grid first all-hands meeting. — Vilnius, 26–28 April 2006.
3. *Zgurovsky M.Z., Petrenko A.I.* National Ukrainian GRID Infrastructure (UGRID) for sciences and educations // Proc. Intern. conf. on computer science and information technologies. — Lviv, Ukraine, October, 2006. — P. 6–12.
4. *Petrenko A.I.* Development of GRID-infrastructure for educational and research segment of information society in Ukraine with focus on Ecological monitoring and Telemedicine // Proc. of 20-th CODATA conf. — Beijing, China. 2006 — P. 145–147.
5. *Петренко А.І.* GRID-технології в науці і освіті // Матеріали 9-ї Міжнар. конф. «Системний аналіз та інформаційні технології». — Київ: ИПСА. — 15–19 травня 2007. — С. 22–23.
6. *Петренко А.* Национальная GRID-инфраструктура для науки и образования. — CD-ROM с материалами конф. «HPC Day: Современные решения для высокопродуктивных вычислений». — Киев: ЮСТАР, 18 октября 2007.
7. *Мартынов Е.С.* Украинская национальная GRID-инициатива: Состояние и перспективы. — CD-ROM с материалами конф. «HPC Day: Современные решения для высокопродуктивных вычислений». — Киев: ЮСТАР, 18 октября 2007.
8. *Свистунов С.Я.* GRID-инфраструктура НАНУ: Опыт создания и первые результаты эксплуатации. — CD-ROM с материалами конф. «HPC Day: Современные решения для высокопродуктивных вычислений». — Киев: ЮСТАР, 18 октября 2007.

Надійшла 01.11.2007.