

Національна Grid - інфраструктура для забезпечення наукових досліджень і освіти

Петренко А.І., проф., доктор технічних наук,
зав. кафедри системного проектування ІІСА НТУУ «КПІ»

Анотація: Grid-технології і всевітня 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 вирішенням складних наукових і інженерних задач. Із розвитком Grid проникає у промисловість і бізнес, претендуючи на роль універсальної інфраструктури для обробки даних, у якій функціонує безліч служб (Grid Services), що не лише дозволяють вирішувати конкретні прикладні задачі, але й пропонують послуги з пошуку необхідних ресурсів, збору інформації про стан ресурсів, зберігання і доставки даних. В статті проведено аналіз концепції Grid, архітектури та ресурсів сучасних Grid систем на прикладі побудови національної Grid інфраструктури для научних досліджень і освіти.

1, Вступ до Grid- технологій

Grid Grid- технології і всевітня 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 вирішенням складних наукових і інженерних задач. Із розвитком Grid проникає в промисловість і бізнес, претендуючи на роль універсальної інфраструктури для обробки даних, у якій функціонує безліч служб (Grid Services), що не лише дозволяють вирішувати 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 в дуже широкому сенсі, тобто ресурсом може бути апаратура (жорсткі диски, процесори), а також системне і прикладне ПО (бібліотеки, додатки).

Якщо перекладати дослівно, Grid означає «грати». Ближче всього по **значенню**, мабуть, power grid —мережа електроживлення, розподілений ресурс **загального** користування, коли кожний може легко підключитися через розетку і вжити електрики, скільки йому вимагається. Аналогічно користувачі комп'ютерних потужностей дістають можливість прямого підключення до **віддаленої** обчислювальної мережі, не цікавлячись, звідки саме приходять що вимагаються для роботи обчислювальні ресурси, які для цього використовуються лінії передачі, паролі або протоколи і т.п. При цьому аналогом інфраструктури електричних мереж (ліній електропередач, підстанцій, трансформаторів і ін.) **виступає** Grid програмне забезпечення (ПГО), або Middleware (рис.1).

Розвиток ПГО починався від базових засобів, що 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 ПГО, так що зараз є платформи, що 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 ПГО, що призначено для підтримки:

- Сервісу аутентифікації і авторизації
- Сервісу брокера.
- Сервісу моніторингу і діагностики.
- Сервісу реплікації даних і програм.
- Сервіс пошуку програмного забезпечення.

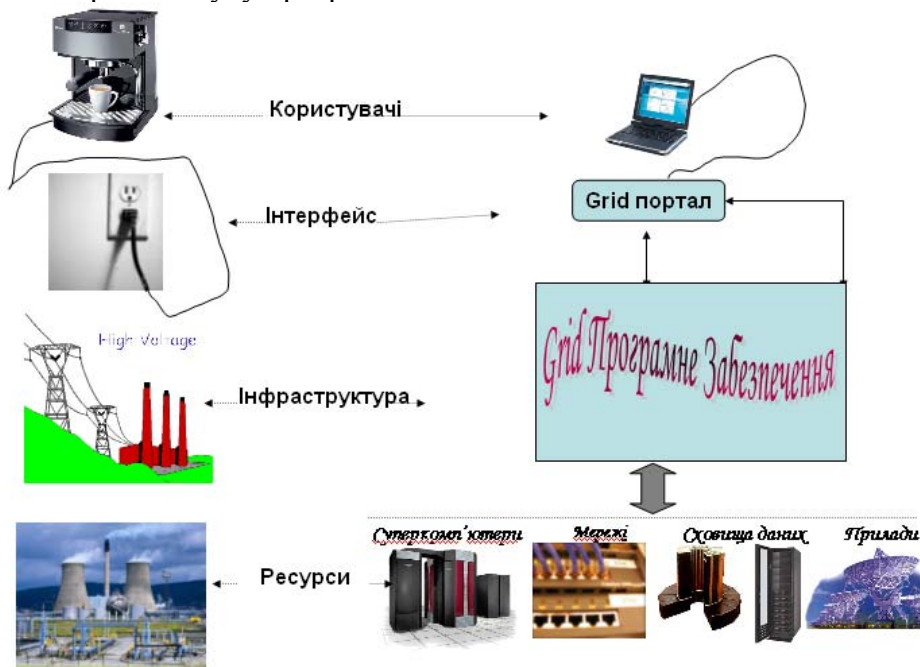


Рис.1. Аналогія електро – і Grid мереж

- Сервісу каталогів даних і додатків
- Сервісу менеджменту ресурсів і завдань
- Сервісу збірки і вирішення задачі
- Сервісу обліку і оплати отриманих послуг

Grid припускає колективний режим доступу до ресурсів і до пов'язаних з ними послуг (сервісів) в рамках глобально розподілених віртуальних організацій (ВО), які спільно використовують загальні ресурси. При цьому Grid визначає без допомоги користувача найбільш відповідне джерело даних і виконує їх аналіз, знайшовши якнайкраще місце для запуску відповідної програми на комп'ютерах, які простоюють. Якщо ж необхідно провести такий аналіз в інтерактивному режимі кількома користувачами з 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 буде хвилюватися про безліч паролів - Grid здатний зрозуміти, хто має право брати участь в 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.

В числі ключових чинників, сприяючих 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 використовувати одну і ту ж інфраструктуру, хоч стосовно України ще потрібно вирішити цілий ряд проблем щодо пропускнуєї спроможності мережі, управління даними і безпеки. Спочатку Grid - технології призначалися для вирішення складних

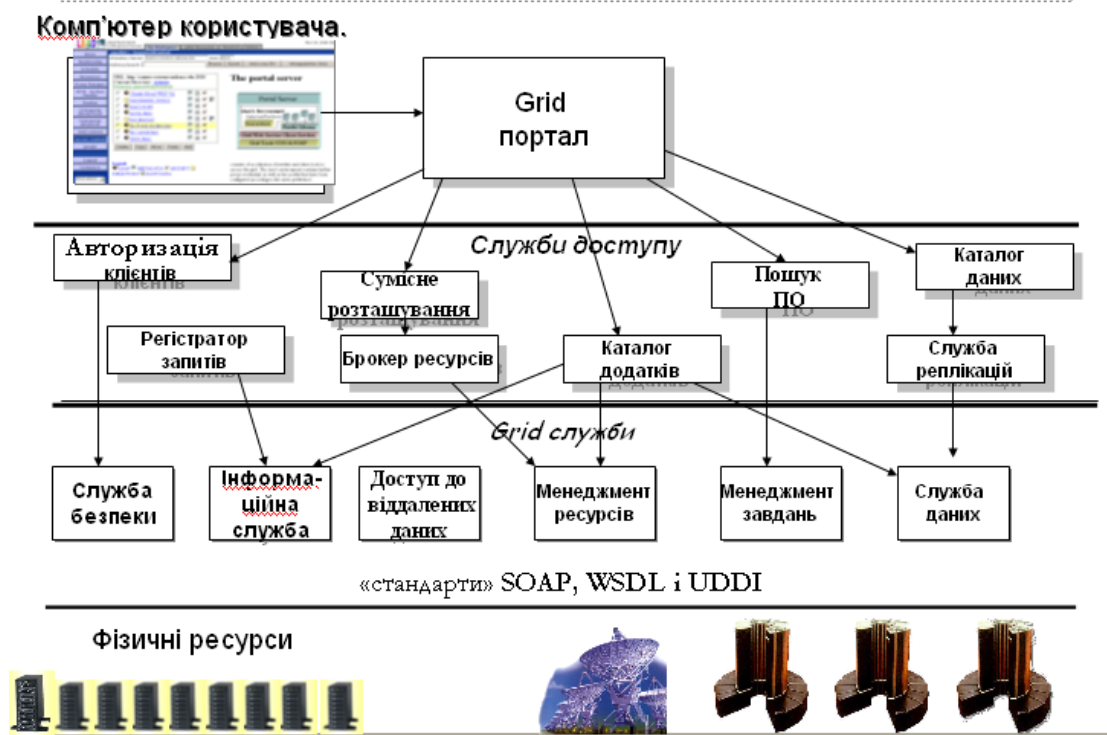


Рис.2. Структура сервісів ПГО

наукових, виробничих і інженерних задач, які неможливо вирішити в розумні терміни на окремих обчислювальних установках. Проте тепер область застосування технологій Grid не обмежується тільки цими типами задач. У міру свого розвитку Grid проникає в промисловість і бізнес, крупні підприємства створюють Grid для вирішення власних виробничих задач. Таким чином, Grid претендує на роль універсальної інфраструктури для обробки даних, в якій функціонує множина сервісів (Grid Services), які дозволяють вирішувати не тільки конкретні прикладні задачі, але і пропонують додаткові послуги: пошук необхідних ресурсів, збір інформації про стан ресурсів, зберігання і доставка даних.

Застосування Grid може дати нову якість рішення наступних 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 охоплює ядерну фізику, захист навколишнього середовища, прогноз погоди і моделювання кліматичних змін, чисельне моделювання в машино- і авіабудуванні, біологічне моделювання, фармацевтику тощо.

2. Grid- системи

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

1. Grid-система координує розрізнені ресурси. Ресурси не мають загального центру управління, а Grid-система займається координацією їх використання, наприклад, балансуванням 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.і.

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

- **обчислювальні ресурси** – окремі комп'ютери, кластери;
- **ресурси зберігання даних** – диски і дискові масиви, стрічки, системи масового зберігання даних;
- **мережеві ресурси**;
- **програмне забезпечення** – яке-небудь спеціалізоване ПО.

Відзначимо різницю між технологією Grid і реалізаціями Grid-систем. Технологія Grid включає 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 ту або іншу Grid-систему, призначену для вирішення того або іншого класу прикладних завдань.

Не слід змішувати технологію Grid з технологією паралельних обчислень. В рамках конкретної Grid-системи, звичайно, можливо організувати паралельні обчислення з використанням існуючих технологій (PVM, MPI), оскільки Grid-систему можна розглядати як якийсь 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 цільну Grid-інфраструктуру для підтримки наукових досліджень. EGEE надасть дослідникам як в академічних кругах, так і в різних областях економіки цілодобовий доступ до самих високопродуктивних обчислювальних ресурсів незалежно від їх географічного положення. Користуються інфраструктурою географічно розподілені співтовариства дослідників, які потребують загальних для них обчислювальних можливостей Grid-систем, готові об'єднати свою власну обчислювальну інфраструктуру і згодні з принципами загального доступу. Проект, впроваджений

ЦЕНР (Європейським центром ядерних досліджень).

підтримують, в основному, фінансуючі установи ЄС, але призначений він для роботи у всьому світі. Значні засоби поступають від США, Росії і інших учасників проекту, що не входять в ЄС.

Сьогодні в EGEE функціонують 240 вузлів в 45 країнах (див. рис. 9.1), в яких задіяні 41,000 процесорів і 5 Пб (PetaBytes) пам'яті. Мережа обслуговує більше 10,000 споживачів і 150 віртуальних організацій з продуктивністю більше 100,000 обчислювальних завдань за день. В мережі експлуатуються додатки з археології, астрономії, астрофізики, захисту навкіл ля, 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 сучасному, досконалішому стану Grid-технологій в EGEE-II частка фінансування SA і NA підвищиться, а JRA – знизиться. Це дозволить включити в інфраструктуру проекту нові країни, додатки і сайти, а також в більшому масштабі

поширювати інформацію, вести навчання і підтримувати додатки. Скорочена розробка програмного забезпечення, оскільки його можна інтегрувати з інших проектів і джерел.

Мережні напрями. Проект включає наступні мережні напрями робіт: NA1 – управління проектом; NA2 – розповсюдження інформації і зв'язок; NA3 – навчання і включення в число користувачів; NA4 – визначення і підтримка грид-додатків; NA5 – політика і міжнародна співпраця.

Служби. Проект включає наступні служби: SA1 – підтримка, експлуатація і управління європейськими Grid; SA2 – забезпечення мережними ресурсами; SA3 – інтеграція, тестування і сертифікація. SA3 – новий напрям; в його рамках об'єднуюватимуться елементи програмного забезпечення з різних джерел з метою отримати інтегровані випуски програмного забезпечення, готові до розміщення в інфраструктурі проекту.

Напрями сумісних досліджень. Проект включає наступні напрями сумісних досліджень: JRA1 – перепроектувало проміжного програмного забезпечення і JRA 2 – забезпечення якості. В рамках JRA1 продовжиться розробка і підтримка gLite. В рамках JRA2 забезпечуватиметься якість робіт протягом всього проекту; зокрема, вестиметься загальна координація заходів безпеки.

З 2007 року 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 ПГО і систему загальної безпеки.

Grid-система стає **національною** (NGI- National Grid Initiative), якщо вона задовольняє наступним критеріям: має державну підтримку, наприклад, шляхом включення проекту створення Grid інфраструктури в Державну програму з гарантованим фінансуванням; представляє інтереси всіх шарів суспільства (вчених, вузівських працівників, промисловців, комерсантів і ін.); має розгалужену структуру з координуючих, регіональних і ресурсних центрів, що забезпечують функціонування базових Grid сервісів, моніторинг і реагування на надзвичайні ситуації, ведення обліку ресурсів і виконаних робіт (accounting), управління і підтримку віртуальних організацій (V), сертифікацію Grid ПО; 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.

3. 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, обчислення за запитом, обчислення з інтенсивним обміном даними і сумісні додатки.

Треба розрізняти поняття Grid-системи і Grid-проектів, здійснювальних, наприклад, ЦЕРН в галузі фізики високих енергій на Європейській EGEE інфраструктурі, в яких сьогодні приймають участь два фізичні інститути (ХФТІ і ІТФ). При чому кластери ІТФ і Київського Національного Університету підключені до мережі AliEn-grid для обслуговування експерименту ALICE на прискорювачі LHC ; 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 НАНУ-НКАУ (по обробці космічних знімків) і Головної астрономічної обсерваторії НАНУ (по задачах астрофізики);

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

Процес реєстрації нового користувача включає два основні кроки:

- Отримання персонального призначеного для користувача сертифікату.
- Реєстрація у Віртуальній організації.

Персональний призначений для користувача сертифікат (Personal Certificate)- це свого роду електронний документ, підтверджуючий 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. Також існує можливість почати роботу в Grid скориставшись учбовими - і демо - проектами. В даному проекті передбачається, що майбутні користувачі можуть спробувати попрацювати з невеликою інфраструктурою Grid, не будучи офіційними членами якої-небудь віртуальної організації. Для більшості користувачів це найпростіший, а іноді і єдиний спосіб навчитися працювати в Grid.

4. Grid – системи України

Сьогодні важко знайти більш-менш розвинену країну, в якій не було б розгорнуто національні Grid-проекти. 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 різнорідних ресурсів, єдина обчислювальна інфраструктура повинна мати єдину архітектуру. Можливість впорядкування різнорідного початкового матеріалу забезпечується попереднім визначенням базових функцій Grid - інфраструктури..

4.1. Grid сегмент НАНУ (www.acadgrid.org.ua)

За ініціативою Інституту теоретичної фізики ім. М.М. Боголюбова (ІТФ) в квітні 2006 року в НАНУ стартувала корпоративна програма «Впровадження Grid- технологій і створення кластерів в Національній академії наук України» (або **Українська Академічна Grid Ініціативи - UAGI**), що не враховує повною мірою загальнонаціональні потреби і наукові інтереси вчених і організацій, що не працюють в системі НАНУ. Науковим керівником проекту є Анатолій Глібович Загородній , директор ІТФ НАНУ, академік НАНУ. У квітні 2007 року Національна академія наук підписала Меморандум про взаєморозуміння з WLCG (Worldwide LHC Computing Grid). у ЦЕРН і обробці і аналізу даних з LHC ((Large Hadron Collider, великого адронного колайдери).

Завершується обладнання нових кластерів в Головній Астрономічній Обсерваторії, Інституті молекулярної біології і генетики, Інституті клітинної біології і генної інженерії, модернізація кластерів в Харківському фізико-технічному інституті, в Інституті фізики конденсованих станів у Львові. Загалом комп'ютерні ресурси цієї групи складають 330 процесорів і біля 53 Терабайтів дискового масиву. Всі ці кластери будуть об'єднані в грід-сегмент НАН України. Принципи побудови цього сегменту, його конфігурації, методи керування ним вже практично відпрацьовано на прообразі цього сегменту - об'єднанні кластерів Інституту теоретичної фізики, Київського Національного Університету і Інституту кібернетики НАНУ, в якому зосереджені найпотужніші супер-

комп'ютери НАНУ, серед яких є СКІТ-3 з параметрами: 280 - процесорів (70 вузлів), Linpack продуктивність 2,235 Тфлопс, система збереження даних обсягом 10 Т байт. Передбачена також побудова кластерів в Інституті фізики (Київ), Інституті металофізики (Київ), Інституті математики (Київ), Радіоастрономічному інституті (м. Харків), Інституті сцинтиляційних матеріалів (м. Харків), Інституті геотехнічної механіки (Інституті технічної механіки), Інституті транспортних систем і технологій, Інституті чорної металургії, Інституті проблем природокористування та екології (Дніпропетровськ).

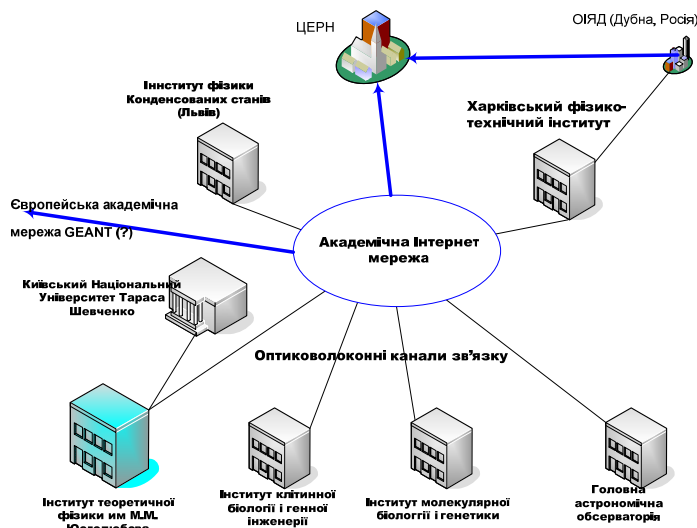


Рис.3. Grid сегмент НАНУ (слайд А.Г.Загороднього)

Слід зауважити, що реалізоване сьогодні в Grid сегменті НАНУ ПГО NorduGrid не забезпечує централізоване розподілення обчислювальних задач (Centralized workload management), загальну організацію каталогу для всій Grid інфраструктури (Global names in file catalogue for whole Grid), глобальну організацію послуг брокера і віртуалізацію сховищ даних (Different kinds of backend for storage elements), і тому планується перехід до ПГО gLite.

4.2. Проект UGrid МОНУ (www.grid.ntu-kpi.kiev.ua)

Згідно Наказу Міністерства науки і освіти (МОН) № 758 від 22 серпня 2007 року в Україні розпочата також 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 команда з 10 : різних українських організацій (2-х академічних, 6-ти освітянських і 2-х промислових), яку очолюють Інститут Системного Аналізу (ІСА) Національного Технічного Університету “ Київський політехнічний Інститут”. Науковим керівником проекту є Михайло Захарович Згуровський, ректор НТУУ”КПІ”, академік НАНУ.

Серед виконавців проекту є Інститут проблем моделювання в енергетиці імені Г.Є. Пухова НАНУ, Харківський національний університет радіоелектроніки (ХНУРЕ), Львівський національний технічний університет „Львівська політехніка”(НУЛП), Запорізький національний технічний університет (ЗНТУ), Донецький національний технічний університет (Доц.НТУ), Дніпропетровський національний гірничий університет (ДНГУ), державне підприємство ”Львівський науково-дослідний радіотехнічний інститут”(ЛНДРІ) і підприємство ЮСТАР.

Проект присвячено побудові 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 вагомим наукових додатків, які використовуються в Grid середовищі. При реалізації проекту вирішуються наступні задачі:

- Забезпечення обслуговування Української філії Світового Центру Даних (УФ СЦД), надавши його клієнтам віддалений доступ до світових сховищ наукових даних, можливості ефективного сумісного використання комп'ютерів, унікальних експериментальних установок і приладів;
- Добудування і об'єднання в співдружності з НАНУ існуючої наукової і освітньої обчислювальної і комунікаційної інфраструктури в національну Grid - інфраструктуру і підключення її до Європейської Grid – інфраструктури, що складається з множини національних Grid – інфраструктур, використання ПГО gLite;
- Розповсюдження в суспільстві знань про Grid технології і навички з їх використання на рівні, що притаманний європейським країнам, які мають більш тривалий досвід розробки та використання Grid;
- Розгортання згідно домовленості з EUGridPMA Сертифікаційного центру відкритих ключів в Україні, враховуючи наявність проекту Національної Grid інфраструктури і визнання мережі УРАН в Європі, і роз починання підготовки користувачів Grid;
- Організація обслуговування національної Grid – інфраструктури шляхом створення регіональних ресурсних-операційних центрів і постійно діючих груп з питань підтримки, експлуатації і управління Grid; розповсюдження інформації про проект і розширення круга користувачів; навчання і підготовки користувачів; визначення областей застосування і підтримки додатків; забезпечення мережних ресурсів тощо;
- Сприяння створенню спільних проектів в межах міжнародних віртуальних організацій (ВО) українських науковців з іноземними колегами з питань сталого розвитку, геофізики і комп'ютеризації проектування, забезпечивши їх інформаційну підтримку.

Передбачається протягом 2007-2008 років створити Grid інфраструктуру з 6-ю ресурсно - операційними центрами (в Києві, Харкові, Донецьку, Дніпропетровську, Запоріжжі і Львові) і розпочати відділене обслуговування майбутніх користувачів - науковців з університетів та наукових установ України. Звичайно, що ця початкова Grid інфраструктура буде добудовуватися, а число її центрів і розробників з організацій МОНУ і академічних інститутів НАНУ буде зростати. Цьому сприяє фрактальність Grid інфраструктури, тобто така її властивість, коли система подібна самої собі 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. В червні 2007 року підписана Угода з європейською організацією DANTE про підключення Національної науково-освітньої мережі УРАН до Європейської мережі GEANT. В вересні місяці очікується остаточне 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 в 2008 році будуть створені і зможуть розпочати реальну співпрацю спільні наукові команди у вигляді ВО з українських і зарубіжних вчених. Згідно з проектом в кожному з 6-ти ресурсно - операційних центрів Національної Grid інфраструктури створюються центри навчання і підтримки користувачів. 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.



Рис.4. Розтошування ресурсо- операційних вузлів UGrid

Крім того, в ПСА НТУУ"КПІ" та ряді інших вузів з вересня 2007 року розпочинається програма магістерської підготовки з напрямку „Grid технології в науці і освіті”. Всі організації - учасники проекту (крім обов’язків з організації ресурсно-операційних центрів, забезпечення їх функціонування, підготовки і підтримки користувачів) запланували проведення наукових досліджень в галузі сумісності (interoperability) ПГО, 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 з задач молекулярної динаміки, ЛНДПІ - Grid додаток з задач комп’ютерної інженерії, НУЛПІ- Grid додаток з задач обробки сигнальної інформації та медицини, ДонНТУ- Grid додаток з методики і засобів паралельного програмування, ХНУРЕ- Grid додаток з задач бізнесу і економіки і так далі. Оскільки більшість вищих навчальних закладів, академічних інститутів і промислових підприємств поки не мають суперкомп’ютерів, то в проекті передбачено розробку заходів віртуалізації ресурсів звичайних персональних комп’ютерів з допомогою Grid технологій. Навкруги кожного ведучого університету регіону групуються місцеві вищі навчальні заклади (ВНЗ), так що реальна кількість ВНЗЄ які задіяні в Ugrid проекті, перевершує 30.

Обчислювальні ресурси . UGrid складаються з суперкомп’ютера НТУУ «КПІ», найпотужнішого комп’ютерів в Україні (згідно останньому рейтингу суперкомп’ютерів країн СНД – www.supercomputers.ru) с параметрами: 84 вузлів, Linkpack продуктивність 3,13 Тфлопс , система збереження даних обсягом 12 Т байт і архівна пам’ять 20 Тб, а також кластерів Харківського національного університету радіоелектроніки, Львівського національного технічного університету, Запорізького національного технічного університету, Донецького національного технічного університету 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 з допомогою ПО (IPStor фірми FalcomStor). Забезпечення

користувачів сучасними даними (з фізики твердої Землі, сонячно-земної фізики, сталого розвитку суспільства та інше) і обчислювальними ресурсами для їх обробки здійснювалося дистанційно через спеціалізований СЦД портал і Національну комп'ютерну мережу для науки і освіти УРАН.

4.3. Сумісний проект UNGI для EGI

Інтереси держави і життя вимагають спільних зусиль інститутів НАНУ і організацій МОНУ з об'єднання існуючих сегментів наукової і освітньої обчислювальної і комунікаційної інфраструктури НАНУ і МОНН в єдину Українську Національну Grid - ініціативу (UNGI) і про інтегрування UNGI в Європейську Grid – інфраструктуру EGEE - EGI. На стороні НАНУ: авторитетне Відділення інформатики, істотні обчислювальні ресурси, практичний досвід об'єднання високопродуктивних обчислювачів в 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 і біології і кінчаючи літературознавством

На стороні вузів МОНУ і промисловості : багаторічний досвід побудови і дослідження розподілених обчислювальних систем, зокрема, системи колективного мережного проектування виробів високих технологій на технологіях Grid; досвід партнерства з європейським проектом BalticGrid; порівняльне випробування ПГО Globus, NorduGrid і gLite; офіційна угода з DANTE про підключення комп'ютерної мережі УРАН до європейської мережі GEANT; домовленість з EUGridPMA про заснування служби СА в Україні; величезний заділ з 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 з вибором пріоритетів і призначенням об'ємів фінансування на державному рівні або рівні підприємств, і кінчаючи моделюванням логічних схем з одно електронними нанотранзисторами; швидко наростаючі обчислювальні ресурси і найсучасніший досвід кластеробудування (фірма ЮСТАР); практично невичерпний резерв талановитої творчої молоді.

В цілях забезпечення повномасштабної участі України в проекті глобальної Європейської комп'ютерної інфраструктури типу Grid – EGI-EGEE Інститут теоретичної фізики НАНУ, провідна організація проекту UAGI, і Інститут прикладного системного аналізу НТУУ"КПІ", провідна організація проекту UGrid, сформувавши в серпні 2007 року спільний проект UNGI (Українська Національна Grid Ініціатива – Ukrainian National Grid Initiative), який вже включено в міжнародну 7-му Рамочну програму. Проект UNGI, згідно прийнятої в програмі EGEE структури, входить в EGI-EGEE як складова регіональної федерації «Росія» («Russia»), нарівні з ЦЕРН і іншими 8-у національними і регіональними федераціями (Франція, Італія, Великобританія/Ірландія, Німеччина/Швейцарія, Скандинавські країни, консорціуми Центрально - європейських країн, Південно-східної Європи і Південно-західної Європи). Тепер на карті Європи залишилося тільки два білих м'ятна (Білорусь і Молдова), не охоплених структурою EGI-EGEE.

Спільний проект UAGI увібрив найкраще з своїх попередників: від UAGI – тісний зв'язок з ЦЕРН і участь в його Grid проектах з фізики високих енергій, від UGrid – потужною інформаційну систему у вигляді Української філії Світового Центру Даних, що входить в мережу 52 СЦД. Величезні обсяги доступної інформації стимулюють розроблення методів її обробки, зберігання і оцінки (Data Mining), особливо в сучасних умовах, коли існуючі обсяги інформації почали перевершувати технічні можливості її зберігання. Передбачено єдиний Сертифікаційний Центр відкритих ключів, єдине ПГО 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 1.

Табл.1. Базові параметри UNGI в документах EGI

№	Параметр	Кількість
1	Як багато ресурсно- операційних центрів в Вашій Grid системі?	8 зараз і ще 9 в найближчі півроку
2	Як багато процесорів в Вашій Grid системі?	Більше 1000
3	Яка ємність існуючих сховищ даних?	Більше 100 TB
4	Фінансує чи Уряд Вашу розробку?	Так, Наказ МОНУ №758 твід 22.08.07
5	Чи існує Концепція впровадження Grid технологій в країні?	Так
6	Коли розпочалися роботи з Grid?	Листопад-04 (UAGI) Січень-06 (UGrsd)
7	Чи є національні дослідницько - освітні мережі?	УРАН і УАРНЕТ
8	Чи є центри суперкомп'ютерних обчислень?	Так, в ІК НАНУ і НТУУ «КПІ»
9	Чи займається Ваша NGI розробкою додатків?	Так
10	Скільки додатків планується створити?	20
11	До яких галузей відносяться ці додатки?	Наука, інженерія, бізнес, наука про Землю
12	Скільки людей заняті Grid розробками?	Більше 100

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

5. 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 тому, як подібним інструментарієм стали персональний комп'ютер і Інтернет.

Україна розпочала „на рівних” співпрацювати з країнами Європейського Союзу по створенню і провадженню технології 21-го століття – Grid, яка забезпечує сумісний доступ до комп'ютерних ресурсів (які змінюються від файлів і даних до 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 та інше.

Наведемо основні співтовариства, які мають потребу сьогодні в застосуванні Grid – технологій:

- Урядові організації, (службовці, експерти і науковці), що 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 , що характеризується універсальним доступом, відносно вузькою спрямованістю, динамічно змінним складом користувачів, децентралізованістю 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 Інституту теоретичної фізики НАНУ в виконанні завдань проектів ALICE, CMS, що проводяться ЦЕРНом в галузі фізики високих енергій.

- Співтовариство, що охоплює весь існуючий ринок обчислень. Цьому співтовариству властиво велике число учасників, 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 ввійти в Європейський дослідницький простір (European Research Area, ERA) повноправним і кваліфікованим партнером.

Література

1. M.Z.Zgurovsky. Development of Educational and Research Segment of Information Society in Ukraine.-Proc.WSIS , Tunis, 2004.
2. M.Z.Zgurovsky, A.I.Petrenko. Grid-infrastructure for supporting education and research in Ukraine.- Baltic Grid First All-Hands Meeting, Vilnius, 26-28 April, 2006
3. Zgurovsky, A.I. Petrenko National Ukrainian GRID Infrastructure (UGRID) for Sciences and Educations.- Proc. of CSIT-2006, Lvov, 28-30 September 2006, pp.6-1
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 . - *Data Science Journal, Volume 6, Supplement, 14 April 2007*
5. Петренко А.І. GRID технології в науці і освіті.-Матеріали 9-ої Міжнародної конференції «Системний аналіз та інформаційні технології», Київ,2007.
6. Петренко А. Национальная GRID инфраструктура для науки и образования, Конференция "HPC Day: Современные решения для высокопродуктивных вычислений", Київ,18 жовтня 2007.
7. Мартынов Е.С. Украинская Национальная GRID инициатива состояние и перспективы, Конференция "HPC Day: Современные решения для высокопродуктивных вычислений", Київ, 18 жовтня 2007.
- 8.Свистунов С.Я. GRID –инфраструктура НАНУ: опыт создания и первые результаты эксплуатации, Конференция "HPC Day: Современные решения для высокопродуктивных вычислений", Київ,18 жовтня 2007.