

Є. Мартинов, О. Смирнова

УКРАЇНСЬКИЙ НАЦІОНАЛЬНИЙ ГРІД — УЧАСНИК МІЖНАРОДНОГО ОБ'ЄДНАННЯ NORDUGRID

Грід-технології належать до порівняно нового типу розподілених обчислень, що виник наприкінці ХХ ст. внаслідок стрімкого розвитку інформаційних технологій, зокрема, високошвидкісного Інтернету. З 2004 р., коли в нашій країні було встановлено перші грід-сервери, вітчизняна академічна грід-інфраструктура значно розширилася, перетворившись в Український національний грід. У травні 2011 р. він став повноправним учасником колаборації NorduGrid, завдяки чому українські фахівці можуть впливати на розвиток європейських грід-технологій і долучатися до спільних міжнародних проектів у цій сфері.

Розподілені обчислення як метод одержання числових наукових результатів цікавили вчених з моменту появи потужних робочих станцій на базі систем типу Unix, однак донедавна саме обмежена швидкість передачі даних перешкоджала повсюдному впровадженню таких технологій. Учені Європейського центру ядерних досліджень (ЦЕРН) першими побачили в грід-технологіях альтернативу набагато дорожчим суперкомп'ютерам: обробка даних експериментів на Великому адронному колайдері потребує не стільки виконання ресурсомістких алгоритмів у якнайкоротший термін, скільки високих швидкостей для передачі великої кількості даних протягом багатьох років роботи прискорювача.

Після кількох років розроблення і тестування різних підходів обчислювальні грід-технології стали повсякденною реальністю. Досить сказати, що вся система обробки й аналізу даних експериментів на Великому адронному колайдері в ЦЕРНі базується саме на них. Для того, щоб практично в ре-

жимі реального часу зберігати й аналізувати безпрецедентний обсяг даних (більше 15 петабайтів/рік у чотирьох головних експериментах), створено спеціальну грід-інфраструктуру Worldwide LHC Computing Grid (WLCG, <http://lcg.web.cern.ch/LCG/>), яка поєднує комп'ютерні центри всіх університетів, інститутів, лабораторій світу, що ведуть дослідження з фізики високих енергій.

Грід з погляду користувача — це величезний ресурс для обчислень практично будь-якої складності, для зберігання даних будь-якого обсягу. Такий об'єднаний обчислювальний ресурс не може собі дозволити жодна окрема організація. Хоча грід, можливо, не оптимальний для всіх типів розрахунків (наприклад, не цілком задовольняє вимоги паралельних обчислень), за несприятливих фінансових умов в Україні його можна розглядати як тимчасову альтернативу надпотужним, але дорогим в обслуговуванні суперкомп'ютерам.

«Вісник НАН України» уже повідомляв про основи грід-технологій і перші кроки

© *МАРТИНОВ Євген Сергійович. Доктор фізико-математичних наук. Член координаційної ради NorduGrid. Завідувач лабораторії грід-обчислень у фізиці Інституту теоретичної фізики ім. М.М. Боголюбова НАН України. Київ.*

СМИРНОВА Оксана Георгіївна. Доктор філософії. Член координаційної ради NorduGrid. Доцент кафедри експериментальної фізики високих енергій Університету Лунда. Швеція, 2011.

побудови грид-інфраструктури в Національній академії наук (Загородній А., Зінов'єв Г., Мартинюк Є. та ін. Грид – нова інформаційно-обчислювальна технологія для науки // Вісник НАН України. – 2005. – № 6. – С. 17–25). Нагадаємо, перші грид-сервери було встановлено в ННЦ Харківського фізико-технічного інституту і незалежно від цього спільними зусиллями фахівців Інституту теоретичної фізики ім. М.М. Боголюбова НАН України і Київського національного університету імені Тараса Шевченка в Києві в 2004 р. Уже наступного року було створено перший сегмент Українського академічного гриду (УНГ), який об'єднав 5 обчислювальних грид-кластерів інститутів НАН України і КНУ.

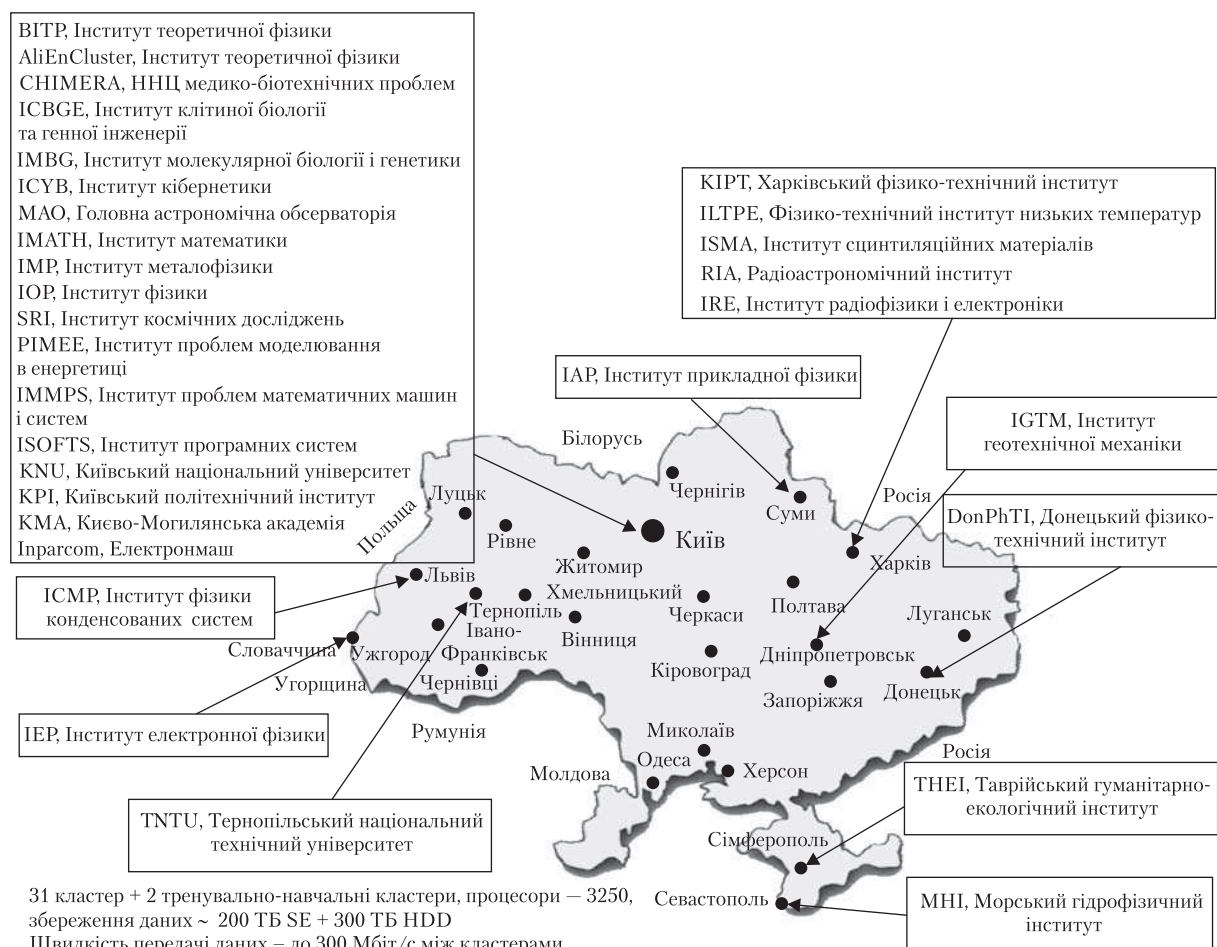
Як і в будь-якій обчислювальній грид-інфраструктурі, тут необхідно було встановити т.зв. проміжне програмне забезпечення (ППЗ), яке інтегрує всі обчислювальні ресурси й організовує їхню узгоджену роботу. Англійською ППЗ звучить як *middleware*. Творці сегмента вибрали пакет Nordugrid ARC (Advanced Resource Connector), розроблений у колаборації північних європейських країн NorduGrid (<http://www.nordugrid.org>). Це було найпростіше ППЗ, яке підтримувало всі необхідні функції для роботи грид-інфраструктури, працювало дуже стабільно і, нарешті, супроводжувалося (і супроводжується) найбільш повною і зрозумілою документацією.

Слід підкреслити, що ініціатори використання грид-технологій в Україні з самого початку не обмежилися своїми вузькими професійними інтересами (фізика високих енергій, участь у грид-інфраструктурі ЦЕРНУ). У будівництво грид-інфраструктури було залучено установи різних напрямів: Інститут молекулярної біології і генетики, Інститут клітинної біології та генетичної інженерії (зараз цей кластер переміщено до Державної установи «Інститут харчової біотехнології та геноміки»), Інститут скінтіляційних матеріалів, Головна астрономічна

обсерваторія (усі – НАН України), Інститут космічних досліджень НАН України та НКА України тощо. Такий мультидисциплінарний підхід сформульовано в академічній Програмі впровадження грид-технологій і будівництва кластерів у НАН України, успішно виконуваний до 2009 р. За ініціативою ІТФ розроблено Державну цільову науково-технічну програму впровадження і застосування грид-технологій на 2009–2013 рр. У вересні 2009 р. її затверджено постановою Кабінету Міністрів України, а фінансування (хоча й не в повному обсязі) почато з 2010 р. (<http://grid.nas.gov.ua>).

До теперішнього часу академічна грид-інфраструктура значно розширилася, об'єднавши більше 30 грид-кластерів не тільки академічних інститутів, а й університетів й інститутів Міністерства освіти і науки, молоді та спорту. Український академічний грид перетворився в Український національний грид. Його склад і географію станом на поч. 2011 р. представлено на рисунку. Цьогорічна Державна програма передбачає збільшити швидкість передачі даних між найпотужнішими кластерами УНГ до 10 Гбіт/с, що суттєво підвищить ефективність грид-обчислень.

Згідно зі стратегією побудови гриду в Україні, обраною 5 років тому, Український національний грид працює на ППЗ ARC. Водночас на трьох кластерах (в Інституті теоретичної фізики, Київському національному університеті імені Тараса Шевченка, Харківському фізико-технічному інституті), що працюють у грид-інфраструктурі ЦЕРНУ, встановлено і сертифіковано ППЗ gLite. Саме завдяки використанню ARC виник взаємний інтерес до контактів між УНГ і колаборацією NorduGrid. У травні 2011 р. УНГ став учасником NorduGrid. Меморандум про це від імені Українського національного гриду підписав керівник відповідної державної програми академік НАН України А.Г. Загородній.



Склад Українського національного ґриду і географічний розподіл ґрид-кластерів

Україна стала одним з нових учасників розширеної колаборації NorduGrid, що об'єднала не тільки скандинавські країни, які стояли біля її витоків, але й науково-дослідні організації з Литви, Словаччини, Словенії, Угорщини, Швейцарії. Історія колаборації почалася в 2001 р., коли інститути з Данії, Норвегії, Фінляндії, Швеції, що брали участь у проектах ЦЕРНу, за підтримки скандинавської мережевої інфраструктури Nordunet взялися до складання проекту створення розподіленого обчислювального центру на базі наявних потужностей. Метою було об'єднати різноманітні ресурси, які належать різним господарчим суб'єктам у низці країн, не порушуючи чин-

них правил і норм. Це суттєво відрізняється від поширеного в Європі підходу, що базується на повсюдному використанні однакової операційної системи і централізованому керуванні. Такі нестандартні вимоги й привели до необхідності створення ППЗ ARC. Правильність обраного шляху доводить стабільна і надійна робота єдиного у світі розподіленого обчислювального центру – скандинавського регіонального Tier-1, інфраструктурного проекту в межах Nordic Data Grid Facility (NDGF).

Скандинавські країни ніколи не обмежували використання ППЗ ARC як безкоштовного програмного забезпечення з відкритим кодом: тестування проводили в ба-

гатьох комп'ютерних центрах від Канади до Австралії, але вузькорегіональний характер колаборації був свого роду стримувальним фактором. До 2010 р. чимраз більше країн бажали приєднатися до колаборації. Це зумовило перегляд базової угоди, що дало можливість УНГ стати повноправним учасником NorduGrid.

Слід зазначити, що буквально з перших кроків колаборація NorduGrid стала залучати до співробітництва вчених з різних галузей. Під час розроблення ППЗ ARC було враховано найрізноманітніші вимоги, тому що жоден з обчислювальних ресурсів Скандинавії не призначений для використання тільки одною групою. У результаті доступ до обчислювальних потужностей за допомогою ARC мають не тільки фізики, але й біокібернетики, хіміки-теоретики, метеорологи, матеріалознавці, екологи.

Розширення NorduGrid практично збіглося з новою європейською ініціативою щодо розробки уніфікованого ППЗ – проектом European Middleware Initiative (EMI). Він поєднує провідних європейських розробників ППЗ ґриду: ARC, gLite, UNICORE, dCache. Дотепер перелічені пакети ППЗ були несумісні, оскільки використовували неоднакові технологічні підходи до впровадження систем типу ґрид. Це створювало безліч труднощів і для вчених, і обчислювальних центрів, змушених підтримувати кілька систем одночасно, задовольняючи вимоги різних дослідницьких груп.

Завдання проекту EMI – створити загальні підходи і впровадити стандартні засоби, поєднуючи різні ППЗ. Наприклад, заплановано розмістити елементи ППЗ ґриду в стандартних репозиторіях операційних систем Linux, роблячи ґрид доступним практично кожному. Це полегшить роботу не тільки науковців, але й операторів обчислювальних центрів і, як наслідок, підвищить ефективність використання ресурсів і якість досліджень. Більше того, нове уніфіковане

ППЗ буде привабливішим для користувачів, які вважали ґрид-технології надто складними для практичного застосування. Розширення колаборації NorduGrid дуже вчасне, зокрема, корисна участь фахівців УНГ з їхнім досвідом застосування ґрид-технологій для найрізноманітніших додатків. До того ж, приєднавшись до колаборації, українські «ґридівці» зможуть впливати на розвиток європейських ґрид-технологій навіть не будучи членами проекту EMI.

У колаборації NorduGrid стали гарною традицією регулярні робочі наради з технічних питань розвитку інфраструктури, удосконалення і створення програмного забезпечення. Такі зустрічі неабияк потрібні для команди розробників, чий офіс розкиданий по всій Європі: у повсякденній роботі сучасні інформаційні технології скорочують відстані, але ніщо не замінить живої дискусії віч-на-віч. Важливо, що наради проводять у різних країнах, щоб залучити найбільшу кількість фахівців.

Формат нарад дещо своєрідний: мінімум формальних доповідей, основний час відведено для неформальних дискусій. Вислів «у суперечці народжується істина» можна вважати неофіційним гаслом колаборації: основні технічні рішення ухвалюють саме за результатами обговорень, причому за участю як фахівців, так і потенційних користувачів. За роки існування NorduGrid робочі наради проводили не тільки в Скандинавії, але й у Росії, Естонії, Литві, Швейцарії, Словенії, Угорщині, як наслідок там зростав інтерес до ґрид-технологій, що найчастіше приводило до створення спільних проектів і розробок.

Цього року, після вступу України в NorduGrid, вирішено провести чергову технічну нараду в Києві. Вона пройшла 5–8 жовтня в Інституті теоретичної фізики ім. М.М. Боголюбова. У роботі взяли участь 20 представників Норвегії, Швеції, Данії, Фінляндії, Італії, Швейцарії, Угорщини,

Литви, а також 16 представників українських інститутів та університетів. Крім того, слухачами на засіданнях були більше 30 учасників різних грид-проектів, виконуваних відповідно до української державної грид-програми. Слід особливо підкреслити, що це були здебільшого молоді вчені і фахівці, а також студенти київських вишів. Засідання транслювали в Інтернеті, про що було попередньо оголошено на сайтах державної програми, ІТФ, Базового координаційного центру УНГ. Таким чином, суттєво розширилась аудиторія наради. Деякі учасники, які не змогли через різні обставини приїхати до Києва, робили доповіді в онлайн-режимі. На нараді найбільше часу зайняли вільні дискусії про розв'язання актуальних проблем у розробленні програмного забезпечення й організації взаємодії в реалізації різних проектів NorduGrid. Завдяки встановленій в ІТФ системі відеоконференцій у дискусіях брали участь і ті, хто перебував далеко за межами України.

Нарада почалася з пленарного засідання, на якому заслухали оглядові доповіді про Український національний грид, його стан і перспективи (Є. Мартинов, Україна), про місце NorduGrid і ППЗ ARC у європейській системі розподілених обчислень (Балаз Конья, Швеція). Далі було виголошено низку повідомлень про розвиток ARC, висловлено пропозиції щодо нових грид-сервісів з його використанням. Дуже гарне враження залишили доповіді українських учасників, зокрема А. Петренка про потокову обробку завдань (workflow) у NorduGrid і М. Барабаша про метапланувальник завдань для гриду (обидва – НТУУ «КПІ»).

Секційні засідання було присвячено обчислювальній функціональності в грид-середовищі, розвитку тематичних додатків, збереженню даних, роботі з їхніми архівами, клієнтським бібліотекам, користувацьким інтерфейсам, інформаційній безпеці в грид-

системах. Обговорювали також документування програмного забезпечення і грид-сервісів, підготовку інструкцій з їх встановлення і використання, навчання користувачів грид-технологій, тренінг системних адміністраторів грид-кластерів. Велику увагу приділили практичним питанням функціонування грид-інфраструктури, зокрема, моніторингу і контролю її роботи. Окремо торкалися взаємодії та сумісності різних ППЗ, які використовують в різних грид-інфраструктурах і національних проектах. У центрі уваги в цій частині наради було виконання проекту створення універсального ППЗ ЕМІ, що, як очікують, об'єднає в собі найкращі риси відомих ППЗ.

Приємно, що високу оцінку одержав внесок українських фахівців (А. Сальників, Є. Слюсар з КНУ, А. Головинський, А. Маланко з ІК ім. В.М. Глушкова, Л. Білоус з ФТІНТ ім. Б.І. Веркіна) у розв'язання цих та інших актуальних проблем функціонування грид-інфраструктури під керуванням ППЗ ARC. Особливий інтерес викликають методи його налаштування й адаптації до конкретних науково-дослідних завдань і специфічних умов у різних обчислювальних центрах. УНГ як велика і різноманітна інфраструктура дає можливість перевірити на практиці різноманітні варіанти використання ППЗ і грид-сервісів. Українські фахівці не тільки виявляють проблеми або відсутні елементи у програмних пакетах і сервісах, але й пропонують власні професійні рішення. За підсумками наради А. Сальникову доручено здійснювати підтримку одного з елементів ARC (nordugridmap – компонента, відповідальна за фактичний доступ користувачів до обчислювального ресурсу), що свідчить про визнання високої кваліфікації українських фахівців з грид-технологій.

Для реалізації планів колаборації NorduGrid і побажань користувачів гриду з різних країн і сфер застосування потрібно залучити якомога більше розробників і випробувачів.

А крім того — вивчити нові варіанти і засоби використання грид-технологій, адаптувати наявне і розробити нове спеціалізоване прикладне програмне забезпечення для використання переваг грид-обчислень у наукових і науково-технічних галузях. Слід широко використовувати вітчизняний потенціал для участі в спільних міжнародних проектах, які може пропонувати й реалізувати не тільки вся колаборація, але й окремі учасники.

УНГ має позитивний досвід у навчанні молодих фахівців. Об'єднання зусиль у цьому напрямі — запорука розвитку грид-технологій у країнах-учасницях NorduGrid. Грид-технології не стоять на місці, їхній розвиток — актуальний і важливий для кожного з учасників колаборації. Київська робоча нарада NorduGrid продемонструвала, що склався міжнародний колектив односторонців, у якого є спільні цілі, а головне науковий і людський потенціал, щоб їх досягти.

Колаборація NorduGrid безумовно захищена в поширенні грид-технологій в Україні й інших країнах. Від імені цієї організації та всіх її учасників просимо керівництво Національної академії наук, її інститутів і установ максимально розширити і підтримати участь українських учених і фахівців, особливо молодих, у розвитку і застосуванні грид-технологій у повсякденній науковій практиці. Слід якнайшвидше розв'язати проблему офіційної участі Українського національного гриду в організації European Grid Infrastructure, покликаний створити і підтримувати єдину грид-мережу, яка складається з європейських національних грид-інфраструктур. Ці зусилля, без сумніву, забезпечать українській науці гідне місце і роль у міжнародній науковій спільноті.

Є. Мартинов, О. Смирнова

УКРАЇНСЬКИЙ НАЦІОНАЛЬНИЙ ГРИД — УЧАСНИК МІЖНАРОДНОГО ОБ'ЄДНАННЯ NORDUGRID

Резюме

У статті висвітлено специфіку застосування грид-технологій, широкі можливості їх використання для науково-технічних досліджень. Наголошено на важливості щорічного вступу України до колаборації NorduGrid, що об'єднує декілька європейських країн, де активно розвиваються грид-технології. Знаковим стало проведення чергової технічної наради NorduGrid у Києві. На ній було порушено практичні питання функціонування грид-інфраструктури, взаємодії та сумісності різних типів програмного забезпечення, обчислювальної функціональності в грид-середовищі, інші актуальні теми. Автори сподіваються, що Національна академія наук і надалі сприятиме участі українських фахівців у розвитку і застосуванні грид-технологій у повсякденній науковій практиці.

Ключові слова: проміжне програмне забезпечення, кластери і сегменти, розподілені обчислення.

Ye. Martynov, O. Smyrnova

UKRAINIAN NATIONAL GRID — MEMBER OF INTERNATIONAL NORDUGRID COLLABORATION

Abstract

The peculiarities of grid technologies usage and wide prospects of their application in scientific and technical researches are pointed. This year Ukrainian national grid has been joined to the NorduGrid collaboration, it is highly appraised. That collaboration unites several European states where grid technologies are developing actively. NorduGrid technical meeting held this time in Kyiv is called significant. The practical problems of grid infrastructure operations, interactions and compatibility in different software, computing functionality in grid medium, other urgent themes were discussed at the meeting. Authors hope National academy of sciences gives further favor to Ukrainian specialists taking part in developing and applying the grid technologies in everyday scientific practice.

Keywords: middleware, clusters and segments, distributed computing.