

МІНІСТЕРСТВО ОСВІТИ І НАУКИ, МОЛОДІ ТА СПОРТУ
УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ»
Навчально-науковий комплекс «Інститут прикладного системного
аналізу»

Петренко А.І., Свістунів С.Я., Кисельов Г.Д

ПРАКТИКУМ З ГРІД- ТЕХНОЛОГІЙ

Призначений для магістерської підготовки за спеціальністю 8.05010103 „Системне проектування”, але може бути корисним і для інших спеціальностей напрямків „Комп’ютерні науки” і „Комп’ютерна інженерія”.

Київ

2011

УДК 004.75; 681.3

Петренко А.І., Свістунов С.Я., Кисельов Г.Д.

Практикум з грід-технологій. – Київ: НТУУ «КПІ», 2011. – 447 с.: іл.

Анотація: Грід-технологій дозволяють об'єднати інформаційні і обчислювальні ресурси шляхом створення комп'ютерної інфраструктури нового типу, яка забезпечує глобальну інтеграцію цих ресурсів на основі мережових технологій і спеціального програмного забезпечення проміжного рівня (між базовим і прикладним ПЗ), а також набору стандартизованих служб для забезпечення надійного спільного доступу до географічно розподілених інформаційних і обчислювальних ресурсів: окремим комп'ютерів, кластерів, сховищ даних.

Метою практикуму з грід-технологій є ознайомлення студентів з концепцією грід-комп'ютингу та поточним станом справ у цій області та набуття практичних навичок з використання грід-технологій для аналізу, оптимізації та проектування інженерних об'єктів. Практикум призначений для проведення лабораторних занять з дисципліни «Грід-технології для розподілених обчислень та обробки даних» магістерської підготовки за спеціальністю „Системне проектування”, але може бути корисним і для інших спеціальностей напрямків „Комп'ютерні науки” і „Комп'ютерна інженерія”. Практикум також може бути рекомендований студентам та аспірантам усіх спеціальностей, які вивчають сучасні інформаційні технології в межах своїх предметних галузей, а також їх викладачам.

Відповідальний редактор В.С. Рогоза, д-р техн.наук, проф.

Рецензенти: С.С. Забара, д-р техн.наук, проф.

С.Г.Стіренко, к-т.техн.наук, доц.

ЗМІСТ

| | |
|--|-----------|
| Передмова..... | 7 |
| <i>Завдання № 1. Вступ до грід-технології.....</i> | <i>12</i> |
| 1 Короткі теоретичні відомості | 12 |
| 1.1 Концепція грід..... | 12 |
| 1.2 Міжнародні грід-проекти..... | 22 |
| 1.3 Грід в Україні | 34 |
| 1.4 Навчальна грід-інфраструктура | 41 |
| 1.5 Порядок виконання завдання: | 47 |
| 1.6 Контрольні запитання | 48 |
| Література..... | 48 |
| <i>Завдання № 2. Робота на обчислювальному кластері з використанням локальної системи управління PBS</i> | <i>50</i> |
| 2.1 Короткі теоретичні відомості | 50 |
| 2.1.1 Багатопроцесорні обчислювальні системи | 50 |
| 2.2 Методика роботи на обчислювальному кластері..... | 61 |
| 2.3 Корисні програми для роботи на кластері: | 69 |
| 2.4 Приклад виконання роботи | 82 |
| 2.5 Порядок виконання лабораторної роботи..... | 88 |
| 2.6 Завдання до лабораторної роботи..... | 89 |
| 2.7 Контрольні запитання | 89 |
| Література..... | 90 |
| <i>Завдання № 3. Отримання сертифікату користувача.....</i> | <i>91</i> |
| 3 Короткі теоретичні відомості | 91 |

| | |
|---|------------|
| 3.1 Інфраструктура відкритого ключа | 91 |
| 3.2 Ідентифікація користувачів і грид вузлів..... | 96 |
| 3.3 Делегування прав і використання довіреності | 98 |
| 3.4 Сервіс управління віртуальними організаціями і авторизація користувачів | 101 |
| 3.5 Користувач у грид | 106 |
| 3.6 Порядок виконання роботи | 108 |
| 3.7 Завдання до лабораторної роботи..... | 121 |
| 3.8 Контрольні питання..... | 122 |
| Література..... | 122 |
| <i>Завдання № 4. Проміжне програмне забезпечення грид - ARC</i> | <i>123</i> |
| 4 Короткі теоретичні відомості | 123 |
| 4.1 Архітектура та основні компоненти ARC..... | 125 |
| 4.2 Система керування завданнями | 137 |
| 4.3. Команди інтерфейсу користувача..... | 143 |
| 4.4 Команди роботи з даними..... | 163 |
| 4.5 Команди тестування | 169 |
| 4.6 Опис завдання | 171 |
| 4.7. Приклад виконання лабораторної роботи..... | 179 |
| 4.8 Завдання до лабораторної роботи..... | 185 |
| 4.9. Контрольні питання..... | 186 |
| <i>Завдання №5. Проміжне програмне забезпечення грид - gLite</i> | <i>188</i> |
| 5 Короткі теоретичні відомості | 188 |
| 5.1. Архітектура та основні компоненти GLITE | 189 |

| | |
|---|------------|
| 5.2 Система управління завданнями..... | 206 |
| 5.3 Опис завдання | 228 |
| 5.4 Приклад виконання лабораторної роботи..... | 246 |
| 5.5 Завдання до лабораторної роботи..... | 255 |
| 5.6 Контрольні питання..... | 256 |
| <i>Завдання № 6. Інформаційна система і моніторинг.....</i> | <i>257</i> |
| 6 Короткі теоретичні відомості | 257 |
| 6.1 Інформаційна система gLite | 257 |
| 6.1.1 Система Globus Monitoring and Discovery Service (MDS) | 258 |
| 6.2 Система (Relational Grid Monitoring Architecture) R-GMA | 271 |
| 6.3. Інформаційна система NorduGrid ARC | 273 |
| 6.4 Моніторинг грід-мереж. Основні положення..... | 275 |
| 6.5 Приклади засобів моніторингу грід-мережі | 279 |
| 6.6 Доступ до даних інформаційної системи gLite | 287 |
| 6.8 Доступ до даних інформаційної системи ARC | 302 |
| 6.9 Завдання до лабораторної роботи | 307 |
| 6.10 Контрольні питання..... | 307 |
| Література..... | 308 |
| <i>Завдання № 7. Реєстрація і отримання користувачем доступу до грід-системи через веб- сервіс.....</i> | <i>310</i> |
| 7.1 Короткі теоретичні відомості. Концепція наукового шлюзу (Science Gateways) | 310 |
| 7.2 Інструментальні засоби побудови грід- порталів застосувань .. | 315 |
| 7.3 Інструментальні засоби на основі Globus Toolkit | 316 |
| 7.4 Інструментальний засіб GridSpeed | 318 |

| | |
|---|-----|
| 7.5 Проект Grid Programming Environment | 320 |
| 7.6 Інструментальний засіб GridSphere Portal Framework..... | 327 |
| 7.7 Приклад виконання практичної роботи | 334 |
| 7.8 Завдання до практичної роботи | 340 |
| 7.9 Контрольні питання..... | 341 |
| Література..... | 341 |
| Додатки до практикуму з грід-технологій | 344 |

Передмова

Грід 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 програмне забезпечення (бібліотеки, програми). Таким чином, грід претендує на роль універсальної інфраструктури для обробки даних, в якій функціонує множина сервісів, які дозволяють дати нову якість рішення різних класів задач. Особливо – таких, які неможливо вирішити у адекватні строки локально на одному, навіть найпотужнішому комп'ютері:

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

В умовах розподіленого програмного забезпечення, а також віддалених робочих місць користувачів і ресурсів виконання операцій у грід ініціюється видачею запиту до деякої служби (сервісу) і здійснюється, як правило, послідовністю взаємодіючих служб, кожна з яких виконує частину необхідної обробки. Надійність обробки запиту забезпечується шляхом моніторингу збоїв і повторного запиту на альтернативних ресурсах.

Грід-технології і всесвітня грід-мережа ідуть на зміну вже звичному Інтернету з його веб - послугами як засіб сумісного використання обчислювальних потужностей та сховищ даних. Тому серед пріоритетних напрямків інформатизації та розвитку сучасних інформаційних технологій в Україні в Постанові Кабінету Міністрів

України №1020 від 23 вересня 2009 року «Державна цільова науково-технічна програма впровадження і застосування грід-технологій на 2009-2013 роки» були зазначені задачі створення Національної грід-інфраструктури для наукових досліджень і розгортання підготовки і перепідготовки фахівців, здатних **забезпечити функціонування грід-середовища і застосування грід-технологій в науці, освіті та інших галузях**, оскільки грід-технології є досить новим напрямком в інформаційних технологіях і в Україні бракує таких фахівців.

Сфера застосування грід-технологій не обмежується лише вирішенням складних наукових і інженерних задач. Із розвитком грід проникає в промисловість і бізнес. Грід-технології вже активно застосовуються у світі як державними організаціями управління, оборони, сфери комунальних послуг, так і приватними компаніями, наприклад, фінансовими і енергетичними. Галузь застосування грід зараз охоплює ядерну фізику, захист навколишнього середовища, прогноз погоди і моделювання кліматичних змін, чисельне моделювання в машино - і авіабудуванні, біологічне моделювання, фармацевтику тощо.

Постановою Ради Міністрів України в 2011 році була з ініціативи ННК «Інститут прикладного системного аналізу» введена нова спеціальність **«Системне проектування»** з напрямку «Комп'ютерні науки» для підготовки фахівців у сфері застосувань розподілених інтелектуальних середовищ і, зокрема, грід-технологій в науці і освіті. Для її реалізації розпочато викладання дисципліни **„Грід – технології для розподілених обчислень та обробки даних ”**, яка узагальнює сьогоденні уявлення про грід - технології та проблеми, які виникають на шляху їх розроблення і впровадження. У результаті вивчення дисципліни студенти повинні:

- **знати:** основи грид – технологій, які дозволяють об’єднати обчислювальні ресурси та ресурси зберігання даних, архітектуру грид – систем, які використовуються в Україні, принципи функціонування основних складових частин грид–системи, технологію підготовки завдань для використання грид – середовища;

- **вміти:** вибирати і використовувати проміжне програмне забезпечення для вирішення науково – практичних завдань, адаптувати пакети прикладних програм до середовища грид, використовувати вхідні мови опису завдання і даних, відслідковувати хід обчислювального процесу під час числового експерименту чи процесу моделювання;

- **набудуть навички:** практичного використання найбільш поширеного проміжного програмного забезпечення gLite та ARC.

Даний практикум містить необхідний теоретичний матеріал, приклади виконання завдань і є навчально-методичним забезпеченням дисципліни «*Грид – технології для розподілених обчислень та обробки даних*». Він призначений для прискорення підготовки фахівців з грид - технологій, навчання їх роботі з цими новими засобами, зокрема, при застосовуванні в чисельних розрахунках, що потребують великих комп’ютерних ресурсів. Він складається з виконання циклу експериментальних досліджень, націлених на набуття практичних навичок використання грид – технології для вирішення науково – практичних завдань:

- **робота на обчислювальному кластері** з використанням локальної системи управління **PBS** для вивчення технології віддаленого доступу до ресурсів багатопроцесорної обчислювальної системи та набуття практичних знань та навичок компіляції та запуску простих програм з використанням системи управління кластера;

- **отримання сертифікату користувача**, для чого користувач вчиться генерувати пару ключів: відкритий ключ (як частина сертифікату) та закритий ключ (для підпису сертифікату); отримувати сертифікат, створювати проксі-сертифікат та користуватися їм;

- **вивчення технології видаленого доступу** до грід-ресурсів, що працює під управлінням проміжного програмного забезпечення **ARC**, коли користувач, використовуючи обчислювальний кластер під управлінням UNIX - сумісною ОС, розробляє файл завдання на мові **xRSL** для запуску через ППЗГ ARC згідно варіанта завдання, запускає задачу з сервера доступу, попередньо створивши проксі-сертифікат, під час виконання задачі перевіряє її статус;

- **вивчення технології видаленого доступу** до грід-ресурсів, що працює під управлінням проміжного програмного забезпечення **gLite**, коли користувач розробляє файл завдання **на мові JDL** для запуску через проміжне програмне забезпечення **gLite** згідно варіанта завдання, запускає задачу з сервера доступу, попередньо створивши проксі-сертифікат, перевіряє її статус під час виконання, отримує результати виконання і аналізує їх;

- **вивчення підходів до побудови інформаційної системи і системи моніторингу** промислових грід – мереж, що працюють під управлінням проміжного програмного забезпечення **gLite** і **ARC** і вирішують задачу збору і управління даними про стан грід, отримуючи інформацію від безлічі розподілених джерел - постачальників. Підсистема призначена для постійного контролю функціонування грід-системи і забезпечення своєчасного реагування на виникаючі проблеми.

Таким чином, за допомогою цього практикуму можна, отримавши сертифікат, готувати свої задачі для запуску, вибирати обчислювальний елемент, на якому виконуватиметься завдання, контролювати процес

виконання, проглядати результати виконання завдання і зберігати вихідні файли на комп'ютері користувача. Передбачається, що користувачі практикуму зможуть під час навчання через українську національну грид- інфраструктуру набути практичний досвід роботи з обчислювальними ресурсами суперкомп'ютерів НТУУ"КПІ" і Інституту теоретичної фізики та інших, об'єднаних в сумісну мережу, а також з інформаційними ресурсами Світового Центру Даних (СЦД) «Геоінформатика і сталий розвиток».