

УДК: 004.457:681.324

ПОЄДНАННЯ ГРІД-ТЕХНОЛОГІЙ ASTRO-WISE I EGEE

студент, Клівода І. В.

Національний технічний університет України «Київський політехнічний інститут», Україна, Київ

В даній статті описано інтеграцію Грід-рішення, що реалізоване в інформаційній системі Astro-WISE з застосуванням EGEE Grid і портування програмних додатків Astro-WISE на EGEE. Astro-WISE це середовище, що складається з апаратного та програмного забезпечення. Описані переваги та недоліки цих технологій. Показано архітектуру Astro-WISE Grid, визначено проблеми для інтеграції Грід-інфраструктур і вирішення цих проблем. Причиною для об'єднання Astro-WISE і EGEE є впровадження міжнародних стандартів, які використовуються EGEE Грід в систему Astro-WISE, який полегшує будь-яке майбутнє підключення до інших мереж.

Ключові слова: Грід-системи, інформаційна система, інтеграція, розширення можливостей, переваги.

студент, Кливода И.В. соединение Грид-технологий ASTRO-WISE и EGEE / Национальный технический университет Украины «Киевский политехнический институт», Украина, Киев

В данной статье описано интеграцию Грид-решения, реализованное в информационной системе Astro-WISE с применением EGEE Grid и портирование приложений Astro-WISE на EGEE. Astro-WISE это среда, состоящая из аппаратного и программного обеспечения. Описанные преимущества и недостатки этих технологий. Показано архитектуру Astro-WISE Grid, определены проблемы для интеграции Грид-инфраструктур и решения этих проблем. Причиной для объединения Astro-WISE и EGEE является

внедрение международных стандартов, используемых EGEE Грид в систему Astro-WISE, который облегчает любое будущее подключения к другим сетям.

Ключевые слова: Грид-системы, информационная система, интеграция, расширение возможностей, преимущества.

student, Klivoda I.V. Confluence Grid Technologies: Astro-WISE and EGEE / National Technical University of Ukraine "Kyiv Polytechnic Institute", Ukraine, Kiev

This paper describes the integration of Grid solutions, implemented in the information system Astro-WISE using EGEE Grid and porting software applications Astro-WISE on EGEE. Astro-WISE is an environment that consists of hardware and software. Describe the advantages and disadvantages of these technologies. Architecture shown Astro-WISE Grid, defined problem to integrate Grid infrastructures and address these problems. The reason for the association Astro-WISE and EGEE is introduction of international standards used EGEE Grid in Astro-WISE, which facilitates any future connection to other networks.

Keywords: Grid system, information system, integration, empowerment, benefits.

Вступ. Зберігання даних та обробки даних проблем в астрономії в порівнянні з тими ж проблемами в галузі ядерної фізики, метеорології, наук про Землю, біології та часто навіть перевищують ці дисципліни в складності структури даних, які повинні бути оброблені і архівовані.

Історично, багато з Грід-ініціатив розроблено самостійно. Вони створили масивне апаратне та програмне забезпечення інфраструктури, і ряд проектів намагаються об'єднати різні Грід-системи. Одним із прикладів є проект ROSSGRID.

Інформаційна система Astro-WISE [1], є прикладом конкретного Грід-рішення спочатку розроблена для зберігання і обробки даних астрономічних зображень. На відміну від CROSSGRID, інтеграція Astro-WISE і EGEE здійснюється на рівні інфраструктури, а не на рівні всієї програми. Кінцева мета цієї інтеграції - зробити можливим для "майстер" Грід-інфраструктури

(Astro-WISE) можливість використовувати ресурси otherGrid (EGEE) без будь-яких змін у вже розроблених користувацьких додатках. Такий підхід значно економить час при перенесенні численних астрономічних програм, запущених на Astro-WISE. У цьому випадку, Astro-WISE підходить EGEE для зовнішнього зберігання і обробки даних об'єктів повністю інтегрованих в систему Astro-WISE, і EGEE [2].

Мета роботи. Основна причина для об'єднання Astro-WISE і EGEE - зробити Astro-WISE доступним для більшої кількості спільноти користувачів, особливо користувачів, які вже використовують EGEE ресурси для зберігання і обробки даних. LOFAR Long Term Archive, наприклад, буде побудований на вершині EGEE Грід, а також на Astro-WISE інфраструктурі. Astro-WISE буде використовуватися для зберігання метаданих, параметрів обробки та організації низки переробних трубопроводів. EGEE буде використовуватися для зберігання частини даних LTA і обробляти його [3]. Щоб зробити це можливим, треба розширити можливості Astro-WISE, в тому числі зберігання EGEE в системі Astro-WISE і дозволити обробку даних Astro-WISE і з Astro-WISE трубопроводів на EGEE. У той же час, потрібно забезпечити Astro-WISE роботу з метаданими та організацією обробки без помилок, тому що вони є ключовими перевагами системи. В результаті, зберігання розподіляється по вузлах Astro-WISE (сервера даних) і EGEE вузлів зберігання.

Виклад основного матеріалу статті. Одне з відмінностей - опис EGEE, що gLite намагається побудувати Грід за міжнародними стандартами, якщо це можливо. Astro-WISE, з іншого боку, була розроблена, щоб служити певній меті, і розроблені свої власні методи, де це необхідно, засновані на стандартних протоколів та інтерфейсах. Він використовує об'єктний підхід і мову Python в загальних рамках, також стандартний протокол HTTP для доступу до сервера даних.

Остаточною і, ймовірно, найбільш важливою відмінністю є те, що EGEE фокусується на основних інфраструктурних компонентів, що забезпечують її обчислення, зберігання та оперативні аспекти [4].

Аутентифікація Astro-WISE відбувається за допомогою імені користувача та пароля, необхідні для інтеграції з сертифікатами X.509, що використовуються в Грід [5]. Тому треба прийняти рішення, яке буде добре інтегруватися з існуючими методами.

Для отримання проксі сертифікату від користувача на сервері Astro-WISE ППН можуть бути використані два методи. Перший використовувати проксі делегації, по-друге, щоб використовувати сервер MyProху в якості проміжної машини. Ми вирішено зробити використання MyProху сервера, тому що це рішення легше реалізувати, а також рекомендований спосіб роботи.

Другим кроком обрано легкий клієнт, який може бути використаний для зберігання проксі в MyProху сервера, який не залежить від великого стека програмного забезпечення для установки. В якості тимчасового рішення ми розгорнули Astro-WISE на установці gLite UI, але було встановлено, що gLite занадто громіздкий для установки в інших місцях, і управління обліковими записами для всіх користувачів централізовано, що теж не варіант.

Тому Astro-WISE це серeda Python і зробити реалізацію клієнтських інструментів MyProху буде кращим рішенням. Враховуючи обмежену робочу силу, відкинуто рішення розробляти проект з нуля. Найбільш перспективним рішенням був розвиток NDG безпеки в OMII-UK36 та НКРЕ даних Grid37 проектів.

Для того, щоб отримати облікові дані Грід Astro-WISE, користувач повинен зберігати проксі на сервері MyProху, використовуючи ті ж ім'я користувача та пароль, що використовується для системи Astro-WISE.

Щоб включити доступ до DPU EGEE обчислень, деякі великі зміни в інтерфейсі сервера повинні бути реалізовані. Основні зміни були в тому, як представити роботу, як контролювати роботу і як зробити Грід-аутентифікацію в доброзичливій формі. Ще одна проблема в тому, що програмне забезпечення Astro-WISE повинно бути встановлено на призначених EGEE обчислювальних вузлах. Для цього, попередньо скомпільованих, саморозпаковуються пакетів були зроблені з програмного забезпечення Astro-WISE, який буде завантажений

в EGEE-вузлах і виконується перед запуском фактичної роботи Astro-WISE. Для EGEE ресурсних центрів, які дозволяють встановити програмне забезпечення Astro-WISE локально, цей крок, звичайно, не потрібен.

Тому Astro-WISE здатний використовувати кілька процесорів завдяки паралельно́ї обробки, паралельні дії можуть бути представлені в EGEE Grid. Застосування паралельного типу роботи має більшу перевагу, ніж автоматична синхронізація часу початку процесів. Це також викликає виконання роботи на одному місці, використовуючи аналогічні ядра процесора. Всі комунікації між частинами працюють за допомоги бази даних, що робить їх незалежними від підтримки бібліотек паралелізму, як MPI або Pthreads.

Насамперед, інтерфейс до Грід-інфраструктурі EGEE необхідний для отримання та зберігання даних з / в Грід.

По-друге, якщо кілька протоколів може бути використано для одержування або збереження файлу, система повинна бути в змозі надати інтерфейс, щоб дозволити користувачам вибрати протокол, щоб використовувати, або відновити або зберегти файл за допомогою своєї логіки за замовчуванням.

Нарешті, з міркувань безпеки, система повинна забезпечувати механізм, який відокремлює право DataObject і пов'язаних файлів даних доступу. Це необхідно в таких додатках, як LOFAR, де різні групи або користувачі можуть мати різні привілеї, щоб отримати відповідні файли. Загальне рішення було розглянути питання про право доступу використання (зберігання) окремої таблиці для зберігання URI файлу. URI, що зберігається в таблиці, повинна бути постійною в часі, так що не доведеться оновлювати це, коли, наприклад, файл переміщується з диска на стрічку. Переваги цього рішення полягають в наступному:

- Таблиця для зберігання може мати кілька записів для одного і того ж об'єкта, тому ми можемо управляти кількома копіями (резервні копії) файлів.

- Привілеї таблиці зберігання можуть відрізнятися від привілеїв інших метаданих, які додає додатковий рівень безпеки в системі. Наприклад, кожен

може побачити метадані файлів, а тільки привілейовані користувачі можуть побачити URI.

Для того, щоб відновити файли з мережі, повинно бути забезпечено або GUID або SRM файли. Таким чином, один з них повинен бути в таблиці зберігання. GUID є унікальним ідентифікатором файлу, в той час як SRM є URL зберіганням копії файлу. Якщо GUID файлу зберігається в таблиці зберігання, LFC необхідно для перегляду кількох копій файлу.

Різні класи протоколу реалізують конкретні дії на отримання і введення даних з або в сховище. Вони виконують вибірку і зберігають завдання відповідно до типу зберігання. Наприклад, користувач хоче отримати файл, пов'язаний з DataObject data_obj. Він може використовувати метод вилучення з класу DataObject, щоб отримати файл, тобто data_obj.retrieve (). Метод, в свою чергу DataObject "retrieve" називає відповідний метод вилучення з об'єкта зберігання, в цьому випадку, LTASStorage. LTASStorage має свою логіку, щоб витягти файл. Воно шукає в таблиці об'єкт файлу, і отримує список об'єктів файлів, які пов'язані з DataObject data_obj. Це сортує список об'єктів файлів, перший FileObject в списку витягується за допомогою відповідного протоколу, в даному випадку Грід-протокол. Витягнений файл потім повертається до DataObject.

Перевага використання цієї структури - вона може бути легко розширена для підтримки інших систем зберігання даних. Те, що ми повинні зробити, це додати відповідний інтерфейс протоколу системи зберігання, яка буде використовуватися для отримання і помістити дані з або в сховище.

Висновки. Найближче майбутнє принесе ряд додатків, що використовують інформаційну систему Astro-WISE і його клонів, крім трьох описаних вище. У Північних Нідерландів велика багатопрофільна ініціатива, на ім'я TARGET, що була запущена для подальшого розгортання типових орієнтованих даних.

Для проекту UMCG Lifelines, геномні та фенотипічні дані 165000 пацієнтів зібрані протягом більше 35 років і зберігаються в системі, що розроблена на

основі Astro-WISE (Life-WISE). Взаємозв'язок з промисловими продуктами, таких як Oracle LifeScience чудовий і потребує подальшого вивчення.

Література:

1. Valentijn, E., et al.: *Astro-WISE: chaining to the universe*. In: Shaw, R.A., Hill, F., Bell, D.J. (eds.) *Proc. of ADASS XVI, ASP Conf. Ser.*, vol. 376, 491 (2007)
2. Taffoni, G., et al.: *Enabling Grid technologies for Planck space mission*. *Future Gener. Comput. Syst.* 23, 129 (2007)
3. Boku, T., et al.: *HMCS-G: Grid-enabled hybrid computing system for computational astrophysics*. *Third IEEE international symposium on cluster computing and the Grid (CCGrid'03)*, p. 558 (2003)
4. Ibarra, A., et al.: *On-the-fly XMM-Newton spacecraft data reduction on the Grid*. *Sci. Program.* 14, 141 (2006)
5. Taffoni, G., et al.: *Bridging the virtual observatory and the Grid with the query element*. *Grid Workshop, astro-ph/0605165* (2006)

References:

1. Valentijn, E., et al.: *Astro-WISE: chaining to the universe*. In: Shaw, R.A., Hill, F., Bell, D.J. (eds.) *Proc. of ADASS XVI, ASP Conf. Ser.*, vol. 376, 491 (2007)
2. Taffoni, G., et al.: *Enabling Grid technologies for Planck space mission*. *Future Gener. Comput. Syst.* 23, 129 (2007)
3. Boku, T., et al.: *HMCS-G: Grid-enabled hybrid computing system for computational astrophysics*. *Third IEEE international symposium on cluster computing and the Grid (CCGrid'03)*, p. 558 (2003)
4. Ibarra, A., et al.: *On-the-fly XMM-Newton spacecraft data reduction on the Grid*. *Sci. Program.* 14, 141 (2006)
5. Taffoni, G., et al.: *Bridging the virtual observatory and the Grid with the query element*. *Grid Workshop, astro-ph/0605165* (2006)