

Дорошенко А.В., Булах Б.В.

ННК “Інститут прикладного системного аналізу” НТУУ “КПІ”, Київ, Україна

Інтеграція і менеджмент даних в семантичному гріді

Метою технології грід є глобальна інтеграція інформаційних і обчислювальних ресурсів на основі мережевих технологій і спеціального програмного забезпечення проміжного рівня, а також набору стандартизованих служб для забезпечення надійного сумісного доступу до географічно розподілених інформаційних і обчислювальних ресурсів: окремих комп'ютерів, кластерів, сховищ даних і мереж. Семантичний грід [1] як розвиток існуючого гріда, об'єктам якого (ресурси, сервіси) надано чітких визначень (придатних для машинної обробки), є інфраструктурою з покращеними можливостями для кооперації людей та програмних агентів [2]. Таким чином, семантичний грід має ефективно вирішувати і задачу інтеграції даних з різних джерел, з розбіжностями у їх моделях, часі отримання, якості тощо.

До небагатьох існуючих програмних рішень, орієнтованих на інтеграцію та управління даними у грід-середовищі, відноситься проект OGSA-DAI [3] — продукт, призначений для вирішення задачі організації доступу до SQL-орієнтованих баз даних на основі сертифікатів стандарту X.509, який використовується в багатьох наукових і бізнес-проектах. Забезпечуючи можливість доступу до неоднорідних ресурсів даних, засоби OGSA-DAI можуть скласти основу для створення спеціалізованих грід-сервісів по роботі з даними та знаннями, а також семантичних грід-додатків.

OGSA-DAI спирається на сервісно-орієнтовану архітектурну модель, як і OGSA-DQP — сервіс-орієнтований обробник розподілених запитів, здатний паралельно виконувати запити до різних ресурсів, доступних через OGSA-DAI, а також використовувати інші веб-сервіси для обробки отриманих даних. OGSA-DAI дозволяє працювати з набором джерел даних як з єдиною «віртуальною БД» та підтримує каскадні проміжні обчислення відповіді на глобальний запит у гріді (у сценаріях аналізу даних тощо), що характерно для наукових досліджень [4].

Зважаючи на потреби українського гріда у високорозвинутих середовищах (програмних комплексах, порталах), орієнтованих на роботу з даними та знаннями, актуальною є задача дослідження можливостей OGSA-DAI для виконання у грід складних сценаріїв обробки (агрегації, трансформації, узгодження та і т.п.) даних та вилучення знань, що залучають різноманітні СУБД та БД, а також шляхів інтеграції OGSA-DAI у архітектуру семантичного гріда. Досліджувана тестова інфраструктура (рис.1) може слугувати основою для подальшої розробки подібних середовищ.

Література. 1. Згуровский М. З. Е-наука на шляху до семантичного Грід. Частина 2: Семантичний Web- і семантичний Грід / М. З. Згуровский, А. І. Петренко // Системні дослідження і інформаційні технології. — К., 2010. — №2. — С. 7–25. 2. De Roure D. The Semantic Grid: Past, Present and Future / D. De Roure, N. R. Jennings, N. R. Shadbolt // Proceedings of the IEEE. — 2005. — 93(3). — P. 669–681. 3. Comito C. A service-oriented system for distributed data querying and integration on Grids / C. Comito, A. Gounaris, R. Sakellariou, D. Talia // Future Generation Computer Systems. — 2009. — Vol.25, Issue 5. — P. 511–524. 4. Ailamaki A. Managing scientific data / A. Ailamaki, V. Kantere, D. Dash // Communications of the ACM. — 2010. — 53(6). — P.68–78.

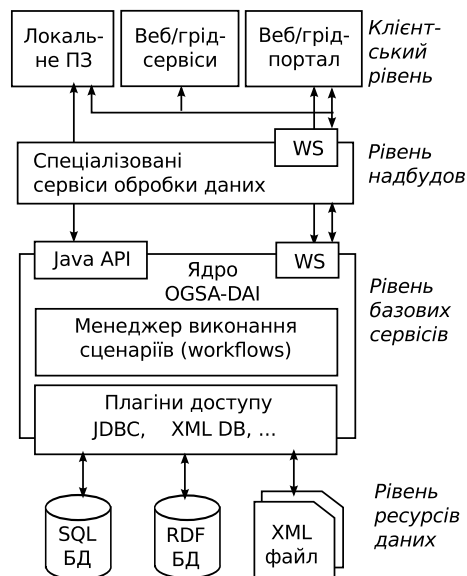


Рис. 1. Обробка даних через OGSA-DAI