

**Е-НАУКА НА ШЛЯХУ ДО СЕМАНТИЧНОГО ГРІД.  
ЧАСТИНА 1: ОБ'ЄДНАННЯ WEB- І ГРІД-ТЕХНОЛОГІЙ**

**М.З. ЗГУРОВСЬКИЙ, А.І. ПЕТРЕНКО**

Досліджено еволюцію Грід-систем на шляху до т.зв. семантичного Грід, в яких інформація і послуги мають чітке визначення, що надає можливість комп'ютерам і людям працювати в кооперації; а також представлено тенденцію до конвенції Грід-сервісів і Web-сервісів на базі використання сервісно-орієнтованої архітектури (SOA), за якої Web використовується як інформаційна інфраструктура Грід.

**1. Е-НАУКА І ГРІД**

Уже кілька років у світі активно вживається новий термін «**е-наука**» (електронна наука, e-Science), який визначає особливості сучасних наукових досліджень, а саме [1,2,3]:

- необхідність обробки потенційно величезних об'ємів інформації, де б вона не була розташована;
- необхідність виконання великого об'єму складних обчислень;
- необхідність спілкування і співпраці під час досліджень.

Ці виклики сьогодення з урахуванням обмеженості ресурсів, що існують в будь-якому суспільстві, можна задовольнити лише спільним і скоординованим їх використанням або «розділенням» ресурсів. Треба зробити доступними для вчених і фахівців різних організацій і країн існуючі комп'ютери, сховища даних, додатки, прилади, мережі з врахуванням різноманіття (гетерогенності) цих ресурсів. Треба розробити і запровадити механізми безпеки і узгодженої політики доступу, умови надання ресурсів, обліку обсягів використання і відповідної оплати. А також створити розподілене обчислювальне середовище, реалізоване за допомогою розгортання заснованої на стандартах інфраструктури, що надає спільне використання ресурсів і координоване вирішення завдань у межах динамічних віртуальних організацій, які є зацікавленими користувачами, крім звичайних організаційних структур.

Таке середовище отримало назву «*Грід-середовище*» (рис. 1) за аналогією з мережами електропостачання [1, 4, 5] Згідно з рис. 1, користувач формує своє завдання на обробку чи обчислення, брокер планує реалізацію

цього завдання і, спираючись на каталоги ресурсів та сервісів, добирає необхідні програми та дані і вільний робочий вузол, ще краще — завдання виконати.

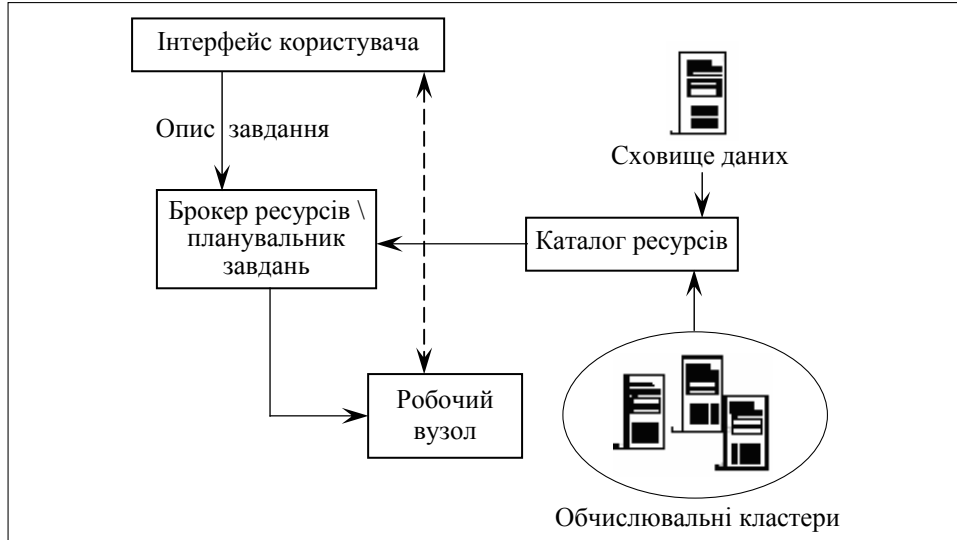


Рис. 1. Спрощена структура Грід-середовища

Грід-середовище забезпечує скоординоване вирішення завдань з інтеграції розподілених ресурсів, ідентифікації користувачів і автоматичного обслуговування їхніх запитів, пошуку для них ресурсів та загального децентралізованого «оркестрування» ресурсами і запитами. Базові функції Грід-середовища реалізуються за допомогою спеціального програмного забезпечення, що отримало назву «*програмне забезпечення проміжного шару*» (ПЗПШ) або middleware.

ПЗПШ використовується для того, щоб приховати різноманітну природу Грід та створити користувачам і додаткам однорідне середовище, забезпечуючи ряд стандартизованих інтерфейсів і безліч сервісів. ПЗПШ перебуває між операційною системою і додатками, забезпечуючи додатки сервісами, необхідними для їх коректного функціонування у розподілених гетерогенних середовищах. Грід-інфраструктура може складатися з будь-яких видів мережевих ресурсів, тобто, як з комп'ютерів (обчислювальних ресурсів), так і сховищ даних та спеціальних наукових приладів.

Результати аналізу і порівняння сучасного стану та тенденцій розвитку існуючих рішень, запроваджених у європейських та світових проектах Грід-середовищ, оцінка їх придатності для використання в Україні, формулювання пропозицій щодо різних компонентів Грід-середовища для національної Грід-інфраструктури представлені в [1, 4, 5, 6].

Спочатку Грід-технології призначалися для вирішення складних наукових, виробничих та інженерних завдань, які неможливо втілити в розумні терміни на окремих обчислювальних установках. Але зараз сфера застосування Грід-технологій не обмежується лише ними. Розвиваючись, Грід проникає в промисловість і бізнес, великі підприємства створюють Грід для вирішення своїх виробничих проблем. Таким чином, Грід претендує на роль універсальної інфраструктури для обробки даних, в якій функціонує безліч

сервісів (*Grid Services*), які сприяють розв'язанню не тільки конкретних прикладних завдань, але й пропонують сервісні послуги — пошук необхідних ресурсів, збір інформації про стан ресурсів, зберігання і доставку даних. Сфера застосування Грід — це ядерна фізика, екологія, метеорологія і моделювання кліматичних змін, чисельне моделювання в машино- і авіабудуванні, біологічне моделювання, фармацевтика.

## 2. ПОКОЛІННЯ ГРІД

Незважаючи на відносну «молодість» Грід (з перших публікацій минуло лише 14 років), можна виділити три стадії його еволюції.

*Грід-системи першого покоління* (I-WAY, FAFNER) дослідили підходи для забезпечення базової інфраструктури Грід на системному рівні. Так, згідно з проектом I-WAY (Information Wide-Area Year), у 1995 році було об'єднано в національну розподілену експериментальну мережу 17 обчислювальних вузлів, пов'язаних за допомогою високошвидкісної магістралі, на основі технології ATM (1,5 Мб/с – 9,6 Мб/с). Було вперше впроваджено розподілену файлову систему і брокер ресурсів, який містив центральний вузол, що керував мережею, та агентів на решті вузлів. Тоді було розроблено також 60 додатків і запроваджено в I-WAY.

Таким чином, уже перше покоління Грід-систем перетворило комп'ютер у пристрій, розрахований на багато користувачів, і створило систему розділення часу; запровадило організацію взаємодії на рівні процесів шляхом виконання загального адресного простору, розподіленої загальної пам'яті (DSM) і передачі повідомлень (MPI, PVM) для підтримки гетерогенних систем і роботи в локальній /глобальній мережі; забезпечило підтримку динамічної конфігурації середовища, при якому метакомп'ютер не має постійного складу і динамічно організовується з географічно розподілених ресурсів, тимчасово делегованих їхніми фактичними власниками (*віртуалізація метакомп'ютера*); а також забезпечило організацію однорідного доступу до обчислювальних ресурсів великої кількості комп'ютерів в локальній або глобальній мережі і вирішення задач, що допускають декомпозицію на велике число невеликих, незалежних підзадач.

*Грід-системи другого покоління* (Globus GT3, Legion, Nimrod, Jini, P2P, NorduGrid, Unicore, gLite) розвивали програмне забезпечення проміжного шару (ПЗПШ), щоб оперувати великомасштабними даними та обчисленнями, а також ключові сервіси, портали і специфічні додатки. Подальший розвиток і узагальнення ідей *метакомп'ютинга* в ширше коло обчислювальних ресурсів і завдань/додатків вимагав невідкладного розв'язання багатьох проблем, пов'язаних з передачею даних, забезпеченням безпеки, управлінням завданнями, доступом до даним, пошуком ресурсів, доступом до них та ін. Підходи до побудови такої Грід-системи були описані І. Фостером у [2]. Друге покоління базового програмного забезпечення для Грід у своєму розвитку перейшло від ранніх систем типу Globus-GT1 і Legion, спеціалізованих для конкретних потреб великих і високопродуктивних додатків, до більш універсальних і відкритих середовищ, таких, як Globus-GT3 і Avaki.

На відміну від традиційних клієнт-серверних систем, в яких ресурси управляються і облік користувачів ведеться централізовано в одному довірчому домені, таке ПЗПШ забезпечує динамічну безліч довірчих доменів, що використовують різні механізми безпеки; динамічну безліч ресурсів, керованих автономно в межах одного з доменів; динамічну безліч користувачів, які мають облікові записи в деяких доменах; доступ в процесі обчислень відразу до декількох ресурсів із різних доменів; запуск процесів на декількох ресурсах; організацію в процесі обчислень взаємодії між ресурсами з різних доменів; взаємодію між запущеними процесами.

Разом з базовим програмним забезпеченням у другому поколінні також був розроблений ряд супровідних інструментальних засобів і утиліт, які здійснюють запуск завдань на видалених ресурсах (суперкомп'ютерах, кластерах) через *вже існуючі* системи управління ресурсами (зазвичай, системи пакетної обробки), контроль стану виконання завдань і управління завданнями, збір і доступ до різноманітної інформації про систему і її компоненти тощо, і за допомогою яких можна надавати послуги (сервіси) користувачам, додаткам, планувальникам ресурсів, брокерам для сервісів верхнього рівня. Протягом цього періоду з'явилися: техніка однорангових систем, предметно-орієнтовані інтерфейси користувачів і портали.

Але друге покоління Грід не забезпечило повну функціональну сумісність створеного Грід програмного забезпечення, що є важливою умовою реалізації великомасштабних обчислень. Щоб побудувати нові Грід-додатки, бажано багаторазово використовувати існуючі компоненти та інформаційні ресурси і гнучко ними оперувати.

У Грід-системах третього покоління (OSGA, OSGI, Web Service, Semantic Web, Globus 4) зроблено наголос на сервісно-орієнтованому підході (відкриті сервісні системи, орієнтовані на архітектуру OGSA [7] і інфраструктуру OSGI [8], глобальне розподілене співробітництво та проблеми інформаційного рівня). Це допускає відому уніфікацію наборів сервісів на основі тотожності їх семантики, а також наявність загальних правил, регламентів і організаційних угод, на які спирається конфігурація сервісів. Таким чином, Грід-середовище — це *розширюваний набір уніфікованих Грід-сервісів* (аутентифікація, авторизація, пошук і розподіл ресурсів, повідомлення про події, облік використання ресурсів, видалений доступ до даних, виявлення відмов і т.ін.), об'єднаних різними способами для задоволення потреб користувачів і не залежних від ресурсів та додатків, при цьому існує можливість злиття базових сервісів в складніші, *високорівневі сервіси*.

Отже, йдеться наразі про інфраструктуру для е-науки, ніж про поліпшення уже розробленої існуючої технології. Грід може забезпечити підтримку віддаленого спілкування вчених у реальному часі. Дуже важливою є інфраструктура для підтримки розподілених ресурсів — тут є багато ключових сервісів: безпека, масштабування і управління, реєстрація і пошук та інтерфейси Web-сервісів, орієнтовані на повідомлення, з метою впровадження дієвих механізмів співробітництва. Всі головні Грід-сервіси і інфраструктура забезпечують співпрацю і є вкрай важливими для суспільства.

Третє покоління Грід-систем націлене, насамперед, на створення «загального дослідницького простору» (collaboratory), у якому національні дослідники можуть діяти безвідносно до географічного місця розташування шля-

хом спілкування з колегами, спільно використовуючи інструментарії, дані й обчислювальний ресурс і звертаючись по інформацію до цифрових бібліотек. Успішними прикладами таких зусиль можуть бути проекти *Grid 2003*, що призначені поєднати великі обчислювальні центри США; проект *EGEE* (Enabling Grids for E-science in Europe), метою якого є об'єднання національних, регіональних і тематичних Грід-систем в єдину цілісну Грід-інфраструктуру для підтримки наукових досліджень; проект *Access Grid*, присвячений колекції ресурсів, які підтримують співробітництво людей через Grid, включаючи великомасштабні розподілені зустрічі та навчання.

Грід-системи третього покоління впевнено просуваються до *семантичних Грід* (рис. 2) [9–11], заснованих на використанні метаданих і онтологій, у яких *інформація тлумачиться як дані, що мають не тільки значення, але і знання*, які здобуваються, використовуються, представляються, публікуються й підтримуються (Data Mining), щоб допомогти е-вченим досягати їхніх специфічних цілей — розділених обчислень, розділених даних, вилучення знань з інформації.

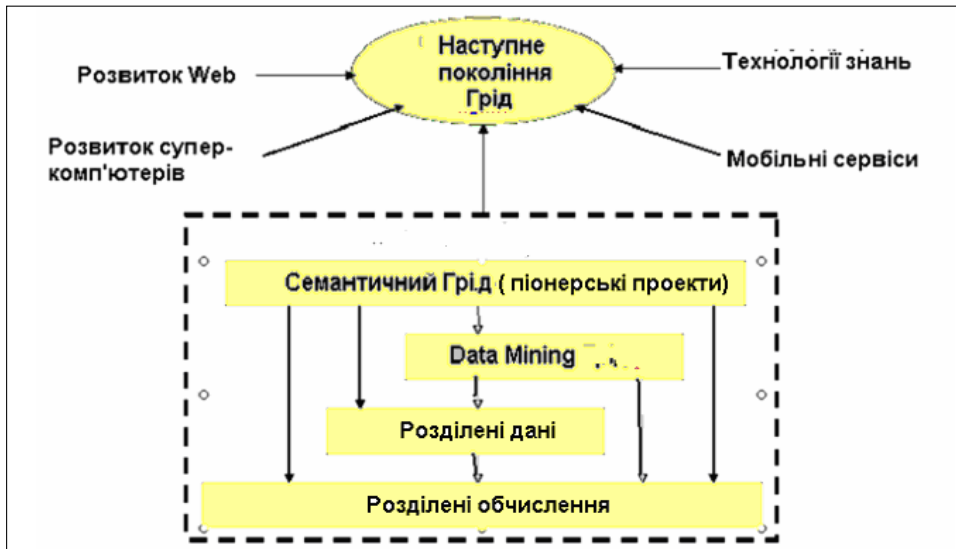


Рис. 2. На шляху до наступних поколінь Грід

Data Mining — це нова технологія інтелектуального аналізу даних з метою виявлення прихованих закономірностей у вигляді значущих особливостей, кореляцій, тенденцій і шаблонів [12]. Сучасні системи добування даних використовують засновані на методах штучного інтелекту засоби уявлення і інтерпретації, що і допомагає знаходити розчинену в терабайтних сховищах не очевидну, але вельми цінну інформацію. Фактично ми говоримо про те, що в процесі Data Mining-система не відштовхується від наперед висунутих гіпотез, а пропонує їх сама на підставі аналізу.

В основу сучасної технології Data Mining покладено концепцію шаблонів (pattern), які відображають фрагменти багатоаспектних взаємозв'язків в даних. Цими шаблонами є закономірності, притаманні підвибіркам даних, що можуть бути стисло виражені у формі, зрозумілій людині. Пошук шаблонів проводиться методами, не обмеженими лише апріорними припущеннями про структуру вибірки і вид розподілів значень аналізованих показників.

Невдовзі стане можливою наступна розмова користувача зі своїм програмним агентом, що пов'язує його з Грід-середовищем:

- знайди, будь ласка, дані про землетруси в Україні за минуле століття та дані про сонячну активність за цей період і оціни зв'язок між ними;
- проаналізуй показники світової торгівлі на початку кризи і спробуй визначити групу товарів, які на першому місці в товарообороті розвинутих країн;
- зайти на сайти вільного програмного забезпечення і завантаж себе програмою вирішення дуже великих за розміром систем лінійних розріджених рівнянь.

Важливо підкреслити: використання Грід-технологій допомагає не тільки вирішувати наукові й практичні завдання, раніше недоступні через занадто великий час, потрібний для одержання відповіді, а й створює основу для нової організації науки, високотехнологічного виробництва, соціального життя, допомагає ефективніше та надійніше управляти ресурсами суспільства. Грід *забезпечує платформу*, за допомогою якої користувачі отримують доступ до об'єднаних обчислювальних ресурсів зберігання даних і мережевих ресурсів, щоб виконувати свої прикладні програми щодо обробки розподілених даних. Це формує *функціональне середовище*, що надає користувачам можливість проаналізувати дані, спільно використовувати результати із співробітниками та інформацію про стан даних через встановлені кордони і географічні межі.

### 3. WEB ЯК ІНФОРМАЦІЙНА ОСНОВА ГРІД

Відомо, що *Інтернет* (мережа комп'ютерів, об'єднаних каналами, які використовують протоколи (TCP/IP) для зв'язку) і *Web* (мережа сайтів, що використовують гіперпосилання для переходів від сторінки до сторінки) виникли раніше *Грід* (обчислювальної інфраструктури, що надає безперервний доступ до обчислювальних потужностей і ресурсів зберігання даних, розподілених по всьому світу).

Певний час, поки Web-мережа була здатна обслуговувати лише людей (рис. 3а), Грід і Web розвивалися відокремлено. Коли ж Web-мережа почала виконувати передачу даних між комп'ютерами (рис. 3б), з'явилася зацікав-

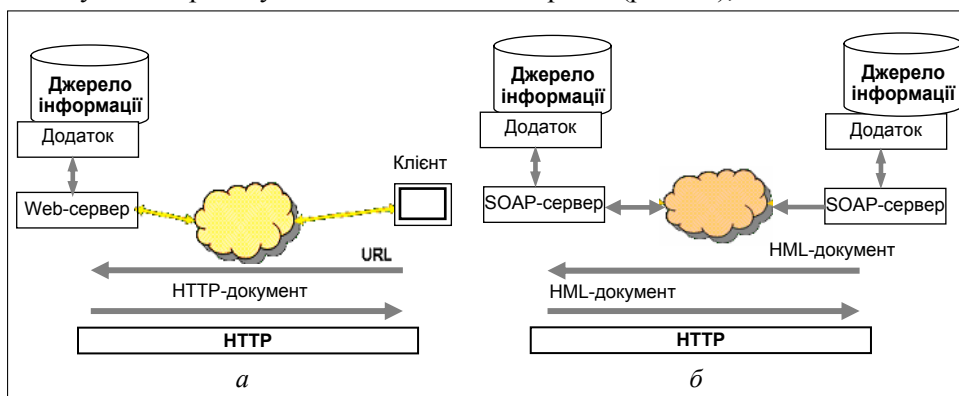


Рис. 3. Схеми отримання документа людиною (а) і обміну документами між комп'ютерами (б) в Web-мережі

леність у використанні здобутків Web-мережі в Гріл-середовищі. Поява Web-сервісів, які забезпечують віддалене виконання певних процедур за параметрами і даними, введеними користувачем, зробили процес об'єднання Web-сервісів і Грід-сервісів неминучим (рис. 4).

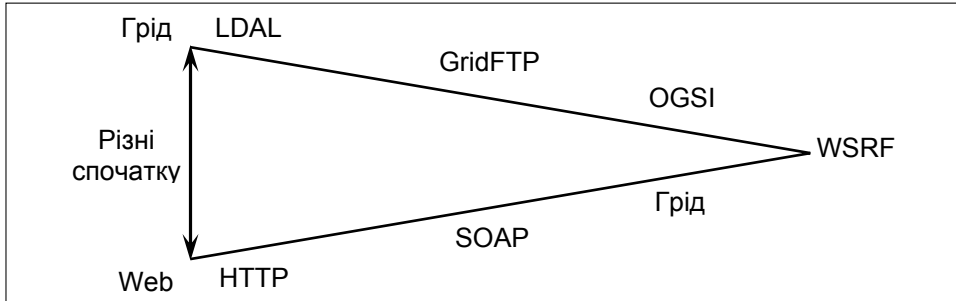


Рис. 4. Процес об'єднання Web- і Грід сервісів

Велике значення мало створення консорціумом W3C стандартів Web-сервісів, таких, як:

- ідентифікатори URI (Uniform Resource Identifier) — одноманітний ідентифікатор ресурсу [13]. Вся безліч URI є децентралізованою. Жодна людина і жодна організація не стежить за тим, хто їх створює або як їх можна використовувати. Тоді, як одні схеми URI (такі, як http:) залежать від централізованих систем (наприклад, систем DNS), інші схеми (такі, як freenet) є повністю децентралізованими;
- мова розмітки документів XML (eXtensible Markup Language) є текстовим форматом, призначеним для зберігання структурованих даних (замість існуючих файлів баз даних), для обміну інформацією між програмами, а також для створення на його основі більш спеціалізованих мов розмітки [14];
- протокол SOAP (Simple Object Access Protocol) [15]. SOAP пропонує пакет, що формує дані XML для передачі через Web-інфраструктуру (наприклад, через HTTP, через кеш та проксі-сервери), з викликом віддалених процедур (Remote Procedure Calls, RPCs);
- схема RDF (Resource Description Framework) [16]. Це модель і синтаксис для опису ресурсів, а саме — метаданих про ресурси. Однією з головних цілей RDF є надання тверджень, зрозумілих і для людей, і для машин. Кожне твердження має вигляд «суб'єкт – предикат – об'єкт» (рис. 5) і в термінології RDF називається *триплетом*. Для обробки RDF є декілька мов запитів. Наприклад, RQL, RDQL, SPARQL;

Суб'єкт:	http://www.w3.org/Home/Foster
Предикат:	http://www.schema.org/#Creator
Об'єкт:	http://www.example.org/ Grid

Рис. 5. Приклад RDF-формату для твердження: «Foster is a creator of Grid»

- мова опису Web-сервісів WSDL (Web Services Description Language) [17]. Описує сервіс в XML, використовуючи XML Schema; є також відображення на RDF (*Resource Description Framework* — структура опису ресурсу);

- сервер Web-сервісів UDDI (Universal Description Discovery and Integration). UDDI, який є переліком доступних Web-сервісів [15];

- мова WSFL (Web Services Flow Language) є пропозицією IBM, що визначає робочий процес як комбінацію Web-сервісів [18];

- WSMF (Web Services Modelling Framework) забезпечує концептуальну модель для розробки та опису Web-сервісів, засновану на принципах максимального роз'єднання та масштабованому сервісі посередництва {18}.

Грід-сервіси, на відміну від Web-сервісів, постійні і з статичним постачальником метаданих з'являються і зникають, зберігають інформацію про свій стан і стан ресурсу та походження інформації, у них динамічні користувачі метаданих. Платформа WSRF (Web Service Resource Framework) [19] — це набір з 6-ти специфікацій, які підтримують Грід-сервіси та інші ресурси, що мають свій стан (stateful). Мета створення таких специфікацій — наближення OGSA і Web-сервісів. За допомогою засобів, що відповідають цим специфікаціям, може стати реальним підхід до моделювання та управління станом у контексті Web-сервісів. Тобто, зроблена спроба реалізувати цей стан — саме те, що відрізняє Грід-сервіси від Web-сервісів, дозволяючи користувачам, які звертаються до Web-сервісів, контролювати і змінювати стан доступних їм ресурсів (WS-Resource).

Назвемо ці специфікації:

- WS-Resource Lifetime визначає механізм припинення існування WS-Resource (включаючи обмін повідомленнями), що дозволяє негайно видалити ресурс або через вказаний час;

- WS-Resource Properties визначає, як WS-Resource може бути асоційований з інтерфейсом, що описує Web-сервіс (включаючи обмін повідомленнями) і витягає, змінює та знищує особливості ресурсу WS-Resource;

- WS-Notification визначає механізм обробки повідомлень про події, що базується на принципах передплати / публікації;

- WS-Renewable References визначає механізм розширення звичайної системи адресації WS-Addressing, прийнятої в Web-сервісі;

- WS-Service Group визначає інтерфейс до набору гетерогенних Web-сервісів;

- WS-Base Faults визначає механізм обробки повідомлень про помилки.

Ці специфікації WSRF є «точками конвергенції» між Web- і Грід-сервісами, їх можна розглядати як природну еволюцію специфікацій OGSI (Open Grid Services Infrastructure). Підхід WSRF ґрунтується на схемі адресації внутрішніх ресурсів, що асоціюються з Web-сервісами. Ця адресна інформація може використовуватися як засіб кореляції і маршрутизації обмінів повідомленнями з цими внутрішніми ресурсами і, нарешті, як засіб забезпечення зв'язку із станом, що зберігається.

Як мережевий показник конструкція *WS-Addressing* [20], що містить спеціальну інформацію про ресурс, слугує тій же меті, що й конструкції CORBA IOR, DCOM OBJREF, Java RMI URL і т. ін. Вона ідентифікує ресурс у всій мережі. Дотримуючись методології існуючих об'єктно-орієнтованих технологій розподілених обчислень, підхід WSRF переносить проблему ідентифікації ресурсу з рівня додатку на рівень проміжного програмного забезпечення і вирішує її в межах стека Web-сервісів. Вимагаючи,



щоб ідентичність ресурсу передавалася як заголовок кожного SOAP-повідомлення, підхід WSRF моделює взаємодії, що зберігають стан за допомогою конкретних ресурсів, а не сервісів. Якщо врахувати наявність додаткових специфікацій, які керують терміном життя експонованих ресурсів і механізмом оновлення посилань на ці ресурси, то можна констатувати, що підхід WSRF має багато спільного з концепціями моделей розподілених об'єктів.

Згідно з концептуальною моделлю підходу WSRF, Web-сервіс — це сутність, яка не має стану, «яка діє, надає доступ або обробляє набір логічних ресурсів (документів), що володіють станом на підставі повідомлень, котрі відправляються або отримуються нею». Така модель допускає явне експонування властивостей (логічних або фізичних) ресурсів через межі сервісу. Представлення стану цих експонованих ресурсів і апарат, за допомогою якого споживачі можуть безпосередньо взаємодіяти з ними, і є головною метою підходу WSRF.

Використовуючи підхід WSRF, схема-доступу, наприклад, до сховища файлів, достатньо прямолінійна. Для конкретного ресурсу отримано *WS-Addressing*-посилання на крайню точку (що здійснене через деякий допоміжний механізм, подібний до реєстрації або фабрики). Це посилання на крайню точку (розширене специфічними WSRF-метаданими) вважають за мережевий показник для ресурсу, веденого Web-сервісом. Отримане посилання на крайню точку використовується і як адреса, на яку можна відправляти повідомлення, і як неявний контекст взаємодії з внутрішнім ресурсом (для цього радимо скористатися вмістом його елементу «ReferenceProperties»).

У випадку із сховищем файлів, ID файла (ідентифікатор вузла або деякий інший дескриптор) може надавати метадані, необхідні сервісу, для маршрутизації звернення до того ж самого ресурсу при кожному повідомленні, переданому сервісу. Щоразу, коли користувач (рис. 6) посилає повідомлення сервісу, його крайня точка ідентифікується *WS-Addressing*-елементом, а метадані, пов'язані з ресурсом, використовуються з метою допомоги сервісу знайти маршрут до правильного внутрішнього ресурсу.

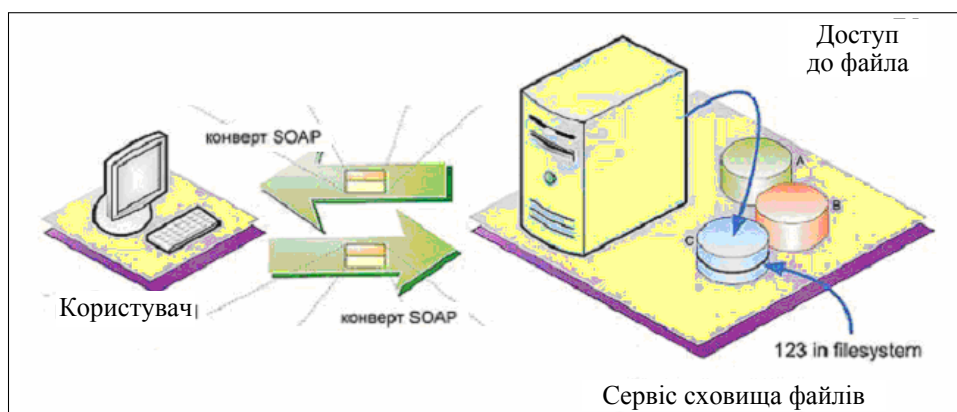


Рис. 6. WSRF-схема доступу до сховища файлів

Основний недолік такого підходу в тому, що за необхідності все ж таки розвивати можливості сервісу (наприклад, коли реалізація сховища файлів

системи мігрує від файлової системи до системи управління базою даних), ідентичність інформації, яка міститься у метаданих посилання на крайню точку, може виявитися застарілою. Як наслідок цього, сесія із збереження стану не досягне успіху. От чому додаткові механізми, такі, як управління терміном життя і відновлення посилань, є необхідними частинами платформи WSRF. Хоча і такі проблеми можна виключити, використовуючи логічні ідентифікатори (які вирішуються сервісом у фізичні ресурси), але це не прописано в специфікаціях WSRF.

Web-сервіси близькі до вимог систем Грід третього покоління. Це підтримка сервісно-орієнтованого підходу і стандартів, що полегшують інформаційні процеси, такі, як опис сервісу (рис. 7).

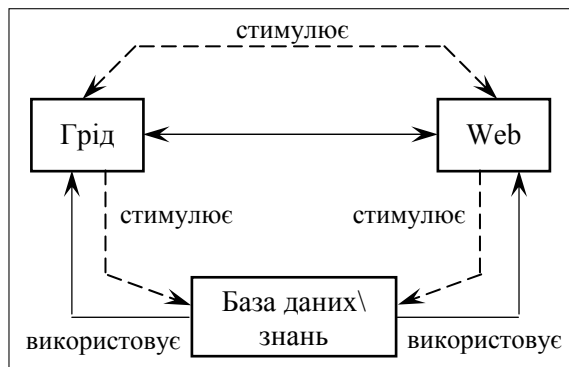


Рис. 7. Взаємодії між Грід і Web

Архітектура OGSA обумовлює створення, обслуговування та додавання ансамблів сервісів, які використовуються у віртуальних організаціях. Тут сервіс визначений як об'єкт із підтримкою мережевої роботи, яка забезпечує такі можливості, як обчислювальні ресурси, ресурси пам'яті, мережі, програми та бази даних. Це пристосовує створення Web-

сервісів до виконання деяких вимог, специфічних для Грід. Нижче наведені стандартні інтерфейси, наявні в OGSA:

- *виявлення (discovery)*: користувачі потребують механізмів, щоб знайти доступні сервіси, визначити їхні особливості та мати можливість конфігурувати сервіси відповідно до запитів, поставлених до них;
- *динамічне створення сервісу (Dynamic service creation)*: стандартний інтерфейс і семантика, які повинні забезпечити створення будь-якого сервісу;
- *керування життєвим циклом (Lifetime management)*: у системі мають бути надані механізми для відновлення сервісів і їхнього стану, пов'язаного з невдалими операціями;
- *повідомлення (Notification)*: безліч динамічних, розподілених сервісів, здатних виявляти один одного та одержувати інформацію про зміну їхнього стану;
- *керуваність (Manageability)*: надані операції, що належать до керування та контролю великої кількості екземплярів класів Грід-сервісів;
- *просте виконавче середовище (Simple hosting environment)*: ряд ресурсів, розташованих у межах окремого адміністративного домена, що підтримує різні засоби керування сервісом: наприклад, сервер додатків J2EE, система Microsoft NET або кластер Linux.

Таким чином, до найбільш вагомих здобутків Грід-систем третього покоління, отриманих дотепер, можна віднести злиття Грід-технологій і технологій Web-сервісів, формування Грід-сервісу як спеціального розширення Web-сервісу шляхом підтримки екземплярів Грід-сервісу, що мають стан і,

можливо, обмежений час життя. З кожним екземпляром Грід-сервісу пов'язані дані сервісу, тобто, інформація, структурована у набір іменованих XML-елементів, що типізуються (service data elements, SDE), а також впровадження платформи WSRF, яка спирається на архітектуру OGSA і на загальноновизнані стандарти Web-сервісів, при цьому зберігається багато елементів OGSI, але використовується інша термінологія і розширюються можливості OGSI.

#### 4. ВРАХУВАННЯ СЕМАНТИКИ ПРИ ОБРОБЦІ ІНФОРМАЦІЇ

Сьогодні інформація Web-мережі пристосована здебільшого для людського сприйняття. Користувач може переходити з одного посилання на інше, давати запити різним пошуковим системам або ж відшукувати сайти, вводячи їхні адреси. І хоча Web-сторінки досить привабливі для людини, але для комп'ютерної програми, яка оброблює їх вміст, вони — лише рядки з випадковими символами.

Комп'ютерна програма не здатна, завантаживши довільний документ, чи то Web-сторінку або інший файл, зрозуміти його зміст. Вона, звичайно, може зробити якісь припущення, ґрунтуючись на HTML- або XML-тегах, але все одно потрібна людина-програміст, яка повинна розібратися в них і зрозуміти сенс або семантику кожного з тегів. З погляду комп'ютера, існуюча Web-мережа — це суцільна плутанина. На щастя, вихід є: це Семантична Web [21–23], яка повинна стати певним доповненням Web-мережі, що складається із зрозумілої комп'ютерам інформації. Реалізація цієї нової мережі стає можливою завдяки ряду нових стандартів, що розробляються WWW-консорціумом (W3C) [22]. Коли семантична Web набере обертів, тоді значна кількість інформаційних ресурсів будуть придатні для використання як людиною, так і програмними агентами. Отже, програмні агенти нарешті навчатися читати Інтернет.

Подібно до того, як семантична Web є розширенням звичайної Web-мережі, семантичні Web-сервіси (SW) розширюють поняття звичайних Web-сервісів. Ідея семантичної Web — це концепція мережі, в якій кожен ресурс, відтворений людською мовою, забезпечений описом, зрозумілим комп'ютеру. Документи в мережі публікуються в форматі XML, який містить семантичні RDF твердження. Відношення, аксіоми, твердження про об'єкти зберігаються в *онтологіях*, що реалізуються мовою *OWL*. Добути дані з RDF можливо за допомогою *SPARQL*

Сьогодні створені програми, здатні відшукати потрібні нам сервіси в UDDI-сервері, й хоча програма може знайти якийсь Web-сервіс без допомоги людини, вона не в змозі зрозуміти, як саме ним користуватися і для чого він призначений. Мова опису Web-сервісів (WSDL) надає інструмент для опису, яким чином взаємодіяти [communicate] з тим чи іншим Web-сервісом, тоді як *семантична розмітка* змісту забезпечує нас інформацією про те, що і як робить даний сервіс.

Вважають, що повний потенціал Грід-обчислень можна досягнути тільки завдяки повній експлуатації функціональних можливостей та можливостей, що надаються рівнем знання. Тому цей рівень необхідний

для автоматизованого безпосереднього доступу до операцій і взаємодій.

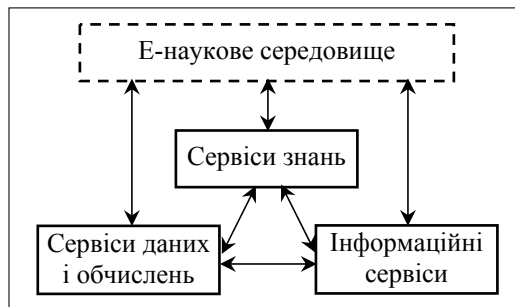


Рис. 8. Триврівнева архітектура Грід-сервісів

Семантичний Грід стане місцем, де дані будуть обладнані багатим контекстом і перетворені на інформацію. Тоді ця інформація буде розділена й оброблена віртуальними організаціями для досягнення певних цілей. Така інформація містить знання (рис. 8). Отже, рівень знання — ключ до наступної стадії в еволюції Грід, до повністю оснащеного семантичного Грід [12].

Семантичний Грід, який базується на семантичній Web (рис. 9), можна вважати спрямованою на сервіси архітектурою, в межах якої об'єкти в певному середовищі обмінюються послугами (сервісами) один з одним на певних умовах.

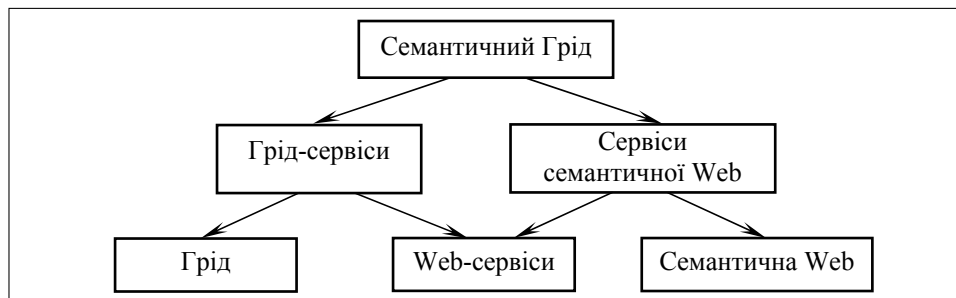


Рис. 9. Взаєморозвиток Web і Грід

Web сьогодні стає інфраструктурою для розподілених застосувань, де наразі відбувається обмін інформацією між програмами, ніж прочитання її людиною. У ідеальному варіанті вся інформація в Інтернеті повинна розміщуватися двома мовами: людською мовою для суспільства і комп'ютерною мовою для розуміння комп'ютером. Основний акцент при цьому зроблено на роботі з *метаданими*, які однозначно характеризують особливості та зміст ресурсів і здійснюють текстовий аналіз документів. Якщо самі ресурси призначені для сприйняття їх людиною, то метадані використовують пошукові роботи та інші інтелектуальні агенти для проведення однозначних логічних висновків про особливості цих ресурсів. Додатки семантичних Grid можуть інтегрувати безліч різномірних джерел інформації і послуг, які залучені і об'єднані, а також людські і обчислювальні послуги та послуги передачі інформації.

## ВИСНОВКИ

Прогресивне поєднання Web- і Grid-технологій дає підстави спростити механізм пошуку, відбору та обробки інформації, зважаючи на нещодавно розпочаті у Web-середовищі проекти побудови семантичних енциклопедій, баз знань, лексичних баз розмовних мов та ін.

Семантичний Грід сьогодні використовує засоби, за допомогою яких знання можуть бути вилучені з даних, а також використані, відновлені, опубліковані і оброблені для того, щоб допомогти е-науковцям досягти своїх конкретних в кожному випадку цілей. Цей підхід дозволяє набуті повної реалізації переваг Грід-технологій з вищим ступенем зручної в роботі і цілісної автоматизації, що забезпечує гнучку співпрацю вчених і обробки інформації в глобальному масштабі.

Докладніше семантична Web і семантичний Грід та їх взаємодії будуть розглянуті у наступній роботі авторів.

## ЛІТЕРАТУРА

1. Згуровський М.З., Петренко А.І. Grid-технології для е-науки і освіти // Наукові вісті НТУУ «КПІ». — 2009. — № 2. — С. 10–17.
2. Foster I., Kesselman C. (Eds.). The Grid: Blueprint for a New Computing Infrastructure. — Morgan — Kaufman, 1998. — 701 p.
3. <http://www.research-councils.ac.uk/escience/documents/gridteam.htm> — UK e-Science Network.
4. Zagorodny A., Zgurovsky M., Zinovjev G., Petrenko A., Martynov E. Integrating Ukraine into European Grid Infrastructure // Системні дослідження і інформаційні технології. — 2009. — № 2. — С. 35–49.
5. Петренко А.І. Вступ до Грід-технологій для науки і освіти: Учб. посіб. — Київ: Політехніка, 2008. — 124 с.
6. Петренко А.І. Застосування Грід-технологій в науці і освіті. — Київ: Політехніка, 2009. — 145 с.
7. Талиа Д. OGSA: где Grid встречается с Web // Открытые системы. — 2003. — № 1. — С. 47–50.
8. <https://forge.gridforum.org/projects/ogsi-wg> — Tuecke, S., et al. (2003). «Open Grid Services Infrastructure (OGSI) — Version 1.0».
9. De Roure D. Jennings N.R., Shadbolt N.R. The Semantic Grid: Past, Present, and Future / Proceedings of the IEEE. — 2006. — 93, № 3. — P. 669–681.
10. Zhaohui Wu., HuaJun Ch. Semantic Grid: Model, Methodology, and Applications. — Springer. — 2008. — 327 p.
11. Петренко А.І. Семантичний Грід для гнучкого оброблення даних: Матер. конф. «Системний аналіз та інформаційні технології». — 2009. — 26–30 травня. — С. 16–17.
12. Петренко А.І. Grid і інтелектуальна обробка даних // Системні дослідження та інформаційні технології. — 2008. — № 4. — С. 97–110.
13. <http://www.w3.org/Addressing/> — Web Naming and Addressing Overview (URIs, URLs).
14. <http://www.w3.org/XML/> — Extensible Markup Language (XML).
15. Ньюкоммер Э. Веб-сервисы. XML, WSDL, SOAP и UDDI. — СПб.: Питер, 2003. — 256 с.
16. <http://www.w3.org/RDF/> — Resource Description Framework (RDF).
17. Christensen E., Cubera F., Meredith G., Weerawarana S. Web Services Description Language (WSDL) 1.1, Technical report, WWW Consortium. — 2001. — 114 с.
18. <http://www.osp.ru/os/2004/12/184882/> — Web-сервисы, Uhl--сервисы и sysis.
19. <http://www.globus.org/wsrfl/> — сайт Web Service Resource Framework (WSRF).
20. <http://msdn.microsoft.com/ws/2003/03/ws-addressi> — Web Services Addressing (WS-Addressing).

Надійшла 10.09.2009