УДК 004.75

А.О. Сальніков, Ю.В. Бойко

Київський національний університет імені Тараса Шевченка, Київ

МЕТОДИКИ ТА ЗАСОБИ РЕПЛІКАЦІЇ ГРІД-СЛУЖБИ ЗАСВІДЧЕННЯ ТА КЕРУВАННЯ УЧАСТЮ В ВІРТУАЛЬНИХ ОРГАНІЗАЦІЯХ

Проведене дослідження грід-служби засвідчення та керування участю в віртуальних організаціях (ВО) показало, що серверна частина не надає гнучких механізмів забезпечення високої доступності. Сформовано методики реплікації бази даних ВО на рівні застосувань (в режимі мульти-майстер). Методики втілено у розробленому програмному забезпеченні – PHP VOMS – Admin. Примірники розробленої служби з підтримкою реплікації впроваджено в роботу віртуальних організацій MolDynGrid, NetworkDynamics та testbed.univ.kiev.ua в Українському національному грід-сегменті (УНГ).

Ключові слова: *грід, віртуальна організація, служба керування участю в віртуальних організаціях.*

Вступ

Служба засвідчення та керування участю в віртуальних організаціях (Virtual Organization Mambership Service, VOMS) є центральним елементом, що забезпечує інтеграцію віртуальних організацій (ВО) в грід–інфраструктури. Від доступності служби VOMS залежить робота всіх ВО, що обслуговуються її примірником.

Переважна більшість віртуальних організацій в Українському національному грід–сегменті (УНГ) обслуговується єдиним національним сервером VOMS Київського національного університету імені Тараса Шевченка [1, 2]. За централізованої реалізації служби VOMS робота всіх віртуальних організацій (наразі 15 ВО обслуговуються сервером) залежить від доступності одного серверу в грід–інфраструктурі.

Розроблення методик та засобів реалізації розподіленої служби засвідчення та керування участю в ВО для підвищення показників її доступності в грід–інфраструктурах є предметом дослідження даної роботи.

Постановка задачі

Для забезпечення високої доступності VOMS, в грід–інфраструктуру необхідно впровадити декілька примірників цієї служби, що надають доступ до тотожної інформації та функцій. Архітектура VOMS представляє собою систему керування базами даних учасників ВО, службу засвідчення участі в ВО та службу керування участю в ВО, що реалізує інтерфейс користувача [3].

Для роботи декількох примірників служби необхідно забезпечити роботу програмних засобів грід з декількома примірниками служби (як для роботи з засвідченням повноважень, так і для взаємодії з службою керування) та забезпечити синхронізацію вмісту бази даних учасників ВО поміж примірників служби.

Стандартні поставки програмного забезпечення проміжного рівня грід надають засоби роботи з декількома примірниками VOMS. Основним елементом методики взаємодії з декількома примірниками служби засвідчення участі в ВО є обслуговування списків VOMS, поміж яких для виконання операції обирається доступний примірник.

Єдиним методом синхронізації вмісту баз даних класичної служби засвідчення та керування участю в ВО – EDG VOMS – Admin є реплікація засобами системи керування базами даних (СКБД) в режимі ведучий – ведений (master–slave) [4]. Такий метод має ряд недоліків:

- необхідність використання однакових СКБД на всіх примірниках;
- режим реплікації master–slave при виході з ладу ведучого примірника дозволяє проводити тільки операції читання, що значно обмежує можливості роботи з VOMS;
- налаштування та контроль реплікації відбувається на рівні всього серверу, а не на рівні кожної ВО;
- відповідальність за забезпечення реплікації лягає на адміністраторів служб VOMS а не на адміністратора ВО.

Забезпечення реплікації вмісту баз даних VOMS без зазначених недоліків потребує винесення функцій взаємодії примірників на рівень застосувань, засобами самої служби. Такий підхід дозволить незалежно від СКБД впровадити функції мульти–майстер реплікації на рівні ВО без зміни конфігурації серверу.

Розробка методик забезпечення мульти-майстер реплікації VOMS на рівні ВО та їх реалізація в службі керування участю в віртуальних організаціях є **метою даної роботи**.

Реплікація VOMS на рівні застосувань

Запропонована модель реплікації на рівні застосувань використовує журнал транзакцій для операцій, що змінюють вміст бази даних (рис. 1). Журнал транзакцій містить імена та параметри функцій,

що були викликані для виконання операцій. Для збільшення швидкості обробки запитів операції читання не записуються до журналу і їх атомарність не забезпечується.

Для синхронізації транзакцій використовується модель з активним опитуванням повідомлень (pull–model). Така модель передбачає ініціацію обміну примірником–клієнтом, який запитує інформацію про нові транзакції серверу. Використання активної моделі дозволяє мінімізувати затримки виконання операцій на сервері (немає необхідності очікувати надсилання транзакції на всі репліки), і у випадку виходу з ладу примірника отримати транзакції за період неактивності при відновленні роботи.

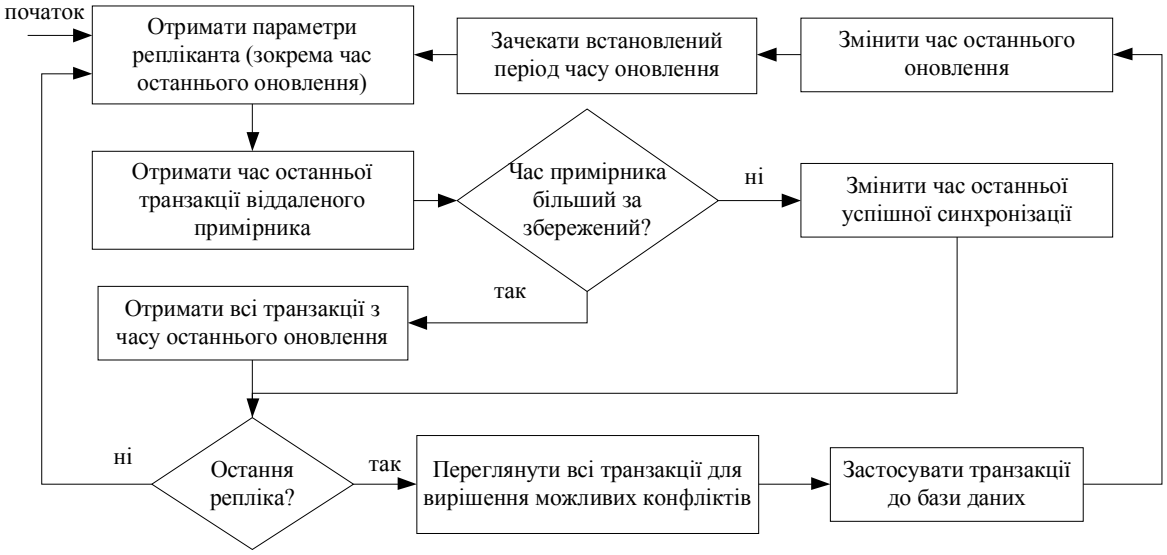

Рис. 1. Блок–схема станів реплікації VOMS

Операції читання з бази даних (засвідчення, отримання списків учасників, тощо) переважають над операціями запису при роботі зі службою VOMS. Для ВО, що не використовують збереження динамічної інформації, період зміни інформації в базі даних зазвичай є тривалим. Період синхронізації бази даних визначається в межах періоду оновлення інформаційної системи (3 – 5 хвилин) та періоду оновлення провайдерів ресурсів (декілька годин).

Для моделі мульти–майстер кожна служба є клієнтом до інших, і інформація про нові транзакції розповсюджується при їх виникненні на будь–якому примірнику. Але не виключена ситуація, коли інформацію про транзакцію буде надіслано до екземпляра служби, який є її джерелом. Для уникнення дублювання операцій, пропонується введення унікальних ідентифікаторів (Universally Unique IDentifier, UUID), що характеризують кожну транзакцію [5].

Для всіх реплікантів в топології, клієнт, який виконує pull–запит повинен отримувати лише ті транзакції, що відбулися на сервері з часу останнього оновлення. Для цього необхідно фіксувати час останньої транзакції на кожному сервері. Для уникнення проблеми різного значення часу на примірниках операції виконуються в одиницях часу кожного віддаленого серверу. Схему роботи реплікації VOMS на рівні застосувань наведено на рис. 1.

Авторизація реплікантів на рівні ВО

Передача даних між примірниками служби керування участю в ВО відбувається по незахищеним каналам зв'язку. Тому необхідно забезпечити цілісність зв'язку та автентифікацію реплікантів. Використання моделі з активним опитуванням повідомлень виключає ситуації за яких примірнику надсилається невірна інформацію від третьої сторони. В розглянутій моделі реплікант, виступаючи ініціатором з'єднання, повинен переконатись в автентичності серверу до якого надсилається запит на отримання інформації. З іншого боку сервер повинен повертати інформацію тільки авторизованим клієнтам.

Забезпечення цілісності та автентифікації учасників передачі даних між грід–сервісами Grid Security Infrastructure (GSI) зазвичай використовує взаємну автентифікацію (Mutual authentication). Для взаємної автентифікації GSI використовує Transport Layer Security (TLS) [6,7], що передбачає обмін сертифікатами. Проте такий алгоритм роботи дозволяє організувати канал зв'язку тільки на рівні всього сервісу і не дозволяє реалізувати керування реплікацією на рівні ВО.

Розміщувати для кожної ВО, яку обслуговує сервер VOMS, пару ключів для взаємної автентифікації суперечить вимогам безпеки, адже приватний ключ не повинен передаватись третій стороні. Особливо у випадку національних служб керування участю в ВО, це б призвело до передачі всіх приватних ключів до центральної служби.

Враховуючи модель передачі даних та вимоги до реплікації на рівні ВО, сформовано такий підхід:

- цілісність та автентичність отримуваних даних забезпечується за допомогою Transport Layer Security – використання цифрового підпису сертифікатом сервера;

- автентифікація клієнта відбувається за допомогою попередньо погодженого спільного ключа, що прив'язаний до IP-адреси клієнта.

Такий підхід дозволяє для кожної ВО окремо визначити спільний ключ, що відповідає репліканту з зазначеною IP-адресою. Ключ контролюється адміністратором ВО та зберігається в базі даних серверу. Робота з базою даних служби дозволяє винести процес узгодження параметрів реплікації на рівень застосувань та відокремити різні віртуальні організації. Цілісність даних для всіх ВО забезпечується єдиним механізмом використання цифрового підпису сервера VOMS в цілому.

Вирішення проблеми автоінкременту мульти-майстер реплікації

Сумісність із службою засвідчення участі в ВО та іншими реалізаціями служби реплікації накладає принципове обмеження – незмінність схеми бази даних ВО. Схема, що використовується, проектувалась без урахуванням мульти–майстер реплікації і містить первинні ключі, доменом яких є множина натуральних чисел. Операції додавання інформації створюють новий запис в відношенні, значення первинного ключа якого визначається як автоінкремент попереднього.

Таблиця 1

Робота з автоінкрементованими ідентифікаторами за умов реплікації

Етап роботи	Служба 1	Служба 2
1. Створення записів	А (10)	Б (10)
2. Синхронізація транзакцій	А (10) Б (11)	А (11) Б (10)
3. Видалення запису	Б (11)	А (11)

Розглянемо приклад, що ілюструє проблему використання автоінкременту первинних ключів (див. табл. 1). Нехай в період між синхронізацією транзакцій на різних примірниках служби було виконано операції додавання учасників ВО. Наступне значення первинного ключа для визначеності вважатимемо рівним 10.

Перший етап – учасника "А" було додано до першого примірника служби, а учасника "Б" до другого. На другому етапі примірники обмінялись новими транзакціями: до першого примірника додано учасника "Б" (значення ключа – 11), до другого примірника додано учасника "А" (значення ключа – 11). В результаті такої операції буде втрачено цілісність даних. Наприклад, при видаленні учасника "А" з першого примірника за його первинним ключем (етап 3), на другому примірнику буде видалено учасника "Б".

Проблема синхронізації первинних ключів вирішується за допомогою використання UUID замість натуральних чисел, але обмеження на незмінність бази даних не дозволяє внести такі зміни.

Для вирішення проблеми пропонується наступний підхід:

- до бази даних додається окреме відношення, що зберігає відповідності між значеннями первинних ключів (ID) та UUID записів;
- всі функції служби керування з реплікацією в першу чергу використовують UUID, проте за його відсутності працюють зі значеннями ID;
- функції служби керування з реплікацією, що додають інформацію до бази даних, зберігають автоінкрементоване значення первинного ключа разом зі згенерованим UUID в відношення відповідності;
- згенерований UUID додається до журналу транзакцій разом з параметрами функції;
- інший примірник служби отримує транзакцію разом з UUID та викликає функцію;
- функція іншого примірника служби зберігає визначене автоінкрементоване значення, та його відповідність до отриманого UUID, що співпадає з першим примірником.

До функцій, що додає інформацію до бази даних, в загальному випадку передаються локальні ID для проведення операцій (наприклад, ідентифікатор групи до якої додається учасник), значення атрибутів та необов'язковий масив значень UUID. Алгоритм роботи функції наступний:

1) визначення ефективних ID з урахуванням масиву UUID та відношення відповідності;

2) операції вибірки даних в стандартній структурі бази даних з використанням ефективних ID;

3) операція додавання запису (визначається автоінкрементований ID);

4) генерація UUID для автоінкрементованого ID;

5) збереження UUID до масиву, який в свою чергу зберігається у тілі транзакції.

Реалізація методики реплікації в PHP VOMS–Admin

Для вирішення проблем масштабованості служби засвідчення та керування участю в ВО було розроблено та впроваджено в Український національний грід програмне забезпечення PHP VOMS–Admin [8]. Саме цю реалізацію інтерфейсу керування використовує сервер Київського національного університету імені Тараса Шевченка для обслуговування українських ВО.

Для реалізації методик реплікації до PHP VOMS–Admin було додано інтерфейс віддаленого виклику процедур, що слугує для обміну транзакціями між примірниками. Інтерфейс підтримує такі операції:

- ping – перевірка підтримки інтерфейсу реплікації примірником;
- status – повертає стан узгодження реплікації на базі даних аутентифікації;
- ltt – повертає час останньої транзакції (last transaction time);
- tdiff – повертає всі транзакції з моменту часу, що запитується;
- alldata – повертає повний вміст бази даних для ініціалізації клієнта.

Щоб розпочати реплікацію необхідно створити узгодження між двома примірниками PHP VOMS–Admin. На цьому етапі вказуються унікальні імена сертифікату сервера та центру сертифікації репліканта, точка входу та спільні ключі для авторизації репліканта. Узгодження реплікації може перебувати в одному з наступних станів:

- НЕ ПІДТВЕРДЖЕНО – вихідний стан нового узгодження; запис в базі даних було створено, проте синхронізація бази даних ще не проводилась;
- ІНІЦІАЛІЗОВАНО – проведено синхронізацію бази даних, але обмін транзакціями ще не відбувся;
- СИНХРОНІЗОВАНО – успішний обмін транзакціями; вказано час останньої синхронізації;
- НЕ СИНХРОНІЗОВАНО – протягом трьох періодів синхронізації обмін транзакціями не відбувся; вказано час останньої успішної операції.

Приклад відображення стану узгодження в інтерфейсі керування наведено на рис. 2.

Для перегляду та зміни параметрів узгодження реплікації та/або виконання синхронізацію вмісту бази даних примірників (один виступає в ролі ведучого і його база даних перезаписується поверх бази примірника) слугує детальний перегляд стану (рис. 3).

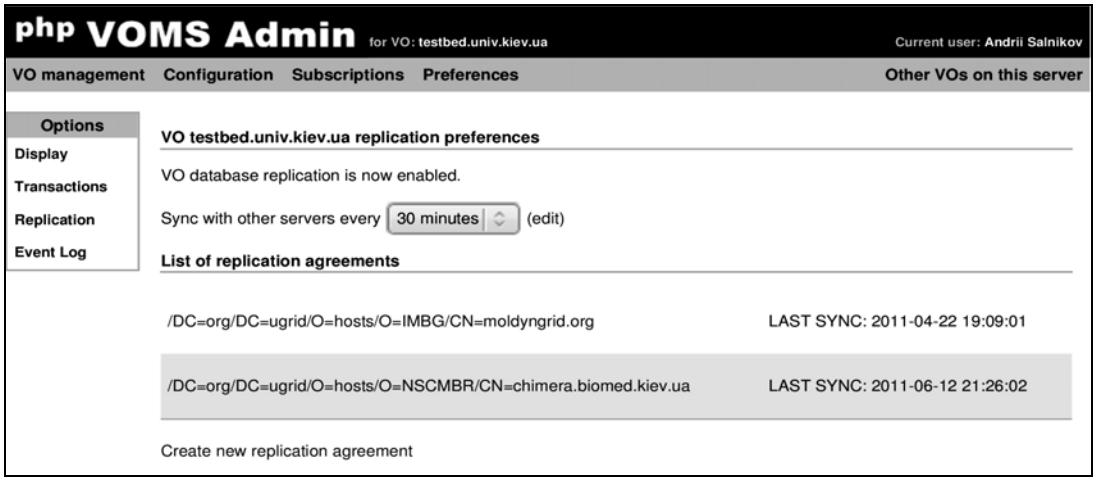

Рис. 2. Стан узгодження реплікації

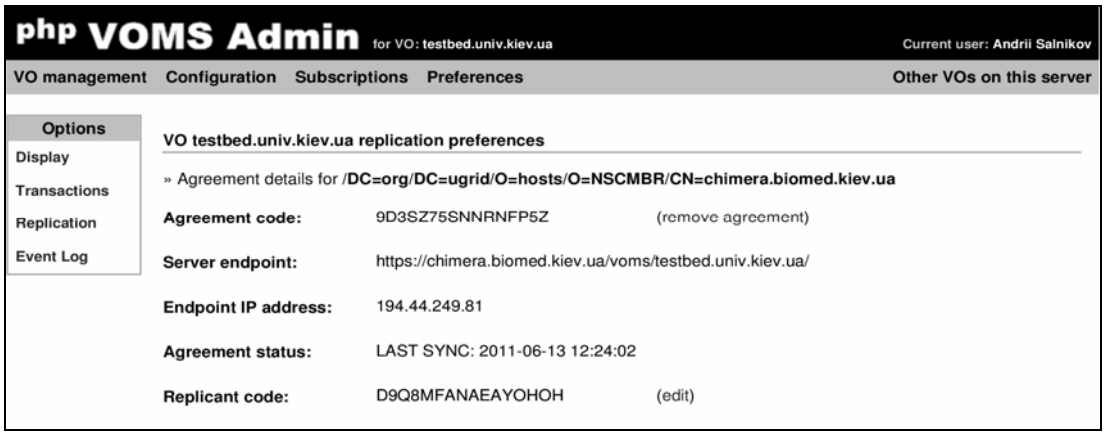

Рис. 3. Параметри узгодження реплікації

Висновки

Проведено дослідження механізмів забезпечення високої доступності грід-служби засвідчення та керування участю в віртуальних організаціях та сформовано критерії впровадження реплікації бази даних учасників на рівні ВО.

Запропоновано методики мульти-майстер реплікації на рівні застосувань, авторизації примірників служби та вирішення проблеми застосування автоінкременту первинних ключів за умови збереження схеми бази даних.

Методики було втілено в програмному продукті «PHP VOMS – Admin», що застосовується для обслуговування національних віртуальних організацій в Українському національному грід–сегменті.

Розроблене рішення впроваджено в роботу віртуальних організацій MolDynGrid [9], NetworkDynamics [10] та testbed.univ.kiev.ua для забезпечення відмовостійкості їх роботи. Примірники служби VOMS в ресурсних центрах віртуальних організацій працюють в режимі мульти-майстер реплікації з примірником центральної служби Київського національного університету імені Тараса Шевченка.

Для ВО testbed.univ.kiev.ua, в рамках якої проводиться навчання студентів університету та розробка програмного забезпечення, налаштовано мульти-майстер реплікацію поміж трьох примірників. Робота сервісів протягом півроку зарекомендувала себе, як відмовостійка конфігурація.

Список літератури

1. Український академічний Грід: досвід створення й перші результати експлуатації / Ю.В. Бойко, М.Г. Зинов'єв, О.О. Судаков, С.Я. Свістунов // Математичні машині і системи. – 2008. – Vol. 1. – P. 67 – 84.

2. Служба засвідчення та керування участю в ВО Українського національного гріду [Електронний ресурс]. – Режим доступу: http://grid.org.ua/voms. – 2012.

3. An Authorization System for Virtual Organizations / R. Alfieri, R. Cecchini, V. Ciaschini et al. // Proceedings of the 1st European Across Grids Conference, Santiago de Compostela. – 2003. – P. 13 – 14.

4. VOMS Replication, EGI Wiki. [Електронний ресурс]. – Режим доступу: https://wiki.egi.eu/wiki/VOMS.

5. Leach P. A Universally Unique IDentifier (UUID) URN Namespace. – RFC 4122 (Proposed Standard) [Електронний ресурс]. – 2005. – July. – Режим доступу: http://www.ietf.org/rfc/rfc4122.txt.

6. Team, The Globus Security. Globus Toolkit Version 4 Grid Security Infrastructure: A Standards Perspective. [Електронний ресурс]. – Режим доступу: http://www/unix.globus.org/toolkit/docs/5.0. — 2005.

7. Dierks, T. The Transport Layer Security (TLS) Protocol Version 1.1. – RFC 4346 (Proposed Standard). – 2006, April. – Obsoleted by RFC 5246, updated by RFCs 4366, 4680, 4681, 5746, 6176 [Електронний ресурс]. – Режим доступу: http://www.ietf.org/rfc/rfc4346.txt.

8. Salnikov A/. PHP VOMS–Admin project development [Електронний ресурс]. – Режим доступу: http://grid.org.ua/development/pva/. – 2011.

9. Virtual Laboratory MOLDYNGRID as a Part of Scientific Infrastructure for Biomolecular Simulations / A.O. Salnikov, I.A. Sliusar et al. // Int. Journal of Comp. – 2010. – Vol. 9, no. 4.

10. Salnikov A. Integrated grid environment for massive distributed computing in neuroscience / A. Salnikov, R. Levchenko, O. Sudakov // Intelligent Data Acquisition and Advanced Computing Systems (IDAACS), 2011 IEEE 6th International Conference on / IEEE. – 2011. – Vol. 1. – P. 198 –202.

Надійшла до редколегії 19.01.2012

Рецензент: д-р техн. наук, проф. С.В. Лістровий, Українська державна академія залізничного транспорту, Харків.

МЕТОДИКИ И СРЕДСТВА РЕПЛИКАЦИИ ГРИД-СЛУЖБЫ УДОСТОВЕРЕНИЯ И УПРАВЛЕНИЯ УЧАСТИЕМ В ВИРТУАЛЬНЫХ ОРГАНИЗАЦИЯХ

А.А. Сальников, Ю.В. Бойко

Проведенное исследование грид-службы удостоверения и управления участием в виртуальных организациях (ВО) показало, что серверная часть не предоставляет гибких механизмов обеспечения высокой доступности. Сформировано методики репликации базы данных ВО на уровне приложений (в режиме мульти–мастер). Методики воплощены в разработанном программном обеспечении – PHP VOMS – Admin. Экземпляры разработанной службы с поддержкой репликации внедрены в работу виртуальных организаций MolDynGrid, NetworkDynamics и testbed.univ.kiev.ua в Украинском национальном грид-сегменте.

Ключевые слова: *грид, виртуальная организация, служба управления участием в виртуальных организациях.*

VIRTUAL ORGANIZATION MEMBERSHIP SERVICE REPLICATION METHODS AND TOOLS

A..O. Salnikov, Yu.V. Boyko

Conducted virtual organization (VO) membership service analyses has shown that server backend does not provide flexible solutions to ensure high availability. Methods for VO database multi-master replication at the application layer has been formed. Methods has been implemented in the developed software named PHP VOMS – Admin. Replication–enabled instances of the developed service has been deployed into operation of MolDynGrid, NetworkDynamics and testbed.univ.kiev.ua virtual organizations in Ukrainian National Grid .

Keywords: *grid, virtual organization, virtual organization membership service (VOMS).*