

ПРОЕКТ РАЗВИТИЯ ЭЛЕКТРОННОЙ ИНФРАСТРУКТУРЫ ГРИД-КОМПЬЮТИНГА В СТРАНАХ ЮГО-ВОСТОЧНОЙ ЕВРОПЫ

*П. Богатенков, А. Андриеш, Г. Секриеру,
В. Сидоренко, А. Алтухов, В. Покотиленко*
Ассоциация RENAM

Интенсивное развитие технологий, связанных с потоками вычислений и электронной связью создает беспрецедентную возможность для инновационных изменений в науке, производстве, обществе и имеет высокий потенциал поддержки формирования новой экономики, основанной на информации и знаниях.

Приложения, развиваемые в настоящее время в этой динамичной среде требуют формирования широкомасштабных распределенных электронных инфраструктур (eInfrastructures), создаваемых путем интеграции существующих и развивающихся исследовательских сетей, рассредоточенных глобальных вычислительных фабрик, и новой вычислительной среды грид [1], основанной на модульном программном обеспечении среднего звена (middleware). Сейчас вложения в проекты, связанные с грид технологиями, во всем мире превышают миллиард долларов. Ожидается, что через 2-3 года эта технология выйдет на твердый уровень промышленных и бизнес приложений. В глобальных системах в качестве middleware используют так называемый Globus Toolkit - инструментарий, разработанный американскими учеными, который фактически стал мировым стандартом. Он включает в себя специальный протокол на основе HTTP для использования вычислительных ресурсов GRAM (Grid Resource Allocation Management); расширенную версию протокола для передачи файлов GridFTP; службу безопасности GSI (Grid Security Infrastructure); распределенный доступ к информации на основе протоколов SOAP и LDAP; удаленный доступ к данным через интерфейс GASS (Globus Access to Secondary Storage), .

Проблема развития и эксплуатации растущим сообществом пользователей огромного потенциала вычислительной среды грид (десятков и сотен тысяч процессоров, систем хранения информации большой емкости) и других электронных ресурсов, связанных между собой широкополосными оптическими сетями, имеет ключевую важность для Европейского Союза, международных и региональных научных и экономических образований.

В этом контексте, региональная инициатива SEE-GRID [2] продемонстрировала тот факт, что общая географически – независимая среда вычислительных ресурсов может иметь существенное научное значение для развития региона Юго-Восточной Европы (ЮВЕ), не владеющего собственными мощными вычислительными ресурсами. Через соединение региональной инфраструктуры с Европейской и всемирной электронной инфраструктурой, новая грид-среда стран ЮВЕ может получить несомненную выгоду от доступа к существующим вычислительной мощностям и развитым грид-приложениям. Потенциал технологий грид уже сейчас оценивается очень высоко: он имеет стратегический характер, и в близкой перспективе грид должен стать вычислительным инструментарием для развития высоких технологий в различных сферах человеческой деятельности, подобно тому, как подобным инструментарием стали персональный компьютер и Интернет.

Таким образом, можно приблизить достижение фундаментальной цели связанной с уменьшением цифрового неравенства в странах ЮВЕ и обеспечения равных возможностей для каждого гражданина региона.

Поддерживаемый ЕС проект SEE-GRID в мае 2006 г. перешел во вторую фазу SEE-GRID-2 и объединил научно-образовательные институты и сетевые организации таких

стран восточной Европы, как Албания, Босния-Герцеговина, Болгария, Хорватия, Македония, Греция, Венгрия, Сербия, Черногория, Румыния, Турция и Молдова.

Проект SEE-GRID-2 рассчитан на два года и призван достичь следующие основные цели:

- Обеспечение гарантий жизнеспособного функционирования и развития среды грид в странах ЮВЕ.
- Развитие региональной инфраструктуры грид в странах ЮВЕ.
- Сплочение рядов региональной сети и национальных сетей человеческих ресурсов.

Национальное научно-образовательное сообщество Молдовы RENAM (участник проекта SEE-GRID-2) приступило к созданию национальной грид-инициативы MD-GRID – координационно-административной инфраструктуры, объединяющей ведущие исследовательские институты и организации Молдовы, заинтересованные в развитии и продвижении технологии и приложений грид. Главная цель Молдовы в реализации программы проекта SEEGRID-2 это - включение сети RENAM и MD-GRID в общую инфраструктуру региональной сети SEE-GRID и обеспечение постоянного ее функционирования как полноценной операционной и функциональной составляющей.

Инициатива MD-GRID предусматривает построение Базового инфраструктурного операционного центра (БИОЦ) и Ресурсных центров (РЦ) организаций-участников MD-GRID для решения следующего комплекса задач:

- Развитие сетевой инфраструктуры RENAM и включение ее в европейские академические сети GEANT, интеграция с ресурсами SEEGRID, EGEE - Enabling Grids for E-science in Europe [3] и другими международными грид-проектами.
 - Создание и наращивание мощностей РЦ, установка и обновление версий ПО грид.
 - Поддержка функционирования БИОЦ и РЦ, обеспечение их бесперебойной работы: создание резервных копий, протоколирование и учет, восстановление после сбоев.
 - Управление функционированием инфраструктуры MD-GRID: регистрация пользователей, мониторинг, аудит.
 - Тестирование и поддержка новых версий ПО и приложений грид (на первом этапе – медицины, ядерной физики, биологии, 3D графики).
 - Поддержка пользователей и виртуальных организаций (VO).
 - Адаптация ПО грид, оценка и выработка рекомендаций по его развитию.
- Проектом предусмотрена установка кластеров грид и таких базовых сервисов, как:
- Брокер ресурсов (RB),
 - Управление виртуальными организациями - Virtual Organization Membership Service (VOMS),
 - Управление базами данных,
 - Система поддержки виртуальных организаций (VO),
 - Система грид-портала (PGRADE).

Научному сообществу Молдовы предстоит в ближайшее время приступить к освоению и развитию технологий грид, объединяющей множество ресурсов разных типов (процессоры, долговременная и оперативная память, хранилища и базы данных, сети), доступ к которым пользователь может получить из любой точки, независимо от места их расположения. Грид предполагает коллективный разделяемый режим доступа к ресурсам и к связанным с ними услугам в рамках глобально распределенных виртуальных организаций, состоящих из предприятий и отдельных специалистов, совместно использующих общие ресурсы.

В каждой виртуальной организации имеется своя собственная политика поведения ее участников, которые должны соблюдать установленные правила. Виртуальная организация может образовываться динамически и иметь ограниченное время существования.

На основе безопасного и надежного удаленного доступа к ресурсам глобально распределенной инфраструктуры грид способствует решению двух главных проблем:

1. Создания распределенных вычислительных систем сверхвысокой пропускной способности из серийно выпускаемого оборудования (показатели производительности: агрегированная мощность более 1 терафлоп, объем обрабатываемых данных более 1 петабайта в год) при одновременном повышении эффективности (до 100%) имеющегося парка вычислительной техники путем предоставления в грид временно простаивающих ресурсов;

2. Создания широкомасштабных систем мониторинга, управления, комплексного анализа и обслуживания с глобально распределенными источниками данных, способных поддерживать жизнедеятельность государственных структур, организаций и корпораций.

По мере своего развития грид проникает в промышленность и бизнес, крупные предприятия создают грид для решения собственных производственных задач. Таким образом, грид претендует на роль универсальной инфраструктуры для обработки данных, в которой функционирует множество служб (Grid Services), которые позволяют решать не только конкретные прикладные задачи, но и предлагают сервисные услуги: поиск необходимых ресурсов, сбор информации о состоянии ресурсов, хранение и доставка данных.

Применение грид может дать новое качество решения следующих классов задач:

- массовая обработка потоков данных большого объема;
- многопараметрический анализ данных;
- моделирование на удаленных суперкомпьютерах;
- реалистичная визуализация больших наборов данных;
- сложные бизнес-приложения с большими объемами вычислений.

Грид-технологии уже активно применяются как государственными организациями управления, обороны, сферы коммунальных услуг, так и частными компаниями, например, финансовыми и энергетическими.

Особое значение придается вопросам подготовки интеграции инфраструктур, созданных в рамках проекта SEEGRID с ресурсами более масштабной системы грид научного сообщества Европы EGEE, которая создается в рамках одного из крупнейших в мире проектов в этой области - проекта создания глобальной грид-инфраструктуры для компьютерной поддержки исследований в самых разных областях науки. В результате его осуществления ученые, работающие в академических институтах и в промышленности, получают доступ к значительным ресурсам вычислительной техники вне зависимости от того, где они находятся.

В настоящее время в инфраструктуре EGEE работает около 20000 CPU, расположенных на более чем 100 грид-сайтах в 27 участвующих в проекте странах. Шесть из этих сайтов поочередно берут на себя оперативное руководство глобальной инфраструктурой распределенных компьютерных ресурсов. Хотя EGEE и европейский проект, в нем участвуют и американские университеты - родоначальники концепции грид.

Один из ключевых факторов успеха проекта SEE-GRID – наличие в грид-инфраструктуре множества приложений, представляющих как можно большее разнообразие областей специализации. Поэтому направление "Определения областей и поддержки приложений" способствует включению в проект новых пользователей, сообществ и виртуальных организаций.

Очень важна стратегия размещения и интеграции приложений. На начальных этапах реализации проекта была разработана процедура подачи заявок на включение приложений в инфраструктуру SEE-GRID и критерии оценки приложений, а также учреждена Консультативная комиссия по базовым приложениям для оценки всех заявок и выработки рекомендаций.

В настоящее время сформирован расширенный комплект из 10-ти приложений, прошедших экспертизу и которые планируется включить в инфраструктуру грид. Они принадлежат разнообразным областям, среди которых – науки о Земле, астрофизика, медицина, вычислительная химия и биоинформатика и другие. В дальнейшем каждая страна-участник имеет возможность подать заявку на участие своего приложения в инфраструктуре.

Жизненно важной стороной проекта SEE-GRID2 является безопасность. Обеспечение безопасности в разделяемых средах коллективного использования всегда представляет трудность, особенно если речь идет о международной работе и географически больших территориях.

Обеспечение безопасности распределенных ресурсов проекта базируется на общих принципах и архитектуре безопасности, определенных ранее группой обеспечения безопасности проекта EGEE, а также составленном своде основополагающих правил, которые носят руководящий характер для всех направлений работы в рамках проекта. Согласованность всех этих сторон заметно усиливает грид-инфраструктуру хорошо упорядоченной системой безопасности и тем, что пользователям будет достаточно единственного входа в систему для получения доступа к разным ресурсам. Архитектура безопасности основана на требованиях, предъявляемых как к пользователям, так и к поставщикам ресурсов. Реализованная и действующая архитектура безопасности действует в рамках программы EUGridPMA, координирующей деятельность всех участников в соответствии с политикой согласованной и подробно изложенной в документах.

В 2004 году Операционный Центр GridAUTH Университета Aristotle г. Фессалоники ввел в эксплуатацию Центр Сертификации SeeGrid Certification Authority [4] – Центральный Сертификационный Орган (CO) проекта SEE-GRID. Целью его создания было предоставление сервисов PKI (Private Key Infrastructure - инфраструктура открытого ключа) в странах ЮВЕ – членах проекта SEE-GRID, которые не имели возможности создать свою PKI инфраструктуру для аутентификации в собственных грид-системах.

Сертификационный Орган SEE-GRID управляется специалистами вышеупомянутого университета под надзором GRNET, координатора проекта SEE-GRID. GRNET находится в собственности и управляется Генеральным Секретариатом по Исследованиям и Технологиям Министерства Развития Греции (General Secretariat of Research and Technology, Greek Ministry of Development).

Одной из целей проекта SEE-GRID2, как идеологического преемника SEE-GRID, является создание федерации независимых друг от друга Сертификационных Органов без ярко выраженной центрально управляемой иерархической структуры. Управляться федерация будет Координационным советом представителей Сертификационных Органов стран участниц. Каждый участник проекта реализует собственную PKI инфраструктуру, функционирующую в рамках политики, оговоренной в Сертификационной Политике CO.

Реализация проекта SEE-GRID2 позволит построить, протестировать и запустить в промышленную эксплуатацию стабильную, надежную и интероперабельную грид-инфраструктуру в странах ЮВЕ, включая Молдову, создать новые коллективы пользователей, оказать помощь этим коллективам в их стремлениях овладения новыми мощными грид-инструментами и службами на благо е-науки. Несомненно, что все это должно благоприятно сказаться на создании интегрированного "виртуального исследовательского грид-пространства", охватывающего Европу и мировое сообщество научного грид-компьютинга.

ЛИТЕРАТУРА

- I. Foster, C. Kesselman, and S. Tuecke, "The Anatomy of the Grid: Enabling Scalable Virtual Organizations," International Journal of Supercomputing Applications, vol. 15, no. 3, 2002.
- South Eastern European Grid Enabled eInfrastructure Development. <http://www.see-grid.eu/>, <http://www.see-grid.org/>

- EGEE (Enabling Grids for E-scienceE). EGEE Project Portal. <http://www.eu-egee.org/>
- SeeGrid Certification Authority <http://www.grid.auth.gr/pki/seegrid-ca/>