

В апреле 2004 года стартовал большой европейский проект EGEE (Enabling Grids for E-science in Europe). В числе российских организаций, образовавших консорциум RDIG (Russian Data Intensive GRID), Институт математических проблем биологии РАН. Благодаря организации гид-инфраструктуры, которая развивается в рамках финансируемого ЕС проекта EGEE, для всего научного сообщества станут доступными не имеющие себе равных вычислительные мощности и объёмы информации.

Что такое EGEE?

Цель проекта EGEE - объединить уже ведущиеся национальные, региональные и тематические гид-разработки в единую цельную гид-инфраструктуру для поддержки научных исследований. EGEE предоставит исследователям как в академических кругах, так и в разных областях экономики круглосуточный доступ к самым высокопроизводительным вычислительным ресурсам независимо от их географического положения. Пользоваться инфраструктурой смогут географически распределённые сообщества исследователей, которые нуждаются в общих для них вычислительных возможностях гридов, готовы объединить свои собственные вычислительные инфраструктуры и согласны с принципами общего доступа. Проект поддерживают, в основном, финансирующие учреждения ЕС, но предназначен он для работы во всём мире. Значительные средства поступают от США, России и других участников проекта, не входящих в ЕС.

В каких научных дисциплинах будет полезен EGEE?

Коротко на этот вопрос можно ответить так: выиграет любая научная дисциплина, где уже сейчас широко применяются передовые вычислительные средства. Для отработки начального уровня внедрения развивающейся гид-инфраструктуры и официальной оценки её эксплуатационных качеств и функциональности выбраны две практические области. Одна - Large Hadron Collider Computing Grid (LCG, вычислительный гид для ЛНС), где гид-инфраструктура обеспечивает хранение и анализ петабайтов (10^{15} байтов) реальных и смоделированных данных из экспериментов в физике высоких энергий, ведущихся в CERN. Другая - биомедицинские гриды, где несколько коллабораций решают одинаково сложные задачи, например, поиск в геномных базах данных и индексирование больничных баз данных, что составляет несколько терабайтов в год для одной больницы.

Как новые научные сообщества смогут получить доступ в EGEE?



Для включения в гид-инфраструктуру EGEE новых пользователей и научных сообществ будет действовать Служба приёма заявок и поддержки (Application Identification and Support Activity), которая будет идентифицировать и поддерживать начинающих пользователей из широкого

круга академических дисциплин и областей экономики. Предусмотрено также включение в число пользователей на основании заявок, подаваемых Консультативной комиссией по тематически классифицированным заявкам (EGAAP, EGEE Generic Applications Advisory Panel) представителями определённых исследовательских сообществ. Комиссия будет оценивать заявки с таких сторон, как научный интерес; повышение эффективности работы, представленной в заявке, от применения гридов; осведомлённость кандидатов о возможностях грид-инфраструктуры. Комиссия передаст свои рекомендации Исполнительному совету проекта EGEE (EGEE Project Executive Board); кандидаты, чьи заявки будут одобрены, получают помощь в приведении своего программного обеспечения в соответствие с требованиями грид-инфраструктуры.

Какое обучение EGEE предложит пользователям?

В рамках проекта будет работать Служба обучения и включения в число пользователей (Training and Induction Activity); она выпустит комплект учебных материалов и курсов на английском языке. Комплект составлен для начального и повышенного уровней. Обучение будет организовано по региональному принципу, и ключевые материалы могут быть переведены на соответствующий европейский язык. Основные типы учебных курсов - следующие: вводные курсы (2 дня, минимум 10 в год); обучение разработчиков приложений (4 дня, минимум 8 в год); курсы повышенного уровня (5 дней, минимум 2 в год) и узкоспециальные курсы по конкретной технической деятельности (2 дня, минимум 6 в год).

Как пользователи смогут держаться в курсе развития EGEE?

В рамках проекта будет работать Служба распространения информации и расширения круга пользователей (Dissemination and Outreach Activity); она будет держать собранные воедино все главные сведения о проекте вместе со ссылками на соответствующие региональные веб-сайты. Кроме того, ключевые пользовательские группы и группы потенциальных пользователей будут получать сообщения из специализированных списков рассылки информации по e-mail. В сети будут проведены несколько принципиально важных для EGEE мероприятий, участие в которых всех заинтересованных научных работников весьма приветствуется.

Каковы временные рамки проекта EGEE?

EGEE - двухлетний проект; официально он начат 1 апреля 2004. Он задуман как часть 4-летней программы. Проект стартовал в исключительно благоприятных условиях, созданных заранее: до его формального начала уже были размещены основные сервисы и начаты разработка промежуточного программного обеспечения и распространение информации. Доступные ресурсы и пользовательские группы быстро разрастутся в ходе выполнения проекта. Ожидается, что преемником EGEE станет второй двухлетний проект, в котором работа и поддержка устойчивой грид-инфраструктуры перейдёт из научного сообщества в разные области экономики - подобно тому, как всеевропейская многогигабитная научно-образовательная сеть для передачи данных GEANT последовательно перешла из общественного сектора в частный.

Каковы критерии оценки успеха EGEE?

Ожидается, что к концу второго года проекта число активных пользователей грид-инфраструктуры возрастёт до 3000; они будут представлять не менее пяти дисциплин. Доступная в рамках EGEE грид-инфраструктура увеличится со свыше 3000 ЦПУ в начале проекта до свыше 8000 к концу второго года. Успех грид-инфраструктуры будет оцениваться по нескольким

показателям качества сервиса. Кроме количественной оценки того, насколько грид-технология повышает эффективность решения разных задач, внимание будет уделяться и качественным достижениям в плане функциональности, которая ранее не была доступна участвующим в проекте научным сообществам.

Кто стоит за проектом EGEE?



В проекте EGEE участвуют 70 ведущих организаций приблизительно из 27 стран. Эти организации объединены в региональные гриды, суммарная максимальная мощность которых составляет свыше 20 тыс. ЦПУ. Это самая большая из всех когда-либо составленных международных грид-инфраструктур. ЕС выделяет на этот проект 32 млн. евро; столько же - участники проекта. Людские ресурсы, задействованные в проекте, составляют приблизительно 600 человеко-лет на протяжении двух лет. Распределение финансирования работ таково: 48% - на грид-сервисы, 24% - на перепроектирование промежуточного программного обеспечения и 28% - на распространение информации, расширение круга пользователей и обучение. В ходе развития EGEE структура мирового грид-сообщества разделилась на регионы - участники проекта. Каждый из них представлен в Правящем совете проекта (Project Management Board). Точно так же там представлены: ведущий участник проекта - CERN, Европейские исследовательские сети (European Research Networks) и LCG. Контактные лица по вашей стране или вашему региону указаны на веб-сайте проекта.

Какое отношение имеет EGEE к существующим гридам и исследовательским сетям?

Инфраструктура EGEE будет построена на исследовательской сети Европейского союза (ЕС) GEANT и в ней будет использован опыт, накопленный в таких проектах, как DataGrid, поддержанный ЕС, других грид-проектах ЕС и в национальных проектах, например, e-Science (Британия), INFN Grid, Nordugrid и американском Trillium (PPDG, GriPhyN and IVDGL). Она обеспечит возможность совместной работы с другими гридами во всём мире, включая США и Азию, что будет способствовать установлению всемирной грид-инфраструктуры.

В чём выражается функциональность EGEE?

- **Упрощённый доступ.** EGEE уменьшит издержки, связанные с разнообразием не связанных между собой систем учёта пользователей. Пользователи смогут объединяться в виртуальные организации с доступом к гриду, содержащему нужные каждому пользователю рабочие ресурсы.
- **Выполнение вычислений по требованию.** Эффективно распределяя ресурсы, грид-технология значительно сокращает время ожидания доступа к ним.
- **Географически распределённый доступ.** Инфраструктура будет доступна отовсюду, где обеспечена хорошая стыкуемость сетей. Ресурсы, таким образом, становятся более широко доступными.
- **Исключительно большой объём ресурсов.** Благодаря согласованности ресурсов и пользовательских групп, в прикладной работе в рамках EGEE будут доступны ресурсы таких объёмов, какие не может предоставить ни один компьютерный центр.
- **Совместное использование программного обеспечения и данных.** Благодаря единой структуре вычислительных средств, в EGEE географически распределённым

сообществам пользователей будет легко совместно использовать программное обеспечение и базы данных и разрабатывать программное обеспечение.

- **Высокий уровень поддержки приложений.** Компетентность всех участников EGEE обеспечит тщательную, всестороннюю поддержку всех основных приложений.

Каков будет уровень сервисов EGEE?

Группы обеспечения эксплуатации гридов будут гарантировать промышленный уровень грид-сервисов. Основными качествами, определяющими такой уровень грид-сервисов, являются их управляемость, ошибкоустойчивость, устойчивость к разного рода неисправностям, единый подход к обеспечению безопасности, а также их расширяемость, необходимая для включения в работу новых ресурсов немедленно по мере их появления. Главные задачи этих групп - поддержка базовых сервисов инфраструктуры, мониторинг и контроль гридов, размещение промежуточного программного обеспечения, подключение ресурсов, поддержка ресурсов и пользователей, общее управление гридами и международное сотрудничество. Оперативный центр (Operations Management Centre) в CERN (Швейцария) будет согласовывать работу пяти центров базовой инфраструктуры (Core Infrastructure Centre) (CERN, Франция, Италия, Россия, Британия) и восьми региональных оперативных центров (Regional Operation Centre). Работа EGEE для массового пользователя вначале будет основана на промежуточном программном обеспечении и сервисах проекта LCG.

Какое промежуточное программное обеспечение будет основой EGEE?

Служба разработки и интеграции промежуточного программного обеспечения для гридов (Grid Middleware Engineering and Integration Activity) будет поддерживать и постоянно совершенствовать набор программных средств, благодаря которому грид-сервисы промышленного уровня будут доступны основному кругу пользователей, причём ожидается, что этот круг стремительно разрастётся и станет намного разнороднее. Деятельность, связанная с промежуточными программными средствами EGEE, заключается, главным образом, в их перепроектировании в плане их функциональности, для чего нынешнее поколение промежуточного программного обеспечения совершенствуется с учётом богатого опыта участников EGEE. Промежуточное программное обеспечение будет развиваться в направлении сервисно-ориентированной архитектуры, основанной на стандартах, разработанных в рамках веб-сервисов и соответствующих сообществ. Центры перепроектирования промежуточного программного обеспечения будут отвечать за следующие ключевые сервисы: доступ к ресурсам (Италия), организация данных (CERN), поиск и сбор информации (Британия), брокерские операции с ресурсами и учёт ресурсов (Италия), безопасность (страны Северной Европы). Непосредственное отношение к этой работе будут иметь группы обеспечения качества и безопасности гридов (Quality Assurance and Grid Security). Центр интеграции и тестирования промежуточного программного обеспечения (Middleware Integration and Testing Centre) будет находиться в CERN. Начать эту работу решено с интеграции части наиболее важного промежуточного программного обеспечения гридов от Virtual Data Toolkit (VDT), европейского проекта DataGrid (European DataGrid Project, EDG) и проекта AliEN.

Контактные адреса EGEE:

- Отдельные вопросы опубликованы на официальном сайте проекта: www.eu-egee.org
- Общие сведения для средств массовой информации - pressoffice@eu-egee.org