



Enabling Grids for
E-science in Europe

В недалёком будущем Вам как научному работнику станут доступны не имеющие себе равных вычислительные мощности и объёмы информации. Это станет возможным благодаря Грид-инфраструктуре, которая развивается в рамках финансируемого ЕС проекта *Enabling Grids for E-science in Europe* ("Развёртывание Грид-систем для развития е-науки в Европе"), более известного как EGEE. Настоящая публикация представляет краткий обзор важнейших сторон проекта EGEE, знание которых понадобится Вам как потенциальному пользователю Грид-инфраструктуры.

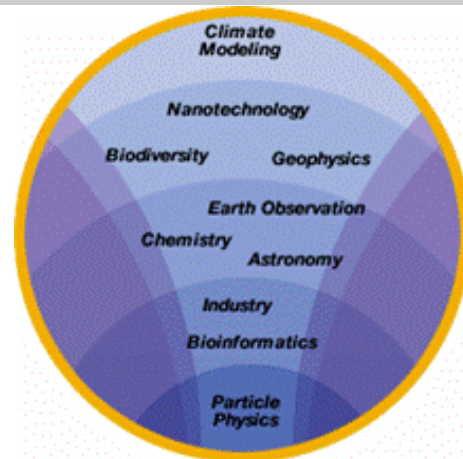
Что такое EGEE?

Цель проекта EGEE – объединить уже ведущиеся национальные, региональные и тематические Грид-разработки в единую цельную Грид-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 Грид-систем, 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 Грид-инфраструктуры и официальной оценки её эксплуатационных качеств и функциональности выбраны две практические области. Одна – Large Hadron Collider Computing Grid (LCG, вычислительная Грид-система для LHC), где Грид-инфраструктура обеспечивает хранение и анализ петабайтов (10^{15} байтов) реальных и смоделированных данных из экспериментов в физике высоких энергий, ведущихся в CERN. Другая – биомедицинские Грид-системы, 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 Грид-инфраструктуру EGEE новых пользователей и научных сообществ будет действовать Служба идентификации и поддержки приложений (Application Identification and Support Activity), задачей которой является определение областей приложений и поддержка пользователей из широкого круга академических дисциплин и промышленности. Предусмотрено также включение в число пользователей на основании заявок, подаваемых Консультативной комиссией по базовым областям приложения (EGAAP, EGEE Generic Applications Advisory Panel) представителями определённых исследовательских сообществ. Комиссия будет оценивать заявки с таких сторон, как научный интерес; повышение эффективности от применения Грид-систем; 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.

С подробностями можно ознакомиться в разделе "Связь по вопросам подачи заявки" (application contact) на www.eu-egee.org

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 будет 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).

С подробностями можно ознакомиться в разделе "Связь по вопросам обучения" (training contact) на www.eu-egee.org

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 (см. www.eu-egee.org). 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.

За подробностями обращайтесь в службу распространения информации (EGEE Dissemination Office) – pressoffice@eu-egee.org

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 перейдет из научного сообщества в разные области промышленности – 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 возрастет до 3000; они будут представлять не менее пяти дисциплин. Доступная в рамках EGEE Грид-инфраструктура составит свыше 3000 ЦПУ в начале проекта и 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 Грид-инфраструктуры, 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 разработки проекта EGEE произошло структурирование мирового Грид-сообщества по регионам – участникам 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.

Подробнее – см. www.eu-egee.org

Какое отношение имеет EGEE к существующим Грид-системам 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,



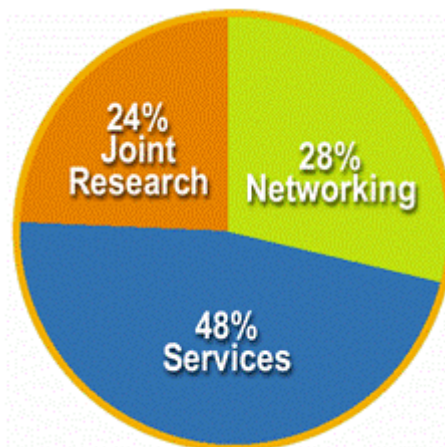
других Грид-проектах ЕС и в национальных проектах, например, e-Science (Британия), INFN Grid, NorduGrid и американском Trillium. Она обеспечит возможность совместной работы с другими Грид-системами во всём мире, включая США и Азию, что будет способствовать установлению всемирной Грид-инфраструктуры.

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 Грид-системе, содержащей 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 и всестороннюю поддержку всех основных приложений.

Каков будет уровень сервисов EGEE?

Группы обеспечения эксплуатации Грид-систем будут гарантировать промышленный уровень Грид-сервисов. Основными качествами, определяющими такой уровень Грид-сервисов, являются их управляемость, надёжность, устойчивость к разного рода отказам, 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 Грид-систем, размещение промежуточного программного обеспечения, подключение ресурсов, поддержка ресурсов и пользователей, общее управление Грид-системами 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.



Подробнее о проекте LCG – см. www.cern.ch/lcg

Какое промежуточное программное обеспечение будет основой EGEE?

Служба разработки и интеграции промежуточного программного обеспечения для Грид-систем (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 Грид-систем (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 Грид-систем от Virtual Data Toolkit (VDT), европейского проекта DataGrid (European DataGrid Project, EDG) и проекта AliEN.

В чём состоит выгода от предоставления ресурсов проекту EGEE?

Поставщики ресурсов получают следующие преимущества:

- **Возможность работы с использованием колоссальных ресурсов.** EGEE обеспечивает координированную крупномасштабную рабочую среду. Для каждого поставщика ресурсов это выразится в значительном снижении расходов и, в то же время, в повышенном уровне сервиса.
- **Выигрыш от высокого уровня компетентности специалистов, лежащего в основе инфраструктуры.** Обеспечивая своим участникам доступ к сервисам, EGEE привлекает ведущих европейских специалистов к созданию и поддержке инфраструктуры.
- **Общение с пользователями.** Организация поддержки со стороны EGEE будет допускать возможность адаптации с учётом региональной специфики и предполагать тесные связи с региональными пользовательскими сообществами.
- **Возможность создания коллабораций между поставщиками ресурсов.** Поставщики ресурсов смогут образовывать коллаборации и открывать новые направления разработки и поддержки, не включённые в настоящую версию проекта. Затраты на исследования и разработки будут в таком случае разделены между участниками.

Какой интерес EGEE представляет для промышленности?

Промышленность извлечёт из EGEE выгоду, выступая по отношению к нему в трёх качествах:

- **Промышленность как участник EGEE.** Сотрудничая с отдельными участниками EGEE, промышленность может участвовать в решении конкретных задач, благодаря чему Грид-технологии получают дальнейшее развитие.
- **Промышленность как пользователь.** Применяя компьютерные сети в своей работе, конкретные отрасли промышленности станут потенциальными пользователями Грид-инфраструктуры – прежде всего, в области исследований и разработок.
- **Промышленность как обеспечивающая сторона.** Создание Грид-систем промышленного уровня потребует долгосрочного участия представителей разных областей промышленности в сопровождении Грид-сервисов: центров обработки вызовов, центров поддержки и центров предоставления компьютеринг-ресурсов.

Главные задачи промышленного форума проекта EGEE – повышение осведомлённости о EGEE в разных отраслях промышленности и поощрение участия фирм в EGEE. Для выполнения этих задач будут развиваться прямые контакты между участниками проекта и представителями разных отраслей промышленности с целью обеспечить соответствие Грид-инфраструктуры требованиям предприятий в плане стандартов и функциональности и, с другой стороны, гарантировать извлечение пользы для проекта из практического опыта фирм.

Подробнее о промышленном форуме см. раздел "Контакты" на сайте проекта, www.eu-egee.org

Ожидается, что EGEE также внесёт значительный вклад в работу учреждений, вырабатывающих стандарты – например, Глобального Грид-форума, и будет определять стратегию развития таких областей, как безопасность в Грид-системах, диспетчеризация ресурсов и деятельность сертифицирующих органов. Всё это повлияет на развитие промышленных Грид-стандартов.

Наконец, вполне вероятно, что EGEE станет разработчиком передовых информационных технологий. Тогда его польза для промышленности и коммерции выйдет далеко за пределы научных расчётов – подобно тому, как World Wide Web, изначально создававшийся для науки, оказал гораздо более широкое влияние на общество в целом.

Контактные адреса EGEE:

- Отдельные конкретные вопросы уже опубликованы для обращений на официальном сайте проекта, www.eu-egee.org
- Общие сведения для средств массовой информации – pressoffice@eu-egee.org
- Правление проекта – project-eu-egee-po@cern.ch