

Приложения, работающие в EGEE

Проект "Развёртывание гридов для развития е-науки" (Enabling Grids for E-science, EGEE) вначале был сосредоточен на двух чётко сформировавшихся областях – физике элементарных частиц и биологических науках, поскольку эти области уже были приобщены к грид-технологиям и с самого начала проекта были готовы к размещению приложений для решения сложных реальных задач.

Позже в EGEE был составлен расширенный комплект приложений из разнообразных областей, среди которых – науки о Земле, астрофизика элементарных частиц, вычислительная химия. На последней конференции проекта в ноябре 2004 Консультативная комиссия по базовым приложениям (EGEE Generic Applications Advisory Panel, EGAAP) рекомендовала включить в инфраструктуру EGEE ещё четыре области: космологию (проект "Планк"); создание медикаментов (разработка препаратов от малярии); продолжение наук о Земле – гидрологию; проект GRACE (Grid enabled seArch and Categorization Engine – грид-система поиска и категоризации).

Приложения в физике высоких энергий

С самого начала проекта EGEE его инфраструктура широко используется в четырёх экспериментах в физике элементарных частиц на LHC (Large Hadron Collider – большой адронный коллайдер): ALICE (A Large Ion Collider Experiment – эксперимент "Большой ионный коллайдер"), ATLAS, CMS (the Compact Muon Solenoid experiment – эксперимент "Компактный мюонный соленоид" и LHC. Эти проекты продолжают использовать инфраструктуру EGEE, готовясь к пуску LHC в CERN в 2007 г. У каждого эксперимента – свои физические цели, но во всех этих экспериментах требуется широкомасштабное, основанное на моделировании исследование событий, происходящих при столкновении высокоэнергетичных пучков протонов или тяжёлых ионов. Примером такого применения инфраструктуры EGEE может служить эксперимент по физике "очарованных" частиц LHCb (the LHC beauty experiment), в котором было генерировано 300 млн. событий на 63 сайтах, географически распределённых по миру. Работа по LHCb велась на 4500 ПК; объём результатов – свыше 100 терабайтов данных, распределённых по европейским научным центрам.

В каждом эксперименте LHC используется своя среда из многих гридов, сложившаяся на основе предыстории эксперимента и его ограничений. В двух крупнейших экспериментах LHC – ATLAS и CMS – участвуют около 200 институтов; значительная часть этих ресурсов находится в США. В обоих экспериментах в дополнение к EGEE использовался проект GRID3 и были созданы рабочие интерфейсы к этим гридам. В эксперименте ATLAS используется также NorduGrid. В экспериментах ALICE и LHCb ещё до EGEE была разработана собственная распределённая технология, подобная гридам, в системах ALIEN и DIRAC. Эта технология управления данными и заданиями до сих пор частично задаёт направление развития программного обеспечения gLite.

Инфраструктура EGEE использовалась и в других международных экспериментах LHC, среди которых – Babar (B and B-bar), CDF (Collider Detector at Fermilab – коллайдерный детектор в

Фермилаб), серия экспериментов D0 на коллайдерах в США, ZEUS – эксперимент на коллайдере в Германии.

В проекте ARDA (A Realization of Distributed Analysis for LHC – реализация распределённого анализа для LHC) разработаны прототипы систем анализа, у которых есть интерфейс с промежуточным программным обеспечением gLite, благодаря чему в физике высоких энергий можно будет перейти к обработке данных отдельными пользователями. Это – децентрализованное использование гридов сотнями отдельных пользователей в противоположность десяткам руководителей работ.

Биомедицинские приложения

С самого начала проекта EGEE в него были включены биомедицинские приложения. Сейчас в EGEE доступны и регулярно используются девять приложений; рассматриваются ещё несколько инициатив.

Сейчас в EGEE доступны следующие приложения:

- **GATE** – планирование радиотерапии и применение медицинской томографии.
- **GPS@** (портал, специализирующийся на геномах) – портал грид-средств анализа белков.
- **CDSS** (Clinical Decision Support System – клиническая система поддержки принятия решений) – система помощи практикующим врачам, выдающая релевантные сведения на основе огромного объёма информации.
- **gPTM3D** (интерактивное средство визуализации и обработки радиологических данных) – интерактивный поиск и анализ медицинских данных в гридах.
- **Mammogrid** – разработка общеевропейской базы данных маммограмм, которая может использоваться важными медицинскими приложениями, а также может значительно способствовать развитию сотрудничества в Европейском союзе между специалистами в этой области.
- **SiMRI 3D** – сервис моделирования изображений, который поможет научным работникам лучше понять сложную технологию магнитного резонанса.
- **xmipp_Mlrefine** (трёхмерный макромолекулярный структурный анализ) – трёхмерный структурный анализ макромолекул на основе изображений с сильным фоном, полученных на электронных микроскопах.
- **GridGRAMM** – простой интерфейс для создания молекулярных цепочек в сети. Его нынешняя версия позволяет с низким и высоким разрешением подбирать молекулы для пар белок – белок и лиганд – рецептор.
- **GROCK** – удобный сетевой сервис отбора межмолекулярных взаимодействий по массе молекул.

В инфраструктуре LCG2 размещены и в настоящее время проходят испытания три новых биомедицинских приложения в области геномики и разработки медикаментов:

- **SPLATCHE** (моделирование эволюции геномов) – моделирование расселения человека в реалистичных ландшафтах, учитывающее пространственно-временную неоднородность окружающей среды, и генерирующее молекулярное разнообразие одной или нескольких выборок генов.
- **Drug Discovery** – докинг-платформа для исследования тропических болезней. Это приложение нацелено на создание принципиально новых препаратов от малярии – болезни, второй по числу умирающих от неё людей (на первом месте – туберкулёз). Оно позволяет рассчитать вероятность присоединения данного химического вещества к данному белку.
- **Xmipp_assign_multiple_CTFs** – средство определения частотно-контрастной характеристики, искажающей изображения, полученные на электронных микроскопах.

Базовые приложения

- Проект **GRACE** (Grid enabled seArch and Categorization Engine) – грид-система поиска и категоризации завершён в феврале 2005. GRACE дал проекту EGEE конкретную и быструю обратную связь по использованию гридов, а также способствовал распространению знаний о EGEE.
- Приложение **MAGIC** моделирует ливни частиц в атмосфере, порождённые высокоэнергетичными космическими лучами. Это моделирование необходимо для изучения происхождения и свойств высокоэнергетичных гамма-лучей на основе анализа данных, полученных на телескопе MAGIC, находящемся на Канарских островах. Работа с этими данными в инфраструктуре EGEE стартовала в начале марта 2005.
- Проект **Planck** Европейского космического агентства – микроволновое картографирование неба, основанное минимум на двух полных съёмках неба с небывалым сочетанием масштаба охваченного пространства, ширины частотного спектра, а также точности, устойчивости и чувствительности измерений. Спутник будет запущен в 2007; его коммерческая нагрузка будет состоять из набора микроволновых и субмиллиметровых детекторов, объединённых в низко- и высокочастотную установки. Будет охвачен спектр частот от 30 до 900 ГГц.

Область наук о Земле разбита на пять доменов: геонауки, гидрология, наблюдения за поверхностью Земли, климатология, физика твёрдой Земли. В каждом домене минимум одно приложение развёрнуто в EGEE и GILDA

Первым промышленным приложением, успешно работающим в EGEE, стал **Geocluster** – ведущее средство обработки промышленных данных в области сейсмологии, часть виртуальной организации Expanding Geosciences on Demand ("Заказные прикладные разработки в области наук о Земле", EGEODE), ведущей исследования для государственных и частных лабораторий – как проектно-конструкторских, так и академических. Geocluster позволяет изучать состав слоёв земной коры.

В домене "**Наблюдения за поверхностью Земли**" размещены результаты съёмки озонного слоя спутником GOME и данные эксперимента ERS/SAR по спутниковой съёмке разливов нефти. Грид

облегчит определение состояния озонового слоя, обеспечивая единую вычислительную среду для разных этапов работы и предоставляя "поставщику" и "потребителю" данных возможность легко ими обмениваться и совместно их использовать.

Приложения в домене **"Физика твёрдой Земли"** специализируются на механизме землетрясений и численном моделировании землетрясений в сложных трёхмерных геологических моделях.

Первое приложение в домене **"Гидрология"** изучает влияние хозяйственной деятельности человека на проникновение морской воды в прибрежный водоносный слой в Средиземноморье. Применяется моделирование Монте-Карло, основанное на трёхмерной модели движения грунтовых вод и переноса соли; в модели учитывается плотность среды.

В домене **"Климатология"** из межгридовой системы отладки (CrossGrid testbed) в EGEE переносится приложение, предсказывающее наводнения. Оно состоит из каскадообразной структуры моделирования, в которую введены экспериментальные данные и применяются метеорологические, гидрологические и гидравлические модели.

Узнать подробнее о приложениях, работающих в EGEE, можно на сайте NA4:

<http://egee-na4.ct.infn.it/index.php> или связавшись с контактным лицом: Флоренс Жак (Florence Jacq), florence.jacq@clermont.in2p3.fr.