

Биологические приложения

Введение: краткие сведения о проекте "Развёртывание Грид-систем для развития е-науки" (Enabling Grids for E-science)

EGEE – европейская инициатива (имеющая связи во всём мире), цель которой – создание для е-науки в течение ближайших двух лет стабильной, надёжной Грид-инфраструктуры промышленного качества. Предусматривается поддержка переноса приложений в Грид-системы и специализированное обучение на всех уровнях. Ожидается, что в первые два года в EGEE будет свыше 3000 пользователей более чем в 15 странах. За это время компьютеринг-ресурсы Грид-систем EGEE должны составить около 9000 сетевых узлов и 1000 терабайт (1 Tb = 1024 гигабайтов) хранимой информации.

Какая польза биокомпьютингу от Грид-технологии?

Многие задачи биологии, анализ медицинских изображений и другие связанные с медициной приложения – это области, где в самой полной мере можно воспользоваться спецификой Грид-компьютинга. По своей природе Грид-система работает распараллеленно (одновременно ведутся похожие вычисления); она объединяет распределённые ресурсы и находящиеся вдали друг от друга пользователей; она даёт доступ к мощнейшим вычислительным ресурсам; она использует колоссальные объёмы данных, которые часто хранятся в разнородных и географически распределённых базах данных; наконец, она позволяет учёным совместно работать, образуя виртуальные сообщества.

Вот несколько примеров биокомпьютинга, где требуются такие свойства Грид-систем:

- Анализ больших объёмов данных, например, анализ геномов, популяционная генетика, испытания лекарств.
- Распределённые приложения, например, доступ к технологиям работы с изображениями для медицинских работников через мобильные средства компьютеринга с возможностью самой сложной обработки информации и вычислений в полевых условиях и доступа к распределённым средствам установления последовательности вычислительных процессов и средствам протеомики (возможна разработка последовательности выполняемых действий).
- Моделирование: например, свёртывание белка, биохимические цепочки, моделирование целых клеток, экологическое моделирование, биомедицинская визуализация (например, моделирование магнитнорезонансных изображений).
- Массивы разнородных данных: например, объединённые данные по геномам, аннотации и информация по биохимическим цепочкам из разных биоинформационных баз данных.

EGEE и биокомпьютинг

В проекте EGEE биокомпьютинг рассматривается как одна из важнейших областей, благодаря которой грид-компьютинг будет распространяться в Европейском союзе: предшественником EGEE был проект European DataGrid (EDG), в ходе выполнения которого был накоплен опыт биомедицинских приложений. Чтобы упростить доступ биологического сообщества к участию в EGEE, часть этого проекта отведена специально для определения областей приложений и размещения приложений биокомпьютинга в Грид-инфраструктуре EGEE.

Обучение в рамках EGEE

Учебная деятельность – отдельное направление работы в рамках EGEE. Она нацелена на упрощение доступа к Грид-компьютерному сообществу, представляющих определённые области его возможного приложения, например, биомедицину и биоинформатику. Обучение может вестись на любом уровне: от элементарного знакомства с Грид-технологией до курса подготовки экспертов – разработчиков приложений. Учебную работу возглавляет Британский национальный центр e-науки в Эдинбурге; на федеративных началах с ним работают группы по всей Европе. Можно связаться с Центром по e-mail, training-support@nesc.ac.uk, или узнать подробности на сайте www.egee.nesc.ac.uk

Какие приложения уже доступны?

С самого начала в проект EGEE были включены биомедицинские приложения. Промежуточное программное обеспечение EGEE будет создано на основе стандартных технологий веб-сервисов, которые хорошо знакомы разработчикам приложений для биокомпьютерного и широко используются ими.

- **GATE:** планирование радиотерапии и приложение для медицинской томографии. См. сайт <http://www-lphe.epfl.ch/~PET/research/gate/>
- **CPS@:** анализ последовательностей белков на основе Грид-технологии, в частности – blast, psi-blast, fasta, search, clustal, proscan, предсказание вторичной структуры, SRS. См. сайт <http://gpsa.ibcp.fr/>
- **CDSS:** поддержка принятия решений в клинической практике. Назначение CDSS – поиск информации для практикующих врачей, ведущийся в огромных объёмах ресурсов. Медицинская база знаний создана на основе данных, доступных в региональной больничной сети Валенсии. См. сайт <http://egee-na4.ct.infn.it/biomed/20040608/NA4>
Biomed Meeting - Pilots CDSS (UPV).ppt
- **PTM3D:** интерактивная визуализация и обработка радиологических изображений. PTM3D предназначен для интерактивной работы в Грид-системе с задачами, связанными с поиском и анализом медицинских данных.
- **Mammogrid:** цель этого проекта – ориентируясь на развитие Грид-технологии, создать всеевропейскую базу данных маммограмм, которые понадобятся для разработки важных медицинских приложений, а также оценить возможности этой Грид-системы в плане повышения эффективности сотрудничества специалистов в области медицины и здравоохранения во всём Европейском союзе. См. сайт <http://mammogrid.vitamib.com/>

EGEE приветствует и готов немедленно рассмотреть и другие приложения. Бланк заявления можно взять на сайте <http://egee-na4.ct.infn.it/biomed/process.html>