

Гридовская ферма Софийского университета

доц., д-р Вл. Димитров,
докторант Р. Горанова
Факультет математики и информатики
Софийского университета
"Св. Климент Охридски"
e-mail: cht@fmi.uni-sofia.bg

Резюме

Гридовские технологии определяют характер современного метакомпьютинга (распределение вычисления и распределение информации). Разработка и стандартизация этих технологий осуществляется в университетах, в научных и промышленных институтах. Уже видны структуры будущего Интернета, основывающегося на Грид-технологиях. Продуктовые линии (железа и софтвера) передовых фирм ориентированы на поставку Гридовых компонентов. В этом контексте особо важно вести адекватное обучение и исследовательскую деятельность в университетах.

В Европе координация по развитию инфраструктуры, разработка и внедрение Грида осуществляется CERN-ом. Современные проблемы в науке требуют больших вычислительных ресурсов и огромного дискового пространства, какие мало университетов и институтов могут себе позволить. Эти требования определяют поиск решения глобального совместного использования ресурсов. Такое решение - Грид.

Резюме

LCG (LHC Computing Grid) – Европейский проект развития вычислительной инфраструктуры для симуляции и анализа данных генерированных LHC (Large Hadron Collider) ускорителя CERN-а. В настоящем докладе рассматриваются место и роль Софийского университета «Св. Климент Охридски» в мировой вычислительной сети (Грид) в обслуживании LHC - экспериментов.

Ключевые слова: Грид, метакомпьютинг, научные вычисления, образование.

Введение

Ученые сегодняшнего дня (физики, биологи, химики) решают сложные проблемы в областях биоинформатики, физики высоких энергий, астрофизики и т.д. Исследования человеческого генома, поиск симметричных частиц, прогноз землетрясений и др. требуют обработки огромного объема информации. Проигрывание сложных моделей можно осуществлять на компьютерных системах, имеющих большую мощность и, по крайней мере, терабайтового дискового пространства. Немного университетов и институтов в состоянии себе позволить такие системы. Поиском глобального решения совместного использования вычислительных ресурсов для научных экспериментов является Грид.

Введение

Грид – это координированное использование ресурсов для решения сложных проблем. Координация осуществляется в контексте виртуальных организаций.

Основной проект по Гриду в Европе - LCG (LHC Computing Grid) [2]. Его назначение - построить вычислительную инфраструктуру для анализа и симуляции данных ,генерированных LHC проектами. LHC (Large Hadron Collider) [1] – это ускоритель тяжелых частиц, который строится в CERN. Детекторы LHC экспериментов будут генерировать петабайты информации.

Введение

Фининсирование LCG осуществляется по проекту EGEE (Enabling Grid for E-science in Europe) [3], цель которого - построить Гридовскую инфраструктуру для научных вычислений научного сообщества европейских стран.

Кафедра «Атомная физика» Физического факультета Софийского университета участвует в LHC проекте CMS (Compact Muon Solenoid) [5]. Дальнейшее участие в этом проекте зависит от наличия Гридовской фермы в университете. Кафедра «Компьютерная информатика» Факультета математики и информатики (ФМИ) совместно с коллегами с Физического факультета стали решать проблему по созданию LCG фермы.

Введение

В 2005 году три специалиста из ФМИ были в долгосрочных командировках в ОИЯИ по Гридовской тематике, другие два специалиста были в CERN по той-же самой тематике. В 2005 году поступило финансирование по договору «Построение и развитие Гридовской инфраструктуры для научных исследований и обучения», № ИО-10/2005 с Фондом «Научные исследования» Министерства образования и науки. Часть необходимой техники была куплена, софтвер инсталлирован, и ферма сертифицирована.

С начало 2006 года ферма Софийского университета включена в Гридовскую сеть и обслуживает в производционном режиме ряд проектов, среди который - LHC эксперименты.

Инсталляция и сертификация фермы

Данные, генерированные LHC экспериментами, будут распределяться на четырех уровнях [4]. На нулевом уровне (Tier-0) необработанные данные будут сохраняться в CERN. После некоторой первоначальной обработки будут поставляться на первый уровень. Первый уровень (Tier-1) состоит из крупных вычислительных комплексов с большими возможностями сохранения терабайтовой информации. Нулевой и первый уровень будут связаны между собой быстрой сетью (2 GB/s, по крайней мере). Центры первого уровня будут поддерживаться в CERN, CNAF (Италия), CCIN2P3 (Франция), PIC (Испания), RAL (Англия). Данные основных LHC экспериментов (ALICE, ATLAS, CMS и LHCb) будут сохраняться на первом уровне.

Инсталляция и сертификация фермы

Первый уровень будет координировать обмен данных со второго уровня. На втором уровне будут расположены вычислительные фермы регионального значения. Они будут обрабатывать данные первого уровня. Фермы такого типа будут расположены в INFN (Италия), NorthGrid (Англия), Russian Tier-2 cluster (Россия). На третьем уровне (Tier-3) будут подключены фермы институтов и университетов. Ферма Софийского университета подключена на третьем уровне.

Фиг. 1. Йерархия ЛНС

Tier 0

CERN

Tier 1

CERN
Tier 1

Italy
Bologna

UK
Didcot

USA
Upton

Germany
Karlsruhe

France
Lyon

Tier 2

Univ. A

Univ. B

Univ. C

Univ. D

Lab. X

Lab. Y

Lab. Z

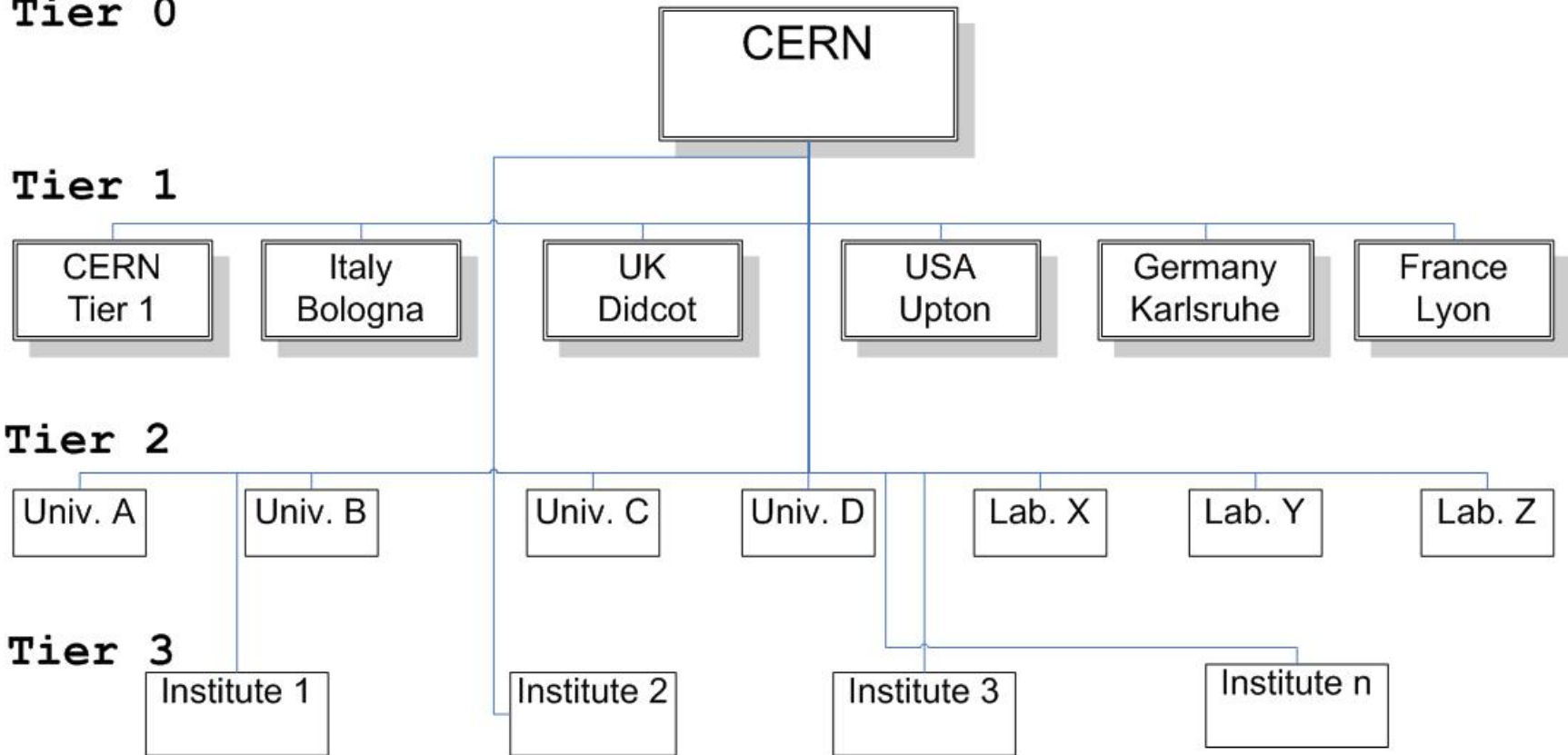
Tier 3

Institute 1

Institute 2

Institute 3

Institute n



Инсталляция и сертификация фермы

На нашей ферме инсталлирована версия LCG-2_7_0 софтвера. Ферма имеет следующие элементы:

- пользовательский интерфейс (UI) – 1 компьютер;
- вычислительный элемент (CE) – 1 компьютер;
- элемент сохранения данных (SE) и мониторинга (MON) – 1 компьютер;
- рабочие узлы (WN) – 6 компьютеров.

Таблица 1. Структура фермы

Hostname	IP	HDD	RAM	Processor	Услуга
ce001.grid.uni-sofia.bg	62.44.127.131	7 GB	384 MB	Intel Dual Pentium II	CE
se001.grid.uni-sofia.bg	62.44.127.132/133	320 GB	1 GB	Intel Dual Xeon	SE/MON
wn001.grid.uni-sofia.bg	62.44.127.135	80 GB	512 MB	AMD Dual Opteron	WN
wn002.grid.uni-sofia.bg	62.44.127.136	160 GB	1 GB	AMD Dual Athlon	WN
wn003.grid.uni-sofia.bg	62.44.127.137	320 GB	1 GB	Intel Dual Xeon	WN
wn004.grid.uni-sofia.bg	62.44.127.138	80 GB	512 MB	Intel Pentium III	WN
wn005.grid.uni-sofia.bg	62.44.127.139	80 GB	512 MB	Intel Pentium III	WN
wn006.grid.uni-sofia.bg	62.44.127.140	60 GB	1 GB	AMD Dual Athlon	WN

Инсталляция и сертификация фермы

Хардвер фермы состоит из следующих компьютеров: два однопроцессорных Pentium III на 1 GHz, два двухпроцессорных Pentium II на 400 MHz, два Xeon на 2.8 GHz, один двухпроцессорный Opteron на 1 GHz и два двухпроцессорных Athlon на 1.7 GHz.

Безопасность инфраструктуры обеспечивается публичными ключами, X.509 сертификатам и SSL протокола. Чтобы пользователи имели доступ к LCG ресурсам, они должны иметь персональный сертификат и регистрацию виртуальной организации (CMS, ALICE, ATLAS, LHCB, DTEAM и OPS).

Инсталляция и сертификация фермы

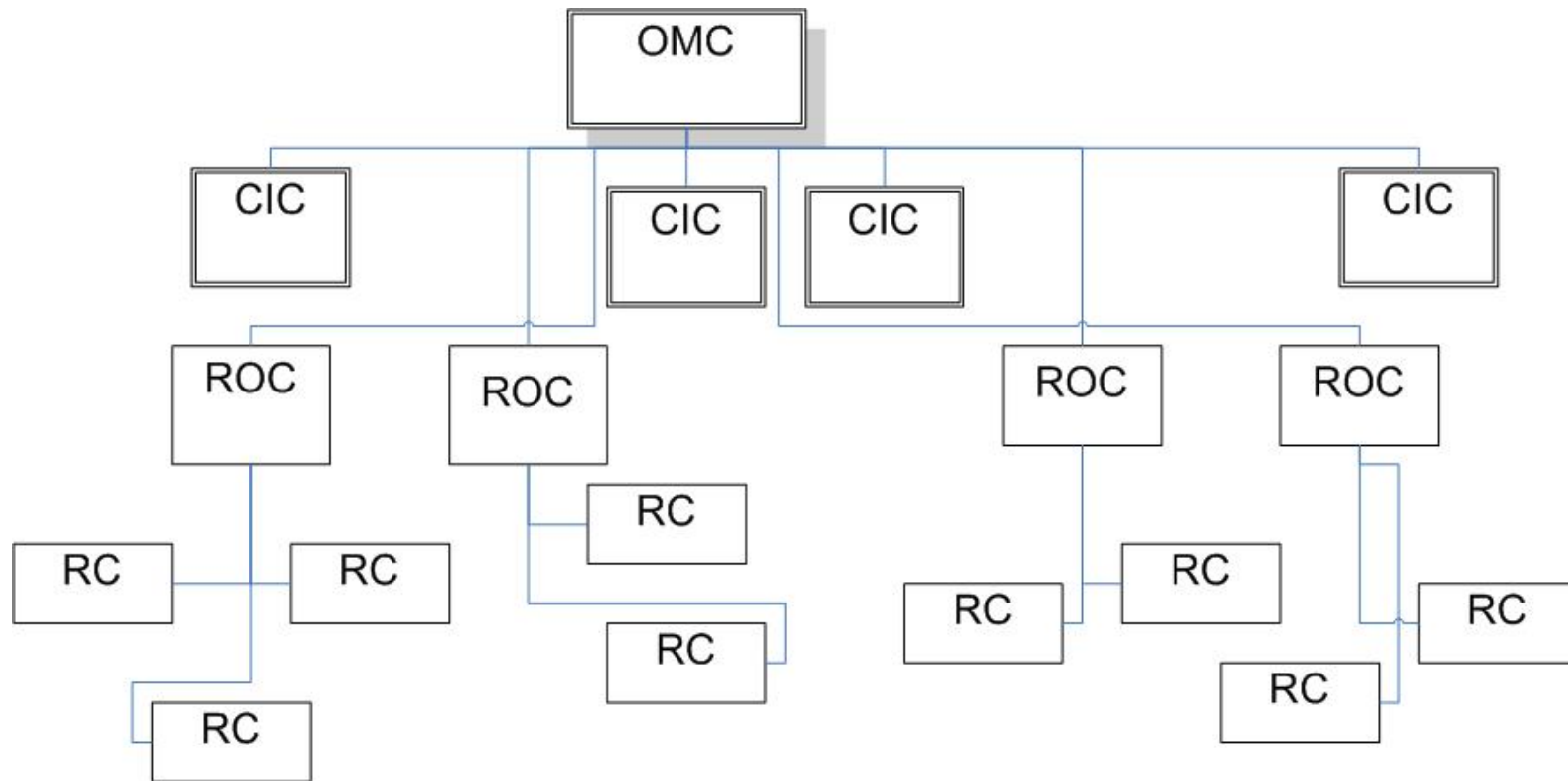
Сертификация пользователей осуществляется локальной сертификационной организацией. Для Болгарии такой организацией является SEE-GRID SA. Региональным оперативным центром Грида в Болгарии является SEE-GRID.

На фиг. 2 изображена иерархия оперативных центров EGEE проекта. Классификация центров в иерархии осуществляется по ролям центров. На самом нижнем уровне находятся Ресурсные центры (RC), которые поставляют вычислительные ресурсы. Ферма Софийского университета является ресурсным центром. Управление ресурсных центров осуществляется Региональными оперативными центрами (ROC). Они способствуют разрешению проблем, возникших в ресурсных центрах.

Инсталляция и сертификация фермы

В Болгарии функции Регионального оперативного центра осуществляются в ИПОИ, БАН. В Юго-восточном регионе Европы Региональным оперативным центром является SEE-GRID. На более высших уровнях находятся Инфраструктурный (СIC) и Управленческие (ОМС) центры. Они отвечают за улучшение и управление Гридovской инфраструктурой.

Фиг. 2. Иерархия оперативных центров

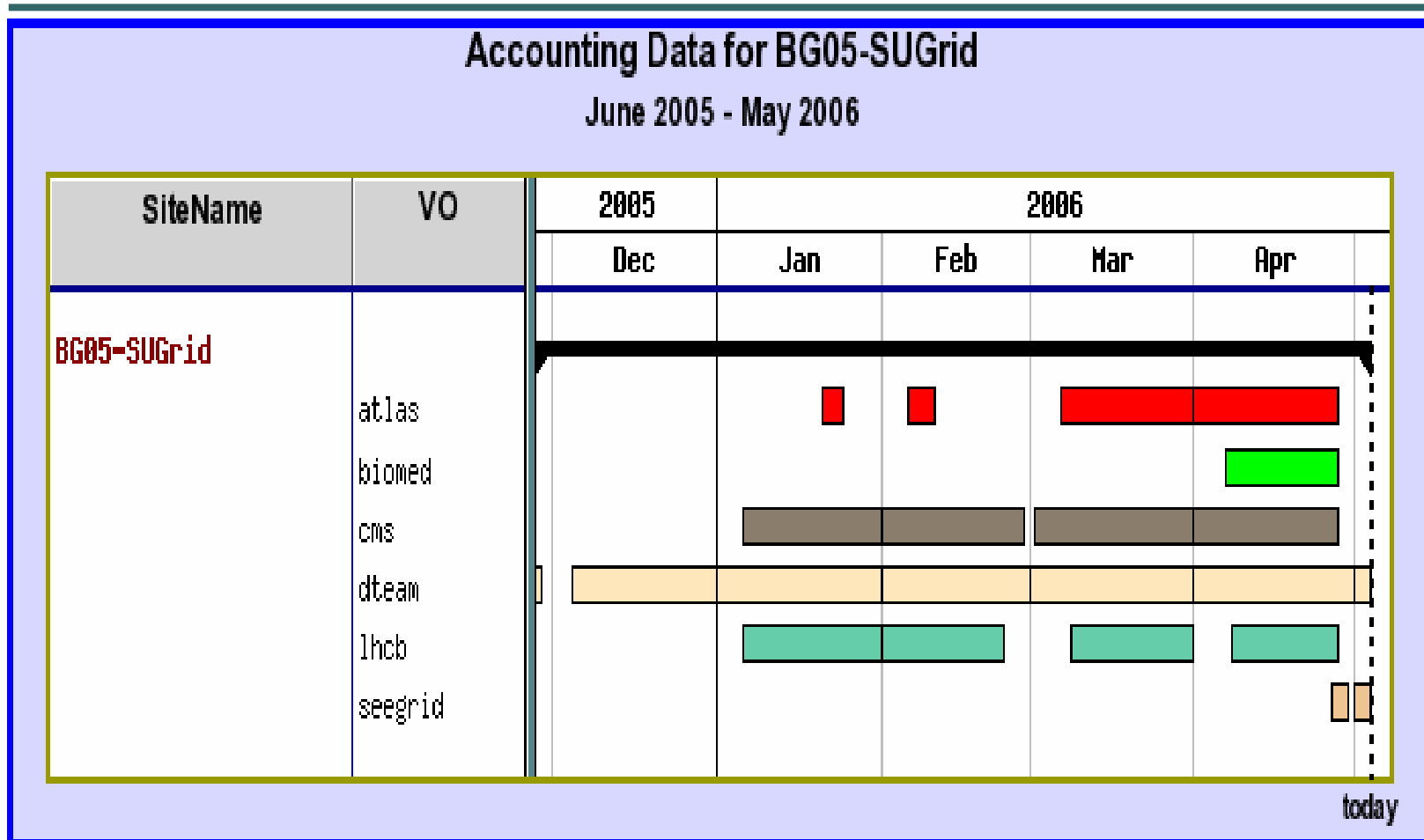


Инсталляция и сертификация фермы

Регистрация пользователей осуществляется на LCG Registration Server в CERN. Это единственный сервер этого типа во всей LCG вычислительной сети. Для регистрации сайта (фермы) по требованиям EGEE процедуры регистрации необходима следующая информация: название института (университета), электронный адрес и телефонный номер администратора фермы, электронный адрес и телефонный номер ответственного за безопасность фермы.

Региональный оперативный центр провел тесты функциональности каждого элемента фермы. Тесты проводились безостановочно в течении двух недель. Ферма успешно переходила из состояния Candidate site в Uncertified site и, наконец, в Certified site.

Фиг. 3. Использование фермы



Инсталляция и сертификация фермы

Теперь ферма сертифицирована и работает в производственном режиме. На ней решаются задачи проектов LHCb, Atlas, CMS, Biomed и SEE-GRID.

Информацию о ферме и статистику ее работы можно увидеть на адресах:

<http://goc.grid.sinica.edu.tw/gstat/BG05-SUGrid>

http://mon.egee-see.org/gridice/site/site_details.php?siteName=BG05-SUGrid

Заключение

Ферма Софийского университета - уже реальность.

Статистика ее работы показывает, что сначала надо повысить скорость связи с GEANT, так как некоторые задачи не могут успешно передать свои данные.

Улучшение связи с другой стороной потребует увеличения вычислительных и дисковых ресурсов.

Ферма дает возможность исследователям физического факультета полностью перейти на Гридовский софтвер. Ферма позволит на ФМИ с начала следующего года вести обучение по Гриду на реальной базе. До сих пор мы это делали на локальной gLite ферме.

Литература

1. The Large Hadron Collider Project <http://lhc.web.cern.ch/lhc>
2. LCG <http://lcg.web.cern.ch/LCG/activities/middleware.html>
3. EGEE <http://public.eu-egee.org/intro/>
4. LCG Tiers
<http://lcg.web.cern.ch/LCG/operations/regionalcenters.html>
5. CMS <http://cms.cern.ch>

Благодарю за внимание!