GRID Activity at Russia and JINR

Korenkov Vladimir (LIT,JINR)

MMCP-2009, Dubna, 07.07.09
Today there are many definitions of Grid computing:

The definitive definition of a Grid is provided by Ian Foster in his article "What is the Grid? A Three Point Checklist"

The three points of this checklist are:

- Computing resources are not administered centrally;
- Open standards are used;
- Non-trivial quality of service is achieved.
Tier 0 at CERN: Acquisition, First pass reconstruction, Storage & Distribution

LHCb ~ 50 MB/sec
ATLAS ~ 320 MB/sec
ALICE ~ 100 MB/sec
CMS ~ 220 MB/sec

1.25 GB/sec (ions)
Tier 0 – Tier 1 – Tier 2

Tier-0 (CERN):
• Data recording
• Initial data reconstruction
• Data distribution

Tier-1 (11 centres):
• Permanent storage
• Re-processing
• Analysis

Tier-2 (>200 centres):
• Simulation
• End-user analysis
Some history

1999 – Monarc Project
  – Early discussions on how to organise distributed computing for LHC

2001-2003 - EU DataGrid project
  – middleware & testbed for an operational grid

  – deploying the results of DataGrid to provide a production facility for LHC experiments

2004-2006 – EU EGEE project phase 1
  – starts from the LCG grid
  – shared production infrastructure
  – expanding to other communities and sciences

2006-2008 – EU EGEE-II
  – Building on phase 1
  – Expanding applications and communities …

2008-2010 – EU EGEE-III

http://www.egee-rdig.ru
Purpose

- Develop, build and maintain a distributed computing environment for the storage and analysis of data from the four LHC experiments
  - Ensure the computing service
  - ... and common application libraries and tools

- **Phase I** – 2002-05 - Development & planning

- **Phase II** – 2006-2008 – Deployment & commissioning of the initial services
LCG depends on two major science grid infrastructures

EGEE - Enabling Grids for E-Science
OSG - US Open Science Grid

http://www.egee-rdig.ru
The Map of OSG Sites (in the US)
The aim of the project is to create a global Pan-European computing infrastructure of a Grid type.

- Integrate regional Grid efforts
- Represent leading grid activities in Europe

10 Federations, 27 Countries, 70 Organizations
Objectives

- Flagship European grid infrastructure project
- Now in 2nd phase with 91 partners in 32 countries

- Large-scale, production-quality grid infrastructure for e-Science
- Attracting new resources and users from industry as well as science
- Maintain and further improve gLite Grid middleware

http://www.egee-rdig.ru
Flagship Grid infrastructure project co-funded by the European Commission

Main Objectives

- Expand/optimise existing EGEE infrastructure, include more resources and user communities
- Prepare migration from a project-based model to a sustainable federated infrastructure based on National Grid Initiatives

Duration: 2 years
Consortium: ~140 organisations across 33 countries
EC co-funding: 32Million €
Infrastructure operation

- Currently includes >350 sites across 55 countries
- Continuous monitoring of grid services & automated site configuration/management
- Support ~300 Virtual Organisations from diverse research disciplines

• Middleware
  - Production quality middleware distributed under business friendly open source licence

• User Support - Managed process from first contact through to production usage
  - Training
  - Expertise in grid-enabling applications
  - Online helpdesk
  - Networking events (User Forum, Conferences etc.)

http://www.egee-rdig.ru
• >300 VOs from several scientific domains
  – Astronomy & Astrophysics
  – Civil Protection
  – Computational Chemistry
  – Comp. Fluid Dynamics
  – Computer Science/Tools
  – Condensed Matter Physics
  – Earth Sciences
  – Fusion
  – High Energy Physics
  – Life Sciences
• Further applications under evaluation

Applications have moved from testing to routine and daily usage
~80-95% efficiency

http://www.egee-rdig.ru
Collaborating e-Infrastructures

Potential for linking ~80 countries by 2008

http://www.egee-rdig.ru
Grid - projects around EGEE

Infrastructures
geographical or thematic coverage

Applications
improved services for academia, industry and the public

Support Actions
key complementary functions

http://www.egee-rdig.ru
Archeology
Astronomy
Astrophysics
Civil Protection
Comp. Chemistry
Earth Sciences
Finance
Fusion
Geophysics
High Energy Physics
Life Sciences
Multimedia
Material Sciences
...
The Future of Grids

- From e-Infrastructures to Knowledge Infrastructures

- Network infrastructure connects computing and data resources and allows their seamless usage via Grid infrastructures

- Federated resources and new technologies enable new application fields:
  - Distributed digital libraries
  - Distributed data mining
  - Digital preservation of cultural heritage
  - Data curation

→ Knowledge Infrastructure
Major Opportunity for Academic and Businesses alike

http://www.egee-rdig.ru

Erwin Laure – gLite and Business – EGEE’08 BT – 22 Sept 2008
The Russian consortium RDIG (Russian Data Intensive Grid, was set up in September 2003 as a national federation in the EGEE project.

**IHEP** - Institute of High Energy Physics (Protvino),  
**IMPB RAS** - Institute of Mathematical Problems in Biology (Pushchino),  
**ITEP** - Institute of Theoretical and Experimental Physics  
**JINR** - Joint Institute for Nuclear Research (Dubna),  
**KIAM RAS** - Keldysh Institute of Applied Mathematics  
**PNPI** - Petersburg Nuclear Physics Institute (Gatchina),  
**RRC KI** - Russian Research Center “Kurchatov Institute”  
**SINP-MSU** - Skobeltsyn Institute of Nuclear Physics, MSU,
Now the RDIG infrastructure comprises 15 Resource Centers with more 7000 kSI2K CPU and more 1850 TB of disc storage.

- RDIG Resource Centres:
  - ITEP
  - JINR-LCG2
  - Kharkov-KIPT
  - RRC-KI
  - RU-Moscow-KIAM
  - RU-Phys-SPbSU
  - RU-Protvino-IHEP
  - RU-SPbSU
  - Ru-Troitsk-INR
  - ru-IMPB-LCG2
  - ru-Moscow-FIAN
  - ru-Moscow-GCRAS
  - ru-Moscow-MEPHI
  - ru-PNPI-LCG2
  - ru-Moscow-SINP

http://www.egee-rgid.ru
The main directions in development and maintenance of RDIG e-infrastructure are as the following:

- support of basic grid-services;
- Support of Regional Operations Center (ROC);
- Support of Resource Centers (RC) in Russia;
- RDIG Certification Authority;
- RDIG Monitoring and Accounting;
- participation in integration, testing, certification of grid-software;
- support of Users, Virtual Organization (VO) and application;
- User & Administrator training and education;
- Dissemination, outreach and Communication grid activities.
Portal www.egee-rdig.ru
Карта сайта

* Пользователям
  * Получение нового сертификата
  * Перерегистрация сертификата в виртуальной организации
  * Корневой сертификат RDIG CA для загрузки в браузер
  * Корневой сертификат RDIG CA в формате PEM
  * Список действительных сертификатов
  * Список отозванных сертификатов (CRL) для загрузки в браузер
  * Список отозванных сертификатов (CRL) в формате PEM
  * Политика выдачи сертификатов
  * Рекомендации по работе с вашим сертификатом
  * Нерешенные вопросы. Может быть ваша проблема уже известна и решена — вглядите.

* Организациям и институтам
  * Добавление нового Registration Authority для вашей организации.
• **Infrastructure VO's (all RC's):**
  – dteam
  – ops

• **Most RC support the WLCG/EGEE VO's**
  – Alice
  – Atlas
  – CMS
  – LHCb

• **Supported by some RC's:**
  – gear
  – Biomed
  – Fusion

• **Regional VO's**
  – Ams, eearth, photon, rdteam, rgstest

*Flagship applications:*
  - LHC, Fusion (toward to ITER), nanotechnology

*Current interests from: medicine, engineering*
RDIG monitoring & accounting
http://rocmon.jinr.ru:8080

- Monitoring – allows to keep an eye on parameters of Grid sites' operation in real time
- Accounting - resources utilization on Grid sites by virtual organizations and single users

http://www.egee-rdig.ru
Production Normalised CPU time per EGEE Region (2008 and June 2009)

PRODUCTION Normalised CPU time per REGION
LHC VO's. December 2008 - July 2009

PRODUCTION Normalised CPU time per REGION
LHC VO's. June 2009 - June 2009
Production Normalised CPU time per Country (Dec 2008 - June 2009)

COUNTRY Normalised CPU time per COUNTRY
LHC VO's. December 2008 - June 2009

- Canada: 8.0%
- Denmark: 1.2%
- France: 2.2%
- Germany: 15.5%
- Italy: 4.8%
- Netherlands: 11.7%
- Poland: 1.3%
- Romania: 2.0%
- Russia: 2.4%
- Slovenia: 11.2%
- Spain: 2.9%
- Switzerland: 6.3%
- United Kingdom: 1.2%
- United States of America: 23.9%
- Others: 1.2%

(C) CESGA 'EGEE View': COUNTRY / normcpu / 2008:12 - 2009:6 / COUNTRY-VO / lhc (x) / ACCBAR-LIN / i
2009-07-06 10:08 UTC

http://www.egee-rdig.ru
Russia and JINR Normalized CPU time per SITE (December 2008 - June 2009)

Normalised CPU time (SpectInt2000*hour = 1000) per Site

http://www.egee-rdig.ru
<table>
<thead>
<tr>
<th>Site</th>
<th>Normalised CPU Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>FZK</td>
<td>8,640,981</td>
</tr>
<tr>
<td>GRIF</td>
<td>6,413,585</td>
</tr>
<tr>
<td>IN2P3-CC-T2</td>
<td>6,274,114</td>
</tr>
<tr>
<td>IN2P3-CC</td>
<td>5,869,444</td>
</tr>
<tr>
<td>NIKHEF</td>
<td>5,479,998</td>
</tr>
<tr>
<td>TRIUMF</td>
<td>5,342,865</td>
</tr>
<tr>
<td>RAL</td>
<td>4,007,746</td>
</tr>
<tr>
<td>JINR</td>
<td>3,970,242</td>
</tr>
</tbody>
</table>

http://www.egee-rdig.ru
The protocol between CERN, Russia and JINR on a participation in LCG Project has been approved in 2003.

The tasks of the Russian institutes in the LCG have been defined as:

- LCG software testing;
- evaluation of new Grid technologies (e.g. Globus toolkit 3) in a context of using in the LCG;
- event generators repository, data base of physical events: support and development.
The tasks of the Russian institutes & JINR in the LCG (2008 years):

• Task 1. MW (gLite) Testsuit
  (supervisor O. Keeble)
• Task 2. LCG vs Experiments
  (supervisor I. Bird)
• Task 3. LCG monitoring
  (supervisor J. Andreeva)
• Task 4/5. Genser/ MCDB
  (supervisor A. Ribon)
A necessary level of all the elements of the JINR telecommunication, network and information infrastructure should be provided:

- High-throughput telecommunication data links,
- JINR local area network (LAN) backbone,
- Central computer complex and Grid segment,
- Software support of the LHC experiments.
Collaboration with RSCC, RIPN, MSK-IX, JET Infosystems, Nortel

http://www.egee-rdig.ru
JINR Central Information and Computing Complex (CICC)

http://www.egee-rdig.ru
**In 2008**, total CICC performance was 1400 kSI2K, the disk storage capacity 100 TB.

**At present**, the CICC performance equals **2300 kSI2K** and the disk storage capacity **400 TB**.

<table>
<thead>
<tr>
<th></th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU (kSI2K)</td>
<td>1250</td>
<td>1750</td>
<td>2500</td>
</tr>
<tr>
<td>Disk (TB)</td>
<td>400</td>
<td>800</td>
<td>1200</td>
</tr>
<tr>
<td>Tapes (TB)</td>
<td>-</td>
<td>100</td>
<td>200</td>
</tr>
</tbody>
</table>

**SuperBlade – 5 BOX**
80 CPU Xeon 5430 2.66 GHz Quad Core (two with InfiniBand)

[http://www.egee-rdig.ru](http://www.egee-rdig.ru)
CICC comprises:

- 65 servers
- 4 interactive nodes
- 960 computing nodes, Xeon 5150, 8GB RAM (GEthernet)
- 160 computing nodes, Xeon X5450, 16GB RAM, InfiniBand.

Site name: JINR-LCG2
Internal CICC network – 1Gbit/sec
Operating system - Scientific Linux CERN 4.6;
Middleware version GLITE-3.1
File Systems – AFS (the Andrew File System) for user
Software and home directories is a world-wide distributed file system. AFS permits to share easily files in an heterogeneous distributed environment (UNIXes, NT) with a unique authentication scheme (Kerberos).
- dCache- for data.
User registration system - Kerberos 5 (AFS use Kerberos 5 for authentication)
JINR provides the following services in the WLCG environment:

Basic services:
- Berkley DB Information Index (top level BDII);
- site BDII;
- 2 x Computing Element (CE);
- Proxy Server (PX);
- 2 x Workload Management System (WMS);
- Logging&Bookkeeping Service (LB);
- RGMA-based monitoring system collector server (MON-box);
- LCG File Catalog (LFC);
- Storage Element (SE), dCache 400 TB, 4 x gridftp door, 14 x pool;
- 4 x User Interface (UI), installed in AFS.

Special Services - VO boxes for ALICE and for CMS; ROCMON;
PPS and testing infrastructure - Pre-production gLite version;

Software for VOs: dCache xrootd door, AliROOT, ROOT, GEANT packages for ALICE; ATLAS packages; CMSSW packages for CMS and DaVinchi, Gauss packages for LHCb.
support and development WLCG/EGEE infrastructure;

participation in middleware testing/evaluation,

participation in Data and Service Challenges,

grid monitoring and accounting system development;

FTS-monitoring and testing

MCDB development;

Participation in ARDA activities in coordination with experiments;

HEP applications;

User & Administrator Training and Education

support of JINR member states in the GRID activities.

http://www.egee-rdig.ru
User Training and Induction

Russian and JINR physicists participants of ATLAS experiment train and practise with Grid and the GANJA

http://www.egee-rdig.ru
Worldwide LHC Computing Grid (WLCG);
Enabling Grids for E-science (EGEE);
RDIG Development (Project of FASI)
CERN-RFBR project “GRID Monitoring from VO perspective”
BMBF grant “Development of the GRID-infrastructure and tools to provide joint investigations performed with participation of JINR and German research centers’
“Development of Grid segment for the LHC experiments” was supported in frames of JINR-South Africa cooperation agreement;
NATO project "DREAMS-ASIA“ (Development of gRid EnAbling technology in Medicine&Science for Central ASIA);
JINR-Romania cooperation Hulubei-Meshcheryakov programme
Project "SKIF-GRID" (Program of Belarussian-Russian Union State).
Project GridNNN (National Nanotechnological Net)
We work in close cooperation and provide support to our partners in Armenia, Belarus, Bulgaria, Czech Republic, Georgia, Germany, Poland, Romania, South Africa, Ukraine, Uzbekistan.

http://www.egee-rdig.ru
Useful References:

- OPEN GRID FORUM:  //www.ogf.org
- GLOBUS:     //www.globus.org
- TERAGRID:   //www.teragrid.org
- LCG:        //lcg.web.cern.ch/LCG/
- EGEE:       //www.eu-egee.org
- EGEE-RDIG:  //www.egee-rdig.ru
- GRIDCLUB:   //www.gridclub.ru
The blind men and the elephant in the room

Cyberinfrastructure

SaaS

Shared Infrastructure/
Shared Services

SOA

Web 2.0

Grids

Automation

Virtualization

http://www.egee-rdig.ru