The LHC Computing Grid Project

LCG Asia Workshop
ASCC, Taipei – 26 July 2004

Les Robertson – LCG Project Leader
CERN – European Organization for Nuclear Research
Geneva, Switzerland
les.robertson@cern.ch
Outline

- What is LCG and why
- Status
  - Applications support
  - The LCG grid service
  - Middleware
  - Fabric & Network
  - ARDA - distributed physics analysis
- LCG timeline
- Conclusion
This is reduced by online computers that filter out a few hundred “good” events per sec.

The LHC accelerator —
• the largest superconducting installation in the world
• 27 kilometres of magnets cooled to – 300° C
• colliding proton beams at an energy of 14 TeV

Which are recorded on disk and magnetic tape at 100-1,000 MegaBytes/sec ~15 PetaBytes per year
CERN Collaborators

CERN has over 6,000 users from more than 450 institutes from around the world

Europe:
- 267 institutes
- 4603 users

Elsewhere:
- 208 institutes
- 1632 users

LHC Computing → uniting the computing resources of particle physicists in the world!
**Aim of the project**
To prepare, deploy and operate the **computing environment** for the **experiments** to **analyse the data from** the LHC detectors

**Applications development environment, common tools and frameworks**

**Build and operate the LHC computing service**

*The Grid is just a tool towards achieving this goal*
LHC Computing Grid Project – a Collaboration

Building and operating the LHC Grid – a collaboration between

- The physicists and computing specialists from the LHC experiment
- The projects in Europe and the US that have been developing Grid middleware
- The regional and national computing centres that provide resources for LHC
- The research networks
Applications Area Projects

- **Software Process and Infrastructure (SPI)** (A.Aimar)
  - Librarian, QA, testing, developer tools, documentation, training, ...
- **Persistency Framework & Database Applications (POOL)** (D.Duellmann)
  - Relational persistent data store, conditions database, collections
- **Core Tools and Services (SEAL)** (P.Mato)
  - Foundation and utility libraries, basic framework services, object dictionary and whiteboard, maths libraries
- **Physicist Interface (PI)** (V.Innocente)
  - Interfaces and tools by which physicists directly use the software. Interactive analysis, visualization
- **Simulation** (T.Wenaus)
  - Generic framework, Geant4, FLUKA integration, physics validation, generator services
- **ROOT** (R.Brun)
  - ROOT I/O event store; analysis package
POOL – Object Persistency

- Bulk event data storage – an object store based on ROOT I/O
  - Full support for persistent references automatically resolved to objects anywhere on the grid
  - Recently extended to support updateable metadata as well (with some limitations)

- File cataloging – Three implementations using –
  - Grid middleware (EDG version of RLS)
  - Relational DB (MySQL)
  - Local Files (XML)

- Event metadata –
  - Event collections with query-able metadata (physics tags etc.)

- Transient data cache –
  - Optional component by which POOL can manage transient instances of persistent objects

- POOL project scope now extended to include the Conditions Database
POOL Component Breakdown

POOL API
- Storage Service
  - ROOT I/O Storage Svc
  - RDBMS Storage Svc
- FileCatalog
  - XML Catalog
- Collections
  - Explicit Collection
  - Implicit Collection
  - EDG Replica Location Service
Simulation Project Organisation

Simulation Project Leader

Subprojects

- Framework
  - Generic interface to multiple simulation engines (G4, FLUKA), building on existing ALICE work (VMC)

- Geant4
  - Part of the GEANT4 collaboration Development aligned with and responding to needs from LHC experiments

- FLUKA integration
  - FLUKA team participating in framework integration, physics validation

- Physics Validation
  - Assess adequacy of simulation and physics environment for LHC, provide the focus for the LHC requirements

- Generator Services
  - Generator librarian, common event files, validation/test suite, development when needed (HEPMC, etc.)
LHC Computing Model (simplified!!)

Tier-0 – the accelerator centre
- Filter → raw data
- Reconstruction → summary data (ESD)
- Record raw data and ESD
- Distribute raw and ESD to Tier-1

Tier-1 –
- Managed Mass Storage – permanent storage raw, ESD, calibration data, meta-data, analysis data and databases → grid-enabled data service
- Data-heavy analysis
- Re-processing raw → ESD
- National, regional support

“online” to the data acquisition process high availability, long-term commitment

Tier-2 –
- Well-managed disk storage – grid-enabled
- Simulation
- End-user analysis – batch and interactive
- High performance parallel analysis (PROOF)
## Current estimates of Computing Resources needed at Major LHC Centres

*First full year of data - 2008*

<table>
<thead>
<tr>
<th></th>
<th>Processing M SI2000**</th>
<th>Disk PetaBytes</th>
<th>Mass Storage PetaBytes</th>
</tr>
</thead>
<tbody>
<tr>
<td>CERN</td>
<td>20</td>
<td>5</td>
<td>20</td>
</tr>
<tr>
<td>Major data handling centres (Tier 1)</td>
<td>45</td>
<td>20</td>
<td>18</td>
</tr>
<tr>
<td>Other large centres (Tier 2)</td>
<td>40</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td><strong>105</strong></td>
<td><strong>37</strong></td>
<td><strong>43</strong></td>
</tr>
</tbody>
</table>

**Current fast processor ~1K SI2000**
The LCG Service

- Service opened on 15 September 2003 - with 12 sites
- Middleware package - components from
  - European DataGrid (EDG)
  - US Virtual Data Toolkit (Globus, Condor, PPDG, iVDGL, GriPhyN)
- About 30 sites by the end of the year
- Upgraded version of the grid software (LCG-2) in February 2004
- Additional VOs being added for other sciences as part of the EGEE project
- Grid Operations Centres at Rutherford Lab (UK) and ASCC (Taiwan)
- User Support Centres at ASCC and Forschunszentrum Karlsruhe
The LCG Service

July 2004 - 64 sites ~ 6,000 processors

Number of LCG Sites
64
Mon July 12 2004
<table>
<thead>
<tr>
<th>Country</th>
<th>Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>U-Innsbruck</td>
</tr>
<tr>
<td>Canada</td>
<td>Triumf</td>
</tr>
<tr>
<td></td>
<td>Alberta</td>
</tr>
<tr>
<td></td>
<td>Carleton</td>
</tr>
<tr>
<td></td>
<td>Montreal</td>
</tr>
<tr>
<td></td>
<td>Toronto</td>
</tr>
<tr>
<td>China</td>
<td>IHEP, Beijing</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>Prague-FZU, Prague-CESNET</td>
</tr>
<tr>
<td>France</td>
<td>CC-IN2P3</td>
</tr>
<tr>
<td></td>
<td>Clermont-Ferrand</td>
</tr>
<tr>
<td>Germany</td>
<td>FZK</td>
</tr>
<tr>
<td></td>
<td>Aachen</td>
</tr>
<tr>
<td></td>
<td>DESY</td>
</tr>
<tr>
<td></td>
<td>Wuppertal</td>
</tr>
<tr>
<td>Greece</td>
<td>HellasGrid</td>
</tr>
<tr>
<td>Hungary</td>
<td>Budapest (KFK)</td>
</tr>
<tr>
<td>Israel</td>
<td>Tel-Aviv</td>
</tr>
<tr>
<td></td>
<td>Weizmann</td>
</tr>
<tr>
<td>Italy</td>
<td>CNAF</td>
</tr>
<tr>
<td></td>
<td>Frascati</td>
</tr>
<tr>
<td></td>
<td>Legnaro</td>
</tr>
<tr>
<td></td>
<td>Milano</td>
</tr>
<tr>
<td></td>
<td>Napoli</td>
</tr>
<tr>
<td></td>
<td>Roma</td>
</tr>
<tr>
<td></td>
<td>Torino</td>
</tr>
<tr>
<td>Japan</td>
<td>Tokyo University (ICEPP)</td>
</tr>
<tr>
<td>Netherlands</td>
<td>NIKHEF</td>
</tr>
<tr>
<td>Pakistan</td>
<td>NCP</td>
</tr>
<tr>
<td>Poland</td>
<td>CYFRONET (Krakow)</td>
</tr>
<tr>
<td>Portugal</td>
<td>LIP</td>
</tr>
<tr>
<td>Russia</td>
<td>SINEP-Moscow</td>
</tr>
<tr>
<td></td>
<td>JINR-Dubna</td>
</tr>
<tr>
<td>Spain</td>
<td>PIC</td>
</tr>
<tr>
<td></td>
<td>Univ. Autonoma de Madrid</td>
</tr>
<tr>
<td></td>
<td>Univ. Santiago de Compostela</td>
</tr>
<tr>
<td></td>
<td>Univ. Barcelona</td>
</tr>
<tr>
<td></td>
<td>IFCA Santander</td>
</tr>
<tr>
<td></td>
<td>CIEMAT Madrid</td>
</tr>
<tr>
<td></td>
<td>IFIC Valencia</td>
</tr>
<tr>
<td>Switzerland</td>
<td>CERN</td>
</tr>
<tr>
<td></td>
<td>CSCS</td>
</tr>
<tr>
<td>Taiwan</td>
<td>ASCC</td>
</tr>
<tr>
<td></td>
<td>NCU</td>
</tr>
<tr>
<td>UK</td>
<td>RAL</td>
</tr>
<tr>
<td></td>
<td>RAL-Physics</td>
</tr>
<tr>
<td></td>
<td>Universities:</td>
</tr>
<tr>
<td></td>
<td>Birmingham</td>
</tr>
<tr>
<td></td>
<td>Cambridge</td>
</tr>
<tr>
<td></td>
<td>Glasgow</td>
</tr>
<tr>
<td></td>
<td>Imperial</td>
</tr>
<tr>
<td></td>
<td>Lancaster</td>
</tr>
<tr>
<td></td>
<td>Manchester</td>
</tr>
<tr>
<td></td>
<td>Queen Mary</td>
</tr>
<tr>
<td></td>
<td>Sheffield</td>
</tr>
<tr>
<td></td>
<td>UCL</td>
</tr>
<tr>
<td>US</td>
<td>BNL</td>
</tr>
<tr>
<td></td>
<td>FNAL</td>
</tr>
<tr>
<td>Hewlett-Packard</td>
<td>Puerto Rico</td>
</tr>
</tbody>
</table>

In preparation/certification –
- India – Tata Institute, Mumbai
- New Zealand
LCG-2 for the 2004 Data Challenges

- Large-scale tests of the experiments’ computing models, processing chains, grid technology readiness, operating infrastructure

- ALICE and CMS data challenges started at the beginning of March, LHCb in May, ATLAS in July

- The big challenge for this year - data & storage -
  - file catalogue,
  - replica management,
  - database access,
  - integrating mass storage ..
    .. and networking
Service Challenges for LCG-2

Confronting the practical issues of setting up a service

- Exercise the operations and support infrastructure
  - Gain experience in service management
  - Uncover problems with long-term operation
  - Explore grid behaviour under load
  - Exercise capability to respond to security incidents, infrastructure failure
  - Develop long-term fixes - not workarounds

- Focus on
  - Data management, batch production and analysis
  - Reliable data transfer
  - Integration of high bandwidth networking
  - Operation with minimal human intervention

- Target by end 2004 -
  - Robust and reliable data management services in continuous operation between CERN, Tier-1 and large Tier-2 centres
  - Sufficient experience with sustained high performance data transfer to guide wide area network planning

- The Service Challenges are a complement to the experiment Data Challenges
Basics

- Getting the data from the detector to the grid requires sustained data collection and distribution -- keeping up with the accelerator
- To achieve the required levels of performance, reliability, resilience -- at minimal cost (people, equipment)
  -- we also have to work on scalability and performance of some of the basic computing technologies –
  - cluster management
  - mass storage management
  - high performance networking
Tens of thousands of disks
Thousands of processors
Hundreds of tape drives
Continuous evolution

Sustained throughput
Resilient to problems
Fabric Automation at CERN

Fault & hardware Management

Configuration Installation

Includes technology developed by DataGrid

Node

Monitoring

LEMON

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid

Configuration Installation

Fault & hardware Management

Fabric Automation at CERN

includes technology developed by DataGrid
CERN and Caltech join forces to smash Internet speed record

CERN* and California Institute of Technology (Caltech) will tomorrow receive an award for transferring over a Terabyte of data across 7,000 km of network at 5.44 gigabits per second (Gbps), smashing the old record of 2.38 Gbps achieved in February between CERN in Geneva and Sunnyvale in California by a Caltech, CERN, Los Alamos National Laboratory and Stanford Linear Accelerator Center team.

The international CERN-Caltech team set this new Internet2® Land Speed Record on 1 October 2003 by transferring 1.1 Terabytes of data in less than 30 minutes, corresponding to 38,420.54 petabits-metres per second. The average rate of 5.44 Gbps is more than 20,000 times faster than a typical home broadband connection and is equivalent to transferring a full CD in 1 second or a full length DVD movie in approximately 7 seconds. The award will be made to Olivier Martin of CERN and Harvey Newman of Caltech on the Lake Geneva Region Stand at the ITU Telecom World event in Geneva live from the Internet2 conference in Indianapolis at 17.30 CET on Thursday 16 October.

We now have to get from an R&D project (DATATAG) to a sustained, reliable service – Asia, Europe, US
From HEP grids to multi-science grids

EGEE: Enabling Grids for E-science in Europe
EU 6th Framework Project

- Create a Grid for European Science
- Supporting many application domains with one large-scale infrastructure
- Providing “round-the-clock” access to major computing resources, independent of geographic location

- Emphasis on grid deployment (rather than development)
- Leverages national and regional Grid programmes
  - building on the results of existing projects
  - National Research Networks and the EU Research Network Backbone “Geant”
EGEE: Partners

- 70 institutions in 27 countries organised into regional federations
EGEE and LCG - Fusion for Evolution

LCG will is working very closely with EGEE

- EGEE has started by using the LCG service
  - One operations manager for both projects

- EGEE will provide the basic middleware for LCG
  - One middleware manager for both projects

- Involvement of US -
  - VDT leader (Miron Livny) is key member of the middleware activity
  - Globus (ISI and Argonne) are also partners
Distributed Physics Analysis
The ARDA Project

ARDA - distributed physics analysis
→ batch to interactive
→ end-user emphasis

- 4 pilots by the LHC experiments (core of the HEP activity in EGEE NA4)
- Rapid prototyping → pilot service
- Providing focus for the first products of the EGEE middleware
- Kept realistic by what the EGEE middleware can deliver
LCG-2 and gLite

**LCG-2**
- Focus on production, large-scale data handling
- The service for the 2004 data challenges
- Provides experience on operating and managing a global grid service
- Development programme driven by data challenge experience
  - Data handling
  - Strengthening the infrastructure
  - Operation, VO management
- Evolves to LCG-3 as components progressively replaced with new middleware
  -- Target is to minimise the discontinuities of migration to the new generation
- Aim for migration plan by end of year

**gLite**
- Focus on analysis
- Developed by EGEE project in collaboration with VDT (US)
- LHC applications and users closely involved in prototyping & development (ARDA project)
- Short development cycles
- Co-existence with LCG-2
- Profit as far as possible from LCG-2 infrastructure, experience
  → Ease deployment - avoid separate hardware
- As far as possible - completed components integrated in LCG-2
  → Improved testing, easier displacement of LCG-2

LCG-2 (EGEE-0)

<table>
<thead>
<tr>
<th>Year</th>
<th>LCG-2</th>
<th>gLite</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>Prototyping</td>
<td>Prototyping</td>
</tr>
<tr>
<td>2005</td>
<td>Product</td>
<td>Product</td>
</tr>
</tbody>
</table>

EGEE-1

---

Last update 05/08/2004 12:29
Les Robertson - CERN-IT-26
Preparing for 2007

- 2003 – has demonstrated event production

- In 2004 we must show that we can also handle the data – even if the computing model is very simple

-- This is a key goal of the 2004 Data Challenges

- Target for end of this year –
  - Basic model demonstrated using current grid middleware
  - All Tier-1s and ~25% of Tier-2s operating a reliable service
  - Validate security model, understand storage model
  - Clear idea of the performance, scaling, operations and management issues
Final Points

- Still early days for operational grids
- There are still many questions about grids & data handling
- The way to grid standards is not clear
  - Standards body - GGF? Oasis?
  - Industrial interests
  - We probably need more practice and experience before standards emerge
- LCG encompasses resources in America, Asia and Europe
- EGEE and LCG are working very closely together
  - to develop an operational grid in an international multi-science context
  - looking for convergence on middleware rather than divergence
- But the LHC clock is ticking - deadlines will dictate simplicity and pragmatism
  - practical “challenges” are essential to get the right focus and keep our feet on the ground