LCG - The Worldwide LHC Computing Grid

openlab
Board of Sponsors
WLCG Status update

CERN
26 April 2007
Les Robertson
LCG Project Leader
LCG Service Hierarchy

Tier-0 - the accelerator centre
- Data acquisition & initial processing
- Long-term data curation
- Distribution of data → Tier-1 centres

Tier-1 - “online” to the data acquisition process → high availability
- Managed Mass Storage - → grid-enabled data service
- Data-heavy analysis
- National, regional support

Tier-2 - ~100 centres in ~40 countries
- Simulation
- End-user analysis - batch and interactive

Canada – Triumf (Vancouver)
France – IN2P3 (Lyon)
Germany – Forschungszentrum Karlsruhe
Italy – CNAF (Bologna)
Netherlands – NIKHEF/SARA (Amsterdam)
Nordic countries – distributed Tier-1
Spain – PIC (Barcelona)
Taiwan – Academia Sinica (Taipei)
UK – CLRC (Oxford)
US – FermiLab (Illinois)
   – Brookhaven (NY)
CERN Tier-0 Progress

ALICE MDC7 target reached

ATLAS T0-2006 Phase 1

Phase 2

Peak transfer of incoming data to tape at over 2GB/s

19-Sep-05

les.robertson@cern.ch
LCG depends on two major science grid infrastructures ....

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

A map of the worldwide LCG infrastructure operated by EGEE and OSG.
CPU Usage – LHC Experiments
March 2007

CERN 20%
11 Tier-1s 60%
140 Tier-2s 20%

Tier-2 Sites Contributing in
March 2007
Grid Activity

- Steady increase in usage of the EGEE and OSG grids
- Example shows LHC experiment jobs run on the EGEE grid
- 3 x increase in past twelve months
- Need a further 5 x increase by mid-2008
- Similar growth on OSG grid
Grid Activity

CPU Time Delivered

Disk Storage Used

Tier-1s and CERN
- CPU usage increased by factor of 2 over past year
- Disk usage by a factor of 4.9
Growth to 2008

Required CPU Growth
CERN + Tier-1s

Required DISK Growth
CERN + Tier-1s
The LHC Grid is all about data.

Experiment computing models define specific data flows between CERN, Tier-1s and Tier-2s.
ATLAS “average” Tier-1 Data Flow (2008)

Real data storage, reprocessing and distribution

Plus simulation & analysis data flow

Tier-0

Tier-2s

Other Tier-1s

Other Tier-1s

RAW
ESD2
AODm2
0.044 Hz
3.74K f/day
44 MB/s
3.66 TB/day

AODm1
500 MB/file
0.04 Hz
3.4K f/day
20 MB/s
0.16 TB/day

AODm2
500 MB/file
0.04 Hz
3.4K f/day
20 MB/s
1.6 TB/day

Tier-2s

Disk storage

CPU farm

Other Tier-1s

EODm1
500 MB/file
0.02 Hz
1.7K f/day
10 MB/s
0.8 TB/day

EODm2
500 MB/file
0.036 Hz
3.1K f/day
18 MB/s
1.44 TB/day

Tier-2s

Disk storage

CPU farm

Other Tier-1s

EODm1
500 MB/file
0.02 Hz
1.7K f/day
10 MB/s
0.8 TB/day

EODm2
500 MB/file
0.036 Hz
3.1K f/day
18 MB/s
1.44 TB/day

Tier-2s

Disk storage

CPU farm

Other Tier-1s

EODm1
500 MB/file
0.02 Hz
1.7K f/day
10 MB/s
0.8 TB/day

EODm2
500 MB/file
0.036 Hz
3.1K f/day
18 MB/s
1.44 TB/day
Data Distribution Tests 2006 – CERN → Tier-1s

April 06 test period
- Sustained operation at the “nominal” rate when LHC is running - 1.3 GB/s
- Peak daily rate 1.6 GB/s

August
- Experiment-driven transfers (ATLAS and CMS) sustained 60% of the SC4 target under much more realistic conditions
  - CMS transferred a steady 1 PByte/month between Tier-1s & Tier-2s during a 90 day period
  - ATLAS distributed 1.25 PBytes from CERN during a 6-week period

19-Sep-05
2007 - CERN → Tier-1 Data Distribution

Daily Report
(VO-wise Data Transfer From CERNCI To All Sites)

Averaged Throughput From 01/01/07 To 23/04/07
VO-wise Data Transfer From CERNCI To All Sites

Data rate required for 2008 run
Data Transfers
Comparison with CSA06 — weekly

CMS PhEDEx - Transfer Rate
52 Weeks from 2006/16 to 2007/15 UTC

LoadTest07

all sites \(\rightleftharpoons\) all sites
Data Analysis on the Grid

Grid middleware creates a general purpose virtual computing service

Logging and bookkeeping

Computer sites

User Interface

Checking job status

Submitting a job

WMS

Register

Information System

VOMS

Getting proxy

File Catalogue

19-Sep-05

Julia Andreeva, CERN

ICHEP06 - Moscow
Data Analysis on the Grid

But there is also a very complex VO specific environment.

Grid middleware creates a general purpose virtual computing service.

How do I make sure that when my job lands at the remote site it will have the same environment as at the local site?

How do I know what I have submitted?

Ten of my thousand jobs crashed – at seven different sites!

Why?

Grid? Application? Data access?

19-Sep-05

Julia Andreeva, CERN    ICHEP06 - Moscow
Each VO needs tools to hide the complexity of the distributed application environment.

- Grid analysis tool
- Experiment Data Management
- LHC specific applications
- Computer sites

- Talks to experiment Data Management to find out where data is and how
- Generates the parallel jobs
- Packages the end-user code with
- Generates the executable shell
- Submits a job
- Generates Grid wrapper
- Performs error recovery
- Provides diagnostic information
- Keeps track of the jobs, eventually retrieving the job output
- Experiment software
- Experiment Data Management
- File Catalogue
- VOMS Catalogue

19-Sep-05 les.robertson@cern.ch
Julia Andreeva, CERN ICHEP06 - Moscow
Ganga

- Job Submission
- Archiving
- Automatic monitoring

Ganga Core

Client

GUI

CLIP

SCRIPT

GPI

Job manager

Job Repository

File Workspace, Input/Output Sandbox

Application Manager

Executable
- Gaudi
- Athena
- ADA
- Other

Job/Application plugins

Monitoring

Localhost

gLite

LSF

DIRAC

Condor

DIAL

LCG

Other

Backend plugins
1) Data location
2) Job preparation
3) Job submission
4) Job status
5) Job output retrieval
Analysis Job Success Rates

CMS CRAB jobs submitted through EGEE Resource Brokers - March 2007

User view – after automatic job re-submission by the Resource Brokers

19-Sep-05 les.robertson@cern.ch
The Worldwide LHC Computing Grid

- The LHC physics data analysis service distributed across the world
  - CERN, 11 large Tier-1 centres, over 100 active Tier-2 centres
- Status at April 2007
  - Established the 10 Gigabit/sec optical network that interlinks CERN and the Tier-1 centres
  - Demonstrated data distribution from CERN to the Tier-1 centres at 1.3 GByte/sec - the rate that will be needed in 2008
  - ATLAS and CMS can each transfer 1 PetaByte of data per month between their computing centres
  - Running ~2 million jobs each month across the grid
  - The distributed grid operation, set up during 2005, has reached maturity, with responsibility shared across 7 sites in Europe, the US and Asia
  - End-user analysis tools enabling “real physicists” to profit from this worldwide data-intensive computing environment