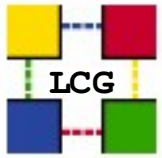


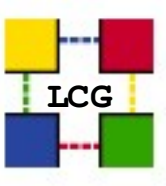
LHC Computing Grid Project - LCG

Report of WG1 to the GDB

6th February 2003

David Foster – LCG Chief Technology Officer
Information Technology Division
CERN – European Organization for Nuclear Research
Geneva, Switzerland

david.foster@cern.ch



Members of WG1

- John Renner Hansen (Atlas/Nordugrid)
- David Foster (LCG)
- Simon Lin (Taiwan computer center Atlas/CMS)
- Laura Perini (Atlas/WP8)
- Yves Schutz (Alice)
- Oxana Smirnova (Atlas/Nordugrid/LCG Apps)
- Eric van Herwijnen (LHCb/WP8)
- Tony Wildish (CMS)
- Flavia Donno (invited) (LCG)

- Many meetings held, 23 revisions of the document circulated.

WG1 Report in Context

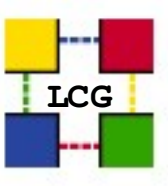
- What it is …
  - A recommendation based on what exists.
  - A specification for the grid technology suppliers of what we need to get started.
  - A starting point to understand the overall picture including things not being treated elsewhere (e.g portals)
- What it is not …
  - Our design for the LHC grid.
  - A complete list of everything we will need.
  - The final solution.
- It is clear that many things will evolve this year and we must be able to take advantage of changes in technology to provide better service.

Key Achievements

- Strategic Technology Advisory Group (STAG) established and provided valuable input.
- Agreement made with VDT team on support structures
  - The VDT team will provide Globus first line communication and support follow-up
- Specific middleware deliverables agreed with the major suppliers.
  - VDT and NMI (Packaged Globus)
  - EDG and Globus (Standard release + fixes)
  - DataTag (Glue enhancements)
- Better understanding with the experiments of the use of LCG-1 middleware
  - Some performance metrics defined
- LCG-Deployment has started to set up the first environments.

Selection Overview

- Specific deliverables identified in the document:
  - Work is starting now with the initial components.
    - VDT 1.1.6
    - EDG 1.4.3
  - Convergence with EDG comes with testbed 2 (April)
    - Globus 2.2.4
    - WP2+RLS
    - WP1+Glue
  - Some "late" components may come but we have to see
    - RGMA
    - WP5
    - VOMS



Selection Overview (2)

- Very basic functionality equivalent to much of the testing done today.
  - Gridmap files
    - Local user accounts
  - Local file storage only
    - Copy input to the worker node.
    - Execute
    - Copy output directly to mass storage (e.g GridFTP to Castor or other systems)
  - Simple interface to job submission
    - No portals considered
    - Experiment "front end" systems

Use of LCG-1

Experiment	July				November			
	Alice	Atlas	CMS	LHCb	Alice	Atlas	CMS	LHCb
% Production Capacity Provided by LCG-1	No production planned only tests	No production planned only tests	10%	No production planned only tests	100%	50%	75%	30%

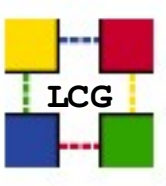
WG1 looked at some simple metrics to understand the scale and scope of the expectation for LCG-1.

Capacity Metrics

Metric	July				November			
	Alice	Atlas	CMS	LHCb	Alice	Atlas	CMS	LHCb
Total Stored Data	?	150GB	10TB	0.5TB	220TB	1TB	140TB	6TB
Total Number of Files	?	2K	100K	2k	200K	10k	850K	24k
Average data read/job	?	300MB	20MB	0.75GB	2GB	50MB	2GB	0.75GB
Average data written/job	?	500MB	200MB	1.5GB	2GB	500MB	1.5GB	1.5GB
Job submission rate (jobs/day)	?	40	4500	10	Variable	200	10K	130

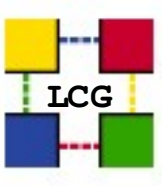
It has been commented that the CMS requirements exceeds what has been possible to date with EDG 1.4.3

It was recommended that capacity be added at a rate that does not significantly reduce efficiency



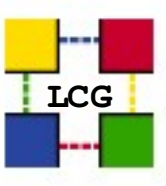
Some Requirements Needing More Work

- Experiment production environment needs to be installed at all sites and maintainable by the experiment.
- Standardized installation of HEP software.
- Tools needed for copying/deleting bulk data between regional centers.
- Worker nodes will need outbound TCP/IP access for now:
  - Wan access to input data sets and writing output data sets
  - Initiating jobs (scripts pull in the jobs)
  - WP2 does not foresee any proxy outbound write technology in the near future.
- Alice will need AliEn gateways at regional centers



Data Management – Past, Present and Future

- Data Access
  - How to open/seek/read/write/close a file in grid storage is still an open question.
    - Local/Remote issues, Authorization
    - Mass storage issues
- Data Management
  - Management of replicas, load balancing, registering/un-registering of replicas.
  - Copying in/out, deleting of replicas etc.
  - No common theme on proxy and 3rd party management of data
- Role of Peer-Peer in the future data management architectures
  - Some interesting opportunities moving away from a strict hierarchical data model.
  - Part of a future architectural modeling.



Data Management Activities

- Work has started on this topic (as requested by the GDB).
  - Information collection is on-going
    - Many issues to the general problem.
    - Need to start simple but must go in the right direction.
  - Many Technologies
    - SRM, SRB, DCache, RFIO, GridFTP, GDMP etc.
  - US CMS Very Active
    - Working to solve this problem. Preliminary work plan produced.
    - Need to understand the activities better and work together.
      - LCG needs to recommend for the LHC Grid as a whole.
  - Small workshop possible with Openlab mid-March.
  - **Initial proposal plus on-going strategy should be created and widely discussed (end March).**

The TCP/IP Issue

- The issue is largely an architectural problem. Todays solution is to have outbound TCP/IP access for:
  - Jobs that use scripts to pull data from remote sites (input data sets and executables)
    - But access to well known nodes could permitted.
    - Jobs could be packaged and pushed but that would require a workflow design change.
  - Jobs that write output datasets directly to remote SE's
    - But data management services should provide:
      - 3rd party transfer initiated by the job. (push)
      - 3rd party transfer initiated externally. (pull)
  - Replica services, for example access to the Metadata Catalog (RMC). RLI/RLS access may be local depending on the infrastructure architecture.
- **But security proxies are still a big block to implementation**
- **Need to fully understand:**
  - **The services that need external access.**
  - **The regional center policies that restrict access to services. (is NAT capability sufficient to allow outbound TCP/IP?)**

3 Levels of Architecture

- An Initial Storage Architecture needs to be defined
  - At least Data Access but also storage services that could be provided (GridFTPd, RFIOd etc…)
  - Storage Use Cases including 3rd party push/pull of data but maybe the security issues are the major problem here.
- A Regional Center Infrastructure needs to be defined
  - Grid services that need to be provided and associated network access policies (e.g. Gatekeeper, AliEn gateway, Storage services etc.)
- A LCG-1 Infrastructure needs to be defined
  - Resource brokers, Information hierarchies, Replica Services (RLS, RLI, RMC and data management services) etc.

- Should be 3 **small** working groups.

Going Forward

- WG1 needs to continue in some form.
  - Proposal discussed at the SC2 was to continue in the context of the LCG Grid Technology Area.
  - Much experience will now be gained by deployment on "what works and what doesn't".
  - Needs to understand some key architectural issues:
    - The experiments workflow's.
    - The storage architecture.
    - OGSA.
  - LCG-1 evolution for the 2004 DC's
    - GAG will provide input. But to be useful must complete sufficiently in August to provide input for the DC04 challenges.