CHIPP Meeting, March 31, 2005

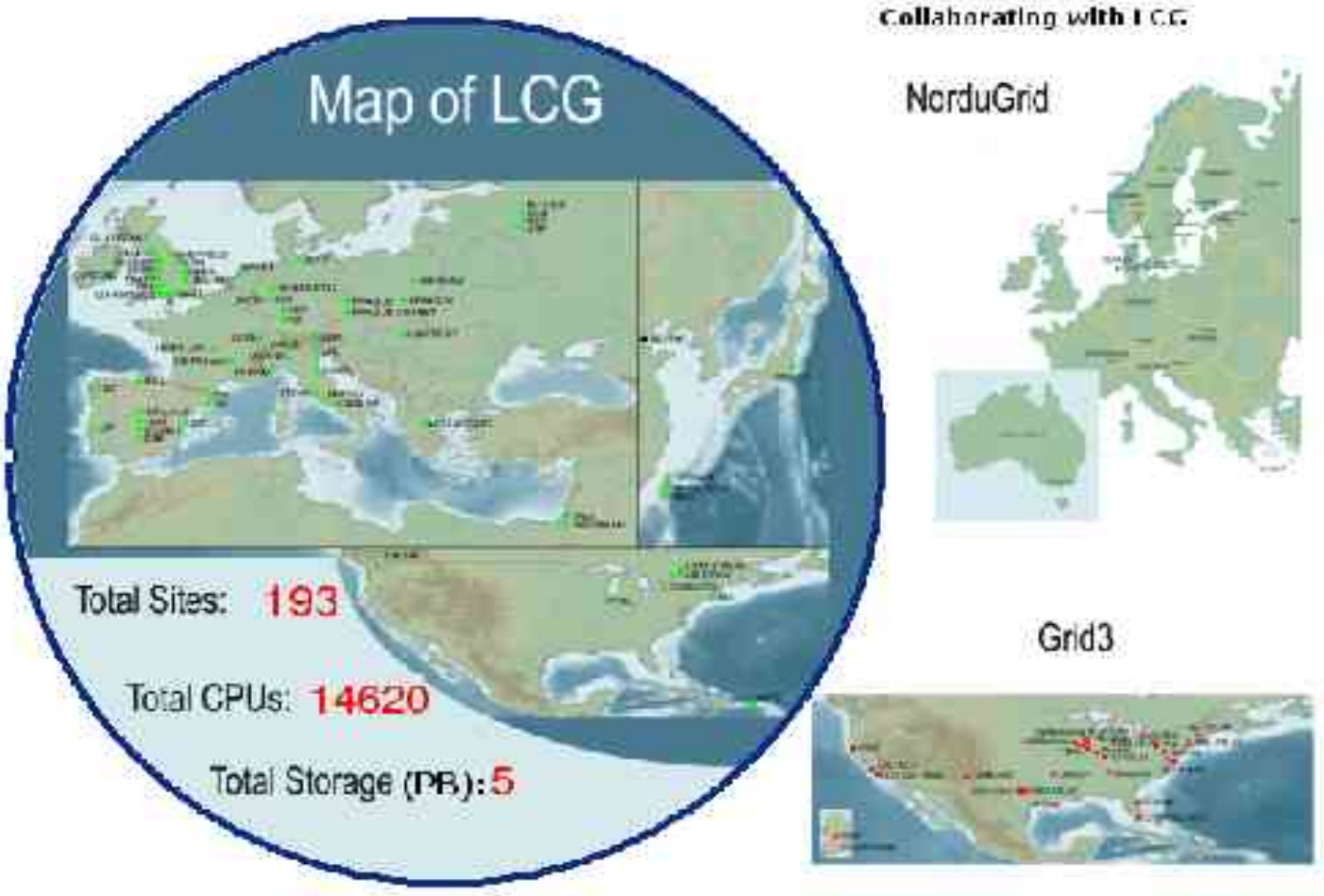
Status of LCG Middleware

Derek Feichtinger

Contents

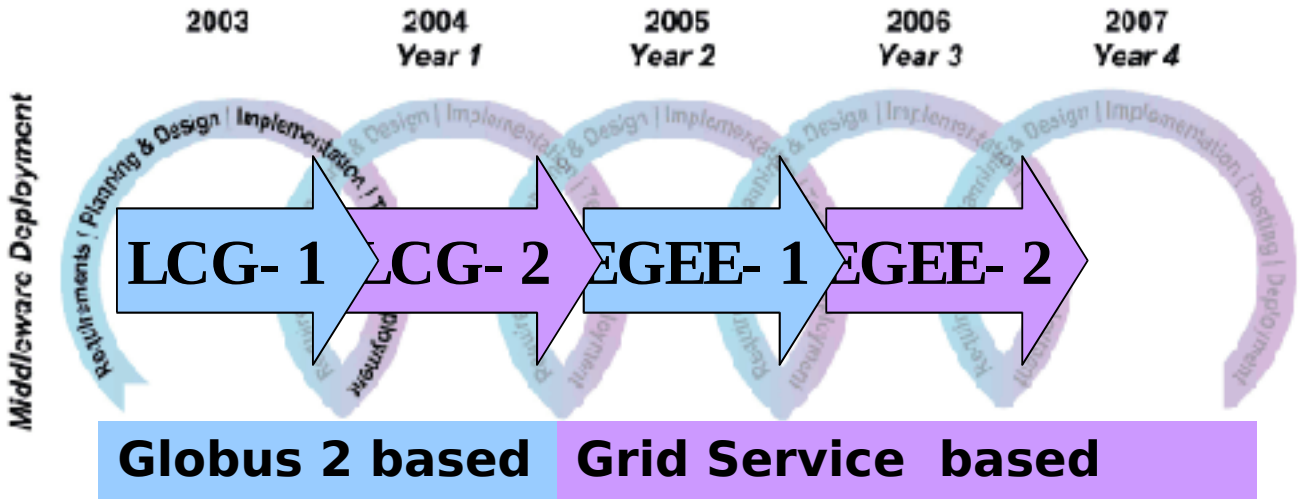
1. LCG / EGEE
2. Grid components in LCG-2
   - usage by experiments
   - developments
3. Conclusion

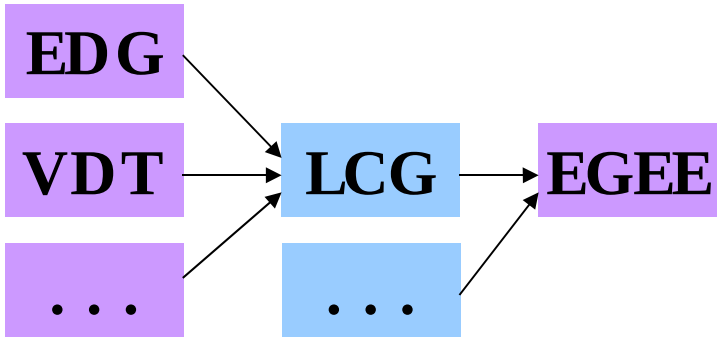
LCG-2 Sites


Map of LCG
Total Sites: 193
Total CPUs: 14620
Total Storage (PB): 5
Collaborating with LCG
NorduGrid
Grid3

EGEE project

EGEE: Enabling Grids for E-sciencE





New Middleware Stack developed in EGEE is called

gLite

EGEE (2)

- EGEE is a project funded by the EU
- Customers are various sciences.
  - Initially the principal customers are HEP experiments.
  - Gradually extending to biochemistry, medicine, ...
- Middleware needs to address needs of all these communities.
  - Requirements can vary substantially, e.g.
    - Security: authorization to patient data
    - connectivity: parallel MPI jobs
  - Solutions have to be adopted which may seem “overkill” for HEP applications
  - Services may be slower due to taking into account more requirements
- This results in a “zoo” of systems (and acronyms) being produced, which hopefully will undergo evolutionary selection.

Grid Components

- A closer look at the following subsystems
  - Workload Management System (WMS)
  - Information Service
  - Data Management
  - Monitoring
- VO-specific services
  - VO-software
  - VO-box

WMS

- The Workload Management System is responsible for accepting and distributing jobs to computing elements (CEs).
- LCG-2

  Improved (more stable) EDG resource broker using a 'push' model. I.e. the broker relies on an efficient information system to base descisions on the Grid's current state.

  Broker able to accept 1 job / every few seconds.

  No bulk jobs possible, no support for job splitting
- Experiments

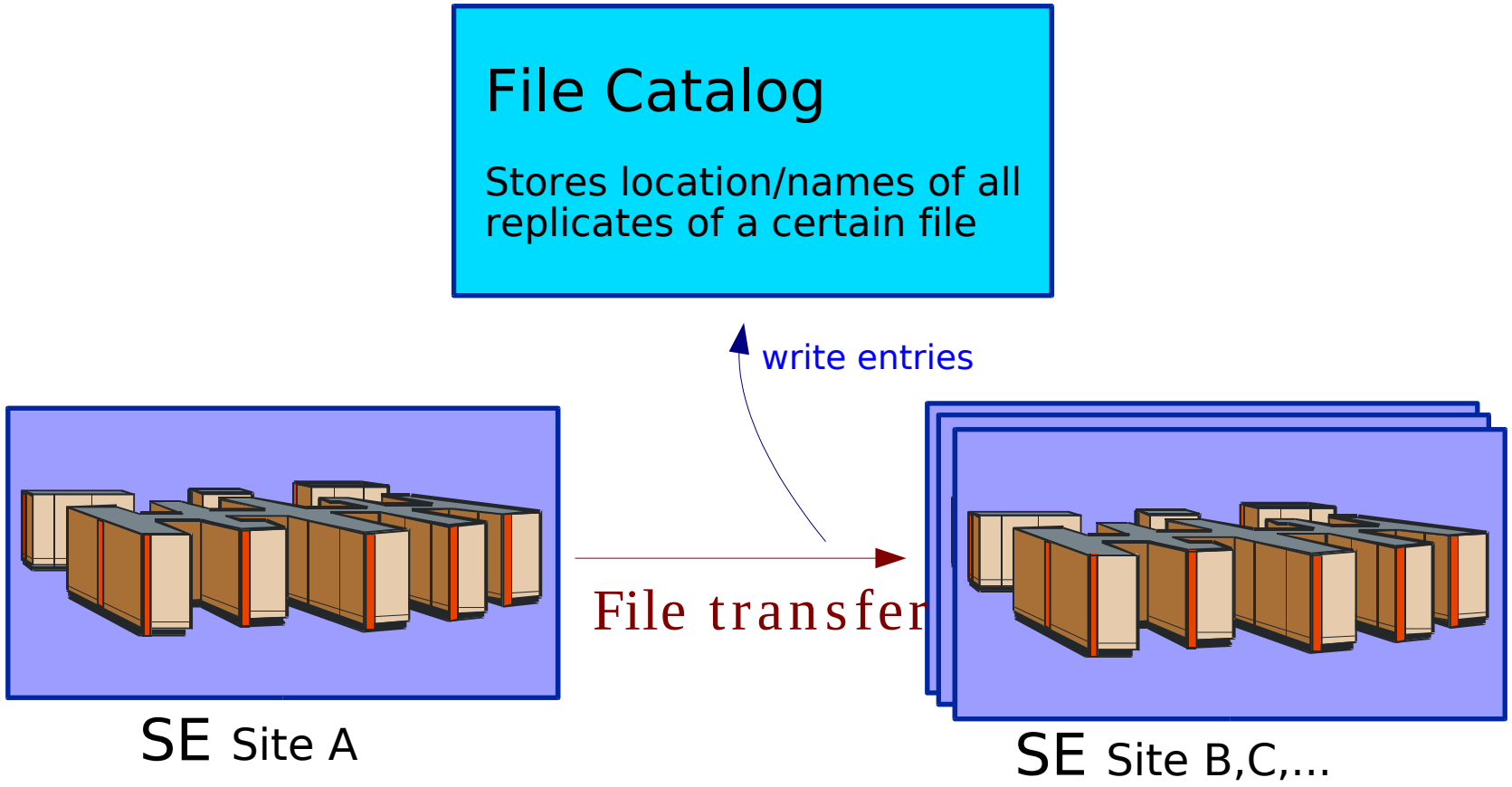
  Missing functionality is made up by
  - Agents (Alice/LHCb): One grid job actually pulls in a whole sequence of jobs from a task queue, always running the maximal possible time.
  - Job splitters submitting the parts as single Jobs, but suffering from the low submission frequency.
- gLite: improved job broker with bulk job support

IS

- The information system is used to propagate status information of grid components upwards to the main information indexes.
- LCG-2
  - Currently still using Globus MDS as a base system. Experience shows lots of scaling and stability issues.
  - The introduction of the BDII component led to much enhanced stability (originally a step-in solution).
  - A new system, R-GMA, is being deployed. Sites need to gain experience with it.
- Experiments
  - mostly use LCG-2 services.
- gLite
  - will be based on R-GMA

Data Management

- Comprises transport of data, physical storage in Storage Elements (SEs), and data access information in catalogs.



Catalogs

- Catalogs store information about the available Grid data (usually on a per file granularity), i.e. locations of files and associated meta data.
- LCG-2

  Centralized catalogues
  - Globus RLS is still used. Several performance and security issues were identified during data and service challenges.
  - LFC, a newly developed LCG catalog is just being deployed.
- Experiments

  ALICE: AliEn Catalog + LFC as local on-site catalog

  ATLAS: RLS

  CMS: RefDB/PubDB

  LHCb: LFC and AliEn, central catalog model
- gLite

  Fireman Catalog, not yet used by experiments

SE

- Storage Elements are used to store files and make them available to grid jobs via entries in file catalogs.
- LCG-2

  Many small sites still use the *classic SE* (gridFTP on a disk array. Local access through rfiod).

  Introduction of SRM standard to talk to storage systems in a generic way.
  - Only MSS supported: Castor
  - Disk Pools: dCache and (new) DPM, the LCG disk pool manager (most interesting for a Tier2).
  - Standards compliance of these tools is an issue (there exist two versions of the SRM standard).

File Transfer

- LCG-2
  - Offers tools to do concurrent file transfers. More mature through introducing timeouts and higher level management tools which prevent corruption of catalog information.
  - No coordinated file transfer service yet (new system developed in EGEE).
- Experiments
  - ALICE: AliEn File Transfer service, xrootd as an additional server/client
  - ATLAS: Don Quijote
  - CMS: PhEDEx service
- gLite
  - File Transfer Service

Monitoring

- Monitoring for resources as well as applications (jobs) and data transfer.
- LCG-2
  - LCG tools (command line) to retrieve R-GMA information
  - GridICE http://gridice2.cnaf.infn.it:50080/gridice/
  - Gstat for resources: http://goc.grid.sinica.edu.tw/gstat/
- Experiments:
  - ALICE: + MonALISA for application monitoring
  - CMS: + MonALISA for transfer/job monitoring

VO specific

- VO Software installation
  - Experiments' software can be installed in shared areas by VO managers with write access to these areas.
  - Installation currently done via normal Grid jobs
- VO-Box
  - Since every experiment wants (needs) to run some extra services/agents, one dedicated host per VO per site will be available (non-root access).

Conclusion

- LCG-2 is a big step in terms of stability
  - LCG-deployment team has fixed many issues
  - a few systems have been replaced or extended
- Functionality not yet enough for experiments
  - leads to homegrown solutions and parallel developments
  - LCG middleware often used in a minimal fashion
- Only way to get to grips with these huge complicated systems are data and service challenges
- To optimise the integration with experiments' requirements, task forces have been formed, which mix experts from middleware providers, ARDA and experiments.
  - Task force leaders determined by the respective experiment
- New middleware (gLite) components will gradually be included