GridPP
UK Computing for Particle Physics

LCG/EGEE Operational Issues

Stephen Burke
RAL

Introduction

- List of problems to initiate discussion
  - A personal selection

- Organisation
- Operation
- Monitoring & Accounting
- User Support
- Design
- Middleware

Organisation

Release management

- Sysadmins would like more information about what will be in future releases
  - With likely timescales
  - Status of dcache? VOMS?
  - Resource implications, migration etc
- Input to developments, site requirements
  - E.g. Tank & Spark, disk pool manager
- How do we upgrade in a safe way?
  - Mixed version systems
- Installation tools – Quattor support?

VO management

- How to add a new VO
  - Who runs services
  - How to add VOs at sites
  - Resource allocation policy
  - User registration policy
- Lightweight system for small VOs
  - Must scale to tens of VOs, hundreds of users
  - NA4 reports 47 VOs already!
- What happens if a VO ends?

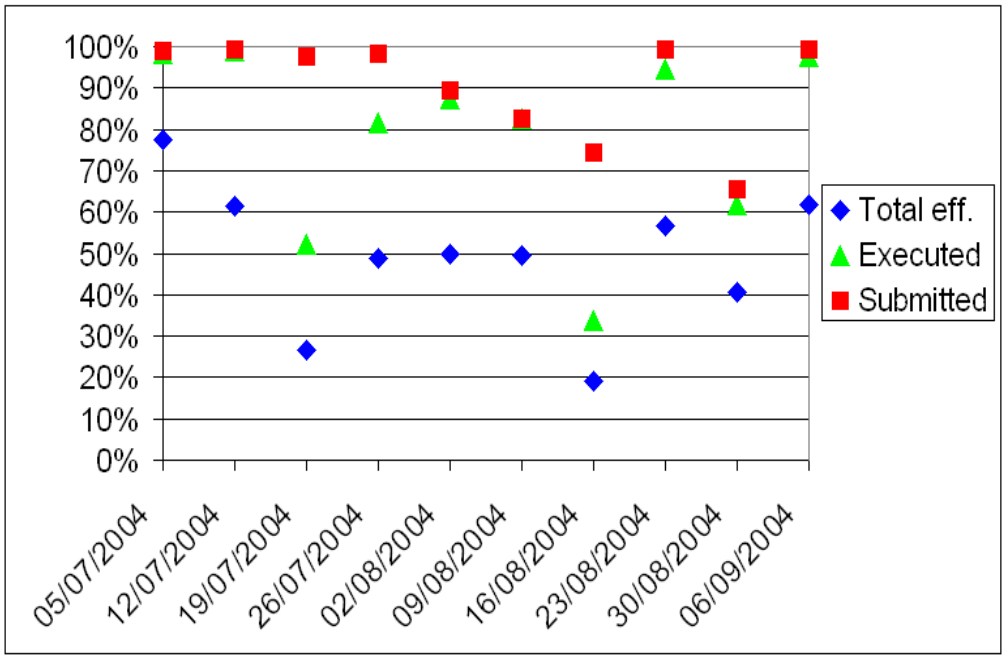
Pre-production system

- How many sites?
- What level of service?
- What middleware?
- Who will use it?
- When?
- SA1 requirements to gLite

Operation

Current LCG Performance

- Efficiency measured with a test job submitted once per hour
- Plot shows weekly average over a ten-week period
- Most problems are site-specific



- Submitted: broker down (x2), myproxy problem, cron failure
- Executed: mainly BDII empty/overloaded
- Total efficiency: many detailed site-specific problems

Site Management

- Tools for sites to check installation
  - Most errors due to site misconfiguration or faults, e.g. NFS, clocks, ssh keys, full disks, …
  - Scaling: 1/century*10,000 = 2/week!
- Ongoing certification
  - Sites seem to “decay” with time
- How to certify for each VO – standard tests?
- Remote control of services?
  - How fast do problems need to be fixed?
  - Small sites not used to 24*7 cover
- Safe service shutdown procedure

Flexibility

- Overall system is very complex, hard to predict the effect of actions
- Need to balance the freedom of sysadmins against usability of the overall system
  - The more flexibility in configuration, the harder to certify/validate
- How do people know what they can do safely, what the consequences are?

Security

- Incident response
  - Several intrusions lately, what if a stolen proxy is used?
  - Are all sites passing on information about incidents?
    - No such thing as a local incident!
  - Shut down sites remotely?
- Logging
  - Can use of proxies be traced?
  - Are the logs secure?
- Care of private keys and proxies
- Outbound IP access

GridPP
UK Computing for Particle Physics

Monitoring and Accounting

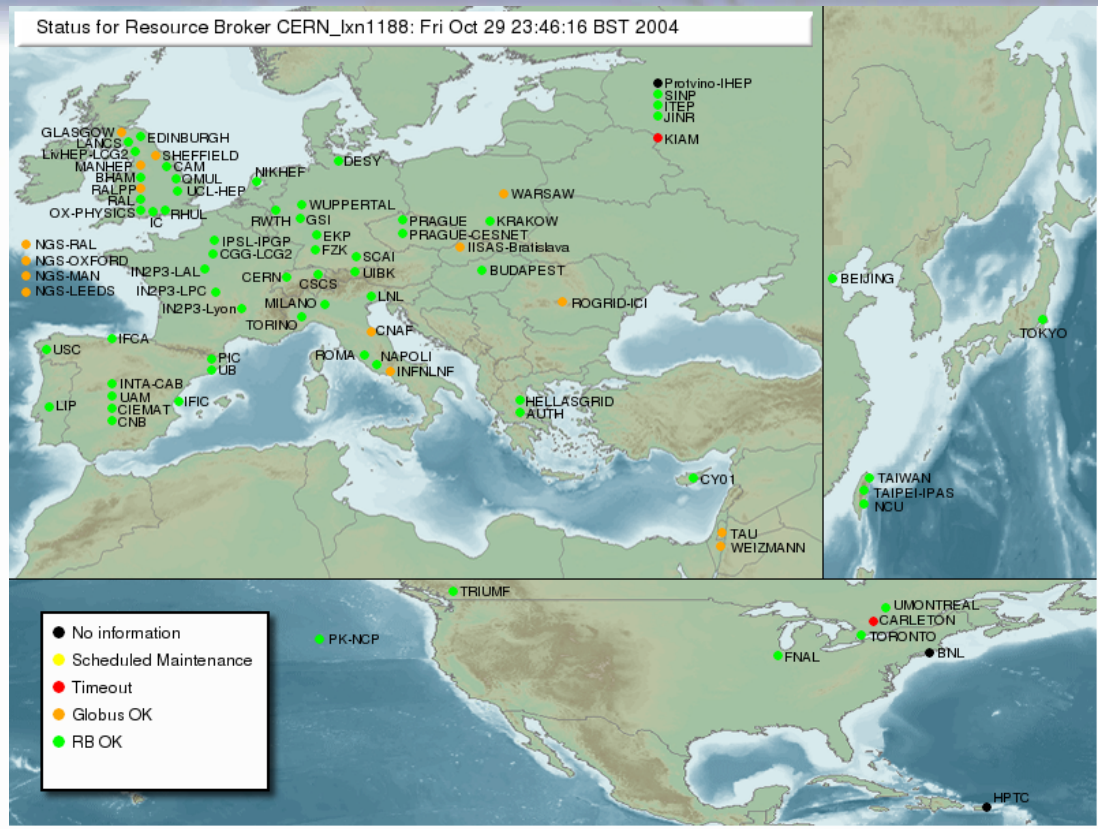
Monitoring

- Lots of tools
  - GridICE, ganglia, MonaLISA, R-GMA, GIIS monitor …
  - Coherency?
- Different clients, purposes
  - Users, sysadmins, CIC/GOC/ROC, funding bodies
  - Routine monitoring, alerts, problem tracing, measuring resources, PR, …
- Info system schema not sufficient?
  - Job information missing
- Must fix problems, not just monitor them!

How many sites?

- Map: 82
- GridICE: 77
- BDII: 84

- CPUs
  - GIIS: 8805
  - GridICE: 34222



Test jobs

- Frequency of job submission
  - Load on CE
  - Time taken to catch problems
- All jobs run as dteam
  - How to test for other VOs?
- Test coverage
  - As problems are found, new tests should be added
- Automatic reporting back to sites?

Accounting

- Several clients
  - VOs, funding bodies, resource providers, SLAs
- What granularity?
  - VOs, users, single jobs?
  - What about failed jobs?
  - Separation of jobs run at low priority?
- Data protection/privacy?
- Normalising different CPUs
  - CPU time or real time? Hyperthreading?
- Accounting for disk space, networking
- Enforcement of quotas?

GridPP
UK Computing for Particle Physics

User Support

Support needed

- Single users, VO managers, sysadmins …
- How do sites and VOs communicate?
  - Or sites and individual users?
- Different kinds of support
  - "How do I …?"
  - Bugs
  - Requirements for extra features
  - Problems with installation/configuration
  - Site/service faults
  - Problems with applications
- Training – pre-emptive support?

User support channels

- Want single point of contact – GGUS?
- Several FAQs – need a single FAQ database
- Many problems very hard, need experts
  - Limited number of experts
  - All reading rollout list …
- Mailing lists may be good for some things
  - ~150 lists, hard to know where to go!
- Use of savannah?
  - 186 open bugs, 46 assigned to “none”
  - Oldest from February

Documentation

- LCG user guide fairly good, but still need EDG manuals for details
  - LCG has modified many things, need updated documentation
- Need sysadmin guide, beyond installation
- LCG web site is complex, hard to find info
  - Use google to find things!
- Many other web sites
  - GridPP, INFN, EGEE, …
  - Lots of info, how do you find it?

System Design

Design

- EDG did not really have a complete system design
- LCG has made short-term decisions on a pragmatic basis, no real overall design
  - Complicated BDII/RB structure
  - Separate services per VO
  - Config for data management tools
- gLite has an architecture, but not a design for a deployed system?
- How much can LCG do?
  - High level design for medium term?

UI configuration

- UI needs to point to RBs, BDIIs etc
  - Will also need VOMS config
  - How does the admin find the info?
    - There is no single list of RBs
  - Need fallback if services are down
    - e.g. can configure multiple RBs but not multiple myproxies, BDIIs
- Get info from info system rather than static configuration?
  - Bootstrap - need a ***single*** info system

Security

- Is VOMS being deployed?
  - If so, how will it be used?
- Secure services needed
- ACLs, ownership for files
- Need VO groups and roles
  - Software managers lose the ability to run as normal users
- Support for multi-VO membership?
  - e.g. access for atlas *and* gridpp

Glue schema

- Several known problems
  - CE can't describe scheduling policies
  - Can't deal with inhomogeneous WNs
  - Various SE problems
  - Clear definition of measurement units
  - Need "failsafe" defaults
- Who controls the schema?
- New objects, e.g. RB?
- Info about running jobs?

Middleware

Error handling

- Services need to be fault tolerant
  - In a large Grid some things are always broken
- Error messages need to say what went wrong
  - Often very hard even for experts to diagnose problems
  - Can't tell if a job exceeded time limit!
- Need consistent logging, remote diagnostics

Job submission

- Passing requirements (memory, CPU time etc) to LRMS
- No simple link from RB job ID to PBS ID
- Some successful jobs get resubmitted, some failed jobs look like successes, some jobs are just lost
  - If an RB fails the jobs are in limbo
- RB doesn't notice if a site keeps failing – black holes
- Local environment with inhomogeneous sites
  - lcg-*, rgma, …
  - Installed software
    - Versions, locations
  - Outbound email
- Ranking algorithm – EstimatedResponseTime!

Data management

- Check consistency between SE and LRC
  - Sites should avoid deleting files!
- "Close SE" concept is not well defined
  - Info system, brokerinfo, edg-rm config, environment variables
- Does the classic SE disappear?
  - Could improve some things
- Control access, e.g. to shut down?
  - iptables -> block SYN packets!
- File ownership?
- Space management on WNs
- Software installation
  - Is Tank & Spark sufficient?
  - What about the gLite solution?

Information system

- Unique service index to locate all services
  - How to find BDIIs?
  - Need to know about RB/BDII/myproxy relations
  - How to understand the structure of the system
- Info provider configuration is complex, hard to diagnose problems
  - Need management tool?
- How to find out service status info?
  - Use web pages as well?
  - Entries in the GOC DB not enough

Conclusion

Priorities

- Security incident response
- Ongoing site validation
  - Problems must be fixed and not just detected
- Pre-production system
- VO management (VOMS?)

Next Steps

- So many problems, so little time …
- Organise targeted working groups
- Develop operations model for next year
  - Have to please many constituencies:
    - VOs, users, sysadmins, middleware developers, funding bodies, …
    - Not always with the same goals
- Who agrees to do what?