

## CZ Cluster Summary

*Aleš Křenek, CESNET*

- Remove bottlenecks in L&B
  - logging via L&B proxy – **done**, more than 1 M jobs/day
  - logging via logd – **not critical**
    - ▶ most traffic goes via L&B proxy
    - ▶ CE with 1000 WN's: 1 M jobs/day if running at least 250 s in average
  - registering large collections – **4× faster**
    - ▶ reduced number of SQL statements during subjob registration
    - ▶ transaction-enabled database engine (InnoDB)
    - ▶ foreseen further 2× improvement with parallel proxy+server registration
  - transparent connection management in client libraries
    - ▶ share connections among threads, avoid unnecessary re-connects
    - ▶ not thoroughly tested so not included in gLite 3.1
  - performance tests wrapped in self-contained scripts

- Proxy renewal within the services (proxy renewal library)
  - available in both gLite 3.0 and 3.1

- Proxy renewal within the services (proxy renewal library)
  - available in both gLite 3.0 and 3.1
- Establishment of trust between the service and myproxy
  - implemented as extension of MyProxy server, allowing to configure authorization via VOMS
  - included in MyProxy 3.6 release, not in VDT yet

- Proxy renewal within the services (proxy renewal library)
  - available in both gLite 3.0 and 3.1
- Establishment of trust between the service and myproxy
  - implemented as extension of MyProxy server, allowing to configure authorization via VOMS
  - included in MyProxy 3.6 release, not in VDT yet
- “Hot standby” configuration of a pool of WMS services.
  - MySQL Cluster too specialised and restricted
    - ▶ lot of data must fit into memory, sequential query processing, difficult schema changes, . . .
  - Oracle – to be evaluated
  - LCG HAGD – doubts about performance, unclear write semantics
  - L&B-level replication
    - ▶ more development effort required
    - ▶ the fastest and most reliable solution

- Decouple bulk-submitted jobs from DAGs (L&B part)
  - work in progress, **prototype expected in October**
  - still unclear issues (minor from L&B viewpoint)

- Decouple bulk-submitted jobs from DAGs (L&B part)
  - work in progress, **prototype expected in October**
  - still unclear issues (minor from L&B viewpoint)
- Allow subscriptions to LB information for all jobs in a VO for selected VO users
  - **requirements need to be clarified**

- Decouple bulk-submitted jobs from DAGs (L&B part)
  - work in progress, **prototype expected in October**
  - still unclear issues (minor from L&B viewpoint)
- Allow subscriptions to LB information for all jobs in a VO for selected VO users
  - **requirements need to be clarified**
- Methods for building statistics from LB
  - L&B support (data export and dump-parsing tools) ready in gLite 3.1
  - issues on purging data from L&B have to be clarified
  - **no feedback from JRA2**



- Decouple bulk-submitted jobs from DAGs (L&B part)
  - work in progress, **prototype expected in October**
  - still unclear issues (minor from L&B viewpoint)
- Allow subscriptions to LB information for all jobs in a VO for selected VO users
  - **requirements need to be clarified**
- Methods for building statistics from LB
  - L&B support (data export and dump-parsing tools) ready in gLite 3.1
  - issues on purging data from L&B have to be clarified
  - **no feedback from JRA2**
- Enable L&B to relate system-level events to active jobs.
  - **still needs more planning**

- Deployment of Job Provenance system on preview test-bed and test by applications
  - **deployment done** ([skurut1.cesnet.cz](http://skurut1.cesnet.cz))
  - tested within the First Provenance Challenge
  - no production-like traffic yet

- Deployment of Job Provenance system on preview test-bed and test by applications
  - **deployment done** ([skurut1.cesnet.cz](http://skurut1.cesnet.cz))
  - tested within the First Provenance Challenge
  - no production-like traffic yet
- Job Provenance in gLite 3.X
  - **ready for gLite 3.1**

- followup of IPAW'06 workshop, 17 participants  
<http://twiki.ipaw.info/bin/view/Challenge>
- comparison of various provenance systems
- the challenge task:
  - sample workflow (5 stages, 15 nodes), medical image processing
  - to be implemented and run in an arbitrary way
  - 9 challenge queries, e.g.  
*“Find all images outputted from workflows where at least one of the input file headers had an entry `global_maximum=4095`.”*
- DAG support in gLite WMS was used to run the workflow
- information was gathered via L&B (including user tags)

- JP was able to answer 8 of the queries
  - unable to work with attributes assigned to data, not jobs
- main differences wrt. other system
  - approach to data
    - ▶ JP (via L&B) tracks the jobs actively and explicitly
    - ▶ no need to derive eg. workflow structure from other job traces
  - low-level system (no GUI etc.)
  - aiming at production quality, not experimental system only
- foreseen possible cooperation
  - JP – reliable and scalable storage and search engine
  - higher level systems – additional processing, user interaction

- finish L&B performance improvements
- support for non-DAG job collections
- evaluate Oracle hot-standby capabilities, make final decision on support in L&B
- support SA3 in developing functional and performance L&B and JP tests
- JP evaluation – we need real usage of Preview testbed
- get rid of JP remaining “prototype” features

- generic L&B
  - separate re-usable and WMS-specific functionality in L&B
  - non-WMS jobs (Condor)
- Job Provenance
  - complex workflow support, relationships among jobs
  - further integration (gather job sandboxes, user interface, statistics)
  - long-history features (file format evolution, purging)
  - complex AuthZ scheme
  - service discovery – how to find a suitable JPIS to query?