

L&B Job Statistics and L&B Monitoring

Jan Pospíšil

www.eu-egee.org



Information Society



Motivation: L&B contains valuable aggregate information on jobs executing in the Grid infrastructure

- JRA2 Job Statistics
- Job Monitoring

Both collect data from our L&B (MySQL) databases

- internal database structure may change
- misinterpreting the values (especially the state)
- possible security holes (remote DB access)

L&B-based tools to support both needs and provide the data *correctly*.

- One dump file per job created on regular basis (cron job)
 - on L&B server, during the purge (JP upload)
- Processing dumps by the `glite-lb-statistics` tool, part of `org.glite.lb.utils`
 - reads a dump and outputs an XML file with particular Job Record
 - <http://egee.cesnet.cz/en/Schema/LB/JobRecord>
 - agreed in Brno, All JRA1 meeting, June 2005
- A strategy of moving files to a statistics server to be agreed.
 - Suggestion: statistics server downloads itself the dump data from the L&B bkserver and is also responsible for removing the data there

- Assuming single statistics server
 - more also possible, but it requires more complicated configuration and administration
- Grid-FTP as the transfer protocol
 - or something else?
- Mutual authorisation necessary
 - L&B must trust the statistics server, because it offers the sensitive data
 - statistics server must trust L&B that the data are *valid*

Summary: statistics server get in principle the same data, consistent with original L&B interpretation

Job Statistics: post-mortem analysis vs.

Job Monitoring: on-line information

- Querying the L&B server on regular basis to obtain data like:
 1. number of jobs in the system grouped by internal status (SUBMITTED, RUNNING, DONE, ...)
 2. number of jobs that reached final state in the last hour and their statistics like average, maximum and minimum times spent in the system
 3. number of jobs that entered the WMS system in the last hour
- Several supporting tools (`glite-lb-mon*`) created, also part of `org.glite.lb.utils`
 - they use L&B API (`edg-wll_QueryJobs`, ...) and
 - new bkindices

Summary: we strongly encourage using L&B-based tools instead of accessing the MySQL database directly.

Job Statistics: post-mortem analysis vs.

Job Monitoring: on-line information

- Querying the L&B server on regular basis to obtain data like:
 1. number of jobs in the system grouped by internal status (SUBMITTED, RUNNING, DONE, ...)
 2. number of jobs that reached final state in the last hour and their statistics like average, maximum and minimum times spent in the system
 3. number of jobs that entered the WMS system in the last hour
- Several supporting tools (`glite-lb-mon*`) created, also part of `org.glite.lb.utils`
 - they use L&B API (`edg_wll_QueryJobs`, ...) and
 - new bkindices

Summary: we strongly encourage using L&B-based tools instead of accessing the MySQL database directly.

Job Statistics: post-mortem analysis vs.

Job Monitoring: on-line information

- Querying the L&B server on regular basis to obtain data like:
 1. number of jobs in the system grouped by internal status (SUBMITTED, RUNNING, DONE, ...)
 2. number of jobs that reached final state in the last hour and their statistics like average, maximum and minimum times spent in the system
 3. number of jobs that entered the WMS system in the last hour
- Several supporting tools (`glite-lb-mon*`) created, also part of `org.glite.lb.utils`
 - they use L&B API (`edg_wll_QueryJobs`, ...) and
 - new bkindices

Summary: we strongly encourage using L&B-based tools instead of accessing the MySQL database directly.