

Communicating Machine Features to Batch Jobs

WLCG MB, March 8th 2011

Tony.Cass@cern.ch

Reminder

- ◆ The HEPiX Virtualisation Working Group proposed the creation of 3 files in `/etc/machinefeatures` with machine information during virtual machine instantiation:
 - `hs06` – the HS06 rating for a single core,
 - `shutdowntime` – timestamp for when the node is expected to be shutdown, and
 - `shutdowncommand` – the full path to a command that can be invoked to arrange early termination of the machine.
- ◆ Passing similar information to jobs in real machines is considered to be useful.

Proposal for real machines

- ◆ Physical machines, just as virtual machines, should have in `/etc/machinefeatures` the files
 - `hs06` – the HS06 rating for a single core
 - `shutdowntime` – timestamp for when the node is expected to be rebooted (may be empty)
- ◆ Additionally, the batch system should provides each job with an environment variable, `$JOBFEATURES`, pointing to a directory with the files
 - `jobstart` timestamp for the start of job execution
 - `cpu_limit` the allowed number of cpu-seconds that can be used by the job. The value must be comparable directly against system reported cpu consumption.
 - `wall_limit` the allowed number of wall clock seconds. The job will be terminated at time `jobstart+wall_limit` whether or not `cpu_limit` has been reached.

With thanks to Steve Traylen
and Ulrich Schwickerath

Implementation Status

◆ LSF

- Initial prototype available in SVN repository; initial tests suggest the implementation works OK, no firm date for deployment as yet, but possible before end-March.
 - » Developed by Ulrich Schwickerath; `$JOBFEATURES` points to `/var/tmp/jobfeatures/<jobid>`

◆ PBS

- NIKHEF asked to deliver an implementation; waiting to see whether or not this request will be accepted.

◆ Other workload management systems

- Which?
- Who?