

Status of proposal and
implementation of a
mechanism to make
environment information
available to jobs on a WN

- Original proposal by HEPiX virtualization working group for virtual machines
- Proposal made on 14/1/2011 by Tony Cass
- Prototype implementation at CERN in production since 4/2011 for LSF
- Found to be useful for physical nodes as well
- Proposal picked up to deal with multi-core jobs
- Redone and documented the specifications over summer 2012

See also the presentation by Tony Cass, GDB June 2012

The specifications can be found here:

<https://twiki.cern.ch/twiki/bin/view/LCG/WMTEGEnvironmentVariables>

Currently still marked as “draft” but very close to final...
(Feedback is still possible)

How does it work ?

For each job two **environment variables** are set pointing to a local directory:

\$MACHINEFEATURES

Points to a directory containing worker node specific information

\$JOBFEATURES

Points to a directory containing job specific information

Files in \$MACHINEFEATURES/

<i>hs06</i> :	host HS06 rating in its current configuration
<i>shutdowntime</i>	shutdowntime, Unix time stamp (in seconds), if the machine is being drained
<i>jobslots</i>	Number of job slots of this machine
<i>phys_cores</i>	Number of physical cores
<i>log_cores</i>	Number of logical cores (if SMT enabled)

Files in \$JOBFEATURES/

<i>cpu_factor_lrms</i>	normalization factor as used by the batch system
<i>cpu_limit_secs_lrms</i>	CPU limit in seconds, normalized
<i>cpu_limit_secs</i>	CPU limit in seconds, real time
<i>wall_limit_secs_lrms</i>	run time limit in seconds, normalized
<i>wall_limit_secs</i>	run time limit in seconds, real time
<i>disk_limit_GB</i>	scratch space limit in GB
<i>jobstart_secs</i>	job start time, unix time stamp (seconds)
<i>mem_limit_MB</i>	memory limit (if any)
<i>allocated_CPU</i>	number of allocated CPUs to the job

Torque 2.5:

<http://www.nikhef.nl/~janjust/wn-env/wlwg-wn-env-torque-0.2-1.rhel5.noarch.rpm>

PBS 2.3:

<http://www.nikhef.nl/~janjust/wn-env/wlwg-wn-env-pbs-0.2-1.rhel5.noarch.rpm>

PBS Professional and Grid Engine:

Contact Manfred Alef from KIT

LSF:

Contact me

Any other implementations already out there?

LSF version limitations:

Can't use batch system queries to retrieve information, hence relies on getting the information from somewhere else

- Parse local LSF configuration files
- Hard-coded convention on the number of job slots per node following site policies
- Get HS06 rating from Quattor created local file
- Patch to CERN specific job starter script to set environment variables
- Currently integrated into the LSF deployment scripts at CERN

Deployment status

CERN: LSF version following latest specs

NIKHEF: PBS/Torque

KIT: GridEngine/PBS professional

What is next ?

- Collect feedback on existing implementations
- Finalize specifications
- Update implementations
- Continue deployment