# Machine/Job Features mechanism

Within the HEPiX virtualization group and the WLCG MJF Task Force, a mechanism has been developed which provides access to detailed information about the current host and the current job to the job itself. This allows user payloads to access meta information, independent of the current batch system or virtual machine model.

At the start of 2016, the mechanism was described in HEP Software Foundation technical note HSF-TN-2016-02 "Machine/Job Features". As the schema is designed in an extensible way, further additions are expected and can be readily accommodated.

## Definitions

On VM-based systems, references to "jobs" in the specification are to be interpreted as "virtual machines" and "machines" as "hypervisors".

When jobs are running within virtual machines, the entity that provides the system level configuration or contextualization of the VM acts as the "resource provider".

## Environment variables

For each job, two environment variables may be set, with the names $MACHINEFEATURES and $JOBFEATURES.

These environment variables are the base interface for the user payload. Their values must be provided for the job by the resource provider.

In the case of virtual machines on IaaS cloud platforms, the virtual machine may discover the values to set for the environment variables from "machinefeatures" and "jobfeatures" metadata keys provided by the resource provider via the cloud infrastructure. These metadata keys should only be accessed once in the lifetime of each virtual machine. Alternatively, the values to set may be supplied as part of the contextualization of the virtual machines.

## Directories

The environment variables point to directories created by the resource provider. Inside, the file name is the key, the contents are the values, so that files can be referred to with expressions like $MACHINEFEATURES/shutdowntime . The directory name should not include the trailing slash. These directories are either local directories in the filesystem or sections of the URL space on an HTTP(S) server. The user positively determines whether the files are to be opened locally or over HTTP(S) by checking for a leading slash or the prefix http:// or https:// respectively. Typically this can achieved using library functions which can transparently handle local files and remote URLs when opening files.

Unlike metadata keys, the key/value files may be accessed multiple times to check for changes in value or in the absence of caching by the user. An HTTP(S) server may provide HTTP cache control and expiration information which the user may use to reduce the number of queries. All files in the directories must be readable by both the user and the resource provider services, and have file names which only consist of lowercase letters, numbers, and underscores.

## $MACHINEFEATURES

total_cpu
hs06
shutdowntime
grace_secs

## $JOBFEATURES

allocated_cpu
hs06_job
shutdowntime_job
grace_secs_job
jobstart_secs
job_id
wall_limit_secs
cpu_limit_secs
max_rss_bytes
max_swap_bytes
scratch_limit_bytes

## Deployment

The role of the task force shifted during 2016 to supporting the deployment of software to create the Machine/Job key/value pairs as described in the technical note.

RPMs are provided for PBS/Torque and HTCondor, developed with the help of sites running these batch systems. These RPMs install scripts which determine the relevant values from the operating system and batch system parameters, potentially overridden by values chosen by the site administrators. It is also possible to install the scripts directly using a configuration management system. In most cases, the RPMs or scripts can be installed without any site-specific customisation other than the measured value of HEPSPEC06 for the hardware.

Machine/Job Features key/value pairs are also provided by default at VM-based sites managed by Vac or Vcycle.

In total, a dozen mainly Tier-1 and larger Tier-2 sites were providing Machine/Job Features using these scripts or systems by the start of CHEP2016.

To monitor the progress of the deployment and to validate the sites as they are enabled, a SAM/ETF monitor has been written and deployed. This is run at all WLCG sites and provides pass, fail and warning outcomes plus a detailed report of the outcome for each key/value pair tested.

The sites participating so far have also been invaluable in studies of the variation in HEPSPEC06 and the DIRAC fast benchmark (DB12) between processor versions and sites.

https://twiki.cern.ch/twiki/bin/view/LCG/MachineJobFeatures

M. Alef, T. Cass, J.J. Keijser, A. McNab, S. Roiser, U. Schwickerath, I. Sfiligoi