

MPI VT report

OMB Meeting
28th February 2012



- Introduction
 - Objectives
 - Motivation
 - Expected Output
- Status Report
 - Open tasks and progress
 - Open issues
- Future work

- MPI VT created in November 2011
 - Supposed to end on May 2012
 - https://wiki.egi.eu/wiki/VT_MPI_within_EGI
- Objectives
 - Collaborate with MPI user communities
 - Promote MPI deployment
 - Enhance MPI experience and support for EGI users.
- Output
 - Improve the communications between MPI users and developers within EGI SA3.
 - Improve / create new MPI documentation for EGI users and administrators.
 - Find and detect current issues and bottlenecks related with MPI jobs.
 - Set up a VO on EGI with site committed to support MPI jobs.

- Motivation:
 - MPI documentation was fragmented in different places.
 - Too much site admin / infrastructure oriented.
- Progress:
 - Documentation reviewed and extended by MPI VT members.
 - MPI Administrator guide: <https://wiki.egi.eu/wiki/MAN03>
 - new sections to clarify MPI doubts
 - Working on an MPI user guide
 - Propose to concentrate documentation under a single endpoint

- Motivation:

- Few sites are publishing correctly MPI tags and flavours
 - Some MPI info is omitted, missing or incorrect.
 - Incoherent information published by queues supporting MPI

- Progress:

- Survey the information published by sites
- Inspect which GLUE 1.3/2.0 variables could be used in the MPI context and ask for its correct implementation
 - GLUE 1.3 “MaxSlotsPerJobs” could be already used but only one site publishes a different value wrt the default “999999999”
 - Detected problems in the CPU / WallClockTime relationship published by sites. May lead to the early killing of parallel jobs

- Motivation:
 - Current SAM nagios probes are quite simple.
 - It does not detect if a MPI sites is really working fine or not.
- Progress:
 - A new set of nagios probes specifications were created by MPI members:
https://wiki.egi.eu/wiki/VT_MPI_within_EGI:Nagios
 - Probe to check information system
 - Set of probes to check the MPI functionality (EGEE MPI WG recommendations)
 - Request the creation of a new GOCDB Service Type
 - GOCDB acting like an ITIL CMDB: all configuration items should be registered
 - Release the dependency from the information system
 - Allows accurate infrastructure information about infrastructure status

- Motivation:

- The only way to recognize a parallel jobs in the accounting system and looking for jobs with efficiency >1
- APEL is not ready to registry MPI usage records.
 - This work is on progress and tracked by JRA1.4 task and EMI.
 - New APEL with MPI capabilities will be ready at the end of summer.

- Progress:

- MPI VT members are in contact with EMI developers to describe the parallel usage record for the different batch systems.
- Hopefully this will lead to some requirements to an accounting system able to correctly record parallel usage

- Motivation:
 - Detect LRMS MPI issues. Check MPI LRMS status support.
- Progress:
 - CREAM GE EMI release fully supports MPI. Tested by MPI VT members.
 - Issue with Torque / Maui MPI support (GGUS ticket #)
 - Not supported by EMI but by a third party.
 - EPEL includes torque server / client packages and an external EMI repo includes MAUI updates.
 - Issue raised to EGI SA1/2 and to EMI

- Motivation:
 - MPI VO bring together sites and users interested in MPI.
 - This VO is NOT intended for everyday use by all users wishing to use MPI.
 - The main reason for its establishment is to collect experience that will be later adopted by regular VOs.

- Progress:
 - A new VOMS service was created and is ready to include new users:
https://egee.cesnet.cz/mpi/registration/prihlaska_priprav.php
 - MPI VO configuration is available in the MPI VT wiki page.
 - At this moment CESNET is supporting this VO for testing purposes.

- Batch system support: It was discussed with EMI.
 - Torque batch system is not supported by EMI. Is a third-party piece of software.
 - Torque server/client are maintained by Steve Trayler (not in EMI). These packages are included into EPEL repo.
 - MAUI scheduler is also maintained by Steve Trayler (not in EMI) and updates an external CERN repo.

- Continue to unify and simplify MPI documentation
- Inspect how the IS schema can adjust to specific LRMS policies, and enforce the correct values
 - May involve interaction with LRMS IP developers
- Tested “new” probes functionality
 - Should be developed by SA3
- Include new sites to support VO MPI
 - For testing purposes. Cesga will include the new VO soon.
- Analyse NGI feedback and surveys.
 - NGI Italy will present a report in the next VT meeting. This is allow the detection of problems from user perspective
- Provide requirements to a new accounting system able to support MPI jobs

- Questions?