MPI Survey Report in the IGI Infrastructure

Costantini Alessandro

INFN - IGI

alessandro.costantini@cnaf.infn.it

Summary

- Survey description and responses
- Needs and requirements
- Test of the MPI Attributes
  - SMPGranularity
  - HostNumber
  - WholeNodes

Aim of the Survey

- Identify the current status of the computational resources made available by those sites that are part of the Italian end EGI infrastructure

- Focus the attention on those resources devoted for parallel calculations

- Official documentation and Customer Care experience

Status of the computational resources

- 29 questionnaires have been collected on a total of 57 production sites in the IT domain
- SL5.5 – gLite3.2 – CE CREAM
- 55% of the sample declare to support MPI (about 3,5K cores)
- 87% of these resources are shared among different projects
- MPICH1, MPICH2, OPENMPI

NAGIOS failures

19% of the Sites supporting MPI declare to hide the correct MPI TAGs due to the following reasons:

- Configuration problems; Nagios failures
- Nagios failures due to multiple libraries support (supporting one library the problem does not appear)

Documentation

The sample has been queried about the official documentation already available for Admins and Users involving the following procedures description:

- MPI-installation
- MPI-configuration
- MPI-getting started

Documentation

Admins seem to use the available documentation but the absence of a “Very Good” or “Excellent” rate indicates that it is unsatisfactory for people that already had experience with MPI in the three aspect described above.

A deep restyling is needed and appreciate specially after the introduction of the new MPI attributes.

Customer Care

The sample has been queried for their customer experience (and the service received) described as follow:

- MPI support as Admin
- MPI support as a User

Customer Care

- Admins who had contact with MPI-US seem to appreciate the effort spent by the people involved in MPI (at the moment we cannot define a proper MPI support team)
- The absence of a “Very Good” or “Excellent” rate indicate that this effort is unsatisfactory
- Greater attention to the customer is needed.

Conclusions

- Restyling of the available documentation
- Resolution of the most common configuration problems and probe failures
- Implementation of new probes able to identify
  - MPI resources effectively available in a site [2]
  - Max allowed MPI resources for a single calculation [3]

Conclusions

- Work out of the MPI related "known issues" in order to distribute a stable version of the package(s) [4,5,6]
  - IGI released new packages with the MAUI v. 3.3-4
- Compilers: some applications have compilation problems with the standard compilers in SL5.7 (gcc 4.1)
- Impossibility to handle more instances of the same MPI library

APPENDIX 1 – test of MPI attributes

In all the job submissions (via WMS and direct to the CE CREAM) the following attribute's values have been used:

- CPUNumber = 8;
- SMPGranularity = 4;
- HostNumber = 2;
- WholeNodes = true;

The above configurations require a CE with 8 CPU and two nodes with SMP>=4 as from the MPI user guides.[7,8]

APPENDIX 1 – test of MPI attributes

LIST-MATCH

No errors have been registered using the specified attributes in the JDL

SUBMISSION

From the results it is clear that a minimum of

- Nodes with SMP>=4
- Hosts>=2

are guaranteed from the batch system of the CE but this does not means that the processes will be equally distributed among the Hosts.