

Requirement from Sys Admin: Application User's Perspective

Chintalapati Janaki
SSDG
C-DAC, Bangalore

Basic understanding

User needs to understand

- a. How to login into the submit node and from there into grid nodes (gssissh).
- b. Clusters available under the Grid (needs to know the commands to check for the available resources)
- c. How logins are mapped in different nodes
- d. The hostname of all available nodes to login
- e. Security mechanism followed
- f. Network connectivity available and the bandwidth & latency details

Grid Node Details

The configuration of each cluster and the specification details

- i. Memory
- ii. Available Disk space
- iii. Number of nodes and CPUs available in each node
- iv. Compilers
- v. Libraries (can be even mathematical libraries or scientific libraries like ESSL)
- vi. OS details
- vii. Architecture details like 64-bit or 32-bit

- Once the grid nodes suitable for their job execution are identified then
 - how to compile and run the jobs on the nodes (may require help from sys admin if any compilation errors come)
 - how to write RSL scripts for job submission
 - Globus commands (for job submission, job status checking, job cancellation, getting back the outputs etc).

General Requirements

- (Non)Dedicated Nodes
- Disk storage
- Memory
- More Swap size for specific applications
- Debuggers, Profilers
- Recent version of all Libraries and compilers
eg:MPI/PVM, Perl, GCC/CC, Fortran etc

Advanced help from sys admin

- Optimization flags for that particular architecture
- Specific Libraries like FFT (Fast fourier transform)
- Benchmark data if any
- Any support for converting the codes into MPI/OpenMP.

Job submission and Execution

- User may want to know
 - Metascheduler available.
 - The job submission mechanism and job tracking using metascheduler.
 - Local schedulers available in the grid nodes for job submission through RSL.
 - Awareness of Local scheduler commands (eg: PBS, Loadleveller, LSF etc)
 - GridFTP protocol and gsiscp for transferring files across nodes

Intimation required from Sys admin

- Any version change of the Libraries/Compilers
- New Grid node availability
- Backup mechanism – how frequently backup of the data is taken
- If any nodes are blocked, as they are given in dedicated mode for some user and for how much time
- Usage Policies if any.Eg:
 - Number of dedicated nodes users can request
 - Number of hours in dedicated mode
 - Jobs running from command line for more than 72hrs will be killed automatically
 - maximum amount of disk space usage allowed etc

Intimation from Sys admin

- Node upgradation with respect to addition of nodes, memory or disk storage at any location
- **Node non-availability:** If any setup change is going to happen due to which the nodes at the location will not be available for few days
Eg: it happened with chennai node and recently with Pune cluster
- With the new SOA architecture going to be available, the services available on the grid nodes to be listed
- License expiration of any tool eg:MOAB

Thank You