egee

Enabling Grids for E-science in Europe

www.eu-egee.org

Issues from current Experience

SA1 Feedback

to JRA1

A. Pacheco

PIC Barcelona

Motivation

- Goal of this talk: to start the discussions on the grid middleware components.

- Today, some experience has been collected thanks to LCG2.

- LCG has reached this year the experience to operate with the number of sites expected at the start of LHC and half of the number of nodes expected in the Tier-0.

- However, the start-up of the EGEE project and its ROC-based structure bring new challenges.

Trends observed in SA1

- In EGEE SA1 we are observing a change in the profile of the site administrators with respect to LCG.

- SA1 experience is that new sites use untrained or recently recruited people to deploy their initial grid middleware services. Some of them have never installed a Linux cluster before.

- Site administrators expect a complete solution pack with documentation, recipes and point-and-click installation kits. They do not expect to edit, copy, cut and paste configuration data nor search for hidden log files.

- The effort expected to install and run the grid is not persistent. Once installed they expect to run unless something is broken or upgrade is needed. No baby-sitting.

Typical scenario for new sites

- One or more Linux administrators with the mandate to install an EGEE grid site by tomorrow. Specially to appear in the GOC site map!

- One PC with email and web access...

- One printer (for the documentation,...)...

- One web address www.eu-egee.org...

- A number 1..n of PC boxes in a rack, shelf, table or even in the floor.

- ...

What do they require or expect...

- A complete set of documentation in electronic format (introduction, installation, operations, user and reference manuals) for all components.

- A complete set of products with the installation kits associated.

- A secure web interface to apply, register and maintain site information, service information and user profiles.

- An script to install the components in the appropiate order with or without automatic installation tools.

- An script to configure the components for the most common scenarios.

- An script to verify the installation and configuration of the components.

What do they require or expect (2)

- A complete management toolkit with an interface to operate the components and services.

- A standard environment for easy log file examination, debugging, understanding of error codes and problem tracing.

- An automatic tool to update the configuration files and to upgrade to the latest version of the components.

- Access to a professional support channel to get solutions in case of need.

- Access to a technical or operation channel to keep informed of new releases, operational issues, etc...

Preliminary conclusions

- Middleware delivery and requirements must be adapted to the new scenarios in EGEE.

- Clients are not longer computer experts located in HEP institutes.

- In the next transparencies a list of issues concerning current deployment of middleware are shown.

- They issues have been discussed at ROC Management level.

- This requirement list is incomplete. A more detailed list can be obtained from the JRA1-SA1 meetings.

Issues

- Need for easy installation of middleware software
  - Current experience is that the released middleware is very difficult to install and integrate in preproduction testbeds.
  - Dependencies and conflicts with other packages should be clearly identified and eliminated.
  - Use of tools like apt4rpm to convert an rpm repository into an apt, yum or metadata repository should be encouraged.
- Need for easy configuration and verification scripts
  - Some components such as VO Support and RGMA are known to be particularly difficult to configure.
  - A complete verification procedure prior to the request for certification should be provided.
  - Configuration files should be concentrated in one directory or in a single file.
  - Configuration management should be centralized using lcfgng/quattor tools.

Issues (2)

- Need for easy upgrade and site maintenance
  - Continuous effort may be scarce. Once installed, automate as much as possible.

- Need for management tools
  - In order to operate and configure correctly, the services should provide a secure interface to allow remote query and management of their operational state

- Use of coherent naming conventions and procedures
  - Important for fast debugging and inter-component problem tracing

- Need for an improved computer element component
  - Difficulties for the middleware to express job requirements to the job scheduler

Issues (3)

- Need for improvements in the storage element
  - No tools to manage the request queue nor redirect the requests.
  - Any reconfiguration or maintenance becomes a loss of service requests.
  - No possibility in gridftp storage elements to request asynchronous copy from tape to HSM disk cache (pre-staging) before really copying the files. Useful for data challenges. Solved with SRM GET.

- Need for a single support channel for middleware
  - Clarify the support interfaces GGUS, Savannah, LCG-ROLLOUT.

- Need for a realistic plan for deploy gLite
  - Experiments are depending on new middleware but deployment plan is difficult with 80 sites and 7000 computers widely located.

Conclusions

- The key of the EGEE project is the spread of the grid technology worldwide.

- Configuration and management areas should be improved.

- A state machine for operational supervision of components and their interdependencies in a local as well as a global perspective is needed. HEP is very experienced in DAQ systems...

- The discussions of this LCG workshop will be very important to prioritize or refocus the work to be done.