



Enabling Grids for E-science

ROCs Top 5 Middleware Issues

Daniele Cesini, Alessandra Forti
WLCG Operation workshop
CERN, 25th January 2007

www.eu-egee.org



- Following the presentation of the ROC UKI Top 5 m/w issues at June 2006 Operation Workshop, each ROC produced its own list of Top 5 issues.
- All the lists from the ROCs were collected and discussed at the ROC managers meetings
- Each ROC voted 3 issues from that list and this resulted in a final list published in the sa1 CERN wiki:
https://twiki.cern.ch/twiki/bin/view/EGEE/SA1_TCG
- This list has been presented and discussed at the TCG by the sites representatives (Alessandra, Jeff, Daniele) on Oct. 18th and Nov. 29th 2006
- Only voted issues were discussed
- This talk is the result of the follow up of that discussion
- The issues are not presented in voting order

Issue number 1:

“Fine grained list of RPMs per service component (and not only per node/service. Those lists should include all dependencies both internal and external, and should be available in a official and stable url). Does the information at this URL: <http://qlite.web.cern.ch/qlite/packages/R3.0/deployment/> satisfy this requirement?”

SA3: the link reported on the issue will be maintained. For a finer grain list SA3 will investigate the possibilities to use the ETICS facilities.

Issue number 3:

“Reduce dependency of OS: e.g. better support for the UI/WN tarball (it should be certified and released at the same time as the WN/UI RPMs are certified and released.)”

For the specific examples of the UI/WN tarballs, the problem is now fixed and in the future they will be deployed with the standard releases.

About the porting to OS different from SLC, this is a big and complex issue and it is already faced by an SA3 activity that will take all the time needed. ETICS could help also in this. JRA1 is already doing some work in release3.1 in order to have them more portable.

Issue number 4:

“Documentation: release notes should contain all information on: required configuration changes; detailed deployment instructions (for example, must service be re-started after upgrade); bugs fixed by the update; new functionality introduced by the update; etc.”

SA3 says that information are published but probably not properly read by system administrators and site managers. Information about patches and upgrades are published since the upgrade to savannah. The link to this info is not very visible on the glite web portal (glite.org) - it is only in the news section - and will be put in greater evidence. (DONE: <http://glite.org> → Packages).

The possibility of making a pdf with the release notes will be investigated.

Issue number 5:

“New services should come with clear and complete error messages. This should be part of the SA3 checklist (if it isn't already); new components should not be accepted if they don't bring this.”

JRA1: the activity of improving error messages is already ongoing inside JRA1, but it doesn't appear yet in the JRA1 work plan, however the TCG asked JRA1 to add some information.

Issue number 10:

“Implement VO quotas on shared disk pools provided by DPM or dcache. Publish the quota and also available space, per VO, based on this quota.”

The publication of VO available space on a SE will be there with the new glue schema. There is some work on VO quotas on SE, but it has not been a priority, because the priority was the implementation of the srm2.2. At the moment can be used some mechanism such as space reservation, but this not a real VO quota implementation.

For what concerns VO quotas we have to wait at least the srm2.2 implementation that will not be there before spring 2007.

Issue number 12:

“Provide better and clearer log system, with a standard logging format. This should be part of the SA3 checklist (if it isn't already); new components should not be accepted if they don't bring this. Also, the logging system must be able to store the logs on a remote host (as the "syslog" system does).”

The issue of logging standardization has been recognized as a valid and urgent issue. Action raised on JRA1 to implement the syslog functionalities for the internal developed components. For what concern external component JRA1 cannot have any control. Using syslog automatically enables the remote logging possibility.

The syslog implementation for logging is not in the workplan as it a general activity for every component and isn't listed for each one, but there was a discussion at the jra1 all-hands in order to understand if syslog is enough for the jra1 requirements.

Issue number 14:

“Provide service management interface for all middleware services (start/stop/status, etc.). Clarifications: We don't mean "start/stop/status" of linux services. We are actually talking about administration tools which are more or less node-specific. For example, add/remove/supend a VO from a node, add/remove/close a CE queue, close a site (on site BDII), close a storage area, etc. We don't want to launch a node configuration. A good example is the FTS, you can dynamically add or remove a transfer channel by the way of a command. The FTS node is easier to administrate.”

The issue has been received. It is decided that in a first phase it will be implemented a system to monitor all the services and components status inside a node. The use of GridICE sensors will be investigated for the monitoring purposes. As site representatives we clarified that the request is not only for a monitoring tool, but for an administrative tool (i.e. enabling/disabling VO). The development of such tools is postponed to a second phase.

Monitoring issues should converge into the monitoring-dashboard group

Issue number 17:

“Failover in clients and user tools. E.g. enable user data management tools to use redundant BDII to look up information. So if the primary BDII specified by LCG_GFAL_INFOSYS does not respond a backup BDII can be used instead. This will eliminate the need to setup HA BDII service which most sites do not have.”

This is an issue and should be analyzed for any service/client. JRA1 is working on this already (clients of wmpoxy are an example). In the end the most complicated thing is to have a failover mechanism for the site and top level IS. Having that all the other failover mechanisms will be easier to implement.

Issue number 18:

“Pass parameters to LRMS: In order to improve the efficiency of LRMS, some information from user's job description should be passed to the CE through the RB as, for instance, the required amount of memory, the required size of scratch space, the required CPU max. time.”

This is an already ongoing activity of JRA1 and the functionality is present in blah.

Issue number 20:

“Test under real conditions. E.g. make it easier to tailor the MW to specific fabrics needs (e.g. Quattor, batch systems, etc.). **CLARIFICATION** - It means: This is not restricted to regional certification, moreover it is raised as a top issue for the middleware specifically design, development and rollout (SA3 and JRA1). E.g. if a middleware distribution would be designed for one single use case, we had a problem.....”

TCG asked for a clarification the first time we presented it, the clarification was added.

This is not restricted to regional certification, moreover it is raised as a top issue for the middleware specifically design, development and rollout (SA3 and JRA1). E.g. if a middleware distribution would be designed for one single use case, we had a problem.

Issue number 21:

“need clear concepts for maintenance and garbage collection for each persistent/core MW node (like RB/WMS and storage services)”

This was recognized a real issue by everybody. For some items JRA1 will provide guidelines about garbage collection, i.e. LB database. For other items it should be agreed somewhere the period we want to store data i.e. sandbox of not collected jobs, logs etc.

It is still not clear were this agreement should be done (SA1?)

About garbage collection the only guidelines that come from the TCG are those regarding security, so 90 days for logs related to security (see the documents regarding this topic). The rest are policies at the operational level to be discussed inside SA1 and the ROC managers.

Erwin proposed that this topic will be discussed inside SA1 to have a proposal. This proposal should than be discussed at the TCG to see if application are happy with it.

- The list is now in the TCG work plan
- A slot for sites issues will be allocated every couple of months in the TCG agenda
- Although it is very difficult to obtain even broad deadlines from JRA1 and SA3 because:
 - **priorities change continuously depending on the number of problems and bugs found**
 - **the move from SL3 to SL4 and from 32 to 64 bit is more complicated than expected**

It is important that the presentation and voting of m/w issues from ROCs is a continuous process.