



Contribution ID: 94

Type: **not specified**

Achievements and perspectives of the biomed technical team

Thursday, 29 March 2012 14:15 (15 minutes)

Description of the Work

The biomed VO uses resources allocated by several NGIs. Although a resource may be considered up and running from an administrator's perspective, it may not be working as expected for users of a specific VO. From this observation arises the need for a support team to "take place in the VO user's seat", in order to monitor - from their perspective - the resource availability. The biomed support team has been set up with this purpose. It currently consists of eight teams of volunteers from the most active user groups in France, Hungary, Spain, Italy and Viet-Nam. The support is organized in shifts to provide sustained support with a limited burden on contributors. A Nagios box, operated by GRIF, monitors storage and computing resources as well as critical services like the LFC and VOMS. The team on duty applies documented daily tasks and procedures, and submits GGUS "team" tickets when appropriate. It identifies and follows up on issues, discusses salient technical problems and investigate solutions. It also represents VO users to discuss technical requirements with EGI-InSPIRE instances.

During the last 2 years, efforts have been devoted to organize the shifts, coordinate teams take-over, document daily tasks and procedures. Lots of existing tools, services and portals are available to assist the support team in its tasks. A task initiated more recently has been to assess how those tools could be used by the support team, and how reliable the data they provide is with regards to the VO, to avoid misinterpretations. This includes but is not be limited to GSTAT, GOCDB, CESGA, MyEGI, VO Admin Portal, and Operations Portal. A continuous work is also being done to improve existing procedures, develop new tools, extend the scope of the monitored resources, and figure out new metrics that allow to assess resource quality of service for biomed users. For the time being, the latter regards mainly the SEs which are the most critical elements from the user's perspective.

Conclusions

The biomed VO is a significant user of the computing and storage resources that are made available to its users. Resources are monitored, and dynamically and seamlessly extended. As a result, computing power and storage is delivered at a production level.

However, current support efforts mainly deal with generic technical issues that remain a heavy burden. Data management definitely concentrates the most important effort, although substantial efforts are needed on the computing resource side, to develop more accurate monitoring tools, assess the need for more resources or ways to use them more efficiently.

Support concerns remain bound to the infrastructure, to the expense of collaborations within biomed and with other LSGC VOs. As a result, little is invested on application-specific issues, on the sharing of community-specific data, knowledge, tools and experiments, while this is certainly a key potential of grids. This remains an exciting challenge to tackle in the next years.

Impact

Typical life-science applications involve computational workflows consuming/producing large data sets. While a single job failure may be acceptable, a storage resource failure may be a show-stopper, in particular when the middleware does not provide a file replication strategy. Therefore, besides critical services (LFC, VOMS), the initial effort of the support team has focused on Storage Elements (SE). Then, the monitoring has progressively spanned to computing services (CE and WMS). To complement Nagios probes, an on-going development effort consists in defining new ways of identifying potential problems. Again, storage resources are the primary focus with attempts to detect erroneous data published in the BDII, detect full SEs and come up with ways to handle them (identify heaviest users, files not registered in any LFC, etc.).

Since its start in April 2010, the VO support team has handled more than 390 GGUS tickets, that is an average of 5.1 tickets per week. This figure is an indicator of the high operational cost for an international VO. Over the last year the biomed VO consumed 12 millions normalized CPU hours. Usage data history shows that the average ratio of waiting jobs over running jobs is 3.5, causing long delays in the jobs queues; possible reasons are currently investigated. During the same period, the used storage space has raised from approximately 1.2 PB to 2 PB, out of 3.6PB.

So far, collected experience has helped identify features that would improve support quality and efficiency, but that evaluated tools do not address. Some features are addressed by the LSGC (redundant VOMS server), while some are being discussed with the User Community Support Team (redundant LFC server, VO Operations Dashboard). The remaining features have been gathered in the specification of the LSGC Dashboard, still to be developed, with features such as the management of VO users life-cycle workflow, support for robot certificates, and VO specific accounting metrics.

URL

<http://wiki.healthgrid.org/Biomed-Shifts:Index>

Overview (For the conference guide)

The Life-Science Grid Community (LSGC) gathers five Virtual Organizations (VOs) related to the Life-Science field. Among those, the biomed VO has set up a technical support team to be the technical interface between VO users and the NGI sites providing computing, storage and infrastructure services.

The support team is a front-line for handling requests from sites and users; it monitors the resources, and enforces a set of pro-active measures to improve the service quality. It liaises with EGI-Inspire instances (UCB, UCST) to report needs, share experience, and learn from other communities' experience and best-practices. This abstract presents those goals in more details, the achievements of the past year, and highlights current actions and challenges. Overall, the handling of technical issues in an international VO still requires substantial manpower, to the expense of domain-specific activities. Exciting challenges remain to bring the full potential of grids to the end users.

Primary author: Mr MICHEL, Franck (CNRS)

Co-authors: Dr MONTAGNAT, Johan (CNRS); Dr GLATARD, Tristan (CNRS)

Presenter: Mr MICHEL, Franck (CNRS)

Session Classification: Virtual Research Communities (VRCs)