

Minutes of Information System Meeting with users, 4th November 2012

Present from IT-ES: Alessandro di Girolamo (AG), Maarten Litmaath (ML), Stefan Roiser (SR), Andrea Sciaba (AS, minutes).

Present from IT-GT: Maria Alandes IT-GT (MA, chair and minutes), Laurence Field (LF)

Present from EGI: Stephen Burke (SB)

Present from OSG: Brian Bockelman (BB), Catalin Dumitrescu (CD)

Agenda available in Indico

<https://indico.cern.ch/conferenceDisplay.py?confId=220247>

1. Introduction

MA presents the goals of this meeting which is to have a place for discussing IS changes and issues with experiments, experts and EGI/OSG representatives. The idea is to meet once per month and define an action list which will contribute to have a closer and regular collaboration.

2. Monitoring the IS

- MA presents that nowadays there are no concrete actions to validate and monitor the published information. That will change for GLUE 2.0. glue-validator (<https://tomtools.cern.ch/confluence/display/IS/Glue+Validator+Guide>) is already there but needs to be extended since it performs only some generic checks. There is also the gstat-validate-sanity-check probe (<https://svnweb.cern.ch/trac/gridinfo/browser/gstat-validation/trunk/bin/gstat-validate-sanity-check>) running in SAM which checks some limited GLUE 1.3 attributes. The error messages raised by this probe are sent to the EGI operations dashboard where SAM notifications are manually monitored and GGUS tickets are opened if necessary. The criteria to open GGUS tickets are not clear and presumably no tickets are opened for gstat-sanity-check error messages. This operational workflow needs to be better understood and will be confirmed with EGI.
- ML points out that glue-validator should indeed run in the resource BDII so that we make sure invalid information is not even published. A discussion takes place since it is sometimes not clear whether an attribute value is valid or not.
- MA continues the presentation mentioning the Profile document for GLUE 2.0 written by SB (<http://go.egi.eu/glue2-profile>). This document will be the base to improve the glue-validator. The idea is to replace the gstat-validate-sanity-check with glue-validator in SAM. ML reminds again that the best approach would be to run glue-validator before publishing. LF agrees that this could be implemented, integrating the glue-validator within the BDII and being more strict by not publishing those objects that do not conform with the schema.

- LHCb use case with wrongly published GlueCEPolicyAssignedJobSlots and GlueHostBenchmarkSI00 is discussed. Ideally this should not be monitored by the experiments but it should be done automatically as part of EGI operation activities.

Action items

- *Improve glue-validator implementing the EGI profile document for GLUE 2.0 (MA)*
- *Replace gstat-validate-sanity-check with glue-validator in SAM (MA)*
- *Understand operational workflow to open GGUS tickets out of SAM errors and see how this could be used for glue-validator errors (MA)*
- *Integrate glue-validator in the resource BDII (MA)*

3. GLUE 2 plans

- Since BB has to leave, this item is moved earlier in the agenda. MA presents the advantages of GLUE 2.0 and the current deployment status in WLCG. EGI sites continue to ramp up. It is confirmed that everything needed to exploit GLUE 2.0 is in place. AS asks whether EMI WMS is also GLUE 2.0. MA will provide a list of the EMI services which can publish GLUE 2.0. 100% deployment would be needed to start considering using GLUE 2.0 and then, probably existing info providers and clients may show in practice some glitches that would need to be fixed.
- BB confirms that there is no requirement from US-CMS and US-ATLAS to use GLUE 2.0 and therefore there are no plans to publish in GLUE 2.0. AG asks whether it is clear how much effort would be needed to publish in GLUE 2.0. BB states that clear use cases need to be identified before the effort could be estimated. ML asks about other stakeholders and BB answers that there are some smaller VOs which use WMS. LF informs that the experience of enabling GLUE 2.0 in the EMI project was very positive, since more than 20 development teams were able to do this without major problems. BB mentions that he personally understands what needs to be done from the technical point of view but that he is waiting for a request coming from the stakeholders. MA asks about interoperability and BB answers that the needed level of interoperability is currently provided with GLUE 1.3. SB mentions that any future changes will only be done in GLUE 2.0 and not in GLUE 1.3 though.
- ML wraps up the discussion mentioning that EGI will gain experience with GLUE 2.0, use cases will be clarified and GLUE 2.0 will prove itself in the upcoming months, something which hasn't happened yet. The item should be kept opened.

Action items

- *Confirm the list of EMI services able to publish GLUE 2.0 information (MA)*
- *Estimate the effort needed by OSG to be able to publish the complete GLUE 2.0 schema (BB)*

4. ginfo

- MA presents ginfo and the fact that it is a client tool focused on the service discovery use case. MA also mentions that there has been already some feedback from AS to support other use cases. LF asks who actually uses tools like lcg-info and lcg-infosites. Experiments reply that many people are using these tools and they are very useful for them because they hide the complexity of ldapsearch which is not easy to learn. AS, ML and SR give examples of how these client tools are used within their experiments. LF asks whether experiments would be interested in extending ginfo, having a new tool for the other use cases or having a GLUE 2.0

version of lcg-info/lcg-infosites. AS replies that this is not relevant as long as the functionality is present. AS, AG, SR and ML agree to compile a list of the most common queries they perform that they would like to see implemented in ginfo. This includes lcg-info/lcg-infosite queries and ldapsearch queries performed by them or any other examples regularly executed within the experiment.

Action items

- *Provide a list of most popular lcg-info/lcg-infosite use cases and ldapsearch queries. (AG, ML, SR and AS)*
- *Provide a development plan to improve ginfo with the implementation of the provided uses cases (LF)*

5. AOB

- AS asks what is the status of the proposed plan to change the IS (static vs dynamic vs metadata) as presented in the June GDB (See Information System Status and Evolution <https://indico.cern.ch/conferenceDisplay.py?confId=155069>). LF replies nothing has been done yet since we first need to identify which are the use cases: what information is important for the experiments. For instance, if dynamic information is important, this could be provided via messaging. This will require a major effort that it is necessary to understand how much this is needed. AG asks whether it would be possible to separate static information from dynamic information by publishing more static data in the GOCDB. LF agrees that this could be a possibility defining policies for what can be published and how it is validated.
- SR asks whether the glue-validator could help to improve attributes like installed capacity and system architecture since it seems to him that sites are not publishing this properly. ML reminds that there are very complete documents provided by EGI where this is explained in detail (https://wiki.egi.eu/wiki/Administrator_Documentation).

Action items

- *Prepare one slide describing the ideal information system for each experiment. How would you like to access the information? What is your ideal information system to be integrated with your experiment needs? It could be very far from current reality. The idea of this exercise is to compare with the existing solutions and identify areas for improvement and provide hints on future evolution. (AG, ML, SR and AS)*

6. Next meeting

January 2013