Minutes of Information System Task Force, 24th September 2015

Local: Maria Alandes (chair, minutes), Julia Andreeva, Laurence Field, Alessandra Forti, Alessandro di Girolamo, Andrea Sciaba, Andrea Valassi.

Remote: Stefano Belforte, Brian Bockelman, Stephen Burke, Tom Byrne, Balasz Konya, Andrew McNab, David Meredith, Stuart Pullinger, Rob Quick, Oxana Smirnova, Vincenzo Spinoso.

Agenda available in Indico

https://indico.cern.ch/event/441747/

---

1. Next Steps

Maria Alandes proposes a schedule to prepare the Future Use Cases document to be presented in October GDB. Andrea Sciaba explains that CMS is planning to discuss all this in the Offline and Computing week that will take place the same week as the GDB (12-16 October). Alessandro di Girolamo also mentions that this will take time as they also want to cover non-grid resources, like HPC or cloud, this means that many people in ATLAS will have to contribute and it will take longer. Maria agrees to delay the proposed deadline and try to present the document in November GDB. Everyone agrees that this is a more realistic deadline.

Maria Alandes asks for feedback on Alessandra's proposal already discussed by mail (see slide). Alessandro di Girolamo mentions that as long as ATLAS is able to retrieve the necessary CE and SE information, it is fine with them. Maria proposes to listen to the scheduled presentations that will help to better shape Alessandra's proposal.

Action items:

- *Send a new schedule to prepare the Future Use Cases document to be presented at the November GDB (Maria Alandes)*
- *Propose a plan to implement a first prototype of Alessandra's proposal:*
    - *Investigate which information would be needed in GOCDB/OIM to be used as service registries (Maria Alandes together with experiment liaison and other activities using the IS).*
    - *Understand what would be needed in GOCDB/OIM to provide the required information (new fields? how will the information be obtained the first time? Only static?) And in which format (GLUE 2? Json?) (Maria Alandes with David Meredith and Brian Bockelman)*

- *Understand whether there is any dynamic information needed and whether it would make sense to query directly the resource BDIIs (Maria Alandes with the experiments where this use case makes sense)*
- *Organise a plan together with experiments and other activities to start querying GLUE 2 instead of GLUE 1 (Maria Alandes)*
- *Write a first version of the WLCG profile for GLUE 2 and implement it in glue-validator for LDAP (Maria Alandes)*
- *Extend glue-validator to be able to also validate other formats, like json or xml (Maria Alandes)*
- *Re-open the discussion on running glue-validator at the resource BDII level (Maria Alandes to contact developers, maybe URT for EGI, Brian for OSG).*

2. OSG Future plans for the Information System

Brian Bockelman presents OSG plans to get rid of the BDII and move towards OSG collectors expressing information in ClassAds. Brian also presents the ClassAd-GLUE 2 translator that is already being implemented in collaboration with CERN IT-PES group. Brian explains that OIM is publicly accessible through MyOSG which exports XML in a custom schema.

Andrea Sciaba asks whether the OSG collector focuses only on HTCondorCE, and Brian confirms that this is the case. Oxana Smirnova asks why ClassAd language is difficult to be translated to GLUE 2 if it's schema-free. Brian explains that ClassAd language is much richer in semantics than *json* or *ldap* and this makes it a bit more difficult to translate. Andrea Sciaba reminds that GLUE 2 has the *GLUE2EntityOtherInfo* attributes that could match any key-value pairs that are not directly defined in GLUE 2 attributes. Maria Alandes asks about the state of the ClassAd-GLUE 2 translator, and Brian explains that this is still a prototype. More effort is needed to finish implementing all the use cases, improve its reliability, etc. Julia Andreeva asks whether the OSG collector mentioned at the presentation is the same one as the collector used to create the CMS VOfeed. Brian explains that it's something different.

Action items:

- *Contact IT-PES to know more details on the ClassAd-GLUE 2 translator (Maria Alandes)*

3. EGI Future plans for the Information System

Vincenzo Spinoso presents EGI plans to keep on supporting BDII and GLUE 2. There is no more development in GLUE 1. It will be published as long as WLCG needs it but it would be good to start planning a transition.

Andrea Sciaba asks whether validation of information interesting for WLCG will be done by EGI. Maria Alandes explains that the glue-validator currently implements the EGI profile to validate information published by EGI sites. WLCG will have to define its own WLCG profile. In principle it should be a subset of EGI profile, although WLCG may want to be stricter in certain allowed thresholds or values. Sites are ultimately responsible for fixing the information but WLCG should benefit as much as possible from the existing automatic validation mechanisms that EGI has put in place.

4. NDGF Future plans for the Information System

Oxana Smirnova presents the different types of schemas currently supported by NDGF: the Nordugrid schema and GLUE 2. Oxana explains that things will be much simpler if only GLUE 2 is supported.

Andrea Sciaba asks whether there is any client to query ARC resources. Oxana says that nothing special is needed, any IS client would work, like ginfo.

5. GOCDB status and future developments

David Meredith presents GOCDB internal technology and a series of new features that could be interesting for the TF. Tags are now available to be associated to each entity in GOCDB. For instance, this will allow tagging all VO resources or all WLCG resources. It is also possible to define any key-value pairs. GOCDB offers a REST API (read only) to be able to export *XML* and *json*.

Maria Alandes asks how many people are supporting GOCDB. David answers that it's mainly him together with Tom Byrne. This is developed and also managed by RAL. Laurence Field explains that GOCDB could make use of a service crawler to gather information from resources in an automatic way. This could be a starting point to gather information for a service registry in GOCDB. Laurence also asks whether GOCDB is open source and people can contribute to the code base. David answers that this is indeed the case.

6. Next meeting

Not decided. To be discussed in the ML. Likely in one month.