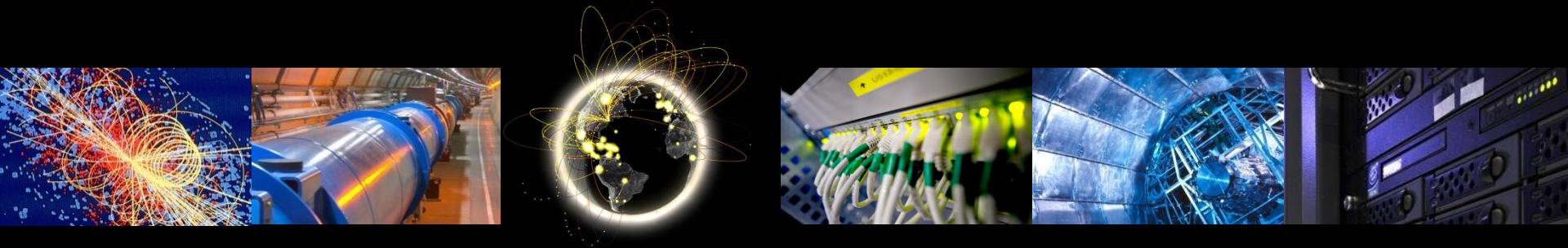


Information System WLCG Issues

Markus Schulz

11 Jan 2012



Overview

- What happened in the last 1.5 years
- What is happening soon
- Current problem

Last 1.5 Years

- WLCG recognized the need of the role of a WLCG Information System Officer
 - The Information System is more than GLUE, Resource BDIIs, Site BDIIs, Top Level BDIIs and clients.
 - Complex problem propagation through the system
 - Quality of data, versions, etc.
 - Communication with users to understand use cases etc.

Information System Officers

- The two consecutive officers followed the same approach (Flavia Donno and Lorenzo Dini):
- Identify, clarify and document current and future use of the information system by the experiments (directly and indirectly)
- Identified and documented main problems of the current Information Systems
- Coordinate short and long term solutions
- Follow up on operational issues
 - to shield developers from trivial config issues etc.

Main Use Cases

- Collected, discussed and documented
 - presented last in June at the GDB
- Main points:
 - Clear shift from on the spot matching to available resources to using it as the source for service discovery
 - Additional information added by the experiments
 - maintained in their own frameworks

Main Problems

- After some iterations and discussions:
 - No major issues with the schema
 - that can't be handled with GLUE-2 or small changes to 1.3
 - The building blocks have no major software issues
- But:
 - System config. doesn't match the current usecases
 - focus on data freshness / stability of data
 - what is there now / what should be there
 - Data Quality
 - wrong or incomplete data being published
 - Data Bloat (load balancing CEs multiply the data volume)
 - Deployment (very different speed in following releases)
 - sometimes more recent versions of LDAP upstream break when fed ill formed data

Improvements

- Not all new, focus on cost/benefit
- BDII → Caching BDII
 - Developed, Tested, Piloted, Packaged
 - Used on a few sites only
- Client configuration to use reliable Top-BDIIs
 - List of agreed Top-Level BDIIs to use
 - Not very successful: 55 sites still use CERN directly...
- Data Quality
 - lead to the development of the Glue Validator
 - can be turned into a sensor for SAM
 - Requires operational follow-up on tickets

Improvements II

- Data Bloat: Focus on CE
- Deprecation of the Glue Location Object
- Release of gLite-Cluster component

- Not rolled out on many sites
 - → no measurable impact

Now/Soon

- The Workload Management TEG
 - collection of current use cases
 - no surprising differences
 - maybe we should reuse the work of the Info Officers
 -
- EMI on planned and communicated trajectory
- ...

Current Problems

- The data volume keeps increasing
 - December 2011 160MB
 - November 2010 102MB
 - Mid 2007 28MB (stress tested to 50MB)
- Operational/config. problems are dealt with on an case by case basis
 - involving mostly a small band of volunteers
 - escalating rapidly to the developer level
 - level-3 support doing often level 1-2 work
- Improvements are not rolled out, no follow-up
- GLUE-2 introduction driven by EMI only

What can be done?

- Existing improvements should be rolled out
 - Maybe as we did it with the Cream-CE/glexec on WN. Action items for the WLCG-MB
 - Caching BDII, Client Config, glite-Cluster, GLue-Validator
 - Main work is to propagate and follow up
- Ensure that work on Use Cases is not lost
 - Presentation to the TEG or appendix to report?
- Align TEG and EMI work on info systems
 - EMI representation in the TEG?
- Less clear: Operational Support for the infosys.