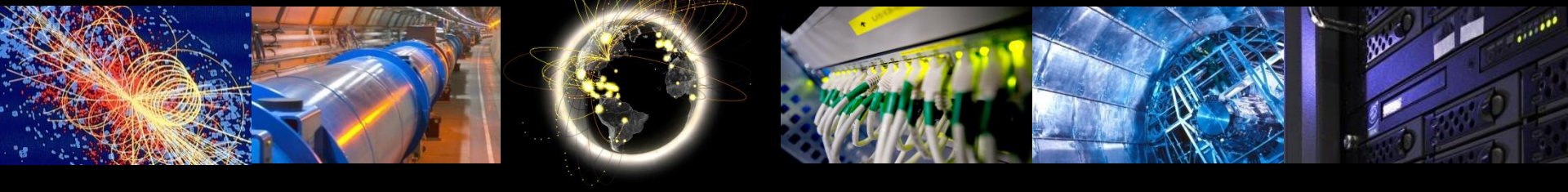
WLCG Information System Status

Maria Alandes Pradillo, CERN

CERN IT Department, Support for Distributed Computing Group

GDB 9th December 2015



WLCG Worldwide LHC Computing Grid

Contents

- BDII development and issues
- Information System Evolution TF
- Future Use Cases Document

BDII development

- No releases since September 2014
- CENTOS7 BDII successfully tested
  - All flavours tested in production (resource, site and top BDII)
  - To be included in UMD-4 this month

BDII issues

- slapd process crashing in top BDII and ARC CEs
  - When upgrading to to SLC6.7/CentOS 6.7
    - Includes a new version of openLDAP: openldap-servers-2.4.40-5
    - WLCG suggests sites not to upgrade to this version of openLDAP
  - Fix provided by Redhat
    - Tested successfully in top BDII and ARC Ces
    - Requested to Redhat to release the new version

Information System Evolution TF

Progress made so far

- Lots of discussion in mailing list and meetings
- Input from EGI, OSG, NDGF, experiments and GOCDB
- IS Use Cases

  https://espace.cern.ch/WLCG-document-repository/Technical_Documents/WLCGISUseCases_1.4.pdf
- IS Future Use Cases

  https://espace.cern.ch/WLCG-document-repository/Technical_Documents/WLCGFutureISUseCases_1.6.pdf

GLUE 2.0

- Sites are currently maintaining two schemas
  - GLUE 1.3
  - GLUE 2.0
- Information providers are also maintaining both schemas
  - Even if in most cases code is reused
- Simplification is needed!
- GLUE 2.0 is the natural evolution
  - Moreover, cloud resources could be defined in GLUE 2
    - Still under discussion within the WG whether this will be included in GLUE 2.1 as an extension of the current schema
- We need to give a step further and make some progress in this direction
  - Resource information is in GLUE 2.0 for a while
  - It's time to start consuming this information

GLUE 2.0

- EGI is planning a transition to GLUE 2.0 for operations tools and monitoring
  https://indico.cern.ch/event/462599/contribution/2/attachments/1195163/1736449/EGI-GLUE2_status.pdf
- OSG is willing to provide a script for those WLCG applications that rely on GLUE 2.0 in JSON-formatted descriptions
  - There has to be a clear motivation of the need
- Experiments have expressed the feasibility of consuming GLUE 2.0
  - It should be pretty easy in most cases
  - But they are waiting for an official decision to put effort on this
- We need an official decision to make further progress
  - Could WLCG agree on having GLUE 2.0 as the official schema to publish and consume information?
  - Can we therefore plan the decommissioning of GLUE 1.3 information?

Definitions

- The task force is working on better definitions so that sites publish information in an homogeneous way
  - For the time being we are focusing on
    - HEPSPEC06: GLUE2ExecutionEnvironmentLogicalCPUs
    - Logical CPUs: GLUE2BenchmarkValue

    https://twiki.cern.ch/twiki/bin/view/EGEE/GLUE2WLCGRoadmap#Definitions
  - Feedback from sys admins is being collected
  - Aligned with MJF TF
- Once definitions are clear and everyone has agreed on them
  - They will be properly documented in a twiki or document
  - Sites will be monitored and validated according to the definitions

Future Use Cases Document

Summary

- A reliable central information system owned by WLCG and presenting information in an homogeneous way is needed
  - Service topology
  - Installed capacities
- Information is primarily collected from OSG, EGI, NDGF for grid and non grid resources (like clouds and HPC)
  - Other information sources may be added if needed
  - MoU and non MoU sites should be included
- Information should be cached and validated
  - Information is correct before it gets published
- Logging information should be recorded as well
  - When, how and by whom information was provided

Proposal

- Develop a WLCG Information System prototype based on AGIS
  - It already implements collectors to gather information from different sources (GOCDB, OIM, BDII, REBUS)
  - It provides caching and validation mechanisms
  - It records logging information
  - It is also a tool successfully used by ATLAS based on well known technology (Django, python)
    - No need to implement a new tool from scratch!
- AGIS details and how it could be used as a generic information system were presented at the TF are available here:
  https://indico.cern.ch/event/454975/contribution/2/attachments/1186493/1720358/2015.11.12.ISTF_AGIS.pdf
- Further discussion and decisions on the different features that need to be available could be still discussed within the TF
  - i.e. the output of the ongoing discussion on definitions will be used to implement validation criteria, etc.