

Use of the IS by WLCG experiments

The baseline for usage of the IS in WLCG comes from the document *WLCG Information System Use Cases* [1].

Updates were then provided primarily at the May 2011 GDB [2].

ALICE

Each ALICE site has an ALICE VO-box installed within the site boundary and responsible for job submissions to local CEs. It queries the resource BDII of each CE to regulate the flow of jobs to the site through the pilot-based AliEn framework. The list of CEs to be used is hard-coded in the AliEn LDAP service. The BDII is also used to identify which CEs are in production mode, looking at the CE status (only production CEs are used to submit jobs).

For SAM tests ALICE depend on the WMS, which relies on the top-level BDII.

ATLAS

ATLAS maintains a cache of experiment-specific info for its software components. (PanDA, dashboards, etc.) This info is collected from the BDII, but also from other sources (GOCDB, the OSG Information Management System, other services) and involves static and quasi-static information like downtimes, queues being set offline, blacklisted sites. BDII is used by PanDA to discover and keep current the list of known endpoints/sites. The most dynamic part is SoftwareRunTimeEnvironment and the status of CEs.

BDII is periodically scanned and its info cached. ATLAS maintains a site configuration db in Oracle site attributes may be added to this database when needed; for example, recently, fields to control many-core queues were defined.

A common source of problems is related to the publication of disk space (it would be desirable to know how much is *in use* and how much is *available* instead; this is not necessarily coincident with how much is *installed*).

ATLAS relies on services such as FTS and, for the time being, the WMS (SW installation/validation jobs, SAM tests); these services query the BDII, which must therefore work reliably.

CMS

The BDII information used by CMS is quasi-static. For example, in CRAB (WMAgent) there are queries for CE status, SoftwareRunTimeEnvironment, CEUniqueId or Close SE [static match for inclusion/exclusion], OS version, but typically with a trust but verify model. For pilot factories, the list of sites is not automatically updated based on BDII info. Site attributes are not auto-updated (and, like ATLAS, CMS may define custom site attributes).

The requirements to the IS are for relatively basic items, which should be easy for the sites to operate and not error-prone. CMS does not use dynamic info (slot utilization, system usage), which were found frequently unreliable / out-of-date. CMS suggests that it is better to have a fast, simple, reliable, quasi-static IS. It is also questionable how much benefit pilots would gain from dynamic info. For example, CMS validates nodes directly before the glideins start on a given node. This avoids most black-hole problems.

CMS also uses services like FTS and WMS, which rely on the BDII.

LHCb

DIRAC, the LHCb software framework, does not basically use the BDII and incorporates its own workload management system. Endpoint info (e.g. list of CEs) is statically defined in the DIRAC Configuration Service.

Like others, LHCb uses services like FTS and WMS, which rely on the BDII.

WMS

Requirements and ranking may be specified in the JDL to an extent that allows the WMS to work without querying the BDII. The matchmaking and its corresponding IS dependency can also be bypassed by providing the "-r" option with the designated CE as argument to the job submission command. Furthermore, the WMS can use its *replanning* feature [3], used to remove a job from a queue after some timeout, and to automatically resubmit it to another queue, to build its own resource ranking without querying the BDII.

Conclusions

WLCG experiments all have developed (totally or in part) their own software frameworks and tend to use the BDII for static/quasi-static information, and in general for limited purposes; the general pattern is often use for bootstrap, then refine with our own heuristics .

Quality control of the IS content is important and needs to be automated. WLCG experiments have learnt by experience that no info is at least not worse than unreliable dynamic information. Reliable storage information would be certainly desirable, but it is currently not available. Having cached info in the IS is considered to be vital to overcome possible services outages.

Other, future services related e.g. to the integration of Cloud resources might possibly use the IS; however, it is still early days in that area and it is difficult to draw conclusions.

It is also not clear to what extent a (more) dynamic IS would benefit pilot-job based frameworks, based on late-binding of jobs to slots. It is likely that in the future WLCG experiments will continue to need mostly a simple discovery service.

A summary of initiatives and ideas for a common service registry across heterogeneous infrastructures can be found in [4].

Notes

[1] WLCG Information System Use Cases,
https://twiki.cern.ch/twiki/pub/LCG/WLCGISArea/WLCG_IS_UseCases.pdf

[2] May 2011 GDB, <https://indico.cern.ch/conferenceDisplay.py?confId=106644>

[3] EMI 1 WMS v.3.3.0, http://www.eu-emi.eu/products/-/asset_publisher/z2MT/content/wms

[4] Towards an Integrated Information System, Amsterdam Dec 1, 2011,
<https://www.egi.eu/indico/conferenceDisplay.py?confId=654>

-- DavideSalomoni - 03-Feb-2012

This topic: LCG > WMTEGInformationServiceSummary

Topic revision: r2 - 2012-02-03 - MaartenLitmaath



Copyright &© 2008-2021 by the contributing authors. All material on this collaboration platform is the property of the contributing authors.
or Ideas, requests, problems regarding TWiki? use [Discourse](#) or [Send feedback](#)