

Consolidating WLCG topology and configuration in the Computing Resource Information Catalogue

Maria Alandes¹, Julia Andreeva¹, Alexey Anisenkov^{2,8}, Giuseppe Bagliesi³, Stephano Belforte⁴, Simone Campana¹, Maria Dimou¹, Jose Flix⁵, Alessandra Forti⁶, A di Girolamo¹, Edward Karavakis¹, Stephan Lammel⁷, Maarten Litmaath¹, Andrea Sciaba¹ and Andrea Valassi¹

¹ CERN, Geneva, Switzerland

² Budker Institute of Nuclear Physics, Novosibirsk, Russia

³ Universita di Pisa and INFN, Pisa, Italy

⁴ Universita di Trieste and INFN, Trieste, Italy

⁵ PIC, Barcelona, Spain

⁶ University of Manchester, Manchester, United Kingdom

⁷ Fermi National Accelerator Laboratory, Chicago, USA

⁸ Novosibirsk State University, Novosibirsk, Russia

Abstract. The Worldwide LHC Computing Grid infrastructure links about 200 participating computing centres affiliated with several partner projects. It is built by integrating heterogeneous computer and storage resources in diverse data centres all over the world and provides CPU and storage capacity to the LHC experiments to perform data processing and physics analysis. In order to be used by the experiments, these distributed resources should be well described, which implies easy service discovery and detailed description of service configuration. Currently this information is scattered over multiple generic information sources like GOCDB, OIM, BDII and experiment-specific information systems. Such a model does not allow to validate topology and configuration information easily. Moreover, information in various sources is not always consistent. Finally, the evolution of computing technologies introduces new challenges. Experiments are more and more relying on opportunistic resources, which by their nature are more dynamic and should also be well described in the WLCG information system. This contribution describes the new WLCG configuration service CRIC (Computing Resource Information Catalogue) which collects information from various information providers, performs validation and provides a consistent set of UIs and APIs to the LHC VOs for service discovery and usage configuration. The main requirements for CRIC are simplicity, agility and robustness. CRIC should be able to be quickly adapted to new types of computing resources, new information sources, and allow for new data structures to be implemented easily following the evolution of the computing models and operations of the experiments.

1. Current WLCG Information System Architecture

The WLCG information system [1] is a mission-critical component in the WLCG grid infrastructure [2]. It provides detailed information about grid services which is needed for various different tasks. As represented in **Figure 1**, currently the grid information system has a hierarchical structure of three levels.



The fundamental building block used in this hierarchy is the Berkley Database Information Index (BDII). Although the BDII has additional complexity, it can be visualized as an LDAP database. The resource level or core BDII is usually co-located with the grid service and provides information about that service. Each grid site runs a site level BDII. This aggregates the information from all the resource level BDII's running at that site. The top level BDII aggregates all the information from all the site level BDII's and hence contains information about all grid services. There are multiple instances of the top level BDII in order to provide a fault tolerant, load balanced service. The information system clients query a top level BDII to find the information that they require.

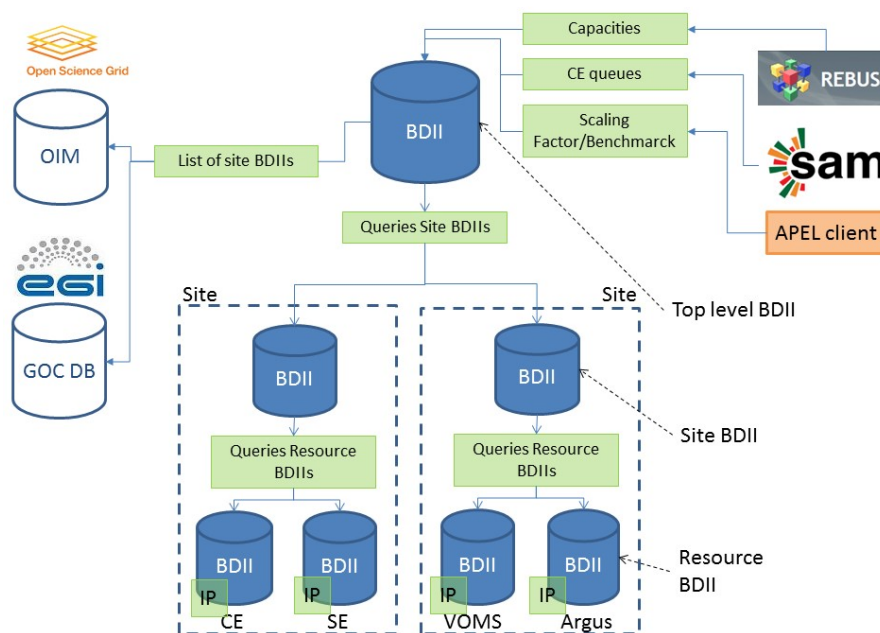


Figure 1. WLCG Information System

The BDII's are populated with information by running information providers. These are scripts which obtain information, format it as LDIF and print the result to standard out. These information providers can also be used to query other BDII's which is how the hierarchy is built. The order in which these information providers are run is random.

The information in the information systems conforms to a schema called the GLUE schema [3] [4]. The GLUE schema started as collaboration effort between European and US grid projects to facilitate interoperation between them. The Open Grid Forum (OGF) is now responsible for the GLUE schema.

The information system is bootstrapped from the information registered in the Operations Databases of EGI and OSG grid infrastructures (GOCDB [5] and OIM [8]). When a site registers, it enters the URL for the site level BDII into the GOCDB/OIM. GOCDB/OIM then generates a list of LDAP URLs for all the sites in the grid and this is downloaded by the information provider running on the top level BDII. These URLs are then used to query all the site level BDII's and the result is used to populate the top level BDII.

1.1. Weaknesses of the Current WLCG Information System

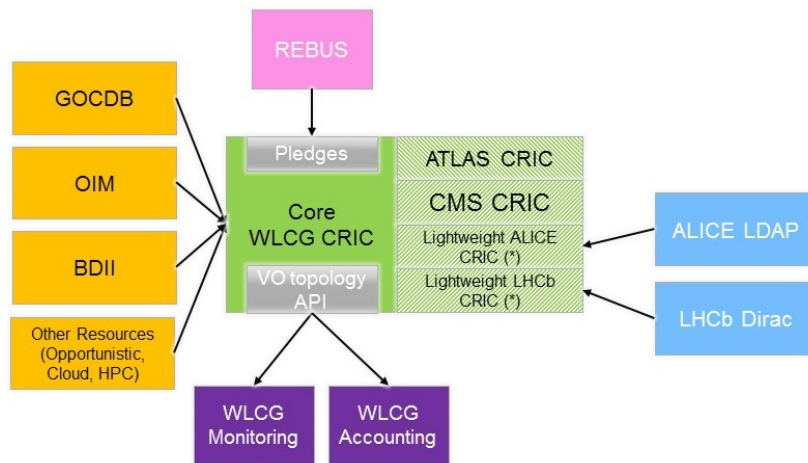
The WLCG Information System presents a set of weaknesses as explained in the following points:

- Lack of flexibility: the main building block of the WLCG information system is the BDII. It's not possible to integrate information from other systems without a major restructuring of its architecture. OSG has decided to stop relying on the BDII as of March 2017. This means that the current information system won't be able to cope with this change, providing only partial information for the overall set of WLCG resources.
- Lack of reliability: information is scattered over multiple information sources like GOCDB, OIM, BDII and experiment specific tools. None of these tools is owned by WLCG making it very difficult to ensure the quality of the information and to influence the policies that manage how the information is described, modified and maintained.
- Incomplete description of available resources: current information is basically coming only from grid resources. However, experiments are more and more relying on other resources like cloud and HPC and currently there is no easy way to integrate these resources in the information system.
- Lack of unified view for experiment topologies and configurations: the current information system is lacking a mechanism to allow experiments to describe their internal topology and configurations. For this reason, each experiment has developed their own information system, duplicating efforts and adding an extra layer of complexity to get experiment specific topology information.

For this reason, a task force dedicated to study the evolution of the information system was set up in September 2015. The goals of the task force were to evolve towards a new information system that would be more flexible, more reliable, more complete and that will provide a single entry point to understand experiment topology.

2. The Computing Resource Information Catalog (CRIC)

The Computing Resource Information Catalog will describe WLCG topology and will be the entry point to consume information about WLCG resources. As represented in **Figure 2**, it is comprised of a core module that is integrated with experiment specific modules, as described further in this paper. The core module consumes information from different information sources and is flexible enough to add or remove information sources and even allow sites to enter directly information about their resources. This flexibility enables a complete picture of not only traditional grid resources through their information sources, like BDII, GOCDB or OIM, but also other resources like Cloud or HPC. Moreover, opportunistic sites do not need to be part of GOCDB or OIM anymore, nor run a BDII to be able to be described in CRIC. This offers a major advantage to small sites who don't have the effort to run extra services.



(*) Maintained by WLCG to store very simple experiment topology information (i.e. experiment names)

Figure 2- Computing Resource Information Catalogue with Information Sources

Experiments like ATLAS and CMS are planning to be fully integrated with CRIC and to provide via CRIC all required configuration for their data management and workload management systems. ALICE and LHCb for the time being are not interested in running an experiment CRIC. In order to provide a single entry point for WLCG topology, lightweight experiment CRIC instances are also envisaged for ALICE and LHCb. Lightweight experiment CRIC instances for ALICE and LHCb will define the set of sites and services which are used by these experiments. This offers the basic functionality which is required for WLCG testing, monitoring and accounting systems. ALICE and LHCb lightweight CRICs will retrieve the required information from the existing experiment topology systems.

With CRIC, WLCG will own the information and will manage the policies to add, remove and modify information, enforcing also some quality criteria to make sure the information is reliable. CRIC will also provide log information about who has modified information, what information has been modified and when it has happened.

Moreover, one of the important operational aspects is who is allowed to do what depending on the role of the particular member of the virtual organization. User role should define read and/or write privileges on a pretty low level, as low as a particular CRIC object instance. For example, the coordinator of the experiment data management team might be allowed to perform any actions on all instances describing storage services defined in the experiment CRIC. While a member of a particular site team supporting a particular experiment, might need read and write access to all instances describing all services hosted by this site. Fine grain access control management system will be based on the federated identity and single sign-on authentication.

2.1. CRIC Architecture

CRIC was inspired by the Atlas Grid Information System (AGIS) [7] which was designed to integrate configuration and status information about resources, services and topology of the ATLAS computing infrastructure. Experience accumulated during AGIS development and many years of successful use in ATLAS is being taken on board for CRIC design and development. Unlike AGIS which has been focused on the needs of a single experiment, CRIC is being developed having in mind more generic use cases. The system should be able to be used on a global WLCG scope, providing a cross-experiment view for the WLCG operations, monitoring and accounting, as well as experiment specific view to

describe in detail experiment computing and operations models. Despite this difference, both systems follow the same paradigm of decoupling physical resources from their logical organization inside experiment computing.

AGIS and CRIC are implemented in the Django [6] open source high-level Python Web framework which follows the model-view-template (MVT) architectural pattern.

2.2 Core CRIC

The purpose of the core CRIC is to describe physical services hosted by the distributed computing sites which are part of the WLCG infrastructure. These services represent resources which are *provided by the WLCG infrastructure* in contrast with the additional configuration information described in the experiment CRICs and indicating how resources are *used by the experiments*. Services and sites are normally declared in the systems like GOCDB and OIM and all this topology description is being periodically collected into core CRIC. Though most of this information is static, there is also dynamic information such as downtime declaration which is performed via GOCDB and OIM and is being imported into core CRIC.

2.3 Experiment CRIC

Experiment CRIC provides experiment specific configuration information for the resources used by the experiments for data storage, data distribution and data processing. It contains all necessary information for organization of the data management and workflow management activities and models the experiment specific concepts. Therefore, experiment CRIC serves the experiment data management and workload management systems as well as various operational tools, monitoring and accounting systems. It plays a key role in the information flow of the experiment offline computing.

Objects described in the experiment CRIC reference objects defining physical services contained in the core CRIC. Both (core and experiment) parts share a common implementation framework. However, experiment CRICs describe concepts which are not necessarily the same for various experiments and therefore, experiment CRICs represent experiment-specific plugins. Below we describe some concepts implemented in the experiment CRICs and correspondingly some important functionality which experiment CRICs provide in addition to the core one.

2.3.1 Experiment Site

GRID sites declared in GOCDB and OIM get an official name. Some of those sites are geographically distributed. From the experiment perspective, a processing site represents a set of services which normally provides a certain storage capacity and some processing resources, and that is handled by a dedicated team of site administrators. From the computing operations point of view, sites as they are declared in GOCDB or OIM, do not necessarily represent a single processing site for a given experiment. For example, the geographically distributed GRIF site which is declared as a single site in GOCDB, represents several sites from the experiment point of view.

Moreover, experiments can name sites following their internal naming convention; therefore, sites get different names in the experiment scope compared to their official names. In order to solve these issues, a concept of the 'Experiment Site' has been introduced. Experiment Site defines the subset of resources which are hosted by a particular GRID site declared in GOCDB/OIM and are used by a given experiment. Experiment site described in the experiment CRIC is mapped to the GRID site described in the core CRIC via set of services which logically belong to the Experiment site and are physically being hosted by the GRID site. It has various properties for example 'tier'. For various experiments a given site can play a role of a different tier. Experiment Site has a name following the internal naming

convention. The concept of the experiment site enables the introduction of resources which are not hosted by sites declared in GOCDB/OIM, for example commercial clouds or HPC.

2.3.2. *Resource Unit*

The CMS computing operations require a structure which represents an intermediate layer between the site and a particular computing or storage service (element). The new concept of the ‘Resource Unit’ has been introduced in order to describe this structure. It can include a set of storage services of different types, not necessarily hosted by a single experiment site. Following the operational needs, resource unit allows to define an association between particular storage and computing services.

2.3.3 *Transfer Matrix*

Understanding of network performance is important in order to organize data transfer and data processing in an efficient way. Therefore, experiments put a lot of effort in testing, debugging and monitoring of the network links. It is foreseen that experiment CRIC will provide a possibility to describe network topology of the experiment infrastructure. This functionality has been implemented in AGIS and will be offered by the CRIC experiment instances.

3. CRIC Roadmap

CRIC development tasks include tasks related to the core CRIC module and experiment specific plugins. The new system should ensure clean decoupling of the core part from the experiment-specific plugins. The system needs to enable modelling of various computing services. Most of this work has been already done in AGIS and can be re-used. However, some of those services and corresponding concepts are more complex than the others and therefore, the limitations in the initial AGIS implementation need to be addressed. Among such services is storage service which has to be modelled considering complexity of having multiple protocols, quota nodes and various permissions. Storage service object should be easily integrated in different experiment-specific data structures and should provide complete description of the storage which is required for operations, data transfers, data access, monitoring and accounting tasks.

Another important goal is to implement fine-grain user permissions policy via federated identity integration and single-sign-on authentication. Both core and experiment-specific components of the system have to provide highly customised UI allowing to perform all administrative and operational tasks of various user categories.

Considering various experiment-specific plugins, most of the work required to complete the CRIC project is required on the CMS CRIC. The CRIC development team works in close collaboration with the CMS experts in order to understand the CMS data structures. For LHCb and ALICE the task is much simpler since they are currently not planning to rely on CRIC for the experiment-specific configurations. Only basic concepts of the experiment sites need to be implemented for those two experiments. Concerning ATLAS, the effort to have an ATLAS CRIC relies on the migration from AGIS to CRIC. This will require substantial work but it is less effort than implementing CMS CRIC, as AGIS and CRIC are based on the same technologies and similar data structures.

References

- [1] Laurence Field and Markus W. Schulz. Grid deployment experiences: The path to a production quality LDAP based grid information system. In the Proceedings of 2004 Conference for Computing in High-Energy and Nuclear Physics, pages 723-726, Interlaken, Switzerland, October 2004.

- [2] Ian Bird. Computing for the large hadron collider. *Annual Review of Nuclear and Particle Science*, 61(1):99-118, November 2011.
- [3] Burke, Laurence Field, and Maarten Litmaath. GLUE schema version 1.3.
- [4] L. Field, B. Konya, and S. Andreozzi. GLUE specification v. 2.0. Technical Report GFD.147, Open Grid Forum, March 2009.
- [5] Gilles Mathieu, Dr Andrew Richards, Dr John Gordon, Cristina Del Cano Novales, Peter Colclough, and Matthew Viljoen. GOCDB, a topology repository for a worldwide grid infrastructure. *Journal of Physics: Conference Series*, 219(6):062021, April 2010.
- [6] Django Project: <https://www.djangoproject.com/>
- [7] Alessandro Di Girolamo, Alexander Senchenko, Dr Alexei Klimentov, Alexey Anisenkov. AGIS: The ATLAS Grid Information System. In the Proceedings of 2012 Conference for Computing in High-Energy and Nuclear Physics, New York, USA, May 2015.
- [8] OIM: <https://oim.opensciencegrid.org>

Reproduced with permission of copyright owner. Further reproduction prohibited without permission.