

# Table of Contents

<b>WLCG Information System Evolution.....</b>	<b>1</b>
Motivation.....	1
Mandate and Goals.....	1
Contact.....	1
Recent tasks and discussions.....	1
Publishing CE configuration in the JSON format.....	2
Latest docs.....	2
Historical docs.....	2
First implementations.....	2
Storage Resource Reporting Implementation.....	2
Automatic JSON Validation.....	2
Completed Task Tracking and Timeline.....	3
Completed.....	4
Documentation.....	7
Use Cases.....	7
Future Use Cases.....	7
Information Sources.....	7
IS clients.....	7
VOfeed.....	7
Types of Information.....	7
Static Information.....	7
Mutable Information.....	8
Dynamic Information.....	8
Classification of WLCG Information per type.....	8
Requirements for the new WLCG Information System.....	9
Validation.....	9
OSG.....	9
EGI.....	9
WLCG.....	10
MW Information Providers.....	10
Roadmap to the new WLCG IS.....	10
Definitions.....	10
CRIC development.....	11
Meetings and presentations.....	12
WLCG Operations Coordination reports.....	12
2016-12-01.....	12
2016-11-03.....	12
2016-09-29.....	13
2016-09-01.....	13
2016-07-07.....	13
2016-06-02.....	13
2016-04-28.....	14
2016-04-07.....	14
2016-03-17.....	14
2016-02-18.....	14
2016-01-21.....	15
2016-01-07.....	15
2015-12-17.....	15
2015-12-03.....	15
2015-11-19.....	16
2015-11-05.....	16
2015-10-01.....	17
2015-09-17.....	17
2015-09-03.....	18

# Table of Contents

## WLCG Information System Evolution

2015-07-30.....	18
Additional material.....	18

# WLCG Information System Evolution

## Motivation

In June 2015, OSG announced their plans to stop using the BDII to publish their computing resources (See [Slides](#) presented at the WLCG Operations Coordination Meeting in 18th of June). This announcement has triggered the review of the current WLCG Information System. It has been decided to create a task force to evaluate how WLCG is going to evolve to be able to cover the existing use cases and finally improve all the existing drawbacks and weaknesses of its current implementation.

## Mandate and Goals

In the scope of WLCG Operations Coordination, the WLCG Information System Evolution Task Force will pursue this objectives:

- Short term goals:
  - ◆ Fix existing issues in REBUS. REBUS is the Resource, Balance and Usage website for the whole of WLCG project, including topology information, resource pledges, and installed capacities. It is the authoritative source of information for WLCG and for this reason information published there should be correct and consistent so that users can trust it.
- Long term goals:
  - ◆ identify the existing use cases of the current WLCG Information System with experiments and other activities within WLCG like monitoring or accounting
  - ◆ define the architecture of the new WLCG Information System deciding how the different types of information need to be provided. Is there a need for a service registry for static information? could we consider messaging to retrieve dynamic information? what about mutable information?
  - ◆ identify the list of requirements for the new WLCG Information System
  - ◆ identify with OSG, EGI and NDGF which services providing information about their resources will be supported
  - ◆ identify the authoritative sources of information for the WLCG Information System. This could rely on the information services provided by OSG, EGI and NDGF or rely on manual methods (like T1s installed capacities in REBUS). In summary, it has to be decided how WLCG wants to collect the needed information to meet the existing use cases so that the information is guaranteed to be correct
  - ◆ plan the implementation of a new WLCG Information System that integrates the information from OSG, EGI and NDGF, providing the information needed by the defined use cases
  - ◆ plan the transition from the current to the new WLCG Information System

## Contact

All members of the task force can be contacted at [Infosys-discuss@cernSPAMNOTSPAMNOTNOSPAMPLEASE.ch](mailto:Infosys-discuss@cernSPAMNOTSPAMNOTNOSPAMPLEASE.ch)

[Infosys-discuss](#) egroup page and membership [↗](#)

## Recent tasks and discussions

## Publishing CE configuration in the JSON format

The proposal consists of publishing CE description and configuration in the agreed JSON format available through HTTP as an alternative or in addition to BDII. The URL for CE description can be attached to CE in GocDB. When CRIC is in place, it can contain the description translating JSON info into CRIC CE model. The idea and workflow are similar to Storage Resource Report proposal discussed in the accounting task force scope, though it also has a topology and IS impact. The CE proposal has been welcome by the members of the task force. The initial proposal from Alessandra Forti can be found here [↗](#).

### Latest docs

Google doc with the latest format and specification [↗](#)

### Historical docs

The initial version of the json format proposal

Google doc which keeps initial discussion and work on the specification [↗](#)

CRR schema version 1.2 (26/4/2019)

Minutes of the meeting when the format was discussed are attached to the agenda of the meeting [↗](#)

### First implementations

- VAC example [↗](#)
- Manchester example [↗](#)

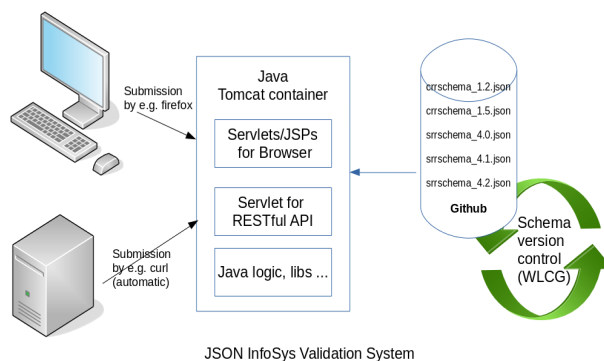
## Storage Resource Reporting Implementation

- Various sites have implemented SRR JSONs for different storage technologies
- Latest SRR specification

## Automatic JSON Validation

Prototype technology to perform automatic JSON validation has been developed and is under version control at: [https://github.com/sjones-hep-ph-liv-ac-uk/json\\_info\\_system](https://github.com/sjones-hep-ph-liv-ac-uk/json_info_system) [↗](#).

- JSON Validation Architecture (basic):



The system has these parts:

- JSONSchema schemas for CRR (compute) v1.5 and SRR (storage) v4.1 (and a draft 4.2 that is very lenient). These depend on version 7 of JSONSchema.
- Java to validate a JSON (whether well formed and in compliance with the relevant schema.)
- Java to parse a valid JSON to do further checks related to data integrity (unique names, valid relationships...)
- A website that allows a user to post a JSON file for validation, returning a status and description.
- A RESTful webservice that does the same validation in a way that can be scripted with (say) curl.
- Some equivalent work in Python that might be used on the command line (incomplete.)

The website is here: <http://hep.ph.liv.ac.uk/JisValidator/JVMain.jsp>. The current options are to test a CRR or an SRR JSON, or to view the schemas. The webservice can also be used; here are some examples.

```
curl -i -F jsonfile=@/root/dev/json_info_system/srr/v4.0/test/storage_service_v4.json https://h
```

```
curl -i -F jsonfile=@/root/dev/json_info_system/crr/v1.5/test/liv.json https://hep.ph.liv.ac.uk
```

You can specify the schema version and whether to check JSON integrity in both Browser and CLI/RESTful interfaces. The RESTful interfaces take the ver and integrity parameters, e.g.

```
curl -i -F jsonfile=@/root/tmp/storagesummary_lanc_edtwork.json https://hep.ph.liv.ac.uk/JisVa
```

You can also use the website to download and read schemas - they are fairly clear.

## Completed Task Tracking and Timeline

Date	Task Name	Deadline	Progress	Affected VOs	Affected Sites	Responsible	Comments
12.05.2016	AGIS to consume static attributes from GOCDB/OIM	-	On hold	ATLAS	A few sites	Maria Alandes, Scott Teige, Alessandro di Girolamo	Understand CRIC requirements
12.05.2016	Report about how CRIC may impact VOfeed/ETF plans	-	On hold	All	-	Julia Andreeva	As soon as CRIC future is more clear, this needs to be evaluated. First prototype will be ready soon
12.05.2016	Check status of LHCb VOfeed	-	Ongoing	LHCb	-	Stefan Roiser	
12.05.2016	Review VO tag validation after the first exercise and decide whether this needs to be done in the future	16.06.2016	On hold	All	All	Maria Alandes	Understand CRIC requirements
08.01.2016	Create a new type in GOCDB (Execution Environment) to publish Logical CPUs and Benchmark values	-	On hold	All	All	Maria Alandes	Understand CRIC requirements

08.01.2016	Create a new type in OIM (Execution Environment) to publish Logical CPUs and Benchmark values	-	On hold	All	All	Maria Alandes	Understand CRIC requirements
24.09.2015	Define a GLUE 2.0 Roadmap	-	On hold	All	All	Maria Alandes	See Roadmap twiki Understand CRIC requirements
12.11.2015	Give examples of wrongly published information	-	-	ATLAS	ATLAS sites	Alessandro di Girolamo	

## Completed

Task Name	Deadline	Progress	Affected VOs	Affected Sites	Responsible	Comments
12.05.2016 : EGI to report on possible security implications of removing BDII publication	16.06.2016	☑ DONE	All	All	Alessandro Paolini, Vincenzo Spinoso	Conclusions presented
12.05.2016: Review the need for REBUS installed capacities view	-	☑ DONE	ATLAS	All	Maria Alandes	Known issues found in REBUS capacities now documented. No changes in REBUS for the time being. Capacities are needed also by WLCG Management and will be included in CRIC. Further discussions on how to obtain them as part of CRIC development effort
12.05.2016: Checkpoint with pic after removing BDII publication for Storage	-	☑ DONE	ATLAS, CMS, LHCb	pic	Maria Alandes, Marc Caubet	All OK after several weeks
31.03.2016: Validate WLCG resources (associated to the WLCG tags) in GOCDB by comparing to the experiments VOfeed	-	☑ DONE	All	All	Aleksandr Berezhnoi	See more in VO Tags Validation
31.03.2016: Find some volunteer sites to start playing with more static information in GOCDB/OIM	-	☑ DONE	ATLAS	Glasgow	Maria Alandes, Gareth Roy	Static info can be easily added to GOCDB as demonstrated by G. Roy

31.03.2016: Study the feasibility of stopping BDII publication for storage resources dedicated to LHC VOS. This includes discussing with EGI about OPS tests		☑ DONE	All	All	Maria Alandes	There has been a test with PIC and their dCache server dedicated to LHC VOs. MORE details in Stop WLCG dependencies on BDII
31.03.2016: Understand the timeline to have a writeable API in GOCDB	-	☑ DONE	-	All	Maria Alandes	First prototype to be expected in ~3months. Regular update of progress in upcoming IS TF meetings
31.03.2016: Inform to TF members whether OSG has now a timeline to decommission BDII		☑ DONE	ATLAS	-	Maria Alandes	Target timeline is 31.03.2017. BDII may still be available in an unsupported manner beyond that date, but right now that will be the last day it will run as a production OSG operational service.
11.02.2016: Work on a CRIC prototype for ATLAS and CMS	2-3 months	☑ DONE	ATLAS, CMS	-	Alexey Anisenkov, Alessandro di Girolamo, Stephan Lammel, Giuseppe Bagliesi	See Evaluation of CRIC by CMS
11.02.2016: Prepare a table of primary information sources	-	☑ DONE	All	-	Maria Alandes	See Information Sources table
11.02.2016: Follow up whether there is any room for collaboration between LHCb and ATLAS for LHCb's plans to improve current collectors	-	☑ DONE	ATLAS, LHCb	-	Maria Alandes	LHCb doesn't see the need to collaborate with ATLAS
Study the proposal of publishing a subset of the current GLUE schema in JSON/HTTPS based on the attributes needed by WLCG	-	☑ DONE	LHCb, ATLAS	All	Andrew McNab	See Vcycle/Vac support for GLUE 2.0 publishing via JSON/HTTP <a href="#">☞</a>
Check validation mechanisms in OSG	-	☑ DONE	All	OSG	Maria Alandes	This is now documented in the Validation section
Understand the status of ClassAd-GLUE 2 translator with IT-PES	-	☑ DONE	All	HTCondor sites	Andrea Manzi	The translator will be distributed as an rpm in the WLCG repository

WLCGISEvolution < EGEE < TWiki

Investigate the use of resource BDII's to get dynamic information using GLUE 2.0	-	☑ DONE	ALICE	All	Maria Alandes, Maarten Litmaath	See minutes of TF meeting on 12.11.2105 <a href="#">↗</a>
Investigate the use of GOCDB/OIM as service registries based on use cases document	-	☑ DONE	All	-	Maria Alandes, David Meredith, Brian Bockelman	GOCDB and OIM developers have provided the necessary details and some VOs are already investigating and exploiting these features
Prepare a Future Use Case document to be presented at the GDB	November	☑ DONE	All	All	Maria Alandes	See Future Use Cases section
Prepare a Use Case document to be presented at the MB	September	☑ DONE	All	All	Maria Alandes	See Use Cases section
Review information providers to match agreed definitions	-	Cancelled	-	-	Maria Alandes	Execution Environment directly in GOCDB/OIM. No need for info providers
Review sites configurations to match agreed definitions	-	Cancelled	-	-	Maria Alandes	Include validation steps already in GOCDB/OIM
REBUS to consume GLUE 2 information	-	Cancelled	-	-	-	New IS will be based on AGIS
REBUS to validate information before it gets published	-	Cancelled	-	-	-	Installed Capacity information will be included only in the new IS, not needed in REBUS
REBUS to include T3 sites	-	Cancelled	-	-	-	REBUS will only include official MoU sites. This would fit in the new IS
REBUS to include pledges per sites	-	Cancelled	-	-	-	REBUS will only collect official pledges per federation. This would fit in the new IS
24.09.2015: Investigate the possibility of integrating glue-validator at resource BDII level	-	Cancelled	All	All	Maria Alandes	No effort will be put in BDII as the idea is to reduce its dependencies
08.01.2016: Agree on clear definitions for Installed Capacities	-	Cancelled	All	All	Maria Alandes	It was decided to leave other more relevant WGs and TFs to work on the definitions
08.01.2016: Prepare a Publishing Tutorial twiki based on the GridPP one <a href="#">↗</a>	-	Cancelled	All	All	Maria Alandes	As stated in the previous item, this should be done within other WGs and TFs



## Documentation

### Use Cases

The WLCG Information System use cases document was presented at the MB on 15.09.2015. It collects input from all LHC experiments and WLCG activities, describing their interactions with the WLCG Information System. Document available in PDF [↗](#)

### Future Use Cases

The WLCG Future Information System use cases is now ready. The document describes future use cases envisaged by experiments and other WLCG activities interacting with the IS. Future use cases include a review of existing use cases (are they still needed?) and new use cases that may be desired. Document available in PDF [↗](#).

### Information Sources

The list of information Sources from which experiments collect information about existing services is summarised in the following documents:

- Information Sources for services defined in GOCDB, OIM/MyOSG, BDII or REBUS: this is basically a summary of the Information System Use Cases ( PDF)

### IS clients

Name	GLUE schema version	Main developer	Status
lcg-info	GLUE 1	Andrea Sciaba	No longer maintained. Best effort in case of problems
lcg-infosites	GLUE 1	Maarten Litmaath	No longer maintained. Best effort in case of problems
ginfo	GLUE 2	IT-SDC	No further developments scheduled, still waiting for user's feedback
ldapsearch	GLUE 1 and GLUE 2	OpenLDAP	Maintained

### VOfeed

VOfeed contains experiment topology information. For more details on VOfeed, please check these slides [↗](#):

- VOfeed documentation
- Experiment VOfeeds:
  - ◆ ALICE [↗](#)
  - ◆ ATLAS [↗](#)
  - ◆ CMS [↗](#)
  - ◆ LHCb [↗](#)

## Types of Information

### Static Information

Static information is information that is constant throughout the lifetime of a service. A collection of this type of information is what we call a service registry. Service registries are used for service discovery. This task

force should decide what sort of service registry is needed to address the existing use cases.

A WLCG service registry could be implemented extending the current OIM/GOCDB implementations, or extending REBUS, where there is already integrated information from OSG, EGI and NDGF. In the past there was an attempt to implement a prototype for the WLCG Global Information Registry. The WLCG Global Information Registry is based on REBUS and brings together information published by different grid infrastructures like EGI and OSG. It shows both information on pledged resources and actual available resources. The WLCG Global Information Registry aims at aiding LHC experiments to configure their own experiment databases for job submissions and storage management.

A policy stating how services are added to and removed from the service registry and in which way this is done (manually or automatically) also needs to be defined.

### Mutable Information

Mutable information may change during the lifetime of the service, mainly due to configuration changes. In order to get mutable information, information could be periodically polled (like it is currently done with the BDII) or could use messaging to propagate updates in an automatic fashion.

Another issue is where to store mutable information. One possibility is to extend the service registry with this information. This task force should decide how mutable information is going to be published and stored to address the existing use cases.

### Dynamic Information

Dynamic information is highly-mutable information, mainly state changes. This is basically monitoring information. Messaging is the technology most suitable to get monitoring information since BDII has shown not to be ideal as it is fairly long to propagate changes. This task force should decide how dynamic information is going to be consumed to address the existing use cases.

## Classification of WLCG Information per type

VO/Project	Information	Type	Comments
ALICE	Status of the CEs	Dynamic	Resource or Site BDII queried once per minute
ALICE	Number of waiting jobs in the VOView	Dynamic	
ALICE	Number of running jobs in the VOView	Dynamic	
ATLAS	List of CEs	Static	Top BDII queried once every 2h
ATLAS	CE submission queues and associated parameters	Mutable	
ATLAS	List of SEs	Static	
ATLAS	SE protocol, storage areas and paths	Mutable	
ATLAS	Site latitude and longitude	Static	
ATLAS	Batch system type and version	Static	
ATLAS	HEPSPEC and Logical CPUs	Mutable	
CMS	List of CEs	Static	Bootstrapping of glideinWMS factory
CMS	Queue name	Static	
CMS	Number of cores, CPU and Wall clock time limits	Mutable	
LHCb	List of CEs	Static	Top BDII queried once every 12h
LHCb		Mutable	

	MaxCPU Time and CPU Scaling Reference		
SAM	Queue name	Mutable	SAM CE tests query the SAM BDII every time they run. 600-800 hits/hour
REBUS	Capacities	Dynamic	Top BDII queried once per hour
GFAL2	SE path	Mutable	Random queries, depending on GFAL configuration and whether full SURL is provided
C5 report	Capacities	Mutable	Once per week
Google Earth Dashboard	Site latitude and longitude	Static	Top BDII queried once per hour
Accounting	Benchmark information	Static	Input needed from APEL developers
Accounting	Message broker discovery	Static	Input needed from APEL developers

## Requirements for the new WLCG Information System

After the experience of running the current WLCG Information System, the new WLCG Information System should also address the following issues:

- **Validation:** even if glue-validator is in place and has helped to improve the overall quality of published information, sites can still publish wrong information into the BDII. It would be good to define validation mechanisms to ensure that the information published is correct and can be trusted.
- **Persistency:** BDII hierarchy relies on three levels from resource, to site and then to top BDII. If one of these levels fails, the information disappears. This has been partially fixed by the cache mechanism. The new WLCG Information System should ensure that information is available as long as it is valid. The validity of the information depends on the type of information. Update and deletion policies need to be defined.
- **Topology:** REBUS is the tool where sites belonging to WLCG are declared. On the other hand, BDII relies on OSG and OIM to get the list of site BDIIs to be published. WLCG has recently suffered from a suspended EGI site which has disappeared from the BDII. This has also impacted the capacity information published in REBUS that comes from the BDII. WLCG should define its own mechanisms to include or reject sites from its information system.
- **Flexibility:** with OSG planning to stop publishing in the BDII, the current WLCG Information System will be unable to provide information for all WLCG sites. The new WLCG Information System should be flexible enough to easily allow disparate information services running by different organisations, speaking different schemas or having different semantics to be integrated.

## Validation

### OSG

MyOSG publishes information stored in OIM (along with other datasources). The validation happens in OIM which is the only place where users can enter or update service details<sup>?</sup>. There aren't many validations performed apart from the required field validations. However, it is possible to extend the existing validations for each service type adding new ones.

### EGI

A Nagios test<sup>?</sup> is executed every 24h. It runs glue-validator against the site BDII of every EGI site. glue-validator is executed with the option to validate the GLUE 2.0 profile for EGI. In case of Errors, the COD opens a GGUS ticket to the site reporting about the errors and asking the site to fix them. See example GGUS ticket<sup>?</sup>. For more details, please check EGI Nagios tests<sup>?</sup>.

## WLCG

A series of validation campaigns have been carried out manually and has been documented in the GLUE Monitoring twiki.

Using the SSB<sup>?</sup>, these monitoring activities could be automated and WLCG has implemented various monitoring campaigns targeting specific GLUE attributes that are used by LHC experiments.

- Maximum CPU time<sup>?</sup> for LHCb
- BDII vs SRM Storage Capacity numbers<sup>?</sup> for LHCb
- WaitingJobs<sup>?</sup> for ALICE

EGI monitoring has contributed to a great extent to improve the quality of published information. However, EGI profile for GLUE 2.0 tests many attributes that are not needed by WLCG. Moreover, when a ticket is opened to a site, it compiles the results of several failed Nagios tests. For this reason, WLCG has considered useful to open GGUS tickets reporting about a particular attribute that is wrongly published.

## MW Information Providers

MW Information Providers do not perform any automatic validation of the information before it gets published. Although both MW developers and site admins are using glue-validator when implementing changes in their information providers or deploying a new service at their site. There are on going discussions to see how to improve the situation by integrating a glue-validator check at the start time of a resource BDII. However, it has to be noted that this could be used to validate static information since dynamic information changes on the fly while the resource BDII is running. For dynamic information, the validation mechanisms implemented by EGI and WLCG are more useful.

## Roadmap to the new WLCG IS

- Information Sources
- Core CRIC functionality

Note that this twiki is now obsolete. It presented a preliminary roadmap for a new WLCG IS before CRIC was designed. Please, use the above links for up to date information.

## Definitions

The definition of the following words is being discussed and agreed in the TF:

Pledge	Installed Capacity	Planned Capacity	Input from
Pledges tell me what funding agencies promised and are for comparison with usage, request making etc. and are relevant on a ~yearly base having pledges by federations, not sites, is bad	installed capacity (or whatever) is what I can use to process data and is relevant for my ops. planning on a weekly base change names as you like, but in practice we need the two numbers	-	Stefano
Used for political monitoring	"installed capacity" is a very outdated concept, valid only in cases of dedicated hardware, strictly speaking. Realistically, "installed capacity" should be replaced by "available capacity" and should be dynamic	-	Oxana
		-	Andrea

<p>Pledges are by definition a promise (for the future or the present), but it cannot be assumed that the installed capacity is 100% of the pledges. They should be used only for high level planning, not in daily operations. It's not clear when pledges become valid (they are given with a yearly granularity in REBUS). There are no pledges for T2s</p>	<p>It is the amount of resources available to a VO under normal operating conditions. This would mean that if the farm is partially off, it's not a normal operating condition and capacity won't change because of that (there will be a downtime in GOCDB/OIM for that). If the resources dynamically change (due to elastically changing cloud resources for example) the installed capacity should change accordingly (it's still normal operating conditions). The site publishes the installed capacity to REBUS via a REST API and the numbers are calculated by whatever means the site chooses (BDII not involved at all unless the site chooses to use its information for that)</p>		
	<p>Capacity refers to HW. While you may say that the virtual machines/containers are created and destroyed "automatically" the hardware still is installed with that capability and virtual machines become not different from job slots with tweaks for fair shares or dedicated nodes. The method is just more dynamic and may happen faster but what really counts is the hardware that can be used whether via a batch system or a cloud is irrelevant.</p>		Alessandra
		<p>Planned capacity is a site s best estimate of what capacity will be available at a given point in the future, given its current plans. (What will be available in 1 second s time is already the future .)</p>	Andrew McNab

It may be interesting to check how these words are defined in Usage of Glue Schema v1.3 for WLCG  
 Installed Capacity information

It has been discussed that it would be good to differentiate between:

- Installed Capacity: Physical HW which is in place
- Available Capacity: HW which is actually usable (i.e. not offline for maintenance) for a period of time longer than i.e. 3 days.

## CRIC development

- Core CRIC
- CMS CRIC
- Authentication-authorization
- Installation

## Meetings and presentations

Task Force meetings take place on Thursdays at 15h30. Meetings are called on a regular basis as needed.

- 16th July 2015 - Proposal to create the TF at the Ops Coord meeting [↗](#)
- 23rd July 2015 [↗](#) ( Doodle [↗](#) )
- 27th August 2015 [↗](#) ( Doodle [↗](#) )
- GDB presentation [↗](#) 9th September 2015
- MB presentation [↗](#) 15th September 2015
- 24th September 2015 [↗](#)
- 12th November 2015 [↗](#)
- 26th November 2015 [↗](#)
- GDB presentation [↗](#) 9th December 2015
- 8th January 2016 [↗](#)
- Information System Discussion at the WLCG workshop [↗](#) 1st february 2016
- 18th February 2016 [↗](#)
- 31st March 2016 [↗](#)
- 12th May 2016 [↗](#)
- 16th June 2016 [↗](#)
- GDB presentation [↗](#) 13th July 2016
- MB presentation [↗](#) 19th July 2016
- 22nd September 2016 [↗](#)
- CHEP 2016 Presentation [↗](#)

## WLCG Operations Coordination reports

### 2016-12-01

- VOfeed management and documentation discussed at the last IS TF meeting.
  - ◆ Working on a twiki page where VOfeed structure will be documented and discussed at the next meeting.
  - ◆ It was decided that VOfeed changes and general strategy will be discussed at the IS TF from now on.
- Ongoing discussions on which syntax based on GLUE 2 should be used to enter more info in GOCDB.
- New person selected to work on the CRIC project. Contract procedure on going. Likely to start in January.
- Next IS TF meeting [↗](#) will take place on 8th December:
  - ◆ GOCDB writeable API
  - ◆ VOfeed Documentation
  - ◆ Proposal to introduce extended attributes in GOCDB
- Reminder on REBUS known issues regarding capacity numbers. Please check them before opening any GGUS ticket or spending time in understanding capacity numbers.

### 2016-11-03

- GOCDB developers and EGI contacted to understand how to add extra information associated to service endpoints with extension properties in GOCDB. It is feasible to consider this feature in GOCDB and it is aligned with EGI plans to add more information in GOCDB.
- Next GOCDB release to be released in the next weeks will contain a writeable API. This is an interesting feature to allow sites to publish more information in GOCDB in an automatic way.
- Feedback from GLUE-WG experts to define storage and computing attributes in GLUE 2 needed for CRIC and storage accounting. This list will be used to document the information needed in the

different information sources queried by CRIC. Discussions with OSG to make sure they can also provide this list are ongoing.

- Recruitment process to hire a new CRIC developer is ongoing this week and a candidate is expected to be selected very soon.
- Next IS TF meeting [↗](#) will take place on 10th November. VOfeed structure and integration with CRIC will be discussed.

## 2016-09-29

- An IS TF meeting [↗](#) took place on 22nd of September. Information sources and main functionality of central CRIC were discussed. Alignment with EGI plans on moving more information to GOCDB was agreed. There is on going progress on the defined actions.
- Next IS TF meeting [↗](#) will take place on 10th November. VOfeed structure and integration with CRIC will be discussed.

## 2016-09-01

- At the last MB [↗](#) a proposal to adopt CRIC as the new Information System was approved. A new project associate will join the development team in the next weeks.
- The next IS TF meeting [↗](#) is scheduled on 22nd of September. Information sources and main functionality of central CRIC will be discussed.

## 2016-07-07

- An IS TF meeting [↗](#) took place on 16.06.2016:
  - ◆ EGI presented their main motivations to keep on relying on the BDII. There were discussions on which areas would be affected if WLCG stops relying on the BDII. Some of them are: MW upgrades and EGI 2nd line support. If WLCG finally decides to stop BDII, the impact of this needs to be better understood.
  - ◆ There was a proposal to drop capacity views from REBUS. There was agreement within the TF to do this. It was decided to present this at the MB for official green light. However, after discussions with I. Bird, it was decided not to do anything for the time being until the new IS is in place. It was agreed that since REBUS capacity known issues have been documented, if sites open tickets complaining about wrong values, they won't be fixed for the time being and sites will be pointed to the known issues page.
- CRIC evaluation for CMS was over and CMS decided to engage further in the project. Developers and CMS people are now discussing the next steps.
- A GDB presentation [↗](#) to report about the status of the TF and CRIC is scheduled for the 13.07.2016.

## 2016-06-02

- Evaluation with REBUS developers and AGIS developers of the capacity view to consider whether it can be dropped from REBUS
- Ongoing discussions with EGI and relevant experts (MW Officer, Security) on evaluating the impact of stopping BDII
- Testing deployment of CRIC prototype as well as playing with first CMS data on it. More details in CRIC Evaluation
- WLCG scope tags in GOCDB in the process to be validated (33 out of 123 tickets still open). Thanks to the sites for their effort. More details in VO Tags validation twiki. There are ongoing discussions to decide whether it makes sense to repeat this validation on a regular basis and whether it makes sense to compare with VOfeeds and maintain a 1 to 1 relationship between tags and resources in VOfeed.
- Next TF meeting [↗](#) scheduled on 16.06.2016

**2016-04-28**

- Working on reducing BDII dependencies:
  - ◆ Dedicated LHC VOs Storage: a recipe is being prepared for sites based on PIC experience to be able to stop publishing BDII for dedicated LHC VOs storage services.
  - ◆ Computing: work ongoing to define static CE attributes in GOCDB/OIM. ATLAS contacted to test this in a few ATLAS sites and AGIS.
- WLCG scope tags in GOCDB are being validated by Aleksandr Berezhnoi. More details in VO Tags validation twiki.
- CRIC prototype for CMS progressing well. More details in CRIC Evaluation
- Ongoing work will be summarised in the next TF meeting [☞](#) scheduled on 12.05.2016

**2016-04-07**

- Targeted timeline for OSG BDII decommissioning is March 31st, 2017. It may be available in a unsupported manner beyond that date, but right now that will be the last day it will be run as a production OSG operational service.
- An IS TF meeting [☞](#) took place on 31st March:
  - ◆ Medium term plans were discussed
  - ◆ A feasibility study to stop dependencies on the BDII is being carried out.
    - ◇ A first test with pic and storage was carried out in the past weeks and this shows SAM dependencies on SAM OPS tests that need to be discussed with EGI.
    - ◇ Discussions with Marian to understand current BDII dependencies on VOfeed generation.
  - ◆ Discussions with GOCDB/OIM developers to understand how to add more static information
  - ◆ WLCG scope tags in GOCDB need to be validated

**2016-03-17**

- List of primary information sources is now summarised in this document.
- CMS and ATLAS agreed to evaluate together a common information system (CRIC). First meetings are taking place. It was agreed to work on a prototype in the next few months.
- The strategy to stop depending on the BDII and using GOCDB/OIM as unique information sources will be evaluated as part of the CRIC work.
- Short and medium term plans within the TF will be discussed at the next meeting taking place on 31st of March.

**2016-02-18**

- Information System discussed at the WLCG workshop:
  - ◆ General agreement that it would be desirable to become independent from the BDII, although in practice this needs to be understood.
  - ◆ No clear outcome about the new IS. There is a general feeling that a new IS is useful, but this needs in any case to be supported by the experiments. As a follow up at the MB on Tuesday, it was agreed to re-visit the experiment needs for this.
- An IS TF meeting [☞](#) took place on 11th February:
  - ◆ In order to define a strategy for the BDII, EGI was invited to present their plans to support the BDII and it was made clear that EGI plans to support the BDII as many VOs rely on it.
  - ◆ It was agreed to assess the feasibility of moving static information to GOCDB/OIM, since experiments like ATLAS are interested in going in this direction.
  - ◆ It was agreed to work on a table where all primary information sources for each experiment will be described and identified. This should be a compact version of the Use Cases document and an easy way to understand where information is defined and where information is



consumed, highlighting possible inconsistencies and also helping to steering the discussion on how to evolve the IS.

- ◆ It was agreed to investigate whether there is room for collaboration between LHCb and ATLAS after LHCb's implementation of multiple information collector plugins for the DIRAC CS.
- ◆ It was decided to stop discussing about definitions since this work fits better within the benchmarking working group and the MJF TF.

## 2016-01-21

- Preparation and discussion of the slides to be presented in the WLG workshop.
- New Execution Environment service in GOCDB/OIM to give logical CPUs and Benchmark information of the resources in a site:
  - ◆ Discussion with GOCDB developer to understand whether a new Execution Environment service could be added to GOCDB. The answer is yes but there is no writeable REST API for the time being. Feedback being collected from sys admins to understand advantages and disadvantages of having this new service defined in GOCDB.
  - ◆ OSG is partially providing the needed information (Benchmark) already. They are planning to add HS06 normalisation constant to be able to derive the number of Logical CPUs from there (Logical cores = (total hs06 / hs06 normalization))
- After the WLCG workshop we hope to have more clear directions on next steps inside the TF, especially for the new IS, that for the time being is on hold.

## 2016-01-07

- IS TF meeting scheduled tomorrow Friday 8th January. ( [Agenda](#) )
  - ◆ Definitions: summary of the proposed definitions and feedback from sys admins.
  - ◆ Status of new IS: news on the feedback given so far by experiments.
  - ◆ Preparation for the WLCG workshop discussion about the IS.

## 2015-12-17

- A proposal for a new WLCG IS based on AGIS was presented at the last GDB .
  - ◆ Ongoing discussions with experiments to understand their interest in this new IS.
  - ◆ The proposal will be presented at the MB next year to see whether it gets approved.
- In the meantime, the following activities are ongoing within the TF:
  - ◆ Ongoing discussion to agree on a better definition of the GLUE 2 attributes defining HS06 (GLUE2BenchmarkValue) and Logical CPUs (GLUE2ExecutionEnvironmentLogicalCPUs): feedback from sys admins is being collected for two possible definitions.
  - ◆ Presented at the last UMD meeting a proposal  to validate information at its source so that we can avoid publishing information that is known to be wrong. A technical solution will have to be worked out together with MW developers.
- Preparing the IS session at the WLCG workshop in February together with Alessandra Forti who will be the chair and who is gathering feedback on what to discuss.
- Next IS TF meeting scheduled on Friday 8th January. ( [Preliminary agenda](#) )

## 2015-12-03

- The Future Use Cases Document is now ready in the WLCG Document Repository ( [PDF](#) ). There is a general agreement that a central information system owned by WLCG is an interesting idea. For some VOs the requirement is stronger than for others, but all VOs agree that they would rely on a central information system that provides good quality information. Activities like WLCG Monitoring and Operations will definitely rely on such tool. The WLCG Information System should:

- ◆ Cache information from heterogeneous resources by regularly collecting information from primary data sources for WLCG service discovery (Now GOCDB, OIM and BDII, but the list of primary resources can evolve in the future).
- ◆ Provide a consistent interface for all interested WLCG clients offering an intermediate layer between the sources of information maintained by EGI and OSG.
- ◆ Include grid and non grid resources, like HPC and Clouds and be flexible enough to be able to include new types of resources.
- ◆ Validate information before it gets published, applying corrective actions if necessary.
- ◆ Logging information, namely when, how, by whom information was provided
- Starting to prepare a Roadmap to GLUE 2.0 so that VOs and WLCG clients start consuming GLUE 2.0 information and we can plan at some point the decommission of GLUE 1.3.
  - ◆ EGI presented their plans to move to GLUE 2.0. Main showstopper is GLUE 2 WMS that was never tested in production. EGI is now trying to understand its actual use.
  - ◆ Waiting for OSG input about their plans to provide information to WLCG once they stop publishing in the BDII and whether we could expect information published in GLUE 2 after the implementation of the ClassAds to GLUE 2 translator.
- Ongoing discussion to agree on a better definition of the GLUE 2 attributes defining HS06 (GLUE2BenchmarkValue) and Logical CPUs (GLUE2ExecutionEnvironmentLogicalCPUs), so that sites understand in a clear way what it is expected from them to be published in these attributes.

## 2015-11-19

- The first draft of the Future Use Cases document is now available for comments. Deadline to provide input is on 24.11. The document will be presented at the December GDB.
- There was a TF meeting on 12.11 ( Minutes [↗](#)). All the experiments presented their plans to move to GLUE 2.0 and proposals to simplify the interactions with the IS. Several action items were defined after the meeting:
  - ◆ Define a roadmap to stop publishing GLUE 1.3 in coordination with EGI and OSG.
  - ◆ Information validation:
    - ◇ Document existing validation mechanisms (this is now documented in the TF wiki)
    - ◇ Actively validate information that is important for WLCG. Feedback from experiments is needed (especially ATLAS). In particular, validation of the Waiting Jobs GLUE attribute for ALICE has been implemented ( SSB [↗](#)).
    - ◇ It was agreed that after the feedback collected so far, it doesn't make sense to define a GLUE 2.0 profile for WLCG.
    - ◇ There are ongoing discussions with MW officer to integrate glue-validator within the different services running a resource BDII and improve information quality before it gets published. This will be proposed at the URT meeting on 14th December.
  - ◆ Study the proposal of publishing a subset of the current GLUE schema that is useful for WLCG in JSON/HTTPS. Andrew McNab presented his work on publishing Vac/Vcycle resources using this approach.
- Next meeting is on 26.11 ( Agenda [↗](#))

## 2015-11-05

- Input for Future Use Case document is being finished within the experiments, some drafts already available waiting for the final green light. A complete first draft will be distributed within the TF in the next days.
- Ongoing discussions with experiments to understand what information needs to be validated for GLUE 2.
  - ◆ Specific actions are being implemented for ALICE.
  - ◆ Waiting for LHCb migration to GLUE 2 to have more details. So far, they are happy with the existing validation.
  - ◆ No specific requirement from CMS for the time being.

- ◆ To be understood for ATLAS.
- GOCDB testing instance<sup>☞</sup> is now able to filter WLCG services and also services per LHC VOs using the scope option. An option to get T1 and T2 downtimes is under development.
- Ongoing discussions with OIM developers to understand the feasibility of adding more information and implementing similar features as in GOCDB.
- IT-PES has developed the OSG ClassAds to GLUE 2 translator for HTCondor, and together with the MW Officer we are planning the distribution of the rpm through the WLCG repository.
- A TF meeting<sup>☞</sup> is scheduled next week where each experiment will present their future interactions with the IS and their plans to migrate to GLUE 2. It will also include a presentation about the GLUE 2 validation status and AGIS.

## 2015-10-01

- There was a TF meeting<sup>☞</sup> where the following presentations were made:
  - ◆ Follow up on MB&GDB presentations:
    - ◇ it was agreed to investigate the possibility of using OIM/GOCDB as service registries, extending the information they currently provide to meet use cases for static/mutable information; and to query the resource BDII/OSG collectors for dynamic information.
    - ◇ it was also agreed to consider the implementation of a WLCG profile to target the validation of information on WLCG use cases. Discussions on going with EGI to consider the integration of glue-validator at resource BDII level.
  - ◆ OSG presented their plans to move to ClassAds and OSG collectors to provide information about their resources, for the time being for HTCondorCE. A translator from ClassAds to GLUE 2 is developed by OSG and CERN IT-PES. MW Officer in contact with developer at CERN to understand how this translator could be distributed to all sites.
  - ◆ EGI presented their plans where the current information system is going to be used with the idea of moving to GLUE 2 and deprecating GLUE 1, as long as WLCG doesn't depend on it. It was agreed to plan for a transition in WLCG so that GLUE 2 information is consumed and we stop relying on GLUE 1.
  - ◆ NDGF presented the way in which they currently publish information, supporting both nordugrid and GLUE 2 schemas. They would prefer if we could move to GLUE 2 to make things simpler.
  - ◆ GOCDB developer made a presentation of technical details and features available in GOCDB that would allow us to move in the service registry direction.
- Ongoing discussions in the mailing list to resurrect ginfo as the GLUE 2 client tool to query the information system.

## 2015-09-17

- WLCG Information System Use Cases document presented at the MB
- MB gave feedback to work on several areas that need further discussion and agreement within the TF:
  - ◆ Future Use Cases: use cases document describes the current interactions with the IS. The TF should now investigate what it is actually needed so that we can better understand how the IS could evolve.
  - ◆ Static vs Dynamic: MB would like to see summarised the types of information actually needed by the experiments. Probably a more elaborated version of what it is already summarised in this twiki under Types of Information and focus only in the future use cases.
  - ◆ "Indicative pledges" per site in REBUS: The TF requested the MB to include "indicative pledges" per site in REBUS. MB would like to understand why this information is needed and have a concrete proposal on how it will be collected.
  - ◆ Installed capacity: a better definition, and maybe also name, is needed for what it is called today "installed capacity". MB would also like to understand why this information is needed and also how it will be collected.

- ◆ T3s and opportunistic resources: it would be good to understand how information is going to be collected from T3s and opportunistic resources.
- OSG, NDGF and EGI will present their plans to provide information about their resources in the future at the next TF meeting. GOCDB will also present the latest features.

## 2015-09-03

- REBUS known issues have been either fixed or are in the to do list of REBUS maintainers.
- Many action items are put on hold until Information System Use Cases presented at the MB
- Draft document describing use cases should be ready on Monday 7th September. It will be presented at the MB on 15th September
- Update on Information System Status also scheduled at next GDB on 9th September

## 2015-07-30

- The first TF meeting took place last week ( [agenda](#), [minutes](#))
  - ◆ It was agreed to implement in REBUS a set of easy fixes. For more details, please check REBUS known issues
  - ◆ A set of action items were defined, for more details, please check Task tracking and timeline. A summary below:
    - ◇ Requirements to remove information (Physical CPU) or change how information is collected (HS06) in REBUS will be followed up
    - ◇ Agree on a better definition of Installed Capacities, or even decide to change this name and better use "Available capacities" or something similar
    - ◇ Discuss at the MB the possibility of adding T3s and also publish pledges per sites in REBUS
- A draft document describing use cases from experiments and project activities relying on the information system has been circulated among TF members for their contribution. This will be presented in the future MB (date to be confirmed) although we are aiming to have the document ready by end August

## Additional material

- WLCG Information System main web page [↗](#)
- Past presentations about the status of WLCG Information System
- Past meetings [↗](#) held between December 2012 and May 2014 discussing issues affecting the WLCG Information System
- All About REBUS
- GOCDB [↗](#)
- OIM [↗](#)
- Summary of requirements for the Information System collected from 2012 TEG reports

-- MariaALANDESPRADILLO - 2015-06-29

- CE-json-proposal.txt: Draft of the CE json format

---

This topic: EGEE > WLCGISEvolution

Topic revision: r90 - 2020-05-20 - JuliaAndreeva



Copyright &© by the contributing authors. All material on this collaboration platform is the property of the contributing authors.

Ideas, requests, problems regarding TWiki? Ask a support question or Send feedback