CRIC Information Sources

Information System Task Force
September 2016

CRIC Information Sources

- The idea is basically to support what the sites are using
  - CRIC is flexible enough to query different information sources as long as there is a client/API to do this
  - We want to support sites saving them to duplicate effort and offering as much automation as possible
- At the same time, CRIC doesn't have a hard dependency on any information source
  - A site could directly enter its site information manually on CRIC
    - So sites who i.e. don't want to support BDII, they don't have to support it for WLCG once CRIC is in place

Supported CRIC Information Sources

- GOCDB/BDII for EGI resources

- OIM for OSG resources

- REBUS for federations and pledges

- CRIC direct input

  - HPC/Opportunistic

  - Any site who doesn't want/need to rely on GOCDB/BDII/OIM

What happens with GLUE?

- BDII currently publishes GLUE 1 and GLUE 2

- OIM doesn't publish GLUE at all

- CRIC can query GLUE 1, GLUE 2 or no GLUE

  - As long as the Information Source API is well structured and documented

- For those information sources that support GLUE, we propose to query GLUE 2

  - To be aligned with EGI plans

Proposal to gather information from the supported information sources

Service attributes

- Service attributes are described in:
  https://twiki.cern.ch/twiki/bin/view/EGEE/CoreCRIC
  - Note that these attributes are under discussion
  - In any case, adding new attributes to CRIC is not a problem

- In this presentation we will focus on
  - Computing and Storage
  - Static attributes for service discovery
    - Let's discuss capacities in a future meeting

Computing	CE endpoint Queue name CE type CE LRMS type Max CPU time Max Wallclock time
Storage	SE Endpoint SE implementation Protocol Quota node

GOCDB Information Source

- GOCDB tags will be required
  - It's the way to know that a site supports a certain VO for a particular service
  - Sites need to make sure they are correct and properly maintained
- CRIC needs extra information that is currently missing
  - Custom properties (key/value pair) should be used
  - See proposal presented in May 2016
    https://indico.cern.ch/event/517084/contributions/2151002/attachments/1270763/1885548/StaticInfoGOCDB.pdf
- GOCDB will offer a writeable API in version 5.8 (Autumn 2016)

GOCDB Service Type	Custom Property (key)
ARC-CE CREAM-CE CE	GLUE2EndpointURL GLUE2ComputingShareMappingQueue GLUE2EndpointImplementationName GLUE2ManagerProductName GLUE2ComputingShareMaxCPUTime GLUE2ComputingShareMaxWallTime
SRM XRootD globus-GRIDFTP globus-GSISSHD	GLUE2EndpointURL GLUE2EndpointImplementationName GLUE2EndpointInterfaceName

BDII Information Source

- The following GLUE 2 attributes are needed

| GLUE 2 Attribute	Custom Property
GLUE2ComputingEndpoint	GLUE2EndpointURL GLUE2EndpointImplementationName
GLUE2ComputingShare	GLUE2ComputingShareMappingQueue GLUE2ComputingShareMaxCPUTime GLUE2ComputingShareMaxWallTime
GLUE2ComputingManager	GLUE2ManagerProductName
GLUE2StorageEndpoint	GLUE2EndpointURL GLUE2EndpointImplementationName GLUE2EndpointInterfaceName
GLUE2StorageShare	GLUE2StorageSharePath
GLUE2MappingPolicy	GLUE2PolicyRule

OIM Information Source

- Already discussed with OIM developers

  - At least for computing, to be followed up on storage

- OIM is very flexible to provide what CRIC needs in different ways

  - For example (just a test, no final output):

  - http://reports.grid.iu.edu/reports/atlas_interop_information

  - http://reports.grid.iu.edu/reports/cms_interop_information

CRIC direct input

- Site admins and CRIC super user will be able to do modifications directly in CRIC
  - Site admins only for their site and services related information
    - This feature doesn't exist in CRIC yet, TBD
  - Super user for everything
    - Super users will be members of WLCG Operations
- We should define a clear policy on how a super user modifies information
  - Notifies the corresponding site
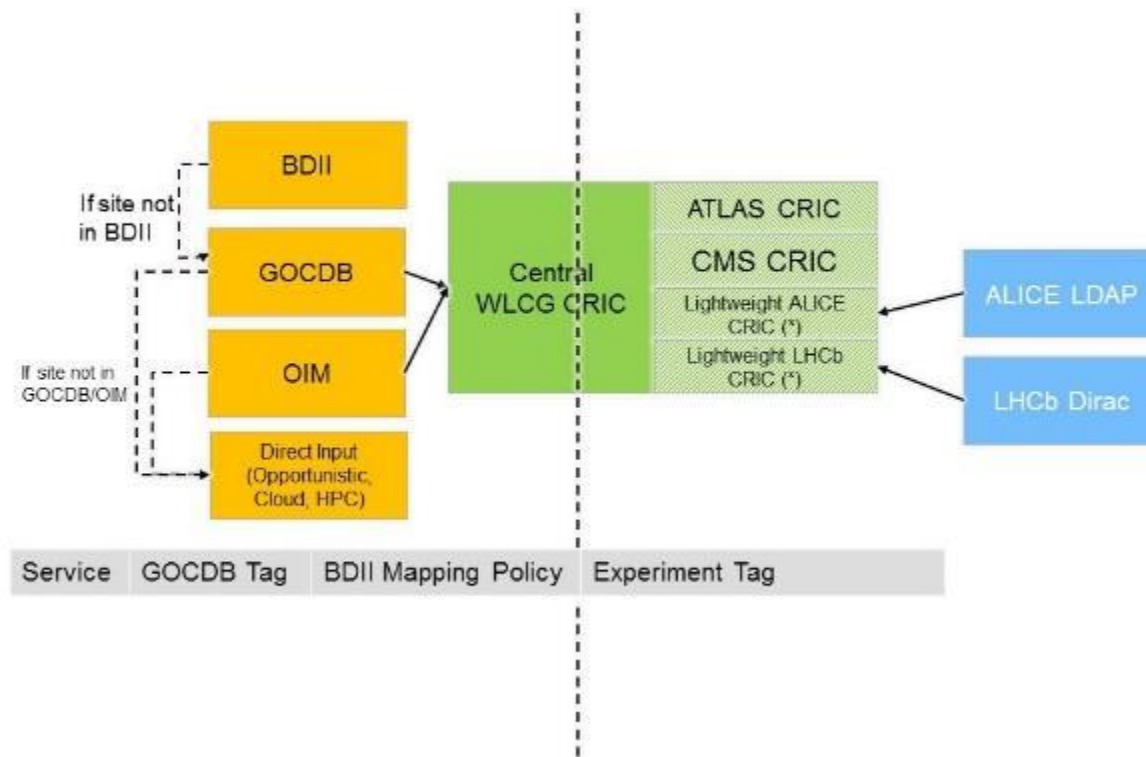 - Notifies WLCG Operations
  - Anything else?

Information sources workflow discussion

Bootstrap
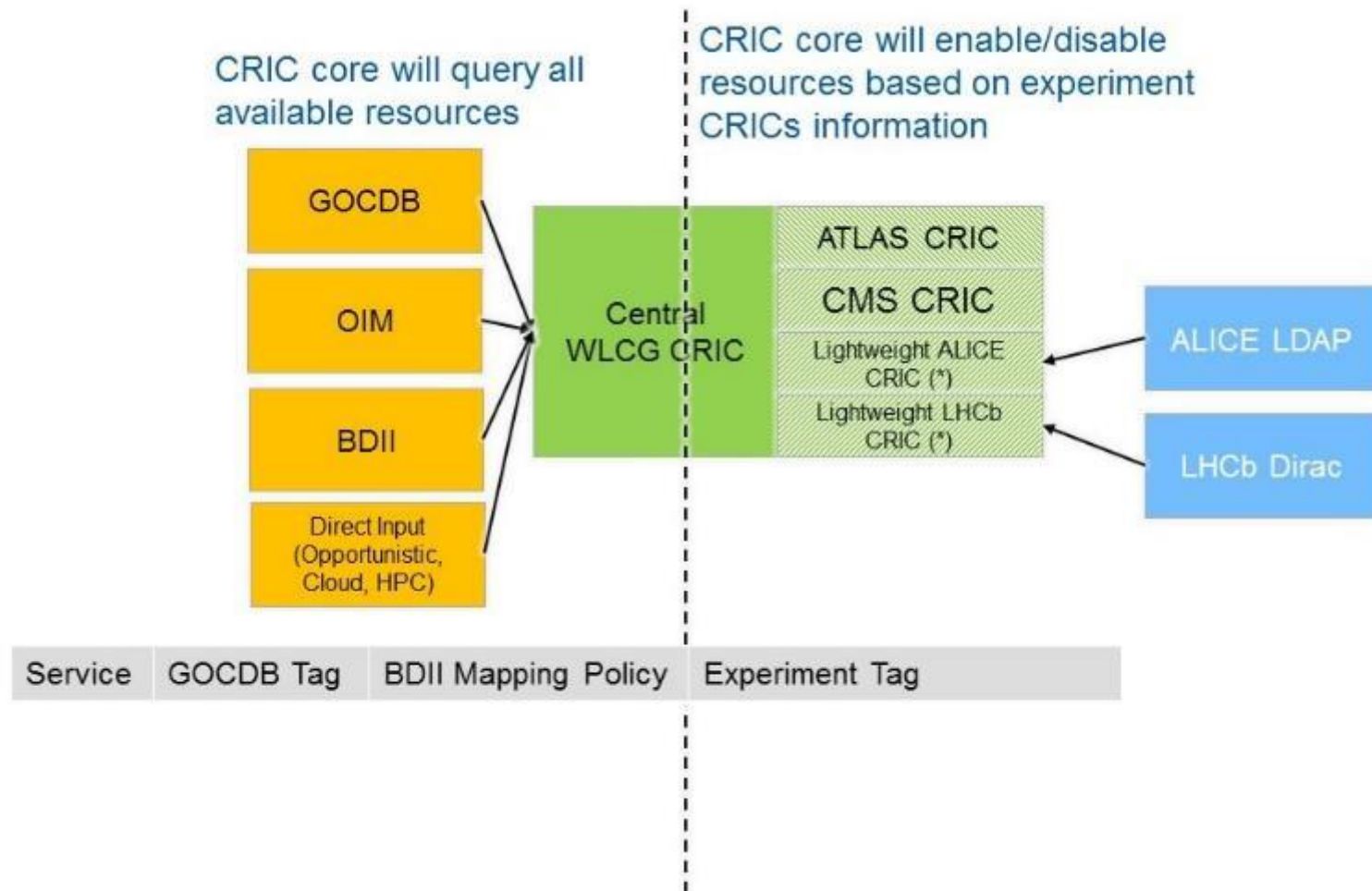
Regular updates

- How often we will query information sources?
- How often we will synchronise with experiment CRICs?
- The workflow below relies on sites publishing only in one information source

Conclusion and next steps

- The message today is that CRIC is very flexible and can accommodate different information sources

- We need to work on the details now

  - Volunteer sites willing to rely on the BDII

  - Volunteer sites willing to get rid of the BDII and rely on GOCDB

  - Volunteer sites willing to rely only on direct input

  - Further work with OSG to integrate computing and storage information