

# ARGO proposal to use GOCDDB as the only source of topology information

## Introduction

The ARGO monitoring engine and the ARGO compute engine are part of ARGO framework. The first one is responsible for monitoring all EGI production services, while the second one is responsible for computing the availability and reliability of the monitored services. Services are monitored by Nagios probes provided by service product teams. Topology information is currently retrieved by using two sources:

- GOCDDB - NGIs, sites, services, contacts and operators
- BDII - service URLs, service attributes.

Drawbacks of the current implementations are:

- ARGO relies on two sources of topology which makes the whole infrastructure more error prone and requires more effort in case of monitoring issues.
- BDII is information system which contains information what is currently available. In case of site or service downtime relevant topology information will be missing and ARGO monitoring and compute engine will result with faulty configuration.
- BDII contains information only for certified sites and thus cannot be used to retrieve information for un-certified sites. Therefore ARGO cannot be used to monitor un-certified sites as part of certification or un-suspension procedures.

## Proposal

We propose to extend GOCDDB information with service URLs to be used alongside (and progressively replace) BDII as a source of topology information for ARGO monitoring engine. This approach is already being used for FedCloud services and is far more reliable than mixed GOCDDB and BDII approach.

Main advantages of this approach are:

- Single source of topology makes ARGO monitoring engine simpler and more reliable, in the same time making issue tracking easier.
- ARGO monitoring engine could monitor all resources in GOCDDB irrelevant of their certification status.
- NGI and site managers could better control what services are being monitored.
- Easier development of new probes for new and existing service types as developers would know which attributes are provided as part of topology or if necessary request new ones.

Procedure:

- Services would be extended with multiple service endpoints, each service endpoint containing a URL required for probe execution (e.g. CREAM URL, SRM URL, etc). There are no limitations to the number of endpoints that can be added to a service in GOCDB. Therefore, endpoints not marked for monitoring will not be used by ARGO. Additional endpoints can be freely entered as needed by other VO-specific tools (e.g. WLCG monitoring, job broker systems, etc).
- Service endpoints to be monitored by ARGO must have:
  - Attribute Monitored set to True (this will require a new boolean 'Monitored' attribute be defined on the ServiceEndpoint)
  - Service URL defined:
    - CREAM-CE: URL should be set to GlueCEUniqueID from BDII (e.g. cream.egi.eu:8443/cream-pbs-ops)
    - SRM: URL should be set to GlueServiceEndpoint from BDII (e.g. http://se.egi.eu:8444/srm/managerv2)
    - Other services: URL should be set to GlueServiceEndpoint from BDII.
- In case when probes require additional attributes that are not part of the URL of the endpoint (e.g. fedcloud image name, SRM paths, etc) extension properties would be used. Currently none of the probes require extension properties.
- In the interim period ARGO would rely on both GOCDB and BDII, with GOCDB information having higher priority than BDII one. This approach would be used until all sites provide relevant information in GOCDB.
  - From the very beginning adding service endpoints to the GOCDB will be mandatory for sites that need monitoring and are not certified (e.g. sites under certification or suspended).
- During the interim period we will collect feedback from NGIs on how is the GOCDB information schema working. If the feedback is positive then BDII as a topology source for ARGO will be dropped after an agreed period.