

# Table of Contents

<b>Topology information integration - Proposal 1.....</b>	<b>1</b>
Introduction.....	1
Information representation and annotation.....	1
Core vocabulary.....	2
Information transport.....	2
Query/response paradigm.....	2
Publish/subscribe paradigm.....	2
Local information caching.....	2
Information integration and equivalence.....	2

# Topology information integration - Proposal 1

## Introduction

This proposal is based on:

- observation of significant number and distributed character of topology information sources,
- current experience with designing and maintaining SAM/GridView database model containing the grid topology information,
- brief research in the area of existing technologies for integrating distributed and multi-domain information systems.

The suggested solution is to use Semantic Web [\[7\]](#) approach or similar technologies to build integration and data exchange platform for all the grid monitoring and operation management tools that need topology information. This is in contrast to existing approach used in SAM/GridView system, which is using a number of protocols and information access methods (HTTP/XML, direct Oracle connections, flat text files, etc.) to build a single and monolithic topology model of the grid.

Consequently, the basic guidelines for the new approach are the following:

- define core vocabulary (namespace or ontology) for concepts that are common for most of the grid tools, like: Service, VO, etc.
- define namespaced vocabularies for individual sources of topology information: BDII (Glue), GOCDB, VO specific etc.
- expose information provided by the topology data sources as RDF
- use messaging system (MSG) to publish and subscribe for instantaneous topology changes
- use local caching wherever possible (local RDF stores or equivalent in monitoring tools)
- use core vocabulary and in future ontology specifications (OWL, reasoning) to 'glue' together information coming from various sources

## Information representation and annotation

The topology information can be easily represented as RDF triples. However, because of different validity lifetime of information, level of authoritativeness, and other factors depending on the source and type of information, a special care has to be taken to provide additional annotation or meta-data. This meta-data should contain at least the following information:

- original source of the information - who produced the information (used to identify authoritativeness)
- assertion time - when the information was actually produced (freshness)
- declared validity time - until when the producer declares the information to be valid
- imposed validity time - for how long from the assertion time the information coming from a given source and of a given type should be considered valid (according to a policy on the ontology level, no matter of declared validity), this type of meta-data can be defined on ontology level as an inferable rule

There are at several ways to represent this kind of meta-data in RDF:

- using RDF reification - quite complex to maintain and query, can be heavy in storage (triple storage bloat)
- using contexts or sub-graphs - RDF store implementation specific
- using 'fake' annotation - additional properties or annotation objects pointing to the resources

## Core vocabulary

## Information transport

## Query/response paradigm

## Publish/subscribe paradigm

## Local information caching

## Information integration and equivalence

-- PiotrNyczyk - 12 Mar 2008

---

This topic: LCG > TopologyDBIntegrationProposal1

Topic revision: r2 - 2008-03-13 - PiotrNyczyk



Copyright &© 2008-2019 by the contributing authors. All material on this collaboration platform is the property of the contributing authors. Ideas, requests, problems regarding TWiki? Send feedback