Editor:
Morris Riedel (FZJ, OMII-Europe)
2/10/2007

GIN – Information Services and Schema (GIN – INFO)

Experiences from interoperation scenarios in production Grids

DRAFT version 1.00

Status of This Memo

This memo provides experimental information to the Grid community regarding interoperation scenarios within production Grids. It does not define any standards or technical recommendations. Distribution is unlimited.

Abstract

Many production Grid projects have begun to offer services to end-users during the past several years with an increasing number of application projects that require access to a wide variety of resources and services in multiple Grids. Therefore, the purpose of the Grid Interoperation Now (GIN) Community Group (CG) of the Open Grid Forum (OGF) is to organize and manage a set of interoperation efforts among production Grid projects that require resources in multiple Grids. The major goals of the group are several interoperation demonstrations held at the Supercomputing 2006 (SC 2006) conference in Tampa, FL, USA.

The focus of this group is to plan and implement interoperation in five specific areas such as authorization and identity management (GIN-AUTH), data management and movement (GIN-DATA), job description and submission (GIN-JOBS), information services and schema (GIN-INFO), and operations experience of pilot test applications (GIN-OPS).

This document focusses on the results of the interoperation efforts within the GIN – INFO area and thus provides feedback into the interoperability efforts being conducted by the working groups of OGF. The GIN – INFO area focus on interoperation of information services and schema to identify appropriate resources within production Grid infrastructures that are typically provided through standard schemas supporting a query mechanism.

The outcome of these interoperation efforts was shown at the SC2006 conference by using a Google Earth demonstrator that relies on information provided by a GIN Berkeley Database Information Index (BDII). In particular, a lot of Grid sites that participate in production Grids were shown with Google Earth based on various pieces of information provided by different information services and schemas.

The efforts include interoperation of three different schemas that are used in production Grids today. Firstly, the GLUE schema that is mainly used within the OSG, EGEE, and TERAGRID production Grids. Secondly, the ARC schema used within the NDGF production Grids and finally the CIM schema with dedicated NAREGI vendor extensions used within the NAREGI production Grids. Finally, several information services were used while the most of them rely on the Lightweight Directory Access Protocol (LDAP).

Contents

1. Introduction

The Grid Interoperation Now (GIN) Community Group (CG) contributes to the broader Grid community in the context of the e-Science Function of the Open Grid Forum (OGF) within the area of Grid Operations.

The major goal of the GIN-CG is to work in the area of interoperation between production Grids. In the following the difference between interoperability and interoperation is as follows: Interoperation is specifically defined as what needs to be done to get production Grids to work together as a fast short-term achievement using as much existing technologies as available today. Hence, this is not the perfect solution and different than interoperability that is defined as the native ability of Grids and Grid middleware to interact directly via common open standards in future.

The focus of the group is to plan and implement interoperation in the following specific areas:

- (1) Job Submission (job description and execution) GIN-JOBS
- (2) Data management (data location and movement) GIN-DATA
- (3) Information services and schema (information about data, sites) GIN-INFO
- (4) Security Discussions (authN/authZ and identity management) GIN-AUTH
- (5) Operations / Initial Applications (applications to test interoperation) GIN-OPS

The 'Information services and schema' area is in the following named as GIN-INFO area.

In order to identify appropriate resources for users within a Grid infrastructure there must be some form of resource information conforming to schemas and accessible via standard query mechanisms. The GIN-INFO area of GIN tries to provide interoperation components (e.g. schema translators, adapters, providers) to solve the wider usage of different information schemas such as GLUE [GLUE-WG], the ARC schema [ARC-SCHEMA] and solutions based on CIM [DMTF-CIM]. The efforts in the GIN-INFO area build upon previous bi-lateral successes between EGEE and OSG [EGEE-OSG], EGEE and NDGF [EGEE-NDGF] and other bi-lateral interoperability efforts in various Grid projects and production Grids. The main idea of the work is to identify a subset of information items that can be used as a common minimum set. Furthermore, several interesting work has been done on translation of these information items between the different schemas. The GIN-INFO area is coordinated by the following individuals:

- Laura Pearlman (ISI)
- Satoshi Matsuoka (TIT)

At the time of writing, the participating Grid projects are as follows:

Project	Contact	Schema
APAC	Gerson Galang	GLUE
DEISA	Thomas Soddemann	GLUE
EGEE	Laurence Field	GLUE
NAREGI	Yuji Saeki	CIM (NAREGI extensions)
NDGF	Balazs Konya	ARC
NGS	Matt Viljoen	MDS2.4
OSG	Shaowen Wang	GLUE
PRAGMA	Sugree Phatanapherom	GLUE
TERAGRID	Laura Pearlman	GLUE

Table 1: Participating Grid sites and their information schemas

Several interoperation efforts where demonstrated at the Supercomputing 2006 by using Google Earth showing all these participating Grid sites. Finally, interested individuals can find the mailing list archives here: http://www.ogf.org/pipermail/gin-info/ .

1.1 Terminology

The keywords "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [**RFC 2119**].

2. Technology Foundations

The wide variety of production Grids developed many different technologies over the last years. The described interoperation scenarios within this document are related to the following technology foundations.

2.1 Basic Technologies

To understand the technology foundations for the interoperation scenarios several basic technologies in the area of directory services have to be introduced in this section.

2.1.1 Directory Services

Directories are specialized databases optimized for reading, browsing and searching of information. In general, directories contain descriptive, attribute-based information and support a wide variety of filtering capabilities. The update of the provided pieces of information are typically based on all-or-nothing changes and thus do not suport transaction or roll-back mechanisms found in Database Management Systems (DBMS).

The major goal of directories is that they are tuned to give very quick responses to high-volume lookups or search operations. Also, directories typical support the replication of information in order to increase availability and reliability. Finally, to provide an example, the Internet Domain Name System (DNS) [RFC 3467] is a rather global distributed directory service.

For more information please refer to [DS].

2.1.2 Lightweight Directory Access Protocol (LDAP)

LDAP [RFC 2251] in version 3 is a lightweight protocol for accessing directory services, specifically X.500-based directory services. [RFC 2251] and other documents comprise the LDAP technical specification [RFC 3377]. An X.500 directory service is named as Open System Interconnection (OSI) directory service and typically accessed via the X.500 Directory Access Protocol (DAP). Typically LDAP clients are used to access gateways to the X.500 directory service. In more detail, the LDAP protocol is used between the LDAP client and gateway and the DAP protocol between the gateway and the X.500 directory server.

Since DAP is a rather heavyweight protocol that operates over a full OSI protocol stack, the usage of it requires a significant amount of computation. On the other hand, LDAP is using TCP/IP and thus provides with much lower computation most of the functionality that DAP provides. A lot of systems rely on LDAP to access an X.500 directory service via Gateways. But more recently, LDAP is more commonly implemented in X.500 servers. Hence, only LDAP is used from the client to the X.500 directory server without DAP. Furthermore, the so-called slapd is a stand-alone LDAP daemon that represents a lightweight X.500 directory server, for instance in the OpenLDAP implementation [OPEN LDAP]. In other words, it does not implement the X.500 DAP and thus provides only a subset of the X.500 DAP by still providing enough functionality for many use cases.

In typical deployment, one or more LDAP servers contain data that together represent the Directory Information Tree (DIT). This data or more precisely, the LDAP information model is based on entries (e.g. a server within the Internet) that provide a certain amount of attributes. Each attribute has a type and one or more values. Each entry has a Distinguished Name (DN) that uniquely identifies it. The DN in turn is formed from a sequence of attribute/value pairs that lays the foundation to arrange the DNs entries into the hierarchical tree-like DIT structure.

Since LDAP is a client-server protocol, a client connects to a server requesting pieces of information. The respond can be the requested information and/or a pointer where the client can get additional information, typically another LDAP server. In this context it is important to understand that it does not matter which LDAP server the clients connect to, because all provide the same view of the directory. To provide an example, the ldapsearch command represents such a LDAP client that search for and retrieve entries provided by a LDAP server.

TBD: Example: LDAP URL Lists?

2.1.3 LDAP Data Interchange Format (LDIF)

The LDIF [RFC 2849] is used to represent LDAP entries in a text format.

TBD: LDIF example

2.2 Grid Information models and schemas

The wide variety of Grid resources that are available within production Grids must be described in a precise and systematic manner to be able to be discovered for subsequent management or use. Over the years several schemas evolved either in scientific-oriented production Grids or within business oriented standard bodies such as the Distributed Management Task Force (DMTF).

2.2.1 Grid Laboratory Uniform Environment (GLUE) Schema 1.3

The GLUE schema [GLUE 1.3] in version 1.3 represents an abstract model for Grid resources and mappings to concrete schemas that can be used by Information services within Grids. Hence, the underlying idea of the schema is to provide an information model that can be used to exchange pieces of information among different knowledge domains. The GLUE schema was initially used within EGEE and becomes one of the most known schemas related to Grid information modeling.

At the time of writing, mappings for LDAP, XML schemas and OldClassAds exist. But in principle, the proposed information model of GLUE can also be mapped on data models that are specific of other Information systems. Note that the proposed information model is not tied to any particular implementation.

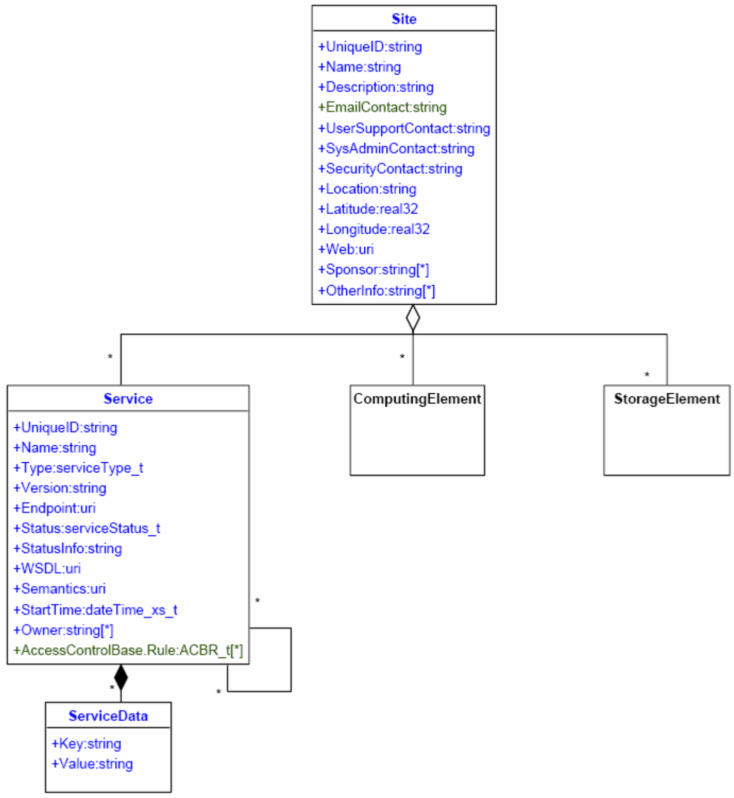

Figure 1: Core Entities of the GLUE schema: Site, Service and ServiceData. In addition there are scheme definitions for a ComputingElement and StorageElement [GLUE 1.3].
TBD: Original UML diagram for better quality → Sergio, Laurence

In more detail, the GLUE schema is defined by using several UML diagramms. These definitions provide a description of core Grid resources at the conceptual level by defining an information model that is an abstraction of the real world into constructs that can be represented in computer systems (e.g. objects, properties, behavior, and relationships). The GLUE schema specification defines so called core entities that are shown in Figure 1, namely the Site, Service and ServiceData entities. In addition to them there are schema definitions for a ComputingElement and StorageElement [GLUE 1.3].

More recently, the GLUE 2.0 working group of OGF emerged with the goal of providing a major revision of the existing GLUE schema that address additional new use cases in that area. Hence, the group tries to incorporate the existing experiences of the current production GLUE schemas and address the requirements imposed by newly collected and evaluated use cases for an information model in Grids. Also, the GLUE 2.0 schema specification will be independent of an information system implementation

For an intensive insight about GLUE 1.3 please refer to [GLUE 1.3] and for more plans about GLUE 2.0 refer to [GLUE-WG]. Note that the GLUE schema is used in EGEE, OSG, and TERAGRID.

2.2.2 NorduGrid / ARC information model

The ARC [ARC NDGF] middleware implements a scalable and dynamic distributed information system. In particular, it is an OpenLDAP [OPENLDAP] – based system that makes use of the OpenLDAP modifications provided by the Globus MDS2. In an LDAP-bases system the pieces of information are stored as attribute-value pairs grouped together in entries that are stored in a hierarchical tree. Using this technology, the ARC information model describes major Grid components of the NDGF such as computing resources with Grid jobs and Grid users as well as storage elements and metadata catalogs. In general, these pieces of information are stored using Object Identifiers (OID) for attributes while the values are normal text (e.g. 1.3.6.1.4.1.11604.2.1.3 OID maps to 'queue objectclass').

For more information about the ARC schema please refer to [ARC NDGF]. Note that this schema is used within the NDGF.

2.2.3 Common Information Model (CIM) Schemas

The CIM schemas [DMTF-CIM] provide a common definition of management information for systems, networks, applications and services that is standardized at the DMTF and thus rather business oriented. CIM is a conceptual information model for describing computing and business entities in enterprise environments. The fundamental goals of CIM are common definitions that enable vendors to exchange semantically rich management information between wide varieties of systems. In particular, CIM consists of a specification and a schema that provides model descriptions. The specification defines the details for integration with other management models

In the context of GIN – INFO it is extremely important that CIM allows for vendor extensions that are used by NAREGI. Furthermore, CIM provides guidelines for CIM-to-LDAP directory mappings. To sum up, CIM does not require any particular instrumentation or repository format since it is only an information model used for unifying data by using object-oriented constructs. Finally, the rather unique design of CIM that distinguishes it from other rather flat data formats is abstraction and classification by object inheritance. On the other hand, CIM is a huge set of classes and thus provides a rather hugh amount of schemas.

For more information about CIM please refer to [DMTF-CIM]. Note that CIM is used with dedicated NAREGI vendor extensions within the NAREGI production Grid.

2.3 Information Services in Grids

An Information Service (IS) provides pieces of information about a Grid infrastructure that consists of a wide variety of Grid resources that all have an individual actual status. These pieces of information are typically used for various Grid operations that reach from resource discovery and monitoring to accounting and job execution purposes. Note that the most of the data that is published at an IS conforms to a dedicated information schema (e.g. the GLUE or CIM schema) that defines a common conceptual data model to be used for Grid resource monitoring and discovery.

Over the last years different information services evolved within production Grids that are shortly introduced in this section. For example, the gLite 3 [GLITE] Grid middleware uses two information services, the Globus-based Monitoring and Discovery Service (MDS) for resource discovery and to publish the resource status and the Relational Grid Monitoring Architecture (R-GMA) for accounting, monitoring and publication of user-level information. While Globus-based systems rely on the MDS by using an LDAP server named as Site Grid Index Information Server (GIIS), the gLite MDS uses another version of the GIIS that works with the Berkeley Database Information Index (BDII). Hence, there are often services that provide the same functionality but base on different technologies.

2.3.1 MDS of gLite 3

The MDS described in this paragraph is one of the IS of gLite 3 that is used for resource discovery and to publish the resource status within the EGEE production Grid infrastructure. In particular, the MDS implements the GLUE schema using OpenLDAP [OPEN LDAP]. In gLite, the computing and storage resources run an Information Provider that generates the relevant information about the resource, for instance static type information about the storage element or dynamic information such as the used space in storage elements. These pieces of information are published via an LDAP server named as Grid Resource Information Service (GRIS) that typically runs on the resource itself. At each Grid site another LDAP server named as the GIIS collects all the pieces of information from the local GRIS servers and repusblishes it. In the context of gLite this GIIS uses the BDII to store data, because at the time of integration into gLite it was more stable than the original Globus GIIS [GLITE USERGUIDE].

In more detail, the BDII consists of two ore more standard LDAP databases that are populated by an update process. The update process obtains LDIF from either doing an ldapsearch on LDAP URL lists or by running a local script that generates LDIF. Afterwards, the LDIF is then inserted int the LDAP database as an entry. In particular, an update process starts with the update of configuration files that contain LDAP URLs. Next, it forks off processes to obtain the LDIF entries from the URLs that are later added into the database.

Finally, there is typically also a top-level-BDII on the top level of the hierarchy that reads data from a specific set of Grid sites and thus provides a view of the overall Grid resources. In detail, these BDIIs query the GIIS systems at every Grid site and act as a cache by storing information about the requested status in their databases. Hence, such a top-level-BDII consists of all the available pieces of information about the defined Grid sites. As a sideremark, information about specific resources can also be requested by directly contacting GIISes or even the GRISes. Within the EGEE infrastructure, the top-level BDIIs obtain information about the Grid sites in from a so-called Grid Operations Centre (GOC). At these GOCs site managers can insert endpoints for GIIS systems as well as other useful information about the Grid site.

For more information please refer to [GLITE USERGUIDE]. Note that this kind of information system is used within the EGEE and OSG production Grids.

2.3.2 Relational Grid Monitoring Architecture (R-GMA) of gLite 3

TBD: to be comprehensive – but not used in GIN – INFO…

2.3.3 NorduGrid / ARC Grid Information System

The ARC middleware provides a dynamic LDAP-based distributed information system via set of coupled lists named as Index Services that use local LDAP databases. In more detail, the overall information system consists of the following three main components: Local Information Tree (LIT), Index Service (IS) and Registration Processes (RP)

The LIT are responsible for resource decriptions of characteristics in the area of computing or storage. The local information is generated on the resource and optionally cached. This information can be queried by clients via LDAP interfaces. In this context, the IS is responsible to maintain dynamic lists or resources that are realized by LDAP URLs of the LITs and other ISs. Hence, ISs can be further coupled together and thus realize a specific hierarchical topology. Finally, the LITs make use of RPs running locally on the Grid resources in order to list themselves in some of the LDAP lists maintained by the ISs. Usually, the information content of these information systems are queried using the LDAP protocol.

For more information about the NorduGrid / ARC information system please refer to [ARC INFOSYS]. Note that this information system is used within the NDGF production Grid.

2.3.4 MDS2.4 of Globus Toolkit 2

TBD: based on Site Grid Index Information Server (GIIS)

For more information about MDS2.4 please refer to [MDS GT2.4]. Note that the MDS2.4 information model is used by the NGS.

2.3.5 Monitoring and Discovery System (MDS) of Globus Toolkit 4

The Monitoring and Discovery System (MDS) described in this paragraph is the monitoring and discovery system of the Globus Toolkit 4 [MDS GT4]. In particular, this MDS represents a whole suite of Web services to monitor and discover resources and services on Grids. In the context of GIN-INFO, the MDS services provide query and subscription interfaces to arbitrarily detailed resoure data and other pieces of information. In addition, it also provides interfaces to third-party monitoring systems.

In more detail, MDS4 includes several OASIS Web Services Resource Framework [WS-RF] based services such as the Index Service. The Index Service collects data from various sources and provides a query/subscription interface to that data. In other words it represents a registry that is similar to the Universal Description, Discovery and Integration (UDDI), but much more flexible. An

Index Service collects information and publishes that information using WS-Resource properties [WS-RF] that provide a standard set of message exchanges and property access for resources. Hence, clients that use these Index Services use the WS-RF-based query message exchanges to retrieve information from an Index.

To provide more complex scenarios, the Index Services can register to each other in a hierarchical fashion and thus aggregate data at several levels within Grids. Note that each entry within the service has a lifetime and will removed from the Index Service if it is not refreshed before the lifetime expires. Typically, each Globus container that has MDS4 installed will automatically provide a default Index Service instance and thus ans other service running in the container will register itself to this Index Service.

In the context of collecting and aggregating data, GT4 also provides the so-called Aggregator Framework that can be used to build aggegator services that collect information via Aggegator Sources. An Aggegator Source implements an interface to collect XML-formatted data and is also able to use an external software compont. These external software components are named as Information Providers.

MDS4 includes, for instance, a Hawkeye Information Provider that gathers Hawkeye data about Condor pool resources using the XML mapping of the GLUE schema. Furthermore, a Ganglia Information Provider that gathers cluster data from resources running Ganglia using also the XML mapping of the GLUE schema. In addition, MDS4 provides the WebMDS Web front-end to the Index Service that displays the results of WS-Resource properties queries. Finally, the MDS4 system is mainly used to discover needed data from services in order to use them in conjunction with job submission. Also, it allows for the evaluation of the status of Grid services within Grids.

For more information please refer to [MDS GT4]. Note that this kind of information service is used within the TERAGRID production Grid.

3. Interoperation Scenarios at Supercomputing 2006

The following projects participated in interoperation scenarios at the Supercomputing 2006 using the following overview of technologies:

Project	Contact	Schema	Datamodel	Interface	Client
APAC	Gerson Galang	GLUE	LDIF	LDAP	ldapsearch
DEISA	Thomas Soddemann	GLUE	XML	WS	WS
EGEE	Laurence Field	GLUE	LDIF	LDAP	ldapsearch
NAREGI	Yuji Saeki	CIM (NAREGI)	???	???	NAREGI API
NDGF	Balazs Konya	ARC	LDIF	LDAP	ldapsearch
NGS	Matt Viljoen	MDS2.4	LDIF	LDAP	ldapsearch
OSG	Shaowen Wang	GLUE	LDIF	LDAP	ldapsearch
PRAGMA	Sugree Phatanapherom	GLUE	LDIF	HTTP	wget
TERAGRID	Laura Pearlman	GLUE	XML	WS	WS-RF query

Kommentar [MRi1]: TBD: GET this information from NAREGI

Table 2: Overview of the production Grids that participate in the GIN – INFO area, including their used schemas, datamodels, interfaces and clients.

3.1 Identified problems and issues for interoperation

This section provides an overview of several encountered problems and questions thar arised during the interoperation efforts.

3.1.1 Identification of Participating Sites

The identification of participating Grid sites was one of the initial challenges during the interoperation efforts. In particular, the identification process tried to to get production Grid involved so that they agree to advertise their Grid sites in their local information schema exposed via an information service.

Also, the amount of identified Grid sites within production Grids raised the demand for a uniqueID. However, common rules to choose such a uniqueID in cross-Grid domains are a real challenge. As a sideremark, this is rather similar to the VO naming problem that came up in the GIN – AUTH area. One possible solution for the uniqueID is the use the actual domain.

3.1.2 Lack of Information

Not all information items or attributes can be found in all Grids that lead to missing data for one Grid site when using it in conjunction with other Grids that provide such information. Therefore, an implementation may have errors using this information. Of course these problems also arise if Grid sites not publish this information correctly (e.g. not schema compliant). In this context it seams reasonable to consider that this information lays the foundation for service discovery of job execution endpoints (e.g. OGSA-BES interfaces [OGSA BES], gLite Ces [GLITE], or UNICORE Atomic Services [UNICORE]). Thus, missing or incorrect data could lead to problems when using this data in real use cases such as data movement between sites or a simple job execution. Hence, the lack of information disturbs the work of e-Scientists. To sum up, the pieces of information required will be defined by the use cases and thus raise the demand to identify pieces of information required for cross-Grid use cases.

3.1.3 Set of Common Minimal Attributes

The participating production Grids evolved over time and thus use a wide variety of technologies that deal with more or less the same pieces of information, but in different sets or named as different attributes using different information schemas. Therefore, when doing interoperation, it is extremely important to agree on a common minimal attribute set that all participating sites must provide (e.g. longitude and latitude information). Hence, the contents of the different attributes and schemas should at least consist of the defined common minimal set of attributes. Of course, this common minimal set depends on the use case, for instance in the area of job execution the endpoint attributes of a procuction Grid that offer a job execution interface MAY be published by an information system.

In addition to the agreement on the attributes for a site entity, an agreement on the attributes for the service entities is needed. This is particularly useful to list services that are hosted at a Grid site that raise the demand to be able to link services to Grid sites.

3.1.4 Information Quality

Not only the quantity of information matters, also the quality of an information system is important that depend on the quality of information. There are many reasons for poor quality of information such as a poor schema design or unsatisfied developed information providers. In addition, incorrect deployments or configurations as well as site problems can have an impact on the quality of the information. Hence, one identified problem within the GIN – INFO area is related to the question of how good information quality can be ensured within Grids. That raises the demand for tests of information by the means of test cases based on use cases. To provide an example, how can it be ensured that the coordinates are correct for a Grid site? To sum up, it must be somehow ensured that all participating Grid sites publish information in a good quality, for instance the common minimal set of attributes.

3.1.5 Translaters between Information Systems

In addition to the requirement of having a common minimal set of attributes for the content of information it is also necessary to map the different schemas to each other in order to put the contents into the right places within other schemas. Hence, translaters must be developed that map the content of one attribute in schema A to the content of one attribute (or even a set of attributes) in schema B.

3.2 Existing Bi-literal Activities in Production Grids

3.2.1 EGEE / OSG Interoperability

The EGEE and the OSG production Grids are already interoperating since autumn 2005. The interoperability is achieved by the usage of LDAP based information systems and the GLUE schema in both Grid infrastructures. However, there is a different boot strapping. While OSG site URLs are generated from OSG GOC DB, the EGEE site URLs are generated from EGEE GOC DB.

For more information about the EGEE / OSG Interoperability please refer to [EGEE OSG].

3.2.2 EGEE / NDGF Interoperability

The EGEE and the NDGF production Grids are working on interoperability since summer 2005. Both Grid infrastructures use an LDAP based information system. But while EGEE uses the GLUE schema, the NDGF uses the ARC information system that bsae on a different schema. This acitvity provides already basic schema translations and an official mapping document and prototype is ready, including reverse translations. The bootstraping is done via site URLs and a top-level GIIS.

For more information about the EGEE / NDGF interoperability please refer to [EGEE NDGF].

3.3 Interoperation of Different Information Services and Schemas

The interoperation of the different information services and schemas within participating production Grids was firstly achieved by an initial architecture that will be shortly presented in this section. Afterwards, this section will focus on the final architecture and the particularly efforts within production Grids for establishing interoperation according to this architecture.

3.3.1 Initial Architecture

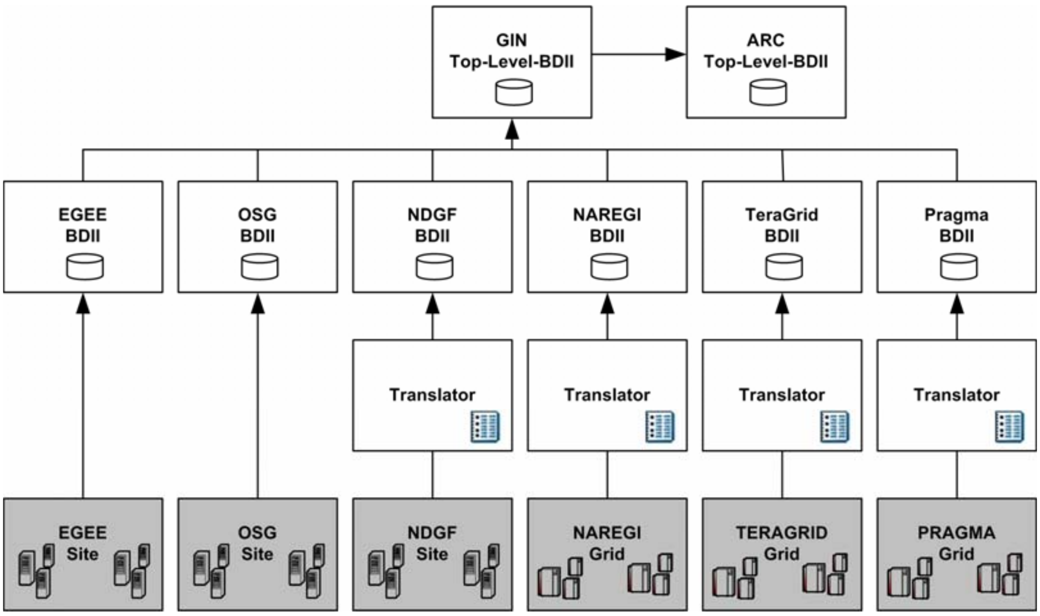

Figure 2: Initial architecture of the BDII-based interoperation between different information services and schemas. The interoperation was achieved using translators for those Grids that not naturally support the BDII system.

There are several different schemas used in production Grid projects today. The idea of having translators evolved in the EGEE / NDGF interoperabilty activity, introduced in the previous section, was the idea in the beginning for the interoperation within GIN – INFO.

As shonw in Figure 3, the GIN-BDII is a top-level BDII with information from all Grids in accordance to the GLUE schema in version 1.3. In particular, the GLUE schema has a site entry that is able to provide such information needed during interoperation. In the context of the GIN – INFO interoperation efforts the following attributes are considered to be MANDATORY:

- Site Location, more precisely the latiture and longitude
- Site Name
- Unique Identifier for the site

In addition to these attributes, the following attributes are considered to be OPTIONAL:

- Site Description
- Site Location in a human readable format
- Site email contact
- Site web page

Hence, all the MANDATORY attributes for a Grid site are provided by the GIN Top-Level-BDII. In addition, information about a Grid site MAY include several of the listed OPTIONAL attributes.

TBD: More about the translators used in the first approach and more about the representation of attributes

3.3.2 Final Architecture

Slightly different as the initial architecture is the final architecture that achieved the interoperation among a wide variety of production Grids at the SC 2006. Also in the final architecture, the GIN Top-Level-BDII contains information of varying quality from all participating Grid sites.

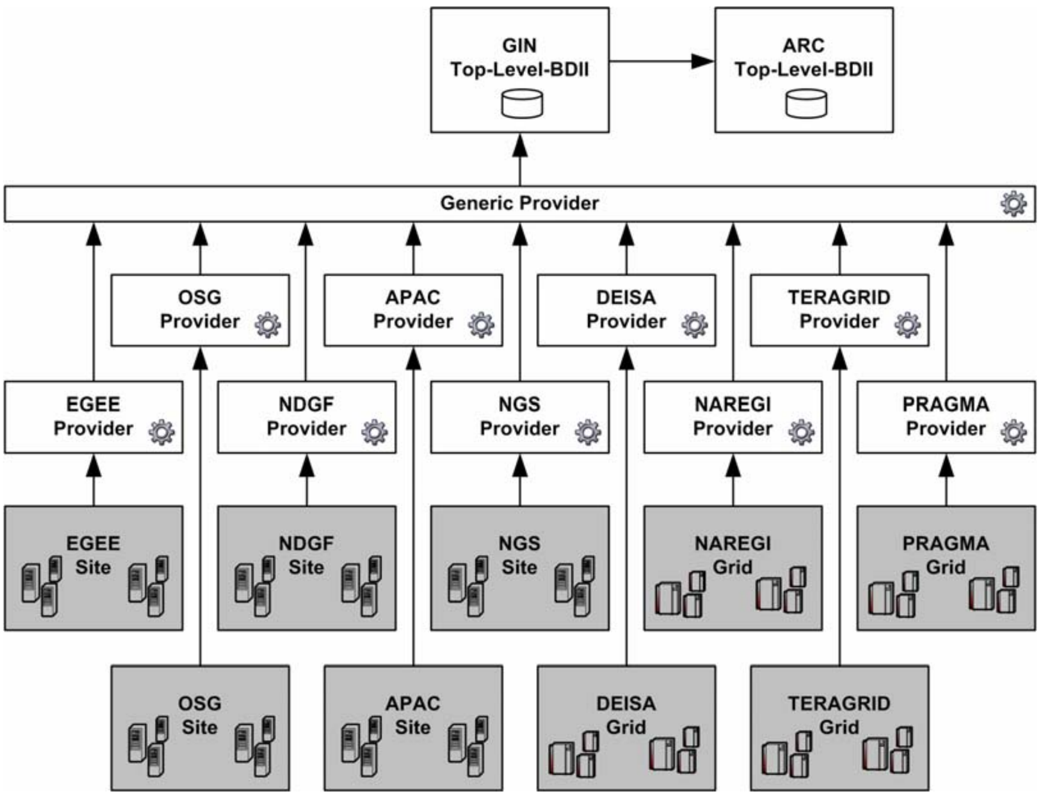

Figure 3: Final architecture of the BDII provider-based interoperation between different information services and schemas. Each Grid Site or participating Grid infrastructure uses a dedicated provider to publish the information via the generic provider to the GIN Top-Level-BDII.

TBD: Integrate the GIN CELL DOMAIN in the figure ad well as LDAP parts and translators

As shown in Figure 4, for each participating production Grid an information provider has been developed. In particular, the information provider queries the correspondent Grid site and outputs the information in LDIF in accordance to the GLUE schema. In turn these information providers are used to provide the GIN Top-Level-BDII with information from all participating infrastructures using a generic provider in between. In turn, these pieces of information lay the foundation for the ARC-BDII that is also a Top-Level-BDII with information from all Grid infrastructures in accordance to the ARC schema.

// TBD: more on exposure of schemas via info services

The next paragraphs provide more in depth details of how the interoperation was achieved.

3.3.3 APAC

The top-level MDS GIIS can be queried in the following way:

```
ldapsearch -x -h ng1.vpac.org -p 2135 -b mds-vo-name=apac,o=grid
```

TBD: More about the APAC provider

3.3.4 DEISA

DEISA relies on the UNICORE Grid Middleware [UNICORE] and offers information that can be queried via Web services that follow a dedicated WSDL.

TBD: Parts of WSDL maybe

By using goap, the C header file can be created and compiled with the following command:

```
wsdl2h -c -o GINService.h
http://gin.deisa.org/ginservice/services/GINService?wsdl
soapcpp2 -C -L -c GINService.h
gcc -O3 -o client Client.c soapClient.c soapC.c -lgsoap
```

Afterwards the Client can be used for queries.

TBD: More about the DEISA provider

3.3.5 EGEE / OSG

The GLUE schema 1.3 is used by OSG [OSG], EGEE [EGEE].

The interoperation uses a list of LDAP URLs that represent EGEE sites and another list of LDAP URLs that represent OSG sites. These URLs lay the foundation for the following ldapsearch query:

```
Site-Name ldap://<host>:<port>/bind
ldapsearch -x -h <hots> -p port -b <bind>
```

Note that for the bind point the forward slash should be change to a comma.

TBD: More about the EGEE/OSG provider

3.3.6 NAREGI

NAREGI [NAREGI] is using the CIM schema with a vendor extension. In particular, the NAREGI information systems are named as ‘cell domains’ that have CIM Providers and expose OGSA-DAI [OGSA DAI] interfaces.

A translator working with the OGSA – DAI interface was developed using the API for querying provided by NAREGI.

The GIN Cell Domain is a NAREGI cell domain with information from all Grids in accordance to CIM and the NAREGI vendor extension.

TBD: More about the NAREGI provider

3.3.7 NDGF

The NDGF [NDGF] are using the ARC schema. In GIN – INFO the work of the EGEE / NDGF interoperability (see Section ….) was further improved. The ARC information system basically relies on LDAP and thus the query for information is quite similar to EGEE/OSG.

The ARC-BDII is a top-level BDII with information from all Grids in accordance to ARC.

The interoperation uses a list of LDAP URLs that represent NDGF sites. These URLs lay the foundation for the following ldapsearch query:

```
Site-Name ldap://<host>:<port>/bind
ldapsearch –x –h <hots> -p port –b <bind>
```

Note that for the bind point the forward slash should be change to a comma.

TBD: More about the NDGF provider

3.3.8 NGS

The top-level MDS GIIS can be queried in the following way:

```
ldapsearch –x –h ngsinfo.grid-support.ac.uk –p 2135 –b ,o=grid
```

TBD: More about the NGS provider

3.3.9 PRAGMA

PRAGMA uses WebSIM.

A prototype translator was developed that generates a Web based LIDF file using wget and ldapadd as follows:

```
wget "http://pragma-goc.rocksclusters.org/cgi-
bin/scmsweb/glue.cgi?format=ldif&compat=1" –O pragma.ldif
```

TBD: More about translator and provider here in PRAGMA

3.3.10 TERAGRID

TERAGRID uses the MDS4 system of the Globus Toolkits 4 that unfortunately follows the rather old version 1.1 of the GLUE schema. Therefore, also a prototype translator was developed. After installing GT4, wsrf-query command can be used as follows:

```
Wsrf-query –a –s
https://mds.teragrid.org:8443/wsrf/services/DefaultIndexService
```

TBD: More about GLUE 1.1 translator. And TERAGRID PRovider

3.4 Initial Use Case: Grid Sites on a Google Earth Map

Some information exposes by the different information services of the production Grids can be shown within Google Earth. This was demonstrated at the SC 2006 conference. In more detail a Google Earth KML file is generated from information in the GIN Top-Level-BDII introduced in the previous sections. All participating Grid sites published an identifier for the site, the latitude and the longitude (the minimal common set) and by using this information it was possible to use Google Earth to show the location of the sites and the Grid affiliation. To sum up, a script that runs every 5 minutes via a cron job is used to generate a KML file for the visualization of Grid sites within Google. This script queries the GIN BDII to get information such as longitude or latitude.

For more details about the KML file refer to [GIN-INFO-KML]

```
<?xml version="1.0" encoding="UTF-8"?>
<kml xmlns="http://earth.google.com/kml/2.0">
<Document>
…
  <Style id="EGEE">
    <IconStyle>
      <color>ffffaaaa</color>
    </IconStyle>
    <LabelStyle>
      <color>ffffaaaa</color>
    </LabelStyle>
    <BalloonStyle>
      <text><![CDATA[<b>$[name]</b><br/><br/>$[description]]]></text>
      <color>ffffaaaa</color>
    </BalloonStyle>
  </Style>
…
</Document>
</kml>
```

Figure 4: Small part of the Google Earth KML file that defines a style for the EGEE Grid.

Beside the definition of the styles for the different production Grids the exact location of the production Grid sites must be provided as follows in the same KML file. Note that each participating Grid MUST provide for the 'Grid-on-a-map' use case a uniqueID, latitude, longitude and its location. Furthermore, a participating Grid site SHOULD have attributes such as Name, Description, Email and Web. In addition it CAN have optional tagging attributes.

```
…
  <Folder>
    <name>Switzerland</name>
    <Placemark>
      <name>CERN-PROD</name>
      <Snippet>EGEE</Snippet>
      <styleUrl>#EGEE</styleUrl>
      <description>EGEE<p/>Location: Geneva, Switzerland
        <br/>Description : LCG Site
      </description>
      <Point>
        <coordinates>6.0458,46.2325,0</coordinates>
      </Point>
    </Placemark>
  </Folder>
…
```

Figure 5 : Small part of the Google Earth KML file that defines the exact location of the production site CERN within the EGEE production Grid.

Note that one <Folder> element can have multiple <Placemark> components. This is in particular useful when one Institution participates in different production Grids or provides more than one Grid resource within one production Grid.

TBD: image of google earth with sites maybe

3.5 Security Considerations

In the context of security it is important to consider which individuals or services are allowed to query which information from the information services. This leads to simple requirements in the area of role-based authorization and authentication that other areas of OGF provide as well as the GIN – AUTH area. In the interoperation scenarios presented here, no security mechanisms were used.

To provide an example, the access to the MDS data used within gLite is typically insecure. In particular, there are no Grid credentials required for reading (clients and users) or writing (services publishes information) [GLITE USERGUIDE].

3.6 Relationship to other areas in GIN

In the GIN – DATA area, the discovery of storage areas usable by a VO is a problem. Therefore it should be part of an information model or exposed by an information service.

The GIN – JOBS area would like to be able to find all the execution services to which members of a GIN VO has access.

The problem of having unique names for the wide variety of VOs across different production Grids raised an issue in GIN – AUTH that is similar to the problem of having a uniqueID for each participating Grid site within GIN – INFO.

Finally, the GIN – OPS area should provide use cases that identify attributes needed by 'cross-Grid' applications

4. Emerging standards used in interoperation and feedback

This chapter focusses on production Grids that not use systems that typically reflects the current status of the work done in OGF/OASIS/DMTF working groups. However, this chapter lists some working groups of standardization bodies that MAY gain some experiences of feedback from this document and thus from the early large-scale interoperation experiences.

4.1 Open Grid Forum

4.1.1 GLUE – WG

TBD: Impact on GLUE-WG → Laurence

4.1.2 OGSA-WG design team for information modeling

TBD: OGSA – WG design team information modeling based on CIM, additions to CIM that fit OGSA-BES and HPC-Profile, visio document…

4.1.3 Info Dissemination working group (INFOD-WG)

This group develops a model for information dissemination that supports asynchronous data and event distribution, data replication and third-party data delivery especially in real-time environments. Hence, this group focusses on real-time environments and thus the impact of GIN – INFO is quite small since interoperation of rather batch-oriented Grid environments was achieved. However, some overlaps MAY occur with GIN – INFO.

For more information about this group please refer to [OGF INFOD].

4.1.4 Grid Information Retrieval WG (GIR-WG)

This group establishes a specific set of requirements, architectures, and detailed specifications for Information Retrieval (IR) within Grids. In particular, the GIR will provide documents for collection management, indexing/searching and query processing services.

The scope of the group is rather close to the work done in GIN – INFO and thus some parts of this work MAY be of interest for this group, especially in the context of queries for information in production Grids (e.g. LDAP queries).

For an intensive insight into the GIR-WG work please refer to [GIR-WG].

4.2 Distributed Management Task Force

4.2.1 Common Information model

More Grid-specific schemas are required and the work of the OGSA – WG design team for information modeling is a good starting point.

TBD: More on CIM in that matter

5. Conclusions

All the described information systems are rather similar in their scope and partly also in their functionality. Therefore, the usage and development of information providers to populate systems was easier than expected, mainly because the query mechanisms to extract the data were often based on LDAP, basic WS calls or open standards such as WS-RF. To sum up, query an information system and populate another was rather straight forward.

This also lays the foundation for closer interoperability in future by the means of joining information systems. In particular, an information provider for one system can be developed that quries another rather easily.

In the context of the output of queries, mainly all information conforms to a dedicated schema. In general, the Grid community can life with different information systems but not with different content schemas (e.g. GLUE, CIM). If there is a use case that needs to be done across all Grid infrastructures, then the information we need for these use cases must be present and in agreement. The GLUE-WG in OGF gathering such use cases.

Another demand that comes out of the GIN – INFO interoperation scenarios is a common interface (connection and query language) to information. Nearly all participants in GIN-INFO agree that something like this is needed; however there was no success in achieving a broadly accepted consensus.

The creation of translators worked, but non-existing information can not be translated. Hence, the translation of information is sometimes a bit tricky. In more detail, moving information from one model to another is straight forward using simple mapping mechanisms, but if information is not able to be mapped, for instance, because of missing attributes, the translations becomes a problem.

The key idea behing GIN – INFO interoperations are information providers (with translators) that were used to setup the GIN Top-Level BDII (gin-bdii.cern.ch) that contains information from all the participating infrastructures. Using this information, different use cases can be tried. The first use case selected was the “site on a map” use case. For this use case, three MANDATORY values are required: An identifier for the Grid site, the latitude and the longitude of the location of the Grid site. Using this information it was possible to use Google Earth to show the location of the sites and the Grid affiliation. In this context, it is easy to handle manually the integration of a small number of Grid sites, but problems with a large number of Grid sites will appear. The exposure of maps becomes time consuming. Furthermore, manual integration with potentially bad geographical knowledge can lead to huge political errors. To conclude, the integration of new information in such systems raises the demand for automation capabilities.

The minimal set of attributes needed depends on the correspondent use case. The limiting factor was the information available in the information systems. The information required will be defined by the use cases and hence we need to identify pieces of information required for ‘cross-Grid’ use cases. But such use cases (e.g. job submission) should come from the relevant experts. In other words information system experts ensure the consistency of information and provide them, but its usage and requirements definition in real use cases should be done by experts of other areas, e.g. GIN – JOBS.

To sum up, to demonstrate the feasibility of the interoperation of different information systems and schemas, the GIN – INFO area demonstrated the Grid site on a world map use case. Hence, the GIN – INFO area has successfully shown that Grids can be interopreable in respect to information sharing, because the GIN Top-Level-BDII contains information about all participating production Grid infrastructures within GIN. Finally, note that the interoperation efforts have a rather proof of concept character and thus the overall quality can still be improved by ‘driving use cases’ in future.

6. Outlook

At the time of writing, a second phase of GIN is in preparation by defining a new charter with new milestones and goals. These milestones could potentially include the interoperation with new production Grids such as CROWN [CROWNGRID] or the collection of more ‘cross-Grid’ use cases. In addition, a closer collaboration with OGF working groups (GLUE-WG, etc.) could be defined in the second charter of GIN to ensure direct feedback to the respective groups.

In the context of GIN – INFO, the basic use case with the map can be extended to show the wide variety of services that are offered at a Grid site. In addition, next steps are also addressing improvement of the quality of information provided by the information systems. This means in future there is again an agreement on the sites and service entries necessary to ensure everyone publishes information correctly.

7. Abbreviations

APAC	Australian Partnership for Advanced Computing
BDII	Berkeley Database Information Index
DAP	Directory Access Protocol
DBMS	DataBase Management System
DEISA	Distributed European Infrastructure for Supercomputing Applications
DMTF	Distributed Management Task Force
DNS	Domain Name System
EGEE	Enabling Grids for E-ScienceE
GIIS	Grid Index Information Server
GLUE	Grid Laboratory Uniform Environment
GOC	Grid Operations Centre
GRIS	Grid Resource Information Service
LDAP	Lightweight Directory Access Protocol
LDIF	LDAP Data Interchange Format
NAREGI	National Research Grid Initiative
NDGF	Nordic DataGrid Facility
NGS	(UK) National Grid Service

OGF	Open Grid Forum
OSG	Open Science Grid
OSI	Open System Interconnection
PRAGMA	Pacific Rim Application and Grid Middleware Assembly
SC	Supercomputing
VO	Virtual Organization

8. Contributor Information

The following list represents the major contributors within the GIN – INFO area. The interoperation among the different production Grids would not have been possible without their invaluable efforts and continous support:

APAC
- Gerson Galang

DEISA
- Thomas Soddemann

EGEE
- Laurence Field
- Gidon Moont

OSG
- Shaowen Wang
- Liegh Grundhofer

NAREGI
- Yuji Saeki
- Hitoshi Sato
- Satoshi Matsuoka

NDGF
- Balazs Konya

NGS
- Matt Viljoen

PRAGMA
- Sugree Phatanapherom
- Somsak Sriprayoonsakul

TERAGRID
- Laura Pearlman

The interoperation efforts are all partly funded by the listed production Grid sites, except the documentation process to foster the feedback capabilities of GIN within OGF that is partly funded by the Open Middleware Infrastructure Institute for Europe (OMII – Europe) project.

Kommentar [MRi2]: Ask the group if we should add general Acknowledgements for each participating project.

9. References

gin-info@ogf.org

[APAC]
http://www.apac.edu.au/

[ARC INFOSYS]
http://www.nordugrid.org/documents/arc_infosys.pdf

[CROWNGRID]
http://www.crown.org.cn/en/

[DEISA]
https://www.deisa.org/

[DMTF CIM]
http://www.dmtf.org/standards/cim

[DS]
http://www.openldap.org/doc/admin22/intro.html#What%20is%20LDAP

[EGEE-NDGF]
http://twiki.cern.ch/twiki/bin/view/EGEE/EGEE-NDGF

[EGEE-OSG]
http://twiki.cern.ch/twiki/bin/view/EGEE/EGEE-OSG

[GIN-INFO-KML]
http://gridportal-ws01.hep.ph.ic.ac.uk/gin/gin-locations.kml

[GIR]
https://forge.gridforum.org/projects/gir-wg

[GLITE]
http://glite.web.cern.ch/glite/

[GLITE USERGUIDE]
https://edms.cern.ch/file/722398//gLite-3-UserGuide.pdf

[GLUE 1.3]
http://glueschema.forge.cnaf.infn.it/

[GLUE WG]
http://forge.gridforum.org/sf/projects/glue-wg

[EGEE-BDII]
https://twiki.cern.ch/twiki/bin/view/EGEE/BDII

[MDS GT4]
http://www.globus.org/toolkit/mds/

[NAREGI]
http://www.naregi.org/index_e.html

[NGS]
http://www.grid-support.ac.uk/

[OGF INFOD]
https://forge.gridforum.org/projects/infod-wg/

[OGSA BES]
https://forge.gridforum.org/projects/ogsa-bes-wg/

[OGSA-DAI]
http://www.ogsadai.org.uk/

[OPEN LDAP]
http://www.openldap.org/

[RFC 2119]
http://www.ietf.org/rfc/rfc2119.txt

[RFC 2251]
http://www.ietf.org/rfc/rfc2251.txt

[RFC 2849]
http://www.ietf.org/rfc/rfc2849.txt

[RFC 3377]
http://www.ietf.org/rfc/rfc3377.txt

[RFC 3467]
http://www.ietf.org/rfc/rfc3467.txt

[TERAGRID]
http://www.teragrid.org/

[UNICORE]
http://www.unicore.eu

[WS-RF]
http://www.oasis-open.org/committees/wsrf/