Grid Information System Interoperability: The Need For A Common Information Model

Laurence Field
European Organization for Nuclear Research, CERN
Geneva, Switzerland
Laurence.Field@cern.ch

Sergio Andreozzi
INFN-CNAF
Bologna, Italy
sergio.andreozzi@cnaf.infn.it

Balázs Kónya
Lund University
Lund, Sweden balazs.konya@hep.lu.se

Abstract

A fundamental building block of any Grid infrastructures is the Grid information service and the information model. The information model describes the entities and relationships between entities within the infrastructure along with their semantics. The realization into a concrete data model defines the syntax by which these concepts can be exchanged. This data model enables consumers of information to efficiently find the information they require and ensures that there is agreement on the meaning with the information producer. The need for a common information model is therefore critical for the seamless interoperation of Grid infrastructures. A number of example interoperation activities are presented which highlight this point and the requirement for a common schema in general. An attempt to achieve interoperability between multiple Grid infrastructures, which was demonstrated at Super Computing 2006, helped motivate work on a common schema within the Open Grid Forum. The result of this effort, GLUE 2.0, which in itself defines the current view of Grid computing, is presented.

1 Introduction

Over recent years, a number of Grid projects have emerged that have built Grid infrastructures which are now the computing backbones for various user communities. A significant number of these communities are limited to using one Grid infrastructure due to the different middleware and procedures used in each Grid. Grid interoperation is trying to bridge these differences and enable virtual organizations to access resources independent of the Grid project affiliation.

This paper highlights an number of bilateral interoperations activities and focuses on methods used for achieving information system interoperability. Section 2 highlights one of the first Grid interoperations activities between European and US Grid initiatives. This exercise lead to the development of the GLUE schema as a solution for information system interoperability. Section 3 describes the interoperation activity between the EGEE project [16] and the Nordugrid [10]. This activity endeavored to achieve interoperation by translating between difference schemata and the problems faced with this approach are highlighted. The interoperations activity between EGEE and OSG [12] is reported in section 4, which showed that even if the same schema is used, differences due to versioning, interpretation and extensions can also create problems for interoperability. The GIN activity, outlined in section 5, attempted to achieve information system interoperability between many Grid infrastructures. This experience lead to the creation of the GLUE Working Group in the Open Grid Forum [4]. The output of this working group has been the definition of GLUE 2.0, which is presented in section 6. The potential impact of GLUE 2.0 as a solution to information system interoperation is discussed in the final section.

2 The Worldwide Grid Testbed

The need for a common schema was first highlighted in 2002 during the DataTAG [7] initiative to create a worldwide Grid testbed [9]. This testbed comprised of eight European sites providing services using the EDG[19] 1.2 release and 9 U.S. sites providing services from the VDT [6] 1.1.3 release.

Globus MDS [20] was at the time the defacto informa-
tion system and was included in the both releases. The MDS information model describes the physical and logical components of a compute resource as a set of hierarchical of elements. There are four fundamental elements: MdsHost, MdsDevice, MdsService and MdsSoftware. The most important of these elements is MDSHost which represents a networked computer and its various service access points e.g. a workstation, SMP parallel system, or the frontend server for a distributed parallel machine. MdsDevice, MdsService and MdsSoftware are used to represent local user-visible physical or logical devices, network-accessible services or an instance of installed software respectively. A set of specific attribute types [3] can be used to add information about the particular instances of hosts, devices, services and software. Grouping objects [3] can be used as collection points for the specific instances and give summary information where applicable.

The EU DataGrid (EDG) project [19], which provided the EDG release, found that the MDS schema did not describe Grid entities in sufficient detail for use with higher level middleware services. Within the project structure, middleware development was divided into six functional areas: workload management, data management, Grid monitoring and information systems, fabric management, mass data storage and network monitoring [19]. As such each area defined its own subschema which described the entities required to enable the higher level middleware functions.

For workload management, a description of the computing resource was required. This was realized as a single entity, ComputingElement, which described the Globus GRAM [15] endpoint, the batch system, the state of the batch system and a simple description of the computing resource which consisted of a homogeneous cluster. Another entity defined was the CloseStorageElement which described the mount point of the local storage resource. The storage resource was defined with the StorageElement entity which described the endpoint and the protocols supported by the storage resource were defined by the StorageElementProtocol entity.

As the EDG project had defined their own information model, the information providers were not available in the VDT release, which relied upon the standard Globus distribution. Since these were essential in order for the Resource Broker [19] to discover a resource and take it into account during the brokering process, the information providers needed to be installed on all the US machines. This approach to interoperability is an example of the parallel deployment approach [13] as there is parallel deployment of the information providers on the resource, one for each project.

3 The GLUE Information Model

The Grid Laboratory Uniform Environment (GLUE) 1.x Schema [8] resulted from a collaboration effort between the DataTAG and US-iVDGL [18], Globus [15], and the EDG [19] projects. The aim was to define an information model and LDAP rendering for a uniform representation of Grid resources. After the definition of the GLUE 1.x schema in 2002, the iVDGL and DataTAG collaborations developed information providers to conform to the schema and they were installed on both the European and the U.S. resources.

Information about GLUE entities is encoded in terms of named objects comprising attribute-value pairs that describe properties of the supported entities. At the top of this hierarchical structure is the ComputingElement which represents the entry point to a queue in a batch system. This entity has a number of associated entities which have a one to one relationship with the ComputingElement. These associated entities are used to describe the State, Policy and other information about the ComputingElement including authorisation which is described using the AccessControlBase attribute. Below the Computing Element in the hierarchy are the Cluster and SubCluster entities. The SubCluster entity is used to describe the hardware and software properties of a homogeneous group of machines by use of the respective associated classes and the Cluster entity is used as an aggregation entity.

The storage model was based around the concept of a StorageElement which provides access to Storage Libraries where files can be stored and retrieved. The Storage Element has associated entities representing the state and access policy for the service. The Storage Library describes the physical properties of the storage system as well as the file system used. A Storage Area represents the logical space which is associated to policy, state and access control entities. The storage spaces are aggregated to the Storage Element entity. The relationship between the Computing Element and the Storage Element is described by the Computing Element Storage Element Bind Schema. This describes the mount point used in the cluster for the Storage Element and is associated to both the Computing Element and Storage Element.

4 The Interoperability Activity between EGEE and Nordugrid

The Nordugrid project [10] started in May 2001 and aimed to build a Nordic testbed for wide-area computing and data handling. As the focus was on deployment it was hoped that the Globus Toolkit would provide a box of tools from which a solution could be developed. The building blocks and libraries offered by the Globus toolkit was then used to develop an innovative Grid middleware known as
the Advanced Resource Connector ARC [11]. The ARC middleware provides an LDAP-based distributed information system which highly leverages components from MDS. One of the main differences with the deployment of MDS in NorduGrid is the information schema. The Nordugrid collaboration also found that the MDS model did not describe Grid entities in sufficient detail so the NorduGrid information model was developed and put in production in 2001. The main entities described in this model are computing resource, Grid jobs, Grid users and storage resources.

The computing resource is described using two primary entities, the nordugrid-cluster and nordugrid-queue. The nordugrid-cluster entity describes the hardware, software and middleware properties of a cluster in addition to its location and ownership. The batch system used to manage the resource as well as dynamic state information, such as number of queued/total jobs, is also part of this description. The nordugrid-queue entity represents either a traditional batch queue of a local resource management system (LRMS) or an entire LRMS when the LRMS does not support conventional queues. Besides the usual queue-specific information queue-level node attributes are also introduced to describe the hardware/software characteristics of computing nodes assigned to a certain queue. The concept of a computing queue plays a central role in Nordugrid as they are the job submission targets for the brokering process.

Within the information model every Grid user authorized to use a resource is described by a nordugrid-authuser entry. These entries are used to present the current state of the computing resource from the specific view point of the user such as the number of running and waiting jobs. A Grid job entity is created for each job which is submitted to the computing resource. This entity describes various metadata and monitoring information about the job. Storage resources are described using the nordugrid-se entity. A storage resource consists of a physical data source (the storage space itself) plus the protocols, policies, services and interfaces which make the storage space available to the clients.

The Enabling Grids for E-ScienceE (EGEE) project [16] aimed to build upon the experience generated by previous EU projects in Grid technology and develop a service Grid infrastructure which is available to scientists 24 hours-a-day. Both infrastructures provide resources to the LHC Computing Grid [17] (LCG), which aims to build and maintain a data storage and analysis infrastructure for the entire high energy physics community that will use the Large Hadron Collider (LHC). Due to the demands of the LHC Virtual Organizations, it has become necessary to ensure that these Grids interoperate so that the VOs can seamlessly use both Grid infrastructures as one resource.

As the middleware distributions for EGEE and Nordugrid were different, it was first necessary to understand the main components of both infrastructures and compare interfaces. In order to help in this task an interoperability matrix [13] was created to highlight differences between the infrastructures. This showed that although both provided an LDAP-based distributed information system, EGEE used the GLUE 1.x schema where as Nordugrid used the Nordugrid schema. To overcome these differences a translation approach [13] was chosen. This required translating the information from the GLUE information model to the Nordugrid Information model and vice-versa. Once this translation was understood, translating information providers were developed to implement this translation. These would query the corresponding information system components in the other infrastructure and translate the information into the native model so that it can be added to the native information system. This method was deployed and has been successfully used in the production environment for LCG.

5 The Interoperability Activity between EGEE and OSG

The Open Science Grid [12] is a US Grid computing infrastructure that supports scientific computing via an open collaboration of science researchers, software developers and resource providers. The OSG Consortium builds and operates the OSG infrastructure, bringing together resources and researchers from universities and national laboratories. The capabilities and schedule of development are driven by U.S. participants that will use the LHC. Due to the demands of this user community, it has become necessary to ensure that OSG can interoperate with EGEE so that the VOs can seamlessly use both Grid infrastructures as one resource.

EGEE and OSG produce software releases for their respective infrastructures that are based on components from VDT. However, VDT contains many components and only a subset are deployed on both infrastructures. In addition EGEE substantially augments VDT with many extra components and services. Even if the same components are deployed, variations in versions, configuration and deployment models can occur.

To overcome these differences, it was again necessary to create an interoperability matrix to highlight differences between the infrastructures. This exercise revealed that there was a great deal of commonality between the infrastructures which was not surprising as they both leveraged components from VDT. With respect to the information system, both infrastructures used the same interface, LDAP, however there were significant differences with the schemata.

EGEE was using GLUE with a number of extensions and OSG was using MDS, GLUE and their own Grid3 schema. The original aim of GLUE was to create a common information model to facilitate interoperability between EU and US Grid projects. As such OSG considered moving to the
GLUE 1.x schema on condition that a minor revision was undertaken to address their additional use cases. A proposal for version 1.2 of the GLUE schema was discussed which, in addition, addressed a number of minor problems with the original version.

Soon after version 1.2 of the schema definition was completed, information providers were created to produce information in accordance with the new schema. The new information providers were included in the respective software releases and rolled out across both infrastructures. This work successfully achieved interoperability between the two infrastructures and demonstrated how to achieve interoperability by moving to agreed common specifications [13].

6 Grid Interoperability Now

Over recent years a number of Grid projects, many of which have a strong regional presence, have emerged to help coordinate the different actors and enable Grid infrastructures. Today, we face a situation where a number of Grid infrastructures exist, each of which is built using different middleware. Due to the regional character of the different projects, this resulted in the existence of largely isolated Grid islands. Grid interoperation is trying to bridge these differences and enable Virtual Organizations to access resources independently of their Grid project affiliation.

Without this VOs are artificially limited by the boundaries created by the different Grid infrastructures or have to separate their entire work-flow by region. Interoperation is usually achieved through bilateral and multilateral interoperability initiatives.

Following the visible successes of a few bilateral interoperation activities, an ad-hoc meeting at Super Computing 2005 in Seattle between representatives of different Grid infrastructures lead to the start-up of the Grid Interoperability Now! [1] (GIN) activity. The GIN Community Group from the Open Grid Forum [4] is trying to build upon these bilateral activities with the aim of building bridges between the islands to create a uniform Grid landscape. The GIN group is a focal point where the infrastructures can come together to share ideas and experiences on Grid interoperation. The first goal was to find short term working solutions which could be demonstrated at Super Computing 2006. The purpose of these initiatives is not only to provide a proof-of-concept for interoperability and potential standards for the future, but also to provide a production-ready solution for the present.

The GIN activities focused on the four corner stones of interoperability: Security, Workload Management, Data Management and Information Systems. For information system interoperability two approaches were discussed, either agree on a common interface and information model or create translators to transfer information from one Grid island to another. As the deadline for a working solution was very short, it would have been impossible to agree on a common solution and migrate all the infrastructures to this solution so the translation approach was attempted.

An interoperability matrix [13] was created to understand the complexity of the problem. This showed the participating infrastructures, the connection protocol, the schema and the data-model. An attempt was made to translate information from all Grids into a common form. As the GLUE schema was designed for interoperability, it was chosen as the information model and LDIF was used as the data model as it was the most widely used. The task to create a translator can be divided into two separate tasks, setting up a communication channel and translating the communication. The first task is achieved by using the native information system client of the Grid to perform the actual query. The second task is achieved by translating the query response (note that if both the schema and the representation are different then this requires two steps).

A simple architecture was created, leveraging the Generic Information Provider (GIP) [14], to support the use of pluggable information translators for the different Grids. This was used to populate a BDII [14] with all resulting translated information which came known as the GIN BDII. The information could then be accessed using the standard LDAP protocols.

The attempt to map the schema was non-trivial and a number of major obstacles were faced. Firstly, the mapping is not one to one. Different naming and semantics of attributes resulted in the need for heuristics to populate certain attributes in the GLUE schema. This lead to many assumptions which may only be accurate through coincidence rather than design. The second problem is that some attributes just do not exist and it is impossible to translate these values which results in missing information. This could sometimes be worked around in the prototype by either hard-coded values in the translator or using default values but in many cases this was not possible. Poorly translated or missing attributes could be problematic for applications which rely on those attributes for successful operation.

An example is the GlueSiteLocation attribute (Format: City, Country). This did not exist in the Nordugrid schema but it does offer a postal code which was used to identify the City and Country. Even if the GLUE 1.x schema was used, these problems are not entirely avoided as different projects may interpret the schema slightly differently and additionally decide not to use some of the optional attributes which may be important for certain use cases.

The information in the GIN BDII was used to create the Google Earth demo for Super Computing 2006. This demo used Google Earth to show the location of all the sites from the Grid infrastructures which were participat-
ing in the GIN activity and to which Grid they belonged. This was achieved by periodically querying the GIN BDII for the GlueSite entry and using this information to generate a KML [2] file for Google Earth. This demonstration was repeated at Super Computing 2007 and showed how the Grid had evolved during the year. It also showed that a few infrastructures have adopted the GLUE schema based on the success of the 2006 demonstration.

7 A Common Grid Information Model

7.1 Motivation

The GIN activity highlighted the requirement for a common schema. While it is possible to do limited translations, if the fundamental models differ or the information required for a particular use case is not present, it is impossible to translate this information. If interoperability is to be achieved there needs to agree on the fundamental model of Grid computing and information required for the cross-grid use cases needs to be identified.

A survey of existing middleware solutions showed the dominant information model was the GLUE schema. Many existing models were too low-level and could not be used to describe abstract Grid services. The Nordugrid model by contrast had a very mature description of a Computing Element but lacked a detailed Storage Schema.

With the agreement of all the main players, the GLUE Schema Working Group was created in the OGF and the GLUE schema activity was moved there to leverage the existing proposals into a community-agreed standard. A plan to conduct a major revision to GLUE was outlined, which would consolidate the Nordugrid Computing schema with the GLUE schema. By enabling rendering for schema in the LDAP, Relational, XML and CIM [?] data models most existing used cases could be met. In addition, the activity benefited from several years of experience in the context of production Grid infrastructures.

7.2 GLUE 2.0

The modeling activity concentrated on capturing the abstract representation of entities, their properties, operations and relationships in an implementation independent information model. Before in-depth work could be undertaken on information modeling a number of fundamental Grid concepts needed to be agreed upon. The GLUE Schema Working Group and the Reference Model Working Group [5] had a number of joint discussions to find this common view. The result of the discussion was captured in the Glue Main Entities model.

At the core of this model is the concept of a Service and as such this demonstrates that a Grid has a Service Oriented Architecture. A Service enables a User from a User Domain to run an Activity via an Endpoint on a Share of a Resource in accordance to an Access Policy. As such, a Service is container which is used to describe a collocation of sub components that are required to fulfill a particular function. Services are related to an Admin Domain which provides and manages the Service. This concept is important for infrastructures as there is a Service Level Agreement between the User Domain and the Admin Domain which places requirements on the quality of service provision. An Activity is a unit of work managed by a service and can have relationships to other activities being managed by different Services.

For many use cases it is necessary to define a more detailed information model about specific Services. The two main Services that require more detailed information are the Computing Service and the Storage Service. The definition of these two services would enable the main entities model to be evaluated.

The Computing Service is a specialization of a Service for creating, monitoring and controlling computational activities more commonly known as jobs. The following entities were identified and required specialization: computing manager, computing resource, computing share, execution environment, application environment and job. A Job is an Activity that runs on a Computing Resource. The Computing Resource is a grouping concept for a set of different types of Execution Environments where the aggregation is defined by the common management scope e.g. a LRMS. The Execution Environment provides a description of the hardware and software characteristics available to a Job. An Execution Environment may also contain one or more Application Environments. The Computing Share is a utilization target for a set of Computing Resources defined by Mapping and Access Policies and characterized by the Status of the Computing Share.

The Storage Service is a specialization of a service for storing and retrieving file based data. The following entities were identified and required specialization: storage manager, storage resource, storage share, storage service capacity and storage share capacity. The Storage Resource is a sufficiently homogeneous storage device providing a storage capacity, managed by a local Storage Manager. The Storage Share is a utilization target for a set of Storage Resources defined by Mapping and Access Policies and characterized by the Status of the Storage Share. The Storage Service Capacity is a description of the size and usage of a homogeneous storage extent; the storage extent is aggregated at the storage service level by the type of Storage Resource.

The information model can be rendered into one or many different data models which represent the information model in a given language. It is the data model which enables the information model to be transmitted, manipulated
leads to the second point, to obtain seamless interoperability between information systems, there must be agreement on the information model. Finally, as information schema are defined in response to the use cases they have to meet, there must be agreement on the cross-grid use cases before a common information model can be defined.

The investigation in the GIN info group demonstrated the need for a common schema for interoperation, and with the agreement of all the main players the GLUE Schema Working Group was created in OGF to leverage the existing proposals into a community-agreed standard. A major revision of GLUE was undertaken, GLUE 2.0, which built upon the experience gained while used the Nordugrid, GLUE 1.x and other schemata in production environments. The activity benefited from several years of experience in the context of production Grid systems and numerous interactions with other OGF working groups. The result was an abstract information model describing the main entities in a Grid infrastructure as well as specializations for the computing and storage services.

The abstract information model is in essence the definition of a Grid and the definition of GLUE 2.0 represents this view based on the experience of running production Grid infrastructures as well as a common consensus between Grid practitioners. GLUE 2.0 brings a number of significant improvements over GLUE 1.x, based on real deployment scenarios which in addition to the elegant abstract main entities will hopefully encourage a broad and timely adoption.

References


506


