Abstract

This survey report is focused on providing a brief description about the architecture of Globus infrastructure’s version of Grid Information Service called Monitoring and Discovery Services (MDS) and its underlying protocol, LDAP.

1. Introduction

Grid computing is a form of distributed computing that involves coordinating and controlled sharing of diverse computing, applications, data, storage, or network resources across dynamic and geographically dispersed multi-institutional virtual organizations [2]. The Globus Toolkit is an open source technology that provides core services for grid enabled applications and tools. The toolkit includes software services and libraries for resource monitoring, discovery, and management, plus security and file management.

The Globus provides four types of services.

<table>
<thead>
<tr>
<th>Services</th>
<th>Enabling Technologies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Security Services</td>
<td>Grid Security Infrastructure</td>
</tr>
<tr>
<td>Data Services</td>
<td>Grid FTP and RFT (Reliable File Transfer)</td>
</tr>
<tr>
<td>Execution Management Services</td>
<td>GRAM(Grid Resource Allocation and Management)</td>
</tr>
<tr>
<td>Information Services</td>
<td>MDS (Monitoring and Discovery System)</td>
</tr>
</tbody>
</table>

The goal of this paper is limited to discussion about the Information services of Globus. The following sections are arranged as follows: MDS overview and architecture, MDS configuration, LDAP, and
security in MDS in section 2, related work in section 3 and conclusion in the section 4 followed by references in the last section.

2.1. MDS Overview and architecture

Information services for a Grid computing infrastructure realized through systems like GLOBUS are essential for providing configuration and adaptation capabilities for heterogeneous resources. In GLOBUS toolkit this is achieved through the information service infrastructure, Monitoring and Discovery Service (MDS). The system not only aids discovery and utilization of the Virtual Organizations that are otherwise unknown to the user/client but also serves many purposes such as super scheduling, troubleshooting, performance diagnosis and application adaptation. Any GIS has the following goals to be met: scalability, performance, access to static and dynamic information, access to multiple information sources, robustness, decentralized maintenance, uniformity, flexible access to information, security and access control.

MDS provides two kinds of services, namely Grid Resource Information Service (GRIS) and Grid Index Information Service (GIIS). GRIS is associated with each resource and answers queries from client for their current configuration, capabilities, and status. GIIS on the other hand is a directory service that pulls information for GRIS’s. It is a “caching” service which provides indexing and searching functions. Thus the MDS is a hierarchical infrastructure with distributed services allowing for scalability and fault tolerance.

The GRIS uses two protocols:
1. Enquiry protocol: Grid Resource Inquiry Protocol (GRIP)
2. Registration protocol: Grid Resource Registration Protocol (GRRP)

These two protocols facilitate interoperability between GIIS and GRIS servers. The GRIP protocol is the core of the resource layer function of GIS; It must supply a query function, must be deployed universally and should support distributed operations. GRIP supports both discovery and enquiry. The GRRP protocol notifies one directory server about the existence of another; it is a Soft-state protocol (relationship by means of notifications, discarded after a long period of silence) and may be implemented on a variety of transport protocol (e.g. LDAP implementation coupled with SASL security).

---

Fig.2. MDS service architecture
2.2. LDAP:

LDAP (Lightweight Directory Access Protocol) is used as the protocol for GRIP. LDAP is a vendor and platform-independent, open, network protocol standard for accessing information from the directories. Each entry has a distinguished name for unique identification resource and associated attributes and their values.

Data in LDAP based directory is based on

- **base DN:** indicates where in the hierarchy to begin the search.
- **filter:** specifies attribute types, assertion values, and matching criteria.
- **scope:** indicates how many levels of the directory tree are searched, relative to the base DN.

LDAP is used by the MDS in the following manner.

1) **Initialization phase**
   - Provides a naming schema for altering properties of a computer and its network and contains tools that help defining and modifying information description.
2) **Population phase**
   - A framework to populate the directory with information. A UNIX script is available to populate the object with information. This modular approach fosters portability between platforms.
3) **Querying phase**
   - A framework to request information from the directory service. LDAP makes it possible to use a variety of application interfaces to communicate with the server. This feature makes querying and updating easy.

![Diagram of LDAP data model](image)

**Fig.3.** The LDAP data model represents information as a set of objects organized in a hierarchical namespace. Each object is tagged with one or more named types. Each object contains bindings of values to named attributes according to the object type(s).

2.3. MDS configuration

MDS can be configured manually or by automated discovery based on hierarchical discovery or based on other services such as SLP[4], DNS [5] or UDDI[6].

Overall procedure is as follows:

1) Design MDS hierarchy and define necessary GIS services,
2) Determine the Distinguished Name(DN) for GIS services.
   - e.g., dc=kisti, dc=kr, o=Grid
3) Install, setup and deploy Globus toolkit on each server
e.g. globus-install, globus-setup, globus-local-deploy
4) Modify a few configuration files manually
   E.g. Configuration file for daemon processes {deploy_dir}/etc/globus-daemons.conf
5) Test and Running
e.g., MDS tools or LDAP browser

2.4 Security
Security is a main concern especially in distributed systems. Consumers and GIS use single sign-on and public key security mechanism to verify credentials and mutually authenticate each other. Access control policies based on GIS name and LDAP access control mechanisms provide authorization.

3. Related Work

Ganglia[8] is also scalable distributed monitoring system for high-performance computing systems such as clusters and Grids. It possesses simple design and engineering to achieve high levels of robustness, ease of management and portability. Hawkeye [9] is another monitoring tool built on Condor technology that provides rich mechanisms for collecting, storing, and using information about computers.

4. Conclusion

Globus toolkits are widely adopted by scientists and researchers all across the globe. The new MDS4 provides new features to query WSRF services for resource property information, execute a program to acquire data and interface with third-party monitoring systems. MDS4 was specifically designed to address the needs of a Grid monitoring system – one that publishes data that is available to multiple people at multiple sites.

5. References


