BEN CLIFFORD

MDS: The Globus Monitoring and Discovery System

MDS is the Monitoring and Discovery System (*www.globus.org/mds*), a component in the Globus Toolkit™ (GT). As the name implies, MDS targets the two problems of monitoring and discovery.

Monitoring is the process of observing resources (compute systems, for example) on the Grid to track their status for purposes such as fixing problems and tracking usage. For example, MDS users might want to see which resources are running low on disk space in order to take corrective action.

Discovery is the process of finding a suitable resource on the Grid to perform your task, such as finding a compute host on which to run a job. This task involves both finding resources that are suitable (e.g. have the correct CPU architecture) and then choosing the best of those (e.g. the one with shortest submission queue).

Although these two activities are distinct, they share a common need for information to be collected from many resources, distributed around a virtual organization (VO), and made available in one place.

MDS provides the infrastructure to collect information from resources on the Grid, tools to allow users to make use of that information, and a number of mechanisms for extensibility. MDS users include administrators of virtual organizations who wish to track the resources in their VO and application users who wish to make use of the Grid to run their jobs.

The Infrastructure

The MDS *Index Service* is the main component concerned with collecting MDS information from resources. Grid resources register with one or more indexes; those indexes then collect information about each resource and make it available to end users. Users need only know the location of a suitable index service in order to discover and monitor the resources that it indexes.

Each index service provides a view of the resources that are likely to be relevant for the users in its VO.

MDS has a number of features that are sometimes surprising to new users, but are necessary due to Grid scalability and policy issues:

Indexes can be arranged in hierarchies, but there is no single global index that provides information about every resource on the Grid. This omission is deliberate, as each virtual organization will have different policies on who can access its resources. No person in the world is part of every virtual organization!

The presence of a resource in an index makes no guarantee about the availability of the resource for users of that index. An index provides an indication that certain resources are likely to be useful, but the ultimate decision about whether the resources can be used is left to direct negotiation between user and resource. A user who has decided on a particular resource to use based on MDS information might still find they are not authorized when they submit a job. This situation may result because the MDS does not keep track of policy information — something that is hard to do concisely. In addition, MDS does not require that resources reveal their policies publicly.

MDS has a soft consistency model. Published information is recent, but not guaranteed to be the absolute latest. This design allows the load caused by information updates to be reduced at the expense of having slightly older information. The delay is generally not a problem For instance, it is generally acceptable to know the amount of free disk space on a system five minutes ago (rather than two seconds ago).

Each registration into an index is subject to soft-state lifetime management. Registrations have expiry times and must be periodically renewed to indicate the continued existence of the resource. This requirement allows each index to be self-cleaning, with outdated entries disappearing automatically when they cease to renew their registrations.

Information

In the default installation, the most useful information collected and made available by MDS is from the Globus Resource Allocation Manager (GRAM, which deals with compute job submissions) and the Globus Reliable File Transfer Service (RFT).

For GRAM, the collected information includes details of each queue to which jobs can be submitted, including name of the queue and number of jobs in the queue. The information also includes information about the underlying cluster or host which will be used to execute the compute job, with details such as CPU type and disk space. This information is presented in the GLUE schema, a collaborative effort to develop a common model for cluster information.

GRAM provides this information to MDS by interfacing with the underlying batch scheduler on a compute resource (such as PBS or LSF) and from cluster monitoring systems such as Hawkeye and Ganglia.

For RFT, a number of simple numerical statistics are made available including the total number of files transferred, total number of

bytes transferred and the number of transfers under way at present.

Getting Information to the User

MDS information can be delivered to a user in a number of ways:

Users can pull information from MDS using either a command-line tool or a web interface; information can be pushed to users by e-mail; or specialized tools such as brokers can collect MDS information and use it as part of a larger process.

From a user's perspective, the easiest interface to MDS is provided by *WebMDS*. This tool presents MDS information in a standard web browser. The interface displays overview information for an entire VO, with hyperlinks giving the ability to drill down and view more detailed information about each monitored resource. Users can see the status of each available resource and pick out the resource they want to use based on the information presented through the web-based interface.

WebMDS is included as a technology preview in the latest GT release.

For command-line users, the Globus Toolkit includes a simple *wsrf-get-property* command that will query MDS information from an index service and output in raw XML format. The same tool can also be used to gather information from WSRF-compliant (Web Services Resource Framework, *www.globus.org/wsrf*) Grid resources directly.

The MDS *Trigger Service* provides a very different approach to user interaction. Instead of a user requesting information when they want to know something (for example, by loading a WebMDS page in their web browser), the Trigger Service will notify a user when something of interest happens.

A user registers one or more trigger rules, consisting of a match expression and an action script.

The match expression is applied to information on Grid resources every time that information is updated. Match expressions are written in XPath, a simple XML querying language. If the information matches, then the action script fires to inform the user by sending e-mail or a page.

Extending MDS

Some users might wish to code their own clients, or extend MDS in other ways. For example, a client might use MDS information to choose the most suitable place to submit a job and then actually submit the job to the discovered resource using the GRAM component of the Globus Toolkit.

The programming model is straightforward. MDS information is made available through the Web Services Resource Framework (WSRF), for which client implementations exist in a number of languages (including C and Java). The underlying data model for MDS information is XML, which allows great flexibility in the kinds of data that can be represented, from simple byte counts (as provided by RFT) to detailed per-node cluster information (as provided by GRAM).

MDS can also be extended to collect information about different kinds of resources on the Grid. Out of the box, it can collect information from any resource that adheres to the WSRF architecture, but information providers can be written (using either executable or Java APIs) to collect information about any other kind of resource. Each information provider can provide arbitrary XML about the resource it is monitoring, which will be published through the Index Service.

A third extensibility point is for Trigger Service action scripts. The trigger service can call arbitrary action scripts, allowing new delivery mechanisms to be implemented to notify users.

WebMDS allows formatting of information to be customized using XSLT style sheets. XSLT (*www.w3.org/TR/xslt*) is a stylesheet language designed to convert XML into HTML. It can be used in WebMDS to present existing MDS information through the web interface in new ways, or can be used to integrate new kinds of MDS information (for example, as collected by new information providers) into the existing displays. Stylesheets are fed MDS information as XML input and are expected to output appropriately formatted HTML.

Future Development

Ongoing MDS development is active in a number of areas. There is always a need for additional information to be published, both about resources which are already monitored and about new kinds of resources. In addition to implementing information providers to collect the information, it is also sometimes necessary to define new schemas to represent that information. We are pursuing mechanisms for collecting information directly from resources and also by interfacing with other Grid monitoring systems.

Prototype work is under way for archiving historical MDS data, with the intention of providing historical records of the same information that is available live through an index.

We are also investigating ways to enhance scalability.

Globus Toolkit is a registered trademark held by the University of Chicago. This work was supported in part by the Mathematical, Information, and Computational Sciences Division subprogram of the Office of Advanced Scientific Computing Research, Office of Science, U.S. Department of Energy, under Contract W-31-109-ENG-38 and under Contract DE-AC03-76SF0098 with the University of California; by the National Science Foundation; by the NASA Information Power Grid program; and by IBM.

Ben Clifford (benc@isi.edu) is a programmer at USC Information Sciences Institute, in Marina Del Rey, California.