

Data grids for storage

http://storageMagazine.techtarget.com/magItem/0,291266,sid35_gci1132545,00.html

by: Ray Lucchesi

Storage Magazine Issue: Oct 2005

For large geographically dispersed companies, data grids offer an ingenious new model to economically share computing power and storage resources

Current grid computing projects manned mostly by scientific teams offer some tantalizing prospects for general corporate computing. Imagine making your organization's data accessible throughout the world or replicating data to multiple, geographically dispersed sites--even sites you don't own or control, but with which you collaborate.

If you use traditional access-control methods, the barriers to this scenario are substantial. You could, for example, set up replicated Web FTP mirror sites with user logins and passwords to all of the sites providing access, or set up VPN access to each site holding the data.

But it isn't easy to replicate data to alternate sites with an FTP site, and user IDs/passwords become a major hassle with multiple sites. VPNs require different passwords and configurations for each data repository site, and users would certainly balk at having to navigate 10 or 100 VPN connections to get one piece of data. Another--and better--solution is to use a data grid.

Data grids

With so many storage vendors touting some sort of grid architecture these days, an accurate definition of a grid may be elusive. For the

**WITH SO MANY STORAGE VENDORS
TOUTING SOME SORT OF GRID
ARCHITECTURE THESE DAYS,
AN ACCURATE
DEFINITION OF A GRID
MAY BE ELUSIVE**

purposes of this article, a grid spans sites, companies and continents with non-proprietary hardware, software and protocols supporting authenticated access, replication and compute services. Clustered file systems don't qualify as data grids because they typically exist at one or two sites and require high bandwidth connections between nodes. Wide-area file systems come closer to a data grid model, but they don't currently offer continent spanning or multicompartment hosted data; they also require proprietary hardware, software and internode protocols.

It's possible to use a grid to securely share your data and compute services. To tap into these capabilities, you need to implement standard, compliant grid services on your systems. These services are available from the open-source community; proprietary grid products are also available from some vendors, including IBM Corp., Oracle Corp., Silicon Graphics Inc. (SGI)/YottaYotta Inc. and Sun Microsystems Inc.

Data grids are perfect for organizations that need a collaborative work environment despite having diverse, distributed resources where data resides across multiple business and/or organizational domains. Data grid services allow users to access and manipulate data residing at sites around the world. Data can be retrieved from any location on the grid, and can be deposited or replicated to any location with space.

Compute grids

A compute grid can schedule computation to occur at one site with the results

transmitted to another (see "Open-source grids," above), and a compute grid may exist with or without a data grid. Together, a compute grid and data grid can interoperate to move data residing throughout the grid to where

System Grid (ESG), administrators limit the amount of data and/or files users can download to effectively govern overall grid activity.

One popular open-source compute grid is Condor-G, which was developed at the University of Wisconsin-Madison. Condor-G administrators can limit the amount of shared compute resources and thereby control remote use of resources, including compute services, disk space and data.

A Directed Acyclic Graph (DAG) structure is used to schedule Condor-G grid work. A DAG can coordinate the steps in a computation and/or the data grid accesses needed to supply the processing/data being requested. Because it may take a number of steps to process the data, they can be

broken down into manageable execution steps so the process can be easily restarted if a failure occurs. Automated resource managers on the grid take the DAG and parcel out the work to grid nodes supplying the requested services.

ESG users access a Web portal to search a directory, and can specify what timeframes and data slices they need. The ESG schedules the data extract and sends it to the requestor; it also provides a list of locations where file replicas may be found, allowing users to choose the one they'd like to download.

Storage on data grids

Storage on data grids can be managed in many ways. One popular approach uses storage resource managers (SRMs) to manage files, disks and archives. An SRM can enforce space quotas and other storage constraints. Files can be pinned (reserved) by an SRM. While pinned, a file can't be removed from an SRM's control, but it can be accessed by multiple users. File pins may be released (unpinned) or left to timeout, after

OPEN-SOURCE GRIDS

Globus is the progenitor of many of today's grids. The Globus Alliance and the Global Grid Forum (GGF) support the Globus Toolkit, and have developed some of the fundamental services required to implement a grid. The GGF is also charged with popularizing the grid by making it easier for all users to participate in grid work.

There are essentially three different modes of Globus support software: the API-based model in Globus Toolkit version 2.0 (GT2), the service model in GT3 and the Web services resource framework in Globus Toolkit version 4.0 (GT4), released last May.

There are many compute-data grids in operation around the world, including AstroGrid, the Biomedical Informatics Research Network, the Enabling Grids for E-science, Grid Physics Network and the Particle Physics Data Grid.

computation can occur and send results wherever required.

For example, animators can publish images on a grid and provide access to other artists to supply the background, foreground and other elements. Further processing can be done on any grid-enabled system with available cycles. Results can be transmitted back to the original location or sent elsewhere for further processing. Computations can be handed from one system to another to take advantage of each node's capabilities.

Security concerns

Security is a major concern for any grid. Only authorized users can access a grid, and data grid transmissions can be encrypted. Strong encryption authentication based on public key infrastructure (PKI) is used. One advantage of grid services is that security is built in from the start--user IDs/ passwords aren't required for every site entered and secure authentication is maintained. Grid administrators can also set up access constraints. For example, on the Earth

which the file is no longer guaranteed to be

**STORAGE ON DATA GRIDS
CAN BE MANAGED
IN MANY WAYS.
ONE POPULAR APPROACH
USES STORAGE RESOURCE
MANAGERS TO MANAGE
FILES, DISKS, AND ARCHIVES**

available.

SRMs also support dynamic space management. A grid user can request, for example, 200GB of space to be reserved by an SRM. As long as space is available this reservation will be held; if additional requests come in, the reservation may be reduced, assuming there aren't already pinned files in the space. The SRM works with DAS, SAN or NAS storage, anything that supports file storage; but because of the unique nature of data grids, software licensing is an issue (see "Grid software licensing issues," this page).

Local grid data and meta data can be backed up with any standard backup package; non-local data has to be retrieved at your site to be backed up locally. Replication of data and meta data across the grid can also be used for backup purposes.

Commercial grids

There are many commercial grid systems available today. Some of the better known include IBM's grid offerings, Sun's N1 Grid Engine, Oracle's 10g and a new offering from a partnership between SGI and YottaYotta.

One of the advantages of commercial packages is that you don't have to rely on the open-source community for support and vendors can provide installation services. Also, most vendor grid offerings are more tightly integrated with other commercial products. Some vendor products may also provide better scalability, reliability and serviceability than open-source versions.

IBM's DataSynapse is fully interoperable with current grid standards. By following Global Grid Forum (GGF) standards, you could plug your IBM compute grid into any grid around the world that supports these standards. IBM claims its grid allows for better throughput and more parallel computation than what's available from open-source grid products.

Sun's N1 Grid Engine (SGE) also provides a standards-compliant compute grid service. GGF-compliant Globus Toolkit version 4.0 (GT4) services can be used to submit jobs to an SGE grid.

Although proprietary and specific to databases, Oracle 10g is probably closest to a data grid among commercial products. 10g supports transportable tablespaces that can be used to move tablespaces among remote databases. Moving tablespaces around could improve performance for remote sites accessing the data. By moving a tablespace, you can free up local database space; tablespaces can also be mounted as read-only to a number of databases. 10g supports "federated databases" that use distributed SQL and gateways to other databases to provide clients with a single, unified view of multiple databases.

GRID SOFTWARE LICENSING ISSUES

Most standard computer-resource accounting models break down when applications run across a distributed grid, so another issue facing grid use is software licensing. It's hard enough ensuring proper software licensing for all of the computers in one data center, but try doing this for 14,000 CPUs across 135 data centers. Grid use has evolved based mainly on open-source applications with generous licensing terms.

To join an open grid service, you need to be pretty computer-savvy and patient. It took me the better part of a day to log on to the Globus data grid. My test installation was limited to data grid services (GridFTP, Reliable File Transfer [RFT] and Replica Location Service [RLS]) even though Globus Toolkit version 4.0 (GT4) supports computational services.

Software components required to install a Globus data grid include Java SDK, Apache Ant, a C compiler, a database, GNU make and GNU tar. A JDBC-compliant database may be needed for some grid Web services. Tomcat can be used as a Web server or the Globus Toolkit has a standalone Web service container that can be used.

Security requirements for data and compute grids are complex for obvious reasons. Globus security infrastructure depends upon public key infrastructure, host and user security certificates, and a certificate authority (CA) to validate them. The best option is to use a currently supported CA where available. If none exists, SimpleCA from the Globus Toolkit can be used for test purposes. Although installing the security infrastructure was complex and time-consuming, the advantages were immediate. Activities that normally took additional logins were authenticated and approved automatically by grid services using security proxies.

GLOBUS ON RED HAT 9

With all prerequisites in place, the build of the source code took over three hours. Setting up certificates for two machines and two users, fixing permissions and other middleware such as proxies and grid-mapfiles took the better part of a day. Afterwards, GridFTP was used to FTP a file from one machine to another. It wasn't until after this completed that I noticed no login was required--GridFTP and grid security provided automatic authentication. RLS was the last service deployed. The key to a successful RLS implementation is to set up the environment and database linkages. RLS provides the mapping between logical file names and physical file locations. One logical file could potentially have hundreds of physical locations on the grid, and RLS can be used to catalog all of the files.

SGI/YottaYotta's offering provides another kind of data grid service using proprietary hardware, software and internode protocols. With the new SGI CXFS and YottaYotta's NetStorager appliance, you can have multiple sites that each have a clustered file system (SGI CXFS) to share and replicate data throughout the WAN at SAN speeds with replication directories maintained automatically. SGI and YottaYotta say they've demonstrated a CXFS cluster reading and writing to a shared file across 2,900 miles at approximately 700MB/sec.

Implementation steps

As a first step toward using data grids, corporate IT departments might

consider implementing GridFTP with security services (see). Any company that has a lot of FTP activity might consider using this to authenticate FTP access automatically and to provide for more efficient file data transfers. If the GridFTP implementation proves helpful, the next step may be to support a Replica Location Service database of replicated data to identify where duplicated data might be found.

For organizations running compute-intensive applications such as video rendering, seismic analysis or protein modeling, using a Condor-G compute grid would allow their local compute cluster to make use of other collaborating sites throughout the world.

GRID WEB SITES:
 Enabling Grids for E-science (EGEE) www.eu-egee.org/
 Earth System Grid (ESG) www.earthsystemgrid.org/
 Globus Alliance www.globus.org/
 Global Grid Forum (GGF) www.gridforum.org/

Ray Lucchesi is president of Silverton Consulting and since 2004 has helped startups to Fortune 500 companies improve storage product development and marketing. Ray has also helped end-users better select and use storage subsystems.

info@SilvertonConsulting.com
<http://www.SilvertonConsulting.com>

Storage magazine is a part of the TechTarget portfolio of enterprise IT-focused media. Copyright 2001-2005, TechTarget. All Rights Reserved. Storage
TechTarget

