



THE GRID

BRINGING **COMPUTING** POWER TO THE MASSES

Grid-based “computing as utility” is poised to take off. Here’s how Oracle is making it possible.

You’re a financial services company. At the end of the quarter, when you’re processing your clients’ reports, you’d like to focus most of your processing power there. Four weeks later, when you’re analyzing returns from a big marketing campaign, you’d like to shift that power to your data warehouse. So how would you reallocate your processing resources? Most likely, you wouldn’t. Each database server would be outfitted for its respective peak load, with one running hot while the other runs cold. That is, unless you’ve set up a grid to dynamically reallocate your computer resources, to share them across departments in your organization.

Easy to say but hard to do—until now. Advances in clustered computing power, faster networks, and shared storage are placing the benefits of grid computing within reach of even the most cost-conscious businesses today. When deployed on Linux and powered by Oracle, IT organizations can now build dynamic industry-standard grid systems well-suited to meet the rapidly shifting needs of twenty-first-century business.

BY MICHAEL MILEY

WHAT IS GRID COMPUTING?

The concept of grid computing—of separate groups of users dynamically sharing computer resources across high-speed networks to meet changing computational needs—is not new. According to Ian Foster, head of the Distributed Systems Lab at Argonne National Laboratory, in his February 2002 *Physics Today* article “The Grid: A New Infrastructure for Twenty-First-Century Science”: “In 1965, MIT’s Fernando Corbató and the other designers of the Multics operating system [the precursor to Linux] envisioned a computer facility operating ‘like a power company or water company.’”

In the U.S., it wasn’t until the early ‘90s, with the gigabit testbeds of CASA and BLANCA (two experimental high-speed networks that linked a number of geographically dispersed laboratories), that researchers had live examples of what could be done with inter-organizational grids. By the late ‘90s, grids had taken off among scientific and research communities, and major grid application and deployment projects are now ongoing among science and engineering organizations around the world.

While some of the most far-reaching deployments of grid technology are still being executed by scientific and academic research organizations, twenty-first-century businesses and some of their top supporting vendors are now taking notice of the

grid. With the rise of the internet and the Web, as well as e-business and outsourcing, the possibilities presented by grid computing are profound. At its highest level, using a grid to meet dynamically changing business needs is all about providing ‘computing as utility’—where, as a user, you don’t want to worry about where your data resides or which computer processes your query. You just want the computing power delivered—as much as you want, whenever you want.

But that’s just how the user sees it. What about the view of the grid from those whose job is to architect and manage it? “Behind the scenes, the grid is about deploying systems that permit dynamic resource allocation, flexible information provisioning, and high availability,” says Benny Souder, vice president for Distributed Database Development, Database and Application Server Technologies at Oracle. “Dynamic resource allocation means ensuring that resources are not standing idle while user requests are going unserved but are dynamically reallocated on demand. Flexible information provisioning means making sure that the information users and applications need is available when and where they need it. And high availability means that the data and computational ability must always be there, just like electric power.”

FACTORS DRIVING GRID ADOPTION

Enterprise adoption of grid computing is being driven by the current business climate. Companies today must be able to meet changes in customer and market demand, meaning they need to build computing infrastructures that are strong, flexible, and scalable enough to respond to surges in traffic and use. They need to deploy new workloads rapidly and to quickly provide users with access to the resources they need, while maintaining security. And these resources need to work together. Finally, IT organizations need to provide this relatively cheaply, tapping into existing resources when possible, while making cost-effective new investments in computing technology that are able to scale as the business grows.

In addition to business needs, several technology trends are working together to push the grid to the forefront of IT architectures. “Every hardware vendor today is delivering low-cost computer blades that cost as much as 80 percent less than SMP, offering highly scalable, industry standard computing power. When assembled into clustered servers, outfitted with fast interconnects and attached disk farms, they provide the building blocks for what we call ‘an industry-standard grid.’ We believe this to be the most likely infrastructure to support utility computing,” says Souder. “Next, in software, there’s the tremendously rapid growth of Linux, the ‘industry standard OS,’

which already runs well on blades today. Since a blade node is typically a maximum of four CPUs, which Linux handles just fine, Linux doesn’t require the ability to scale to the 64-CPU range of SMP. Thus, it’s the most likely OS for a standard grid.”

Additionally, grid momentum is building because of vendors like Oracle, which provides grid-enabling technology. And now that the Global Grid Forum (GGF), the grid standards body, is in place and has the support of all the major vendors, including Oracle, IT organizations are beginning to grasp how the grid will maximize their computing resources. “Oracle sees the grid poised to take off like the Web, as a logical extension of what the Web provides,” explains Souder. “While the Web was about internet-enabled information presentation, the grid, among other things, is about internet-enabled information processing. We believe that all these factors will combine to make the grid the next Big Thing.”

WHAT A GRID LOOKS LIKE

The first grids have resources being shared across the network and are deployed across heterogeneous platforms ranging from supercomputers to mainframes and from SMP machines to workstations. But the version that will ultimately resonate with the business community is an industry-standard grid, running on clustered

GLOSSARY OF GRID-RELATED TECHNOLOGIES

BLADES. A computing system that includes processors and memory on a single board, but where other resources such as power, cooling, network access, and storage services are shared among a collection of blades. Blades are easily installed and removed and are smaller than rack-optimized servers.

CLUSTERING. Connecting two or more computers together in such a way that they behave like a single computer. Clustering is used for parallel processing, load balancing, and fault tolerance. Clustering is a popular strategy for implementing parallel processing, distributed computing, and grid applications because it enables companies to leverage the investment already made in computers. In addition, it’s relatively easy to add new CPUs simply by adding a new PC to the network; or, in the case of blade servers, a new blade or a server composed of blades.

FIBRE CHANNEL. A serial data transfer architecture developed by a consortium of computer

and mass storage device manufacturers, now standardized by ANSI. The most prominent Fibre Channel (FC) standard is Fibre Channel Arbitrated Loop (FC-AL). FC-AL was designed for new mass storage devices and other peripheral devices that require very high bandwidth. Using optical fiber to connect devices, FC-AL supports full-duplex data transfer rates of 100Mbps. FC-AL is compatible with, and is expected to eventually replace, SCSI for high-performance storage systems.

GIGABIT AND 10-GIGABIT ETHERNET. Ethernet is the world’s most pervasive networking technology. Gigabit Ethernet is 100 times faster than the original Ethernet (10Mbps) and 10 times faster than Fast Ethernet (100Mbps). It offers 1,000Mbps (1Gbps) raw bandwidth and is compatible with existing Ethernets. Gigabit Ethernet is already being deployed in large numbers in both corporate and public data networks. Meanwhile, an even faster 10-Gigabit Ethernet standard is nearing completion.

GRID (GRID ARCHITECTURE). A virtualization of an enterprise’s computational resources, including servers, networks, storage, and information, into one or more large pools of resources. These resources can be dynamically provisioned on demand to various enterprise applications and users, allowing enterprises to dynamically align their IT resources to their business needs.

INFINIBAND. Both an I/O architecture and a specification for the transmission of data between processors and I/O devices that has been gradually replacing the PCI bus in high-end servers and PCs. Instead of sending data in parallel, InfiniBand sends data in serial and can carry multiple channels of data at the same time in a multiplexing signal. InfiniBand channels are created by attaching host channel adapters (HCAs) and target channel adapters (TCAs) through InfiniBand switches. HCAs are I/O engines located within a server. TCAs enable remote storage and network connectivity into the InfiniBand interconnect infrastructure, called a fabric. InfiniBand architecture is

capable of supporting tens of thousands of nodes in a single subnet.

LDAP. Short for Lightweight Directory Access Protocol, LDAP is a set of protocols for accessing information directories. LDAP is based on the standards contained within the X.500 directory standard, but it is significantly simpler. And unlike X.500, LDAP supports TCP/IP, which is necessary for any type of internet access. LDAP makes it possible for almost any application running on virtually any computer platform to obtain directory information, such as e-mail addresses and public keys.

NAS. A network-attached storage device is a server that is dedicated to nothing more than file sharing. NAS does not provide any of the activities a server in a server-centric system typically provides, such as e-mail, authentication, or file management. NAS allows more hard-disk storage space to be added to a network that already utilizes servers without shutting them down for maintenance and upgrades.

NODE. In networks, a node is simply a processing location. A node can be a computer, a set of clustered blades, or some other device, such as a printer. Every node has a unique network address, sometimes called a Data Link Control (DLC) address or a Media Access Control (MAC) address.

RAID. An acronym first used in a 1988 paper by Berkeley researchers Patterson, Gibson, and Katz to describe a redundant array of independent disks for providing fault tolerance (redundancy) and improved access rates. RAID provides a method of accessing multiple individual disks as if the array were one larger disk, spreading data access out over these multiple disks, thereby reducing the risk of losing all data if one drive fails, and improving access time.

SAN. A storage area network is a high-speed subnetwork of shared storage devices. A storage device is a machine that contains nothing but a disk or disks for storing data. A

SAN’s architecture works in a way that makes all storage devices available to all servers on a LAN or WAN. As more storage devices are added to a SAN, they too will be accessible from any server in the larger network.

SMP. Symmetric multiprocessing is a computer architecture that provides fast performance by making multiple CPUs available to complete individual processes simultaneously (multiprocessing). Unlike with asymmetrical processing, any idle processor can be assigned any task, and additional CPUs can be added to improve performance and handle increased loads.

VPN. Short for virtual private network, a VPN is a private network that is constructed by using public wires to connect nodes. There are several systems that enable you to create networks using the internet. These systems use encryption and other security mechanisms to ensure that only authorized users can access the network.

server blades, and in the case of Oracle users, running Oracle Database with Real Application Clusters (RAC).

Oracle already has first-hand experience with this version of the grid. The company developed a grid on which to build and test Oracle Database 10g. Explains Bob Thome, group product manager of Grid Computing at Oracle: "We went through three phases when deploying a grid at Oracle to develop Oracle Database, which included scavenging, sharing, and pooling resources throughout our development groups. In Phase One, which began in 1999, we tried to scavenge more cycles for development by running tests on idle workstations. When the engineers would come into work in the morning, they'd see their CPUs cycling away and they'd kill the process to free their machines. In Phase Two, which began in 2000, we tried to harness disparate SMP machines by sharing them throughout the developmental groups, but couldn't overcome managerial obstacles as to who got what resource and when. In Phase Three, which began in 2001, we first tried

pooling machines in a central location to implement a 'server farm,' but found that departments were only willing to give up older devices for the pool. So that didn't work either," continues Thome. "That's when we installed server blade farms. The small granularity of the blade resources facilitates their dynamic apportionment to the various development workloads as our development needs change. Developers can harness the power of multiple blade servers to more quickly complete larger development tasks—tasks that would not have been feasible without the grid."

The optimal infrastructure for an industry-standard grid will consist of interconnected clusters of blade servers (from one to four CPUs per server), with shared storage attached, configured as network-attached storage (NAS) or a storage area network (SAN). The blade servers and the shared storage will constitute the hardware layer. High-speed interconnects between the nodes and the shared storage will constitute the

network layer, using Gigabit Ethernet or InfiniBand between the nodes and Gigabit Ethernet, Fibre Channel, or InfiniBand for the disk farms.

ORACLE TECHNOLOGIES THAT ENABLE THE GRID

Oracle has been working for years on technologies that support and enable this vision of grid computing. Some of those technologies are new with Oracle Database 10g, while others are inherent in the Oracle database and have existed for many years. Let's examine these features in the context of the capabilities they enable in a grid.

Computing resource provisioning is one of the most important capabilities of a grid. This enables computing resources to be dynamically provisioned to applications as required. Resources must be appropriately allocated based on business priorities and demand. Oracle provides a number of features for computing resource provisioning, including:

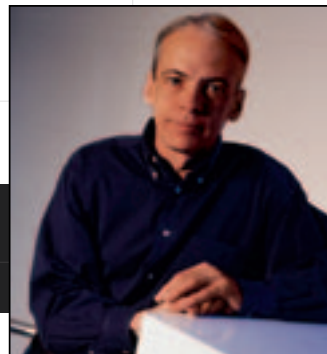
Real Application Clusters. Oracle RAC is the linchpin of a standard grid. RAC is a cluster database with a shared cache architecture that runs on multiple machines, attached through a cluster interconnect, and a shared storage subsystem. An Oracle RAC database not only appears like a single standard Oracle Database 10g to users, but the same maintenance tools and practices used for a single Oracle Database 10g can be used on the entire cluster. All of the standard backup and recovery operations, including the use of Recovery Manager, work transparently with RAC. All SQL operations, including data definition language and integrity constraints, are also identical for both configurations. The most important part of RAC, however, is the ability to manage your workload—to add nodes or relinquish nodes on demand—based on your business processing needs.

Automatic Storage Management. Oracle recommends using Automatic Storage Management (ASM) for your database files and a cluster file system for the Oracle home. "ASM is a new Oracle Database 10g feature that simplifies the administration of Oracle database files," says Rich Long, senior development manager for ASM. "Instead of managing many database files, ASM requires you to manage only a small number of disk groups. You can define a particular disk group as the default disk group for a database."

"ASM provides many of the same benefits of storage technologies such as RAID," adds Angelo Pruscino, Oracle vice president of Cluster and Parallel Storage Technology. "Like these technologies, ASM enables you to create a single disk group from a collection of individual disk devices. Unlike RAID, ASM implements striping and mirroring at the file level."

Oracle Resource Manager. Though Oracle Database 10g is largely a self-managing database, administrators

Benny Souder, Oracle Vice President for Distributed Database Development (left), and Brajesh Goyal, Oracle Principal Product Manager for Grid Computing (right)



need to exercise control over the manner in which database resources are utilized. One of the key tools used to manage resources within an Oracle database is the Database Resource Manager, which allows resource administrators to influence how the Oracle database resources are allocated to grid users. **Oracle Scheduler.** Oracle Database 10g also introduces Oracle Scheduler, which provides many capabilities to schedule and perform business and IT tasks, called jobs, in a grid.

Information provisioning is another important capability supported by Oracle features, which means delivering information to users whenever they need it, regardless of where it resides on the grid. To process information on any available resource, the grid must efficiently share information across distributed systems. The grid must also provide access to data residing on heterogeneous systems—database systems from multiple vendors and file systems. Oracle provides a broad set of features and tools for information provisioning on a grid, including the following: **Oracle Transportable Tablespaces.** Transportable Tablespaces allows Oracle datafiles to be unplugged from a database, moved or copied to another location, and then plugged into another database. Unplugging or plugging a datafile involves reading or loading only a small amount of metadata. Transportable Tablespaces also supports simultaneous mounting of read-only tablespaces by two or more databases.

Oracle Streams. Some data needs to be shared as it is created or changed, rather than occasionally shared in bulk. Oracle Streams can stream data between databases, nodes, or blade farms in a grid and can keep two or more copies in sync as updates are applied. It also provides a unified framework for information sharing, combining message queuing, replication, events, data warehouse loading, notifications, and publish/subscribe into a single technology.

With the Oracle Database 10g release, an additional combination Streams/Transportable Tablespace feature enables what Brajesh Goyal calls a "self-propelled database." Goyal, Oracle principal product manager for Grid Computing, says that with a single command, you can take a tablespace from one database, ship the tablespace to another database, reformat the tablespace if the second database is on a different OS, plug this tablespace into the second database, and start syncing the tablespace with the changes happening in the first database. "If the second data-

base is on a grid," explains Goyal, "what you have just done is migrate your application to a grid with a single command."

Easy OCI/JDBC Install. With Oracle Database 10g, any application using Oracle Call Interface (OCI) can now easily install a small-footprint version of the Oracle client files without installing and configuring the entire client. These applications include Java database connectivity (JDBC) type-2 driver applications. OCI and the JDBC type-2 driver (which requires OCI) can easily be installed by downloading a small subset of the Oracle client files and updating a few environment variables to point to the location of the downloaded libraries. This means your grid client doesn't need to install any Oracle software and yet you get easy, secure access to data from an Oracle database running on a grid. **Distributed SQL and Distributed Transactions.** Oracle Distributed SQL allows grid users to efficiently access and integrate data stored in multiple Oracle and non-Oracle databases.

Transparent remote data access with Distributed SQL allows grid users to run their applications against any other database without making any code change to the applications. While integrating data and managing transactions across multiple data stores, the Oracle database intelligently optimizes the execution plans to access data in the most efficient manner.

Ultra Large Database Support. Oracle Database 10g introduces an important change for enabling Ultra Large Databases (ULDB), the Bigfile Tablespace, which allows Oracle Database to contain tablespaces made up of single large files rather than numerous smaller ones. This allows Oracle Database to utilize the ability of 64-bit systems to create and manage ultralarge files. The consequence of this is that Oracle Database 10g can now scale up to 8 exabytes in size, 8 times the capacity of Microsoft SQL Server 2000 and over 1,000 times the capacity of IBM DB2 Version 8.

MANAGING THE GRID

Since some of the key goals of a grid are to provide high availability, scalability, and service performance optimization with minimal costs and complexity, an integrated approach to grid management is necessary. Oracle has worked hard over the years to provide extensive support for managing distributed resources within an enterprise. These tools and features include: **Grid Management with Oracle Enterprise Manager.** Enterprise Manager provides a simplified, centralized management frame-

THE ROLE OF LINUX IN THE GRID

Linux is the OS of choice for an industry-standard grid for several reasons. First, it's cheap, with a variety of vendors providing a range of low-cost supporting services. It's also a multiuser, multitasking networked operating system that's extremely stable and can run for years without a crash. It provides very high performance on lower-cost commodity systems; companies have built Linux clusters out of cheap PCs and blades that rival the computational power of multimillion-dollar supercomputers. And finally, Linux developers are fanatics for unwavering compliance with public and industry standards—which is a critical feature. Any business building an internal grid will need to talk to other external grids in the future.

Oracle supports all the components and standards for a standard grid architecture—the blades, the interconnects, and the Linux OS (whether Red Hat or United Linux)—in addition to other possible grid components within heterogeneous settings; for example, Sun 64-CPU SMP servers running Solaris. When using Oracle, the platform is not an issue.

LEFT PHOTO: JULE NESTGEN; RIGHT PHOTO: MARKOT HARTFORD



Bob Thome, Group Product Manager of Grid Computing (left), and Angelo Pruscino, Vice President of Cluster and Parallel Storage Technology at Oracle (right).

work for managing enterprise resources and analyzing a grid's performance. With Enterprise Manager, grid administrators can manage the grid environment through a Web browser throughout the system's lifecycle, front to back, from any location on the network. With Oracle Database 10g, this includes integrated management and monitoring of RAC databases as well as standby Data Guard systems, for failover or switchover scenarios.

Managing Security in the Grid. The dynamic environment in a grid makes security extremely important. Oracle makes managing security easy for you by centralizing security management for a distributed enterprise using the Lightweight Directory Access Protocol (LDAP)-compliant Oracle Internet Directory (OID). In this regard, Enterprise User Security provides the ability to create and manage privileges of a user globally—across all enterprise databases. The enterprise user privilege administration is done in the OID, thus avoiding the need to create the same user in multiple databases across a grid. Additionally, with the Oracle Database 10g release, you now have the ability to store a Secure Sockets Layer (SSL) Certificate in a smart card, for roaming access to the grid.

Virtual Private Database. "VPD [Virtual Private Database] provides server-enforced, fine-grained access control and a secure application context that can be used within a grid setting to enable multiple customers, partners, or departments utilizing the same database to have secure access to mission-critical data," says Sudha Iyer, principle product manager, Oracle Database Security. "VPD enables per-user and per-customer data access within a single database, with the assurance of physical data separation. VPD is enabled by associating one or more security policies with tables or views."

Oracle Label Security. Oracle Label Security gives administrators an out-of-the-box row-level—and now column-level—security solution for controlling access to data based on its sensitivity, eliminating the need to manually write such policies. "Using the GUI tool Oracle Policy

Manager, administrators can quickly create and assign Oracle Label Security policies to rows and columns within application tables," says Paul Needham, Oracle director of Product Management, Database Security. "Moreover, Oracle Database 10g now adds integration of Oracle Label Security with OID, allowing policies to be managed centrally within a dynamically changing grid setting."

THE FUTURE OF GRID COMPUTING

Grid computing is poised to dramatically lower the cost of computing and to extend the availability and utility of computing resources across departments and organizations. Within the enterprise, the various strategies of sharing and pooling resources will be tried, often with the attendant problems that derive from sharing fixed CPU cycles, as well as those that derive from human reluctance to share limited resources. Between enterprises, security issues will remain a big stumbling block for some time, though Oracle believes that the proper use of its own security technology will provide a strong basis for extending internal grids outside the firewall. And finally, if distributed resources are to be made available on demand, with the response

time needed for the most demanding applications, network speeds must be more than just adequate for participating partners.

Dedicated server blade clusters now provide a way for businesses to sidestep many of the barriers to building a grid. The new Oracle Database is up to the task of provisioning resources on an industry-standard grid. Plus, all the operational goals for a grid—such as flexibility, portability, high availability, security, and scalability—are met by an Oracle grid, and system managers have a full arsenal of Oracle tools to manage it. ■

Michael Miley (mmiley@pacbell.net) is a freelance writer living in Sonoma, California.

nextSTEPS

READ more about Oracle's grid technologies

otn.oracle.com/products/oracle9i/grid_computing

LEARN more about Oracle Real Application Clusters

otn.oracle.com/products/database/clustering

DISCOVER what Oracle Consulting is doing around grid computing

oracle.com/consulting/offerings/platform

VISIT the Global Grid Forum

www.gridforum.org

EXPLORE

Ian Foster's listing of key grid projects

www.mcs.anl.gov/~foster/grid-projects

CERN's LHC computing grid

otn.oracle.com/oramag/oracle/cerngrid