

WHITE PAPER

Oracle 10g: Putting Grids to Work

Sponsored by: Oracle

Dan Kusnetzky

Carl W. Olofson

April 2004

IDC OPINION

IDC believes that organizations wishing to move in an evolutionary way to Grid/Utility Computing architectures would be well advised to consider Oracle's Oracle 10g family of products. Applications currently written for Oracle's application server or data management software will be able to take advantage of this distributed computing architecture with little or no change. Organizations using applications using Oracle application server and data management software are very likely to find the transition straightforward.

IN THIS WHITE PAPER

Grid Computing represents an evolutionary phase in the history of distributed computing architectures. Some would call it a new disruptive technology even though it is clearly an outgrowth of established distributed computing models. IDC has been tracking distributed computing architectures, such as Grid Computing, and the software that supports this approach for quite a number of years. This paper will discuss Grid Computing as the next phase of computing architectures, IDC's model of distributed computing software, called the virtual environment, and how it is implemented in the form of Grid Computing. It will present Oracle's Oracle 10g family of Grid Computing software and then evaluate it in terms of IDC's model.

Grid Computing represents an evolutionary phase in the history of distributed computing architectures.

SITUATION OVERVIEW

For over a decade Oracle has been providing software that allows organizations to adopt distributed computing architectures as part of their business transactional applications, content management applications, and business intelligence applications. Grid or utility Computing is based upon a well-established foundation — a distributed computing architecture often utilized in the world of high-performance technical computing. This approach allows organizations to harness together the power of the newest generation of high-performance and low-cost systems. This paper presents Oracle's Oracle 10g family of products and how Oracle 10g allows organizations to deploy Grid Computing as a foundation for business-oriented transactional, content management, and business intelligence applications.

Grid Computing — The Next Phase of Computing

IDC expects that the adoption of advanced distributed computing architectures, also known as "Grid Computing," as a platform for business-oriented applications will follow the same pattern as technological innovations and begin to supplant more traditional styles of computing for many types of applications. This is likely to first occur for applications that are relatively straightforward to segment into components or functions that can be distributed. As both networking technology and distributed computing software improve, this approach will be applied to more applications.

End-user organizations are likely to take an evolutionary, not a revolutionary approach to adopting this technology. Established applications are likely to continue on as before. Grid Computing will increasingly be considered as a platform for new application adoption. Oracle's approach allows a mix of both approaches. It's possible to add Grid Computing technology to an established application that is based upon Oracle's application server or data management software products without also requiring an extensive rewrite.

Grid Computing will increasingly be considered as a platform for new application adoption.

Market Drivers Leading to Adoption

IDC's market studies have repeatedly indicated that organizations are seeking approaches that are able to address their concerns including the following:

- The ever-growing complexity of their IT infrastructure.
- The desire to lower hardware, software, and staff-related costs of their IT infrastructure.
- An increasing need for robust, reliable systems, which also allow the organizations to be able to respond rapidly to a changing environment.
- An increasing need to protect the organizations' IT infrastructure from outside intrusion or misuse.
- The organizations also wish to use the newest technology, when and where it is appropriate, without also having to abandon investment in older technologies.

Complexity!

Over time, organizations have adopted new information technologies to support new applications and placed them alongside older technology rather than abandoning their previous investment in systems, software, and processes. This has often led to an IT infrastructure that resembles a layer cake. The foundation often is 20 years old. Layered on top of that foundation are layers of new technology.

Complexity clearly can increase an organization's staff-related costs. It may also decrease an organization's ability to respond to rapidly moving market forces.

Complexity clearly can increase an organization's staff-related costs.

Driving Down Costs

Declining or stagnant revenue has forced many organizations to seek ways to trim costs. Their IT budgets either have been reduced while the IT organization is expected to continue supporting all application systems and business initiatives or have remained static while the IT organization is expected to provide expanded levels of service to the organization. Thus, the IT staff faces the paradox of having to do more with less. To address these issues, IT organizations have turned to the following techniques:

- ☒ Consolidating multiple workloads on fewer, larger system configurations in a smaller number of datacenters
- ☒ Increasing focus on the use of multisystem configurations (blade computing, clustering, or grids) constructed of high-volume, low-cost systems
- ☒ Reducing the number of system architectures utilized throughout the network to make development, administration, operations, and support easier and, thus, less costly

Need for Reliable Systems

Organizations rely more heavily on their IT infrastructure than ever before. The systems no longer can be considered "field," "front-office," or "back-office" systems. Increasingly, IDC's studies indicate that an organization's information systems are accessed by customers, partners, suppliers, and mobile employees as well as the traditional in-house users. This means that if an organization's systems are not available continuously, it can lose business to competitors. The customers lost in this fashion may never come back. Although the costs of lost productivity are hard to measure, partners, suppliers, mobile employees, and in-house users certainly are not productive when supporting systems that are unavailable for any reason.

Organizations rely more heavily on their IT infrastructure than ever before.

Availability of New Technology

New information technology is becoming available all of the time. Organizations, looking to lower their costs, are trying to adopt high-volume and, thus, lower-cost, hardware and software. Their challenge is bringing in new technology without also being forced to abandon years of investment in application systems, tools, processes, and training that are currently supporting their business.

Organizations, looking to lower their costs, are trying to adopt high-volume and, thus, lower-cost, hardware and software.

Better Utilization of IT Infrastructure

Organizations have installed a number of different computing solutions and are now finding that these systems are not fully utilized nor do they interoperate as a single computing resource. IT management is often seeking ways to increase the overall utilization of its systems and integrate application systems.

Since organizations are facing a rapidly changing environment, they are also concerned about agility. That is, they are seeking ways to align their system resources to address today's need, knowing that tomorrow they may need to realign to address a different set of needs.

Distributed Computing: Where Does Grid Computing Fit?

IDC has observed five generations of distributed computing architectures over time. Grid Computing is merely the newest generation of distributed computing. Table 1 lists them. The industry is clearly entering the fifth generation now.

TABLE 1

Five Generations of Distributed Computing

Generation	Characteristics
First (Host-based Computing)	<ul style="list-style-type: none"> • Dumb Terminal • Single Server • Monolithic Applications
Second (Remote Access)	<ul style="list-style-type: none"> • Single Client supporting only terminal emulation functions • Single Server
Third (Client/Server)	<ul style="list-style-type: none"> • Single Client supporting rules processing as well as user interface • Up to two servers
Fourth (Multitier)	<ul style="list-style-type: none"> • Single Client supporting rules processing as well as user interface • More than two server tiers
Fifth (Grid Computing)	<ul style="list-style-type: none"> • Virtual Environment where all systems are considered a pool of resources • N-tier • Service-Oriented Architectures

Source: IDC, 2004

Entering the Fifth Generation — Grid Computing

Grid Computing is the result of several trends coming together. Some of these are the following:

- ☒ New standards for object-to-object communications making it easier to build multivendor, multiapplication networks.
- ☒ High-performance microprocessors have become available, making it possible to deploy large applications on a number of low-cost systems rather than a single mid-range system.
- ☒ High-speed networking technology is becoming both less costly and readily available, offering higher levels of performance when deploying distributed application architectures.

As these trends combine, applications are likely to be segmented by function or instance of a function. This approach will allow each function to be hosted on the most cost-effective platform. In some cases, both types of segmentation will be used. In the end, an organization's systems can be considered a pool of shared resources that adapt automatically to changing conditions and failures based upon rules of the organization's choosing.

Many Buzzwords for Distributed Computing

As each evolutionary step was taken, suppliers created new catch phrases or buzzwords in the hopes of creating greater market awareness and some level of competitive advantage. Grid Computing is the newest addition to the list. Even though each of these steps can be considered just another evolutionary step, the ever-increasing list of buzzwords leaves the impression that it is something new, exciting, and somehow fundamentally different from previous steps. Being able to apply this technology to mainstream business applications, not the technology itself, is the true source of the excitement this time.

Being able to apply this technology to mainstream business applications, not the technology itself, is the true source of the excitement this time.

Buzzwords such as "adaptive computing," "autonomic computing," "on-demand computing," and "N1" all discuss the same basic concept as Grid Computing. So, in this paper, IDC will use the term Grid Computing to cover all of these different approaches to the same concept. The emergence of all of these different terms has left an unnecessary impression that this technology is complex and difficult to understand, each supplier is doing something different, and thus, the whole concept is not ready to support critical applications. This simply isn't true.

Virtual Environment Software

Grid Computing is built upon the foundation of a stack of software IDC calls the "virtual environment" (see Figure 1). This software creates an image of a single computing resource to the end user even though the actual computing environment might be made up of distributed systems housed in datacenters all over the globe.

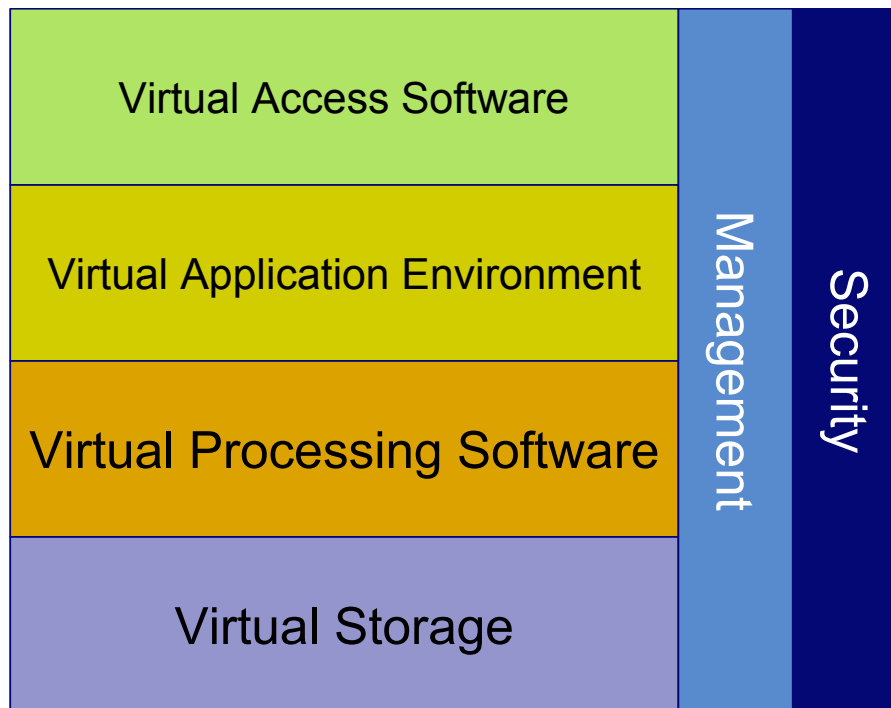
Virtual environment software can break the link between a given function and the underlying systems. This means that functions may survive the loss of their original host system. In case of a failure or slowdown, some forms of virtual environment software will either start the function on another system or pass the request to another instance of that application or function. The newest generation of virtual environment software allows organizations to increasingly see their systems as a pool of shared resources that appear to be both self-healing and self-managing.

The newest generation of virtual environment software allows organizations to increasingly see their systems as a pool of shared resources that appear to be both self-healing and self-managing.

In the end, this allows organizations not only to protect their investment in hardware and software but also to optimize this investment. A completely virtual environment allows established applications or functions to access features of newer systems and to be more reliable, more powerful, more scalable, or enhanced in some other way.

FIGURE 1

IDC's Virtual Environment Software Model



Source: IDC, 2004

The following bullets describe each of the segments of the model:

- ☒ **Virtual Access Software** — This software allows applications to be accessed from nearly any intelligent access point device over just about any network without the application having to have been architected to support that device or network. Virtual user interface software and various types of portal software fit in this category.
- ☒ **Virtual Application Environments** — This software creates an application development and deployment environment that allows properly developed applications to be more robust and reliable and also allows them to be unaware of the underlying operating environment and hardware platform. These benefits are only available to applications written for this environment. Application server software and parallel database software fit in this category.
- ☒ **Virtual Processing Software** — This category of software ranges from virtual machine software making a single system appear to be many systems, each supporting its own operating environment, to single system image clustering software, which makes many systems appear to be a single computing resource, running a single operating environment. This category also includes parallel processing software, load balancing software, and data and application availability software.

- ☒ **Virtual Storage Software** — This software allows applications to be unaware of where and how application and data files are actually stored. This category includes storage replication and file system software. This software supports both storage area network (SAN) and network attached software (NAS) hardware configurations.
- ☒ **Server Provisioning and Management** — This software makes it possible for operators and administrators to load, manage, and operate multisystem configurations regardless of whether any of the other virtual environment software categories are present. This is a primary component of "on-demand" or "adaptive environment" approaches to application deployment.
- ☒ **Security** — As applications are decomposed into components (sometimes called services), identity management and access control become increasingly important. Without a strong security layer, "black hats" could pick apart the distributed architecture and commandeer application components or functions to the organization's detriment.

Grid Computing: Putting Virtual Environment Software to Work

Grid Computing is a specific implementation of distributed computing using virtual environment software. The goal is typically one of the following: raw application performance, scalability, reliability, and, only recently, making the entire computing infrastructure appear to be a pool of resources that can be used as a single, self-managing, self-healing system. Grid Computing allows applications and the constituent components or functions to be segmented both horizontally (by functional layer) or vertically (by iteration or instance of that function or component). Table 2 lists IDC's model of Grid Computing.

Grid Computing is a specific implementation of distributed computing using virtual environment software.

TABLE 2

IDC's Model of Grid Computing

Type of Grid	Goals	Challenges	Description
Computational	<ul style="list-style-type: none"> • Raw application performance • Scalability • Reliability 	<ul style="list-style-type: none"> • Not all applications can be segmented in this fashion • Emerging standards (OGSA) 	<ul style="list-style-type: none"> • Multisystem configuration focused on high computational performance
Application	<ul style="list-style-type: none"> • Scalability • Robustness • Reliability • Availability 	<ul style="list-style-type: none"> • Established applications may need to be rewritten • Not all applications can be segmented in this fashion 	<ul style="list-style-type: none"> • Multisystem configuration focused on application scalability, reliability, and availability
Storage	<ul style="list-style-type: none"> • Scalability of storage • Reliability/availability of storage • Optimal use of storage devices • Increase interoperability 	<ul style="list-style-type: none"> • Single vendor approaches • Single operating system approaches • Single data management approach 	<ul style="list-style-type: none"> • Multisystem configuration focused on storage scalability, reliability, and availability
Optimization	<ul style="list-style-type: none"> • Scalability • Reliability • Optimal use of organization's resources allowing all systems to appear to be a single pool of available resources 	<ul style="list-style-type: none"> • Sophisticated server provisioning and management software needed • Multisystem workload management required • Single vendor approaches • Single operating system approaches • Single hardware architecture approaches 	<ul style="list-style-type: none"> • Multisystem configuration focused on dynamic resource allocation, scalability, reliability, and manageability • A complete implementation of the Grid Computing concept

Source: IDC, 2004

Challenges to Grid Adoption

Although Grid Computing has been utilized in academic, research, and scientific applications for a number of years, it is only now becoming possible to deploy commercial, transactional applications upon this foundation. As the industry moves forward in "normalizing" Grid Computing, it will need to overcome the following challenges:

- ☒ **Performance concerns** — Network latency can degrade performance as well as application function interdependency, which makes it difficult or impossible to segment the functions to run on several systems.
- ☒ **Security concerns** — Once application functions or components are exposed to the network, will it be easier for the "black hats" to find and exploit vulnerabilities? Identity management, access control, and other security issues will extend all the way down to the application function or component level.
- ☒ **Management concerns** — Complex system configurations may increase costs of installation; provisioning; operations; administration; and problem isolation, identification, and resolution.

☒ **Expertise concerns** — Will an organization's IT staff be able to fully utilize this technique? While it is likely that medium-sized and large organizations will have the expertise on staff to take full advantage of this technology, smaller organizations will have to rely on service providers or outside experts.

IDC believes that these concerns, in some cases, are more perception than reality. As organizations have more experience with this distributed approach, their concerns will be laid to rest.

IDC believes that these concerns, in some cases, are more perception than reality.

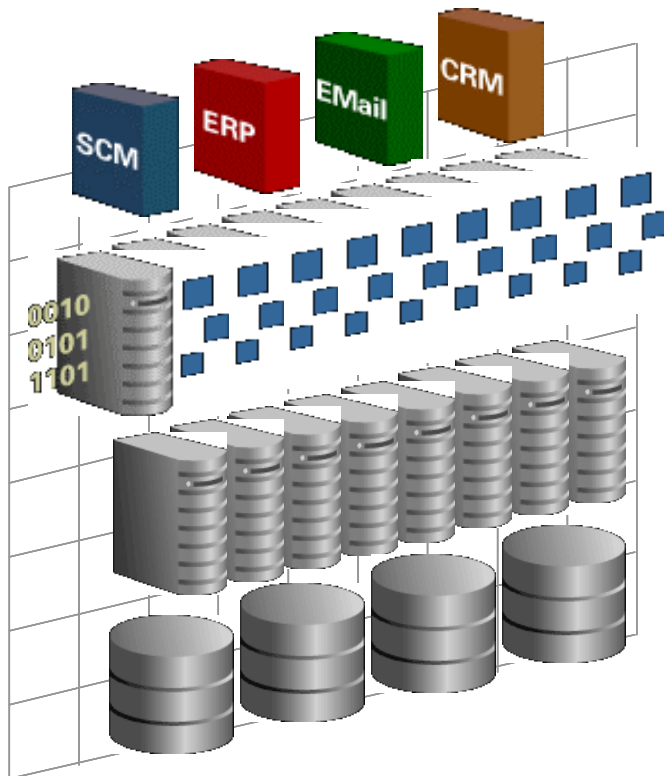
Oracle 10g Family of Software Products

Oracle announced Oracle 10g, the newest edition of its family of software products, on September 8, 2003. Oracle announced Oracle Database 10g, Oracle Application Server 10g, and Oracle Enterprise Manager 10g. The focus of this new edition is enabling organizations to simplify the process of putting Grid Computing to work outside its traditional base of academic, engineering, research, and scientific computing. This software is an evolutionary step based upon Oracle's years of experience both in distributed computing architectures and in supporting multivendor environments. Oracle's previous generations of technology are clearly the foundation for this software. Figure 2 shows a conceptual diagram of this type of configuration.

The focus of this new edition is enabling organizations to simplify the process of putting Grid Computing to work ...

FIGURE 2

Grid Computing with Oracle 10g



Source: Oracle and IDC, 2004

As described earlier in this paper, concerns about performance, security, and manageability provide a challenge to grid adoption. A variety of enhancements have been made to the Oracle 10g family of software products to address these concerns.

Oracle Database 10g

Oracle Database 10g includes a variety of Grid Computing features that provide efficient utilization of enterprise resources and dynamically realign enterprise resources based on business needs.

Server Virtualization and Provisioning

A key feature of Oracle Database 10g is Real Application Clusters, introduced in Oracle9i and in use at hundreds of customer sites. Oracle Real Application Clusters enables a single database to run across multiple clustered nodes in a grid, pooling the processing resources of several standard machines. Oracle can flexibly provision workload across machines because it does not require data to be partitioned and distributed along with the work.

A key feature of Oracle Database 10g is Real Application Clusters ...

In Oracle 10g, the database can immediately begin balancing workload across a new node with new processing capacity as it gets reprovisioned from one database to another and can relinquish a machine when it is no longer needed, thus providing capacity on demand. Multiple databases running on the same cluster automatically load balance based on the defined policies, thus matching resource allocation with business needs.

New integrated clusterware, the software that provides needed clustering services, in Oracle 10g reduces the challenges organizations face when deploying clustered systems by reducing or eliminating the need to purchase, install, configure, and support third-party clusterware. Servers can be easily added and dropped to an Oracle cluster with no downtime. Oracle includes clusterware for all supported operating systems, which is expected to reduce the opportunities for failure in a clustered environment.

Servers can be easily added and dropped to an Oracle cluster with no downtime.

Oracle Database 10g can also treat individual systems or groups of systems as a pool of resources. A node or a blade can be added or dropped while the application is running, allowing the system configuration to be tuned to best meet current application requirements. Low-cost systems in a cluster can also be allocated to specific tasks to better support service level agreements. This capability combined with the software's support of policy-based management means that the configuration can be viewed as both self-managing and self-healing.

Storage Virtualization and Provisioning

Oracle Database 10g also introduced Automatic Storage Management (ASM), software designed to simplify storage configuration and management for the database. ASM hides the underlying complexity of how the database deals with data files and storage subsystems. It also automatically distributes the storage workload to get the best possible system performance. ASM eliminates the need to constantly monitor storage systems for "hot spots" or performance bottlenecks that often slow down data processing.

Oracle has worked with several industry storage leaders such as EMC, Hitachi, HP, Network Appliance, and XioTech to enable ASM to work seamlessly with their network attached storage (NAS) and storage area network (SAN) offerings.

Oracle Database 10g is designed to be effectively deployed on many different hardware configurations including small servers, SMP systems, clusters, and enterprise grids. This new software also provides automated tuning and management capabilities to simplify database operations and lower administration costs.

Information Virtualization and Provisioning

In addition to the provisioning of work across multiple nodes and the provisioning of storage across multiple disks, another type of provisioning happens within Oracle Database 10g — the provisioning of information itself. Depending on the volume of information and the frequency of access, it may be necessary to move data from where it currently resides or to share data across multiple databases. Oracle 10g includes various facilities to provide access to information when and where it's needed, matching information providers and information requestors. The most fine-grained and real-time of these facilities is Oracle Streams, which can migrate data from one database to another while both are online. Bulk data transfers are more suitable in some circumstances, for which Oracle provides Data Pump and Transportable Tablespaces. In Oracle 10g, all information-provisioning facilities can move data to databases running on different operating systems, which is particularly useful for migrating databases to a grid environment; for example, blade servers running Linux.

Oracle 10g includes various facilities to provide access to information when and where it's needed ...

Structured and Nonstructured Data

Oracle Database 10g can manage data in many forms, including traditional structured business information for transactional applications or online analytical processing, XML documents, or spatial/location information.

Oracle Database 10g can manage data in many forms ...

When viewed from IDC's Virtual Environment Software model, Oracle's Oracle Database 10g offers an integrated set of capabilities that includes features normally found in the management, security, virtual application environment software, virtual processing software, and virtual storage software categories.

Oracle Application Server 10g

Oracle Application Server 10g, the next generation of Oracle's *integrated suite of application infrastructure software*, has been specifically designed to run enterprise applications on computing grids. It is designed to run enterprise applications on pools of low-cost servers and storage with very high performance, scalability, and availability while reducing the costs of systems and applications monitoring and management. Further, customers can deploy all their existing Oracle9i Application Server applications on Application Server 10g without any changes and take advantage of the new grid technology features such as policy-based resource management, metric-based workload management, and centralized user provisioning. This software works in conjunction with the Oracle Database 10g to provide a grid-enabled infrastructure.

Infrastructure Platform

Oracle Application Server 10g provides a complete infrastructure platform for developing and deploying enterprise applications, integrating many functions including a J2EE and Web services runtime environment, an enterprise portal, an enterprise integration broker, business intelligence, Web caching, and identity management services.

Tightly integrated with Oracle Application Server 10g are the development tools that enable companies to quickly develop custom Internet applications and then easily deploy those applications to Oracle Application Server. Enterprise applications do not need to be redesigned to exploit the availability, scalability, and performance benefits of enterprise grids. When applications are deployed to an application server in a grid, those applications benefit immediately from the transparent workload distribution, load balancing, and scheduling necessary to efficiently coordinate work across multiple servers.

To gain additional benefits from Grid Computing, however, enterprise applications can expose their behavior to other applications and to management tools through standardized interfaces in a service-oriented architecture. Oracle Developer Suite 10g, which includes JDeveloper 10g, enables developers to create dynamic Web sites, J2EE applications, and Web services and to make these services accessible through enterprise portals and wireless devices.

Application Virtualization and Provisioning

Oracle Application Server 10g runtime services can be pooled and virtualized via application server clusters. Every service within the Oracle Application Server — HTTP, J2EE, Web cache, Web Services, LDAP, portal, and others — can be distributed across multiple machines in a grid. New features in Oracle 10g enable performance thresholds to be defined beyond which new application server instances can automatically be added and started (or relinquished) to process additional work on new nodes of a grid, delivering capacity on demand.

With Oracle 10g, an administrator can define a set of policies or business rules that affects how individual work is provisioned across multiple machines. Specifically, workload allocation can be influenced by resource consumption metrics, such as CPU or memory usage, or application-specific metrics, such as transaction throughput or JDBC connections, or workload can be provisioned based on schedules, such as peak times of day or end of quarter.

Application Server Integration with Database Server

Oracle Application Server 10g provides out-of-the-box instrumentation that captures these various metrics and creates advisories based on historical and real-time information to help administrators make the best policy choices. Oracle Application Server 10g also provides several availability enhancements. Because Oracle 10g includes clustering of every service within the application server, there is no single point of failure. Both planned and unplanned downtime of an individual instance will simply cause requests to be routed to another node. Because Application Server 10g includes efficient session replication, any type of failure (even that of a J2EE application holding state) will remain transparent to the user.

Enterprise applications do not need to be redesigned to exploit the availability, scalability, and performance benefits of enterprise grids.

Oracle Application Server 10g runtime services can be pooled and virtualized via application server clusters.

Because Oracle 10g includes clustering of every service within the application server, there is no single point of failure.

Oracle Application Server 10g allows organizations to view nearly every system service, including HTTP, cache, Java, LDAP, and others, as highly scalable. As with Oracle Database 10g, this software can automatically add or drop nodes in a cluster. It can also automatically distribute loads in a cluster. This capability also provides a self-managing, self-healing capability.

This capability also provides a self-managing, self-healing capability.

Applications designed to a service-oriented architecture can leverage a set of standards-based Internet protocols to communicate with other applications and heterogeneous resources across a grid. Designing to a service-oriented architecture enables companies to reduce development time and integration costs.

Applications designed to a service-oriented architecture can leverage a set of standards-based Internet protocols ...

Application Server 10g further improves application reliability through its interaction with Oracle Real Application Clusters. If an instance in the back-end database goes down, Application Server 10g is notified to reconnect. Without notification from a failed instance, an application server would wait for an IP time-out, which takes several minutes, but the multitier failover notification feature reduces recovery time in such cases to mere seconds, and both failure and recovery remain transparent to the user.

As with Oracle Database 10g, Oracle Application Server 10g integrates capabilities found in several of IDC's virtual environment software categories, including management, security, virtual access software, virtual application environment software, and the load balancing capabilities typically found in virtual processing software. These capabilities would allow Oracle Application Server 10g to be the application platform for application grids and optimization grids.

Grid Security

The dynamic nature of the grid makes security extremely important. Oracle Database 10g provides a variety of mechanisms that enable enterprises to manage their security needs in a grid environment. Enterprise User Security centralizes the management of user credentials and privileges in a directory, eliminating the need to create the same user across multiple databases in a grid.

The dynamic nature of the grid makes security extremely important.

Virtual Private Database (VPD) and Oracle Label Security can be used to ensure that only authorized users have access to selected data in a grid environment. VPD provides server-enforced, fine-grained access control and a secure application context that can be used in a grid setting to enable multiple database users to have secure access to data in a single database, with the assurance of physical data separation. Oracle Label Security provides access control to data at the row and column levels based on the sensitivity of the data.

Centralized application user administration becomes even more important in a grid environment. Identity management features within Application Server 10g simplify and centralize account creation, suspension, and deletion and privilege modification, all of which lower administration costs and reduce security vulnerabilities.

Centralized application user administration becomes even more important in a grid environment.

Oracle provides centralized user provisioning and single sign-on for users across all applications deployed to the Oracle Application Server. Access privileges for all applications can be created and revoked through a single interface. Identities can be managed through Oracle Internet Directory, a standards-based LDAP directory that benefits from the availability and scalability of being built on the Oracle Database.

Oracle Enterprise Manager 10g

Oracle Enterprise Manager 10g with new Grid Control feature (Oracle Grid Control) is Oracle's comprehensive, integrated, central management console and underlying framework that automates administrative tasks across sets of systems in a grid environment. Oracle Grid Control helps reduce administration costs through automation and policy-based standardization. With Oracle Grid Control, IT professionals can group multiple hardware nodes, databases, application servers, and other targets into single logical entities. By executing jobs, enforcing standard policies, monitoring performance, and automating many other tasks across a group of targets instead of on many systems individually, Oracle Grid Control enables IT staff to scale with a growing grid. Because of this feature, the existence of many small computers in a grid infrastructure does not increase management complexity.

Oracle Grid Control helps reduce administration costs through automation and policy-based standardization.

Because of the potentially large number of physical nodes, it's especially important in a grid environment that installation and configuration of the software running on those nodes occur quickly and require no human intervention. Manually installing software on hundreds of nodes would be time consuming and cumbersome.

With Oracle Grid Control, Oracle 10g automates installation, configuration, and cloning of Application Server 10g and Database 10g across multiple nodes. Oracle Enterprise Manager provides a common framework for software provisioning and management, allowing administrators to create, configure, deploy, and utilize new servers with new instances of the application server and database as they are needed. This framework is used not only to provision new systems but also to apply patches and upgrade existing systems.

With Oracle Grid Control, Oracle 10g automates installation, configuration, and cloning of Application Server 10g and Database 10g across multiple nodes.

In Oracle Application Server 10g, applications can be deployed once to a single application server instance, registered with the central repository, then automatically deployed to all relevant nodes in the grid. As changes are made to the application and as new nodes are added to the grid, nodes can be kept in sync.

Oracle Grid Control views the availability and performance of the grid infrastructure as a unified whole, as a user would experience it, rather than as isolated storage units, processing boxes, databases, and application servers. An administrator can trace a performance or availability problem as experienced by a user from end to end — from the user visible Web page, through external and internal networks, to application code, application server, and database access. Oracle Grid Control then allows an administrator to trace the root cause of the problem down to the individual Java class, for example, or the individual system configuration parameter.

As with other members of the Oracle 10g family of software products, this software offers functions that typically can be found in the management and security software categories of IDC's Virtual Environment Software model.

Oracle's Vision of Grid Computing

Oracle's 10g technology is designed to be deployed in several different types of grids. Oracle's view of Grid Computing will be presented here.

Oracle's 10g technology is designed to be deployed in several different types of grids.

Oracle's announcement discusses several components of Grid Computing:

- ☒ **Automatic Storage Management** — Oracle Database 10g's Automatic Storage Management supports a virtual storage function by automatically mirroring and striping data for performance and reliability. It can manage all storage for the database without requiring a volume manager or global file system. Storage can be added or removed online. It also provides "self-management" functions so that it is not necessary to constantly tune the storage configuration to avoid hot spots.
- ☒ **Oracle Database** — Oracle Database 10g's clustering capability allows large data management tasks to be effectively handled by clusters or grids of high-volume, low-cost industry-standard systems rather than requiring organizations to turn to midrange or large single computing systems. Organizations can add processors, blades, or complete systems as required to increase capacity as the demand increases, without forcing them to turn to a "rip and replace" strategy.
- ☒ **Oracle Application Server** — Oracle Application Server 10g allows organizations to deploy cluster or grid configurations to support complex or large applications. System services such as HTTP, cache, Java, LDAP, and others can be hosted on these configurations. This allows the supported applications to be highly reliable and to offer scalable performance. Cluster or grid nodes (or blades in a blade computer) can be automatically added or removed as needed to support a dynamic and changing workload. This configuration also supports end-to-end failover and fast fault recovery.
- ☒ **Oracle Enterprise Manager** — Oracle Enterprise Manager 10g with Oracle Grid Control makes it possible for organizations to manage and monitor Grid Computing environments from a central console.

Oracle Application Server 10g allows organizations to deploy cluster or grid configurations to support complex or large applications.

Comparing Oracle 10g to IDC's Model of Grid Computing

It is clear that Oracle's primary focus is to bring the power, scalability, and reliability that multisystem configurations offer to organizations conducting business transactional or business intelligence functions in an easy-to-use, easy-to-manage, integrated approach. This approach makes it possible for Oracle's customers to see Grid Computing as an evolutionary step rather than something that is a disruption to their overall computing environment.

Both Oracle's Application Server and its Database are implementations of IDC's application grid category. Applications built using Oracle's application development and deployment foundation would be able to utilize the multisystem configuration. The users of these applications would be able to realize much of the benefit of Grid Computing today without having to make many changes or do a great deal of work. Applications constructed some other way, on the other hand, would not see a grid at all. They would see a networked environment and would not be able to take full advantage of the grid.

Both Oracle's Application Server and its Database are implementations of IDC's application grid category.

Since Oracle does not offer a general-purpose virtual processing software product, directly serving the virtual processing needs of the high-performance computing community appears not to be Oracle's primary goal. It is possible, however, to construct computational grids using Oracle's software. Oracle's Oracle Database 10g clearly could be the data repository for a computational grid constructed using virtual processing software from any number of software suppliers. The leaders in this category are all Oracle partners.

Oracle's software allows Oracle-centric optimization grids to be constructed today. Creating a broadly defined optimization grid is also not a goal of this release of software, although it clearly could be on Oracle's road map for future releases. Oracle appears to be focused on bringing its customers into the grid computing world with this release of software.

Oracle's software allows Oracle-centric optimization grids to be constructed today.

Oracle's storage grid clearly belongs in IDC's storage grid category. Applications built upon Oracle's Database 10g would be able to see a completely virtual storage environment, one that is easy to manage, powerful, and reliable. Applications not based upon Oracle's database would not be able to utilize all of the features of the grid. As with Oracle's other grid offerings, it is designed to bring the benefits of Grid Computing to users of Oracle's foundation software, not a general-purpose virtual storage offering.

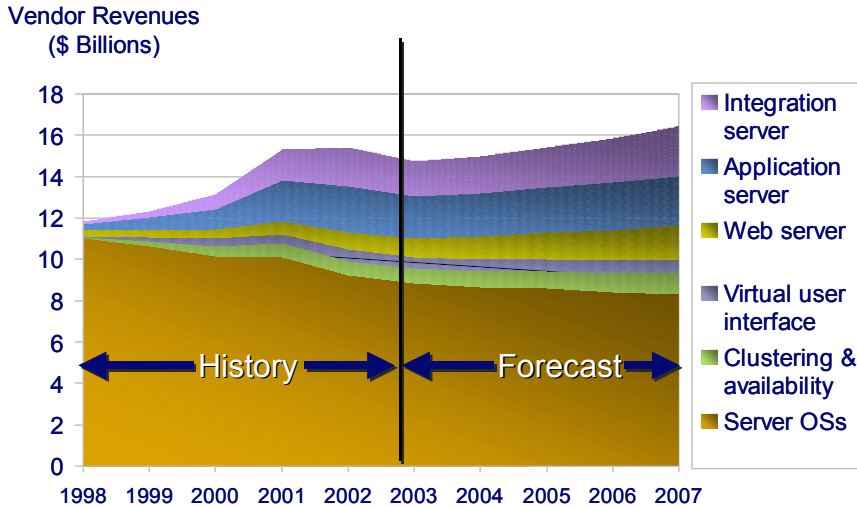
FUTURE OUTLOOK

IDC has been forecasting the virtual environment software markets to grow at a faster rate than the operating system and subsystem market. This is, in part, due to its belief that Grid Computing is going to be increasingly important to organizations.

One can see from actual IDC data that server operating environments are a declining source of revenue while there has been rapid growth in various types of virtualization software (see Figure 3). Currently, the combined worldwide vendor revenue of virtualization products is roughly equal to the revenue of server operating environments. One should note that most of these virtualization products barely existed in 1998, so this represents very sizeable growth.

FIGURE 3

The Worldwide Software Revenue Is Already Shifting Toward Virtualization



Source: IDC, 2004

CHALLENGES/OPPORTUNITIES

Oracle's product offerings are focused on Oracle-based solutions, not generalized offerings. This being said, a great deal of what organizations are seeking when they consider Grid Computing architectures can be supported quite well using Oracle's products.

Competitive products exist at each level of IDC's Virtual Environment Software model (refer back to Figure 1). What Oracle is offering is a high level of integration that would make Grid Computing easier to adopt and use.

IDC believes the largest challenge to Oracle is the fact that Oracle's partners, as well as HP, IBM, Microsoft, Sun, and others, are putting forward their own marketing messages about Grid Computing that may leave the impression that they're different from the approach Oracle is putting forward. In reality, Oracle's approach works well with the approach being offered by the other suppliers. Overcoming the market confusion on this topic is likely to be a challenge for Oracle.

CONCLUSION

Organizations wanting improved reliability and performance at a lower cost for their transactional, business intelligence, or knowledge management applications can deploy Oracle's Grid Computing solutions today. This is especially true for Oracle's current customers, as the adoption of this software would require adopting just the next generation of the software they're already successfully using. IDC also believes that this software is worthy of consideration by others with similar needs.

CASE STUDY: FANNIE MAE

Fannie Mae is a New York Stock Exchange company and the largest non-bank financial services company in the world. It operates pursuant to a federal charter and is the nation's largest source of financing for home mortgages. Since its inception, Fannie Mae has provided a constant source of mortgage funds for homebuyers by investing in mortgages from institutions that originate them such as savings and loan companies, mortgage companies, and commercial banks.

Fannie Mae is the second largest mortgage organization in the United States when viewed by assets, managing a portfolio made up of over \$900 billion in assets. Fannie Mae uses its systems to manage its large portfolio of assets in the most effective way possible.

Hoosh Afsar is part of the group that provides applications and tools to support various internal organizations within Fannie Mae. The applications help decision makers evaluate various options so they can make the best decision possible. The application runs on a specialized grid. This grid uses Oracle's software, but it is not based solely on Oracle's grid software.

Afsar's mandate is to develop a comprehensive monitoring and management system for the variety of applications and the grid itself. The grid requires a great deal of monitoring and management to offer the most comprehensive service on a 24 x 7 basis. Oracle Enterprise Manager Grid Control, which comes with Oracle 10g, provides such an extensive set of tools that it met the majority of Fannie Mae's requirements.

Fannie Mae has experienced higher levels of application reliability and availability since deploying Oracle 10g. The deployment has provided Fannie Mae access to a high level of detail and data from the systems and from the applications. It has also made it much easier to provide high-level information to Fannie Mae executives.

Fannie Mae is now developing a new portfolio management application. The tools Oracle 10g offers make it much easier to monitor the application performance and health. This means that Fannie Mae has a much better understanding of system utilization, allowing it to better control hardware and software acquisition costs.

The result is that Fannie Mae could experience less risk of system outages and higher levels of performance at a lower cost.

There are always problems with any project. Fannie Mae has always been able to resolve them when they appear, either with Oracle support or on its own. The company was running a beta version of the software and eagerly awaiting the final version.

Afsar advises companies to plan carefully before implementation. It would be wise to run a pilot project so that the software can be tested and any configuration issues found and resolved prior to putting it into production. This process also develops a good working relationship with the supplier.

Copyright Notice

External Publication of IDC Information and Data — Any IDC information that is to be used in advertising, press releases, or promotional materials requires prior written approval from the appropriate IDC Vice President or Country Manager. A draft of the proposed document should accompany any such request. IDC reserves the right to deny approval of external usage for any reason.

Copyright 2004 IDC. Reproduction without written permission is completely forbidden.