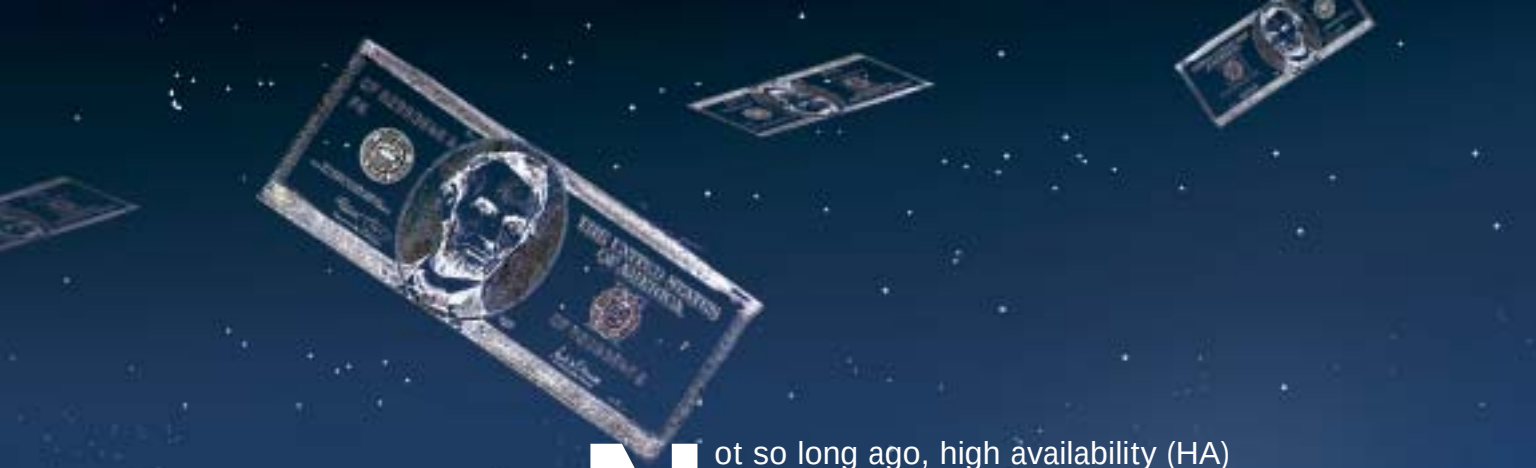
DOWN FROM THE STRATOSPHE

Highly available systems, once afforded by only large enterprises, are now within reach of most businesses

By David Baum

FPO



ERE

Not so long ago, high availability (HA) computing was restricted to truly mission-critical information systems, like stock-trading applications and missions to Mars. Today, both the price of HA solutions and the definition of mission-critical have changed, bringing HA into mainstream business use. As online business applications grow in popularity and importance, keeping systems up and running continuously is part of the expected norm. "Nearly every business is looking for more reliable infrastructures, says Charlie Garry, an analyst with META Group. "Customers expect your systems to be online 24/7, whether you're a small retailer or a global conglomerate."

HA computing can be achieved via many different techniques involving both hardware and software. For example, it can be based on a cluster of processors with failover software. Or, alternatively, HA can be achieved with a single processor with sufficient redundancy, where failover is controlled by a combination of the operating system and special hardware features, such as dual disk drives.

'We could not have done it any other way and still have met our price/performance objectives."—Marc West

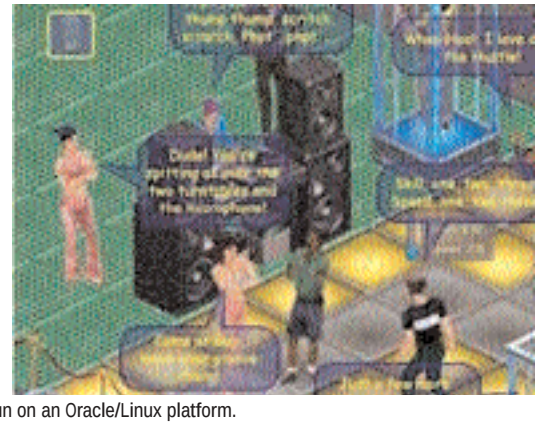
For Oracle database customers, HA is becoming more affordable and approachable than ever, thanks to a software technology for distributing a single database across a cluster of servers. The technology is called Oracle9*i* Real Application Clusters (RAC), and it works with any Oracle database application on just about any computing platform—including the Red Hat Linux Advanced Server operating system, which can be run on comparatively inexpensive Intel processors. "When people look at their needs for high availability, they're really looking at the busines cost of downtime," continues Garry. "They calculate their risk and then try to pick a solution that's priced appropriately. RAC is changing the dynamics of these decisions by making highly available Oracle systems more attractive because it offers both availability and some measure of scalability versus traditional redundant HA infrastructures.

Oracle9*i* RAC runs on top of a hardware cluster—a group of independent servers that cooperate as a single system. The nodes share access to the storage subsystem and resources that manage data, but they do not physically share main memory in their respective nodes. Emerging Storage Area Networks (SANs) provide sophisticated schemes for disk connectivity, circumventing the prior limitations of direct-attached disks and allowing each cluster node to be connected to a large number of storage devices.

According to Sohan Demel, group product manager for Oracle9*i* RAC, most of the leading relational database vendors have developed clustering solutions for data warehouse applications, but until now the industry lacked a general-purpose database solution that could scale beyond a single machine. "While some vendors offer proprietary solutions for clustering OLTP systems, Oracle is the first vendor to introduce a solution for scaling any kind of database application on any type of platform—whether it is an OLTP system, a packaged application, a homegrown database program, or a data warehouse," Demel says.

In practical terms, that means an application does not have to be rewritten, redesigned, or reprogrammed to run concurrently on multiple servers. Oracle9*i* recognizes the application and transparently reconfigures and balances the load across all nodes.

"IBM, HP, and a couple of other vendors offer proprietary hardware/software solutions for clustering databases, but with Oracle9*i* RAC you can select whatever type of hardware you want," says Demel. "Oracle combines the memory in the individual nodes to provide a single view of the distributed cache memory for the entire database system."



The Sims Online from Electronic Arts will run on an Oracle/Linux platform.

STEVEN LYONS

CLUSTERING MAGIC

A hardware cluster comprises two or more nodes (or servers) tightly coupled together to share resources such as disks and processors across all nodes in a cluster. While clustered database systems share some of these characteristics, they differ widely in architecture, implementation, and administrative requirements. This is true for the hardware cluster systems as well as the clustered database software that runs on these clustered systems.

The magic behind Oracle9*i* RAC is *Cache Fusion*, a shared-cache architecture that overcomes the limitations of traditional shared-nothing and federated database approaches. Prior to Oracle9*i* RAC, server farms and clustered configurations treated their respective data caches as local resources, requiring constant synchronization of those caches to ensure system integrity. This approach came with a costly performance penalty and resulted in limited scalability. With Oracle9*i* RAC and today's ubiquitous low-latency, high-speed interconnect technology, the caches of every node in the cluster are "fused" across the cluster, thus the label: Cache Fusion.

"While Oracle Parallel Server used the disk subsystem to synchronize activities between the different database

Electronic Arts' chief information officer Marc West (left) and lead architect for RAC/Intel/Linux design Mark Rizzo (right).

SNAPSHOT

ELECTRONIC ARTS
www.ea.com
Oracle Software:
Oracle Database
Oracle9*i* RAC
Oracle9*i* Application Server

instances on the different cluster nodes, the Cache Fusion architecture creates a single virtual cache across all the nodes of the cluster that satisfies all database requests from an application," Demel explains. "The resulting single view of all buffer caches minimizes disk I/O, because any database request can be served by any node in the cluster. This is the key to the transparent scalability, since cache coherency is maintained even when multiple nodes are involved."

A BOOST FOR E-COMMERCE

Oracle9*i* RAC technology is ideal for highly available e-commerce sites, such as the enormously successful Sales Operational Data Store (ODS) that Dell Computer customers use. Based on the concept of selling directly to customers, Dell offers custom-configured, built-to-order systems at competitive prices. By eliminating retailers, Dell can introduce new technology more quickly than its competitors and turn over inventory an average of once every four days.

To make this model work, sales representatives must carefully coordinate orders with a network of suppliers to ensure that products are produced and delivered on time. Dell's operation relies heavily on the Sales ODS, which gives sales representatives real-time access to customer information and up-to-the-minute order-status information.

Dell's Sales ODS originally ran on a four-node symmetric mulitprocessing (SMP) system from Sun Microsystems that was expensive to maintain and difficult to expand. Any problems with the system would result in significant downtime—with a potentially serious impact on Dell customer service. "Every server has an expected mean time between failure, so after you get beyond a certain point in the life of a machine, you're living on borrowed time," says Charlie McMurtry, vice president of Dell IT.

"Traditional SMP systems use redundant hardware architectures to avoid single points of failure and ensure fault resilience," says McMurtry. "You can add additional CPUs to the existing cabinets, to a point, or add additional servers for redundancy. But this approach is very costly, and you'll have all your eggs in one basket," he adds. "If your system goes down, the business essentially stops."

Oracle9*i* RAC technology, by contrast, enables Dell to run a single database across a cluster of Linux servers that incorporate high-performance Intel processors. If one server in the cluster fails, the other servers will take over automatically and instantaneously, which means that the overall cluster provides very high levels of availability—at a comparatively low cost.

GRACEFUL GROWTH CURVE

Furthermore, because Oracle9*i* RAC supports asymmetric clusters, customers no longer need to have the exact same servers in the cluster. They can utilize new server technology as it becomes available, and instead of paying upfront for unused capacity, they can add processing power one node at a time. This type of infrastructure not only simplifies capacity planning but also enables an organization to take advantage of the continually decreasing cost of computer hardware, since businesses don't have to plan for additional capacity years in advance. As traffic increases, system administrators simply add more hardware to an existing pool, allowing a business to tailor the hardware configuration to any projected growth pattern. When a new node is introduced, Oracle9*i* RAC identifies it and

Upwardly Mobile Linux: EA Bets the Worlds on It

For years, IT shops have stretched their IT dollars with Linux deployments. Today, thanks to high-availability clustering solutions such as Oracle9*i* Real Application Clusters, they are stretching their MIPs as well.

For years, the Linux operating system (OS) has been adopted for niche applications and to run small business systems. Today, Linux is starting to play a major role in enterprise information systems as well. According to industry analyst firm International Data Corp. (IDC), Linux is the fastest growing OS platform and is expected to grow 174 percent to US$5.9 billion by 2006.

New products and services from Oracle are having a major impact on Linux adoption in the enterprise—improving cluster scalability, availability, security, and customer support. "Oracle is among the first vendors to really focus on asynchronous I/O and other optimizations to improve database performance," says Stacey Quandt, an analyst with Giga Information Group who follows the Linux market closely.

Oracle is working with Red Hat and other Linux vendors to ensure that Oracle products and the Linux kernel are optimally tuned and configured to the underlying hardware. According to Quandt, their efforts have motivated a growing number of companies to choose Linux as a platform for database management. "These companies use Linux because they want independence from one platform provider and the benefits that come from building an open-source community within their own companies," she says. "In the context of the current economic climate, with many companies facing tight IT budgets, Linux is an appealing technology."

JOINING THE LINUX STAMPEDE

Since its release in June 2002, more than 68,000 copies of Oracle9*i* Database Release 2 on Linux have been downloaded from OTN, representing a 12-fold increase over the previous year. Wim Coekaerts, a principal member of the Oracle technical staff, credits this rapid adoption to the decisive improvements that Oracle and Red Hat have made to the Linux kernel to improve its support for database processing. "We've added the ability to process database requests asynchronously in the I/O subsystem, support for very large memory capabilities, and the ability to run Linux applications on the 64-bit Intel Itanium processor, to name a few," he says.

These technical improvements motivated Dell Computer to migrate its Oracle-based order status management system from UNIX to Linux. Dell is now using the system to respond to 125,000 order-status transactions per hour—at a lower cost (see "A Boost for E-Commerce").

Electronic Arts (EA) had a similar experience. "We deployed Oracle9*i* RAC on an Intel-based Linux platform for about half the cost of a Sun Solaris configuration," reports Marc West, chief information officer at EA.

EA's online gaming service is the fourth largest internet site in the world. About 15 million active players enjoy EA's Persistent State Worlds such as Sims, which recently overtook Myst as the most successful PC game title ever. The upcoming debut of The Sims Online motivated the shift to the Oracle/Linux platform. "The Sims Online is a massively parallel, multiplayer world that is designed to support millions of players," says West. "The database is an integral part of the game design."

According to Mark Rizzo, lead architect for the RAC/Intel/Linux design, the Sims Online will present more demanding requirements, much more complexity, and many more pieces of information to manage than ever before. "We estimate we're going to have better than 250,000 concurrent players running rich sessions, each with different data object sizes," Rizzo predicts. "That's hard to manage because there is no order, nothing is uniform, and it has to scale by a million. Oracle9*i* lets us scale nonuniform object sizes—including all calling and linking—with very high performance."

ORACLE ON CALL

Charles Garry, senior program director of server infrastructure strategies at META Group, believes Oracle has made some bold moves in its support for Linux, particularly its willingness to take Linux support calls. "That's huge for companies that might be hesitant to deploy major systems on the Linux platform; they know they will be backed by a multibillion-dollar company like Oracle," he says.

West concedes that Oracle's willingness to field Linux support calls made a difference in his decision to adopt a relatively new operating system, although EA's primary motivation was to boost clustering performance and reduce cost. "We had an existing architecture but wanted a better way to support a large scale launch while still meeting our price performance objectives," he says. "Oracle and Intel have been really good partners—they are focused on helping us achieve a successful implementation."

Quandt warns that although Linux can improve efficiency and drive down costs, it is not as vertically scalable as high-end UNIX systems. For this reason, she says, users need to understand the performance requirements of their IT infrastructure prior to evaluating Linux. "Most Linux customers are running either two-way or four-way systems, in contrast to a UNIX system that can scale up to 128-way," she says. With Oracle9*i* RAC, however, customers can scale out horizontally instead of vertically, using low-cost commodity servers running Linux.

And while some customers have concerns about Linux reliability, META Group's Garry says that is becoming less and less of a problem as the OS matures. "One of the driving issues they had with Linux is reliability, but technologies such as Oracle9*i* RAC go a long way in terms of allaying their fears," Garry says. "The move toward lower-cost servers is going to take some time, but I think that's when you'll see a real sweet spot emerge for Oracle9*i* RAC technology."

According to EA's West, those days are already here. "Oracle9*i* RAC works," he concludes. "We're pleased with the RAC software, and we are considering using it and Linux on other platforms to run core systems in the future."

dynamically redistributes the workload to take advantage of the added computing resources. Since the physical nodes run independently, the failure of one or more nodes does not affect other nodes in the cluster.

"It used to be that if you wanted to expand a database beyond a single machine—or if your database demands went beyond a single machine—you needed to make application-level changes," Demel emphasizes. "Not so with Oracle9*i* RAC. Oracle customers now have a scalable solution that addresses how different machines coordinate with each other to maintain cache coherency."

THE SUITE SPOT FOR CLUSTERS

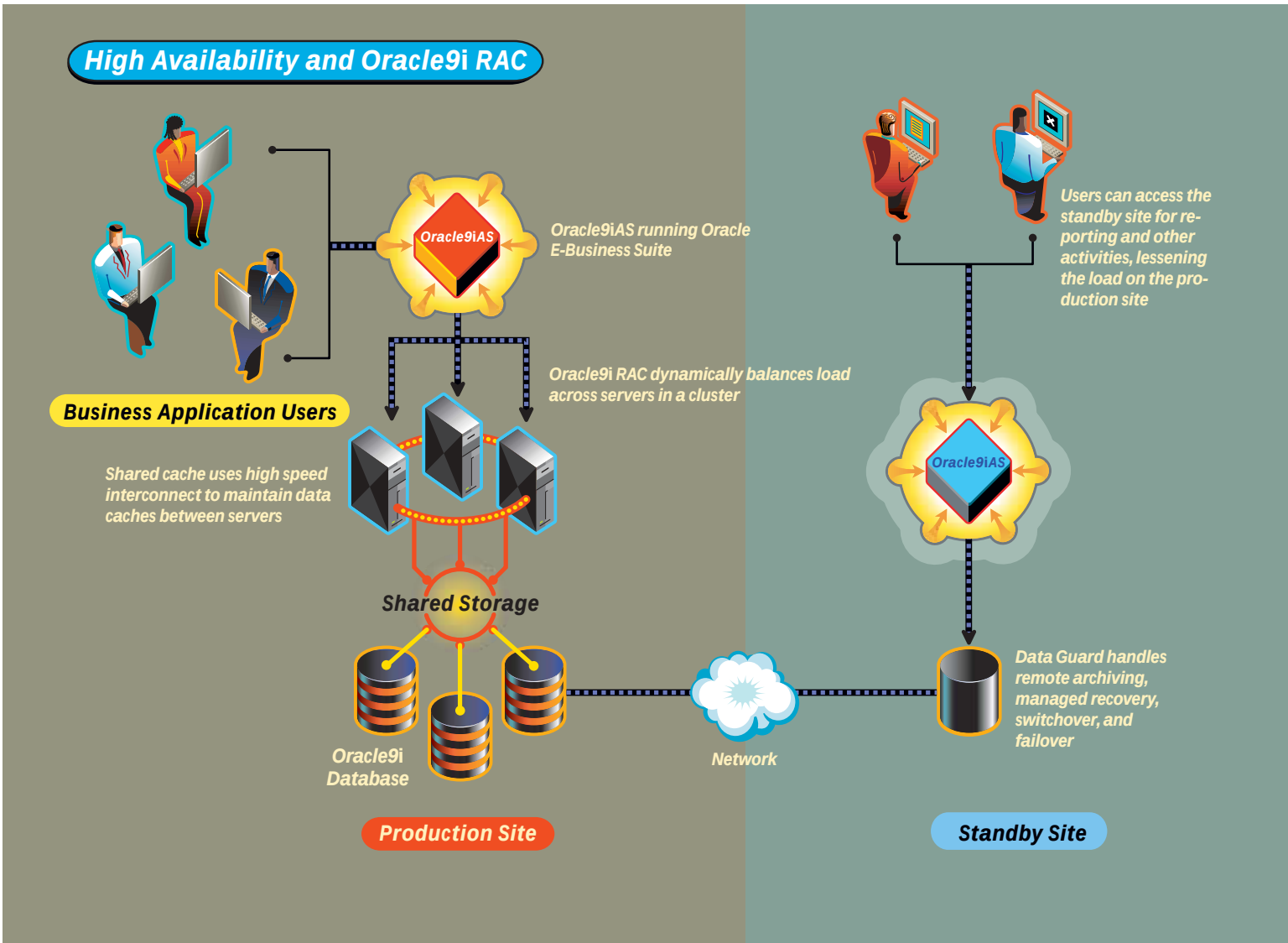
In addition to powering consolidated business-critical applications such as the Dell Sales ODS, Oracle9*i* RAC can be configured to support the various modules of an application suite. Each server in a cluster provides built in redundancy for other servers in a cluster.

META Group's Garry expands the point. "In the past, to ensure a reliable infrastructure, you had to have a lot of redundancy, which meant deploying extra equipment that you don't really use unless you have an outage," he says. "Today's companies are looking to have their cake and eat it too. Deploying a solution like Oracle9*i* RAC means they are covered for unplanned outages yet they are building an infrastructure that has value day-to-day."

For example, a company implementing the Oracle E-Business Suite could host individual modules—such as Payroll, General Ledger, and Receivables—on distinct servers and yet use Oracle9*i* RAC as a way of gaining HA for the suite as a whole. "Each server has an active or primary function, yet one can take up the slack for another if the first one goes down," says Garry. "And because the cache is warmed, failover will take only a few seconds."

Garry believes this type of shared-disk environment simplifies application maintenance tasks as well. "With Oracle9*i* RAC you can deploy distinct modules on separate machines, yet with hooks to share the same database infrastructure. Because each module has its own memory and a measure of separation, you can bring a module down for maintenance while all other modules remain online."

McMurtry also sees the wisdom in this arrangement.



More Than Database Administration

Introducing Enterprise Manager 4.0, a comprehensive platform for managing IT assets.

When Oracle released Oracle9*i*, one of its primary design goals was to simplify management tasks and enhance the tools available to DBAs. In addition to new packaging and deployment options for Oracle9*i* and Oracle9*i* Application Server, Oracle expanded the toolset in Oracle Enterprise Manager to provide coverage of the entire computing platform. Oracle continues to enhance Oracle9*i* management capabilities, culminating in the imminent release of Enterprise Manager (EM) 4.0.

"Although the cost of hardware and software continues to drop, the relatively fixed costs of human administrators loom proportionately larger," notes Martin Peña, director of product management for Oracle's Enterprise Manager group. "With an eye toward more-effective systems management, we've introduced a whole new set of features—both within Enterprise Manager and within Oracle9*i* Database itself."

Peña describes Oracle Enterprise Manager as a state-of-the-art control panel for fixing any Oracle system. "There are a lot of different components it can monitor, administer, and diagnose—not just for Oracle software but for the surrounding infrastructure on which that software is based."

One of EM 4.0's many monitorring screens.

A HOLISTIC APPROACH

Oracle's objective with EM 4.0 is to allow administrators to manage the complete application lifecycle—development and deployment, change management and configuration, day-to-day administration, and diagnostics and monitoring. Using a centralized repository and Oracle Portal, EM 4.0 provides customizable HTML interfaces to supply management information to anyone with an interest in how systems are running. "It's not just DBAs and IT managers who are interested in application performance," suggests Sushil Kumar, senior group manager for database development and product management at Oracle. "As databases and Web-based applications grow in importance, the audience for these applications grows as well."

IT professionals require vastly different levels of detail than corporate executives, yet all of them are interested in knowing how their systems are running on a day-to-day basis. EM 4.0 uses portlet technology to create predefined views of the management operation that can be customized for the different audiences and inserted into a Web page of any portal application. For example, DBAs might want information about I/O statistics and cache ratios, while managers are concerned with overall service levels.

At the same time, Enterprise Manager is becoming more comprehensive, enabling not just database monitoring but application-performance monitoring—right down to the end-user level. "Whenever you click on a button to request data from an application, Enterprise Manager can track how much time you spent in the database, in the application server, even in the different network components," Peña explains. "This helps organizations track realistic service levels and determine where the sources of bottlenecks may be."

EM 4.0 portlets can reveal an aggregate view of the various IT resources that make up each application. This enables users to define a logical view of the application environment, irrespective of the physical components it includes. For example, if a financial application consists of a Sun Solaris host, two Oracle databases, and three instances of Oracle9*i* Application Server, an administrator can define an aggregate target that compiles information from these particular components so he or she can monitor the application as a whole. "With DBAs enabled to observe the performance of a given stack of products, they can manage service levels on the appropriate IT resources," Peña adds. These management activities can be conducted from any location, thanks to a new HTML-based management console.

FROM BACK ROOM TO BOARDROOM

At a macro level, Enterprise Manager incorporates built-in intelligence to simplify the process of configuring and managing database resources. The trend is to gradually elevate the management interface so that administrators can express their needs in business terms rather than technical ones. This enables less-technical users to track application performance and, in some cases, to make changes to the infrastructure.

For example, if a certain application can't be down for more than one minute, DBAs formerly had to translate that business requirement into a set of complex parameters and configuration settings. Now they can simply say, "set it so it will come back up in one minute," and all of the cache levels and other settings will be adjusted automatically. Similarly, instead of manually defining rollback segments, DBAs simply tell the database how much space they would like to use to undo data, and the database figures out the rest.

ORACLE9*i* RAC ADMINISTRATION

According to Peña, proper administration and maintenance of clustering systems is one of the biggest challenges in the mid-market today. Now that organizations are depending heavily on e-business systems, there is a huge revenue impact associated with these systems going down or running poorly. Enterprise Manager can monitor a wide range of facilities in Oracle9*i* Real Application Clusters (RAC) environments, including virtual IP addresses, processes bound to local nodes, database instances, service attractors, service members, and so forth. "DBAs can treat cluster resources as basic management units or 'composite resources,' automating their start, stop, and monitoring as well as failover, relocation, and restart," he says.

"By liberating DBAs from mundane management tasks, we are allowing them to become more strategic players in their business," Peña concludes.

“Oracle9*i* RAC gives us a high degree of flexible scalability for the long run.”

—Charlie McMurtry

“This was our first deployment of Oracle9*i* Real Application Clusters, so we intentionally picked a key system within Dell to put the technology to the test,” he explains. “Now we're in a good position to put this concept to work in other parts of our operation.”

ELIMINATING DOWNTIME

Dell made the switch to the new clustered platform over a weekend in May 2002, incurring less than an hour of downtime during the transition. The Sales ODS and its supporting database now run on two Dell PowerEdge 6450 servers with four Intel Pentium III Xeon processors in each server. Downtime due to planned maintenance has been largely eliminated. “You can work on the system one node at a time,” says McMurtry. “If we need to flash BIOS across the board because of some newly discovered issue that could affect our production system, we can take one node, flash the BIOS, and reinsert the node back into the cluster. Through it all, the cluster will continue to run, and there will be no impact on our production capability.”

Dell's four-node cluster cost approximately 25 percent less than a comparable UNIX SMP solution. And with the use of industry-standard servers and a tested configuration, future upgrades will be simpler and less expensive as well. “If we decide down the road that we need more power to support growth in the business, we can just add more processors or more servers, without having to shut down the whole production system,” McMurtry explains. “Oracle9*i* RAC gives us a high degree of flexible scalability for the long run.”

LAYING A FOUNDATION FOR WEB SERVICES

Some visionaries see technologies such as Oracle9*i* RAC playing a key role in the adoption of Web services—a complementary set of standards for implementing distributed information systems. Today's software developers are adopting these standards to make it easier for applications to request information from each other, even when those applications reside on different types of platforms or are separated by wide-area networks.

“Once you have composite or virtual applications that are bound by Web services, each application must take on the availability characteristics of the most highly available application in the set,” Garry theorizes. “When that happens, there's going to be even more interest in highly available computer solutions such as Oracle9*i* RAC.”

Garry believes the industry will begin to see Web-services-aware databases deployed in a virtual format. “Instead of having five different customer tables for five different applications, it might be possible to have one customer table that serves them all,” he says. “Highly available clustered systems are going to become increasingly important in this context.”

Indeed, as people in all types of businesses come to depend on their computer systems almost as much as on the electricity that powers them, they will be less and less tolerant of outages of any kind. Keeping applications online and databases available is an essential aspect of doing business. Thanks to Oracle9*i* Real Application Clusters, building this level of availability for Oracle database applications is easier than ever before. ■

David Baum (david@dbaumcomm.com) is a freelance writer specializing in business and technology.

nextSTEPS

TEST DRIVE Oracle9*i* Application Server
In OracleWorld's DEMOgrounds

ATTEND Oracle9*i* RAC, Enterprise Manager, and Linux sessions at OracleWorld

Managing Your Enterprise, Using Enterprise Manager, Martin Peña (Oracle): Monday November 11, 1:00 p.m., Moscone Room 134

Why Oracle Built an Oracle Cluster File System on Linux, Wim Coekaerts (Oracle): Monday November 11, 1:00 p.m., Marriott Salon 7

Oracle9*i* RAC: The Present, the Future, but Not Science Fiction, Angelo Pruscino (Oracle): Monday November 11, 1:00 p.m., Moscone Room 103

Using the New Enterprise Manager to Manage the Oracle Environment, Ana Hernandez and Nicole Haba (Oracle): Tuesday November 12, 11:00 a.m., Moscone Room 134

Oracle9*i* RAC on Linux in a Production Environment, Mark Clark (Merrill Lynch): Tuesday November 12, 11:00 a.m., Marriott Salon 8

Oracle's Road Map for Linux, Jamshed Patel (Oracle): Wednesday November 13, 8:30 a.m., Moscone Room 103

Red Hat Linux Road Map for Oracle Today and in the Future, Larry Woodman (Red Hat): Wednesday November 13, 4 p.m., 4:00 p.m., Marriott Salon 7

Achieving Scalability, Performance, and Availability on Linux with Oracle9i Release 2 - Oracle9*i* RAC, Grant McAlister (Amazon.com): Wednesday November 13, 5:30 p.m., Moscone Room 309

Oracle9*i*AS/Enterprise Manager: Best Practices for Managing Oracle9*i*AS and Clusters, Valerie Kane (Oracle): Wednesday November 13, 5:30 p.m., Marriott Salon 2

DISCOVER what Oracle Consulting offers
oracle.com/consulting/offerings/platform/index.html?hot_availability.html

TAKE AN INTERNET SEMINAR on high availability
oracle.com/iseminars/index.html?ondemand_archive.html

LEARN about Linux from Oracle Technology Network
otn.oracle.com/tech/linux/content.html