How Cisco IT Virtualized Oracle RAC Databases on Unified Computing System with VMware

Nearly 300 nonproduction and production databases are virtualized, increasing resiliency while lowering costs.

EXECUTIVE SUMMARY
CHALLENGE
• Increase business agility by accelerating provisioning of new Oracle and Oracle RAC databases • Increase resiliency • Lower infrastructure costs for Oracle databases
SOLUTION
• Virtualized production and nonproduction Oracle databases on Cisco Unified Computing System with VMware • Automated backup processes
RESULTS
• Accelerated provisioning from 12 weeks to 3 days, and as little as 3 hours • Lowered TCO by approximately 45-50 percent • Improved recovery time from several hours to a few minutes
NEXT STEPS
• Virtualize 95 percent of Oracle databases

Challenge

The Cisco workforce accesses approximately 1300 Oracle databases for business processes ranging from order processing to customer care. "While Oracle database servers are a small fraction of the total server footprint at Cisco, they are among the largest and most critical, and they incur significant infrastructure and maintenance costs," says Malathi Pinnamaneni, Cisco IT architect.

Cisco IT has been systematically virtualizing applications and had already virtualized the Oracle web and application tiers. The next step in the journey would be virtualizing Oracle databases.

"Originally we focused on standalone, nonproduction databases," says Paul Wiltsey, Cisco IT engineer. "When backup solutions became more sophisticated and we gained more experience with virtualization, we were ready to virtualize our Oracle RAC databases."

Goals for virtualizing the Cisco IT Oracle environment included:

- Increasing business agility by accelerating provisioning: Fulfilling requests for new Oracle databases on standalone servers often took 12 weeks or longer. To increase business agility, Cisco IT wanted the ability to fulfill urgent requests in a few days.

- Increasing resiliency: Cisco uses Oracle databases for revenue-generating activities, including customer support. "Previously, if a server failed, restoring the database on another server generally took several hours, and people were concerned that virtualizing production databases would increase downtime," says Todd Glenn, Cisco IT service manager. "Today, database failover is actually faster in a virtualized environment, improving the user experience and reducing lost productivity."

- Lowering costs: The fewer the servers, the lower the space, power, cooling, cabling, and management costs. When Cisco hosted its Oracle databases on bare-metal servers, utilization was only about 25 percent. Other sources of high costs included high availability requirements and the need for specialized servers requiring scarce management skills. These servers also required expensive, dedicated network interconnects for clustered database hosts.

Solution

After carefully evaluating the risks, Cisco IT management decided to virtualize Oracle applications and databases on VMware even though standard industry practice was to use physical servers. "Without the vision of senior management, virtualization of the Oracle environment would not have proceeded as quickly as it did," says Jag Kahlon, Cisco IT architect. "Their leadership accelerated benefits such as faster provisioning, lower TCO, and improved disaster recovery."

Virtualizing the Oracle environment simplified infrastructure and processes. "Because VMs are relatively independent of hardware, we can dynamically adjust compute and bandwidth resources without physically visiting the data center," Wiltsey says.

Design Principles

A guiding principle for Cisco IT was to use the existing shared infrastructure, Cisco Unified Computing System™ (Cisco UCS®) and Cisco Nexus® switches, instead of building special clusters that would require separate support and disaster recovery processes. Cisco UCS service profiles further simplify the environment by enabling IT staff to accurately configure server identity with a few clicks.

The other guiding principle was not introducing new technology. Therefore, Cisco IT decided to continue using VMware datastores. The other option considered, raw device mapping (RDM), would increase throughput by mapping logical unit numbers (LUNs) directly to virtual machines. "We decided to not use RDM, because it would prevent us from freely moving data on the backend," says Nagarajan R, Cisco IT architect. By using VMware datastores instead, Cisco IT was able to treat the virtualized Oracle database environment as part of the same Cisco UCS cluster used for other enterprise applications. "We used CPU and memory reservations to control oversubscription," says Nagarajan. "The more reservation, the less contention in the environment."

Selecting Databases to Virtualize

When Cisco IT initially began virtualizing Oracle databases, the database team targeted only standalone, nonproduction Oracle databases that were under 2 TB, with fewer than 200 transactions per second (TPS) and bandwidth requirements under 100 Mbps. "Now that hardware and software vendors provide better support for Oracle 11g RAC on VMware, we can support 1000–2000 TPS, a 400 percent increase," says Pinnamaneni.

Virtualization Policy

Cisco IT is virtualizing all standalone Oracle databases in all environments: development, stage/test, load/test performance, disaster recovery, and production. Oracle RAC databases are virtualized if the VM sizing and I/O characteristics in Table 1 can support the workload. If not, the databases are hosted on bare-metal servers in the same Cisco® UCS. In nonproduction environments, multiple databases are consolidated onto a single virtual machine when possible. The methodology to arrive at this table is described in the section entitled "Workload Testing Methodology and Results" later in this case study.

Table 1.   Standard VM Profiles: Criteria for Determining Whether a Database Qualifies for Virtualization

Database	Number of Nodes	CPU Cores and RAM	Server Performance (TPS)	IOPS	Bandwidth (Mbps)
Oracle 11g	Standalone VM	4 x 16	536	20,746	162
	2-node RAC VMs	4 x 16	932	30,418	238
	4-node RAC VMs	4 x 16	1577	48,421	378
Oracle RAC 11g – D-NFS	Standalone VM	4 x 16	697	33,126	259
	2-node RAC VMs	4 x 16	1292	49,528	387
	4-node RAC VMs	4 x 16	2077	74,259	580

Figure 1 illustrates the deployment topology at Cisco, and Table 2 lists solution components.

Figure 1.   Virtualized Oracle RAC Environment
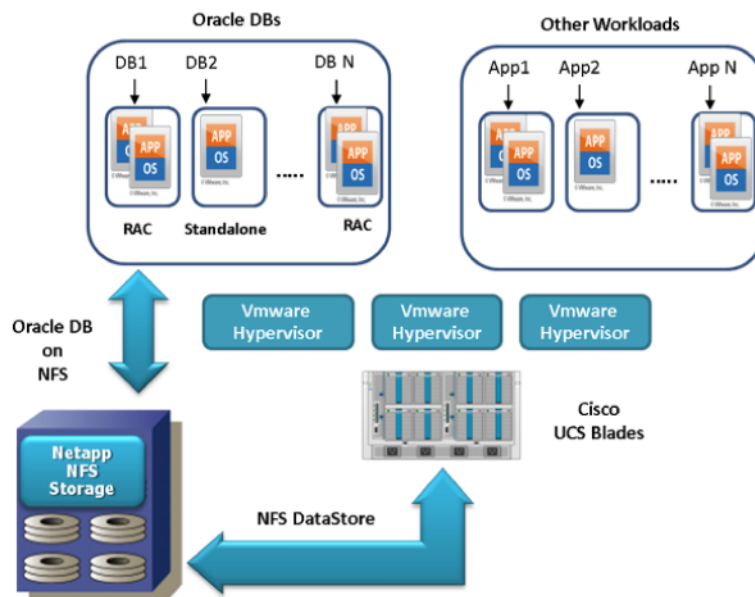
Table 2.   Oracle RAC Databases are Hosted on Standard Cisco Unified Data Center Environment

Server	Virtual Switch	VMware Environment	Storage
• 40 Cisco B200 M2 Servers with 12 Cores and 96 GB • Connected to Cisco Nexus 7000 Switches	• Cisco Nexus 1000V with three interfaces: public, private, and monitoring (heartbeat)	• VMware Vsphere versions 4 and 5 • 20-node clusters • 10 TB NFS datastores • Largest configuration: 8 cores with 64 GB RAM	• NetApp FAS 6280 • ONTAP version 8 • 10 GB Link Aggregation Control Protocol (LACP) Port Channel to network

Cisco IT configures Cisco UCS blade servers to support both the Oracle database and the underlying Oracle Grid

Infrastructure, which is the Oracle software that provides volume management, file system, and automatic restart capabilities. The standard VM configuration includes at least 2.5 GB of RAM for Oracle Grid Infrastructure plus swap space, which depends on available RAM (Table 3). Cisco IT also uses resource reservations to guarantee resources to specific service tiers.

Table 3.   Swap Space Calculation

Available RAM	Swap Space
2–8 GB	2 x RAM
8–32 GB	1.5 x RAM
32+ GB	32 GB

An Oracle RAC cluster can contain a combination of virtual and physical servers. Cisco IT currently does not use a combination, but might in the future. "If a cluster provides sufficient capacity for all periods other than year-end or month-end, we could quickly add virtual machines, or temporarily add another Cisco UCS blade server," says Nagarajan.

File System Standards

Cisco IT followed Oracle recommendations for mount options (support note 359515.1). That is, Oracle database file systems, including data, redo, and archive log files, are mounted to VMs by way of D-NFS, a new feature introduced with Oracle 11g. All non-Oracle database file systems are mounted locally as a VMware Disk (VMDK) file.

Based on the size of a given Oracle database, including the standard LUN and store, Cisco IT provisions dedicated database volumes on NFS. "This is the only situation where we dedicate resources," says Nagarajan. "With NFS, a database can be mounted on a physical and virtual server. If we encounter a surprise during migration, we can easily back out from the virtual server to the physical server."

Migrating Existing Databases to VMware

To migrate physical databases to virtual machines, Cisco IT first installed database software on the virtual machine and then migrated the data from the physical server. "This approach minimized disruption for our application teams, because we were able to upgrade the database, migrate to the new data center, and consolidate storage at the same time instead of sequentially," Pinnamaneni says.

Failover and Disaster Recovery

"When Cisco's Oracle databases were hosted on physical servers, failover took several hours," says Sunil Rajesh, Cisco IT database administrator, "Now failover takes just a few minutes." Cisco IT built several layers of resiliency into the solution:

- Each Oracle RAC cluster contains from two to eight nodes. If one node is down, Cisco users automatically connect to one of the other nodes without having to take any action.

- Each Cisco UCS has multiple blade servers. If one server develops issues, Cisco IT can quickly move its virtual machines to any other available blade server, with a few clicks.

- Cisco IT enabled the Oracle Data Guard Fast-Start Failover feature. If all the nodes in Oracle RAC cluster

fail, services automatically fail over to another location in the metro virtual data center (MVDC).

- In the event of a regional disaster that takes down production, Cisco IT can use VMware Site Recovery Manager (SRM) in conjunction with Oracle Data Guard for push-button failover to the disaster recovery data center in Research Triangle Park, North Carolina.

Backup and Cloning

To simplify IT, the team wanted to use just one backup application, although many would work in the environment. Therefore, Cisco IT decided to continue using Oracle Recovery Manager (RMAN), a built-in component of the Oracle database server that provides block-level corruption detection during backup and recovery. RMAN is integrated with Symantec NetBackup, which Cisco IT uses for tape backups.

Cisco IT developed a custom tool to automate much of the end-to-end backup process. The tool dynamically generates the RMAN script, copies the script to primary databases, and executes the NetBackup policies. Cisco IT defined two policies for each database, one for data files and the other for archive log files. Each backup policy has its own schedule, and archive logs are backed up more frequently than the database. The script, which runs on the NetBackup server, directs RMAN on the primary database host to back up the database, controlfile, and parameter file to tape. Cisco IT can tweak the script for different databases, specifying full or incremental backup, direct backup or proxy copy backup, and throttle parameters to optimize backup speed.

To make sure that development and test teams work with up-to-date data and data structures, Cisco IT uses RMAN to regularly refresh the nonproduction database with production data. To keep these large file transfers, which might be 10 TB or larger, from adversely affecting service levels for time-sensitive and latency-sensitive traffic, the OS team enabled file transfer over a nonstandard port. The network team forces traffic over that port to a lower quality of service (QoS) level so that production traffic receives priority.

Workload Testing Methodology and Results

During the proof of concept, Cisco IT measured CPU load using an open-source tool called Swingbench, and measured I/O throughput using an internally developed storage stress test. Tests were conducted on a four-node, 2.5 TB Oracle RAC database. Each VM was configured with four cores and 16 GB RAM. Cisco IT conducted multiple test runs to validate that results were consistent.

Test criteria included:

- Normal CPU utilization over 90 percent and run queues with more than 30 processes
- High disk I/O and network I/O stress, but not high enough to interfere with CPU
- I/O wait not exceeding 30-40 percent
- Recovery from destructive failure tests

Note that the following results are specific to Cisco IT's environment, and other organizations may experience different results.

Test Scenario 1: CPU Utilization and TPS

Cisco IT executed 300 sessions of the Swingbench Calling Circle benchmark tool for 30 minutes, comparing TPS results using NFS and D-NFS. TPS performance for NFS and D-NFS was similar, and scaled linearly with the number of Oracle RAC nodes (Figure 2).

Figure 2. CPU Utilization for NFS (Left) and D-NFS (Right); Near Linear Performance Increases
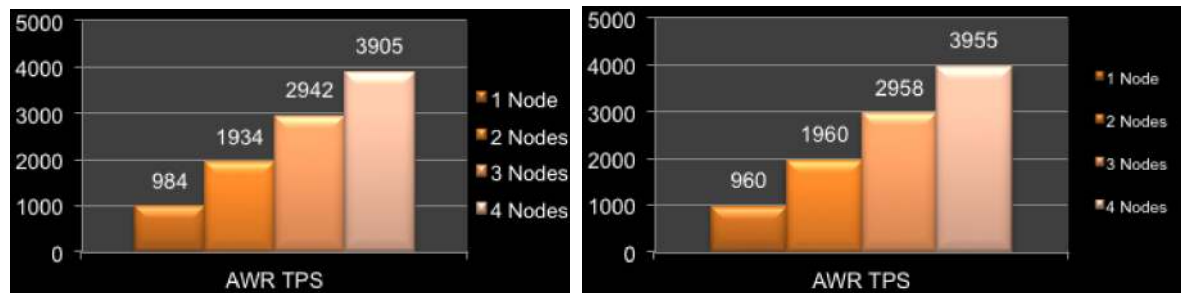
Test Scenario 2: Maximum I/O Throughput

To test I/O utilization, Cisco IT conducted 10 sessions using the internal storage tool as well as 16 users' random select statements for 30 minutes. The I/O throughput measures were high for NFS and even higher for D-NFS (Figure 3).

Figure 3. I/O Results for NFS (Left) and D-NFS (Right)
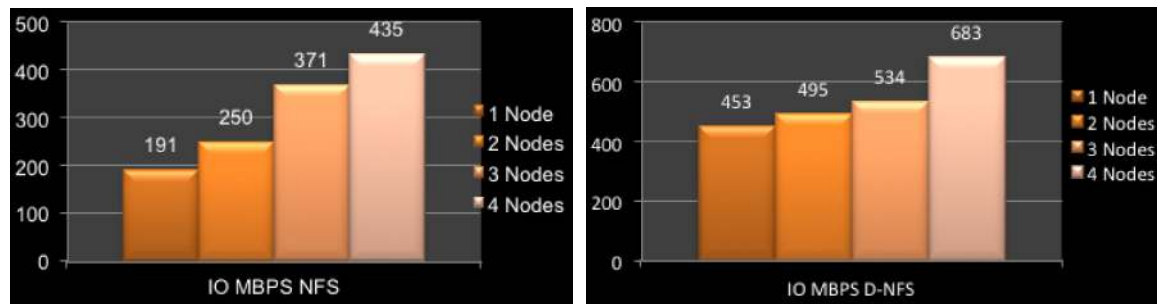
Scenario 3: Combined CPU and I/O Utilization Testing

Finally, to simulate real-world workload patterns, Cisco IT executed both tests in parallel, for 30 minutes. Again, D-NFS delivered higher transaction processing and bandwidth performance than NFS (Figure 4).

Figure 4. D-NFS Performed Better than NFS in Cisco Environment
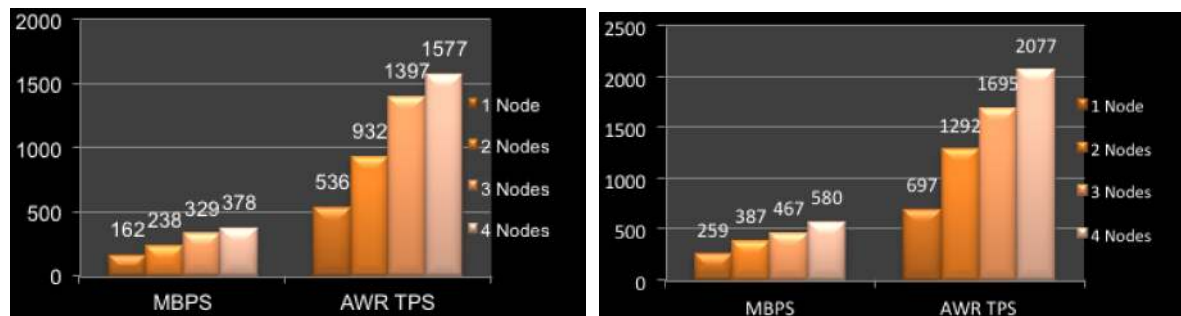
Results

"When we used physical servers, provisioning a new database was time-consuming, complex, and expensive," says Adwait Samant, a senior manager with Cisco IT who oversees virtualization for Oracle databases. "Virtualizing our Oracle databases enabled Cisco IT to quickly create test and development environments with the desired service-level agreements, helping us respond quickly to changing business needs."

Increased Business Agility Through Faster Provisioning

Cisco IT now routinely fulfills requests for new Oracle database servers in three days, and in as little as three hours for urgent requests. Cisco business users request virtual servers using the CITEIS online catalogue. Whether requests are approved automatically or manually depends on the required service level. After receiving their servers, business users can self-provision additional storage, compute, and memory resources using CITEIS.

New VLAN requests are also fulfilled more quickly, a benefit of the Cisco Nexus 1000V Virtual Switch. Previously, a database administrator (DBA) who needed a new VLAN had to open a ticket with three separate Cisco IT teams: servers, VMware, and networking. "The Nexus 1000V eliminates the need for the DBA to coordinate new or changed VLANs with the VMware team," says Wiltsey. "That's one less change request window to consider, which can make the VLAN available days or even weeks sooner."

Increased Resiliency

"Virtualizing the Oracle environment on the Cisco UCS means that failure of a single node no longer disrupts critical business processes that depend on access to Oracle databases," says Anil Nileshwar, director of global infrastructure services for Cisco IT. If a server fails, Cisco IT can use Cisco UCS Manager server profiles to provision any other available Cisco UCS blade server in any chassis in a few minutes.

During the proof of concept, Cisco IT measured the time to recover from various destructive events with vMotion. The tests simulated host failure, loss of votedisk on a node, deletion of a voting disk, and reboot of one node with a full load. "Test results for all events met expectations for Oracle RAC databases," says Rajesh.

Lower Cost Scalability

When Cisco IT hosted the Oracle environment on physical servers, adding capacity required purchasing more hardware. "We used to overprovision, to avoid delays in procuring new hardware when needed," Glenn says. "Now, we can right-size the configuration for a particular database. If the workload later increases, we can either add new servers or migrate the workload to a larger host." For example, use of certain Oracle databases spikes at the end of each quarter. Rather than sizing servers to accommodate peak load, Cisco IT can now size them for moderate load, and then temporarily assign more resources at quarter-end. When these resources are not needed, they are available for other applications.

Lower TCO for Oracle Databases

Cisco IT estimates that increasing the density of Oracle databases lowered total cost of ownership by 45-50 percent. Factors contributing to these cost savings include infrastructure, licensing, and support.

No Compromise to Performance

The Oracle D-NFS client introduced with Oracle 11g provides better scalability and performance than traditional NFS v3 clients. "For most workloads, we are experiencing near-native performance, similar to the performance we would expect from a physical server," Pinnamaneni says.

Next Steps

Cisco IT is taking advantage of continual improvements in the Cisco Unified Data Center solution to virtualize ever-larger databases. "Larger virtual machines become feasible as the Cisco UCS and Nexus 1000V continue to scale," Wiltsey says. "The next standard will be 20 virtual cores and up to 256 GB RAM."

Cisco IT expects to virtualize 95 percent of its Oracle databases, including business-critical databases, by the end of the 2013 calendar year. "This is a critical milestone in Cisco IT's platform-as-a-service strategy," Pinnamaneni says.

Lessons Learned

Cisco IT offers the following suggestions for organizations planning to virtualize their Oracle database environment.

- Disable the hang-check timer for standalone instances. The hang-check timer is required only for Oracle RAC instances, to monitor the Linux kernel for extended operating system hangs that could affect the reliability of the RAC node and corrupt the database. Before Cisco IT disabled the hang-check timer, VMware ESX host kept rebooting because the vMotion load-balancing process introduced just enough temporary latency to trigger the timer.
- Be prepared to change the Oracle System Global Area (SGA) size. "On bare metal servers, all RAM is available for a single host," says Rajesh. "But when we migrated to virtual machines, high SGA caused host hangs." Cisco IT overcame the issue by resizing SGA to match VM reservation. In the Cisco IT environment, the optimal settings were shmax at 50 percent of RAM and shmall at 75 percent of RAM.

For More Information

How Cisco IT Migrated a 34TB Oracle RAC Database from HP Superdome to Cisco Unified Data Center

How Cisco IT Plans and Executes a Large-Scale Data Center Application Migration

To read additional Cisco IT case studies on a variety of business solutions, visit Cisco on Cisco: Inside Cisco IT www.cisco.com/go/ciscoit.

Note

This publication describes how Cisco has benefited from the deployment of its own products. Many factors may have contributed to the results and benefits described; Cisco does not guarantee comparable results elsewhere.

CISCO PROVIDES THIS PUBLICATION AS IS WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING THE IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Some jurisdictions do not allow disclaimer of express or implied warranties, therefore this disclaimer may not apply to you.

CISCO