Scaling an Oracle RAC System with

Disk-based Backup and Recovery

Data growth is inevitable in most IT environments, and yesterday's cutting-edge system can become today's bottleneck. Tested and validated Dell/Oracle industry-standard database components and Dell/EMC storage platforms can help enterprises expand their IT capacity to accommodate their increasing data. One way to scale is to add storage for disk-based backup and recovery, which can help minimize backup times, speed database recovery, add recovery points, and help meet service-level agreements.

BY MAHMOUD AHMADIAN, UJJWAL RAJBHANDARI, DAVID MAR, AND CHETHAN KUMAR

Related Categories:

Backup

Backup to disk

Business continuity

Clustering

Dell/EMC storage

Disaster recovery

Microsoft Windows Server 2003

Oracle

Red Hat Enterprise Linux

Scalable enterprise

Visit www.dell.com/powersolutions for the complete category index.

Dell best practices strongly recommend deploying tested and validated configurations for databases because of the complexity and business-critical nature of such environments. Administrators can deploy these configurations with confidence that they have been thoroughly tested and stressed.

Although database configurations can be custom-tailored to an enterprise's needs, such configurations are typically unique in their combination of software, hardware, and driver versions. Consequently, they may not have been thoroughly tested or stressed in that particular combination. To alleviate such issues and to provide comprehensive support for the enterprise, Dell defines and tests stacks of database components, helping to identify interoperability issues before they reach the enterprise. These tested and validated Dell™ configurations are defined by a configuration matrix that includes hardware and software components and the process by which to deploy these components (see Figure 1).

Deploying tested and validated Dell/Oracle database configurations can hold several advantages for an enterprise. Such an approach can help eliminate deployment problems and enable system support and fast deployment. A deployment CD is provided with Dell/Oracle configurations to help ensure that drivers, kernel configurations, and Oracle prerequisites are optimally fulfilled for proper implementation. Furthermore, Dell promotes a scalable enterprise strategy, which calls for a phased deployment approach that can scale to meet the needs of any size enterprise. Dell/Oracle configurations use industry-standard components that can easily be upgraded to meet business requirements—providing a predictable foundation on which enterprises can later expand.

Achieving high data availability with Dell/Oracle configurations

Enterprise databases may range from a single-instance database consisting of a few gigabytes of data to multiple

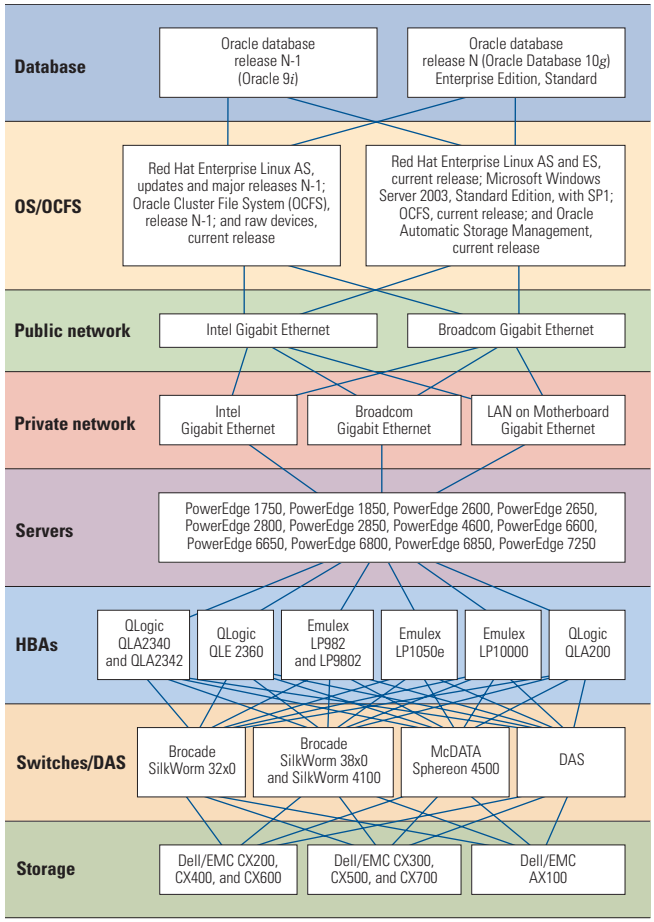

Figure 1. Dell component test matrix

databases storing hundreds of terabytes of data. Regardless of their size, databases require high data availability to support mission-critical deployments. Integrating the Dell/EMC AX100 storage array into an Oracle® automatic disk-based backup and recovery system can enable enterprises to achieve online data recovery at a reasonable cost.

Configuring Oracle's flash recovery area on an economical storage unit such as the AX100 can help ease management through retention policies and enable automated management of backup space, redo logs, and other recovery-related files on disk-based storage. This article describes two example configurations that combine Dell/EMC storage arrays with existing data storage systems to support the Oracle Database 10*g* flashback area, thereby achieving cost-effective online data backup. The first configuration combines Oracle Database 10*g* Standard Edition (SE) Real Application Clusters (RAC), the Microsoft® Windows Server™ 2003 OS, and a Dell/EMC AX100 storage array; the second configuration invokes a larger, enterprise-class system with a Dell/EMC CX300 storage array, Oracle Database 10*g* Enterprise Edition, and the Red Hat® Enterprise Linux® AS 3 Quarterly Update 4 (QU4) OS.

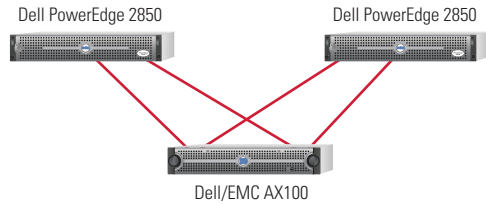

Figure 2. Dell/EMC AX100 in a DAS configuration

Dell/Oracle configuration 1: entry-level database

The Dell/EMC AX100 is designed to provide easy-to-use, low-cost storage in a direct attach storage (DAS) configuration, as shown in Figure 2, or in a storage area network (SAN) configuration. The AX100 is well suited for small workgroups, medium-sized businesses, and branch offices of large corporations. Tested and validated Dell configurations using the AX100 and a Microsoft Windows® OS can be excellent for small and medium-sized databases. Additionally, enterprises starting with these solutions can achieve low cost per unit of storage capacity for a cluster.

For small-to-medium businesses and budget-conscious enterprises, Dell has introduced a tested and validated configuration for Oracle Database 10*g* SE RAC on Windows Server 2003, Standard Edition, with Service Pack 1 (SP1).[1] The configuration uses two Dell PowerEdge™ 2850 servers and the Dell/EMC AX100 (see Figure 3). This configuration was selected for the test environment because it characterizes an entry-level offering wherein cost and ease of use are paramount. Although this configuration represents an entry-level scenario, the existing hardware can be integrated into a more complex, scaled design.

Servers	Two Dell PowerEdge 2850 servers
OS	Microsoft Windows Server 2003, Standard Edition, with SP1
Oracle database software	Oracle Database 10*g* SE, Release 1 base (10.1.0.2) plus the 10.1.0.4 patch set
Fibre Channel switches	None (servers directly attached to storage)
HBAs	Two QLogic QLA2340 Fibre Channel adapters
Storage	Dell/EMC AX100

Figure 3. Components of configuration 1

[1] For more information about this configuration, visit www.dell.com/oracle and click "Dell Supported Configurations." Then select the "Oracle Database 10g" tab, and in the "Microsoft Windows 2003 SP1" section, click "Oracle Database 10*g* Standard Edition with Real Application Clusters."

Dell/Oracle configuration 2: database cluster

Combining Oracle Database 10*g* Enterprise Edition, Red Hat Enterprise Linux AS 3, and the Dell/EMC CX300 storage array can yield a tested and validated configuration suitable for a larger enterprise than that envisioned for the first configuration. For its database, this configuration relies on a Dell/EMC CX300 storage unit (see Figure 4). Designed for slightly larger and more performance-critical databases than those for which the AX100 is intended, the Dell/EMC CX300 is an entry-level RAID storage system with 2 Gbps Fibre Channel host interfaces and capacity for as many as 60 Fibre Channel disk drives. Capable of operating as DAS, in a SAN, or attached to a Dell PowerVault™ network attached storage (NAS) system, the CX300 can provide up to 18 TB of FC2 raw storage capacity or up to 27 TB of Serial ATA (SATA) storage capacity using Dell/EMC 2 GB Disk Array Enclosures.

Tested and validated Dell configurations using the AX100 and a Microsoft Windows OS can be excellent for small and medium-sized databases. Additionally, enterprises starting with these solutions can achieve low cost per unit of storage capacity for a cluster.

Scaling the tested and validated Dell/Oracle configurations

Configurations 1 and 2 differ significantly. Although the end goal involves a similar configuration of Dell/Oracle and Dell/EMC products, the motivation to scale these configurations is dissimilar.

Scaling configuration 1: enhanced performance

Configuration 1 represents an entry-level Oracle database that is easy to deploy and to which storage can be added at a relatively low cost. However, improving Oracle cluster performance requires hardware additions. Consequently, to scale configuration 1, best practices recommend maximizing the use of initial hardware while seeking to improve performance and response times.

To accomplish these requirements, a Dell/EMX CX300 was added to the initial hardware configuration. The database that originally resided on the AX100 was then moved onto the CX300, while the flashback area remained on the AX100. This modification improved database performance because the CX300 provided greater storage speed and cache size compared to the AX100, and the initial hardware could then be used for disk-based recovery features.

If an enterprise currently uses the QLogic QLA200 host bus adapter (HBA) card in a fabric environment, Dell best practices recommend that administrators employ an HBA architecture that can offload tasks from the host node. They should also select a card with higher buffer credit levels than the QLA200 can provide—although that setup was not part of these test designs. Often the AX100 is coupled with QLA200 cards, which are designed for entry-level systems. Although moving to new HBAs requires initial hardware repurposing, replacing the entry-level HBAs can bring worthwhile performance improvement.

Scaling configuration 2: optimized cost

Configuration 2 assumes that a Dell/EMC CX300 storage system has already been integrated into an Oracle cluster. Enterprises can help reduce storage costs by adding an AX100 array to the cluster for disk-based backup and recovery.

Although an extra disk array enclosure (DAE) can be added to the cluster to increase storage capacity, this test design chose the goal of cost optimization. To minimize costs and add disk-based recovery, Dell engineers designed configuration 2 to leverage existing components by keeping the database logical units (LUNs) on the CX300. In contrast, the flashback and recovery area was moved to the AX100. This modification anticipated database performance at levels comparable to those of the initial configuration 2 while leveraging the added AX100 for disk-based recovery features, thereby dedicating disk space and storage processor cycles entirely to the database files.

The cost trade-offs for adding a DAE to the CX300 versus adding an AX100 were simple to analyze:

- Cost of DAE + (number of drives × cost of drives for DAE)
- Cost of Dell/EMC AX100 + (number of drives × cost of SATA drives for AX100)

Configuration 2 represents a scenario in which the addition of an AX100 can provide cost advantages. Because the Oracle flashback

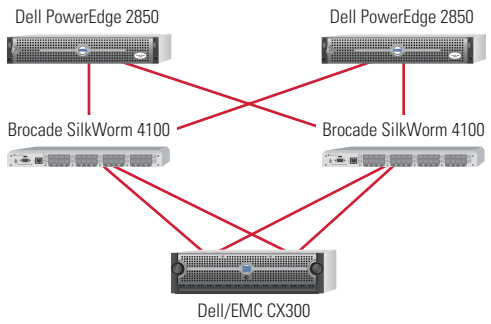

Figure 4. Dell/EMC CX300 in a SAN configuration

Servers	Two Dell PowerEdge 2850 servers
OS	Red Hat Enterprise Linux AS 3 QU4
Oracle database software	Oracle Database 10*g*, Release 1 base (10.1.0.2) plus the 10.1.0.4 patch set
Fibre Channel switches	Two Brocade SilkWorm 4100 switches running at 2 Gbps
HBAs	Two QLogic QLA2340 Fibre Channel adapters per node
Additional software	EMC PowerPath software
Storage	Dell/EMC AX100 and CX300

Figure 5. Components of configuration 2 scaled for cost and performance

recovery and backup functions do not require high random I/O reads and writes, backup and flashback storage can be placed on the AX100, which is slower than the CX300, without performance degradation to the database. Additionally, the cache on the CX300 can be leveraged to greater potential than in its initial configuration because sequential reads and writes do not fill up the cache.

In addition, Oracle enhanced the disk-based backup and recovery capability of its database software in Oracle 10*g*. Oracle added several features to take advantage of disk-system economics. These features allow leveraging additional cost-effective storage to help reduce expensive system outage time and associated costs.

Analyzing the optimal Dell/Oracle configuration

To optimize both cost and performance, a model configuration would involve both an AX100 and a CX300. In such a configuration, the CX300 stores the data files while the AX100 houses the flashback recovery area. To accomplish this, the Dell engineers placed the database disk groups on the CX300 LUNs. To enable flashback, they allocated a flash recovery area on the AX100. In tests performed by Dell engineers in October 2005, the AX100 was exclusively designated for the retention and backup components of the data file images, redo logs, and control files.

The scaled configuration comprised three layers: the server layer with two Dell PowerEdge 2850 servers; the switch layer with two Fibre Channel Brocade SilkWorm 4100 switches running in 2 Gbps mode; and the storage layer with both switches connected to the CX300 and AX100 arrays (see Figure 5).

High-availability design principles

Dell's Best Practices Program for Oracle encourages high-availability design principles. Consequently, the scaled design described in this article used redundant components at each level. Because the initial configuration already employed two PowerEdge 2850 servers, no additional servers were required. The SAN fabric environment used multiple Fibre Channel switches that allowed the use of more than one back-end storage unit while adding redundancy. Additionally, the presence of EMC PowerPath® software and multiple QLogic QLA2340 adapters in each PowerEdge 2850 server helped ensure that, if an HBA failed, alternate paths would be available from each node.

Figure 6 shows a two-node Oracle RAC cluster deployed in a high-availability configuration. Such a configuration should minimally contain the following components:

- Two servers
- Two HBAs per server
- Two Fibre Channel switches
- Multiple storage systems

Similar to a DAS configuration, each node's HBA in this high-availability configuration can access both storage processors; preconfigured zoning allows the HBAs to do so. In addition, with redundant switches, each server can still reach virtual disks if a switch fails because both storage processors are connected to alternate switches.

Configuration 1 versus the optimized scaled implementation

The scalability goals of the first configuration included reuse of initial hardware coupled with increased performance. Because the database was moved to the CX300, performance of the scaled configuration was significantly better than that of the AX100 alone, as test results show.[2] The Dell test team employed a benchmark that was similar to TPC-C from the Transaction Processing Performance Council (TPC), and these benchmark results revealed a significant performance increase for the Oracle cluster after the CX300 was added.[3]

Configuration 2 versus the optimized scaled implementation

In configuration 2, adding an AX100 easily achieved the goal of cost optimization for expanded storage capacity. However, the

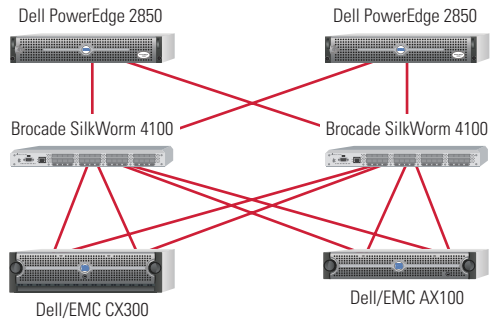

Figure 6. Dell/EMC CX300 and AX100 in a scaled configuration

[2] For more information and a detailed comparative performance analysis of the two systems, please see the associated white paper on this topic at www.dell.com/oracle. Search for "Dell Technology White Papers," and then click on the link. From the list of articles, click on "A Detailed Look into Scaling an Oracle RAC System with Disk-Based Backup and Recovery."

[3] See the associated white paper for test results.

performance of the scaled configuration needed to be comparable to that of the CX300 alone. To verify this, the Dell engineers first tested the CX300 alone and then tested the AX100 and CX300 combined configuration.

For the benchmark test of configuration 2, four LUNs were created in the CX300 and shared between the cluster nodes. A 1 GB LUN was assigned for storing Oracle Cluster Recovery (OCR) and voting-disk information. A 133 GB LUN was used for storing database files. Two 20 GB LUNs were used for flash recovery. The same storage processor was assigned as the default owner of the database and the flash recovery LUNs (to analyze the effect of offloading backup and flash recovery cycles to the AX100). Tests results showed that the CX300 performed relatively well by itself.

After the initial configuration 2 system was benchmarked, it was migrated to the optimized scaled configuration, which included both the AX100 and CX300. Two LUNs (1 GB and 133 GB) were created in the CX300 and two LUNs (20 GB) were created in the AX100 for the cluster nodes. The 1 GB LUN in the CX300 was assigned for storing OCR and voting-disk information; the 133 GB LUN was used for storing database files. The 20 GB LUNs in the AX100 were used for flash recovery.

Next, the Dell testing team subjected the scaled configuration 2 to a TPC-C benchmark–like workload—a series of tests using 200, 400, and 600 users. The tests showed that the AX100 and CX300 combination performed better in some cases than the CX300 alone, and that no serious degradation resulted from introducing the AX100 system to a cluster with a CX300.

The slight performance increase may have been attributable to the CX300 cache. The CX300 divides the cache into two components—one area is designated as a share cache for all LUNs, and a second area is broken into smaller segments wherein each segment is dedicated to a specific LUN. Because the flashback LUN was removed from the CX300, the entire cache was dedicated to the database. Consequently, small performance improvements might have resulted until the SAN cache was filled.

Balancing performance and cost in Dell/Oracle environments

The trade-off between speed and cost-effectiveness of disk-based backup and recovery is a factor that every enterprise must constantly weigh. Dell engineers analyzed two Dell/Oracle tested and validated configurations and scaled them to a common solution configuration. The goal in scaling the first configuration was to increase the solution's transaction rate. The goal in the second configuration—where database transaction rate was acceptable—was to add cost-effective disk-based recovery without sacrificing performance. In both cases, a common goal was to retain as much of the initial hardware as possible.

An ideal storage system provides high performance (like that provided by the Dell/EMC CX300) at a low price point (like that of the AX100). The final scaled database approach chosen for both initial configurations was to leverage two distinct storage units. Enterprises that originally deploy the AX100 alone and need to scale out can add a CX300 and continue to use the AX100. In the Dell tests, this configuration proved to add significant performance to the database with minimal hardware reprovisioning. Conversely, enterprise deployments that start with a CX300 can augment storage for disk-based recovery at a low price point with the addition of an AX100. This scaled solution demonstrated the ability to maintain the performance level of a CX300 system while enjoying the cost advantages of the entry-level AX100 storage system.

By combining an AX100 and a CX300 within an IT environment, enterprises can scale tested and validated Dell/Oracle configurations without sacrificing performance or cost. This cost-effective, phased scalability strategy can be instrumental to an enterprise's growth.

Mahmoud Ahmadian is an engineering consultant with the Database and Application Engineering Department of the Dell Product Group. Mahmoud has an M.S. in Computer Science from The University of Houston, Clear Lake.

Ujjwal Rajbhandari is a systems engineer in the Database and Application Engineering Department of the Dell Product Group. He is responsible for validating Oracle RAC solutions on Dell PowerEdge servers and Dell/EMC Fibre Channel storage. Ujjwal has a B.E. in Electrical Engineering from the Indian Institute of Technology, Roorkee, and an M.S. in Electrical Engineering from Texas A&M University.

David Mar is a senior software engineer in the Database and Application Engineering Department of the Dell Product Group. His principal focus is developing deployment strategies for Oracle-based solutions on Dell PowerEdge servers, PowerVault storage units, and Dell/EMC storage units. David has a B.S. in Computer Engineering and an M.S. in Computer Science from Texas A&M University.

Chethan Kumar is a systems engineer and advisor in the Database and Applications Group at Dell. He has an M.S. in Computer Science and Engineering from The University of Texas at Arlington.

FOR MORE INFORMATION

Dell/Oracle:
www.dell.com/oracle

Dell/EMC storage arrays:
www.dell.com/storage

