Implementing Oracle Database 10*g*

Maximum Availability Architecture

on Dell PowerEdge Servers and Dell/EMC Storage over Wide Area Networks

The high cost of downtime has prompted many organizations to view business continuity and high availability as critical IT concerns. This article explores how components of Oracle® Maximum Availability Architecture, including Oracle Real Application Clusters and Oracle Data Guard, can be implemented on Dell™ PowerEdge™ servers and Dell/EMC storage to help build the foundation of a scalable, end-to-end architecture.

BY MICHAEL SMITH, PAUL RAD, AND ASHISH RAY

Related Categories:

Clustering

Database

Dell PowerEdge servers

Dell/EMC storage

Disaster recovery

High availability (HA)

Oracle

Storage

Wide area network (WAN)

Visit www.dell.com/powersolutions for the complete category index to all articles published in this issue.

If a critical application, server, or storage subsystem fails—or an unforeseen disaster strikes—the ensuing downtime can place an enterprise in jeopardy. Even carefully planned downtime can affect user productivity when mission-critical applications must be taken offline. Whether planned or unplanned, downtime can translate into lost business opportunities and increased costs.

To help address business continuity requirements, engineering teams from Dell and Oracle conducted a joint project to demonstrate best practices for building a high-availability architecture that enables enterprises to minimize downtime for business applications and build resiliency within their IT infrastructure. This architecture used the Real Application Clusters (RAC) and Data Guard features of Oracle Database 10*g* in conjunction with Dell PowerEdge servers and Dell/EMC storage to help provide high availability, scalability, and disaster protection.

This article explores the high-availability features enabled by Oracle Database 10*g*, Dell PowerEdge servers, and Dell/EMC storage. In addition, the article describes best practices used in the Dell and Oracle joint project.

Integrated high-availability features

One challenge in designing a high-availability IT infrastructure is examining and addressing all possible causes of downtime. Downtime can be classified into two primary categories: unplanned and planned. IT organizations should consider potential causes of both unplanned and planned downtime when designing a fault-tolerant and resilient IT infrastructure. Unplanned downtime primarily results from system failures or data failures. Planned downtime is typically caused by system changes or data changes that must be applied to the production system.

As shown in Figure 1, Oracle Database 10*g*, Dell PowerEdge servers, and Dell/EMC storage offer an integrated set of high-availability features designed to help organizations minimize the various kinds of downtime that can affect their businesses.[1]

[1] For an overview of the high-availability features of Oracle Database 10*g*, visit www.oracle.com/technology/deploy/availability. For Oracle high-availability architecture and best-practices documentation, visit download-west.oracle.com/docs/cd/B14117_01/server.101/b10726/toc.htm and www.dell.com/oracle.

Minimizing unplanned downtime

To help protect against server failures, Oracle RAC allows multiple Dell PowerEdge servers to access a single Oracle database in a clustered environment. This approach can offer the benefit of scalability without requiring changes in application code.

To help protect against downtime caused by various problems—including storage failure, human error, data corruption, and site disruption—Oracle Database 10*g* offers a suite of features. Among them, Automatic Storage Management (ASM) offers file system and volume manager capabilities integrated with Oracle Database 10*g* as well as native mirroring of database files for enhanced protection.[2] To help protect against human error, Oracle Database 10*g* offers the Flashback suite of features. For example, Flashback Database and Flashback Table enable administrators to easily "rewind" the state of the database or database objects to a known, safe point in time. This approach can help undo the effects of human error without incurring much downtime.

To help protect data from various media failures, Oracle offers Recovery Manager (RMAN), which is designed to provide comprehensive backup, restore, and recovery capabilities for the Oracle database. With RMAN, Oracle database backups can be performed online to avoid incurring expensive downtime. Furthermore, Oracle Database 10*g* offers the Flash Recovery Area, which is a unified disk-based storage location for all recovery-related files and activities in an Oracle database. RMAN and Flash Recovery Area are designed to work together to allow an enhanced, automated approach to disk-based backup and recovery. This approach enables administrators to perform fast backups and restores of the Oracle database to help shrink maintenance windows.

Finally, Oracle offers Data Guard to help protect against site disruptions or storage system failures that could result from localized or regional disasters such as fires, earthquakes, hurricanes, and malicious acts. Data Guard enables multiple *standby* databases to be connected to the production, or *primary,* database over a network, keeping the standby databases transactionally consistent with the primary database. If an unforeseen disaster occurs at the primary data center, Data Guard allows the production role to be switched easily to a chosen standby database, and Data Guard can be configured to avoid data loss during this process. Data Guard also allows the standby databases to be used for activities such as reporting and backup.

Dell/EMC storage arrays such as the Dell/EMC CX300, CX500, and CX700 are designed to help improve availability and data integrity by removing single points of failure.[3]

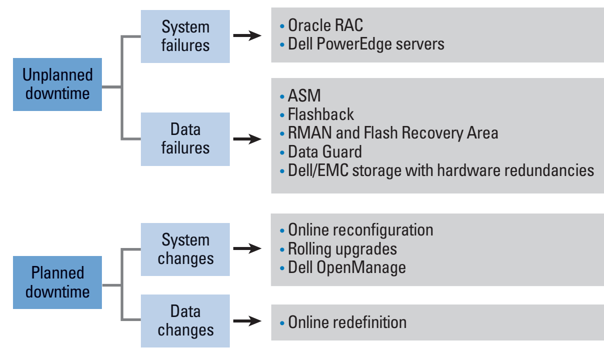

Figure 1. Dell and Oracle integrated high-availability features

Minimizing planned downtime

Planned downtime, which includes activities such as routine maintenance and new deployments, can be just as disruptive to operations as unplanned downtime—especially in enterprises that support users across several time zones. Besides helping to minimize unplanned downtime, Oracle Database 10*g* and Dell PowerEdge servers offer a suite of capabilities that can minimize or eliminate planned downtime.

For example, the Oracle database dynamically accommodates changes to hardware configurations such as adding and removing nodes in a RAC cluster, dynamically growing and shrinking shared memory allocation, and adding and removing online database disks without disturbing database activities. Moreover, the rolling upgrades capability uses Data Guard SQL Apply to permit upgrades of database patch sets or major releases[4] in rotation, which also helps minimize application downtime.

Dell OpenManage™ infrastructure provides administrators with comprehensive, one-to-one systems management capabilities for Dell PowerEdge servers to help eliminate planned hardware downtime and avoid hardware failures that could lead to unplanned downtime.[5]

Finally, to guard against problems caused by data changes, the online redefinition capability enables Oracle Database 10*g* to support many data maintenance operations without disrupting database operations or preventing users from updating or accessing data. For example, the online redefinition capability allows administrators to redefine database tables—including changing table types; adding, dropping, or renaming columns; and changing storage parameters—without interrupting end-user activities such as viewing and updating the underlying data.

[2] For more information about best practices for ASM on Dell/EMC storage, see "Best Practices for Oracle Database 10*g* Automatic Storage Management on Dell/EMC Storage" by Paul Rad, Ramesh Rajagopalan, Tesfamariam Michael, and Jay Kozak in *Dell Power Solutions,* October 2004.

[3] For more information about Dell/EMC storage, visit www.dell.com/emc.

[4] This capability enables upgrades for database patch sets or major releases from Oracle Database 10*g* Release 1 and later.

[5] For more information about Dell OpenManage, visit www.dell.com/openmanage.

Best-practices guidelines for configuring Oracle MAA on Dell servers

The goal of the Dell and Oracle project was to provide best-practices guidelines for configuring Oracle Maximum Availability Architecture (MAA) using Oracle Database 10*g* on Dell servers and Dell/EMC storage. The objective of MAA—Oracle's blueprint for Oracle high-availability technologies—is to remove the complexity of designing an optimal high-availability architecture and to help maximize systems availability.[6]

Two fundamental technologies that enable MAA are RAC[7] and Data Guard,[8] which help the system architecture provide end-to-end support for high availability and disaster protection. RAC is designed to provide IT organizations with near-instantaneous server failover when server failures occur, while Data Guard—especially if deployed over a wide area network (WAN)—helps protect production data from disasters that could otherwise severely affect data center operations. The following sections provide further details on the joint Dell and Oracle project.

System and network considerations

Conducted in the fourth quarter of 2004 by Dell and Oracle engineers, the joint Dell and Oracle project involved setting up three sites with identical hardware and software configurations: two sites in the Dell Engineering Lab in Austin, Texas, and one site in the Oracle System Technology Data Center in Redwood Shores, California (see Figure 2).

In the Dell and Oracle implementation configured for this study, each site consisted of redundant hardware and software components designed to ensure that all requests were serviced, even if a failure occurred. The primary site was located in Austin, with the primary database hosted on a two-node Dell cluster comprising Dell PowerEdge 2650 servers. The two servers were configured with Oracle Database 10*g* (10.1.0.2), including the RAC option, and a Dell/EMC CX500 Fibre Channel storage array.

Dell and Oracle engineers configured the implementation described in this article for demonstration purposes. However, many different implementations are possible using Dell and Oracle supported and validated configurations;[9] administrators should implement the configuration that best suits their organization's specific needs.

The primary database in the Dell and Oracle implementation described in this study was configured with two standby databases: a logical standby database and a physical standby database. The LAN-attached logical standby database was hosted in Austin and maintained using Data Guard to enable synchronous transport. The WAN-attached physical standby database was hosted in Redwood Shores and maintained using Data Guard to enable asynchronous transport. To allow the standby sites to process application workloads at the same performance level as the primary site during a role transition, both standby databases were configured identically to the database at the primary site.

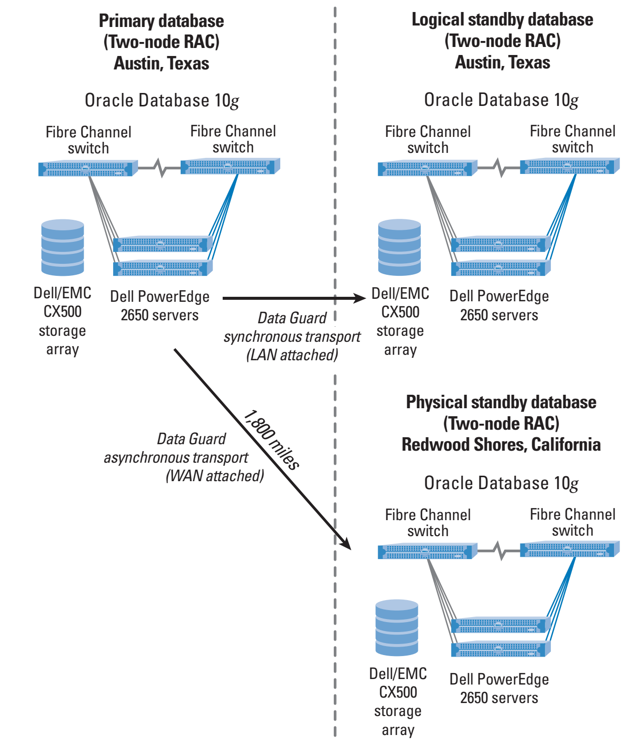

Figure 2. Dell and Oracle MAA configuration

Oracle database considerations

When building an MAA environment, administrators organize the major tasks into four main categories: preparing the cluster for the primary and standby sites, building and configuring the primary database, building and configuring the physical standby database, and building and configuring the logical standby database. The following sections provide an overview of steps that were completed in each category to implement the example Dell and Oracle configuration described in this article.[10]

Preparing the cluster to run RAC. Dell and Oracle engineers completed the following steps on each node in the server cluster for the primary database site in Austin, the physical standby

[6] For more information about MAA best-practices publications, visit www.oracle.com/technology/deploy/availability/htdocs/maa.htm.
[7] For more information about Oracle RAC, visit www.oracle.com/technology/products/database/clustering/index.html.
[8] For more information about Oracle Data Guard, visit www.oracle.com/technology/deploy/availability/htdocs/DataGuardOverview.html.
[9] For a comprehensive list of Dell and Oracle supported and validated configurations, visit www.dell.com/downloads/global/solutions/Dell-Oracle10g%20Software%20Product%20Matrix%20version%201.1.pdf.
[10] For detailed MAA deployment information, visit www.oracle.com/technology/deploy/availability/htdocs/maa.htm.

database site in Redwood Shores, and the logical standby database site in Austin:

- Installed required OS packages
- Created OS groups and users
- Configured system networking
- Configured Secure Shell (SSH) and Secure Copy (SCP)
- Set OS kernel parameters
- Created the Oracle Cluster Registry (OCR) and voting disks[11]
- Installed Cluster Ready Services (CRS)
- Installed Oracle Database 10*g* software

Creating the RAC primary database. Once the cluster was operational and all Oracle software had been installed, Dell and Oracle engineers created and configured the primary database at the Austin site as follows:

- Created a RAC database using Oracle Database Configuration Assistant (DBCA)
- Configured the database for MAA best practices, which included setting up multiple control files, setting up multiple online redo log groups, enabling archivelog mode, using automatic undo management, and using locally managed table spaces
- Set appropriate initialization parameters
- Enabled Flashback Database

Creating the RAC physical standby database. Once the primary database was configured, Dell and Oracle engineers created the physical standby database at the Redwood Shores site as follows:

- Performed an online backup of the primary database using RMAN
- Configured Oracle Net Services on each node of the standby database
- Created the standby database from the previously performed backup
- Set appropriate initialization parameters for the standby database
- Added the standby database and instances into the OCR
- Configured the primary database for redo transport
- Verified that the standby database was receiving redo generated by the primary database

Creating the RAC logical standby database. To support a logical standby database in the Data Guard configuration, Dell and Oracle engineers created the logical standby database at the Austin site as follows, taking care to ensure that the tables in the primary database included only data types that were supported by Data Guard SQL Apply:[12]

- Prepared the primary database to support a logical standby database
- Created a physical standby database by following the steps recommended in the "Creating the RAC physical standby database" section in this article
- Created a logical standby control file at the primary database
- Started up and mounted the standby database using the control file
- Activated the logical standby database and assigned it a unique name using the Oracle DBNEWID utility
- Changed the database type in the OCR to indicate it is a logical standby database
- Started SQL Apply and verified that the logical standby database was performing properly

Network configuration implications for Data Guard

For the Data Guard configuration described in this study, Dell and Oracle engineers created the logical standby database in the same data center as the primary database, and connected the logical standby database to the primary database over a LAN. The physical standby database was located 1,800 miles away, connected to the primary database over a WAN.

The objective behind such a configuration, which involved two standby databases, was to use the local logical standby database for reporting activities while enabling the remote physical standby database to be used for recovery from disasters at the primary site. Because an Ethernet LAN enables high reliability and low latency, the logical standby destination was configured using the attributes `LGWR SYNC AFFIRM` on the primary database. However, in view of the possible latencies associated with an IP-based WAN, the physical standby destination was configured with the attributes `LGWR ASYNC`. This asynchronous configuration helped minimize the impact of latency on production throughput while helping to contain the exposure to data loss in the event of a disaster.[13]

Demonstration of MAA capabilities in a high-availability implementation

After the databases and hardware were configured at the sites in Austin and Redwood Shores, Dell and Oracle engineers tested and validated the example configuration in various ways. This

[11] For more information about the OCR and voting disks, visit download-west.oracle.com/docs/cd/B14117_01/rac.101/b10765/toc.htm.

[12] For a complete list of data types not supported by Data Guard SQL Apply, visit download-west.oracle.com/docs/cd/B14117_01/server.101/b10823/toc.htm.

[13] For a comprehensive discussion of network configuration best practices for Data Guard, visit www.oracle.com/technology/deploy/availability/htdocs/maa.htm.

section highlights the major validation steps completed by the Dell and Oracle team.

Measuring load generation and network throughput

To properly demonstrate the MAA environment, Dell and Oracle engineers generated a transaction load in the primary database using a simple SQL*Loader script that inserted 20 million records into a test table. This generated redo on the primary database, and Data Guard transmitted that redo to the standby databases to keep the standby databases synchronized with the primary database. Because the redo was transmitted over the network, engineers verified that the network infrastructure was adequate to support the Data Guard configuration. This was accomplished by determining the peak database redo generation rate and comparing that rate against the network throughput. The peak redo generation rate was observed by generating Oracle Automatic Workload Repository (AWR) reports.[14] The effective network throughput rate was measured using a utility called Iperf.[15] Using these tools, engineers measured the throughput for the WAN between Austin and Redwood Shores. This approach can help an organization evaluate whether the network in its own specific Data Guard configuration can provide adequate bandwidth to support the transaction load of a particular primary database.

> The objective of MAA—Oracle's blueprint for Oracle high-availability technologies—is to remove the complexity of designing an optimal high-availability architecture and to help maximize systems availability.

Using the physical standby database in read-only mode

The physical standby database can be used to perform read-only reporting in addition to providing disaster protection for the primary database. In the example implementation described in this article, engineers enabled the read-only reporting capability by temporarily suspending Redo Apply operations on the physical standby database at Redwood Shores and then opening the database in read-only mode. Once the physical standby database was open in read-only mode, engineers ran simple queries on the tables in the physical standby database to verify that the standby database was functioning properly.

After they completed this exercise, Dell and Oracle engineers restarted Redo Apply operations in the physical standby database. The Redo Apply operations automatically applied to the physical standby database all the redo that had accumulated in the physical standby database server while the physical standby database was open in read-only mode—thereby bringing the physical standby database up-to-date with the primary database.

Using the logical standby database for reporting

The logical standby database can provide advanced reporting capabilities because it is designed to be opened in read-write mode while changes from the primary database are applied to it. In the example configuration described in this article, such advanced reporting capabilities indicate that the logical standby database at the Austin site could be used as a real-time reporting system when the logical standby destination is configured with `LGWR SYNC` in the primary database and SQL Apply is running in Real-Time Apply mode[16] on the logical standby database.

Dell and Oracle engineers validated the real-time reporting capabilities of the example configuration described in this article by verifying that updates occurring on the primary database were instantaneously observed on the logical standby database. To help ensure a transactionally consistent view of the data as it was applied on the logical standby database, the SQL Apply parameter `TRANSACTION_CONSISTENCY` was set to `FULL`.

Performing a switchover

The fundamental benefit of a Data Guard configuration is that any standby database can be chosen to take over as the primary database. This approach enables organizations to resume business operations without incurring significant downtime if an outage affects the primary data center. Data Guard offers two types of role transitions: *switchover,* to be invoked for planned maintenance, and *failover,* to be invoked after unplanned outages or disasters at the primary site.

To simulate a planned outage using the example implementation and demonstrate the ease with which Data Guard can perform such role transitions, Dell and Oracle engineers executed a switchover across the WAN between the primary database in Austin and the physical standby database in Redwood Shores. Before performing the switchover, the test team verified that the standby database had recovered all available redo and that the managed recovery process was running on the physical standby database. Also, all primary and standby instances, with the exception of the ones in which the switchover was performed, were shut down.[17]

---

[14] For more information about AWR, visit download-west.oracle.com/docs/cd/B14117_01/server.101/b10739/toc.htm.

[15] For more information about Iperf, visit dast.nlanr.net/Projects/Iperf.

[16] For more information about Real-Time Apply, visit download-west.oracle.com/docs/cd/B14117_01/server.101/b10823/toc.htm.

[17] For more information about Data Guard switchover and failover best practices, visit www.oracle.com/technology/deploy/availability/htdocs/maa.htm.

The switchover commands—one command converts the primary database to a standby database and another command converts the standby database to the primary database—were respectively entered on the remaining instances of the primary and standby databases. The `database_role` column in the `v$database` view confirmed the status of the databases after the switchover was complete. The remaining instances of the new primary database and the new standby database were restarted, after which the database at the Redwood Shores location became the full-fledged primary database. Applications could then be directed to this new primary database in Redwood Shores, which would, in turn, start generating redo and transmit this redo to the two standby databases at the Austin site.

No data was lost in the process of the database switchover. Because the Real-Time Apply feature of Oracle Database 10*g* was used in the example configuration, the Data Guard switchover process did not have to wait for accumulated archived logs to be applied, as was the case in previous Oracle database releases. In addition, appropriate settings had been configured in all three databases before the switchover occurred, in anticipation of possible role transitions. For example, engineers set up the Oracle Database 10*g* `VALID_FOR` attributes for the `log_archive_dest_n` parameters and turned on supplemental logging at the physical standby database. Thus, the logical standby database in Austin was automatically brought along as a logical standby database to the new primary database in Redwood Shores.

Note: To revert to the original configuration following a switchover—that is, to perform a *switchback*—administrators can simply perform another switchover operation.

Monitoring Data Guard performance

The performance of a Data Guard configuration depends on whether the Redo Apply and SQL Apply operations on the physical and logical standby databases are able to keep up with the redo generation rates of the primary database.

The Oracle database offers various views that allow administrators to track the progress of the standby databases. Some relevant views to monitor the progress of the physical standby database are `v$managed_standby` and `v$dataguard_status`. Similarly, the logical standby database can be monitored using the views `v$logstdby` and `dba_logstdby_progress`.

The information provided by these views was used in the study described in this article to help engineers determine that the redo data generated by the primary database was indeed being transmitted to the standby servers across the network, and applied to the standby databases according to the configuration parameters set forth by the project team.

Understanding redo transport and networking best practices

The extent of data protection in a Data Guard configuration is controlled by the Data Guard protection mode in effect, which in turn determines whether redo is sent to the standby database synchronously or asynchronously over the network. To maintain high primary redo generation rates with the least impact on application throughput at the primary database, administrators should allow sufficient low-latency network bandwidth between the primary and standby sites and properly configure relevant OS network settings (such as TCP send and receive buffer sizes). In general, MAA best practices recommend `LGWR SYNC` (Maximum Protection or Maximum Availability protection mode) for low-latency LAN or metropolitan area network (MAN) environments in which the standby site may be located up to a few hundred miles from the primary site. `LGWR ASYNC` or ARCH-based redo transport (Maximum Performance protection mode) is recommended for more distant standby databases connected by an IP-based WAN.

Characteristics of a highly resilient IT architecture

A global enterprise must be based on a highly resilient IT architecture to provide mission-critical levels of service to customers and stakeholders around the world. Such an architecture must be complete, integrated, easy to manage, and flexible enough to serve multiple purposes. At the same time, the technology to implement a highly resilient IT architecture should be cost-effective to enable businesses to derive optimal value from their IT investments.

Oracle Database 10*g* offers an integrated suite of high-availability capabilities that can help meet these demanding business and technical requirements. The implementation discussed in this article demonstrates the ease with which Oracle Database 10*g* MAA can be deployed over a WAN on Dell servers and a Dell/EMC storage platform to provide a robust degree of high availability. Using the configuration best practices discussed in this article, enterprises can deploy a combined Dell and Oracle platform to build an end-to-end, highly available and scalable infrastructure.

Michael Smith is a senior member of technical staff with the Oracle Database High Availability Group. His principal focus is Oracle Data Guard. He has a bachelor's degree in Computer Science.

Paul Rad is a senior software engineer in the Dell Database and Application Engineering Department of the Dell Product Group. He has master's degrees in both Computer Engineering and Computer Science from The University of Texas at San Antonio.

Ashish Ray is a group product manager with the Oracle Database High Availability Group. His principal focus is Oracle Data Guard. He has bachelor's and master's degrees in Computer Science, and an M.B.A.