www.ProcaseConsulting.com
Phone: (905) 856-7479
Fax: (416) 352-1334

Oracle RAC, Data Guard, and Recovery Manager Concepts

www.ProcaseConsulting.com
Phone: (905) 856-7479
Fax: (416) 352-1334

Table of Contents

Revision History

REVISION	Description	By	Date
1.0	Concepts of Oracle RAC, Data Guard and RMAN	Nasser Kazi	March 30, 2003

Introduction

This document provides an overview of the high availability options that are available under Oracle 9i. The concepts of Oracle Real Application Clusters (RAC), Data Guard and Recovery Manager (RMAN) are also discussed. Comparisons are made between raw-devices and cluster file systems (CFS). An overall summary is provided at the end of the document for an appropriate solution using the outlined technology.

www.ProcaseConsulting.com
Phone: (905) 856-7479
Fax: (416) 352-1334

Raw Devices and Cluster File Systems

Oracle Real Application Clusters requires that each instance be able to access a set of unformatted devices on a shared disk subsystem. These shared disks are referred to as raw devices. Oracle instances in Real Application Clusters use raw devices to update control files, server parameter files, datafiles, and online redo logs. All instances in the cluster share these files. Raw devices must be pre-created using OS specific commands before the database can be created. Archive logs are written to the local filesystem and should not be stored on raw devices.

Raw Devices

Each raw device is created from a set of unformatted partitions referred to as raw partitions. A raw partition is a portion of a physical disk that is accessed at the lowest possible level. A raw device has no attached filesystem to it.

Cluster File Systems

A Cluster File System (CFS) is a filesystem that may be accessed by all members in a cluster at the same time. This implies that all members of a cluster have the same view.

Oracle has developed the Oracle Cluster File System (OCFS) that was designed specifically for Oracle Real Application Clusters. OCFS eliminates the requirement for Oracle database files to be linked to logical drives and enables all nodes to share a single Oracle home instead of requiring each node to have its own local copy. OCFS volumes can span one shared disk or multiple shared disks for redundancy and performance enhancements.

The following files can be placed on an Oracle Cluster File System:

- Oracle files (controlfiles, datafiles, redologs, bfiles)
- Shared configuration files (spfile)
- Files created by Oracle during runtime

Technology Comparisons

The table below outlines the benefits for each type of technology:

CFS	Raw Devices
Simpler management	Always required when CFS is not available or not supported by Oracle
Use of Oracle Managed Files with RAC	Raw devices offer best performance without any intermediate layer between Oracle and the disk
Single Oracle Software installation	
Enables the use of autoextend capabilities for datafiles	
Uniform accessibility to archive logs in case of physical node failure	
With Oracle home on CFS, CFS guarantees that patches applied to the Oracle Home on one of the nodes is visible to all nodes in the cluster.	
Logical Storage Managers or Logical Volume Managers can ease the work with raw devices by the ability to modify (extend, add, delete) raw devices online	

The table below outlines the limitations for each type of technology:

CFS	Raw Devices
Additional software licensing costs may be incurred	Autoextend fails on raw devices if space is exhausted in the partition
	Complex maintenance process.

Although both types of technologies can coexist, Oracle does not advise using both raw devices and CFS together due to manageability reasons.

Reference

http://metalink.oracle.com
Note: 183408.1 - Raw Devices and Cluster Filesystems With Real Application Clusters

High Availability (HA) Options

High availability systems are configured in order to provide 24X7 operations that avoid having single points-of-failure. This is accomplished by having hardware and/or software redundancy that take over the tasks performed by the failed nodes or components.

Oracle offers a wide variety of options providing high availability, on top of the Oracle 9i kernel. The HA features provided by Oracle 9i are RAC, RAC Guard, Data Guard, Failsafe, and Advanced Replication.

Considerations about the cost, the needed availability, the downtime allowed in case of scheduled or unscheduled downtime due to any disaster would help in making a choice between the above options.

Oracle Real Application Clusters (RAC)

Oracle RAC relies on clustered hardware and permits multiple instances to share a single database. RAC offers high availability with Transparent Application Failover (TAF) in which active sessions on a failed node or instance can be automatically failed over to surviving instances on other nodes. RAC also provides scalability. The workload can be spread among all nodes, while maintaining constant response times. Based on Oracle's Cache Fusion architecture, Oracle9i Real Application Clusters enables sharing of frequently accessed data across all the servers in a cluster. When a remote cache server a query request, the block is transferred across the high-speed cluster interconnect from one node's cache to another. This "fusing of the caches" happens automatically and is transparent to the application. This transparency is the key technology that provides the fast, efficient scaling of Oracle9i Real Application Clusters. This block shipping through the High Speed Interconnect was true in Oracle 8i for read-read pings; it is now also true for write-write pings in Oracle 9i.

Oracle RAC Configuration Options

Oracle RAC Guard / Oracle Failsafe	RAC with additional HA features
Data Guard	Provides protection against data loss
Advanced Replication	Provides sharing of information across differing hardware platforms or database versions

Oracle RAC Guard

Oracle Real Application Clusters Guard is an enhanced configuration of Real Application Clusters. RAC Guard connects all clients to one Oracle instance in normal operation. In the event of a failure, Oracle's fail over and monitoring software together with a partner's cluster framework will detect the problem and gracefully switch over clients to the second Oracle instance, ensuring continued data access. RAC Guard has been designed to quickly and automatically recover from many failures, including hardware, operating system, or Oracle instance faults. It provides fast and bounded recovery, enforced primary/secondary node access, automatic reconnect of failed sessions, and capture of diagnostics data after a failure.

Oracle RAC Guard Configuration Options

Data Guard	Provides protection against data loss
Advanced Replication	Provides sharing of information across differing hardware platforms or database versions

Oracle Failsafe

Oracle Failsafe is a core feature of Oracle9i, Oracle iAS, and Oracle Applications Release 11i that provides high availability for e-business solutions deployed on Windows clusters. Oracle Fail Safe release 3.2 works with Microsoft Cluster Server to ensure that if a failure occurs on one cluster node, then the Oracle databases and applications running on that node will automatically fail over to a surviving node, therefore minimizing downtime. Oracle Failsafe is optimized for Windows customers with database and application workloads that can be handled by a single system. Oracle Failsafe solutions can be deployed on Windows NT and Windows 2000 clusters. It supports up to 4 servers in a cluster with Windows 2000.

Oracle Failsafe supports the fail over for the following products and services:

- Oracle Databases (Standard and Enterprise Editions of Oracle9i and Oracle8i)
- Oracle Applications release 11i
- Oracle iAS components, including:
- Oracle Forms Services
- Oracle Reports Services
- Oracle HTTP Server
- Oracle Intelligent Agent
- Oracle Service for Microsoft Transaction Server (for Oracle8i release 8.1.7)

Oracle Failsafe Configuration Options

Data Guard	Provides protection against data loss
Advanced Replication	Provides sharing of information across differing hardware platforms or database versions

Oracle Data Guard

Oracle9i Data Guard provides both physical and logical standby database (Oracle 9.2+ only) protection and disaster recovery features. These features would guard against operation mistakes, corruptions, and other disasters that may destroy a database. Data Guard broker protects critical data by automating the creation, management, and monitoring aspects of a standby database environment.

Oracle 8i provided a physical standby database that mirrored the primary database. The archived redo logs of the primary database were shipped across to the standby database. If the primary database went down, the standby database would take over the primary role and the interruption of service were minimized.

In Oracle 9.2, the logical standby database was introduced. This functionality enabled the archived redo logs to be mined for SQL statements that were then applied to the logical standby database. An advantage of this feature is the capability of reading and writing (to create indexes or do some maintenance) to the logical standby database while applying logs through SQL statements.

The standby database can be configured on a Single-Node or on a Cluster system. In Oracle 9i (9.0.1 and 9.2.0) the Data Guard Manager does not support RAC. Therefore, the standby configuration requires to be done manually.

Standby Database on a Single-Node System

In the case of a standby database on a Single-Node, all the instances have to archive the logs to the service of the Single-Node standby system. The standby system creates a "pool" of archive logs and on the basis of the database System Change Number (SCN), determines which archive log to apply in sequence.

Clustered RAC Standby Database

It is also possible to create a standby system on a RAC environment as well. This provides more scalability and performance improvements. If the systems were identical, there would be minimal performance and availability degradation during a switchover or fail over process.

Typically, the same number of instances should exist on the standby system as the primary system.

Oracle Data Guard Configuration Options

RAC and RAC Guard	Will provide additional HA and scalability
Advanced Replication	Provides sharing of information across differing hardware platforms or database versions

Oracle Advanced Replication

Replication is the process of copying and maintaining database objects in multiple databases that make up a distributed database system. There are three ways of replicating data:

- Read-only materialized views: One master table and multiple remote copies being refreshed at intervals by pulling changes from the "master site" to the "slave site".
- Updateable materialized views: One master table and multiple partial or entire copies on snapshot sites (updateable).
- N-way masters replication: Multiple master tables being updateable from all sites. Changes applied at one site are captured and stored locally before being forwarded and applied at each of the remote locations.

In addition to providing users with fast, local access to shared data, applications benefit from increased availability because alternate data access options exist. Even if one site becomes unavailable, users can continue to query or even update the remaining locations.

Oracle Replication is suitable for applications that typically require data to be periodically synchronized between central systems and very large numbers of small, remote sites often operating in a disconnected manner. Oracle provides a sophisticated conflict detection mechanism and a comprehensive set of automated conflict resolution routines to ensure data convergence throughout the replicated environment.

Oracle Replication facilitates information sharing across multiple hardware platforms and database versions.

The replication option works on an object level, whereas the other HA options work at a database level. Furthermore, some operations like adding objects in a replication group require quiescing the replication group, making the replicated objects unavailable. A quiesced state prevents users from executing any transactions against a replicated object in the quiesced master group. Oracle Replication has been enhanced in Oracle 9i and now it is no longer necessary to quiesce a replication group when adding or removing a master database.

Oracle Data Guard Configuration Options

RAC and RAC Guard	Will provide additional HA and scalability
Advanced Replication	Provides sharing of information across differing hardware platforms or database versions

Reference

Note:179312.1 - A Short Description of HA Options Available in 9i
Note 132987.1 - High Availability - Real Application Clusters Library Page Index
Note 146388.1 - High Availability - Oracle Fail Safe Library Index
Note 146387.1 - Backup and Recovery - Standby Databases and Data Guard Library
Note 132256.1 - Distributed - Advanced Replication Library Index
Real Application Clusters Concepts
Real Application Clusters Installation and Configuration
Real Application Clusters Administration
Real Application Clusters Deployment and Performance
Oracle Real Application Clusters Guard Administration and Reference Guide
Oracle Real Application Clusters Guard Installation Guide
Oracle Real Application Clusters Guard Concepts and Administration Guide
Oracle Services for MSCS Error Messages
Oracle Fail Safe Installation Guide
Oracle Fail Safe Concepts and Administration Guide
Oracle9i Data Guard Concepts and Administration
Oracle9i Data Guard Broker
Oracle9i Replication
Oracle9i Replication Management API Reference
Note:180031.1 - Creating a Data Guard Configuration
Note:150584.1 - Data Guard 9i Setup with Guaranteed Protection Mode
Oracle9i Data Guard Concepts and Administration, Release 2 (9.2), A96653-01
Oracle9i Database Reference, Release 2 (9.2), A96536-01
Oracle9i SQL Reference, Release 2 (9.2), A96540-01
Note:203326.1 - Data Guard 9i Log Transportation on RAC

Oracle Recovery Manager

Recovery Manager (RMAN) is an Oracle utility that can back up, restore, and recover database files. The product is a feature of the Oracle database server and does not require separate installation. Recovery Manager is a client/server application that uses database server sessions to perform backup and recovery. It stores metadata about its operations in the control file of the target database and, optionally, in a recovery catalog schema in an Oracle database. RMAN can be invoked as a command-line executable from the operating system prompt or use some RMAN features through the Enterprise Manager GUI.

Most production database systems impose stringent requirements on backup and recovery. RMAN facilitates the following tasks:

- Manage the complexity of backup and recovery operations
- Minimize the possibility of human error
- Make backups scalable and reliable
- Utilize all available media hardware
- Make backups proportional to the size of transactional changes, not to the size of database
- Make recovery time proportional to the amount of data recovered
- Have a single centralized repository of metadata for all the existing databases in the organization

Reference

Oracle 9i Recovery Manager User's Guide Release 2 (9.2)

Conclusion

When making a decision about high availability, it will be important to weigh the costs vs. the cost of downtime, the various features available, and the implementation of these features.

Based on hardware acquired by Keifer-SCARS a two-node cluster is proposed. A single-instance physical standby database would be used in conjunction with RMAN database. The backup strategy of RMAN database would be to configure RMAN repository database located on the same server as the standby database and have a backup placed on one of the nodes of the RAC database.

The architecture proposed for the implementation of RAC and Data Guard coupled with RMAN is illustrated below.

Oracle RAC / Data Guard / RMAN configuration