An Oracle White Paper
November 2010

Oracle Real Application Clusters (RAC) 11g Release 2

ORACLE

Introduction

Oracle Real Application Clusters (RAC) allows Oracle Database to run any packaged or custom application, unchanged across a server pool. This provides the highest levels of availability and the most flexible scalability. If a server in the pool fails, the Oracle database continues to run on the remaining servers. When you need more processing power, simply add another server to the pool without taking users offline. To keep costs low, even the highest-end systems can be built out of standardized, commodity parts.

Oracle Real Application Clusters provides a foundation for Oracle's Private Cloud Architecture. Oracle RAC technology enables a low-cost hardware platform to deliver the highest quality of service that rivals and exceeds the levels of availability and scalability achieved by more expensive mainframe SMP computers. By dramatically reducing administration costs and providing new levels of administration flexibility, Oracle RAC enables private clouds. Oracle RAC 11g Release 2 in addition enables customers to build a dynamic private cloud infrastructure.

This paper provides a technical overview of Oracle Real Application Clusters 11g with the emphasis on the features and functionality that can be implemented to provide the highest availability and scalability for enterprise applications.

"We've been able to save over $5 million dollars a year by re-platforming from our mainframe to Oracle Real Application Clusters."

— Eugene Park, Senior Director of Platform Services, PG&E

What is Oracle Real Application Clusters?

Oracle Real Application Clusters is an option to the Oracle Database Enterprise Edition that was first introduced with Oracle 9*i*. Oracle Real Application Clusters is now proven technology used by thousands of customers in every industry for every type of application. Oracle RAC provides options for scaling applications beyond the capabilities of a single server. This allows customers to take advantage of lower cost commodity hardware to reduce their total cost of ownership and provide a scaleable computing environment that supports their application workload. Oracle RAC enables the Oracle Database to run mainstream business applications of all kinds on server pools including popular packaged products (such as Oracle Applications, Peoplesoft, SAP), in-house developed applications, which can be either OLTP, DSS, or a mixed workload.

Oracle Real Application Clusters is a key component of Oracle's Maximum Availability Architecture (MAA)[1], which provides direction to architect the highest availability for applications. Oracle RAC provides the ability to remove the server as a single point of failure in any database application environment.

Oracle Real Application Clusters Architecture

An Oracle RAC database is a clustered database. A cluster can be described as a pool of independent servers that cooperate as a single system. Server Pools provide improved fault resilience and modular incremental system growth over single symmetric multi-processor (SMP) systems. In the event of a system failure, clustering ensures high availability to users. Access to mission critical data is not lost. Redundant hardware components, such as additional servers, network connections, and disks, allow the cluster to provide high availability. Such redundant hardware architectures avoid a single point-of-failure and provide exceptional fault resilience.

[1] For more information on Oracle's Maximum Availability Architecture see: http://www.oracle.com/technetwork/database/features/availability/maa-090890.html

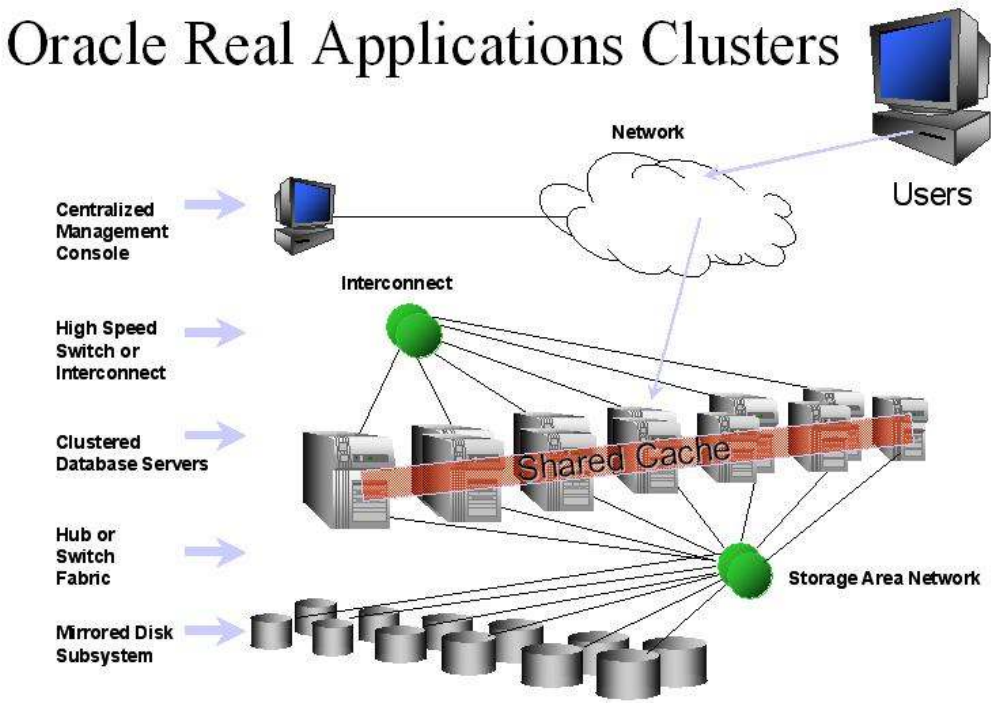

Figure 1 Oracle Real Application Clusters Architecture

With Oracle Real Application Clusters, Oracle de-couples the Oracle Instance (the processes and memory structures running on a server to allow access to the data) from the Oracle database (the physical structures residing on the storage storing the data, commonly referred to as the datafiles). A clustered database is a single database that can be accessed by multiple instances. Each instance runs on a separate server in the server pool. When additional resources are required, additional servers and instances can easily be added to the server pool with no downtime. Once the new instance is started, applications using services can immediately take advantage of it with no changes to the application or application server.

Oracle Real Application Clusters is an extension of the Oracle Database and therefore benefits from the manageability, reliability and security features built into Oracle Database 11g.

Oracle Clusterware

Starting with Oracle Database 10g, Oracle provides Oracle Clusterware, a portable clusterware solution that is integrated and designed specifically for the Oracle Database. Oracle Clusterware provides a complete clustering solution and supports any application. Oracle Clusterware is a pre-requisite for all Oracle RAC implementations. Support is made easier as there is one support organization to deal with for the clusterware and cluster database. One can choose to run Oracle

RAC on top of additional, selected third party cluster solutions, in which case, Oracle Clusterware must still manage all Oracle RAC databases on the system.

Oracle Clusterware monitors and manages Oracle Real Application Cluster databases. When a server in the server pool is started, all instances, listeners and services are automatically started. If an instance fails, the Oracle Clusterware will automatically restart the instance so that the service is often restored before the administrator notices it was down.

With Oracle Database 10g Release 2, Oracle added a High Availability API so that non-Oracle applications can be managed by the high availability framework provided by Oracle Clusterware. When registering the application with Oracle Clusterware, information must be provided on how to start, stop, and monitor the process. In addition, one can specify candidate servers to take over a resource once a failure occurred. With Oracle Database 11g Release 2, managing applications was made even easier through the graphical interface provided by Oracle Enterprise Manager (EM). Using Oracle EM and the enhanced High Availability Framework provided by Oracle Clusterware in conjunction with even more dependency options, Oracle Clusterware can now be use to easily model even complex failover and recovery scenarios.

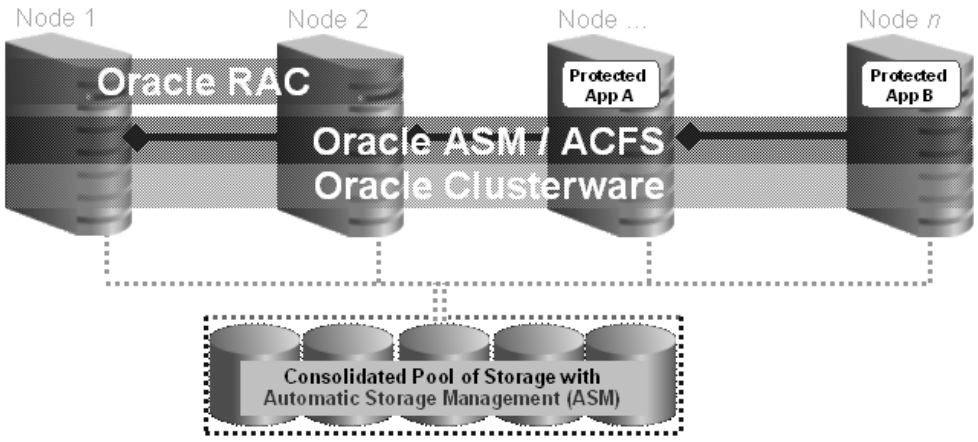

Figure 2: Oracle Clusterware Overview

Hardware Architecture

Oracle Real Application Clusters is a shared everything architecture. All servers in the server pool share all storage used for an Oracle RAC database. The type of storage pool used can be network attached storage (NAS), storage area network (SAN), or SCSI disks. Your storage choice is dictated by the server hardware choice and the hardware supported by your hardware vendor. The key to choosing an appropriate storage pool is choosing a storage system that will provide scaleable I/O for your application and an I/O system that will scale as additional servers are added to the pool.

An Oracle RAC Database requires a network connection to the Local Area Network (LAN) that a database server is attached to for application connections. A server pool also requires a private network commonly known as "the interconnect". Oracle recommends redundant interconnects for high availability purposes. With Oracle Real Application Clusters 11g Release 2, patch set 1 (11.2.0.2), Oracle provides a native solution to support redundant interconnects as past of the Oracle Grid Infrastructure solution. With former versions, an external, mostly Operating System (OS) dependent network redundancy solution (bonding / teaming) should be used to provide failover and load balancing. The interconnect is used by Oracle Clusterware for inter-node messaging. The interconnect is also used by Oracle RAC to implement the cache fusion technology. Oracle recommends the use of UDP over GigE for the cluster interconnect. The use of crossover cables as the interconnect is not supported for a production Oracle RAC database.

A server pool is made up of 1 or more servers, each having a public LAN connection, an interconnect connection, and must be connected to a shared pool of storage. Oracle Clusterware and Oracle Real Application Clusters support up to 100 nodes in the cluster. Each server in the cluster does not have to be exactly the same but it must run the same operating system, and the same version of Oracle. All servers must support the same architecture; e.g. all servers must be either 32bit or 64bit.

Current and detailed information on certifications and technology restrictions related to Oracle Real Application Clusters can be obtained through My Oracle Support (http://support.oracle.com)

File Systems and Volume Management

Since Oracle RAC is a shared everything architecture, the volume management and file system used must be cluster-aware. Oracle recommends using Oracle Automatic Storage Management (ASM), which is a feature included with Oracle Database 11g to automate the management of various storage pools for the database. ASM provides the performance of an async I/O storage subsystem with the ease of management of a file system. ASM distributes I/O load across all available resource to optimize performance while removing the need for manual I/O tuning. With Oracle Database 11g Release 2 ASM comes with a dynamic volume manager and a general purpose file system.

Alternatively, Oracle supports certain, certified cluster file systems such as the Oracle Cluster File System (OCFS) that is available on Windows and Linux (called OCFS2). With Oracle Database 11g Release 2, the Oracle Universal Installer and the Database Configuration Assistant do not support the use of raw or block devices for database files anymore. Databases that are currently using raw devices can continue to use those during or after an upgrade to 11g Release 2 as the command line interfaces continue to support the use of raw devices.

Oracle Grid Infrastructure

With Oracle RAC 11g Release 2, Oracle introduces Oracle Grid Infrastructure. With Oracle Grid Infrastructure, Oracle integrated Oracle ASM, the proven storage pool management solution for the Oracle Database, and Oracle Clusterware in one software bundle. Oracle has thereby combined two of its strongest products for cluster environments to form a universal grid foundation. In an Oracle RAC environment, Oracle Grid Infrastructure provides the necessary foundation with respect to volume management, file system, and server pool management to run an Oracle RAC database. The system administrator and systems administration team should therefore manage oracle Grid Infrastructure.

Virtual Internet Protocol Address (VIP)

Oracle Real Application Clusters 11g requires a virtual IP address for each server in the cluster. The virtual IP address is an IP address managed by Oracle Clusterware and taken from the same subnet as the Local Area Network (LAN). This address is used by applications to connect to the Oracle RAC database. If a node fails, the Virtual IP is failed over to another node in the cluster to provide an immediate "node down"-response to incoming connection requests. This increases the availability for applications, as they no longer have to wait for network timeouts before the connection request fails over to another instance in the server pool.

Grid Naming Service

Oracle RAC 11g Release 2 introduces the Grid Naming Service (GNS), which makes it easier to scale by automating the VIP management for Oracle RAC. In order to use GNS, one will have to work with the Network Administrator to set up a delegated domain in the Domain Name Service (DNS) and a Virtual IP address for the Grid Naming Service (GNS). However, once this part is set up, Oracle will manage the VIPs in the cluster automatically. Regardless of whether or not you add or remove servers from the pool, you will not have to return to the network administrator for additional IP addresses. To further automate the network management in the cluster, the use of the Grid Naming Service (GNS) requires a DHCP server on the public network. Oracle uses DHCP to dynamically allocate the required Virtual IP addresses as servers join the cluster. Overall, GNS will reduce unnecessary, manual re-configuration with respect to the network in dynamic environments.

Single Client Access Name (SCAN)

Oracle RAC 11g Release 2 also introduces the new Single Client Access Name (SCAN) to simplify client access to Oracle RAC databases. SCAN provides a single name to be used in the client connection requests that does not change as the cluster expands or if any of the nodes in the cluster changes over time. This allows the use of simplified connect strings such as EZConnect (sqlplus system/manager@sales1-scan:1521/oltp - or for JDBC: jdbc:oracle:thin:@sales1-scan:1521/oltp). When using GNS, you only need to supply the name and listener port for your SCAN. If you choose not to use GNS, you must define the SCAN in your DNS as a single name that round robins to 3 IP Addresses. These IP addresses must be on the same subnet as the public network for the cluster.

Cluster Verification Utility

Oracle Grid Infrastructure includes a cluster configuration verification tool (CVU). The cluster verification tool eliminates errors through pre- and post-validation of installation steps and/or configuration changes. It can also be used for ongoing cluster validation, for which reason the toll has been incorporated into Oracle Grid Infrastructure 11g Release 2, patch set 1 (11.2.0.2) The tool can invoked through a command line interface. With Oracle RAC 11g Release 2, the Cluster Verification Utility (CVU), is integrated with the OUI to validate configuration input during the installation interview as well as the system pre-requisites for Oracle Clusterware, ASM, and Oracle Database. If your server(s) do not meet certain necessary pre-requisites, a fix-up script is created by CVU. In order to automatically fix those pre-requisites, you can then run the fix-up script on each node that applies.

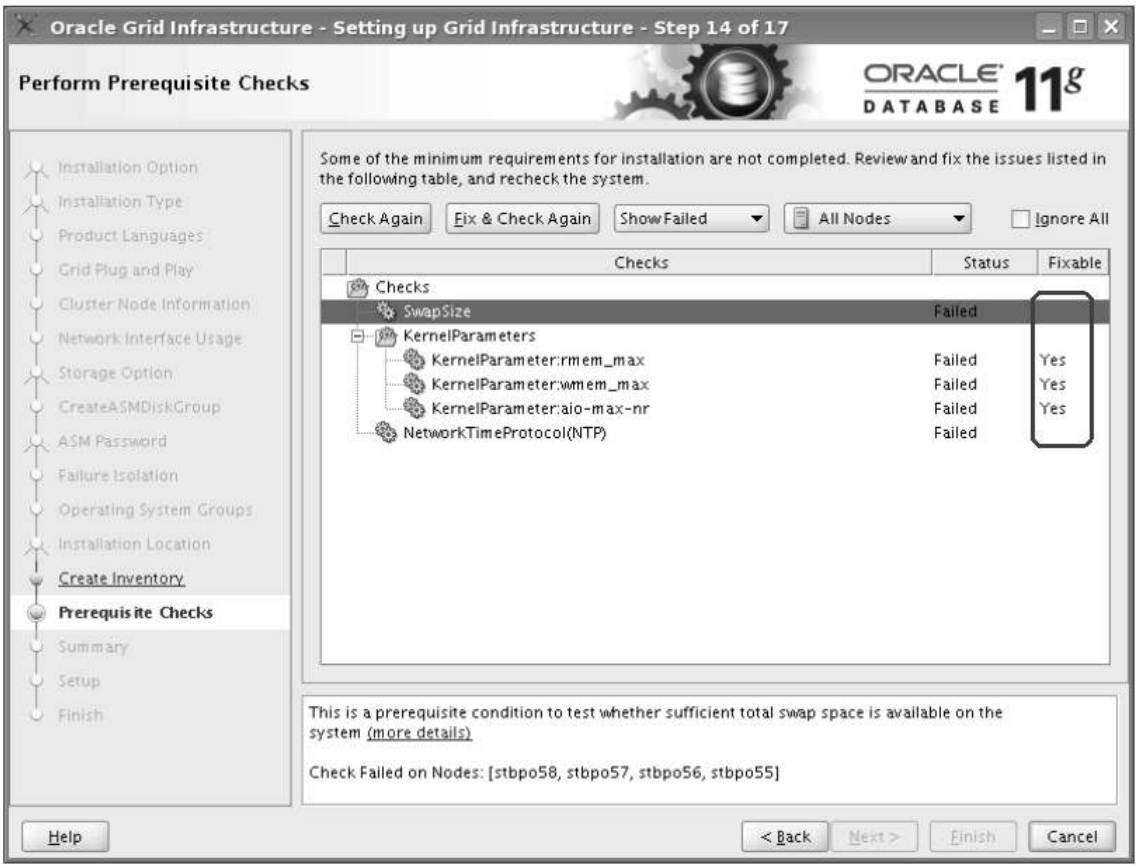

Figure 3: Oracle Universal Installer CVU Integration

"High availability is absolutely essential for us…we now use Oracle RAC for instance failover, data guard for site failover, ASM to manage our storage, and Oracle clusterware to hang the whole thing together."

— Jon Waldron, Executive Architect, Commonwealth Bank of Australia

Oracle RAC on Extended Distance Clusters

Oracle RAC on Extended Distance Clusters is an architecture where servers in the cluster reside in locations that are physically separate. Oracle RAC on Extended Distance Clusters provides extremely fast recovery from a site failure and allows for all servers, in all sites, to actively process transactions as part of a single database cluster. While this architecture creates great interest and has been successfully implemented, it is critical to understand where this architecture best fits especially in regards to distance, latency, and degree of protection it provides.

The high impact of latency, and therefore distance, creates some practical limitations as to where this architecture can be deployed. This architecture fits best where the 2 datacenters are located relatively close (<~100km) and where the extremely expensive costs of setting up direct cables with dedicated channels between the sites have already been taken.

Oracle RAC on Extended Distance Clusters provides greater availability than local Oracle RAC but it may not fit the full Disaster Recovery requirements of every organization. Feasible separation is great protection for some disasters (local power outage, airplane crash, server room flooding) but not all. Disasters such as earthquakes, hurricanes, and regional floods may affect a greater area. One should analyze their situation in order to determine if both sites are likely to be affected by the same disaster.

For comprehensive protection against disasters including protection against corruptions and regional disasters, Oracle recommends the use of Oracle Data Guard in combination with Oracle RAC as described in the Oracle Maximum Availability Architecture guidelines. Oracle Data Guard also provides additional benefits such as support for rolling upgrades across Oracle versions.

Configuring an extended distance cluster is more complex than configuring a local cluster. Specific focus needs to go into node layout, voting disks, and data disk placement. Implemented properly, this architecture can provide greater availability than a local Oracle RAC database, but should be considered with Oracle's Maximum Availability Architecture in mind.

Benefits of Oracle Real Application Clusters

High Availability

Oracle Real Application Clusters 11g provides the foundation for data centre-high availability. It is also an integral component of Oracle's Maximum Availability Architecture, which provides best practices to provide the highest availability for your data center. Oracle Real Application provides the following ke characteristics essential for a high available data management::

Reliability – The Oracle Database is known for its reliability. Oracle Real Application Clusters takes this a step further by removing the database server as a single point of failure. If an instance fails, the remaining instances in the server pool remain open and active. Oracle Clusterware monitors all Oracle processes and immediately restarts any failed component.

Recoverability – The Oracle Database includes many features that make it easy to recover from all types of failures. If an instance fails in an Oracle RAC database, it is recognized by another instance in the server pool and recovery will start automatically. Fast Application Notification (FAN) and Fast Connection Failover (FCF) or Transparent Application Failover (TAF) make it easy for applications to mask component failures from the user.

Error Detection – Oracle Clusterware automatically monitors Oracle RAC databases as well as other Oracle processes (ASM, listener, etc) and provides fast detection of problems in the environment. It also automatically recovers from failures often before users noticed that a failure has occurred. Fast Application Notification (FAN) provides the ability for applications to receive immediate notification of cluster component failures in order to re-issue the transaction before the failure surfaces.

Continuous Operations – Oracle Real Application Clusters provides continuous service for both planned and unplanned outages. If a server (or an instance) fails, the database remains open and the application is able to access data. Most database maintenance operations can be completed without downtime and are transparent to the user. Many other maintenance tasks can be done in a rolling fashion so application downtime is minimized or removed. Fast Application Notification and Fast Connection Failover assist applications in meeting service levels.

> **"We successfully completed a 'Big Bang' migration from a mainframe system to a clustered server environment. The Oracle environment is highly stable and offers powerful performance and easy scalability, which has enhanced user satisfaction and customer service."**
>
> **— NoCheol Park, CIO & Senior VP,** NGM Program Manager, SK Telecom

Scalability

Oracle Real Application Clusters provides a unique technology for scaling applications. Traditionally, when database servers ran out of capacity, they were replaced with new and larger servers. As servers grow in capacity, they are more expensive. For databases using Oracle RAC, there are alternatives for increasing the capacity. Applications that have traditionally run on large SMP servers can be migrated to run on pools of small servers. Alternatively, you can maintain the investment in the current hardware and add a new servers to the pool (or to create a server pool) to increase the capacity. Adding servers to a server pool with Oracle Clusterware and Oracle RAC does not require an outage and as soon as the new instances are started, the application can take advantage of the extra capacity. All servers in the server pool must run the same operating system and the same version of Oracle, but they do not have to be of exactly the same capacity. Customers today run server pools that fit their needs often using servers of (slightly) different characteristics.

The Oracle Real Application Clusters architecture automatically accommodates rapidly changing business requirements and the resulting workload changes. Application users, or mid tier application server clients, connect to the database by way of a service name. Oracle automatically balances the user load among the multiple nodes in the server pool. The Oracle Real Application Clusters database instances on the different nodes subscribe to all or some subset of database services. This provides the Database Administrator witjh the flexibility of choosing whether specific application clients that connect to a particular database service can connect to some or all of the database nodes. Administrators can painlessly add processing capacity as application requirements grow. The Cache Fusion architecture of Oracle RAC immediately utilizes the CPU and memory resources of the new node(s). DBAs do not need to manually re-partition the data.

Another way of distributing workload in an Oracle database is through the Oracle Database parallel execution feature. Parallel execution (I.E. parallel query or parallel DML) divides the work of executing a SQL statement across multiple processes. Using Oracle's cost-based optimizer intelligent decisions are made with respect to intra-node and inter-node parallelism in an Oracle RAC environment.

For example, if a particular query requires six query processes to complete the work and six CPUs are idle on the local node (the node that the user connected to), then the query is processed using only local resources. This demonstrates efficient intra-node parallelism and eliminates the query coordination overhead across multiple nodes. However, if there are only two CPUs available on the local node, then those two CPUs and four CPUs of another node are used to process the query. In this manner, both inter-node and intra-node parallelism are used to provide speed up for query operations.

Managing Your Oracle Real Application Clusters Database

Oracle Real Application Clusters provides a single system image for easy configuration and management. The Oracle RAC database can be installed, configured, and managed from a single location. All tools and utilities provided to manage the database are cluster-aware from the Oracle Universal Installer (OUI), to Enterprise Manager including the database configuration assistant (DBCA), the database upgrade assistant (DBUA), the network configuration assistant (NETCA), and the command line interfaces such as srvctl.

Oracle Enterprise Manager

Oracle Enterprise Manager Grid Control is the recommended management interface for an Oracle environment. Oracle Grid Control delivers centralized management functionality for the complete Oracle IT infrastructure, including systems running Oracle and non-Oracle technologies. With a broad set of administration, configuration management, provisioning, end-to-end monitoring, and security capabilities, Oracle Grid Control reduces the cost and complexity of managing private clouds, while helping customers maintain their IT infrastructure service levels.

Oracle Enterprise Manager Database Control is the graphical management tool provided by Oracle to manage your Oracle Database. Oracle Enterprise Manager Database Control can be automatically configured by the DBCA when a database is created. Alternatively, the database can be automatically registered with Oracle Enterprise Manager Grid Control at the time of creation. Both Enterprise Manger products are cluster-aware and can be used to manage Oracle RAC and server pools created with Oracle Clusterware.

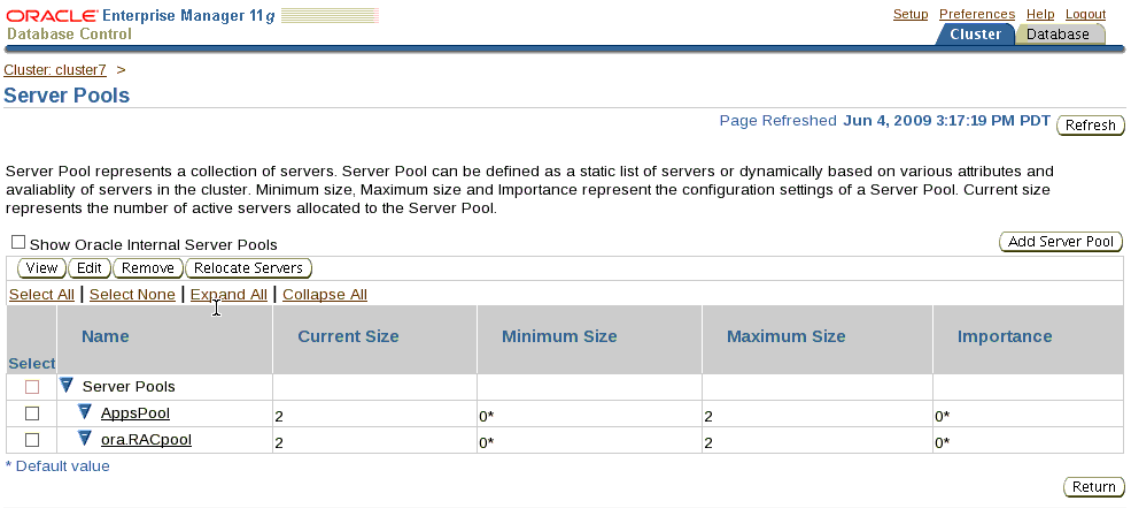

Figure 4: Oracle Enterprise Manager Database Control - Server Pool Management

Oracle Enterprise Manager Grid Control simplifies the day to day management tasks to be performed by an Oracle Database Administrator as well as Oracle Database deployments and monitoring. It assists:

- When migrating Oracle Databases from a file system to Oracle ASM
- When converting an Oracle Single Instance into an Oracle RAC Database
- When monitoring various HA areas (e.g. backup & recovery, replication, disaster recovery)
- When defined thresholds are reached (cluster-wide) so that an alert must be issued.

Starting with Oracle Enterprise Manager 10g Release 2, Grid Control provides additional capabilities to make the provisioning of Oracle Real Application Clusters databases easier. The initial creation of a server pool, including deploying the required Oracle software (a.k.a. Oracle Homes) and the configuration of Oracle clusterware can easily be performed using Oracle Enterprise Manager.

The Oracle Home-software can be kept in Oracle Enterprise Manager as the known "Gold Image" or sourced for further deployments. The "Gold Image" is created from a copy of a known, good implementation of Oracle Clusterware or Oracle Real Application Clusters. Thereby, Oracle Enterprise Manager supports the complete end-to-end creation of new Oracle RAC and Oracle Grid Infrastructure environments, including the execution of super-user actions (root.sh) and customizable pre- and post-installation steps. Similar support is provided when extending the server pools with more servers.

For Linux operating systems, Oracle can also provision an "image" to a bare metal node. The image can consist of the Operating System, the Oracle Enterprise Manager agent, Oracle Grid Infrastructure, and Oracle Database with Oracle Real Application Clusters. This image can be associated with a hardware profile. All the components for this image are stored as "Gold Images" in Oracle Enterprise Manager.

Oracle Enterprise Manager Database Control introduces a new area for server pool and cluster management. The "cluster-tab" allows you to monitor and manage a server pool created with Oracle Clusterware. From the cluster-tab, you can create and manage resources for both Oracle as well as user created resources. Users can create resources with dependencies to allow the cluster to monitor and managed any process on any server in the cluster. Oracle Clusterware includes sophisticated dependency options to allow you to easily model all your business critical applications. Oracle Enterprise Manager Database Control 11g Release 2 requires an Oracle RAC 11g Release 2 database. However, it can manage Oracle Clusterware and its resources even when the database is down.

Rolling Patch Application

Oracle supports the application of patches to the nodes of an Oracle RAC database in a rolling fashion with no downtime. Patches are applied one node at a time while the other nodes in the Oracle RAC system are up and operational. This requires that each node has a separate Oracle Home. Patches will be labeled as being qualified for installation as rolling upgradeable, or not, depending on the changes being made by the patch. Some patches that modify common structures shared between instances, or the contents of the database, will not be.

This capability is supported beginning with Oracle 9.2.0.2. All Oracle Clusterware patches can be applied in a rolling fashion. With Oracle Database 11g, Automatic Storage Management is rolling upgradeable. Oracle Grid Infrastructure 11g Release 2, which includes Oracle Clusterware and Oracle ASM, is rolling upgradeable for all sorts of patches, including patch sets. Also, with Oracle Database 11g Release 2, Oracle has streamlined the application of patches in a cluster. The patch utility Opatch, will apply patchsets and patch bundles in a few easy steps.

Rolling Release Upgrade Support

Oracle Grid Infrastructure 11g Release 2 supports rolling upgrades; so does Oracle Clusterware, which has offered this feature starting with Oracle Clusterware 10g. This provides the ability to upgrade without taking the entire cluster down and therefore enables 24x7 operation of business. Once you have upgraded Oracle Automatic Storage Management to 11g, you will be able to do rolling upgrades with ASM, too.

Oracle RAC 11*g* supports database software upgrades (from Oracle Database 10g Release 1 Patchset 1 onwards) in a rolling fashion – with nearly zero database downtime, by using Oracle Data Guard SQL Apply. The steps involve upgrading the logical standby database to the next release, running in a mixed mode to test and validate the upgrade, doing a role reversal by switching over to the upgraded database, and then finally upgrading the old primary database. While running in a mixed mode for testing purpose, the upgrade can be aborted and the software downgraded, without data loss. For additional data protection during these steps, a second standby database may be used.

By supporting rolling upgrades with minimal downtimes, Oracle Data Guard reduces the large maintenance windows typical for many administrative tasks, and enables 24x7 operation of business.

Workload Management with Oracle Real Application Clusters

Applications using an Oracle RAC database need to manage workload across a server pool. Oracle Real Application Clusters includes innovative technologies to manage workloads while providing the best application throughput given the configuration and high availability for the application. Oracle RAC 11g Release 2 introduces a simplified management of Oracle RAC databases within server pools to ease the consolidation of applications into a pool of servers, while maintaining resource allocation and role separated management at the same time.

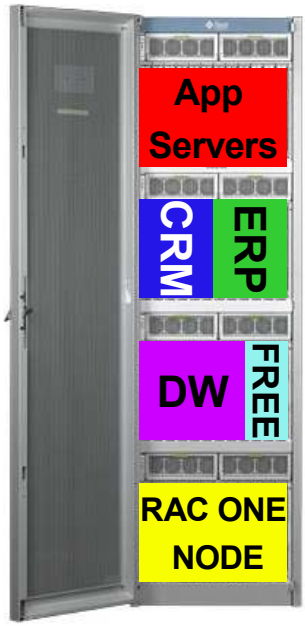

Figure 5 Consolidate into Low Cost Servers using Server Pools

Server Pools

With Oracle RAC 11g Release 2, a database can be defined to run in a server pool. A server pool is a logical entity in the cluster that allows the administrator to allocate resources to specific applications. A server pool is defined by 3 attributes: min (the minimum number of servers in the pool, max (the maximum number of servers in the pool), and importance (providing the ability to give a relative importance to different pools in the cluster). Oracle Clusterware will allocate servers to user defined pools when a cluster reconfiguration takes place. Oracle Clusterware will allocate servers in order of the importance. The number of instances maintained for the database is defined by the cardinality of the server pool.

The database configuration assistant (DBCA) provides the option to create a policy-managed database by defining the cardinality and the name. A server pool will automatically be created trying to allocate as many servers and therefore instances of the database as specified by the cardinality parameters requested. Oracle Clusterware will maintain the servers in the server pool up to the maximum based on availability of servers in the cluster.

Services

Workload Management relies on the use of Services, a feature of the Oracle Database. Services hide the complexity of an Oracle RAC database by providing a single system image to manage workload. Services allow applications to benefit from the reliability of a server pool. Traditionally, a database provided a single service and this name was used in the connect data for SQL*NET. With Oracle Database 11g, a DBA can define more than a hundred database services for a single database.

This allows you to breakup workloads from applications into manageable components based on business requirements such as service levels and priorities. Services are integrated with many features of the Oracle Database. Application users can – for example – automatically be assigned to a Resource Manager consumer group, which can limits their consumed resources such as CPU, depending on a service. Batch jobs can be assigned to specific job class based on their service. The use of services also achieves location transparency for queues when using Oracle Streams Advanced Queuing, while inter-node parallel query will be restricted to the instances where the service is active with Oracle Database 11g.

A service can span one or more instances of an Oracle database and an instance can support multiple services. The number of instances offering a service is managed dynamically by the DBA independently of the application. When outages occur, services are automatically restored to surviving instances. When instances are restored, any services that are not running are restored automatically.

With Policy Managed Databases, a service can only run in one server pool and is defined as either uniform (service is offered by all instances in the server pool) or as a singleton (runs only on one instance in the server pool).

Connection Load Balancing

Oracle Net Services provide connection load balancing for database connections. Client-side load balancing, which balances connection requests across all SCAN listeners in the cluste, is achieved by using the SCAN on the address list of the client connect string. SQL*NET will randomly select one of the SCAN ip addresses. If the server chosen is not available, the next server in the list is tried. Server side load balancing is achieved using the SCAN listener. Each SCAN listener is aware of all instances in the cluster providing each service. Based on goal defined for the service, the listener chooses the instance that will best meet the goal and the connection is routed to that instance through the local listener.

"The clustered database architecture that we've deployed has satisfied our business and performance requirements. It also offers us the flexibility we need for future growth. The newly re-architected Global Data Warehouse features, in addition to more powerful processors, improved performance and provide additional capacity for applications and users."

— Matthew Schroeder, Manager, Business Information & Technologies, Alcoa

Fast Application Notification

Fast Application Notification (FAN) provides integration between the Oracle RAC database and the application. It allows the application to be aware of the current configuration of the server pools at any given time so that application connections are only made to those instances that are currently able to respond to the application requests. The Oracle RAC HA framework posts a FAN event immediately when a state change occurs within the cluster.

Integrated clients receive these events and immediately react. For DOWN events, application interruption is minimized by cleaning up connections to the failed instance, in-flight transactions are interrupted with an error returned to the application. Applications making connections are directed to active instances only. Server side callouts can be used to log trouble tickets or page administrators alerting them of the failure.

For UP events, new connections are created to allow the application to immediately take advantage of the extra resources available. Oracle JDBC, the Oracle Universal Connection Pool, ODP.NET and OCI clients are integrated with FAN. Applications with their own connection pools can take advantage of FAN by using either the Oracle RAC FAN API available with the Oracle Database 11g Release 2 JDBC driver or the Oracle Call Interface callback functionality.

Load Balancing Advisory

Since Database workloads can change over time as well as the server pool configuration, it is important to create and allocate database connections based on the most current information. Oracle Real Application Clusters provides a load balancing advisory. Oracle RAC constantly monitors the workload being executed for each service by each instance providing the service. This information is published to the Automatic Workload Repository and published to the application using FAN events. The FAN event includes the current service level provided and a recommendation of what percentage of connections to be directed to each instance.

The integrated Oracle Clients use these events to provide intelligent load balancing of application requests. Most connection pools use a random or round robin algorithm to select an idle connection from the pool when the application attempts to connect. Using FAN events from the load balancing advisory, the connection pool will select the connection currently providing the best service. Oracle JDBC, Oracle UCP, OCI, and ODP.NET provide runtime connection load balancing through integration with the load balancing advisory.

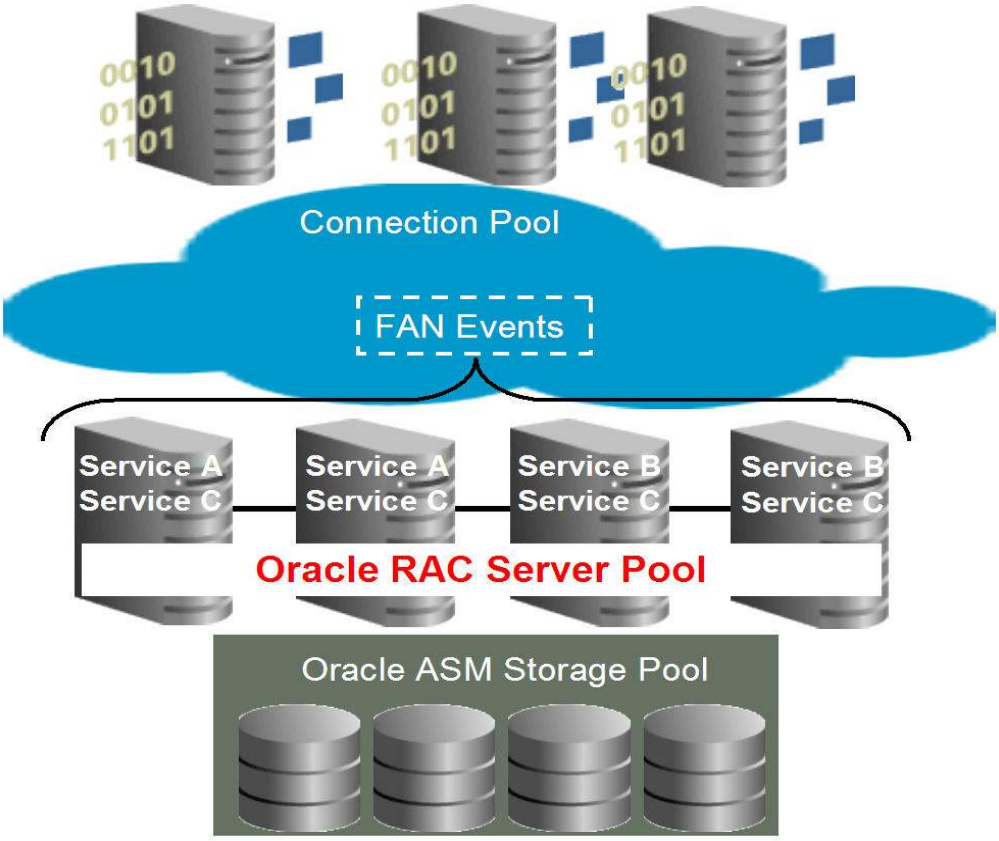

Figure 6: Load Balancing in an Oracle RAC stack

Conclusion

Oracle Real Application Clusters has been designed for high availability and scalability. By providing protection from hardware and software failures, Oracle Real Application Clusters provides system availability ensuring continuous data access. Its scale out and scale up features offer a platform, which can grow in any direction allowing enterprises to grow their businesses. Existing applications as well as newly developed applications benefit from the transparency Oracle Real Application Clusters provides.

Application development as well as administration and change management thus become much easier, allowing reduction in total cost of ownership. Oracle Real Application Clusters is unique to the market with its offering and capabilities. Oracle RAC is used by thousands of customers worldwide in all industries in mission critical and many other application environments and a crucial component when setting up or managing private clouds.

Oracle Real Application Clusters (RAC)
11g Release 2
November 2010

Author: Barb Lundhild, Markus Michalewicz
Contributing Authors:

Oracle Corporation
World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065
U.S.A.

Worldwide Inquiries:
Phone: +1.650.506.7000
Fax: +1.650.506.7200
oracle.com

Oracle is committed to developing practices and products that help protect the environment

0109