Technical Report

Virtualizing Oracle RAC on Red Hat Enterprise Virtualization 3.1 and NetApp Clustered Data ONTAP

Jon Benedict, NetApp
August 2013 | TR-4216

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 Executive Summary

This NetApp® solution enables companies to optimize their virtualized Oracle® Real Application Clusters (Oracle RAC) database infrastructure by providing advanced storage and data management capabilities. NetApp's industry-leading storage solutions provide the scalability, flexibility, and availability necessary for a dynamic organization.

This solution guide provides guidelines and best practices for architecting, deploying, and managing Oracle RAC on NetApp clustered Data ONTAP® storage systems. Example scenarios, validation procedures, and implementation guidelines are described in detail. Best practices for implementation and operation are provided as needed.

Additionally, Red Hat Enterprise Virtualization (RHEV) was used as the virtualization platform, providing a predictable high-performance experience.

1.1 Purpose

This document describes the design, implementation, and validation of an Oracle on NetApp clustered Data ONTAP solution in detail. It can be used in both presale and postsale scenarios. It is not intended to be one definitive solution to how such an environment can be implemented.

To enable all of the customer's needs to be met and to realize the entire value of Oracle on NetApp clustered Data ONTAP, it's imperative to reference a detailed survey of a customer's operation.

1.2 Objective

The objective is to deliver an overall solution for Oracle that utilizes all the features of NetApp clustered Data ONTAP 8.1 to address challenges such as:

- Providing the same level of resiliency at the storage layer that Oracle RAC provides at the database layer.
- Avoiding operational disruption caused by tech refreshes, upgrades, and component failures.
- Making effective use of all storage resources by redistributing data and production loads.
- Consolidating storage and database environments.

This solution is designed to consider various use cases to support different customer requirements. The solution incorporates key technology components and products from Red Hat, NetApp, and Oracle, where appropriate and effective.

1.3 Audience

The target audience for this document includes sales engineers, field consultants, professional services personnel, IT managers, partner engineering personnel, and customers who want to deploy Oracle solutions on clustered Data ONTAP 8.1.

1.4 Previous Work

The design and validation performed in this document are based on a similar project: TR-3979: Red Hat Enterprise Linux 6, KVM, and NetApp Storage: Best Practices Guide for Clustered Data ONTAP. In some instances, the work performed was identical, except for the hypervisor and virtualization management used. Therefore, this document uses much of the previous material.

1.5 Requirements and Assumptions

For the methods and procedures described in this document to be useful, the reader must have basic understanding of the following:

- NetApp administration skills and administrative access to the storage systems using the command-line interface
- Data ONTAP clustered configurations
- Oracle products and technologies
- Red Hat Enterprise Linux® (RHEL) and Red Hat Enterprise Virtualization (RHEV)

2 Scope

The scope of this document is to provide architectural, deployment, and management guidelines for those who are planning or have already decided to virtualize Oracle RAC on RHEV and clustered Data ONTAP. The document provides a brief overview of the clustered Data ONTAP technology concepts; key solution architecture considerations for implementation; storage estimation and data layout recommendations; and solution, deployment, and management guidelines.

Note: Raw performance testing is not within the scope of this document. Representative loads were generated in order to test the features of clustered Data ONTAP 8.1 and record the effects, if any, while operating Oracle. For performance information, see TR-3961: Oracle Database 11g Release 2 Performance Using Data ONTAP 8.1 Operating in Cluster-Mode, although it might not be completely representative of the configuration that was used in this technical report.

2.1 Summary of Use Cases

The following use cases are identified as key to leveraging the primary benefits of operating Oracle Database products in a clustered Data ONTAP environment. These use cases will be used to drive validation testing and documentation of an Oracle on clustered Data ONTAP solution.

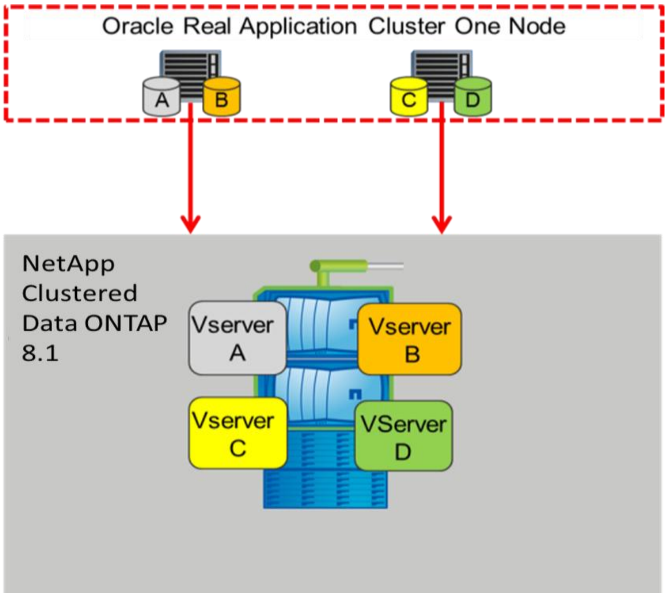
Storage Migration and Consolidation

As companies grow, multiple Oracle Database and application environments can develop within the organization; each has management, availability, and growth needs. With the advent of virtualization, many organizations are undertaking consolidation projects to reduce cost and complexity and increase the efficiency of the IT infrastructure. This consolidation has resulted in increased service level and growth demands. Clustering, including Oracle Real Application Clusters, has addressed the availability and scalability needs at the server level. The same consolidation needs to take place at the storage layer for the same reasons as those previously mentioned. NetApp clustered storage systems provide the flexibility, efficiency, and scalability to meet and exceed those needs.

Database Consolidation

NetApp clustered Data ONTAP technologies can combine with Oracle Real Application Cluster One Node to provide a scalable, resilient environment into which databases can be consolidated.

Figure 1) Database consolidation.

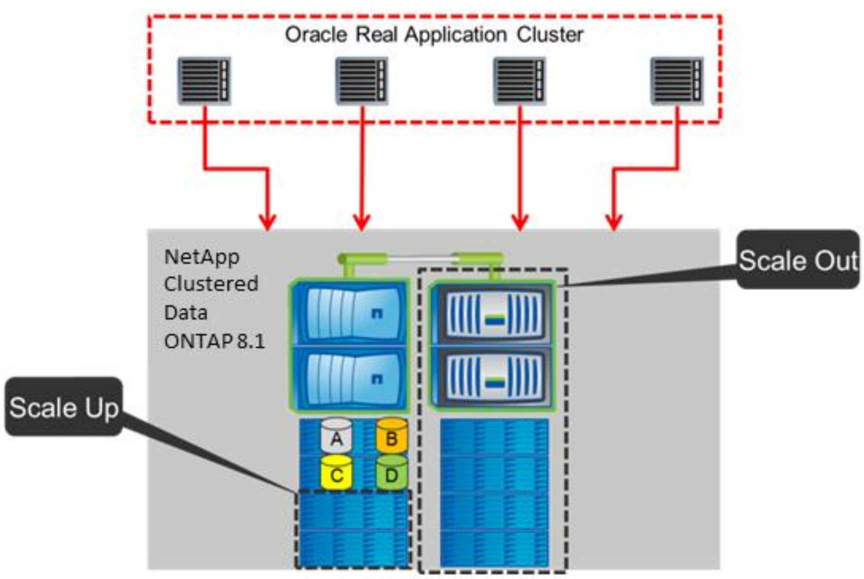


Dynamic Scalability with Online Scale-Up and Scale-Out

A critical consideration that arises as a result of storage consolidation is that the availability and resiliency of that consolidated storage now become even more vital. When many departments and applications are relying on the availability of their resources, the ability to scale storage up and out is not only desirable but required. With clustered Data ONTAP, scaling up is achieved by adding disk arrays to existing controllers to increase IOPS and capacity. Additional and larger storage systems can be added to the existing storage cluster to meet scaling-out demand. The online nondisruptive storage expansion and reconfiguration capability of clustered Data ONTAP makes it the industry-leading scale-out storage platform.

Although not part of this use case, Flash Cache™ cards can also be applied for better read performance with most random access workloads, of which approximately 80% tend to be read operations.

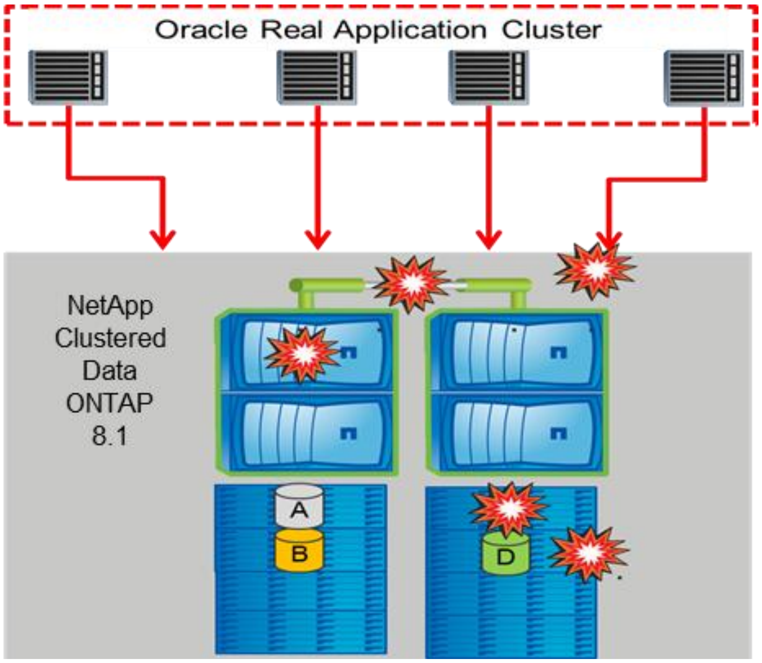
Figure 2) Scale-up and scale-out.



Seamless Failover Protection

Storage failover protection is a core attribute of clustered Data ONTAP. A high-availability (HA) pair of controllers are the building blocks that together form the storage cluster. This architecture enables transparent controller clustering and failover capability in which a failed storage controller causes its partner node to take over its disk arrays, volumes, and running services to provide continuous operation.

Figure 3) Failover protection.



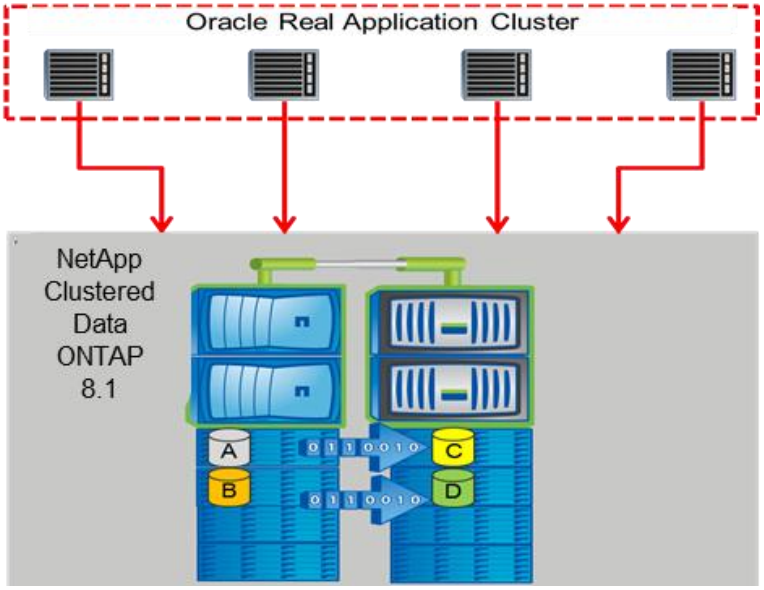
Live Migration of Data Volumes

Clustered Data ONTAP is equipped with DataMotion™ for Volumes software, which is a nondisruptive transparent volume move feature. Volumes can be transparently moved between storage systems within the storage pool without requiring coordination between applications or application downtime. This single feature has many benefits, including:

- Eliminating application outages
- Providing always-on data access
- Enabling load and resource tiering
- Improving resource utilization
- Facilitating dynamic scaling
- Phased upgrades (tech refresh)

The DataMotion for Volumes feature of clustered Data ONTAP can also be used to facilitate online load balancing, to adjust the service-level demand, and to manipulate resource utilization. The logical interface migration capability of Data ONTAP 8.1 allows the IP address to be moved with the volume enabling continued primary local access of the data.

Figure 4) Data volume migration.



Operational Flexibility

DataMotion for Volumes in clustered Data ONTAP provides the following:

- Better average utilization across the storage pool by selectively moving and redesignating volumes to different aggregates and controllers within the cluster. Increased demand of input/output (I/O) can be met by relocating volumes to aggregates made up of a higher number of disks or faster drive types.
- Performance and capacity gains to meet new service levels can be achieved by relocating volumes to higher performance aggregates or controllers. The controllers can be of a different class, have higher specifications, or be equipped with SSD and Flash Cache.
- Some storage systems might suffer from sustained high load, or certain volumes might be specifically hard hit, which could result in significant hot spots and unsatisfactory levels of service to applications. Moving the subject volumes to different storage nodes could alleviate the problems.

If Oracle Clusterware and RAC need to be migrated to a new storage system, DataMotion for Volumes can be used to move not only the database volumes but also the crucial Clusterware volumes that store OCR and voting disks, without any disruption to the operation of Oracle Clusterware and RAC databases.

Best Practices

Customers rely on NetApp to provide all the guidelines and tools necessary for installation and configuration of an optimum Oracle environment in terms of both performance and availability. This would include:

- Database layout
- Configuration parameters
- Monitoring methods

Protocol Performance

NetApp support for multiple protocols provides customers with the flexibility necessary for a dynamic environment. Quite often in the implementation planning stages, customers rely on NetApp to not only support multiple protocols but also provide data on how these protocols might perform in their environment in various situations (moving a volume, failover, and so on). In this version of the solution the storage protocol used is NFSv3. However, it should be noted that dNFS (Oracle Direct NFS) and LUNs (FC, FCoE, or iSCSI) are all supported and valid options. For performance information, see [TR-3961: Oracle Database 11g Release 2 Performance Using Data ONTAP 8.1 Operating in Cluster-Mode](#).

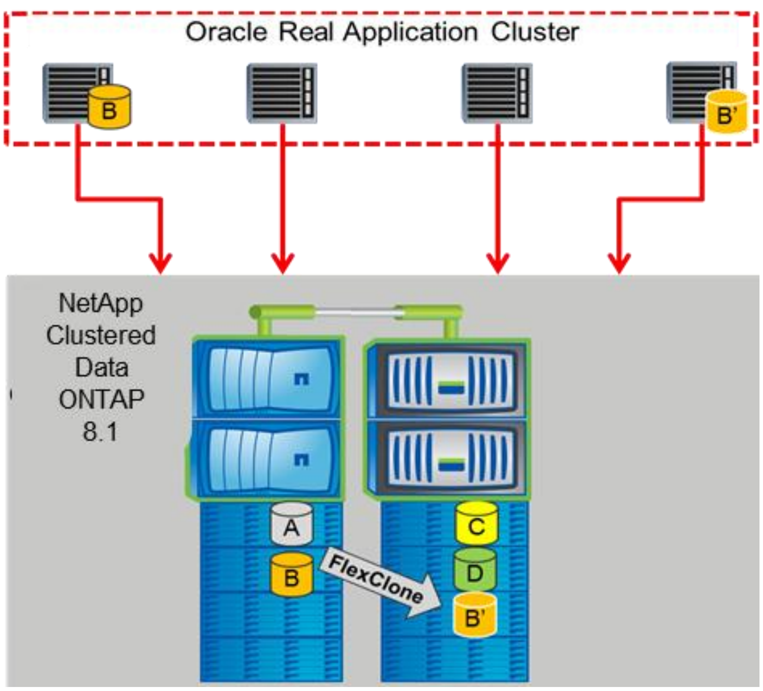
Application Development and Test

Application development and test environments are faced with challenges such as:

- Testing upgrades without disruption to production
- Limited storage for test and development purposes
- Environmental considerations for extra hardware needed
- Lengthy application development and test cycles
- Copying data can take a lengthy amount of time
- Large number of copies of datasets is required

NetApp offers unique advantages when it comes to cloning of a production Oracle environment for development and test. With space-efficient and point-in-time Snapshot™ copies and cloning, the amount of time to deploy and the storage required for application development and test are greatly reduced. This enables multiple space-efficient clones to be created for other purposes such as reporting or training.

Figure 5) Oracle RAC on clustered Data ONTAP.



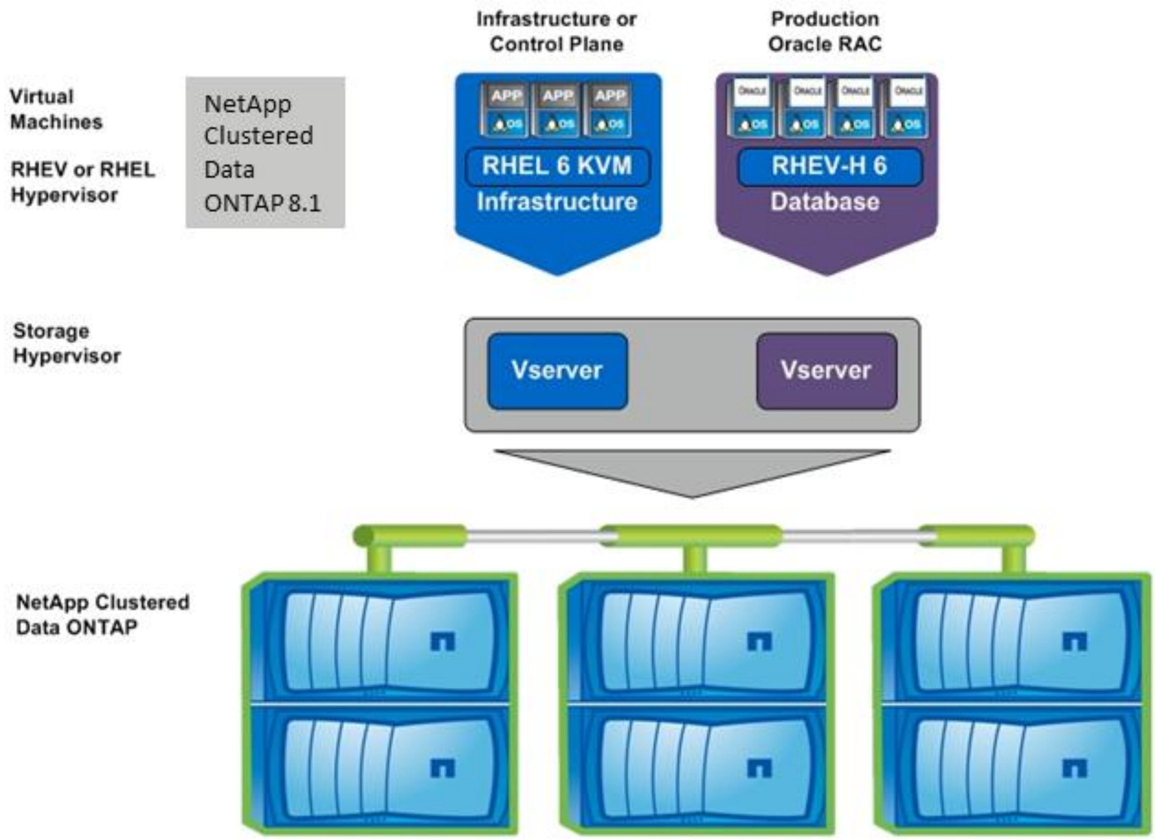
3 Solution Overview

3.1 General Description

The intent of the configuration described in the following sections is to approximate a production environment as realistically as possible, especially in the context of virtualizing Oracle 11g RAC with Red Hat Enterprise Virtualization. The solution architecture shown in sections that follow.

Figure 6 can be broken down into the areas described in the sections that follow.

Figure 6) Solution architecture.



Oracle

A four-node Oracle 11g R2 Real Application Cluster (RAC) was used to provide the scalability and availability necessary at the Oracle RDBMS level.

Physical Servers

Several high-density blade servers based on Intel® were used to host the hypervisors in both the “production Oracle RAC environment” and the “control plane.”

Virtual Servers

In this solution the four Oracle RAC nodes are each hosted in a virtual machine running Red Enterprise Linux 6 update 4. These virtual machines in turn were hosted by Red Hat Enterprise Virtualization 3.1.

Storage

The storage controllers used for this solution were multiple FAS32470 nodes running Data ONTAP 8.1 operating in clustered Data ONTAP. Internal to the cluster were two redundant dedicated interconnect switches used for the cluster network. The storage type used was DS2246 disk shelves using SAS disk drives. This environment could easily use FAS32xx or FAS62xx storage controllers, depending on the requirements.

Software

Outside of the Oracle software the following software was used:

- NetApp SnapManager® for Oracle to facilitate ease of backup and cloning for development/test purposes
- Quest Benchmark Factory to facilitate load generation

3.2 Components

Table 1) Software components.

Vendor	Product/Solution	Version	Description
Oracle	RDBMS	11g R2 (11.2.0.3)	Highlighted application for scale-up/scale-out demonstration
Oracle	Real Application Cluster	11g R2	
NetApp	Data ONTAP	8.1.2	
NetApp	OnCommand System Manager	2.0.1	
NetApp	SnapManager for Oracle	3.2	Used for backup, recovery, and cloning
Red Hat	Enterprise Linux	6.4	Guest OS for Oracle nodes

The NetApp FAS3240, the NetApp storage platform used in this solution, is part of the FAS3000 series.

Table 2 lists the details of the storage.

Table 2) Storage components.

NetApp Fabric-Attached Storage (FAS) 3200 Series	
Component	4 FAS3240s
Number of disks	24 per controller
Disk size	450GB
Drive type	SAS
Shelf type	DS2246
Number of shelves	1 per controller

3.3 NetApp Benefits

NetApp enables its customers to move to a fully virtualized compute and storage platform to provide the market's best database consolidation platform with the highest level service level agreements (SLAs) and operational flexibility and efficiency for the Oracle Database environment. Benefits include:

- Elimination of business disruptions due to infrastructural upgrades and maintenance
- Reduction of capital expense by consolidation of storage and databases
- Nondisruptive movement of workloads across the physical storage layer to meet demanding SLAs
- Facilitation of migration from NetApp Data ONTAP 8.0 operating in 7-Mode to clustered Data ONTAP 8.1 through the use of Oracle Data Guard
- Seamless failover protection against component failure

Storage Efficiency

- **RAID-DP®** technology, a high-performance double-parity version of RAID 6, protects against double disk failure. It is integrated with the WAFL® file system to prevent dedicated parity drives from becoming a performance bottleneck. It provides higher availability than RAID 1 (mirroring), RAID 0+1 (striping + mirroring), and RAID 5 while incurring no performance penalty.

- **Snapshot** technology is available from a variety of data storage vendors, but not all snapshot copies are created equal. NetApp Snapshot technology enables IT administrators to create point-in-time copies of virtual machines or entire datastores. Using SnapRestore® technology, you can restore from these backup copies at any level of granularity—single files, directories, or entire volumes—simply and quickly when required. Copies can be made at any time increment in less than one second with no performance impact, no matter how many Snapshot copies are taken. These are not full copies of data; they only track changes and are very efficient in terms of overall storage capacity. Restores can be done rapidly from any of the copies, providing customers with an exceptional recovery time objective (RTO).
- NetApp **FlexVol®** thin provisioning helps customers:
  - Avoid the costly overprovisioning and time-consuming reconfiguration that are typical with other storage solutions.
  - Reduce storage provisioning time from days to minutes.
  - Allocate storage on demand for LUNs, volumes across all protocols (FCP, iSCSI, FCoE, and NAS), and operating systems according to application needs.

  With traditional storage provisioning, disk capacity is allocated to applications regardless of how much data is actually being written. Space requirements are rarely known early in a deployment; therefore, to make sure that their applications will have adequate capacity, users often present inflated estimates of capacity requirements. This forces administrators to overprovision the larger storage infrastructure.

  As your capacity needs evolve, FlexVol volumes can dynamically grow and shrink to meet exactly those needs. As you decrease volumes, their capacity returns to the pool for allocation to other volumes as needed.
- Using NetApp **FlexClone**® technology, administrators can create instant writable Snapshot copies to support application testing. Using FlexClone, you can affordably create as many clones as needed to speed product development. Unlike full copies from mirrored production data, FlexClone copies:
  - Can be created almost instantly.
  - Take very little space.
  - Have negligible performance impact.
  - With FlexClone, you can affordably create as many clones as needed to speed product development.
- With NetApp **deduplication**, customers can store just one copy of each unique data object, substantially reducing capacity requirements. Deduplication automatically removes duplicate data blocks on a 4KB level across an entire volume, reclaiming wasted storage to achieve significant space savings.
- Manageability:
  - A single point of management
  - Policy driven
  - Tracking and monitoring
  - User interfacing tools for backup/restore/cloning/mirroring

Large enterprises utilizing Oracle Database products (Single Instance, Real Application Clusters, and RAC One Node) are faced with meeting a diverse set of application service levels and a continual growth in the amount of data. Competitive pressures and economic constraints only add to the challenges.

These enterprises can benefit from a flexible, scalable, and efficient storage infrastructure for Oracle that allows workloads to be relocated around the physical storage layer based on capacity and/or performance requirements without any disruption to business operations.

4 Solution Implementation

4.1 Overview

Implementation was broken down into the following general areas:

- Servers
- Storage
- Database
- Software

4.2 Server/Host Installation

All servers were Red Hat Enterprise Linux x86_64 running inside virtual machines hosted by Red Hat Enterprise Virtualization 3.1. The following steps were taken while setting up the environment.

1. Install and configure 2 “thick” RHEL 6 hypervisors, as described in TR-4104: Best Practices for RHEL 6, KVM, and Clustered Data ONTAP. These will be used to host the “infrastructure servers” (control plane), including RHEV-M.
2. Install and configure an RHEL 6 virtual machine on one of the thick hypervisors, as described in TR-4104: Best Practices for RHEL 6, KVM, and Clustered Data ONTAP. Install RHEV-M on this virtual machine, as described in Red Hat’s official documentation.
3. Install and configure two or more RHEV-H “thin” hypervisors, as described in Red Hat’s official documentation.
4. From RHEV-M, create virtual machines that will host the Oracle servers. For purposes of this test, four virtual machines were created.
5. Install and configure Red Hat Enterprise Linux on each of the virtual machines created, as described in TR-4104: Best Practices for RHEL 6, KVM, and Clustered Data ONTAP.
6. Install any necessary RPMs required by the Oracle installation. Refer to the “Oracle Linux Installation guide” for specific RPMs and procedures for installation.

Note: All thick hypervisors and virtual machines must be registered to Red Hat Network (RHN) and must be updated. All thick hypervisors, thin hypervisors, and virtual machines must be configured to synchronize their times using the Network Time Protocol (NTP).

4.3 Storage Installation

Tables 3 through 6 describe the storage systems and layout used in the validation of this solution.

Table 3) Storage components.

Type	Purpose	Notes
FAS3240	HA pair	Initial two-node cluster
FAS3240	HA pair	Add-on nodes for cluster

Table 4) Storage layout.

Controller	Aggregate	Aggr Size	Volume
FVL1-cluster-01	aggr0_FVL1-cluster-01	492GB	vol0
	aggr01	7.2TB	vol0
	“”		Infra_Boot

Controller	Aggregate	Aggr Size	Volume
	“”		root_vol
	“”		root_vol_m01 (SVM 1)
	“”		root_vol_m01 (SVM 2)
	“”		oraarch_flexdb
	“”		orabin_flexdb
	“”		oracntrl_flexdb
	“”		oracrs_flexdb_ocr
	“”		oracrs_flexdb_vote
	“”		oradata_flexdb_sm
	“”		oraredo_flexdb
	“”		oratemp_flexdb
FVL1-cluster-02	aggr0_FVL1-cluster-012	492GB	vol0
	aggr02	7.2TB	Root_vol
	“”		Prod_Boot (SVM 1)
	“”		root_vol_m02 (SVM 1)
	“”		root_vol_m02 (SVM 2)
FVL1-cluster-03	aggr0_FVL1-cluster-012	492GB	vol0
	aggr03	7.2TB	root_vol
	“”		Infra_NFS (SVM 1)
	“”		root_vol_m03 (SVM 1)
	“”		root_vol_m03 (SVM 2)
	“”		Prod_ISO (SVM 2)
FVL1-cluster-04	aggr0_FVL1-cluster-012	492GB	vol0
	aggr04	7.2TB	vol0
	“”		root_vol_m04 (SVM 1)
	“”		root_vol_m04 (SVM 2)

Note: All Oracle related volumes are part of the “production” storage virtual machine (SVM 1).

Table 5) Storage virtual machines.

Storage Virtual Machine	Type	LIF	Port	Notes
FVL1-cluster	admin	cluster_mgmt	e0a	
FVL1-cluster-01	node	clus1	e1a	

Storage Virtual Machine	Type	LIF	Port	Notes
	node	clus2	e2a	
	admin	mgmt1	e0b	
FVL1-cluster-02	node	clus1	e0a	
	node	clus2	e1a	
	admin	mgmt1	e0b	
FVL1-cluster-03	node	clus1	e1a	
	node	clus2	e2a	
	admin	mgmt1	e0b	
FVL1-cluster-04	node	clus1	e1a	
	node	clus2	e2a	
	admin	mgmt1	e0b	
Infrastructure (SVM 1)	cluster	nfs_lif01	e1b	NFS server
“”	cluster	nfs_lif02	i0a-3081	NFS server
“”	cluster	nfs_lif03	i0a-3081	NFS server
“”	cluster	nfs_lif04	i0a-3081	NFS server
“”	admin	vsmgmt	e0a	
Production (SVM 2)	cluster	nfs_lif01	e1b	NFS server
	cluster	nfs_lif02	i0a-3081	NFS server
	cluster	nfs_lif03	i0a-3081	NFS server
	cluster	nfs_lif04	i0a-3081	NFS server
	admin	vsmgmt	e0a	

Table 6) NFS exports.

Vserver	Volume	Export Name	Options
Infrastructure	/Infra_NFS	/Infra_NFS	rw,vers=3,rsize=65536,wsize=65536,hard,proto=tcp,timeo=600,addr=172.20.81.13
Production	/Prod_ISO	/Prod_ISO	rw,vers=3,rsize=65536,wsize=65536,hard,proto=tcp,timeo=600,addr=172.20.81.17 (through .20)
Production	/Prod_NFS	/Prod_NFS	rw,vers=3,rsize=65536,wsize=65536,hard,proto=tcp,timeo=600,addr=172.20.81.17 (through .20)
Production	/oratemp_flexdb	/oracle/oratemp	rw,bg,hard,vers=3,proto=tcp,timeo=600,rsize=65536,wsize=65536,nointr,actimeo=0,addr=172.20.81.15
Production	/oraredo_flexdb	/oracle/oralog	rw,bg,hard,vers=3,proto=tcp,timeo=600,rsize=65536,wsize=65536,nointr,actimeo=0,addr=172.20.81.15

Vserver	Volume	Export Name	Options
Production	/oradata_flexdb	/oracle/oradata	rw,bg,hard,vers=3,proto=tcp,timeo=600,rsize=65536,wsize=65536,nointr,actimeo=0,addr=172.20.81.15
Production	/oradata_flexdb_sm	/oracle/oradata_sm	rw,bg,hard,vers=3,proto=tcp,timeo=600,rsize=65536,wsize=65536,nointr,actimeo=0,addr=172.20.81.15
Production	/oracrs_flexdb_vote	/oracle/oracrsvote	rw,bg,hard,vers=3,proto=tcp,timeo=600,rsize=65536,wsize=65536,nointr,actimeo=0,addr=172.20.81.15
Production	/oracntrl_flexdb_ocr	/oracle/oracrsocr	rw,bg,hard,vers=3,proto=tcp,timeo=600,rsize=65536,wsize=65536,nointr,actimeo=0,addr=172.20.81.15
Production	/oracntrl_flexdb	/oracle/oracntrl	rw,bg,hard,vers=3,proto=tcp,timeo=600,rsize=65536,wsize=65536,nointr,actimeo=0,addr=172.20.81.15
Production	/orabin_flexdb	/u01	rw,bg,hard,vers=3,proto=tcp,timeo=600,rsize=65536,wsize=65536,nointr,actimeo=0,addr=172.20.81.15
Production	/oraaarch_flexdb	/oracle/oraarch	rw,bg,hard,vers=3,proto=tcp,timeo=600,rsize=65536,wsize=65536,nointr,actimeo=0,addr=172.20.81.15

In the steps that follow, it is assumed that the installation and configuration of all hardware are complete.

1. Install Data ONTAP 8.1 if not already installed.
2. Install the following licenses:
   - GX_Base
   - GX_Mirror
   - GX_FlexVol_HPO
   - GX_SnapRestore
   - NFS
   - SnapMirror_DP
   - FlexClone

```
FVL1-cluster::system> license show
Feature          Cluster SN   Limit    Description
---------------  -----------  -------  -----------
Base             1-80-000011    4      Base License w/cluster size limit (nodes)
Mirror           1-80-000011    4      Mirror License
CIFS             1-80-000011    4      CIFS License
SnapRestore      1-80-000011    4      SnapRestore License
FlexClone        1-80-000011    4      FlexClone License
NFS              1-80-000011    4      NFS License
SnapMirror_DP    1-80-000011    4      SnapMirror Data Protection License
7 entries were displayed.
```

Note: Make sure the licenses indicate a node count that will accommodate the nodes to be added later.

Note: Configure aggregates Vservers and volumes according to Table 6.

3. Add the NFS exports as outlined in Table 6.
4. Configure each Linux server to NFS mount the appropriate volumes using the options specified in section “Linux Configuration” in the appendix.
5. Configure the pair for HA.

A two-node cluster is a special case for the storage failover (SFO) mechanism and requires an additional setting. If you have more than two nodes in a cluster, each node must be rebooted prior to testing/using storage failover capabilities.

Note: The cluster HA mode (two-node only) enables storage failover and also the auto-sendhome feature by default. In clusters that have more than two nodes, the auto-sendhome feature defaults to off.

4.4 Database Installation

Prerequisites

Before deploying Oracle 11g R2 RAC, check that the following conditions are met:

- The hardware requirements meet or exceed Oracle's requirements, including CPU, memory, networking, and out-of-band management.
- The network requirements meet or exceed Oracle's requirements, including one public IP address, one private IP address, and three single client access name (SCAN) addresses per RAC node.

Best Practice

Sizing for Oracle on Data ONTAP 8.1 is the same as for Data ONTAP 7G or 8.0 operating in 7-Mode with one caveat:

- In a cluster with mixed platforms, it is best to size for the smallest platform in the cluster. By doing this, a potential performance degradation caused by moving Oracle from a more powerful node to a less powerful node can be avoided.

Oracle RAC User Deployment Steps

1. Create the necessary Oracle groups, including `oinstall`, `oracle`, and `dba`.
2. Create the necessary Oracle users, including `grid` and `oinstall`, and add the groups created in the previous step to the oinstall user.
3. Set the password for the `oracle` and `grid` user accounts.

Oracle RAC Networking Deployment Steps

1. Determine the cluster name, public host names, public virtual host names, and private host name for each of the RAC nodes, all according to Oracle requirements.
2. Determine a SCAN DNS name that resolves to the three SCAN IP addresses.
3. Add the public, private, and virtual IP addresses to the /etc/hosts file. Repeat this step on all RAC nodes.

Note: Do not put the SCAN IP address in the /etc/hosts file.

4. Configure all RAC nodes to synchronize time using the NTP.

Tuning Kernel Parameters and Shell Limits

1. Follow Oracle's requirements on kernel parameters and make sure that the changes survive a reboot.
2. Repeat the kernel parameter changes on all RAC nodes.
3. Add or increase parameters in the /etc/security/limits.conf /etc/pam.d/login and /etc/profile files per Oracle requirements.

Creating Required Files and Directories

1. Create Oracle inventory files and directories.
2. Create the Oracle Grid Infrastructure home directory.
3. Create the Oracle Base directory.
4. Create the Oracle RDBMS home directory.

Installing the Oracle Grid Infrastructure

1. Make sure that the shared storage is mounted.
2. As the `grid` user, start the `runInstaller` script.
3. Continue through the installer, entering information as prompted.
4. When prompted for the installation type, select "Install and Configure Oracle Grid Infrastructure for a Cluster."
5. Continue through the installer, entering information as prompted, until the installation is complete.

Install the RDBMS Software

1. Switch to the oracle user account.
2. Run the installer script to deploy the Oracle RDBMS, entering information as prompted.
3. When prompted for the installation type, select "Oracle Real Application Clusters database installation" and select all of the Oracle RAC nodes created earlier.
4. Continue the installation, entering information as prompted.
5. When the RDBMS installation is complete, run the Database Configuration Assistant (DBCA) to create the necessary databases.

This section summarizes those NetApp products that have some type of integration component with other vendors' products.

4.5 Snap Creator Framework

Snap Creator™ Framework leverages proven NetApp technologies to support multiple applications using a single interface:

- **Data protection.** The NetApp Snap Creator Framework addresses the needs and challenges of administrators and developers by providing a centralized, consistent solution for backing up critical information, one that integrates with existing application architectures to reduce costs and enable a higher ROI.
- **OS independence.** Our design approach delivers functionality that is OS independent to help alleviate the need to manage interoperability issues.
- **Extensibility.** Snap Creator offers fast integration and greater storage abstraction to leverage our policy-based automation features and better integrate with your current and future application needs.

Snap Creator Framework provides plug-ins to integrate NetApp features with popular third-party applications such as Oracle, MySQL, and DB2. And if you have special applications that are not part of our supported offering, it's no problem. Snap Creator Framework allows the development of custom plug-ins to provide backup procedures for applications not covered by existing supported applications.

See the NetApp Snap Creator Framework 4.0 Installation and Administration Guide available at the NetApp Support site.

4.6 SnapManager for Oracle

NetApp SnapManager for Oracle (SMO) automates and simplifies the complex, manual, and time-consuming processes associated with the backup, recovery, and cloning of Oracle Databases. SnapManager leverages NetApp technologies, such as Snapshot, SnapRestore, and FlexClone, while integrating with the latest Oracle Database releases. SnapManager also integrates seamlessly with native Oracle technologies such as Oracle Real Application Clusters (Oracle RAC), Automatic Storage Management (ASM), and Direct NFS and across FC, iSCSI, and NFS protocols.

This integration allows IT organizations to scale their storage infrastructure, meet increasingly stringent SLA commitments, and improve the productivity of database and storage administrators across the enterprise. Backups created using SMO can optionally be cataloged with Oracle Recovery Manager (RMAN) to preserve the value of RMAN functions, such as block-level restore or tablespace point-in-time recovery.

For more information, refer to TR-3761: Best Practices for SnapManager 3.2 for Oracle.

4.7 Selecting the Right Backup and Recovery Application

In order to determine which backup and recovery solution is best suited for your environment, here are some key points to consider:

Snap Creator

- A free offering.
- Integrates with SnapMirror® and SnapVault®.
- Flexible configuration to handle simple and complex backup, cloning, and recovery actions.
- Single pane of glass to manage and schedule backups.
- Can integrate with SnapDrive® for UNIX® or OnCommand® Unified Manager, if required.

SnapManager for Oracle

- Deep backup and recovery integration (SnapDrive for UNIX, OnCommand Unified Manager, and so on)
- Cloning automation
- Catalog of backups
- Integrates seamlessly with Oracle technologies
- Additional license cost

5 Solution Validation

The procedures used to validate the functionality of the solution are described in this section. A summary is provided in Table 7.

Table 7) Test cases.

Test Case	Description
UC01	Live migration of data volumes
UC02	Live migration of RHEV storage domain
UC03	Load balancing

5.1 Live Migration of Data Volumes (UC01)

This test involves the live migration of a NetApp volume that houses an Oracle Database. This happens transparently to the running database and users and is most relevant to scaling out the database across storage nodes or consolidating databases to fewer storage nodes.

Summary	Clustered Data ONTAP includes DataMotion for Volumes, which is a nondisruptive transparent volume move feature. Volumes can be transparently moved between storage systems within the storage pool, much like live migrating a virtual machine between hypervisors. This single feature has many benefits such as: • Eliminating application outages • Providing always-on data access • Enabling load and resource tiering • Improving resource utilization • Facilitating dynamic scaling • Providing phased upgrades (tech refreshes) • Load balancing to meet service levels
Hardware setup	• Four-node FAS3240 clustered Data ONTAP deployment
Preconditions	• Existing fully configured NetApp Data ONTAP 8 operating in clustered Data ONTAP cluster containing at least one HA pair (additional pairs possibly necessary for extensions) • Proper licenses installed on NetApp cluster on operational Oracle Database
Test plan	1. An existing database in the clustered Data ONTAP environment 2. Move the database volumes and data LIFs together 3. Move the database volumes when the LIF is local and remote 4. All operations are carried out while the Oracle Database load is running
Proof points	• Capture the performance impact while the database volume move is in progress. • Capture the statistics from the storage side to see the impact on these operations. • Capture the database server performance while the operation is in progress.
Results/capture issues	• Database applications operate as before without interruptions. • Volumes are moved and redesignated to different controllers within the cluster to drive a better average utilization across the storage pool. • Volume move is not only achieved for database volumes, but also for the crucial Clusterware volumes that store OCR. Voting disks are moved without any disruption to the operation of Oracle Clusterware and databases. • The performance impact on volume move was minimal (less than 10% in terms of TPS and response time for a 500GB volume) while the Oracle Database load was running. • Recommendation is to carry out volume move operation (one at a time) during nonpeak hours.
Workload	OLTP (TPC-C) tests: Benchmark Factory (BMF)

5.2 Live Migration of RHEV Storage Domain (UC02)

This test involves the live migration of a NetApp volume that houses the virtual disks for the virtualized Oracle RAC nodes. This happens transparently to the running database and users and is most relevant to load balancing the Oracle RAC nodes across the storage cluster.

Summary	Clustered Data ONTAP includes DataMotion for Volumes, which is a nondisruptive transparent volume move feature. Volumes can be transparently moved between storage systems within the storage pool, much like live migrating a virtual machine between hypervisors. This single feature has many benefits such as: • Eliminating application outages • Providing always-on data access • Enabling load and resource tiering • Improving resource utilization • Facilitating dynamic scaling • Providing phased upgrades (tech refreshes) • Load balancing to meet service levels
Hardware setup	• Four-node FAS3240 clustered Data ONTAP deployment
Preconditions	• Existing fully configured operational clustered Data ONTAP 8 cluster containing at least one HA pair (additional pairs possibly necessary for extensions) • Proper licenses installed on NetApp cluster • An existing Red Hat Enterprise Virtualization 3.1 environment using NetApp storage for data storage domains (volume that contains the virtual disks for the virtualized Oracle RAC nodes)
Test plan	1. Leverage an existing RHEV deployment that virtualizes the Oracle RAC nodes. 2. An existing database in the clustered Data ONTAP environment as deployed on RHEV. 3. Move the storage domain volume across different storage controllers. 4. Move the storage domain volumes and data LIFs together. 5. Move the storage domain volumes when the LIF is local and remote. 6. All operations are carried out while the Oracle Database load is running.
Proof points	• Capture the performance impact while the database volume move is in progress. • Capture the statistics from the storage side to analyze the impact on these operations. • Capture the database server performance while the operation is in progress. • Visually monitor the health of the virtualized RAC nodes in RHEV.
Results/capture issues	• Database applications operate as before without interruptions. • Volumes are moved and redesignated to different aggregates and controllers within the cluster to drive a better average utilization across the storage pool. • Volumes move is achieved for RHEV storage domain volumes. • The performance impact on volume move was minimal (less than 10% in terms of TPS and response time for a 500GB volume) while the Oracle Database load was running. • The virtualized RAC nodes as well as the RHEV storage domain continued to operate without disruption. • Recommendation is to carry out volume move operation (one at a time) during off-peak hours.
Workload	OLTP (TPC-C) tests: Benchmark Factory (BMF)

5.3 Load Balancing (UC03)

This test involves the live migration of a NetApp volume that houses an Oracle Database. This happens transparently to the database and users and is most relevant to load balancing the database(s) across multiple storage nodes.

Summary	• DataMotion for Volumes in clustered Data ONTAP provides the following: • Volumes can be selectively moved and redesignated to different aggregates and controllers within the cluster to drive a better average utilization across the storage pool. • Increased demand in input/output (I/O) can be met by relocating volumes to aggregates made up of a higher number of disks or faster drive types. • Performance and capacity gains to meet new service levels can be achieved by relocating volumes to higher performance aggregates or controllers. The controllers can be of a different class, have higher specifications, or be equipped with SSD and Flash Cache. • Some storage systems might suffer from sustained high load, or certain volumes might be specifically hard hit, which could result in significant hot spots and unsatisfactory levels of service to applications. Moving the subject volumes to different storage nodes could alleviate the problems. • If Oracle RAC needs to be migrated to a new storage system, DataMotion for Volumes can be used to move database volumes, Clusterware volumes that store OCR, and voting disks. This can be done without any disruption to the operation of Oracle RAC databases. • Network load can be balanced automatically through LIF load balancing.
Hardware setup	• Four-node FAS3240 clustered Data ONTAP deployment
Preconditions	• Existing fully configured operational NetApp clustered Data ONTAP 8 operating in cluster containing at least one HA pair • Proper licenses installed on NetApp cluster • Operational Oracle Database
Test plan	1. Leverage an existing database setup in clustered Data ONTAP environment. 2. Get the baseline performance with initial database layout. 3. Modify the database volume layout to span across multiple nodes in the storage cluster. 4. All operations are carried out while the Oracle Database load is running.
Proof points	• Capture the performance impact with the database volumes spread across multiple storage nodes. • Capture the statistics from the storage side to see the impact on these operations. • Capture the database server performance while the operation is in progress.
Results/capture issues	• Database applications operate as before without interruptions. • Database data volumes were moved seamlessly across different storage nodes. • Logical components tied up with Vservers, that is, data LIFs, are also moved along with volumes without any disruptions to Oracle Clusterware and databases. • Most of the Oracle scale-up and scale-out features are compatible with NetApp storage clustered Data ONTAP scale-up and scale-out features.
Workload	OLTP (TPC-C) tests: Benchmark Factory (BMF)

6 Summary

The ability to manage multiple physical storage systems as a single storage entity is a true benefit of clustered Data ONTAP. Any object or resource (such as disks, aggregates, volumes, networks, I/O ports, and services) within the storage cluster can be managed, administered, and monitored from a single point of management. Resources can be accessed and shared cluster-wide. Secure virtual containers called Storage Virtual Machines (SVMs) consisting of storage resources can be created per application where volumes and data access are isolated and restricted to the virtual containers.

Database and storage consolidation can be achieved by virtualizing Oracle RAC with Red Hat Enterprise Virtualization (RHEV) and Data ONTAP 8.1 clustered Data ONTAP, a scale-out and dynamic storage platform. RHEV provides an open source virtualization solution that includes record performance and ease of management. With clustered Data ONTAP, scaling up is achieved by adding disk arrays to existing controllers to increase IOPS and capacity. Additional and larger storage systems can be added to the existing storage cluster to meet scaling-out demands. The online nondisruptive storage expansion and reconfiguration capability of clustered Data ONTAP make it an industry-leading scale-out storage platform.

NetApp support for multiple protocols provides customers with the flexibility necessary for a dynamic environment. Quite often in the implementation and planning stages, customers rely on NetApp not only to support multiple protocols but also to provide data on how these protocols might perform in their environment in various situations (moving a volume, failover, and so on).

NetApp offers unique advantages when it comes to cloning of a production Oracle environment for development and test through Snap Creator. With space-efficient point-in-time Snapshot copies and cloning, the amount of time to deploy and the storage required for application development and test are greatly reduced. This enables multiple space-efficient clones to be created for other purposes such as reporting or training.

Overall, the architecture, features, and value proposition within NetApp clustered Data ONTAP are a complement to RHEV, Oracle RAC, as well as Oracle RAC solutions that are built on RHEV and clustered Data ONTAP.

7 References

7.1 Recommended Support Documents

Refer to the following documents for more information on this solution.

- TR-3961: Oracle Database 11g Release 2 Performance Using Data ONTAP 8.1 Operating in Cluster-Mode
  www.netapp.com/us/system/pdf-reader.aspx?pdfuri=tcm:10-59751-16&m=tr-3961.pdf
- TR-3633: Oracle Best Practices Guide
  www.netapp.com/us/system/pdf-reader.aspx?pdfuri=tcm:10-60340-16&m=tr-3633.pdf
- Data ONTAP 8.1 Software Setup Guide
- Data ONTAP 8.1 System Administration Guide
- Data ONTAP 8.1 Storage and Data Protection Guide
- Data ONTAP 8.1 Network and File Access Management Guide
- Snap Creator Framework 4.0 Installation and Administration Guide
  https://library.netapp.com/ecm/ecm_get_file/ECMP1149421
- Oracle 11g documentation
  http://www.oracle.com/technetwork/indexes/documentation/index.html

- TR-4104: Best Practices for RHEL 6, KVM, and Clustered Data ONTAP
  www.netapp.com/us/media/tr-4104.pdf
- Red Hat Enterprise Linux product documentation
  https://access.redhat.com/site/documentation/Red_Hat_Enterprise_Virtualization/

Appendix

Oracle Configuration

Table 8) Oracle initialization parameters (init.ora).

Parameter Name	Value	Description
_in _memory _undo	FALSE	Make in-memory undo for top-level transactions
_undo_autotune	FALSE	Enable autotuning of undo _retention
Compatible	11.2.0.0.0	Database is completely compatible with this software version
cursor _space _for _time	TRUE	Use more memory for faster execution
db _block _size	8192	Size of database block in bytes
db _cache _size	5251268608	Size of DEFAULT buffer pool for standard block-size buffers
db _file _multiblock _read _count	128	Database block to be read each I/O
db_files	123	Maximum allowable number of database files
db_name	tpcc	Database name specified in CREATE DATABASE
disk_asynch_io	TRUE	Use async I/O for random access devices
dml _locks	500	One dml locks for each table modified in a transaction
filesystemio _options	setall	I/O operations on file system files
log_buffer	1048576	Redo circular buffer size
parallel_max_servers	100	Maximum parallel query servers per instance
plsql _optimize _level	2	PL/SQL optimize level
processes	1500	User processes
recovery_parallelism	40	Number of server processes to use for parallel recovery
sessions	1000	User and system sessions
shared_pool_size	393216000	Size in bytes of shared pool
statistics_level	Typical	Statistics level
transactions	2000	Maximum number of concurrent active transactions
undo_management	AUTO	If TRUE, instance runs in SMU mode; otherwise in RBU mode
undo _retention	10800	Undo retention in seconds
undo_table space	undo_1	Use or switch undo tablespace

Linux Configuration

Table 9) Linux nondefault kernel parameters.

Parameter Name	Value	Description
kernel.sem	250 32000 100 128	Semaphores
net.core.rmem_max	4194304	Maximum TCP receive window size (maximum buffer size)
net.ipv4.tcp_rmem	4096 16384 4194304	Memory reserved for TCP receive buffers
net.core.rmem_default	4194304	Default TCP receive window size (default buffer size)
net.core.wmem_max	4194304	Maximum TCP send window size (maximum buffer size)
net.ipv4.tcp_wmem	4096 16384 4194304	Autotuning for TCP send window size (default and max. values are overridden by wmem_default wmem_max)
net.core.wmem_default	4194304	Default TCP send window size (default buffer size)
net.ipv4.ip_local_port_range	9000 65500	Local port range used by TCP and UDP
net.core.netdev_max_backlog	300000	Maximum number of packets, queued on the INPUT side, when the interface receives packets more quickly than kernel can process them
fs.file-max	6815744	Maximum number of file handles that the Linux kernel allocates
fs.aio-max-nr	1048576	Maximum number of allowable concurrent requests
sunrpc.tcp_slot_table_entries	128	Maximum number of (TCP) RPC requests that can be in flight

Table 8) Other Linux configuration settings.

Configuration File	Settings	Description
/etc/security/limits.conf	oracle hard nofile 65536	Shell limits for Oracle
/etc/sysconfig/cpuspeed	governor=performance	
/etc/selinux/config	SELINUX=disabled	
	Service irqbalance stop chkconfig irqblance off	Disable irqbalance

Best Practices

RHEV Best Practices

- Use VLANs to separate data traffic from storage traffic. Use additional VLANs to separate public and private data, as well as different storage types.
  - If VLANs are used, you must explicitly configure a VLAN for RHEV-M, or channel bonding will be difficult.
  - After the initial VLAN is created on an interface, it must be synchronized before channel bonding can be configured.
- Oracle Databases are not usually heavy CPU consumers and therefore are not characterized as CPU-bound applications. Therefore, NetApp recommends that you use fewer vCPUs while creating a VM. Ideally the vCPU for a VM is not more than the number of physical CPUs in the host. However,

NetApp suggests that you start with two or three vCPUs and later add more if required. If you add too many vCPUs to an Oracle VM, then the unused vCPUs will consume time interrupts, which will be an overhead.

- Use the latest VirtIO network adapter and VirtIO disk driver.
- Use thick provisioning on the virtual machines, but thin provision the storage volumes.
- Do not use RHEV to manage database volumes. Mount all database-related volumes directly from the virtualized RAC nodes to the NetApp storage.
- Use 10GbE wherever possible.
- Virtualize RHEV-M and other management applications on two or more RHEL 6 KVM hypervisors attached to NetApp storage. This will allow for maximum flexibility and easier backup and recovery, as well as site failover.

Virtual machines running Oracle should be in their own RHEV cluster, with as few hypervisors as possible, in order to mitigate Oracle licensing and maintenance costs.

Oracle Best Practices

- Place Oracle Database data and log volumes separately on different aggregates depending on disk types. For example, place the data and log volumes on faster disk type aggregates and place the archive and other files on slower disk type aggregates.
- Database volumes can be moved seamlessly within the storage cluster without disrupting the complete database stack. A best practice is to move one volume at a time at nonpeak work hours.
- kNFS and dNFS are tested features of NetApp Data ONTAP storage operating in clustered Data ONTAP for database applications. A dNFS best practice is to configure multiple paths from database hosts to storage on different subnets.

NFS Mounting Best Practices

- It is always recommended to mount a logical interface (LIF) that is local to the Oracle volumes. This provides better performance improvements.
- The NFS clients need to follow the recommended mount options as documented at https://kb.netapp.com/support/index?page=content&id=3010189, while mounting the Oracle volumes.
- In case of volume mobility using “vol move” done on any Oracle volume, it is recommended that the corresponding LIF that was mounting that volume be migrated manually, as listed in the section “Linux Configuration” in the appendix.
- Sizing for Oracle on Data ONTAP 8.1 is the same as for Data ONTAP 7G or 8.0 operating in 7-Mode with one caveat: In a cluster with mixed platforms it is best to size for the smallest platform in the cluster. By doing this, you can avoid a potential performance degradation caused by moving Oracle from a more powerful node to a less powerful node.

For other best practices, refer to TR-3633: Best Practice Guidelines for Oracle 11*g* Databases.

Version History

Date	Version	Description	Author(s)
September 2013	0.1	Initial release	Jon Benedict

Refer to the Interoperability Matrix Tool (IMT) on the NetApp Support site to validate that the exact product and feature versions described in this document are supported for your specific environment. The NetApp IMT defines the product components and versions that can be used to construct configurations that are supported by NetApp. Specific results depend on each customer's installation in accordance with published specifications.