



Technical Report

# Oracle Database 11g Release 2 Performance Using Data ONTAP 8.1 Operating in Cluster-Mode

Saad Jafri, Chris Lemmons, NetApp  
October 2011 | TR-3961

## A PERFORMANCE COMPARISON

This technical report compares the performance of the Oracle<sup>®</sup> 11g R2 Real Application Cluster (RAC) and single-instance databases using an OLTP workload running in NetApp<sup>®</sup> Data ONTAP<sup>®</sup> 8.1 7-Mode and Cluster-Mode environments in both bare metal and virtualized configurations. This report demonstrates the readiness of Data ONTAP operating in Cluster-Mode environments to provide a high-performance environment for Oracle databases.

## TABLE OF CONTENTS

<b>1</b>	<b>INTRODUCTION AND EXECUTIVE SUMMARY</b>	<b>4</b>
1.1	INTRODUCTION	4
1.2	AUDIENCE	4
<b>2</b>	<b>TEST CONFIGURATIONS</b>	<b>4</b>
<b>3</b>	<b>SUMMARY OF TEST RESULTS</b>	<b>5</b>
<b>4</b>	<b>RESULTS AND ANALYSIS</b>	<b>7</b>
4.1	OLTP WORKLOAD	7
4.2	DETAILED TEST RESULTS	7
<b>5</b>	<b>CONFIGURATION DETAILS</b>	<b>9</b>
5.1	GENERIC CONFIGURATION	9
5.2	ORACLE SINGLE-INSTANCE CONFIGURATION USING DATA ONTAP OPERATING IN 7-MODE	9
5.3	ORACLE RAC CONFIGURATION USING DATA ONTAP 8.1 OPERATING IN 7-MODE	11
5.4	ORACLE RAC CONFIGURATION USING DATA ONTAP 8.1 OPERATING IN CLUSTER-MODE	17
<b>6</b>	<b>CONCLUSION</b>	<b>22</b>
<b>7</b>	<b>APPENDIXES</b>	<b>22</b>
7.1	BEST PRACTICE SUMMARY FOR DATA ONTAP 8.1 OPERATING IN 7-MODE	22
7.2	HARDWARE	23
7.3	STORAGE LAYOUTS FOR ALL TESTING CONFIGURATIONS	24
7.4	ORACLE INITIALIZATION PARAMETERS FOR SINGLE-INSTANCE CONFIGURATION TESTING	27
7.5	ORACLE INITIALIZATION PARAMETERS FOR RAC CONFIGURATION TESTING	28
7.6	LINUX KERNEL PARAMETERS FOR SINGLE-INSTANCE DATABASE CONFIGURATION TESTING	29
7.7	LINUX KERNEL PARAMETERS FOR RAC CONFIGURATION TESTING	30
7.8	OTHER LINUX OS SETTINGS	31
<b>8</b>	<b>ACKNOWLEDGMENTS</b>	<b>31</b>

## LIST OF TABLES

Table 1)	Hardware for the DNFS network using a 10GbE NetApp UTA	10
Table 2)	Hardware for the Oracle RAC configuration using Data ONTAP 8.1 operating in 7-Mode with either two or four physical servers using DNFS over 10GbE	13
Table 3)	Hardware for the Oracle RAC configuration using Data ONTAP 8.1 operating in Cluster-Mode with either two or four physical servers using DNFS over 10GbE	19
Table 4)	Best practice summary for Data ONTAP operating in 7-Mode	22
Table 5)	Database server hardware specifications for Oracle single-instance testing	23
Table 6)	Database server hardware specifications for Oracle RAC testing	23
Table 7)	NetApp FAS3270 storage system specifications	23

Table 8) Oracle initialization parameters for single-instance database configuration testing. ....	27
Table 9) Oracle initialization parameters for RAC configuration testing. ....	28
Table 10) Linux nondefault kernel parameters for single-instance database configuration testing. ....	29
Table 11) Linux nondefault kernel parameters for RAC configuration testing. ....	30
Table 12) Linux shell limits for Oracle. ....	31

## LIST OF FIGURES

Figure 1) OETs and db file sequential read latencies. ....	8
Figure 2) OETs and CPU utilization by test configuration. ....	9
Figure 3) DNFS network using a 10GbE NetApp UTA. ....	10
Figure 4) Oracle RAC configuration using Data ONTAP 8.1 operating in 7-Mode with four physical servers. ....	12
Figure 5) Oracle RAC configuration using Data ONTAP 8.1 operating in 7-Mode with two physical servers. ....	13
Figure 6) Oracle RAC configuration using Data ONTAP 8.1 operating in Cluster-Mode with four physical servers. ....	18
Figure 7) Oracle RAC configuration using Data ONTAP 8.1 operating in Cluster-Mode with two physical servers. ....	19
Figure 8) Volume layout for single-instance testing. ....	24
Figure 9) DNFS volume layout for Oracle RAC testing using Data ONTAP 8.1 operating in 7-Mode. ....	25
Figure 10) DNFS volume layout for Oracle RAC testing using Data ONTAP 8.1 operating in Cluster-Mode. ....	26

# 1 INTRODUCTION AND EXECUTIVE SUMMARY

## 1.1 INTRODUCTION

To help NetApp customers select the most appropriate solutions for their IT infrastructure, NetApp provides up-to-date documentation describing its products. This technical report contains performance information and tuning recommendations for the Oracle Database 11g Release 2 (Oracle 11g R2) running in a Data ONTAP 8.1 Cluster-Mode environment. Customers can use this updated information to make informed decisions about the optimal storage configuration for their Oracle database deployments.

The test configuration environments described in this report consist of the following elements:

- The Oracle 11g R2 RAC on Red Hat Enterprise Linux® 5 U4
- NetApp storage systems running Data ONTAP 8.1 operating in 7-Mode and in Cluster-Mode
- Bare metal and virtualized configurations using VMware® ESX 5.0

The protocols used in this technical report are kernel Network File System (kNFS) and Oracle Direct Network File System (DNFS) running over 10GbE connections. This report makes a distinction between two different Network File System (NFS) clients. The NFS client provided by the Red Hat operating system is referred to as kNFS. The second NFS client, referred to as DNFS, is provided by Oracle 11g R2. It runs as part of the Oracle database and is optimized specifically for database workloads. Therefore, DNFS can be used only to access the Oracle database files.

This report demonstrates that NetApp storage systems provide a solid environment for running the Oracle RAC database independently of whether a bare metal or virtualized environment is used as part of a data center consolidation. This report also demonstrates that a Data ONTAP 8.1 Cluster-Mode environment is capable of providing the same excellent performance that NetApp customers expect from Data ONTAP operating in 7-Mode, but with all the benefits that a clustered storage environment delivers.

## 1.2 AUDIENCE

The audiences for this report are:

- NetApp customers and employees who are investigating whether to deploy their Oracle databases using NetApp storage running Data ONTAP 8.1 in Cluster-Mode
- Customers who want to deploy their Oracle database servers in VMware virtualized hosts

This report can also be used as a performance-tuning reference.

# 2 TEST CONFIGURATIONS

Our tests focused on measuring and comparing the performance of the Oracle 11g R2 single-instance and RAC databases that use an online transaction processing (OLTP) workload connected to NetApp FAS3270 storage controllers running Data ONTAP 8.1 in either 7-Mode or Cluster-Mode configurations. Additionally, we investigated the performance impact of running the Oracle RAC configurations in a virtualized environment by configuring the RAC nodes in virtual machines running VMware vSphere™ 5.0.

All test configurations used the NFS protocol between the database servers and the NetApp FAS3270 storage controllers. Additionally, the Oracle DNFS protocol was used from within the database servers. Oracle DNFS is an optimized NFS client that provides faster and more scalable access to NFS storage located on network-attached storage (NAS) devices. DNFS is accessible over TCP/IP.

DNFS is built directly into the database kernel and provides better performance than the NFS driver in the host system by bypassing the operating system and generating only the requests required to complete the tasks at hand.

All tested configurations used QLogic 8152 10GbE converged network adapters (CNAs) in the database servers connected to NetApp 10GbE unified target adapters (UTAs) installed in the FAS3270A storage controllers.

The tests evaluated the following configurations:

- DNFS using a single-instance Oracle 11g R2 physical database server accessing the NetApp FAS3270A storage controllers configured in 7-Mode.
- DNFS using a 4-node Oracle RAC implementation configured on four physical servers accessing the NetApp FAS3270 storage controllers configured in 7-Mode.
- DNFS using a 4-node Oracle RAC implementation configured on four physical servers accessing the NetApp FAS3270A storage controllers configured in Cluster-Mode.
- DNFS using a 4-node Oracle RAC implementation configured in a vSphere 5.0 environment using four physical servers. In this case, each physical server contained a single virtual machine (VM) in which the Oracle RAC software was installed. The FAS3270A storage controllers were configured using Data ONTAP 8.1 operating in 7-Mode.
- DNFS using a 4-node Oracle RAC implementation configured in a vSphere 5.0 environment using four physical servers. In this case, each physical server contained a single VM in which the Oracle RAC software was installed. The FAS3270A storage controllers were configured using Data ONTAP 8.1 in Cluster-Mode.
- DNFS using a 4-node Oracle RAC implementation configured in a vSphere 5.0 environment using two physical servers. In this case, each physical server contained two VMs in which the Oracle RAC software was installed. The FAS3270A storage controllers were configured using Data ONTAP 8.1 in 7-Mode.
- DNFS using a 4-node Oracle RAC implementation configured in a vSphere 5.0 environment using two physical servers. In this case, each physical server contained two VMs in which the Oracle RAC software was installed. The FAS3270A storage controllers were configured using Data ONTAP 8.1 operating in Cluster-Mode.

The NetApp Data ONTAP 8.1 operating system, the latest release in the Data ONTAP family, was used in conjunction with 64-bit aggregates to demonstrate the performance capabilities it offers to NetApp customers. All tests were conducted using jumbo frames.

An OLTP workload was used for the performance tests. This workload simulated 250 users interacting with 8,000 product warehouses in an order-processing application. The client processes for the OLTP application were executed on a separate application server (client-server mode). We used the number of order entry transactions (OETs) completed per minute as the primary metric to measure application throughput. The input/output (I/O) mix for the OLTP workload was approximately 60% reads and 40% writes.

### 3 SUMMARY OF TEST RESULTS

As Data ONTAP 8.1 Cluster-Mode configurations become available, it is likely that the Oracle user base will have questions about the performance implications of moving to Cluster-Mode. The test results in this report can help answer the following questions about the transition of Oracle databases from Data ONTAP 8.1 7-Mode to Cluster-Mode and about the virtualization of Oracle database environments that use both Data ONTAP 8.1 7-Mode and Cluster-Mode:

**What is the performance difference between the Oracle single-instance and the Oracle 4-node RAC databases that use 7-Mode?**

Answer: The overall performance generated using the 4-node RAC configuration on bare metal servers was comparable to the performance generated using the same workload against the single-instance configuration.

**What is the performance impact of migrating a 4-node RAC implementation from 7-Mode to Cluster-Mode?**

Answer: The performance in Cluster-Mode was comparable to the performance in 7-Mode after installing and configuring Data ONTAP 8.1 in Cluster-Mode in the previously used FAS3270A supporting the 4-node RAC implementation in 7-Mode.

**What is the performance impact of migrating a 4-node RAC implementation into a virtualized environment that uses four ESX hosts running in 7-Mode?**

Answer: When running the 4-node RAC implementation in a virtualized environment on Data ONTAP 8.1 in 7-Mode there was a 10% delta in performance in terms of OETs processed.

**What is the performance impact of migrating a 4-node RAC implementation into a virtualized environment that uses four ESX hosts running in Cluster-Mode?**

Answer: When running the 4-node RAC implementation in a virtualized environment on Data ONTAP 8.1 in Cluster-Mode there was a 16% delta in performance in terms of OETs processed.

**What is the performance impact of migrating a 4-node RAC implementation from a virtualized environment that uses four ESX hosts to a virtualized environment that uses two ESX hosts running in 7-Mode?**

Answer: Deploying the 4-node RAC implementation from a virtualized environment with four physical servers into a virtualized environment with two physical servers reduced the ESX host CPU and memory resources available to the RAC nodes and resulted in a 27% delta in performance in terms of OETs processed. However, although the performance delta was 27%, the number of servers required to support the RAC implementation was reduced by 50%.

**What is the performance impact of migrating a 4-node RAC implementation from a virtualized environment that uses four ESX hosts to a virtualized environment that uses two ESX hosts running in Cluster-Mode?**

Answer: Deploying the 4-node RAC implementation from a virtualized environment with four physical servers into a virtualized environment with two physical servers reduced the ESX host CPU and memory resources available to the RAC nodes and resulted in a 25% delta in performance in terms of OETs processed. However, although the performance delta was 25%, the number of servers required to support the RAC implementation was reduced by 50%.

**What is the performance impact of migrating a 4-node RAC implementation from a virtualized environment that uses four ESX hosts running in 7-Mode to Cluster-Mode?**

Answer: When running the previously used 4-node RAC implementation in a virtualized environment using four physical servers and FAS3270A running Data ONTAP 8.1 in Cluster-Mode, there was an 8% delta in performance in terms of OETs processed.

## **What is the performance impact of migrating a 4-node RAC implementation from a virtualized environment that uses two ESX hosts running in 7-Mode to Cluster-Mode?**

Answer: When running the previously used 4-node RAC implementation in a virtualized environment using two physical servers and FAS3270A running Data ONTAP 8.1 in Cluster-Mode, there was a 6% delta in performance in terms of OETs processed.

In summary, the test results show that Data ONTAP 8.1 is an excellent option for providing a high-performance storage environment to support Oracle databases, regardless of whether the implementation uses 7-Mode or Cluster-Mode, or whether the hosts are physical or virtual.

The remainder of this report presents the details of the test configurations and the full set of test results and analysis.

## **4 RESULTS AND ANALYSIS**

Before analyzing the test results, it's important to understand the testing methodology and the workload employed. We employed a consistent testing methodology for all test cases. This methodology used an OLTP workload to demonstrate the capabilities of the configurations using Data ONTAP 8.1 in 7-Mode and in Cluster-Mode.

### **4.1 OLTP WORKLOAD**

The database created for the OLTP workload uses a data model designed for OET processing. The OLTP database was approximately 2.5TB in size and contained approximately 8,000 warehouses. We used a workload that simulated 250 users and 8,000 active warehouses. The client processes for the OLTP application were executed on a separate application server (client-server mode).

To determine the workload, we ran a series of tests using different numbers of users and, through an iterative process, increased the load generated against the database until either the database server CPU resources were exhausted or we observed that the read latencies reported by the database exceeded 10ms. Either of these events indicates an upper limit on acceptable performance. As previously mentioned, we settled on a load consisting of 250 users accessing the database. We held this load constant for all tests in order to have an apples-to-apples comparison of application throughput. These tests did not have any bottlenecks on the NetApp storage platform.

A mix of different types of transactions was used during each OLTP test run. These transaction types included order entries, payments, order status, delivery, and stock level. The number of OETs completed per minute was the primary metric used to measure application throughput.

The I/O mix for the OLTP workload was approximately 60% reads and 40% writes.

### **4.2 DETAILED TEST RESULTS**

Figure 1 shows the total number of OETs and the average db file sequential read wait time as reported by the Oracle database for each test. The db file sequential read wait time reflects the I/O latency seen at the database layer and indicates the amount of time, in milliseconds, that the database took to read a single database block from storage.

The results of these tests show that the performance of the 4-node Oracle RAC environment with Data ONTAP 8.1 operating in Cluster-Mode was comparable to its performance in 7-Mode. Such results indicate that Data ONTAP 8.1 Cluster-Mode is an excellent storage platform to support this Oracle environment. The tests further show that the performance in the 4-node RAC environment dropped approximately 10% and 16% after the RAC nodes were moved into a virtualized environment supported by Data ONTAP 8.1 operating in 7-Mode and in Cluster-Mode, respectively.

In general, database latencies observed with the test cases using Data ONTAP 8.1 in Cluster-Mode were slightly higher compared to the same tests using Data ONTAP 8.1 in 7-Mode.

Figure 1) OETs and db file sequential read latencies.

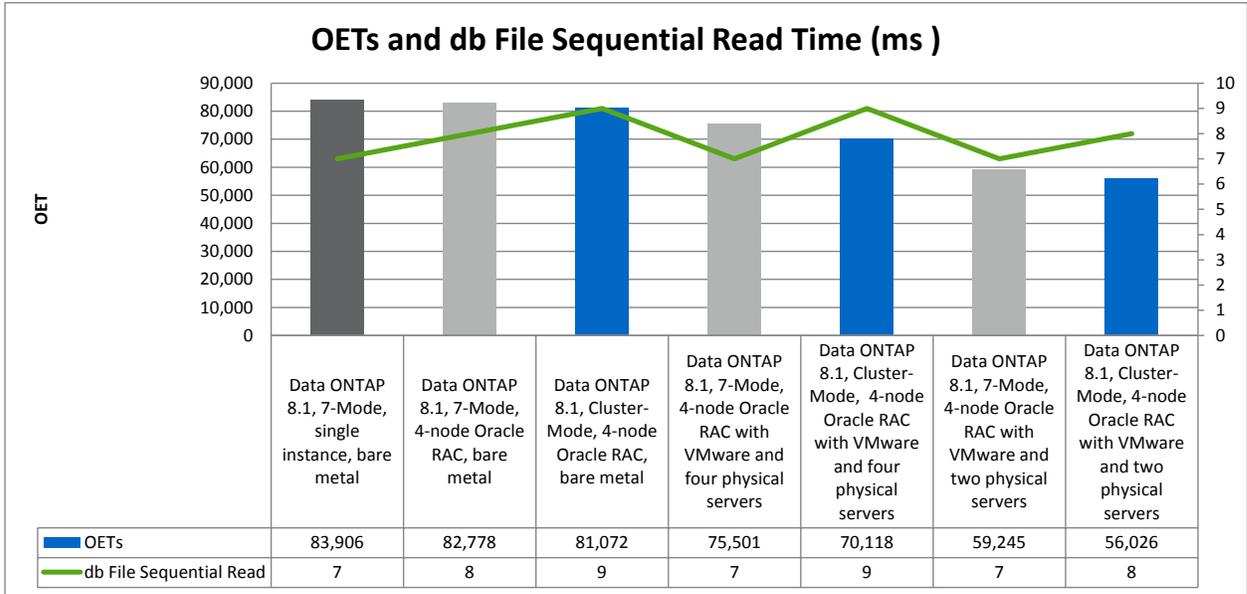
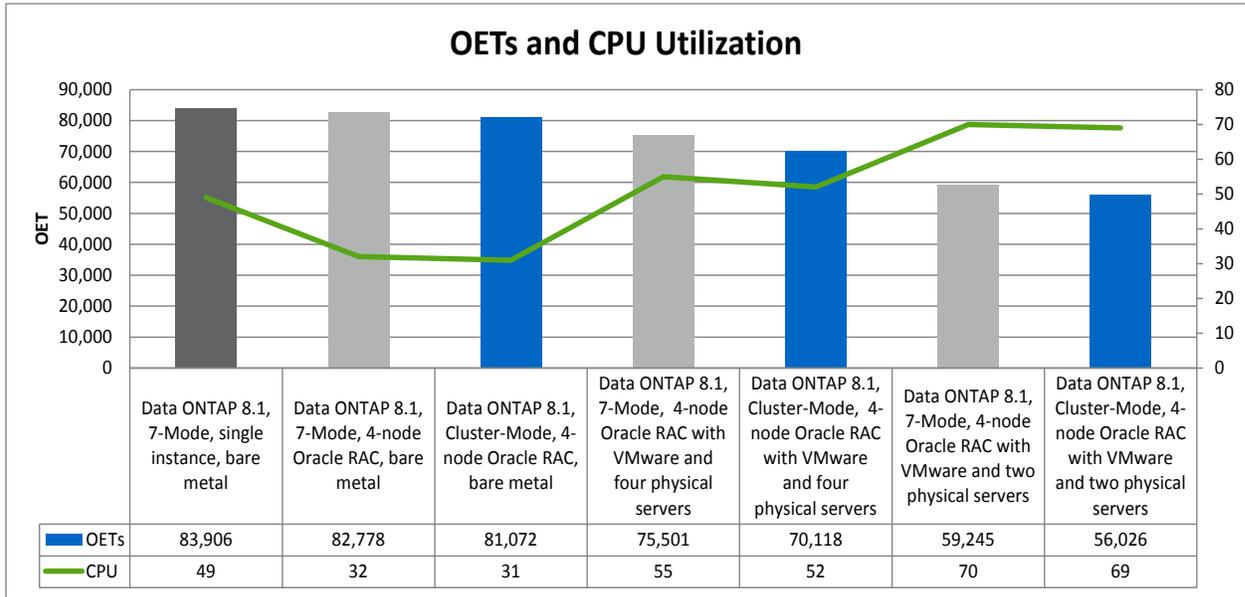


Figure 2 shows the typical amount of host CPU used to generate the corresponding OETs per minute for each protocol. We observed the following results:

- For the bare metal RAC configurations, the levels of RAC node CPU utilization were similar for the 7-Mode and the Cluster-Mode results. This indicates that running the RAC configuration in Cluster-Mode does not introduce any additional CPU utilization in the RAC node.
- For the virtualized RAC configurations using four physical nodes, the RAC node CPU utilization was approximately 70% higher compared to the bare metal test case using four physical servers.
- For the virtualized RAC configurations using two physical nodes, the RAC node CPU utilization was approximately 27% higher compared to the test case where the four RAC nodes were virtualized using four physical servers. This is expected because only two physical servers were hosting the four RAC nodes (that is, two VMs per physical server).

In all test cases, we found that running the four RAC nodes in virtualized environments resulted in higher levels of RAC node CPU utilization and slightly lower performance as compared to running the four RAC nodes in a bare metal environment.

Figure 2) OETs and CPU utilization by test configuration.



## 5 CONFIGURATION DETAILS

### 5.1 GENERIC CONFIGURATION

Wherever possible, we used similar configurations for the test environments, including:

- Identical hardware
- Similar storage layouts
- Similar Oracle initialization parameters
- Similar Linux OS settings

The remainder of section 5 explains the test configurations in detail. Unless otherwise specified, the configurations followed the best practices put forth by NetApp, Oracle, and VMware.

### 5.2 ORACLE SINGLE-INSTANCE CONFIGURATION USING DATA ONTAP OPERATING IN 7-MODE

The tests for this configuration used standard 10GbE for DNFS with a QLogic 8152 10GbE CNA in the server connected to NetApp 10GbE UTAs. The adapters were installed in the FAS3270A storage controllers and were connected through a Cisco Nexus<sup>®</sup> 5010 switch. The tests using the Oracle single-instance configuration used only Data ONTAP 8.1 configured in 7-Mode.

#### NETWORK CONFIGURATION

Figure 3 shows the network configuration for the tests using 10GbE. Table 1 describes the storage network hardware depicted in the diagram.

Figure 3) DNFS network using a 10GbE NetApp UTA.

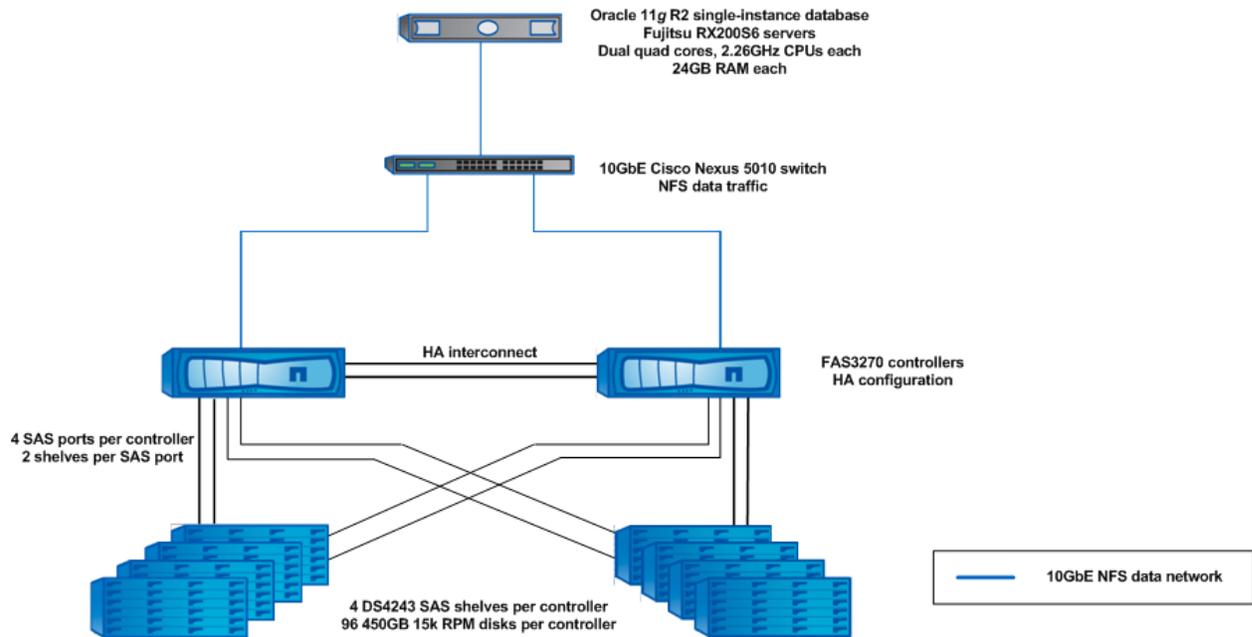


Table 1) Hardware for the DNFS network using a 10GbE NetApp UTA.

Hardware	Description
Database server	Fujitsu Primergy RX200 S5 dual quad core CPU @ 2.26GHz, 24GB RAM, RHEL 5 U4, and 1 single-port QLogic 8152 10GbE CNA
Switch	Cisco Nexus 5010 10GbE switch configured into a single private VLAN
NetApp controllers	FAS3270 controllers, active-active configuration using multipath high availability (MPHA) and 1 NetApp 10GbE UTA per storage controller
Storage shelves	4 DS4243 SAS shelves per controller, 450GB disks @15k RPM

## STORAGE NETWORK CONFIGURATION

We used jumbo frames in this network configuration. During these tests, a maximum transmission unit (MTU) size of 9,000 was set for all storage interfaces on the host, for all interfaces on the NetApp controllers, and for the ports involved on the switch. We used a single private subnet to segment Ethernet traffic between the host and the storage controllers.

## NFS MOUNTS AND DNFS CONFIGURATION

With Oracle DNFS, the NFS mount points and network paths are specified in a new configuration file, `oranfstab`. However, the NFS mounts must still be specified in the `/etc/fstab` file because Oracle cross-checks the entries in this file with the `oranfstab` file. If any NFS mounts do not match between `/etc/fstab` and `oranfstab`, then DNFS does not use those NFS mount points. We maintained the set of NFS mounts from the `/etc/fstab` file on the Linux host.

The following excerpt is from the `/etc/fstab` file on the Linux host.

```

192.168.1.10:/vol/nfsdata1 /mnt/nfsdata1 nfs
rw,nointr,hard,tcp,bg,rsize=65536,wsiz=65536,vers=3,timeo=0 0 0

192.168.1.10:/vol/nfslog1 /mnt/nfslog1 nfs
rw,nointr,hard,tcp,bg,rsize=65536,wsiz=65536,vers=3,timeo=0 0 0

192.168.1.10:/vol/nfsfra1 /mnt/nfsfra1 nfs
rw,nointr,hard,tcp,bg,rsize=65536,wsiz=65536,vers=3,timeo=0 0 0

192.168.1.11:/vol/nfsdata_f2_1 /mnt/nfsdata_f2_1 nfs
rw,nointr,hard,tcp,bg,rsize=65536,wsiz=65536,vers=3,timeo=0 0 0

192.168.1.11:/vol/nfslog_f2_1 /mnt/nfslog_f2_1 nfs
rw,nointr,hard,tcp,bg,rsize=65536,wsiz=65536,vers=3,timeo=0 0 0

192.168.1.11:/vol/nfsfra_f2_1 /mnt/nfsfra_f2_1 nfs
rw,nointr,hard,tcp,bg,rsize=65536,wsiz=65536,vers=3,timeo=0 0 0

```

The mount options that we used came from the NetApp Knowledge Base (KB) article [kb3010189: What are the mount options for databases on NetApp NFS?](#)

We made one change to the mount options: The [KB article kb3010189](#) uses an `rsiz` and a `wsiz` of 32,768 bytes. However, we decided to use an `rsiz` and a `wsiz` of 65,536 bytes. Using these settings enabled us to take advantage of the larger TCP window sizes, which are now available on the storage controllers.

The following excerpt is from the `oranfstab` file that was used during the DNFS tests.

```

server: filer1
path: 192.168.1.10
export: /vol/nfsdata1 mount: /mnt/nfsdata1
export: /vol/nfslog1 mount: /mnt/nfslog1

server: filer2
path: 192.168.1.11
export: /vol/nfsdata_f2_1 mount: /mnt/nfsdata_f2_1
export: /vol/nfsfra_f2_1 mount: /mnt/nfsfra_f2_1

```

For more information about DNS configuration, refer to the [Oracle Database Installation Guide](#) for Oracle 11g R2.

## DATA ONTAP TUNING OPTIONS

The following tuning option increases the size of the TCP receive window for NFS. The TCP receive window on the client should be adjusted to match the size of the NetApp controller because TCP reduces the size of the larger TCP receive window to match the size of the smaller TCP window.

```
options nfs.tcp.recvwindowsize 262144
```

## STORAGE CONFIGURATION

We used Data ONTAP 8.1 operating in 7-Mode for the storage controllers and for all 7-Mode tests. For details on the storage layout, see Figure 8 in section 7.3, “Storage Layouts for All Testing Configurations.”

### 5.3 ORACLE RAC CONFIGURATION USING DATA ONTAP 8.1 OPERATING IN 7-MODE

The tests for this configuration used standard 10GbE for DNFS with a QLogic 8152 10GbE CNA in the Oracle RAC servers connected to NetApp 10GbE UTAs. The adapters were installed in the FAS3270A storage controllers and were connected through a Cisco Nexus 5020 switch.

For the tests using Data ONTAP 8.1 operating in 7-Mode, we used the following configurations:

- DNFS using a 4-node Oracle RAC implementation configured on four physical servers accessing the NetApp FAS3270A storage controllers.
- DNFS using a 4-node Oracle RAC implementation configured in a vSphere 5.0 environment using four physical servers. In this case, each physical server contained a single VM in which the Oracle RAC software was installed.
- DNFS using a 4-node Oracle RAC implementation configured in a vSphere 5.0 environment using two physical servers. In this case, each physical server contained two VMs in which the Oracle RAC software was installed.

For the tests that used vSphere 5.0 to virtualize the Oracle RAC environment, we configured a separate server running VMware vCenter™ 5.0 to manage the environment. We installed Red Hat Enterprise Linux 5.6 on each of the VMs created to host the four RAC nodes. Each of the VMs contained a virtual disk for the OS partition that was mapped through the ESX 5.0 NFS stack to the FAS3270A storage controllers. To access the Oracle database and related files using DNFS, we configured each of the VMs to directly mount the volumes on the FAS3270A storage controllers by using the NFS client in the guest OS to effectively bypass the NFS services of the ESX servers that hosted the RAC nodes.

## NETWORK CONFIGURATIONS

Figure 4 and Figure 5 show the network configurations used for the tests involving a 4-node RAC configuration deployed on two and on four physical servers in a Data ONTAP 8.1 7-Mode environment. Table 2 describes the storage network hardware depicted in the diagrams. The only difference in the configurations is the number of physical server systems on which the four RAC nodes were deployed.

Figure 4) Oracle RAC configuration using Data ONTAP 8.1 operating in 7-Mode with four physical servers.

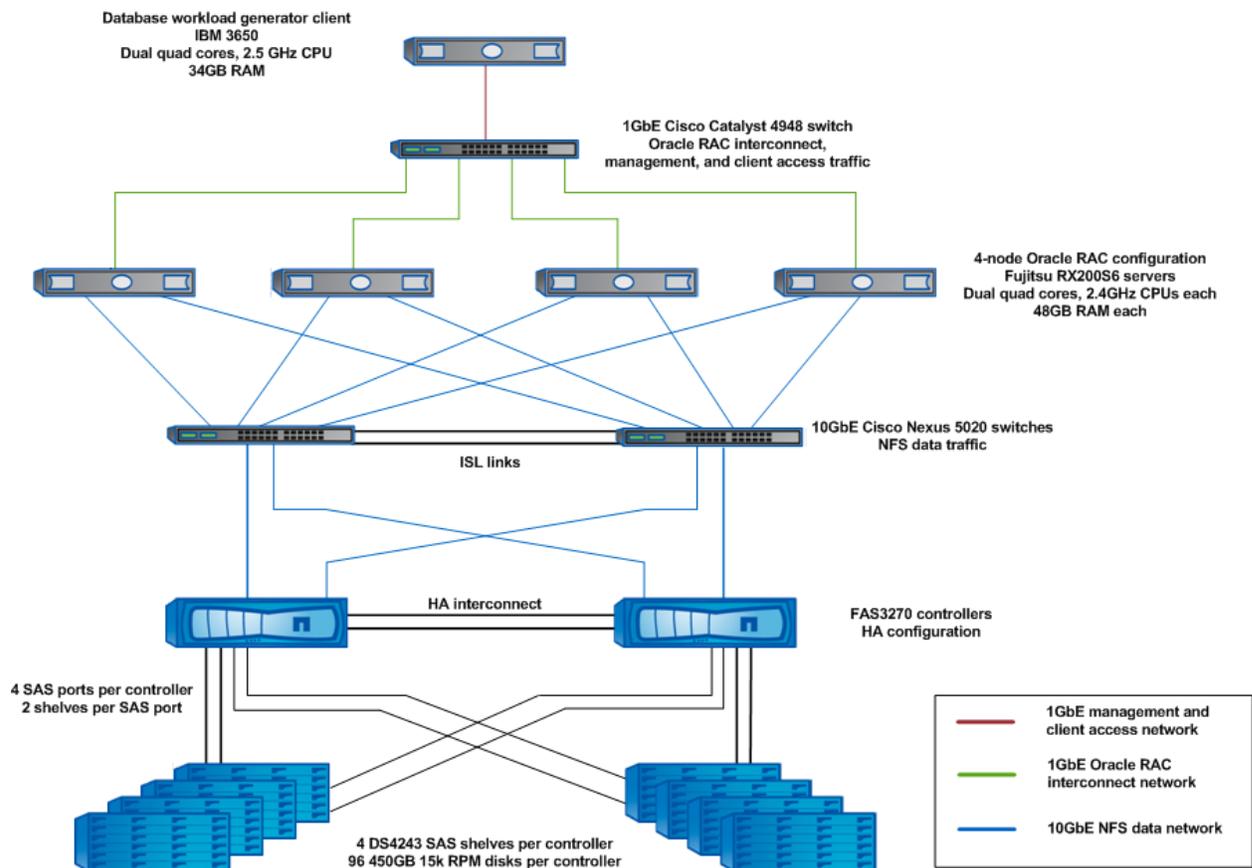


Figure 5) Oracle RAC configuration using Data ONTAP 8.1 operating in 7-Mode with two physical servers.

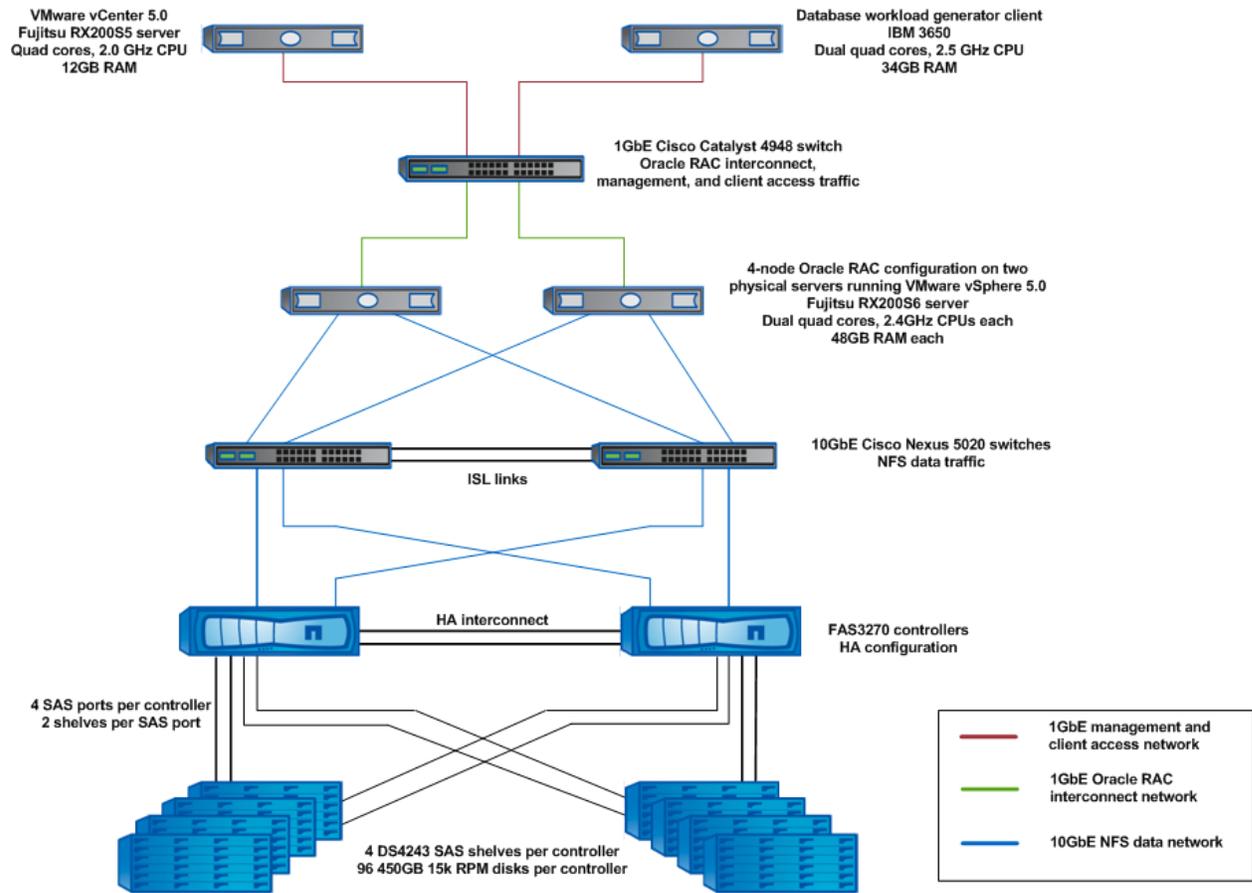


Table 2) Hardware for the Oracle RAC configuration using Data ONTAP 8.1 operating in 7-Mode with either two or four physical servers using DNFS over 10GbE.

Hardware	Description
Database server	2 or 4 Fujitsu Primergy RX200 S6 dual quad core CPU @ 2.4GHz, 48GB RAM, RHEL 5 U6, and 1 dual-port 10 GbE Intel® 82599EB NIC per server
Switch	Cisco Nexus 5020 10GbE switch configured into a single private VLAN
NetApp controllers	FAS3270 controllers, active-active configuration using multipath high availability (MPHA) and 1 NetApp 10GbE UTA per storage controller
Storage shelves	4 DS4243 SAS shelves per controller, 450GB disks @15k RPM

## STORAGE NETWORK CONFIGURATION

We used jumbo frames in this network configuration. During these tests, an MTU size of 9,000 was set for all storage interfaces on the host, for all interfaces on the NetApp controllers, and for the ports involved on the switch. We used a single private subnet to segment Ethernet traffic between the host and the storage controllers.

## NFS MOUNTS AND DNFS CONFIGURATION

With Oracle DNFS, the NFS mount points and network paths are specified in a new configuration file, `oranfstab`. However, the NFS mounts must still be specified in the `/etc/fstab` file because Oracle cross-checks the entries in this file with the `oranfstab` file. If any NFS mounts do not match between `/etc/fstab` and `oranfstab`, then DNFS does not use those NFS mount points.

When conducting tests using Oracle RAC and Data ONTAP 8.1 operating in 7-Mode, we created a set of NFS mount points on the RAC database nodes that allowed the RAC nodes to access the database files uniformly across the FAS3270A storage controllers. The following mount points were defined in the `/etc/fstab` file on the Linux hosts supporting RAC nodes 1 through 4.

### RAC Node 1

```
stl-3270-1-e1a:/vol/orabin1 /u01
rw,bg,hard,rsi=32768,ws=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-1-e1a:/vol/oraocr1 /u02
rw,bg,hard,rsi=32768,ws=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-2-e1a:/vol/oraocr2 /u03
rw,bg,hard,rsi=32768,ws=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-1-e1a:/vol/oraocr3 /u04
rw,bg,hard,rsi=32768,ws=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-2-e1a:/vol/oravote1 /u05
rw,bg,hard,rsi=32768,ws=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-1-e1a:/vol/oravote2 /u06
rw,bg,hard,rsi=32768,ws=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-2-e1a:/vol/oravote3 /u07
rw,bg,hard,rsi=32768,ws=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-1-e1a:/vol/oradata1 /u08
rw,bg,hard,rsi=32768,ws=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-2-e1a:/vol/oradata2 /u09
rw,bg,hard,rsi=32768,ws=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-1-e1a:/vol/oralog1 /u10
rw,bg,hard,rsi=32768,ws=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-2-e1a:/vol/oralog2 /u11
rw,bg,hard,rsi=32768,ws=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp
```

### RAC Node 2

```
stl-3270-2-e1b:/vol/orabin2 /u01
rw,bg,hard,rsi=32768,ws=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-1-e1b:/vol/oraocr1 /u02
rw,bg,hard,rsi=32768,ws=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-2-e1b:/vol/oraocr2 /u03
rw,bg,hard,rsi=32768,ws=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-1-e1b:/vol/oraocr3 /u04
rw,bg,hard,rsi=32768,ws=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-2-e1b:/vol/oravote1 /u05
rw,bg,hard,rsi=32768,ws=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp
```

```

stl-3270-1-elb:/vol/oravote2 /u06
rw,bg,hard,rsize=32768,wsiz=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-2-elb:/vol/oravote3 /u07
rw,bg,hard,rsize=32768,wsiz=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-1-elb:/vol/oradata1 /u08
rw,bg,hard,rsize=32768,wsiz=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-2-elb:/vol/oradata2 /u09
rw,bg,hard,rsize=32768,wsiz=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-1-elb:/vol/oralog1 /u10
rw,bg,hard,rsize=32768,wsiz=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-2-elb:/vol/oralog2 /u11 nfs
rw,bg,hard,rsize=32768,wsiz=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

```

### RAC Node 3

```

stl-3270-1-ela:/vol/orabin3 /u01
rw,bg,hard,rsize=32768,wsiz=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-1-ela:/vol/oraocr1 /u02
rw,bg,hard,rsize=32768,wsiz=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-2-ela:/vol/oraocr2 /u03
rw,bg,hard,rsize=32768,wsiz=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-1-ela:/vol/oraocr3 /u04
rw,bg,hard,rsize=32768,wsiz=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-2-ela:/vol/oravote1 /u05
rw,bg,hard,rsize=32768,wsiz=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-1-ela:/vol/oravote2 /u06
rw,bg,hard,rsize=32768,wsiz=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-2-ela:/vol/oravote3 /u07
rw,bg,hard,rsize=32768,wsiz=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-1-ela:/vol/oradata1 /u08
rw,bg,hard,rsize=32768,wsiz=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-2-ela:/vol/oradata2 /u09
rw,bg,hard,rsize=32768,wsiz=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-1-ela:/vol/oralog1 /u10
rw,bg,hard,rsize=32768,wsiz=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-2-ela:/vol/oralog2 /u11
rw,bg,hard,rsize=32768,wsiz=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

```

### RAC Node 4

```

stl-3270-2-elb:/vol/orabin4 /u01
rw,bg,hard,rsize=32768,wsiz=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-1-elb:/vol/oraocr1 /u02
rw,bg,hard,rsize=32768,wsiz=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

```

```

stl-3270-2-elb:/vol/oraocr2 /u03
rw,bg,hard,rsize=32768,wsiz=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-1-elb:/vol/oraocr3 /u04
rw,bg,hard,rsize=32768,wsiz=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-2-elb:/vol/oravote1 /u05
rw,bg,hard,rsize=32768,wsiz=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-1-elb:/vol/oravote2 /u06
rw,bg,hard,rsize=32768,wsiz=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-2-elb:/vol/oravote3 /u07
rw,bg,hard,rsize=32768,wsiz=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-1-elb:/vol/oradata1 /u08
rw,bg,hard,rsize=32768,wsiz=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-2-elb:/vol/oradata2 /u09
rw,bg,hard,rsize=32768,wsiz=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-1-elb:/vol/oralog1 /u10
rw,bg,hard,rsize=32768,wsiz=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

stl-3270-2-elb:/vol/oralog2 /u11 nfs
rw,bg,hard,rsize=32768,wsiz=32768,vers=3,actimeo=0,nointr,suid,timeo=600,tcp

```

The following excerpts are from the `oranfstab` file that was used during the RAC performance testing to define the mount points for use by the Oracle DNFS client. The `oranfstab` file excerpts that follow were used by RAC nodes 1 and 3.

```

server: stl-3270-1
path: 192.168.20.221
export: /vol/oradata1 mount: /u08
export: /vol/oralog1 mount: /u10

server: stl-3270-2
path: 192.168.30.222
export: /vol/oradata2 mount: /u09
export: /vol/oralog2 mount: /u11

```

The `oranfstab` file excerpts that follow were used by RAC nodes 2 and 4.

```

server: stl-3270-1
path: 192.168.30.231
export: /vol/oradata1 mount: /u08
export: /vol/oralog1 mount: /u10

server: stl-3270-2
path: 192.168.20.232
export: /vol/oradata2 mount: /u09
export: /vol/oralog2 mount: /u11

```

For more information about DNS configuration, refer to the [Oracle Database Installation Guide](#) for Oracle 11g R2.

## DATA ONTAP TUNING OPTIONS

The following tuning option increases the size of the TCP receive window for NFS. The TCP receive window on the client should be adjusted to match the size of the NetApp controller because TCP reduces the size of the larger TCP receive window to match the size of the smaller TCP window.

```
options nfs.tcp.recvwindowsize 262144
```

## STORAGE CONFIGURATION

We used Data ONTAP 8.1 operating in 7-Mode for the storage controllers and for all tests. For details on the storage layout for DNFS testing, see Figure 9 in section 7.3, “Storage Layouts for All Testing Configurations.”

### 5.4 ORACLE RAC CONFIGURATION USING DATA ONTAP 8.1 OPERATING IN CLUSTER-MODE

The tests for this configuration used standard 10GbE for DNFS with an Intel dual-port 10 GbE NIC in the server connected to NetApp 10GbE UTAs. The adapters were installed in the FAS3270A storage controllers and were connected through a Cisco Nexus 5020 switch.

For the tests using Data ONTAP 8.1 operating in Cluster-Mode, we used the following configurations:

- DNFS using a 4-node Oracle RAC implementation configured on four physical servers accessing the NetApp FAS3270A storage controllers.
- DNFS using a 4-node Oracle RAC implementation configured in a vSphere 5.0 environment using four physical servers. In this case, each physical server contained a single VM in which the Oracle RAC software was installed.
- DNFS using a 4-node Oracle RAC implementation configured in a vSphere 5.0 environment using two physical servers. In this case, each physical server contained two VMs in which the Oracle RAC software was installed.

For the tests using vSphere 5.0 to virtualize the Oracle RAC environment, we configured a separate server running VMware vCenter 5.0 to manage the environment. We installed Red Hat Enterprise Linux 5.6 on each of the VMs created to host the four RAC nodes. Each of the VMs contained a virtual disk for the OS partition that was mapped through the ESX 5.0 NFS stack to the FAS3270A storage controllers. To access the Oracle database and related files by using DNFS, we configured each of the VMs to directly mount the volumes on the FAS3270A by using the NFS client in the guest OS to effectively bypass the NFS services of the ESX servers that hosted the RAC nodes.

## NETWORK CONFIGURATIONS

Figure 6 and Figure 7 show the network configurations used for the tests involving a 4-node RAC configuration deployed on two and on four physical servers in a Data ONTAP 8.1 Cluster-Mode environment. Table 3 describes the storage network hardware depicted in the diagrams. The only difference in the configurations is the number of physical server systems on which the four RAC nodes were deployed.

Figure 6) Oracle RAC configuration using Data ONTAP 8.1 operating in Cluster-Mode with four physical servers.

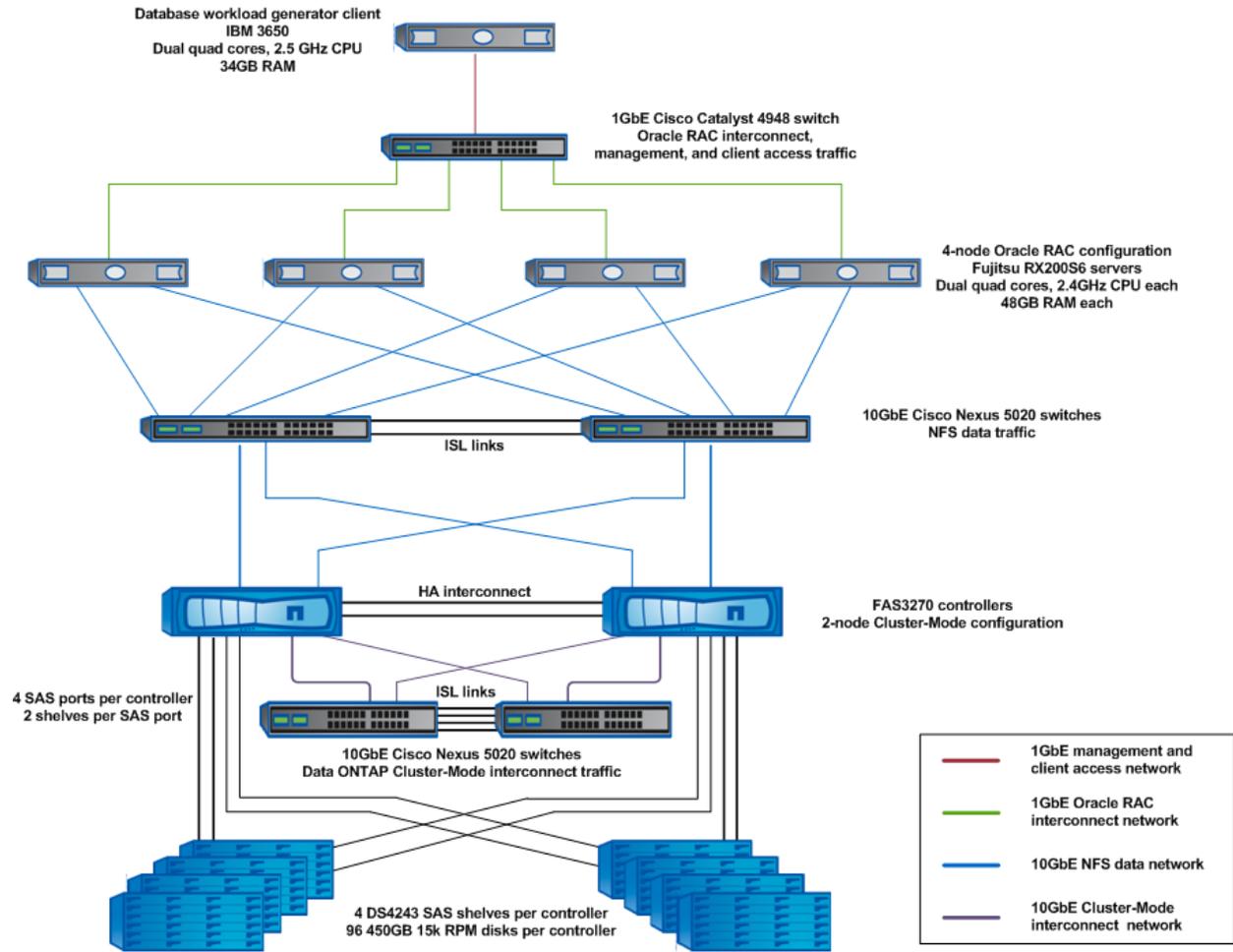


Figure 7) Oracle RAC configuration using Data ONTAP 8.1 operating in Cluster-Mode with two physical servers.

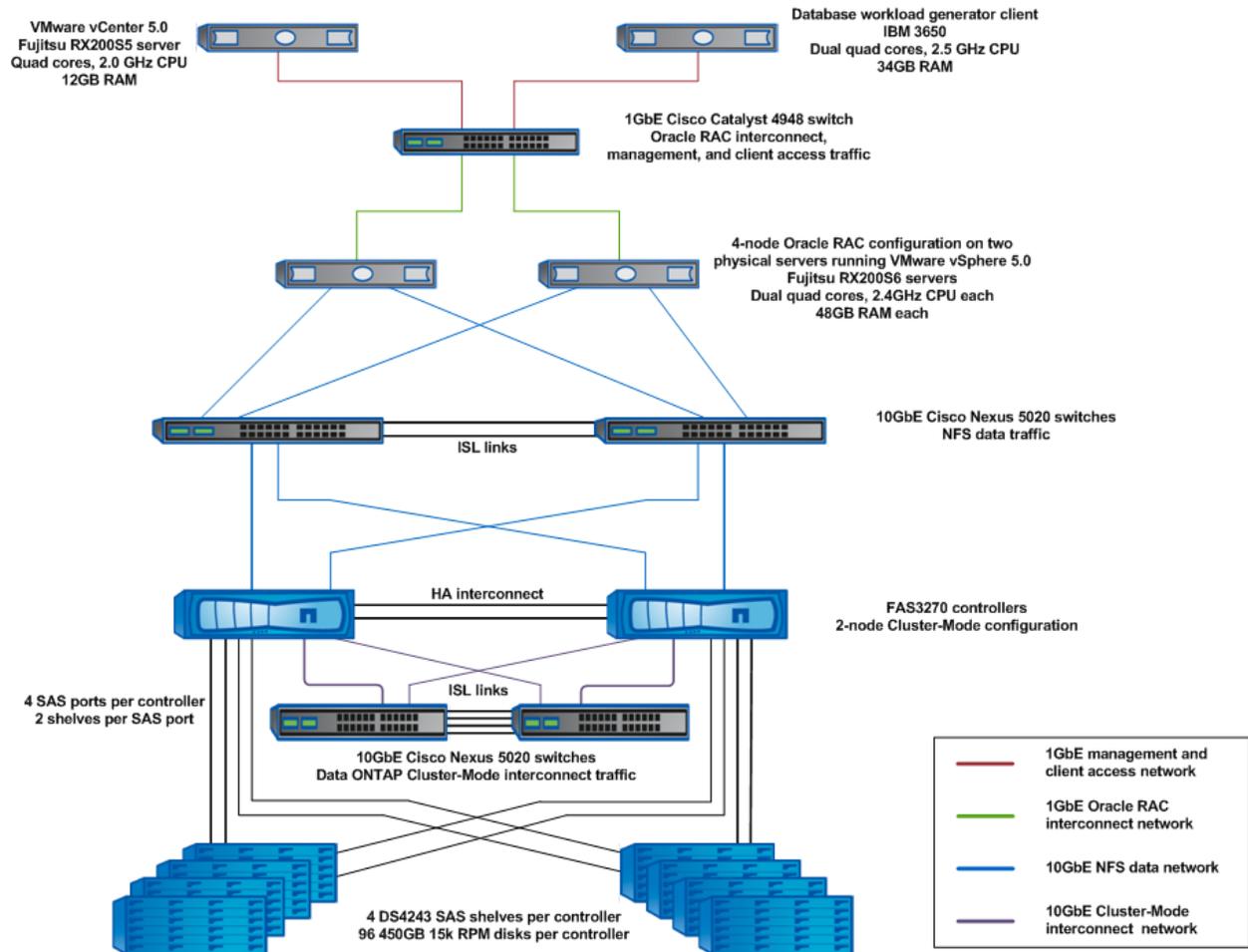


Table 3) Hardware for the Oracle RAC configuration using Data ONTAP 8.1 operating in Cluster-Mode with either two or four physical servers using DNFS over 10GbE.

Hardware	Description
Database server	2 or 4 Fujitsu Primergy RX200 S6 dual quad core CPU @ 2.4GHz, 48GB RAM, RHEL 5 U6, and 1 dual-port 10GbE Intel 82599EB NIC per server
Switch infrastructure for data traffic	2 Cisco Nexus 5020 10GbE switches
Switch infrastructure for ONTAP cluster traffic	2 Cisco Nexus 5020 10GbE switches
NetApp controllers	FAS3270 controllers, Cluster-Mode configuration using multipath high availability (MPHA) and 1 NetApp 10GbE UTA per storage controller
Storage shelves	4 DS4243 SAS shelves per controller, 450GB disks @15k RPM

## STORAGE NETWORK CONFIGURATION

We used jumbo frames in this network configuration. During these tests, an MTU size of 9,000 was set for all storage interfaces on the host, for all interfaces on the NetApp controllers, and for the ports involved on the switch. We used a single private subnet to segment Ethernet traffic between the host and the storage controllers.

## NFS MOUNTS AND DNFS CONFIGURATION

With Oracle DNFS, the NFS mount points and network paths are specified in a new configuration file, `oranfstab`. However, the NFS mounts must still be specified in the `/etc/fstab` file because Oracle cross-checks the entries in this file with the `oranfstab` file. If any NFS mounts do not match between `/etc/fstab` and `oranfstab`, then DNFS does not use those NFS mount points.

When conducting tests using Oracle RAC and Data ONTAP 8.1 operating in Cluster-Mode, we created a set of NFS mount points on the RAC database nodes that allowed the RAC nodes to access the database files uniformly across the FAS3270A storage controllers.

In Data ONTAP 8.1 operating in Cluster-Mode, the NFS mount points on the RAC nodes are specified using a combination of the logical network interface and Cluster-Mode junction path that provides access to the different Oracle configuration and database files.

The following mount points were defined in the `/etc/fstab` file on the Linux hosts supporting RAC nodes 1 through 4.

### RAC Node 1

```
vsln1data1:/orabin1 on /u01 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.20.221)
vsln1data1:/oraocr1 on /u02 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.20.221)
vsln2data1:/oraocr2 on /u03 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.30.222)
vsln1data1:/oraocr3 on /u04 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.20.221)
vsln2data1:/oravotel on /u05 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.30.222)
vsln1data1:/oravote2 on /u06 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.20.221)
vsln2data1:/oravote3 on /u07 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.30.222)
vsln1data1:/oradata1 on /u08 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.20.221)
vsln2data1:/oradata2 on /u09 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.30.222)
vsln1data1:/oralog1 on /u10 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.20.221)
vsln2data1:/oralog2 on /u11 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.30.222)
```

### RAC Node 2

```
vsln2data2:/orabin2 on /u01 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.20.232)
vsln1data2:/oraocr1 on /u02 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.30.231)
vsln2data2:/oraocr2 on /u03 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.20.232)
vsln1data2:/oraocr3 on /u04 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.30.231)
vsln2data2:/oravotel on /u05 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.20.232)
vsln1data2:/oravote2 on /u06 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.30.231)
vsln2data2:/oravote3 on /u07 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.20.232)
```

```

vslndata2:/oradata1 on /u08 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.30.231)
vslndata2:/oradata2 on /u09 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.20.232)
vslndata2:/oralog1 on /u10 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.30.231)
vslndata2:/oralog2 on /u11 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.20.232)

```

### RAC Node 3

```

vslndata1:/orabin3 on /u01 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.20.221)
vslndata1:/oraocr1 on /u02 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.20.221)
vslndata1:/oraocr2 on /u03 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.30.222)
vslndata1:/oraocr3 on /u04 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.20.221)
vslndata1:/oravotel on /u05 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.30.222)
vslndata1:/oravote2 on /u06 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.20.221)
vslndata1:/oravote3 on /u07 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.30.222)
vslndata1:/oradata1 on /u08 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.20.221)
vslndata1:/oradata2 on /u09 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.30.222)
vslndata1:/oralog1 on /u10 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.20.221)
vslndata1:/oralog2 on /u11 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.30.222)

```

### RAC Node 4

```

vslndata2:/orabin4 on /u01 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.20.232)
vslndata2:/oraocr1 on /u02 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.30.231)
vslndata2:/oraocr2 on /u03 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.20.232)
vslndata2:/oraocr3 on /u04 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.30.231)
vslndata2:/oravotel on /u05 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.20.232)
vslndata2:/oravote2 on /u06 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.30.231)
vslndata2:/oravote3 on /u07 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.20.232)
vslndata2:/oradata1 on /u08 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.30.231)
vslndata2:/oradata2 on /u09 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.20.232)
vslndata2:/oralog1 on /u10 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.30.231)
vslndata2:/oralog2 on /u11 type nfs
(rw,bg,hard,rsize=32768,wsiz=32768,nfsvers=3,actimeo=0,nointr,timeo=600,tcp,addr=192.168.20.232)

```

The following are excerpts from the `oranfstab` file that was used during the RAC performance testing to define the mount points for use by the Oracle DNFS client. The `oranfstab` file excerpts that follow were used by RAC nodes 1 and 3.

```

server: vslndata1
path: 192.168.20.221
export: /oradata1 mount: /u08
export: /oralog1 mount: /u10

```

```

server: vs1n2data1
path: 192.168.30.222
export: /oradata2 mount: /u09
export: /oralog2 mount: /u11

```

The oranfstab file excerpts that follow were used by RAC nodes 2 and 4

```

server: vs1n1data2
path: 192.168.30.231
export: /oradata1 mount: /u08
export: /oralog1 mount: /u10

server: vs1n2data2
path: 192.168.20.232
export: /oradata2 mount: /u09
export: /oralog2 mount: /u11

```

For more information about DNS configuration, refer to the [Oracle Database Installation Guide](#) for Oracle 11g R2.

## DATA ONTAP TUNING OPTIONS

For these tests, we used only default values for the Data ONTAP 8.1 Cluster-Mode NFS parameters that are related to transfer and window sizes.

## STORAGE CONFIGURATION

We used Data ONTAP 8.1 operating in Cluster-Mode for the storage controllers and for all tests. For details on the storage layout for DNFS testing, see Figure 10 in section 7.3, “Storage Layouts for All Testing Configurations.”

## 6 CONCLUSION

NetApp has a long history of providing high performance storage systems for Oracle database environments. With the advent of Data ONTAP Cluster-Mode, NetApp continues to develop leading edge storage systems that provide the ability to scale out both capacity and performance in support of NetApp customers’ current and future Oracle database environments.

The test results presented in this report show that both Data ONTAP 8.1 7-Mode and Cluster-Mode provide excellent performance in Oracle DNFS environments.

## 7 APPENDIXES

### 7.1 BEST PRACTICE SUMMARY FOR DATA ONTAP 8.1 OPERATING IN 7-MODE

This section summarizes the best practices listed in this report for each of the configurations that we tested. Table 4 contains the parameter name, value, and description, and the protocol for which it was used.

Table 4) Best practice summary for Data ONTAP operating in 7-Mode.

Parameter Name	Protocol	Value	Description
rsize	DNFS	65535	NFS mount option
wsizer	DNFS	65535	NFS mount option

Parameter Name	Protocol	Value	Description
nfs.tcp.recvwindowsize	DNFS	262144	Storage controller TCP receive window size
nfs.tcp.xfersize	DNFS	65536	TCP transfer window size

## 7.2 HARDWARE

Table 5) Database server hardware specifications for Oracle single-instance testing.

Component	Details
System type	Fujitsu Primergy RX200 S5
Operating system	RHEL 5 U4
Processor	2 quad core CPU @ 2.26GHz
Memory	24GB
Oracle database version	11.2.0.1.0
Network connectivity	1 dual-port QLogic 8152 10GbE CNA

Table 6) Database server hardware specifications for Oracle RAC testing.

Component	Details
System type	Fujitsu Primergy RX200 S6
Operating system	RHEL 5 U6
Processor	2 quad core CPU @ 2.4GHz
Memory	48GB
Oracle database version	11.2.0.1.0
Network connectivity	1 dual-port 10GbE Intel 82599EB NIC

Table 7) NetApp FAS3270 storage system specifications.

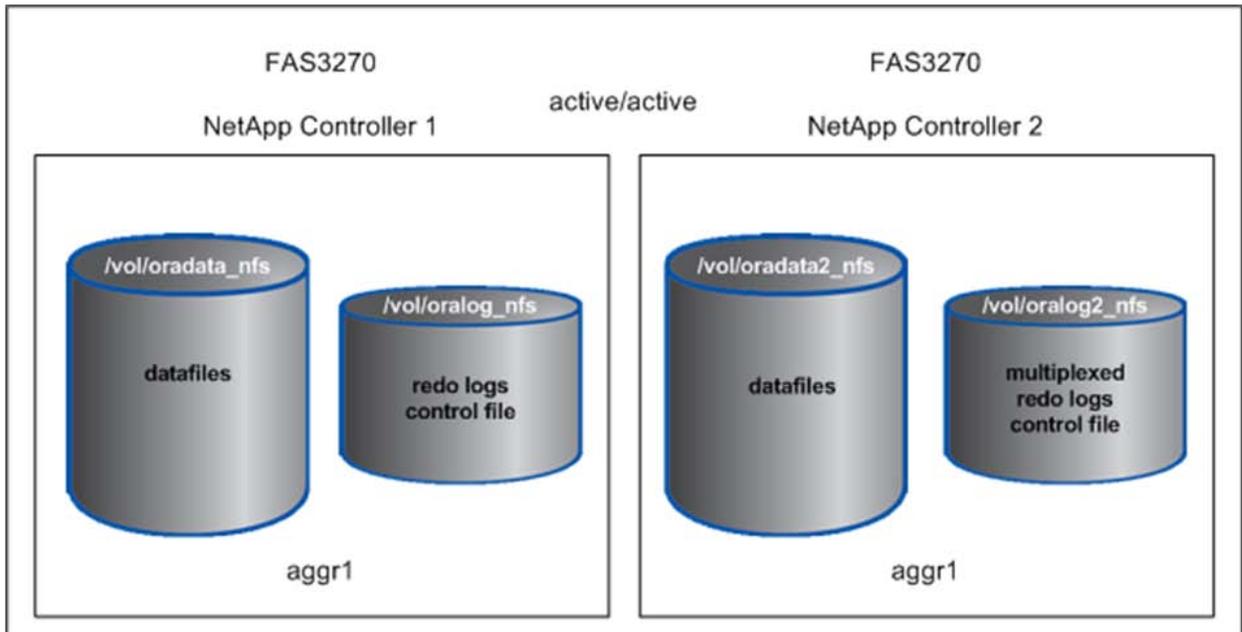
Component	Details (in each storage controller)
System type	NetApp FAS3270 HA pair
Operating system	Data ONTAP 8.1 7-Mode and Cluster-Mode
Processor	2 dual core Intel CPU @ 3GHz
Memory	16GB
NVRAM	4GB
Disks	4 DS4243 SAS shelves per controller, 450GB disks @15k RPM
Network devices	10GbE UTA

### 7.3 STORAGE LAYOUTS FOR ALL TESTING CONFIGURATIONS

#### DNFS LAYOUT FOR SINGLE-INSTANCE TESTING

Figure 8 shows the layout used for DNFS. The Oracle data files are distributed evenly across both NetApp storage controllers in the `oradata_nfs` and `oradata2_nfs` volumes. The Oracle online redo logs are also balanced across the two controllers. The primary redo log members are stored on NetApp controller 1, and all multiplexed redo log members are stored on NetApp controller 2.

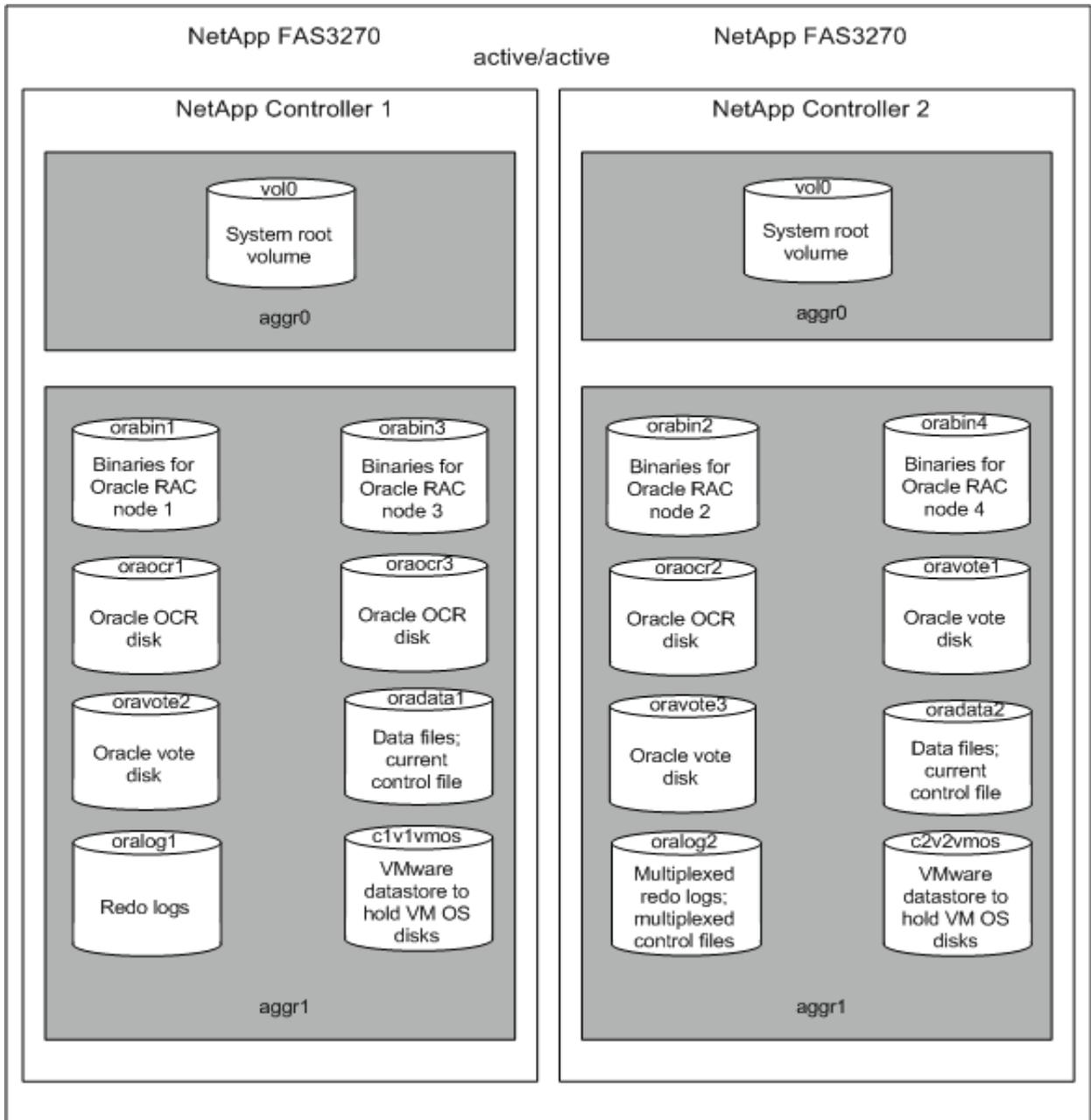
Figure 8) Volume layout for single-instance testing.



#### DNFS LAYOUT FOR 7-MODE TESTING USING ORACLE RAC

Figure 9 shows the layout used for DNFS. The Oracle data files are distributed evenly across both NetApp storage controllers in the `oradata1` and `oradata2` volumes. The Oracle online redo logs are also balanced across the two controllers. The primary redo log members are stored on NetApp controller 1, and all multiplexed redo log members are stored on NetApp controller 2.

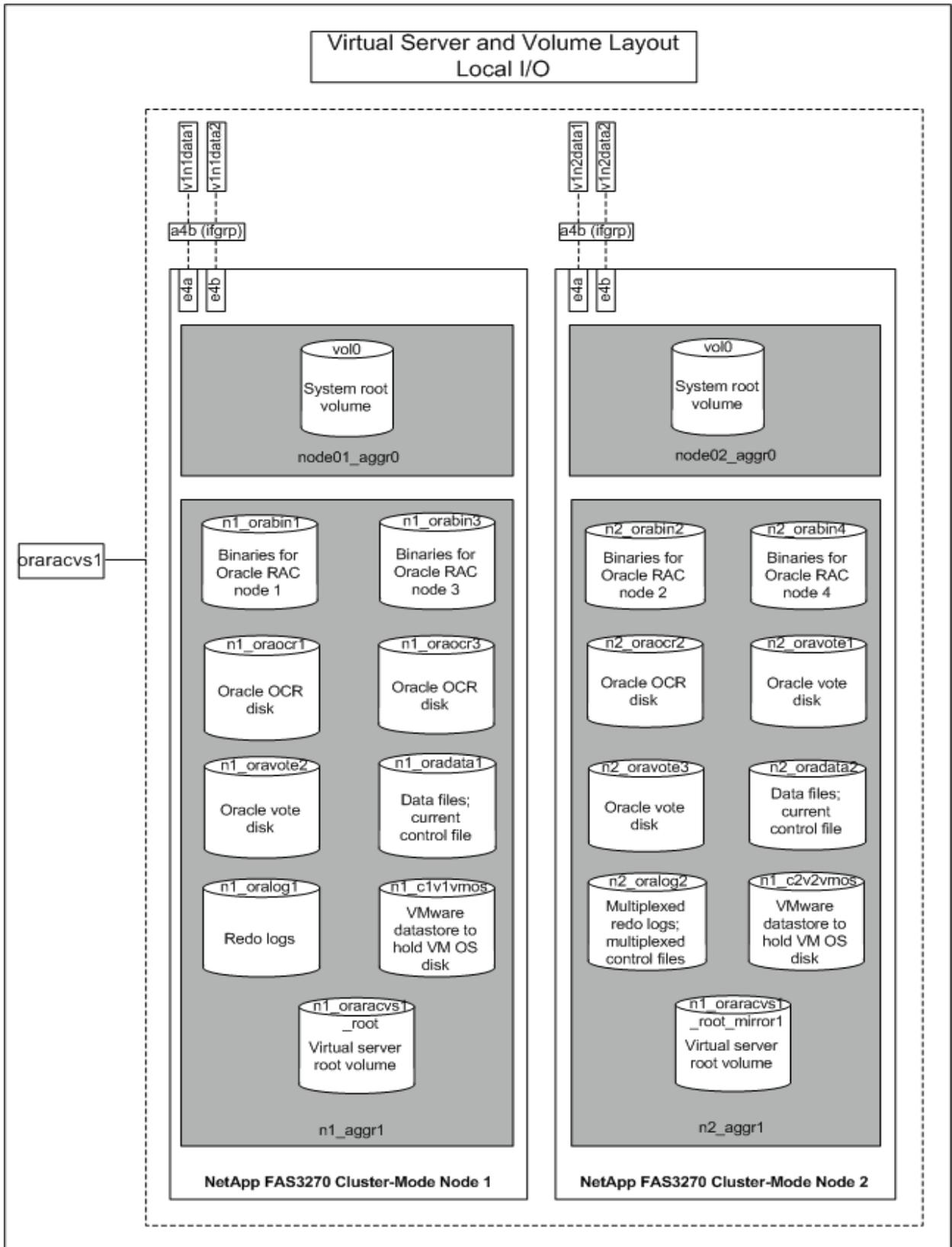
Figure 9) DNFS volume layout for Oracle RAC testing using Data ONTAP 8.1 operating in 7-Mode.



### DNFS LAYOUT FOR CLUSTER-MODE TESTING USING ORACLE RAC

Figure 10 shows the layout used for DNFS. The Oracle data files are distributed evenly across both NetApp storage controllers in the n1\_oradata1 and n2\_oradata2 NFS volumes. The Oracle online redo logs are also balanced across the two controllers. The primary redo log members are stored on NetApp controller 1, and all multiplexed redo log members are stored on NetApp controller 2.

Figure 10) DNFS volume layout for Oracle RAC testing using Data ONTAP 8.1 operating in Cluster-Mode.



## 7.4 ORACLE INITIALIZATION PARAMETERS FOR SINGLE-INSTANCE CONFIGURATION TESTING

Table 8) Oracle initialization parameters for single-instance database configuration testing.

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 to get faster execution
db_block_size	8192	Size of database block in bytes
db_cache_size	5301600256	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 asynch I/O for random access devices
dml_locks	500	One DML lock for each table modified in a transaction
filesystemio_options	setall	I/O operations on file system files
log_buffer	1703936	Redo circular buffer size
parallel_max_servers	100	Maximum parallel query servers per instance
plsql_optimize_level	2	PL/SQL optimize level
processes	350	User processes
recovery_parallelism	40	Number of server processes to use for parallel recovery
sessions	560	User and system sessions
shared_pool_size	402653184	Size in bytes of shared pool
statistics_level	Typical	Statistics level

Parameter Name	Value	Description
transactions	350	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_tablespace	undo_1	Use or switch undo tablespace

## 7.5 ORACLE INITIALIZATION PARAMETERS FOR RAC CONFIGURATION TESTING

Table 9) Oracle initialization parameters for RAC configuration testing.

Parameter Name	Value	Description
_allocate_creation_order	FALSE	During allocation, files should be examined in the order in which they were created
_in_memory_undo	FALSE	Make in-memory undo for top-level transactions
_undo_autotune	FALSE	Enable autotuning of undo_retention
cluster_database	TRUE	RAC database
compatible	11.2.0.0.0	Database is completely compatible with this software version
control_files	/u08/oradata1/control_001, /u09/oradata2/control_002	Control file location
db_16k_cache_size	5368709120	Size of cache for 16K buffers
db_2k_cache_size	268435456	Size of cache for 2K buffers
db_4k_cache_size	268435456	Size of cache for 4K buffers
db_block_size	8192	Size of database block in bytes
db_cache_size	536870920	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	500	Maximum allowable number of database files
db_name	tpcc	Database name specified in CREATE DATABASE
disk_asynch_io	TRUE	Use asynch I/O for random access devices

Parameter Name	Value	Description
dml_locks	500	One DML lock for each table modified in a transaction
filesystemio_options	setall	I/O operations on file system files
instance_number	1	DB instance number
log_buffer	16777216	Redo circular buffer size
open_cursors	1024	Maximum number of open cursors
parallel_max_servers	100	Maximum parallel query servers per instance
plsql_optimize_level	2	Optimization level that will be used to compile PL/SQL library units
processes	1000	User processes
recovery_parallelism	40	Number of server processes to use for parallel recovery
remote_listener	oracscan:1521	Name of remote listener
remote_login_passwordfile	EXCLUSIVE	The password file can be used by only one database
sessions	1536	User and system sessions
shared_pool_size	4294967296	Size in bytes of shared pool
spfile	/u08/oradata1/spfile.ora	Location of SPFILE
statistics_level	typical	Statistics level
thread	1	Database threads
undo_management	AUTO	If TRUE, instance runs in SMU mode; otherwise in RBU mode
undo_retention	10800	Undo retention in seconds
undo_tablespace	undo_1	Use or switch undo tablespace

## 7.6 LINUX KERNEL PARAMETERS FOR SINGLE-INSTANCE DATABASE CONFIGURATION TESTING

Table 10) Linux nondefault kernel parameters for single-instance database configuration testing.

Parameter Name	Value	Description
sunrpc.tcp_slot_table_entries	128	Maximum number of outstanding asynch I/O calls

Parameter Name	Value	Description
kernel.sem	250 32000 100 128	Semaphores
net.ipv4.ip_local_port_range	9000 65500	Local port range used by TCP and UDP
net.core.rmem_default	4194304	Default TCP receive window size (default buffer size)
net.core.rmem_max	4194304	Maximum TCP receive window size (maximum buffer size)
net.core.wmem_default	262144	Default TCP send window size (default buffer size)
net.core.wmem_max	1048576	Maximum TCP send window size (maximum buffer size)
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

## 7.7 LINUX KERNEL PARAMETERS FOR RAC CONFIGURATION TESTING

Table 11) Linux nondefault kernel parameters for RAC configuration testing.

Parameter Name	Value	Description
sunrpc.tcp_slot_table_entries	128	Maximum number of outstanding asynch I/O calls
kernel.sem	250 32000 100 128	Semaphores
net.ipv4.ip_local_port_range	9000 65500	Local port range used by TCP and UDP
net.core.rmem_default	262144	Default TCP receive window size (default buffer size )
net.core.rmem_max	16777216	Maximum TCP receive window size (maximum buffer size)
net.core.wmem_default	262144	Default TCP send window size (default buffer size)
net.core.wmem_max	16777216	Maximum TCP send window size (maximum buffer size)
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

Parameter Name	Value	Description
net.ipv4.tcp_rmem	4096 262144 16777000	Receive buffer size params
net.ipv4.tcp_wmem	4097 262144 16777000	Send buffer size params
net.ipv4.tcp_window_scaling	1	Allow increase of the TCP receive window size
net.ipv4.tcp_syncookies	0	TCP syncookies disabled
net.ipv4.tcp_timestamps	0	TCP timestamps disabled
net.ipv4.tcp_sack	0	TCP selective ACKs disabled
net.ipv4.tcp_no_metrics_save	1	Don't save characteristics of the last connection in the flow cache
net.ipv4.tcp_moderate_rcvbuf	0	Disable TCP received buffer autotuning
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

## 7.8 OTHER LINUX OS SETTINGS

Table 12) Linux shell limits for Oracle.

Configuration File	Settings
/etc/security/limits.conf	oracle hard nofile 65536

## 8 ACKNOWLEDGMENTS

Special thanks to the following people for their contributions:

- John Elliott, NetApp
- Allan Wilson, NetApp
- Keith Griffin, NetApp
- Tom Paisley, NetApp
- Antonio Jose Rodrigues Neto, NetApp

NetApp provides no representations or warranties regarding the accuracy, reliability or serviceability of any information or recommendations provided in this publication, or with respect to any results that may be obtained by the use of the information or observance of any recommendations provided herein. The information in this document is distributed AS IS, and the use of this information or the implementation of any recommendations or techniques herein is a customer's responsibility and depends on the customer's ability to evaluate and integrate them into the customer's operational environment. This document and the information contained herein may be used solely in connection with the NetApp products discussed in this document.

Go further, faster®

© 2011 NetApp, Inc. All rights reserved. No portions of this document may be reproduced without prior written consent of NetApp, Inc. Specifications are subject to change without notice. NetApp, the NetApp logo, Go further, faster, and Data ONTAP are trademarks or registered trademarks of NetApp, Inc. in the United States and/or other countries. Cisco and Nexus are registered trademarks of Cisco Systems. Intel is a registered trademark of Intel Corporation. Linux is a registered trademark of Linus Torvalds. Oracle is a registered trademark of Oracle Corporation. VMware is a registered trademark and vCenter and vSphere are trademarks of VMware, Inc. All other brands or products are trademarks or registered trademarks of their respective holders and should be treated as such. TR-3961-0911



[www.netapp.com](http://www.netapp.com)