Validated Reference Architecture for Oracle Database with Oracle Linux

A Dell Reference Architecture with Oracle Validated Configuration Certification
A pre-tested and validated reference architecture using two socket Dell PowerEdge R720, Dell Compellent SC8000 SAN Storage, Oracle Linux 6.4 with the Unbreakable Enterprise Kernel and Oracle Database 11g Release 2

Dell and Oracle - Joint Infrastructure Solutions
September 2013

Revisions

Date	Description
September 2013	Initial release
October 2013	Dell | Oracle Logo update

Table of contents

Revisions 2

Executive Summary 5

1 Introduction 6

1.1 Oracle Validated Configurations (OVC) 6

2 Audience 8

3 Reference Architecture Components Overview 9

3.1 Dell PowerEdge R720 Server 9

3.2 Dell Compellent SC8000 Storage Controller 9

3.3 Networking Switches 10

3.3.1 Dell Networking S4810 10GbE Switch 10

3.3.2 Dell Networking S55 1GbE Switch 10

3.3.3 Brocade 6510 16Gbps Fibre Channel Switch 11

3.3.4 Oracle Linux with the Unbreakable Enterprise Kernel 11

4 Reference Architecture Configuration 12

4.1 Hardware Configuration Overview 13

4.2 Software Configuration Overview 13

5 Test Methodology 15

5.1 OLT steps 15

5.2 OLT Test Cases 15

6 Results, Analysis and Recommendations 17

7 Conclusion 18

8 References 19

A Hardware configuration details 20

A.1 Server Configuration 20

A.2 Storage Configuration 20

B Software Configuration Details 21

B.1.1 Operating System Configuration 21

B.1.2 Kernel Parameter Settings 21

B.1.3 Limits.conf Settings 21

B.1.4 Oracle ASM RPMS and Oracle Pre-validated RPM 22

B.1.5 Oracle ASM Configuration File Settings 22

B.1.6 Multipath Configuration File Settings 22
C Sales and Support 24

Executive Summary

Dell and Oracle have partnered to deliver end-to-end solutions that integrate data and processes across organizations of various sizes. This new x86 reference architecture continues a long history of joint support and cooperative engineering that has resulted in a stable and pervasive portfolio of database implementation guidance, automation, and optimization making both Dell and Oracle a forerunner in x86 database solutions.

The x86 reference architecture covered in this white paper is targeted for deploying the Oracle Database 11g Release 2 on Dell's latest 12th generation x86 servers optimized with Oracle Linux with the Unbreakable Enterprise Kernel (UEK) operating system. The reference architecture is validated using the Oracle Validated Configuration (OVC) program - a partner vendor program from Oracle. The solution combines Dell's best-in-class hardware with Oracle's industry-leading software, enabling organizations to deploy and manage applications faster than ever before with increased performance, flexibility, and value.

This white paper describes the following:

- Architectural design of the offered reference architecture
- Overview of the Oracle Linux Test (OLT) test suite used to stress and validate the reference architecture
- Best practices and recommendations for implementing the reference architecture
- Details of the hardware, software, test suite configuration, and tuning details that are relevant to this study

1 Introduction

The Active Infrastructure for Oracle is a new solution offering from Dell that is built on the longtime relationship with Oracle. This solution now offers a bundle that includes:

- **Validated Hardware and Software Stack:**
  - Certified Dell x86 hardware infrastructure with Oracle Linux with the Unbreakable Enterprise Kernel and Oracle VM operating systems
  - Pre-tested, validated reference architectures on Oracle Linux and Oracle VM along with best practices and recommendations for Oracle Databases
- **Sales Enablement**: One stop-shop sales enablement to streamline the ordering of the certified and validated hardware and software stack
- **Joint Support**: Joint support from Dell and Oracle on the certified and validated hardware and software stack

This white paper describes a reference architecture that is built for end-users looking to deploy a database on Dell's end-to-end hardware stack that includes servers, storage, and network components along with Oracle's software stack that includes Oracle Linux with the Unbreakable Enterprise Kernel operating system and Oracle Database 11g Release 2.

The key objectives of this reference architecture are to:

- Simplify deployment for end-users with a pre-tested and validated configuration.
- Verify that the system is installed and configured correctly as per Dell and Oracle's best practices.
- Verify the solution robustness against various hardware failures by performing fault injection testing to validate the availability, performance and integrity at each and every layer of the hardware and software stack.
- Simulate a data load on the cluster database and ensure that the cluster performs as expected.

This reference architecture is validated using the Oracle Validated Configuration (OVC) program. The following sub-section introduces the OVC program. Succeeding sections describe the overview of the products used in the reference architecture and the test methodology including the test cases. This is followed by the results and recommendations. It ends with a conclusion.

1.1 Oracle Validated Configurations (OVC)

Oracle Validated Configurations is a partner program offered by Oracle to provide pre-tested, validated reference architectures that include software, hardware, storage, and network components along with documented best practices for running Oracle Databases. In order to validate the reference architectures Oracle provides an Oracle Linux Test (OLT) tool kit that simulates different workloads, performs stress tests, regression tests, system verification tests and destructive tests.

Dell engineers collaborated and worked closely with Oracle engineers to design the reference architecture covered in this white paper and to test it using the OLT tool kit. The OLT tool kit provides an automated mechanism to define, execute and analyze the tests results. It verifies the Linux kernel functionality and

stability essential for the Oracle Database. The OLT kit is used for running tests on Oracle Linux with the Unbreakable Enterprise Kernel and Oracle VM.

For more details on the OVC program, visit http://www.oracle.com/technetwork/topics/linux/validated-configurations-085828.html

2 Audience

The target audiences who will benefit from this white paper are:

- Oracle database administrators
- Oracle solution architects
- Storage administrators
- System administrators
- Oracle database deployment services
- Sales representatives
- Technical support

3 Reference Architecture Components Overview

The following are the critical components that were used to build and test the solution:

- Dell PowerEdge R720 server
- Dell Compellent SC8000 storage controller
- Networking switches:
  - Dell Networking Force 10 rack switches:
    - Dell Networking Force 10 S4810 10GbE switch
    - Dell Networking Force 10 S55 1GbE switch
  - Brocade rack switches:
    - Brocade 6510 16Gbps Fibre Channel switch
- Oracle Linux running Unbreakable Enterprise Kernel

3.1 Dell PowerEdge R720 Server

The Dell PowerEdge R720 uses the Intel Xeon E5-2600 series processor and the Intel chipset architecture in a 2U rack form factor. The PowerEdge R720 is a powerful general purpose platform with highly expandable memory and I/O capabilities. Its extensive scalability and balanced design allows it to handle very demanding workloads. The server features two processor sockets and 24 memory slots supporting 2, 4, 8, 16, or 32 GB DIMMs to meet the memory demands of a virtualized infrastructure.

Energy-efficient design features include power-supply units sized appropriately for system requirements, innovative system-level design efficiency, policy-driven power and thermal management, and highly efficient standards-based Energy Smart components.

Note: For more information on the PowerEdge R720 system, see the PowerEdge R720 Technical Guide.

PowerEdge R720 is certified on Oracle Linux with the Unbreakable Enterprise Kernel and Oracle VM.

Note: For more details refer to the following website:
http://linux.oracle.com/pls/apex/f?p=117:1:1202773984821014:::::

3.2 Dell Compellent SC8000 Storage Controller

The Dell Compellent SC8000 is a 2U rack form factor storage controller with dual six-core, 2.5 GHz (with Turbo) Intel Xeon E5-2640 series processors. It can support full high availability and failover capabilities with a dual controller system and provides enhanced diagnostics capability with the Integrated Dell Remote Access Controller (iDRAC).

A Storage Center Storage Area Network (SAN) built with the Compellent SC8000 scales up to 960 SAS drives per dual controller system and scales-out to multiple systems across multiple sites, monitored by a

single console. It supports simultaneous iSCSI, Fibre Channel (FC) and Fibre Channel over Ethernet (FCoE) front-end interconnects to provide flexibility in the datacenter.

The SC8000 offers exceptional power efficiency with dual redundant 80 Plus® Platinum rated hot-swappable, low wattage power supplies and has six redundant hot-plug fans. Using Fresh Air™ technology, the SC8000 can operate at higher temperatures or even chiller-free environments to help reduce datacenter costs.

Providing up 99.999% availability, Storage Center keeps critical customer data at the ready. With resilient hardware and software and world-class Copilot Support, continuous data availability is ensured.

Note: For more information on the Compellent SC8000 system, see the Compellent SC8000 Spec Sheet.

3.3 Networking Switches

The networking switches used in this reference architecture are Dell Force10 Ethernet switches and Brocade Fibre Channel switches.

High-performance Dell Networking S-Series Managed Ethernet switches are designed for non-blocking access and aggregation while lowering data center costs and improving manageability at the network edge.

Brocade switches meet the demands of hyper-scale, private cloud storage environments by delivering market-leading 16 Gbps Fibre Channel technology and capabilities that support highly virtualized environments.

3.3.1 Dell Networking S4810 10GbE Switch

The Dell Networking Force10 S-Series 4810 (S4810) switch is an ultra-low-latency 10/40 Gigabit Ethernet (GbE) Top-of-Rack (ToR) switch that is designed for applications in high-performance data centers and computing environments. The compact S4810 design provides 48 dual-speed 1/10 GbE (SFP+) ports as well as four 40 GbE QSFP+ uplinks in a 1 rack unit (1RU) package to conserve rack space and simplify the migration to 40 Gbps in the data center core. Each 40 GbE QSFP+ uplink can optionally support four 10 GbE ports with a breakout cable. In addition, the S4810 incorporates multiple architectural features that optimize data center network flexibility, efficiency, and availability, with features like reversible front-to-back or back-to-front airflow for hot/cold aisle environments, and redundant, hot-swappable power supplies and fans.

3.3.2 Dell Networking S55 1GbE Switch

The Dell Networking S-Series Force 10 S55 1/10GbE top-of-rack (ToR) switch is optimized for lowering operational costs while increasing scalability and improving manageability at the network edge. Optimized for high-performance data center applications, the S55 leverages a non-blocking architecture that delivers low-latency L2 and L3 switching to eliminate network bottlenecks

3.3.3 Brocade 6510 16Gbps Fibre Channel Switch

Designed to enable maximum flexibility and investment protection, the Brocade 6510 is configurable in 24, 36, or 48 ports and supports 2, 4, 8, 10, or 16 Gbps speeds in an efficiently designed 1U package. It also provides a simplified deployment process and a point-and-click user interface making it both powerful and easy to use.

3.3.4 Oracle Linux with the Unbreakable Enterprise Kernel

Oracle Linux with the Unbreakable Enterprise Kernel (UEK) brings the latest Linux innovations to market, delivering extreme performance, advanced scalability, and reliability for enterprise applications and systems along with worldwide, enterprise-class, low-cost support.

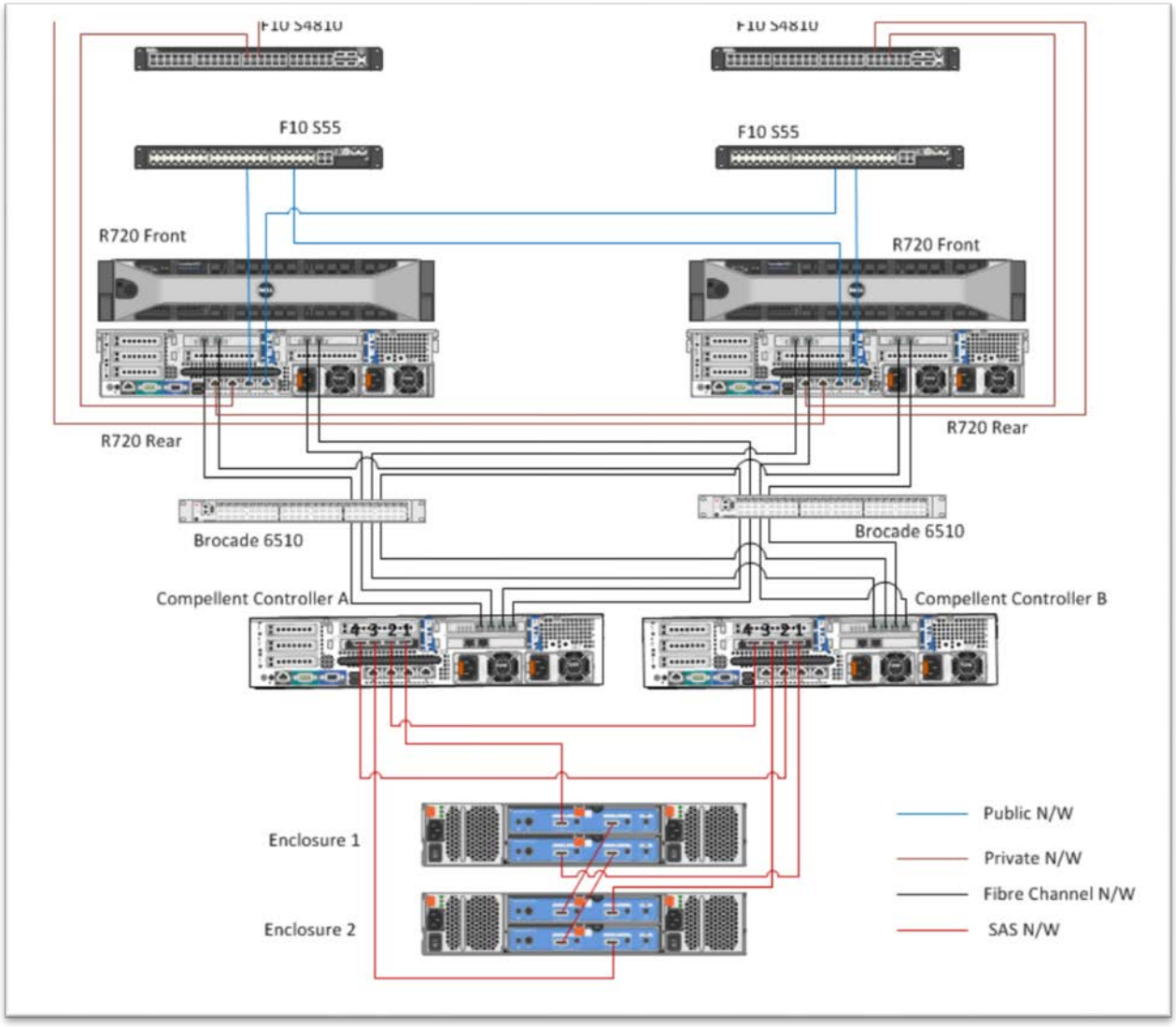
Note: For more information on Oracle Linux with the Unbreakable Enterprise Kernel, see the Oracle Linux with the Unbreakable Enterprise Kernel spec sheet.

4 Reference Architecture Configuration

The reference architecture discussed in this white paper is configured using Dell PowerEdge servers, Dell Compellent Storage, Dell Force10 networking, and Oracle Linux with the Unbreakable Enterprise Kernel. This section provides the architectural diagram and provides an overview of the hardware and software configuration of the reference architecture.

Figure 1 below shows the architectural diagram of the reference architecture covered in this paper.

Figure 1. Architectural Diagram



4.1 Hardware Configuration Overview

The features of the hardware configuration are:

- **Hardware Components High Availability**: As seen in Figure 1, the architecture is designed in such a way that there is no single point of failure and redundancy is incorporated into every mission critical component of the solution that includes servers, storage controllers, networking components, and the database. The design also includes high availability at the sub-component level.
- **Database High Availability**: For database high availability, Oracle Real Application Cluster (RAC) Database is configured on two Dell PowerEdge R720 servers. Each R720 hosts a Brocade quad port Ethernet rack Network Daughter Card (rNDC), in that two are 1GbE Base-T ports and two are 10GbE SPF+ ports. For network high availability, from each server, the two 1GbE Base-T ports are used for public network and are connected to two separate Dell Force 10 S55 1GbE rack switches. Similarly, from each server, the two 10GbE SPF+ ports are used for Oracle private interconnect traffic and are connected to two separate Dell Force 10 S4810 10GbE rack switches.
- **SAN High Availability**: For SAN high availability, each R720 that consists of two Qlogic QLE2662 16Gbps Host Bus Adapters (HBAs) is connected to two Dell Compellent SC8000 storage controllers using two Brocade 6510 16Gbps Fibre Channel rack switches. The storage controllers are connected to two Compellent expansion enclosures in the backend.
  As part of Compellent best practice, the four fibre channel front end ports on each of the controllers are configured as virtual ports. This allows for high availability at the controller port level. If any one of the four physical ports becomes inaccessible, then the virtual IP of that port fails over to one of the three remaining active ports.
  Similarly, as shown in Figure 1, the cabling best practice between the front end ports of the two controllers enables IPC at the controller level. If one controller were to go down, the database will continue to access the storage enclosures using the other active controller.

For more details on the hardware configuration, refer to *Hardware configuration details* in the appendix section. For additional Compellent best practice guides, refer to the *References* section

4.2 Software Configuration Overview

The features of the software configuration are:

- **Operating System**: The Oracle database servers were installed with Oracle Linux 6 Update 4 running the Unbreakable Enterprise Kernel.
- **RPMs**: The Oracle users, groups, permissions, and the recommended kernel and database parameter were set using the Oracle's validated pre-install RPM for Oracle Database 11g Release 2. For details of these settings refer to the *Software Configuration* section in the Appendix.
- **Testing Software**: The Oracle Linux Test (OLT) tool kit was used to stress test and validate the reference architecture covered in this paper. The OS configuration, network and the storage requirements needed to conduct the OLT tests were configured following the OLT user's guide. Oracle Grid Infrastructure and Real Application Cluster (RAC) Database 11g Release 2 was installed using the silent install test from the OLT tool kit.

- **Storage Volumes**: The necessary volumes needed for the OLT tests were created as thin provisioned volumes on the Compellent storage array. The Compellent Storage Center was configured with the default policy settings with no data progression enabled. The disks used for Oracle Clusterware and datafiles were configured as Oracle Automatic Storage Management (ASM) disks. ASMLib was used to configure these ASM disks using the OLT user's guide. For recommended best practice on the Oracle ASM configuration refer to *Oracle ASM* Appendix section.
- **Multipath configuration:** On the Linux host, the storage volumes were configured using the device-mapper. The multipath settings were configured using Compellent's best practice for Linux. For details on the recommended settings refer to *Multipath Configuration File Settings* section in the appendix.

Note: For best practice on how to configure multipath in Enterprise Linux 6.x for Dell Compellent storage solution, refer to http://en.community.dell.com/techcenter/enterprise-solutions/w/oracle_solutions/4983.how-to-configure-multipath-on-enterprise-linux-6-x-for-dell-compellent-storage.aspx

5 Test Methodology

This section provides the high level steps for installing the Oracle Linux Test tool Kit and some of the high level test cases performed using this tool kit.

5.1 OLT steps

The following high level steps are involved to setup the OLT tool kit and to execute the tests:

1. Verification of hardware requirements for OLT testing
2. OLT pre-installation
    a. Operating system configuration
    b. Storage configuration
3. OLT installation and configuration
4. OLT test execution

Note: For more details on the above individual steps, refer to OLT user's guide.

5.2 OLT Test Cases

The OLT toolkit comprises of the following tests that were used in this validation:

1. **Stress Tests for RACs:** These tests stress the CPU, memory and the I/O sub-system across all the nodes in the RAC.
2. **RAC Destructive Tests:** These tests simulate ASM instance crash, Cluster Ready Services (CRS) crash, and database instance crash on one of the nodes and verify the database functionality on the other node. These tests also verify the Virtual IP (VIP) failover, CRS services failover to the available node.
3. **Multipath tests:** In this, various fault injection tests are conducted at each layer of the hardware stack.
    a. **Host side Testing:** In this test case, with IO stress test running, single and multiple FC HBA cables are pulled from one of the nodes to simulate HBA failure. This test verifies the high availability of the host side HBAs and ensures that the multiple paths are functioning as expected.
    b. **Switch side Testing:** In this, one of the test cases is to disable a port on the FC switch that connects the switch to the host HBA to simulate the switch port failure. Another test case is to shut down one of the FC switches to simulate a switch failure. Both these test cases verify the high availability of the FC switches and ensures that the multiple paths are functioning as expected.
    c. **Storage side Testing:** In this, one of the test cases is to unplug the cable from one of the front end ports on the controller to simulate storage port failure. This test case verifies the virtual port failover within the Compellent controller. Another test case is to shut down one of the Compellent controllers to simulate the controller failure. This test case verifies the failover of the volumes on to the other Compellent controller.

Note: For more details on the different OLT test cases, refer to OLT test coverage guide.

6 Results, Analysis and Recommendations

During the stress testing two issues were discovered. The details of the two issues along with the recommended workarounds are listed below:

Summary	IO Performance degradation due to transparent hugepages (THP) enabled, when the system is under memory pressure.
Affects	2.6.39-100 and upwards
Symptom	With THP enabled by default, RAC nodes may reboot due to slow I/O, when the system is under moderate memory pressure
Solution/ Workaround	Disable hugepages by the command `echo never > /sys/kernel/mm/transparent_hugepage/enabled`, and `echo never > /sys/kernel/mm/transparent_hugepage/defrag`

Summary	The parameter `path_checker` when set to `readsector0` in the `multipath.conf` file causes the RAID controller driver to reset and reboot the OS
Affects	OL6.4 with storage controller (PERC)H710P, H710P Adapter, H710P mini
Symptom	In the `/etc/multipath.conf` file, under the `defaults` section, setting the '`path_checker readsector0`' option causes the `megaraid_sas` driver to reset with messages in `dmesg` as "`test kernel: megaraid_sas: Reset successful`"
Solution/ Workaround	Uncomment or remove the '`path_checker readsector0`' option in the `defaults` section in `/etc/multipath.conf` file

Results from all the testing, including the ones above after the workaround was applied, were successful. During all the tests the Oracle Database RAC performed as expected.

7 Conclusion

With this reference architecture successfully completing and passing the comprehensive suite of OLT test cases, we have validated the proper functionality of Oracle Database 11g Release 2 RAC solution running with Oracle Linux with the Unbreakable Enterprise Kernel. This reference architecture is validated and approved by Oracle as an Oracle Validated Configuration. This validated reference architecture can be found at the following link:

http://linux.oracle.com/pls/apex/f?p=102:2:299028432187263::NO::P2_VC_ID:626

All the best practices and recommendations provided in this white paper serve as a proof-point of a robust Oracle database hardware and software solution that is specifically integrated and optimized for running on Dell's enterprise hardware and on Oracle Linux with the Unbreakable Enterprise Kernel.

With this pre-tested and validated reference architecture jointly supported by Dell and Oracle, customers can now purchase with confidence and help reduce risks and save on cost for running their mission critical database applications in the datacenters.

With the new worldwide alliance between Dell and Oracle, customers can expect more innovations on the horizon, as Dell and Oracle will continue to combine forces to drive systems management integration and innovation and further improving the manageability and enterprise class performance of the overall solution.

8 References

- Visit Dell Compellent's Knowledge Center: http://kc.compellent.com to access the following documents:
  - CT-SC040 & SC8000 Connectivity Guide
  - Dell Compellent - Oracle Best Practices Guide
- Dell Server Hardware Certification List (HCL) certified with Oracle Linux and Oracle VM: http://linux.oracle.com/pls/apex/f?p=117:1:2526069251147525::NO:RP::
- Dell Storage Hardware Certification List (HCL) certified with Oracle VM: http://linux.oracle.com/pls/apex/f?p=117:3:46019236825768::NO:RP::
- Dell best practices, deployment and how-to guides, deployment automation tools and white papers for Oracle Databases: www.delltechcenter.com/oracle

A Hardware configuration details

A.1 Server Configuration

Server Configuration	
Server	PowerEdge R720
BIOS	v1.6.0
iDRAC7 ESM FW	V1.35.35
LC2	v1.1.1.1.8
CPU	2 x Intel Xeon E5-2690 2.90GHz; cache size 20MB
Memory	128GB (8x16GB); 1600MHz; Populated DIMM Slots: A1-A4; B1-B4
PERC H710 Adapter	FW: 21.2.0-0007 Driver: 06.505.02.00 (megaraid_sas)
Broadcom BCM57800 10-Gigabit Ethernet (2 numbers for private)	Driver name: bnx2x Driver version: 1.76.54 Firmware : 7.0.47(bc 7.0.49)
Broadcom BCM57800 1-Gigabit Ethernet (2 numbers for public)	Driver name: bnx2x Driver version: 1.76.54 Firmware : 7.0.47(bc 7.0.49)
Qlogic QLE 2662 16 Gbps	Driver name: qla2xxx Driver version: 8.05.00.03.39.0-k
Open Manage Server Assistant (OMSA)	v7.1.2 (prev v7.0.0)
Team Interfaces	Bond0 (public) Bond1 (private)

A.2 Storage Configuration

Storage Configuration	
Storage	Compellent SC8000
Firmware	Storage Center 6.3
Controllers	2
Enclosures	2
Disks	SAS 15k 300GB
Tiering	Tier 1

B Software Configuration Details

B.1.1 Operating System Configuration

Software Configuration	
OS	Oracle Linux 6 Update 4 – 2.6.39-400.109.4.el6uek.x86_64
Oracle	Oracle Database 11g Release 2 Grid Infrastructure and Real Application Clusters for x86-64

B.1.2 Kernel Parameter Settings

Parameter Name	Value
net.ipv4.ip_forward	0
net.ipv4.conf.default.rp_filter	1
net.ipv4.conf.default.accept_source_route	0
kernel.sysrq	0
kernel.core_uses_pid	1
net.ipv4.tcp_syncookies	1
kernel.msgmnb	65536
kernel.msgmax	65536
fs.file-max	6815744
kernel.sem	250 32000 100 128
kernel.shmmni	4096
kernel.shmall	1073741824
kernel.shmmax	4398046511104
net.core.rmem_default	262144
net.core.rmem_max	4194304
net.core.wmem_default	262144
net.core.wmem_max	1048576
fs.aio-max-nr	1048576
net.ipv4.ip_local_port_range	9000 65500

B.1.3 Limits.conf Settings

Parameter Name	Value
oracle soft nofile	1024
oracle hard nofile	65536
oracle soft nproc	2047
oracle hard nproc	16384
oracle soft stack	10240
oracle hard stack	32768
oracle soft core	unlimited
oracle hard core	unlimited
oracle soft memlock	50000000
oracle hard memlock	50000000

B.1.4 Oracle ASM RPMS and Oracle Pre-validated RPM

oracleasmlib-2.0.4-1.el6.x86_64
oracleasm-support-2.1.8-1.el6.x86_64
oracle-rdbms-server-11gR2-preinstall-1.0-7.el6.x86_64

B.1.5 Oracle ASM Configuration File Settings

The recommended settings in the `/etc/sysconfig/oracleasm` configuration file are as follows:

Parameter Name	Value
ORACLEASM_ENABLED	true
ORACLEASM_UID	oracle
ORACLEASM_GID	oinstall
ORACLEASM_SCANBOOT	True
ORACLEASM_SCANORDER	"dm"
ORACLEASM_SCANEXCLUDE	""
ORACLEASM_CLEARBOOT	True
ORACLEASM_USE_LOGICAL_BLOCK_SIZE	True

B.1.6 Multipath Configuration File Settings

The recommended settings in the `/etc/multipath.conf` file are as follows:

```
blacklist_exceptions {
   wwid "36000d31000ece2000000000000000051"
   wwid "36000d31000ece2000000000000000050"
   }
defaults {
   user_friendly_names yes
}
multipaths {
   multipath {
         wwid              36000d31000ece2000000000000000051
         alias             OCRVOTE
```

```
    }
    multipath {
                wwid                    36000d31000ece2000000000000000050
                alias                   DATA
        }
}
blacklist {
}
```

C Sales and Support

Customers interested in purchasing this pre-tested and validated reference architecture featuring Oracle Database 11g Release 2 running on Oracle Linux with the Unbreakable Enterprise Kernel, call the Dell sales representative and enquire about Dell Active Infrastructure for Oracle.

The entire Dell Active Infrastructure for Oracle hardware and software stack is jointly supported by Dell and Oracle.

For more information, see **dell.com/contactdell**.