Oracle Real Application Cluster on Cisco Unified Computing System and EMC VNX Storage

A Cisco Validated Design for Oracle 11g R2 (11.2.0.2) RAC Running on Oracle Linux and Cisco UCS B-Series Servers

Last Updated: May 22, 2012

Cisco Validated Design

Building Architectures to Solve Business Problems

About the Authors

Tushar Patel, Principle Engineer, Data Center Group, Cisco

Tushar Patel is a Principle Engineer for the Cisco Systems Data Center group. Tushar has nearly fifteen years of experience in database architecture, design and performance. Tushar also has strong background in Intel x86 architecture, storage and virtualization. He has helped a large number of enterprise customers evaluate and deploy various database solutions. Tushar holds a Master's degree in Computer Science and has presented to both internal and external audiences at various conferences and customer events.

Niranjan Mohapatra, Technical Marketing Engineer, Data Center Group, Cisco

Niranjan Mohapatra is a Technical Marketing Engineer for Cisco Systems Data Center group and specialist on Oracle RAC RDBMS. He has over 13 years of extensive experience with Oracle RDBMS and associated tools. Niranjan has worked as a TME and a DBA handling production systems in various organizations. He holds a Master of Science (MSc) degree in Computer Science and is also an Oracle Certified Professional (OCP) and Storage Certified Professional.

About Cisco Validated Design (CVD) Program

The CVD program consists of systems and solutions designed, tested, and documented to facilitate faster, more reliable, and more predictable customer deployments. For more information visit http://www.cisco.com/go/designzone.

Deploying Oracle 11gR2 (11.2.0.2) Real Application Clusters on Oracle Linux 5.7 Using Cisco Unified Computing System and EMC VNX Storage

Overview

This Cisco Validated Design describes how the Cisco Unified Computing System™ can be used in conjunction with EMC® VNX™ storage systems to implement an Oracle Real Application Clusters (RAC) solution that is an Oracle Certified Configuration. The Cisco Unified Computing System provides the compute, network, and storage access components of the cluster, deployed as a single cohesive system. The result is an implementation that addresses many of the challenges that database administrators and their IT departments face today, including needs for a simplified deployment and operation model, high performance for Oracle RAC software, and lower total cost of ownership (TCO). This document introduces the Cisco Unified Computing System and provides instructions for implementing it; it concludes with an analysis of the cluster's performance and reliability characteristics.

Introduction

Data powers essentially every operation in a modern enterprise, from keeping the supply chain operating efficiently to managing relationships with customers. Oracle RAC brings an innovative approach to the challenges of rapidly increasing amounts of data and demand for high performance. Oracle RAC uses a horizontal scaling (or scale-out) model that allows organizations to take advantage of the fact that the price of one-to-four-socket x86-architecture servers continues to drop while their processing power increases unabated. The clustered approach allows each server to contribute its processing power to the overall cluster's capacity, enabling a new approach to managing the cluster's performance and capacity.

Corporate Headquarters:
Cisco Systems, Inc., 170 West Tasman Drive, San Jose, CA 95134-1706 USA

Leadership From Cisco

Cisco is the undisputed leader in providing network connectivity in enterprise data centers. With the introduction of the Cisco Unified Computing System, Cisco is now equipped to provide the entire clustered infrastructure for Oracle RAC deployments. The Cisco Unified Computing System provides compute, network, virtualization, and storage access resources that are centrally controlled and managed as a single cohesive system. With the capability to centrally manage both blade and rack-mount servers, the Cisco Unified Computing System provides an ideal foundation for Oracle RAC deployments.

Historically, enterprise database management systems have run on costly symmetric multiprocessing servers that use a vertical scaling (or scale-up) model. However, as the cost of one-to-four-socket x86-architecture servers continues to drop while their processing power increases, a new model has emerged. Oracle RAC uses a horizontal scaling, or scale-out, model, in which the active-active cluster uses multiple servers, each contributing its processing power to the cluster, increasing performance, scalability, and availability. The cluster balances the workload across the servers in the cluster, and the cluster can provide continuous availability in the event of a failure.

Oracle Certified Configuration

All components in an Oracle RAC implementation must work together flawlessly, and Cisco has worked closely with EMC and Oracle to create, test, and certify a configuration of Oracle RAC on the Cisco Unified Computing System. Cisco's Oracle Certified Configurations provide an implementation of Oracle Database with Real Application Clusters technology consistent with industry best practices. For back-end SAN storage, the certification environment included an EMC VNX storage system with a mix of SAS drives and state-of-the-art Flash drives (FDs) to further speed performance.

Benefits of the Configuration

The Oracle Certified Configuration of Oracle RAC on the Cisco Unified Computing System offers a number of important benefits.

Simplified Deployment and Operation

Because the entire cluster runs on a single cohesive system, database administrators no longer need to painstakingly configure each element in the hardware stack independently. The system's compute, network, and storage-access resources are essentially stateless, provisioned dynamically by Cisco® UCS Manager. This role-based and policy-based embedded management system handles every aspect of system configuration, from a server's firmware and identity settings to the network connections that connect storage traffic to the destination storage system. This capability dramatically simplifies the process of scaling an Oracle RAC configuration or rehosting an existing node on an upgrade server. Cisco UCS Manager uses the concept of service profiles and service profile templates to consistently and accurately configure resources. The system automatically configures and deploys servers in minutes, rather than the hours or days required by traditional systems composed of discrete, separately managed components. Indeed, Cisco UCS Manager can simplify server deployment to the point where it can automatically discover, provision, and deploy a new blade server when it is inserted into a chassis.

The system is based on a 10-Gbps unified network fabric that radically simplifies cabling at the rack level by consolidating both IP and Fiber Channel traffic onto the same rack-level 10-Gbps converged network. This "wire-once" model allows in-rack network cabling to be configured once, with network

features and configurations all implemented by changes in software rather than by error-prone changes in physical cabling. This Oracle Certified Configuration not only supports physically separate public and private networks; it provides redundancy with automatic failover.

High-Performance Platform for Oracle RAC

The Cisco UCS B-Series Blade Servers used in this certified configuration feature Intel Xeon 5670 series processors that deliver intelligent performance, automated energy efficiency, and flexible virtualization. Intel Turbo Boost Technology automatically boosts processing power through increased frequency and use of hyperthreading to deliver high performance when workloads demand and thermal conditions permit.

The Cisco Unified Computing System's 10-Gbps unified fabric delivers standards-based Ethernet and Fiber Channel over Ethernet (FCoE) capabilities that simplify and secure rack-level cabling while speeding network traffic compared to traditional Gigabit Ethernet networks. The balanced resources of the Cisco Unified Computing System allow the system to easily process an intensive online transaction processing (OLTP) and decision-support system (DSS) workload with no resource saturation.

Safer Deployments with Certified and Validated Configurations

Cisco and Oracle are working together to promote interoperability of Oracle's next-generation database and application solutions with the Cisco Unified Computing System, helping make the Cisco Unified Computing System a simple and reliable platform on which to run Oracle software. In addition to the certified Oracle RAC configuration described in this document, Cisco, Oracle and EMC have certified single-instance database implementations of Oracle Database 11gR2 on Oracle Linux.

Implementation Instructions

This Cisco Validated Design introduces the Cisco Unified Computing System and discusses the ways it addresses many of the challenges that database administrators and their IT departments face today. The document provides an overview of the certified Oracle RAC configuration along with instructions for setting up the Cisco Unified Computing System and the EMC VNX storage system, including database table setup and the use of flash drives. The document reports on Cisco's performance measurements for the cluster and a reliability analysis that demonstrates how the system continues operation even when commonly encountered hardware faults occur.

Introducing the Cisco Unified Computing System

The Cisco Unified Computing System addresses many of the challenges faced by database administrators and their IT departments, making it an ideal platform for Oracle RAC implementations.

Comprehensive Management

The system uses an embedded, end-to-end management system that uses a high-availability active-standby configuration. Cisco UCS Manager uses role and policy-based management that allows IT departments to continue to use subject-matter experts to define server, network, and storage access policy. After a server and its identity, firmware, configuration, and connectivity are defined, the server, or a number of servers like it, can be deployed in minutes, rather than the hours or days that it typically

takes to move a server from the loading dock to production use. This capability relieves database administrators from tedious, manual assembly of individual components and makes scaling an Oracle RAC configuration a straightforward process.

Radical Simplification

The Cisco Unified Computing System represents a radical simplification compared to the way that servers and networks are deployed today. It reduces network access-layer fragmentation by eliminating switching inside the blade server chassis. It integrates compute resources on a unified I/O fabric that supports standard IP protocols as well as Fiber Channel through FCoE encapsulation. The system eliminates the limitations of fixed I/O configurations with an I/O architecture that can be changed through software on a per-server basis to provide needed connectivity using a just-in-time deployment model. The result of this radical simplification is fewer switches, cables, adapters, and management points, helping reduce cost, complexity, power needs, and cooling overhead.

High Performance

The system's blade servers are based on the Intel Xeon 5670 and 7500 series processors. These processors adapt performance to application demands, increasing the clock rate on specific processor cores as workload and thermal conditions permit. These processors, combined with patented Cisco Extended Memory Technology, deliver database performance along with the memory footprint needed to support large in-server caches. The system is integrated within a 10 Gigabit Ethernet-based unified fabric that delivers the throughput and low-latency characteristics needed to support the demands of the cluster's public network, storage traffic, and high-volume cluster messaging traffic.

Overview of the Certified Configuration

The Cisco Unified Computing System used for the certified configuration is based on Cisco B-Series Blade Servers; however, the breadth of Cisco's server and network product line suggests that similar product combinations will meet the same requirements. The Cisco Unified Computing System uses a form-factor-neutral architecture that will allow Cisco C-Series Rack-Mount Servers to be integrated as part of the system using capabilities planned to follow the product's first customer shipment (FCS). Similarly, the system's core components -- high-performance compute resources integrated using unified fabric -- can be integrated manually today using Cisco C-Series servers and Cisco Nexus™ 5000 Series Switches.

The system used to create the Oracle Certified Configuration is built from the hierarchy of components illustrated in Figure 1.

- The Cisco UCS 6120XP 20-Port Fabric Interconnect provides low-latency, lossless, 10-Gbps unified fabric connectivity for the cluster. The fabric interconnect provides connectivity to blade server chassis and the enterprise IP network. Through an 8-port, 4-Gbps Fiber Channel expansion card, the fabric interconnect provides native Fiber Channel access to the EMC VNX storage system. Two fabric interconnects are configured in the cluster, providing physical separation between the public and private networks and also providing the capability to securely host both networks in the event of a failure.

- The Cisco UCS 2104XP Fabric Extender brings the unified fabric into each blade server chassis. The fabric extender is configured and managed by the fabric interconnects, eliminating the complexity of blade-server-resident switches. Two fabric extenders are configured in each of the cluster's two blade server chassis. Each one uses two of the four available 10-Gbps uplinks to connect to one of the two fabric interconnects.
- The Cisco UCS 5108 Blade Server Chassis houses the fabric extenders, up to four power supplies, and up to eight blade servers. As part of the system's radical simplification, the blade server chassis is also managed by the fabric interconnects, eliminating another point of management. Two chassis were configured for the Oracle RAC described in this document.

  The blade chassis supports up to eight half-width blades or up to four full-width blades. The certified configuration uses eight (four in each chassis) Cisco UCS B200 M2 Blade Servers, each equipped with two quad-core Intel Xeon 5670 series processors at 2.933 GHz. Each blade server was configured with 96 GB of memory. Larger memory footprint can be accommodated through the use of a Cisco UCS B250 M2 Extended Memory Blade Server.
- The blade server form factor supports a range of mezzanine-format Cisco UCS network adapters, including a 10 Gigabit Ethernet network adapter designed for efficiency and performance, the Cisco UCS M81KR Virtual Interface Card designed to deliver outstanding performance and full compatibility with existing Ethernet and Fiber Channel environments. These adapters present both an Ethernet network interface card (NIC) and a Fiber Channel host bus adapter (HBA) to the host operating system. They make the existence of the unified fabric transparent to the operating system, passing traffic from both the NIC and the HBA onto the unified fabric. This certified configuration uses a Cisco UCS M81KR Virtual Interface Network Adapter that provides 20-Gbps of performance per blade server.

Figure 1 Cisco Unified Computing System Components

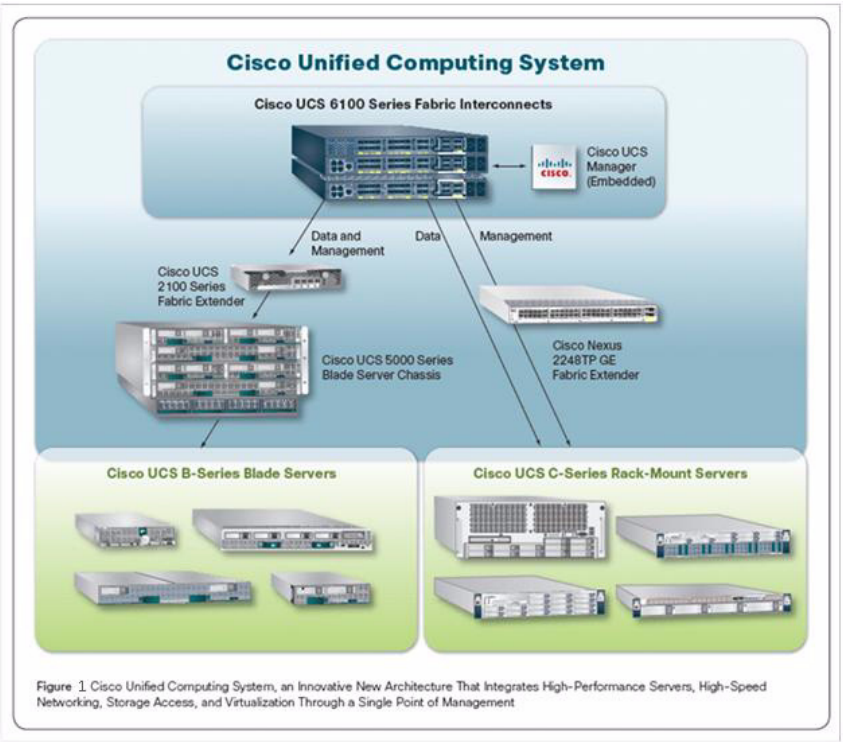


Figure 1 Cisco Unified Computing System, an Innovative New Architecture That Integrates High-Performance Servers, High-Speed Networking, Storage Access, and Virtualization Through a Single Point of Management

Figure 2 Cisco Unified Computing System Component Descriptions

The Solution Infrastructure Components

Figure 3 Cisco Nexus 5548UP Switch

The Cisco Nexus 5548UP switch delivers innovative architectural flexibility, infrastructure simplicity, and business agility, with support for networking standards. For traditional, virtualized, unified, and high-performance computing (HPC) environments, it offers a long list of IT and business advantages, including:

Architectural Flexibility

- Unified ports that support traditional Ethernet, Fiber Channel (FC), and Fiber Channel over Ethernet (FCoE)
- Synchronizes system clocks with accuracy of less than one microsecond, based on IEEE 1588
- Supports secure encryption and authentication between two network devices, based on Cisco TrustSec IEEE 802.1AE
- Offers converged Fabric extensibility, based on emerging standard IEEE 802.1BR, with Fabric Extender (FEX) Technology portfolio, including:
  - Cisco Nexus 2000 FEX

  - Adapter FEX
  - VM-FEX

Infrastructure Simplicity

- Common high-density, high-performance, data-center-class, fixed-form-factor platform
- Consolidates LAN and storage
- Supports any transport over an Ethernet-based fabric, including Layer 2 and Layer 3 traffic
- Supports storage traffic, including iSCSI, NAS, FC, RoE, and IBoE
- Reduces management points with FEX Technology

Business Agility

- Meets diverse data center deployments on one platform
- Provides rapid migration and transition for traditional and evolving technologies
- Offers performance and scalability to meet growing business needs

Specifications at a Glance

- A 1 -rack-unit, 1/10 Gigabit Ethernet switch
- 32 fixed Unified Ports on base chassis and one expansion slot totaling 48 ports
- The slot can support any of the three modules: Unified Ports, 1/2/4/8 native Fiber Channel, and Ethernet or FCoE
- Throughput of up to 960 Gbps

EMC VNX Unified Storage System

EMC VNX series unified storage systems deliver uncompromising scalability and flexibility for the mid-tier while providing market-leading simplicity and efficiency to minimize total cost of ownership.

Based on the powerful new family of Intel Xeon-5600 processors, the EMC VNX implements a modular architecture that integrates hardware components for block, file, and object with concurrent support for native NAS, iSCSi, Fiber Channel, and FCoE protocols. The unified configuration includes the following rack mounted enclosures:

- Disk processor enclosure (holds disk drives) or storage processor enclosure (requires disk drive tray) plus stand-by power system to deliver block protocols.
- One or more data mover enclosures to deliver file protocols (required for File and Unified configurations)
- Control station (required for File and Unified configurations)

A robust platform designed to deliver five 9s availability, the VNX series enables organizations to dynamically grow, share, and cost-effectively manage multi-protocol file systems and multi-protocol block storage access. The VNX series has been expressly designed to take advantage of the latest innovation in Flash drive technology, maximizing the storage system's performance and efficiency while minimizing cost per GB.

Finally, Cisco and EMC are collaborating on solutions and services to help build, deploy, and manage IT infrastructures that adapt to changing needs. Industry-leading EMC information infrastructure and intelligent Cisco networking products, including the Cisco Unified Computing System, will reduce the complexity of data centers.

Together, EMC and Cisco provide comprehensive solutions that can benefit customers now and in the future, including:

- High-performance storage and SANs that reduce total cost of ownership
- Disaster recovery to protect data and improve compliance
- Combined computing, storage, networking, and virtualization technologies

Figure 4 EMC VNX Storage

Leveraging EMC software creates additional benefits which can be derived when using products such as:

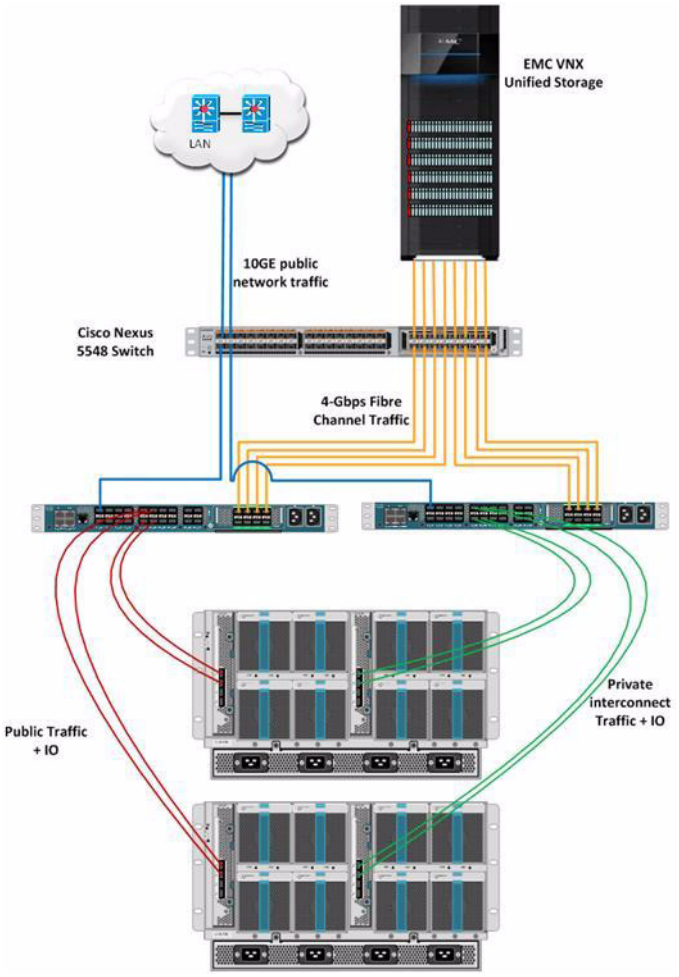
- Fast Cache: Dynamically absorbs unpredicted spikes in system workloads.
- FAST VP: Tiers data from high-performance to high-capacity drives in one-gigabyte increments, with Fully Automated Storage Tiering for Virtual Pools, resulting in overall lower costs, regardless of application type or data age.
- FAST Suite: Automatically optimizes for the highest system performance and the lowest storage cost simultaneously (includes FAST VP and FAST Cache).
- EMC PowerPath®: Provides automated data path management and load-balancing capabilities for heterogeneous server, network, and storage deployed in physical and virtual environments. For additional information reference: http://www.emc.com/collateral/software/data-sheet/l751-powerpath-ve-multipathing-ds.pdf.
- EMC Unisphere™: Delivers simplified management via a single management framework for all NAS, SAN, and replication needs. For additional information on Unisphere, reference http://www.emc.com/collateral/software/data-sheet/h7303-unisphere-ds.pdf.

For additional information on the EMC VNX Series reference: http://www.emc.com/storage/vnx/vnx-series.htm.

Solution Overview

The configuration presented in this Cisco Validated Design is based on the Oracle Database 11g Release 2 with Real Application Clusters technology certification environment specified for an Oracle RAC and EMC VNX storage system (Figure 5).

Figure 5 Oracle Database 11gR2 with Real Application Clusters technology on Cisco Unified Computing System and EMC VNX Storage



In Figure 5, the red lines indicate the public network connecting to Fabric Interconnect A, and the green lines indicate the private interconnects connecting to Fabric Interconnect B. For Oracle RAC environments, it is a best practice to keep all private interconnect (intra-blade) traffic to one Fabric interconnect. The public and private VLANs spanning the fabric interconnects help ensure the connectivity in case of link failure. Note that the FCoE communication takes place between the Cisco Unified Computing System chassis and fabric interconnects (red and green lines). This is a typical configuration that can be deployed in a customer's environment. The best practices and setup recommendations are described in subsequent sections of this document.

Note For this setup, we used a single Nexus 5548UP switch as the focus is on Oracle database testing. For a typical datacenter setup, the best practice scenario is to setup a pair of Layer 2 network switches to avoid a single point of failure.

Detailed Topology

Figure 6 shows two chassis, housing four blades each, that were used for this eight-node Oracle RAC solution. Tables 1 through 5 list the configuration details for all the server, LAN, and SAN components that were used for testing.

Figure 6 Detailed Topology of the Public Network and Oracle RAC Private Interconnects

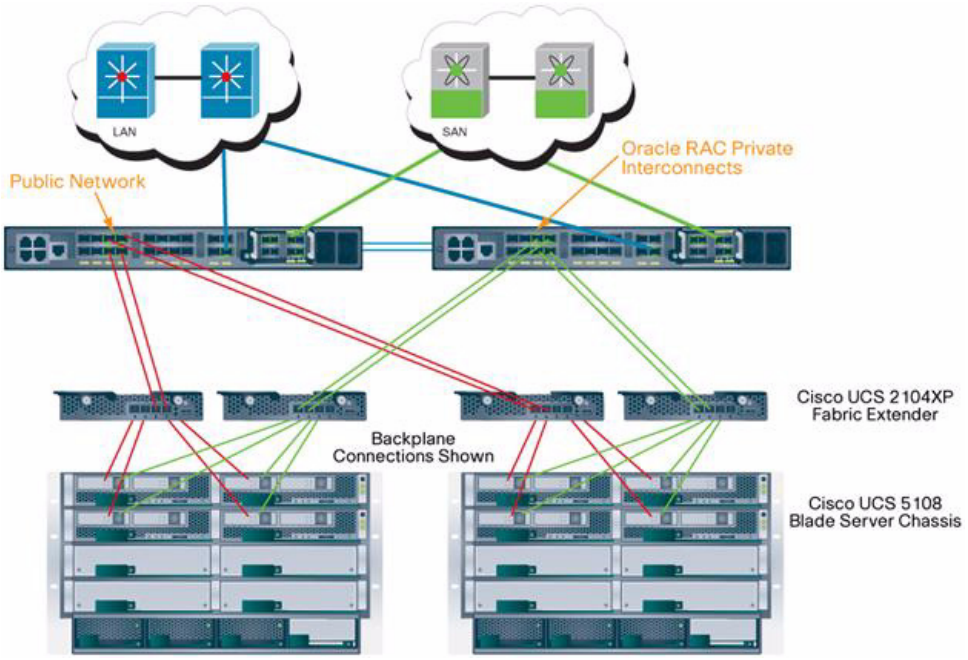


Table 1 Physical Cisco Unified Computing System Server Configuration

Quantity	Description
2	Cisco UCS 5108 Blade Server Chassis, with 4 power supply units, 8 fans, and 2 fabric extenders
8	Cisco UCS B200 M2 Half width Blade Servers
16	Quad-core Intel Xeon 5670 series 2.933-GHz processor (2 per blade server)
96	8-GB DDR3 DIMM, 1333 MHz (12 per server, totaling 96 GB per blade server)
8	Cisco UCS M81KR Virtual Interface Network Adapter, PCIe, 2 ports, and 10 Gigabit Ethernet (1 per blade server)
16	146-GB SAS, 10,000 RPM, Small Form-Factor (SFF) hard disk drive (HDD) hot pluggable (1 or 2 per blade)
2	Cisco UCS 6120XP 20-Port Fabric Interconnect with 2 power supply units and 2 fans
2	8-port, 4-Gbps Fiber Channel expansion port module
8	4-Gbps Fiber Channel Small Form-Factor Pluggable Plus (SFP+)
8	10GBASE-CU SFP+ cable (5 meters)
8	Fiber cables for connectivity to Fiber Channel and 10 Gigabit Ethernet

Note This test used only one SAN and LAN switch. Use of redundant SAN and LAN switches is highly recommended to avoid a single-point failure.

Table 2 LAN and SAN Components

Quantity	Description
1	Cisco Nexus 5548UP Switch (A pair strongly recommended for Data Center deployments)
VLAN ID	**VLAN Configuration**
134	Public VLAN
10	Private traffic VLAN (Private traffic VLAN must be configured on this switch to ensure traffic flow in partial link faiure as discussed in later section.
11	If using Oracle HAIP, one VLAN for each additional HAIP interface.
VSAN ID	**VLAN Configuration**
9	Oracle database VSAN

Table 3 Storage Configuration

Quantity	Description
1	EMC VNX5500™ Storage System
105	600-GB, 15,000 RPM, SAS disk drives
15	200-GB EFDs

Table 4 Operating System and RPM Component (installed on all Oracle Cluster Nodes)

Description	OS and RPM(s)
Operating System (64 bit)	Oracle Linux 5.7 x86_64 (2.6.18-274.el5)
Required RPM(s) by Oracle	oracleasm-2.6.18-164.el5-2.0.5-1.el5, oracleasm-support-2.1.3-1.el5 oracle-validated-1.0.0-18.el5, oracle-logos-4.9.17-6
Required RPM(s) by EMC (To be installed on ALL Oracle Cluster Nodes to support EMC PowerPath and Hostagent)	EMCpower.LINUX-5.5.0.00.00-275.3.1.00.00-111.rhel5.x86_64.rpm HostAgent-Linux-64-x86-en_US-1.1.32.1.0042
Required RPM(s) on a DNS server (To be installed on the DNS Server to support DNS and DHCP and in turn to support Oracle GNS)	bind-9.3.6-4.P1.el5.x86_64.rpm , bind-libs-9.3.6-4.P1.el5.x86_64.rpm , bind-utils-9.3.6-4.P1.el5.x86_64.rpm , dhcp-3.0.5-21.el5.rpm

Table 5 Software Components

Description
Oracle Database 11*g* Release 2 (11.2.0.2.0) with GRID Infrastructure Oracle Patch 9974223 Oracle Patch 6880880 Oracle Patch 10425672

Configuring Cisco Unified Computing System for the Eight-Node Oracle RAC

Configuring the Cisco UCS 6120XP Fabric Interconnect

The Cisco UCS 6120XP Fabric Interconnect is configured in a cluster pair for redundancy. It provides resiliency and access to the system configuration data in the rare case of hardware failure.

For fabric interconnects, the configuration database is replicated from the primary switch to the standby switch. All operations are transaction-based, keeping the data on both switches synchronized.

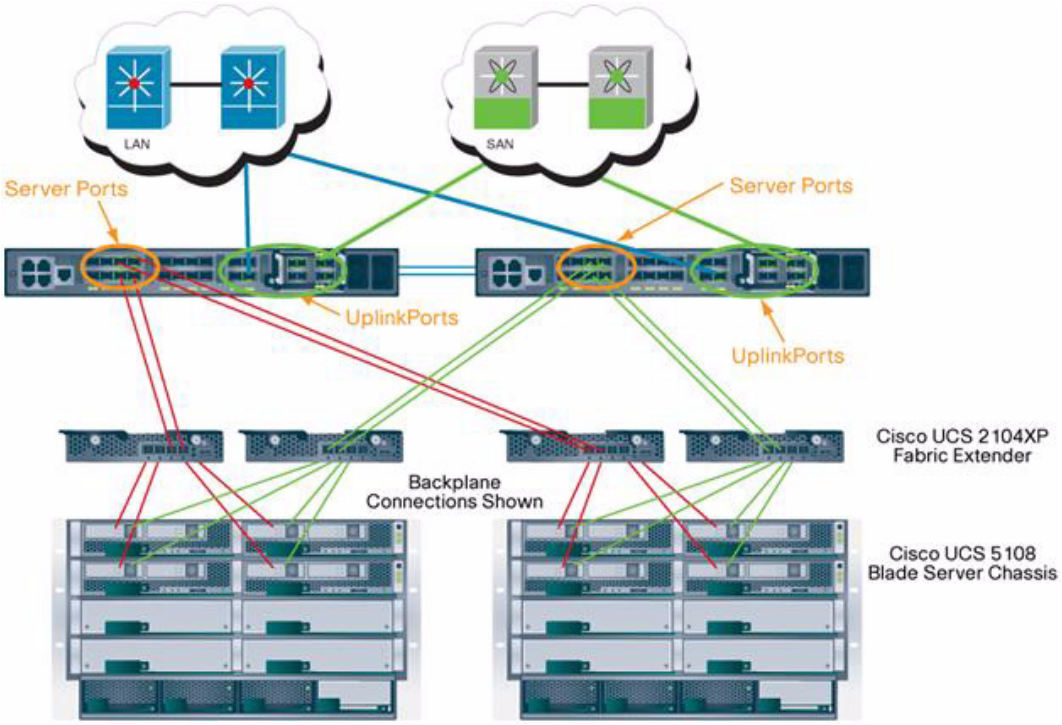
Detailed information about the fabric interconnect configuration is beyond the scope of this document. For more information, refer to the Cisco Unified Computing System documentation at http://www.cisco.com/en/US/docs/unified_computing/ucs/sw/gui/config/guide/b_GUI_Config_Guide.html.

Configuring the Server Ports

The first step is to establish connectivity between the blades and fabric interconnects. As shown in Figure 7, four public (two per chassis) links go to Fabric Interconnect "A" (ports 5 through 8). Similarly, four private links go to Fabric Interconnect B. In normal conditions, we recommend to keep all private interconnects on a single Fabric interconnect. In such case, the private traffic will stay local to that fabric interconnect and will not go to northbound network switch. In other words, all inter blade (or RAC node private) communication will be resolved locally at the fabric interconnect.

These ports should be configured as server ports as shown in Figure 7.

Figure 7 Physical Connectivity and Port Connectivity

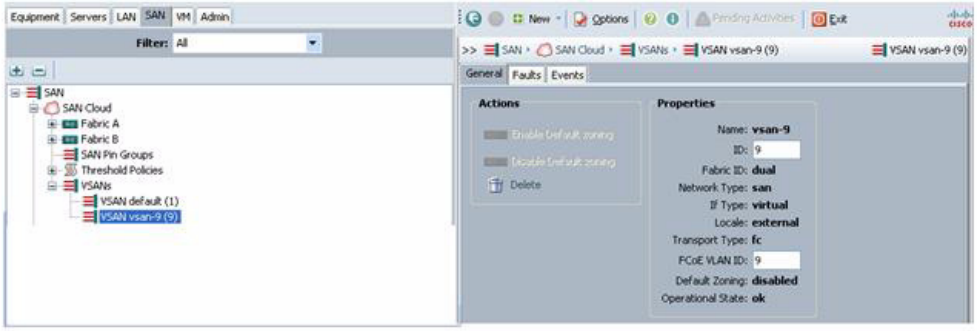


Configuring the SAN and LAN on Cisco UCS Manager

Before configuring the service profile, perform the following steps:

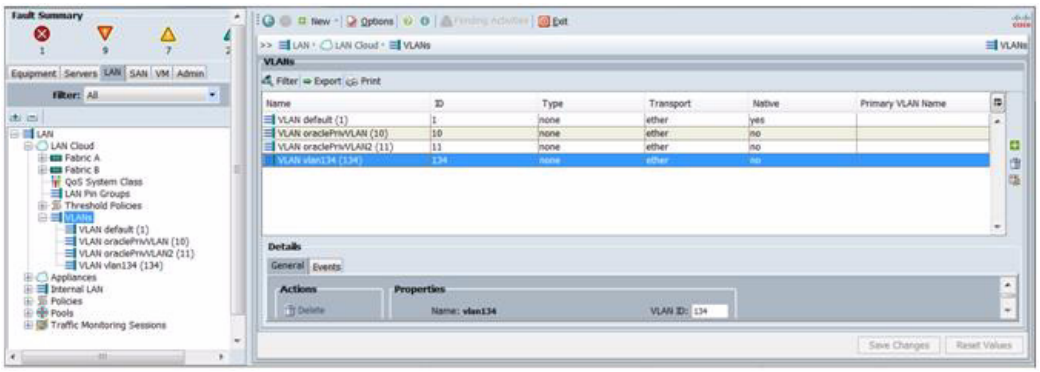
1. Configure the SAN: On the SAN tab, create and configure the VSANs to be used for database. In our setup, we used vSAN 9 for database. You should also set up pools for World Wide Names (WWNs) and World Wide Port Names (WWPNs) for assignment to the blade server virtual HBAs (vHBAs).

The following screenshot shows the vSAN-9.



2. Configure LAN: On the LAN tab, create VLANs and that will be used later for virtual NICs (vNICS) configured for private and public traffic. You can also set up MAC address pools for assignment to vNICS. For this setup, we used VLAN 134 for public interfaces and VLAN 10 for Oracle RAC private interconnect interfaces. It is also very important that you create both VLANs as global across both fabric interconnects. This way, VLAN identity is maintained across the fabric interconnects in case of failover. If you are planning to use Oracle HAIP, please create additional VLANs as required.

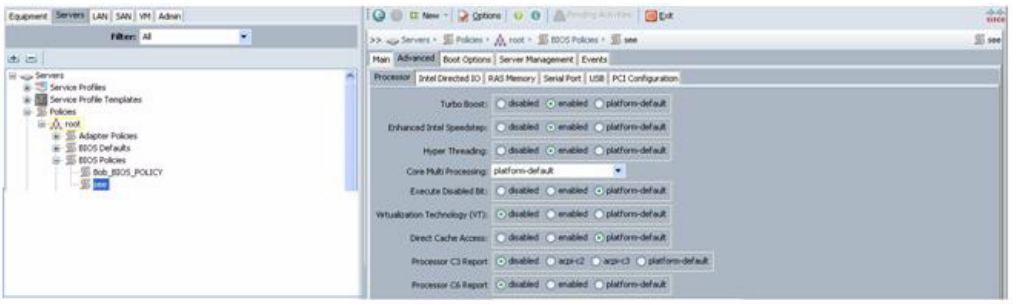
The following screenshot shows Public (134) and Private (10) VLANs for association with public (eth0) and private (eth1, eth2) network Interfaces of each Oracle RAC node. Oracle HAIP allows configuring up to four private interfaces for private interconnect traffic. Create one private VLAN for each additional HAIP interface.



Note Even though private VLAN traffic stays local within UCS domain, it is necessary to configure entries for these private VLANs in northbound network switch. This will allow the switch to route interconnect traffic appropriately in case of a partial link failure.

3. Enable Jumbo Frames (for Oracle interconnect private traffic) as shown below.

4. BIOS Settings: Go to BIOS Policies under Servers tab and create a new BIOS policy to disable C6 power state as shown below.



After these preparatory steps have been completed, you can generate a service profile template for the required hardware configuration. You can then create the service profiles for all eight nodes from the template.

Note You only need to create this BIOS policy if you are using M81KR Virtual interface Network Adapter. All other BIOS options are default.

This BIOS policy (Disable C6) is only necessary to address a rarely occurring but a known conflict between legacy interrupts and Intel Oplin chipset. More details about the issue are documented at:

http://developer.intel.com/design/network/specupdt/321040.pdf

Service Profiles: Cisco Unified Computing System Foundation Technology

Before creating a service profile template and service profiles for Oracle RAC configuration, let us look at service profiles and templates.

What is a Service Profile?

Conceptually, a service profile is an extension of the virtual machine abstraction applied to physical servers. The definition has been expanded to include elements of the environment that span the entire data center, encapsulating the server identity (LAN and SAN addressing, I/O configurations, firmware versions, boot order, network VLAN physical port, and quality-of-service [QoS] policies) in logical "service profiles" that can be dynamically created and associated with any physical server in the system within minutes rather than hours or days. The association of service profiles with physical servers is performed as a simple, single operation. It enables migration of identities between servers in the environment without requiring any physical configuration changes and facilitates rapid bare metal provisioning of replacements for failed servers. Service profiles also include operational policy information, such as information about firmware versions.

This highly dynamic environment can be adapted to meet rapidly changing needs in today's data centers with just-in time deployment of new computing resources and reliable movement of traditional and virtual workloads. Data center administrators can now focus on addressing business policies and data access on the basis of application and service requirements, rather than physical server connectivity and configurations. In addition, using service profiles, Cisco® UCS Manager provides logical grouping capabilities for both physical servers and service profiles and their associated templates. This pooling or grouping, combined with fine-grained role-based access, allows businesses to treat a farm of compute blades as a flexible resource pool that can be reallocated in real time to meet their changing needs, while maintaining any organizational overlay on the environment that they want. Figure 8 shows the major elements of a service profile.

Figure 8 Service Profile

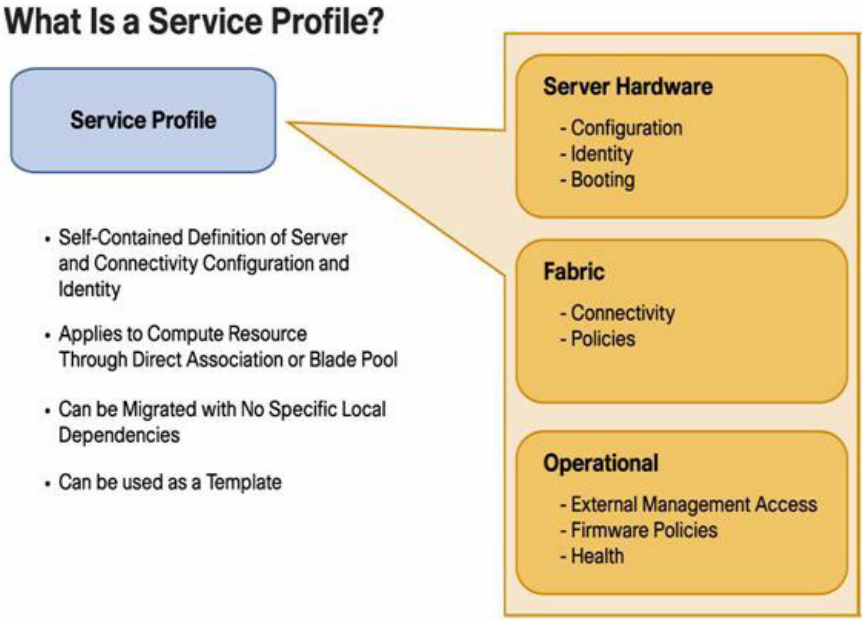
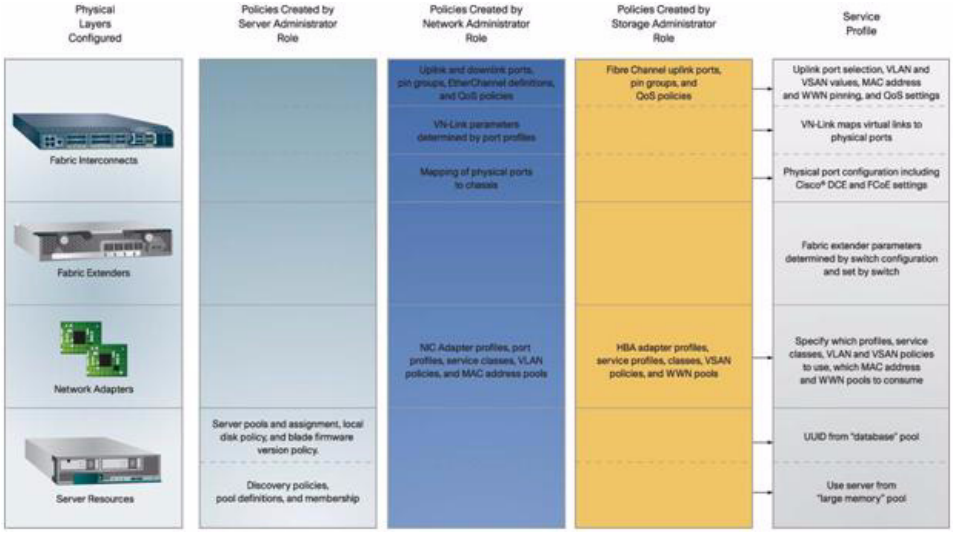


Figure 9 Service Profile Elements



In summary, service profiles represent all the attributes of a logical server in Cisco UCS data model. These attributes have been abstracted from the underlying attributes of the physical hardware and physical connectivity. Using logical servers that are disassociated from the physical hardware removes many limiting constraints around how servers are provisioned. Using logical servers also makes it easy to repurpose physical servers for different applications and services.

Understanding Service Profile Template

A lot of time can be lost between the point when a physical server is in place and when that server begins hosting applications and meeting business needs. Much of this lost time is due to delays in cabling, connecting, configuring, and preparing the data center infrastructure for a new physical server. In

addition, provisioning a physical server requires a large amount of manual work that must be performed individually on each server. In contrast, the Cisco UCS Manager uses service profile templates to significantly simplify logical (virtual) server provisioning and activation. The templates also allow standard configurations to be applied to multiple logical servers automatically, which reduces provisioning time to just a few minutes.

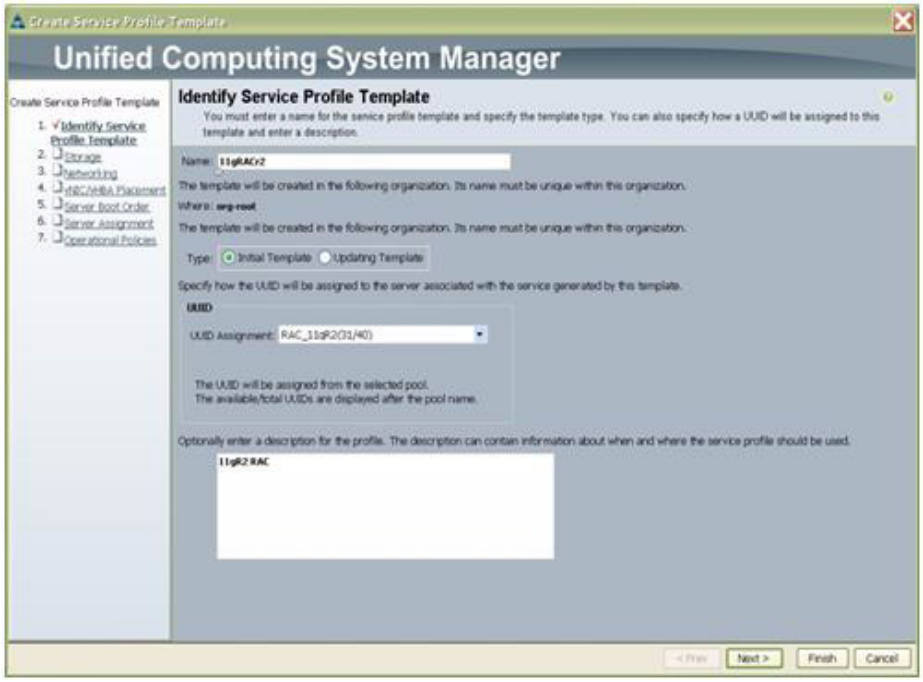
Logical server profiles can be created individually or as a template. Creating a service profile template allows rapid server instantiation and provisioning of multiple servers. The Cisco UCS data model (e.g., pools, policies, and isolation security methods) also creates higher-level abstractions such as virtual network interface cards (VNICs) and virtual host bus adapters (VHBAs). Ultimately, these service profiles are independent of the underlying physical hardware. One important aspect of the Cisco UCS data model is that it is highly referential. This means you can easily reuse and refer to previously define objects and elements in a profile without having to repeatedly redefine their common attributes and properties.

Creating the Service Profile Template

Follow these steps to create the service profile template:

1. From the Service Profile Templates screen:
   a. Click the Servers tab.
   b. Select **Service Profile Template** from the Filter.

2. From the Identify Service Profile Template screen:
   a. In the Name field, enter the template name (such as 11gRACr2).
   b. For the template type, select **Initial Template**.
   c. Click **Next**.



3. From the Storage screen (to create vHBAs for SAN storage):
   a. In the How would you like to configure SAN storage? Options, select **Expert**.
   b. Click **Add** to add vHBA.
   c. Click **Next**.

4. From the Create vHBA screen:
   a. In the Name field, enter **vHBA1**.

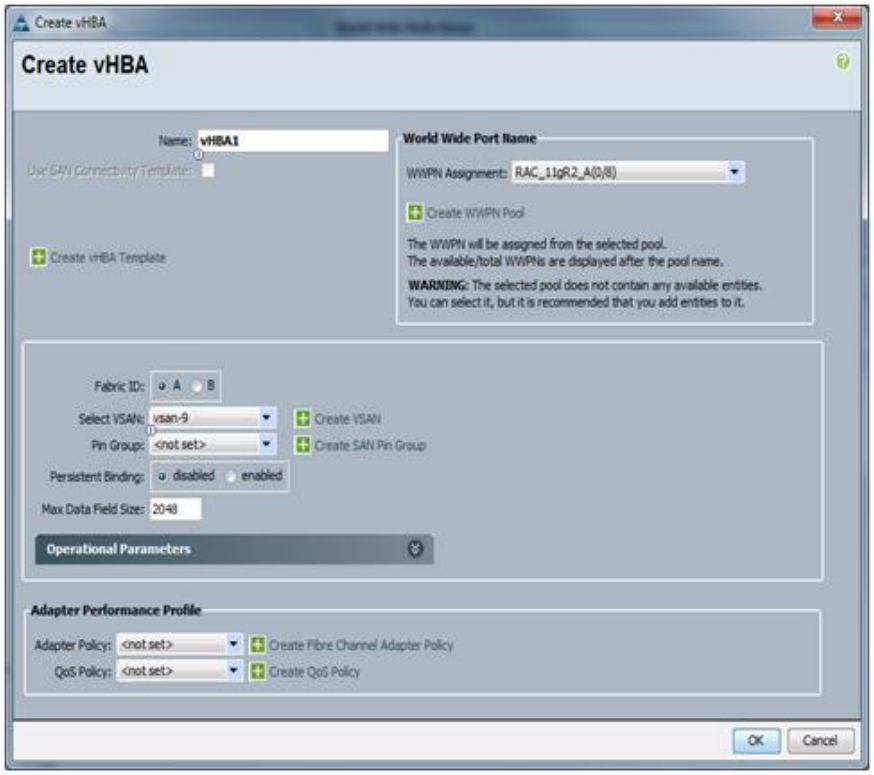
b. In the WWPN Assignment drop-down list, select the appropriate WWPN pool.

c. In the Select VSAN drop-down list, choose **vsan-9**. In this setup, we have used vsan-9 for both HBAs. You may need to make a different selection depending on what is appropriate for your configuration per your infrastructure.

d. If you have created SAN pin groups for pinning Fiber Channel traffic to a specific Fiber Channel port, specify appropriate pin groups, using the Pin Group drop-down list.

e. If you have created Adapter Policy and QoS Policy you may assign it according to your settings otherwise leave it as <not set>. In our setup, we used default Adapter Policy.

The screenshot below shows the configuration for vHBA1 assigned to Fabric Interconnect A.



f. Click **OK**.

5. Create the second HBA from the storage screen from the Create vHBA screen:

a. In the Name field, enter **vHBA2**.

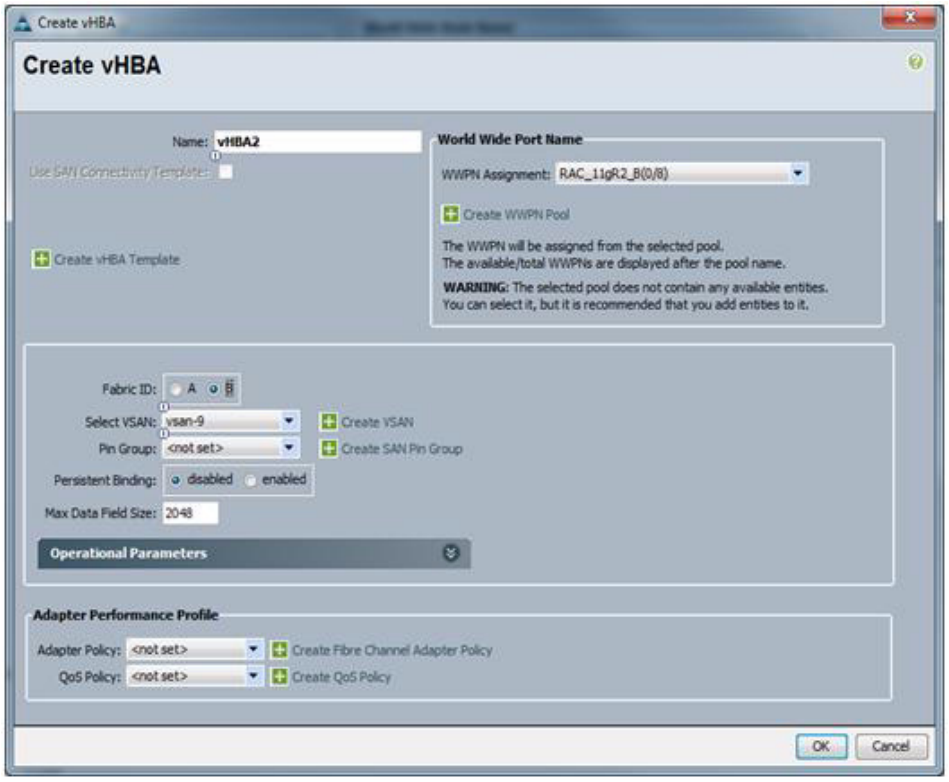
b. In the WWPN Assignment drop-down list, select the appropriate WWPN pool.

c. In the Select VSAN drop-down list, choose **vsan-9**. In this setup, we have used vsan-9 for both HBAs. You may need to make a different selection depending on what is appropriate for your configuration per your infrastructure.

d. If you have created SAN pin groups for pinning Fiber Channel traffic to a specific Fiber Channel port, specify appropriate pin groups, using the Pin Group drop-down list.

e. If you have created Adapter Policy and QoS Policy you may assign it according to your settings otherwise leave it as <not set>. In our setup, we used default Adapter Policy.

The screenshot below shows the configuration for vHBA2 assigned to Fabric Interconnect B.



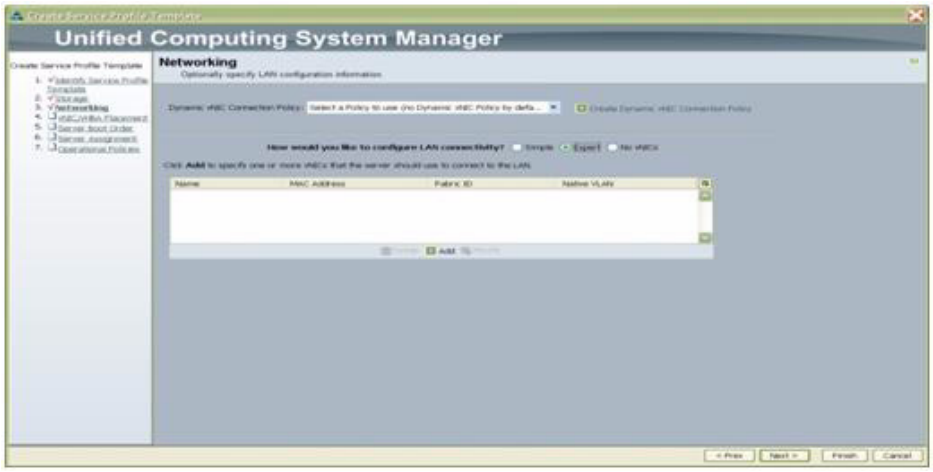
f. After creation of vHBA(s) you will be brought back to the Storage screen, click **Finish**.

The following screenshot shows that two vHBAs have been created and this completes the SAN configuration.

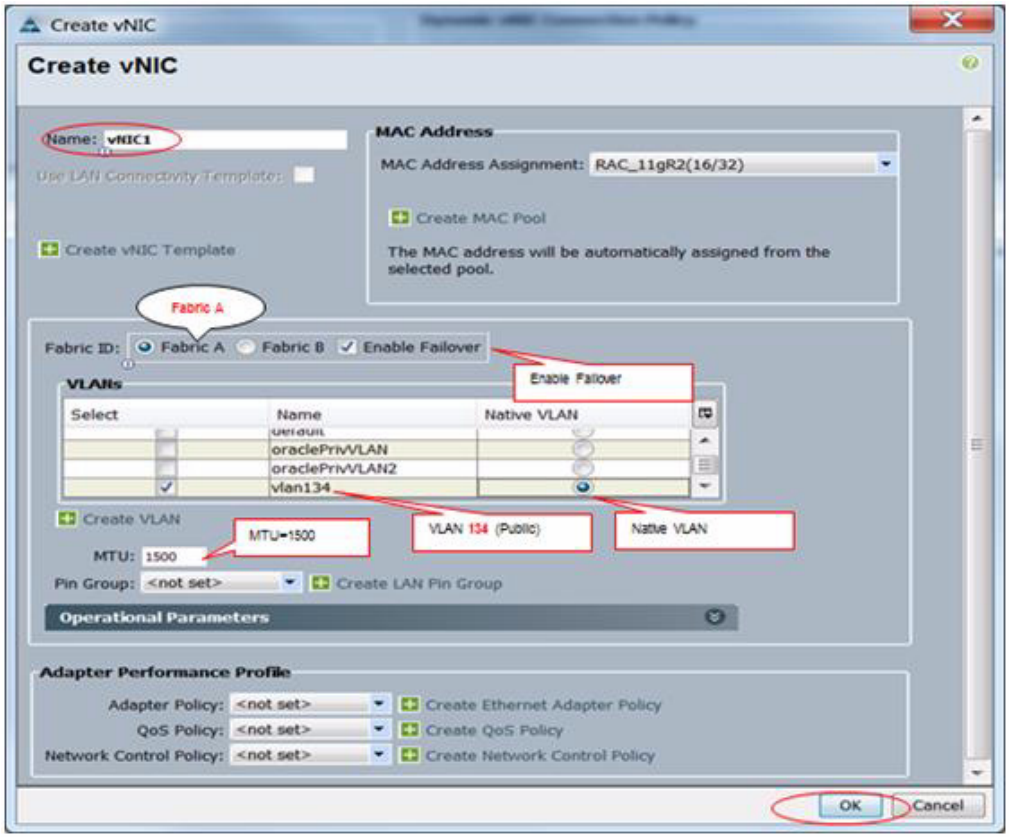
Creating and Associating the vNICs With VLANs

Follow these steps to create the vNICs and then associate them with the appropriate VLANs:

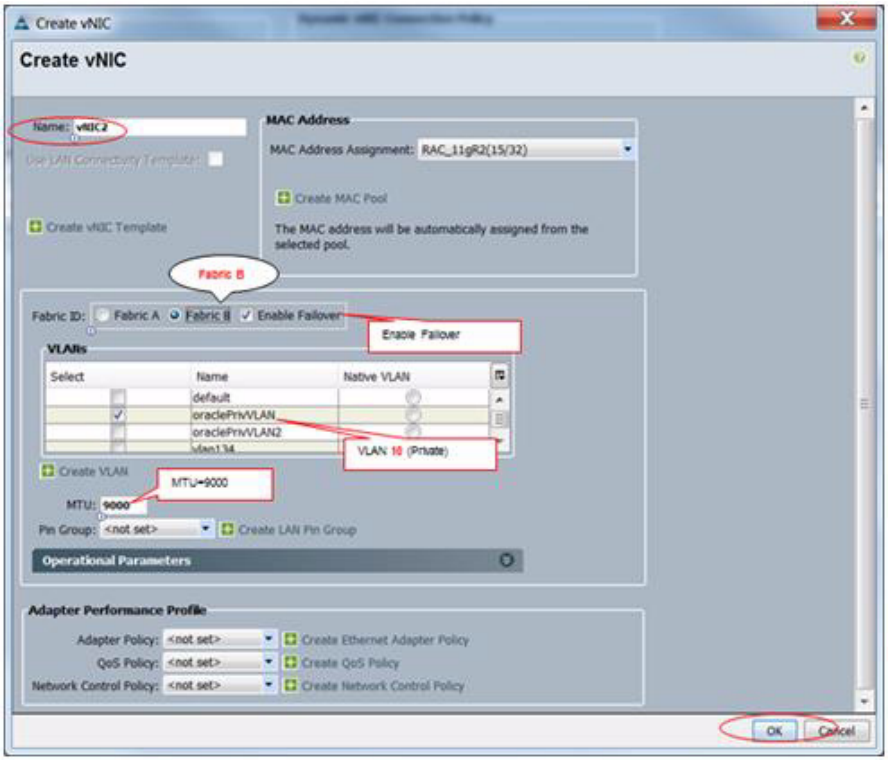
1. From the Networking screen:
   a. In the How would you like to configure LAN connectivity? options, select **Expert**.
   b. Click **Add**.



2. From the Create vNICs Networking screen:
   a. In the Name field, enter **vNIC1**.
   b. For MAC Address Assignment, select the appropriate pool.
   c. For the Fabric ID options, select **Fabric A** and Enable Failover.
   d. For Select VLAN, select **VLAN 134** (Public VLAN).
   e. Make sure that Native VLAN is selected.
   f. Select appropriate MTU values for your vNIC.
   g. If you have created Adapter Policy and QoS Policy you may assign it according to your settings otherwise leave it as <not set>.
   h. Click **OK**.
   i. VNIC1 is now assigned to Fabric A and VLAN 134.



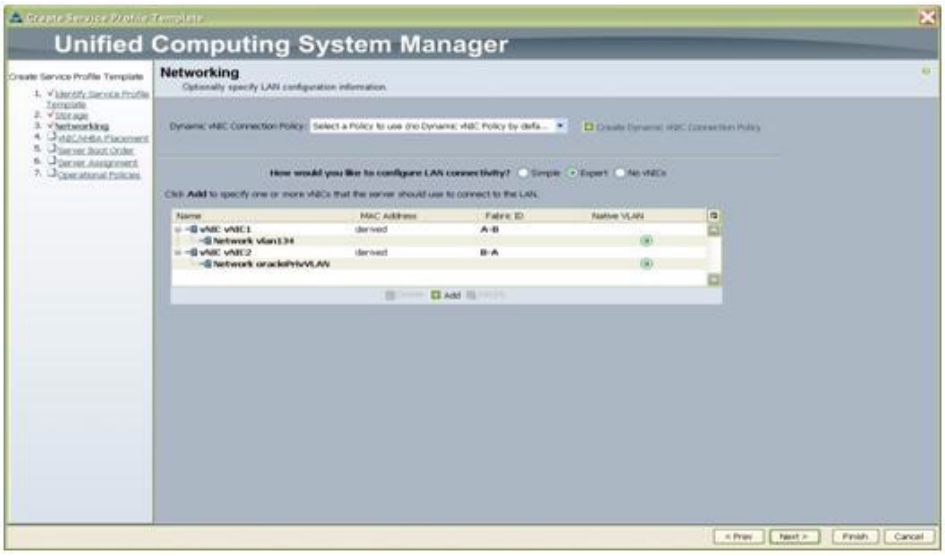
3. From the Create vNICs Networking screen:
   a. In the Name field, enter **vNIC2**.
   b. For the Fabric ID options, select **Fabric B** and Enable Failover.
   c. For Select VLAN, select **oraclePrivVLAN** (Private VLAN 10).
   d. For Private interconnect, we recommend to use jumbo frames (MTU=9000).
   e. If you have created Adapter Policy and QoS Policy you may assign it according to your settings otherwise leave it as <not set>.



Note Oracle Grid Infrastructure can activate a maximum of four private network adapters for availability and bandwidth requirements. Please note that Grid infrastructure can use only four network adapters even if more are defined. If you want to configure HAIP for GRID Infrastructure, you will need to create additional vNICs. We strongly recommend to use separate vLANs for each private vNIC. For example, create a vNIC3 using VLAN 11 (Private VLAN) as mentioned in the above step for vNIC2 creation.

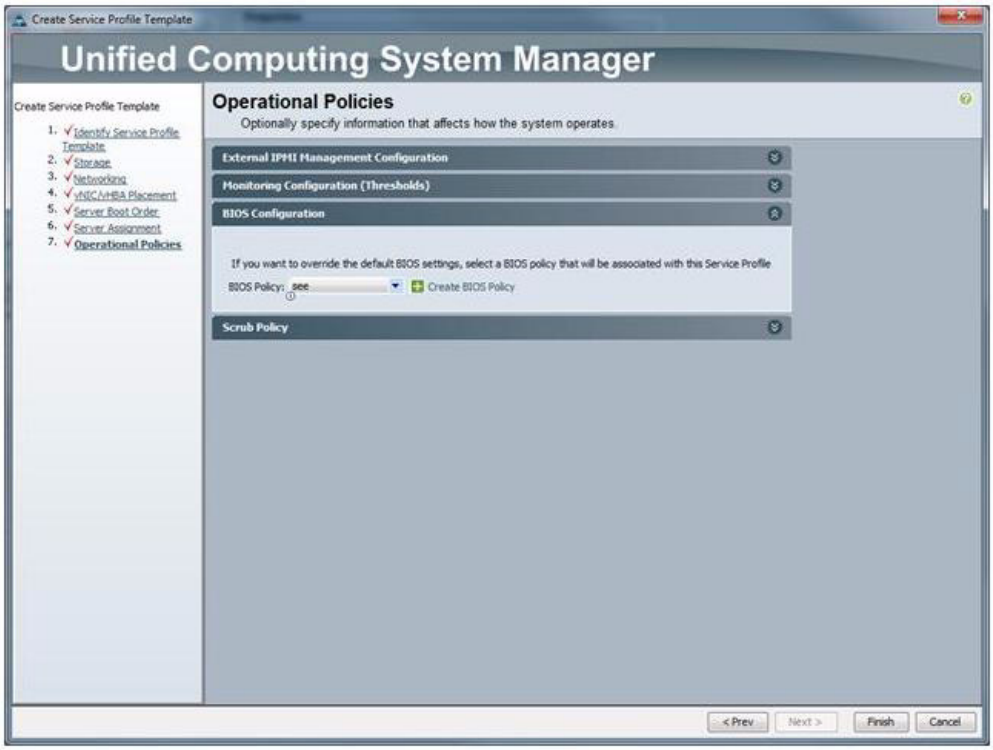
In our testing, we observed that a single Cisco UCS 10GE private vNIC configured with failover did not require HAIP configuration from bandwidth and availability perspective. As a general best practice, it is a good idea to localize all private interconnect traffic to single fabric interconnect. For more information on Oracle HAIP, please refer to Oracle metalink note 1210883.1

4. The Networking screen shows the newly created vNIC(s). Click **Next**.



The only additional policy we used in this setup is "BIOS configuration". After the network configuration, you can skip "vNIC/vHBA placement", "Server Boot Order", "Server Assignment" and directly select "Operational Policies".

5. From the Operational Policies screen:
   a. In the BIOS Policy field, select the policy named **see**.
   b. Click **Finish**.

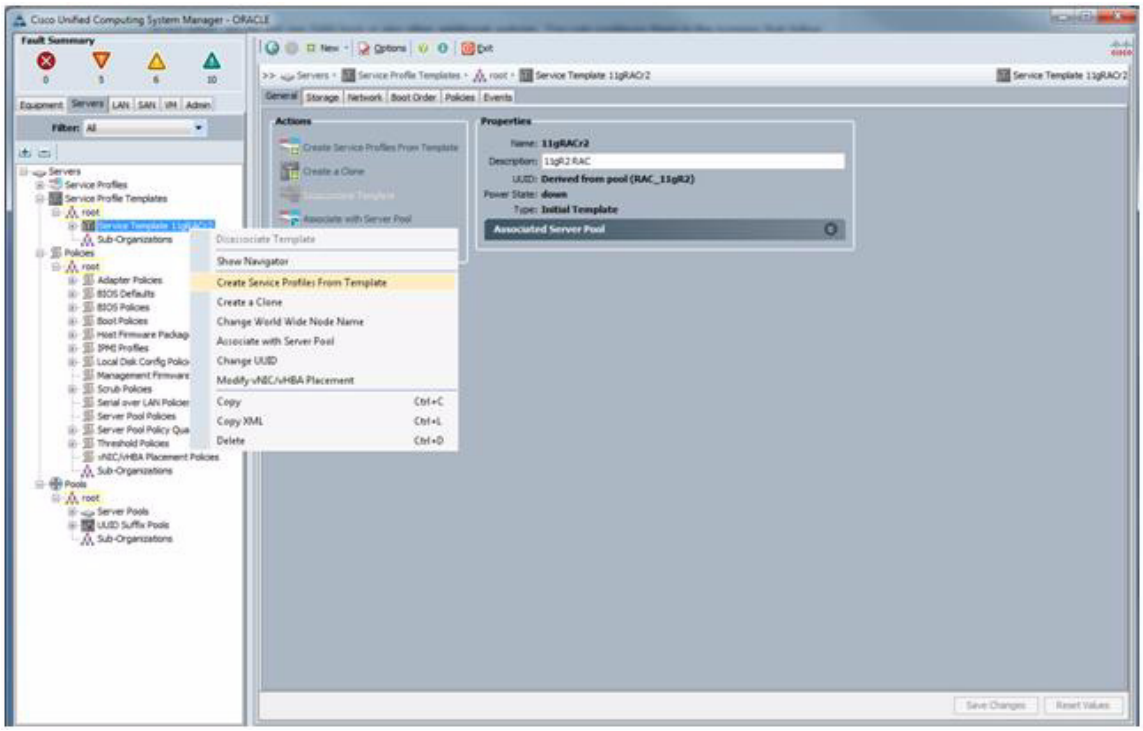


This completes the service profile template creation.

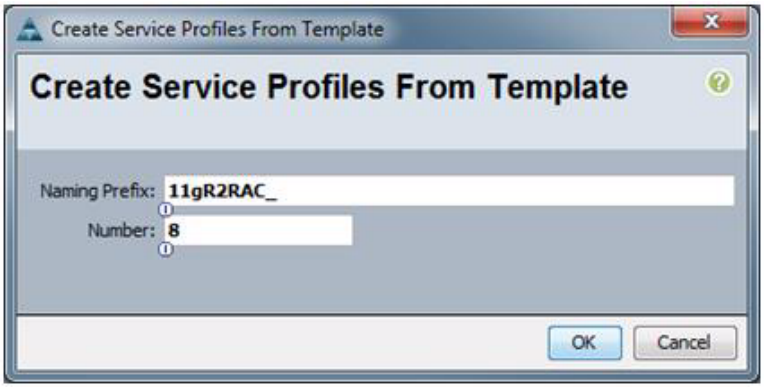
Creating and Associating Service Profiles With Blade Servers

Follow these steps to create eight service profiles and then associate them with individual blade servers:

1. From the Cisco Unified Computing System Manager screen:
   a. Right-click **Service Template 11gRACr2**.
   b. Select **Create Service Profiles From Template**.



   c. In the Naming Prefix field, enter **11gR2RAC_**.
   d. In the Number field, enter **8**.

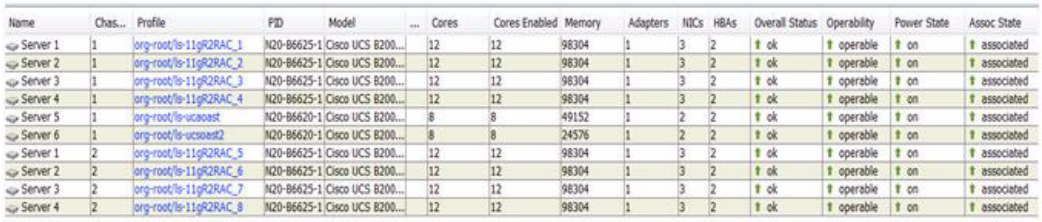


   e. Click **OK**.

This step creates service profiles for all eight blade servers. When the service profiles are created, they will select a unique MAC address, WWN, and WWPN values from the resource pools created earlier.

Now you can associate the profiles with the appropriate blade servers in the chassis, as shown below.

Check the Server Assignments with the newly created Service Profiles from template “11gR2RAC_”



Name	Chas...	Profile	PID	Model	...	Cores	Cores Enabled	Memory	Adapters	NICs	HBAs	Overall Status	Operability	Power State	Assoc State
Server 1	1	org-root/ls-11gR2RAC_1	N20-B6625-1	Cisco UCS B200...		12	12	98304	1	3	2	ok	operable	on	associated
Server 2	1	org-root/ls-11gR2RAC_2	N20-B6625-1	Cisco UCS B200...		12	12	98304	1	3	2	ok	operable	on	associated
Server 3	1	org-root/ls-11gR2RAC_3	N20-B6625-1	Cisco UCS B200...		12	12	98304	1	3	2	ok	operable	on	associated
Server 4	1	org-root/ls-11gR2RAC_4	N20-B6625-1	Cisco UCS B200...		12	12	98304	1	3	2	ok	operable	on	associated
Server 5	1	org-root/ls-ucaoast	N20-B6620-1	Cisco UCS B200...		8	8	49152	1	2	2	ok	operable	on	associated
Server 6	1	org-root/ls-ucsoast2	N20-B6620-1	Cisco UCS B200...		8	8	24576	1	2	2	ok	operable	on	associated
Server 1	2	org-root/ls-11gR2RAC_5	N20-B6625-1	Cisco UCS B200...		12	12	98304	1	3	2	ok	operable	on	associated
Server 2	2	org-root/ls-11gR2RAC_6	N20-B6625-1	Cisco UCS B200...		12	12	98304	1	3	2	ok	operable	on	associated
Server 3	2	org-root/ls-11gR2RAC_7	N20-B6625-1	Cisco UCS B200...		12	12	98304	1	3	2	ok	operable	on	associated
Server 4	2	org-root/ls-11gR2RAC_8	N20-B6625-1	Cisco UCS B200...		12	12	98304	1	3	2	ok	operable	on	associated

Configuring the SAN and Zoning on Nexus 5548UP Switch

After the service profiles are associated with servers, you need to configure Nexus 5548 unified port switch for SAN and establish the connectivity between fabric interconnect and storage.

In Appendix E—Sample Nexus 5548UP Configuration, we have shown some commands for this SAN configuration but it is beyond the scope of this document to provide whole configuration. For details about additional configuration and best practices, please refer to N5548UP documentation.

Table 6 below lists the zones and their associated members that are used in the testing and discussed in this document.

Table 6 Zones for Oracle RAC Setup

Zone name	Host (HBA)	Storage visible from Zone
RAC1_hba1	RAC1_HBA1	SPA0, SPA4, SPB2, SPB6
RAC2_hba1	RAC2_HBA1	
RAC3_hba1	RAC3_HBA1	
RAC4_hba1	RAC4_HBA1	
RAC5_hba1	RAC5_HBA1	
RAC6_hba1	RAC6_HBA1	
RAC7_hba1	RAC7_HBA1	
RAC8_hba1	RAC8_HBA1	
RAC1_hba2	RAC1_HBA2	SPA2, SPA6, SPB0, SPB4
RAC2_hba2	RAC2_HBA2	
RAC3_hba2	RAC3_HBA2	
RAC4_hba2	RAC4_HBA2	
RAC5_hba2	RAC5_HBA2	
RAC6_hba2	RAC6_HBA2	
RAC7_hba2	RAC7_HBA2	
RAC8_hba2	RAC8_HBA2	

After you complete the zoning, you are ready to configure storage.

Setting Up EMC VNX Storage

This document provides a general overview of the storage configuration for the database layout. However, it is beyond the scope of this document to provide details about host connectivity and logical unit number (LUNs) in RAID configuration. For more information about Oracle database best practices for deployments with EMC VNX storage, refer to http://www.emc.com/oracle.

The following are some generic recommendations for EMC VNX storage configuration with mixed drives.

- Turn off the read and write caches for flash drive-based LUNs. In most situations, it is better to turn off both the read and write caches on all the LUNs that reside on flash drives, for the following reasons:
  - The flash drives are extremely fast: When the read cache is enabled for the LUNs residing on them, the read cache lookup for each read request adds more overhead compared to SAS drives. This scenario occurs in an application profile that is not expected to get many read cache hits at any rate. It is generally much faster to directly read the block from the flash drives.
  - Typically, the storage array is also shared by several other applications along with the database. In some situations, the write cache may become fully saturated, placing the flash drives in a force-flush situation. This adds unnecessary latency. This typically occurs particularly when storage deploys mixed drives and consists of slower Near Line SAS drives. Therefore, it is better in these situations to write the block directly to the flash drives than to the write cache of the storage system.
  - Distribute database files for flash drives. Refer to the table below for recommendations about distributing database files based on the type of workload.

Table 7 Distribution of Data Files Based on the Type of Workload

Flash drive friendly database files	Not as cost-effective on flash drives
User Tablespace Based Data Files **Random Reads** B-Tree leaf access ROWID look up into Table Access to out-of-line LOB Access to overflowed row Index scan over Unclustered Table Compression: Increases I/O intensity (IOPS/GB) **Serial Reads** **Random Writes** Row update by PK Index maintenance Reduce checkpoint interval **Temp Tablespace Files - Sort Areas and Intermediate Tables** Sequentially read and written *but* I/O done in 1 MB units: not enough to amortize seeks. Lower Latency: Get In, Get Out	**Redo Log Files** Sequential I/O Read and write *and* commit latency already handled by cache in storage controller **Undo Tablespace** Sequential writes and randomly read by Oracle Flashback. Generally, reads are for recently written data that is likely to be in the buffer cache **Large Table Scans (if single stream)**

The configuration described here employs most of EMC's best practices and recommendations for LUN distribution in the database. It also adopts the layout for a mixed storage environment consisting of SAS disks and Flash drives.

For more information about Oracle database best practices for flash-drive-based EMC VNX storage, refer to the document “Leveraging EMC VNX5500 with Enterprise Flash Drives for Oracle Database Deployments” at http://www.emc.com/products/series/vnx-series.htm .

Applying Patches, Environment, and OS Settings

The following is the summary for the 8-node Oracle RAC setup.

Oracle 11gR2 - Cisco UCS Certification test bed setup:

1. 8 Node RAC Cluster on CISCO UCS B200-M2 Half width blades
2. EMC VNX5500 Storage
3. FCoE (Carries 10GigE Ethernet and Fiber Channel Traffic)
4. CISCO M81KR Virtual interface Card.
5. Oracle Linux 5.7 , 64-bit
6. 11gR2 Grid Infrastructure and Oracle Real Application Clusters
7. 3 ASM Disk Groups
   a. OCRVD: Contains Voting and OCR disks.

Note It is not mandatory to create a separate diskgroup for OCR and Voting disks.

   b. DATADG: Contains All Data and Temp tablespaces using Flash drives.
   c. REDODG: Contains All Redo Logs

After completing the configuration of the Cisco Unified Computing System, the SAN, and storage, you are ready to install the OS.

Install OS and Setup the Environment

Note Due to the nature of failure tests in this certification, we used local disk for OS boot. However, we highly recommend SAN boot configuration in order to fully realize the benefits of stateless computing.

Follow these steps to install the OS and other required packages to enable the RAC environment settings:

1. Install 64-bit Oracle Linux 5.7, on all eight nodes (during OS install, select packages shown below).
2. Install the following RPM(s) by choosing from the options:
   - Oracle Validated Configuration RPM package

   The Oracle Validated Configuration RPM sets and verifies system parameters based on recommendations from the Oracle Validated Configurations program, and installs any additional packages needed for Oracle Clusterware and database installation. It also updates sysctl.conf settings, system startup parameters, user limits, and driver parameters to Oracle recommended values.

   - Linux ASMLib RPMs

   Oracle recommends that you use ASM for database file management, and install the Linux ASMLIB RPMs to simplify administration. ASMLib 2.0 is delivered as a set of three Linux RPM packages:

   - oracleasmlib-2.0 - the ASM libraries
   - oracleasm-support-2.0 - utilities needed to administer ASMLib
   - oracleasm - a kernel module for the ASM library

Please refer to Oracle documentation to identify the kernel and appropriate RPM packages.

3. Install the EMC hostagent RPM package on all the nodes.

   The EMC host agent is server-based software that communicates with client applications, such as the Navisphere Command Line Interface (CLI) and Manager. A host agent automatically registers hosts and HBAs and also provides drive mapping information for the UI and CLI.

4. Install the EMC PowerPath RPM package on all nodes.

   EMC PowerPath software is a server-resident, performance and application availability enhancing software solution. PowerPath combines multiple path I/O capabilities, automatic load balancing, and path failover functions into one integrated package. Here are the basic commands to install and configure PowerPath. Please refer to PowerPath documentation for detailed information.

   Install the EMC PowerPath rpm

   ```
   rpm -ivh EMCpower.LINUX-5.5.0.00.00-275.3.1.00.00-111.rhel5.x86_64.rpm
   ```

   Install the license key

   ```
   emcpreg -add xxxx-xxxx-xxxx
   ```

   Confirm license key installation

   ```
   powermt check_registration
   ```

   Save the License

   ```
   powermt save
   ```

   Start powerpath

   ```
   /etc/init.d/PowerPath start
   ```

   Checking if powerpath has been installed

   ```
   powermt display dev=all
   ```

5. With multipathing software installed, we are ready to create disk partitions for Oracle GRID Clusterware disks. The command below shows for one disk. Repeat the same sequence to create partitions on all Grid control disks.

   ```
   [root@orarac1 install]# fdisk /dev/emcpowera
   Device contains neither a valid DOS partition table, nor Sun, SGI or OSF
   disklabel
   Building a new DOS disklabel. Changes will remain in memory only,
   until you decide to write them. After that, of course, the previous
   content won't be recoverable.
   Warning: invalid flag 0x0000 of partition table 4 will be corrected by
   w(rite)
   Command (m for help): n
   Disk /dev/emcpowera: 1073 MB, 1073741824 bytes
   34 heads, 61 sectors/track, 1011 cylinders
   Units = cylinders of 2074 * 512 = 1061888 bytes
   Device Boot      Start         End      Blocks   Id System
   Command (m for help): n
   Command action
      e   extended
      p   primary partition (1-4)
   p
   ```

```
Partition number (1-4): 1
First cylinder (1-1011, default 1):
Using default value 1
Last cylinder or +size or +sizeM or +sizeK (1-1011, default 1011):
Using default value 1011
Command (m for help): w
The partition table has been altered!
Calling ioctl () to re-read partition table.
Syncing disks.
```

Now the newly partitioned disk is available to Oracle "oracleasm" utility for further ASM configuration.

```
[root@orarac1 ~]# ls -l /dev/emcpowera1
brw-r----- 1 root disk 120, 1 Aug 20 14:07 /dev/emcpowera1
```

Create the partition for ASM-managed DATA and REDO disks at the offset of 1 MB (or 2048 sectors).

The 1MB partition header offset is recommended for VNX storage due to the fact that when doing parity RAID (e.g. RAID-5 or RAID-6), VNX uses a patented parity rotation scheme that shifts the parity data to a different drive after multiple stripes of data as opposed to every data stripe. With this parity rotation scheme, 1MB worth of data, that actually represents multiple default parity RAID stripes worth of data, would have the data pieces all stored consecutively on all but one (R5) or two (R6) drives, and with the parity for those multiple full data stripes stored completely on the same drive (R5) or drives (R6). As long as application is storing data in 1MB units, aligned on 1MB offset within the LUNs, all 1MB data chunk accesses will only involve reading data from the data spindles, and the drives holding the parity will not be involved. If the reads are not aligned on 1MB LUN boundaries, the opportunity would then exist that the same 1MB of data would have to be satisfied from reading uneven amount of data from all the drives in that RAID group.

Use the following setup to create partitions for ASM-managed data disks, for example:

```
[root@orarac1 install]# fdisk /dev/emcpowerf
n             new partition
p             primary partition
1             partition 1
(CR)          start from beginning of device or LUN
(CR)          use all the available sectors
x             go into EXPERT mode
b             adjust partition header data begin offset
1             for partition 1
2048          to sector 2048 from beginning of LUN, or 1MB
w             commit changes
```

Note The above step will create partition /dev/emcpowerf1 which can be used for DATA_1 for Oracle ASM. Repeat this procedure to create partitions for the rest of the DATA, REDO disks as well.

After all the partitions are created, you are ready to configure the ASM kernel module and stamp the ASM disks.

Note These are sample commands and the purpose is to highlight logical steps for the Oracle configuration. Please refer to Oracle documentation for step-by-step install and detailed configuration.

- With the ASMLib RPMs installed, configure the ASM kernel module on all the nodes.

```
[root@orarac1 tmp]# /usr/sbin/oracleasm configure -i
Configuring the Oracle ASM library driver.

This will configure the on-boot properties of the Oracle ASM library
driver.  The following questions will determine whether the driver is
loaded on boot and what permissions it will have.  The current values
will be shown in brackets ('[]').  Hitting <ENTER> without typing an
answer will keep that current value.  Ctrl-C will abort.

Default user to own the driver interface []: grid

Default group to own the driver interface []: asmadmin

Start Oracle ASM library driver on boot (y/n) [n]: y

Scan for Oracle ASM disks on boot (y/n) [y]: y

Writing Oracle ASM library driver configuration: done
```

- ASM cannot handle seeing the same disk twice and will generate an error if a disk appears multiple times. In its default configuration, ASMLib will choose the first path for the disk it finds during the scan. Normally, this is the first path as reported by the operating system and that path could be the multipath, or it could be one of the single paths. For multipath based configurations, we want ASMLib to always use the multipath disk. There is no way, however, for ASMLib to know what a multipath looks like. So, the disk scan order and type of disks to scan parameters must be configured. ASMLib allows two modifications to the disk scan order. First, it allows exclusion of certain disks. In other words, ASMLib will ignore those disks completely. Second, one can specify the disks that are to be scanned first. Disks in this list are scanned before the rest of the disks in the system. To set the scan order and disk exclusion

  Please modify the /etc/sysconfig/oracleasm file as shown below.

```
# ORACLEASM_SCANORDER: Matching patterns to order disk scanning
ORACLEASM_SCANORDER="emcpower"
# ORACLEASM_SCANEXCLUDE: Matching patterns to exclude disks from scan
ORACLEASM_SCANEXCLUDE="sd"
```

- Now you are ready to stamp (or label) the partitions created earlier as ASM disks. This is required for only on one node.

```
[root@orarac1 tmp]# /usr/sbin/oracleasm createdisk CLUSTERGRID_VOL1
/dev/emcpowerq1
Writing disk header: done
```

```
Instantiating disk: done
[root@orarac1 tmp]# /usr/sbin/oracleasm createdisk CLUSTERGRID_VOL2
/dev/emcpowerr1
Writing disk header: done
Instantiating disk: done
[root@orarac1 tmp]# /usr/sbin/oracleasm createdisk CLUSTERGRID_VOL3
/dev/emcpowers1
Writing disk header: done
Instantiating disk: done
[root@orarac1 tmp]# /usr/sbin/oracleasm createdisk CLUSTERGRID_VOL4
/dev/emcpowert1
Writing disk header: done
Instantiating disk: done
[root@orarac1 tmp]# /usr/sbin/oracleasm createdisk CLUSTERGRID_VOL5
/dev/emcpoweru1
Writing disk header: done
Instantiating disk: done
```

These commands are only shown for Grid Control disks. Please create all remaining disks (Data, logs) as shown in Table 8.

Table 8 LUNS for 8-Node Oracle RAC Cluster

Disk Type	SP Path Details of EMC VNX	Mapped PowerPath Devices	Oracle ASM DG Label	Purpose
FC	Owner:default=SP A, failover mode: 1, policy=CLAROpt	/dev/emcpowera1	CLUSTERGRID_VOL1	ASM (Voting Disk, OCR)
FC	Owner:default=SP B, failover mode: 1, policy=CLAROpt	/dev/emcpowerb1	CLUSTERGRID_VOL2	ASM (Voting Disk, OCR)
FC	Owner:default=SP B, failover mode: 1, policy=CLAROpt	/dev/emcpowerc1	CLUSTERGRID_VOL3	ASM (Voting Disk, OCR)
FC	Owner:default=SP A, failover mode: 1, policy=CLAROpt	/dev/emcpowerd1	CLUSTERGRID_VOL4	ASM (Voting Disk, OCR)
FC	Owner:default=SP B, failover mode: 1, policy=CLAROpt	/dev/emcpowere1	CLUSTERGRID_VOL5	ASM (Voting Disk, OCR)
FD	Owner:default=SP A, failover mode: 1, policy=CLAROpt	/dev/emcpowerf1	DATA_1	OASTOLTP, OASTDE OASTDSS, STRAC
FD	Owner:default=SP B, failover mode: 1, policy=CLAROpt	/dev/emcpowerg1	DATA_2	OASTOLTP, OASTDE OASTDSS, STRAC
FD	Owner:default=SP A, failover mode: 1, policy=CLAROpt	/dev/emcpowerh1	DATA_3	OASTOLTP, OASTDE OASTDSS, STRAC
FD	Owner:default=SP B, failover mode: 1, policy=CLAROpt	/dev/emcpoweri1	DATA_4	OASTOLTP, OASTDE OASTDSS, STRAC
FD	Owner:default=SP B, failover mode: 1, policy=CLAROpt	/dev/emcpowerj1	DATA_5	OASTOLTP, OASTDE OASTDSS, STRAC
FD	Owner:default=SP A, failover mode: 1, policy=CLAROpt	/dev/emcpowerk1	DATA_6	OASTOLTP, OASTDE OASTDSS, STRAC
FD	Owner:default=SP B, failover mode: 1, policy=CLAROpt	/dev/emcpowerl1	DATA_7	OASTOLTP, OASTDE OASTDSS, STRAC
FD	Owner:default=SP A, failover mode: 1, policy=CLAROpt	/dev/emcpowerm1	DATA_8	OASTOLTP, OASTDE OASTDSS, STRAC
FD	Owner:default=SP A, failover mode: 1, policy=CLAROpt	/dev/emcpowern1	DATA_9	OASTOLTP, OASTDE OASTDSS, STRAC
FD	Owner:default=SP B, failover mode: 1, policy=CLAROpt	/dev/emcpowero1	DATA_10	OASTOLTP, OASTDE OASTDSS, STRAC
FD	Owner:default=SP A, failover mode: 1, policy=CLAROpt	/dev/emcpowerp1	DATA_11	OASTOLTP, OASTDE OASTDSS, STRAC
FD	Owner:default=SP B, failover mode: 1, policy=CLAROpt	/dev/emcpowerq1	DATA_12	OASTOLTP, OASTDE OASTDSS, STRAC
FC	Owner:default=SP B, failover mode: 1, policy=CLAROpt	/dev/emcpowerr1	REDO_1	OASTOLTP, OASTDE OASTDSS, STRAC
FC	Owner:default=SP A, failover mode: 1, policy=CLAROpt	/dev/emcpowers1	REDO_2	OASTOLTP, OASTDE OASTDSS, STRAC
FC	Owner:default=SP B, failover mode: 1, policy=CLAROpt	/dev/emcpowert1	REDO_3	OASTOLTP, OASTDE OASTDSS, STRAC
FC	Owner:default=SP A, failover mode: 1, policy=CLAROpt	/dev/emcpoweru1	REDO_4	OASTOLTP, OASTDE OASTDSS, STRAC

Configure the private and public NICs with the appropriate IP addresses.

Note If you have planned to configure HAIP during GRID infrastructure installation, make sure you have created two vNIC's in previous steps for private IP configuration.

Make sure you have enough IP addresses available in your DHCP server apart from Public IP, for the use of Virtual IP, SCAN IP.

Public IP, Private IP and GNS IP can be either defined in "/etc/hosts" file or can be stored in dns file with the appropriate names.

You are now ready to install the Oracle Clusterware and the database.

Oracle 11gR2 Clusterware and the Database

The following are the new features and requirements for the Oracle 11gR2 RAC:

- A new OS user called "grid" must be created for installing GRID infrastructure and associated with oinstall, asmadmin, asmdba, dba groups.
- 2 separate Oracle Homes are required for grid and oracle.
- Choice of Time Synchronization
  - Oracle Cluster Time Synchronization Service (CTSS) or
  - NTP with slewing (option -x prevents time from being adjusted backward)
- New SCAN (Single Client Access Name) Listener for Client connection uses 3 VIP(s) hence making the possibility of using a "dedicated single hostname" to access the cluster services by the clients. This is highly beneficial in the event of any changes to the cluster since the clients will not be affected and will not require any changes to the "dedicated single hostname". It can move around cluster nodes and helps to register all database instances and services. It also uses the load balance advisory to distribute the loads across cluster nodes. In addition, you can have local listeners too. Both SCAN and Local listeners are managed by GRID OS user.
- For 11gR2 RAC, requirement of DNS and GNS (Grid Naming Service) configuration helps the VIP(s) to be managed efficiently for GPnP (Grid Plug 'n' Play) and makes any cluster nodes could plug-in / plug-out at any time and without any manual intervention of VIP configuration. In a RAC cluster environment, there is a high chance of nodes being added and removed dynamically as per certain customer requirements. In this situation, the complete administration with ip address management and name resolution management is done by the cluster itself using GNS. There is no need to do any manual work in updating connection strings, configuring Virtual IP numbers etc. If you decide to use GNS, it is required that a GNS sub-domain must be created and the DNS must be configured such that each request for this sub-domain will be delegated to the GNS sub-domain for request handling. The GNS VIP address is the ip address of the server that will host the GNS. You need to make sure that one GNS VIP is available for use. Single SCAN hostname resolves to 3 VIPs and during installation, DNS resolution provides 3 VIPs which are used to create 3 SCAN-VIP / Listener pairs scattered across the cluster.
- Policy-based cluster and Server Pool capacity management allows efficient allocation of resources to all kinds of applications in the cluster. This way, the applications can be hosted on a shared infrastructure while being isolated regarding their resource consumption via policies and therefore behave as if they were deployed in dedicated system environments. These features can be summarized as following:

  Database Policy Management:

  - Allows defining resource requirements of workload.
  - Oracle Clusterware allocates and reassigns capacity based on defined policies. This enables faster resource failover and dynamic capacity assignment.
  - Enough instances are started to support workload requirements
  - Eliminates need to hard code a service to a specific instance or node.

  Server Pool:

  - Logical division of the cluster into pools of servers
  - Managed by crsctl (applications), srvctl (Oracle)
  - Defined by 3 attributes (min, max, importance) or a defined list of nodes.

  - Oracle HAIP: With Oracle Database 11gR2 (11.2.0.2 and later), you can identify multiple interfaces to use for the cluster private network, without the need of using bonding or other technologies. When you define multiple interfaces for private network, Oracle Clusterware creates from one to four (depending on the number of interfaces you define) highly available IP (HAIP) addresses, which Oracle Database and Oracle ASM instances use to ensure highly available and load balanced communications. With HAIP, by default, interconnect traffic is load balanced across all active interconnect interfaces, and corresponding HAIP address are failed over transparently to other adapters if one fails or becomes non-communicative. The Oracle software (including Oracle RAC, Oracle ASM, and Oracle ACFS, all Oracle Database 11g release 2 (11.2.0.2), or later), by default, uses these HAIP addresses for all of its traffic allowing for load balancing across the provided set of cluster interconnect interfaces. If a defined cluster interconnect interface fails or becomes non-communicative, then Oracle Clusterware transparently moves the corresponding HAIP address to one of the remaining functional interfaces.
  - If you install Oracle Clusterware using OUI, then the public interface names associated with the network adapters for each network must be the same on all nodes, and the private interface names associated with the network adaptors should be the same on all nodes. This restriction does not apply if you use cloning, either to create a new cluster, or to add nodes to an existing cluster.

    For example: With a two-node cluster, you cannot configure network adapters on node1 with eth0 as the public interface, but on node2 have eth1 as the public interface. Public interface names must be the same, so you must configure eth0 as public on both nodes. You should configure the private interfaces on the same network adapters as well. If eth1 is the private interface for node1, then eth1 should be the private interface for node2.
- The following Metalink notes provide additional details about these new features:
  - 1053147.1: Oracle 11gR2 Clusterware and Grid Home (What you need to know) details
  - 887522.1 Oracle 11gR2 Grid Infrastructure SCAN Explained.
  - 946452.1 11gR2 Grid Infrastructure GNS Explained
  - 1210883.1Grid Infrastructure Redundant Interconnect and ora.cluster_interconnect.haip

Multicast Requirements for Oracle Grid Infrastructure

With Oracle Grid Infrastructure release 2 (11.2), on each cluster member node, the Oracle mDNS daemon uses multicasting on all interfaces to communicate with other nodes in the cluster. With Oracle Grid Infrastructure release 2 patch set 1 (11.2.0.2) and later releases, multicasting is required on the private interconnect. For this reason, at a minimum, you must enable multicasting for the cluster:

- Across the broadcast domain as defined for the private interconnect
- On the IP address subnet ranges 224.0.0.251/24 and/or 230.0.1.0/24. Multicast configuration support is required only one of this subnet ranges.
- If you decide to use 230.0.1.0 for multicast, a multicast querier configuration at Layer 3 network switch is required. These details are discussed in later section.

Note You do not need to enable multicast communications across routers.

Initially for Oracle 11.2, Oracle Grid Infrastructure for a cluster mandated multicast communication on 230.0.1.0 addresses working properly between the cluster nodes. When first attempting to start an 11.2.0.2 cluster node, either during a fresh install or after an upgrade attempt, it is possible that the node will be unable to join the cluster. The Oracle Cluster Synchronization Services (OCSSD) will report a failure to join the existing cluster members. If multicasting is not enabled or upstream IGMP queries does not exist, you may see the following symptoms where as root.sh succeeds on the first node and fails on subsequent nodes with:

```
CRS-4402: The CSS daemon was started in exclusive mode but found an active
CSS daemon on node prittoprfdb1, number 1, and is terminating

An active cluster was found during exclusive startup, restarting to join the
cluster

Failed to start Oracle Clusterware stack

Failed to start Cluster Synchronisation Service in clustered mode at
/u01/app/crs/11.2.0.2/crs/install/crsconfig_lib.pm line 1016.

/u01/app/crs/11.2.0.2/perl/bin/perl -I/u01/app/crs/11.2.0.2/perl/lib
-I/u01/app/crs/11.2.0.2/crs/install
/u01/app/crs/11.2.0.2/crs/install/rootcrs.pl execution failed
```

The problem occurs because the clusterware in 11.2.0.2 uses multicast messages to initiate communications on the 230.0.1.0 address. If multicast on these addresses is not enabled for the private interfaces, then the clusterware will be unable to connect to it's peer nodes. To address this issue, Oracle has released Patch: 9974223 on top of 11.2.0.2. This patch must be applied before executing root.sh on each node in the cluster. This patch makes use of the 224.0.0.251 (port 42424) multicast network address in addition to the 230.0.1.0 (port 42424) multicast address. Multicast must be enabled on one of these two addresses to allow for Oracle Grid Infrastructure to successfully start on all cluster nodes.

Please refer to Oracle metalink note 1212703.1 for additional details.

Installing Oracle Clusterware and the Database

The next step is to install the Oracle Clusterware and database software. For more information about the Oracle RAC installation, refer to the Oracle install documentation.

1. Download the Oracle Database 11g Release 2 (11.2.0.2.0) software.
2. Install Oracle Database 11g Release 2 Grid Infrastructure.
3. Install Oracle Database 11g Release 2 Database “Software Only”; do not create the database.

Database and Workload Configuration

After the database software is installed, We Installed OAST kit and configured it. Created four different databases (OASTOLTP, OASTDB, OASTDSS, STRAC) using OAST kit and populated data using OAST kit. By default databases were created as admin managed. Perform the following steps to convert all four databases to policy managed.

1. Shutdown all four databases.
2. Create the server pool using the following commands as an Oracle user:

```
oracle@orarac1:>srvctl add srvpool -g oastdbsvp  -l 1 -u 8  -i 999 -n
"orarac1,orarac2,orarac3,orarac4,orarac5,orarac6,orarac7,orarac8"

Convert all four databases to policy managed database using as oracle user.

oracle@orarac1:>srvctl modify database -d oastdb -g oastdbsvp
```

```
oracle@orarac1:>srvctl modify database -d oastoltp -g oastdbsvp
oracle@orarac1:>srvctl modify database -d oastdss -g oastdbsvp
oracle@orarac1:>srvctl modify database -d strac -g oastdbsvp
```

All four databases were created for the following purposes.

- OASTOLTP database—200 Users performing smaller and random transactions.
- OASTDSS database—8 Users with larger transactions
- OASTDB database—220 users performing smaller and random transactions and stressing CPU.
- STRAC databse—16 users to generate high amount of cache fusion traffic.

For load generation, we used Oracle Automated Stress/System Testing (OAST) as the primary tool. OAST is an Oracle Partner Standard Stress Test Suite intended to test Oracle on a particular system by simulating a real world environment. OAST is comprised of a set of test drivers for stress testing platform specific CPU, I/O, IPC, and Memory subsystem when running 11g Release 2 Oracle database server software. OAST workloads can be tuned to maximize the loads on all nodes within a cluster or a single node, when running a single instance database.

For initial configuration, OAST takes user input about the system and feeds them into its built-in sizing engine. The sizing engine produces an input parameter (number of warehouses) to the stress OLTP kit generator, and creates a stress OLTP kit. This OLTP kit then is used to prepare the database, create the schema, populate the tables and views, and create indexes. The run script also prepares database backup and restore and performs stress test runs, and transaction driver recompilation. Once the kit is recompiled, we are ready to create the schema and populate the databases.

OASTOLTP Database

Using OAST toolkit, we populated an OASTOLTP database (1500 Warehouses) for OLTP workload as shown below.

```
oracle@orarac1:~> sqlplus oastoltp/oastoltp
SQL*Plus: Release 11.2.0.2.0 Production on Tue May 1 xx:xx:xx xxxx
Copyright (c) 1982, 2010, Oracle.  All rights reserved.
Connected to:
Oracle Database 11g Enterprise Edition Release 11.2.0.2.0 - 64bit Production
With the Partitioning, Real Application Clusters, Automatic Storage
ManagemOLAP, Data Mining and Real Application Testing options


OASTOLTP@oastoltp_1 > select table_name, num_rows from user_tables;
```

TABLE_NAME	NUM_ROWS
WARE	1500
DIST	15,889
ITEM	1,04,235
NORD	14,470,544
CUST	45,685,156
HIST	46,322,879
ORDR	44,835,952
STOK	150,173,385
ORDL	459,096,209

OASTDB Database

Using OAST toolkit, we populated another OLTP database (OASTDB database) for CPU Stress (1500 Warehouses) as shown below.

```
oracle@orarac1:~> sqlplus oastoltp/oastoltp
SQL*Plus: Release 11.2.0.2.0 Production on Tue May 1 xx:xx:xx xxxx
Copyright (c) 1982, 2010, Oracle.  All rights reserved.
Connected to:
Oracle Database 11g Enterprise Edition Release 11.2.0.2.0 - 64bit Production
With the Partitioning, Real Application Clusters, Automatic Storage
Management, OLAP, Data Mining and Real Application Testing options
OASTOLTP@oastdb_1 > select table_name, num_rows from user_tables;
```

TABLE_NAME	NUM_ROWS
WARE	1500
DIST	15,889
ITEM	1,04,235
NORD	14,929,506
CUST	46,002,852
HIST	46,327,704
ORDR	44,507,991
STOK	150,026,291
ORDL	460,723,314

DSS Database

Next, the DSS database (750 Warehouses) was populated as shown below.

Note Both OLTP and DSS databases have same user named oastoltp.

```
oracle@orarac1:~> sqlplus oastoltp/oastoltp
SQL*Plus: Release 11.2.0.2.0 Production on Tue May 1 xx:xx:xx xxxx
Copyright (c) 1982, 2010, Oracle.  All rights reserved.
Connected to:
Oracle Database 11g Enterprise Edition Release 11.2.0.2.0 - 64bit Production
With the Partitioning, Real Application Clusters, Automatic Storage
Management, OLAP, Data Mining and Real Application Testing options
OASTOLTP@oastdss_1 > select table_name, num_rows from user_tables;
```

TABLE_NAME	NUM_ROWS
WARE	750
DIST	7,500
ITEM	1,04,237
NORD	7,053,265
CUST	22,479,675
HIST	23,603,890
ORDR	22,491,678
STOK	74,673,399
ORDL	234,400,943

STRAC Database

This is another customized Database that generates cache fusion traffic. This is useful to flood the Private Interconnect with more traffic while the OLTP and DSS workloads are being generated.

Additional Configuration

After the databases are created and populated, use Oracle srvctl utility to modify the instance configuration. Use of srvctl modify preserves the environment in the OCR configuration that would otherwise need to be re-entered. The following are the commands to modify the instance configuration.

```
Assign the respective DG group to the Databases
srvctl modify database -d oastdb  -a DATADG,REDODG
srvctl modify database -d oastdss -a DATADG,REDODG
srvctl modify database -d oastoltp -a DATADG,REDODG
srvctl modify database -d strac   -a DATADG,REDODG

Assign the respective Database Instances to the respective nodes

OASTDB
srvctl modify instance -d oastdb -n orarac1 -i oastdb_1
srvctl modify instance -d oastdb -n orarac2 -i oastdb_2
srvctl modify instance -d oastdb -n orarac3 -i oastdb_3
srvctl modify instance -d oastdb -n orarac4 -i oastdb_4
srvctl modify instance -d oastdb -n orarac5 -i oastdb_5
srvctl modify instance -d oastdb -n orarac6 -i oastdb_6
srvctl modify instance -d oastdb -n orarac7 -i oastdb_7
srvctl modify instance -d oastdb -n orarac8 -i oastdb_8
OASTDSS
srvctl modify instance -d oastdss -n orarac1 -i oastdss_1
srvctl modify instance -d oastdss -n orarac2 -i oastdss_2
srvctl modify instance -d oastdss -n orarac3 -i oastdss_3
srvctl modify instance -d oastdss -n orarac4 -i oastdss_4
srvctl modify instance -d oastdss -n orarac5 -i oastdss_5
```

```
srvctl modify instance -d oastdss -n orarac6 -i oastdss_6
srvctl modify instance -d oastdss -n orarac7 -i oastdss_7
srvctl modify instance -d oastdss -n orarac8 -i oastdss_8
OASTOLTP
srvctl modify instance -d oastoltp -n orarac1 -i oastoltp_1
srvctl modify instance -d oastoltp -n orarac2 -i oastoltp_2
srvctl modify instance -d oastoltp -n orarac3 -i oastoltp_3
srvctl modify instance -d oastoltp -n orarac4 -i oastoltp_4
srvctl modify instance -d oastoltp -n orarac5 -i oastoltp_5
srvctl modify instance -d oastoltp -n orarac6 -i oastoltp_6
srvctl modify instance -d oastoltp -n orarac7 -i oastoltp_7
srvctl modify instance -d oastoltp -n orarac8 -i oastoltp_8
STRAC
srvctl modify instance -d strac -n orarac1 -i strac_1
srvctl modify instance -d strac -n orarac2 -i strac_2
srvctl modify instance -d strac -n orarac3 -i strac_3
srvctl modify instance -d strac -n orarac4 -i strac_4
srvctl modify instance -d strac -n orarac5 -i strac_5
srvctl modify instance -d strac -n orarac6 -i strac_6
srvctl modify instance -d strac -n orarac7 -i strac_7
srvctl modify instance -d strac -n orarac8 -i strac_8
Add services to the instances.
srvctl add service -d oastdb -s oastdbfan -B THROUGHPUT -j SHORT -q TRUE -g
oastdbsvp -c uniform

srvctl add service -d oastdss -s oastdssfan -B THROUGHPUT -j SHORT -q TRUE
-g oastdbsvp -c uniform

srvctl add service -d oastoltp -s oastoltpfan -B THROUGHPUT -j SHORT -q TRUE
-g oastdbsvp -c uniform

srvctl add service -d strac -s stracfan -B THROUGHPUT -j SHORT -q TRUE -g
oastdbsvp -c uniform

srvctl start service -s oastdbfan -d oastdb

srvctl start service -s oastdssfan -d oastdss

srvctl start service -s oastoltpfan -d oastoltp

srvctl start service -s stracfan -d strac
```

Testing Workload Performance

After the databases are created and configured, we calibrated the loads to establish baseline for the stress tests as per criteria listed below.

- Normal CPU utilization should be close to 100% and run queues higher than 30 should be achieved. External tools such as CPU busy scripts can be used as a supplemental tool to maximize CPU.
- Memory utilization greater than 90 percent should be sustained and spikes in the workload should cause memory utilization going over 100% occasionally to force, memory paging, swapping and defragmentation.
- Disk IO stress should be high, but not as high that it would limit the CPU utilization on the system. If not enough disk stress can be achieved by just running the database workloads, artificial stress should be added to the system using disk stress tools.
- Network IO stress should be high, but not as high that it would limit the CPU utilization on the system. If not enough network stress can be achieved by just running the database workloads, artificial stress should be added to the system using network stress tools.

After the appropriate baseline is established for the above stress level criteria, the load must be run for 24 hours. We established the baseline with:

- Database workloads:
  - OLTP workload with 2400 users (Approx 300 users/node)
  - DSS workload with 160 users (Approx. 20 users/node)
  - STRAC cache fusion workload scripts were run to flood private interconnect
- CPU stress Program: This program runs high CPU cycle oriented functions and saturates CPU usage. This program runs on all 8 nodes.

The following are the general observations from the stress tests followed by workload specific details.

- Observations:
  - Achieved very consistent resource utilization around 90 percent on all 8 nodes.
  - No saturation levels or resource stalling on any subsystems (CPU, disk, I/O, or networking)
  - Sustained FCoE-based I/O ranging between 2.0 GBytes/Sec and 2.4 GBytes per second.
  - No performance degradation over time or occurrences of bottlenecks or wait times.
  - Excellent I/O service times from the EMC VNX storage.

This consistent workload performance without any significant tuning can be attributed to:

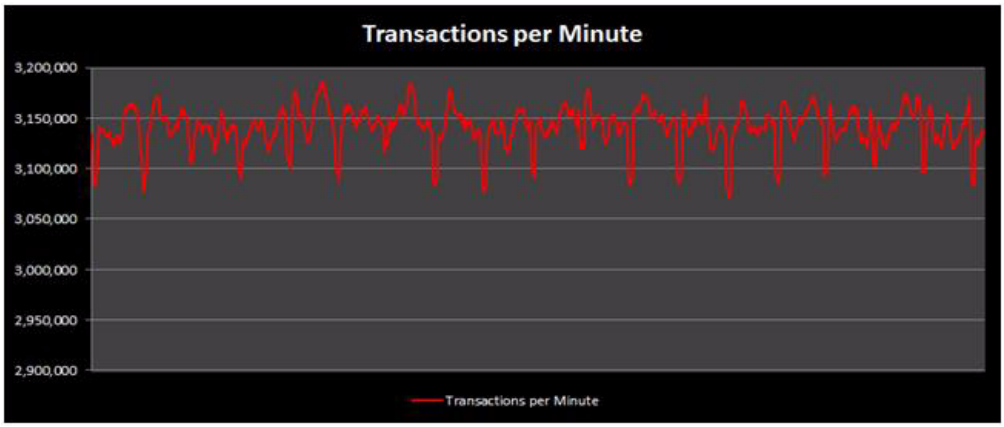
- The simplified, excellent architectural design of the Cisco Unified Computing System based on a 10-Gbps unified fabric.
- The pairing of the Cisco Unified Computing System with EMC VNX storage with high-performance Flash drives.

Note This is a testing, not a performance benchmarking, exercise. The numbers presented here should not be used for comparison purposes. The intent here is to look at the Cisco Unified Computing System supporting a sustained load over a long time period. Note that no tuning was performed and the lack of resource saturation without external load generation tools indicate that significant headroom is available to support greater performance than that shown in this exercise.

OLTP Workload

The figure below shows a three hour sample of all OLTP transactions during the stress test.

Figure 10 OLTP Transaction per Minute



For OLTP workload, we observed the transactions ranging between 3 Million to 3.2 Million.

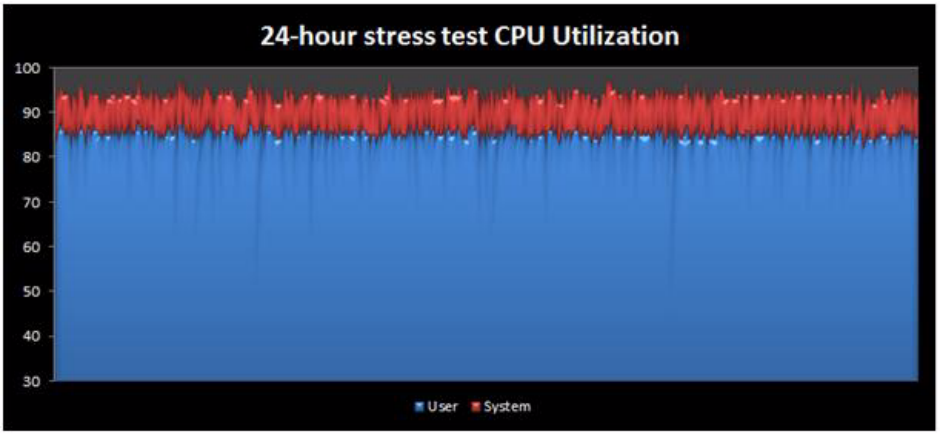
DSS Workload

DSS workloads are sequential in nature, read intensive and typically exercise larger IO patterns. For DSS workloads, it is common practice to set the parallel queries and the degree of parallelism on heavily read tables. This practice was followed in the test environment and achieved excellent performance. The AWR report section below indicates that DSS workload is generating approximately 900 Mbytes/Sec of IO throughout the stress run(24-hours).

	Reads MB/sec			Writes MB/sec				Reads requests/sec			Writes requests/sec			
I#	Total	Data File	Temp File	Total	Data File	Temp File	Log File	Total	Data File	Temp File	Total	Data File	Temp File	Log File
1	101.57	101.52	0.00	0.02	0.01	0.00	0.00	242.38	239.38	0.00	1.44	0.67	0.00	0.42
2	115.82	115.77	0.00	0.02	0.01	0.00	0.00	194.59	191.62	0.00	1.61	0.82	0.00	0.46
3	106.16	106.11	0.00	0.01	0.01	0.00	0.00	173.82	170.86	0.00	1.12	0.57	0.00	0.22
4	122.71	122.66	0.00	0.01	0.01	0.00	0.00	209.81	206.84	0.00	1.14	0.55	0.00	0.25
5	114.46	114.41	0.00	0.01	0.01	0.00	0.00	197.04	194.07	0.00	1.13	0.56	0.00	0.25
6	133.65	133.61	0.00	0.01	0.01	0.00	0.00	228.33	225.37	0.00	1.02	0.47	0.00	0.22
7	107.05	107.01	0.00	0.01	0.01	0.00	0.00	185.08	182.12	0.00	1.03	0.49	0.00	0.21
8	92.40	92.36	0.00	0.01	0.01	0.00	0.00	144.41	141.45	0.00	1.07	0.52	0.00	0.21
Sum	893.83	893.46	0.00	0.12	0.05	0.00	0.02	1,575.47	1,551.73	0.01	9.58	4.63	0.00	2.26
Avg	111.73	111.68	0.00	0.01	0.01	0.00	0.00	196.93	193.97	0.00	1.20	0.58	0.00	0.28

System Resource Utilization

Figure 11 Average CPU Utilization for Each Node



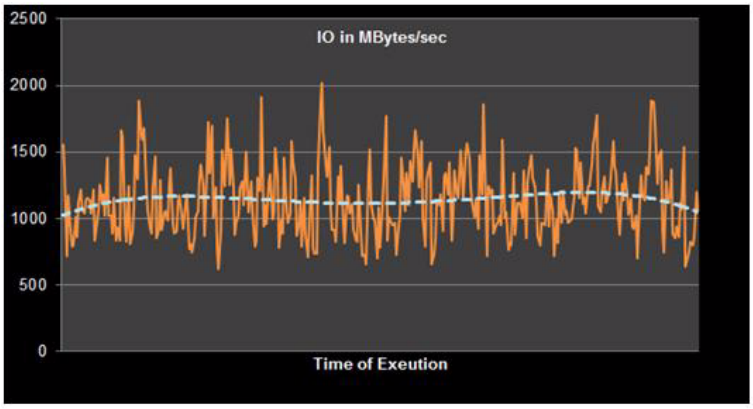
During the stress test run, we observed CPU utilization consistently above 90 percent across all 8 nodes of the cluster, as indicated from a snippet of the Oracle AWR report.

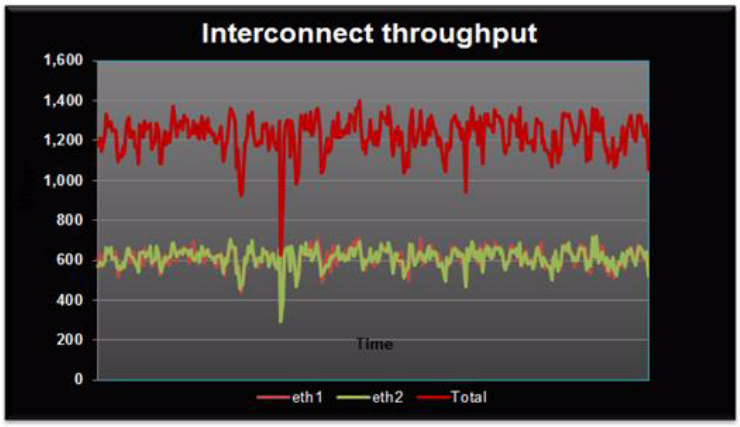
I#	Num CPUs	CPU Cores	CPU Sckts	Load Begin	Load End	% Busy	% Usr	% Sys	% WIO	% Idl	Busy Time (s)	Idle Time (s)	Memory (M)
1	24	12	2	13.12	14.16	94.03	85.54	6.07	1.84	5.97	244,726.02	15,533.41	96,684.48
2	24	12	2	8.70	4.98	97.38	89.75	5.43	0.56	2.62	252,756.23	6,790.14	96,684.48
3	24	12	2	9.38	6.51	95.68	87.17	5.89	0.81	4.32	248,280.81	11,213.06	96,684.48
4	24	12	2	15.69	14.46	93.80	86.18	5.39	1.88	6.20	244,297.41	16,158.60	96,684.48
5	24	12	2	13.78	11.58	92.92	84.65	5.74	2.09	7.08	242,199.30	18,457.37	96,684.48
6	24	12	2	10.06	13.06	89.77	81.49	5.79	3.19	10.23	234,146.81	26,675.51	96,684.48
7	24	12	2	15.23	16.86	92.16	84.39	5.44	1.94	7.84	239,996.70	20,430.00	96,684.48
8	24	12	2	15.65	16.25	94.41	86.43	5.56	1.54	5.59	245,273.72	14,511.86	96,684.48

IO and Network Traffic

The majority of network traffic in this testing consists of Oracle RAC interconnect traffic and IO traffic is SAN traffic to databases.

Figure 12 Network and IO Traffic During the Stress Test





Improved Availability with Cisco UCS Failover

Previous sections described Cisco Unified Computing System installation, configuration, and performance. This section examines the Cisco Unified Computing System's nearly instant failover capabilities to show how they can improve overall availability after unexpected, but common, hardware failures attributed to ports and cables.

Figure 13 shows the failure scenarios (indicated by numbers) that were tested under the stress conditions described in the preceding section, Testing Workload Performance.

Figure 13 Oracle RAC Private Interconnect Traffic Localized in a Cisco UCS Domain Under Normal Operations

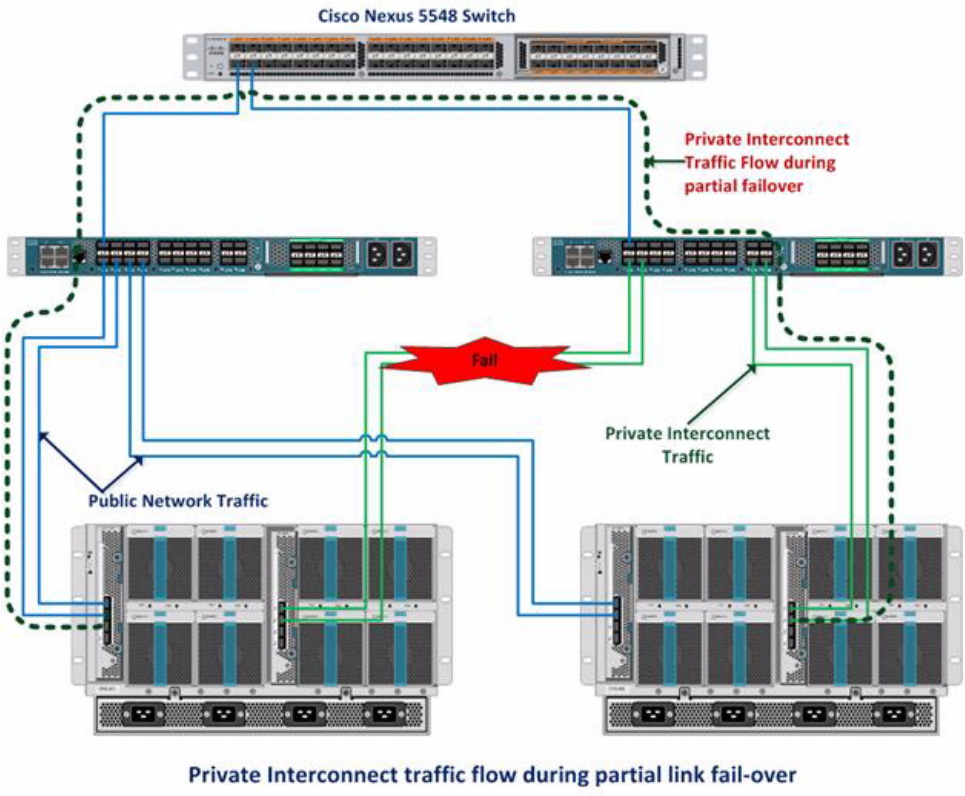


Figure 14 Oracle RAC Partial Failure Interconnect Traffic Traversing to Northbound Nexus 5548 Switch

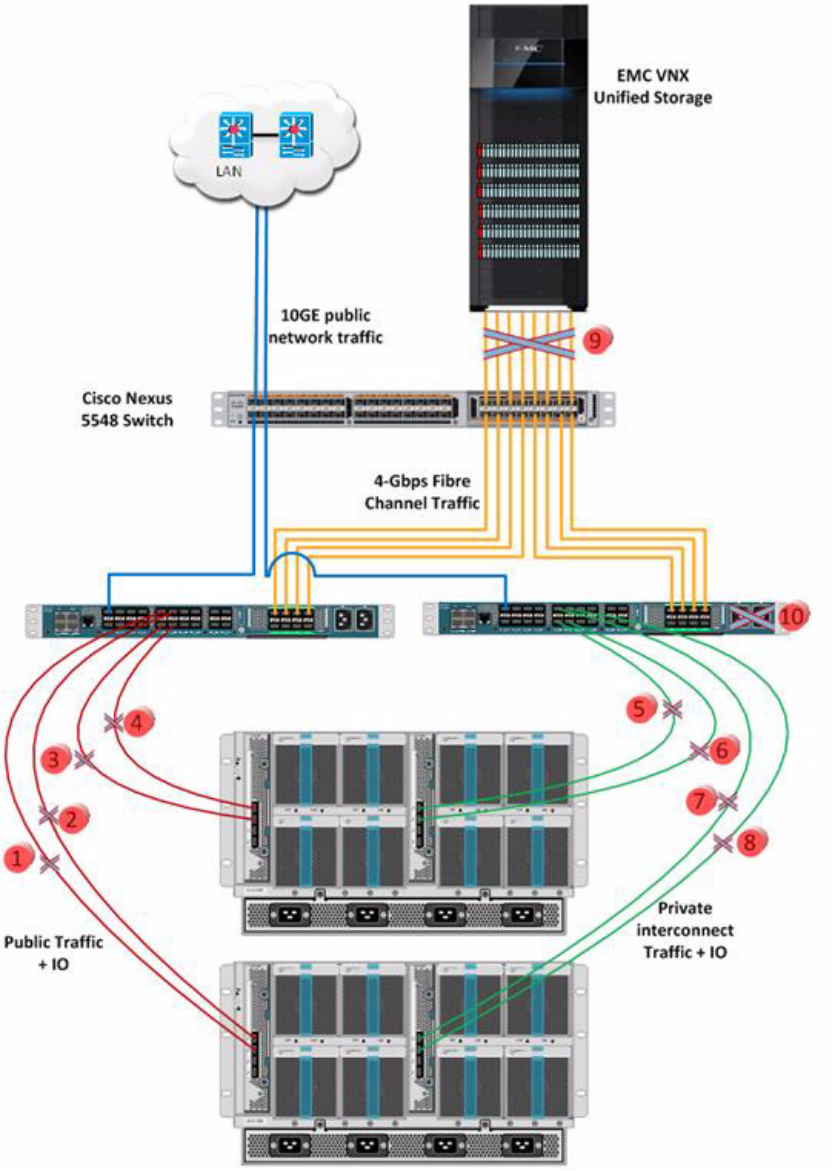


Table 9 summarizes the failure scenarios (each indicated by a number in Figure 10) and describes how the Cisco Unified Computing System architecture sustains unexpected failures related to ports, links, and the fabric interconnect (a rare occurrence).

Table 9 Failure Scenarios and the Cisco Unified Computing System Response

Failure Scenario (Figure 10)	Description	Response
1 or 2 or 3 or 4	Single-link failure (public interface)	Represents port failure; nodes should continue to work
1,2, 3 and 4	All-links failure (public interface)	Nodes should continue to work through other I/O module
5 or 6 or 7 or 8	Single-link failure (private interconnect)	Represents port failure; nodes should continue to work
5,6,7 and 8	All-links failure (private interconnect)	Nodes should continue to work through other I/O module
9	Storage path failure	Single I/O path loss should have no effect; if all I/O paths fail, nodes should reboot
(1,2,7 and 8) or (3,4,5 and 6)	Failure of both I/O modules on a single chassis	All nodes in failed chassis should reboot; nodes in other chassis should continue to work
10	Single Fabric interconnect failure	All nodes should continue to work through other fabric interconnect

Conclusion

Designed using a new and innovative approach to improve data center infrastructure, the Cisco Unified Computing System unites compute, network, storage access, and virtualization resources into a scalable, modular architecture that is managed as a single system.

For the Cisco Unified Computing System, Cisco has partnered with Oracle because Oracle databases and applications provide mission-critical software foundations for the majority of large enterprises worldwide. In addition, the architecture and large memory capabilities of the Cisco Unified Computing System connected to the industry-proven and scalable VNX storage system enable customers to scale and manage Oracle database environments in ways not previously possible.

Both database administrators and system administrators will benefit from the Cisco Unified Computing System combination of superior architecture, outstanding performance, and unified fabric. They can achieve demonstrated results by following the documented best practices for database installation, configuration, and management outlined in this document.

In summary, the Cisco Unified Computing System is a new computing model that uses integrated management and combines a wire-once unified fabric with an industry-standard computing platform.

- The platform
  - Optimizes database environments
  - Reduces total overall cost of the data center
  - Provides dynamic resource provisioning for increased business agility
- The benefits of the Cisco Unified Computing System include:
  - Reduced TCO—Enables up to 20 percent reduction in capital expenditures (CapEx) and up to 30 percent reduction in operating expenses (OpEx)
  - Improved IT productivity and business agility—Enables IT to provision applications in minutes instead of days and shifts the focus from IT maintenance to IT innovation
  - Increased scalability without added complexity—Managed as a single system, whether the system has one server or 320 servers with thousands of virtual machines
  - Improved energy efficiency—Significantly reduces power and cooling costs

For More Information

Please visit http://www.cisco.com/en/US/netsol/ns944/index.html#.

Appendix A—Cisco Unified Computing System Kernel Settings (/etc/sysctl.conf) and user limits settings /etc/security/limits.conf

This Appendix provides the parameters for the Cisco Unified Computing System with 24 GB of RAM.

It is highly recommended that you use the Oracle Validated RPM to derive kernel settings that are most suitable for your system.

```
# Kernel sysctl configuration files for Red Hat Linux
# for binary values, 0 is disabled, 1 is enabled.  See sysctl(8) and
# sysctl.conf(5) for more details.
# Controls IP packet forwarding
net.ipv4.ip_forward = 0
# Controls source route verification
net.ipv4.conf.default.rp_filter = 1
# Do not accept source routing
net.ipv4.conf.default.accept_source_route = 0
# Controls the System Request debugging functionality of the kernel
kernel.sysrq = 1
# Controls whether core dumps will append the PID to the core filename
# Useful for debugging multi-threaded applications
kernel.core_uses_pid = 1
# Controls the use of TCP syncookies
net.ipv4.tcp_syncookies = 1
# Controls the maximum size of a message, in bytes
kernel.msgmnb = 65536
# Controls the default maximum size of a message queue
kernel.msgmax = 8192
# Controls the maximum shared segment size, in bytes
kernel.shmmax = 4398046511104
# Controls the maximum number of shared memory segments, in pages
kernel.shmall = 1073741824
kernel.msgmni = 2878
kernel.sem = 250 32000 100 142
kernel.shmmni = 4096
net.core.rmem_default = 262144
# For 11g recommended value for net.core.rmem_max is 4194304
net.core.rmem_max = 4194304
```

```
net.core.wmem_default = 262144
fs.aio-max-nr = 3145728
fs.file-max = 6815744
net.ipv4.ip_local_port_range = 9000 65500
net.core.wmem_max = 1048576

Update /etc/security/limits.conf
     ## added ##
oracle  soft    nproc   2047
oracle  hard    nproc   16384
oracle  soft    nofile  1024
oracle  hard    nofile  65536
grid hard nofile 65536
```

Appendix B—Oracle RAC 11gR2 Database Instances Configuration

Check database figuration about the services

```
srvctl config database -d oastoltp -a
Database unique name: oastoltp
Database name: oastoltp
Oracle home: /u01/app/oracle/product/11.2.0/ucsrac
Oracle user: oracle
Spfile: +DATADG/oastdb/spfileoastoltp.ora
Domain:
Start options: open
Stop options: immediate
Database role: PRIMARY
Management policy: AUTOMATIC
Server pools: oastdbsvp
Database instances:
oastoltp_1,oastoltp_2,oastoltp_3,oastoltp_4,oastoltp_5,oastoltp_6,oastoltp_7,
oastoltp_8
Disk Groups: DATADG,REDODG
Services: oastoltpfan
Database is enabled
Database is policy managed

srvctl config database -d oastdb -a
Database unique name: oastdb
Database name: oastdb
Oracle home: /u01/app/oracle/product/11.2.0/ucsrac
```

```
Oracle user: oracle
Spfile: +DATADG/oastdb/spfileoastdb.ora
Domain:
Start options: open
Stop options: immediate
Database role: PRIMARY
Management policy: AUTOMATIC
Server pools: oastdbsvp
Database instances:
oastdb_1,oastdb_2,oastdb_3,oastdb_4,oastdb_5,oastdb_6,oastdb_7,oastdb_8
Disk Groups: DATADG,REDODG
Services: oastdbfan
Database is enabled
Database is policy managed
srvctl config database -d oastdss -a
Database unique name: oastdss
Database name: oastdss
Oracle home: /u01/app/oracle/product/11.2.0/ucsrac
Oracle user: oracle
Spfile: +DATADG/oastdss/spfileoastdss.ora
Domain:
Start options: open
Stop options: immediate
Database role: PRIMARY
Management policy: AUTOMATIC
Server pools: oastdbsvp
Database instances:
oastdss_1,oastdss_2,oastdss_3,oastdss_4,oastdss_5,oastdss_6,oastdss_7,oastdss
_8
Disk Groups: DATADG,REDODG
Services: oastdssfan
Database is enabled
Database is policy managed
srvctl config database -d strac -a
Database unique name: strac
Database name: strac
Oracle home: /u01/app/oracle/product/11.2.0/ucsrac
Oracle user: oracle
Spfile: +DATADG/strac/spfilestrac.ora
Domain:
Start options: open
Stop options: immediate
Database role: PRIMARY
```

```
Management policy: AUTOMATIC
Server pools: oastdbsvp
Database instances:
strac_1,strac_2,strac_3,strac_4,strac_5,strac_6,strac_7,strac_8
Disk Groups: DATADG,REDODG
Services:
Database is enabled
Database is policy managed
```

Check Cluster Nodes Status [olsnodes -n -s -i]

Node Name	Node Number	Node VIP Address (managed by Oracle GNS)	Node Status
orarac1	1	10.29.134.74	Active
orarac2	2	10.29.134.76	Active
orarac3	3	10.29.134.75	Active
orarac4	4	10.29.134.78	Active
orarac5	5	10.29.134.77	Active
orarac6	6	10.29.134.74	Active
orarac7	7	10.29.134.80	Active
orarac8	8	10.29.134.73	Active

Cluster Ready Services Resource Status in tabular format [crsctl status resource -t]

NAME	TARGET	STATE	SERVER	STATE_DETAILS
Local Resources				
ora.ASMGRID.dg				
	ONLINE	ONLINE	orarac1	
	ONLINE	ONLINE	orarac2	
	ONLINE	ONLINE	orarac3	
	ONLINE	ONLINE	orarac4	
	ONLINE	ONLINE	orarac5	
	ONLINE	ONLINE	orarac6	
	ONLINE	ONLINE	orarac7	
	ONLINE	ONLINE	orarac8	
ora.DATADG.dg				
	ONLINE	ONLINE	orarac1	
	ONLINE	ONLINE	orarac2	
	ONLINE	ONLINE	orarac3	
	ONLINE	ONLINE	orarac4	
	ONLINE	ONLINE	orarac5	
	ONLINE	ONLINE	orarac6	
	ONLINE	ONLINE	orarac7	

	ONLINE	ONLINE	orarac8	
ora.REDODG.dg				
	ONLINE	ONLINE	orarac1	
	ONLINE	ONLINE	orarac2	
	ONLINE	ONLINE	orarac3	
	ONLINE	ONLINE	orarac4	
	ONLINE	ONLINE	orarac5	
	ONLINE	ONLINE	orarac6	
	ONLINE	ONLINE	orarac7	
	ONLINE	ONLINE	orarac8	
ora.LISTENER.lsnr				
	ONLINE	ONLINE	orarac1	
	ONLINE	ONLINE	orarac2	
	ONLINE	ONLINE	orarac3	
	ONLINE	ONLINE	orarac4	
	ONLINE	ONLINE	orarac5	
	ONLINE	ONLINE	orarac6	
	ONLINE	ONLINE	orarac7	
	ONLINE	ONLINE	orarac8	
ora.asm				
	ONLINE	ONLINE	orarac1	Started
	ONLINE	ONLINE	orarac2	Started
	ONLINE	ONLINE	orarac3	Started
	ONLINE	ONLINE	orarac4	Started
	ONLINE	ONLINE	orarac5	Started
	ONLINE	ONLINE	orarac6	Started
	ONLINE	ONLINE	orarac7	Started
	ONLINE	ONLINE	orarac8	Started
ora.eons				
	ONLINE	ONLINE	orarac1	
	ONLINE	ONLINE	orarac2	
	ONLINE	ONLINE	orarac3	
	ONLINE	ONLINE	orarac4	
	ONLINE	ONLINE	orarac5	
	ONLINE	ONLINE	orarac6	
	ONLINE	ONLINE	orarac7	
	ONLINE	ONLINE	orarac8	
ora.gsd				
	OFFLINE	OFFLINE	orarac1	
	OFFLINE	OFFLINE	orarac2	
	OFFLINE	OFFLINE	orarac3	
	OFFLINE	OFFLINE	orarac4	
	OFFLINE	OFFLINE	orarac5	
	OFFLINE	OFFLINE	orarac6	

	OFFLINE	OFFLINE	orarac7	
	OFFLINE	OFFLINE	orarac8	
ora.net1.network				
	ONLINE	ONLINE	orarac1	
	ONLINE	ONLINE	orarac2	
	ONLINE	ONLINE	orarac3	
	ONLINE	ONLINE	orarac4	
	ONLINE	ONLINE	orarac5	
	ONLINE	ONLINE	orarac6	
	ONLINE	ONLINE	orarac7	
	ONLINE	ONLINE	orarac8	
ora.ons				
	ONLINE	ONLINE	orarac1	
	ONLINE	ONLINE	orarac2	
	ONLINE	ONLINE	orarac3	
	ONLINE	ONLINE	orarac4	
	ONLINE	ONLINE	orarac5	
	ONLINE	ONLINE	orarac6	
	ONLINE	ONLINE	orarac7	
	ONLINE	ONLINE	orarac8	
ora.registry.acfs				
	ONLINE	ONLINE	orarac1	
	ONLINE	ONLINE	orarac2	
	ONLINE	ONLINE	orarac3	
	ONLINE	ONLINE	orarac4	
	ONLINE	ONLINE	orarac5	
	ONLINE	ONLINE	orarac6	
	ONLINE	ONLINE	orarac7	
	ONLINE	ONLINE	orarac8	
Cluster Resources				
ora.LISTENER_SCAN1.lsnr				
1	ONLINE	ONLINE	orarac7	
ora.LISTENER_SCAN2.lsnr				
1	ONLINE	ONLINE	orarac8	
ora.LISTENER_SCAN3.lsnr				
1	ONLINE	ONLINE	orarac2	
ora.gns				
1	ONLINE	ONLINE	orarac3	
ora.gns.vip				
1	ONLINE	ONLINE	orarac3	
ora.oastdb.db				
1	ONLINE	ONLINE	orarac1	Open

2	ONLINE	ONLINE	orarac7	Open
3	ONLINE	ONLINE	orarac3	Open
4	ONLINE	ONLINE	orarac4	Open
5	ONLINE	ONLINE	orarac5	Open
6	ONLINE	ONLINE	orarac6	Open
7	ONLINE	ONLINE	orarac2	Open
8	ONLINE	ONLINE	orarac8	Open

ora.oastdb.oastdbfan.svc				
1	ONLINE	ONLINE	orarac1	
2	ONLINE	ONLINE	orarac6	
3	ONLINE	ONLINE	orarac3	
4	ONLINE	ONLINE	orarac4	
5	ONLINE	ONLINE	orarac2	
6	ONLINE	ONLINE	orarac5	
7	ONLINE	ONLINE	orarac7	
8	ONLINE	ONLINE	orarac8	
ora.oastdss.db				
1	ONLINE	ONLINE	orarac1	Open
2	ONLINE	ONLINE	orarac6	Open
3	ONLINE	ONLINE	orarac3	Open
4	ONLINE	ONLINE	orarac4	Open
5	ONLINE	ONLINE	orarac2	Open
6	ONLINE	ONLINE	orarac5	Open
7	ONLINE	ONLINE	orarac7	Open
8	ONLINE	ONLINE	orarac8	Open
ora.oastdss.oastdssfan.svc				
1	ONLINE	ONLINE	orarac1	
2	ONLINE	ONLINE	orarac6	
3	ONLINE	ONLINE	orarac2	
4	ONLINE	ONLINE	orarac3	
5	ONLINE	ONLINE	orarac5	
6	ONLINE	ONLINE	orarac4	
7	ONLINE	ONLINE	orarac7	
8	ONLINE	ONLINE	orarac8	
ora.oc4j				
1	OFFLINE	OFFLINE		
ora.orarac1.vip				
1	ONLINE	ONLINE	orarac1	
ora.orarac2.vip				
1	ONLINE	ONLINE	orarac2	
ora.orarac3.vip				
1	ONLINE	ONLINE	orarac3	

ora.orarac4.vip				
1	ONLINE	ONLINE	orarac4	
ora.orarac5.vip				
1	ONLINE	ONLINE	orarac5	
ora.orarac6.vip				
1	ONLINE	ONLINE	orarac6	
ora.orarac7.vip				
1	ONLINE	ONLINE	orarac7	
ora.orarac8.vip				
1	ONLINE	ONLINE	orarac8	
ora.scan1.vip				
1	ONLINE	ONLINE	orarac7	
ora.scan2.vip				
1	ONLINE	ONLINE	orarac8	
ora.scan3.vip				
1	ONLINE	ONLINE	orarac2	
ora.strac.db				
1	ONLINE	ONLINE	orarac1	Open
2	ONLINE	ONLINE	orarac3	Open
3	ONLINE	ONLINE	orarac2	Open
4	ONLINE	ONLINE	orarac4	Open
5	ONLINE	ONLINE	orarac5	Open
6	ONLINE	ONLINE	orarac7	Open
7	ONLINE	ONLINE	orarac6	Open
8	ONLINE	ONLINE	orarac8	Open

Appendix C—DNS and GNS Setup and Configuration

If you would like to use Oracle GNS and it requires the DNS server, DHCP server to do the range of VIP(s) to be published to the respective 8 Node(s) VIP(s), 3 SCAN VIP(s).

```
DNS Settings (/etc/named.conf)
options {
        // DNS tables are located in the /var/named directory

 directory "/var/named";

        // Forward any unresolved requests to our ISP's name server
        // (this is an example IP address only -- do not use!)
        // forwarders {
        //      10;
        // };

        /*
```

```
         * If there is a firewall between you and nameservers you want
         * to talk to, you might need to uncomment the query-source
         * directive below.  Previous versions of BIND always asked
         * questions using port 53, but BIND 8.1 uses an unprivileged
         * port by default.
         */
        // query-source address * port 53;
        allow-query { any; };
};

// Enable caching and load root server info
zone "named.root" {
        type hint;
        file "";
};

zone "0.0.127.in-addr.arpa" {
        type master;
        file "127_0_0.rev";
        notify no;
};

zone "ucs.cisco.com" IN {
        type master;
        file "ucs_cisco_com.db";
};
```

```
DNS ' Oracle GNS Settings (/var/named/ucs_cisco_com.db)

; This is the Start of Authority (SOA) record.  Contains contact
; & other information about the name server.  The serial number
; must be changed whenever the file is updated (to inform secondary
; servers that zone information has changed).
$TTL 86400

$ORIGIN ucs.cisco.com.

@ IN SOA ucs.cisco.com. root.ucs.cisco.com. (
        19990811        ; Serial number
        3600            ; 1 hour refresh
        300             ; 5 minutes retry
```

```
        172800          ; 2 days expiry
        43200 )         ; 12 hours minimum

; List the name servers in use.  Unresolved (entries in other zones)
; will go to our ISP's name server ucs.cisco.com
        NSucs.cisco.com.
        TXT            "UCS Domain"
localhost       A       127.0.0.1
ucs.cisco.com.  A       10.29.134.85

; Oracle RAC - 11GR2 - The below server runs DNS and DHCP
ucsoast.ucs.cisco.com.     IN  A 10.29.134.85

; Oracle 11gR2 GNS related.
; Point of delegation (NS) Name Server Record to the RAC GNS Cluster Domain.
ucscluster.ucs.cisco.com. NS gns.ucscluster.ucs.cisco.com.
gns.ucscluster.ucs.cisco.com.  IN  A 10.29.134.82

DNS Settings (/var/named/127_0_0.rev)

$TTL 86400
0.0.127.in-addr.arpa. IN SOA ucs.cisco.com.  root.ucs.cisco.com. (
        19990811        ; Serial number
        3600            ; 1 hour refresh
        300             ; 5 minutes retry
        172800          ; 2 days expiry
        43200 )         ; 12 hours minim

0.0.127.in-addr.arpa.   IN      NS      ucs.cisco.com.
$ORIGIN 0.0.127.in-addr.arpa.
1 IN PTR localhost.ucs.cisco.com.

DHCP Settings (/etc/dhcpd.conf)

#
# DHCP Server Configuration file.
#   see /usr/share/doc/dhcp*/dhcpd.conf.sample
#
ddns-update-style interim;
ignore client-updates;

subnet 10.29.134.0 netmask 255.255.255.0 {
        option routers                  10.29.134.1;
```

```
        option ip-forwarding            off;
        option subnet-mask              255.255.255.0;
        option broadcast-address        10.29.134.255;

        option domain-name              "ucs.cisco.com";
        option domain-name-servers      10.29.134.85;

        option time-offset              -28800; # Pacific Standard Time
        option ntp-servers              171.68.10.80, 171.68.10.150;

        range 10.29.134.71 10.29.134.81;
        default-lease-time 604800; # 7 Days Lease Time
        max-lease-time 604800;
}

"Stop and Start of DNS, DHCP Servers

First make sure the DNS (named) and DHCP (dhcpd) is switched ON at various run
levels.

root@ # chkconfig --list | egrep 'named|dhcpd'
dhcpd           0:off   1:off   2:off   3:on    4:off   5:on    6:off
named           0:off   1:off   2:on    3:on    4:on    5:on    6:off
```

(a) **DNS MANAGEMENT**

service named stop

service named start

(b) **DHCP MANAGEMENT**

service dhcpd stop

Check and find out the RANGE IP (10.29.134.71 ?' 10.29.134.81) addresses given to DHCP are now leased to ALL respective RAC Services [VIP(s) and SCAN(s)].

```
[root@ucsoast ~]# cat /var/lib/dhcpd/dhcpd.leases
# All times in this file are in UTC (GMT), not your local timezone. This is not a
# bug, so please don't ask about it.  There is no portable way to store leases in
# the local timezone, so please don't request this as a feature. If this is
# Inconvenient or confusing to you, we sincerely apologize. Seriously, though
- don't # ask. The format of this file is documented in the dhcpd.leases(5)
manual page.
```

```
# This lease file was written by isc-dhcp-V3.0.5-RedHat
lease 10.29.134.79 {
  starts 5 2010/08/13 17:41:00;
  ends 5 2010/08/20 17:41:00;
  binding state active;
  next binding state free;
  hardware ethernet 00:00:00:00:00:00;
  uid "\000ucscluster-scan3-vip";
}
lease 10.29.134.72 {
  starts 5 2010/08/13 17:41:07;
  ends 5 2010/08/20 17:41:07;
  binding state active;
  next binding state free;
  hardware ethernet 00:00:00:00:00:00;
  uid "\000orarac6-vip";
}
lease 10.29.134.71 {
  starts 5 2010/08/13 17:41:12;
  ends 5 2010/08/20 17:41:12;
  binding state active;
  next binding state free;
  hardware ethernet 00:00:00:00:00:00;
  uid "\000ucscluster-scan1-vip";
}
lease 10.29.134.80 {
  starts 5 2010/08/13 17:41:12;
  ends 5 2010/08/20 17:41:12;
  binding state active;
  next binding state free;
  hardware ethernet 00:00:00:00:00:00;
  uid "\000orarac7-vip";
}
  lease 10.29.134.81 {
  starts 5 2010/08/13 17:41:16;
  ends 5 2010/08/20 17:41:16;
  binding state active;
  next binding state free;
  hardware ethernet 00:00:00:00:00:00;
  uid "\000ucscluster-scan2-vip";
}
lease 10.29.134.73 {
  starts 5 2010/08/13 17:41:16;
```

```
  ends 5 2010/08/20 17:41:16;
  binding state active;
  next binding state free;
  hardware ethernet 00:00:00:00:00:00;
  uid "\000orarac8-vip";
}
lease 10.29.134.77 {
  starts 5 2010/08/13 17:41:38;
  ends 5 2010/08/20 17:41:38;
  binding state active;
  next binding state free;
  hardware ethernet 00:00:00:00:00:00;
  uid "\000orarac5-vip";
}
lease 10.29.134.75 {
  starts 5 2010/08/13 17:41:43;
  ends 5 2010/08/20 17:41:43;
  binding state active;
  next binding state free;
  hardware ethernet 00:00:00:00:00:00;
  uid "\000orarac3-vip";
}
lease 10.29.134.78 {
  starts 5 2010/08/13 17:41:43;
  ends 5 2010/08/20 17:41:43;
  binding state active;
  next binding state free;
  hardware ethernet 00:00:00:00:00:00;
  uid "\000orarac4-vip";
}
lease 10.29.134.74 {
  starts 5 2010/08/13 19:09:40;
  ends 5 2010/08/20 19:09:40;
  binding state active;
  next binding state free;
  hardware ethernet 00:00:00:00:00:00;
  uid "\000orarac1-vip";
}
lease 10.29.134.76 {
  starts 5 2010/08/13 19:45:51;
  ends 5 2010/08/13 19:45:51;
  binding state free;
  hardware ethernet 00:00:00:00:00:00;
```

```
  uid "\000orarac2-vip";
}
```

Appendix D—Multicast Ranges and Cisco UCS Firmware Recommendations

Recommended Cisco UCS Firmware version for Multicast Range 230.0.1.0/24 (IGMP Querier required)

Adapter	UCS Firmware
M81KR (Palo)	1.3 / 1.4
M71KR (Menlo Qlogic or Emulex)	1.3 / 1.4

Recommended Cisco UCS Firmware version for Multicast Range 224.0.0.251/24 (IGMP Querier NOT required).

Adapter	UCS Firmware
M81KR (Palo)	1.4
M71KR (Menlo Qlogic or Emulex)	1.3 / 1.4

Appendix E—Sample Nexus 5548UP Configuration

In our configuration, the first 32 ports are left as 10G Ethernet ports(default configuration) and modified Slot 2 ports (1-16) as FC ports as shown below.

Configure Slot 2 (1-16) ports as FC ports on N5548UP

```
n5548# configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
n5548(config)# slot 2
n5548(config-slot)# port 1-16 type fc
n5548(config-slot)# exit
n5548(config)# reload
WARNING: There is unsaved configuration!!!
WARNING: This command will reboot the system
Do you want to continue? (y/n)y
```

When you have configured the ports and appropriate connectivity with storage, to configure the SAN zoning configuration. We have shown some commands for this SAN configuration but it is beyond the scope of this document to provide whole configuration. For details about additional configuration and best practices, please refer to N5548UP documentation.

```
Zoning information:

N5548# show zoneset active vsan 9
zoneset name OraRAC_11g vsan 9
  zone name RAC1_hba1 vsan 9
    pwwn 20:00:00:25:b5:25:01:a7
```

```
    pwwn 50:06:01:65:3e:a0:24:bd
    pwwn 50:06:01:67:3e:a0:24:bd
    pwwn 50:06:01:6d:3e:a0:24:bd
    pwwn 50:06:01:6f:3e:a0:24:bd

  zone name RAC1_hba2 vsan 9
    pwwn 20:00:00:25:b5:25:01:b7
    pwwn 50:06:01:64:3e:a0:24:bd
    pwwn 50:06:01:66:3e:a0:24:bd
    pwwn 50:06:01:6c:3e:a0:24:bd
    pwwn 50:06:01:6e:3e:a0:24:bd

  zone name RAC2_hba1 vsan 9
    pwwn 20:00:00:25:b5:25:01:a6
    pwwn 50:06:01:65:3e:a0:24:bd
    pwwn 50:06:01:67:3e:a0:24:bd
    pwwn 50:06:01:6d:3e:a0:24:bd
    pwwn 50:06:01:6f:3e:a0:24:bd

  zone name RAC2_hba2 vsan 9
    pwwn 20:00:00:25:b5:25:01:b6
    pwwn 50:06:01:64:3e:a0:24:bd
    pwwn 50:06:01:66:3e:a0:24:bd
    pwwn 50:06:01:6c:3e:a0:24:bd
    pwwn 50:06:01:6e:3e:a0:24:bd

  zone name RAC3_hba1 vsan 9
    pwwn 20:00:00:25:b5:25:01:a5
    pwwn 50:06:01:65:3e:a0:24:bd
    pwwn 50:06:01:67:3e:a0:24:bd
    pwwn 50:06:01:6d:3e:a0:24:bd
    pwwn 50:06:01:6f:3e:a0:24:bd

  zone name RAC3_hba2 vsan 9
    pwwn 20:00:00:25:b5:25:01:b5
    pwwn 50:06:01:64:3e:a0:24:bd
    pwwn 50:06:01:66:3e:a0:24:bd
    pwwn 50:06:01:6c:3e:a0:24:bd
    pwwn 50:06:01:6e:3e:a0:24:bd

  zone name RAC4_hba1 vsan 9
    pwwn 20:00:00:25:b5:25:01:a4
    pwwn 50:06:01:65:3e:a0:24:bd
```

```
  pwwn 50:06:01:67:3e:a0:24:bd
  pwwn 50:06:01:6d:3e:a0:24:bd
  pwwn 50:06:01:6f:3e:a0:24:bd

zone name RAC4_hba2 vsan 9
  pwwn 20:00:00:25:b5:25:01:b4
  pwwn 50:06:01:64:3e:a0:24:bd
  pwwn 50:06:01:66:3e:a0:24:bd
  pwwn 50:06:01:6c:3e:a0:24:bd
  pwwn 50:06:01:6e:3e:a0:24:bd

zone name RAC5_hba1 vsan 9
  pwwn 20:00:00:25:b5:25:01:a3
  pwwn 50:06:01:65:3e:a0:24:bd
  pwwn 50:06:01:67:3e:a0:24:bd
  pwwn 50:06:01:6d:3e:a0:24:bd
  pwwn 50:06:01:6f:3e:a0:24:bd

zone name RAC5_hba2 vsan 9
  pwwn 20:00:00:25:b5:25:01:b3
  pwwn 50:06:01:64:3e:a0:24:bd
  pwwn 50:06:01:66:3e:a0:24:bd
  pwwn 50:06:01:6c:3e:a0:24:bd
  pwwn 50:06:01:6e:3e:a0:24:bd

zone name RAC6_hba1 vsan 9
  pwwn 20:00:00:25:b5:25:01:a2
  pwwn 50:06:01:65:3e:a0:24:bd
  pwwn 50:06:01:67:3e:a0:24:bd
  pwwn 50:06:01:6d:3e:a0:24:bd
  pwwn 50:06:01:6f:3e:a0:24:bd

zone name RAC6_hba2 vsan 9
  pwwn 20:00:00:25:b5:25:01:b2
  pwwn 50:06:01:64:3e:a0:24:bd
  pwwn 50:06:01:66:3e:a0:24:bd
  pwwn 50:06:01:6c:3e:a0:24:bd
  pwwn 50:06:01:6e:3e:a0:24:bd

zone name RAC7_hba1 vsan 9
  pwwn 20:00:00:25:b5:25:01:a1
  pwwn 50:06:01:65:3e:a0:24:bd
  pwwn 50:06:01:67:3e:a0:24:bd
```

```
    pwwn 50:06:01:6d:3e:a0:24:bd
    pwwn 50:06:01:6f:3e:a0:24:bd

  zone name RAC7_hba2 vsan 9
    pwwn 20:00:00:25:b5:25:01:b1
    pwwn 50:06:01:64:3e:a0:24:bd
    pwwn 50:06:01:66:3e:a0:24:bd
    pwwn 50:06:01:6c:3e:a0:24:bd
    pwwn 50:06:01:6e:3e:a0:24:bd

  zone name RAC8_hba1 vsan 9
    pwwn 20:00:00:25:b5:25:01:a0
    pwwn 50:06:01:65:3e:a0:24:bd
    pwwn 50:06:01:67:3e:a0:24:bd
    pwwn 50:06:01:6d:3e:a0:24:bd
    pwwn 50:06:01:6f:3e:a0:24:bd

  zone name RAC8_hba2 vsan 9
    pwwn 20:00:00:25:b5:25:01:b0
    pwwn 50:06:01:64:3e:a0:24:bd
    pwwn 50:06:01:66:3e:a0:24:bd
    pwwn 50:06:01:6c:3e:a0:24:bd
    pwwn 50:06:01:6e:3e:a0:24:bd
```