Deploying Oracle RAC 11gR2 (11.2.0.3) on Oracle Linux 6.3 using Cisco Unified Computing System 2.1 and EMC VNX7500

Last Updated: June 5, 2013

Cisco Validated Design

Building Architectures to Solve Business Problems

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 RAC 11gR2 (11.2.0.3) on Oracle Linux 6.3 using Cisco Unified Computing System 2.1 and EMC VNX7500

Overview

This Cisco Validated Design describes how the Cisco Unified Computing System,Ñ¢ can be used in conjunction with EMC¬Æ VNX¬Æ 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.

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

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

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¬Æ 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 Cisco Validated 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 E7- 4870 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 hyper threading 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. 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 Cisco Unified Computing System

Cisco Unified Computing System unites computing, networking, storage access, and virtualization resources into a single cohesive system. When used as the foundation for Oracle RAC database and software the system brings lower total cost of ownership (TCO), greater performance, improved scalability, increased business agility, and Cisco's hallmark investment protection.

The system represents a major evolutionary step away from the current traditional platforms in which individual components must be configured, provisioned, and assembled to form a solution. Instead, the system is designed to be stateless. It is installed and wired once, with its entire configuration-from RAID controller settings and firmware revisions to network configurations-determined in software using integrated, embedded management.

The system brings together Intel Xeon processor-powered server resources on a 10-Gbps unified fabric that carries all IP networking and storage traffic, eliminating the need to configure multiple parallel IP and storage networks at the rack level. The solution dramatically reduces the number of components needed compared to other implementations, reducing TCO, simplifying and accelerating deployment, and reducing the complexity that can be a source of errors and cause downtime.

Cisco UCS is designed to be form-factor neutral. The core of the system is a pair of Fabric Interconnects that link all the computing resources together and integrate all system components into a single point of management. Today, blade server chassis are integrated into the system through Fabric Extenders that bring the system's 10-Gbps unified fabric to each chassis.

The Fibre Channel over Ethernet (FCoE) protocol collapses Ethernet-based networks and storage networks into a single common network infrastructure, thus reducing CapEx by eliminating redundant switches, cables, networking cards, and adapters, and reducing OpEx by simplifying administration of these networks (Figure 1). Other benefits include:

- I/O and server virtualization
- Transparent scaling of all types of content, either block or file based
- Simpler and more homogeneous infrastructure to manage, enabling data center consolidation

Fabric Interconnects

The Cisco Fabric Interconnect is a core part of Cisco UCS, providing both network connectivity and management capabilities for the system. It offers line-rate, low-latency, lossless 10 Gigabit Ethernet, FCoE, and Fibre Channel functions.

The Fabric Interconnect provides the management and communication backbone for the Cisco UCS B-Series Blade Servers and Cisco UCS 5100 Series Blade Server Chassis. All chassis, and therefore all blades, attached to the Fabric Interconnects become part of a single, highly available management domain. In addition, by supporting unified fabric, Fabric Interconnects support both LAN and SAN connectivity for all blades within their domain. The Fabric Interconnect supports multiple traffic classes over a lossless Ethernet fabric from a blade server through an interconnect. Significant TCO savings come from an FCoE-optimized server design in which network interface cards (NICs), host bus adapters (HBAs), cables, and switches can be consolidated.

The Cisco UCS 6248 Fabric Interconnect is a one-rack-unit (1RU), 10 Gigabit Ethernet, IEEE Data Center Bridging (DCB), and FCoE interconnect built to provide 960 Gbps throughput with very low latency. It has 48 high density ports in 1RU including one expansion module with 16 unified ports. Like its predecessors, it can be seamlessly managed with Cisco UCS manager.

Fabric Extenders

The Cisco Fabric Extenders multiplex and forward all traffic from blade servers in a chassis to a parent Cisco UCS Fabric Interconnect from 10-Gbps unified fabric links. All traffic, even traffic between blades on the same chassis, is forwarded to the parent interconnect, where network profiles are managed efficiently and effectively by the Fabric Interconnect. At the core of the Cisco UCS Fabric Extender are application-specific integrated circuit (ASIC) processors developed by Cisco that multiplex all traffic.

The Cisco UCS 2208XP Fabric Extender has eight 10 Gigabit Ethernet, FCoE-capable, enhanced small Form-Factor Pluggable (SFP+) ports that connect the blade chassis to the fabric interconnect. Each Cisco UCS 2208XP has thirty-two 10 Gigabit Ethernet ports connected through the midplane to each half-width slot in the chassis. Typically configured in pairs for redundancy, two fabric extenders provide up to 160 Gbps of I/O to the chassis. Each fabric extender on either sides of the chassis are connected through 8 x 10 Gb links to the fabric interconnects and offer:

- Connection of the Cisco UCS blade chassis to the Fabric Interconnect
- Eight 10 Gigabit Ethernet, FCoE-capable SFP+ ports
- Built-in chassis management function to manage the chassis environment (the power supply and fans as well as the blades) along with the Fabric Interconnect, eliminating the need for separate chassis management modules
- Full management by Cisco UCS Manager through the Fabric Interconnect
- Support for up to two Fabric Extenders, enabling increased capacity as well as redundancy
- Up to 160 Gbps of bandwidth per chassis

Blade Chassis

The Cisco UCS 5100 Series Blade Server Chassis is a crucial building block of Cisco UCS, delivering a scalable and flexible blade server chassis.

Cisco UCS Manager

Cisco UCS Manager provides unified, embedded management of all software and hardware components of the Cisco Unified Computing System (Cisco UCS) across multiple chassis, rack-mount servers, and thousands of virtual machines. Cisco UCS Manager manages Cisco UCS as a single entity through an intuitive GUI, a command-line interface (CLI), or an XML API for comprehensive access to all Cisco UCS Manager functions.

Cisco UCS VIC 1280 Adapters

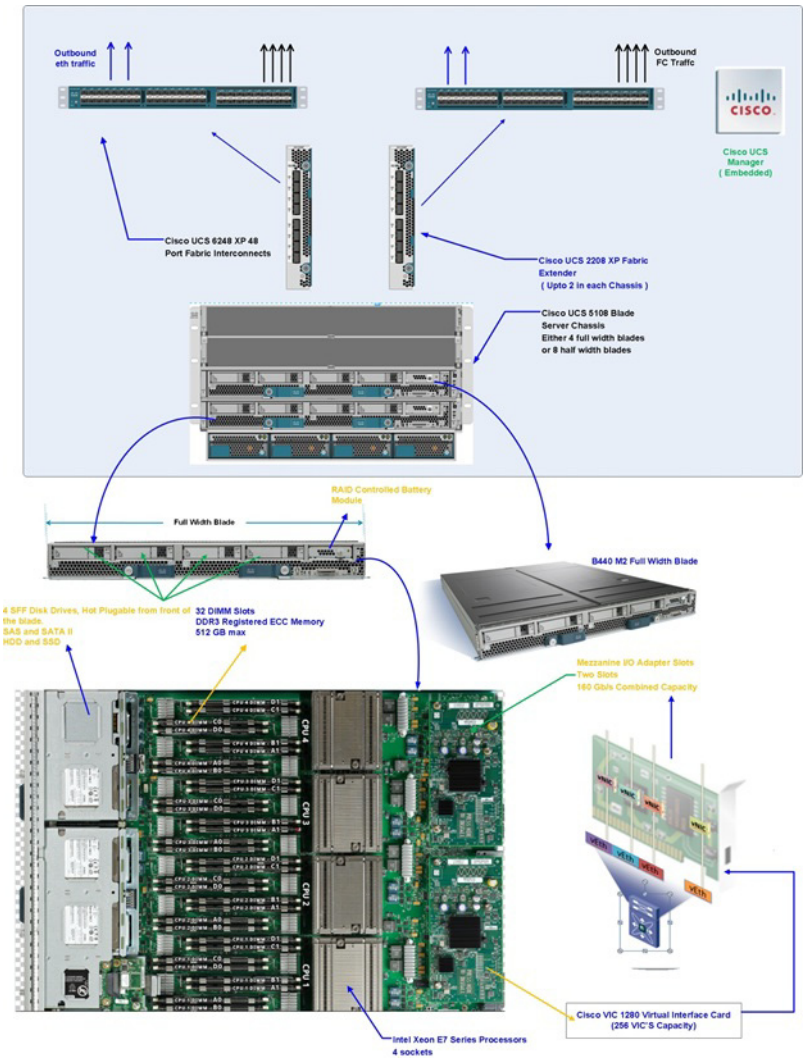
Cisco VIC 1280 is the second generation of Mezzanine Adapters from Cisco. VIC 1280 supports up to 256 PCI-e devices and up to 80 Gbps of throughput. Compared with its earlier generation of Palo Adapters it had doubled the capacity in throughput and PCI-e devices and is complaint with many OS and storage Vendors.

Cisco UCS B440 M2 High-Performance Blade Servers

The Cisco UCS B440 M2 High-Performance Blade Servers are full-slot, 4-socket, high-performance blade servers offering the performance and reliability of the Intel Xeon processor E7-4800 product family and up to 512 GB of memory. The Cisco UCS B440 supports four Small Form Factor (SFF) SAS and SSD drives and two converged network adapter (CNA) mezzanine slots up to 80 Gbps of I/O throughput. The Cisco UCS B440 blade server extends Cisco UCS by offering increased levels of performance, scalability, and reliability for mission-critical workloads.

The Cisco UCS components used in the certified configuration are shown in Figure 1.

Figure 1 Cisco Unified Computing System Components



Service Profiles: Cisco Unified Computing System Foundation Technology

Cisco UCS resources are abstract in the sense that their identity, I/O configuration, MAC addresses and worldwide names (WWNs), firmware versions, BIOS boot order, and network attributes (including quality of service (QoS) settings, pin groups, and threshold policies) are all programmable using a just-in-time deployment model. The manager stores this identity, connectivity, and configuration information in service profiles that reside on the Cisco UCS 6200 Series Fabric Interconnects. A service profile can be applied to any blade server to provision it with the characteristics required to support a specific software stack. A service profile allows server and network definitions to move within the management domain, enabling flexibility in the use of system resources. Service profile templates allow different classes of resources to be defined and applied to a number of resources, each with its own unique identities assigned from predetermined pools.

Service Profile

Description

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.

The 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.

Overview

A service profile typically includes four types of information:

- Server definition: It defines the resources (e.g. a specific server or a blade inserted to a specific chassis) that are required to apply to the profile.
- Identity information: Identity information includes the UUID, MAC address for each virtual NIC (vNIC), and WWN specifications for each HBA.
- Firmware revision specifications: These are used when a certain tested firmware revision is required to be installed or for some other reason a specific firmware is used.
- Connectivity definition: It is used to configure network adapters, fabric extenders, and parent interconnects, however this information is abstract as it does not include the details of how each network component is configured.

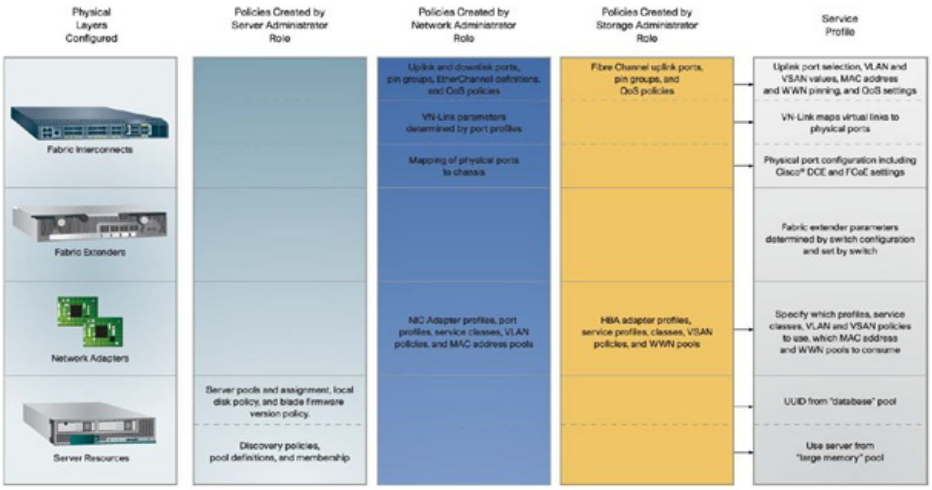
A service profile is created by the UCS server administrator. This service profile leverages configuration policies that were created by the server, network, and storage administrators. Server administrators can also create a Service profile template which can be later used to create Service profiles in an easier way. A service template can be derived from a service profile, with server and I/O interface identity information abstracted. Instead of specifying exact UUID, MAC address, and WWN values, a service template specifies where to get these values. For example, a service profile template might specify the standard network connectivity for a web server and the pool from which its interface's MAC addresses can be obtained. Service profile templates can be used to provision many servers with the same simplicity as creating a single one.

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.

Figure 2 below figure represents how Server, Network, and Storage Policies are encapsulated in a service profile.

Figure 2 Service Profile inclusions



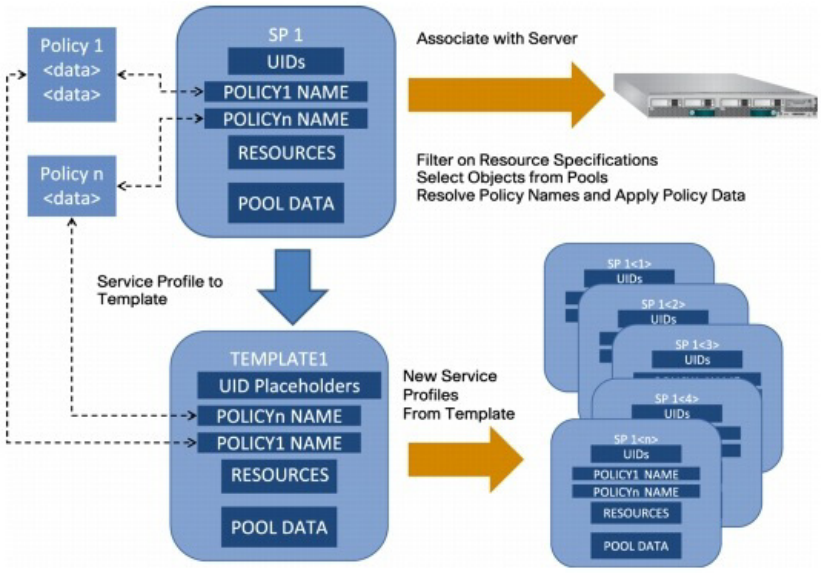
Understanding the Service Profile Template

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.

Figure 3 represents the relationship between the Service Profile and Templates.

Figure 3 Service Templates and Service profiles



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

- The Cisco UCS 6248 XP 48-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 a 16-port, 8-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 2208XP 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.
- The Cisco UCS 5108 Blade Server Chassis houses the fabric extenders, up to four power supplies, and up to four full width 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 used four (two in each chassis) Cisco UCS B440 M2 full width Blade Servers, each equipped with four 8-core Intel Xeon E7-4870 series processors. Each blade server was configured with 256 GB of memory.
- The blade server form factor supports a range of mezzanine-format Cisco UCS network adapters, including a 80 Gigabit Ethernet network adapter designed for efficiency and performance, the Cisco UCS VIC 1280 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 used Cisco UCS VIC 1280 Virtual Interface Network Adapters (2 adapters per blade) that provide 160 Gbps of performance per blade server.

Cisco Nexus 5548UP Switch

Figure 4 shows the Cisco Nexus 5548UP Switch

Figure 4 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, while providing market-leading simplicity and efficiency to minimize total cost of ownership.

Based on the powerful 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 enable 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

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). For additional information on this please refer link http://www.emc.com/collateral/hardware/white-papers/h8242-deploying-oracle-vnx-wp.pdf
- EMC PowerPath¬Æ: Provides automated data path management and load-balancing capabilities for heterogeneous server, network, and storage deployed in physical and virtual environments. For additional information refer: http://www.emc.com/collateral/software/data-sheet/l751-powerpath-ve-multipathing-ds.pdf.
- EMC Unisphere¬Æ: Delivers simplified management via a single management framework for all NAS, SAN, and replication needs. For additional information on Unisphere, refer: http://www.emc.com/collateral/software/data-sheet/h7303-unisphere-ds.pdf.

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

For details regarding EMC VNX Series Software Suites and the resulting value in performance, protection, and TCO that can be derived, please refer:

http://www.emc.com/collateral/software/data-sheet/h8509-vnx-software-suites-ds.pdf

Figure 5 EMC VNX Storage System

For additional detail regarding VNX Family Data Sheet to learn more about features available in the VNX product line enabling value in your Oracle deployment environment, please refer to http://www.emc.com/collateral/hardware/data-sheets/h8520-vnx-family-ds.pdf

Cisco Certified Configuration Inventory and 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.

Inventory of the Certified Configuration

The inventory of the components used in the certification stack is listed in Table 1.

Table 1 Inventory of the Certified Configuration

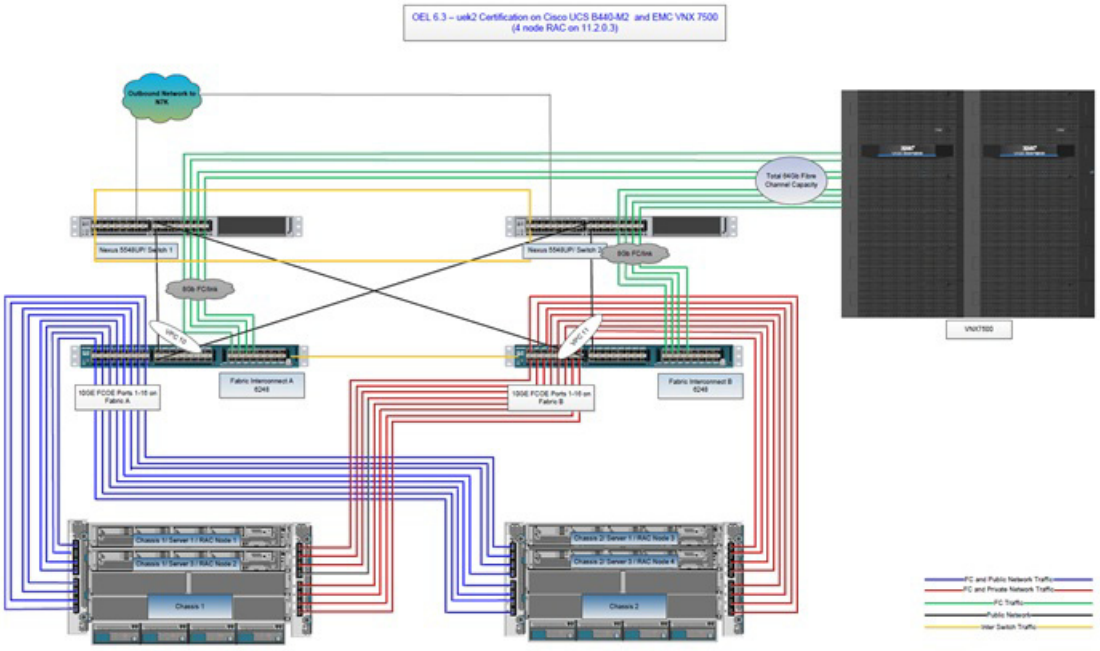
Physical Cisco Unified Computing System Server Configuration	
Description	Quantity
Cisco UCS 5108 Blade Server Chassis, with 4 power supply units, 8 fans and 2 fabric extenders	2
Cisco UCS B440-M2 full width blades	4
Four Socket - Eight Core Intel Xeon E7-4870 series 2.40 GHz processors	128
8 GB DDR3 DIMM, 1333 MHz (32 per server, totalling 256 GB per blade server)	128
Cisco UCS VIC 1280 Virtual Interface Card, 256 PCI devices, Dual 4 x 10G (2 per server)	8
Hard Disk Drives (SAN Boot Install)	0
Cisco UCS - 6248XP 48 port Fabric Interconnect	2
16 port 8 Gbps Fiber Channel expansion module	2

LAN and SAN Components	
Description	Quantity
LAN and SAN Components	
Cisco Nexus 5548 UP Switch	2
VLAN Configuration	VLAN ID
Public VLAN	134
Private Network VLAN (Private traffic VLAN must be configured on the Nexus switches to ensure traffic flow in partial link failure as discussed later	10
VSAN Configuration	VSAN ID
Oracle database VSAN	15

Storage Configuration	
Description	Quantity
EMC VNX 7500 Storage System	1
600 GB 15k RPM, SAS disk drivers	290
200 GB flash drives	25

Operating System and RPM Components (installed on all Oracle nodes)	OS and RPMs
Operating System (64 bit)	Oracle Linux 6.3 x86_64(2.6.39-200.24.1.el6uek.x86_64)
Required RPMs by EMC (to be installed on all Cluster nodes to support EMC Power Path and Host agent)	EMCpower.LINUX-5.7.1.00.00-029.ol6_uek2_r2.x86_64 HostAgent-Linux-64-x86-en_US-1.0.0.1.0474-1.x86_64

Figure 6 Oracle Database 11gR2 with Real Application Clusters technology on Cisco Unified Computing System and EMC VNX Storage



In Figure 6, the blue lines indicate the public network connecting to Fabric Interconnect A, and the red 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 (blue and red lines). The Fiber channel traffic leaves the UCS Fabrics through their own N5k Switches to EMC (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.

Configuring Cisco Unified Computing System for the 4 node Oracle RAC

Detailed information about configuring the Cisco Unified Computing System is available at http://www.cisco.com/en/US/products/ps10281/products_installation_and_configuration_guides_list.html.

It is beyond the scope of this document to cover all of these. However an attempt is made to include as many and as much as possible.

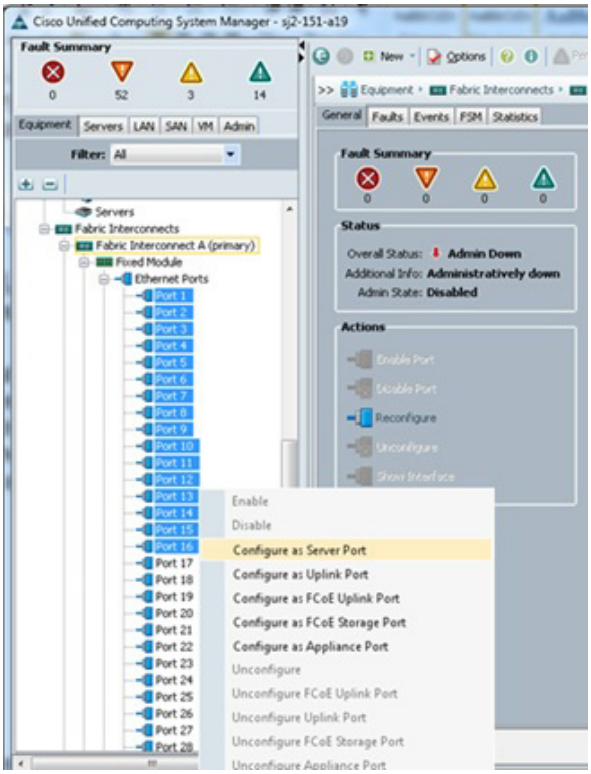
Configuring Fabric Interconnects

Two Cisco UCS 6248 UP Fabric Interconnects are configured for redundancy. It provides resiliency in case of failures.

The first step is to establish connectivity between the blades and fabric interconnects. As shown in Figure 7, sixteen public (eight per chassis) links go to Fabric Interconnect "A" (ports 1 through 16). Similarly, sixteen private links go to Fabric Interconnect B. It is recommended 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.

Configure Server Ports

Figure 7 Configuring Server ports

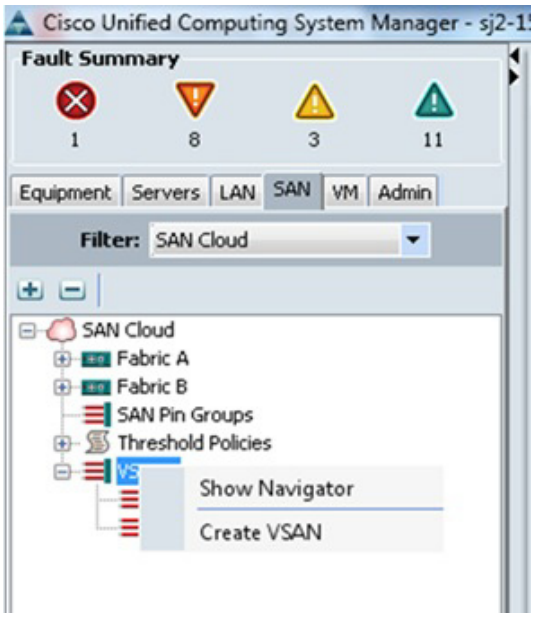


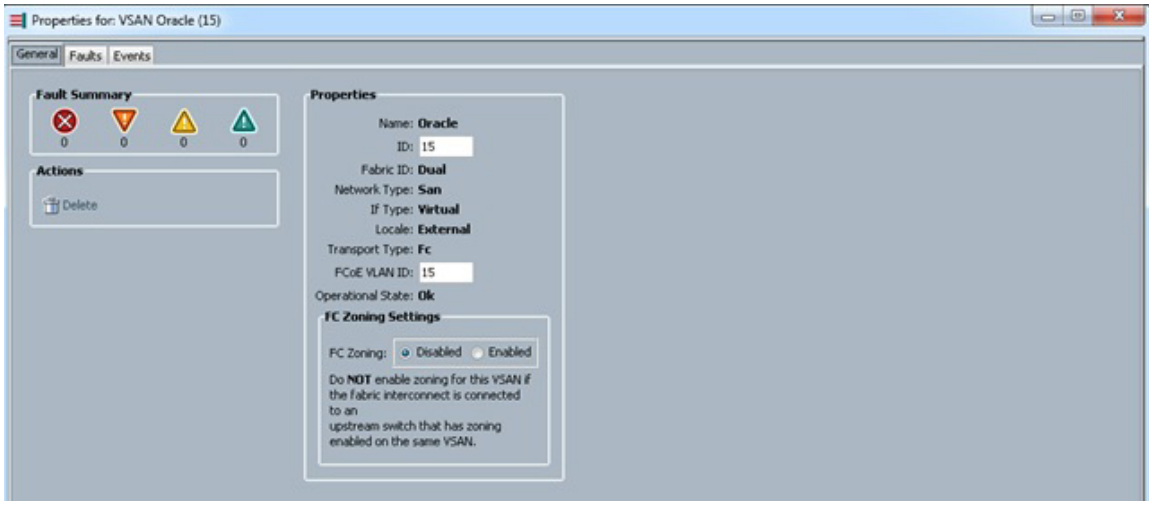
Configuring SAN and LAN on the Cisco UCS Manager

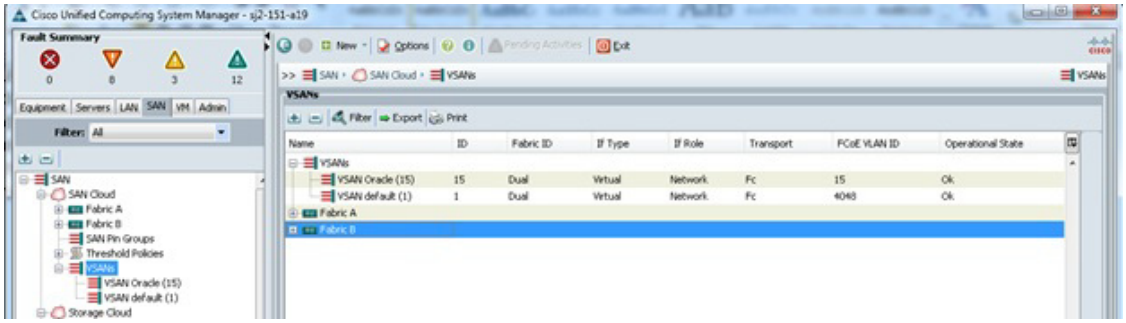
Configure SAN

On the SAN tab, create and configure the VSANs to be used for database as shown in Figure 8. On the test bed, we used vSAN 15 for database.

Figure 8 Configuring SAN on the Cisco UCS Manager



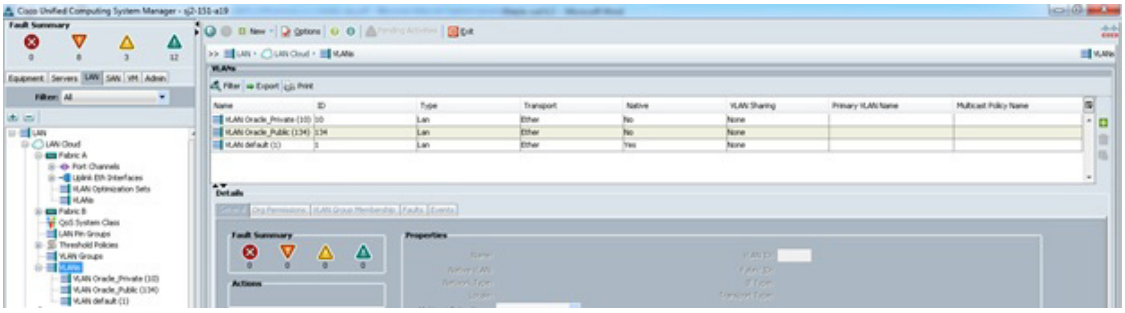




Configure LAN

On the LAN tab as shown in Figure 9, create VLANs, that will be used later for virtual NICs (vNICS) configured for private and public traffic for Oracle RAC. 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 interface. 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.

Figure 9 Configuring LAN on the Cisco UCS Manager

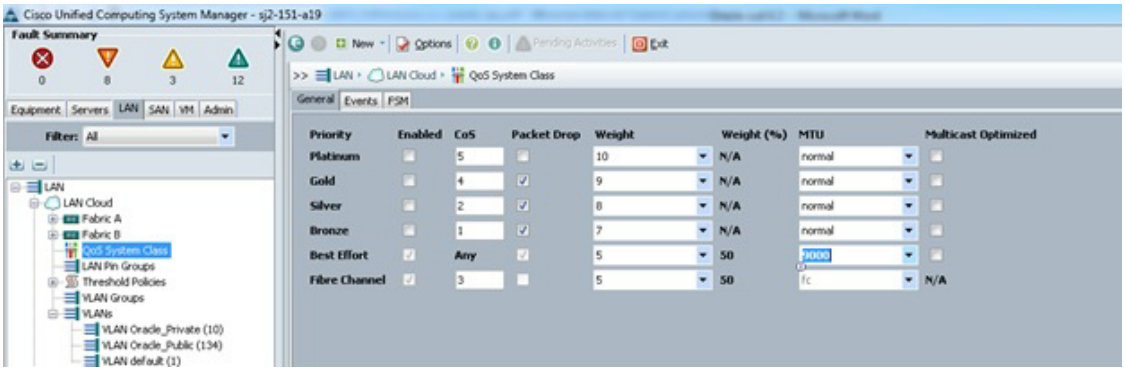


Even though private VLAN traffic stays local within Cisco UCS domain, it is necessary to configure entries for these private VLANS in northbound network switch. This will allow the switch to route interconnect traffic appropriately in case of partial link or IOM failures.

Configure Jumbo Frames

Enable Jumbo Frames for Oracle Private Interconnect traffic.

Figure 10 Configuring Jumbo Frames



After these initial setups we can setup UCS service profile templates for the hardware configuration.

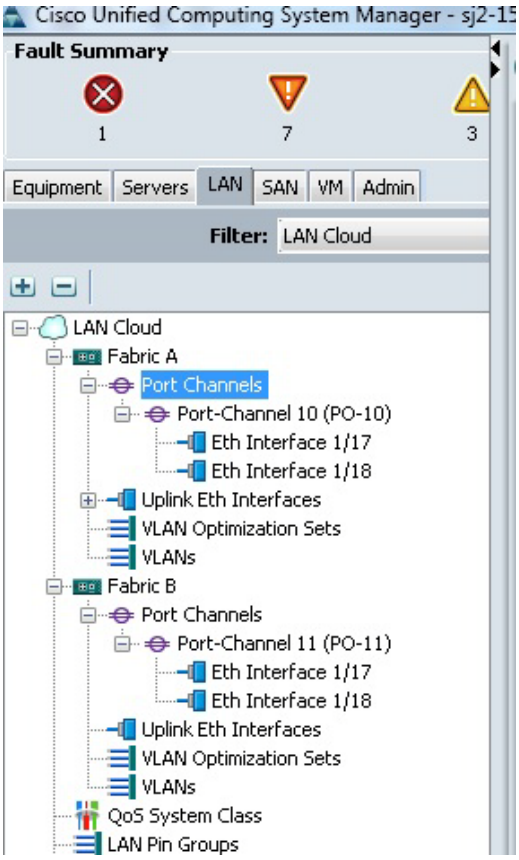
Configure Ethernet Port-Channels

For configuring Port-Channels, login to Cisco UCS Manager and LAN tab, filter on LAN cloud as shown in Figure 11.

Select Fabric A, right click on port-channels and create port-channel. In the current Oracle RAC setup ports 17 and 18 on Fabric A were selected to be configured as port channel 10.

Similarly ports 17 and 18 on Fabric B were selected as port channel 11.

Figure 11 Configuring Port Channels



Port Channel 10 Details

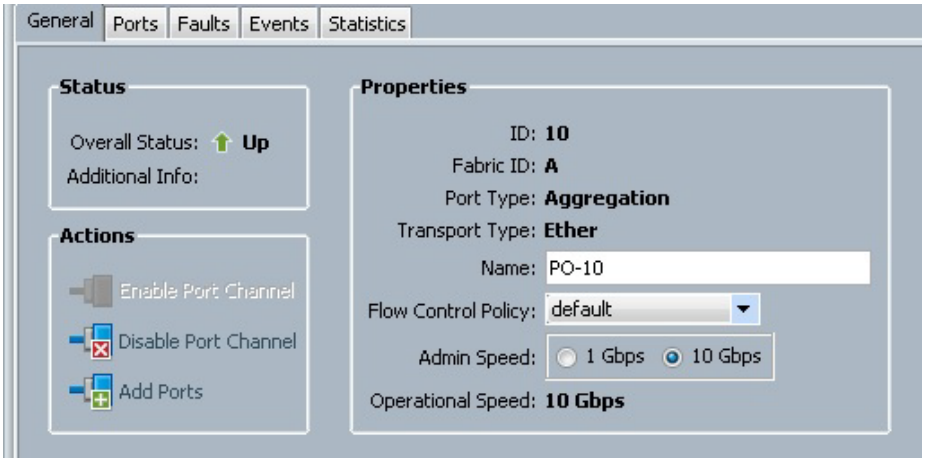


Figure 12 Port Channels on Fabric A

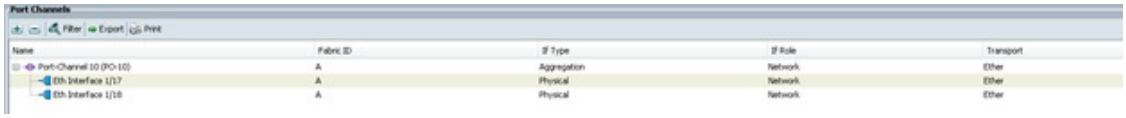
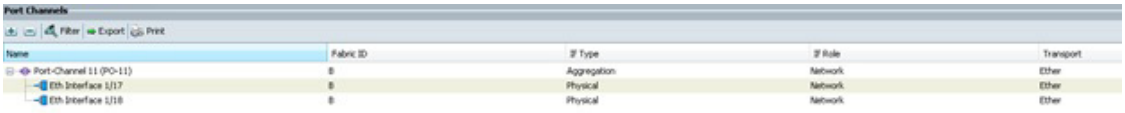


Figure 13 Port Channels on Fabric B



The next step is to set up VPC in n5k. This is covered in the n5k section.

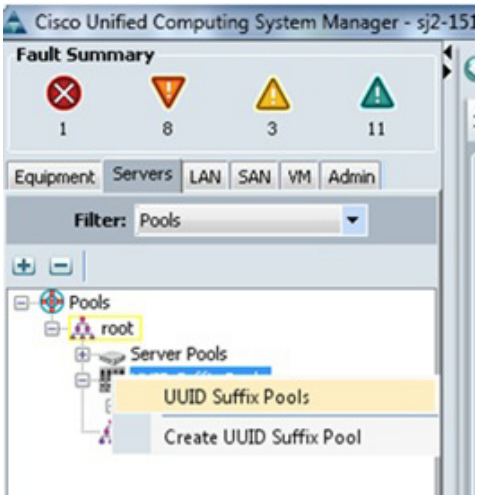
Preparatory Steps Before Creating Service Templates

First create the UUID, IP, MAC, WWNN and WWPN pools and keep them handy in case they are not pre-created. If already pre-created make sure that you have enough of them free and unallocated.

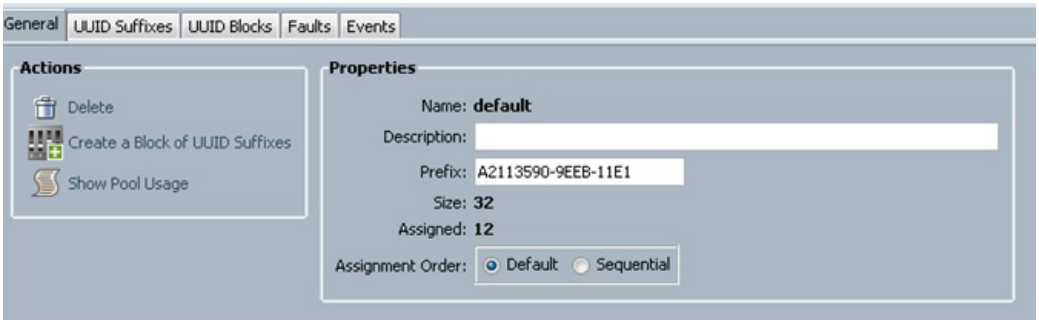
UUID Pool

To create the UUID, do the following steps:

1. Click Servers tab.
2. Filter on pools.
3. Expand UUID suffix pools and create a new pool.



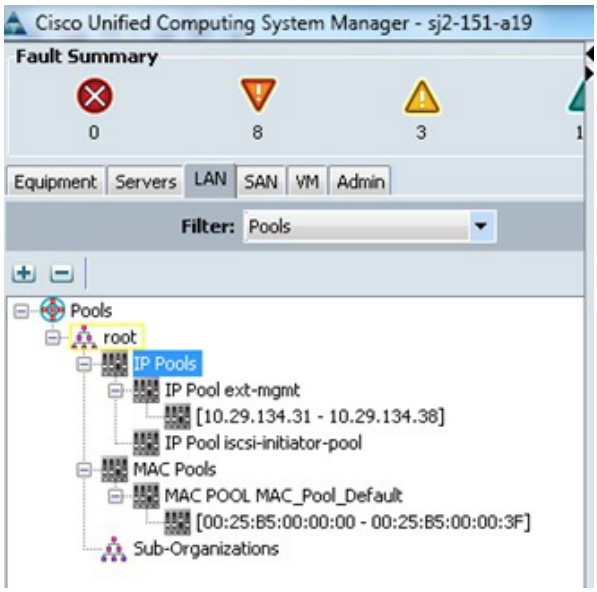
Note You may create a default pool as shown below.



IP and MAC Pools

To create IP and MAC pools, do the following steps:

1. Click the LAN tab.
2. Filter on pools.
3. Create IP and MAC pools.

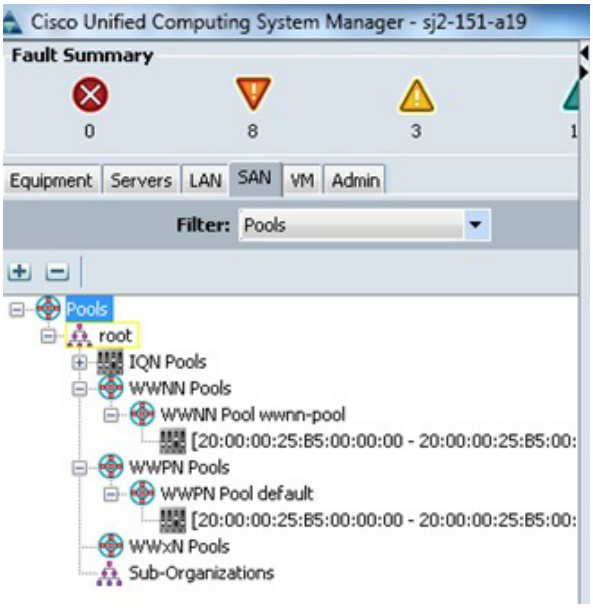


The IP pools will be used for console management, while the MAC addresses for the vNICs will be created later in the process.

WWNN and WWPN Pools

To create WWNN and WWPN pools, do the following step:

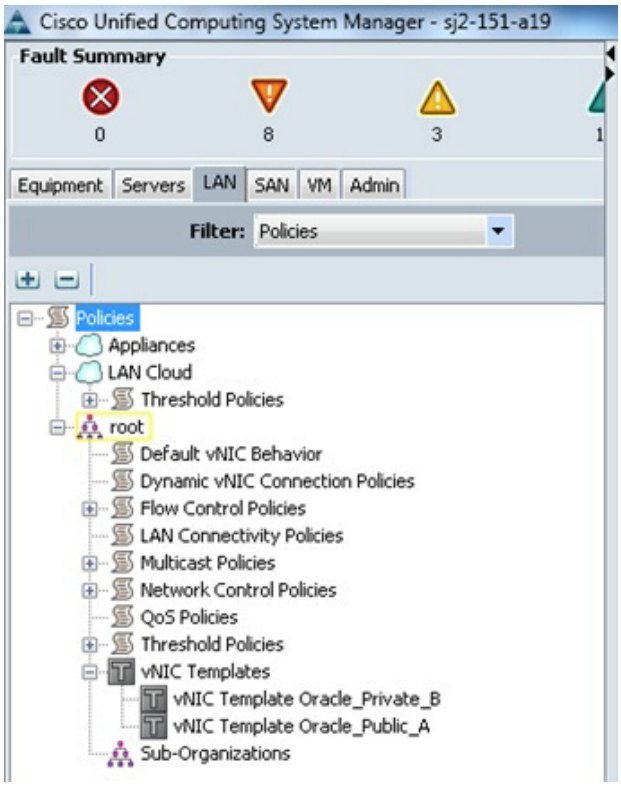
1. Click on SAN tab filter on pools and create the pools.



Configure vNIC Templates

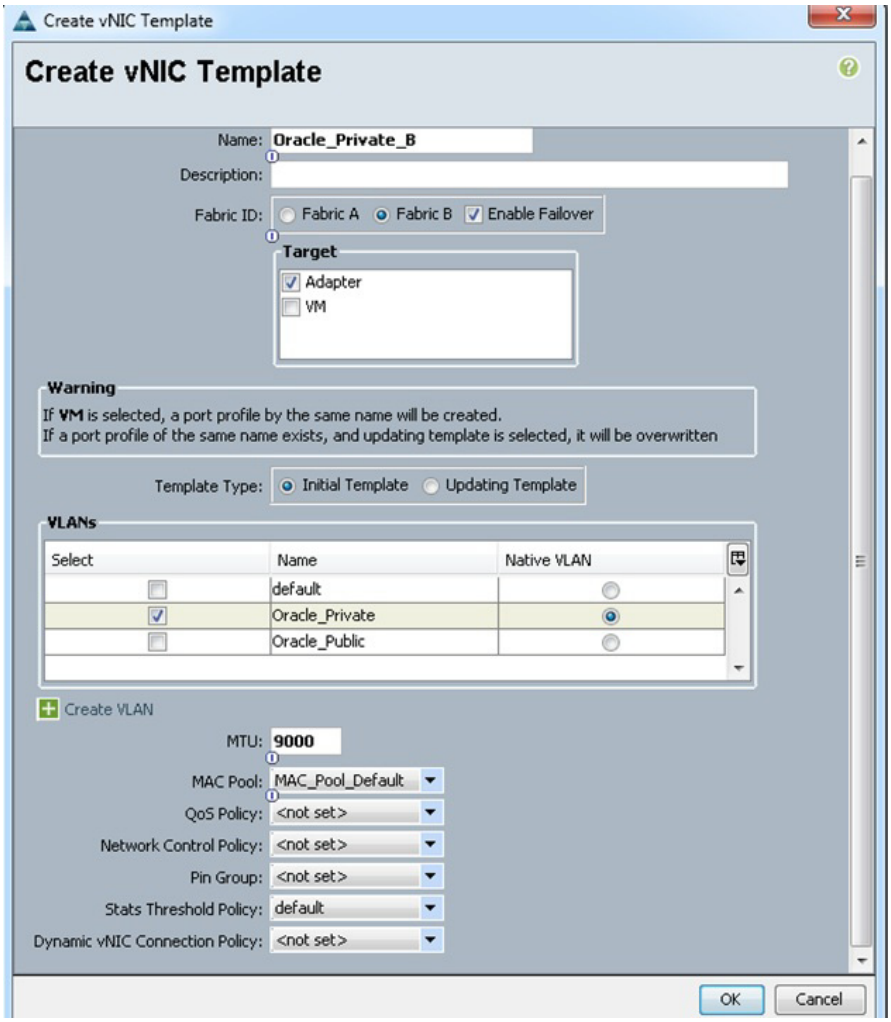
To configure the vNIC template, do the following steps:

1. Click the LAN tab.
2. Filter on policies and select vNIC templates. Two templates are created; one for Public network and one for Private network. The Private network is for the internal Heart Beat and message transfers between Oracle Nodes while Public network for external clients like middle tiers and ssh sessions to the Oracle database hosts.

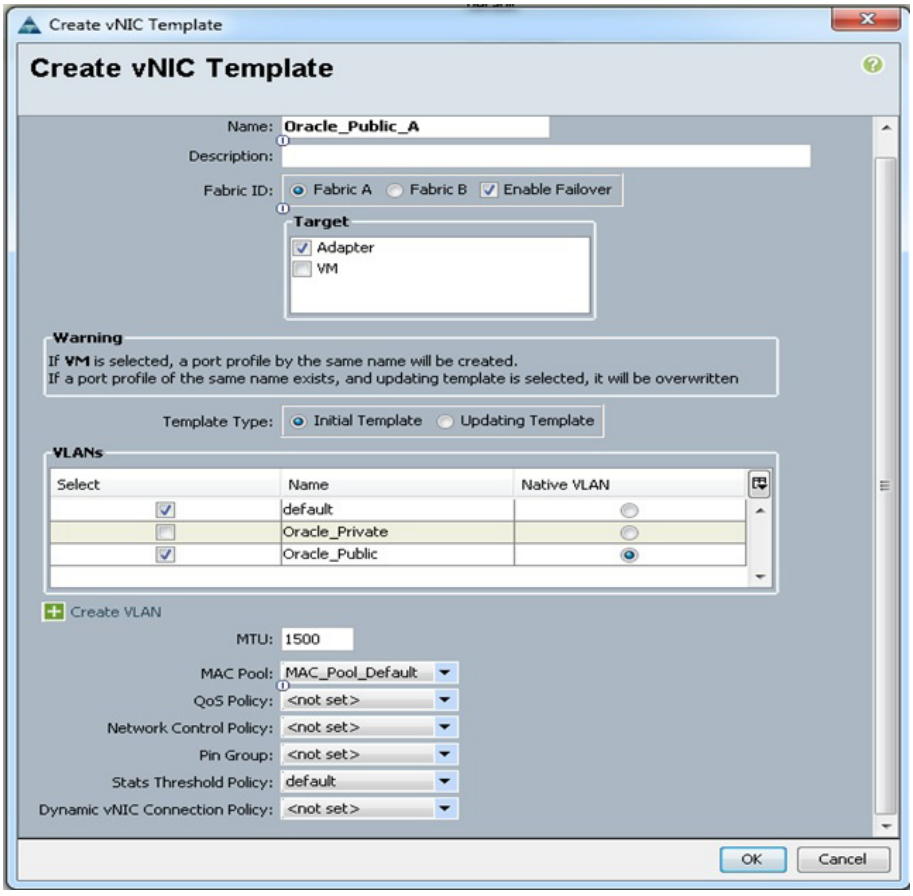


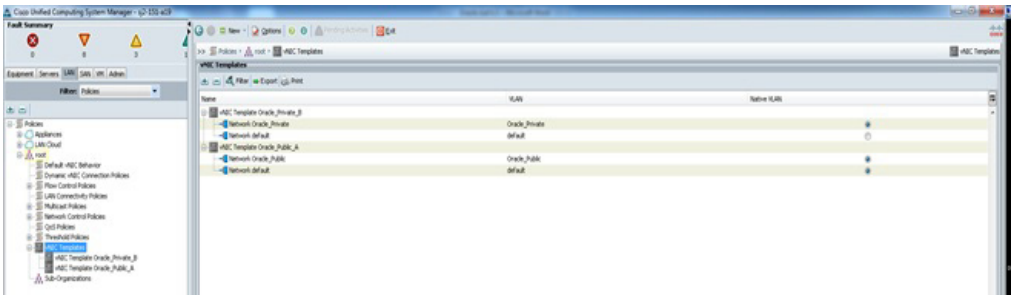
The vNIC template for Oracle Private link is set at 9000 MTU and pinned to Fabric B. However, the failover is enabled. This allows the vNIC to failover to Fabric A, in case of failures of Fabric B.

Create a Private vNIC Template



Create a Public vNIC Template

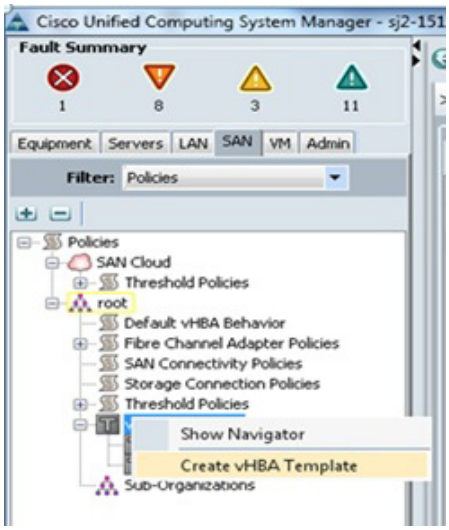


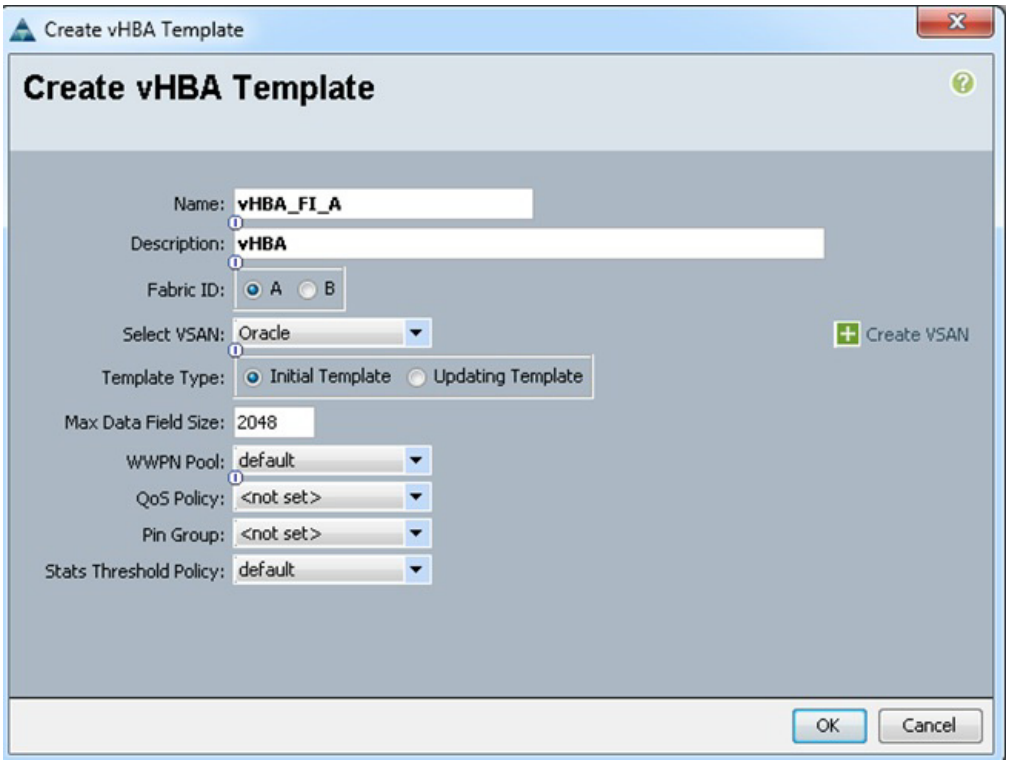


Create a HBA Template

To create a HBA template, do the following steps:

1. Click on the SAN tab.
2. Filter out policies, right-click the vHBA templates and create a template.



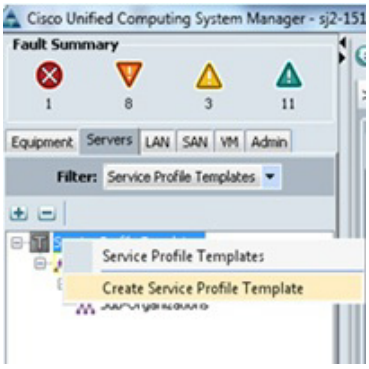


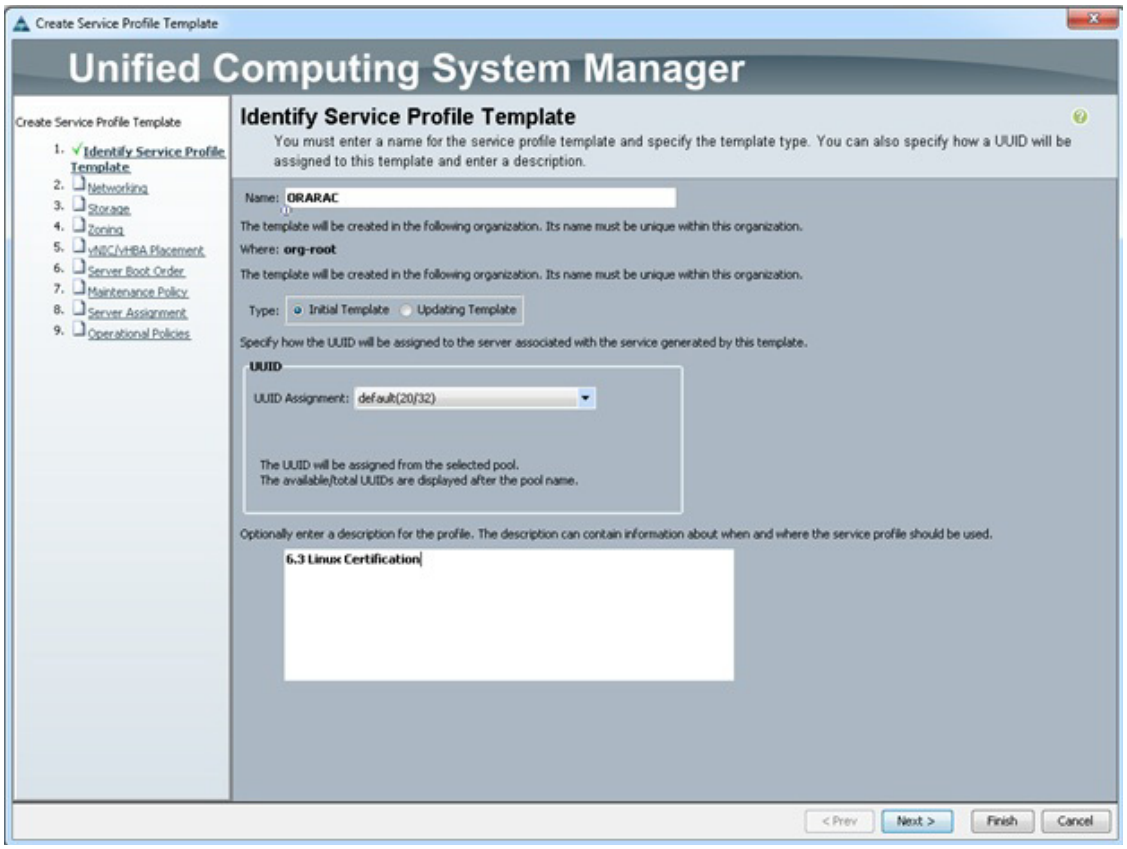
When the preparatory steps are complete, create a service template for the service profiles.

Create Service Profile Template

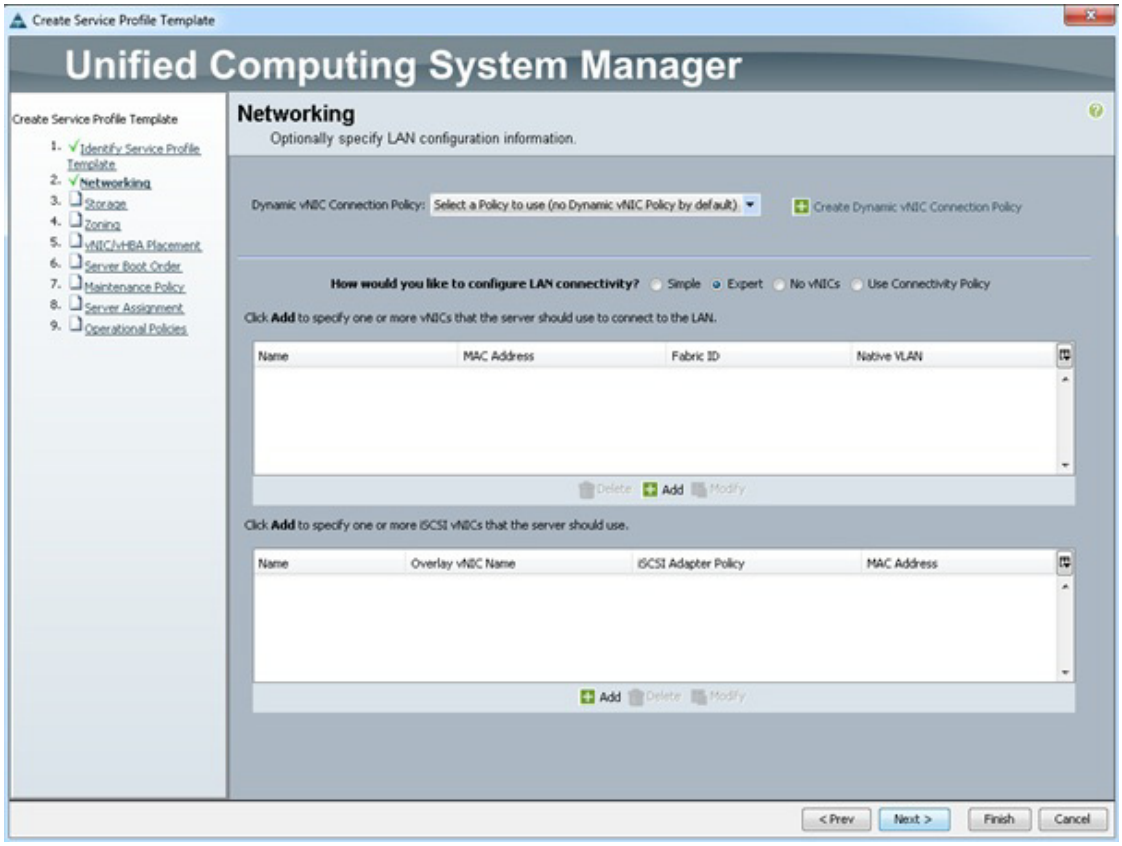
Create a service profile template before forking the service profiles that will be allocated to the servers.

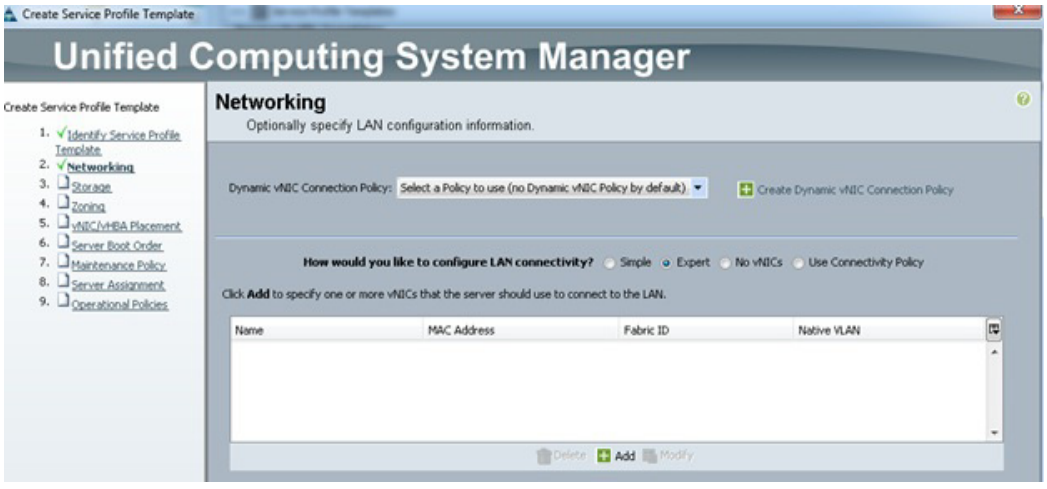
1. Click the Servers tab in the Cisco UCS Manager.
2. Filter out the Service Profile Templates and select Create Service Profile Template.



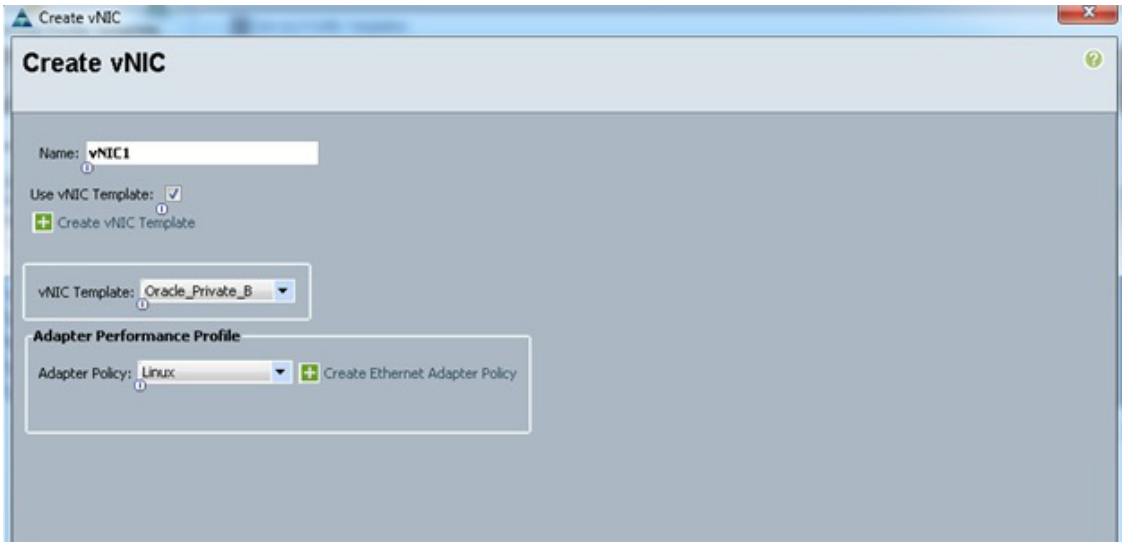


3. Enter the name, select the default UUID created earlier and click Next.
4. In the Networking page create one nNIC on each fabric and associate them with the VLAN policies created earlier.
5. Select Expert mode and click Add to connect to the LAN.

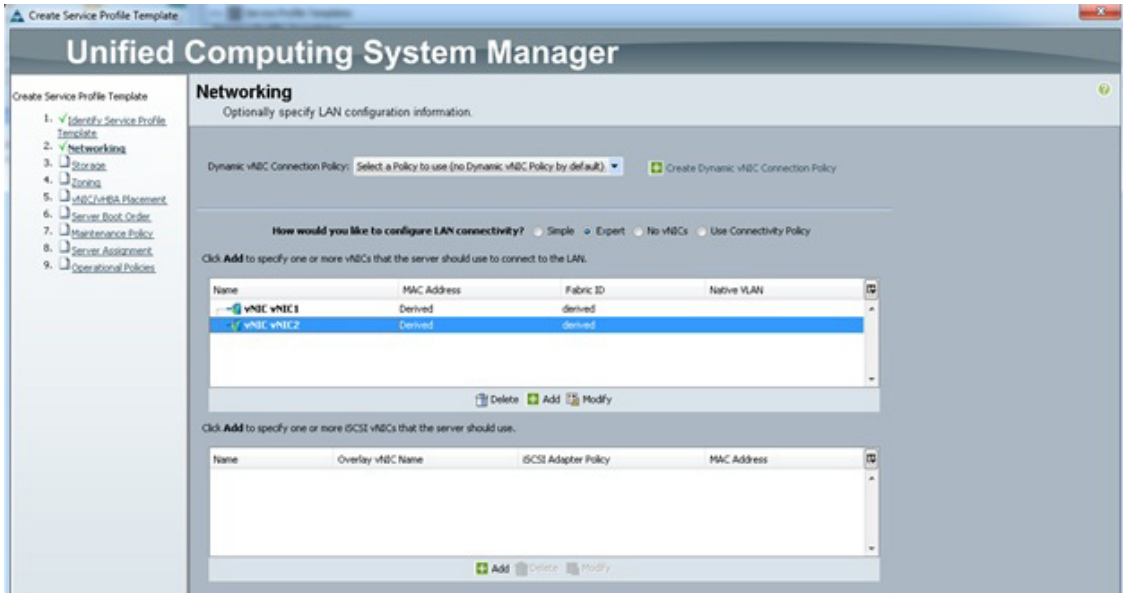

Create Service Profile Template
Unified Computing System Manager
Create Service Profile Template
1. Identify Service Profile Template
2. Networking
3. Storage
4. Zoning
5. vNIC/vHBA Placement
6. Server Boot Order
7. Maintenance Policy
8. Server Assignment
9. Operational Policies
Networking
Optionally specify LAN configuration information.
Dynamic vNIC Connection Policy: Select a Policy to use (no Dynamic vNIC Policy by default)
Create Dynamic vNIC Connection Policy
How would you like to configure LAN connectivity? Simple Expert No vNICs Use Connectivity Policy
Click Add to specify one or more vNICs that the server should use to connect to the LAN.
Name
MAC Address
Fabric ID
Native VLAN
Delete
Add
Modify
Click Add to specify one or more iSCSI vNICs that the server should use.
Name
Overlay vNIC Name
iSCSI Adapter Policy
MAC Address
Add
Delete
Modify
< Prev
Next >
Finish
Cancel


Create Service Profile Template
Unified Computing System Manager
Create Service Profile Template
1. Identify Service Profile Template
2. Networking
3. Storage
4. Zoning
5. vNIC/vHBA Placement
6. Server Boot Order
7. Maintenance Policy
8. Server Assignment
9. Operational Policies
Networking
Optionally specify LAN configuration information.
Dynamic vNIC Connection Policy: Select a Policy to use (no Dynamic vNIC Policy by default)
Create Dynamic vNIC Connection Policy
How would you like to configure LAN connectivity? Simple Expert No vNICs Use Connectivity Policy
Click Add to specify one or more vNICs that the server should use to connect to the LAN.
Name
MAC Address
Fabric ID
Native VLAN
Delete
Add
Modify

6. In the create vNIC page, select Use vNIC template and adapter policy as Linux. We selected vNIC1 for the Oracle private network.



Create vNIC2 for Public



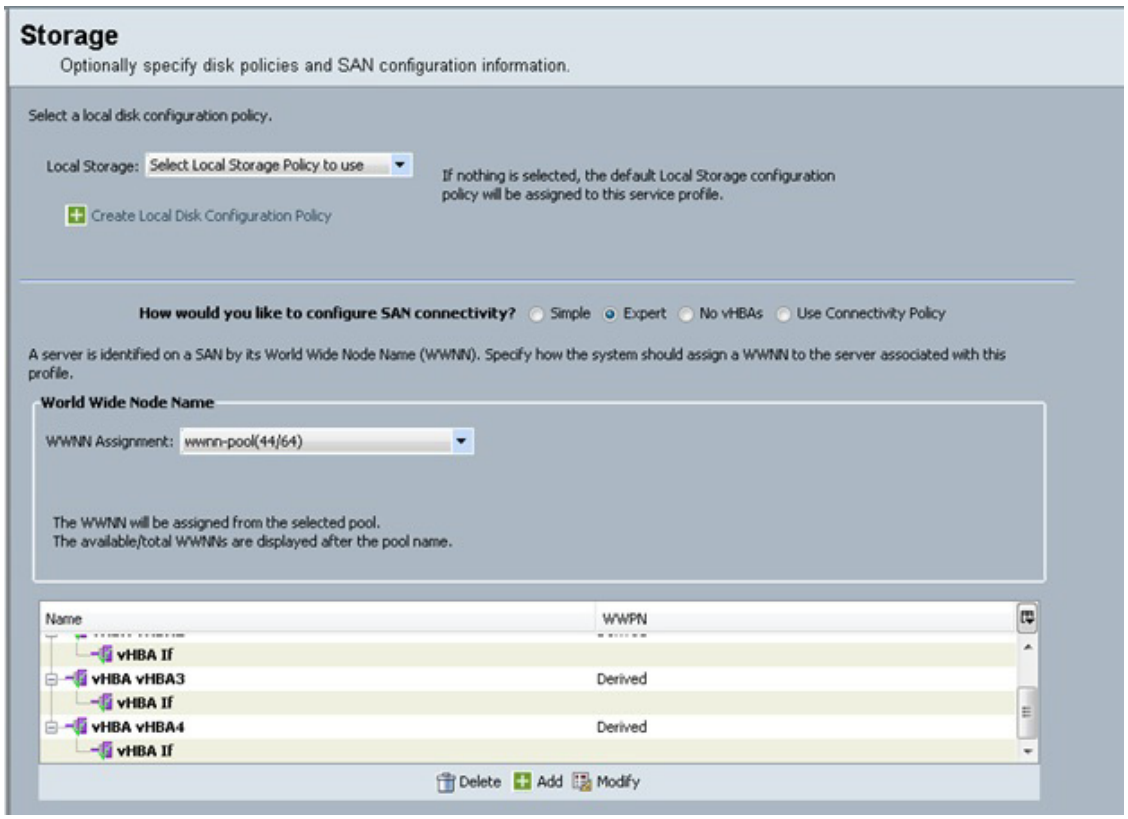
1. In the Storage page, select Expert mode in adapter.
2. Choose the WWNN pool created earlier and click Add to create vHBA's. We select 4xvHBA's that are shown below.

Create vHBA1 using template vHBA_FI_A.

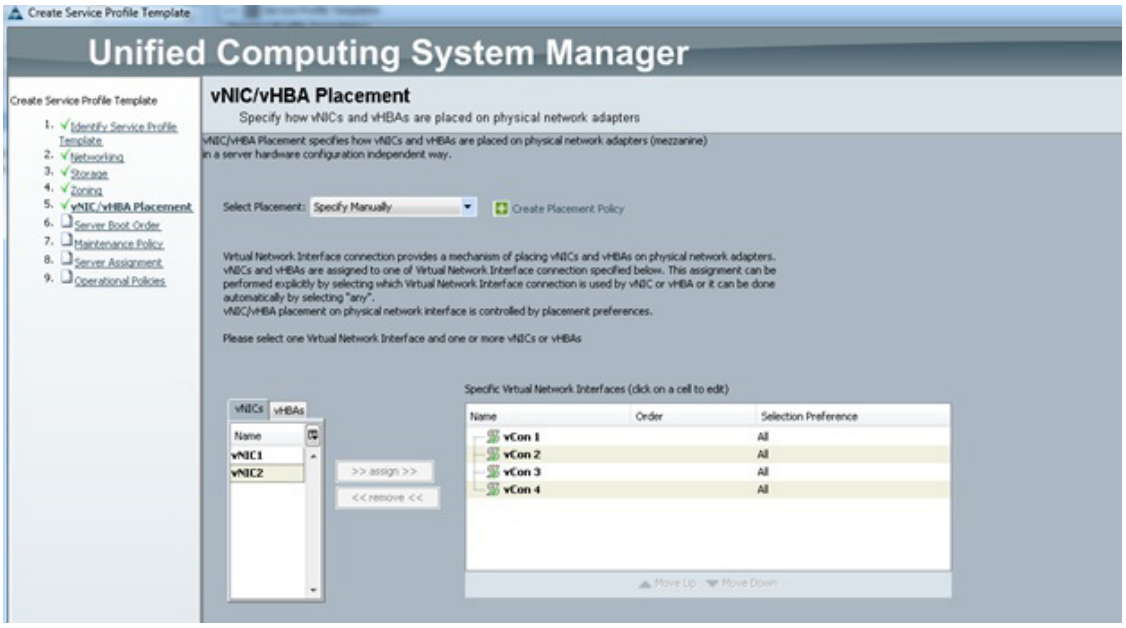
Create vHBA2 using template vHBA_FI_B.

Create vHBA3 using template vHBA_FI_A.

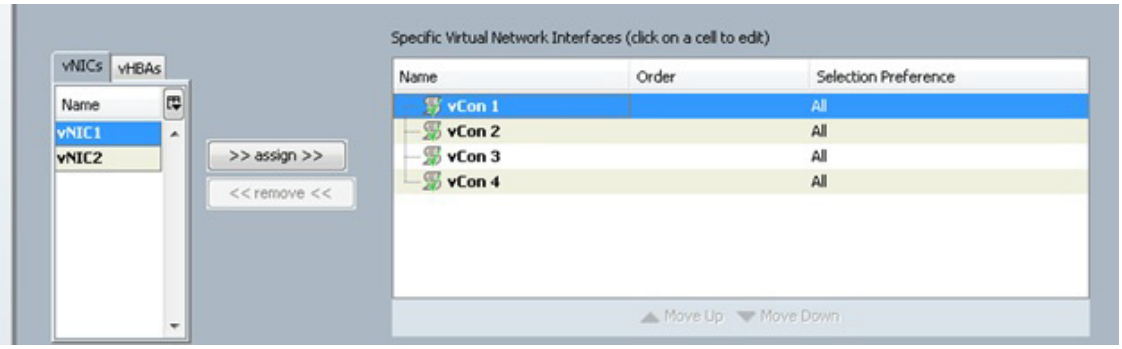
Create vHBA4 using template vHBA_FI_B.



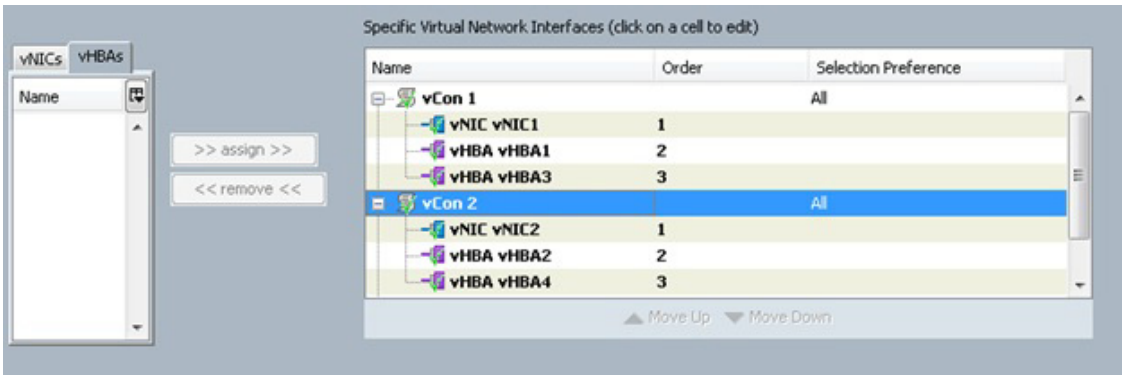
3. Skip the zoning section and go to vNIC/vHBA placement.
4. In the next screen select Manually.



5. Select vNICS, highlight vNICs on the left side and vCONs on the right and click Assign. The vNICs will be placed on the chosen adapter vCONs.



6. Assign vNIC2 on vCon2. Repeat the above procedure for vHBAs.



We allocated vNIC1, vHBA1 and vHBA3 to the first vic1280, with the rest of vNIC2, vHBA2 and vHBA4 to the second.

Server Boot Policy

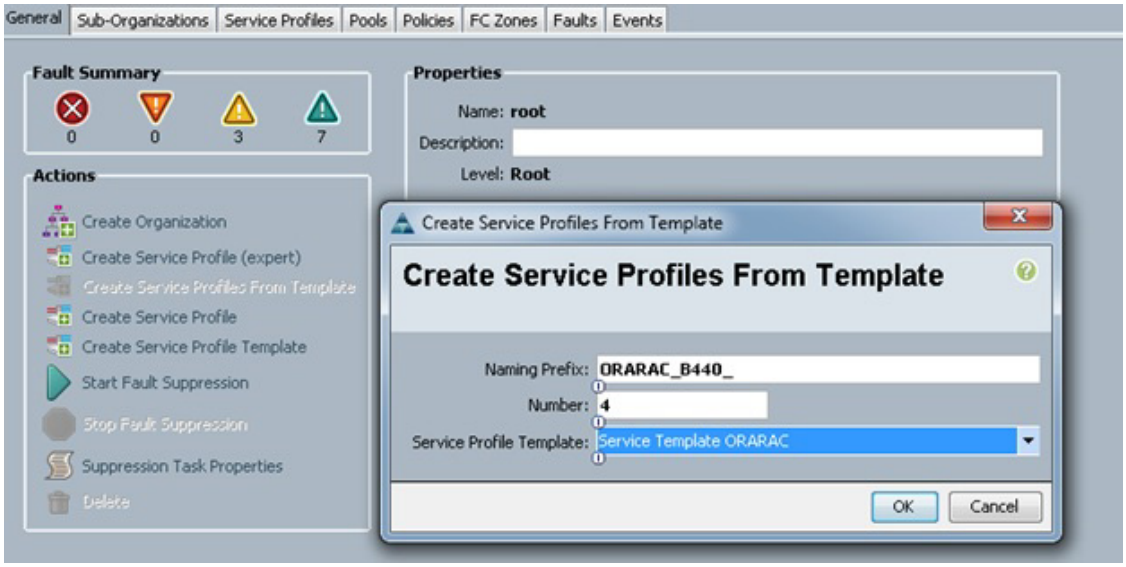
Leave this as Default since the initiators may vary from one server to the other.

The rest of the maintenance and assignment policies were left as Default in the test. But they may be selected and may vary from site to site, depending on your workloads, best practices and policies.

Create Service Profiles from Service Profile Templates

To create service profiles from templates, do the following steps:

1. Click the Servers tab.
2. Right-click the root and select Create Service Profile from Templates.

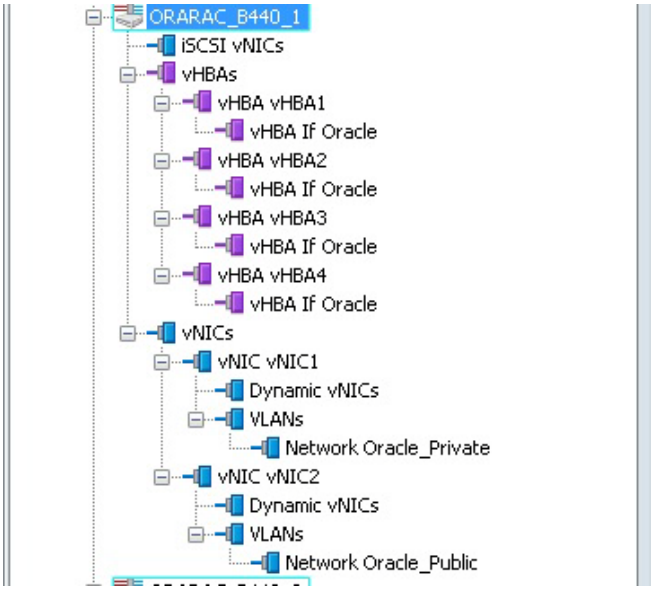


This will create 4 service profiles with the name prefix as shown below:

ORARAC_B440_1, ORARAC_B440_2, ORARAC_B440_3, ORARAC_B440_4

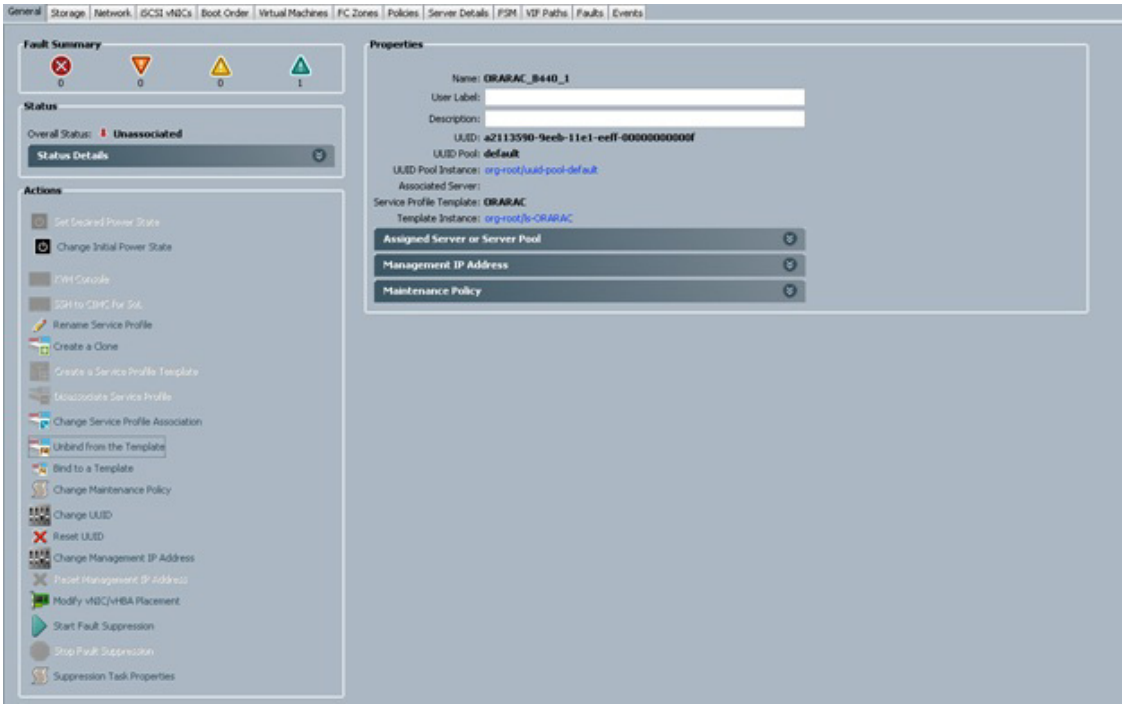
Associating Service Profile to the Servers

Make sure that, a few of the entries in the service profile appear as shown below before associating them to a server.



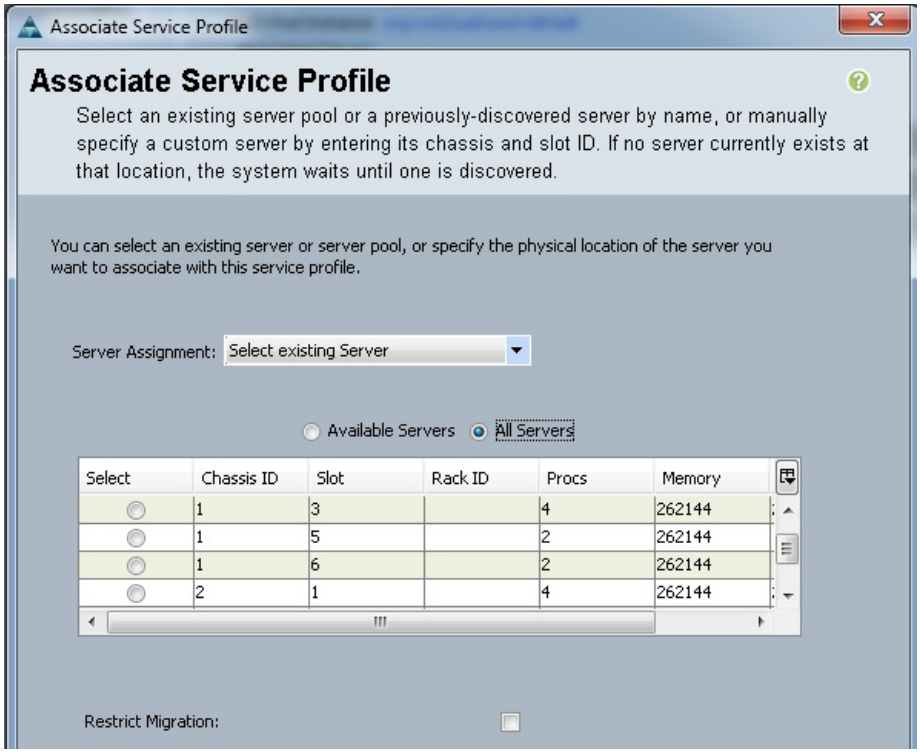
In order to associate this service profile to a server, perform the following steps:

1. Under the servers tab, select the desired service profile, and select change service profile association



Now the service profile is unassociated and can be assigned to a server in the pool.

1. Click Change Service Profile Association.
2. From the Server Assignment drop-down list, select the existing server that you would like to assign, and click OK.



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 Table 2 for recommendations about distributing database files based on the type of workload.

While it is out of scope to cover all the aspects of VNX storage here, a brief overview is given below. Two databases were created, one for Online transactions processing (OLTP) and another for a Decision support system (DSS).

Storage Pool for OLTP was created with a mix of SAS and Flash drives while RAID groups with SAS disks were created for the DSS system. The redo logs for both the databases were created from the same RAID group.

The following table illustrates the distribution of Luns carved out from a VNX7500 for the setup.

Storage Configuration

Table 2 lists the storage configuration

Table 2 Storage Configuration

Purpose	OLTP Database data and temp files	DSS Database data and temp files	Redo Log Files for OLTP and DSS databases
Disk Type	Mixed (SAS and Flash)	SAS	SAS
RAID Type	RAID 5 Storage Pool	RAID 5 RAID Groups	RAID 1/0 RAID Groups
SAS Disks	80	300	32
Flash Disks	10	0	0
Total Luns	16	32	8
Lun Size	600GB	200GB	100GB

Purpose	Boot Luns and Oracle RAC OCR/Voting Luns
Disk Type	SAS
RAID Type	RAID 5
SAS Disks	5
Flash Disks	0
Total Luns	4 Boot Luns and 5 RAC Luns
Lun Size	Boot Luns - 100GB RAC Luns - 20GB

Hardware Storage Processors Configuration

A total of eight ports were used from storage processors and were equally distributed between SPA and SPB as shown in Table 3 and were connected to the respective N5K's.

Table 3 Service Processor Distribution

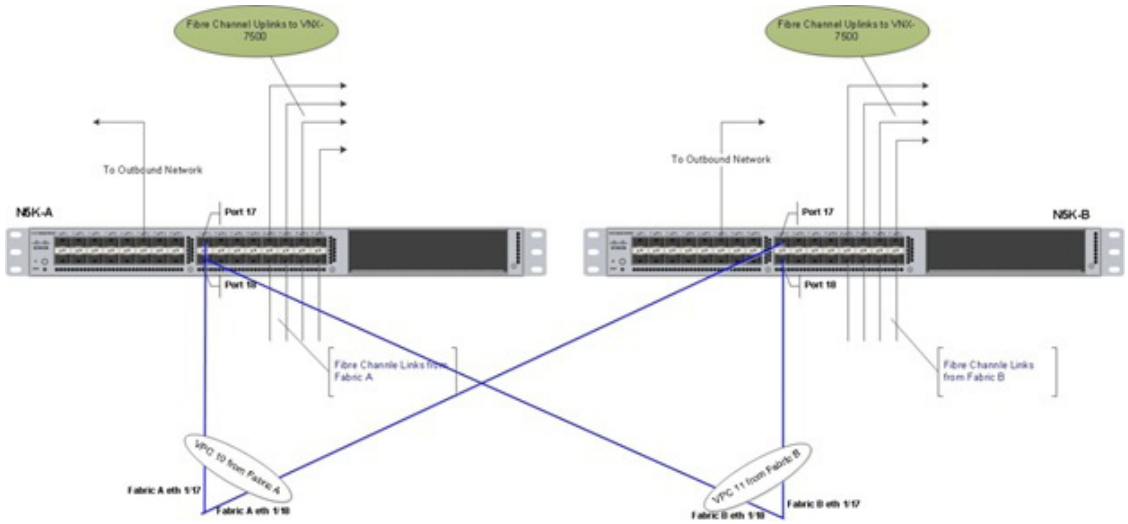
Processor	Slot/Port	WWPN
SPA	A2	5006:0160:3de0:21f6
	A3	5006:0161:3de0:21f6
SPB	B2	5006:0168:3de0:21f6
	B3	5006:0169:3de0:21f6

In the later sections of N5K zoning, we will cover how these WWPN' s will be used in zoning, boot policies, and in achieving high availability in case of failures.

Configure SAN zoning on N5K 5548 UP Switches

Two N5K 5548 UP switches were configured.

Figure 14 N5K Configuration with vPC



The above figure diagrammatically represents how the N5K UP switches are connected to North bound switches and storage while connected to the underlying Cisco UCS Fabrics. The N5K switches form a core group in controlling SAN zoning.

Fibre Channel Zoning

Before going to the zoning details, decide how many paths are needed for each LUN and extract the WWPN numbers for each of the HBA's.

For details about the WWPN's for each of the HBA's, login to the Cisco UCS Manager.

1. Click Equipment, chassis, servers and the desired server. On the right hand menu, click the Inventory tab and HBA's, sub-tab.

The WWPN numbers for all four HBA's for server 1, as an example, is illustrated above. In the current setup, it was decided to have a total of 8 paths, 4 paths from each Fabrics and N5K's to the storage.

The zoning for Server1, HBA1 is setup as follows:

* fcid 0x3a01ef [pwwn 50:06:01:62:47:20:2c:af] [A2P2]

* fcid 0x3a02ef [pwwn 50:06:01:68:47:20:2c:af] [B2P0]

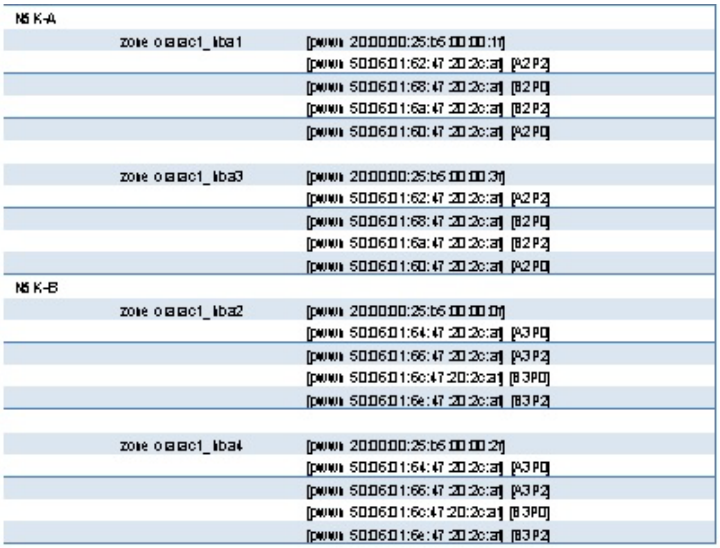
* fcid 0x3a03ef [pwwn 50:06:01:6a:47:20:2c:af] [B2P2]

* fcid 0x3a00ef [pwwn 50:06:01:60:47:20:2c:af] [A2P0]

* fcid 0x3a0022 [pwwn 20:00:00:25:b5:00:00:1f] < Extracted from the above figure for HBA1.

The WWPN's from storage are distributed between both storage processors providing distribution and redundancy in case of a failure.

Table 4 Example for Server 1



2. Login through ssh and issue the following:

The following is an example for one zone on one N5K:

```
conf term
zoneset name ORARAC_FI_A vsan 15
zone name orarac1_hba1
member device-alias A2P0
member device-alias A2P2
member device-alias B2P0
member device-alias B2P0
member pwwn 20:00:00:25:b5:00:00:1f ( orarac1 hba1 wwpn )
exit
exit
zoneset activate name ORARAC_FI_A vsan 15
copy running-config startup-config
```

Repeat the steps for the HBA's. A detailed list of zones added in the setup is provided in the Appendix.

Setup VLAN and VSAN on Both N5K's

```
conf term
vlan 134
  name Oracle_RAC_Public_Traffic
exit
vlan 10
  name Oracle_RAC_Private_Traffic
  no ip igmp snooping
exit
vsan database
vsan 15
exit
```

Setting Up Device Aliases for Storage Initiators

```
device-alias database
  device-alias name A2P0 pwwn 50:06:01:60:47:20:2c:af
  device-alias name A2P2 pwwn 50:06:01:62:47:20:2c:af
  device-alias name A3P0 pwwn 50:06:01:64:47:20:2c:af
  device-alias name A3P2 pwwn 50:06:01:66:47:20:2c:af
  device-alias name B2P0 pwwn 50:06:01:68:47:20:2c:af
  device-alias name B2P2 pwwn 50:06:01:6a:47:20:2c:af
  device-alias name B3P0 pwwn 50:06:01:6c:47:20:2c:af
  device-alias name B3P2 pwwn 50:06:01:6e:47:20:2c:af

device-alias commit
exit
```

Setting Up VPC on N5K‚Äôs

From Figure 14, both N5K's port 17 receives traffic from UCS Fabric A, that has port-channel 10 defined. Similarly both N5K's port 18 receives traffic from UCS Fabric B, that has port-channel 11 configured.

Login into N5K-A as admin.

```
conf term
feature vpc

vpc domain 1
peer-keepalive destination  <IP Address of peer-N5K>
exit

interface port-channel  1
  switchport mode trunk
  vpc peer-link
  switchport trunk allowed vlan 1,10,134
  spanning-tree port type network
exit

interface port-channel 10
  description Oracle RAC port-channel
  switchport mode trunk
  vpc 10
  switchport trunk allowed vlan 1,10,134
  spanning-tree port type edge trunk
exit

interface port-channel 11
  description Oracle RAC port-channel
  switchport mode trunk
  vpc 11
  switchport trunk allowed vlan 1,10,134
  spanning-tree port type edge trunk
exit

interface eth 1/17
switchport mode trunk
switchport trunk allowed vlan 1,10,134
```

```
channel-group 10 mode active
no shut

interface eth 1/18
switchport mode trunk
switchport trunk allowed vlan 1,10,134
channel-group 11 mode active
no shut

copy running-config startup-config
```

Repeat the above on both N5K's.

The Show VPC status will display the following for a successful configuration.

```
vPC Peer-link status
---------------------------------------------------------------------
id   Port   Status Active vlans
--   ----   ------ --------------------------------------------------
1    Po1    up     1,10,134

vPC status
----------------------------------------------------------------------------
id     Port        Status Consistency Reason                Active vlans
------ ----------- ------ ----------- ------------------------ -----------
10     Po10        up     success     success                  1,10,134
11     Po11        up     success     success                  1,10,134

show interface port-channel 10-11 brief

--------------------------------------------------------------------------------
Port-channel VLAN  Type Mode   Status  Reason               Speed  Protocol
Interface
--------------------------------------------------------------------------------
Po10         1     eth  trunk  up      none                 a-10G(D)  lacp
Po11         1     eth  trunk  up      none                 a-10G(D)  lacp
```

Setting Up Jumbo Frames on N5K

Jumbo frames with an mtu=9000 have to be setup on N5K. Oracle Interconnect traffic under normal conditions does not go to the northbound switch like N5K's as all the private vinc's are configured in Fabric B. However if there is a partial link or IOM failure, the private interconnect traffic has to go to the immediate northbound switch (N5K in our case) to reach Fabric B.

Use the command shown below to configure the Jumbo frames Nexus 5K Fabric A Switch:

```
sj2-151-a19-n5k-FI-A# conf terminal
Enter configuration commands, one per line.  End with CNTL/Z.
sj2-151-a19-n5k-FI-A(config)# class-map type network-qos class-platinum
sj2-151-a19-n5k-FI-A(config-cmap-nq)# exit
```

```
sj2-151-a19-n5k-FI-A(config)# policy-map type network-qos jumbo
sj2-151-a19-n5k-FI-A(config-pmap-nq)# class type network-qos class-default
sj2-151-a19-n5k-FI-A(config-pmap-nq-c)# mtu 9216
sj2-151-a19-n5k-FI-A(config-pmap-nq-c)# multicast-optimize
sj2-151-a19-n5k-FI-A(config-pmap-nq-c)# exit
sj2-151-a19-n5k-FI-A(config-pmap-nq)# system qos
sj2-151-a19-n5k-FI-A(config-sys-qos)# service-policy type network-qos jumbo
sj2-151-a19-n5k-FI-A(config-sys-qos)# exit
sj2-151-a19-n5k-FI-A(config)# copy running-config startup-config
[########################################] 100%
sj2-151-a19-n5k-FI-A(config)#
```

Enable this on both N5K setups.

Installing the Operating System, Additional RPM's and Preparing the System for Oracle RAC and Database

Note For our testing purposes, Oracle Linux 6.3 was installed. However, the tests were done with both uek2 kernel and the Red Hat Compatible kernel. Where necessary, information is provided on RHEL compatible kernels.

Preparatory Steps

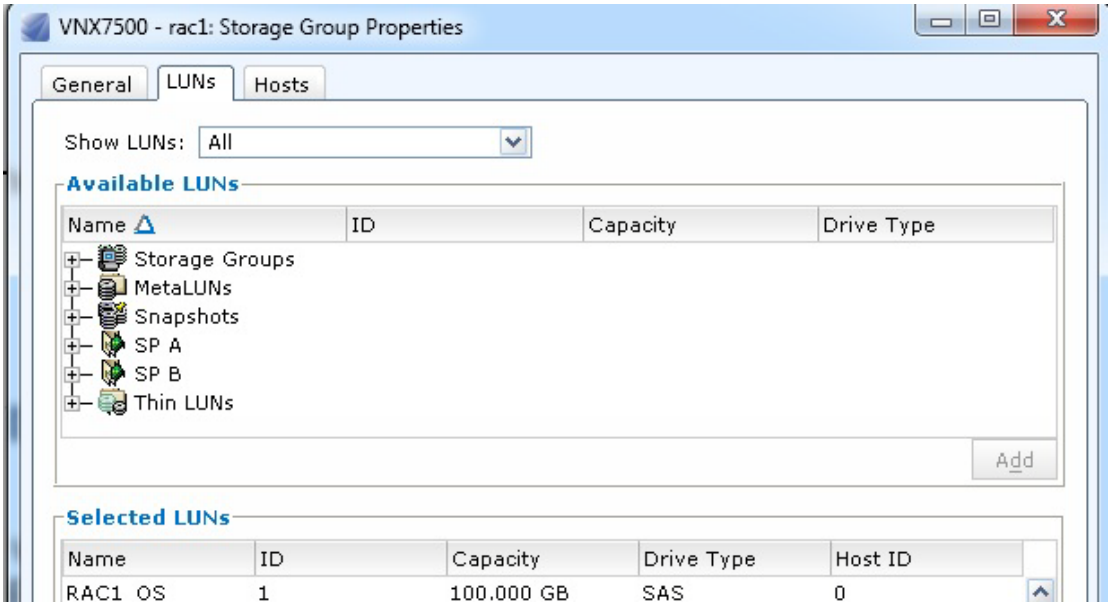
A few changes may have to be done on the storage and on N5K in order to install Oracle Linux 6.3 with boot LUNs, configured on EMC PowerPath. More detailed steps are provided in EMC PowerPath for Linux ver 5.7 Installation and Administration guide.

Cisco UCS Manager allows you to define boot policies for each server that can be configured to present the boot lun.

Storage Boot LUN Configuration

Make sure that the boot LUN for the server is presented to the host first from the storage side. Four storage groups were defined, one for each Cisco UCS B440. For server 1, the boot LUN was added to the first storage group. Also make a note of host id (preferably 0 as this is the first LUN presented to the host) before moving further.

Figure 15 Storage Group Properties



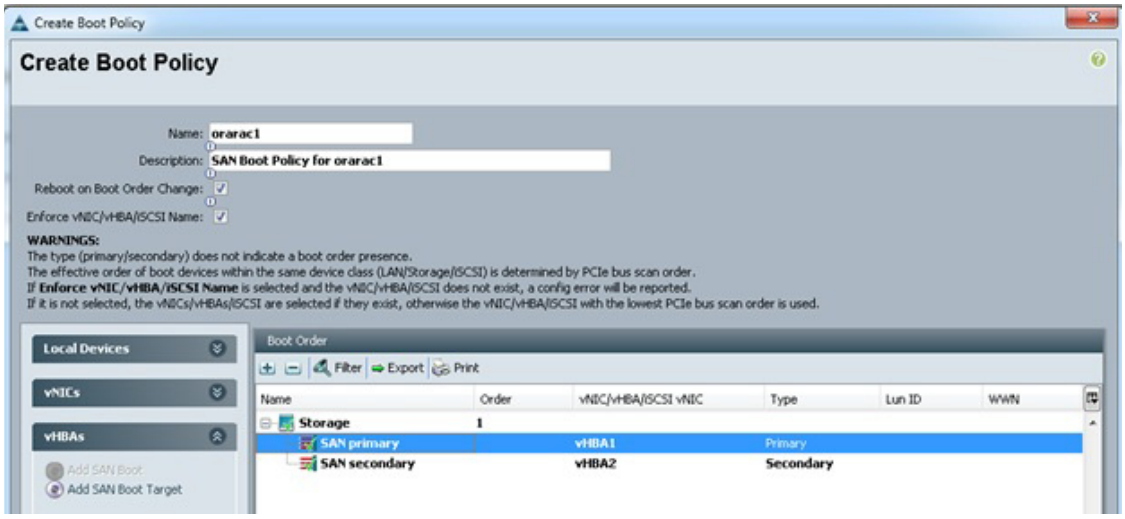
SAN Zoning Changes on N5K for Boot

Change the zoning policy on N5K's so that only one path is available during the boot time. Disable the zones say on N5k-B and enable only on N5k-A. Also make sure that only one path is available before install. Once the installation is complete and PowerPath is completely setup, this may be reverted back to it's full paths. As an example for server 1 (orarac A) only one zone is made available before install as below.

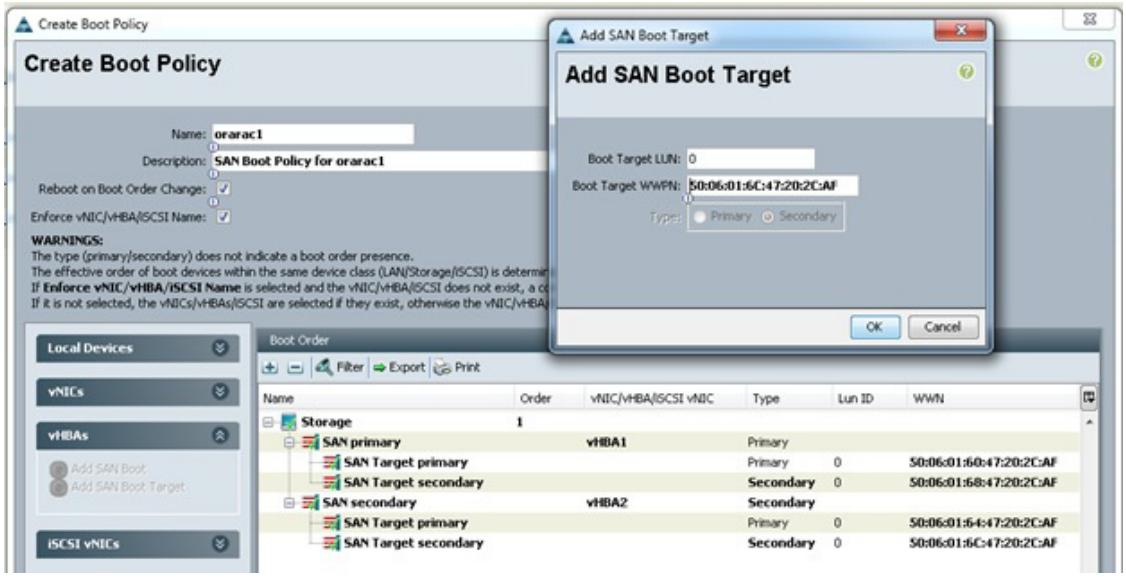
```
  zone name orarac1_hba1 vsan 15
* fcid 0x3a0022 [pwwn 20:00:00:25:b5:00:00:1f]
* fcid 0x3a00ef [pwwn 50:06:01:60:47:20:2c:af] [A2P0]
```

Configure Boot Policies on Cisco UCS Servers

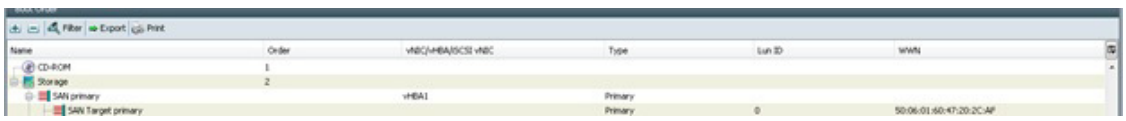
1. Define boot policy for the server 1.
2. Login to Cisco UCS Manager, servers tab, filter on policies, and right-click on Boot Policy.



For both SAN Primary and SAN secondary add the SAN Boot targets as shown below. The Boot Target LUN ID should match with Host ID from VNX as mentioned earlier.



3. Click OK to create the boot policy for the server. This has to be repeated for all the Oracle RAC servers.
4. To help ensure that you do not have multiple paths during boot time, temporarily disable all the paths and enable only one as below.



When the hardware boots up, since it has only one path now, you may see only WWN as shown above.

This completes the preparatory step for the OS install.

Install Oracle Linux 6.3 from Image

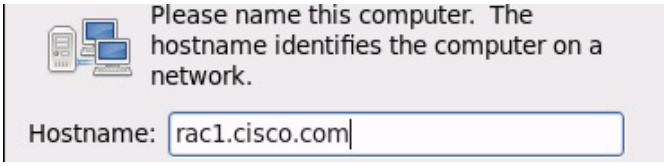
To install Oracle Linux 6.3, do the following steps:

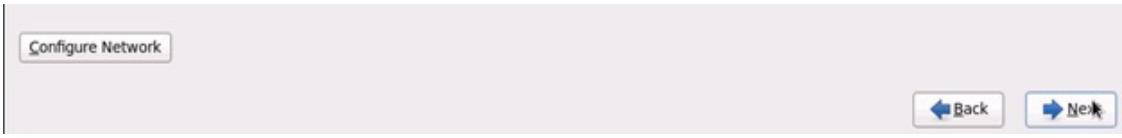
1. Download Oracle Linux 6.3 images from https://edelivery.oracle.com/linux or as appropriate. Mount the image and launch the installer.
2. Launch KVM console for the desired server, click on virtual media, add image and reset the server. When the server comes up, it launches the Oracle Linux Installer.

Note Only a few of the screen shots for the install are provided below.

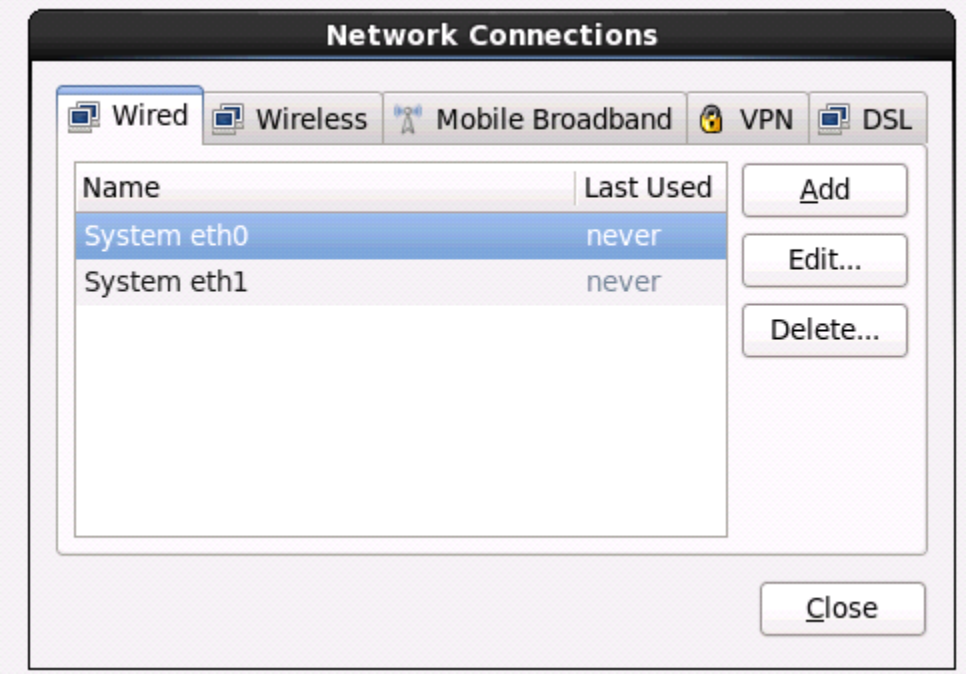
3. Select your language and installation.

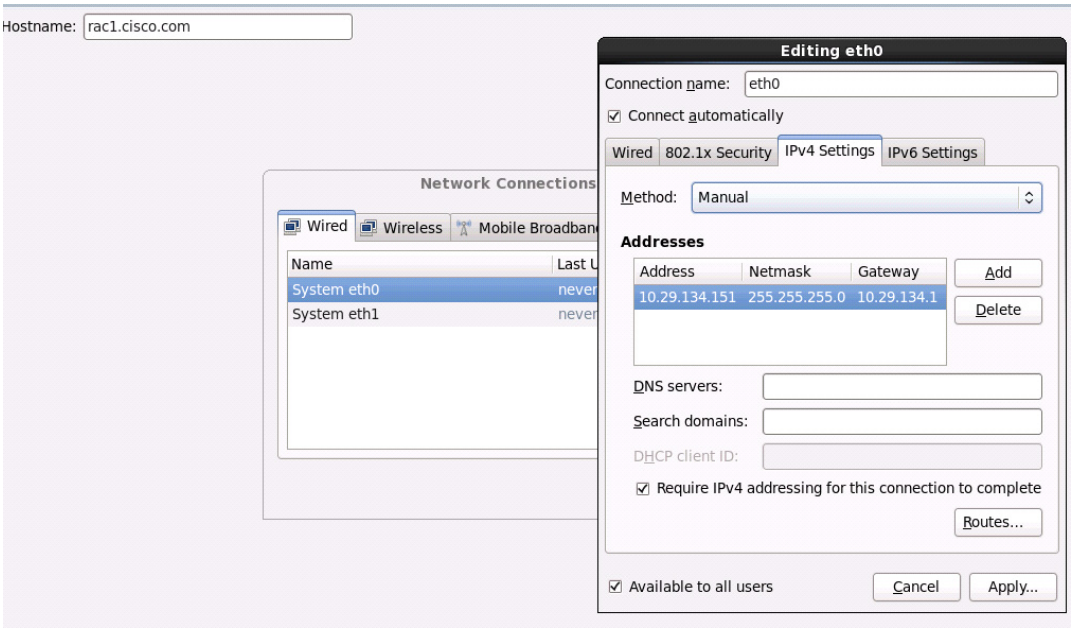
4. Select the hostname, and click Configure Network to configure both your private and public networks.



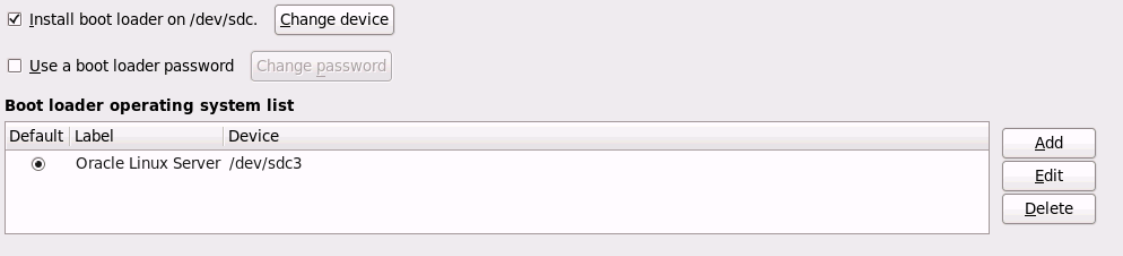


5. Edit each network interface and populate with appropriate entries.

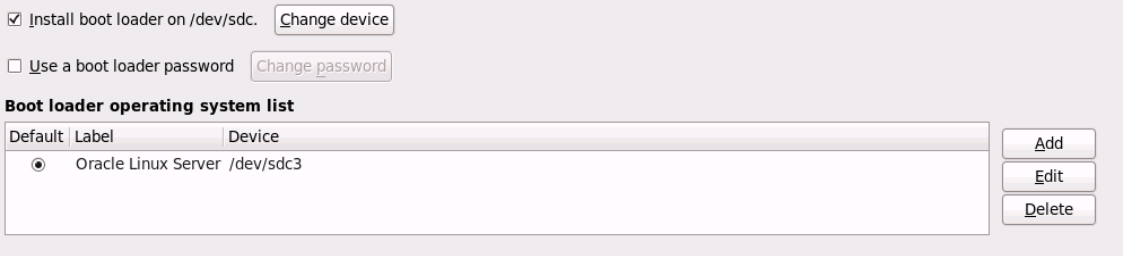




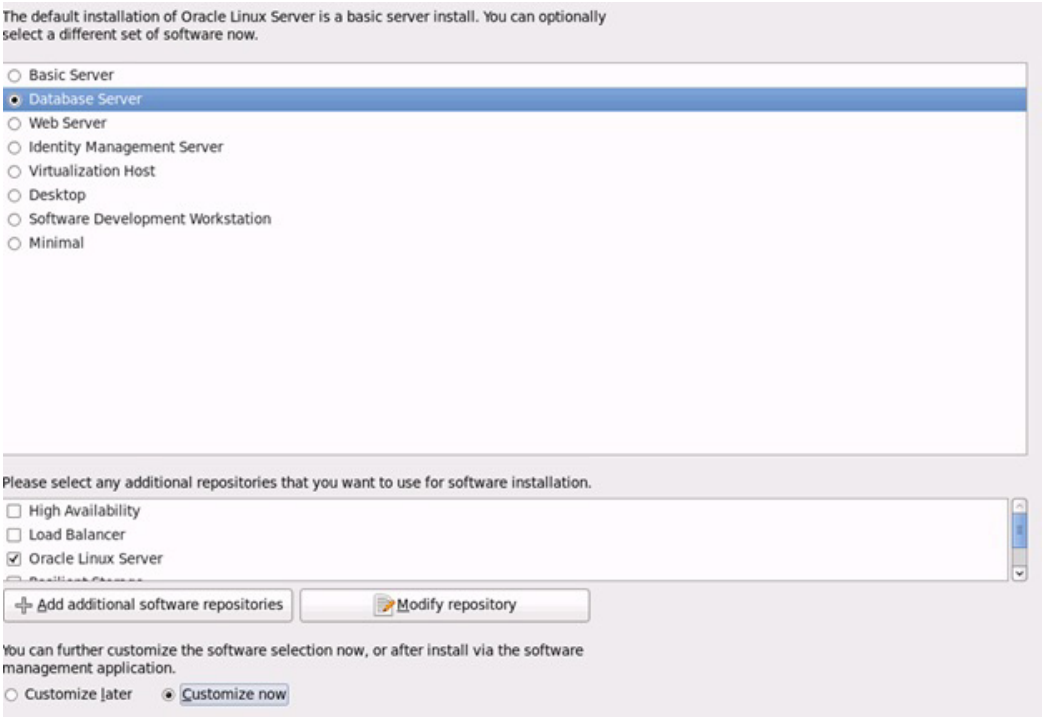
6. Select an appropriate Time Zone for your environment and enter the root password.
7. Click Review and modify partitioning layout.
8. Click Next.
9. Select the appropriate devices and size them.
10. Select to change the boot loader device.



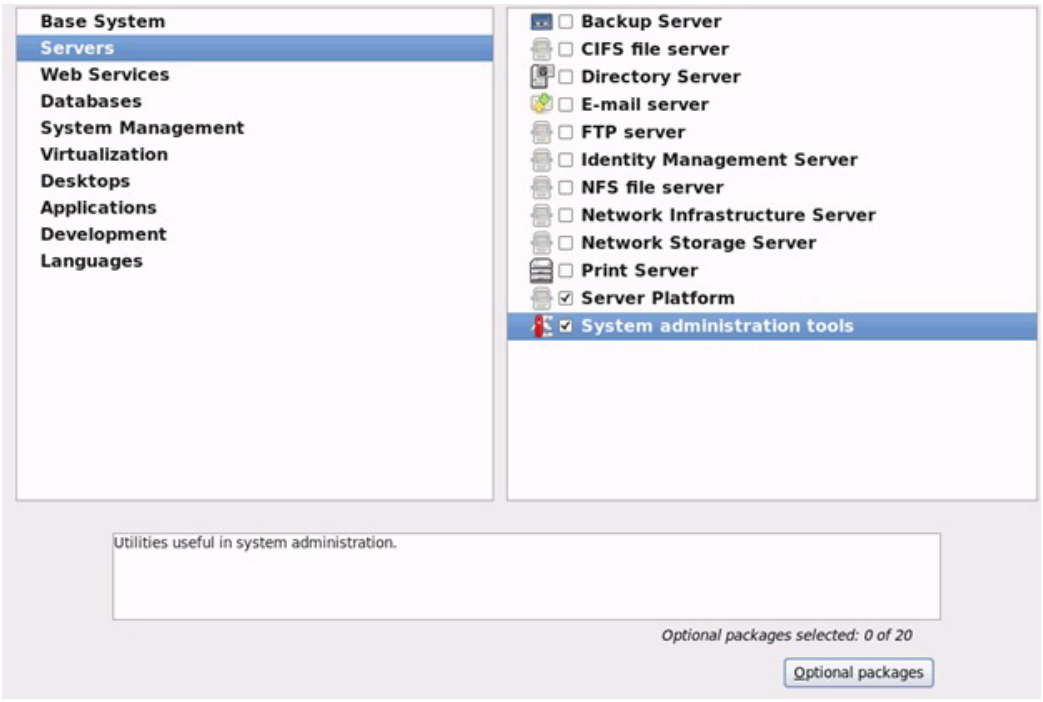
11. Select Install boot loader on dev/sdc and highlight the boot loader.

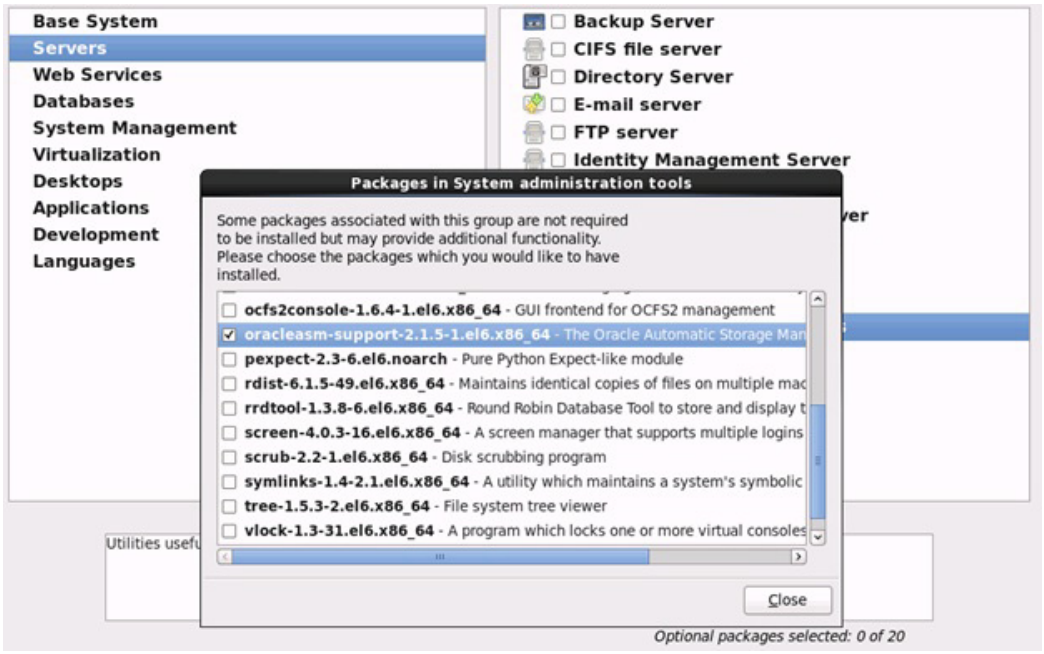


12. Select Database Server and click Customize now.

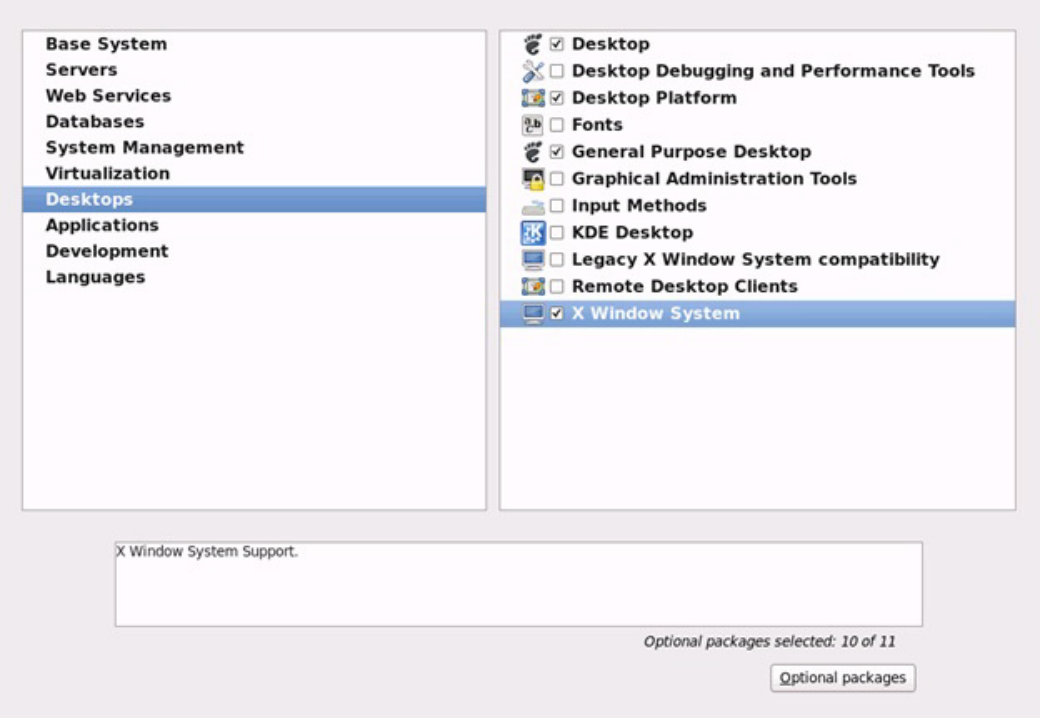


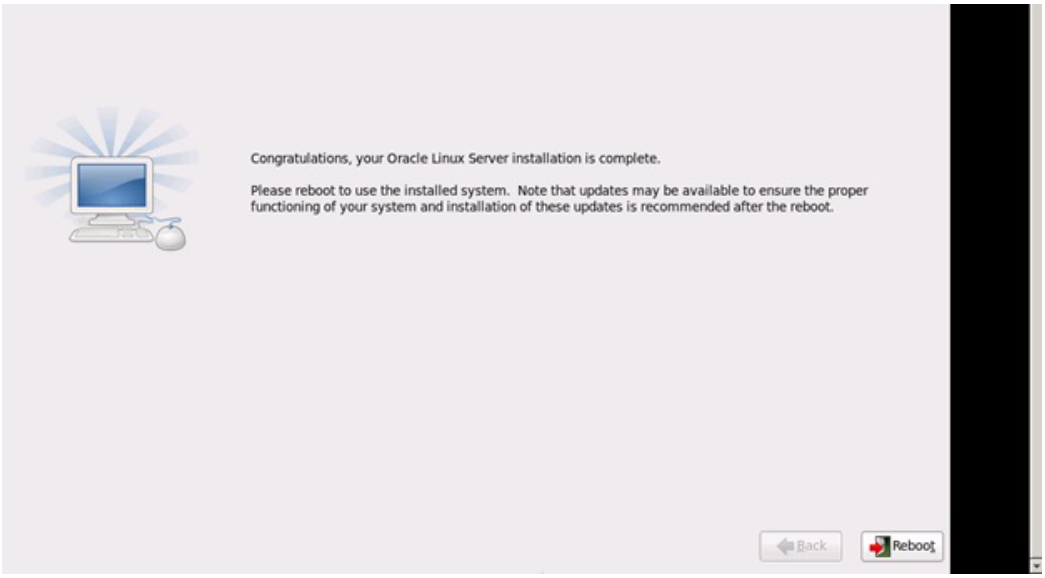
13. In the servers menu, select system administration tools and Oracle asm support tools.





14. Check X Windows System for Desktops.





15. Reboot the server and accept license information.
16. Register the system as needed and synchronize the time with NTP. If NTP is not configured, Oracle RAC cluster synchronization daemon starts in Oracle RAC node to sync up the time between the cluster nodes and maintaining the mean cluster time. Both NTP and OCSSD are mutually exclusive.

This completes the OS install.

Miscellaneous Post-install Steps

The following changes were made on the test bed where Oracle RAC install was done.

Disable selinux

It is recommended to disable selinux.

Edit /etc/selinux/config and change to

SELINUX=disabled

#SELINUXTYPE=targeted

Modify/create the dba group if needed

groupmod -g 500 dba

Change sshd_config file

RSAAuthentication yes

PubkeyAuthentication yes

AuthorizedKeysFile .ssh/authorized_keys

AuthorizedKeysCommand none

UsePAM yes

X11Forwarding yes

Subsystem sftp /usr/libexec/openssh/sftp-server

Disable firewalls

service iptables stop

service ip6tables stop

chkconfig iptables off

chkconfig ip6tables off

Make sure /etc/sysconfig/network has an entry for hostname. Preferably add NETWORKING_IPV6=no

Configure ssh trust for Oracle user

Preferably configure trust between nodes for Oracle user. This can be done by Oracle Installer during run time also.

ssh-keygen -t rsa

cd $HOME/.ssh.

cat id_rsa.pub >> authorized_keys

ssh <server name > should login back to the host.

Setup yum.repository

cd /etc/yum.repos.d

wget http://public-yum.oracle.com/public-yum-ol6.repo

edit the downloaded file public-yum-ol6.repo and change status as enabled=1

Run yum update.

You may have to set up http_proxy environment variable in case the server accesses internet via a proxy.

The yum update will not only bring the latest packages, but also brings and installs

oracle-rdbms-server-11gR2-preinstall-1.0-6.el6.x86_64 rpm. This rpm sets the some of the kernel parameters like in /etc/sysctl.conf, /etc/security/limits.conf etc

Install Linux driver for Cisco 10G FCOE HBA

Go to http://software.cisco.com/download/navigator.html

In the download page, select servers-Unified computing. On the right menu select your class of servers say Cisco UCS B-Series Blade Server software and then select Cisco Unified Computing System (UCS) Drivers in the following page.

Select your firmware version under All Releases, say 2.1 and download the ISO image of

Cisco UCS-related drivers for your matching firmware, for example ucs-bxxx-drivers.2.1.1a.iso.

Extract the fnic rpm from the iso.

Alternatively you can also mount the iso file. You can use KVM console too and map the iso.

After mapping virtual media - Login to host to copy the rpm

[root@rac1 ~]# mount -o loop /dev/cdrom /mnt

[root@rac1 ~]# cd /mnt

[root@rac1 mnt]# cd /mnt/Linux/Storage/Cisco/1280/Oracle/OL6.3

[root@rac1 OL6.3]# ls

dd-fnic-1.5.0.18-oracle-uek-6.3.iso

README-Oracle Linux Driver for Cisco 10G FCoE HBA.docx

Extract the rpm from iso.

Follow the instructions in README-Oracle Linux Driver for Cisco 10G FCoE HBA. In case you are running this on Oracle Linux Redhat compatible kernel, the appropriate driver for your linux version should be installed.

Here are the steps followed for uek2 kernel.

```
[root@rac2 fnic]# rpm -ivh kmod-fnic-1.5.0.18-1.el6uek.x86_64.rpm
Preparing...                ########################################### [100%]
   1:kmod-fnic              ########################################### [100%]

[root@rac2 fnic]# cd /lib/modules/2.6.39-200.24.1.el6uek.x86_64/extra/fnic/
[root@rac2 fnic]# ls -l
total 3448
-rw-r--r-- 1 root root 3524389 Oct 25 00:14 fnic.ko

[root@rac2 fnic]# cd
/lib/modules/2.6.39-200.24.1.el6uek.x86_64/kernel/drivers/scsi/fnic/
[root@rac2 fnic]# ls -l
total 124
```

```
-rwxr--r--. 1 root root 122008 Jun 23  2012 fnic.ko ? This was the original
driver
```

As this was a SAN Boot install, rmmod did not work.

```
[root@rac2 fnic]# pwd
/lib/modules/2.6.39-200.24.1.el6uek.x86_64/kernel/drivers/scsi/fnic
[root@rac2 fnic]# mv fnic.ko fnic.ko.old
cp fnic.ko /lib/modules/2.6.39-200.24.1.el6uek.x86_64/kernel/drivers/scsi/fnic

[root@rac2 fnic]# modprobe fnic
[root@rac2 fnic]# modinfo fnic
filename:       /lib/modules/2.6.39-200.24.1.el6uek.x86_64/extra/fnic/fnic.ko
version:        1.5.0.18
license:        GPL v2
author:         Abhijeet Joglekar <abjoglek@cisco.com>, Joseph R. Eykholt
<jeykholt@cisco.com>
description:    Cisco FCoE HBA Driver
srcversion:     24F8E443F0EEDBDF4802F20
alias:          pci:v00001137d00000045sv*sd*bc*sc*i*
depends:        libfc,libfcoe,scsi_transport_fc
vermagic:       2.6.39-200.24.1.el6uek.x86_64 SMP mod_unload modversions
parm:           fnic_log_level:bit mask of fnic logging levels (int)
parm:           fnic_trace_max_pages:Total allocated memory pages for fnic trace
buffer (uint)
```

In general it is good practice to install the latest drivers. In case you are planning to run RHEL compatible kernel, you may have to check for any additional drivers in enic/fnic category to be installed.

Reboot the host after making the changes and verify.

Configure PowerPath

After reboot, configure PowerPath as it is only with single path now. Please contact EMC for the appropriate version of PowerPath for the operating system.

The Oracle Linux 6.3 installs with 2 kernels

```
Uek2 kernel - 2.6.39-200.24.1.el6uek.x86_64 which is the default.
Red Hat binary compatible kernel - 2.6.32-279.el6.x86_64.
```

Note that the PowerPath versions are different with these kernels. As the tests were done in the test bed with both the kernels (by flipping the grub entries), the powerpath versions also had to be changed accordingly.

Obtain the following rpm's from EMC directly.

```
HostAgent-Linux-64-x86-en_US-1.0.0.1.0474-1.x86_64
EMCpower.LINUX-5.7.1.00.00-029.ol6_uek2_r2.x86_64 ( power path rpm for uek2
kernel )
EMCPower.LINUX-5.7.1.00.00-029.RHEL6.x86_64 (PowerPath rpm for Red Hat
Compatible kernel ).
```

For the actual list of PowerPath and Linux Kernel versions, refer to http://powerlink.emc.com

Make sure that multipath is not running.

```
[root@rac1 powerpath]# service --status-all  | grep multipath
[root@rac1 powerpath]# multipath -ll
-bash: multipath: command not found

[root@rac1 powerpath]# rpm -ivh
HostAgent-Linux-64-x86-en_US-1.0.0.1.0474-1.x86_64.rpm
Preparing...                ########################################### [100%]
   1:HostAgent-Linux-64-x86-########################################### [100%]

[root@rac1 powerpath]# rpm -ivh
EMCPower.LINUX-5.7.1.00.00-029.OL6_UEK2_R2.x86_64.rpm
Preparing...                ########################################### [100%]
   1:EMCpower.LINUX         ########################################### [100%]
All trademarks used herein are the property of their respective owners.

[root@rac1 powerpath]# service hostagent start
Starting Navisphere agent:                                 [  OK  ]

[root@rac1 powerpath]# service PowerPath start
Starting PowerPath:  done

[root@rac1 powerpath]# powermt check_registration
There are no license keys now registered.

[root@rac1 powerpath]#  emcpreg  -add  < power path key here >
1 key(s) successfully added.

[root@rac1 powerpath]# powermt set policy=co
[root@rac1 powerpath]# powermt config
[root@rac1 powerpath]# powermt save
[root@rac1 powerpath]#

[root@rac1 powerpath]# powermt display dev=all
Pseudo name=emcpowera
VNX ID=APM00120902426 [rac1]
Logical device ID=600601605DB02600E00053566AAFE111 [RAC1_OS]
state=alive; policy=CLAROpt; queued-IOs=0
Owner: default=SP A, current=SP A       Array failover mode: 4
==============================================================================
--------------- Host ---------------   - Stor -  -- I/O Path --   -- Stats ---
###  HW Path               I/O Paths    Interf.  Mode     State   Q-IOs Errors
==============================================================================
   3 fnic                   sda         SP A0    active   alive      0      0
```

Note Only one path is currently active.

Reconfigure Zoning and Boot Policies

When PowerPath is installed, make necessary changes both in boot policies and zoning info as mentioned earlier to revert back to all the paths.

The zoning attributes for each HBA (hba1 as an example below) needs to be reverted back to what was planned earlier

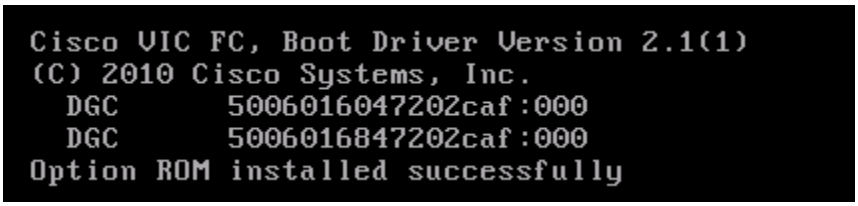
```
[pwwn 20:00:00:25:b5:00:00:1f]
```

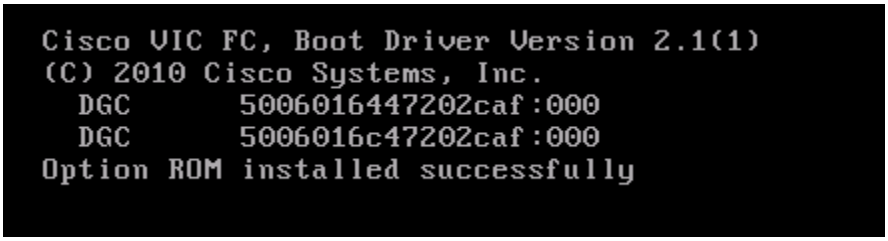
```
[pwwn 50:06:01:62:47:20:2c:af] [A2P2]
[pwwn 50:06:01:68:47:20:2c:af] [B2P0]
[pwwn 50:06:01:6a:47:20:2c:af] [B2P2]
[pwwn 50:06:01:60:47:20:2c:af] [A2P0]
```

1. Change the boot policy of the server to multiple paths.

2. Reboot the server.

When hardware boots up, it provides information on these paths:





After reboot all the paths should be active.

After activating, powermt will display information, for example:

```
[root@rac1 powerpath]# powermt display dev=all
Pseudo name=emcpowera
VNX ID=APM00120902426 [rac1]
Logical device ID=600601605DB02600E00053566AAFE111 [RAC1_OS]
state=alive; policy=CLAROpt; queued-IOs=0
Owner: default=SP A, current=SP A       Array failover mode: 4
==============================================================================
--------------- Host ---------------   - Stor -   -- I/O Path --    -- Stats ---
###  HW Path                I/O Paths    Interf.  Mode     State   Q-IOs Errors
==============================================================================
   2 fnic                    sdq         SP A4    active   alive       0      0
   4 fnic                    sdr         SP A6    active   alive       0      0
   2 fnic                    sdp         SP A6    active   alive       0      0
```

```
4 fnic                  sdo         SP A4    active   alive      0      0
2 fnic                  sdm         SP B4    active   alive      0      0
4 fnic                  sdn         SP B4    active   alive      0      0
2 fnic                  sdl         SP B6    active   alive      0      0
4 fnic                  sdk         SP B6    active   alive      0      0
1 fnic                  sdj         SP A0    active   alive      0      0
1 fnic                  sdi         SP A2    active   alive      0      0
3 fnic                  sdh         SP A2    active   alive      0      0
3 fnic                  sdg         SP B0    active   alive      0      0
1 fnic                  sdf         SP B0    active   alive      0      0
3 fnic                  sde         SP B2    active   alive      0      0
1 fnic                  sdd         SP B2    active   alive      0      0
3 fnic                  sda         SP A0    active   alive      0      0
```

Configuring Boot LUN

Follow the instructions from the EMC PowerPath Install and Administration guide. A few of the steps are mentioned below.

Powermt command above shows that emcpowera is the pseudo device for RAC1_OS lun.

- Capture the partitions from /proc/partitions

  root@rac1 dev]# cat /proc/partitions | grep emcpower

  120 0 104857600 emcpowera

  120 1 512000 emcpowera1 <- Boot Partition

- Backup /etc/fstab file and change the entries

  /dev/mapper/vg_rac1-lv_root / ext3 defaults 1 1

  #UUID=6ec54694-f657-4078-a151-f45d93e125cd /boot ext3 defaults 1 2

  /dev/emcpowera1 /boot ext3 defaults 1 0

  # fsck disabled for /boot partition

  /dev/mapper/vg_rac1-lv_swap swap swap defaults 0 0

  tmpfs /dev/shm tmpfs defaults 0 0

  devpts /dev/pts devpts gid=5,mode=620 0 0

  sysfs /sys sysfs defaults 0 0

  proc /proc proc defaults 0 0

- Change to pseudo devices entries in fstab

- Unmount and mount boot partition

  [root@rac1 ~]# umount /boot

  [root@rac1 ~]# mount /boot

- Check emcpower devices for system partitions now

  [root@rac1 ~]# df -k

  Filesystem 1K-blocks Used Available Use% Mounted on

  /dev/mapper/vg_rac4-lv_root

```
                    82545328  26613832  51738424  34% /
tmpfs                132278816    743332 131535484   1% /dev/shm
/dev/emcpowera1         495844     79767    390477  17% /boot
```

- Make lvm changes

  Take backup of /etc/lvm.conf and make changes to filter as below.

```
#    filter = [ "a/.*/" ]
   filter = [ "a/emcpower.*/", "r/sd.*/", "r/disk.*/" ] # New values
```

- Run vgscan and lvmdiskscan to flush out cache

```
[root@rac1 lvm]# vgscan -v
Wiping cache of LVM-capable devices
Wiping internal VG cache
Reading all physical volumes.  This may take a while...

[root@rac1 lvm]# lvmdiskscan
  /dev/ram0       [      16.00 MiB]
  /dev/ram1       [      16.00 MiB]
  /dev/emcpowera1 [     500.00 MiB]
  /dev/ram2       [      16.00 MiB]
  /dev/emcpowera2 [      19.53 GiB]
  ,Ä¶,Ä¶,Ä¶,Ä¶,Ä¶,Ä¶,Ä¶,Ä¶,Ä¶,Ä¶,Ä¶,Ä¶,Ä¶..
```

- Create new image file

```
cd /boot
[root@rac1 boot]# dracut /boot/initramfs-PP-$(uname -r).img $(uname -r)

[root@rac1 boot]# ls -l initramfs*
-rw-r--r--. 1 root root 16188188 Jan 15  2013 initramfs-2.6.32-279.el6.x86_64.img
-rw-r--r--. 1 root root 16082045 Jan 15  2013 initramfs-2.6.39-200.24.1.el6uek.x86_64.img
-rw-r--r--  1 root root 16138643 Jan 15 14:58 initramfs-PP-2.6.39-200.24.1.el6uek.x86_64.img
```

- Backup grub.conf and replace the entries pointing to new PowerPath initramfs.
- Reboot the server

This completes the SAN boot install items.

Repeat the above steps on all hosts.

Configure Oracle ASM

Oracle ASM is installed as part of the install in OEL 6. It just needs to be configured.

```
[root@rac1 ~]# /etc/init.d/oracleasm configure
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 []: oracle
Default group to own the driver interface [dba]: dba
Start Oracle ASM library driver on boot (y/n) [y]: y
Scan for Oracle ASM disks on boot (y/n) [y]: y
Writing Oracle ASM library driver configuration: done
Initializing the Oracle ASMLib driver:                     [  OK  ]
Scanning the system for Oracle ASMLib disks:               [  OK  ]

[root@rac1 ~]# cat /etc/sysconfig/oracleasm | grep -v '^#'
ORACLEASM_ENABLED=true
ORACLEASM_UID=oracle
ORACLEASM_GID=dba
ORACLEASM_SCANBOOT=true
ORACLEASM_SCANORDER="emcpower" <Add this entry
ORACLEASM_SCANEXCLUDE="sd"  < Add this entry
```

This will create a mount point /dev/oracleasm/disks

Configure ASM LUNS and Create Disks

Mask the LUNS and Create Partitions

Configure Storage LUNs

Add the necessary luns to the storage groups and provide connectivity to the hosts. Reboot the hosts so that scsi is scanned and the luns are visible.

ls /dev/emcpower* or powermt display dev=all should reveal that all devices are now visible on the host.

Partition LUNs

Partition the luns with an offset of 1MB. While it is necessary to create partitions on disks for Oracle ASM (just to prevent any accidental overwrite), it is equally important to create an aligned partition. Setting this offset aligns host I/O operations with the back end storage I/O operations.

Use host utilities like fdisk to create a partition on the disk.

Create a input file, fdisk.input as shown below:

d

n

p

1

x

b

1

2048 <- 2048 for EMC VNX.

p

w

Execute as fdisk /dev/emcpower[name] < fdisk.input. This makes partition at 2048 cylinders. In fact this can be scripted for all the luns too.

Now all the pseudo partitions should be available in /dev as emcpowera1, emcpowerb1, emcpowerab1 etc.

Create ASM Disks

When the partitions are created, create ASM disks with oracleasm API's.

oracleasm createdisk -v DSS_DATA_1 /dev/emc[partition name]

This will create a disk label as DSS_DATA_1 on the partition. This can be queried with oracle supplied kfed/kfod tools that are covered later.

Repeat the process for all the partitions and create ASM disks for all your database and RAC files.

Scan the disks with oracleasm and these should be visible under /dev/oracleasm/disks mount point created by oracleasm earlier as shown below:

```
[root@rac1 disks]# oracleasm scandisks
Reloading disk partitions: done
Cleaning any stale ASM disks...
Scanning system for ASM disks...

[root@rac1 disks]# cd /dev/oracleasm/disks/
[root@rac1 disks]# ls
DELMARCLUSDG_0000  DSS_DATA_17  DSS_DATA_29   OLTP_DATA_11  OLTP_DATA_9
DELMARCLUSDG_0001  DSS_DATA_18  DSS_DATA_3    OLTP_DATA_12  REDO01
DELMARCLUSDG_0002  DSS_DATA_19  DSS_DATA_30   OLTP_DATA_13  REDO02
DELMARCLUSDG_0003  DSS_DATA_2   DSS_DATA_31   OLTP_DATA_14  REDO03
DELMARCLUSDG_0004  DSS_DATA_20  DSS_DATA_32   OLTP_DATA_15  REDO04
DSS_DATA_1         DSS_DATA_21  DSS_DATA_4    OLTP_DATA_16  REDO05
DSS_DATA_10        DSS_DATA_22  DSS_DATA_5    OLTP_DATA_2   REDO06
DSS_DATA_11        DSS_DATA_23  DSS_DATA_6    OLTP_DATA_3   REDO07
DSS_DATA_12        DSS_DATA_24  DSS_DATA_7    OLTP_DATA_4   REDO08
DSS_DATA_13        DSS_DATA_25  DSS_DATA_8    OLTP_DATA_5
DSS_DATA_14        DSS_DATA_26  DSS_DATA_9    OLTP_DATA_6
DSS_DATA_15        DSS_DATA_27  OLTP_DATA_1   OLTP_DATA_7
DSS_DATA_16        DSS_DATA_28  OLTP_DATA_10  OLTP_DATA_8
```

Now the system is ready for the Oracle installation.

Oracle RAC and Database Installation

RAC and Database Setup

We are not presenting the detailed steps of creating 4 Node Oracle RAC and database in this section. However, a few changes that were done to spfile etc. were noted and are presented in the Appendix.

It was a default Oracle 11.2.0.3 install after which PSU2 patchset was applied. A standard way of running runInstaller from oui to configure the components.

Few points from the install:

- 5 luns used for ASM Diskgroup for hosting OCR and Voting disks with normal redundancy.
- 32 luns were used by DSS database for datafiles.
- 16 luns were used by OLTP database for datafiles.
- 8 luns were used for Redo Log files for both the databases.
- 4 Diskgroups were created in ASM
  - DELMARCLUSDG
  - DSSDG
  - OLTPDG
  - REDODG

Swingbench Setup

Swingbench client (http://www.dominicgiles.com) was installed on another client host. Order Entry(oe) and Sales history (sh) load generators were created in oltp and dss databases respectively.

soe schema in OLTP

Table Name	Number of Rows
CUSTOMERS	4,955,836,186
WAREHOUSES	1,000
ORDER_ITEMS	19,695,700,442
ORDERS	6,889,511,481
INVENTORIES	901,334
PRODUCT_INFORMATION	1,000
LOGON	7,691,179,347
PRODUCT_DESCRIPTIONS	1,000
ORDERENTRY_METADATA	4

sh schema in DSS

Table Name	Number of Rows
PROMOTIONS	503
PRODUCTS	72
CHANNELS	5
CUSTOMERS	3,319,173,792
SUPPLEMENTARY_DEMOGRAPHICS	3,319,173,792
SALES	16,595,869,056
COUNTRIES	23
TIMES	6,209

Both Order entry and Sales history were run same time, collecting performance data from database and swingbench output. Also data was captured from EM Grid Control that is presented in the subsequent sections.

Performance Data from the Test Bed

OLTP and DSS work loads were run in steady state for 24 hrs. First OLTP work loads were run, followed by DSS and then a combination of both. While the tests were run for 24 hrs a snippet of snapshots is presented below. The data was extracted from Oracle AWR reports, oratop, Operating system utilities etc. and were compared

OLTP Workload

This is OLTP workload only. Order entry tool from swingbench was used to run on the system. EM Grid generated graphs for OLTP Work Load on the setup with 700 users is shown in Figure 16.

Figure 16 OLTP Workload

Interconnect Traffic Extracted from OS Utilities (eth1 traffic)

Table 5 Interconnect Traffic in OLTP Workload

Interconnect Traffic	Rx	Tx
Node 1	129,211	140,750
Node 2	107,666	100,249
Node 3	109,797	111,874
Node 4	103,506	97,222
Total KB/sec	450,180	450,095

Oratop

Table 6 Oratop

```
oratop 1: 7685 oltp 12:54:29 up 0.3h, 4 ins, 5.15G mt, 724 sn, 3 us, 78% db
```

There were around 724 user sessions and 78% of database is busy.

ID	%CU	HLD	MBPS	IORL	%FR	PGAU	ASC	ASI	ASW	ASP	AAS	USN	TPS	UCPS	SSRT	DBC	DBW
3	18	17	132	1m	45	1G	13	11	34	0	73	194	4137	5466	2m	20	80
1	18	16	124	1m	45	1G	16	17	18	0	56	175	4055	5361	1m	26	74
2	21	17	131	1m	45	1G	17	15	22	0	65	185	4105	5428	2m	26	74
4	20	13	113	1m	48	1G	9	11	27	0	57	170	3816	5047	2m	28	72

The SSRT is the SQL service response time in milliseconds.

EVENT	AVG: TOT WAITS	TIME(s)	AVG_MS	PCT	WAIT_CLASS
DB CPU		41139		57	
db file sequential read	9710917	15957	1.6	22	User I/O
gc current grant 2-way	4223916	5564	1.3	8	Cluster
log file sync	1891433	4566	2.4	6	Commit
gc current block 3-way	2829359	4394	1.5	6	Cluster

ID	SID	SPID	USR	PROG	PGA	OPN	SQLID/BLOCKER	E/T	STATUS	STE	WAIT EVENT	WT
4	5765	56738	SOE	DEDI	3M	PLS	0w2qpro5t2zsp	0	ACTIVE	WAI	library ca	11m
3	7459	43420	SOE	DEDI	3M	PLS	0w2qpro5t2zsp	0	ACTIVE	WAI	library ca	10m
4	2714	56682	SOE	DEDI	3M	INS	0yas01k2p9ch4	0	ACTIVE	WAI	gc current	10m
3	6444	43564	SOE	DEDI	3M	SEL	4tx09gq85z630	0	ACTIVE	I/O	db file se	8m
2	3958	52286	SOE	DEDI	4M			0	ACTIVE	WAI	library ca	8m
1	4408	27742	SOE	DEDI	3M	SEL	4tx09gq85z630	0	ACTIVE	I/O	db file se	7m
2	3393	52276	SOE	DEDI	3M	PLS	apgb2g9q2zjk1	0	ACTIVE	WAI	library ca	7m
4	6782	56752	SOE	DEDI	4M	SEL	0mk367arf7gbw	0	ACTIVE	I/O	db file se	4m
3	2941	43662	SOE	DEDI	4M	PLS	147a57cxq3w9y	0	ACTIVE	WAI	library ca	4m
4	7911	56772	SOE	DEDI	3M	SEL	c13sma6rk27c	0	ACTIVE	WAI	gc cr req	4m
3	681	43622	SOE	DEDI	3M	SEL	7kk2m27t02na0g	0	ACTIVE	I/O	db file se	3m
1	4072	27957	SOE	DEDI	4M	UPD	3katpjzrqurh7	0	ACTIVE	I/O	db file se	3m
2	229	52166	SOE	DEDI	3M	SEL	c13sma6rk27c	0	ACTIVE	WAI	gc cr req	3m
3	7912	43592	SOE	DEDI	3M	INS	0bzdqkkj9mpaa	0	ACTIVE	CPU	wait for cpu	3m
2	8704	52370	SOE	DEDI	3M	SEL	0mk367arf7gbw	0	ACTIVE	WAI	gc current	3m

The system was generating around 500,000 TPM as recorded by swingbench (around 16,000 AWR TPS) with 700 Swingbench OE users for OLTP workload while performing 500 MB/sec of throughput at 8k database block size and was running around 50K to 60K of IOPS. The CPU utilization was well below 30% mark for the above load.

DSS Workload

This was DSS workload only. The SH from swingbench was used to generate DSS load.

EM generated graphs for DSS workload with 22 users is shown in Figure 17.

Figure 17 DSS Workload

The system was doing an IO of around 3400 MBPS with 22 users for DSS workload. The CPU utilization was well below 30% mark for the above load as well.

Mixed Workload (OLTP and DSS)

The tests were repeated with both the databases up and running to inject a mixed work load, OLTP with 8k block size and DSS with 1MB block size. The following is an analysis.

Performance Data Gathered from the Test Bed

OLTP Database

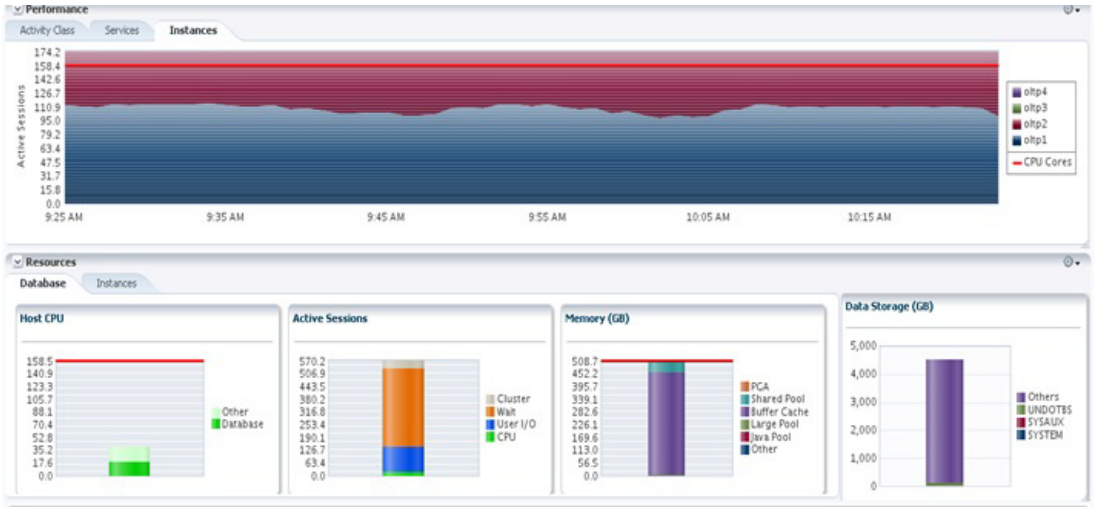
OLTP system was run with 500 users (reduced from 700 of standalone OLTP testing). Data was extracted from Oracle Enterprise Manager. The host CPU utilization was around 35% for the combined workload.

The swingbench log file shows the following at steady state.

Time	Users	TPM	TPS	User	System	Wait	Idle
10:24:59	[500/500]	207,330	3,510	0	0	0	0
10:25:04	[500/500]	209,980	4,065	0	0	0	0
10:25:09	[500/500]	212,466	4,205	0	0	0	0

When 500 users were run on OLTP database along with other DSS database work load, system was doing around 4200 Swing TPS or 210K TPM.

Figure 18 OLTP Performance Data from Mixed Workload



The above graph extracted from EM Grid control for OLTP database when combined workload was running. The host CPU utilization was around 30-35% with around 500 active sessions in the database.

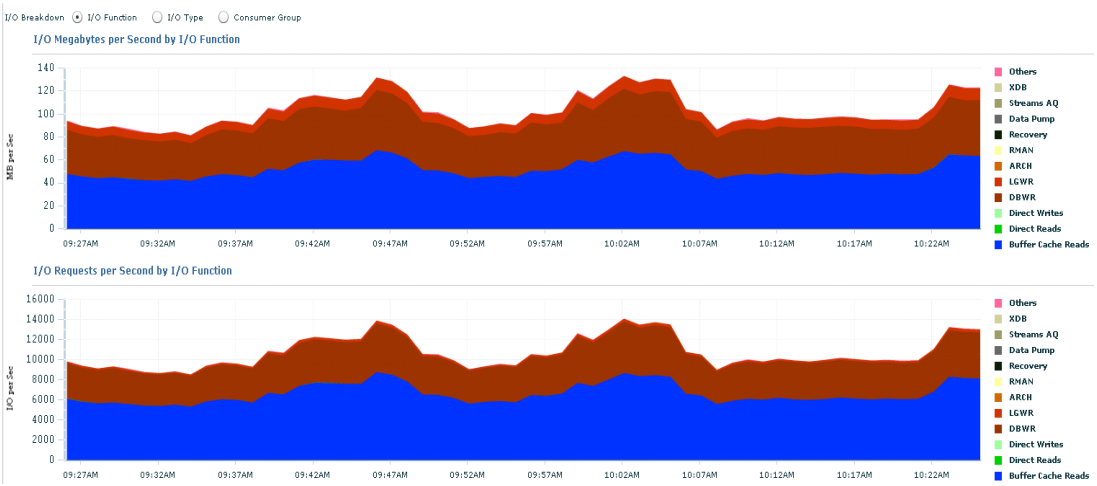
System Statistics - Per Second

Table 7 AWR Data from OLTP Database in Mixed Workload

Instance Number	Logical Reads/s	Physical Reads/s	Physical Writes/s	Redo Size (k)/s	Block Changes/s	User Calls/s	Execs/s	Parses/s	Logons/s	Txns/s
1	75,901.55	1,945.63	1,608.52	2,469.75	17,595.03	1,611.09	7,687.63	1.67	0.41	1,212.26
2	71,419.56	1,842.36	1,499.44	2,310.06	16,408.50	1,503.00	7,168.70	1.74	0.4	1,130.51
3	68,001.78	1,731.53	1,415.96	2,156.95	15,276.06	1,395.33	6,663.68	1.22	0.39	1,050.02
4	73,800.66	1,531.45	1,239.56	1,901.05	13,617.69	1,236.80	5,895.46	1.79	0.41	928.94
Sum	**289,123.55**	**7,050.97**	**5,763.48**	8,837.82	62,897.28	5,746.23	27,415.48	6.42	1.6	**4,321.73**
Avg	72,280.89	1,762.74	1,440.87	2,209.45	15,724.32	1,436.56	6,853.87	1.61	0.4	1,080.43
Std	3,389.77	177.25	155.65	242.04	1,693.76	159.66	763.55	0.26	0.01	120.78

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	15.27	15.2	0	15.08	12.6	0	2.51	1,949	1,944	0	1,190	1,130	0	59
2	14.47	14.4	0	14.06	11.7	0	2.34	1,847	1,842	0	1,099	1,041	0	57
3	13.59	13.52	0	13.26	11.1	0	2.19	1,735	1,731	0	1,099	1,042	0	56
4	12.13	11.95	0	11.63	9.68	0	1.93	1,541	1,530	0	988	933	0	54
Sum	**55.46**	55.06	0	**54.04**	45	0	8.97	**7,072**	7,047	0	4,375	**4,146**	0	225
Avg	13.87	13.77	0	13.51	11.3	0	2.24	1,768	1,762	0	1,094	1,036	0	56

From AWR RAC report, OLTP system was doing around 4300 tps and around 110 MBPS with 11,500 IOPS with almost 50% writes. This is also evident from the EM graphs captured for OLTP database.

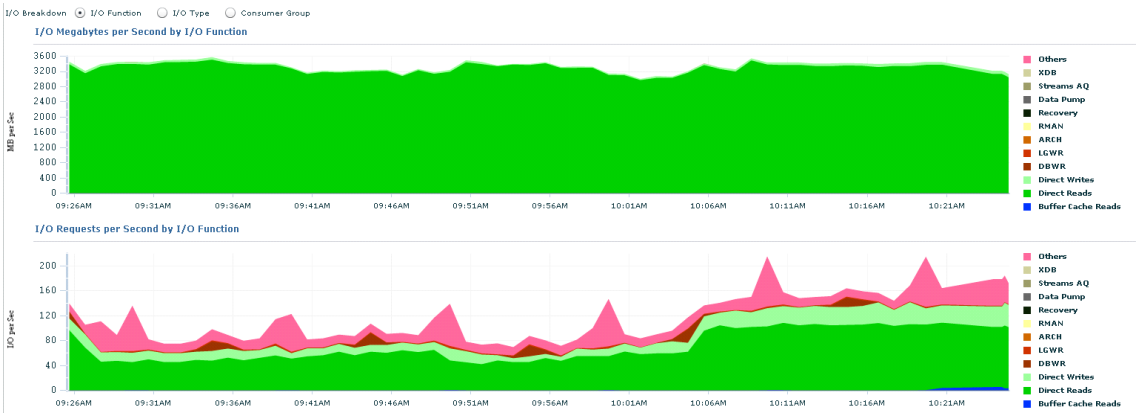


DSS Database

DSS tests were run with 18 users sales history load.

From Enterprise Manager graph system was doing around 3200 MBPS of read intensive work load.

Figure 19 DSS Performance Data from Mixed Workload



Also, data captured from AWR report for the DSS database reported the following stats.

	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	794.95	789.05	5.82	2.42	0.01	2.40	0.00	819	790	24	12	1	10	0
2	796.78	796.11	0.61	13.47	0.01	13.44	0.00	804	797	3	57	1	56	0
3	826.29	812.96	13.27	6.00	0.01	5.97	0.00	873	814	55	26	1	25	0
4	834.52	832.61	1.73	6.92	0.01	6.90	0.00	854	835	7	30	1	29	0
Sum	**3,252.55**	3,230.72	21.43	28.80	0.03	28.72	0.01	3,350	3,235	89	125	3	119	1
Avg	813.14	807.68	5.36	7.20	0.01	7.18	0.00	837	809	22	31	1	30	0

DSS System was doing an IO of around 3,250 MBPS.

Mixed Workload Operating System Statistics

Statistics were collected at host level across all the nodes and average values are reported below.

vmstat output

						Vmstat	output							
	free	buff	cache	si	so	bi	bo	in	cs	us	sy	id	wa	st
Node1	50274924	284516	2478272	0	0	910,723	56,905	72776	79612	10	5	66	19	0
Node2	52929856	236164	3008196	0	0	961,427	19,328	65389	69720	8	3	72	17	0
Node3	50067184	243516	1954484	0	0	947,715	33,560	67437	70712	10	4	68	19	0
Node4	50575280	242464	2781848	0	0	402,355	17,862	62990	69874	7	4	73	16	0
						3,222,220	127,655					69.8		
						Total	**3,271**	**MB/sec**						

Mpstat output

				mpstat output						

	CPU	%usr	%nice	%sys	%iowait	%irq	%soft	%steal	%guest	%idle
Node1	all	9.58	0	2.8	20	0	0.87	0	0	67
Node2	all	8.44	0	2.3	19	0	0.72	0	0	70
Node3	all	9.31	0	2.6	18	0	0.56	0	0	70
Node4	all	6.69	0	2.3	17	0	0.47	0	0	74

The average CPU idle% from mpstat was around 70%, thus making the utilization around 30%.

Five minutes load average

Node 1	40.7
Node 2	39.86
Node 3	39.54
Node 4	34.91

Destructive and Hardware Failover Tests

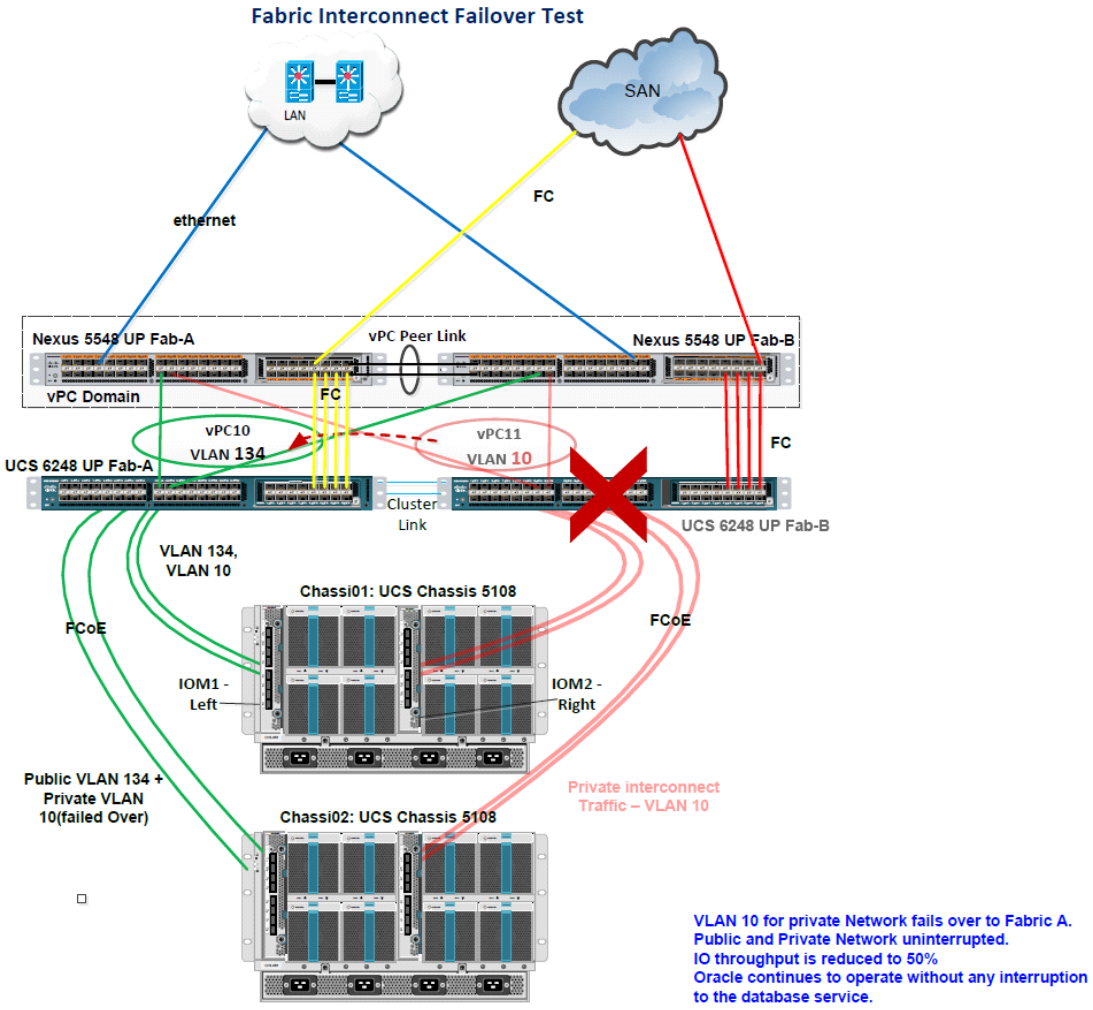
A few of the following destructive and hardware failover tests were conducted on a fully loaded system (with both OLTP and DSS workload running) to check on the high availability and recoverability of the system when faults were injected. The test cases are listed in Table 8.

Table 8 Destructive Test Cases

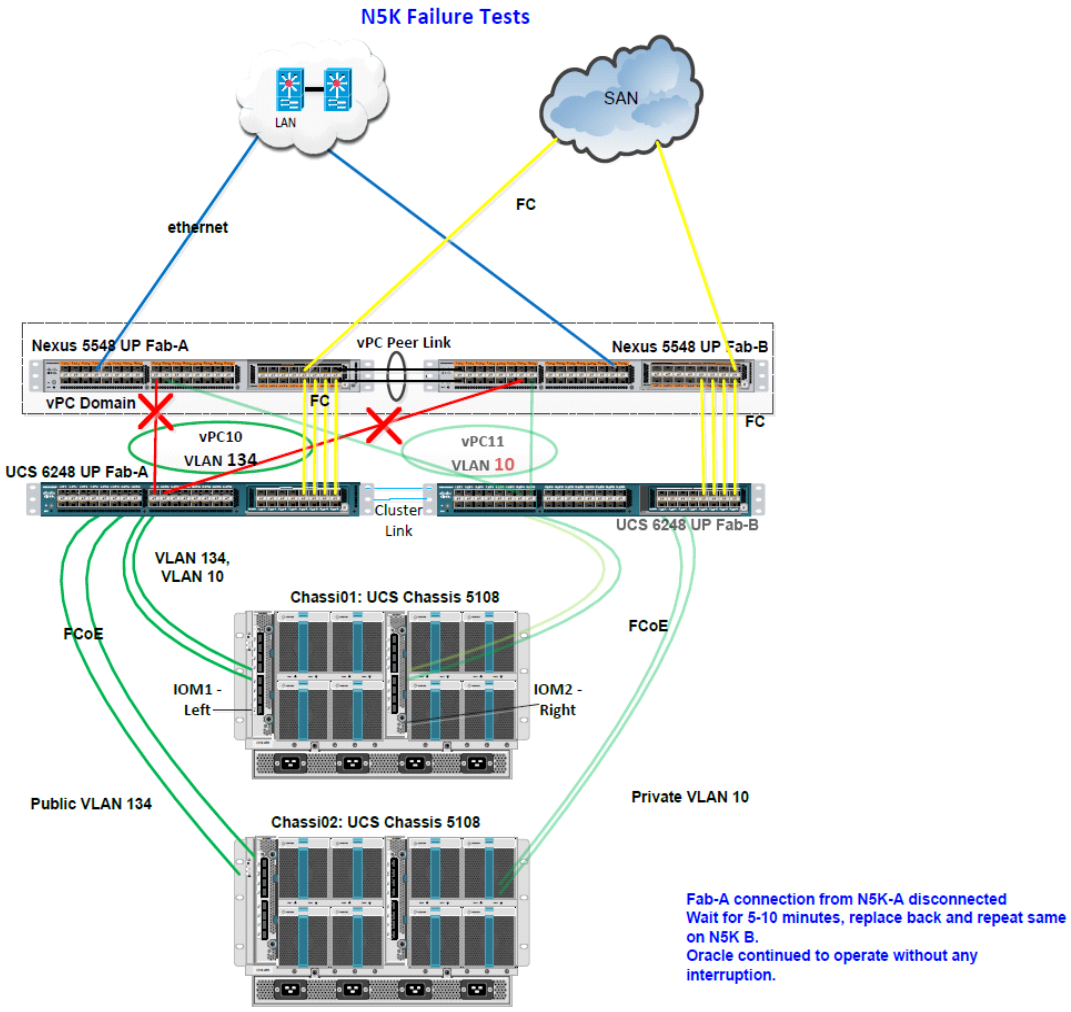
	Test	Status
Test 1 ,Äì Multiple Network Connection Failures	Run the system on full mixed work load. Disconnect the public links from first chassis and private links from second chassis one after other and reconnect each of them after 5 minutes.	Only second chassis servers rebooted. They joined the cluster back with a successful reconfigurations.
Test 2 ,Äì Single Network failure between FCoE and Corporate network	Run the system on full mixed work load. Disconnect connection to N5K-A from Fabric A, wait 5 minutes, connect it back and repeat for N5K-B.	No disruption.
Test 3 ,Äì Fabric Failover tests	Run the system on full load as above. Reboot Fabric A, let it join the cluster back and then Reboot Fabric B.	Fabric failovers did not cause any disruption to ethernet and/or FC traffic.
Test 4 ,Äì Disconnect all FC Storage paths	Run the system on full load and disconnect the storage paths on the fabrics.	All nodes went for a reboot because of inaccessibility of voting files. All instances joined the cluster back later.
Test 5 ,Äì Multi host failure across chassis	Run the system on full load and pull out one blade from each chassis. Put them back after 10 minutes	Instances reconfiguration happened both times while pulling and putting back the blades with no interruption to clients.

A few of the destructive tests done shown above are diagrammatically represented below.

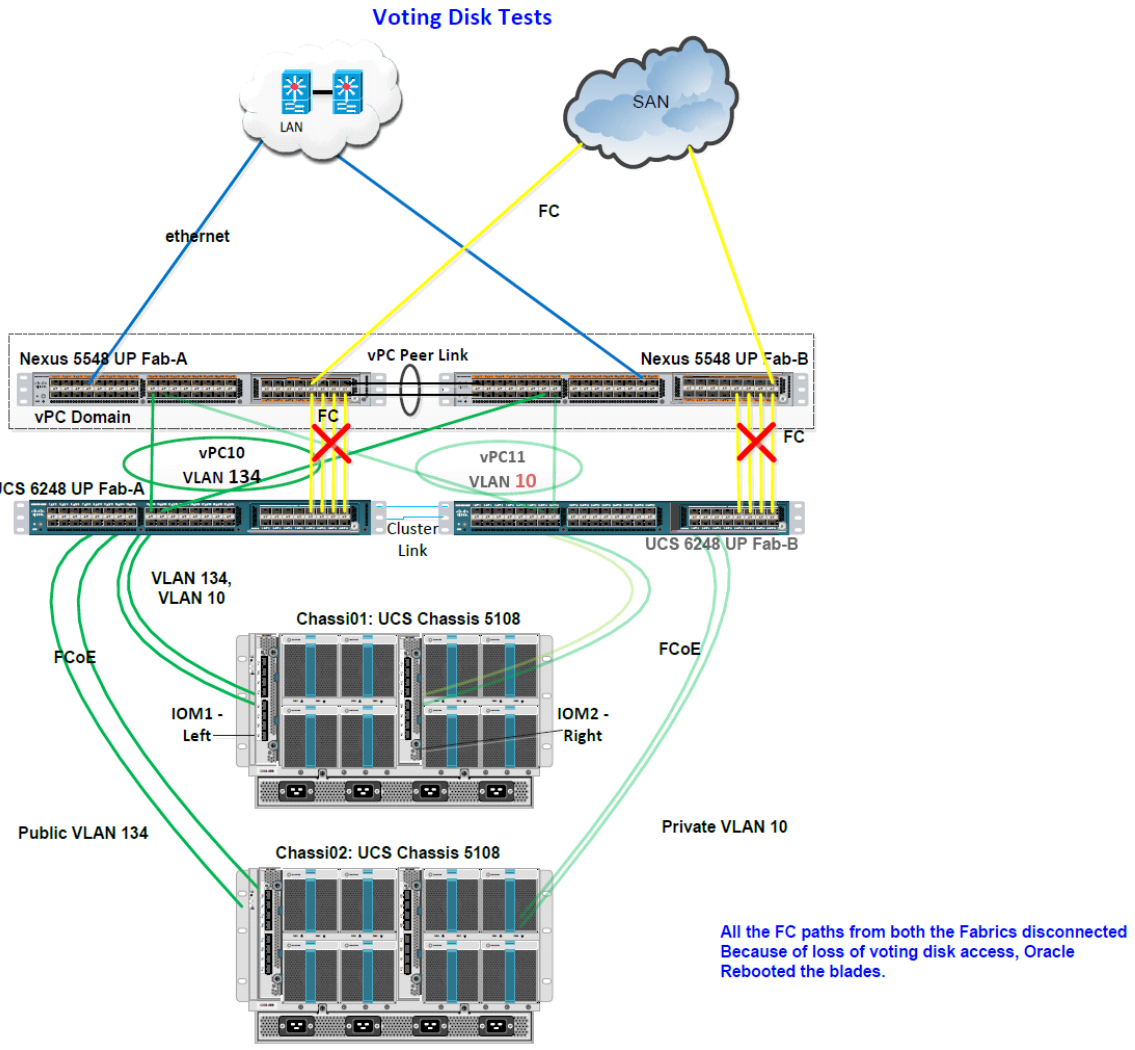
Test 3 - Fabric Failover Tests



Test 2 - Network Connectivity Failures



Test 4: Storage Path Failure Tests



Migrating ASMLIB to udev

ASMLIB to udev

As mentioned earlier, a few of the tests were done on both uek2 kernel and the Red Hat compatible kernel. We observed that Oracle ASM gets disabled while reverting to Red Hat Compatible Kernel. Hence there was a need to configure udev for Oracle ASM LUNs. An attempt was made to migrate Oracle LUNs from OracleASM to udev and that's what is provided below.

Note The following is only provided for reference. Please exercise caution and preferably work with Oracle if this is the desired direction on your production system. Because of its unique nature and the way installs could differ from site to site, it is strongly recommended to proof read, verify in a test and development systems before attempting this in a production.

For moving from uek2 kernel to Red Hat Compatible kernel, you may have to do a minimum of the following.

- Configure boot lun - A new image may have to be created with dracut. Refer boot lun section and EMC PowerPath 5.7 Install and Administration guide for details on configuring boot luns with PowerPath on RHEL 6 Kernels.
- The PowerPath version installed for uek2 kernel is not compatible with Oracle Linux Red Hat Compatible Kernel. Install the appropriate version of PowerPath.
- Check Miscellaneous post-install steps above for fnic drivers install on uek2 kernel. Please follow similar steps if any, for enic and fnic drivers for the appropriate Red Hat kernel version.
- Make changes in the grub.conf, pointing it to the right version and the boot ramfs created.

Create a Mapping Table Between Storage LUNs, PowerPath Pseudo Devices and Oracle ASM Disks

In a system that is up and running on Oracle Linux uek2 kernel, with PowerPath configured and OracleASM enabled, create the mapping table for all the luns in /dev/oracleasm/disks

List out Oracle ASM luns and Disk groups

Use kfod utility to capture the diskgroup details.

Setup your Oracle CRS/ASM environment.

Run kfod disk=all to list the ASM disks. In the below table, the path below identifies the label that was given to oracleasm createdisk command earlier. In the test bed, the lun name on the storage and on the ASM side are same, but need not be always same.

Table 9 Oracle ASM Disk Details

Disk	Size	Path	User	Group
1	20479 Mb	ORCL:DELMARCLUSDG_0000	oracle	dba
2	20479 Mb	ORCL:DELMARCLUSDG_0001	oracle	dba
3	20479 Mb	ORCL:DELMARCLUSDG_0002	oracle	dba
4	20479 Mb	ORCL:DELMARCLUSDG_0003	oracle	dba
5	20479 Mb	ORCL:DELMARCLUSDG_0004	oracle	dba
6	204796 Mb	ORCL:DSS_DATA_1	oracle	dba
7	204796 Mb	ORCL:DSS_DATA_10	oracle	dba
8	204796 Mb	ORCL:DSS_DATA_11	oracle	dba
9	204796 Mb	ORCL:DSS_DATA_12	oracle	dba
10	204796 Mb	ORCL:DSS_DATA_13	oracle	dba
11	204796 Mb	ORCL:DSS_DATA_14	oracle	dba
12	204796 Mb	ORCL:DSS_DATA_15	oracle	dba
13	204796 Mb	ORCL:DSS_DATA_16	oracle	dba
14	204796 Mb	ORCL:DSS_DATA_17	oracle	dba
15	204796 Mb	ORCL:DSS_DATA_18	oracle	dba
16	204796 Mb	ORCL:DSS_DATA_19	oracle	dba
17	204796 Mb	ORCL:DSS_DATA_2	oracle	dba
18	204796 Mb	ORCL:DSS_DATA_20	oracle	dba
19	204796 Mb	ORCL:DSS_DATA_21	oracle	dba
20	204796 Mb	ORCL:DSS_DATA_22	oracle	dba
21	204796 Mb	ORCL:DSS_DATA_23	oracle	dba

22	204796 Mb	ORCL:DSS_DATA_24	oracle	dba
23	204796 Mb	ORCL:DSS_DATA_25	oracle	dba
24	204796 Mb	ORCL:DSS_DATA_26	oracle	dba
25	204796 Mb	ORCL:DSS_DATA_27	oracle	dba
26	204796 Mb	ORCL:DSS_DATA_28	oracle	dba
27	204796 Mb	ORCL:DSS_DATA_29	oracle	dba
28	204796 Mb	ORCL:DSS_DATA_3	oracle	dba
29	204796 Mb	ORCL:DSS_DATA_30	oracle	dba
30	204796 Mb	ORCL:DSS_DATA_31	oracle	dba
31	204796 Mb	ORCL:DSS_DATA_32	oracle	dba
32	204796 Mb	ORCL:DSS_DATA_4	oracle	dba
33	204796 Mb	ORCL:DSS_DATA_5	oracle	dba
34	204796 Mb	ORCL:DSS_DATA_6	oracle	dba
35	204796 Mb	ORCL:DSS_DATA_7	oracle	dba
36	204796 Mb	ORCL:DSS_DATA_8	oracle	dba
37	204796 Mb	ORCL:DSS_DATA_9	oracle	dba
38	614398 Mb	ORCL:OLTP_DATA_1	oracle	dba
39	614398 Mb	ORCL:OLTP_DATA_10	oracle	dba
40	614398 Mb	ORCL:OLTP_DATA_11	oracle	dba
41	614398 Mb	ORCL:OLTP_DATA_12	oracle	dba
42	614398 Mb	ORCL:OLTP_DATA_13	oracle	dba
43	614398 Mb	ORCL:OLTP_DATA_14	oracle	dba
44	614398 Mb	ORCL:OLTP_DATA_15	oracle	dba
45	614398 Mb	ORCL:OLTP_DATA_16	oracle	dba
46	614398 Mb	ORCL:OLTP_DATA_2	oracle	dba
47	614398 Mb	ORCL:OLTP_DATA_3	oracle	dba
48	614398 Mb	ORCL:OLTP_DATA_4	oracle	dba
49	614398 Mb	ORCL:OLTP_DATA_5	oracle	dba
50	614398 Mb	ORCL:OLTP_DATA_6	oracle	dba
51	614398 Mb	ORCL:OLTP_DATA_7	oracle	dba
52	614398 Mb	ORCL:OLTP_DATA_8	oracle	dba
53	614398 Mb	ORCL:OLTP_DATA_9	oracle	dba
54	102398 Mb	ORCL:REDO01	oracle	dba
55	102398 Mb	ORCL:REDO02	oracle	dba
56	102398 Mb	ORCL:REDO03	oracle	dba
57	102398 Mb	ORCL:REDO04	oracle	dba
58	102398 Mb	ORCL:REDO05	oracle	dba
59	102398 Mb	ORCL:REDO06	oracle	dba
60	102398 Mb	ORCL:REDO07	oracle	dba
61	102398 Mb	ORCL:REDO08	oracle	dba

```
--------------------------------------------------
ORACLE_SID ORACLE_HOME
==============================
     +ASM1 /oracle/product/grid_home
     +ASM2 /oracle/product/grid_home
     +ASM3 /oracle/product/grid_home
     +ASM4 /oracle/product/grid_home
```

Map ASM LUNs with emc pseudo devices

Each of the above luns will be associated with emc pseudo power path device. Use Oracleasm querydisk to determine as below

```
[root@rac1 disks]# for i in *
> do
> emcname=`oracleasm querydisk -p $i | grep emcpower | awk -F ":|/" '{print
$3}'`
> echo "$i $emcname"
> done
```

This will give us the mapping between emcpower device and the ASM Lun. A sample is provided below.

DSS_DATA_1	emcpowerh1
DSS_DATA_10	emcpowerq1
DSS_DATA_11	emcpowerr1
DSS_DATA_12	emcpowers1
DSS_DATA_13	emcpowert1
DSS_DATA_14	emcpoweru1
DSS_DATA_15	emcpowerv1
DSS_DATA_16	emcpowerw1
DSS_DATA_17	emcpowerx1
DSS_DATA_18	emcpowery1

Document ASM Lun Headers data

```
[root@rac1 disks]# kfed read DSS_DATA_1 | egrep "provstr|dskname|grpname|fgname"
kfdhdb.driver.provstr:ORCLDISKDSS_DATA_1 ; 0x000: length=18
kfdhdb.dskname:               DSS_DATA_1 ; 0x028: length=10
kfdhdb.grpname:                    DSSDG ; 0x048: length=5
kfdhdb.fgname:                DSS_DATA_1 ; 0x068: length=10
kfed is oracle utility in CRS home directory. Hence set up your environment
before issuing kfed.
```

The above data will be handy if at all a mismatch happens with name of the emcpower device later. Kfed can be used to query the device and check whether it is the same lun for which it is being provisioned or not.

Storage LUNs and PowerPath pseudo device relationship

Run powermt display dev=all to relate the storage lun and the PowerPath device

Pseudo name=emcpowerac

```
VNX ID=APM00120902426 [rac1]
Logical device ID=600601605DB026004F6295C36DAFE111 [DSS_DATA22]
state=alive; policy=CLAROpt; queued-IOs=0
Owner: default=SP B, current=SP B       Array failover mode: 4
==============================================================================
--------------- Host ---------------   - Stor -  -- I/O Path --   -- Stats ---
###  HW Path               I/O Paths    Interf.  Mode     State   Q-IOs Errors
```

Identify the scsi ID's for the emcpower devices

Get the scsi Lun ID's on the system for each of the emcpower devices gathered above.

```
[root@rac1 ~]# /sbin/scsi_id -u -g -d /dev/emcpowerh1
3600601605db02600f47957226cafe111
```

Prepare a table of contents as below that can be used to build the udev file.

Lun Name	ASM Lun	EMC device	scsi id
DSS_DATA_1	DSS_DATA_1	/dev/ emcpowerh1	3600601605db02600f47957226cafe111
DSS_DATA_2	DSS_DATA_2	/dev/emcpoweri1	3600601605db02600f57957226cafe111

ASM disk string parameter

By default the parameter is null. Either issue gpnptool get or asmcmd dsget to query from CRS/ASM. As the disks reside in /dev/oracleasm/disks, the path in udev mapping has to be adjusted to /dev/oracleasm/disks too.

Create a copy of spfile and keep it handy too.

```
SQL> create pfile='?/dbs/init+ASM4.ora' from spfile;

File created.
```

Udev Migration

Change only on one node to make sure that it works fine, before propagating the change to other nodes.

Alter the asm_diskstring parameter with an alter system command to '/dev/oracleasm/disks' from null on one node saying sid='+ASM1';

crsctl disable crs on all the nodes so that it does not attempt to bring up the cluster during boot.

Shutdown CRS and ASM on all the nodes and make changes on one node.

```
Disable oracleasm,
/etc/init.d/oracleasm stop
/etc/init.d/oracleasm disable
```

Reboot and install the PowePath etc as mentioned in the beginning of this section.

```
cd /etc/udev/rules.d and create a file say 99-asmudev.rules.
```

Use the information gathered above to build the udev rules file.

```
KERNEL=="emcpowerh1", PROGRAM=="/sbin/scsi_id -u -g -d --whitelisted
--replace-whitespace --device=/dev/emcpowerh1",
RESULT=="3600601605db02600f47957226cafe111" OWNER="oracle", GROUP="dba",
MODE="660", NAME+="oracleasm/disks/DSS_DATA_1"
KERNEL=="emcpoweri1", PROGRAM=="/sbin/scsi_id -u -g -d --whitelisted
--replace-whitespace --device=/dev/emcpoweri1",
RESULT=="3600601605db02600f57957226cafe111" OWNER="oracle", GROUP="dba",
MODE="660", NAME+="oracleasm/disks/DSS_DATA_2"
KERNEL=="emcpowerj1", PROGRAM=="/sbin/scsi_id -u -g -d --whitelisted
--replace-whitespace --device=/dev/emcpowerj1",
RESULT=="3600601605db02600f67957226cafe111" OWNER="oracle", GROUP="dba",
MODE="660", NAME+="oracleasm/disks/DSS_DATA_3"
```

Either reboot the machine or issue start_udev to check the disks are visible under /dev/oracleasm/disks

Issue crsctl start crs to start the CRS resources and ASM.

start the databases on this node and do a sanity check.

Please note that asm_diskstring parameter is stored in spfile and also in profile.xml, locally on the host. Any mismatch will cause CRS voting not to come up.

${CRS_HOME}/gpnp/<node>/profiles/peer/profile.xml

Follow metalink notes 1077094.1 and 1410243.1 for further details on how to recover from such errors.

Repeat the above steps on the remaining nodes in the cluster, enable CRS, and reboot to complete udev setup.

Udev to ASMLIB

The process of migrating to ASMLIB is very similar to the above. After making the preparatory steps as above:

- Shutdown CRS and databases on all the nodes.
- Stop udev and rename the udev rules file for ASM. Reboot the host
- Install and configure oracleasm and issue oracleasm scan disks that should bring the disks in its default location /dev/oracleasm/disks.
- Restart the CRS, ASM and databases.
- Repeat the process on other nodes.

One issue observed was if the disk provstring(obtained from kfed read as documented above) does not match, the disks may not be visible.

kfdhdb.driver.provstr:ORCLDISKDSS_DATA_1 ; 0x000: length=18

In the above, ORCLDISK means it was created with oracleasm disk api. We had to issue oracleasm rename disk command in one of the setups to alter provstring as ORCLDISK+DISK NAME. The disks were discovered under mount point after this rename.

Appendix A: Cisco UCS Service Profiles

```
sj2-151-a19-A# show fabric-interconnect

Fabric Interconnect:
    ID   OOB IP Addr     OOB Gateway     OOB Netmask     Operability
    ---- --------------- --------------- --------------- -----------
    A    10.29.134.10    10.29.134.1     255.255.255.0   Operable
    B    10.29.134.11    10.29.134.1     255.255.255.0   Operable

sj2-151-a19-A# show fabric version
Fabric Interconnect A:
    Running-Kern-Vers: 5.0(3)N2(2.11)
    Running-Sys-Vers: 5.0(3)N2(2.11)
    Package-Vers: 2.1(1)A
    Startup-Kern-Vers: 5.0(3)N2(2.11)
```

```
    Startup-Sys-Vers: 5.0(3)N2(2.11)
    Act-Kern-Status: Ready
    Act-Sys-Status: Ready
    Bootloader-Vers:   v3.5.0(02/03/2011)

Fabric Interconnect B:
    Running-Kern-Vers: 5.0(3)N2(2.11)
    Running-Sys-Vers: 5.0(3)N2(2.11)
    Package-Vers: 2.1(1)A
    Startup-Kern-Vers: 5.0(3)N2(2.11)
    Startup-Sys-Vers: 5.0(3)N2(2.11)
    Act-Kern-Status: Ready
    Act-Sys-Status: Ready
    Bootloader-Vers:   v3.5.0(02/03/2011)



sj2-151-a19-A# show server inventory
Server  Equipped PID Equipped VID Equipped Serial (SN) Slot Status      Ackd
Memory (MB) Ackd Cores
------- ------------ ------------ -------------------- ----------------
---------------- ----------
1/1     B440-BASE-M2 V01          FCH16177D9Y          Equipped
262144          40
1/2                                                     Equipped Not Pri
1/3     B440-BASE-M2 V01          FCH161772YT          Equipped
262144          40
1/4                                                     Equipped Not Pri
1/5     UCSB-B200-M3 V01          FCH16197RE2          Equipped
262144           8
1/6     UCSB-B200-M3 V01          FCH16337DD4          Equipped
262144          16
1/7                                                     Empty
1/8                                                     Empty
2/1     B440-BASE-M2 V01          FCH1617730R          Equipped
262144          40
2/2                                                     Equipped Not Pri
2/3     B440-BASE-M2 V01          FCH161772V1          Equipped
262144          40
2/4                                                     Equipped Not Pri
2/5     UCSB-B200-M3 V01          FCH16207MXA          Equipped
262144          16
2/6                                                     Empty
2/7                                                     Empty
2/8                                                     Empty



sj2-151-a19-A(nxos)# show interface port-channel 10 brief

--------------------------------------------------------------------------------
Port-channel VLAN  Type Mode   Status  Reason                    Speed  Protocol
Interface
--------------------------------------------------------------------------------
Po10         1     eth  trunk  up      none                       a-10G(D)  lacp
```

```
sj2-151-a19-A# show service-profile inventory
Service Profile Name Type              Server  Assignment Association
-------------------- ----------------- ------- ---------- -----------
B200_M3              Instance          1/5     Assigned   Associated
B200_M3_LSI          Instance          2/5     Assigned   Associated
B200M2_VMW           Instance                  Unassigned Unassociated
B230_OVM             Instance                  Unassigned Unassociated
b440_1_clone         Instance          1/1     Assigned   Associated
b440_2_clone         Instance          1/3     Assigned   Associated
b440_3_clone         Instance          2/1     Assigned   Associated
b440_4_clone         Instance          2/3     Assigned   Associated
ORARAC               Initial Template          Unassigned Unassociated
ORARAC_B440_1        Instance                  Unassigned Unassociated
ORARAC_B440_2        Instance                  Unassigned Unassociated
ORARAC_B440_3        Instance                  Unassigned Unassociated
ORARAC_B440_4        Instance                  Unassigned Unassociated
```

```
Service Profile Name: b440_1_clone
Type: Instance
Server: 1/1
Description:
Assignment: Assigned
Association: Associated
Power State: On
Op State: Ok
Oper Qualifier: N/A
Conf State: Applied
Config Qual: N/A
Current Task:
    Server 1/1:
        Overall Status: Ok
        Operability: Operable
        Oper Power: On

        Adapter 1:
            Threshold Status: N/A
            Overall Status: Operable
            Operability: Operable
            Power State: On
            Thermal Status: N/A
            Voltage Status: N/A

        Adapter 2:
            Threshold Status: N/A
            Overall Status: Operable
            Operability: Operable
            Power State: On
            Thermal Status: N/A
            Voltage Status: N/A

        Motherboard:
            Threshold Status: OK
            Overall Status: N/A
```

```
                Operability: N/A
                Oper Power: On
                Power State: Ok
                Thermal Status: OK
                Voltage Status: OK
                CMOS Battery Voltage Status: Ok
                Mother Board Power Usage Status: Ok

                Motherboard Temperature Statistics:
                    Motherboard Front Temperature (C): N/A
                    Motherboard Rear Temperature (C): N/A

                Memory Array 1:
                    Threshold Status: N/A
                    Overall Status: N/A
                    Operability: N/A
                    Power State: N/A
                    Thermal Status: N/A
                    Voltage Status: N/A

                    DIMMs:

                    DIMM Threshold Status Overall Status  Operability     Power
State   Thermal Status  Voltage Status
                    ---- ---------------- --------------- ---------------
------------- --------------- --------------
                      1 N/A              Operable        Operable        N/A
OK              N/A
                      2 N/A              Operable        Operable        N/A
OK              N/A
                      3 N/A              Operable        Operable        N/A
OK              N/A
                      4 N/A              Operable        Operable        N/A
OK              N/A
                      5 N/A              Operable        Operable        N/A
OK              N/A
                      6 N/A              Operable        Operable        N/A
OK              N/A
                      7 N/A              Operable        Operable        N/A
OK              N/A
                      8 N/A              Operable        Operable        N/A
OK              N/A
                      9 N/A              Operable        Operable        N/A
OK              N/A
                     10 N/A              Operable        Operable        N/A
OK              N/A
                     11 N/A              Operable        Operable        N/A
OK              N/A
                     12 N/A              Operable        Operable        N/A
OK              N/A
                     13 N/A              Operable        Operable        N/A
OK              N/A
                     14 N/A              Operable        Operable        N/A
OK              N/A
                     15 N/A              Operable        Operable        N/A
OK              N/A
                     16 N/A              Operable        Operable        N/A
OK              N/A
```

```
                 17 N/A              Operable        Operable        N/A
OK               N/A
                 18 N/A              Operable        Operable        N/A
OK               N/A
                 19 N/A              Operable        Operable        N/A
OK               N/A
                 20 N/A              Operable        Operable        N/A
OK               N/A
                 21 N/A              Operable        Operable        N/A
OK               N/A
                 22 N/A              Operable        Operable        N/A
OK               N/A
                 23 N/A              Operable        Operable        N/A
OK               N/A
                 24 N/A              Operable        Operable        N/A
OK               N/A
                 25 N/A              Operable        Operable        N/A
OK               N/A
                 26 N/A              Operable        Operable        N/A
OK               N/A
                 27 N/A              Operable        Operable        N/A
OK               N/A
                 28 N/A              Operable        Operable        N/A
OK               N/A
                 29 N/A              Operable        Operable        N/A
OK               N/A
                 30 N/A              Operable        Operable        N/A
OK               N/A
                 31 N/A              Operable        Operable        N/A
OK               N/A
                 32 N/A              Operable        Operable        N/A
OK               N/A

            CPU 1:
                Threshold Status: N/A
                Overall Status: Operable
                Operability: Operable
                Power State: N/A
                Thermal Status: OK
                Voltage Status: N/A

            CPU 2:
                Threshold Status: N/A
                Overall Status: Operable
                Operability: Operable
                Power State: N/A
                Thermal Status: OK
                Voltage Status: N/A

            CPU 3:
                Threshold Status: N/A
                Overall Status: Operable
                Operability: Operable
                Power State: N/A
                Thermal Status: OK
                Voltage Status: N/A

            CPU 4:
```

```
                    Threshold Status: N/A
                    Overall Status: Operable
                    Operability: Operable
                    Power State: N/A
                    Thermal Status: OK
                    Voltage Status: N/A
```

Appendix B: N5K Zone Definitions

N5K-A

```
zoneset name ORARAC_FI_A vsan 15
  zone name orarac1_hba1 vsan 15
  * fcid 0x3a01ef [pwwn 50:06:01:62:47:20:2c:af] [A2P2]
  * fcid 0x3a02ef [pwwn 50:06:01:68:47:20:2c:af] [B2P0]
  * fcid 0x3a03ef [pwwn 50:06:01:6a:47:20:2c:af] [B2P2]
  * fcid 0x3a0022 [pwwn 20:00:00:25:b5:00:00:1f]
  * fcid 0x3a00ef [pwwn 50:06:01:60:47:20:2c:af] [A2P0]

  zone name orarac1_hba3 vsan 15
  * fcid 0x3a0023 [pwwn 20:00:00:25:b5:00:00:3f]
  * fcid 0x3a00ef [pwwn 50:06:01:60:47:20:2c:af] [A2P0]
  * fcid 0x3a01ef [pwwn 50:06:01:62:47:20:2c:af] [A2P2]
  * fcid 0x3a02ef [pwwn 50:06:01:68:47:20:2c:af] [B2P0]
  * fcid 0x3a03ef [pwwn 50:06:01:6a:47:20:2c:af] [B2P2]

  zone name orarac2_hba1 vsan 15
  * fcid 0x3a0025 [pwwn 20:00:00:25:b5:00:00:1e]
  * fcid 0x3a00ef [pwwn 50:06:01:60:47:20:2c:af] [A2P0]
  * fcid 0x3a01ef [pwwn 50:06:01:62:47:20:2c:af] [A2P2]
  * fcid 0x3a02ef [pwwn 50:06:01:68:47:20:2c:af] [B2P0]
  * fcid 0x3a03ef [pwwn 50:06:01:6a:47:20:2c:af] [B2P2]

  zone name orarac4_hba3 vsan 15
  * fcid 0x3a00ef [pwwn 50:06:01:60:47:20:2c:af] [A2P0]
  * fcid 0x3a01ef [pwwn 50:06:01:62:47:20:2c:af] [A2P2]
  * fcid 0x3a02ef [pwwn 50:06:01:68:47:20:2c:af] [B2P0]
  * fcid 0x3a03ef [pwwn 50:06:01:6a:47:20:2c:af] [B2P2]
  * fcid 0x3a0021 [pwwn 20:00:00:25:b5:00:00:3c]

  zone name orarac3_hba1 vsan 15
  * fcid 0x3a00ef [pwwn 50:06:01:60:47:20:2c:af] [A2P0]
  * fcid 0x3a01ef [pwwn 50:06:01:62:47:20:2c:af] [A2P2]
  * fcid 0x3a02ef [pwwn 50:06:01:68:47:20:2c:af] [B2P0]
  * fcid 0x3a03ef [pwwn 50:06:01:6a:47:20:2c:af] [B2P2]
  * fcid 0x3a0024 [pwwn 20:00:00:25:b5:00:00:1d]

  zone name orarac3_hba3 vsan 15
  * fcid 0x3a00ef [pwwn 50:06:01:60:47:20:2c:af] [A2P0]
  * fcid 0x3a01ef [pwwn 50:06:01:62:47:20:2c:af] [A2P2]
  * fcid 0x3a02ef [pwwn 50:06:01:68:47:20:2c:af] [B2P0]
  * fcid 0x3a03ef [pwwn 50:06:01:6a:47:20:2c:af] [B2P2]
```

```
  * fcid 0x3a0027 [pwwn 20:00:00:25:b5:00:00:3d]

  zone name orarac4_hba1 vsan 15
  * fcid 0x3a00ef [pwwn 50:06:01:60:47:20:2c:af] [A2P0]
  * fcid 0x3a01ef [pwwn 50:06:01:62:47:20:2c:af] [A2P2]
  * fcid 0x3a02ef [pwwn 50:06:01:68:47:20:2c:af] [B2P0]
  * fcid 0x3a03ef [pwwn 50:06:01:6a:47:20:2c:af] [B2P2]
  * fcid 0x3a0020 [pwwn 20:00:00:25:b5:00:00:1c]

  zone name orarac2_hba3 vsan 15
  * fcid 0x3a0026 [pwwn 20:00:00:25:b5:00:00:3e]
  * fcid 0x3a00ef [pwwn 50:06:01:60:47:20:2c:af] [A2P0]
  * fcid 0x3a01ef [pwwn 50:06:01:62:47:20:2c:af] [A2P2]
  * fcid 0x3a02ef [pwwn 50:06:01:68:47:20:2c:af] [B2P0]
  * fcid 0x3a03ef [pwwn 50:06:01:6a:47:20:2c:af] [B2P2]
```

N5K-B

```
sj2-151-a19-n5k-FI-B(config)# show zones active
zoneset name ORARAC_FI_B vsan 15
  zone name orarac1_hba2 vsan 15
  * fcid 0xe50023 [pwwn 20:00:00:25:b5:00:00:0f]
  * fcid 0xe500ef [pwwn 50:06:01:64:47:20:2c:af] [A3P0]
  * fcid 0xe501ef [pwwn 50:06:01:66:47:20:2c:af] [A3P2]
  * fcid 0xe502ef [pwwn 50:06:01:6c:47:20:2c:af] [B3P0]
  * fcid 0xe503ef [pwwn 50:06:01:6e:47:20:2c:af] [B3P2]

  zone name orarac1_hba4 vsan 15
  * fcid 0xe50024 [pwwn 20:00:00:25:b5:00:00:2f]
  * fcid 0xe500ef [pwwn 50:06:01:64:47:20:2c:af] [A3P0]
  * fcid 0xe501ef [pwwn 50:06:01:66:47:20:2c:af] [A3P2]
  * fcid 0xe502ef [pwwn 50:06:01:6c:47:20:2c:af] [B3P0]
  * fcid 0xe503ef [pwwn 50:06:01:6e:47:20:2c:af] [B3P2]

  zone name orarac3_hba2 vsan 15
  * fcid 0xe500ef [pwwn 50:06:01:64:47:20:2c:af] [A3P0]
  * fcid 0xe501ef [pwwn 50:06:01:66:47:20:2c:af] [A3P2]
  * fcid 0xe502ef [pwwn 50:06:01:6c:47:20:2c:af] [B3P0]
  * fcid 0xe503ef [pwwn 50:06:01:6e:47:20:2c:af] [B3P2]
  * fcid 0xe50020 [pwwn 20:00:00:25:b5:00:00:0d]

  zone name orarac4_hba2 vsan 15
  * fcid 0xe500ef [pwwn 50:06:01:64:47:20:2c:af] [A3P0]
  * fcid 0xe501ef [pwwn 50:06:01:66:47:20:2c:af] [A3P2]
  * fcid 0xe502ef [pwwn 50:06:01:6c:47:20:2c:af] [B3P0]
  * fcid 0xe503ef [pwwn 50:06:01:6e:47:20:2c:af] [B3P2]
  * fcid 0xe50021 [pwwn 20:00:00:25:b5:00:00:0c]

  zone name orarac4_hba4 vsan 15
  * fcid 0xe500ef [pwwn 50:06:01:64:47:20:2c:af] [A3P0]
  * fcid 0xe501ef [pwwn 50:06:01:66:47:20:2c:af] [A3P2]
  * fcid 0xe502ef [pwwn 50:06:01:6c:47:20:2c:af] [B3P0]
  * fcid 0xe503ef [pwwn 50:06:01:6e:47:20:2c:af] [B3P2]
  * fcid 0xe50022 [pwwn 20:00:00:25:b5:00:00:2c]
```

```
zone name orarac3_hba4 vsan 15
* fcid 0xe500ef [pwwn 50:06:01:64:47:20:2c:af] [A3P0]
* fcid 0xe501ef [pwwn 50:06:01:66:47:20:2c:af] [A3P2]
* fcid 0xe50027 [pwwn 20:00:00:25:b5:00:00:2d]
* fcid 0xe502ef [pwwn 50:06:01:6c:47:20:2c:af] [B3P0]
* fcid 0xe503ef [pwwn 50:06:01:6e:47:20:2c:af] [B3P2]

zone name orarac2_hba2 vsan 15
* fcid 0xe50025 [pwwn 20:00:00:25:b5:00:00:0e]
* fcid 0xe500ef [pwwn 50:06:01:64:47:20:2c:af] [A3P0]
* fcid 0xe501ef [pwwn 50:06:01:66:47:20:2c:af] [A3P2]
* fcid 0xe502ef [pwwn 50:06:01:6c:47:20:2c:af] [B3P0]
* fcid 0xe503ef [pwwn 50:06:01:6e:47:20:2c:af] [B3P2]

zone name orarac2_hba4 vsan 15
* fcid 0xe50026 [pwwn 20:00:00:25:b5:00:00:2e]
* fcid 0xe500ef [pwwn 50:06:01:64:47:20:2c:af] [A3P0]
* fcid 0xe501ef [pwwn 50:06:01:66:47:20:2c:af] [A3P2]
* fcid 0xe502ef [pwwn 50:06:01:6c:47:20:2c:af] [B3P0]
* fcid 0xe503ef [pwwn 50:06:01:6e:47:20:2c:af] [B3P2]
```

Appendix C: Oracle spfile PARAMETERS

Linux Huge Pages were setup on each host

```
ASM
asm_diskgroups='OLTPDG','DSSDG','REDODG'
asm_power_limit=1
memory_target=1023M
large_pool_size=12M
sessions=400

OLTP
sga_max_size=128G
sga_target=128G
db_name='oltp'
cluster_database=TRUE
log_buffer= 183623680
processes=3000

DSS
sga_max_size=64G
sga_target=64G
db_name='dss'
cluster_database=TRUE
processes=3000
```