Oracle Database 11*g* Release 2 Enterprise Edition using Oracle Real Application Clusters on IBM Power Systems with AIX 6.1

Ravisankar Shanmugam
IBM Oracle International Competency Center
May 2010

Table of Contents

Table of Contents **2**
Abstract **1**
Prerequisites **1**
Introduction **1**
Oracle Database 11*g* Release 2 new features 1
High Availability 1
Performance and scalability 2
Security 2
Clustering 2
Manageability 2
About Oracle Real Application Clusters 11*g* Release 2 3
About IBM Power Systems 3
About IBM System Storage DS6000 5
Hardware requirements **6**
Oracle Real Application Clusters requirements 6
Server CPU 7
Server memory 7
Network 7
Shared storage 8
High availability considerations 9
Software requirements **9**
Operating system 9
Storage System Manager 10
Subsystem Device Driver Path Control Module (SDDPCM) 10
Oracle Database 11*g* Release 2 10
Automatic Storage Management 11
Configuring the system environment **11**
Virtual IO Server (VIOS) and Logical Partitions (LPARs) 11
Hardware Management Console (HMC) 13
Installing the AIX operating system 14
Installing Oracle Grid Infrastructure 11.2.0.1 ..... **16**
Pre-Installation tasks 16
Configuring OS kernel parameters 16
Configuring network parameters 17
Network time protocol 17
Creating users and groups 19
Setting Oracle inventory location 19
Setting up network files 20

Configuring SSH on all cluster nodes 20
Configuring shared disks for OCR, voting and database 20
Running Cluster Verification Utility (CVU) 21
Performing Oracle Clusterware installation and Automatic Storage Management installation 28
Performing post-installation tasks 41
Installing Oracle Database 11*g* Release 2 (11.2.0.1) 43
Pre-Installation tasks 43
Running Cluster Verification Utility 43
Preparing Oracle home and its path 44
Performing database installation 45
Post-installation tasks 54
Summary 55
References 56
Oracle documentation 56
IBM documentation 56
IBM and Oracle Web sites 56
About the author 58
Appendix A: List of common abbreviations and acronyms 59
Trademarks and special notices 60

Abstract

The purpose of this white paper is to assist those who are implementing Oracle Database 11g Release 2 with Oracle Real Application Clusters (RAC) on IBM Power Systems™ servers with AIX® 6.1 and IBM System Storage™ disk systems products. The information provided herein is based on experiences with test environments at the IBM Oracle International Competency Center and is based on available documentation from IBM and Oracle.

This paper does not cover the installation of AIX, the Virtual I/O Server (VIOS) or the IBM Systems Storage™ Management Software used to configure the IBM System Storage DS6800 used in our tests.

Prerequisites

- Good knowledge of Oracle Database
- Knowledge of the AIX, Virtual IO Server and IBM Systems Storage™.

Introduction

This white paper will discuss the necessary steps to prepare the AIX nodes with shared disks for installing Oracle Grid Infrastructure 11*g* Release 2 and Oracle Database 11*g* Release 2 with RAC.

An implementation of Oracle Real Application Clusters consists of three main steps:

1. Planning the hardware for Oracle Real Application Clusters implementation
2. Configuring the servers and storage disk systems
3. Installing and configuring Oracle Grid Infrastructure 11*g* Release2 and Oracle Database 11*g* Release 2 with RAC

Oracle Database 11*g* Release 2 new features

There are many new features found in Oracle Database 11*g* Release 2. They can be found in Oracle 11*g* Release 2 documentation available on the Oracle web site. According to *Oracle Database New Features Guide 11g Release 2*, the main highlights are as follows:

High Availability

- Automatically repair corrupted blocks on the primary database or physical standby database (which must be in real-time query mode).
- As part of Oracle Cloud Computing offering, databases can be backed up to Amazon S3.
- With connection to a catalog and auxiliary database, DUPLICATE command in RMAN can be executed without any connection to the target database.
- Tables with compression are supported in logical standby databases and Oracle LogMiner.
- A primary database can support up to 30 standby databases.

- Whenever the host computer restarts, Oracle Restart will automatically restart the database instance, the ASM instance, the listener, and other components. Oracle Restart is a separate installation from Oracle Database.

Performance and scalability

- Oracle RAC has integrated with Universal Connection Pool (UCP) which is the new Java™ connection pool. With UCP, Java applications can easily manage connections to an Oracle RAC database.
- UCP for JDBC enhances performance and stabilization, and provides connection labeling and harvesting.
- Database Smart Flash Cache is a transparent extension of the database buffer cache using solid state device (SSD) technology. This SSD acts as a Level 2 cache to the SGA (Level 1). SSD can reduce the amount of disk I/O at a much lower cost than adding same amount of memory.
- Oracle ASM can migrate a disk group with 512 byte sector drives to 4 KB sector drives.
- Oracle RAC One Node is a new option to Oracle Database 11*g* Release 2 Enterprise Edition. It can easily upgrade to a full multi-node Oracle RAC database without downtown or disruption.

Security

- New encryption key management can update the master key associated with transparent data encryption (TDE) encrypted tablespaces.
- New package for audit data management can clean up audit trail records after backup and control the size and age of the audit files.

Clustering

- Oracle Universal Installer has integrated with the Cluster Verification Utility (CVU) in the pre-installation steps of Oracle RAC installation.
- In order to have successful installation of Oracle RAC, a synchronized system time across the cluster is a requirement. Cluster Time Service will be responsible to synchronize the system time on all nodes in the cluster.
- The high redundancy option for storing OCR has increased to 5 copies so as to improve the cluster availability.
- OCR can now be stored in Automatic Storage Management (ASM).
- Oracle Clusterware is installed into a separate home from Oracle Database home.

Manageability

- Single Client Access Name (SCAN) provides a single name for clients to access an Oracle Database running a cluster. It provides load balancing and failover of client connections to the database.
- The Clusterware administrator can delegate specific tasks on specific servers to different people based on their roles. This is called role-separated management.
- Patch sets for Oracle Clusterware and Oracle RAC can be applied to the servers as out-of-place upgrades to the Oracle Grid infrastructure without bringing the entire cluster down.
- The new Enterprise Manager GUI can monitor and manage the full lifecycle of Oracle Clusterware resources. It also introduces procedures to scale up or scale down Oracle Clusterware and Oracle Real Application Clusters easily.

- Complete deinstallation and deconfiguration of Oracle RAC databases and listeners can be done by Database Configuration Assistant (DBCA), Database Upgrade Assistant (DBUA), and Net Configuration Assistant (NETCA).
- Oracle Universal Installer can help to clean up a failed Oracle Clusterware Installation by advising you the places to clean and steps to change prior to reattempting the installation again. During the installation, it also consists of several recovery points for you to retry and rollback to the closet recovery point once the problem has been fixed.
- Database administrator can limit Oracle instance's CPU usage by setting the CPU_COUNT initialization parameter. This is called Instance Caging.
- E-mail notifications can be sent to users on any job activities.

About Oracle Real Application Clusters 11*g* Release 2

Oracle Real Application Clusters (RAC) is an option of Oracle Database that allows a database to be installed across multiple servers (RAC nodes). According to Oracle, RAC uses the shared disk method of clustering databases. Oracle processes running in each node access the same data residing on shared data disk storage. Oracle RAC uses a "shared everything" data architecture. This means that all data storage needs to be globally available to all RAC nodes. First introduced with Oracle Database 9*i*, RAC provides high availability and flexible scalability. If one of the clustered nodes fails, Oracle continues processing on the other nodes. If additional capacity is needed, nodes can be dynamically added without taking down the cluster.

In Oracle Database 11*g* Release 2, Oracle provides Oracle Clusterware, which is designed specifically for Oracle RAC. You do not need a third party Clusterware product to implement Oracle RAC. Since storage is shared, the file system and volume management must be cluster-aware.

Starting with Oracle Database 11*g* Release 2, Oracle Clusterware files can be stored in Oracle ASM. Oracle Clusterware and Oracle ASM are installed into a single home directory called grid home.

About IBM Power Systems

The IBM Power 570 mid-range server with POWER6 and available POWER6+ processor cards delivers outstanding price/performance, mainframe-inspired reliability and availability features, flexible capacity upgrades, and innovative virtualization technologies to enable management of growth, complexity, and risk. The Power 570 leverages your existing investments by supporting AIX, IBM i, and Linux for Power, and x86 Linux applications on a single server. It is available in 2-, 4-, 8-, 12-, and 16-core and 32-core configurations. As with the p5 570, the POWER6-based 570s modular symmetric multiprocessor (SMP) architecture is constructed using 4U (EIA units), 4-core or 8-core building block modules (also referred to as nodes, or CECs). Each of these nodes supports four POWER6 3.5, 4.2 or 4.7 GHz dual-core processors, and new POWER6 4.2 GHz dual-core processors, or POWER6+ 4.4, and 5.0 GHz four-core processors along with cache, memory, media, disks, I/O adapters, and power and cooling to create a balanced, extremely high-performance rack-mount system.

This design allows up to four modules to be configured in a 19-inch rack as a single SMP server, allowing clients to start with what they need and grow by adding additional building blocks. A fully configured 570

server may consist of 32 processor cores, 768 GB of DDR2 memory, four media bays, integrated ports for attaching communications devices, 24 mixed PCI-X and PCI Express adapter slots, and 24 internal SAS (Serial Attached SCSI) drives accommodating up to 7.2 TB of internal disk storage. The 64-bit POWER6 processors in this server are integrated into a dual-core single chip module and a dual-core dual chip module, with 32 MB of L3 cache, 8 MB of L2 cache, and 12 DDR2 memory DIMM slots. The unique DDR2 memory uses a new memory architecture to provide greater bandwidth and capacity. This enables operating at a higher data rate for large memory configurations. Each new processor card can support up to 12 DDR2 DIMMs running at speeds of up to 667 MHz.

As with the POWER5™ processor, simultaneous multithreading enabling two threads to be executed at the same time on a single processor core is a standard feature of POWER6 technology. Introduced with the POWER6 processor design is hardware decimal floating-point support improving the performance of the basic mathematical calculations of financial transactions that occur regularly on today's business computers. The POWER6 processor also includes an AltiVec SIMD accelerator, which helps to improve the performance of high performance computing (HPC) workloads.

All Power Systems servers can utilize logical partitioning (LPAR) technology implemented using System p virtualization technologies, the operating system (OS), and a hardware management console (HMC). Dynamic LPAR allows clients to dynamically allocate many system resources to application partitions without rebooting, allowing up to 16 dedicated processor partitions on a fully configured system. In addition to the base virtualization that is standard on every System p server, two optional virtualization features are available on the server: PowerVM Standard Edition (formerly Advanced POWER Virtualization (APV) Standard) and PowerVM Enterprise Edition (formerly APV Enterprise).

PowerVM Standard Edition includes IBM Micro-Partitioning™ and Virtual I/O Server (VIOS) capabilities. Micro-partitions can be defined as small as 1/10th of a processor and be changed in increments as small as 1/100th of a processor. Up to 160 micro-partitions may be created on a 16-core 570 system. VIOS allows for the sharing of disk and optical devices and communications and Fibre Channel adapters. Also included is support for Multiple Shared Processor Pools and Shared Dedicated Capacity.PowerVM Enterprise Edition includes all features of PowerVM Standard Edition plus Live Partition Mobility, newly available with POWER6 systems. It is designed to allow a partition to be relocated from one server to another while end users are using applications running in the partition. Other features introduced with POWER6 processor-based technology include an Integrated Virtual Ethernet adapter standard with every system, the Processor Instruction Retry feature automatically monitoring the POWER6 processor and, if needed, restarting the processor workload without disruption to the application, and a new HMC (Hardware Management Console) graphical user interface offering enhanced systems control.

Figure 1: IBM Power System p570 with four Modules in a rack

Figure 2: IBM Power System p570 one Module

About IBM System Storage DS6000

The IBM System Storage DS6800 is a DS6000 series of Storage server, flexible, high-performance storage for medium and large enterprises. It is an innovative storage system designed to provide high availability and high performance in a small, space-saving, power-efficient modular package. This series, along with the DS8000 series, offers an enterprise-class continuum of storage systems with shared replication services and common management interfaces.

As part of the IBM System Storage DS® Family, the DS6800 is designed to provide medium and large businesses with a low-cost, enterprise-class storage solution to help simplify data management and to provide comprehensive data protection and recovery capabilities and easy scalability for both mainframe and open system storage needs.

The IBM System Storage DS6800 is designed to provide over 1600 MBps performance for high throughput applications.

The DS6800 can help simplify IT infrastructure by supporting a wide range of servers, both mainframe and open systems, including IBM Power Systems, System x, System z®, and non-IBM platforms running UNIX®, Linux®, and Windows® operating systems.

Figure 3: IBM System Storage DS6800

The DS6800 is designed to help consolidate server storage into a centrally-managed, shared or storage area network (SAN) environment. With its modular design, the DS6800 system can scale from 292GB up to 57.6TBof physical storage capacity by adding storage expansion enclosures, each of which can contain up to 16 hard disk drives (HDD). Non-disruptive storage capacity expansion helps businesses maintain high data availability while accommodating rapid data growth.

The DS6800 features enterprise class data backup and disaster recovery capabilities. IBM FlashCopy® can create point-in-time copies of data that allow users to have nearly instantaneous access to information on both the source and target volumes. IBM System Storage Metro Mirror and Global Mirror options can generate and maintain data-consistent copies of data on separate storage systems installed either locally or at a geographically dispersed location. These functions are designed to help protect data and to provide failover and failback capabilities to support business continuance strategies and operations.

As shown in Figure 3, IBM System Storage DS4800 has 2 U rack-mount enclosures with 12 easily accessible drive bays. It supports dual-ported and hot-swappable SAS disk at 10,000 and 15,000 rpm speeds. It is also scalable to 3.6 TB of storage capacity with 300 GB hot-swappable SAS disks.

Hardware requirements

Oracle Real Application Clusters requirements

An Oracle Real Application Clusters Database environment consists of the following components:

- Cluster nodes - 2 to *n* nodes or hosts, running Oracle Database server(s)
- Network interconnect - a private network used for cluster communications and Cache Fusion
- Shared storage - used to hold database's system and data files and accessed by the cluster nodes
- Production network - used by clients and application servers to access the database

Figure 4 below is an architecture diagram for Oracle Real Application Clusters:

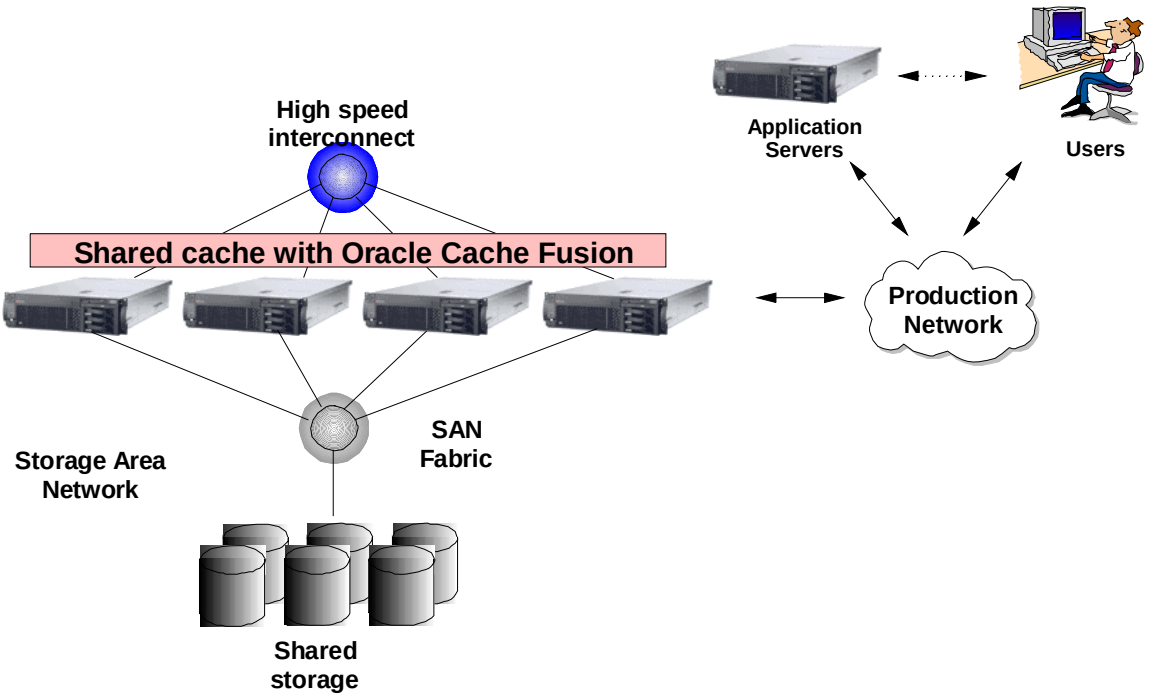


Figure 4: Oracle Real Application Clusters architecture

Server CPU

There should be enough server CPU capacity in terms of speed and number of CPU's to handle the workload. Generally speaking, there should be enough CPU capacity to have an average CPU utilization of 65%. This will allow the server absorb peak activity more easily.

Server memory

An Oracle Database may require a lot of memory that depends on the activity level of users and the nature of the application or workload. As a rule of thumb, the server should have more memory than it actually uses because performance will be greatly degraded, heavy disk swapping and node eviction may occur when there is insufficient of memory.

It is important to select servers that are available with the amount of memory required plus room for growth. Memory utilization should be around 75-85% maximum of the physical memory in production environment. Otherwise, heavy disk swapping may occur and server performance will decrease.

Network

Servers in an Oracle Real Application Clusters need at least two separate networks, a public network and a private network. The public network is used for the communication between the clients or applications servers and the database. The private network, sometimes referred to as "network interconnect" is used for cluster node communication. It is used for monitoring the heartbeat of the

cluster and by Oracle Real Application Clusters for Cache Fusion. At least 2 physical or Logical or virtual Ethernet adapters are needed on each of the RAC nodes, one for public network and another one of private RAC interconnection.

IBM Power Systems POWER6 and POWER7 processor-based systems offers Integrated Virtual Ethernet adapter (IVE), which gives integrated High-speed Ethernet Adapter ports (Host Ethernet Adapter (HEA)) with hardware assisted virtualization capabilities. IVE also includes special hardware features that provides the logical Ethernet adapter or otherwise called Logical Host Ethernet Adapter (LHEA). These LHEA adapters can directly assigned to the LPARs without configuring through the POWER Hypervisor (PHYP). This eliminates the need to move the pockets between Logical Partitions (LPARs) through the Shared Ethernet Adapter (SEA). IVE replaces the need of virtual Ethernet and SEA in Virtual IO server (VIOS) environment and LPARs can share the HEA ports with improved performance.

InfiniBand networking for Oracle RAC interconnecting is supported with Oracle Database 11*g* on AIX.

Shared storage

Shared storage for Oracle Real Application Clusters devices can be logical drives or LUNs from a Storage Area Network (SAN) controller or a Network File System (NFS) from a supported Network Attached Storage (NAS) device. IBM sells NAS products such as IBM System Storage N3000, N3700, N5000 and N7000.

For SAN products, IBM offers enterprise disk systems such as DS6000™ and DS8000®, mid-range disk systems such as DS3400, DS4000 and DS5000 series. Check to ensure the System Storage product you are using is supported with Oracle Real Application Clusters implementations. Third party storage subsystem can also be used with AIX servers. Please refer to third party documentation or contact a third party representative for product certification information.

To use a shared file system for Oracle Clusterware, Oracle ASM and Oracle RAC Database files, the file system must comply with the following requirements:.

1. A certified cluster file system is required.

This is a file system that may be accessed (read and write) by all members in a cluster at the same time, with all cluster members having the same view of the file system. It allows all nodes in a cluster to access a device concurrently via the standard file system interface. IBM General Parallel File System (GPFS version 3.2.1.8 or later) is an example. GPFS can be used for placing shared Oracle Home for Grid Infrastructure software files (Clusterware and ASM) and database software and database files.

2. Oracle Automatic Storage Management (ASM)

ASM is a simplified database storage management and provisioning system that provides file system and volume management capabilities in Oracle. It allows database administrators (DBA) to reference disk groups instead of individual disks and files which ASM manages internally. ASM is included in Oracle Database 11*g* and is designed to handle Oracle Database files, control files and log files.

In Oracle 11*g* Release 2, Oracle Automatic Storage Management Cluster File System (Oracle ACFS) is introduced. It is a multi-platform, scalable file system which supports database and application files like executables, database trace files, database alert logs, application reports, BFILEs, and

configuration files. However, it does not support any file that can be directly stored in Oracle ASM as well as any files for the Oracle grid infrastructure home.

In the lab test for Oracle Database 11*g* Release 2 with RAC on AIX, voting, OCR disks and database files are created in ASM disk groups. The Oracle Grid Infrastructure software files and Oracle Database software files are placed in the local file systems (JFS2).

High availability considerations

High availability (HA) is a key requirement for many clients. From a hardware configuration standpoint, this means eliminating single point of failure. IBM products are designed for high availability, with such standard features as redundant power supplies and cooling fans, hot-swappable components, and so on.

For high availability environments, the following recommendations should also be taken into consideration when selecting the server:

1. Configure additional network interfaces and use AIX Etherchannel to combine at least two network interfaces for each of the two Oracle RAC networks. This reduces the downtime due to a network interface card (NIC) failure or network component failure. Multi-port adapters provide network path redundancy, however the adapter will be a single point of failure. In this case, redundant multi-port adapters are the best solution. In addition, NICs used for Etherchannel should be on separate physical network cards and connected to different network switches.
2. There should be at least two fibre channel host bus adapters (HBA) on each node to provide redundant I/O paths to the storage subsystem. Multi-port HBAs and Storage Area Network (SAN) with redundant components like SAN switches and cabling will provide higher availability of the servers.

Finally, an Oracle RAC implementation requires at least two network interfaces. Nevertheless, up to four network interfaces are recommended, two for public, two for private. The more redundancy of hardware architectures and software components, the less downtime databases and applications will experience.

Software requirements

In an Oracle Real Application Clusters implementation, there are additional AIX file sets need to be installed in the cluster nodes. A few of them are optional and may not required in the RAC nodes. If the optional file sets are missing the Cluster Verification tool may show a failure report, which can be ignored.

Operating system

AIX 6.1 (6100-04-03-1009) is the operating system used in the tests described in this paper.

Storage System Manager

IBM System Storage DS6000 Storage Manager is used to manage the DS6800 via the graphical user interface. The DS6000 Storage Manager Host software is required for managing the DS6800 with controller firmware version 06.17.xx.xx. IBM System Storage DS6800 also can be managed by the Command Line interface tool DSCLI.

The DS6000 Storage Manager comes along with the DS6000 Microcode bundle, which can be downloaded from the IBM DS Storage Systems support Web site. It needs a registered Used ID and Password.

The IBM DS Storage Manager Software packages are available for AIX, Microsoft® Windows (32-bit and 64-bit version), Linux, and other platforms.

Subsystem Device Driver Path Control Module (SDDPCM)

The SDDPCM is loadable path control module designed to support the multipath configuration environment in the IBM System Storage Enterprise Storage Server, the IBM System Storage SAN Volume Controller and the IBM System Storage DS family. When the supported devices are configured as MPIO-capable devices, SDDPCM is loaded and becomes part of the AIX MPIO FCP (Fiber Channel Protocol) device driver. The AIX MPIO device driver with the SDDPCM module enhances the data availability and I/O load balancing.

SDDPCM manages the paths to provide High Availability and Load Balancing of storage I/O, automatic path-failover protection, prevention of a single-point-failure caused by host Bus Adapter (HBA). Fibre channel cables or host-interface adapters on supported storage.

To download SDDPCM driver for AIX and documentation, follow the link:

http://www-01.ibm.com/support/docview.wss?uid=ssg1S4000201#DS6K

Oracle Database 11*g* Release 2

Oracle Database 11*g* Release 2 (11.2.0.1) is the current release of Oracle's database product. For both RAC and non-RAC installations AIX should be running in 64-bit kernel mode only. For the latest information on Oracle product certifications, please visit My Oracle Support web site:

https://support.oracle.com/CSP/ui/flash.html

The Oracle Database software can be downloaded from the Oracle Technology Network (OTN) or the DVDs can be requested from Oracle Support. Oracle RAC is a separately licensed option of Oracle Enterprise and Standard Editions. For additional information on pricing, please refer to:

http://www.oracle.com/corporate/pricing/technology-price-list.pdf

Automatic Storage Management

Automatic Storage Management (ASM) provides volume and cluster file system management where the I/O subsystem is directly handled by the Oracle kernel. Oracle ASM will have each LUN mapped as a disk. Disks are then grouped together into disk groups. Each disk group can be segmented in one or more fail groups. ASM automatically performs load balancing in parallel across all available disk drives to prevent hot spots and maximize performance.

Starting with Oracle Database 11*g* Release 2, Oracle Clusterware OCR and voting disk files can be stored in Oracle ASM disk group.

In Oracle Database 11*g* Release 2 , ASM becomes a complete storage management solution for both Oracle Database and non-database files and has many extended functions for not only storing database files, but also storing binary files, report files, trace files, alert logs and other application data files.

ASM Cluster File Systems (ACFS) extends ASM by providing cluster file system scaled to a large number of nodes and uses extend-based storage allocation for improved performance. ACFS can be can be exported to remote clients through NFS and CIFS.

ASM Dynamic Volume Manager (DVM), ASM FS Snapshot, ASM Intelligent Data Placement, ASM Storage Management Configuration Assistant (ASMCA), ASM File Access Control and ASMCMD are some of the extended functions of ASM.

For more information on ASM new features , refer to the Oracle document "*Oracle Database New Features Guide 11g Release 2 (11.2)"*

Configuring the system environment

Virtual IO Server (VIOS) and Logical Partitions (LPARs)

The VIOS is part of the IBM Power Systems server machine's Advanced Power Virtualization hardware feature. VIOS allows sharing of physical resources between LPARs including virtual SCSI, virtual networking and virtual fibre channel adapters. This allows more efficient utilization of physical resources through sharing between LPARs and facilitates server consolidation. This allows a single machine to run multiple operating system (AIX or Linux on POWER) images at the same time while each is isolated from the others.

VIOS itself a logical partition (LPAR) on a IBM Power System machine, which has the OS and command line to manage hardware resources. VIOS is controlled by the Hardware Management Console (HMC) that owns hardware adapters like SCSI disks, Fibre-Channel disks, Ethernet or CD/DVD optical devices but allows other LPARs to access them or a part of them. This allows the device to be shared. The LPAR with the resources is called the VIO Server and the other LPARs using it are called VIO Clients. For example, instead of each LPAR (VIO client) having a SCSI adapter and SCSI disk to boot from they can share one disk on the VIO Server. This reduces costs by eliminating adapters, adapter slots and disks. This client - server access is implemented over memory within the machine for speed.

In the lab test for installing Oracle Database 11*g* Release 2 with RAC, two LPARs were used as RAC nodes. Each LPAR is created on one of the two Power 570 servers. One virtual Ethernet adapter was

used for public connectivity and one 10GigE Logical Host Ethernet Adapter (LHEA) was used for interconnection between LPARs. IBM POWER6 processor-based servers like p570 used in our tests offer Integrated Virtual Ethernet adapter (IVE), which gives integrated high-speed Ethernet adapter ports (Host Ethernet Adapter (HEA)) with hardware assisted virtualization capabilities.

Host Ethernet Adapter can be used for external Ethernet connectivity for LPARs using dedicated ports without the need of a Virtual IO Server. An HEA adapter directly connects to the internal GX+ bus of a POWER6 processor-based server instead of connecting to PCIe or PCI-X bus.

IVE also includes special hardware features that provide the logical Ethernet adapter, also called the Logical Host Ethernet Adapter (LHEA). These LHEA adapters can be directly assigned to the LPARs without being configured through the POWER Hypervisor (PHYP). This eliminates the need to move the sockets between LPARs through the Shared Ethernet Adapter (SEA). IVE replaces the need for virtual Ethernet and SEA in Virtual IO server (VIOS) environment and LPARs can share the HEA ports with improved performance.

Each LPAR has a physical Host Base Adapter (HBA) and is connected to a switched SAN for storage from the IBM System storage DS6800. The disks for installing AIX and Oracle Database on each LPAR were supplied from VIOS local SCSI disks.

The following diagram shows the setup of LPARs for the Oracle Database 11*g* Release 2 with RAC environment.

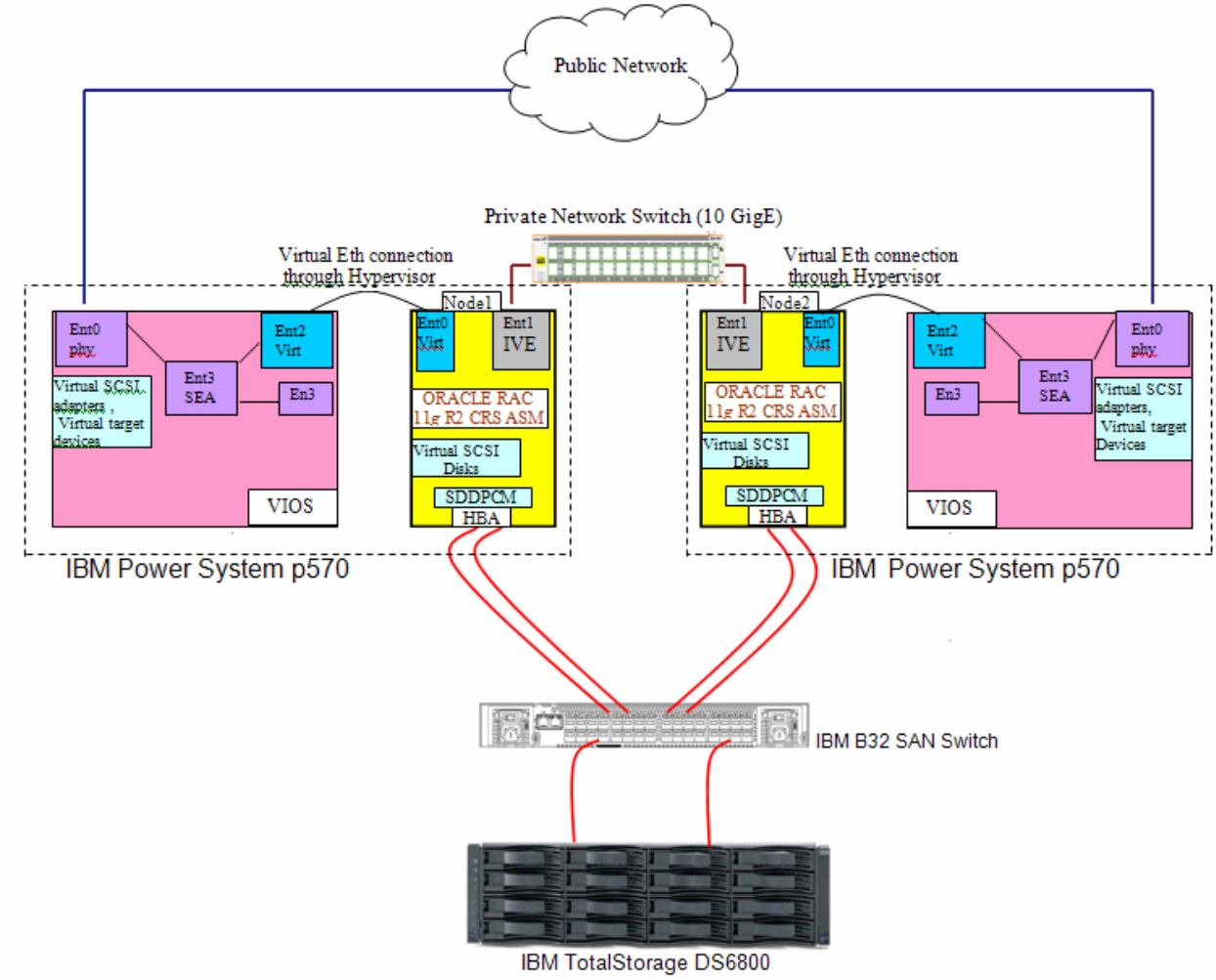


Figure5: 2 Node Oracle RAC setup in the test lab.

Hardware Management Console (HMC)

The HMC is based on the IBM System x hardware architecture running dedicated applications to provide partition management for one or more IBM Power Systems servers called managed systems. The HMC is used for creating and maintaining a multiple partition (LPAR) environment. The HMC acts as a virtual operating system session terminal for each partition. It is used for detecting, reporting and storing changes in hardware conditions, managing system power ON/OFF for the server, acts as a service focal point, is used to upgrade the Power Systems server micro code, and for activating Capacity on Demand.

The major functions that the HMC provides are server hardware management and virtualization management. Using the HMC, dynamic LPAR operations can be done to change resource allocation such as processor, memory, physical I/O and virtual I/O for a specific partition.

The HMC can be accessed through a web-based client using web browsers and command line interfaces. The web interface uses a tree-style navigation model that provides hierarchical views of system resources and tasks using drill-down and launch-in-context techniques to enable direct access to

hardware resources and task management capabilities. The HMC manages advanced PowerVM virtualization features of POWER5, POWER6, and POWER7 servers.

Two LPARs were created on the p570 servers for the two node Oracle Database with RAC testing described in this paper.

Installing the AIX operating system

Installation of the operating systems will not be discussed in detail in this paper.

AIX 6.1 TL04 (6100-04-03-1009) is installed on the RAC nodes through the Network Installation Manager server (NIM).

Prior to Oracle software installation, please make note of the following:

- Be sure to create sufficient swap space appropriate for the amount of physical memory on your servers (use “lsps -a” command). Oracle recommends that the swap space be 1.5 times the size of RAM if the RAM size is between 1.5 GB and 2 GB. The amount of swap space should equal the amount of RAM if RAM size is between 2 GB and 8 GB. For more than 8 GB of RAM, the swap space should be 0.75 times the size of RAM.

Note: On AIX systems with 1.5 GB or more of memory, Oracle recommends that you set the paging space to an initial setting of half the size of RAM plus 4 GB, with an upper limit of 32 GB. During installation, to optimize paging, monitor the paging space use in a separate window. Use the command *chps* to increase or decrease the paging space size. The output of *chps* should indicate paging space use of less than 25 percent on a properly configured system. Refer to Oracle Database Administrator’s Reference for AIX for more information about configuring paging space.

For listing the real memory and the available swap space, use the following commands:

- # /usr/sbin/lsattr -E -l sys0 -a realmem
- # lsps –s

To find out the disk size use:

- #bootinfo -s hdisk<#>

The above command displays the size in MB.

It is strongly recommended that every node of the cluster have an identical hardware configuration, although it is not mandatory.

Oracle publishes the following as a minimal set of hardware requirements for each server.

Hardware	Minimum	Recommended
Physical memory	1.5GB	Depends on applications and usage
CPU	1 CPU per node	2 or more CPUs per node (a processor type that is certified with Oracle 11*g* Release 2)
Interconnect network	1Gb	2 teamed Gb
External network	100Mb	1Gb
Backup network	100Mb	1Gb
HBA or NIC for SAN, iSCSI, or NAS	1Gb HBA	Dual-pathed storage vendor certified HBA
Oracle Database single instance	4 GB	4 GB or more
Oracle Grid home (includes the binary files for Oracle Clusterware and Oracle ASM and their associated log files)	4.5GB	5 GB (with sample schemas)
Temporary disk space	1 GB	1 GB or more (and less than 2TB)

Table 1: Hardware requirements

Prior to installing the Oracle products, you should install the required OS packages, otherwise the Oracle Universal Installer will provide you with the list of packages that you need to install before you can proceed.

- Install openssh and openssl RPM from AIX operating system CD pack “Linux Toolkit for AIX” or download it from http://sourceforge.net/projects/openssh-aix/.
- bos.adt.base
- bos.adt.lib
- bos.adt.libm
- bos.perf.libperfstat
- bos.perf.perfstat
- bos.perf.proctools
- rsct.basic.rte
- rsct.compat.clients.rte
- xlC.aix61.rte 10.1.0.0 (or later)

You must have the IBM XL C/C++ runtime filesets for installation, but you do not need the C/C++ compilers. You do not require a license for the XL C/C++ runtime filesets. IBM XL C/C++ Enterprise Edition for AIX, V9.0 September 2008 PTF is the recommended version.

All AIX 6L 6.1 Authorized Problem Analysis Reports (APARs) for AIX 5L 5.3 ML06, and AIX fixes IZ41855, IZ51456 and IZ52319 should be installed:

Make sure these APAR are applied. If you are using very latest AIX TL and SP like 6100-04-03-1009, the above issues are fixed. But the Cluster Verification tool will fail and show these specific APARs as not applied, ignore such failures.

The PTF files can be downloaded from the IBM fix central http://www-933.ibm.com/support/fixcentral/.

Installing Oracle Grid Infrastructure 11.2.0.1

Before installing Oracle Grid Infrastructure 11.2.0.1 on both servers, there are several important tasks that need to be done on all of the cluster nodes.

Pre-Installation tasks

Configuring OS kernel parameters

Make sure the "aio_maxreqs" is set to 65536 (64K) by issuing "*ioo -a |grep aio_maxreqs*". If it is not 64K, set it by "*ioo –a aio_maxreqs=65536*".

Keep the default values of the virtual memory parameter values in AIX 6.1 and make sure the following values are set:

- minperm%=3
- maxperm%=90
- maxclient%=90
- lru_file_repage=0
- strict_maxclient=1
- strict_maxperm=0

If these values are not set, use the command "*vmo -p –o <parameter=new value>*"

Edit the following lines to the /etc/security/limits file, -1 represents "unlimited", the default values are:

- fsize = -1
- core = 2097151
- cpu = -1
- data = -1
- rss = -1
- stack = -1
- nofiles = -1

Set the "maxuproc" parameter to 16384 by using the command "*/usr/sbin/chdev -l sys0 -a maxuproc = 16384*".

Verify that the maximum number of processes allowed for each user is set to 2048 using "*# smit chgsys*".

Configuring network parameters

The recommended values for the network parameters in AIX when running the Oracle Database are:

- ipqmaxlen=512
- rfc1323=1
- sb_max=1310720
- tcp_recvspace=65536
- tcp_sendspace=65536
- udp_recvspace=65536
- udp_sendspace 65536

Find out the current values for the above parameters using the "*no –a*" command. To set the values, determine whether the system is running in compatibility mode or not by using the command "*# lsattr -E -l sys0 -a pre520tune*".

If the output is "pre520tune disable Pre-520 tuning compatibility mode True" ,the system is not running in compatibility mode. If the system is not running in compatibility mode you can set the values using the following commands:

- For setting ipqmaxlen use "*/usr/sbin/no -r -o ipqmaxlen=512*"
- For setting other parameters use "*/usr/sbin/no -p -o parameter=value"*

If the system is running in compatibility mode, then the output is similar to the following, showing that the value of the pre520tune attribute is enabled: "*pre520tune **enable** Pre-520 tuning compatibility mode True*".

For compatibility mode, set the values by using the command "*no -o parameter_name=value"* and make the following entries in */etc/rc.net* file.

```
if [ -f /usr/sbin/no ] ; then
             /usr/sbin/no -o udp_sendspace=65536
             /usr/sbin/no -o udp_recvspace=655360
             /usr/sbin/no -o tcp_sendspace=65536
             /usr/sbin/no -o tcp_recvspace=65536
             /usr/sbin/no -o rfc1323=1
             /usr/sbin/no -o sb_max=2*655360
             /usr/sbin/no -o ipqmaxlen=512
fi
```

Network time protocol

Oracle Clusterware 11*g* Release 2 (11.2) requires time synchronization across all Oracle RAC nodes within a cluster when Oracle RAC is deployed. There are two ways the time synchronization can be

configured. An operating system configured network time protocol (NTP) or Oracle Cluster Time Synchronization Service (CTSS). Oracle Cluster Time Synchronization Service is designed for organizations whose cluster servers are unable to access NTP services. In the lab test setup, NTP is used for the time synchronization. If the NTP is used, then the Oracle Cluster Time Synchronization daemon starts up in observer mode.

Add the following entry in the Time Server, the system acting as the Time Server should preferably be the server (for example, in the lab setup the time server was 9.38.158.208) that is running all of the time, but not one of the Oracle RAC nodes.

```
/etc/ntp.conf:
          server 127.127.1.0
          #broadcastclient
          driftfile /etc/ntp.drift
         tracefile /etc/ntp.trace
```

Restart the "*ntpd*" after updating *ntp.conf* file on the NTP server and all client nodes.

In each of the Oracle RAC nodes update as below.

```
/etc/ntp.conf:
          server 9.38.158.208
          #broadcastclient
          driftfile /etc/ntp.drift
         tracefile /etc/ntp.trace
```

In each of the Oracle RAC nodes, edit the entry for *ntpd* in the file */etc/rc.tcpip*:

```
/etc/rc.tcpip:
        start /usr/sbin/xntpd "$src_running" "-x"
```

Restart the *xntpd* on NTP server and RAC nodes by issuing:

- # stopsrc -s xntpd
- # startsrc -s xntpd -a "-x" where, -x is used for making small time adjustments (SLEWING).

After the times in the Oracle RAC nodes are synchronized with the time server, check the status of time synchronization on Oracle RAC nodes by running the command:

```
# ntpq –p
    remote    refid   st t when poll reach   delay   offset    disp
```

```
==============================================================

*dyn9038158208.s LOCAL(0)4 u   52 1024  377   0.23   -0.052    0.26
```

Time synchronization on the nodes will take a few minutes after restarting the *xntp* daemon on the nodes.

Creating users and groups

Oracle recommends creating the following operating system groups and users for all installations where separate software installation needs its own user. The operating system groups are *oinstall*, *dba*, *asmdba*, *asmadmin* and *asmoper*. The users are *grid* and *oracle*.

On each of the Oracle RAC nodes the group ID and user ID number should be the same.

- # mkgroup -'A' id='1000' adms='root' oinstall
- # mkgroup -'A' id='2000' adms='root' dba
- # mkgroup -'A' id='3000' adms='root' asmadmin
- # mkgroup -'A' id='4000' adms='root' oper
- # mkuser id='1100' pgrp='oinstall' groups='dba,oper,asmadmin' home='/home/grid' grid
- # mkuser id='1101' pgrp='oinstall' groups='dba,oper,asmadmin' home='/home/oracle' oracle

For the lab test, the user *grid* is used for installing the Oracle Grid Infrastructure software and the user *oracle* is used for installing the Oracle RAC software.

Set values for the attribute “capabilities” of the user *grid* as follows

/usr/bin/chuser capabilities=CAP_NUMA_ATTACH,CAP_BYPASS_RAC_VMM, CAP_PROPAGATE grid

Create a separate filesystem for the Oracle Grid Infrastructure software and the Oracle Database software. Create two directories (ORACLE_HOME), one for the Oracle Grid Infrastructure and another one for the Oracle Database software.

- #mkdir –p /u01/app/112/grid
- #chown –R grid:oinstall /u01/app/112/grid
- #mkdir –p /d01/app/112/dbhome
- #chown –R oracle:oinstall /u01/app/112/oracle

With 11*g* Release 2 of both products there are two separate ORACLE_HOME directories: one home for Oracle Grid Infrastructure; and the other home for Oracle Database. To execute commands like ASMCA for Oracle ASM configuration or DBCA for database configuration, you will need to change the ORACLE_HOME environment variable to the Oracle Database home directory.

Setting Oracle inventory location

When you install Oracle software on the system for the first time, a file called *oraInst.loc* will be created under the */etc* directory. The file stores information about where the Oracle inventory directory is and the name of the Oracle Inventory group.

```
inventory_loc=/u01/app/oraInventory
inst_group=oinstall
```

If a previous inventory directory exists, please make sure that the same Oracle inventory directory is used and all Oracle software users have write permissions to this directory.

Setting up network files

The following network addresses are required for each node:

- Public network address
- Private network address
- Virtual IP network address (VIP)
- Single Client Access Name (SCAN) address for the cluster

The interfaces and IP addresses for both public and private networks need to be set up. After setting up the IP addresses for public and private IP addresses, */etc/hosts* will look like the following.

```
9.38.158.206    dyn9038158206   dyn9038158206.sanmateo.ibm.com

9.38.158.207    dyn9038158207   dyn9038158207.sanmateo.ibm.com

10.0.0.10       racnode1_pvt

10.0.0.20       racnode2_pvt

9.38.158.204    racnode1_vip    dyn9038158204.sanmateo.ibm.com

9.38.158.228    racnode2_vip    dyn9038158228.sanmateo.ibm.com
```

In Oracle Database 11*g* Release 2, SCAN is introduced. It needs a static IP address which should be resolved in the Domain Name Server (DNS). The SCAN IP address needs to not be placed in */etc/hosts*. Oracle recommends three IP addresses for SCAN. It is a single DNS entry with three IP addresses attached to a name and set to round robin. For the lab test, only one IP addresses is used:

- SCAN IP address: 9.38.158.229

All of the public IP, VIP and the SCAN IP addresses should be in the same subnet.

Configuring SSH on all cluster nodes

Oracle Database 11*g* Release 2 needs SSH and it should be setup on the Oracle RAC nodes to login each other without the password. This is done either manually or with an Oracle provided script "*sshUserSetup.sh*" or by the Oracle Universal Installer (OUI).

In the lab test, "*sshUserSetup.sh*" was used to configure the SSH on the nodes. This should be done for each of the users, for instance *grid* and *oracle*.

Configuring shared disks for OCR, voting and database

Starting with version 11*g* Release 2, the Oracle Clusterware voting disk and OCR can be stored in ASM. Oracle strongly recommends storing Oracle Clusterware disks in ASM. However, Oracle Clusterware binaries and files cannot be stored in an Oracle ASM Cluster File System (ACFS). Oracle recommends 280 MB minimum for each voting disk and OCR file. The total required values are cumulative and it depends on the level of redundancy you choose during the installation.

In our example, Oracle Clusterware disks (OCR and voting disks) will be stored in the Oracle ASM diskgroup. Oracle ASM disks will need to be created prior to installation with correct ownership and permission. All of the disks should be shared across Oracle RAC nodes.

Oracle RAC OCR, voting disks and database are created in ASM diskgroup.

```
ASM diskgroup for OCR, voting disks

#chown grid:oinstall /dev/rhdisk<#>
#chmod 660 /dev/rhdisk<#>

For Diskgroup for the database

#chown oracle:oinstall /dev/rhdisk<#>
#chmod 660 /dev/rhdisk<#>
```

Running Cluster Verification Utility (CVU)

The Cluster Verification Utility (CVU) can be used to verify that the systems are ready to install Oracle Clusterware 11*g* Release 2. The Oracle Universal Installer will use CVU to perform all pre-requisite checks during the installation interview. Login as user *grid* and run the following command:

```
$./runcluvfy.sh stage -pre crsinst -n dyn9038158207,dyn9038158206 -verbose

Performing pre-checks for cluster services setup

Checking node reachability...

Check: Node reachability from node "dyn9038158207"
  Destination Node                      Reachable?
  ------------------------------------  ------------------------
  dyn9038158207                         yes
  dyn9038158206                         yes
Result: Node reachability check passed from node "dyn9038158207"

Checking user equivalence...

Check: User equivalence for user "grid"
  Node Name                             Comment
  ------------------------------------  ------------------------
  dyn9038158207                         passed
  dyn9038158206                         passed
Result: User equivalence check passed for user "grid"

Checking node connectivity...

Checking hosts config file...
  Node Name     Status                    Comment
  ------------  ------------------------  ------------------------
  dyn9038158207  passed
  dyn9038158206  passed

Verification of the hosts config file successful

Interface information for node "dyn9038158207"
 Name   IP Address      Subnet          Gateway         Def. Gateway    HW
Address        MTU
 ------ --------------- --------------- --------------- --------------- ---------
-------- ------
```

```
 en0    10.0.0.20       10.0.0.0        10.0.0.20       9.38.158.129
00:1A:64:28:3F:B2 1500
 en1    9.38.158.207    9.38.158.128    9.38.158.207    9.38.158.129
72:53:C6:57:79:02 1500


Interface information for node "dyn9038158206"
 Name   IP Address      Subnet          Gateway         Def. Gateway    HW
Address        MTU
 ------ --------------- --------------- --------------- --------------- ---------
-------- ------
 en0    10.0.0.10       10.0.0.0        10.0.0.10       9.38.158.129
00:1A:64:28:4D:72 1500
 en1    9.38.158.206    9.38.158.128    9.38.158.206    9.38.158.129
5A:53:FB:6D:D6:02 1500


Check: Node connectivity of subnet "10.0.0.0"
  Source                          Destination                     Connected?
  ------------------------------  ------------------------------  ---------------
-
  dyn9038158207:en0               dyn9038158206:en0               yes
Result: Node connectivity passed for subnet "10.0.0.0" with node(s)
dyn9038158207,dyn9038158206


Check: TCP connectivity of subnet "10.0.0.0"
  Source                          Destination                     Connected?
  ------------------------------  ------------------------------  ---------------
-
  dyn9038158207:10.0.0.20         dyn9038158206:10.0.0.10         passed
Result: TCP connectivity check passed for subnet "10.0.0.0"


Check: Node connectivity of subnet "9.38.158.128"
  Source                          Destination                     Connected?
  ------------------------------  ------------------------------  ---------------
-
  dyn9038158207:en1               dyn9038158206:en1               yesResult: Node
connectivity passed for subnet "9.38.158.128" with node(s)
dyn9038158207,dyn9038158206


Check: TCP connectivity of subnet "9.38.158.128"
  Source                          Destination                     Connected?
  ------------------------------  ------------------------------  ---------------
-
  dyn9038158207:9.38.158.207      dyn9038158206:9.38.158.206      passed
Result: TCP connectivity check passed for subnet "9.38.158.128"


Interfaces found on subnet "9.38.158.128" that are likely candidates for VIP are:
dyn9038158207 en1:9.38.158.207
dyn9038158206 en1:9.38.158.206

Interfaces found on subnet "10.0.0.0" that are likely candidates for a private
interconnect are:
dyn9038158207 en0:10.0.0.20
dyn9038158206 en0:10.0.0.10

Result: Node connectivity check passed
```

```
Check: Total memory
  Node Name     Available                 Required                  Comment
  ------------  ------------------------  ------------------------  ----------
  dyn9038158207  12GB (1.2582912E7KB)      1.5GB (1572864.0KB)       passed
  dyn9038158206  12GB (1.2582912E7KB)      1.5GB (1572864.0KB)       passed
Result: Total memory check passed

Check: Available memory
  Node Name     Available                 Required                  Comment
  ------------  ------------------------  ------------------------  ----------
  dyn9038158207  6.9351GB (7271952.0KB)    50MB (51200.0KB)          passed
  dyn9038158206  7.5801GB (7948348.0KB)    50MB (51200.0KB)          passed
Result: Available memory check passed

Check: Swap space
  Node Name     Available                 Required                  Comment
  ------------  ------------------------  ------------------------  ----------
  dyn9038158207  12.375GB (1.2976128E7KB)  9GB (9437184.0KB)         passed
  dyn9038158206  12.375GB (1.2976128E7KB)  9GB (9437184.0KB)         passed
Result: Swap space check passed

Check: Free disk space for "dyn9038158207:/tmp/"
  Path              Node Name     Mount point   Available     Required
Comment
  ----------------  ------------  ------------  ------------  ------------  -----
-------
  /tmp/             dyn9038158207  /tmp          1.1501GB      1GB
passed
Result: Free disk space check passed for "dyn9038158207:/tmp/"

Check: Free disk space for "dyn9038158206:/tmp/"
  Path              Node Name     Mount point   Available     Required
Comment
  ----------------  ------------  ------------  ------------  ------------  -----
-------
  /tmp/             dyn9038158206  /tmp          1.013GB       1GB
passed
Result: Free disk space check passed for "dyn9038158206:/tmp/"

Check: User existence for "grid"
  Node Name     Status                    Comment
  ------------  ------------------------  ------------------------
  dyn9038158207  exists(1100)              passed
  dyn9038158206  exists(1100)              passed

Checking for multiple users with UID value 1100
Result: Check for multiple users with UID value 1100 passed
Result: User existence check passed for "grid"

Check: Group existence for "oinstall"Node Name     Status
Comment
  ------------  ------------------------  ------------------------
  dyn9038158207  exists                    passed
  dyn9038158206  exists                    passed
Result: Group existence check passed for "oinstall"

Check: Group existence for "dba"
  Node Name     Status                    Comment
  ------------  ------------------------  ------------------------
  dyn9038158207  exists                    passed
```

```
  dyn9038158206  exists                    passed
Result: Group existence check passed for "dba"

Check: Membership of user "grid" in group "oinstall" [as Primary]
  Node Name         User Exists   Group Exists  User in Group  Primary
Comment
  ----------------  ------------  ------------  ------------  ------------  -----
-------
  dyn9038158207     yes           yes           yes           yes
passed
  dyn9038158206     yes           yes           yes           yes
passed
Result: Membership check for user "grid" in group "oinstall" [as Primary] passed

Check: Membership of user "grid" in group "dba"
  Node Name         User Exists   Group Exists  User in Group  Comment
  ----------------  ------------  ------------  ------------  ----------------
  dyn9038158207     yes           yes           yes           passed
  dyn9038158206     yes           yes           yes           passed
Result: Membership check for user "grid" in group "dba" passed

Check: Run level
  Node Name     run level                 Required                  Comment
  ------------  ------------------------  ------------------------  ----------
  dyn9038158207  2                         2                         passed
  dyn9038158206  2                         2                         passed
Result: Run level check passed

Check: Hard limits for "maximum open file descriptors"
  Node Name         Type          Available     Required      Comment
  ----------------  ------------  ------------  ------------  ----------------
  dyn9038158207     hard          9223372036854775807  65536         passed
  dyn9038158206     hard          9223372036854775807  65536         passed
Result: Hard limits check passed for "maximum open file descriptors"

Check: Soft limits for "maximum open file descriptors"
  Node Name         Type          Available     Required      Comment
  ----------------  ------------  ------------  ------------  ----------------
  dyn9038158207     soft          9223372036854775807  1024          passed
  dyn9038158206     soft          9223372036854775807  1024          passed
Result: Soft limits check passed for "maximum open file descriptors"

Check: Hard limits for "maximum user processes"
  Node Name         Type          Available     Required      Comment
  ----------------  ------------  ------------  ------------  ----------------
  dyn9038158207     hard          16384         16384         passed
  dyn9038158206     hard          16384         16384         passed
Result: Hard limits check passed for "maximum user processes"

Check: Soft limits for "maximum user processes"
  Node Name         Type          Available     Required      Comment
  ----------------  ------------  ------------  ------------  ----------------
  dyn9038158207     soft          16384         2047          passed
  dyn9038158206     soft          16384         2047          passed
Result: Soft limits check passed for "maximum user processes"

Check: System architecture
  Node Name     Available                 Required                  Comment
  ------------  ------------------------  ------------------------  ----------
dyn9038158207  powerpc                   powerpc                   passed
dyn9038158206  powerpc                   powerpc                   passed
Result: System architecture check passed
```

```
Check: Kernel version
  Node Name     Available                 Required                  Comment
  ------------  ------------------------  ------------------------  ----------
  dyn9038158207  6.1-6100.04.03.1009       6.1-6100.02.01            passed
  dyn9038158206  6.1-6100.04.03.1009       6.1-6100.02.01            passed
Result: Kernel version check passed

Check: Kernel parameter for "SEM_NSEMS_MAX"
  Node Name     Configured                Required                  Comment
  ------------  ------------------------  ------------------------  ----------
  dyn9038158207  32768                     256                       passed
  dyn9038158206  32768                     256                       passed
Result: Kernel parameter check passed for "SEM_NSEMS_MAX"

Check: Kernel parameter for "SEM_VALUE_MAX"
  Node Name     Configured                Required                  Comment
  ------------  ------------------------  ------------------------  ----------
  dyn9038158207  32767                     100                       passed
  dyn9038158206  32767                     100                       passed
Result: Kernel parameter check passed for "SEM_VALUE_MAX"

Check: Kernel parameter for "ncargs"
  Node Name     Configured                Required                  Comment
  ------------  ------------------------  ------------------------  ----------
  dyn9038158207  256                       128                       passed
  dyn9038158206  256                       128                       passed
Result: Kernel parameter check passed for "ncargs"

Check: Package existence for "bos.adt.base-..."
  Node Name     Available                 Required                  Comment
  ------------  ------------------------  ------------------------  ----------
  dyn9038158207  bos.adt.base-6.1.4.1-0    bos.adt.base-...          passed
  dyn9038158206  bos.adt.base-6.1.4.1-0    bos.adt.base-...          passed
Result: Package existence check passed for "bos.adt.base-..."

Check: Package existence for "bos.adt.lib-..."
  Node Name     Available                 Required                  Comment
  ------------  ------------------------  ------------------------  ----------
  dyn9038158207  bos.adt.lib-6.1.2.0-0     bos.adt.lib-...           passed
  dyn9038158206  bos.adt.lib-6.1.2.0-0     bos.adt.lib-...           passed
Result: Package existence check passed for "bos.adt.lib-..."

Check: Package existence for "bos.adt.libm-..."
  Node Name     Available                 Required                  Comment
  ------------  ------------------------  ------------------------  ----------
  dyn9038158207  bos.adt.libm-6.1.4.0-0    bos.adt.libm-...          passed
  dyn9038158206  bos.adt.libm-6.1.4.0-0    bos.adt.libm-...          passed
Result: Package existence check passed for "bos.adt.libm-..."

Check: Package existence for "bos.perf.libperfstat-6.1.2.1"
  Node Name     Available                 Required                  Comment
  ------------  ------------------------  ------------------------  ----------
  dyn9038158207  bos.perf.libperfstat-6.1.4.2-0  bos.perf.libperfstat-6.1.2.1
passed
  dyn9038158206  bos.perf.libperfstat-6.1.4.2-0  bos.perf.libperfstat-6.1.2.1
passed
Result: Package existence check passed for "bos.perf.libperfstat-6.1.2.1"

Check: Package existence for "bos.perf.perfstat-..."
  Node Name     Available                 Required                  Comment
  ------------  ------------------------  ------------------------  ----------
  dyn9038158207  bos.perf.perfstat-6.1.4.1-0  bos.perf.perfstat-...     passed
```

```
  dyn9038158206  bos.perf.perfstat-6.1.4.1-0  bos.perf.perfstat-...
passedResult: Package existence check passed for "bos.perf.perfstat-..."

Check: Package existence for "bos.perf.proctools-..."
  Node Name     Available                 Required                  Comment
  ------------  ------------------------  ------------------------  ----------
  dyn9038158207  bos.perf.proctools-6.1.4.1-0  bos.perf.proctools-...    passed
  dyn9038158206  bos.perf.proctools-6.1.4.1-0  bos.perf.proctools-...    passed
Result: Package existence check passed for "bos.perf.proctools-..."

Check: Package existence for "rsct.basic.rte-..."
  Node Name     Available                 Required                  Comment
  ------------  ------------------------  ------------------------  ----------
  dyn9038158207  rsct.basic.rte-2.5.4.0-0  rsct.basic.rte-...        passed
  dyn9038158206  rsct.basic.rte-2.5.4.0-0  rsct.basic.rte-...        passed
Result: Package existence check passed for "rsct.basic.rte-..."

Check: Package existence for "rsct.compat.clients.rte-..."
  Node Name     Available                 Required                  Comment
  ------------  ------------------------  ------------------------  ----------
  dyn9038158207  rsct.compat.clients.rte-2.5.4.0-0  rsct.compat.clients.rte-...
passed
  dyn9038158206  rsct.compat.clients.rte-2.5.4.0-0  rsct.compat.clients.rte-...
passed
Result: Package existence check passed for "rsct.compat.clients.rte-..."

Check: Package existence for "xlC.aix61.rte-10.1.0.0"
  Node Name     Available                 Required                  Comment
  ------------  ------------------------  ------------------------  ----------
  dyn9038158207  xlC.aix61.rte-10.1.0.2-0  xlC.aix61.rte-10.1.0.0    passed
  dyn9038158206  xlC.aix61.rte-10.1.0.2-0  xlC.aix61.rte-10.1.0.0    passed
Result: Package existence check passed for "xlC.aix61.rte-10.1.0.0"

Check: Operating system patch for "Patch IZ41855"
  Node Name     Applied                   Required                  Comment
  ------------  ------------------------  ------------------------  ----------
  dyn9038158207  missing                   Patch IZ41855             failed
  dyn9038158206  missing                   Patch IZ41855             failed
Result: Operating system patch check failed for "Patch IZ41855"

Check: Operating system patch for "Patch IZ51456"
  Node Name     Applied                   Required                  Comment
  ------------  ------------------------  ------------------------  ----------
  dyn9038158207  missing                   Patch IZ51456             failed
  dyn9038158206  missing                   Patch IZ51456             failed
Result: Operating system patch check failed for "Patch IZ51456"

Check: Operating system patch for "Patch IZ52319"
  Node Name     Applied                   Required                  Comment
  ------------  ------------------------  ------------------------  ----------
  dyn9038158207  missing                   Patch IZ52319             failed
  dyn9038158206  missing                   Patch IZ52319             failed
Result: Operating system patch check failed for "Patch IZ52319"

Checking for multiple users with UID value 0
Result: Check for multiple users with UID value 0 passed

Check: Current group ID
Result: Current group ID check passed
Checking Core file name pattern consistency...
Core file name pattern consistency check passed.
```

```
Checking to make sure user "grid" is not in "root" group
  Node Name     Status                    Comment
  ------------  ------------------------  ------------------------
  dyn9038158207  does not exist            passed
  dyn9038158206  does not exist            passed
Result: User "grid" is not part of "root" group. Check passed
 Check default user file creation mask
  Node Name     Available                 Required                  Comment
  ------------  ------------------------  ------------------------  ----------
  dyn9038158207  022                      0022                      passed
  dyn9038158206  022                      0022                      passed
Result: Default user file creation mask check passed

Starting Clock synchronization checks using Network Time Protocol(NTP)...

NTP Configuration file check started...
The NTP configuration file "/etc/ntp.conf" is available on all nodes
NTP Configuration file check passed

Checking daemon liveness...

Check: Liveness for "xntpd"
  Node Name                             Running?
  ------------------------------------  ------------------------
  dyn9038158207                         yes
  dyn9038158206                         yes
Result: Liveness check passed for "xntpd"

Checking NTP daemon command line for slewing option "-x"
Check: NTP daemon command line
  Node Name                             Slewing Option Set?
  ------------------------------------  ------------------------
  dyn9038158207                         yes
  dyn9038158206                         yes
Result:
NTP daemon slewing option check passed

Checking NTP daemon's boot time configuration, in file "/etc/rc.tcpip", for
slewing option "-x"

Check: NTP daemon's boot time configuration
  Node Name                             Slewing Option Set?
  ------------------------------------  ------------------------
  dyn9038158207                         yes
  dyn9038158206                         yes
Result:
NTP daemon's boot time configuration check for slewing option passed

Result: Clock synchronization check using Network Time Protocol(NTP) passed

Result: User ID < 65535 check passed

Result: Kernel 64-bit mode check passed

Pre-check for cluster services setup was unsuccessful on all the nodes.
```

The cluster verification test shows unsuccessful because the required OS patches are not found as indicated above. It doesn't mean that the current installed OS level does not have the fixs. These OS patches are required to the lower level (6100-03 or less) of current installed AIX level (6100-04-03-1009). The current level has all of the fixes so we can ignore the failure report.

Performing Oracle Clusterware installation and Automatic Storage Management installation

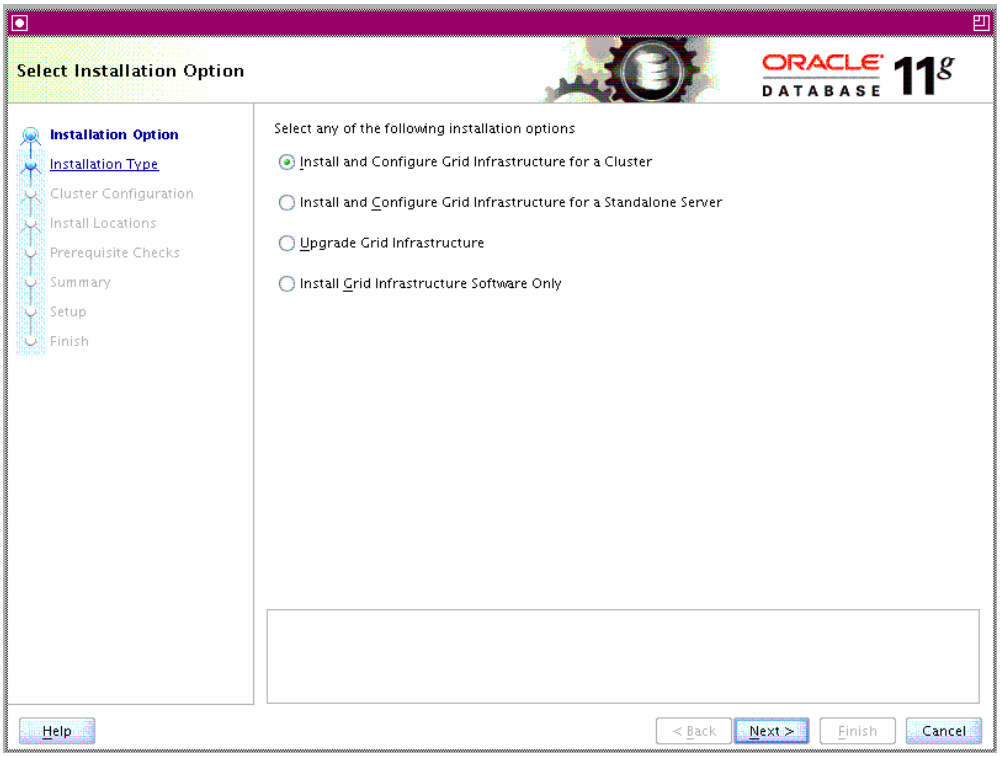
To install Oracle Clusterware 11*g* Release 2, Oracle Database 11*g* Release 2 Grid Infrastructure (11.2.0.1) for AIX, needs to be downloaded. After that, unzip "*aix.ppc64_11gR2_grid.zip*" and run the Oracle Universal Installer (OUI) from one node (local node). For the most part, OUI handles the installation of the other cluster nodes. There are a number of steps that need to be done on the other cluster nodes and these are called out by OUI at various points during the process.

Running the installation from the system console will require an XWindows session, or you can run it from vncserver on the node and use XWindows on the workstation to start the OUI.

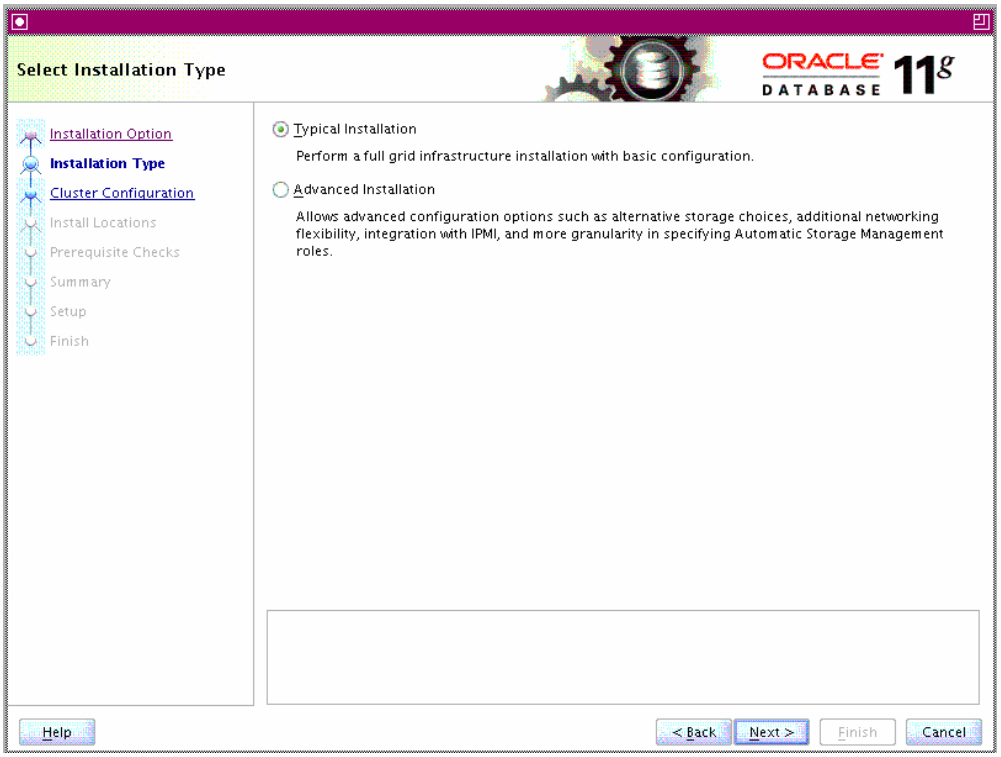
1. Execute "*rootpre.sh*"

```
# ./rootpre.sh
./rootpre.sh output will be logged in /tmp/rootpre.out_10-03-24.21:20:57
Saving the original files in /etc/ora_save_10-03-24.21:20:57....
Copying new kernel extension to /etc....
Loading the kernel extension from /etc

 Oracle Kernel Extension Loader for AIX
       Copyright (c) 1998,1999 Oracle Corporation
 Kernel Extension /etc/pw-syscall.64bit_kernel already loaded, unloading it
 Unconfigured the kernel extension successfully
 Unloaded the kernel extension successfully
 Successfully loaded /etc/pw-syscall.64bit_kernel with kmid: 0x509f9000
 Successfully configured /etc/pw-syscall.64bit_kernel with kmid: 0x509f9000
The kernel extension was successfuly loaded.

Checking if group services should be configured....
Nothing to configure.
```

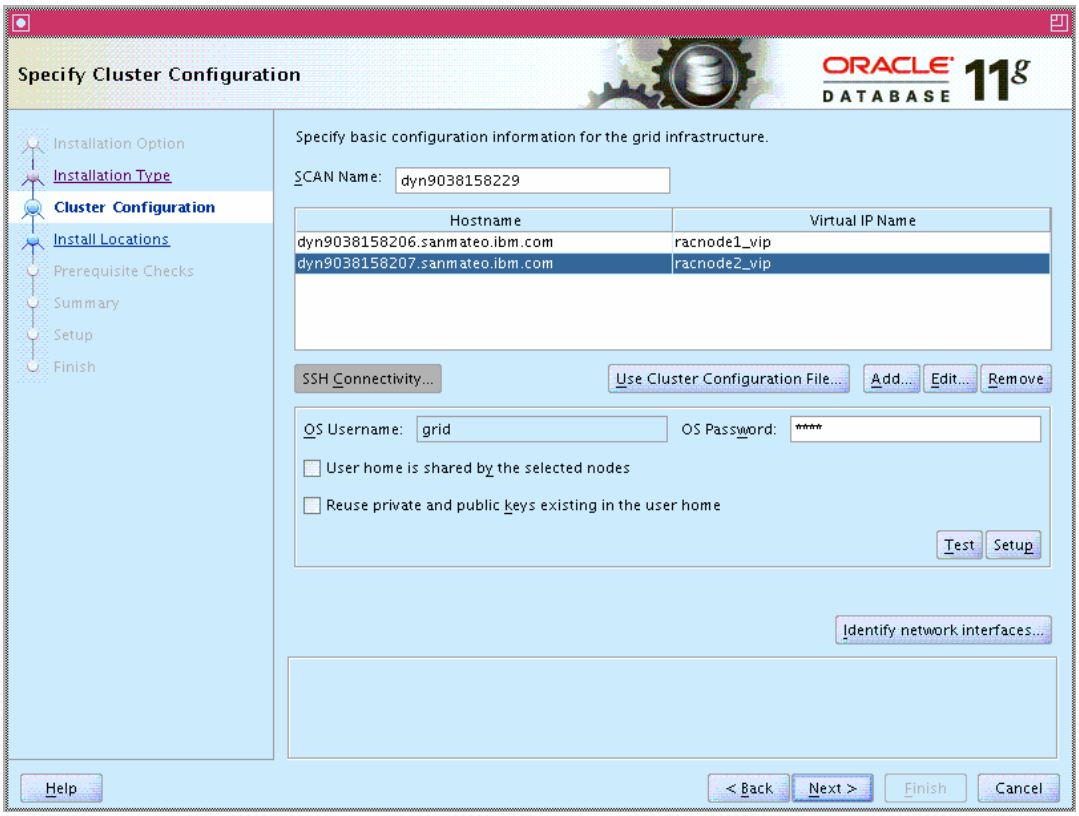
2. Run "*./runInstaller*", the first screen will ask you select one of the installation options. In this example, we select "Install and Configure Grid Infrastructure for a Cluster".



3. The next screen will ask if this is a typical or advanced installation. We select typical installation.

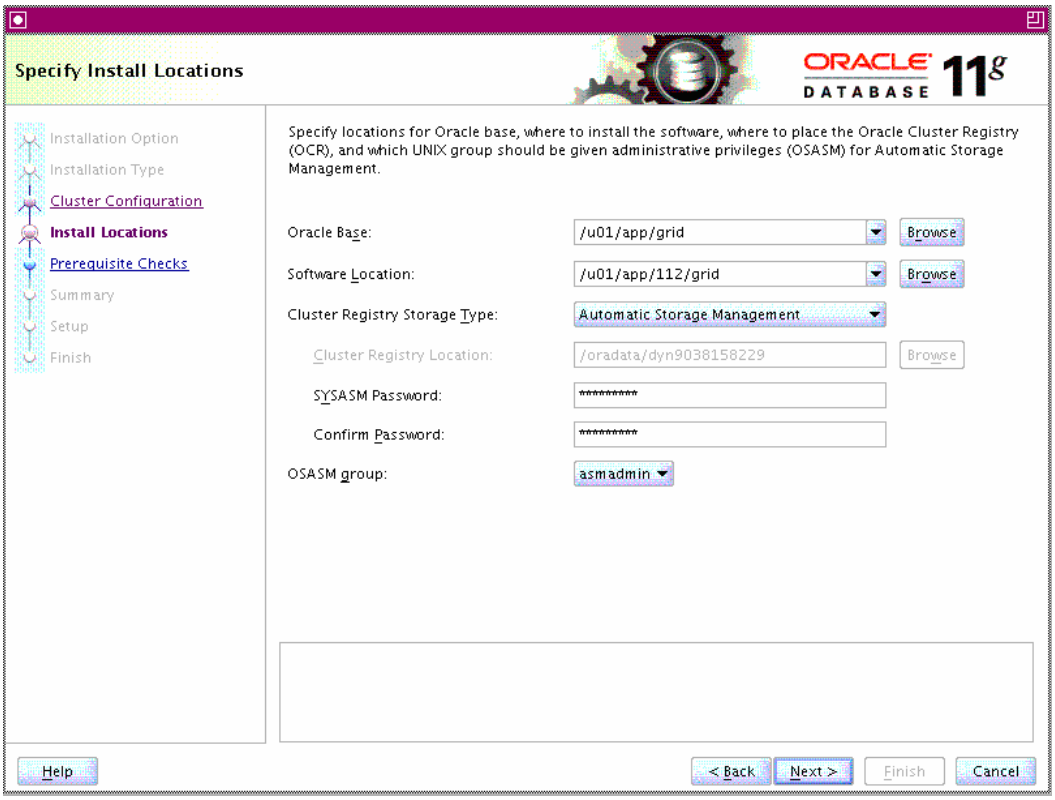


4. The next screen asks for the SCAN and its cluster node names and virtual IP addresses. If this is the first installation, put in the OS password for user *oracle* and click setup. Since SSH is setup using the script on the Oracle RAC nodes, enter the password for the user *grid*. While moving to the next screen OUI automatically tests the SSH setup. Also, you can click “Test” to make sure that the SSH worked properly between the nodes. Since we have chosen “Typical”, it needs a SCAN name that can resolve with the DNS. If you choose “Advanced installation”, a Grid Naming Services (GNS) and its associated information will be required.

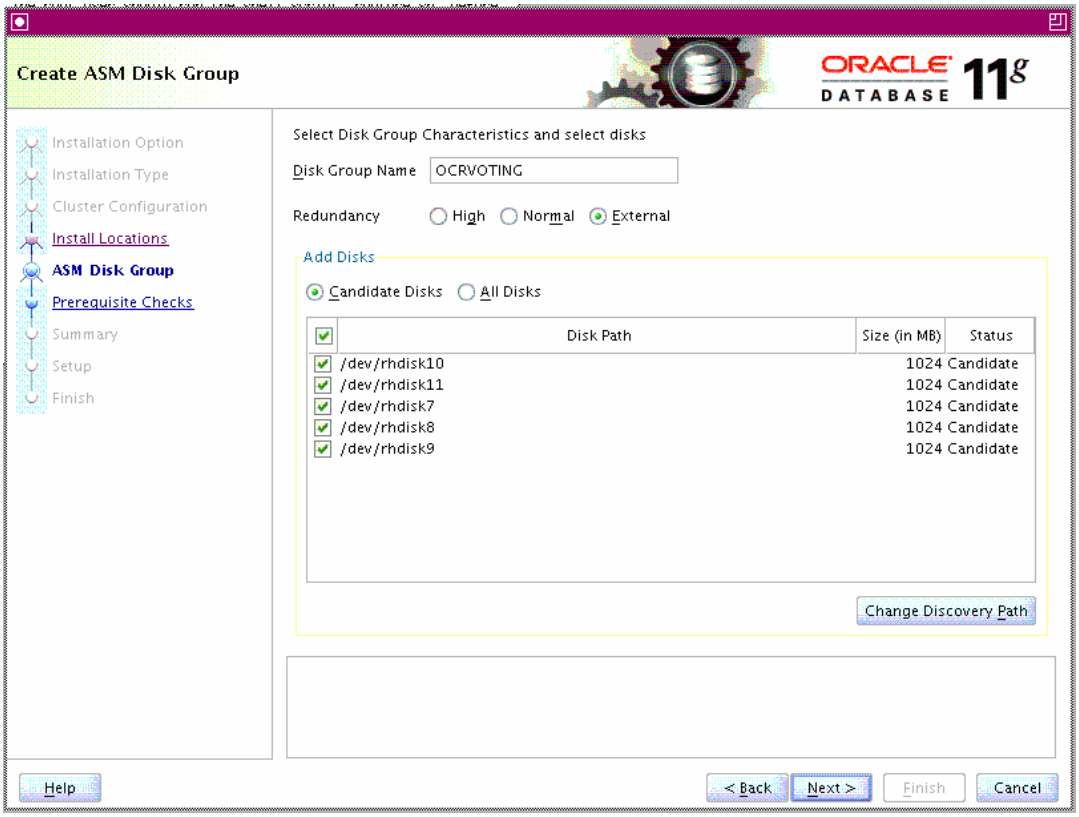


After pressing “Next”, OUI validates the nodes and SSH setup. If you haven’t setup SSH before starting the Oracle Grid Infrastructure installation, you can run “setup” from this screen.

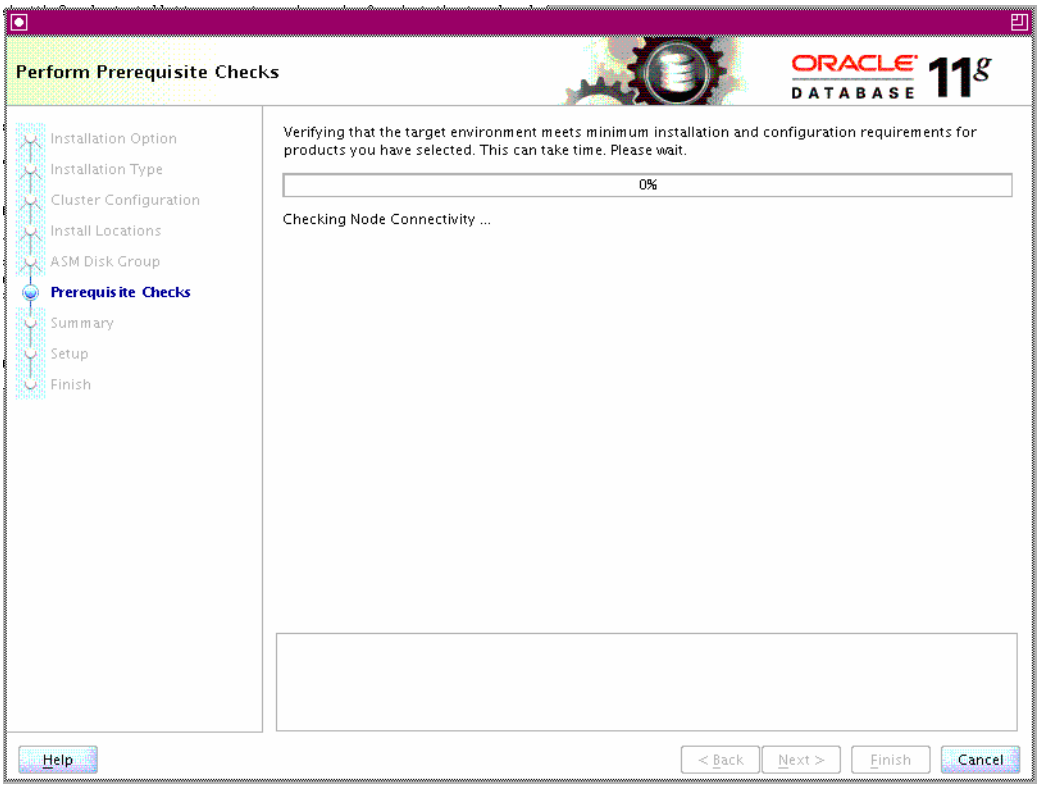
5. The next screen will ask you for the Oracle base and software directories. In this example, all Oracle Clusterware files are going to be stored in ASM. Then, enter the password for SYSASM. Oracle would like the password to conform to specific rules. If you did not follow these rules, errors will be shown at the bottom of the screen.



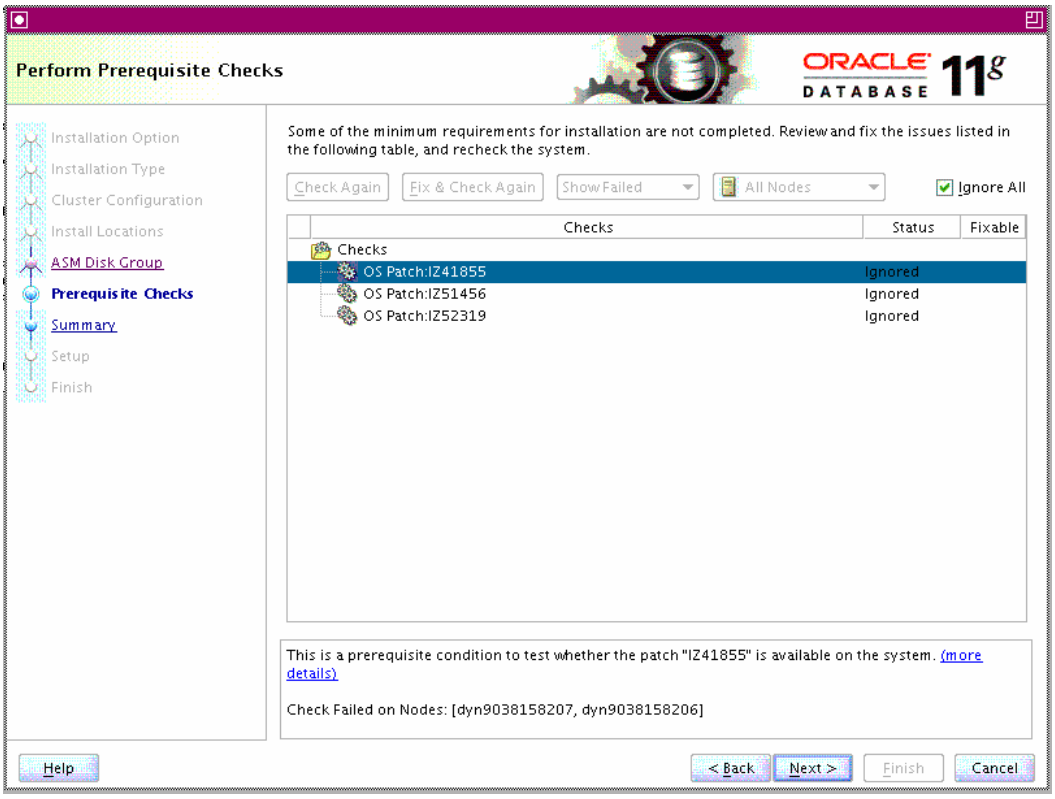
6. Since ASM is chosen to be the storage type for the Clusterware files, the install process then asks for the names of the ASM disks and it will create the Disk Group Name with the selected ASM disks to store the OCR and voting disks. The number of disks needed for installation depends on the redundancy level you picked. For high redundancy five disks are required, for normal redundancy three disks are required, and for external redundancy one disk is required. If you do not select enough disks an error message will be given. The minimum size of each disk is 280 MB. For this example, external redundancy has been chosen.



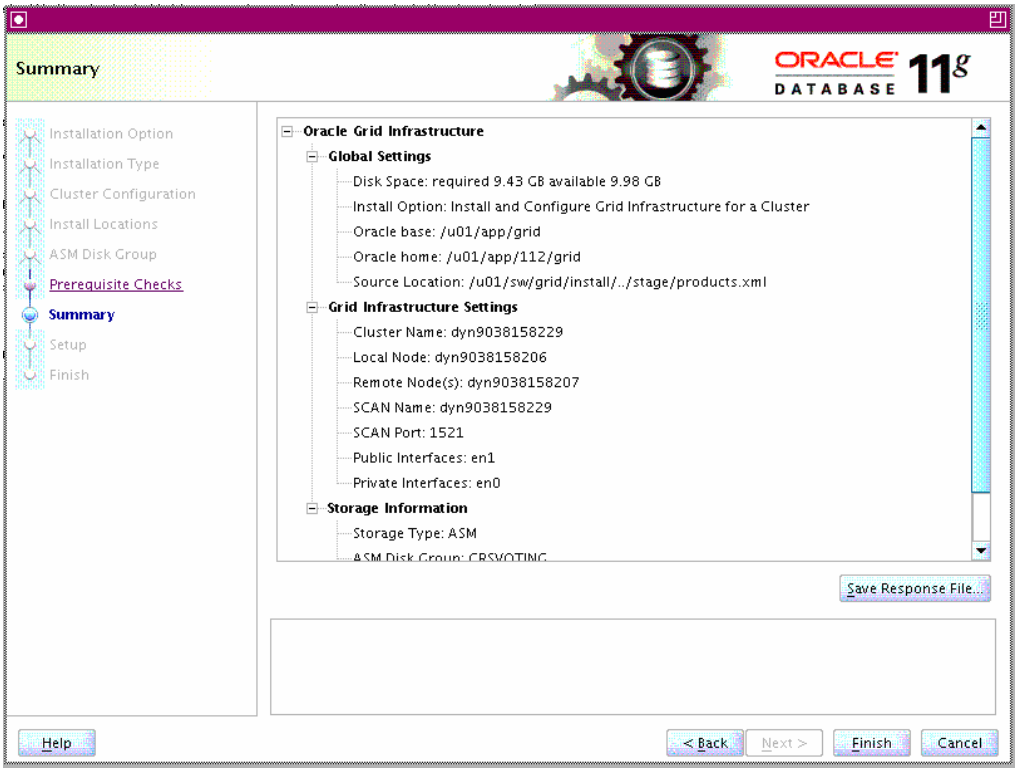
7. Next the Cluster Verification Utility is run to check if the cluster nodes have met all the prerequisites. If not, the installation will stop and show you the errors. You can fix the errors and ask run the check again. At the bottom of the screen, you can click on more details where suggestions on how to fix the errors will be shown.



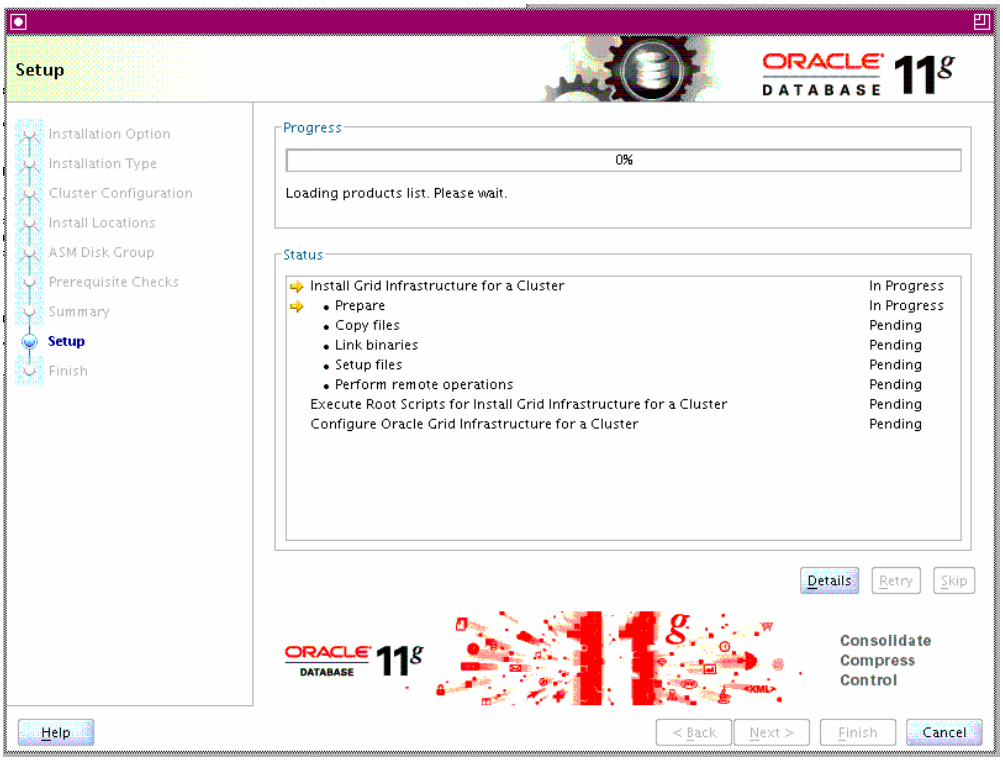
The failure seen earlier by the Cluster Verification Utility for the AIX OS patches will also be shown by the OUI at this point. Again we can ignore it since the OS patches are included in the installed OS level.



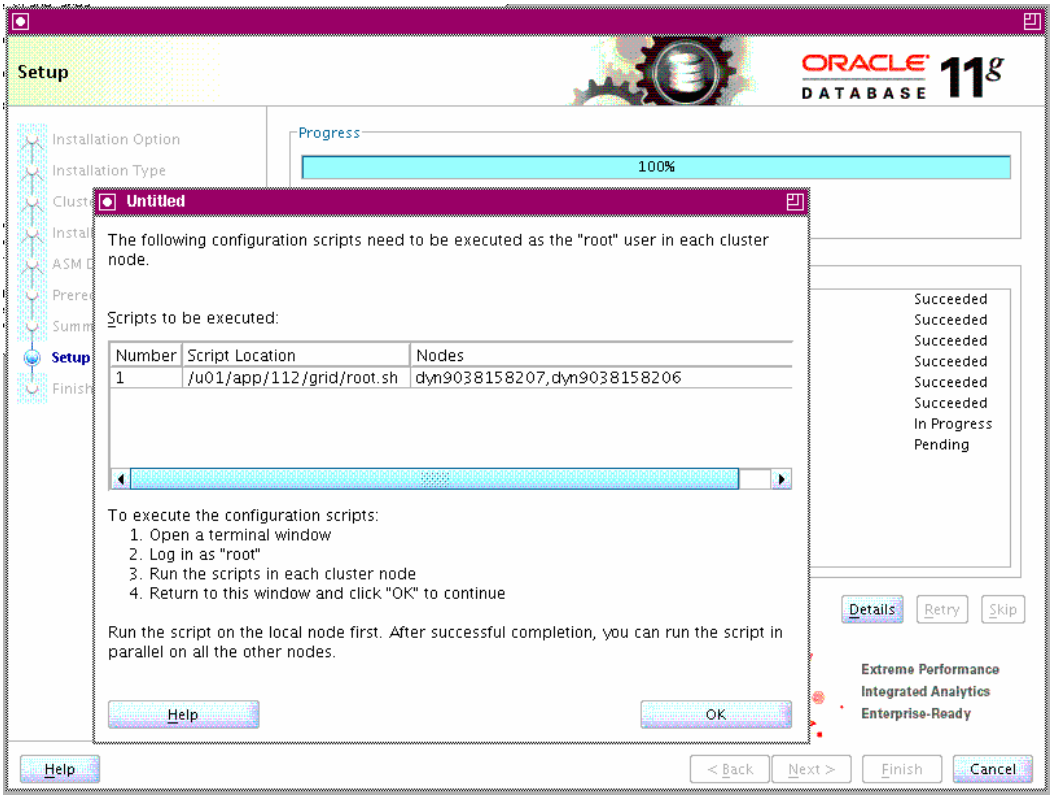
After fixing all the errors and passing the prerequisites tests the installation summary is shown. You can save the response file for future silent installation if desired.



8. This is the screen showing the installation process.



9. After Oracle has installed the binary files on all cluster nodes, it will ask you to run "*root.sh*" as user root. It is very important to run *root.sh* on the local node first and allow it to successfully complete. Do not run *root.sh* on other nodes until *root.sh* on the local node has completed; otherwise, errors will occur on the other cluster nodes.



This is the output from *root.sh* on the local node which is *blade1* in this example:

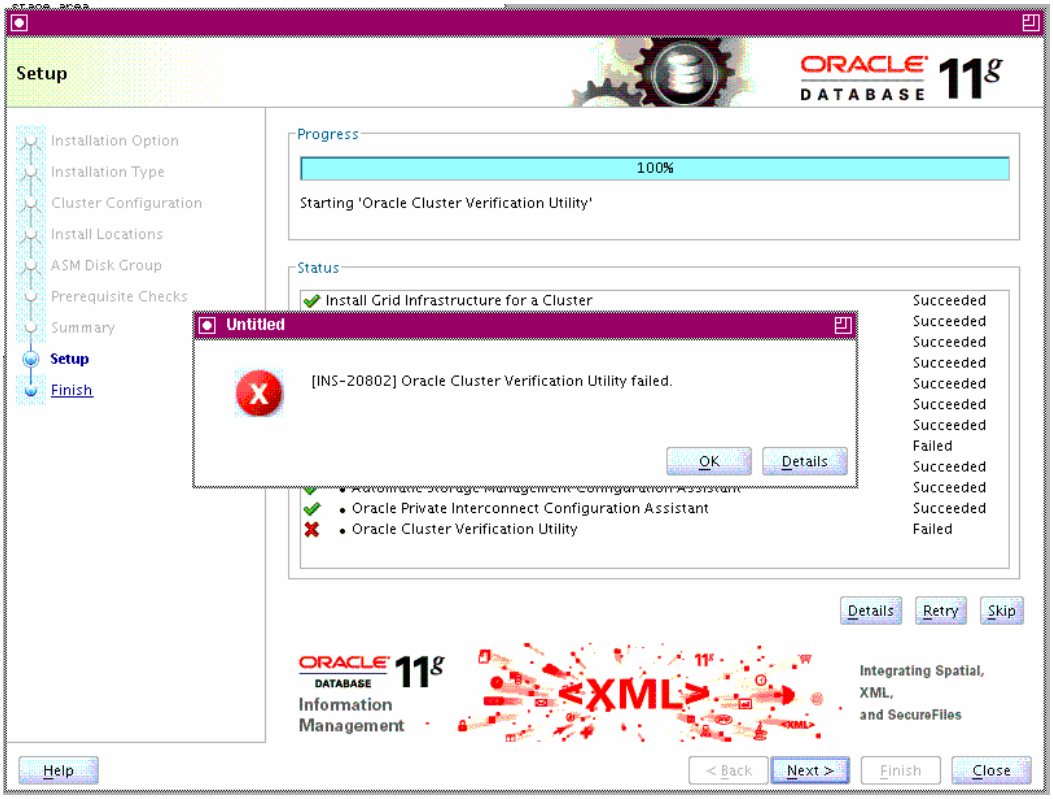
```
  # ./root.sh
Running Oracle 11g root.sh script...

The following environment variables are set as:
    ORACLE_OWNER= grid
    ORACLE_HOME=  /u01/app/112/grid

Enter the full pathname of the local bin directory: [/usr/local/bin]:
The file "dbhome" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]:
The file "oraenv" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]: y
   Copying oraenv to /usr/local/bin ...
The file "coraenv" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]: y
   Copying coraenv to /usr/local/bin ...
Entries will be added to the /etc/oratab file as needed by
Database Configuration Assistant when a database is created
Finished running generic part of root.sh script.
Now product-specific root actions will be performed.
2010-03-03 15:47:33: Parsing the host name
2010-03-03 15:47:33: Checking for super user privileges
2010-03-03 15:47:33: User has super user privileges
Using configuration parameter file:
/u01/app/112/grid/crs/install/crsconfig_params
```

```
Creating trace directory
User grid has the required capabilities to run CSSD in realtime mode
LOCAL ADD MODE
Creating OCR keys for user 'root', privgrp 'system'..
Operation successful.
  root wallet
  root wallet cert
  root cert export
  peer wallet
  profile reader wallet
  pa wallet
  peer wallet keys
  pa wallet keys
  peer cert request
  pa cert request
  peer cert
  pa cert
  peer root cert TP
  profile reader root cert TP
  pa root cert TP
  peer pa cert TP
  pa peer cert TP
  profile reader pa cert TP
  profile reader peer cert TP
  peer user cert
  pa user cert
Adding daemon to inittab
CRS-4123: Oracle High Availability Services has been started.
ohasd is starting
CRS-2672: Attempting to start 'ora.gipcd' on 'dyn9038158207'
CRS-2672: Attempting to start 'ora.mdnsd' on 'dyn9038158207'
CRS-2676: Start of 'ora.gipcd' on 'dyn9038158207' succeeded
CRS-2676: Start of 'ora.mdnsd' on 'dyn9038158207' succeeded
CRS-2672: Attempting to start 'ora.gpnpd' on 'dyn9038158207'
CRS-2676: Start of 'ora.gpnpd' on 'dyn9038158207' succeeded
CRS-2672: Attempting to start 'ora.cssdmonitor' on 'dyn9038158207'
CRS-2676: Start of 'ora.cssdmonitor' on 'dyn9038158207' succeeded
CRS-2672: Attempting to start 'ora.cssd' on 'dyn9038158207'
CRS-2672: Attempting to start 'ora.diskmon' on 'dyn9038158207'
CRS-2676: Start of 'ora.diskmon' on 'dyn9038158207' succeeded
CRS-2676: Start of 'ora.cssd' on 'dyn9038158207' succeeded
CRS-2672: Attempting to start 'ora.ctssd' on 'dyn9038158207'
CRS-2676: Start of 'ora.ctssd' on 'dyn9038158207' succeeded

ASM created and started successfully.

DiskGroup OCRVOTING created successfully.

clscfg: -install mode specified
Successfully accumulated necessary OCR keys.
Creating OCR keys for user 'root', privgrp 'system'..
Operation successful.
CRS-2672: Attempting to start 'ora.crsd' on 'dyn9038158207'
CRS-2676: Start of 'ora.crsd' on 'dyn9038158207' succeeded
CRS-4256: Updating the profile
Successful addition of voting disk 0bf0564ccce94f30bf951b946e9a9c44.
Successfully replaced voting disk group with +OCRVOTING.
CRS-4256: Updating the profile
CRS-4266: Voting file(s) successfully replaced
##  STATE    File Universal Id               File Name Disk group
--  -----    -----------------               --------- ---------
 1. ONLINE   0bf0564ccce94f30bf951b946e9a9c44 (/dev/rhdisk10) [OCRVOTING]
```

```
Located 1 voting disk(s).
CRS-2673: Attempting to stop 'ora.crsd' on 'dyn9038158207'
CRS-2677: Stop of 'ora.crsd' on 'dyn9038158207' succeeded
CRS-2673: Attempting to stop 'ora.asm' on 'dyn9038158207'
CRS-2677: Stop of 'ora.asm' on 'dyn9038158207' succeeded
CRS-2673: Attempting to stop 'ora.ctssd' on 'dyn9038158207'
CRS-2677: Stop of 'ora.ctssd' on 'dyn9038158207' succeeded
CRS-2673: Attempting to stop 'ora.cssdmonitor' on 'dyn9038158207'
CRS-2677: Stop of 'ora.cssdmonitor' on 'dyn9038158207' succeeded
CRS-2673: Attempting to stop 'ora.cssd' on 'dyn9038158207'
CRS-2677: Stop of 'ora.cssd' on 'dyn9038158207' succeeded
CRS-2673: Attempting to stop 'ora.gpnpd' on 'dyn9038158207'
CRS-2677: Stop of 'ora.gpnpd' on 'dyn9038158207' succeeded
CRS-2673: Attempting to stop 'ora.gipcd' on 'dyn9038158207'
CRS-2677: Stop of 'ora.gipcd' on 'dyn9038158207' succeeded
CRS-2673: Attempting to stop 'ora.mdnsd' on 'dyn9038158207'
CRS-2677: Stop of 'ora.mdnsd' on 'dyn9038158207' succeeded
CRS-2672: Attempting to start 'ora.mdnsd' on 'dyn9038158207'
CRS-2676: Start of 'ora.mdnsd' on 'dyn9038158207' succeeded
CRS-2672: Attempting to start 'ora.gipcd' on 'dyn9038158207'
CRS-2676: Start of 'ora.gipcd' on 'dyn9038158207' succeeded
CRS-2672: Attempting to start 'ora.gpnpd' on 'dyn9038158207'
CRS-2676: Start of 'ora.gpnpd' on 'dyn9038158207' succeeded
CRS-2672: Attempting to start 'ora.cssdmonitor' on 'dyn9038158207'
CRS-2676: Start of 'ora.cssdmonitor' on 'dyn9038158207' succeeded
CRS-2672: Attempting to start 'ora.cssd' on 'dyn9038158207'
CRS-2672: Attempting to start 'ora.diskmon' on 'dyn9038158207'
CRS-2676: Start of 'ora.diskmon' on 'dyn9038158207' succeeded
CRS-2676: Start of 'ora.cssd' on 'dyn9038158207' succeeded
CRS-2672: Attempting to start 'ora.ctssd' on 'dyn9038158207'
CRS-2676: Start of 'ora.ctssd' on 'dyn9038158207' succeeded
CRS-2672: Attempting to start 'ora.asm' on 'dyn9038158207'
CRS-2676: Start of 'ora.asm' on 'dyn9038158207' succeeded
CRS-2672: Attempting to start 'ora.crsd' on 'dyn9038158207'
CRS-2676: Start of 'ora.crsd' on 'dyn9038158207' succeeded
CRS-2672: Attempting to start 'ora.evmd' on 'dyn9038158207'
CRS-2676: Start of 'ora.evmd' on 'dyn9038158207' succeeded
CRS-2672: Attempting to start 'ora.asm' on 'dyn9038158207'
CRS-2676: Start of 'ora.asm' on 'dyn9038158207' succeeded
CRS-2672: Attempting to start 'ora.OCRVOTING.dg' on 'dyn9038158207'
CRS-2676: Start of 'ora.OCRVOTING.dg' on 'dyn9038158207' succeeded

dyn9038158207     2010/03/03 15:57:30
/u01/app/112/grid/cdata/dyn9038158207/backup_20100303_155730.olr
Configure Oracle Grid Infrastructure for a Cluster ... succeeded
Updating inventory properties for clusterware
Starting Oracle Universal Installer...

Checking swap space: must be greater than 500 MB.   Actual 12672 MB    Passed
The inventory pointer is located at /etc/oraInst.loc
The inventory is located at /u01/app/oraInventory
'UpdateNodeList' was successful.

This is the output of the second node which is “dyn9038158206”.

# ./root.sh
Running Oracle 11g root.sh script...

The following environment variables are set as:
    ORACLE_OWNER= grid
    ORACLE_HOME=  /u01/app/112/grid
```

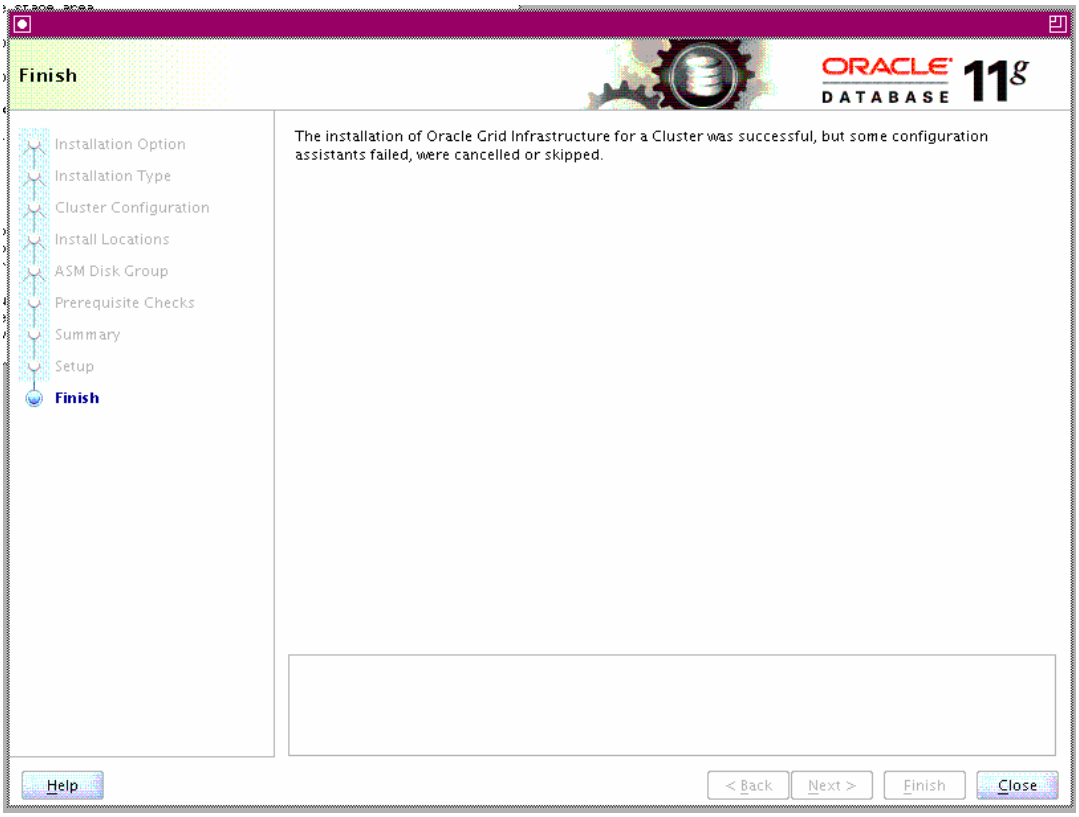
```
Enter the full pathname of the local bin directory: [/usr/local/bin]:
The file "dbhome" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]: y
   Copying dbhome to /usr/local/bin ...
The file "oraenv" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]: y
   Copying oraenv to /usr/local/bin ...
The file "coraenv" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]: y
   Copying coraenv to /usr/local/bin ...

Entries will be added to the /etc/oratab file as needed by
Database Configuration Assistant when a database is created
Finished running generic part of root.sh script.
Now product-specific root actions will be performed.
2010-03-03 15:58:35: Parsing the host name
2010-03-03 15:58:35: Checking for super user privileges
2010-03-03 15:58:35: User has super user privileges
Using configuration parameter file:
/u01/app/112/grid/crs/install/crsconfig_params
Creating trace directory
User grid has the required capabilities to run CSSD in realtime mode
LOCAL ADD MODE
Creating OCR keys for user 'root', privgrp 'system'..
Operation successful.
Adding daemon to inittab
CRS-4123: Oracle High Availability Services has been started.
ohasd is starting
CRS-4402: The CSS daemon was started in exclusive mode but found an active CSS
daemon on node dyn9038158207, number 1, and is terminating
An active cluster was found during exclusive startup, restarting to join the
cluster
CRS-2672: Attempting to start 'ora.mdnsd' on 'dyn9038158206'
CRS-2676: Start of 'ora.mdnsd' on 'dyn9038158206' succeeded
CRS-2672: Attempting to start 'ora.gipcd' on 'dyn9038158206'
CRS-2676: Start of 'ora.gipcd' on 'dyn9038158206' succeeded
CRS-2672: Attempting to start 'ora.gpnpd' on 'dyn9038158206'
CRS-2676: Start of 'ora.gpnpd' on 'dyn9038158206' succeeded
CRS-2672: Attempting to start 'ora.cssdmonitor' on 'dyn9038158206'
CRS-2676: Start of 'ora.cssdmonitor' on 'dyn9038158206' succeeded
CRS-2672: Attempting to start 'ora.cssd' on 'dyn9038158206'
CRS-2672: Attempting to start 'ora.diskmon' on 'dyn9038158206'
CRS-2676: Start of 'ora.diskmon' on 'dyn9038158206' succeeded
CRS-2676: Start of 'ora.cssd' on 'dyn9038158206' succeeded
CRS-2672: Attempting to start 'ora.ctssd' on 'dyn9038158206'
CRS-2676: Start of 'ora.ctssd' on 'dyn9038158206' succeeded
CRS-2672: Attempting to start 'ora.asm' on 'dyn9038158206'
CRS-2676: Start of 'ora.asm' on 'dyn9038158206' succeeded
CRS-2672: Attempting to start 'ora.crsd' on 'dyn9038158206'
CRS-2676: Start of 'ora.crsd' on 'dyn9038158206' succeeded
CRS-2672: Attempting to start 'ora.evmd' on 'dyn9038158206'
CRS-2676: Start of 'ora.evmd' on 'dyn9038158206' succeeded

dyn9038158206     2010/03/03 16:01:27
/u01/app/112/grid/cdata/dyn9038158206/backup_20100303_160127.olr
Configure Oracle Grid Infrastructure for a Cluster ... succeeded
Updating inventory properties for clusterware
Starting Oracle Universal Installer...

Checking swap space: must be greater than 500 MB.   Actual 12672 MB    Passed
The inventory pointer is located at /etc/oraInst.loc
The inventory is located at /u01/app/oraInventory
```

10. After executing *root.sh* on all cluster nodes, OUI will continue to configure the Oracle Grid Infrastructure for a cluster.



There are two places that show as “Failed”. The reason was “Oracle Cluster Verification Uutility” shown under the title “Configure Oracle Grid Infrastructure for a cluster”. Since “Oracle Cluster Verification Utility” is failed, in general it will also show that the “Configure Oracle Grid Infrastructure for a cluster” is failed. This can be ignored.

11. After you press OK and continue, the Oracle Grid Infrastructure installation has completed.



Please check the configuration log file for more details if there are any other failures during the installation and configuration process. The configuration log file is located in the Oracle Inventory location.

Performing post-installation tasks

To confirm Oracle Clusterware is running correctly, use this command:

$CRS_HOME/bin/crsctl status resource –w "TYPE co 'ora'" -t

```
$ ./crsctl status resource -w "TYPE co 'ora'" -t
--------------------------------------------------------------------------------
NAME           TARGET  STATE        SERVER                   STATE_DETAILS
--------------------------------------------------------------------------------
Local Resources
--------------------------------------------------------------------------------
ora.DATA.dg
               ONLINE  ONLINE       dyn9038158206
               ONLINE  ONLINE       dyn9038158207
ora.LISTENER.lsnr
               ONLINE  ONLINE       dyn9038158206
               ONLINE  ONLINE       dyn9038158207
ora.OCRVOTING.dg
               ONLINE  ONLINE       dyn9038158206
               ONLINE  ONLINE       dyn9038158207
ora.asm
               ONLINE  ONLINE       dyn9038158206            Started
```

```
               ONLINE  ONLINE       dyn9038158207            Started
ora.eons
               ONLINE  ONLINE       dyn9038158206
               ONLINE  ONLINE       dyn9038158207
ora.gsd
               OFFLINE OFFLINE      dyn9038158206
               OFFLINE OFFLINE      dyn9038158207
ora.net1.network
               ONLINE  ONLINE       dyn9038158206
               ONLINE  ONLINE       dyn9038158207
ora.ons
               ONLINE  ONLINE       dyn9038158206
               ONLINE  ONLINE       dyn9038158207
--------------------------------------------------------------------------------
Cluster Resources
--------------------------------------------------------------------------------
ora.LISTENER_SCAN1.lsnr
      1        ONLINE  ONLINE       dyn9038158207
ora.dyn9038158206.vip
      1        ONLINE  ONLINE       dyn9038158206
ora.dyn9038158207.vip
      1        ONLINE  ONLINE       dyn9038158207
ora.oc4j
      1        OFFLINE OFFLINE
ora.orcl.db
      1        ONLINE  ONLINE       dyn9038158207            Open
      2        ONLINE  ONLINE       dyn9038158206            Open
ora.scan1.vip
      1        ONLINE  ONLINE       dyn9038158207
```

Another command, "*crsctl check cluster -all*", can also be used for cluster check.

```
 # $ crsctl check cluster -all
**************************************************************
dyn9038158206:
CRS-4537: Cluster Ready Services is online
CRS-4529: Cluster Synchronization Services is online
CRS-4533: Event Manager is online
**************************************************************
dyn9038158207:
CRS-4537: Cluster Ready Services is online
CRS-4529: Cluster Synchronization Services is online
CRS-4533: Event Manager is online
**************************************************************
```

Finally, the command, "*crsctl check crs*", can also be used for a less detailed system check.

```
[oracle@blade1 bin]$ $. /crsctl check crs
CRS-4638: Oracle High Availability Services is online
CRS-4537: Cluster Ready Services is online
CRS-4529: Cluster Synchronization Services is online
CRS-4533: Event Manager is online
```

After the installation of Oracle Clusterware and Oracle Database software, a backup is made for the contents of *root.sh* and *emkey.ora* for future use. *Emkey.ora* is located in the $ORACLE_HOME/<*node_name*>_<*database_name*>/sysman/config directory. In this example, *emkey.ora* is located under the /d01/app/112/dbhome/dyn9038158207_orcl/sysman/config directory on all of the nodes. This file contains the encryption key for all enterprise manager data.

Pre-Installation tasks

All of the pre-installation tasks for Oracle Database 11*g* Release 2 are done before installing the Oracle Grid Infrastructure software. No other specific tasks are needed except the cluster verification test for pre-database configuration.

If you have decided to use ASM for storing database files, create the diskgroup using the *asmca* utility before starting to install and create the database. The ASM diskgroup name for the database files will be asked for in one of the following screens.

Running Cluster Verification Utility

The Cluster Verification Utility can be used to verify if the systems are ready to install Oracle Database 11*g* Release 2 with Oracle RAC.

The command "*cluvfy.sh stage –pre dbcfg –n nodelist –d $ORACLE_HOME*" is used to pre-check requirements for an Oracle Database with Oracle RAC installation. Login as user *oracle* and run the *cluvfy* command.

```
$ ./runcluvfy.sh stage -pre dbcfg -n dyn9038158207,dyn9038158206 -d
/d01/app/112/dbhome

Performing pre-checks for database configuration

Checking node reachability...
Node reachability check passed from node "dyn9038158207"


Checking user equivalence...
User equivalence check passed for user "oracle"
Total memory check passed
Available memory check passed
Swap space check passed
Free disk space check passed for "dyn9038158207:/d01/app/112/dbhome"
Free disk space check passed for "dyn9038158206:/d01/app/112/dbhome"
Free disk space check passed for "dyn9038158207:/u01/app/112/grid"
Free disk space check passed for "dyn9038158206:/u01/app/112/grid"
Free disk space check passed for "dyn9038158207:/tmp/"
Free disk space check passed for "dyn9038158206:/tmp/"
Check for multiple users with UID value 1101 passed
User existence check passed for "oracle"
Group existence check passed for "oinstall"
Group existence check passed for "dba"
Membership check for user "oracle" in group "oinstall" [as Primary] passed
Membership check for user "oracle" in group "dba" passed
Run level check passed
Hard limits check passed for "maximum open file descriptors"
Soft limits check passed for "maximum open file descriptors"
Hard limits check passed for "maximum user processes"
Soft limits check passed for "maximum user processes"
System architecture check passed
Kernel version check passed
Kernel parameter check passed for "SEM_NSEMS_MAX"
Kernel parameter check passed for "SEM_VALUE_MAX"
Kernel parameter check passed for "ncargs"
```

```
Package existence check passed for "bos.adt.base-..."
Package existence check passed for "bos.adt.lib-..."
Package existence check passed for "bos.adt.libm-..."
Package existence check passed for "bos.perf.libperfstat-6.1.2.1"
Package existence check passed for "bos.perf.perfstat-..."
Package existence check passed for "bos.perf.proctools-..."
Package existence check passed for "rsct.basic.rte-..."
Package existence check passed for "rsct.compat.clients.rte-..."
Package existence check passed for "xlC.aix61.rte-10.1.0.0"
Operating system patch check failed for "Patch IZ41855"
Check failed on nodes:
        dyn9038158207,dyn9038158206
Operating system patch check failed for "Patch IZ51456"
Check failed on nodes:
        dyn9038158207,dyn9038158206
Operating system patch check failed for "Patch IZ52319"
Check failed on nodes:
        dyn9038158207,dyn9038158206
Check for multiple users with UID value 0 passed
Current group ID check passed

Checking CRS integrity...

CRS integrity check passed

Checking node application existence...

Checking existence of VIP node application (required)
Check passed.

Checking existence of ONS node application (optional)
Check passed.

Checking existence of GSD node application (optional)
Check ignored.

Checking existence of EONS node application (optional)
Check passed.

Checking existence of NETWORK node application (optional)
Check passed.

Checking time zone consistency...
Time zone consistency check passed.

Pre-check for database configuration was unsuccessful on all the nodes.
```

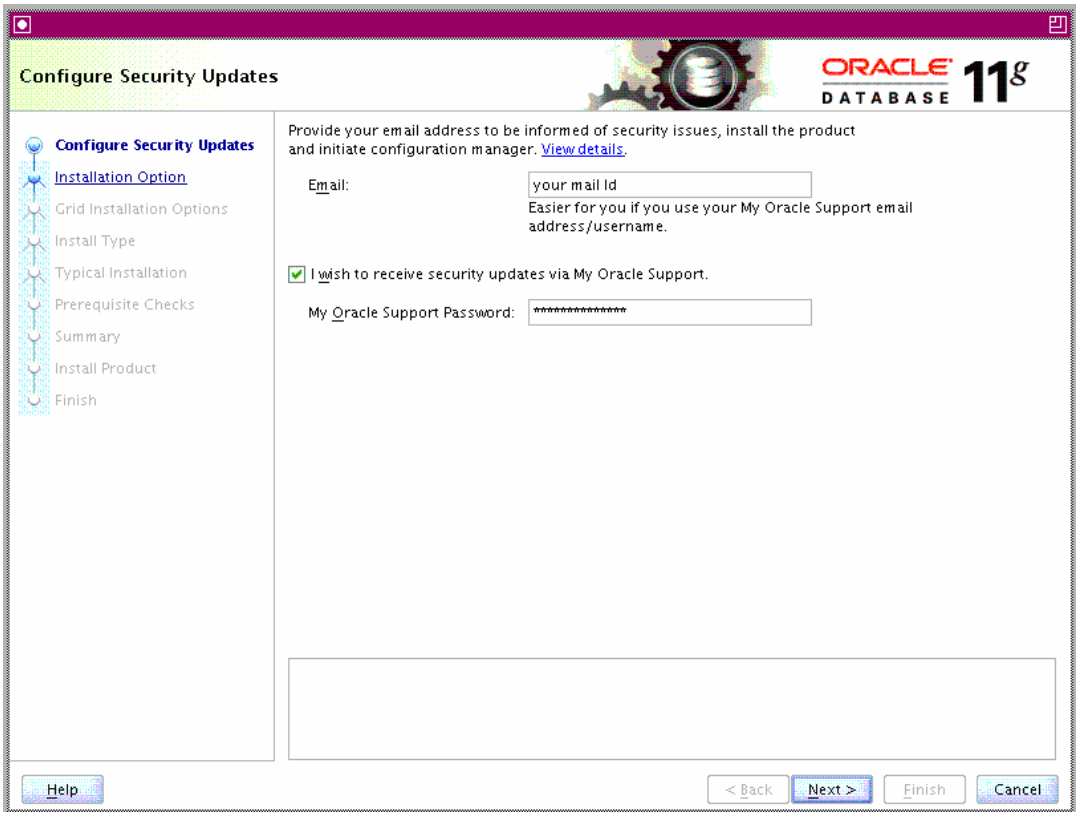
The cluster verification is unsuccessful for the same reason as seen in the Oracle Grid Infrastructure installation. The required OS patches are included in the currently installed higher version of the OS level. Again we can ignore the failures.

Preparing Oracle home and its path

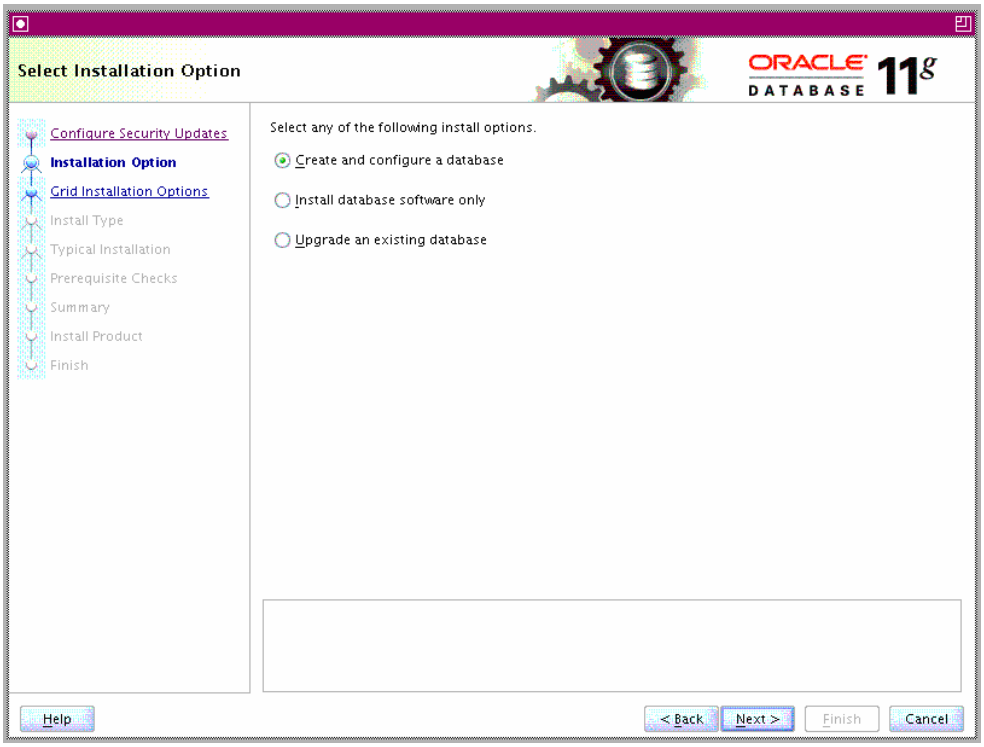
The Oracle home path must be different from the Oracle Clusterware home. In other words, Oracle Database 11*g* Release 2 with RAC cannot be installed onto the same home as the Oracle Clusterware software.

Performing database installation

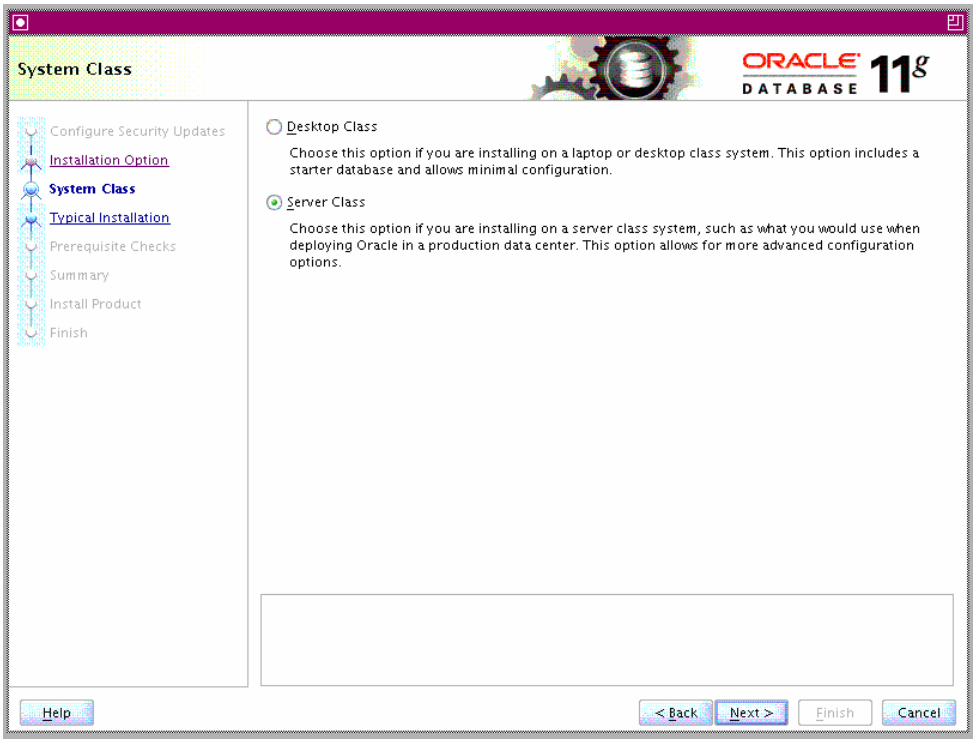
1. Download and unzip "*aix.ppc64_11gR2_database_1of2.zip*" and "*aix.ppc64_11gR2_database_2of2.zip*" from technet.oracle.com and go to the database directory and execute *./rootpre.sh*.
2. Login as user *oracle*. The installation needs to be run in XWindows or through vnc.
3. Execute *./runInstaller*, the first screen asks for your email address. You have to provide your email address in order to proceed. If you want to receive security updates from My Oracle Support, you will need to provide the password of your email address (username) for the My Oracle Support web site.



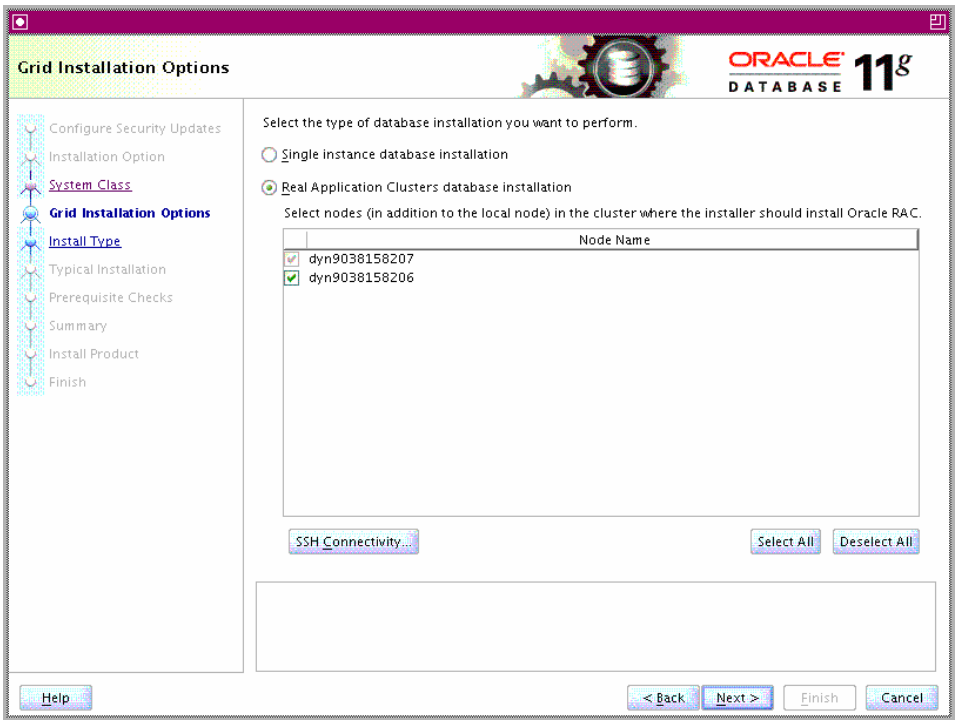
4. The next screen provides the user different installation options. In this example, we will be creating and configuring a database.



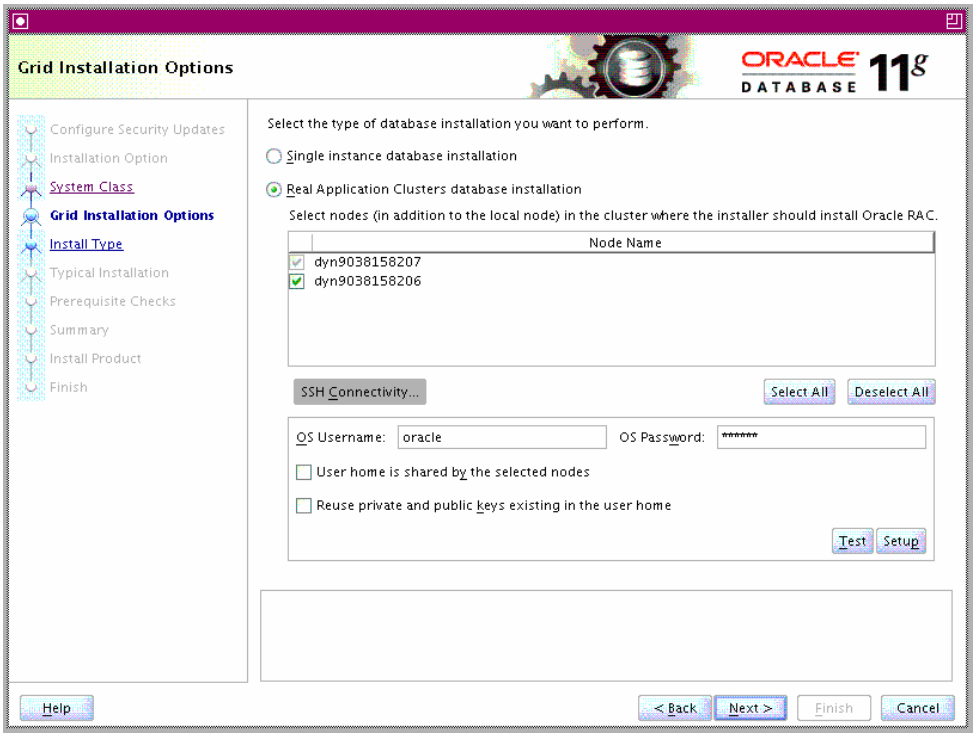
5. The next screen asks for the class of the database server. For this example, Server Class will be selected.



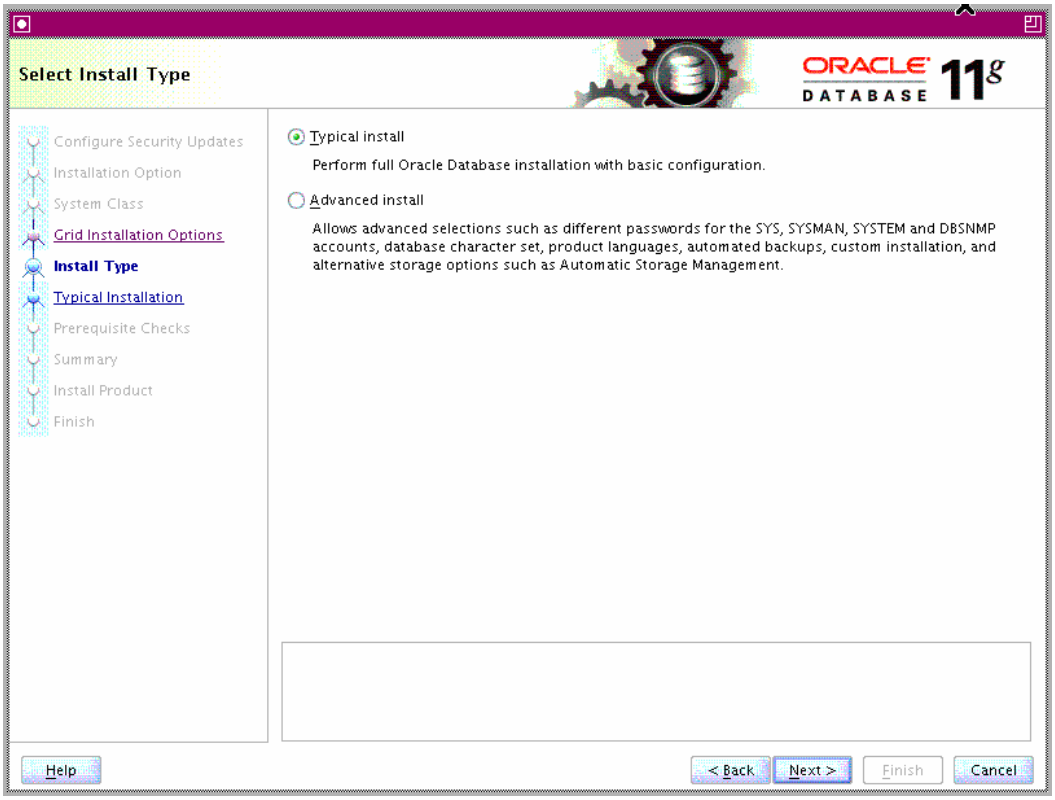
6. The next screen asks if you want to install and configure a single instance or an Oracle Database with RAC. In this example, we are going to install the Oracle Database with RAC on *blade1* and *blade2*.



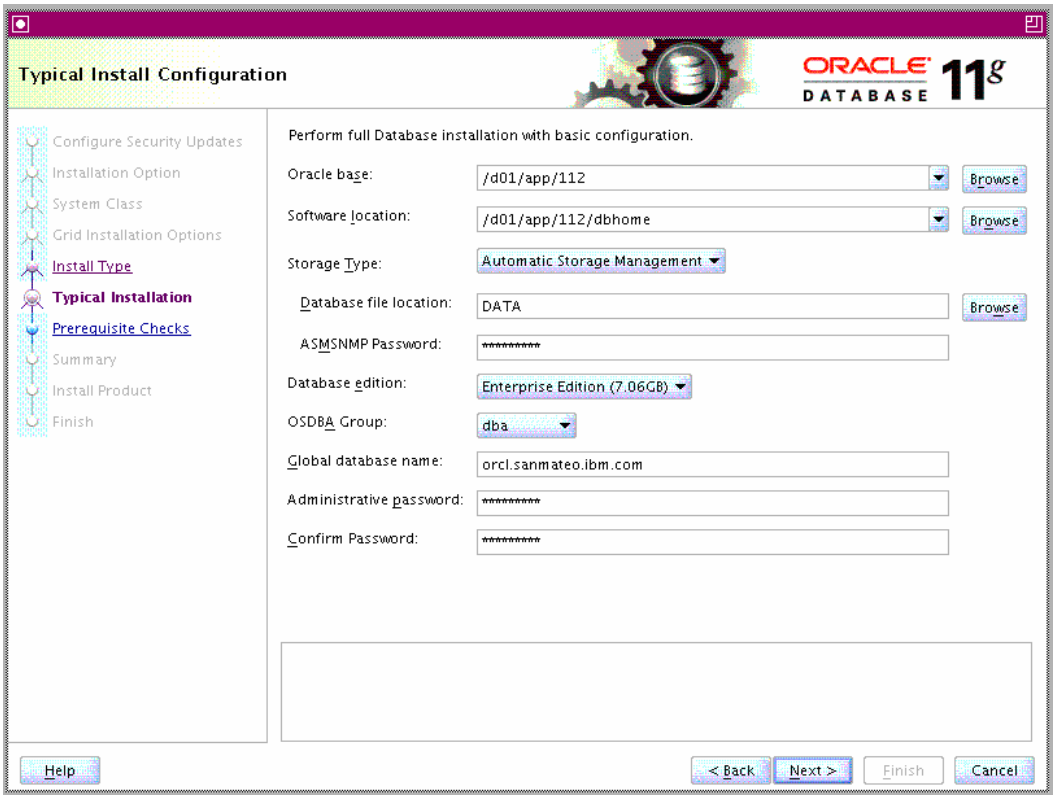
7. The database software installation also checks the SSH setup while clicking “Next”. Give password for the user *oracle* in the password field in the SSH connectivity section of the screen.



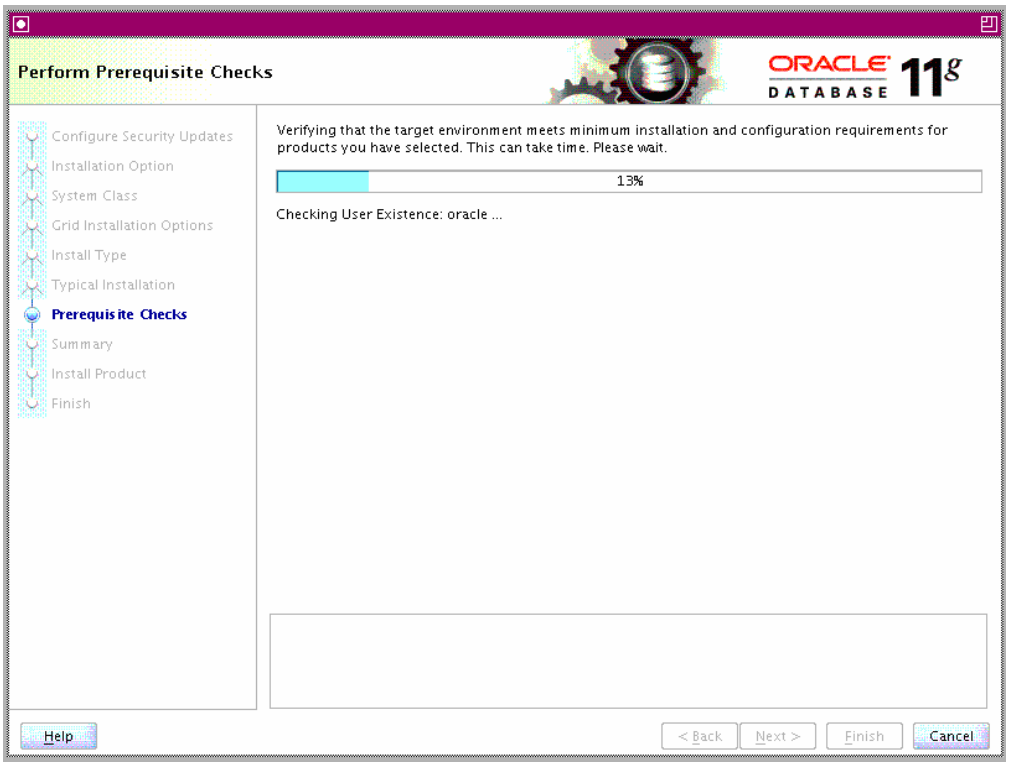
8. The next screen asks for the type of installation.



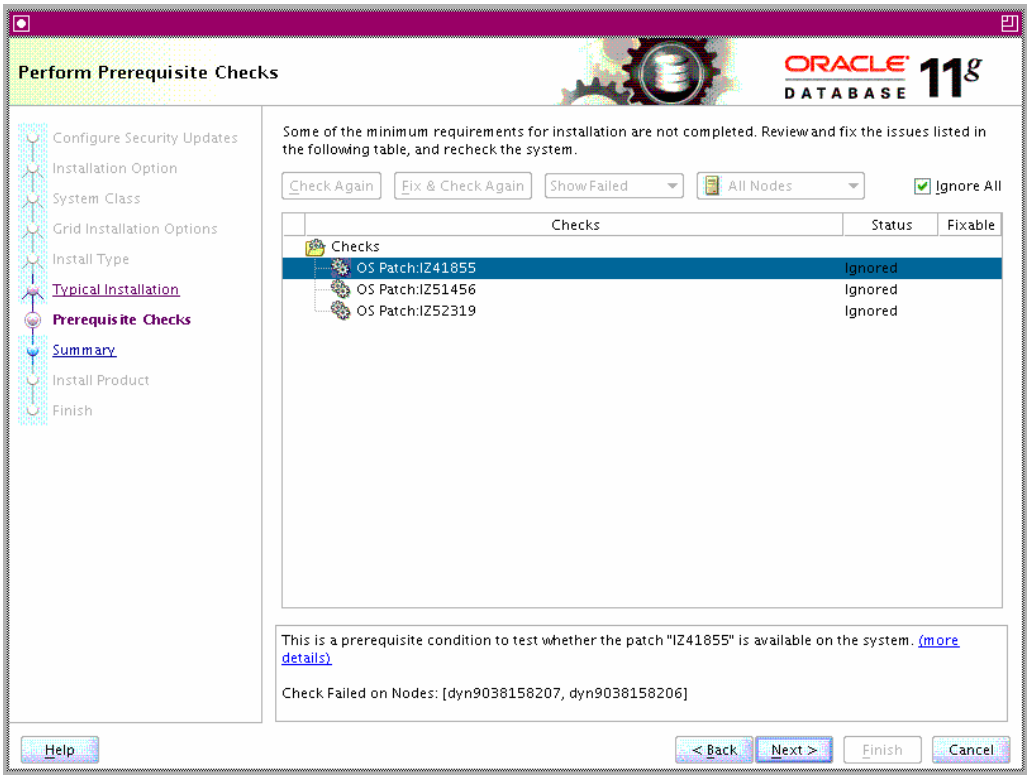
9. The next screen asks for the configuration details of the database installation. The software location must be different from the software location of the grid infrastructure. If the storage type is ASM, the ASM disk group needs to be provided in the space “Database file locations”. If you have not done so, please create ASM disk groups by using the Oracle ASM configuration assistant (ASMCA). You can use the diskgroup created to store the OCR and voting disks during the grid infrastructure install.



10. This screen performs all the prerequisite checks on all cluster nodes before installation.

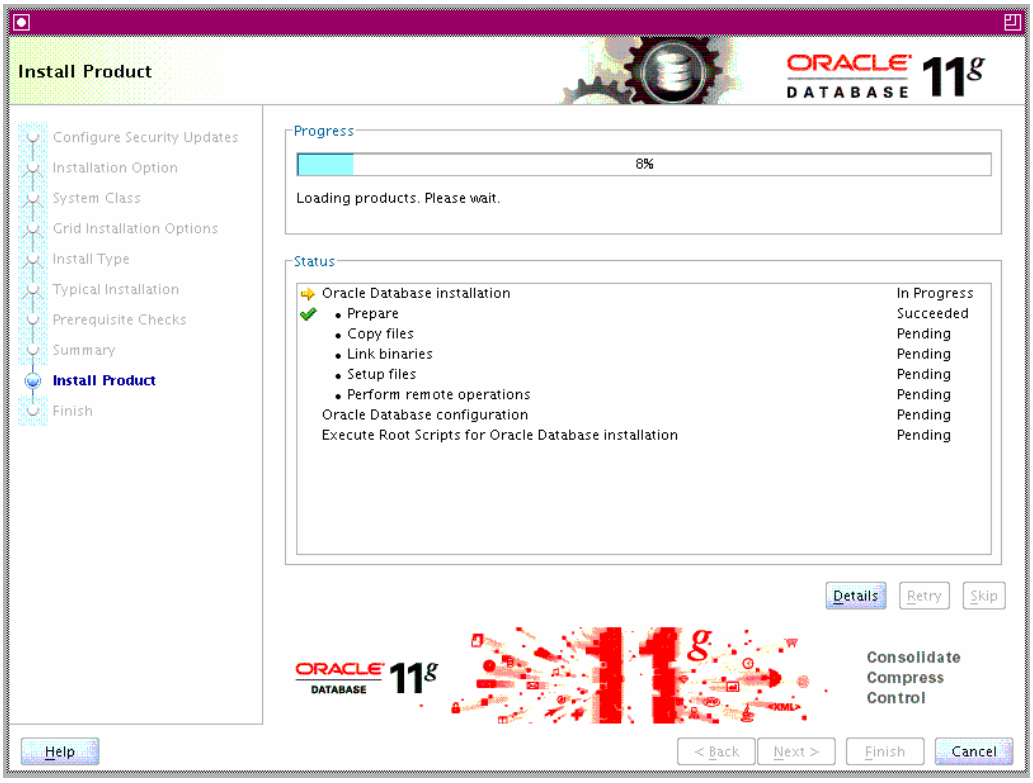


As seen in the Oracle Grid Infrastructure installation and output of the Cluster Verification tool, the OUI also shows the missing OS patches. This can be safely ignored.

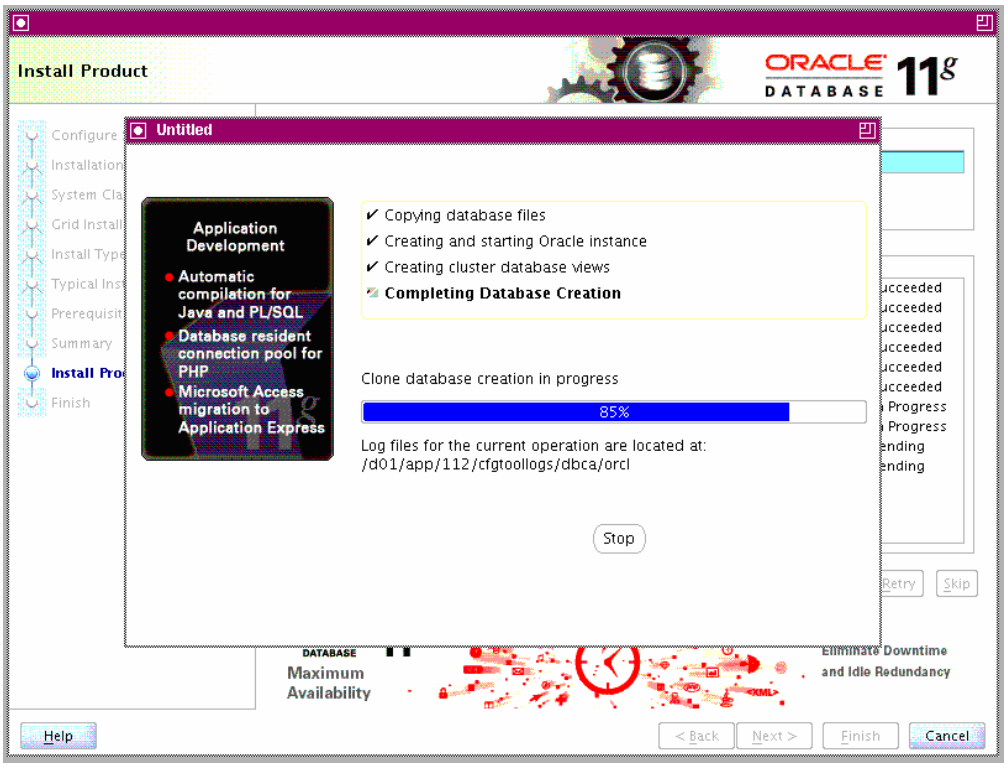


11. The next screen shows the installation summary for the database install.

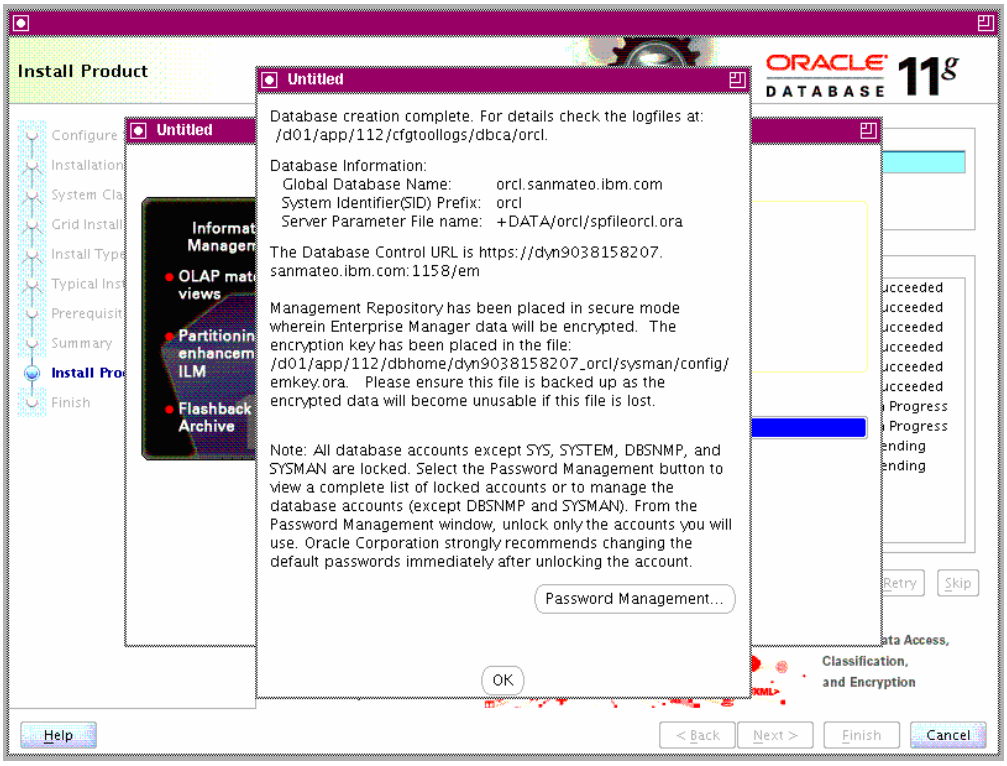
12. This screen shows the installation process of the Oracle RAC installation.



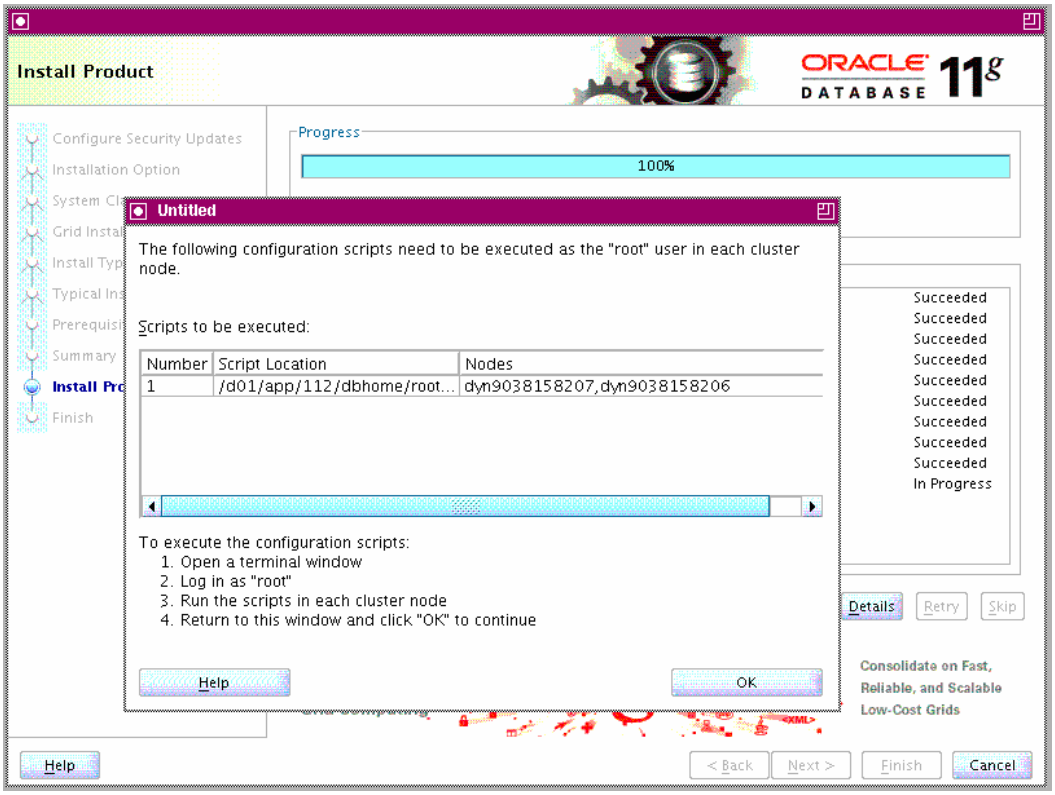
13. This screen shows the progress of the Oracle Database configuration.



14. This screen shows the completion of the Oracle Database configuration.



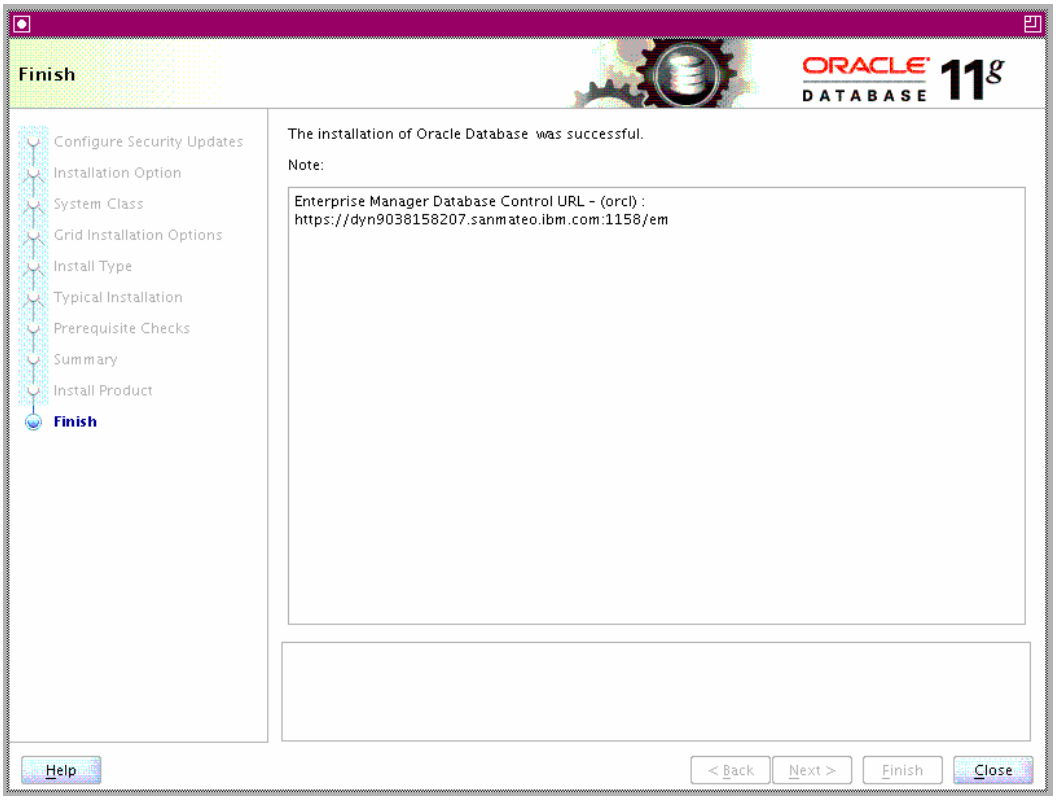
15. This is the last step of the database installation process. Execute *root.sh* from the software location that you provided previously on all cluster nodes as user *root*.



The outputs from all the cluster nodes should be the same. This is the output from running *root.sh* on node “dyn9038158207”. The output looked the same on node “dyn9038158206”.

```
# ./root.sh
Running Oracle 11g root.sh script...

The following environment variables are set as:
    ORACLE_OWNER= oracle
    ORACLE_HOME=  /d01/app/112/dbhome

Enter the full pathname of the local bin directory: [/usr/local/bin]:
The file "dbhome" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]: y
   Copying dbhome to /usr/local/bin ...
The file "oraenv" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]: y
   Copying oraenv to /usr/local/bin ...
The file "coraenv" already exists in /usr/local/bin.  Overwrite it? (y/n) [n]: y
   Copying coraenv to /usr/local/bin ...

Entries will be added to the /etc/oratab file as needed by
Database Configuration Assistant when a database is created
Finished running generic part of root.sh script.
Now product-specific root actions will be performed.
Finished product-specific root actions.
```

16. This is the end of the database installation process.



Post-installation tasks

1. Oracle recommends the *root.sh* script to be backed up after completing the database installation. If the information is needed in the future, the original *root.sh* script can be easily recovered.
2. After upgrading or creating databases, it is recommended that utlrp.sql be executed to compile or re-compile all PL/SQL modules that might be in an invalid state including packages, procedures and types. This script is located in the $ORACLE_HOME/rdbms/admin directory.
3. Finally, user accounts need to be created for the database and system. Most of the administrator accounts in the new database have been locked except sys, system. They will need to be unlocked if the modules for the administrators are going to be implemented.
4. The port numbers of several Web-based applications including Oracle Enterprise Manager Database Control are recorded in $ORACLE_HOME/install/portlist.ini. Make a note of these port numbers for future reference.

Summary

Oracle Database 11*g* Release 2 offers many new features. Many of the new features further optimize the performance, scalability and failover mechanisms of Oracle Real Application Clusters. These new features make implementing Oracle RAC easier and give you the flexibility to add nodes. Integrated with Oracle Fusion Middleware, Oracle RAC can fail over connections in the connection pools and immediately take appropriate recovery action.

It is important to make sure that the Oracle Clusterware installation is successful and functional before proceeding to the Oracle Database installation. This is because Oracle Clusterware daemons make sure that all applications startup during system startup and any failed applications will be started automatically to maintain the high availability aspect of the cluster.

Last but not least, choosing the hardware, operating systems and storage for the Oracle RAC deployment is a very significant step. Having the right combination of all options will contribute to the success of the installation and implementation on the IBM Power Systems, AIX and IBM System Storage platforms.

References

Oracle documentation

- Oracle Database New Features Guide 11*g* Release 2 (11.2), E10881-03
- GPFS for Oracle RAC, refer "My Oracle support" document 302806.1.
- Oracle® Grid Infrastructure Installation Guide 11g Release 2 (11.2) for IBM AIX on POWER Systems (64-Bit), E10814-01
- Oracle® Real Application Clusters Installation Guide 11g Release 2 (11.2) for Linux and UNIX, E10813-04
- For more information on Oracle ACFS, please refer to Oracle Database Storage Administrator's Guide 11*g* Release 2 (11.2), Part Number E10500-02.
- Oracle 11*g* Release2 Document library for more documents

IBM documentation

- For more information on HMS, visit the following link, "*Hardware Management Console V7 Handbook*".
- Oracle Database 11*g* Release 2 Enterprise Edition using Oracle Real Application Clusters on IBM BladeCenter running Red Hat Enterprise Linux 5 and IBM System Storage DS4800 by Betty Lee. Document ID: WP101608

IBM and Oracle Web sites

These Web sites provide useful references to supplement the information contained in this document:

- IBM Power Systems p570:
  http://www-03.ibm.com/systems/power/hardware/570/index.html

- IBM System Storage product offerings:
  http://www-03.ibm.com/systems/storage/disk

- IBM System Storage DS6000 series:
  http://www-03.ibm.com/systems/storage/disk/ds6000/index.html

- Interoperability matrix for IBM System Storage DS6800:
  http://www-03.ibm.com/systems/resources/systems_storage_disk_ds6000_pdf_interop.pdf

- IBM NAS offerings such as IBM System Storage N3000, N3700, N5000 and N7000:
  http://www-03.ibm.com/systems/storage/nas

- IBM SDDPCM Multipath driver:
  http://www-01.ibm.com/support/docview.wss?uid=ssg1S4000201#DS6K

- IBM RedBooks
  http://www.redbooks.ibm.com

- IBM Techdocs (White Papers)
  http://www-03.ibm.com/support/techdocs/atsmastr.nsf/Web/WhitePapers

- IBM ISV Solutions for Oracle
  http://www-03.ibm.com/systems/storage/solutions/isv/#oracle

- Oracle Real Application Clusters
  http://www.oracle.com/technology/products/database/clustering

- Technology supported by Oracle with Oracle Real Application Clusters, please visit:
  http://www.oracle.com/technology/products/database/clustering/certify/tech_generic_unix_new.html

- Oracle Automatic Storage Management (ASM)
  http://www.oracle.com/technology/products/database/asm

- My Oracle Support (formerly Oracle Metalink)
  https://support.oracle.com/CSP/ui/flash.html

About the author

Ravisankar Shanmugam is a Senior IT Specialist with IBM Advanced Technical Support and works in the IBM Oracle International Competency Center based in San Mateo, CA. He provides Power Systems and System x platform support for projects at the Competency Center and for enablement activities at Oracle Corporation in Redwood Shores, CA.

Appendix A: List of common abbreviations and acronyms

ASM	**Automatic Storage Management** A feature of Oracle Database 11*g* that provides an integrated cluster file system and volume management capabilities.
FC	**Fibre Channel** A gigabit-speed network technology primarily used for storage networking.
GHz	**Gigahertz** Represent computer processor speed.
HBA	**Host bus adapter** It connects a host system to other network and storage devices.
HDD	**Hard Disk Drive** A non-volatile storage device which stores digitally encoded data on rapidly rotating platters with magnetic surfaces.
I/O	**Input / Output** The communication between an information processing system and the outside world.
iSCSI	**Internet Small Computer System Interface** An Internet Protocol (IP)-based storage networking standard for linking data storage facilities developed by the Internet Engineering Task Force (IETF).
LUN	**Logical Unit Number** It is a subnet of a larger physical disk or disk volume. It can be a single disk drive, or a partition of a single disk drive or disk volume from a RAID controller. It represents a logical abstraction or virtualization layer between the physical disk device/volume and the applications.
MB	**Megabyte** For processor storage, real and virtual storage, and channel volume, 2 to the 20th power or 1,048,576 bytes. For disk storage capacity and communications volume, 1 000 000 bytes.
Mb	**Megabit** For processor storage, real and virtual storage, and channel volume, 2 to the 20th power or 1 048 576 bits. For disk storage capacity and communications volume, 1 000 000 bits.
NAS	**Network-attached storage** File-level data storage connected to a computer network providing data access to heterogeneous network clients.
NIC	**Network interface controller** Hardware that provides the interface control between system main storage and external high-speed link (HSL) ports.
OCFS	**Oracle Cluster File System** A consistent file system image across the servers in a cluster.
OCFS2	**Oracle Cluster File System Release 2** The next generation of the Oracle Cluster File System for Linux. It is a general-purpose file system that can be used for shared Oracle home installations.
OCR	**Oracle Cluster Registry** A file that contains information pertaining to instance-to-node mapping, node list and resource profiles for customized applications in the Clusterware.
RAC	**Real Application Cluster** A cluster database with a shared cache architecture that supports the transparent deployment of a single database across a cluster of servers.
RDAC	**Redundant Disk Array Controller** It provides redundant failover/failback support for the logical drives of the storage server.
RHEL5	**Red Hat Enterprise Linux 5** Linux operating systems released in March 2007 and it is based on the Linux 2.6.18 kernel.
SAN	**Storage area network** A dedicated storage network tailored to a specific environment, combining servers, storage products, networking products, software, and services.
SAS	**Serial Attached SCSI** A communication protocol for direct attached storage (DAS) devices. It uses SCSI commands for interacting with SAS End devices.
SCSI	**Small Computer System Interface** (1) An ANSI-standard electronic interface that allows personal computers to communicate with peripheral hardware, such as disk drives, tape drives, CD-ROM drives, printers, and scanners faster and more flexibly than previous interfaces.
SLES	**SUSE Linux Enterprise Server** A Linux distribution supplied by Novell.

Trademarks and special notices

References in this document to IBM products or services do not imply that IBM intends to make them available in every country.

IBM. the IBM logo, ibm.com, AIX, BladeCenter, DS4000, DS6000, DS8000, POWER, System Storage, Power Systems and System x are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both:

Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Intel and Xeon are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

AMD and AMD Opteron are trademarks of Advanced Micro Devices, Inc.

Red Hat, the Red Hat "Shadow Man" logo, and all Red Hat-based trademarks and logos are trademarks or registered trademarks of Red Hat, Inc., in the United States and other countries.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.

The information provided in this document is distributed “AS IS” without any warranty, either express or implied.

The information in this document may include technical inaccuracies or typographical errors.

Information concerning non-IBM products was obtained from a supplier of these products, published announcement material, or other publicly available sources and does not constitute an endorsement of such products by IBM. Sources for non-IBM list prices and performance numbers are taken from publicly available information, including vendor announcements and vendor worldwide homepages. IBM has not tested these products and cannot confirm the accuracy of performance, capability, or any other claims related to non-IBM products. Questions on the capability of non-IBM products should be addressed to the supplier of those products.

All statements regarding IBM future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only. Contact your local IBM office or IBM authorized reseller for the full text of the specific Statement of Direction.

Some information addresses anticipated future capabilities. Such information is not intended as a definitive statement of a commitment to specific levels of performance, function or delivery schedules with respect to any future products. Such commitments are only made in IBM product announcements. The information is presented here to communicate IBM's current investment and development activities as a good faith effort to help with our customers' future planning.

Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput or performance improvements equivalent to the ratios stated here.

Photographs shown are of engineering prototypes. Changes may be incorporated in production models.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.