

Oracle Real Application Clusters on Oracle VM Environments

ORACLE WHITE PAPER | JANUARY 2017



Table of Contents

Table of Contents	2
Executive Summary	4
Introduction	5
What is Oracle Real Application Clusters (RAC)?	5
What is Oracle RAC One Node?	7
What are “Flex Cluster” and “Flex ASM”?	9
What is Oracle VM?	10
Why run Oracle RAC on Oracle VM?	13
Business Continuity, High Availability and Scalability	14
Live Migration, Flex Cluster and Online Database Relocation Support	15
Extended Distance Oracle RAC on Oracle VM	16
When to use <i>Live Migration</i> or Oracle RAC One Node <i>Database Relocation</i> ?	19
Deployment Methods	20
Hardware and Software Requirements	23
Hardware requirements for Production Environments	23
Hardware requirements for Development Environments	24
Oracle RAC Software	24
General Oracle VM	24
Oracle VM Guest Configuration	25
Oracle VM VCPU configuration for Oracle RAC	25
VCPU allocation examples	25
Oracle VM supported storage configurations for Oracle RAC	27
Dom-0 managed storage	27
Guest managed storage	27
Oracle RAC on Oracle VM – Best Practices	29
Oracle Installation requirements and recommendations	29
Linux Kernel parameters for Oracle VM guest environment	29
Time synchronization	29
Storage Configuration	30
Network Configuration	31



General HugePages and Transparent HugePages	33
Orachk utility and best practices	33
Sizing the Oracle VM Environment for Oracle RAC	34
CPUs and cores	34
Memory	34
Network	34
Storage	35
Recommended Installation: Oracle VM Templates	35
Oracle RAC on Oracle VM Environments: manual installation	36
Live Migration and Online Database Relocation Best Practices	37
Oracle VM Live Migration Best Practices	37
Oracle RAC One Node Online Database Relocation Best Practices	38
Summary	39
Appendix A – Disk persistence with Oracle VM Manager	40
Appendix B – Bonding Setup Example	41
Appendix C – References Official Oracle Documentation - General	43
Official Oracle Documentation – Oracle Real Application Cluster	43
My Oracle Support Notes	43

Executive Summary

Running today's non-critical business applications in virtualized environments has shown to be efficient and cost saving. More sophisticated or highly available applications on the other hand were most likely incompatible with commonly used software based virtualization solutions.



Note

Oracle VM and Oracle Real Application Clusters (RAC) enable the benefits of a virtualized datacenter infrastructure for highly available applications.

The availability of Oracle VM overcomes this obstacle. Providing software based virtualization infrastructure (Oracle VM) and the market leading high availability solution Oracle Real Application Clusters (RAC), Oracle now offers a highly available, cloud-ready virtualization solution for your data center, combining all the benefits of a fully virtualized environment.

The combination of Oracle VM and Oracle RAC enables a better server consolidation (Oracle RAC databases with underutilized CPU resources or peaky CPU utilization can often benefit from consolidation with other workloads using server virtualization) sub-capacity licensing, and rapid provisioning.

Oracle RAC on OVM also supports the creation of non-production virtual clusters on a single physical server for product demos, educational settings and test environments.

This deployment combination permits dynamic changes to pre-configured database resources for agile responses to changing service level requirements common in consolidated environments.

For x86-64 hardware platform, Oracle VM is the only software based virtualization solution that is fully certified for Oracle Real Application Clusters on Linux.

This paper discusses various Oracle RAC deployment scenarios and provides best practices for an optimized Oracle RAC deployment on Oracle VM environments for Oracle VM versions 3.2 and above. Version specific information is noted accordingly.



Note

Oracle constantly tests additional and advanced Oracle VM features with Oracle RAC. This paper is updated regularly as new test results are available.

Introduction

Oracle Real Application Clusters (RAC)¹ is an option to the award-winning Oracle Database Enterprise Edition. Oracle RAC is a cluster database with a shared cache architecture that overcomes the limitations of traditional shared-nothing and shared-disk approaches to provide highly scalable and available database solutions for all your business applications. Oracle RAC has evolved over the years and, based on a comprehensive high availability (HA) stack that, on top of Oracle VM Server, can be used as the foundation of a database cloud system as well as a shared infrastructure that can ensure high availability, scalability, flexibility and agility for any application in your data center.

Oracle Real Application Clusters family of Solutions includes **Oracle Clusterware**² which provides the group membership, high availability, and application resource management. Oracle Clusterware is the technology that transforms a server farm into a cluster. A cluster in general is a group of independent servers (virtual or physical) that cooperate as a single system for robust high availability and agile scalability. Oracle Clusterware is the intelligence in this system that ensures the required cooperation and is a key component of Oracle enterprise grid architecture.

Oracle RAC utilizes **Oracle Automatic Storage Management (ASM)**³ for efficient shared storage access. ASM acts as the underlying, clustered volume manager. It provides the database administrator with a simple storage management interface that is consistent across all server and storage platforms. As a vertically integrated file system and volume manager, purpose-built for Oracle database files, ASM provides the performance of raw I/O with the easy management of a file system. Oracle Automatic Storage Management provides the basis for a shared storage pool in Oracle enterprise Grid Architectures.

Engineered for open cloud infrastructure, **Oracle VM**⁴ is a platform that provides a fully equipped environment to better leverage the benefits of virtualization technology. Oracle VM enables deployment of operating systems and application software within a supported virtualization environment. Oracle VM complements the foundation of a database cloud system by providing an Oracle RAC certified virtualization environment.

What is Oracle Real Application Clusters (RAC)?

An Oracle RAC Database is a clustered database. A cluster is a group of independent servers that cooperate as a single system. Clusters provide improved fault resilience and modular incremental system growth over single symmetric multi-processor (SMP) systems.



Note

Oracle RAC enables the Oracle Database to run mainstream business applications of all kinds on clusters, including popular packaged products (such as Oracle Ebusiness Suite, Peoplesoft, Siebel, SAP) and in-house developed applications generating OLTP, DSS, or mixed workload.

¹ [Oracle Real Application Clusters \(RAC\) homepage](#)

² [More information on Oracle Clusterware](#)

³ [More information on Oracle Automatic Storage Management \(ASM\)](#)

⁴ [Oracle VM homepage](#)

In the event of a system failure, clustering ensures the highest availability to users and access to mission critical data is not lost. Redundant hardware components, such as additional nodes, interconnects, and disks, allow the cluster to provide high availability. Such redundant hardware architectures avoid single points-of-failure and provide exceptional fault resilience.

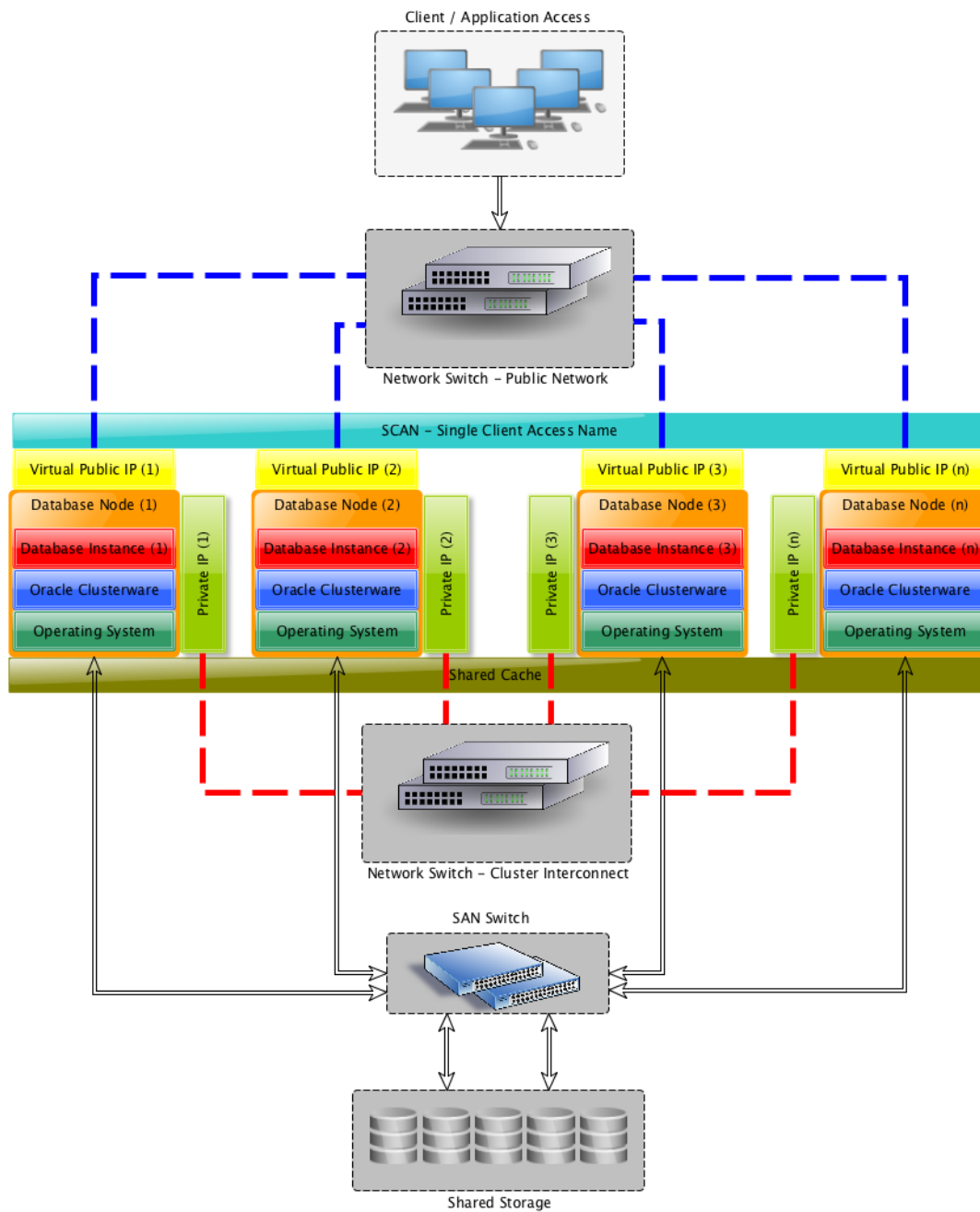


Figure 1: Oracle Real Application Clusters Overview

Oracle RAC allows users to access their data stored on shared storage from any instance of the cluster

However, a clustered database (using more than one instance) differs from a single instance database in a way that the database can be accessed by multiple instances concurrently. Each instance runs on a separate server in the cluster (formed by Oracle Clusterware).

When additional resources are required, additional nodes and instances can easily be added to the cluster with no downtime. Once a new instance has been started, applications using services can immediately take advantage of it with no changes to the application or application server.

Oracle Real Application Clusters is an extension to the Oracle Database and therefore benefits from the manageability, reliability, and security features built into the Oracle Database.

What is Oracle RAC One Node?

Oracle Real Application Clusters (RAC) One Node⁵ is an option introduced from the Oracle Database 11g Release 2 Enterprise Edition. It improves upon many of the benefits of server virtualization and extends them to databases running in physical server environments. Oracle RAC One Node enables:

- Better server consolidation
- Enhanced protection from failures
- Greater flexibility and workload management
- Better online maintenance

In addition, it allows customers to virtualize database storage, standardize their database environment, and, should the need arise, upgrade to a full multi-node Oracle RAC database without downtime or disruption. Further, it is fully compatible with and complementary to Oracle Virtual Machine (VM), allowing customers to build environments that leverage the strengths of both Oracle RAC One Node and Oracle VM Server virtualization.

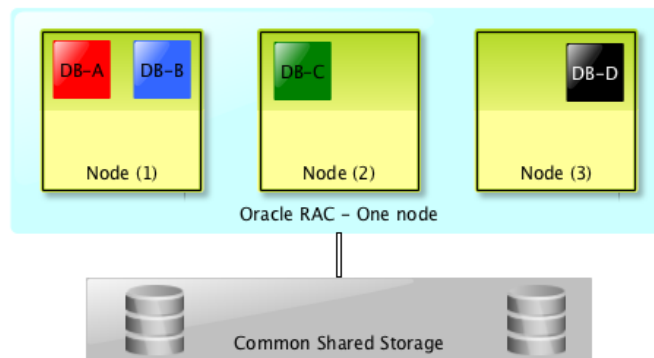


Figure 2: Oracle RAC One Node - typical architecture overview

In the configuration shown in [Figure 2](#), four single-instance Oracle RAC One Node databases are running in a cluster of three servers. Server A is hosting Oracle RAC One Node databases DB-A and DB-B, server B is hosting database DB-C and server C is hosting databases DB-D. Each server runs one OS. In servers A and C above, multiple databases are consolidated onto a single OS.

⁵ [More information on Oracle RAC One Node](#)

What is an Extended Distance Oracle RAC ?

Oracle RAC on Extended Distance (Stretched) Clusters is an architecture that provides extremely fast recovery from a site failure and allows for all nodes, in all sites, to actively process transactions as part of a single database cluster.

Oracle RAC on Extended Distance Clusters provides a greater high availability than a local Oracle RAC implementation. However, it does not provide full Disaster Recovery. Feasible separation is a great protection for some disasters (local power outage, , server room flooding) but not all. Disasters such as earthquakes, hurricanes, and regional floods may affect a greater area. Oracle RAC on Extended Distance Clusters does not protect from human errors or corruptions in the shared storage, either, as an Oracle RAC system, even on Extended Distance Clusters, is still a tightly coupled and enclosed system.

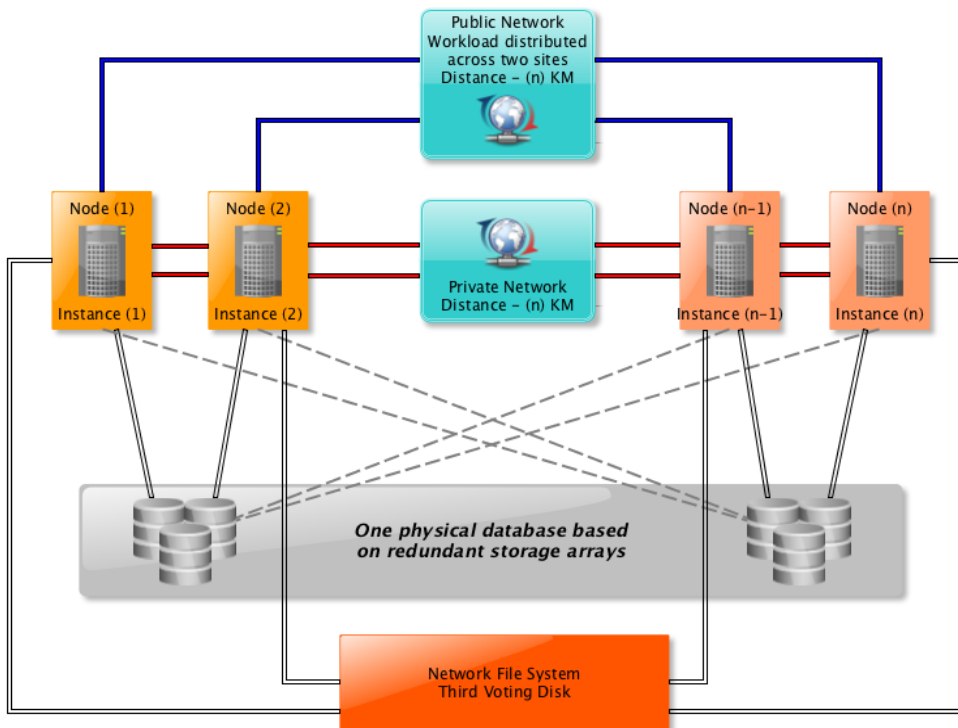


Figure 3: Oracle Extended RAC - architecture overview

For comprehensive data protection, including protection against corruptions and regional disasters, Oracle recommends using Oracle Data Guard together with Oracle RAC as building blocks of Oracle's Maximum Availability Architecture (MAA). For active-active databases deployed across geographically separated data centers, the MAA recommendation is to use Oracle GoldenGate. Note that Data Guard and GoldenGate also provide additional benefits such as minimizing downtime for maintenance activities such as upgrades and migrations.

Clusters of this kind have been referred to by many names, including “campus clusters”, “metro clusters”, “geo clusters”, “stretched clusters” and “extended clusters”.

Oracle RAC One Node completes the Extended Distance Cluster offering. While using the same infrastructure like Oracle RAC (i.e. Oracle Clusterware and Oracle ASM), Oracle RAC One Node is best used in environments or for applications that do not require the degree of HA or scalability that an active / active Oracle RAC database provides, but which benefit from an integrated and fully supported failover solution for the Oracle database.



Note

It's important to know that while Oracle VM Clustered Pool⁶ is not supported as stretched/extended configuration, Oracle RAC Extended can be deployed on top of two different Oracle VM Pools, where each one resides into a data center.

What are “Flex Cluster” and “Flex ASM”?

Oracle RAC is constantly improving on every version and this time was not different. 12c Release incorporates two properties called “Flex ASM” and “Flex Cluster” that gives support to the demand requirements on Cloud Computing oriented environments. Oracle RAC 12c introduces two new concepts:

- **Hub Nodes:** They are connected among them via private network and have direct access to the shared storage just like previous versions. These nodes are the ones that access the Oracle Cluster Registry (OCR) and Voting Disk (VD) directly.
- **Leaf Nodes:** These nodes are lighter and are not connected among them, neither access the shared storage like the Hub Nodes. Each Leaf Node communicates with the Hub Node that is attached to, and its connected to the cluster via the Hub Node that is linked to.

Tightly coupled servers are Hub Servers that share storage for database, OCR and Voting devices as well as peer-to-peer communication with other Hub Servers in the cluster. A loosely coupled server is a Leaf Server that has a loose communication association with a single Hub Server in the cluster and does not require shared storage nor peer-to-peer communication with other Hub or Leaf Servers in the cluster, except to communicate with the Hub to which it is associated. In 12.1, Leaf Servers are designed for greater application high availability and multi-tier resource management.

Architecturally Oracle Flex Cluster comprises of a Hub and Leaf architecture where only the Hub nodes will only have direct access to Oracle Cluster Registry (OCR) and Voting Disk (VD). However application can access the database via Leaf nodes without ASM instance NOT running on Leaf nodes. The connection to the database is through Hub making it transparent for the application.

⁶ [Oracle VM Concepts Guide: understanding Server Pools and Oracle VM Servers](#)

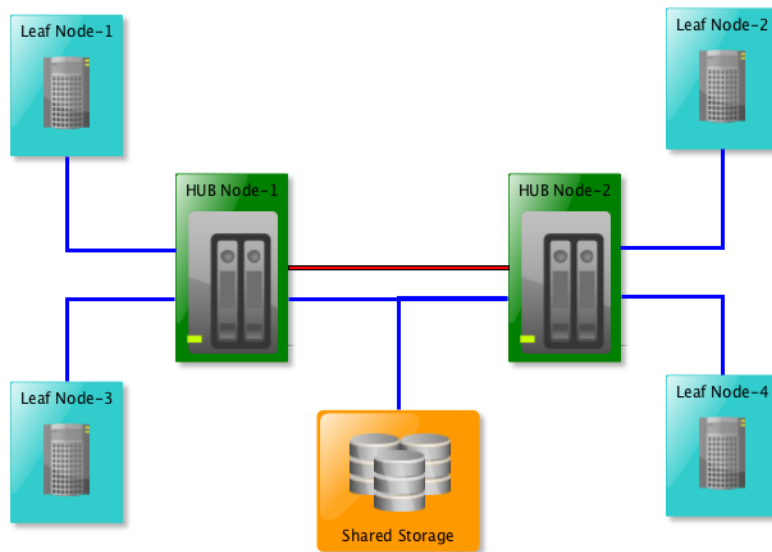


Figure 4: Oracle Flex Cluster - architecture overview

Prior to Oracle 12c, for a database instance to use ASM it is expected that the ASM instance must be up and running on all nodes before the database instance is brought up. Failure of ASM instance to come-up means that database instance using ASM at the storage level cannot be brought up. This literally means that the database instance is not accessible immaterial of the technologies put in use i.e. RAC, ASM and Shared Storage.


With the launch of Oracle 12c the above constraint has been addressed with the feature called Oracle Flex ASM which primarily has a feature to fail over to another node in the cluster. Essentially a Hub and Leaf architecture, the connection of a failed node is seamlessly transferred to another participating node by way of a replacement ASM instance by Oracle Clusterware. The number of ASM instances running in a given cluster is called ASM cardinality with a default value of 3. However the cardinality value can be amended using the Clusterware commands.

What is Oracle VM?

Oracle VM⁷ is a platform that provides a fully equipped environment with all the latest benefits of virtualization technology. Oracle VM enables you to deploy operating systems and application software within a supported virtualization environment. Oracle VM insulates users and administrators from the underlying virtualization technology and allows daily operations to be conducted using goal-oriented GUI interfaces. The components of Oracle VM are:

- **Oracle VM Manager:** Oracle VM Manager controls the virtualization environment, creating and monitoring Oracle VM servers and the virtual machines. Oracle VM Manager serves as the only administrative interface to the Oracle VM servers. Oracle VM Manager 3 is an Oracle Fusion Middleware application, based on the Oracle Weblogic Server application server. Oracle VM Manager uses MySQL Database Enterprise Edition as the management repository. The Oracle VM Manager runs on 64-bit Oracle Linux.

⁷ [Oracle VM homepage](#)

- 
- **Oracle VM Server:** A managed virtualization environment providing a lightweight, secure, server platform which runs virtual machines, also known as domains. At least one Oracle VM Server is required, but several are needed to take advantage of clustering. Oracle VM Server is installed on a bare metal computer, and contains the Oracle VM Agent to manage communication with Oracle VM Manager. On x86-based systems, Oracle VM Server is based upon an updated version of the underlying Xen hypervisor technology and the Linux kernel is run as dom0 to manage one or more domU virtual machines, each of which could be Oracle Linux, Oracle Solaris, Microsoft Windows, or other Linux distributions.
 - **Client Applications:** Various user interfaces to Oracle VM Manager are provided, either via the graphical user interface (GUI) accessible using a web-browser; the command line interface (CLI) accessible using an SSH client; or external applications, such as Oracle Enterprise Manager, custom built applications or scripts that use the Web Services API (WS-API). All communications with Oracle VM Manager are secured using either a key or certificate based technology.

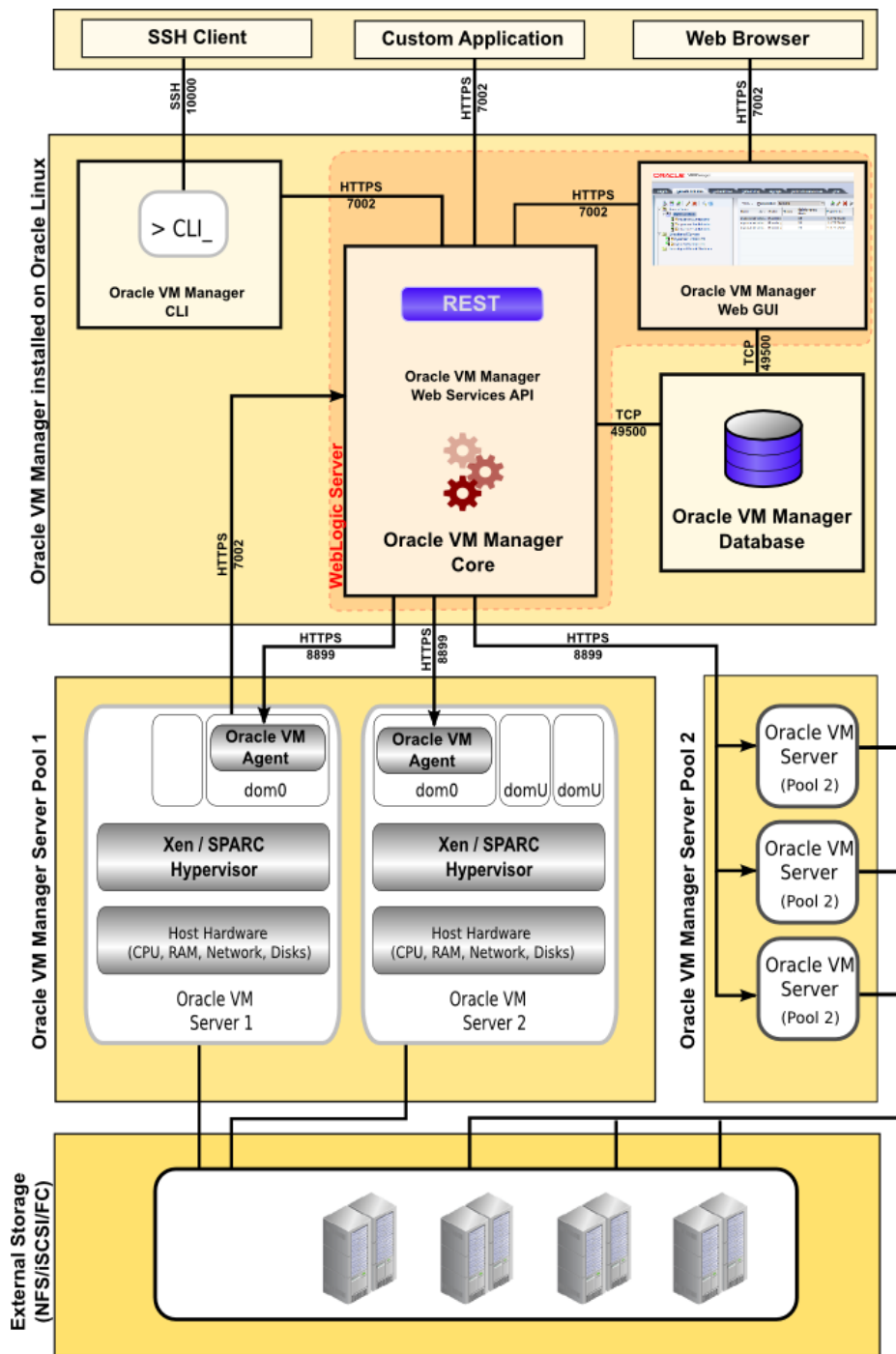


Figure 5: Oracle VM overview

Why run Oracle RAC on Oracle VM?

There are several reasons why customers may want to run Oracle RAC in an Oracle VM environment. Some of the more common ones are:

- **Server Consolidation:** Oracle RAC, RAC One Node or single instance databases with underutilized CPU resources or variable CPU utilization can often benefit from consolidation with other workloads using server virtualization. A typical use case for this scenario would be the consolidation of several Oracle databases into a single Oracle RAC database or multiple Oracle RAC databases where the hosting Oracle VM guests have pre-defined resource limits configured for each VM guest, e.g. CPU resources. Each Oracle VM hosting an Oracle RAC instance can be set up with pre-determined, fixed and limited number of resources maximizing aggregate capacity across all Oracle VMs and Oracle RAC instances. Each Oracle VM and hosted Oracle RAC instance sufficiently isolated to not interfere with other Oracle instances sharing the same physical hardware. A new option for Oracle Database 12c, Oracle Multitenant, delivers a new architecture that allows a multitenant container database to hold many pluggable databases. Oracle Multitenant fully complements other options, including Oracle Real Application Clusters and Oracle Active Data Guard.
- **Sub-capacity licensing:** The current Oracle licensing model requires the Oracle RAC database to be licensed for all CPUs on each server in the cluster. Sometimes customers wish to use only a subset of the CPUs on the server for a particular Oracle RAC database. Oracle VM can be configured in such way that it is recognized as a hard partition. Hard partitions allow customers to only license those CPUs used by the partition instead of licensing all CPUs on the physical server. For more information on using hard partitioning with Oracle VM refer to the “Hard Partitioning with Oracle VM Server for x86”⁸ white paper.
- **Create a virtual cluster:** Oracle VM enables the creation of a virtual cluster on a single physical server. This use case is particularly interesting for product demos, educational settings, and test environments. This configuration should never be used to run production Oracle RAC environments. The following are valid deployments for this use case:
 - Test / development cluster
 - Demo cluster
 - Education cluster



Note

More examples of this can be found as HOL (Hands-On Labs) usually presented at Oracle Open World conference in San Francisco at following link:
<http://www.oracle.com/technetwork/server-storage/vm/downloads/hol-oraclevm-2368799.html>

- **Rapid Provisioning:** The provisioning time of a new application consists of the server (physical or virtual) deployment time, and the software install and configuration time. Oracle VM can help reduce the deployment time for both of these components. Oracle VM supports the ability to create deployment templates. These templates can then be used to rapidly provision new (Oracle RAC) systems. Oracle

⁸ [Hard Partitioning with Oracle VM Server for x86](#)

provides a rich portfolio of pre-configured, certified gold image, deployment templates⁹ for Oracle VM 3.2 and above for Oracle Real Application Cluster, applications and operating systems.

Business Continuity, High Availability and Scalability

Business continuity is a key element in today's business and while Oracle RAC is still the ultimate solution when it comes to Oracle Databases, there are more options to choose from when considering virtualization technologies.

Two different kinds of High Availability (HA) mechanisms can be distinguished in an Oracle virtual environment: Oracle VM HA (external HA) and Oracle Clusterware based, internal HA (used for Oracle RAC) as illustrated in [Figure 6](#).

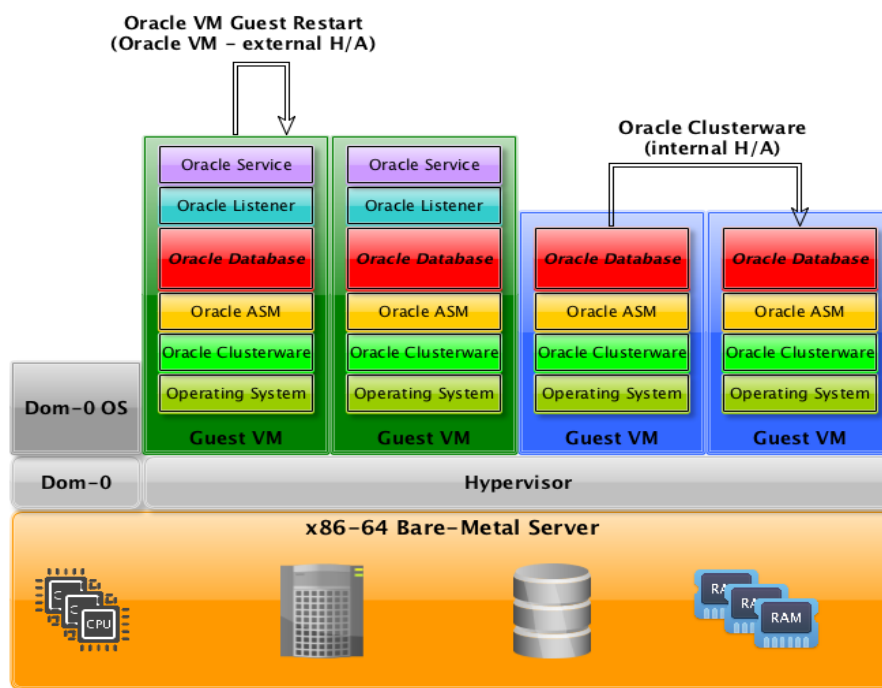


Figure 6: Oracle VM HA and Oracle RAC


In general, Oracle VM HA enables the restart of an Oracle VM guest on either the same physical machine or a different machine, if more than one physical machine is available in the server pool. However, in any case, Oracle VM would operate on the VM guest as a whole, restarting whatever is running within it.

The reason is that the Oracle VM guest is usually unaware of the applications running in the virtualized environment as much as the application is typically unaware of the Oracle VM guest it is running in.

Oracle Clusterware provides faster, fine grained recovery from a process/application failure within a guest.

Utilizing application specific agents, these solutions will perform corrective actions particular to the failure without the overhead of restarting the whole Oracle VM guest

⁹ [Oracle VM Templates for Oracle Database - Single Instance, Oracle Restart and Oracle RAC \(KM 1185244.1\)](#)



While a combination of both an internal application HA and the external VM HA serves different HA functions, it needs to be noted that the Oracle VM HA and Oracle VM policy-based resource management, Distributed Resource Scheduling, (DRS) must not interfere with inherent Oracle RAC instance placement rules. Specifically, for mission-critical, production deployments it is unsupported to co-locate Oracle VM guests hosting instances of the same Oracle RAC database on a single OVS physical server as a result of Oracle VM guest failover or automated DRS placement policy. Any Oracle VM guest failover or DRS placement policies must respect this fundamental Oracle RAC instance placement rule. For this reason, it's always suggested to configure Oracle VM "Anti-Affinity Groups": Anti-Affinity Groups specify that certain virtual machines should never run on the same Oracle VM Server. An anti-affinity group applies to all the Oracle VM Servers in a server pool.

Possible HA combinations include:

- Oracle Real Application Clusters High Availability and Scalability
 - Without additional Oracle VM guest HA
 - In conjunction with Oracle VM guest HA, where Oracle RAC instance placement policies are respected.
- Oracle Clusterware (failover cluster) provided High Availability
 - Without additional Oracle VM guest HA
 - In conjunction with Oracle VM guest HA, where Oracle RAC instance placement policies are respected.
- Standalone Oracle VM guest HA (single-instance database or Oracle RAC One Node database)

Production environments have to be based on solutions listed above.

For Oracle Clusterware based solutions regardless of Oracle VM, the current support status can be found in [Oracle Metalink Note 790189.1 - Oracle Clusterware and Application Failover Management](#). More information can be found on <http://www.oracle.com/goto/clusterware>.

Live Migration, Flex Cluster and Online Database Relocation Support

Additional solutions that improve business continuity and availability for Oracle RAC in an Oracle Virtual Machine environment include Oracle VM *Live Migration* and Oracle RAC One Node *Online Database Relocation*.

Oracle VM *Live Migration* moves a virtual machine from one physical node to another, within the same pool of servers. Oracle RAC One Node *Online Database Relocation* moves an Oracle database instance from one server to another within the same cluster. On Oracle VM environments, these servers are virtual machines, which host an Oracle Clusterware based cluster.

As of now, Oracle supports mixed cluster environments, in which some cluster members are based on virtual machines while others are based on physical servers, for migration purposes (going from bare metal to Oracle VM or vice versa). Oracle RAC One Node Online Database Relocation can therefore be used to migrate between physical and virtual environments based on Oracle VM. Mixed cluster environments are also supported with Oracle Flex Cluster where "Leaf Nodes" are VMs running on Oracle VM and "Hub Nodes" are based on bare-metal systems.

However, Oracle VM *Live Migration* offers a controlled mechanism for manually moving Oracle VM guests between physical nodes in the server pool. Since Oracle VM *Live Migration* allows for the relocation of an Oracle VM guest while respecting the placement and CPU allocation rules for Oracle RAC, it can be applied to Oracle VM guests that host an active Oracle RAC instance or instances.

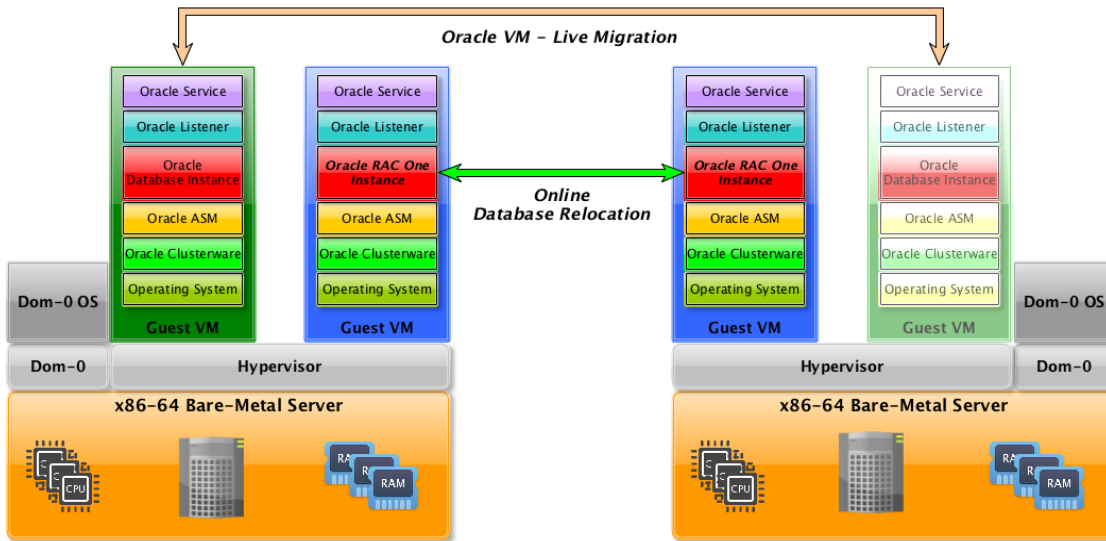


Figure 7: Live Migration and Oracle RAC One Node Database Relocation Illustrated

Extended Distance Oracle RAC on Oracle VM

An Oracle VM server pool consists of one or more Oracle VM Servers, and represents a logical grouping of the servers where a particular set of virtual machines can run. It is a requirement of all server pools that the Oracle VM Servers within them share the same CPU architecture.

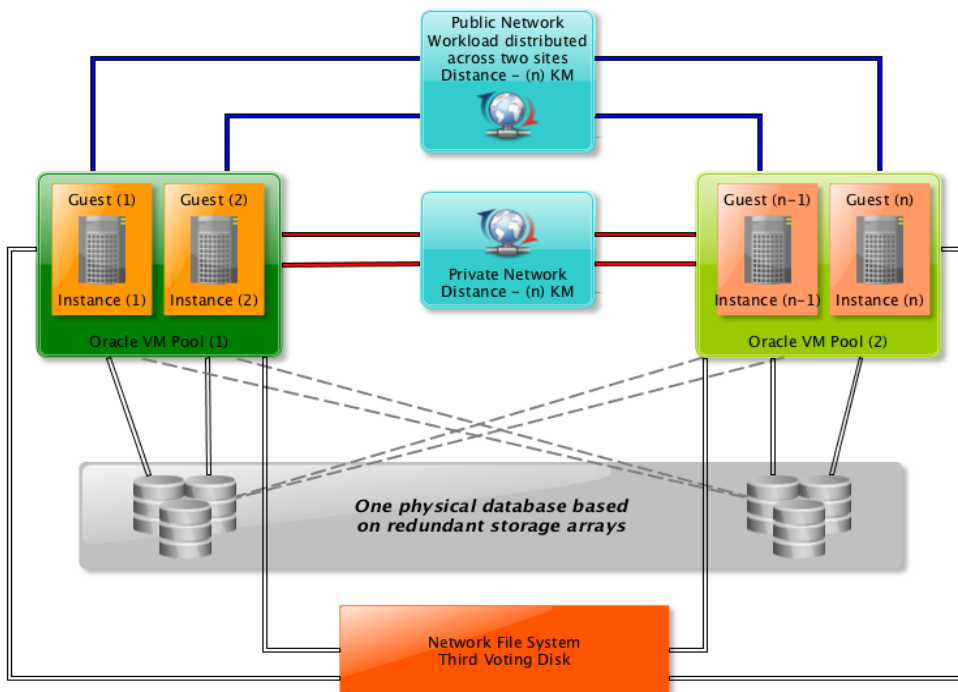


Figure 8: Oracle Extended RAC on Oracle VM - architecture overview

All servers within a server pool must be in the same physical location. Stretching Oracle VM server pools across geographical locations is not supported.

While Oracle VM Server Pool cannot be stretched on two (or more) different data centers, Oracle RAC can be deployed on Extended Distance on top of two or more Oracle VM Pools, one for each site (see [Figure 8](#) above).

On this architecture, Oracle VM Manager installation/configuration can have different options:

- Unique Oracle VM Manager on one data center

Oracle VM Manager will be installed and configured on one of the data centers available and will manage all the Oracle VM Pools deployed locally and remotely. On the remote site, one other Oracle VM Manager installation, with same UUID¹⁰, is needed: in case of disaster the “Passive” Oracle VM Manager will get in charge the local Oracle VM Pool, as well as, the remote Oracle VM Pool when available by re-discovering them and refreshing all components like Servers and Repositories.

Oracle VM Manager can be also installed as a clustered application using Oracle Clusterware.¹¹

¹⁰ [Oracle VM Manager recovery installation](#)

¹¹ [Oracle VM 3: using Oracle Clusterware to protect Oracle VM Manager](#)

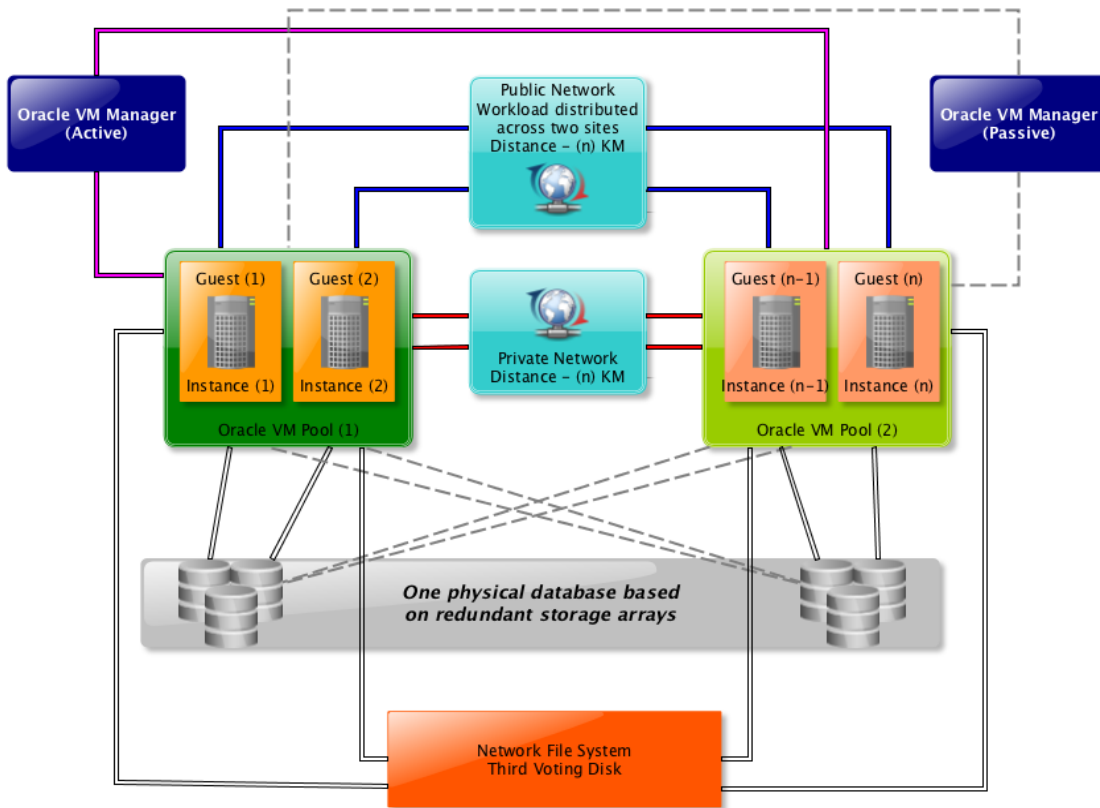


Figure 9: Oracle Extended RAC on Oracle VM – Single Oracle VM Manager interface

- One Oracle VM Manager for each data center (valid also for Private Cloud Appliance¹² Engineered System)

Oracle VM Manager will be installed and configured on each data center; each Oracle VM Manager installations will have one unique UUID and configuration. On top of Oracle VM Manager installations, Oracle Enterprise Cloud Control 13c Release 2 will act as one unique point of deployment and management for the entire architecture.

¹² [More information on Oracle Private Cloud Appliance](#)

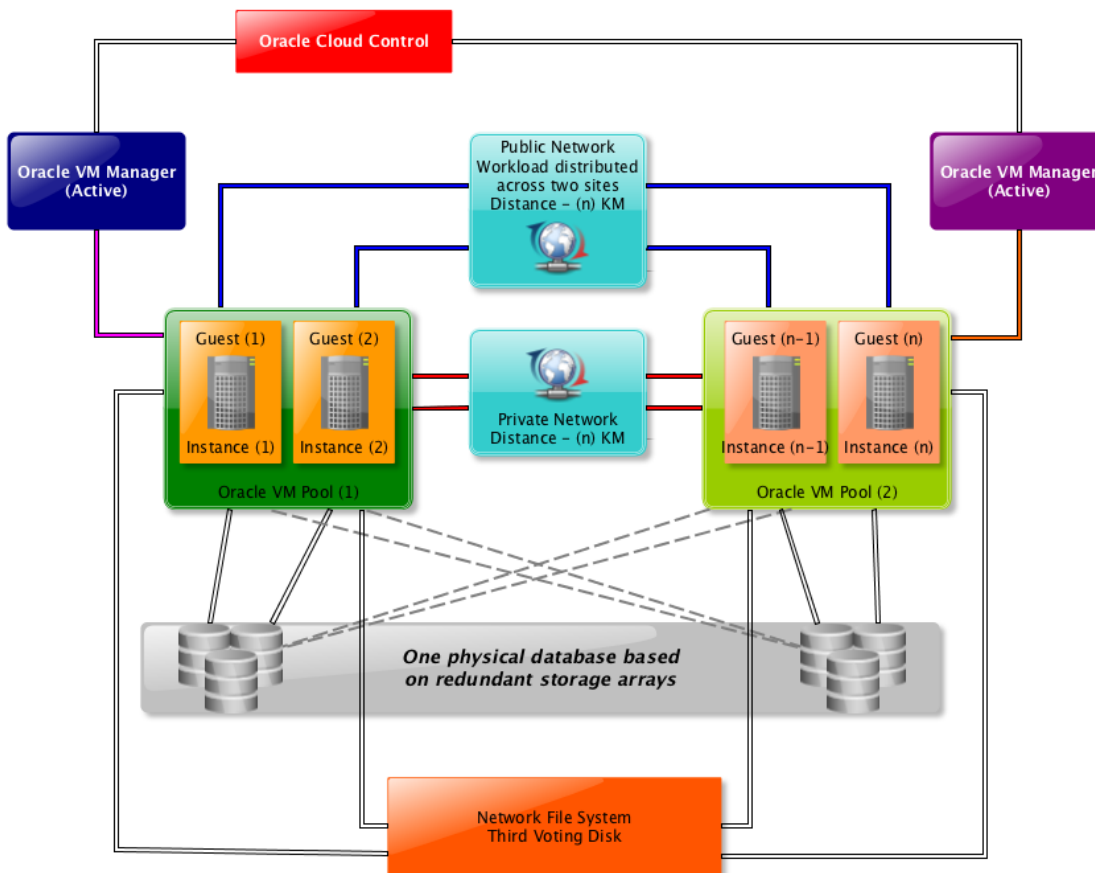


Figure 10: Oracle Extended RAC on Oracle VM – Oracle Cloud Control


On architecture above further options related to Disaster & Recovery / Business Continuity architecture for Oracle Cloud Control are available¹³.

When to use *Live Migration* or Oracle RAC One Node *Database Relocation*?

Oracle VM Live Migration and Oracle RAC One Node, together, offer complimentary solutions which support continuous, non-interrupted HA.

Live Migration of Oracle VMs offers operating system (OS) level migration to facilitate workload management, fault management, and system maintenance operations, for example, Live Migration permits server level maintenance while not losing access to the Oracle RAC database VM. Oracle Live Migration basically consists of iteratively pre-copying the contents of a source VM's memory from one physical host to another target VM on a remote host with no interruption to service.

¹³ [More information on Enterprise Manager Cloud Control 13.2 Disaster & Recovery](#)



Oracle RAC One Node and *Online Database Relocation* offer instance level migrations between properly configured VMs to facilitate maintenance and workload management within VMs. Using Oracle RAC One Node *Online Database Relocation*, an Oracle RAC One Node instance can be relocated to another server, if the current server is resource constrained or requires maintenance operations such as OS patches for example.

Online Database Relocation will also migrate Oracle RAC One Node databases to target VMs to facilitate database maintenance on the source VM without interruption of service to other database instances that may be running in the source or target VM. A typical example for such an operation would be a rolling patch upgrade for a database.

Deployment Methods

A good example for the universal use and applicability of virtualized solutions, based on Oracle VM Server, are Oracle Engineered Systems like [Virtualized Exadata Database Machine](#) or [Oracle Database Appliance](#).



Note

Oracle VM is the ideal server infrastructure for Oracle RAC used in virtual development, test, demo, or education clusters, based on only one Oracle VM host.

Easy to re-install by re-deploying a formerly saved base image or template, once the current environment is 'worn out' by various test runs.

Oracle VM provides those benefits as part of the virtual cluster creation and the rapid provisioning approach, in conjunction with Oracle RAC, as described in the previous section.

Figure 11 shows a typical Oracle RAC deployment for a development or test environment based on Oracle VM. It should be noted that **for virtual test or development clusters** Oracle allows having the 2 Oracle VM Guest Domains, each hosting 1 Oracle RAC database instance, running on only 1 Oracle VM host.

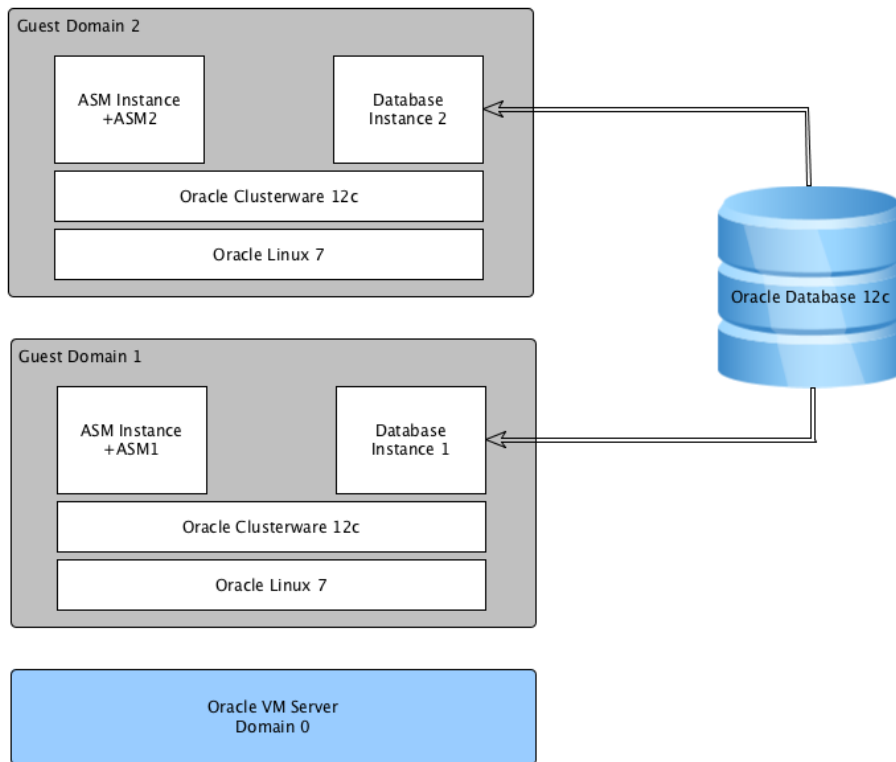


Figure 11: Oracle RAC on Oracle VM: Development Deployment example

Based on certification tests, Oracle has **certified Oracle RAC on Oracle VM for production environments from Oracle VM 3.2.x through Oracle VM 3.4.x**. This is the only certified x86 server virtualization technology for Oracle Real Application Clusters on Linux and thereby enables these technologies to be used beyond the scope of pure development or test environments.



Note

Oracle VM together with Oracle RAC enable the use of fully virtualized environments as the basis for server consolidation and highly available applications. In order to avoid a single point of failure, a minimum of 2 Oracle VM Servers should be used

Unlike test and development environments, typical production environments use more than one Oracle VM Server for each Oracle VM Guest Domain (minimum 2), each, again, hosting an Oracle RAC database instance, as illustrated in [Figure 12](#). These configurations are most likely used in server consolidation environments as well.

Having more than one Oracle VM Server as the underlying hardware platform to host the Oracle VM Guest domains with the Oracle RAC database instances eliminates the host hardware from becoming the single point of failure.

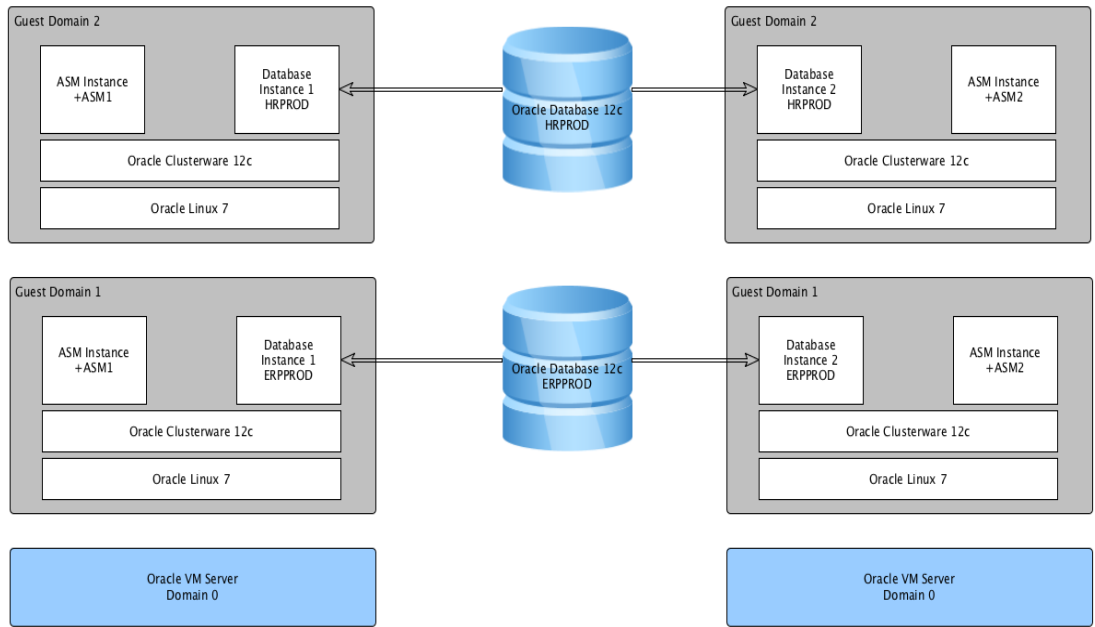


Figure 12: Oracle RAC on Oracle VM: Production Deployment example

Hardware and Software Requirements

Oracle has certified Oracle RAC on Oracle VM environments based on the following hardware and software requirements. Configurations that do not fulfill these requirements are currently not supported.

Note: As a minimum, the hardware and software requirements listed in the Oracle RAC documentation must be met to deploy Oracle RAC on Oracle VM in production environments.¹⁴

Note: Oracle is constantly in the process of testing additional and advanced Oracle VM features with Oracle RAC. This paper will be updated regularly as new test results are available.

Hardware requirements for Production Environments

- Minimum of 2 Oracle VM Servers are strongly recommended
 - Each host must provide a minimum of two physical Ethernet NICs for Oracle VM “Management Network”, two physical Ethernet NICs for Oracle RAC public communication and two physical Ethernet NICs for Oracle RAC private (interconnect) communication. A minimum of six physical Ethernet NICs is required, or a minimum of eight physical Ethernet NICs if network based storage connectivity is used.
 - Using two physical Ethernet 10GbE NICs (instead of 6/8 suggested above) for all roles (“Management Network – OVM”, “Public Network – RAC”, “Private Network – RAC”) is a supported option only when an external solution is able to create vNICs (minimum 3 for each NIC) with dedicated bandwidth and QoS enabled.
 - Each pair of Ethernet NICs will be configured in H/A by bond driver on Oracle VM Server (dom0) unless for Oracle RAC Private Interconnect that can rely on Oracle RAC Redundant Interconnect Usage or HAIP¹⁵.
 - 10GbE network interfaces¹⁶ are fully supported and it is strongly recommended that these NICs support IEEE Data Center Bridging if server VLAN tagging is to be configured.
 - For the Oracle Clusterware interconnect traffic or ASM storage traffic 1GbE is required and 10GbE is suggested.
 - For the Oracle Clusterware interconnect traffic it is suggested to implement Jumbo Frames (MTU=9000) on NICs that will be dedicated to Oracle RAC private communication; the same change will also have to be applied and verified within Virtual Machine configuration (/etc/sysconfig/network-scripts).
- Storage access redundancy requirements will require a minimum of two HBAs/SCSI controllers or the same level of redundancy for iSCSI/NFS storage based solutions.

¹⁴ [Support questions: How to use Oracle RAC in a Cloud Solution](#)

¹⁵ [Oracle® Clusterware Administration and Deployment Guide 12c Release 1 \(12.1\)](#)

¹⁶ [Oracle VM 3: 10 GbE Network Performance Tuning](#)

- Storage access timeout needs to react to given limits when Oracle VM clustered Pool is in place; recommended timeout values must come from the storage vendors and a possible path failover has to react within the given limits (OCFS2 timeout/Voting Disk timeout)¹⁷
- Using single-server Oracle VM Pool relaxes the requirement to have two physical Ethernet NICs dedicated to Oracle VM “Management Network”; Oracle VM “Management Network” and Oracle RAC “public communication” can rely on the same two physical Ethernet NICs.

Hardware requirements for Development Environments

- For Development or non-production environments on a single Oracle VM Server, external storage is recommended, but not required.
- Virtual network connectivity can be established within the Oracle VM Server, if no external client connects are required.

Oracle recommends using identical hardware configurations for development and pre-production Q&A environments accurately modeling production environments.

Oracle RAC Software

- Oracle RAC and Oracle Clusterware **11.2.0.3**, **11.2.0.4**, **12.1.0.2** in 64-bit software versions are currently supported. Oracle recommends to keep these releases always updated with latest CPU (Critical Patch Update¹⁸) and PSU (Patch Set Update¹⁹).

General Oracle VM

Oracle VM release: 3.2.x through Oracle VM release 3.4.x

Older releases are not supported with Oracle RAC or Oracle Clusterware for new implementations.

When using Oracle RAC:

- Dynamic vCPU and memory changes are not supported.
- Pause/Unpause Oracle RAC virtual machine is not supported.
- Live-migration of an Oracle RAC VM is not supported with:
 - Oracle VM releases earlier than 3.2.10
 - Oracle RAC One Node is supported with Oracle Database 11.2.0.4 and later

¹⁷ [Oracle VM: clustering for x86 Server Pools](#)

¹⁸ [Critical Patch Updates, Security Alerts and Third Party Bulletin](#)

¹⁹ [Oracle Recommended Patches - Oracle Database \(Doc ID 756671.1\)](#)

Oracle VM Guest Configuration

- The guest must be an Hardware Virtualized guest with PV drivers (PVHVM) or a Para-Virtualized (PVM) Guest
- Oracle Linux 6 and Oracle Linux 7 (only HVM with PV drivers)
- 64-bit Linux is currently supported for the Oracle VM Guest for Oracle Database
- Oracle VM Server release 3.2 and above requires 64-bit x86 hardware

Oracle VM VCPU configuration for Oracle RAC

Testing has shown that a small amount of incidental over-committing of physical CPUs will not greatly diminish the overall stability of the system or the cluster stack where the competing workloads for oversubscribed resources are well-understood and concurrent demand does not exceed physical capacity. It is still recommended, however, not **to over-commit CPUs**. Multi-core CPUs will treat every core as a virtual CPU (vCPU) and multi-core CPUs with multi-thread support **will treat each thread** as a vCPU. **vCPU allocation** in Oracle RAC / VM environments **must therefore adhere to the following rules**:

- **VCPU** = vCPU calculations are based on physical cores and **not** threads. For CPUs that are not multi-thread “1 vCPU=1 core”.
- If the hyper-threading is enabled on the bare-metal system “1 vCPU = 2 threads”.
- The total amount of vCPUs allocated to guest domains, should not exceed two times (2x) the amount of real cores in the Oracle VM server.
- The amount of vCPUs allocated to a single guest domain should not exceed the amount of real cores / threads in the Oracle VM server.
- CPU pinning is recommended for hard partitioning but also for NUMA awareness. For further details see [Optimizing Oracle VM Server for x86 Performance](#).
- Oracle VM Server dom0 default vCPU allocation should not be changed

VCPU allocation examples

On all following examples you can also consider defining CPU capacity and CPU priority using Oracle VM Manager web interface.

CPU Priority: The higher the priority value, the more physical CPU cycles are given to the virtual machine.

Processor Cap %: Increase or decrease the percentage to which the virtual CPUs can receive scheduled time. This parameter defines the maximum percentage to which the virtual CPUs can receive scheduled time.

Valid configuration examples

A server with 2 Socket(s), 4 core(s) each (8 cores total), 16 threads with HT enabled.

Virtual Machine	vCPUs allocated
Guest1	6 vCPUs (6 threads, 3 cores)
Guest2	6 vCPUs (6 threads, 3 cores)
Guest3	4 vCPUs (6 threads, 2 cores)

**Note**

This is a valid configuration because the total amount of vCPUs allocated to guest domains does not exceed two times (2x) the amount of real cores in the Oracle VM Server. Total Guest vCPU = 16 (equal two times the number of cores available)

A server with 4 Sockets, 10 cores each (40 cores total), 80 threads with HT enabled.

Virtual Machine	vCPUs allocated
Guest1	20 vCPUs (20 threads, 10 cores)
Guest2	20 vCPUs (20 threads, 10 cores)
Guest3	20 vCPUs (20 threads, 10 cores)

**Note**

This is a valid configuration because the total amount of vCPUs allocated to guest domains does not exceed two times (2x) the amount of real cores in the Oracle VM Server. Total Guest vCPU = 60 (less than two times the number of cores available)

Invalid configuration examples

A server with 2 Sockets, 4 cores each (8 cores total), 16 threads with HT enabled.

Virtual Machine	vCPUs allocated
Guest1	12 vCPUs (12 threads, 6 cores)
Guest2	12 vCPUs (12 threads, 6 cores)

**Stop!**

This is an invalid configuration because the total amount of vCPUs allocated to guest domains exceeds two times (2x) the amount of real cores in the Oracle VM Server. Total Guest vCPU = 24 (more than two times the number of cores available)

A server with 4 Sockets, 10 cores each (40 cores total), 80 threads with HT enabled.

Virtual Machine	vCPUs allocated
Guest1	40 vCPUs (20 threads, 10 cores)
Guest2	40 vCPUs (20 threads, 10 cores)
Guest3	20 vCPUs (20 threads, 10 cores)

**Stop!**

This is an invalid configuration because the total amount of vCPUs allocated to guest domains does not exceed two times (2x) the amount of real cores in the Oracle VM Server. Total Guest vCPU = 100 (more than two times the number of cores available)

Oracle VM supported storage configurations for Oracle RAC

Oracle VM itself allows configuring storage in many different ways. Oracle RAC and Oracle Clusterware are not supported on all possible storage configurations that Oracle VM offers.

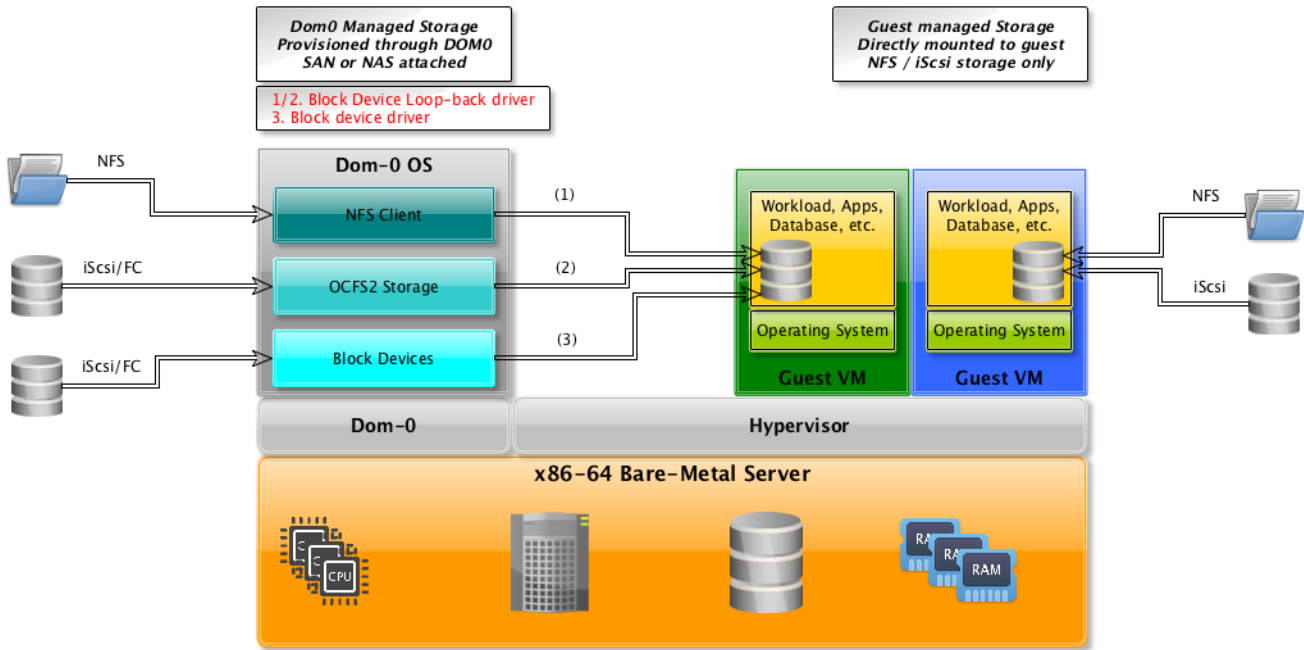


Figure 13: Storage configuration options on Oracle VM

Generally, storage configurations can be divided into two groups:

Dom-0 managed storage

In this method of storage configuration, the storage is made available in dom0 and then made available in the guest domains using the guest's configuration file. The storage available in dom0 can be configured as an Oracle VM repository or presented, by an Oracle VM Manager interface, to the guest.

Guest managed storage

In this method of storage configuration, the storage is made available directly in the guest domain. It is not visible in dom-0; options for this method are: iSCSI and NFS.

Storage Configuration Method	Supported with Oracle Real Application Cluster
Dom0 Managed Storage	
SAN / block device (physical disk mapped to the guest)	Yes (Recommended)
iSCSI / block device (physical disk mapped to the guest)	Yes (Recommended)

SAN / OCFS2 (virtual disk on repository)	Yes (see Note below)
iSCSI / OCFS2 (virtual disk on repository)	Yes (see Note below)
NFS (virtual disk on repository)	Yes (see Note below)
Guest Managed Storage	
iSCSI	Yes (Recommended)
NFS	Yes (Recommended)

Table 1: Storage options supported with Oracle Real Application Cluster



Note

Oracle Supports Oracle RAC for Production environments, built on Oracle VM, with both physical and virtual disks for OCR, voting and database files; virtual disks may be used for Oracle RAC Production environment on Oracle VM when I/O performance is not a critical requirement²⁰; otherwise, physical disks are recommended for Oracle RAC deployment on Oracle VM. Each Oracle VM Server (physical-server) cannot guest more than one Oracle Database 12c Hub-Node or one Oracle Database 11gR2 Node, part of the same Oracle Real Application Cluster. Building Oracle RAC with more virtual machines on a single Oracle VM Server (physical server) is allowed only for demo or educational purposes.

Examples of demo environment are available, as Hands-On Lab at:

<http://www.oracle.com/technetwork/server-storage/vm/downloads/hol-oraclevm-2368799.html>

²⁰ [How to Troubleshoot I/O Performance under Oracle VM \(dom0 / domU \) \(Doc ID 2212880.1\)](#)

Oracle RAC on Oracle VM – Best Practices

Based on tests in the course of the Oracle RAC for Oracle VM certification, Oracle has developed some best practices recommendations in order to run Oracle RAC in an optimized virtual environment. Below you will find some of the recommendations as a result of those tests.

Oracle Installation requirements and recommendations

Linux Kernel parameters for Oracle VM guest environment

Set the kernel parameter `vm.min_free_kbytes=51200` (not required for UEK kernel) which will reserve 512MB to allow the OS to reclaim memory faster and avoid LowMem pressure. See following "My Oracle Support" documents for additional information:

- [Linux Kernel Low-Mem Pressure Issues and Related Kernel Structures \(Doc ID 452326.1\)](#)
- [Linux: Out-of-Memory \(OOM\) Killer \(Doc ID 452000.1\)](#)
- [Top 5 Issues That Cause Node Reboots or Evictions or Unexpected Recycle of CRS \(Doc ID 1367153.1\)](#)

On kernel revision 2.6.18 and below set the kernel parameter `vm.swappiness=100`. Stress testing has shown that `vm.swappiness = 100` (default = 60) on kernel version 2.6.18 or lower can reduce or delay node evictions under conditions of heavy memory pressure due to large numbers of client connections or during login storms.

On Oracle Linux 6.x, you may find that `limits.conf` appears to be ignored when establishing `nproc` - see [Document 1487773.1](#) for additional details.

Time synchronization

By default, the Oracle VM Guest time is automatically synchronized with Domain-0. This time synchronization can allow some time drift in the guest domains running Oracle RAC. Therefore, the following recommendations should be implemented when running Oracle RAC on Oracle VM:

- Starting from 11gR2, Cluster Time Synchronization Daemon (CTSSD) can be used in place of NTPD. CTSSD will synchronize time with a reference node in the cluster when an NTPD is not found to be configured. Should you require synchronization from an external time source you must use NTPD which will cause CTSSD to run in "observer" mode. However, if NTP is running, then it must to be configured with the slewing option as documented in [Document 551704.1](#).
- Use an external configured NTPD as server for your VMs

Storage Configuration

Oracle VM Guest limits

Oracle RAC deployments on Oracle VM should also consider maximum-limits²¹ imposed by the solution itself; so, while evaluating Oracle RAC installation on Oracle VM, it becomes fundamental to evaluate them.

Best solution possible should be to consider database size, hypothetical growth of the same and, based on this, define a good balance between the number of disks and their size.

If the storage used is guest-managed (NFS shares or iSCSI devices directly presented to the VM), configuration limits for Oracle VM, mostly related to the number of disk devices, can be safely ignored; best practices valid for bare-metal system are still valid for this type of configuration on top of Oracle VM.

Multipath and Device Persistency in Domain-0

For Oracle RAC on Oracle VM environments, multipath access to the (SAN) storage is highly recommended. Multipath is, as a default, already configured in Domain-0 and not in the guest domains.

Device persistency is granted by Oracle VM Server Domain-0 itself relying on storage WWID of devices presented. There is no need to set up device persistency in the Guest VMs. The mapping of Oracle VM disks to guest disks can be dynamically managed using Oracle VM Manager / CLI interface (see Appendix A for an example).

Starting with Oracle Database / Clusterware 11.2, Oracle ASM is Oracle's recommended storage management solution that provides an alternative to conventional volume managers, file systems, and raw devices.

The Oracle ASM volume manager functionality provides flexible server-based mirroring options. The Oracle ASM normal and high redundancy disk groups enable two-way and three-way mirroring respectively. You can use external redundancy to enable a Redundant Array of Independent Disks (RAID) storage subsystem to perform the mirroring protection function.

Starting with Oracle Database / Clusterware 12.1, Oracle ASM Filter Driver (Oracle ASMFD) is available as a storage option for guest VM.

Oracle ASMFD is a kernel module that resides in the I/O path of the Oracle ASM disks. Oracle ASM uses the filter driver to validate write I/O requests to Oracle ASM disks.

Oracle ASMFD simplifies the configuration and management of disk devices by eliminating the need to rebind disk devices used with Oracle ASM each time the system is restarted.

Oracle ASM Filter Driver rejects any I/O requests that are invalid. This action eliminates accidental overwrites of Oracle ASM disks that would cause corruption in the disks and files within the disk group. For example, the Oracle ASM Filter Driver filters out all non-Oracle I/Os which could cause accidental overwrites (for example `dd if=/dev/zero of=/dev/oracledisk` will simply fail when ASMFD is used).

If you have an existing Oracle ASM library driver (Oracle ASMLIB) configuration, then depending on whether you want to use Oracle ASMLIB or Oracle ASMFD, make the necessary changes based on the Oracle Database 12c documentation²².

²¹ [Oracle VM 3: Configuration Limits for Release 3.4](#)

²² [Oracle Database 12.1: Oracle ASMLIB and Oracle ASMFD](#)

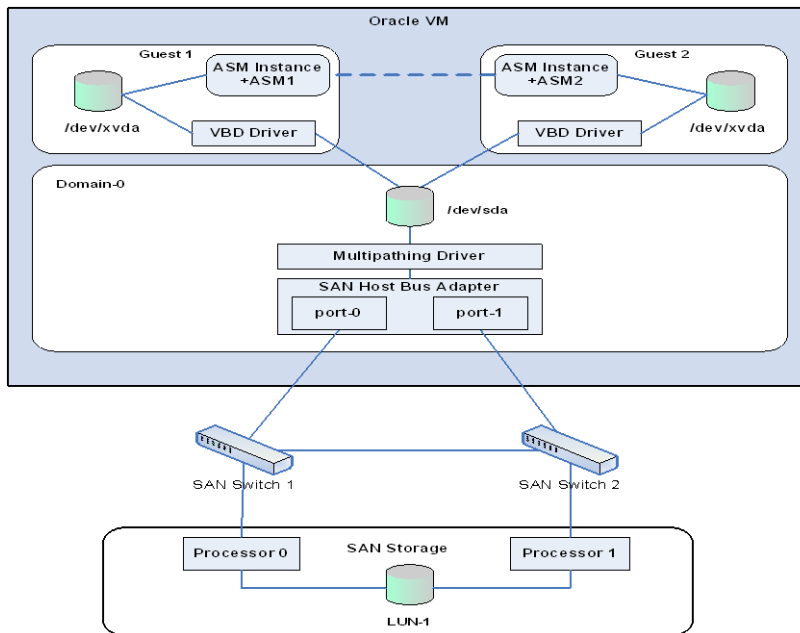


Figure 14: SAN Storage on Oracle VM

Cluster file system requirements

Some applications require running the Oracle RAC database on a cluster file system or relying on some files that need to be available on every node. For these applications the Oracle ASM Cluster File System (ACFS – 11g Rel. 2 or later) should be used in the guest domains.



Note

For updated and further best practices follow-up on “My Oracle Support” KM document: [Oracle RAC and Oracle Clusterware Best Practices and Starter Kit \(Linux\) \(Doc ID 811306.1\)](#)

Network Configuration

Required Networks

The Oracle VM Server should have a sufficient number of NICs to provide at least two distinct bridges and virtual NICs to the guest domains. For production systems the number of NICs per Oracle VM Server is specified in the “Hardware and Software Requirements” section of this paper. The actual number of NICs depends on the overall configuration and level of redundancy desired.

As for any other Oracle RAC configuration, a separation of networks for:

- Oracle VM management network
- Public network



Note

On Virtualized Environment, like on Oracle VM Server, network components may be virtualized; as a result, sometimes network failures won't be reflected to guest VM; so even if the HW network card fails, the "Public Network" virtual NIC, available within the guest, remains up. Starting with release 12.1.0.2 "Ping Targets"²³ enhancement allows overcoming this problem.

- Private network (cluster interconnect)
- Storage network (if applicable)

is required for production environments. Therefore, the Oracle VM Server must ensure that the respective communication would not interfere with each other.

These networks, however, can be shared with multiple Oracle RAC VM Guests in an Oracle VM Server. Consolidation of network traffic is only supported for 'like' traffic, e.g. private interconnect traffic may be consolidated with private interconnect traffic from other Oracle VM guests, public traffic may be consolidated with public traffic from other Oracle VM guests, where both traffic types are arbitrated through their respective virtual bridge. Oracle RAC public and private traffic must be physically and logically separate. When consolidating networks, latency, bandwidth and capacity must be carefully measured and monitored (see "Sizing the Oracle VM environment for Oracle RAC" on this document for recommendations).

The Oracle VM management network must be on its own physical network segment or on its own bandwidth dedicated vLAN with QoS enabled (see "Hardware and Software Requirements").

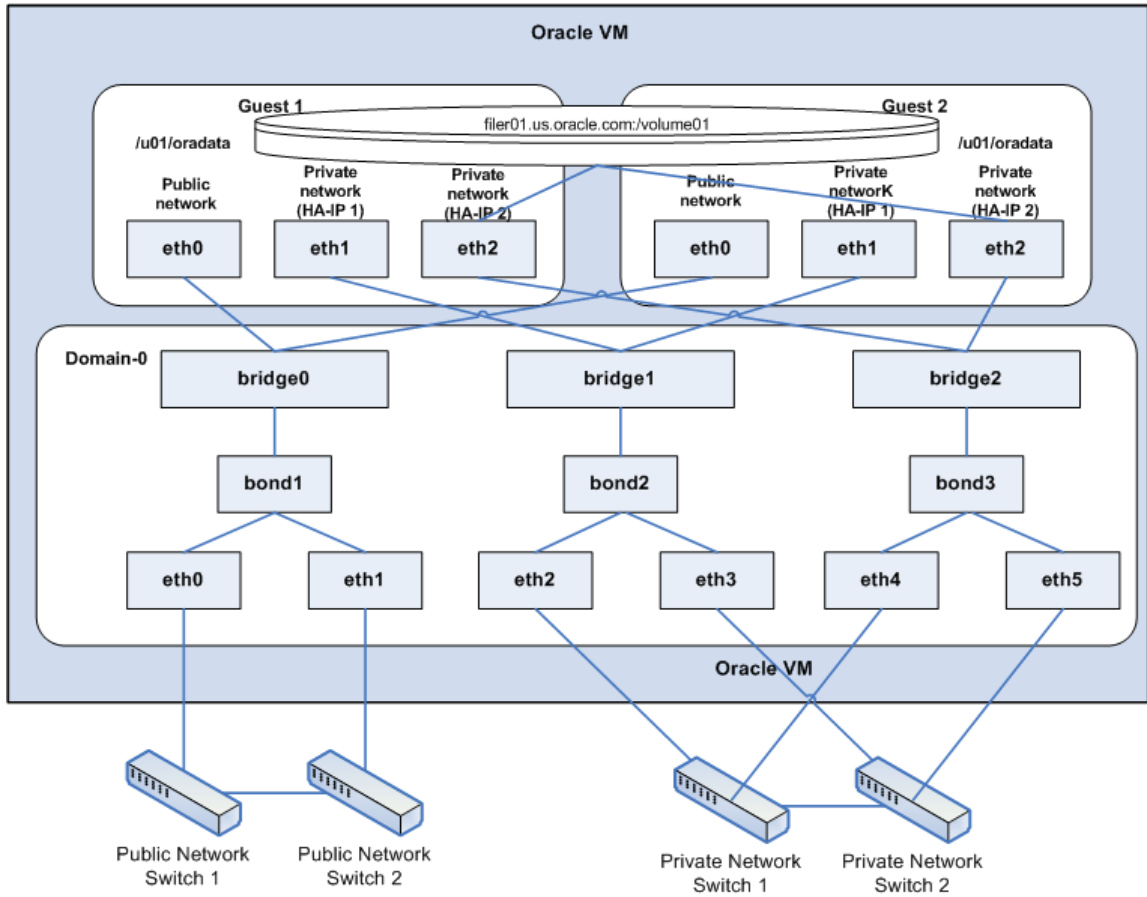
Network Bonding

For production environments it is strongly recommended to have two Network Interface Controllers (NICs) for each network and use Linux bonding configured in Dom0 to make them highly available. This results in an Oracle RAC network HA requirement of a minimum of four NICs (for public and private network) per Oracle VM Server running Oracle RAC, as illustrated in [Figure 15](#). Appendix B explains setting up bonding devices step-by-step.

For the Oracle Clusterware private interconnect, in place of Linux bonding in Dom0, you may present two to four vNICs for the private interconnect to each of the configured Oracle VM guests. Each vNIC would map to a distinct physical network interface, either on a single subnet or multiple subnets for Oracle Clusterware interconnect high availability. This feature, Redundant Interconnect Usage²⁴ or HAIP is failover model with multiple communication endpoints, managed by the Oracle Clusterware in each of the Oracle VM guests and uses automatically assigned, non-routed link local addresses for the Oracle Clusterware private interconnect communication. An acceptable configuration would be a combination of Linux bonding of the clusterware interconnects in Dom0 with HAIP configured in the Oracle VM guests.

²³ [12.1.0.2 Ping Target for Oracle Clusterware VIP \(Doc ID 1958241.1\)](#)

²⁴ [Oracle® Clusterware Administration and Deployment Guide 12c Release 1 \(12.1\)](#)



NB: bond0 is dedicated, for default, to Oracle VM Management Network

Figure 15: Full network bonding on Oracle VM

General HugePages and Transparent HugePages

If you are running Oracle Linux 6/7, RedHat Enterprise 6/7 or SLES 11/12 or UEK2 kernels and above, be sure to disable Transparent HugePages to prevent performance problems and Node/Instance evictions. Our long-standing recommendation of using general HugePages still stands (Transparent Huge Pages is a different implementation than general HugePages). See following documents for additional details:

- [Oracle VM 3.4: How is the HugePages Feature Enabled for Virtual Machines?](#)
- [Disable Transparent HugePages on SLES11, RHEL6/7, OL6/7, and UEK2 and above \(Doc ID 1557478.1\)](#)

Orachk utility and best practices

Oracle provide the ORAchk tool to audit the configuration of RAC, CRS, ASM, GI etc. It supports database versions from 10.2-12.1, making it a useful starting point for most analysis. Further Information are available at:

- [ORAchk - Health Checks for the Oracle Stack \(Doc ID 1268927.2\)](#)
- [ORAchk and EXAchk User's Guide](#)

Sizing the Oracle VM Environment for Oracle RAC

As a general rule of thumb, you should “size your Oracle VM environment in a similar way as you would size a non-VM environment.” Following this rule when sizing an Oracle VM system Oracle RAC should cover most of the optimization. Particular attention should be paid to the following parameters:

CPUs and cores

Follow the requirements as stated in this document for assigning vCPUs to dom-0 and guest domains. It is supported, but not recommended, to over-commit by a factor of 2.

If multiple databases have been configured into a single Oracle RAC virtual machine, please evaluate Oracle Database Resource Manager configuration and “Instance Caging” best practices²⁵.

Memory

Memory is not shared between domains on Oracle VM. When sizing memory for your Oracle RAC VM system, make sure your system memory is large enough to accommodate the memory needed for Domain-0, the Linux Operating systems, Oracle Clusterware, and the Oracle RAC databases in each of the guest domains.

Evaluation of the amount of memory needed for Domain-0 can be easily obtained with following formula:

$$\text{dom0_mem} = 502 + \text{int}(\text{physical_mem} * 0.0205)$$

Further details are available on Oracle VM documentation²⁶.

Network

For an Oracle RAC on Oracle VM environments the general recommendation is to install a minimum amount of 3 NICs per Oracle VM Server machine. For a production environment the recommendation is to have at least 6 network cards installed in the Oracle VM Server and use bonding; for demo purposes the environment can be built also on a single physical NIC.

Keep in mind that in a shared network bridge, bandwidth is shared between the different guest domains installed on one Oracle VM server. In production environments, these network interfaces must be at a minimum 1 Gbit Ethernet, for example 1000BASE, however, 10GbE, 10GBASE is fully supported and recommended. As noted earlier, similar network traffic may be consolidated but the physical and logical segregation between public/private/management and storage networks is required. That would mean dedicated 10GbE cards for each traffic type where consolidated traffic on this NIC may be VLAN tagged for logical separation. This configuration should be performed in Dom0 through the OVM Manager or Oracle VM CLI interface.

When configuring a larger number of guest VMs on one host, or when deploying guest VMs that require a high amount of network bandwidth, conscientious monitoring for oversubscription events resulting in packet loss, either on the NIC or the switch, and/or increased latencies and bandwidth degradation must be implemented.

²⁵ [Oracle Database 12cR1: Managing Multiple Database Instances on a Single Server](#)

²⁶ [Oracle VM Official Documentation – Dom0 Memory configuration](#)

Storage

When using SAN storage for your Oracle RAC on Oracle VM deployment, more than one HBA can be used to obtain a higher bandwidth to the storage in the Oracle VM host. The multipath driver usually combines the bandwidth of each path automatically.

It is recommended to use a dedicated network interface when using NAS or iSCSI storage.

The dedicated storage network should be configured in the same way as the public or interconnect network. Again, multiple NICs can be used to eliminate a single NIC as a Single-Point-Of-Failure and appropriate bonding modes can be used to increase bandwidth²⁷

Recommended Installation: Oracle VM Templates

Using the Oracle VM Templates for Oracle Database provide an easy and fast way of deploying any one of three configurations: Oracle single-instance, Oracle Grid Infrastructure for Standalone Server (formerly known as Oracle Restart) and Oracle RAC cluster on Oracle VM **for test** as well as **production environments**.

The Oracle VM Templates for Oracle Database are deployed using an automated tool called Deploycluster. The Deploycluster tool automatically connects to the Oracle VM Manager, starts the VMs and initiates the automatic deployment which then runs to completion on the desired configuration, be it single instance or 20 node Oracle RAC Cluster..Any number of nodes or instances can be added or removed from the cluster using a single command.

The entire installation is automated, requesting minimal information (e.g. node names, IP addresses, etc.) from the user. The final n-node cluster (or Single Instance Database) configuration adheres to the best practices for production systems as described in this paper.

Oracle VM Templates for Oracle Database templates uses updated Oracle Linux system image for 64-bit, which can also be updated in real-time at deployment time (see Yum@Deploy support). The OS image contains a minimal installation of Oracle Linux, only basic RPM packages are installed with the use of the preinstall rpms described earlier in this paper.

Configuration and benefits at a glance:

- Support for three modes of operations: Single Instance, Grid Infrastructure for Standalone Server (Oracle Restart) and Oracle RAC
- Templates available for both Enterprise Edition as well as Standard Edition 2
- Templates released as fully operational “kit” with OS as well as Oracle disk, however one can “mix and match” thus creating any possible combination for deployment.
- Full support for majority of Database and Oracle RAC features like Oracle Flex Cluster and Flex ASM, pluggable databases, container databases, Admin Managed or Policy Managed databases, automated ACFS Filesystem creation including Database in ACFS; support for dedicated ASM network interface and much more.
- Allows OS real-time update at deployment, see Yum@Deploy support in Template’s documentation.
- The environment comes loaded with Swingbench, OS Watcher, and more tools.

²⁷ [How is Network Bonding used on Oracle VM Server](#)

- The exact same Template or VM created from the template can be re-purposed to become Single Instance or a member of an Oracle RAC Cluster via a simple cleanup and rebuilding of that VM.

To download and use the templates and the Deploycluster tool to configure an Oracle Database on Oracle VM in test or production systems see the following OTN page:

<http://www.oracle.com/technetwork/server-storage/vm/database-templates-12c-11gr2-1972804.html>

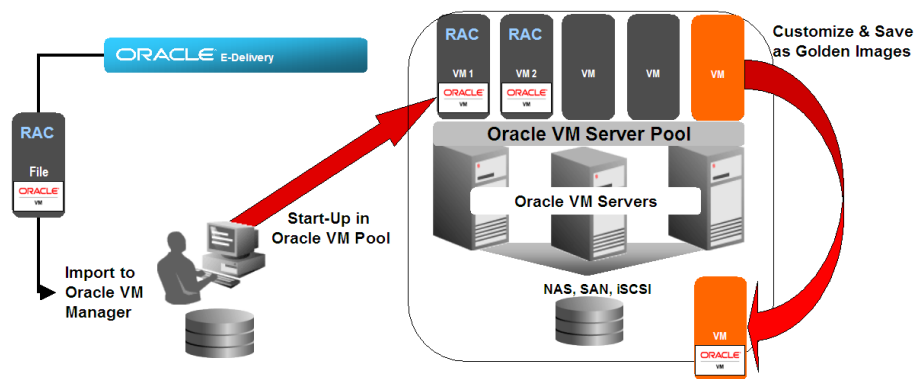


Figure 16: Oracle Template Deployment

Oracle RAC on Oracle VM Environments: manual installation

The following steps describe the complete installation of Oracle RAC in an Oracle VM environment:

- Plan your installation carefully. Consider the requirements for your new environment and plan accordingly using the guidelines and best practices in this white paper and the official Oracle documentation.

Install Oracle VM Server binaries on the machine(s) on which you want to perform the Oracle RAC installation.

Follow the Oracle VM Server Installation Guide²⁸ for the details on how to do this.

- From the OVM Manager GUI, discover the servers and configure storage, networking and all components related to the Oracle VM Server.
- Create the “hardware-virtualized with PV drivers” – HVM with PV Drivers - (Oracle Linux 6 and 7) Guests domains for the Oracle RAC installation using a virtual-disks for both the Operating-System and the Oracle Database/Clusterware software.
- Install and configure the guest operating system (Oracle Linux 6/7). Make all necessary configuration changes to the guest operating system as described in the requirements for installing Oracle Clusterware and Oracle RAC; for Oracle Linux you also have the option to rely on “preinstall” rpms, dedicated to automate the OS setup for Oracle Database / Oracle Clusterware; see further details:

Linux OS Installation with Reduced Set of Packages for Running Oracle Database Server without ULN/RHN (Doc ID 579101.1)

²⁸ [Oracle VM Installation Guide](#)

- Install Oracle Clusterware and Oracle RAC 11g Rel. 2 / 12c Rel. 1 as for non-virtualized environments, following the installation guides for these products.

Live Migration and Online Database Relocation Best Practices

Oracle VM Live Migration Best Practices

Live Migration of VMs offers operating system (OS) level migration to facilitate workload management, fault management and system maintenance operations. Oracle Live Migration basically consists of iteratively pre-copying the contents of a source VM's memory from one physical host to another target VM on a remote host with minimal interruption to running services.

This iterative pre-copy of the memory continues until it is determined that the dirty page rate on the source machine is consistently higher than the memory transfer rate to the target machine. At this point, the VM is stopped and all remaining dirty pages are copied from the source VM to the target VM.

The stop-and-copy operation to complete the Live Migration is a suspend time blackout of the VM. For most applications, the dirty page rate will be low and the suspend time will measure in milliseconds to seconds. But for large highly active applications with potential dirty page sets of greater than 2.7GB, the suspend time of the VM could be considerably longer, potentially violating clusterware heartbeat thresholds which would trigger node fencing at the Oracle Clusterware level.

To avoid these potential failures during a Live Migration of VMs the following best practices should be observed:

- Prior to initiating a Live Migration in an Oracle RAC production environment redefine the Oracle Clusterware misscount from the default of 30 seconds to 60. Issue the following command as root:

```
`crsctl set css misscount 60`
```

This will allow a Live Migration for an application with a potential dirty page set of greater than 2.7GB. Assuming a conservative maximum throughput for a single 1GbE shared link provides 90MB/sec., 60 seconds misscount would allow a suspend period of 5.4GB (90MB/sec * MC) where MC= misscount set to 60 seconds).

- When the Live Migration is complete, set the Oracle Clusterware misscount back to the default. Issue the following command as root:

```
`crsctl unset css misscount`
```



Note

There is a known issue that Live Migration may fail when ACFS is used, see MOS Note: 2202282.1

Only one Live Migration should be executed at a time. Multiple, simultaneous Live Migrations within the clustered environment may violate the suspend time thresholds described above.

A dedicated, high bandwidth network has to be used for Live Migrations and the VM guests should be on the same network and IP subnet. The default network used by Live Migration is the OVM management network that may be the public network. If Live Migration is used as a frequent maintenance operation, it is recommended to deploy bonded interfaces used to separate the Dom0/ovs-agent/OCFS2/Live Migration traffic from other application traffic.

While Secure Live Migration is enabled by default, it should be considered that a secure connection to the remote guest (using `-ssl`) adds overhead to the Live Migration operation. It is therefore recommended to avoid such secure connections, if permitted. In general, a secure connection can be avoided, if the Live Migration network is inherently secure.



Note

The pause, suspend and resume OVM management functionality for Guest domains must be avoided when the Guest Domain contains an actively running production Oracle RAC instance. Any pause or suspend operation could trigger Oracle Clusterware to fence the suspended Oracle VM virtual machine or, in worst cases, lead to a block corruption.

Oracle RAC One Node Online Database Relocation Best Practices

The best practices for Oracle RAC One Node Online Database Relocation remain the same when performing an Online Database Relocation operation within an Oracle VM based environment as opposed to executing this procedure on physical machines. More information on the Oracle RAC One Node Online Database Relocation can therefore be found in the Oracle RAC One Node User's Guide²⁹. When performing an Online Database Relocation between two Oracle VM guest, consider the following:

- The target VM for an Online Database Relocation event must have adequate OS/system resources to allow for an Oracle RAC One Node instance to run.
- Only one Online Database Relocation event at a time should be executed.
- The target VM must be a configured member of the cluster.

²⁹ [Oracle RAC / RAC One Node 12c Rel. 1 \(12.1.0.2\) installation guide](#)

Summary

Having certified Oracle RAC on Oracle VM in release 3.2 or higher, and following the recommendations made in this paper; it is possible to implement, using these standardized technologies in (Oracle) Enterprise Grid Infrastructures configurations like the one illustrated in fig. 13 or even more sophisticated configurations in the future.

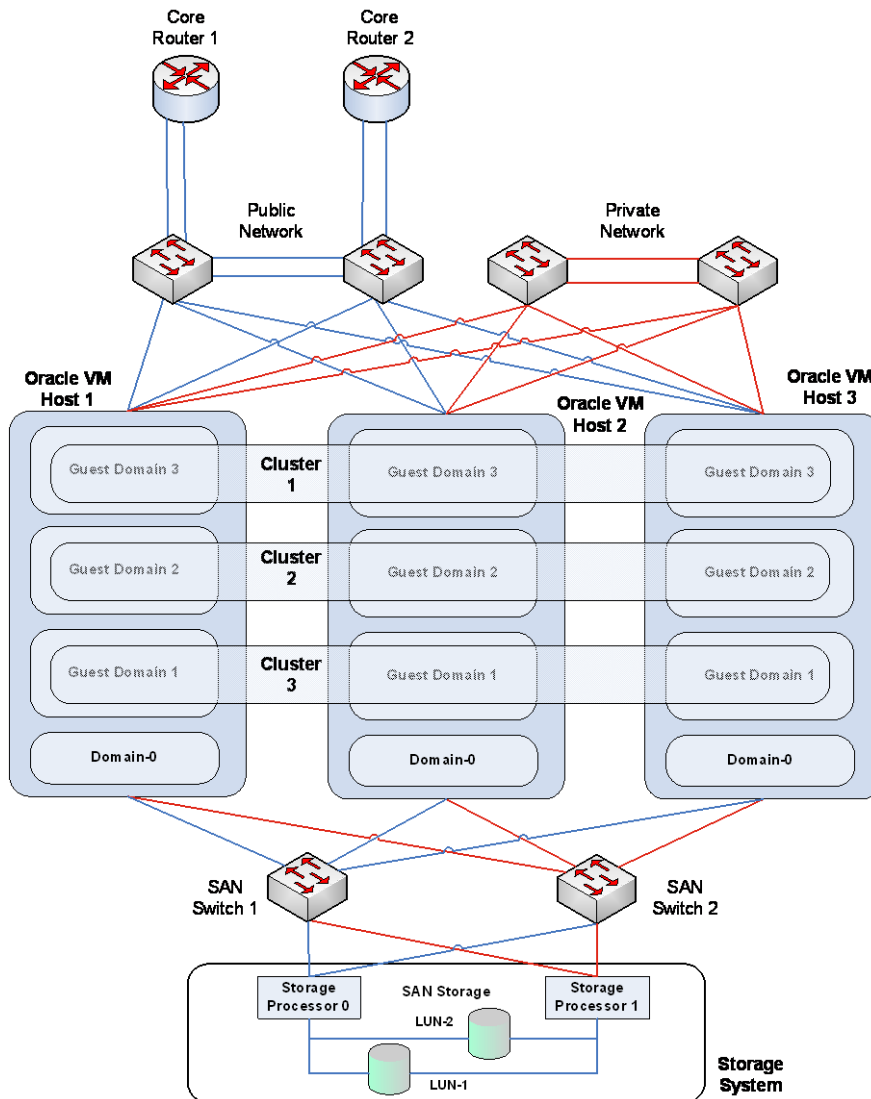


Figure 17: Production Architecture Example

Appendix A – Disk persistence with Oracle VM Manager

Oracle VM Manager allows editing VMs and associate the correct virtual or physical-disks and obtains configuration similar to this one:

Name	Status	Tag(s)	Event Severity	Server	Max. Memory (MB)	Memory (MB)	Max. Processors	Processors	Keymap	Operating System
racnode0.2	Stopped		Informational	ovs10471.ooowl...	3072	3072	1	1	en-us	Oracle Linux 6
racnode0.1	Stopped		Informational	ovs10471.ooowl...	3072	3072	1	1	en-us	Oracle Linux 6

Slot	Disk Type	Name	Size (GiB)	Repository	Location
0	Virtual Disk	64bc0463f63424099f9f9e48a9d9393.img (2)	12.0	repo01	/devmapper/LATA_VBOX_HARDDISK_VB432dccc7f-896e621f
1	Virtual Disk	4d78d4f4550a47b28b399cd8915f1274.img (2)	45.0	repo01	/devmapper/LATA_VBOX_HARDDISK_VB432dccc7f-896e621f
12	Virtual Disk	racasm1	5.0	repo01	/devmapper/LATA_VBOX_HARDDISK_VB432dccc7f-896e621f






Note

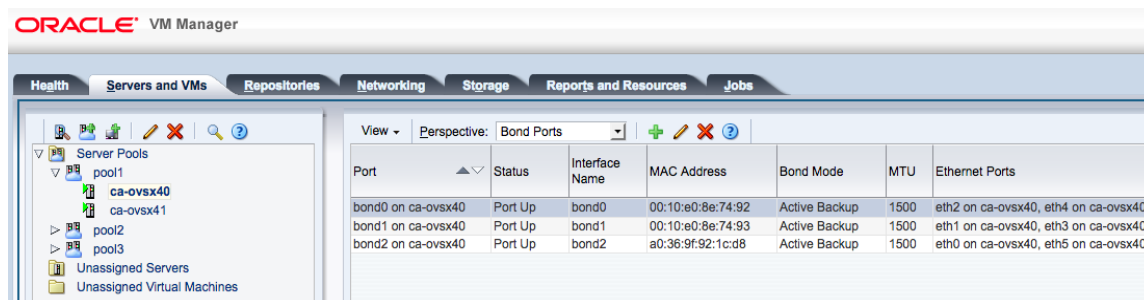
One virtual or physical disk must be edited and configured as “shared” to be presented to more VMs. For Production/Quality environments physical-disks for database files are recommended.

Appendix B – Bonding Setup Example

Bonding configuration needs to be applied on all Oracle VM Server (dom0) part of the Oracle VM Pool where Oracle RAC virtual machines will be created. All network configuration related to the Oracle VM Server (dom0) must be executed using Oracle VM Manager WEB or CLI interface.

Using buttons like  or  or  it's possible to create / edit and delete Bond interfaces on each single Oracle VM Server. If the same operation will be executed on many Oracle VM Server, it is possible to automate this configuration using [Oracle VM Manager CLI interface](#).

Each bond interface (one for Oracle VM Management, one for Oracle RAC public network and one for Oracle RAC private network) needs, at least, to rely on two different NICs, like in the following example:



The screenshot shows the Oracle VM Manager interface with the 'Networking' tab selected. The 'Bond Ports' perspective is active, displaying a table with the following data:

Port	Status	Interface Name	MAC Address	Bond Mode	MTU	Ethernet Ports
bond0 on ca-ovsx40	Port Up	bond0	00:10:e0:8e:74:92	Active Backup	1500	eth2 on ca-ovsx40, eth4 on ca-ovsx40
bond1 on ca-ovsx40	Port Up	bond1	00:10:e0:8e:74:93	Active Backup	1500	eth1 on ca-ovsx40, eth3 on ca-ovsx40
bond2 on ca-ovsx40	Port Up	bond2	a0:36:9f:92:1c:d8	Active Backup	1500	eth0 on ca-ovsx40, eth5 on ca-ovsx40

- Bond0 = eth2 and eth4 (dedicated to Oracle VM Management)
- Bond1 = eth1 and eth3 (dedicated to Oracle RAC Public-Network – VMs)
- Bond 2 = eth0 and eth5 (dedicated to Oracle RAC Private-Network – VMs)

The bonding mode used in the example above is mode 1, active-backup. This is the easiest mode and will work on most switches without any problems. Any other modes might require specific switches and a certain switch configuration.

Once all bond configurations are defined, you have to correctly associate these interfaces to a Network Role, using Oracle VM Manager; under the “Networking” tab of Oracle VM Manager you can create two (or more) different Network Roles that will be dedicated to Virtual Machines:

ORACLE VM Manager

Health Servers and VMs Repositories **Networking** Storage Tools and Resources Jobs

Networks VLAN Groups Virtual NICs

View +

Name	Intra-Network Server	Network Channels					VLAN Segment	Description
		Server Management	Cluster Heartbeat	Live Migrate	Storage	Virtual Machine		
192.168.140.0		√		√				
192.168.40.0			√		√			
COAXM_eth0_public						√	network.VlanSegment (1153) in vm_public_vlan_vgrp	
COAXM_eth1_storage						√	network.VlanSegment (1155) in vm_public_vlan_vgrp	
OVMPM_DR_chgnetid_net1						√	network.VlanSegment (921) in vm_public_vlan_vgrp	
OVMPM_DR_chgnetid_net2						√	network.VlanSegment (922) in vm_public_vlan_vgrp	
OVMPM_DR_chgnetid_net3						√	network.VlanSegment (923) in vm_public_vlan_vgrp	This network has a description
OVMPM_DR_chgnetid_net4						√	network.VlanSegment (924) in vm_public_vlan_vgrp	
VLAN99_PVI						√	network.VlanSegment (99) in vm_private_vgrp	PCA Internal network VLAN 99
VM_eth0_public						√	network.VlanSegment (521) in vm_public_vlan_vgrp	
VM_eth1_storage						√	network.VlanSegment (523) in vm_public_vlan_vgrp	
VM_eth2_backup						√	network.VlanSegment (912) in vm_public_vlan_vgrp	
VM_eth3_private						√	network.VlanSegment (913) in vm_public_vlan_vgrp	
lab_storage					√	√	lab_storage_vsegs_1	
mgmt_public_eth						√		
vm_private						√	vm_private_vsegs_1	
vm_public_vlan						√	vm_public_vlan_vsegs_1	

With Networks created you can proceed to create your own Virtual Machines that will be dedicated to Oracle Real Application Cluster.

While creating Virtual Machines you will have to associate one vNIC for each role needed. In the example below you can see a VM with two different vNICs, one dedicated to the public network and one dedicated to the private:

myrac13b Stopped Normal N/A 8192 8192 4 2 en-us Oracle Linux 7

Configuration Networks Disks

VNIC	Ethernet Network	IP Addresses
00:21:f6:cb:ef:0c	VM_public	
00:21:f6:cb:ef:0b	VM_private	

Starting with Oracle RAC 12c Release 1 it is possible to create two different vNICs and configure the Oracle Database with HA-IP as described in this document.

If the intention is to use HA-IP then it is not needed to create the bond interface on Oracle VM Server (dom0) dedicated to Private-Network, instead, it is needed to associate each NIC to a specific network role and then associate two different vNICs to the VMs:

- Eth0 → Network Role: "HA-IP 1of2" → vNIC (1) for Private Network
- Eth1 → Network Role: "HA-IP 2of2" → vNIC (2) for Private Network

Appendix C – References

Official Oracle Documentation - General

Product / Reference	URL
Oracle VM on oracle.com	http://www.oracle.com/virtualization
Oracle VM Official Documentation	http://www.oracle.com/technetwork/documentation/vm-096300.html
Oracle RAC on oracle.com	https://www.oracle.com/database/real-application-clusters/index.html
Oracle VM Templates for Oracle Database	http://www.oracle.com/technetwork/server-storage/vm/database-templates-12c-11gr2-1972804.html
Oracle Clusterware on oracle.com	http://www.oracle.com/technetwork/database/database-technologies/clusterware/overview/index.html
Oracle ASM	http://www.oracle.com/technetwork/articles/database/asm-odb12c-enhancements-2206102.html
Oracle ASMLib	http://www.oracle.com/technetwork/topics/linux/asmlib/index-101839.html
Oracle ASMLib vs ASMFD	https://docs.oracle.com/database/121/OSTMG/GUID-9C4245E0-279B-4832-A2FA-00E57B34D604.htm#OSTMG95908

Official Oracle Documentation – Oracle Real Application Cluster


Product / Reference	URL
Oracle RAC 11.2 – Installation Guide	https://docs.oracle.com/cd/E11882_01/install.112/e41962/toc.htm
Oracle RAC 12c Release 1 – Installation Guide	https://docs.oracle.com/database/121/nav/portal_16.htm

My Oracle Support Notes

[Information Center: Oracle Scalability Grid Infrastructure / Clusterware and Real Application Clusters \(RAC\) \(Doc ID 1452965.2\)](#)

[RAC Cluster is Experiencing Node Evictions after Kernel Upgrade to OL6.6 \(Doc ID 2011957.1\)](#)

[Requirements for Installing Oracle Database 12.1 on RHEL6 or OL6 64-bit \(x86-64\) \(Doc ID 1529864.1\)](#)



[ALERT: Disable Transparent HugePages on SLES11, RHEL6, RHEL7, OL6, OL7, and UEK2 and above \(Doc ID 1557478.1\)](#)

[Information Center: Known Issues for Oracle Cluster Verification Utility \(CLUVFY\) for Oracle Database - RAC/Scalability \(Doc ID 1605365.2\)](#)

[Top 5 Issues That Cause Node Reboots or Evictions or Unexpected Recycle of CRS \(Doc ID 1367153.1\)](#)

[Oracle Clusterware and Application Failover Management \(Doc ID 790189.1\)](#)

[12.1.0.2 Ping Target for Oracle Clusterware VIP \(Doc ID 1958241.1\)](#)

[Oracle VM Templates for Oracle Database - Single Instance, Oracle Restart \(SIHA\) and Oracle RAC \(Doc ID 1185244.1\)](#)



CONNECT WITH US



[Blogs.oracle.com/virtualization](https://blogs.oracle.com/virtualization)



Facebook.com/OracleVirtualization



Twitter.com/ORCL_Virtualize



oracle.com

Oracle Corporation, World Headquarters

500 Oracle Parkway
Redwood Shores, CA 94065, USA

Worldwide Inquiries

Phone: +1.650.506.7000
Fax: +1.650.506.7200

Hardware and Software, Engineered to Work Together

Copyright © 2017, Oracle and/or its affiliates. All rights reserved. This document is provided for information purposes only, and the contents hereof are subject to change without notice. This document is not warranted to be error-free, nor subject to any other warranties or conditions, whether expressed orally or implied in law, including implied warranties and conditions of merchantability or fitness for a particular purpose. We specifically disclaim any liability with respect to this document, and no contractual obligations are formed either directly or indirectly by this document. This document may not be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose, without our prior written permission.

Oracle and Java are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.

Intel and Intel Xeon are trademarks or registered trademarks of Intel Corporation. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. AMD, Opteron, the AMD logo, and the AMD Opteron logo are trademarks or registered trademarks of Advanced Micro Devices. UNIX is a registered trademark of The Open Group. 0117

Oracle Real Application Clusters on Oracle VM Environments
January 2017

Author: Simon Cotter, Markus Michalewicz

Contributors: Anil Nair, Saar Maoz, John McHugh, Honglin Su, Sebastian Solbach

Revision: 1.0



Oracle is committed to developing practices and products that help protect the environment