

Deploying Oracle VM 2.2 on EMC CLARiiON CX4 Series Storage

Applied Technology

Abstract

This white paper examines the deployment considerations of Oracle VM on EMC[®] CLARiiON[®] storage, as well as discusses the best practices for deploying Oracle VM.

February 2010

Copyright © 2010 EMC Corporation. All rights reserved.

EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

THE INFORMATION IN THIS PUBLICATION IS PROVIDED “AS IS.” EMC CORPORATION MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND WITH RESPECT TO THE INFORMATION IN THIS PUBLICATION, AND SPECIFICALLY DISCLAIMS IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.

For the most up-to-date listing of EMC product names, see EMC Corporation Trademarks on EMC.com

All other trademarks used herein are the property of their respective owners.

Part Number h6904.1

Table of Contents

Executive summary	4
Introduction	4
Audience	4
Terminology	4
EMC CLARiiON CX4.....	5
Oracle VM	5
Oracle VM Server	6
Oracle VM deployment configuration.....	7
Oracle VM Server on a CLARiiON CX4 architecture	8
Deploying Oracle VM with CLARiiON CX4.....	9
Oracle VM Manager installation and configuration	10
Oracle VM Server installation and configuration.....	10
Network configuration.....	11
EMC Navisphere Host Agent and CLI installation	11
Device-Mapper Multipath tool configuration for storage	12
Creating shared storage and clusters on CLARiiON CX4	13
Shared repository	13
Configuring shared storage on the multipath device.....	14
Configuring the cluster and shared repository using OCFS2 on the multipath device	14
Managing server pool(s), servers, and resources	15
Creating and configuring virtual machines.....	16
Configuring storage for the VM	17
Live domain migration and High Availability setup	17
Live migration of virtual machines	17
Enabling High Availability (HA).....	18
OVM test scenarios.....	18
Storage considerations for PVM and HVM.....	19
Conclusion	19
References	19

Executive summary

Virtualization is the ability to run multiple virtual machines on a single piece of hardware. The hardware runs software that enables you to install multiple operating systems capable of running simultaneously and independently, in their own secure environment, with minimal impact on performance. Each virtual machine has its own virtual CPU, network interfaces, storage, and operating system. Virtualization has become a key technology for enterprises to enable server consolidation, reduce physical infrastructure complexity, and help make data centers more energy-efficient. With increased server provisioning in the data center, several factors play a role in stifling growth. Increased power and cooling costs, physical space constraints, administrative resources, and interconnection complexity all contribute significantly to the costs and feasibility of continued expansion. Virtualization can result in lower total cost of ownership (TCO) while maintaining availability and increasing resource utilization.

Deploying Oracle VM on EMC® CLARiiON® CX4 storage allows you to benefit from the advanced capabilities that CX4 software provides, including Virtual Provisioning™, local and remote replication, Navisphere® Quality of Service management, and five 9s availability.

Introduction

This white paper is intended to help you understand the deployment guidelines and considerations of Oracle VM 2.2 on x86-based Linux servers with EMC CLARiiON CX4 storage. This paper covers the important deployment aspects of Oracle VM in the context of CLARiiON CX4 storage based on the engineering testing done at EMC labs. It is not intended to cover every feature of Oracle VM. For specific Oracle VM features not covered, please refer to the appropriate Oracle VM documentation (see “References”).

The rest of this white paper will provide the reader with a detailed view of deploying Oracle VM 2.2 on EMC CLARiiON CX4 SAN storage, best practices for configuring the hardware and software components, and pointers for obtaining more information.

Audience

This white paper is intended for Oracle VM administrators, storage architects, customers, and EMC field personnel who want to deploy and configure Oracle VM on EMC CLARiiON CX4 storage arrays.

Terminology

Dom0	Domain zero is the management domain.
DomU	User Domain is an unprivileged domain with no direct access to the hardware or device drivers. Each domU is started by Oracle VM Server in dom0.
Guest OS	A guest OS is a virtualized operating system running within a domain.
HVM	Hardware Virtualized Machine or fully virtualized guests: The unmodified guest operating system runs on the virtual machine. It traps and emulates every I/O and hardware instruction. To apply the fully virtualized mode, either Intel-VT or AMD-V should be available on the host. For example, it is possible to run Microsoft Windows as a guest operating system unmodified, using the fully virtualized method provided by Intel-VT technology or AMD-V technology.
Hypervisor	A small, lightweight, software virtual machine monitor, for x86-compatible computers that controls only the most basic resources of the system, including CPU and memory usage, privilege checks, and hardware interrupts.
Oracle VM Manager	Provides the user interface (a web application) to manage Oracle VM Servers and to manage a virtual machine lifecycle.
OVA	The Oracle VM Agent communicates with Oracle VM Manager to create and manage guests on an Oracle VM Server.

OVM / Oracle VM	A platform that provides a fully equipped environment for better leveraging the benefits of virtualization technology.
OVS	The Oracle VM Server is an updated version of the underlying Xen hypervisor technology and the Oracle VM Agent.
PVM	Para-Virtualized Machine or paravirtualized guests: The guest operating system is modified and recompiled before being installed on a virtual machine. Also, the virtual machine does not need to trap privileged instructions. Trapping is a method used to handle unexpected or unallowable conditions, which is time-consuming and can impact operating system performance. Without trapping privileged instructions, the para-virtualized operating system runs at near native speed.
Storage / Disk Array	A disk array is a disk storage system that contains multiple disk drives. It is differentiated from a disk enclosure, in that an array has cache memory and advanced functionality, like RAID and virtualization.
Virtualization	The ability to run multiple virtual machines on a single piece of hardware.
Virtual Machine	A virtual machine is a guest operating system and its associated application software.
Xen	Xen is a virtual machine monitor/hypervisor for x86, x86-64, Itanium, and PowerPC 970 architectures. It allows several guest operating systems to be executed on the same computer hardware concurrently. Xen was initially created by the University of Cambridge Computer Laboratory and is now developed and maintained by the Xen community as free software.
xvd	xvd is the Xen virtual block device. The xvd interface does not truly emulate IDE or SCSI, but you can think of it as similar to hd (for IDE) and sd (for SCSI).

EMC CLARiiON CX4

The EMC CLARiiON CX4 series with UltraFlex™ technology is based on a new breakthrough architecture and extensive technological innovation, providing a midrange solution that is highly scalable, meeting the price points of most midrange customers. The CX4 is the fourth-generation CX series, and continues EMC's commitment to maximizing customer's investments in CLARiiON technology by ensuring that existing resources and capital assets are optimally utilized as customers adopt new technology. The CX4 delivers immediate support for the latest generation of disk drive technologies, such as Enterprise Flash Drives (EFDs). EFDs are supported on all four models of EMC CX4 (CX4-120, CX4-240, CX4-480, and CX4-960).

The CLARiiON CX4 series introduces thin LUN technology that builds on CLARiiON virtual LUN capabilities and seamlessly integrates with CLARiiON management and replication software. With CLARiiON Virtual Provisioning, you can choose between traditional LUNs, metaLUNs, and thin LUNs. The ability to nondisruptively migrate data to different LUN and disk types allows you to deploy the best solution without incurring downtime. Virtual Provisioning enables organizations to reduce costs by increasing utilization without overprovisioning of storage capacity, simplifying storage management, and reducing application downtime.

CLARiiON also provides the Navisphere management software, which is a suite of tools that allows centralized management of CLARiiON storage systems. Navisphere provides a centralized tool to monitor and configure CLARiiON storage. The Navisphere suite includes Navisphere Manager, which has a web-based UI, Navisphere Secure CLI (Command Line Interface). CLARiiON provides functional capabilities like point-in-time local replicas and remote replication options for business continuity using the Navisphere management tool.

Oracle VM

Oracle VM enables data centers to run multiple virtual machines on a single piece of hardware. The bare-metal hardware runs software that enables you to install multiple instances of operating systems that are able to run simultaneously and independently, in their own secure environment, with minimal reduction in

performance. Each virtual machine has its own virtual CPU, network interfaces, storage, and operating system. EMC Fibre Channel SAN storage systems are ideal choices for deploying highly reliable virtualization solutions including Oracle VM.

Oracle VM has the following components:

- **Oracle VM Manager** provides the web user interface, which is a standard web application, to manage Oracle VM Servers, server pools, virtual machines, and resources.
- **Oracle VM Server** is a self-contained virtualization environment based on Xen hypervisor technology designed to provide a lightweight, secure, server-based platform to run virtual machines.
- **Oracle VM Agent** communicates with Oracle VM Manager to manage virtual machines and is installed with Oracle VM Server.

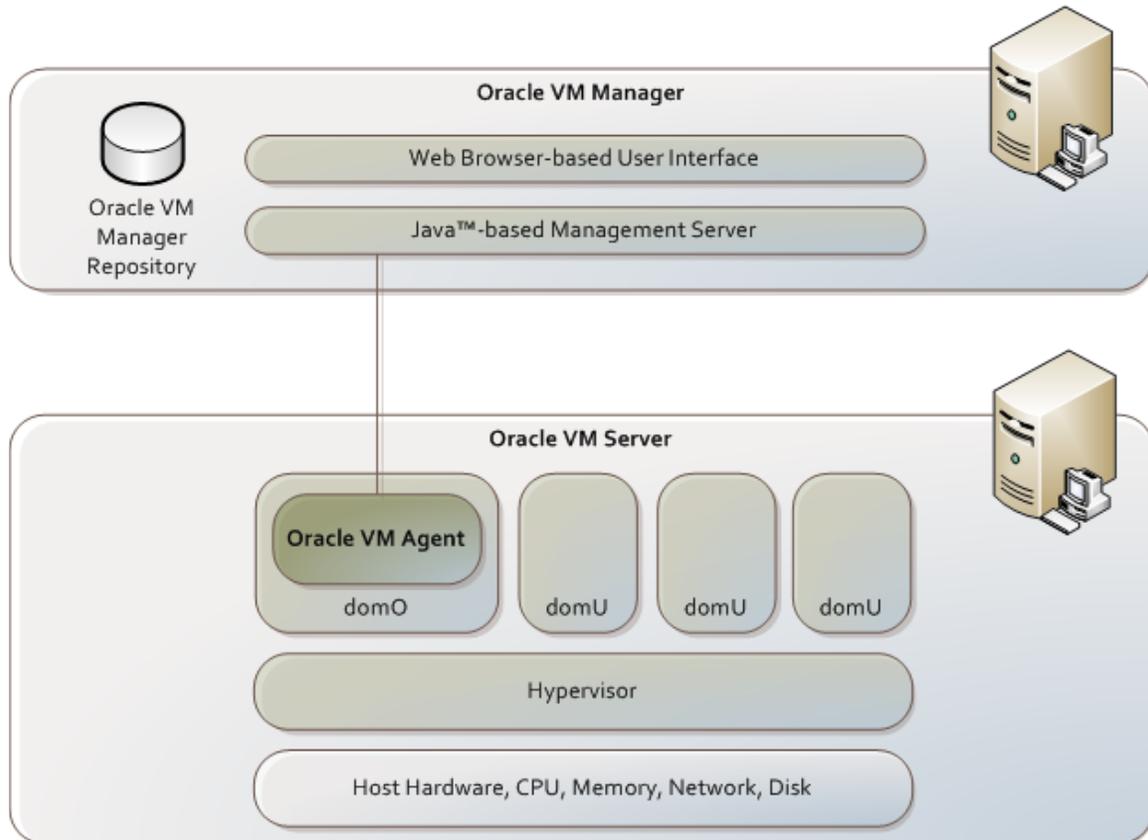


Figure 1. Oracle VM high-level architecture¹

Figure 1 shows the Oracle VM architecture.

Oracle VM Server

Oracle VM Server is a central component of Oracle VM. It includes an updated version of the underlying Xen hypervisor technology, and the Oracle VM Agent. It also includes a Linux kernel, with support for a broad array of devices, as dom0 to manage one or more domU virtual machines, each of which could be Linux or Microsoft Windows. The terms *domain*, *guest* and *virtual machine* are often used interchangeably, but they have subtle differences. A domain is a configurable set of resources, including memory, virtual CPUs, network devices and disk devices, in which virtual machines run. A domain is granted virtual resources and can be started, stopped, and restarted independently. Oracle VM Server guest operating systems may run in one of two modes, paravirtualized or hardware virtualized. In paravirtualized

¹ [Oracle VM Server User's Guide - Release 2.2](#)

mode, the kernel of the guest operating system is recompiled to be made aware of the virtual environment. This allows the paravirtualized guest to run at near native speed, since most memory, disk, and network accesses are optimized for maximum performance. If support for hardware virtualization is available (either Intel VT or AMD-V), the guest operating system may run completely unmodified. This hardware virtualized guest is carefully monitored and trapped by Oracle VM Server when any instruction is executed, which would violate the isolation with other guests or dom0. In the current implementation, there may be a performance penalty for certain types of guests and access types, but hardware virtualization also allows many Microsoft Windows operating systems and legacy operating systems to run unmodified.

Oracle VM deployment configuration²

Figure 2 shows the various deployment configurations of Oracle VM and the Oracle VM Manager.

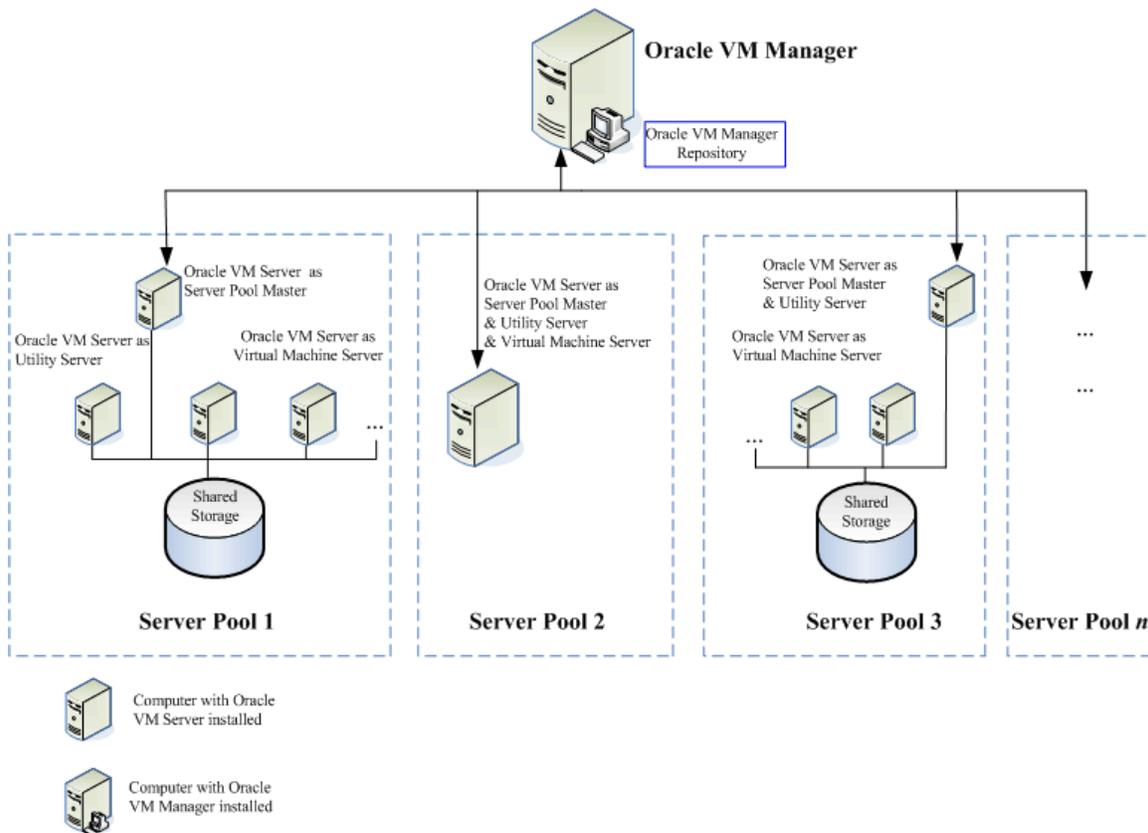


Figure 2. Deployment view of an Oracle VM infrastructure³

The following describes the components in the deployment view of an Oracle VM infrastructure:

- **Oracle VM Manager host:** The host machine on which Oracle VM Manager is installed is known as an Oracle VM Manager host. It provides the interface where most of the virtual machine management tasks are performed.
- **Servers:** An Oracle VM Server, once added to a server pool, is assigned with one, two, or even all three functions, namely, Server Pool Master function, Utility Server function, and Virtual Machine Server function.

² [Oracle VM Manager User's Guide - Release 2.2](#)

³ [Oracle VM Manager User's Guide - Release 2.2](#)

-
- **Server Pool Master:** The Server Pool Master is the core of the server pool operations. It acts as the contact point of the server pool to the outside world, and also as the dispatcher to other servers within the server pool.
 - **Utility Server:** The Utility Server is responsible for I/O-intensive operations such as copying, or moving files.
 - **Virtual Machine Server:** The primary function of the Virtual Machine Server is to run virtual machines, thus acting as a hypervisor. Oracle VM Agent is set up on the Virtual Machine Server to establish communication between the Server Pool Master, other Utility Servers, and Virtual Machine Servers.
 - **Server pool:** A server pool contains a collection of one or more Oracle VM Servers. A server pool presents a unified view of the storage in which the virtual machines reside. Each server pool must have its own shared storage.

The server functions described here can be deployed in a server pool in a variety of ways as shown in Figure 2. For example, in Server Pool 1, each one of the three functions is implemented on an individual Oracle VM Server. In Server Pool 2, all three functions are performed by a single Oracle VM Server.

In medium to large-scale environments with more than just a few virtual machines in a server pool, it is recommended that the Server Pool Master and Utility Server functions reside together or individually on a separate and dedicated physical server that does not host any guest virtual machines, as illustrated in Server Pool 3. This is to prevent any significant Server Pool Master or Utility Server usage from impacting the performance of the workloads hosted in the guest virtual machines.

- **Storage:** A storage resource is mounted to store virtual machines, external resources, and other data files that are shared among Oracle VM Servers in the server pool. In order to perform Live Migration of virtual machines between separate physical machines in the server pool, each machine involved must have shared access to storage. In our case we are using EMC CLARiiON CX4 SAN storage for this purpose.

Oracle VM Server on a CLARiiON CX4 architecture

Figure 3 shows an architectural overview of Oracle VM 2.2 using the EMC CX4 storage system that was set up in the EMC lab. The figure depicts a simple configuration and can be extended to incorporate failover scenarios with redundant hardware. This architecture is comprised of the following components:

- Two-node Dell R900 (Intel) server pools with one server acting as the Server Pool Master and Utility Server, as well as the Oracle VM Server, and the other R900 server acting as the Oracle VM Server host. Both servers are running Oracle VM 2.2.
- Oracle VM Manager running on a Dell R900 (Intel) server running OEL 4.5
- Gigabit Ethernet switch for the network
- Fibre Channel switches and an EMC CLARiiON CX4 storage system in a redundant configuration

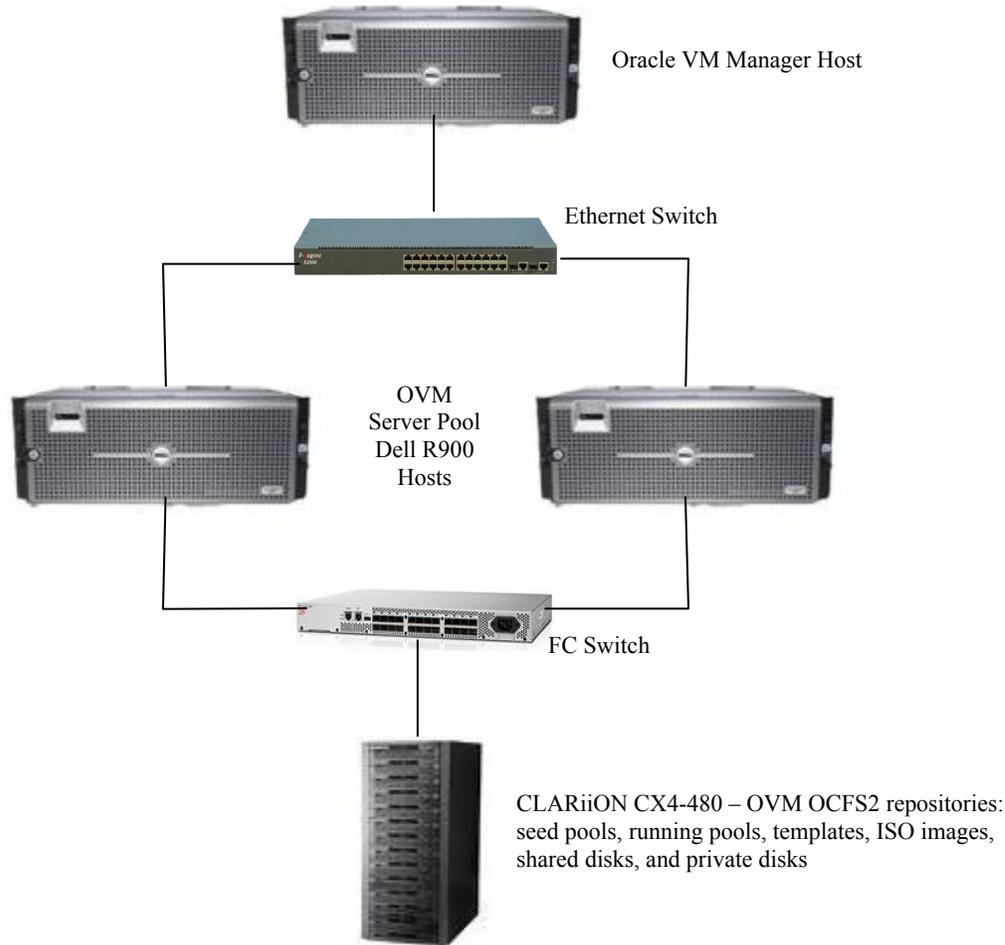


Figure 3. Architectural overview of Oracle VM Server pools attached to Dell/EMC storage

Deploying Oracle VM with CLARiiON CX4

After deciding the right deployment configuration, Oracle VM can be deployed with CLARiiON CX4 storage using the following steps:

1. [Oracle VM Manager installation and configuration](#)
2. [Oracle VM Server installation and configuration](#)
3. [EMC Navisphere Host Agent and CLI installation](#)
4. [Device-Mapper Multipath tool configuration for storage](#)
5. [Creating shared storage and clusters on CLARiiON CX4](#)
6. [Managing server pool\(s\), servers, and resources](#)
7. [Creating and configuring virtual machines](#)
8. [Live domain migration and High Availability setup](#)

Oracle VM Manager installation and configuration

To install/upgrade Oracle VM Manager, follow the detailed steps outlined in the *Oracle VM Manager Installation Guide*. Oracle VM Manager needs to be installed on a separate physical machine running Oracle Enterprise Linux 4 Update 5 or later or Red Hat Enterprise Linux Release 4 or later. After successful installation you can access Oracle VM Manager by either using `http://hostname:8888/OVS` or `https://hostname:port/OVS` URL format in the web browser.

Oracle VM Server installation and configuration

Oracle VM Server can be installed from the CD in the following ways.

- Installing onto the internal server hard drive
- Installing onto the SAN storage as a Boot from SAN (BFS)

To install the Oracle VM Server to the local hard drive, check the detailed installation steps provided in the *Oracle VM Server Installation Guide*.

Booting the servers from SAN with multiple paths can provide significant benefits for the complex, modern data center environment. To boot your Oracle VM Server from CLARiiON CX4 storage, you'll need to set up a SAN boot operating system to your storage with multipath driver support; Boot from SAN can also simplify data center administration, improve disaster tolerance, and reduce TCO through diskless servers.

On the storage side, there is no difference between a local boot and a SAN boot. On the host side, you need to change a few things to differentiate the SAN boot from the local boot:

1. Set up a host bus adapter (HBA) to enable the BIOS and select the boot device from the CLARiiON CX4 storage array.
2. Disable the local hard disk through the server BIOS, or physically remove it.
3. Set the server BIOS to keep the SAN OS volume as the first boot candidate.
4. Enable Device-Mapper MultiPath (DMMP) in the operating system.

It's always a good idea to keep only one path alive between your server and CLARiiON CX4 *before* you start to install Oracle VM Server onto the CLARiiON LUN. To do this, disable one HBA port of your server. It's better to assign one volume as the OS install volume before installation. That way, you can always find the right volume to install Oracle VM Server because there will only be one volume available at this time. The installation is much the same as the steps you follow to install Oracle VM Server to a local hard drive — except that you will need the `linux mpath` option during the initial installation, because Oracle VM Server supports enabling DMMP during installation.

To install the Oracle VM Server on to the CLARiiON SAN storage as BFS, please follow the following steps.

1. Make sure you are using the latest HBA BIOS and zoning is done correctly so that the VM Server host can see the LUNs from the CLARiiON storage array.
2. Present only the boot LUN to the host using the Navisphere management tool.
3. Disable one HBA port out of the two connected HBA ports by pulling out the HBA cable.
4. Enable HBA BIOS, and make sure that you can see your target LUN. Refer to your HBA vendor's (QLogic, Emulex) documentation for configuration of the correct settings.
5. On the server system BIOS, disable the local hard disk and make sure the boot LUN is at the top of the boot sequence.

-
6. Start installation from the Oracle VM Server software CD by and input **linux mpath** as your installation option at the **boot:** prompt in the first screen.
 7. Follow the installation steps provided in the *Oracle VM Server Installation Guide* to install and boot your OVM Server.

After completing the Oracle VM Server installation, the SAN LUN will be set as the primary candidate disk for booting up when the server powers on and runs through its self-check. In addition, as DMPP loads during the OS boot right after the HBA driver, your Linux will have multipath enabled. Now you can reconnect the previously disconnected HBA ports.

To scale the Oracle VM system as illustrated in Figure 2 (Server Pool 3), you may have to install the same version of Oracle VM Server on different physical hosts, which can be used as Virtual Machine Servers, and/or Utility Servers.

To enable advanced features like domain live migration or High Availability (HA) you have to install the same version of Oracle VM Server on other identical physical hosts.

Network configuration

An Oracle VM bridge is equivalent to a virtual switch presented to the virtual machines. Each physical server can have several bridges according to the availability of the physical network interface cards (NICs). Each physical NIC is assigned a bridge. One can have multiple guest domains attached to each bridge or vice versa. By default, Oracle VM presents the number of bridges equal to the physical network interface cards on the host. Each bridge is named xenbrx (where *x* corresponds to the number of the physical network interface). The host domain (dom0) and new domains (domUs) are attached to the default xenbr0 bridge unless otherwise specified while creating the guest domains using Oracle VM Manager.

Bridges typically consist of a connection to a physical interface and a virtual interface. The physical interface can be any network adapter in the server. The virtual interface corresponds to the domain Ethernet interface. To verify the available bridges on your host, use the command: `brctl show`.

EMC Navisphere Host Agent and CLI installation

Navisphere Host Agent communicates with the CLARiiON storage management software over a local area network (LAN). You must install the Navisphere Host Agent on all Oracle VM Server hosts. Navisphere Manager, the storage management software, provides a web-based UI that lets you configure the storage systems connected to the servers on the LAN.

After completing the Oracle VM Server installation, log in to the server and create a file called *agentID.txt* (case-sensitive) in the root directory, with the following two lines of information. This file ensures that the Navisphere Agent binds to the correct HBA/NIC for registration and therefore registers the host with the correct storage system.

Line 1: Fully-qualified hostname of the host (for example, ovmserver1.emc.com)

Line 2: IP address of the HBA/NIC port that you want Navisphere Agent to use (for example, 10.15.160.30)

Now install and start the EMC Navisphere Host Agent (NaviHostAgent-Linux-32-x86-en_US-6.29.0.6.1-1.i386.rpm) using the following commands.

```
# rpm -Uvh NaviHostAgent-Linux-32-x86-en_US-6.29.0.6.1-1.i386.rpm
# /etc/rc.d/init.d/naviagent start
```

Make sure that `/var/log/HostIdFile.txt` has the IP address you have entered in to the `agentID.txt` file.

Optionally you may need to disable the firewall using the following command if Navisphere is unable to see the updated information about this host.

```
# /usr/bin/system-config-securitylevel
```

Install the Navisphere Host Agent on the other nodes in the server pool after creating agentID.txt.

After installing and starting EMC naviagent, now you can add additional LUNs using the CLARiiON Navisphere management tool and assign the LUNs to the nodes in the server pool.

Please refer to the *EMC Navisphere Server Software for Linux Administrator's Guide* for additional details and the latest version of the software that can be downloaded from the EMC [Powerlink](#)[®] website.

Device-Mapper Multipath tool configuration for storage

The connection from the server through the HBA to the storage controller is referred to as a path. When multiple paths exist to a storage device (LUN) on a storage subsystem, it is referred to as multipath connectivity. It is an enterprise-level storage capability. The main purpose of multipath connectivity is to provide redundant access to the storage devices, that is, to have access to the storage devices when one or more of the components in a path fail. Another advantage of multipathing is the increased throughput by way of load balancing.

Each multipath device has a World Wide Identifier (WWID), which is guaranteed to be globally unique and unchanging. When new devices are brought under the control of DM-Multipath, the new devices may be seen in three different places under the /dev directory: /dev/mapper/mpath*n*, /dev/mpath/mpath*n*, and /dev/dm-*n*. By default, the multipathed devices are named with the WWID of the device, which one accesses through /dev/mpath/<wwid_name>. When the user_friendly_names configuration option in /etc/multipath.conf is set to yes, devices will be named as mpath0, mpath1, and so on, which may meet one's needs. Users also have an option to define an alias in /etc/multipath.conf for each device. Once multipath.conf file is changed, the multipath tools need to be rerun for those configuration values to be effective. One has to kill multipathd, run multipath -F, and then restart *multipathd* and *multipath*. The names (WWID_names or mpath names or alias names) that appear in /dev/mpath are persistent across boots, and the names /dev/dm-0, /dev/dm-1, and so on, can change between reboots. So, it is advisable to use the device names that appear under /dev/mpath and avoid using the dm-N names directly. To configure device partitions, create device partitions before running multipath. Partitions on device mpath0 appear as /dev/mpath/mpath0p1, /dev/mpath/mpath0p2, and so on.

If you install the Oracle VM Server to the CLARiiON CX4 storage array then the Device-Mapper Multipath tool is enabled automatically, but if you installed the Oracle VM Server to a local hard drive then follow the steps below to enable DMPP. Once enabled, DMPP will assign the persistent user-friendly names to the LUNs exposed by the storage array. The persistent user-friendly names stay the same across the reboots or when the new LUN gets added.

Edit the device mapper multipath configuration file /etc/multipath.conf and make sure that it has all the right settings.

Add the multipath-tools startup scripts to the startup list as follows:

```
# chkconfig --level 35 multipathd on
```

Verify that it's set up correctly:

```
# chkconfig --list | grep multipath
```

Edit the */etc/multipath.conf* file and make the necessary changes to add support for storage devices.

Start the multipath daemon service by typing the following command:

```
# service multipathd start
```

Check the multipath device names list and if the list is not populated then rescan the HBAs or reboot the system:

```
# multipath -l (or)
```

```
# dmsetup ls
```

Perform the above procedure on the other nodes in the server pool. The devices appear in the folder /dev/mpath. We can now proceed to create desired partitions on the multipath devices and then create the OCFS2 file system for shared access by all nodes in the server pool. To create the partitions, type:

```
# fdisk /dev/mpath/mpath1
```

Use KPARTX to view and add the newly created partitions to the multipath configuration:

```
# kpartx -a /dev/mpath/mpath1
```

On the other nodes in the pool, perform:

```
# kpartx -l /dev/mpath/mpath1
```

```
# kpartx -a /dev/mpath/mpath1
```

Please refer to Red Hat's [Using Device-Mapper Multipath](#) document for additional information about configuring multipath.

Creating shared storage and clusters on CLARiiON CX4

Shared storage is used for the shared repository, which is used to store VM resources, and to create shared virtual disks. To enable High Availability (HA), or to perform live migration of domains to other identical computers, you must create a shared virtual disk to be used as shared storage, and then create the cluster.

Shared repository⁴

Oracle VM uses the shared storage to create storage repositories and shared virtual disks. The concept of storage repositories is used to define where Oracle VM resources may reside. Resources include guest virtual machines, virtual machines templates (guest seed images), ISO images, shared virtual disks, and so on. A storage repository is used for live migration of guests, HA, and for local storage. You can have multiple storage repositories for Oracle VM resources.

The /OVS directory is the location of the default storage repository created when you install Oracle VM Server with the default partition layout. When you create a server pool in Oracle VM Manager, the repository is automatically mounted with the source type, for example, NFS, OCFS2, or ext3. The repository directory structure is also automatically created under the /OVS directory on the repository. These directories are listed in Table 1. The /OVS directory is the cluster root and is a symbolic link mounted to the /var/ovs/mount/*uuid* directory.

Table 1. /OVS directory contents

Directory Name	Description
iso_pool	Contains ISO files imported using Oracle VM Manager
publish_pool	Contains guest virtual machines deployed as public
seed_pool	Contains guest virtual machine templates and initial configuration files
sharedDisk	Contains shared virtual disks
running_pool	Contains virtual machine images and configuration files

Storage repositories are managed by Oracle VM Agent. Oracle VM Agent is installed and configured during the installation of Oracle VM Server. You can change the default configuration of Oracle VM Agent by running a configuration script (`#service ovs-agent configure`) as a root user. When you create a storage repository, Oracle VM Agent saves the configuration information in the Oracle VM Agent database. You can have multiple storage repositories for Oracle VM resources. If you create a server pool using Oracle VM Manager, the storage repository is automatically propagated to the other nodes

⁴ [Oracle VM Manager User's Guide - Release 2.2](#)

(Oracle VM Servers) in the cluster. You can manage storage repositories using the `/opt/ovs-agent-2.3/utills/repos.py` script. This script should be run on the Server Pool Master. The `repos.py` script supports the following operations.

- Listing the Storage Repositories
- Adding a Storage Repository
- Deleting a Storage Repository
- Initializing a Storage Repository
- Setting the Cluster Root

Configuring shared storage on the multipath device

To create a shared virtual disk as shared storage from a multipath device partition you must manually configure the multipath device on each Oracle VM Server in the server pool, and then create the cluster. The prerequisite steps to use a multipath device as a shared virtual disk are:

1. Review the multipath devices by checking the output of the `multipath -l` command and identify the multipath device's WWID that needs to be configured as shared storage.
2. Make sure the multipath device has the same device identifier and device path on each Oracle VM Server in the server pool. For example, if you want to use a multipath device with the device identifier `mpath-3600601604ca02400e04717e64ea1de11` it must have the same device path name on each Oracle VM Server in the server pool, for example, `/dev/mpath/ovmsr1`. (You should not use the path `/dev/mapper/mpathN` for multipath devices; only use the path `/dev/mpath/mpathN`). To configure the same device path name use the *alias* setting the `/etc/multipath.conf` file.
3. Make sure the multipath device is not used for any other application to avoid any possible data corruption. When you have added the shared storage, create the cluster using the steps below.
4. Partition the device using `fdisk`, update the device maps using `kpartx` and create an OCFS2 file system on the device using at least a 1 MB cluster size. In the following commands `ovmsr1` refers to the alias name of the DMMP device.

```
# fdisk /dev/mpath/ovmsr1
# kpartx -a /dev/mpath/ovmsr1
# mkfs.ocfs2 -Tdatafiles -N8 /dev/mpath/ovmsr1p1
```

Perform the following on all the other nodes in the server pool so that all the nodes have updated device maps for the new partitions.

```
# kpartx -l /dev/mpath/ovmsr1
# kpartx -a /dev/mpath/ovmsr1
```

Configuring the cluster and shared repository using OCFS2 on the multipath device

In addition to creating shared storage for HA and live migration, you must also create an Oracle VM Server cluster to ensure the safety of guest data, and protect from runaway nodes that may become unreachable. You can create a cluster using OCFS2 shared storage. You must have shared storage created and ready to use as described in the previous section before you create a cluster.

Before you create a cluster, the *hostname* of every Oracle VM Server in the server pool should be correctly configured with its public IP address. To make sure the hostname is correctly set, ping the hostname of all Oracle VM Servers in the server pool to display their public IP addresses.

To create a cluster and storage repository use these steps:

-
1. Add the storage repository: On the Server Pool Master, create a storage repository by using the following script. You can use the following command to create/add one or more storage repositories

```
# /opt/ovs-agent-2.3/utils/repos.py --new /dev/mpath/ovssr1p1
```

2. List the current repositories: List the UUID for the storage repository. You need this to set the cluster root in the next step. Retrieve the UUID for the storage repository using the command:

```
# /opt/ovs-agent-2.3/utils/repos.py --list
```

This command lists the storage repositories and the UUID for each. Copy the UUID for the storage repository you want to use as the cluster root.

3. Set the cluster root: You can select one storage repository as the cluster root in a server pool. Use the following command and paste the UUID for the storage repository from the previous step to set the cluster root. Alternatively, if you have only one Oracle VM Server in the server pool, you can select your local disk as the storage repository and set it as the cluster root.

```
# /opt/ovs-agent-2.3/utils/repos.py --root UUID
```

The shared storage, storage repository, and cluster are configured and ready to use. Log in to the Oracle VM Manager and create a server pool. Please refer to the next section on how to create a server pool. The shared storage configuration is then propagated to all Virtual Machine Servers in the server pool and updates */etc/ocfs2/cluster.conf* automatically on all the nodes. In case you created the server pool before creating the shared repository then you can use the following command to propagate the shared configuration as well.

```
# /opt/ovs-agent-2.3/utils/repos.py --init
```

You can create multiple storage repositories in a server pool. If you want to add to an additional storage repository to an existing server pool, you must first add it, and then initialize it to create the storage repository directories, and mount it on all the nodes in the cluster.

Managing server pool(s), servers, and resources⁵

A server pool is a logical entity that contains one or more Oracle VM Servers. It presents a unified view of the storage in which the virtual machines reside. A server pool consists of one Server Pool Master, and at least one Utility Server, and one or more Virtual Machine Servers. You can configure one or more servers to perform these roles. In a server pool there is only one Server Pool Master at one time.

Before creating a server pool, you need to consider how many physical Oracle VM Servers are to be included in the server pool, and what function(s) each physical Oracle VM Server is to perform. The more guest virtual machines you run in the server pool, the more resources these guest virtual machines will consume, and so the more physical Oracle VM Servers may be required to provide sufficient resources for the server pool. A server pool is scalable. If you find a server pool does not have sufficient resources, such as CPU or memory, to run the virtual machines, you can expand the server pool by adding more Virtual Machine Servers using Oracle VM Manager. Make sure you have the following before creating server pool:

- Oracle VM Servers to deploy as the Server Pool Master, the Utility Server, or the Virtual Machine Server
- A repository that is used for live migration of virtual machines and for storing OS images and templates on the Oracle VM Servers
- IP addresses or hostnames for the Oracle VM Servers

⁵ [Oracle VM Manager User's Guide - Release 2.2](#)

-
- Passwords to access Oracle VM Agent installed on these Oracle VM Servers (password can be changed as a root user with the command `#service ovs-agent configure`).

After installation, you can log in to Oracle VM Manager, and follow the wizard to create a server pool containing a physical server that will act as the Server Pool Master, Virtual Machine Server, and Utility Server.

Oracle VM Manager is also used to manage resources. Resources include virtual machine templates, virtual machine images, ISO files, virtual disks, and converted virtual machines.

You can create shared virtual disks, and use them to expand the storage capacity of your virtual machines. The available virtual disks can also be used by other users in the group. Oracle VM supports using files as virtual disks. In this case, the prefix of the disk in the configuration file is *file*. The following is an example of an entry for a file-based disk in the virtual machine configuration file (vm.cfg):

```
disk = [ 'file:/OVS/seed_pool/system.img,hda,w', ]
```

Physical devices can be used but cannot be changed through Oracle VM Manager. You can share your physical devices (multipath devices, normal SAN LUNs, and so on) with all Virtual Machine Servers in the server pool. In this case, the prefix of the disk in the configuration file is *phy*. The following is an example of an entry for a multipath disk in the virtual machine configuration file (vm.cfg).

```
disk = [ 'phy:/dev/mpath/mpath1,hdd,w', ]
```

You must manually configure the multipath device on each Virtual Machine Server in a server pool, then you can add the multipath device to the server pool as a shared virtual disk. The prerequisite steps to use a multipath device as a shared virtual disk are:

- Manually set up the SAN connection to the multipath device on each Virtual Machine Server in the server pool.
- Manually configure the multipath device on each Virtual Machine Server in the server pool.
- Make sure the multipath device has the same device identifier and device path on each Virtual Machine Server in the server pool. For instance if you wanted to set up the `/dev/mpath/mpath-3` device as a shared device, make sure this path points to the same LUN on all the nodes in the server pool. Do not use the path `dev/mapper/x` for multipath devices, only use the path `dev/mpath/x`.
- Make sure the multipath device is not used for any other purpose to avoid any possible data corruption.

For detailed information on how to manage server pools, please refer to the *Oracle VM Manager User's Guide*.

Creating and configuring virtual machines

A virtual machine contains a guest operating system and some preinstalled applications. It runs on a Virtual Machine Server. An operating system installed inside a virtual machine is known as a guest operating system. Oracle VM supports a variety of guest operating systems.

Before you create a new virtual machine, ensure that the following resources are available:

- A server pool
- A Virtual Machine Server to run the virtual machine
- Imported ISO files for creating virtual machines from installation media, or imported virtual machine templates for creating virtual machines based on templates

Typically, the life cycle of a virtual machine consists of the following process and is supported by Oracle VM Manager:

- Creating a virtual machine
- Starting a virtual machine

-
- Connecting to a virtual machine's console (to access the virtual machine itself)
 - Shutting down a virtual machine (to reduce system resource consumption when the virtual machine is not in use)
 - Deleting a virtual machine (when you no longer require the virtual machine)

You can use Oracle VM Manager to create a virtual machine and to perform other life cycle operations. See the [Oracle VM Manager User's Guide](#) for information on creating a virtual machine Oracle VM Manager.

Alternatively, you can create a virtual machine and perform other life cycle operations from the command line on an Oracle VM Server using **xm**.

When creating a virtual machine, you can configure some of the parameters, such as the number of virtual CPUs, the size of memory, and so on; the other parameters, such as boot source, virtual disk, and the type of virtual network interfaces (VIFs), are assigned default settings by Oracle VM Manager, and you cannot configure them when creating the virtual machine. If needed, you can configure such parameters *after* the virtual machine is created using Oracle VM Manager. In this paper we are going limit our focus to the storage aspects of the VM configuration

Configuring storage for the VM

There are two ways to expand the storage capacity of a virtual machine: by creating *non-sharable virtual disks*, or by adding *shared virtual disks*. Shared virtual disks can only be shared among virtual machines within the same server pool. Oracle VM supports using files as virtual disks. In this case, the prefix of the disk in the configuration file is *file*. The following is an example of an entry for a file-based disk in the virtual machine configuration file:

```
disk = [ 'file:/OVS/seed_pool/system.img,hda,w', ]
```

In addition to file-based shared virtual disks, you can share your physical devices (multipath devices, normal SAN LUN devices, and so on) with all Virtual Machine Servers in the server pool. In this case, the prefix of the disk in the configuration file is *phy*. Multipath devices can be used, but cannot be changed or configured through Oracle VM Manager. The following is an example of an entry for a multipath disk in the virtual machine configuration file.

```
disk = [ 'phy:/dev/mpath/mpath1,hdd,w', ]
```

You must manually configure the multipath device on each Virtual Machine Server in a server pool, then you can add the multipath device to the server pool as a shared virtual disk.

For a Hardware Virtualized Machine (HVM), you can attach up to four IDE disks (including CD-ROM and shared virtual disks), and seven SCSI disks. A paravirtualized machine (PVM) has no such limitation.

Live domain migration and High Availability setup

Live migration of virtual machines

Live migration is a process to migrate a running virtual machine from one Virtual Machine Server to another, while applications on the existing virtual machine continue to run. Live migration ensures high availability of virtual machines. This feature is important, and useful, when the existing Virtual Machine Server may be out of commission, or on a planned shutdown for maintenance purposes. Cross-server-pool live migration is not allowed. You can only migrate virtual machines from one Virtual Machine Server to another within the same server pool. You must use identical computers to perform live migrations, that is, the computer make and model number of both the source computer and the destination computer must be identical.

If you want to perform live migration of domains to other, identical computers, you must create a shared virtual disk to be used as shared storage and used during the live migration. The live migration feature will require device consistency on all nodes in the server pool and makes guest OS management more involved.

You can use Oracle VM Manager to migrate a virtual machine. See the *Oracle VM Manager User's Guide* for information on migrating virtual machines using Oracle VM Manager.

Alternatively, you can migrate a virtual machine from the command line on an Oracle VM Server. To migrate a virtual machine from the Oracle VM Server command line, on the Oracle VM Server that contains the virtual machine, use the following command:

```
# xm migrate mydomain myremotecomputer
```

To perform live migration of the virtual machine, use the command:

```
# xm migrate -l mydomain myremotecomputer
```

In our lab tests, we have successfully migrated the live virtual machines using the setup shown in Figure 3.

Enabling High Availability (HA)

The High Availability (HA) feature in Oracle VM guarantees the availability of virtual machines if the Virtual Machine Server they are running on fails or restarts. When a Virtual Machine Server is restarted or shut down, the virtual machines running on it are either restarted on, or migrated to, another Virtual Machine Server.

Oracle VM Manager is required to manage HA of a system. To implement HA, first you must create a cluster of Virtual Machine Servers in a server pool and have them managed by Oracle VM Manager. HA cannot be implemented with Oracle VM Server alone. To use HA, you must first enable HA on the server pool, then on all virtual machines using Oracle VM Manager. You can set up HA for the virtual machine by selecting **Virtual Machines > Configure > Policies > Enable High Availability** and for server pool by selecting **Server Pools > Edit > Enable High Availability** in the Oracle VM Manager. If HA is not enabled for both server pool and for virtual machines, then HA will be disabled.

HA is set up in Oracle VM to guarantee the availability of virtual machines if the Virtual Machine Server they are running on fails or restarts. When a Virtual Machine Server is restarted or shut down, the virtual machines running on it are either restarted on, or migrated to, another Virtual Machine Server which is selected using the preferred server setting for the server pool when you create a virtual machine in the Oracle VM Manager. If there is no preferred Virtual Machine Server or Virtual Machine Server available, the virtual machines shut down (Power Off) and are restarted when a Virtual Machine Server becomes available.

If the Server Pool Master fails, another Oracle VM Server is selected from the server pool to act as the Server Pool Master. The Oracle VM Server with the most resources available is the one chosen to take over the Server Pool Master role. To use the Server Pool Master failover feature, you should make sure the Oracle VM Agent password is identical on all Oracle VM Servers in the server pool.

In our lab tests, we have verified that HA feature works well based on the setup shown in Figure 3.

OVM test scenarios

In addition to the current Oracle VM 2.2 version, EMC has been involved in testing earlier versions of OVM in our lab. The following scenarios were successfully tested using Oracle VM in our lab:

- ORION workload simulation testing
- Shared storage using CLARiiON CX4 and DMMP
- Boot from SAN (BFS) testing
- Virtual machine Live Migration testing
- High Availability testing

Storage considerations for PVM and HVM

Based on these mentioned test scenarios, a PVM guest OS is recommended where available for better performance. Use of PVM (over HVM) in Oracle VM environments for a better performance⁶ is also verified by Oracle.

Conclusion

Oracle VM 2.2 configured on EMC CLARiiON CX4 storage as described in this paper is designed to help optimize performance and simplify the deployment and operations of applications in a virtualized environment. Oracle VM 2.2 using CLARiiON storage provides the flexibility of adding additional resources as needed as the workload demands increase by adding additional storage to the server pools level or at the VM guest level. CLARiiON CX4 storage simplifies storage operations, improves utilization, and cost-effectively scales as your needs grow over time. This white paper provides the deployment guidelines for setting up an Oracle VM 2.2 platform using x86 servers and CLARiiON CX4 storage. The best practices described here are intended to help achieve optimal performance of applications running in an Oracle VM virtualized environment.

References

EMC documentation can be found on Powerlink.

The following on the Oracle website may be helpful:

- [*Oracle VM Supported guests matrix*](#)
- [*Oracle VM Features*](#)
- [*Oracle VM Documentation Release 2.2*](#)
- [*Architect Center - Virtualization*](#)
- [*Oracle VM Wiki*](#)
- [*Oracle VM Templates*](#)
- [*Oracle E-Delivery Web site for Enterprise Oracle VM*](#)
- [*Oracle VM Cheat Sheet*](#)
- [*Multipath Usage Guide*](#)
- [*Using Device-Mapper: Multipath*](#)
- [*Multipath Device Identifiers*](#)
- [*Oracle VM Server Guest Configuration*](#)
- [*Creating a Guest Virtual Machine*](#)
- [*Oracle VM Quick Start Guide*](#)
- [*Oracle VM Manager Release Notes*](#)
- [*Oracle VM Server Release Notes*](#)
- [*Oracle VM Manager Installation Guide*](#)
- [*Oracle VM Server Installation Guide*](#)
- [*Oracle VM Template Builder Installation and User's Guide*](#)
- [*Oracle VM Manager User's Guide*](#)
- [*Oracle VM Server User's Guide*](#)
- [*Oracle VM Upgrade Guide*](#)

⁶ Chapter 2, "Domains, Guests and Virtual Machines" section of the [Oracle VM Server User's Guide](#)

-
- [Oracle VM Windows Paravirtual Drivers Installation Guide for Microsoft Windows](#)
 - [Oracle VM Manager Web Services API Reference](#)
 - [Virtualization Management with Oracle Enterprise Manager](#)