InfoScale Deployment on Virtual Machines in VMware vSphere ESXi Environments

Last updated: 2020-09-17

VERITAS™

The truth in information.

Contents

Introduction

When deploying VMware virtual machines, the variants of InfoScale seamlessly operate within the vSphere hyper-converged infrastructure. Achieving higher levels of availability for business-critical applications while reducing overall cost is now much simpler with the vSphere ESXi hypervisor. vSphere makes it possible to mitigate both planned and unplanned downtime by leveraging advanced features such as VMware vMotion™ (vMotion), Storage vMotion, and vSphere HA. vSphere HA specifically reduces unplanned downtime by leveraging multiple VMware ESX® and VMware ESXi™ hosts configured as a cluster, thereby providing rapid recovery from outages as well as cost-effective high availability (HA) for individual virtual machines.

VMware Virtual SAN (vSAN) is a distributed object storage platform, which depends on IP network connectivity to provide shared storage resources and infrastructure management services. vSAN requires that each of the participating ESXi hosts are members of the same vSphere cluster. Using a software-defined approach, the vSAN architecture creates shared datastores by virtualizing local HDD and flash resources from each ESXi host. By abstracting the local nature of the ESXi storage resources, direct-attached devices can now be incorporated into storage pools. Such storage pools that can be divided and assigned to virtual machines as well as applications according to their quality-of-service requirements, irrespective of which ESXi host the guest resides on. vSAN is implemented directly in the ESXi hypervisor, and therefore, presents itself as a typical datastore when instantiating a new VMware guest or migrating an existing one.

Veritas InfoScale Enterprise provides application-level HA and storage management within the virtual clusters themselves. In case of virtual machine failure, advanced VMware features can still be used. Thus, this combined approach not only guarantees an optimized infrastructure but also provides the granularity needed to account for any constituent applications or services.

Benefits of InfoScale and VMware vSphere integration

InfoScale and VMware vSphere integration offers following benefits:

- **Simplified management**
  - Automated application resiliency through monitoring and orchestration
  - Application-aware operations
  - Accelerated private cloud and public cloud adoption
- **Resiliency**
  - Protection for mission-critical applications against failures and disasters
  - Protection for data with scale and performance
  - Migration of workloads between platforms—permanently, or in the event of a disaster
  - Adaptive HA with automated application failover using InfoScale and virtual workloads using vSphere
  - Synchronous and asynchronous replication with DR orchestration between sites for near-zero RTO and zero RPO

Overview

InfoScale Enterprise offers an extensive catalog of enterprise-grade features and functionalities, such as: Cluster File System, Veritas Oracle Disk Manager (ODM), and server-based I/O fencing (CP Server). These features ensure that performance expectations, resiliency, and stability are each maintained in response to a failure or a service disruption.

Veritas InfoScale Enterprise

The InfoScale Enterprise solution provides an end-to-end solution for enterprise storage management. It virtualizes the heterogeneous storage over heterogeneous servers into logical objects. The following components and features are included:

- **Veritas Volume Manager (VxVM)** virtualizes any block-based storage that is visible to the system. Users can then choose the desired resiliency level or RAID parity. VxVM can even replicate data across remote sites with its Veritas Volume Replicator (VVR) feature. VxVM features such as Dynamic Multi-Pathing (DMP), snapshots, Fast Mirror Resync (FMR), and SmartMove migration provide further resiliency, scale, and faster recovery of your critical services.
- **Veritas Cluster File System (CFS)** is a POSIX-compliant and highly resilient derivative of the Veritas File System (VxFS). CFS provides a global namespace across multiple x86, SPARC, or RISC-based servers. The resulting topology allows an application to concurrently access data from any of the nodes within the cluster. CFS tuning can further improve performance with application-aware caching and data flow to the underlying storage. With checkpoints and FSCK enhancements, data and metadata can be quickly recovered. CFS achieves linear scalability of application performance for a range of common workloads, thus guaranteeing the scale-out compute power for your application.
- **Veritas Cluster Server (VCS)** provides a sophisticated cluster membership framework such that your critical applications are monitored in real time for any number of factors such as state changes or unplanned disruptions. InfoScale Availability (known as VCS in versions prior to 7.0) enables the deployment of large-scale clusters, with 100+ participating nodes. Furthermore, VCS allows for seamless failover of not just an individual application but also groups of applications—which form entire services—from one node to another. Upon doing so, VCS also considers intelligent failover policies when determining the target on which to move said application or service. These include memory and CPU utilization as well as the existence of less critical workloads. VCS is also extensible, automating recovery across several different failover topologies, including local, metro/stretched, and wide-area DR.
- **I/O fencing** is a core component of VCS that is focused on properly handling a cluster partition event that occurs due to the loss of cluster communication. I/O fencing consists of two distinct components, *membership arbitration* and *data protection*. These two components together can deliver maximum data integrity in a cluster environment. Membership arbitration is necessary to ensure that when the cluster members are unable to communicate over the cluster heartbeat network, only a single subcluster is active.

  In vSAN environments, SCSI3-PR support is not available, because RDM-P disks assignment is not supported. The data protection aspect of non-SCSI3 based fencing is implemented through the use of judicious timing. When a fencing race occurs, a minimal time gap is put in place before attempting to bring the application online on the subcluster that wins the race. This is to ensure that there is enough time for the losing node to panic and reboot. In environments that do not support

SCSI3-PR, Veritas recommends the deployment of non-SCSI3-based fencing. Non-SCSI3-based fencing can be configured using coordination point servers (CP servers), that are placed outside the client InfoScale clusters.

- **Application availability with VCS**

  Using VCS virtual-to-virtual or in-guest clustering in VMware environments provides HA of applications inside the guest. This is achieved by providing protection from host failures, hardware failures, OS crashes, and application failures at the software layer. For example, in cases of application hang, file-level corruption at the OS level cannot be resolved with a reboot.
  Since there is a cost involved in maintaining standby virtual machines, you may choose to protect only specific applications by using VCS in-guest and to protect the remaining applications using VMware HA. By using VMware-HA in conjunction with VCS in the guest, when a host fails, the standby VCS nodes running on that host are automatically restarted by VMware HA on a new host. There is no need for user intervention, which potentially eliminates the need to maintain multiple standbys.

- **VCS support for live migration**

  VCS in-guest clustering continues to provide HA of applications on virtual machines, in live migration scenarios initiated by the virtualization technology. You can use live migration to perform a stateful migration of a virtual machine in a VCS environment. During this period, you may see notifications if the migrating node is unable to heartbeat with its peers within the default peer inactive timeout of LLT.
  To avoid false failovers, determine how long the migrating node is unresponsive in your environment. If that time is less than the default LLT peer inactive timeout of 16 seconds, VCS operates normally. If not, increase the peer inactive timeout to an appropriate value on all the nodes in the cluster before beginning the migration. Reset the value back to the default after the migration is complete.

VMware vSAN

You can configure vSAN to work as either a hybrid or an all-flash cluster. In hybrid clusters, flash devices are used for the cache layer and magnetic disks are used for the storage capacity layer. In all-flash clusters, flash devices are used for both cache and capacity.

vSAN aggregates all local capacity devices into a single datastore that is shared by all the hosts in the vSAN cluster. You can expand the datastore by adding capacity devices or hosts with capacity devices to the cluster. vSAN works best when all the ESXi hosts in the cluster share similar or identical configurations across all the cluster members, including similar or identical storage configurations. This consistent configuration balances virtual machine storage components across all the devices and hosts in the cluster. Hosts without any local devices can also participate and run their virtual machines on the vSAN datastore. If a host contributes its local storage devices to the vSAN datastore, it must provide at least one device for flash cache and at least one device for capacity. Capacity devices are also called data disks. The devices on the contributing host form one or more disk groups. Each disk group contains one flash cache device and one or multiple capacity devices for persistent storage. Each host can be configured to use multiple disk groups.

The details of the setup used for the sanity qualification are as follows:

Hardware components

- HP ProLiant DL380 Gen9: Intel(R) Xeon(R) CPU E5-2690 v4 @ 2.60GHz, 512 GB Memory

Software components

- VMware ESXi, 6.7.0, 8169922
- VMware vSAN, 6.7.0
- InfoScale version: 7.4.1 (3-node CFS cluster)
- Guest OS: Red Hat 7.6, Red Hat 6.9

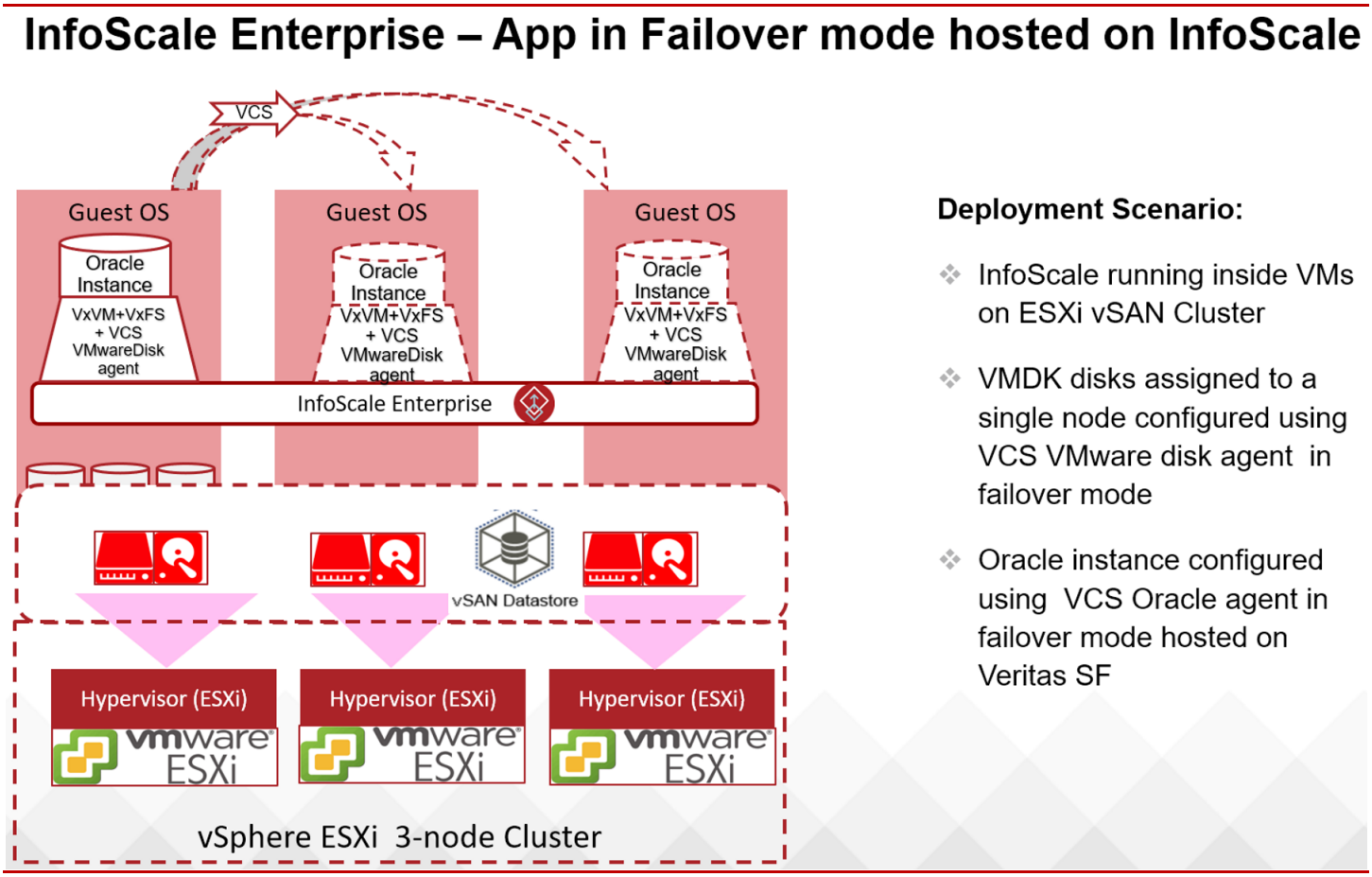
Use cases coverage

This document covers the use cases that are commonly used and are required in any customer environment to provide HA, performance, and reliability. These use cases assume a setup where:

- VCS monitors the application using the VCS agents and other components.
- VxVM, along with either VxFS or CFS, provides the storage management for the application data.

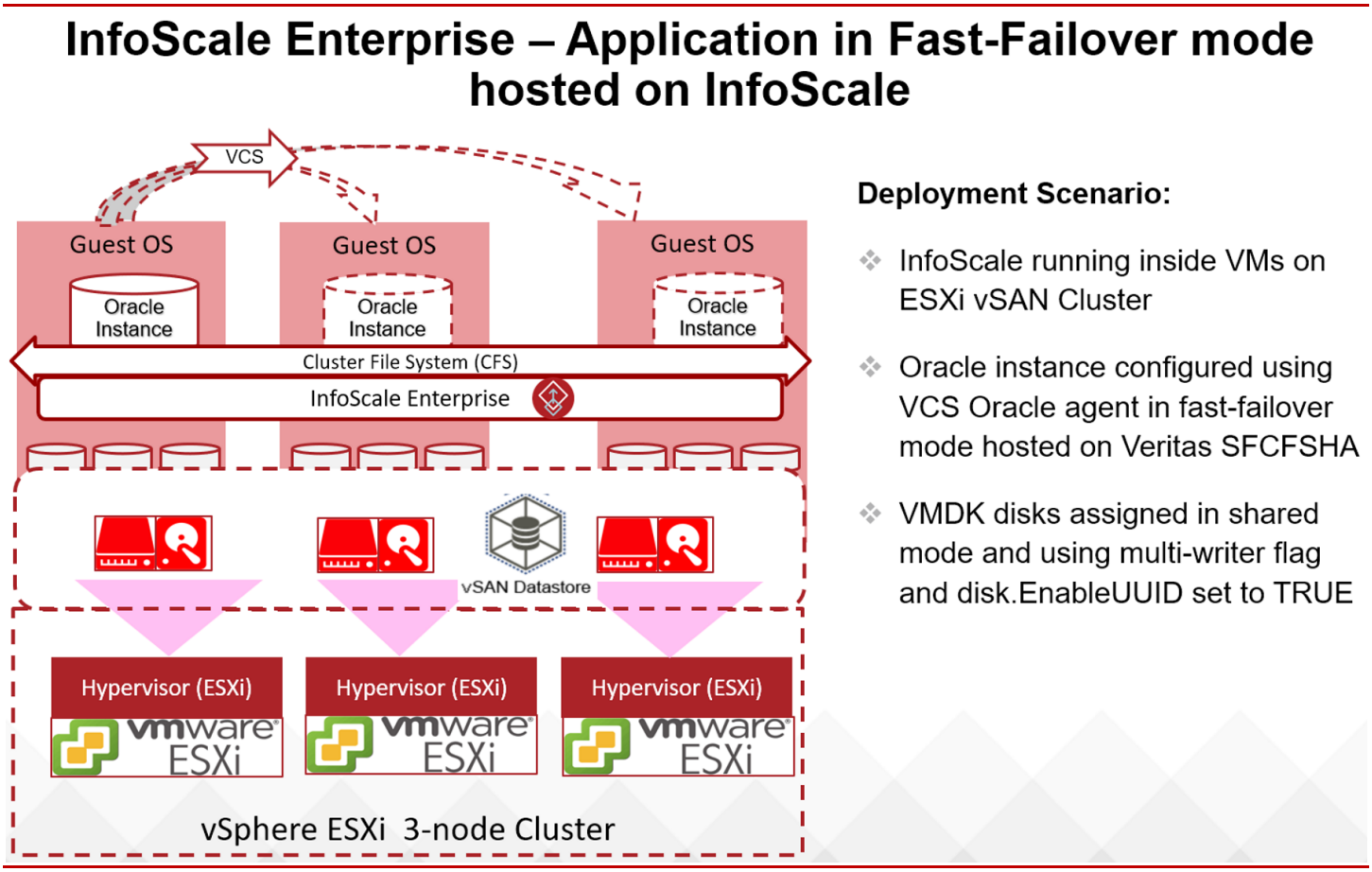
Use case 1

In this use case, the application data itself resides on VxFS, which allows it to be available on one of the cluster nodes at any given time in an Active-Passive configuration. In case of an InfoScale cluster node failure, the storage infrastructure uses VxFS to re-associate each application volume to a surviving node. The VMDK disks on which the volumes reside will also be detached and subsequently re-attached to the same targeted node. Upon completing this storage resource transition, any configured service or application impacted by said outage will be brought back online.



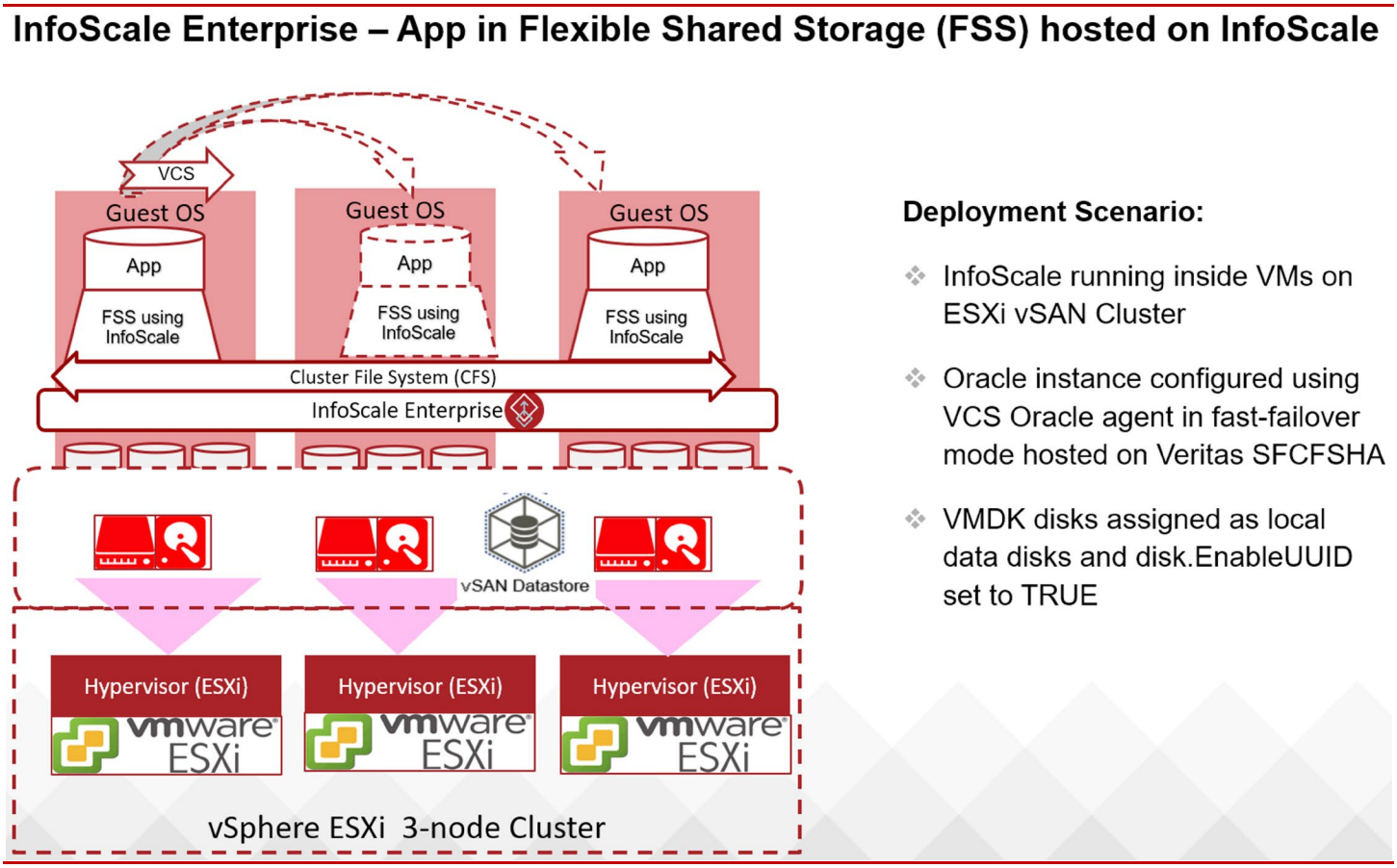
Use case 2

In this use case, the application data resides on CFS, which enables it to be available on all the cluster nodes concurrently, as an Active-Active configuration. In case of a node failure, only those applications that are monitored by a VCS agent will fail over to an available node in the cluster. Thus, application downtime is significantly reduced as shown in the following graphic:



Use case 3

In this use case, the application is configured in fast failover mode using the Veritas Flexible Storage Sharing (FSS) functionality, which eliminates the use of traditional shared storage. Cluster nodes with unused local disks can share these resources across the cluster heartbeat network, thereby creating a shared namespace that can be leveraged by all the participating nodes.



Prerequisites and setup recommendations

To set up an InfoScale cluster on VMware virtual machines, ensure that the following prerequisites are met:

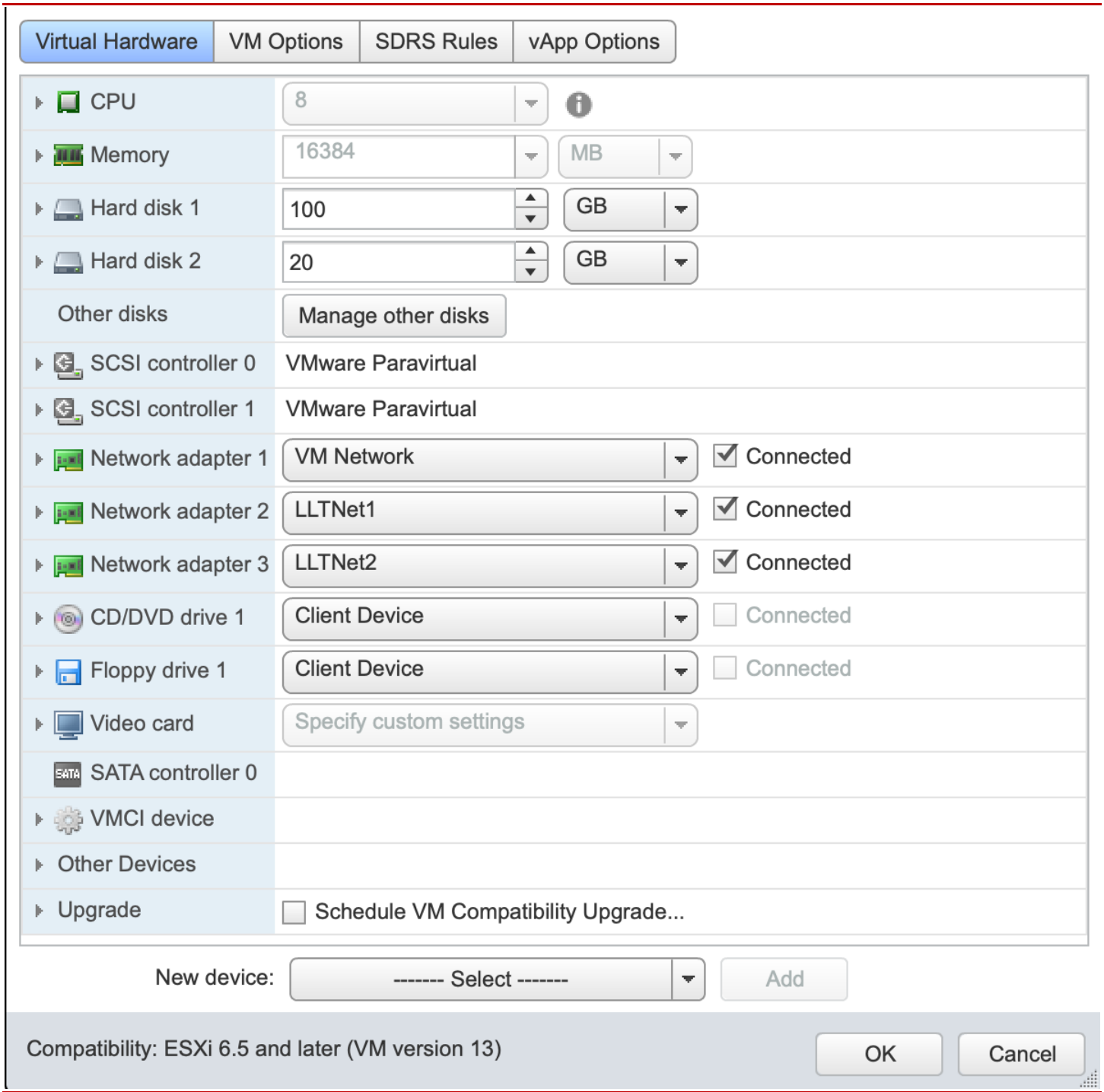
- A VMware vSAN 3-node cluster is set up and a vSAN datastore is available for use.
- The required OS images are configured in the datastore or the required OS ISO files are available.
- Each virtual machine is hosted on a different ESXi node of an ESXi cluster
- Multi-writer enabled VMDKs are created on the vSAN datastore, by following the procedure described in the VMware knowledge base article at (for details, contact VMware Support): https://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=2121181
- The disks are configured. EnableUUID is set to TRUE in the configuration file of each virtual machine.
- Virtual network details for the public network and the private network are available.

Creating virtual machines in VMware vSphere

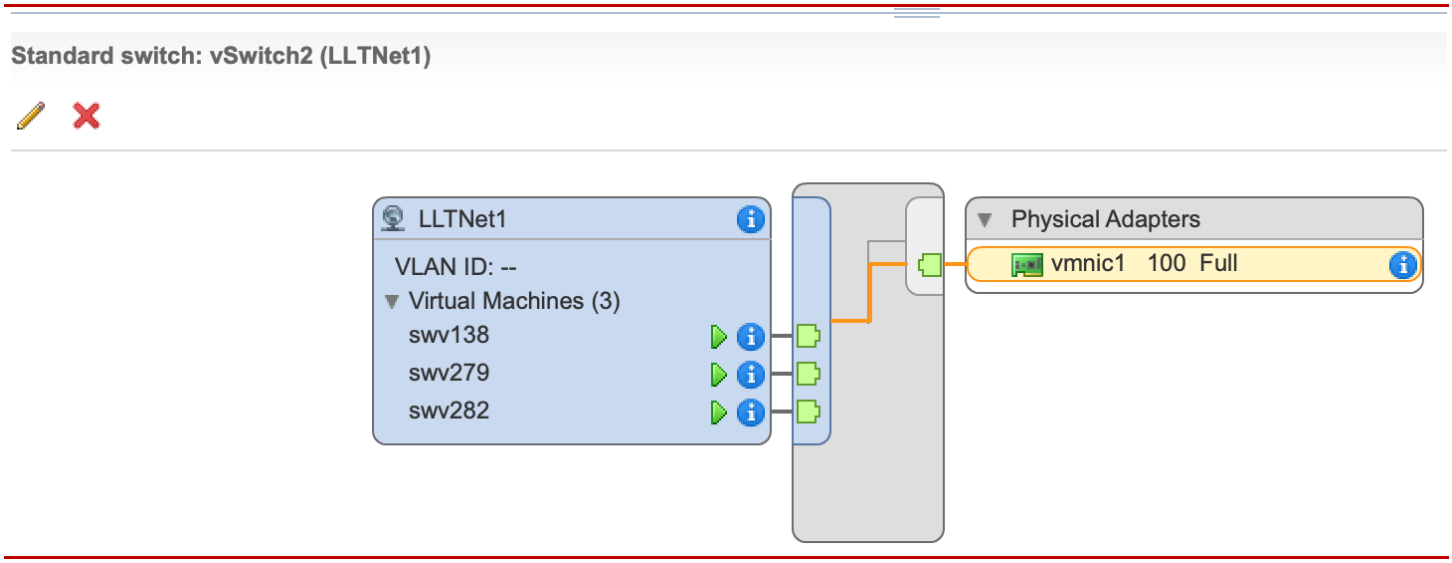
You can create a VMware virtual machine using the steps listed in the VMware documentation at:

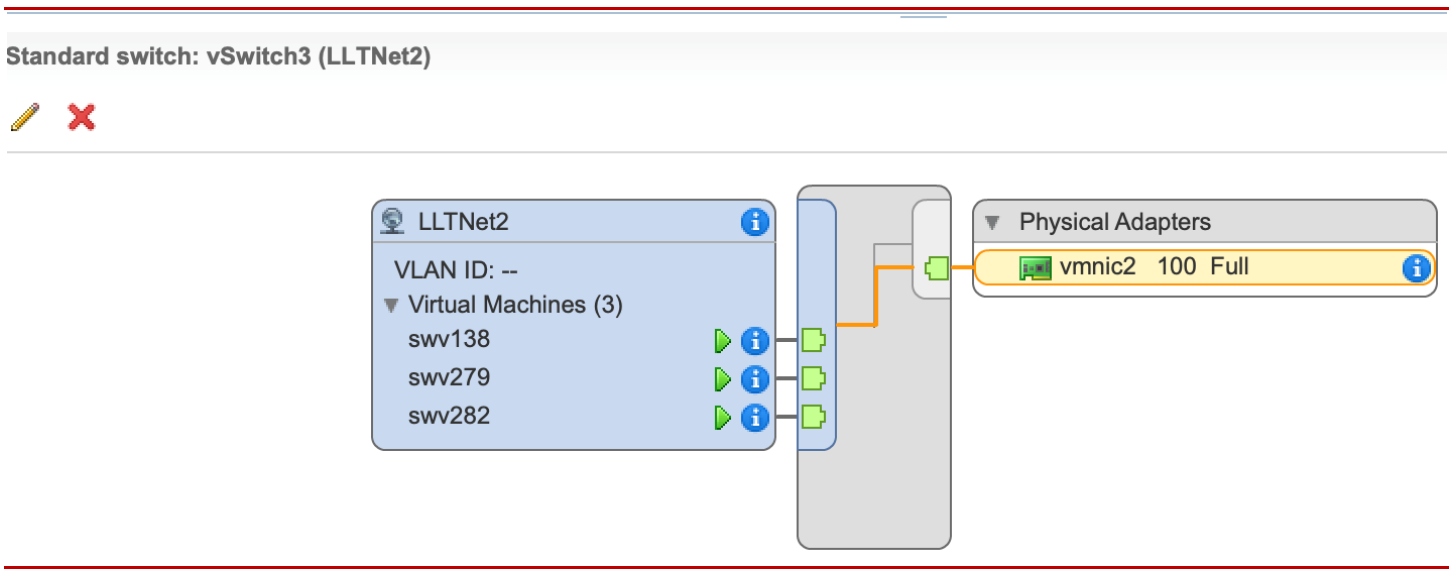
https://docs.vmware.com/en/VMware-vSphere/6.7/com.vmware.vsphere.vm_admin.doc/GUID-55238059-912E-411F-A0E9-A7A536972A91.html

The following screenshot depicts the resources (CPU, memory, boot disk, and vNICs assignments) that must be specified when you click the **Create VM** button on the VM pane:



After the successful creation of virtual machines, the vNICs assignments for the cluster heartbeat (LLT network) should look like the ones depicted in the following screenshots:





Creating storage infrastructure in VMware vSphere

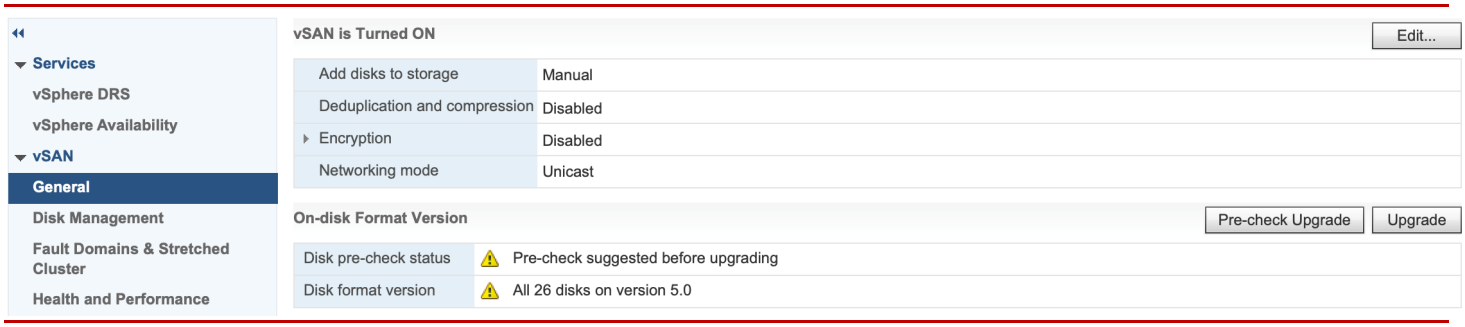
To configure storage for a virtual machine in the shared mode, you must first create a shared datastore; for example, a vSAN datastore. VMware vSAN is a distributed layer of software that runs natively as a part of the ESXi hypervisor. vSAN aggregates local or direct-attached capacity devices of a host cluster and creates a single storage pool that is shared across all the hosts in the vSAN cluster.

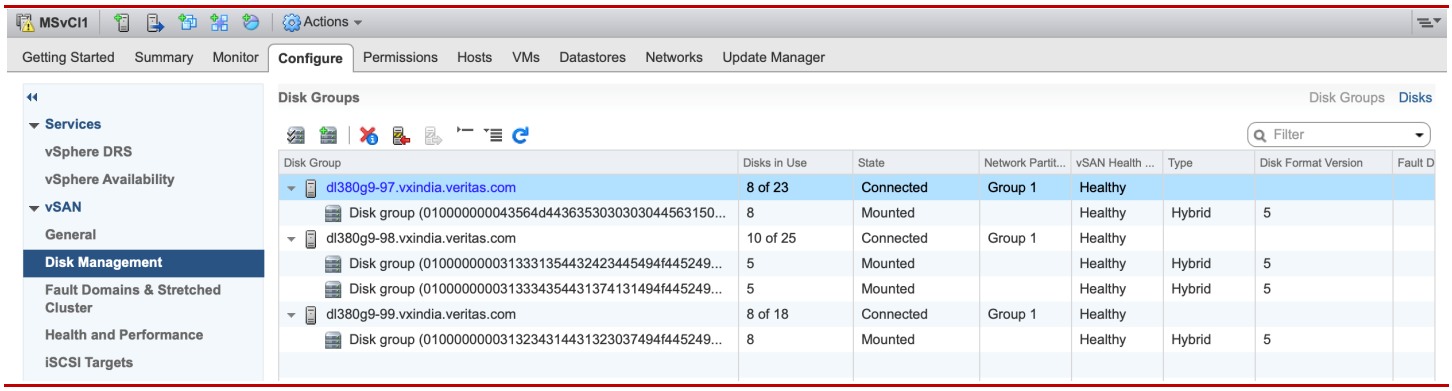
While supporting VMware features that require shared storage, such as HA, vMotion, and DRS, vSAN eliminates the need for external shared storage and simplifies storage configuration and virtual machine provisioning activities.

For details, refer to the VMware documentation at:

https://docs.vmware.com/en/VMware-vSphere/6.7/com.vmware.vsphere.virtualsan.doc/GUID-B2847022-281A-458F-80D7-E936F75E160B.html

The following screenshot depicts the vSAN configuration examples for the disk group after its successful creation and the vSAN cluster network configuration:





For instructions to create VMDK disks with the Multi-Writer flag enabled, refer to the VMware knowledge base article at:

https://kb.vmware.com/s/article/2121181

The following steps are performed to create the VMDK disks in this technical paper:

Creating eager-zeroed disks using the ESXi command line

1. Navigate to the directory of the first virtual machine in the Oracle RAC cluster:

```
cd /vmfs/volumes/vsan_datastore/VM_Name

For example:
cd /vmfs/volumes/vsanDatastore/RAC_0
```

2. Create an eager-zeroed thick (EZT) virtual disk to be shared using `vmkfstools`:

```
vmkfstools -c size -W vsan -d eagerzeroedthick `pwd`/vmdk_ile_name

For example:
vmkfstools -c 12G -W vsan -d eagerzeroedthick `pwd`/RAC_0_1.vmdk
```

3. (Optional) As EZT disks created on vSAN are not zeroed automatically, you must use the `vmkfstools -w <path_to_vmdk>` command to zero out all the blocks, if zeroing is required. Be aware of the additional I/O workload on vSAN during zeroing.

```
vmkfstools -w `pwd`/vmdk_file_name
For Example:
vmkfstools -w `pwd`/RAC_0_1.vmdk
```

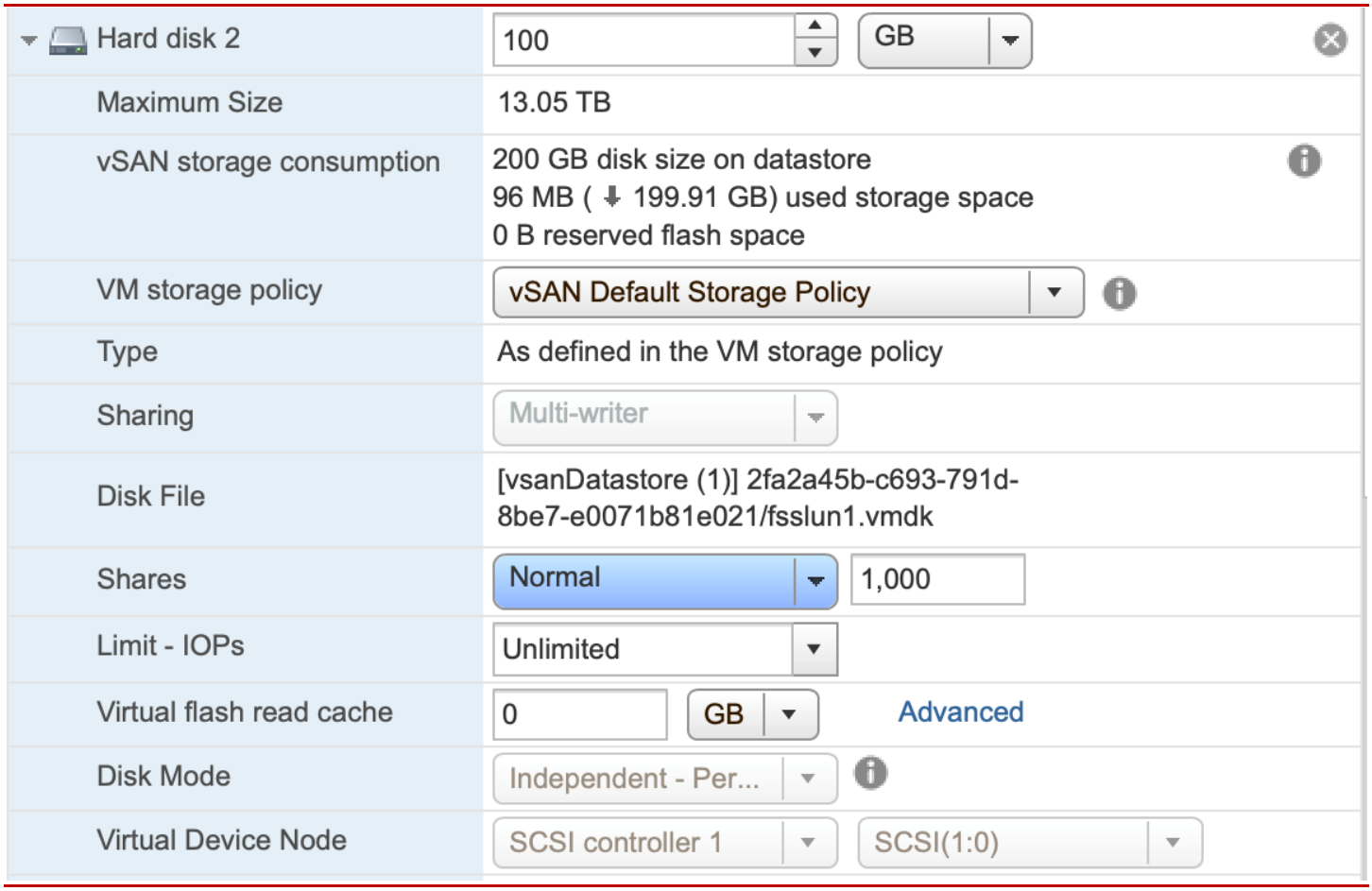
4. Repeat step 2 for each shared disk that needs to be created.
5. After all the VMDKs are created, use the vSphere web client to add VMDKs as shared disks to each of virtual machine that will run InfoScale Enterprise.

Using the vSphere web client

To add shared disks to one or more virtual machines using the vSphere Web Client

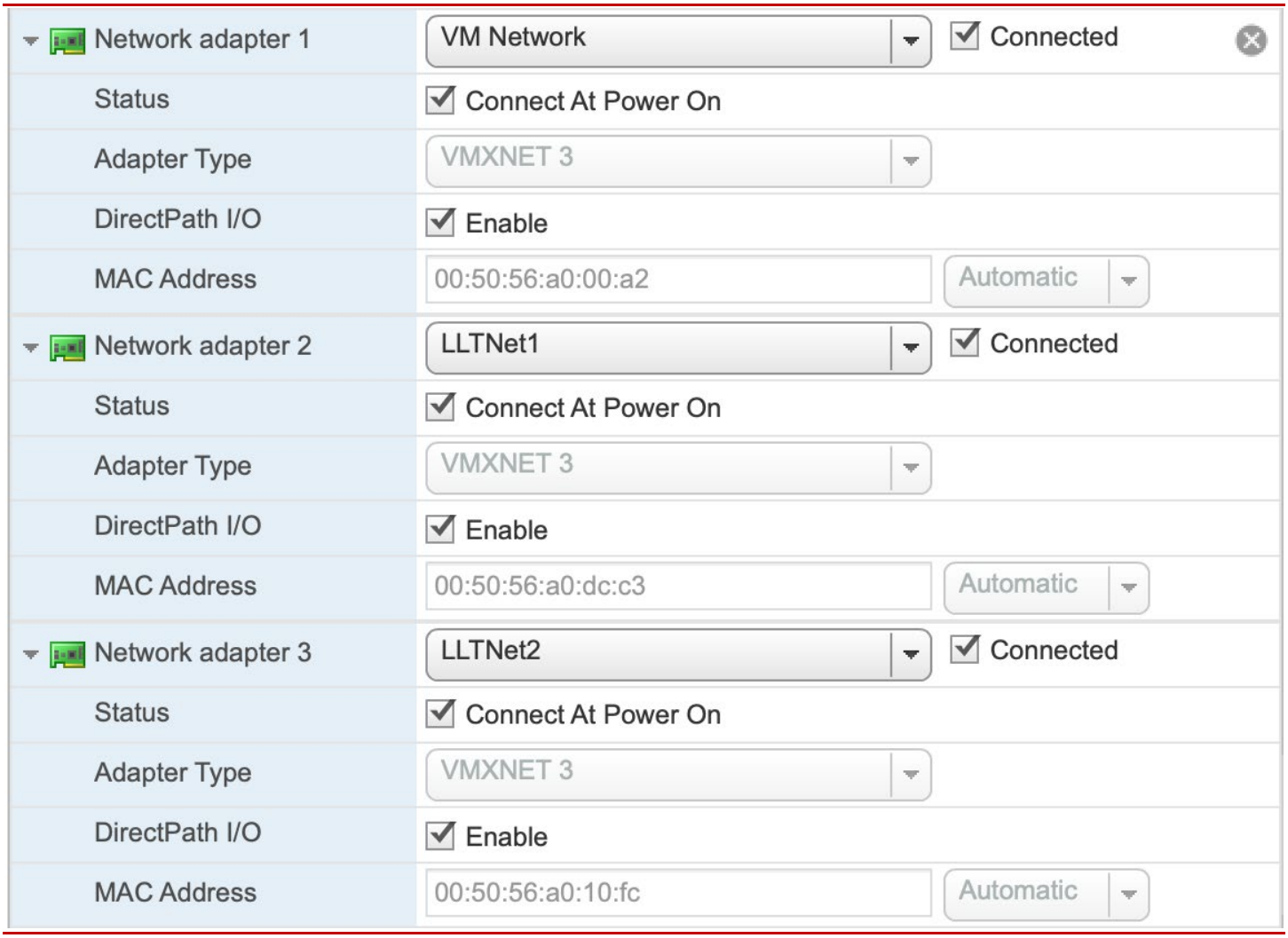
1. Right-click the appropriate virtual machine and select **Edit Settings**.
2. Select **Existing Hard Disk** from the **New Device** drop-down menu, and then click **Add**.
3. Navigate to the applicable directory and select the disk; then click **OK**.
4. Expand the **New Hard Disk** entry and modify the **Virtual Device Node**, as appropriate.
5. In the **Sharing** drop-down menu, select the **Multi-writer** option.
6. Change the **Disk Mode** to **Independent-Persistent**.

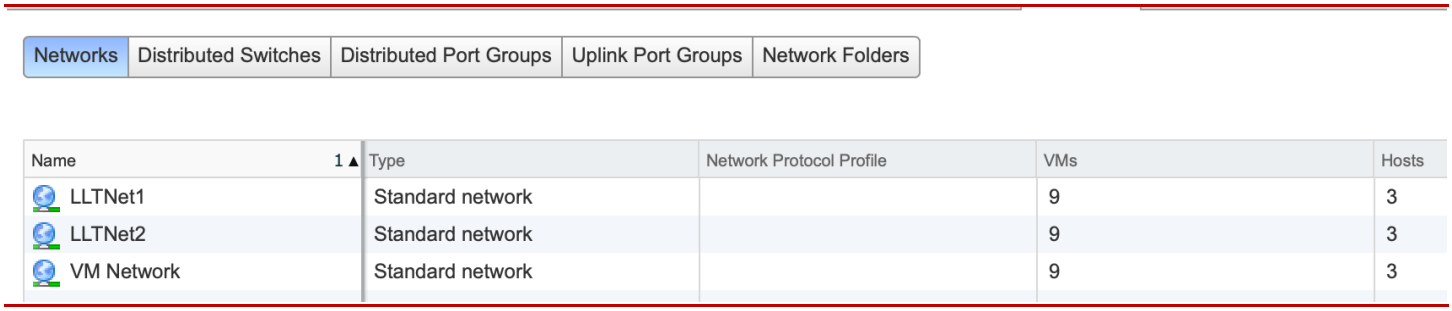
The following screenshot depicts the appropriate settings for each VMDK disk to be done while adding it to the virtual machine:



Creating networking infrastructure in VMware vSphere

To assign public and private vNICs to a virtual machine, you must first create the respective vSwitches and then create the vNIC under the corresponding vSwitches as depicted in following images:





For more details on network management, refer to VMware documentation at:

https://docs.vmware.com/en/VMware-vSphere/6.7/com.vmware.vsphere.networking.doc/GUID-35B40B0B-0C13-43B2-BC85-18C9C91BE2D4.html

Creating the InfoScale Enterprise cluster

Use the Common Product Installer (CPI) to install InfoScale Enterprise on the systems that you want to configure as the InfoScale cluster nodes. For details on InfoScale installation, refer to the *Veritas InfoScale Installation Guide*.

After the successful installation and configuration of InfoScale Enterprise and the application, you can obtain the details about all the service groups, including the application service groups.

You can also use Veritas Information Operation Manager to manage the InfoScale system and to perform the supported operations. For details, refer to the InfoScale Operations Manager documentation at:
https://sort.veritas.com/documents/doc_details/vom/7.4/Windows%20and%20UNIX/Documentation/

During the installation, the installer displays an option to select the InfoScale product:

```
    Veritas InfoScale Storage and Availability Solutions 7.4.1 Install Program
          swv279 swv280 swv281
     1)  Veritas InfoScale Foundation
     2)  Veritas InfoScale Availability
     3)  Veritas InfoScale Storage
     4)  Veritas InfoScale Enterprise
     b)  Back to previous menu
Select a product to install: [1-4,b,q,?] 4
Would you like to configure InfoScale Enterprise after installation? [y,n,q] (n) y
```

After successful installation and configuration, you can check the cluster heartbeat status as shown in the following snippet:

```
swv280.vxindia.veritas.com:/root>lltstat -nvv active
LLT node information:
    Node                 State    Link  Status  Address
     0 swv279            OPEN     ens224   UP      00:50:56:A0:DC:C3
                                  ens256   UP      00:50:56:A0:10:FC
   * 1 swv280            OPEN
                                  ens224   UP      00:50:56:A0:3B:88
                                  ens256   UP      00:50:56:A0:8D:18
     2 swv281            OPEN
                                  ens224   UP      00:50:56:A0:51:F3
                                  ens256   UP      00:50:56:A0:F1:35
swv280.vxindia.veritas.com:/root>
```

You can also check the Global Atomic Broadcast (GAB) status of the cluster, which may appear similar to the following snippet:

```
swv280.vxindia.veritas.com:/root>gabconfig -a
GAB Port Memberships
===============================================================
Port a gen   19fa02 membership 012
Port b gen   19fa14 membership 012
Port d gen   19fa01 membership 012
Port f gen   19fa22 membership 012
Port h gen   19fa26 membership 012
Port m gen   19fa1a membership 012
Port u gen   19fa20 membership 012
Port v gen   19fa1c membership 012
Port w gen   19fa1e membership 012
Port y gen   19fa1b membership 012
swv280.vxindia.veritas.com:/root>
```

When you check the VCS fencing configuration status, the details may look similar to the following:

```
swv280.vxindia.veritas.com:/root>vxfenadm -d
I/O Fencing Cluster Information:
================================
 Fencing Protocol Version: 201
 Fencing Mode: Customized
 Fencing Mechanism: cps
 Cluster Members:
     0 (swv279)
   * 1 (swv280)
     2 (swv281)
 RFSM State Information:
   node   0 in state  8 (running)
   node   1 in state  8 (running)
   node   2 in state  8 (running)
swv280.vxindia.veritas.com:/root>vxfenconfig -L
```

```
I/O Fencing Configuration Information:
======================================
 Single-CP-Flag            : 0
 Server-Count              : 3
 Security-Flag             : 0
 FIPS-mode                 : 0
 Host-Name                 : 10.209.84.163
Port                       : 443
 UUID                      : {418fc89c-78ef-11e7-9e7b-cc07b571e8f4}
 Host-Name                 : 10.209.84.164
 Port                      : 443
 UUID                      : {e706acaa-78ef-11e7-b956-56055249f9bc}
 Host-Name                 : 10.209.84.165
 Port                      : 443
 UUID                      : {f2115974-78ef-11e7-878f-e2c2cd55d365}
swv280.vxindia.veritas.com:/root>
```

One of the use cases describes the configuration of the VMware disk agent. This agent enables vMotion and VMware Distributed Resource Scheduler (DRS) in InfoScale Enterprise clusters that are configured and deployed on virtual machines in VMware environments. For details, refer to the Veritas documentation at:

https://www.veritas.com/content/support/en_US/doc/79620650-79620654-0/v87705353-79620654

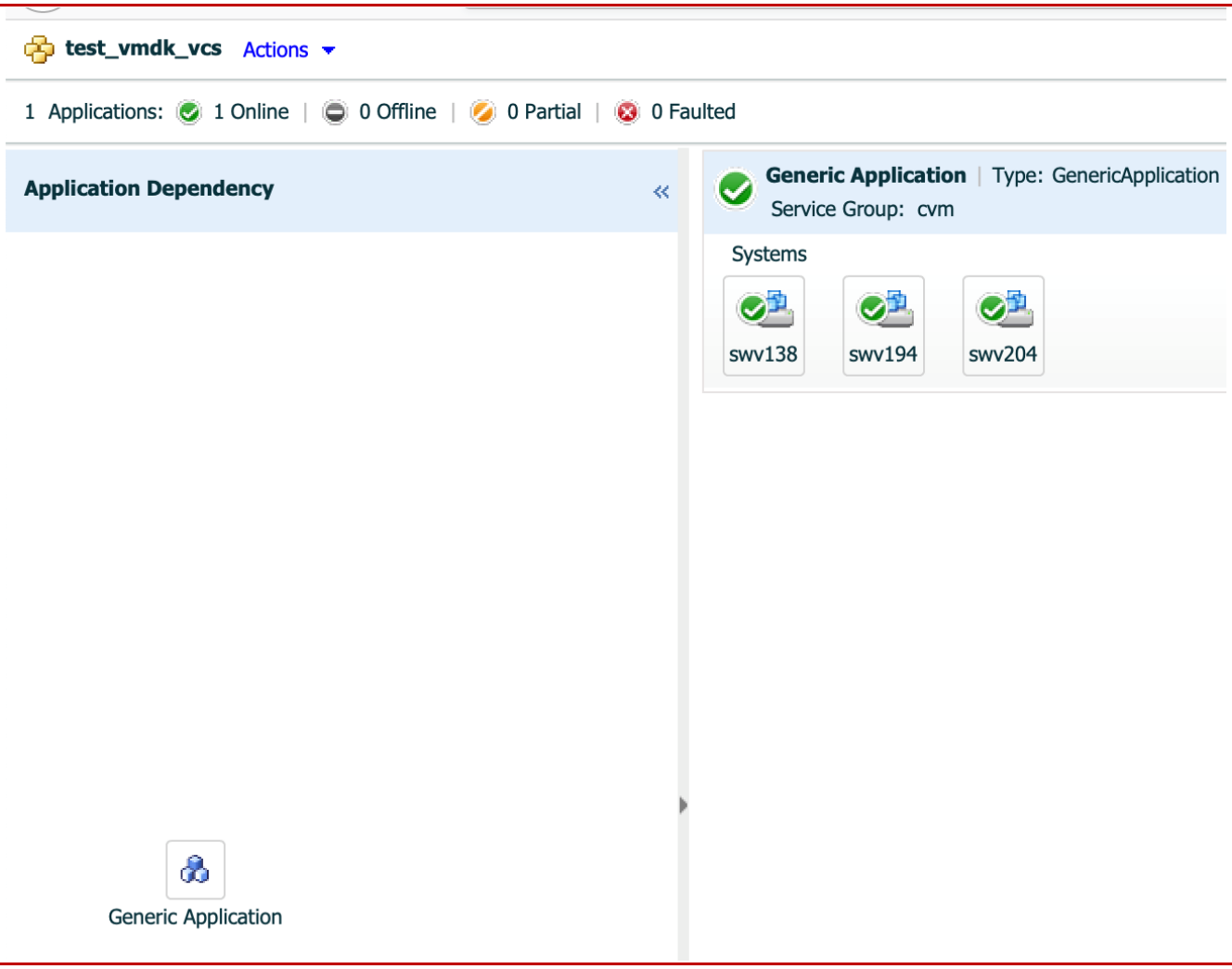
You can configure the VCS VMwareDisk agent by using the VCS Application Configuration wizard as depicted in the following screenshots. After the InfoScale installation and configuration is complete and after the required application (for example, Oracle) is configured on one of the cluster nodes, start the Application Configuration wizard by using the following URL:

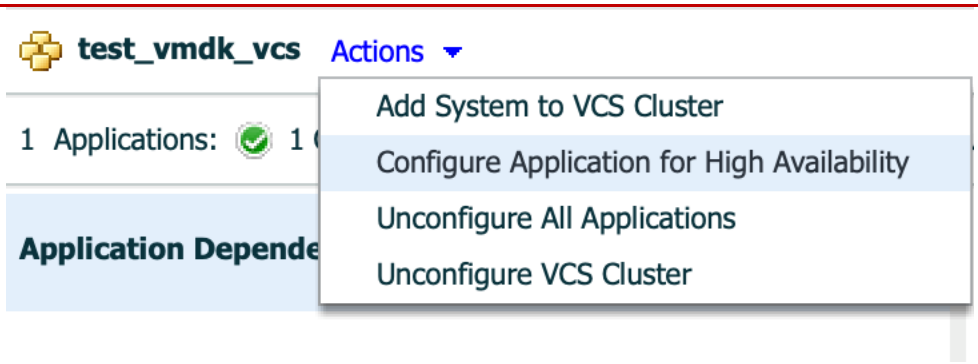
`https://<virtual_machine_IP_or_hostname>:5634/vcs/admin/application_health.html`

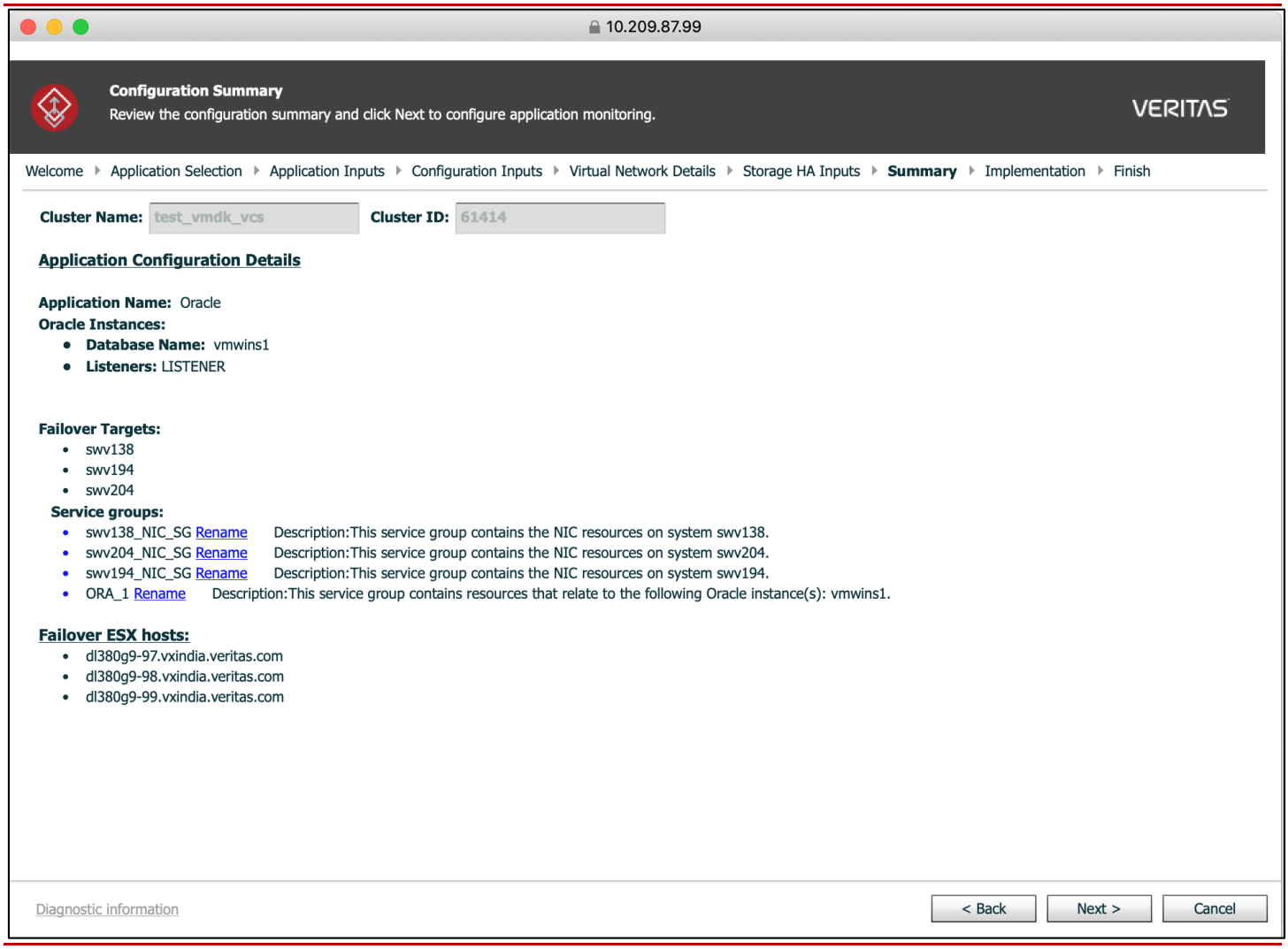
Follow the on-screen instructions. For details, refer to the following Veritas documentation:

https://www.veritas.com/content/support/en_US/doc/79630152-79630229-0/v101600189-79630229

The following screenshots depict how to configure of the VMware Disk agent and the Oracle agent using the VCS Cluster Configuration wizard:



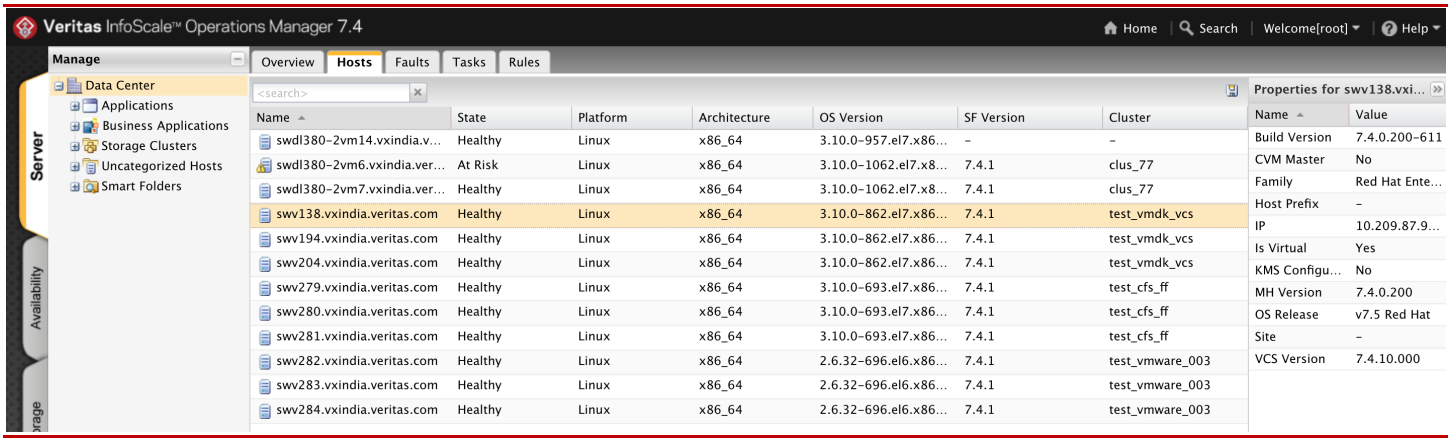



10.209.87.99
Configuration Summary
Review the configuration summary and click Next to configure application monitoring.
VERITAS
Welcome ▸ Application Selection ▸ Application Inputs ▸ Configuration Inputs ▸ Virtual Network Details ▸ Storage HA Inputs ▸ Summary ▸ Implementation ▸ Finish
Cluster Name: test_vmdk_vcs
Cluster ID: 61414
Application Configuration Details
Application Name: Oracle
Oracle Instances:
Database Name: vmwins1
Listeners: LISTENER
Failover Targets:
swv138
swv194
swv204
Service groups:
swv138_NIC_SG Rename Description:This service group contains the NIC resources on system swv138.
swv204_NIC_SG Rename Description:This service group contains the NIC resources on system swv204.
swv194_NIC_SG Rename Description:This service group contains the NIC resources on system swv194.
ORA_1 Rename Description:This service group contains resources that relate to the following Oracle instance(s): vmwins1.
Failover ESX hosts:
dl380g9-97.vxindia.veritas.com
dl380g9-98.vxindia.veritas.com
dl380g9-99.vxindia.veritas.com
Diagnostic information
< Back
Next >
Cancel

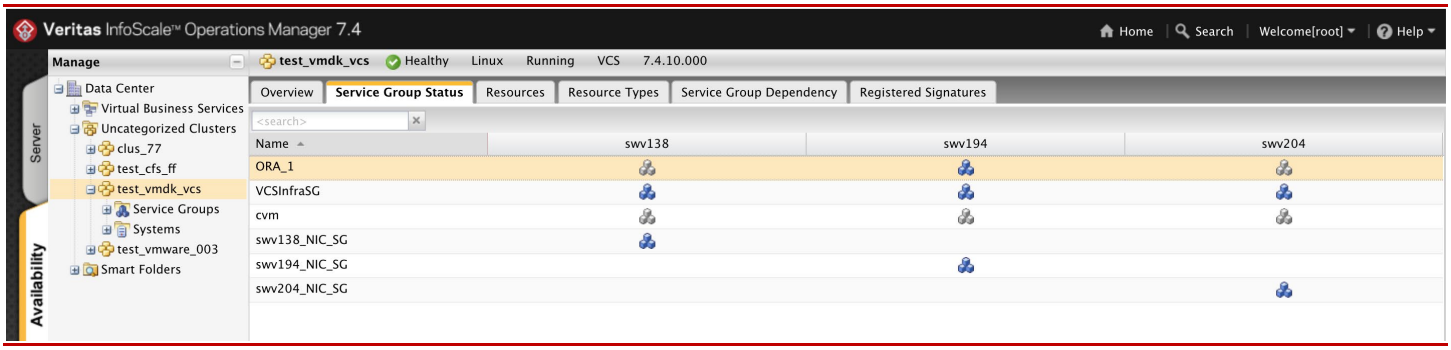
After you successfully configure the product and the application, you can also check the status of the entire cluster. The following snippet provides a sample:

```
[root@swv138 ~]# hastatus -summ
-- SYSTEM STATE
-- System               State                Frozen
A  swv138               RUNNING              0
A  swv194               RUNNING              0
A  swv204               RUNNING              0
-- GROUP STATE
-- Group           System               Probed     AutoDisabled    State
B  ORA_1           swv138               Y          N               OFFLINE
B  ORA_1           swv194               Y          N               ONLINE
B  ORA_1           swv204               Y          N               OFFLINE
B  VCSInfraSG      swv138               Y          N               ONLINE
B  VCSInfraSG      swv194               Y          N               ONLINE
B  VCSInfraSG      swv204               Y          N               ONLINE
B  cvm             swv138               Y          N               OFFLINE
B  cvm             swv194               Y          N               OFFLINE
B  cvm             swv204               Y          N               OFFLINE
B  swv138_NIC_SG   swv138               Y          N               ONLINE
B  swv194_NIC_SG   swv194               Y          N               ONLINE
B  swv204_NIC_SG   swv204               Y          N               ONLINE
[root@swv138 ~]#
```

If you use InfoScale Operations Manager for the configuration, the interface displays the host information—similar to the following screenshot—after the full discovery of its managed hosts is complete:



You can use InfoScale Operations Manager to monitor the system and to perform various operations for the application service groups, like online, offline, and switch. InfoScale Operations Manager displays the cluster information as follows:



Test cases

The following tests were run successfully to qualify a VMware Guest cluster with InfoScale Enterprise. These tests demonstrate that InfoScale Enterprise runs seamlessly on VMware virtual machines and the assigned disks that are backed by the vSAN datastore.

Scenario	Remark
Installing and configuring InfoScale Enterprise (SFCFSHA) 7.4.1 Update 1 on a 3-node Guest cluster.	Install and configure InfoScale Enterprise using the CPI.
Exporting vDisks using MWF to all the three guests in the Sharing mode.	Establish iSCSI connections using the `iscsiadm` command options to assign vDisks.
Configuring disk-based non-SCSI3-PR server-based fencing for node arbitration.	Configure disk-based SCSI3 fencing using the CPI.
Configuring the VCS VMwareDisk agent to configure HA for the VMDK disks that are assigned to a single virtual machine in the InfoScale cluster.	Configure the VCS VMwareDisk agent using the VCS Application Configuration wizard along with the Application agent (for example, Oracle).
Installing and configuring the Oracle database on the CFS mounts and the VxFS mounts.	To test application-based HA, use the Oracle aatabase for qualification testing.
Configure the VCS agent for Oracle in the fast failover mode.	To reduce the duration for failover and failback, configure the VCS Oracle agent in the recommended mode.
Test the failover and failback of VCS Oracle agent.	Test the failover and failback of Oracle service while the Database workload is running.

Test network split-brain.	Simulate the VCS cluster failure by disabling the LLT (cluster interconnects) and verify that the Database service fails over to an available node.
Test virtual machine live migration.	Test the live migration of a virtual machine that is part of the InfoScale cluster without affecting the cluster failure.
Test ESXi hypervisor node failure.	Configure VMware–HA and test the hypervisor node failure hosting virtual machine where the application is online. Ensure that VMware–HA starts the failed virtual machine on another hypervisor host, the VMDK disks are available, and the virtual machine joins the VCS cluster.

References

https://docs.vmware.com/en/VMware-vSphere/6.7/com.vmware.vsphere.vsan-planning.doc/GUID-ACC10393-47F6-4C5A-85FC-88051C1806A0.html

ABOUT VERITAS TECHNOLOGIES LLC

Veritas Technologies empowers businesses of all sizes to discover the truth in information—their most important digital asset. Using the Veritas platform, customers can accelerate their digital transformation and solve pressing IT and business challenges including multi-cloud data management, data protection, storage optimization, compliance readiness and workload portability—with no cloud vendor lock-in. Eighty-six percent of Fortune 500 companies rely on Veritas today to reveal data insights that drive competitive advantage. Learn more at veritas.com or follow us on Twitter at @veritastechllc.

Veritas World Headquarters
2625 Augustine Drive
Santa Clara, CA 95054
+1 (866) 837 4827
www.veritas.com

For specific country offices
and contact numbers,
please visit our website.

VERITAS
The truth in information.