



P r o f e s s i o n a l E x p e r t i s e D i s t i l l e d

Mastering VMware vSphere Storage

Monitor and optimize the storage capabilities of your vSphere environment

Victor Wu

Eagle Huang

[PACKT] enterprise 
PUBLISHING professional expertise distilled

www.it-ebooks.info

Mastering VMware vSphere Storage

Monitor and optimize the storage capabilities of your vSphere environment

Victor Wu

Eagle Huang



BIRMINGHAM - MUMBAI

Mastering VMware vSphere Storage

Copyright © 2015 Packt Publishing

All rights reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, without the prior written permission of the publisher, except in the case of brief quotations embedded in critical articles or reviews.

Every effort has been made in the preparation of this book to ensure the accuracy of the information presented. However, the information contained in this book is sold without warranty, either express or implied. Neither the authors, nor Packt Publishing, and its dealers and distributors will be held liable for any damages caused or alleged to be caused directly or indirectly by this book.

Packt Publishing has endeavored to provide trademark information about all of the companies and products mentioned in this book by the appropriate use of capitals. However, Packt Publishing cannot guarantee the accuracy of this information.

First published: July 2015

Production reference: 1290715

Published by Packt Publishing Ltd.
Livery Place
35 Livery Street
Birmingham B3 2PB, UK.

ISBN 978-1-78217-485-1

www.packtpub.com

Credits

Authors

Victor Wu
Eagle Huang

Reviewers

Renjith V C
Anthony Fortunato
Jan Gerrit Kootstra
Jason Langer

Commissioning Editor

Dipika Gaonkar

Acquisition Editor

Meeta Rajani

Content Development Editor

Arwa Manasawala

Technical Editors

Ryan Kochery
Tanmayee Patil

Copy Editor

Vikrant Phadke

Project Coordinator

Shweta Birwatkar

Proofreader

Safis Editing

Indexer

Rekha Nair

Production Coordinator

Melwyn D'sa

Cover Work

Melwyn D'sa

About the Authors

Victor Wu has over 10 years of IT experience. He currently works as a solution architect at BoardWare Information System in Macau. He is a team leader of the system deployment team at his current company and is responsible for storage implementation, architecture, upgrades, and migration (EMC Clariion/VNX, HP 3PAR StoreServ 7200/7400, HP P-Series, IBM DS-Series, and more).

Victor Wu has good experience in virtualization solutions. These include VMware vSphere/View, Microsoft Hyper-V, Novell PlateSpin, Double-Take, Citrix XenServer, Citrix XenApp, and Citrix XenDesktop. He is interested in some deployments of virtualization solutions and troubleshooting, such as VMware version upgrades, storage data migration, and so on.

His professional qualifications include EMCIE, EMCPE, EMCTA, vExpert 2014/2015, VCAP5-DCD, VCAP4/5-DCA, VCP5-DT, VCP-Cloud, VCP-NT, VCP3/4/5, CDCUCSS, CDCUCDS, CCA, MCITP, and MCP.

Eagle Huang has over 10 years of IT experience. He currently works as a technical instructor of the virtualization technology for one of the world's leading technology companies in Shanghai. Since 2006, he has been to several cities in China. There, he has carried out numerous training programs for IBM, Cisco, Dell, and a number of original factories in NetApp special agencies, leading several phases and customized training projects. In his career, he has been responsible for many projects and has carried out many virtualization projects. Huang has served industries such as finance, securities, health care, energy, and so on. He is good at project design and project planning in line with industry standards. Also, according to the use of virtualization architecture, he has built optimal configurations.

His professional qualifications include CCNA, VCI, CCI, CCEE, MCP, VCAP5-DCD, VCAP4/5-DCA, VCP5-DT, VCP-Cloud, VCP-NT, VCP3/4/5, CCA, and MCITP.

About the Reviewers

Renjith V C is a senior system administrator with over 11 years of experience in enterprise system administration. He is currently working for an international health care group in Dubai, UAE. Before he decided to move to Dubai back in 2008, he was working for a top MNC in India. He holds several certifications from most vendors.

At work, on a daily basis, he deals with virtualization technologies, backup and replication, most Microsoft Server roles, and so on. He likes tech blogging and learning new technologies.

Renjith started his blogging journey with www.hackstacks.com back in 2007 and has recently started his personal blog as well (<http://www.renjithmenon.com/>). You can follow him on Twitter at @vcrenjith.

Anthony Fortunato is a technical infrastructure center manager for the Pima County Government in Arizona, USA. He leverages his 20 years of experience in data center infrastructure, virtualization, and information security leadership to meet the ever-changing technological needs of the public sector. He delivers expert guidance on infrastructure implementation and Agile methodologies with his in-depth knowledge of system architecture, policy development, and operational overhead, shaping the future of Pima County.

Anthony has been involved in VMware virtualization using VMware's vCloud Enterprise platform to construct a software-defined data center, allowing the adoption of agile system deployment on demand. He leads the implementation of server and storage hardware platform standardization to allow Pima County to formally observe hardware life cycle management for system replacements.

Additionally, he serves as an information security program adjunct lecturer for the Department of Management Information Systems at the Eller College of Management at the University of Arizona. Anthony has developed and taught graduate and senior-level undergraduate coursework in systems security and information assurance. He uses VMware as a platform to engage students and provide real-world experiences with virtualization and information security. This information security program is designated by the National Information Assurance Education and Training Program (NIETP) office, under the scrutiny of the United States National Security Agency as a National Center for Academic Excellence in Information Assurance Education. U.S. News & World Report ranked the program in the top five among America's Best Colleges 2015 and America's Best Graduate Schools 2016.

Before coming to Pima County, Anthony worked as a leader for the IT department of the Eller College of Management, where he gained valuable experience by implementing VMware virtualization technology, leading security initiatives, and building resiliency for the organization.

I would like to take this opportunity to thank my beautiful wife, Claudia, for her unwavering love and support. I would not have been the person I am today without you. To my daughter, Rory: thank you for bringing so much joy to our lives. I love you both with all of my heart.

Jan Gerrit Kootstra studied mathematics and became an engineer in applied mathematics from the University of Groningen in 1993. He works as a Linux and Unix system administrator. His major hobby is traveling all over the world with his wife and two children.

His employer is a major Dutch telecommunication company with a large IT subsidiary.

He has not worked on any books yet, but he is an author of a scientific article called *H.W. Hoogstraten, J.G. Kootstra, B. Hillen, J.K.B. Krijger, and P.J.W. Wensing: Numerical simulation of blood flow in an artery with two successive bends*, *Journal of Biomechanics*, Volume 29 (8) 1996 1075-1083. It is related to his master's thesis.

I want to thank my wife, children, colleagues, and friends for their support and knowledge sharing.

Jason Langer works as a solutions architect for a VMware partner in the Pacific Northwest, helping customers achieve their data center server virtualization and end user computing goals. He has obtained multiple levels of certification from both Microsoft (MCSE/MCSA) and VMware (VCP/VCAP), and he brings over 15 years of IT experience to the table. When not in his daily job, Jason is active in the VMware community, as a member of the Seattle VMUG Steering Committee, and writes content for his blog at <http://www.virtuallanger.com/>.

He was a technical reviewer on *VMware ESXi Cookbook*, *Troubleshooting vSphere Storage*, *VMware Horizon View 5.3 Design Patterns and Best Practices*, and *Getting Started with VMware Virtual SAN*, all by Packt Publishing.

www.PacktPub.com

Support files, eBooks, discount offers, and more

For support files and downloads related to your book, please visit www.PacktPub.com.

Did you know that Packt offers eBook versions of every book published, with PDF and ePub files available? You can upgrade to the eBook version at www.PacktPub.com and as a print book customer, you are entitled to a discount on the eBook copy. Get in touch with us at service@packtpub.com for more details.

At www.PacktPub.com, you can also read a collection of free technical articles, sign up for a range of free newsletters and receive exclusive discounts and offers on Packt books and eBooks.



<https://www2.packtpub.com/books/subscription/packtlib>

Do you need instant solutions to your IT questions? PacktLib is Packt's online digital book library. Here, you can search, access, and read Packt's entire library of books.

Why subscribe?

- Fully searchable across every book published by Packt
- Copy and paste, print, and bookmark content
- On demand and accessible via a web browser

Free access for Packt account holders

If you have an account with Packt at www.PacktPub.com, you can use this to access PacktLib today and view 9 entirely free books. Simply use your login credentials for immediate access.

Instant updates on new Packt books

Get notified! Find out when new books are published by following [@PacktEnterprise](#) on Twitter or the *Packt Enterprise* Facebook page.

Table of Contents

Preface	v
Chapter 1: Getting Started with vSphere 5.x and vCenter 5.x	1
Physical and virtual architecture	2
A comparison between physical and virtual machines	3
Installing the VMware ESXi host	4
Installing VMware vCenter Server	9
Prerequisites	9
Installing the VMware vCenter Server Appliance	17
Prerequisites	17
Connecting to vCenter Server with vSphere Client	21
Connecting to vCenter Server with vSphere Web Client	21
Summary	22
Chapter 2: Getting Started with vSphere Management Assistant	23
Deploying vMA	24
Configuring ESXi technical support mode	30
Enabling and accessing ESXi Shell	30
Enabling ESXi Shell from vSphere Client	32
Accessing ESXi Shell from DCUI	33
Using the VMware commands	34
Reviewing ESXi and vCenter Server logs	44
The location of the vCenter Server log	45
The location of the vSphere Server log	45
Exporting the vm-support log file from vSphere Client	46
Exporting the vm-support log file from ESXi Shell	48
Summary	50

Chapter 3: Using the Virtual Machine Monitor	51
The VMware vSphere ESXi architecture	51
Understanding the VMM	55
Software and hardware virtualization techniques	57
Using vSphere performance monitoring tools	61
Summary	70
Chapter 4: Storage Scalability	71
vSphere storage APIs for array integration and storage awareness	71
Virtual machine storage profile	74
VMware vSphere Storage DRS	87
VMware vSphere Storage I/O Control	92
Summary	95
Chapter 5: Optimizing Storage	97
Concepts of storage virtualization	97
Monitoring vSphere storage	101
vSphere storage management using the command line	105
Troubleshooting vSphere storage performance problems	111
First scenario	112
Second scenario	112
Summary	115
Chapter 6: vSphere Storage Configuration Settings	117
vSphere storage components	117
LUN masking	119
vSphere 5 storage maximums	125
Identifying the vSphere log used to troubleshoot a storage problem	126
Summary	134
Chapter 7: Analyzing vSphere Storage by CLI	135
Analyzing PSA and multipathing using esxcli	136
Applying VMFS volume copies resignaturing	142
Troubleshooting VMware snapshots and VMFS resignaturing	146
VMFS DataStore volume unmounting	148
Identifying and tagging SSD devices	149
Summary	152
Chapter 8: Troubleshooting vSphere FC Storage	153
The vSphere Fibre Channel storage component	153
A vSphere Fibre Channel storage troubleshooting example	154
vRealize displays the WorkLoad Badge Metrics	157
Summary	161

Chapter 9: Troubleshooting vSphere iSCSI Storage	163
vSphere iSCSI storage components	163
vSphere iSCSI storage troubleshooting examples	165
Summary	169
Chapter 10: Troubleshooting vSphere NFS Storage	171
vSphere NFS storage components	171
A vSphere NFS storage case study	172
A vSphere NFS storage troubleshooting example	175
Summary	177
Chapter 11: vSphere Storage Design	179
Storage design key points	179
The vSphere storage architecture	180
Getting started with vSphere storage design	181
Share storage size – how many datastore requirements fit?	182
Virtual machines per LUN	182
Datastore types – how many types fit your environment?	184
Storage I/O Control – is it needed to enable the feature?	185
A Shared I/O Control settings explanation	185
RDM – does it need to be used in your environment?	187
How to design ESXi multipathing policies	188
How to design zoning and masking	191
Summary	192
Chapter 12: ESXi Host Design	193
ESXi host design key points	193
CPU capacity	193
Number of hosts	194
Host hardware types	195
Host naming conventions	196
An ESXi host design example	196
Management cluster specifications	198
Oracle cluster specifications	198
VMware ESXi physical specifications	199
Oracle cluster VMware ESXi physical specifications	199
Summary	200
Chapter 13: Virtual Machine Design	201
Virtual machine design key points	201
Number of virtual CPUs	201
Ensuring the memory's performance	202
VM resource setting – limit, reservation, and share	203

Table of Contents

Virtual machine hard disk – how to deploy and which types?	203
Multiple virtual disks	204
Virtual disk location	204
Swap file location	204
Virtual SCSI HBA type – which one fits your OS?	205
Virtual NICs	205
Virtual machine hardware compatibility	206
Considering guest OS	207
A virtual machine design example	207
Summary	208
Chapter 14: vSphere Virtual Datacenter Design	209
vSphere virtual datacenter design key points	209
VMware vCenter Server	210
Which platform do you choose?	210
Deploying on a physical server or a virtual machine?	210
How to deploy VMware vCenter Server DB	211
vSphere clusters	211
Number of vSphere clusters	211
vSphere HA	213
vSphere FT	215
vSphere DRS	215
Convention for naming a host	218
A vSphere virtual datacenter design example	219
Summary	220
Index	221

Preface

This book is for users who already have some experience with the VMware vSphere platform. You will want to learn and design VMware vSphere storage solutions and how to troubleshoot vSphere storage issues. Also, you will want to know how to design ESXi hosts, virtual machines, and virtual data centers in a better way.

What this book covers

Chapter 1, Getting Started with vSphere 5.x and vCenter 5.x, shows you the difference between physical and virtual hosts and the benefits of using a virtualized host, and describes how to set up VMware ESXi Server and VMware vCenter Server.

Chapter 2, Getting Started with vSphere Management Assistant, explains what VMware vSphere Management Assistant is, how to set up and configure it with vSphere Server, and how to enable ESXi TSM during troubleshooting.

Chapter 3, Using the Virtual Machine Monitor, introduces the components of the VMware vSphere ESXi architecture and the VMM. You will also learn the techniques to configure software and hardware with VMM in the virtual machine.

Chapter 4, Storage Scalability, explains the vSphere Storage API for Array Integration (VAAI) and Storage Awareness (VASA). In this chapter, you also learn about the vSphere administrator and how to configure the virtual machine storage profile and assign it to the ESXi data store.

Chapter 5, Optimizing Storage, teaches you the concept of storage virtualization and how to monitor the storage metric using vSphere performance charts and esxtop/resxtop. You will also learn about the management of vSphere storage using the command line.

Chapter 6, vSphere Storage Configuration Settings, discusses what a storage component is, for example, the LUN name, identifier name, and runtime name. You also learn how to set up LUN masking on a vSphere host using `esxcli` commands, and setting up vSphere Syslog Collector to collect the ESXi host's log.

Chapter 7, Analyzing vSphere Storage by CLI, provides more commands for managing vSphere storage. Here, you also read about the difference between the existing VMFS signature and resignaturing.

Chapter 8, Troubleshooting vSphere FC Storage, focuses on troubleshooting vSphere FC storage, and explains the FC storage infrastructure. This chapter lists some examples that show you how to troubleshoot.

Chapter 9, Troubleshooting vSphere iSCSI Storage, is concerned with analyzing the vSphere iSCSI storage infrastructure, and lists some examples that show how to troubleshoot.

Chapter 10, Troubleshooting vSphere NFS Storage, describes some examples that tell you how to carry out NFS troubleshooting.

Chapter 11, vSphere Storage Design, tells the story of the vSphere storage architecture design and its main point. This chapter also illustrates the design method.

Chapter 12, ESXi Host Design, contains the story of the key points of the ESXi host design, according to the actual environment, for example, how to design the host.

Chapter 13, Virtual Machine Design, analyzes the main points of the virtual machine design in detail. At the same time, according to the actual environment, we see, for example, how to design the virtual machine.

Chapter 14, vSphere Virtual Datacenter Design, describes in detail the main points of the virtual data center design, and as an example, we see how to design a virtual data center.

What you need for this book

This book requires the following software: VMware vSphere 5, VMware vCenter Server 5, VMware vSphere Management Assistant 5 (vMA), and vCenter Operation Manager.

Who this book is for

This book is for users such as data center administrators, people at IT help desks, and IT architects who have already worked with the VMware vSphere platform and who want to design a VMware vSphere Storage solution and troubleshoot issues.

Conventions

In this book, you will find a number of text styles that distinguish between different kinds of information. Here are some examples of these styles and an explanation of their meaning.

Code words in text, database table names, folder names, filenames, file extensions, pathnames, dummy URLs, user input, and Twitter handles are shown as follows: "Open a web browser and enter this URL: `https://<vcenter FQDN or vcenter IP address>: 9443/vsphere-client/` (port 9443 is the default)."

Any command-line input or output is written as follows:

```
esxstop -a -b > "output-name.csv"
```

New terms and **important words** are shown in bold. Words that you see on the screen, for example, in menus or dialog boxes, appear in the text like this: "To use the bundled database, click on **Install a Microsoft SQL Server 2008 Express instance.**"



Warnings or important notes appear in a box like this.



Tips and tricks appear like this.

Reader feedback

Feedback from our readers is always welcome. Let us know what you think about this book—what you liked or disliked. Reader feedback is important for us as it helps us develop titles that you will really get the most out of.

To send us general feedback, simply e-mail feedback@packtpub.com, and mention the book's title in the subject of your message.

If there is a topic that you have expertise in and you are interested in either writing or contributing to a book, see our author guide at www.packtpub.com/authors.

Customer support

Now that you are the proud owner of a Packt book, we have a number of things to help you to get the most from your purchase.

Downloading the color images of this book

We also provide you with a PDF file that has color images of the screenshots/diagrams used in this book. The color images will help you better understand the changes in the output. You can download this file from: https://www.packtpub.com/sites/default/files/downloads/B04074_48510S_Graphics.pdf.

Errata

Although we have taken every care to ensure the accuracy of our content, mistakes do happen. If you find a mistake in one of our books – maybe a mistake in the text or the code – we would be grateful if you could report this to us. By doing so, you can save other readers from frustration and help us improve subsequent versions of this book. If you find any errata, please report them by visiting <http://www.packtpub.com/submit-errata>, selecting your book, clicking on the **Errata Submission Form** link, and entering the details of your errata. Once your errata are verified, your submission will be accepted and the errata will be uploaded to our website or added to any list of existing errata under the Errata section of that title.

To view the previously submitted errata, go to <https://www.packtpub.com/books/content/support> and enter the name of the book in the search field. The required information will appear under the **Errata** section.

Piracy

Piracy of copyrighted material on the Internet is an ongoing problem across all media. At Packt, we take the protection of our copyright and licenses very seriously. If you come across any illegal copies of our works in any form on the Internet, please provide us with the location address or website name immediately so that we can pursue a remedy.

Please contact us at copyright@packtpub.com with a link to the suspected pirated material.

We appreciate your help in protecting our authors and our ability to bring you valuable content.

Questions

If you have a problem with any aspect of this book, you can contact us at questions@packtpub.com, and we will do our best to address the problem.

1

Getting Started with vSphere 5.x and vCenter 5.x

This chapter describes the basic concepts of virtualization and introduces VMware vSphere ESXi 5 and VMware vCenter 5. It also covers how to connect to vCenter Server by VMware vSphere Client and configure the ESXi settings.

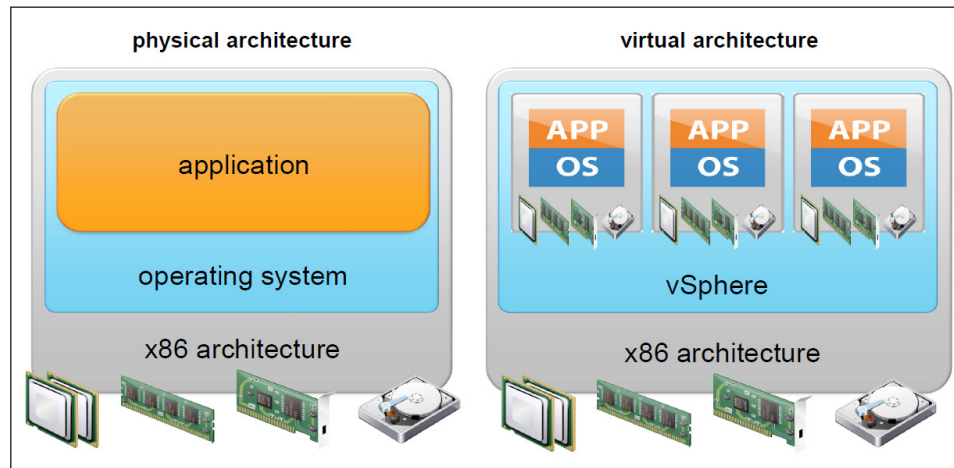
VMware vCenter Server allows you to centrally manage many VMware ESXi hosts and their virtual machines. For large deployments, you can deploy virtual machines from the VM template, which can result in reduced system administrator operations.

In this chapter, you will learn the following topics:

- Comparing physical and virtual architecture
- Deploying a VMware ESXi host
- Deploying a VMware vCenter Server
- Deploying a VMware vCenter Server Appliance
- Connecting to VMware vCenter Server by vSphere Client
- Connecting to VMware vCenter Server by vSphere Web Client

Physical and virtual architecture

Before learning about the VMware vSphere Server, you should know what the difference between traditional physical architecture and virtual architecture is.



The preceding diagram describes the differences between a virtualized and a non-virtualized host. In traditional architectures, the operating system is directly installed on hardware devices, for example, a rack-mount server, a blade server, and so on. The operating system, which is a Microsoft Windows platform or Linux platform, can only allocate the physical CPU and memory resources. It sends and receives data on a physical network adapter. It is required to upgrade the hardware if the administrator wants to allocate more physical resources, for example, the CPU core, memory, number of **Host Bus Adapters (HBAs)** and network adapters, and so on. Also, it is required to schedule the service down during hardware upgrades on each ESX host.

In virtual architectures, the operation system is installed on the hardware through a thin layer of software, called the virtualization layer or hypervisor. VMware vSphere is a hypervisor that can dynamically allocate physical hardware resources to each virtual machine. For example, suppose that a vSphere Server has 64 GB memory and three virtual machines are running on this vSphere host. Each VM is allocated 4 GB, which shares the physical memory (64 GB) of that vSphere host. It is not required to upgrade the hardware if the administrator wants to allocate more memory resources to the virtual machine, because the vSphere still has 52 GB of free memory available. To sum up, virtual architecture is more flexible than traditional architecture.

A comparison between physical and virtual machines

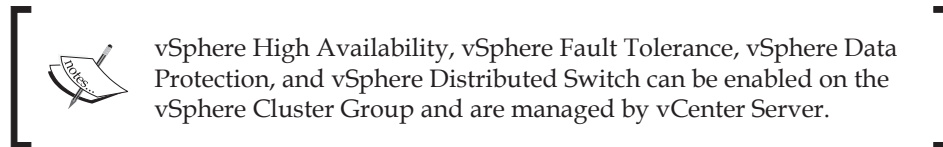
The following table is a comparison of physical and virtual machines:

	Physical machine	Virtual machine
Relocation	Difficult to relocate. Requires service downtime. Specific to physical hardware.	Easy to relocate. Encapsulated into files. Independent of physical hardware.
Management	Difficult to manage. Hardware failures cause service downtime.	Easy to manage. Isolated from other virtual machines.
Cost	High.	Virtualizing physical system saves cost.
Hardware limitations	Changes in hardware limit application support.	Changes in hardware cannot affect application support.
Resource sharing	This is not supported.	Virtualization can share multiple virtual machines on a single physical host.
Memory usage	Operating system assumes that it owns all of the physical memory in the system.	Allow the hypervisor to run multiple virtual machines simultaneously.
Virtual networking	This is not supported.	A virtual machine can be configured with one or more virtual Ethernet adapters (vSwitch or dvSwitch).
Filesystems	NTFS and ext3.	VMFS3 and VMFS5.
Operating system deployment	This is a time-consuming task.	It is easy to deploy virtual machines from the VM template.
Backup	Use third-party backup software.	It is easy to create a VM snapshot or clone the virtual machine.

According to the preceding table, you know that virtual architecture is more flexible than traditional architecture, for example, with OS deployment, management, and relocation.

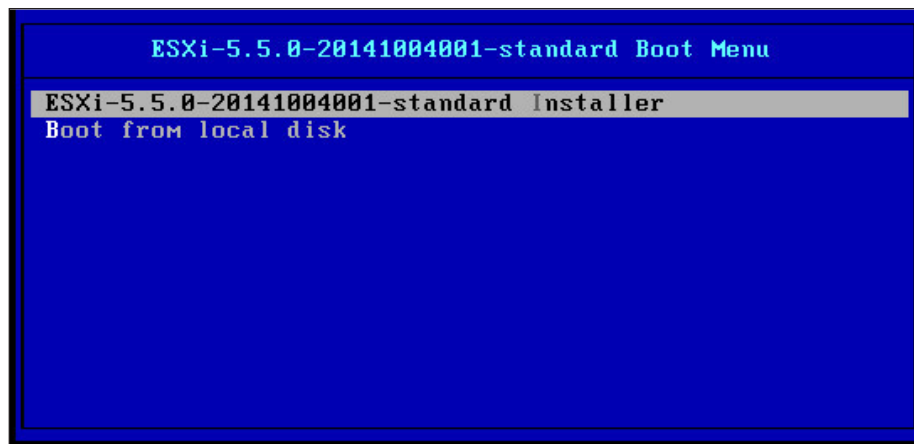
Installing the VMware ESXi host

VMware vSphere is a virtualization platform for building cloud infrastructure. It provides a high-performance virtualization layer. Multiple virtual machines can share hardware resources, for example, CPU, memory, storage, and so on. According to its licensed edition, vSphere can enable different features, such as vSphere vMotion, vSphere Storage vMotion, **vSphere High Availability (HA)**, **vSphere Fault Tolerance (FT)**, vSphere Data Protection, vSphere Distributed Switch, and so on.



The following is the procedure for installation of VMware ESXi:

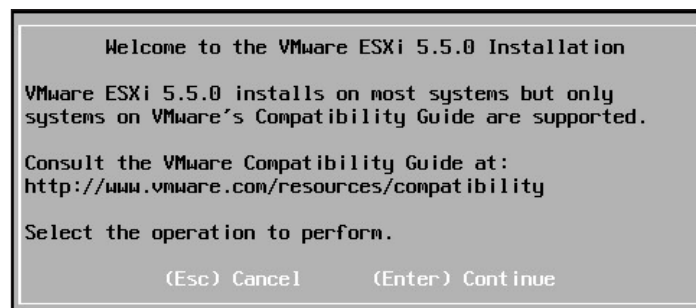
1. Download the VMware vSphere Server installer from VMware website at <https://my.vmware.com/web/vmware/downloads>.
2. Ensure that your server is configured to boot from the CD-ROM drive.
3. Then ensure that the VMware ESXi installation media are available for the server:
 - If it is a local installation, insert the VMware ESXi installation CD into the optical drive
 - If it is a remote installation, map an image of the installation media, known as an ISO image, to a virtual optical drive
4. Power up your server.




5. Press *Enter* to boot the ESXi installer, which will boot up and stop, showing a welcome message. Press *Enter* again to continue.

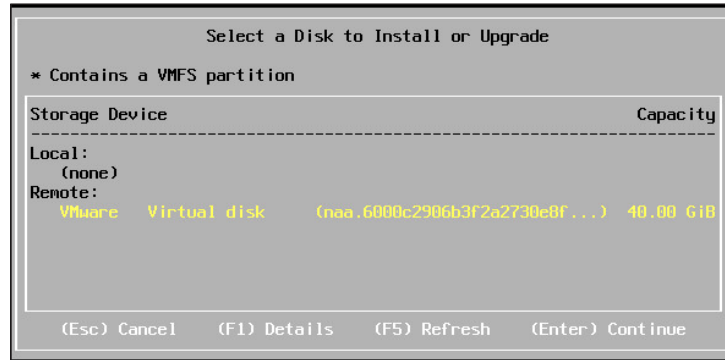


6. At the **End User License Agreement (EULA)** screen, press *F11* to accept the EULA and continue with the installation.

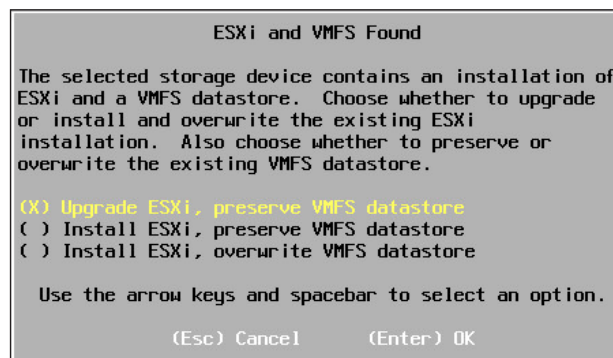


7. Next, the installer will display a list of available disks on which you can install or upgrade ESXi.

[ SAN LUNs are listed as remote devices, and local disks are listed as local devices.]



8. Select the device on which you are going to install ESXi, and press *Enter*.
9. If the selected device includes a VMFS data store or an installation of ESXi, you'll be prompted to decide what action you want to take. Select the desired action and press *Enter*.
10. The available actions are as follows:
 - **Upgrade ESXi, preserve VMFS datastore:** This option upgrades to ESXi 5 and preserves the existing VMFS data store
 - **Install ESXi, preserve VMFS datastore:** This option installs a fresh copy of ESXi 5 and preserves the existing VMFS data store
 - **Install ESXi, overwrite VMFS datastore:** This option overwrites the existing VMFS data store with a new one, and freshly installs ESXi 5



11. Select the desired keyboard layout and press *Enter*.

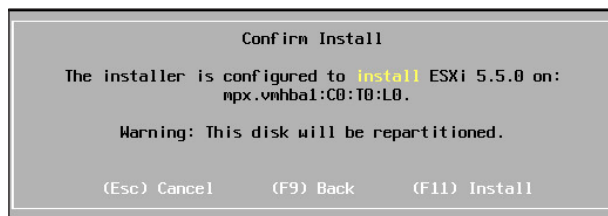


12. Enter a password for the root account. Press *Enter* when you are ready to continue with the installation:

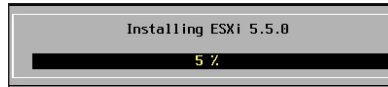


13. When you reach the final confirmation screen, press *F11* to proceed with the installation of ESXi.

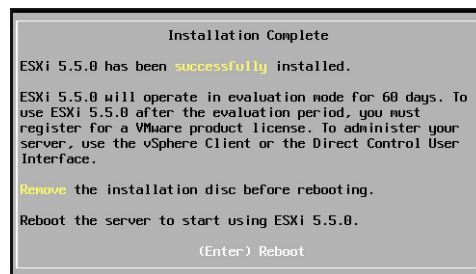
Once the installation process begins, it takes a few minutes to install ESXi on the selected storage device.



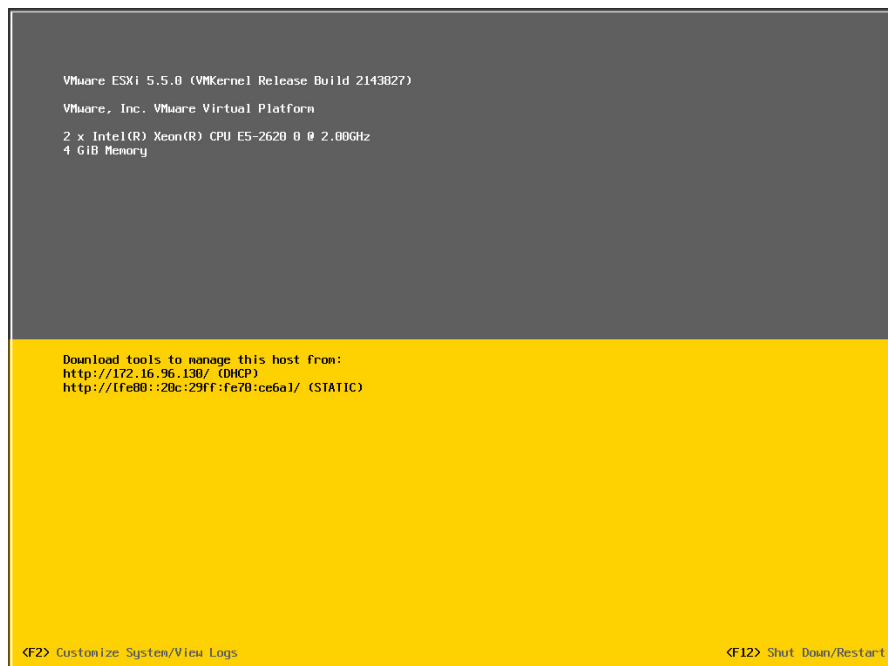
14. The installation of ESXi will start after you have clicked on *Install*:



15. Press *Enter* to reboot the host at the **Installation Complete** screen:

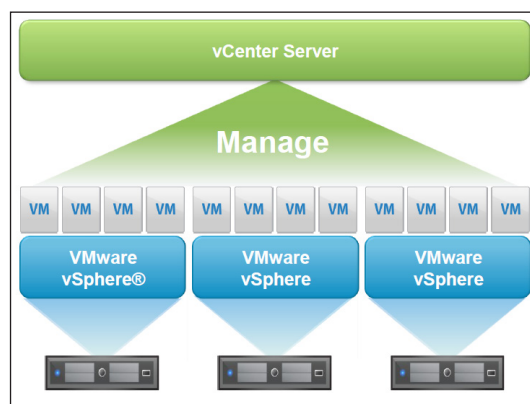


The ESXi installation is completed after reboot.



Installing VMware vCenter Server

VMware Center Server is a central management server that is used to manage VMware ESXi Server and virtual machines, for example, VM cloning, creating a VM template, vMotion, storage vMotion, and so on. We can connect to vCenter by vSphere Client or vSphere Web Client. You can set up vCenter Server or deploy vCenter Server Appliance to manage your virtual environment. vCenter Server Appliance is a preconfigured Linux-based virtual machine optimized for running vCenter Server's services. The following figure shows vCenter and ESXi management:



The following is the procedure for the installation of VMware vCenter Server 5.1 on Microsoft Windows 2008 R2.

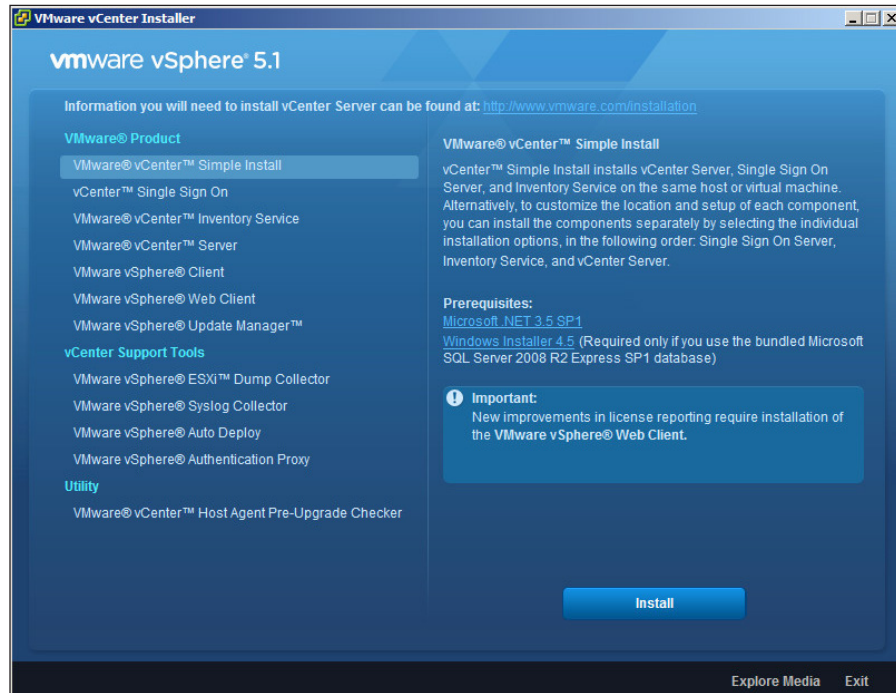
Download the vCenter Server installer from VMware website, choose the product as **VMware vSphere version 5.1**, and select **VMware vCenter Server 5.1** from <https://my.vmware.com/web/vmware/downloads>.

Prerequisites

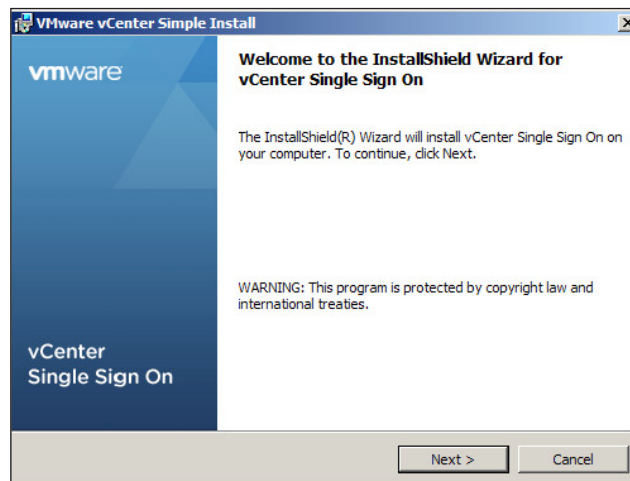
Microsoft .NET 3.5 SP1 and Windows Installer 4.5 (required only if you use the bundles Microsoft SQL Server 2008 R2 Express SPA database):

1. Ensure that your server is configured to boot from the CD-ROM drive.
2. Ensure that the VMware vCenter installation media are available for the server:
 - If it is a local installation, insert the VMware ESXi installation CD into the optical drive
 - If it is a remote installation, map an image of the installation media, known as an ISO image, to a virtual optical drive

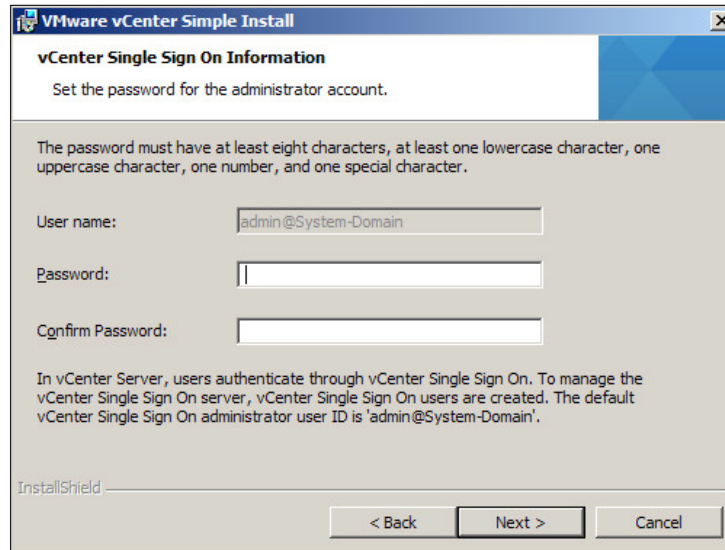
3. Select **VMware vCenter Simple Install**, and then click on **Install**, as shown here:



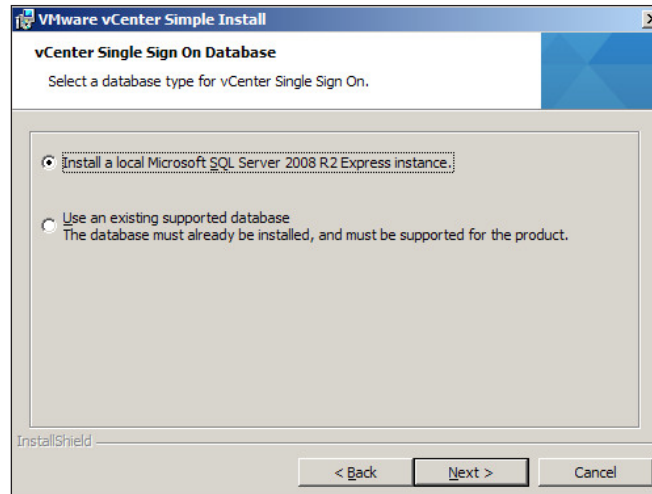
4. The installer will install **vCenter Single Sign-On**, **vCenter Inventory Service**, and **vCenter Server Service**.



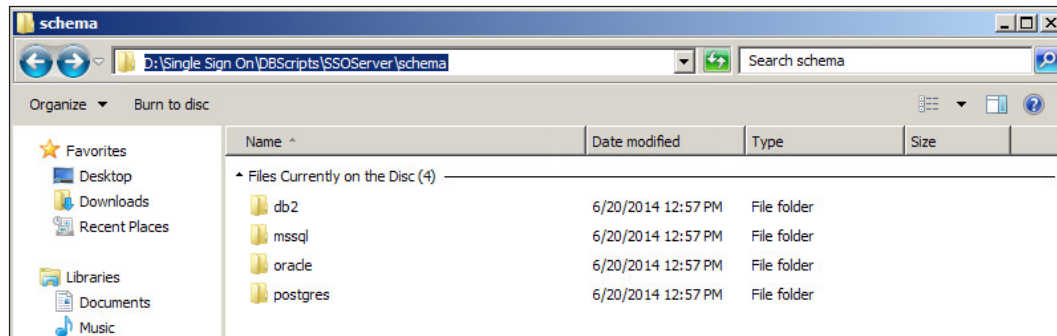
5. Start installing **vCenter Single Sign On (SSO)**. Then set the password for the administrator of vCenter Single Sign On:



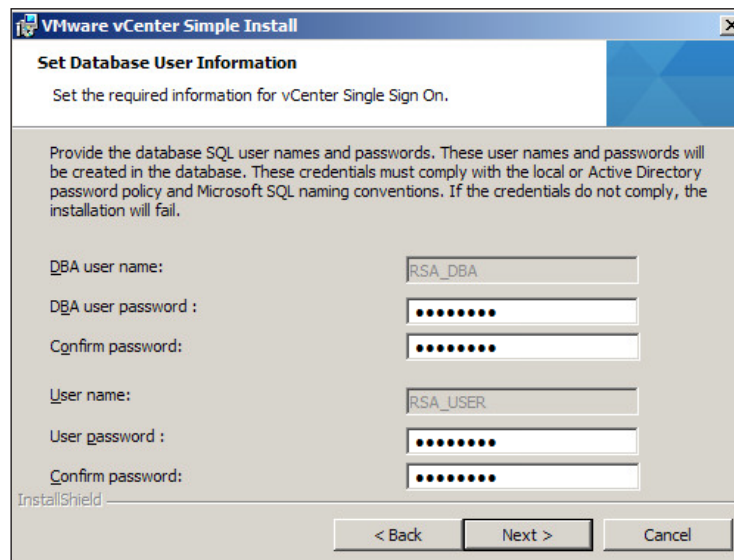
6. Select the database type for vCenter SSO:



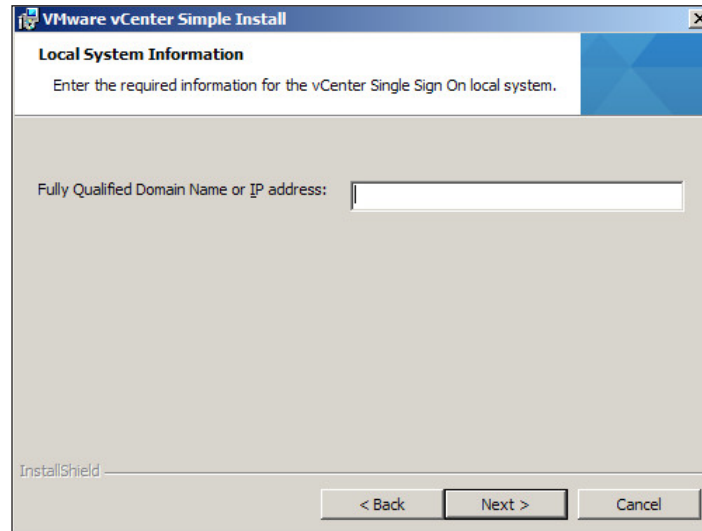
7. If you are using an existing database, it is required to create for the database by the script. The script is located at vCenter Server Installation directory\Single SignOn\DBScripts\SSOserver\Schema\. The SSO database is created by the SetupTablespaces.sql script. Then create a database user (RSA_USER) and database administrator (RSA_DBA) by the SetupUsers.sql script.



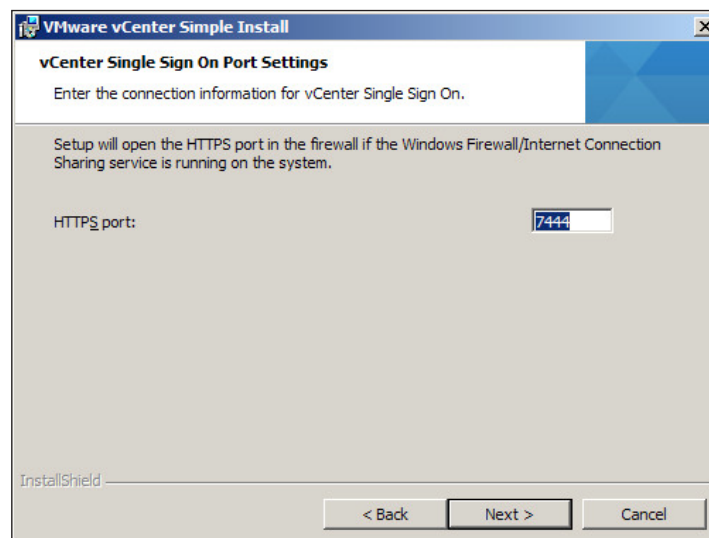
8. If you are using the bundled Microsoft SQL Server 2008 R2 Express database, enter the passwords for an SSO database administrator (RSA_DBA) and database user (RSA_USER). The installer uses these credentials to create the users in the database:



9. Enter **Fully Qualified Domain Name or IP address** for the vCenter SSO:

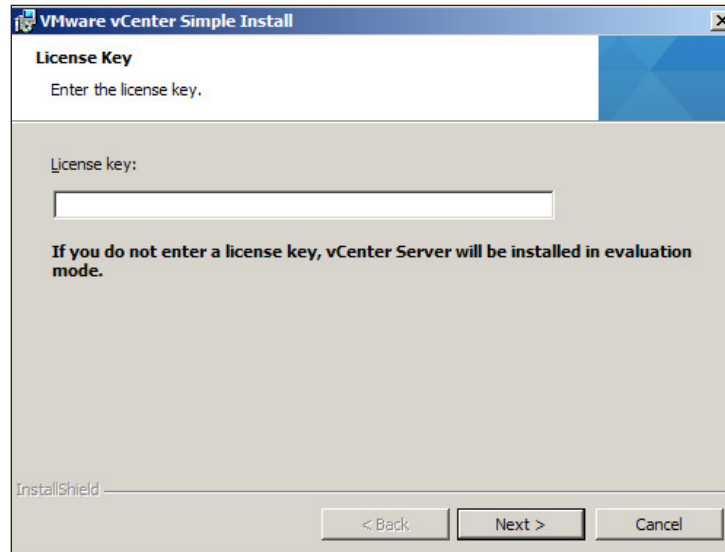


10. Accept the default **HTTPS port** for vCenter SSO, and click on Install:




The installer will install vCenter Inventory Service after it has finished installing the vCenter SSO.

11. Also, it will install vCenter Server after it has finished installing the vCenter Inventory Service. Enter the vCenter **License key**:

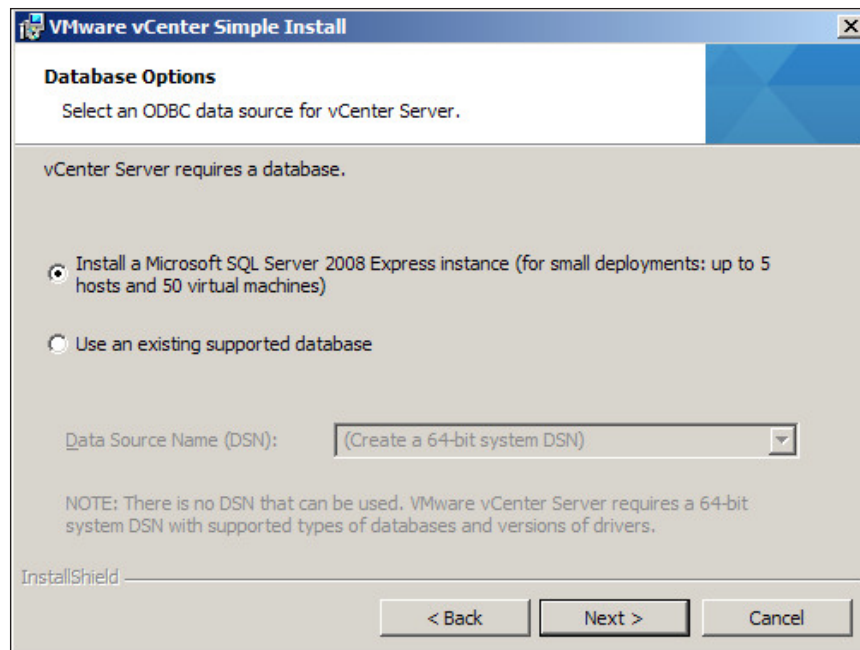


12. Select the type of database that you want to use:
- To use the bundled database, click on **Install a Microsoft SQL Server 2008 Express instance**. This database is limited to five hosts and 50 virtual machines.
 - To use an existing database, click on **Use an existing supported database**. Select your database from the list of available DSNs. Enter the username and password for the DSN.

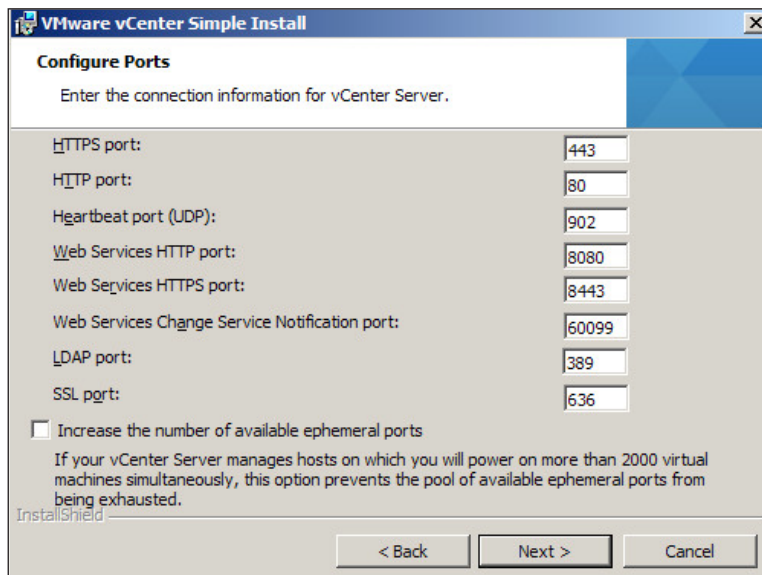
- [


 - **Option 1:** You can choose this option if your vSphere environment is small in size and limited to five hosts and 50 virtual machines
 - **Option 2:** You can choose this option if your vSphere environment is large in size and has hundreds of hosts and virtual machines

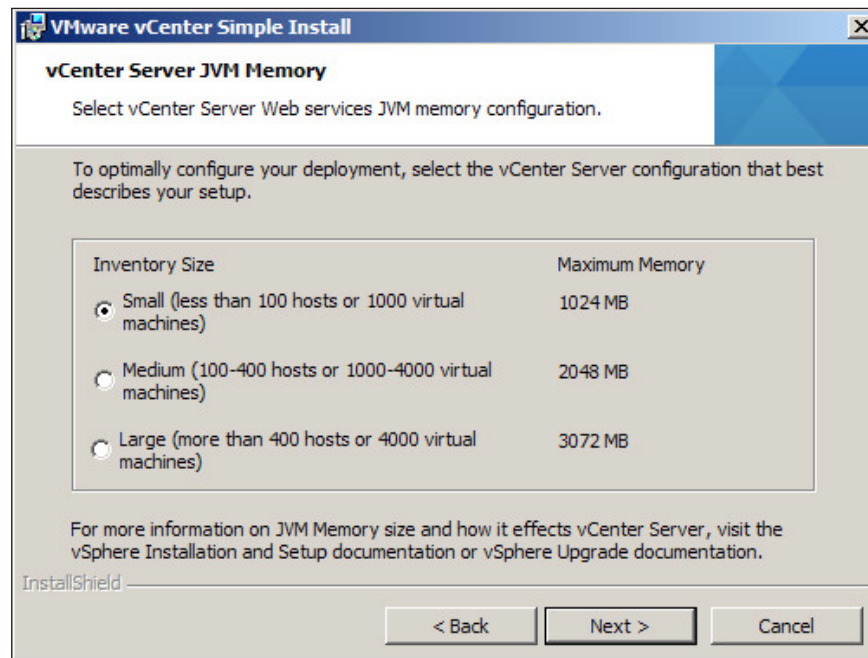
]



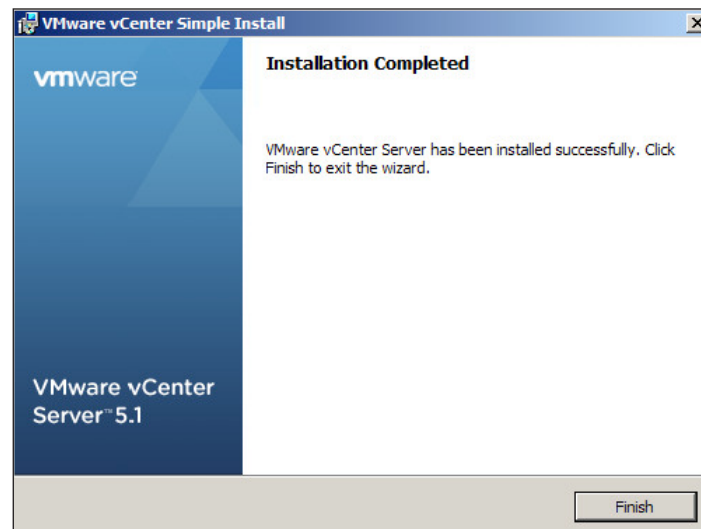
12. Enter the FQDN of the system that you are installing vCenter Server on:



13. Select the **vCenter Server JVM Memory** configuration, as shown in the following screenshot, and click on Install:



Installation completed!



Installing the VMware vCenter Server Appliance

VMware vCenter installation can be deployed as a virtual appliance. You can deploy VMware vCenter Server Appliance in your virtual environment if you are not spending much of your time configuring vCenter Server on a Microsoft Windows platform. VMware vCenter Server Appliance is a preconfigured Linux-based virtual machine optimized for running vCenter Server and associated services.



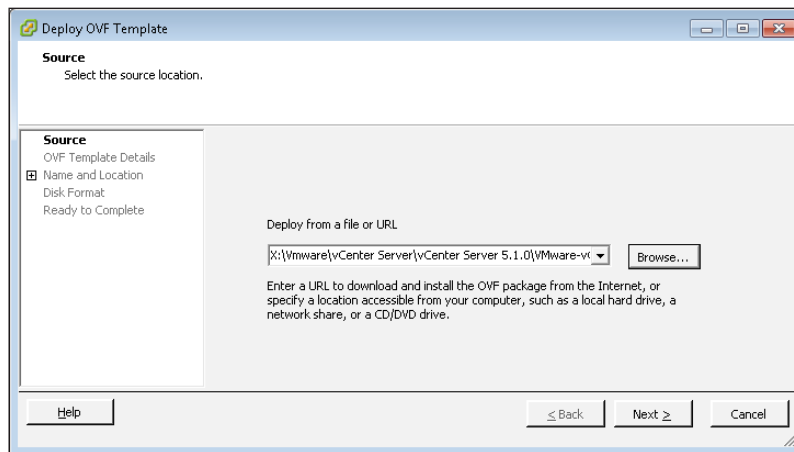
The embedded database is not configured to manage more than 5 hosts and 50 virtual machines. vCenter Server will stop responding if it exceeds these limits. Also, the vCenter Server Appliance does not support **Linked Mode** configuration. The following is the procedure required for installing VMware vCenter Server Appliance.

Download the vCenter Server Appliance (OVA) from the VMware website. Choose the product as **VMware vSphere version 5.1**, and select **VMware vCenter Server 5.1 Appliance** from <https://my.vmware.com/web/vmware/downloads>.

Prerequisites

The hosts should be running VMware ESX version 4.x, ESXi version 4.x, or later:

1. Use the vSphere Client or the vSphere Web Client to deploy the vCenter Server Appliance from the OVF template. If you don't want to allocate thick disk space for the deployment, you can deploy the vCenter Server Appliance with thin provisioning.



Deploying VMware vCenter Server Appliance from the OVF template



- Create a virtual disk in the default thick format. The space required for the virtual disk is allocated during its creation.
- Create a virtual disk in thin provisioning. At first, a thin provisioned disk uses only as much data store space as it initially needs.

2. Configure the IP address of vCenter Server Appliance and power it up:

```
proc on /proc type proc (rw)
sysfs on /sys type sysfs (rw)
debugfs on /sys/kernel/debug type debugfs (rw)
[ 53.204446] loop: module loaded
udev on /dev type tmpfs (rw,mode=0755)
tmpfs on /dev/shm type tmpfs (rw,mode=1777)
devpts on /dev/pts type devpts (rw,mode=0620,gid=5)
[ 53.245190] kjournald starting. Commit interval 15 seconds
[ 53.245676] EXT3-fs (sda1): using internal journal
[ 53.245872] EXT3-fs (sda1): mounted filesystem with ordered data mode
/dev/sda1 on /boot type ext3 (rw,nosuid,nodev,acl,user_xattr)
[ 53.254435] kjournald starting. Commit interval 15 seconds
[ 53.254858] EXT3-fs (sdb1): using internal journal
[ 53.255052] EXT3-fs (sdb1): mounted filesystem with ordered data mode
/dev/sdb1 on /storage/core type ext3 (rw,nosuid,nodev)
[ 53.262104] kjournald starting. Commit interval 15 seconds
[ 53.262515] EXT3-fs (sdb2): using internal journal
[ 53.262715] EXT3-fs (sdb2): mounted filesystem with ordered data mode
/dev/sdb2 on /storage/log type ext3 (rw,nosuid,nodev)
[ 53.271439] kjournald starting. Commit interval 15 seconds
[ 53.271910] EXT3-fs (sdb3): using internal journal
[ 53.272096] EXT3-fs (sdb3): mounted filesystem with ordered data mode
/dev/sdb3 on /storage/db type ext3 (rw,nosuid,nodev)
Check if the profiles matches the system
-
done
done
```

3. Follow the instructions on the welcome screen to open a browser window and go to the URL shown. The default username is root and password is vmware. It is highly recommended to change the password immediately after the first login:

VMware vCenter Server Appliance

Login

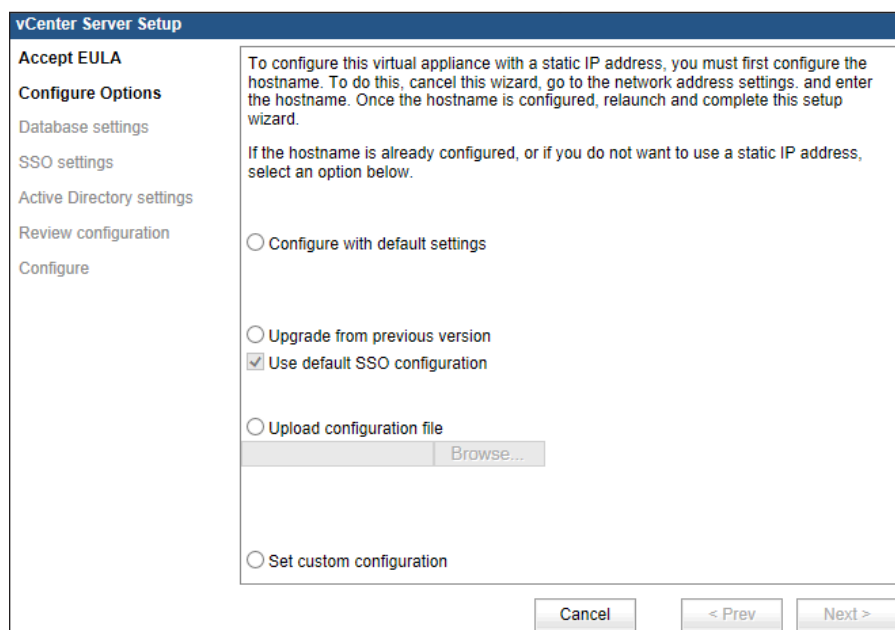
User name:

Password:

Login

The Login page of VMware vCenter Server Appliance

4. When you log in, the **vCenter Server Setup** wizard starts. You can select three options for installation:




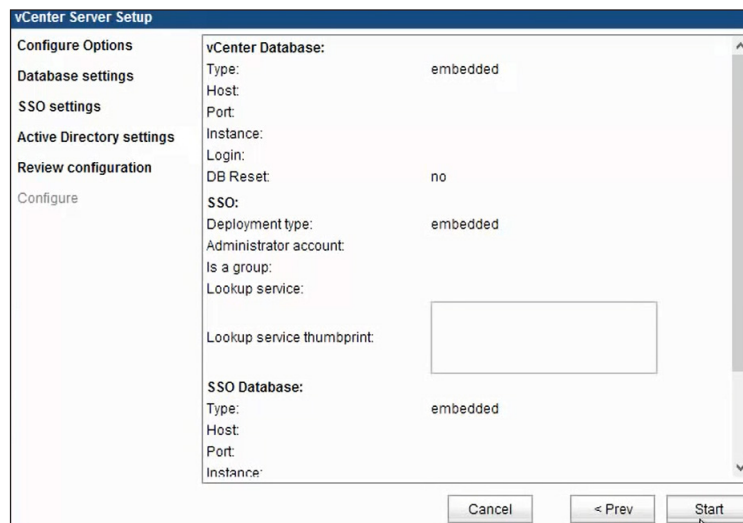
vCenter Server Appliance setup

The options in the previous screenshot are explained as follows:

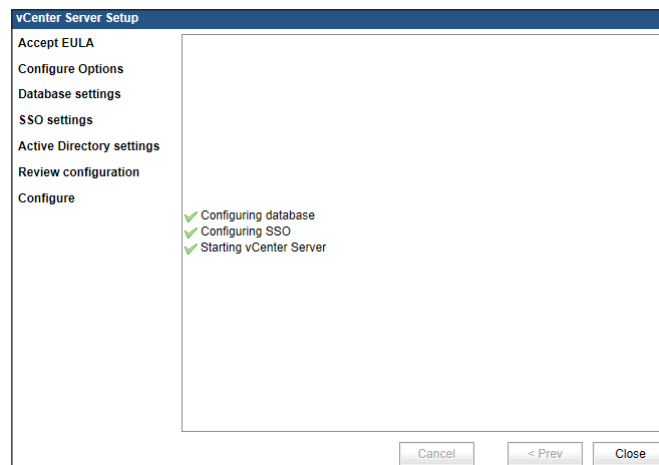
Option	Description
Configure with default settings	This sets up embedded vCenter Server and vCenter SSO databases in the vCenter Server Appliance, and configures the databases and Active Directory with default settings.
Upgrade from previous version	This is used to configure the vCenter Server Appliance from a prepared configuration file.
Set custom configuration	This is used to customize the configuration of the vCenter Server Appliance. Configure the appliance to be embedded, or the external vCenter Server and vCenter SSO databases, and to configure custom Active Directory settings .

5. For this installation, we choose **Configure with default settings**. This will configure the database and the SSO type will be embedded. Then click on the **Start** button.

[ For a small environment, you should use the embedded database. However, if your environment is large in size, you should use the other database type.]



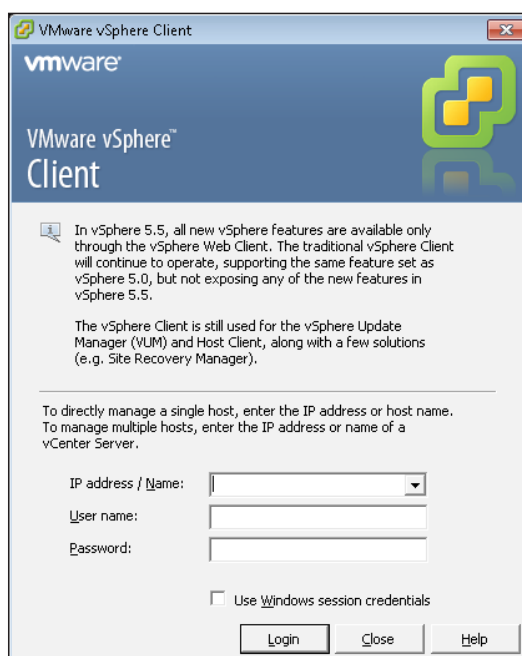
6. After finishing the configuration of the database and SSO, it will start the service of vCenter Server:



The installation of vCenter Server Appliance is complete

Connecting to vCenter Server with vSphere Client

You can connect to vCenter Server by vSphere Client or vSphere Web client if you want to manage an individual ESXi host or virtual machine. It can also log in to the ESXi host directly to manage an individual host using the vSphere Client or Web client. Enter the IP address, username, and password for your vCenter Server:

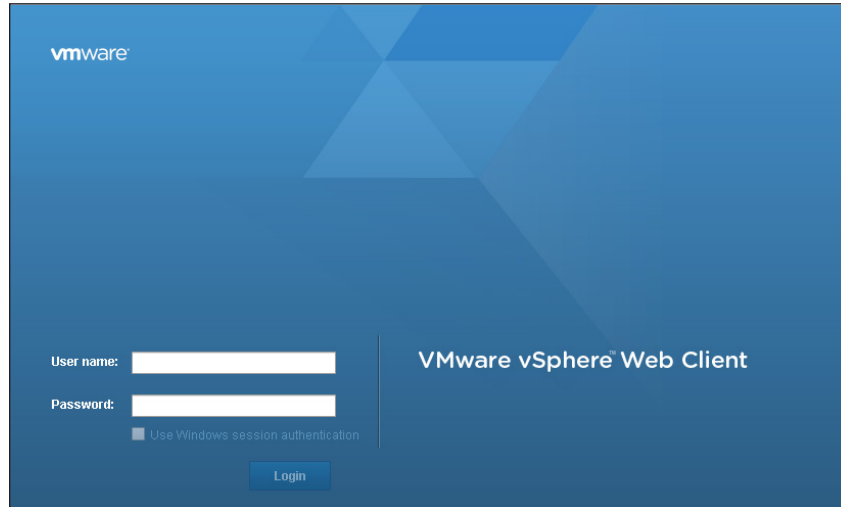


Connecting to vCenter Server with vSphere Web Client

Follow these steps to connect to vCenter Server by vSphere Web Client:

1. Open a web browser and enter the URL `https://<vcenter FQDN or vcenter IP address>: 9443/vsphere-client/` (port 9443 is the default). An example of this is `https://vcenter55u2.vmlab.com:9443/vsphere-client/`.

2. Enter the IP address, username, and password for your vCenter Server:



Summary

In this chapter, you learned the differences between physical and virtual hosts, and the benefits of using a virtualized host. We described how to set up VMware ESXi Server and VMware vCenter Server, and how to connect to vCenter Server using vSphere Client and vSphere Web Client. We also saw what vCenter Server Appliance is and how to deploy it on our VMware virtualized environment.

In the next chapter, we will take a look at vSphere Management Assistant.

2

Getting Started with vSphere Management Assistant

After vSphere deployment, ESXi configuration and troubleshooting are the most common operations performed by vSphere Client. VMware administrators often choose the vSphere Client GUI to configure the ESXi host. However, they sometimes choose the command line for configuration or maintenance tasks, for example, `esxcli`, `vicfg`, and `vmware-cmd` commands. This is because some logs or information about the vSphere host cannot be found using vSphere Client. VMware **vSphere Management Assistant (vMA)** is a useful tool, and it is used to interact with ESXi hosts and vCenter Server. For example, you can add the ESXi host or vCenter Server to vMA, and then manage the individual ESXi host or virtual machine using vMA's commands.

During troubleshooting, we need to know how to gather the log files of ESXi and vCenter Server, and where we can find the related system log, for example, the VMkernel, warning message, management agent, the `esxupdate` log, and so on.

In this chapter, we will cover the following topics:

- Deploying and configuring vMA
- Configuring the ESXi technical support mode and SSH access
- Use the `esxcli`, `vicfg`, and `vmware-cmd` commands
- Reviewing the ESXi and vCenter Server logs

Deploying vMA

In the first chapter, we saw that we can manage the ESXi host or virtual machine by VMware vCenter Server. In some situations, some operations and configurations cannot execute in vCenter Server; for example, if one virtual machine stops responding, the VMware administrator can normally reboot that virtual machine using vSphere Client. However, they cannot reboot and shut down this virtual machine in vCenter Server. The solution is to force a reboot or shut it down using the command line. VMware vMA is a useful tool, which is a virtual appliance. It is a SUSE Linux Enterprise Server 11-based virtual machine that includes the VMware CLI and vSphere SDK for Perl. System administrators can run scripts that interact with ESXi hosts and vCenter Server systems.

The following is the procedure for deploying VMware vMA:

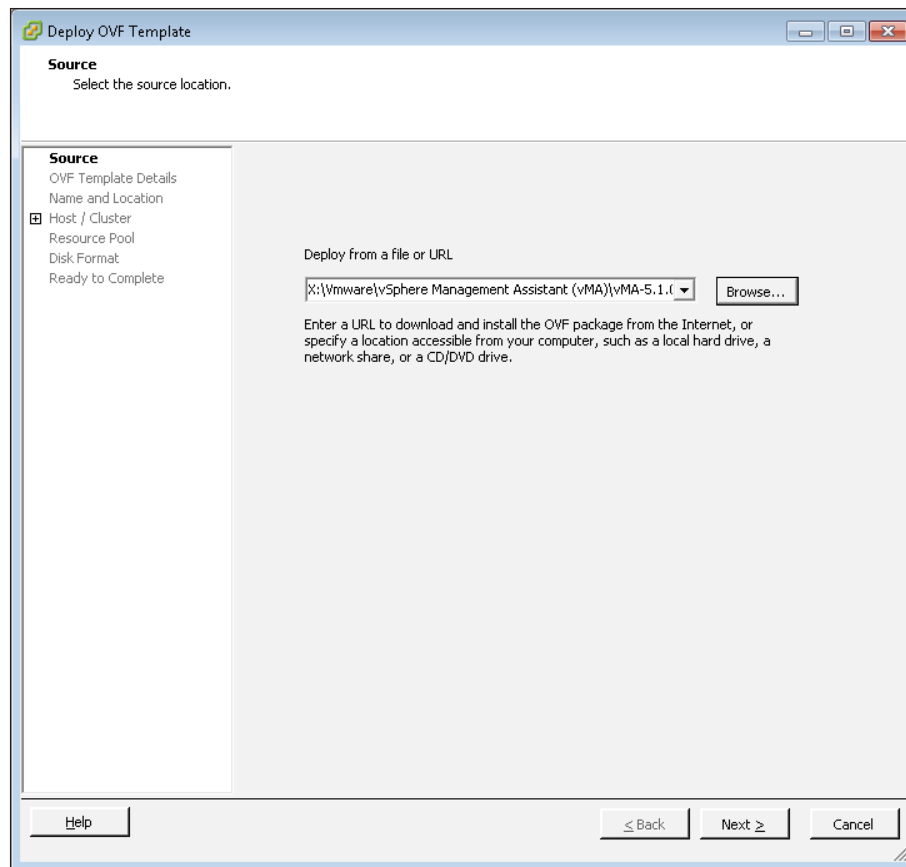
1. Download vMA from the VMware website (<https://my.vmware.com/web/vmware/downloads>). Choose **VMware vSphere version 5.1** as the product, go to the **Drivers & Tools** tab, and select **VMware vSphere Management Assistant**.



Requirements

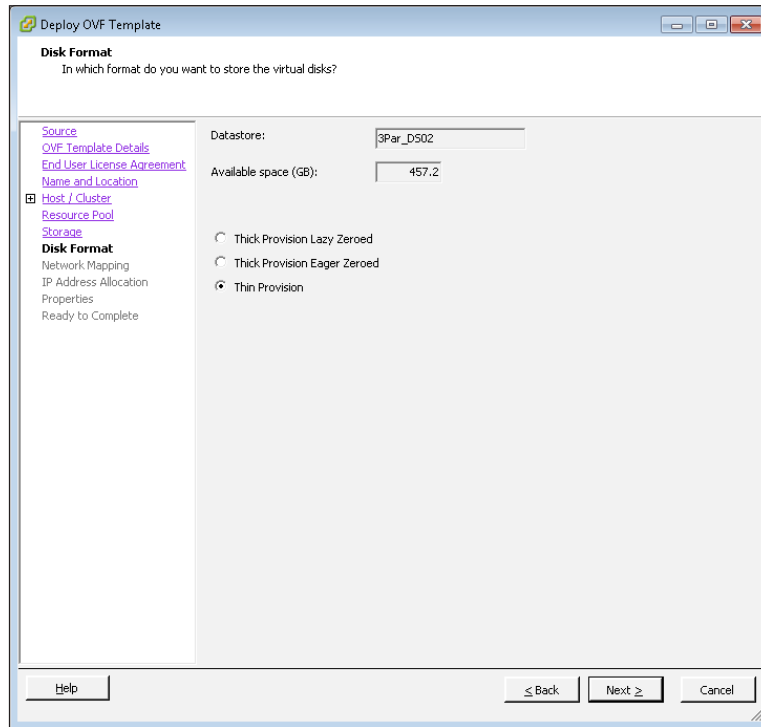
You can deploy vMA on VMware vSphere 4.1 or later versions.

2. Deploy the OVF template in VMware vSphere Server or VMware vCenter Server, as shown in the following screenshot:



3. Accept the license agreement and click on **Next**.
4. Enter a name for vMA.
5. Select the store location and resource pool for vMA.

6. Select the disk format for vMA, as shown in this screenshot:



You should choose the disk in the **Thin Provision** format if the destination data store is the SAS/FC drive. It has three options for the virtual disk, which are detailed in the following table:

Option	Description
Thick Provision Lazy Zeroed	Creates a virtual disk in a default thick format. The space required for the virtual disk is allocated during its creation. Any data left over on the physical device is not erased during creation, but is zeroed out on demand at a later time, during the first write from the virtual machine.
Thick Provision Eager Zeroed	Creates a thick disk that supports clustering features, such as fault tolerance. The space required for the virtual disk is allocated at the time of creation. In contrast to the Thick Provision Lazy Zeroed format, the data left over on the physical device is zeroed out during creation. It might take longer to create disks in this format than other types of disks.

Option	Description
Thin Provision	Uses the thin provisioned format. At first, a thin provisioned disk uses only as much data store space as the disk initially needs. If the thin disk needs more space later, it can grow to the maximum capacity allocated to it.

1. Select the network mapping and click on **Next**.



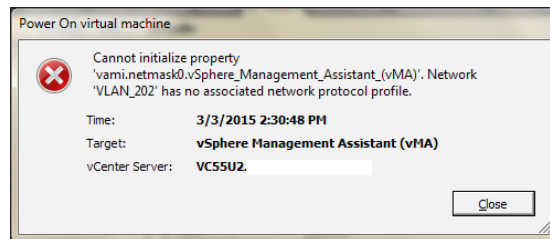
Ensure that vMA is connected to the management network on which the vCenter Server system and the ESXi hosts reside. The vMA appliance is designed to help the administrator manage the ESXi host and vCenter, so it needs to be connected to the same management network.

The screenshot shows the 'Deploy OVF Template' wizard window. The title bar says 'Deploy OVF Template'. The main heading is 'IP Address Allocation' with the subtext 'Which allocation scheme should be used to allocate IP addresses?'. On the left is a navigation pane with links: Source, OVF Template Details, End User License Agreement, Name and Location, Host / Cluster (selected), Resource Pool, Storage, Disk Format, IP Address Allocation (current step), Properties, and Ready to Complete. The main area contains the text 'Choose the IP allocation policy to use:' followed by three radio button options: 'Fixed' (selected), 'Transient', and 'DHCP'. Each option has a description: 'Fixed' (IP addresses are manually configured. No automatic allocation is performed.), 'Transient' (IP addresses are automatically allocated from the vCenter managed IP network range at power-on, and released at power-off.), and 'DHCP' (A DHCP server is used for IP allocation.). Below these is a label 'Choose the IP protocol to use:' and a dropdown menu showing 'IPv4'. At the bottom are buttons for 'Help', '< Back', 'Next >', and 'Cancel'.

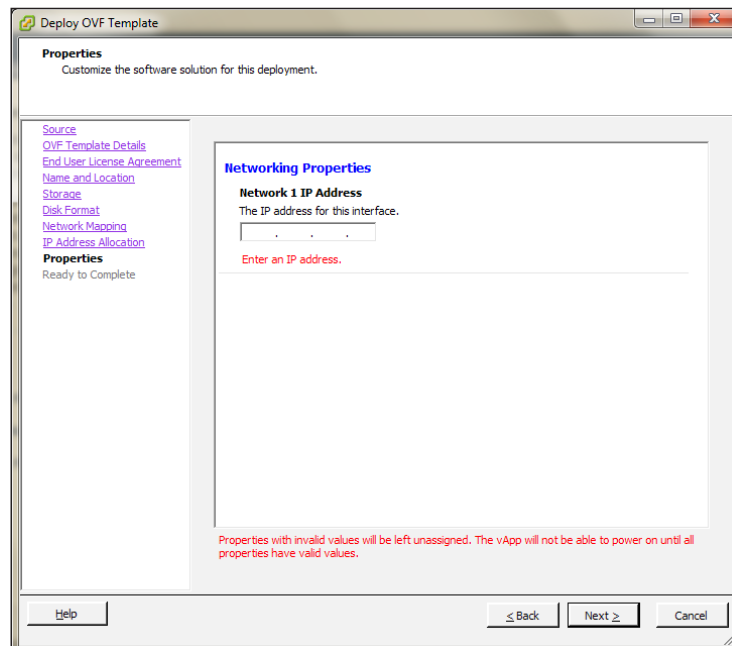
The following table lists the IP address allocation policy to use:

Option	Description
Fixed	Set up the IP address manually
Transient	IP addresses are automatically allocated from the vCenter-managed IP network range at power on and released at power off
DHCP	A DHCP server is used for IP allocation

You need to create an IP pool on vCenter Server and associate it with vMA's network label if you select the **Fixed** option. The vMA appliance cannot be powered on if it doesn't associate with the correct network label. Then it displays the following error:



2. Input one IP address for vMA.





The vMA will not be able to power on if you don't input the IP address.

3. Click on **Finish**.
4. Power on the vMA after the deployment is finished.
5. Configure the network setting on **Main Menu**. Choose the options numbered 2 and 6.

```

Main Menu
0) Show Current Configuration (scroll with Shift-PgUp/PgDown)
1) Exit this program
2) Default Gateway
3) Hostname
4) DNS
5) Proxy Server
6) IP Address Allocation for eth0
Enter a menu number [0]: _

```



You can configure only one network adapter (vNIC) in vMA. You cannot add and configure multiple network adapters. This is vMA's default configuration; you cannot change it.

6. After completing the configuration, it is required that you input a password for **vi-admin**. Input the default password as **vmware**, and enter a new password for **vi-admin**. It is highly recommended that you change the password for **vi-admin** upon the first login.

```

Proxy Server:
Main Menu
0) Show Current Configuration (scroll with Shift-PgUp/PgDown)
1) Exit this program
2) Default Gateway
3) Hostname
4) DNS
5) Proxy Server
6) IP Address Allocation for eth0
Enter a menu number [0]: 1

Starting password configuration ...
The root account is disabled in this vMA virtual machine, which means no one can
log in as root. The administrator account for vMA is called "vi-admin". In orde
r to log in to vMA, you need to log in as this user. This user has been pre-crea
ted in the vMA, and its password needs to be set now. Please enter a secure pass
word for the account now.

Please provide a password for the vi-admin user. If you are prompted for an old
password for this user, enter vmware.
Old Password: _

```



The new password must have at least nine characters, one uppercase character, one lowercase character, one numeral, and one symbol.

Configuring ESXi technical support mode

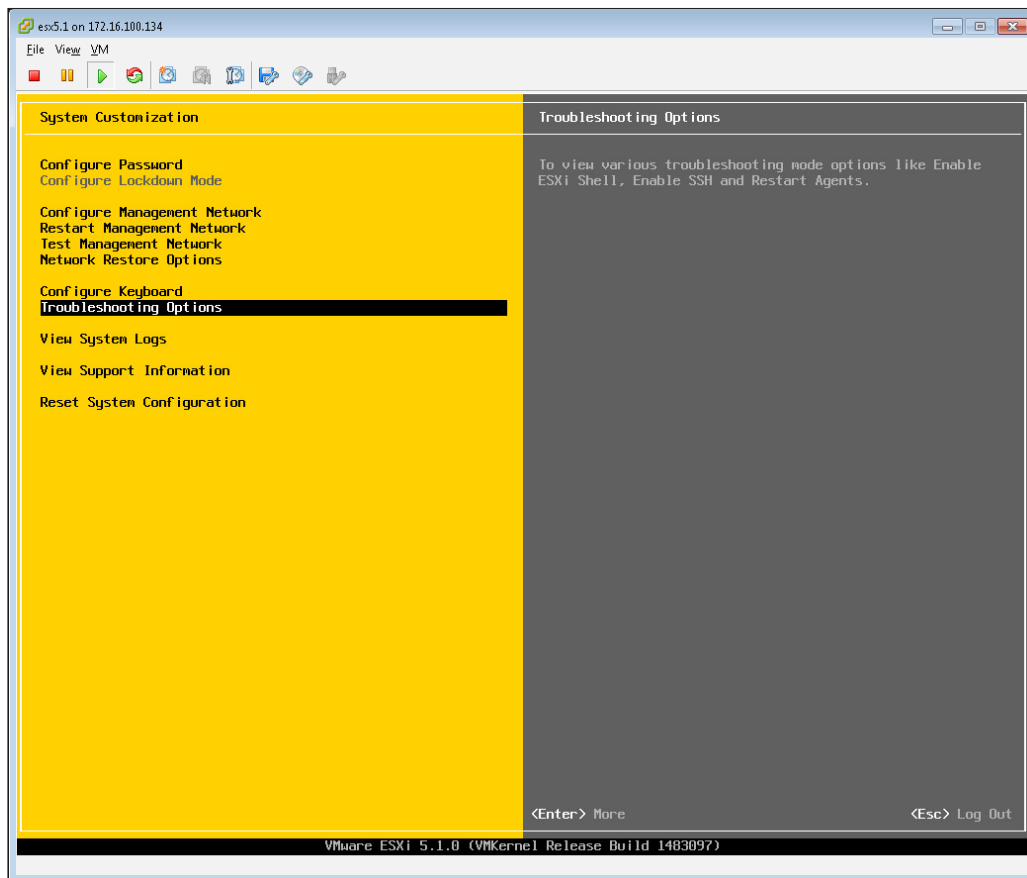
For security reasons, by default, the ESXi Shell of VMware vSphere is disabled. The VMware administrator needs to enable it manually if they want to access it locally or remotely, due to some maintenance tasks, for example, ESXi service patch upgrade tasks, HBA or NIC driver upgrade and parameter configuration, and so on. The ESXi Shell can be used by administrators to troubleshoot on VMware ESXi hosts. We can access it in two ways:

- Logging in directly on the console of the ESXi host
- Logging in remotely by SSH

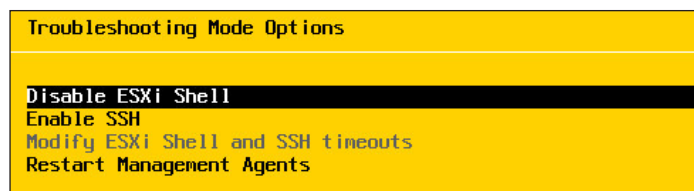
Enabling and accessing ESXi Shell

When you want to access ESXi Shell, you should enable ESXi Shell first. Then you can execute `esxcli`. To enable local or remote ESXi Shell from the **Direct Console User Interface (DCUI)**, we have to follow these steps:

1. From the DCUI of the ESXi host, press *F2*. Provide the credentials when prompted.
2. Select **Troubleshooting Options**. You can enable local TSM or remote TSM.
3. To enable local TSM allows users to log on to the Virtual Console ESXi hosts.



4. Select **Enable SSH** and press *Enter*.



In ESXi 5.x, select **Enable ESXi Shell** and press *Enter*.
In ESXi 4.1, select **Local Tech Support** and press *Enter*.

5. To enable remote TSM allowing users to log in via SSH on the virtual console of the ESXi host.



In ESXi 5.x, select **Enable SSH** and press *Enter*.

In ESXi 4.1, select **Remote Tech Support (SSH)** and press *Enter*.

Enabling ESXi Shell from vSphere Client

If you cannot access the DCUI used to enable ESXi Shell, you can enable ESXi Shell using vSphere Client. To enable local or remote TSM from vSphere Client, we can follow these steps:

1. Select the host and click on the **Configuration** tab.
2. Click on **Security Profile** and then on **Properties**, as shown in this screenshot:

The screenshot shows the vSphere Client interface with the Configuration tab selected. The left sidebar shows the Hardware and Software sections. The main pane displays the Security Profile section, which includes a list of services and a firewall table.

Incoming Connections		
CIM Server	5988 (TCP)	All
CIM SLP	427 (UDP,TCP)	All
vSphere Web Access	80 (TCP)	All
CIM Secure Server	5989 (TCP)	All
SSH Server	22 (TCP)	All
Fault Tolerance	8100,8200,8300 (TCP,UDP)	All
SNMP Server	161 (UDP)	All
DHCPv6	546 (TCP,UDP)	All
vMotion	8000 (TCP)	All
DHCP Client	68 (UDP)	All
NFC	902 (TCP)	All
vSphere High Availability Agent	8182 (TCP,UDP)	All
vSphere Client	902,443 (TCP)	All

3. Select **SSH** and click on **Options**. Then start up the service.

Label	Daemon
vSphere High Availability Agent	Running
vpaa	Running
ESXi Shell	Running
xorg	Stopped
Local Security Authentication Serv...	Stopped
NTP Daemon	Running
vprobed	Stopped
SSH	Running
Direct Console UI	Running
CIM Server	Running



In ESXi 5.x, select **SSH** or **ESXi Shell** and click on **Options**. Then start up the service.

Accessing ESXi Shell from DCUI

If you can remotely connect to the ESXi host directly using the remote management console (for example, HP **Integrated Lights-Out (iLO)**, IBM RSA, and so on), you access ESXi Shell from DCUI. The following is the procedure.

Press **Alt + F1** at the main DCUI screen and provide the credentials.

```

ESXi 5.1.0 http://www.vmware.com
Copyright (c) 2007-2014 VMware, Inc.

localhost.boardware.com.mo login: root
Password:
The time and date of this login have been sent to the system logs.

VMware offers supported, powerful system administration tools. Please
see www.vmware.com/go/sysadmintools for details.

The ESXi Shell can be disabled by an administrative user. See the
vSphere Security documentation for more information.
~ # _

```

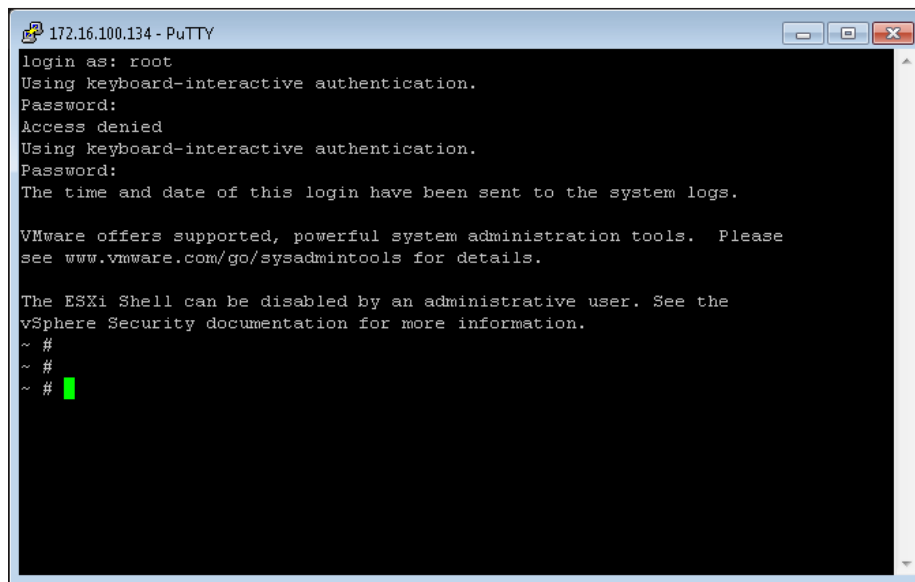


Press **Alt + F2** to return to the main DCUI screen of the ESXi host.

Accessing the remote ESXi Shell from the SSH client

If you cannot access ESXi Shell directly by DCUI, you should access ESXi Shell using SSH. The following is the procedure required:

1. Open an SSH client.
2. Input the IP address or domain name of the ESXi host and provide the credentials.



```
172.16.100.134 - PuTTY
login as: root
Using keyboard-interactive authentication.
Password:
Access denied
Using keyboard-interactive authentication.
Password:
The time and date of this login have been sent to the system logs.

VMware offers supported, powerful system administration tools. Please
see www.vmware.com/go/sysadmintools for details.

The ESXi Shell can be disabled by an administrative user. See the
vSphere Security documentation for more information.
~ #
~ #
~ # █
```

[ By default, SSH works on TCP port 22.]

Using the VMware commands

We have described the configuration of vMA in the first section. In this section, you will learn how to manage the supported vCenter Server or ESXi host using vMA. Firstly, you will learn how to connect the vCenter or ESXi to vMA. The **vifp** interface enables the administrator to add, list, and remove the target host (ESXi host or vCenter Server) from vMA. The following are some examples of ESXi host maintenance tasks using vMA.

vMA has included many commands, for example, `esxcli`, `resxstop`, `vicfg`, `esxcfg`, `vmware-cmd`, and so on.

The following table lists a description of each command:

Command	Description
<code>esxcli</code>	This is used to manage the vSphere host remotely or in the ESXi Shell.
<code>resxstop</code>	This is used to remotely monitor the resource of the vSphere host.
<code>vicfg- commands</code>	You should use this when deploying a vMA virtual machine and targeting the vSphere host.
<code>esxcfg- commands</code>	These are available in the ESXi Shell. Most commands are replaced by <code>esxcli</code> .
<code>vmware-vmd</code>	The operation includes creating a snapshot, power on, power off, and information about the virtual machine.

It is required to add vSphere host and vCenter Server to vMA using vMA commands before we manage the vSphere host and vCenter Server. This table lists the vMA management commands:

vMA command	Description
<code>vifp addserver</code>	Add the target ESX host or vCenter Server to vMA
<code>vifp removeserver</code>	Remove the target ESX host or vCenter Server from vMA
<code>vifp reconfigure</code>	Reconfigure the authentication policy or the users for the target hosts
<code>vifp rotatepassword</code>	Change the password for <code>vi-admin</code> and <code>vi-user</code> at the target hosts
<code>vifp listservers</code>	Verify that the target hosts have been added to vMA successfully
<code>vifptarget --set -s</code>	Set the target host as default for the current session
<code>vifptarget --clear -c</code>	Clear the target host for the current session
<code>vifptarget --display -d</code>	Display the initialized fast-pass target host
<code>vicfg-hostops</code>	Shut down and reboot the vSphere host
<code>vicfg-hostops</code>	Place the vSphere host to enter and exit maintenance mode
<code>vicfg-cfgbackup</code>	Backup and restore the host configuration

vMA command	Description
vicfg-authconfig	Add ESXi hosts to AD
vicfg-nics -l	Display information about the physical NIC of the ESXi host
vicfg-vswitch -l	Display information about the vSwitch of the ESXi host

The following table lists the common connection options for vMA commands:

vMA command option	Description
--vihost	Name of the ESXi host
--username	Login username
--url	Connect to the specified vSphere web service SDK URL
--server	The ESXi or vCenter Server
--savesessionfile	Saves the session in the specified file
--config	The path to a configuration file
--password	Login password
--portnumber	Specified port to connect to

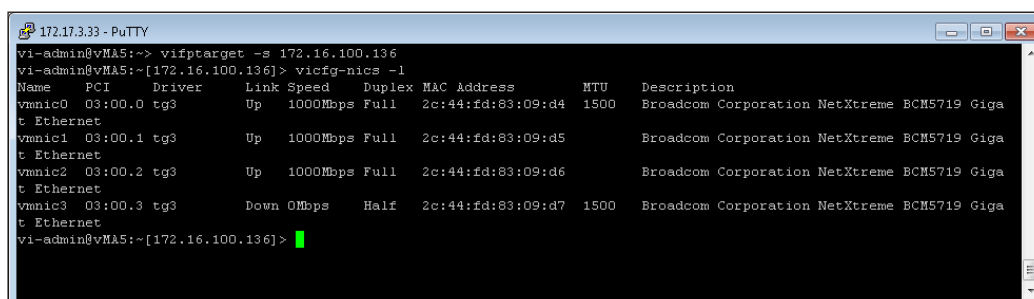
The following is the procedure for adding an ESXi host to vMA and collecting information about the physical network interface on an ESXi host:

```
vi-admin@vMA5:~> vifp addserver 172.16.100.136
root@172.16.100.136's password:
vi-admin@vMA5:~> vifp listservers
172.16.100.136 ESXi
vi-admin@vMA5:~>
```

We will follow these steps to add the ESXi Server to vMA using the vifp commands:

1. Log in as vi-admin (the default administrator).
2. Run `vifp addserver <VMware ESXi host>`.
3. Run `vifp listservers` to verify that the system that has been added as a target.
4. Run `vifptarget -s <VMware ESXi host>`.
5. Then run `vicfg-nics -l`.

The VMware administrator has the right to access this ESXi host after adding the 172.16.100.136 ESXi host to vMA in step 2. They can then manage this ESXi host using the `vifptarget` command in step 4. At this moment, we can collect information about that ESXi host's physical NIC adapter using the `vicfg-nics` command. The output is shown in the following screenshot:



```

172.17.3.33 - PuTTY
vi-admin@vMA5:~> vifptarget -s 172.16.100.136
vi-admin@vMA5:~[172.16.100.136]> vicfg-nics -l
Name      PCI      Driver    Link Speed Duplex MAC Address MTU Description
vmnic0    03:00:0  tg3      Up   1000Mbps Full  2c:44:fd:83:09:d4 1500 Broadcom Corporation NetXtreme BCM5719 Giga
t Ethernet
vmnic1    03:00:1  tg3      Up   1000Mbps Full  2c:44:fd:83:09:d5      Broadcom Corporation NetXtreme BCM5719 Giga
t Ethernet
vmnic2    03:00:2  tg3      Up   1000Mbps Full  2c:44:fd:83:09:d6      Broadcom Corporation NetXtreme BCM5719 Giga
t Ethernet
vmnic3    03:00:3  tg3      Down 0Mbps  Half  2c:44:fd:83:09:d7 1500 Broadcom Corporation NetXtreme BCM5719 Giga
t Ethernet
vi-admin@vMA5:~[172.16.100.136]>

```

Let's discuss an example: the VMware administrator plans to upgrade the physical memory to one ESXi host, A, which is managed by vCenter Server. This ESXi host (for whom lockdown mode is enabled) is required to shut down before the hardware upgrade, but the VMware administrator cannot connect to vCenter Server and shut down the host A. Finally, they discover is that the capacity of vCenter's database is full, so the vCenter Server service cannot start up. At this moment, they cannot put A into maintenance mode and shut it down, but if there is one vMA in their virtual environment or if they can deploy one vMA in their virtual environment, then they can add A to vMA and shut it down using the `vicfg-hostops` command.

You cannot manage the ESXi host directly using vSphere Client or the web client if lockdown mode of ESXi is enabled.

The maintenance mode can be explained as follows: when you need to upgrade the resources of a vSphere host, such as the memory for example, the vSphere host must enter or leave maintenance mode.

You can put the ESXi host into maintenance mode and exit this mode using `vicfg-hostops`. You can check whether the ESXi host is in maintenance mode or in entering maintenance mode using the **info** option after adding the target host to vMA. The following result proves that the ESXi host doesn't enter maintenance mode:

1. `vicfg-hostops --server <esx host> --username <username> --password <password> --operation info`

```
vi-admin@vMA5:~> vicfg-hostops --server 172.16.100.77 --username root --password
password --operation info

Host Name           : esx51f.testlab.com
Manufacturer        : VMware, Inc.
Model               : VMware Virtual Platform
Processor Type      : Intel(R) Xeon(R) CPU E5-2640 0 @ 2.50GHz
CPU Cores           : 2 CPUs x 2493 GHz
Memory Capacity     : 4095.48828125 MB
VMotion Enabled     : no
In Maintenance Mode : no
Last Boot Time      : 2014-11-04T14:19:28.486739Z

vi-admin@vMA5:~> █
```

2. You can put the ESXi host into maintenance mode using the following command. Then you will see that the status of the maintenance mode of the host has changed to **yes**:

```
vicfg-hostops --server <esx host> --username <username> --password
<password> --operation enter
```

```
vi-admin@vMA5:~> vicfg-hostops --server 172.16.100.77 --username root --password
password --operation enter
Host esx51f.testlab.com entered into maintenance mode successfully.
vi-admin@vMA5:~> vicfg-hostops --server 172.16.100.77 --username root --password
password --operation info

Host Name           : esx51f.testlab.com
Manufacturer        : VMware, Inc.
Model               : VMware Virtual Platform
Processor Type      : Intel(R) Xeon(R) CPU E5-2640 0 @ 2.50GHz
CPU Cores           : 2 CPUs x 2493 GHz
Memory Capacity     : 4095.48828125 MB
VMotion Enabled     : no
In Maintenance Mode : yes
Last Boot Time      : 2014-11-04T14:19:28.486739Z


vi-admin@vMA5:~> █
```

3. You can place the ESXi host into exit maintenance mode using the following command, if you have completed the hardware upgrade. You will see that the status of maintenance mode of the host has changed to **no** after executing the command:

```
vicfg-hostops --server <esx host> --username <username> --password
<password> --operation exit
```

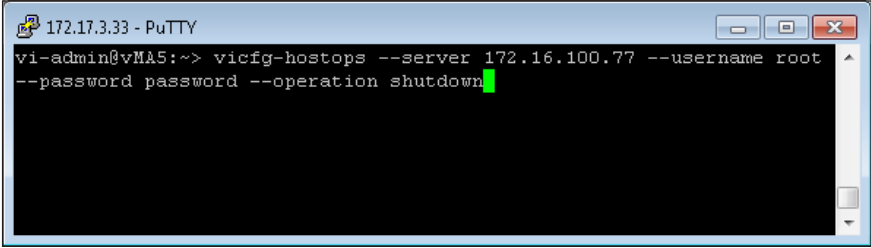
```
vi-admin@vMA5:~> vicfg-hostops --server 172.16.100.77 --username root --password
password --operation exit
Host esx51f.testlab.com exited from maintenance mode successfully.
vi-admin@vMA5:~> vicfg-hostops --server 172.16.100.77 --username root --password
password --operation info

Host Name           : esx51f.testlab.com
Manufacturer        : VMware, Inc.
Model               : VMware Virtual Platform
Processor Type      : Intel(R) Xeon(R) CPU E5-2640 0 @ 2.50GHz
CPU Cores           : 2 CPUs x 2493 GHz
Memory Capacity     : 4095.48828125 MB
VMotion Enabled     : no
In Maintenance Mode : no
Last Boot Time      : 2014-11-04T14:19:28.486739Z
vi-admin@vMA5:~>
```

 The ESXi host does not enter maintenance mode until all the virtual machines running on the host have been shut down, migrated, or suspended.

4. Then shut down and reboot the ESXi host using the following command:

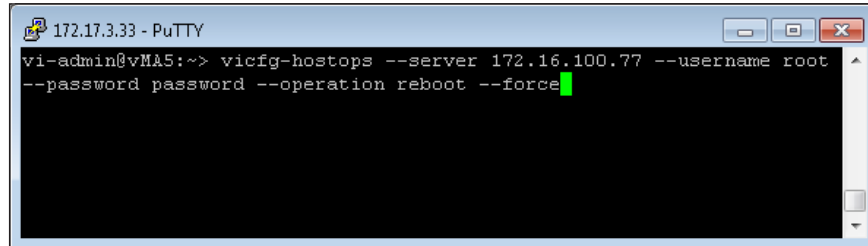
```
vicfg-hostops --server <esx host> --username <username> --password
<password> --operation shutdown
```




```
172.17.3.33 - PuTTY
vi-admin@vMA5:~> vicfg-hostops --server 172.16.100.77 --username root
--password password --operation shutdown
```


5. Force reboot the ESXi host using the following command:

```
vicfg-hostops --server <esx host> --username <username> --password  
<password> --operation reboot --force
```

A screenshot of a PuTTY terminal window titled '172.17.3.33 - PuTTY'. The terminal shows a command prompt 'vi-admin@vMA5:~>' followed by the command 'vicfg-hostops --server 172.16.100.77 --username root --password password --operation reboot --force'. A green cursor is visible at the end of the command line.

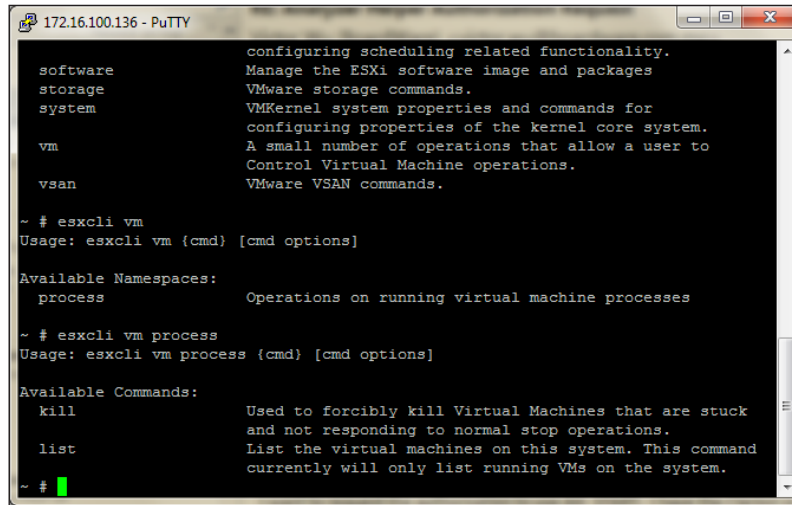
```
172.17.3.33 - PuTTY  
vi-admin@vMA5:~> vicfg-hostops --server 172.16.100.77 --username root  
--password password --operation reboot --force
```

 It is recommended that the ESXi host enters maintenance mode before rebooting the ESXi host.

According to the previous example, you can learn how to place an ESXi host into maintenance mode and shut it down by vMA. Then, complete the hardware upgrade and reboot the ESXi host. Finally, place the ESXi host into exit maintenance mode. Thus, all the maintenance tasks will be completed.

Let's discuss another example: one virtual machine is running on ESXi host A, which is managed by the vCenter Server. The VMware administrator discovers that the virtual machine is not responding. They decide to restart this virtual machine through vCenter Server, but this virtual machine still doesn't reboot automatically, even after executing the reboot action. In this situation, you can forcibly reboot and shut down this virtual machine using the VMware CLI in ESXi's **Technical Supported Mode (TSM)** or vMA. The following example is for reference.

You can power down one virtual machine, `VC_TEST`, using the following `esxcli` in ESXi's TSM:



```

172.16.100.136 - PuTTY
software      configuring scheduling related functionality.
storage       Manage the ESXi software image and packages
system        VMware storage commands.
              VMKernel system properties and commands for
              configuring properties of the kernel core system.
vm            A small number of operations that allow a user to
              Control Virtual Machine operations.
vsan          VMware VSAN commands.

~ # esxcli vm
Usage: esxcli vm {cmd} [cmd options]

Available Namespaces:
  process      Operations on running virtual machine processes

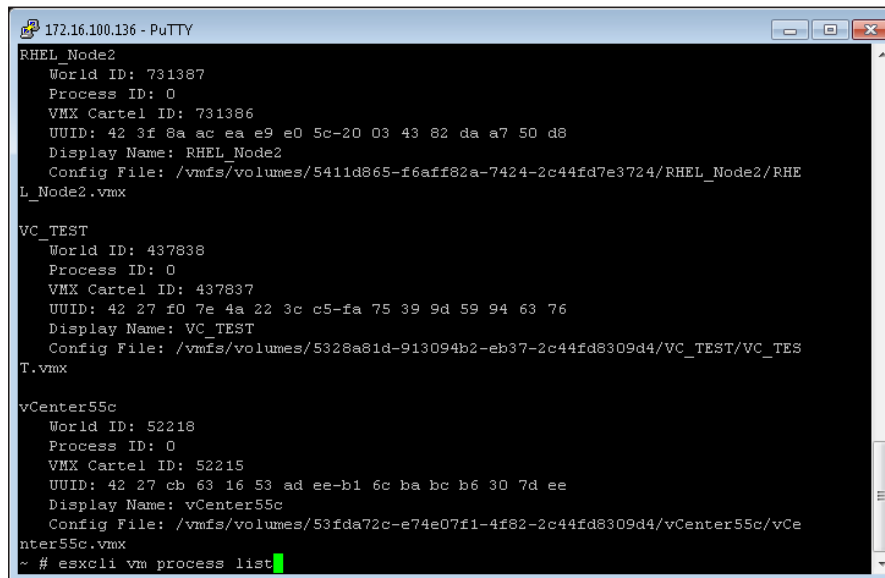
~ # esxcli vm process
Usage: esxcli vm process {cmd} [cmd options]

Available Commands:
  kill         Used to forcibly kill Virtual Machines that are stuck
              and not responding to normal stop operations.
  list         List the virtual machines on this system. This command
              currently will only list running VMs on the system.

~ #

```

1. List all the virtual machines using `esxcli vm process list`. You will notice the **World ID: 437838** of the `VC_TEST` virtual machine:



```

172.16.100.136 - PuTTY
RHEL_Node2
  World ID: 731387
  Process ID: 0
  VMX Cartel ID: 731386
  UUID: 42 3f 8a ac ea e9 e0 5c-20 03 43 82 da a7 50 d8
  Display Name: RHEL_Node2
  Config File: /vmfs/volumes/5411d865-f6aff82a-7424-2c44fd7e3724/RHEL_Node2/RHE
L_Node2.vmx

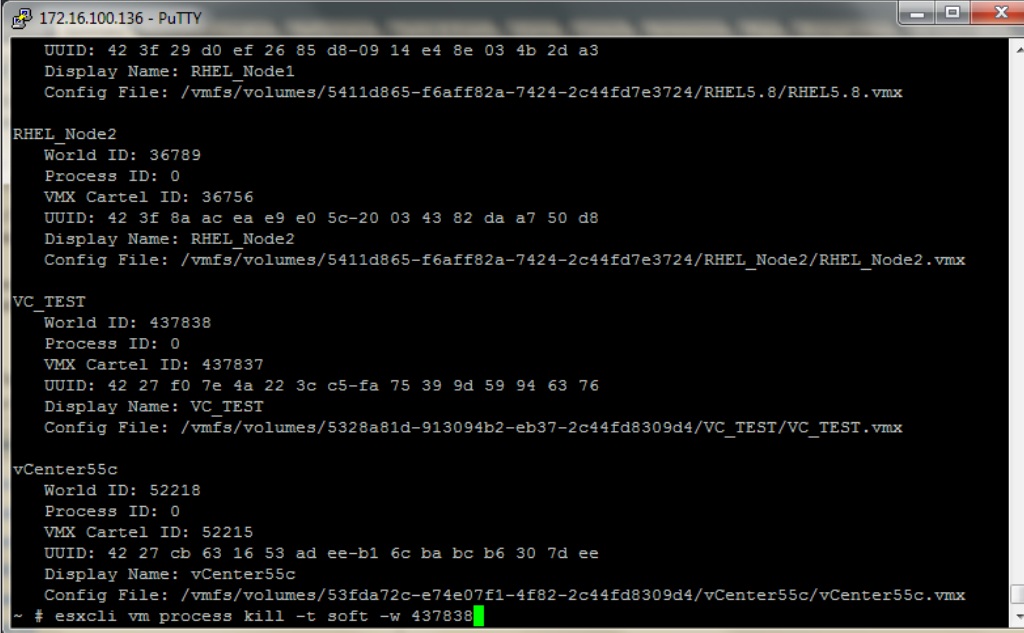
VC_TEST
  World ID: 437838
  Process ID: 0
  VMX Cartel ID: 437837
  UUID: 42 27 f0 7e 4a 22 3c c5-fa 75 39 9d 59 94 63 76
  Display Name: VC_TEST
  Config File: /vmfs/volumes/5328a81d-913094b2-eb37-2c44fd8309d4/VC_TEST/VC_TES
T.vmx

vCenter55c
  World ID: 52218
  Process ID: 0
  VMX Cartel ID: 52215
  UUID: 42 27 cb 63 16 53 ad ee-b1 6c ba bc b6 30 7d ee
  Display Name: vCenter55c
  Config File: /vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/vCenter55c/vCe
nter55c.vmx

~ # esxcli vm process list

```

2. After you have found the virtual machine's world ID, you can power down the VM using the `esxcli vm process kill -t soft -w 437838` command. You may execute the command using the **force** option if the virtual machine still does not power down.



```
172.16.100.136 - PuTTY
UUID: 42 3f 29 d0 ef 26 85 d8-09 14 e4 8e 03 4b 2d a3
Display Name: RHEL_Node1
Config File: /vmfs/volumes/5411d865-f6aff82a-7424-2c44fd7e3724/RHEL5.8/RHEL5.8.vmx

RHEL_Node2
World ID: 36789
Process ID: 0
VMX Cartel ID: 36756
UUID: 42 3f 8a ac ea e9 e0 5c-20 03 43 82 da a7 50 d8
Display Name: RHEL_Node2
Config File: /vmfs/volumes/5411d865-f6aff82a-7424-2c44fd7e3724/RHEL_Node2/RHEL_Node2.vmx

VC_TEST
World ID: 437838
Process ID: 0
VMX Cartel ID: 437837
UUID: 42 27 f0 7e 4a 22 3c c5-fa 75 39 9d 59 94 63 76
Display Name: VC_TEST
Config File: /vmfs/volumes/5328a81d-913094b2-eb37-2c44fd8309d4/VC_TEST/VC_TEST.vmx

vCenter55c
World ID: 52218
Process ID: 0
VMX Cartel ID: 52215
UUID: 42 27 cb 63 16 53 ad ee-b1 6c ba bc b6 30 7d ee
Display Name: vCenter55c
Config File: /vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/vCenter55c/vCenter55c.vmx
~ # esxcli vm process kill -t soft -w 437838
```



soft: These kills will give the VMX process a chance to shut down cleanly.

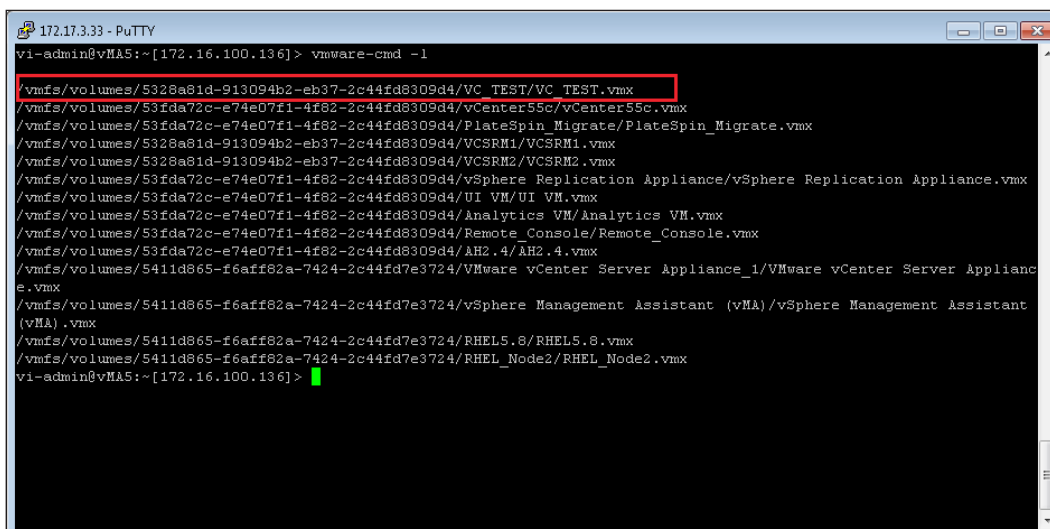
hard: These kills will shut down the process immediately.

force: This should be used as a last resort to kill the VM. If all three of these fail, then the ESXi host reboot is required.

You can also power down the `VC_TEST` virtual machine in vMA using the following remote CLI.

If there is one vMA in the virtual environment, you can connect the ESXi host by vMA and instruct the remote CLI to shut down the virtual machine. For a detailed connection procedure, you can refer to the previous section. You can see which virtual machines are running on this ESXi host using `vmware-cmd -l`. Then you can find the location of the `VC_TEST` virtual machine and check its current status.

1. Firstly, you need to find the location of `VC_TEST.vmx` using the `vmware-cmd -l` command. The result is shown in the following screenshot:



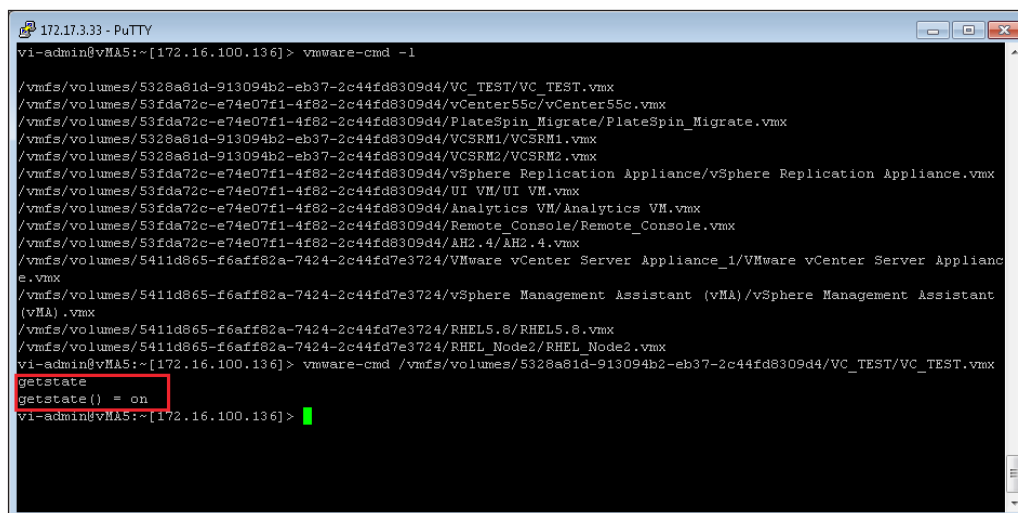
```

172.17.3.33 - PuTTY
vi-admin@vMA5:~[172.16.100.136]> vmware-cmd -l
/vmfs/volumes/5328a81d-913094b2-eb37-2c44fd8309d4/VC_TEST/VC_TEST.vmx
/vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/vCenter55c/vCenter55c.vmx
/vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/PlateSpin_Migrate/PlateSpin_Migrate.vmx
/vmfs/volumes/5328a81d-913094b2-eb37-2c44fd8309d4/VCSRM1/VCSRM1.vmx
/vmfs/volumes/5328a81d-913094b2-eb37-2c44fd8309d4/VCSRM2/VCSRM2.vmx
/vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/vSphere Replication Appliance/vSphere Replication Appliance.vmx
/vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/UI VM/UI VM.vmx
/vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/Analytics VM/Analytics VM.vmx
/vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/Remote_Console/Remote_Console.vmx
/vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/AH2.4/AH2.4.vmx
/vmfs/volumes/5411d865-f6aff82a-7424-2c44fd7e3724/VMware vCenter Server Appliance_1/VMware vCenter Server Appliance.vmx
/vmfs/volumes/5411d865-f6aff82a-7424-2c44fd7e3724/vSphere Management Assistant (vMA)/vSphere Management Assistant (vMA).vmx
/vmfs/volumes/5411d865-f6aff82a-7424-2c44fd7e3724/RHEL5.8/RHEL5.8.vmx
/vmfs/volumes/5411d865-f6aff82a-7424-2c44fd7e3724/RHEL_Node2/RHEL_Node2.vmx
vi-admin@vMA5:~[172.16.100.136]>

```

2. Then check the current status of VM `VC_TEST` using the following command. It shows `getstate() = on`:

```
vmware-cmd /vmfs/volumes/5328a81d-913094b2-eb37-2c44fd8309d4/VC_TEST/VC_TEST.vmx getstate
```



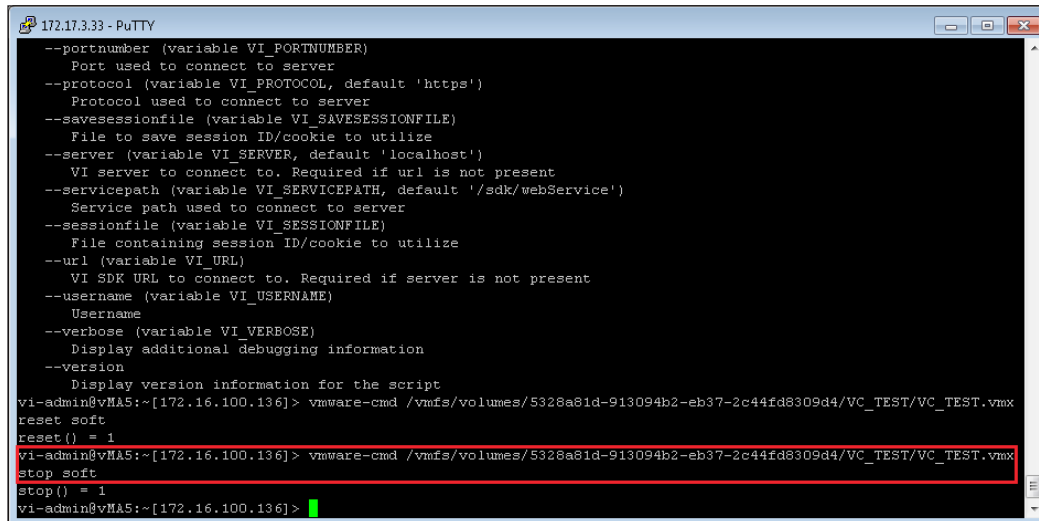
```

172.17.3.33 - PuTTY
vi-admin@vMA5:~[172.16.100.136]> vmware-cmd -l
/vmfs/volumes/5328a81d-913094b2-eb37-2c44fd8309d4/VC_TEST/VC_TEST.vmx
/vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/vCenter55c/vCenter55c.vmx
/vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/PlateSpin_Migrate/PlateSpin_Migrate.vmx
/vmfs/volumes/5328a81d-913094b2-eb37-2c44fd8309d4/VCSRM1/VCSRM1.vmx
/vmfs/volumes/5328a81d-913094b2-eb37-2c44fd8309d4/VCSRM2/VCSRM2.vmx
/vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/vSphere Replication Appliance/vSphere Replication Appliance.vmx
/vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/UI VM/UI VM.vmx
/vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/Analytics VM/Analytics VM.vmx
/vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/Remote_Console/Remote_Console.vmx
/vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/AH2.4/AH2.4.vmx
/vmfs/volumes/5411d865-f6aff82a-7424-2c44fd7e3724/VMware vCenter Server Appliance_1/VMware vCenter Server Appliance.vmx
/vmfs/volumes/5411d865-f6aff82a-7424-2c44fd7e3724/vSphere Management Assistant (vMA)/vSphere Management Assistant (vMA).vmx
/vmfs/volumes/5411d865-f6aff82a-7424-2c44fd7e3724/RHEL5.8/RHEL5.8.vmx
/vmfs/volumes/5411d865-f6aff82a-7424-2c44fd7e3724/RHEL_Node2/RHEL_Node2.vmx
vi-admin@vMA5:~[172.16.100.136]> vmware-cmd /vmfs/volumes/5328a81d-913094b2-eb37-2c44fd8309d4/VC_TEST/VC_TEST.vmx getstate
getstate()
getstate() = on
vi-admin@vMA5:~[172.16.100.136]>

```

- Next, execute the following command to shut down the VM:

```
vmware-cmd /vmfs/volumes/5328a81d-913094b2-eb37-2c44fd8309d4/VC_
TEST/VC_TEST.vmx stop soft
```



The screenshot shows a PuTTY terminal window titled '172.17.3.33 - PuTTY'. It displays the output of the 'vmware-cmd' command. The output includes a list of variables and their values, followed by the command execution. The command 'vmware-cmd /vmfs/volumes/5328a81d-913094b2-eb37-2c44fd8309d4/VC_TEST/VC_TEST.vmx stop soft' is highlighted with a red box. The output shows 'stop()' = 1, indicating the command was successful.

```
--portnumber (variable VI_PORTNUMBER)
Port used to connect to server
--protocol (variable VI_PROTOCOL, default 'https')
Protocol used to connect to server
--savefile (variable VI_SAVESESSIONFILE)
File to save session ID/cookie to utilize
--server (variable VI_SERVER, default 'localhost')
VI server to connect to. Required if url is not present
--servicepath (variable VI_SERVICEPATH, default '/sdk/webService')
Service path used to connect to server
--sessionfile (variable VI_SESSIONFILE)
File containing session ID/cookie to utilize
--url (variable VI_URL)
VI SDK URL to connect to. Required if server is not present
--username (variable VI_USERNAME)
Username
--verbose (variable VI_VERBOSE)
Display additional debugging information
--version
Display version information for the script
vi-admin@vMA5:~[172.16.100.136]> vmware-cmd /vmfs/volumes/5328a81d-913094b2-eb37-2c44fd8309d4/VC_TEST/VC_TEST.vmx
reset soft
reset() = 1
vi-admin@vMA5:~[172.16.100.136]> vmware-cmd /vmfs/volumes/5328a81d-913094b2-eb37-2c44fd8309d4/VC_TEST/VC_TEST.vmx
stop soft
stop() = 1
vi-admin@vMA5:~[172.16.100.136]>
```

From the preceding example, you can learn how to shut down a virtual machine by TSM or vMA.

Reviewing ESXi and vCenter Server logs

In some configuration and troubleshooting tasks, the system log is a very important piece of information for finding the root cause of service corruption. vSphere logs and vCenter logs are at different locations, and can be gathered by different ways. In this section, you will learn how to gather and locate each log. You can generate log bundles for support requests. The methods of gathering logs are listed in the following table:

System logs	Gathering methods
vCenter Server	From vSphere Client
vCenter Server	From a command on vCenter Server itself
vSphere Server	Export the log directly to vCenter Server using vSphere Client
vSphere Server	Use the vm-support command in ESXi's TSM with the help of SSH

The location of the vCenter Server log

The location paths of the vCenter Server log are listed in the following table:

vCenter platform	Location path
vCenter Server 5.x Linux Virtual Appliance	/var/log/vmware/vpx/
vCenter Server 5.x Linux Virtual Appliance UI	/var/log/vmware/vami
vCenter Server 5.x and earlier versions on the Windows 2008 platform	C:\ProgramData\VMware\VMware VirtualCenter\Logs\
vCenter Server 5.x and earlier versions on the Windows 2003 platform	%ALLUSERSPROFILE%\Application Data\VMware\VMware VirtualCenter\Logs\

The descriptions of vCenter logs are given in the following table:

Name	Description
vpzd.log	The main vCenter Server logs and includes vSphere Client and vCenter WebServices events
vpzd-profiler.log, profiler.log, and scoreboard.log	In order to diagnose a vCenter Server issue, you can find these logs in the VPX Operational Dashboard (VOD)
cim-diag.log and vws.log	The communication log between vCenter Server and vSphere Server
vpzd-alert.log	Warning alert for a vpzd process
drmdump\:	The ESXi cluster's DRS log with vCenter Server

The location of the vSphere Server log

This table shows the location paths, with descriptions:

Location path	Description
/var/log/auth.log	ESXi Shell authentication logs
/var/log/dhclient.log	The DHCP Client log
/var/esxupdate.log	ESXi patch and driver update logs
/var/log/shell.log	ESXi Shell usage logs
/var/log/sysboot.log	VMkernel startup and module loading
/var/log/boot.gz	A compressed file that contains a boot log
/var/log/usb.log	Information about the ESXi USB device

Location path	Description
/var/log/syslog.log	This log includes the ESXi service and vCenter schedule tasks
/var/log/vobd.log	VMkernel observation events
/var/log/vmkkernel.log	ESX's core log, which includes all of vSphere's process services
/var/log/vmkwarning.log	The warning message of ESXi's service
/var/log/vmksummary.log	The log summary of all VMkernel, for example, the status of the virtual machine and the ESXi host



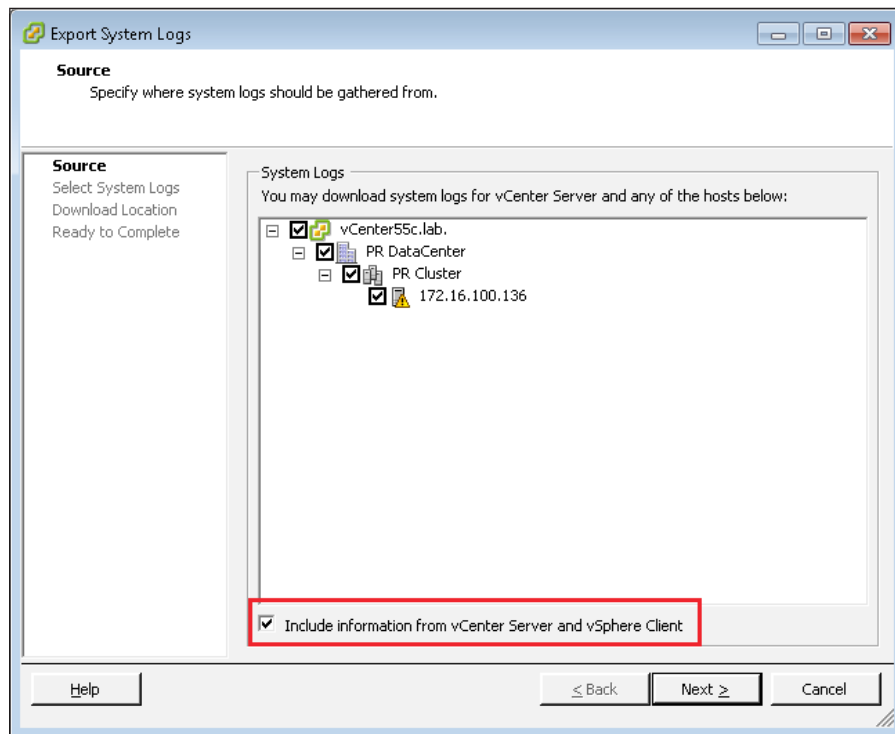
vSphere Server is a hypervisor itself, and runs on a Linux platform. Most of the concerned commands are Linux-based CLI.


Exporting the vm-support log file from vSphere Client

In most situations, the VMware support engineer requests you to provide ESXi's vm support log to analyze when you open a VMware service request. The following procedure shows you how to export the ESXi host's support log from vSphere Client:

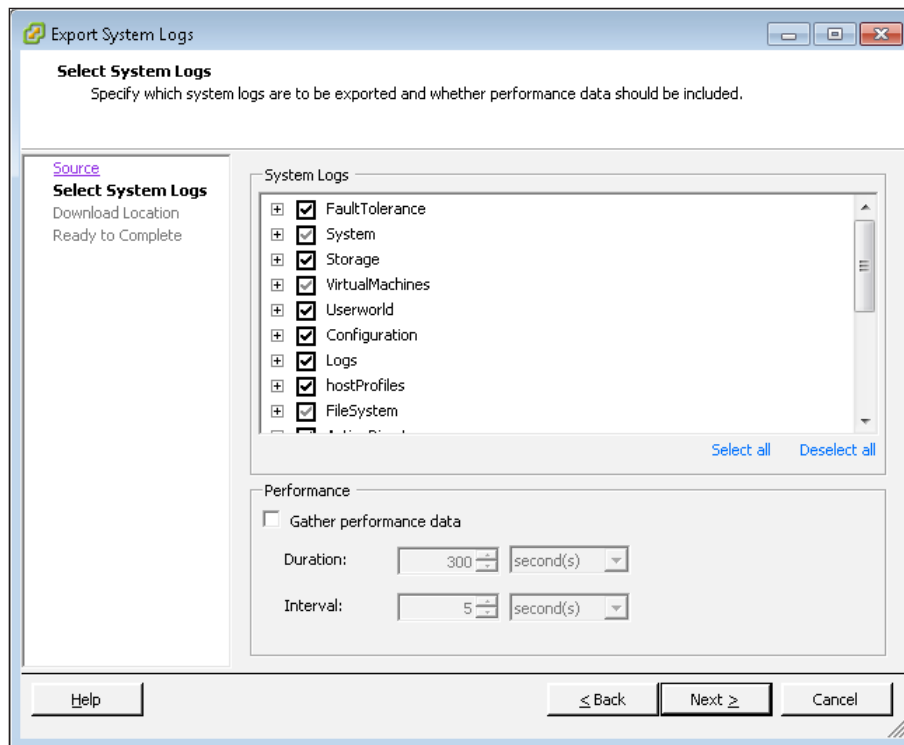
1. Connect to vCenter Server from vSphere Client.
2. Select **File** and choose **Export**.

3. Go to **Export System Logs** and choose the relevant ESXi host log, as shown here:



 It is required to select **Include information from vCenter Server and vSphere Client** if you want to export a vCenter Server log.

4. Select the system logs to export and click on **Next**, as shown in this screenshot:



5. Choose the download location and then select **Finish**.

Exporting the vm-support log file from ESXi Shell

ESXi Shell has other options for collecting vm-support. For specific information, enter `vm-support -help` in the ESXi Shell. Additionally, you can describe some of the common options that can be used. VMware usually requests specific information for troubleshooting in the event of a crash. The following procedure describes how to export ESXi's support log using ESXi commands:

Summary

In this chapter, you learned what the VMware vSphere Management Assistant is, and how to set up and configure with vSphere Server. We also saw how the VMware administrator is able to enable the ESXi Shell during troubleshooting. Then we saw how to use VMware commands (such as `esxcli`, `vifcfg`, and `vmware-cmd`) to manage the vSphere host and virtual machine in vMA or TSM. You also learned how the VMware administrator gathers the vSphere Server log and vCenter Server log.

In the next chapter, we will look at the Virtual Machine Monitor (VMM) and see how to monitor a virtual machine's performance.

3

Using the Virtual Machine Monitor

Performance monitoring in a virtual machine is different from that in a traditional physical server. The **virtual machine monitor (VMM)** is a key component of the VMware vSphere environment, which is a thin layer that provides virtual x86 hardware for virtual machines' (VM) guest operating systems. VMware administrators are needed in order to know how to find performance problems using vSphere monitoring tools. VMware vCenter Server's performance charts and `resxtop` commands, which are common tools, can help you find performance problems in the vSphere Server.

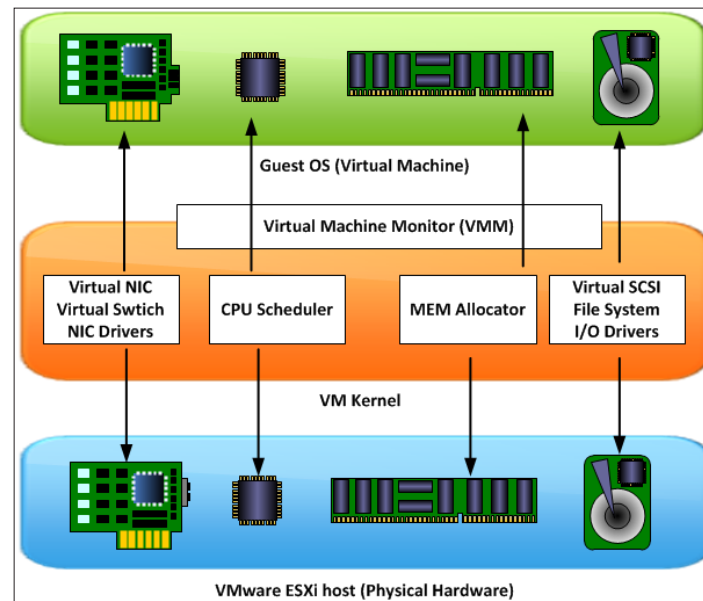
In this chapter, we will cover the following topics:

- The components of the VMware vSphere ESXi architecture
- What the virtual machine monitor is
- The difference between software and hardware virtualization techniques
- Using vSphere Performance monitoring tools

The VMware vSphere ESXi architecture

Firstly, you should understand the VMware vSphere architecture before considering how to optimize the performance of VMware vSphere ESXi. VMware vSphere ESXi (a hypervisor) installs the virtualization layer on an x86-based platform. vSphere ESXi is a platform for running some virtual machines on a single physical machine. Each VM runs on the VMM, which can provide virtual hardware for the guest operating system in the VM. VMware vSphere ESXi uses the VMM to share the physical hardware (for example, CPU, memory, network, storage devices, and so on) with each virtual machine in the VMkernel. Each virtual machine has only one VMM for physical resource sharing.

The following diagram shows the components of the vSphere ESXi host and the guest OS:



There are many different factors responsible for performance issues in a VMware vSphere environment. These factors depend on different hardware and software components, for example, CPU, memory, network, disk I/O, and so on. Multiple virtual machines run on a physical server, with each virtual machine sharing the resources. vSphere ESXi schedules the virtual machine to access the physical server's CPU, memory, network, and storage. If you overcommit any of these resources, you might see performance issues in the virtual machine. For example, an ESXi host has one physical server with a two-core CPU, and three virtual machines are running on it. If you assign four virtual CPUs on each virtual machine, it might have performance issues because all the virtual machines must share the underlying physical CPU. In another example, one virtual machine is required for two physical servers with a two-core CPU, and the running ESXi host has four physical processors with an eight-core CPU, though the physical host can provide total of 32 processing cores. If the virtual disk of the virtual machine is a SQL Server that is running on a SATA 7.2k drive, you might notice that it has disk I/O performance issues on this virtual machine. This is because of a design problem in vSphere. The virtual disk should be stored in an SAS or FC 10/15K drive that can provide enough disk I/O for the virtual machine. Finally, design and configuration issues can affect performance in a VMware vSphere environment.

To get the best results from a VMware ESXi environment, you should follow the best practices of the traditional physical server, for example, the number of vCPUs, the memory, virtual NIC adapters, the RAID level in the virtual disk, disk I/O, and so on. Use the newest hardware to improve virtualization performance, if possible, as newer hardware platforms have the latest features (for example, vCPU, memory, FT, and so on) for improving the system's performance in a VMware ESXi environment.

Once you have identified a performance problem in your VMware ESXi environment, you first need to identify which factor has caused that performance issue, and then find a workaround, or solution. The following table explains the basic troubleshooting options for a VMware ESXi host:

Problem/symptom	Checking	Recommendation
High CPU load	<ul style="list-style-type: none"> • The running status of VMware VM tools on a virtual machine • The number of vCPU assigned to the virtual machine • The number of physical CPUs installed on the ESXi host • The CPU allocation on the ESX resource pool 	<ul style="list-style-type: none"> • Decrease the number of vCPU on the running virtual machine • Install and upgrade the VMware VM tools to the latest version on the virtual machine • According to VMware best practices, upgrade the virtual hardware to the recommended version on the virtual machine
Memory high loading	<ul style="list-style-type: none"> • The running status of the VMware VM tools on the virtual machine • The amount of Memory assigned to the virtual machine • The total amount of memory installed on the ESXi host • The memory allocation on the ESX resource pool 	<ul style="list-style-type: none"> • Reserve the required memory on the virtual machine or ESXi's resource pool • Install and upgrade the VMware VM tools to the latest version on the virtual machine • According to VMware best practices, upgrade the virtual hardware to the recommended version on the virtual machine

Problem/symptom	Checking	Recommendation
Network response time is low	<ul style="list-style-type: none">• The current status of the VMware VM tools on the virtual machine• What is the version of virtual hardware on the virtual machine?• The network speed of the physical NIC on the ESXi host• What is the type of virtual NIC on the virtual machine?• Does it have good performance, for example, CPU, memory, and disk I/O, on the virtual machine?• Identify the bottleneck in the network if it is a network issue	<ul style="list-style-type: none">• According to VMware best practices, choose the recommended virtual NIC adapter on the virtual machine• Install and upgrade the VMware VM tools to the latest version• Add two or more NIC uplinks onto the virtual standard switch or virtual distributed switch• Setup on network I/O control if using a virtual distributed switch• Separate the production network and management and the vMotion network into a different virtual standard switch or virtual distributed switch

Problem/symptom	Checking	Recommendation
Storage disk I/O response time is low	<ul style="list-style-type: none"> • What is RAID level, disk type is using on the ESXi data store? • What path policy of ESXi the data store is using • The ESXi host's connectivity of FC/iSCSI SAN storage • The status of the read/write cache on the SAN storage • Does it have good performance (for example, CPU and memory) on the virtual machine? • Identify the bottleneck in the SAN storage if it is a storage issue • Does it create a VM snapshot on the virtual machine? 	<ul style="list-style-type: none"> • One VMDK is mapping to one disk drive at the OS system level • Use single initiator zoning on the SAN switch if it is using FC SAN storage • Enable the jumbo frame on the virtual switch on the ESXi host and LAN switch if using iSCSI SAN storage • According to best practices of the storage vendor, select the recommended path policy of the data store on the ESXi host • Don't allocate too many virtual machines running on a single ESXi data store • According to the disk I/O, allocate the virtual machine to different ESXi data stores • Don't use a thin disk on the virtual machine if thin provisioning is enabled on the LUN in SAN storage

Understanding the VMM

The VMM is a key component of the VMware vSphere environment. It is a thin layer that provides virtual x86 hardware for a virtual machine's guest operating system. This hardware includes CPU, memory, network, storage, and so on. Each VMM is serviced to one virtual machine; it shares the CPU, memory, NIC driver, and I/O driver with the VM. The VMM can be configured using hardware virtualization, software virtualization, and paravirtualization techniques. Hardware virtualization and software virtualization will be discussed in the next section. Paravirtualization allows software running on a virtual machine to bypass the virtual interface. It runs the operation on the physical hardware.

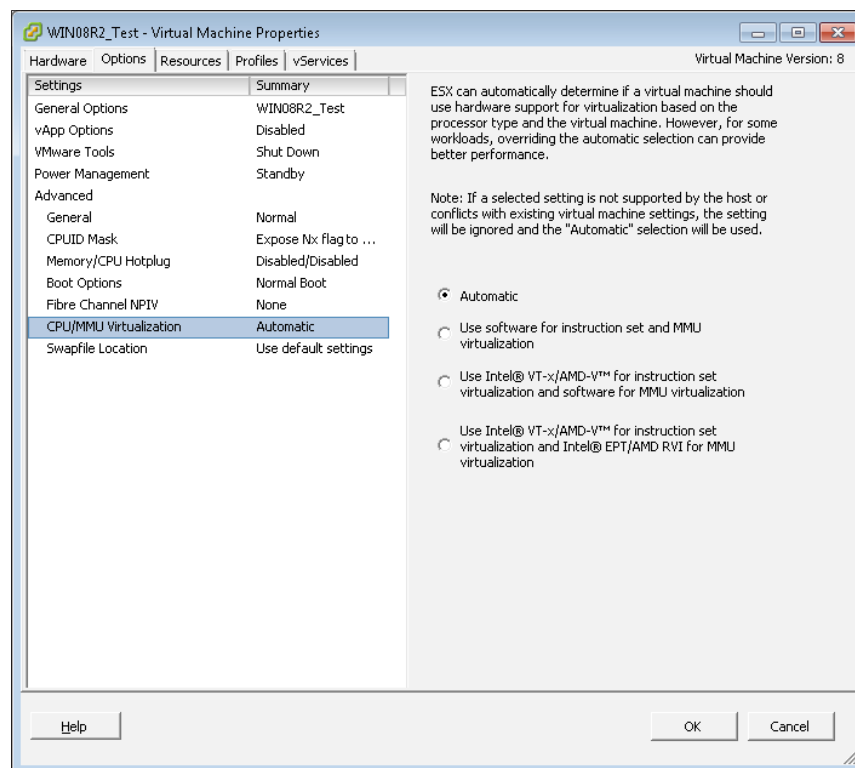
The VMM can choose from three monitor modes:

- Binary translation and shadow page tables
- Intel VT-x or AMD-V and shadow page tables
- AMD-V with RVI or Intel VT-x with EPT



Paravirtualization techniques are not commonly implemented in the production environment because they require modifications to the operation system kernel. This technique is not supported by unmodified operation systems, for example, the Microsoft Windows platform. The details of the options you just saw will be discussed in the next section.

Here is a screenshot showing the default mode of the virtual machine chosen by the VMM:



Software and hardware virtualization techniques

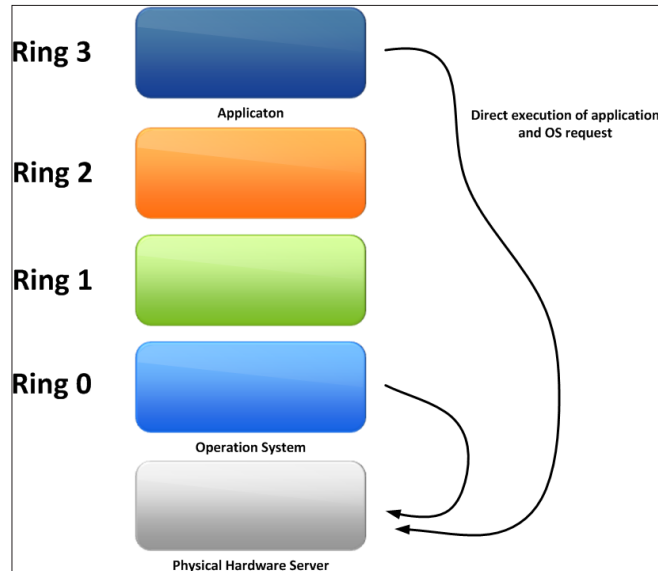
In VMware ESXi vSphere, the virtual CPU consists of the virtual instruction set and virtual **memory management unit (MMU)**. The virtual instruction set is a list of CPU-executed instructions. The virtual MMU is mapped between the virtual address and physical address in the physical memory. The combination of the virtual instruction set and memory is called monitor mode. The VMM can identify the ESXi host hardware and its CPU model. Then it chooses monitor mode for the virtual machine on that hardware platform. It can also configure the monitor using software techniques, hardware techniques, or a combination of both techniques.

The following table lists the difference between CPU software virtualization and hardware virtualization:

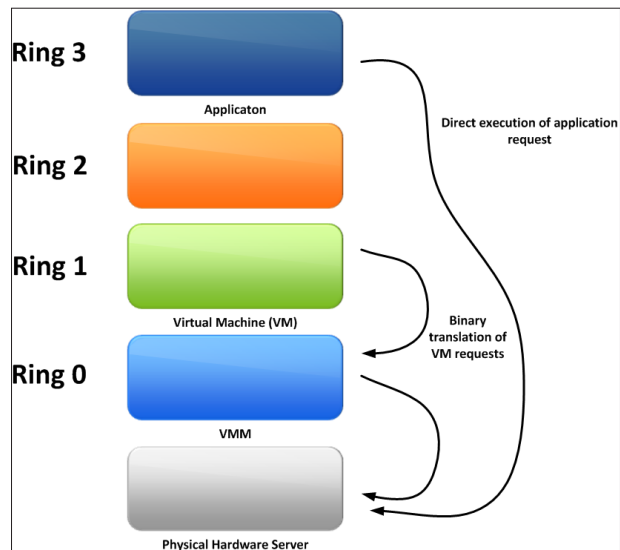
	Advantage	Disadvantage
CPU software virtualization	<ul style="list-style-type: none">• When this doesn't have enough CPU resources, performance is unaffected	<ul style="list-style-type: none">• There is an increase in host CPU utilization• There is an increase in virtual machine latency
CPU hardware virtualization	<ul style="list-style-type: none">• Virtual machines can directly request or access the hardware with VMM	<ul style="list-style-type: none">• This has a large cost if the CPU doesn't support hardware virtualization

The architecture of an x86 operating system has four levels for operation of the system: **Ring 0** to **Ring 3**. The user application runs in **Ring 3**. The operating system needs to directly access physical hardware; it executes in **Ring 0**.

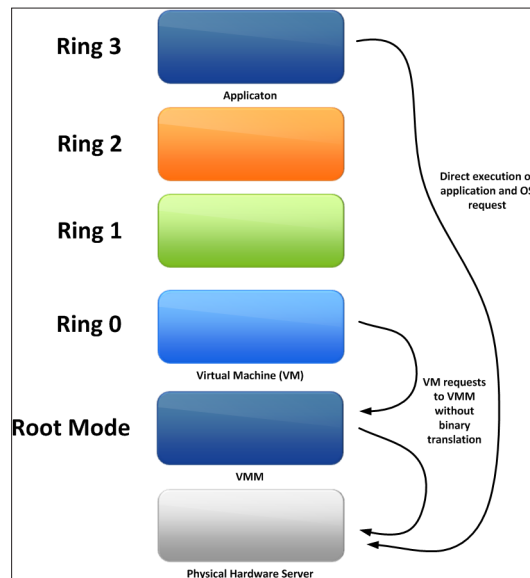
The following diagram shows the x86 architecture:



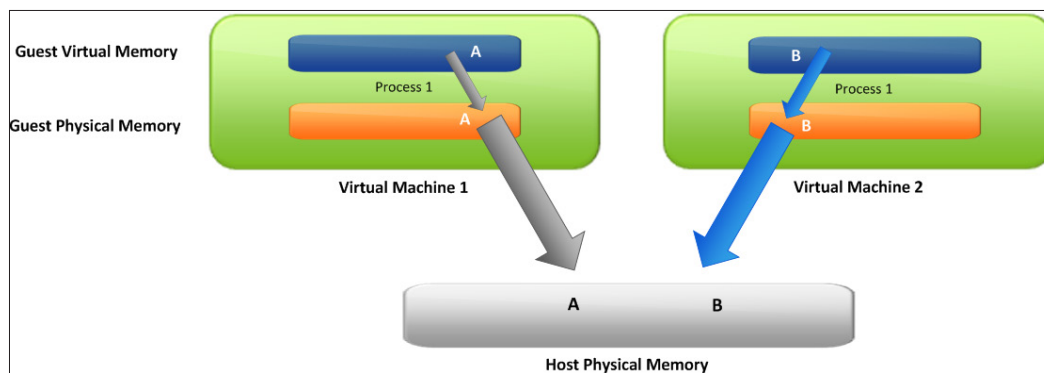
The VMware ESXi host virtualizes the x86 operating system using a combination of binary translation and direct execution techniques. Binary translation allows the VMM to run in **Ring 0** for isolation and performance. The virtual machine runs in **Ring 1**. The VMM translator all virtual machine instruction and caches the results for the execution. The user application continues to run in **Ring 3**. It is directly executed on the physical processor. This technique is called CPU software virtualization. The following diagram shows the details of this architecture:



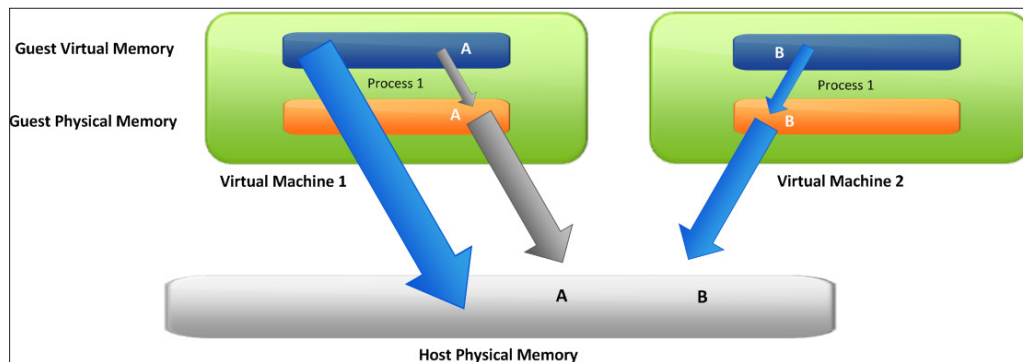
Now, hardware virtualization is available on **Intel Virtualization Technology (Intel VT-x)** and **AMD Virtualization (AMD-V)**. Both of these technologies allow a VMM to fully control the execution of a virtual machine, without binary translation, and both allow the VMM to run in root mode below **Ring 0**. The **Virtual Machine (VM)** can run in **Ring 0**. The virtual machine's state is stored in the virtual machine structure (Intel VT-x) or virtual machine blocks (AMD-V). This technique is called CPU hardware virtualization. The following diagram shows this architecture in detail:



After discussing CPU virtualization, the next topic to be covered is MMU. Memory virtualization shares the physical memory and dynamically allocates it to the virtual machine. The virtual machine uses page tables to map the virtual memory address to the physical memory address. The MMU translates the virtual address into a physical address. The **translation look-aside buffer (TLB)** is a cache that speeds up these translations in MMU. When multiple virtual machines are running on a single host, the virtual machine controls the mapping of the virtual address to the physical address, but it cannot directly access the host physical memory. The MMU maps the guest virtual memory (vRAM) to the host physical memory. To support the VM, the MMU must be virtualized by both the software technique and the hardware technique. This diagram depicts the architecture of the MMU:



MMU virtualization uses shadow page tables, which are composed of two mappings. The first mapping is of the **Guest Virtual Memory** to the **Guest Physical Memory**. This mapping is obtained from the primary page table. The second mapping is of the **Guest physical memory** to the **Host Physical Memory**. This mapping is defined by the VMM and VMkernel. The VMM can directly point to the hardware MMU at the shadow page tables. This allows the virtual machine to access the host physical memory. The following diagram shows the architecture of software MMU virtualization:



Both the platforms – Intel and AMD – support hardware MMU virtualization. **Extended Page Tables (EPT)** is a feature of the Intel platform. **Rapid Virtualization Indexing (RVI)** is a feature included in the AMD platform. Both EPT and RVI are not required for the VMM to synchronize the shadow page table with the guest page table.

The following table lists the differences between software MMU virtualization and hardware MMU virtualization:

	Advantage	Disadvantage
Software MMU virtualization	Improves memory performance by using shadow page tables	Shadow page tables consume extra memory and increase CPU overhead
Hardware MMU virtualization	Performance is better than that of software MMU virtualization	Increased costs

Using vSphere performance monitoring tools

VMware performance charts and **esxtop** are common monitoring tools in VMware vSphere. You can display performance charts in either the VMware ESXi host or VMware vCenter Server. Performance charts can provide a lot of useful information for analysis of performance problems, for example, CPU and memory utilization, disk I/O, network bandwidth, and so on. VMware performance charts can display two types of charts: overview charts and advanced charts. The overview performance shows the host performance statistics.

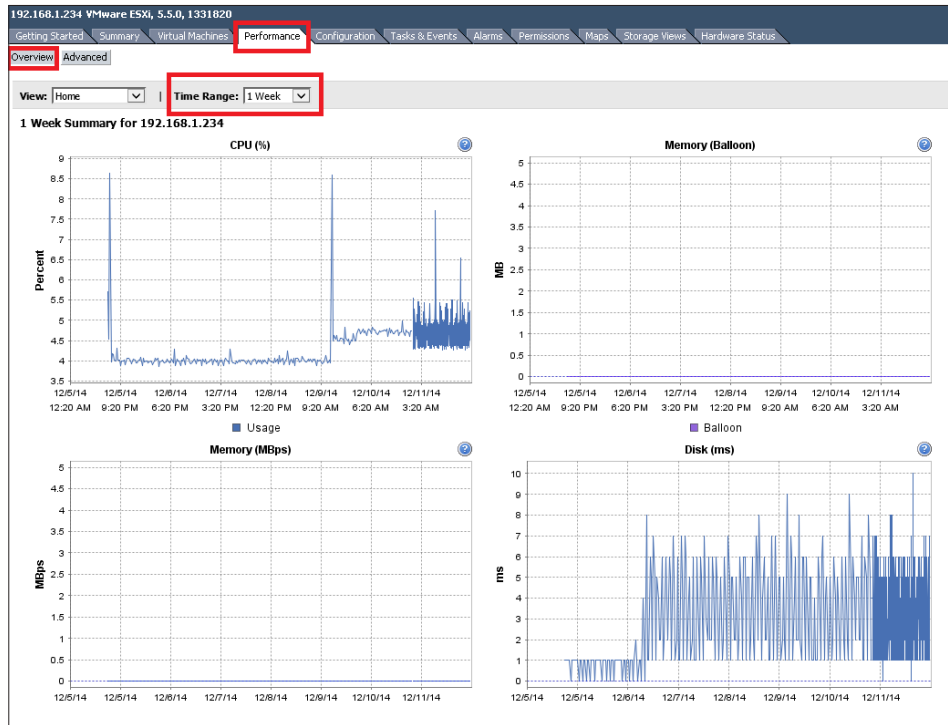
Here is a screenshot of an overview performance chart on VMware vCenter Server:



The following procedure describes how to view an overview performance chart on VMware vCenter Server using the vSphere Client:

1. Log in to vCenter Server using the vSphere Client.
2. Choose the **Performance** tab at the top, and then select **Overview**.

3. You can view these overview performance charts by selecting **Time Range**.

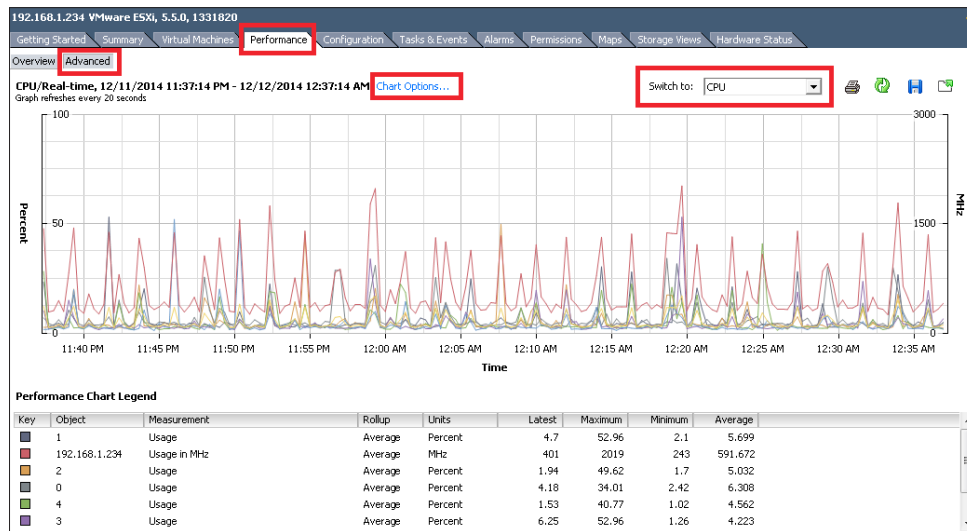


Advanced performance charts can display graphical statistical data for the VMware ESXi host or vCenter Server. You can collect data about ESXi hosts, clusters, the virtual machine, the resource pool, and data stores. Also, you can create and save customized charts in real time or for time frames of the past.

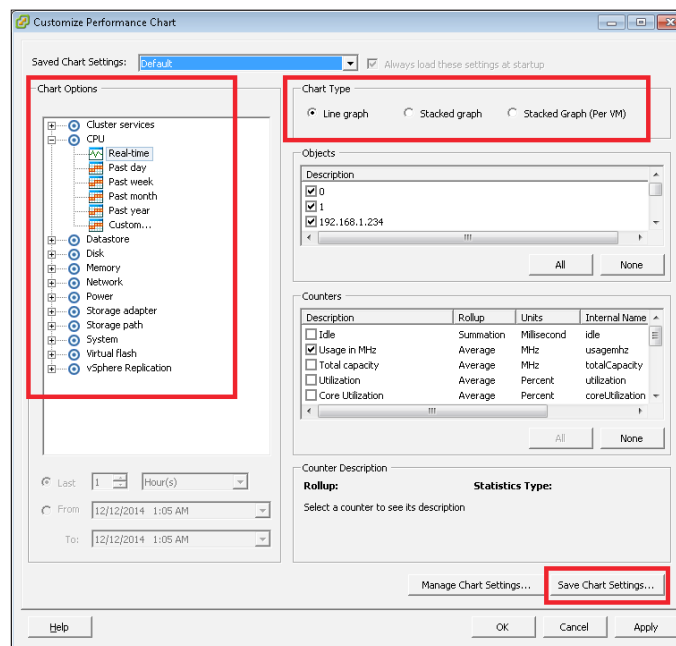
The following procedure describes how to view advanced performance charts and save customized charts in VMware vCenter Server:

1. Log in to vCenter Server using the vSphere Client.
2. Choose the **Performance** tab at the top, and then select **Advanced**.


- You can view the advanced performance chart by selecting different options from the **Switch to** menu, for example, **CPU**, **Memory**, **DataStore**, or **Network**.



- Choose the **Chart Options...** link. Then select the option settings shown as follows:



5. Choose **Chart Type**, **Chart Options**, and **Object**. Then click on the **Save Chart Settings...** button to enter the name of the charts.

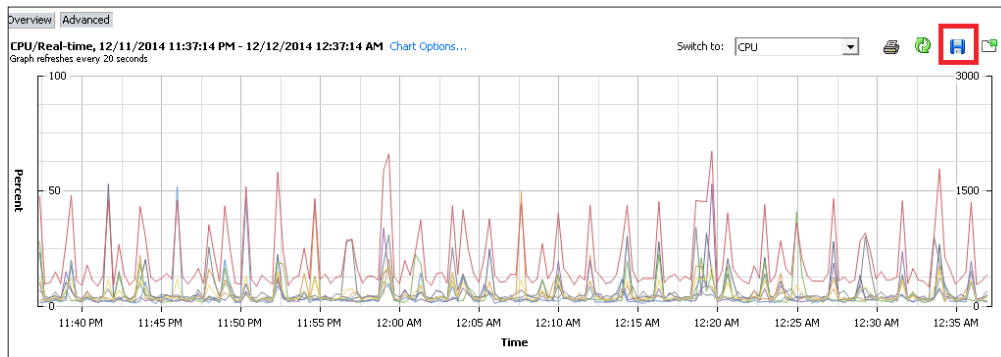



Line graph: Displays the metrics for a single inventory object

Stacked graph: This is used to compare data across virtual machines

Stacked graph (Per VM): This is used to compare data between virtual machines

6. Next, click on **OK** to come back to the advanced performance chart.
7. Finally, export the advanced performance chart by clicking on the disk button, as shown here:





The advanced performance chart can be exported in these formats: PNG, JPEG, BMP, GIF, and CSV.

The `esxtop` utility and the `resxtop` utility are commands that can monitor and collect resource data from the VMware ESXi host, for example, CPU, memory, disk, network, and so on. This data includes some metrics that can't be accessed using overview or advanced performance charts. The `esxtop` utility runs only on the ESXi host's service console and executes it using root user privileges. The `resxtop` must execute by the vSphere CLI or vMA. It has three modes for collecting the data, as follows:

- **Default mode (interactive mode):** It can collect all statistical data for the VMware ESXi host in real time
- **Batch mode:** Collects the statistic data and saves it in a file
- **Replay mode:** The data is collected using the `vm-support` command

The following procedure describes how to execute `esxtop`:

1. Log in to the ESXi host by SSH with root user privileges.



The SSH service has to be enabled.

In vSphere ESXi 5.x, you can use SSH to run the `esxtop` command; however, the `resxtop` command requires use of the vSphere CLI.

2. Input the `esxtop` command in the ESXi host's service console. Then you can enter the following screen:

```

12:49:40am up 31 days 9:14, 614 worlds, 7 VMs, 14 vCPUs; CPU load average: 0.03, 0.03, 0.03
PCPU USED(%): 2.3 0.0 4.8 0.2 4.9 0.0 0.0 0.4 0.5 0.0 0.2 37 6.1 0.5 57 0.1 0.1 2.5 2.6 1.6 0.7 0.0 0.2 0.1 AVG: 5.
PCPU UTIL(%): 10 0.0 10 0.0 11 0.0 0.0 0.0 0.5 0.0 0.0 61 12 1.6 100 0.0 0.0 6.5 6.3 0.0 1.9 0.0 0.0 5.5 AVG: 9.
CORE UTIL(%): 10 11 12 0.0 0.5 71 13 100 6.7 8.9 1.9 0.0 AVG: 1
  
```

ID	GID	NAME	NWLD	%USED	%RUN	%SYS	%WAIT	%VMWAIT	%RDY	%IDLE	%OVRLP	%CSTP	%ML
15314474	15314474	esxtop.7962250	1	40.95	49.74	0.00	0.00	-	0.00	0.00	0.01	0.00	0
2001	2001	hostd.34009	26	22.95	40.40	0.00	1403.12	-	0.01	0.00	0.03	0.00	0
7792935	7792935	VC TEST	8	3.15	6.43	0.08	428.84	0.00	0.56	102.00	0.05	0.00	0
1322298	1322298	RHEL_Node1	9	2.85	5.45	0.28	463.81	0.71	0.64	97.65	0.05	0.00	0
1327745	1327745	RHEL_Node2	9	2.81	4.44	0.16	545.96	0.98	0.35	116.71	0.04	0.00	0
10137251	10137251	vCenter55c	8	2.51	4.48	0.02	388.58	0.03	0.08	93.68	0.03	0.00	0
2	2	system	146	1.69	2.90	0.10	10115.20	-	0.72	0.00	0.06	0.00	0
6221	6221	AH2.4	11	0.35	0.75	0.00	503.15	16.79	0.20	165.48	0.01	0.00	0
7787435	7787435	PlateSpin_Migra	7	0.19	0.25	0.00	322.37	4.48	0.01	41.34	0.00	0.00	0
8	8	helper	163	0.16	0.34	0.00	11038.22	-	0.15	0.00	0.00	0.00	0
6407	6407	vSphere Managem	6	0.16	0.17	0.02	270.71	7.26	0.03	37.73	0.00	0.00	0
915	915	vmssyslogd.33323	3	0.14	0.31	0.00	181.25	-	0.04	0.00	0.00	0.00	0
15313117	15313117	sshd.7961574	1	0.06	0.06	0.00	54.70	-	0.03	0.00	0.00	0.00	0
2437	2437	rhttpproxy.3423	12	0.01	0.02	0.00	648.44	-	0.06	0.00	0.00	0.00	0
3326	3326	vpca.34696	12	0.01	0.02	0.00	630.60	-	0.06	0.00	0.00	0.00	0
4550	4550	sfcb-ProviderMa	10	0.01	0.02	0.00	467.68	-	0.01	0.00	0.00	0.00	0
1350	1350	net-lacp.33581	3	0.01	0.01	0.00	173.21	-	0.01	0.00	0.00	0.00	0
1	1	idle	24	0.00	2288.37	0.00	0.00	-	2400.00	0.00	1.81	0.00	0
9	9	drivers	13	0.00	0.00	0.00	869.51	-	0.00	0.00	0.00	0.00	0
10	10	ft	5	0.00	0.00	0.00	333.92	-	0.00	0.00	0.00	0.00	0
11	11	vmotion	1	0.00	0.00	0.00	66.74	-	0.00	0.00	0.00	0.00	0
392	392	init.33084	1	0.00	0.00	0.00	61.35	-	0.00	0.00	0.00	0.00	0



The `-a` option shows all the statistics.

The `-b` option is executed in batch mode.

The `-h` option is used for help on the `esxtop` command.

In batch mode, `esxtop` runs until it produces the number of iterations requested (which was done using the `-n` option), or until you end the process by pressing `Ctrl + C`.

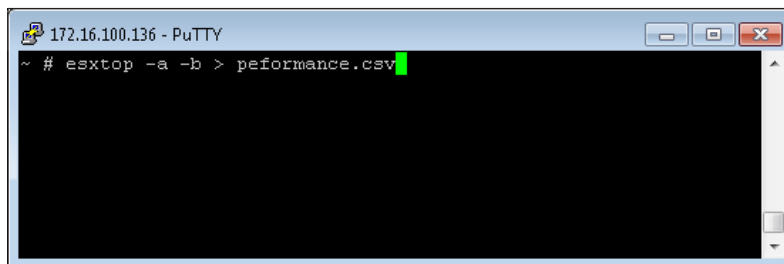
You can type a character to go to a different screen. These characters are listed in the following table:

Characters	Description
c	Switch to the CPU resource utilization screen
m	Switch to the memory resource utilization screen
d	Switch to ESXi's storage adapter resource utilization screen
u	Switch to the storage device (path) resource utilization screen
v	Switch to the virtual disk resource utilization screen
V	Display the virtual machines only
n	Switch to the network resource utilization screen
h	Display the help screen
q	Quite the esxtop utility mode


The following command is used to execute `esxtop` in batch mode:

```
esxtop -a -b > "output-name.csv"
```

Here is a screenshot showing the command being typed:



The following procedure is used to execute `esxtop` in replay mode:

-  The SSH service has to be enabled.

- ```
172.16.100.136 - PuTTY
~ # vm-support -p -d 300 -i 40
05:47:26: Creating /var/tmp/esx-esxi55b.boardware.com.mo-2014-12-12--05.47.tgz
05:48:56: Gathering output from /bin/cim-diagnostic.sh
```



The `-p` option is used to collect performance data

The `-d` option gives the duration of performance monitoring (in seconds)

The `-i` option gives the interval between performance snapshots (in seconds)

The `-h` option is used for help on `vm-support`

3. After collecting the `vm-support` performance file, the file is stored at the location marked in this screenshot:

```
172.16.100.136 - PuTTY
05:51:21: Adding /vmfs/volumes/5411d865-f6aff82a-7424-2c44fd7e3724/VMware vCenter Server Appliance_1/vmware.1
05:51:21: Adding /vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/PlateSpin_Migrate/vmware-0.log
05:51:22: Adding /vmfs/volumes/5411d865-f6aff82a-7424-2c44fd7e3724/vSphere Management Assistant (vMA)/vmware.
05:51:23: Adding /vmfs/volumes/5411d865-f6aff82a-7424-2c44fd7e3724/vSphere Management Assistant (vMA)/vmware.
05:51:23: Adding /vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/vSphere Replication Appliance/vmware.log
05:51:25: Adding /vmfs/volumes/5411d865-f6aff82a-7424-2c44fd7e3724/VMware vCenter Server Appliance_1/VMware v
05:51:25: Adding /vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/PlateSpin_Migrate/PlateSpin_Migrate.vmx
05:51:25: Adding /vmfs/volumes/5411d865-f6aff82a-7424-2c44fd7e3724/vSphere Management Assistant (vMA)/vSphere
05:51:25: Adding /vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/vSphere Replication Appliance/vSphere Repl
05:51:25: Adding /vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/AH2.4/AH2.4.vmx
05:51:25: Adding /vmfs/volumes/5411d865-f6aff82a-7424-2c44fd7e3724/VMware vCenter Server Appliance_1/VMware v
05:51:25: Adding /vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/PlateSpin_Migrate/PlateSpin_Migrate.vmx
05:51:25: Adding /vmfs/volumes/5411d865-f6aff82a-7424-2c44fd7e3724/vSphere Management Assistant (vMA)/vSphere
05:51:25: Adding /vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/vSphere Replication Appliance/vSphere Repl
05:51:25: Adding /vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/AH2.4/AH2.4.vmx
05:51:25: Adding /vmfs/volumes/5411d865-f6aff82a-7424-2c44fd7e3724/VMware vCenter Server Appliance_1/VMware v
05:51:25: Adding /vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/PlateSpin_Migrate/PlateSpin_Migrate.vmad
05:51:25: Adding /vmfs/volumes/5411d865-f6aff82a-7424-2c44fd7e3724/vSphere Management Assistant (vMA)/vSphere
05:51:25: Adding /vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/vSphere Replication Appliance/vSphere Repl
05:51:25: Adding /vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/AH2.4/AH2.4.vmsd
05:51:25: Gathering output from /bin/ls -la /vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/Remote_Console
05:51:25: Gathering output from /bin/ls -la /vmfs/volumes/5411d865-f6aff82a-7424-2c44fd7e3724/VMware vCenter
05:51:25: Gathering output from /bin/ls -la /vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/PlateSpin_Migra
05:51:25: Gathering output from /bin/ls -la /vmfs/volumes/5411d865-f6aff82a-7424-2c44fd7e3724/RHEL5.8
05:51:25: Gathering output from /bin/ls -la /vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/Analytics VM
05:51:25: Gathering output from /bin/ls -la /vmfs/volumes/5411d865-f6aff82a-7424-2c44fd7e3724/vSphere Managem
05:51:25: Gathering output from /bin/ls -la /vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/vSphere Replica
05:51:25: Gathering output from /bin/ls -la /vmfs/volumes/53fda72c-e74e07f1-4f82-2c44fd8309d4/AH2.4
05:56:57: Done.
Please attach this file when submitting an incident report.
To file a support incident, go to http://www.vmware.com/support/sr/sr_login.jsp
To see the files collected, run: tar -tzf '/var/tmp/esx-esxi55b.boardware.com.mo-2014-12-12--05.47.tgz'
```

4. Unzip or untar the `vm-support` file. Finally, run this:  
`esxtop -R/var/tmp/esx-esxi55b.boardware.com.mo-2014-12-12--05.47.tgz`

## Summary

In this chapter, you learned about the components of the VMware vSphere ESXi architecture, what VMM is, and how to configure software and hardware techniques with VMM in the virtual machine. We then covered how to view the vCenter performance chart and use the esxtop utility to analyze any performance problems.

In the next chapter, we will look at VMware vSphere APIs for array integration (VAAI) and VMware vSphere APIs for Storage Awareness (VASA). Also, you will learn how to configure the vSphere Storage DRS and storage I/O control in a vSphere ESXi host.

# 4

## Storage Scalability

SAN storage is a key component of a VMware vSphere environment. We can choose different vendors and types of SAN storage to deploy on a VMware Sphere environment. The advanced settings of each storage can affect the performance of the virtual machine, for example, FC or iSCSI SAN storage. It has a different configuration in a VMware vSphere environment. Host connectivity of Fibre Channel storage is accessed by **Host Bus Adapter (HBA)**. Host connectivity of iSCSI storage is accessed by the TCP/IP networking protocol. We first need to know the concept of storage. Then we can optimize the performance of storage in a VMware vSphere environment.

In this chapter, you will learn these topics:

- What the vSphere storage APIs for **Array Integration (VAAI)** and **Storage Awareness (VASA)** are
- The virtual machine storage profile
- VMware vSphere Storage DRS and VMware vSphere Storage I/O Control

### vSphere storage APIs for array integration and storage awareness

In *Chapter 3, Using the Virtual Machine Monitor*, you gained knowledge about performance monitoring in a VMware vSphere environment. VMware vMotion is a key feature in vSphere hosts. An ESXi host cannot provide the vMotion feature if it is without shared SAN storage. SAN storage is a key component in a VMware vSphere environment. In large-scale virtualization environments, there are many virtual machines stored in SAN storage. When a VMware administrator executes virtual machine cloning or migrates a virtual machine to another ESXi host by vMotion, this operation allocates the resource on that ESXi host and SAN storage.



In vSphere 4.1 and later versions, it can support VAAI. The vSphere storage API is used by a storage vendor who provides hardware acceleration or offloads vSphere I/O between storage devices. These APIs can reduce the resource overhead on ESXi hosts and improve performance for ESXi host operations, for example, vMotion, virtual machine cloning, creating a virtual machine, and so on. VAAI has two APIs: the hardware acceleration API and the array thin provisioning API.

The hardware acceleration API is used to integrate with VMware vSphere to offload storage operations to the array and reduce the CPU overload on the ESXi host.

The following table lists the features of the hardware acceleration API for block and NAS:

| Array integration | Features                | Description                                                                                                                       |
|-------------------|-------------------------|-----------------------------------------------------------------------------------------------------------------------------------|
| Block             | Fully copy              | This blocks clone or copy offloading.                                                                                             |
|                   | Block zeroing           | This is also called write same. When you provision an eagerzeroedthick VMDK, the SCSI command is issued to write zeroes to disks. |
|                   | Atomic Test & Set (ATS) | This is a lock mechanism that prevents the other ESXi host from updating the same VMFS metadata.                                  |
| NAS               | Full file clone         | This is similar to Extended Copy (XCOPY) hardware acceleration.                                                                   |
|                   | Extended statistics     | This feature is enabled in space usage in the NAS data store.                                                                     |
|                   | Reserved space          | The allocated space of virtual disk in thick format.                                                                              |

The array thin provisioning API is used to monitor the ESXi data store space on the storage arrays. It helps prevent the disk from running out of space and reclaims disk space. For example, if the storage is assigned as 1 x 3 TB LUN in the ESXi host, but the storage can only provide 2 TB of data storage space, it is considered to be 3 TB in the ESXi host. Streamline its monitoring LUN configuration space in order to avoid running out of physical space.

When vSphere administrators delete or remove files from the data store that is provisioned LUN, the storage can reclaim free space in the block level.

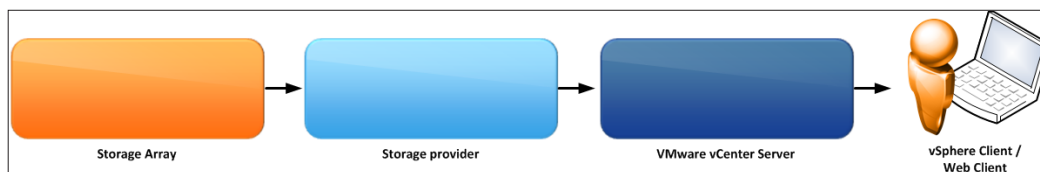


In vSphere 4.1 or later, it can support VAAI features.

In vSphere 5.5, you can reclaim the space on thin provisioned LUN using `esxcli`.

VMware VASA is a piece of software that allows the storage vendor to provide information about their storage array to VMware vCenter Server. The information includes storage capability, the state of physical storage devices, and so on. vCenter Server collects this information from the storage array using a software component called VASA provider, which is provided by the storage array vendor. A VMware administrator can view the information in VMware vSphere Client / VMware vSphere Web Client. The following diagram shows the architecture of VASA with vCenter Server. For example, the VMware administrator requests to create a 1 x data store in VMware ESXi Server. The flow of this operation is shown in the following main components in this diagram.

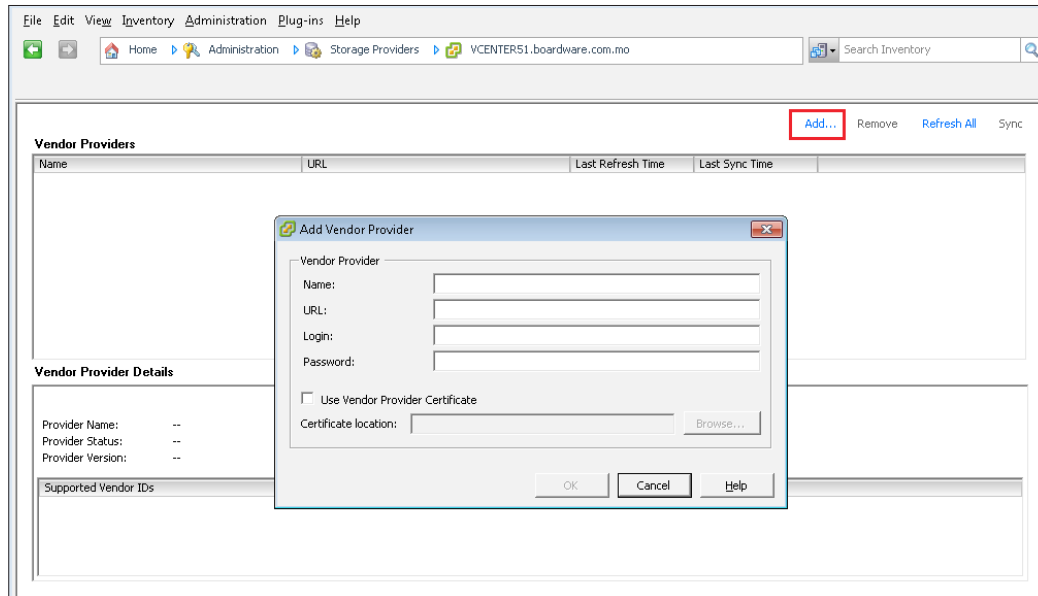
It has three main components: the storage array, the storage provider and VMware vCenter Server.



The following is the procedure to add the storage provider to vCenter Server:


1. Log in to vCenter by vSphere Client.
2. Go to **Home | Storage Providers**.
3. Click on the **Add** button.

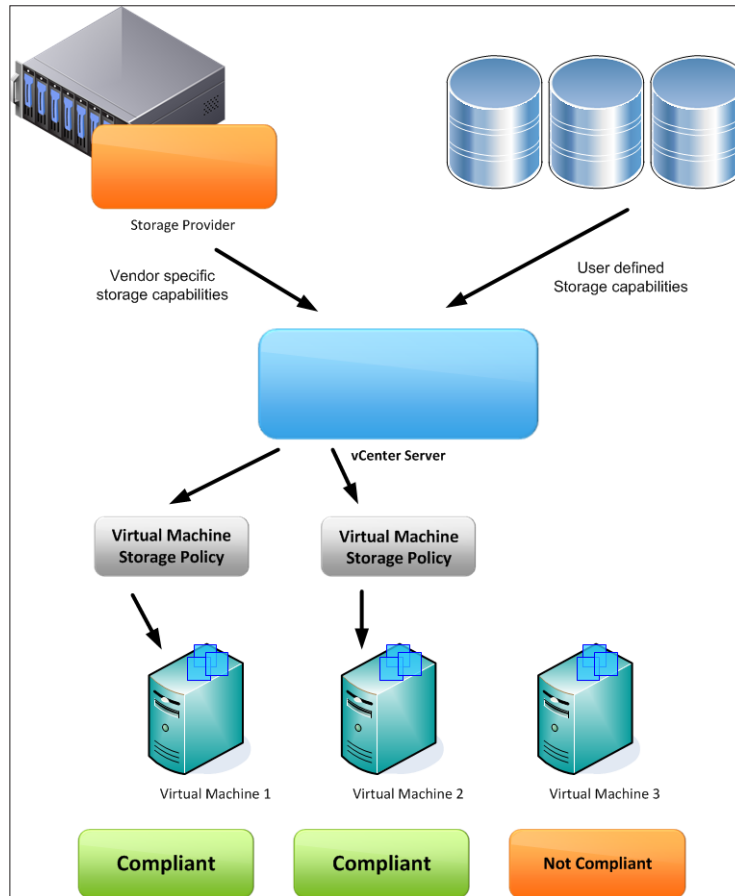
4. Input information about the storage vendor name, URL, and credentials.




## Virtual machine storage profile

The storage provider can help the vSphere administrator know the state of the physical storage devices and the capabilities on which their virtual machines are located. It also helps choose the correct storage in terms of performance and space by using virtual machine storage policies. A virtual machine storage policy helps you ensure that a virtual machine guarantees a specified level of performance or capacity of storage, for example, the SSD/SAS/NL-SAS data store, spindle I/O, and redundancy. Before you define a storage policy, you need to specify the storage requirement for your application that runs on the virtual machine. It has two types of storage requirement, which is storage-vendor-specific storage capability and user-defined storage capability. Storage-vendor-specific storage capability comes from the storage array. The storage vendor provider informs vCenter Server that it can guarantee the use of storage features by using storage-vendor-specific storage capability. vCenter Server assigns vendor-specific storage capability to each ESXi data store. User-defined storage capability is the one that you can define and assign storage profile to each ESXi datastore. The following diagram shows the architecture of a virtual machine storage policy.

 In vSphere 5.1/5.5, the name of the storage policy is VM storage profile.

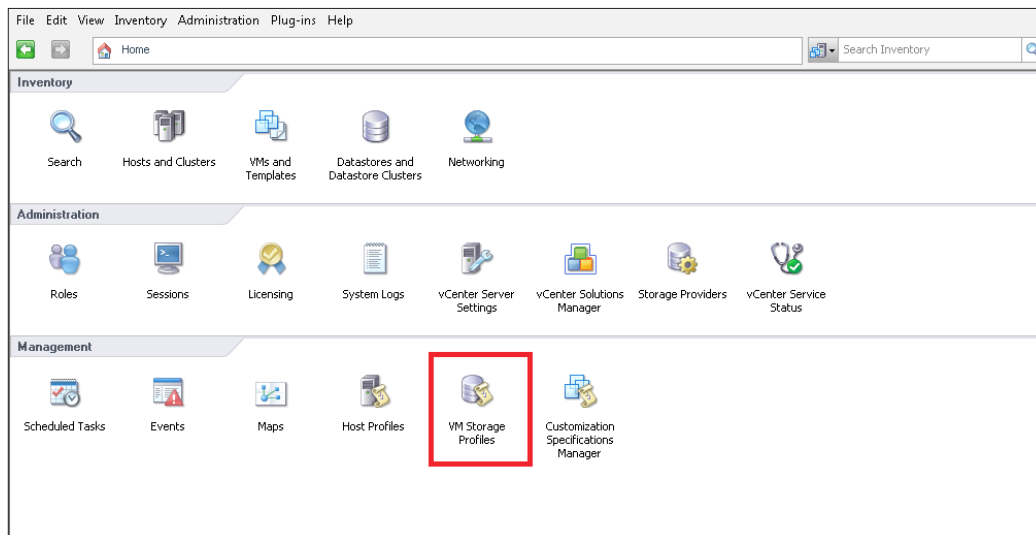


Virtual machine storage policies can include one or more storage capabilities and assign to one or more VM. The virtual machine can be checked for storage compliance if it is placed on compliant storage. When you migrate, create, or clone a virtual machine, you can select the storage policy and apply it to that machine. The following procedure shows how to create a storage policy and apply it to a virtual machine in vSphere 5.1 using user-defined storage capability:

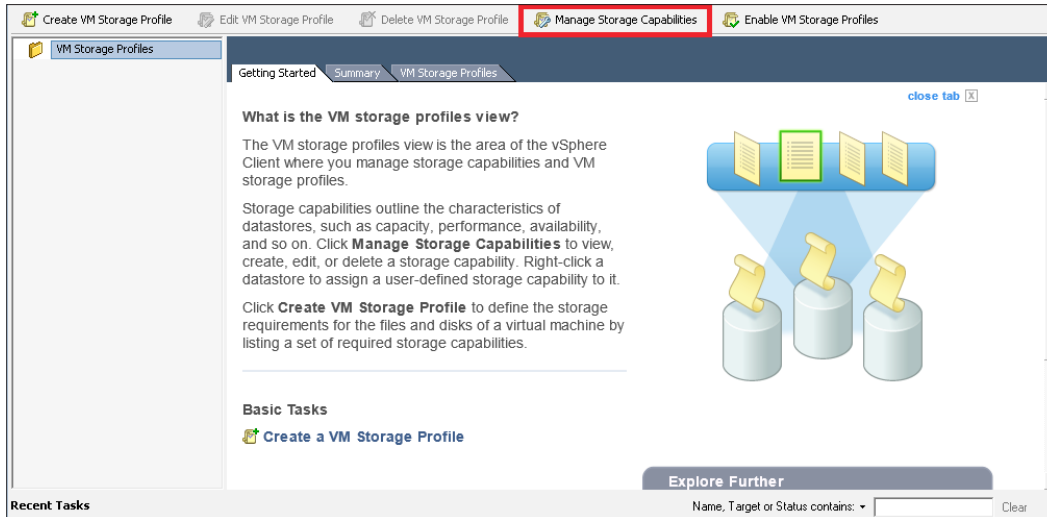
 The vSphere ESXi host requires the license edition of Enterprise Plus to enable the VM storage profile feature.

The following procedure is adding the storage profile into vCenter Server:

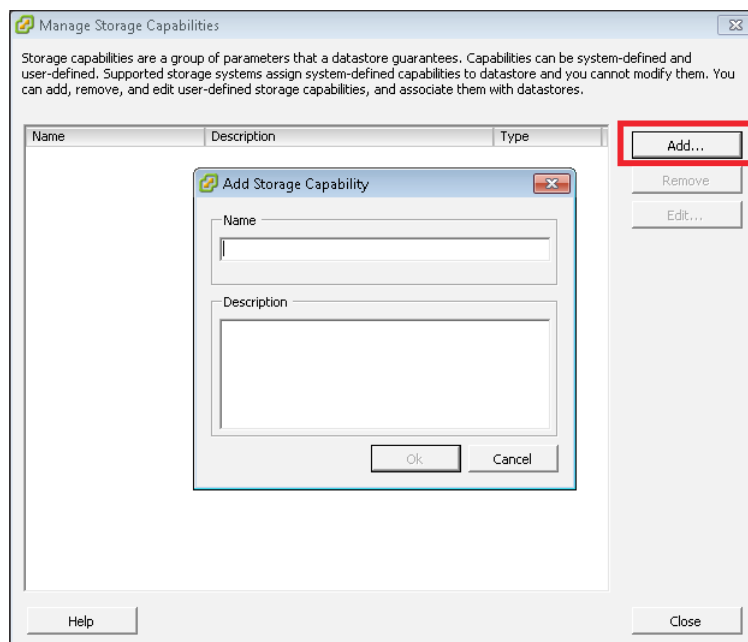
1. Log in to vCenter Server using vSphere Client.
2. Click on the **Home** button in the top bar, and choose the **VM Storage Profiles** button under **Management**. This screenshot is given for reference:



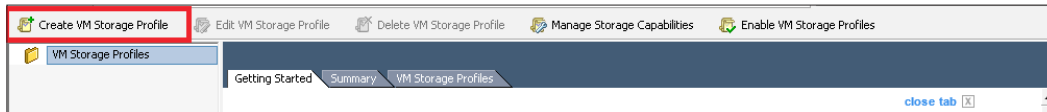
- Click on the **Manage Storage Capabilities** button to create user-defined storage capability. The following screenshot is given for reference:



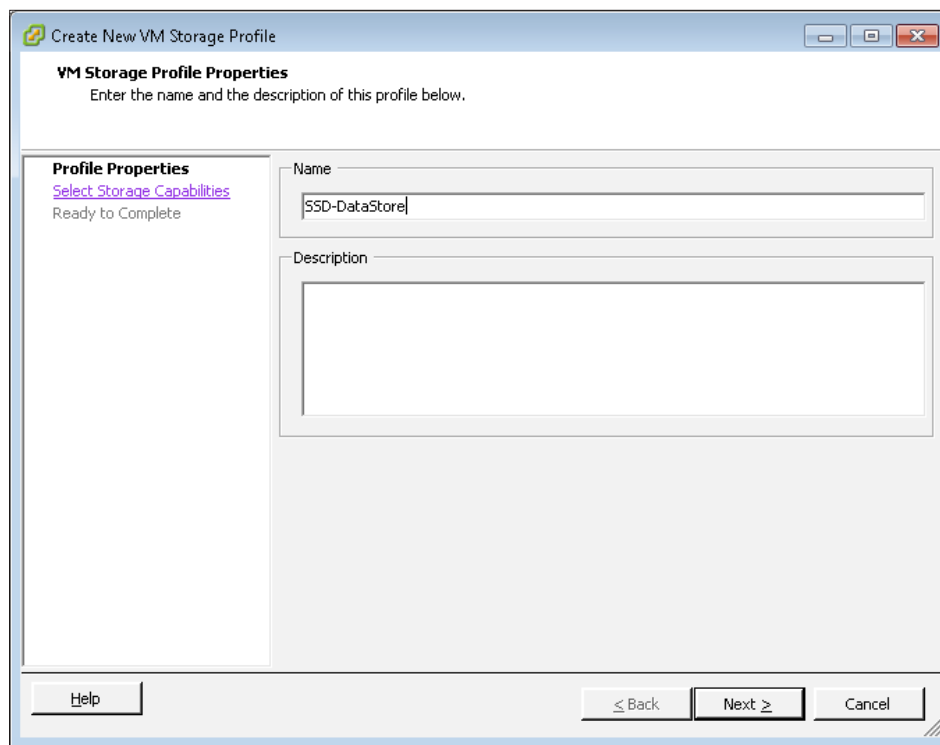
- Click on the **Add** button to create the name of the storage capacity, for example, SSD Storage, SAS Storage, or NL-SAS Storage. Then click on the **Close** button, as shown in this screenshot:



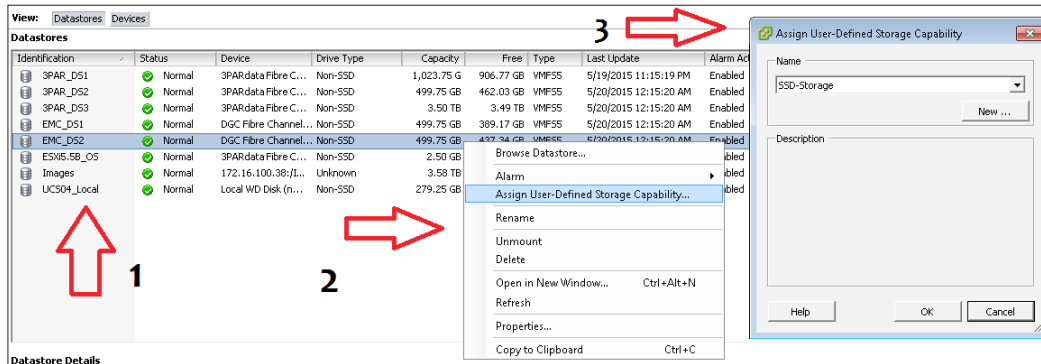
- Click on the **Create VM Storage Profile** button to create the storage policy, as shown in the following screenshot:



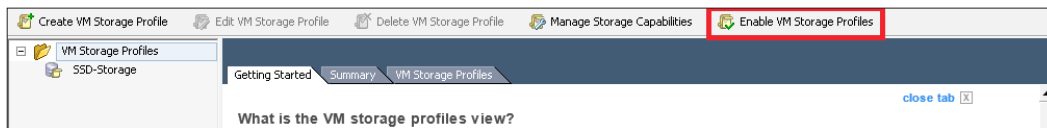
- Input the name of the VM storage profile, as shown in the following screenshot, and then click on the **Next** button to select the user-defined storage capability, which is defined in step 4. Click on the **Finish** button.



7. Assign the user-defined storage capability to your specified ESXi data store. Right-click on the data store that you plan to assign the user-defined storage capability to. This capability is defined in step 4. This screenshot is for reference:

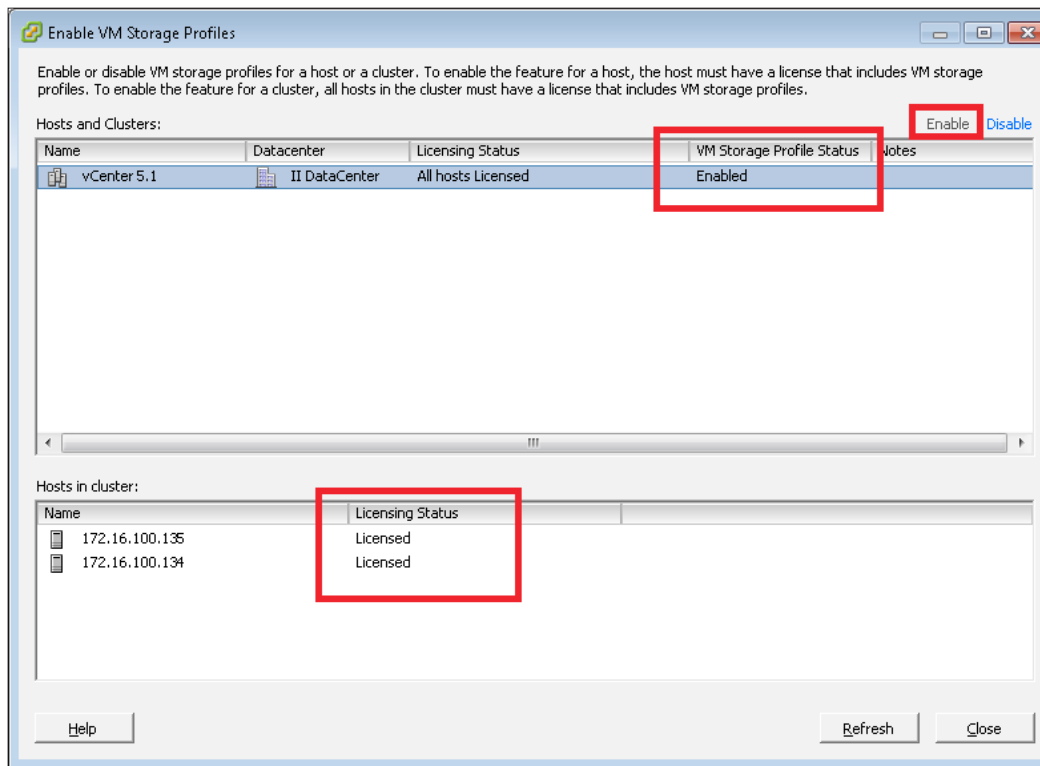


8. After creating the VM storage profile, click on the **Enable VM Storage Profiles** button. Then click on the **Enable** button to enable the profiles. The following screenshot shows **Enable VM Storage Profiles**:

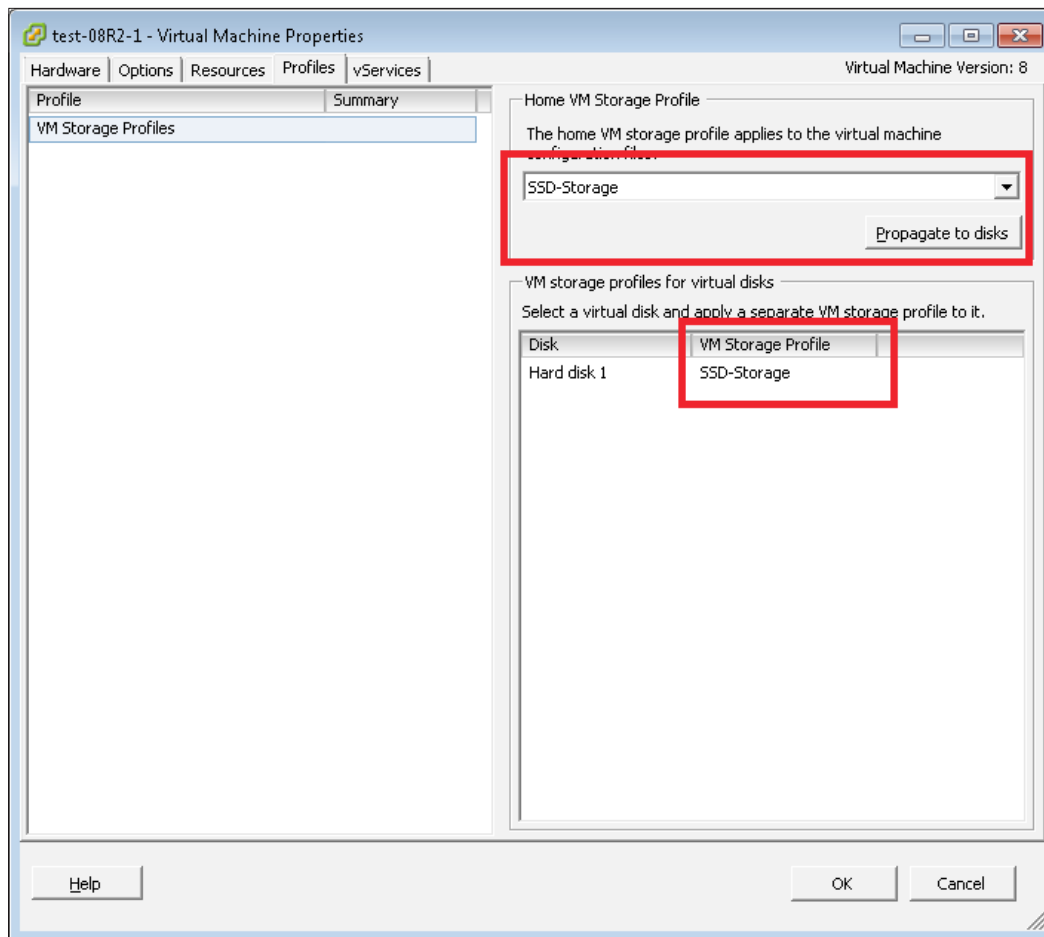




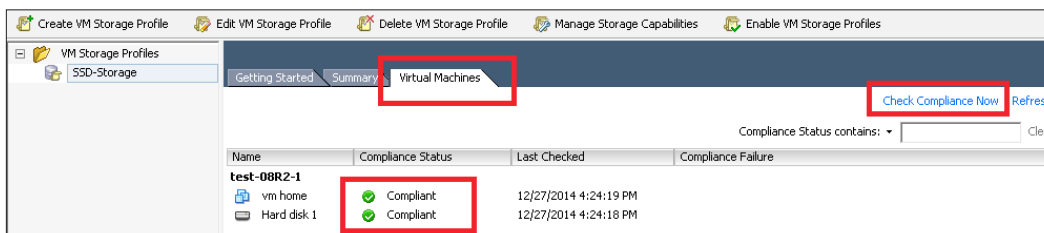
9. After enabling the VM storage profile, you can see **VM Storage Profile Status** as **Enabled** and **Licensing Status** as **Licensed**, as shown in this screenshot:



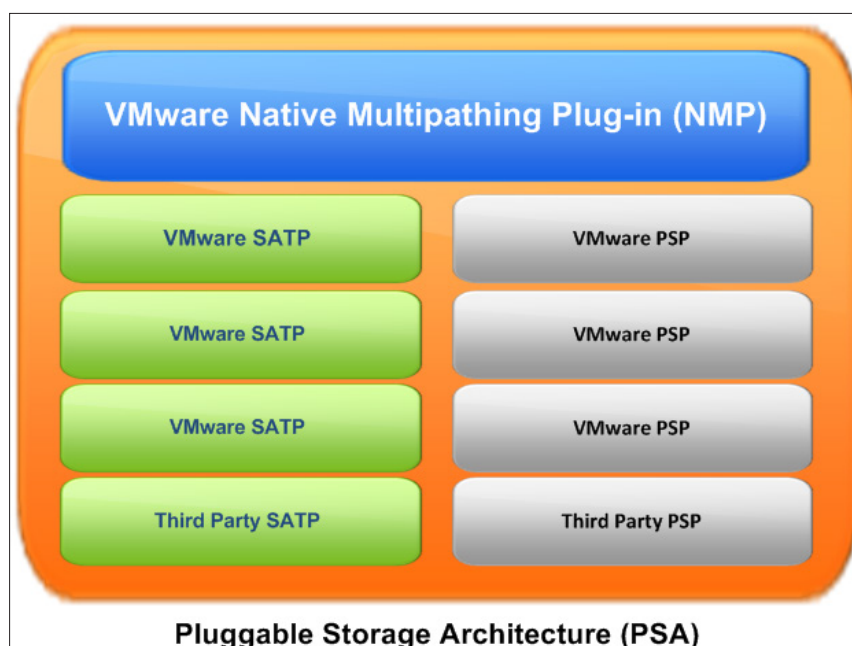
10. We have successfully created the VM storage profile. Now we have to associate the VM storage profile with a virtual machine. Right-click on a virtual machine that you plan to apply to the VM storage profile, choose **VM Storage Profile**, and then choose **Manage Profiles**.
11. From the drop-down menu of **VM Storage Profile** select your profile. Then you can click on the **Propagate to disks** button to associate all virtual disks or decide which virtual disks you want to associate with that profile by setting manually. The following screenshot is for reference. Click on **OK**.



12. Finally, you need to check the compliance of **VM Storage Profile** on this virtual machine. Click on the **Home** button in the top bar. Then choose the **VM Storage Profiles** button under **Management**. Go to **Virtual Machines** and click on the **Check Compliance Now** button. The **Compliance Status** will display **Compliant** after compliance checking, as follows:



**Pluggable Storage Architecture (PSA)** exists in the SCSI middle layer of the VMkernel storage stack. PSA is used to allow thirty-party storage vendors to use their failover and load balancing techniques for their specific storage array. A VMware ESXi host uses its multipathing plugin to control the ownership of the device path and LUN. The VMware default **Multipathing Plugin (MPP)** is called **VMware Native Multipathing Plugin (NMP)**, which includes two subplugins as components: **Storage Array Type Plugin (SATP)** and **Path Selection Plugin (PSP)**. SATP is used to handle path failover for a storage array, and PSP is used to issue an I/O request to a storage array. The following diagram shows the architecture of PSA:



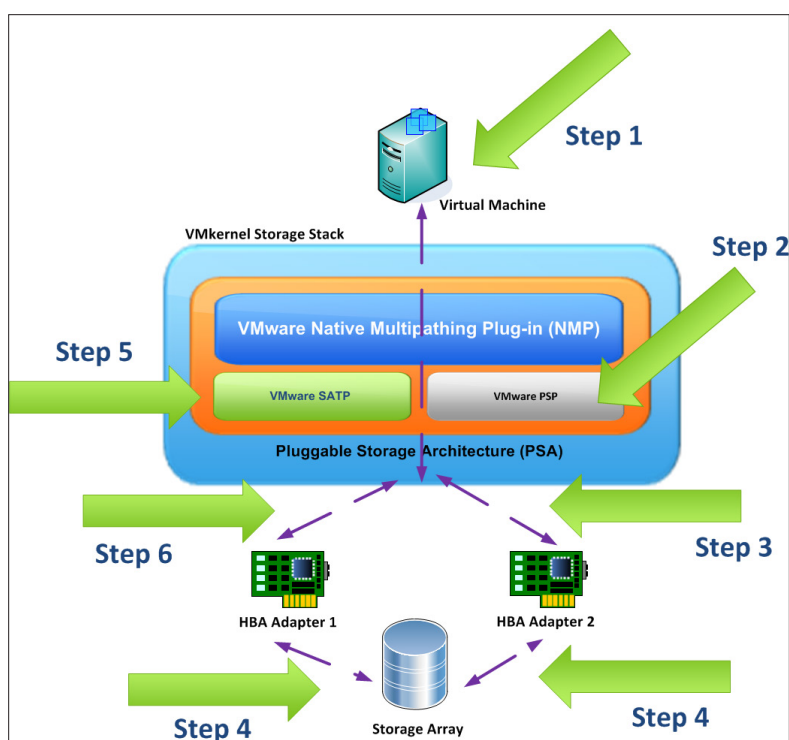
This table lists the operation tasks of PSA and NMP in the ESXi host:

|                 | PSA                                                                  | NMP                                                 |
|-----------------|----------------------------------------------------------------------|-----------------------------------------------------|
| Operation tasks | Discovers the physical paths                                         | Manages the physical path                           |
|                 | Handles I/O requests to the physical HBA adapter and logical devices | Creates, registers, and deregisters logical devices |
|                 | Uses predefined claim rules to control storage devices               | Selects an optimal physical path for the request    |

The following is an example of operation of PSA in a VMkernel storage stack:

1. The virtual machine sends out an I/O request to a logical device that is managed by the VMware NMP.
2. The NMP requests the PSP to assign to this logical device.
3. The PSP selects a suitable physical path to send the I/O request.
4. When the I/O operation is completed successfully, the NMP reports that the I/O operation is complete. If the I/O operation reports an error, the NMP calls the SATP.
5. The SATP fails over to the new active path.
6. The PSP selects a new active path from all available paths and continues the I/O operation.

The following diagram shows the operation of PSA:

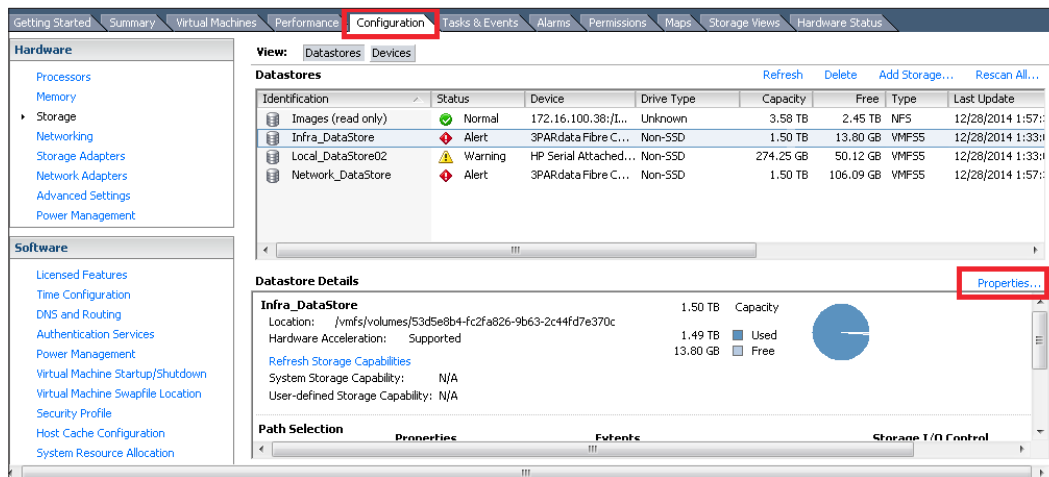


VMware vSphere provides three options for the path selection policy. These are **Most Recently Used (MRU)**, **Fixed**, and **Round Robin (RR)**. The following table lists the advantages and disadvantages of each path:

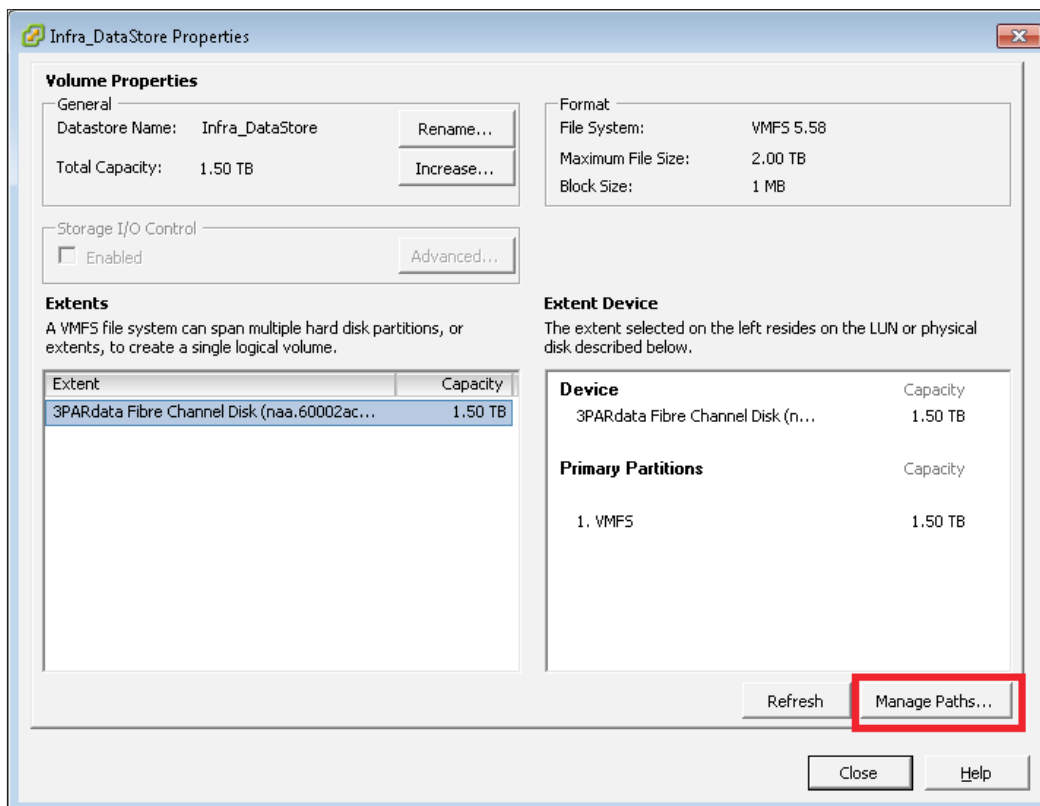
| Path selection | Description                                                                                                                                           | Advantage                                                                                      | Disadvantage                                                                                                                              |
|----------------|-------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------|
| MRU            | The ESXi host selects the first preferred path at system boot time. If this path becomes unavailable, the ESXi host changes to the other active path. | You can select your preferred path manually in the ESXi host.                                  | The ESXi host does not revert to the original path when that 1 path becomes available again.                                              |
| Fixed          | You can select the preferred path manually.                                                                                                           | The ESXi host can revert to the original path when the preferred path becomes available again. | If the ESXi host cannot select the preferred path, it selects an available preferred path randomly.                                       |
| RR             | The ESXi host uses automatic path selection.                                                                                                          | The storage I/O across all available paths and enable load balancing across all paths.         | The storage is required to support ALUA mode. You cannot know which path is preferred because the storage I/O across all available paths. |

The following is the procedure of changing the path selection policy in an ESXi host:

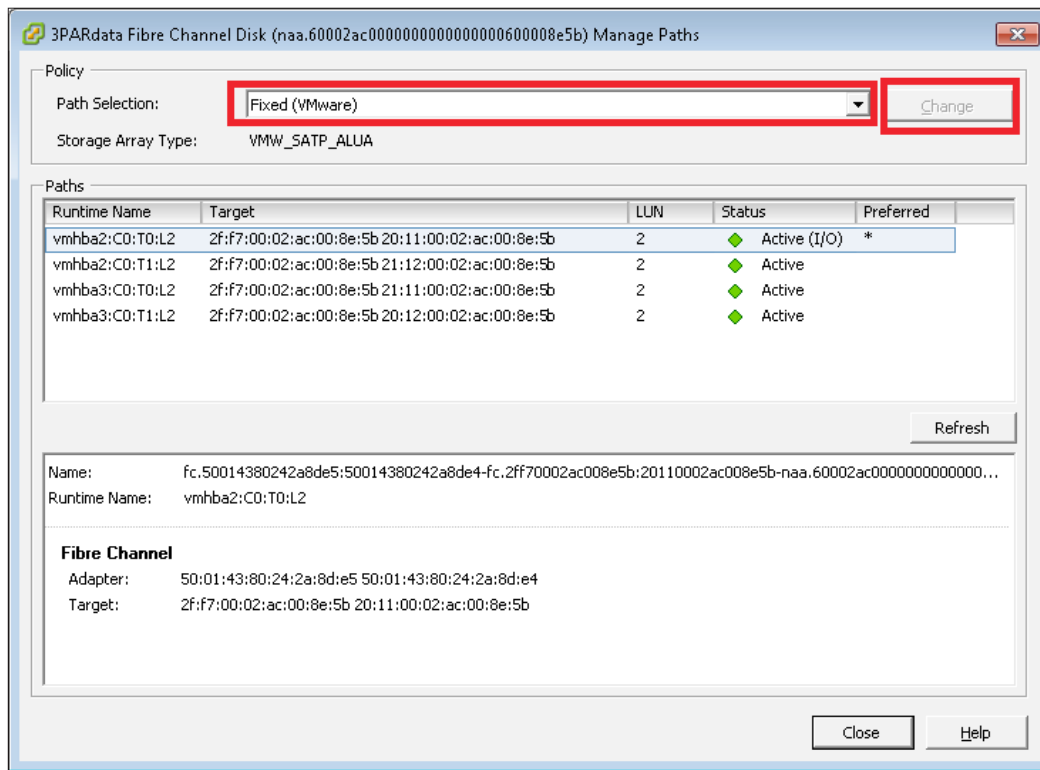
1. Log in to vCenter Server using vSphere Client.
2. Go to the configuration of your selected ESXi host, choose the data store that you want to configure, and click on the **Properties...** button. as shown in the following screenshot:



3. Click on the **Manage Paths...** button, as shown here:



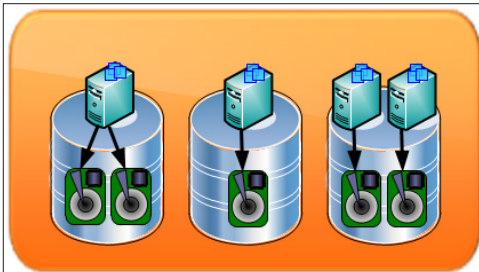
4. Select the drop-down menu and click on the **Change** button, as shown in this screenshot:



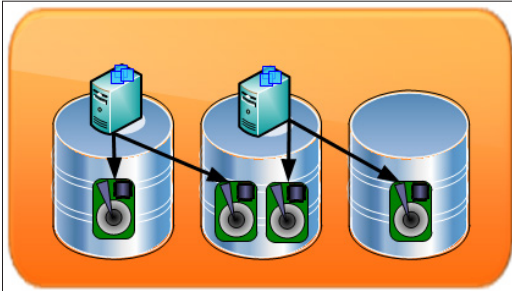
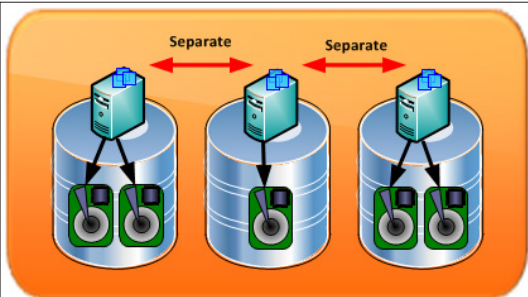
If you plan to deploy a third-party MPP on your ESXi host, you need to follow up the storage vendor's instructions for the installation, for example, EMC PowerPath/VE for VMware that it is a piece of path management software for VMware's vSphere server and Microsoft's Hyper-V server. It also can provide load balancing and path failover features.

## VMware vSphere Storage DRS

VMware vSphere **Storage DRS (SDRS)** is the placement of virtual machines in an ESX's data store cluster. According to storage capacity and I/O latency, it is used by VMware storage vMotion to migrate the virtual machine to keep the ESX's data store in a balanced status that is used to aggregate storage resources, and enable the placement of the virtual disk (VMDK) of virtual machine and load balancing of existing workloads. What is a data store cluster? It is a collection of ESXi's data stores grouped together. The data store cluster is enabled for vSphere SDRS. SDRS can work in two modes: manual mode and fully automated mode. If you enable SDRS in your environment, when the vSphere administrator creates or migrates a virtual machine, SDRS places all the files (VMDK) of this virtual machine in the same data store or different a data store in the cluster, according to the SDRS affinity rules or anti-affinity rules. The VMware ESXi host cluster has two key features: VMware vSphere **High Availability (HA)** and VMware vSphere **Distributed Resource Scheduler (DRS)**. SDRS is different from the host cluster DRS. The latter is used to balance the virtual machine across the ESXi host based on the memory and CPU usage. SDRS is used to balance the virtual machine across the SAN storage (ESX's data store) based on the storage capacity and IOPS. The following table lists the difference between SDRS affinity rules and anti-affinity rules:

| Name of SDRS rules                                                                                                                     | Description                                                                                                                         |
|----------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------|
| VMDK affinity rules<br><br><b>DataStore Cluster</b> | This is the default SDRS rule for all virtual machines. It keeps each virtual machine's VMDKs together on the same ESXi data store. |

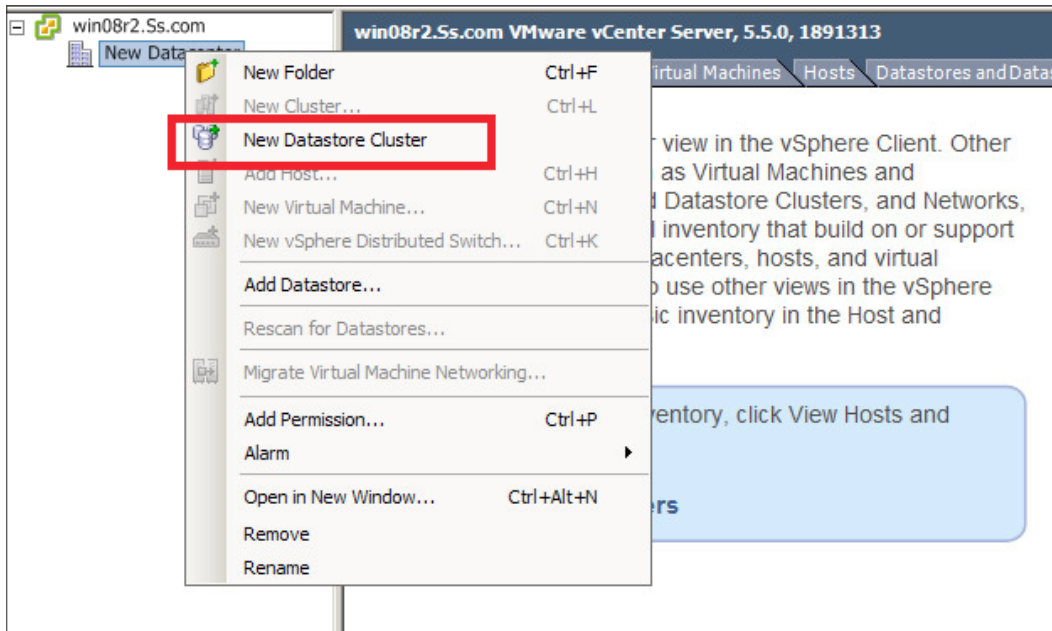


| Name of SDRS rules                                                                                                                                                            | Description                                                                                                                                                               |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>VMDK anti-affinity rules</p>  <p style="text-align: center;"><b>DataStore Cluster</b></p> | <p>Keep each virtual machine's VMDKs on different ESXi data stores. You can apply this rule into all virtual machine's VMDKs or to dedicated virtual machine's VMDKs.</p> |
| <p>VM anti-affinity rules</p>  <p style="text-align: center;"><b>DataStore Cluster</b></p>  | <p>Keep the virtual machine on different ESXi data stores. This rule is similar to the ESX DRS anti-affinity rules.</p>                                                   |

The following is the procedure to create a storage DRS in vSphere 5:


1. Log in to vCenter Server using vSphere Client.

2. Go to home and click on the **Datastores and Datastore Clusters** button.  
Right-click on the data center and choose **New Datastore Cluster**, as shown in the following screenshot:



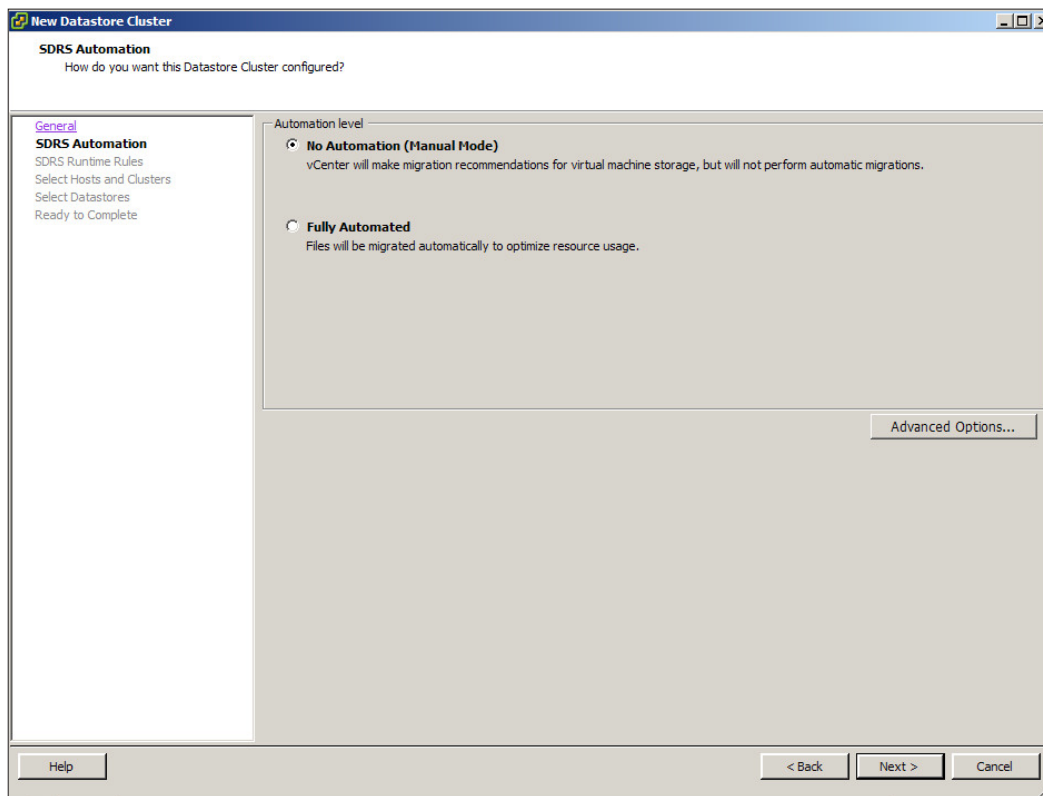
3. Input the name of the SDRS and then click on the **Next** button.

4. Choose **Storage DRS** mode, **Manual Mode** and **Fully Automated Mode** as shown in the next screenshot.



**Manual Mode:** According to the placement and migration recommendation, the placement and migration of the virtual machine are executed manually by the user.

**Fully Automated Mode:** Based on the runtime rules, the placement of the virtual machine is executed automatically.



5. Set up **SDRS Runtime Rules**. Then click on the **Next** button, as shown here:

**New Datastore Cluster**

**SDRS Runtime Rules**  
How do you want this Datastore Cluster configured?

**General**  
**SDRS Automation**  
**SDRS Runtime Rules**  
Select Hosts and Clusters  
Select Datastores  
Ready to Complete

**I/O Metric Inclusion**  
Select this option if you want I/O metrics considered as a part of any SDRS recommendations or automated migrations in this datastore cluster. This will also enable Storage I/O Control on all datastores in this cluster.  
☒ **Enable I/O metric for SDRS recommendations**  
I/O load balancing functionality is available only when all hosts connected to the datastores in this datastore cluster are of version 5.0.

**Storage DRS Thresholds**  
Runtime thresholds govern when storage DRS performs or recommends migrations (based on your selected automation level). Utilized space dictates the minimum level of consumed space that is the threshold for action, and I/O latency dictates the minimum I/O latency below which I/O load balancing moves will not be considered.

Utilized Space: 50% 100% **80** %

I/O Latency: 5ms 100ms **15** ms

[Hide Advanced Options](#)

**Advanced Options**  
No recommendations until utilization difference between source and destination is: 1% 50% **5** %  
Check imbalances every: **8** Hours  
I/O imbalance threshold: **Aggressive** Conservative  
The I/O imbalance threshold determines the amount of imbalance that Storage DRS should tolerate. Aggressive setting would make Storage DRS correct small imbalances, if possible and moving it toward conservative would make Storage DRS produce recommendations only in cases when the imbalance across datastores is very high.

Help < Back Next > Cancel



**Enable I/O metric for SDRS recommendations** is used to enable I/O load balancing.


**Utilized Space** is the percentage of consumed space allowed before the storage DRS executes an action.

**I/O Latency** is the percentage of consumed latency allowed before the storage DRS executes an action. This setting can execute only if the **Enable I/O metric for SDRS recommendations** checkbox is selected.

**No recommendations until utilization difference between source and destination is** is used to configure the space utilization difference threshold.

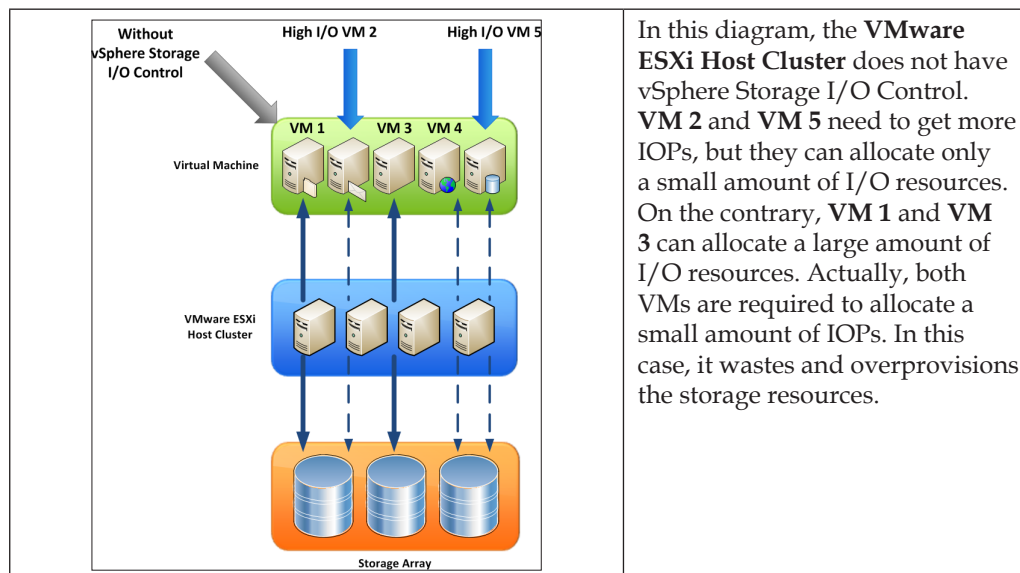
**I/O imbalance threshold** is used to define the aggressive of IOPs load balancing. This setting can execute only if the **Enable I/O metric for SDRS recommendations** checkbox is selected.

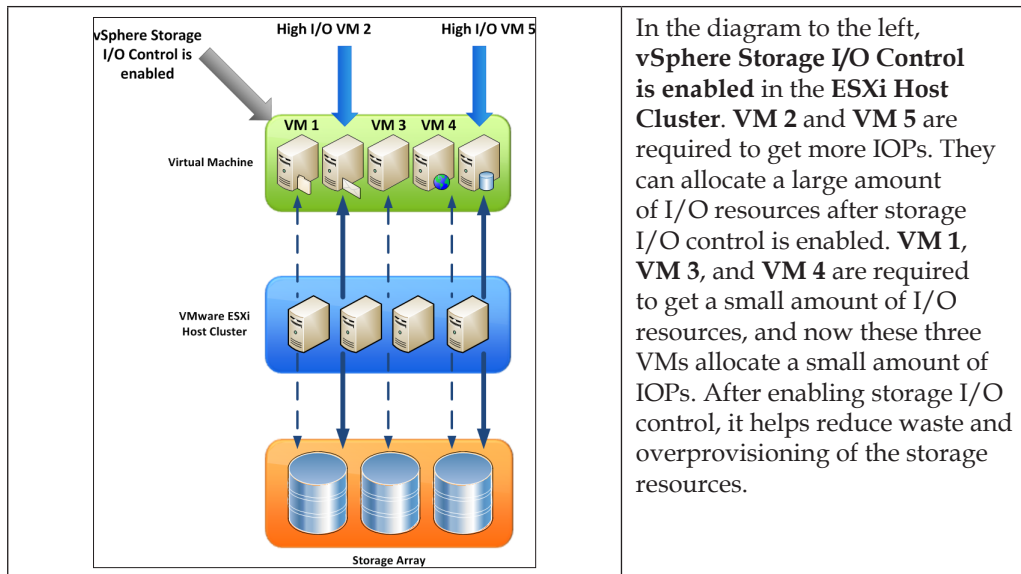
6. Select the ESXi host that is required to create SDRS. Then click on the **Next** button.
7. Select the data store that is required to join the data store cluster, and click on the **Next** button to complete.
8. After creating SDRS, go to the vSphere Storage DRS panel on the **Summary** tab of the data store cluster. You can see that **Storage DRS is Enabled**.
9. On the **Storage DRS** tab on the data store cluster, it displays the recommendation, placement, and reasons. Click on the **Apply Recommendations** button if you want to apply the recommendations.

[  Click on the **Run Storage DRS** button if you want to refresh the recommendations. ]

## VMware vSphere Storage I/O Control

What is VMware vSphere Storage I/O Control? It is used to control in order to share and limit the storage of I/O resources, for example, the IOPS. You can control the number of storage IOPs allocated to the virtual machine. If a certain virtual machine is required to get more storage I/O resources, vSphere Storage I/O Control can ensure that that virtual machine can get more storage I/O than other virtual machines. The following table shows example of the difference between vSphere Storage I/O Control enabled and without vSphere Storage I/O Control:

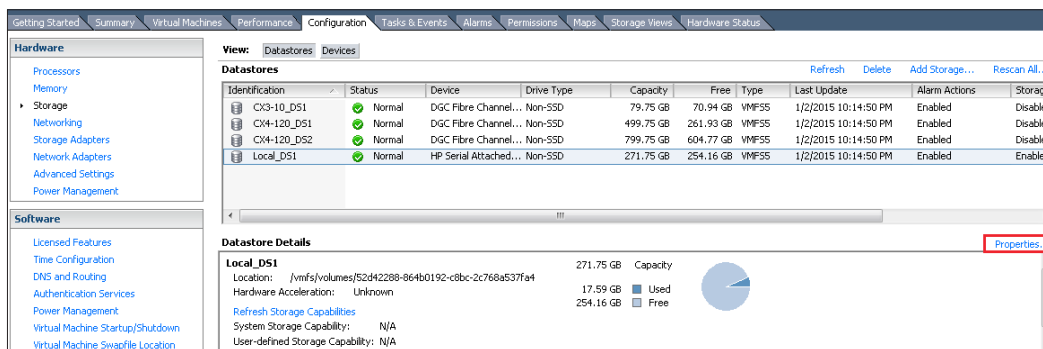




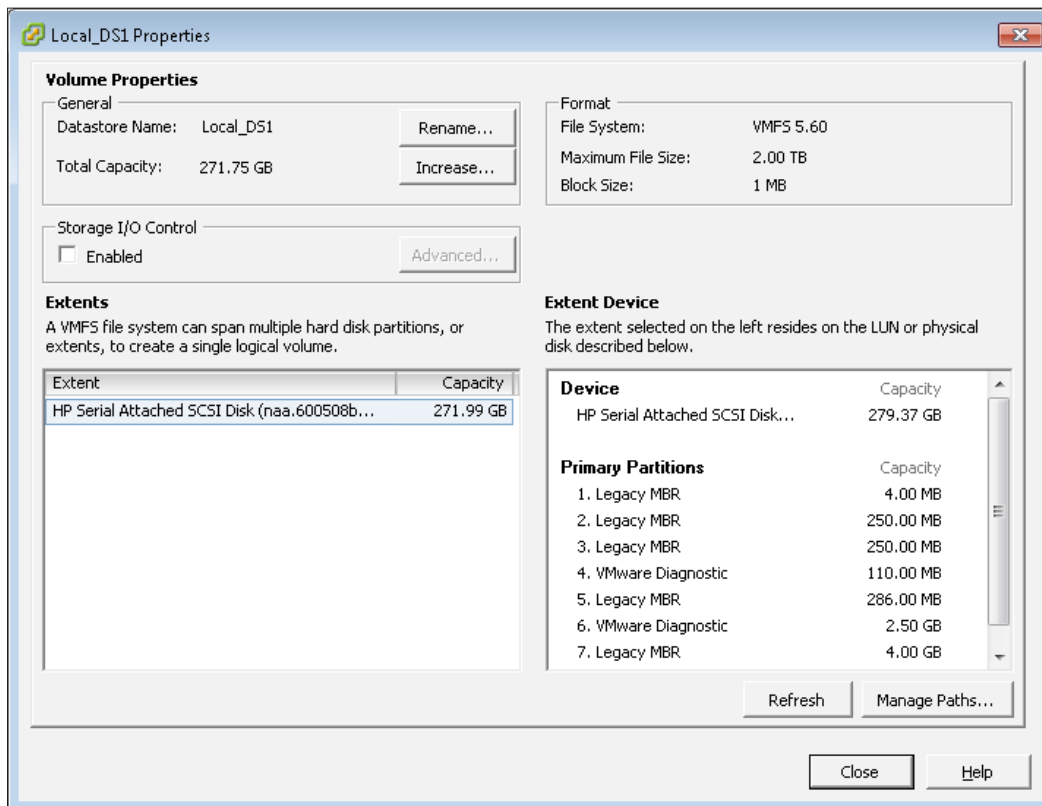
When you enable VMware vSphere Storage DRS, vSphere Storage I/O Control is automatically enabled on the data stores in the data store cluster.

The following is the procedure to be carried out to enable vSphere Storage I/O control on an ESXi data store, and set up storage I/O shares and limits using vSphere Client 5:

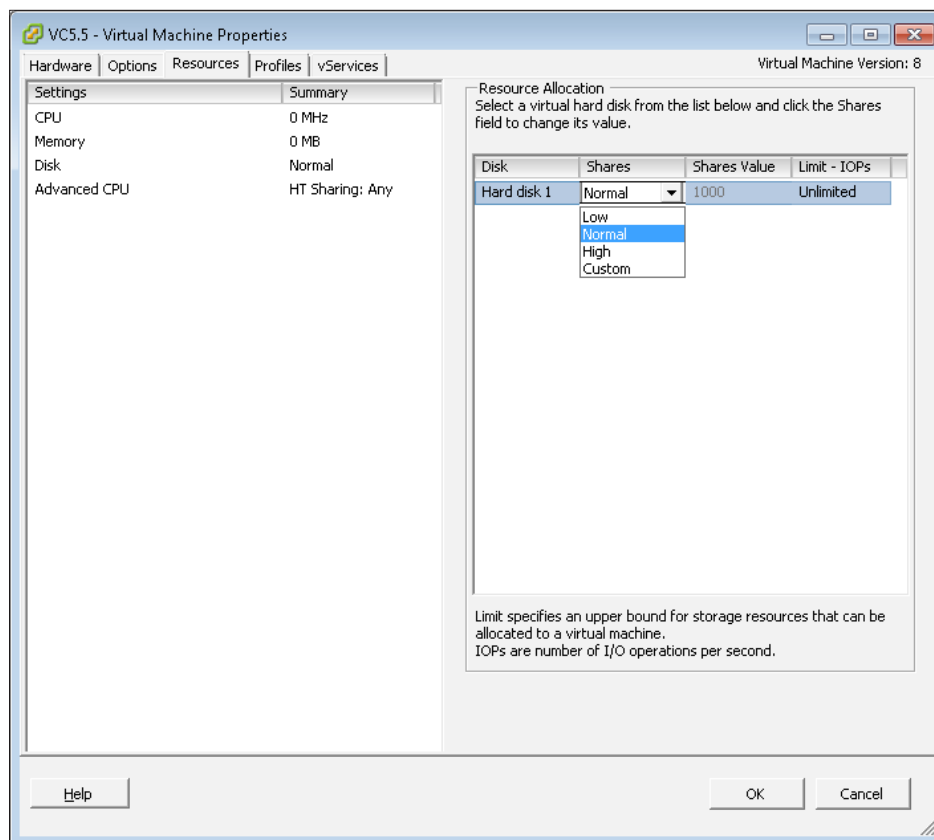
1. Log in to vCenter Server using vSphere Client.
2. Go to the **Configuration** tab of the ESXi host, select the data store, and then click on the **Properties...** button, as shown in the following screenshot:




3. Select **Enabled** under **Storage I/O Control**, and click on the **Close** button, as shown here:



4. After Storage I/O Control is enabled, you can set up the storage I/O shares and limits on the virtual machine. Right-click on the virtual machine and select **Edit Settings**.
5. Click on the **Resources** tab in the virtual machine properties box, and select **Disk**. You can individually set each virtual disk's **Shares** and **Limit** field, like this:




 By default, all virtual machine shares are set to **Normal** and with **Unlimited** IOPs.

## Summary

In this chapter, you learned what VAAI and VASA are. In a vSphere environment, the vSphere administrator learned how to configure the storage profile in vCenter Server and assign to the ESXi data store. We covered the benefits of vSphere Storage I/O Control and vSphere Storage DRS.

In the next chapter, we will see how to monitor the ESXi host's storage performance and manage the storage in a vSphere host using the command line. When you found that it has a storage performance problem in the vSphere host, we saw how to troubleshoot the performance problem, and found out the root cause.





# 5

## Optimizing Storage

The utilization and performance of SAN storage can affect the performance of every virtual machine in the virtualization environment. The vSphere administrator should know how to monitor the host's storage IOPS and troubleshoot performance using vSphere tools and the command line. They should also know the property of SAN storage and operation flow between the vSphere host and SAN storage. Then they can solve performance problems easily.

In this chapter, we will learn:

- Concepts of storage virtualization
- Monitoring vSphere storage
- vSphere storage management using the command line
- Troubleshooting vSphere storage performance problems

### Concepts of storage virtualization

In *Chapter 4, Storage Scalability*, you learned about the advanced features and setting of storage in a VMware vSphere environment. In this chapter, you will learn how to optimize and monitor vSphere's storage. Firstly, you should know the concept of vSphere's storage clearly, and then know how to tune performance of vSphere's storage. We know that the filesystem of the VMware vSphere ESXi host is VMFS, which can support most storage protocols, for example, FC, **Fibre Channel over Ethernet (FCoE)**, hardware iSCSI, software iSCSI, and NFS. It has a different configuration based on the storage protocol during the vSphere's storage configuration. Storage performance problems are caused due to many factors, for example, the hardware (HBA adapter, SAN storage, and SAN switch), RAID level, cache size, and queue depth. ESX's path selection policy can directly affect the performance of each virtual machine in vSphere environment. For ESX's path selection policy, you can refer to *Chapter 4, Storage Scalability*.

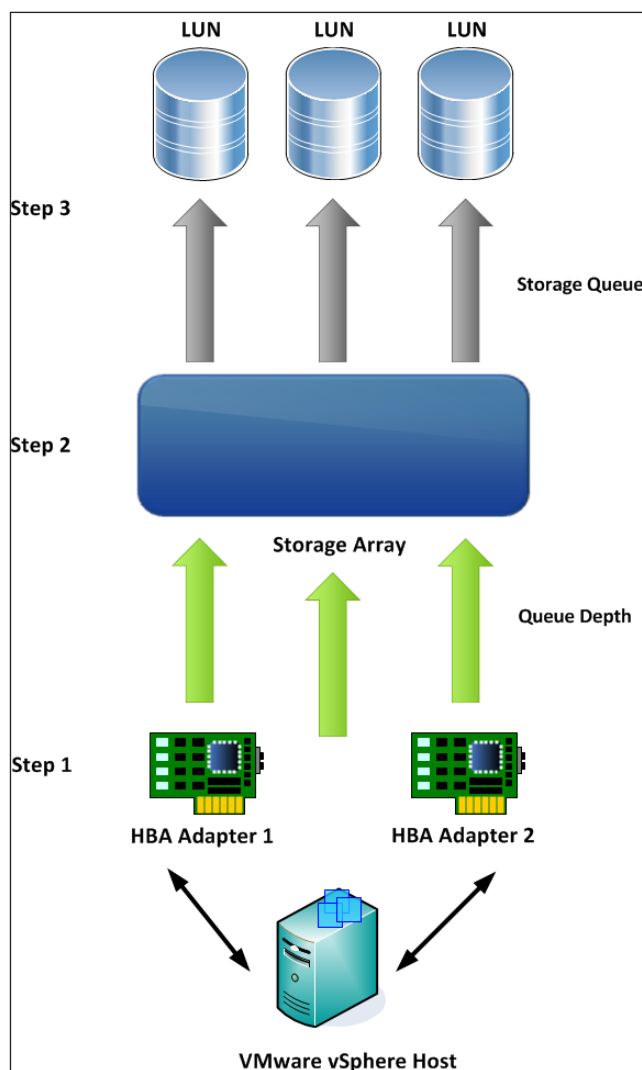
When you set up your vSphere host to the storage array using the fibre channel protocol, HBA is a key component between the vSphere host and storage array. QLogic and Emulex HBA is the common branch chosen to be used. What is queuing at the host and storage array? It controls the number of active commands on a LUN.

The next diagram lists the high-level operation flow of the vSphere host access storage array. There are two types of queues in a VMware vSphere host; they are the **device driver queue** and the **kernel queue**. The device driver queue is used to control the number of active commands that can be on a LUN at the same time. A kernel queue is the overflow queue for a device driver queue and it optimizes storage. The device driver queue has a configuration parameter called **queue depth** that decides how many SCSI commands can be active at one time on a LUN. If the total number of outstanding SCSI commands from all virtual machines exceeds this queue depth on one ESXi host, the excess SCSI commands are queued in the ESXi VMkernel; this increases latency. When the number of active SCSI commands on a LUN is too high, queuing occurs in the storage array. To prevent this performance issue, VMware recommends configuring the queue depth of HBA in the ESXi host. The default value of queue depth for some brand HBA is 32. If one ESXi host generates more SCSI commands to a LUN than this queue depth, it will have performance problems. As best practice, the recommended value of queue depth is 64.



Each brand of HBA has a different default value of queue depth. You need to verify the vendor of HBA. The default queue depths of the QLogic HBA and Emulex HBA are different.

The following diagram shows the high-level operation flow of a vSphere host access storage array:



After discussing the performance problem with Fibre Channel SAN storage, we will discuss iSCSI and NFS SAN storage. Most performance problems in iSCSI and NFS are related to the network bottleneck. Make sure that the storage network is isolated from other networks. This is best practice of iSCSI and NFS configuration in an ESXi host. If your ESXi host has enough NIC adapters, we recommend that you configure dual uplinks for the iSCSI or NFS port group in ESX's virtual switch (vSphere standard vSwitch and distributed vSwitch).



For software iSCSI and NFS storage protocols, it is necessary to use a CPU resource on the ESXi host.

VMFS is a filesystem in the VMware vSphere ESXi host. In vSphere 4.x, the filesystem of the ESX data store is VMFS3. In vSphere 5.0 and above, it was upgraded to VMFS5. VMFS5 can provide performance improvements over VMFS3; for example, the maximum ESXi's datastore can be up to 64 TB, and the block size of the filesystem changes to 1 MB. If you upgrade your vSphere host from version 4.x to 5.x, ESXi's data store upgrade is one part of the upgrade plan. It has one important thing during the data store upgrade; a newly created VMFS5 and a VMFS5 system file upgrade are different. Most vSphere administrators make mistakes here. If you upgrade VMFS3 to VMFS5, it continues to use the previous file block size (2 MB). During a vSphere upgrade, most vSphere administrators upgrade VMFS3 to VMFS5 directly if there is no new ESXi's data store as extra. In this situation, the data store uses the previous file size (2 MB).

In VMFS5, its block size has changed to 1 MB, and it wastes unused disk space when there are many small files on the ESXi data store if you upgrade VMFS3 to VMFS5. In some hands-on experience, the VMware administrator creates a new data store (VMFS5) and then migrates all virtual machines that are running on the VMFS3 data store to this new data store, using VMware Storage vMotion.

There are three ways of accessing a virtual disk in a virtual machine: virtual disk in a VMFS, physical **raw device mapping (RDM)**, and virtual raw device mapping. Each type of virtual disk can give a different kind of performance in a virtual machine. For example, if your application has to provide sequential reads/writes for a large I/O block size, VMFS is a better choice than RDM. The following table lists the difference between a virtual disk in VMFS and RDM:

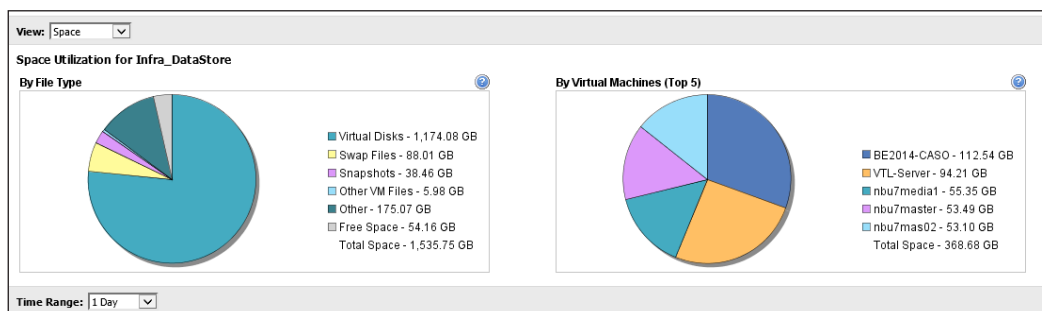
| Disk access type       | I/O characteristic                                                                                                                        | Use case                                                                                                                           |
|------------------------|-------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------|
| Virtual disk in a VMFS | <ul style="list-style-type: none"><li>• Random reads/writes</li><li>• Sequential reads/writes at small or large I/O block sizes</li></ul> | Not supported in the Microsoft Cluster Service (MSCS)                                                                              |
| Physical RDM           | <ul style="list-style-type: none"><li>• Random reads/writes</li><li>• Sequential reads/writes at small I/O block sizes</li></ul>          | <ul style="list-style-type: none"><li>• Supported in MSCS</li><li>• Not supported for creating snapshots in physical RDM</li></ul> |

| Disk access type | I/O characteristic                                                                                                              | Use case                                                                                                                     |
|------------------|---------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------|
| Virtual RDM      | <ul style="list-style-type: none"> <li>Random reads/writes</li> <li>Sequential reads/writes at small I/O block sizes</li> </ul> | <ul style="list-style-type: none"> <li>Supported in MSCS</li> <li>Supported for creating snapshots in virtual RDM</li> </ul> |

According to the type of virtual disk in virtual machine is used that has different I/O performance figure, vSphere ESXi host can support three types of virtual disks: eagerzeroed thick, lazy zeroed thick, and lazy zeroed thin. You can refer to *Chapter 2, Getting Started with vSphere Management Assistant*, for more details.

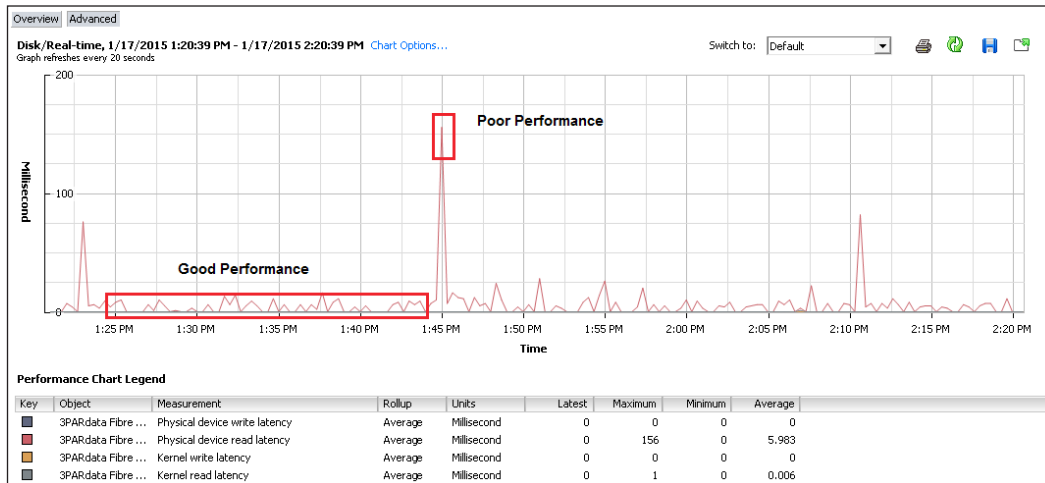
## Monitoring vSphere storage

vSphere performance charts are very useful for monitoring the performance of the ESX data store. In a performance chart, by default, you can view the space utilization for the data store. It includes two pie charts: **By File Type** and **By Virtual Machines (Top 5)**. The first pie chart (**By File Type**) can display the portions of space utilized by **Virtual Disk**, **Swap Files**, **Snapshots**, **Other VM Files**, **Other** (files), and **Free Space**. The other pie chart (**By Virtual Machines**) can display the space utilization of the top five virtual machines stored in the data store. The following charts, which are located under the vSphere **Performance** tab of each data store, are for your reference:



VMware vSphere performance charts have a lot of statistics metrics that can help you identify the storage or disk problem, for example, disk read/write latency, number of commands queued, number of active disk commands, and number of aborted disk commands. Disk latency is the time taken to complete an I/O request. In a vSphere host environment, the I/O request passes from the VMkernel to the physical storage device. Disk latency can occur when a lot of SCSI commands are queued, either at the VMkernel or at the SAN storage.

Some performance problems can easily be identified by monitoring the physical device read/write latency metric and the kernel read/write latency metric in the vSphere performance charts. The following performance chart is for your reference:



Physical device read/write latency is the average time for the physical device to complete an SCSI command. If this value is greater than 15 milliseconds, we can interpret that the storage array might be slow. Kernel read/write latency is the average time for which the VMkernel runs each SCSI command. For best performance, this value should be 0 to 1 milliseconds. If it is greater than 4 milliseconds, it means that this ESXi host might have performance problem.

If you want to collect detailed information of the read/writes per second of the disk in a vSphere host, you can use `resxtp`/`esxtp` to monitor it. For the procedure of enabling `esxtp`, you can refer to *Chapter 3, Using the Virtual Machine Monitor*. In the following example of the ESXi's storage adapter resource utilization screen, you can see the **READS/s**, **WRITES/s**, **MBREAD/s**, and **MBWRTN/s** of each storage adapter. The sum of **READS/s** and **WRITES/s** equals to I/O operations/s (IOPs), which is a common benchmark for storage. If you want to monitor the throughput, you can check the metrics of **MBREAD/s** and **MBWRTN/s**. **MBREAD/s** is the number of megabytes read per second, and **MBWRTN/s** is the number of megabytes written per second. The following table gives the details of each metric:

| Metric   | Description                                |
|----------|--------------------------------------------|
| READS/s  | Number of read commands issued per second  |
| WRITES/s | Number of write commands issued per second |
| MBREAD/s | Megabytes read per second                  |
| MBWRTN/s | Megabytes written per second               |

```
11:30:05am up 58 days 7:02, 554 worlds, 19 VMs, 34 vCPUs; CPU load average: 0.10, 0.10, 0.10
```

| ADAPTR  | PATH | NPTH | CMDS/s | READS/s | WRITES/s | MBREAD/s | MBWRTN/s | DAVG/cmd | KAVG/cmd | GAVG/cmd | QAVG/cmd |
|---------|------|------|--------|---------|----------|----------|----------|----------|----------|----------|----------|
| vmhba0  | -    | 0    | 0.00   | 0.00    | 0.00     | 0.00     | 0.00     | 0.00     | 0.00     | 0.00     | 0.00     |
| vmhba1  | -    | 1    | 5.72   | 1.53    | 4.20     | 0.05     | 0.06     | 0.05     | 0.02     | 0.07     | 0.00     |
| vmhba2  | -    | 8    | 390.63 | 222.78  | 167.85   | 46.20    | 1.66     | 1.71     | 0.01     | 1.72     | 0.00     |
| vmhba3  | -    | 8    | 3.81   | 0.00    | 3.81     | 0.00     | 0.02     | 0.26     | 0.03     | 0.29     | 0.00     |
| vmhba32 | -    | 0    | 0.00   | 0.00    | 0.00     | 0.00     | 0.00     | 0.00     | 0.00     | 0.00     | 0.00     |

Listed here are some examples of monitoring disk throughput using `esxtop`. In the following example 1, it displays all the I/O metrics of `vmhba2` on the following screen. You need to ensure that the I/O statistics and overall latency statistics fields are selected. You can find the real time of **READS/s**, **WRITES/s**, **DAVG/cmd**, **KAVG/cmd**, **DAVG/rd**, and **KAVG/rd**. **DAVG/cmd** is the average response time of each command sent to the device in milliseconds, and **KAVG/cmd** is the time the command spends in the VMkernel. **DAVG/rd** and **KAVG/rd** is the read latency statistics. The following table gives the details of each metric:

| Metric   | Description                                                                            |
|----------|----------------------------------------------------------------------------------------|
| DAVG/cmd | The average response time in milliseconds per command being sent to the storage device |
| KAVG/cmd | The amount of time the command spends in the VMkernel                                  |
| DAVG/rd  | The read latency stats in milliseconds per command being sent to the storage device    |
| KAVG/rd  | The read latency time that the command spends in the VMkernel                          |

The following is the screenshot for example 1:

```
3:37:46am up 60 days 23:10, 546 worlds, 19 VMs, 34 vCPUs; CPU load average: 0.12, 0.11, 0.11
```

| ADAPTR  | PATH | NPTH | CMDS/s | READS/s | WRITES/s | MBREAD/s | MBWRTN/s | DAVG/cmd | KAVG/cmd | GAVG/cmd | QAVG/cmd | DAVG/rd | KAVG/rd | GAVG/rd | QAVG/rd |
|---------|------|------|--------|---------|----------|----------|----------|----------|----------|----------|----------|---------|---------|---------|---------|
| vmhba0  | -    | 0    | 0.00   | 0.00    | 0.00     | 0.00     | 0.00     | 0.00     | 0.00     | 0.00     | 0.00     | 0.00    | 0.00    | 0.00    | 0.00    |
| vmhba1  | -    | 1    | 2.10   | 0.00    | 2.10     | 0.06     | 0.06     | 0.06     | 0.02     | 0.08     | 0.00     | 0.00    | 0.00    | 0.00    | 0.00    |
| vmhba2  | -    | 8    | 276.57 | 75.15   | 201.42   | 19.70    | 2.61     | 1.86     | 0.01     | 1.87     | 0.00     | 2.87    | 0.02    | 5.89    | 0.00    |
| vmhba3  | -    | 8    | 0.00   | 0.00    | 0.00     | 0.00     | 0.00     | 0.00     | 0.00     | 0.00     | 0.00     | 0.00    | 0.00    | 0.00    | 0.00    |
| vmhba32 | -    | 0    | 0.00   | 0.00    | 0.00     | 0.00     | 0.00     | 0.00     | 0.00     | 0.00     | 0.00     | 0.00    | 0.00    | 0.00    | 0.00    |





If the value of **DAVG/cmd** is greater than 20, it means that that ESXi host has a performance problem.

The value of **KAVG/cmd** should be close to zero. If this value is greater than 1, it means that that ESXi host has a performance problem.

[illegible]

```

4:48:44am up 61 days 21 min, 548 worlds, 19 VMs, 34 vCPUs; CPU load average: 0.12, 0.11, 0.11

```

| GID     | VMNAME          | VDEVNAME | NVDISK | CMDS/s | READS/s | WRITES/s | MBREAD/s | MBURTN/s | LAT/rd | LAT/wr |
|---------|-----------------|----------|--------|--------|---------|----------|----------|----------|--------|--------|
| 38457   | UCS Platform Em | -        | 4      | 5.34   | 0.00    | 5.34     | 0.00     | 0.04     | 0.00   | 0.29   |
| 38726   | 20i2            | -        | 1      | 0.00   | 0.00    | 0.00     | 0.00     | 0.00     | 0.00   | 0.00   |
| 143305  | VTL-Server      | -        | 1      | 0.00   | 0.00    | 0.00     | 0.00     | 0.00     | 0.00   | 0.00   |
| 144425  | OpsCenter7      | -        | 2      | 0.38   | 0.00    | 0.38     | 0.00     | 0.00     | 0.00   | 0.33   |
| 147701  | IICDS01         | -        | 1      | 0.00   | 0.00    | 0.00     | 0.00     | 0.00     | 0.00   | 0.00   |
| 154269  | be2014c1nt2     | -        | 1      | 0.00   | 0.00    | 0.00     | 0.00     | 0.00     | 0.00   | 0.00   |
| 1641664 | Redhat6.3 64bit | -        | 1      | 0.00   | 0.00    | 0.00     | 0.00     | 0.00     | 0.00   | 0.00   |
| 2058455 | UCS Platform Em | -        | 4      | 218.20 | 0.00    | 218.20   | 0.00     | 2.83     | 0.00   | 0.33   |
| 4096813 | Treeman's Serve | -        | 1      | 0.19   | 0.00    | 0.19     | 0.00     | 0.00     | 0.00   | 0.39   |
| 4688934 | nbu7master      | -        | 3      | 13.16  | 0.00    | 13.16    | 0.00     | 0.10     | 0.00   | 0.34   |
| 6112811 | IICEX13CAS-01   | -        | 1      | 1.14   | 0.00    | 1.14     | 0.00     | 0.04     | 0.00   | 0.44   |
| 6114317 | IICEX13MBX-01   | -        | 1      | 0.76   | 0.00    | 0.76     | 0.00     | 0.01     | 0.00   | 0.15   |
| 6256315 | IICEX13MBX-02   | -        | 1      | 2.67   | 0.00    | 2.67     | 0.00     | 0.02     | 0.00   | 0.43   |
| 6297497 | nbu7c1nt1       | -        | 1      | 0.00   | 0.00    | 0.00     | 0.00     | 0.00     | 0.00   | 0.00   |
| 6472735 | O8r2 vc5.5      | -        | 2      | 0.57   | 0.00    | 0.57     | 0.00     | 0.01     | 0.00   | 0.37   |
| 7312873 | rhel_Vsp-4.3.0. | -        | 1      | 2.67   | 0.00    | 2.67     | 0.00     | 0.02     | 0.00   | 0.39   |
| 7580487 | esx             | -        | 1      | 0.00   | 0.00    | 0.00     | 0.00     | 0.00     | 0.00   | 0.00   |
| 8321809 | winxp           | -        | 1      | 0.00   | 0.00    | 0.00     | 0.00     | 0.00     | 0.00   | 0.00   |
| 8466995 | nbu7c1nt3       | -        | 2      | 2.29   | 0.00    | 2.29     | 0.00     | 0.02     | 0.00   | 1.70   |

[illegible][illegible]

## vSphere storage management using the command line

In this section, you will learn how to manage vSphere storage using the command line. During vSphere storage configuration, the detailed information about the HBA adapter is very important. Before you define zoning on the SAN switch, you should know the **World Wide Number (WWN)** of each host bus adapter for the ESXi host. Then you can define the zoning access for the ESXi host and SAN storage correctly. The WWN is a unique identifier that is hardcoded in a FC device. The `esxcli storage` command is used to manage different storage management tasks. For example, you can list the WWN of HBA, list the information of the LUN, manage storage paths from the vSphere host, and so on. The following section lists an example of storage management tasks using the `esxcli storage` commands.

In the following example (which displays the LUNs on ESXi host), you can execute an `esxcli` command to collect the result, which is given for your reference. You can see that this LUN is **3PARdata Fibre Channel Disk** and its device name is **naa.60002ac0000000000000000600008e5b**. All of this is done using the following `esxcli` command:

```
esxcli storage core device list
```

The output is shown in the following screenshot:

```
~ # esxcli storage core device list
naa.60002ac00000000000000000600008e5b
Display Name: 3PARdata Fibre Channel Disk (naa.60002ac00000000000000000600008e5b)
Has Settable Display Name: true
Size: 1572864
Device Type: Direct-Access
Multipath Plugin: NMP
Devfs Path: /vmfs/devices/disks/naa.60002ac00000000000000000600008e5b
Vendor: 3PARdata
Model: VV
Revision: 3123
SCSI Level: 6
Is Pseudo: false
Status: on
Is RDM Capable: true
Is Local: false
Is Removable: false
Is SSD: false
Is Offline: false
Is Perennially Reserved: false
Queue Full Sample Size: 0
Queue Full Threshold: 0
Thin Provisioning Status: yes
Attached Filters:
VAAI Status: supported
Other UUIDs: vml.020002000060002ac0000000000000000600008e5b565620202020
Is Local SAS Device: false
Is Boot USB Device: false
```

In the next example (which displays information about HBA on the ESXi host), you can execute the esxcli command shown to collect the result, which is given for your reference. You can find two HBAs installed on this ESXi host. It is the QLogic FC adapter and the WWN of each HBA. The **vmhba2** is **QLogic Corp ISP2532-based 8Gb Fibre Channel to PCI Express HBA**, and its WWN is **5001438022429fe94**. The **vmhba3** is **QLogic Corp ISP2532-based 8Gb Fibre Channel to PCI Express HBA**, and its WWN is **5001438022429fe96**:

**esxcli storage core adapter list**

```
~ # esxcli storage core adapter list
HBA Name Driver Link State UID Description

vmhba0 ata_piix link-n/a sata.vmhba0 (0:0:31.2) Intel Corporation Patsburg 4 port SATA IDE Controller
vmhba1 hpsa link-n/a sas.500143802798b030 (0:2:0.0) Hewlett-Packard Company Smart Array P420i
vmhba2 qla2xxx link-up fc.500143802429fe95:500143802429fe94 (0:7:0.0) QLogic Corp ISP2532-based 8Gb Fibre Channel to PCI Express HBA
vmhba3 qla2xxx link-up fc.500143802429fe97:500143802429fe96 (0:7:0.1) QLogic Corp ISP2532-based 8Gb Fibre Channel to PCI Express HBA
vmhba32 ata_piix link-n/a sata.vmhba32 (0:0:31.2) Intel Corporation Patsburg 4 port SATA IDE Controller
~ #
```

This example displays information about the vSphere storage path. You can execute the following `esxcli` command to collect the result, which is given for your reference. You will notice that the path status of this device is active, and find the target WWPN of this device. The adapter WWPN is **50:01:43:80:24:29:fe:96** and the target WWPN is **20:12:00:02:ac:00:8e:5b**:

```
esxcli storage core path list
```

```
~ # esxcli storage core path list
fc.500143802429fe97:500143802429fe96-fc.2ff70002ac008e5b:20120002ac008e5b-naa.60002ac0000000000000000200008e5b
UID: fc.500143802429fe97:500143802429fe96-fc.2ff70002ac008e5b:20120002ac008e5b-naa.60002ac0000000000000000200008e5b
Runtime Name: vmhba3:C0:T1:L0
Device: naa.60002ac0000000000000000200008e5b
Device Display Name: 3PARdata Fibre Channel Disk (naa.60002ac0000000000000000200008e5b)
Adapter: vmhba3
Channel: 0
Target: 1
LUN: 0
Plugin: NMP
State: active
Transport: fc
Adapter Identifier: fc.500143802429fe97:500143802429fe96
Target Identifier: fc.2ff70002ac008e5b:20120002ac008e5b
Adapter Transport Details: WWNN: 50:01:43:80:24:29:fe:97 WWPN: 50:01:43:80:24:29:fe:96
Target Transport Details: WWNN: 2f:f7:00:02:ac:00:8e:5b WWPN: 20:12:00:02:ac:00:8e:5b
Maximum IO Size: 33553920
```

The following example lists the devices controlled by VMware NMP. You can execute the following `esxcli` command to collect the following result. You can find out what the path selection policy is that this device is using, this device **naa.60002ac0000000000000000600008e5b** is using **VMW\_PSP\_RR** (round robin) now:


```
esxcli storage nmp device list
```

```
~ # esxcli storage nmp device list
naa.60002ac0000000000000000600008e5b
Device Display Name: 3PARdata Fibre Channel Disk (naa.60002ac0000000000000000600008e5b)
Storage Array Type: VMW_SATP_ALUA
Storage Array Type Device Config: (implicit_support=on;explicit_support=off; explicit_allow=on;alua_followover=on;(TPG_id=256,TPG_state=A0))
Path Selection Policy: VMW_PSP_RR
Path Selection Policy Device Config: (policy=rr,iops=1000,bytes=10485760,useANC=0;lastPathIndex=2; NumIOsPending=9,numBytesPending=503808)
Path Selection Policy Device Custom Config:
Working Paths: vmhba2:C0:T0:L2, vmhba2:C0:T1:L2, vmhba3:C0:T0:L2, vmhba3:C0:T1:L2
Is Local SAS Device: false
Is Boot USB Device: false
```

In the following example (set the path selection policy for a device), you can execute the `esxcli` command shown to display the following result, which is given for your reference. You can set **Path Selection Policy** for a device `naa.60002ac00000000000000000000000008e5b` to a fixed path selection:

```
esxcli storage nmp device set --device <device name> --psp VMW_PSP_FIXED
```

```
~ # esxcli storage nmp device set --device naa.60002ac00000000000000000000000008e5b --psp VMW_PSP_FIXED
~ # esxcli storage nmp device list
naa.60002ac00000000000000000000000008e5b
 Device Display Name: SPARdata Fibre Channel Disk (naa.60002ac00000000000000000000000008e5b)
 Storage Array Type: VMW_SATP_ALUA
 Storage Array Type Device Config: (implicit_support=on;explicit_support=off; explicit_allow=on;alua_followover=on;(TPG_id=256,TPG_state=AO))
 Path Selection Policy: VMW_PSP_FIXED
 Path Selection Policy Device Config: (preferred=vmhba2:C0:T0:L2;current=vmhba2:C0:T0:L2)
 Path Selection Policy Device Custom Config:
 Working Paths: vmhba2:C0:T0:L2
 Is Local SAS Device: false
 Is Boot USB Device: false
```

|                                                                                   |                                                           |
|-----------------------------------------------------------------------------------|-----------------------------------------------------------|
|  | <b>VMW_PSP_FIXED:</b> Fixed path selection                |
|                                                                                   | <b>VMW_PSP_MRU:</b> The most recently used path selection |
|                                                                                   | <b>VMW_PSP_RR:</b> Round robin path selection             |

There is another very useful command in vSphere shell that is used to collect data at the virtual SCSI device level in the kernel. It is the `vscsiStats` command. It creates ESXi disk I/O workload characterization per virtual disk, and can collect the I/O size, seek distance, outstanding I/Os, and latency. Because this command is unaware of the storage implementation, we can use it to collect latency statistics for all storage configurations. The following is the procedure for collecting a virtual machine's virtual SCSI disk data using the `vscsiStats` command:

1. Suppose I want to collect the redhat6\_3 virtual machine's data to list the world group ID using the following command. The ID is 3917848:

```
vscsiStats -l
```

```
~ # vscsiStats -l
Virtual Machine worldGroupID: 1574483, Virtual Machine Display Name: NSDU, Virtual Machine Config File: /vmfs/volumes/52d54416-2616153d-7bd0-2c768a537fa4/NSDU/NSDU.vmx, (
 Virtual SCSI Disk handleID: 8198 (ide0:0)
 Virtual SCSI Disk handleID: 8199 (scsi0:0)
)
Virtual Machine worldGroupID: 3917848, Virtual Machine Display Name: redhat6_3, Virtual Machine Config File: /vmfs/volumes/531ffd97-1318d660-1559-2c768a537fa4/redhat6_3/redhat6_3.vmx, (
 Virtual SCSI Disk handleID: 8200 (scsi0:0)
 Virtual SCSI Disk handleID: 8201 (scsi1:0)
)
Virtual Machine worldGroupID: 3917870, Virtual Machine Display Name: AH2.4, Virtual Machine Config File: /vmfs/volumes/52d54416-2616153d-7bd0-2c768a537fa4/AH2.4/AH2.4.vmx, (
 Virtual SCSI Disk handleID: 8202 (scsi0:0)
 Virtual SCSI Disk handleID: 8203 (scsi0:1)
)
Virtual Machine worldGroupID: 3917903, Virtual Machine Display Name: Unisphere-Central, Virtual Machine Config File: /vmfs/volumes/531ffd97-1318d660-1559-2c768a537fa4/Unisphere-Central/Unisphere-Central.vmx, (
 Virtual SCSI Disk handleID: 8204 (scsi0:0)
 Virtual SCSI Disk handleID: 8205 (scsi0:1)
)
Virtual Machine worldGroupID: 3917923, Virtual Machine Display Name: UCSPE, Virtual Machine Config File: /vmfs/volumes/531ffd97-1318d660-1559-2c768a537fa4/UCSPE/UCSPE.vmx, (
 Virtual SCSI Disk handleID: 8206 (ide0:0)
 Virtual SCSI Disk handleID: 8207 (ide0:1)
 Virtual SCSI Disk handleID: 8208 (ide1:0)
 Virtual SCSI Disk handleID: 8209 (ide1:1)
)
```


The following table lists the details of each option for `vscsiStats`:

| Related options | Description                                                                                                                                                                                                                                                                                                          |
|-----------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| -h              | Print the usage.                                                                                                                                                                                                                                                                                                     |
| -l              | List the available virtual machines and their virtual disks.                                                                                                                                                                                                                                                         |
| -r              | Reset the stats.                                                                                                                                                                                                                                                                                                     |
| -s              | Start the <code>vscsiStats</code> collection; exclusion of -x.                                                                                                                                                                                                                                                       |
| -x              | Stop the <code>vscsiStats</code> collection; exclusion of -s.                                                                                                                                                                                                                                                        |
| -w              | Specifies a wordID to use for this operation.                                                                                                                                                                                                                                                                        |
| -i              | Specifies a <code>vscsi</code> handleID to use for this operation.                                                                                                                                                                                                                                                   |
| -p              | This prints the current histograms for the specified. It may be used in conjunction with -w and -i. The <code>histoType</code> must be only one of these: <code>all</code> , <code>ioLength</code> , <code>seekDistance</code> , <code>outstandingIOs</code> , <code>latency</code> , or <code>interarrival</code> . |

2. After finding the world ID, **3917848**, start collecting data for that ID using the following command:


```
vscsiStats -s -w 3917848
```

```
~ # vscsiStats -x -w 3917848
vscsiStats: Stopping all Vscsi stats collection for worldGroup 3917848, handleID 8200 (scsi0:0)
Success.
vscsiStats: Stopping all Vscsi stats collection for worldGroup 3917848, handleID 8201 (scsi1:0)
Success.
~ #
```

[  -s: This option is used to start the vscsiStats data collection  
-w: This option is used to specify the world ID  
The collection will automatically stop after about 30 minutes. ]

3. To display the data collection, use this command:

```
/usr/lib/vmware/bin/vscsiStats -p all -c
```

[  -p: Display histograms and select the histogram type  
-c: Display the result, which is separated by a comma ]

The following result is for your reference. You can change the I/O size, seek distance, outstanding I/Os, and latency.

```
~ # /usr/lib/vmware/bin/vscsiStats -p all -c
Histogram: IO lengths of commands,virtual machine worldGroupID,3917848,virtual disk handleID,8200 (scsi0:0)
min,4096
max,28672
mean,5006
count,1372
Frequency,Histogram Bucket Limit
0,512
0,1024
0,2048
0,4096
1190,4096
0,8191
133,8192
18,16383
12,16384
19,32768
0,49152
0,65535
0,65536
0,81920
0,131072
0,262144
0,524288
0,524288
Histogram: IO lengths of Read commands,virtual machine worldGroupID,3917848,virtual disk handleID,8200 (scsi0:0)
min,0
max,0
```

4. After viewing the data collection, you need to stop it manually and reset the counters using the following command, because it will cause some performance degradation. The following is the result of stopping data collection of the **3917848** world ID using this command:

```
vscsiStats -x -w 3917848
```

```
~ # vscsiStats -s -w 3917848
vscsiStats: Starting Vscsi stats collection for worldGroup 3917848, handleID 8200 (scsi0:0)
Success.
vscsiStats: Starting Vscsi stats collection for worldGroup 3917848, handleID 8201 (scsi1:0)
Success.
~ #
```



The -x option is used to stop the vscsiStats data collection

## Troubleshooting vSphere storage performance problems

Storage performance problems can be caused by different factors. In most cases, problems exist at the storage level, for example, overloaded storage and slow storage. We will list the main causes of these two problems. The main cause of overloaded storage is that the vSphere administrator doesn't know what the IOPs requirement of an application that is being run in a virtual machine is, if some application performs sequential I/O and another application performs random I/O. The application vendor or end user needs to give this information to the vSphere administrator or storage administrator. As best practice, don't mix these two different types of I/O in the same RAID group, because it will cause some performance degradation. In other situations, entry-level storage is the most common cause of performance problems in a vSphere environment. It can provide limited disk IOPs, high latency of disk can cause performance problems. You can monitor the device latency using vSphere performance charts or esxtop/resxtop tools; or you can refer to the previous section.

In the following section, we have listed some examples of performance problems in virtual machines.



## First scenario

In scenario 1, some VMware users have found that powering on their virtual machine is very slow. It takes more than 4 to 5 minutes to power on. Firstly, you might want to check the ESXi host's disk latency in the disk metrics using vSphere performance charts or esxtop/resxtop tools. If you find that the disk latency is very high, you then need to identify whether it is a software issue (configuration) or a hardware issue. If the ESXi host is connected to the iSCSI storage, you should check the connection of the storage and the ESXi host; it may be a software configuration issue. For details, you can refer to *Chapter 9, Troubleshooting vSphere iSCSI Storage*. If it is not a software configuration issue, you might focus on the SAN configuration level. Our recommendation is that you migrate this virtual machine to other data stores (LUNs) using storage vMotion, don't mix the high-I/O virtual machine runs on the same RAID group, and check the path selection policy of the data store based on the storage vendor's recommendation.

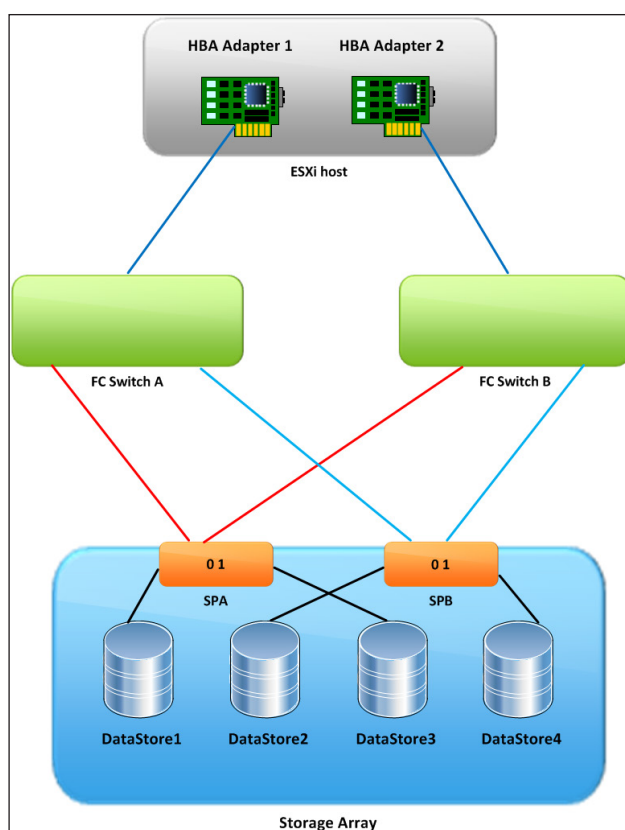
## Second scenario

In scenario 2, a VMware user has found that the response time of the virtual machine suddenly becomes very slow. This behavior might relate to a resource problem. Firstly, you need to identify what the status of resource of this virtual machine is, and whether it has enough resources for that virtual machine to power on and provide enough resource. In most cases, the response time of a virtual machine can get slower if the virtual machine is suddenly out of the resources. You need to identify whether it is a CPU or physical memory resource issue using vSphere performance charts or esxtop/resxtop tools. Our recommendation is to increase the resources (CPU and memory) in this virtual machine, and set up the ESX's resource pool to control each virtual machine in the vSphere environment.

Since the storage infrastructure has a different configuration, providing one prefect and good-performance solution is difficult. If you want to provide one good-performance storage solution for the virtual environment, you need to follow the storage vendor's configuration recommendations for configuring the vSphere host connectivity of storage. To optimize storage performance, separate the I/O loading across multiple storage processors. Make sure that each vSphere host has a sufficient number of HBAs to allow maximum throughput. In the rest of the section, we will share the best practices for vSphere host connectivity of the storage array. Install 2 x single-port FC adapters on each vSphere host. 2 x single-port FC adapters are better than a 1 x dual-port FC adapter, which is a VMware recommendation. It has redundancy at the adapter and port levels. To define zoning, you can define the 2 x zoning of HBA1, the array's SPA0 and HBA1, and its SPB1 on **FC Switch A**. You can also define the other 2 x zoning of HBA2, the array's SPA1 and HBA2, and its SPB0 on **FC Switch B**. Make sure that each HBA can access the storage array.

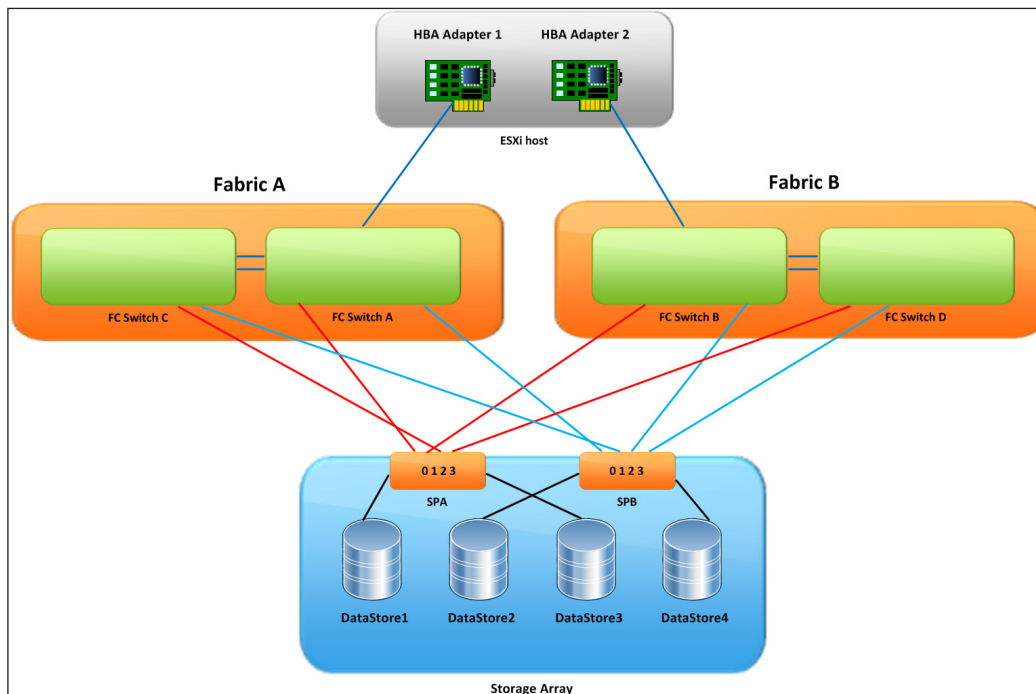
This is single-initiator zoning, which is a VMware recommendation setting. According to this configuration, I/O loading can across the HBA into four logical paths and has load balancing based on the ESXi path selection policy.

The following diagram shows the storage infrastructure; it's for your reference:



In another scenario, the existing SAN switch topology has two large fabrics that include two members. As best practice, it is installed 2 x single-port FC adapters on each vSphere host. Define the 4 x zoning of HBA1, the array's SPA0 and HBA1, its SPA1 and HBA1, its SPB0 and HBA1, and only its SPB0 on **Fabric A**. Then define the other 4 x zoning of HBA2, the array's SPA2 and HBA2, its SPA3 and HBA2, then its SPB2 and HBA2, and finally its SPB3 on **Fabric B**. Make sure that each HBA can access the storage array.

According to this configuration, the I/O loading also can across the HBA into eight logical paths, and has load balancing based on the ESXi path selection policy.



As VMware's best practice, if the SAN storage is an active-passive array, it is recommended that the path selection policy of vSphere be MRU. If the SAN storage is active-active array, it is recommended that the path selection policy of vSphere be round robin.

Best practices are actually recommendations for configuration and operation. In the following section, we have listed the best practices for storage performance:

- According to the requirement of the application in a virtual machine, configure each LUN with the correct RAID level and storage characteristics
- Do not mix the Fibre Channel or iSCSI HBA from different vendors on the same vSphere host
- Only allow one vmdk file map to one system drive in the virtual machine
- Install two single ports, Fibre Channel or iSCSI HBA, on each vSphere host
- Configure the value of queue depth for the HBA based on the vendor's recommendation

- Create a single VMFS volume per LUN on each vSphere host
- For performance, define single-initiator zoning on the Fibre Channel Switch
- Present the same LUN target ID number to all vSphere hosts
- In SAN configurations, separate the I/O loading over the available path to storage devices

## Summary

In this chapter, you learned what the concepts of storage virtualization are and how to monitor storage metrics using vSphere performance charts and esxtop/resxtop. We also covered management of vSphere storage using the command line. Then you looked at a list of some scenarios that describe existing performance problems, and provided solutions for those performance problems. We also saw what VMware's recommended configuration for the storage array and vSphere hosts is, according to the best practices.

In the next chapter, we will see how to find the location of VMware ESXi and vCenter, and identify the log used to troubleshoot storage performance problems.



# 6

## vSphere Storage Configuration Settings

In the previous chapter, you learned how to handle the problem of storage performance using vSphere tools. Now we will go through some troubleshooting procedures to troubleshoot the paths in vSphere hosts, for example, configuring storage path masking. You will also learn where the vSphere storage log is located.

In this chapter, we will cover the following topics:

- vSphere storage components
- LUN masking
- vSphere 5 storage maximums
- Identifying vSphere log used to troubleshoot storage problem

### vSphere storage components

In *Chapter 4, Storage Scalability*, you learned about the **Pluggable Storage Architecture (PSA)**, VMware SATP, and VMware PSP. What is the relationship between each component in vSphere hosts? During vSphere storage troubleshooting, you should understand the difference between device names, identifiers, and runtime names. The device name can usually be defined by the vSphere administrator, storage identifiers are unique IDs associated with devices, and the runtime name is a path. In the following screenshot, you can see vSphere's data store configuration. There are two types of identifier: 1 x local disk and 4 x FC channel disk. The `naa` identifier is established by the Network Address Authority (NAA).

This identifier is similar to the MAC identifier in the network interface adapter. In the following screenshot, you can also see two types of drive. One is SAN storage's disk, which is labeled **Fibre Channel**. The other is a local disk. It is labeled **Block Adapter**.

**View:** Datastores Devices

**Devices**

| Name                                        | Identifier                             | Runtime Name    | Operational State | Transport     |
|---------------------------------------------|----------------------------------------|-----------------|-------------------|---------------|
| 3PARdata Fibre Channel Disk (naa.60002a...  | naa.60002ac00000000000000000600008e... | vmhba2:C0:T0:L2 | Mounted           | Fibre Channel |
| 3PARdata Fibre Channel Disk (naa.60002a...  | naa.60002ac00000000000000000c00008e... | vmhba2:C0:T0:L3 | Mounted           | Fibre Channel |
| 3PARdata Fibre Channel Disk (naa.60002a...  | naa.60002ac00000000000000000200008e... | vmhba2:C0:T0:L0 | Mounted           | Fibre Channel |
| 3PARdata Fibre Channel Disk (naa.60002a...  | naa.60002ac00000000000000000b00008e... | vmhba2:C0:T0:L1 | Mounted           | Fibre Channel |
| HP Serial Attached SCSI Disk (naa.600508... | naa.600508b1001c0bd40a7e1a4da8c2c2...  | vmhba1:C0:T0:L1 | Mounted           | Block Adapter |

The following screenshot shows LUN, which is the iSCSI LUN (**iSCSI Software Adapter**):

**Details**

**vmhba36** [Properties...](#)

Model: iSCSI Software Adapter

iSCSI Name: iqn.1998-01.com.vmware:mfmrs006-6b4f37fc

iSCSI Alias:

Connected Targets: 6    Devices: 3    Paths: 6

**View:** Devices Paths

| Name                               | Runtime Name     | Operational State | LUN | Type | Drive Type | Transport | Cap |
|------------------------------------|------------------|-------------------|-----|------|------------|-----------|-----|
| EQLOGIC iSCSI Disk (naa.64ed2a6... | vmhba36:C1:T2:L0 | Mounted           | 0   | disk | Non-SSD    | iSCSI     | 2.  |
| EQLOGIC iSCSI Disk (naa.64ed2a6... | vmhba36:C1:T1:L0 | Mounted           | 0   | disk | Non-SSD    | iSCSI     | 1.  |
| EQLOGIC iSCSI Disk (naa.64ed2a6... | vmhba36:C0:T0:L0 | Mounted           | 0   | disk | Non-SSD    | iSCSI     | 2.  |

Runtime names are basically path names that include an adapter, a channel, a target, and an LUN number. These are shown in the following screenshot:

**View:** Devices Paths

| Name                                       | Identifier                             | Runtime Name    | Operational State | LUN | Type |
|--------------------------------------------|----------------------------------------|-----------------|-------------------|-----|------|
| 3PARdata Fibre Channel Disk (naa.60002a... | naa.60002ac00000000000000000200008e... | vmhba2:C0:T0:L0 | Mounted           | 0   | disk |
| 3PARdata Fibre Channel Disk (naa.60002a... | naa.60002ac00000000000000000b00008e... | vmhba2:C0:T0:L1 | Mounted           | 1   | disk |
| 3PARdata Fibre Channel Disk (naa.60002a... | naa.60002ac00000000000000000600008e... | vmhba2:C0:T0:L2 | Mounted           | 2   | disk |
| 3PARdata Fibre Channel Disk (naa.60002a... | naa.60002ac00000000000000000c00008e... | vmhba2:C0:T0:L3 | Mounted           | 3   | disk |

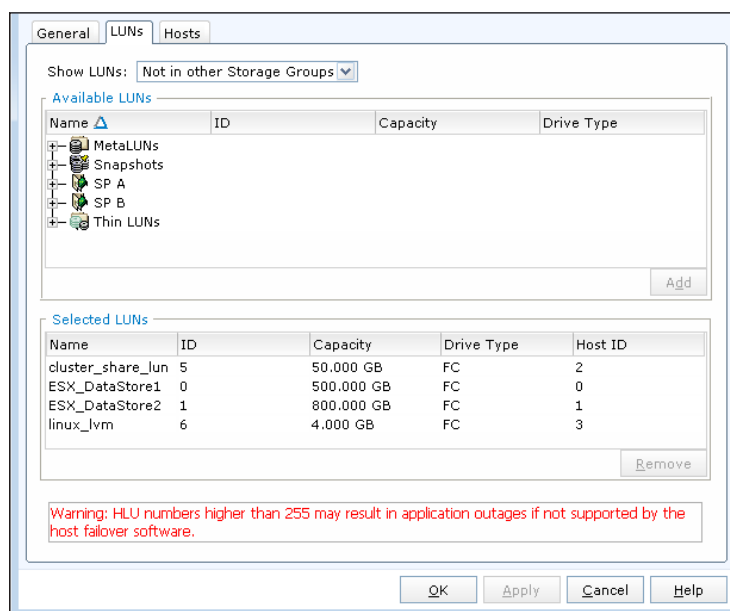
This table lists the description of each item for **Runtime Name**:

| Term     | Description                                                                                                                     |
|----------|---------------------------------------------------------------------------------------------------------------------------------|
| vmhba<n> | This is the physical storage HBA adapter on the host.                                                                           |
| C<n>     | This is the storage channel number. Software iSCSI initiators use the channel number to show multiple paths to the same target. |
| T<n>     | The target number that is shared by different ESXi hosts might not have the same target numbers.                                |
| L<n>     | This is the LUN number that shows the location of the LUN with the target. This number is provided by the storage array.        |

## LUN masking

**LUN masking** can control which LUNs are visible to each vSphere host. This is the opposite of zoning, where the storage array configuration determines which LUNs are visible to a host. This feature allows multiple vSphere hosts to be connected to a storage with multiple LUNs, while allowing only one vSphere host, which you specify, to see some particular LUNs. This feature is the same as EMC CLARiON or VNX provide LUN masking in the storage group at the array level. You can add the host and LUNs to a storage group, and then the host will only be able to see those LUNs.


Here is the GUI meant for providing LUN masking to the Storage Group in the EMC array:





Now we'll go through an example showing how you can operate this on an ESXi host. Here is the procedure:

1. First, we should find out which LUN we want to mask. We need to display the LUN with VMFS volumes using the `esxcfg-scsidevs -m` command. In the following example, we find the device ID (it starts with naa). Our device ID is `naa.60002ac00000000000000000600008e5b`.


 The `-m` option displays only the LUN with VMFS volumes. 

| Option | Description                                                                                |
|--------|--------------------------------------------------------------------------------------------|
| -l     | List all paths in the system with their detailed information.                              |
| -L     | List all paths with abbreviated information.                                               |
| -m     | List all paths with adapter and device mappings.                                           |
| -b     | List all devices with their corresponding paths.                                           |
| -s     | Set the state for a specific LUN path. This requires the path UID or path Runtime Name in. |
| -G     | List all multipathing plugins loaded into the system.                                      |
| -d     | Used to filter the list commands to display only a specific device.                        |
| -r     | Restore path setting to configured values on system start.                                 |

Here is the result that collects the device ID using the `esxcfg-scsidevs -m` command:

```
~ # esxcfg-scsidevs -m
naa.60002ac00000000000000000600008e5b:1 /vmfs/devices/disks/naa.60002ac000000000
0000000600008e5b:1 53d5e8b4-fc2fa826-9b63-2c44fd7e370c 0 Infra_DataStore
naa.60002ac00000000000000000200008e5b:1 /vmfs/devices/disks/naa.60002ac000000000
0000000200008e5b:1 5303059b-41d7c70d-6d6a-2c44fd7e370c 0 Network_DataStore
naa.600508b1001cd9a83d03ef4f03c79972:3 /vmfs/devices/disks/naa.600508b1001cd9a8
3d03ef4f03c79972:3 523e498f-df33bb36-fa20-2c44fd7e370c 0 Local_DataStore02
~ #
```

2. After identifying the device ID, we have to find the path (or paths) of that LUN using the `esxcfg-mpath -L | grep naa.60002ac0000000000000000000600008e5b` command. We can see that there are four paths to this LUN, which are `C0:T1:L2`, `C0:T0:L2`, `C0:T1:L2`, and `C0:T0:L2`.

 The `-L` option displays a list of paths for this device ID. 

This screenshot shows collection of the number of paths using the `esxcfg-mpath -L | grep naa.60002ac00000000000000000600008e5b` command:

```
~ # esxcfg-mpath -L | grep naa.60002ac00000000000000000600008e5b
vmhba3:C0:T1:L2 state:active naa.60002ac00000000000000000600008e5b vmhba3 0 1 2 NMP active san fc.50014380242a8de
7:50014380242a8de6 fc.2ff70002ac008e5b:20120002ac008e5b
vmhba3:C0:T0:L2 state:active naa.60002ac00000000000000000600008e5b vmhba3 0 0 2 NMP active san fc.50014380242a8de
7:50014380242a8de6 fc.2ff70002ac008e5b:21110002ac008e5b
vmhba2:C0:T1:L2 state:active naa.60002ac00000000000000000600008e5b vmhba2 0 1 2 NMP active san fc.50014380242a8de
5:50014380242a8de4 fc.2ff70002ac008e5b:21120002ac008e5b
vmhba2:C0:T0:L2 state:active naa.60002ac00000000000000000600008e5b vmhba2 0 0 2 NMP active san fc.50014380242a8de
5:50014380242a8de4 fc.2ff70002ac008e5b:20110002ac008e5b
~ #
```

- Now we can create a new claim rule for LUN masking. We should check what claim rules exist in order not to use an existing claim rule number before creating a new claim rule. The following screenshot shows the output that displays the existing claim rule using of the `esxcli storage core claimrule list` command:

```
~ # esxcli storage core claimrule list
Rule Class Rule Class Type Plugin Matches

MP 0 runtime transport NMP transport=usb
MP 1 runtime transport NMP transport=sata
MP 2 runtime transport NMP transport=ide
MP 3 runtime transport NMP transport=block
MP 4 runtime transport NMP transport=unknown
MP 101 runtime vendor MASK_PATH vendor=DELL model=Universal Xport
MP 101 file vendor MASK_PATH vendor=DELL model=Universal Xport
MP 65535 runtime vendor NMP vendor=* model=*
~ #
```

- We can create a new claim rule by any rule number that isn't in the preceding list (the new claim rule number is set to 300 in this example). Now let's create the new claim rule the first path, C0:T1:L2, which is on the **vmhba3** adapter in the vSphere host. The command is `esxcli storage core claimrule add -r 300 -t location -A vmhba3 -C 0 -T 1 -L 2 -P MASK_PATH`.



- r option = Claim rule number
- t option = Indicates which type of matching is used for claim/unclaim or claimrule
- A option = Indicates the adapter of the paths to be used in this operation
- C option = Indicates the channel of the paths to be used in this operation
- T option = Indicates the target of the paths to be used in this operation
- L option = Indicates the LUN of the paths to be used in this operation
- P option = Indicates which PSA plugin should be used for this operation

- Due to this, the LUN that has four paths will still allow the LUN to be seen on other paths in the vSphere host. So, we need to mask the other paths. We'll use 301/302/303 for the rule number and C0:T0:L2/ C0:T1:L2/ C0:T0:L2 as the path. The adapter will still be vmhba3 and vmhba2. Use the following commands on the screen:

```
esxcli storage core claimrule add -r 301 -t location -A vmhba3 -C 0 -T 0 -L 2 -P MASK_PATH

esxcli storage core claimrule add -r 302 -t location -A vmhba2 -C 0 -T 1 -L 2 -P MASK_PATH

esxcli storage core claimrule add -r 303 -t location -A vmhba2 -C 0 -T 0 -L 2 -P MASK_PATH
```

```
~ # esxcli storage core claimrule list
Rule Class Rule Class Type Plugin Matches

MP 0 runtime transport NMP transport=usb
MP 1 runtime transport NMP transport=sata
MP 2 runtime transport NMP transport=ide
MP 3 runtime transport NMP transport=block
MP 4 runtime transport NMP transport=unknown
MP 101 runtime vendor MASK_PATH vendor=DELL model=Universal Xport
MP 101 file vendor MASK_PATH vendor=DELL model=Universal Xport
MP 65535 runtime vendor NMP vendor=* model=*
~ # esxcli storage core claimrule add -r 300 -t location -A vmhba3 -C 0 -T 1 -L 2 -P MASK_PATH
~ # esxcli storage core claimrule add -r 301 -t location -A vmhba3 -C 0 -T 0 -L 2 -P MASK_PATH
~ # esxcli storage core claimrule add -r 302 -t location -A vmhba2 -C 0 -T 1 -L 2 -P MASK_PATH
~ # esxcli storage core claimrule add -r 303 -t location -A vmhba2 -C 0 -T 0 -L 2 -P MASK_PATH
~ #
```

- Now you can execute the `esxcli storage core claim rule list` again, and you will see the new rules, 300/301/302/303, in the claim rule list. But you will notice that the class for these new rules is shown as `file`, which means that it is required to load our new rules into the runtime.

The following result displays the new claim rule 300/301/302/303 using the `esxcli storage core claimrule list` command:

```
~ # esxcli storage core claimrule list
Rule Class Rule Class Type Plugin Matches

MP 0 runtime transport NMP transport=usb
MP 1 runtime transport NMP transport=sata
MP 2 runtime transport NMP transport=ide
MP 3 runtime transport NMP transport=block
MP 4 runtime transport NMP transport=unknown
MP 101 runtime vendor MASK_PATH vendor=DELL model=Universal Xport
MP 101 file vendor MASK_PATH vendor=DELL model=Universal Xport
MP 300 file location MASK_PATH adapter=vmhba3 channel=0 target=1 lun=2
MP 301 file location MASK_PATH adapter=vmhba3 channel=0 target=0 lun=2
MP 302 file location MASK_PATH adapter=vmhba2 channel=0 target=1 lun=2
MP 303 file location MASK_PATH adapter=vmhba2 channel=0 target=0 lun=2
MP 65535 runtime vendor NMP vendor=* model=*
```

7. Load the new claim rule into the runtime using the `esxcli storage core claimrule load` command. Then we execute the `esxcli storage core claimrule list` command again. You can see that each new claim rule has class displayed as runtime.



The new rule is loaded in `/etc/vmware/esx.conf`, but it isn't loaded into the runtime. You execute the `esxcli storage core claimrule load` command to replace the current rules in the VMkernel with the modified rules from the configuration file.

The following result shows that each new claim rule has runtime as its class, using the `esxcli storage core claimrule list` command:

```
~ # esxcli storage core claimrule load
~ # esxcli storage core claimrule list
```

| Rule     | Class   | Type      | Plugin    | Matches                                 |
|----------|---------|-----------|-----------|-----------------------------------------|
| MP 0     | runtime | transport | NMP       | transport=usb                           |
| MP 1     | runtime | transport | NMP       | transport=sata                          |
| MP 2     | runtime | transport | NMP       | transport=ide                           |
| MP 3     | runtime | transport | NMP       | transport=block                         |
| MP 4     | runtime | transport | NMP       | transport=unknown                       |
| MP 101   | runtime | vendor    | MASK_PATH | vendor=DELL model=Universal Xport       |
| MP 101   | file    | vendor    | MASK_PATH | vendor=DELL model=Universal Xport       |
| MP 300   | runtime | location  | MASK_PATH | adapter=vmhba3 channel=0 target=1 lun=2 |
| MP 300   | file    | location  | MASK_PATH | adapter=vmhba3 channel=0 target=1 lun=2 |
| MP 301   | runtime | location  | MASK_PATH | adapter=vmhba3 channel=0 target=0 lun=2 |
| MP 301   | file    | location  | MASK_PATH | adapter=vmhba3 channel=0 target=0 lun=2 |
| MP 302   | runtime | location  | MASK_PATH | adapter=vmhba2 channel=0 target=1 lun=2 |
| MP 302   | file    | location  | MASK_PATH | adapter=vmhba2 channel=0 target=1 lun=2 |
| MP 303   | runtime | location  | MASK_PATH | adapter=vmhba2 channel=0 target=0 lun=2 |
| MP 303   | file    | location  | MASK_PATH | adapter=vmhba2 channel=0 target=0 lun=2 |
| MP 65535 | runtime | vendor    | NMP       | vendor=* model=*                        |

- The current path of these LUNs is claimed by the NMP plugin (rule 65535). We need to disassociate with the NMP plugin and associate with the new plugin (MASK\_PATH). The last procedure is to unclaim all paths of that device ID (naa.60002ac00000000000000000600008e5b) and then reclaim them in new claim rules using the following commands:

```
esxcli storage core claiming reclaim -d
naa.60002ac00000000000000000600008e5b
```

```
~ # esxcli storage core claiming reclaim -d naa.60002ac00000000000000000600008e5b
```

Execute `esxcfg-mpath -L | grep naa.60002ac00000000000000000600008e5b` again after reclaiming the device. Now it will show 0 path, as shown in this screenshot:

```
~ # esxcfg-mpath -L | grep naa.60002ac00000000000000000600008e5b
~ #
```

The following is a brief list of commands you will need to run if you want to unmask those four paths to that LUN in vSphere host, you can reference as follows:

Steps 1 to 4 remove `claimrule` (ID 300/301/302/303). Step 5 involves replacement of the current rules in the VMkernel with the modified rules from the configuration file. Steps 6 to 9 are used to change the rule and unclaim the paths. Steps 10 and 11 are used to rescan `vmhba3` and `vmhba2`:

1. `esxcli storage core claimrule remove -r 300`
2. `esxcli storage core clamrule remove -r 301`
3. `esxcli storage core claimrule remove -r 302`
4. `esxcli storage core claimrule remove -r 303`
5. `esxcli storage core claimrule load`
6. `esxcli storage core claiming unclaim -t location -A vmhba3 -C 0 -T 1 -L 2`
7. `esxcli storage core claiming unclaim -t location -A vmhba3 -C 0 -T 0 -L 2`
8. `esxcli storage core claiming unclaim -t location -A vmhba2 -C 0 -T 1 -L 2`
9. `esxcli storage core claiming unclaim -t location -A vmhba2 -C 0 -T 0 -L 2`
10. `esxcli storage core adapter rescan -A vmhba3`
11. `esxcli storage core adapter rescan -A vmhba2`

## vSphere 5 storage maximums

In some cases, the vSphere administrator isn't aware of the maximum configuration or limitations of the vSphere storage, for example, the maximum volume size of VMFS volume, or number of targets per HBA adapter. It gives some configuration errors during vSphere storage configuration. It has a different maximum configuration and limitation in each edition of vSphere. The following table shows the common maximum configuration settings in a vSphere 5.1 host:

| Items                                   | Maximum |
|-----------------------------------------|---------|
| LUNs per ESXi host (FC Channel)         | 256     |
| LUN size (FC Channel)                   | 64 TB   |
| LUN ID (FC Channel)                     | 255     |
| Number of HBAs of any type (FC Channel) | 8       |

| Items                                                                                                                                        | Maximum              |
|----------------------------------------------------------------------------------------------------------------------------------------------|----------------------|
| HBA ports (FC Channel)                                                                                                                       | 8                    |
| Targets per HBA (FC Channel)                                                                                                                 | 256                  |
| Raw device mapping size (virtual mode) – VMFS3                                                                                               | 2 TB minus 512 bytes |
| Raw device mapping size (physical mode) – VMFS3<br>(Note that if the presented LUN is greater than 2 TB)                                     | 2 TB minus 512 bytes |
| Block size – VMFS3                                                                                                                           | 8 MB                 |
| File size (1 MB block size) – VMFS3                                                                                                          | 256 GB               |
| File size (2 MB block size) – VMFS3                                                                                                          | 512 GB               |
| File size (4 MB block size) – VMFS3                                                                                                          | 1 TB                 |
| File size (8 MB block size) – VMFS3                                                                                                          | 2 TB minus 512 bytes |
| Raw device mapping size (physical mode) – VMFS5                                                                                              | 64 TB                |
| Raw device mapping size (virtual mode) – VMFS5                                                                                               | 2 TB minus 512 bytes |
| Block size – VMFS5<br>(Note that 1 MB is the default block size. Upgraded VMFS5 volumes will inherit the VMFS3 block size value)             | 1 MB                 |
| File size – VMFS5<br>(Note that the maximum file size for upgraded VMFS5 is 2 TB minus 512 bytes, irrespective of the filesystem block size) | 2 TB minus 512 bytes |



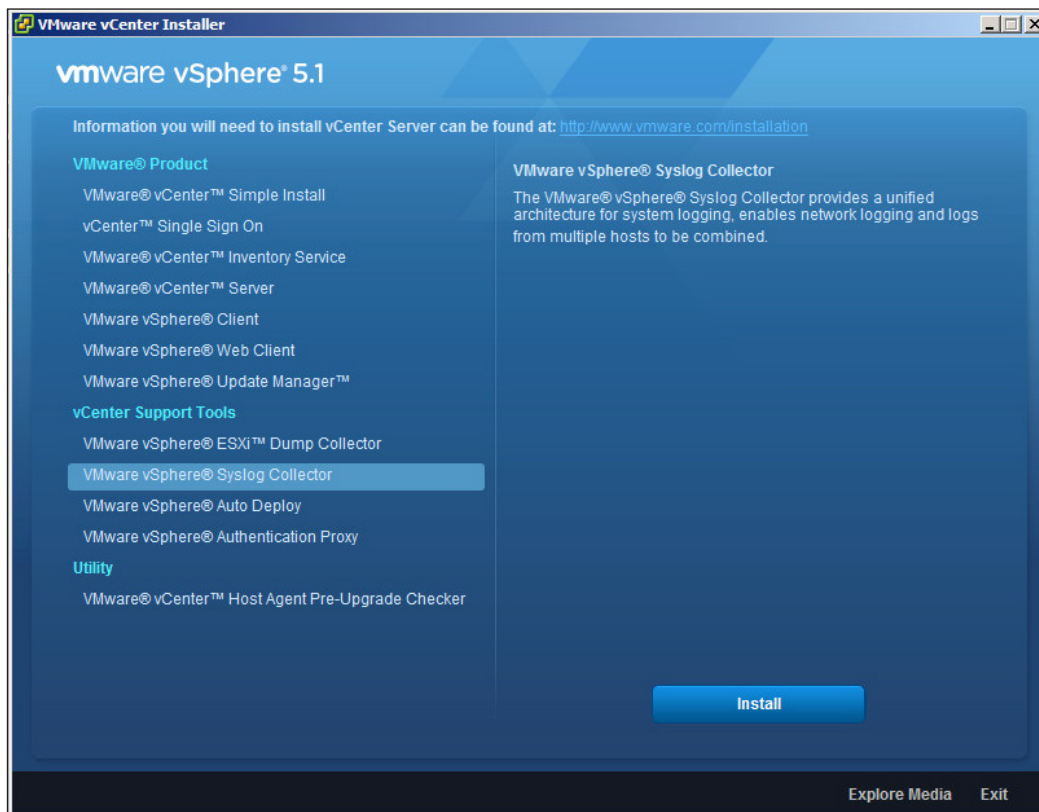
For other maximum configurations in each edition of vSphere, you can refer to the official VMware website at <http://www.vmware.com/support/pubs/vsphere-esxi-vcenter-server-pubs.html>.

## Identifying the vSphere log used to troubleshoot a storage problem

In *Chapter 2, Getting Started with vSphere Management Assistant*, you learned about the location of vSphere and vCenter Server's logs and how to export the bundle log file. In most vSphere environments, the system administrator installs one syslog server to store the vSphere server's log for any troubleshooting required.

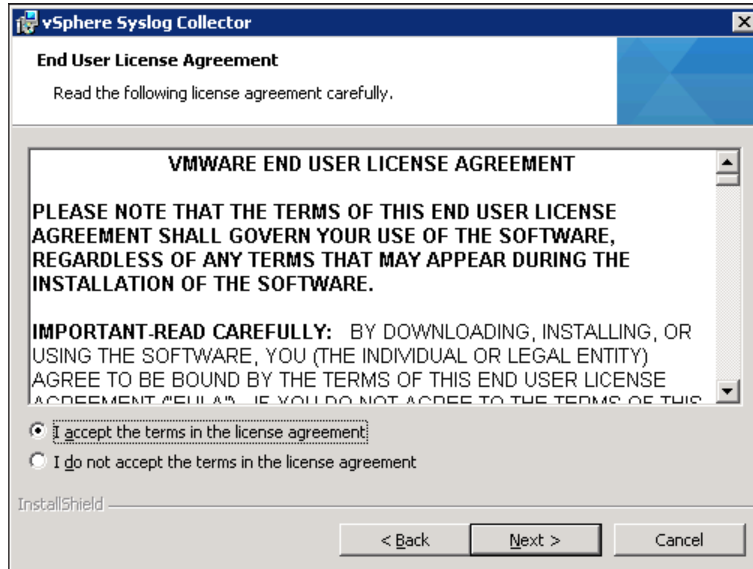
You can set up a third-party syslog server (for example, a Linux Syslog server or Windows-based Syslog server) or VMware vSphere Syslog Collector. You can also set up a vSphere Syslog Collector on the vCenter server, or on a different machine that has a network connection with the vCenter Server. The following is the installation procedure of the vSphere Syslog Collector on the vCenter Server, which is installed on a Windows platform and is not using the vCenter appliance:

1. Select **VMware vSphere Syslog Collector** in the VMware **vCenter Support Tools** menu and click on the **Install** button, as shown here:

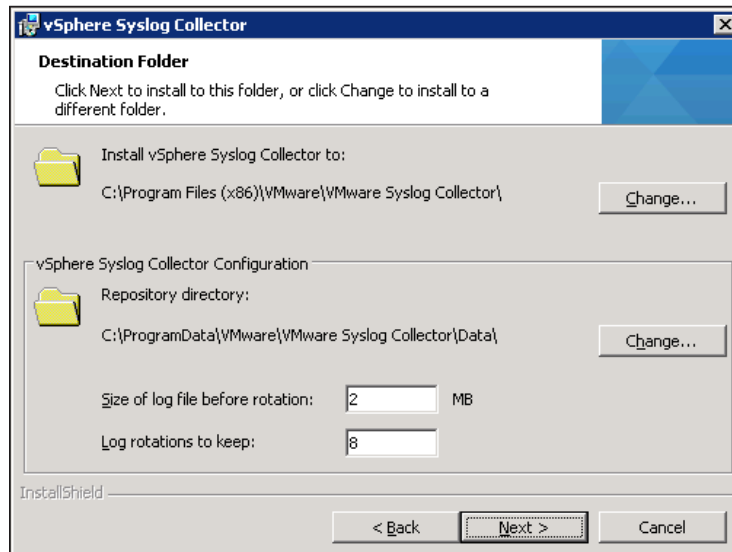




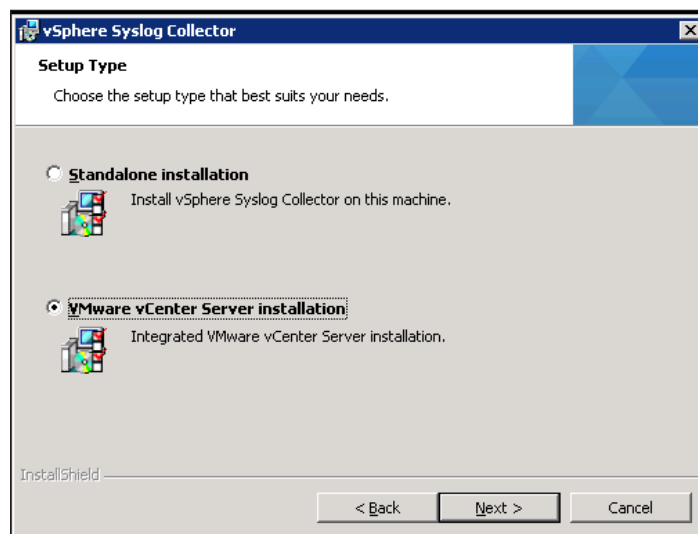
2. Accept the license agreement, as shown in this screenshot:



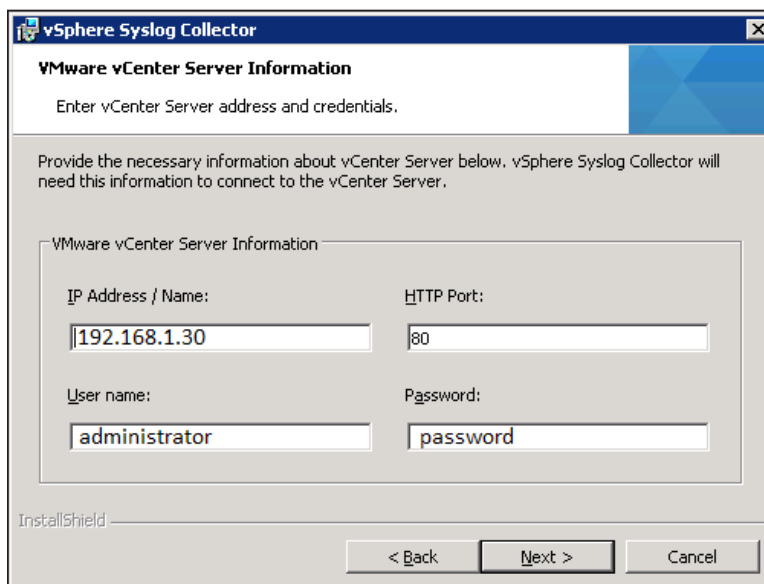
3. You can configure the size of the log file and log rotation. Then click on the **Next** button, as shown in the following screenshot:



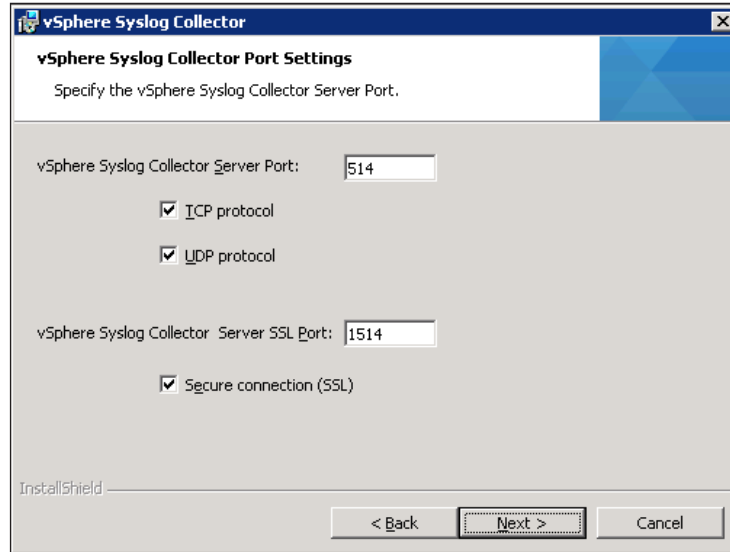
4. Select **Integrated VMware vCenter Server Installation** and click on the **Next** button, like this:



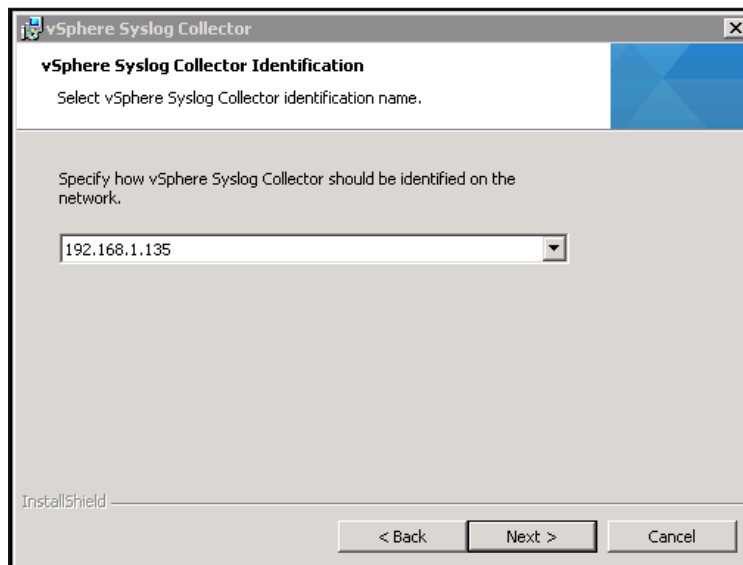
5. Input the **IP Address** of vCenter, its **User name**, and the **Password**. Then click on the **Next** button, as shown in this screenshot:



6. You can change **Syslog Collector Server Port** and **Syslog Collector Server SSL Port**. After that, click on **Next**, as shown in the following screenshot:



7. Identify the **vSphere Syslog Collector** in the network. Then click on the **Next** button. Finally, click on the **Install** button.



8. After you have finished installing the **vSphere Syslog Collector**, go to the vCenter Server's Windows services. Make sure that the **VMware vSphere Syslog Collector Service** has started after finishing the installation of the **vSphere Syslog Collector**.

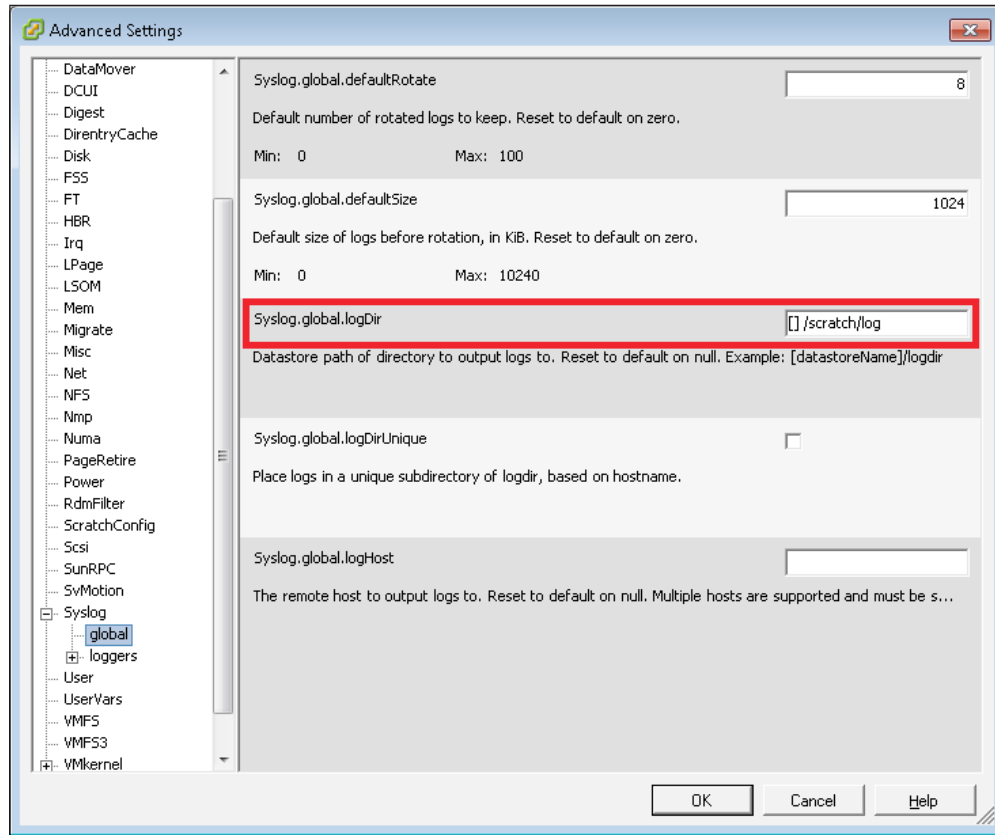
|                                    |                |         |                 |              |
|------------------------------------|----------------|---------|-----------------|--------------|
| VMware vCenter Inventory Service   | Provides c...  | Started | Automatic       | Local System |
| VMware vCenter Orchestrator C...   | VMware vC...   |         | Manual          | Local System |
| VMware vCenter Orchestrator S...   | Hosts the ...  |         | Manual          | Local System |
| VMware VirtualCenter Managem...    | Allows conf... | Started | Automatic (D... | Local System |
| VMware VirtualCenter Server        | Provides c...  | Started | Automatic (D... | Local System |
| VMware vSphere Profile-Driven S... | VMware vS...   | Started | Automatic       | Local System |
| VMware vSphere Syslog Collector    | Enables su...  | Started | Automatic       | Local System |
| VMware vSphere Web Client          | VMware vS...   | Started | Automatic       | Local System |

9. Go to the **Configuration** tab of the vSphere host that you plan to collect the syslog from. Then choose **Security Profile**. Make sure that the ports of the syslog are opened in the **Firewall** properties.

The screenshot shows the vSphere Configuration tab with the Security Profile selected. The Firewall properties are expanded, showing a table of incoming and outgoing connections. A red box highlights the 'Properties...' link for the Firewall, and a red arrow points down to the Firewall configuration table.

| Label                                                                    | Incoming Ports       | Outgoing Ports    | Protocols | Daemon  |
|--------------------------------------------------------------------------|----------------------|-------------------|-----------|---------|
| <input checked="" type="checkbox"/> vSphere High Availability Agent      | 8182                 | 8182              | TCP,UDP   | Running |
| <input checked="" type="checkbox"/> HBR                                  |                      | 31031,44046       | TCP       | N/A     |
| <input checked="" type="checkbox"/> gdbserver                            | 1000-9999,50000-5... |                   | TCP       | N/A     |
| <input checked="" type="checkbox"/> Fault Tolerance                      | 8100,8200,8300       | 80,8100,8200,8300 | TCP,UDP   | N/A     |
| <input checked="" type="checkbox"/> syslog                               |                      | 514,1514          | UDP,TCP   | N/A     |
| <input checked="" type="checkbox"/> VMware vCenter Agent                 |                      | 902               | UDP       | Running |
| <input checked="" type="checkbox"/> IKED                                 | 500                  | 500               | UDP       | N/A     |
| <input checked="" type="checkbox"/> VM serial port connected over net... | 23,1024-65535        | 0-65535           | TCP       | N/A     |
| <input checked="" type="checkbox"/> httpClient                           |                      | 80,443            | TCP       | N/A     |
| <input checked="" type="checkbox"/> DNS Client                           | 53                   | 53                | UDP,TCP   | N/A     |

10. Configure the path of syslog in the **Advanced Settings** of the vCenter Server.  
Input `tcp://<vCenter IP address>:514` in `Syslog.global.logDir`.



If there are some fibre channel's array errors that are logged to ESXi, you can find the detailed logs in `/var/log`. You can display all the logs containing SCSI in `/var/log` using the `"grep -s SCSI /var/log | more"` command. The following are some of the possible warning messages:

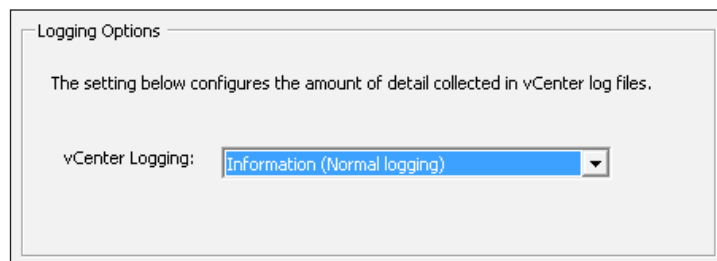
```

192.168.1.234 - PuTTY
(handle 40150, token 0x412e80190540): No connection
/var/log/vmkwarning.log:2014-12-30T10:58:11.892Z cpu5:36601) WARNING: VSCSI: 205:
(handle 40150, token 0x412e80101280): No connection
/var/log/vmkwarning.log:2014-12-30T10:58:11.892Z cpu5:36601) WARNING: VSCSI: 205:
(handle 40150, token 0x412e80101280): No connection
/var/log/vmkwarning.log:2014-12-30T10:58:11.892Z cpu5:36601) WARNING: VSCSI: 205:
(handle 40150, token 0x412e80101280): No connection
/var/log/vmkwarning.log:2014-12-30T10:58:11.892Z cpu5:36601) WARNING: VSCSI: 205:
(handle 40150, token 0x412e80101280): No connection
/var/log/vmkwarning.log:2014-12-30T10:58:11.892Z cpu5:36601) WARNING: VSCSI: 205:
(handle 40150, token 0x412e80101280): No connection
/var/log/vobd.log:2014-12-29T01:40:31.136Z: [netCorrelator] 28847959us: [vob.net
.firewall.config.changed] Firewall configuration has changed. Operation 'add' fo
r rule set iSCSI succeeded.
/var/log/vobd.log:2014-12-29T01:40:31.136Z: [netCorrelator] 28848056us: [esx.aud
it.net.firewall.config.changed] Firewall configuration has changed. Operation 'a
dd' for rule set iSCSI succeeded.
/var/log/vobd.log:2015-01-02T01:24:04.897Z: [netCorrelator] 28612078us: [vob.net
.firewall.config.changed] Firewall configuration has changed. Operation 'add' fo
r rule set iSCSI succeeded.
/var/log/vobd.log:2015-01-02T01:24:04.897Z: [netCorrelator] 28612174us: [esx.aud
it.net.firewall.config.changed] Firewall configuration has changed. Operation 'a
dd' for rule set iSCSI succeeded.
~ #

```

The default settings of vCenter Server logging will only record the normal message log. In order to troubleshoot, a detailed log collection method has to be enabled. By following the procedure given here, you can configure the **Verbose** setting in the logging options in the vCenter Server to collect the detailed log:

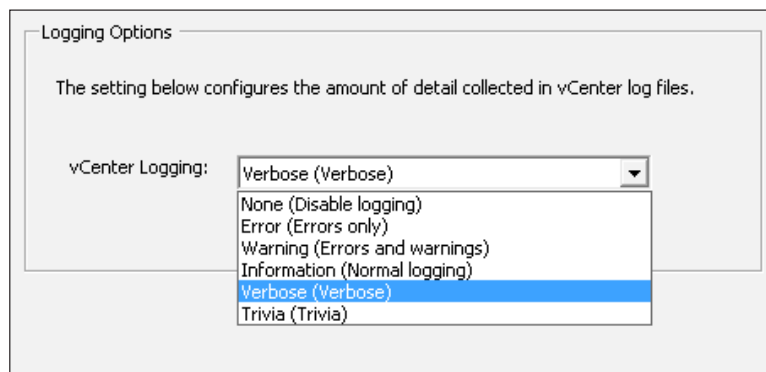
1. Go to the **vCenter Server Settings** and choose **Logging Options**. The default setting is **Information (Normal logging)**, as shown here:



2. Select **Verbose** in the **vCenter Logging** menu, and click on the **OK** button. The verbose mode setting displays detailed information, errors, warnings, and verbose log entries for troubleshooting.

The following table shows the options of the vCenter Server logging list:

| Option                               | Description                                                           |
|--------------------------------------|-----------------------------------------------------------------------|
| <b>None (Disable logging)</b>        | Disables logging                                                      |
| <b>Error (Errors only)</b>           | Displays only error log entries                                       |
| <b>Warning (Errors and warnings)</b> | Displays warning and error log entries                                |
| <b>Information (Normal logging)</b>  | Displays information, error, and warning log entries                  |
| <b>Verbose (Verbose)</b>             | Displays information, error, warning, and verbose log entries         |
| <b>Trivia (Trivia)</b>               | Displays information, error, warning, verbose, and trivia log entries |



## Summary

In this chapter, you learned what a storage component is, for example, the LUN name, identifier name, runtime name, and so on. We saw how to set up LUN Masking on the vSphere host using the `esxcli` commands, and set up the vSphere Syslog Collector to collect the ESXi host's log.

In the next chapter, we will see how to analyze the vSphere storage path, and identify performance issues with storage using `esxtop`.

# 7

## Analyzing vSphere Storage by CLI

In the previous chapter, you learned about storage components, such as the LUN name, identifier name, and runtime name; how to set up LUN masking in a vSphere host using `esxcli` commands; and setting up the vSphere Syslog Collector to collect the ESXi host's log. During configuration and installation, the vSphere administrator may collect some information (such as the WWN of the host bus adapter of the vSphere host) and verify the path policy of the vSphere host. Executing the command is more convenient than using the vSphere client.

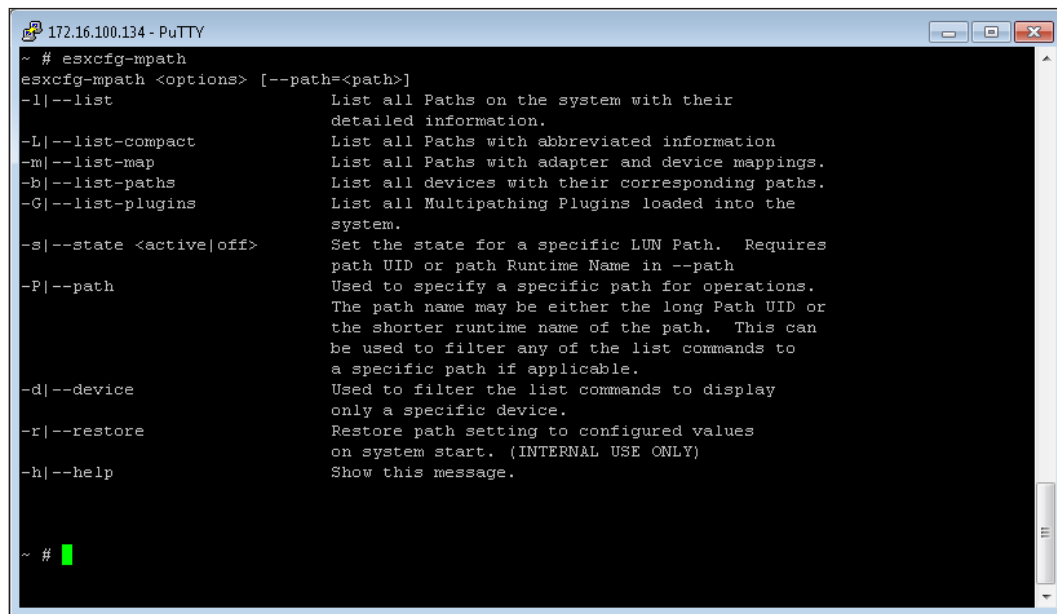
In this chapter, we will cover these topics:

- Analyzing PSA and multipathing by `esxcli`
- Applying VMFS volume copies resignaturing
- Troubleshooting VMware snapshots and VMFS resignaturing
- VMFS data store volume unmounting
- Identifying and tagging SSD devices



## Analyzing PSA and multipathing using esxcli

In *Chapter 5, Optimizing Storage*, we covered some examples of managing vSphere storage using esxcli. Now you will learn some more commands used to manage vSphere storage in this chapter. First, let's read about a useful command called esxcfg-mpath. It is used to list all paths or paths with the HBA adapter and device mappings, and so on. In the following screenshot, we can see all the options for the esxcfg-mpath command:



```
172.16.100.134 - PuTTY
~ # esxcli esxcfg-mpath
esxcli esxcfg-mpath <options> [--path=<path>]
-l|--list List all Paths on the system with their
 detailed information.
-L|--list-compact List all Paths with abbreviated information
-m|--list-map List all Paths with adapter and device mappings.
-b|--list-paths List all devices with their corresponding paths.
-G|--list-plugins List all Multipathing Plugins loaded into the
 system.
-s|--state <active|off> Set the state for a specific LUN Path. Requires
 path UID or path Runtime Name in --path
-P|--path Used to specify a specific path for operations.
 The path name may be either the long Path UID or
 the shorter runtime name of the path. This can
 be used to filter any of the list commands to
 a specific path if applicable.
-d|--device Used to filter the list commands to display
 only a specific device.
-r|--restore Restore path setting to configured values
 on system start. (INTERNAL USE ONLY)
-h|--help Show this message.

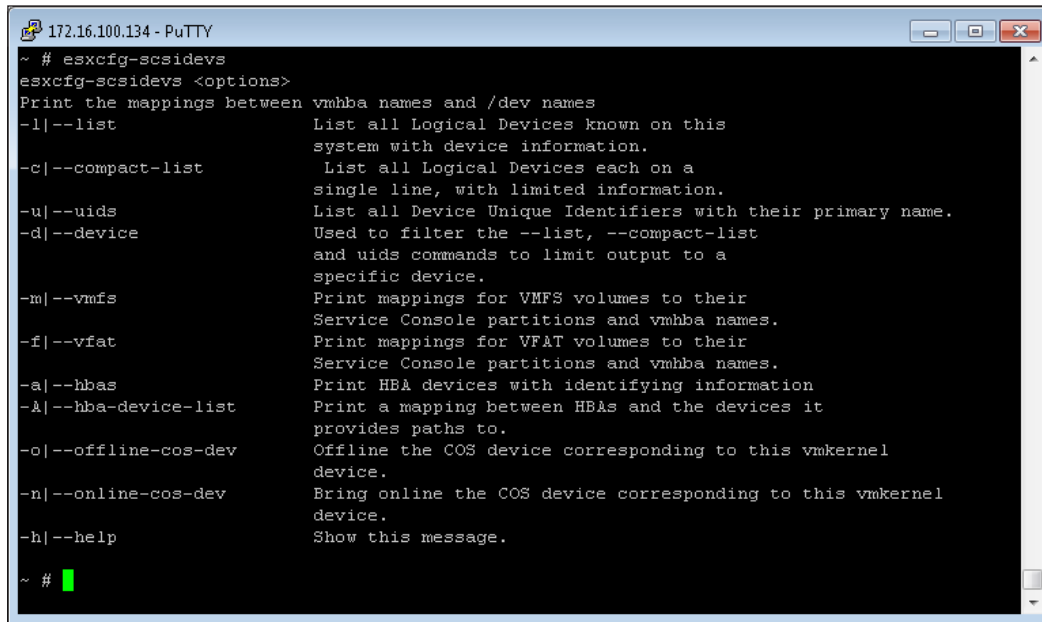
~ #
```

This table lists the description of options of the esxcli esxcfg-mpath command:

| Option | Description                                                             |
|--------|-------------------------------------------------------------------------|
| -l     | List all the paths on the system, with detailed information about them. |
| -L     | List all the paths with abbreviated information.                        |
| -m     | List all the paths with adapter and device mappings.                    |
| -b     | List all the devices, with their corresponding paths.                   |



Let's talk about another useful command, `esxcfg-scsidevs`. It is used to list all paths or paths with the HBA adapter, device mappings, and so on. Listed here are all the options for the `esxcfg-scsidevs` command:



```
172.16.100.134 - PuTTY
~ # esxcfg-scsidevs
esxcfg-scsidevs <options>
Print the mappings between vmhba names and /dev names
-l|--list List all Logical Devices known on this
 system with device information.
-c|--compact-list List all Logical Devices each on a
 single line, with limited information.
-u|--uids List all Device Unique Identifiers with their primary name.
-d|--device Used to filter the --list, --compact-list
 and uids commands to limit output to a
 specific device.
-m|--vmfs Print mappings for VMFS volumes to their
 Service Console partitions and vmhba names.
-f|--vfat Print mappings for VFAT volumes to their
 Service Console partitions and vmhba names.
-a|--hbas Print HBA devices with identifying information
-A|--hba-device-list Print a mapping between HBAs and the devices it
 provides paths to.
-o|--offline-cos-dev Offline the COS device corresponding to this vmkernel
 device.
-n|--online-cos-dev Bring online the COS device corresponding to this vmkernel
 device.
-h|--help Show this message.

~ #
```

The following table lists the options description for the `esxcfg-scsidevs` command:

| Option | Description                                                                                                                                           |
|--------|-------------------------------------------------------------------------------------------------------------------------------------------------------|
| -l     | List all logical devices known on this system, with device information                                                                                |
| -m     | Print the mappings for VMFS volumes to their service console partitions and vmhba names                                                               |
| -d     | This is used to filter the <code>-list</code> , <code>--compact-list</code> , and <code>uids</code> commands to limit the output to a specific device |
| -u     | List all device unique identifiers with their primary name                                                                                            |
| -f     | Print the mappings for VFAT volumes to their service console partitions and vmhba names                                                               |
| -a     | Print HBA devices with their identifying information                                                                                                  |
| -c     | List all logical devices, each on a single line, with limited information                                                                             |
| -A     | Print a mapping between the HBAs and the devices it provides paths to vSphere host                                                                    |

If you want to list all the LUNs on a vSphere host, that is, the device mount point and LUN UID, you can execute the `esxcfg-scsidevs` command with the `-c` option.

If you want to list all the mappings of VMFS and HBA names on the vSphere host—that is, the device mount point and VMFS volume name—you can execute the `esxcfg-scsidevs -m` command. The result we get after execution is shown here:

If you want to list the current SATP plugins with their information about the default PSP, you can execute the `esxcli` command with the `storage nmp satp list` namespace. The result that we get after execution is shown here.

```

~ # esxccli storage nmp satp list
Name Default PSP Description

VMW SATP_ALUA VMW_PSP_MRU Supports non-specific arrays that use the ALUA protocol
VMW SATP_MSA VMW_PSP_MRU Placeholder (plugin not loaded)
VMW SATP_DEFAULT_AP VMW_PSP_MRU Placeholder (plugin not loaded)
VMW SATP_SVC VMW_PSP_FIXED Placeholder (plugin not loaded)
VMW SATP_EOL VMW_PSP_FIXED Placeholder (plugin not loaded)
VMW SATP_INV VMW_PSP_FIXED Placeholder (plugin not loaded)
VMW SATP_EVA VMW_PSP_FIXED Placeholder (plugin not loaded)
VMW SATP_ALUA_CX VMW_PSP_RR Placeholder (plugin not loaded)
VMW SATP_SYMM VMW_PSP_RR Placeholder (plugin not loaded)
VMW SATP_CX VMW_PSP_MRU Placeholder (plugin not loaded)
VMW SATP_LSI VMW_PSP_MRU Placeholder (plugin not loaded)
VMW SATP_DEFAULT_AA VMW_PSP_FIXED Supports non-specific active/active arrays
VMW SATP_LOCAL VMW_PSP_FIXED Supports direct attached devices
~ #

```

This table lists the description of each SATP plugin in the preceding output:

| Name                       | Description                                                                                                                                                                                                                                                    |
|----------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>VMW_SATP_ALUA</b>       | <b>VMW_SATP_ALUA</b> is assigned to a specific storage device, but the device is not ALUA aware. No claim rule match occurs for this device. The device is claimed by the default SATP based on its transport type. Its plugin is not loaded into this system. |
| <b>VMW_SATP_MSA</b>        | Supports HP MSA arrays. Its plugin is not loaded into this system.                                                                                                                                                                                             |
| <b>VMW_SATP_DEFAULT_AP</b> | Supports non-specific active/passive arrays. Its plugin is not loaded into this system.                                                                                                                                                                        |
| <b>VMW_SATP_SVC</b>        | Supports IBM SVC arrays. Its plugin is not loaded into this system.                                                                                                                                                                                            |
| <b>VMW_SATP_EQL</b>        | Supports Dell EquaLogic arrays. Its plugin is not loaded into this system.                                                                                                                                                                                     |
| <b>VMW_SATP_INV</b>        | Supports the EMC Invista array family. Its plugin is not loaded into this system.                                                                                                                                                                              |
| <b>VMW_SATP_EVA</b>        | Supports HP EVA arrays. Its plugin is not loaded into this system.                                                                                                                                                                                             |
| <b>VMW_SATP_ALUA_CX</b>    | Supports EMC CX that do not use the ALUA protocol. Its plugin is not loaded into this system.                                                                                                                                                                  |
| <b>VMW_SATP_SYMM</b>       | Supports the EMC Symmetrix array family. Its plugin is not loaded into this system.                                                                                                                                                                            |
| <b>VMW_SATP_CX</b>         | Supports EMC CX that do not use the ALUA protocol. Its plugin is not loaded into this system.                                                                                                                                                                  |
| <b>VMW_SATP_LSI</b>        | Supports LSI arrays. Its plugin is not loaded into this system.                                                                                                                                                                                                |
| <b>VMW_SATP_DEFAULT_AA</b> | Supports non-specific active/active arrays. Its plugin is loaded into this system.                                                                                                                                                                             |
| <b>VMW_SATP_LOCAL</b>      | Supports directly attached devices. Its plugin is loaded into this system.                                                                                                                                                                                     |

If you want to change the default path selection policy for any new storage for a storage array type plugin, you can execute the `esxcli storage nmp satp set -P <Default PSP> -s <SATP Name>` command. The result we get after execution is shown here:

```
~ # esxcli storage nmp satp list
Name Default PSP Description

VMW_SATP_ALUA VMW_PSP_MRU Supports non-specific arrays that use the ALUA protocol
VMW_SATP_MSA VMW_PSP_MRU Placeholder (plugin not loaded)
VMW_SATP_DEFAULT_AP VMW_PSP_MRU Placeholder (plugin not loaded)
VMW_SATP_SVC VMW_PSP_FIXED Placeholder (plugin not loaded)
VMW_SATP_EQL VMW_PSP_FIXED Placeholder (plugin not loaded)
VMW_SATP_INV VMW_PSP_FIXED Placeholder (plugin not loaded)
VMW_SATP_EVA VMW_PSP_FIXED Placeholder (plugin not loaded)
VMW_SATP_ALUA_CX VMW_PSP_RR Placeholder (plugin not loaded)
VMW_SATP_SYMM VMW_PSP_RR Placeholder (plugin not loaded)
VMW_SATP_CX VMW_PSP_MRU Placeholder (plugin not loaded)
VMW_SATP_LSI VMW_PSP_MRU Placeholder (plugin not loaded)
VMW_SATP_DEFAULT_AA VMW_PSP_FIXED Supports non-specific active/active arrays
VMW_SATP_LOCAL VMW_PSP_FIXED Supports direct attached devices
~ # esxcli storage nmp satp set -P VMW_PSP_MRU -s VMW_SATP_ALUA
```

In this example, changes the default PSP (**VMW\_PSP\_MRU**) associated with a given SATP (**VMW\_SATP\_ALUA**) to **VMW\_PSP\_PP**.

The following table lists the description of path selection policies (PSP):

| Name                 | Description          |
|----------------------|----------------------|
| <b>VMW_PSP_MRU</b>   | For MRUmode          |
| <b>VMW_PSP_FIXED</b> | For Fixed mode       |
| <b>VMW_PSP_RR</b>    | For Round Robin mode |

```
~ # esxcli storage nmp satp list
Name Default PSP Description

VMW_SATP_ALUA VMW_PSP_MRU Supports non-specific arrays that use the ALUA protocol
VMW_SATP_MSA VMW_PSP_MRU Placeholder (plugin not loaded)
VMW_SATP_DEFAULT_AP VMW_PSP_MRU Placeholder (plugin not loaded)
VMW_SATP_SVC VMW_PSP_FIXED Placeholder (plugin not loaded)
VMW_SATP_EQL VMW_PSP_FIXED Placeholder (plugin not loaded)
VMW_SATP_INV VMW_PSP_FIXED Placeholder (plugin not loaded)
VMW_SATP_EVA VMW_PSP_FIXED Placeholder (plugin not loaded)
VMW_SATP_ALUA_CX VMW_PSP_RR Placeholder (plugin not loaded)
VMW_SATP_SYMM VMW_PSP_RR Placeholder (plugin not loaded)
VMW_SATP_CX VMW_PSP_MRU Placeholder (plugin not loaded)
VMW_SATP_LSI VMW_PSP_MRU Placeholder (plugin not loaded)
VMW_SATP_DEFAULT_AA VMW_PSP_FIXED Supports non-specific active/active arrays
VMW_SATP_LOCAL VMW_PSP_FIXED Supports direct attached devices
~ # esxcli storage nmp satp set -P VMW_PSP_MRU -s VMW_SATP_ALUA
```



Preceding vSphere host has only enabled three plugins, that is, **VMW\_SATP\_ALUA**, **VMW\_SATP\_DEFAULT\_AA**, and **VMW\_SATP\_LOCAL**.

It is required to reboot the ESXi/ESX host to apply the changes after changing the policy.

## Applying VMFS volume copies resignaturing

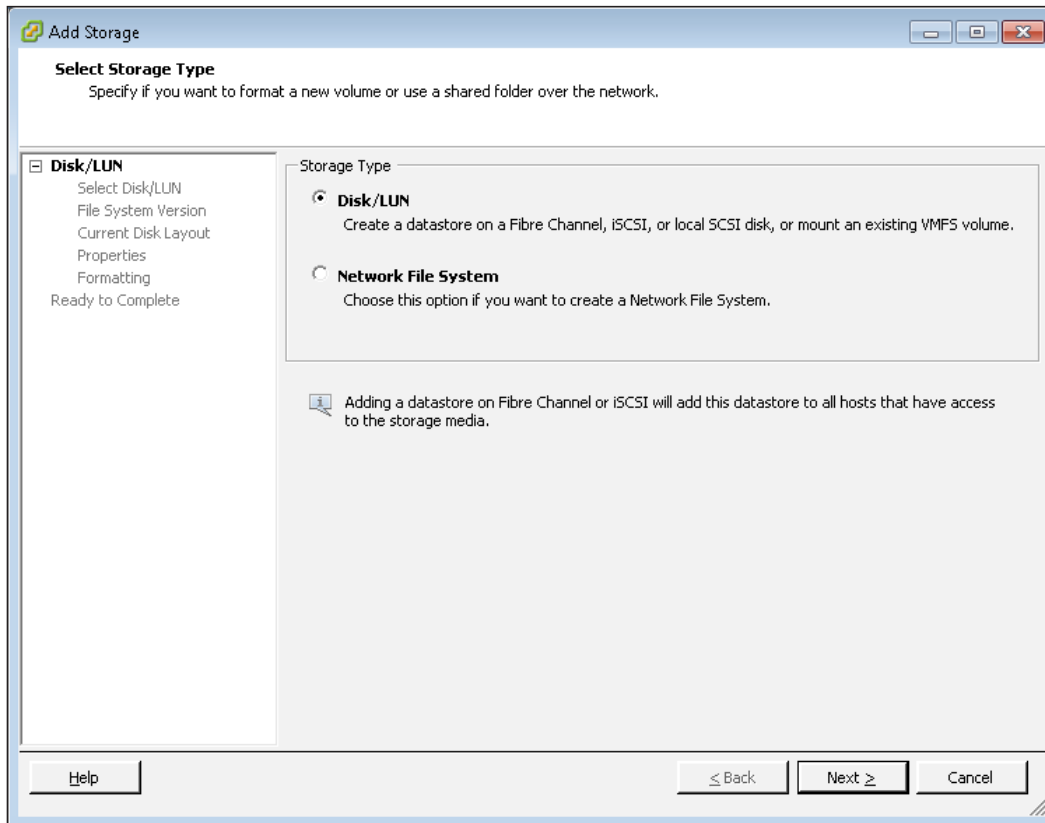
When you create the data store in a vSphere host, the vSphere host mounts the data store with an existing signature or assigns a new signature. Each VMFS data store filesystem has a unique UUID. In this situation, you can specify VMFS by this option. During a disaster recovery plan, for example, VMware Site Recovery Manager, you need to mount the VMFS data store copy with its original UUID and the existing signature in the vSphere host on a secondary site. When you mount the VMFS data store copy at that site, the original VMFS data store is required to be offline at the primary site. Here is the procedure of specifying the data store with the existing signature:

1. Log in to vCenter Server using the vSphere Client. Click on the **Configuration** tab and then on **Storage** under the **Hardware** panel. The result we get after execution is shown here:

The screenshot shows the vSphere Client interface with the **Configuration** tab selected. Under the **Hardware** panel, the **Storage** section is expanded. The **Datastores** table is displayed, showing the following data:

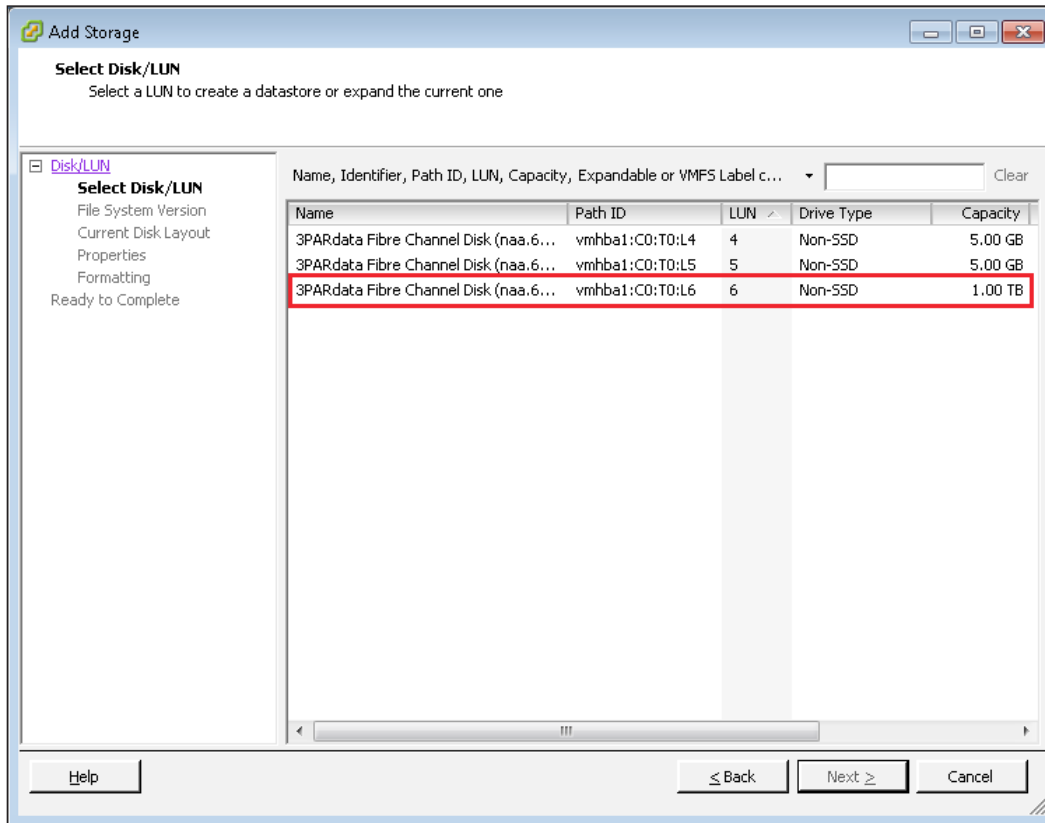
| Identification     | Status | Device              | Drive Type | Capacity   | Free       | Type  | Last Update           | Alarm Act |
|--------------------|--------|---------------------|------------|------------|------------|-------|-----------------------|-----------|
| 3PAR_DS1           | Normal | 3PARdata Fibre C... | Non-SSD    | 1,023.75 G | 1,022.80 G | VMFS5 | 2/19/2015 11:44:31 AM | Enabled   |
| 3PAR_DS2           | Normal | 3PARdata Fibre C... | Non-SSD    | 499.75 GB  | 498.80 GB  | VMFS5 | 2/19/2015 11:44:31 AM | Enabled   |
| 3PAR_DS3           | Normal | 3PARdata Fibre C... | Non-SSD    | 3.50 TB    | 3.50 TB    | VMFS5 | 2/19/2015 11:44:31 AM | Enabled   |
| ESXi5.5A_OS        | Normal | 3PARdata Fibre C... | Non-SSD    | 2.50 GB    | 1.92 GB    | VMFS5 | 2/19/2015 11:44:31 AM | Enabled   |
| Images (read only) | Normal | 172.16.100.38/1...  | Unknown    | 3.58 TB    | 2.41 TB    | NFS   | 2/19/2015 11:44:31 AM | Enabled   |
| UCS03_Local        | Normal | Local WD Disk (n... | Non-SSD    | 279.25 GB  | 271.11 GB  | VMFS5 | 2/19/2015 11:44:31 AM | Enabled   |

2. Select the **Add Storage...** link and you will get a **Add Storage** screen.
3. Then select **Disk/LUN** under **Storage Type** and click on **Next**, as shown in this screenshot:

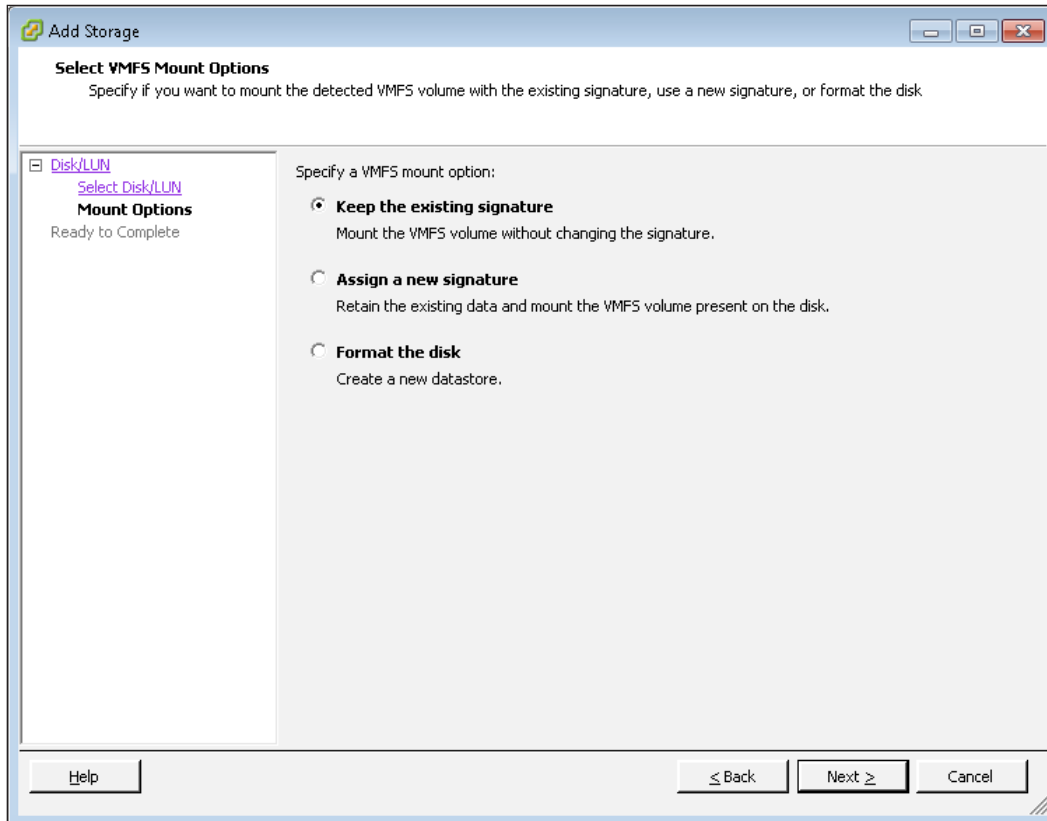





4. Select the LUN that has the data store name displayed in the LUN list, and then click on **Next**, as shown in the following screenshot. In this example, select LUN 6 (**vmhba1:C0:T0:L6, 1.00 TB**) which is the data store copy.



- Under **Mount Options**, you can select **Keep the existing signature**. Click on **Next**. Here is a screenshot showing this, for your reference:




 You cannot select this option if the VMFS volume is not a data store copy.

- On the complete page, you can check out the data store configuration information and click on **Finish**.

When you create the VMFS volume by data store resignaturing, ESXi assigns a new UUID and new label to that data store copy. This data store is different from the original data store. This operation is the same as the disk format on a Windows platform.

## Troubleshooting VMware snapshots and VMFS resignaturing

You can execute the VMFS resignaturing operation using the `esxcfg-volume` command. Listed here are all the options for this command:

```
~ # esxcfg-volume
esxcfg-volume <options>
-l|--list List all volumes which have been
 detected as snapshots/replicas.
-m|--mount <VMFS UUID|label> Mount a snapshot/replica volume, if
 its original copy is not online.
-u|--umount <VMFS UUID|label> Umount a snapshot/replica volume.
-r|--resignature <VMFS UUID|label> Resignature a snapshot/replica volume.
-M|--persistent-mount <VMFS UUID|label> Mount a snapshot/replica volume
 persistently, if its original copy is
 not online.
-U|--upgrade <VMFS UUID|label> Upgrade a VMFS3 volume to VMFS5.
-h|--help Show this message.
~ #
```

You can list all the VMFS volumes that have been detected as snapshots or replicas using the `esxcfg-volume` command with the `-l` option. You can find one **VMFS UUID**, **54e2a884-74597529-015b-2c44fd8309d4**, that has been created as a snapshot, as shown in the following screenshot:

```
~ # esxcfg-volume
esxcfg-volume <options>
-l|--list List all volumes which have been
 detected as snapshots/replicas.
-m|--mount <VMFS UUID|label> Mount a snapshot/replica volume, if
 its original copy is not online.
-u|--umount <VMFS UUID|label> Umount a snapshot/replica volume.
-r|--resignature <VMFS UUID|label> Resignature a snapshot/replica volume.
-M|--persistent-mount <VMFS UUID|label> Mount a snapshot/replica volume
 persistently, if its original copy is
 not online.
-U|--upgrade <VMFS UUID|label> Upgrade a VMFS3 volume to VMFS5.
-h|--help Show this message.
~ # esxcfg-volume -l
VMFS UUID/label: 54e2a884-74597529-015b-2c44fd8309d4/3PAR_DS1
Can mount: No (the original volume is still online)
Can resignature: Yes
Extent name: naa.60002ac000000000000000001500008e5b:1 range: 0 - 1048319 (MB)
~ #
```



The preceding command is used to list the VMFS volumes that have been detected as snapshots or replicas in vSphere 4.x.

You need to list the VMFS volume snapshots using the `esxcli storage vmfs snapshot list` command, if the vSphere host is version 5 or above. The following result is for your reference:

```
~ # esxcli storage vmfs snapshot list
54e2a884-74597529-015b-2c44fd8309d4
 Volume Name: 3PAR_DS1
 VMFS UUID: 54e2a884-74597529-015b-2c44fd8309d4
 Can mount: false
 Reason for un-mountability: the original volume is still online
 Can resignature: true
 Reason for non-resignaturability:
 Unresolved Extent Count: 1
~ #
```

If you plan to mount the VMFS volume snapshot, you can execute the operation using the `esxcli storage vmfs snapshot mount` command. Listed here are all the options for `esxcli storage vmfs snapshot mount`:

```
~ # esxcli storage vmfs snapshot mount --help
Usage: esxcli storage vmfs snapshot mount [cmd options]

Description:
 mount Mount a snapshot/replica of a VMFS volume.

Cmd options:
 -n|--no-persist Mount the volume non-persistently; the volume will not be automounted
 after a restart.
 -l|--volume-label=<str>
 The VMFS volume label of the snapshot to mount.
 -u|--volume-uuid=<str>
 The VMFS volume uuid of the snapshot to mount.
~ #
```

According to the preceding result, to mount the VMDS volume copy that will not be automated after reboot, we use this command:

```
esxcli storage vmfs snapshot mount -n -u "54e2a884-74597529-015b-2c44fd8309d4"
```

```
~ # esxcli storage vmfs snapshot mount --help
Usage: esxcli storage vmfs snapshot mount [cmd options]

Description:
 mount Mount a snapshot/replica of a VMFS volume.

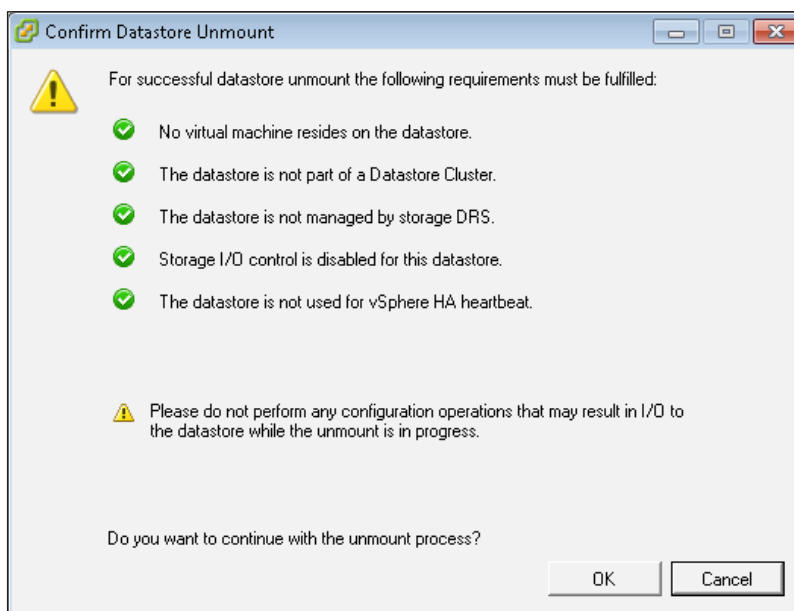
Cmd options:
 -n|--no-persist Mount the volume non-persistently; the volume will not be automounted
 after a restart.
 -l|--volume-label=<str>
 The VMFS volume label of the snapshot to mount.
 -u|--volume-uuid=<str>
 The VMFS volume uuid of the snapshot to mount.
~ # esxcli storage vmfs snapshot mount -n -u "54e2a884-74597529-015b-2c44fd8309d4"
```

## VMFS DataStore volume unmounting

When you unmount a data store in one vSphere host, make sure that the following prerequisites are met:

- No virtual machines are residing on the data store
- The data store is not managed by the vSphere storage DRS
- Storage I/O control is disabled for this data store
- The data store is not used for vSphere HA heartbeating
- The data store is not part of a data store cluster

The following result is the requirement checking during a data store unmount:



## Identifying and tagging SSD devices

If you plan to use an SSD device in your vSphere host, ensure that the device is tagged as SSD. You can identify it in the vSphere client or using commands. The information about the drive type is displayed as **SSD** in the vSphere client if the data store is SSD. Both of the following vSphere hosts have different ESX data store configurations for your reference.

This vSphere host has no SSD data store. It has four FC drive data stores, one local drive data store, and one NFS data store.

| Identification     | Status | Device              | Drive Type | Capacity   | Free       | Type  | Last Update          | Alarm Actions | Storage I/O Control | Hardware Acceleration |
|--------------------|--------|---------------------|------------|------------|------------|-------|----------------------|---------------|---------------------|-----------------------|
| 3PAR_DS1           | Normal | 3PARdata Fibre C... | Non-SSD    | 1,023.75 G | 1,022.80 G | VMFS5 | 2/19/2015 1:44:34 PM | Enabled       | Disabled            | Supported             |
| 3PAR_DS2           | Normal | 3PARdata Fibre C... | Non-SSD    | 499.75 GB  | 498.80 GB  | VMFS5 | 2/19/2015 1:44:34 PM | Enabled       | Disabled            | Supported             |
| 3PAR_DS3           | Normal | 3PARdata Fibre C... | Non-SSD    | 3.50 TB    | 3.50 TB    | VMFS5 | 2/19/2015 1:44:34 PM | Enabled       | Disabled            | Supported             |
| ESXi5_SA_OS        | Normal | 3PARdata Fibre C... | Non-SSD    | 2.50 GB    | 1.92 GB    | VMFS5 | 2/19/2015 1:44:34 PM | Enabled       | Disabled            | Supported             |
| Images (read only) | Normal | 172.16.100.38:/L... | Unknown    | 3.58 TB    | 2.41 TB    | NFS   | 2/19/2015 1:44:34 PM | Enabled       | Disabled            | Not supported         |
| UCS03_Local        | Normal | Local WD Disk (n... | Non-SSD    | 279.25 GB  | 271.11 GB  | VMFS5 | 2/19/2015 1:44:34 PM | Enabled       | Disabled            | Unknown               |

The following vSphere host has one SSD data store, which is displayed as SSD under **Drive Type**. The other data stores are FC drives.

| Datastores     |        |                       |            |           |           |       |                       |               |                     |                       |  |
|----------------|--------|-----------------------|------------|-----------|-----------|-------|-----------------------|---------------|---------------------|-----------------------|--|
| Identification | Status | Device                | Drive Type | Capacity  | Free      | Type  | Last Update           | Alarm Actions | Storage I/O Control | Hardware Acceleration |  |
| CX4-120-D51    | Normal | DGC Fibre Channel...  | Non-SSD    | 499.75 GB | 190.99 GB | VMFS5 | 6/20/2015 6:10:11 PM  | Enabled       | Disabled            | Supported             |  |
| CX4-120-D52    | Normal | DGC Fibre Channel...  | Non-SSD    | 799.75 GB | 234.31 GB | VMFS5 | 6/18/2015 10:40:09 PM | Enabled       | Disabled            | Supported             |  |
| CX4-120-D53    | Normal | DGC Fibre Channel...  | Non-SSD    | 299.75 GB | 295.65 GB | VMFS5 | 6/11/2015 10:56:37 AM | Enabled       | Disabled            | Supported             |  |
| CX4-120-D54    | Normal | DGC Fibre Channel...  | Non-SSD    | 299.75 GB | 295.59 GB | VMFS5 | 6/11/2015 6:13:10 PM  | Enabled       | Disabled            | Supported             |  |
| ESXi6a_local   | Normal | HP Serial Attached... | Non-SSD    | 271.75 GB | 270.80 GB | VMFS5 | 5/27/2015 2:39:26 PM  | Enabled       | Disabled            | Unknown               |  |
| SSD_D51        | Normal | DGC Fibre Channel...  | SSD        | 91.50 GB  | 71.16 GB  | VMFS5 | 6/10/2015 11:13:08 PM | Enabled       | Disabled            | Supported             |  |

You can also list all storage devices using the following command. The value of **Is SSD** is **true**. The result shown is for your reference:

```
esxcli storage core device list
```

```
~ # esxcli storage core device list
naa.50000c0f02ecba5c
 Display Name: Local WD Disk (naa.50000c0f02ecba5c)
 Has Settable Display Name: true
 Size: 286102
 Device Type: Direct-Access
 Multipath Plugin: NMP
 Devfs Path: /vmfs/devices/disks/naa.50000c0f02ecba5c
 Vendor: WD
 Model: WD3001BKHG-33D1
 Revision: CS05
 SCSI Level: 6
 Is Pseudo: false
 Status: on
 Is RDM Capable: false
 Is Local: true
 Is Removable: false
 Is SSD: false
 Is Offline: false
 Is Perennially Reserved: false
 Queue Full Sample Size: 0
 Queue Full Threshold: 0
 Thin Provisioning Status: unknown
 Attached Filters:
 VAAI Status: unknown
 Other UIDs: vml.020000000050000c0f02ecba5c574433303031
 Is Local SAS Device: false
 Is USB: false
 Is Boot USB Device: false
 No of outstanding IOs with competing worlds: 32
```

If the vSphere host cannot automatically identify the storage device as SSD, you need to tag the device as SSD using SATP claim rules, that is, the LUN masking procedure. For more details of this procedure, you can refer to *Chapter 6, vSphere Storage Configuration Settings*.

Here are the instructions for tagging the SSD device:

1. Firstly, identify the device to be tagged and its SATP. You can execute the `esxcli storage nmp device list` option. The following screenshot shows the result, which is given for your reference:

```
~ # esxcli storage nmp device list
naa.50000c0f02ecba5c
 Device Display Name: Local WD Disk (naa.50000c0f02ecba5c)
 Storage Array Type: VMW_SATP_LOCAL
 Storage Array Type Device Config: SATP VMW_SATP_LOCAL does not support device configuration.
 Path Selection Policy: VMW_PSP_FIXED
 Path Selection Policy Device Config: (preferred=vmhba0:CO:T2:L0;current=vmhba0:CO:T2:L0)
 Path Selection Policy Device Custom Config:
 Working Paths: vmhba0:CO:T2:L0
 Is Local SAS Device: false
 Is USB: false
 Is Boot USB Device: false

naa.60002ac00000000000000000e00008e5b
 Device Display Name: 3PARdata Fibre Channel Disk (naa.60002ac00000000000000000e00008e5b)
 Storage Array Type: VMW_SATP_ALUA
 Storage Array Type Device Config: (implicit_support=on;explicit_support=off; explicit_allow=on;alua_
followover=on;(TPG_id=256,TPG_state=AO))
 Path Selection Policy: VMW_PSP_MRU
 Path Selection Policy Device Config: Current Path=vmhba1:CO:T0:L5
 Path Selection Policy Device Custom Config:
 Working Paths: vmhba1:CO:T0:L5
 Is Local SAS Device: false
 Is USB: false
 Is Boot USB Device: false

naa.60002ac00000000000000000900008e5b
 Device Display Name: 3PARdata Fibre Channel Disk (naa.60002ac00000000000000000900008e5b)
 Storage Array Type: VMW_SATP_ALUA
 Storage Array Type Device Config: (implicit_support=on;explicit_support=off; explicit_allow=on;alua_
followover=on;(TPG_id=256,TPG_state=AO))
```

2. Mark the SATP associated with the device.
3. Add a PSA claim rule to the device as SSD. The following command is given for your reference:

```
esxcli storage nmp satp rule add -s SATP --device <<device_
name>> --option=enable_ssd
```

4. Unclaim the device:

```
esxcli storage core claiming unclaim --type device --device
<<device_name>>
```

5. Reclaim the device by running the following commands:

```
esxcli storage core claimrule load
esxcli storage core claimrule run
```

6. Check whether the devices are tagged as SSD:

```
esxcli storage core device list --device <device_name>
```



The value of **Is SSD** for a listed device is **true** if that device is tagged as SSD.

```
[root@esxi5a:~] esxcli storage core device list --device naa.60060160085122009a883c81dbe4e411
naa.60060160085122009a883c81dbe4e411
 Display Name: DGC Fibre Channel Disk (naa.60060160085122009a883c81dbe4e411)
 Has Settable Display Name: true
 Size: 93894
 Device Type: Direct-Access
 Multipath Plugin: NMP
 Devfs Path: /vmfs/devices/disks/naa.60060160085122009a883c81dbe4e411
 Vendor: DGC
 Model: RAID 10
 Revision: 0430
 SCSI Level: 4
 Is Pseudo: false
 Status: on
 Is RDM Capable: true
 Is Local: false
 Is Removable: false
 Is SSD: true
 Is VVOL PE: false
 Is Offline: false
 Is Perennially Reserved: false
 Queue Full Sample Size: 0
 Queue Full Threshold: 0
 Thin Provisioning Status: unknown
```



If the SSD device that you want to tag is shared among multiple vSphere hosts, make sure that you tag the device from all the hosts that share that device.

## Summary

In this chapter, you learned about more commands that are used to manage vSphere storage. You also read about the difference between existing VMFS signatures and resignaturing, and learned how to collect information about VMware snapshots and VMFS resignaturing using the `esxcli` command. Then we covered the prerequisite operation to be performed before data store unmounting in the vSphere host, and identifying and tagging SSD devices in a vSphere client.

In the next chapter, we will see how to troubleshoot vSphere FC storage.

# 8

## Troubleshooting vSphere FC Storage

In the previous chapter, you learned how to troubleshoot storage performance problems using vSphere commands. Now we will go through some troubleshooting procedures to troubleshoot the Fibre Channel storage, and list some examples focusing on FC storage troubleshooting.

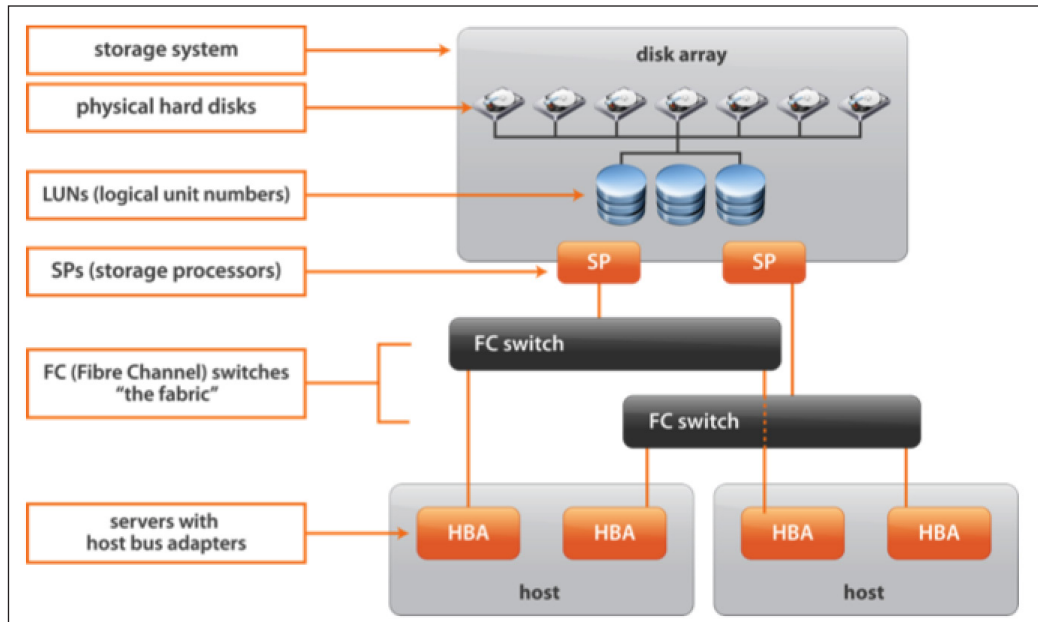
In this chapter, you will learn about the following topics:

- vSphere Fibre Channel storage components
- vSphere Fibre Channel storage troubleshooting examples

### The vSphere Fibre Channel storage component

In the previous chapter, we discussed storage troubleshooting by different types of common tools in detail. Now we will focus on Fibre Channel storage. So, let's learn more about the FC storage architecture. A Fibre Channel storage area network is a specialized high-speed network that connects computer systems or host servers to high-performance storage subsystems. The SAN's components include host bus adapters (HBAs) in the host servers, switches that help route storage traffic, cables, storage processors (SPs), and storage disk arrays. A SAN topology with at least one switch present in the network forms a SAN fabric. To transfer traffic from host servers to shared storage, the SAN uses the FC protocol, which packages SCSI commands into Fibre Channel frames.

In the vSphere environment, the FC SAN – as shown in the following diagram – includes virtual machines, ISO images, and templates in a data store that supports 16 Gbps throughput:

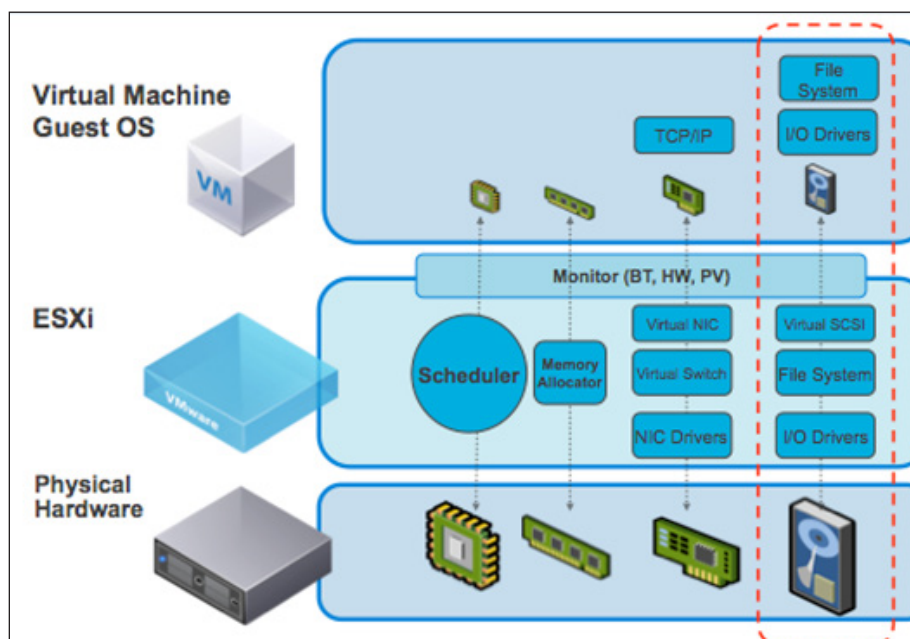


## A vSphere Fibre Channel storage troubleshooting example

In the previous chapter, you learned about the Fibre Channel storage architecture. Now you will learn how to troubleshoot problems with FC SAN. In FC SAN, general problems occur in storage throughput. The general problems that can occur include storage throughput and latency in read and write operations. Storage stack, follow the stack to troubleshoot the latency, and so on.

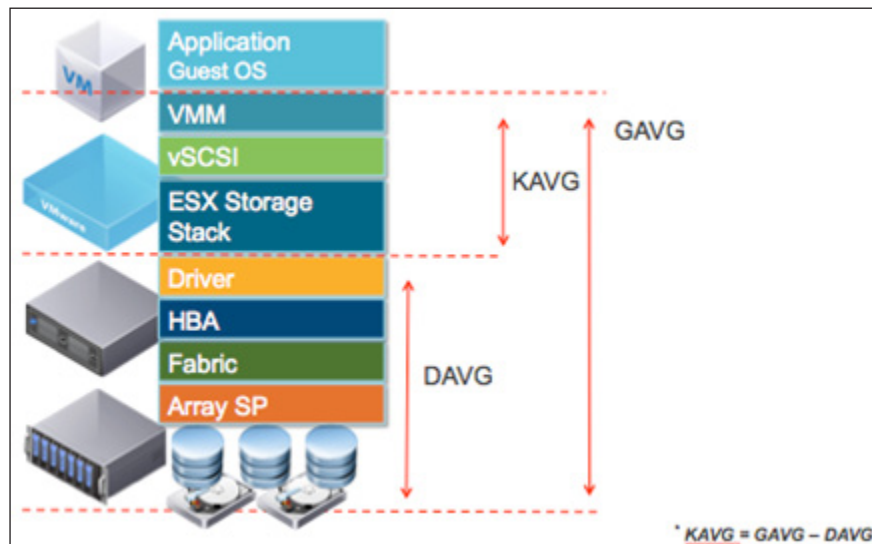
The ESX architecture has lots of components. Here, we will see the physical components at the bottom: CPU, memory, networking, and storage devices. Then, on the upper layer, we have the ESXi VMkernel software layer, and in it, we have various components, such as the scheduler that schedules the VMs on the physical CPUs, and the memory allocator that divides the physical memory and allocates it to the VMs. We also have various drivers for connecting to physical devices, such as networking and storage drivers. The most important part of the VMkernel is the monitor, which sits right between the VM and the VMkernel.

The monitor presents the physical devices as virtual devices to the VM. Then, inside the guest operating system, we have the virtual devices that are presented by the ESX monitor. After that, there are various drivers, filesystems, and so on that connect and operate with these virtualized devices, in the physical hardware layer as shown in the following diagram:



In a vSphere environment, the storage problem of the top issue is generally latency, so you need to learn what layers there are in the storage stack and where latency can build up. At the topmost layer is the application running in the guest operating system. That is ultimately the place where we care about latency the most. This is the total amount of latency that the application sees, and it includes the latencies of the total storage stack including the guest OS, VMkernel virtualization layers, and physical hardware.

Okay, now **Guest Average Latency (GAVG)** is made up of two major components: **Kernel Average Latency (KAVG)** and DAVG. DAVG basically means how much time is spent on the device from the driver HBA and storage array. KAVG means how much time is spent in the ESXi kernel. From ESXi, we see three main latencies and these are reported in esxtop and vCenter.



So how do we see three metrics? Well, they are visible in vCenter. It is recommended that you use vCenter as your first stop to debugging any problem of FC storage in your virtualization environment. vCenter provides nice alarms that will trigger if the device latency is too high or if a VM is using a large amount of bandwidth for the device. There are also charts that we can use to analyze the values and see historical trends. The only thing to be aware of is that vCenter uses a 20-second sample period, so it might miss some intermittent performance issues, but it is a good first start. Also, the metrics are set to the static threshold, which may not always indicate an actual performance problem. We need to analyze historical data.

VMware also has vRealize operations manager that provide a more robust total view of the health of your system, and has the intelligent engine, which relies on dynamic thresholds to analyze the data and identifies changed behaviors. So, the FC storage, if it has any problems, will show in the GUI in vRealize operations manager. You can download the production from the VMware website.



## vRealize displays the WorkLoad Badge Metrics

Now let's learn how to use the vCenter for troubleshooting an FC storage problem.

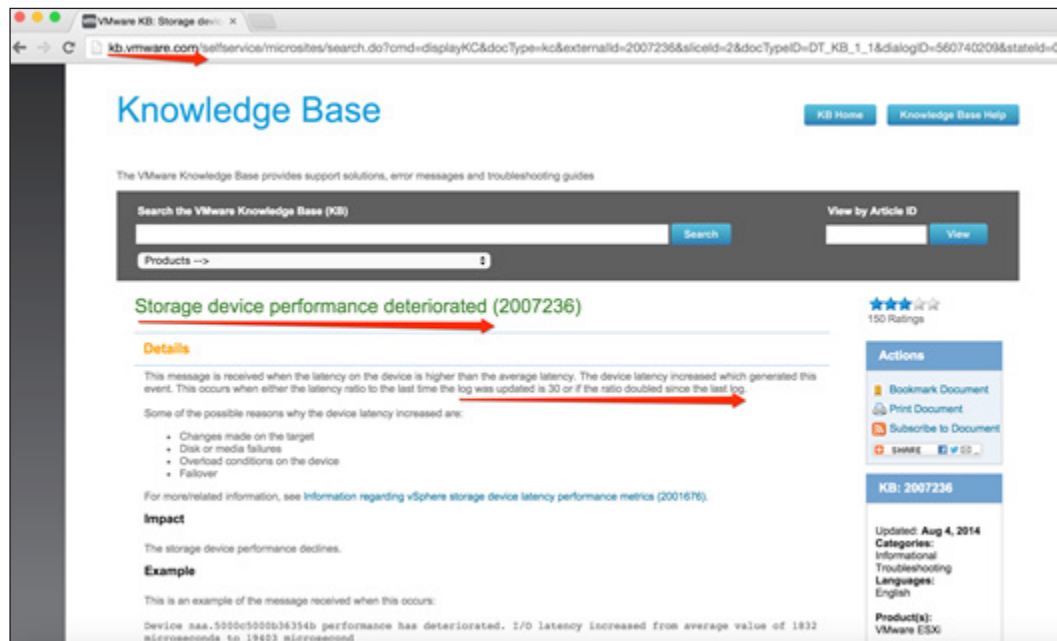
How can IT admin know that the FC storage has a problem in the virtualization environment? When do they? Let's cover that here.

As an example, the IT admin will log in to vCenter everyday to check whether any alarm has been triggered, and if they find a trigger, they will look for it and find out what problem has occurred.

First, if we suspect an FC storage problem, we have to log in to vCenter and check the events log, as shown here:

| Description                                                                                                                                                     | Type    | Date Time           | Task | Target | User   |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|---------------------|------|--------|--------|
| Device naa.60e9800042473274752b445648334734 performance has deteriorated. I/O latency increased from average value of 1897 microseconds to 236720 microseconds. | warning | 5/9/2014 2:08:24 PM |      |        |        |
| Device naa.60e9800042473274752b445648334734 performance has deteriorated. I/O latency increased from average value of 1897 microseconds to 111191 microseconds. | warning | 5/9/2014 2:08:17 PM |      |        |        |
| Device naa.60e9800042473274752b445648334734 performance has deteriorated. I/O latency increased from average value of 1896 microseconds to 54470 microseconds.  | warning | 5/9/2014 2:08:16 PM |      |        |        |
| Virtual machine ADC disks consolidation succeeded.                                                                                                              | info    | 5/9/2014 9:38:43 AM |      |        | vpuser |
| Device naa.60e9800042473274752b445648334734 performance has deteriorated. I/O latency increased from average value of 1898                                      | warning | 5/9/2014 9:10:13 AM |      |        |        |

The log shows that the I/O latency has increased. Then we search the website at <http://kb.vmware.com>, as shown in this screenshot:



Now, to troubleshoot the problem, let's follow these steps:

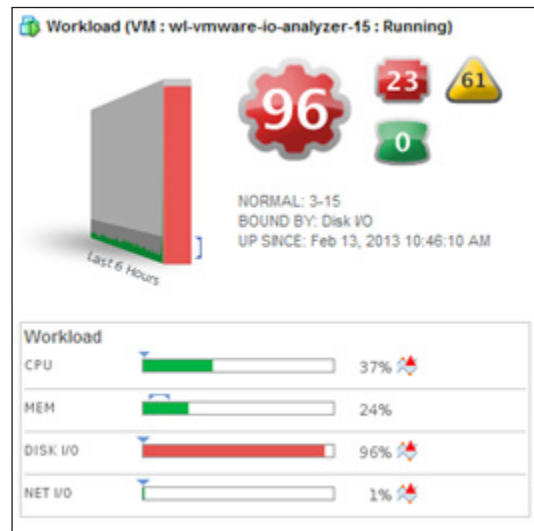
1. Check the log, includes `vm-support bundle`, `vmkwarning.log`, and `vmkernel.log`.
2. Look for the website [kb.vmware.com](http://kb.vmware.com) and then search for a keyword — storage latency. It will display a similar solution to solve the problem, the article ID **2076267**, **2007236**.
3. Check which job is associated with more I/O is currently running, such as VM backup, LUN replication, kill virus, unified patch to all VMs, and so on.
4. Check when the problem was triggered and for how long. Look for the rate of alarm.
5. Check the maximum queue of storage and other SCSI error return values.
6. Check whether only one host or an entire cluster has the problem.
7. Verify the level of the I/O latency.

Finally, determine the delay was caused by the database system at night time homework tasks! This is done so that we can adjust the operating time or migrate the other high-workload VM to another host.

The last example covers how to troubleshoot an FC storage problem using the vCenter log. This theory needs us to analyze more metrics from vCenter and logs, so we will take more time to solve it.

Now let me use another tool, vRealize operations manager, to troubleshoot the FC storage problem, and consider the differences between them.

vRealize operations manager is designed to take a holistic approach to analyzing all metrics. This gives the ability to monitor large environments without the need to track many individual counters. Let's begin our diagnosis by looking at the machine details in the vRealize operations manager, as follows:

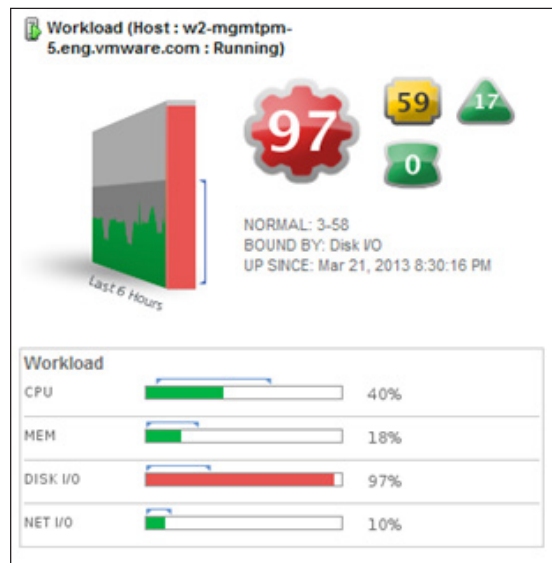


From the preceding badge, we can find the following observations about the VM:

- The VM is demanding 96 percent, much more than its entitlement
- Analytics has determined that the disk I/O is much more than normal behavior



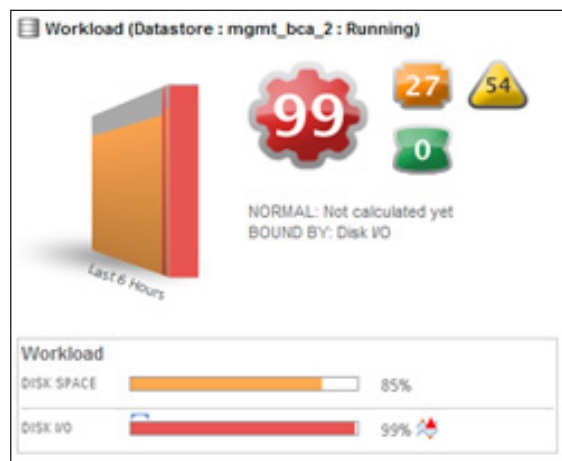
Now let's check the behavior of the host disk I/O by double-clicking on the workload badge under the related objects.



From the preceding host badge, we can gather these observations in regards to the host:

- The disk I/O demand of the host is 97 percent of its maximum disk capability
- Analytics has determined that the disk I/O is much above its normal

Next, let's look at **Datastore**:



From the workload badge of **Datastore**, we can note the following observations:

- 99 percent of datastore's disk I/O capability is being consumed
- Analytics has determined that the disk I/O is much heavier than its normal behavior

Based on the workflow and analysis, this is the conclusion:

- There is a throughput problem
- There are abnormal storage metrics
- The host is affected with the throughput
- Datastore is affected with disk I/O

The recommended next action is to distribute the VMS to other data stores.

## Summary

In this chapter, you learned about the FC storage stack in a vSphere environment. You also came to understand how to use the vCenter and vRealize operations manager, and how to troubleshoot the FC storage problem. The difference between these two tools is that vCenter gives the static metrics and confirms the problem that we need for analysis. vRealize has the smart engine, which will analyze the metrics itself and show us the root cause. So what is a very useful tool?

In the next chapter, you will learn how to troubleshoot the vSphere iSCSI storage.

Here are some metrics available with common thresholds that can be referred to, in order to troubleshoot any storage problems.

If the virtual disk read latency or write latency is more than 15 ms or 20 ms, then this will impact the performance of the VM. If the device's kernel disk command latency is more than 1 ms or the physical device command latency is more than 15 ms or 20 ms, then the host or array will face a base issue.



# 9

## Troubleshooting vSphere iSCSI Storage

In the previous chapter, you learned how to troubleshoot FC storage problems. Now we will go through some procedures used to troubleshoot iSCSI storage, and list some examples focusing on iSCSI storage troubleshooting.

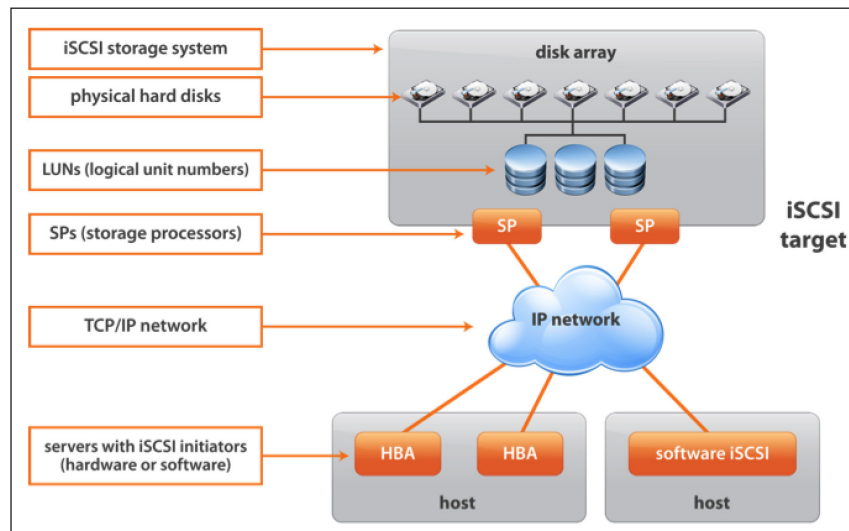
In this chapter, you will learn about:

- vSphere iSCSI storage components
- vSphere iSCSI storage troubleshooting examples

### vSphere iSCSI storage components

In the previous chapter, we discussed the details of FC storage troubleshooting. Now we will focus on iSCSI storage, so let's learn more about the iSCSI storage architecture. **iSCSI**, which stands for **I**nternet **S**mall **C**omputer **S**ystem **I**nterface, works on top of the **T**ransport **C**ontrol **P**rotocol (TCP) and allows the SCSI command to be sent from end to end over local area networks (LANs), wide area networks (WANs), or the Internet. It works by transporting block-level data between an iSCSI Initiator on a server and an iSCSI target on a storage device. The iSCSI protocol encapsulates SCSI commands and assembles the data in packets for the TCP/IP layer. Packets are sent over the network using a point-to-point connection. Upon arrival, the iSCSI protocol disassembles the packets, separating the SCSI commands so that the OS will see the storage as a local SCSI device that can be formatted as usual. Today, some of iSCSI's popularity in small- to medium-size businesses (SMBs) has to do with the way server virtualization makes use of storage pools.

In a virtualized environment, the storage pool is accessible by all the hosts within the cluster and the cluster nodes. Nodes communicate with the storage pool over the network through the use of the iSCSI protocol. In a vSphere environment, iSCSI storage architecture is used to store VM, template, and ISO images, as shown in this diagram:



Many users are concerned about performance and compatibility with storage networks, so iSCSI SANs took several years to catch up as mainstream alternatives to the FC. An FC SAN transmits data without dropping packets, and has traditionally supported higher bandwidths, but FC technology is expensive and requires a specialized skill base to install and configure properly. An iSCSI SAN, on the other hand, can be implemented with existing Ethernet network interface cards and switches, and run on an existing network. Instead of learning, building, and managing two networks – an Ethernet LAN for user communication and an FC SAN for storage – an organization can use its existing knowledge and infrastructure for both LANs and SANs.

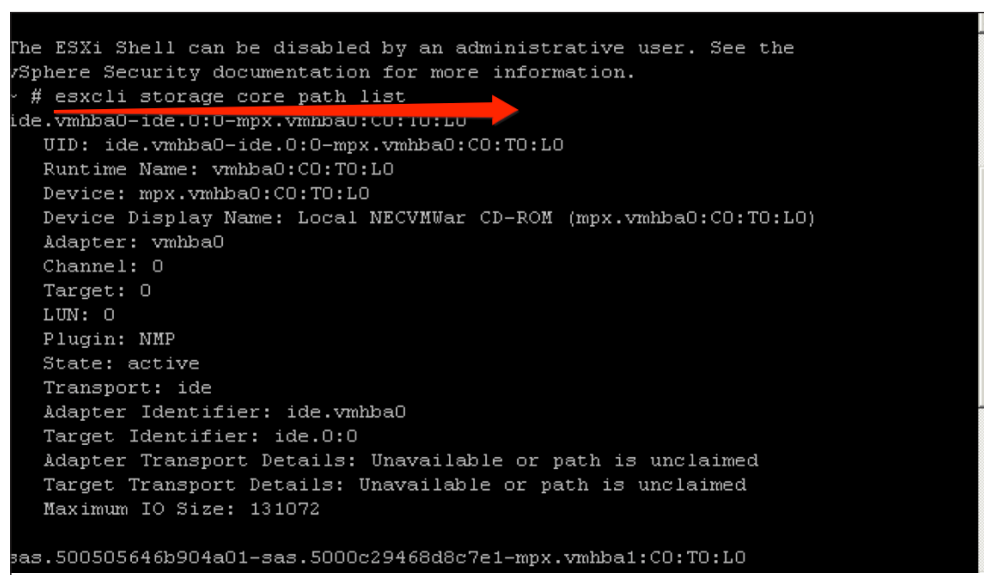
## vSphere iSCSI storage troubleshooting examples

In the last section, you learned about the iSCSI storage architecture. Now you will learn how to troubleshoot the problems in iSCSI storage. In a vSphere environment, the iSCSI storage problem that is most associated with the network configuration settings is that the storage can't connect or the performance does not keep up. In the following example, we're going to look at it!

The ESXi host can't access the iSCSI storage, so what is the problem?

1. First, check whether the ESXi host can see the LUN via this command:

**Esxcli storage core path list**

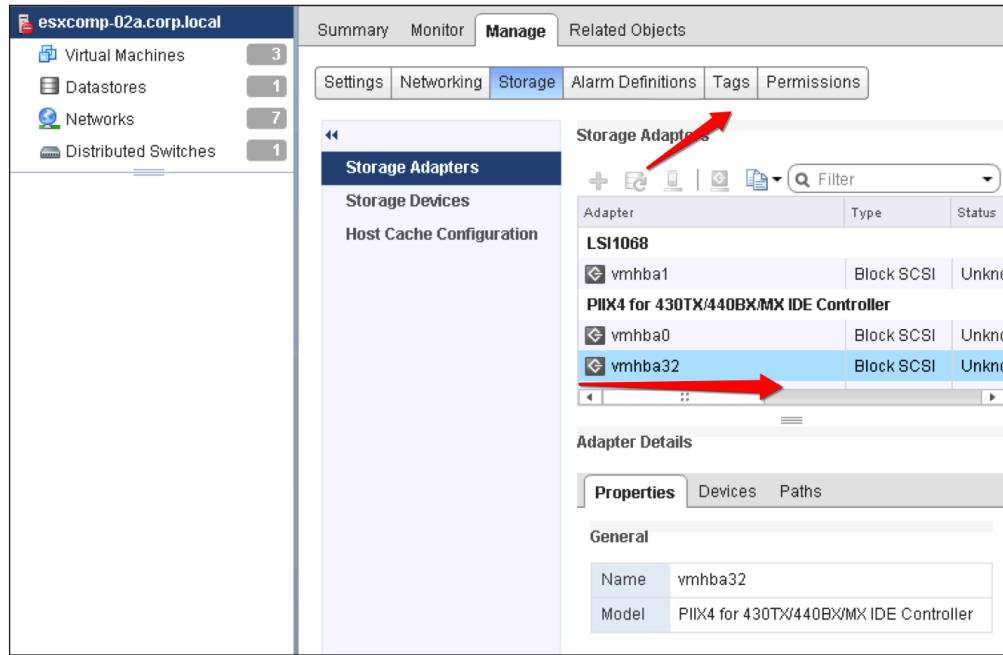


```

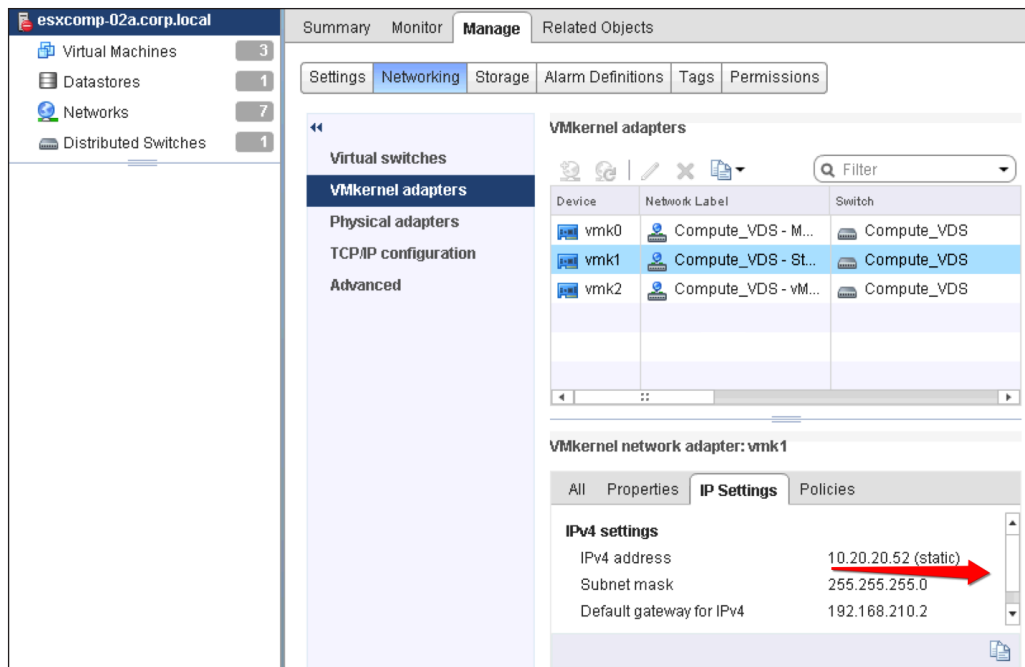
The ESXi Shell can be disabled by an administrative user. See the
vSphere Security documentation for more information.
> # esxcli storage core path list
ide.vmhba0-ide.0:0-mpx.vmhba0:CO:TO:LO
 UID: ide.vmhba0-ide.0:0-mpx.vmhba0:CO:TO:LO
 Runtime Name: vmhba0:CO:TO:LO
 Device: mpx.vmhba0:CO:TO:LO
 Device Display Name: Local NECVMWar CD-ROM (mpx.vmhba0:CO:TO:LO)
 Adapter: vmhba0
 Channel: 0
 Target: 0
 LUN: 0
 Plugin: NMP
 State: active
 Transport: ide
 Adapter Identifier: ide.vmhba0
 Target Identifier: ide.0:0
 Adapter Transport Details: Unavailable or path is unclaimed
 Target Transport Details: Unavailable or path is unclaimed
 Maximum IO Size: 131072
sas.500505646b904a01-sas.5000c29468d8c7e1-mpx.vmhba1:CO:TO:LO

```

2. Then check whether rescanning for storage has restored the visibility of the LUN using the GUI.



3. Next, check whether the ESXi host access IP was stored in the past, and whether you have changed the host configuration recently.  
You can use a bottom-up approach to troubleshoot, as described here.
4. Check for the iSCSI storage VMkernel interface configuration errors.
5. In the ESXi host, use this command: `ping 10.20.20.53` (the iSCSI target).



If the ping fails, check the IP address for the vmk that accesses the storage as shown previously. Ensure that the IP address is correct.

6. Check whether port 3260 is open between the ESXi host and SCSI storage:
  - Verify that the iSCSI storage array is configured properly and can be used normally. Ask the storage admin for support and confirm it.
  - Verify that the firewall interference is set to deny the iSCSI traffic, and ask the network admin for support and to confirm it.
7. Check out iSCSI to know whether the storage initiator is correct:
  - Verify the name of the iSCSI Initiator
  - Check whether the iSCSI target address and port number are correct or not
  - Verify if the CHAP setting is correct or not
8. Check out the hardware problem.
9. Then check out the VMware compatibility guide. Verify the iSCSI HBA or check whether the iSCSI storage arrays are supported.



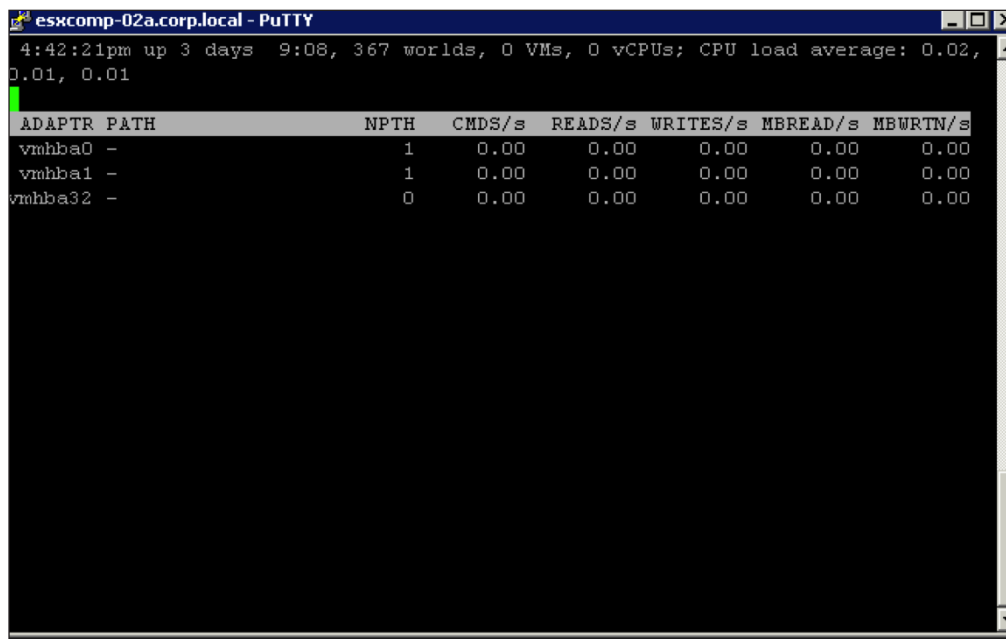
10. Then check whether the ESXi host of the LUN is correct:

- Verify all the ESXi hosts of the LUN, and also verify that it is located in the same storage group
- Check whether the LUN is configured correctly and whether it can be used with the ESXi host
- Verify that the LUN is not set to read-only mode on the array
- Then verify that the ESXi LUN host ID is below 255

If you have followed all these steps but still haven't found any problems, then maybe the storage is busy or is disconnected from the host. Now you can just use the monitor storage tools for troubleshooting.

In the ESXi Shell console, input this command:





**ESXTOP**



The screenshot shows a PuTTY terminal window titled 'esxcomp-02a.corp.local - PuTTY'. The terminal output displays system statistics at the top: '4:42:21pm up 3 days 9:08, 367 worlds, 0 VMs, 0 vCPUs; CPU load average: 0.02, 0.01, 0.01'. Below this is a table with storage performance metrics. The table has seven columns: ADAPTR, PATH, NPTH, CMDS/s, READS/s, WRITES/s, MBREAD/s, and MBWRIT/s. There are three rows of data for vmhba0, vmhba1, and vmhba32, all showing zero activity.

| ADAPTR  | PATH | NPTH | CMDS/s | READS/s | WRITES/s | MBREAD/s | MBWRIT/s |
|---------|------|------|--------|---------|----------|----------|----------|
| vmhba0  | -    | 1    | 0.00   | 0.00    | 0.00     | 0.00     | 0.00     |
| vmhba1  | -    | 1    | 0.00   | 0.00    | 0.00     | 0.00     | 0.00     |
| vmhba32 | -    | 0    | 0.00   | 0.00    | 0.00     | 0.00     | 0.00     |

In the command window, input the `d` command and list the storage adapter as shown before. Checkout the **DAVG**, **KAVG**, and **GAVG** to know whether all the metrics are okay or not. Follow the table values shown here:

| Object Level | Metric                                   | Threshold                                                                                      | Description                                                                                                                                    |
|--------------|------------------------------------------|------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------|
| VM           | Virtual Disk:Read Latency (ms)           | > 15-20ms<br> | The I/O read response time for a virtual disk. High latency here has an immediate impact on performance.                                       |
| VM           | Virtual Disk:Write Latency (ms)          | > 15-20ms<br> | The I/O write response time for a virtual disk. High latency here has an immediate impact on performance. Write is typically slower than read. |
| Host         | Devices:Kernel Disk Command Latency (ms) | > 1ms<br>     | (known in esxtop as KAVG) The latency incurred passing the I/O through the hypervisor. Latency here typically indicates a host-based issue.    |
| Host         | Physical Device Command Latency (ms)     | > 15-20ms<br> | (known in esxtop as DAVG) The latency incurred for an I/O headed to the array and back. Latency here typically indicates an array-based issue. |

## Summary

In this chapter, you learned about the iSCSI IP storage stack in vSphere environments, and how to troubleshoot iSCSI storage problems.

In the next chapter, you will learn how to troubleshoot vSphere NFS storage.



# 10

## Troubleshooting vSphere NFS Storage

In the previous chapter, you learned how to troubleshoot iSCSI storage problems. Now we will go through the procedure of troubleshooting **Network File Storage (NFS)** storage, and will list some examples that focus on NFS storage troubleshooting.

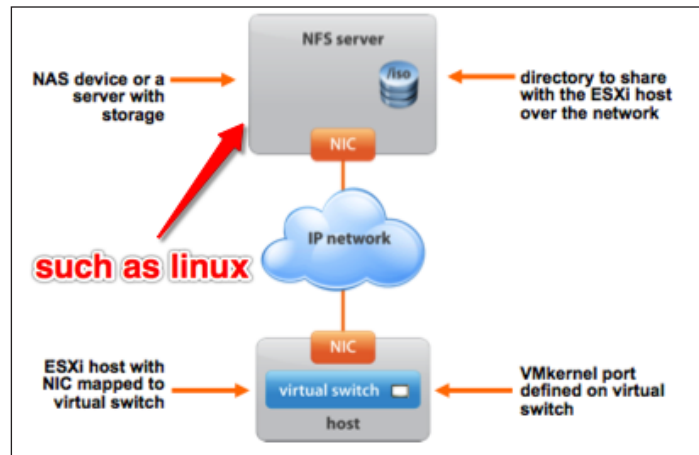
In this chapter, you will learn about:

- vSphere NFS storage components
- A vSphere NFS storage case study
- A vSphere NFS storage troubleshooting example

### vSphere NFS storage components

In the previous chapter, we discussed the details of iSCSI storage troubleshooting. Now we will focus on NFS storage, so let's learn more about the NFS storage architecture. NAS devices are storage arrays or gateways that support file-based storage protocols, such as NFS and **Common Internet File System (CIFS)**, and are typically connected via an IP network. These file-based protocols provide clients shared access to storage resources. This centralization of shared storage resources reduces management complexity, minimizes stranded disk capacity, improves storage utilization rates, and eliminates file server sprawl.

Many customers are expanding the use of NAS to include storage for relational databases such as Oracle and MySQL, server virtualization environments such as VMware VSphere, and virtual desktop solutions such as VMware Horizon View. In a vSphere environment, the ESXi host supports the NFS protocols. An NFS filesystem is located on a NAS device. This device is called the NFS server. The NFS server contains one or more directories that are shared with the ESXi host over a TCP/IP network. An ESXi host accesses the NFS server via one VMkernel port, as shown in the following diagram:

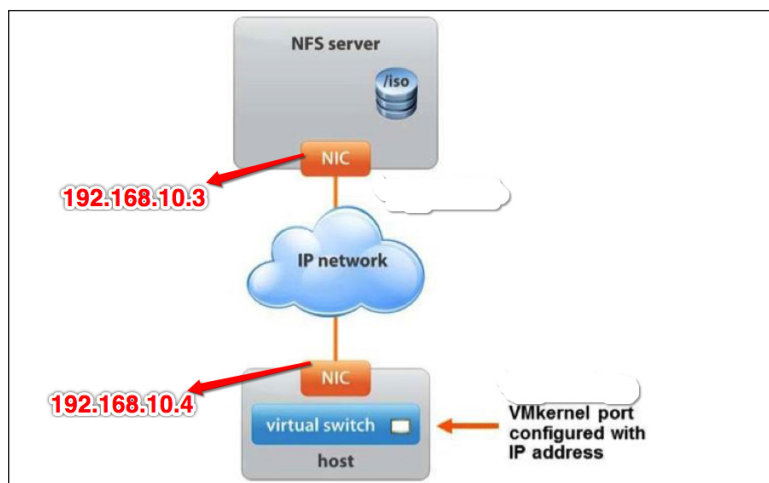


If you want to set up the vSphere ICM lab, just install one Linux server with NFS services installed and configure the NFS services. If you are running a production application, buy storage vendor products, such as Dell, NetApp, HP, and IBM.

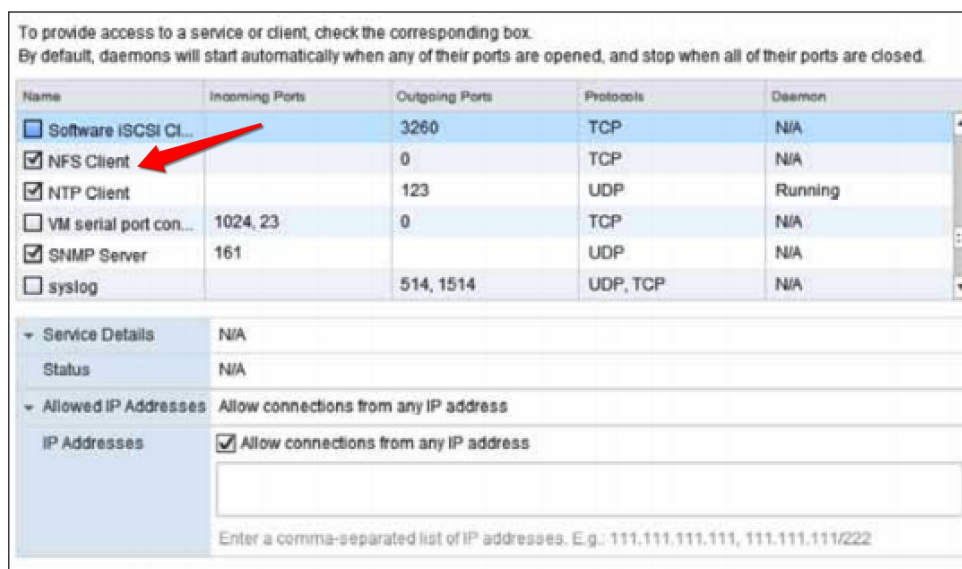
## A vSphere NFS storage case study

In the previous chapter, you learned about the NFS storage architecture. Now you will learn how to configure NFS storage. In the following steps, we will see how to mount an NFS data store in ESXi.

Create a new VMkernel port group for IP storage on an already existing virtual switch (vSwitch), or on a new vSwitch when it is configured. The vSwitch can be a **vSphere Standard Switch (VSS)** or a **vSphere Distributed Switch (VDS)**. You must create a new VMkernel port group to configure the vSwitch for IP storage access. You must also populate the network access information.

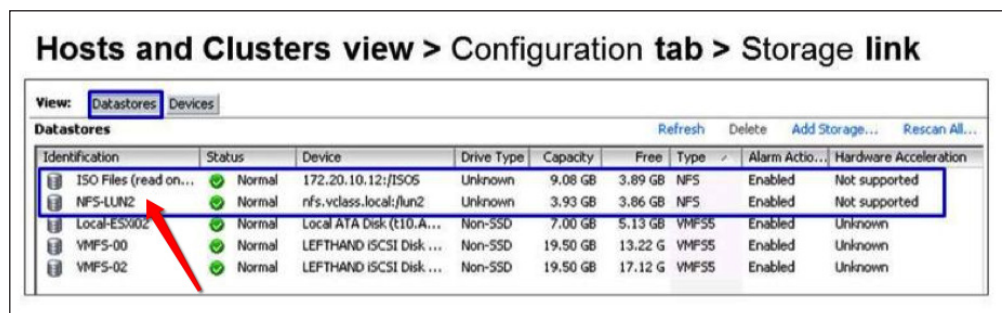


1. Ensure that the NFS client on the vSphere host (or hosts) is enabled. You must open the firewall port for the NFS client on all hosts to configure the NFS client on the vSphere host. To check whether the port is open, go to the ESXi host manage tab in VMware® vCenter™, select **Settings**, and then select **Security Profile**. Click on **Edit** and scroll down to the NFS client, as shown in this screenshot:



Configuring NFS client on the vSphere host

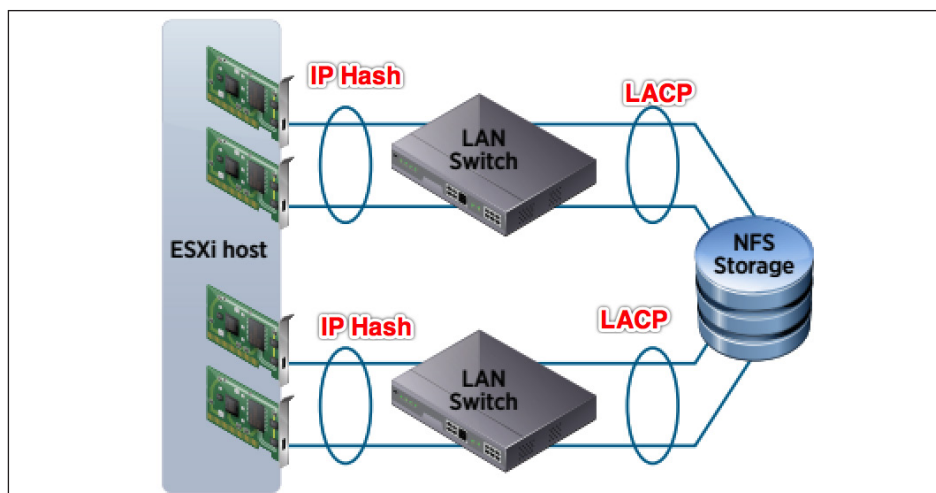
2. Ensure that the NFS storage is configured to export a mount point accessible to the vSphere hosts on a trusted network. VMware currently supports NFS version 3 over TCP only. You must also ensure that the vSphere host has root access to the data store. This is typically done on NFS servers using the `no_root_squash` option in the `/etc/exports` file. Different storage vendors have different methods of enabling this functionality. If you do not grant root access, you might still be able to mount the NFS data store on the vSphere host. However, you will not be able to create any virtual machines on the data store. It will fail with an **unable to access file** error message. For more details on NFS storage options and setup, consider the best practices for VMware provided by your storage vendor.



Mounting NFS datastore

To create a highly available NAS architecture, you must avoid single points of failure. The best approach is to have multiple NIC adapters that are configured as NIC teams installed on the ESXi host. Whether to apply a load balancing algorithm or not depends on whether your external switches support 802.3ad or you use Cisco switches or EtherChannel. NIC teams should be configured on separate external switches, with each NIC pair configured as a team at the respective external switch.

An even higher level of performance and high availability can be achieved with cross-stack EtherChannel-capable switches. With certain network switches, you can team ports across two separate physical switches that are managed as one logical switch. NIC teaming across virtual switches provides additional resilience, as well as some performance optimization. While using LACP or EtherChannel may provide additional redundancy, it will not provide any performance gain. VMware will create a single session between the source and target IP address, and all traffic will cross that single session. To provide additional throughput, multiple TCP sessions will need to be created using additional IPs at both the ESXi host and the NAS server/appliance.



## A vSphere NFS storage troubleshooting example

Most customers use NFS storage to store the virtual machine and ISO. Because it is based on IP access, generally, an NFS problem is that the network connection or NFS is under the attack of reading and writing. Let's learn two examples; one problem is a general error accessing, reading, and writing, and the other problems in the use of software backup.

One user wants to use the NFS data store to store the virtual machines, but wants to mount. It shows an error message similar to this:

**NFS Error: Unable to connect to NFS server**

**WARNING: NFS: 983: Connect failed for client 0xb613340 sock 184683088:  
I/O error**

**WARNING: NFS: 898: RPC error 12 (RPC failed) trying to get port for Mount  
Program (100005) Version (3) Protocol (TCP) on Server (192.168.10.4)**

**Network cable is unplugged**



Now, how to troubleshoot the issue? Let's follow the steps:

1. Check the MTU size on the port group that the NFS VMkernel used. You can use the `vmkping` command. You can also use the `esxconfig-vmknic` command to show all `vmknic` connection messages. If the MTU size is the same as end to end, that's okay,
2. Check the NFS server and ensure that it is accessible through the firewall. Follow the suggestion of the storage vendor.
3. Run the `netcat` command to verify the NFS server's `nfsd` port status.
4. Verify that the NFS host can ping the VMkernel IP of the ESXi host.
5. Then verify that the storage array is listed in the *Hardware Compatibility Guide*.
6. Verify that the physical hardware functions correctly.

If troubleshooting with the preceding steps does not resolve the issue, collect the ESXi host log and send it to the VMware support team.

Another use case is the Veeam software, which shows the error message as follows:

**Error during the configuration of the host: NFS Error: Unable to Mount filesystem: Unable to connect to NFS server**

Then the NFS data store can't mount, and the vPower NFS-based restorations fail.

Now let's troubleshoot this issue:

1. Check whether the Veeam vPower NFS service is running, and if not, start it.
2. Then check whether the Veeam server is connected to the VMkernel port on the ESXi host. If not, ask for help from the network team to verify that the subnet is okay.
3. Check whether the VMware environment is preventing the NFS data store from being mounted, such as the VMkernel port on the ESXi host is configured correctly, or the firewall on the ESXi host is open for the NFS client.
4. Cloud test that you can add another data store. If you can add one, the issue is the backup software, so ask the backup software companies' support to solve the problem.

## Summary

In this chapter, you learned about NFS storage in a vSphere environment.

The NFS storage problem is similar to iSCSI. Most issues are focused on network connectivity, so configure the NFS storage and be careful while setting the network. This lesson is over. By now, you have learned storage troubleshooting on FC, iSCSI, and NFS. In the next chapter, we will see how to design storage in a vSphere environment. Storage design includes the size of LUN, the IOPS of the data store, the multipath algorithm of storage, and more.



# 11

## vSphere Storage Design

In the previous chapter, we saw how to troubleshoot different types of storage problems. Now we will go through vSphere storage design. The focus will be on discussing how to design the storage comfortably within a vSphere environment.

In this chapter, you will learn about:

- Storage design key points
- The vSphere storage architecture
- Getting started with vSphere storage design

### Storage design key points

In this chapter, we are going to discuss how to design the storage that meets the needs of the business. In general, we purchase storage mainly by considering its high availability, performance, capacity, and cost! So let's talk about the important factors:

- The high availability of storage is critical. A high availability of storage is meant to ensure that in any case – whether the hard disk, controller, link switches, or the host's HBA is broken – the storage can still provide resources for the host to use and you can access the storage. So, a storage design must consider the redundancy of all components. First, hard disk redundancy can be achieved according to the demands of the hard disk, the corresponding RAID, controller to double alive, the trunk switch to redundancy, and the HBA to redundancy so that the entire link is redundant. As different types of storage, with greater levels of availability, are factored in, the cost obviously rises. However, the importance of availability should be overriding in almost any other storage design key points.

- Performance is generally less well understood than availability, but it has a much greater impact. We can use several metrics, such as **input/output operations per second (IOPS)**, MB/s, and latency, to accurately measure performance. The idea behind this IOPS calculator is to illustrate the importance of choosing the proper RAID and storage type for your workload. Once you understand how to measure performance, you can use it effectively to underpin a successful storage design.
- Capacity is what everyone thinks of as the focus of a storage array's principle specifications. Capacity needs to be managed on an ongoing basis and predicted and provisioned as required. The design should also consider the capacity of storage according to the application of space need to statistics to disk space. If the space is not satisfied, it is also likely to affect the business negatively. By the way, the design of capacity must also consider the development of the business and an emergency environment.
- Cost can be easy or difficult to factor in, depending on the situation. In storage designs, cost has a very big effect! If the clients have a very high budget, we can follow the best practices of the technology to design the storage and achieve the very best design goal. But if the customer's budget is limited, we have to balance the requirement from customers and budget accordingly. To master the scale of the balance, a simple technique is to meet the highest customer requirements with the lowest cost. This is the best design solution. Apart from the key points we've discussed, there are other storage design considerations, such as watts/IOPS, rack usage, management overhead, and flexibility.

## The vSphere storage architecture

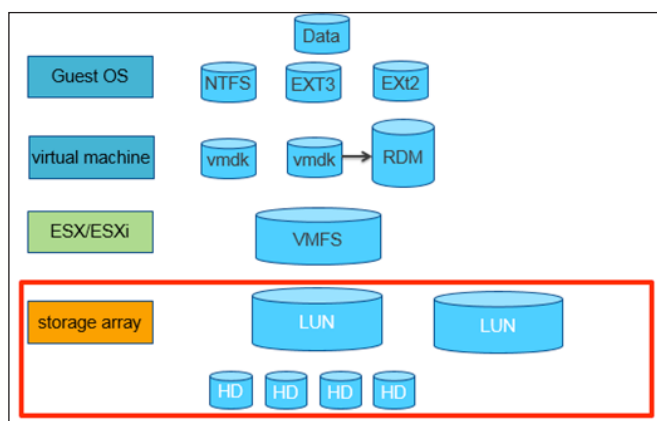
Data is what we normally use in operating systems, applications, documents, and so on.

A virtual machine is a pile of discrete files encapsulated in the form of the folder and then stored in the ESXi host's data store.

A VMware ESXi host's data stores are logical containers that hide the specifics of each storage device and provide a uniform model for storing virtual machine files. Depending on the type of storage that you use, data stores can be formatted with VMware vSphere VMFS, or mounted with a filesystem via the NFS protocol.

ESXi supports many more storage types, such as DAS, FC, FCOE, iSCSI, and NAS.

The DAS, FC, FCOE, and iSCSI storages are based on the block to read and write. Mainly, is formatted for the VMFS data store. The NAS is based on the file to read and write. The virtual machine's hard disk is written to the virtual disk file VMDK. The virtual machine's operating system, applications, and data, are written to the disk file. Shared storage is a requirement for a vSphere cluster. Shared storage means that all ESXi hosts in a vSphere cluster can see the same storage LUNs. This is necessary because all ESXi hosts need to be able to access the files of the virtual machines that run on top of the vSphere cluster. vSphere features, such as vMotion and high availability, work only if shared storage is available to all hosts. If a file is written in the guest OS, the file first writes in the operating system of the filesystem, such as NTFS or Ext3, then to the ESXi filesystem, and then to the storage system. The final data instructions are written to the physical hard disk. This is the vSphere storage infrastructure. Then, based on this architecture, we should consider how to design, as shown in the following diagram:



## Getting started with vSphere storage design

In the previous section, you learned how to use storage in a vSphere environment. Now, according to the following rule, we'll design our vSphere storage architecture.

**Service-level agreement (SLA)** defines the standard (storage-driven profile or virtual machine profile). Let the virtual machine store in the storage that meets the demand of its need.

## Share storage size – how many datastore requirements fit?

The most commonly discussed aspect of shared storage (LUN / data store) sizing is what limit should be implemented on the number of VMs per data store. The reason for limiting this number is to minimize the potential for excessive SCSI locking. Most mainstream storage vendors will provide VMware-specific guidelines for this limit, and VMware recommends an upper limit of five to eight VMs per VMFS, regardless of the storage platform. In many cases, it is forgotten that the number of VMs per LUN is also influenced by the size and I/O requirements of the VM, but perhaps more importantly the selected storage solution and even the disk types.

When considering the number of VMs to place on a single data store, some of the following factors should be considered in conjunction with any recommended VMs-per-LUN ratio:

- The average VM workload/profile (in particular, the amount of I/O)
- The typical VM size (including configuration files, logs, swap files, and snapshots)
- VMFS metadata
- The size of the LUN (include defined by the VM size and number of VMs per LUN)

## Virtual machines per LUN

When the VM configuration is defined, working assumptions can be created regarding the number of VMs that can be hosted on a VMFS data store. The sizing requirements for data stores are dependent on a lot more than just the VM `.vmdk` files. The following list summarizes the VM files that are key to sizing the LUN:

- The `.vmss` file is created when a VM is put into suspension, and is used to save the suspended state. In essence this is a copy of the VM's memory.
- The `-flat.vmdk` file is a raw disk file that is created for each virtual disk allocated to a given VM. It will be of the same size as that of the virtual disks added to the VM at the time of creation.
- The `-delta.vmdk` file is a differential disk file created when a snapshot of a VM is taken. After a snapshot is taken, data is no longer written to the base `.vmdk` file. Instead changes are written to the differential disk. A differential disk file will be created every time a snapshot is taken.

- The `.vswp` file is a swap file dedicated to a given VM to allow for memory overcommitment on an ESXi server host. This file is created when a VM is powered on, and will be equal in size to the unreserved memory configured on the VM. When VMs are created, the default memory reservation is 0 MB. Hence, the `.vswp` file will be equal to the amount of memory allocated to the VM. If a VM is configured with a 1,024 MB memory reservation, then the `.vswp` file will be equal in size to the amount of memory allocated to the VM, minus the 1024 MB reservation. It is essential that the size of the `.vswp` files be considered when defining LUN size requirements, because if there is insufficient space for these files to be created, then the VM cannot be powered on.

The sum of the predicted sizes of all of these files for an average VM within a vSphere deployment can be multiplied by the number of VMs to be stored per LUN to provide an estimate for the required size of an LUN.

For example, one VM has a total need of 50 GB (C:\20GB, D:28GB, Swap 2G) space. In order to define the maximum number of VMs per VMFS data store, it is necessary to consider a number of assumptions or working principles regarding the size of VM files. In this case, the following principles should be adhered to:

- Virtual machine disk files will be based on the mean used capacity + 50 percent
- Virtual machine swap files will be based on the allocated RAM
- Virtual machine configuration and log files will remain under 1 GB
- Virtual machine snapshots will include up to 20 percent changes and not exceed 7 days in age
- No more than three concurrent snapshots per VMFS data store
- Virtual machines will not be suspended
- Safety margin of 20 percent

An often used method to come up with a generic LUN size is the following equation:

$$\text{maxAmountOfVMs} = \text{LUNsize} / ((\text{avgSize} + 30\%) + \text{avgMEMSize})$$

$$\text{avgSize} = 50 \text{ GB}$$

$$\text{avgMEMSize} = 2 \text{ GB}$$

$$\text{VMSize} + = ((\text{avgSize} + 30\%) + \text{avgMEMSize}) = 67 \text{ GB}$$



The preceding reference is the recommended one. Some cases need consideration of your environment.





## Datastore types – how many types fit your environment?

VMware, NetApp, and EMC recommend that an application with high I/O requirements or one that is sensitive to latency variation requires a storage design that focuses on a particular VM, and it should be isolated from other datasets. Ideally, the data will reside on a VMDK stored on a data store that is connected to multiple ESX servers, yet is only accessed by a single VM. So, we need to consider how many data stores are needed.

These are the types of data stores. Follow your environment for this example:

- **Production VM:** Tier 1 VM, Tier 2 VM, Tier 3 VM, and single VM. Each Tier may have multiple data stores.
- **IT VM:** For example, controller VM
- **Staging VM :** From a P2V process or moving from non-production to production.
- **Isolated VM:** Test or DMZ.
- **Template and ISO:** Installer image – about 500 GB.
- **Desktop VM:** Mostly local data store on an ESXi host, backed by SSD.
- **SRM placeholder:** A VMware SRM solution – about 10 GB.

You should always know where a key VM is stored. Data store corruption, although rare, is possible.

*1 data store = 1 LUN*

Here are some other rules:

- Use thin provisioning at the array level, not the ESXi level. If you want to use it at the ESXi level, set up an alarm at the data store.
- Separate production and non-production. Add a process to migrate to production.
- The RAID level does not matter much if the array has sufficient cache (with the battery backed, obviously).
- There should be 30 percent free capacity for VM swap files, snapshots, logs, thin volume growth, and storage vMotion (interior).

## Storage I/O Control – is it needed to enable the feature?

SIOC monitors the latency of I/Os to data stores at each ESXi host sharing that device. When the average normalized data store latency exceeds a set threshold (30 ms by default), the data store is considered to be congested, and SIOC steps in to distribute the available storage resources to the virtual machines in proportion to their share. This is done to ensure that low-priority workloads do not monopolize or reduce the I/O bandwidth of high-priority workloads. SIOC accomplishes this by throttling back the storage access of low-priority virtual machines by reducing the number of I/O queue slots available for them. Depending on the mixture of virtual machines running on each ESXi server and the relative I/O shares they have, SIOC may need to reduce the number of device queue slots that are available on a given ESXi server.


Storage I/O Control is enabled. This allows cluster-wide storage I/O prioritization, providing the ability to control the amount of storage I/O that is allocated to the virtual machines during periods of I/O congestion. The shares are set per virtual machine and can be adjusted for each virtual machine based on the requirement. Even when no shares or limits are set per virtual machine, enabling Storage I/O Control will ensure that all VMs on a particular data store play fair in the event of a contention. This is true even if those VMs are running on different hosts.

## A Shared I/O Control settings explanation

Following are the explanation points for Shared I/O Control:

- **Storage I/O enabled:** Storage I/O is enabled per data store. Navigate to **Configuration** tab | **Properties** to verify that the feature has been enabled.
- **Storage I/O shares:** Storage I/O shares are similar to VMware CPU and memory shares. Shares define the hierarchy of the virtual machines for distribution of storage I/O resources during periods of I/O congestion. Virtual machines with higher shares have higher throughput and lower latency.
- **Limit IOPs:** By default, unlimited IOPs are allowed for a virtual machine. By allocating storage I/O resources, you limit the IOPs allowed to a virtual machine. If a virtual machine has multiple disks, you must set the same IOPs value for all the disks that access that virtual machine.


It is recommended that you use Storage I/O Control shares over limits whenever possible.

[  Storage IO Control is at the data store level, and there is no control at the RDM level. ]

## vStorage APIs for Array Integration

vStorage APIs for Array Integration is a feature introduced in ESX/ESXi 4.1 that provides the hardware acceleration functionality. It enables your host to offload specific virtual machines and storage management operations to the compliant storage hardware. With storage hardware assistance, your host performs these operations faster and consumes less CPU, memory, and storage fabric bandwidth. The following are the fundamental operations that VAAI helps improve. The main functions of these three are also given:

- **Hardware offloaded copy up:** This can provide up to 10 times faster VM deployment, cloning, and Storage vMotion. VAAI offloads the copy task at the array, enabling the use of a native storage-based mechanism, which results in the decrease of deployment time. But what is equally important is that it reduces the amount of data flowing between the array and the server.
- **Write same / Zero 10:** This can provide x times less I/O for common tasks. Take, for instance, a zero-out process. It typically sends the same SCSI command several times. By enabling this option, the same command will be repeated by the storage platform, resulting in reduced utilization of the server but also decreasing the time span of the action.
- **Hardware offloaded locking SCSI reservation conflicts:** VAAI solves the issue by offloading the locking mechanism to the array. This should reduce latency in an environment where thin provisioned disks are used, and even where VMware-based snapshots are used. ATS (the locking mechanism of the array) removes the need to lock the full VMFS volume. Instead, it locks a block when an update needs to occur.

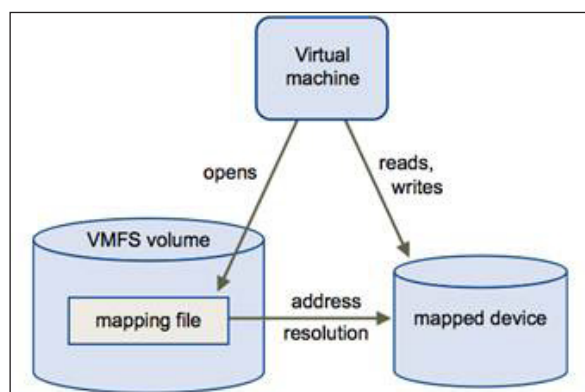
[  It is recommended that you install the VAAI plugin in your environment. ]

## RDM – does it need to be used in your environment?

A Raw Device Mapping (RDM) is a mapping file in a separate VMFS volume that acts as a proxy for a raw physical storage device. An RDM allows a virtual machine to directly access and use the storage device. It contains metadata for managing and redirecting disk access to the physical device.

This file gives you some of the advantages of direct access to a physical device, while keeping some advantages of a virtual disk in VMFS. As a result, it merges VMFS manageability with raw device access.

RDMs can be described in terms such as mapping a raw device to a data store, mapping a system LUN, or mapping a disk file to a physical disk volume. All of these terms refer to RDMs.



Although VMware recommends that you use VMFS data stores for most of your virtual disk storage, on certain occasions, you might need to use raw LUNs or logical disks located in a SAN.

For example, you need to use raw LUNs with RDMs in the following situations:

- When SAN snapshots or other layered applications run in the virtual machine. The RDM better enables scalable backup offloading systems using features inherent to the SAN.
- In any MSCS clustering scenario that spans physical hosts – virtual-to-virtual clusters as well as physical-to-virtual clusters. In this case, the cluster data and quorum disks should be configured as RDMs rather than virtual disks on a shared VMFS.

Think of an RDM as a symbolic link from a VMFS volume to a raw LUN. The mapping makes LUNs appear as files in a VMFS volume. The RDM, not the raw LUN, is referenced in the virtual machine's configuration. The RDM contains a reference to the raw LUN.

There are two compatibility modes available for RDMs:

- **Virtual compatibility mode:** This allows an RDM to act exactly like a virtual disk file, including the use of snapshots
- **Physical compatibility mode:** This allows direct access to the SCSI device for those applications that need lower level control

Please follow the preceding rules in your environment, and it is recommended that you use the VMDK instead of RDMs where possible.

## How to design ESXi multipathing policies

The default multipathing functionality for ESXi hosts is provided by the generic NMP. The VMware vStorage APIs for multipathing provide a framework to integrate third-party multipathing plugins (MPPs) into the ESXi platform. The following is an explanation, with recommendations, of the different multipathing policies available with ESXi.

**Most Recently Used (MRU):** Selects the first working path discovered at system boot time. If this path becomes unavailable, the ESXi host switches to an alternative path and continues to use the new path as long as it is available. This is the default policy for LUNs presented from an active/passive array. ESX/ESXi does not return to the previous path if or when it returns. It remains on the working path until it, for any reason, fails. This policy helps eliminate LUN trashing or trespassing of the LUNs in active/passive arrays.

- **Fixed (Fixed):** Uses the designated preferred path flag, if it has been configured. Otherwise, it uses the first working path discovered at system boot time. If the ESXi host cannot use the preferred path or if it becomes unavailable, ESXi selects an alternative available path. The host automatically returns to the previously-defined preferred path as soon as it becomes available again. This is the default policy for LUNs presented from an active/active storage array.
- **Round robin (RR):** Uses an automatic path selection rotating through all available paths, enabling distribution of the load across the configured paths:
  - For active/passive storage arrays, only the paths to the active controller will be used in the round robin policy

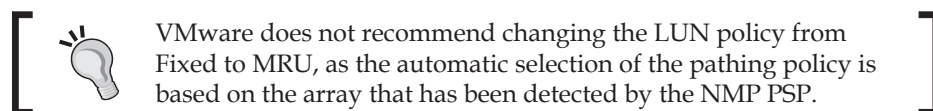
- For active/passive storage arrays, all paths will be used in the round robin policy



Third-party MPPs, such as PowerPath, run in parallel with the NMP, and for specified arrays, they replace the default NMP behavior by taking control of the path failover and load balancing operations. PowerPath/VE works with VMware vSphere to give enhanced path management capabilities to ESXi hosts. Having multiple paths enables the ESXi host to access a storage device even if a specific path is unavailable. Multiple paths can also share the I/O traffic to a storage device.

PowerPath/VE uses redundant physical path components—host based adapters (HBAs), switches, storage processors (SPs), and cables—between an ESXi host and an external storage device to provide fault tolerance. If one or more path components fail, the ESXi host can use a viable alternate path to access an external storage device. The process of detecting a failed path and switching to another path is called a path failover. A path failover helps ensure uninterrupted I/O between an ESXi host and external storage devices, allowing applications to continue accessing their data.

PowerPath/VE also redistributes the I/O load across multiple paths between an ESXi host and an external storage device. This process is called load balancing. Load balancing improves a host's ability to manage heavy I/O loads by continually balancing the load across all paths, eliminating the need for repeated static reconfigurations as workloads change.



For example, FC-SAN multipathing.

VMware recommends four paths:

- The path is point-to-point. The switch in the middle is not part of the path as far as vSphere is concerned.
- Ideally, they are all active-active for a given data store.
- Fixed means one path is active and three are idle.
- One zone per HBA port. The zone should see all the target ports.

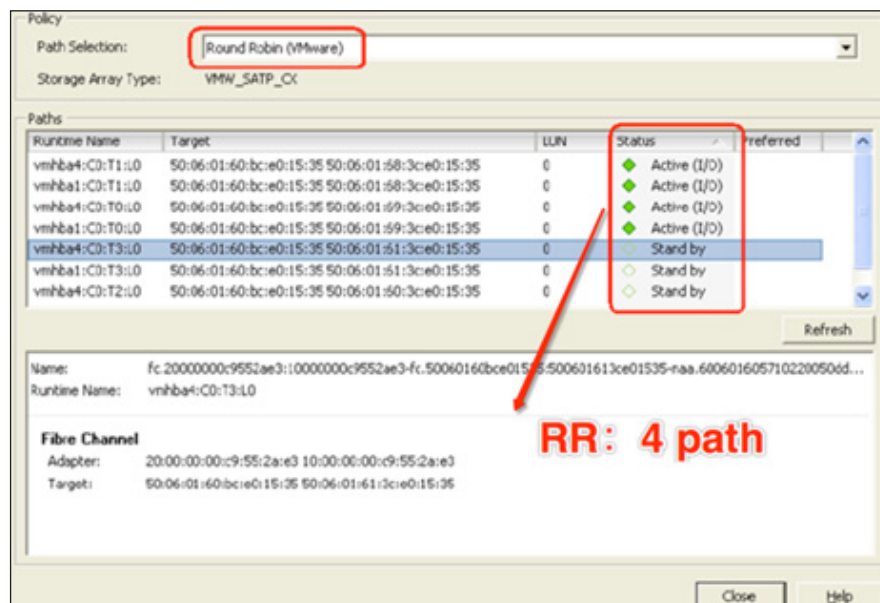
If you are buying new SAN switches, consider the direction for the next 3 years:

Whatever you choose will likely be in your data center for the next 5 years.

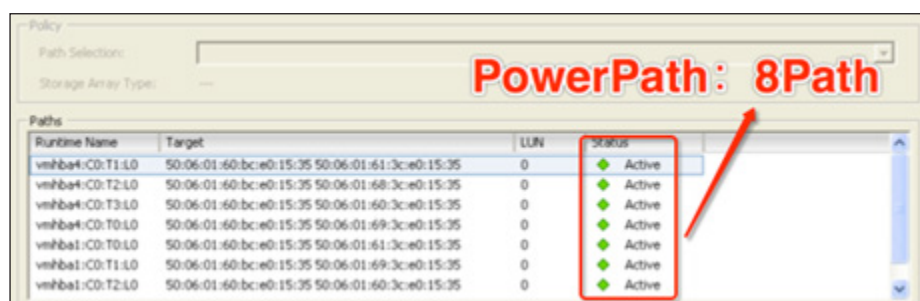
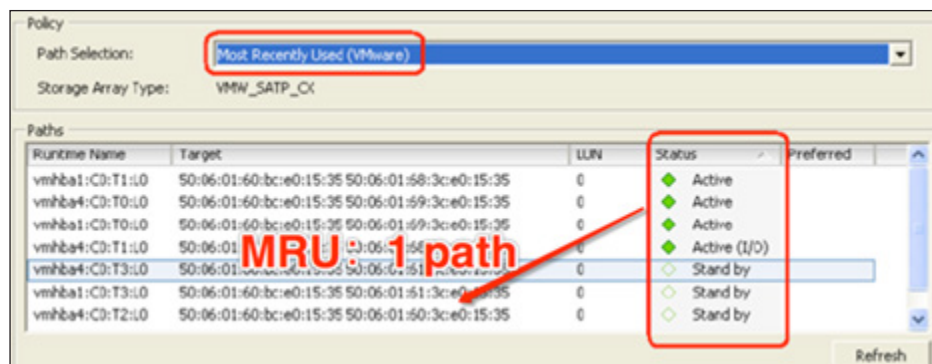
- If you are buying a director-class, then consider for the next 5 years. Upgrading director is a major task, so plan for 5 years of usage. Consider both the EOL and EOSL dates.
- Discuss with SAN switch vendors and understand their road map.
- 8 GB refers to FC, and FCoE should be 10, 40, or 100 GB depending on your infrastructure.

The RR policy states the following:

- It is per data store, not per HBA:
  - One ESXi host typically has multiple data stores
  - One array certainly has multiple data stores
  - All of these data stores share the same sp, cache, ports, and possibly spindles
- It is active/passive at a given data store.
- Leave the default setting at 1000. There is no need to set `iooperationslimit=1`.
- Choose this over MRU. MRU needs manual failback after a path failure.







As shown in the preceding screenshot, it is recommended that you use the vendor's Storage MultiPath Policy.

## How to design zoning and masking

Zoning is implemented in switches and controls, which HBA ports have access to which storage processor ports. Port zoning allows devices attached to particular ports on the switch to communicate only with devices attached to other ports in the same zone. The SAN switch (<http://searchtelecom.techtarget.com/definition/switch>) keeps a table indicating which ports are allowed to communicate with each other. Port zoning is more secure than WWN zoning, but it creates a number of problems because it limits the flow of data to connections between specific ports on the fabric.

Masking is done at the storage controller or server levels. Make a LUN "invisible" when a target is scanned.



Following are the tips for zoning:

- Implementing zoning:
  - One zone per HBA port
  - One HBA port does not need to know the existence of others
  - This eliminates the registered state change notification
- Use soft zoning, not hard zoning:
  - **Hard zone:** This zone is based on the SAN switch port. Any HBA that connects to this switch port gets this zone. So, this is more secure. But be careful when re-cabling things into the SAN switch.
  - **Soft zone:** This zone is based on the HBA port. The switch port is irrelevant.
  - **Situations that need rezoning in the soft zone are as follows:** Changing the HBA, replacing the ESXi server (which comes with a new HBA), and upgrading the HBA.
  - **Situations that need rezoning in the hard zone are the following:** Reassigning the ESXi to another zone and port failure in the SAN switch.
- Implementing LUN Masking:
  - Complement zoning. Do this at the switch level, not the ESXi level.
  - Mask on the array, not on each ESXi host.
  - Masking done at the ESXi host level is often based on the controller, target, and LUN numbers, all of which can change with the hardware configuration.

## Summary

In this chapter, you learned about vSphere storage design. When creating the storage design of a vSphere environment, you should consider its capacity to meet the requirements of the concerned applications and services. It should also satisfy the entire link redundancy, and meet the performance requirements for applications, within the budget limits. Keep in mind to put the VM in the correct storage.

In the next chapter, we will see how to design VMware ESXi hosts.

# 12

## ESXi Host Design

In the previous chapter, we saw how to design different types of storage in a vSphere environment. Now we will go through vSphere ESXi host design, with a focus on discussing how to design the ESXi host best.

In this chapter, you will learn about:

- ESXi host design key points
- An ESXi host design example

### ESXi host design key points

In this chapter, we will discuss how to design the ESXi host and make it meet the requirements of the business. In general, when we purchase an ESXi host, we mainly consider the data center CPU type and CPU capacity, data center memory capacity, host hardware type, and so on.

### CPU capacity

ESXi hosts are the fundamental compute building blocks of a virtual data center.

ESXi host resources are distributed in order to run virtual machines. They are aggregated to build clusters of highly available pools of compute resources. Now, if you want to design the ESXi, the first step is to identify useful information from the customer, including the CPU type, CPU capacity, memory capacity, and the customer's preferred hardware vendor.

Now how can we go ahead and determine all of the CPU capacity required? Just follow these steps:

1. Choose the appropriate tools to analyze the current system CPU capacity required. Many tools can help do that, such as the VMware capacity planner and vRealize operation manager. Without these tools, the OS's task manager can also be used for analysis, but it needs a lot of data collection, statistical analysis, and manpower to do all the work.
2. Refer to the operating system and application of the manufacturer's suggested document. Check whether the VMware data center CPU capacity is accurate.
3. Ask for the customer to determine the future capacity requirement, or requirements.
4. We can describe the total CPU capacity as:  $Total\ CPU = application\ requirement + headroom$ . The headroom is generally 30-50 percent. It is for HA, host maintenance, short-term usage increase, and so on.

## Number of hosts

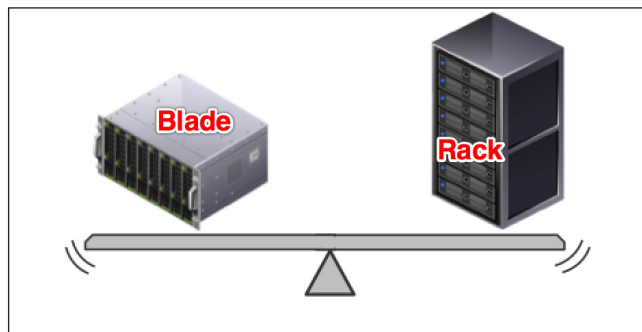
Follow the preceding steps to determine the total CPU capacity. Now how do we determine how many hosts the data center needs? This is not as simple as just dividing the total required CPU capacity by the planned CPU capacity per host. So, follow these instructions:

1. Check whether the customer has already designated a specific supplier. If it is, then according to the customer request to confirm the number of hosts.
2. Confirm that each host CPU and memory match, and check whether the host supports enough memory to keep all the CPUs busy.
3. Consider the number of I/O ports available per host.
4. Consider the vSphere cluster HA setting. Generally, it will keep one more host for the total capacity, just like N+1.
5. Decide which cluster you want to deploy, scale out or scale up. A scale out cluster's advantage is that when an HA failover occurs, the virtual machine is relatively affected less, and better use of DRS function to balance resources. A scale up cluster's advantage is that makes full use of host memory resources.
6. Consider the total memory requirement and each host's memory requirement.
7. Compared with the CPU capacity calculation of count and memory capacity calculation of count host and pick up at least as a host number.

## Host hardware types

Now let's see how to choose the type of server. We have two choices: blade and rack.

If a specific blade or rack-mount server meets the requirements of the design, then it is a great solution. The choice between a blade or rack-mount server might not be due to focus on any technical decision, but might depend on other factors. For example, the customer's relationship with a specific vendor might determine the choice. Eliminating this factor, will consider this instead from the point of view of the technology. Let's take a look at these two types, as shown in the following screenshot, and see how to choose one:



The blade server modularity offers both advantages and disadvantages.

These are its advantages:

- Easy to manage and replace the hardware quickly
- Saves operation costs
- Shares components such as power supplies, DVD-player, fans and so on

The following are its disadvantages:

- The chassis is shared, so it will become a single point of failure
- It needs more powering and cooling than the rack type
- The entire chassis must be supplied by the same vendor


The rack server modularity also offers both advantages and disadvantages.

Listed here are its advantages:

- Supplies more I/O expansion capability
- Easy heat dissipation
- The rack server can be supplied by different vendors

The following are its disadvantages:

- Takes up a lot of space, so more cable management is required
- Higher cost of DVDs and fans

[  Now, after seeing their strengths and weaknesses, design can be considered according to the demand. But no matter what you choose—blade or rack architecture—it must satisfy the VMware hardware compatibility list. ]

## Host naming conventions

Many technical consultants don't pay attention to considering the design scheme to a hostname, but the hostname is very important for operational management. So you should name the host according to the following rules.

Use simple, descriptive, and easy-to-understand hostnames. A standardized naming convention results in easier management, troubleshooting, and use, all of which reduce operational expenses. An example is `esxi-  
<locationcode-##>.<domain.name>`, `esxi-shanghai02.test.com`.

## An ESXi host design example

The following is our case background: customer X has invested in an enterprise-designed virtual infrastructure for the x86 environment within its big company centers. These investments, over time, have provided customer X with both financial and operational benefits, in addition to a more highly available and reliable platform for delivering virtual server services to the core business functions and faculties. Through this investment in vSphere, customer X is continuing to standardize all x86 services on the VMware vSphere platform to maintain and maximize their capital, operational, and availability benefits. Within this context, the current Sun E25K servers are up for retirement, and due to the high cost of maintaining the environment, these assets have been selected for migration from the current Oracle on SPARC Solaris, which is on the Sun E25K servers, to Oracle on Red Hat Linux, which is on vSphere.

Customer X aims to achieve the following benefits:

- Effectively utilize data center resources
- Maximize availability of the Oracle environment at lower cost

- Rapidly deploy Oracle database servers for development, testing, and production
- Maximize uptime during planned maintenance
- Minimize Oracle data center costs of floor space, energy, cooling, hardware, and labor
- Minimize planned and unplanned downtime
- Automate testing and failover of Oracle data center environments for disaster recovery and business continuity
- Achieve IT compliance

Customer x's ESXi design scenario:

- The solution will provide a platform that is horizontally scalable on commodity hardware.
- It will support recovery of application services within a 4-hour RTO for production systems.
- Also, it will be available 24x7 and allow for non-disruptive infrastructure maintenance.
- The solution will be easily scalable horizontally so as to handle any future increases in the number or size of virtual machines.
- Dell R910 servers will be purchased for management hosts.
- The maximum memory and CPU utilization per host should not exceed 85 percent during normal operations for production hosts and 90 percent for non-production hosts.
- The solution and the VMs must be able to easily scale horizontally and vertically to allow for future growth.
- Vendor Y will provide the server hardware.
- QLogic HBAs will be used in the ESXi hosts.
- Each host will have a single dual-port QLogic QLE2562 8 GB/s HBA card for access to the fibre channel shared storage infrastructure.
- Each ESXi host will contain a single dual-port HBA. This is a single point of failure for access to the fibre channel shared storage infrastructure, and if the HBA card of PCI slot fails will cause loss of storage access to all VM's on the host.

According to the customer's requirements and restrictions, the design is as explained in the following section:

## Management cluster specifications

Here is a table providing the specifications:

| Attribute                   | Specification                   |
|-----------------------------|---------------------------------|
| Host type and version       | ESXi 5.5                        |
| Number of CPUs              | 2                               |
| Number of cores             | 6                               |
| Total number of cores       | 12                              |
| Processor speed             | 2.1 GHz                         |
| HT and turbo boost enabled? | Yes                             |
| Memory                      | 96 GB (6 x 16 GB 1333 MHz DDR3) |
| Number of NIC ports         | 4                               |
| Number of HBA ports         | 2                               |

## Oracle cluster specifications

Here is a table providing the specifications:

| Attribute                   | Specification                |
|-----------------------------|------------------------------|
| Host type and version       | ESXi 5.5                     |
| Number of CPUs              | 4                            |
| Number of cores             | 6                            |
| Total number of cores       | 24 (48 LCPU with HT)         |
| Processor speed             | 6.92 GHz                     |
| HT and turbo boost enabled? | Yes                          |
| Memory                      | 192 GB (12 x 16 GB 1333 MHz) |
| Number of NIC ports         | 6                            |
| Number of HBA ports         | 2                            |

## VMware ESXi physical specifications

Here is a table providing the specifications:

| Attribute                       | Specification                 |
|---------------------------------|-------------------------------|
| Vendor and model                | Dell PowerEdge R910           |
| Processor type                  | Hex core x64 Intel Xeon X5770 |
| Total number of cores           | 12 (24 LCPU with HT)          |
| On-board NIC vendor and model   | Broadcom BMC5709C             |
| On-board NIC ports x speed      | 4 x Gigabit Ethernet          |
| Number of attached NICs         | 1 (excluding on-board)        |
| NIC vendor and model            | Intel X520-2                  |
| Number of ports/NIC x speed     | 2 x 10-gigabit Ethernet       |
| Total number of NIC ports       | 6                             |
| Storage HBA vendor and model    | QLogic QLE2562                |
| Storage HBA type                | Fibre Channel                 |
| Number of HBAs                  | 1                             |
| Number of ports/HBA x speed     | 2 x 8 GB                      |
| Total number of HBA ports       | 2                             |
| Number and type of local drives | N/A                           |
| RAID level                      | N/A                           |
| Total storage                   | N/A                           |
| System monitoring               | IPMI-based BMC                |

## Oracle cluster VMware ESXi physical specifications

Here is a table providing the specifications:

| Attribute             | Specification                 |
|-----------------------|-------------------------------|
| Vendor and model      | Dell PowerEdge R910           |
| Processor type        | Hex core x64 Intel Xeon X5690 |
| Total number of cores | 24 (48 LCPU with HT)          |



| Attribute                       | Specification           |
|---------------------------------|-------------------------|
| On-board NIC vendor and model   | Broadcom BCM5709C       |
| On-board NIC ports x speed      | 4 x Gigabit Ethernet    |
| Number of attached NICs         | 1 (excluding on-board)  |
| NIC vendor and model            | Intel X520-2            |
| Number of ports/NIC x speed     | 2 x 10-gigabit Ethernet |
| Total number of NIC ports       | 6                       |
| Storage HBA vendor and model    | QLogic QLE2562          |
| Storage HBA type                | Fibre Channel           |
| Number of HBAs                  | 1                       |
| Number of ports/HBA x speed     | 2 x 8GB                 |
| Total number of HBA ports       | 2                       |
| Number and type of local drives | N/A                     |
| RAID level                      | N/A                     |
| Total storage                   | N/A                     |
| System monitoring               | IPMI-based BMC          |

The configuration and assembly process for each system is standardized, with all the components installed the same on each host. Standardizing not only the model but also the physical configuration of the ESXi hosts is critical for providing a manageable and supportable infrastructure—it eliminates variability. A consistent PCI card slot location, especially for network controllers, is essential for the accurate alignment of physical-to-virtual I/O resources. All of the ESXi host hardware, including the CPUs, were selected by following the vSphere hardware compatibility lists.

## Summary

In this chapter, you learned about ESXi host design.

The next chapter will teach you how to design a virtual machine, and we will continue following customer X's request to build the VM design example.

# 13

## Virtual Machine Design

In the previous chapter, you learned how to design the ESXi host in a vSphere environment. Now we will go through vSphere virtual machine design, focusing on discussing how VM is designed to work comfortably with the applications' needs.

In this chapter, you will learn:

- Virtual machine design key points
- Virtual machine design example

### Virtual machine design key points

Virtual machines run applications and services that support individual users and entire lines of businesses. They must be properly designed, provisioned, and managed for efficient operation of the relevant applications and services. In the following section, we will discuss the key points, that is, how to design a virtual machine.

### Number of virtual CPUs

The number of virtual CPUs (vCPUs) required for a virtual machine depends on the operating system, application, and workload. You must configure a single vCPU, unless the need for more is clear. If more vCPUs are needed, the operating system and application must support them. The operating system must also support **symmetric multiprocessing (SMP)**, and the application must be multithreaded.

If the application is not multithreaded, use a scale out strategy by installing multiple virtual machines, each providing the same service or application for different users. If the workload requires multiple vCPUs, configure as few of them as possible. The more the vCPUs a virtual machine has, the more the CPU and memory overhead necessary to manage it. This additional overhead needlessly wastes resources if the additional vCPUs are not required.

It must be possible for all vCPUs to be scheduled on physical processors at the same time when the virtual machine is scheduled to run. This, however, does not mean that every vCPU is scheduled to run on a physical processor when the virtual machine is scheduled to run. The number of vCPUs that run at once depends on the operation being performed at that moment.

Here are some tips:

- Going from one vCPU to many is okay  
Windows XP and Windows Server 2003 automatically upgrade to the ACPI Multiprocessor HAL
- Going from many to one is not okay
- To change from one vCPU to two, you must change the kernel to SMP

In Windows 2000, you can change to any listed **hardware abstraction layer (HAL)** type. However, if you select an incorrect HAL, the computer may not start correctly. Therefore, only the compatible HALs are listed in Windows Server 2003 and Windows XP.

With some operating systems, scheduling inefficiencies negatively affect performance if a multiprocessor HAL runs on a uniprocessor virtual machine. For information about operating systems related to HAL, look up the guest operating system documentation.

## Ensuring the memory's performance

To ensure the virtual machine memory's performance, you have to keep a virtual machine's active memory in physical RAM:

- Limit the host memory over commitment, or configure the virtual machine reservations, or both.
- If reservations are configured, set them slightly above the virtual machine's average active memory size. You can use the vCenter performance chart to monitor the active memory.
- Virtual machine reservations increase administrative overhead, so it might be better to design a consolidation ratio that does not overcommit active memory.
- Verify that the virtual machine is installed with VM Tools. If installed, the ballooning driver could adjust the memory.

## VM resource setting – limit, reservation, and share

To simplify the configuration and administration, use the default virtual machine CPU and memory shares, reservation, and limit, unless a clear reason to do otherwise exists. Are there critical applications or services that must continue to receive resources even during periods of resource contention? If so, you can use reservations to guarantee the resources and shares to set relative priorities.

For example, if the consolidation ratio is too high, you might have to configure a virtual machine's memory reservation equal to the average active memory to maintain acceptable application performance.

If a department or group requires, or has paid for, a certain amount of CPU or memory resources, use a resource pool and not individual virtual machine reservations to reserve these resources.



In the very rare case that you want to save or are concerned about data store space, raise the virtual machine memory's reservation to the virtual machine memory limit. This memory setting prohibits the creation of a swap file, effectively disables ballooning, and prevents overcommitment. If the resource is enough, keep the VM resource setting in default state, If the resource is a contention, adjust the share values, not the reservation.

## Virtual machine hard disk – how to deploy and which types?

Deploy a system disk and a separate application data disk:

- A separate system disk simplifies provisioning.
- A separate application data disk simplifies backup.
- A separate application data disk can easily be increased in size if necessary.
- Separate disks help distribute the I/O load:
  - Do not place all the system disks in a single data store and all the data disks on another.
  - Place a virtual machine's system and data disks on the same data store, unless they have widely varying I/O characteristics.
- Configure one partition per virtual disk. If there are multiple partitions, you can extend the last partition only if you increase the size of the disk.

- Consider thin provisioned disks if data growth is slow or static, such as in a web server.
- How the operating system initially formats its disks can negate the benefits of thin provisioning. If the format operation writes zeroes to all sectors, the disk will be prematurely inflated. Use the operating system vendor's documentation to find the format options that do not write zeroes to all disk sectors during format operations.

## **Multiple virtual disks**

If a virtual machine has multiple disks, keep them on the same data store if possible. This simplifies configuration and administration. For example, if you plan to replicate the virtual machine's files to another LUN or data center for availability, replicating a single LUN replicates all of the virtual machine's files. This configuration simplifies the configuration and administration of products such as VMware vCenter Site Recovery Manager.

There are reasons for placing a virtual machine's disks on separate LUNs. If the virtual disks have different I/O characteristics, you might place them on separate LUNs that accommodate those characteristics. Another reason might be that a data disk is exceptionally large. Such a large disk might be kept on another data store or on another LUN accessed using raw device mapping.

## **Virtual disk location**

In most cases, storing virtual machine disks together on the network storage rather than on the local storage makes sense. Nearly all of the more important benefits of vSphere, such as virtual machine migration and availability, depend on network storage.

In rare cases, local storage might be required. For example, local storage might be more secure. Local storage might also be less expensive and might be used by smaller organizations. Local storage can also be used to create a virtual SAN environment.

## **Swap file location**

There are four main alternatives for swap file location:

- On shared storage, with the virtual machine files
- On local storage, with the virtual machine files on shared storage
- On dedicated shared storage, not with the virtual machine files
- On local storage, with the virtual machine files

A virtual machine swap file location may affect the following:

- vMotion performance
- Ease of administration and provisioning
- Data store replication performance

Consider placing swap files on solid-state drives to reduce performance issues caused by actively using swap files.

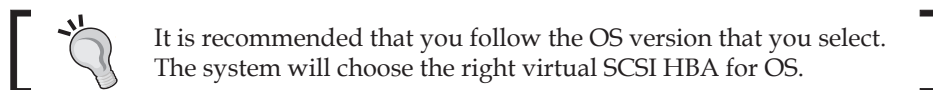


## Virtual SCSI HBA type – which one fits your OS?

Use the default choice, unless it does not support a required feature. For example, Microsoft Windows 2008 cluster services require a serial attached SCSI device.

If you consistently need to configure a nondefault choice, create a template to simplify virtual machine provisioning. Verify that you have sufficient template space and template access, and ready the driver to install by yourself.

One exception for using the default choice might be the configuration of a **Paravirtual SCSI (PVSCSI)** HBA; its performance is best.



## Virtual NICs

If the OS supports VMXNET3, choose the newest driver. The bandwidth will be 10 GB/s.

In all cases, use a virtual network adapter that is supported by the guest operating system and VMware Tools.

## Virtual machine hardware compatibility

With each new version of ESXi, new features and capabilities are added to the virtual machine hardware. Depending on the virtual hardware level and the required feature set, the virtual machine might not run on older hardware. This introduces potential compatibility problems if newer versions of virtual hardware are used but there are older hosts in the clusters, as a VM will not be able to run on older hardware if using a newer hardware version.

With ESXi 5.5, virtual hardware version 10 is introduced, with several new features.

As reference, the compatibility levels supported in vSphere 5.5 are shown in the following table:

| vSphere release            | Virtual machine hardware version | Compatibility                 |
|----------------------------|----------------------------------|-------------------------------|
| Virtual Infrastructure 3.5 | Version 4                        | VMware ESX/ESXi 3.5 and later |
| vSphere 4.0                | Version 4                        | VMware ESX/ESXi 4.0 and later |
| vSphere 4.1                | Version 7                        | VMware ESX/ESXi 4.0 and later |
| vSphere 5.0                | Version 8                        | VMware ESXi 5.0 and later     |
| vSphere 5.1                | Version 9                        | VMware ESXi 5.1 and later     |
| vSphere 5.5                | Version 10                       | VMware ESXi 5.5 and later     |



Keep the VM use the latest version. Then, it will support more features and improve performance.



## Considering guest OS

Let us consider the following points:

- Keep the variations minimal for each guest operating system. This approach simplifies administration and troubleshooting.
- Use standardized templates for the installation of every key application.
- Use standard sizing for virtual machines. This approach simplifies administration, troubleshooting, and chargeback.
- Disable unnecessary devices from the guest OS.
- For Windows VM, stagger an anti-virus scan. The performance will degrade significantly if you scan all VMs simultaneously.
- Use a 64-bit OS if possible.

## A virtual machine design example

The initial, or base, design of virtual machines will have a standard configuration in terms of virtual machine hardware configuration settings and virtual disk sizes, based on a small number of different template sizes. The *customer x* configuration standard for each of the virtual machine standard template sizes is shown in this table:

| Hardware configuration option | Small                                                                 | Medium                | Large                  | Extra large             |
|-------------------------------|-----------------------------------------------------------------------|-----------------------|------------------------|-------------------------|
| SCSI controller               | LSI logic parallel: OS<br>PVSCSI (x3): app and redo, data, and backup |                       |                        |                         |
| Virtual CPU                   | 2vCPU                                                                 | 4 vCPU                | 6 vCPU                 | 8 vCPU                  |
| Virtual memory                | 8 GB (default reserved)                                               | 16 GB (4 GB reserved) | 32 GB (10 GB reserved) | 128 GB (48 GB reserved) |
| Virtual NICs                  | Two virtual NICs: VMXNET3                                             |                       |                        |                         |
| CD-ROM                        | One:- IDE                                                             |                       |                        |                         |
| Floppy                        | 0 (removed)                                                           |                       |                        |                         |
| Windows OS virtual HDD size   | 30 GB (C) and 40 GB (D)                                               |                       |                        |                         |
| Linux /home                   |                                                                       | 5 GB                  |                        |                         |
| Linux /tmp                    |                                                                       | 10 GB                 |                        |                         |



| Hardware configuration option                    | Small | Medium                                                                  | Large | Extra large |
|--------------------------------------------------|-------|-------------------------------------------------------------------------|-------|-------------|
| Linux swap                                       |       | 10 GB                                                                   |       |             |
| Linux /                                          |       | 20 GB                                                                   |       |             |
| Linux /var                                       |       | 10 GB                                                                   |       |             |
| Linux /redolog                                   |       | 10 GB                                                                   |       |             |
| Linux /ora_archive                               |       | 50 GB                                                                   |       |             |
| Linux /oracle                                    |       | 40 GB                                                                   |       |             |
| Linux /u01 or /d01<br>(Oracle data file storage) |       | Multiple VMDKs, as required in multiples of 60 GB up to 180 GB per VMDK |       |             |
| Linux /backup_store                              |       | Combined size of Oracle data file storage VMDKs                         |       |             |

## Summary

In this chapter, you learned about virtual machine design.

In the next chapter, you will learn how to design a virtual data center and continue to follow this *customer x* request to make the data center design example.

# 14

## vSphere Virtual Datacenter Design

In the previous chapter, you learned how to design a virtual machine in a vSphere environment. Now we will go through vSphere virtual datacenter design, and focus on discussing how to design a virtual datacenter that is comfortable within a vSphere environment.

In this chapter, you will learn about:

- vSphere virtual datacenter design key points
- A vSphere virtual datacenter design example

### **vSphere virtual datacenter design key points**

In this chapter, we will discuss how to design a vSphere Virtual Datacenter to meet the needs of the business. In general, a Virtual datacenter is the logical boundary of the highest level. A virtual datacenter may be used to delineate separate physical sites/locations or vSphere infrastructures with completely independent purposes. Within vSphere datacenters, allocating compute resources to the virtual machine and protecting the application and services is done by VMware vCenter server, virtual cluster.

So let's design the best virtual datacenter.

## VMware vCenter Server

vCenter Server is a service that acts as a central administration point for ESXi hosts and their virtual machines. It has two editions: Windows and Linux Appliances. Which one do you want to deploy? Let's take a look at the answers to some common questions in the following sections.

### Which platform do you choose?

The Windows edition and VMware just support the installer software. You must install it by yourself every step. You need more than one license for the Windows OS and support for SQL server and Oracle DB. vCenter Server is supported on a few 64-bit Windows operating systems. On a Windows system, it scales well, supporting environments with up to 10,000 powered-on virtual machines.

The Linux-based vCenter Server Appliance is a new management option. The vCenter Server Appliance is made to support Auto Deploy. The appliance comes with the Auto Deploy software installed, as well as its own DHCP server and **Trivial File Transfer Protocol (TFTP)** server. Because it is a new technology, the vCenter Server Appliance has limitations. Certain features, such as vCenter Linked Mode and IPv6, are not yet supported by it. Limited plug-in support is available. Only Oracle is supported as a remote database. In addition, no migration path exists from vCenter Server on a Windows system to a vCenter Server Appliance. But in future, the two editions will small different with each other.

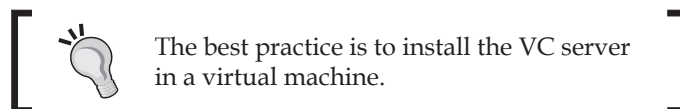


In a demo environment, you can deploy the Linux Appliance easily and show all the functions. In a production environment, you can deploy the Windows edition – the edition that you can fit in your very big environment – and scaling is easy.

### Deploying on a physical server or a virtual machine?

vCenter Server on a Windows system runs well on both a physical and a virtual machine. Many organizations are more comfortable with the management software being separate from the managed environment. So, deploy on the physical server that separates the production environment. However, we need to properly size the physical machine resources, and the resources cannot modify when you have bought the physical server. Purchase a third-party clustering solution to protect vCenter Server, such as a Microsoft cluster.

A virtual machine can protect by vSphere HA or the Data Protector backup software, To improve the virtual machine's flexibility, use vSphere vMotion and DRS. If the virtual machine crashes, you can easily recover the service. You just need to give some minutes to vSphere HA; it is very quick.



The best practice is to install the VC server in a virtual machine.

## How to deploy VMware vCenter Server DB

The Windows edition supports SQL databases and Oracle databases. Linux Appliance can support a remote Oracle database. All of the configuration information of vCenter will be stored in a database, so a good database backup means that vCenter Server is prepared to protect. Each database system in the production environment has the corresponding protection strategy, so the only thing to do in the database system is to create a DB instance for the VC Server to connect.

## vSphere clusters

A VMware cluster is a collection of multiple ESXi hosts, mainly for running the virtual machine in the cluster and application to provide high availability and enhance performance. VMware ESXi hosts are typically organized into clusters. Clusters group similar hosts into a logical unit of virtual resources, enabling technologies such as:

- VMware High Availability (HA)
- VMware vSphere Fault Tolerance (FT)
- VMware vSphere Distributed Resource Scheduler (DRS)

To address customer requirements, several design options were proposed during design workshops. Now let's design the vSphere cluster.

## Number of vSphere clusters

vSphere clusters are created to aggregate hosts. The number of hosts per cluster depends on the number of VMs you are planning to host on it as well as on the type of availability and performance expected. All customer sites vary in size, the types of clusters, and availability requirements. This section will discuss guidelines for establishing these clusters at different sites.

## **A tier one cluster**

This has the following features:

- HA setup to handle failure of two hosts to allow for a host failure during another host maintenance
- Better host hardware specifications to enhance a VM's performance and availability running on it
- Ensure maintaining a low level of overcommitment

## **A tier two cluster**

This has the following features:

- HA setup for handling one host's failure
- A higher level of overcommitment than tier one is allowed

## **A management cluster**

This has the following features:

- A three-host cluster (some customers' sites might require larger management clusters, whereas other sites might not be large enough to justify the cost)
- HA setup for handling failure of one host
- Host only management VMs, including vCenter, vCenter Operation Manager, AD,DNS, printer server SSO, and other management components

The preceding three cluster types will be the dominant types for each site where some sites might require more than one cluster of a certain type, to handle any one of these: a large number of VMs, multiple locations, or compliance. Certain customers' sites will require a separate cluster for certain applications to meet compliance or for special configurations.

The size of the clusters at the customer will vary from site to site, depending on the number of VMs it will need to host and their setup hardware capabilities.



For customer sites where VAAI is not supported, it is recommended that you limit the cluster size so as not to cause VMFS locking contention. It is good practice to cut clusters in a way that will allow them to grow it in the future. It is recommended to start with no more than 16 hosts per cluster so as to allow enough room to grow the cluster as the VMs workload increase. It is also recommended to stop adding VMs to the cluster after 24 hosts to keep enough room for future growth of the currently hosted VMs. So, we design how many clusters will implement. Just focus on the infrastructure limit, application HA and performance demand, and the company management policy.

## vSphere HA

VMware HA is a mechanism within an ESXi cluster that provides high availability for virtual machines. VMware HA continuously monitors the availability of all ESXi hosts in the cluster. If it detects that an ESXi host is not available, it will ensure that the virtual machines that were on that host will be restarted on different hosts in the cluster. This is made possible by the fact that all ESXi hosts in the cluster make use of shared storage. The files of the virtual machines are located on the shared storage. All ESXi hosts can, at all times, these virtual machines and, if necessary, take control of these files. This makes sure that VMware HA can restart a virtual machine on another ESXi host. Also note that all ESXi hosts can supply the same network to a virtual machine.

However, we are designing the vSphere HA cluster. So don't just consider whether the host setting is fit for HA, but also consider the resource and check whether it fits the HA needs.

vCenter HA uses "admission control" to make sure that sufficient resources are available within a cluster to ensure failover. Also, "admission control" ensures that reservations that are issued to virtual machines can be fulfilled in the event of a failover.

In order to guarantee this, HA makes use of an admission control policy. This policy calculates the capacity that should be available at all times within the cluster. If this guarantee cannot be met, the HA cluster will not allow a virtual machine to be started.

The following is a brief description of the admission control policies that are available within vSphere. It is according to this formula: *Production + Failover = Total (cluster resource)*.

| Admission control policy                 | Description                                                                                                                                                                                                                      |
|------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Host failures cluster tolerates          | This policy ensures that there is always sufficient capacity available for the failure of X hosts. Here, X is the number of hosts that can be configured. The slot as one unit to calculate the CPU and memory resource consume. |
| Percentage of cluster resources reserved | The specified percentage of the total number of resources available for a cluster has to be reserved for HA.                                                                                                                     |
| Failover host                            | Here, a specific host is designated to act as a failover host in the event of a failure. This host is active in the cluster, but no virtual machines will be placed on it.                                                       |

## Useful tips

With "host failover", only one host in the cluster can be assigned the failover capacity. For small cluster sizes, this should not be a problem, but if the cluster size increases, the need for more failover capacity will also increase in order to be able to deal with failures.

The "host cluster tolerates failures" admission control policy is able to reserve more host capacity to accommodate HA in the event of a failure. However, the disadvantage of this policy is the way the failover capacity is calculated. The calculation assumes a worst-case scenario that, in most cases, results in the following: admission control no longer allows running virtual machines to the cluster because of the guarantees it wants to give the benefit of HA, while there are still enough resources available. The policy is to use the Slot as a unit to calculate the CPU or memory. We must be very clear of each application to the actual memory and CPU, and then set the slot. So it is very accurate statistics about the number of the virtual machine in the cluster, and clear the current slot can use, has been in use for how many, how much in total. Resource granularity is small. not flexible resources adjustment in future.

The "Percentage of Cluster Resources" admission control policy uses a more flexible way of HA failover resource calculation. Due to the fact that percentages are used in this calculation, it will also grow if the cluster size grows. Hence, the choice was made to use this admission control policy. In most production environments, all the virtual machines have different reservation settings, so we cannot exactly calculate the standard resource consume unit. Just choose that policy.

The last policy for "admission control" is failover host, is specified failover host in cluster, normally don't have any virtual machine is running. Once the machine has crashed, the virtual machine restarts, like the traditional Lord for cluster pattern. This is a waste of resources.

So use this policy just follow the company enforce policy.

## vSphere FT

If VMs in a particular cluster require being FT protected, then an extra distributed switch with two uplinks and an FT port group will need to be added, where VMkernel FT virtual adapters will be attached.

All virtual machines to be protected by FT have only one vCPU and disks configured as eager-zeroed, also called thick provisioned (not thin provisioned). An eager-zeroed thick disk has all of the space allocated and zeroed out at the time of creation. This takes somewhat longer for its creation.

FT with DRS are enabled. This process allows fault-tolerant virtual machines to benefit from better initial placement and to be included in the cluster's load balancing calculations.



Enable the **Enhanced vMotion Compatibility (EVC)** feature. Because a pair of Gigabit Ethernet ports can support four FT-protected virtual machines per host on average, you will need to watch out for the capacity of your FT-enabled clusters. If your environment has the application requirement to protect as use MSCS solution, you can use the FT to protect is simple.

## vSphere DRS

ESXi hosts are bundled together in a cluster under VMware vSphere to make better use of the capacity. A cluster can be seen as a logical collection of resources, which can then be distributed over the virtual machines. Distributed Resource Scheduler (DRS) is a mechanism, together with vMotion technology, that ensures that the available resources of the ESXi hosts in the cluster are effectively distributed over the hosted virtual machines.



DRS is an infrastructure service within vCenter Server. vCenter Server is a requirement for the use of DRS. DRS communicates with all the hosts in the cluster and looks at the resource requirements of various virtual machines hosted within the cluster. It tries to satisfy the resource needs of the virtual machine. If this fails on a particular host, the virtual machine will be moved using vMotion to another host that has more resources available. In this way, DRS tries to ensure that the virtual machines are balanced over the available ESXi hosts and, therefore, the resources are evenly distributed within the cluster.

## DRS automation levels

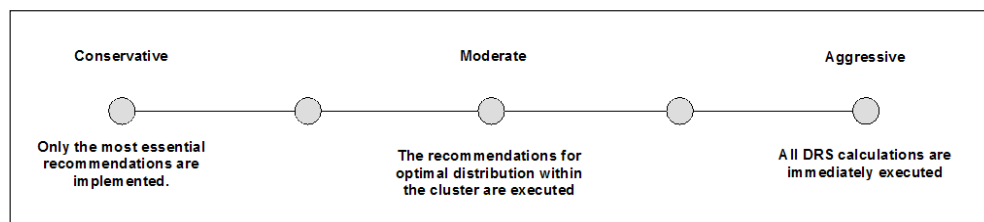
DRS is a feature that can operate autonomously within the ESXi cluster. With the Automation Level, the level of autonomy can be configured for DRS. DRS looks at set intervals at the resources that are used within the cluster. Using resource calculation, it then generates recommendations for moving virtual machines between ESXi hosts. The Automation Level indicates the degree of autonomy that DRS has to execute the recommendations in the cluster.

The following table mentions the automation levels and what they mean:

| Automation level    | Description                                                                                                                                                                                                             |
|---------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Manual              | DRS generates recommendations for the initial placement of virtual machines within the cluster, and generates migration recommendations for virtual machines within the cluster. However, all of this is done manually. |
| Partially automated | DRS makes the initial placement of the virtual machines on the ESXi hosts. Migration recommendations are not executed automatically; they must be implemented manually.                                                 |
| Fully automated     | All recommendations by DRS, for both initial placement and migration, are automatically executed.                                                                                                                       |

Besides the automation level, the migration threshold can also be set. This configuration sets the "expectation for performance gains", which can be achieved by the migration of virtual machines.

The following figure and table show how the migration threshold works:



| Migration threshold   | Description                                                                                                                                                                                                         |
|-----------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Conservative          | DRS will only redistribute virtual machines over ESXi hosts based on cluster requirements, such as maintenance and host affinity rules                                                                              |
| Moderate-conservative | DRS will redistribute virtual machines over those ESXi hosts where the expected performance improvement is at least very good or the observance meets the conditions set for conservative migration                 |
| Moderate              | DRS will redistribute virtual machines over ESXi hosts where the expected performance improvement is at least good or the observance of the observance meets the conditions set for moderate-conservative migration |
| Moderate-aggressive   | DRS will redistribute virtual machines over ESXi hosts where the expected performance improvement is at least reasonable or the observance meets the conditions set for moderate migration                          |
| Aggressive            | All DRS migration recommendations will be executed                                                                                                                                                                  |



Manual execution of the migration recommendations generated by DRS is a very labor-intensive task. Each recommendation should be approved within the cluster. If your production environment has a large number of VMs, it becomes a very labor-extensive task to manage. So, we'd better set the cluster to fully automated, and individual machines can be set to manual or partially automated. The migration threshold should be set to *Moderate*.

## DRS rules

DRS has two types of rules that can be configured alongside policies. Rules dictate that a certain dependency must be met in the cluster, as stated here:

| DRS rule                  | Specification                                                                                                                                                                                                                                                                                                                     |
|---------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| The VM-VM affinity rule   | This is a dependency between two virtual machines. It can be a rule indicating that both virtual machines always have to be on the same ESXi host. This is the so-called affinity rule. But it may also indicate that two virtual machines are never to be placed on the same ESXi host. Then it is called an anti-affinity rule. |
| The VM-host affinity rule | The VM-host affinity rule relates to a virtual machine or group of virtual machines that always need to be hosted on a specific ESXi host or group of ESXi hosts.                                                                                                                                                                 |

The VM-VM affinity rule is set to keep the VM in the same host because of heavy communication, just like a web server that always has to be on the same ESXi host as the database server, because of performance. An anti-affinity rule is set to separate the VM on different hosts because of HA, just similar to the exchange server and database server that never may be hosted on the same ESXi host. All of the chassis must be supplied by the same vendor.


Rack server modularity offers both advantages and disadvantages.

Here are its advantages:

- Supplies more I/O expansion capability
- Easy to heat dissipation
- The rack server can be supplied by a different vendor

The following are its disadvantages:

- Takes up a lot of space, and a lot of cable management is needed
- More cost on DVD and fans

 Now that you know their strengths and weaknesses, the design can be considered according to the demand. But no matter what you choose—blade or rack architecture—it must satisfy the VMware hardware compatibility list.

## Convention for naming a host

Many technical consultants don't pay attention to consider the design scheme of a host name, but host name for the operational management is very important, so please name the host according to the following rule:

Use simple, descriptive, easy-to-understand host names. A standardized naming convention results in easier management, troubleshooting, and use, all of which reduce operational expenses. For example, `esxi-<locationcode>-##>.<domain.name>`, `esxi-shanghai02.test.com`.

## A vSphere virtual datacenter design example

The case background is as follows: A customer need software life cycle management wants to move to version 5 of the VMware vSphere platform. All of vSphere's features will be used to create a flexible, dynamic, and solid virtual infrastructure. This virtual infrastructure will be used as the foundation for the private cloud. The customer is challenged by several business initiatives when it comes to delivering IT services to their customers. The customer needs to create business agility while keeping the IT budget under control – "do more with less." This requires a virtual infrastructure that is flexible and scalable and can grow at the same pace as the business requirements. An architectural design that takes future growth into account needs to be delivered.

Virtualization is specifically designed for creation of an environment that can meet these needs. By using virtualization technology, it is possible for the different layers within the IT infrastructure to work independently. The customer has been doing this for quite some time now, and has built a foundation with their virtual infrastructure.

This allows the **Infrastructure-as-a-Service (IaaS)** to supply the end user organization. IaaS is next to **Software-as-a-Service (SaaS)** and **Platform-as-a-Service (PaaS)**, one of the three layers of the cloud model.

The design is as follows.

vSphere HA cluster.

The "percentage of cluster resources reserved" admission control policy has been selected for this customer. This policy is preferred over the other two admission control policies for flexible implementation of the policy control mechanism.

The customer will get a standard cluster size of 16 ESXi hosts per cluster. The functional requirement is a failure capacity reservation that is a minimum of two ESXi hosts per cluster. For the standard cluster size, that means 12.5 percent. However, it is not possible within vCenter to configure half a percentage, so there will be a standard of 13%, which is specified as the admission control policy percentage configuration.

The HA configuration:

| Attribute                | Specification                       |
|--------------------------|-------------------------------------|
| Enable host monitoring   | Yes                                 |
| Admission control        | Enable                              |
| Admission control policy | The percentage of cluster resources |
| Percentage HA failover   | 13 percent                          |
| VM restart priority      | Medium                              |
| Isolation response       | Leave the power on                  |
| VM monitoring            | VM monitoring only                  |
| Monitoring sensitivity   | High                                |

vSphere DRS cluster.

The following table lists the DRS configuration:

| Attribute               | Specification   |
|-------------------------|-----------------|
| Automation level        | Fully automated |
| Migration threshold     | Moderate        |
| Virtual machine options | Enabled         |

## Summary

In this chapter, you learned how to design a vSphere virtual data center. When you have all of the knowledge about this, you can continue to learn much more about the new IT infrastructure, SDDC.

# Index

## A

**AMD Virtualization (AMD-V)** 59

**automation levels, vSphere DRS**

- fully automated 216
- manual 216
- partial automated 216

## B

**blade server**

- advantages 195
- disadvantages 195

## C

**Common Internet File System (CIFS)** 171

**compatibility modes, RDM**

- physical mode 188
- virtual mode 188

**CPU hardware virtualization**

- advantages 57
- disadvantages 57

**CPU software virtualization**

- advantages 57
- disadvantages 57

## D

**datastore types, vSphere storage design**

- about 184
- desktop VM 184
- ESXi multipathing policies 188
- isolated VM 184
- IT VM 184
- masking 191
- production VM 184

RDM 187

rules 184

Shared I/O Control settings 185

SRM placeholder 184

Staging VM 184

Storage I/O Control 185

Template and ISO 184

zoning 191

**design key points, ESXi host**

about 193

CPU capacity 193, 194

hardware type 195, 196

host count 194

**design key points, virtual machine**

about 201

limit 203

memory's performance, ensuring 202

reservation 203

resource, setting 203

shares 203

virtual CPUs, count 201, 202

virtual SCSI HBA type 205

**device driver queue** 98

**Direct Console User Interface (DCUI)** 30

**Distributed Resource Scheduler (DRS)** 87

## E

**End User License Agreement (EULA)** 5

**Enhanced vMotion Compatibility (EVC)** 215

**esxcfg-mapth command**

options 136, 137

**esxcfg-scsidevs command**

options 138

**esxcli**  
multipathing 136-142

**ESXi host**  
about 193  
design, example 196, 197  
design key points 193  
naming convention 196  
reviewing 44

**ESXi host design**  
example 196, 197  
management cluster specifications 198  
Oracle cluster specifications 198  
Oracle cluster VMware ESXi physical specifications 199, 200  
VMware ESXi physical specifications 199

**ESXi multipathing policies**  
designing 188

**ESXi Shell**  
accessing 30-32  
accessing, from DCUI 33  
accessing, from SSH client 34  
enabling 30-32  
enabling, from vSphere Client 32, 33  
vm-support log file, exporting 48, 49

**ESXi technical support mode**  
configuring 30

**esxtop** 61

**Extended Page Tables (EPT)** 61

## **F**

**FC-SAN multipathing** 189  
**Fibre Channel over Ethernet (FCoE)** 97

## **G**

**Guest Average Latency (GAVG)** 156  
**Guest Physical Memory** 60  
**Guest Virtual Memory** 60

## **H**

**hard disk, virtual machine**  
deploying 203, 204  
disk location 204  
multiple virtual disks 204  
swap file location 204, 205  
**hardware abstraction layer (HAL)** 202

**hardware acceleration API**  
for block 72  
for NAS 72

**hardware MMU virtualization**  
advantages 61  
disadvantages 61

**hardware virtualization**  
CPU 57-60  
MMU 61  
techniques 57-60

**High Availability (HA)** 87

**Host Bus Adapter (HBA)** 71

**Host Physical Memory** 60

## **I**

**Infrastructure-as-a-Service (IaaS)** 219  
**input/output operations per second (IOPS)** 180

**installations**  
VMware ESXi host 4-8  
VMware vCenter Server 9-15  
VMware vCenter Server Appliance 17

**Integrated Lights-Out (iLO)** 33

**Intel Virtualization Technology (Intel VT-x)** 59

**Internet Small Computer System Interface (iSCSI)** 163

## **K**

**Kernel Average Latency (KAVG)** 156  
**kernel queue** 98

## **L**

**LUN masking**  
about 119-125  
implementing 192  
options 120, 121

## **M**

**masking** 191  
**memory management unit (MMU)** 57  
**migration threshold**  
conservative 217  
moderate 217

- moderate-aggressive 217
- moderate-conservative 217
- working 216, 217

**Most Recently Used (MRU) 84, 188**

**Multipathing Plugin (MPP)**

- about 82
- Path Selection Plugin (PSP) 82
- Storage Array Type Plugin (SATP) 82

## N

**naming convention, ESXi host 196**

**Native Multipathing Plugin (NMP) 82**

**Network File Storage (NFS)**

- about 171
- case study 172-174
- component 171, 172
- troubleshooting example 175, 176

## P

**physical architecture**

- differentiating, with virtual architecture 2

**physical machine**

- comparing, with virtual machines 3

**Platform-as-a-Service (PaaS) 219**

**Pluggable Storage Architecture (PSA)**

- about 82, 117
- analyzing 136-142

**PowerPath/VE 189**

## Q

**queue depth 98**

## R

**rack server modularity**

- advantages 195, 218
- disadvantages 196, 218

**Rapid Virtualization Indexing (RVI) 61**

**Raw Device Mapping (RDM)**

- about 187
- compatibility modes 188

**resxtop utility**

- batch mode 65

- default mode (interactive mode) 65
- replay mode 65

**Round Robin (RR) 84**

**rules, vSphere DRS**

- VM-host affinity rule 217
- VM-VM affinity rule 217

## S

**service-level agreement (SLA) 181**

**settings, Storage I/O Control**

- Limit IOPs 185
- Storage I/O enabled 185
- Storage I/O shares 185

**Software-as-a-Service (SaaS) 219**

**software MMU virtualization**

- advantages 61
- disadvantages 61

**software virtualization**

- CPU 57-61
- MMU 61
- techniques 57-61

**SSD devices**

- identifying 149-152
- tagging 149-152

**storage design**

- key points 179, 180

**Storage I/O Control**

- about 185
- settings 185

**storage performance problems, vSphere**

- scenarios 112-114
- troubleshooting 111

**storage virtualization**

- concepts 97-101

**symmetric multiprocessing (SMP) 201**

## T

**Technical Supported Mode (TSM) 40**

**translation look-aside buffer (TLB) 60**

**Transport Control Protocol (TCP) 163**

**Trivial File Transfer Protocol (TFTP) 210**

**troubleshooting example,**

- NFS storage 175, 176



## V

### **vCenter Server**

- connecting, vSphere Client used 21
- connecting, vSphere Web Client used 21
- logging, options 134
- logs 45

### **vCenter Single Sign On (SSO) 11**

### **vCenter troubleshooting FC storage**

- using 157-161

### **vifp interface 34**

### **virtual architecture**

- differentiating, with physical architecture 2

### **virtual disk**

- in RDM 100
- in VMFS 100

### **virtual machine**

- comparing, with physical machine 3
- design, example 207, 208
- design key points 201
- hard disk, deploying 203
- hard disk, types 203
- storage profile 74-86

### **virtual machine monitor. See VMM**

### **virtual SCSI HBA type**

- about 205
- guest OS 207
- hardware compatibility 206
- virtual NICs 205

### **vMA**

- about 23
- connection options 36
- deploying 23-30
- IP address allocation policy 28
- management commands 35, 36
- Thick Provision Eager Zeroed option 26
- Thick Provision Lazy Zeroed option 26
- Thin Provision option 27
- URL 24

### **VMFS**

- about 100
- datastore volume, unmounting 148
- resignaturing 142-148
- virtual disk, accessing 100
- volume copies, applying 142-145

### **vmhba3 adapter 121**

### **vmkping command 176**

### **VMM**

- about 51-55
- monitor modes 56

### **vm-support log file**

- exporting, from ESXi Shell 48, 49
- exporting, from vSphere Client 46-48

### **VMware**

- recommended paths 189, 190
- URL 126

### **VMware commands**

- connection options 36
- esxcfg- commands 35
- esxcli 35
- management commands 35
- resxtp 35
- using 34-44
- vicfg- commands 35
- vmware-vmd 35

### **VMware ESXi host**

- installing 4-8

### **VMware snapshots**

- troubleshooting 146-148

### **VMware vCenter Server**

- about 210
- installing 9
- platform, selecting 210
- prerequisites 9-16
- URL 9

### **VMware vCenter Server Appliance**

- about 17
- installing 17
- prerequisites 17-20
- URL 17

### **VMware vCenter Server DB**

- deploying 211

### **VMware vSphere ESXi**

- architecture 51, 52
- troubleshooting 53-55

### **VMware vSphere Storage DRS (SDRS)**

- about 87
- creating 88-92
- VM anti-affinity rules 88
- VMDK affinity rules 87
- VMDK anti-affinity rules 88

### **VMware vSphere Storage I/O Control 92-94**

### **VPX Operational Dashboard (VOD) 45**

**vSphere**

- log, locations 45, 46
- storage architecture 180, 181
- storage performance problems, troubleshooting 111

**vSphere Client**

- used, for connecting vCenter Server 21
- vm-support log file, exporting 46-48

**vSphere clusters**

- about 211
- management cluster 213
- tier one cluster 212

**vSphere Distributed Switch (VDS) 172****vSphere DRS**

- about 215
- automation levels 216
- rules 217, 218

**vSphere Fibre Channel storage**

- about 153, 154
- troubleshooting example 154-156

**vSphere FT 215****vSphere HA**

- about 213
- useful tips 214

**vSphere iSCSI storage**

- components 163, 164
- troubleshooting example 165-169

**vSphere log**

- used, for troubleshooting vSphere storage problem 126-133

**vSphere Management Assistant. *See* vMA****vSphere performance monitoring tools**

- characters 67
- using 61-69

**vSphere Standard Switch (VSS) 172****vSphere storage**

- components 117-119
- issue, troubleshooting 126-134
- managing, with command line 105-111
- monitoring 101-105

**vSphere storage APIs**

- for array integration 71-74
- for storage awareness 71-74

**vSphere storage design**

- about 181
- datastore types 184
- share storage size 182

**vSphere virtual datacenter design**

- example 219, 220
- key points 209

**vSphere Web Client**

- used, for connecting vCenter Server 21

**vStorage APIs, for Array Integration**

- about 186
- hardware offloaded copy up 186
- hardware offloaded locking SCSI reservation conflicts 186
- Write same / Zero 10 operation 186

**W****World Wide Number (WWN) 105****Z****zoning**

- about 191
- tips 192





## Thank you for buying **Mastering VMware vSphere Storage**

### **About Packt Publishing**

Packt, pronounced 'packed', published its first book, *Mastering phpMyAdmin for Effective MySQL Management*, in April 2004, and subsequently continued to specialize in publishing highly focused books on specific technologies and solutions.

Our books and publications share the experiences of your fellow IT professionals in adapting and customizing today's systems, applications, and frameworks. Our solution-based books give you the knowledge and power to customize the software and technologies you're using to get the job done. Packt books are more specific and less general than the IT books you have seen in the past. Our unique business model allows us to bring you more focused information, giving you more of what you need to know, and less of what you don't.

Packt is a modern yet unique publishing company that focuses on producing quality, cutting-edge books for communities of developers, administrators, and newbies alike. For more information, please visit our website at [www.packtpub.com](http://www.packtpub.com).

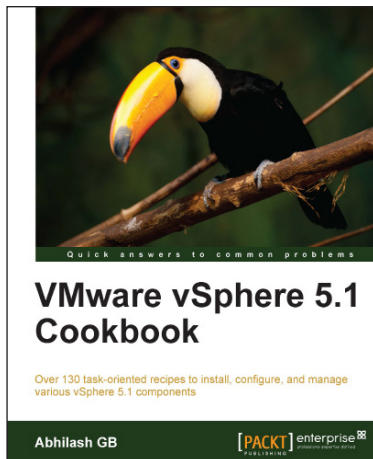
### **About Packt Enterprise**

In 2010, Packt launched two new brands, Packt Enterprise and Packt Open Source, in order to continue its focus on specialization. This book is part of the Packt Enterprise brand, home to books published on enterprise software – software created by major vendors, including (but not limited to) IBM, Microsoft, and Oracle, often for use in other corporations. Its titles will offer information relevant to a range of users of this software, including administrators, developers, architects, and end users.

### **Writing for Packt**

We welcome all inquiries from people who are interested in authoring. Book proposals should be sent to [author@packtpub.com](mailto:author@packtpub.com). If your book idea is still at an early stage and you would like to discuss it first before writing a formal book proposal, then please contact us; one of our commissioning editors will get in touch with you.

We're not just looking for published authors; if you have strong technical skills but no writing experience, our experienced editors can help you develop a writing career, or simply get some additional reward for your expertise.



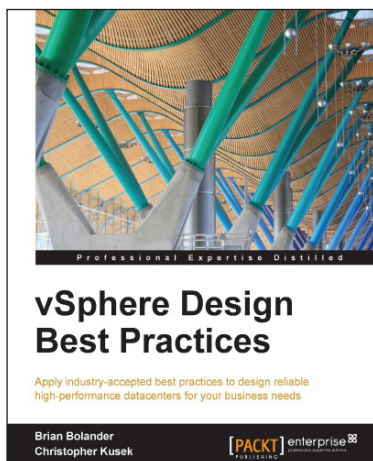
## VMware vSphere 5.1 Cookbook

ISBN: 978-1-84968-402-6

Paperback: 466 pages

Over 130 task-oriented recipes to install, configure, and manage various vSphere 5.1 components

1. Install and configure vSphere 5.1 core components.
2. Learn important aspects of vSphere such as administration, security, and performance.
3. Configure vSphere Management Assistant(VMA) to run commands/scripts without the need to authenticate every attempt.



## vSphere Design Best Practices

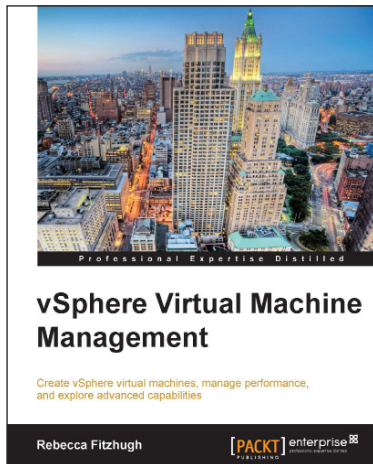
ISBN: 978-1-78217-626-8

Paperback: 126 pages

Apply industry-accepted best practices to design reliable high-performance datacenters for your business needs

1. Learn how to utilize the robust features of VMware to design, architect, and operate a virtual infrastructure using the VMware vSphere platform.
2. Customize your vSphere Infrastructure to fit your business needs with specific use-cases for live production environments.
3. Explore the vast opportunities available to fully leverage your virtualization infrastructure.

Please check [www.PacktPub.com](http://www.PacktPub.com) for information on our titles



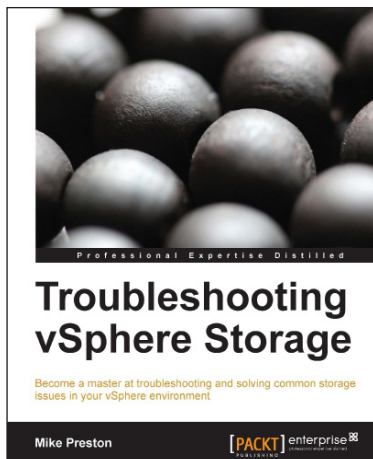
## **vSphere Virtual Machine Management**

ISBN: 978-1-78217-218-5

Paperback: 326 pages

Create vSphere virtual machines, manage performance, and explore advanced capabilities

1. Create virtual machines using the wizard, cloning, deploying from a template, and using OVF templates.
2. Manage multi-tiered applications using vApps.
3. Learn how to optimise virtual machine performance and resource allocation.



## **Troubleshooting vSphere Storage**

ISBN: 978-1-78217-206-2

Paperback: 150 pages

Become a master at troubleshooting and solving common storage issues in your vSphere environment

1. Identify key issues that affect vSphere storage visibility, performance, and capacity.
2. Comprehend the storage metrics and statistics that are collected in vSphere.
3. Get acquainted with the many vSphere features that can proactively protect your environment.

Please check [www.PacktPub.com](http://www.PacktPub.com) for information on our titles

