USENIX®

THE ADVANCED COMPUTING SYSTEMS ASSOCIATION

ISBN-13: 978-1-931971-72-0

90000

9 781931 971720

21

John Arrasjid, Duncan Epping, and Steve Kaplan

Foundation for Cloud Computing with VMware vSphere 4

21

Short Topics in
System Administration

Jane-Ellen Long, Series Editor

Foundation for Cloud Computing with VMware vSphere 4

John Arrasjid,
Duncan Epping, and
Steve Kaplan

USENIX®

THE ADVANCED COMPUTING SYSTEMS ASSOCIATION

21 *Short Topics in*
System Administration

Jane-Ellen Long, Series Editor

Foundation for Cloud Computing with VMware vSphere 4

John Arrasjid, Duncan Epping, and Steve Kaplan

Published by the USENIX Association
2010

ISBN 978-1-931971-72-0

To purchase additional copies, see http://www.sage.org/pubs/short_topics.html.

The USENIX Association
2560 Ninth Street, Suite 215
Berkeley, CA USA 94710

http://www.usenix.org/

Contents

Figures and Tables

Figures

Table

Acknowledgments

The authors work for VMware and INXi. The opinions expressed here are the authors' personal opinions. Content published was not read or approved in advance by VMware and does not necessarily reflect the views and opinions of VMware. This is the authors' book, not a VMware book.

Thank you to my good friends Duncan Epping and Steve Kaplan for making this book an enabler for virtualization advocates and those interested in learning more about VMware virtualization. Thank you to Daniel Conde for material reused and updated for this edition. Thank you to my VMware management team (Peter Giordano, Director; Enis Konuk, Senior VP; Steve Herrod, CTO) for supporting me on this and other projects.

Thank you for other contributors and reviewers of this work. These include Phil Anthony, Ryan K. Baker, David Baldwin, Matthew S. Barnes, Mark Burgess, Philip Callahan, Ken Cline, Richard Damoser, Girish Manmadkar, Andrew Mitchell, Craig Risinger, Rupen Sheth, Lee Dilworth, Ben Lin, and Ken Brady.

Thank you to Jane-Ellen Long for her editorial support in making this a reality. Thank you to the USENIX Association for inviting me for this second VMware Infrastructure Short Topics booklet and for inviting me to participate in the virtualization track at USENIX/LISA conferences since 2005. USENIX continues to hold a high standard for both the research and administration aspects of system administration.

A special thanks to my family and friends for support of this project and the time spent.

I dedicate this book to Amy, Catherine, Sofi, Lila, Mom, and Dad, whose love and support have gotten me through the weekends, nights, and holidays to complete this project.

This book is also dedicated to virtualization advocates around the world!

John Y. Arrasjid
VCDX, VMware, Inc.

Foreword

If you're reading this, you're likely already taking part in the IT revolution that is driven by server and desktop virtualization. The revolution has been unfolding since VMware was founded in 1998, but its speed has clearly accelerated over just the past two to three years. Helping people keep up with this rapid pace of change is what makes education, training, and reference materials such as this book so valuable.

A lot has changed since the first edition of *Deploying the VMware Infrastructure* was published in 2008, most notably the release of VMware vSphere 4 in May of 2009. VMware vSphere was the most ambitious software product release ever undertaken at our company, and it certainly ranks among the biggest releases in our industry as a whole. From a numbers standpoint, the development of vSphere encompassed more than 3,000,000 engineering hours by well over 1,000 talented engineers over a three-year period. And from a technical standpoint, vSphere's new capabilities clearly mark the arrival of the fourth generation of virtualization.

As the first "cloud computing operating system," vSphere was designed to help transform IT-owned and operated datacenters into efficient and safe *private clouds*. VMware vSphere was also designed for service providers, hosters, and other companies looking to transform their datacenters into enterprise-friendly *public clouds*. Lastly, this release creates the foundation for connecting these different cloud types into a *hybrid cloud* that is based on open standards and that provides the flexibility needed to truly enable the delivery of IT as an efficient and customer-friendly service.

While VMware vSphere includes more than 150 new features when compared to Virtual Infrastructure 3.5, it is the significant advancements in performance, scalability, availability, security, and management that are particularly exciting. For example, vSphere includes the ability to apply security and policies that follow virtual machines as they migrate across the network, obviating previous regulatory and organizational challenges with virtualizing certain tier-one applications. Taken together, these advances give organizations the capability and confidence to virtualize 100% of their datacenters.

Furthermore, a rich set of APIs have enabled hundreds of other companies to integrate their products directly into vSphere. The industry leaders in compute, storage, security, and networking are taking advantage of these APIs by creating solutions that go far beyond the capabilities possible in a physical infrastructure. At the same time, a large number of start-ups and smaller companies are creating entirely new products and offerings that exploit the many capabilities that VMware vSphere delivers.

vCenter Orchestrator, and Site Recovery Manager help automate IT processes, while vCenter Chargeback assists in the organizational transformation of IT infrastructure to a service. The just-released VMware View 4 also builds upon vSphere to deliver the most efficient virtual desktop infrastructure on the market.

A lot of important considerations go into designing and optimizing a virtualized datacenter. A business plan typically must first be created to sell the use case to senior management. A comprehensive architectural design follows, encompassing not only the deployment and configuration of vSphere, but also optimizing compute, storage, and network resources. Effective implementation of security, management, and automation tools are essential to a successfully virtualized datacenter.

In *Foundation for Cloud Computing with vSphere 4*, authors Arrasjid, Epping, and Kaplan met the challenge of covering these topics in under 120 pages. It is an easy and valuable read whether for virtualization novices or experienced IT consultants with years of ESX experience.

I'll close by wishing you pleasant reading and a great experience on your virtualization journey.

Dr. Stephen Alan Herrod
Palo Alto, California
December 2009

1. Introduction

What is virtualization? How does it benefit my organization? What is VMware vSphere? What other technologies are used to build a virtual infrastructure? These questions are answered in this Short Topics book along with use cases for performance, optimization, and return on investment (ROI) for deploying VMware Virtual Infrastructure.
This book is a follow-on to volume 18 of the USENIX Short Topics in System Administration book *Deploying the VMware Infrastructure*. While the previous work covered VI 3.5 and related technologies, this new edition covers vSphere 4 and add-on technologies that can be used for a full Service-Oriented Architecture based on the Virtual Datacenter.

Virtualization applies to many different areas in the computer world, including graphics, sound, and computing. This book focuses on the server virtualization space. Server virtualization allows multiple operating systems to run concurrently on the same hardware by logically partitioning the hardware and presenting a standardized set of resources and devices to the running operating systems. VMware has extended server virtualization solutions to include management capabilities and tools for tasks such as server provisioning and live migration of virtual machines and virtual machine disk files.

Virtualization benefits the datacenter by reducing hardware and infrastructure costs, reducing power and cooling costs, increasing utilization of hardware, and simplifying provisioning and budgeting processes. Virtualization also provides the underpinning technology for cloud computing. VMware has additional technologies that can extend your datacenter further using this Virtual Datacenter with various service catalog items, including chargeback, capacity planning, security, business resiliency, application performance, resource management, and lifecycle automation.

Virtualization is recognized as a foundation for Cloud Computing. Most of the existing VMware technologies will fit into Cloud architecture design. VMware offers recommendations on approaching the Cloud categories of Infrastructure-as-a-Service, Platform-as-a-Service, and Software-as-a-Service. Other Cloud categories such as DR-as-a-Service and Desktop-as-a-Service are extensions to the base three. This book is an overview of the VMware technologies and how they can support the various services and management pieces required for Cloud architecture. Although this is not a deep dive into the specific design patterns, it provides insight into the tools to fit your design criteria.

VMware Technology Overview

VMware was founded in 1998 with the goal of putting mainframe-level virtualization technology and the associated resource partitioning capabilities on an x86 platform. VMware software provides hardware virtualization capabilities that present an x86/ x64 platform and associated devices to a guest operating system running in a virtual machine.

The suite of products from VMware includes virtualization platforms to run virtual machines along with migration and conversion tools, assessment tools, and management tools to support the VMware Virtual Infrastructure. For interoperability information, please check the relevant software and hardware compatibility lists for each product. This suite has the technologies and associated products described below.

Hosted Virtualization Software

Hosted virtualization software runs on top of a standard operating system. Several of these technologies provide support for Open Virtualization Format (OVF).

- ❖ **VMware Workstation**—Desktop virtualization product designed for end-users and developers to create and run virtual machines on Windows- or Linux-based systems.
- ❖ **VMware Player**—Free virtualization product for running (but not creating) multiple virtual machines on Windows or Linux systems.
- ❖ **VMware Server**—Free entry-level server virtualization product for creating and running multiple virtual machines on existing physical Windows or Linux servers (formerly GSX Server).
- ❖ **VMware Fusion**—Virtualization product for Intel-based Mac OS X systems.
- ❖ **VMware ACE**—Virtualization product for enterprise desktop deployments, providing a highly configurable, secure, and portable PC environment.

Native Virtualization Software

- ❖ **VMware ESX/ESXi**—ESX and ESXi are both hypervisors which install directly on the server hardware. Although the deployment and management methods are slightly different, both solutions provide maximum performance and availability. Classic ESX installs with a Linux-based Service Console to assist with management functions. ESXi removes the Service Console, reducing the attach surface due to a smaller footprint and allowing the functionality to be embedded within the server hardware.

Virtualization Management Software and Scalability

VMware vCenter Server manages all components of VMware vSphere, spanning multiple clusters and datacenters through one centralized interface. The following virtualization tools are managed through VMware vCenter Server:

- **VMware Virtual SMP**—Enables multiprocessor virtual machines.
- **VMware VMotion**—Enables live migration of virtual machines from one physical server to another with no impact on end-users and without rebooting or changing device drivers. Regardless of the underlying hardware, a migration between different ESX host hardware will not impact the operating system and its applications.
- **VMware Storage VMotion**—Enables live migration of virtual machine disk files across storage locations while maintaining service availability.
- **VMware Distributed Resource Scheduler (DRS)**—Dynamically allocates and balances workloads across hosts in a cluster.
- **VMware Distributed Power Management (DPM)**—Dynamically starts up and shuts down ESX host hardware to reduce power consumption.
- **VMware High Availability (HA)**—Provides automated restart of failed virtual machines, regardless of the guest operating system technology. See Business Continuity and Disaster Recovery, below.
- **VMware Fault Tolerance (FT)**—Enables active clustering of virtual machines without requiring applications to be cluster aware. See Business Continuity and Disaster Recovery, below.
- **Hot add/plug/extend of devices**—Hot add of CPU and memory, hot plug of virtual storage and network devices, and hot extend of virtual disks have been included. This provides the ability to add/plug/extend virtual machine resources without disruption or downtime. The associated guest OS requires support for this feature.
- **VMware Host Profiles**—Enables the definition and application of standardized host configurations. Also supports compliance checks against the defined standards.
- **vSphere Management Assistant (vMA)**—Linux appliance with pre-built management tools and the vCLI interface.
- **vNetwork Distributed Switch (vDS)**—Provides a switch that acts at a datacenter level across multiple ESX hosts, which offers centralized provisioning, administration, and monitoring. Simplifies network management by moving the virtual network configuration and management from the host level to the datacenter level.
- **vNetwork Standard Switch (vSS)**—A software program emulating a physical switch to enable one virtual machine to communicate with another. It is a basic Layer 2 switch without routing.

Migration Tools

The following technologies allow the migration to a VMware Virtual Infrastructure. Support for Open Virtualization Format (OVF) imports is included. Please refer to the respective product documentation for how to import and use OVF systems.

- **VMware Guided Consolidation**—Used for planning physical-to-virtual machine migrations by utilizing VMware Capacity Planner Converter technology. VMware Guided Consolidation is an optional vCenter component and is designed for small-scale consolidation.
- **VMware Converter**—Used for physical-to-virtual machine migrations, as well as importing virtual machines from other virtualization vendors. VMware Converter can import multiple machines concurrently and non-disruptively. Designed for large-scale consolidation, VMware Converter can be used with or without VMware vCenter Server.

Security Enablers

The following provide tools to support security requirements and regulatory compliance guidelines for a company. They also have an important role in deploying Cloud architectures, including Private Cloud, Public Cloud, and Hybrid Cloud.

- **VMware ACE**—See Hosted Virtualization Software, above.
- **VMware VMsafe**—Provides an open approach to security through an application program interface (API). This enables selected partners to develop security products for VMware environments. VMsafe gives fine-grained visibility over virtual machine resources, making it possible to monitor every aspect of the execution of the system and stop previously undetectable viruses, rootkits, and malware before they can infect a system. VMsafe provides inspection of virtual machine memory pages and CPU states; filtering of network packets inside hypervisors as well as within the virtual machine itself; and in-guest, in-process APIs that enable complete monitoring and control of process execution. Guest virtual machine disk files can be mounted, manipulated, and modified as they persist on storage devices.
- **VMware vShield Zones**—Enforces corporate security policies at the application level in a shared environment, while still maintaining trust and network segmentation of users and sensitive data. Provides a mechanism to monitor, log, and block inter-VM traffic with an ESX/ESXi host or between hosts in a cluster. This includes the ability to firewall, bridge, or isolate virtual machines between multiple pre-defined zones. All activities, blocked as well as allowed, are logged and can be graphed.

Desktop Virtualization Software

Desktop virtualization is an important and cost-saving advantage for companies and fits the Cloud concept of "access everywhere." Access from a personal computer, laptop, smartphone, or thin client makes this possible.

- **VMware View**—A system for managing connectivity, security, and administration of centralized virtual desktop computers hosted on ESX clusters. VMware View Manager supports the connection brokering for the virtual desktop infrastructure (VDI), while View Composer provides advanced desktop image management.

Application Virtualization Software

VMware ThinApp is another technology that fits Cloud deployments for a company. It permits highly portable applications that can be easily deployed within a Cloud. ThinApp simplifies the upgrade path of applications deployed in a Cloud architecture.

- **VMware ThinApp**—An application virtualization platform that enables complex software to be delivered as self-contained, executable (EXE) files which can run instantly with zero installation from any data source. The core of the technology is the Virtual Operating System, a small, lightweight component embedded with each ThinApp-compiled application. Applications packaged by ThinApp require a compatible MS Windows operating system to run.

Capacity Management/Assessment

Cloud architectures rely on capacity planning, financial accountability (charge-back or show-back of costs), and elasticity dynamism. The tools included here are designed to fill this Cloud requirement.

- **VMware Capacity Planner**—An agentless data collection and "what if" scenario building tool that identifies server inventories and resource utilization to determine virtual machine candidates, server consolidation ratios, and resource requirements for migrating to a VMware Infrastructure based on target ESX host platform resources.
- **VMware vCenter CapacityIQ**—Identifies server resource inventories including used and unused capacity. This can be used for capacity planning, budgeting, and lifecycle management of resources. VMware CapacityIQ is used for cost avoidance and justification, availability and risk mitigation, and project planning and decision-making.
- **VMware vCenter Chargeback**—Provides cost measurement, analysis and reporting to provide cost transparency and accountability for the virtual machines and the supporting virtual infrastructure. IT costs may be mapped to business units, cost center, or external customers to provide a better understanding of resource costs. This can further be used to determine optimization for cost reduction.

Software Lifecycle Automation

Software Lifecycle Automation includes tools that allow workflow management, lifecycle management, and dynamic deployment of machines. These are fundamental tools that are a starting point of technologies supporting Cloud architecture.

- **VMware vCenter Orchestrator**—For a description, see Workflow Management, below. This tool can be used for software lifecycle management.
- **VMware vCenter Lab Manager**—Provides a self-service portal for real-time provisioning, managing, and collaboration of virtualized development and testing environments. VMware vCenter Lab Manager allows developers and testers to create and share libraries of virtualized application environments used in software development and testing. Applications can be moved through lifecycle stages until they reach production state.
- **VMware vCenter Lifecycle Manager**—Manages the lifecycle of virtual machines from request through provisioning and eventual archiving or destruction. VMware vCenter Lifecycle Manager provides a self-service portal for virtual machine requests, routed through a predefined workflow, streamlining provisioning, reducing overhead, and providing consistent management of the virtual machine lifecycle.

Workflow Management

Workflow management tools support task automation. vCenter Orchestrator used with the vCloud API can work together in deploying a Private Cloud.

- **VMware vCenter Orchestrator**—Provides out-of-the-box workflows to help automate existing manual tasks. Workflows can be created, modified and extended to meet custom needs.
- **vSphere PowerCLI**—See Command Line Interfaces, below.

Business Continuity and Disaster Recovery

Multiple technologies are available for supporting both Business Continuity (strategic) and Disaster Recovery (tactical). All can be used to support various Cloud deployments for both the Enterprise and a Cloud Provider.

- **VMware vCenter Site Recovery Manager (SRM)—**Provides disaster recovery workflow automation through a centralized management interface. SRM automates the setup, testing, failover, and failback of virtual infrastructures between protected and recovery sites.
- **VMware High Availability (HA)**—Provides fault tolerance in the event of an ESX failure or a VM operating system failure. VMware HA enables the automated restart of virtual machines on other hosts in a cluster upon host failure, minimizing downtime without the cost of application clustering.
- **VMware Fault Tolerance (FT)**—Provides clustering support of single vCPU VMs without requiring the embedded application to be cluster aware. FT utilizes VMware vLockstep technology. This technology uses an active secondary VM that runs in virtual lockstep with the primary VM. VMware vLockstep establishes and maintains this secondary VM. The secondary VM runs on a different host and executes the same set of instructions, in the same sequence, as the primary VM.

- **VMware Consolidated Backup (VCB)**—Provides the capability to perform SAN-based backup and recovery of virtual machines using a backup proxy server without any network or virtual machine overhead.
- **VMware Data Recovery (vDR)**—Provides a backup solution for virtual machines for smaller sites. VMware Data Recovery is fully integrated with vCenter Server and includes data de-duplication to save on disk storage for full virtual machine backups. Includes file level restore or entire images as needed.
- **vCenter Heartbeat**—Protects the vCenter Server, License Server, and Database against hardware, OS, application, and network downtime. Failover and failback are provided for each. Protection is important especially when using VMware View, vCenter Lab Manager, and vCenter SRM, which require vCenter to be running at all times.

Application Management, Analysis, and Performance

The tools listed here can be used for benchmarking (VMmark), application performance monitoring and latency analysis (VMware AppSpeed), and application performance Service Level Agreements (vApp). Enterprise Clouds and Cloud Providers can benefit from these tools to take proactive steps to manage application performance.

- **VMmark**—A benchmark tool specifically designed for measuring scalability of virtualization host systems. Provides an accurate measurement of application performance in virtualized environments. Measures virtual machine performance, determines how different hardware and virtualization platforms will affect performance, and enables "best fit" choices for hardware. VMware is working with the Standard Performance Evaluation Corporation (SPEC®) and members of the SPEC Virtualization subcommittee to develop standard methods of comparing virtualization performance for virtualized applications running on hypervisors.
- **VMware vCenter AppSpeed**—An application performance monitoring tool engineered specifically for multi-tiered applications. AppSpeed passively listens to traffic flowing over a vSwitch (including the Nexus 1000V), which permits discovery of transactions, application mapping, performance monitoring against SLAs, and root cause analysis. Provides a method to evaluate performance of an application before and after virtualization to ensure that performance remains consistent. This tool offers breadth in latency analysis for an application.
- **VMware (SpringSource) Hyperic HQ**—Provides complete discovery, monitoring, analysis, and control of all application, system, and network assets both inside and outside the virtual machines. Hyperic HQ includes full VMware ESX and VMware Server support, analysis of utilization and performance within a VM, correlation of events between hosts and guest OSes, and control of VMs. This tool gives detailed analysis of how the virtual machine is performing and depth in latency analysis for an application.

- **vApp**—Provides a logical entity, or object, comprising one or more virtual machines using the OVF (Open Virtualization Format) to specify and encapsulate all components of a multi-tier application. In addition, policies and SLAs can be associated with the object as an attribute. The vApp construct is designed for interoperability of a multi-tiered application on the virtual datacenter as well as the ability to move the application between internal or external clouds while maintaining the same SLAs.

Application Programming Interfaces (APIs)

- **vCloud API**—Supplies an interface for providing and consuming virtual resources within a VMware-based cloud by enabling deployment and management of virtualized workloads by working with vApps. This API is based on OVF standards providing platform independence and multi-tenancy in a purely virtual infrastructure. Includes functions for Inventory Listing, Catalog Management, Upload/Download/Provisioning Operations, vApp Configuration Operations, Resource Entities Operations, vApp State Operations, and other operations. Also includes administrative functions, including Cloud, Org, vDC, Catalog, User, Group, and Role Administration.
- **vStorage API**—Provides integration of advanced capabilities from storage vendors with the vSphere Cloud OS from VMware. This API enables customers to leverage array-based capabilities such as support for multi-pathing control, which enables advanced load balancing algorithms.
- **vNetwork API**—Provides integration with the virtual networking capabilities of vSphere to enable the development of advanced network tools.
- **VMsafe API**—Allows vendors to develop advanced security products.
- **CIM Interfaces**—Designed for hardware management tool development.
  - **Server Management API**—CIM SMASH interface to monitor and manage virtualization server platforms.
  - **Storage Management API**—CIM SIMI-S interface to monitor and manage virtual storage.
- **VMware Orchestrator API**—Allows for the programming of workflows for execution by VMware Orchestrator.
- **VMware vCenter Site Recovery Manager API**—Provides an interface to SRM, which allows external management systems to initiate tests or failovers and record results.
- **vCenter Chargeback API**—Provides an interface for Chargeback functionality. This includes management of the hierarchy, cost configurations, and reporting.
- **VIX API**—Allows you to write programs and scripts that automate virtual machine operations, as well as the guests within virtual machines. This API is high-level, easy to use, and practical for both script writers and application programmers. It runs on either Windows or Linux and supports management of VMware Workstation, VMware Server, and VMware vSphere, including

ESX/ESXi and vCenter Server. Bindings are provided for C, Perl, and COM (Visual Basic, VBscript, C#).

Software Development Kits (SDKs)

- **vSphere SDK**—Interface for ESX/ESXi, vCenter, and VMware Server to extend the management of the virtual datacenter. Programming languages supported include Perl, .NET, and Java.
- **Virtual Disk Development Kit (VDDK)**—Interface to allow ISVs to use VMDK as a native format when developing virtual disk tools through the use of the VMware Virtual Disk Libraries (VixDiskLib and ViMntapi).
- **vSphere Guest SDK**—Enables development of applications that will run within a virtual machine using C or Java libraries. Enables customers to write smart applications that respond to changes at the virtualization environment layer. Included with VMware Tools.
- **vSphere Web Services SDK**—Provides a Web service accessible through the vSphere API to develop client applications.
- **vSphere SDK for Java**—Supports simplified vSphere Management applications by defining client-side data models. These models provide utility functions to simplify data access to servers.
- **Lab Manager SDK**—Enables development of applications that use Lab Manager Web service data, automate tasks, or integrate VMware Lab Manager with software testing tools.

Command Line Interfaces (CLIs)

- **vSphere Command Line (vCLI)**—Uses the vSphere SDK for Perl to provide commands to control vSphere and ESX/ESXi functionality. Previously known as RCLI or VI CLI. This tool is similar to command line functionality within the ESX Service Console and is useful for scripting and automating a repetitive task or pulling information out of the vCenter database.
- **Power CLI (PowerCLI)**—Allows you to manage your Virtual Infrastructure using Windows PowerShell. This allows you to script and automate actions you would normally do in vCenter. There are approximately 200 cmdlets (PowerShell exposed procedures) to manage vSphere and ESX/ESXi functionality. There are many pre-built scripts available online that can provide functionality such as finding all VM snapshots, finding orphaned VMs, or even creating reports. Previously known as the VI ToolKit.
- **VIX**—Allows development of programs and scripts to automate virtual machine and guest OS operations. VIX runs on Windows or Linux platforms. It manages VMware vSphere, ESX, ESXi, VMware Server, and VMware Workstation through the use of C, Perl, and COM bindings. COM bindings include Visual Basic, VBscript, and C#.

- **vSPhere Management Assistant**—Allows scripting and agents to manage ESX, ESXi, and vCenter Server systems. vMA is a virtual appliance that includes the vSphere SDK and the vSphere CLI, logging capabilities, and authentication mechanism.

The VMware vSphere suite is a collection of software providing management of a dynamic environment, cost reduction, and significant improvement to the life-work balance of IT professionals.

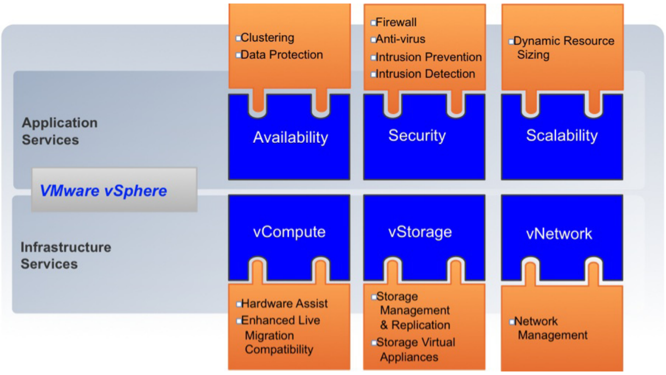


Figure 1: vSphere application and infrastructure service categories

2. What Is Virtualization?

A common definition of virtual is "something that exists in essence or effect but not in actual fact" or "performing the function of something that isn't really there." Virtual machines are servers or desktops that exist in essence and perform the function of an actual physical server or desktop, but that do not physically exist in a traditional sense. A virtual machine is made up of processes and files. They share physical hardware and are prevented from monopolizing it by the virtual machine layer.

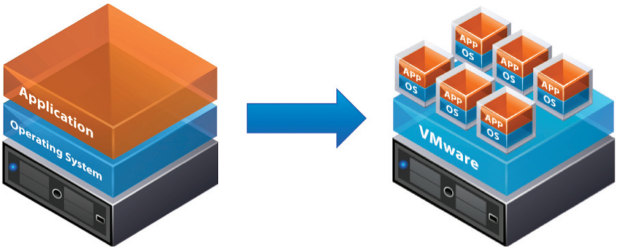


Figure 2: Virtualization groups App/OS pairs into virtual machines.

VMware achieves this by inserting, directly on the computer hardware or on a host operating system, a thin layer of software providing the illusion of actual hardware devices to multiple virtual machines. The same virtualized hardware devices are presented in the virtual machines regardless of the underlying physical hardware. This allows operating systems to install in virtual machines without any knowledge of the actual physical hardware. The virtual resources include CPU, memory, disk drives, and network interfaces.

VMware software allows multiple virtual machines to share the resources of a physical server known as the host. Multiple computer workloads are able to execute simultaneously without the limitations of tying a single operating system to specific hardware. VMware's virtualization solutions support the scaling of server virtualization across hundreds of host servers running thousands of virtual machines to create an entire virtual infrastructure.

Virtualization is a concept that has been familiar in the computer industry since the 1970s. It leverages resources such as computing, storage, and networking by abstracting them from the underlying hardware. For example, virtual memory is abstracted from the physical memory of the computer system, enabling virtual memory to be oversubscribed when presented to an application. Networking is virtualized through the use of VLANs (virtual local area networks) or VPNs (virtual private networks), storage through the use of storage and I/O virtualization, desktop and server hardware through virtual machines, and applications through application abstraction from the underlying operating system. In addition, vSphere 4 provides the concept of a vNetwork Distributed Switch (vDS) that spans a single virtual switch across multiple hosts and supports Private VLANs (PVLANs) and bi-directional traffic shaping.

VMware combines several aspects of virtualization—compute, storage, and networking—to create an underlying foundation for IT deployment and management. At the core of a VMware virtual infrastructure is the virtual machine.

Virtual Machines

The term "virtual machine" has many meanings, depending on which system layer is virtualized. Most prominent are system virtual machines and process virtual machines.

The phrase "System virtual machines" refers to a form of virtualization whereby the underlying physical computer resources are mapped into one or more different virtual machines (tightly isolated software containers that behave exactly like a physical computer).

Process virtual machines, also called *application virtual machines*, provide an abstraction of a high-level computer programming language runtime environment. The Java Virtual Machine (JVM) and the Microsoft .NET Framework's Common Language Runtime (CLR) are the two most popular process virtual machines.

This book focuses on system virtual machines. They are a representation of a real machine based on a software implementation providing an environment that can run or host an operating system such as Microsoft Windows or Linux. Each virtual machine contains its own resources, including CPU, memory, hard disk, video adapter, or USB controllers, but each resource is virtual, meaning that it is a software-based abstraction and contains no actual hardware components. Think of a virtual machine as an environment that appears to the operating system to be a physical computer.

How does a virtual machine operate? Running on the physical hardware underneath a virtual machine is a layer of software called a virtual machine monitor (VMM). The VMM has many distinct advantages over physical hardware by providing a layer of abstraction between the hardware and the virtual machine. Abstraction is a mapping of a virtual to physical resource similar to the way telephone call forwarding enables call receipt on one telephone number while ringing on a different phone number. The VMM is a layer below the virtual machine operating system and is invisible to it. The operating system and its system administrator are not aware of the VMM; it simply runs and provides the services needed by the virtual machine.

An operating system running in a virtual machine is called a *guest operating system*. The VMM layer quietly provides mapping between the physical and virtual resources. Each virtual machine guest OS acts as though it is installed on physical hardware and behaves exactly like a physical system.

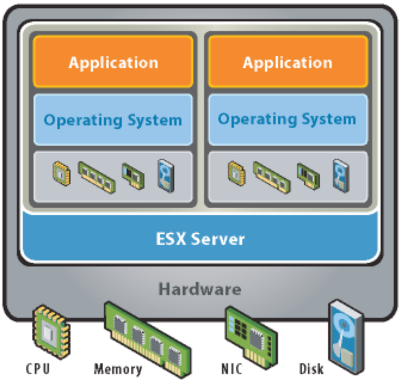


Figure 3: Mapping between physical and virtual hardware

Multiple virtual machines can run on a single physical computer and end-users can run multiple operating systems on a shared computer (*partitioning*). This creates logical partitions of the underlying computer. Partitioning is a key benefit of virtualization, but it's not the only one.

VMware ESX is able to schedule the workloads of multiple virtual machines on a physical server. In fact, a virtual machine, given a certain set of compute resources and shared storage, is free to migrate around the datacenter while ensuring the same performance to the end-user. This is one of the primary benefits of abstracting the hardware, as virtual machines are no longer bound to physical servers.

Characteristics of a Virtual Machine

Several inherent and fundamental characteristics of virtual machines are responsible for much of the flexibility and benefits of virtual infrastructures. These characteristics are a recurring theme throughout this book.

Compatibility

Virtual machines have all the components expected of a physical computer, including the CPU, RAM, and video graphics adapter. This compatibility with standard x86 computers enables standard *unmodified* operating systems, device drivers, and applications to run on the virtual machine. Operating systems and device drivers do not know they are running in a virtualized environment.

Isolation

Although many virtual machines can share a single physical computer, they are as completely isolated from each other as if they were separate physical machines. An operating system crash in one virtual machine cannot affect the operation of other virtual machines or the host server. A program running in one virtual machine cannot peek into the memory of another virtual machine. Unlike the vulnerabilities resulting from stacking applications on top of a single operating system, virtual machines provide secure and reliable consolidation of multiple applications on one physical server.

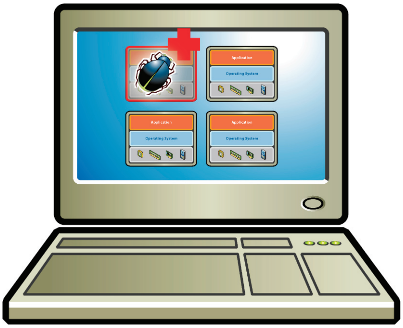


Figure 4: Isolation of virtual machines

Encapsulation

A virtual machine is essentially a software container that encapsulates a complete set of virtual hardware resources, the enclosed operating system, and the applications installed on the OS. In other words, a server becomes a set of files. The bulk of a virtual machine's size is due to a large file that represents the virtual machine's disk. A virtual x86 computer with a 60GB disk in many cases has a file that represents the 60GB disk within the virtual machine. Depending on the configuration, it is possible to create a single large virtual disk that spans multiple physical disks installed on the server. Alternatively, the virtual disk can reside on shared storage systems such as Storage Area Network (SAN–Fibre Channel or iSCSI) or Network Attached Storage (NAS) arrays.

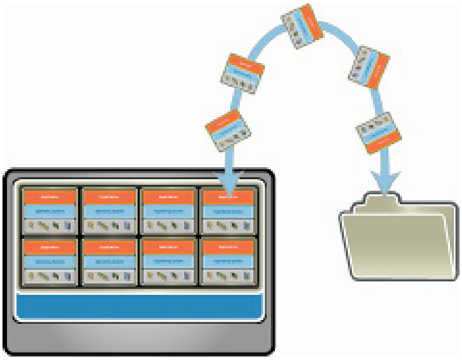


Figure 5: Encapsulation of a virtual machine into files

Hardware Independence

Virtual machine operating systems are completely independent from their underlying physical hardware due to the virtualization abstraction layer and are instead tied to a standard set of virtual hardware devices. This concept is similar to application independence from server hardware achieved by a standard operating system sitting between the application and the physical server. Virtual machines on the same physical server can even run different kinds of operating systems at the same time. This characteristic is provided by hardware virtualization.

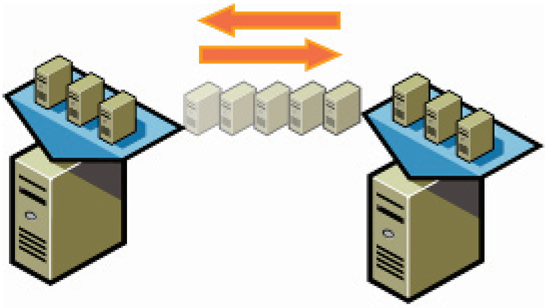


Figure 6: The virtualization abstraction layer provides hardware independence and portability.

What's in a VMware Virtual Machine?

A virtual machine consists of several files that define the virtual machine and encapsulate the virtual disks. These files include a virtual machine configuration file (*.vmx), an NVRAM (non-volatile RAM) file representing the BIOS settings, one or more virtual disks (*.vmdk), and one or more log files (*.log). Several additional files may also be used for operations such as use of virtual machine snapshots or saving the state of a suspended VM. Suspending a virtual machine is similar to suspending a physical machine but is handled differently from an operating system suspending itself by going into, in the case of Windows, *hibernation mode*. Because the virtual machine monitor has complete control over the virtual machine, it is possible to suspend a virtual machine outside the control of the operating system. Indeed, it is possible to place any operating system, regardless of its support for power management or hibernation, into a suspended state from which it can later be resumed.

The virtual hardware platform presented to a virtual machine is standardized regardless of the underlying physical hardware. This provides independence from the underlying physical hardware. The following are the components found within a VMware virtual machine (as implemented in VMware vSphere 4):

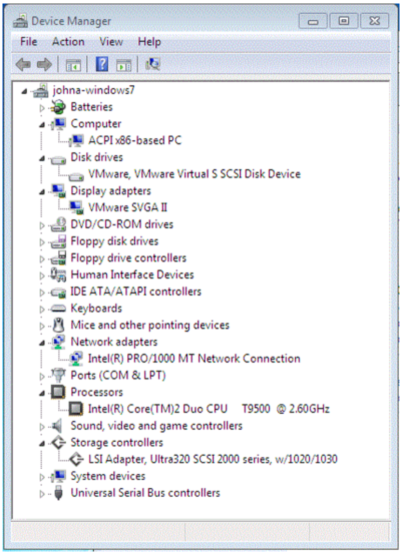
- An Intel 440BX-based virtual motherboard
- Virtual Phoenix BIOS 4.0 Release 6
- NS338 SIO chip
- Up to eight virtual CPUs (vCPUs)—same processor type as host
- Up to 255GB of RAM
- Up to four CD/DVD-ROM drives
- Up to two parallel ports and up to four serial/COM ports
- Up to two 1.44MB floppy drives
- SVGA graphics adapter
- Up to ten virtual network adapters
- VMware Accelerated AMD PCNet Adapter (PCI II compatible)
- VMXNET Generation 3—paravirtualized NIC adapter with MSI/MSI-X support, Receive Side Scaling, IPv6 checksum and TCP Segmentation Offloading (TSO) over IPv6, VLAN offloading, and large TX/RX ring sizes (configured within VM).

- Two DirectPath PCI/PCIe devices with up to 60 SCSI targets
- Up to four SCSI controllers with up to 15 devices each
- Generic SCSI device support
- Mylex (BusLogic) SCSI Host Adapter Parallel (BT-358)
- LSI Logic SCSI Host Adapter Parallel (Ultra 320)
- LSI Logic SAS (Serial Attached SCSI)
- VMware Paravirtualized SCSI (PVSCSI)—high-performance storage adapters offering greater throughput and lower CPU utilization for virtual machines (for applications that are I/O intensive)
- Hot additions of hard drives, CPU, and memory for guest operating systems supporting this as a feature

Note: VMware vSphere 4 does not provide a maximum for the number of PCI or PCIe adapters. The specific maximum on each of the adapter types will be hit without maxing out on the PCI/PCIe bus.

The VMware virtual machine *remote console* provides access to the console of the virtual machine, including boot options and configuration. Remote console access is the same as accessing the local console of a physical server through the directly connected keyboard, video, and mouse. At a maximum, 40 remote console sessions can be attached to a single virtual machine. For resource conservation, minimize the number of open remote console sessions.

Figure 7 shows the devices as seen from the Windows Computer Management views.



- johna-windows7
  - Batteries
  - Computer
    - ACPI x86-based PC
  - Disk drives
    - VMware, VMware Virtual S SCSI Disk Device
  - Display adapters
    - VMware SVGA II
  - DVD/CD-ROM drives
  - Floppy disk drives
  - Floppy drive controllers
  - Human Interface Devices
  - IDE ATA/ATAPI controllers
  - Keyboards
  - Mice and other pointing devices
  - Network adapters
    - Intel(R) PRO/1000 MT Network Connection
  - Ports (COM & LPT)
  - Processors
    - Intel(R) Core(TM)2 Duo CPU T9500 @ 2.60GHz
  - Sound, video and game controllers
  - Storage controllers
    - LSI Adapter, Ultra320 SCSI 2000 series, w/1020/1030
  - System devices
  - Universal Serial Bus controllers

Figure 7: Windows computer management view of devices

vSphere Components and Plug-Ins

VMware vSphere consists of multiple components allowing for server virtualization of an entire IT infrastructure architecture.

The components of a virtual infrastructure include the underlying virtualization system and a comprehensive suite of technology management utilities, creating a complete computing virtualization solution. Ultimately, the performance, capacity, and reliability of the solution are enabled or limited by the underlying hardware.

A VMware virtual infrastructure consists of one or more ESX hosts, a VMware vCenter Server, and a set of tools to provide additional functionality. These tools include DRS for distributed resource scheduling, VMware HA for server hardware recovery, and VMotion for live migration of virtual machines.

VMware ESX

ESX is the core server virtualization platform in the VMware product line, a server virtualization product that fits into the larger datacenter virtualization product space. ESX is designed for maximum performance, availability, security, scalability, and reliability. ESX abstracts the most essential devices: processors, memory, storage, and network resources.

ESXi 4.0 has a thin 32MB footprint and is managed by vCenter Server. It provides the exact same performance and functionality as ESX 4.0 but strips out the Service Console. All control and configuration is managed through one of four mechanisms: VMware vCenter, the VI Client, the VI Management Appliance (VIMA), or a Web interface for this particular version.

ESX offers a bare-metal that includes the following characteristics:

- CPU virtualization, providing time-sharing between multiple virtual machines and direct pass-through execution of ring 3 commands on the processors.
- Storage virtualization, supporting FC SAN/iSCSI SAN/NAS devices that feature virtual disk files, the VMFS cluster file system, a logical volume manager, direct mapping to raw SAN LUNs, Fibre Channel HBA consolidation, write-through I/O, and boot-from-SAN capabilities.
- Network virtualization, featuring 802.3ad link aggregation, virtual NICs, virtual switches with port configuration policies, VLAN support, and vNetwork Distributed Switches (vDS) that span multiple ESX Servers.

VMware vCenter Server

VMware vCenter Server provides centralized management and configuration of multiple ESX hosts as a single unit in a VMware vSphere environment. It is used for configuring, provisioning, and managing virtual machines, networking, and storage, as well as providing centralized license management. Support for managing both ESX and VMware Server is included with appropriate licensing.

VMware Server

VMware Server is a free hosted solution that runs on supported versions of Microsoft Windows and Linux operating systems. It is designed to assist in software development efforts, to run virtual machines requiring specialized hardware devices, and to provide an entry point for those new to virtualization.

VMware Server is designed for maximum hardware compatibility and can present any hardware device supported by the underlying operating system to the virtual machines.

VMware Virtual SMP

Virtual SMP enables allocation of multiple virtual CPUs (vCPUs) to a virtual machine. Up to eight vCPUs can be allocated to virtual machines in VMware vSphere 4.0. Before assigning multiple vCPUs to a virtual machine, a CPU utilization assessment should be performed to determine whether the operating system and associated applications would benefit from virtual SMP. The best practice is to start with a single vCPU, then scale up as necessary.

Multi-threaded applications that benefit from multiple CPUs in a physical environment can also benefit from multiple vCPUs. Many organizations make multiple CPUs a standard in physical servers. When moving to VMware vSphere, this might be an inefficient use of virtual resources. You must determine whether the underlying ESX host hardware will adequately support the total number of vCPUs allocated to virtual machines. For a two-way, single-core CPU platform, two vCPUs on a virtual machine could potentially limit the total number of virtual machines due to the way VMware allocates CPU cycles. When a two-vCPU virtual machine runs in its time slice, two physical CPUs are typically locked. An exception to this is when one vCPU is executing a “real” workload while the other vCPU is executing an idle loop. In this case the ESX host only needs to schedule on physical core for use. Although this is the case, the use of multiple vCPU virtual machines on ESX can reduce the overall consolidation ratio and can in some cases prove to be a bottleneck in an ESX host. vSphere 4 provides better handling of multiple vCPU virtual machines. Design with the VM resource requirements in mind to reduce the overhead on the host platform.

VMware VMotion

VMotion offers a technology to migrate running virtual machines between ESX hosts. VMotion migration requires either the same processor family on both the source and target ESX hosts, or “Enhanced VMotion Compatibility” (EVC) on a cluster of hosts with technologies enabling VMotion-compatibility with older servers. Hosts need to be grouped within the same vCenter datacenter. The shared storage holding the VM virtual disk is presented to both the source and target hosts.

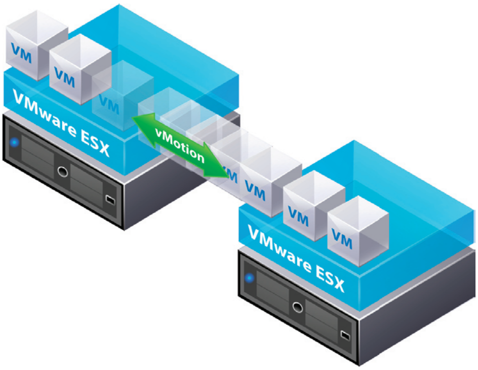


Figure 8: VMotion

More information on VMotion CPU validation can be found in the VMware whitepaper *VMware VMotion and CPU Compatibility,* available at http://www.vmware.com/files/pdf/vmotion_info_guide.pdf.

VMware Storage VMotion

Storage VMotion enables live migration of virtual machine disk files across storage locations while maintaining service availability. Storage VMotion utilizes VMotion technology to optionally move the VM to an alternate ESX host which has access to both the source and target storage locations. Storage VMotion can move the storage location of a virtual disk as long as the target is visible to the source and destination ESX hosts(s). The processes of the corresponding VM can stay on the same host, or the VM can be simultaneously VMotioned to a new host. Limitations are enumerated in the documentation for Storage VMotion.

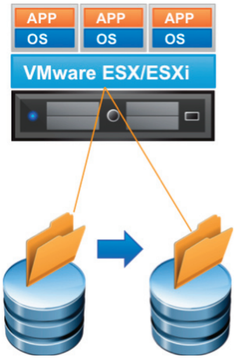


Figure 9: Storage VMotion

VMware vCenter Host Profiles

VMware vCenter Host Profiles enables more fully automated host provisioning through the use of host templating. Host Profiles facilitates quickly adding new hardware to the datacenter by enabling easy application of consistent configurations to the hosts, including NFS, firewalls, NTP settings, DNS settings, user profiles, security profiles, and advanced settings in host configurations. Additionally, Host Profiles provides monitoring to ensure hosts remain in compliance as configurations change over time.

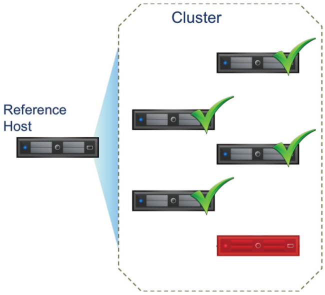


Figure 10: vCenter Host Profiles

vCenter Server Heartbeat

vCenter Server Heartbeat enables monitoring, replication, and rollback for instances of vCenter Server, enabling effective disaster recovery in the event of loss of the primary datacenter. It is also particularly useful for protecting vCenter Server when installed on a physical server.

vCenter Linked Mode

vCenter Linked Mode allows up to 10 vCenter instances (or datacenters) to be aggregated into a single vSphere client, and it includes an advanced search capability designed to assist in finding VMs throughout the linked vCenter instances. As with the other components of vSphere, third-party manufacturers can enhance vCenter with SDK and toolkits.

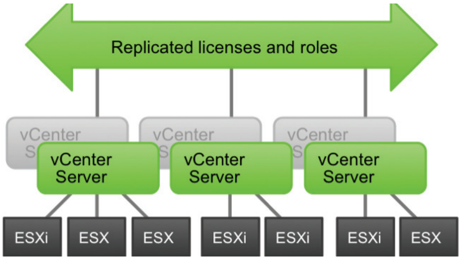


Figure 11: vCenter Linked Mode

vCenter for Linux (Technology Preview)

vCenter for Linux is particularly beneficial for organizations not running Windows, meaning they do not need to purchase and support a single copy of Windows just to manage their virtual infrastructure. vCenter for Linux is a technology preview, pre-beta, which can be downloaded for free on the VMware Community Forums: http://communities.vmware.com/community/beta/vcserver_linux.

VMware vCenter Chargeback

vCenter Chargeback provides the ability for organizations to gauge the cost of providing and maintaining virtual machines, including rack space, power, cooling, software licenses, and associated maintenance. It includes report generation for both internal departments and external customers, encouraging more efficient use of virtual infrastructure.

VMFS Volume Grow (Dynamic VMFS Expansion)

Dynamic VMFS Expansion is provided by the VMFS Volume Grow mechanism. This enables the VMFS volume extent to dynamically increase in size along with dynamic LUN expansion without disrupting the virtual machine.

VMware vStorage Thin Provisioning

VMware vStorage Thin Provisioning allows oversubscription of storage, enabling the number of allocated "disks" to add up to much more storage than the amount owned. Advanced management capabilities ensure that over-provisioning does not adversely affect the most important workloads and that it takes advantage of technologies such as disk duplication. An open API enables leveraging the thin provisioning offered by storage array manufacturers by enabling them to write vSphere plug-ins.

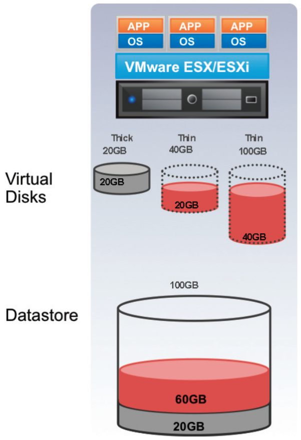


Figure 12: Thin Provisioning

vNetwork Distributed Switch (vDS)

Defined in vCenter Server, the vDS is an abstract representation of multiple hosts defining the same vSwitch and port groups. This provides the illusion that the VM is connected to the same network as it migrates between multiple hosts. Network management is simplified by moving the virtual network configuration from the host level to the datacenter level. The configuration plane and I/O plane have been split, with the I/O continuing to run on the host level.

This type of switch presents ports to virtual machines and connects to DV Uplink ports that correspond to physical NIC ports. Policies can be programmed on both DV Port Groups (VM side) and DV Uplink Port Groups (NIC side).

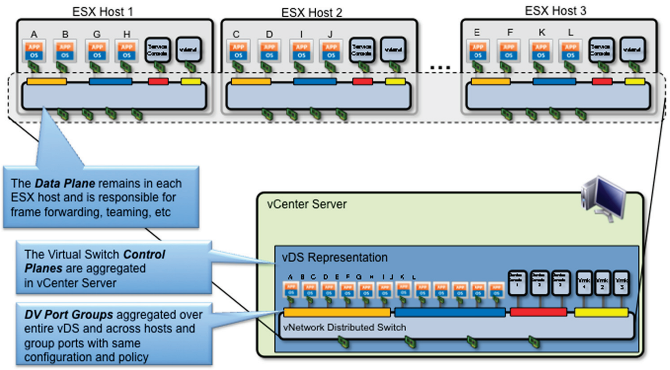
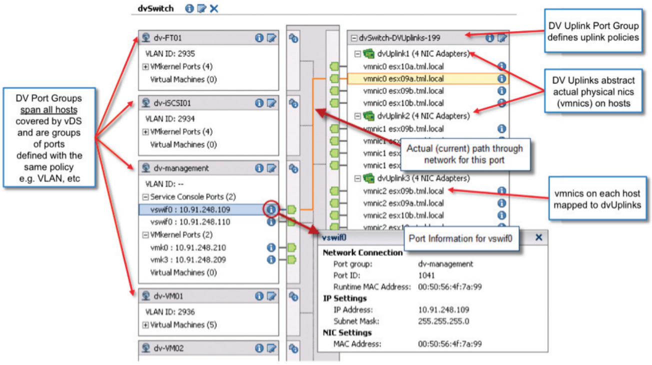

ESX Host 1
ESX Host 2
ESX Host 3
A
B
G
H
C
D
I
J
E
F
K
L
APP
OS
Service Console
vmkernel
...
The Data Plane remains in each ESX host and is responsible for frame forwarding, teaming, etc
The Virtual Switch Control Planes are aggregated in vCenter Server
DV Port Groups aggregated over entire vDS and across hosts and group ports with same configuration and policy
vCenter Server
vDS Representation
A B C D E F G H I J K L
vNetwork Distributed Switch

Figure 13: vNetwork Distributed Switch (vDS) (diagram from Guy Brunsdon of VMware)



vswif0

Network Connection

```
Port group:            dv-management
Port ID:               1041
Runtime MAC Address:   00:50:56:4f:7a:99
IP Settings
IP Address:            10.91.248.109
Subnet Mask:           255.255.255.0
NIC Settings
MAC Address:           00:50:56:4f:7a:99
```

Figure 14: vNetwork Distributed Switch Connections (diagram from Guy Brunsdon of VMware)

VMware High Availability (HA)

VMware High Availability (HA) is a clustering technology that provides virtual machine high availability on VMware Infrastructure. If one host server fails, all the virtual machines configured for VMware HA can be restarted on an alternate ESX host. The ESX hosts must all be members of a VMware HA cluster and must have access to the shared data stores. Data stores are file systems that store the virtual machine configuration files and virtual disks, formatted with either VMFS or NFS file system formats.

VMware Fault Tolerance (FT)

VMware Fault Tolerance (FT) provides zero downtime and zero data loss for a virtual machine by keeping a second instance of the FT-enabled virtual machine in lockstep. If the host server running an FT-enabled virtual machine fails, a failover to the secondary virtual machine will take place. The failover is completely transparent to the end-user. The initial releases of vSphere 4 limit a maximum of one vCPU for a VM running in FT mode.

VMware Distributed Resource Scheduling (DRS)

VMware Distributed Resource Scheduling (DRS) is another ESX clustering technology that works in conjunction with VMotion to manage load balancing across ESX hosts while providing a guaranteed quality of service for groups of hosted virtual machines. When the load on one node in the cluster is unbalanced, DRS uses VMotion to rebalance it, minimizing resource contention and guaranteeing resource levels for virtual machines. DRS works with VMware HA to ensure that loads are balanced and resource guarantees are respected in the event of virtual machine redistribution after a host failure.

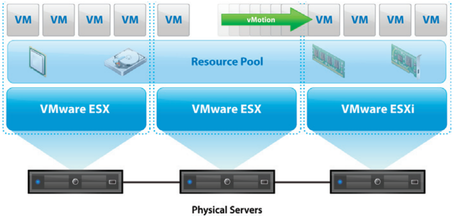


Figure 15: VMware Distributed Resource Scheduling

VMware Distributed Power Management (DPM)

VMware Distributed Power Management (DPM) optimizes power consumption in a VMware DRS cluster by consolidating workloads on fewer servers and powering off unneeded hosts. DRS is an integral part of DPM.

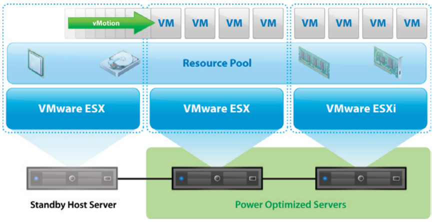


Figure 16: VMware Distributed Power Management

VMware vShield Zones

VMware vShield Zones provides monitoring and control of network traffic based on corporate security policies defined through the use of vShield virtual appliances. These policies can be used to align with regulatory compliance guidelines. vShield Zones allows multi-tenancy (internal or external customers on a shared computing resource pool) while maintaining trust and network segmentation of users and data. This technology allows monitoring, logging, and blocking of inter-VM traffic within one ESX host or across multiple ESX hosts. Bridges and firewalls can be created to create multiple zones based on trust boundaries.

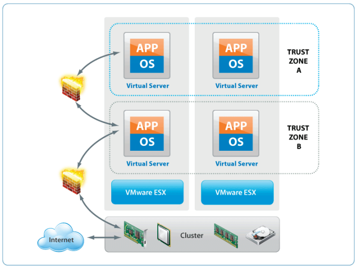


Figure 17: vShield Zones

The management component of this technology is vShield Manager, which is used to create a firewall for virtual machines. The rules used within this firewall can manage down to the TCP/UDP port level, operating in Layers 2, 3, and 4 of the network stack.

VMware VMsafe

VMsafe is both an API for security management and control, and a partner program. VMsafe provides programming interfaces that allow developers to build security solutions, such as antivirus or intrusion detection tools, for the VMware virtual infrastructure. A virtual machine is protected by these security solutions through inspection of virtual components, including CPU, memory, network, and storage, meaning protection at both the host and network layers. By creating tools at this level, there is less resource load generated within the virtual machine. Similar to how VMware Consolidated Backup offloads backup traffic from the VM networks, VMsafe offloads security management and control from the virtual machines.

VMware Consolidated Backup (VCB)

VMware Consolidated Backup (VCB) provides a mechanism to offload virtual machine backup traffic from the VM network, the guest OS, and the underlying host by using a proxy server with direct access to the shared storage. VCB uses virtual machine snapshot technology, allowing the VCB proxy to back up images of running virtual machines directly from the storage arrays. A file-level method is provided for daily recovery, such as for lost files, and a full image method is provided for disaster recovery purposes. The full image includes the virtual disk(s), configuration file, NVRAM file, log file(s), and any additional files that represent the state of the machine.

VMware Data Recovery (vDR)

VMware Data Recovery (vDR) offers a disk-based backup and recovery solution that includes data de-duplication to save on disk storage for backups. It is centrally managed by VMware vCenter Server. This technology supports restoration of individual files or entire VM images and automatically monitors VMs that move within the VMware environment to ensure backup continuity. It is designed for datacenters requiring backup of no more than 100 VMs. Full VM backups are taken but allow for individual file restores.

vCenter Orchestrator

vCenter Orchestrator provides a predefined set of workflows that operate with vCenter and specific features. The version provided with vCenter offers predefined workflows that can be modified and extended. It allows drag-and-drop workflow creation, using both the out-of-the-box workflows and access to the vCenter Server API (800+ actions possible).

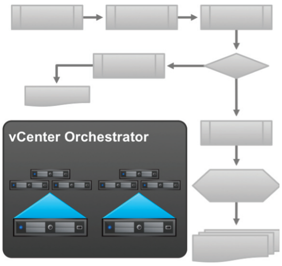


Figure 18: vCenter Orchestrator—Workflow Orchestration

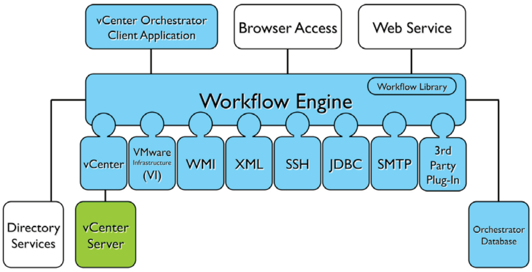


Figure 19: vCenter Orchestrator—Workflow Engine

VMware VMDirectPath

VMware VMDirectPath enables connecting the virtual machine directly to a physical hardware component such as a network card. VMDirectPath enables the virtual machines to experience native storage and network performance characteristics when required for particularly demanding workloads.

VMware Record and Replay Virtual Machine Execution

vSphere provides the ability to record and replay the execution of a virtual machine for forensic or debugging purposes. APIs enable third parties to control this functionality.

The VMware vCenter vCloud Plug-in

The vCenter vCloud Plug-in allows management of, and authentication to, both internal and external clouds. The vCloud Plug-in uses a variety of techniques to bridge the VM and cloud and offers a private cloud using resources from both services.

Working with Virtual Appliances

VMware vSphere is a new layer in the traditional infrastructure architecture. In addition to the required infrastructure hardware and the VMware product suite, many virtual environments benefit from additional value-added components that extend the native VMware functionality or provide new features for integrating VMware vSphere with existing systems and processes. In particular, there are aspects of working with virtualized applications (Virtual Appliances, OVF, vApp, and ThinApp), which are discussed in the following section.

Virtual Appliances

Virtual appliances emulate hardware application appliances in that they include a prebuilt and preconfigured application and operating system with simplified management designed for a specific solution. The operating system and application are packaged using the industry-standard Open Virtualization Format (OVF) rather than a physical server. The OVF-formatted virtual appliance can be downloaded and deployed on ESX or other virtualization platforms. Some virtual appliances emulate traditional physical hardware appliances such as routers, firewalls, and backup appliances. Others are built to distribute test or evaluation software or to provide unique functionality such as inline packet-level patching or storage virtualization.

Whether physical or virtual, the appliance method of application distribution has several advantages over distributing software applications to be installed by customers on standard OS x86/x64 machines. The appliance typically includes a slimmed-down version of the operating system tailored to optimally manage the specific application. This utilizes the computing resources more efficiently while improving reliability, simplifying troubleshooting, and enhancing security through a standardized application environment. Appliances generally require less frequent patching. They also eliminate problems resulting from customers using incompatible hardware or incorrectly installing the application.

The downside to hardware appliances is that they are costly, take up rack space, and require power both to operate and to cool. Hardware appliances result in more under-utilized servers and also can result in non-standardized hardware in the datacenter. They have parts that can fail, so duplicate devices might be required to guarantee redundancy. Additionally, redundant devices might also be required at disaster recovery sites.

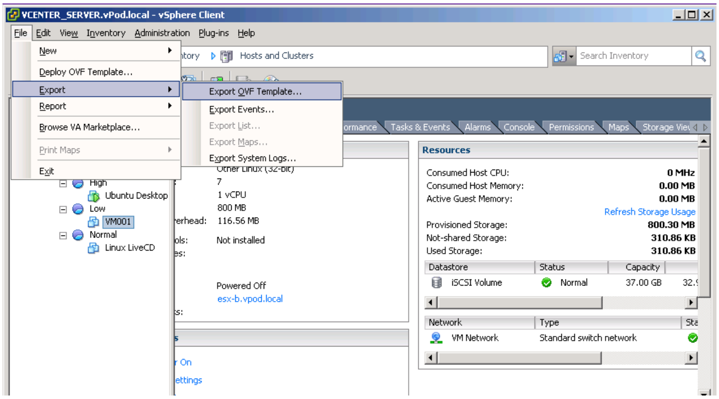
Virtual appliances take the application appliance concept to a new level, providing all of the advantages of hardware appliances and utilizing a specialized operating system in a controlled environment, but without requiring a dedicated hardware server. Deployment is simplified, costs are reduced, and high availability is enabled without requiring duplicate hardware. The virtual appliance can even be replicated off-site, along with the virtual infrastructure, for disaster recovery without requiring additional appliance licensing or dedicated hardware at the recovery site. Downloading, deploying, evaluating, and replacing virtual appliances is much quicker and easier.

For these reasons, software manufacturers are increasingly delivering their applications as virtual appliances. By packaging a database as a virtual appliance, the software manufacturer no longer needs to be concerned about what hardware, drivers, and OS version the application is being installed on. The manufacturer's best practices are already incorporated into the virtual machine, ensuring that it is configured correctly. Complexity is reduced while reliability is improved, and the resource burden on the customer is also greatly reduced. In addition, the ability to quickly and easily reproduce a problem through access to the real VM, reproduced in a testing environment at the vendor site, can help the vendor provide a high support response and customer satisfaction.

Another advantage of virtual appliances is that an organization can collapse more of the network support services into the virtual infrastructure along with the application servers. Virtual appliances relying on the network transport, such as firewalls, gain significant performance advantages by keeping the connections close to the virtual application servers inside the physical hosts using virtual switches. The virtual appliances will see benefits at the Layer 2 level, the level at which VMware networking operates. There is less latency within a virtual infrastructure than routing traffic from the virtual infrastructure out to the physical network and back. These factors make custom-built virtual software appliances a natural extension of a virtual infrastructure. As more of the infrastructure is virtualized, the portability of the entire infrastructure increases, enabling simpler disaster recovery planning, which makes virtual appliances even more beneficial. The increased availability of virtual appliances for all types of software will continue to simplify the deployment and support of virtual IT infrastructures.

Open Virtualization Format (OVF)

Open Virtualization Format (OVF) provides an open, secure, portable, efficient, and extensible format for virtual machines that enables virtualized appliances to be moved between hypervisors and architectures. The OVF package format provides a complete description of a single VM or complex multi-VM environments. The package is optimized for distribution. VM authors can describe "portable hardware" for maximum interoperability or optimum performance. Both compression and package signatures can be applied. Version 1.0 of OVF has been ratified by the DMTF. For further information on OVF, see http://www.vmware.com/appliances/learn/ovf.html. VMware vCenter supports both importing and exporting of virtual machines in an OVF format.



Other Linux (32-bit)
7
1 vCPU
800 MB
erhead: 116.56 MB
ols: Not installed
es:
Powered Off
esx-b.vpod.local
s:

Resources

Consumed Host CPU:	0 MHz
Consumed Host Memory:	0.00 MB
Active Guest Memory:	0.00 MB
	Refresh Storage Usage
Provisioned Storage:	800.30 MB
Not-shared Storage:	310.86 KB
Used Storage:	310.86 KB

Datastore	Status	Capacity	
iSCSI Volume	✓ Normal	37.00 GB	32.9

Network	Type	Sta
VM Network	Standard switch network	✓

Figure 20: OVF Template creation

VMware vApp

VMware vApp is a logical entity comprising one or more virtual machines. These virtual machines use OVF, and all components of a multi-tiered application are included in a package. In addition, operational policies and service levels can be set as attributes for vApp. The vApp provides a standardized method for describing these operational policies for the associated application. vApps enable the movement of the application(s) between internal or external clouds while still supporting the same policies.

The vApp construct was developed to help with the growth of cloud computing by enabling ease of migration between internal and external clouds and across different virtualization technologies supporting OVF. vApps provide more flexibility than virtual appliances, because they allow defining and enforcing policies on an entire application stack and not just on a single virtual machine. This provides fine-grained control to support performance aspects of the embedded application(s).

VMware Studio is a tool used to author, configure, deploy, and customize vApps and virtual appliances. VMware Studio is a free virtual appliance downloadable from the VMware site. The tool packages software applications so that they may be easily run within VMware product platforms and a cloud platform supporting OVF. Both Linux-based and Windows-based VMs, vApps, and/or virtual appliances can be built with multiple application tiers included. For further information, see the product description at http://www.vmware.com/appliances/learn/vmware_studio.html.

VMware ThinApp

VMware ThinApp allows provisioning of applications abstracted from the underlying operating systems. The application is packaged as a standard EXE or MSI file, including registry settings or drivers. The application is executed in a sandbox environment. ThinApp applications can either be stored locally in the operating system or may be placed on a network share for simplified use and deployment. ThinApp packages reduce complexity and dependencies during the development and update management process. You can find more information at http://www.vmware.com/products/thinapp/using.html.

3. The Benefits of Infrastructure Virtualization

The first edition of this book described how virtual infrastructure reduces capital and operational expenses while improving datacenter agility. While these same benefits continue to be true, and are indeed still larger today, technological advances enable organizations to improve performance, management, and security while reducing risk by virtualizing their datacenters.

VMware vSphere provides the scalability, availability, security, and management capabilities required to enable complete datacenter virtualization. When combined with the latest generation of multicore CPUs designed to optimize virtualization performance, nearly every workload can now run as fast or faster as a virtual machine, including Microsoft Exchange, SQL Server, large Web servers, and even Oracle OLTBs. Deploying virtual machines on robust shared storage and encompassing the enhanced networking capabilities of vSphere should result in overall virtual datacenter performance superior to that of a physical environment.

Beyond the huge capital and operational expense savings, a completely virtualized datacenter provides levels of reliability, agility, and recoverability unattainable in the physical realm at any price. New vSphere capabilities combined with published APIs enabling plug-ins from leading industry manufacturers enable a virtualized datacenter that is more secure and better managed than its physical counterpart.

A virtual datacenter is also the platform for implementing robust private and public cloud strategies, thereby enabling still greater resource utilization efficiency and further economic benefits.

Capital Expense Reduction

- Huge reduction in server hardware expenses resulting from increased server and storage utilization (the average physical MS Windows server utilization is less than 15%)
- Reduced facilities-related capital costs such as for air conditioners, generators, PDUs, racks, and cabling
- Reduced network and storage switches and cards
- Reduced OS licensing costs
- Reduced storage costs with VMware vSphere Thin Provisioning
- Reduced hardware requirements for high availability and disaster recovery
- Reduced cost to purchase PCs, laptops, and remote office infrastructure
- Potential elimination of software and hardware requirements for clustering by using VMware Fault Tolerance

Operational Expense Reduction

- Reduced facility costs for rack space and electricity
- Reduced procurement costs concomitant to reduced server purchases
- Reduced resources required for deployment, replacing, and upgrading physical servers
- Reduced IT staff time spent troubleshooting, patching, capacity planning, budgeting, and performance monitoring of servers
- Reduced downtime expenses resulting from hardware failure, maintenance, or upgrades and servicing (High Availability, Fault Tolerance, VMotion, Maintenance Mode, and Update Manager)
- Reduced network, storage, and disaster recovery management costs
- Reduced cost to manage, support, and secure the desktop environment
- Reduced power costs of ESX hosts through VMware Distributed Power Management
- Potential reduced cost to manage remote office network infrastructures through their reduction or elimination entirely

Improved Agility

- Ability to provision a private cloud complete with chargeback of IT services to individual departments
- Ability to federate resources between private and public clouds, still further enhancing both performance and efficiency
- Enhanced management, security, and automation capabilities using advanced VMware features and products such as DRS, Storage VMotion, Update Manager, vCenter Orchestrator, Host Profiles, VMsafe, vShield Zones, VMware Record and Replay, LifeCycle Manager, and Lab Manager, among others
- Increased ability for disaster recovery through VMware encapsulation and portability as well as advanced features such as VMware Consolidated Backup, VMware Data Recovery, and VMware Site Recovery Manager (more cost-effective savings)
- Increased service levels from reduced deployment time and cost and from increased responsiveness
- The ability to remove legacy hardware while remaining fully operational
- Greater service delivery speed through provisioning applications via virtual appliances
- Better support for testing applications and operating systems (via VMware snapshots and live clones)
- Greater alignment of IT with business through increased responsiveness to business requirements
- Ability to enhance organizational "green" initiatives by slashing electricity usage

Reduced Risk

A 2009 webtorials.com IT survey showed server failure as the number one cause of outage, followed by human error. Virtualization, through VMware High Availability and Fault Tolerance, enables such fast recovery that server failure is no longer an operational risk. It also reduces the risk of human error by enabling all testing, patching, and upgrades to be performed in a virtual sandbox prior to a production rollout. Utilizing the VMware management, security, and automation products makes a virtualized datacenter both simpler and less risky to operate than its physical counterpart.

Summary of Benefits

The following is a summary of the benefits that can be achieved through virtualization of a datacenter.

- Reduced risk—Reduces datacenter risk by incorporating superior virtualization capabilities in performance, management, security, and automation
- Server consolidation—Reduces the number of physical servers and associated costs
- Server containment—Prevents server sprawl
- Legacy applications—Allows older software applications and operating systems to run on the latest hardware platforms
- Server uptime—Eliminate server failure as an operational concern
- Simplified disaster recovery—Encapsulates an operating system, its applications, and its state within a small set of files, eases recovery, and eliminates typical plug ’n’ play issues
- Standardized hardware for operating system stability—Abstracts the underlying hardware from the virtual machine, reducing risk and easing underlying hardware changes
- Production stability—Simplifies capacity planning and quickens response to the needs of the business through standardized hardware and encapsulation, defined roles and responsibilities, and a robust management interface with resource trending information
- Ease of use—Enhances management and security capabilities
- Desktop management—Uses VMware View to ease desktop management, upgrades, and access

More benefits continue to be found as the technology develops and additional use cases are identified.

The Business and Operational Case for Virtualization

We described earlier how hardware virtualization enables operating systems consolidation onto fewer hardware platforms. Several additional benefits are created by the defining characteristics inherent in virtual machines. In the following sections, we identify benefits enabled by the following key characteristics of VMware virtual machines:

- Compatibility
- Isolation

- Encapsulation
- Hardware independence

Compatibility

Virtual machine compatibility created by standardization is an essential requirement for successful VMware implementations. Incompatibility of virtual machines with operating systems or applications causes instability for some workloads and makes it impossible to virtualize others. Standardized and compatible virtual machines provide great benefits. Availability is increased and troubleshooting is simplified when virtual machines present identical hardware profiles to the operating systems regardless of the underlying hardware.

Isolation

The isolation characteristic of virtual machines is critical in production and software development environments. Virtual machine isolation has significant availability and security benefits. This feature allows users to safely run multiple virtual machines on one host. Security-conscious organizations need confidence that information from one virtual machine cannot leak into another and that a compromised virtual machine cannot have access to the memory, data, or network traffic of any other virtual machine on the host. The isolation property of virtual machines ensures that this does not happen.

The availability and stability benefits of isolation are necessary because in a production or software testing environment you cannot afford to have one virtual machine disturb the operation of other virtual machines. In a software development environment, test programs may crash or misbehave, but this does not affect the operation of other environments that share the same computer.

Encapsulation

Encapsulation makes virtual machines portable and easy to manage by representing physical hardware in a set of files on disk. A virtual machine is backed up by copying the set of files. Of course, a backup program can also be run within the virtual machine to perform incremental or full backups.

As simple as this seems, this is a major benefit. Having a whole operating system instance in a set of files allows for virtual machine portability. A system administrator does not have to worry that a particular configuration is composed of an operating system install, set of updates and patches, a set of registry settings, and appropriate documents and settings directory files, all configured for specific hardware. A virtual machine can be replicated, copied to portable media, or archived just like any standard data files, then executed on any supported virtualization platform at a later time. This characteristic, along with hardware independence, provides one of the greatest benefits of a virtual infrastructure: the virtual machine mobility and portability that enable simplified disaster recovery and hardware migrations.

Hardware Independence

When coupled with the properties of encapsulation and compatibility, hardware independence provides the freedom to move a virtual machine from one type of x86 computer to another of the same instruction-set family without making any changes to the device drivers, operating system, or applications. Hardware independence also means that you can run a heterogeneous mixture of operating systems and applications on a single physical computer. In your datacenter you can mix and match servers from multiple vendors with ease, and also use a single group of servers to run both Windows and Linux without any alteration or reinstallation of operating systems. However, you might want to standardize on a few hardware vendors for other reasons, such as spare parts, standardized hardware configuration, and simplified purchasing and discounts. The flexibility of changing vendors using the layer of virtualization provides for a form of freedom that has seldom been seen before.

Return on Investment (ROI)

It is hard to imagine a more clear and compelling ROI than is obtainable from implementing a VMware Virtual Infrastructure. The primary areas of savings result from virtualizing production datacenters, backup and disaster recovery, and desktop systems.

Production Datacenters

The most obvious savings result from consolidating servers. While consolidation ratios can vary widely depending on the type of servers being virtualized, VMware DRS enables effective load balancing of workloads among multiple ESX servers. It is not uncommon to achieve average consolidation ratios of 20 or more virtual machines per two-CPU late-generation multicore servers. An organization virtualizing its production environment, even allowing for redundant ESX hosts and resource intensive tier 1 applications, may be able to consolidate 85% or more of its physical servers, along with their associated power consumption, cooling costs, and maintenance expenses. The demand for less power can lead to further savings from reduced requirements for air conditioners, PDUs (Power Distribution Units), UPS (Uninterruptible Power Supply) devices, and generators. Slashing the number of servers similarly slashes expenses for rack space and other maintenance-related costs. This can be particularly significant for organizations running out of datacenter space or using outsourced datacenters that charge by the square foot or rack.

Storage

Virtual storage becomes a reality in a VMware virtual infrastructure. Features that SAN manufacturers sometimes require to be purchased separately, such as server snapshots, migrations, and multipath software, are provided through VMware's disk abstraction. The thin provisioning capability of vSphere provides a higher level of utilization. Resource consolidation provided by ESX allows all virtual servers to benefit from high-performance SAN storage without incurring the per-host connectivity fees traditionally associated with network storage. Purchasing and maintenance costs for host bus adaptors (HBAs) and storage switches can be greatly reduced because multiple virtual machines of all types and priorities can be connected to high-end shared storage through *storage* *virtualization*. Also, because LUNs can span many virtual machines, managing the storage environment is easier as well.

Network

VMware's virtual switches provide the ability to seamlessly extend the physical network into the virtual network. Network virtualization expands the network infrastructure without requiring additional switch ports. Consolidating 100 servers on six ESX hosts can result in saving 60 network switch ports while providing multiple link redundancy for each virtual server. This same configuration can provide SAN connectivity to the 100 virtual servers while avoiding the cost of 172 FC switch and HBA ports.

Additional virtual switch features such as 802.1q (VLAN tagging) and 802.1ad (port trunking or link aggregation) provide increased flexibility and functionality to the virtual and physical network infrastructure. The VMware vNetwork Distributed Switch (vDS), included with VMware vSphere Enterprise Plus, enables application of policy-based network capabilities that move with a virtual machine during live migration. This allows virtualization of more servers while still meeting security and regulatory policies. Consolidation of the virtual machine management reduces the amount of time spent on virtual network administration.

Management

The ability to manage all VMware vSphere components from a single point of view (pane of glass) makes virtual infrastructure administration less complex than physical. An IDC and VMware study showed the ability for administrators to effectively manage over three times more servers in a virtualized environment than in a physical environment.

Software Licensing

Savings frequently result from operating system licensing policies in a virtual environment. For example, licensing Microsoft Windows Server Datacenter Edition for the underlying CPUs of the ESX host enables an organization to run unlimited instances of Windows Server on the host. The same is true for SQL Server Enterprise and for BizTalk, as well as other applications from various vendors.

Backups and Disaster Recovery

Virtualization enables backups to be performed at the virtual disk (*.vmdk) level as well as standard file and block-level backups. This results in less effort and money spent by organizations trying to meet shrinking backup availability windows.

Disaster recovery in a virtual environment is not only functionally much more effective, it is more cost-effective as well. Far fewer servers are required at the recovery facility, and they do not need to be the same brand or type of servers as the production servers. Both data and virtual machines can be continuously replicated to the recovery site, enabling inexpensive recovery of a VMware virtual infrastructure. Disaster recovery testing can also be done much less expensively and without requiring personnel to go on-site at the recovery site facility. VMware's Site Recovery Manager (SRM) automates the disaster recovery process, enabling still further increased effectiveness with reduced staffing requirements.

Desktops

Virtualizing the desktop with VMware View (VDI) enables another realm of savings and very positive ROI. These savings result from reducing the frequency of PC and laptop upgrades and from reduced power and operational costs when using thin-client or zero-client terminals or lower-cost PCs or laptops. Maintenance costs are lowered and administrative requirements are reduced. An IDC analysis (Michael Rose and Randy Perry, September 2009) concluded that organizations deploying VMware View saved on average over $610 per supported end-user per year compared with organizations using unmanaged PCs, with an additional $122 per year when using the capabilities of VMware View Premier. The IDC white paper can be found by registering for the desktop virtualization kit on www.vmware.com. An example of a desktop ROI can be found on http://www.bythebell.com/2009/09/the-desktops-may-be-virtual-but-the-roi-is-real.html.

ROI Case Study: A Community Bank

The following example of ROI savings is based on a VMware deployment by a major San Francisco Bay Area community bank. This bank consolidated 102 of its 105 servers onto six ESX hosts. The bank realized a cost savings of $1.6 million over a five-year period versus an initial investment of $270K, resulting in an investment payback period of only 10 months.

While the majority of savings result from server consolidation and containment, substantial savings also result from reducing DR costs and electricity requirements, shown in the figures on the following page.

ROI/TCO Calculator

The VMware ROI/TCO Calculator provides a method for determining ROI and TCO (total cost of ownership) based on server consolidation scenarios. It provides support for measuring the savings or estimating potential savings when deploying any combination of VMware vSphere, VMware Lab Manager, and VMware View.

The tool focuses on comparing the costs of an existing physical computing environment against the same environment after being converted into a VMware virtual infrastructure environment. The tool asks five to ten questions about the existing environment to determine the cost to deploy and maintain the infrastructure. The questions are about industry, business, and technology drivers; location; current assets, including the number of servers to be virtualized; and time required for a typical new server deployment, among others.

This information is used to set default values for 200 additional metrics used to determine the cost to deploy and maintain the existing and future virtualized infrastructure. These values can be modified to better represent your organization or to create what-if scenarios.

Initially, the calculator quantifies the cost of maintaining the existing physical infrastructure. Then it calculates the costs for converting the environment into a VMware virtual infrastructure. The cost savings are then quantified for management overhead, productivity improvements, risk reduction, and other technical and business benefits. The ROI/TCO Calculator produces final outputs, including the cost investment, savings, and additional benefits to set up, deploy, and support the new virtualized infrastructure.

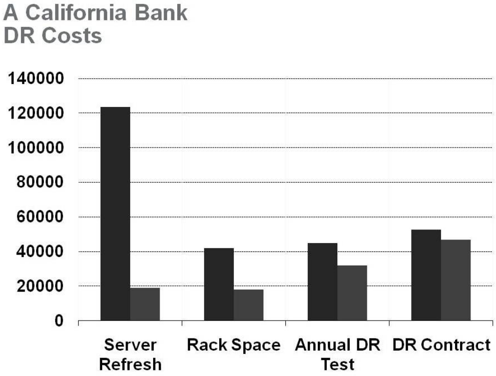


Figure 21: Physical versus virtual DR costs

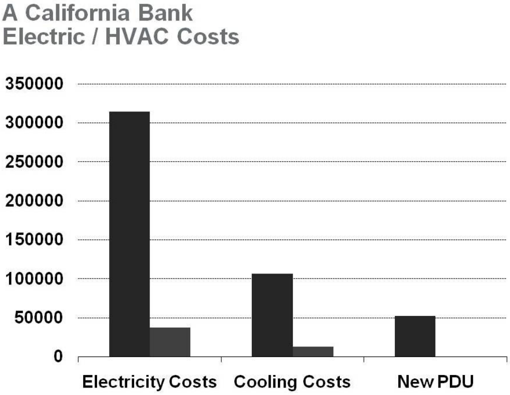


Figure 22: Physical versus virtual electric costs

Results of the analysis can be viewed comprehensively or individually for VMware vSphere 4, Lab Manager, and VMware View deployments. The calculator results can be used to analyze different scenarios by adjusting variables and saving multiple result sets for comparison. Results can be reviewed online or in export PDF, RTF, or raw data (XLS) formats. The VMware ROI/TCO Calculator is a valuable tool if you are considering or trying to justify a large-scale server, desktop, or lab virtualization project. The calculator is also an excellent tool for verifying ongoing savings with accurate values after a deployment.

The online ROI/TCO calculator can be found at http://www.vmware.com/products/vi/calculator.html.

4. Use Cases for Virtualization

VMware is used by 96% of the Fortune 1000 in a production capacity, but the average virtualization penetration within a datacenter is only around 30%. While effective as a point solution, the full benefits of virtualization are realized when it is deployed as an encompassing platform for the entire datacenter, incorporating both storage and network along with servers. VMware vSphere provides the scalability, availability, management, and security requirements enabling a 100% virtualized datacenter, which in turn provides the foundation for both effective cloud computing and service-oriented architecture.

Production Environments

A virtualized datacenter reduces risk by providing high availability for any server, independent of traditional clustering technologies like MSCS and VCS. It enables more effective disaster recovery by being able to continuously replicate all servers off-site. It maximizes reductions in equipment and facility expenses while simplifying administration.

Cloud Computing

VMware vSphere transforms a datacenter into an internal cloud by enabling a pool of virtualized resources within a datacenter. It also enables seamless resource shifting between internal and external clouds.

Service-Oriented Architecture

A virtualized datacenter is a key enabler of service-oriented architecture (SOA) by automatically provisioning IT infrastructure in response to application requirements. Dynamically allocated virtualization pool resources optimize service requests. Additionally, as applications increasingly are written to communicate directly with the ESX hypervisor, the requirement for a traditional full operating system layer with its inherent greater costs and inefficiencies is eliminated. The better the integration with the hypervisors, the lower the total overhead.

Software Test/Development and Testing

ESX got its start in the test/development area. Because ESX enables administrators to easily make a snapshot copy of a server, the snapped copy can then be patched, upgraded, or migrated in a separate virtual environment that uses VLANs to emulate the production network. Also, VMware Lab Manager takes software development to a new level by enabling developers to provision multiple sets of servers in seconds, as well as the ability to share and archive environments for QA testing or troubleshooting.

Disaster Recovery

Research organizations such as Meta and Gartner emphasize the high probability of business failure if an organization suffers a disaster such as a fire or flood. This results from the lack of an effective disaster recovery plan or the failure of a disaster recovery plan. Disaster recovery plans for traditional physical computing infrastructures are very complex and expensive to implement and maintain, usually requiring a mirror of the production infrastructure at a remote site. These plans are difficult to test effectively before they are needed and they tend not to work well, if they work at all. Conversely, VMware Site Recovery Manager enables disaster recovery that is significantly less complex, more affordable, and testable. All virtual machines, not just a small subset deemed as mission-critical, can be included in a VMware Infrastructure disaster recovery solution. The virtual machines with their networking and storage can be *continuously replicated* from the production datacenter to the recovery facility where they are ready to be activated with a single click. Virtual desktops and virtualized applications can also be replicated, enabling users to connect to their applications and data from anywhere they can access the Internet using a browser.

Remote Offices

Small, remote offices frequently can be inexpensively virtualized using VMware vSphere Essentials, which is a lower-cost version of vSphere designed for smaller deployments, making virtualization in small offices affordable while providing the encapsulation and mobility features that simplify backup and disaster recovery from remote offices.

Desktops

VMware View extends virtualization to the client by running desktops on centralized vSphere hosts. This computing model gives employees far more flexibility by being able to work from any PC, Mac, Microsoft Windows terminal, or zero-client device. When employees log on, they are directed to their virtual desktop, applications, and data. Unlike server-based computing that relies on a shared Windows kernel to achieve multiuser access, VDI truly provides users with their own customized desktops while reducing the cost of hardware and simplifying administration.

5. Designing a Virtual Infrastructure

As you virtualize your IT infrastructure, you need to address the planning, design, implementation, management, and technical details of VMware vSphere. The VMware Server Consolidation Methodology provides a proven approach to virtual infrastructure.

VMware Server Consolidation Methodology

The VMware Server Consolidation Methodology consists of several phases designed to support a smooth transition from your existing environment to a virtualized infrastructure. Small tactical deployments of VMware vSphere can be done without using the full methodology, but skipping any of these phases for a larger infrastructure deployment usually leads to a more expensive and under-optimized infrastructure. All of the phases are recommended and are typically used to complete a full end-to-end virtualization strategy.

Phase 1—Assessment and Planning Phase

The assessment and planning phase includes conducting a virtualization assessment to identify virtualization candidates and their resource requirements. Application requirements and interdependencies are also determined. During this phase, inventory and performance data is collected to develop implementation options. A gap analysis provides detailed information about missing components, staffing, and operational procedures. Phase 1 can include training, identification of options for the design, and a proof-of-concept implementation for demonstration purposes.

Phase 2—Design Phase

In the design phase, customized tactical implementation details, technical blueprints, management and assembly guides, and test plans are created for deployment. These are used for the deployment of VMware vSphere during the third phase. Phase 2 can include a pilot program to test aspects of the design to validate alignment with business requirements.

Phase 3—Build/Implementation Phase

Deployment and testing occur during the build and implementation phase. Virtual machines are built or imported using VMware Converter or other physical-to-virtual (P2V) migration tools. Both unit-level testing of individual components and testing of multiple system components working together are conducted to ensure that new virtual machines fit correctly within their allocated resources and environment. User acceptance tests are conducted at the end of this phase.

Phase 4—Management Phase

The management phase is when maintenance and ongoing refinement to the infrastructure occur. Performance is analyzed and tuning is performed to ensure compliance with service level agreements. Service level agreements can cover availability, performance, and other areas essential to your business requirements. A check is typically performed once a year by internal teams or external consultants to identify deviations from current VMware best practices and supported configurations. Over time, best practices evolve based on technology updates and feedback from users.

Identifying Virtualization Candidates

When beginning a server virtualization project, the first question is often which servers can or should be virtualized. Large-scale server virtualization projects done without accurate visibility into the existing infrastructure can undermine predictability and efficiency and increase risk. A systematic approach to identifying virtualization candidates and computing capacity is required to realize the maximum benefits from a server virtualization and consolidation project.

Almost all operating systems and applications are potential candidates for virtualization. However, some may require analysis to determine the resources required to maintain service levels. When determining VMware vSphere resource requirements, Virtualization Assessments consider many factors, including utilization of the original physical machine and its hardware, operating system, and applications.

Good candidates can maintain or improve performance levels in a virtual machine post-virtualization. The following are items that need to be considered or are typical qualifiers when determining virtualization candidates.

Vendor Support

Does the vendor support their application running within a virtual machine?

- If yes, then proceed with migration.
- If no, then follow up with vendor to determine plans for support.
- If the vendor has no plans, determine whether using virtual to physical (V2P) migration during troubleshooting will satisfy internal and vendor support needs. This allows reproducing a problem on a physical platform.

Resource Requirements

- If more than eight CPUs are required to maintain service levels, the server may not be a candidate for virtualization.
- If USB ports are required, a USB-over-IP hub might be considered to provide the resource. Otherwise, VMware Server, which provides greater hardware compatibility, may be an alternative.

Real-time Data Acquisition

- If the application and associated devices providing data to the application can ensure no data loss, then the virtual machine can typically be virtualized. This typically requires a cache mechanism at the device site.

- If not, then the application stays physical.

Conducting a Virtualization Assessment

There are several phases in a comprehensive virtualization assessment. First, you must inventory the hardware and software components of a system. The second phase involves analyzing application resource utilization for the target physical servers over the course of several weeks. During the final phase of the assessment, the resource utilization of the virtualization candidates is analyzed to determine the aggregate CPU, memory, network, and disk I/O requirements. This enables proper sizing of the target VMware Infrastructure.

Manually gathering inventory data from a large number of servers can be time-consuming and difficult. Some tools that can be used to help with assessments include VMware Capacity Planner, IBM CDAT, Microsoft Msginfo32, Microsoft Srvinfo, the UNIX System Activity Reporter (SAR) tool, and Platespin PowerRecon. These tools provide support for one or more operating systems. VMware Capacity Planner is used during a VMware Virtualization Assessment to gather an inventory and performance metrics. The typical data collection requires 30 days to account for any workload variance in the month. Afterwards, optimization scenarios are run against the target hardware platform to produce consolidation ratios.

Inventory

The first part of the virtualization assessment is the gathering of inventory data for the systems that are to be virtualized. Specific items to consider include:

- Type, speed, number of CPUs, and number of cores per CPU
- CPU average and peak usage
- RAM available and peak usage
- Disk space available and used
- Network I/O used
- Type and number of NICs
- Operating system and version (include patches and service packs)
- Applications installed
- Special hardware used, which determines whether or not it can be virtualized on ESX

Application Resource Considerations

Application resource utilization must be considered and analyzed. Analyzing the resource utilization of each operating system and application pairing over a 30-day business cycle makes the expected resource load evident. Interactions between the existing resource utilization and additional overhead caused by virtualization must be considered. CPU virtualization incurs the least amount of overhead, but disk I/O increases CPU utilization and context switching.

The subtle interactions between these various components, combined with translating individual application performance to a consolidated platform on a new hardware cluster, makes virtualization planning a complex task.

Tracking resource utilization over a one- to three-month period provides important data to determine candidates for virtualization. The components to measure include CPU, RAM, disk I/O, and network I/O. Average and peak usage, base characteristics of devices, and the target hardware platforms influence the migration and deployment design.

To manually conduct this analysis, a programmer could use statistics gathered using the Perfmon for Microsoft Windows operating systems or the System Activity Reporter (SAR) tool for UNIX or Linux operating systems, but these tools cannot translate the data to useful virtualization scenarios. Using a capacity analysis tool such as VMware Capacity Planner provides a more accurate model and greatly simplifies the consolidation planning process.

CPU

ESX 4 supports a maximum of 64 cores on its host platform, and up to eight vCPUs in a virtual machine. It is recommended not to provision a virtual machine with more virtual CPUs than the host has CPU cores.

Additional metrics which influence the target virtualization architecture include CPU queue depth, %READY (the percentage of time an instruction was ready to execute but could not run due to lack of CPU resources), user and system time, as well as other items. The %READY measure is valid for both physical and virtual machines.

When sizing the target host hardware, it is important to anticipate the number of vCPUs required for the virtual machines. It is a best practice to initially provision a single vCPU per virtual machine. After monitoring performance and utilization, additional vCPUs can be added if necessary. Multi-threaded applications often benefit from multiple vCPUs.

An estimate of four vCPUs per core can be used for capacity planning. Actual consolidation ratios vary based on performance data and target ESX hardware platform. When selecting CPUs for host systems, larger cache size can provide performance gains. Large cache sizes (both level 2 and level 3 caches) typically yield higher performance gains.

Virtualizing CPU resources creates lower overhead than memory, disk, or network I/O virtualization. CPU virtualization adds varying amounts of overhead, depending on how much of the virtual machine workload can be run in direct execution.

RAM

VMware ESX 4.0 supports a maximum of 1TB of RAM on its host platform. Each guest OS can be allocated a maximum of 255GB of RAM (guest OS limitations apply).

Faster memory makes a significant difference to application performance as consolidation ratios are increased.

All operating systems use page tables to map virtual memory to physical memory. The VMware hypervisor allocates physical memory to the guest OS using shadow page tables. It then uses a separate table to map those allocations to physical memory. This process allows memory sharing between guest operating systems and can reduce memory latency. We also use CPU techniques such as Intel's Extended Page Tables™ (EPT)and AMD Rapid Virtualization Indexing (RVI). Both of these provide support for Memory Management Unit (MMU) virtualization increasing performance, particularly for

MMU-intensive workloads. Some benefits include higher throughput and lower CPU utilization.

DISK I/O

VMware ESX 4.0 can support up to 64TB (minus 512B overhead) of storage per host using a file block size of 256MB. Each guest OS can be allocated a maximum of 2TB (minus 512B overhead) of storage.

The analysis of disk I/O is a critical part of resource planning for a server consolidation project. Disk access can sometimes be the constraining resource for application performance in both physical and virtual computing platforms. In a VMware Infrastructure, storage resources cannot be dynamically migrated for performance optimization or have the quality of service features that CPU and memory provide. Storage VMotion can be utilized to manually move virtual disks to assist in optimizing disk performance. Disk I/O is the virtualization resource that requires the most overhead, consuming CPU time and creating additional interrupts.

Depending on throughput requirements, application servers can rely heavily on disk I/O for performance. It is best to map SAN LUNs to application servers according to their required performance levels. A file server may be optimally configured for sequential access utilizing a RAID 5 LUN built on 7200 RPM Fibre Channel or SATA disks. A typical LUN for high-performance applications such as databases would be built on RAID 10 using many 15,000 RPM disks for optimal random access speeds. High disk I/O performance can come at a price.

When using iSCSI, TCP/IP can consume large amounts of CPU cycles unless paired with an iSCSI TOE (TCP Offload Engine) card. Also, Fibre Channel requires overhead, as the host has to use CPU cycles to compute disk I/O. However, with current hardware and improvements in the vSphere iSCSI software initiation, CPU overhead should not be a concern.

vSphere provides support for NFS Jumbo Frames when using 1Gbps and 10Gbps NICs. NAS, like iSCSI, can benefit from TOE cards. Some vendors claim performance gains up to 30% depending on load.

Measuring the anticipated disk I/O profile of the servers to be consolidated is essential for accurate disk I/O planning and maintenance of service levels. A virtualization assessment provides the level of detail necessary to plan for thoughtful I/O consolidation.

In the case of large Exchange servers or database servers, performance considerations for I/O may be important. Without the benefit of a virtualization assessment, users often create a LUN and continue to add virtual machines to it until performance degrades or the capacity is exhausted. To recover performance on the oversubscribed LUN, virtual machines need to be migrated off the congested LUN, costing additional administrative time and SAN overhead. This approach rarely leads to optimal performance and capacity utilization. Support for N_Port ID Virtualization (NPIV) has been added to allow each virtual machine to have its own World Wide Name (WWN), allowing for VM-specific QoS. Additionally, full support for round-robin HBA load balancing is included, as well as support for third-party multipathing plug-ins like EMC's Powerpath/VE.

When sizing VMFS storage, measure disk throughput in inputs/outputs per second (IOPS) for each candidate server collected during the virtualization assessment. A good rule of thumb is that a Fibre Channel disk can support approximately 100 IOPS RAID groups. LUNs should be built with enough disks to support the anticipated IOPS for the candidate servers. This planned approach can lead to predictable performance with fewer migrations to reallocate virtual machines from oversubscribed LUNS.

NETWORK I/O

VMware vSphere 4.0 can support Fast Ethernet (100 Mbps), GigE (1000 Mbps), and 10G (10,000 Mbps, with support for TCP Segmentation Offload (TSO) and jumbo frames.

Virtualized network I/O on ESX hosts, like disk I/O, creates greater overhead than CPU or memory virtualization while simultaneously adding to the CPU load for each I/O. Much of network performance (like disk performance) is external to the ESX host and out of the control of the VMkernel and the Virtual Machine Monitor (VMM). These factors should be considered when identifying candidates for server virtualization.

These additional layers of overhead need to be taken into account when sizing host servers and considering which systems to use for virtualization. Several improvements in networking efficiency have been introduced that reduce VMM overhead and decrease network I/O-induced CPU utilization. The latest version of the VMXNET virtual network driver includes TCP Segmentation Offload (TSO) support which allows previously CPU-intensive TCP segmentation operations to be off-loaded to the physical network adapter. Jumbo Ethernet frames (frames with a size greater than 1500 bytes) are also supported; they greatly reduce the number of packets needed to transmit data. Each I/O operation removed from the CPU workload through TSO and jumbo frame support reduces virtualization overhead and increases available CPU time for virtual machine execution.

When considering a server's workload for virtualization, it is important to consider the VMM and VMkernel overhead created by network I/O, and whether TSO or jumbo frames can be used to reduce this overhead.

Agile Approach to Designing the Virtual Datacenter

Designing a virtual datacenter requires significant consideration for integration with hardware, software, policies, and procedures within a business. A design includes more than configuration of vCenter and ESX/ESXi. In reality it starts with the underlying requirements and constraints from the business.

In many situations, a waterfall design approach does not always fit due to time, budget, or staffing constraints. An agile approach provides an opportunity for starting a design while the planning phase is underway, and for starting the implementation during the design phase. A waterfall approach is sequential while an agile approach supports parallelism in the process.

As part of the design process, factor in the justification and impact of each design choice. This information provides a clear picture during the review process with different business units and the core IT teams. It also ensures better integration with other VMware and third-party technologies.

6. Building a VMware vSphere Environment

Selection and configuration of hardware components are critical to the performance, scalability, and reliability of a VMware infrastructure. The details of the host hardware don't matter to the guest operating systems in a virtual infrastructure, but the underlying hardware architecture and configuration have considerable effect on the performance, scalability, and reliability of the VMware vSphere clusters.

The VMware vSphere hardware cluster consists of two main components: the host server hardware and the storage network and its subsystems. Two additional hardware components, the data network and the backup infrastructure, play an important part, but are usually external to the virtual infrastructure. The most important requirement for system stability and VMware support of ESX or storage hardware is compliance with the VMware Hardware Compatibility List (HCL). The tight coupling between the hypervisor and hardware depends on using tested and compatible hardware. The VMware HCL can be found on the VMware Web site: http://www.vmware.com/resources/compatibility/.

Server Hardware

In a virtual infrastructure, the function of the host server hardware is to provide a container (processing, memory, and I/O) for guest operating system execution. The nature of consolidated resource sharing on an ESX host or cluster makes selection and optimization of its underlying hardware critical to efficiency and performance. The ROI and success of a VMware vSphere project depend largely on hardware selection.

When selecting ESX host hardware, CPU core density, memory, host server motherboard, I/O, and scalability must be considered for maximum cost-efficiency, scalability, and performance.

CPU Counts and Core Density

As CPU speed and core density increase, host servers are able to provide processing power for more guest operating systems. Multiple multicore processors provide low latency performance during the simultaneous execution of multiple guest operating systems within ESX. A typical processor configuration consists of at least two quad-core 64-bit compatible processors with Intel VT or AMD-V virtualization extensions for maximum compatibility and performance. High CPU core densities and large processor caches provide greater ESX scalability and performance than increased CPU MHz alone.

The high core counts in today's CPUs allow for high consolidation ratios on servers with only two physical processors. Four or more physical processors are recommended if the host will be running virtual machines with quad-processor SMP. To run ESX 4.0, 64-bit processors are required.

Memory

Memory is a critical component of virtual infrastructure hardware. Today's latest generation of 64-bit-capable servers feature large memory capacity and high-speed front side buses. The high consolidation ratios supported by servers with high processor core counts require servers with large amounts of memory to fully utilize the available processor power. Even with ESX memory sharing and memory reclaiming capabilities, ESX needs additional memory available to satisfy Distributed Resource Scheduling and High Availability services.

It is not uncommon for ESX to reach memory capacity while still having unused processor power available, making memory the limiting resource to scalability. A dual-processor eight-core server hosting 20 to 25 guest operating systems can often utilize 32–48GB of RAM when accounting for Distributed Resource Scheduling and overhead.

Host Server Motherboard Considerations

The continuing increase in processor speeds only translates to incremental increases in overall system speeds without fully optimized subsystems to support data movement within the host server. The latest server generation features up to 1600 MHz front side bus speeds to support the newest PCIe and memory speeds and processor core densities. ESX benefits from parallel processing, increasing guest operating system scalability with the addition of cores and processors. However, it is the internal bus, memory, peripheral speeds, and system cache that have the greatest effect on the application and guest operating system speeds.

PCIe Slots

Due to the scalability enabled by quad-core processors and 64-bit memory addressing, it is desirable to configure servers with a large number of PCIe slots for network and storage cards. PCIe buses enable dedicated I/O paths, allowing for linear performance scalability. It is not unusual for 8- or 16-core ESX hosts to use six or more network interfaces and at least two storage I/O cards. The maximum number of supported cards is provided in the "Configuration Maximums" paper on the VMware Web site at http://www.vmware.com/pdf/vsphere4/r40/vsp_40_config_max.pdf.

Processing virtualized storage and network I/O can consume a large number of CPU cycles, creating context switches and a considerable amount of FSB (front side bus) traffic. In a physical system, data can be moved across the system buses five times after entering the server before it gets to the CPU. Processing 1GB/s of network traffic can consume up to a GHz of CPU cycles. This situation is exacerbated by time-sharing between a large number of guest operating systems on an ESX host. This results in a large percentage of host server CPU power dedicated to processing network and storage I/O, ultimately limiting system scalability. ESX supports TCP segmentation offload, jumbo frames, and the use of TCP Offload Engine (TOE) network cards. These NICs offload

TCP processing from the CPU, reducing I/O overhead and freeing CPU cycles. New technologies like Infiniband connectivity overcome these limitations by using remote direct memory access and low latency 10GB/s consolidated into a single link (two links for redundancy). However, with today's processors, you will need to keep in mind that investing in additional processors or processor speed might be more cost-effective than, for instance, buying iSCSI Initiators.

Storage Hardware

Storage networks and shared storage arrays are critical to a well-architected virtual infrastructure. Storage design, configuration, and performance are the most important and complex hardware subsystem considerations in virtual infrastructures. Virtual machines are stored in raw LUNs, NFS shares, or VMFS file systems and are executed on any of the servers in a VMware vSphere cluster. Networked storage is necessary for all of the advanced VMware vSphere management features, including VMotion, HA, FT, DRS, vDR, and VCB. Storage-related issues are the most common and problematic VMware vSphere support issues.

The full potential and savings of virtualization can only be realized when the storage system is optimized and managed as a single functional unit within the virtual environment. There are a wide variety of storage arrays tested and approved for compatibility with VMware vSphere 4, from inexpensive cluster solutions and mid-range arrays to large enterprise storage systems. Across this range of arrays, feature sets and performance vary widely. All storage arrays on the VMware HCL can be integrated into a VMware vSphere environment, but there are some storage system characteristics more optimized for VMware vSphere. Several storage vendors offer VMware vCenter plug-ins to enable monitoring and managing the virtual environment from a single pane of glass.

VMware vCenter Server provides interfaces for dynamically managing all aspects of a VMware vSphere environment, including hosts, clusters, guest operating systems, and virtual networks. Storage and VMware vSphere environments are often managed by different teams, with storage systems generally being less agile and dynamic than the rest of the VMware virtual infrastructure. Storage as an extension of a VMware virtual infrastructure must be able to quickly react to change if it is to add to, and not detract from, the overall value of the VMware virtual infrastructure. The ideal situation is to have the storage system managed as a part of the VMware virtual infrastructure, extending the boundary of the virtual infrastructure to include the physical and logical storage. Newer storage systems that provide storage virtualization features allow a more dynamic and manageable storage layer for ESX clusters.

Infrastructure performance issues that are not caused by oversubscribing the server's capacity are most often traced back to storage architecture. Server processor speeds have increased approximately 12 times faster than storage performance has, fostering the perception that simply throwing more or faster processors at a performance issue will cure it. One vital but often underappreciated requirement for VMware vSphere storage systems is the need for linear performance per capacity scaling.

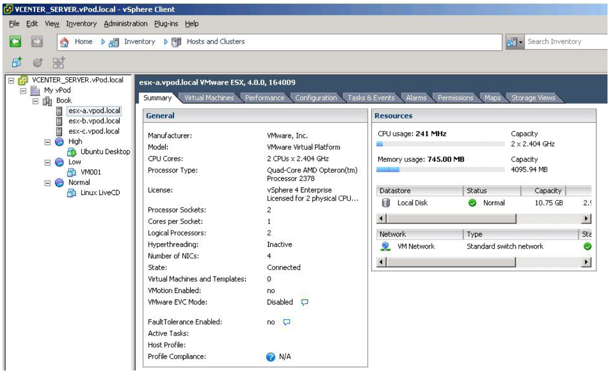
Virtual infrastructures have a very different storage profile from those of most traditional workloads. Capacity expansion is driven by adding virtual servers, often with higher-load virtual machines being added in later phases of the projects. Maintaining multiple virtual disk files in each LUN combined with the resource time-sharing inherent in virtualization results in increased random disk I/O with each virtual server added to the infrastructure. A storage system selected for a VMware vSphere environment should be able to scale performance linearly with virtual server growth to the upper limits of its capacity growth.

Managing quality of service is essential when consolidating several different workloads with various service level needs on a VMware vSphere cluster. VMware vCenter Server supplies robust quality of service features to maintain desired service levels across different priority workloads running on shared hardware. The Distributed Resource Scheduler provides effective grouped and granular controls over server CPU and memory resource shares and limits; however, there is no effective way to manage disk access performance levels from within the standard VMware vSphere management systems today.

To maintain true quality of service within a VMware Infrastructure, the storage systems should be able to easily differentiate and manage LUN performance for different virtual server workloads. VMware infrastructure administrators should be able to manage storage quality of service as easily and non-disruptively as CPU and memory QoS administration.

ESX

ESX is the primary datacenter virtualization platform within VMware Infrastructure. ESX is a thin layer of software that runs on bare-metal x86/x64 hardware. It abstracts available hardware resources, such as processor, memory, storage, and networking, and multiplexes them among virtual machines running unmodified commodity operating systems.



esx-a.vpod.local VMware ESX, 4.0.0, 164009

General

Manufacturer:	VMware, Inc.
Model:	VMware Virtual Platform
CPU Cores:	2 CPUs x 2.404 GHz
Processor Type:	Quad-Core AMD Opteron(tm) Processor 2378
License:	vSphere 4 Enterprise Licensed for 2 physical CPU...
Processor Sockets:	2
Cores per Socket:	1
Logical Processors:	2
Hyperthreading:	Inactive
Number of NICs:	4
State:	Connected
Virtual Machines and Templates:	0
VMotion Enabled:	no
VMware EVC Mode:	Disabled
FaultTolerance Enabled:	no
Active Tasks:	
Host Profile:	
Profile Compliance:	N/A

Resources

CPU usage: **241 MHz** Capacity 2 x 2.404 GHz

Memory usage: **745.00 MB** Capacity 4095.94 MB

Datastore	Status	Capacity	
Local Disk	✓ Normal	10.75 GB	2.!

Network	Type	Sta
VM Network	Standard switch network	✓

Figure 23: vSphere Client view of an ESX Server

Prerequisites

ESX can be deployed on a variety of server hardware platforms. The exact configuration of server hardware that is certified to run ESX is frequently updated. The VMware Compatibility Guide contains the most current list of systems certified to run ESX (http://www.vmware.com/resources/compatibility).

VMware ESX 4.0 will only install on a server with 64-bit x86 CPUs.

Installation

The high-level steps for installing ESX are:

1. Collect information about the system on which ESX is to be installed.
2. Decide on an installation method from the following options:
   a. Local CD-ROM or DVD-ROM drive.

   b. A remote CD/DVD drive using the remote system software provided by the manufacturer of the hardware system.

   c. Automatic Kickstart-based scripted installation. (See the section Kickstart Scripted Installation, below.)

   d. Download ESX software ISO image and burn a CD or DVD.
3. Choose and identify the correct network interface, as the ESX Service Console requires this. By default, ESX uses the first interface discovered (vmnic0) as the Service Console interface. On a system with multiple network interfaces, it is important to know which NIC shows up as vmnic0. Viewing the BIOS and identifying the MAC address of the NIC can determine this.
4. Install ESX.

Local CD or DVD Drive Installation

After the target server is booted from an ESX install CD, two options are available for continuing the installation:

- A mouse-based graphical installation is available. This is the easiest and recommended method of installing ESX.
- For cases where the mouse, the keyboard, or the video adapter does not function properly, a text-based installation interface is also available. Some system administrators prefer the text-based installation to the GUI-based one.

Follow these high-level steps to install ESX from CD using either the graphical or the text-based installation option:

1. Insert the ESX boot CD into the CD-ROM drive and power on the system; ensure the system is configured to boot from CD. The system commences its boot process and the installation mode selection page is displayed.
2. Press Enter to start the graphical installer or type “esx text” and press Enter to start the text-based installation.
3. Accept the VMware license agreement.
4. Select the appropriate mouse and keyboard.
5. Add any necessary custom drivers.

6. Select “Yes” to load the system drivers.
7. Enter a serial number, or leave blank to centrally administer serial numbers in vCenter Server. The system will run in evaluation mode for 60 days.
8. Select the correct network adapter and specify a VLAN ID if required.
9. Configure the network settings on the next screen. Configure the ESX host IP address by entering a static IP address or a DHCP-based IP address. (VMware recommends using a static IP address.) Specify a host name and an optional VLAN ID as well. It is recommended to test the chosen settings.
10. Select the Advanced setup type. This enables customization of storage partitions.
11. ESX is installed on the first drive that is seen by the system by default, either a local SCSI drive or a Fibre Channel/iSCSI LUN. If selecting the latter, you perform a boot-from-SAN installation. This drive (or LUN) is initialized and partitioned during the installation process. Keep in mind that during initialization the contents of the disk will be erased.
12. Each VMware ESX Server will store its Service Console VMDK on a VMFS volume. This VMFS volume will by default be named “Storage1”. We recommend preceding Storage1 with the host name of the server. This makes the volume unique within vCenter Server.
13. The default partitioning scheme is basic and only contains partitions for swap, /var/log, and root. Our preferred disk-partitioning scheme for ESX is illustrated in Table 1.

/dev/sda (Primary Partition)			
Mount Point	Partition Type	Size	Description
/boot	ext3	1100 MB	Holds the VMkernel binaries
N/A	vmkcore	150 MB	Pre-configured
N/A	VMFS	Remaining	For local VMFS volume
esxconsole.vmdk			
/	ext3	5120 MB	The root partition for the Service Console
/tmp	ext3	2048	Used to store temporary files
N/A	swap	1600 MB	2x maximum Service Console memory (800 MB)
/home	ext3	2048 MB	Used for storage by individual users
/opt	ext3	2048 MB	Used for log files from 3rd-party agents
/var/	ext3	5120 MB	Separate to avoid overfilling root with log and core files

Table 1: ESX Server 4.x partition table

14. Partitioning changed with VMware vSphere. The Service Console is isolated in a VMDK which is hosted on a VMFS volume. During installation only three physical partitions will be created: physical /boot,

vmkcore, and a VMFS volume. This VMFS volume will hold the Service Console VMDK which will, in turn, hold all remaining virtual partitions.

15. Specify on the next screen that the ESX is to boot from the drive.
16. Provide boot options for the installer:
    a. General kernel parameters—Use this option to add default boot options to the boot command. These options are passed to the ESX kernel every time it boots.

    b. Force LBA32—This option is used to exceed the 1024 cylinder limit for the /boot partition. This option is only needed for legacy hardware.
17. Select the time zone.
18. Manually set the time or configure NTP servers.
19. Choose a secure root password and enter it twice in the fields provided. (Secure passwords usually consist of at least eight characters of which a combination of letters, numbers, and symbols is preferred.)
20. Confirm your installation configuration and click Next. The installer starts the installation of the ESX software. Progress bars appear to show the status of the installation, and a dialog box provides a notification of completion.
21. Click “Finish” to exit the installation.

Kickstart-Scripted Installation

ESX software can also be installed by leveraging Red Hat’s Kickstart installation method. Kickstart allows for the quick installation of ESX without having to go through all the installation screens. The workflow for a Kickstart-based install of ESX is straightforward.

1. Boot the server on which ESX is to be installed, either over the network using PXE (Pre-boot Execution Environment), a boot floppy, or a boot CD.
2. The boot program reads a specially created Kickstart configuration file and determines the preferred installation method (FTP, NFS, or HTTP). The Kickstart file automatically answers prompts for information such as network parameters and partition sizing.
3. The installation procedure continues unattended until ESX is completely installed.

We generally recommend using a solution that is easy to manage and repeatable. Examples of this would be the Ultimate Deployment Appliances and the ESX Deployment Appliance. EDA can be found at http://www.vmware.com/appliances/directory/89313/, and UDA can be found at http://www.ultimatedeployment.org/. Both are appliances that can be deployed within a matter of minutes. Examples of scripted installations can be found at http://communities.vmware.com/thread/90424?tstart=120.

See Host Profiles for a vSphere advanced tool for provisioning ESX hosts and maintaining host consistency within the infrastructure.

VMware vCenter Installation

A VMware vCenter Server configured with the hardware minimums can support 15 concurrent clients, 50 ESX hosts, and 250 powered-on virtual machines. A single 64-bit

Windows OS-based quad-processor VMware vCenter Server with 8GB RAM can scale to 30 concurrent client connections, 300 ESX hosts, and 3000 virtual machines. With Linked Mode enabled, each environment can scale up to 10,000 powered-on virtual machines and 15,000 registered virtual machines. The current limit is 10 vCenter Servers.

VMware vCenter Server version 4 has the following prerequisites for successful installation and usage (see the documentation for the version you plan to install).

Hardware Requirements

- Processor: 2GHz or faster Intel or AMD x86 processor. Processor requirements can be larger if your database is run on the same hardware.
- Memory: 3GB RAM minimum. RAM requirements can be larger if your database is run on the same hardware.
- Disk storage: 2GB minimum, 3GB recommended. You must have 601MB free on the destination drive for installation of the program, and you must have 375MB free on the drive containing your temp directory. Additionally, you will need 1.13GB free on the c:\ drive for Microsoft .NET 3.0 SP1, Microsoft ADAM, and Microsoft Visual C++ 2005 Redistributable.
- Database storage requirements: The demonstration database, using Microsoft SQL Server Express 2005, requires up to 2GB free disk space to decompress the installation archive. However, approximately 1.5GB of these files are deleted after the installation is complete. Note that MS SQL and Oracle are recommended for production use. SQL Express is supported for production use but should only be used for small environments (up to five hosts).
- Networking: Gigabit recommended.

Software Requirements

The VMware vCenter Server is supported as a service on these operating systems:

- Windows XP Pro SP2
- Windows Server 2003 (all releases except 64-bit), Windows Server 2003 R2
- Windows 2008

Database Configuration

VMware vCenter Server supports the following database formats:

- Microsoft SQL Server 2005 (SP1, SP2, and SP3 only) and Microsoft SQL Server 2008
- Oracle 9iR2, 10gR1 (versions 10.1.0.3 and higher only), 10gR2, 11g
- Microsoft SQL Server 2005 Express
- With Update 1, vSphere includes support for IBM DB2

Each database requires some configuration adjustments in addition to the basic installation. Please see the ESX and vCenter Server installation guide for further details.

The VMware vCenter Server database can coexist with the VMware vCenter Server on the same physical or virtual machine provided there are sufficient compute resources to handle the load. However, in an enterprise production setting, it is advisable to host the VMware vCenter Server database on an enterprise database server that is monitored, maintained, and regularly backed up. This ensures prompt recovery in case of data loss or database failure.

License Server Configuration

The license server as it existed in VI3 has been deprecated and is only needed for backwards compatibility in mixed environments. VMware vSphere licenses are maintained within vCenter and locally cached on each host. This avoids the unnecessary downtime that occurred in the VI3 timeframe when the license server was unavailable for longer than 14 days. An evaluation license is bundled with vSphere which provides 60 days of use.

Installation

The VMware vCenter Server is a Windows executable and is installed as a Windows service. The server can be installed on a physical or virtual machine. While installing the VMware vCenter Server, you may opt to install the VMware vCenter Server database on the same physical or virtual machine as the VMware vCenter Server.

The user installing VMware vCenter Server requires administrative privileges on the machine. For detailed installation instructions, see the ESX and vCenter installation guide. After the VMware vCenter Server is installed, the features specific to VMware vCenter Server can only be enabled by adding the appropriate licenses. Unlike ESX, which can be enabled by a host license file, the VMware vCenter Server features can only be used with licenses served by vCenter itself.

A default installation of VMware vCenter includes VMware vCenter Orchestrator. vCenter Update Manager, vCenter Guided Consolidation, vCenter Converter Enterprise, and the vCenter client can be installed from the autorun.exe application available on the installation CD.

The look and feel of VMware vCenter Server through the vSphere Client has been vastly improved by introducing a new home screen which simplifies administration.

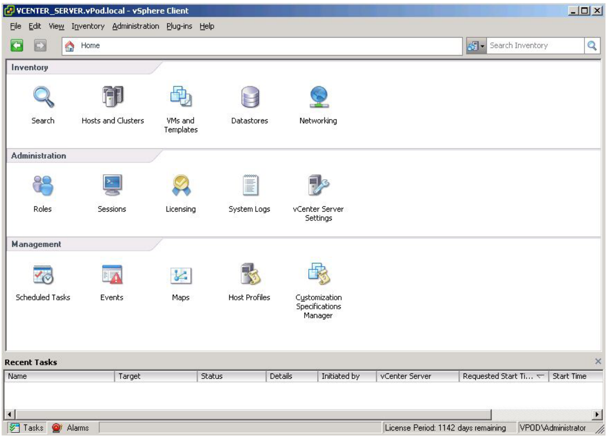


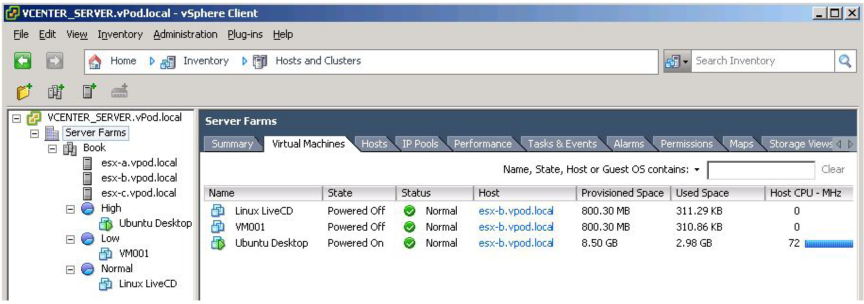
Figure 24: vSphere Client view of inventory, administration, and management

7. Managing VMware vSphere

VMware vCenter Server

VMware vCenter Server is the VMware vSphere management software. It provides a single point of control for all VMware vSphere resources including ESX hosts, virtual machines, virtual networking, and virtual storage.

VMware vCenter Server is accessed through the VMware vSphere Client. The interface provides administrators with a view of the entire virtualized datacenter and the means to manage its resources.



Name	State	Status	Host	Provisioned Space	Used Space	Host CPU - MHz
Linux LiveCD	Powered Off	Normal	esx-b.vpod.local	800.30 MB	311.29 KB	0
VM001	Powered Off	Normal	esx-b.vpod.local	800.30 MB	310.86 KB	0
Ubuntu Desktop	Powered On	Normal	esx-b.vpod.local	8.50 GB	2.98 GB	72

Figure 25: vSphere Client view of vCenter Server operations

ESX provides a robust virtualization layer, but it is VMware vCenter Server that provides the features required for a complete infrastructure management system. VMware vCenter Server enables virtual machine mobility, centralized management, centralized monitoring and alerting, advanced quality of service, and high availability functionality which moves ESX from a tactical solution to an enterprise strategic infrastructure platform. The following sections describe the VMware vCenter Server features in greater detail.

VMotion

A virtual machine can be migrated between two ESX hosts in one of two ways. The first option is to power off the virtual machine, move it to another host, then power it back on. This is called a *cold migration*. The second option is to move a virtual machine between hosts using VMotion.

VMotion offers huge benefits to datacenter administrators, who can now move running virtual servers to another host when it is necessary to perform maintenance on an ESX host. With the help of this technology, each virtual machine can be considered as a workload, and the workloads can be load-balanced across physical hosts. DRS technology does this automatically, utilizing VMotion to load-balance a VMware vSphere cluster.

One of the prerequisites for VMotion is that the virtual disk and the other files that encapsulate the virtual machine exist on shared storage such as SAN, iSCSI, or NAS. Also, these files must be visible to both physical hosts between which the virtual machine is being migrated.

Successful migration of virtual machines between hosts using VMotion requires identical CPU instruction sets on both hosts, accomplished with either utilizing hosts with the same family of CPUs or using Enhanced VMotion Compatibility (EVC) on a cluster of hosts with technologies enabling VMotion compatibility with older servers. More information on this and other VMotion requirements is available on the VMware Web site.

VMware vCenter enables VMotion migration or cold migration by dragging the virtual machine and dropping it on the destination host. A wizard opens to provide guidance through the migration process. Another option is to use the context menu by right-clicking on the virtual machine you want to move and selecting “migrate.” This also opens the migration wizard for guidance through the process.

Virtual Machine Provisioning

Physical servers can be virtualized using a variety of P2V tools and techniques. Administrators can also opt to build a virtual server from scratch by creating a new virtual machine, installing the operating system, installing all required services and applications, and patching the operating system and applications.

After creating a virtual machine by virtualizing a physical server or by building a new virtual server in the virtual environment, VMware vCenter Server enables marking a virtual machine as a template, which is a virtual machine that cannot be turned on. Template creation takes place after installation of the required applications is complete and is accomplished by creating a VM instance, right-clicking on the virtual machine in the VMware vCenter Server, and selecting “Convert to Template.” An unlimited number of virtual machines can be created from this template by right-clicking on the template and choosing “Clone to Virtual Machine.” This method of creating a virtual machine is called *provisioning from a template.* The guest operating system (computer name, domain, and IP address) can be customized when provisioning from a template by manually entering the customization information or by using Sysprep. For more information, see the *Basic System Administration* guide.

Provisioning from a template is an invaluable VMware vCenter feature and significantly reduces the time required to create a new server. Administrators can create different templates for different purposes. For example, a Windows 2003 Server template might be created for the finance department, a Red Hat Linux template for the engineering department, and a Windows 2008 Server template for the marketing department. The administrator can quickly provision a correctly configured virtual server on demand.

The ease and flexibility of template creation bring the risk of virtual machine sprawl, where virtual machines are provisioned so rapidly that documenting and managing the virtual machine lifecycle becomes a challenge. Many VMware vSphere 4 customers are implementing change management processes in the virtual environment. It is helpful to have virtual machine naming conventions that help identify each machine’s purpose and users. VMware Lifecycle Manager can significantly ease the burden with both workflow approval processes and by assigning limited lifecycles to certain VMs or categories of VMs.

Virtual Machine and ESX Monitoring

VMware vCenter Server provides a way to monitor the performance statistics of the virtual machines and ESX hosts. The CPU, memory, network, and disk parameters displayed on the Performance tab of the vSphere Client can be customized and graphed from short-term to yearly time frames.

Performance monitoring from within the virtual machine yields incorrect information, because each virtual machine sees only its share of the resources as dictated by the ESX kernel (hypervisor). Consequently, any performance monitoring should be done through ESX. Viewing the performance at a VMware vCenter Server level enables viewing of metrics for all of the ESX hosts (and their virtual machines) in the environment. The level of detail and estimated dataset size for performance monitoring can be set by selecting the Statistics setting in the VMware vCenter Server Configuration window accessible from the Administration menu in the VMware vSphere Client.

Alarms can be set to trigger when certain metrics exceed a set threshold. For example, an alarm can be defined to be triggered when CPU utilization on a host reaches 90% or when the health of the host hardware has changed. New alarms can be created by selecting the object (ESX host or virtual machine), clicking on the Alarms tab, right-clicking on the Alarms panel, and selecting New Alarm.

The action that VMware vCenter Server should take in case an alarm is triggered can also be defined. The available actions for an ESX host are to send a notification email message, send a notification trap, enter maintenance mode, exit maintenance mode, enter standby, exit standby, reboot, or run a script. When creating an Alarm at a virtual machine level, the virtual machine can be suspended, powered off or reset. Using the Alarms feature avoids the need for third-party tools that run in the virtual machine or in the service console of ESX, consuming CPU cycles.

Intelligent Clusters

VMware vCenter Server 4 comes with the ability to create a cluster, a logical group of servers that can be managed together and work in tandem for load balancing or disaster recovery.

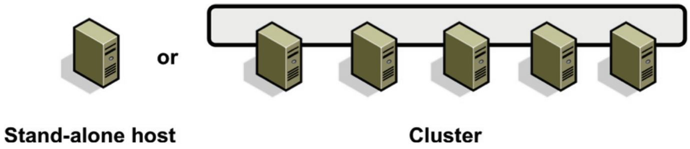


Figure 26: Stand-alone host versus a vCenter ESX cluster

As servers are added to a cluster, the available resources in the cluster accumulate. Virtual machines created on an ESX host are displayed in the cluster, not in the individual host. This is an important distinction. Although a virtual machine in a cluster can exist on only one host at a time, the virtual machine can execute on any ESX host in that cluster. VMware vCenter Server can move the virtual machine (or workload) using VMotion initiated by DRS, or VMware HA, to a different host for the purpose of load balancing or recovering from a physical host failure.

A cluster in VMware vCenter Server can be configured to balance its load among hosts by moving its virtual machines as necessary. VMware DRS uses VMotion to distribute the workload in the cluster. DRS can be configured to automatically use VMotion to migrate its virtual machines between hosts when the need to load balance occurs or to prompt the administrator with a recommendation to move the virtual machine over to the new host. The former configuration is called *fully automated mode* and the latter is called *manual mode.*

VMware High Availability is another feature of VMware vCenter Server clusters. A cluster with VMware HA enabled automatically reacts to a host isolation or host failure by powering on a virtual machine on an alternate good host. This ensures that, in case of an unplanned host failure, the virtual machines are brought back to life without manual intervention.

VMware HA does not use VMotion to migrate virtual machines to a good host. The virtual machines are restarted on the alternate host, so the virtual machine users will notice a temporary service interruption. VMware HA is a reactive feature triggered by ESX isolation or failure, while DRS is a proactive approach to keep the load on the hosts balanced at all times. VMware HA, once configured, does not require VMware vCenter Server to operate. The ESX hosts communicate directly with the other ESX hosts in the cluster to determine whether a particular ESX host is offline. VMware vCenter Server can be run within a virtual machine and the VMware HA technology will still function and restart the VMware vCenter Server along with the other virtual machines.

VMware HA responds to host isolations and host failures. VMware vSphere offers two high availability features which enable per-virtual-machine-level high availability. VM Monitoring is an integral part of VMware HA. VM Monitoring restarts individual virtual machines when VMware Tools heartbeats are not received within the given threshold. Like VMware HA, this is a form of stateless clustering, meaning that the virtual machine will be restarted instead of a seamless failover.

VMware FT provides a stateful solution; that is, when VMware FT is enabled on a specific machine and the host that is running the primary virtual machine fails, the secondary virtual machine seamlessly takes over. VMware FT provides continuous availability by keeping a shadow copy of a virtual machine in lockstep with the primary virtual machine. Like VMware HA and VM Monitoring, VMware FT also uses heartbeats to monitor the status of a virtual machine and instantly replaces the primary with the secondary when heartbeats are lost.

A VMware vCenter Server cluster can be configured for both VMware DRS (including DPM) and VMware HA (including VM Monitoring) and can recover from unplanned host isolation and host failures. Within a VMware vCenter Server cluster, VMware FT can be enabled for virtual machines which meet all prerequisites and which need continuous availability.

Managing Resource Pools

VMware vCenter Server enables carving of resources for a set of virtual machines. For example, 25% of the CPU resources and 30% of the memory resources might be assigned in a cluster to a set of staging virtual machines. This reservation is called a *resource pool*. An additional resource pool can be created on the same cluster with 70% of the CPU resources and 60% of the memory resources for production virtual machines.

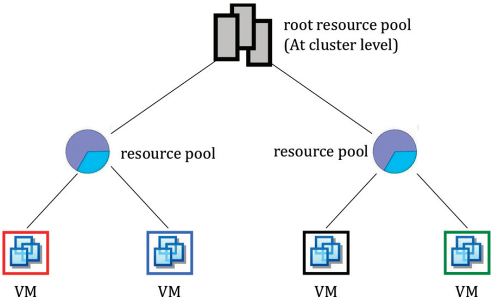


Figure 27: VMs assigned to resource pools representing shares of the physical resources

This configuration ensures that the staging virtual machines never use more than 25% of the CPU resources on the cluster. A rogue staging virtual machine that is consuming excess CPU cycles cannot affect the production virtual machines.

Resource pools also help organizations that use an IT chargeback model to limit VM resource use by a particular department to only those resources allocated to that department.

The combined effect of resource pools, DRS, DPM, and VMware HA is a radical shift in how administrators think about their datacenter. The clusters in a datacenter can now be considered a pool of compute resources. Administrators can carve out portions of these resources for a set of virtual machines depending on their service level requirements.

VMware Capacity Planner

VMware Capacity Planner is described in the section “Identifying Virtualization Candidates” in Chapter 5. This is an add-on module for VMware vCenter Server 4 which collects comprehensive resource utilization data in heterogeneous IT environments. It compares the information gathered with industry-standard reference data to provide both analysis and decision support modeling. Scenario modeling provides guidance in the design of a virtual infrastructure.

VMware Update Manager

VMware Update Manager is part of the VMware vSphere 4 product suite. It supports a streamlined, automated update mechanism for ESX hosts and virtual machines. See the VMware Update Manager section later in this chapter.

VMware Guided Consolidation and VMware Converter

VMware Converter is described in the section VMware Guided Consolidation and VMware Converter in Chapter 8. This is an add-on module for VMware vCenter Server 4.

Distributed Power Management

VMware Distributed Power Management (DPM) is a new technology as of ESX 3.5 which supports dynamic utilization of ESX hosts. The initial release was considered experimental; the current version is fully supported. Based on virtual machine resource requirements using DRS, ESX hosts are powered on or off to support virtual machine requirements and maximize power savings during periods of lower resource requirements. This enables a minimal number of ESX hosts to be running to support the resource requirements.

Infrastructure Management with VMware vCenter

Effective management tools and processes are essential components of a VMware vSphere environment. Hardware selection and configuration determine, to a large extent, the performance and reliability of the infrastructure.

Ongoing change is common in virtualized environments. One of the strengths of a VMware vSphere environment is the ability to create, change, and reallocate virtual resources in real time independent of the management of the underlying physical hardware. Some of these changes may happen automatically in order to maintain the environment. This dynamic environment and ease of change demand a comprehensive tool set capable of providing administrators with intuitive interfaces that enhance and integrate with standard support processes and tools.

VMware vCenter Server provides administrators with a single management system for controlling and monitoring the virtual servers, networks, storage, and physical host servers as a unified infrastructure. The ability to consolidate and standardize infrastructure management into unified tool and set processes is one of the driving features for infrastructure agility. Integrating all infrastructure management tasks through the use of roles and views helps drive the long-term ROI and change in an organization's service support and management. A poorly managed virtual infrastructure can erode the ROI and undermine the stability of the infrastructure.

The following are some of the VMware vCenter interfaces that are routinely used for VMware vSphere management tasks.

Datacenters and Clusters

The Hosts and Clusters page is available from the home page in VMware vSphere Client. This interface is where the physical and logical organization of the VMware virtual infrastructure occurs. Datacenters, folders, and clusters are containers for organizing VMware vSphere components. The datacenter object is used to represent a physical datacenter or location. Each datacenter created is the root of a hierarchy of folders or containers organizing the hosts, clusters, virtual machines, networks, and datastores within it.

The datacenter is a logical isolation boundary for a set of hosts, networks, datastores, and datacenters online with VMotion or offline with a cold migration. Templates are specific to a datacenter and cannot be shared, and all object names within a datacenter must be unique.

VMware vSphere clusters are created within a datacenter object to aggregate the resources and management of a number of ESX hosts. Clusters are the main management containers in VMware vCenter. Virtual machines and resources are distributed and managed within clusters. DRS (including DPM) and VMware HA services are managed and activated at the cluster level. Hosts are placed in a cluster, where they are managed as a logical unit. Host resources are aggregated within the cluster.

Roles and Permissions

Centralized security and permissions for the VMware vSphere infrastructure are stored and managed in the VMware vSphere Client through the Permissions tab. User accounts and permissions for local host access can be assigned to the service console or the VMware vSphere Client when connected directly to an ESX host.

Roles are used to define privileges—that is, a set of individual rights to read properties or to perform actions within VMware vCenter. Users and groups from a Windows Active Directory are placed in roles. Permissions are created by assigning roles to VMware vCenter Server objects. Members of the VMware vCenter Server local Windows administrators group have the administrator role in the VMware vCenter Server.

The propagation of permissions through hierarchies is optional and is set for each permission rule. Permissions set on a VMware vCenter Server object override permissions set on the parent object.

VMware vCenter Server and ESX offer the following predefined roles:

- No access
- Read-only user
- Administrator
- Virtual machine user
- Virtual machine power user
- Resource pool administrator
- VMware Consolidated Backup user
- Datastore consumer
- Network consumer

Roles can be created, edited, removed, renamed, or cloned using the VMware vSphere Client through the Roles page on via the Home view by right-clicking in the Roles panel and choosing Add Role.

The combination of VMware vCenter roles and the nested hierarchies of VMware vSphere and DRS makes it simple to delegate management of virtual resources to groups of users. This is an effective way to provide localized resource and service-level control to groups of users while reaping the benefits of infrastructure consolidation.

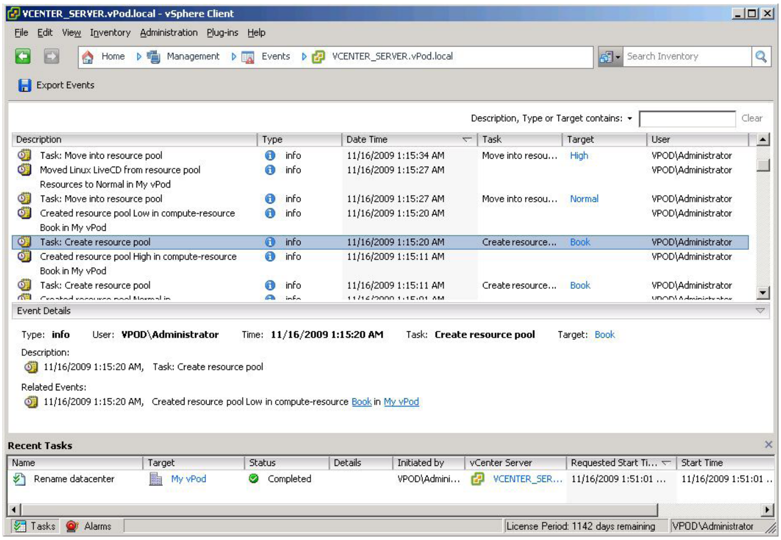
The most effective way to manage permissions through VMware vCenter is to create resource pools for groups of virtual machines for management and resource sharing, and Active Directory groups for each resource pool. By assigning each Active Directory group permissions to the appropriate resource pools, VMware vCenter permissions can be managed through standard Active Directory security practices.

Events

Any event of interest in a VMware vCenter Server or an ESX host triggers an *event message.* Event messages are stored in the VMware vCenter Server database and are viewable in the VMware vSphere Client. The Events page, accessed via the Home page, on the displays all of the events in the VMware vSphere infrastructure, and the Events tab in the inventory for any object in the hierarchy enables you to see the events for that particular virtual machine, host, or datacenter.

Events are displayed in scrolling chronological order and are color-coded for Information, Error, or Warning messages. Event details can be seen by selecting the specific event and viewing the Event Details window.

The VMware vSphere Client enables greater visibility and event management by allowing sorting of event columns or searching for events. These features are necessary when reporting on events in large, consolidated environments. Events are searched for using the event-filtering feature in the VMware vSphere Client, allowing individual events to be identified in a contextual view.



Description	Type	Date Time	Task	Target	User
Task: Move into resource pool	info	11/16/2009 1:15:34 AM	Move into resou...	High	VPOD\Administrator
Moved Linux LiveCD from resource pool Resources to Normal in My vPod	info	11/16/2009 1:15:27 AM			VPOD\Administrator
Task: Move into resource pool	info	11/16/2009 1:15:27 AM	Move into resou...	Normal	VPOD\Administrator
Created resource pool Low in compute-resource Book in My vPod	info	11/16/2009 1:15:20 AM			VPOD\Administrator
Task: Create resource pool	info	11/16/2009 1:15:20 AM	Create resource...	Book	VPOD\Administrator
Created resource pool High in compute-resource Book in My vPod	info	11/16/2009 1:15:11 AM			VPOD\Administrator
Task: Create resource pool	info	11/16/2009 1:15:11 AM	Create resource...	Book	VPOD\Administrator
Created resource pool Normal in	info	11/16/2009 1:15:01 AM			VPOD\Administrator

Type: **info** User: **VPOD\Administrator** Time: **11/16/2009 1:15:20 AM** Task: **Create resource pool** Target: Book

Description:
11/16/2009 1:15:20 AM, Task: Create resource pool

Related Events:
11/16/2009 1:15:20 AM, Created resource pool Low in compute-resource Book in My vPod

Name	Target	Status	Details	Initiated by	vCenter Server	Requested Start Ti...	Start Time
Rename datacenter	My vPod	Completed		VPOD\Admini...	VCENTER_SER...	11/16/2009 1:51:01 ...	11/16/2009 1:51:01 ..

License Period: 1142 days remaining VPOD\Administrator

Figure 28: vSphere Client view of events

Alarms

VMware vCenter Server and ESX provide a comprehensive set of tools for monitoring and triggering alarms in response to conditions or events within the virtual infrastructure. Default alarms are included in VMware vCenter, but alarms are not directly configurable through the VMware vSphere Client connected to ESX. In addition to the default alarm settings in the VMware vCenter Server, a rich interface is available for creating additional alarms that can alert on some or all inventory objects by being applied to objects or containers in the VMware vSphere hierarchy. VMware vSphere introduced many new events, triggers, and actions which are invaluable for system administrators. VMware vSphere makes it possible to initiate maintenance mode when the hardware health status of a host degrades. When a host is placed in maintenance mode, all virtual machines will automatically be VMotioned to all other available hosts in the cluster and avoid possible downtime because of hardware outage.

VMware vCenter Server Alarms Tab

VMware vCenter Server provides two main views for setting and monitoring alarms, the Alarms view and the Definitions view. These views are available from the VMware vSphere Client connected to a VMware vCenter Server, Home: Hosts & Clusters > Host > Alarms tab: Alarms or Definitions buttons.

The Alarms view displays any active alarms that are triggered. The Definitions view is where existing alarms are listed and where alarms can be edited or added.

Alarms trigger actions when the event defined in the alarm occurs on a host or virtual machine. Alarms are applied to datacenters, folders, clusters, resource pools, hosts, or virtual machines to display or notify regarding the status of one or more of these objects. Alarms defined within the hierarchy are inherited by all of the object's child items. This cascading of alarms within the hierarchy cannot be prevented.

Alarms can be defined or edited by a user with the appropriate permissions on datacenters, hosts, and any virtual machines within the scope of the alarm definition, that is, the object the alarm is applied to and the flows to its child objects.

Alarm triggers define conditions that, when met, cause an event to be activated. Alarm triggers can be defined as percentages above or below host or virtual machine memory or CPU utilization, or by a virtual machine's heartbeat. Alarms can also be triggered based on the state of a host or virtual machine. State alarms define a state as "Is" or "Is Not" a particular condition.

Examples of virtual machine and host condition states are:

- Virtual machine powered on
- Virtual machine powered off
- Virtual machine suspended
- Host powered on
- Host powered off
- Host suspended

New condition states can be defined for datastores:

- Connected to all hosts
- Disconnected from all hosts

Alarm notifications are actions taken when an alarm is triggered. Notification actions available in VMware vCenter Server include the following:

- Send an email notification message.
- Send an SNMP notification trap.
- Run a command.
- Enter maintenance mode.
- Exit maintenance mode.
- Enter standby.

- Exit standby.
- Reboot host.
- Suspend a virtual machine.
- Power off a virtual machine.
- Reset a virtual machine.

VMware vCenter Server alarms and notifications provide a rich set of tools for notifying and reacting to alarm events within the infrastructure, and they allow seamless integration of a VMware vSphere infrastructure with IT Service Management tools and practices.

Virtual Machine Deployment

Virtual machine deployment and management are among the most common management tasks for VMware vSphere administrators. The VMware vSphere Client provides interfaces for virtual machine management that are integrated with the alerting, monitoring, and organizing views.

Virtual machines can be created by manually installing an operating system, by using a preconfigured template, by importing an appliance, by cloning an existing virtual machine, or by importing a physical server or a virtual server from another hosting platform.

A simple wizard-based interface is used to create a new virtual machine manually from a template or by cloning an existing virtual machine. Virtual machine deployment is an area where updating organizational processes to accommodate the VMware vSphere infrastructure is important to maintaining availability and provides significant operational advantages.

Resource groups and distributed permissions give customers or groups within the organization the ability to create and manage virtual machines within their resource group. Moving these tasks closer to the owners by making the Create New Virtual Machine wizard available (through the VMware vSphere Client or, preferably, through an online service catalog or portal) reduces overhead on the infrastructure management team and increases the speed and accuracy of server request fulfillment.

Reducing server deployment lead time and distributing access to the infrastructure require enhanced capacity and change management functions. The VMware vSphere Client provides simple centralized monitoring of the VMware vSphere infrastructure capacity and the ability to non-disruptively add and distribute additional capacity throughout the infrastructure. Organizations with large virtual infrastructures need to merge this manageability and agility with their service and support processes for maximum benefit.

The information needed to build a virtual server is similar to the information gathered in project planning for physical server acquisition. New virtual servers can be created from the Summary tab or from the context menu of host servers, clusters, or resource pools in the VMware vSphere Client. The Create New Virtual Machine wizard performs the following information-collecting activities for virtual server creation:

- Selects a virtual machine name. This is not the domain name of the server, but they can be the same. The name must be unique within the datacenter context; it is used to name the virtual machine's files.

- Selects a location for the virtual machine in a folder, resource pool, or datacenter.
- Selects a datastore for the virtual machine files.
- Identifies the guest operating system to be installed on the virtual machine.
- Sets the virtual disk size. A virtual disk can be from 1MB to almost 2TB in size and either thin or thick. Additional virtual disks can be added later. The size of existing virtual disks can be changed through the VMware vSphere Client. When cloning an existing virtual machine or creating one from a template, only the information that differentiates the new virtual machine from the template or existing server is required.

Creating a virtual machine using the custom option includes several more detailed options: amount of vCPUs, amount of memory, amount and type of network adapters, type of SCSI adapter, LSI Logic Parallel or SAS, BusLogic Parallel, or VMware Paravirtual. Select the type of virtual disk: new, existing, or a mapped LUN from a storage network. When creating a new disk, virtual disk mode can be specified. The disk modes allow for *persistent* disks, which commit changes to disk immediately, and *non-persistent* disks, on which changes are discarded when the server is powered off or reverted to a snapshot.

To map a raw SAN LUN to a virtual machine:

- Select a target LUN visible from the SAN.
- Choose a datastore.
- Choose physical or virtual disk compatibility mode.
  - **Physical compatibility mode** gives the guest operating system direct access to the disk hardware. Disks in physical compatibility mode have the restrictions of not being able to be cloned, migrated, or used for templates. This is used for application clustering technologies (MSCS and VCS) and for utilizing vendor-based storage management tools.
  - **Virtual compatibility mode** maps a SAN LUN to the guest OS, but retains the functionality of a virtual disk.

After completing the Create New Virtual Machine wizard, you can choose to edit the virtual machine setting before completion. This enables adding or removing specific devices or setting reservations and limits.

Guest operating systems can be installed into a new virtual machine by mapping a CD, DVD, or disk image to the virtual machine and booting to that device. After it is booted to the installation disk, operating system installation can continue as usual.

The boot order for a virtual machine can be altered by pressing F2 during initial post-test of the machine or by checking the “Enter BIOS on next boot” box in the virtual machine options. This might be necessary to enable boot from CD or DVD functionality.

After installation of the guest operating system, the VMware Tools service should be installed to allow full manageability of the virtual machine. VMware Tools includes drivers for the virtualized hardware and scripts to enable shutting down, suspending, and restarting of the virtual machine from the VMware vSphere Client.

Migration of Virtual Machines to Alternate Platforms

It is possible to migrate virtual machines from one virtualization platform to another. There are a number of ways to do this, depending on the source and target virtualization platforms.

VMware Converter handles migrations between ESX hosts, VMware Server, and VMware Workstation. VMware Converter can also import from other Intel/AMD-based virtualization platforms such as Microsoft Virtual Server and Virtual PC virtual machines from Symantec Backup Exec System Recovery, Symantec LiveState, StorageCraft ShadowProtect, Norton Ghost images, Acronis True Image, and VMware Consolidated Backup full backup images.

Hot Migrations (VMotion)

Virtual machines can be migrated between hosts within a datacenter for performance, resource balancing, or maintenance reasons. VMotion migrations, also known as *hot migrations*, allow a virtual machine to be moved from one virtualization platform to another with no downtime. The virtual machine migration wizard can be started from the context menu or the Summary page of a virtual machine. VMotion migrations can also be triggered by automated DRS operations to balance resource usage in resource pools.

VMware vCenter Server validates the virtual machine and its target destination host and resource pool to ensure a successful migration. Hot migrations of virtual machines can occur only when certain conditions are met, such as when the start host and destination host have compatible CPUs or Enhanced VMotion Compatibility has been enabled. Both hosts must have access to the same datastores and adequate resources to perform the migration without violating any resource pool admission requirements.

Migration validation warnings are issued if the virtual networks on the host and destination sites do not match. Non-matching networks during a VMotion migration cause the guest OS to be disconnected from the network when the migration is complete, requiring a new network and NIC to be assigned and connected at the destination side of the migration.

VMotion migrations fail validation if removable devices such as CDs, DVDs, or floppy drives are mounted and connected to host or client hardware. Disconnecting or unloading these devices eliminates the problem and allows VMotion migration to occur if all other requirements are met.

The speed and success of VMotion migrations depend upon the active load of the host and destination servers and the activity of the virtual machine to be migrated.

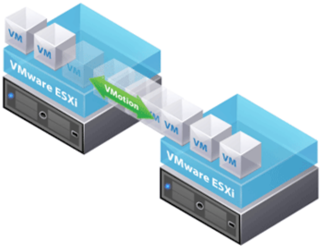


Figure 29: VMotion

Storage VMotion

Storage VMotion is based on the hot migration technology of VMotion, but also enables moving a virtual machine's virtual disk to a different datastore. This is available with ESX 3 and newer versions with the appropriate version of VMware vCenter Server. An example is migrating a running virtual machine from one ESX host to another while also migrating the location of the virtual machine's virtual disk to another storage location. Another example would be migrating the virtual disk for the virtual machine from one LUN to another without moving the virtual machine to another ESX platform.

Storage VMotion is supported for virtual machines stored on Fibre Channel and iSCSI SAN shared storage or NFS shares. It performs proactive, non-disruptive storage migrations to simplify array storage maintenance without virtual machine downtime, and it enables changing from thin-provisioned to thick-provisioned disks or vice versa. Storage VMotion can be used to eliminate storage I/O bottlenecks without impacting virtual machines, just as the standard VMotion eliminates CPU bottlenecks. Storage VMotion uses existing VMware technologies such as changed block tracking, introduced with VMware vSphere. The flow of a Storage VMotion migration is as follows:

1. The virtual machine's home directory (containing virtual machine configuration, swap, and log files) is copied to the new storage location.
2. Optionally: The virtual machine is migrated, using VMotion, to the target ESX host.
3. Change block tracking is enabled on the source VMDKs.
4. The parent virtual disk is copied from the old storage device to the new storage device via the Data Mover, which enhances performance and decreases overhead on both source and destination storage devices.
5. All remaining changed blocks are copied from the old storage device to the new storage device.
6. The source VM is suspended and the destination VM is resumed when the state (checkpoint data) is copied to start running on the new home and VMDKs.
7. The source VM home and VMDKs are deleted.

The entire process takes several minutes, depending on the size of the VMDKs of the virtual machine being migrated.

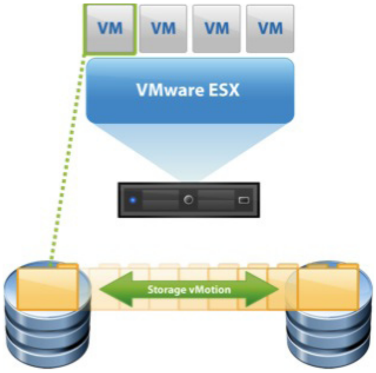


Figure 30: Storage VMotion

Cold Migrations

Cold migrations perform the same function as Storage VMotion for virtual machines that are powered off. Cold migrations allow the movement of a virtual machine that is powered off from one physical ESX host to an alternate physical ESX host. Cold migrations, unlike VMotion, allow movement between ESX hosts utilizing different vendor types and CPU families without masking specific CPU characteristics.

VMware Update Manager

VMware Update Manager provides an automated patch management system for VMware vSphere. The software manages the tracking and patching of ESX. It also handles select Windows and Linux virtual machines.

The VMware Update Manager eliminates the risk of manually tracking and patching ESX and virtual machines by the following strategies:

- Scanning the state of the physical VMware ESX hosts.
- Scanning the state of select guest operating systems running within the virtual machines.
- Scanning the state of select applications running within the virtual machines.
- Comparing the current state of each with baselines set by the IT team.
- Applying updates and patches for compliance with business requirements.
- Allowing for a snapshot prior to patching to enable a method to back out of the patch, if needed.
- Allowing patching of offline virtual machines.
- Working with DRS to eliminate disruptions during ESX host patching:
  - Hosts are placed in maintenance mode one at a time.
  - Virtual machines are migrated to other hosts prior to patching and are migrated back after patching is complete.

Updates to the ESX do not introduce virtual machine downtime.

VMware Update Manager provides notification of host patches and allows update scheduling. It also handles all of the virtual machine migration, patching, and rebooting, eliminating typical manual steps and downtime during updates in a physical environment.

VMware vCenter Orchestrator

VMware vCenter Orchestrator is an orchestration engine simplifying management with task automation. VMware vCenter Orchestrator relies on workflows to execute automated processes. These workflows can contain specific actions or decisions based on output of a specific task and can be reused at any given time, providing consistency in operational tasks. VMware vCenter Orchestrator provides a library of example workflows for common management tasks and individual actions.

VMware vCenter Linked Mode

VMware vCenter Server 4 introduces a new feature called Linked Mode. Linked Mode provides a single point of authentication for a multi-vCenter Server environment as it logs in simultaneously to all vCenter Servers. Linked Mode simplifies management of remote offices or multi-datacenter environment by offering a single pane of glass. Roles, permissions, and licenses are replicated across the virtual infrastructure by utilizing Microsoft ADAM. The most powerful function, however, is the search option as there are no boundaries for exploring your environment.

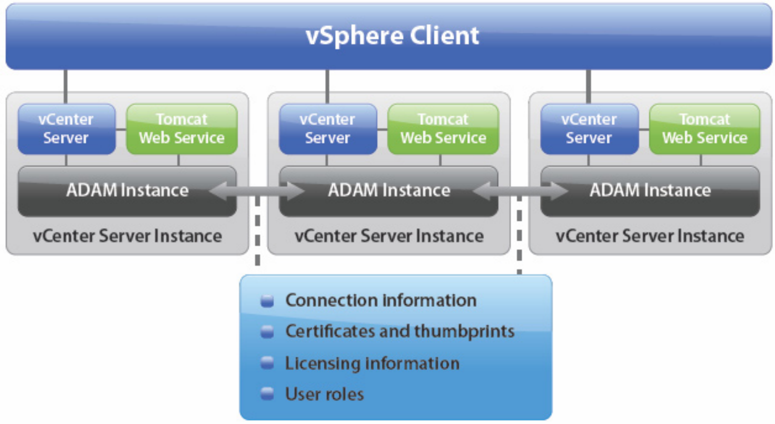


Figure 31: vCenter Linked Mode

VMware Host Profiles

VMware Host Profiles enables centralized compliance monitoring and reporting for host configuration. VMware Host Profiles can be compared to VMware Update Manager, as it provides a baseline for configuration items to which the environment needs to comply. Likewise, VMware Host Profiles can be used to quickly make host configuration changes to any environment. A single host can be reconfigured and used to capture a blueprint, which, in its turn, can be applied to a datacenter, cluster, or a single host. VMware Host Profiles also assists in maintaining the integrity of the host servers in the face of normal changes over time.

Host Profiles consists of several sub-profiles. Each sub-profile contains policies which describe the configuration for a specific profile. Examples of these policies are:

- Firewall rule sets
- DNS settings or updates
- Service console memory
- NTP configuration
- Virtual switch configuration, including NIC teaming and security policies
- Additional users or user groups

8. Migrating Virtualization Candidates

Conducting a virtualization assessment helps ensure virtualization project success. After candidate identification, the migration time can be estimated based on network bandwidth, storage requirements, and application complexities.

VMware Physical-to-Virtual Process

The physical-to-virtual (P2V) process is the method used to convert and migrate a physical server to a virtual machine. VMware uses P2V migration software to handle data migration and hardware reconfiguration. The process is similar to migrating an OS-application pairing from one physical platform to another, but instead of changing the hardware devices and drivers to another physical set, the migration changes them to a virtual set.

For native virtualization, the virtual set consists of CPU, RAM, SCSI disks, AMD PC-Net Adaptive NICs, parallel and serial devices, CD-ROM, and floppy devices.

For hosted virtualization, the virtual set consists of all the hardware that the underlying host's OS can see. This provides a greater number of devices, including USB and Wintel modems. Physical machines that are better suited for a hosted virtualization deployment include fax systems and RAS servers, although over-IP solutions are typically used in virtualized environments.

Migration checklist:

1. Migrate the OS disk.
2. Reconfigure the OS with new devices.
3. Install VMware Tools.
4. Check for and remove unused devices.
5. Check for and disable unused services.
6. Clean up legacy log information but not warnings or errors that may assist in troubleshooting application or OS issues once the migration is complete.
7. Migrate the data disk(s).
8. Reboot.
9. Check logs.
10. If clean, then done. If not, go to step 4.

VMware Guided Consolidation and VMware Converter

VMware Guided Consolidation is an optional vCenter Server component and is used for planning physical-to-virtual machine migrations. VMware Guided Consolidation leverages VMware Capacity Planner and VMware Converter technology and is designed for small-scale consolidation.

VMware Converter can convert and migrate physical machines, as well as third-party disk image formats, to VMware virtual machines. It also converts and migrates virtual machines from one VMware platform to another.

Local and remote physical machines can be migrated hot or cold. When performing hot conversions or migrations, there is no downtime for the original physical machine. Multiple simultaneous conversions and migrations are supported with VMware Converter.

The VMware Converter supports migration from Microsoft Virtual PC and Microsoft Virtual Server to VMware virtual machines. It also supports migrating from backup images such as Symantec Backup Exec LiveState Recovery or Symantec Norton Ghost, as well as restoring VMware Consolidated Backup images to running virtual machines.

VMware Converter supports the import of OVF-based virtual machine images.

VMware Converter Starter and VMware Converter Enterprise both support hot cloning, local conversions, and remote conversions. The differences between the starter and enterprise versions of VMware Converter include:

- The enterprise version allows for cold cloning with a special bootable CD, multiple simultaneous conversions, automation and management of large-scale conversions, and remote conversion to all destinations.
- The starter version is free, allows single conversions, and limits remote conversion to VMware Workstation, VMware Server, VMware Player, and VMware GSX Server.

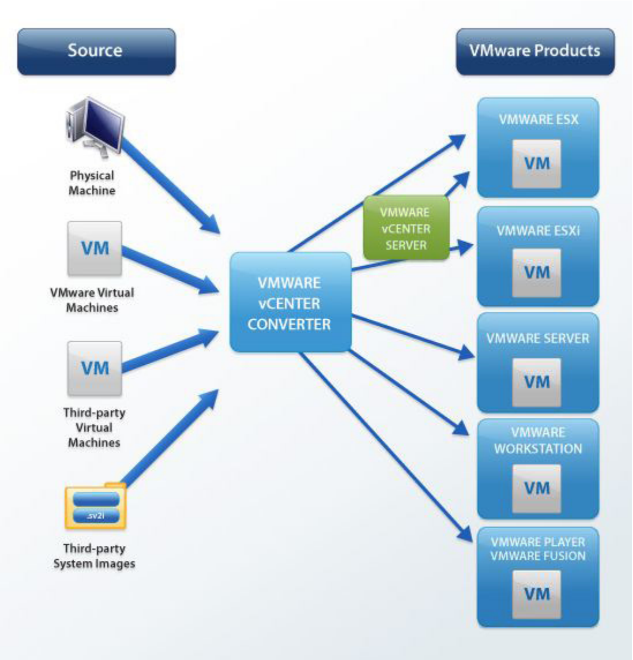


Figure 32: VMware vCenter Converter

Third-Party Migration Tools

Several third-party migration tools can be used for migrations. These include commercial products such as Platespin PowerConvert, Leostream P>V Direct, and Ultimate-P2V (free P2V migration tool based on BartPE Boot-CD).

Manual Migration

Manual migration from physical to virtual machines is possible. Due to the time involved, most businesses choose to utilize automation tools. The manual process is documented in white papers.

A manual migration follows these high-level steps:

1. Create a helper virtual machine to assist in the migration.
2. Connect additional blank virtual disks to the helper virtual machine to hold the migrated physical disks.
3. Using imaging or backup software, migrate the physical disk data to the virtual disks.
4. Modify the boot information to ensure the correct boot order for starting the virtual machine.
5. Power off the helper virtual machine.
6. Configure a new virtual machine that uses the new virtual disks.
7. Boot the system.
8. Install and configure VMware Tools.

Considerations for Successful Migrations

What constitutes a successful migration? These are the criteria customers rank highest:

- A functional guest OS and associated application(s)
- Network connectivity
- Sign-off by key stakeholders:
  - Migration staff
  - OS administrators
  - Application owners and/or administrators
  - End users

Virtual-to-Physical Process

The V2P process enables the migration of a virtual machine to a physical server. Microsoft Sysprep 1.1 is used to handle the conversion process for Microsoft guest operating systems. Other options are possible for non-Microsoft operating systems. The Kudzu imaging tool works very well for Linux environments.

See the V2P TechNote at http://www.vmware.com/support/v2p/doc/V2P_TechNote.pdf.

The note refers to the Kudzu imaging tool but does not specify the functionality. The following is a short set of steps for migrating a Linux physical machine:

1. Create a new VM and configure appropriate CPU, memory, storage, and network.

2. Boot the physical machine with a real Linux Live-CD disk (e.g., Knoppix).
3. Boot the new VM connected to the Linux Live-CD image.
4. Use “dd” to migrate the physical machine’s disk data to the new VM.
5. Boot the new VM in rescue mode.
6. On the new VM use “kudzu” or “yast” to remove old devices and identify new ones presented to the VM.
7. Update the new VMs init using “mkinitrd.”
8. Modify “grub” on the new VM.
9. Reboot the VM and do final cleanup.

Virtual-to-Virtual Process

VMware Converter allows for migrating from one virtual technology platform to another. It provides a one-way migration to a VMware virtualization technology such as ESX, VMware Server, or VMware Workstation.

VMware Converter allows source virtual machines to run on ESX, VMware Server, VMware Workstation, Microsoft Virtual PC/Server, and Xen.

VMware Converter supports the import of OVF-based virtual machine images.

9. Optimization

The consolidation benefits of virtualization result from significantly boosting resource optimization from the low levels of a physical environment. Often represented as a percentage such as %CPU or the I/O bandwidth of a system, utilization is the amount of used resources. Utilization levels normally can be increased without affecting performance, which is measured in the number of transactions that can execute in a given amount of time, such as per-second Web page views or I/O operations. An exception, as shown in the following graph, is a negative performance impact as resource capacity reaches its maximum. The sharing of resources on an ESX host pushes the resource utilization higher when consolidation ratios increase, making it important to understand the subsequent performance impact.

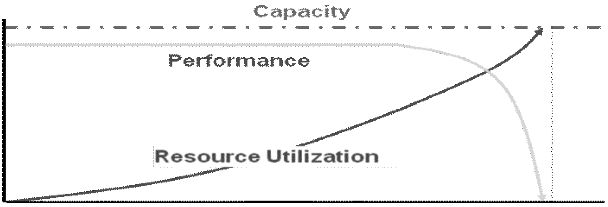


Figure 33: Performance impact as resource utilization reaches capacity

Performance monitoring enables resource demand balancing to achieve an optimal, or at least acceptable, performance level.

Applications requiring precise timing, should not be run in virtual machines. This includes performance monitoring tools, lab measurement applications, and instrument controllers. The variations introduced to the vCPU through time slicing of the real CPU can lead to millisecond delays.

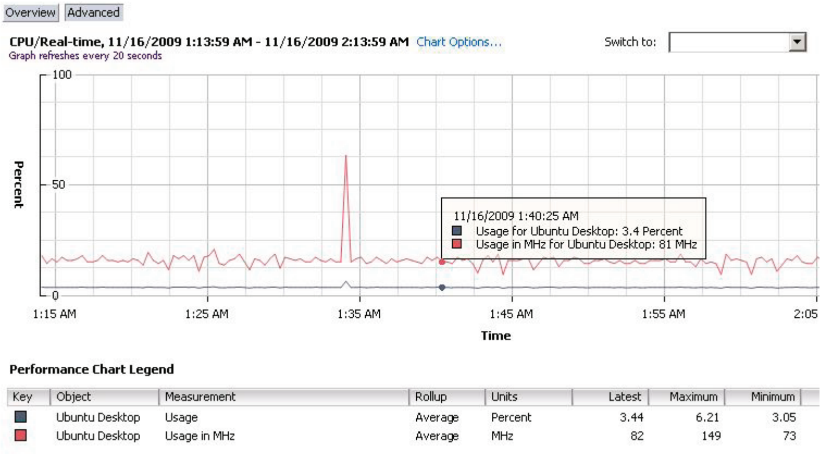
ESX Optimization

An individual virtual machine can impact other virtual machines based upon how it uses resources. Successful consolidation of multiple operating systems and applications under one set of hardware requires constant monitoring of resource loads and rebalancing of virtual machines when required, both within and across ESX hosts.

Monitoring

In a physical machine environment, resources are monitored using tools such as Perfmon and Disk Manager under MS Windows, and top and vmstat under UNIX.

In a virtual machine environment, guest OS–based performance monitoring tools may not represent reality, but can nonetheless be useful for understanding application issues. VMware vSphere tools such as VMware vCenter Server, esxtop, and resxtop (via the rCLI) represent actual performance. Figure 34 is a screenshot of a VMware vCenter client showing monitoring of CPU utilization of virtual machines.



CPU/Real-time, 11/16/2009 1:13:59 AM - 11/16/2009 2:13:59 AM Chart Options…

Graph refreshes every 20 seconds

11/16/2009 1:40:25 AM
🟦 Usage for Ubuntu Desktop: 3.4 Percent
🟥 Usage in MHz for Ubuntu Desktop: 81 MHz

Performance Chart Legend

Key	Object	Measurement	Rollup	Units	Latest	Maximum	Minimum
🟦	Ubuntu Desktop	Usage	Average	Percent	3.44	6.21	3.05
🟥	Ubuntu Desktop	Usage in MHz	Average	MHz	82	149	73

Figure 34: Monitoring virtual machine resource utilization and status

VMware vCenter Server is used to view resource utilization and performance metrics. VMware vCenter Server contains multiple default resource views that can be displayed on multiple objects such as total CPU utilization and total disk and network I/O. Custom performance charts can also be created and are invaluable when troubleshooting performance-related problems.

VMware vCenter provides detailed metrics in real-time for the past hour. Typically a much reduced set of metrics is available for data more than an hour old. For more detailed metrics, esxtop, vm-support -S, or third-party tools can be used to collect data.

Resource Pools

Resource pools are hierarchical logical containers created within VMware vSphere DRS clusters. They segment the total cluster CPU and memory resources and/or delegate administrative control over the resources and virtual machines within the pool.

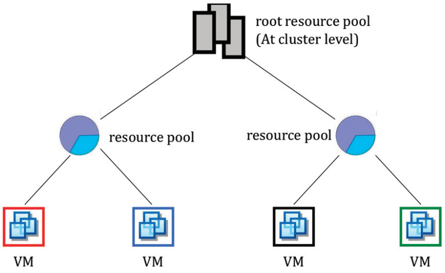


Figure 35: VMs assigned to resource pools representing shares of the physical resources

In a VMware vSphere cluster utilizing VMware HA and DRS, all CPU and memory resources are treated as a logical whole because virtual machines can be executed on any of the ESX hosts in the cluster, depending on current circumstances. Applying resource pools to the infrastructure separates the resources from a cluster of servers. Abstracting the resource pools from the physical hardware creates a form of virtualized resource management areas.

Using resource pools within a VMware vSphere cluster provides significant flexibility in managing resources and distributed administration of virtual resources. VMware vSphere administrators can delegate complete control of a resource pool to a resource manager, allowing functional or departmental groups the ability to manage both the virtual machines and the resource distribution within their managed pool—in essence, giving the pool administrators a small VMware Virtual Infrastructure of their own. By allowing both resource sharing and isolation, this feature can be especially useful when hosting groups of virtual machines for other IT departments.

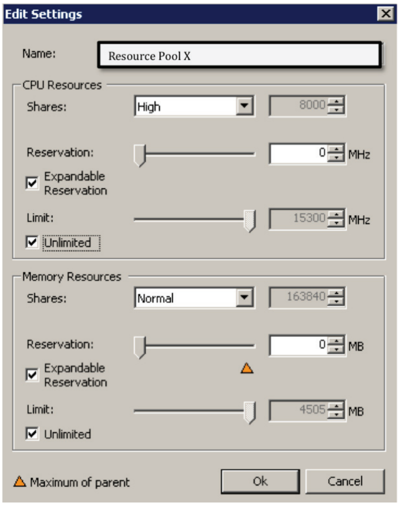
By default, each ESX host or resource cluster includes an invisible resource group containing the total CPU and memory resources of that host or of all hosts in the DRS cluster. A resource pool can contain virtual machines or child resource pools utilizing a portion of the parent pool's resources. Nesting resource pools within pools creates a parent, child, and sibling resource pool relationship that provides structure for resource distribution. Using nesting resource pools to delegate administration enables the localized management necessary for large, consolidated IT environments.

Each resource pool is assigned a *reservation,* a *limit,* and a *number of shares* for CPU and memory resources. Each resource pool must also be designated as either having an *expandable reservation* or not. Combining these settings provides isolation between resource pools while allowing sharing within resource pools.

- Reservation—A guaranteed amount of CPU or memory resources set aside for use by the resource pool. Reserved resources are not available for use by virtual machines in other resource pools. By default this is set to 0.
- Limit—The maximum amount of CPU and memory resources available to the virtual machines and child pools assigned to a resource pool. The default value for resource pool limits is unlimited.
- Shares—The number of CPU or memory shares allocated to the pool from the parent pool. Resource sharing within sibling pools is done relative to their share of the parent pool's total shares. Numerical share values may be compared only among siblings. For example, the number of shares of a VM in one Resource Pool can be compared to the shares for other VMs in the same pool but not to VMs in other pools. A pool's share value is constrained by its reservation and limit. Share values can be set to Low, Normal, High, or Custom. A custom selection allows a specific share value to be entered instead of selecting one of the default relative values.
- Expandable reservation—This is set by default for newly created resource pools. Using expandable reservations allows borrowing of resources from a parent (or ancestor) resource pool to start a new virtual machine when the pool's existing reservation is exhausted.

Shares, Reservations, and Limits can be applied at the Resource Pool level, the VM level, or both. Resources are divided at the Resource Pool level first and then subdivided among the VMs in that pool. For example, if two Resource Pools are at the same level, one with 100 VMs and the other with 2 VMs, resources are divided between the pools and then subdivided among 100 VMs for one pool and among 2 VMs for the other.

Placing application systems or functionally similar servers in resource pools simplifies resource management by avoiding the need to individually manage resource allocation to each server. Resource pools use an admission control system to ensure that the pool limit is not violated and the reservation is met for all running virtual machines and child pools contained within it. When a virtual machine is powered on or a child resource pool is created, admission control checks for available resources in the pool. Insufficient pool resources, combined with a reservation type set to Fixed prevents powering on the virtual machine. When the reservation type is set to Expandable and the current pool has insufficient resources, the parent and ancestor pools are checked for available resources. If sufficient resources are available in one of the parent pools, the resources are borrowed and reserved by the child pool and the action is allowed.


Edit Settings

Name: Resource Pool X

CPU Resources

Shares: High 8000

Reservation: 0 MHz

☑ Expandable Reservation

Limit: 15300 MHz

☑ Unlimited

Memory Resources

Shares: Normal 163840

Reservation: 0 MB

☑ Expandable Reservation

Limit: 4505 MB

☑ Unlimited

△ Maximum of parent

Figure 36: Resource pool settings

Using expandable reservations requires coordination between the pool administrators, as a child pool can borrow and reserve all of the resources in a parent or ancestor group.

Resource pools are easily managed from a series of tabbed interfaces in the VMware vSphere client, similar to the management of hosts or virtual machines.

The combination of fixed and expandable reservations with limits and shares allows resource pools to validate and ensure service levels for all virtual machines within a VMware vSphere cluster.

Distributed Resource Scheduling

DRS clusters enable load balancing of virtual machines across multiple hosts by creating logical pools of resources and distributing virtual machines across hosts as required. This load balancing can be manual, partially automated, or fully automated.

- *Manual* load balancing only provides recommendations for new virtual machine placement or migrations.
- *Partially automated* provides automatic placement for new virtual machines being deployed and makes recommendations on migrating virtual machines for load balancing.

- *Fully automated* provides automated placement of new virtual machines and automatic migration of virtual machines for load balancing. Five modes are available for this automation level: conservative (5) to aggressive (1).

Virtual Machine Optimization

VMware best practices for single and multi-threaded application CPU allocation recommend starting with the lowest number of vCPUs when either deploying new virtual machines or migrating from physical servers to virtual machines. Multi-threaded applications can benefit from multiple vCPUs, although they require all vCPUs to be scheduled to perform an execution at the same time. This can reduce the overall computing resources available. Single-threaded applications require only a single vCPU scheduled for a 4 vCPU VM. The scheduler has been vastly improved with ESX 3.x and again enhanced with ESX 4.x. However, unnecessary vCPUs still waste physical CPU resources. Detailed information can be found here: http://www.vmware.com/files/pdf/perf-vsphere-cpu_scheduler.pdf.

Application and Operating System Tuning

Although a physical server has full access to all resources, it frequently is over-provisioned in order to account for future capacity requirements. An application-OS pair needing one gigabyte of memory often requires a physical server with two to four gigabytes of RAM. The resources of an ESX host, on the other hand, are shared among all the virtual machines running on that host. Higher resource demands by one virtual machine are offset by lower demands from another. Utilizing resource pooling between ESX hosts further maximizes resource utilization.

Streamlining applications and operating systems enables further virtual infrastructure savings. Tuning an application and operating system for a virtual environment has many benefits. Examples of tuning include:

- Minimizing wasted resources (CPU, RAM, and disk)
- Generating warning and error logs when application and operating system issues arise
- Minimizing the use of snapshots
- Reducing unneeded layered applications
- Scheduling antivirus software and guest OS–based backups to minimize contention
- Using VMware Consolidated Backup or VMware Data Recovery instead of guest OS–based backups, where appropriate
- Tuning code that accesses your databases to minimize contention by doing more operations with fewer connections
- Minimizing the use of screen savers, including the pre-login screensaver
- Splitting data and log writes to different virtual disks and/or VMFS volumes to reduce I/O contention and provide better DR capabilities
- Disabling unused services
- Disabling visual effects
- Minimizes use of defragmentation tools
- Avoiding mass scheduling of defragmentation tools, antivirus, or backup solutions

VMware VMmark

VMware VMmark is a scalable benchmark tool for virtualized systems. This benchmarking tool is used to quantify the performance of virtualized environments through the use of diverse workload sets found in datacenters.

Traditional benchmarking tools focus on single workloads. In some cases, vendors have tried to measure multiple workloads on the same platform but still use the benchmarking tool in a one-to-one relationship with each workload being run. A virtual environment consists of multiple server instances all running different workloads on the same hardware platform. Traditional benchmarks do not factor in load change among active virtual machines, and they present the following challenges for a virtual environment:

- The benchmark specification must remain platform-neutral.
- The benchmark must capture the key performance characteristics found within virtualized servers and the supporting infrastructure.
- The metric must be easy to understand.
- Benchmarks must provide a way to measure scalability that applies to small as well as large servers.

The VMmark benchmarking tool measures the diverse virtual machine workloads and generates an output that scales as resource capability and resource requirements change.

VMmark calls the work unit measuring these workloads a *tile*. A tile includes the collection of virtual machines executing diverse workloads on one virtualization platform. A system is benchmarked based on the total number of tiles that a physical system and virtualization layer can accommodate.

Examples of tiles include a mail server, Web server, database server, file server, and application server. These tiles are each sub-tests that are derived from standard load generators. Determining the maximum number of tiles that fit on one virtualization platform provides a measure of its capabilities.

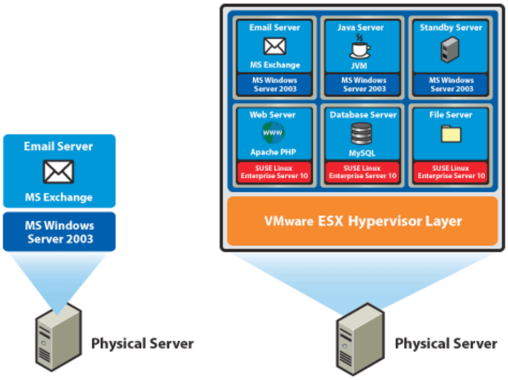


Figure 37: VMmark tiles

A typical VMmark benchmark test runs over several hours (at least three), with the actual workload metrics reported every 60 seconds. Typical trending benchmarks for scalability run for at least one month.

VMmark is typically used by hardware vendors to compare results of the same hardware platform to those of competitors. Results can be used for hardware selection based on performance requirements (http://www.vmware.com/products/vmmark/results.html).

10. Business Resiliency

Business continuity is a strategic initiative in which the business owners determine and outline requirements to ensure successful recovery in the event of a disaster. This includes determining what personnel and resources are necessary for the IT infrastructure to survive should a disaster occur. The outcome is a business continuity plan.

Disaster recovery is a tactical initiative for an IT department to meet business continuity requirements. As with business continuity, the goal is to determine who and what is necessary for the IT infrastructure to survive a disaster. The outcome is a disaster recovery plan.

Staffing and facilities do not depend on whether a solution is physical or virtual other than the need for training and the facilities components to support the disaster recovery requirements.

Depending on the level of interdependencies between systems and applications, it might make sense to keep related applications in the same physical location. Both RTOs (recovery time objectives) and RPOs (recovery point objectives) need to be considered. Consistency groups and dependency groups describe a collection of servers that perform a single business function. For example, a consistency group could be made up of Web, application, and database servers used to host an e-commerce front-end, where another consistency group might be made up of an accounting file server and the application server that relies on the data from it. It is important to plan your consolidation strategy so that business functions are recoverable across applications rather than individual servers.

Redundancy

Component redundancy for each of the infrastructure and hardware components is just as important in the virtual infrastructure as it is in a physical infrastructure. Multipathing for storage, redundant network links, redundant hardware devices, and an N+1 server architecture are all methods for providing redundancy in the environment.

Backup and Recovery Strategies

Several data-protection strategies are available, including the use of guest OS–based software, VMware Consolidated Backup, VMware Data Recovery, VMware snapshots, storage snapshots, and storage replication. VMware Site Recovery Manager is a business continuity and disaster recover process workflow product.

Guest Operating System-Based Backup

Guest OS backups operate identically to physical machine backups. A backup agent is installed within the virtual machines and is registered with a backup server. Recovery, just as on a physical machine, can be done in either of two ways:

- ❖ Bare-metal restoration, if supported by the vendor and used for the backup process, is one option. This is optimal for guest operating system–based backups, since it does not require an OS installation.
- ❖ Installation of an OS and backup agent prior to performing restoration is the other option.

VMware Consolidated Backup/vStorage API

VMware Consolidated Backup enables offloaded and impact-free backup for virtual machines on an ESX host by allowing traditional file-based backup software to leverage VMware virtual machine snapshot technology and efficient SAN-based data transfer.

Consolidated Backup is a backup enabler that allows for LAN-free backups handled by a SAN-connected proxy server. The Consolidated Backup proxy server handles backup processing, which frees resources on the virtual machines and the ESX host.

Consolidated Backup performs two types of backup: file-level backups and full virtual machine method (file system level) backups. File-level backups are similar to traditional backup methods in that full and incremental backups can be completed for Microsoft Windows–based virtual machines. Full virtual machine backups are similar to bare-metal restore backups and include all of the virtual disks, configuration files, NVRAM file, and logfiles for a virtual machine. Full virtual machine backups are supported for all guest OS types. The guest OS system state (such as registry information, system files, and boot files) is not included when using Consolidated Backup–enabled backups.

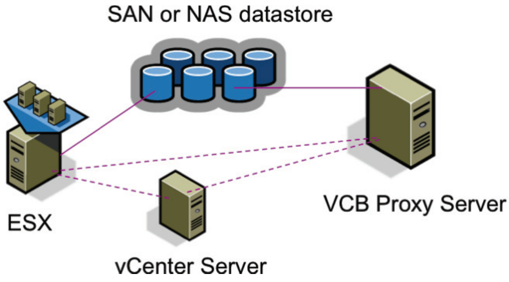


Figure 38: Consolidated Backup infrastructure

Consolidated Backup requires installation on a Windows 2003 physical server. The backup vendor client must be installed on this Consolidated Backup proxy server.

File-level Backups

Here are the high-level steps for file-level backups using third-party software with Consolidated Backup:

1. The Consolidated Backup server requests that a snapshot be added to a virtual machine's virtual disk.

2. The block list for the virtual disk snapshot is provided to the Consolidated Backup proxy server.
3. A driver on the Consolidated Backup proxy server mounts the block list.
4. The third-party backup software performs backup of the mounted virtual disk.
5. The Consolidated Backup server requests that the snapshots be removed from the base disks.

Full Virtual Machine Backups

Here are the high-level steps for completing a full virtual machine backup with third-party software for Consolidated Backup:

1. The Consolidated Backup server requests that a snapshot be added to a virtual machine's disks.
2. The disk is exported to sparse format and stored on the Consolidated Backup proxy server holding tank. The configuration file, NVRAM file, and logfiles are copied to the same location of the holding tank.
3. The third-party backup software performs backup of the data in the holding tank.
4. The Consolidated Backup server requests that the snapshot be removed from the base disk.

Several components are involved in the Consolidated Backup–enabled backup process, including the following:

- Hostd—A small application that runs on the ESX host and performs commands on virtual machines on behalf of software such as VMware vCenter Server and Consolidated Backup.
- VM to be backed up—VMware Tools is involved with the backup and is used to make the virtual machines' disks quiescent.
- Backup Proxy Server—This system contains the third-party backup software as well as the Consolidated Backup framework.
- VMware Consolidated Backup framework—This consists of the generic components:
  - vcbMounter
  - vLUN driver
  - Common API across all supported backup software
  - Integration module for specific third-party backup software

Full virtual-machine-method backups provide an option for disaster recovery. A full virtual-machine-method backup can be set to export to a storage location presented to the Consolidated Backup proxy server from a remote location. In the event of a disaster, VMware Converter can be used at the remote site to import, register, and power on the new virtual machine based on the original system. The RPO achieved is the time at which the full virtual-machine-method backup occurred. This provides a bare-metal recovery mechanism that works well for both ESX and VMware Server, enabling small to large businesses varying options based on budget and business continuity requirements.

VMware Data Recovery (vDR)

VMware Data Recovery is a virtual machine backup solution designed for smaller sites and can support a limited number of virtual machines. vDR is fully integrated with VMware vCenter and includes data de-duplication to save on disk storage for your full virtual machine backups, and it offers file-level restore or entire images as needed. vDR supports both full and incremental backups by utilizing the changed block tracking feature that the vStorage API provides.

vDR uses the vStorage API and supports the use of Volume Shadow Copy Service (VSS) used by certain Windows operating systems. For virtual machines with Windows operating systems that do not support VSS, VMware Tools uses the default LGTO SYNC driver. vDR will work with SAN or NAS devices. It also supports CIFS-based storage such as Samba.

vDR supports de-duplication stores that are up to one terabyte in size. Each backup appliance supports the use of two de-duplication stores. There is no limit defined for the number of de-duplication stores or the size of a store. However, when using more than two stores or when the size of a store exceeds one terabyte, performance will be affected.

Virtual Machine Snapshots

Virtual machine snapshots take a point-in-time copy of a virtual machine, including disk, RAM, and the virtual machine configuration. In the event of a guest OS failure, an administrator can revert to one of the previously created snapshots.

An important consideration when utilizing snapshots is that Consolidated Backup prefers that a virtual machine have no snapshots. It is also important to keep in mind that a snapshot requires disk space to store a copy of the virtual disks and other information from the snapshot. Snapshots can be used by setting variables to determine how to proceed during the VCB operations. The variables are PREEXISTING_MOUNTPOINT (removes a VM mount point used during a file-level VCB backup) and PREEXISTING_VCB_SNAPSHOT (can be used to fail the VCB job or to delete the snapshot before continuing, with "fail" as the default option).

For disaster recovery, snapshot technology can be used similarly to Consolidated Backup, in that the snapshot can be placed on alternate storage locations. The snapshot can be used to recover the virtual machine in the event of a disaster.

Snapshots can be used to assist in testing updates to the guest OS and applications running within a virtual machine. A snapshot captures the state of the virtual machine at the moment the snapshot was taken. The snapshot data files subsequently collect any changes made to the virtual machine since the initial snapshot was taken.

The snapshot includes the virtual disk, configuration information, and BIOS configuration. The size of a snapshot can grow to be as large as the virtual disk it represents.

Several other files are created to support the snapshot created. X represents the number of the snapshot taken in relation to previous snapshots:

- A file storing the state of the virtual machine when the snapshot was taken, with the filename <VM name>-Snapshot<###>.vmsn
- A file storing the state of the virtual machine memory when the snapshot was taken, with the filename <VM name>-Snapshot<###>.vmem

- A file acting as a write cache for changes to the virtual machine since the snapshot was taken, with the filename <VM name>-<nnnnnn>.vmdk

The Snapshot Manager within the VMware vCenter Server manages the snapshots associated with each virtual machine.

Snapshots allow you to test new software updates and provide a simple process to back out of the change, if required.

Storage Replication

Replication of storage lies at the heart of any virtual infrastructure disaster recovery plan. The relative ease of virtual server disaster recovery is enabled by the hardware independence of virtual servers and the mobility created by encapsulating them in a small set of files.

Planning and designing the storage and replication for a VMware vSphere infrastructure is closely related to larger disaster recovery planning efforts. Several replication strategies provide different levels of service and associated costs. For most organizations, a multi-tiered replication strategy provides the most cost-effective VMware vSphere infrastructure replication solution. Organizations typically classify approximately 80% of their servers and data with a 24-hour RTO and the remaining systems with stricter RTOs, ranging from several hours down to continuous data protection (CDP) with synchronous wide area replication.

Some companies with zero data loss (or as little data loss as possible) requirements use data replication technologies such as synchronous replication to ensure that every write is replicated before accepting the next write. Data protection technologies such as CDP can also be used when faced with more comprehensive data recovery requirements.

As with any disaster recovery plan, the RPO and RTO are constrained by budget. VMware vSphere enables extremely short RTOs, because servers can be restored and placed in service very quickly. Indeed, it is often possible to have virtual machines up and running within minutes following a disaster declaration. RPO, on the other hand, is constrained by bandwidth. Replicating a large amount of changed data over a small data pipe increases the RPO. This makes the selection of a storage replication strategy critical to a successful VMware vSphere infrastructure disaster recovery plan.

STORAGE REPLICATION STRATEGIES

Several replication strategies for VMware virtual infrastructures align nicely with backup and recovery system tiers and technologies.

The 80% of servers and data that are backed up daily with a traditional 24-hour RPO can be replicated through a variety of means from the main site to the recovery site. This type of RPO is often best met by replicating backups. After the nightly backups are completed (file- or vmdk-based), the data needs to be replicated to the recovery site before the next backup window begins. When replicating a large volume of data, it can become a challenge to maintain even a 24-hour RPO at a level of bandwidth that does not exceed the budget.

The 20% of remaining systems can have much stricter RPOs, some even requiring continuous data protection and transactional replication. In these cases, replicating nightly backups is insufficient. Either asynchronous or synchronous replication between storage mediums is required.

Several strategies exist for enabling storage replication, depending upon RPO and budget. The primary differentiation between these replication strategies is where they sit in the storage stack.

Embedding storage intelligence directly into the storage fabric creates the highest performance and transparency. Virtual SANs, LUN virtualization, remapping, and replication can all be enabled transparently to the VMware Infrastructure. Synchronous remote replication of some or all of the storage traffic is accomplished by using a variety of storage replication protocols.

When the higher-performing fabric-based replication solution is not practical or economic, the recommended solution can be a man-in-the-middle replication server or appliance to provide asynchronous in-band or side-band replication of LUNs, and virtual machine exports. These solutions include software- or appliance-based storage virtualization or replication systems. They can also be used for asynchronous mirroring of operating system or vmdk-based backups on disk, virtual tape library, RDM LUNs, or NAS storage. This replication strategy is still transparent to the VMware Infrastructure but not to the storage infrastructure.

The man-in-the-middle replication scheme is most effective when combined with data de-duplication solutions. Data de-duplication can dramatically reduce the amount of replicated data, particularly when replicating backups that involve a great deal of duplicate data. Adding data bandwidth compression further enables rapid replication with less bandwidth required.

The last replication method involves inserting an operating system or application shim to enable continuous data protection for individual applications. This is common in highly transactional, business-critical applications, including financial, database, and messaging servers. This method is very effective for application protection but needs to be integrated with the application and operating system, creating a more complex solution.

Uncompressed, only about 14GB of data can be transmitted over a T1 line in a 24-hour period. Determining the bandwidth required must be included in the RPO. The total data, rate of data changes, and replication technology utilized must all be considered when calculating replication time. RPOs can also vary widely depending upon the type of server or application. When the necessary bandwidth is calculated, a cost-benefit analysis can be performed to determine whether bandwidth-saving techniques such as data compression or data de-duplication are viable options for optimizing replication speed.

Fault Tolerance Technologies

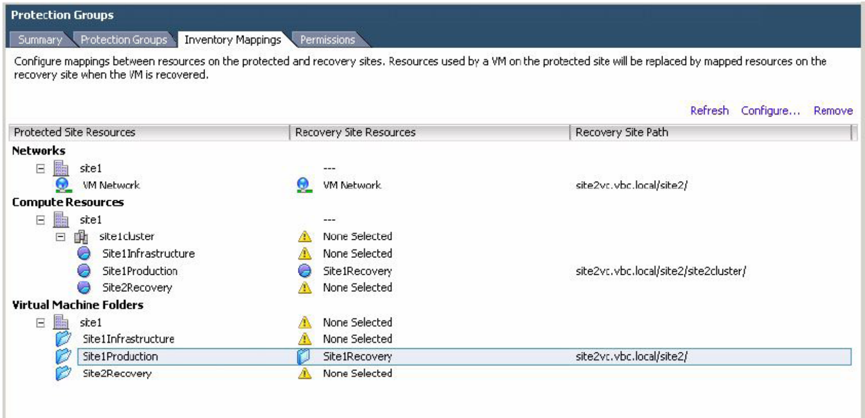
vCenter Site Recovery Manager (SRM)

VMware Site Recovery Manager (SRM) provides end-to-end disaster recovery workflow management and automation for a VMware Infrastructure. The workflow steps can be automated to ensure that the recovery processes are consistently followed both for testing and for real disaster recovery situations. SRM is built upon a storage replication foundation.

Recovery plans are created within the tool to define the order of virtual machine failover and startup. Automated, non-disruptive, disaster recovery testing allows validation testing of the workflow in a fenced environment. This means that testing of the disaster recovery workflow can be carried out without affecting the running production systems that the workflow protects. This also minimizes testing costs and errors in manual steps (such as breaking replication, rescanning LUNs, and re-registering the virtual machines on the failover VMware Infrastructure environment).

The workflow acts as a working document of the recovery plans and provides the instructions for recovery. Frequent testing followed by any necessary adjustments to the workflow results in higher success rates in the event of a true disaster recovery situation.

SRM provides sites that fall under regulatory compliance guidelines with a very cost-effective method to define a workflow and complete testing. The testing process outputs a report that can be used to support regulatory compliance auditor requirements.



Configure mappings between resources on the protected and recovery sites. Resources used by a VM on the protected site will be replaced by mapped resources on the recovery site when the VM is recovered.

Protected Site Resources	Recovery Site Resources	Recovery Site Path
Networks		
site1	---	
VM Network	VM Network	site2vc.vbc.local/site2/
Compute Resources		
site1	---	
site1cluster	⚠ None Selected	
Site1Infrastructure	⚠ None Selected	
Site1Production	Site1Recovery	site2vc.vbc.local/site2/site2cluster/
Site2Recovery	⚠ None Selected	
Virtual Machine Folders		
site1	⚠ None Selected	
Site1Infrastructure	⚠ None Selected	
Site1Production	Site1Recovery	site2vc.vbc.local/site2/
Site2Recovery	⚠ None Selected	

Figure 39: VMware SRM inventory mappings

vCenter Heartbeat

VMware vCenter Server Heartbeat is designed to protect against failure of the OS or hardware that the vCenter Server operates on. This protection occurs for both physical and virtual vCenter servers through the following protection levels: server protection, network protection, application protection, performance protection, and data protection. This operates in a passive clustering model. Should any of the protection levels fail on the primary server, a secondary server takes over.

VMware vCenter Heartbeat provides high availability and disaster recovery for VMware vCenter Server and related components such as licensing and plug-ins. vCenter Server Heartbeat enables monitoring, replication, and rollback for instances of vCenter Server enabling effective disaster recovery in the event of loss of the primary datacenter. The secondary server uses the same name, same network address configuration, and the same file and data structure and can run the same applications/services as the primary server.

The secondary server is activated if a dedicated network connection between the two loses a heartbeat. The heartbeat is maintained using keep-alive messages. When the secondary server detects the loss of a heartbeat, then a packet filtering mechanism changes state to allow the secondary server to communicate with the primary node network configuration.

VMware Fault Tolerance

VMware Fault Tolerance eliminates downtime due to hardware failures, by creating a live shadow virtual machine to run on a second ESX host, receiving and executing each memory operation and each CPU instruction in virtual lockstep, using vLockstep technology, with the primary VM. If the primary host and protected VM fail, processing is seamlessly cut over to the secondary VM running on an alternative ESX host without requiring a reboot. A new shadow VM is then automatically configured to run on another ESX host within the cluster, which then runs in virtual lockstep with the new primary VM. Fault Tolerance is a simple, selectable feature that can be implemented for specific VMs deemed to be critical to the operation. Both primary and secondary VMs utilize shared storage and can be VMotioned among ESX hosts. Fault Tolerance eliminates the need to purchase and manage traditional clustering offerings dependent upon stand-by hardware.

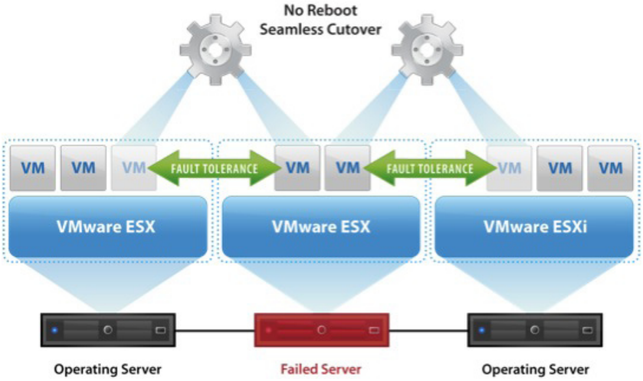


Figure 40: VMware FT provides zero downtime of applications and OS

VMware High Availability

VMware High Availability provides quick and automated failover of applications in the event of ESX host failure. When a host fails, HA quickly detects the failure, sorts all the VMs on the host based on their restart priority, and then places them on alternative ESX hosts within the cluster. The virtual machines reboot very quickly compared to physical servers, although the in-memory state of the virtual machines at the time of host failure is lost.

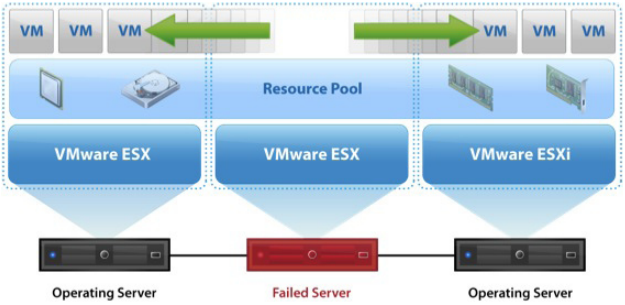


Figure 41: VMware HA protects and restarts VMs from host failure or VM failure

Networking Strategies for Disaster Recovery

While storage replication is the foundation of VMware Infrastructure disaster recovery, enabling effective replicated networking is a requirement that is often overlooked. In the event of a disaster declaration, users may not be able to get access to their personal computer and may only be able to work from another office or remote location. Both application and network access strategies are essential to enabling quick and effective recovery.

Thin Remote Application Access

A crucial piece of the business continuity plan is to re-establish user access along with server recovery. Architecting a datacenter failover that is transparent to application users requires that the primary site network scheme be moved to the failover site along with the applications and processing. Providing thin network-based application delivery using common methods such as application virtualization or streaming Web-based applications, or connectivity to remote virtual or thin desktops through a centralized application gateway hub or portal, enables continued access to recovered applications without reconfiguring client access.

Network Infrastructure Extension

When some form of remote application access is available, networking name and address spaces must be transferred from the primary site to the disaster recovery site to allow the systems to resume activity without application or network reconfiguration. When possible, the primary network is virtually extended not only into the VMware Infrastructure, but also out to the remote recovery site using CWDM or MPLS WAN technologies, creating an extended virtual grid or clustered datacenters. Network and port address translation schemes are used to protect the internal address space.

Datacenter Access Cutover

When the network infrastructure and applications are replicated and recovered, a method must be available to cut over application access from the primary datacenter to one of the disaster recovery datacenters within the RTO without reconfiguring clients. The most common way of achieving this is through the use of dynamic DNS to change address resolution for the application delivery points, redirecting user access to the replicated and restored disaster recovery site. Another way is through the use of network load balancers such as those from Cisco or F5 Networks.

The use of consolidated application delivery points simplifies the redirection of users, requiring only redirection of DNS resolution to restore application access. This consolidation and redirection can occur easily with application portals or hubs for Web access, secure SSL VPN gateway products for streamed or virtualized applications, or virtual desktop connection brokers for organizations hosting virtual desktops in the datacenter.

Security Considerations for Disaster Recovery

Many virtualization deployments must meet specific regulatory compliance guidelines for security. Consider any firewalls, packet filters, or intrusion detection systems that can affect the virtual machine after it is consolidated. For example, a physical machine that is not protected by a firewall might end up on an ESX host that is behind a firewall. Also, IDS systems might not be prepared for the types of traffic that are seen coming from new virtual machines on a network segment.

Supportability is often a customer concern. Some software vendors say that they will not support their applications within a virtual machine. If this is the case, the customer should contact the ISV directly and state the need for support. If future license revenues are at stake, the ISV might be more inclined to provide some sort of support statement. In the past, VMware has participated in these sorts of conversations. See VMware's partner catalog Web page for details on existing agreements with various ISVs.

Using VMware vCenter Server roles and permissions is a best practice to support regulatory compliance and change management requirements. Changes made by individuals are logged for use by an audit team.

Creation of security pods is one method used by customers to satisfy audit requirements. A specific grouping of virtual machines using the same VMFS volumes and virtual switches is an example of a security pod. Add specific roles and permissions to complete the setup. Remember that the roles and permissions are tied to the accounts on the VMware vCenter Server.

11. Security

Security Considerations

As with physical infrastructures, virtual infrastructure architecture needs to reflect security considerations at every juncture to ensure an effective, useful, and secure environment. In many cases, specific regulatory guidelines must be met. VMware vSphere incorporates an ability to build a virtual infrastructure that not only addresses virtualization-specific security concerns but is more secure than the physical alternative.

As with physical infrastructure, security design must incorporate firewalls, packet filters, and intrusion detection systems. The flexibility of virtualization, though, requires different considerations. For example, a virtual machine protected by a firewall might VMotion to an unprotected ESX host. Or an Intrusion Detection System (IDS) may be unprepared for traffic coming from new virtual machines on a network segment.

Some software manufacturers refuse to support their applications within a virtual machine, sometimes claiming a lack of security as justification. A properly architected virtual infrastructure leveraging the capabilities of vSphere enables an environment at least as secure as the physical equivalent. A customer can often contact the ISV and obtain support—VMware has participated in these types of discussions. See VMware's partner catalog Web page for details on existing agreements with various ISVs. If no policy is in place, consider working with VMware to get the ISV to establish a reasonable support policy.

Utilizing VMware vCenter Server roles and permissions is a best practice to support regulatory compliance and change management requirements. Changes made by individuals are logged for use by an audit team.

Creation of security pods is a method to satisfy audit requirements. A specific grouping of virtual machines using the same VMFS volumes and virtual switches is an example of a security pod. Add specific roles and permissions to complete the setup. Roles and permissions are tied to the accounts on the VMware vCenter Server. Other considerations are the type of hypervisor used in the virtualized infrastructure. Both ESX and ESXi are bare-metal hypervisors and are fully supported, with each having its own benefits. VMware is moving towards an ESXi-only solution, where security risks are reduced due to a smaller attack surface. Where ESX contains a modified version of Red Hat Linux, in VMware terms called the Service Console, ESXi does not. The functionality of the Service Console, however, has been moved to an appliance which is optional. Currently many of the patches released are Service Console related. Not only will using ESXi decrease security risks, it will also simplify patch management and reduce all related operational costs.

Enhancements

VMware vSphere 4.0 introduces many new security-related enhancements. These enhancements include memory hardening, kernel module integrity, and extended permissions for vCenter. As introduced in Chapter 2, VMware vSphere also includes enhanced security capabilities with VMsafe, vShield Zones, and the VMware vNetwork Distributed Switch.

Memory hardening can best be described as a method for storing drivers, libraries, and applications in random memory locations. This location is non-predictable, which makes the hypervisor less vulnerable to memory exploits and malicious code.

Authenticity of modules, drivers, and applications is assured by a digital signature. This signature will be verified when the module, driver, or application is loaded by the VMkernel.

One of the areas where vCenter permissions have been vastly improved is datastore privileges. Datastore privileges enable you to control who has access to a datastore and what kind of operations they are allowed. A good example of a new option is “allocate space.” With “allocate space” one can control who can create new virtual machines, snapshots, or clone or create a new virtual disk on a specific datastore.

Privileges can also be controlled for vNetwork Distributed Switches and vNetwork Distributed Port Groups. With Distributed vSwitches, setting the correct permissions has become much more important because of the risk and impact changing a Distributed vSwitch or Port Group has.

VMsafe

VMware VMsafe is a new security technology for VMware environments. VMware VMsafe enables, through published APIs, third-party vendors to provide virtual appliances that automatically protect everything running within the vSphere virtual machines, thus no longer requiring the deployment of individual agents. Virtual security can be assigned specifically to individual VMs. This makes it possible to monitor within a virtual machine and stop the execution of, for example, a virus. Other threats/risks that can easily be identified are rootkits and malware, and VMsafe even monitors for disallowed process execution.

vShield Zones

VMware vShield Zones is an application-aware firewall technology that is built for virtualization and uses VMsafe APIs. vShield Zones builds a firewall “fence” around a VM cluster by creating security zones internal to the virtual machine that follow it as it migrates between hosts. This ability enhances security by providing firewall appliance functionality without the complexity of network management. The following figure shows an example of securing specific zones in your virtual network with vShield Zones.

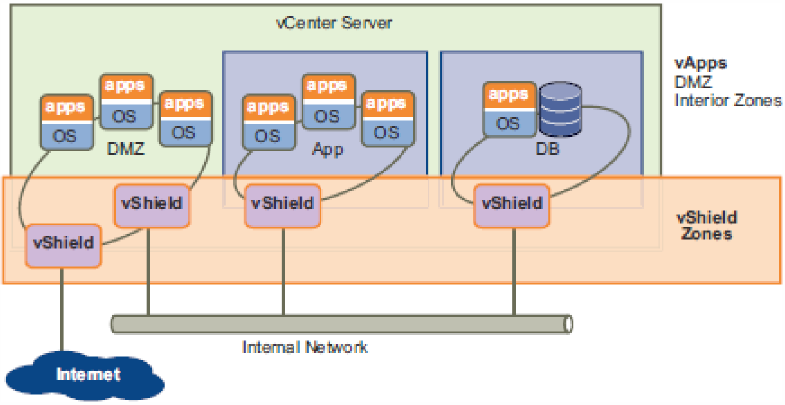


Figure 42: vShield Zones

vNetwork Distributed Switch (vDS)

The vDS enables monitoring of traffic over a virtual switch, thereby resolving one of the largest security concerns regarding virtualization. It enables application of policies such as availability, encryption, and maximum latency to a virtual machine that then follow the virtual machine as it migrates between ESX hosts. A VMware vDS also provides an open API, thus allowing third-party manufacturers to provide their own switch plug-ins. The Cisco Nexus 1000V virtual switch, for example, provides a full suite of Cisco security features, including access control lists, private VLANs, IP Sourceguard, Netflow, etc.

12. VMware View

The benefits of a virtual infrastructure are well documented. Consolidation and standardization of resources enables greater physical and human resource efficiency. Virtual machine encapsulation provides mobility, enabling replication and recovery of logical servers in ways unheard of in a physical datacenter. High availability, distributed resource scheduling, and consolidated backup services add additional value to virtualized servers.

Organizations implementing comprehensive VMware-based virtualized datacenters often find a divide between the manageability and simplicity in their virtual server infrastructures and the cumbersome distributed management of traditional physical desktop computers. This separation between virtualized servers and physical desktop computers becomes problematic when planning for disaster recovery—a replicated, recovered virtual server infrastructure without access to the desktop computers cannot provide complete business continuity. Early adopters of VMware looked for a way to extend the scope and benefits of their VMware Virtual Infrastructure to include the client-side computing resources to provide a consistent and complete virtualized infrastructure, and VMware responded by launching a new industry—Virtual Desktop Infrastructure (VDI).

VMware View delivers desktops as a managed service as part of the virtualized infrastructure. A VMware View VDI solution is built by accessing ESX-hosted virtual desktop images using a high-performance PCoIP remote display protocol through the View Manager virtual desktop connection broker. Individual desktops can be assigned to each user or allocated from a shared pool of desktops: persistent pools for logging users into the same desktop each time and non-persistent pools for generic desktop access. VMware View can suspend virtual desktops not in use to free up compute resources and to save power. Additional components manage printing, client hardware, application distribution, and patching services.

Combining View and vSphere enables the server, desktop, disaster recovery, network, and storage computing infrastructures to be deployed, managed, replicated, and recovered as a unified whole throughout the enterprise. The centralization of management, administration, and resources significantly lowers costs, as shown in the ROI section of this book. Security is enhanced as the desktops move to the datacenter where they are backed up, managed, and replicated off-site for disaster recovery purposes. The desktop computing experience improves through ubiquitous access from any device anywhere a user can get a browser. And business agility and productivity increase through rapid provisioning and a reduction in hardware-related user downtime.

Among the PCoIP enhancements included are multimedia redirection, dynamic host- or client-side media rendering, dynamic Adobe Flash control, USB redirection for seamless peripheral support, dynamic bandwidth adjustments and network optimizations, dynamic audio adjustment and audio redirection support for audio peripherals, and multi-monitor support for up to four monitors, along with 1920x1200 resolution per display.

VMware View Product Suite

Whereas VMware formerly referred to its desktop virtualization products as VDI, they are now part of the VMware View portfolio. The standard edition of View enables VDI with View Manager, while the Premier Edition enables storage optimization with View Composer, application virtualization with ThinApp, and client virtualization with Offline Desktop. Figure 43 represents a comprehensive VMware Infrastructure environment, including the VMware View product suite components.

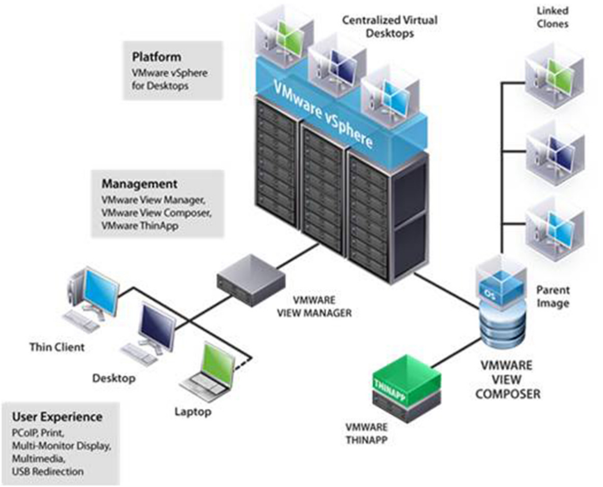

Centralized Virtual Desktops
Linked Clones
Platform
VMware vSphere for Desktops
VMware vSphere
Management
VMware View Manager, VMware View Composer, VMware ThinApp
OS
Parent Image
VMWARE VIEW MANAGER
Thin Client
Desktop
Laptop
THINAPP
VMWARE VIEW COMPOSER
User Experience
PCoIP, Print, Multi-Monitor Display, Multimedia, USB Redirection
VMWARE THINAPP

Figure 43: VMware View VDI infrastructure

View Manager

View Manager, formerly called Virtual Desktop Manager, provides a single console to provision, deploy, and manage virtual desktops and includes View Connection Server, View Agent, View Client, View Portal, View Administrator, View Printing, and Unified Access.

- View Connection Server works with vCenter Server to provide advanced management capabilities. It installs on a Microsoft Windows server that is part of an Active Directory domain.
- View Agent runs on each virtual desktop and enables session management and single sign-on. View Agent can be installed on a virtual machine template, enabling automatic inclusion for desktops created from the template.
- View Client runs on a Windows PC or a thin client and connects users to their virtual desktops through the View Connection Server.

- View Portal provides alternative Web browser access to virtual desktops.
- View Administrator provides a Web-based console for administrators to manage virtual desktops, including configuration and policy settings.
- View Printing is installed on the virtual machines and configures them for printing. It receives, decompresses, and decrypts print data, then converts the common data format into printer-specific formats for sending to the print device. It automatically makes all necessary printers available while eliminating the need to install and maintain printer drivers on virtual desktops. The advanced print stream compression can reduce network utilization by 98% and enables high quality printing even over WAN connections.
- Unified Access leverages View Manager's secure connection brokering for other platforms accessible by RDP, including terminal servers, blade PCs, and physical PCs. It enables load balancing of multiple terminal servers along with monitoring and auditing from within View Manager.

View Composer

VMware View Composer provides image management for environments with large segments of homogeneous desktops. It is a complementary solution to traditional management technologies, and while it can dramatically reduce storage, it also provides a new method of management and control.

View Composer uses VMware Linked Clone technology to rapidly provision desktops. A linked clone shares virtual disks with a master virtual machine image, enabling administrative tasks such as patching to take place at the master image level. It also can reduce storage requirements. The clones can be powered on, suspended, snapshot, and reconfigured independently of the parent. User profiles are separated from the desktop image and stored as user personality disks, enabling independent administration.

View Composer image management provides the ability to add more storage as required and to retire an existing storage array with three primary techniques: refresh, recompose, and re-balance.

- Refresh enables resetting a linked clone desktop back to the state of initial roll out.
- Recompose alters a deployed desktop state such as assigning it a snapshot of a different master virtual machine.
- Re-balance balances virtual machine disks across available datastores to enable efficient usage of available storage.

VMware ThinApp

VMware ThinApp is an agentless application virtualization solution that decouples the application from the OS. It encapsulates each application and all components required for it to run in an isolated container. Applications can be packaged once and deployed to multiple devices across both virtual and physical systems, minimizing the desktop images to be managed. User-specific configurations and changes are stored separately.

VMware ThinApp enables VDI administrators to reuse virtual machine templates and reduce storage requirements by storing and streaming application and user-specific configuration data from a central server. A pool of users can utilize the same central application. Multiple versions of the same application can run side by side without conflict, because resources are unique to each application. ThinApp applications can also be linked together to share common independent components (e.g., Java, .Net), minimizing package size. Application patches and updates are streamlined and can take place while applications are in use.

Offline Desktop—Experimental Use

Remote connectivity is increasingly becoming ubiquitous as even airlines are beginning to outfit their planes with WiFi. For users requiring an ability to work while disconnected, Offline Desktop allows the movement of complete virtual desktops between the datacenter and the laptop or other physical devices, with administrator-defined security levels and encryption policies intact. Offline Desktop enables designated end-users to relocate their View virtual machines to a local physical computer and back. The processes are termed *check out* and *check in*, respectively.

- A check-out operation is initiated by the client and ends with a local copy of the online desktop on the user's local computer. A user can check out after connecting and authenticating to View Manager at least once, when a policy allows offline access and when the desktop is not currently checked out.
- A check-in operation is initiated by the client to push changes made to a previously checked-out desktop up to the datacenter. A user can check in when connected and authenticated to View Manager, assuming an open offline session exists on the client device and the desktop retains policy permission for offline usage.

When checked out, the virtual machine has a "heartbeat" back to the datacenter allowing administrators to deactivate it if necessary. When the user reconnects, only the delta is checked in. The Offline Desktop must be managed by vCenter Server and must either be an individual desktop or in a persistent desktop pool. ThinApp applications stored on a shared network drive will not be checked out.

The administrator can optionally encrypt communication and data transfer between the offline desktop and View Connection Server, although the data on an offline desktop will always be encrypted.

Policies are used to assert control over product components at the global, desktop pool, and desktop user level. By default, each user-level policy inherits its setting from a pool-level policy that, in turn, inherits its setting from a global policy. Policies include the ability to limit the amount of time an offline desktop can run without successfully contacting the View Manager server for policy updates.

VMware View Infrastructure—Design Considerations

The virtual desktop workload is very different from a virtual server workload; a View implementation can create hundreds to thousands of virtual machines, each with a smaller resource footprint than a virtual server. Proper sizing is critical to ensure both adequate capacity and performance, as the end-user experience will make or break a VDI project.

The first step is to assess the physical desktop usage. VMware Capacity Planner along with tools such as Perfmon or LiquidWare Labs Stratosphere can be used to measure desktop performance, including CPU usage, memory consumption, storage throughput, storage IOPS, and network throughput. The results will assist in designing an optimal VMware View infrastructure. Of course, a trade-off exists between performance and resource dedication. VDI technology is advancing at a rapid rate and sizing considerations are similarly rapidly changing.

Architecture Planning

In planning the View architecture, we recommend taking a conservative approach that errs on the side of excess performance and capacity. It is easier to increase consolidation densities later than to salvage a project that has garnered a reputation for poor performance. The environment should be designed to accommodate usage spikes. It is important for the storage and network groups to participate along with the server group and desktop team in the planning.

An important architectural consideration is whether or not to run View on the same ESX instances as the server virtual machines. VMware recommendations are generally to use a set of ESX servers in a cluster dedicated to desktops, although utilizing vCenter Server to manage both environments. A large environment will likely have multiple clusters for both View and servers.

Some other architectural considerations include:

- ❖ Protocol: FC, iSCSI, NFS. If using VMware View Composer, NFS may be a more desirable protocol, as it is not restricted by the VMFS limitation of an eight-host maximum.
- ❖ Storage: Type of shared storage and drives and whether or not to utilize a tiered solution (e.g., FC for system drives, NAS for data drives). Special considerations also need to be taken into account when planning storage for virtual desktop infrastructures. Virtual desktop images are generally much smaller than server images, creating a larger vmdk-to-physical disk or LUN ratio. VDI storage architecture and management practices need to deal with concurrent access to thousands of vmdk files. Resource requirements must be factored into activities such as desktop search tools. Storage should be sized for peak usage, not average. Enough IOPS and bandwidth must be available or performance will suffer. The storage footprint, however, can be minimized with View Composer, virtualizing the applications with ThinApp, utilizing thin provision storage, using nLite and vLite to streamline the guest OS, and with array-based snapshots. Activities such as anti-virus scans, OS patching, and file updates should be staggered and deployed at the least busy times.
- ❖ Servers: Considerations include whether to use traditional rack-mount servers, blade servers, or new, specifically designed virtual infrastructure hosting platforms such as the Cisco Unified Computing System. When sizing, it is important to plan for virtual machine memory overhead—84MB for each 32-bit VM with 1GB RAM. Page sharing should be maximized. Over-commit should be effectively leveraged and ESX host-level memory swapping avoided.
- ❖ Network bandwidth: Just as in storage systems, activities such as anti-virus scans, OS patching, and file updates should be staggered and deployed at the least busy times.

- WAN bandwidth: When accessing virtual desktops over a WAN, extraneous services and graphical add-ons such as themes and screen savers should be disabled. Bandwidth requirements depend upon the display protocol being utilized (PCoIP or RDP), the usage cases, and the amount and type of printing. A good rule of thumb is to allow at least 128K per user while keeping latency at less than 150 ms.
- View Composer: The heterogeneity of the environment is important in determining whether or not to use Linked Clones. View Composer supports only Windows desktop images. Similar or superior storage consolidation can be achieved with SAN array products such as NetApp FlexClones or EMC Snap. Today the clones must be manually registered in View Manager, and automatic pool sizing cannot be utilized; however, storage array manufacturers are working to integrate their products into View Composer in order to provide enhanced functionality and scalability.

View Composer/Linked Clones Deployment Considerations

- Keep the number of VMs per LUN under 64 for best performance.
- Each desktop can only have one user data disk.
- Store the OS disk and the user data disk on different data stores.
- Keep the system disk from growing too big. Updates directly to a linked clone system disk can potentially cause inefficiency by increasing their size.
- A SAN is required to support View Composer in a cluster; it will not work on a local disk. View Composer supports a maximum of eight hosts in a cluster unless running NFS as noted above.
- Be conservative with storage over commitment.
- Available storage needs to be managed regularly by the administrator, with used space kept under 95%, as exceeding this level can result in loss of performance.

Capacity Planning

Capacity planning is difficult but critical. It is facilitated by evaluating the results from the physical desktop environment assessment. Typical desktop workloads for CPU, memory, network, and disk requirements can be modeled for different categories of users, but should be validated against resource oversubscription. Alternatively, the View environment might be sized according to worst-case performance requirements.

- Estimate the CPU requirements. VMware studies show approximately eight typical knowledge workers per virtual desktop per core, although a 2009 in-depth study of evaluating the scalability of VMware View on a Cisco Unified Computing System (UCS) showed an average of 20 users per core (two 4-core Intel 5500 series 2.93 GHz CPUs per UCS blade and 96 GB RAM). See http://www.emc.com/collateral/software/white-papers/view-ucs-vmax-whitepaper.pdf.
- Estimate the memory requirements. The ability of vSphere to utilize page sharing significantly reduces the amount of memory normally required when aggregating virtual desktops. VMware recommends provisioning a minimum of 512

MB RAM for a Windows XP virtual machine without memory sharing, or half that with memory sharing.

- Estimate the network requirements. In addition to examining the physical network usage patterns, virtualization considerations include the remote display protocol being used, the printing requirements, shared/redirected folders, and multimedia requirements.
- Estimate storage capacity requirements. A formula guideline is (size of vmdk) + (VM RAM) + (VM suspend/resume) + 100MB per VM for logs.
- Estimate storage performance requirements. Storage performance is the most difficult and important element to design correctly. In addition to examining the physical usage patterns, virtualization considerations include whether other systems or virtual machines are sharing the same spindles and the VMware ESX disk I/O. Also important are usage patterns such as boot periods, user initiated desktop search, defrag, virus scan times, etc.

Guest OS Sizing Considerations

Unnecessary services, device drivers, and add-ons, including graphical screen savers, should be disabled. Desktop images should be eliminated by using centralized file storage, roaming/virtualized profiles, de-coupled applications, and redirected application data, cookies, favorites, and templates.

VMware View Infrastructure—Deployment Considerations

Careful piloting is essential for a successful enterprise View deployment. It is important to simulate the anticipated production VDI workloads as accurately as possible and then to evaluate the performance in actual, though limited, usage scenarios. This will either validate the capacity planning or allow a reevaluation of resource requirements.

Another important area when considering VDI is management of user expectations and change management. Users who are assigned virtual desktops should know about policy changes such as capped storage, no administrator privileges, and so on. The VDI pilot will assist in enabling a smooth transition to View from the physical desktop.

A pilot program of 25 to 100 virtual desktop users is typically adequate to test-drive the policy decisions and better understand user requirements. During the pilot, group policies can be adjusted and finalized. User satisfaction surveys can identify productivity issues. The virtual desktop templates (virtual machine image) can be finalized for use in the full-scale implementation.

It is important to understand that moving an organization from physical desktops to virtual desktops is as much a culture change as a technology change. If managed properly, a View VDI implementation can save your organization considerable capital, operational, and management expense, as well as enabling increased user productivity. A View VDI deployment puts your desktop assets in a position to benefit from all the current and future innovations that virtualization has to offer.

Appendix A. Virtualization Technologies

Virtualization technology[1] has a long history. We'll describe some different virtualization technologies, compare these approaches, and provide a context for VMware virtualization. We'll describe only system-level virtualization, as opposed to process-level virtual machines such as Java virtual machines.

Virtualization was first developed in the 1960s and later popularized on IBM mainframes in the 1970s. The most famous version implementation was the VM/370. These systems partitioned hardware on a single computer into virtual machines. The goal was to enable mainframe computers to perform different tasks at the same time. Mainframe computers were expensive; virtual machines maximized their resource utilization by running concurrent tasks for different users.

UNIX vendors adopted hardware-level virtualization and partitioning technologies, and then later adopted software virtualization.

Operating System Virtualization

Operating system virtualization splits the operating system of a single computer into multiple partitions, which can run multiple instances of the same operating system. Examples of this include chroot jails in UNIX.

Logical partitioning, also known as LPAR, is found in mainframe computers such as IBM System z (and less commonly on other IBM systems), as well as on computer systems from other vendors. In logical partitioning the resources of a physical computer are partitioned so the computer's memory may be split, allocating a specific range to each partition. Hardware assistance is often used to partition a system but is not necessary for operating system virtualization in general.

Hardware Virtualization

In this model, which VMware technology uses, one or more abstractions of a computer are created. This enables more flexibility, since it allows one computer to run several different operating systems.

Hardware virtualization can be performed using two different methods: hosted or hypervisor. Examples of VMware products in each category are given.

[1] Terminology for virtualization technologies varies by vendor. Where different terminologies can be used, we have defaulted to using the VMware terminology in this document. Concepts included can be applied to other virtualization implementations.

Hosted

Hosted virtualization (VMware Workstation, VMware Player, VMware ACE, VMware-Server, and VMware Fusion) relies on having a standard operating system between the physical computer and the virtualization layer. This requires installation of an operating system such as Microsoft Windows, and then installation virtualization software such as VMware Workstation on top of it. Finally, a guest operating system such as Windows or Linux is installed in one or more virtual machines running within VMware Workstation.

The hosted virtualization platform depends upon the host operating system for resources and is also impacted by any issues that might affect the host operating system. If the host operating system gets busy or if it crashes, the virtual machines are affected. A benefit is that hosted virtualization systems can run on computers that support common OSes such as Microsoft Windows, increasing compatibility.

Hypervisor or Bare Metal

Hypervisor virtualization (VMware ESX, VMware ESXi, part of VMware vSphere) platforms have a partitioning layer that runs directly on top of the hardware and below higher-level virtualization services that provide a virtual machine abstraction. The hypervisor is installed on the computer, just as though it is an operating system. It provides the capability to create virtual machine partitions, with a virtual machine monitor running within each partition.

Hypervisor virtualization platforms eliminate the overhead of typical operating systems and have direct access to and control of the actual hardware resources. The performance of virtual machines operating on top of bare metal virtualization is closer to native performance than a hosted approach.

Virtual Machine Monitor

The VMM is a layer of software that runs between the hypervisor or host operating system and a virtual machine. It manages the resources and their allocation to the virtual machines running on the system.

The VMM decouples the software from the hardware underneath. As a famous quote from computer scientist David Wheeler states, “All problems in computer science can be solved by another level of indirection.” VMM’s decoupling capability provides substantial control over how the guest operating system accesses the hardware.

VMM may be implemented in many ways. Some computer architectures in the past were designed to be virtualized, but many CPUs, including the x86 family (except for recent editions), are not, thus requiring techniques such as binary translation to work around this limitation.

VMMs have the primary task of executing instructions on the virtual CPU and emulating virtual devices.

CPU Virtualization

The VMM can follow one of several techniques for CPU virtualization. These examples are specific to x86 virtualization.

Full Virtualization

With full virtualization (also known as *native virtualization*), the guest OS is presented with a virtual hardware abstraction that represents a physical machine. This does not mean that the virtual machine is identical to the underlying hardware. A virtual machine can be thought of as a VMware-brand PC that is standardized to the VMware Infrastructure architecture but is different from the underlying hardware. The virtual machine is recognized and accessible to the operating system or applications software just as if it were a physical machine, so no modification to the software is necessary. VMware has traditionally used a binary translator to overcome some limitations in the x86 architecture that made virtualization difficult. Note that hardware-assisted virtualization is a variant of this, where x86 hardware designed for virtualization assists in this task. It is important to remember that in the case of full virtualization, a standard operating system such as Windows or Linux, without modifications, will run in a virtual machine.

Para-Virtualization

With para-virtualization (also known as *OS-assisted virtualization*), the guest OS is presented with a modified hardware abstraction. This requires operating systems to be modified and ported to this particular virtualization platform. This reduces operating system compatibility, but that is a trade-off against potential increases in performance of certain CPU-bound applications that run on systems without virtualization hardware. This performance increase is achieved by *hypercalls*, a communication method that occurs between the guest OS and the hypervisor, but the performance advantage can vary greatly, depending on the workload. However, each guest operating system, such as Linux, needs to be modified. VMware has traditionally offered full virtualization, but aspects of para-virtualization have been offered as an option for enhanced device drivers that increase the efficiency of guest operating systems.

Hardware-Assisted Virtualization

Recent CPUs from Intel (Intel VT) and AMD (AMD-V) implement hardware assistance for CPU virtualization; the first generation of these CPUs was released in 2005 and 2006. This method overcomes some of the problems in x86 virtualization that originally led companies such as VMware to pursue full virtualization with a binary translator. Although the binary translator can outperform first-generation hardware-assisted virtualization, future enhancements are expected to improve performance and flexibility in the programming model. Hardware-assisted virtualization can be considered a special aid to enable virtualization, and it gives the x86 architecture some capabilities of mainframe CPUs that were lacking in the original x86 CPUs. After all, most people didn't expect x86 computers to be powerful enough to run multiple operating systems at once, but now they are capable of running many virtual machines at once.

Device Virtualization

Note that CPU virtualization is not sufficient to create a fully functional virtual machine. VMM provides many other critical components, such as the computer's memory (memory management unit, also known as MMU), devices, and I/O, which are required for a fully functional x86 computer. The complexity of creating fully virtualized memory, devices, and I/O subsystems can be as great as the effort required for core CPU utilization itself. Hardware virtualization virtualizes the underlying hardware as *virtual devices* and presents them to the guest operating systems, and the virtual hardware presented is consistent for all virtual machines regardless of the underlying physical hardware. This means that a virtual machine running on one hardware platform can easily be moved to another platform, as the virtual devices are the same.

For example, let us take a Brand A computer, installed with virtual machine software and configured with Linux to run within a virtual machine. That instance of Linux is not configured for the Brand A computer (in the device drivers, size of disk, etc.). Instead, it is configured to run against the configuration of the virtual machine (a "VMware brand PC"). If you choose to replace the Brand A computer with Brand B, you simply move the virtual machine (which is just a set of files) to the Brand B computer, and the instance of Linux you installed earlier will run without any need to reconfigure its drivers. This greatly simplifies the difficulties associated with hardware upgrades. Furthermore, VMware Infrastructure also provides VMotion capability, which migrates a running virtual machine from one computer to another (provided they share the same storage where the virtual machine resides) without downtime.

Here are some specific examples. In a VMware virtualization environment, network adapters are presented as AMD PCNet devices, storage devices are presented as SCSI (even if the underlying physical devices are SAN, iSCSI, or SATA devices), and CPUs are presented as the underlying CPU architecture type. Other computer devices such as DVD/CD-ROM drives and floppy drives are presented using either physical devices or device images as the source.

Other Forms of Virtualization

There are other methods of virtualization not directly related to VMware Infrastructure. Some of the commonly used terms for these methods are given below.

Emulation

A virtual machine simulates the hardware needed in a way that allows it to run on a platform with a different CPU than it was originally designed to work with, e.g., PearPC, a PowerPC platform emulator; Microsoft Virtual PC, a 32-bit x86 emulator for Apple PowerPC Macintosh; Bochs, a 32-bit x86 emulator project.

Storage Virtualization

The process of abstracting a logical storage device from a physical device. These are found in many areas, from virtual disks in VMware products to Fibre Channel or IP network storage devices such as IBM's SAN Volume Controller (SVC), EMC Invista, or LeftHand Networks Virtual SAN Appliance (VSA). The physical location of the storage can be on a SAN, but the representation might be iSCSI. The software handles the mapping between the physical storage and the logical storage.

Network Virtualization

VLANs (virtual LANs) are used to segment networks on physical networks. This is different from the virtualized network devices available in VMware, although these two technologies can coexist.

Summary

This has been a quick overview of virtualization, spanning virtual machines, virtualization technologies (software- and hardware-level virtualization), approaches to CPU virtualization (full, para-, and hardware-assisted), what virtual machine monitors provide, and four key properties of virtual machines.

In your daily work and conversations, it's unlikely that these terms will come up, but understanding these concepts might become useful if you ever get into discussions where some parties are confused about basic concepts in virtualization. A four-way server built with dual-core CPUs and 32GB RAM, for example, is going to be used as a virtualization host running VMware ESX. A beginner's misconception might be that the four-CPU system will be logically partitioned into four virtual machines so that each partition is a two-core CPU, and you partition the RAM into 8GB partitions each. This type of misconception can stem from knowledge of LPARs, but you know that hardware-level virtualization is quite different from the flexible software-level virtualization, where you can create a wide variety of virtual machines to run on this system. Examples include:

- Eight virtual machines, each with a single virtual CPU and 4GB RAM
- Four virtual machines, each with a dual virtual CPU and 8GB RAM
- Four virtual machines, each with a dual virtual CPU and 16GB RAM

You might say, "Wait! Four times 16GB is 64GB, which is more than the 32GB of RAM on the physical system. How is that possible?" The answer is that some virtualization systems, such as VMware ESX, use advanced techniques (transparent page sharing, ballooning, and swapping) to enable over-commit of memory. This is somewhat of a mental somersault, because each virtual machine is probably managing memory inside the operating system, and the virtual machine software is doing even more memory management at the virtualization layer. This is all possible, although there might be some performance issues when there are too many virtual machines sharing a limited pool of physical memory.

Another misconception can be that hardware-assisted virtualization, such as that provided by the new class of CPUs, will make virtualization software obsolete, but CPU virtualization is not the whole story. Device and I/O virtualization, as well as the rich set of functionality provided by VMware Infrastructure, are critical to providing a complete virtualization solution.

Over time, most people in the IT industry will come to understand these distinctions, and you, as someone interested in deploying VMware Infrastructure, probably already know the differences and capabilities, but having background information is always useful.

Appendix B. Virtualization Ecosystem

A vast virtualization ecosystem has grown along with VMware, encouraged by published VMware vSphere APIs to enable direct integration by manufacturers in security, management, storage, and networking. For example, leading security manufacturers such as Symantec, McAfee, Check Point, Trend Micro, RSA, and Altor include VMsafe virtual appliance plug-ins that automatically protect everything running within the vSphere virtual machines without requiring the deployment of individual agents.

The three items included here are examples of these technologies. We have included specific examples from the host platform, networking, and software categories.

Hardware

Cisco is one of VMware's key partners and produces two products specifically designed to enhance a virtual infrastructure: the Nexus 1000V virtual switch and Unified Computing System (UCS).

Cisco Nexus 1000V

The Nexus 1000V, also discussed in Chapter 11 on security, is a software switch built on Cisco NX-OS specifically for VMware vSphere. It extends virtual networking to include IP/port security rules, multi-host PVLAN, flow statistics, IGMP snooping, 802 1Q tagging, rate limiting, and quality of service. The vSwitch and Port Groups are provisioned along with the physical network infrastructure, using the familiar NX-OS CLI, enabling more efficient configuration by the network team.

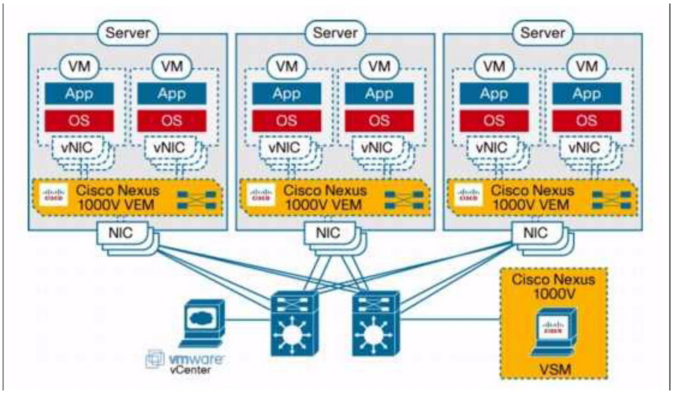


Figure 44: Cisco Nexus 1000V architecture

The Nexus 1000V facilitates a higher percentage of server virtualization. DMZ applications, for example, can be virtualized with the help of private VLAN isolation and security policy enforcement with ACL. Regulatory applications can be virtualized with Netflow, ERSPAN, and port statistics that persist after vMotion, and Tier-1 applications can be virtualized with increased visibility and I/O optimization with LACP, vPC host mode.

Cisco Unified Computing System (UCS)

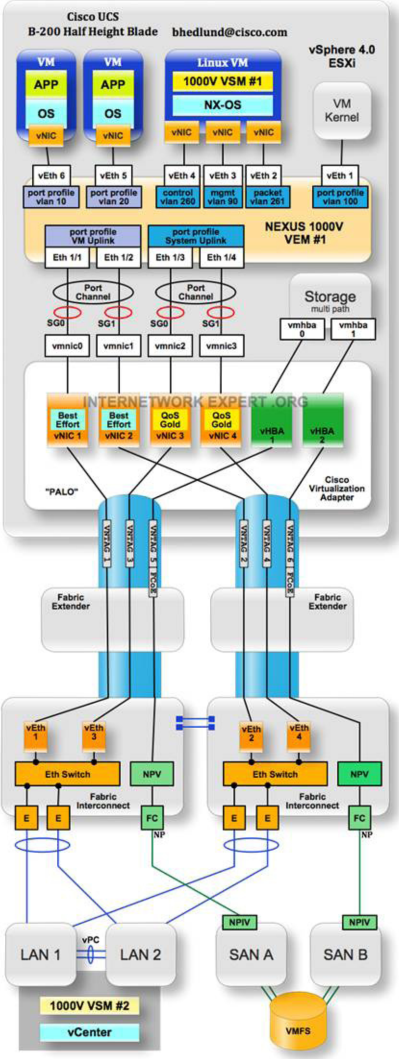


Figure 45: Cisco Unified Computing System B-Series (courtesy Brad Hedlund, Cisco)

Cisco UCS is an example of an ecosystem computing platform architected from the ground up to optimize the hosting of virtual infrastructure. It includes a hardware version of Cisco's Virtual Network Link (VN-Link), enabling optimized I/O performance while simplifying management with its tight integration with VMware vCenter Server. Other innovations in extended memory capabilities, unified fabric support, and stateless computing accelerate and simplify virtual infrastructure provisioning, performance, and management. UCS is also the compute foundation for Acadia, the partnership between VMware, Cisco, and EMC.

Software

There are a number of commercial and open source tools to manage a virtual infrastructure, including those that support a number of virtualization platforms. An example of an open source tool is included here.

Cfengine

Cfengine was the first open source system configuration system, developed originally in 1993 as a research project by university researcher and professor Mark Burgess. It was recently overhauled in 2008 as Cfengine 3 to take advantage of user experience and to simplify its "promise language." In 2009, a commercial version called Cfengine Nova also became available, with enhancements and support.

Cfengine is a useful player in a virtualization strategy because it can configure systems "hands-free" both on physical hosts and virtual guests, and its architecture can be fully decentralized, allowing it to scale up to tens of thousands of machines, the kind of scenario one expects to find in modern hosting centers and cloud computing dens. Cfengine can recover from network outages and clock drift, so it can work well even when systems are suspended for a length of time. Cfengine is used for deployments, change control, and consistency or security checks, and it can also perform detailed performance monitoring and compliance auditing.

Cfengine 3 has a few strengths that are useful for virtualization. It performs system configuration and monitoring in a very lightweight fashion, with very little memory and CPU overhead. This is especially useful in a virtualized environment where you want to pack many virtual hosts onto limited hardware. Second, Cfengine's now widespread term *convergence* or *computer immunology* can bring a system from basically any initial state into a state that is compliant with a defined policy. A long history has made Cfengine available on all kinds of operating systems, including pretty much all UNIX variants, Linux, Macintosh, and Windows, and the latest commercial versions of Cfengine Nova even run natively on Windows, supporting registry management and other Windows oddities.

At the time of writing, there are plans to support VMware as well as other virtualization engines directly in Cfengine, to simplify the deployment and management of virtual machines. There are even ambitions to bring adaptive power-saving issues into the repertoire of ordinary mortals, as discussed in Green IT initiatives.

Manage Large Networks (MLN)

Manage Large Networks (MLN) is a management tool specialized for rapid deployment and management of large numbers of virtual machines. It is vendor-agnostic, supporting

VMware, Xen, and other virtualization technologies. Instead of using a graphical interface, MLN utilizes a configuration language. This language allows accurate descriptions of the virtual machines along with their internal configuration, such as users and network configuration. A group of virtual machines is called a project, and the configuration file describing the project can make use of super classes and inheritance, along with variables, to shorten the amount of redundant text and remove errors. An extendable plug-in framework allows users to extend the functionality of MLN.

MLN offers command-line management with the possibility of managing a project as an atomic unit, e.g., starting, stopping, or polling status information on all virtual machines belonging to the same project with one command. With the MLN daemon running, projects can span several servers, allowing complicated scenarios to be realized effectively.

MLN was developed by Kyrre M. Begnum, PhD, associate professor at Oslo University College, as part of his PhD program.

Glossary

In this glossary, multi-work terms that start with “VMware,” “vSphere,” or “vCenter” are listed alphabetically using the word that follows in the term. “VMware vCenter term” identifies technologies that are integrated with VMware vCenter, some of which are considered add-on products requiring additional licensing. “VMware term” identifies technologies that are stand-alone products. Where applicable, definitions are taken from the formal VMware glossary guide.

A

VMware ACE—Virtualization product for enterprise desktop deployments providing a highly configurable, secure, and portable PC environment.

ACE instances—The virtual machines that ACE administrators create, associate to virtual rights management (VRM) policies, and package for deployment to users. An ACE instance is an ACE.

VMware vCenter AppSpeed—An application performance monitoring tool engineered specifically for multi-tiered applications. AppSpeed passively listens to traffic flowing over a vSwitch (including the Nexus 1000V), which allows for discovery of transactions, application mapping, performance monitoring against SLAs, and root-cause analysis. Provides a method to evaluate performance of an application before and after virtualization to ensure that performance remains consistent. This tool provides *breadth* in latency analysis for an application.

C

VMware vCenter CapacityIQ—Identifies server resource inventories including used and unused capacity. This can be used for capacity planning, budgeting, and lifecycle management of resources. VMware vCenter Capacity IQ is used for cost avoidance and justification, availability and risk mitigation, and project planning and decision-making.

VMware Capacity Planner—An agentless data collection and “what if” scenario building tool that identifies server inventories and resource utilization to determine virtual machine candidates, server consolidation ratios, and resource requirements for migrating to a VMware Infrastructure based on target ESX host platform resources.

VMware vCenter Chargeback—Provides cost measurement, analysis, and reporting to provide cost transparency and accountability for the virtual machines and the supporting virtual infrastructure. IT costs may be mapped to business units, cost centers, or external customers to provide a better understanding of resource costs. This can further be used to determine optimization for cost reduction.

VMware vCenter Chargeback API—Provides an interface for vCenter Chargeback functionality. This includes management of the hierarchy, cost configurations, and reporting.

CIM Interfaces—Software interfaces designed for hardware management tool development. This includes **Server Management API**—CIM SMASH interface to monitor and manage virtualization server platforms and **Storage Management API**—CIM SIMI-S interface to monitor and manage virtual storage.

Cluster—A server group in the virtual environment that enables a high-availability solution.

vCLI (vSphere Command Line Interface)—Allows you to manage your Virtual Infrastructure using Windows PowerShell. This allows you to script and automate actions you would normally do in vCenter. There are approximately 200 cmdlets (PowerShell exposed procedures) to manage vSphere and ESX/ESXi functionality. There are many pre-built scripts available online that can provide functionality such as finding all VM snapshots, finding orphaned VMs, or even creating reports. Previously known as the VI ToolKit.

VCB (VMware Consolidated Backup)—Provides the capability to perform SAN-based backup and recovery of virtual machines using a backup proxy server without any network or virtual machine overhead.

VMware Converter—Used for physical-to-virtual machine (P2V) migrations, as well as importing virtual machines from other virtualization vendors (V2V). VMware Converter can import multiple machines concurrently and non-disruptively. Designed for large-scale consolidation, VMware Converter can be used with or without VMware vCenter Server.

D

Datacenter—In the context of vCenter usage, an optional inventory grouping structure contained within the datacenter structure. A vCenter Server supports multiple datacenter folders. Datacenter folders can contain only datacenters and other datacenter folders.

vDR (VMware Data Recovery)—Provides data protection for virtual machines. VMware Data Recovery is fully integrated with vCenter Server and includes data de-duplication to save on disk storage for full virtual machine backups. Includes file-level restore or entire images as needed.

Data Source Name (DSN)—An ODBC (Open Database Connectivity) object that you must configure to enable vCenter Server to access a database.

DPM (Distributed Power Management)—Dynamically starts up and shuts down ESX host hardware to reduce power consumption.

DRS (Distributed Resource Scheduler)—Dynamically allocates and balances workloads across hosts in a cluster.

DV Port Group—A port group associated with a Distributed Virtual Switch (DVS). It specifies port configuration options for each member port. It defines how connections are made through the DVS to the network.

DV Uplink (DV Uplinks)—Physical uplinks attached to a vDS to enable VMs and virtual network adapters connected to vNetwork Distributed Switch to connect to networks outside the hosts on which they reside.

DV Uplink Groups—Defines uplink policies for the associated DV Uplinks.

DV Port (Distributed Virtual Port)—A port on a DVS that connects to a host's Service Console or VMkernel or to a virtual machine's network adapter.

DVS (Distributed Virtual Switch)—An abstract representation of multiple hosts defining the same vSwitch (same name and network policy) and port group. These representations are needed to explain the concept of a virtual machine being connected to the same network as it migrates among multiple hosts.

E

Elastic Sky—The acronym ESX was created from the term Elastic Sky, a marketing label that was created but not used by VMware. The initials plus the letter X became the official product name. Elastic Sky was used as the name for a VMware band that was founded by Jeff Hanson, John Arrasjid, Melinda Marks, David Haberman, Doug Clark, Drew Kramer, and Ken Watson. Additional band members have participated with Elastic Sky, including Tim Mann, Linda Pak, Norman Malonzo, Arisa Amano, Vittorio Viarengo, and Robert Noth.

Emulation—In which a virtual machine simulates the hardware needed in a way that allows it to run on a platform with a different CPU than it was originally designed to work with.

VMware ESX/ESXi Server—ESX and ESXi are both hypervisors which install directly on the server hardware. Although the deployment and management methods are slightly different, both solutions provide better performance and availability than other methods. VMware ESX Server was initially released in 2002.

VMware ESX Server (VMware ESX Classic)—Classic ESX installs with a Linux-based Service Console to assist with management functions.

VMware ESXi Server—A version of VMware ESX Server that may be installed like VMware ESX Classic, on USB flash storage, or on an internal device. ESXi removes the Service Console, reducing the attach surface due to a smaller footprint and allowing the functionality to be embedded within the server hardware.

F

VMware Fault Tolerance or **VMware FT**—Provides clustering support of single vCPU VMs without requiring the embedded application to be cluster aware. FT utilizes VMware vLockstep technology. This technology uses an active secondary VM that runs in virtual lockstep with the primary VM. VMware vLockstep establishes and maintains this secondary VM. The secondary VM runs on a different host and executes the same set of instructions, in the same sequence, as the primary VM.

Full Virtualization (Native Virtualization)—The guest OS is presented with a virtual hardware abstraction that represents a physical machine. The virtual machine is recognized and accessible to the operating system or applications software just as if it were a physical machine, so no modification to the software is necessary. With full virtualization, a standard operating system such as Windows or Linux, without modifications, will run in a virtual machine.

VMware Fusion—A virtualization product for Intel-based Mac OS X systems.

G

GSX Server—Original name for VMware Server. GSX existed as a commercial product from 2001 to 2006. The acronym is created from the term Ground Storm, a marketing label that was created but not used by VMware. The initials were used and the letter X was added to create the official product name.

GOS (Guest Operating System)—An operating system that runs within a virtual machine.

vSphere Guest SDK—Enables development of applications that will run within a virtual machine using C or Java libraries. Enables customers to write smart applications that respond to changes at the virtualization environment layer. Included with VMware Tools.

VMware Guided Consolidation—Used for planning physical-to-virtual machine migrations by utilizing VMware Capacity Planner Converter technology. VMware Guided Consolidation is an optional vCenter component and is designed for small-scale consolidation.

H

Hardware-assisted Virtualization—CPUs from Intel (Intel VT) and AMD (AMD-V) implement hardware assistance for CPU virtualization; the first generation of these CPUs was released in 2005 and 2006.

vCenter Heartbeat—Protects the vCenter Server, License Server and Database against hardware, OS, application, and network downtime. Failover and failback are provided for each. Protection is important especially when using VMware View, vCenter Lab Manager, and vCenter SRM, which require vCenter to be running at all times.

VMware High Availability (VMware HA)—Provides automated restart of failed virtual machines, regardless of the guest OS technology. Provides fault tolerance in the event of an ESX host failure. VMware HA enables the automated restart of virtual machines on other hosts in a cluster upon host failure, minimizing downtime without the cost of application clustering.

Host—A compute platform supporting the execution of virtual machines. Includes standard physical servers as well as platforms specifically designed to support virtual infrastructure such as Cisco UCS.

VMware Host Profiles—Enables the definition and application of standardized host configurations. Also supports compliance checks against the defined standards.

Hosted Virtualization—Relies on having a standard operating system between the physical computer and the virtualization layer. This requires installation of an operating system such as Microsoft Windows or Red Hat Linux, and then installation of virtualization software such as VMware Workstation on top of it. Finally, a guest operating system such as Windows or Linux is installed in one or more virtual machines running within VMware Workstation.

VMware Hyperic HQ—Provides complete discovery, monitoring, and analysis and control of all application, system, and network assets both inside and outside the virtual machines. Hyperic HQ includes full VMware ESX and VMware Server support, analysis of utilization and performance within a VM, correlation of events between hosts and guest OSes, and control of VMs. This tool provides detailed analysis of how the virtual machine is performing. This tool provides *depth* in latency analysis for an application.

Hypervisor—Hypervisor virtualization platforms have a partitioning layer, which runs directly on top of the hardware and below higher-level virtualization services, that provide a virtual machine abstraction. The hypervisor is installed on the computer, just as though it is an operating system. It provides the capability to create virtual machine partitions, with a virtual machine monitor running within each partition.

The Hypervisors—A VMware band.

L

VMware vCenter Lab Manager—Provides a self-service portal for real-time provisioning, managing, and collaboration of virtualized development and testing environments. VMware vCenter Lab Manager allows developers and testers to create and share libraries of virtualized application environments used in software development and testing. Applications can be moved through lifecycle stages until they reach production state.

VMware vCenter Lab Manager SDK—Enables development of applications that use Lab Manager Web service data, automate tasks, or integrate VMware Lab Manager with software testing tools.

VMware Lifecycle Manager or **VMware LCM**—Manages the lifecycle of virtual machines from request through provisioning and eventual archiving or destruction. VMware vCenter Lifecycle Manager provides a self-service portal for virtual machine requests, routed through a predefined workflow—streamlining provisioning, reducing overhead, and providing consistent management of the virtual machine lifecycle.

Logical Partitioning (LPAR)—Found in mainframe computers such as IBM System z (and, less commonly, on other IBM systems), as well as on computer systems from other vendors. In logical partitioning the resources of a physical computer are partitioned so that the computer's memory may be split, allocating a specific range to each partition. Hardware assistance is often used to partition a system but is not necessary for operating system virtualization in general.

M

vSphere Management Assistant (vMA)—A Linux appliance with pre-built management tools and the vCLI Interface. Allows scripting and agents to manage ESX, ESXi, and vCenter Server systems. vMA is a virtual appliance that includes the vSphere SDK and the vSphere CLI, logging capabilities, and an authentication mechanism.

Multipathing Policy—When connecting an ESX host to a Fibre Channel SAN, the multipathing policy enables multipathing support to maintain a constant connection between the ESX host and the storage device in the event that a critical connecting component (e.g., host bust adapter, storage controller, storage processor, or Fibre Channel cable) fails.

N

Network virtualization—VLANs (Virtual LANs) are used to segment networks on physical networks. This is different from the virtualized network devices available in VMware, although these two technologies can coexist.

O

VMware vCenter Orchestrator (vCO)—Provides out-of-the-box workflows to help automate existing manual tasks. Workflows can be created, modified and extended to meet custom needs.

VMware vCenter Orchestrator API—Allows for the programming of workflows for execution by VMware vCenter Orchestrator.

P

Para-virtualization (OS-assisted virtualization)—The guest OS is presented with a modified hardware abstraction. This requires operating systems to be modified and ported to this particular virtualization platform. This reduces operating system compatibility, but that is a trade-off against potential increases in performance of certain CPU-bound applications that run on systems without virtualization hardware. This performance increase is achieved by hypercalls, a communication method that occurs between the guest OS and the hypervisor, but the performance advantage can vary greatly, depending on the workload. However, each guest operating system, such as Linux, needs to be modified. VMware has traditionally offered full virtualization, but aspects of para-virtualization have been offered as an option for enhanced device drivers that increase the efficiency of guest operating systems.

PC-over-IP (PCoIP)—PCoIP was a purpose-built protocol designed by Teradici to deliver a rich desktop experience consisting of content such as application windows, Web pages, graphics, text, and streaming video and audio over both the LAN and the WAN. PCoIP recognizes the different types of content and uses different compression algorithms based on content type. PCoIP delivers multi-monitor support with up to 1920 x 1200 resolution, clear-type fonts, and 32-bit color per monitor for up to four monitors. It includes Auto Display scaling, dynamic resizing, and support of monitor pivoting. It enables multimedia redirection, USB support, support for host-based rendering of Flash, and bi-directional audio.

VMware Player—Enables creating virtual machines and running virtual appliances.

Port Group—Specifies port configuration options, including VLAN tagging policies and bandwidth limitations for each member port. Network services connect to vSwitches through port groups. Port groups define how a connection is made through the vSwitch to the network. In typical use, one or more port groups is associated with a single vSwitch.

Port Groups—A construct for configuring virtual network options such as bandwidth limitations and VLAN tagging policies for each member port. Virtual networks that are connected to the same port group share network policy configuration.

vSphere PowerCLI—Allows you to manage your Virtual Infrastructure using Windows PowerShell. This allows you to script and automate actions you would normally do in vCenter. There are approximately 200 cmdlets (PowerShell exposed procedures) to manage vSphere and ESX/ESXi functionality. There are many pre-built scripts available online that can provide functionality such as finding all VM snapshots, finding orphaned VMs, or even creating reports. Previously known as the VI ToolKit.

S

vCenter Server (VMware Server)—Free entry-level server virtualization product for creating and running multiple virtual machines on existing physical Windows or Linux servers. Formerly titled GSX Server.

Service Console—The command-line interface for an ESX server system that enables administrators to configure the system. The Service Console is installed as the first component and is used to bootstrap the ESX server installation and configuration. The Service Console also boots the system and initiates execution of the virtualization layer and resource manager. You can open the Service Console directly on an ESX server system. If the ESX server system's configuration allows Telnet or SSH connections, you can also connect remotely to the Service Console.

VMware vCenter Site Recovery Manager (vCenter SRM)—Provides disaster recovery workflow automation through a centralized management interface. SRM automates the setup, testing, failover, and failback of virtual infrastructures between protected and recovery sites.

VMware vCenter Site Recovery Manager API—Provides an interface to SRM which allows external management systems to initiate tests or failovers and record results.

Storage Virtualization—The process of abstracting a logical storage device from a physical device. These are found in many areas, from virtual disks in VMware products to Fibre Channel or IP network storage devices such as IBM's SAN Volume Controller (SVC), EMC Invista, or LeftHand Networks Virtual SAN Appliance (VSA). The physical location of the storage can be on a SAN, but the representation might be iSCSI. The software handles the mapping between the physical storage and the logical storage.

Storage VMotion—Storage VMotion enables live migration of virtual machine disk files across storage locations while maintaining service availability. Storage VMotion utilizes VMotion technology to optionally move the VM to an alternate ESX host which has access to both the source and target storage locations. Storage VMotion can move the storage location of a virtual disk as long as the target is visible to the source and destination ESX host(s). The processes of the corresponding VM can stay on the same host, or the VM can be simultaneously VMotioned to a new host.

T

VMware ThinApp—Enables application virtualization, encapsulating the applications from both the OS and each other. This eliminates conflicts from badly behaving applications as well as the requirement for regression testing.

V

vApp—Provides a logical entity, or object, comprising one or more virtual machines using the OVF (Open Virtualization Format) to specify and encapsulate all components of a multi-tier application. In addition, policies and SLAs can be associated with the object as an attribute. The vApp construct is designed for interoperability of a multi-tiered application on the virtual datacenter as well as for the ability to move the application between internal or external clouds while maintaining the same SLAs.

Virtual Machine (VM)—This book focuses on "system virtual machines," a form of virtualization whereby the underlying physical computer resources are mapped into one or more different virtual machines (tightly isolated software containers that behave exactly like a physical computer).

VMDirectPath—Offloads I/O processing from the hypervisor by allowing virtual machines to directly access the underlying hardware devices.

VMware Certified Professional or **VCP**—A certification designed for individuals wishing to demonstrate expertise in virtual infrastructure and increase potential for career advancement.

VMware Certified Design Expert or **VCDX**—The highest level of VMware certification, evidencing an exceptional proficiency in designing and implementing successful VMware infrastructures. John Arrasjid, lead author of this booklet, is VCDX #1 and Duncan Epping is VCDX #7.

vCLI—The vSphere Command-Line Interface command set enables running common system administration commands against ESX/ESXi systems from any machine with network access to those systems. vSphere CLI commands are especially useful for ESXi hosts that do not include a Service Console.

vCloud—The VMware vCloud initiative consists of technology from VMware as well as the ecosystem of technology and cloud service providers to enable application delivery on a common VMware vSphere platform. Nearly 1,000 validated applications are easily deployed to an on-premise environment or to a cloud without requiring recoding or rebuilding.

vCloud API—Provides an interface for providing and consuming virtual resources within a VMware-based cloud by enabling deployment and management of virtualized workloads by working with vApps. This API is based on OVF standards providing platform independence and multi-tenancy in a purely virtual infrastructure. Includes functions for Inventory Listing, Catalog Management, Upload/Download/Provisioning Operations, vApp Configuration Operations, Resource Entities Operations, vApp State Operations, and other operations. Also includes administrative functions including Cloud, Org, vDC, Catalog, User, Group, and Role Administration.

vCloud Express—An Infrastructure as a Service (IaaS) offering providing pay-as-you-go infrastructure that ensures compatibility with internal VMware environments.

vCPU—A virtual central processing unit is similar to the CPU in a traditional physical machine. A virtual processor is assigned (either one or in multiples) to a virtual machine.

VMware View—A system for managing connectivity, security, and administration of centralized virtual desktop computers hosted on ESX clusters. VMware View Manager supports the connection brokering for the Virtual Desktop Infrastructure (VDI), while View Composer provides advanced desktop image management.

VMware View Composer—Provides advanced desktop image management for the Virtual Desktop Infrastructure (VDI).

VMware View Manager—Supports the connection brokering for the virtual desktop infrastructure (VDI).

VIX—Allows development of programs and scripts to automate virtual machine and guest OS operations. VIX runs on Windows or Linux platforms. It manages VMware vSphere, ESX, ESXi, VMware Server, and VMware Workstation through the use of C, Perl, and COM bindings, including Visual Basic, VBscript, and C#.

VIX API—Provides a programming interface to manage and automate functionality of OS guests. Allows a script to be executing through vCenter using VMware Tools rather than being sent across the VM networks.

Virtual Desktop Infrastructure (VDI)—Virtual Desktop Infrastructure is the term coined by VMware to describe the process of running desktop operating systems and applications as virtual machines on an ESX host.

Virtual Disk—A virtual disk simulates a physical disk in memory and is usable although not physically present in the computer.

Virtual Disk Development Kit (VDDK)—Interface to allow ISVs to use VMDK as a native format when developing virtual disk tools through the use of the VMware Virtual Disk Libraries (VixDiskLib and ViMntapi).

Virtual Machine Monitor (VMM)—The VMM is a layer of software that runs between the hypervisor or host operating system and a virtual machine. It manages the resources and their allocation to the virtual machines running on the system. The VMM decouples the software from the hardware underneath. As computer scientist Butler Lampson famously said, "All problems in computer science can be solved by another level of indirection." VMM's decoupling capability provides substantial control over how the guest operating system accesses the hardware.

Virtual Operating System—A small, lightweight component embedded with ThinApp compiled applications.

Virtual SMP—Virtual SMP allows a single virtual machine to use multiple physical processors simultaneously. It enables virtualization of even very resource-intensive applications such as databases, large Web servers, and ERP.

Virtual Switch—A software program emulating a physical switch to enable one virtual machine to communicate with another. See *vNetwork Standard Switch* and *vNetwork Distributed Switch.*

vLockstep—Virtual Lockstep technology is utilized as part of vSphere Fault Tolerance. It facilitates zero downtime and zero data loss for a virtual machine by keeping a second instance of the FT-enabled virtual machine in lockstep with the primary instance**.**

VMDK—The Virtual Machine Disk Format is a file format used to store a virtual machine image and the contents of a virtual machine hard drive.

VMFS—Virtual Machine File System is a cluster file system enabling storage of virtual machine disk images, including snapshots.

VMkernel—The VMkernel is a microkernel handling CPU and memory directly as part of CPU instructions. It provides a hardware simulation interface to guest systems.

VMmark—A benchmark tool specifically designed for measuring scalability of virtualization host systems. Provides an accurate measurement of application performance in virtualized environments. Measures virtual machine performance, determines how different hardware and virtualization platforms will affect performance, and enables "best fit" choices for hardware. VMware is working with the Standard Performance Evaluation Corporation (SPEC®) and members of the SPEC Virtualization subcommittee to develop standard methods of comparing virtualization performance for virtualized applications running on hypervisors.

VMotion—VMware VMotion enables the live migration of running virtual machines from one physical host to the other with zero downtime, continuous service availability, and complete transaction integrity. VMotion migration requires either the same processor family on both the source and target ESX hosts, or "Enhanced VMotion Compatibility" (EVC) on a cluster of hosts with technologies enabling VMotion compatibility with older servers. Hosts need to be grouped within the same vCenter datacenter. The shared storage holding the VM virtual disk is presented to both the source and target hosts.

VMware VMsafe—Provides an open approach to security through an application program interface (API). This enables selected partners to develop security products for VMware environments. VMsafe gives fine-grained visibility over virtual machine resources, making it possible to monitor every aspect of the execution of the system and stop previously undetectable viruses, root kits, and malware before they can infect a system. VMsafe provides inspection of virtual machine memory pages and CPU states, filtering of network packets inside hypervisors as well as within the virtual machine itself, and in-guest, in-process APIs that enable complete monitoring and control of process execution. Guest virtual machine disk files can be mounted, manipulated, and modified as they persist on storage devices.

VMsafe API—Allows vendors to develop advanced security products.

vNetwork—vNetwork refers to the collection of networking technologies enabling optimal integration of networking and I/O functionality into vSphere. The vNetwork enhancements include the vNetwork Distributed Switch, VMXNET3 (third-generation para-virtualized NIC), IPv6 support extended to VMkernel and Service Console ports, bi-directional traffic shaping, network VMotion, and VMDirectPath.

vNetwork API—Provides integration with the virtual networking capabilities of vSphere to enable the development of advanced network tools.

vNetwork Distributed Switch (vDS)—Provides a switch that acts at a datacenter level across multiple ESX hosts, providing centralized provisioning, administration, and monitoring. Simplifies network management by moving the virtual network configuration and management from the host level to the datacenter level.

vNetwork Standard Switch (vSS)—A software program emulating a physical switch to enable one virtual machine to communicate with another. It is a basic Layer 2 switch without routing.

vnicvNIC—A virtual network interface card is a virtual representation of a physical NIC and is configured on top of a system's physical network adapter.

VMware vShield Zones—Enforces corporate security policies at the application level in a shared environment, while still maintaining trust and network segmentation of users and sensitive data. Provides a mechanism to monitor, log, and block inter-VM traffic with an ESX/ESXi host or between hosts in a cluster. This includes the ability to firewall, bridge, or isolate virtual machines between multiple pre-defined zones. All activities, blocked as well as allowed, are logged and can be graphed.

VMware vSphere—A collection of software providing management of a dynamic environment, cost reduction, and significant improvement to the life-work balance of IT professionals. VMware vSphere is the next generation of the Virtual Infrastructure 3 product. Known also as the first cloud operating system, vSphere 4 was one of the most complex software development projects of all time, consisting of over 3,000,000 engineering hours by over 1,000 engineers over a three-year period.

vSphere SDK—A software development kit that acts as an interface ESX/ESXi, vCenter, and VMware Server to extend the management of the Virtual Datacenter. Programming languages supported include Perl, .NET, and Java.

vSphere SDK for Java—Supports simplified vSphere Management applications by defining client-side data models. These models provide utility functions to simplify data access to servers.

vStorage—Provides integration of advanced capabilities from storage vendors with the vSphere Cloud OS from VMware. This API enables customers to leverage array-based capabilities such as support for multipathing control, which enables advanced load balancing algorithms.

vStorage API—A collection of software providing management of a dynamic environment, cost reduction, and significant improvement to the life-work balance of IT professionals.

vswif#— The VMware vCenter Server Service Console NIC.

vSwitch—See Virtual Switch.

W

vSphere Web Services SDK—Provides a Web service accessible through the vSphere API to develop client applications.

VMware Workstation—Enables running multiple virtual machines on a PC, each with its own operating system. VMware Workstation resides on top of a host operating system and supports over 200 guest operating systems. It includes clones, multiple snapshots, and secured "virtual rights management" features. See also *VMware Player*.

Booklets in the Series

#21: **Cloud Computing with VMware vSphere,** by John Arrasjid, Duncan Epping, and Steve Kaplan

#20: **Running the Numbers: System, Network, and Environment Monitoring,** by Daniel V. Klein and John Sellens

#19: **Configuration Management with Bcfg2,** by Narayan Desai and Cory Lueninghoener

#18: **Deploying the VMware Infrastructure,** by John Arrasjid, Karthik Balachandran, Daniel Conde, Gary Lamb, and Steve Kaplan

#17: **LCFG: A Practical Tool for System Configuration,** by Paul Anderson

#16: **A System Engineer's Guide to Host Configuration and Maintenance Using Cfengine,** by Mark Burgess and Æleen Frisch

#15: **Internet Postmaster: Duties and Responsibilities,** by Nick Christenson and Brad Knowles

#14: **System Configuration,** by Paul Anderson

#13: **The Sysadmin's Guide to Oracle,** by Ben Rockwood

#12: **Building a Logging Infrastructure,** by Abe Singer and Tina Bird

#11: **Documentation Writing for System Administrators,** by Mark C. Langston

#10: **Budgeting for SysAdmins,** by Adam Moskowitz

#9: **Backups and Recovery,** by W. Curtis Preston and Hal Skelly

#8: **Job Descriptions for System Administrators, Revised and Expanded Edition,** edited by Tina Darmohray

#7: **System and Network Administration for Higher Reliability,** by John Sellens

#6: **A System Administrator's Guide to Auditing,** by Geoff Halprin

#5: **Hiring System Administrators,** by Gretchen Phillips

#4: **Educating and Training System Administrators: A Survey,** by David Kuncicky and Bruce Alan Wynn

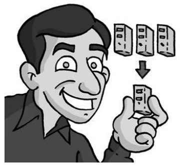
#3: **System Security: A Management Perspective,** by David Oppenheimer, David Wagner, and Michele D. Crabb, and edited by Dan Geer

#2: **A Guide to Developing Computing Policy Documents,** edited by Barbara L. Dijker

About the Authors

John Y. Arrasjid is a Principal Architect in the Technical Services Advanced Services group at VMware. He specializes in virtualization aspects of business continuity, disaster recovery, and enterprise cloud. His experience includes work with AT&T, Amdahl, 3Dfx Interactive, Kubota Graphics, Roxio, and his own company, WebNexus Communications. As an architect, John develops IP for customers, field consultants, and partners on topics including enterprise cloud, business continuity, disaster recovery, performance, security, and virtualized datacenters. John's early contributions with the vmsnap/vmres scripts led to the development of VMware Consolidated Backup. John regularly presents at VMworld, VMware Partner Exchange, USENIX Annual Technical Conference and USENIX LISA, LinuxWorld, and other conferences. John is also a founding member of the two VMware music bands Elastic Sky and The Hypervisors. John is a VMware Certified Professional and one of the first VMware Certified Design Experts (VCDX 001). John holds a Bachelor of Science in Computer Science degree from SUNY at Buffalo. He can be followed on Twitter at http://twitter.com/vcdx001.

Duncan Epping is a Senior Consultant and the Datacenter Practice Lead for VMware. Duncan works primarily with enterprise customers. He is focused on designing virtual infrastructures and specializes in business continuity, disaster recovery, and VMware HA. Duncan is a VMware Certified Professional and among the first VMware Certified Design Experts (VCDX 007). Duncan is the owner of Yellow-Bricks.com, one of the leading VMware/virtualization blogs worldwide and an active contributor and moderator on the VMTN Community Forums. He can be followed on twitter at http://twitter.com/DuncanYB.



Steve Kaplan is Vice President, Data Center Virtualization Practice for INX. Steve has authored scores of articles, white papers, and books on different aspects of virtual infrastructure and is the author of the *VirtualMan* comic book series. He has spoken on datacenter and desktop virtualization at venues around the globe and maintains Bythebell.com, a blog site emphasizing the economics of virtualization. Steve holds a Bachelor of Science degree in business administration from the University of California, Berkeley, and an MBA from Northwestern University's J.L. Kellogg Graduate School of Management. He can be contacted at Steve.Kaplan@inxi.com or followed on Twitter at http://twitter.com/roidude.