THE USENIX SIG FOR

[sage]

SYSADMINS

ISBN-13: 978-1-931971-62-1

18 Arrasjid, Balachandran, Conde, Lamb, and Kaplan Deploying the VMware Infrastructure

18

Short Topics in
System Administration

Jane-Ellen Long, Series Editor

Deploying the VMware Infrastructure

John Arrasjid, Karthik Balachandran, Daniel Conde, Gary Lamb, and Steve Kaplan

THE USENIX SIG FOR

[sage]

SYSADMINS

Booklets in the Series

#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

18 *Short Topics in*
System Administration

Jane-Ellen Long, Series Editor

Deploying the VMware Infrastructure

John Arrasjid, Karthik Balachandran,

Daniel Conde, Gary Lamb, and Steve Kaplan

Published by the USENIX Association
2008

About SAGE
SAGE is a Special Interest Group of the USENIX Association. Its goal is to serve the system administration community by:

- Offering conferences and training to enhance the technical and managerial capabilities of members of the profession.
- Promoting activities that advance the state of the art or the community.
- Providing tools, information, and services to assist system administrators and their organizations.
- Establishing standards of professional excellence and recognizing those who attain them.

SAGE offers its members professional and technical information through a variety of programs. Please see http://www.sage.org for more information.

Authors' Note: We have changed some names for VMware products from the first printing of this booklet in order to adhere to product name changes that the company has undertaken. These include changes from VirtualCenter to VMware vCenter and from Virtual Desktop Manager to VMware View Manager.

ISBN 978-1-931971-62-1

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/

Second Printing 2008

Contents

Acknowledgments

Thank you to John Gannon and Shridhar Deuskar for content contribution. Thank you to editor extraordinaire and writer guru, Matthew Wood, for his attention to detail, document flow, and support. Thank you to Evelyn Eldridge of VMware and Bryan Dickson of INX for editorial assistance. I would like to thank my co-authors in creating this SAGE Short Topics booklet. This document has also benefited from review by several individuals, and we would like to thank them for their efforts. They are Gretchen Phillips, Todd Massey, Karen Zeller, Jarrod Swetland, Cheryl Eagan, Philip Callahan, Kris Boyd, Mark Broda, and Lance Owen. Thank you to my VMware management team (Lawrence Rupp and Jason Martin) for supporting the creation and release of this material to the SAGE community.

Thank you to Amy, Catherine, Sofi, and Lila, whose love and support have enabled me to complete this project, and to my parents, Dorine and Harun, for years of encouragement and sacrifice to ensure that family always came first. This book is dedicated to my family.

Thank you to the USENIX Association for inviting me to create and deliver tutorials around VMware virtualization technology for the past few years, and for encouraging the publication of this booklet for the SAGE community. USENIX continues to hold a high standard for both the research and administration aspects of system administration.

This booklet is dedicated to Diane Greene and Mendel Rosenblum for their leadership, support, and friendship.

This booklet gives you a starting point for understanding the VMware Infrastructure and deploying it for cost reduction, quicker deployments of systems, and better control of resource utilization, as well as datacenter management and high availability. Welcome to VMware Infrastructure.

John Y. Arrasjid
VCP, VMware, Inc.

Foreword

After years of working in the computer industry, we have come to realize that although almost everything has been done before, approaches to problems have changed.

The evolution in processing power and declining costs has driven many of these changes. In the early days of computing, hardware was expensive. In the 1960s, virtual machines were developed on mainframe computers, notably from IBM, to enable multiple users to share expensive resources. In the 1970s came mini-computers, such as those from Digital Equipment Corporation, which were cheaper and enabled decentralized computing at a departmental level. In both cases, remote terminals accessed these shared systems.

In the 1980s, personal computers (IBM, Apple) and UNIX engineering workstations (Sun, DEC, SGI) further popularized decentralized computing, as microprocessors became faster and cheaper. Instead of having to share a departmental computer, each user had his or her own, and virtual machines became less popular.

In parallel, during the 1980s, graphical user interfaces became popular in PCs. Client-server computing arose as a method to meld interactive local user-interfaces (either on a PC or a thin client) with a central server. The desktop PC architecture evolved into business-class servers, offering inexpensive commodity-based pricing for systems that have many of the capabilities of mini or mainframe computers. Client-server computing faded away as centralized Web-based systems started to take over, and the pendulum started to swing back towards centralized computing.

At the same time, the processors were increasingly underutilized as servers had multiple processors, each with multiple cores, and software demands flagged behind hardware capabilities. The proliferation of computers spread the workload to ever more distributed yet underutilized systems. Even desktop PCs were underutilized, because they also used faster processors.

Systems proliferated, but centralized computing continued to regain popularity as Web-based computing took hold. But unlike the central mainframes of the 1970s, the new model of centralization often consisted of a complex mesh of servers, frequently configured as multi-tiered systems.

Yet these systems were becoming harder to manage. Not only were there more of them, but they came from many vendors (not just IBM anymore), and the pace of change accelerated, which put higher demands on the need to rapidly test and deploy systems. The proliferation of desktop PCs continues to pose complex management problems.

Some concepts from the past have returned to help with this situation. Virtual machines, which were almost forgotten in the 1990s, re-emerged as a method to install and consolidate many server systems into one physical machine. Many virtual machines can run on

one host. Fast networks and CPUs now enable PCs to be hosted as virtual machines in central servers accessed from thin clients, just like the old time-sharing terminals, which improves the manageability of desktop computing. Complex, multi-tiered systems can be tested and deployed using virtual machines, which helps bring a wide array of configurations to the fingertips of developers, with no need to configure them by hand. Virtual machines present a common platform that simplifies software distribution for software developers—this helps tame the headaches of software installation and returns us to the simplicity of the "good old days" when there were few target platforms to worry about. Finally, the rise of Web-based systems has led to another reason to move back to the datacenter: cloud computing. Virtual machines offer a great solution for treating an entire cluster of servers as a single shared resource, sliced and diced according to the computing needs of the moment, not unlike the central mainframes of the past.

For the first-time user of virtualization, John Arrasjid and his co-authors have outlined and discussed the world of virtualization in layman's terms. Along with explaining how virtualization can be applied to today's businesses, they have shown how incredible the return on investment can be in terms of resource utilization and staff productivity. For Privacy Networks, as a software company, the ability to use virtualization has increased our sales-demo capabilities, engineering test productivity, and deployment of email archiving software for customers as a virtual appliance in a VMware environment. I (Todd Massey) am continually amazed at the uses we come up with for virtualization in our company. As you learn more about deploying enterprise-class virtualization, think outside the box—for business today, virtualization can increase productivity in ways that seem almost limitless.

Todd Massey
CTO, Privacy Networks

Daniel Conde
VCP, VMware, Inc.

1. Introduction

What is virtualization? How does it benefit me? What is the VMware Infrastructure? These questions are answered in this short topics booklet in relation to the VMware technologies used in the enterprise.

Virtualization applies to many different areas in the computer world, including graphics, sound, and computing. In this booklet we focus on the server virtualization space. Server virtualization allows multiple operating systems to run at the same time on the same hardware. It does this by logically partitioning the hardware and presenting a standardized set of resources and devices to the running operating systems. VMware has extended server virtualization to include management capabilities for server provisioning and management with tools such as VMware VMotion, which can migrate a running virtual machine to different hardware platforms without rebooting or changing device drivers. Regardless of the underlying hardware, an operating system can be confident that migration between different ESX host hardware will not impact the system and its applications.

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.

The VMware Infrastructure consists of ESX and ESX host hardware, which allows the server virtualization platform to run multiple operating systems; the VMware vCenter Server, which offers monitoring, provisioning, and management of the environment; and a set of tools to provide additional functionality. These tools include DRS for distributed resource scheduling, VMware HA for distributed availability services, VMotion for migrating virtual machines without downtime, and a number of other tools described later. We also provide use cases for virtualization, performance, and optimization, and the return on investment (ROI) for deploying the VMware Infrastructure.

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, migration and conversion tools, assessment tools, and management tools to support the VMware Infrastructure. This suite has the following technologies and associated products:

Virtualization Management Software

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

- VMware Virtual SMP—Enables multiprocessor virtual machines.
- VMware Distributed Resource Scheduler (DRS)—Dynamically allocates and balances resources across multiple virtual machines.
- VMware High Availability (HA)—Provides automated recovery of any applications running in a virtual machine, regardless of the underlying operating system or hardware configuration.
- VMware VMotion—Enables live migration of virtual machines from one physical server to another with no impact on end-users.
- VMware Storage VMotion—Enables live migration of virtual machine disk files across storage locations while maintaining service availability.
- VMware DPM—Distributed Power Management for the VMware Infrastructure, allowing dynamic startup and shutdown of ESX host hardware.

Native Virtualization Software

VMware ESX—ESX is the core enabling technology of the VMware Infrastructure. ESX version 3 (ESX 3) runs directly on physical hardware, not on top of an operating system, and is designed for maximum performance and availability. ESX version 3i (ESX 3i) can be embedded on a motherboard in a 32MB footprint.

Hosted Virtualization Software

Hosted virtualization software is virtualization software that runs on top of a standard operating system, including:

- VMware Workstation—Desktop virtualization product designed for end-users and developers to create and run virtual machines on their own Windows- or Linux-based desktop computers.
- VMware Player—Free virtualization product for running (but not creating) multiple virtual machines on Windows or Linux desktops.
- VMware Server—Free server virtualization product for running (but not creating) multiple virtual machines on existing physical Windows or Linux servers.
- VMware Fusion—Virtualization product for Intel-based Mac OS X systems.
- VMware ACE—Virtualization product for enterprise desktop deployments providing a highly configurable, secure, portable PC environment with centralized administrative control of software versioning and updates, and all facets of security and admission control.

Migration Tools

VMware Converter—Used for physical-to-virtual machine migrations, as well as importing virtual machines based on other virtualization vendors. VMware Converter can

import multiple machines concurrently and non-disruptively while the servers are running, or offline using a converter boot disk.

Security Enablers

- VMware ACE—See preceding page.
- VMware VMsafe—Provides an open approach to security through an application program interface (API) sharing program. 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.

Desktop Virtualization Software

VMware Virtual Desktop Infrastructure (VDI)—A system for managing connectivity, security, and administration of centralized virtual desktop computers hosted on ESX clusters. VMware Desktop Manager (VDM) is the VMware technology supporting connection brokering for VDI.

Application Virtualization Software

VMware Application Virtualization (formerly Thinstall)—An application virtualization platform that enables complex software to be delivered as self-contained 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 which is embedded with each "thinstalled" application.

Virtualization Assessment

VMware Capacity Planner—Identifies server inventories and resource utilization to determine virtual machine candidates, server consolidation ratios, and resource requirements for migrating to a VMware Infrastructure.

Software Lifecycle Automation

- VMware Lab Manager—Provides a self-service portal for real-time provisioning, managing, and collaboration of virtualized development and testing environments. Lab Manager allows developers and testers to create and share a library of virtualized application environments used in software development and testing.
- VMware Stage Manager—Automates the management of service transition and release management of preproduction resources. Production application systems can be captured for testing or updating, then promoted or demoted in and out of production through a predefined release management workflow.

Workflow Management

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

Disaster Recovery

- VMware Site Recovery Manager (SRM)—Provides disaster recovery automation and workflow management for a VMware Infrastructure. SRM automates setup, testing, failover, and failback of virtual infrastructures between primary and disaster recovery sites, as well as simplifying and centralizing the management of disaster recovery plans.
- VMware High Availability (HA)—Provides fault tolerance in the event of an ESX host failure. VMware HA allows the definition of rules for the automated restart of virtual machines on other hosts in a cluster upon host failure, providing minimal downtime during hardware failure without the cost of OS-level clustering.
- VMware Consolidated Backup (VCB)—Provides the capability to support SAN-based backup and recovery of virtual machines using a backup proxy server without any network or virtual machine overhead.

Benchmarking

- VMmark—A benchmark tool specifically designed for measuring scalability of virtualization host systems.

This collection of software, used together, can provide a dynamic environment that can reduce costs and provide significant improvement to the life-work balance of IT professionals.

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 or resources, as discussed in this booklet, 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.

VMware achieves this by inserting, directly on the computer hardware or on a host operating system, a thin layer of software that presents virtual machines or computers containing CPU, memory, hard disks, network interfaces, and other peripherals, which then appear as standard hardware devices to operating systems and applications. The same virtual devices are presented in the virtual machines regardless of the underlying physical hardware. This allows operating systems to install in virtual machines unaware that the resources are virtual and without any knowledge of the actual physical hardware.

In addition to virtualizing resources, VMware software acts as a broker, allocating physical resources to multiple machines. This enables multiple virtual machines to share one computer's resources, thus allowing the execution of many computer workloads simultaneously on one physical server known as the host server. This frees you from 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 VMware Infrastructure.

Although virtualization has become increasingly popular only recently, it is a concept that has been familiar in the computer industry for many years.

Virtualization generally refers to the separation of resources (such as networking and storage). It also separates storage, as well as a virtual machine's requests for service, from the underlying physical resources. An example is the use of virtual memory, where more memory is presented to a computer program than is physically installed on the computer system. In this case, the memory resources are virtualized from the computer program using them. In a similar way, other resources may be virtualized, such as networking through the use of VLANs (virtual local area networks) or VPNs (virtual private networks), storage through the use of storage virtualization, computer hardware (virtual machines), or applications (application virtualization).

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

Virtual Machines

The term "virtual machine" has many meanings, depending on which systems layer is virtualized. The two most prominent types of virtual machines are system virtual machines and process virtual machines.

System virtual machines map the underlying physical computer resources into one or more different virtual machines, which are 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 booklet focuses on system virtual machines. They are a representation of a real machine, based on a software implementation which provides 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, there is a layer of software called a virtual machine monitor (VMM). The VMM is a layer of indirection between the hardware and the virtual machine, allowing it to provide many distinct advantages over physical hardware. Indirection is a mapping of a virtual to a physical resource, much the way telephone call forwarding enables you to receive a call on one telephone number and have it ring 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 it—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. In turn, each Windows or Linux virtual machine acts as though it is installed on a physical computer, because a virtual machine behaves exactly like a physical computer.

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.

Characteristics of a Virtual Machine

We described earlier how hardware virtualization enables you to consolidate operating systems onto fewer hardware platforms. 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 makes them compatible with standard x86 computers so that standard *unmodified* operating systems, device drivers, and applications can run on the virtual machine. Operating systems and device drivers do not know they are running in a virtualized environment.

Isolation

Virtual machines are completely isolated from each other as if they were separate physical machines, even though they reside on a single physical computer. An operating system crash in a virtual machine cannot affect the operation of other virtual machines or the host server. Nor can a program running in one virtual machine peek into the memory of another virtual machine. This is a secure and reliable way to consolidate multiple applications on one physical machine, as opposed to stacking applications within one operating system to achieve server consolidation.

Encapsulation

A virtual machine is essentially a software container that bundles or encapsulates a complete set of virtual hardware resources, the enclosed operating system, its applications, and related settings. In other words, a computer becomes a set of files. Most of the virtual machine is a large file that represents the virtual machine's disk. A virtual x86 computer with a 60GB disk 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 computer, or the virtual disk can reside on a network storage system such as Storage Area Network (SAN), iSCSI, or network attached storage (NAS).

Hardware Independence

Virtual machine operating systems are completely independent from their underlying physical hardware, due to the virtualization abstraction layer. Instead, they are tied to the standard virtual machine platform. This is similar to the way applications are independent from server hardware because a standard operating system sits between the application and the physical server. For example, you can configure a virtual machine with virtual components (CPU, network card, SCSI controller, etc.) that are completely different from the physical components that are present on the underlying hardware. 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.

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 (nonvolatile 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 from 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.

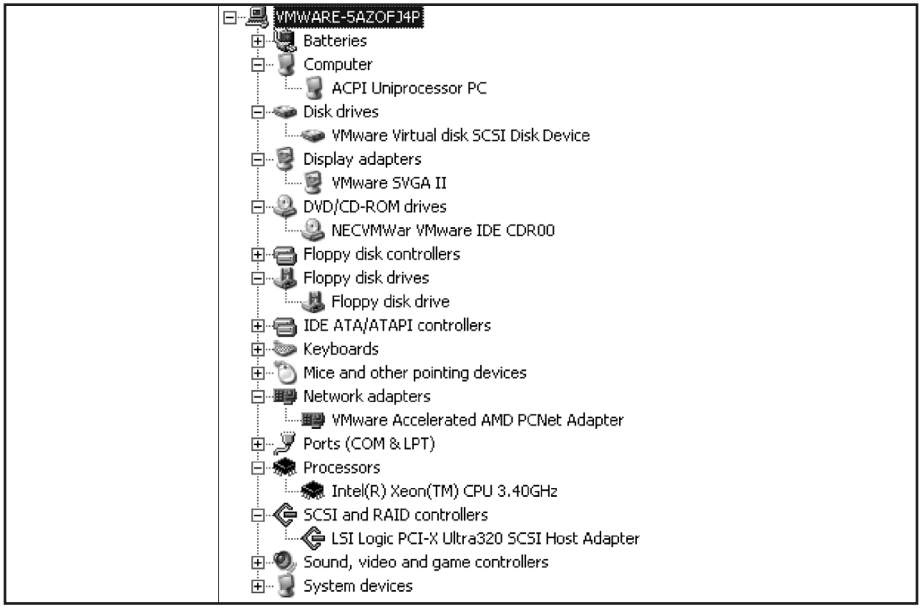
The virtual hardware platform presented to a virtual machine is standardized regardless of the underlying physical hardware. This simplifies operations and increases stability. The following are the components found within a VMware virtual machine (as implemented in VMware Infrastructure 3):

- An Intel 440BX-based virtual motherboard
- Virtual Phoenix BIOS 4.0 Release 6
- NS338 SIO chip
- Up to four virtual CPUs (vCPUs)—same processor type as host
- Up to 16GB of RAM (64GB for virtual machines under ESX 3.5)
- 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 four network adapters
- VMware Accelerated AMD PCNet Adapter
- Up to four SCSI controllers with up to 15 devices each
- BusLogic SCSI Host Adapter (BT-358)
- LSILogic SCSI Host Adapter (53c1030)
- Hot adds of hard drives for guest operating systems supporting this as a feature

Note: Up to six total PCI adapters are supported (SCSI+Network+Video combined).

The VMware virtual machine *remote console* provides access to the console of the virtual machine, providing access to the 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.

The figure on the facing page shows the devices as seen from the Windows Computer Management view of devices.



Windows Computer Management View of Devices

Components of the VMware Infrastructure

The VMware Infrastructure consists of multiple components that, when combined, allow the virtualization of an entire IT infrastructure architecture, from storage through the computing and networking layers.

The components of the VMware 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.

The VMware Infrastructure consists of one or more ESX hosts, the 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, VMotion for migrating virtual machines without downtime, and a number of other tools.

VMware ESX

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

With ESX 3i, this virtualization platform has a thin 32MB footprint and can be directly managed by VMware vCenter. It provides ESX 3.x functionality without requiring a service console for management. All control is managed through VMware vCenter or a Web interface for this particular version.

ESX offers a bare-metal architecture (requiring no underlying operating system on the physical server) that includes the following characteristics:

- CPU virtualization, providing time-sharing between multiple virtual machines and direct pass-through execution of guest OS commands on the processors.
- Storage virtualization, supporting SAN/iSCSI/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, and VLAN support.

VMware vCenter

The VMware vCenter Server provides centralized management in the VMware Infrastructure. It is used for configuring, provisioning, and managing the 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 one of the no-cost VMware-hosted virtualization platforms that run on supported versions of Microsoft Windows and Linux operating systems. It is designed to assist in software development efforts, to run virtual machines that require access to devices that are not supported by ESX, and to provide a starting point for those wanting to try virtualization through the use of virtual appliances.

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

VMware Virtual SMP

Virtual SMP enables allocation of multiple virtual CPUs (vCPUs) to a virtual machine. Two or four vCPUs can be allocated to virtual machines in VMware Infrastructure 3. Although multiple vCPUs can be allocated, it is not always best practice to do so. Before assigning multiple vCPUs to a virtual machine, a virtualization assessment should be performed to determine whether the virtual machine operating system and associated applications would benefit from virtual SMP.

Multithreaded 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 a VMware Infrastructure, 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 locked. Therefore, 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.

VMware VMotion

VMotion offers a technology to migrate running virtual machines between different ESX hosts. VMotion migration requires the same processor family on both the source and target ESX hosts, a GigE network connection configured for VMotion between the hosts, and that hosts be grouped within the same VMware Infrastructure cluster. For example, virtual machines running on AMD-based hosts can be migrated to other AMD-based hosts as long as the AMD CPU families on source and destination are the same. This is to ensure that the instruction sets are the same. Moving a running operating system from one CPU to another that uses a different instruction set will result in an operating system crash. Before starting the VMotion process, VMware vCenter Server validates that the target host is in the same processor family. 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 High Availability

VMware High Availability (HA) is a clustering technology that provides virtual machine high availability on VMware Infrastructure. If one host server fails, all of the virtual machines that are registered on that host and 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 same shared VMFS volumes. VMFS volumes are the file systems that store the virtual machine configuration files and virtual disks.

VMware Distributed Resource Scheduling

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 (QoS) for groups of hosted virtual machines. When the load on one node in the cluster is out of balance with others, DRS uses VMotion to rebalance the server load, minimizing resource contention and guaranteeing resource levels for hosted 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.

VMware Consolidated Backup

VMware Consolidated Backup provides a mechanism to support LAN- and server-free virtual machine backups by using a proxy server with direct access to the shared storage. Consolidated Backup uses virtual machine snapshot technology, allowing the Consolidated Backup 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. 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.

Additional VMware and Third-Party Components

VMware Infrastructure 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 the VMware Infrastructure with existing systems and processes.

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. However, rather than a physical server, the operating system and application are packaged using the industry-standard Open Virtual Machine Format (OVF). 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 operating systems on 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 underutilized 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 VMware 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.

Another advantage of virtual appliances in a virtualized infrastructure is that an organization can collapse more of the network support services into the VMware 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 of the physical hosts using virtual switches. There is less latency within the VMware Infrastructure than routing traffic from the VMware Infrastructure out to the physical network and back. These factors make custom-built virtual software appliances a natural extension of a VMware Infrastructure. As more of the infrastructure is virtualized, the portability of the entire infrastructure is increased, enabling simpler disaster recovery planning that 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.

3. The Benefits of Infrastructure Virtualization

Infrastructure virtualization provides significant reductions in capital and operational expenses, and improved agility in the datacenter.

Capital Expense Reduction

- Significant reduction in server hardware expenses.
- Reduced electricity consumption and costs for datacenter cooling.
- Reduced network and storage network switch ports, network and storage cards, and cabling expenses.
- Increased server and storage utilization (the average MS Windows server utilization is less than 20% without virtualization).
- Potential reduction in OS or application licensing costs.
- Reduction in hardware requirements for high availability and disaster recovery.

Operational Expense Reduction

- Reduction in procurement costs concomitant to reduced server purchases.
- Reduction in resources required for physical server deployments.
- Reduction in resources required for replacement and upgrading of physical servers.
- Increased service levels from reduced deployment time and cost, and increased responsiveness.
- Greater alignment of IT with the business through increased responsiveness to business requirements.
- Savings from reduced downtime for hardware maintenance or upgrades and servicing (VMotion, Maintenance Mode, and Update Manager).
- Reduced network, storage, and disaster recovery management costs.

Improved Agility

- The ability to remove legacy hardware while remaining fully operational.
- Greater service delivery speed through provisioning applications via virtual appliances.
- Greater stability and supportability using standardized virtual machines and appliances that are identical across VMware virtualization platforms.
- Support of merger and acquisition activity by providing increased infrastructure flexibility and compatibility.

- Increased ability for disaster recovery due to the encapsulation and portability provided by VMware virtualization.
- Increased ability for disaster recovery through advanced features such as VMware HA, VMware Consolidated Backup, and VMware Site Recovery Manager (more cost-effective savings).
- Better support for testing applications and operating systems (via VMware snapshots).

Summary of Benefits

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

- Server Consolidation—Reduce the number of physical servers.
- Server Containment—Prevent server sprawl.
- Legacy Applications—Allow older software applications and operating systems to run on the latest hardware platforms.
- Simplified Disaster Recovery—Encapsulation of 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—Virtualization abstracts the underlying hardware from the virtual machine, reducing risk and easing underlying hardware changes.
- Production Stability—Use of standardized hardware, encapsulation, and defined roles and responsibilities, combined with a robust management interface with resource trending information, allows for simplified capacity planning and quicker response to the needs of the business.
- Desktop Management—Use of VDI can 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 you to consolidate operating systems onto fewer hardware platforms. There are several additional benefits 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 Infrastructure 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. You can back up a virtual machine by copying the set of files. Of course, you can also run a backup program 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. This means that 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 VMware 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 gives you the freedom to move a virtual machine from one type of x86 computer to another 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 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 15 to 20 or more virtual machines per two-CPU quad-core servers. This means that an organization virtualizing its production environment may be able to consolidate 75–80% or more of its physical servers, along with their respective 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 a major savings for organizations running out of datacenter space or organizations that use outsourced datacenters that charge by the square foot or rack.

Storage

Virtual storage becomes a reality in a VMware 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. 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 additional flexibility and functionality to the virtual and physical network infrastructure.

Management

Managing a VMware Infrastructure is much easier than managing a physical environment, because all VMware Infrastructure components can be administered from a single point of view (pane of glass). While reports vary widely about the level of resulting savings, organizations commonly report administrators managing approximately 3–10 times more virtual machines than in physical environments.

Software Licensing

Savings frequently come from operating system licensing policies in a virtual environment. For example, purchasing Microsoft Windows Server 2003 Data Center Edition enables an organization to run unlimited instances of Windows Server on a single ESX host. Similarly, an unlimited number of instances of SQL Server 2005 Enterprise Edition can be run by licensing for the number of physical CPUs on the ESX host. You can run eight instances of SQL Server Enterprise Edition on a two-CPU multi-core ESX host, but only need to purchase licenses for two CPUs.

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 the VMware Infrastructure. Disaster recovery testing can also be done much less expensively and without requiring personnel to go on-site at the recovery site facility.

Desktops

Virtualizing the desktop with Virtual Desktop Infrastructure (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 terminals or lower-cost PCs. Maintenance costs are lowered and administrative requirements are reduced.

For example, an organization with 1,000 PCs and a three-year refresh cycle spends $1,000 per PC, including administrative costs, taxes, shipping, and set-up. An additional $150 per PC per year is spent for maintenance and troubleshooting expenses. Using VDI, the organization can either replace the PCs with inexpensive Windows terminals

(with no moving parts, local data or maintenance contracts) or lock down PCs to act like Windows terminals (if the PC breaks, it is simply replaced with a Windows terminal), enabling it to operate these devices an average of six years.

Using traditional desktops over a five-year period, the organization will spend $2.12 million: (1,000 PCs x $1,000 purchase price)/3-year life x 5 years = $1.67M + 1,000 x $450 maintenance = $2.12M.

During the same five-year period, a virtual desktop infrastructure would cost the organization $833 thousand: (1,000 Windows terminals x $1,000 purchase price)/6-year life x 5 years = $833K.

In other words, this organization could save $1.3 million over five years using a virtual desktop infrastructure based on hardware and maintenance savings alone.

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. They realized a cost savings of $1.6 million over a five-year period versus an initial investment of $270K, resulting in a ROI 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 facing page.

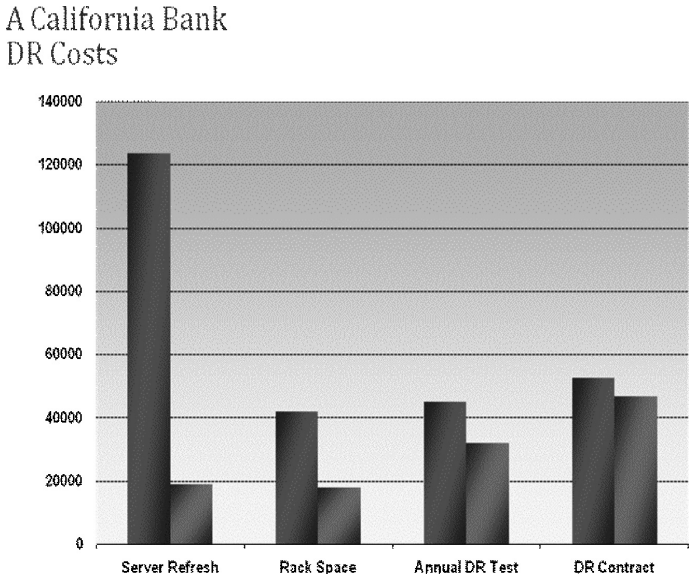
ROI/TCO Calculator

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 the VMware Infrastructure, VMware Lab Manager, and Virtual Desktop Infrastructure (VDI).

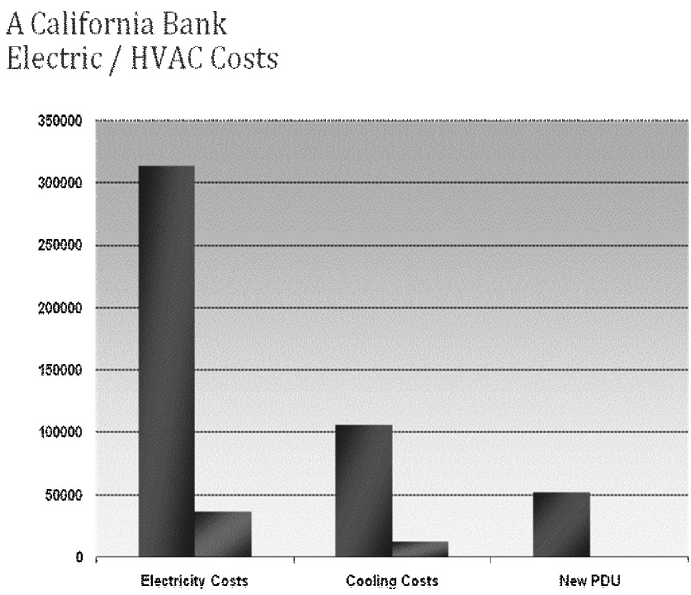
The tool focuses on comparing the costs of an existing physical computing environment against the same environment after being converted into a VMware Infrastructure environment. The tool asks five to ten questions about the existing environment to determine the cost to deploy and maintain the infrastructure. The questions are about your 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, as well as other areas.

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 Infrastructure. The cost savings are then quantified for management overhead, productivity improvements, risk reduction, and other technical and business benefits. The



Physical Versus Virtual DR Costs Comparison



Physical Versus Virtual Electric Costs Comparison

ROI/TCO Calculator produces final outputs including the cost investment, savings, and additional benefits to set up deploy, and support the new virtualized infrastructure.

Results of the analysis can be viewed comprehensively or individually for VMware Infrastructure 3, Lab Manager, and VDI deployments. The calculator results can be used to analyze different scenarios by adjusting variables and saving multiple result sets for comparison. You can review the results online or 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

Here are some of the major use cases for virtualization. More are discovered every day.

Production Environments

The advent of multi-core processors combined with ESX 3.5 pushed VMware solidly into the production mainstream. Today all 100 of the world's largest companies run VMware in some production capacity. As the section on ROI shows, VMware significantly reduces the cost of IT capital and operations, and also reduces risk by providing high availability for all servers, not just a minority that utilize expensive and difficult-to manage clustering software. The integrated development, testing, and disaster recovery capabilities further reduce IT infrastructure risks.

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 almost impossible to test before they are needed and they tend not to work well, if they work at all. Conversely, VMware Infrastructure 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 within minutes. 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 Foundation. VMware Foundation provides a lower-cost version of ESX designed for smaller deployments. This makes virtualization in small offices affordable while providing the encapsulation and mobility features that simplify backup and disaster recovery from remote offices.

Desktops

VMware Virtual Desktop Infrastructure (VDI) extends virtualization to the client. A user's desktop is hosted on centralized ESX hosts instead of distributing them to PCs. This gives employees far more flexibility by being able to work from any PC, Microsoft Windows terminal, or even kiosks. 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 multi-user access, VDI truly provides users with their own customized desktops while reducing the cost of hardware and simplifying administration.

5. Virtualizing Your IT Infrastructure

As you virtualize your IT infrastructure, you need to address the planning, implementation, management, and technical details of the VMware Infrastructure. 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 Infrastructure 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 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 the VMware Infrastructure during the third phase.

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 is 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 Infrastructure resource requirements, VMware Virtualization Assessments consider many factors, including utilization of the original physical machine and its hardware, operating system, and applications.

Good virtualization candidates can maintain performance levels equivalent to those on a physical machine in a virtual machine. 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 four 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. Next you need to look at the operating systems and applications to determine the hardware requirements and resource utilization. During the final phase of the assessment, the resource utilization of the virtualization candidates is analyzed to determine the aggregate CPU, RAM, 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. The VMware Capacity Planner is used during a VMware Virtualization Assessment to gather an inventory of hardware and software and metrics pertaining to resource utilization and performance. The typical data acquisition schedule is 30 days to identify trending, and the trending is then compared with industry numbers to identify virtualization considerations.

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 help in qualifying a system for virtualization include:

- Type, speed, number of CPUs, and number of cores per CPU.
- RAM available and peak usage.
- Disk space available and 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 recorded for analysis. 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. Attempting this without much experience or the assistance of a tool like the Capacity Planner is a difficult 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.

VMware uses a tool called Capacity Planner to track resource utilization and plan a virtualized environment. This tool is used to conduct a Virtualization Assessment. Other tools can be used but may not provide the detailed analysis necessary to provide a solid migration guideline.

To manually conduct this analysis, a programmer could use statistics gathered using the System Activity Reporter (SAR) tool for UNIX or Linux operating systems, or Perfmon for Microsoft Windows operating systems, but these tools cannot translate the data to useful virtualization scenarios. Using an automated tool with an understanding of virtualization technologies like the Capacity Planner can provide a more accurate model from the data gathered.

CPU

ESX 3.5 supports a maximum of 64 cores on its host platform, and one, two, or four virtual CPUs in a virtual machine. A virtual machine cannot have more virtual CPUs than the host has CPU cores.

Average and peak CPU utilization over an analysis window are not the only items for which to collect data. 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, influence the target virtualization architecture. 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 that virtual machines will be using. It is a best practice to default virtual machines to a single vCPU. After monitoring performance and utilization, additional vCPUs can be added if necessary. Multithreaded applications and applications written specifically to require more than one CPU often benefit from multiple vCPUs.

An estimate of three vCPUs per core can be used for capacity planning. Actual consolidation ratios vary based on application types, load, and number of SMP virtual machines on an ESX host and resource availability in an ESX cluster. When selecting CPUs for host systems, opt for larger cache size over CPU speed. Large cache sizes (both level 2 and level 3 caches) usually yield higher performance gains than minor CPU speed increases.

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. This is because the remaining instructions cannot be executed directly.

RAM

VMware ESX 3.x supports a maximum of 256GB of RAM on its host platform. Each guest OS can be allocated a maximum of 64GB 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.

Disk I/O

VMware ESX 3.5 can support up to 64TB of storage per host. Each guest OS can be allocated a maximum of 2TB of storage.

The analysis of disk I/O is a critical part of resource planning for a server consolidation project. Disk access can often 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. Disk I/O is the most expensive resource to virtualize, 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 7500RPM FC or SATA disks. A typical LUN for high-performance applications such as databases would be built using RAID 10 with 15,000RPM 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.

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.

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) for Fibre Channel SAN has been added to allow each virtual machine to have its own World Wide Name (WWN). Additionally, experimental support for round-robin HBA load balancing is included.

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, and 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 ESX 3.5 can support Fast Ethernet (100 Mbps), GigE (1000 Mbps), and 10G (10,000 Mbps, with support for TCP Segmentation Offload (TSO) and jumbo frames support.

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 in ESX 3.5 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 that allows previously CPU-intensive TCP segmentation operations to be managed in silicon on the physical network adapter. Jumbo Ethernet frames (frames with a size greater than 1500 bytes) are also supported if the external networking infrastructure is compatible, greatly reducing the number of network operations 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 or not TSO or jumbo frames can be used to reduce this overhead.

6. Building a VMware Infrastructure

The 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 VMware Infrastructure, but the underlying hardware architecture and configuration have considerable effect on the performance, scalability, and reliability of the VMware Infrastructure clusters.

The hardware components in a VMware Infrastructure cluster consist of two main components: the host server hardware, and the storage network and its subsystems. There are additional hardware components that play an important part, but they are usually external to the VMware Infrastructure. These external components include the physical data network and the backup infrastructure.

When selecting ESX or storage hardware, the most important requirement for system stability and VMware support is that the hardware is compliant with the VMware Hardware Compatibility List (HCL). The tight coupling between the hypervisor and hardware depends on using tested, compatible hardware. The VMware HCL can be found on the VMware Web site.

Server Hardware

In a VMware 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 Infrastructure project depend largely on hardware selection.

When selecting ESX host hardware, several features 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 multi-core 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 servers with quad-processor SMP. To run 64-bit guest operating systems, 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 busses. 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.
- I/O performance and scalability considerations—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 eight- or sixteen-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/vi3_301_201_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 one GB per second 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).

Storage Hardware

Storage networks and shared storage arrays are critical to a well-architected VMware Infrastructure. Storage design, configuration, and performance are the most important and complex hardware subsystem considerations in VMware Infrastructures. Virtual machines are stored in raw LUNs or VMFS file systems and are executed on any of the servers in a VMware Infrastructure cluster. Networked storage is necessary for all of the advanced VMware Infrastructure management features, including VMotion, HA, DRS, and Consolidated Backup. Storage-related issues are the most common and problematic VMware Infrastructure 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 VMware Infrastructure. There are a wide variety of storage arrays tested and approved for compatibility with VMware Infrastructure 3, 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 Infrastructure, but there are some storage system characteristics more optimized for VMware Infrastructures.

VMware vCenter Server provides interfaces for dynamically managing all aspects of a VMware Infrastructure, including hosts, clusters, guest operating systems, and virtual networks. Storage and VMware Infrastructures are often managed by different teams, with storage systems generally being less agile and dynamic than the rest of the VMware Infrastructure. Storage as an extension of a VMware 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 Infrastructure. The ideal situation is to have the storage system managed as a part of the VMware 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 Infrastructure storage systems is the need for linear performance per capacity scaling.

VMware 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 Infrastructure should be able to scale performance linearly with virtual server growth to the upper limits of its capacity growth.

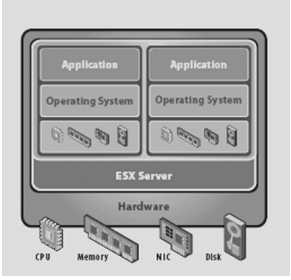
Managing quality of service is essential when consolidating several different workloads with various service level needs on a VMware Infrastructure cluster. VMware vCenter Server supplies robust quality of service features to maintain desired service levels across

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 Infrastructure management systems.

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 Access to Resources

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 *Systems Compatibility Guide for ESX Server 3.0.x* contains the most current list of systems certified to run ESX (http://www.vmware.com/pdf/vi3_systems_guide.pdf).

VMware recommends a minimum 1.5GHz (1500MHz) CPU clock speed to run ESX.

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:
   - Local CD–ROM or DVD-ROM drive.
   - A remote CD/DVD drive using the remote system software provided by the manufacturer of the hardware system.
   - Automatic kickstart-based scripted installation (see the section Kickstart Scripted Installation on page 36).
   - 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. 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. Select the appropriate mouse and keyboard.
4. Accept the VMware license agreement.

   ESX is installed on the first drive that is seen by the system, 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.
5. When the partitioning page is displayed, click the Recommended partitioning option to configure default partitions based on the capacity of the drive, and use the Edit option to change the automatic partitioning settings. One such disk-partitioning scheme for ESX is illustrated in the following table.

/dev/sda (Primary Partition)			
Mount Point	***Partition Type***	***Size***	***Description***
/boot	ext3	250 MB	Holds the VMkernel binaries
/	ext3	5120 MB	The root partition for the service console
N/A	swap	800 MB	2x service console memory (assumes 400 MB physical memory assigned to the service console)
/dev/sda (Extended Partition)			
/var/log	ext3	5120 MB	Separate to avoid overfilling root with log and core files
N/A	vmkcore	100 MB	Pre-configured
N/A	VMFS	Remaining	For local VMFS volume

ESX Server 3.x Partition Table

Some system administrators choose to have separate partitions for /var as well as for /tmp. The reasoning behind this is that if these partitions fill up, the rest of the system is unaffected.

6. Specify on the next screen that the ESX is to boot from the drive.
7. Provide boot options for the installer:
   - ❖ 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.
   - ❖ Force LBA32—This option is used to exceed the 1024 cylinder limit for the /boot partition. This option is only needed for legacy hardware.
8. Configure the network settings on the next screen. Configure the ESX host IP address by entering 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.
9. Select the time zone.
10. Choose a secure root password and enter it twice in the fields provided.
11. 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.
12. 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.

VIRTUAL APPLIANCE FOR KICKSTART DEPLOYMENTS

Taking the kickstart installation a step further is the Ultimate Deployment Appliance (UDA), an open source tool for creating kickstart scripts and associated pre- and post-install scripts. This appliance is available from the VMware Virtual Appliance area of the VMware Web site.

VMware vCenter Installation

A VMware vCenter Server configured with the hardware minimums can support 20 concurrent clients, 50 ESX hosts, and over 1000 virtual machines. A dual-processor

VMware vCenter Server with 3GB RAM can scale to 50 concurrent client connections, 100 ESX hosts, and over 2000 virtual machines.

VMware vCenter Server version 2.5 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: 2GB RAM minimum. RAM requirements can be larger if your database is run on the same hardware.
- Disk storage: 560MB minimum, 2GB recommended. You must have 245MB free on the destination drive for installation of the program, and you must have 315MB free on the drive containing your %temp% directory.
- Database storage requirements: The demonstration database, using MSDE, 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. MSDE is not supported for production use.
- Networking: 10/100 Ethernet adapter minimum (Gigabit recommended).

Software Requirements

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

- Windows 2000 Server SP4 with Update Rollup 1
- Windows XP Pro SP2
- Windows Server 2003 (all releases except 64-bit), Windows Server 2003 R2

VMware vCenter Server 2.0 installation is not supported on 64-bit operating systems. The VMware vCenter installer requires Internet Explorer 5.5 or higher.

Database Configuration

VMware vCenter Server supports the following database formats:

- Microsoft SQL Server 2000 (SP 4 only) and Microsoft SQL Server 2005
- Oracle 9iR2, 10gR1 (versions 10.1.0.3 and higher only), and 10gR2
- Microsoft MSDE (not supported for production environments)

Each database requires some configuration adjustments in addition to the basic installation. Please see the VMware Infrastructure 3 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 can be installed on the same physical or virtual machine that hosts the VMware vCenter Server. When installing VMware vCenter, the wizard asks if you want to install the license server.

The license server manages the license pool for virtualization products managed by the VMware vCenter Server.

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 and the license server (to serve VMware Infrastructure 3 feature licenses) on the same physical or virtual machine as the VMware vCenter Server.

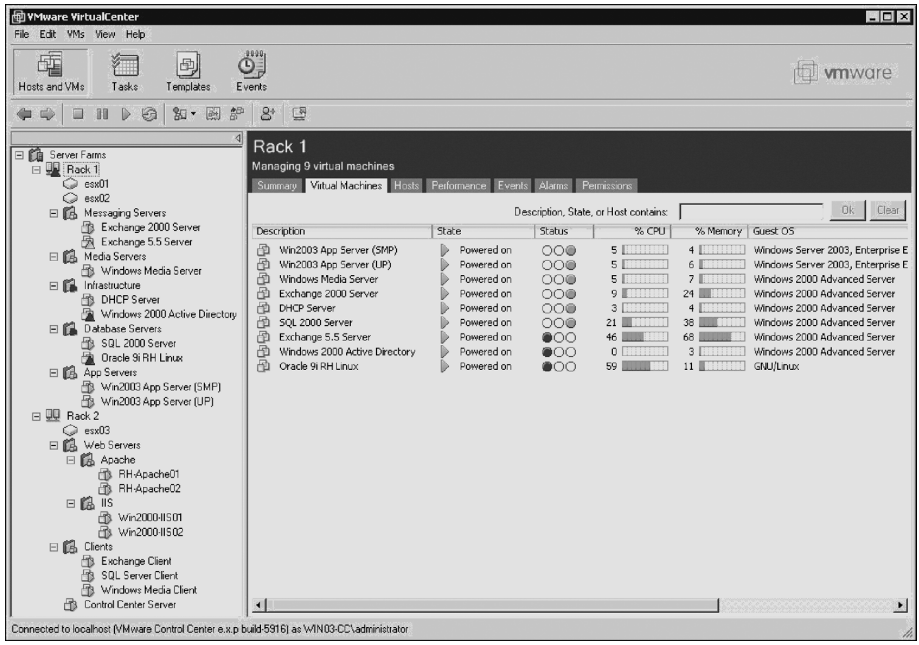
The user installing VMware vCenter Server requires administrative privileges on the machine. For detailed installation instructions, see the VMware Infrastructure 3 installation guide. After the VMware vCenter Server is installed, the features specific to VMware vCenter Server can only be enabled by using the license server. Unlike ESX, which can be enabled by a host license file, the VMware vCenter Server features can only be used with licenses served by the license server.

7. Managing the VMware Infrastructure

VMware vCenter Server

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

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



VMware vCenter Server Cluster View for Rack 1 Cluster

ESX provides a robust virtualization layer, but it is the VMware vCenter Server that provides the features required for a complete infrastructure management system. VMware vCenter Server enables virtual machine mobility, centralized monitoring and alerting, advanced quality of service, and high availability functionality that 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 Infrastructure 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 having the same family of CPUs on each host so that CPU instruction sets a re identical on both hosts. More information on this and other VMotion requirements are available on the VMware Web site.

VMotion migration or cold migration can be accomplished from VMware vCenter by dragging the virtual machine and dropping it on the destination host. This action opens a wizard that guides you 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 opens the migration wizard to guide you 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 you to create and mark a virtual machine as a template. Do this when installation of the required applications is complete. To create a template, right-click on the virtual machine in the VMware vCenter Server and select Convert to Template. A template is essentially a virtual machine that cannot be powered on. You can create an unlimited number of virtual machines 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.* You can choose to customize the guest operating system (computer name, domain, and IP address) when provisioning from a template. This is accomplished 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. It significantly reduces the time required to create a new server. Administrators can create different templates for different purposes. For example, you can create a Windows 2003 Server template for the finance department, a Red Hat Linux template for the engineering department, and a Windows 2003 Server template for the marketing department. This enables the administrator to quickly provision a correctly configured virtual server on demand.

This ease and flexibility bring with them the problem of virtual machine sprawl, where virtual machines are provisioned so rapidly that documenting and managing the virtual machine lifecycle becomes a challenge. Many VMware Infrastructure 3 customers are implementing change management processes in the virtual environment. It is also helpful to have virtual machine naming conventions that help identify each machine's purpose and users.

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, and network and disk parameters displayed on the Performance tab of the VI 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). Thus, any performance monitoring should be done through ESX. Viewing the performance at a VMware vCenter Server level enables you to look at 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 Infrastructure Client.

Additionally, alarms can be set to trigger when certain metrics exceed a set threshold. For example, you can define an alarm that is triggered when CPU utilization on a host reaches 90%. 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.

You can also define the action VMware vCenter Server should take in case an alarm is triggered. The available actions for an ESX host are to send a notification email message, send a notification trap, or run a script. If you are creating an alarm at a virtual machine level, you can also choose to suspend, power off, or reset the virtual machine.

Using the alarm 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 2.x comes with the ability to create a cluster. A cluster is a logical group of servers that can be managed together and can work in tandem for load balancing or disaster recovery.

As you add servers 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. 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. This means that the VMware vCenter Server could be run within a virtual machine and the VMware HA technology would still function and restart the VMware vCenter Server along with the other virtual machines.

A VMware vCenter Server cluster can be configured for both DRS and VMware HA and recover from unplanned host isolation and host failures.

Managing Resource Pools

VMware vCenter Server enables you to carve out resources for a set of virtual machines. For example, you can choose to assign 25% of the CPU resources and 30% of the memory resources in a cluster to a set of staging virtual machines. This reservation is called a *resource pool*. You can create an additional resource pool on the same cluster with 70% of the CPU resources and 60% of the memory resources for production virtual machines.

This configuration ensures that the staging virtual machines never use more than 25% of the CPU resources on the cluster. Therefore, 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 resource use by the virtual machines created for a particular department to only those resources allocated to that department.

The combined effect of resource pools, DRS, 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" on page 26. This is an add-on module for VMware vCenter Server 2.5 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 the VMware Infrastructure.

VMware Update Manager

VMware Update Manager is part of the VMware Infrastructure 3.5 product suite. It supports a streamlined, automated update mechanism for ESX hosts and virtual machines. See the "VMware Update Manager" section on page 51.

VMware Converter

VMware Converter is described in the section "VMware Converter" on page 53. This is an add-on module for VMware vCenter Server 2.5.

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 is considered experimental. 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 VMware Infrastructure.

Infrastructure Management with VMware vCenter

Effective management tools and processes are essential components of a VMware Infrastructure. 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 Infrastructure 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 VMware 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 Infrastructure management tasks.

Datacenters and Clusters

The Hosts and Clusters view is available from the Inventory view in Virtual Infrastructure Client. This interface is where the physical and logical organization of the VMware Infrastructure occurs. Datacenters, folders, and clusters are containers for organizing VMware Infrastructure 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 Infrastructure 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 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 independent of individual hosts as host resources are aggregated within the cluster.

Roles and Permissions

Centralized security and permissions for the VMware Infrastructure are stored and managed in the Virtual Infrastructure Client through the Permissions tab. User accounts and permissions for local host access can be assigned to the service console or the Virtual Infrastructure 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
- Datacenter administrator
- Virtual machine administrator

Roles can be created, edited, removed, renamed, or cloned using the VMware Virtual Infrastructure Client through the Roles tab on the Admin view by right-clicking in the roles panel and choosing Add Role.

The combination of VMware Infrastructure roles and the nested hierarchies of VMware Infrastructure 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.

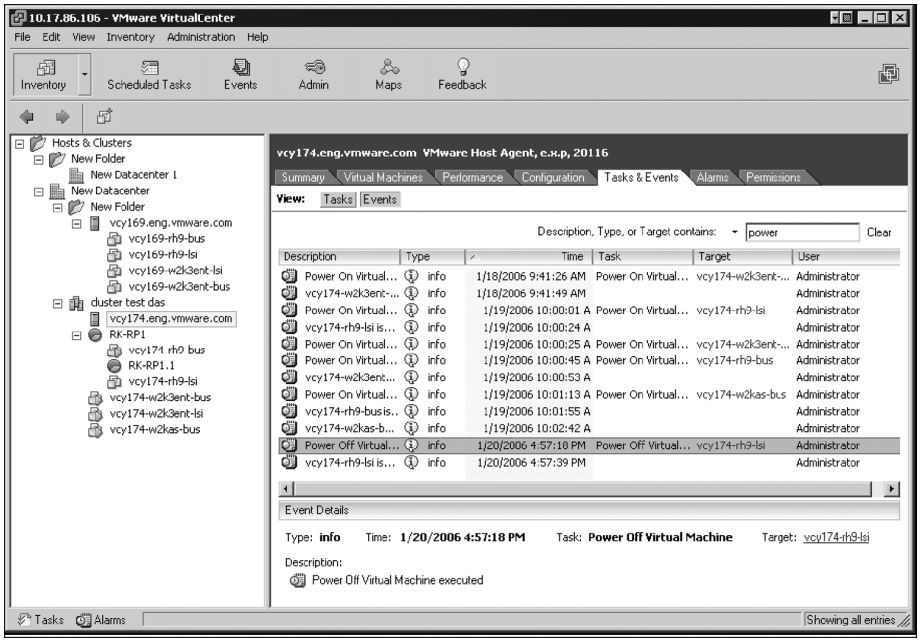
The most effective way to manage permissions through VMware vCenter is to create resource pools for groups of virtual machines for management and resource sharing, and Active Directory groups for each resource pool. By assigning each Active Directory group with permissions to the appropriate resource pools, VMware Infrastructure 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 Infrastructure Client. The Events tab on the navigation bar displays all of the events in the VMware 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.

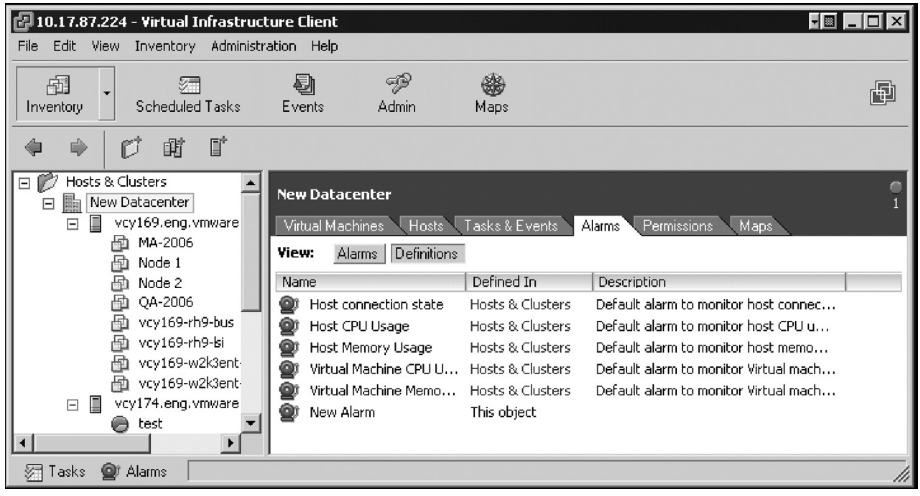
The VMware Infrastructure Client enables greater visibility and event management by allowing sorting of event columns or searching for events. These features are necessary when reporting on events in large, consolidated environments. Events are searched for using the event-filtering feature in the VMware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



Tasks and 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 VMware Infrastructure. Default Alarms are included in VMware vCenter, but alarms are not directly configurable through the VMware Infrastructure 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 them to objects or containers in the VMware Infrastructure hierarchy.



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 Infrastructure Client connected to a VMware vCenter Server, Inventory: 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.

The following virtual machine and host condition states are available for alarm triggers.

Virtual Machine Conditions

Creating	Migrating	Connecting	Disconnecting	VMotion
Reconnecting	Removing	Resetting	Resuming	Starting
Stopping	Suspending	Disconnected	Initial	Orphaned
Powered Off	Powered On	Suspended		

Available Host Conditions

Connecting	Disconnecting	Reconnecting	Removing	Shutting Down
Connected	Disconnected			

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 script.
- Suspend a virtual machine.
- Power off a virtual machine.
- Reset a virtual machine.

Using VMware vCenter Server alarms and notifications provides a rich set of tools for notifying and reacting to alarm events within the infrastructure, and it allows seamless integration of a VMware 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 Infrastructure administrators. The VMware Infrastructure 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 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 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 Infrastructure 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 Infrastructure Client provides simple centralized monitoring of the VMware Infrastructure capacity and the ability to non-disruptively add and distribute additional capacity throughout the infrastructure. Organizations with large VMware 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 Infrastructure Client. The Create New Virtual Machine wizard collects the following information for virtual server creation:

- Select 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.
- Select a location for the virtual machine in a folder, resource pool, or datacenter.
- Select a datastore for the virtual machine files.
- Identify the guest operating system to be installed on the virtual machine.
- Select the number of virtual CPUs presented to the virtual machine.
- Determine the virtual machine memory size. The wizard presents a minimum, maximum, and recommended memory size for this virtual machine based on available memory and the guest OS selected.
- Select the number of NICs, which virtual networks to connect to, and the connectivity of the networks.

- Set the virtual disk size. A virtual disk can be from 1MB to 2TB in size. Additional virtual disks can be added later. The size of existing virtual disks can be changed through the VMware Infrastructure 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.

If you are creating a virtual machine using the custom option, you have several more detailed options: type of SCSI adapter, LSI Logic, or BusLogic. Select the type of virtual disk: new, existing, or a mapped LUN from a storage network. When creating a new disk there are options for disabling write caching, allocating all disk space now, and splitting the disk into 2GB files. The virtual disk device node and 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:

1. Select a target LUN visible from the SAN.
2. 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 save the choices you made to create a custom specification for reuse when creating virtual machines in the future.

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 must 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 Infrastructure 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 and Norton Ghost 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.

Hot migrations of virtual machines can occur only when certain conditions are met. VMware vCenter Server validates the virtual machine and its target destination host and resource pool to ensure a successful migration. VMotion migrations can only occur when the starting and destination host have compatible CPUs, 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, thus 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.

Storage VMotion

Storage VMotion is based on the hot migration technology of VMotion, but adds the changing of the storage location holding the virtual disk of a virtual machine. 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 SAN shared storage. It performs proactive, non-disruptive storage migrations to simplify array storage maintenance without virtual machine downtime. This 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 disk snapshots and REDO logs. 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 moved to the new storage location.
2. The virtual machine is migrated, using VMotion, to the target ESX host.
3. The virtual disks are moved to the new storage location.
   - A child disk is created on the target storage device for each virtual disk to hold all disk writes, just like a write cache.
   - The parent virtual disk is copied from the old storage device to the new storage device.
   - The child disk pointer is changed from its attachment with the original parent disk to the new parent disk.
   - The write operations applied to the child disk are consolidated to the new parent disk.

The entire process takes less than two seconds in most circumstances.

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.

VMware Update Manager

VMware Update Manager provides an automated patch management system for VMware Infrastructure. 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 you to schedule updates. It also handles all of the virtual machine migration, patching, and rebooting, eliminating typical manual steps and downtime during updates in a physical environment.

8. Migrating Candidates

Conducting a virtualization assessment helps make your virtualization projects a success. After candidates are identified, migrations can begin, and the time needed for migration 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 machine to a virtual machine. VMware uses P2V migration software (VMware Converter) 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, and parallel, serial, 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.

VMware Converter

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 9.

VMware Converter also supports restoring VMware Consolidated Backup images to running virtual machines.

VMware Converter Starter is free and allows single conversions. VMware Converter Enterprise allows automation and management of large-scale conversions. The differences between the starter and enterprise versions of VMware Converter include:

- Both versions support hot cloning, local conversions, and remote conversions.
- The enterprise version allows for cold cloning with a special bootable CD, multiple simultaneous conversions, and remote conversion to all destinations.
- The starter version limits remote conversion to VMware Workstation, VMware Server, VMware Player, and VMware GSX Server.

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, a 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. For Microsoft guest operating systems, Microsoft Sysprep 1.1 is used to handle the conversion process. For non-Microsoft operating systems, other options are possible. For Linux environments, the Kudzu imaging tool works very well.

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

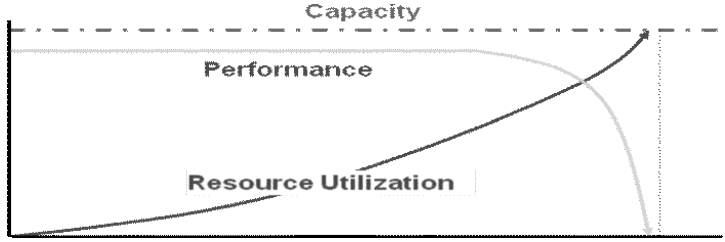
Virtual-to-Virtual Process

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.

9. Optimization

In terms of resources, *utilization* is the amount of a given resource that is being used. It is often represented as a percentage such as %CPU or the I/O bandwidth of a system. Performance is the number of transactions that can execute in a given amount of time, such as per-second Web page views, loan applications, or I/O operations. Performance does not equal utilization, as either high or low resource utilization can yield the same performance. Increases in resource utilization do impact performance the closer you are to the capacity of the resource, as shown in the following graph. This is especially important to understand in relation to virtualization, because the sharing of resources on an ESX host pushes the resource utilization higher when consolidation ratios increase.



Performance Impact as Resource Utilization Reaches Capacity

Performance can be impacted by various resource constraints including hardware, configurations, and application limits, as well as systemwide resources. Given resource constraints and competition for those resources, performance monitoring and tuning strive to reach an optimal or acceptable level. Do not run in a virtual machine applications that require precise timing such as performance monitoring tools, lab measurement applications, or instrument controllers. An example is an application that polls data every 10 milliseconds—the data may not be checked every 10 milliseconds, due to the variations introduced to the vCPU through time slicing of the real CPU.

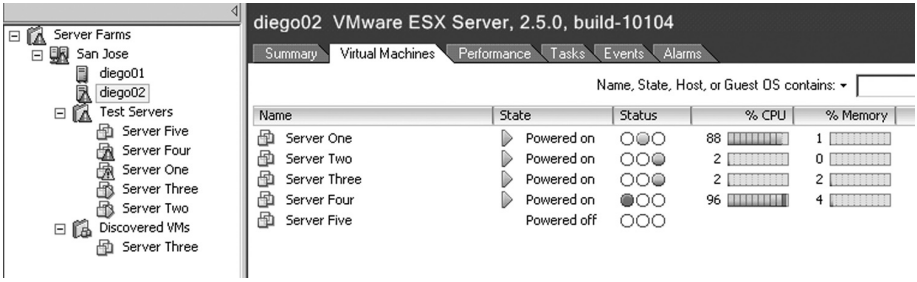
ESX Optimization

It is important to understand that an individual virtual machine can impact other virtual machines based on 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 they can be useful for understanding application issues. VMware Infrastructure tools such as VMware vCenter Server and esxtop represent actual performance. The following figure is a screenshot of a VMware vCenter client showing monitoring of CPU and memory utilization for a set of virtual machines.



Monitoring Virtual Machine Resource Utilization and Status

VMware vCenter Server is used to view resource utilization. The following table lists the default resource views that can be displayed. Custom performance charts can also be created.

- CPU
  - Total CPU utilization
  - Total CPU utilization by virtual machine
- Memory
  - Total memory utilization
  - Total memory utilization by virtual machine
- Disk and Network I/O
  - Total disk and network I/O
  - Total disk and network I/O per device
  - Total disk and network I/O by virtual machine
  - Individual disk and network I/O

The host CPU chart has one usage line. The virtual machine CPU chart has two lines. The first line shows usage as the percentage of CPU time used by one or all virtual machines on the ESX host. The second line shows the percentage of CPU time guaranteed to be available to the virtual machines.

The memory chart has two lines. The first line shows the active memory, including the memory recently used by one or all virtual machines on the host. The second line is the memory allocated to the virtual machines(s).

The disk charts report the amount of read and write activity occurring to and from the disk. The data is defined by shares of the disk resources per volume. Two aggregated

charts show combined read and write activity. The virtual machine disk chart has three lines: the first shows the total I/O; the second, the disk reads; and the third, the disk writes.

The network chart reports the number of bits per second of input and output to network interface cards (NICs). The virtual machine aggregate chart shows all input and output activity per virtual machine. The virtual machine network chart has three lines: the first shows the total I/O; the second, the network traffic received; and the third, the network traffic transmitted.

The charts are broken down to show aggregate charts for daily, weekly, monthly, and yearly views.

Resource Pools

Resource pools are a new feature introduced with VMware Infrastructure 3 and VMware vCenter 2. Resource Pools are hierarchical logical containers created within VMware Infrastructure 3 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.

In a VMware Infrastructure 3 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 Infrastructure provides significant flexibility in managing resources and distributed administration of virtual resources. VMware Infrastructure 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 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 contains 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 the distribution of resources. It enables the delegation of administration to localized administration, which is necessary for management of 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. 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.

Placing application systems or functionally similar servers in resource pools simplifies resource management by avoiding the need to manage resource allocation to each server individually. 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. If the pool has insufficient resources and the reservation type is set to Fixed, the virtual machine cannot be powered on because of the pool's reservation and limit. 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.

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 Infrastructure 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 the VMware Infrastructure.

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. Two modes are available for this automation level: conservative and aggressive.

Virtual Machine Optimization

VMware best practices for single and multithreaded application CPU allocation recommend starting with the lowest number of vCPUs when deploying new, or when migrating from physical machines to virtual machines. Multithreaded applications can benefit from multiple vCPUs. However, a side effect of this can be a reduction in overall computing resources, as all vCPUs must be scheduled to perform an execution at the same time even if only one is performing the computations.

Application and Operating System Tuning

The resources for one ESX host are shared with all the virtual machines that are running on that host, compared to a physical server, which has full access to all resources. Virtualization maximizes resource utilization better than physical machines. Streamlining your applications and operating systems provides a better TCO and ROI in your virtual deployment. A typical deployment of a physical machine takes future capacity requirements into consideration differently from virtual machines. An application-OS pair needing one gigabyte of memory implies that a physical machine with more than one gigabyte is required. In many cases, this means installing two or four gigabytes of RAM.

Tuning an application and operating system for a virtual environment has many benefits, some of which are similar to tuning a virtual machine:

- Minimizes wasted resources (CPU, RAM, disk).
- Generates warning and error logs when application and operating system issues arise.
- Minimizes the use of snapshots.
- Reduces unneeded layered applications.
- Schedules antivirus software and guest OS-based backups to minimize contention.
- Uses VMware Consolidated Backup instead of guest OS-based backups where appropriate.
- Tunes code that accesses your databases to minimize contention by doing more operations with fewer connections.
- Minimizes the use of screen savers. Splits data and log writes to different virtual disks and/or VMFS volumes to reduce I/O contention and provide better DR capabilities.
- Disables unused services.

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. As active virtual machines change their load, with some being added and some removed, traditional benchmarks do not factor in all of these variables.

A virtual environment consists of multiple server instances all running different workloads on top of the same hardware platform.

The use of traditional benchmarking tools for a virtual environment presents the following challenges:

- 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 provides the ability to measure these diverse workloads and generates an output that scales as resource capability and resource requirements change.

VMmark calls the work unit measuring these workloads a *tile*. This 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.

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.

10. Disaster Recovery and Security

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 in the event of a disaster. 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 consist of Web, application, and database servers used to host an eCommerce front-end, where another consistency group might consist 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.

Backup and Recovery Strategies

Several data-protection strategies are available, including the use of guest OS–based software, VMware Consolidated Backup, VMware snapshots, storage snapshots, and storage replication. A new technology from VMware, released in 2008, is Site Recovery Manager, a business continuity and disaster recovery 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.

VMware Consolidated Backup

VMware Consolidated Backup is a backup enabler, not a backup product. Consolidated Backup allows for a SAN-based or iSCSI-based solution. As such, it allows backups to be offloaded from the data network. It utilizes a proxy server that is connected to the shared storage. To be truly LAN-free, the backup media server must be installed on the proxy server also.

The use of a proxy server minimizes the resource load on the guest OS and requires almost no load on the ESX host during backups. Based on the restoration method, recovery may require some resources on the network and on the ESX host.

Snapshots

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 in time that 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>-<SnapshotX.vmsn>.
- A file storing the state of the virtual machine memory when the snapshot was taken, with the filename <VM name>-<Snapshot.X.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 VMware 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 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 multitiered replication strategy provides the most cost-effective VMware 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 requirements (or as little data loss as possible) 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 Infrastructure enables extremely short RTOs, because they can be restored and 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 Infrastructure disaster recovery plan.

STORAGE REPLICATION STRATEGIES

Several replication strategies for VMware 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.

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 reestablish user access along with server recovery. Architecting a datacenter failover that is transparent to application users requires that the primary site network scheme must 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

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. Advanced Capabilities

The following sections cover capabilities that are part of the VMware Infrastructure 3 Enterprise release, including:

- VMware HA for high availability of ESX hosts.
- VMware Consolidated Backup for LAN-free and ESX host and guest OS off-loaded backups via a shared storage infrastructure.
- Virtual Machine Snapshots for reversion in the event of a guest OS failure.
- Site Recovery Manager for end-to-end disaster recovery workflow management and automation.

VMware High Availability

VMware HA is a high-availability solution designed to handle ESX host failure and enable the virtual machines running on a failed system to be migrated to an alternate ESX host in a VMware HA cluster and, optionally, powered on.

When implementing virtual machines that are part of a traditional cluster (such as MSCS, OpenMosix, or Veritas Cluster Services), it is important that the virtual machines in that cluster are excluded from VMware HA. If this is not done, an ESX failure results in both the clustering software and the VMware HA software trying to do a failover. The clustering software fails the virtual machine over to another cluster node and the VMware HA server restarts that virtual machine on another ESX host, resulting in two operational virtual machines with the same characteristics (IP address, host name, and so on).

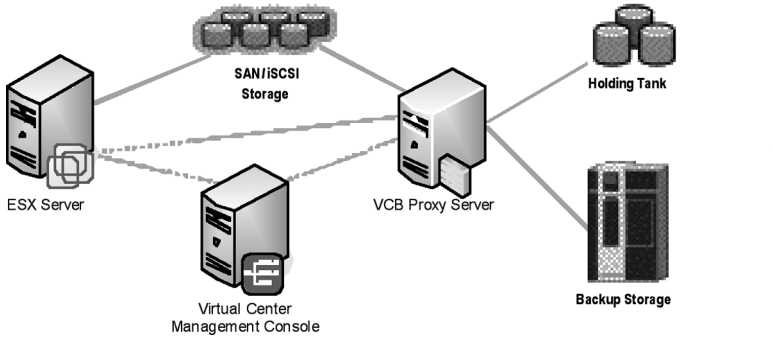
VMware Consolidated Backup

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 (filesystem-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

log files for a virtual machine. Full virtual machine backups are supported for all guest OS types. The guest OS system state is not included when using Consolidated Backup–enabled backups.



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 log files 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.

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 a 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 similar 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.

Site Recovery Manager

VMware Site Recovery Manager 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.

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 3 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 method to define a workflow and complete testing in a very cost-effective manner. The output of the testing process is a report that can be used to support regulatory compliance auditor requirements.

12. Virtual Desktop Infrastructure

The benefits of a virtual computing infrastructure are well documented. Greater physical and human resource efficiency is enabled through consolidation and standardization of resources. Encapsulation of virtual systems provides mobility which enables the replication and recovery of logical servers in ways unheard of under a physical computing paradigm. Inherent in VMware Infrastructure are high availability, distributed resource scheduling, and consolidated backup services, which add additional value to all virtualized systems.

Organizations that embrace server virtualization and have implemented comprehensive datacenter VMware Infrastructures 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 because a replicated, recovered virtual server infrastructure without access to the desktop computers is only a partial solution, which cannot provide efficient business continuity. Early adopters of VMware Infrastructure looked for a way to extend the scope and benefits of their VMware Infrastructure to include the client-side computing resources to provide a consistent and complete virtualized infrastructure.

VMware Virtual Desktop Infrastructure (VDI) assists in incorporating desktop computing into the virtualized infrastructure. VDI collapses desktop computing, storage, and processing into the VMware Infrastructure, extending its benefits throughout the enterprise.

VDI Overview and Planning

VDI is not a product; it is a specific use case for VMware Infrastructure built on ESX and enabled by VMware or third-party products to complete the solution. VDI maintains security, control, and performance by keeping the desktop systems, data, and execution in the datacenter and close to application servers.

A VDI solution is built by accessing ESX hosted virtual desktop images using a remote display protocol through a virtual desktop connection broker. Additional components can be used to manage printing, client hardware, application distribution, and patching services. Combining these components into a comprehensive VDI solution enables the server, desktop, and storage computing infrastructures to be deployed, managed, replicated, and recovered as a unified whole.

VDI requires a connection broker, ESX, and VMware vCenter Server for a basic implementation. An organization's existing VMware Infrastructure can host a virtualized desk-

top OS and allow user access through a remote display protocol or the VMware vCenter Client. The main considerations for implementing VDI in a VMware Infrastructure are capacity and performance planning and management.

The workload created by VDI is very different from a virtual server workload. A VDI implementation can create hundreds to thousands of virtual machines, each with a smaller resource footprint than a virtual server. A VMware Performance Study white paper, *VDI Server Sizing and Scaling* (http://www.vmware.com/pdf/vdi_sizing_vi3.pdf) reported that a dual-processor, dual-core HP DL385 with 8GB RAM and local SCSI storage could host 42 *light worker* Windows XP virtual machines or 26 *heavy worker* virtual machines while maintaining a consistent guest OS performance profile. Newer-generation servers with quad-core processors dramatically increase the virtual desktop-to-CPU ratio in the host servers. High numbers of low-utilization virtual desktops benefit more from the simultaneous processing provided by increased core counts than from higher processor speeds. Virtual desktops have much lower memory requirements, and hosting many similar virtual machines usually creates greater memory-sharing ratios than with server workloads.

Special considerations also need to be taken into account when planning SAN 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.

Desktop numbers, configurations, and usage vary considerably between and within organizations. Accurate capacity planning for a VDI implementation, such as a large-scale server consolidation project, requires usage monitoring and analysis as well as a pilot implementation to determine typical desktop usage profiles.

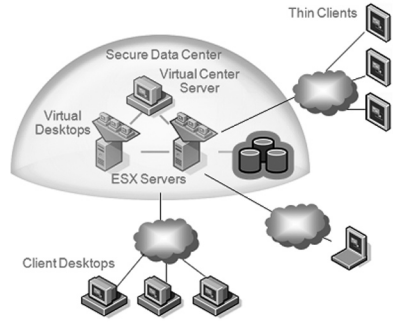
Connection Brokering

A Virtual Desktop Infrastructure is built upon ESX and VMware vCenter, but a critical component at the center of a VDI is the connection broker. VDI has several architectural patterns. All but the most basic include a connection-brokering component to manage dynamic connectivity between users and virtual desktops. A connection broker is a software component that acts as connection concentrator and negotiator between users and virtual desktops.

The following sections describe the most common VDI architectural patterns and vendor-specific implementations for VDI connection brokering.

Basic VDI Connectivity

For small VDI implementations where dynamic connections are not required, basic connectivity architecture can be used. In this model users are given the address of a specific virtual desktop host on the local network and they connect using a standard protocol such as RDP.



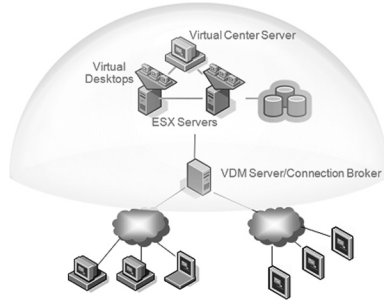
Simple VDI Implementation

The figure above shows a simple VDI implementation where client desktops (thick clients) and thin clients connect to their respective virtual desktops through a remote connection protocol. This setup requires that the users at the client devices know the IP address or hostname of the virtual desktop. The IT administrator cannot enforce that users connect only to the assigned desktops. Such a basic implementation is suitable for smaller infrastructures but becomes unmanageable as the number of clients increases or advanced features are implemented.

These limitations warrant a VDI connection broker implementation. A connection broker also helps to add rich features to your virtual desktop provisioning and administration.

Simple Brokering

With this simple brokering architecture, users authenticate with a connection broker through a local Web address. The broker provides the authenticated user with a list of virtual desktops authorized for the user's access. Users connect to the selected virtual desktop resource using a third-party remote display protocol or one native to the OS such as Microsoft's RDP. Several vendors offer this type of simple virtual desktop connection brokering, which works well for non-distributed organizations with LAN client connectivity.



VDI Connection Broker Implementation

Tunneled Brokering

A tunneled brokering architecture is recommended when users require secure access to desktops from outside the trusted corporate network. In this configuration the connection broker sits at the public-private network broker (usually behind the firewall) presenting a Web interface for user authentication. Authenticated users are presented with an approved list of virtual desktop resources for connectivity via RDP or another available protocol. The tunneling connection broker creates an encrypted VPN tunnel, passing only the approved remote protocol between the client and the virtual desktop.

A tunneled brokering architecture increases VDI flexibility by allowing secure access to roaming users from public or external private networks. Some form of tunneled brokering (implemented as a virtual server) is recommended for VMware Infrastructure disaster recovery implementations, allowing distributed client connectivity to the recovered VMware Infrastructure.

Vendor-Specific Implementations

Several vendor-specific architectures integrate one of the basic designs with their existing solution architectures. A number of vendors, including Citrix, Sun Microsystems, and VMware, offer VDI connection brokers that integrate into and extend existing solutions.

Connectivity Protocols

Connectivity protocols must be selected for a VDI implementation. Some connection brokers use a specific protocol, while others work with more than one protocol.

The most common remote display protocols used for VDI connectivity are Microsoft's Remote Desktop Protocol (RDP), Virtual Network Computing (VNC), Citrix Independent Computing Architecture (ICA), and Hewlett-Packard Remote Graphics Software (RGS), which offer the following features:

- Remote Desktop Protocol (RDP)
  - Standard Windows remote protocol included with Windows XP and Vista
  - Good performance
  - Supports remote audio
  - Multi-platform support
  - Multi-monitor support
  - Available as a browser plug-in
  - Required for VMware VMware View Manager connection broker
- Virtual Network Computing (VNC)
  - Wide platform support
  - Both free and commercial versions available
  - Variable performance
  - No audio support
- Independent Computing Architecture (ICA)
  - Standard Citrix protocol, very mature implementation
  - Excellent performance and feature set

- Bi-directional audio support
- Requires Citrix Presentation Server and per-user licensing
- Extensive platform support
- Browser plug-in available

- Remote Graphics Software (RGS)
  - Provides the best graphics performance
  - Least used protocol
  - Requires per-node licensing

The selection of a remote display protocol largely depends on the specific use case and the combination of desktop OS and connection broker used.

Printing and Peripheral Hardware Management

Management of peripheral hardware and client printing is problematic for server-based computing environments. VDI, like server-side application virtualization, supports several printing configurations.

Virtual desktop computers can be configured for printing, just like physical desktop machines, by installing a network printer near the client location with print drives installed locally in the virtual desktop. Print jobs are created in the virtual machine and sent across the network to a printer near the client machine. This configuration can work well for local corporate networks where security and bandwidth are not concerns.

Third-party universal print drivers can be used to simplify print client driver configurations. The universal print driver is installed in the virtual desktops, and a client component is installed on the client machine (not supported on most thin clients). The universal print driver on the virtual machine transmits the print job to the client component, which then sends the job to the requested printer.

Third-party brokered printing is used for printing from public or wide-area connections where security is a concern. With brokered printing, the virtual desktop uses a print driver to create a standard print file like a PDF. The PDF print jobs are staged for user pickup with the connection broker. The user can view and select their available print jobs on the connection broker, which is downloaded and printed from the client machine.

In addition to printing devices, clients may need locally attached USB peripheral hardware such as scanners, bar code readers, or other input devices. Third-party hardware and software products are available to enable USB connectivity over Ethernet. By placing a USB-over-Ethernet hub near the client location and installing the remote USB driver software in the virtual desktop, any USB device connected to the hub appears and functions as a locally installed USB device in the virtual desktop machine. The USB-over-Ethernet hub is a solution that is client-independent and works well with thin client devices. Software-only solutions are available for use with local PC clients.

Patching and Application Management

Desktop patching and application management can be handled in VDI identically to existing processes used for physical clients. However, additional options simplify applica-

tion and patch deployment on virtual desktops. Combining virtual desktops with client- or server-side application virtualization (application publishing or streaming) creates an extremely manageable and flexible desktop solution where replicated virtual desktop images are dynamically populated with the appropriate applications from a central location managed in the datacenter.

Implementation Planning

One important consideration in implementing VDI is network bandwidth. The network bandwidth consumption increases with the number of virtual desktop users. While it is desirable to have dedicated bandwidth or network resources for VDI, this is not practical, due to cost considerations. As a rule of thumb, plan for 25kbps per virtual desktop user. Remember that as the number of users grows, the per-user bandwidth requirement drops, because not everyone generates heavy loads at the same time.

Network latency is a critical success factor for any VDI implementation. Any latency over 150 milliseconds is noticeable in mouse and keyboard responses. To ensure good end-user experience, maintain network latency between the datacenter and client terminals below 150 milliseconds.

With virtual desktops, virtual disks are normally kept on shared storage such as a SAN. Use of shared storage helps deliver the benefits of DRS and VMware HA in the virtual desktop implementation. This also increases the cost of resources used by each virtual desktop, because shared storage is more expensive than local desktop storage. Most organizations are unable to provide upwards of 40GB storage per virtual desktop.

To overcome this limitation, consider provisioning the virtual desktops on a virtual system disk of minimal size (e.g., 10GB) on more expensive shared storage where only the operating system and applications are installed. This disk should be locked down and users should not have access to use this space. A second virtual disk, for user consumption, is provided on less expensive storage such as a NAS device.

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. To enable a smooth transition to VDI from physical desktops, a pilot VDI program is recommended.

A pilot program of 25 to 100 virtual desktop users seems adequate to test-drive the policy decisions and better understand user requirements. During the pilot, adjust and finalize group policies, conduct user satisfaction surveys to identify productivity issues, and finalize the virtual desktop templates (virtual machine image) that are to be used 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 VDI implementation can save your organization operational and management expenses. Most importantly, VDI can put your desktop assets in a position to benefit from all the current and future innovations that virtualization has to offer.

Appendix: 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 Syste*f*m 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 enables 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:

- Hosted: VMware Workstation, VMware Player, VMware ACE, VMware Server, and VMware Fusion.

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 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, 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: VMware ESX, VMware ESXi, part of VMware Infrastructure
  - Hypervisor virtualization 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 (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 which 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 (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 were 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 that 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:

- *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 Infrastructure, although these two technologies can coexist.

We provided 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, it's unlikely that these terms will come up in your conversations, 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. In software-level virtualization, 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, can over-commit memory using a swap file, just as a regular operating system has virtual 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 are 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 is 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.

About the Authors

John Y. Arrasjid has over 20 years of expertise in the computer science field. His experience includes work with AT&T, Amdahl, 3Dfx Interactive, Kubota Graphics, Roxio, and his own company, WebNexus Communications. John is currently a senior member of the VMware Professional Services Organization (PSO) as a Consulting Architect. He is the Worldwide Business Continuity and Disaster Recovery Practice Lead and is a developer of consulting engagements, including Performance, Security, Disaster Recovery and Backup, VMware Consolidated Backup, and Site Recovery Manager. He was the original developer of the vmsnap/vmres scripts that evolved into Consolidated Backup. John regularly presents at VMworld, USENIX, LISA, LinuxWorld, and other conferences. John is also a founding member of the two VMware music bands Elastic Sky and The Hypervisors.

Karthik Balachandran is a Senior Consultant specializing in helping IT organizations design, deploy, and manage their technology efficiently. Karthik is currently developing best practices for operational readiness using VMware Infrastructure. This enables companies to adopt and expand their virtual environment while adapting processes and existing technology to the environment. Karthik has worked in software development, release engineering, and enterprise consulting for Verizon and VMware, among other companies. In the Virtual Desktop Infrastructure (VDI) area, Karthik worked on product development and requirement definition for VMware's connection broker (VMware View Manager) technology. His expertise in VDI and ITIL methodologies has helped his clients implement scalable, low-cost solutions.

Daniel Conde is a Senior Technical Account Manager at VMware. His experience includes work at Digital Equipment Corporation, Microsoft, NetIQ, and Rendition Networks in software development, product management, business development, and professional services. He is a coauthor of USENIX conference papers "An Experimental Symmetric Multiprocessor Ultrix Kernel" and "Ultrix Threads." He has authored articles and books and spoken at conferences on a variety of information technology and computer science topics.

Gary Lamb is Senior Director, Data Center Virtualization Practice for INX, a VMware Premier Partner. Gary has been working with VMware since ESX 1.0. During his 23 years in IT, he has served as a MVS systems programmer, LAN/WAN design and integration architect, and an enterprise security assessment practitioner (2+ years with ISS). Gary also has an extensive background in disaster recovery. He designed the branch ATM backup network for one of the nation's largest banks and spent two years on a disaster recovery team for a Fortune 1000 Company. Gary can be contacted at Gary.Lamb@inxi.com.

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 data-

center and desktop virtualization at venues around the globe, including delivering the keynote at 2006 ThinPower in Norway. Steve holds a Bachelor of Science degree in business administration from U.C. Berkeley and an MBA from Northwestern University's J.L. Kellogg Graduate School of Management. He can be contacted at Steve.Kaplan@inxi.com.