State of California
Office of the State Chief Information Officer
Infrastructure Consolidation Program
Server, Virtualization, Backup and Storage Workgroup

TRM 4.17.865.002, Server Virtualization Practices

Revision History

Date	Revision	Author	Comments	Reviewers
2/26/2010	.1		Initial draft	
3/8/2010	.5		Draft	
3/12/2010	.7		Revisions based on feedback on 3/10/2010 meeting – P2V flowchart	
3/22/2010	1		Final Draft	
10/12/2010	2		Combined Practices Documents	

Review History

Date	Revision	Author	Comments	Reviewers

Contents

List of Tables 7
List of Figures 8
Virtualization Practices 9
Chapter 1: VMware Virtualization Practices 10
Purpose and Overview 11
VMware vSphere platform 11
vSphere Datacenter/Cluster 11
Sites/Locations 12
Naming conventions 12
vSphere Clusters 12
VMware HA 13
VMware HA Considerations 15
VMware Fault Tolerance 15
VMware Distributed Resource Scheduler 16
VMware ESX Host 17
VMware ESX Host Hardware Specifications 17
Host Device Placement 19
Local Storage 19
VMware vCenter Management Server 19
VMware vCenter Server System Specifications 19
vCenter Database 20
Licenses 20
vSphere Network Architecture 21
Network Architecture Logical Design 21
Network Redundancy Considerations 23
vSphere Shared Storage Design 23
Shared Storage Logical Design 24
Shared Storage Requirements 25
Datastore Configuration Specifications 25
Storage Path Redundancy Design 27

vSphere Infrastructure Security 27
vSphere Host Security 27
VMware ESX Service Console Security Specifications 28
Authentication 28
SUDO 28
vCenter and Virtual Machine Security 28
Security Considerations with multiple security zones 29
vSphere Network Port Requirements 30
Virtual Machines 30
"Virtual Machine First" 30
Virtual Machine Templates 31
vSphere Infrastructure Monitoring 32
Overview 32
vSphere Monitoring 33
Virtual Machine Monitoring 33
vSphere Infrastructure Patch/Version Management 34
Overview 34
vCenter Update Manager 34
vCenter Server and vSphere Client Updates 36
Backup/Restore Considerations 36
VMware ESX Server Host Backup 37
VMware ESX Server Host Recovery 37
Virtual Infrastructure Backup 37
Special vSphere Architecture Design Considerations 37
vSphere Architecture Redundancy 39
Assumptions 39
Hardware 39
External Dependencies 40
Reference Documents 41
Supplemental White Papers and Presentations 41
Supplemental VMware Knowledgebase Articles 42
Chapter 2: Hyper-V Virtualization Practices 43

Document Purpose 44
Microsoft Server Virtualization Overview 44
Workgroup Recommendations 44
Management Tools 46
In The Box Tools 46
System Center Virtual Machine Manager 46
Workgroup Recommendations 46
Monitoring 46
PRO and SCOM Integration 47
Workgroup Recommendations 47
Hyper-V Host Sizing and Configuration 47
Server Type 47
Commodity Server 47
Blade Server 47
Processor 48
Memory 48
Networking 48
Storage 48
Pass-through Disk 49
Physical Storage Option 50
Direct Attached Storage 50
Shared Storage 50
Workgroup Recommendations 51
Server Types 51
Processors 51
Memory 51
Network 52
Storage 52
High Availability 52
Workgroup Recommendations 53
Security 53
Protecting Host Servers 53

Server Core 53
Other Host Considerations 54
Protecting virtual machines 55
Workgroup Recommendations 56
Backup and Recovery of Hosts, Virtual Machines and Datasets 56
Traditional File-Level Backups 57
Block-Level Backups 57
Guest Sizing and Configuration 59
Basic Configuration 59
Guest Virtual Processors 60
Guest Memory 60
Networking 60
4 Networking Types 61
General Best Practices 62
Advanced Options 62
Exhibits 64
Performance Monitoring 64
Use Case 65
Virtualization and the California State Lands Commission (CSLC) 65
The bottom line 68
References 69
Appendix A - ESX Service Console Firewall Settings 70
Appendix B – Port Requirements 72
Appendix C – Monitoring Configuration 76
Appendix D – P2V or VM Suitability Flowchart 78
Acknowledgements 79

List of Tables

Table 1 VMware HA Cluster Configuration 14

Table 2 VMware DRS Cluster Configuration 16

Table 3 vSwitch Security Settings 22

Table 4 ESX Server Hosts 23

Table 5 Shared Storage Logical Design Specifications 24

Table 6 Storage Configuration Specifications 26

Table 7 Storage Path Redundancy 27

Table 8 vCenter Update Manager Specifications 35

Table 9 Noteworthy Items 37

Table 10 Potential Failure Points and Measures for Redundancy 39

Table 11 Sources of Technical Assumptions for this Design 39

Table 12 VMware Infrastructure External Dependencies 40

Table 13 VMware ESX Service Console Firewall Settings 70

Table 14 ESX/ESXi Port Requirements 72

Table 15 vCenter Server Port Requirements 73

Table 16 vCenter Converter Standalone Port Requirements 74

Table 17 vCenter Update Manager Port Requirements 75

Table 18 Physical to Virtual Windows Performance Monitor (Perfmon) Counters 76

Table 19 Modifications to Default Alarm Trigger Types 77

List of Figures

Figure 1 VMware ESX Clusters spanned multiple HP c7000 Chassis 13

Figure 2 VMware HA 13

Figure 3 VMware FT 16

Figure 4 VMware DRS load balancing 17

Figure 5 Network separation 21

Figure 6 VMware ESX Logical Network Design 22

Figure 7 ESX with SAN attached storage 24

Figure 8 Logical SAN Diagram 25

Figure 9 Virtualization with multiple security zones 30

Figure 10 Distributed Power Management 33

Figure 11 vSphere Update Manager 34

Figure 12 P2V Decision flowchart 78

Virtualization Practices

This document is intended to provide basic information and recommended practices for the VMware VMWare vSphere 4 environments and Microsoft server virtualization environments including Hyper-V and System Center Virtual Machine Manager. This document identifies and highlights recommended practices only.

It is not intended to be a substitute for thorough research and adequate training by State staff working with server virtualization products.

Chapter 1: VMware Virtualization Practices

Content Contributions by:

Sergio Guiterrez CalEPA	Doug Novak DFG	Brian Amos DHCS
Richard Harmonson CALEMA	Robert Stuart DFG	Casey Evans DMV
Tony Woo Energy	Dan Marksbury CalFire	Robert Dolliver Water
Vince Leong EDD	Richard Rogers EDD	Jerry Lee EDD
Ru Ma Food and Agriculture	Chris Dove DOF	Wesley Major DOF
Blake Rushworth Conservation	Kevin Hudgens BOE	Scott MacDonald CDCR

Purpose and Overview

This document is designed to give Agencies within the State of California some fundamental recommendations in designing a VMware vSphere 4 environment. This documentation is to be used as a guide when planning, designing, purchasing, and implementing VMware in State of California Agencies.

This document is targeting "enterprise" environments with 250 or more servers. The fundamental principles of these practices are:

- High Availability
- Consistency
- Flexibility for future growth

This document assumes that organizations will review and apply these practices in accordance to their own security and business continuity requirements.

VMware vSphere platform

Both VMware ESX and VMware ESXi are the current hypervisor platforms for running virtual machines. Both ESX and ESXi are available as the foundation of an organizations virtual infrastructure.

The major difference between VMware ESX and VMware ESXi is the inclusion of a Linux based "Service Console" in ESX versus a more appliance like configuration interface in ESXi. Traditionally the ESX service console has been used for initial configuration as well as a convenient environment for running third-party management or monitoring agents. Since the introduction of ESXi in 2007, VMware has been actively working to eliminate dependencies on the service console from both their own and supported third-party software.

The elimination of the service console simplifies and reduces the amount of software code, reducing complexity and providing more of the CPU, memory, network and storage resources to the virtual machines.

VMware has announced its intention to replace ESX with ESXi at some point in the future.

Organizations should design for ESXi, and even if they deploy ESX, they should manage their vSphere environment as if it was running on ESXi.

vSphere Datacenter/Cluster

In VMware vSphere implementation it is possible to group resources logically and hierarchically. In this logical hierarchy datacenters and clusters are defined. At minimum a single datacenter is required and typically that object encompasses all datacenter and cluster objects for an entire organization.

Sites/Locations

In VMware vSphere, a Datacenter is the highest-level logical boundary and is typically used to delineate separate physical sites/locations or vSphere infrastructures with completely independent purposes.

Naming conventions

Use defined, documented, and consistent naming conventions for all objects in a virtual datacenter. This provides order to the virtual infrastructure and helps administrators readily and correctly identify virtual infrastructure objects

vSphere Clusters

Within vSphere datacenters, ESX/ESXi hosts are organized into clusters. Clusters group similar hosts into a logical unit of virtual resources enabling such technologies as VMware VMotion, VMware High Availability (HA), and VMware Fault Tolerance (FT).

- The maximum number of ESX/ESXi hosts in a HA cluster is 32. But if the configuration exceeds 40 VMs per host, the maximum number of hosts in a HA cluster is only 8. The recommended number of ESX/ESXi hosts in a single cluster is between 4-16.
- Make sure that the size of the current environment plus future growth is within the specified limits in order to avoid performance and supportability issues.
- VMware ESX clusters should be located on physical hardware in a manner that takes into account the degradation of capabilities will preserving the viability of the virtual machines running on the cluster.
    - When using blade servers, make sure that sufficient ESX hosts will remain running if one or more components in the blade chassis fail. For example a fully populated HP c7000 chassis has 6 power supply/fan modules. As these modules fail, the remaining take up the load if possible, at the point where 4 modules have failed ½ of the device bays will lose power.

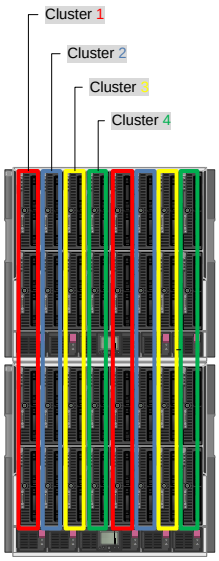


Figure 1 VMware ESX Clusters spanned multiple HP c7000 Chassis

VMware HA

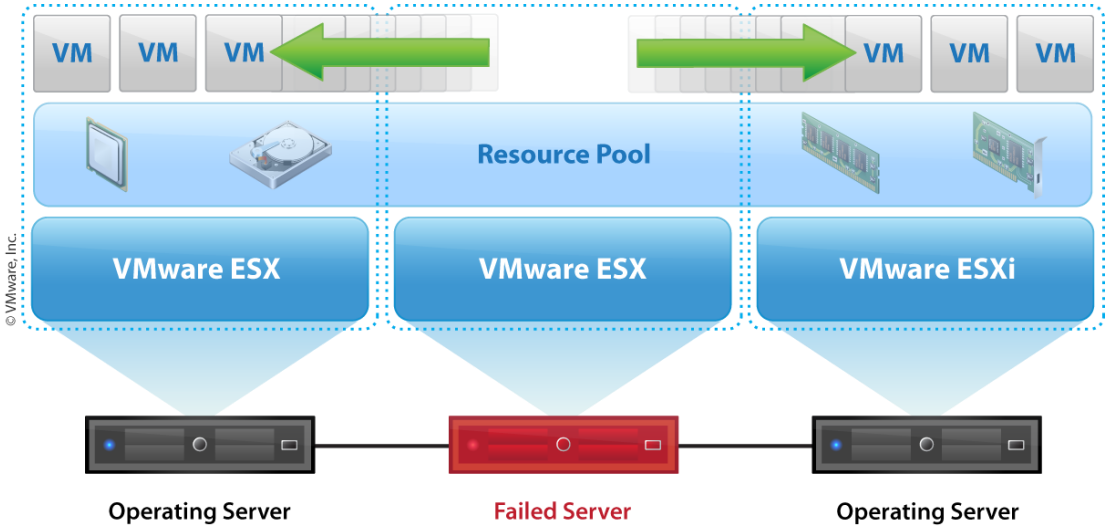


Figure 2 VMware HA

Each VMware ESX cluster should be configured with VMware High Availability (HA) to provide automatic recovery of VMs in the event of either an ESX Server host failure. A host is declared failed if the other hosts in the cluster cannot communicate with the host. If an ESX Server host in an HA enabled cluster fails, the VMs running on that server will go down, but will be restarted on another host within a few minutes. While there would be a service interruption perceptible to users in the event of an ESX Server host failure, the impact is minimized by the automatic restarting of these virtual machines on other hosts.

The configuration settings for VMware HA will be as follows:

Table 1 VMware HA Cluster Configuration

Attribute	Specification
Enable host monitoring	Enable
Admission control	Prevent VMs from being powered on if they violate availability constraints
Admission control policy	Cluster tolerates 1 host failure
Default VM restart priority	High (critical VMs) Medium (majority of VMs) Disabled (non-critical VMs)
Host isolation response	Power off VM

Setting Explanations

- **Enable host monitoring**. When HA is enabled, hosts in the cluster are monitored and in the event of a host failure, the virtual machines on a failed host are restarted on alternate running hosts in the cluster.
- **Admission control.** This enforces availability constraints and preserves host failover capacity. Any operation on a virtual machine that decreases the unreserved resources in the cluster and violates availability constraints is not permitted.
- **Admission control policy.** Each HA cluster can support as many host failures as specified.
- **Default VM restart priority.** The priority level specified here is relative. VMs will need to be assigned a relative restart priority level for HA. VMs will be organized into four categories: high, medium, low and disabled. It is presumed the majority of systems will be satisfied by the medium setting and therefore will be left at default. VMs identified as high priority, such as the Active Directory VMs, will be started before the medium priority VMs, which in turn will be restarted before the VMs configured with low priority. If insufficient cluster resources are available, it is conceivable that VMs configured with low priority will not be restarted. To help prevent this situation, non-critical systems such as QA and test VMs should be set to disabled. In the event of a host failure, these VMs will not be restarted, saving critical cluster resources for higher priority VMs.

- **Host isolation response.** Host isolation response determines what happens when a host in a VMware HA cluster loses its service console/management network connection but continues running. A host is deemed isolated when it stops receiving heartbeats from all other hosts in the cluster and it is unable to ping its isolation addresses. When this occurs, the host executes its isolation response to prevent multiple instances of each virtual machine from running if a host becomes isolated from the network (causing other hosts to believe it has failed and automatically restart the host's VMs), the VMs will automatically be powered off upon host isolation.

VMware HA Considerations

The configuration of VMware ESX host networking and name resolution, is critical to optimizing VMware HA operation.

DNS must be configured to resolve fully qualified domain names for the VMware ESX hosts. In order to configure VMware HA, a DNS server is required to resolve host names. However, once configured, VMware HA caches the name resolution and does not require DNS lookup in order to perform failover operations.

VMware Fault Tolerance

VMware Fault Tolerance (FT) can be enabled on VM that need to an even higher degree of resiliency. VMware FT provides continuous protection for a VM by creating a secondary copy of that VM on a different ESX host.

Fault Tolerance uses the VMware vLockstep technology on the ESX host to provide continuous availability. This is done by ensuring that the states of the Primary and Secondary VMs are identical at any point in the instruction execution of the virtual machine. vLockstep accomplishes this by having the Primary and Secondary VMs execute identical sequences of x86 instructions. The Primary VM captures all inputs and events -- from the processor to virtual I/O devices -- and replays them on the Secondary VM. The Secondary VM executes the same series of instructions as the Primary VM, while only a single virtual machine image (the Primary VM) is seen executing the workload.

If either the host running the Primary VM or the host running the Secondary VM fails, a transparent failover occurs whereby the host that is still functioning seamlessly becomes the host of the Primary VM. With transparent failover, there is no data loss and network connections are maintained. After a transparent failover occurs, a new Secondary VM is automatically respawned and redundancy is re-established. The entire process is transparent and fully automated and occurs even if vCenter Server is unavailable.

VMware FT does impose very strict configuration requirements and restricts the availability of some advanced features, so VMware FT will only be implemented for specific VM's where the uptime requirements warrant.

All VMs to be protected by VMware FT will have only one vCPU and virtual disks configured eager-zeroed, also called thick-provisioned (not thin-provisioned). An eager-zeroed thick disk has all space allocated and zeroed out at creation time; this takes a bit longer for the creation time, but facilitates optimal performance and better security.

FT traffic will be supported with a pair of Gigabit Ethernet ports (see vSphere Network Architecture section). Since a pair of Gigabit Ethernet ports can support on average 4 to 5 FT-protected VMs per host, there is capacity for additional VMs to be protected by VMware FT.

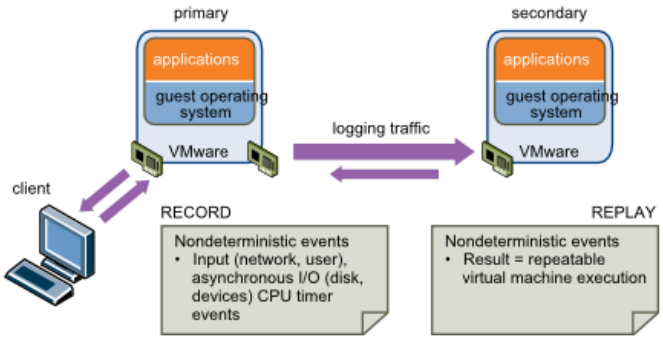


Figure 3 VMware FT

VMware Distributed Resource Scheduler

VMware Distributed Resource Scheduler (DRS) technology should be used to dynamically distribute virtual machines on the available ESX Server hosts to provide consistent utilization, and performance across the clusters. When a virtual machine is powered on, DRS will choose an ESX Server host with adequate resources to run that virtual machine. If an ESX Server host's utilization reaches a level that negatively affects the performance of running virtual machines, one or more virtual machines will be migrated to a more suitable host. The load balancing with VMware DRS leverages VMware VMotion to migrate the virtual machines without any service interruption to users.

Table 2 VMware DRS Cluster Configuration

Attribute	Specification
Automation Level	Fully Automated
Migration Threshold	Three Stars (default)

Setting Explanations

- **Automation Level**. The DRS automation level setting controls the initial placement and ongoing load balancing aspects of DRS. When DRS is fully automated, vCenter Server performs admission control, checking that there are enough resources in the cluster to

support the virtual machine and starting the VM on a host that has sufficient resources without creating too large of an imbalance. vCenter Server also migrates running virtual machines between hosts as needed to ensure efficient use of cluster resources.

- **Migration Threshold.** The DRS migration threshold allows you to specify which recommendations are generated and then applied (when the virtual machines involved in the recommendation are in fully automated mode) or shown (if in manual mode). This threshold is also a measure of how much cluster imbalance across host (CPU and memory) loads is acceptable. The Three Star threshold provides for a well balanced CPU and memory load across the hosts in the DRS cluster without a high number of VMotion migrations.

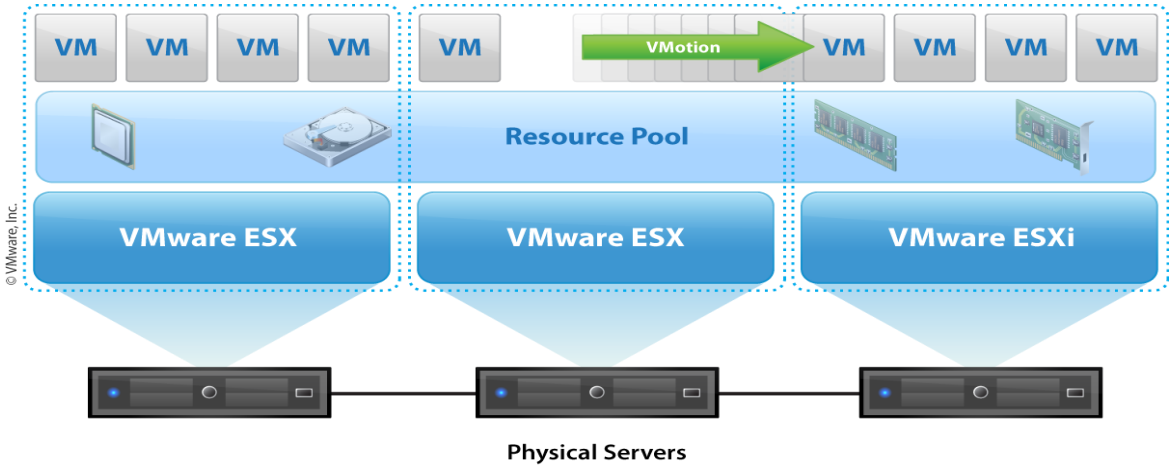


Figure 4 VMware DRS load balancing

VMware ESX Host

The VMware ESX server is the heart of the virtualized environment, providing the platform for running virtual machines. Like any other hardware in an enterprise environment, care must be taken to ensure reliability, consistency and resiliency.

VMware ESX Host Hardware Specifications

The configuration and assembly process for each system should be standardized, with all components installed the same on each host. Standardizing not only the model but also the physical configuration of the ESX hosts is critical to providing a manageable and supportable infrastructure—it eliminates variability.

Servers

- Blade servers and modular servers could be used as virtualization hosts. However, where the cost/benefit analysis makes sense it is recommended that blade servers be used. Blade

servers assist in lowering overall costs, maintenance and support of server hardware because power, networking, storage connectivity, and management interfaces are centralized in a chassis. In addition, rack space requirements are significantly reduced because of the small form factor design of blade architectures.

- It is recommended that all servers in a VMware ESX/ESXi cluster are exactly the same.
- Purchase servers that include an Intelligent Platform Management Interface (IPMI) Such as HP's iLO or Dell's RAC.

CPU's

- It is recommended that all servers have a minimum of two dual-core or quad-core processors.
    - More VM CPUs can be assigned to each core of a dual-core processor than quad-core processors. The numbers are lower with quad-core processors because of bus sharing between the cores.
- Buy the best (fastest, largest cache) CPU you can afford.
    - Often the CPU model just below the top of the line offers the best price to performance ratio.
- Buy CPUs with second-generation hardware virtualization assist and hardware-assisted MMU functionality. These include:
    - Intel CPUs with VT-x and Extended Page Tables (EPT) support
    - AMD CPUs with AMD-V and Rapid Virtualization Indexing (RVI) support
- All processors in a VMware ESX/ESXi cluster must be from the same manufacturer and should be of the same model and speed.

Memory

Typically memory is the limiting factor for virtual machines on an ESX host. As with CPU, budget is likely the primary consideration when deciding how much RAM for each system. Similar to CPU the option just below the top of the line likely offers a better price to performance ratio than the highest density memory modules.

- It is recommended that every server have a minimum of 32 GB of RAM.
- All servers in a VMware ESX/ESXi cluster should have equal amounts of memory.

Network Adapters

- It is recommended to utilize a minimum of 6 GB Ethernet ports or 2 10GB Ethernet ports.
    - This allows for separation of ESX management network traffic to be segregated from VM network traffic while providing redundancy.
    - Additional Ethernet adapters/ports may be necessary to segregate connections to and from multiple physical Ethernet switches to support DMZ/security zone segmentation.

Storage Adapters (HBA's)

- It is recommended to utilize a minimum of 2 Fibre Channel or 2 iSCSI ports for access to shared storage.

In order to improve overall performance, disconnect or disable unused or unnecessary physical hardware devices, such as:

- COM ports
- LPT ports
- USB controllers
- Floppy drives

Disabling hardware devices (typically done in BIOS) can free interrupt resources. Additionally, since devices, such as USB controllers, operate on a polling scheme that consumes extra CPU resources. Lastly, some PCI devices reserve blocks of memory, making more memory unavailable to ESX/ESXi

Host Device Placement

Consistent PCI card slot location, especially for network controllers, is essential for accurate alignment of physical to virtual I/O resources.

The unique nature of blade server chassis, server blades and the associated IP network and Fibre Channel interconnects requires the consistent placement and configuration of all interface ports.

Local Storage

VMware recommends that ESX hosts boot from local storage if available.

When installing ESX (not ESXi), set the “swap” partition to 1600 MB. This will provide sufficient virtual memory swap space to support the maximum service console memory configuration.

VMware vCenter Management Server

VMware vCenter Server System Specifications

It is recommended to run VMware vCenter Management server as a VM on an ESX cluster that has VMware HA configured.

For optimal performance, use the following guidelines:

- Up to 200 hosts, you can use a 32-bit Windows OS for vCenter Server, but a 64-bit Windows OS is preferred
- When you have more than 200 hosts, a 64-bit Windows OS is required
- For up to 50 hosts and 250 powered on VMs, vCenter Server should have at least 2 CPUs, 4 GB memory
- For up to 200 hosts and 2000 powered on VMs, vCenter Server should have at least 4 CPUs, 4 GB memory

- For up to 300 hosts and 3000 powered on VMs, vCenter Server should have at least 4 CPUs, 8 GB memory

Consider not only the current size of the infrastructure, but also the future growth possibilities.

It is recommended that the vCenter database and other vCenter components and add-ons (Converter, Update Manager, SRM Server etc) be installed on separate servers/VMs for larger environments.

vCenter Database

VMware supports Oracle, Microsoft SQL Server and IBM DB2 databases for vCenter.

If an Enterprise has an existing database standard that is supported by VMware, then the enterprise should use their supported standard platform.

For those enterprises that do not have a standard, VMware documentation and support favors Microsoft SQL Server, making this the recommended default.

Install the database system separate from the VMware vCenter Management Server.

Create separate databases for the vCenter Server and other VMware vSphere management products like Update Manager.

Database storage requirements can be estimated using VMware provided calculators:

MS SQL Server http://www.vmware.com/support/vsphere4/doc/vsp_4x_db_calculator.xls

Oracle http://www.vmware.com/support/vsphere4/doc/vsp_4x_db_calculator_oracle.xls

Licenses

The vCenter Server will be configured with the issued 25-character license keys and will automatically install the appropriate license on the ESX hosts as they are added to inventory.

Purchase sufficient licenses for the planned number of CPU's for the datacenter.

Purchase the license version that supports the planned number of cores per socket.

http://www.vmware.com/products/vsphere/buy/editions_comparison.html

vSphere Network Architecture

Network Architecture Logical Design

Following best practices, the network architecture will meet these requirements:

- Separate networks for vSphere management, VM connectivity, VMotion traffic, VMware Fault Tolerance (FT) and IP storage (NFS/NAS or iSCSI)

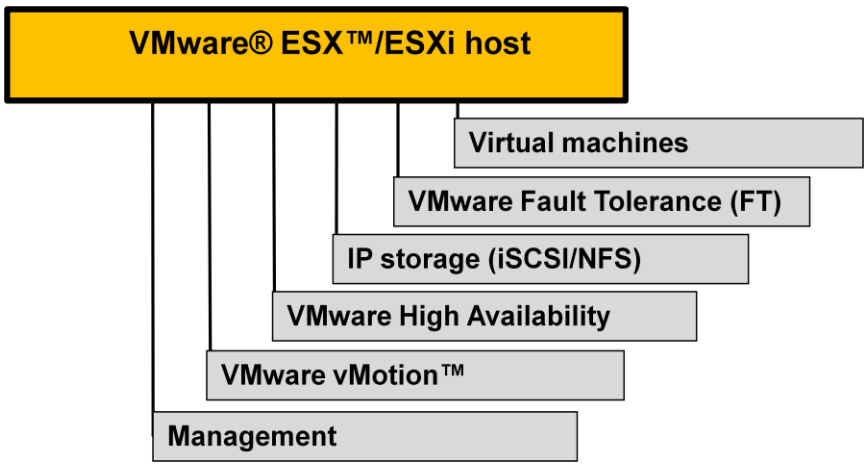


Figure 5 Network separation

- Redundant vSwitch uplinks with at least 2 active physical adapter ports per vSwitch.
- Redundancy at the physical switch level

All VMware ESX hosts will have at least two vSwitches, configured with virtual switch port groups and 802.1q VLAN tagging to segment traffic into VLANs. All physical network switch ports connected to these adapters will be configured as trunk ports with “portfast” enabled. The trunk ports will be configured to pass traffic for all VLANs used by the virtual switch.

The example below illustrates the use of two virtual switches to segment ESX/ESXi management traffic from virtual machine network traffic.

- vSwitch0 supports vSphere management and VMotion. For each host supporting FT, a total of two VMkernel Gigabit NICs is needed—one dedicated to FT logging and one dedicated to VMotion, and both need to be on different subnets. Additional NICs are recommended for VM and management traffic.
- vSwitch1 supports VM network connectivity To support the network demands of up to 60 VMs per host, this vSwitch is configured to use four active Gigabit Ethernet adapters.

The physical NIC ports will be connected to redundant physical switches.

The following diagrams depict the virtual network infrastructure designs:

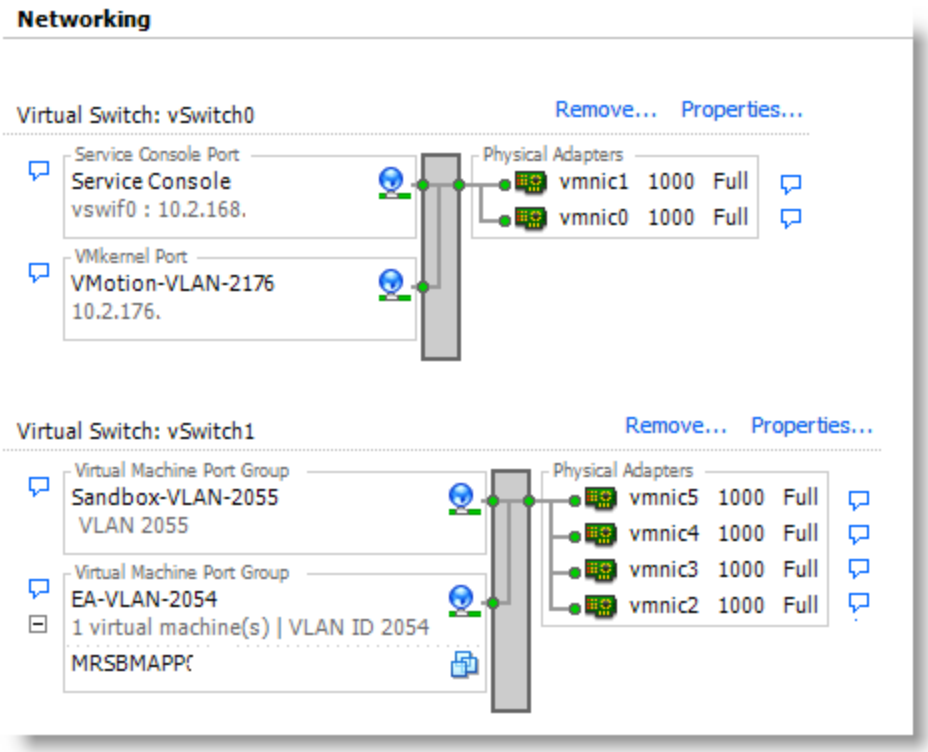


Figure 6 VMware ESX Logical Network Design

Table 3 vSwitch Security Settings

Parameter	Setting
Promiscuous mode	Reject (default)
MAC address changes	Reject
Forged transmits	Reject

vSwitch Security Setting Explanations

- **Promiscuous Mode.** Setting to Reject at the vSwitch level protects against virtual machine virtual network adapters. Placing a VM virtual network adapter in promiscuous mode has no effect on which frames are received by the adapter.
- **MAC Address Changes.** Setting to Reject at the vSwitch level protects against MAC address spoofing. If the guest OS changes the MAC address of the adapter to anything other than what is in the .vmx configuration file, all inbound frames are dropped If the guest OS

changes the MAC address back to match the MAC address in the .vmx configuration file, inbound frames are sent again.

- **Forged Transmits**. Setting to Reject at the vSwitch level protects against MAC address spoofing. Outbound frames with a source MAC address that is different from the one set on the adapter are dropped.

Network Redundancy Considerations

Potential failure points and measures for redundancy identified include the following.

Table 4 ESX Server Hosts

Failure Point	Redundancy
Service console connection	Minimum of 2 physical adapters on "Management" vSwitch
VMotion connection	Minimum of 2 physical adapters on "Management" vSwitch
IP Storage connection (Full height server blades)	Minimum of 2 physical adapters on "IP Storage" vSwitch
VM connection	Minimum of 2 physical adapters on "Virtual Machine" vSwitch
VMware HA Heartbeat connection	2 physical adapters on vSwitch0 provides redundancy at the NIC level. Optionally, can add VMkernel port if additional redundancy is required. Configure DAS isolation addresses.

vSphere Shared Storage Design

ESX storage is storage space on a variety of physical storage systems, local or shared, that a host uses to store virtual machine disks.

A virtual machine uses a virtual hard disk to store its operating system, program files, and other data associated with its activities. A virtual disk is a large physical file, or a set of files, that can be copied, moved, archived, and backed up as easily as any other file. To store virtual disk files and manipulate the files, a host requires dedicated storage space.

The host uses storage space on a variety of physical storage systems, including your host's internal and external devices, or networked storage, dedicated to the specific tasks of storing and protecting data.

The host can discover storage devices to which it has access and format them as datastores. The datastore is a special logical container, analogous to a file system on a logical volume, where ESX places virtual disk files and other files that encapsulate essential components of a virtual machine. Deployed on different devices, the datastores hide specifics of each storage product and provide a uniform model for storing virtual machine files.

To enable the use of the clustered load distribution, high availability and virtual machine fault tolerance, virtual machine disk files should be located on shared storage.

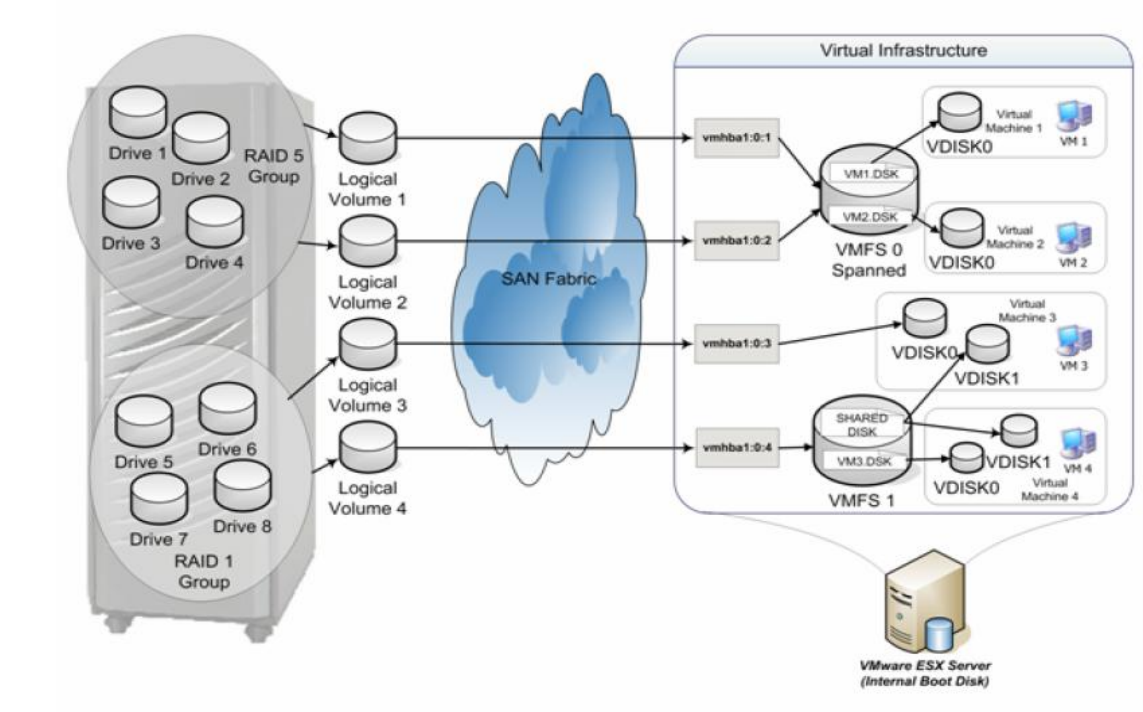


Figure 7 ESX with SAN attached storage

Shared Storage Logical Design

Table 5 Shared Storage Logical Design Specifications

Attribute	Specification
Number of storage processors per storage array	2 (redundant)
Number of switches	2 (redundant)
Number of ports per host per switch	2
VMFS datastores per LUN	1

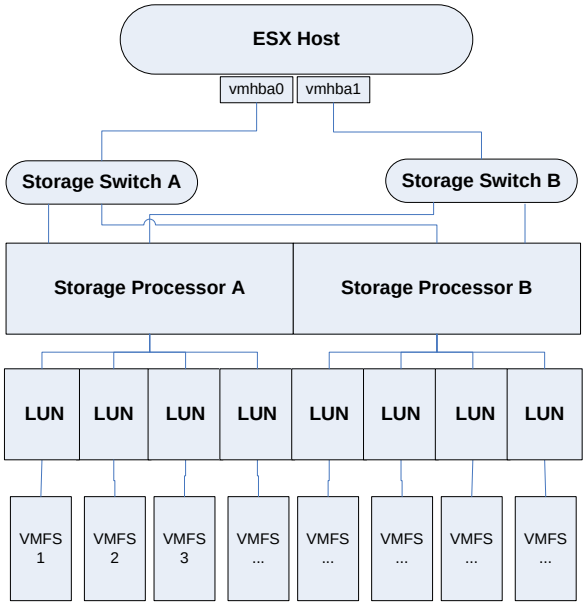


Figure 8 Logical SAN Diagram

Shared Storage Requirements

The consumption of each storage volume will be monitored in production with alarms configured to alert if any approach capacity to provide sufficient time to source and provision additional disk.

Maintain at least 20% free capacity on each VMFS volume for VM swap files, snapshots, logs, and thin volume growth.

Additional LUNS are required for the storage of virtual machine templates, guest operating system installation CD images (ISOs), and to provide administrators second-tier storage for log and VM archival and infrastructure test purposes. The separation of such files from VM files was done recognizing that these non-VM files can often have different I/O characteristics.

For large applications (Exchange, Sharepoint, Oracle, ArcGIS, etc…), the sizing, layout and best practices for storage for large databases or workloads are not dissimilar to physical deployments and may be a good choice for application specific VMFS datastores or Raw Device Mappings. It is recommended to leverage joint reference architectures and sizing tools from the application vendors, VMware and storage vendors. These types of workloads require implementation specific planning and storage allocation.

Datastore Configuration Specifications

Datastore sizing needs to take into account I/O requirements of application and operating systems running in virtual machines and the recovery requirements for these VMs. A balance between performance and manageability will help to determine the number of virtual disks per datastore. Limiting the number of VM disks on a particular datastore helps maintains a reasonable RTO and reduces the risks associated with losing a single LUN.

SAN storage specifications are typical of high load random access environments. To balance the manageability, performance, and availability requirements, the table below provides the following set of configuration specifications.

Table 6 Storage Configuration Specifications

Item	Specification	Reasoning
LUN Size	300 GB to 1TB	Although significantly larger LUNs are possible, this size was chosen for several reasons. For manageability, it allows an adequately large portion of disks to better use resources and limit storage sprawl. A smaller size maintains a reasonable RTO and reduces the risks associated with losing a single LUN. In addition, the size limits the number of VMs that remain on a single LUN.
Block Size	8 MB	Setting the block VMFS block size to 8 MB will allow for VMDK files up to 2 TB if required. Even though the beginning LUN and datastore size is less than 2 TB initially, these can be dynamically expanded without requiring the relocation of the VM.
LUN RAID	5 1+0	Based on the high read to write ratio, RAID-5 was chosen for VM operating systems and general file data disks in order to ensure availability. RAID 1+0 will be used for disks with higher write performance/high availability needs such as database tables or log drives.
VMs per Datastore Hosts per Datastore	Approx 16 per Datastore Approx 16 per Datastore	Although up to 32 hosts can be simultaneously attached to a datastore, it is best to minimize the number of active VMs accessing a given datastore to reduce the amount of I/O contention that might be introduced. Since each VM can be hosted by different servers, the number of hosts can be the same as the number of VMs. In addition, the impact of losing one datastore is reduced.
Template/ ISO LUN	Separate LUNs	To ensure optimal performance, it is recommended to separate VM files from other files such as templates and ISO files that have higher (more sustained) I/O characteristics. A best practice is to dedicate separate datastores/LUNs for VM templates and for ISO/FLP files, separate from the VMs themselves.

Item	Specification	Reasoning
Fibre Channel SAN Zoning	Datacenter	Zoning of the storage should include all VMware ESX hosts in a particular cluster in addition to any VCB proxy server. This provides protection and management of the workload across the VI. Each of the zones will have a single initiator (ESX HBA port) and a single target (each SP from a single storage array).

Storage Path Redundancy Design

Table 7 Storage Path Redundancy

Failure Point	Redundancy
VMware ESX Host	2 “Storage” ports per host
Storage Switch	2 physical switches
Storage Array	2 storage processors per storage array

vSphere Infrastructure Security

Security is critical in this environment, and any security vulnerability or risk exposed by the new vSphere infrastructure would have a negative impact on future adoption of virtualization technology. To protect the business, existing security policies and procedures were considered and leveraged. Microsoft Active Directory users and groups are used to govern access to vCenter roles and privileges.

End users and application administrators will continue to access the virtual machines through the guest OS or application mechanisms and will not have access through VMware vSphere components or the vSphere Client directly. No access will be granted that is not required to perform a specific, authorized job function.

vSphere Host Security

The ESX service console is a limited distribution of Linux based on Red Hat Enterprise Linux 5 (RHEL5). The service console provides an execution environment to monitor and administer the entire ESX host.

Although you can install and run certain types of programs designed for RHEL 5 in the service console, this usage can have serious security consequences and is not supported unless VMware explicitly states that it is.

VMware ESX Service Console Security Specifications

ESX Server includes a firewall between the service console and the network. The allowed ports are used for basic communication with ESX Server. This setting enforces a high level of security for your ESX Server host.

Note: The firewall also allows Internet Control Message Protocol (ICMP) pings and communication with DHCP and DNS (UDP only) clients.

Firewall security settings for VMware ESX are detailed in Appendix A of this document.

By default, SSH access to the ESX service console with root account is disabled.

Function	Setting
Remote login as root using ssh	Disabled

Authentication

ESX uses the Pluggable Authentication Modules (PAM) structure for authentication when users access the ESX host using the vSphere Client, vSphere Web Access, or the service console. The PAM configuration for VMware services is located in `/etc/pam.d/vmware-authd`, which stores paths to authentication modules.

The default installation of ESX uses `/etc/passwd` authentication, just like Linux does.

Users authorized to work directly on an ESX host will have local user accounts. Initially this will be limited to the vSphere Enterprise Administrators.

Enable Active Directory authentication on all ESX hosts to streamline the validation of individual user credentials. The integration with AD will only provide authentication. Authorization will still require local accounts.

Future versions of ESX/ESXi (starting with 4.1) will have built in mechanisms for AD integration.

SUDO

Whether you access the service console locally or through a remote connection such as SSH, users must log in using a user name and password recognized by the ESX host.

When logging onto the ESX host to perform activities that require root privileges, users will be required to log in to the service console with their own account and acquire root privileges through the `sudo` command. The `sudo` command enhances security because it grants root privileges only for select activities in contrast to the `su` command, which grants root privileges for all activities. Using `sudo` also provides superior accountability because all `sudo` activities are logged, whereas if you use `su`, ESX only logs the fact that the user switched to root by way of `su`.

vCenter and Virtual Machine Security

By default, any user or group who is a member of the local Administrators group of the Windows Server running vCenter Server will have full administrative control of vCenter Server (and the virtual

infrastructure). This can allow other system administrators that are not virtual infrastructure administrators access to the virtual infrastructure.

Use the appropriate vCenter Server roles and assign them to the appropriate vCenter Administrators AD group to ensure access is limited to virtual infrastructure administrators.

Before removing users or groups from vCenter Server, make sure that you create and test access to vCenter Server for the new users and groups.

End users of VMs will not need direct access using the VI Client. Remote Desktop can be provided if needed.

Security Considerations with multiple security zones

In this example, shown in Figure 9, you use a combination of physical and virtual technology to enforce trust zone separation. As a result, you can locate virtual servers with different trust levels on the same ESX host. Although physical security devices are part of the configuration, this approach consolidates all virtual machines on the same hosts, thus requiring substantially fewer physical servers. By achieving full server consolidation, you generate significant cost savings for your IT organization. Enforcement of the security zones at the network level takes place in both virtual and physical realms. You use virtual switches to enforce which virtual servers are connected to which zone, but you use physical hardware to enforce the network security between the zones. For this reason, virtual servers must use the physical network and pass through physical security devices to communicate between trust zones.

Because the trust zones in this configuration are enforced in the virtualization layer, you should have clearly named network labels that identify the security zone and VLAN and audit virtual switches regularly for consistent policy and settings to mitigate the potential for a virtual machine to be placed on the wrong network.

Although Figure 9 shows separate virtual switches for each zone, you can accomplish the same goal by using 802.1q VLANs. The most important factor in determining which configuration option to choose is typically the number of physical NICs present in the hardware. You should always dedicate at least one physical NIC to the virtualization management network. If possible, use two physical NICs for the virtualization management network to provide redundancy.

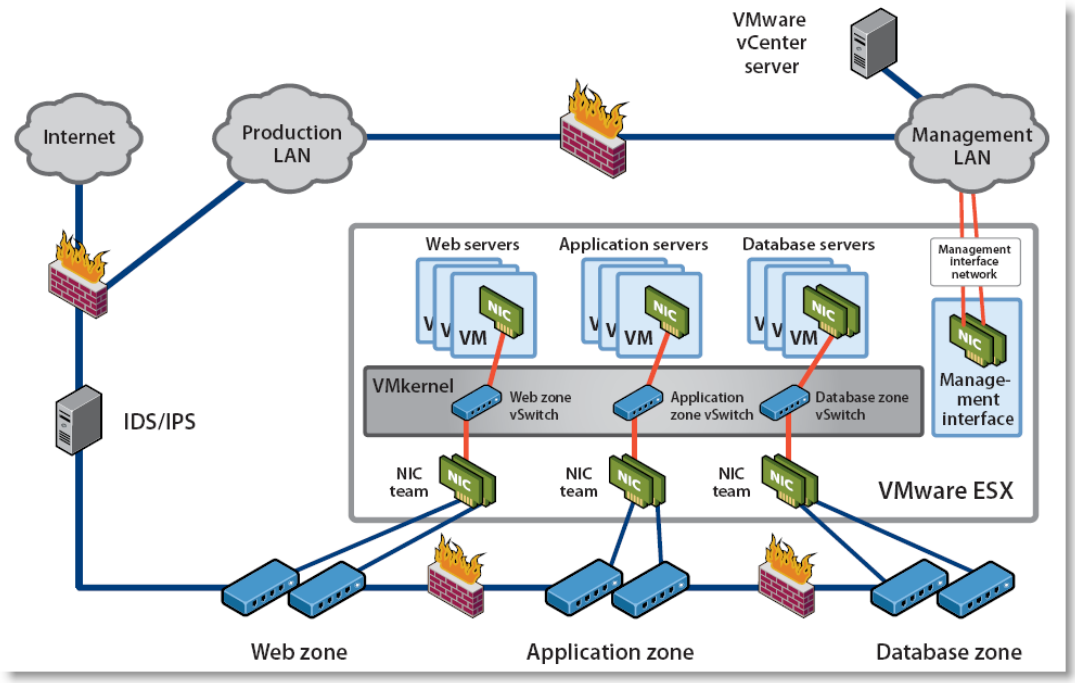


Figure 9 Virtualization with multiple security zones

In the above configuration, separation and hardening of the service console management network is accomplished with physical separation. Further, the virtual switches/portgroups have layer 2 security controls configured as listed in Table 4.

vSphere Network Port Requirements

Appendix B contains detailed information regarding what ports need to be opened for communication between vSphere-related services.

Virtual Machines

Provisioning virtual machines (VMs) is different from provisioning physical machines and needs to be approached differently. It is the over-provisioning and underutilization of servers that has lead to consolidation in order to actually gain the benefit of investment in hardware resources.

In the physical environment, most servers are provisioned based on the maximum that may be needed over the entire lifetime of the server since the intended workload for the server may shift during its lifetime. These physical machines tend to be provisioned with more CPU and more RAM than they really need.

In the virtual environment, machines need to be provisioned with the resources they *really* need. Additional resources can be added later should the workload require them, often without downtime.

"Virtual Machine First"

Organizations that have been most successful with virtualization have adopted a "Virtual Machine First" policy. A virtual machine first policy states that any new servers are provisioned as VM's unless

there is a valid a technical (or business) reason. A “virtual machine first” policy mandates that all new servers are specified as VM’s and tested. This process should follow the organization’s normal application lifecycle model through development, testing and then production.

Virtual Machine Templates

Deploying virtual machines from templates is quick and reduces costly human error. Deploying highly standardized virtual machines and guest operating systems simplifies configuration and troubleshooting.

To simplify administration and to reduce security concerns, templates should only include the software necessary to support application operation.

A typical template should include the following:

- Properly aligned virtual disk (vmdk) file(s)
- The base operating system
- The latest service pack and/or applicable patches
- Any required management or backup agents
- VMware tools

Application software or agents that include host specific configuration that cannot easily be changed or replaced during deployment or may interfere with configuration changes should be left out of the templates and added after the VM has been deployed.

To support the creation and deployment of virtual machines from templates, appropriate volume license media for supported operating systems in ISO format will be placed in a shared VMware ESX datastore. The accessibility of this install media will also eliminate the need for placing copies of the “i386” or “AMD64” directories onto the system drives of templates and by extension the VM’s that are cloned from them.

Use the correct virtual SCSI hardware (e.g. BusLogic Parallel, LSILogic SAS/Parallel, VMware Paravirtual)

Use proper Guest OS type when configuring VMs

- Not setting the proper Guest OS type to match a VM will result in the incorrect VMware Tools installer to be used and changes the default settings specific to each Guest OS (e.g. default memory and disk size).

Develop and use methodology and guidelines for CPU, RAM, network settings

- Having standard guidelines in building VMs can help make resource utilization of VI more predictable.

Understand limitations of serial and parallel devices

- Support for serial and parallel devices is limited to only one connected device per ESX Server host per running VM.

- VMs with serial or parallel devices are tied to their hosts and cannot be migrated using VMotion.
- VMs with serial or parallel devices are often the product of P2V.

Configure and use CD-ROMs and Floppy devices properly

- Remove or disable Floppy drives.
- Configure CD/DVD drives to use ISO images on shared storage.

Avoid using screen savers prior to login

Screen Savers are unnecessary for preventing monitor damage as there are no monitors to burn out, and they waste CPU cycles.

- Use a blank screen saver with password protect instead.
- On Windows VM's remove the "logon.scr" screen saver. This screen saver runs before any user is logged in.

Turn on display hardware acceleration when configuring VMware tools

- Hardware acceleration to full can alleviate mouse jitteriness.

Ensure HAL matches configuration

- A mismatched HAL (HAL type does not match the number of vCPUs) can lead to performance problems in the Guest OS is a common problem with P2V'ed VMs.
- Prior to performing a P2V migration change the Windows HAL to "ACPI"

Power off VMs completely when not in use

- VMs that have their Guest OS shut down but the VM not powered off will still consume resources.
- Non-ACPI VMs will not power off completely when their Guest OSes are shut down.

Understand CPU affinity use

- CPU affinity is an optional setting that pegs a VM to run on certain CPUs on a host. This setting is not normally recommended because doing so will prevent the VM from being migrated using VMotion.

Understand use of resource shares, reservations/minimums, and limits/maximums

- Resource settings can help shape allocation of resources and curtail usage of VMs or give access to minimum amounts of resources.

vSphere Infrastructure Monitoring

Overview

As the uptime and health of the entire technology infrastructure is paramount.

New physical servers purchased to run VMware ESX/ESX should be outfitted with IPMI Baseboard Management Controllers (BMC) used by enterprise monitoring systems to monitor system hardware status such as processor temperature, fan speed, etc.

In environmentsw with sufficient capacity, vSphere Distributed Power Management (DPM) should be considered, and can use these IPMI BMC's to automatically power ESX/ESXi Hosts on and off based on demand to help further power and cooling savings.

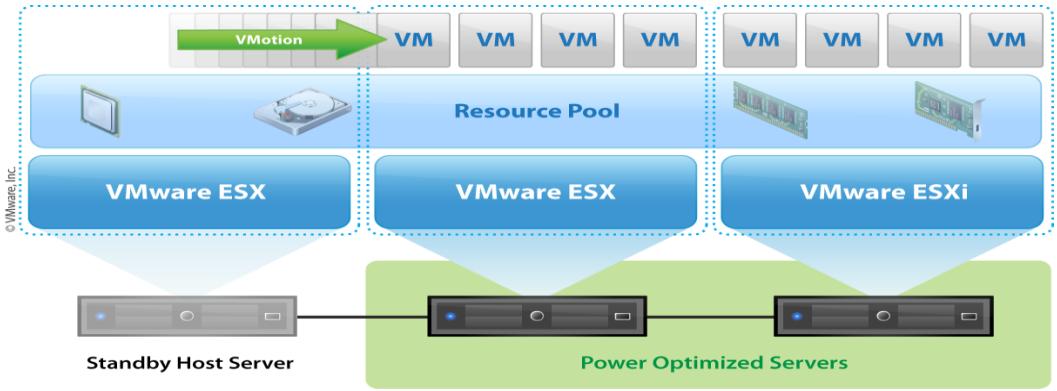


Figure 10 Distributed Power Management

vSphere Monitoring

Leveraging the event monitoring and alarm system in vSphere, vCenter Server can be configured to monitor the health and performance of all critical virtual infrastructure components including the ESX hosts, the clusters, VMware HA and Fault Tolerance, virtual machine operations such as VMotion, and the health of the vCenter Server itself. The events and conditions to be monitored and configured to alert are detailed in Appendix C.

Upon the triggering of an alert, vCenter Server will be configured to send SNMP traps to the enterprise management system's SNMP receiver. Although the same system is primarily responsible for event correlation and email alerting across the enterprise, vCenter Server will also be configured to send email alerts for all triggered events to the vSphere Enterprise Administration group.

vSphere administrators group will be responsible for routinely reviewing and managing the health and system logs generated by the ESX/ESXi hosts, vCenter Server and the virtual machines. These logs will be groomed and archived following corporate log retention policies and procedures.

Virtual Machine Monitoring

The current management system provides monitoring of the systems to be virtualized and will continue performing this task once the systems are converted to vSphere VMs. The monitoring system primarily requires network connectivity to the virtual machines which will be impacted by the conversion, as their IP addresses and host names are being changed. However, the mechanism that monitors the performance of Windows virtual machines utilizes Windows Performance Monitor (Perfmon) counters and will need to be reconfigured to use new, virtualization-specific Windows Perfmon counters provided by VMware Tools. These counters, unlike their counterparts for physical components, are tuned for accurate assessment of *virtualized* Windows performance.

Appendix C provides detailed monitoring system configuration information, including SNMP and SMTP settings, the list of alarms and events to be leveraged, and the Windows Performance Monitor counters to be used by the enterprise monitoring system to monitor virtual machines.

vSphere Infrastructure Patch/Version Management

Overview

Maintaining an up-to-date IT infrastructure is critical. The health, performance and security of the datacenter depends on the health, performance and security of its supporting technology. Maintaining an up-to-date infrastructure can be a daunting task for IT administrators, but if not performed dependably and routinely, the infrastructure is at risk.

VMware vCenter Update Manager, VMware's enterprise patch automation tool, will be implemented as part of the new vSphere infrastructure to keep the vSphere ESX hosts and virtual machines' VMware Tools up-to-date. Administrators will also evaluate patches/updates to VMware vCenter Server and the vSphere Client and update those manually as required.

vCenter Update Manager

VMware vCenter Update Manager is an automated patch management solution that applies patches and updates to VMware ESX hosts, Microsoft Windows virtual machines, and select Linux virtual machines. vCenter Update Manager can also update VMware Tools and VMware virtual hardware in virtual machines. In addition to securing the datacenter against vulnerabilities and reducing or eliminating downtime related to host patching, automated ESX host updates provide a common, installed version for hosts. This is vital for the health of VMware Fault Tolerance which requires the same build to be installed on all hosts supporting FT-protected VMs.

vCenter Update Manager should be installed on a separate system as vCenter Server and configured to patch/upgrade the ESX hosts and VMware Tools installed within the virtual machines. It will, however, not be used to automatically update virtual machine hardware. Virtual machine hardware updates will be evaluated and performed manually as needed.

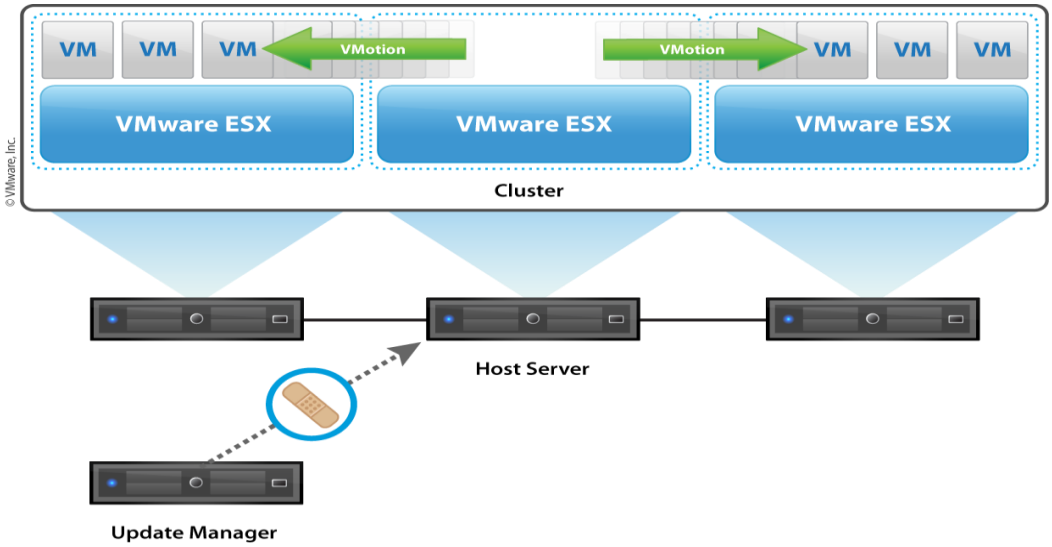


Figure 11 vSphere Update Manager

The Update Manager Server should be run on a separate dedicated system for performance reasons, so as not to overburden the vCenter Server.

Like vCenter, a dedicated database for Update Manager should be created on the database server also housing the database for vCenter Server. Using the vCenter Update Manager Sizing Estimator from vmware.com, and using the following assumptions:

Table 8 vCenter Update Manager Specifications

Attribute	Specification
Patch download sources	Select **Download ESX 4 patches, Download Linux VM patches, Download Windows VM patches** Unselect **Download ESX 3 patches**
Shared repository	D:\VMware Update Manager\Data
Patch download schedule	Every Sunday at 12:00AM EST
Update Manager baselines to leverage	Critical and non-critical ESX/ESXi host patches, Critical guest os patches VMware Tools upgrade to match host
Virtual machine settings	Select **Snapshot virtual machines before remediation to enable rollback** Select **Don't delete snapshots**
ESX host settings	Host maintenance mode failure: **Retry** Retry interval: **30 Minutes** Number of retries: **3**
vApp settings	Select **Enable smart reboot after remediation**

Setting Explanations

- **Patch download sources**. What patches to download.
- **Shared repository**. The vCenter Update Manager Server has been configured with a data disk to be used for storing patches at this location.
- **Proxy settings.** Settings for a proxy server if one is used to access the internet from the datacenter.

- **Patch download schedule.** This is the time and frequency to download new patches.
- **Email notification.** Who Update Manager will automatically notify when new patches have been downloaded.
- **Update Manager baselines to leverage.** Update Manager baselines define a level of patches and updates to monitor for and download.
- **Virtual machine settings.** Initially, Update Manager will not be used to update virtual machines. However, this setting will be configured per best practices to prepare for the event that when VM patching is activated, a snapshot will be taken of each virtual machine prior to performing any remediation operations. This will enable rollback of patches that are applied by vCenter Update Manager, if necessary. These snapshots will not be automatically deleted by Update Manager. Members of the vSphere Administration group will delete the snapshots after determining that the patches have been successfully applied and are functioning correctly.
- **ESX host settings.** Update Manager will place a host into maintenance mode before applying patches. Maintenance mode automatically triggers the migration of any VMs running on the host to other hosts in the cluster to avoid VM downtime. In the event Update Manager and vCenter encounter problems putting a host into maintenance mode, this setting specifies what to do and how many times within what interval between attempts before abandoning attempts to apply patches to a particular host.
- **vApp settings.** vApps are logical groups of VMs. This setting will use the start order of VMs as defined with a vApp when powering on VMs. vApps often require powering on virtual machines in a specific order due to dependencies, and this is configured within the vApps properties.

vCenter Server and vSphere Client Updates

vSphere Administrators will routinely check for and evaluate new vCenter Server and vSphere Client updates. These will be installed manually in a timely fashion following release and proper testing. VMware vSphere Client updates should be manually installed whenever vCenter Server is updated. Using conflicting versions of the vSphere Client and vCenter Server can cause unexpected results. vSphere Client will automatically check for and download an update if it exists when connecting to an updated vSphere Server.

Backup/Restore Considerations

The architecture for VMware ESX provides a very safe and recoverable solution in the case of a VMware ESX host outage or corruption. Since all crucial vmdk files are stored on the SAN as VMFS, they will continue to be available to other VMware ESX hosts within the cluster in the event that a VMware ESX host is taken offline.

VMware ESX Server Host Backup

Backing up VMware ESX is not a necessary practice since a typical build takes minutes from start to finish. Since all critical data is stored on the SAN, it is not necessary to back up the Service Console.

VMware ESX Server Host Recovery

The recovery process for an ESX host is reinstallation with a scripted install. The scripted install process completes in minutes.

Virtual Infrastructure Backup

VMware Consolidated Backup (VCB) can be used to facilitate backups of VMs, both file-level and full VM backups can be performed. If required, a traditional backup agent may be installed into a virtual machine to facilitate backup of databases or other applications with specific requirements beyond simple filesystem backup.

Special vSphere Architecture Design Considerations

Noteworthy items of consideration in design, decision, justification, and impact are listed here.

Table 9 Noteworthy Items

Area	Item	Design Impact
ESX Server Host	Platform choice	The selection of Blade servers, 2-quad core CPUs were chosen.
	Local Storage	All hosts will boot from local storage.
	Storage adapter	For each ESX/ESXi host both HBA ports are active, but for each LUN, only one HBA is active at a time, while a second HBA is a failover adapter.
	Number of NICs	Minimum of 6 GB NIC ports will allow for segregation of VM traffic from management and/or IP storage traffic with sufficient ports for redundancy.
vCenter Management Server	Platform choice	Virtual machine
vCenter Database	Location	Separate database server.

Area	Item	Design Impact
Networking	Segmentation	Production VM networks and their associated security zones are segmented from the VMware Service Console zone.
	Security	• VLANs will be used. • Root will not be given remote ssh access as per VMware recommended security best practices.
	Redundancy	Each vSwitch will have at least 2 active NIC ports.
Storage	Platform choice	MRU/Round Robin failover policy will be needed to match the Active/Passive storage arrays. Fixed/Round Robin failover policy will be needed to match the Active/Active storage arrays.
	LUN allocation	• Separate LUNs for VM OS disks and data disks • Separate LUNs for VM log and database disks • Separate LUNs for VM Templates/ISO's.
VirtualCenter Datacenter Architecture	Service Level Agreements	VMware HA and DRS settings need to reflect SLAs for specific requirements on host load and failover.
P2V Architecture	--	Some ESX Server hosts will have access to the same networks as the existing servers to be consolidated. This will allow converting the machines directly into VMs.
Monitoring Architecture	--	VirtualCenter and ESX hosts should be configured to forward Logs and SNMP traps to a central monitoring server.
Other Customer-Specific Requirements	Policy Requirements and Limitations	A change record must be kept for running production VMs.

vSphere Architecture Redundancy

Potential failure points and measures for redundancy identified include the following.

Table 10 Potential Failure Points and Measures for Redundancy

Failure Point	Redundancy
ESX Server Host	Multiple ESX Server hosts organized into VMware HA Clusters
Blade Server Chassis	vCenter ESX Clusters span multiple chassis
vCenter Server	vCenter Server VM is located on an HA enabled cluster. Backups of vCenter server
vCenter Database	Backups of database taken daily.
Storage	See Storage Redundancy details
Networking	See Networking Redundancy details
VM	• VMware HA • VMware FT

Assumptions

Hardware

Hardware deployment must meet technical requirements for each product. The technical assumptions for this document are listed below.

Table 11 Sources of Technical Assumptions for this Design

Element	Reference
ESX and vCenter Server	ESX/ESXi and vCenter Installation Guide ESX/ESXi Configuration Guide Basic System Administration Guide vSphere 4.0 Configuration Maximums Guide

Element	Reference
ESX host Hardware	vSphere Hardware Compatibility Lists
ESX I/O Adapters	vSphere Hardware Compatibility Lists
ESX SAN Compatibility	Fibre Channel SAN Configuration Guide iSCSI SAN Configuration Guide
VMotion, HA, Fault Tolerance	vSphere 4.0 Availability Guide

External Dependencies

External dependencies address other systems or technologies that depend on or could be affected by vSphere Infrastructure. External Dependencies are different from Assumptions in that they clearly identify dependent factors and the consequent implications.

Table 12 VMware Infrastructure External Dependencies

Item	Requirements
Active Directory	Active Directory is required to implement and operate the VMware Infrastructure.
DNS	DNS must be configured for connectivity between vCenter, Active Directory, VMware ESX and the virtual machines.
DHCP	DHCP must be configured to support the scripted deployment of ESX hosts, automated deployment of virtual machines and for automated addition of Windows VM's into Active Directory.
Network	Network congestion or failure will prevent VMotion from migrating virtual machines.
Network	Network congestion or failure will affect the ability of vCenter to manage VMware ESX hosts.
Storage Area Network	Stability and performance of the SAN will affect the virtual machines.
Time synchronization	Accurate time keeping and time synchronization is critical for a healthy vSphere infrastructure. All components including ESX/ESXi hosts, vCenter Server, the SAN, physical network infrastructure and virtual machine guest operating systems must have accurate time keeping. This is especially critical for virtual machines protected by FT.

Item	Requirements
Staff	Properly trained IT staff is critical for the proper implementation, operation, support and enhancement of the vSphere infrastructure.
Policies and procedures	The policies and procedures governing the use of information technology must be revised to properly incorporate the unique properties and capabilities of virtualization as implemented through this design.
Backup and Recovery	The ability to restore a virtual machine is dependent on the availability and proper function of backup and recovery systems.

Reference Documents

Supplemental White Papers and Presentations

- VMware Infrastructure Architecture Overview: http://www.vmware.com/pdf/vi_architecture_wp.pdf
- Virtualization Overview: http://www.vmware.com/pdf/virtualization.pdf
- What's New in VMware vSphere 4: Performance Enhancements: http://www.vmware.com/files/pdf/VMW_09Q1_WP_vSpherePerformance_P13_R1.pdf
- What's New in VMware vSphere 4:Virtual Networking: http://www.vmware.com/files/pdf/VMW_09Q1_WP_vSphereNetworking_P8_R1.pdf
- What Is New in VMware vSphere 4: Storage: http://www.vmware.com/files/pdf/VMW_09Q1_WP_vSphereStorage_P10_R1.pdf
- Network Throughput in a VMware Infrastructure: http://www.vmware.com/pdf/esx_network_planning.pdf
- Network Segmentation in Virtualized Environments: http://www.vmware.com/files/pdf/network_segmentation.pdf
- The vSphere Availability Guide: http://www.vmware.com/pdf/vsphere4/r40/vsp_40_availability.pdf
- Protecting Mission-Critical Workloads with VMware Fault Tolerance: http://www.vmware.com/resources/techresources/1094
- VMware Infrastructure in a Cisco Network Environment: http://www.cisco.com/application/pdf/en/us/guest/netsol/ns304/c649/ccmigration_09186a00807a15d0.pdf

- Network Segmentation in Virtualized Environments: http://www.vmware.com/files/pdf/network_segmentation.pdf
- VMware ESX 3 802.1Q VLAN Solutions: http://www.vmware.com/pdf/esx3_vlan_wp.pdf
- CLARiiON Integration with VMware ESX: http://www.vmware.com/pdf/clariion_wp_eng.pdf
- Using VMware vSphere with EMC Symmetrix Storage: http://www.emc.com/collateral/hardware/white-papers/h6531-using-vmware-vsphere-with-emc-symmetrix-wp.pdf
- Recommendations for Aligning VMFS Partitions: http://www.vmware.com/pdf/esx3_partition_align.pdf
- Security Design of the VMware Infrastructure 3 Architecture: http://www.vmware.com/pdf/vi3_security_architecture_wp.pdf
- Making Your Business Disaster Ready with VMware Infrastructure: http://www.vmware.com/pdf/disaster_recovery.pdf
- Automating High Availability (HA) Services with VMware HA: http://www.vmware.com/pdf/vmware_ha_wp.pdf
- ESX 4 Patch Management Guide: http://www.vmware.com/pdf/vsphere4/r40/vsp_40_esxupdate.pdf
- Best Practices for Patching ESX: http://www.vmware.com/resources/techresources/1075
- Managing VMware VirtualCenter Roles and Permissions: http://www.vmware.com/resources/techresources/826
- VMware vCenter Update Manager Performance and Best Practices http://www.vmware.com/pdf/Perf_UpdateManager40_Best-Practices.pdf

Supplemental VMware Knowledgebase Articles

- VMotion CPU Compatibility Requirements for Intel Processors: http://kb.vmware.com/selfservice/microsites/search.do?cmd=displayKC&externalId=1991
- VMotion CPU Compatibility - Migrations Prevented Due to CPU Mismatch - How to Override Masks: http://kb.vmware.com/selfservice/microsites/search.do?cmd=displayKC&docType=kc&externalId=1993&sliceId=1&docTypeID=DT_KB_1_1&dialogID=23256056&stateId=0 0 2325069
- Installing ESX 4.0 and vCenter 4.0 best practices: http://kb.vmware.com/selfservice/microsites/search.do?cmd=displayKC&docType=kc&externalId=1009080&sliceId=2&docTypeID=DT_KB_1_1&dialogID=23256161&stateId=0%200%2023250853

- VMware High Availability slot calculation: http://kb.vmware.com/selfservice/microsites/search.do?cmd=displayKC&docType=kc&externalId=1010594&sliceId=1&docTypeID=DT_KB_1_1&dialogID=23256209&stateId=0%200%2023250906
- VMware vSphere 4.0 Software Compatibility Matrix: http://partnerweb.vmware.com/comp_guide/docs/vSphere_Comp_Matrix.pdf
- Processors and Guest Operating Systems that Support VMware Fault Tolerance: http://kb.vmware.com/kb/1008027

Chapter 2: Hyper-V Virtualization Practices

Content Contributions by:

Casey Evans
DMV

David Derks
EDD

Russ Leong
Office of Technology Services

Laura Lichtenberger
EDD

Steve Sax
EDD

Ron Souther
Office of Technology Services

Chris Staniar
EDD

Cliff Stayton
HCD

Erich Cress
State Lands Commission

Jason Johnson
Microsoft

Microsoft and Hyper-V are trademarks of the Microsoft group of companies.

Document Purpose

This document is intended to provide basic information and recommended practices for the Microsoft server virtualization products including Hyper-V and System Center Virtual Machine Manager. This document identifies and highlights recommended practices only. It is not intended to be a substitute for thorough research and adequate training by State staff working with Microsoft server virtualization products.

Microsoft Server Virtualization Overview

The Microsoft server virtualization suite consists of Hyper-V, the enterprise bare metal hypervisor and System Center Virtual Machine Manager (or SCVMM), the enterprise server virtualization management product. SCVMM will be explained in the Management Tools section of this document.

Hyper-V runs between the server hardware layer and the operating system layer allowing multiple instances of a supported OS to run unmodified on a single physical server. Hyper-V is a feature available in the Standard, Enterprise and Datacenter editions of Windows Server 2008 and Windows Server 2008 R2. Hyper-V is also available in a standalone product edition called Windows Hyper-V Server 2008 and Windows Hyper-V Server 2008 R2.

The edition of Hyper-V selected will determine the functionality available in the product as well as the inherent virtualization licensing rights granted, if any, with the product. The Standard edition of Windows Server 2008 includes the right to run one instance of the OS in either a virtual or a physical installation. The Enterprise edition license includes the right to 4 instances of the OS in a virtual operating system environment plus one on the physical server. The Datacenter edition includes the right to run an unlimited number of instances of the OS in a virtual operating system environment. The Hyper-V Server 2008 products have most of the functionality of the Enterprise and Datacenter editions but lack a GUI interface and have no included guest virtualization licensing rights.

Workgroup Recommendations

Each Department or State government entity should select the most cost effective edition of Hyper-V best suited to their needs and environment. Generally, when deploying more than 12 VM's per 2 socket systems the Datacenter edition is the most cost effective. When deploying a pure virtual desktop (VDI) solution or a pure test lab solution, the standalone Hyper-V Server 2008 products can be a good choice. The workgroup recommends the Microsoft Windows Server Virtualization Calculators and the table listed below be used as guidance in determining the correct edition and the correct number of editions to use. The calculator can be found here http://www.microsoft.com/Windowsserver2008/en/us/hyperv-calculators.aspx.

The following table showcases specific scenarios where each Windows Server 2008 edition should be used and where Hyper-V Server 2008 edition should be used:

	Virtualization Needs	Microsoft Hyper-V Server 2008 R2	Windows Server 2008 R2 Standard	Windows Server 2008 R2 Enterprise	Windows Server 2008 R2 Datacenter
Scenarios	Server Consolidation	✔	✔	✔	✔
	Test and Development	✔	✔	✔	✔
	Branch Server Consolidation	✔	✔	✔	✔
	Virtual Desktop Infrastructure (VDI)	✔		✔	✔
	Mixed OS virtualization (Linux and Windows)	✔	✔	✔	✔
	Dynamic Data Center			✔	✔
Features	Host Clustering	✔		✔	✔
	Live Migration	✔		✔	✔
	Large Memory support (Host OS) > 32GB	✔		✔	✔
	Support for >4 Processors (Host OS)	✔		✔	✔
	Local Graphical User Interface		✔	✔	✔
	Ability to Add Additional Server Roles		✔	✔	✔
	Guest Virtualization Rights Included in Host Server License		✔	✔	✔
	Application Failover			✔	✔

Source: http://www.microsoft.com/hyper-v-server/en/us/default.aspx

Management Tools

In The Box Tools

Microsoft Hyper-V ships as a feature of the Windows Server 2008 and Windows Server 2008 R2 server products. As an included feature of the OS, all required tools for full management of Hyper-V are included out-of- the- box. The Hyper-V management console (MMC) is included for full support of all Hyper-V features including Cluster Shared Volumes (CSV). The Windows Failover Cluster management console is included for full support of clustered Hyper-V hosts and highly available (HA) virtual machines. The Windows Server Back tool is included for full support of host and virtual machine backups.

System Center Virtual Machine Manager

SCVMM is a solution for managing the entire feature set of Hyper-V in a single tool. SCVMM leverages a GUI interface connecting to PowerShell commands to manage both Hyper-V hosts and guests. In addition to full management of Hyper-V and guests SCVMM provides the following features:

- Physical-to-virtual migration and virtual-to-virtual migration support
- Intelligent virtual machine placement
- Centralized resource optimization
- Ability to concurrently manage the older Microsoft Virtual Server 2005 product
- Support for management of VMware ESX infrastructure
- Rapid provisioning of virtual machines
- Performance and Resource Optimization
- SCVMM Resource Library for templates, support files, and disk images
- Full support for PowerShell

Workgroup Recommendations

The Workgroup has found that for small or limited deployments of Hyper-V “In the Box” tools may be sufficient although even these deployments tend to mature into environments that can benefit from SCVMM as well as other Microsoft product offerings. The California State Lands Commission, a small State Agency, began using Hyper-V approximately two years ago and is now experienced using both “In the Box” tools and SCVMM. Their experience can be found under “Use Case” below.

Monitoring

The Windows Server operating system includes several tools for monitoring the health of features and roles installed on the server. The Hyper-V feature is fully supported with these “In the Box” tools. For more advanced and automated monitoring of Hyper-V the SCVMM product includes PRO with System Center Operations Manager (SCOM) integration.

PRO and SCOM Integration

Performance and Resource Optimization, or PRO, is used to create a dynamic management environment of virtual resources in a Hyper-V and SCVMM deployment. By utilizing SCOM PRO enabled management packs, SCVMM events can be automated to address situations such as hardware, operating system, or application performance issues or failures. PRO can be enabled to automatically take corrective actions or wait for review and approval by system administrators.

Workgroup Recommendations

Monitoring is a critical part of managing any server environment. The higher density of servers in a server virtualization environment makes monitoring even more critical. The Hyper-V feature can be fully monitored through the standard Windows Server tools such as Performance Monitor and host hardware generally includes hardware monitoring tools provided by the manufacturer. While these tools may be sufficient for small deployments of Hyper-V the Workgroup suggests utilizing SCVMM PRO with SCOM for most environments, large and small.

The resources that you might wish to consider monitoring on an ongoing basis using built-in performance monitoring tools available in all Windows Server 2008 products are presented below under Exhibits.

Hyper-V Host Sizing and Configuration

Server Type

Windows Server 2008 and Hyper-V support any 64-bit (x64) server hardware that is on the Windows Hardware Compatibility list. No special hardware configurations are required other than support for hardware-assisted virtualization (Intel VT or AMD-V) technology. A list of supported hardware can be found here http://www.windowsservercatalog.com

Commodity Server

Traditional rack mounted and tower server configurations are generally referred to as commodity servers. These server types include a broad range of configurations from small 1U platforms with 1 or 2 processor sockets and 4 memory slots to large 7U platforms with 8 processor sockets and 64 memory slots. These types of servers have the most configuration choices and can achieve the highest density of virtual machines per server.

Blade Server

A blade server is a compact modular server design that leverages blade enclosure to provide resources such as power, cooling, network and storage connectivity. Common blade enclosures range in size from 6U to 10U and hold from 4 to 16 blade servers. Blade servers come in two sizes known as full height and half height. Blade servers generally have fewer configuration options, are less expandable and achieve lower virtual machine density then the larger commodity type servers. Blades do however, provide reduced power and cooling requirements as well as substantially reduced cabling requirements in a simple to expand modular platform.

Processor

Hyper-V supports all modern x64 processors from Intel and AMD. These processors generally have from 2 to 8 cores per socket package. These cores are referred to as logical processors. Intel processors may also include a technology known as Hyper-Threading which creates the functionality of additional logical processors for each physical core in the processor socket package. Hyper-V presents virtual processors to a virtual machine and maps these virtual processors back to logical processors through a process known as time slicing. Microsoft supported limit is 8 virtual processors per logical processor. The number of logical processors in a given server will determine the maximum number of possible virtual processors that can be created on that server. As an example, a 4 socket server with 4 logical processors per socket, has been validated and will be supported by Microsoft with up to 128 virtual processors. This is important to consider when sizing server hardware for Hyper-V.

Memory

Hyper-V assigns memory to virtual machines in an exclusive manor. This means the amount of memory assigned to a virtual machine will be reserved for the assigned virtual machine to the exclusion of other processes or virtual machines. The Hyper-V parent partition generally requires at least 512 MB of RAM available after RAM is assigned to virtual machines as virtual memory. Each virtual machine on a host will create a certain amount of memory overhead for the parent partition. A rule of thumb for parent partition overhead related to virtual machine memory is 32 MB of RAM for each VM with up to 1 gigabyte of virtual memory plus 8 MB of RAM for each additional gigabyte of assigned virtual memory. As an example, if a host has 10 virtual machines each with 2 gigabytes of virtual memory assigned the host would see about 400 MB ((10 x 32) +(10 x 8)) of overhead.

(**NOTE:** Early reports from Microsoft mention that with the future release of Service Pack 1 for Windows Server 2008 R2 will add Dynamic Memory to the feature set of Hyper-V. Since little is currently known about this no recommendations can be made at this point.)

Networking

Hyper-V supports two different types of virtual network adapters; emulated network adapters and synthetic network adapters. Synthetic adapters were designed to provide substantially improved network performance with lower processor overhead then emulated adapters can provide. Emulated adapters, while slower, are required for certain scenarios such as PXE boot and when running virtual machines with an unsupported OS. Hyper-V supports multiple network adapters per virtual machine when necessary.

Storage

As with networking adapters, Hyper-V supports two different types of virtual storage adapters; emulated network adapters and synthetic network adapters. Synthetic storage adapters were designed to provide substantially improved performance with lower processor overhead then emulated adapters can provide. Hyper-V storage adapters can connect to 3 different types of virtual

hard disks (VHD's) as well as pass-through disks. Below is a summary of the VHD types from the Microsoft Performance and Tuning Guidelines for Windows Server 2008 R2:

Dynamically expanding VHD

Space for the VHD is allocated on demand. The blocks in the disk start as zeroed blocks but are not backed by any actual space in the file. Reads from such blocks return a block of zeros. When a block is first written to, the virtualization stack must allocate space within the VHD file for the block and then update the metadata. This increases the number of necessary disk I/Os for the write and increases CPU usage. Reads and writes to existing blocks incur both disk access and CPU overhead when looking up the blocks' mapping in the metadata.

Fixed-size VHD

Space for the VHD is first allocated when the VHD file is created. This type of VHD is less apt to fragment, which reduces the I/O throughput when a single I/O is split into multiple I/Os. It has the lowest CPU overhead of the three VHD types because reads and writes do not need to look up the mapping of the block. Fixed-sized VHDs can easily be resized (larger) as additional drive capacity is required. Fixed-size VHDs can be decreased in size though substantially more effort is required to do so.

Differencing VHD

The VHD points to a parent VHD file. Any writes to blocks never written to before result in space being allocated in the VHD file, as with a dynamically expanding VHD. Reads are serviced from the VHD file if the block has been written to. Otherwise, they are serviced from the parent VHD file. In both cases, the metadata is read to determine the mapping of the block. Reads and writes to this VHD can consume more CPU and result in more I/Os than a fixed-sized VHD.

For a deep analysis of VHD performance see the document here:
http://download.microsoft.com/download/0/7/7/0778C0BB-5281-4390-92CD-EC138A18F2F9/WS08_R2_VHD_Performance_WhitePaper.docx

Pass-through Disk

Hyper-V also supports pass-through disks, where the virtual hard disk maps directly to physical storage without encapsulation in a VHD file. This type of storage can reduce CPU overhead and increase I/O to the storage system. Pass-through disks also allow for virtual hard drives of more than 2 TB which is the maximum size limit of a VHD file.

Physical Storage Option

Direct Attached Storage

Direct attached storage is utilized in all cases save those where boot from SAN is leveraged. Even in the event that the primary data store is located on shared storage the parent partition will generally reside on direct attached storage, though the amount of storage will often be only large enough to support the parent.

In many cases ranging from branch office server builds to budgetary considerations and the technical strength of IT staff, it is often desirable and highly functional to use direct attaches storage in your virtual operations.

In circumstances when high availability is not required or possible, or where it is prohibited by the virtual workload, the direct attached storage configuration may be drastically expanded to accommodate storage needs.

Some examples of workloads that would prohibit Hyper-V high-availability configuration, thereby necessitating the use of direct attached storage include:

- Network load balanced web servers,
- SQL database mirroring
- Any Microsoft Failover Cluster workload
- Exchange 2007 CCR (Continuous Cluster Replication)
- Exchange 2010 DAG (Database Availability Groups)
- Software with licensing which may prohibit frequent server relocation (as in Live Migration or Quick Migration)

Shared Storage

Widely considered a best-practice in all virtual environments, shared storage is required in order to take advantage of Hypervisor high availability, dynamic resource allocation (PRO Tips), and Live/Quick migration.

Many shared storage solutions also provide advanced tools that allow for SAN-based backups and offsite replication of data thus providing for additional backup and operational recovery options.

Modern shared storage solutions are available in many configurations consisting of low cost SATA disks, common 10k rpm disks, SAS disks, and extremely fast, high I/O fiber channel disks. Connectivity options include 1/10 gig iSCSI, 4/8 gig fiber channel and 10 gig fiber channel over Ethernet. These varying configurations provide an opportunity to map storage to the appropriate workload. Slow SATA drives may be appropriately used for low I/0 virtual machines while faster 15k SAS or Fiber channel drives should be used for virtual machines requiring higher I/Os such as database servers.

Workgroup Recommendations

In many cases the cost per VM is very similar in commodity servers and blade servers. Each department should evaluate their particular needs when deciding what server type to use for Hyper-V however the Workgroup makes the following general recommendations, many of which are discussed in more detail later in this document:

Server Types

- Modular design requirements, power and cooling reduction goals, white floor size and design, and cabling requirements are should be considered when deciding between blades and commodity servers for Hyper-V deployments.
- Select blade servers when your department requires lower overall power, cooling, cabling and rack space requirements and you have the white floor design to support increased power and heat density.
- Select Blade servers when you require high CPU density but can work with lower I/O density.
- Select large commodity servers when you are looking to maximize the number of virtual machines per physical server.

Processors

- Do not exceed 8 virtual processors per logical processor
- Select processors with at least 4 cores
- Choose more processors cores over higher clocked processors when budgets constrains your options
- Utilize Hyper-threading when using Intel processors
- Understand the workloads that will be run on Hyper-V virtual machines and select processors accordingly.
- Assigning multiple virtual processors to virtual machines can be beneficial to many virtualized workloads but never assign more virtual processors than the virtualized application can consume efficiently. Using a single virtual processor for a VM when appropriate avoids unnecessary loads on the physical host server.
- Do not select more cores then can benefit from the memory capacity of the server; generally a single processor core per 4 gigabytes of RAM sufficient.

Memory

- Understand the workload requirements of the workloads to be virtualized and assign only the amount of RAM necessary for proper virtual machine performance.
- Account for host memory reserve requirements when calculating how much memory to choose when sizing a host server
- A good rule of thumb when sizing a server for memory is to use at least 4 gigabytes of RAM per physical server core. Solutions that require large amounts of memory may push the configuration to 8 gigabytes or more of RAM per physical server core.
- When possible, select higher density memory chips in order to maximize the amount of available memory per memory slot.

- Use standard performance measurement tools to confirm that you have not committed too much or too little memory to your virtual machines. When in doubt error on the side of over-committing and not under-committing RAM.

Network

- Use synthetic adapters when possible
- Use emulated adapters when required but replace them with synthetic adapters as soon as possible (such as when PXE is required for deployment)
- Use dedicated physical and virtual network adapters for iSCSI storage presented directly to the virtual machine.
- Use physical network adapters with network offload capabilities
- Leverage network teaming solutions for the server vendor when necessary to provide highly available network adapters or when additional bandwidth is required

Storage

- Use synthetic storage adapters for optimum performance
- Use fixed size VHD's for optimum performance
- Use a VHD of adequate size for the guest OS. Some example sizes are:
    - 20 gigabytes Windows 2003 Server guest OS
    - At least 30 and up to 40 gigabytes for a Windows Server 2008 guest OS
- Use synthetic SCSI virtual adapters when hot-add virtual machine storage is required or when more than 4 VHD's are required per virtual machine.
- Use pass through disks when SAN tools require it, when near physical disk performance is required or when the target storage LUN presented to the virtual machine exceeds 2 TB.
- Size all storage used by virtual machines the same way you would size the storage for a physical machine.
- Make sure the underlying physical storage solution can meet the combined performance demands of all virtualized workloads that will be placed on the storage device.
- Dynamic and Differencing VHD's can provide a highly efficient platform for building lab and test environments on Hyper-V but should be avoided in most production workloads

High Availability

Hyper-V supports host high availability through the use of Windows Fail-over Clustering. This is the same fail-over clustering technology used with other Microsoft products like SQL server and Exchange server. No special server hardware requirements are required for Windows fail-over clusters other than shared storage. Any server or component on the Windows Server 2008 or Windows Server 2008 R2 hardware compatibility list is supported in a Hyper-V fail-over cluster. When a node in an HA cluster fails, the virtual machines that had been hosted on that node will be restarted on the most appropriate remaining node. High availability in Hyper-V can leverage, but does not require cluster shared volumes. Cluster shared volumes allow multiple hosts in a cluster to access the same storage volume through the use of file level locks instead of volume level locks.

When cluster shared volumes are used, multiple HA virtual machines can be placed on the same LUN and failed-over to another host independently of each other. When cluster shared volumes are not used all virtual machines on a LUN must be owned by the same host and will all fail-over together in the case of a host failure. The Hyper-V Live migration and Quick migration features require that the host servers are part of a fail-over cluster.

Workgroup Recommendations

- Use fail-over clusters when live migration or quick migration is desired
- Use fail-over clusters for virtual machines that require a high level of resiliency to hardware failure
- Use cluster shared volumes when more than one highly available virtual machine will be placed on the same LUN.
- Enable PRO in SCVMM to further extend the capabilities of virtual environments on fail-over clusters
- Use maintenance mode in Windows Server 2008 R2 Hyper-V to evacuate all VM's off of a host when performing scheduled system maintenance
- Host HA does not protect the guest operating system instance from failure or provide the guest OS with clustering features.
- When the guest OS requires additional availability such as patching without downtime guest clustering should be implemented
- Always use a dedicated virtual adapter mapped to a dedicated physical adapter when configuring guest clustering through iSCSI.
- Never combine host HA with guest HA on the same host. As an example don't place a SQL server guest that is part of a SQL mirror group on a host that is part of a fail-over host cluster without excluding the guest from the host HA functionality.

Security

Securing host systems and virtual machines are critical considerations when building and maintaining your virtual environment. While both host and virtual servers should generally be secured much as any other physical server some additional steps are advised.

Please refer to the Hyper-V™ Security Guide, Version 1.0, published: March 2009 for an excellent primer on the security of both host and virtual servers. The document can be found at http://www.microsoft.com/downloads/details.aspx?familyid=2220624B-A562-4E79-AA69-A7B3DFFDD090&displaylang=en

Protecting Host Servers

Server Core

Microsoft recommends the use of the Windows 2008 Server Core because it reduces:

- servicing requirements
- management requirements
- attack surface
- disk space usage

More details on server core can be found here: http://msdn.microsoft.com/en-us/library/ee391626(VS.85).aspx.

In an environment where there are a sufficient number of highly experienced administrations where highly reliable and highly specialized servers are required the use of Server Core may be appropriate and should be considered. In smaller organizations or organizations lacking sufficient administrative experience caution should be taken before implement Server Core.

According to Microsoft (http://msdn.microsoft.com/en-us/library/ee391631(VS.85).aspx), "The minimal nature of Server Core creates limitations":

- There is no Windows shell and very limited GUI functionality (the Server Core interface is a command prompt).
- There is limited managed code support in Server Core.
- There is limited MSI support (unattended mode only).

What does this mean practically speaking? Most server administrators can comfortably and competently manage servers using a GUI. Unless administrators are comfortable using the command line to manage their servers a steep learning curve may be required to adequately manage servers. The time to recover from driver or system failures may also be protracted when using Server Core.

Those considering the use of Server Core can refer to the Server Core Installation Option Getting Started Guide at http://technet.microsoft.com/en-us/library/cc753802(WS.10).aspx

Other Host Considerations

- Run only the Hyper-V role on the host server
    - *"The root partition should be dedicated to the virtualization server role. Additional server roles can adversely affect the performance of the virtualization server, especially if they consume significant CPU, memory, or I/O bandwidth. Minimizing the server roles in the root partition has additional benefits such as reducing the attack surface and the frequency of updates. System administrators should consider carefully what software is installed in the root partition because some software can adversely affect the overall performance of the virtualization server."* http://download.microsoft.com/download/A/2/F/A2F199C0-672E-44E6-BF1D-878E233C3F08/ProvisioningHyper-VVirtualMachineinHostingEnvironment.docx
- Shut down all unnecessary services

  - This may seem an old saw but this as true today as it has ever been. Unnecessary services consume resources needlessly and increase the server's attack surface.
- Use the Windows Server® 2008 Security Compliance Management Toolkit (http://www.microsoft.com/downloads/details.aspx?FamilyID=5534bee1-3cad-4bf0-b92b-a8e545573a3e&displaylang=en)
  - *The Windows Server 2008 Security Compliance Management Toolkit provides you with an end-to-end solution to help you plan, deploy, and monitor the security baselines of servers running Windows Server® 2008 in your environment.*
- Use discrete NICs for management of host vs. operation of virtual machines (see Networking section for additional details).
- Each Host needs its own firewall, antivirus, and intrusion detection software
- Host machines should be added to the appropriate organizational units (OUs) so that Group Policy settings apply correctly.

Protecting virtual machines

- Use Offline Virtual Machines Servicing Tool
  - *"…offline machines do not automatically receive operating system, antivirus, or application updates that would keep them compliant with current IT policy. An out-of-date virtual machine may pose a risk to the IT environment. If deployed and started, the out-of-date virtual machine might be vulnerable to attack or could be capable of attacking other network resources.*
    Therefore, IT groups must take measures to ensure that offline virtual machines remain up-to-date and compliant. At present, these measures involve temporarily bringing the virtual machine online, applying the necessary updates, and then storing it again." http://technet.microsoft.com/en-us/library/cc501231.aspx

- Use private or internal network to prevent test virtual machines from accessing other network resources and conversely to prevent other networking resources from accessing test virtual machines (see Networking for details)
- Use BitLocker™ Drive Encryption on the Hyper-V host
  - The use of BitLocker generally produces a small and often indiscernible performance degradation of the host server. However, this is offset by the fact that its use prevents worries and potentially expensive reporting requirements in the event that virtual machines are stolen or improperly accessed. The cost of such reporting, as has been repeatedly reported in the news over the last several years, can be prohibitive, not only as to financial costs but staff time lost to the effort as well.
  - Administrators should properly plan for and test BitLocker implementations before installing and running it in a production environment. Windows BitLocker Drive Encryption Design and Deployment Guides can be downloaded at http://www.microsoft.com/downloads/details.aspx?familyid=41BA0CF0-57D6-4C38-9743-B7F4DDBE25CD&displaylang=en
- Limit physical administrative access to Host servers

    - Maintain a clear separation of duties between those administrators who are responsible for the operation of host verses virtual servers. Poor administrative decisions/actions on a virtual server will impact that server. Poor administrative decisions/actions on a host can impact every virtual server running on that host.
    - You can use Authorization Manager (AzMan), a snap-in for the Microsoft® Management Console (MMC), to assign selected users and groups to the Hyper-V Administrator role so they can use Hyper-V Manager without being administrators of the physical computer itself.
- Audit access to all virtual machines
    - Virtual machines access should be audited just like physical servers. The use of security auditing and logging, for example, is highly recommended for both physical and virtual servers.
- Delegate virtual machine management (SCVMM2008)
    - The Delegated Administrator profile grants administrative access to a defined set host group(s) and library server(s). Users whom belong to a Delegated Administrator role can use the VMM Administrator Console to modify the configuration of all virtual machines defined on any Hyper-V hosts that they control.
- Leverage Web-based Virtual Machine Manager Self-Service Portal
    - The Self-Service Portal grants administrative access to a defined set of virtual machines through the Web-based Virtual Machine Manager Self-Service Portal.
    - Self-service users cannot use the SCVMM 2008 console to manage virtual machine resources.
    - This practice is consistent with limiting physical and virtual access to physical hosts.

Workgroup Recommendations

Security is an expensive and often double-edged sword. The amount of time, energy and resources committed to securing your virtual environment will rely on a variety of factors which only you can effectively evaluate. The recommendation of the workgroup, then, is to thoroughly evaluate the circumstances of your organization including the nature (i.e. sensitivity) of the data that your organization generates, domiciles and manages and then implement those practices above most suited to your organization and the resources that it can commit to securing its computing environment.

Backup and Recovery of Hosts, Virtual Machines and Datasets

While traditional file-level and application specific back-ups of virtual machine content still plays an important role in the management of data, block level backups of the entire Hyper-V server, as are performed by programs like Windows Server Backup, have now become an invaluable operational recovery tool. A common practice, then, is to incorporate both back-up types into your operational schema.

Traditional File-Level Backups

File level backups are required for certain disk setups. EX Physical disks that are directly attached to a virtual machine and host-level backups of iSCSI volumes in guest VMs can't be backed up using the Hyper-V VSS writer.

The use of commercial back-up software like Symantec's Back-up Exec or Microsoft's Data Protection Manager can and arguably should continue to be used to perform file-level back-ups. This traditional method of backups allows for the rapid recovery of data at the file level and provides for an expanded set of targets for the storage of back-ups not available in Windows Server Backup.

Traditional backups do have their failings in the virtual world, however.

- In the event that you must rebuild/recreate a virtual machine prior to the recovery of files it is imperative that you have documented all VM configuration information to facilitate the rebuild process. Obviously configuration details of virtual machines cannot be restored if they aren't being backed-up.

- Restores can be done at a more granular level but are also more cumbersome. Instead of restoring the entire Hyper-Visor, including the volumes on which virtual machines and their configurations reside, you must first rebuild the virtual machine and then properly configure the operating system. In the case of a file server, for example, numerous file shares may need to be created before data can be restored. This process can be both time consuming and tedious and creates the opportunity for human error.

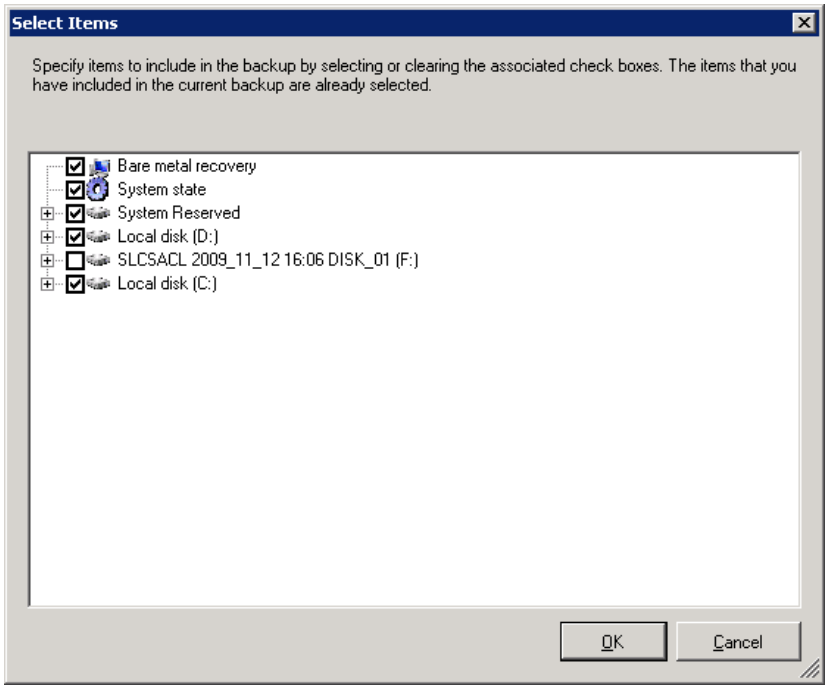
- This approach generally requires the installation of back-up software agents on each virtual machine.

Block-Level Backups

Windows Server Backup is a feature built into all versions of Server 2008 including Server 2008R2. This simple, free backup solution allows you to easily and inexpensively schedule the back-up of the entire host server. Other commercial back-up software that supports VSS Writer can perform the same function but as a rule such software is neither simple nor free.

This type of back-up can be configured to back-up the entire drive system of the host server to facilitate bare-metal recoveries in the event of a catastrophic server failure.

In the example below you can see that “Bare metal recovery”, “System state”, “System Reserved”, and both local disks were selected for back-up.



In the event of a catastrophic server failure one would simply need to attach the external hard drive to which backups were created to the new server, start the server, choose Recover Windows, and the entire contents of the first host can be quickly installed on the new host. The time from server failure to complete recovery can be measured in minutes***. Note that the new server should be substantially similar in architecture to the server it is replacing.***

Like traditional backups this back-up type has its shortcomings as well:

- The number of backup targets is limited and include:
    - File shares (only complete backups will be performed, each time overwriting your previous backups (in other words no differential backups possible)
    - Optical Media or Removable Media
    - Internal Hard Disk
    - External hard disk
- If off-site storage of data is a concern multiple external hard disks can be used as part of the back-up schema. This process may become cumbersome and would involve shipping physical hard disks for storage at off-site locations with all of the incumbent concerns with doing so.
- There are other functional considerations including: Be aware that virtual machines that contain dynamic disks cannot be backed up by WSB. Virtual machines that do not include support for VSS backups—such as Windows 2000, Windows XP, or virtual machines that do

not have Integration Components installed—cannot be backed up using WSB. Lastly, be cautious with the use of snapshots, as virtual machines that contain two or more snapshots will fail to restore. http://technet.microsoft.com/en-us/magazine/2009.06.geekofalltrades.aspx

Remember!

No matter what your backup method is. It is imperative to test your backups by performing restores. Backup and restoration methodologies and practices should be well-documented and tested.

Guest Sizing and Configuration

Many of the administrative and performance guidelines that apply to physical servers also apply to virtual servers. However, administrators must consider additional factors when managing virtual servers.

Basic Configuration

The following items apply to virtual machines of all types and purposes:

- Install and use Integration Components on all supported operating systems
  - Integrations components (ICs) are sets of drivers and services that help your Virtual Machines have more consistent state and perform better by enabling the guest to use synthetic devices.
- A good rule of thumb for virtual machine performance overhead on the host is 110% to 125% of assigned resources. As an example:
  - 16 GB or RAM assigned to virtual machines would be equal to 20 GB of RAM utilization on the host (16 x 1.25)
- Avoid running roles, features or services that are not required by the virtual machine OS to perform its required functions
- Use Server Core for Windows Server 2008 virtual machine operating systems when possible
- For stability and performance use 64-bit guest operating systems when possible
- Apply consistent server naming conventions
  - SCOM will complain if the NetBIOS name of your virtual server is different from the VM name you have assigned that virtual server in SCVMM2008.
- Use proper OS type when building VMs (Standard, Enterprise, Datacenter)
- When possible, use Synthetic Network Adapters, not Emulated (legacy) Network Adapters.
  - This is not always possible as addressed in the Networking section of this document. Synthetic adapters will provide for greater throughput and should be used when possible.
- Virtual Server drives should be sized to the specific needs of the server and organization
  - There is no hard and fast rule regarding the proper size of a virtual hard drive for the System volume or for any other volume on your virtual servers.

    - Oversized hard drives needlessly waste valuable storage space. Undersized hard drives can impair performance.
        - Remember, though, that undersized hard drives can easily be resized. It is far more complex and time consuming to reduce the size of a virtual drive and requires the use of third party products like Partition Magic.
- It is recommended that applications and databases be installed on separate partitions.
- Understand and configure automatic power on options.

Guest Virtual Processors

Each virtual processor creates an overhead for the host. Always configure the virtual machine with the correct number of virtual processors for the workload and avoid arbitrarily creating multi-processor virtual machines if a single processor would suffice.

Guest Memory

Guests should be provided the amount of memory required by the services they provide.

As a general rule start by assigning your guests 2 gigs of RAM but you should be prepared to increase or decrease this amount based on your observations about performance.

- Monitor RAM usage patterns. Standard monitoring tools go a long way to determining whether appropriate RAM has been committed.
    - Create a data collector set in Server 2008 that monitors critical performance data like RAM activity. Collect the data for a reasonable period of time to help determine the sufficiency of RAM committed to the guest.
- Rely on your personal observations and the observations of your constituency.
- Use pass through disks when optimum virtual machine performance is required, when the virtual machine requires more the 2 TB of storage on a single volume, or when you want to leverage native SAN business continuity tools that are not compatible with virtual disks.
- With synthetic adapters there is no disk performance difference between virtual IDE and virtual SCSI attached disks
- Use synthetic SCSI adapters to support hot add of virtual machine virtual hard disks
- Always use dedicated synthetic network adapters to support iSCSI initiators in a virtual machine.

Networking

Networking in Hyper-V is thoroughly addressed in the document *Understanding Networking with Hyper-V* available at: http://www.microsoft.com/downloads/details.aspx?displaylang=en&FamilyID=3fac6d40-d6b5-4658-bc54-62b925ed7eea

While much of this section has been excerpted from that document a number of principals are worth addressing here.

4 Networking Types

There are 4 types of networking available in Hyper-V; private, internal, external and Dedicated External Virtual Network. Each type is described and circumstances under which you might use each has been provided.

- Private: Virtual machines connected to this type of network can communicate among themselves. The management OS has no direct network connectivity with the virtual machines.
    - This networking type allows administrators/developers to create a completely isolated environment for testing and development purposes. No threats to the host server or internal network(s) exist because this networking type creates no communication channels between the virtual machines and the host/internal network(s).
- Internal: Virtual machines connected to this type of network can communicate among themselves and the management OS. There is no connectivity with the physical network.
    - Two interesting uses for an Internal network include:
        - You might want to isolate a virtual machine from your domain network to test a Dynamic Host Configuration Protocol (DHCP) server or a domain controller. Because the virtual machine is isolated, you cannot move files from the virtualization server to the virtual machine by using the domain network. To overcome this limitation, you can use an internal virtual network. http://technet.microsoft.com/en-us/library/ee256061(WS.10).aspx
        - Since you cannot directly bind Hyper-V machines to the 802.11 wireless network connection you can create an internal network and share the wireless connection with your virtual machines. http://bryantlikes.com/SettingUpHyperVVirtualNetworking.aspx

- External: An external virtual network binds to miniports which may exist in the form of multiple miniports for a single physical NIC, a single miniport representing multiple physical NICs, or a single miniport representing a single physical NIC, allowing both virtual machines and the management OS to access the physical network.
    - In English – create an external network when you want your virtual servers to be able to fully interact with other network resources.
- Modified External: The dedicated virtual network is a modified form of the external virtual network offered by Hyper-V. This type of virtual network allows VMs to communicate with other VMs on the same machine, as well as with VMs on other systems. They can also access the external network, although these VMs do not have direct access to the management OS as in the external virtual network configuration. Removing this direct path eliminates many of the drawbacks of the external virtual network type.

    - If two internal or private virtual networks are created in Hyper-V and two virtual machines are created on a separate IP subnet, they cannot communicate with each other. The virtual switch operates at layer 2 of the ISO/OSI Network Model. To achieve routing at higher levels, a router needs to be used the same as would be done in a physical environment. ISA 2006 or RRAS may be used to achieve this functionality.

General Best Practices

- Have at least two physical NICs. If additional services are required, add additional physical network adapters as needed.
    - Use a dedicated NIC for the Hyper-V parent partition.
- Develop and use standard/consistent network adapter naming conventions
- Each virtual machine can have a total of 12 virtual network adapters. Eight network adapters can be assigned to a high-speed synthetic adapter and four network adapters can be assigned to a legacy adapter.
- Whenever possible, use high-speed devices in the virtual machines by enabling the integration services.

Advanced Options

- Use NIC teaming if the added redundancy fault tolerance is required for the application. NIC teaming must be supported by the NIC vendor.
    - NIC teaming is the process of grouping together several physical NICs into one single logical NIC, which can be used for network fault tolerance and transmit load balance. The process of grouping NICs is called teaming. Teaming has two purposes:
    - Fault Tolerance: By teaming more than one physical NIC to a logical NIC, high availability is maximized. Even if one NIC fails, the network connection does not cease and continues to operate on other NICs.
    - Load Balancing: Balancing the network traffic load on a server can enhance the functionality of the server and the network. Load balancing within network interconnect controller (NIC) teams enables distributing traffic amongst the members of a NIC team so that traffic is routed among all available paths.
        - http://www.howtonetworking.com/networking/nicteam1.htm
- Large Send Offload (LSO) and Checksum Offload (CSO). Ensure that LSO and CSO are enabled where they are supported.
    - Large Send Offload (IPv4) and Large Send Offload (IPv6) enable the adapter to offload the task of segmenting TCP messages into valid Ethernet frames. Because the adapter hardware is able to complete data segmentation much faster than operating system software, this feature may improve transmission performance. In addition, the adapter uses fewer CPU resources.
        - http://www.intel.com/support/network/adapter/pro100/sb/CS-029402.htm#lso

    - TCP Checksum Offload (IPv4) and TCP Checksum Offload (IPv6) enable the adapter to compute (Tx) or verify (Rx) the TCP checksum of packets. TCP Checksum Offload is configured under TCP/IP Offloading Options properties when Intel® PROSet for Windows Device Manager is installed. This feature may improve performance and reduce CPU utilization. With Offloading enabled, the adapter computes or verifies the checksum for the operating system.
        - http://www.intel.com/support/network/adapter/pro100/sb/CS-029402.htm#tco
- Use Jumbo Frames for NICS configured to use iSCSI communications.
    - In computer networking, jumbo frames is a feature that allows Ethernet hardware to send, receive, or transport Ethernet frames above 1518 bytes in size. The most common deployments of jumbo frames have an MTU of 9000 bytes.
    - Not all networking equipment supports Jumbo Frames so proceed with care when considering the use of Jumbo Frames

- Enable TCP chimney if it is supported by the NIC manufacturer.
    - TCP Chimney Offload is a networking technology that helps transfer the workload from the CPU to a network adapter during network data transfer. In Windows Server 2008, TCP Chimney Offload enables the Windows networking subsystem to offload the processing of a TCP/IP connection to a network adapter that includes special support for TCP/IP offload processing.
        - http://support.microsoft.com/kb/951037
- Enable Virtual Machine Queuing if it is supported by the NIC manufacturer.
    - Virtual Machine Device Queues (VMDq) is a silicon-level technology that offloads network I/O management burden from the hypervisor. Multiple queues and sorting intelligence in the silicon support enhanced network traffic flow in the virtual environment, freeing processor cycles for application work.
        - http://software.intel.com/file/1919
- Network binding order should be sequenced as follows:
    - The adapter used for managing the Hyper-V parent partition
    - The adapters used for ISCSI, Live Migration and Clustered Shared Volume communications
    - The private network used for used for cluster heartbeat
    - All the adapters associated with virtual networks

Exhibits

Performance Monitoring

Overall health	Hyper-V Virtual IDE Controller	Memory
• Hyper-V Virtual Machine Health Summary • Hyper-V Hypervisor	• Read Bytes / Sec • Write Bytes / Sec • Read Sectors / Sec • Write Sectors / Sec	• Available Bytes • Pages /Sec • Hyper-V Hypervisor Partition • Hyper-V Root Partition • Hyper-V VM Vid Partition
Storage	**Processor**	**Networking**

• Physical Disk
• Hyper-V Virtual Storage Device
  ○ Error Count
  ○ Flush Count
  ○ Read Bytes / Sec
  ○ Write Bytes / Sec
  ○ Read Count
  ○ Write Count
	• Processor
	• Hyper-V Hypervisor Logical Processor
	  ○ %Guest Run
	  ○ %Hypervisor Run Time
	  ○ %Idle Run Time
	  ○ %Total Run Time
	  ○ Hyper-V Hypervisor Root Virtual Processor
	  ○ Hyper-V Hypervisor Virtual Processor
		• Network Interface
		  ○ Bytes Total / Sec
		  ○ Offloaded Connections
		  ○ Packets / Sec
		  ○ Packets Outbound Errors
		  ○ Packets Receive Errors
		• Hyper-V Virtual Switch
		  ○ Bytes/Sec
		  ○ Packets/Sec
		• Hyper-V Legacy Network Adapter
		  ○ Bytes Dropped
		  ○ Bytes Sent / Sec
		  ○ Bytes Received / Sec
		• Hyper-V Virtual Network Adapter
		  ○ Bytes / Sec
		  ○ Packets / Sec

Use Case

Virtualization and the California State Lands Commission (CSLC)

The process of virtualizing servers can seem a daunting task. With dwindling IT staffing levels and what seems to be a never ending stream of demands from our users the idea of wholesale changes to a traditional approach to servers and server management can appear overwhelming. This need not be the case, however.

We have shared the experiences of the CSLC to illustrate that a slow and measured approach to virtualization and the implementation of tools supporting virtualization need not cause too many nights of lost sleep and that even small organizations can reap substantial benefits from deploying other tools designed to support and facilitate the administration of the virtual environment.

Background

The CSLC began virtualizing its server environment approximately two and 1/2 years ago. This was shortly after Microsoft released Hyper-V. In addition to Hyper-V we evaluated several competing products and concluded, based on several factors, that Hyper-V was the best choice for use. Those factors included cost, ease of use, cost of training, cost of maintenance and our ability to identify and hire consultants when system support was required.

Our decision to enter the virtualization world was not predicated on political mandate but rather on many of the now traditional factors currently driving virtualization efforts; the under-utilization of hardware resources, the need to control hardware expense, IT staffing levels and competencies and the desire to reduce the organization's carbon footprint.

Our sometimes slow adoption of other technologies supporting virtual operations, many of which are describe below, was the direct result of circumstances known well by every IT staff member in the State and included competing workloads, a lack of staff with skills and experiences needed to support the products, and not because of a lack of desire on the part of the organization.

Tools and Timelines

Hyper-V MMC Snap-In

For something on the order of 12-15 months we relied exclusively on the MMC to manage all virtual machines. With 4 hypervisors located in three offices in northern and southern California this required us to remote into a physical server to use the MMC to manage local virtual servers. This was the exclusive means by which all virtual servers were managed for more than a year. While not a perfect solution or one promoted or suggested by Microsoft it was practical given our circumstances and served our needs nicely.

System Center Virtual Machine Manager (SCVMM)

Shortly after becoming aware of SCVMM we downloaded and installed the product, still in Beta, and began using it almost daily.

In our case we made very limited use of the product. The program was used to create and configure virtual servers and as a means, although an incomplete one, of quickly determining the health of our virtual servers, and for little more.

In the last 4 months the way we use the product has changed rather significantly. We are now preparing to install the SCVMM 2008R2 administrative console on administrator workstations and to limit administrative access to virtual servers exclusively through the product. This will allow us to make effective use of the administrative roles native to the product including the delegated administration role and the self-service role (discussed elsewhere in this exhibit). The exclusive use of SCVMM to access and manage virtual machines, coupled with the use of native roles, will allow us to create a more secure environment as we will no longer be forced to rely on Remote Desktop to access and administer virtual servers. The use of roles native to SCVMM will allow us to even further restrict administrative access to specific Host Groups.

We have also begun using the product to migrate virtual machines between host servers. As our experiences with virtualization have grown and our needs changed, we found the simplicity of migration offered in the product highly useful and convenient. This process is far easier and more convenient than was our traditional approach which involved copying and pasting VHD files from one server to the next and then creating new virtual machines.

System Center Operations Manager (SCOM)

At the same time our view and use of SCVMM began to change, or perhaps because of it, we installed and have begun using System Center Operations Manager (SCOM). This tool has allowed us greater insight into the health of our networking operations and paid handsome dividends to us within 24 hours of deployment. We were immediately made aware of network latency and disk IOP issues, issues to that point in time we of which we were unaware. SCOM has also provided us early warning about disk utilization as well as the degradation of disk performance due to fragmentation. It also brought to our attention problems we were suffering with our Active Directory structure and provided us the information necessary to take immediate corrective actions.

As part of our effort to deploy SCOM we also performed the suggested SCOM/SCVMM integration. This was a seamless effort and has provided all users of SCVMM ready access to critical performance measures of all physical hosts and virtual servers. The information provided has helped us tune our virtual servers to maximize hardware resources. In some cases we had over-committed resources, particularly memory. In other cases resources had been under-committed. Without the ready reporting made available by the integration we would not have been able to so readily identify issues and take the necessary corrective actions.

Self-Service Portal

We have recently built a self-service portal for use by our administrators. Even in a small shop with so few administrators, or perhaps because of it, I was spending far too much time creating virtual servers for development and production purposes. While the actual time to stand up new servers was measured in minutes I was still required to adjust an already overburdened work calendar in

order to accommodate the requests for new servers. Once the deployment of the self-service portal has been competed I will be freed of the tedium and demands on my time as I am no longer required to stand up new virtual servers. The use of the self-service portal will also allow us the opportunity to limit the ability of our administrators to build a limited number of servers based on templates and hardware profiles that are secure and consistent with internal requirements.

Windows Server Backup

Because of prior issues arising from our backup efforts we have become sensitive, perhaps overly sensitive, to ensuring that we have reliable backups of our servers, both physical and virtual. With the release of Server 2008 Backup as a standard feature on the operating system we are comfortable that we have solid and reliable backups at our disposal.

Until such time that we are able to purchase SANs for placement in our two principal offices, and until we have sufficient bandwidth to support site-to-site replication of data, we have come to rely on attached USB drives as targets for backups of our Hypervisors. We still perform file level backups as well but the use of Windows Server Backup has proven to be an invaluable addition to our backup strategies.

Next Steps

Data Protection Manager (DPM)

As is the case with many organizations server/data backups have always proven problematic. Although we are currently using a commercial software package to manage and conduct backups issues with the product have arisen. In particular, there has often been a predictable lag between the release of new products like Server 2008R2 and the ability of our backup software to actually backup such releases. At the moment none of our R2 servers are being backed up by our current backup software. This has resulted from a combination of the time lag in the backup software vendor's release of an agent that supports R2 and our internal ability to install, test and deploy the new agents.

We believe that our next logical step is to investigate the functionality of Data Protection Manager. We not only expect that this product will meet our immediate and future backup needs but we anticipate that it may alleviate long standing issues with our current back-up software.

SAN Storage

It has become clear to the organization that the most prudent and desirable next step is the purchase and deployment of an iSCSI SAN coupled with the deployment of a cluster server pair. This physical architecture will not only provide the organization and its staff consistent and guaranteed access to critical networking resources but will also alleviate many of the challenges the organization has faced for years vis-à-vis the creation of current, dependable backups as well as the ability to perform necessary restores from these back-ups.

Live Migration

Beyond addressing long-standing backup issues and providing a highly available environment for our

uses we are looking forward to using Live Migration.

As a server administrator there is nothing more frustrating than knowing that every reboot of a physical host is going to take multiple virtual servers off-line, even if only for a few minutes as the host server is rebooted. The only effective way to mitigate the impact on users is to work or require my IT staff to work after hours every time simple administrative tasks like patching is performed.

The ability to rapidly migrate a virtual server from one host to another, so that administrative tasks can be performed on that server with no noticeable impact to our staff is something that we have aspired to for years. The fact that our overworked system administrators will no longer be required to work odd hours to manage their servers would also be welcome.

The bottom line

By using Hyper-V we were able to take a slow and measured approach to virtualizing our environment while containing costs. The product has proven highly stable, affordable, and easy to maintain. More important to us are the facts that the product is well laid-out and highly intuitive to use. We did not have the staff or resources to suffer a long and steep learning curve and with Hyper-V we didn't need to.

We have also been able to deploy and use products in support of our virtualization as time and financial resources have allowed.

Erich Cress
California State Lands Commission

References

Performance Tuning Guidelines for Windows Server 2008

http://www.microsoft.com/whdc/system/sysperf/Perf_tun_srv.mspx

Performance Tuning Guidelines for Windows Server 2008 R2

http://www.microsoft.com/whdc/system/sysperf/Perf_tun_srv-R2.mspx

Microsoft Virtualization Calculators

http://www.microsoft.com/Windowsserver2008/en/us/hyperv-calculators.aspx.

Hyper-V TechNet Library

http://technet.microsoft.com/en-us/library/cc753637(WS.10).aspx

System Center Virtual Machine Manager TechNet Library

http://technet.microsoft.com/en-us/library/bb740737.aspx

Microsoft Server Virtualization Website

http://www.microsoft.com/virtualization/en/us/products-server.aspx

Hyper-V Server 2008 product page

http://www.microsoft.com/hyper-v-server/en/us/default.aspx

Microsoft VHD Performance Analysis

http://download.microsoft.com/download/0/7/7/0778C0BB-5281-4390-92CD-EC138A18F2F9/WS08_R2_VHD_Performance_WhitePaper.docx

Appendix A - ESX Service Console Firewall Settings

To ensure the integrity of the service console, VMware has reduced the number of firewall ports that are open by default. At installation time, the service console firewall is configured to block all incoming and outgoing traffic except for that on ports marked "Yes" in the allow column of the table below.

In the following table, **Underlined** text denotes changes from default settings. Note that the listing of a known service or application in this table does not mean that network traffic is automatically allowed.

Table 13 VMware ESX Service Console Firewall Settings

Access	Incoming Port	Outgoing Port	Protocols	Allow
Secure Shell SSH Client	N/A	22	TCP	No (Default)
Secure Shell SSH Server	22	N/A	TCP	Yes (Default)
SNMP	**161**	**162**	**UDP**	**Yes**
Common Information Model (CIM) SLP	427	427	UDP, TCP	Yes (Default)
VNC Server	5900-5964	N/A	TCP	No (Default)
VMware vCenter Agent	N/A	902	UDP	Yes (Default)
Commvault Dynamic	8600-8619	8600-8619	TCP	No (Default)
Kerberos	N/A	749, 88	TCP	No (Default)
NFS Client	N/A	111, 2049	UDP, TCP	No (Default)
Tivoli Storage Manager Agent	1500	1500	TCP	No (Default)
NTP Client	**N/A**	**123**	**UDP**	**Yes**
SMB Client	N/A	137-139, 445	TCP	No (Default)
CIM Server	5988	N/A	TCP	Yes (Default)
Commvault Static	8400-8403	8400-8403	TCP	No (Default)
CIM Secure Server	5989	N/A	TCP	Yes (Default)
VMware License Client	N/A	27000, 27010	TCP	Yes (Default)
Active Directory Kerberos	N/A	464, 88	TCP	No (Default)
Software iSCSI Client	N/A	3260	TCP	No (Default)
Symantec NetBackup Agent	13732, 13783, 13720, 13734	N/A	TCP	No (Default)
FTP Client	N/A	21	TCP	No (Default)
EMC AAM Client	2050-5000, 8042-8045	2050-5000, 8042-8045	TCP, UDP	Yes (Default)
Telnet Client	N/A	23	TCP	No (Default)
FTP Server	21	N/A	TCP	No (Default)

Access	Incoming Port	Outgoing Port	Protocols	Allow
NIS Client	N/A	111, 0-65535	UDP, TCP	No (Default)
Symantec Backup Exec Agent	10000-10200	N/A	TCP	No (Default)
VI Web Access	80, 443	80, 443	TCP	Yes (Default)
Converter Access	443	443	TCP	Yes (Default)
VM Console	902,903	902,903	UDP	Yes (Default)
VMotion	8000	8000	TCP	Yes (Default)

Appendix B – Port Requirements

Table 14 ESX/ESXi Port Requirements

Description	Port(s)	Protocol	Direction
vSphere Client to ESX/ESXi host	443, 902, 903	TCP	Incoming
VM Console to ESX/ESXi host	903	TCP	Incoming
ESX/ESXi host and vCenter Heartbeat	902	UDP	Incoming/ Outgoing
ESX/ESXi host DNS client	53	UDP	Outgoing
ESX/ESXi host NTP client to NTP server	123	UDP	Outgoing
ESX/ESXi host NFS	111, 2049	TCP, UDP	Outgoing
VMotion between ESX/ESXi hosts	8000	TCP	Incoming/ Outgoing
HA between ESX/ESXi hosts	2050-2250, 8042-8045	TCP, UDP	Incoming/ Outgoing
ESX/ESXi host to Update Manager	80, 443, 9034	TCP	Outgoing
Update Manager to ESX/ESXi host	902, 9000-9010	TCP	Incoming
ESX/ESXi host CIM Client to Secure Server	5988, 5989	TCP	Incoming
ESX/ESXi host CIM service location protocol	427	TCP, UDP	Incoming/ Outgoing

Table 15 vCenter Server Port Requirements

Description	Port(s)	Protocol	Direction
vSphere Client to vCenter Server	443	TCP	Incoming
vSphere Web Access to vCenter Server	443	TCP	Incoming
VM Console to vCenter Server	902, 903	TCP	Incoming
ESX/ESXi host and vCenter Heartbeat	902	UDP	Incoming/ Outgoing
LDAP	389	TCP	Incoming
Linked Mode SSL	636	TCP	Incoming
ESX/ESXi 2.x/3.x host to legacy License Server	27000, 27010	TCP	Incoming/ Outgoing
Web Services HTTP	8080	TCP	Incoming
Web Services HTTPS	8443	TCP	Incoming
vCenter SNMP server polling	161	UDP	Incoming
vCenter SNMP client trap send	162	UDP	Outgoing
vCenter DNS client	53	UDP	Outgoing
vSphere Active Directory integration	88, 445	UDP, TCP	Outgoing
ODBC to MS SQL Server database	1433	TCP	Outgoing
Oracle Listener port to Oracle database	1521	TCP	Outgoing

Table 16 vCenter Converter Standalone Port Requirements

Description	Port(s)	Protocol	Direction
Converter Client (GUI) to Converter Server	443 (configurable)	TCP	Incoming
Converter Server to remote Windows powered-on Machine – remote agent deployment, Windows file sharing	445 and 139	TCP	Incoming
Converter Server to remote Windows powered-on Machine – remote agent deployment, Windows file sharing	137 and 138	UDP	Incoming
Converter Server to remote Windows powered-on machine – agent connection	9089	TCP	Incoming
Converter Server/Linux agent to remote Linux powered-on machine	22	TCP	Incoming
Converter Server/Agent to managed destination – VM creation/management (includes VM Helper creation/management)	443	TCP	Incoming
Windows powered-on machine to managed destination – hot clone – access (vCenter/ESX/ESXi)	443	TCP	Incoming
Windows powered-on machine to managed destination – hot clone – copy (ESX/ESXi)	902	TCP	Incoming
Windows powered-on machine to hosted destination – hot clone – Windows file sharing	445 and 139	TCP	Incoming
Windows powered-on machine to hosted destination – hot Clone – Windows file sharing	137 and 138	UDP	Incoming
Helper VM to Linux powered-on machine – hot clone	22	TCP	Outgoing

Description	Port(s)	Protocol	Direction
Converter Server/Agent to managed source/destination – VM import – access (vCenter/ESX/ESXi)	443	TCP	Incoming
Converter Server/Agent to managed source/destination – VM import – copy from/to ESX/ESXi (Traffic from ESX/ESXi to ESX/ESXi direct for disk-based cloning only)	902	TCP	Incoming
Converter Server/Agent to hosted source/destination – VM import – Windows file sharing	445 and 139	TCP	Incoming
Converter Server/Agent to Hosted Source/Destination – VM Import – Windows file sharing	137 and 138	UDP	Incoming

Table 17 vCenter Update Manager Port Requirements

Description	Port(s)	Protocol	Direction
Update Manager to vCenter Server	80	TCP	Incoming
Update Manager to external sources (to acquire metadata regarding patch updates from VMware	80, 443	TCP	Outgoing
Update Manager client to Update Manager server	8084	TCP	Incoming
Listening ports for the web server, providing access to the plug-in client installer and the patch depot	9084, 9087	TCP	Incoming
Update Manager to ESX/ESXi host (for pushing virtual machine and host updates/patches)	902	TCP	Incoming

Appendix C – Monitoring Configuration

Table 18 Physical to Virtual Windows Performance Monitor (Perfmon) Counters

Old Physical Hardware Counter	New Virtualization Aware Counter
Processor - % Processor Time	VM Processor - % Processor Time
-	Effective VM Speed in MHz (new)
% Committed Bytes in Use	Memory Active in MB
-	Memory Ballooned in MB (new)
% Committed Bytes	Memory Used in MB

Default vSphere Host Alarms to be Used

- Host hardware system board status
- Host power state
- Host memory status
- Host processor status
- Host disk status
- Host network status
- Host connection and power state
- Host memory usage
- Host CPU usage
- Host disk usage
- Host network usage
- Host storage status
- License error
- Cannot connect to network
- Cannot connect to storage

Default vSphere Cluster Alarms to be Used

- All HA hosts isolated
- Cluster deleted
- Cluster overcommitted
- HA admission control disabled
- HA agent unavailable
- HA disabled

- HA host failed
- HA host isolated
- Host resource overcommitted
- Insufficient failover resources
- No compatible host for secondary VM
- Virtual machine Fault Tolerance state changed
- Timed out starting secondary VM
- Cluster High Availability error
- Migration error

Default vSphere Datastore Alarms to be Used

- Datastore disk usage (%)
- Datastore state to all hosts

Table 19 Modifications to Default Alarm Trigger Types

Trigger Type	Condition	Warning	Condition Length	Alert	Condition Length
Host Memory Usage	Is above	80	For 10 minutes	90	For 5 minutes
Host CPU Usage	Is above	80	For 10 minutes	90	For 5 minutes
Datastore Disk Usage	Is above	85	For 30 minutes	90	For 5 minutes

Appendix D – P2V or VM Suitability Flowchart

The following flowchart can be used to determine whether an existing physical server is a candidate for virtualization. By extension this can also be applied to a new implementation to help determine whether the workload can be run as a VM or not.

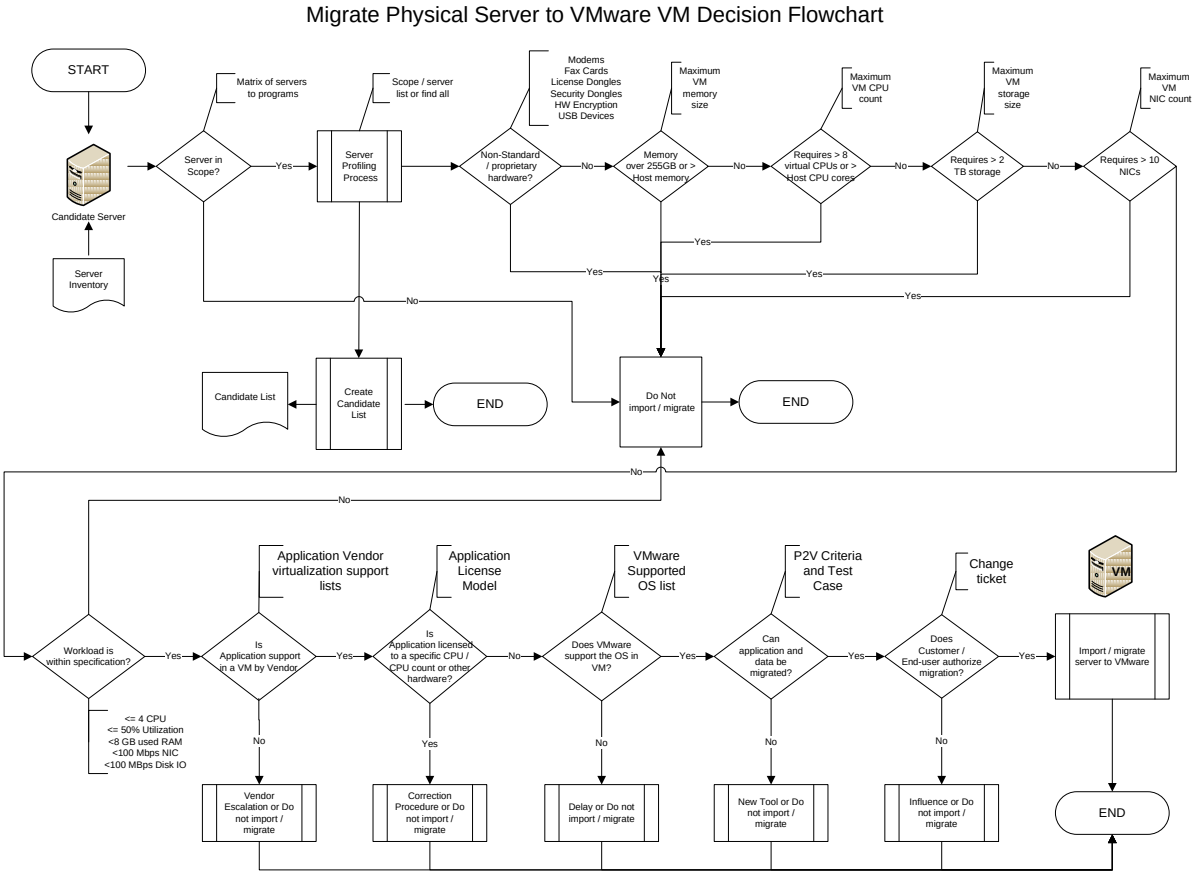


Figure 12 P2V Decision flowchart

Acknowledgements

Microsoft and Hyper-V are trademarks of the Microsoft group of companies.

VMware, the VMware “boxes” logo and design, Virtual SMP and VMotion are registered trademarks or trademarks of VMware, Inc. in the United States and/or other jurisdictions. All other marks and names mentioned herein may be trademarks of their respective companies.