Dell Hyper-V Cloud Fast Track Reference Architecture for vStart200

Reference Architecture and Validation Guide

Release 1.3 for Dell 12th generation servers

Dell Virtualization Solutions Engineering

Revision: A00

March 2012

This document is for informational purposes only and may contain typographical errors and technical inaccuracies. The content is provided as is, without express or implied warranties of any kind.

March 2012| Rev A00

Table of Contents

Tables

Figures

1 Introduction

1.1 Validation Criteria

🛑 **Mandatory A:** This is a mandatory best-practice and is required to pass Microsoft validation. May use Microsoft or non-Microsoft technology

Mandatory B: This is a mandatory best-practice and is required to pass Microsoft validation. Must use Microsoft technology. No technology replacements will be allowed

⚠ **Recommended:** This is a recommended best-practice but may deviated from as appropriate

🟢 **Optional:** This is optional but important - called out for reference

2 Technical Overview

2.1 Cloud Attributes

Note: This section contains a verbatim copy of the NIST Definition of Cloud Computing v15.

On-demand Self-Service. A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service's provider.

Broad network access. Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g., mobile phones, laptops, and PDAs).

Resource Pooling. The provider's computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. There is a sense of location independence in that the customer generally has no control or knowledge over the exact location of the provided resources but may be able to specify location at a higher level of abstraction (e.g., country, state, or datacenter). Examples of resources include storage, processing, memory, network bandwidth, and virtual machines.

Rapid Elasticity. Capabilities can be rapidly and elastically provisioned, in some cases automatically, to quickly scale out and rapidly released to quickly scale in. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be purchased in any quantity at any time.

Measured Service. Cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service (e.g., storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported providing transparency for both the provider and consumer of the utilized service.

2.2 Hyper-V Cloud Architecture Principles

2.2.1 Resource Pooling

Resource optimization is a principle that drives efficiency and cost reduction and is primarily achieved through resource pooling. Abstracting the platform from the physical infrastructure enables optimization of resources through shared use. Allowing multiple consumers to share resources results in higher resource utilization and a more efficient and effective use of the infrastructure. Optimization through abstraction enables many of the Hyper-V Cloud principles and ultimately helps drive down costs and improve agility.

2.2.2 Elasticity and Perception of Infinite Capacity

From a consumer's perspective, cloud services appear to have infinite capacity. The consumer can use as much or as little of the service as needed. Using the "electric utility provider" as a metaphor, the consumer consumes as much as they need. This utility mindset requires that capacity planning be paramount and must be proactive so that requests can be satisfied on demand. Applying this principle reactively and in isolation often leads to inefficient use of resources and unnecessary costs. Combined with other principles, such as incenting desired consumer behavior, this principle allows for a balance between the cost of unused capacity and the desire for agility.

2.2.3 Perception of Continuous Availability

From the consumer's perspective, cloud services appear to always be available when needed. The consumer should never experience an interruption of that service, even if failures occur within the Hyper-V Cloud environment. To achieve this perception, a provider must have a mature service management approach combined with inherent application resiliency and infrastructure redundancies in a highly automated environment. Much like the perception of infinite capacity, this principle can only be achieved in conjunction with the other Hyper-V Cloud principles.

2.2.4 Drive Predictability

Predictability is a fundamental principle from all cloud perspectives whether you are a consumer or provider. From the vantage point of the consumer, cloud services should be consistent; they should have the same quality and functionality any time they are used.

A provider must deliver an underlying infrastructure which assures a consistent experience to the hosted workloads in order to achieve this predictability. This consistency is achieved through the homogenization of underlying physical servers, network devices and storage systems.

From the provider's service management perspective, this predictability is driven through the standardization of service offerings, as well as through standardization of processes. The principle of predictability is necessary for driving service quality.

2.2.5 Metering/Chargeback (Service Provider's Approach to Delivering IT)

Historically, when IT has been asked to deliver a service to the business, they purchase the necessary components and then build an infrastructure specific to the service requirements. This results in longer time to market, increased costs due to duplicate infrastructure, and often does not meet the business expectations of agility and cost reduction. Further compounding the problem, this model is often used when an existing service needs to be expanded or upgraded.

The principle of taking a Service Provider's approach to delivering infrastructure transforms IT's approach. If infrastructure is provided as a service, IT can now leverage a shared resource model that enables economies of scale and combined with the other principles, achieves greater agility in providing services.

2.2.6 Multi-Tenancy

Multi-tenancy refers to the ability of the infrastructure to be logically subdivided and provisioned to different organizations or organizational units. The traditional example is a hosting company which provides servers to multiple customer organizations. Increasingly, this is also a model being utilized by individual organization with a centralized IT organization providing services to multiple business or organizational units within the organization and treating each as a customer or tenant.

2.2.7 Security and Identity

Security for the Hyper-V Cloud is founded on three pillars: Protected Infrastructure, Application Access, and Network Access.

Protected Infrastructure leverages security technologies as well as identity technology to ensure that hosts, information, and applications are secured across all scenarios in the datacenter, including physical (on premises) and virtual (on premises and cloud).

Application Access ensures that IT can extend vital application access not only to internal users but also to vital business partners and cloud users.

Network Access uses an identity-centric approach to ensure that users, whether they're internal employees or in remote locations, have secure access on numerous devices to ensure that business gets done the way it should to maintain productivity.

Most important is that the Secure Datacenter leverages a common integrated technology to ensure that users have simple access with a common identity and that management is integrated across physical, virtual, and cloud environments so that business can take advantage of all capabilities without requiring additional significant financial investments.

2.3 Conceptual Architecture

One of the key drivers of the layered approach to infrastructure architecture that is presented here is to enable complex workflow and automation to be developed over time by creating a collection of simple automation tasks, assembling them into procedures that are managed by the management layer, and then creating workflows and process automation that are controlled by the orchestration layer.

2.3.1 Scale Units

In a modular architecture, the concept of a scale unit refers to the point where a module in the architecture can scale to before another module is required. For example, an individual server is a scale unit, it can be expanded to a certain point in terms CPU and RAM but beyond its maximums, an additional server is required to continue scaling. Each scale unit also has an associated amount of physical installation labor, configuration labor, etc. With large scale units such as a pre-configured full rack of servers, the labor overhead can be minimized.

It is critical to know the scale limits of all components, both hardware and software, to determine the most optimum scale units as input to the overall architecture. Scale units enable the documentation of all the requirements needed (space, power, HVAC, connectivity, etc.) required for implementation.

2.4 Servers

The hardware-architecture choices that are available to datacenter architects are constantly evolving. Choices range from rack-mounted servers to tightly integrated, highly redundant blade systems to container models. The same spectrum exists for storage and networking equipment.

Server scale limits are well published and include number and speed of CPU cores, maximum amount and speed of RAM, number and type of expansion slots, etc. Particularly important are the number and type of onboard IO ports as well as the number and type of supported IO cards. Both Ethernet and Fiber Channel expansion cards often provide multi-port options where a single card can have 4 ports. Additionally, in blade server architectures, there are often limitations in the amount of IO card and/or supported combinations. It is important to be aware of these limitations as well as the oversubscription ratio between blade IO ports and any blade chassis switch modules.

A single server is not typically a good scale unit for a Hyper-V Cloud solution due to the amount of overhead required to install and configure and individual server.

2.5 Storage

Storage architecture is a critical design consideration for Hyper-V Cloud solutions. The topic is challenging as it is rapidly evolving in terms of new standards, protocols, and implementations. Storage and supporting storage networking are critical to the overall performance of the environment but also the overall cost as storage tends to be one of the more costly items.

Storage architectures today have several layers including the storage array(s), the storage network, the storage protocol, and for virtualization, the clustered file system utilizing the physical storage.

One of the primary objectives of the Private Cloud solution is to enable rapid provisioning and de-provisioning of virtual machines. Doing so at large scale requires tight integration with the storage architecture and robust automation. Provisioning a new virtual machine on an already existing LUN is a simple operation, however, provisioning a new LUN, adding it to a Host cluster, etc. are relatively complicated tasks that also greatly benefit from automation.

2.6 Networking

Many network architectures include a tiered design with three or more tiers such as Core, Distribution, and Access. Designs are driven by the port bandwidth and quantity required at the edge, as well as the ability of the Distribution and Core tiers to provide higher speed uplinks to aggregate traffic. Additional considerations include Ethernet broadcast boundaries and limitations, spanning tree and or other loop avoidance technologies, etc.

A dedicated management network is a frequent feature of advanced datacenter virtualization solutions. Most virtualization vendors recommend that hosts be managed via a dedicated network so that there is not competition with guest traffic needs and to provide a degree of separation for security and ease of management purposes. This typically implies dedicating a NIC per host and port per network device to the management network.

With advanced datacenter virtualization, a frequent use case is to provide isolated networks where different owners such as particular departments or applications are provided their own dedicated networks. Multi-tenant networking refers to using technologies such as VLANs or IPSec isolation techniques to provide dedicated networks that utilize a single network infrastructure or wire.

Managing the network environment in an advanced datacenter virtualization solution can present challenges that must be addressed. Ideally, network settings and policies are defined centrally and applied universally by the management solution. In the case of IPSec based isolation, this can be accomplished using Active Directory and group policy to control firewall setting across the hosts and guest as well as the IPSec policies controlling network communication.

For VLAN based network segmentation, several components including the host servers, host clusters, VMM, and the network switches must be configured correctly to enable both rapid provisioning and network segmentation. With Hyper-V and host clusters, identical virtual networks must be defined on all nodes in order for a virtual machine to be able to failover to any node and maintain its connection to the network. At large scale, this can be accomplished via PowerShell scripting.

2.7 Virtualization

The virtualization layer is one of the primary enablers in environments with greater IT maturity. The decoupling of hardware, operating systems, data, applications, and user state opens a wide range of options for better management and distribution of workloads across the physical infrastructure. The ability of the virtualization layer to migrate running VMs from one server to another with zero downtime and many other features that are provided by hypervisor-based virtualization technologies providing a rich set of capabilities. These capabilities can be utilized by the automation, management, and orchestration layers to maintain desired states (such as load distribution) or to proactively address decaying hardware or other issues that would otherwise cause faults or service disruptions.

As with the hardware layer, the virtualization layer must be able to be managed by the automation, management, and orchestration layers. The abstraction of software from hardware that virtualization provides moves the majority of management and automation into the software space, instead of requiring people to perform manual operations on physical hardware.

2.8 Automation

The ability to automate all expected operations over the lifetime of a hardware or software component is critical. Without this capability being embedded in a deep way across all layers of the infrastructure, dynamic processes will grind to a halt as soon as user intervention or other manual processing is required.

Windows PowerShell and several other foundational technologies, including WMI and WS-Management, provide a robust automation layer across nearly all of Microsoft's products, as well as a variety of non-Microsoft hardware and software. This evolution provides a single automation framework and scripting language to be used across the entire infrastructure.

The automation layer is made up of the foundational automation technology plus a series of single-purpose commands and scripts that perform operations such as starting or stopping a VM, rebooting a server, or applying a software update. These atomic units of automation are combined and executed by higher-level management systems. The modularity of this layered approach dramatically simplifies development, debugging, and maintenance.

2.9 Management

The management layer consists of the tools and systems that are utilized to deploy and operate the infrastructure. In most cases, this consists of a variety of different toolsets for managing hardware, software, and applications. Ideally, all components of the management system would leverage the automation layer and not introduce their own protocols, scripting languages, or other technologies (as it increases complexity and may require additional staff expertise).

The management layer is utilized to perform activities such as provisioning the storage-area network (SAN), deploying an operating system, or monitoring an application. A key attribute is its abilities to manage and monitor every single component of the infrastructure remotely and to capture the dependencies among all of the infrastructure components.

2.10 Orchestration

The orchestration layer leverages the management and automation layers. In much the same way that an enterprise resource planning (ERP) system manages a business process such as order fulfillment and handles exceptions such as inventory shortages, the orchestration layer provides an engine for IT-process automation and workflow. The orchestration layer is the critical interface between the IT organization and its infrastructure. It is the layer at which intent is transformed into workflow and automation.

Ideally, the orchestration layer provides a graphical interface in which complex workflows that consist of events and activities across multiple management-system components can be combined, so as to form an end-to-end IT business process such as automated patch management or automatic power management. The orchestration layer must provide the ability to design, test, implement, and monitor these IT workflows such as with System Center Opalis.

2.11 Service Management

The Service Management layer provides the means for automating and adapting IT service management best practices, such as those found in Microsoft Operations Framework (MOF) and the IT Infrastructure Library (ITIL), to provide built-in processes for incident resolution, problem resolution, and change control. By providing an integrated service management platform, Service Manager can reduce costly downtime and improve the quality of the services in the datacenter.

2.12 Tenant / User Self-Service

The Tenant / User Self-Service layer provides an interface for Hyper-V Cloud tenants or authorized users to request, manage, and access the services, such as virtual machines, provided by the Hyper-V Cloud architecture. Using role-based access control and authorization, the Self-Service layer provides the ability to delegate certain aspects of administration (such as starting/stopping VMs) to designated “tenant administrators”.

3 Reference Architecture

3.1 Workload Categories

3.1.1 Server Virtualization and Consolidation

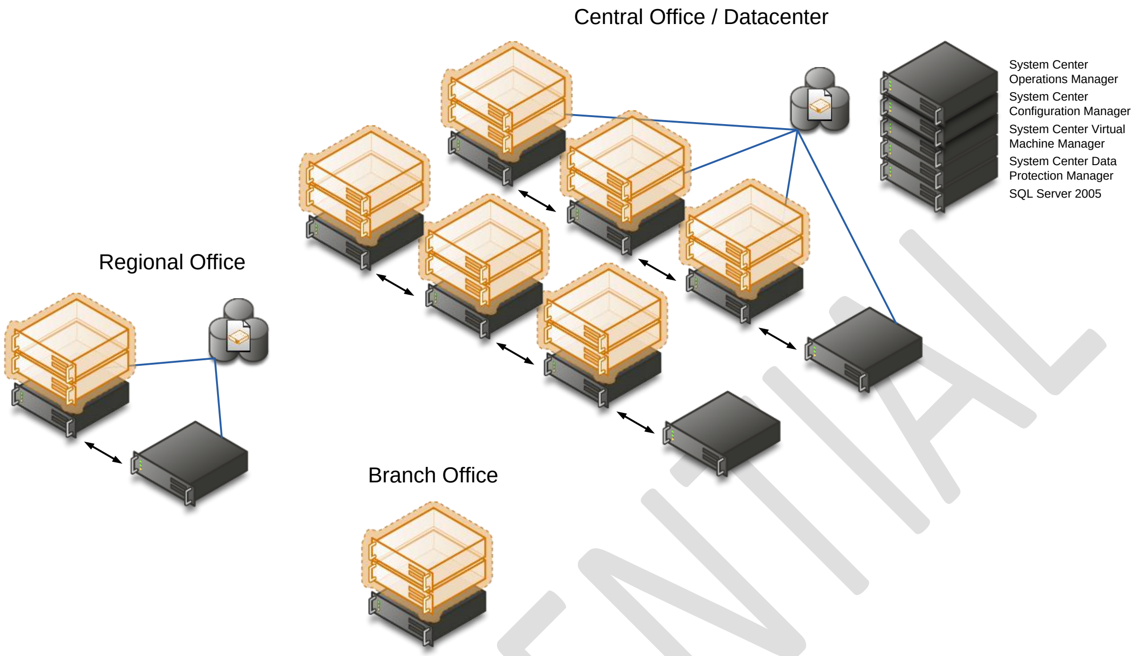
Server virtualization is based on the abstraction of physical system resources so that multiple logical partitions can be created and host a heterogeneous set of operating systems running simultaneously on a single physical server.

Rather than paying for many under-utilized servers, each dedicated to a specific workload, server virtualization allows those workloads to be consolidated onto a smaller number of more efficiently utilized physical systems. Server Virtualization provides the following benefits:

- Consolidates multiple, under-utilized physical servers on a single host, running Virtual Machines
- Reduces workforce/space/kilowatts by leveraging virtualization for server consolidation and agility
- Helps save money because less managing, less space and less kilowatt hours are needed

Virtualization can also help you simplify and accelerate provisioning. The acquisition of workload resources and hardware can be decoupled. Adding the capability required for a particular business process (say, a web commerce engine) becomes streamlined and immediate. In a more advanced virtualized environment, workload requirements can be self-provisioning, resulting in dynamic resource allocation.

Figure 1. Server Virtualization and Consolidation

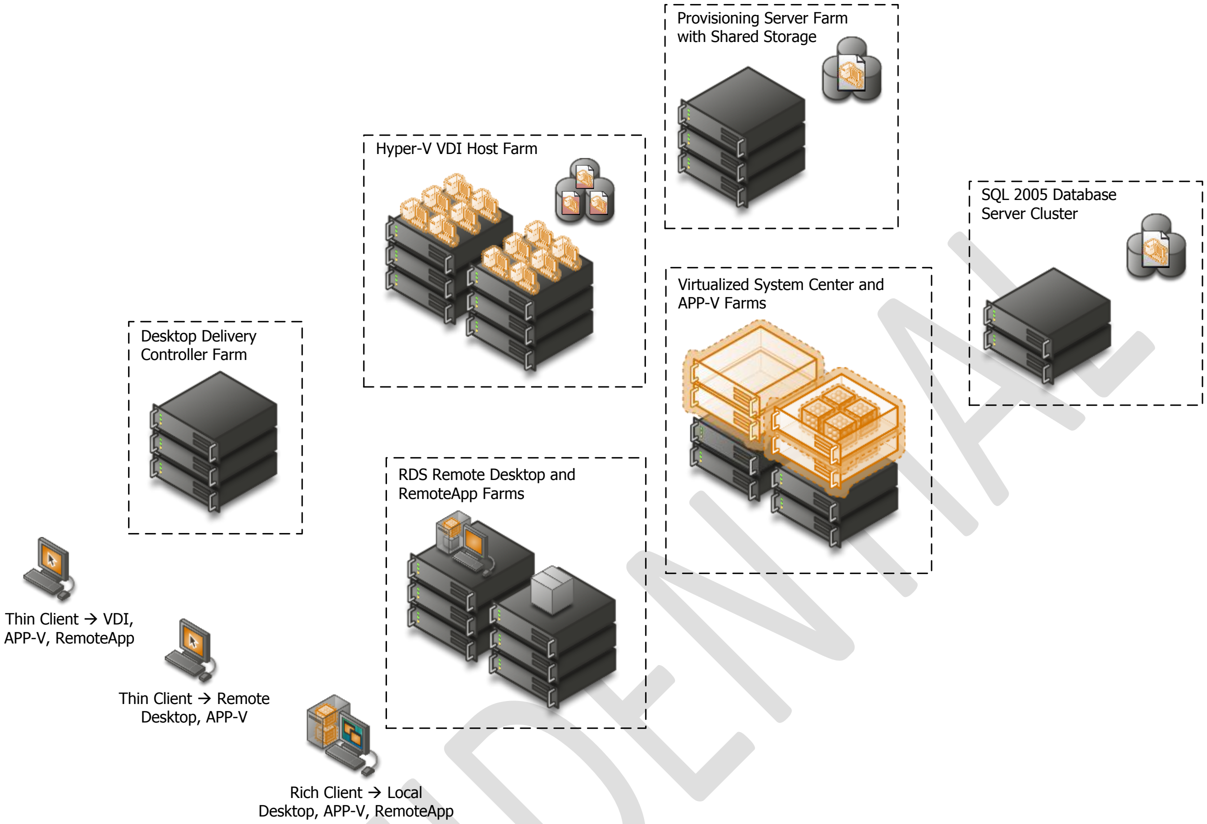

Central Office / Datacenter
System Center Operations Manager
System Center Configuration Manager
System Center Virtual Machine Manager
System Center Data Protection Manager
SQL Server 2005
Regional Office
Branch Office

While virtualization-based server consolidation can provide many benefits, it can also add complexity if the environment is not managed properly. The savings from hardware consolidation could be offset by increases in IT management overhead. Because creating virtual machines is so easy, an unintentional and unnecessary virtual sprawl can result that far exceeds physical server sprawl and that outpaces the tools used to manage virtual machines. A properly managed virtual infrastructure, however, automatically determines which servers are the best candidates for virtualization allowing administrators to initiate automated process to convert them to virtual machines, and provision them to the right hosts in minutes, rather than the weeks or months it takes to procure and configure physical servers manually.

3.1.2 Virtual Desktop Infrastructure

Virtual Desktop Infrastructure (VDI) enables IT staff to deploy desktops in virtual machines on secure and centralized hardware. A centralized and optimized virtual desktop enables users to access and run their desktop and applications wherever they may be, while IT is able to build a more agile and efficient IT infrastructure. Flexible Windows desktop scenarios give organizations the ability to choose the client computing scenarios that best meet the unique needs of their businesses.

Figure 2. Microsoft VDI and App-V

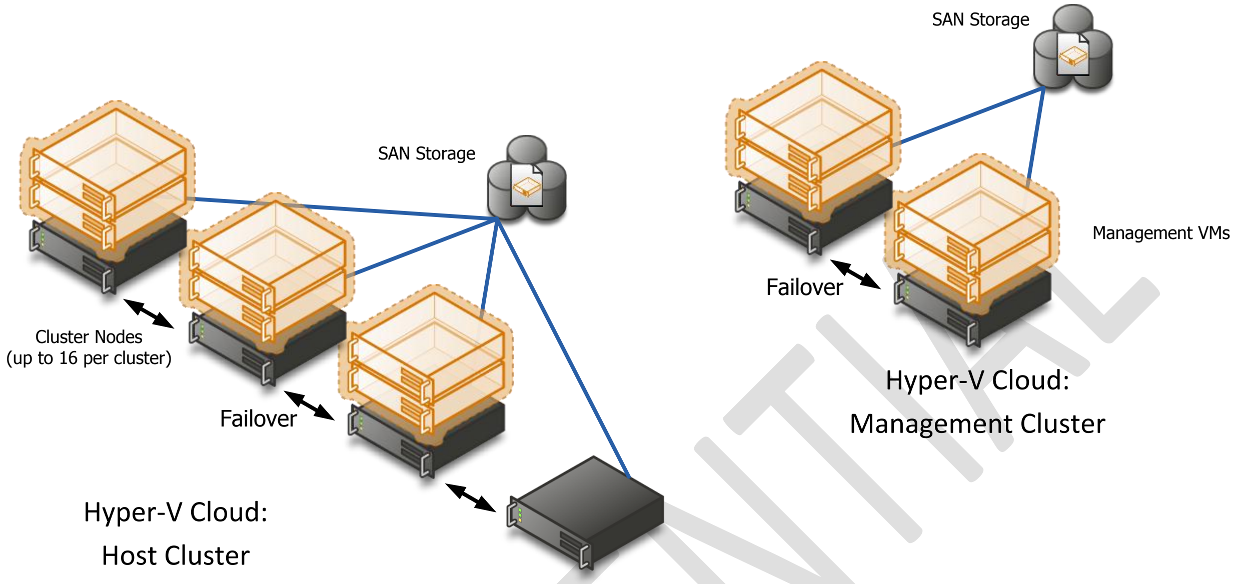

Provisioning Server Farm with Shared Storage
Hyper-V VDI Host Farm
SQL 2005 Database Server Cluster
Virtualized System Center and APP-V Farms
Desktop Delivery Controller Farm
RDS Remote Desktop and RemoteApp Farms
Thin Client → VDI, APP-V, RemoteApp
Thin Client → Remote Desktop, APP-V
Rich Client → Local Desktop, APP-V, RemoteApp

When an organization is managing its virtual infrastructure with the same tools it uses to manage its physical assets, it can reduce system complexity and streamline changes made to the overall infrastructure. By using some or all of these technologies together, organizations can provide very flexible solutions to support many user scenarios, including mobile knowledge workers, corporate knowledge workers, contract and offshore developers, contract employees, and end users in branch locations.

Virtualizing the entire computing infrastructure provides tremendous time and cost savings, as well as flexibility benefits. However, attempting to separately manage each layer of the stack and each instance within those layers (such as individual virtual machines) creates a much more complex situation than is necessary. Using different tools for virtualized resources can result in duplicate or competing processes for managing resources, adding complexity to the IT infrastructure. This can undermine the benefits of virtualization. A virtualized world that isn't well managed can be less reliable and perhaps even more expensive than its non-virtualized counterpart.

3.2 Logical Architecture

Figure 3. Hyper-V Cloud Fast Track Logical Architecture



A Private Cloud is far more than highly available infrastructure providing computing resources to higher level applications. A fundamental shift of Cloud computing is that of IT moving from server operator to Service Provider. This means a set of services accompany the infrastructure such as reporting, usage metering, self-service provisioning, etc. If these services are unavailable, the Cloud “service layer” is unavailable and we’re back to providing little more than a traditional datacenter. For this reason, when managing 8 compute nodes or greater, high availability must also be provided to the management systems. We achieve this via a management Host cluster typically with 2 nodes.

Mandatory: > Dedicated 2-node or more host cluster for Hyper-V Cloud management (if Hyper-V Cloud is 8 compute-nodes or larger)
> All management products deployed in HA VMs on the management cluster. One or more 4 to 16-node host cluster(s) for Hyper-V Cloud tenant VMs
> A Storage Area Network (SAN) and storage array compatible with Windows Failover Clustering
> Gigabit Ethernet or above switched network infrastructure.

Note: In well controlled environments with homogeneous workloads (single workload type private cloud) it can be acceptable to host the management stack on the host cluster instead of creating a separate management cluster. This configuration variant provides a way to further consolidate a private cloud environment; however the same general architectural principals that are laid out in this document apply. Specifically it is important to continue to look at the management stack as a separate logical unit which has to be implemented in a highly available fashion. From a practical perspective this means additional infrastructure or reserves will need to be put in place to ensure this. As management and automation is at the heart of a cloud solution, it is critical that the management resource reserves be configured properly. Without proper configuration and reserves, the solution will lose its cloud attributes as potentially no insight, management, metering or new deployment is possible.

Mandatory: If consolidating the management stack onto the host cluster, the following requirements are to be met

> All management products deployed in HA VMs on the host cluster

> The clustered virtual machines hosting SQL server are guaranteed to run on 2 different physical nodes

> A Storage Area Network (SAN) and storage array compatible with Windows Failover Clustering

> A separate or dedicated storage path for the management VMs. This can be achieved through physical separate storage connections or for example by means of bandwidth reserves on a shared infrastructure

> Networking: The host cluster needs to provide a strong isolated/independent network for the management stack network traffic to ensure bandwidth independence from the running workloads. This can be done for example via separate physical network connections or through shared networks with bandwidth guarantees

> Processing/CPU: The management stack VMs need to be configured with the highest possible CPU priority as well as with a reserve of 100% to increase the resource availability for the management VMs

> Memory reserves: The host cluster capacity needs to be defined so that the management stack will have enough memory resources to function even after loss of a cluster node. To improve the probability of the management VMs to be able to restart after failure the VMs should use the highest possible memory priority as well as the dynamic memory feature provided through Windows Server 2008 R2 SP1

3.3 Server Architecture

The host server architecture is a critical component of the virtualized infrastructure, as well as a key variable in the consolidation ratio and cost analysis. The ability of the host server to handle the workload of a large number of consolidation candidates increases the consolidation ratio and helps provide the desired cost benefit.

The system architecture of the host server refers to the general category of the server hardware itself. Examples include rack mounted servers, blade servers, and large symmetric multiprocessor servers (SMP). The primary tenet to consider when selecting system architectures is that each Hyper-V host will contain multiple guests with multiple workloads. Processor, RAM, Storage, and Network capacity are critical, as well as high I/O capacity and low latency. It is critical to ensure that the host server is able to provide the required capacity in each of these categories.

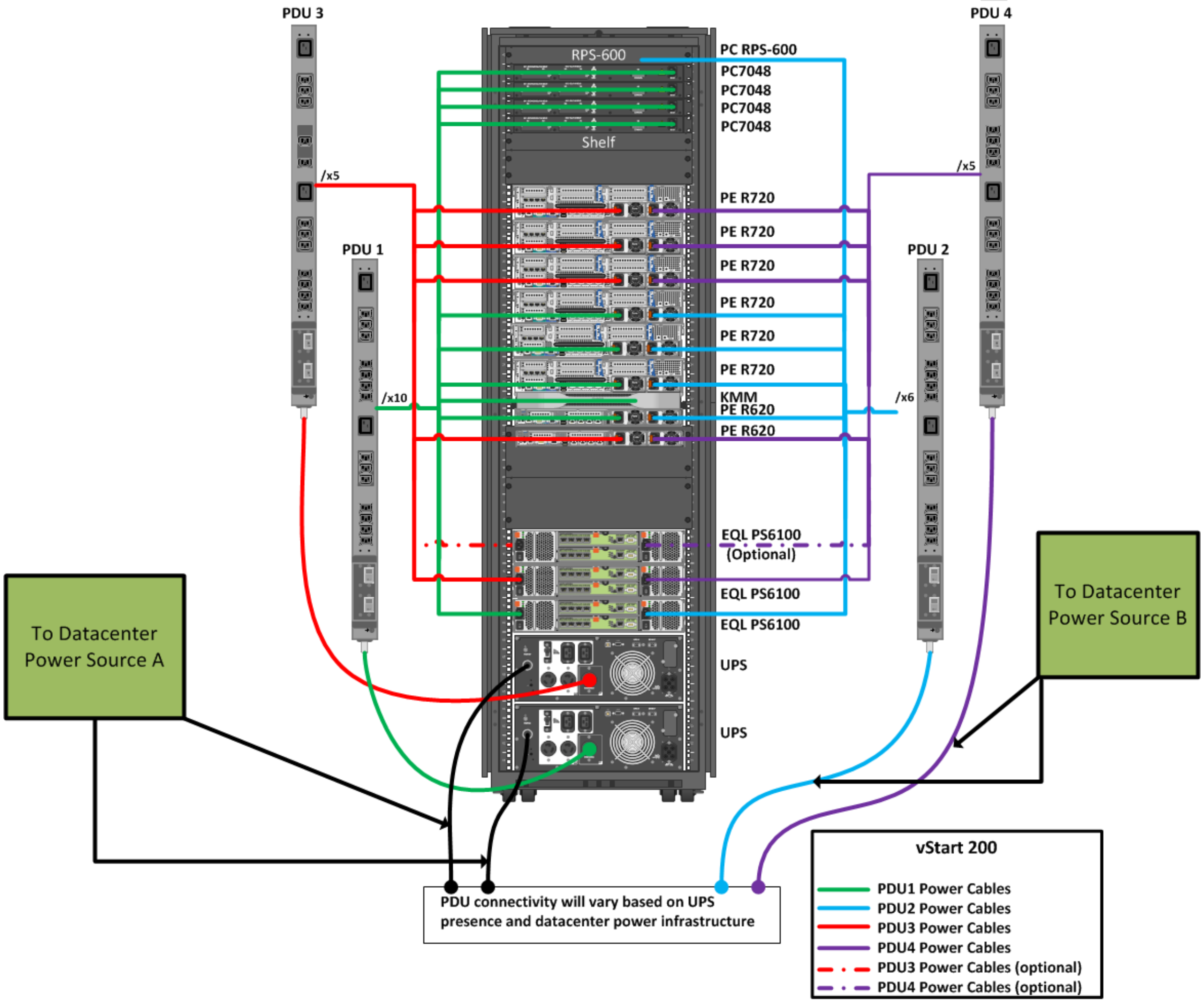
Note: There is a program available for assisting customers in selecting appropriate hardware: Windows Server Catalog contains all servers, storage, and other hardware devices that are certified for Windows Server 2008 R2 (including SP1) and Hyper-V. However, while the Windows Server Catalog lists all the positive tested hardware, it will not help to address sizing or configuration of hardware.

Windows Server Catalog: Go to www.windowsservercatalog.com. Click **Certified Servers.**

3.3.1 Rack Design

The Dell vStart 200 rack configuration includes four Power Distribution Units (PDU) and two Uninterruptable Power Supplies (UPS). The vStart PDUs are designed to be connected with two different datacenter power sources. The PDUs on the left side are connected to the UPS in the rack, while the PDUs on the right side are connected to another datacenter power source. This design ensures that there is no single point of failure in the power architecture. The redundant architecture is illustrated below.

Figure 4 . vStart 200 12G Power Cabling



Mandatory: > The Rack or Blade chassis design must provide redundant power connectivity (multiple PDU capability for racks or multiple hot-swappable power supplied for blade chassis)

3.3.2 Server Design

The cloud solution utilizes different model servers for the host and the management cluster. The Dell PowerEdge R720 servers comprise the Host Cluster while the R620 servers are serving as the Management Cluster hosts.

The Dell PowerEdge R720 server is a 2-socket, 2U rack server designed to excel at running a wide range of applications and virtualization environments for both mid-size and large enterprises. It has highly expandable memory and impressive I/O capabilities. There are six PowerEdge R720 servers included in the vStart configuration. Each is configured with two Intel® Xeon E5-2660 2.2GHz 8-core processors and 96GB memory. Each of the R720 servers also includes a PERC H710 RAID controller along with two 146GB 15K RPM SAS hard drives configured in a RAID 1 for the local storage.

The R620 is a 1U rack server designed with a dual-socket and multi-core processor architecture, a dense memory configuration, and redundant local drives configurable in a RAID. The vStart configuration for Hyper-V Private Cloud Fast Track requirement includes two PowerEdge R620 servers. Each of the R620s includes one Intel® Xeon® E5- E2609 2.4GHz 4-core Processors and 24GB memory. They also each include a PERC H710 RAID controller configured as RAID 1.

Mandatory: > Server with 2 Processor Sockets or more, max of 64 logical processors enabled
> 64-bit CPU with AMD Virtualization or Intel Virtualization Technology support
Minimum of 64 GB RAM
> Min 40 GB local RAID 1 or 10 hard disk space for OS partition
For more information see **Installing Windows Server 2008 R2**
http://technet.microsoft.com/en-us/library/dd379511(WS.10).aspx

Recommended: > Use Processors with support for Second Level Address Translation (SLAT) which is also known as EPT on Intel and NPT/RVI on AMD processors.

Table 1. Dell vStart Server Configurations

Component	Details
Compute Server Configuration	
Server Model	PowerEdge R720
Processor	(2) x Intel Xeon E5-2660, 2.2Ghz, 8-core, 20M Cache, Turbo, HT
Memory	96 GB (12 x 8 GB, DDR3 dual rank DIMMs, 1333MHz or 1600MHz)
Local storage and controller	(1) x PERC H710 Integrated mini RAID controller (2) x 146GB 15K RPM SAS drives configured in a RAID 1
Management Server Configuration	
Server Model	PowerEdge R620

Processor	(1) x Intel Xeon E5-2609, 2.4GHz, 4-core processor, HT
Memory	24 GB (8 x 4 GB, DDR3 dual rank DIMMs, 1333MHz)
Local storage and controller	(1) x PERC H710 Integrated mini RAID controller (2) x 146GB 15K RPM SAS drives configured in a RAID 1

3.3.3 Server Storage Connectivity

Storage connectivity from the PowerEdge R720 Servers are through four 1Gb NIC ports including two ports provided by the Broadcom BCM5720 rack Network Daughter Card (rNDC) and two ports provided by the add-in Broadcom 5719 PCIe card. Storage connectivity from PowerEdge R620 are through two 1Gb Broadcom BCM5720 rNDC NIC ports. Both R720 and R620 storage connections are distributed to two dedicated PowerConnect switches. These NICs are all dedicated to iSCSI traffic and further segmented logically using Virtual LANs (VLAN) on the PowerConnect switches. NIC connectivity can be seen in the diagrams below.

Figure 5. PowerEdge R620 SAN Connectivity

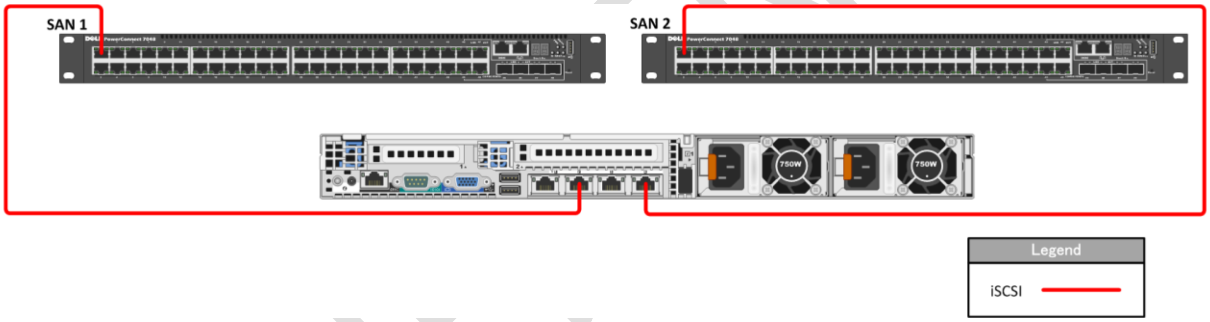
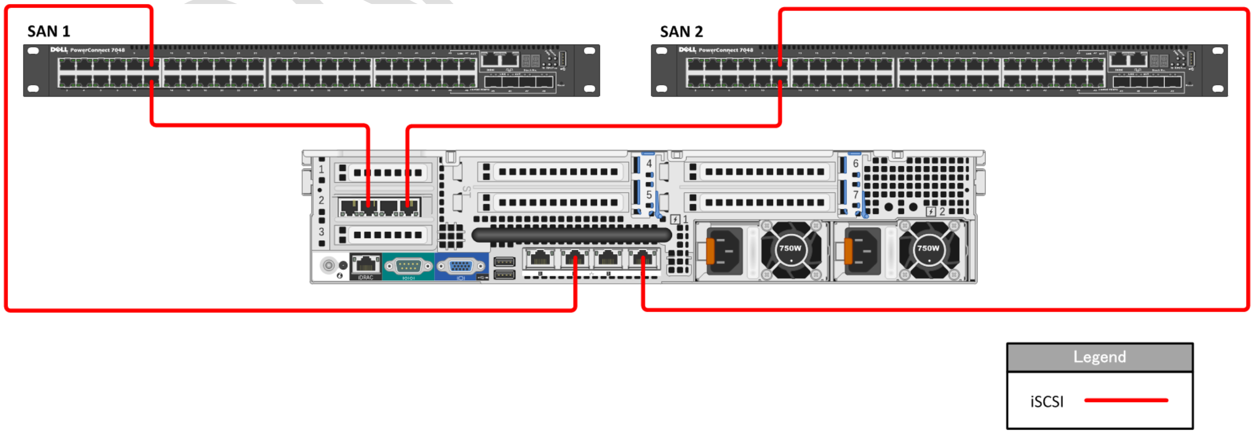


Figure 6. PowerEdge R720 SAN Connectivity



Mandatory: > Internal SATA or SAS controller for direct attached storage unless design is 100% SAN-based including boot from SAN for the Host OS

> If using Fiber Channel SAN, two or more 4 to 8 GFC HBAs

> If using iSCSI, two or more 10 Gb-Ethernet NICs or HBAs

> If using FCOE, two or more 10 Gb-CNAs

3.3.4 Server Network Connectivity

For Network or LAN connectivity, the PowerEdge R720 servers use four 1Gb Ethernet ports including two Broadcom BCM5720 (rNDC) NIC ports and two Broadcom 5719 (Add-in PCIe) NIC ports, while the R620 servers use two 1Gb Broadcom BCM5720 rNDC NIC ports. These connections are distributed to two PowerConnect switches that are configured in a stack. The ports on both the R720 and R620 servers are configured to use VLANs to segregate traffic on the host and provide the segmentation necessary for Hyper-V management, Live Migration (LM), cluster private, VM, and other traffics as described in Table 2. The VLAN configuration used in the vStart configuration is listed in Table 3. Network connectivity for the R620 and R720 are also illustrated below in Figure 7 and 8.

Table 2. Traffic Description

Traffic Type	Use
Hypervisor Management	Supports virtualization management traffic and communication between the host servers in the cluster.
Live Migration	Supports migration of VMs between the host servers in the cluster.
VM	Supports communication between the VMs hosted on the cluster and external systems.
Cluster Private	Supports internal cluster network communication between the servers in the cluster.
Out-of-Band Management	Supports configuration and monitoring of the servers through the iDRAC management interface, storage arrays, and network switches.
iSCSI Data	Supports iSCSI traffic between the servers and storage array(s). In addition, traffic between the arrays is supported.
Management VM	Supports the virtual machine traffic for the management virtual machines.
SQL Cluster Private	Supports private cluster traffic for the SQL Cluster of the management databases.

Table 3. Sample VLAN and subnet configuration

Traffic Type	Sample VLAN	Sample Subnet
Out-of-Band Management	10	192.168.10.X /24
Management	20	192.168.20.X /24
Live Migration	30	192.168.30.X /24
Cluster Private	40	192.168.40.X /24
Management VM	60	192.168.60.X /24
SQL Cluster Private	70	192.168.70.X /24
VM	100	192.168.100.X /24

Figure 7. PowerEdge R620 LAN Connectivity

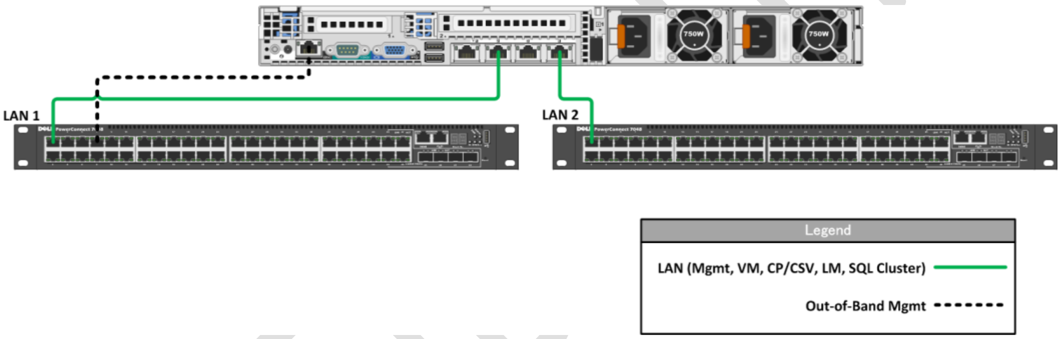
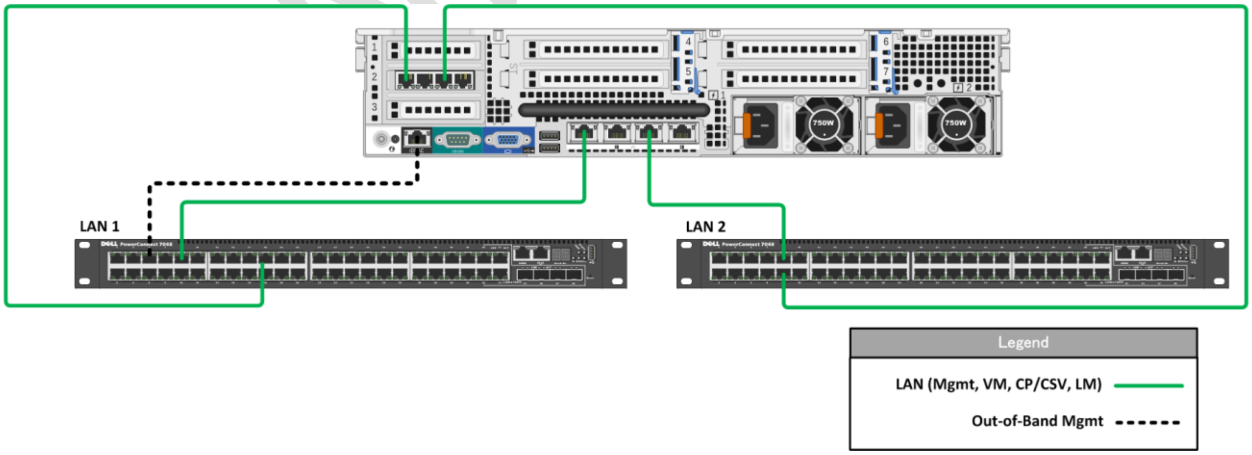


Figure 8. PowerEdge R720 LAN Connectivity



Use multiple network adapters and/or multi-port network adapters on each host server. For converged designs, network technologies that provide teaming or virtual NICs can be utilized provided that two or more physical adapters can be teamed for redundancy and multiple virtual NICs and/or VLANs can be presented to the hosts for traffic segmentation and bandwidth control. The following networks are required:

- One network dedicated to the host machine only for management purposes
- One network dedicated to the CSV/Cluster Communication network
- One network dedicated to the Live Migration network
- One or more networks dedicated to the guest virtual machines (use 10 Gbps network adapters for highest consolidation)
- If using iSCSI, one network dedicated to iSCSI with Multipath I/O (MPIO)

The following reference documents the recommended configuration by quantity and type of NIC: **Hyper-V: Live Migration Network Configuration Guide** http://technet.microsoft.com/en-us/library/ff428137(WS.10).aspx

Mandatory: > If using a 10Gb-Ethernet network backbone, each host must have two or more 10Gb-E NICs and the ability to present multiple teamed and/or virtual NICs to the Windows OS.

> If using a 1Gb-Ethernet network backbone, each host must have five 1Gb-Ethernet NICs (1 for Mgmt, 1 for CSV, 1 for LM, 2 for VM traffic).

> If using a 1Gb-Ethrnet network backbone and iSCSI storage, each host most have two additional 1Gb-Ethernet NICs for a minimum total of seven.

> For more configuration information see **Hyper-V: Live Migration Network Configuration Guide** http://technet.microsoft.com/en-us/library/ff428137(WS.10).aspx.

3.3.5 Server HA and Redundancy

The design of the PowerEdge R620 and R720 servers chosen for the vStart configuration include high availability and redundant features such as redundant fans and power supplies that are distributed to independent power sources. The servers also include RAID configurations to prevent server crashes in the event of single disk failures.

Mandatory: > If using rack mounted servers, each server must have redundant power supplies

> If using rack mounted servers, each server must have redundant fans

> If using blade servers, each chassis must have redundant power supplies

> If using blade servers, each chassis must have redundant fans

> If Hyper-V Host system partition uses direct attached storage, each server must provide a SAS or SATA RAID capability

3.4 Storage Architecture

The storage design for any virtualization-based solution is a critical element which is typically responsible for a large percentage of the solution's overall cost, performance, and agility.

3.4.1 Storage Options

While many storage options exist, there is product-range sweet-spot for datacenter virtualization. These devices are typically modular and flexible mid and high-end SANs. Modular mid-range SANs are procured independently and can be chained together to provide large capacity and high performance. They are efficient and can grow with the environment as needed but require less up-front investment. Large enterprise environments may have more storage demands and need to serve a larger set of customers and workloads. In this case high-end SANs can provide the highest performance and capacity and typically enable more advanced features such as continuous data availability through technologies like metropolitan-area clustering.

3.4.2 SAN Storage Protocols

(1) iSCSI vs. FC vs. FCoE

Fiber Channel has historically been the storage protocol of choice for enterprise datacenters for a variety of reasons, including performance and low latency. These considerations have offset Fiber Channel's typically higher costs. In the last several years, Ethernet's continually advancing performance from 1 GB/s to 10 GB/s and eventually beyond have led to great interest in storage protocols leveraging the Ethernet transport such as iSCSI and recently, Fiber Channel over Ethernet (FCoE).

A key advantage of the protocols leveraging the Ethernet transport is the ability to use a "converged" network architecture where a single Ethernet infrastructure serves as the transport for both local-area network (LAN) and storage traffic. This can reduce costs in several ways such as the elimination of dedicated Fiber Channel switches and a reduction in cabling which can also be a significant cost in large datacenter environments.

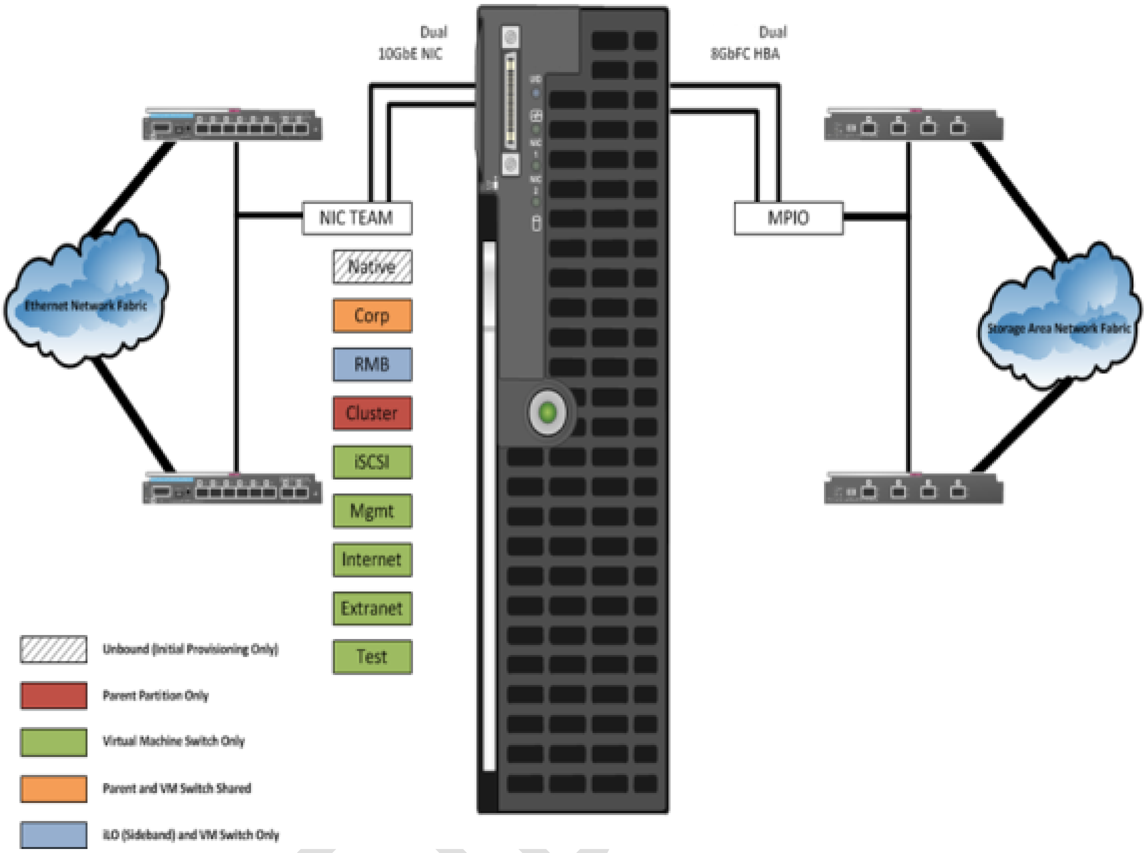
Fiber Channel over Ethernet (FCoE) is an emerging technology, now standardized, which brings the benefits of leveraging an Ethernet transport while retaining the advantages of the Fiber Channel protocol and the ability to leverage Fiber Channel storage arrays.

Several enhancements to standard Ethernet are required for FCoE. This is commonly referred to as Enhanced Ethernet or Data Center Ethernet. These enhancements require Ethernet switches capable of supporting enhanced Ethernet.

For Hyper-V, iSCSI-capable storage provides an advantage in that it is the protocol to be utilized by Hyper-V guest virtual machines for guest clustering.

A common practice in large-scale Virtualization deployments is to utilize both Fiber and iSCSI. Fiber provides the Host storage connectivity, and iSCSI is used only by guests that require in-OS iSCSI connectivity such as a guest cluster. In this case, although Ethernet and some Storage I/O will be sharing the same pipe, segregation is achieved by VLANs and QoS can be further applied by the OEM's networking software.

Figure 9. Example: Blade Server Host Design

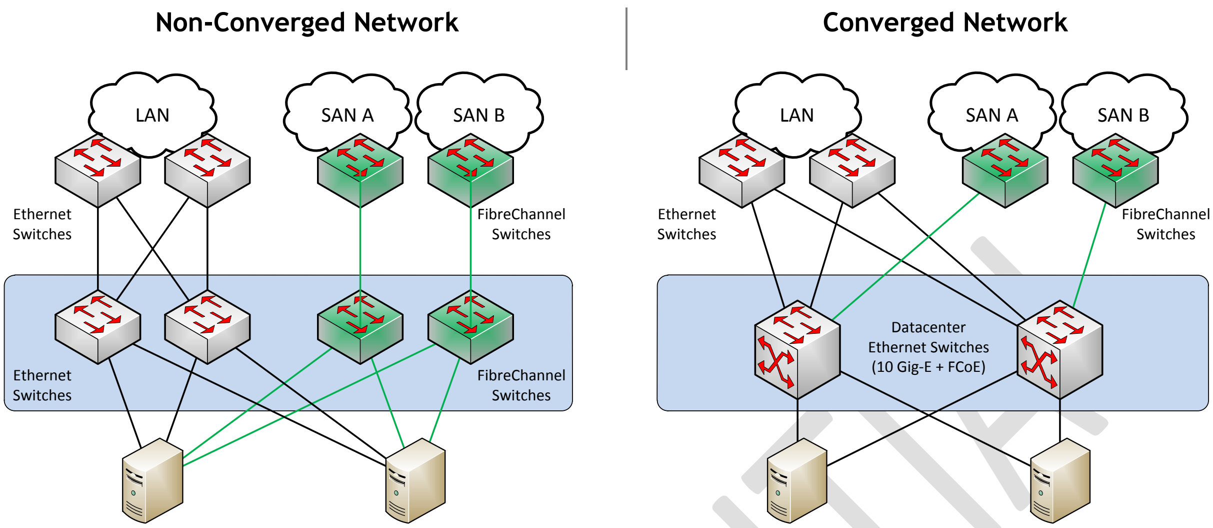

Dual
10GbE NIC
Dual
8GbFC HBA
NIC TEAM
MPIO
Ethernet Network Fabric
Storage Area Network Fabric
Native
Corp
RMB
Cluster
iSCSI
Mgmt
Internet
Extranet
Test
Unbound (Initial Provisioning Only)
Parent Partition Only
Virtual Machine Switch Only
Parent and VM Switch Shared
iLO (Sideband) and VM Switch Only

(2) Storage Network

Both iSCSI and FCoE utilize an Ethernet transport for storage networking. This provides another architecture choice in terms of whether to utilize a dedicated Ethernet network with separate switches, cables, paths etc. or whether to leverage a “converged” network where multiple traffic types are run over the same cabling and infrastructure.

The diagram below illustrates the differences between a traditional architecture on the left with separate Ethernet and Fiber Channel switches, each with redundant paths compared to a converged architecture where both Ethernet and Fiber Channel (via FCoE) utilize the same set of cables while still providing redundant paths. The converged architecture requires fewer switches and cables. In the converged architecture, the switches must be capable of supporting enhanced Ethernet.

Figure 10. Converged Network vs. Non-Converged Network



Mandatory: > Storage solution must provide logical or physical isolation between storage and Ethernet I/O

> If a converged network, QoS must be provided to guarantee storage performance

> Storage solution must provide iSCSI connectivity for guest clustering

> There must be fully redundant, independent paths for storage I/O

Recommended: > For FCoE, utilize standards-based converged network adapters, switches, and Fiber Channel storage arrays. Ensure that the selected storage arrays also provide iSCSI connectivity over standard Ethernet so that Hyper-V guest clusters can be utilized.

> If using iSCSI or Fiber Channel, ensure that there are dedicated network adapters/HBAs, switches, and paths for the storage traffic.

3.4.3 Clustered File Systems (3rd Party)

The choice of file system to run on top of the storage architecture is another critical design factor. While not strictly required to support Live Migration and other advanced features, use of a clustered file systems or Cluster Shared Volumes (CSV) as part of Windows Server 2008 R2 can provide significant manageability benefits. CSV and clustered file systems enable the use of larger LUNs (logical unit numbers) to store multiple virtual machines while providing the ability for each virtual machine to be live migrated independently. This is enabled by providing all nodes the ability to read and write from the shared LUN at the same time.

The decision of using CSV or a 3rd party solution that is compatible with Hyper-V and Windows Failover Clustering should be made by carefully weighting the advantages and disadvantages of one vs. the other vs. the actual environment requirements.

3.4.4 Cluster Shared Volumes

Windows Server 2008 R2 includes the first version of Windows Failover Clustering to offer a distributed file access solution. Clustered Shared Volumes (CSV) in R2 is exclusively for use with the Hyper-V role and enables all nodes in the cluster to access the same cluster storage volumes at the same time. This enhancement eliminates the 1 VM per LUN requirement of previous Hyper-V versions without using a 3rd party filesystem. CSV uses standard NTFS and has no special hardware requirements, from a functional standpoint if the storage is suitable for Failover Clustering, it is suitable for CSV.

CSV provides not only shared access to the disk, but also storage path I/O fault tolerance (*dynamic I/O redirection*). In the event the storage path on one node becomes unavailable, the I/O for that node will be rerouted via Server Message Block (SMB) through another node. There is a performance impact while running this state; it is designed for use as a temporary failover path while the primary dedicated storage path is brought back online. This feature can use any Cluster Communications Network and further increases the need for high-speed networks.

CSV maintains metadata information about the volume access and requires that some I/O operations take place over the cluster communications network. One node in the cluster is designated as the coordinator node and is responsible for these disk operations. Virtual Machines, however, have direct I/O access to the volumes and only use the dedicated storage paths for disk I/O, unless a failure scenario occurs as described above.

(1) CSV Limits

The below limitations are actually imposed by the NTFS file system and are inherited by CSV.

Table 4. CSV Limitations

CSV parameter	Limitation
Maximum Volume Size	256 TB
Maximum # Partitions	128
Directory Structure	Unrestricted
Maximum Files per CSV	4+ Billion
Maximum VMs per CSV	Unlimited

(2) CSV Requirements

- All cluster nodes must use Windows Server 2008 R2
- All cluster nodes must use the same drive letter for the system disk
- All cluster nodes must be on the same logical network subnet. Virtual LANs (VLANs) are required for multi-site clusters running CSV

- NT LAN Manager (NTLM) must be enabled on cluster nodes
- SMB must be enabled for each network on each node that will carry CSV cluster communications
- “Client for Microsoft Networks” and “File and Printer Sharing for Microsoft Networks” must be enabled in the network adapter’s properties to enable all nodes in the cluster to communicate with the CSV.
- The Hyper-V role must be installed on any cluster node that may host a VM

(3) CSV Volume Sizing

Because all cluster nodes can access all CSV volumes simultaneously, we can now use standard LUN allocation methodologies based on performance and capacity requirements of the expected workloads. Generally speaking, isolating the VM Operating System I/O from the application data I/O is a good start, in addition to application-specific considerations such as segregating database from logging I/O and creating SAN volumes and/or Storage Pools that factor in the I/O profile itself (i.e., random read and write operations vs. sequential write operations).

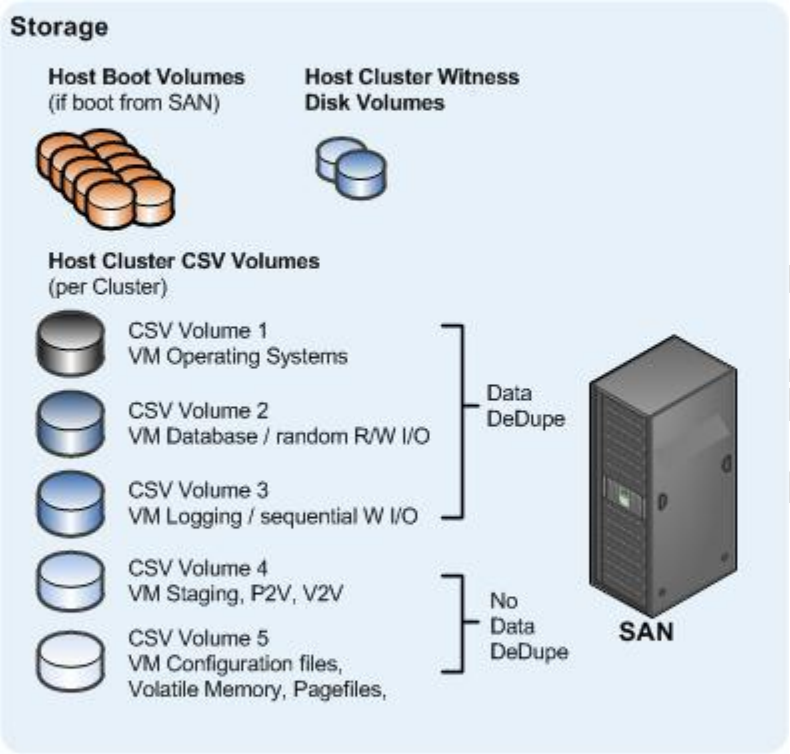
CSV’s architecture differs from other traditional clustered file systems which frees it from common scalability limitations. As a result, there is no special guidance for scaling the number of Hyper-V Nodes or VMs on a CSV volume. The important thing to keep in mind is that all VM’s virtual disks running on a particular CSV will contend for storage I/O.

Also worth noting is that individual SAN LUNs do not necessarily equate to dedicated disk spindles. A SAN Storage Pool or RAID Array may contain many LUNs. A LUN is simply a logic representation of a disk provisioned from a pool of disks. Therefore, if an enterprise application requires specific storage IOPS or disk response times you must consider all the LUNs in use on that Storage Pool. An application which would require dedicated physical disks where not virtualized may require dedicate Storage Pools and CSV volumes running within a VM.

Recommended:

> For maximum flexibility, configure LUNs for CSV with a single volume so that 1 LUN equals 1 CSV
> At least 4 CSVs per Host Cluster are recommended for segregating Operating System I/O, Random R/W I/O, Sequential I/O, and other VM-specific data
> Follow the storage vendor’s recommendations for use with CSVs
> Create a standard size and IOPS profile for each type of CSV LUN to utilize for capacity planning. When additional capacity is needed, provision additional standard CSV LUNs.
> Consider prioritizing the network used for CSV traffic:
Designating a Preferred Network for Cluster Shared Volumes Communication
http://technet.microsoft.com/en-us/library/ff182335(WS.10).aspx

Figure 11. Example: Common CSV design for large Hyper-V Cluster



3.4.5 SAN Design

The vStart configuration uses the iSCSI protocol for the SAN storage fabric. The network supporting the SAN is a 1 Gb Ethernet network with redundancy throughout. Each host in the configuration connects to the SAN though multiple 1 Gb Ethernet ports. The EqualLogic PS6100 arrays used in the vStart provide a high performance and high capacity, cost-effective SAN solution while PowerConnect 1 Gb Ethernet switches are used for SAN connectivity. Details of the components and an overview of the SAN are shown below in Table 5 and Figure 12.

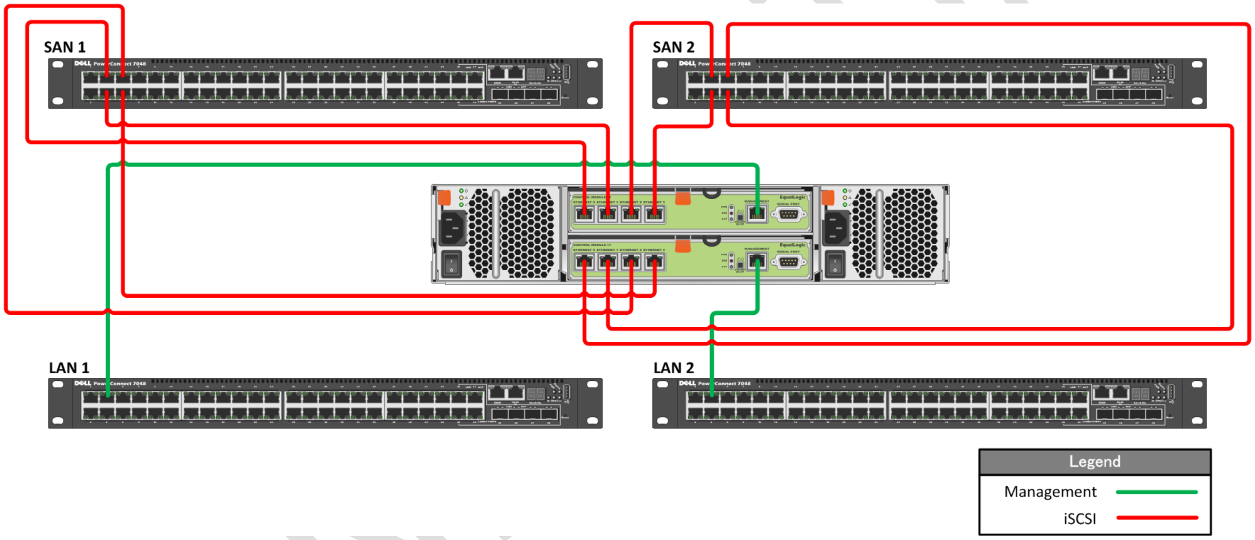
The EqualLogic SAN includes features such as automatic tiering, thin provisioning and a scriptable interface for programmatic control of the arrays. EqualLogic also actively manages iSCSI connections through the EqualLogic Host Integration Tools (HIT) and manages connections to the SAN based upon where data exists on the system.

Table 5. SAN Configurations

Component	Details
Storage Configuration	
Storage Device	EqualLogic PS6100
Drives	• (2) or (3) x PS6100 24 drive Enclosures • (48) or (72) 600 GB 10K RPM SAS Drives
RAW Storage Capacity	~28.8TB or ~43.2 TB

RAID Configuration	RAID 5, 6, 10 or 50
Controller	Dual Controllers per Enclosure, each with 4GB memory for PS Series firmware. Each controller provides cache to flash for data protection. (4) 1 Gb Ethernet Ports per controller and (1) dedicated 100Mb Ethernet port for management.
Network Switches	PowerConnect 7048 or 6248
Network Redundancy	Microsoft MPIO
Device Specific Module	EqualLogic Host Integration Tools (HIT)
Host Interface (R720)	(2) Broadcom BCM5720 and (2) BCM5719 1GbE ports
Host Interface (R620)	(2) Broadcom BCM5720 1GbE Ports

Figure 12. EqualLogic PS6100 Connectivity



(1) High Availability

The SAN's high availability is taken into consideration from multiple standpoints depending on the device.

The PS6100 includes core high availability features such as redundant fans, controllers, network ports, and power supplies. It also includes hot-swappable components such as controllers, power supplies and hard drives. The controllers are configurable to provide volume high availability through the use of RAID. Each PS6100 is connected to the SAN via ten 1 Gb network connections. Half of each array's connections are in a standby state and are bound to the inactive controller. Of the active connections, four are dedicated to SAN traffic and one is dedicated to management.

The PowerConnect network switches also use redundant fans and use an optional external Redundant Power Supply (RPS) for redundant power. The PowerConnect switches are further configured with a

LAG for the Inter Switch Link (ISL) resulting in increased bandwidth and link high availability between switches.

Mandatory: > Describe the SANs approach to High Availability
> Diagram needed
> A highly-available SAN design should have no single points of failure including:
- redundant power from independent PDUs, redundant storage controllers,
- redundant target ports of NICs per controller, redundant FC or IP network switches, etc.
- data storage redundancy such as with volume mirroring, synchronous or asynchronous replication

(2) Performance

Storage performance is a complicated mix of drive, interface, controller, cache, protocol, SAN, HBA, driver, and operating system considerations. The overall performance of the storage architecture is typically measured in terms of Maximum Throughput, Maximum IO operations per second (IOPS), and Latency or Response Time. While each of the three factors is important, IOPS and Latency are the most relevant to server virtualization.

Most modern SANs use a combination of high-speed disks, slower-speed disks, and large memory caches. Storage controller cache can improve performance during burst transfers or when the same data is accessed frequently by storing it in the cache memory and can be several orders of magnitude faster than the physical disk I/O. However, it is not a substitute for adequate disk spindles because caches are ineffective at aiding in heavy Write operations.

The EqualLogic family of PS6100 arrays provide a high performance iSCSI SAN through features such as high speed drives, volumes that span multiple arrays thereby improving IOPS as a result of additional spindle count, and a linear scaling peer architecture where performance will scale as a direct relationship to the number of arrays in a group.

Mandatory: > Describe the SAN's performance characteristics

(3) Drive Types

The type of hard drive utilized in the host server or the storage array the host servers will have the most significant impact on the overall storage architecture performance. The critical performance factors for hard disks are the interface architecture (for example, U320 SCSI, SAS, SATA), the rotational speed of the drive (7200, 10k, 15k RPM), and the average latency in milliseconds (ms). Additional factors, such as the cache on the drive, and support for advanced features, such as Native Command Queuing (NCQ), can improve performance. As with the storage connectivity, high IOPS and low latency are more critical than maximum sustained throughput when it comes to host server sizing and guest performance. When selecting drives, this translates into selecting those with the highest rotational speed and lowest latency possible. Utilizing 15k RPM drives over 10k RPM drives can result in up to 35% more IOPS per drive.

The base EqualLogic PS6100 array used in the vStart configuration uses 600GB 10K RPM SAS drives. The EqualLogic PS6100 family is also capable of supporting SSD, 15K RPM SAS, and 7K RPM SATA drives in a variety of sizes, however those drives are not part of the base solution.

Mandatory: > Describe the SAN's supported and recommended drive types

(4) RAID Array Design

The RAID type should provide both High Availability and high performance even in the event of disk failures and RAID parity rebuilds. In general, RAID 10 (0+1), or a proprietary hybrid RAID type (i.e. NetApp RAID DP) are recommended for Virtual Machine volumes, RAID 1 is also acceptable for Host Boot Volumes.

The RAID level for EqualLogic PS6100 in the vStart configuration can be configured per customer requirement. The PS6100 is capable of a RAID 5, 6, 10, or 50. RAID 6 provides the greatest amount of disk space while allowing for the failure of up to 4 drives (including hot spares).

Mandatory: > Describe the SAN's RAID Array Design

(5) Multi-Pathing

In all cases, multipathing should be used. Generally storage vendors will build a DSM (device-specific module) on top of Microsoft's Windows Server 2008 R2 MPIO software. Each DSM and HBA will have its own unique multipathing options, recommended number of connections, etc.

Each host in the configuration uses the Windows Server 2008 R2 MPIO feature and is combined with the EqualLogic Host Integration Toolkit (HIT) to provide the Device Specific Module (DSM) extension into the MPIO module. The HIT makes intelligent decisions to connect each host to the most appropriate iSCSI target from all available targets. The default and preferred MPIO Load Balance Policy is Least Queue Depth.

Mandatory: > Describe the SAN's Multi-pathing solution

(6) Fiber (if FC is used)

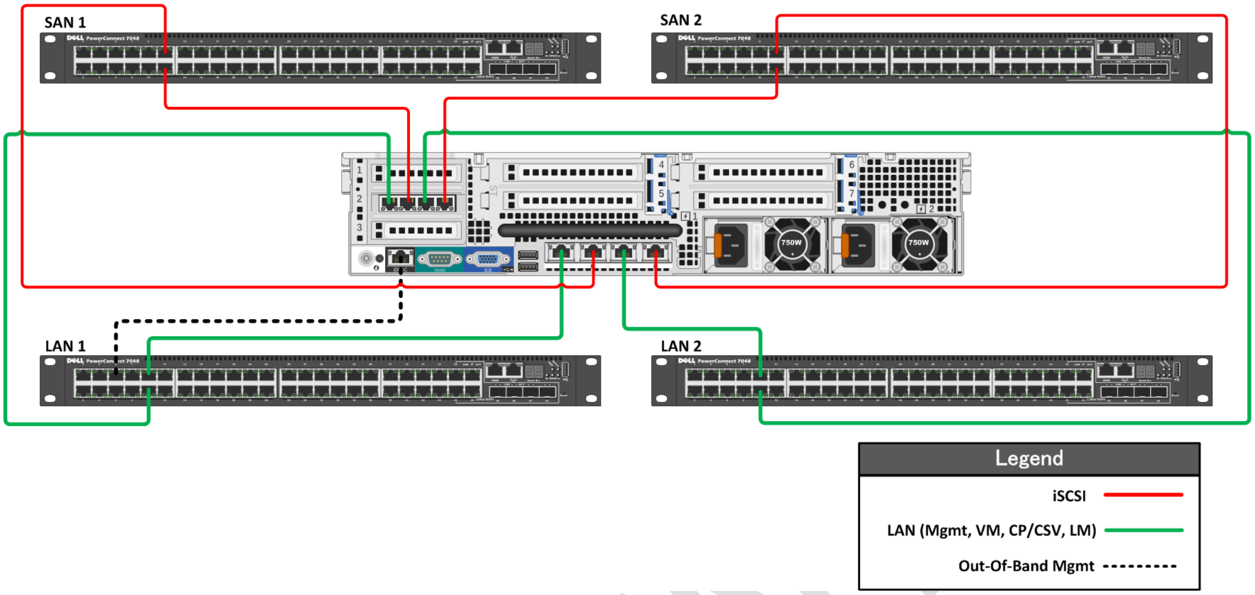
(a) Zoning, Masking, NPIV

Mandatory: > Describe the SAN's Fiber networking, zoning and masking configuration

(7) iSCSI

The vStart configuration uses a dedicated 1GbE iSCSI SAN network with dedicated host network ports and network switches as shown in Figure 13. The SAN network is further isolated through the use of VLANs as noted earlier in Section 3.3.3.

Figure 13. PowerEdge R720 Network Connectivity



Mandatory:

> The iSCSI SAN must be on an isolated network, both for security and performance. Any networking standard practice method for achieving this end is acceptable, including:

> A physically separate, dedicated storage network.

> A physically shared network with the iSCSI SAN running on a private virtual local area network (VLAN). The switch hardware must provide Class or Service (CoS) or Qualify of Service (QoS) guarantees for the private VLAN

(a) Encryption and Authentication

Mandatory:

> If multiple clusters and/or systems are used on the same SAN, proper segregation or device isolation must be provided. In other words, the storage used by cluster A must be visible only to cluster A, and not to any other cluster, nor to a node from a different cluster.

> The use of session authentication (Challenge Handshake Authentication Protocol (CHAP) minimum) is suggested. This provides a degree of security as well as segregation.

> Mutual CHAP or Internet Protocol Security (IPSec) can also be used.

(b) Jumbo Frames

If supported at all points in the iSCSI network, Jumbo Frames can increase throughput by up to 20%. Jumbo frames are supported in Hyper-V at the Host and Guest levels.

In the vStart configuration, Jumbo Frames is enabled for all devices of the SAN fabric. This includes the server network interface ports, the network switch interfaces and the EqualLogic interfaces. The EqualLogic ports are configured to use Jumbo Frames by default however the PowerEdge and PowerConnect devices are not. The PowerEdge server SAN network ports are set to 9000 byte frames by using the Broadcom BACS utility while the PowerConnect switches interfaces are configured for 9216 byte frames by using the command line interface of the switch.

⚠ **Recommended:** > Describe the SAN's implementation of Jumbo Frames

(8) Data De-duplication

Data De-duplication can yield significant storage cost savings in Virtualization environments. Some common considerations are performance hits during the de-duplication cycle, and achieving maximum efficiency by locating similar data types on the same volume, LUN, etc.

No Data-Dedupe is used in the Dell vSTART configuration.

⚠ **Recommended:** > Describe the SAN's use of Data De-duplication and related settings

(9) Thin Provisioning

Particularly in Virtualization environments, thin provisioning is a common practice. This allows for efficient use of the available storage capacity. The LUN and corresponding CSV may grow as needed, typically in an automated fashion to ensure availability of the LUN (auto-grow). However, as storage becomes over-provisioned in this scenario very careful management and capacity planning is critical.

Thin provisioning is a feature of the EqualLogic PS6100 array that allows a volume to be presented to the server that is greater than the capacity on disk. The EqualLogic thin provisioning will consume approximately 10% of the stated size and the space will be consumed as data is written to disk. The EqualLogic will alert the administrator via SNMP trap or email when the free space of the volume is less than 60%; however this is a configurable parameter of the EqualLogic group.

⚠ **Recommended:** > Describe the SAN's use of Thin Provisioning and related settings

(10) Volume Cloning

Volume Cloning is another common practice in Virtualization environments. This can be used for both Host and VM volumes dramatically increasing Host installation times and VM provisioning times.

The EqualLogic PS6100 provides volume cloning capabilities to allow volumes to be duplicated bit by bit. An additional feature of the PS6100 is thin cloning where thin clone volumes are matched with template volume. In the PS Series array architecture, template volumes act as a read-only gold image copy of a volume from which writeable thin clones can be created. The thin clones initially consume little to no storage space on the array and as data is written to the thin clone, only new writes to a thin clone are recorded to storage space thus reducing the amount of storage consumption.

⚠ **Recommended:** > Describe the SAN’s Cloning capabilities and recommendations

(11) Volume Snapshots

SAN Volume snapshots are a common method of providing a point-in-time, instantaneous backup of a SAN Volume or LUN. These snapshots are typically block-level and only utilize storage capacity as blocks change on the originating volume. Some SANs provide tight integration with Hyper-V integrating both the Hyper-V VSS Writer on Hosts and Volume Snapshots on the SAN. This integration provides a comprehensive and high-performing backup and recovery solution.

The EqualLogic PS6100 provides volume snapshots and is capable of automating this through the use of the EqualLogic Auto Snapshot Manager / Microsoft Edition (ASM/ME). The ASM/ME is a tool included with the EqualLogic HIT Kit that will allow administrators to take point-in-time snapshots of NTFS Volumes as well as Windows 2008 R2 VMs. ASM/ME is also integrated with the Microsoft Volume Shadow Copy Service (VSS) to provide application aware snapshots.

⚠ **Recommended:** > Describe the SAN’s Snapshotting capabilities and Hyper-V integration

(12) Storage Tiering

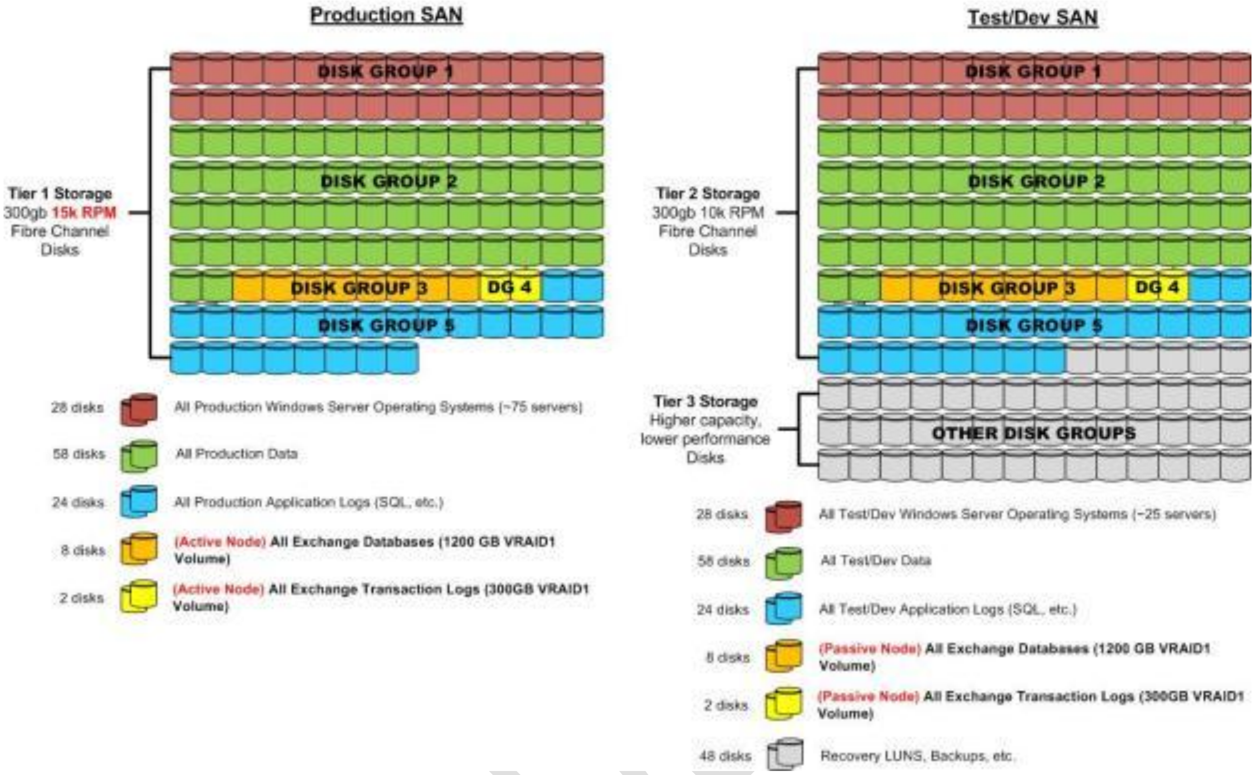
Tiering storage is the practice of physically partitioning data into multiple distinct classes based on price, performance or other attributes. Data may be dynamically moved among classes in a tiered storage implementation based on access activity or other considerations.

This is normally achieved through a combination of varying types of disks which are used for different data types. (i.e. Production, non-production, backups, etc.)

The EqualLogic tiering tools allow IT managers to determine where to apply PS Series tiering heuristics and to what degree. This allows certain volumes to be manually assigned to specific arrays, and others to be managed through the automated tiering of the PS Series SAN. The volumes of an EqualLogic PS Series SAN can be moved across pools as data access patterns change.

⚠ **Recommended:** > Describe the storage tiering strategy & recommendations

Figure 14. Example: Tiered Storage design


Production SAN
Test/Dev SAN
DISK GROUP 1
DISK GROUP 2
DISK GROUP 3
DG 4
DISK GROUP 5
OTHER DISK GROUPS
Tier 1 Storage
300gb 15k RPM
Fibre Channel
Disks
Tier 2 Storage
300gb 10k RPM
Fibre Channel
Disks
Tier 3 Storage
Higher capacity,
lower performance
Disks
28 disks All Production Windows Server Operating Systems (~75 servers)
58 disks All Production Data
24 disks All Production Application Logs (SQL, etc.)
8 disks (Active Node) All Exchange Databases (1200 GB VRAID1 Volume)
2 disks (Active Node) All Exchange Transaction Logs (300GB VRAID1 Volume)
28 disks All Test/Dev Windows Server Operating Systems (~25 servers)
58 disks All Test/Dev Data
24 disks All Test/Dev Application Logs (SQL, etc.)
8 disks (Passive Node) All Exchange Databases (1200 GB VRAID1 Volume)
2 disks (Passive Node) All Exchange Transaction Logs (300GB VRAID1 Volume)
48 disks Recovery LUNS, Backups, etc.

3.4.6 Storage Automation

One of the objectives of the Hyper-V Cloud solution is to enable rapid provisioning and de-provisioning of virtual machines. Doing so at large scale requires tight integration with the storage architecture and robust automation. Provisioning a new virtual machine on an already existing LUN is a simple operation however provisioning a new CSV LUN, adding it to a host cluster, etc. are relatively complicated tasks that must be automated.

Historically, many storage vendors have designed and implemented their own storage management systems, APIs, and command line utilities. This has made it a challenge to leverage a common set of tools, scripts, etc. across heterogeneous storage solutions

For the robust automation that is required in an advanced datacenter virtualization solution preference is given to SANs supporting standard and common automation interfaces such as PowerShell.

Mandatory: > Describe the SAN's Automation interfaces, capabilities and recommendations

> The Storage Solution must provide mechanisms to achieve automated provisioning at a minimum, and ideally automation of all common administrative tasks

The EqualLogic HIT Kit provides a PowerShell cmdlet interface for administrators to manage one or many PS series arrays from a command line or through a scripted process. Through the PowerShell interface an administrator can perform tasks such as Group administration, volume administration, snapshotting, replication, and scheduling.

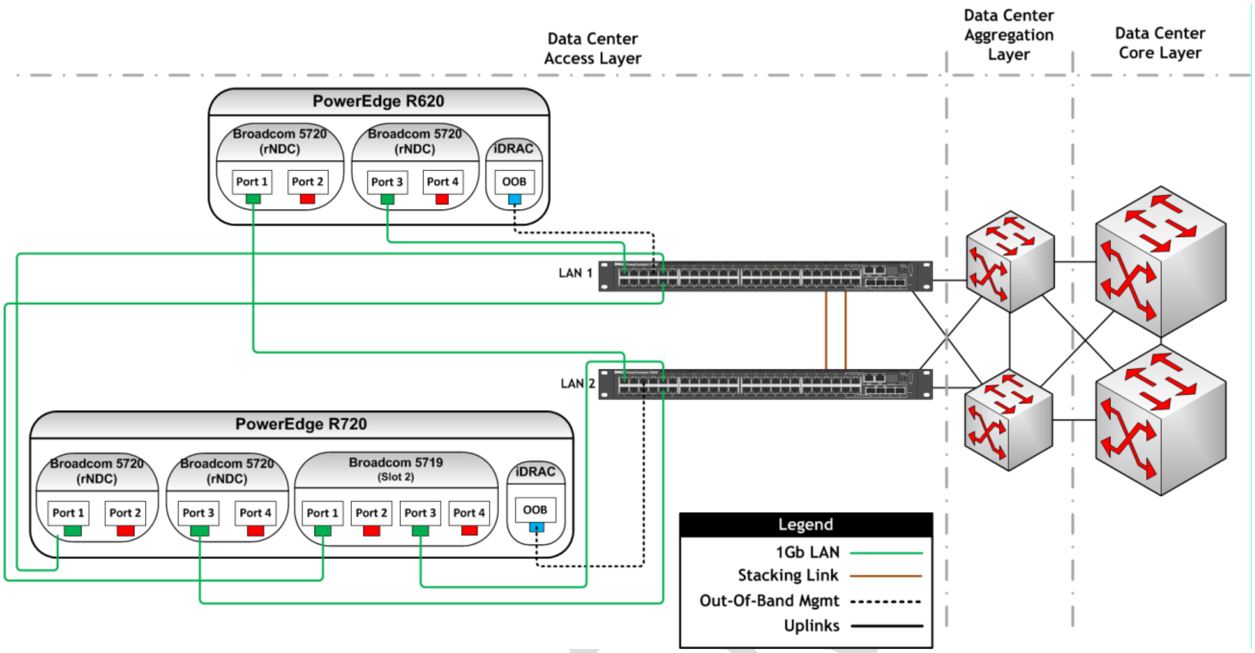
3.5 Network Architecture

3.5.1 Core, Distribution, and Access Network Design

Many network architectures include a tiered design with three or more tiers such as Core, Distribution, and Access. Designs are driven by the port bandwidth and quantity required at the edge, as well as the ability of the Distribution and Core tiers to provide higher speed uplinks to aggregate traffic. Additional considerations include Ethernet broadcast boundaries and limitations, spanning tree and or other loop avoidance technologies, etc.

The diagram in Figure 15 illustrates the network design of the vStart configuration and how its network design fits in to the Data Center Architecture.

Figure 15. vStart LAN Network Architecture



Note: The Aggregation and Core Layers are out of scope of this document.

Mandatory:

> The network switches must support 802.1q VLAN Trunks

> The network switches must support an Ethernet link aggregation standard compatible with the rack or blade server NICs such that NIC teams can span two or more switches.

> The network switches must support Ethernet link aggregation such that multiple uplink ports can be bonded together for high bandwidth.

Dell PowerConnect 7048 and 6048 in vStart 200 support IEEE 802.1q VLAN trunk and the Link Aggregation Control Protocol (LACP).

3.5.2 HA and Redundancy

The PowerConnect network switches used in the vStart LAN network configuration are stacked. Based upon the switch used (PowerConnect 7048 or 6248) the stacking modules provide a 64 or 48 Gbps back-end connection allowing traffic to flow between the two switches. These stacking modules have redundant connections to provide a fault-tolerant design.

The PowerConnect switches also have redundant uplinks to the core network that allows for a failure in either switch path to the core network so that connectivity will be maintained. Host connectivity is maintained during a switch failure through the teaming interfaces of each host.

The PowerConnect switches in the SAN network are connected with two 10Gb links using CX4 modules on each switch. When using the CX4 modules, additional configuration steps must be taken to set the default stacking mode into Ethernet mode, and then configure a port-channel inter switch link (ISL). The port-channel configuration allows individual link failure between the two switches such that SAN connectivity between switches is not impacted. The distribution of each host's NICs and EqualLogic's NICs between two switches provides high availability during a switch failure. If the SAN switches are uplinked to an existing SAN network for replication, spanning-tree link costs must be monitored and properly configured if needed to ensure that switch paths are optimized for local SAN traffic.

Mandatory: > The network design must allow for the loss of any switch module or switch without dropping host server connectivity.

3.6 Virtualization Architecture

3.6.1 Windows Server 2008 R2 and Hyper-V Host Design

The following recommendations in this section adhere to the support statements in the following article:

Requirements and Limits for Virtual Machines and Hyper-V in Windows Server 2008 R2 SP1
http://technet.microsoft.com/en-us/library/ee405267(WS.10).aspx

Licensing

Certain versions of Windows Server 2008 R2 (namely Standard, Enterprise, and Datacenter editions) include "virtualization use rights," which is the right and license to run a specified number of Windows-based virtual machines. Windows Server 2008 R2 Standard edition includes use rights for one running virtual machine. Windows Server 2008 R2 Enterprise Edition includes use rights for up to four virtual machines. This does not limit the number of guests that the host can run; it means that licenses for four Windows guests are included. To run more than four you simply need to ensure you have valid Windows Server licenses for the additional virtual machines.

Windows Server 2008 R2 Datacenter Edition includes unlimited virtualization use rights, which allows you to run as many guests as you like on the physical server running Windows Server 2008 R2 Datacenter edition.

Recommended: > Windows Server 2008 R2 SP1 Enterprise or Datacenter editions are both acceptable editions with Datacenter being preferred due to use-rights advantages.

(1) OS Configuration

The following outlines the general considerations for the Hyper-V Host Operating system. Note that these are not meant to be installation instructions but rather the process requirements and order.

- Use Windows Server 2008 R2 SP1, either Full or Server Core installation option. Note: there is no upgrade path from Server Core to Full or vice-versa, make this selection carefully.
- Use the latest hardware device drivers

- Hyper-V Parent Partition OS should be Active Directory Domain-joined
- Hyper-V Server Role and Failover Clustering Features are required
- Apply relevant Windows Updates, including OOB updates not offered on Microsoft Update
- All Nodes, Networks, and Storage must pass the Cluster Validation Wizard

Performance Settings

The following Hyper-V R2 network performance improvements should be tested and considered for production use:

- TCP Checksum Offload is recommended and benefits both CPU and overall network throughput performance, and is fully supported by Live Migration.
- Support for Jumbo Frames was also introduced with Windows Server 2008. Hyper-V in Windows Server 2008 R2 simply extends this capability to VMs. So just like in physical network scenarios, Jumbo Frames add the same basic performance enhancements to virtual networking. That includes up to six times larger payloads per packet, which improves not only overall throughput but also reduces CPU utilization for large file transfers.
- VMQ essentially allows the host's single NIC card to appear as multiple NICs to the VMs by allowing the host's network interface card (NIC) to Direct Memory Access (DMA) packets directly into individual VM memory stacks. Each VM device buffer is assigned a VMQ, which avoids needless packet copies and route lookups in the virtual switch. The result is less data in the host's buffers and an overall performance improvement to I/O operations.

Recommended: > Enable TCP Checksum Offload

> Enable Jumbo Frames

> Enable Virtual Machine Queue (VMQ) for 10 GbE networks

IP Addressing

The following table shows as an example the IP address settings that could be used in a vStart Fast Track.

Table 6. Example of IP Address Settings in vStart

Setting	Value
Management network IP Address	192.168.20.101
Management network Subnet Mask	255.255.255.0
Management network Gateway	192.168.20.1
Management network Primary DNS	192.168.20.10
Management network Secondary DNS	192.168.20.11
Cluster Private (Heartbeat) network IP Address	192.168.40.101
Cluster Private (Heartbeat) network Subnet Mask	255.255.255.0
Cluster Private network Gateway	N/A
Cluster Private network DNS	N/A
Live Migration network IP Address	192.168.30.101
Live Migration Subnet Mask	255.255.255.0
Live Migration network Gateway	N/A
Live Migration network DNS	N/A
Virtual Machines network IP Address	N/A
Virtual Machines network Subnet Mask	N/A
Virtual Machines network Gateway	N/A
Virtual Machines network DNS	N/A
ISCSI1 network IP Address	192.168.50.101
ISCSI1 network Subnet Mask	255.255.255.0
ISCSI1 network Gateway	N/A
ISCSI1 network DNS	N/A
ISCSI2 network IP Address	192.168.50.161
ISCSI2 network Subnet Mask	255.255.255.0

Mandatory: > Cluster Heartbeat Network must be on a distinctly separate subnet from the Host Management Network
> The Virtual Machine network adapter should not be shared with the Host operating system and therefore should not have an IP address
> The iSCSI network must be on a distinctly separate and isolated network, with a dedicated IP range used only for storage

(2) Fiber Channel / iSCSI HBA Configuration

Mandatory: > OEM to define the Host adaptor's recommended FC and iSCSI configuration, settings

The iSCSI connections from each host are balanced across the nNDC and the add-in NIC on the PowerEdge R720. This provides fault tolerance for the connections at the rNDC and NIC level. These connections are distributed across both the PowerConnect switches of the SAN network fabric. The SAN fabric also takes advantage of Ethernet Jumbo Frames.

(3) MPIO Configuration

Microsoft MPIO architecture supports iSCSI, Fiber Channel and serial attached storage (SAS) SAN connectivity by establishing multiple sessions or connections to the storage array.

Multipathing solutions use redundant physical path components — adapters, cables, and switches — to create logical paths between the server and the storage device. In the event that one or more of these components fails, causing the path to fail, multipathing logic uses an alternate path for I/O so that applications can still access their data. Each network interface card (in the iSCSI case) or HBA should be connected by using redundant switch infrastructures to provide continued access to storage in the event of a failure in a storage fabric component.

Failover times vary by storage vendor, and can be configured by using timers in the Microsoft iSCSI Software Initiator driver, or modifying the Fiber Channel host bus adapter driver parameter settings.

The EqualLogic HIT provides tuning for MPIO options such as the number of connections per volume, the load balance policy, and networks for iSCSI traffic. Of significant importance are the Max sessions per volume slice and Max Sessions per entire volume. Max sessions per volume slice specifies the maximum sessions permitted for a volume per group member. Max sessions per entire volume specifies in the maximum sessions permitted to the entire volume. Typically, configure the max sessions per entire volume as three times the Max sessions per volume slice, because volumes generally span up to three members. Since the vStart configuration is not limited by the maximum number of connections per group, the settings are maximized to provide the greatest number of connections to the volumes. The following are the parameters used for the vStart configuration:

Table 7. MPIO Settings in vStart

Setting	Value
Subnets included for MPIO	Allowed: 192.168.50.0 – 255.255.255.0 All others Excluded.
Default Load Balancing Policy	Least Queue Depth
Maximum Sessions per volume slice	4
Maximum Sessions per entire volume	8*
Use MPIO for snapshots	Checked
Use IPv4/IPv6	IPv4 Selected
Minimum Adapter Speed	1Gbps

Note: The settings for Maximum Sessions per entire volume is 8 for two PS6100 arrays.

Mandatory: > MPIO must be used all storage adapters, both iSCSI and Fiber Channel
> Follow MPIO best practices as documented in the MPIO Whitepaper http://www.microsoft.com/downloads/en/details.aspx?FamilyID=cbd27a84-23a1-4e88-b198-6233623582f3 Appendix B – MPIO & DSM Configuration and best practices
> OEM to define the adaptor’s recommended MPIO settings

(4) NIC Teaming Configuration

NIC Teaming or Link Aggregation (IEEE 802.3ad) enables network maintenance to occur at all points within the datacenter network topology without affecting applications. This technology essentially bonds physical NICs together to form one or more logical network team that sends traffic to all NICs in the team. This allows a single NIC, cable, or switch to sustain a planned or unplanned outage without disrupting the Host’s Ethernet traffic. The NIC manufacturer is also the software provider for the NIC Teaming software. Each software will have its own unique set of requirements, features, Teaming Modes, and configuration recommendations. (Note: NIC Teaming should never be used for storage traffic in conjunction with iSCSI or FCoE. Storage networking should leverage MPIO.)

Mandatory: > NIC Teaming must be used to provide high-availability to the Virtual Machine Networks
> OEM to define the NIC Teaming Solution

As shown in Figures 16 and 17, the vStart configuration uses Broadcom’s Smart Load Balancing with Failover (SLB) for the team. This team type provides a team that does not require a teaming configuration on the switches. The R720 team is configured with four ports; they are distributed with two ports across the rNDC’s ASICs, and two ports across the add-in NIC’s ASIC. These four ports form a single team. The SLB team provides fault tolerance as well as load balancing for TX and RX traffic. The traffic is load balanced on the TX side by using a hash of IP and TCP/UDP port to pick a team member to bind traffic to while the RX traffic is load balanced by analyzing RX traffic flows on each team member and responding to ARP requests to bind traffic to different team members.

Figure 16. PowerEdge R620 NIC Teaming Configuration

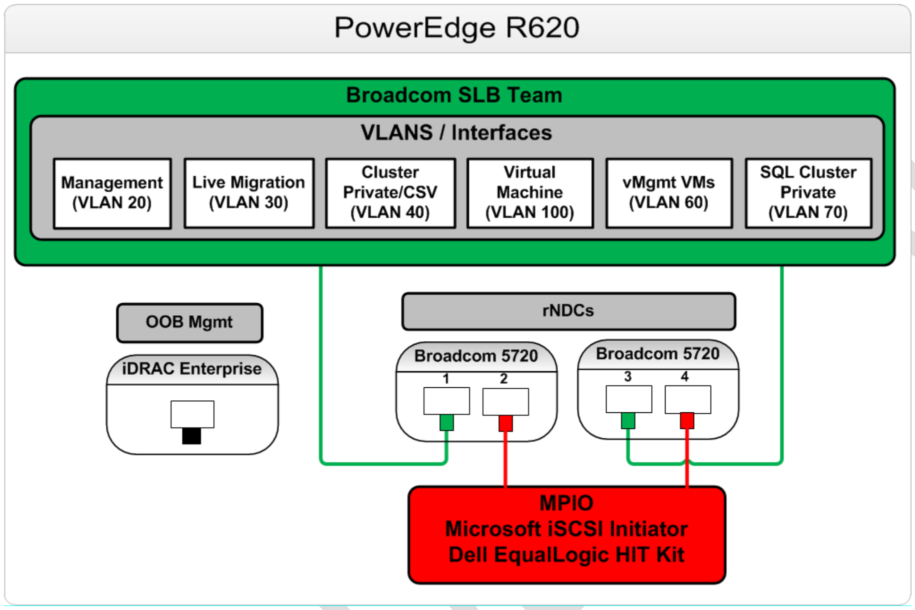
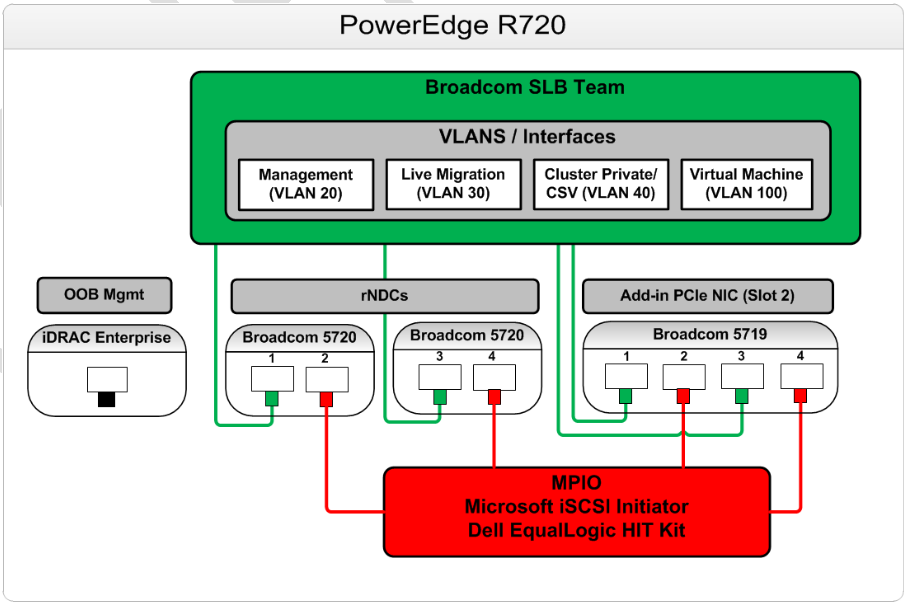


Figure 17. PowerEdge R720 NIC Teaming Configuration



VLANs are created above the team layer. This allows each VLAN interface to take advantage of the high availability and the load balancing the team layer provides. On the host nodes (R720 servers), the following VLANs are created on the NIC team: Management, Live Migration, Cluster Private, and VM traffic. On the management nodes (R620 servers), two additional VLANs are created: Management VM and SQL Cluster Private.

3.6.2 Hyper-V Host Cluster Design

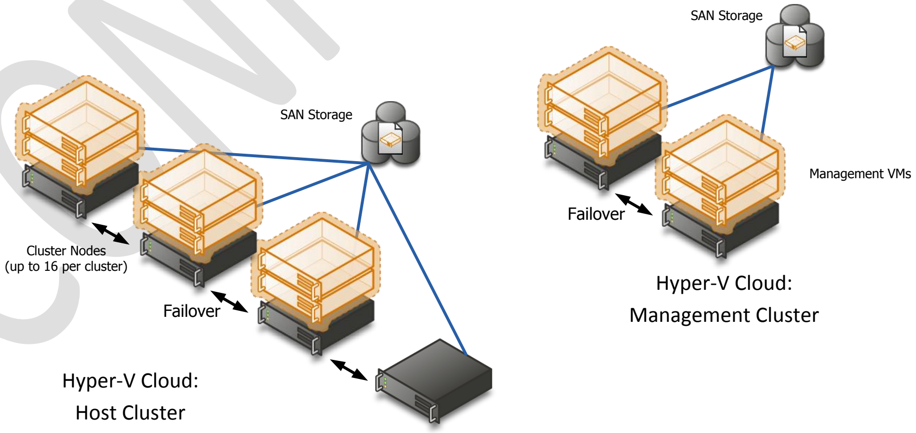
A Hyper-V host cluster is a group of independent servers that work together to increase the availability of applications and services. The clustered servers (called nodes) are connected by physical cables and by software. If one of the cluster nodes fails, another node begins to provide service (a process known as failover) which means the virtual machines on the failing node will be automatically restarted on another node in the cluster. In case of a planned migration (called Live Migration) , one or more virtual machines are moved from one host to another host in the cluster, users experience no perceptible service interruption.

The host servers are one of the critical components of a dynamic, virtual infrastructure. Consolidation of multiple workloads onto the host servers requires that those servers be highly available. Windows Server 2008 R2 provides advances in failover clustering that enable high availability and Live Migration of virtual machines between physical nodes.

(1) Server Topology

The Hyper-V Cloud consists of at least two Hyper-V host clusters (Please note exception to this rule outlined earlier in this document in Section 4.2 “Logical Architecture”). The first cluster will be at least two nodes and is referred to as the Management Cluster. The second and any additional clusters will be referred to as host clusters. Each host cluster can contain up to 16 nodes. Host clusters require some form of shared storage such as a Fiber Channel or iSCSI SAN.

Figure 18. Hyper-V Cloud Topology



(a) Management Network

A dedicated management network is required so hosts can be managed via a dedicated network such that there is not competition with guest traffic needs. A dedicated network provides a degree of separation for security and ease of management purposes. This typically implies dedicating a network adapter per host and port per network device to the management network. This network is used for remote administration of the Host, communication to Management Systems (i.e. System Center Agents), and so on.

Reference: Hyper-V: Live Migration Network Configuration Guide http://technet.microsoft.com/en-us/library/ff428137(WS.10).aspx

Additionally, most server manufacturers also provide a separate out of band management capability that enables remote management of server hardware outside of the host operating system.

Mandatory: > Implement a dedicated network for management of the infrastructure. Ensure that all Hyper-V hosts have a dedicated network adapter connected to the management network for exclusive use by the parent partition.

Recommended: > If the chosen server hardware supports an out-of-band management adapter, establish a dedicated LAN for these adapters.

(b) iSCSI Network

Hyper-V clouds provide an infrastructure to also host highly available/guest clustered workloads leveraging iSCSI connected shared storage. For this reason a dedicated iSCSI network is required which is designed so that storage traffic is not in contention with any other traffic. This typically implies dedicating two network adapters per host and ports per network device to the management network.

Mandatory: > Implement a dedicated iSCSI network or VLAN. If using 1 Gb Ethernet NICs, ensure two NICs are dedicated to iSCSI traffic to ensure redundancy. If using 10 Gb NICs, ensure a teamed, virtual NIC is presented to the parent partition for iSCSI traffic to ensure redundancy.

(c) CSV/Cluster Communication Network

Usually, when the cluster node that owns a virtual hard disk (VHD) file in CSV performs disk input/output (I/O), the node communicates directly with the storage, for example, through a storage area network (SAN). However, storage connectivity failures may prevent a given node from communicating directly with the storage. To maintain function until the failure is corrected, the node redirects the disk I/O through a cluster network (the preferred network for CSV) to the node where the disk is currently mounted. This is called CSV redirected I/O mode.

Mandatory: > Implement a dedicated CSV/Cluster Communication network. If using 1 Gb Ethernet NICs, ensure that all Hyper-V hosts have a dedicated network adapter connected to the CSV network for exclusive use by the parent partition. If using 10 Gb NICs, ensure a teamed, virtual NIC is presented to the parent partition for CSV traffic to ensure redundancy.

(d) Live Migration Network

During Live Migration, the contents of the memory of the VM running on the source node need to be transferred to the destination node over a LAN connection. To ensure high speed transfer, a dedicated Live Migration network is required.

Mandatory: > Implement a dedicated Live Migration network. If using 1 Gb Ethernet NICs, ensure that all Hyper-V hosts have a dedicated network adapter connected to the LM network for exclusive use by the parent partition. If using 10 Gb NICs, ensure a teamed, virtual NIC is presented to the parent partition for LM traffic to ensure redundancy.

Recommended: > Utilize a dedicated or shared 10 Gb Ethernet connection for the Live Migration network. This significantly reduces the time require to evacuate the VMs off of a host with zero downtime during maintenance or update windows

(e) Virtual Machine Network(s)

The Virtual Machine network(s) are dedicated to virtual machine LAN traffic. The VM network can be two or more 1 Gb Ethernet networks, one or more network created via NIC teaming, or virtual networks created from shared 10 Gb Ethernet NICs.

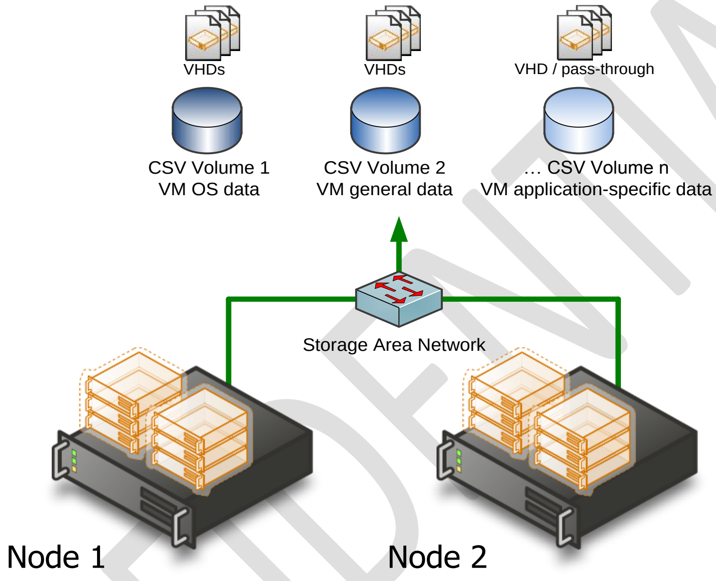
Mandatory: > Implement one or more dedicated Virtual Machine networks. If using 1 Gb Ethernet NICs, ensure that all Hyper-V hosts have two or more dedicated network adapters connected to the VM network for exclusive use by the guest VMs. If using 10 Gb NICs, ensure a teamed, virtual NIC is presented to the guest VMs to ensure redundancy.

(2) Storage Topology

Cluster Shared Volumes (CSV) is a feature that simplifies the configuration and management of Hyper-V virtual machines in failover clusters. With CSV, on a failover cluster that runs Hyper-V, multiple virtual machines can use the same LUN (disk) yet fail over (or move from node to node) independently of one another. CSV provides increased flexibility for volumes in clustered storage—for example, it allows you to keep system files separate from data to optimize disk performance, even if the system files and the data are contained within virtual hard disk (VHD) files. If you choose to use live migration for your clustered virtual machines, CSV can also provide performance improvements for the live migration process. CSV is available in versions of Windows Server® 2008 R2 and of Microsoft® Hyper-V™ Server 2008 R2 that include failover clustering.

The diagram below illustrates a design where a Hyper-V host cluster is utilizing three CSV LUNs to store different data types. One stores the guest VM OS partitions, another stores the guest VM data partition, and the third stores guest VM application data such as database files.

Figure 19. Example: Common CSV design for large Hyper-V Cluster



Mandatory: > Cluster Shared Volumes (CSV) must be enabled and able to be utilized for storing multiple virtual machines on a single LUN.

In the vStart 200 Hyper-V Cloud Fast Track solution, 1 single LUN is created on the PS6100 storage arrays and configured as a CSV on the failover Host Cluster.

3.6.3 Hyper-V Guest VM Design

Standardization is a key tenant of Private Cloud architectures and is driven by the principles that drive predictability and IT usage optimization. This also applies to Virtual Machines. As shown in Table 8, a standardized collection of Virtual Machine templates can both drive predictable performance and greatly improve capacity planning capabilities.

Table 8. Guest VM Templates

Template	Specs	Network	OS	Unit Cost
Template 1 - Small	1 vCPU, 2gb Memory, 50gb Disk	VLAN x	WS 2003 R2	1
Template 2 - Med	2 vCPU, 4gb Memory, 100gb Disk	VLAN x	WS 2003 R2	2
Template 3 - Large	4 vCPU, 8gb Memory, 200gb Disk	VLAN x	WS 2003 R2	4
Template 4 - Small	1 vCPU, 2gb Memory, 50gb Disk	VLAN x	WS 2008 R2	1
Template 5 - Med	2 vCPU, 4gb Memory, 100gb Disk	VLAN x	WS 2008 R2	2
Template 6 - Large	4 vCPU, 8gb Memory, 200gb Disk	VLAN x	WS 2008 R2	4

Mandatory: > Use documented, standardized Virtual Machine configurations for all VMs, management and tenants

(1) VM Storage

Dynamically Expanding Disks

Dynamically expanding virtual hard disks provide storage capacity as needed to store data. The size of the VHD file is small when the disk is created and grows as data is added to the disk. The size of the VHD file does not shrink automatically when data is deleted from the virtual hard disk. However, you can compact the disk to decrease the file size after data is deleted by using the Edit Virtual Hard Disk Wizard.

Fixed Size Disks

Fixed virtual hard disks provide storage capacity by using a VHD file that is in the size specified for the virtual hard disk when the disk is created. The size of the VHD file remains 'fixed' regardless of the amount of data stored. However, you can use the Edit Virtual Hard Disk Wizard to increase the size of the virtual hard disk, which increases the size of the VHD file. By allocating the full capacity at the time of creation, fragmentation at the host level is not an issue (fragmentation inside the VHD itself must be managed within the guest).

Differencing Disks

Differencing virtual hard disks provide storage to enable you to make changes to a parent virtual hard disk without altering that disk. The size of the VHD file for a differencing disk grows as changes are stored to the disk.

Pass-Through Disks

Hyper-V enables virtual machine guests to directly access local disks or SAN LUNs that are attached to the physical server without requiring the volume to be presented to the host server. The virtual machine guest accesses the disk directly (utilizing the disk's GUID) without having to utilize the host's file system. Given that the performance difference between Fixed-Disk and Pass-through Disks is now negligible, the decision is now based on manageability. For instance, if the data on the volume will be very large (hundreds of gigabytes), a VHD is hardly portable at that size given the extreme amounts of time it takes to copy. Also, bear in mind the backup scheme. With pass-through disks, the data can only be backed up from within the Guest.

When utilizing pass-through disks, there is no VHD file created; the LUN is used directly by the guest. Since there is no VHD file, there is no dynamic sizing capability or snapshot capability.

In-guest iSCSI Initiator

Hyper-V can also utilize iSCSI storage by directly connecting to iSCSI LUNs utilizing the guest's virtual network adapters. This is mainly used for access to large volumes, volumes on SANs which the Hyper-V Host itself is not connected to, or for Guest-Clustering. Guests cannot boot from iSCSI LUNs accessed through the virtual network adapters without utilizing a third-party iSCSI initiator.

Recommended: > **Utilize Fixed disks** for production environments which provide better performance and ease the monitoring of storage availability. Utilizing fixed disks allocates the full size of the disk upon creation.

> **Dynamically Expanding disks** are also a viable options for production use. However, they carry other risks such as storage oversubscription and fragmentation, so use with caution.

> **Differencing disks** are never recommended for production server workloads

> **Use pass-through disks** only in cases where absolute maximum performance is required and the loss of features such as snapshots and portability is acceptable. Since the performance difference between pass-through and fixed-disks is minimal there should be very few scenarios where pass-through disks are required.

> **For in-guest iSCSI**, ensure that a separate virtual network is utilized for access to the iSCSI storage to obtain acceptable performance. If the VM iSCSI network is shared with Ethernet traffic, utilize QoS to provide performance guarantees to the different networks. Consider using Jumbo Frames within the Guest to improve iSCSI performance.

(2) VM Networking

Hyper-V Guests support two types of virtual network adapters: Synthetic and Emulated. Synthetic makes use of the Hyper-V VMBUS architecture and is the high-performance, native device. Synthetic devices require the Hyper-V Integration Components be installed within the guest. Emulated adapters are available to all guests even if Integration Components are not available. They are much slower performing and only should be used if Synthetic is unavailable.

Recommended: > Always use Synthetic Virtual Network Adapters when possible.
Use Emulated Network Adapters only for unsupported Guest OSs or in special circumstances such as if the Guest needs to PXE boot.

You can create many virtual networks on the server running Hyper-V to provide a variety of communications channels. For example, you can create networks to provide the following:

- Communications between virtual machines only. This type of virtual network is called a private network.
- Communications between the Host server and virtual machines. This type of virtual network is called an internal network.
- Communications between a virtual machine and a physical network by creating an association to a physical network adapter on the host server. This type of virtual network is called an external network.

Recommended: > For the private cloud scenario, the recommendation is to use one or more External networks per VM, and segregate the networks with VLANs and other network security infrastructure as needed.

(3) Virtual Processors

Please reference the below table for supported number of virtual processors in a Hyper-V guest.

Table 9. Virtual Processors in Supported Guest OS

Supported Operating Systems	Virtual Processors		
	1	2	4
Windows Server 2008 R2	x	x	x
Windows Server 2003 x86x64 SP2	x	x	
Windows 2000 Server & Advanced Server SP4	x		
Windows HPC Server 2008	x	x	x
SUSE Linux Enterprise Server 10 & 11 x86x64	x	x	x
Red Hat Enterprise Linux 5.2, 5.3, 5.4 x86x64	x	x	x
Windows® 7 x86/x64	x	x	x
Windows Vista® x86/x64 w/ SP1	x	x	
Windows XP Pro x64 w/ SP2 & x86 w/ SP3	x	x	
Windows XP Pro x86 w/ SP2	x		

Hyper-V supports a maximum ratio of eight Virtual Processors (VPs) per one Logical Processor. A Logical Processor is defined as a processing core seen by the Host operating system or Parent Partition. i.e. in the case of Intel Hyper-Threading, each thread is considered an LP.

Therefore a 16 LP server supports a maximum of 128 VPs. That would in turn equate to 128 single-proc VMs, 64 dual-proc VMs, or 32 quad-proc VMs. The 8:1 VP/LP ratio is a maximum supported limit, but actual real-work experience has resulted in a best practice of approximately 2.75:1 VP/LP for production server workloads.

⚠ **Recommended:** > Use a Virtual to Logical processor ratio of approximately 2.75:1 for production server workloads

🛑 **Mandatory:** > Hyper-V's maximum supported ratio is 8:1, this must not be exceeded

3.7 Management Architecture

3.7.1 Management Scenarios

(1) Infrastructure Deployment

When deploying several to hundreds of hosts, choosing the deployment model is a critical decision. Choices range from manual installation which is highly inefficient through varying degrees of automation up to enterprise class management systems. To achieve the architecture principle of driving predictability, all infrastructure component should be able to be deployed and configured in a repeatable and automated fashion. Examples include configuring network infrastructure, storage infrastructure, provisioning servers, creating cluster, and so on.

The key components are the Microsoft Deployment Toolkit 2010 (MDT) and Windows Deployment Services (WDS). These are complimented by standard Windows Server roles such as Active Directory Domain Services, DNS, and DHCP.

Using these technologies it is possible to provide a robust deployment infrastructure using standard in box solutions and toolkits. This infrastructure can later be extended with System Center or an equivalent 3rd party management system.

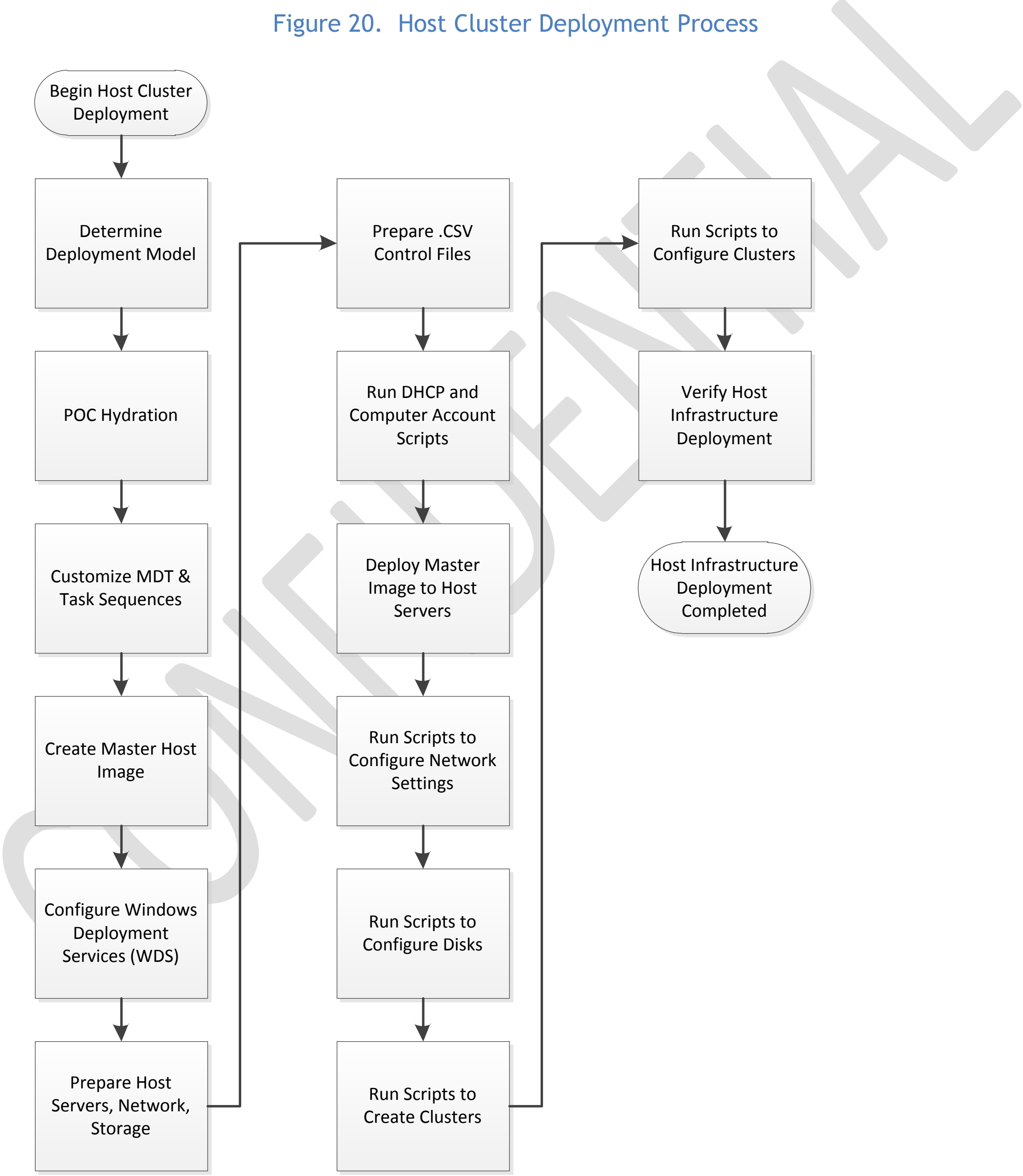
The process and automation detailed in this document are based on the following choices:

- Windows Server 2008 R2 SP1 Datacenter Core x64 installation option for all Hyper-V hosts
- The Windows Management Framework (PowerShell 2.0) is available on all hosts
- PowerShell Remoting (WinRM) will be enabled on all hosts
- All hosts will be domain joined
- All hosts will be joined to 16-node host clusters using Cluster Shared Volumes
- DHCP and DHCP Reservations will be used for all IP addressing
- Microsoft Deployment Toolkit 2010 will be used to create master images
- Windows Deployment Services (WDS) will be used to deploy all hosts
- A network infrastructure supporting PXE boot and infrastructure deployment is available
- A hardware management system is available which can remotely power off, power on, and reboot servers is available

- The MAC address of every NIC in every server is documented and associated to the physical network it is connected to (i.e. if iSCSI is used, the MAC address of each NIC connecting to the iSCSI network and the server it belongs to is documented)

Note: The PowerEdge R720 and R620 servers in the Dell Fast Track vStart configurations are shipped with the factory-installed Windows Server 2008 R2 SP1 with the Hyper-V role enabled. To deploy the failover cluster, it requires these servers to join an Active Directory domain. Contact Dell services representatives if further deployment help is needed.

Figure 20. Host Cluster Deployment Process


Begin Host Cluster Deployment
Determine Deployment Model
POC Hydration
Customize MDT & Task Sequences
Create Master Host Image
Configure Windows Deployment Services (WDS)
Prepare Host Servers, Network, Storage
Prepare .CSV Control Files
Run DHCP and Computer Account Scripts
Deploy Master Image to Host Servers
Run Scripts to Configure Network Settings
Run Scripts to Configure Disks
Run Scripts to Create Clusters
Run Scripts to Configure Clusters
Verify Host Infrastructure Deployment
Host Infrastructure Deployment Completed

Mandatory: > Describe the OEM solution for automatically configuring Hyper-V Cloud hardware and deploying the Windows and Hyper-V clusters and management solutions.

Optional: > Utilize built-in Windows Server components and free Microsoft toolkits

(2) VM Provisioning and De-provisioning

One of the primary cloud attributes is user self-service, or providing the consumer of a service the ability to request that service and have it be automatically provisioned for them. In the Hyper-V Cloud solution, this refers to the ability for the user to request one or more virtual machines or to delete one or more of their existing virtual machines. The infrastructure scenario supporting this capability is the VM Provisioning and De-provisioning process. This process is initiated from the self-service portal or tenant user interface and triggers and automated process or workflow in the infrastructure through System Center Virtual Machine Manager to either create or delete a virtual machine based on the authorized settings input but the user or tenant. Provisioning could be template-based such as requesting a small, medium, or large VM template or a series of selections could be made by the user (vCPUs, RAM, etc.). If authorized, the provisioning process should create a new VM per the user's request, add the VM to any relevant management products in the Hyper-V cloud (such as System Center) and enable access to the VM by the requestor.

(3) Infrastructure Monitoring

The Hyper-V Cloud must enable the ability monitor every major component of the solution and generate alerts based on performance, capacity, and availability metrics. Examples include monitoring server availability, CPU, and storage utilization.

(4) Infrastructure Maintenance

The Hyper-V Cloud must enable the ability to perform maintenance on any component of the solution without impacting the availability of the solution. Examples include the need to update or patch a host server, add additional storage to the SAN, etc. During maintenance the system should ensure that unnecessary alerts or events are not generated in the management systems during planned maintenance.

(5) Resource Optimization

Elasticity, Perception of Infinite Capacity, and Perception of Continuous Availability are Hyper-V Cloud architecture principles that relate to resource optimization. This management scenario deals with optimizing resources by dynamically moving workloads around the infrastructure based on performance, capacity, and availability metrics. Examples include the option to distribute workloads across the infrastructure for maximum performance or consolidating as many workloads as possible to the smallest number of hosts for a higher consolidation ratio.

(6) Backup and Disaster Recovery

The Hyper-V Cloud solution must provide a mean of backing up and recovering both the virtual machines as well as the host infrastructure.

(7) Reporting (used by chargeback, capacity, service management, health, performance)

The Hyper-V Cloud solution must provide a centralized reporting capability. The reporting capability should provide standard reports detailing capacity, utilization, and other system metrics. The reporting functionality serves as the foundation for capacity or utilization-based billing and chargeback to tenants.

3.7.2 Virtualization

(1) Storage Virtualization

Storage virtualization is a concept in IT System Administration, referring to the abstraction (separation) of logical storage from physical storage so that it may be accessed without regard to physical storage or heterogeneous structure. This separation allows the Systems Admin increased flexibility in how they manage storage for end users.

http://en.wikipedia.org/wiki/Storage_virtualization

(2) Network Virtualization

http://www.snia.org/education/storage_networking_primer/stor_virt/

In computing, Network Virtualization is the process of combining hardware and software network resources and network functionality into a single, software-based administrative entity, a virtual network. Network virtualization involves platform virtualization, often combined with resource virtualization.

Network virtualization is categorized as either external, combining many networks, or parts of networks, into a virtual unit, or internal, providing network-like functionality to the software containers on a single system. Whether virtualization is internal or external depends on the implementation provided by vendors that support the technology.

Various equipment and software vendors offer network virtualization by combining any of the following:

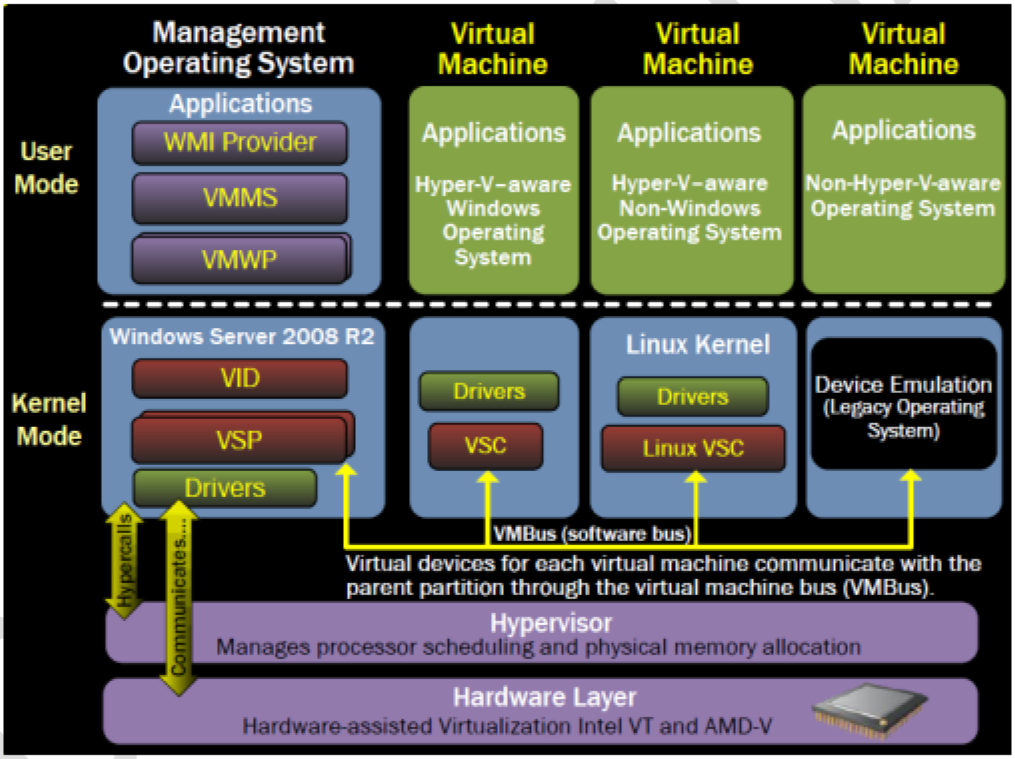
- Network hardware, such as switches and network adapters, also known as network interface cards (NICs)
- Networks, such as virtual LANs (VLANs) and containers such as virtual machines (VMs) and Solaris Containers
- Network storage devices
- Network media, such as Ethernet and Fiber Channel

(3) Server Virtualization

Hardware Virtualization uses software to create a Virtual Machine (VM) that emulates a physical computer. This creates a separate OS environment that is logically isolated from the host server. By providing multiple VMs at once, this approach allows several operating systems to run simultaneously on a single physical machine.

Hyper-V technology is based on a 64-bit hypervisor-based microkernel architecture that enables standard services and resources to create, manage, and execute virtual machines. The Windows Hypervisor runs directly above the hardware and ensures strong isolation between the partitions by enforcing access policies for critical system resources such as memory and processors. Unlike Windows operating systems such as Windows Server 2003 and earlier versions, the Windows Hypervisor does not contain any third-party device drivers or code, which minimizes its attack surface and provides a more secure architecture.

Figure 21. Hyper-V Architecture



Management Operating System
Virtual Machine
Virtual Machine
Virtual Machine
User Mode
Applications
WMI Provider
VMMS
VMWP
Applications
Hyper-V-aware Windows Operating System
Applications
Hyper-V-aware Non-Windows Operating System
Applications
Non-Hyper-V-aware Operating System
Kernel Mode
Windows Server 2008 R2
VID
VSP
Drivers
Drivers
VSC
Linux Kernel
Drivers
Linux VSC
Device Emulation (Legacy Operating System)
Hypercalls
Communicates...
VMBus (software bus)
Virtual devices for each virtual machine communicate with the parent partition through the virtual machine bus (VMBus).
Hypervisor
Manages processor scheduling and physical memory allocation
Hardware Layer
Hardware-assisted Virtualization Intel VT and AMD-V

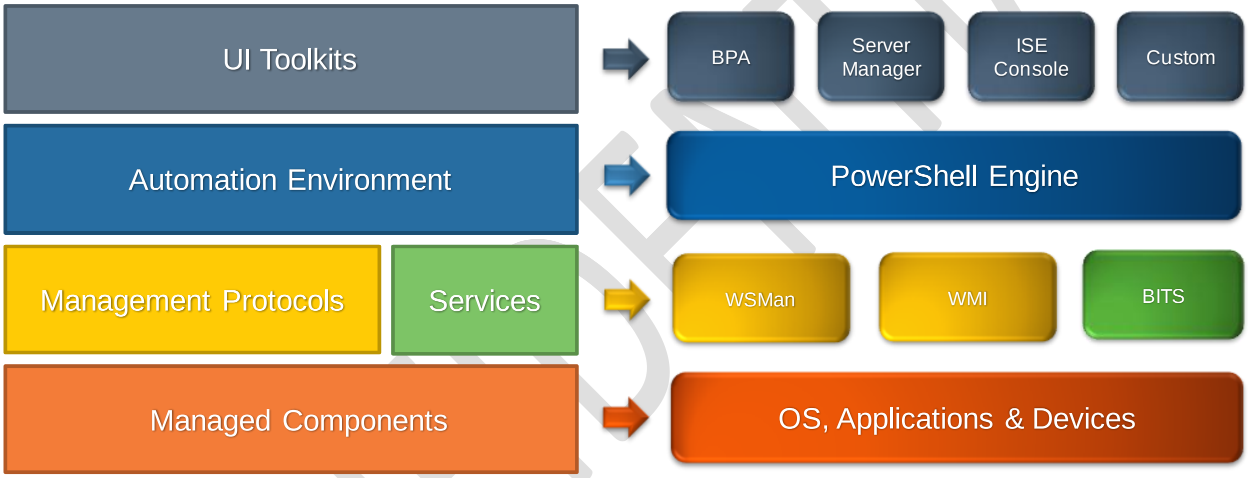
In addition to the Windows Hypervisor, there are two other major elements to consider in Hyper-V: a parent partition and child partition. The parent partition is a special virtual machine that runs Windows Server 2008 R2, controls the creation and management of child partitions, and maintains direct access to hardware resources. In this model, device drivers for physical devices are installed in the parent partition. In contrast, the role of a child partition is to provide a virtual machine environment for the installation and execution of guest operating systems and applications.

3.7.3 Automation

The Automation layer is made up of the foundational automation technology plus a series of single-purpose commands and scripts that perform operations such as starting or stopping a VM, rebooting a server, or applying a software update. These atomic units of automation are combined and executed by higher-level management systems. The modularity of this layered approach dramatically simplifies development, debugging, and maintenance.

The Windows Management Framework is a core set of infrastructure technologies which combine to provide advanced local and remote automation and system management. Key underpinning technologies such as WMI, WSMan, and BITS are leveraged by higher level layers such as PowerShell, which itself is utilized by higher level layers such as scripts, user interfaces, and suites such as System Center.

Figure 22. Windows Management Framework

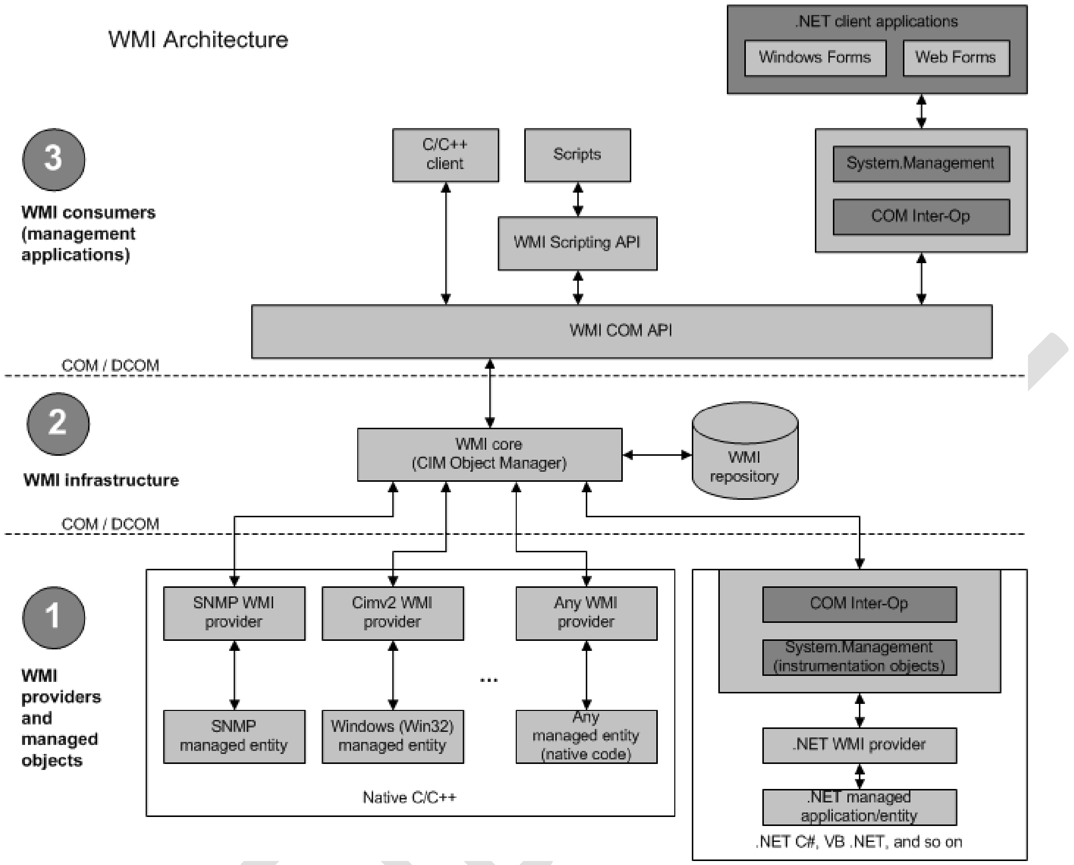


The Windows Management Framework Core package provides updated management functionality for IT Professionals. This package includes the following components: Windows PowerShell 2.0 and Windows Remote Management (WinRM) 2.0.

Windows Management Instrumentation (WMI) is the Microsoft implementation of Web-based Enterprise Management (WBEM), which is an industry initiative to develop a standard technology for accessing management information in an enterprise environment. WMI uses the Common Information Model (CIM) industry standard to represent systems, applications, networks, devices, and other managed components. CIM is developed and maintained by the Distributed Management Task Force (DMTF).

WS-Management is now the recommended protocol for accessing WMI properties and methods over DCOM for the following reasons: single port which is more versatile for traversing networks that are secured, simpler firewall configuration both locally and in the network, more structured (XML/SOAP) data output, and better diagnostics when connectivity is not working.

Figure 23. WMI Architecture



The Windows Remote Management (WinRM) is the Microsoft implementation of WS-Management Protocol, a standard Simple Object Access Protocol (SOAP)-based, firewall-friendly protocol that allows hardware and operating systems, from different vendors, to interoperate.

The WS-Management protocol specification provides a common way for systems to access and exchange management information across an IT infrastructure. WinRM and Intelligent Platform Management Interface (IPMI), along with the Event Collector are components of the Windows Hardware Management features.

3.7.4 Private Cloud Management

The management layer consists of the tools and systems that are utilized to deploy and operate the infrastructure. In most cases, this consists of a variety of different toolsets for managing hardware, software, and applications. The System Center family has comprehensive, broad set of capabilities to enable management of the cloud fabric and integrate with toolsets.

(1) SQL Server 2008 SP1

Microsoft System Center components are database driven applications. This makes a highly-available and well-performing database platform critical to the overall of management the environment.

SQL Server Configuration

- 2 Non-HA VMs on different Hyper-V Hosts
- Windows Server 2008 R2 SP1 Enterprise Edition
- 4 vCPUs
- 8GB Memory (do not use Dynamic Memory)
- 4 vNICs (1 client connections, 1 in-guest cluster communications, 2 iSCSI)
- Storage: 1 OS VHD, 4 x Dedicated iSCSI LUNs

Table 10. SQL Data Locations

LUN	Purpose	Size
LUN 1, iSCSI	VM Operating System	30GB VHD
LUN 2, iSCSI	SQL Databases	Varies
LUN 3, iSCSI	SQL Logging	Varies
LUN 4, iSCSI	SQL Cluster Quorum	1GB
LUN 5, iSCSI	DTC	1GB

Table 11. Databases

DB Client	Instance Name	DB name	Authentication
VMM SSP	<Instance 1>	<SCVMMSSP>	Win Auth
WSUS	<Instance 1>	<WSUS_DB>	Win Auth
Ops Mgr	<Instance 1>	<Ops Mgr_DB>	Win Auth
Ops Mgr	<Instance 2>	<Ops Mgr_DW_DB>	Win Auth
Ops Mgr	<Instance 2>	<Ops Mgr_Report_DB>	Win Auth
VMM	<Instance 1>	<VMM_DB>	Win Auth

(2) System Center Virtual Machine Manager (SCVMM) 2008 R2

Virtual Machine Manager 2008 R2 helps enable centralized management of physical and virtual IT infrastructure, increased server utilization, and dynamic resource optimization across multiple virtualization platforms. It includes end-to-end capabilities such as planning, deploying, managing, and optimizing the virtual infrastructure.

Scope

SCVMM will be used to manage only Hyper-V Cloud Fast Track Hosts and Guests in a single Datacenter. No virtualization infrastructure outside of the solution should be managed by SCVMM. The SCVMM configuration is only considering the scope of this architecture and therefore may suffer performance and health issues if that scope is changed.

Servers

- 1 HA VM
- Windows Server 2008 R2 SP1
- 2 vCPU
- 4gb Memory
- 1 vNICs
- Storage: 1 OS VHD, 1 x Data VHD or Pass-through volume

Roles

The following roles are required by VMM. If a role is not listed, it will not be installed.

- VMM Server
- Administrator Console
- Command Shell
- VMM Library
- SQL Database (remote)

Operations Manager Integration

In addition to the built-in Roles, VMM with be integrated with Operations Manager. VMM uses System Center Operations Manager 2007 to monitor the health and availability of the virtual machines and virtual machine hosts that VMM is managing. VMM also uses Operations Manager to monitor the health and availability of the VMM server, database server, library servers, and self-service Web servers, and to provide Diagram views of the virtualized environment in the VMM Administrator Console. To enable these features, you must integrate Operations Manager with VMM. Integration with Operations Manager is also a prerequisite for enabling Performance and Resource Optimization (PRO) in VMM and for configuring reporting in VMM.

Configuring Operations Manager Integration with VMM http://technet.microsoft.com/en-us/library/cc956099.aspx

VMM Library Placement

Libraries are the repository for VM Templates, VHDs, ISOs, etc. and therefore serve a very important role. The Library Share itself will reside on the VMM Server, however, the share is to be placed on its own logical partition and corresponding VHD or Pass-through disk whose underlying disk subsystem is robust enough to service the provisioning demands.

VM Provisioning

Standard provisioning is used in VMM.

Performance and Resource Optimization (PRO)

Performance and Resource Optimization (PRO) is a feature of System Center Virtual Machine Manager 2008 that enables dynamic management of virtualized infrastructure.

The host-level PRO actions in the VMM 2008 Management Pack recommend migrating the virtual machine with the highest resource usage on the host whenever the CPU or memory usage on the host exceeds the threshold defined by a PRO monitor. The virtual machine is migrated to another host in the host group or host cluster that is running the same virtualization software. If you have a workload that is not suitable for migration running in a virtual machine, you can exclude that virtual machine from host-level PRO actions. The virtual machine will not be migrated even if it has the highest usage of the elevated resource on the host.

This is accomplished by using *intelligent placement* in VMM. During intelligent placement, VMM chooses the most suitable host for the virtual machine based on the virtual machine’s configuration, and determines the placement preference (via load balancing or resource maximization) that is in effect, and then migrates the virtual machine.

Under automatic implementation of PRO tips, remediation actions are performed in the order in which PRO tips are received. If multiple hosts in a host group or cluster have exceeded their PRO threshold for the same resource, this can cause multiple migrations.

If the host group or cluster is configured to receive but not auto implement PRO tips, the VMM administrator can review all PRO tips and implement the PRO tip of choice. The PRO tip details in the **PRO Tips** window indicate the virtual machine that will be migrated and the host to which it will be migrated based on resource usage when the PRO tip was initiated. Because a PRO tip is dismissed when the host’s resource usage falls back to within the performance threshold, the virtual machine and target host are unlikely to change over the life of a PRO tip.

Mandatory:

> VMM must be on a dedicated VM and use the Remote SQL instance
> The VMM Library must be on a dedicated, high-performing VHD or Pass-through disk
> VMM must be integrated with Operations Manager
> PRO must be used with automatically-implemented Tips

Dell PRO-enabled Management Packs

PRO provides an open and extensible framework for the creation of management packs for virtualized applications or associated hardware. In building these management packs enabled for PRO, Dell has created a customized solution combining Dell monitoring offerings with the comprehensive monitoring and issue-resolution capabilities of PRO.

Dell has incorporated awareness of system resources through the use of Open Manage Server Administrator in the PRO-enabled management packs. Watch points include but are not limited to server temperature, local RAID status, and power supply state. With these pre-determined watch points and resolution steps, PRO can react dynamically to adverse situations.

Note:

Dell PRO Management Packs and OpenManage Server Administrator can be downloaded at support.dell.com.

⚠ **Recommended:** > OEM describe the implemented Partner PRO Packs

(3) System Center Operations Manager (OpsMgr) 2007 R2

OpsMgr agents will be deployed to the Fabric Management Hosts and VMs, and to Scale Unit Hosts and VMs. These in-guest agents are used to provide Performance and Health of the operating system only.

Scope of the OpsMgr instance is for Hyper-V Cloud infrastructure monitoring only. Application-level monitoring is out of scope for this OpsMgr instance.

Servers

- 1 HA VM
- Windows Server 2008 R2 SP1
- 2 vCPU
- 4GB Memory
- 1 vNICs
- Storage: 1 OS VHD

Roles

The following roles are required by Ops Manager. If a role is not listed, it will not be installed.

- Root Management Server
- Reporting Server (DB will reside on SQL server)
- Data Warehouse (DB will reside on SQL server)
- Operator Console
- Command Shell

Management Packs

The following OpsMgr Management Packs are required:

- Virtual Machine Manager 2008 R2
- Windows Server Base Operating System
- Windows Server Failover Clustering
- Windows Server 2008 Hyper-V
- Microsoft SQL Server Management Pack
- Microsoft Windows Server Internet Information Services (IIS) 2000/2003/2008
- System Center MPs

- Dell Management Packs for PowerEdge Servers and EqualLogic PS Arrays

Reporting

The following reports are made available as a result of VMM & Operations Manager Integration. In addition, many customizable reports are available with the above Operations Manager Management Packs.

- Virtualization candidates
  - Identify physical computers that are good candidates for conversion to VM's
  - Uses server performance metrics available in Operations Manager
- VM utilization
  - Report resource utilization by your virtual machines
  - Report under-utilized or over-utilized virtual machines
- VM allocation
  - Calculate chargeback to cost centers
  - Report CPU, memory, disk, and network usage
- Host utilization
  - Report the number of virtual machines running on each host
  - Report average usage and total or maximum values for processors, memory, and disk space
- Host utilization growth
  - Report the percentage of change in resource usage and number of running VM's

Mandatory: > OpsManager must be on a dedicated VM and use the Remote SQL instance
> OpsManager must be VMM-integrated

Recommended: > Utilize System Center Virtual Machine Manager and Operations Manager Reporting.

(4) Maintenance and Patch Management

(a) Windows Server Update Services (WSUS)

Windows Server Update Services (WSUS) enables information technology administrators to deploy the latest Microsoft product updates to computers that are running the Windows operating system. By using WSUS, administrators can fully manage the distribution of updates that are released through Microsoft Update to computers in their network.

Microsoft Windows Server Update Services 3.0 SP2 Deployment Guide

http://www.microsoft.com/downloads/en/details.aspx?FamilyID=113d4d0c-5649-4343-8244-e09e102f9706&displaylang=en

(b) System Center Configuration Manager

System Center Configuration Manager 2007 R2 comprehensively assesses, deploys, and updates servers, client computers, and devices-across physical, virtual, distributed, and mobile environments. Optimized for Windows, it is the best choice for gaining enhanced insight into and control over IT systems. Configuration Manager has capabilities such as:

- Deploying operating systems
- Deploying software applications
- Deploying software updates
- Metering software usage
- Assessing variation from desired configurations
- Taking hardware and software inventory
- Remotely administering computers

(c) Virtual Machine Servicing Tool

VMST 3.0 helps customers reduce IT costs by making it easier to update their offline virtual machines, templates, and virtual hard disks with the latest operating system and application patches—without introducing vulnerabilities into their IT infrastructure.

Version 3.0 of the tool works with System Center Virtual Machine Manager 2008 R2, System Center Configuration Manager 2007 SP2, and Windows Server Update Services 3.0 SP2. The tool also supports updating the Windows® 7 and Windows Server® 2008 R2 operating systems.

http://technet.microsoft.com/en-us/library/cc501231.aspx

⚠ **Recommended:** > Describe the OEM solution for maintenance and patch management OR REMOVE SECTION

(5) Backup and Disaster Recovery

In a Virtualized datacenter, there are 3 commonly used backup types: Host-based, Guest-based, and SAN-based. The below table contrasts these types:

Table 12. Comparison of Common Backup Solutions

Capability	Host Based	Guest Based	SAN Snapshot
Protection of VM configuration	X		X*
Protection of Host & Cluster configuration	X		X*
Protection of Virtualization-specific data such as VM snapshots	X		X
Protection of data inside the VM	X	X	X
Protection of data inside the VM stored on pass-through disks		X	X

Support for VSS-based backups for supported operating systems and applications	X	X	X*
Support for Continuous Data Protection	X	X	
Ability to granularly recover specific files or applications inside the VM		X	

*Depends on storage vendor’s level of Hyper-V Integration

(a) Data Protection Manager 2010

Microsoft System Center Data Protection Manager 2010 (DPM) provides continuous data protection for virtual machines hosted on servers running Microsoft’s Hyper-V. This protection includes online backup of supported guest virtual machines hosted on clustered or standalone systems; protection of virtual machines during the live migration process; and, item level recovery from host-level backup. DPM 2010 offers disk-to-disk, disk-to-tape, and disk-to-disk-to-tape technologies; all of which maintain the business value of a virtualized infrastructure by ensuring that it is better protected and always available. Other backup solutions are permitted as long as they support proper online backup of Hyper-V virtual machines in the designed solution.

How to protect Hyper-V with DPM 2010 whitepaper
http://www.microsoft.com/downloads/en/details.aspx?displaylang=en&FamilyID=c9d141cf-c839-4728-af52-928f61bebdca

🛑 **Mandatory:** > Capability to support the Hyper-V VSS writer for host side backup (Required)
> Backup storage separate from the SAN (SAN snap technology can be used in conjunction with a mechanism to move the backup off the production SAN

⚠ **Recommended:** > Capability to restore individual files from the VM backup

🟢 **Optional:** > Application Awareness in backup solution

(6) Tenant / User Self Service Portal

The Tenant / User Self-Service layer provides an interface for Hyper-V Cloud tenants or authorized users to request, manage, and access the services, such as virtual machines, provided by the Hyper-V Cloud architecture. Using role-based access control and authorization, the Self-Service layer provides the ability to delegate certain aspects of administration (such as starting/stopping VMs) to designated “tenant administrators”.

The vStart configuration uses the System Center Virtual Machine Manager Self-Service Portal (VMM SSP) for user self-service.

(a) System Center Virtual Machine Manager Self-Service Portal v2

The System Center Virtual Machine Manager Self-Service Portal is an extensible Web-based application that provides a way for groups within an organization (referred to as business units) to manage the self-service provisioning of IT infrastructures while the physical resources (servers, networks, storage devices, and related hardware) remain in a centralized pool, referred to as the datacenter. Instead of using physical servers and related hardware to build an IT infrastructure, a business unit IT (BUIT) administrator uses the self-service portal to build an IT infrastructure from virtual machines. The self-service portal has three components:

VMMSSP Website. A Web-based component that provides a user interface to the self-service portal. Through the VMMSSP website, users can perform various tasks such as pooling infrastructure assets in the self-service portal, extending virtual machine actions, creating business unit and infrastructure requests, validating and approving requests, and provisioning virtual machines (using the self-service virtual machine provisioning feature). Users can also use the VMMSSP website to view information related to these tasks.

VMMSSP Database. A SQL Server database that stores information about configured assets, information related to business units and requests, and information about what has been provisioned to various business units.

VMMSSP Server. A Windows service that runs default and customized virtual machine actions that the user requests through the VMMSSP website.

Microsoft System Center Virtual Machine Manager Self-Service Portal 2.0 Release Candidate
http://www.microsoft.com/downloads/en/details.aspx?FamilyID=fef38539-ae5a-462b-b1c9-9a02238bb8a7&displaylang=en

🛑 **Mandatory:** > Describe the OEM solution for providing Self-Service access enabling on-demand provisioning, de-provisioning, reporting, monitoring, and metering/billing.

⚠ **Recommended:** > Utilize the Virtual Machine Manager Self-Service Portal 2.0 (VMM SSPv2)

(7) Dell EqualLogic Storage Management

EqualLogic Group Manager is a SAN management tool integrated with the EqualLogic Firmware that provides you with detailed information on your SAN configuration, and provides you with an easy-to-use tool for storage provisioning, replication scheduling and array management. The EqualLogic Group Manager is available as a command-line interface (CLI) or graphical-user interface (GUI) that can be accessed from Microsoft® Internet Explorer® or Mozilla® Firefox® web browsers with a connection to the EqualLogic SAN. The EqualLogic Group Manager GUI enables IT Administrators to manage various components:

- Modify the group configuration and set up accounts, event notification, network services, authentication, and SNMP.
- Monitor iSCSI connections to the group, snapshot and replication schedules, volume replication configurations and activity, administrative sessions and login history, and in-progress member and volume move operations.

- Display events in the group.
- Create and manage pools in the group.
- Monitor and manage group members, including configuring network interfaces, etc.
- Monitor and manage volumes, snapshots, replicas, and schedules.
- Create and manage collections of volumes. Organizing multiple, related volumes into a collection enables you to create snapshots or replicas of the volumes in a single operation or schedule.
- Monitor and manage replication partners.

The EqualLogic HIT also includes the PowerShell snap-in for managing the storage which can be used for automating tasks such as volume creation or permission assignments.

Dell EqualLogic SAN HeadQuarters (SAN HQ) enables you to monitor multiple PS Series groups from a single graphical user interface (GUI). It gathers and formats performance data and other vital group information. Analyzing the data can help you improve performance and more effectively allocate group resources. Using SAN HQ, you can:

- Monitor multiple EqualLogic SAN Groups
- Launch the EqualLogic Group Manager management console for any monitored storage group
- Get a consolidated view of alerts
- Perform trend analysis, with SAN HQ's customizable views according to storage manager-defined timelines
- Manage multiple versions of firmware

Mandatory: > Describe the OEM solution for Storage Management

(8) Dell PowerConnect Network Management

Like the storage architecture, the vStart network can be managed by either a serial port of each switch device, or over the IP network by a web browser, SSH, or telnet client. The embedded Dell OpenManage Switch Administrator provides a Web browser-based graphical user interface (GUI) to manage and monitor Dell PowerConnect switch systems. Using OpenManage Switch Administrator, IT administrators can configure the switch ports' VLAN settings, jumbo frames, LAG membership, etc.

Mandatory: > Describe the OEM solution for Network Management

(9) Dell PowerEdge Server Management Utilities

Dell PowerEdge servers provide several interfaces from which the server's hardware can be monitored and managed. Interfaces such as the front panel LCD screen report simple values such as hostname and power consumption or tools such as the Dell OpenManage™ Server Administrator provides a more comprehensive systems management solution.

(a) Dell iDRAC Out-of-Band Management

The iDRAC7 provides the out-of-band management interface in the vStart configuration for each Dell PowerEdge Server. The iDRAC7 provides an out-of-band interface into server health and provides a remote console for administrators to use to access the in-band system. The iDRAC7 can manage the power state of the server allowing remote operations to be performed such as power up, power down and power cycle. It also provides remote media capabilities allowing media such as CD/DVD ROM, USB, or ISO images to be presented to the host as a local drive. This functionality enables administrators to remotely boot, mount media, and ultimately install a system in a hands-off fashion.

(b) Dell OpenManage Server Administrator

Dell OpenManage Server Administrator (OMSA) provides a comprehensive, one-to-one systems management solution in two ways: from an integrated, Web browser-based GUI and from a command line interface (CLI) through the operating system. OMSA is designed for system administrators to manage systems and its attached storage locally and remotely on a network. It provides system configuration and health information directly to the operating system and can further integrate with external systems such as System Center Operations Manager 2007.

Mandatory: > Describe the OEM solution for Server Management and OOB Management

3.7.5 Orchestration

The orchestration layer is the critical interface between the IT organization and its infrastructure. It is the layer at which intent is transformed into workflow and automation.

Ideally, the orchestration layer provides a graphical interface in which complex workflows that consist of events and activities across multiple management-system components can be combined, so as to form an end-to-end IT business process such as automated patch management or automatic power management. The orchestration layer must provide the ability to design, test, implement, and monitor these IT workflows.

(1) Opalis

Opalis is an automation platform for orchestrating and integrating IT tools to drive down the cost of datacenter operations, while improving the reliability of IT processes. It enables IT organizations to automate best practices, such as those found in Microsoft Operations Framework (MOF) and Information Technology Infrastructure Library (ITIL). This is achieved through workflow processes that coordinate System Center and other management tools to automate incident response, change and compliance, and service-lifecycle management processes.

Through its workflow designer, Opalis automatically shares data and initiates tasks in System Center Operations Manager, System Center Configuration Manager, System Center Service Manager, Virtual Machine Manager, Active Directory and 3rd party tools. Opalis workflow automates IT infrastructure tasks, while System Center Service Manager workflow provides automation of human workflow. The combined offering ensures repeatable, consistent results by removing the latency associated with manual coordination service delivery. System Center and Opalis enable integration, efficiency, and business alignment of the datacenter IT services by:

- Automating cross-silo processes and enforcing best practices for incident, change, and service lifecycle management.
- Reducing unanticipated errors and service delivery time by automating tasks across vendor and organization silos.
- Integrating System Center with non-Microsoft tools to enable interoperability across the datacenter.
- Orchestrating tasks across systems for consistent, documented, compliant activity.

Mandatory: > Describe the OEM solution for orchestrating the various automation and management technologies in the Hyper-V Cloud.

Optional: > Utilize System Center Opalis as the orchestration solution for the Hyper-V Cloud.

3.7.6 Security

The three pillars of IT security are confidentiality, integrity, and availability (CIA).

IT infrastructure threat modeling is the practice of considering what attacks might be attempted against the different components in an IT infrastructure. Generally, threat modeling assumes the following conditions:

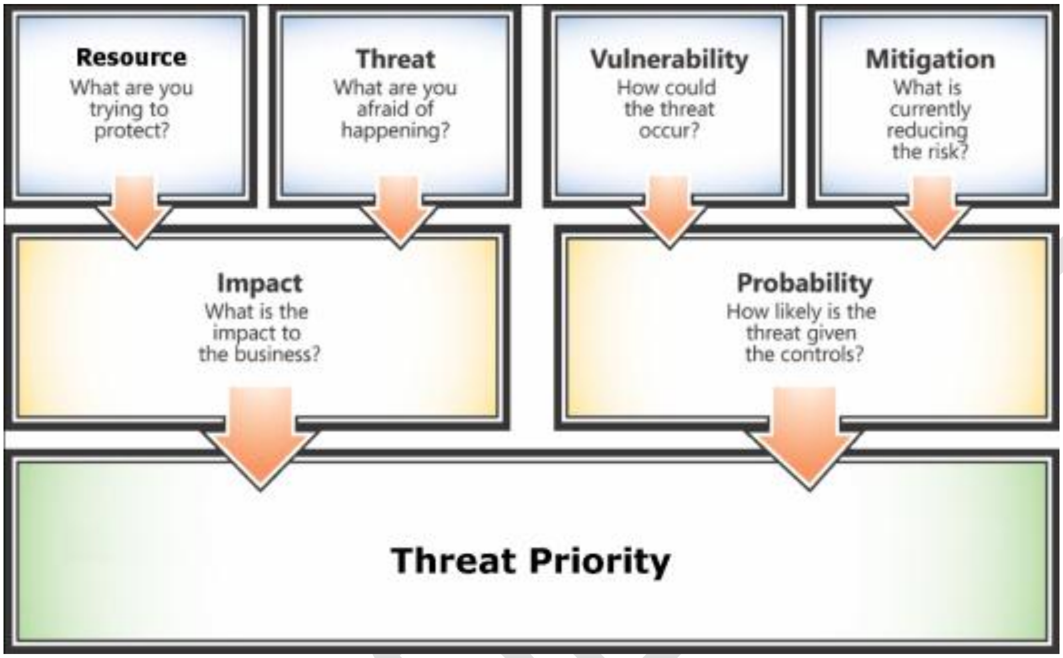
- Organizations have resources (in this case, IT components) that they wish to protect.
- All resources are likely to exhibit some vulnerabilities.
- People might exploit these vulnerabilities to cause damage or gain unauthorized access to information.
- Properly applied security countermeasures help mitigate threats that exist because of vulnerabilities.

The IT infrastructure threat modeling process is a systematic analysis of IT components that compiles component information into profiles. The goal of the process is to develop a threat model portfolio, which is a collection of component profiles.

One way to establish these pillars as a basis for threat modeling IT infrastructure is through Microsoft Operations Framework (MOF) 4.0, a framework that provides practical guidance for managing IT practices and activities throughout the entire IT lifecycle.

The Reliability Service Management Function (SMF) in the Plan Phase of MOF addresses creating plans for confidentiality, integrity, availability, continuity, and capacity, The Policy SMF in the Plan Phase provides context to help understand the reasons for policies, their creation, validation, and enforcement, and includes processes to communicate policy, incorporate feedback, and help IT maintain compliance with directives. The Deliver Phase contains several SMFs that help ensure that project planning, solution building, and the final release of the solution are accomplished in ways that fulfill requirements and create a solution that is fully supportable and maintainable when operating in production.

Figure 24. IT Infrastructure Threat Modeling


Resource
What are you trying to protect?
Threat
What are you afraid of happening?
Vulnerability
How could the threat occur?
Mitigation
What is currently reducing the risk?
Impact
What is the impact to the business?
Probability
How likely is the threat given the controls?
Threat Priority

IT Infrastructure Threat Modeling Guide
http://www.microsoft.com/downloads/en/details.aspx?FamilyID=e1d53e3f-a512-4668-85b3-169a777fc58f&displaylang=en

Security Risk Management Guide
http://technet.microsoft.com/en-us/library/cc163143.aspx

Security for the Hyper-V Cloud is founded on three pillars: Protected Infrastructure, Application Access, and Network Access.

(1) Protected Infrastructure

A defense in depth strategy is utilized at each layer of the Hyper-V Cloud architecture. Security technologies and controls must be implemented in a coordinated fashion.

An entry point represents data or process flow that traverses a trust boundary. Any portions of an IT infrastructure in which data or processes traverse from a less-trusted zone into a more-trusted zone should have a higher review priority.

Users, processes, and IT components all operate at specific trust levels that vary between fully trusted and fully untrusted. Typically, parity exists between the level of trust assigned to a user, process, or IT component and the level of trust associated with the zone in which the user, process, or component resides.

Malicious software poses numerous threats to organizations, from intercepting a user's logon credentials with a keystroke logger to achieving complete control over a computer or an entire network by using a rootkit. Malicious software can cause Web sites to become inaccessible, destroy or corrupt data, and reformat hard disks. Effects can include additional costs such as to disinfect computers, restore files, re-enter or re-create lost data. Virus attacks can also cause project teams to miss deadlines, leading to breach of contract or loss of customer confidence. Organizations that are subject to regulatory compliance can be prosecuted and fined.

A defense-in-depth strategy, with overlapping layers of security, is the best way to counter these threats, and the least-privileged user account (LUA) approach is an important part of that defensive strategy. The LUA approach ensures that users follow the principle of least privilege and always log on with limited user accounts. This strategy also aims to limit the use of administrative credentials to administrators, and then only for administrative tasks.

Mandatory: > Describe the OEM solution for securing each architecture layer of the Hyper-V Cloud.

(2) Application Access

Active Directory provides the means to manage the identities and relationships that make up the Hyper-V Cloud. Integrated with Windows Server 2008 R2, Active Directory provides out-of-the-box functionality needed to centrally configure and administer system, user, and application settings.

Windows Identity Foundation enables .NET developers to externalize identity logic from their application, improving developer productivity, enhancing application security, and enabling interoperability. Enjoy greater productivity, applying the same tools and programming model to build on-premises software as well as cloud services. Create more secure applications by reducing custom implementations and using a single simplified identity model based on claims.

Note: Contact Dell services representative for the latest security solution offerings on datacenter.

Mandatory: > Describe the OEM solution for role or claims based access to each architecture layer of the Hyper-V Cloud.

(3) Network Access

Windows Firewall with Advanced Security combines a host firewall and Internet Protocol security (IPsec). Unlike a perimeter firewall, Windows Firewall with Advanced Security runs on each computer running this version of Windows and provides local protection from network attacks that might pass through your perimeter network or originate inside your organization. It also provides computer-to-computer connection security by allowing you to require authentication and data protection for communications.

Network Access Protection (NAP) is a platform that allows network administrators to define specific levels of network access based on a client's identity, the groups to which the client belongs, and the degree to which the client complies with corporate governance policy. If a client is not compliant, NAP provides a mechanism for automatically bringing the client into compliance (a process known as remediation) and then dynamically increasing its level of network access. NAP is supported by Windows Server 2008 R2, Windows Server 2008, Windows 7, Windows Vista®, and Windows® XP with Service Pack 3 (SP3). NAP includes an application programming interface that developers and vendors can use to integrate their products and leverage this health state validation, access enforcement, and ongoing compliance evaluation.

You can logically isolate server and domain resources to limit access to authenticated and authorized computers. You can create a logical network inside an existing physical network, where computers share a common set of requirements for secure communications. In order to establish connectivity, each computer in the logically isolated network must provide authentication credentials to other computers in the isolated network to prevent unauthorized computers and programs from gaining access to resources inappropriately. Requests from computers that are not part of the isolated network will be ignored.

(a) End-point Protection (AntiVirus & AntiMalware)

The Dell solution recommends configurations utilize an existing Anti-Virus and Anti-Malware solution or utilize the Microsoft ForeFront solutions.

Mandatory: > Describe the OEM security solution
> AntiVirus solution on the host (needs to support Hyper-V and be configured according to KB961804)
> Host-based Firewall must be enabled and configured

(a) Microsoft Forefront

Microsoft Forefront delivers end-to-end security and access to information through an integrated line of protection, access and identity management products.

Forefront Security products deliver protection, access, and management solutions, built around user identity and integrated with a highly secure, interoperable platform. Our solutions help to deliver a more contextual and user-centric security solution aligned to the needs of our customers.

Multi-layered Protection – Forefront delivers leading malware protection solutions across endpoints, messaging and collaboration application servers, and the network.

Identity-based Access – Microsoft's identity-based access technologies and Forefront solutions build upon Active Directory's infrastructure to enable policy-based user access to applications, devices, and information.

Microsoft Identity and Security solutions and the Forefront product line help provide seamless protection of your IT systems through integration with the Windows platform, applications, and infrastructure. The Identity and Security platform supports heterogeneous environments, enabling 3rd party partners to share and utilize capabilities to help deliver greater value across the organization.

Forefront Identity Manager (FIM) 2010 changes the current state of identity management by providing powerful end user self-service capabilities. IT professionals are also given more tools to solve day-today tasks such as delegating administration and creating workflows for common identity management tasks. In addition, FIM 2010 is built on a .NET and WS-* based foundation for developers to build more customized and extensible solutions.

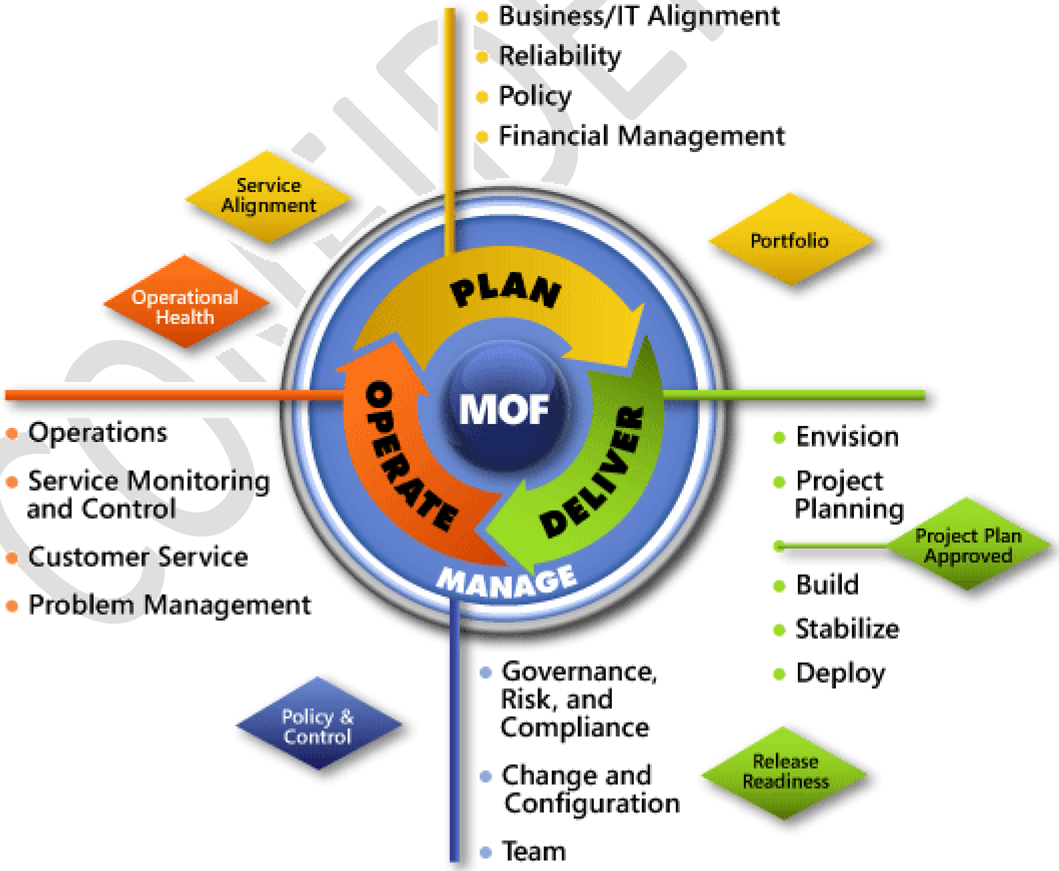
Optional: Utilize the Microsoft Forefront security solutions for the Hyper-V Cloud.

3.7.7 Service Management

The Service Management layer provides the means for automating and adapting IT service management best practices, such as those found in Microsoft Operations Framework (MOF) and the IT Infrastructure Library (ITIL), to provide built-in processes for incident resolution, problem resolution, and change control.

Microsoft Operations Framework (MOF) 4.0 provides relevant, practical, and accessible guidance for today's IT pros. MOF strives to seamlessly blend business and IT goals while establishing and implementing reliable, cost-effective IT services. MOF is a free, downloadable framework that encompasses the entire service management lifecycle. Read MOF online.

Figure 25. Microsoft Operations Framework (MOF)



Microsoft® Operations Framework (MOF) consists of integrated best practices, principles, and activities that provide comprehensive guidelines for achieving reliability for IT solutions and services. MOF provides question-based guidance that allows you to determine what is needed for your organization now, as well as activities that will keep the IT organization running efficiently and effectively in the future.

The guidance in the Microsoft Operations Framework encompasses all of the activities and processes involved in managing an IT service: its conception, development, operation, maintenance, and—ultimately—its retirement. MOF organizes these activities and processes into Service Management Functions (SMFs), which are grouped together in phases that mirror the IT service lifecycle. Each SMF is anchored within a lifecycle phase and contains a unique set of goals and outcomes supporting the objectives of that phase. An IT service's readiness to move from one phase to the next is confirmed by management reviews, which ensure that goals are achieved in an appropriate fashion and that IT's goals are aligned with the goals of the organization.

(1) System Center Service Manager 2010

System Center Service Manager 2010 delivers an integrated platform for automating and adapting IT Service Management best practices to your organization's requirements.

Service Manager can help your organization to increase productivity, reduce costs, improve resolution times, and meet compliance standards. Its built-in processes are based on industry best practices such as those found in Microsoft Operations Framework (MOF) and the IT Infrastructure Library (ITIL).

Included in Service Manager are the core process management packs for incident and problem resolution, change control, and configuration and knowledge management. In addition, our partner Provance will deliver a process management pack for IT asset management.

Through its configuration management database (CMDB) and process integration, Service Manager automatically connects knowledge and information from System Center Operations Manager, System Center Configuration Manager, and Active Directory Domain Services.

Service Manager delivers multiple benefits in the following key areas:

- User-centric support. It can improve user productivity and satisfaction while reducing support costs with its Self-Service Portal.
- Datacenter management efficiency. With its CMDB and management packs, it helps reduce downtime and improve the reliability of IT services running within your datacenter.
- Business alignment. It helps the organization align to business goals and adapt to new requirements through asset management, compliance and risk management, and automated reporting and analysis.

🛑 **Mandatory:** > Describe the OEM solution for service management including CMDB, change, configuration, release, incident, and problem management in the Hyper-V Cloud OR REMOVE SECTION

⚠ **Recommended:** > Utilize System Center Service Manager 2010.

4 Validation Checklist

Cases