FlexPod with Microsoft Private Cloud

Architecture Overview for FlexPod with Microsoft Windows Server 2008 R2 and Microsoft System Center 2012

April 2012

© 2012 Cisco. All rights reserved. This document is Cisco Public Information.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 Introduction

Microsoft® Private Cloud Fast Track is a joint effort between Microsoft and its hardware partners that helps organizations quickly develop and implement private clouds while reducing both cost and risk. The solution provides a reference architecture that combines Microsoft software, consolidated guidance, and validated configurations with partner technology—such as computing power, network, and storage architectures—as well as value-added software components.

The private cloud model provides much of the efficiency and agility of cloud computing along with the increased control and customization achieved through dedicated private resources. With the FlexPod™ with Microsoft Private Cloud solution, Microsoft and its hardware partners provide organizations both the control and the flexibility required to reap the full benefits of the private cloud.

2 FlexPod with Microsoft Private Cloud Solution Description

The FlexPod with Microsoft Private Cloud solution is a joint reference implementation for building private clouds that combines Microsoft software, consolidated guidance, and validated configurations with hardware partner technology, including computing power, network and storage architectures, and value-added software components. FlexPod with Microsoft Private Cloud is a joint submission to the Fast Track program by NetApp and Cisco. This document describes the solution architecture in detail and includes content from NetApp, Cisco, and Microsoft.

FlexPod with Microsoft Private Cloud utilizes the core capabilities of Microsoft Windows Server®, Microsoft Hyper-V™, and System Center to deliver a private cloud infrastructure-as-a-service (IaaS) offering. The key software components of every reference implementation are Windows Server 2008 R2 SP1, Hyper-V, and System Center 2012. This solution also includes hardware and software from Cisco and NetApp to form a complete solution that is ready for your enterprise.

2.1 Business Value

The FlexPod with Microsoft Private Cloud solution provides reference architecture for building private clouds on each organization's unique terms. Each FlexPod with Microsoft Private Cloud solution helps organizations implement private clouds with increased ease and confidence. Among the benefits of the solution are faster deployment, reduced risk, and a lower cost of ownership.

Risk is reduced, because the solution:

- Is tested for end-to-end interoperability of compute, storage, and network
- Offers predefined, out-of-box solutions based on a common cloud architecture that has already been tested and validated
- Provides a high degree of service availability through automated load balancing

Lower cost of ownership benefits include:

- A cost-optimized platform and software-independent solution for rack system integration
- High performance and scalability with Windows Server 2008 R2 operating system's advanced platform editions of Hyper-V technology
- Minimized backup times and fulfilled recovery time objectives for each business-critical environment

2.2 Technical Benefits

The FlexPod with Microsoft Private Cloud solution integrates best-in-class hardware implementations with Microsoft software to create a reference implementation. This solution was co-developed by Microsoft, NetApp, and Cisco, and it has gone through a validation process. Because this is a reference implementation, Microsoft and its hardware partners have taken the work of building a private cloud that is ready to meet a customer's needs.

Faster deployment is made possible by:

- End-to-end architectural and deployment guidance
- Streamlined infrastructure planning due to predefined capacity
- Enhanced functionality and automation through deep knowledge of infrastructure
- Integrated management for virtual machine and infrastructure deployment
- Self-service portal for the rapid and simplified provisioning of resources

2.3 Program Requirements and Validation

The FlexPod with Microsoft Private Cloud solution is composed of three pillars: engineering, marketing, and enablement. These three pillars drive the creation of reference implementations, making them public and finally making them available for customers to purchase. This reference architecture is one step in the engineering phase of the program toward the validation of a reference implementation.

3 Technical Overview

The Fast Track program is administered by Microsoft. Because of this, the solution is a fully validated design. All three parties have validated their components to verify that the system is high quality and consistent in operation.

3.1 NIST Definition of Cloud Computing

Note: The following information is copied verbatim from the NIST Definition of Cloud Computing v15.

> Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model is composed of five essential characteristics, three service models, and four deployment models.
>
> **Essential Characteristics:**
>
> *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 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, tablets, laptops, and workstations).
>
> *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, and network bandwidth.
>
> *Rapid elasticity*. Capabilities can be elastically provisioned and released, in some cases automatically, to scale rapidly outward and inward commensurate with demand. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be appropriated in any quantity at any time.

Measured service. Cloud systems automatically control and optimize resource use by leveraging a metering capability[1] 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.

Service Models:

Software as a Service (SaaS). The capability provided to the consumer is to use the provider's applications running on a cloud infrastructure[2]. The applications are accessible from various client devices through either a thin client interface, such as a web browser (e.g., web-based email), or a program interface. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings.

Platform as a Service (PaaS). The capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created or acquired applications created using programming languages, libraries, services, and tools supported by the provider.[3] The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly configuration settings for the application-hosting environment.

Infrastructure as a Service (IaaS). The capability provided to the consumer is to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications. The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, and deployed applications; and possibly limited control of select networking components (e.g., host firewalls).

Deployment Models:

Private cloud. The cloud infrastructure is provisioned for exclusive use by a single organization comprising multiple consumers (e.g., business units). It may be owned, managed, and operated by the organization, a third party, or some combination of them, and it may exist on or off premises.

Community cloud. The cloud infrastructure is provisioned for exclusive use by a specific community of consumers from organizations that have shared concerns (e.g., mission, security requirements, policy, and compliance considerations). It may be owned, managed, and operated by one or more of the organizations in the community, a third party, or some combination of them, and it may exist on or off premises.

Public cloud. The cloud infrastructure is provisioned for open use by the general public. It may be owned, managed, and operated by a business, academic, or government organization, or some combination of them. It exists on the premises of the cloud provider.

Hybrid cloud. The cloud infrastructure is a composition of two or more distinct cloud infrastructures (private, community, or public) that remain unique entities, but are bound together by standardized or

---

[1] Typically this is done on a pay-per-use or charge-per-use basis.

[2] A cloud infrastructure is the collection of hardware and software that enables the five essential characteristics of cloud computing. The cloud infrastructure can be viewed as containing both a physical layer and an abstraction layer. The physical layer consists of the hardware resources that are necessary to support the cloud services being provided, and typically includes server, storage, and network components. The abstraction layer consists of the software deployed across the physical layer which manifests the essential cloud characteristics. Conceptually, the abstraction layer sits above the physical layer.

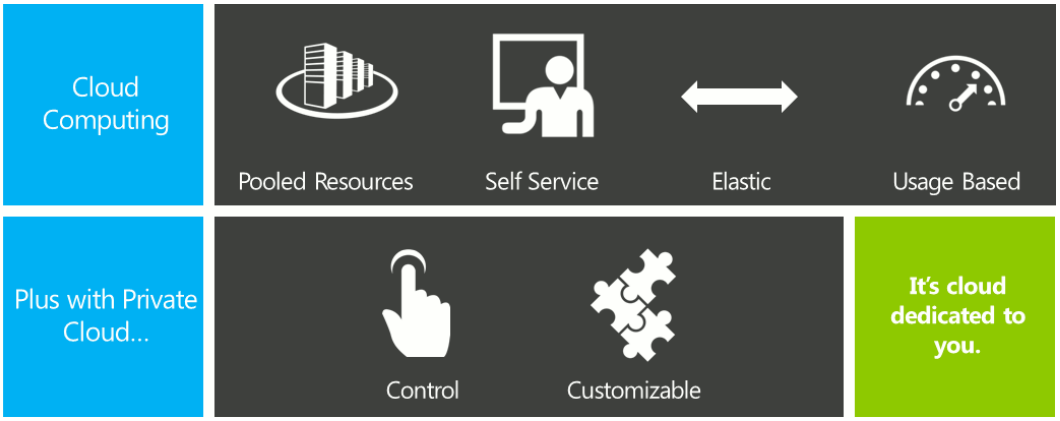
[3] This capability does not necessarily preclude the use of compatible programming languages, libraries, services, and tools from other sources.

proprietary technology that enables data and application portability (e.g., cloud bursting for load balancing between clouds).

3.2 Microsoft Private Cloud Overview

Private cloud is a computing model that uses resources that are dedicated to your organization. A private cloud shares many of the characteristics of public cloud computing including, resource pooling, self-service, elasticity, and metered use, and it is delivered in a standardized manner with additional control and customization available from dedicated resources.

Figure 1) Private cloud attributes (graphic provided by Microsoft).



Although virtualization is an important technological component of private cloud, the key difference is the continued abstraction of computing resources from infrastructure and the machines (virtual or otherwise) used to deliver those resources. Only by delivering this abstraction can customers achieve the benefits of private cloud, which include improved agility and responsiveness, reduced total cost of ownership (TCO), and increased business alignment and focus. Most importantly, a private cloud promises to exceed the cost effectiveness of a virtualized infrastructure through higher workload density and greater resource utilization.

The Microsoft Private Cloud is a unique and comprehensive offering that is built on four key pillars:

- **All about the app.** An application-centric cloud platform that helps you focus on business value.
- **Cross-platform from the metal up**. Cross-platform support for multiple hypervisor environments, operating systems, and application frameworks.
- **Foundation for the future.** Microsoft Private Cloud lets you go beyond virtualization to a true cloud platform.
- **Cloud on your terms.** The ability to consume cloud on your terms, providing you the choice and flexibility of a hybrid cloud model through common management, virtualization, identity, and developer tools.

For more information, refer to the Microsoft Private Cloud Overview.

3.3 Private Cloud Architecture Principles

In accordance with the NIST model core, the required elements must be present for this solution to be considered a private cloud. Implementation of the required elements is discussed in this section.

Resource Pooling

Resource optimization is a principle that drives efficiency and cost reduction. It 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 Microsoft Private Cloud principles and ultimately helps drive down costs and improve agility.

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 proactive so that requests can be satisfied on demand. Applying this principle reactively and in isolation often leads to the inefficient use of resources and unnecessary costs. Combined with other principles, such as incenting desired consumer behavior, this principle allows there to be a balance between the cost of unused capacity and the desire for agility.

Perception of Continuous Availability

From the consumer's perspective, cloud services appear to be continuously available when needed. The consumer shouldn't experience an interruption of that service, even if failures occur within the 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 Microsoft Private Cloud principles.

Drive Predictability

Predictability is a fundamental principle from all cloud perspectives, whether you are a consumer or a 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.

To achieve this predictability, a provider must deliver an underlying infrastructure that assures a consistent experience to the hosted workloads. 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.

Metering and 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 and increased costs due to duplicate infrastructure, and often it 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 deliver infrastructure transforms IT's approach. Because infrastructure is provided as a service, IT can now leverage a shared resource model that enables economies of scale and when, combined with the other principles, achieves greater agility in providing services.

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 that provides servers to multiple customer organizations. Increasingly, this is also a model being used by an individual

organization with a centralized IT organization that provides services to multiple business or organizational units within the organization and that treats each as a customer or tenant.

Security and Identity

Security for the Microsoft Private Cloud is founded on three pillars: protected infrastructure, application access, and network access.

Protected infrastructure leverages security technologies as well as identity technology to make sure that hosts, information, and applications are secured across all scenarios in the data center, including physical (on premises) and virtual (on premises and in the cloud).

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

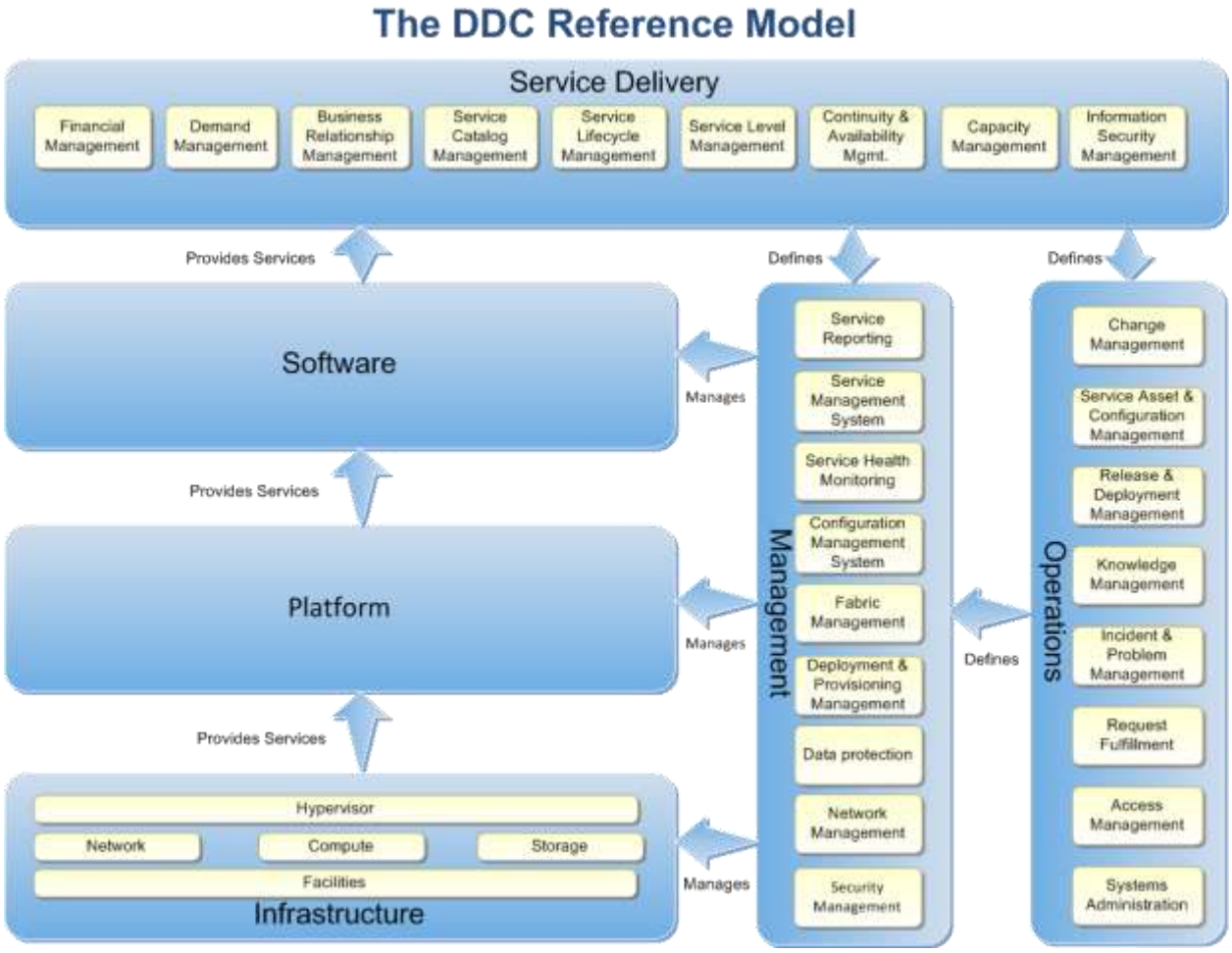
Network access uses an identity-centric approach to make sure that users in local or remote locations have secure access on numerous devices to maintain productivity.

The most important aspect is that the secure data center leverages a common integrated technology to make sure that users have simple access with a common identity and that management is integrated across physical, virtual, and cloud environments so business can take advantage of all capabilities without requiring additional significant financial investments.

3.4 Dynamic Datacenter IaaS Reference Model

IaaS is the application of private cloud architecture principles to deliver infrastructure. As the cloud ecosystem matures, products broaden and deepen in features and capabilities. We can use the dynamic datacenter (DDC) reference model in Figure 2 as a beacon to make sure we are delivering a holistic solution that spans all the layers required for mature IaaS. The model is a guide to assist architects' efforts to holistically address the development of private cloud architecture. This model is for reference only. Some pieces are emphasized more than others in the technical reference architecture based on experience with operating private clouds in real-world environments.

Figure 2) DDC reference model: IaaS view (graphic provided by Microsoft).



The reference model is split into the following layers:

- **Software, platform, and infrastructure.** The software, platform, and infrastructure layers represent the technology stack in which each layer provides services to the layer above.
- **Operations and management.** The operations and management layers represent the process perspective and include the management tooling required to implement aspects of the process.
- **Service delivery.** The service delivery layer represents the alignment between business and IT.

Technology and process perspectives, such as the Information Technology Infrastructure Library (ITIL) and Microsoft Operations Framework (MOF), are deliberately blended, because cloud computing is as much about the service management as it is about the technologies involved in it.

For more information, refer to the Private Cloud Reference Model.

3.5 Conceptual Architecture

Enabling complex workflow and automation to be developed over time is one of the key drivers of the layered approach to infrastructure architecture. This is accomplished 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.

Fabric

In this solution, we refer to the core infrastructure elements, such as servers, storage, and network, as fabric. These elements are discussed in this section.

Scale Units

In a modular architecture, the concept of a scale unit refers to the point to which a module in the architecture can scale before another module is needed. For example, an individual server is a scale unit. It can be expanded to a certain point in terms of CPU and RAM, but beyond its maximum capacities, an additional server is required to continue scaling. Each scale unit also requires a certain amount of physical installation labor, configuration labor, and so on. With large-scale units, such as a preconfigured 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 optimum scale units as input to the overall architecture. Scale units enable the documentation of all of the requirements that are required for implementation, such as space, power, HVAC, connectivity, and so on.

Servers

The hardware-architecture choices available to data center 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 the number and speed of CPU cores, the maximum amount and speed of RAM, the number and type of expansion slots, and so on. Particularly important are the number and type of onboard input/output (I/O) ports, as well as the number and type of supported I/O cards. Both Ethernet and Fibre Channel (FC) expansion cards often provide multiport options, whereas a single card can have only four ports. Additionally, in blade server architectures, there are often limitations on the amount of I/O cards and supported combinations used.

Note: Be aware of these limitations, as well as the oversubscription ratio between blade I/O ports and any blade chassis switch modules.

A single server is not typically a good scale unit for a private cloud solution, because of the amount of overhead required to install and configure an individual server.

Storage

Storage architecture is a critical design consideration for private cloud solutions. The topic is challenging, because it is rapidly evolving in terms of new standards, protocols, and implementations. Storage and supporting storage networking are critical to both the overall performance of the environment and the overall cost, because storage tends to be one of the more costly items.

Storage architectures today have several layers, including the storage arrays; the storage network; the storage protocol; and, for virtualization, the file system using 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 on a large scale requires tight integration with the storage architecture and robust automation. Provisioning a new virtual machine on an already existing logical unit number (LUN) is a simple operation. However, provisioning a new LUN, adding it to a host cluster, and so on are relatively complicated tasks that can also greatly benefit from automation.

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 other loop avoidance technologies.

A dedicated management network is a common feature of advanced data center virtualization solutions. Most virtualization vendors recommend that hosts be managed through a dedicated network to avoid competition with guests' traffic needs and to provide a degree of separation for security and ease of management. This typically means dedicating a network interface card (NIC) for each host and port per network device in the management network.

With advanced data center virtualization, a frequent use case is to provide isolated networks in which different owners, such as particular departments or applications, are provided with their own dedicated networks. Multi-tenant networking refers to using technologies such as virtual LANs (VLANs) or Internet Protocol Security (IPsec) isolation techniques to provide dedicated networks that utilize a single network infrastructure or wire.

Managing the network environment in an advanced data center 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 guests, as well as the IPsec policies that control network communication.

For VLAN-based network segmentation, several components, including the host servers, host clusters, virtual machine manager (VMM), and 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 for a virtual machine to be able to fail over to any node and maintain its connection to the network. On a large scale, this can be accomplished through Windows PowerShell™ scripting.

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 virtual machines from one server to another with zero downtime and many other features that are provided by hypervisor-based virtualization technologies provides a rich set of capabilities. These capabilities can be used 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.

Management

The core difference between virtualization and private cloud is management. Private clouds require a level of automation and self-service that goes beyond what is required for traditional virtualized environments. The overall management architecture is described in this section.

Fabric Management

Fabric management is the concept of treating discrete capacity pools of compute (servers), storage, and networks as a single fabric. The fabric is then subdivided into capacity clouds or resource pools, which carry certain characteristics, such as delegation of access and administration, service-level agreements (SLAs), cost metering, on so on. This enables the centralization and automation of complex management

functions to be carried out in a highly standardized and repeatable fashion that increases availability and lowers operational costs.

Process Automation and Orchestration

Orchestration that manages all of the automation and management components must be implemented as the interface between the IT organization and the infrastructure. Orchestration provides the bridge between IT business logic, such as "deploy a new Web server virtual machine when capacity reaches 85%," and the dozens of steps in an automated workflow that are required to actually implement such a change.

Ideally, the orchestration layer provides a GUI in which complex workflows, which consist of events and activities across multiple management system components, can be combined 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.

Service Management System

A service management system is a set of tools designed to facilitate service management processes. Ideally, these tools should be able to integrate data and information from an entire set of tools found in the management layer. It should also be able to process and present the data as needed. At a minimum, the service management system should be linked to the configuration management system (CMS), commonly known as the configuration management database (CMDB), and it should log and track incidents, problems, and changes. It is also preferred that the service management system be integrated with the service health modeling system so that incident tickets can be automatically generated.

User Self-Service

Self-service capability is a characteristic of private cloud computing and must be available in any implementation. The intent is to permit users to approach a self-service capability and be presented with options for provisioning in an organization. The capability can be basic, such as provisioning a virtual machine with a predefined configuration, or it can be more advanced to allow more complex configuration scenarios such as Web-farm provisioning or the deployment of best practice guidance.

Self-service capability is a critical business driver that enables members of an organization to respond quickly to business needs with IT capabilities and to meet those needs in a way that aligns and conforms with internal business IT requirements and governance.

This means the interface between IT and the business is abstracted to a simple, well-defined, and approved set of service options that are presented as a menu in a portal or are available from the command line. The business can select these services from the catalog, start the provisioning process, and be notified when the services complete. The business is then charged only for what it actually uses.

This is analogous to the capability available on public cloud platforms.

Service Delivery

Unlike traditional virtualized environments, private clouds are delivered as a service. This implies that you have a managed service delivery capability. What services do you offer, and what features or elements do they contain? Management of service delivery is discussed in this section.

Service Catalog

Service catalog management involves defining and maintaining a catalog of services offered to consumers. The catalog lists the following:

- The classes of service available
- The requirements to be eligible for each service class

- The service-level attributes and targets that are included with each service class, such as availability targets
- The cost model for each service class

The service catalog can also include specific virtual machine templates, such as a high-compute template, that are designed for different workload patterns. Each template defines the virtual machine configuration specifics such as the amount of allocated CPU, memory, and storage.

Capacity Management

Capacity management defines the processes that are necessary to achieve the perception of infinite capacity. Capacity must be managed to meet existing and future peak demand while controlling underutilization. Business relationship and demand management are key inputs into effective capacity management, and they require a service provider's approach. Predictability and optimization of resource use are primary principles in achieving capacity management objectives.

Availability Management

Availability management defines the processes that are necessary to achieve the perception of continuous availability. Continuity management defines how risk will be managed in a disaster scenario to make sure minimum service levels are maintained. The principles of resiliency and automation are fundamental here.

Service-Level Management

Service-level management is the process of negotiating SLAs and making sure the agreements are met. SLAs define target levels for cost, quality, and agility by service class, as well as the metrics for measuring actual performance. Managing SLAs is necessary to achieve the perception of infinite capacity and continuous availability. This, too, requires a service provider's approach by IT.

Service Lifecycle Management

Service lifecycle management takes an end-to-end management view of a service. A typical journey starts with the identification of a business need through business relationship management and continues to the time when the service becomes available. Service strategy drives service design. After launch, the service is transitioned to operations and refined through continual service improvement. Taking a service provider's approach is the key to successful service lifecycle management.

Operations

After the system is deployed, it must be operated correctly. The processes and tools described in this section help support the proper post-deployment operation of the overall system.

Change Management

The change management process is responsible for controlling the lifecycle of all changes. The primary objective of change management is to eliminate or minimize disruption while desired changes are made to services. Change management focuses on understanding and balancing the cost and risk of making the change versus the benefit of the change to either the business or the service. Driving predictability and minimizing human involvement are the core principles for achieving a mature service management process and making sure changes can be made without affecting the perception of continuous availability.

Incident and Problem Management

The goal of incident management is to resolve events that affect, or threaten to affect, services as quickly as possible with minimal disruption. The goals of problem management are to identify and resolve root

causes of incidents that have occurred, as well as to identify and prevent or minimize the effect of incidents that might occur.

Configuration Management

Configuration management is the process that verifies that the assets required to deliver services are properly controlled and that accurate and reliable information about those assets is available when and where it is needed. This information includes details about how the assets have been configured and the relationships between assets.

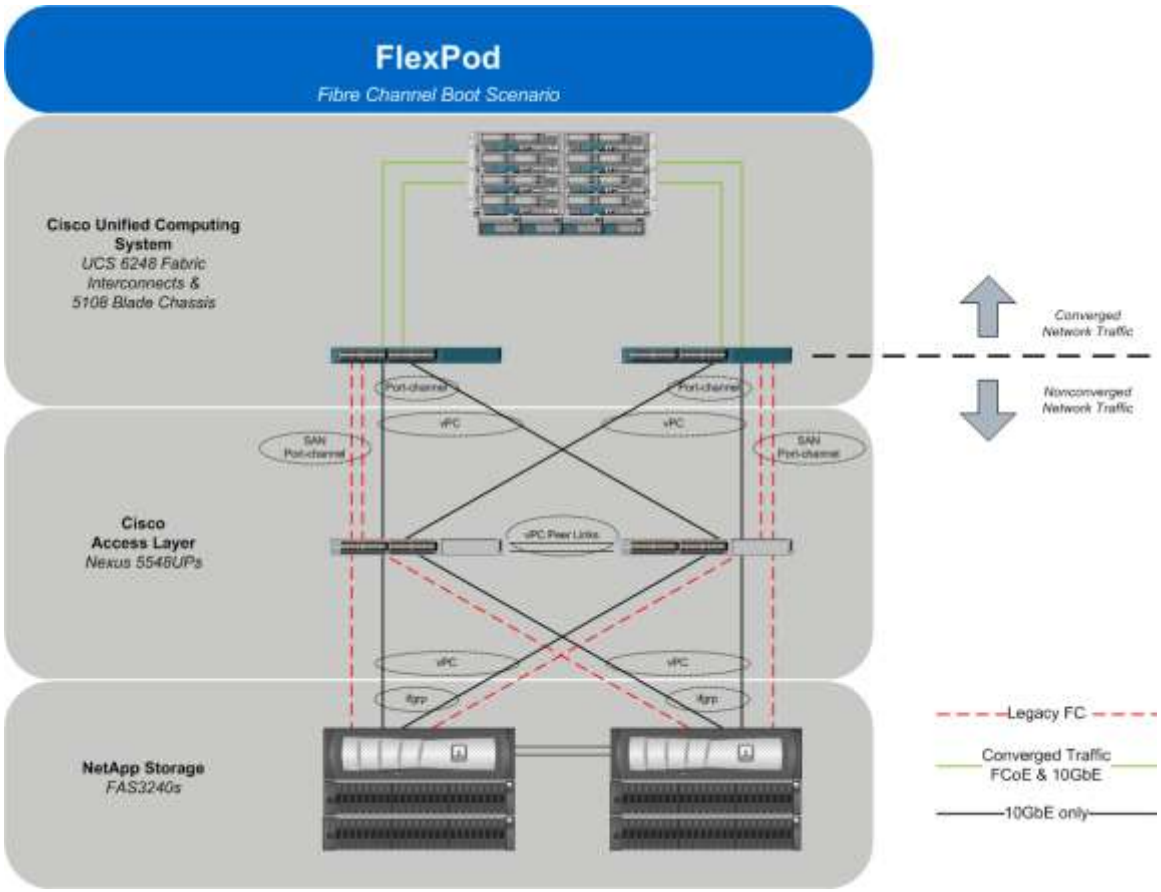
This typically requires a CMDB, which is a database that is used to store configuration records throughout their lifecycle. The CMS maintains one or more configuration management databases. Each database stores the attributes of configuration items and their relationships with other configuration items.

4 Reference Architecture

FlexPod architecture is highly modular, or pod-like. Although each customer's FlexPod unit might vary in its exact configuration, after a FlexPod unit is built, it can easily be scaled as requirements and demands change. This includes both scaling up (adding additional resources within a FlexPod unit) and scaling out (adding additional FlexPod units).

Specifically, FlexPod is a defined set of hardware and software that serves as an integrated foundation for all virtualization solutions. FlexPod validated with Microsoft Private Cloud includes NetApp® FAS3200 series storage, Cisco Nexus® 5500 Series network switches, the Cisco Unified Computing Systems™ (Cisco UCS™) platforms, and Microsoft virtualization software in a single package. The computing and storage can fit in one data center rack with networking residing in a separate rack or deployed according to a customer's data center design. Due to port density, the networking components can accommodate multiple configurations of this kind.

Figure 3) Architecture overview (graphic provided by Cisco).



The reference configuration shown in Figure 3 includes:

- Two Cisco Nexus 5548 switches
- Two Cisco UCS 6248 fabric interconnects
- One chassis of Cisco UCS blades with two fabric extenders per chassis
- One FAS3240A (HA pair)

Storage is provided by a NetApp FAS3240A (HA configuration within a single chassis) with accompanying disk shelves. All systems and fabric links feature redundancy and provide end-to-end high availability. For server virtualization, the deployment includes Hyper-V. Although this is the base design, each of the components can be scaled flexibly to support specific business requirements. For example, more (or different) blades and chassis could be deployed to increase compute capacity, additional disk shelves could be deployed to improve I/O capacity and throughput, or special hardware or software features could be added to introduce new features.

Note: This is a sample bill of materials (BoM) only. This solution is certified for the hardware series (that is, FAS3200 series) rather than for a specific model. FlexPod and Fast Track programs allow customers to choose from within a model family to make sure that each FlexPod for Microsoft Private Cloud solution meets the customers' requirements.

The remainder of this document guides you through the low-level steps for deploying the base architecture, as shown in Figure 3. This includes everything from physical cabling, to compute and storage configuration, to configuring virtualization with Hyper-V.

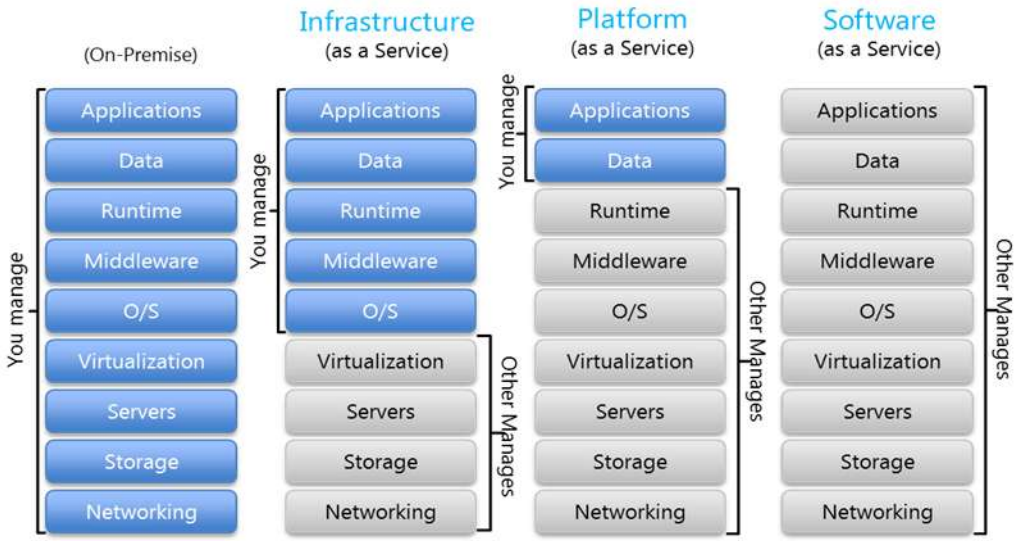
4.1 Use Cases

With any private cloud, the number of workloads that can be run is nearly unlimited. Part of the flexibility of clouds is the ability to host any random workload required by the business. However, there are some key scenarios or use cases that can be used for planning purposes.

Service Models

Figure 4 shows the taxonomy of cloud services and defines the separation of responsibilities for each service model. For more information on the service models, refer to section 3.1.

Figure 4) Taxonomy of cloud services (graphic supplied by Microsoft).



Infrastructure as a Service

IaaS abstracts hardware (such as servers, storage, and network infrastructure) into a pool of computing, storage, and connectivity capabilities that are delivered as services for a usage-based (metered) cost. The goal of IaaS is to provide a flexible, standard, and virtualized operating environment that can become a foundation for PaaS and SaaS.

IaaS usually provides a standardized virtual server. The consumer takes responsibility for the configuration and operations of the guest operating system, software, and database. Compute capabilities such as performance, bandwidth, and storage access are also standardized.

Service levels cover the performance and availability of the virtualized infrastructure. The consumer takes on the operational risk that exists above the infrastructure.

FlexPod with Microsoft Private Cloud primarily aims to deliver IaaS and to enable PaaS and SaaS.

Data Center Consolidation and Virtualization

Data center consolidation and virtualization allow enterprise customers to migrate physical machines and virtual machines to Hyper-V technology virtualization and cloud environments based on Hyper-V technology to reduce capital and operational expenses. Data center consolidation and virtualization also improve the manageability of both virtual and physical environments by using the Microsoft System Center family of products.

The goals of consolidating and virtualizing the data center include:

- Reducing the cost of facilities, hardware, and the licensing of alternative solutions by deploying a standardized Hyper-V network and storage infrastructure
- Reducing server sprawl and implementing a more holistic and robust management solution
- Transitioning from organically grown virtualized environments to a private cloud solution to implement new capabilities and business

Virtual Desktop Infrastructure

A 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 are. VDI also enables IT to build a more agile and efficient IT infrastructure. Flexible Windows® desktop scenarios give organizations the ability to choose the client computing scenarios that meet the unique needs of their businesses.

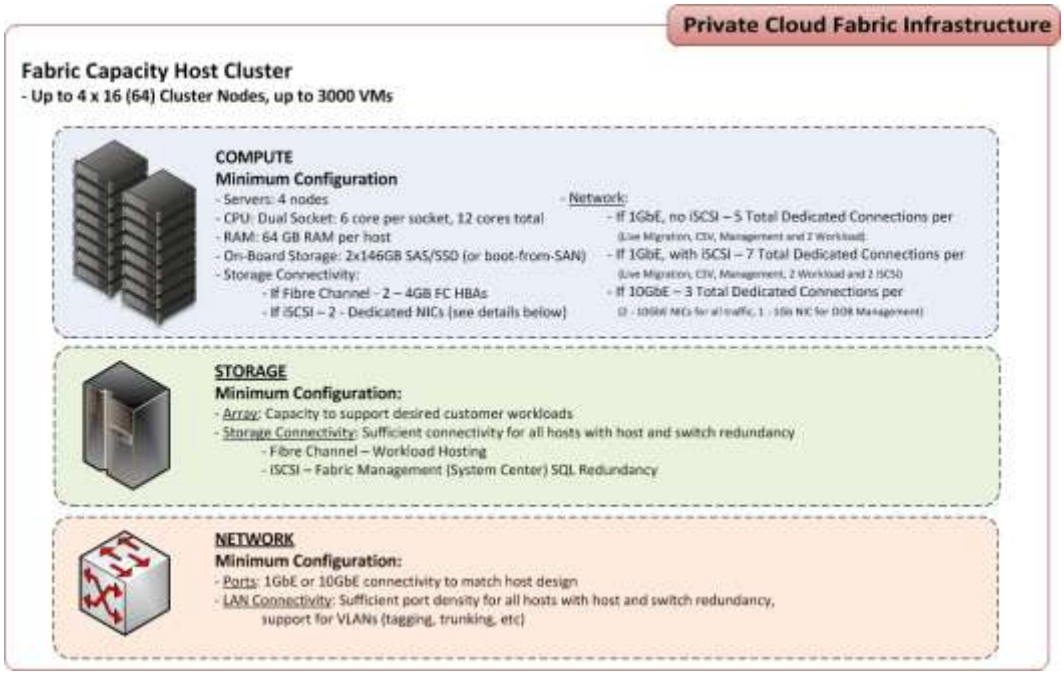
4.2 Fabric Logical Architecture

The logical architecture is composed of two parts. The first is the fabric, which is the physical infrastructure (servers, storage, and network) that hosts and runs all customer and consumer virtual machines. The second is fabric management, which is a set of virtual machines that make up the Microsoft SQL Server® and System Center management infrastructure. The joint best practice is to have two or more Hyper-V host servers in a dedicated host cluster for the fabric management virtual machines and separate clusters for the fabric. For smaller scale deployments, the fabric management virtual machines can be hosted on the fabric itself.

Fabric

Figure 5 shows minimum high-level requirements for the fabric. The requirements are categorized in compute, storage, and network layers. The minimum requirements and recommendations are designed to balance cost versus density and performance.

Figure 5) Fabric capacity host cluster minimum requirements (graphic provided by Microsoft).



4.3 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 (SMP) servers. The primary tenet to consider when selecting system architectures is that each Hyper-V host contains multiple guests with multiple workloads. Processor, RAM, storage, and network capacity are critical, as well as high I/O capacity and low latency.

Note: Capacity planning and management are required to verify that the services running within the cloud do not exceed the platform's capacity.

Server and Blade Design

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.

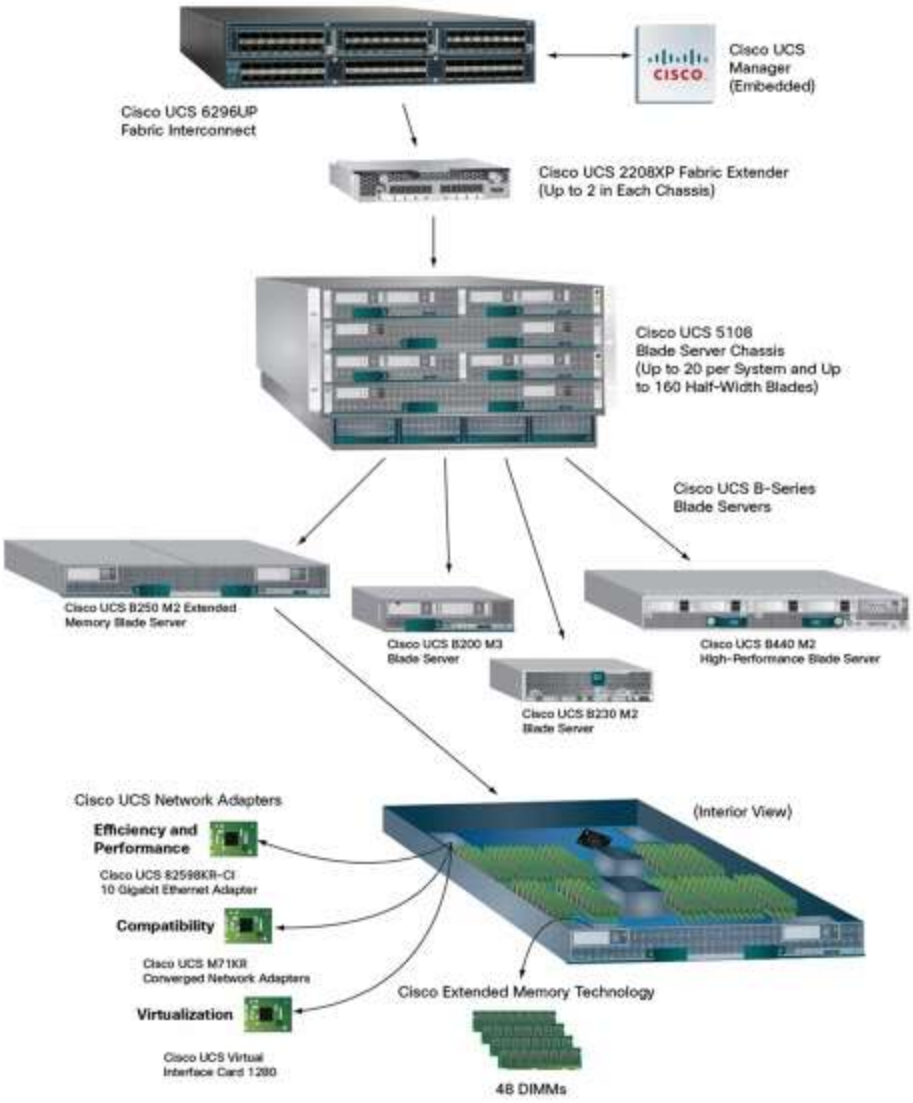
The system architecture of the host server refers to the general category of the server hardware itself. The primary factor to consider when selecting system architectures is that each Hyper-V host contains multiple guests with multiple workloads. The critical factors are processor, RAM, storage, network capacity, high I/O capacity, and low latency. The host server must be able to provide the required capacity in each of these categories.

Cisco Unified Computing System Overview

The Cisco Unified Computing System is a next-generation data center platform that unites computing, networking, storage access, and virtualization resources into a cohesive system designed to reduce TCO and increase business agility. The system integrates a low-latency, lossless 10 Gigabit Ethernet (10GbE) unified network fabric with enterprise-class, x86-architecture servers. The system is an integrated, scalable, multichassis platform in which all resources participate in a unified management domain (Figure 6).

Figure 6) The Cisco Unified Computing System is a highly available cohesive architecture (graphic provided by Cisco).

Figure 7) Microsoft Private Cloud logical architecture (graphic provided by Cisco).



Product Overview

Cisco UCS 6200 Series fabric interconnects are a core part of the Cisco Unified Computing System, providing both network connectivity and management capabilities for the system (Figure 7). The Cisco UCS 6200 Series offers line-rate, low-latency, lossless 10GbE; Fibre Channel over Ethernet (FCoE); and FC functions.

The Cisco UCS 6200 Series provides the management and communication backbone for the Cisco UCS B-Series blade servers and the 5100 Series blade server chassis. All chassis, and therefore all blades, attached to the Cisco UCS 6200 Series fabric interconnects become part of a single, highly available management domain. In addition, by supporting unified fabric, the Cisco UCS 6200 Series provides both LAN and storage area network (SAN) connectivity for all blades within its domain.

From a networking perspective, the Cisco UCS 6200 Series uses a cut-through architecture that supports deterministic low-latency, line-rate 10GbE on all ports; switching capacity of 2 terabits (Tb); and 320 Gbps bandwidth per chassis, independent of packet size and enabled services. The product family supports Cisco® low-latency, lossless 10GbE unified network fabric capabilities that increase the reliability, efficiency, and scalability of Ethernet networks. The fabric interconnect supports multiple traffic classes

over a lossless Ethernet fabric from the blade through the interconnect. Significant TCO savings come from an FCoE-optimized server design in which NICs, host bus adapters (HBAs), cables, and switches can be consolidated.

Unified Fabric with FCoE: I/O Consolidation

The Cisco UCS 6200 Series is built to consolidate LAN and SAN traffic onto a single unified fabric, which saves the capital expenditures (capex) and operating expenses (opex) associated with multiple parallel networks, different types of adapter cards, switching infrastructure, and cabling within racks. Unified ports support allows either base or expansion module ports in the interconnect to support direct connections from Cisco UCS to existing native FC SANs. The capability to connect FCoE to native FC protects existing storage system investments while dramatically simplifying in-rack cabling.

Cisco UCS Manager

The Cisco UCS 6200 Series hosts and runs Cisco UCS Manager in a highly available configuration that enables the fabric interconnects to fully manage all Cisco UCS elements. Connectivity to the Cisco UCS 5100 Series blade chassis is maintained through the Cisco UCS 2100 or 2200 Series fabric extenders in each blade chassis.

The Cisco UCS 6200 Series interconnects support out-of-band management through a dedicated 10/100/1000Mbps Ethernet-management port, as well as in-band management. Cisco UCS Manager typically is deployed in a clustered active-passive configuration on redundant fabric interconnects that are connected through dual 10/100/1000 Ethernet clustering ports.

Optimization for Virtualization

For virtualized environments, the Cisco UCS 6200 Series supports Cisco virtualization-aware networking and Cisco Data Center Virtual Machine Fabric Extender (VM-FEX) architecture. Cisco Data Center VM-FEX allows the interconnects to provide policy-based VM connectivity with network properties that move with the virtual machine and a consistent operational model for both physical and virtual environments.

Cisco UCS 6248UP 48-Port Fabric Interconnect

The Cisco UCS 6248UP 48-port fabric interconnect (Figure 8) is a one-rack unit (1RU) 10GbE, FCoE, and FC switch that offers up to 960Gbps throughput and up to 48 ports. The switch has 32 1/10Gbps fixed Ethernet, FCoE, and FC ports and one expansion slot.

Figure 8) Cisco UCS 6248UP 48-port fabric interconnect (graphic provided by Cisco).

Cisco UCS 6296UP 96-Port Fabric Interconnect

The Cisco UCS 6296UP 96-port fabric interconnect (Figure 9) is a 2RU 10GbE, FCoE and native FC switch that offers up to 1920Gbps throughput and up to 96 ports. The switch has 48 1/10Gbps fixed Ethernet, FCoE, and FC ports and three expansion slots.

Figure 9) Cisco UCS 6296UP 96-port fabric interconnect (graphic provided by Cisco).

Expansion Module Option for Cisco UCS 6200 Series Fabric Interconnects

The Cisco UCS 6200 Series supports an expansion module that can be used to increase the number of 10GbE, FCoE, and FC ports (Figure 10). This unified port module provides up to 16 ports that can be configured for 10GbE, FCoE, and 1/2/4/8Gbps native FC using the SFP or SFP+3 interface for transparent connectivity with existing FC networks.

Figure 10) Unified port 16-port expansion module (graphic provided by Cisco).

Figure 11) Cisco UCS 5108 blade server chassis with blade servers front and back (graphic provided by Cisco).

Server Design Principles

The server design must provide a high level of availability throughout. This includes, for example, features such as redundant power distribution units (PDUs), storage path, networking, and disks. To provide this level of availability, it is necessary to use two or more storage controllers to support multipathing on the I/O side.

Use multiple network adapters, multiport network adapters, or both on each host server. For converged designs, network technologies that provide teaming or virtual NICs can be used if redundancy is provided through the use of NIC teaming or a hardware-level failover solution and if multiple virtual NICs or VLANs can be presented to the hosts for traffic segmentation and bandwidth control. The following network connections are required:

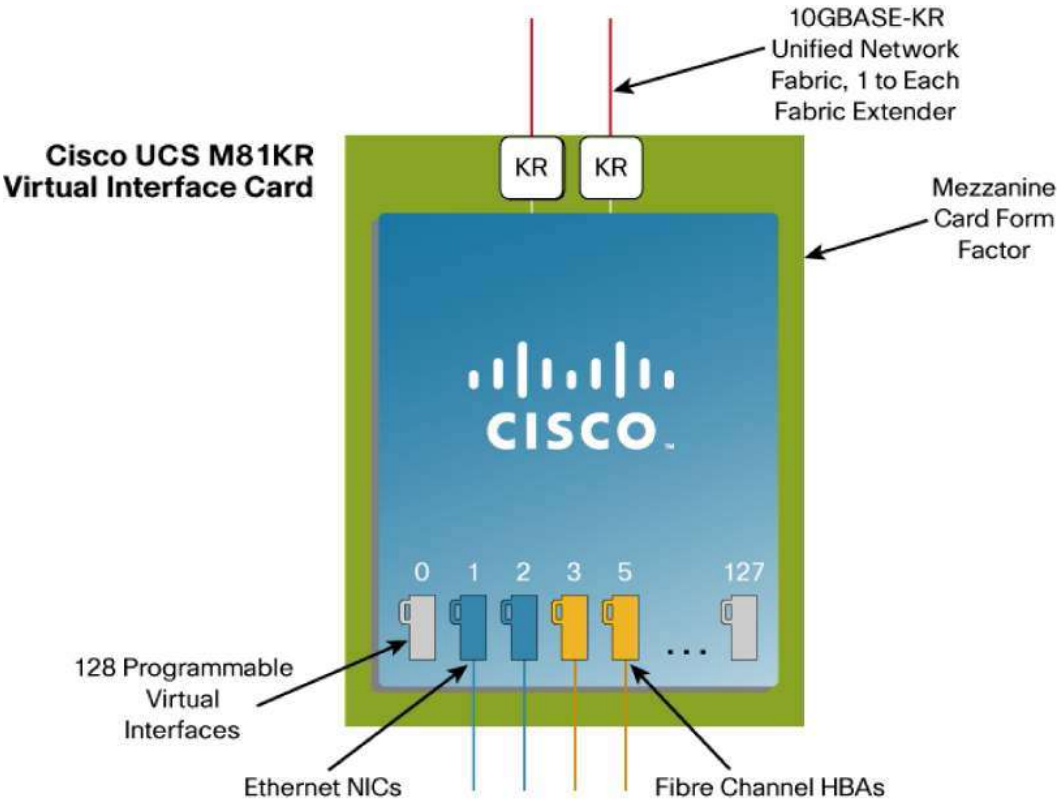
- One network dedicated to management purposes on the host machine
- One network dedicated to the clustered shared volumes (CSVs) and cluster communication network
- One network dedicated to the live-migration network
- One or more networks dedicated to the guest virtual machines (use 10Gbps network adapters for highest consolidation)
- One network dedicated to iSCSI with multipath I/O (MPIO)

For the recommended configuration by quantity and type of NIC, refer to the Hyper-V: Live Migration Network Configuration Guide in the Microsoft TechNet Library.

The Cisco Unified Computing System leverages a 10GbE unified fabric as the underlying I/O transport mechanism. Paired with the unified fabric in this architecture, the Cisco Virtual Interface Card (VIC)

technology has been specified for this solution. The Cisco VIC is a standards-compliant converged network adapter (CNA) that allows both traditional Ethernet as well as FC traffic to share a common physical transport. An added advantage to using this card is the ability to dynamically carve out and assign large numbers of NICs and HBAs from the same physical card. By using two fully redundant 10GbE upstream ports, up to 128 I/O devices can be presented to the PCIe bus and to the operating system running on top of it. These I/O devices can be mixed between NICs and HBAs at the administrator's discretion.

Figure 12) Cisco VIC (graphic provided by Cisco).

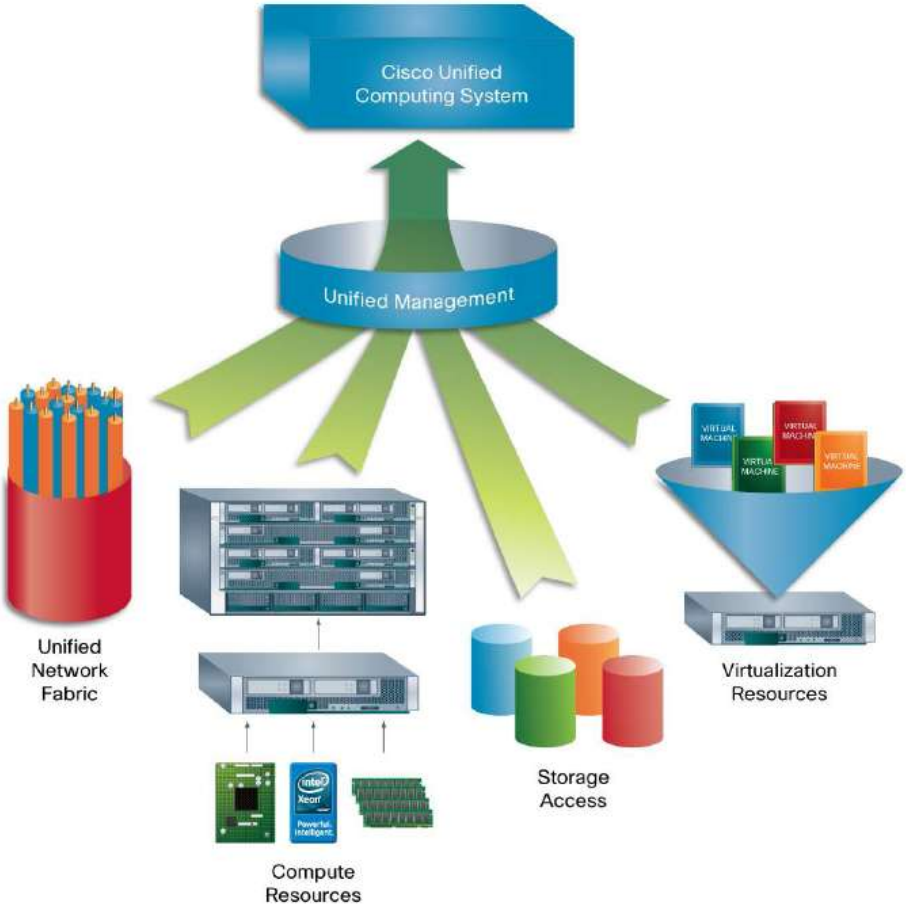


Fabric failover, which allows NIC redundancy to be managed below the operating system, is an additional innovation in the Cisco UCS solution. In the event of an upstream fabric failure, the NICs assigned to the failed fabric automatically fail over to use the remaining fabric. This occurs without the need for special teaming drivers, which frequently introduce instability to a system.

In addition to redundancy, you need a platform that has optimal virtualization support through hardware virtualization capabilities, such as Intel® virtualization (Intel-VT) as the baseline, along with the Second Level Address Translation (SLAT) capabilities of Windows Server 2008 R2 Hyper-V to maximize performance. All of these features are available with the Cisco UCS compute platform.

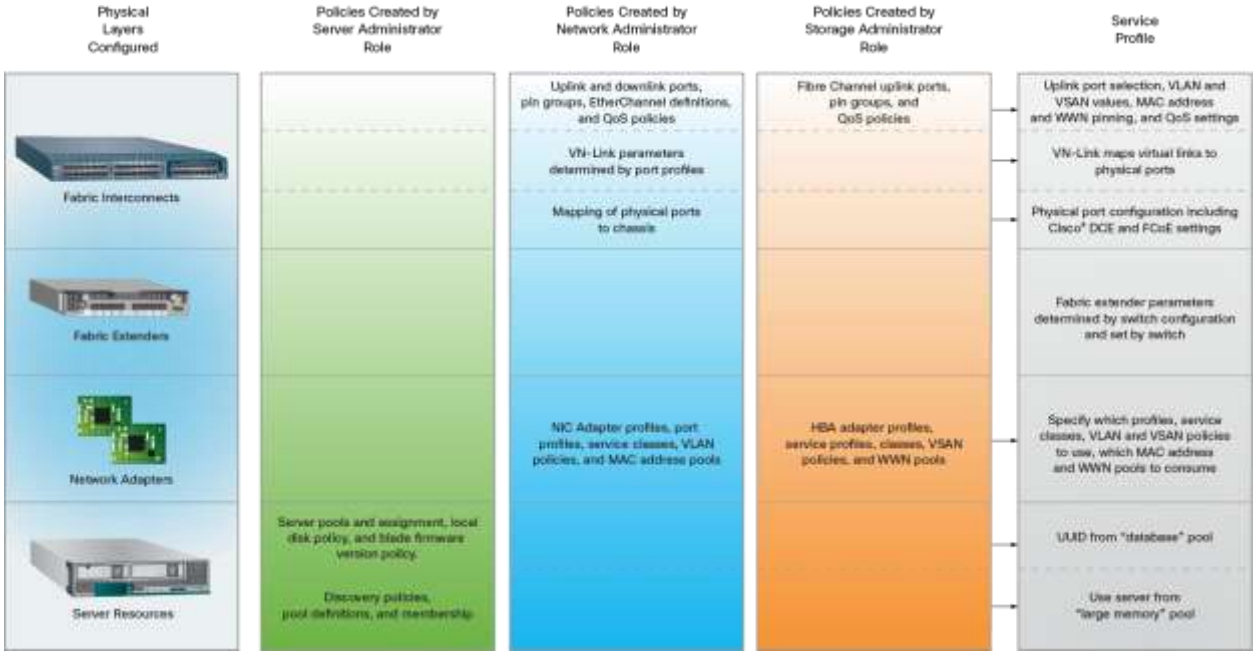
The implementation of a stateless computing model is a unique feature of the Cisco UCS architecture. By building a single point of management for all system components across as many as 320 server blades, Cisco has created a single-pane-of-glass view for the entire compute infrastructure. This unified view of the Cisco UCS is shown in Figure 13.

Figure 13) Cisco Unified Computing System (graphic provided by Cisco).



That capability is then enhanced by abstracting away the configuration of individual services into atomic containers called service profiles. Service profiles contain all the components that identify a physical server to an operating system or hypervisor instance. Figure 14 shows the types of settings contained within a service profile.

Figure 14) Settings contained within a service profile (graphic provided by Cisco).



Because the service profile is an atomic container, the entire personality of a server can be dynamically moved around the compute farm as needed for maximum flexibility.

This logical state abstraction can be managed entirely by using the Cisco UCS Manager API. The API supports every function within the Cisco UCS Manager management schema, and it is fully documented on Cisco.com. Additionally, the API is used by a broad variety of industry partners, which enables rich integration. This architecture features integration through the Microsoft System Center Operations Manager (SCOM) management pack for Cisco UCS, as well as through the Cisco UCS Manager Windows PowerShell toolkit.

Server Storage Connectivity

To achieve maximum flexibility and agility during both host and virtual machine provisioning events, a diskless architecture has been specified. This architecture relies on boot-from-SAN technology for the host operating systems and the use of iSCSI for the guest virtual machine data volumes.

This reference architecture uses a unified fabric architecture that allows both FCoE and 10GbE iSCSI over a common network fabric. To provide the proper quality of service (QoS), advanced QoS features are core to the fabric implementation of this architecture.

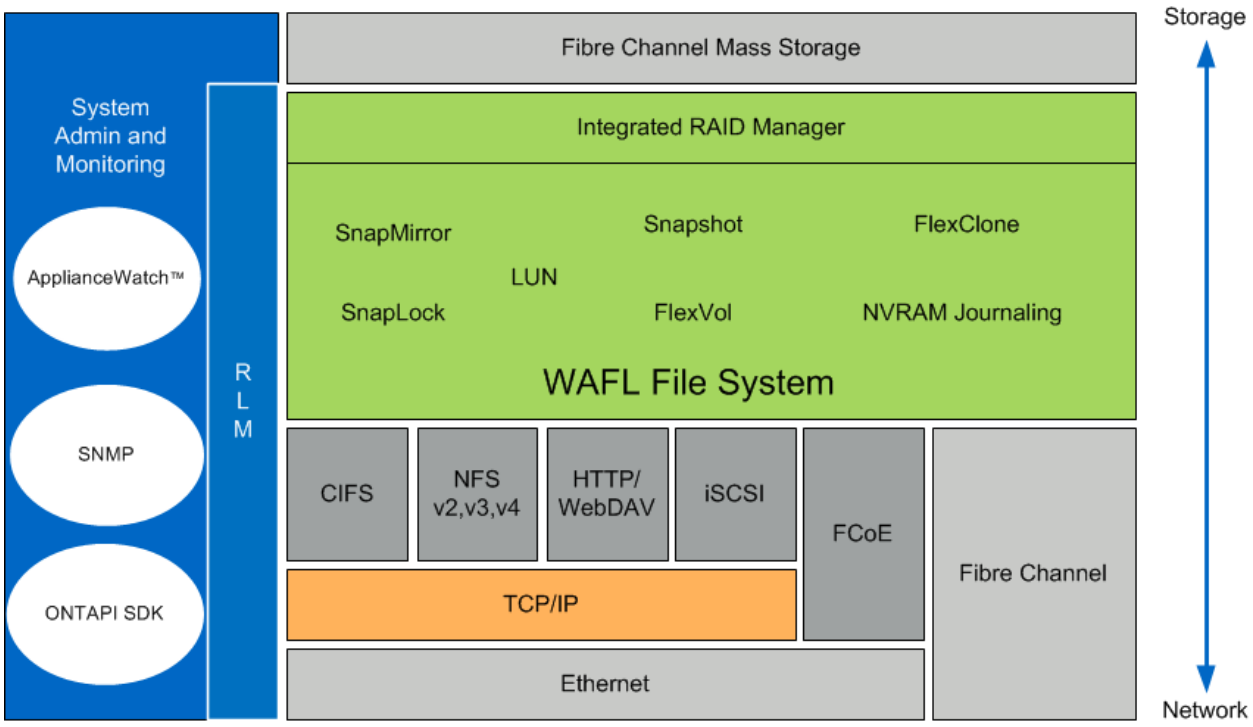
4.4 Storage Architecture

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

The basic architecture of the storage system's software is shown in Figure 15. A collection of tightly coupled processing modules handles CIFS, FCP, FCoE, HTTP, iSCSI, and NFS requests. A request starts in the network driver and moves up through network protocol layers and the file system, eventually generating disk I/O, if necessary. When the file system finishes the request, it sends a reply back to the network. The administrative layer at the top supports a command line interface (CLI) similar to UNIX® that monitors and controls the modules below. In addition to the modules shown, a simple real-time kernel provides basic services such as process creation, memory allocation, message passing, and interrupt handling.

The networking layer is derived from the same Berkeley code used by most UNIX systems, with modifications made to communicate efficiently with the storage appliance's file system. The storage appliance provides transport-independent seamless data access using block- and file-level protocols from the same platform. The storage appliance provides block-level data access over an FC SAN fabric using FCP and over an IP-based Ethernet network using iSCSI. File access protocols such as NFS, CIFS, HTTP, or FTP provide file-level access over an IP-based Ethernet network.

Figure 15) Storage architecture.

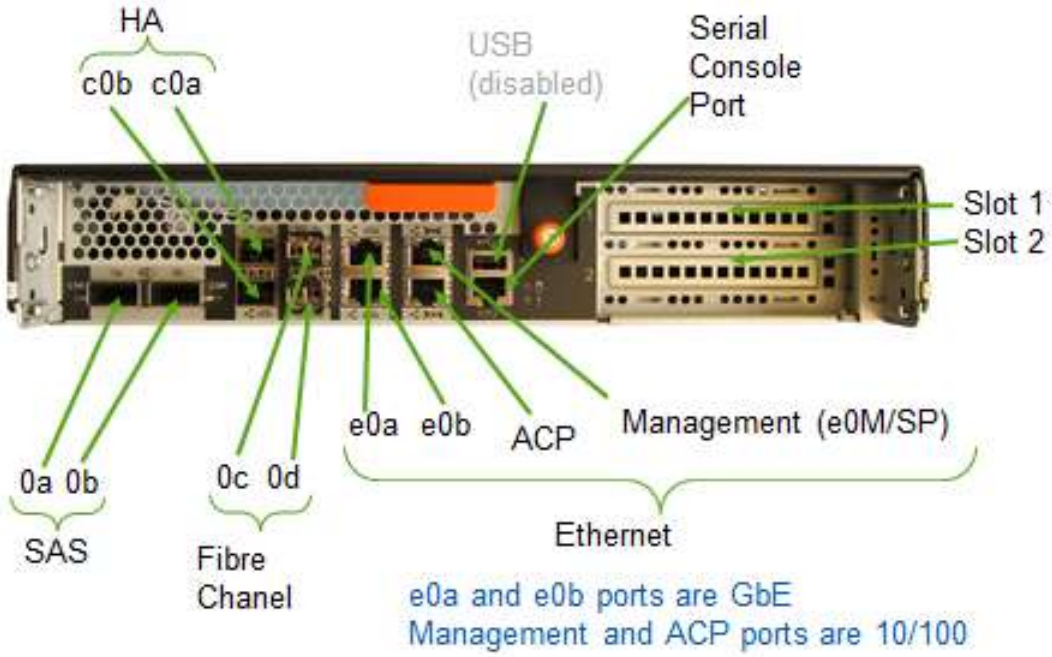


The FlexPod with Microsoft Private Cloud solution uses the FAS3200 series controllers (Figure 16).

Figure 16) Rear view of FAS3240 HA pair.

Each chassis consists of two power supplies and two controllers. Each controller is configured as shown in Figure 17.

Figure 17) FAS3200 controller I/0.



Storage Options

Not all workloads have the same availability requirements, nor do they achieve their requirements in the same way. In the case of data center architecture, we classify workloads as either stateful or stateless. A stateful workload has data that is specific to that virtual machine and if lost would become unavailable. A stateless workload uses data stored elsewhere in the data center and can achieve high availability through resiliency in the application. An example of a stateless workload is a front-end Web server farm.

After the workload type is determined, the performance and availability characteristics of the specific workload should be analyzed to determine the storage characteristics required (performance, redundancy, and so on).

The NetApp FAS series unified architecture has all of these capabilities including replication and clustering support for all FAS controllers from the smallest FAS2000 to the largest FAS6200 series controller. In addition, the FAS3240 controller used in this solution can be licensed for the full range of NetApp Data ONTAP® features.

SAN Storage Protocols

Although many storage options exist, organizations should choose their storage devices based on their specific data management needs. Storage devices are typically modular and flexible midrange and high-end SANs. Modular midrange SANs are procured independently and can be chained together to provide greater capacity. They are efficient, can grow with the environment as needed, and require less up-front investment than high-end SANs. Large enterprises might have larger storage demands and might need to serve a larger set of customers and workloads. In this case, high-end SANs can provide the highest performance and capacity. High-end SANs typically include more advanced features such as continuous data availability through technologies such as replication and clustering.

The NetApp FAS series unified architecture has all of these capabilities, including replication and clustering support for all FAS controllers, from the smallest FAS2000 to the largest FAS6000 series controller. In addition, the FAS3240 controller used in this solution can be licensed for the full range of Data ONTAP features.

Comparing iSCSI, FC, and FCoE

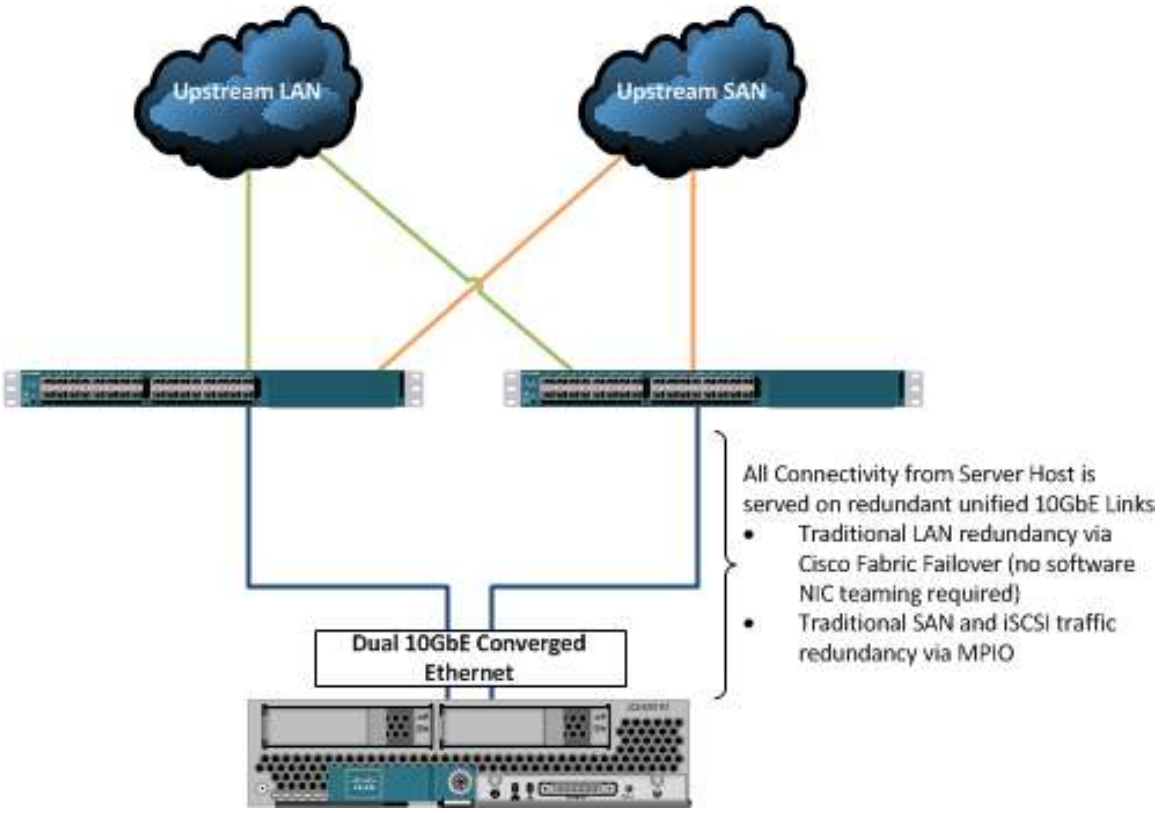
FC has historically been the storage protocol of choice for enterprise data centers for a variety of reasons, including good performance and low latency. Over the past several years, however, the advancing performance of Ethernet from 1Gbps to 10Gbps and beyond has led to great interest in storage protocols that use Ethernet transport, such as iSCSI and, more recently, FCoE.

A key advantage of the protocols that use the Ethernet transport is the ability to use a converged network architecture in which a single Ethernet infrastructure serves as the transport for both LAN and storage traffic. FCoE is an emerging technology that brings the benefits of using an Ethernet transport while retaining the advantages of the FC protocol and the ability to use FC 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 that are capable of supporting enhanced Ethernet.

A common practice in large-scale Hyper-V deployments is to use both FC and iSCSI. FC and iSCSI can provide the host storage connectivity. In contrast, iSCSI is directly used by guests, for example, for the shared disks in a guest cluster. In this case, although Ethernet and some storage I/O share the same pipe, segregation is achieved by VLANs, and QoS can be further applied by the OEM's networking software.

Figure 18) Example of blade server host design (graphic provided by Cisco).



Storage Network

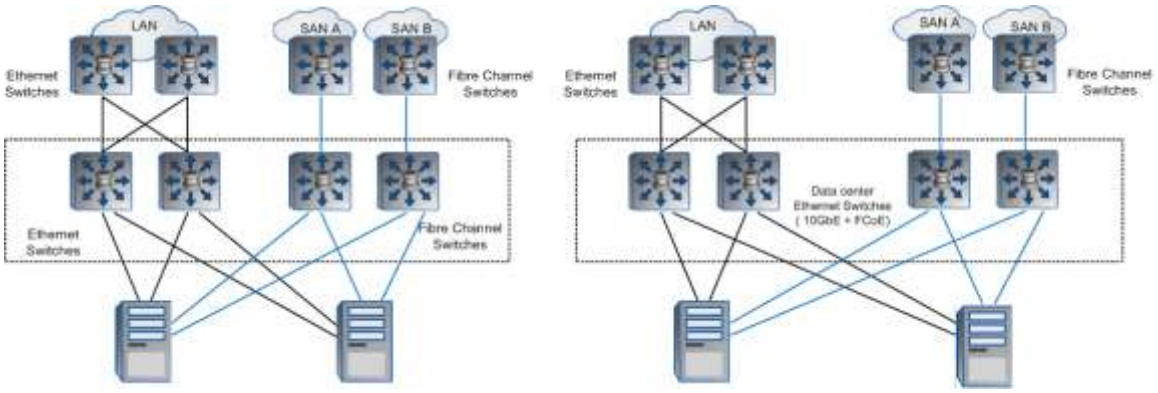
Both iSCSI and FCoE use Ethernet transport for storage networking. This provides another architecture choice in terms of whether to use a dedicated Ethernet network with separate switches, cables, paths, and other infrastructure or, instead, to use a converged network in which multiple traffic types are run over the same cabling and infrastructure.

The storage solution must provide logical or physical isolation between storage and Ethernet I/O. If it's a converged network, QoS must be provided to assure storage performance. The storage solution must provide iSCSI connectivity for guest clustering and fully redundant independent paths for storage I/O.

Standards-based converged network adapters, switches, and FC storage arrays should be used for FCoE. The selected storage arrays should also provide iSCSI connectivity over standard Ethernet so that Hyper-V guest clusters can be used. If iSCSI or FC is used, make sure that there are dedicated network adapters, switches, and paths for the storage traffic.

Figure 19 illustrates the differences between a traditional architecture (shown on the left) with separate Ethernet and FC switches, each with redundant paths, and a converged architecture (shown on the right) in which both Ethernet and FC (through FCoE) use 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 19) Traditional architecture compared to converged architecture.



CSVs

Windows Server 2008 R2 includes the first version of Windows failover clustering to offer a distributed file access solution. 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. CSV uses standard NTFS and has no special hardware requirements beyond supported block-based shared storage.

CSV provides not only shared access to the disk, but also storage path I/O fault tolerance (dynamic I/O redirection). In the event that the storage path on one node becomes unavailable, the I/O for that node is rerouted by Server Message Block (SMB) protocol through another node. There is a performance effect 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.

CSV Limits

Table 1 contains the limitations that are actually imposed by the NTFS file system and which are inherited by CSV.

Table 1) Limitations imposed by NTFS file system.

CSV Parameter	Limitation
Maximum volume size	256TB
Maximum number of partitions	128
Directory structure	Unrestricted
Maximum files per CSV	4+ billion
Maximum virtual machines per CSV	Unlimited

CSV Requirements

The following are the requirements for using CSVs:

- All cluster nodes must use Windows Server 2008 R2 or later.
- All cluster nodes must use the same drive letter for the system disk.
- All cluster nodes must be on the same logical network subnet. VLANs are required for multisite clusters running CSV.
- NT LAN Manager (NTLM) authentication in the local security policy must be enabled on cluster nodes.
- SMB must be enabled for each network on each node that carries 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 can host a virtual machine.

CSV Volume Sizing

Because all cluster nodes can access all CSV volumes simultaneously, we can use standard LUN allocation methodologies that are based on performance and capacity requirements of the workloads running within the virtual machines themselves. Generally speaking, isolating the virtual machine operating system I/O from the application data I/O is a good start, in addition to application-specific I/O considerations, such as segregating databases and transaction logs and creating SAN volumes or storage pools that factor in the I/O profile itself (that is, random read and write operations versus sequential write operations).

CSV's architecture differs from that of 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 virtual machines on a CSV volume other than verifying that the overall I/O requirements of the expected virtual machines running on the CSV are met by the underlying storage system and storage network. Although rare, disks and volumes can enter a state in which a checkdisk is required. With large disks, running checkdisk can take a long time to complete, which can cause downtime on the volume that is roughly proportional to the volume's size.

Each enterprise application you plan to run within a virtual machine can have unique storage recommendations and perhaps even virtualization-specific storage guidance. That guidance applies to use with CSV volumes as well. The important thing to remember is that all virtual machines' virtual disks running on a particular CSV 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 can contain many LUNs. A LUN is simply a logical representation of a disk provisioned from a pool of disks. Therefore, if an enterprise application requires specific storage input/output operations per second (IOPS) or disk response times, you must consider all the LUNs in use

on that storage pool. An application that would require dedicated physical disks, were it not virtualized, might require dedicate storage pools and CSV volumes running within a virtual machine.

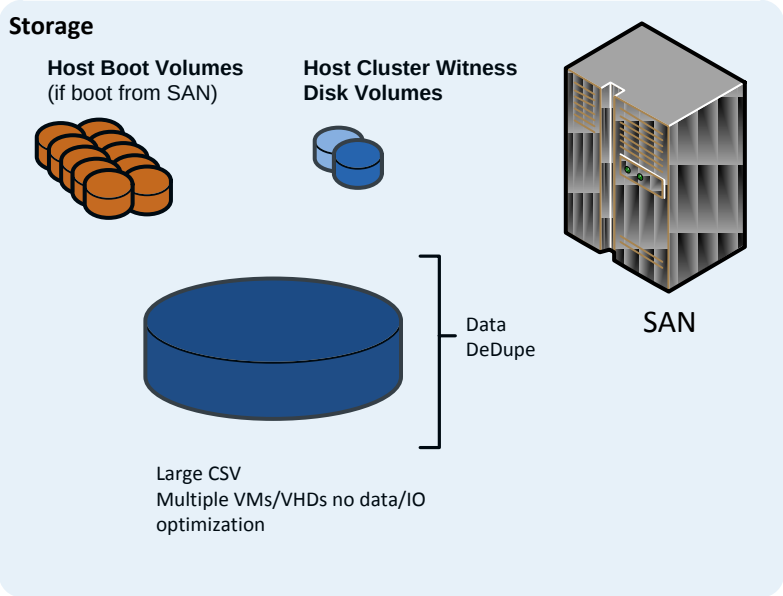
CSV Design Patterns

There are several CSV design patterns that are common in Hyper-V deployments. They are discussed in this section.

Single CSV per Cluster

In the design pattern that includes a single CSV per cluster, the SAN is configured to present a single large LUN to all the nodes in the host cluster. The LUN is configured as a CSV in failover clustering. All virtual machine-related files (virtual hard drives [VHDs], configuration files, and so on) that belong to the virtual machines hosted on the cluster are stored on the CSV. Optionally, data deduplication functionality provided by the SAN can be used (if it is supported by the SAN vendor).

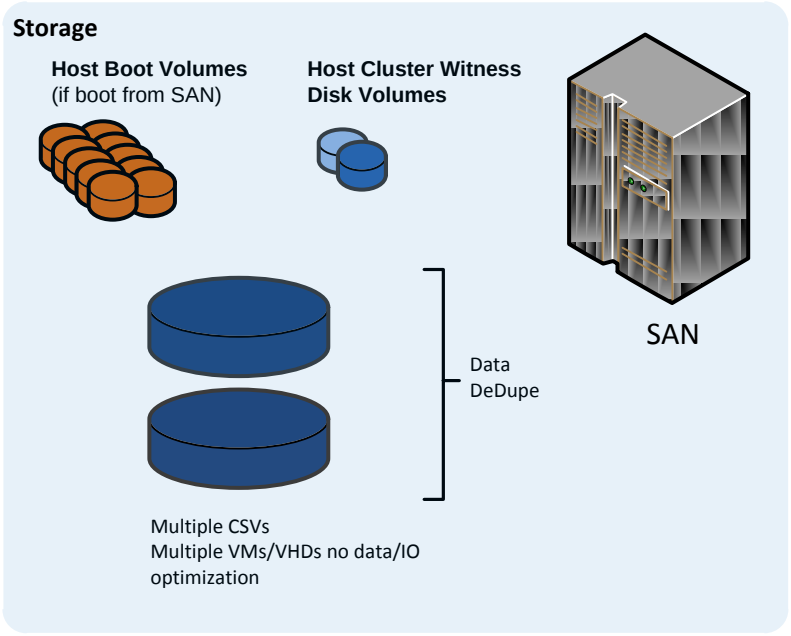
Figure 20) Single CSV per cluster (graphic provided by Microsoft).



Multiple CSVs per Cluster

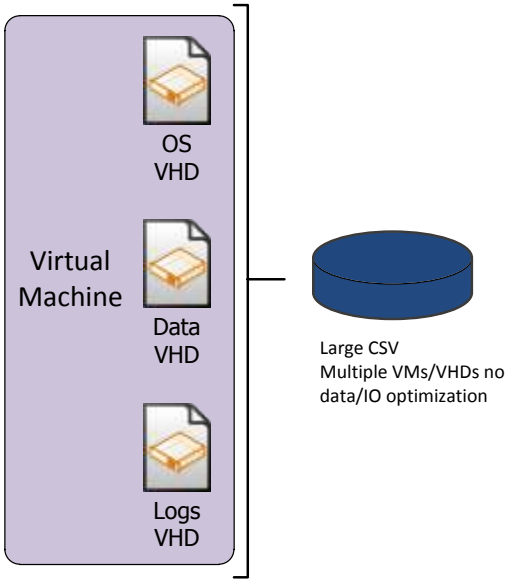
In the design pattern that includes multiple CSVs per cluster, the SAN is configured to present two or more large LUNs to all the nodes in the host cluster. The LUNs are configured as a CSV in failover clustering. All virtual machine-related files (VHDs, configuration files, and so on) that belong to the virtual machines hosted on the cluster are stored on the CSVs. Optionally, data deduplication functionality provided by the SAN can be used (if it is supported by the SAN vendor). NetApp natively supports data deduplication; using deduplication is the joint best practice recommendation for this solution.

Figure 21) Multiple CSVs per cluster (graphic provided by Microsoft).



For both the single and multiple CSV patterns, each CSV has the same I/O characteristics. Therefore, each individual virtual machine has all of its associated VHDs stored on one of the CSVs.

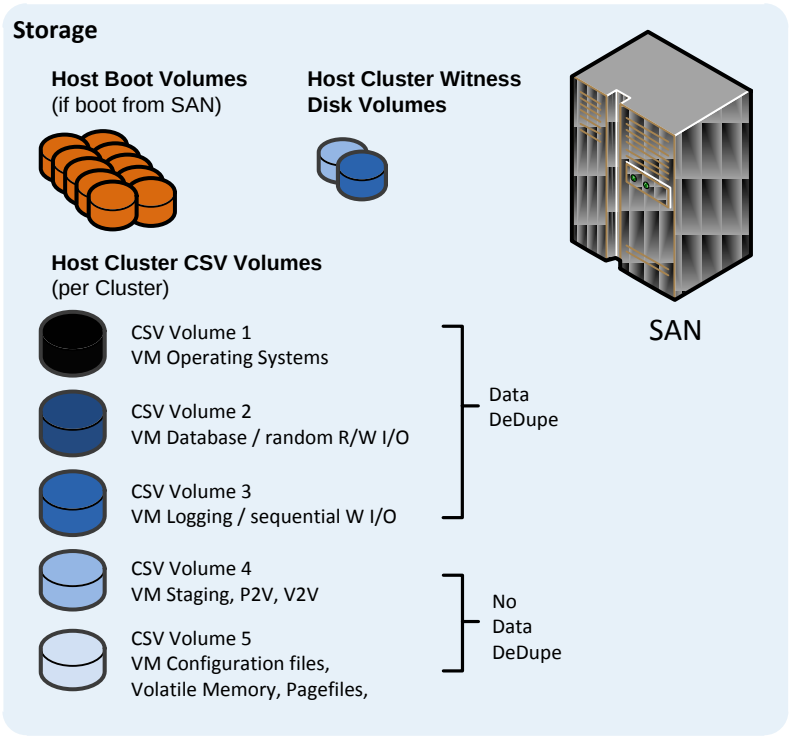
Figure 22) Consolidated VHDs on one CSV (graphic provided by Microsoft).



Multiple I/O-Optimized CSVs per Cluster

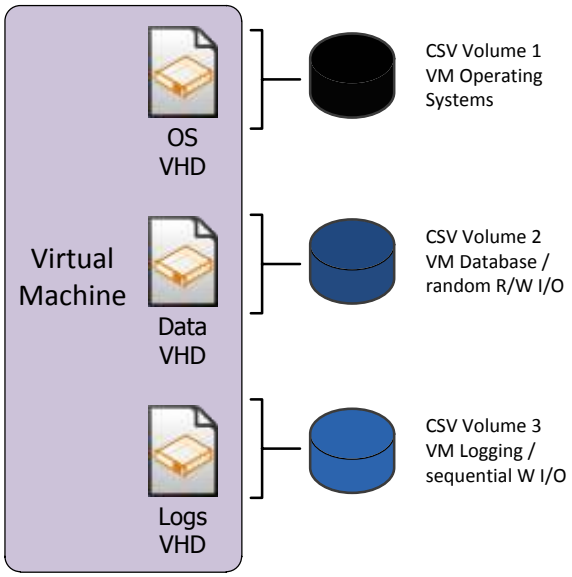
In the design pattern that includes multiple I/O-optimized CSVs per cluster, the SAN is configured to present multiple LUNs to all of the nodes in the host cluster, but the LUNs are optimized for particular I/O patterns, such as fast sequential read performance or fast random write performance. The LUNs are configured as CSVs in failover clustering. All VHDs belonging to the virtual machines hosted on the cluster are stored on the CSVs, but they are targeted to the most appropriate CSV according to the I/O needs.

Figure 23) Multiple I/O-optimized CSVs per cluster (graphic provided by Microsoft).



In this design pattern, each individual virtual machine has all of its associated VHDs stored on the appropriate CSV as per the required I/O requirements.

Figure 24) Distributed VHDs on multiple CSVs (graphic provided by Microsoft).



Note: A single virtual machine can have multiple VHDs, and each VHD can be stored on a different CSV (provided all CSVs are available to the host cluster on which the virtual machine is created).

SAN Design

A highly available SAN design should have no single points of failure, including:

- Redundant power from independent PDUs
- Redundant storage controllers
- Redundant storage paths that are supported, for example, by redundant target ports of NICs per controller, redundant FC or IP network switches, and redundant cabling
- Data storage redundancy similar to what occurs with volume mirroring, or synchronous or asynchronous replication

Address the following elements when designing or modifying your SAN as the basis of your Microsoft Private Cloud storage infrastructure:

- Performance
- Drive types
- Multipathing
- FC SAN
- iSCSI SAN
- Data deduplication
- Thin provisioning
- Volume cloning
- Volume snapshot

Performance

Storage performance is a complex 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 and/or maximum IOPS for a given latency or response time. Although each of these performance measurements is important, IOPS for a given latency are the most relevant to server virtualization.

NetApp virtual storage tiering (VST) leverages NetApp Flash Cache technology. This deduplication-aware technology uses Flash Cache to intelligently store large numbers of recently accessed blocks. The NetApp VST model can significantly increase the performance of an array in servicing the I/O load (or challenge) of a boot storm or a steady-state event.

NetApp FAS controllers use two techniques to optimize both write and read performance. Write performance is optimized by NetApp WAFL® (Write Anywhere File Layout), which delivers writes to the disk that is in the most advantageous rotational position to accept inbound blocks. This provides maximum disk utilization by minimizing write latency. FAS controllers also use Flash Cache to optimize read operations.

Write Anywhere File Layout

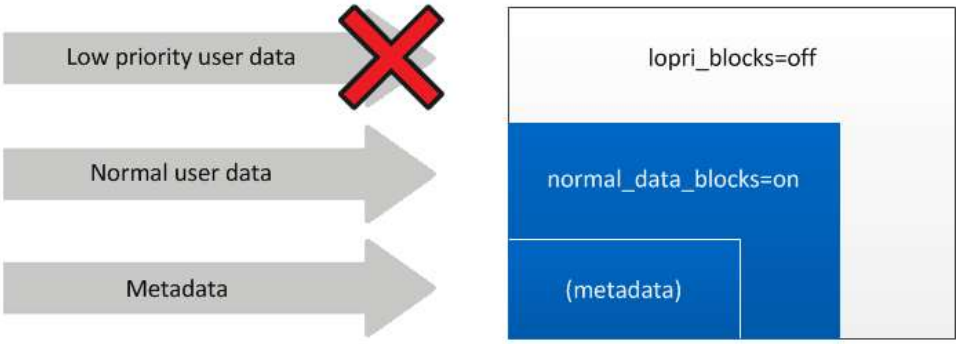
WAFL uses files to store metadata. The three most important WAFL metadata files are the inode file (which contains all inodes), a free block bitmap file, and a free block count file. Keeping metadata in files allows metadata blocks to be written anywhere on the disk. WAFL has complete flexibility in its write allocation policies, because no blocks are permanently assigned to fixed disk locations, because they are in the Berkeley Fast File System (FFS). WAFL uses this flexibility to optimize write performance for the storage system's RAID-DP® features.

Flash Cache

Flash Cache allows Data ONTAP to automatically and dynamically move data from traditional hard drives to flash storage based on actual usage patterns. This allows NetApp controllers to support much higher IOPS levels than traditional storage systems.

- **Modes of operation.** Flash Cache and PAM have three modes of operation. These modes provide the ability to tune the caching behavior of the module to match the storage system's workload. As we move through each of the modes, Data ONTAP allows a broader amount of data to be stored in the module.
- **Normal user data caching.** The simplest performance improvement is made by caching data as it is accessed from disk. On successive accesses, instead of going to disk with higher latency, the data is served from the memory onboard the Flash Cache. This is the default mode.

Figure 25) Normal user data caching.



The next two modes are not the default operation modes. If there is a specific requirement to configure Flash Cache in these modes, refer to TR-3832: Flash Cache and PAM Best Practices Guide.

Metadata Caching

In this mode, the metadata is placed into a read cache, which allows low-latency access to the metadata. It also provides higher speed access to the application data. Because of the much larger size of Flash Cache, this mode is more applicable to the original PAM card.

Low-Priority Data Caching

In this mode, the performance is improved by capturing application data that normally would have been forced to disk or not cached at all. These are called low-priority buffers and exist in a couple of forms.

The first form is write data. Normally, writes are buffered in RAM and logged to NVRAM. After they are committed to disk, they are flushed from NVRAM and retain a lower priority in RAM to avoid overrunning the system memory. In other words, recently written data is the first to be ejected. In some workloads, recently written data might be immediately accessed after being written. For these workloads, Flash Cache improves performance by caching recently written blocks in memory rather than flushing them to disk and forcing a disk access for the next read.

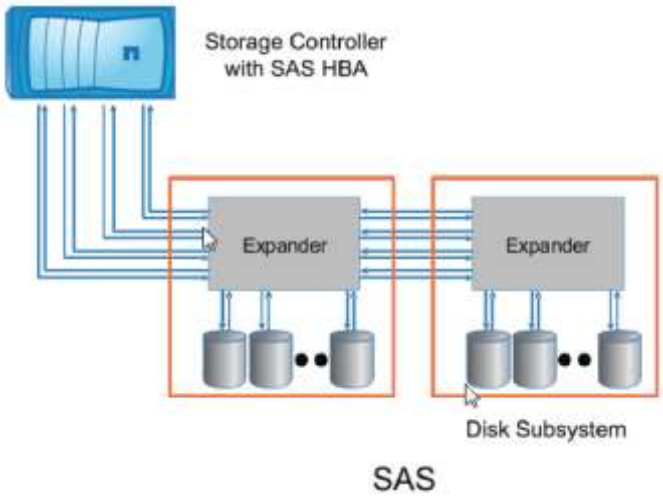
The second form is long sequential read blocks. In this situation, reads can overrun the system memory by overflowing it with a large amount of data that is only accessed once. By default, Data ONTAP does not keep this data, but holds data that is more likely to be reused. The large amount of memory space provided by Flash Cache allows sequential reads to potentially be stored without negatively affecting other cached data. If these blocks are reaccessed, end users see the performance benefit of Flash Cache in contrast to going to disk.

Drive Types

The type of hard drive used in the host server or the storage array has the most significant effect on the overall storage architecture performance. As with 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. Using 15K RPM drives over 10K RPM drives can result in up to 35% more IOPS per drive.

NetApp FAS controllers support FC, SAS, and SATA disks. For this solution, NetApp used four DS2246 shelves (24x 10K RPM, 450GB, 6GB SAS). They were configured using direct SAS drops to both controllers, similar to the configuration shown in Figure 26.

Figure 26) SAS controller with SAS HBA.



Multipathing

In all cases, multipathing should be used to make sure host connectivity is highly available. NetApp provides a device-specific module (DSM) on top of Windows Server 2008 R2 MPIO software that supports the NetApp storage platform. The Data ONTAP DSM provides advanced active-active policies while providing granular failover and path recovery for NetApp LUNs. The Microsoft native DSM is also supported.

FC SAN

FC is an option, because it is a supported storage connection protocol. FC is a robust and well-developed storage protocol that supports multipathing through Microsoft MPIO and NetApp DSM.

iSCSI SAN

As with FC-connected SAN, which is naturally on its own isolated network, the iSCSI SAN must be on an isolated network, for both security reasons and performance considerations. Any networking standard practice method for achieving this goal is acceptable, including a physically separate, dedicated storage network and a physically shared network with the iSCSI SAN running on a private VLAN. The switch hardware must provide class-of-service (CoS) or QoS guarantees for the private VLAN.

- **Encryption and authentication.** If multiple clusters 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 may it be visible to a node from a different cluster. We recommend the use of a session authentication protocol such as Challenge Handshake

Authentication Protocol (CHAP). This provides a degree of security as well as segregation. Mutual CHAP or IPsec can also be used.

- **Jumbo frames.** If supported at all points in the entire path of 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.
- **NetApp MultiStore.** NetApp MultiStore® implements specific, isolated NetApp vFiler® units, each with its own management content and credentials. vFiler units support VLAN tagging, and, when used with iSCSI, they support end-to-end isolation of I/O over a shared Ethernet infrastructure.

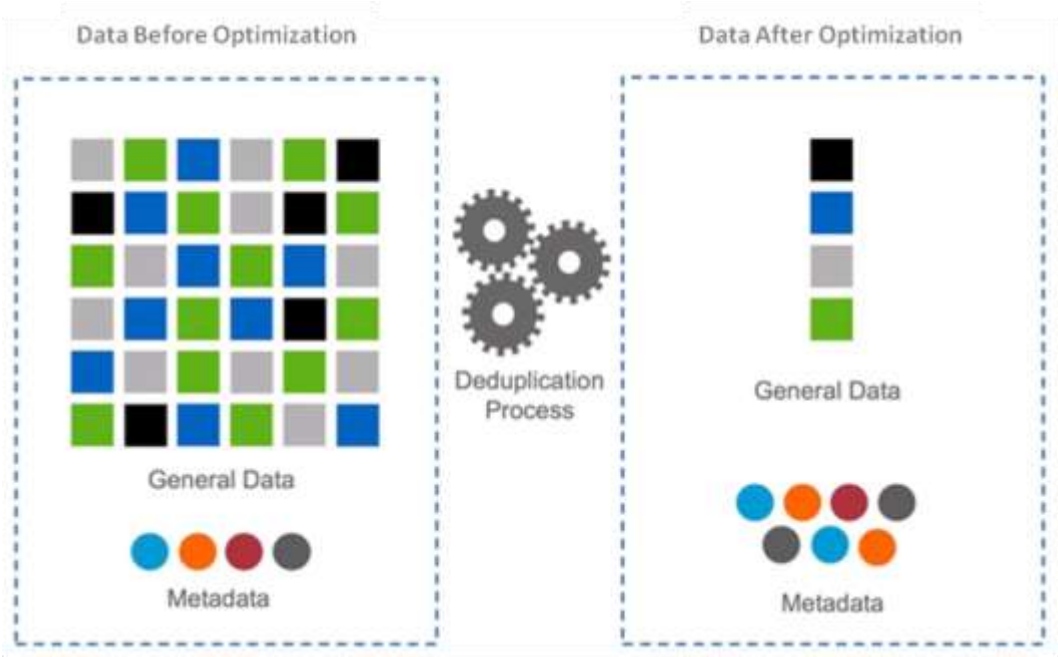
Data Deduplication

Data deduplication can yield significant storage cost savings in virtualization environments. Some common considerations are, for example, performance implications during the deduplication cycle and achieving maximum efficiency by locating similar data types on the same volume or LUN.

How Deduplication for FAS Works

As part of the NetApp storage efficiency offerings, NetApp deduplication for FAS provides block-level deduplication within the entire flexible volume on NetApp storage systems. Beginning with Data ONTAP 7.3, V-Series also supports deduplication. NetApp V-Series is designed to be used as a gateway system that sits in front of third-party storage, and it allows NetApp storage efficiency and other features to be used on third-party storage. Figure 27 shows how NetApp deduplication for FAS works at the highest level.

Figure 27) How NetApp deduplication for FAS systems works.



Essentially, deduplication stores only unique blocks in the flexible volume and creates a small amount of additional metadata in the process. Deduplication works with a high degree of granularity, that is, at the 4KB block level. It operates on the active file system of the flexible volume. Any block referenced by a NetApp Snapshot® copy is not made available until the Snapshot copy is deleted. Deduplication is a background process that can be configured to run automatically on a schedule or manually through the CLI. It is application transparent; therefore, it can be used for the deduplication of data originating from any application that uses the NetApp system. The feature is enabled and managed by using a simple CLI. It can be enabled and can deduplicate blocks on flexible volumes with new and existing data.

In summary, deduplication works like this: Newly saved data on the FAS system is stored in 4KB blocks as usual by Data ONTAP. Each block of data has a digital fingerprint that is compared to all other fingerprints in the flexible volume. If two fingerprints are found to be the same, a byte-for-byte comparison is done of all bytes in the block and, if there is an exact match between the new block and the existing block on the flexible volume, the duplicate block is discarded and the disk space is reclaimed.

Thin Provisioning

Thin provisioning is a common practice, particularly in virtualization environments. This allows available storage capacity to be used efficiently. The LUN and corresponding CSV can grow as needed, typically in an automated fashion, to provide availability of the LUN (autogrow). However, storage can become overprovisioned in this scenario. Therefore, careful management and capacity planning are critical.

How to Set Up SAN Zero-Fat Provisioning

In zero-fat provisioning, primary data and its Snapshot copy space are allocated on demand. This variant achieves the optimal ratio of storage efficiency when provisioning applications from scratch. The joint best practice is for customers to choose the zero-fat provisioning method to increase storage efficiency. Zero fat follows a 100% allocate-on-demand concept.

The zero-fat method has the following characteristics:

- Volumes are created without space guarantee.
- LUNs are created without space guarantee.
- The size of the volume follows the formula X – N + Δ.

  Where X is the size of the primary data equal to the sum of all LUN capacities within the volume, Δ is the amount of space needed to hold Snapshot copy data, and N is the amount of unused blocks within a given LUN.

Table 2) Zero-fat provisioning volume options.

Volume Snapshot Option	Recommended Value	Notes
Reserve	0	
Schedule	Switched off	For network-attached storage (NAS) volumes, a Snapshot copy reserve area and configuration Snapshot copy schedules are a common setup. For SAN volumes, this needs to be switched off according to NetApp best practices (for more information, refer to the Fibre Channel and iSCSI Configuration Guide).
Autodelete	On	To allow use with Provisioning Manager, snap autodelete is turned on. Deleting Snapshot copies might be an option when the volume can no longer be resized because the maximum configured size has been reached or when the aggregate's free space becomes low.

Table 3) Zero-fat provisioning volume Snapshot options.

Volume Snapshot Option	Recommended Value	Notes
Reserve	0	
Schedule	Switched off	For NAS volumes, a Snapshot copy reserve area and configuration Snapshot copy schedules are a common setup. For SAN volumes, this needs to be switched off according to NetApp best practices (for more information, refer to the Fibre Channel

Volume Snapshot Option	Recommended Value	Notes
		and iSCSI Configuration Guide).
Autodelete	On	To allow use with Provisioning Manager, snap autodelete is turned on. Deleting Snapshot copies might be an option when the volume can no longer be resized because the maximum configured size has been reached or when the aggregate's free space becomes low.

Table 4) Zero-fat provisioning LUN options.

LUN Option	Recommended Value	Notes
Reservation	Disable	No pre-allocation of blocks for LUN

Even with a 100% block-usage ratio on primary data, zero-fat provisioning is the preferred method and has many advantages, including:

- The aggregate's free space is a global pool that can serve space for volumes. This provides more flexibility than volumes with their own dedicated free space.
- For SAN volumes, the block consumption can be easily monitored.
- Deduplication savings go directly into the global pool of free space, which is the aggregate or the resource pool in which it belongs.
- Monitoring is needed only on the aggregate level. Volumes grow on demand.

Table 5) Comparison of provisioning methods.

Characteristics	Full Fat	Low Fat	Zero Fat
Space consumption	2X + Δ	X + Δ	X – N + Δ^2
Space efficient	No	Partially, for Snapshot copies	Yes
Monitoring	Optional	Required on volume and aggregate level	Required on aggregate level
Notification and mitigation process required	No	Optional in most cases	Yes
Pool benefitting of dedupe savings	Volume fractional reserve area	Volume free space area	Aggregate free space area
Risk of an out-of-space condition on primary data	No	No, as long as autodelete is able to delete any Snapshot copies	Yes, when monitoring and notification processes are missing
Typical use cases	• Small installations • None or few storage management skills (no monitoring infrastructure)	• Large database environments	• Shared storage infrastructure • Test and development environments • Storage pools for virtualized servers

Volume Size Considerations

Because physical allocation of data within a zero-fat-provisioned volume is done on demand, theoretically the volume size can be set to a very high value that can easily keep all application data and Snapshot copies. The unallocated space in the volume is not exclusively reserved for the volume itself; therefore, all other applications can benefit from the shared pool of unallocated storage. However, the joint best practice recommendation is to size the volume to the anticipated size of its containing objects and to use the autogrow option to allow it to grow on demand. The advantage is that the commitment rate acts as a metric for data consolidation. The commitment rate reflects the amount of logical data consolidation. This metric is suitable for deciding when data should be left for organic growth.

Additionally, the volume size limits when using deduplication should be taken into account, because the maximum sizes depend on the storage controllers.

Consequences for Monitoring

When using zero-fat provisioning, a very high data consolidation can be achieved. Because this effect depends on the usage characteristics of the corresponding applications, monitoring the aggregate is critical.

Volume Cloning

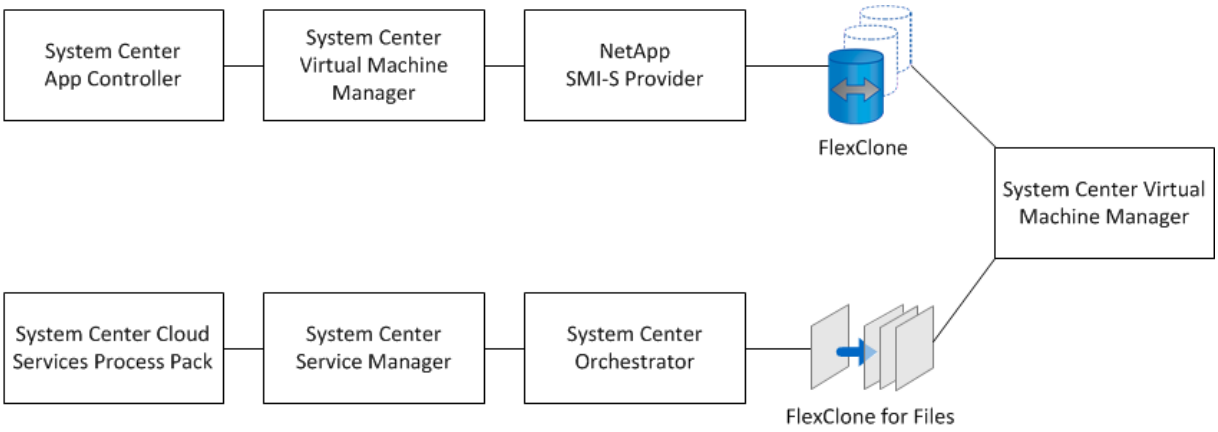
Volume cloning is another common practice in virtualization environments. This can be used for both host and virtual machine volumes to dramatically decrease host installation and virtual machine provisioning times.

Rapid provisioning is a common feature for private cloud implementations. In these environments, the expectation is that end users or departmental administrators will deploy virtual machines. Because of this, the system needs to respond rapidly to provisioning requests and must scale those requests to accept large numbers of simultaneous requests. The joint best practice recommendation is to use clone-based provisioning for this use case. Clone-based provisioning has several key advantages over traditional copy-based provisioning. Specifically, clone-based provisioning provides the following:

- **Very rapid provisioning.** Creating a new clone within Data ONTAP normally takes only a few seconds. This operation can be performed hundreds of times and does not degrade over time.
- **Inherently thin provisioning.** Clone-based provisioning makes sure that only new blocks are stored on the storage controller. This means that virtual machines that have very little unique data do not take up space on the system. This makes sure that only unique data is stored for each virtual machine. In addition, volumes that are deduplicated prior to deployment are retained in their deduplicated state. This prevents unnecessary duplication of data during the provisioning process.
- **Low network impact.** Copy operations can be network intensive. Cloning prevents copy traffic from hitting the back-end network by keeping the operation within the storage controller.
- **Low server impact.** Cloning allows your physical Hyper-V hosts and your management nodes to scale farther by removing I/O-intensive copy operations from the hosts and moving the operation to the storage controller. This allows higher utilization of your existing server infrastructure.
- **Support for fixed-size VHDs.** Although Microsoft recommends fixed VHDs for performance reasons, many sites use dynamic VHDs for performance and efficiency reasons. Cloning enables highly efficient provisioning while still using fixed VHDs. Cloning performance is not dependent on VHD size, so very large VHDs can be deployed efficiently.

Figure 28 shows a high-level process view of the rapid provisioning feature.

Figure 28) Provisioning.



As shown in Figure 28, there are two primary paths to request virtual machine provisioning. One is through System Center App Controller, which uses System Center Virtual Machine Manager's SMI-S-based storage connections to perform LUN clones or traditional file copies. This approach is simple to configure, but does not support advanced scenarios such as sub-LUN cloning through NetApp FlexClone® for files, and it is not recommended for this solution. However, SMI-S is fully supported. For more information on provisioning with SMI-S, refer to the NetApp technical report "System Center Virtual Machine Manager 2012 and NetApp Data ONTAP SMI-S Agent."

The other path is to accept inbound change requests from Service Manager. Organizations that have a robust IT ticketing system use this approach or similar approaches. Because this method is based on tight Orchestrator integration, the Service Manager component can be removed and replaced with most ticketing systems.

In either case, a clone of the source virtual machine is created by Data ONTAP, and it is provisioned to the target physical host. In the case of Service Manager and Orchestrator, this is a file granular clone. That is to say, a single VHD file is cloned within a running virtual machine. This is sometimes referred to as a FlexClone file and was first introduced in Data ONTAP 7.3.1. This cloning technology uses the same block-level file reference mechanism that is used in the Data ONTAP deduplication feature. This allows us to perform cloning very quickly at the sub-LUN level, which reduces deployment time from hours to minutes. For more information about FlexClone files, refer to TR-3742: Using FlexClone to Clone Files and LUNs.

For SMI-S only, LUN granular cloning is supported. This means that each new VHD clone request creates a new LUN that is provisioned to the Hyper-V host. This more traditional method uses FlexClone copies based on Snapshot. This operation can be performed at the LUN or volume level, and it provides an alternative method for rapid cloning.

Volume Snapshot Copies

SAN volume Snapshot copies are a common method of providing a point-in-time instant backup of a SAN volume or LUN. These Snapshot copies are typically block level and only use storage capacity as blocks change on the originating volume. Some SANs provide tight integration with Hyper-V, integrating both the Hyper-V Volume Shadow Copy Service (VSS) writer on hosts and volume Snapshot copies on the SAN. This integration provides a comprehensive and high-performing backup and recovery solution.

Snapshot integration with Hyper-V is achieved through the use of NetApp SnapManager® for Hyper-V (SMHV). SMHV has the following features:

- Allows system administrators to create hardware-assisted backups and restores of Hyper-V VMs running on NetApp storage.

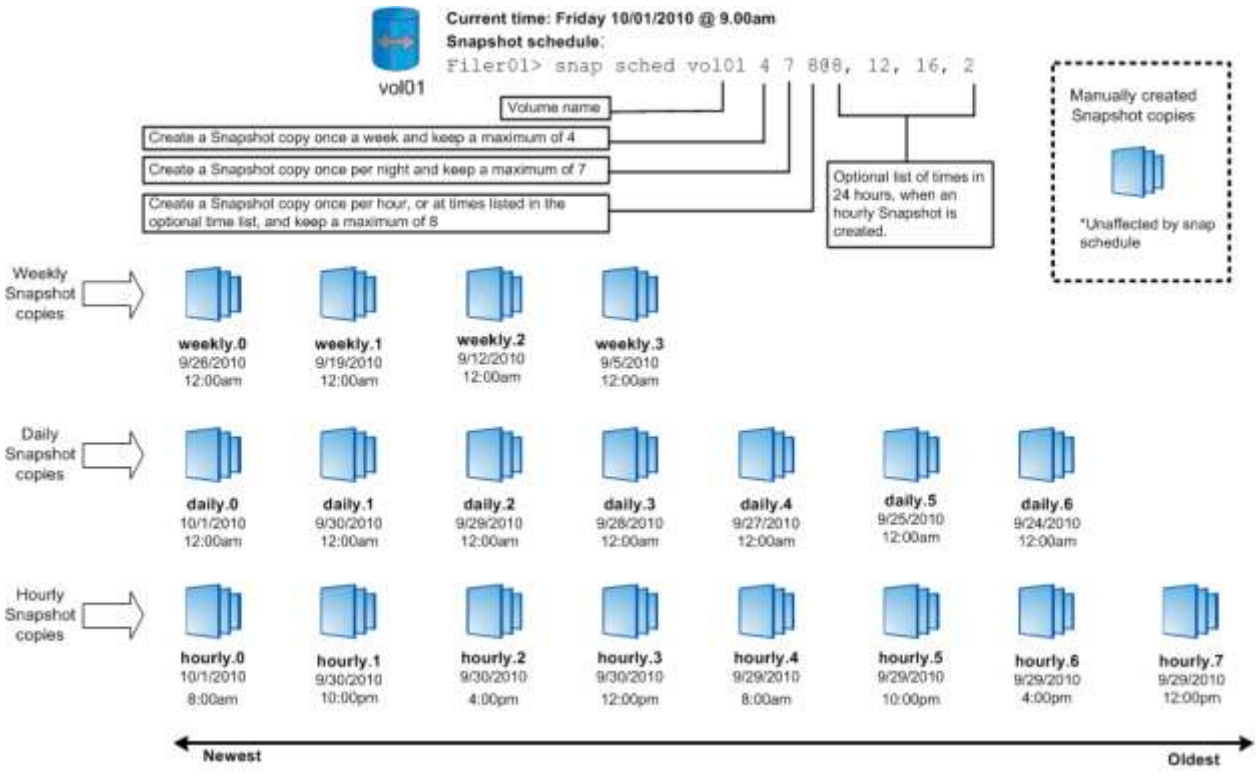
- Provides integration with Microsoft Hyper-V VSS writer to quiesce the Hyper-V VMs before creating an application-consistent Snapshot copy of the virtual machine.
- Allows an administrator to create application-consistent backups of Hyper-V VMs, if the customer has Microsoft Exchange, SQL Server, or any other VSS-aware application running on virtual VHDs in the virtual machine.
- Mirrors backup sets to secondary locations for disaster recovery (DR) planning.
- Supports the backup and restore of shared virtual machines configured for high availability using Microsoft Windows failover clustering (WFC) and also on Microsoft CSVs; SMHV makes sure that the scheduled virtual machine backups happen seamlessly regardless of any virtual machine failovers.
- Supports management of multiple remote Hyper-V parent systems from one console.

A Snapshot copy is a read-only image of a traditional volume, a NetApp FlexVol® volume, or an aggregate that captures the state of the file system at a point in time.

For information about traditional volumes, FlexVol volumes, or aggregates, refer to the Data ONTAP Storage Management Guide. Data ONTAP maintains a configurable Snapshot copy schedule that creates and deletes Snapshot copies automatically for each volume. Snapshot copies can also be created and deleted manually using the `snap create` and `snap delete` commands.

Figure 29 shows an example volume (`vol01`) and the corresponding Snapshot copies that are created by the schedule shown in the diagram.

Figure 29) Snapshot copy example using the snap sched command.



Guidelines and Restrictions

Avoid scheduling creation of Snapshot copies to occur at the same time as NetApp SnapMirror® updates or NetApp SnapVault® activities. If these schedules conflict, creation of Snapshot copies might not occur.

Stagger the Snapshot copy update schedules so that SnapMirror activity does not begin or end at the exact minute a scheduled Snapshot operation attempts to create a Snapshot copy. Additionally, if

scheduled Snapshot copies conflict with SnapVault activity, customer administrators should use the `snapvault snap sched` command to configure equivalent schedules.

Snapshot Copies in a SAN Environment

Administrators can use Snapshot technology to make copies in the SAN environment when the data within a Data ONTAP LUN is in a consistent state. However, Data ONTAP cannot determine if the data within a LUN is in a consistent state; that is, Data ONTAP does not know whether an application is accessing the data inside the LUN. Therefore, before creating a Snapshot copy, administrators need to quiesce the application or file system using the LUN. This action flushes the host file system buffers to disk. Quiescing makes sure that Snapshot copies are consistent.

One way to maintain consistency is to use batch files and scripts on a host that has administrative access to the system. The NetApp SnapDrive® and SnapManager products also quiesce LUNs before creating Snapshot copies and should be used whenever possible.

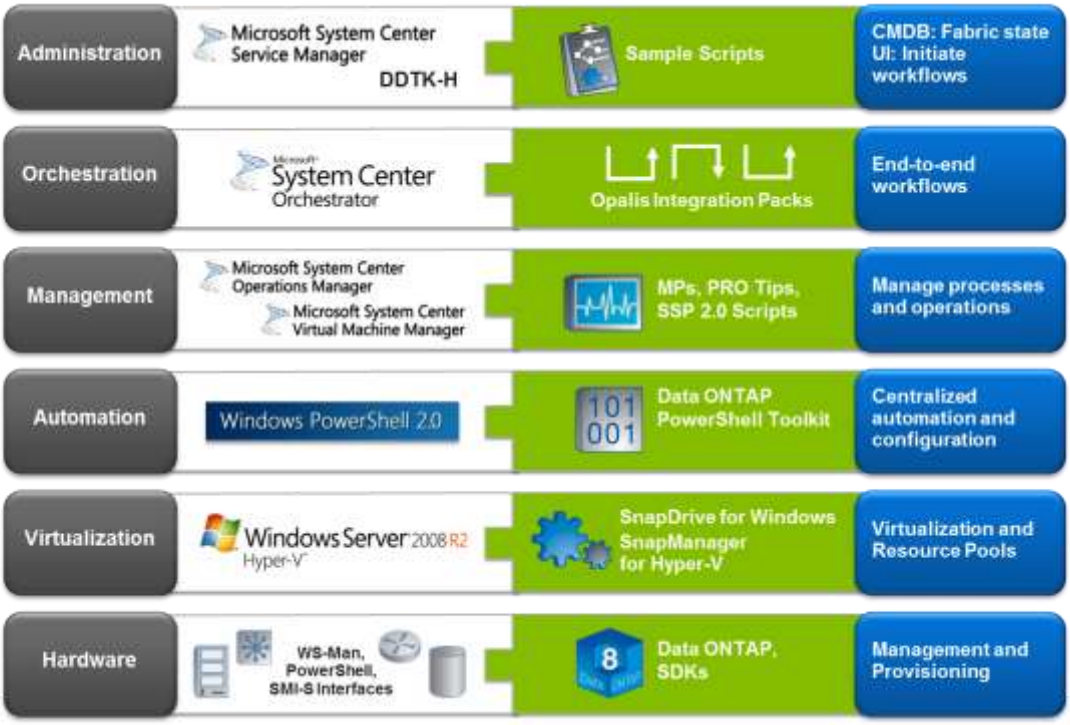
Storage Automation

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

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

To resolve this issue, NetApp supports the range of Microsoft management tools and APIs shown in Figure 30. Specifically, NetApp has shipped the Data ONTAP Windows PowerShell Toolkit, which allows the management of NetApp controllers from Windows PowerShell. In addition, as part of the OnCommand™ plug-in for Microsoft Environments 3.0, NetApp provides a native Opalis integration pack (OIP) that allows common operations such as provisioning and DR.

Figure 30) NetApp solutions for Microsoft Private Cloud.



5 Network Architecture

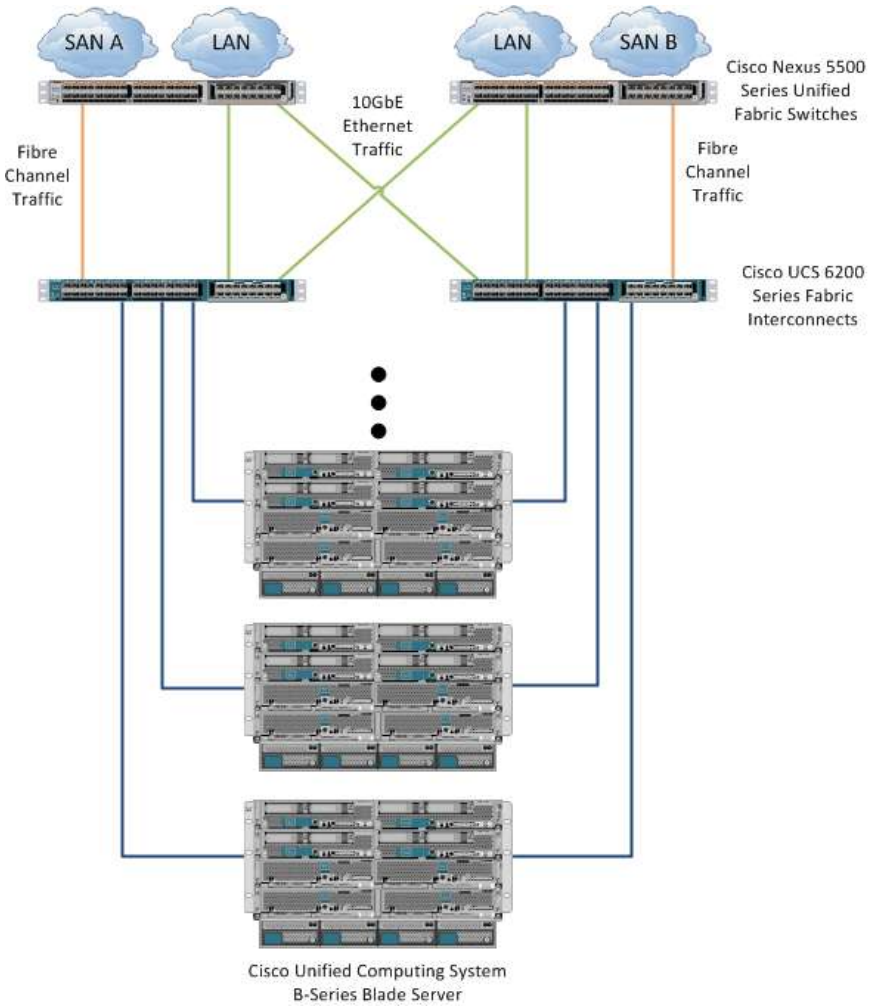
Many network architectures include a tiered design with three or more tiers such as core, distribution, and access. Designs are driven by the port bandwidth and quantity required at the edge, as well as the ability of the distribution and core tiers to provide higher speed uplinks to aggregate traffic.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

Cisco UCS offers a unique perspective on server-focused network architecture. An integrated network strategy at the core of Cisco UCS provides 10GbE connectivity to all components. Coupling this fabric with the stateless, policy-driven server architecture described previously in this document allows vast simplification of the physical infrastructure typically deployed in a new server buildout.

Rather than including localized Ethernet and FC switching in each chassis, all fabric aggregation is performed at a top-of-rack (ToR) type of device called the fabric interconnect. Cisco UCS 6200 Series fabric interconnects are a family of line-rate, low-latency, lossless 10GbE, DCB, and FCoE interconnect switches that consolidate I/O at the system level. Based on the same switching technology as the Cisco Nexus 5500 Series switches, the Cisco UCS 6200 Series fabric interconnects provide the additional features and management capabilities that make up the core of the Cisco UCS.

The fabric interconnects supply a unified fabric that connects every server in the system through wire-once 10GbE and FCoE downlinks and flexible 10GbE and 1/2/4/8Gbps FC uplinks (as shown in Figure 31). Out-of-band management, including switch redundancy, is supported through dedicated management and clustering ports. The interconnects feature front-to-back cooling, redundant front-plug fans and power supplies, and rear cabling that facilitates efficient cooling and serviceability. Typically deployed in active-active redundant pairs, the fabric interconnects provide uniform access to both networks and storage, which eliminates the barriers to deploying a fully virtualized environment based on a flexible, programmable pool of resources.

Figure 31) Fabric interconnects (graphic provided by Cisco).



As shown in Figure 31, the Cisco UCS can join any standards-compliant existing SAN and LAN infrastructure as it leaves the fabric interconnects.

5.1 High Availability and Resiliency

Providing redundant paths from the server through all the network tiers to the core is this solution's joint best practice recommendation for high availability and resiliency. A variety of technologies (NIC teaming, Spanning Tree Protocol, and so on) can be used to create redundant path availability without looping.

Each network tier should include redundant switches. With redundant pairs of access-tier switches, individual switch resiliency is slightly less important, and therefore, the expense of redundant power supplies and other component redundancy might not be required. At the aggregation and core tiers, full hardware redundancy, in addition to device redundancy, is recommended due to the critical nature of those tiers.

However, sometimes devices fail, become damaged, or get misconfigured. In these situations, remote management and the ability to remotely power cycle all devices become important factors in restoring service rapidly.

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

Cisco Unified Computing System

The Cisco UCS platform provides the following features that support high availability and resiliency:

- Redundant LAN and SAN fabrics
- Fabric failover hardware-based NIC teaming
- Port channel link aggregation, load balancing, and link fault tolerance
- Hot-swappable field-replaceable power supplies, fan modules, and expansion modules
- 1+1 power redundancy
- N+1 fan module redundancy

Cisco Nexus 5548UP

The Cisco Nexus platform provides the following high availability and resiliency features:

- Redundant LAN and SAN fabrics
- Virtual port channel
- Spanning Tree Protocol
- Link aggregation control protocol
- Cisco fabric path
- Hot-swappable field-replaceable power supplies, fan modules, and expansion modules
- 1+1 power redundancy
- N+1 fan module redundancy

5.2 Network Security and Isolation

The network architecture must enable both security and isolation of network traffic. A variety of technologies can be used individually or in concert to assist with security and isolation, including:

- **VLANs.** VLANs enable traffic on one physical LAN to be subdivided into multiple virtual LANs or broadcast domains. This is accomplished by configuring devices or switch ports to tag traffic with specific VLAN IDs. A VLAN trunk is a network connection that is able to carry multiple VLANs, with each VLAN tagged with specific VLAN IDs.
- **ACLs.** Access control lists (ACLs) enable traffic to be filtered or forwarded based on a variety of characteristics such as protocol, source and destination port, and many other characteristics. ACLs can be used to prohibit certain traffic types from reaching the network or to enable or prevent traffic from reaching specific endpoints.
- **IPsec.** IPsec enables both authentication and encryption of network traffic to protect from both man-in-the-middle attacks as well as network sniffing and other data collection activities.
- **QoS.** QoS enables rules to be set based on traffic type or attributes so that one form of traffic does not block all others (by throttling it) or to make sure critical traffic has a certain amount of bandwidth allocated.

Additional security and isolation technologies include:

- Standard and extended Layer 2 ACLs (MAC addresses, protocol type, and so on)
- Standard and extended Layer 3 to Layer 4 ACLs (IPv4 and IPv6, Internet Control Message Protocol [ICMP], and TCP)
- User Datagram Protocol (UDP) and so on
- VLAN-based ACLs (VACLs)
- Port-based ACLs (PACLs)
- Named ACLs

- Optimized ACL distribution
- ACLs on virtual terminals (VTYs)
- Dynamic Host Configuration Protocol (DHCP) snooping with option 82
- Dynamic Address Resolution Protocol (ARP) Inspection
- IP source guard
- DHCP relay
- Cisco CTS (authentication and policy download from ACS)
- Ethernet port security

5.3 Network Automation

Remote interfaces and management of the network infrastructure through SSH or a similar protocol are important to both the automation and the resiliency of the data center network. Remote access and administration protocols can be used by management systems to automate complex or error-prone configuration activities. For example, adding a VLAN to a distributed set of access tier switches can be automated to avoid the potential for human error.

Cisco UCS Manager

Cisco UCS Manager offers the following features:

- A unified embedded management interface that integrates server, network, and storage access
- Policy and model-based management with service profiles that improve agility and reduce risk
- Autodiscovery to detect, inventory, manage, and provision system components that are added or changed
- A comprehensive open XML API that facilitates integration with third-party systems management tools
- Role-based administration that builds on existing skills and supports collaboration across disciplines

Cisco Nexus 5548UP

The Cisco Nexus platform offers the following related features:

- Switch management using 10/100/1000Mbps management or console ports
- CLI-based console to provide detailed out-of-band management
- In-band switch management
- Locator and beacon LEDs on Cisco Nexus 2000 Series
- Port-based locator and beacon LEDs
- Configuration synchronization
- Module pre-provisioning
- Configuration rollback
- Secure Shell version 2 (SSHv2)
- Telnet
- AAA
- AAA with RBAC
- RADIUS
- TACACS+
- Syslog (8 servers)
- Embedded packet analyzer

- SNMPv1, v2, and v3 (IPv4 and IPv6)
- Enhanced SNMP MIB support
- XML (NETCONF) support
- Remote monitoring (RMON)
- Advanced Encryption Standard (AES) for management traffic
- Unified username and passwords across CLI and SNMP
- Microsoft Challenge Handshake Authentication Protocol (MS-CHAP)
- Digital certificates for management between switch and RADIUS server
- Cisco Discovery Protocol versions 1 and 2
- RBAC
- Switched Port Analyzer (SPAN) on physical, PortChannel, VLAN, and FC interfaces
- Encapsulated Remote SPAN (ERSPAN)
- Ingress and egress packet counters per interface
- Network Time Protocol (NTP)
- Cisco GOLD
- Comprehensive bootup diagnostic tests
- Call Home
- Smart Call Home
- Cisco Fabric Manager
- Cisco DCNM
- CiscoWorks LAN Management Solution (LMS)

5.4 Virtualization Architecture

Virtualization is the heart of any private cloud. In this section, we discuss the virtualization architecture that supports this solution.

Storage Virtualization

Storage virtualization is a concept in IT system administration that refers to the abstraction (separation) of logical storage from physical storage so it can be accessed without regard to physical storage or heterogeneous structure. This separation gives system administrators increased flexibility in how they manage storage for end users. For more information on storage virtualization, refer to http://en.wikipedia.org/wiki/Storage_virtualization.

Network Virtualization

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

Network virtualization is categorized as either external, in which many networks or parts of networks are combined into a virtual unit, or internal, which provides network-like functionality to the software containers on a single system. Whether network 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 components:

- Network hardware, such as switches and network adapters, also known as NICs

- Networks, such as VLANs and virtual machines
- Network storage devices
- Network media, such as Ethernet and FC

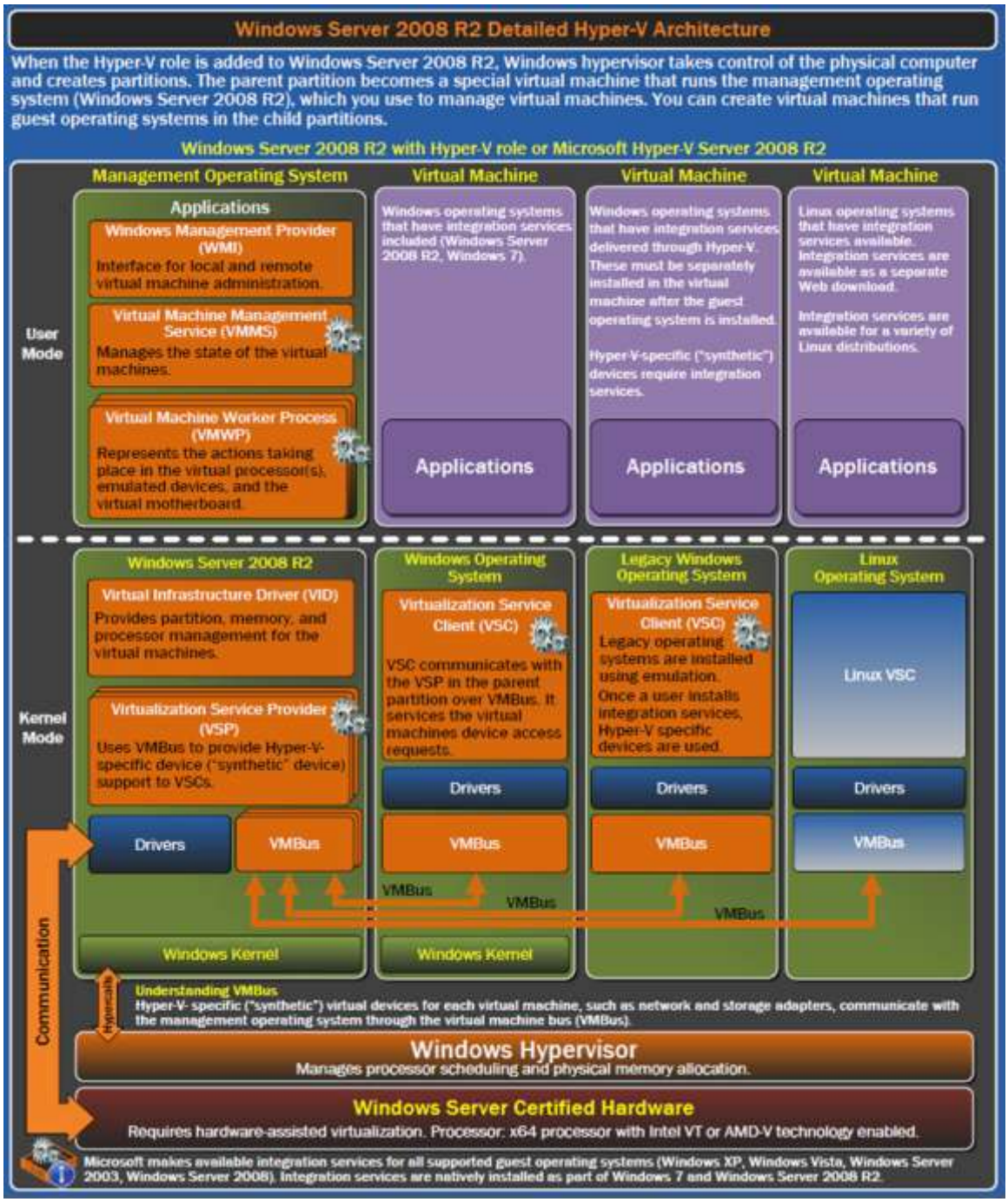
For more information about network virtualization, refer to www.snia.org/education/storage_networking_primer/stor_virt/.

Server Virtualization

Hardware virtualization uses software to create a virtual machine that emulates a physical computer. This creates a separate operating system environment that is logically isolated from the host server. By providing multiple virtual machines 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 effectively isolates the partitions by enforcing access policies for critical system resources such as memory and processors. Hyper-V does not contain any third-party device drivers or code, which minimizes its attack surface and provides a more secure architecture.

Figure 32) Hyper-V architecture (graphic provided by Microsoft).



In addition to the Windows hypervisor, there are two other major elements to consider in Hyper-V: a parent partition and a 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. By 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.

For more information, refer to Windows Server 2008 R2: Hyper-V Component Architecture.

5.5 Windows Server 2008 R2 SP1 and Hyper-V Host Design

The recommendations in this section adhere to the support statements made in Requirements and Limits for VMs and Hyper-V in Windows Server 2008 R2.

Licensing

Certain versions of Windows Server 2008 R2 SP1 (namely, Standard, Enterprise, and Datacenter editions) include virtualization use rights, which include a license to run a specified number of virtual machines based on Windows. Windows Server 2008 R2 SP1 Standard Edition includes use rights for one running virtual machine. Windows Server 2008 R2 SP1 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 Server guests are included. To run more than four, you must have valid Windows Server licenses for the additional virtual machines.

In contrast to the other two Windows Server editions, Windows Server 2008 R2 SP1 Datacenter Edition includes unlimited virtualization use rights, which, from a licensing standpoint, allows you to run as many Windows Server guests as you like on the licensed physical server.

Operating System Configuration

This section outlines the general considerations for the Hyper-V host operating system.

Note: These are not meant to be installation instructions, but rather the process requirements.

To install and use Hyper-V, the following hardware is required:

- **An x64-based processor.** Hyper-V is available in 64-bit editions of Windows Server 2008, specifically, the 64-bit editions of Windows Server 2008 Standard, Windows Server 2008 Enterprise, and Windows Server 2008 Datacenter. Hyper-V is not available for 32-bit (x86) editions or for Windows Server 2008 for Itanium-Based Systems. However, Hyper-V management tools are available for 32-bit editions.
- **Hardware-assisted virtualization.** This is available in processors that include a virtualization option, specifically, processors with Intel Virtualization Technology (Intel VT) or AMD Virtualization (AMD-V) technology.

Hardware-enforced data execution prevention (DEP) must be available and enabled. You must enable Intel XD bit (execute disable bit) or AMD NX bit (no execute bit).

- Use Windows Server 2008 R2, either the Full or Server Core installation option.

  Note: There is no upgrade path from Server Core to Full or from Full to Server Core, so make this selection carefully.

- Use the latest hardware device drivers.
- Join the Hyper-V parent partition operating system to a domain.
- Use the required Hyper-V server roles and failover clustering features.
- Apply relevant Windows updates, including out-of-band updates that aren't offered on Microsoft Update. For more information on Windows updates, refer to Hyper-V Update List for Windows Server 2008 R2.
- Validate all nodes, networks, and storage by using the Cluster Validation wizard.

Memory and Hyper-V Dynamic Memory

Dynamic Memory is a Hyper-V feature that helps you use physical memory more efficiently. With Dynamic Memory, Hyper-V treats memory as a shared resource that can be reallocated automatically among running virtual machines. Dynamic Memory adjusts the amount of memory available to a virtual machine based on changes in memory demand and values that you specify. Dynamic Memory is

available for Hyper-V in Windows Server 2008 R2 SP1. You can make the Dynamic Memory feature available by applying the service pack to the Hyper-V role in Windows Server 2008 R2 or in Microsoft Hyper-V Server 2008 R2.

For a complete description of Dynamic Memory features, settings, and design considerations, refer to the Hyper-V Dynamic Memory Configuration Guide. This guide provides the specific operating system, service pack, and integration component levels for supported operating systems. The guide also contains the minimum recommended startup RAM settings for all supported operating systems.

In addition to the general guidance provided, specific applications or workloads, particularly those with built-in memory management capability (such as SQL Server or Exchange), can provide workload-specific guidance. SQL Server 2008 R2 and the SQL Server product group have published best practice guidelines for Dynamic Memory in Running SQL Server with Hyper-V Dynamic Memory. The three companies recommend following the SQL Server best practices in this solution.

Storage Adapters

Unlike network adapters, storage adapters are certified by both the operating system and the storage provider. In this solution, the Cisco UCS platform has been carefully tuned to be compatible with both Windows Server and the Data ONTAP platform.

FC, iSCSI, and CNE HBA Configuration

Cisco UCS servers with the M81KR VIC adapter can present multiple HBAs to the operating system. The HBA attributes are configured as part of the service profiles. The adapter can operate at wire speed and is pinned to fabric A or fabric B. The adapters are capable of booting from a SAN or iSCSI target array. iSCSI adapters use the Microsoft Software Initiator. High availability is provided by MPIO software running in the operating system.

MPIO Configuration

Microsoft MPIO architecture supports iSCSI, FC, and 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 NIC (in the iSCSI case) or HBA should be connected using redundant switch infrastructures to provide continued access to storage in the event of a failure in a storage fabric component.

The joint best practice recommendation is to use the NetApp DSM in all cases.

For more information about MPIO best practices, refer to TR-3702: NetApp Storage Best Practices for Microsoft Virtualization and NetApp SnapManager for Hyper-V.

Network Adapters

Network adapters and the way in which the network is configured have a direct correlation to the health and stability of your Windows failover cluster. This solution provides a Microsoft best practice networking environment designed for maximum performance and availability.

Protocol Bindings

Each network in the solution has a different requirement for network bindings. Table 6 contains the recommended settings for this solution.

Table 6) Recommended network bindings.

Setting	Management Network Adapter	Heartbeat Network Adapter	Live Migration Network Adapter	iSCSI Network Adapter	Guest Virtual Machine Network Adapter
Client for Microsoft networks	Yes	No	Yes	No	No
File and printer sharing	Yes	No	Yes	No	No
Microsoft Virtual Network Switch Protocol	No	No	No	No	Yes
Internet Protocol Version 6	Optional	Optional	Optional	Optional	No
Internet Protocol Version 4	Yes	Yes	Yes	Yes	No
Link-layer topology discovery mapper I/O driver	Yes	No	No	No	No
Link-layer topology discovery responder	Yes	No	No	No	No

Performance Settings

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

- TCP checksum offload benefits both CPU and overall network throughput performance, and it is fully supported by live migration.
- Jumbo frames capability is extended to VMS with Hyper-V in Windows Server 2008 R2. Just as 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 overall throughput and also reduces CPU utilization for large file transfers.
- Virtual machine queue (VMQ) allows the host's single NIC card to appear as multiple NICs to the virtual machines by allowing the host's NIC direct memory access (DMA) packets directly into individual virtual machine memory stacks. Each virtual machine device buffer is assigned a VMQ, which eliminates needless packet copies and route lookups in the virtual switch. The results are less data in the host's buffers and an overall performance improvement for I/O operations.

5.6 Hyper-V Host Failover Cluster Design

A Hyper-V host failover 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 software. If one of the cluster nodes fails, another node begins to provide service in a process known as failover. In case of a planned migration, also called live migration, 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.

Host Failover Cluster Topology

In FlexPod with Microsoft Private Cloud, there are two standard design patterns. It is recommended that the server topology consists of at least two Hyper-V host clusters. The first cluster, which in the following

examples is referred to as the management cluster, should have at least two nodes. The second and any additional clusters are called the fabric host clusters in the examples that follow.

In some cases, such as in smaller scale scenarios or specialized solutions, the management and fabric clusters can be consolidated onto the fabric host cluster. Special care must be taken in those cases to have sufficient resource availability for the virtual machines that host the various parts of the management stack. Each host cluster can contain up to 16 nodes. Host clusters require some form of shared storage such as an FC or iSCSI SAN.

Host Cluster Networks

A variety of host cluster networks are required for a Hyper-V failover cluster. The network requirements enable high availability and high performance. For more information about the specific requirements and recommendations for network configuration, refer to Hyper-V: Live Migration Network Configuration Guide.

The FlexPod with Microsoft Private Cloud configurations are composed of the following elements:

- **Management network.** A dedicated management network is required so that hosts can be managed through a dedicated network to avoid competition with guest traffic needs. A dedicated network provides a degree of separation for security and ease of management. This typically implies dedicating one network adapter per host and one port per network device to the management network. This network is used for remote administration of the host, communication to management systems (System Center agents), and other administrative tasks.

  Additionally, Cisco offers out-of-band management for Cisco UCS. For more information, refer to Cisco's Web site. For more information about Hyper-V networking requirements, refer to Hyper-V: Live Migration Network Configuration Guide.
- **iSCSI network.** If using iSCSI, a dedicated iSCSI network is required so that storage traffic is not competing with any other traffic. This typically implies dedicating two network adapters per host and two ports per network device to the storage network. For all iSCSI storage connections, an MPIO configuration with two independent physical ports is required.
- **CSV and cluster communication network.** Usually when the cluster node that owns a VHD file in CSV performs disk I/O, the node communicates directly with the storage devices, through a SAN, for example. However, storage connectivity failures sometimes prevent a given node from communicating directly with the storage device. To maintain functionality 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.
- **Live migration network.** During live migration, the contents of the memory of the virtual machine running on the source node need to be transferred to the destination node over a LAN connection. To provide a high-speed transfer, a dedicated redundant 10Gbps live migration network is required. In this case, all Ethernet networking uses the 10GB converged network that is provided by the Cisco 5548 switches that are specified as part of this solution. This significantly reduces the time required to evacuate the virtual machines from a host with zero downtime during maintenance or Windows updates. QoS can be used so that sufficient bandwidth is reserved for this network.
- **Virtual machine networks.** The virtual machine network or networks are dedicated to VM LAN traffic. The virtual machine network can be based on two or more 10GbE networks, one or more networks created using NIC teaming, or virtual networks created from shared 10GbE NICs. Implement one or more dedicated virtual machine networks.

Host Failover Cluster Storage

CSV is a feature that simplifies the configuration and management of Hyper-V VMs in failover clusters. With CSV on a failover cluster that runs Hyper-V, multiple virtual machines can use the same LUN and still 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 VHD files. If you 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 Microsoft Hyper-V Server 2008 R2 that include failover clustering.

5.7 Hyper-V Guest Virtual Machine Design

Standardization is a key tenet of private cloud architectures. This also applies to virtual machines. A standardized collection of virtual machine templates can both drive predictable performance and greatly improve capacity planning capabilities.

Table 7 contains information about what a basic virtual machine template library can look like.

Table 7) Example of a basic virtual machine template library.

Template	Specs	Network	Operating System	Unit Cost
Template 1: small	1 vCPU, 2GB memory, 50GB disk	VLAN 20	WS 2003 R2	1
Template 2: medium	2 vCPUs, 4GB memory, 100GB disk	VLAN 20	WS 2003 R2	2
Template 3: large	4 vCPUs, 8GB memory, 200GB disk	VLAN 20	WS 2003 R2	4
Template 4: small	1 vCPU, 2GB memory, 50GB disk	VLAN 10	WS 2008	1
Template 5: medium	2 vCPUs, 4GB memory, 100GB disk	VLAN 10	WS 2008	2
Template 6: large	4 vCPUs, 8GB memory, 200GB disk	VLAN 10	WS 2008	4

Note: Use standard documented virtual machine configurations for all virtual machines, management, and tenants used for fabric management or for workload deployment by tenants.

Virtual Machine Storage

This section discusses the different types of Hyper-V disks.

Microsoft recommends using only fixed VHDs for production. The following quote is from Windows TechNet:

> "Why are fixed VHDs recommended for production?
>
> Fixed VHDs are recommended for production instead of dynamically expanding or differencing VHDs for the following reasons:
>
> - The I/O performance is highest for fixed VHDs because the file is not dynamically expanded.
> - When a dynamically expanding disk is expanded, the host volume could run out of space and cause the write operations to fail. Using fixed VHDs prevents this from happening.
> - The file data will not become inconsistent due to lack of storage space or power loss. Dynamically expanding and differencing VHDs depend on multiple write operations to expand the file. The internal block allocation information can become inconsistent if all I/O operations to the VHD file and the host volume are not complete and persist on the physical disk. This can happen if the computer suddenly loses power."

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 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. When the full capacity is allocated 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 (using the disk's Global Unique ID [GUID]) without having to use the host's file system. The performance difference between fixed disk and pass-through disks is negligible, so the decision about which type to use is based on manageability. For example, if the data on the volume is very large (hundreds of gigabytes), a VHD is hardly portable at that size given the extreme amount of time it takes to copy. Also consider the backup scheme. With pass-through disks, the data can only be backed up from within the guest. When using pass-through disks, there is no VHD file created; the LUN is used directly by the guest. Because there is no VHD file, there is no dynamic sizing capability.

In-Guest iSCSI Initiator

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

Virtual Machine Networking

Hyper-V guests support two types of virtual network adapters: synthetic and emulated. Synthetic adapters use Hyper-V BUS architecture and are the high-performance, native device. Synthetic adapters require that Hyper-V integration services be installed within the guest. Emulated adapters are available to all guests even if integration services are not available. They are perform much more slowly and only should be used if synthetic adapters are unavailable.

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

- **Private network.** This type of virtual network allows communications only between virtual machines.
- **Internal network.** This type of virtual network allows communications between the host server and virtual machines.
- **External network.** This type of virtual network allows communications between a virtual machine and a physical network by creating an association to a physical network adapter on the host server.

Virtual Processors

Table 8 contains information about the number of virtual processors in a Hyper-V guest.

Note: The information in Table 8 and Table 9 is somewhat dynamic. Improvements to the integration services for Hyper-V are periodically released to add support for additional operating systems. For the most current information, refer to the Supported guest operating systems topic on TechNet.

Table 8) Supported guest operating systems (table data provided by Microsoft).

Server Guest Operating System	Editions	Virtual Processors
Windows Server 2008 R2 with SP1	Standard, Enterprise, Datacenter, and Web editions	1, 2, 3, or 4
Windows Server 2008 R2	Standard, Enterprise, Datacenter, and Windows Web Server 2008 R2	1, 2, 3, or 4
Windows Server 2008	Standard, Standard without Hyper-V, Enterprise, Enterprise without Hyper-V, Datacenter, Datacenter without Hyper-V, Windows Web Server 2008, and HPC Edition	1, 2, 3, or 4
Windows Server 2003 R2 with SP2	Standard, Enterprise, Datacenter, and Web	1 or 2
Windows Home Server 2011	Standard	1, 2, or 4
Windows Storage Server 2008 R2	Essentials	1, 2, or 4
Windows Small Business Server 2011	Essentials	1 or 2
Windows Small Business Server 2011	Standard	1, 2, or 4
Windows Server 2003 R2 x64 Edition with SP2	Standard, Enterprise, and Datacenter	1 or 2
Windows Server 2003 with SP2	Standard, Enterprise, Datacenter, and Web	1 or 2
Windows Server 2003 x64 Edition with SP2	Standard, Enterprise, and Datacenter	1 or 2
Windows 2000 Server with SP4	Server, Advanced Server	1
CentOS 6.0 and 6.1	x86 edition and x64 edition	1, 2, or 4
CentOS 5.2-5.7	x86 edition and x64 edition	1, 2, or 4
Red Hat Enterprise Linux® 6.0 and 6.1	x86 edition and x64 edition	1, 2, or 4
Red Hat Enterprise Linux 5.7	x86 edition and x64 edition	1, 2, or 4
Red Hat Enterprise Linux 5.6	x86 edition and x64 edition	1, 2, or 4

Server Guest Operating System	Editions	Virtual Processors
Red Hat Enterprise Linux 5.5	x86 edition and x64 edition	1, 2, or 4
Red Hat Enterprise Linux 5.4	x86 edition and x64 edition	1, 2, or 4
Red Hat Enterprise Linux 5.3	x86 edition and x64 edition	1, 2, or 4
Red Hat Enterprise Linux 5.2	x86 edition and x64 edition	1, 2, or 4
SUSE Linux Enterprise Server 11 with SP1	x86 edition and x64 edition	1, 2, or 4
SUSE Linux Enterprise Server 10 with SP4	x86 edition and x64 edition	1, 2, or 4

Table 9) Supported client guest operating systems (table data provided by Microsoft).

Client Guest Operating System	Editions	Virtual Processors
Windows 7 with SP1	Enterprise, Ultimate, and Professional. This applies to 32-bit and 64-bit editions, as well as N and KN editions.	1, 2, 3, or 4
Windows 7	Enterprise, Ultimate, and Professional. This applies to 32-bit and 64-bit editions, as well as N and KN editions.	1, 2, 3, or 4
Windows Vista®	Business, Enterprise, and Ultimate, including N and KN editions	1 or 2
Windows XP with SP3 **Note:** Performance might be degraded on Windows XP with SP3 if the server running Hyper-V uses an AMD processor. For more information, refer to Degraded I/O Performance Using a Windows XP Virtual Machine with Windows Server 2008 Hyper-V.	Professional	1 or 2
Windows XP with SP2 **Note:** Support for this operating system ends on July 13, 2010.	Professional	1
Windows XP x64 Edition with SP2	Professional	1 or 2

Hyper-V supports a maximum ratio of 8 virtual processors (VPs) to 1 logical processor (LP) for server workloads, and 12 VPs per 1 LP for VDI workloads. A logical processor is defined as a processing core that is seen by the host operating system or parent partition. In the case of Intel hyperthreading, each thread is considered an LP.

Therefore, a 16-LP server supports a maximum of 128 VPs. That equates to 128 single-processor virtual machines, 64 dual-processor virtual machines, or 32 quad-processor virtual machines. The 8:1 or 12:1 VP-to-LP ratios represent the maximum supported limits. The joint best practice recommendation is for lower limits to be used rather than the maximum.

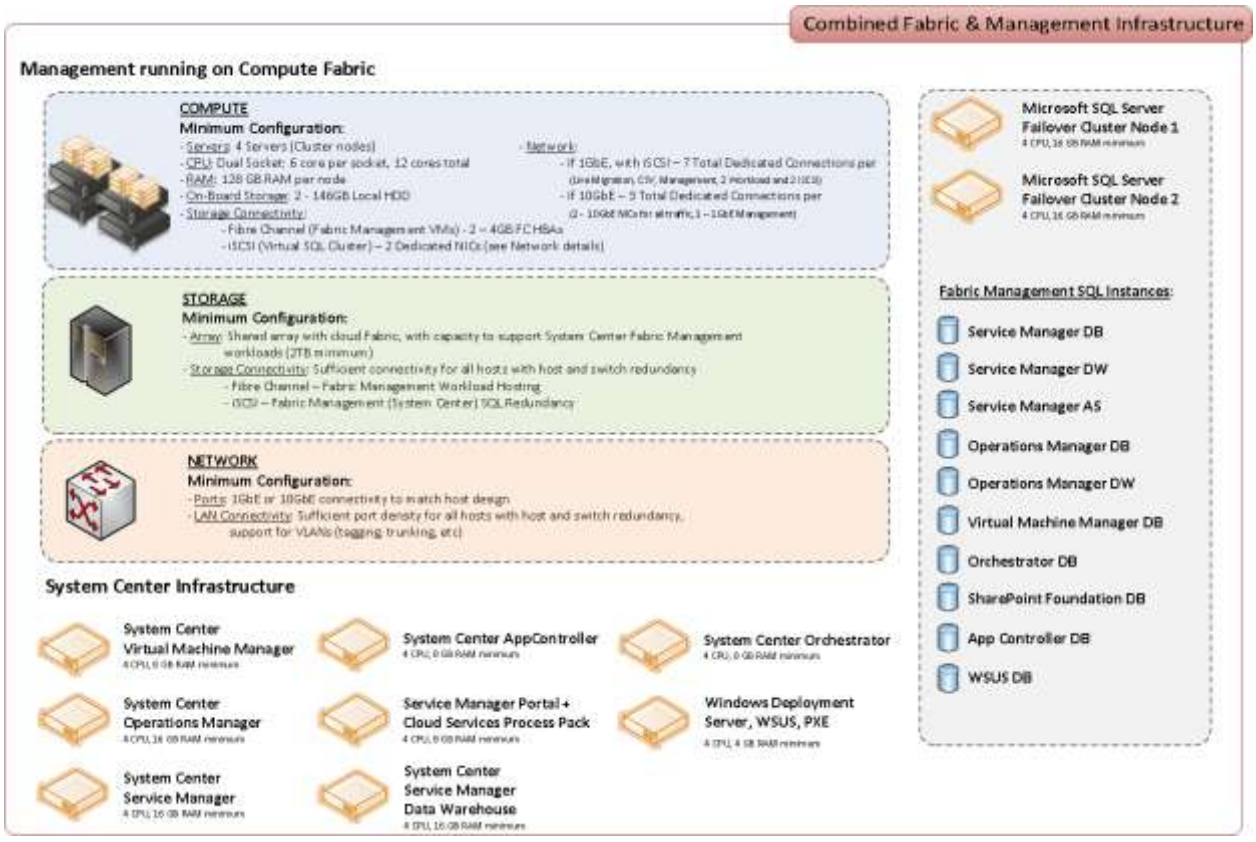
5.8 Management Architecture

In this section, we discuss the details of the underling management architecture that is included in the solution.

Management Logical Architecture

Figure 33 depicts the management logical architecture running the management systems directly on the fabric cluster.

Figure 33) Management running on compute fabric (graphic provided by Microsoft).



The management architecture consists of a minimum of four physical nodes in a failover cluster with SAN-attached storage that supports iSCSI and redundant network connections to provide a highly available platform for the management systems and capacity for workloads. In this scenario, we have scaled down the high-availability options for management to facilitate a smaller management footprint on the fabric.

The management systems include:

- 9 SQL Servers instances in a guest cluster configuration
- 1 VMM server
- 1 Operations Manager server

- 1 Orchestrator server
- 1 Service Manager server
- 1 Service Manager data warehouse
- 1 Service Manager self-service portal with Cloud Service Process Pack
- 1 application controller
- 1 deployment server providing WDS, PXE, and WSUS

Management System Architecture

This solution provides a complete management stack for your private cloud. This section discusses the architecture in detail.

Prerequisite Infrastructure

The following section outlines the management system architecture and its dependencies within a customer environment.

Active Directory Domain Services

Active Directory Domain Services (AD DS) are a required foundational component of the solution. The solution provides support for Windows Server 2008 and Windows Server 2008 R2 SP1 AD DS customer deployments.

Previous versions are not directly supported for all workflow provisioning and de-provisioning automation. It is assumed that AD DS deployments exist at the customer site, and deployment of these services is not within the scope of the typical deployment.

- **Forests and domains.** The preferred approach is to integrate into an existing AD DS forest and domain. This is not a difficult requirement; a dedicated resource forest or domain can also be employed as an additional part of the deployment. The solution supports multiple domains and multiple forests in a trusted environment by using two-way forest trusts.
- **Trusts.** The solution enables multi-domain support within a single forest where two-way forest (Kerberos) trusts exist between all domains. This is referred to as multi-domain or interforest support. Also supported are interforest or multiforest scenarios, as well as intraforest environments.

DNS

DNS name resolution is a required element for System Center 2012 components and the process automation solution.

DNS integrated with Active Directory is required for automated provisioning and de-provisioning components within the System Center Orchestrator runbook as part of the solution. We provide full support and automation for Windows Server 2008 and Windows Server 2008 R2 SP1 DNS deployments integrated with Active Directory.

Using solutions that are not based on Microsoft or DNS integrated with Active Directory might be possible, but do not automatically create or remove DNS records that are related to virtual machine provisioning and de-provisioning processes. Using solutions outside of DNS integrated with Active Directory would require either manual intervention for these scenarios or modifications to Cloud Services Process Pack Orchestrator runbooks.

DHCP

To support the dynamic provisioning and management of physical and virtual compute capacity within the IaaS infrastructure, use DHCP for all physical machines and virtual machines by default to support runbook automation. For physical hosts, such as the fabric management cluster nodes and the scale-unit cluster nodes, DHCP reservations are recommended so that physical servers and NICs have known IP addresses while providing centralized management of those addresses through DHCP.

Windows DHCP is required for automated provisioning and de-provisioning components within System Center Orchestrator runbooks as part of the solution. This is used to support host cluster provisioning, DHCP reservations, and other areas supporting dynamic provisioning of compute within the infrastructure. We provide full support and automation for Windows Server 2008 and Windows Server 2008 R2 SP1 versions of the DHCP server role. Use of solutions that are outside of the Windows DHCP server role requires additional testing and validation activities.

SQL Server

Two SQL Server virtual machines will be deployed as a guest failover cluster to support the solution (with an option to scale to a 4 node cluster). This multi-node SQL Server failover cluster will contain all the databases for each System Center product in discrete instances by product and function. This separation of instances allows for division by unique requirements and scale over time as the needs of each component scales higher. Note that not all features are supported for failover cluster installations, some features cannot be combined on instances and some allow configuration only at initial installation. As a general rule, Database Engine Services and Analysis Services will be hosted in separate instances within the failover cluster. Because of the support for SQL Server Reporting Services (SSRS) in a failover cluster, SSRS will be installed on the hosting System Center component server (the Operations Manager Reporting Server). This installation however will be "Files Only" and the SSRS configuration will configure remote Reporting Services databases hosted on the component instance on the SQL cluster. The exception to this is the SCOM Analysis Services and Reporting Services configuration. For this instance, Analysis Services and Reporting Services must be installed with the same server and with the same instance to support VMM and OM integration. All instances are required to be configured with Windows Authentication. The table below outlines the options required for each instance.

Database Instances and Requirements

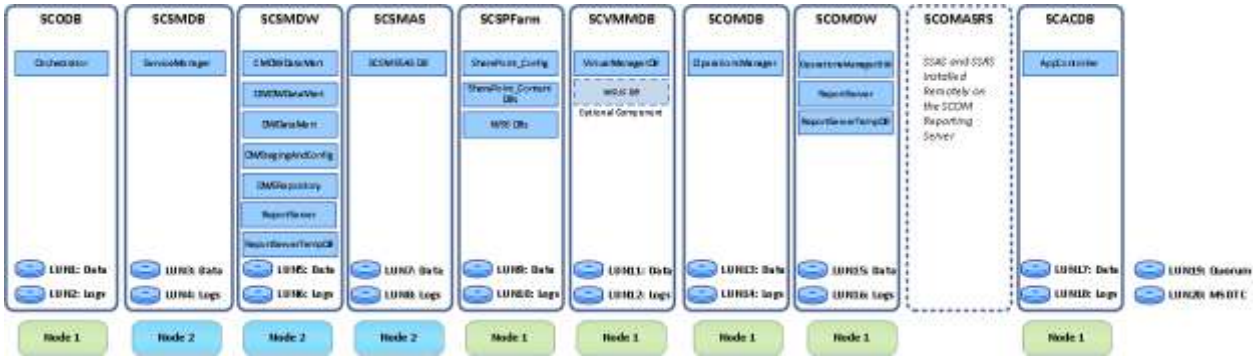
Fabric Management Component	Instance Name (Suggested)	Components	Collation[1]	Storage Requirements
Virtual Machine Manager	SCVMMDB	Database Engine	SQL_Latin1_General_CP1_CI_AS	2 LUNs
Operations Manager	SCOMDB	Database Engine, Full-Text Search	SQL_Latin1_General_CP1_CI_AS	2 LUNs
Operations Manager Data Warehouse	SCOMDW	Database Engine, Full-Text Search	SQL_Latin1_General_CP1_CI_AS	2 LUNs
Service Manager	SCSMDB	Database Engine, Full-Text Search	SQL_Latin1_General_CP1_CI_AS	2 LUNs
Service Manager Data Warehouse	SCSMDW	Database Engine, Full-Text Search	SQL_Latin1_General_CP1_CI_AS	2 LUNs
	SCSMAS	Analysis Services	SQL_Latin1_General_CP1_CI_AS	2 LUNs
	SCSPFarm	Database Engine	SQL_Latin1_General_CP1_CI_AS	2 LUNs
Orchestrator	SCODB	Database Engine	SQL_Latin1_General_CP1_CI_AS	2 LUNs
App Controller	SCACDB	Database Engine	SQL_Latin1_General_CP1_CI_AS	2 LUNs

[1] The default SQL collation settings are not supported for multi-lingual installations of the Service Manager component. Only use the default SQL collation if multiple languages are not required. Note that the same collation must be used for all Service Manager databases (Management, Data Warehouse and Reporting Services).

Windows Server Update Services (optional)	SCWSUSDB	Database Engine	SQL_Latin1_General_CP1_CI_AS	2 LUNs

Each SQL Server virtual machine will be configured with four vCPUs, at least 16 GB of RAM (32 GB is recommended for large scale configurations), and four vNICs (for LAN, Cluster Communications, iSCSI). Each SQL Server virtual machine will access iSCSI-based shared storage with two LUNs configured for each hosted database. Should the needs of the Solution exceed what two SQL virtual machines are able to provide, additional virtual machines can be added to the virtual SQL Server cluster and each SQL Server instance moved to its own virtual machine in the cluster. This configuration requires SQL Server 2008 R2 Enterprise Edition. In addition, where organizations can support SSD storage, it should be used to provide the necessary I/O for these databases. The instances and associated recommended node placement is outlined below:

Figure 34) SQL Server instances and recommended placement (graphic provided by Microsoft).



SQL Server Configuration

- 2 HA virtual machines on different Hyper-V hosts
- Windows Server 2008 R2 SP1 Enterprise Edition
- 4 vCPUs
- 16GB RAM

  Note: Do not use dynamic memory.

- 4 vNICs (1 for client connection, 1 for cluster communication, and 2 for iSCSI fabric connections)
- Storage: 1 operating system VHD and 20 iSCSI LUNs

Table 10) SQL Server data locations.

LUN	Purpose	Size
LUN 1, CSV	Virtual machine operating system	100GB VHD
LUN 2, iSCSI	MSDTC	1GB
LUN 3, iSCSI	SQL Server cluster quorum	1GB
LUN 4, iSCSI	VMM data	Varies
LUN 5, iSCSI	VMM logs	Varies

LUN	Purpose	Size
LUN 6, iSCSI	SCOM data	Varies
LUN 7, iSCSI	SCOM logs	Varies
LUN 8, iSCSI	SCO data	Varies
LUN 9, iSCSI	SCO logs	Varies
LUN 10, iSCSI	SCSM data	Varies
LUN 11, iSCSI	SCSM logs	Varies
LUN 12, iSCSI	SCAC data	Varies
LUN 13, iSCSI	SCAC logs	Varies
LUN 14, iSCSI	SCOM DW data	Varies
LUN 15, iSCSI	SCOM DW logs	Varies
LUN 16, iSCSI	SCSMAS data	Varies
LUN 17, iSCSI	SCSMAS logs	Varies
LUN 18, iSCSI	SCSM DW data	Varies
LUN 19, iSCSI	SCSM DW logs	Varies
LUN 20, iSCSI	SCSP farm data	Varies
LUN 21, iSCSI	SCSP farm logs	Varies

Table 11) Databases.

DB Client	Instance Name	DB Name	Authentication
VMM	<Instance 1>	<VMM_DB>	Win Auth
Ops Mgr	<Instance 2>	<Ops_Mgr_DB>	Win Auth
Ops Mgr	<instance 3>	<Ops_Mgr_DW>	Win Auth
Svc Mgr	<Instance 4>	<Svc Mgr_DB>	Win Auth
Svc Mgr	<Instance 5>	<Svc Mgr_DW>	Win Auth
Orchestrator	<Instance 6>	<Orchestrator_DB>	Win Auth
App Controller	<Instance 7>	<AppController_DB>	Win Auth
Svc Mgr	<Instance 8>	<Svc Mgr_AS>	Win Auth
SC SharePoint	<Instance 9>	<SCSP_DB>	Win Auth

Virtual Machine Manager

System Center Virtual Machine Manager 2012 is required. VMM is deployed against a dedicated SQL Server instance running on the virtualized SQL Server cluster.

Servers

The following hardware configurations are used:

- 1 HA virtual machine
- Windows Server 2008 R2 SP1
- 4 vCPUs
- 8GB RAM
- 3 vNICs (1 for client access, 2 for iSCSI)
- Storage: 1 operating system VHD and 1 iSCSI LUN

Operations Manager

System Center Operations Manager 2012 is required. Operations Manager is deployed using a dedicated SQL Server instance on the virtualized SQL Server cluster. An Operations Manager agent gets installed on every guest virtual machine, as well as on every management host and scale unit cluster node, to support health-monitoring functionality.

Note: Operations Manager gateway servers and additional management servers are supported for custom solutions. However, for the base-reference implementation, these additional roles are not implemented.

The Operations Manager installation uses a dedicated SQL Server instance on the virtualized SQL Server cluster. The installation follows a split SQL Server configuration: SQL Server Reporting Services (SSRS) and OpsMgr components reside on the OpsMgr VM, while the SSRS and OpsMgr databases use a dedicated instance on the virtualized SQL Server cluster. The estimated SQL Server database sizes are:

- 72GB Operations Manager Database, 2.1TB Operations Manager Data Warehouse Database

The following hardware configurations are used for Operations Manager management servers:

- 1 HA virtual machine
- Windows Server 2008 R2 SP1
- 4 vCPUs
- 16GB RAM
- 1 vNIC
- Storage: 1 operating system VHD

The following Operations Manager management packs are required:

- Virtual Machine Manager 2012
- 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 management pack
- Server OEM third party management pack

Service Manager

System Center Service Manager 2012 is optional for Service Manager. The Service Manager management server is installed on two virtual machines. A third virtual machine hosts the Service Manager data warehouse server. Both the Service Manager database and the data warehouse database use a dedicated SQL Server instance on the virtualized SQL Server cluster. The Service Manager portal

is hosted on a fourth virtual machine with the portal. The following virtual machine configuration is used for Service Manager management servers:

- 1 HA virtual machine
- Windows Server 2008 R2 SP1
- 4 vCPUs
- 16GB RAM
- 1 vNIC
- Storage: 1 operating system VHD

The following virtual machine configuration is used for Service Manager data warehouse servers:

- 1 HA virtual machine
- Windows Server 2008 R2 SP1
- 4 vCPUs
- 16GB RAM
- 1 vNIC
- Storage: 1 operating system VHD

The following virtual machine configuration is used for Service Manager portal servers:

- 1 HA virtual machine
- Windows Server 2008 R2 SP1
- 4 vCPUs
- 8GB RAM
- 1 vNIC
- Storage: 1 operating system VHD

The following virtual machine configuration is used for Service Manager estimated SQL Server database sizes:

- 40GB Service Manager database and 80GB Service Manager data warehouse

Orchestrator

The Orchestrator installation uses a dedicated SQL Server instance on the virtualized SQL Server cluster.

The following configuration is used for Orchestrator servers:

- 1 HA virtual machine
- Windows Server 2008 R2 SP1
- 4 vCPUs
- 8GB RAM
- 1 vNIC
- Storage: 1 operating system VHD

App Controller

System Center App Controller is optional unless the Service Manager Portal is used, and then App Controller must be the installer. App Controller uses a dedicated SQL Server instance on the virtualized SQL Server cluster. A single App Controller server is installed on the host cluster.

Service Manager provides the service catalog and service request mechanism, Orchestrator provides the automated provisioning, and App Controller provides the end-user interface to connect to and manage post-provisioning workloads.

The following configuration is used for the App Controller server:

- 1 HA virtual machine
- Windows Server 2008 R2 SP1
- 4 vCPUs
- 8GB RAM
- 1 vNIC
- Storage: 1 operating system VHD

Management Scenarios

The following are the primary management scenarios addressed in Fast Track, although the management layer can provide many more capabilities:

- Fabric management
- Fabric provisioning
- IT service provisioning (including platform and application provisioning)
- Virtual machine provisioning and de-provisioning
- Fabric and IT service maintenance
- Fabric and IT service monitoring
- Resource optimization
- Service management
- Reporting (used by chargeback, capacity, service management, health, and performance)
- Backup and DR
- Security

Fabric Management

Fabric management is the act of pooling together multiple disparate computing resources and subdividing, allocating, and managing them as a single fabric. There are various fabric management methods.

Hardware Integration: Storage

In VMM, you can discover, classify, and provision remote storage on supported storage arrays through the VMM console. VMM fully automates the assignment of storage to a Hyper-V host or Hyper-V host cluster, and it tracks the storage that is managed by VMM.

To enable the new storage features, VMM uses the new Microsoft storage management service to communicate with external arrays through a storage management initiative-specification (SMI-S) provider. The storage management service is installed by default during the installation of VMM. You must install a supported SMI-S provider on an available server and then add the provider to VMM management.

However, SMI-S is designed to be compatible with all storage arrays. This means that SMI-S cannot support the deep integration provided by NetApp tools such as OnCommand plug-in for Microsoft (OCPM). For this reason, the joint best practice recommendation is for storage and virtual machine provisioning to be performed using NetApp OCPM 3.1.

Hardware Integration: Network

Networking in VMM includes several enhancements that enable administrators to efficiently provision network resources for a virtualized environment. The networking enhancements include the following:

- **The ability to create and define logical networks.** A logical network together with one or more associated network sites is a user-defined named grouping of IP subnets, VLANs, or IP subnet and VLAN pairs that is used to organize and simplify network assignments. Some possible examples include back end, front end, lab, management and backup. Logical networks represent an abstraction of the underlying physical network infrastructure, which enables you to model the network based on business needs and connectivity properties. After a logical network is created, it can be used to specify the network on which a host or a virtual machine (standalone or part of a service) is deployed. Users can assign logical networks as part of virtual machine and service creation without having to understand the network details.
- **Static IP address and MAC address pool assignment.** If you associate one or more IP subnets with a network site, you can create static IP address pools from those subnets. Static IP address pools enable VMM to automatically allocate static IP addresses to virtual machines based on Windows that are running on any supported and managed Hyper-V, VMware® ESX® or Citrix XenServer host. VMM can automatically assign static IP addresses from the pool to standalone virtual machines, to virtual machines that are deployed as part of a service, and to physical computers when you use VMM to deploy them as Hyper-V hosts. Additionally, when you create a static IP address pool, you can define a reserved range of IP addresses for load balancers' virtual IP (VIP) addresses. VMM automatically assigns a virtual IP address to a load balancer during the deployment of a load-balanced service tier.

Fabric Provisioning

End-to-end fabric provisioning can be accomplished through Windows PowerShell, through System Center Orchestrator, or manually. Automated provisioning of storage, compute, and software resources is possible due to the combined commitment of NetApp, Cisco, and Microsoft to standardized management interfaces such as Windows PowerShell.

VMM Private Clouds

After you have configured the fabric resources (such as storage, networking, library servers and shares, host groups, and hosts), you can subdivide and allocate them for self-service consumption through the creation of VMM private clouds. During private cloud creation, select the underlying fabric resources that will be available in the private cloud, configure library paths for private cloud users, and set the capacity for the private cloud. For example, you might want to create a cloud for use by the finance department. In such a scenario you can:

- Name the cloud (finance, for example).
- Scope it to one or more host groups.
- Select which logical networks, load balancers, and VIP templates are available to the cloud.
- Specify which storage classifications are available to the cloud.
- Select which library shares are available to the cloud for virtual machine storage.
- Specify granular capacity limits to the cloud (virtual CPUs, memory, and so on).
- Select which capability profiles are available to the cloud. Capability profiles match the type of hypervisor platforms that are running in the selected host groups. The built-in capability profiles represent the minimum and maximum values that can be configured for a virtual machine for each supported hypervisor platform.

Virtual Machine Provisioning and Deprovisioning

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 Microsoft Private Cloud

solution, this refers to the ability of 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 virtual machine provisioning and deprovisioning process. This process is initiated from the self-service portal or tenant user interface, and it triggers an automated process or workflow in the infrastructure through system center VMM to either create or delete a virtual machine based on the authorized settings input by the user or tenant. Provisioning could be template based, such as requesting a small, medium, or large virtual machine template, or a series of selections could be made by the user (vCPUs, RAM, and so on). If authorized, the provisioning process should create a new virtual machine at the user's request, add the virtual machine to any relevant management products in the Microsoft Private Cloud (such as System Center), and enable access to the virtual machine by the requestor.

IT Service Provisioning

In VMM, a service is a set of virtual machines that are configured and deployed together and are managed as a single entity, for example, a deployment of a multitier line-of-business application.

In the VMM console, you use the Service Template Designer to create a service template that defines the configuration of the service. The service template includes information about the virtual machines that are deployed as part of the service, which applications to install on the virtual machines, and the networking configuration needed for the service, including the use of a load balancer.

Resource Optimization

Elasticity, perception of infinite capacity, and perception of continuous availability are Microsoft Private 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.

VMM Dynamic Optimization migrates virtual machines to perform resource balancing within host clusters that support live migration according to the settings you enter.

Dynamic Optimization corrects three possible scenarios, in priority order:

1. Virtual machines that have configuration problems on their current host
2. Virtual machines that are causing their host to exceed configured performance thresholds
3. Unbalanced resource consumption on hosts

VMM Power Optimization is an optional feature of Dynamic Optimization, and it is only available when a host group is configured to migrate virtual machines through Dynamic Optimization. Through Power Optimization, VMM helps to save energy by turning off hosts that are not needed to meet resource requirements within a host cluster and by turning the hosts back on when they are needed again.

By default, VMM performs Power Optimization all of the time when the feature is turned on. However, you can schedule the hours and days during the week when Ppower Optimization is performed. For example, you might initially schedule Power Optimization only on weekends, when you anticipate low resource usage on your hosts. After observing the effects of Power Optimization in your environment, you might increase the hours.

For Power Optimization, computers must have a baseboard management controller (BMC) that enables out-of-band management.

Fabric and IT Service Maintenance

The Microsoft Private Cloud must enable the performance of maintenance on any component of the solution without affecting the availability of the solution. Examples include the need to update or patch a host server, add additional storage to the SAN, and so on. During maintenance, the system should make

sure that unnecessary alerts or events are not generated in the management systems during planned maintenance.

VMM 2012 includes the built-in ability to maintain the fabric servers in a controlled, orchestrated manner.

Fabric servers include the following physical computers managed by VMM: Hyper-V hosts and Hyper-V clusters, library servers, preboot execution environment (PXE) servers, the Windows Server Update Management (WSUS) server, and the VMM management server.

VMM supports on-demand compliance scanning and remediation of the fabric. Administrators can monitor the update status of the servers. They can scan for compliance and remediate updates for selected servers. Administrators also can exempt resources from installation of an update.

VMM supports orchestrated updates of Hyper-V host clusters. When a VMM administrator performs update remediation on a host cluster, VMM places one cluster node at a time in maintenance mode and then installs updates. If the cluster supports live migration, intelligent placement is used to migrate virtual machines off the cluster node. If the cluster does not support live migration, VMM saves state for the virtual machines.

The feature requires the use of a WSUS server.

Fabric and IT Service Monitoring

The Microsoft Private Cloud must enable monitoring of 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.

Monitoring of the fabric is performed by integrating Operations Manager and VMM. Enabling this integration allows Operations Manager to automatically discover, monitor, and report on essential characteristics of any object managed by VMM, such as:

- Health and performance of all VMM-managed hosts and virtual machines
- Diagram views in Operations Manager reflecting all VMM deployed hosts, services, virtual machines, private clouds, IP address pools, storage pools, and more
- Performance and resource optimization (PRO), which can now be configured at a granular level and delegated to specific self-service users
- Monitoring of and automated remediation of physical servers, storage, and network devices

Note: For additional in-guest workload and application-specific monitoring, simply deploy an Operations Manager agent within the virtual machine operating system and enable the desired management pack. This is not considered fabric monitoring, but awareness of it is important.

Reporting

The 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.

In a service-oriented IT model, reporting serves the following purposes:

- Systems performance and health
- Capacity metering and planning
- Service-level availability
- Usage-based metering and chargeback
- Incident and problem reports that help IT focus efforts

As a result of VMM and Operations Manager integration, several reports are created and are made available by default. However, metering and chargeback reports and incident and problem reports are enabled by the use of Service Manager and Cloud Services Process Pack.

Service Management System

The goal of Service Manager 2012 is to support IT service management in a broad sense. This includes implementing Information Technology Infrastructure Library (ITIL) processes, such as change management and incident management, and it can also include processes for other things, such as allocating resources from a private cloud.

Service Manager 2012 maintains a configuration management database (CMDB). The CMDB is the repository for nearly all configuration and management-related information in the System Center 2012 environment. With the System Center Cloud Services Process Pack, this information includes VMM 2012 resources, such as virtual machine templates, virtual machine service templates, and so on, which are copied regularly from the VMM 2012 library into the CMDB.

This allows objects such as virtual machines and users to be tied to Orchestrator runbooks for automated request fulfillment, metering, chargeback, and more.

User Self-Service

The Microsoft user self-service solution consists of three elements:

- Service Manager Self-Service Portal
- Cloud Services Process Pack
- App Controller

Service Manager 2012 provides its own self-service portal. Using the information in the CMDB, Service Manager 2012 can create a service catalog that shows the services available to a particular user. For example, a user wants to create a virtual machine in the group's cloud. Instead of passing the request directly on to VMM 2012 as System Center App Controller 2012 does, Service Manager 2012 starts a workflow to handle the request. The workflow contacts the user's manager to get an approval for this request. If the request is approved, the workflow then starts a System Center Orchestrator 2012 runbook.

The Service Manager Self-Service Portal consists of two parts and has the prerequisite of a Service Manager Server and database. The two parts of the Service Manager Self-Service Portal include:

- Web content server
- SharePoint® Web part

These parts should be colocated on a single dedicated server.

The Cloud Services Process Pack is an add-on component that enables IaaS capabilities through the Service Manager Self-Service Portal and Orchestrator runbooks. It provides the following benefits:

- Standardized and well-defined processes for requesting and managing cloud services, including the ability to define projects, capacity pools, and virtual machines
- Natively supported request, approval, and notification to enable businesses to effectively manage their own allocated infrastructure capacity pools

App Controller is the portal a self-service user uses to connect to and manage their virtual machines and services after the request has been fulfilled. App Controller connects directly to VMM using the credentials of the authenticated user to display that user's virtual machines and services and to provide a configurable set of actions. In this case, App Controller is primarily responsible for post-provisioning maintenance tasks.

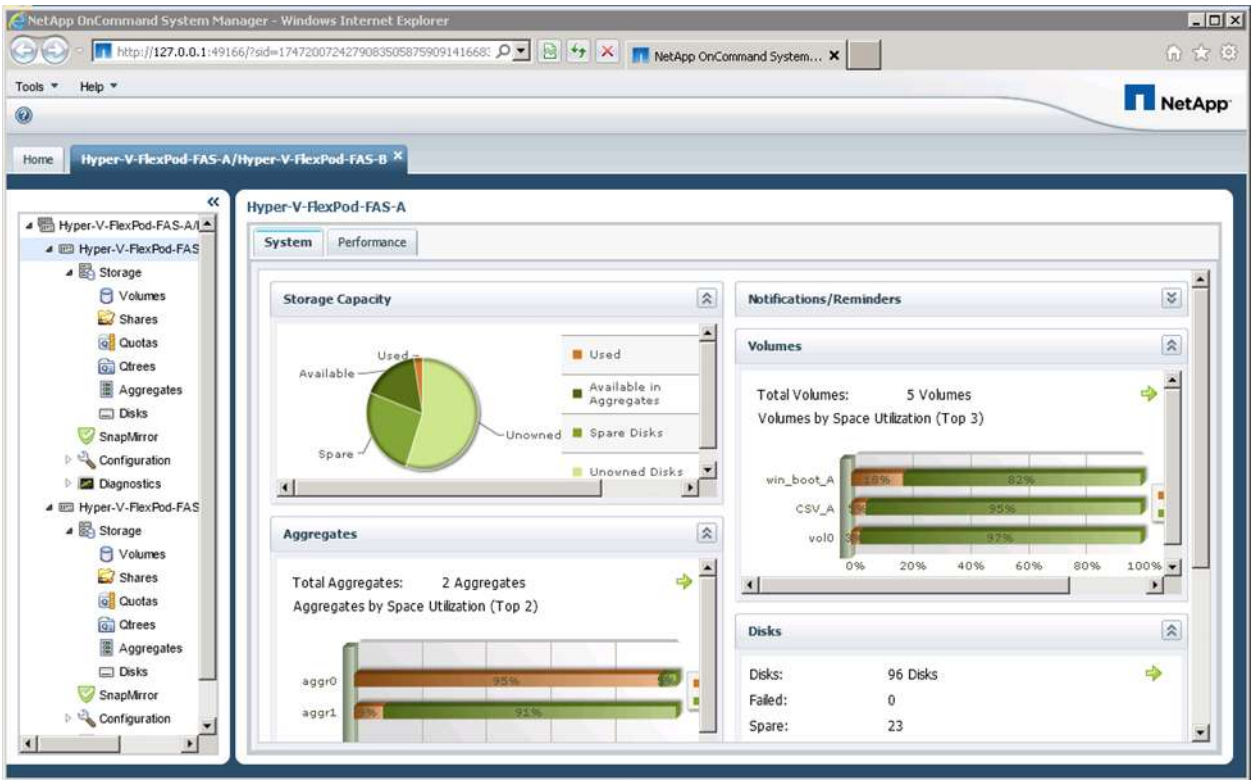
NetApp Storage Management

NetApp OnCommand System Manager (System Manager) is an application that enables you to manage NetApp storage systems and storage objects, such as disks, volumes, and aggregates. System Manager is a Web-based graphical management interface used to perform common storage administration tasks related to NetApp storage systems.

System Manager enables you to perform many common tasks, including:

- Configure and manage storage objects, such as disks, aggregates, volumes, qtrees, and quotas.
- Configure protocols such as CIFS and NFS and provision file sharing.
- Configure protocols such as FC and iSCSI for block access.
- Create and manage vFiler units.
- Set up and manage SnapMirror relationships.
- Manage HA configurations and perform takeover and giveback operations.
- Perform cluster management and storage node management in a cluster environment.

Figure 35) NetApp OnCommand System Manager GUI.



System Manager is installed on a Windows client (using Windows XP or later) that an administrator uses to perform storage system configuration changes and storage provisioning to support the deployment of virtual machines and data storage (LUNs or file shares). System Manager connects to the controllers using credentials that have administrative rights on the NetApp storage systems.

Cisco Network Management

The Cisco Nexus series of switches provides a unified management layer of Ethernet, IP, and FCP combined into a single management platform.

Consistent Management for Cisco Products

The switch platform's network features can be managed using the Cisco CLI. The FC and FCoE features can be managed through the Cisco Fabric Manager suite. Cisco Data Center Network Manager (DCNM) also supports the Cisco Nexus 5500 platform. The capability to manage Ethernet and FCoE features independently with existing Cisco tools preserves existing management models, best practices, and investments in staff training. In addition, simple network management protocol (SNMP) MIBs, XML, and the Cisco CLI are made available to customers for switch management through third-party and custom-developed tools. The switch platform uses Cisco NX-OS for superior operating efficiency, pervasive security, and continuous operation even through software upgrades.

Cisco Data Center Network Manager

The Cisco Nexus 5000 is supported in Cisco DCNM. Cisco DCNM is designed for hardware platforms enabled for Cisco NX-OS, which are in the Cisco Nexus Family of products. Cisco DCNM is a Cisco management solution that increases overall data center infrastructure uptime and reliability to improve business continuity. Focused on the management requirements of the data center network, Cisco DCNM provides a robust framework and a comprehensive feature set that meets the routing, switching, and storage administration needs of current and future data centers. In particular, Cisco DCNM automates the provisioning process, proactively monitors the LAN by detecting performance degradation, secures the network, and streamlines the diagnosis of dysfunctional network elements.

Server Management Utilities

The Cisco UCS platform allows the solution’s compute resources to be managed manually or automatically.

Managing a Cohesive System

Cisco UCS Manager provides unified, embedded management of all software and hardware components of the Cisco Unified Computing System (Cisco UCS) across multiple chassis, Cisco UCS B-Series Blade Servers, Cisco UCS C-Series Rack-Mount Servers, and thousands of virtual machines.

By enabling better automation of processes, Cisco UCS Manager allows data center managers to achieve greater agility and scale in their server operations while reducing complexity and risk. Cisco UCS Manager provides flexible role- and policy-based management using service profiles and templates, and it facilitates processes based on ITIL concepts.

It also provides integration with industry-leading systems management solutions to ease adoption of Cisco UCS within IT departments. Cisco UCS Manager also enables custom development with an extensive XML API that exposes more than 9,000 objects and provides increased system visibility and control.

The Cisco UCS management pack for SCOM graphically depicts Cisco Unified Computing System hardware, service profiles, host operating systems, and virtual machines. The correlation of events with the blades and service profiles they affect simplifies identification of root causes and accelerates problem resolution.

Cisco UCS PowerTool is a flexible and powerful command line toolkit that includes Windows PowerShell cmdlets. By using the flexible and powerful scripting environment provided by Windows PowerShell, customers have an efficient, cost-effective, and easy way to integrate and automate Cisco UCS management with Microsoft products and many third-party products.

Flexible, Role-Based Management

Cisco UCS Manager offers role-based management that helps organizations efficiently use their limited administrator resources. Server, network, and storage administrators maintain responsibility and

accountability for their domain policies within an integrated management environment. Roles and privileges in the system can be easily modified, and new roles can be created quickly.

Administrators can focus on defining the policies needed to provision computing infrastructure and network connectivity. They can also collaborate on strategic architectural issues, because implementation of basic server configurations is now highly accelerated and automated.

Policy-Based Provisioning of Server, Network, and Storage Access Resources

Cisco UCS Manager uses service profiles to provision and manage Cisco UCS blade servers, rack-mount servers, and their I/O properties within a single management domain.

Service profiles are created by server, network, and storage administrators. Infrastructure policies needed to deploy applications are encapsulated in the service profile. The policies coordinate and automate element management at every layer of the hardware stack (Figure 36), including RAID levels, BIOS settings, firmware revisions and settings, adapter identities and settings, VLAN and VSAN network settings, network QoS, and data center connectivity.

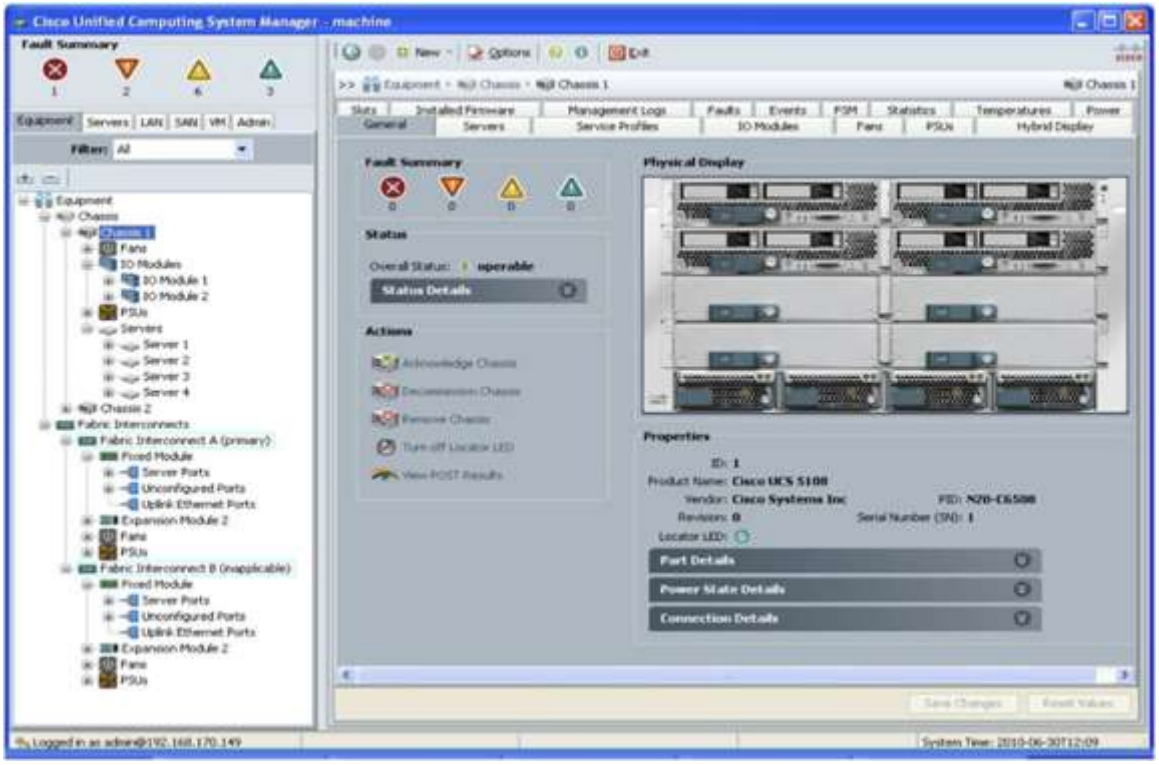
Service profile templates are used to simplify the creation of service profiles, which promotes consistent policies within the system for a given service or application. This approach makes it just as easy to configure one server or hundreds of servers with thousands of virtual machines.

Multiple Interface Options

Cisco UCS Manager has a GUI and a CLI for server, network, and storage administrators to use. Cisco UCS Manager also provides a powerful XML API for integration with existing data center systems' management tools. Some examples of additional management interfaces are Intelligent Platform Management Interface (IPMI); keyboard, video, and mouse (KVM); serial over LAN (SoL); and SNMP. The XML interface allows the entire system to be monitored or configured externally by higher level systems management tools from Cisco's many ecosystem partners.

Figure 36 shows a view of the Cisco UCS 5108 server chassis equipment on the Cisco UCS Manager GUI.

Figure 36) Cisco UCS Manager GUI (graphic provided by Cisco).



Server Out-of-Band Management Configuration

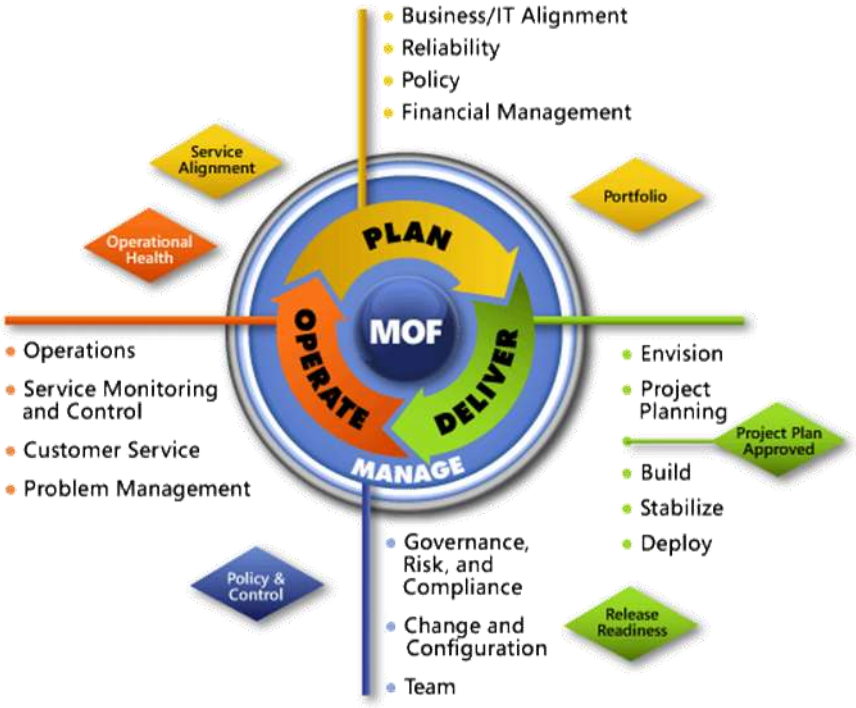
The Cisco UCS 6200 Series hosts and runs Cisco UCS Manager in a highly available configuration that enables the fabric interconnects to fully manage all Cisco UCS elements. Connectivity to the Cisco UCS 5100 Series blade chassis is maintained through the Cisco UCS 2100 or 2200 Series fabric extenders in each blade chassis. The Cisco UCS 6200 Series interconnects support out-of-band management through a dedicated 10/100/1000Mbps Ethernet management port as well as in-band management. The Cisco UCS Manager typically is deployed in a clustered active-passive configuration on redundant fabric interconnects connected through dual 10/100/1000 Ethernet clustering ports.

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 ITIL, to provide built-in processes for incident resolution, problem resolution, and change control.

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 and cost-effective IT services. MOF is a free, downloadable framework that encompasses the entire service management lifecycle. For more information on MOF, refer to Microsoft Operations Framework 4.0.

Figure 37) MOF (graphic provided by Microsoft).



Backup and Disaster Recovery

In a virtualized data center, there are three commonly used backup types: host-based backup, guest-based backup, and SAN-based backup. Table 12 shows the distinctions between the three backup types.

Table 12) Differences in types of backups.

Capability	Host Based	Guest Based	SAN Based
Protection of virtual machine configuration	X		X*
Protection of host and cluster configuration	X		X*
Protection of virtualization-specific data such as virtual machine Snapshots copies	X		X
Protection of data inside the virtual machine	X	X	X
Protection of data inside the virtual machine stored on pass-through disks		X	X
Support for VSS-based backups for supported operating systems and applications	X	X	X*
Ability to granularly recover specific files or applications inside the virtual machine		X	X*

* Supported with SnapManager configured according to NetApp best practice guidelines.

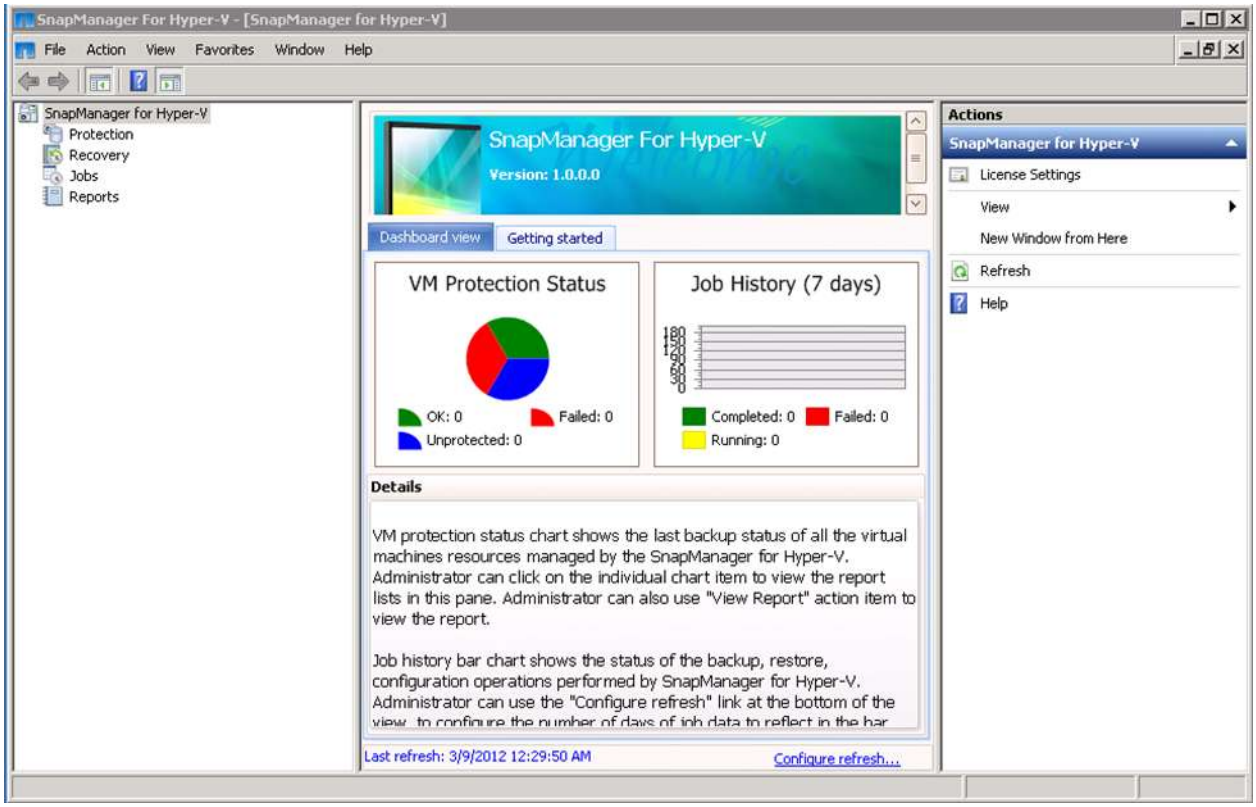
SnapManager for Hyper-V provides a solution for data protection and recovery for Microsoft Hyper-V virtual machines. SnapManager for Hyper-V enables you to perform application-consistent dataset backups according to the protection policies set by your storage administrator. You can also restore

virtual machine backups from those application-consistent backups. You can apply policies to the datasets to automate backup tasks such as scheduling, retention, and replication.

You can perform the following tasks with SnapManager for Hyper-V:

- Group virtual machines into datasets that have the same protection requirements and apply policies to those datasets.
- Back up dedicated and clustered virtual machines on storage systems running Data ONTAP software.
- Backup and restore virtual machines running on clustered shared volumes.
- Automate dataset backups using scheduling policies.
- Perform on-demand backups of datasets.
- Retain dataset backups for as long as you need them using retention policies.
- Update the SnapMirror destination location after a backup finishes successfully.
- Specify custom scripts to run before or after a backup.
- Restore virtual machines from backups.
- Monitor the status of all scheduled and running jobs.
- Manage hosts remotely from a management console.

Figure 38) SnapManager for Hyper-V.



SnapManager for Hyper-V provides integration with Microsoft Hyper-V VSS writer to make an application-consistent Snapshot copy of the virtual machine. SnapManager for Hyper-V is a VSS requestor and coordinates the backup operation to create a consistent Snapshot copy using VSS Hardware Provider for Data ONTAP.

SnapManager for Hyper-V enables you to make application-consistent backups of a virtual machine, if you have Microsoft Exchange, Microsoft SQL Server, or any other VSS-aware application running on VHDs in the virtual machine. SnapManager for Hyper-V coordinates with the application writers inside the virtual machine to verify that application data is consistent when the backup occurs.

You can also restore a virtual machine from an application-consistent backup. The applications that exist in the virtual machine restore to the same state they were in at the time of the backup. SnapManager for Hyper-V restores the virtual machine to its original location.

SnapManager for Hyper-V is installed on each Hyper-V parent host. NetApp SnapDrive for Windows Server is also installed on the Hyper-V parent host. The SnapDrive application installs the NetApp Data ONTAP VSS hardware provider, which is used by SnapManager for Hyper-V to create an application-consistent Snapshot copy of the virtual machine.

The virtual machines must be located on a NetApp LUN. SnapManager for Hyper-V only creates backups of virtual machines located on NetApp LUNs connected to Hyper-V parent hosts where SnapDrive and SnapManager for Hyper-V are installed.

SnapManager for Hyper-V can also be installed on a management host to allow an administrator centralized management of Hyper-V datasets, policies, backup operations, and restore operations. SnapManager for Hyper-V datasets consists of virtual machines located on NetApp LUNs connected to Hyper-V parent hosts.

SnapManager for Hyper-V policies are applied to datasets to specify the configuration of the backup of the virtual machines. SnapManager for Hyper-V policies control backup retention, backup schedule, permissions, and backup options (SnapMirror after backup, allow save state for virtual machine backups) and execute post backup scripts.

SnapManager for Hyper-V creates backup reports after every backup. An administrator can use these reports as well as Windows application events and diagnostic logs created by SnapDrive to troubleshoot SnapManager for Hyper-V backups.

Security

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

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 want 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.

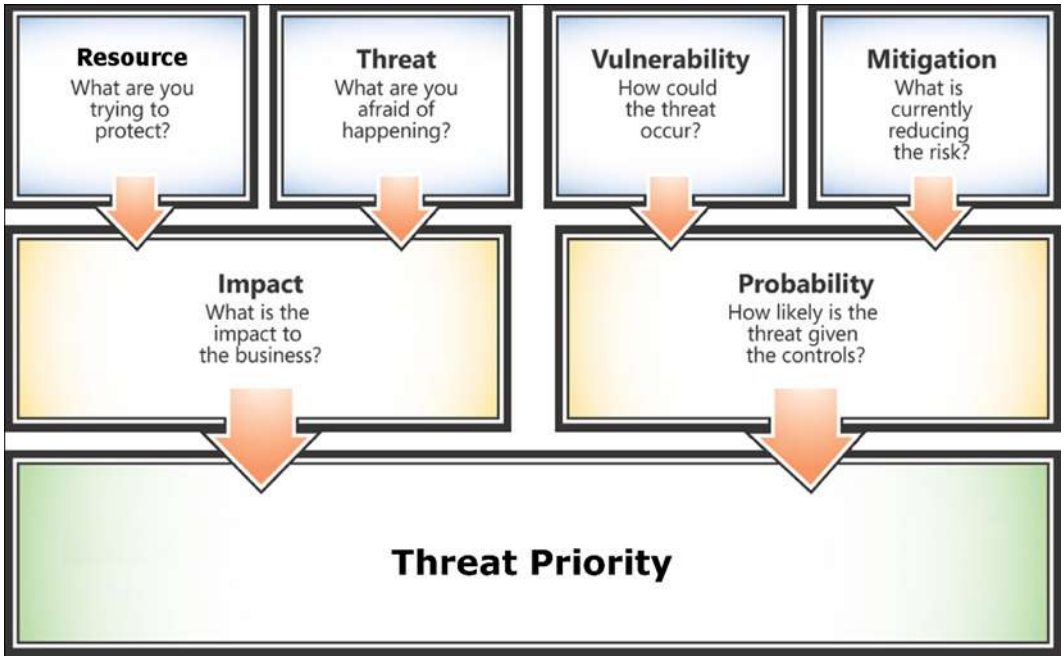
The IT infrastructure threat-modeling process is a systematic analysis of IT components that compiles component information into profiles. The goal of the process is to develop a threat-model portfolio, which is a collection of component profiles.

One way to establish these pillars as a basis for threat modeling IT infrastructure is through 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 planning phase of MOF addresses creating plans for confidentiality, integrity, availability, continuity, and capacity. The policy SMF in the planning phase provides context to help understand the reasons for policies along with their creation, validation, and enforcement and includes processes to communicate policy, incorporate feedback, and help IT maintain compliance with directives. The delivery phase contains several SMFs that help make sure 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 39) Microsoft threat model (graphic provided by Microsoft).



For more information on threat modeling, refer to the IT Infrastructure Threat Modeling Guide.

Security Risk Management

Security for the Microsoft Private Cloud is founded on three pillars: protected infrastructure, application access, and network access. For more information, refer to the Security Risk Management Guide.

Protected Infrastructure

A defense-in-depth strategy is utilized at each layer of the Microsoft Private 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. The effects of malicious software can create additional costs to disinfect computers, restore files, and reenter or recreate lost data. Virus attacks can also cause project teams to miss deadlines, which could lead 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 safest way to counter these threats. The least-privileged user account (LUA) approach is an important part of that defensive strategy. The LUA approach makes users follow the principle of least privilege and log on with limited user accounts. This strategy also limits the use of administrative credentials to administrators, and then only for administrative tasks.

Application Access

Active Directory provides the means to manage the identities and relationships that make up the Microsoft Private Cloud. Integrated with Windows Server 2008 R2, Active Directory provides the 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 to improve developer productivity, enhance application security, and enable interoperability. Windows Identity Foundation also allows developers to enjoy greater productivity by applying the same tools and programming model to build on-premises software as well as cloud services and to create more secure applications by reducing custom implementations and using a single simplified identity model based on claims.

Network Access

Windows Firewall with Advanced Security combines a host firewall and 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. 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 are ignored.

Endpoint Protection (Antivirus and Antimalware)

Microsoft Forefront delivers comprehensive, end-to-end solutions, both on premises and in the cloud, to help protect users and to enable secure access virtually anywhere. You can secure your environment and manage access across data, users, and systems by using our integrated portfolio of protection, identity, and access products.

For more information, refer to www.microsoft.com/en-us/server-cloud/forefront/default.aspx.

System Center Endpoint Protection

Desktop management and security have traditionally existed as two separate disciplines, yet both play central roles in keeping users safe and productive. Management makes sure that systems are properly configured, that patches are deployed against vulnerabilities, and that necessary system updates are

delivered. Security provides critical threat detection, incident response, and remediation of system infection.

System Center 2012 Endpoint Protection makes it easier to protect critical desktop and server operating systems against viruses, spyware, rootkits, and other threats. The key features of System Center 2012 Endpoint Protection include:

- **Single console for endpoint management and security.** Configuration Manager provides a single interface for managing and securing desktops that reduces complexity and improves troubleshooting and reporting insights.
- **Central policy creation.** Administrators have a central location for creating and applying all client-related policies.
- **Enterprise scalability.** Use of the Configuration Manager infrastructure in System Center 2012 Endpoint Protection makes it possible to efficiently deploy clients and policies in the largest organizations around the globe. By using Configuration Manager, distribution points, and an automatic software deployment model, organizations can quickly deploy updates without relying on WSUS.
- **Highly accurate and efficient threat detection.** The antimalware engine in System Center 2012 Endpoint Protection protects against the latest malware and rootkits with a low false-positive rate and keeps employees productive with scanning that has a low impact on performance.
- **Behavioral threat detection.** System Center 2012 Endpoint Protection uses system behavior and file reputation data to identify and block attacks on client systems from previously unknown threats. Detection methods include behavior monitoring, the cloud-based dynamic signature service, and dynamic translation.
- **Vulnerability shielding.** System Center 2012 Endpoint Protection blocks exploitation of endpoint vulnerabilities with deep protocol analysis of network traffic.
- **Automated agent replacement.** System Center 2012 Endpoint Protection automatically detects and removes the most common endpoint security agents, which dramatically lowers the time and effort needed to deploy new protection.
- **Windows Firewall management.** System Center 2012 Endpoint Protection verifies that Windows Firewall is active and working properly to protect against network-layer threats. It also enables administrators to more easily manage these protections across the enterprise.

5.9 Service Delivery

As the primary interface with the business, the service delivery layer is expected to know or obtain answers to the following questions:

- What services does the business want?
- For what level of service are business decision makers willing to pay?
- How can private cloud move IT from being a cost center to becoming a strategic partner to the business?

With these questions in mind, there are two main problems within the service layer that IT must address:

- How do we provide a cloudlike platform for business services that meets business objectives?
- How do we adopt an easily understood usage-based cost model that can be used to influence business decisions?

Figure 40) Components of service delivery (graphic provided by Microsoft).

The components of the service delivery layer are:

- **Financial management.** Financial management incorporates the functions and processes used to meet a service provider's budgeting, accounting, metering, and charging requirements. The primary concerns around financial management in a private cloud are providing cost transparency to the business and structuring a usage-based cost model for the consumer. Achieving these goals is a basic precursor to achieving the principle of encouraging desired consumer behavior.
- **Demand management.** Demand management involves understanding and influencing customer demands for services, in addition to the provisioning of capacity to meet these demands. The principles of perceived infinite capacity and continuous availability are fundamental to stimulating customer demand for cloud-based services. A resilient, predictable environment and predictable capacity management are necessary to adhere to these principles. Cost, quality, and agility factors influence consumer demand for these services.
- **Business relationship management.** Business relationship management is the strategic interface between the business and IT. If an IT department adheres to the principle that it must act as a service provider, mature business relationship management is critical. The business should define the functionality of required services and partner with IT on solution procurement. The business must also work closely with IT to define future capacity requirements to continue adhering to the principle of perceived infinite capacity.
- **Service catalog.** The output of demand and business relationship management is a list of services or service classes offered and documented in the service catalog. This catalog describes each service class, the eligibility requirements for each service class, service-level attributes, targets included with each service class (such as availability targets), and cost models for each service class. The catalog must be managed over time to reflect changing business needs and objectives.
- **Service lifecycle management.** Service lifecycle management takes an end-to-end management view of a service. A typical journey starts with identifying a business need and moves through business relationship management to the time when that service becomes available. Service strategy drives service design. After launch, the service is transitioned to operations and refined through continual service improvement. Taking a service provider's approach is the key to successful service lifecycle management.
- **Service-level management.** Service-level management is the process of negotiating SLAs and making sure the agreements are met. SLAs define target levels for cost, quality, and agility by service class, as well as the metrics for measuring actual performance. Managing SLAs is necessary for achieving the perception of infinite capacity and continuous availability. This, too, requires a service provider's approach by IT.
- **Continuity and availability management.** Availability management defines the processes necessary to achieve the perception of continuous availability. Continuity management defines how risk will be managed in a disaster scenario to make sure minimum service levels are maintained. The principles of resiliency and automation are fundamental here.
- **Capacity management.** Capacity management defines the processes necessary to achieve the perception of infinite capacity. Capacity must be managed to meet existing and future peak demand while controlling underutilization. Business relationship and demand management are key inputs into effective capacity management, and they require a service provider's approach. Predictability and optimization of resource usage are primary principles in achieving capacity management objectives.

- **Information security management.** Information security management strives to make sure that all requirements are met for the confidentiality, integrity, and availability of the organization's assets, information, data, and services. An organization's particular information security policies drive the architecture, design, and operations of a private cloud. Resource segmentation and multi-tenancy requirements are important factors to consider during this process.

5.10 Operations

The operations layer defines the operational processes and procedures necessary to deliver IT as a service (ITaaS). This layer uses IT service management concepts that can be found in prevailing best practice such as ITIL or MOF.

The main focus of the operations layer is to execute the business requirements defined at the service delivery layer. Cloudlike service attributes cannot be achieved through technology alone; mature IT service management is also required.

The operations capabilities are common to all three services: IaaS, PaaS, and SaaS.

Figure 41) Components of the operation layer (graphic provided by Microsoft).

The components of the operations layer include:

- **Change management.** Change management is responsible for controlling the lifecycle of all changes. Its primary objective is to implement beneficial changes with minimum disruption to the perception of continuous availability. Change management determines the cost and risk of making changes and balances them against the benefits to the business or service. Driving predictability and minimizing human involvement are the core principles behind a mature change management process.
- **Service asset and configuration management.** Service asset and configuration management maintains information on the assets, components, and infrastructure needed to provide a service. Accurate configuration data for each component and its relationship to other components must be captured and maintained. This data should include historical and expected future states, in addition to the current state, and be easily available to those who need it. Mature service asset and configuration management processes are necessary for achieving predictability.
- **Release and deployment management.** Release and deployment management is responsible for seeing that changes to a service are built, tested, and deployed with minimal disruption to the service or production environment. Change management provides the approval mechanism (determining what will be changed and why), but release and deployment management is the mechanism for determining how changes are implemented. Driving predictability and minimizing human involvement in the release and deployment process are the keys to achieving cost, quality, and agility goals.
- **Knowledge management.** Knowledge management is responsible for gathering, analyzing, storing, and sharing information within an organization. Mature knowledge management processes are necessary to achieve a service provider's approach and a key element of IT service management.
- **Incident and problem management.** The goal of incident and problem management is to resolve disruptive or potentially disruptive events with maximum speed and minimum disruption. Problem management also identifies root causes of past incidents and seeks to identify and prevent (or minimize the effect of) future ones. In a private cloud, the resiliency of the infrastructure helps make sure that when faults occur, they have minimal effect on service availability. Resilient design

promotes rapid restoration of service continuity. Driving predictability and minimizing human involvement are necessary to achieve this resiliency.

- **Request fulfillment.** The goal of request fulfillment is to manage user requests for services. As IT adopts a service provider's approach, it should define available services in a service catalog based on business functionality. The catalog should encourage desired user behavior by exposing cost, quality, and agility factors to the user. Self-service portals, when appropriate, can assist the drive toward minimal human involvement.
- **Access management.** The goal of access management is to deny access to unauthorized users while making sure that authorized users have access to needed services. Access management implements security policies defined by information security management at the service delivery layer. Maintaining smooth access for authorized users is critical to achieving the perception of continuous availability. Adopting a service provider's approach to access management also makes sure that resource segmentation and multi-tenancy are addressed.
- **Systems administration.** The goal of systems administration is to perform the daily, weekly, monthly, and as-needed tasks required for system health. A mature approach to systems administration is required for achieving a service provider's approach and for driving predictability. The vast majority of systems administration tasks should be automated.

6 Conclusion

FlexPod with Microsoft Private Cloud provides a highly scalable and reliable platform for a wide variety of workloads. The goal of this program is to help you deploy a private cloud environment in your enterprise without the expense or risk associated with designing your own custom solution.

References

This document cites the following references:

- NIST, Special Publication 800-145, "The NIST Definition of Cloud Computing," September 2011; available from http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf
- Windows TechNet Library

Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 527-0883

NetApp provides no representations or warranties regarding the accuracy, reliability, or serviceability of any information or recommendations provided in this publication, or with respect to any results that may be obtained by the use of the information or observance of any recommendations provided herein. The information in this document is distributed AS IS, and the use of this information or the implementation of any recommendations or techniques herein is a customer's responsibility and depends on the customer's ability to evaluate and integrate them into the customer's operational environment. This document and the information contained herein may be used solely in connection with the NetApp products discussed in this document.

© 2012 NetApp, Inc. All rights reserved. No portions of this document may be reproduced without prior written consent of NetApp, Inc. Specifications are subject to change without notice. NetApp, the NetApp logo, Go further, faster, Data ONTAP, FlexClone, FlexPod, FlexVol, MultiStore, OnCommand, RAID-DP, SnapDrive, SnapManager, SnapMirror, Snapshot, SnapVault, vFiler, and WAFL are trademarks or registered trademarks of NetApp, Inc. in the United States and/or other countries. Cisco and Cisco Nexus are registered trademarks and Cisco UCS and Cisco Unified Computing System are trademarks of Cisco Systems, Inc. ESX and VMware are registered trademarks of VMware, Inc. Intel is a registered trademark of Intel Corporation. Linux is a registered trademark of Linus Torvalds. Active Directory, Microsoft, SharePoint, SQL Server, Windows, Windows Server, and Windows Vista are registered trademarks and Hyper-V and Windows PowerShell are trademarks of Microsoft Corporation. UNIX is a registered trademark of The Open Group. All other brands or products are trademarks or registered trademarks of their respective holders and should be treated as such. TR-4058-0412