Microsoft Windows Azure Pack Reference Architecture for Dell XC Series

Contents

Tables

Figures

1. Executive Summary

This document makes recommendations for the design, optimization and scaling of Windows Azure Pack and Microsoft® Hyper-V® on Dell™ XC Series appliances powered by Nutanix software. It shows the scalability of the Dell XC Platform and provides detailed integration and configuration information on the scale-out capabilities of the cluster when leveraged for Azure Pack and Hyper-V deployments. This document is based upon generalized assumptions and Windows Azure design best practices.

Extensive functional testing has been performed using PowerShell and Orchestrator workflows to simulate real-world workloads and conditions of an Azure Pack environment in a Dell- Nutanix environment. The sizing data and recommendations made in this document are based upon multiple testing iterations and thorough technical validation.

This solution is a certified Microsoft Private Cloud Fast Track v4 Reference Architecture.

2. Audience and Purpose

This reference architecture document is intended for architecting, designing, managing, and/or supporting Dell XC Series infrastructures. Consumers of this document should be familiar with Microsoft Hyper-V, Windows Azure Pack, Dell PowerEdge servers and Nutanix software.

We have broken down this document to address key items for each role focusing on the enablement of a successful design, implementation, and transition to operation.

This document will cover the following subject areas:

- Overview of the Dell XC Series solution
- Overview of Windows Azure Pack and its use cases
- The benefits of Windows Azure Pack on Dell XC Series appliances
- Architecting a complete Windows Azure Pack solution on the Dell platform
- Design and configuration considerations when architecting a Windows Azure Pack solution on Dell XC Series appliances
- Functional validation of Windows Azure Pack performance on Dell XC Series appliances

3. Solution Overview

What is the Dell XC Series Architecture?

The Dell XC Platform is a scale-out cluster of high-performance nodes, or servers, each running a standard hypervisor and containing processors, memory and local storage (consisting of SSD Flash and high capacity SATA disk drives). Each node runs virtual machines just like a standard virtual machine host.

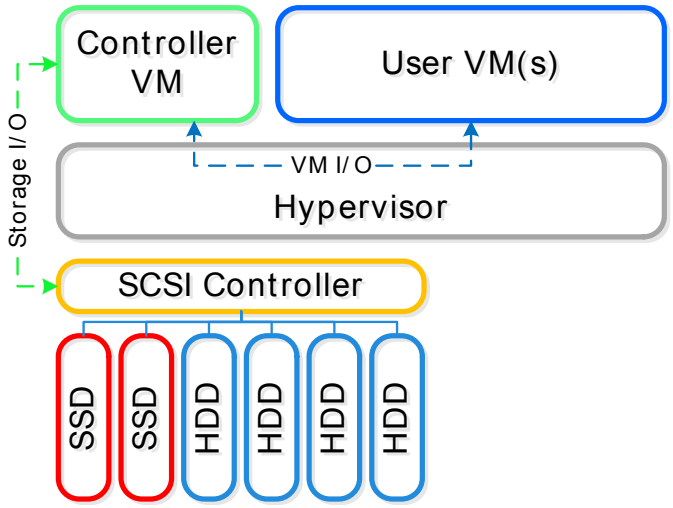


Figure 1. Nutanix Node Architecture

In addition, local storage from all nodes is virtualized into a unified pool by the Nutanix Distributed File System (NDFS). In effect, NDFS acts like an advanced NAS that uses local SSDs and disks from all nodes to store virtual machine data. Virtual machines running on the cluster write data to NDFS as if they were writing to shared storage.

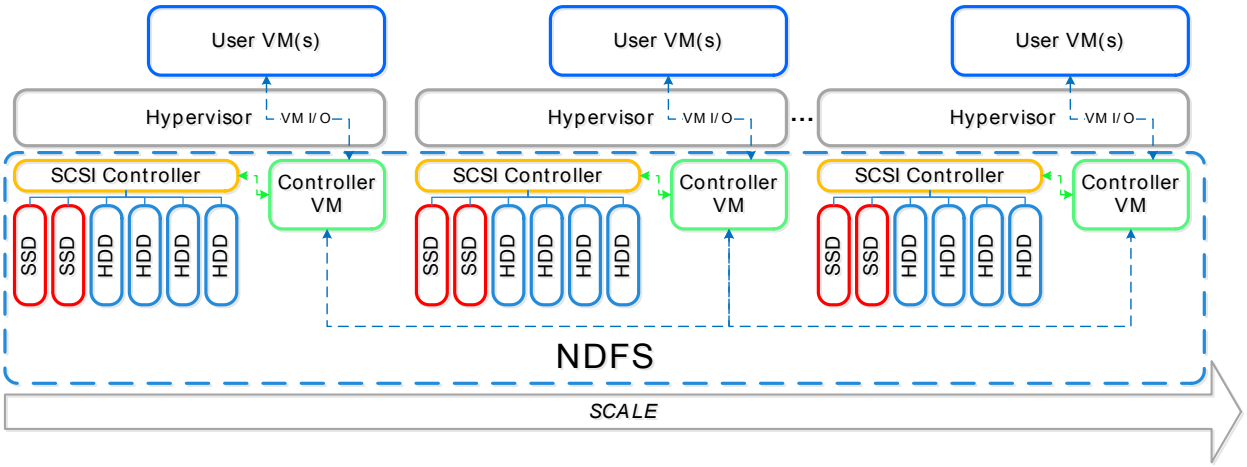


Figure 2. Dell XC Series Architecture

NDFS is VM aware and provides advanced data management features. It brings data closer to virtual machines by storing the data locally on the system, resulting in higher performance at a lower cost.

The Dell XC Platform can horizontally scale from as few as three nodes to a large number of nodes, enabling organizations to scale their infrastructure as their needs grow.

The Nutanix Elastic Deduplication Engine is a software-driven, massively scalable and intelligent data reduction technology. It increases the effective capacity in the disk tier, as well as the RAM and flash cache tiers of the system, by eliminating duplicate data. This substantially increases storage efficiency, while also improving performance due to larger effective cache capacity in RAM and flash. Deduplication is performed by each node individually in the cluster allowing for efficient and uniform deduplication at scale. This technology is increasingly effective with full/persistent clones or P2V migrations.

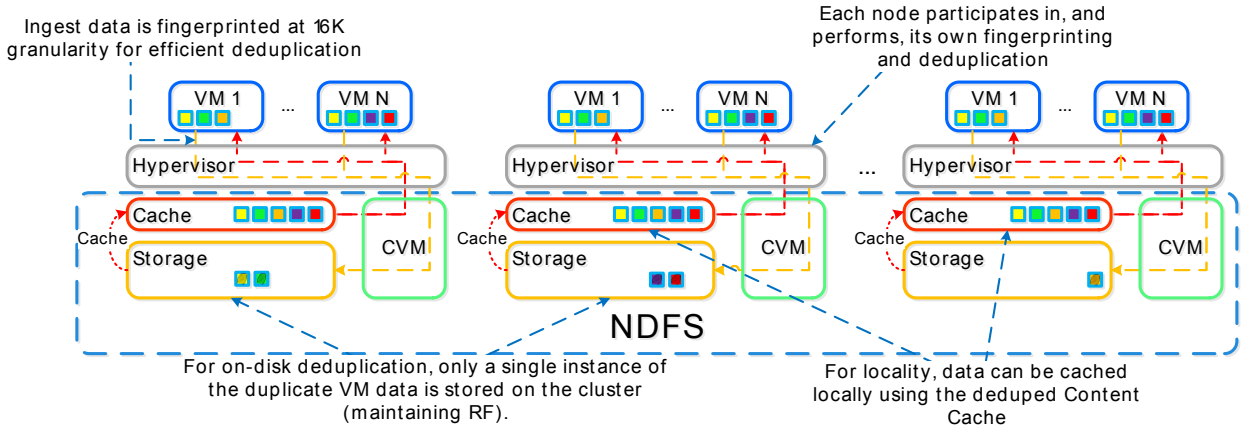


Figure 3. Elastic Deduplication Engine

The NDFS Shadow Clone feature allows for distributed caching of vDisks or VM data which is in a 'multi-reader' scenario. This will allow VMs on each node to read the Base VM's vDisk locally instead of forwarding read requests to a master 'Base VM'. In the case of VDI, this means the base disk can be cached by each node and all read requests for the base will be served locally. In the case where the Base VM is modified the Shadow Clones will be dropped and the process will start over.

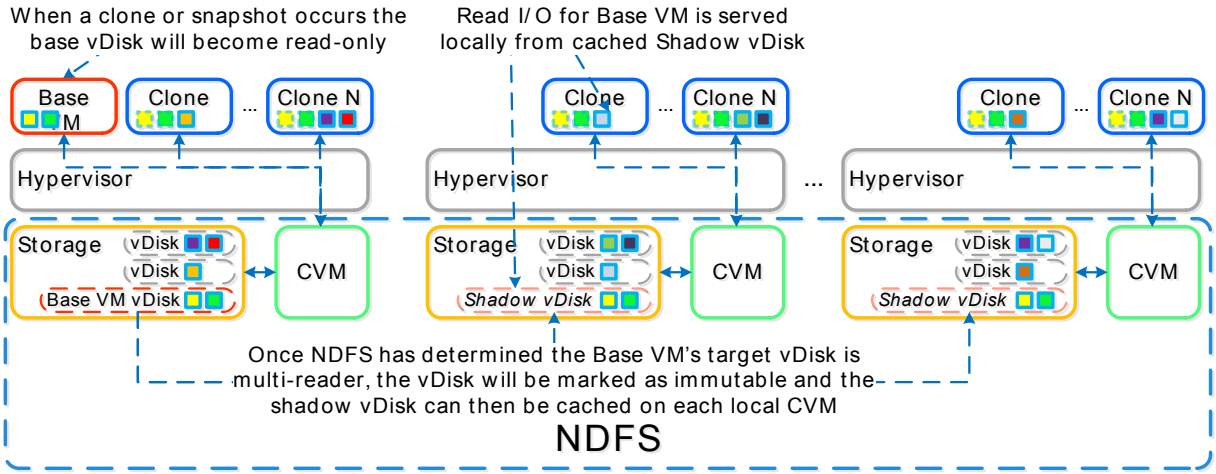
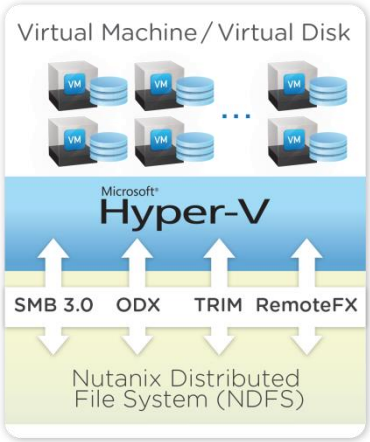


Figure 4. NDFS Shadow Clones

Hyper-V + Dell XC Series

Microsoft server virtualization delivers tremendous IT efficiencies – turning datacenter infrastructure into a flexible and scalable asset. The Dell XC platform is ideal for Hyper-V based virtualization and private cloud deployments, providing high-performance infrastructure for running any Microsoft enterprise application, such as Exchange, SharePoint and SQL Server or test/development environment.

As a converged infrastructure solution, the Dell XC Series simplifies the datacenter and eliminates the complexity of legacy storage architectures.



- The Server Core Installation option of Windows Server 2012 R2 (or Hyper-V Server 2012 R2) allows IT teams to rapidly provision virtual machines
- Support for Microsoft System Center Virtual Machine Manager (SCVMM) for seamless management
- Support for key Hyper-V virtualization capabilities, including Live Migration, Dynamic Optimization, Windows Offloaded Data Transfers (ODX), Failover Clustering and TRIM
- Native SMB 3.0 protocol support
- Provides enterprise-class storage features, including Thin Provisioning, Dynamic Disks, Snapshots, Fast Clones, Compression and De-duplication

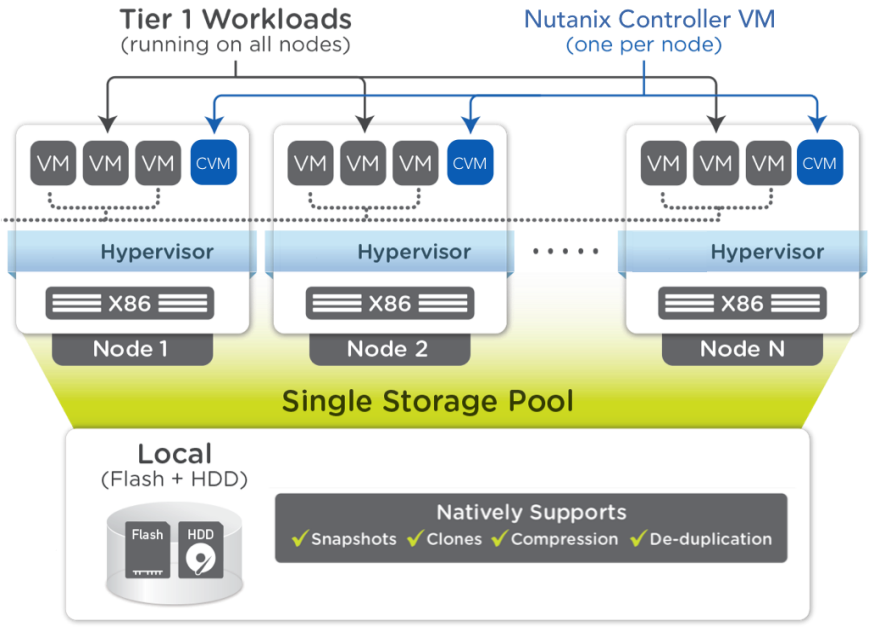


Figure 5. Microsoft Private Cloud Fast Track v4 with Windows Server 2012 R2, System Center 2012 R2, and Dell XC Platform

The Microsoft Private Cloud Fast Track is a joint effort between Microsoft and Nutanix to deliver a preconfigured virtualization and private cloud solution. The program focuses on new technologies and services in Microsoft Windows Server in addition to investments in Microsoft System Center.

The Dell - Nutanix Business-Ready Configuration is pre-engineered, tested, and optimized for virtualization based on the Microsoft Private Cloud Fast Track v4 program. It supports the operating system, virtualization, and management capabilities offered by Windows Server 2012 R2, Hyper-V, and System Center 2012 R2. Organizations using these configurations can benefit from both the control and the flexibility that are required to reap the potential benefits of the private cloud.

For more information on the BOM, please refer to the appendix.

What is Windows Azure Pack?[1]

Windows Azure Pack for Windows Server is a collection of Windows Azure technologies, available to Microsoft customers at no additional cost for installation into your data center. It runs on top of Windows Server 2012 R2 and System Center 2012 R2 and, through the use of the Windows Azure technologies, enables you to offer a rich, self-service, multi-tenant cloud, consistent with the public Windows Azure experience.

Windows Azure Pack provides a complete on-premise cloud delivery system by integrating several distributed components with advanced configuration tools that simplify the creation and real-time management of the VMs and cloud services.

The core components of Azure Pack are:

- Management Portal for Tenants: A Windows Azure-consistent, customizable self-service portal experience for provisioning, monitoring and management of services such as Web Sites, Virtual Machines, and Service Bus.
- Management Portal for Administrators: A portal for administrators to configure and manage resource clouds, user accounts, tenant offers, quotas, and pricing.
- Service Management API: The foundation for the capabilities in the management portal, the service management API is an OData REST API that helps enable a range of integration scenarios including custom portals and billing systems.
- Web Site Clouds: Consistent with Windows Azure Web Sites, this service helps provide a high-density, scalable shared web hosting platform for ASP.NET, PHP, and Node.js web applications. It includes a customizable web application gallery of popular open source web applications and integration with source control systems for custom-developed websites and applications.
- Virtual Machine Clouds: Consistent with Windows Azure Virtual Machines, this service helps provide Infrastructure-as-a-Service (IaaS) capabilities for Windows and Linux virtual machines (VMs). It includes a VM template gallery, scaling options, and virtual networking capabilities.
- SQL and MySQL: Services that provide on-demand database instances. These databases can be used in conjunction with the Web Sites service.
- Service Bus: Consistent with Windows Azure Service Bus, this service helps provide reliable messaging services between distributed applications. It includes queued and topic-based publish/subscribe capabilities.

[1] http://www.microsoft.com/en-us/server-cloud/products/windows-azure-pack/default.aspx?nv1if0=1#fbid=tcHikQKNgLd?hashlink=s1section5

- Automation and Extensibility: The Windows Azure Pack also includes capabilities for automating and integrating additional custom services into the services framework, including a runbook editor and execution environment.

The sections below provide a brief overview of the various Windows Azure Pack deployment architectures. For the sake of this document we will focus on the distributed deployment.

Deployment Scenario – Express Deployment

You can install the **Windows Azure Pack: Portal and API Express** from the Web Platform Installer to create a proof of concept deployment. For instructions on how to do this, see *Install an express deployment of Windows Azure Pack*. In an Express deployment, all of the Windows Azure Pack required components are installed on the same machine. You should then install at least one optional component on additional machines. The express deployment should not be used in a production environment.

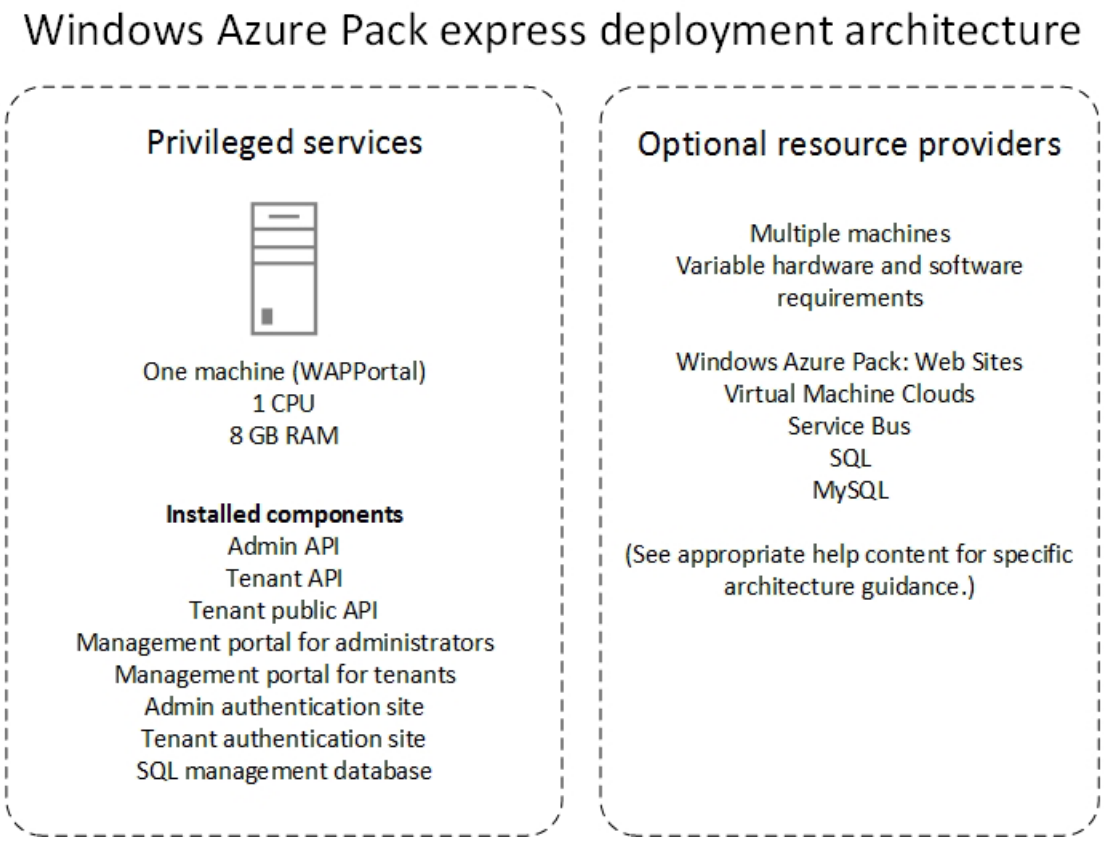


Figure 6. Azure Pack Express Architecture

Deployment Scenario – Basic Distributed Deployment

In a distributed deployment, you can install the required components on up to 8 machines. For instructions on how to install a distributed deployment, see *Install a distributed deployment of Windows Azure Pack*. You should then install at least one optional component on additional machines. A distributed deployment can be used in a production environment. The following diagram shows a basic distributed architecture of required components for a system designed to provide services to external customers.

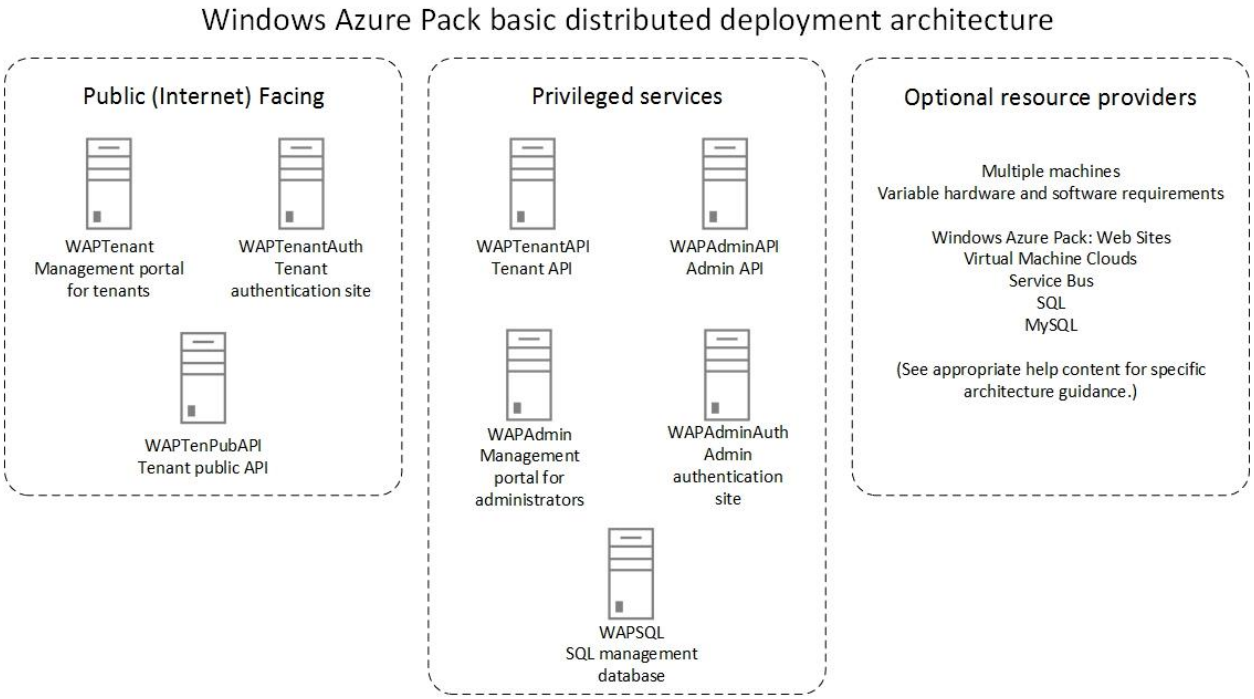


Figure 7. Azure Pack Basic Distributed Architecture

Deployment Scenario – Minimal Distributed Deployment

The following diagram depicts the suggested minimal architecture for a distributed deployment. For instructions on how to install a distributed deployment, see *Install a distributed deployment of Windows Azure Pack*. You should then install at least one optional component on additional machines.

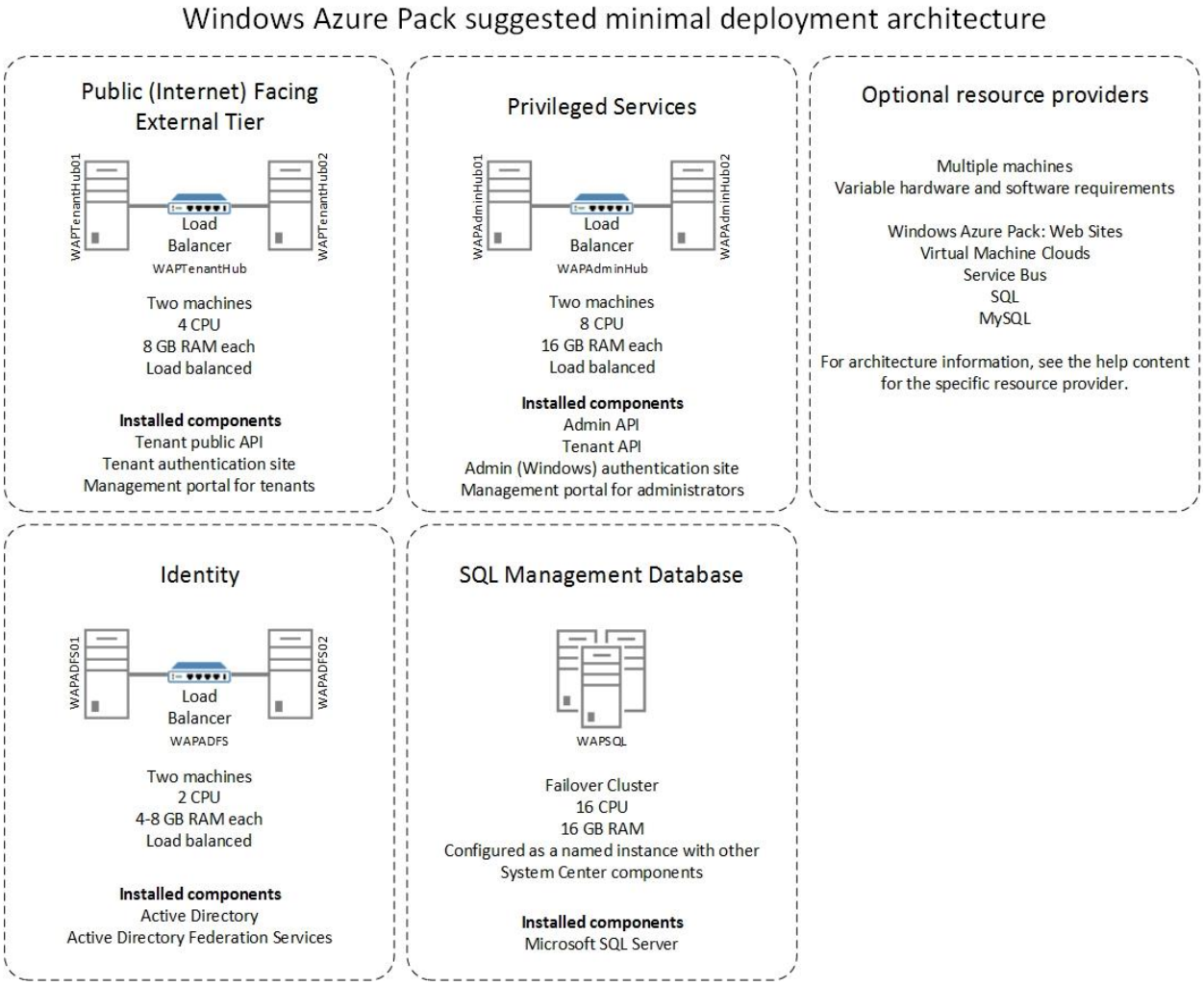


Figure 8. Azure Pack Min Distributed Architecture

Deployment Scenario – Scaled Distributed Deployment

The following diagram shows a distributed deployment using load balancers and failover clusters with scaling information. For instructions on how to install a distributed deployment, see *Install a distributed deployment of Windows Azure Pack*. You should then install at least one optional component on additional machines.

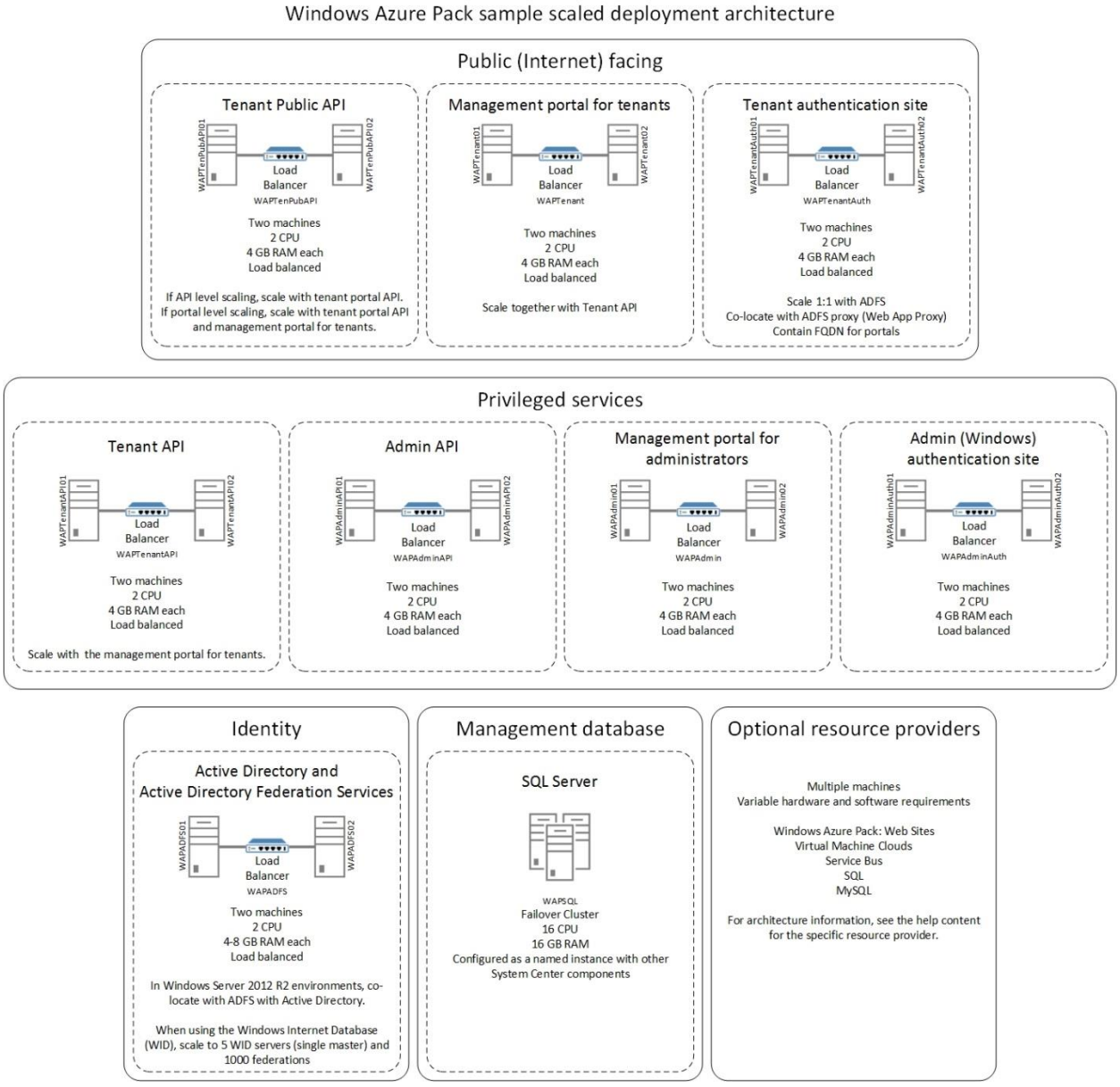
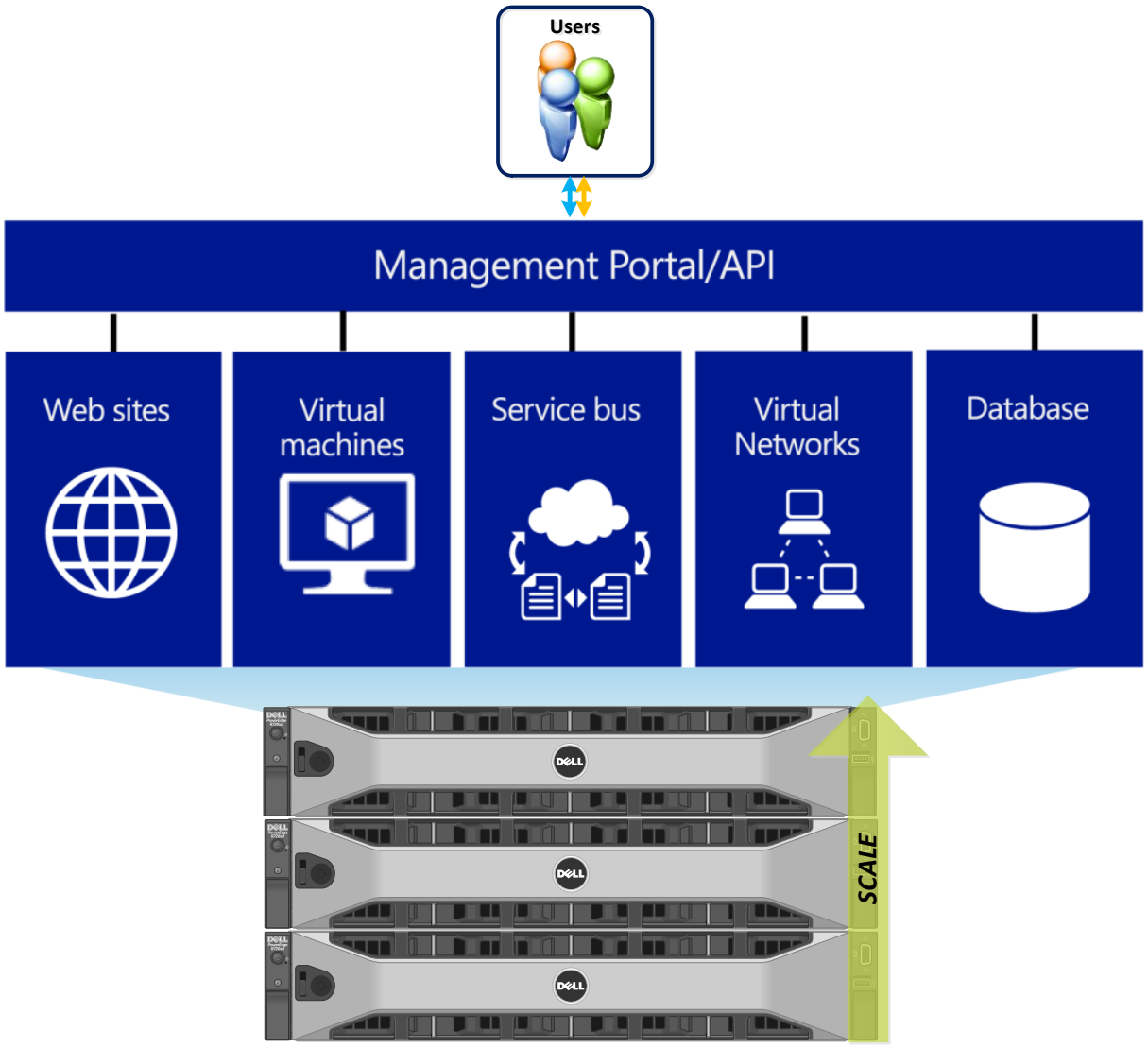


Figure 9. Azure Pack Scaled Distributed Architecture

Resource Providers

Windows Azure Pack allows customers to provision and manage a variety of services and resources including Virtual Machines, Web Sites, SQL Databases, etc. Dell XC Series appliances provide a simple, single platform to easily scale and handle any of the hosting requires for these services.



Virtual Machine Cloud

There are two parts to the VM Clouds in Windows Azure Pack – management portal for administrators and management portal for tenants. The management portal for administrators enables hosting or enterprise service providers to set up the infrastructure against which virtual machines can be provisioned. End users use the management portal for tenants to sign up for plans that include the VM Clouds service, enabling them to provision virtual machines. In this section we look at how the management portal for administrators is associated with the underlying System Center 2012 R2 and Service Provider Foundation to enable the VM Clouds in Windows Azure Pack.

Web Sites

Windows Azure Pack: Web Sites enables an on-premises, high-density, multi-tenant web hosting service for service providers and enterprise IT. Windows Azure Pack: Web Sites provides an experience similar to Windows Azure Web Sites. It is a scalable, shared, and secure web hosting platform that supports both template web applications and a broad range of programming languages like ASP.NET, PHP and Node.js. In addition to a web sites service, it includes a self-service management portal, uses

both SQL and MySQL database servers, integrates with popular source control systems, and offers a customizable web application gallery of popular open source web applications.

SQL Database

Windows Azure Pack (WAP) includes a *SQL Server resource provider* that allows tenants to deploy and manage databases on a shared SQL Server fabric.

Out of the box, and in the spirit of a Platform as a Service approach, tenants do not control on which backend server the database is created. They are just guaranteed that only the user they choose is created and has access to this database, providing security isolation. They eventually know the target server by looking at the connection string provided back, but WAP handles the placement and quota allocation as needed, based on the plans the tenant has signed up for.

Windows Azure Pack the Dell XC Way

Dell XC Series appliances operate and scales on-premise hybrid clouds using the Windows Azure Pack. The figure below shows the high-level Azure Pack on Nutanix solution:

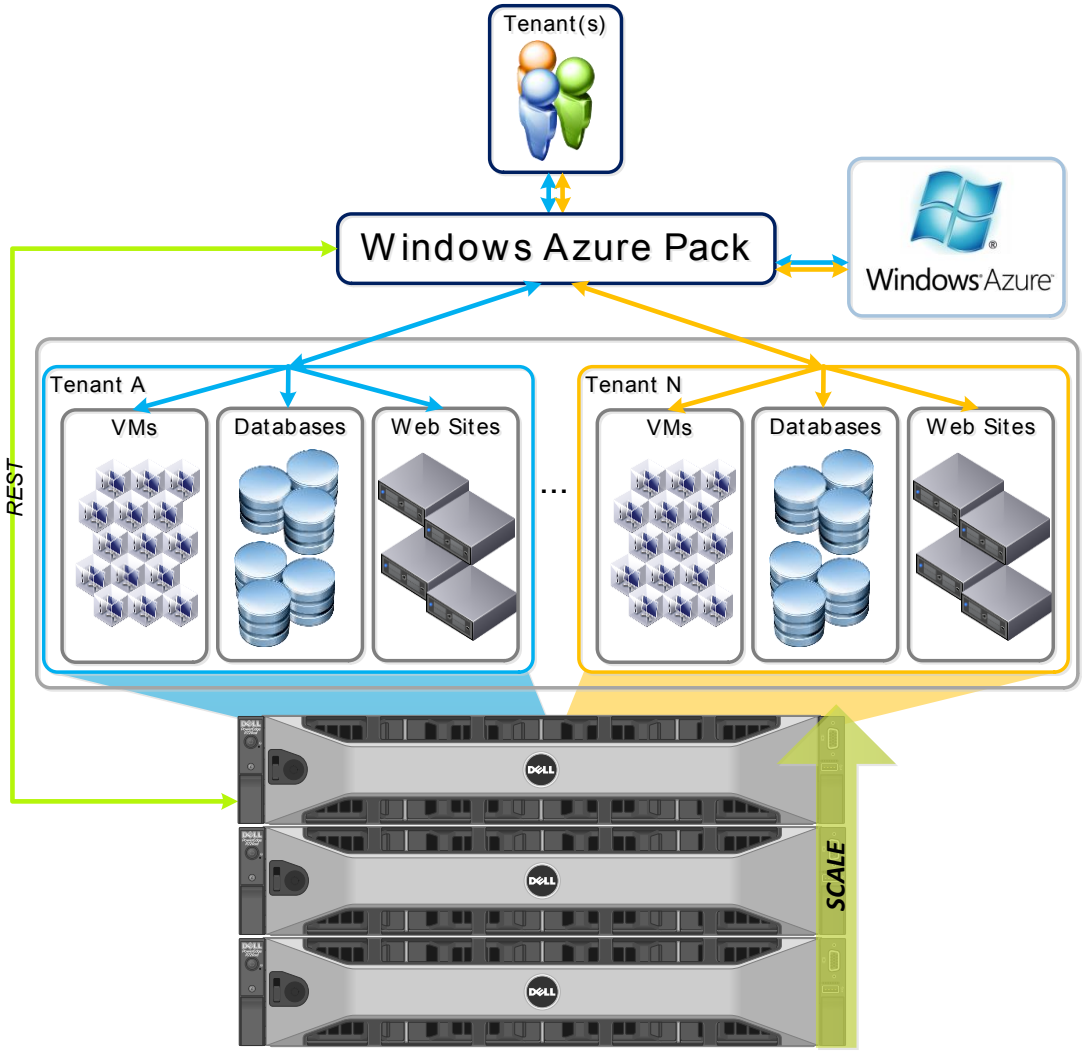


Figure 10. Azure Pack on Nutanix Conceptual Arch

The Dell XC approach of modular scale-out enables customers to select any initial deployment size and grow in more granular data and desktop increments. This removes the hurdle of a large up-front infrastructure purchase that a customer will need many months or years to grow into, ensuring a faster time-to-value for the a hybrid-cloud Azure implementation.

The Dell XC solution is fully integrated with the Microsoft storage APIs (ODX/SMI-s) and provides high performance SSD flash to enable you to provide the best possible experience to the end user with the flexibility of a single modular platform.

Why run Windows Azure on Dell XC Series appliances?

XC Series appliances enable you to run multiple workloads all on the same scalable converged infrastructure:

- Modular incremental scale: With the Dell XC solution you can start small and scale. A single Del XC block provides up to 20TB storage and up to 20 cores in a compact 2U footprint. Given the modularity of the solution, you can granularly scale per-node (up to ~10TB/20 cores or with multiple nodes giving you the ability to accurately match supply with demand and minimize the upfront CapEx.
- Integrated: The Dell XC Series appliances provide full support for ODX and SMI-s allowing you to leverage all the latest advancements from Microsoft and taking your solution to the next level.
- High performance: Up to 150,000 plus random read IOPS and up to 4 GB/s of sequential throughput in a compact 4-node cluster which scales linearly as new nodes are added.
- Elastic Deduplication: Nutanix Map Reduce Deduplication provides granular deduplication of data to increase cache efficiency. The engine will utilize the unique fingerprints of data and only bring one copy up into the Nutanix Content Cache as well as store only one copy on disk. This allows for the highest possible cache utilization and higher performance for VM's accessing common data as well as larger usable storage capacity.
- Data efficiency: The Dell XC solution is truly VM-centric for all compression policies. Unlike traditional solutions that perform compression mainly at the LUN level, the Dell XC solution provides all of these capabilities at the VM and file level, greatly increasing efficiency and simplicity. These capabilities ensure the highest possible compression/decompression performance on a sub-block level. TRIM support also provides data reclamation capabilities for deleted data.
- Business continuity and data protection: Cloud services and VMs are mission critical and need enterprise-grade data management features including backup and DR. With the Dell XC solution these are provided out-of-the-box and can be managed the same as they would be for virtual environments.
- Enterprise-grade cluster management: A simplified and intuitive Apple-like approach to managing large clusters, including a converged GUI that serves as a single pane of glass for servers and storage, alert notifications, and bonjour mechanism to auto-detect new nodes in the cluster. Spend more time enhancing your environment, not maintaining it.

4. Solution Design

Design Decisions

With the Windows Azure Pack on Dell XC solution, you have the flexibility to start small with a minimum of three nodes and scale up incrementally a node, a block, or multiple blocks at a time. This provides the best of both worlds–the ability to start small and grow to massive scale without any impact to performance.

In the following section we cover the design decisions and rationale for the Azure Pack deployments on the Dell XC Platform.

Table 1. Platform Design Decisions

Item	Detail	Rationale
General		
Minimum Size	3 x Dell XC nodes (3 Hyper-V hosts)	Minimum size requirement
Scale Approach	Incremental modular scale	Allow for growth from PoC to massive scale
Scale Unit	Node(s) or Pod(s)	Granularly scale to precisely meet capacity demands Scale in node increments
Infrastructure Services	Small deployments: Shared cluster Large deployments: Dedicated cluster	Dedicated infrastructure cluster for larger deployments (best practice)
Microsoft Hyper-V		
Cluster Size	Up to 12-24 vSphere hosts (Minimum of 4 hosts)	Isolated fault domains
Clusters per SCVMM	Up to 2x24 or 4x12 host clusters	Task parallelization Under 8,000 VM limit
Datastore(s)	1 x Nutanix DFS datastore per pod (Azure Pack Server VMs, VM clones, etc.)	Nutanix handles I/O distribution/localization n-Controller model
Dell XC Series		
Cluster Size	Up to 24-48-nodes	Isolated fault domains
Storage Pool(s)	1 x Storage Pool per cluster	Standard practice ILM handles tiering
Container(s)	1 x Container for VMs 1 x Container for SCVMM Library 1 x Container per Tenant (if required)	Standard practice
Features/ Enhancements	Enable Shadow Clones Increase CVM Memory to 32GB+	Best practice

The table below shows the external and/or public facing components of the Windows Azure Pack. For internal deployments these resources can be on an "internal" network if only intended to be utilized internally, if intended to be accessible externally they should be placed on a DMZ and only specific ports should be forwarded for security purposes.

NOTE: For small and/or test environments these roles can be co-located within VMs. However for larger deployments it is recommended that each role is deployed individually

Table 2. Azure Pack Public Design Decisions

Item	Detail	Rationale
Tenant Public API		
VM Quantity	Min: 2 (n+1) Scale: scale with tenant portal API	HA for Tenant Public API
Server Sizing	vCPU: 2 Memory: 4GB Disk: 40GB	Standard sizing practice
Users per VM	Unknown	Not currently known
Network Connectivity	DMZ (if public, else customer facing)	Connectivity for tenants (external/internal)
Load Balancing	Yes – Load Balancer	Ensures availability of Public API VMs Balances load between API VMs & pods
Management Portal for Tenants		
VM Quantity	Min: 2 (n+1) Scale: scale with tenant API	HA for Tenant Management Portal
Server Sizing	vCPU: 2 Memory: 4GB Disk: 40GB	Standard sizing practice
Users per VM	Unknown	Not currently known
Network Connectivity	DMZ (if public, else customer facing)	Connectivity for tenants (external/internal)
Load Balancing	Yes – Load Balancer	Ensures availability of Tenant Portal VMs Balances load between Tenant Portal VMs & pods
Tenant Authentication Site		
VM Quantity	Min: 2 (n+1) Scale: scale with ADFS	HA for Tenant Authentication VMs
Server Sizing	vCPU: 2 Memory: 4GB Disk: 40GB	Standard sizing practice
Users per VM	Unknown	Not currently known
Network Connectivity	DMZ (if public, else customer facing)	Connectivity for tenants (external/internal)
Load Balancing	Yes – Load Balancer	Ensures availability of Tenant Authentication VMs Balances load between Tenant Authentication VMs & pods

The table below shows the privileged components of the Windows Azure Pack. These resources should be on a infrastructure or non-tenant network with specific ports enabled between them and the public facing components.

NOTE: For small and/or test environments these roles can be co-located within VMs. However for larger deployments it is recommended that each role is deployed individually.

Table 3. Azure Pack Privileged Design Decisions

Item	Detail	Rationale
Tenant API		
VM Quantity	Min: 2 (n+1) Scale: scale with Management Portal for Tenants	HA for Tenant API VMs
Server Sizing	vCPU: 2 Memory: 4GB Disk: 40GB	Standard sizing practice
Users per VM	Unknown	Not currently known
Network Connectivity	DMZ (if public, else customer facing) Internal (privileged network)	Connectivity for tenants (external/internal) Connectivity to back-end resources
Load Balancing	Yes – Load Balancer	Ensures availability of Tenant API VMs Balances load between Tenant API VMs & pods
Admin API		
VM Quantity	Min: 2 (n+1) Scale: 1 per additional pod	HA for Admin API VMs
Server Sizing	vCPU: 2 Memory: 4GB Disk: 40GB	Standard sizing practice
Users per VM	Unknown	Not currently known
Network Connectivity	DMZ (if public, else customer facing) Internal (privileged network)	Connectivity for tenants (external/internal) Connectivity to back-end resources
Load Balancing	Yes – Load Balancer	Ensures availability of Admin API VMs Balances load between Admin API VMs & pods
Management Portal for Admins		
VM Quantity	Min: 2 (n+1) Scale: scale with ADFS	HA for Admin Portal VMs
Server Sizing	vCPU: 2 Memory: 4GB Disk: 40GB	Standard sizing practice
Users per VM	Unknown	Not currently known
Network Connectivity	DMZ (if public, else customer facing) Internal (privileged network)	Connectivity for tenants (external/internal) Connectivity to back-end resources
Load Balancing	Yes – Load Balancer	Ensures availability of Admin Portal VMs Balances load between Admin Portal VMs & pods
Admin Authentication Site		
VM Quantity	Min: 2 (n+1) Scale: scale with ADFS	HA for Admin Authentication VMs
Server Sizing	vCPU: 2 Memory: 4GB Disk: 40GB	Standard sizing practice
Users per VM	Unknown	Not currently known
Network Connectivity	DMZ (if public, else customer facing) Internal (privileged network)	Connectivity for tenants (external/internal) Connectivity to back-end resources
Load Balancing	Yes – Load Balancer	Ensures availability of Admin Authentication VMs Balances load between Admin Authentication VMs & pods

Table 4. Infrastructure Design Decisions

Item	Detail	Rationale
Microsoft SCVMM		
VM Quantity	Min: 2 (n+1) Scale: scale per Services Pod	HA for SCVMM
Server Sizing	vCPU: 6 Memory: 16GB Disk: 40GB	Standard sizing practice
Managed VMs per Instance	8,000	Standard sizing practice Task Parallelization
Network Connectivity	Internal (privileged network)	Connectivity to back-end resources
Load Balancing	Yes – Native	Ensures availability of SCVMM environment
Active Directory		
Global Catalog/DNS Server(s)	Min: 2 (n+1) per site Scale: scale with Tenant Authentication Site	HA for GC/DNS Microsoft Best Practice
Additional Roles	ADFS	Required for Tenant and Admin Authentication Services
DHCP (if leveraged)*		
DHCP Server(s)	Min: 2 (n+1) per site	HA for DHCP Servers
Load Balancing	DHCP Server Failover Relationship	Ensures availability of DHCP Servers Balances load between DHCP Servers in operation
File Services		
DFS Server(s)	Min: 2 (n+1) per site	HA for DFS Servers
Load Balancing	Lowest Cost	Ensures availability of DFS Balances load between DFS Servers
SQL Server (Infra)		
SQL Server(s)	Min: 2 (n+1) per site Scale: 2 per additional pod	HA for SQL Servers
Server Sizing	vCPU: 8 Memory: 16-32GB Disk: 80GB (OS) + 6 x nGB (Data)	Standard sizing practice Multiple disks for TempDB, Data and Logs
Databases	SCVMM Azure Pack DBs Orchestrator DB SMA DB SCOM DB(s) – (if used)	Required databases for SCVMM & Azure Pack
Data Protection	SQL AlwaysOn Availability Group	Ensures availability of SQL Servers

*Static IPs and status DHCP reservation should be used for all infrastructure services. Tenant VMs can leverage DHCP Servers in the case that SCVMM IP Pools aren't leveraged.

Table 5. Network Design Decisions

Item	Detail	Rationale
Virtual Switches		
InternalSwitch	Use: Hyper-V to CVM local communication Uplink(s): N/A	Nutanix Default

Item	Detail	Rationale
ManagementSwitch	Use: All external Hyper-V Management traffic Uplink(s): MGMTAdapterTeam	Dedicated dual-1GbE network for Hyper-V host management traffic **NOTE**: can use LogicalSwitch if 10GbE only
Logical Switches		
LogicalSwitch	Use: All external and tenant VM communication. failover, migration and Nutanix traffic Uplink(s): NetAdapterTeam	Azure Pack Best Practice
NIC Teaming		
MGMTAdapterTeam	NIC(s): 2 x 1Gb Teaming mode: Switch Independent or LACP (if possible) Load balancing mode: Hyper-V Port (for 'Switch Independent' mode only)	Utilize both 1Gb adapters active/active N/A if using LogicalSwitch for Hyper-V host management traffic
NetAdapterTeam	NIC(s): 2 x 10Gb Teaming mode: Switch Independent or LACP (if possible) Load balancing mode: Hyper-V Port (for 'Switch Independent' mode only)	Utilize both 10Gb adapters active/active
VMQ	Enabled for Nutanix CVM (default) Can be disabled for Non-Nutanix VMs if VMQ limits are reached	Hyper-V network performance best practice
Network Site(s)		
Pool(s)	Default MAC address pool	Default
MAC Address Pools		
Pool(s)	Default MAC address pool	Default
Logical Networks		
Cloud Network	Name: Varies Network Site(s): Management Cloud	Segmentation between management and cloud Simplicity
VLANs		
Management VLAN	Network Site: Management ID: Varies Mask: /24 Capability: Routable Components: Hyper-V Hosts Nutanix CVMs SCVMM Azure Pack VMs SQL Servers AD/DHCP/DFS Servers	Dedicated infrastructure VLAN Best Practice
Live-Migration VLAN	Network Site: Management ID: Varies Mask: /24 Capability: Non-Routable Components: Hyper-V Hosts	Microsoft Best Practice

Item	Detail	Rationale
Front-end VLAN(s)	Network Site: Management ID: DMZ (for external) Mask: Varies Capability: Routable, Firewall/NAT (if-required) Components: Tenant Public API Management Portal for Tenants Tenant Authentication Site	Network segmentation for front-end or external services
Tenant Public VLAN(s)*	Network Site: Cloud ID: Multiple Mask: Multiple Capability: Routable Components: Tenant VM(s)	Used for external tenant networks (if applicable)
VM Network(s)		
General	One, or more, VM Network per tenant (if segmented) Logical Network: CloudNetwork Isolation: Hyper-V network virtualization VM subnets: Varies	Dedicated infrastructure VLAN Best Practice
Gateway	Recommended to leverage a 3rd party gateway	Best Practice

Resource Provider Sizing

Dell XC Series nodes can co-host VM Clouds, Hosted SQL Databases and/or Websites services, however for sizing we've chosen to size by VMs per node and use assumed SQL and/or website VM sizes for non-VM sizing.

NOTE: These densities are based upon Azure Pack/Database sizing best practices and in-house validation. Densities will vary based upon specific images and workload. Resources for the XC Series CVM have been accounted for.

The workload and vCPU over-subscription has been characterized by the following classifications:

- Light: 8:1 vCPU/pCore Ratio
- Medium: 6:1 vCPU/pCore Ratio
- Heavy: 4:1 vCPU/pCore Ratio

Table 6. Node vCPU Density

Available vCPUs	Workload/vCPU Density		
Node Model	**Light**	**Medium**	**Heavy**
XC720xd-B7 (dual 10 core)	160	120	80
XC720xd-B5 (dual 8 core)	128	96	64

More information on node models can be found on the Dell XC Tech Specs,

VM Sizing

The sizing for VMs has been characterized by the following classifications:

Table 7. VM Scenario Sizing

Scenario	Resources		
	vCPU	MEM	Storage
Small	1	4	1 x 40GB (OS) + 1 x 100 GB (APP)
Medium 1	2	8	1 x 40GB (OS) + 2-4 x 100 GB (APP)
Medium 2	4	16	1 x 40GB (OS) + 2-4 x 100 GB (APP)
Large	8	24	1 x 40GB (OS) + 6 x 100 GB (APP)

The sizing for Databases and Websites has been characterized by the following assumptions:

- Database VMs:
    - 1 x Large VM Size (x2 if AlwaysOn used)
- Websites VMs:
    - 1 x Small VM size for Web Workers
    - 2 x Medium 2 VM size for Database and File Server

Below we apply these assumptions to show the estimated VM density per node:

Table 8. Node Sizing Estimates

Node Type	Workload/VM Density		
	Light	Medium	Heavy
VM – Small	128	96	64
VM – Medium 1	64	48	32
VM – Medium 2	32	24	16
VM – Large	16	12	8
Websites	14	10	7
Database	16	12	8

Figure 11 is an example of an Azure Pack node providing virtual machines.

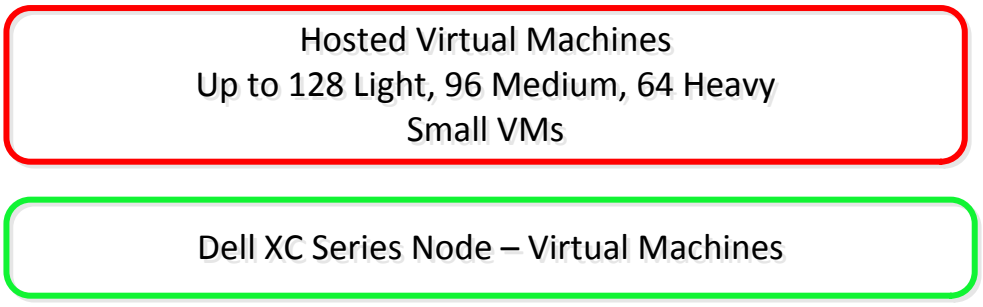


Figure 11. Hosted VM Node

Figure 12 is an example of a SQL Database node providing hosted shared SQL instances.

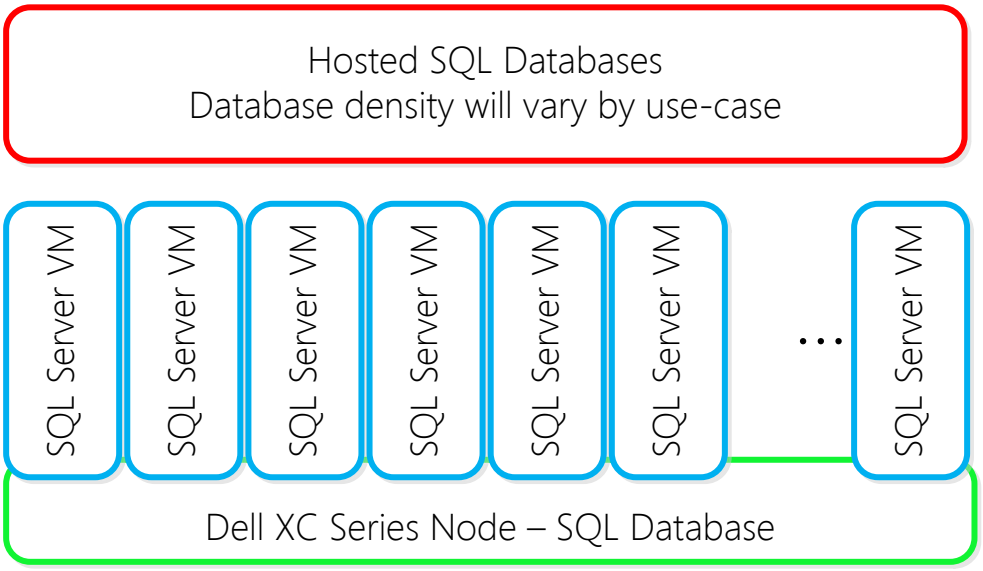


Figure 12. SQL Database Node

Figure 13 is an example of a Web Sites node providing hosted web site services.

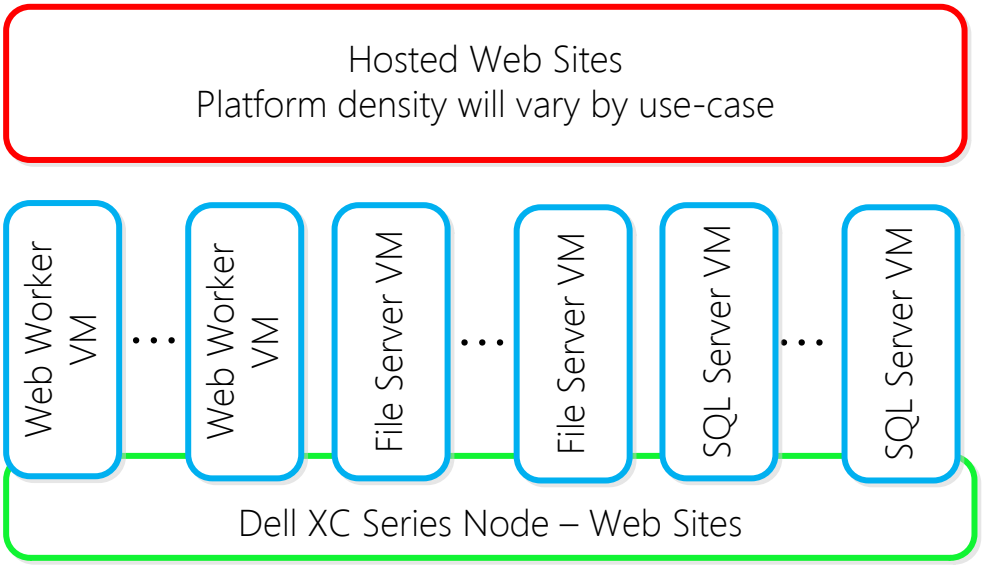


Figure 13. Web Sites Node

Site Planning

For large scale cloud deployments it is important to leverage a delivery topology that will be distributed and meet the requirements of the end-users while providing flexibility and locality.

We've leveraged a two-tier strategy where the sites are grouped into Regions and Availability Zones. The Azure Pack infrastructure components would be spread across Availability Zones and abstracted using a Load Balancer (GSLB).

Region: A geographic landmass or area where multiple Availability Zones are located. These can include regions like US-Northwest or US-West.

Availability Zone: A specific site or datacenter location where cloud services are hosted. These can include co-located facilities or on-site hosting facilities.

Below we show a logical representation of the relationship between a Region and Availability Zone:

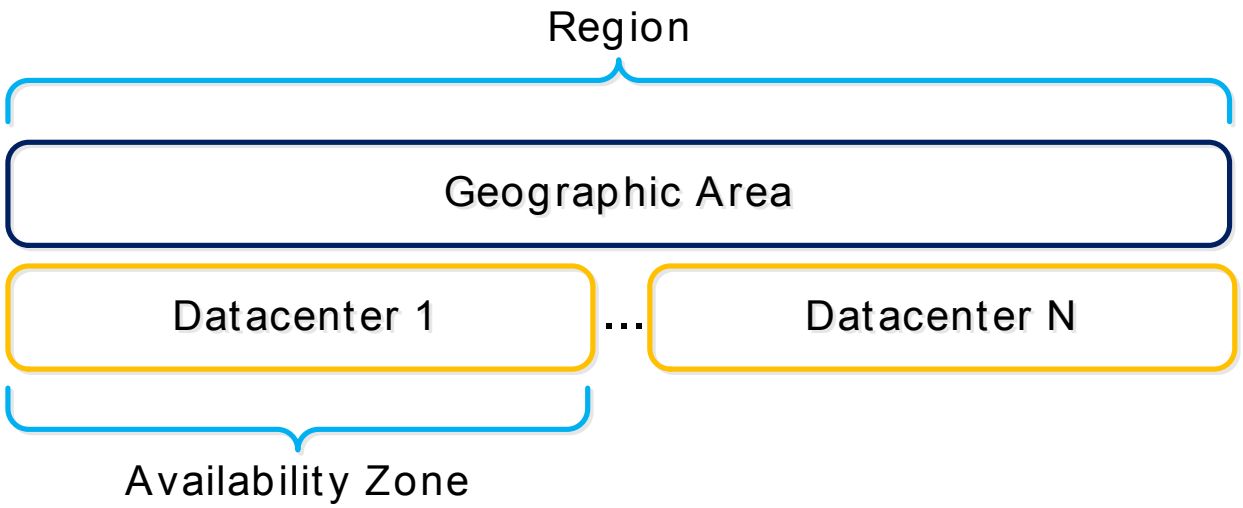


Figure 14. Region Mapping

Figure 15 is an example application of this concept.

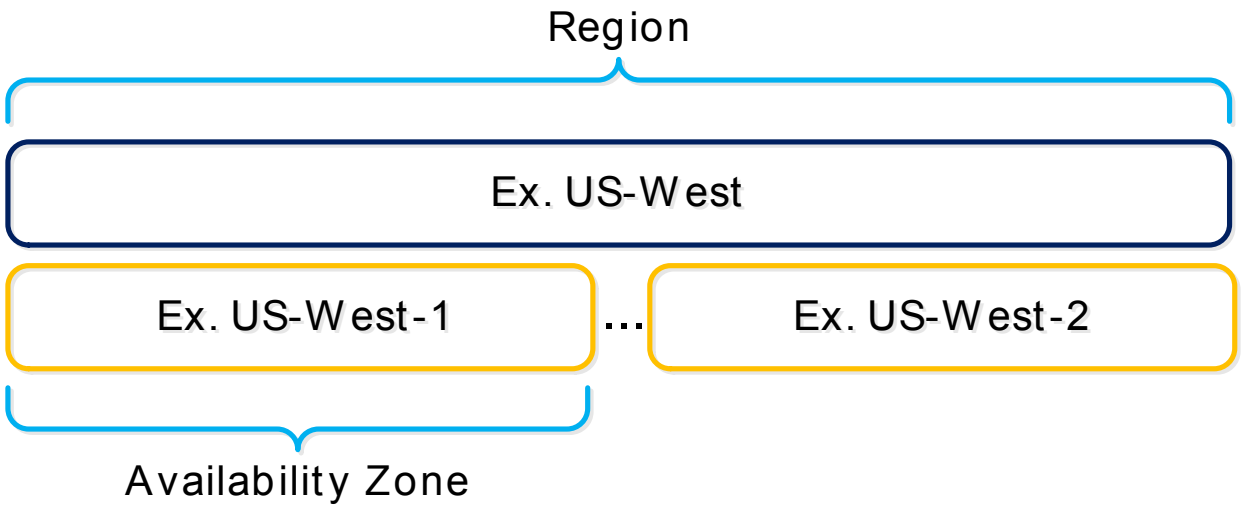


Figure 15. Region Mapping Example

While this construct is recommended for larger and/or global deployments it is not required for smaller deployments. For larger deployments it is recommended that a Region contain a minimum of two Availability Zones for HA and DR purposes.

Azure Pack Pod

We've broken the Azure Pack components into a modular pod based architecture for easy scaling and management. The pod consists of two modular pods: Control Pod and Services Pod. The Control Pod contains all of the Azure Pack management components and VMs while the Services Pod hosts the Azure Pack resources (such as VMs, Databases, Web Sites).

For smaller deployments the resources will be shared between the Control and Services Pods, however for larger deployments it is recommended that a dedicated cluster be used for hosting these components.

NOTE: The term block is utilized to specified 4 Dell XC nodes.

Azure Pack Pod Design

Table 9 is a high-level snapshot of the Windows Azure Pack on Dell XC highlights.

Table 9. Azure Pack Pod Detail

Item	Qty
Control Pod	
# of SCVMM Server(s)	2
# of Public Tenant API Server(s)	2
# of Tenant API Server(s)	2
# of Management Portal for Tenant Server(s)	2
# of Tenant Authentication Site Server(s)	2
# of Admin API Server(s)	2
# of Management Portal for Admin Server(s)	2
# of Admin Authentication Site Server(s)	2
# of Infra SQL Server(s)	2
# of Active Directory Server(s)	2
# of DFS/DHCP Servers(s)	2
Services Pod	
# of Dell XC Node(s)	Up to 48
# of Hyper-V Hosts	Up to 48
# of Dell XC Cluster(s)	Up to 2
# of Datastore(s)	Varies

Table 10 is a breakdown of the resource requirements for the Azure Pack Components.

Table 10. Azure Pack Pod Resources

Component	Resources			
	Qty	vCPU	Memory(GB)	Disk(GB)
SCVMM Server(s)	2	4	16	80
Public Tenant API Server(s)	2	2	4	40
Tenant API Server(s)	2	2	4	40
Management Portal for Tenant Server(s)	2	2	4	40
Tenant Authentication Site Server(s)	2	2	4	40
Admin API Server(s)	2	2	4	40
Management Portal for Admin Server(s)	2	2	4	40
Admin Authentication Site Server(s)	2	2	4	40
Infra SQL Server(s)	2	8	32	400
Active Directory Server(s)	2	2	4	40
DFS/DHCP Servers(s)	2	2	4	Variable
TOTAL	22	60	168	1600

Below we show a logical relationship between the Control and Services Pods:

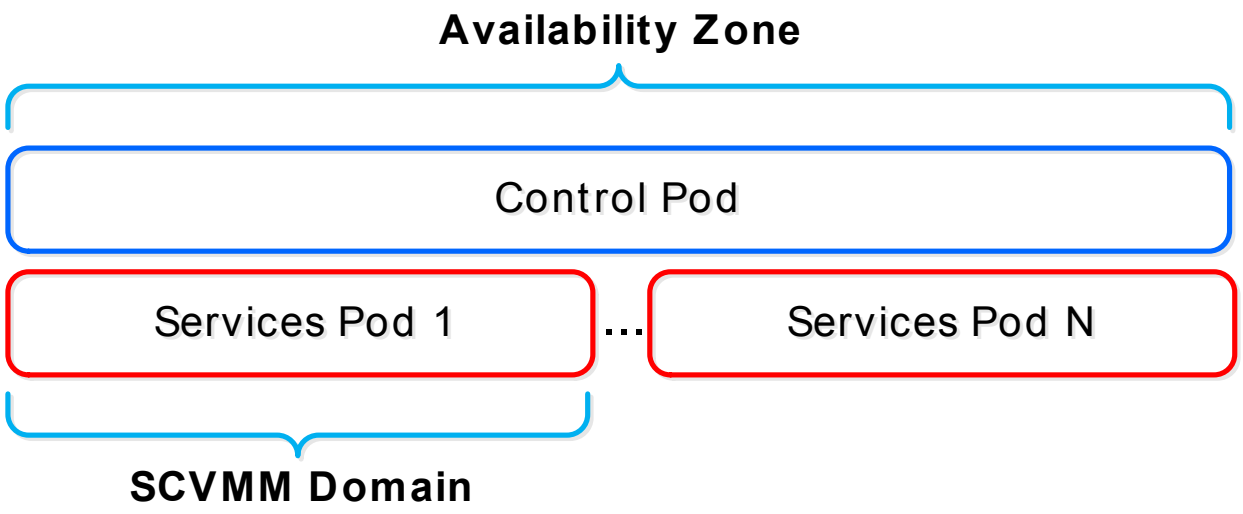


Figure 16. Logical relationship

Figure 17 shows a logical representation of the Control Pod.

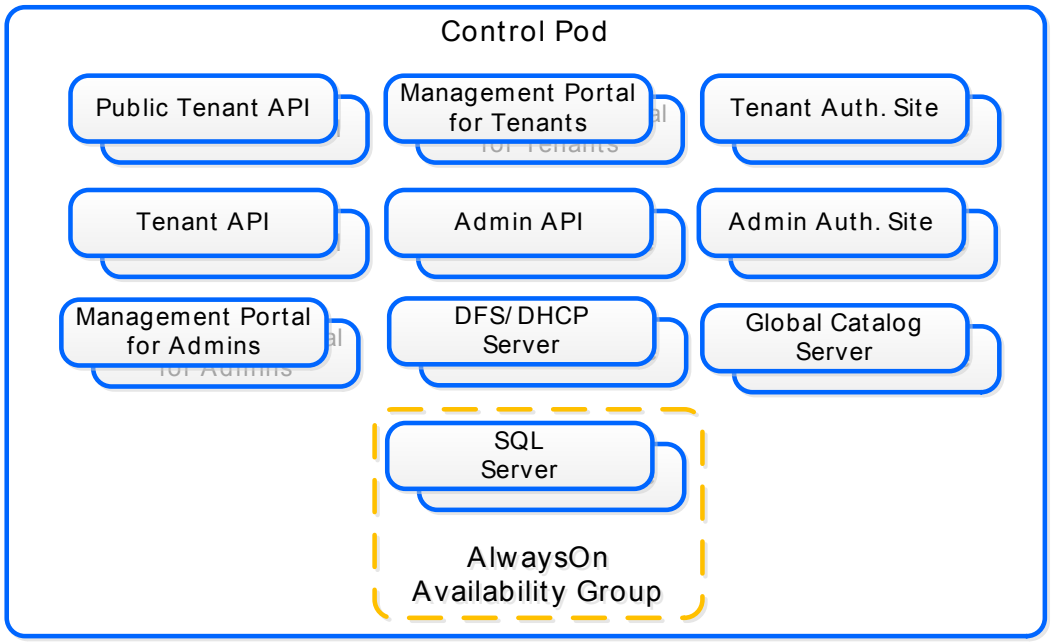


Figure 17. Azure Pack Control Pod Detail

Figure 18 shows a logical representation of the Services Pod.

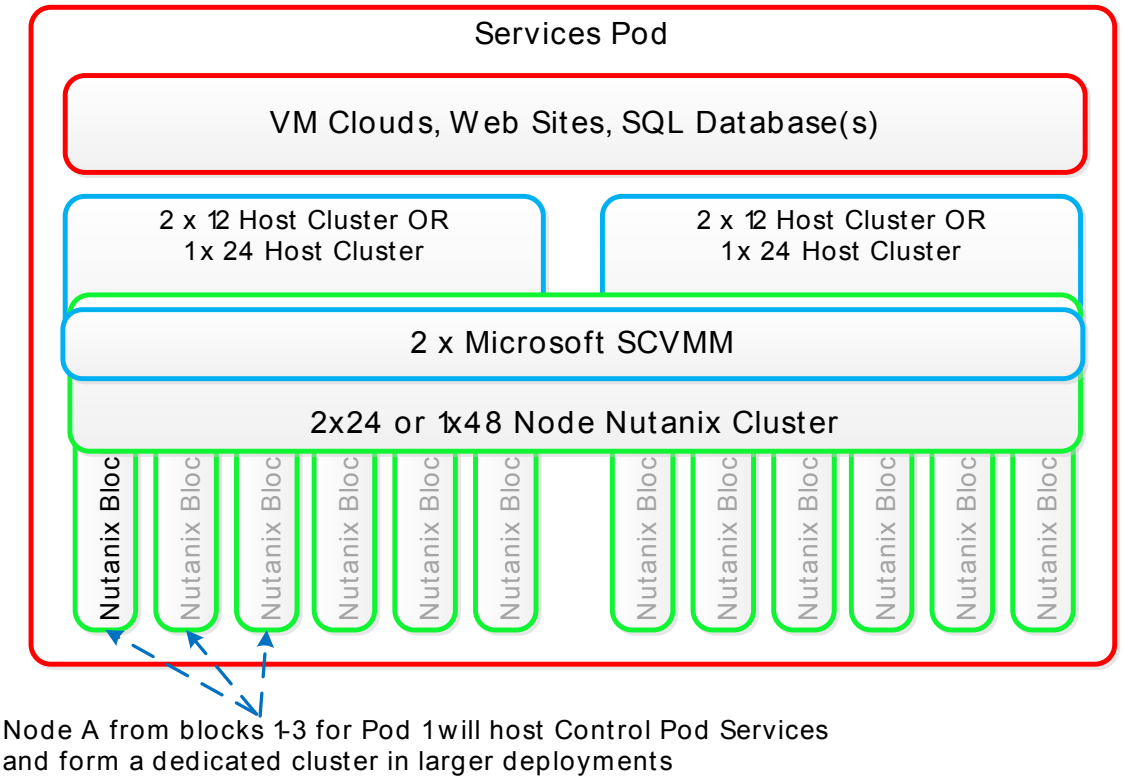


Figure 18. Azure Pack Services Pod Detail

Single-Site Deployment

For smaller deployments only a single site might be used within a region. In this deployment, the Azure Pack and SCVMM components will be hosted by a single site.

Figure 19 shows an example of single-site deployment.

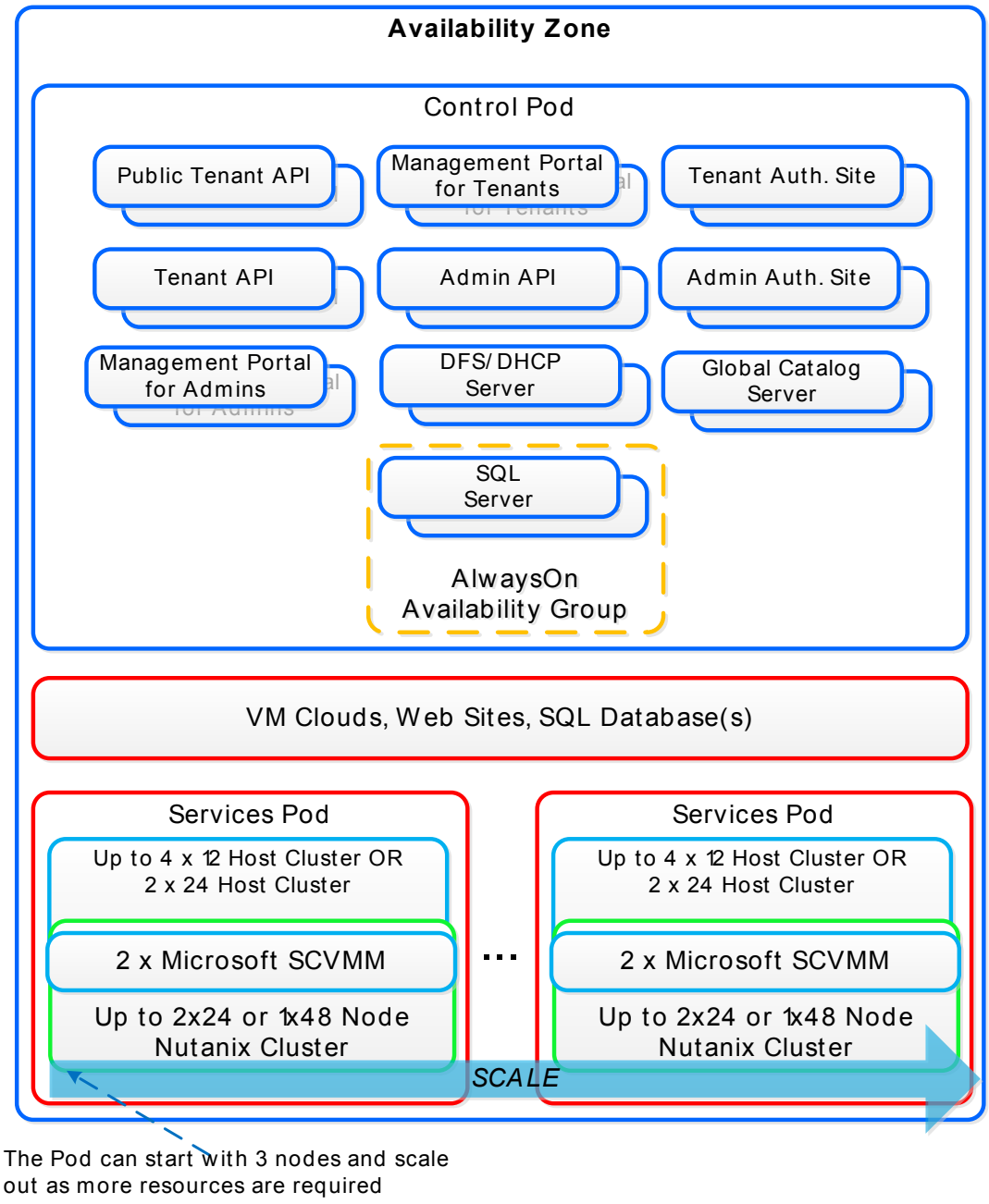


Figure 19. Azure Pack Single Site

Multi-Site Deployment

For larger deployments it is recommended to utilize multiple Availability Zones within a region for availability. In this deployment the Azure Pack components will be distributed between the two sites and leverage a load balancing for distribution.

Figure 20 is an example of multi-site deployment.

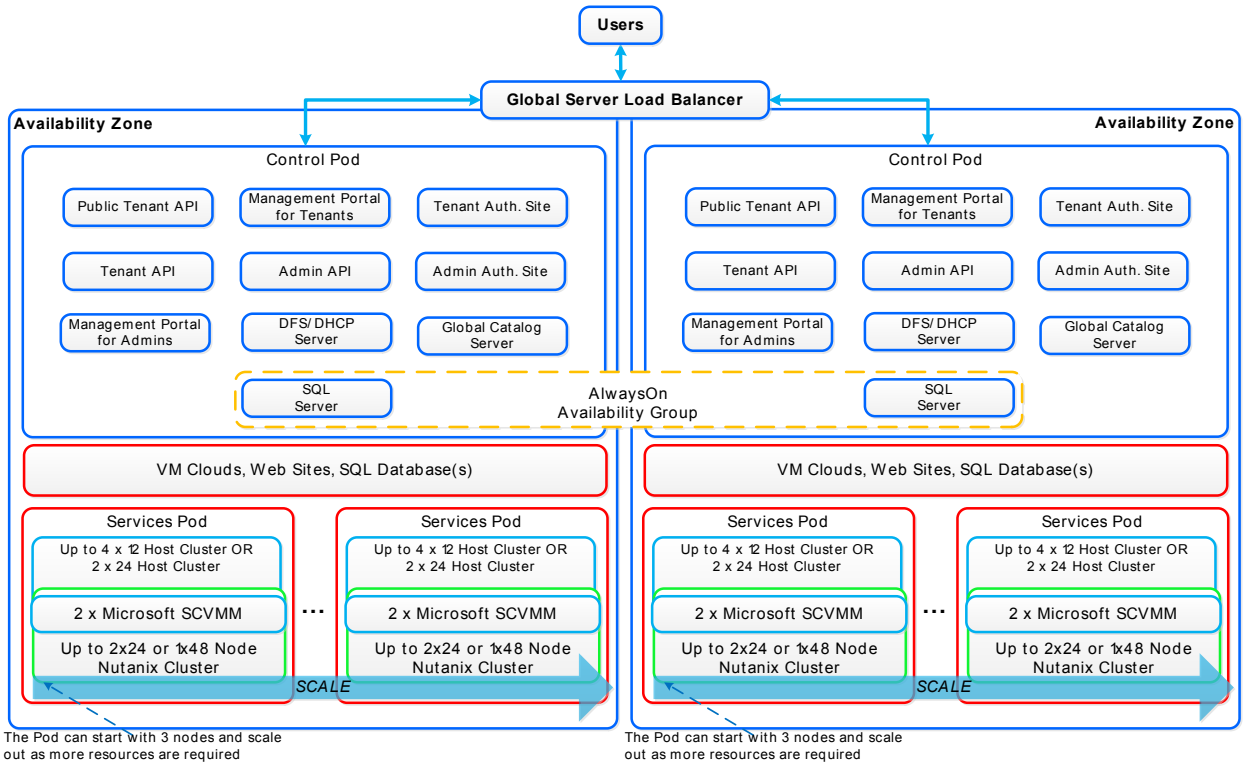


Figure 20. Azure Pack Multi-Site

Dell XC Series – Compute/Storage

The Dell XC Platform provides an ideal combination of both high-performance compute with localized storage to meet any demand. True to this capability, this reference architecture contains zero reconfiguration of or customization of the Nutanix software to optimize for this use case.

Figure 21 shows a high-level example of the relationship between a Dell XC block, node, storage pool and container

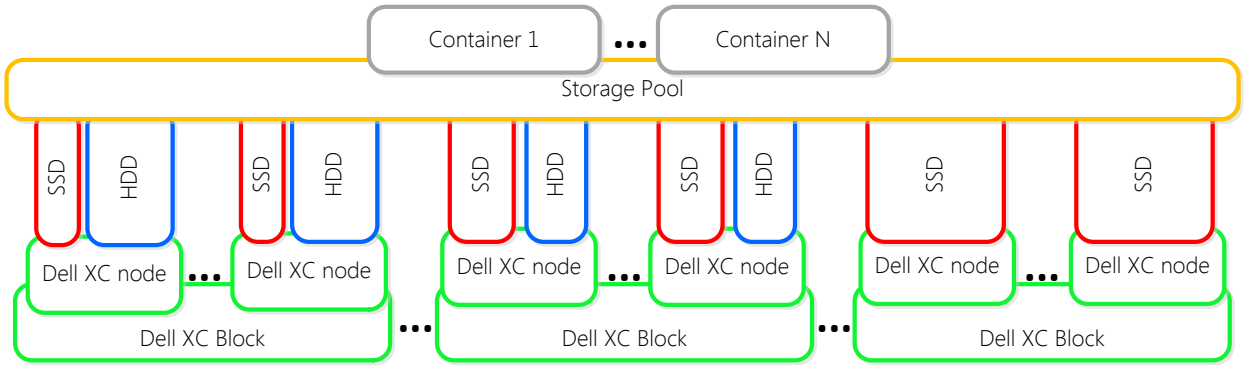


Figure 21. Nutanix Component Architecture

Table 11 shows the Dell XC storage pool and container configuration.

Table 11. Nutanix Storage Configuration

Name	Role	Details
SP01	Main storage pool for all data	All Disks
CTR-LIBRARY-01	Container for Library Server Share and VM Templates	SCVMM – Library Share
CTR-INFRA-01	Container for all VMs	Hyper-V – Datastore
CTR-TEN-001*	Container for all Tenant VMs	Hyper-V – Datastore

*For tenant storage it is recommended to have a single NDFS SMB share backing the VMs and data. However, where logical data separation is required a NDFS Container should be created for each tenant.

Network

Designed for true linear scaling, Dell recommends a Leaf Spine network architecture. A Leaf Spine architecture consists of two network tiers: an L2 Leaf and an L3 Spine based on 40GbE and non-blocking switches. This architecture maintains consistent performance without any throughput reduction due to a static maximum of three hops from any node in the network.

Figure 22 shows a design of a scale-out Leaf Spine network architecture which provides 20Gb active throughput from each node to its L2 Leaf and scalable 80Gb active throughput from each Leaf to Spine switch providing scale from 1 Nutanix block to thousands without any impact to available bandwidth

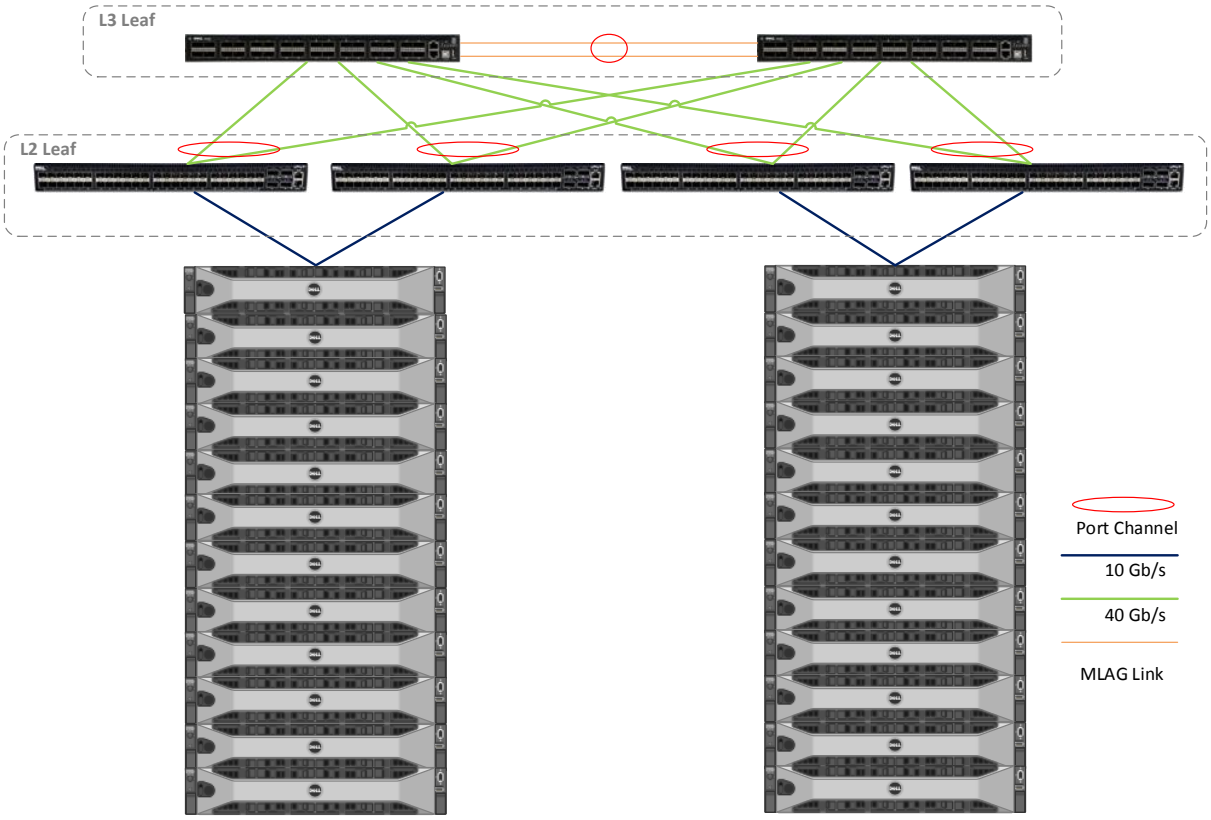


Figure 22. Leaf Spine Network Architecture

Logical Network Design

Each Hyper-V host has two default switches for internal and external communication. The ExternalSwitch is utilized for external node communication and VM traffic and has 10GbE uplinks in a LBFO team. The InternalSwitch is utilized for SMB I/O between the Hyper-V host and the Dell XC CVM. For SQL AlwaysOn and for Control and Service PODs, the internal switch can be leveraged to use the NDFS file share as a witness.

Figure 23 shows a logical network representation of the network segments used in the solution and corresponding components attached.

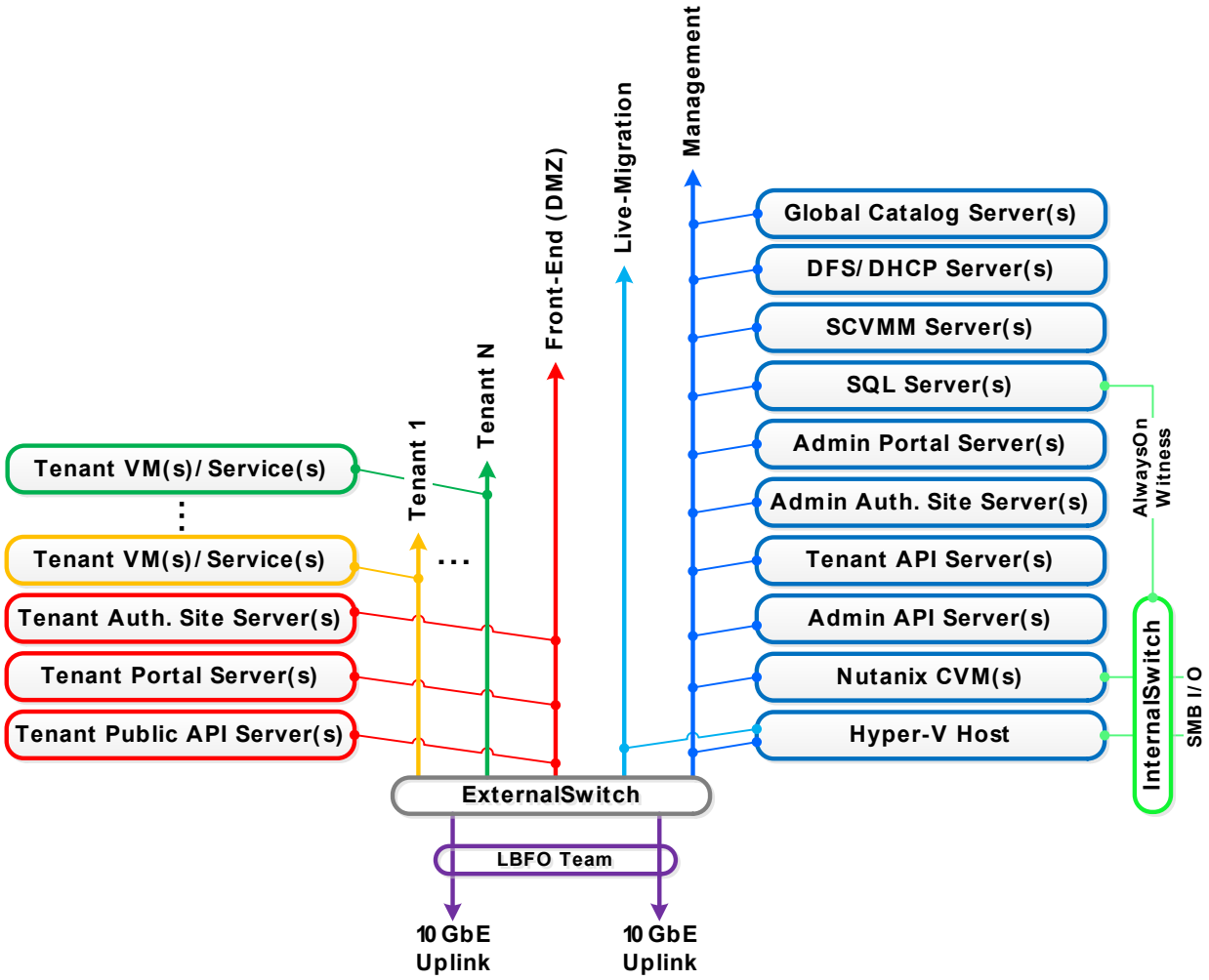


Figure 23. Logical Network Connectivity

5. Solution Application

This section applies this pod based reference architecture to real world scenarios and outlines the sizing metrics and components. The applications below assume a standard medium user desktop/workload however will vary based upon utilization and workload.

NOTE: Detailed hardware configuration and product models can be found in the appendix.

Scenario: 4 Nodes

Table 12. Detailed Component Breakdown – 4 Nodes

Item	Value	Item	Value
Components		**Infrastructure**	
# of Nutanix Services Pods	1/12 (partial)	# of SCVMM Servers	1
# of Dell XC Nodes	4	# of Hyper-V Hosts	4
# of RU (Nutanix)	8	# of Hyper-V Clusters	1
# of 10GbE Ports	8	# of SMB Share(s)	3
# of 100/1000 Ports (IPMI)	4	# of vCPU:	512 Light
# of L2 Leaf Switches	2		384 Medium
# of L3 Spine Switches	1		256 Heavy

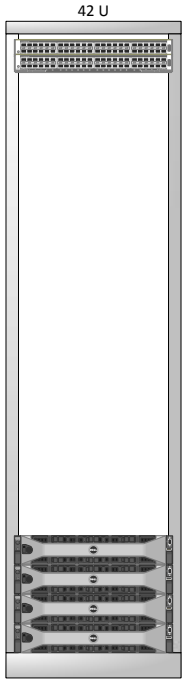


Figure 24. Rack Layout – 4 Nodes

Scenario: 1/4 Pod – 12 Nodes

Table 13. Detailed Component Breakdown – ¼ Pod – 12 Nodes

Item	Value	Item	Value
Components		**Infrastructure**	
# of Nutanix Services Pods	1/4 (partial)	# of SCVMM Servers	1
# of Dell XC Nodes	12	# of Hyper-V Hosts	12
# of RU (Nutanix)	24	# of Hyper-V Clusters	1
# of 10GbE Ports	24	# of SMB Share(s)	3
# of 100/1000 Ports (IPMI)	12	# of vCPU:	1,536 Light
# of L2 Leaf Switches	2		1,152 Medium
# of L3 Spine Switches	1		768 Heavy

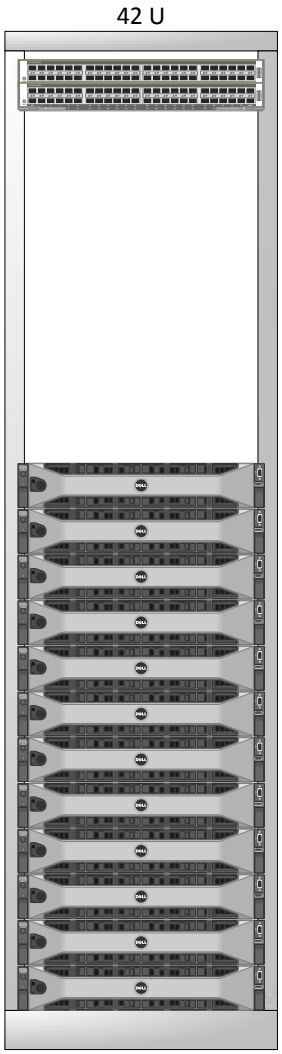


Figure 25. Rack Layout – ¼ Pod – 12 Nodes

Scenario: 1 Pod – 48 Nodes

Table 14. Detailed Component Breakdown – 1 Pod – 48 Nodes

Item	Value	Item	Value
Components		**Infrastructure**	
# of Nutanix Services Pods	1	# of SCVMM Servers	2
# of Dell XC Nodes	48	# of Hyper-V Hosts	48
# of RU (Nutanix)	96	# of Hyper-V Clusters	2-4
# of 10GbE Ports	96	# of SMB Share(s)	3
# of 100/1000 Ports (IPMI)	48	# of vCPU:	6,144 Light
# of L2 Leaf Switches	2		4,608 Medium
# of L3 Spine Switches	1		3,072 Heavy

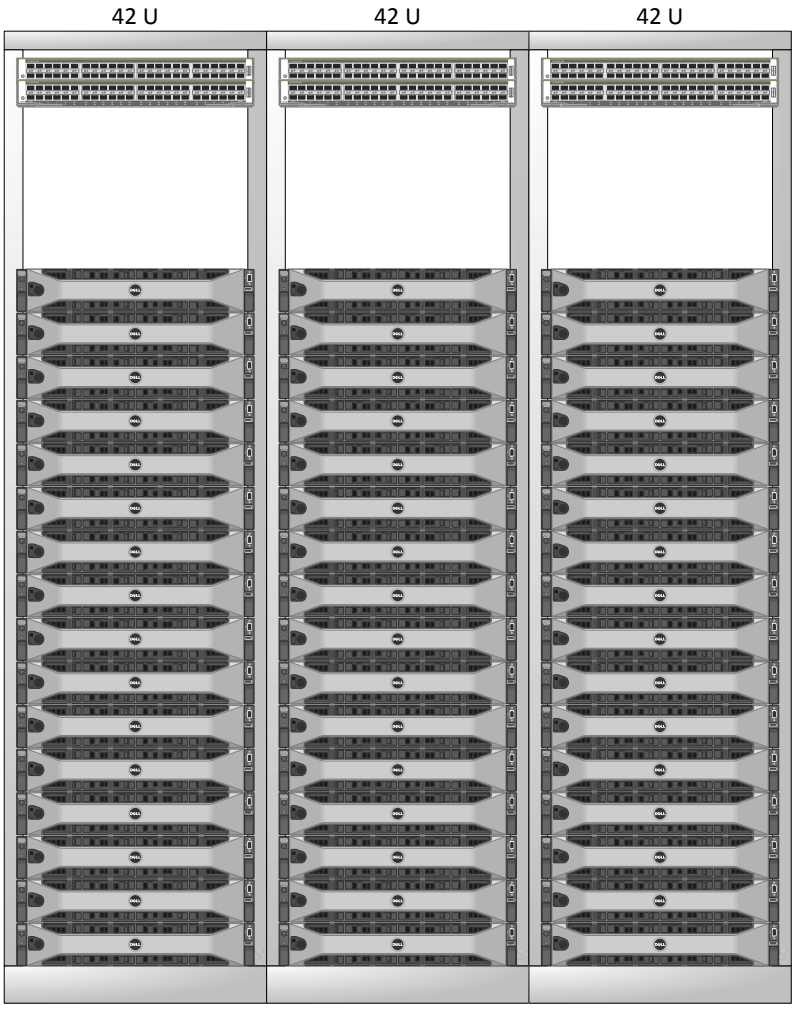


Figure 26. Rack Layout – 1 Pod – 48 Nodes

Scenario: 2 Pods – 96 Nodes

Table 15. Detailed Component Breakdown – 2 Pods – 96 Nodes

Item	Value	Item	Value
Components		**Infrastructure**	
# of Nutanix Services Pods	2	# of SCVMM Servers	4
# of Dell XC Nodes	96	# of Hyper-V Hosts	96
# of RU (Nutanix)	192	# of Hyper-V Clusters	4-8
# of 10GbE Ports	192	# of SMB Share(s)	6
# of 100/1000 Ports (IPMI)	96	# of vCPU:	12,288 Light
# of L2 Leaf Switches	4		9,216 Medium
# of L3 Spine Switches	2		6,144 Heavy

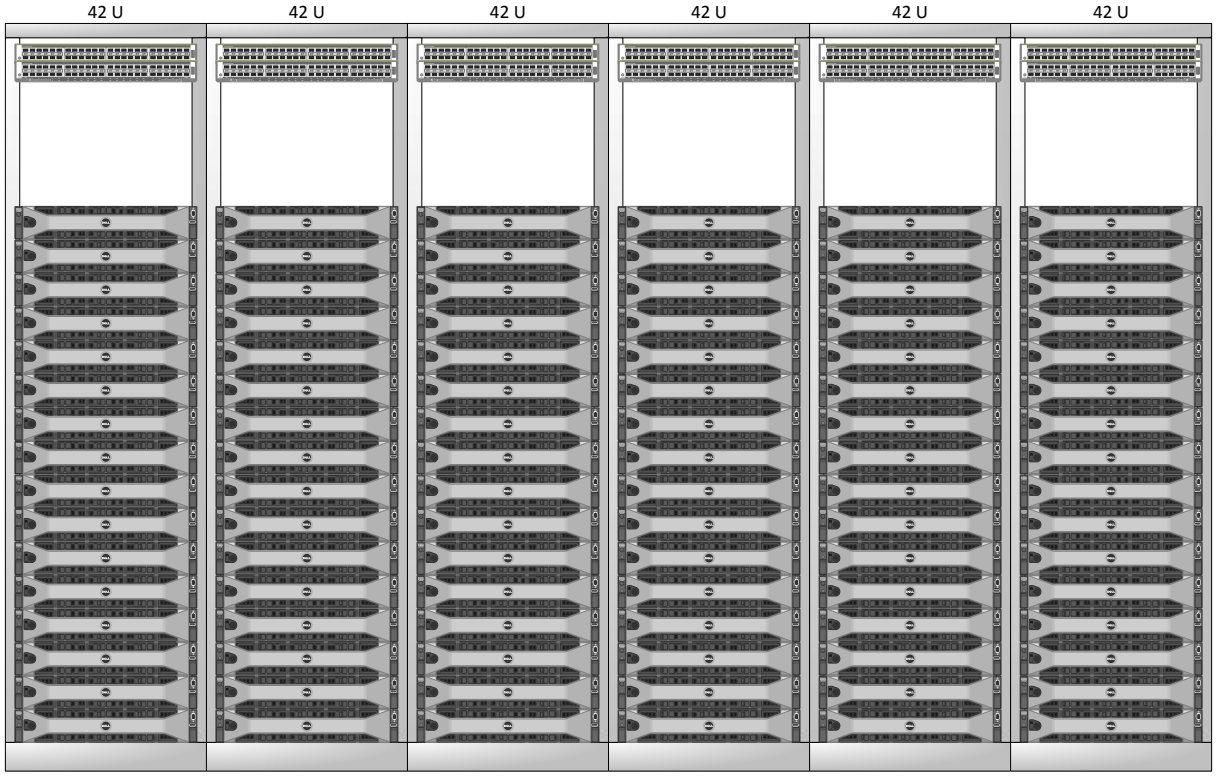


Figure 27. Rack Layout – 2 Pods – 96 Nodes

6. Further Research

As part of its continuous determination to deliver the best possible solutions, Dell and Nutanix will continue to research into the following areas:

- Windows Azure Pack integrations and/or custom Resource Provider
- Scale testing
- Detailed use-case application
- Joint solutions with partners

7. Conclusion

Sizing for the pods was determined after careful consideration of performance as well as accounting for additional resources for N+1 failover capabilities. Assumed vCPU densities per node were 128 light, 96 medium, or 64 heavy based upon assumed vCPU to pCore ratios.

NOTE: these will vary per deployment based upon the customer's over-subscription standards.

This validated deployment was based upon a large scale distributed deployment of Windows Azure Pack. At this time the scalability limits of the components are unknown so a conservative approach was taken in regards to users/sessions per component. It is recommended to have redundancy for each component, ideally across two availability zones (site). SQL Server AlwaysOn should be used for SQL Server HA.

Sizing for the Services Pod was based upon the standard limits of 8,000 VMs per SCVMM instance. The 48 node Dell XC/Hyper-V cluster per Services Pod would be theoretically able to run up to ~6,144 VMs which fits nicely under the 8,000 VM limit. Multiple Services Pods can be under a single Control Pod and a full Pod is not required. Deployments can start at with minimum of 3 nodes.

Given the current nature of Windows Azure Pack and SCVMM's utilization of ODX which performs a offloaded full copy, it is recommended leveraging the Nutanix On-Disk Deduplication for capacity savings and performance. Further research and integration will be done to allow Windows Azure Pack to leverage more of the Nutanix system's capabilities.

The Windows Azure Pack and Microsoft Hyper-V on Dell XC solution provides a single high-density platform for desktop and application delivery. This modular pod based approach enables these deployments to easily be scaled.

8. Appendix: Configuration

Hardware

- Storage / Compute
  - Dell XC720xd-B7
    - Per node specs (2U):
      - CPU: 2x Intel® Xeon® E5-2680v2
      - Memory: 128-768 GB Memory
  - Dell XC720xd-C5
    - Per node specs (2U):
      - CPU: 2x Intel Xeon E5-2690v2
      - Memory: 128-768 GB Memory
  - Dell XC720xd-C7
    - Per node specs (2U):
      - CPU: 2x Intel Xeon E5-2690v2
      - Memory: 128-768 GB Memory
- Network
  - Dell S-series 40Gb - L3 Spine
  - Dell S-series 10Gb - L2 Leaf

Software

- Nutanix
  - NOS 4.0
- Azure Pack Extension
- Infrastructure
  - Windows Server 2012 R2 – Hyper-V Role
  - System Center Virtual Machine Manager 2012 R2

Sample Bill of Materials

Table 16. Sample BOM - XC720xd-B7

XC720xd-B7 Example		
Qty	**P/N**	**Description**
3 - n	XC720xd-B7	Dell XC Platform node **Compute** ▪ 2 x Intel 10-core Xeon E5-2680v2 processors per node **Storage** ▪ Nutanix Controller VM ▪ Nutanix Virtualization-ready Platform ▪ Nutanix Information Lifetime Management ▪ Nutanix Data Protection and Availability

XC720xd-B7 Example		
		▪ Nutanix Automatic Storage Provisioning ▪ Nutanix Advanced Storage Management ▪ Nutanix Capacity Optimization ▪ Nutanix Seamless Scale-out ▪ 800GB SSD raw capacity per node (for High Priority Data and MDS) ▪ 4TB HDD raw capacity per node (for Data Storage) **Dell XC Platform - Management Accessories/Block** ▪ 1 x Bezel ▪ 1 x Rail Kit ▪ 1 x Rail Kit Adapter Kit (Round to square holes rail adapters and rail screw bag) ▪ 2 x C13/C14 Cables ▪ 1 x OEM Documentation Kit
Variable	Variable	128-768GB Memory per node
Variable	Variable	SFP+ to SFP+ cable

Table 17. Sample BOM - XC720xd-C5

XC720xd-C5 Example		
Qty	**P/N**	**Description**
3 - n	XC720xd-C5	Dell XC Platform node **Compute** 2 x Intel 10-core Xeon E5-2690v2 processors per node **Storage** Nutanix Controller VM Nutanix Virtualization-ready Platform Nutanix Information Lifetime Management Nutanix Data Protection and Availability Nutanix Automatic Storage Provisioning Nutanix Advanced Storage Management Nutanix Capacity Optimization Nutanix Seamless Scale-out

XC720xd-C5 Example		
		1.6TB SSD raw capacity per node (for High Priority Data and MDS) 20TB HDD raw capacity per node (for Data Storage) **Dell XC Platform - Management** **Accessories/Block** 1 x Bezel 1 x Rail Kit 1 x Rail Kit Adapter Kit (Round to square holes rail adapters and rail screw bag) 2 x C13/C14 Cables 1 x OEM Documentation Kit
Variable	Variable	128-768GB Memory per node
Variable	Variable	SFP+ to SFP+ cable

Table 18. Sample BOM XC720xd-C7

XC720xd-C7 Example		
Qty	**P/N**	**Description**
3 - n	XC720xd-C7	2-node Dell XC Platform **Compute** 8 x Intel 10-core Xeon E5-2690v2 processors **Storage** Nutanix Controller VM Nutanix Virtualization-ready Platform Nutanix Information Lifecycle Management (ILM) Nutanix Data Protection and Availability Nutanix Online Cluster Grow and Shrink

XC720xd-C7 Example		
		Nutanix Advanced Storage Management Nutanix Capacity Optimization Nutanix Seamless Scale-out 3.2TB SSD raw capacity (for High Priority Data and MDS) 32TB HDD raw capacity (for Data Storage) **Dell XC Platform - Management** **Accessories/Block** 1 x Bezel 1 x Rail Kit 1 x Rail Kit Adapter Kit (Round to square holes rail adapters and rail screw bag) 2 x C13/C14 Cables 1 x OEM Documentation Kit
Variable	Variable	128-768GB Memory per node
Variable	Variable	SFP+ to SFP+ cable