Dell XC Web-Scale Converged Appliance for VMware® Horizon View™

Dell Wyse Technical Marketing

March 2015

A Dell Appliance Architecture

Revisions

Date	Description
March 2015	Initial release

Contents

1 Introduction

1.1 Purpose

This document addresses the architecture design, configuration and implementation considerations for the key components required to deliver full-clone persistent virtual desktops via VMware® Horizon View™ on VMware vSphere® 5.5.

1.2 Scope

Relative to delivering the virtual desktop environment, the objectives of this document are to:

- Define the detailed technical design for the solution.
- Define the hardware requirements to support the design.
- Define the constraints which are relevant to the design.
- Define relevant risks, issues, assumptions and concessions – referencing existing ones where possible.
- Provide a breakdown of the design into key elements such that the reader receives an incremental or modular explanation of the design.
- Provide solution scaling and component selection guidance.

2   Solution Architecture Overview

2.1   Introduction

The Dell XC series delivers an out-of-the-box infrastructure solution for virtual desktops that eliminates the high cost, variable performance, and extensive complexity of conventional infrastructure. The Nutanix™-based Dell XC web-scale converged infrastructure is a turnkey solution that comes ready to run your VDI solution of choice. The Nutanix platform's unique architecture allows enterprises to scale their virtual desktops starting from 200 to tens of thousands of desktops in a linear fashion, providing customers with a simple path to enterprise deployment with the agility of public cloud providers.

2.2   Nutanix Architecture

The Nutanix web-scale converged infrastructure is a scale-out cluster of high-performance nodes (or servers). Each node runs a standard hypervisor and contains 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.
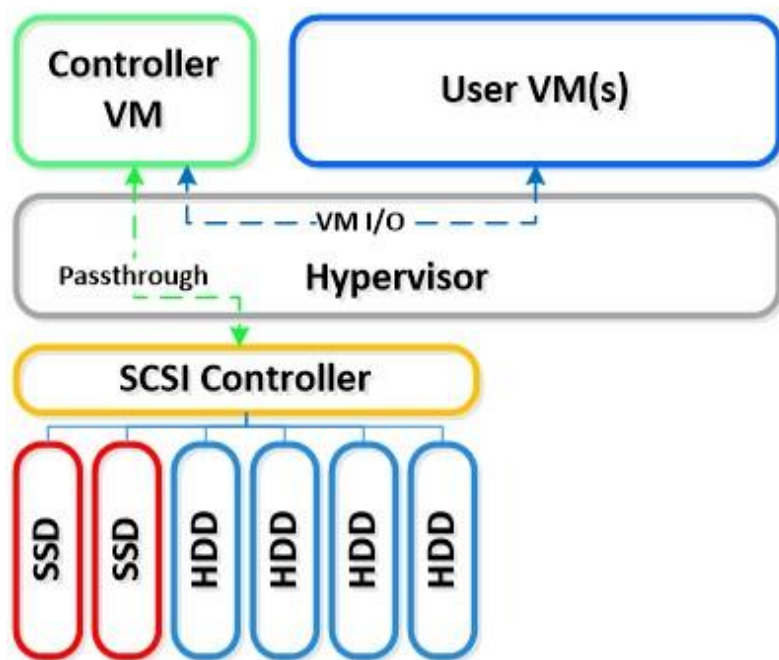
In addition, local storage from all nodes is virtualized into a unified pool by the Nutanix Distributed File System (NDFS). In effect, NDFS acts like an advanced NAS that uses local SSDs and disks from all nodes to store virtual machine data. Virtual machines running on the cluster write data to NDFS as if they were writing to shared storage.
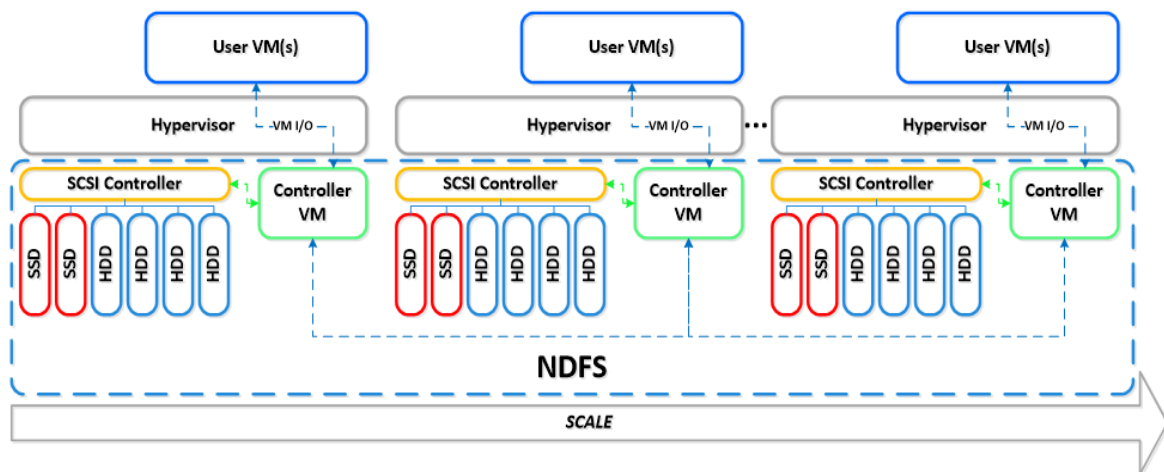
NDFS understands the concept of a virtual machine and provides advanced data management features. It brings data closer to virtual machines by storing the data locally on the system, resulting in higher performance at a lower cost. Nutanix platforms can horizontally scale from as few as three nodes to 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.
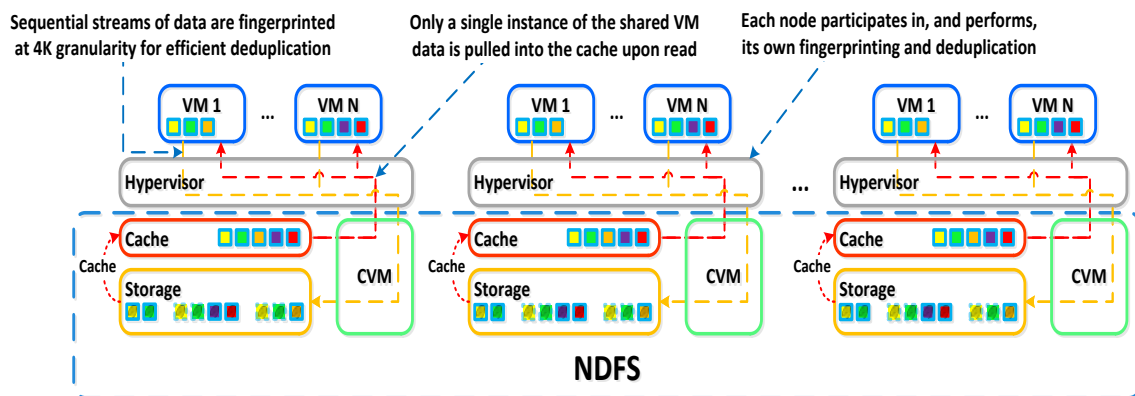
Nutanix Shadow Clones delivers distributed localized caching of virtual disks performance in multi-reader scenarios, such as desktop virtualization using VMware Horizon View with Shadow Clones; the CVM actively monitors virtual disk access trends. If there are read I/O requests originating from the local CVM and two or more remote CVMs, the virtual disk will be marked as immutable. Once the disk has been marked immutable, the virtual disk is then cached locally by each CVM, so read operations are now satisfied locally by local storage.

NOTE: When a clone or snapshot occurs the base vDisk will become read-only

Read I/O for Base VM is served locally from cached Shadow vDisk

Once NDFS has determined the Base VM's target vDisk is multi-reader, the vDisk will be marked as immutable and the shadow vDisk can then be cached on each local CVM

NDFS
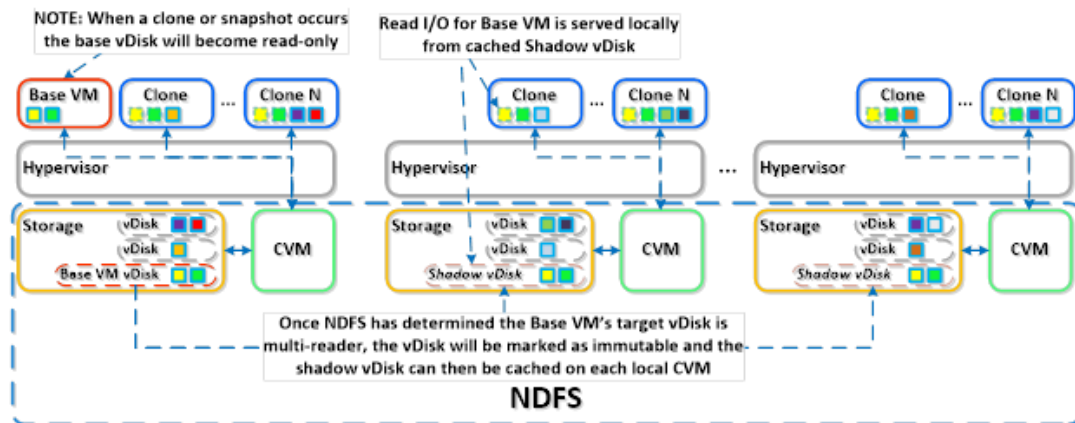
The benefits of the Nutanix Platform are now exposed to scale out vSphere or Hyper-v deployments:
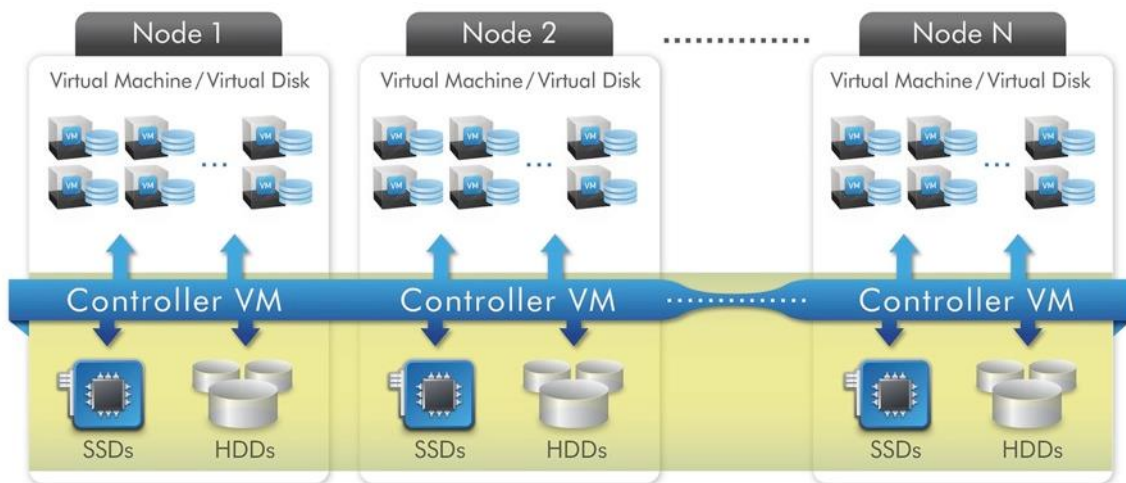
Nutanix Web-scale Converged Infrastructure

The Nutanix web-scale converged infrastructure 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 to the Nutanix product to optimize for this use case.

The next figure shows a high-level example of the relationship between a Nutanix block, node, storage pool, and container:
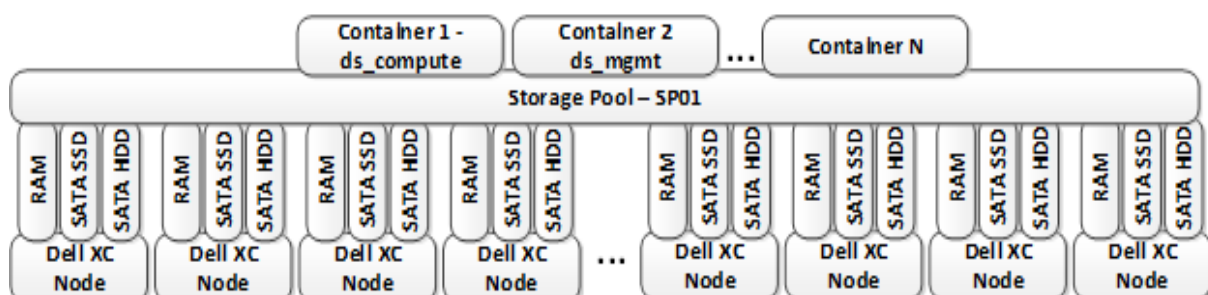
This appliance allows organizations to deliver virtualized or remote desktops and applications through a single platform. It also supports end users with access to all of their desktops and applications in a single location.
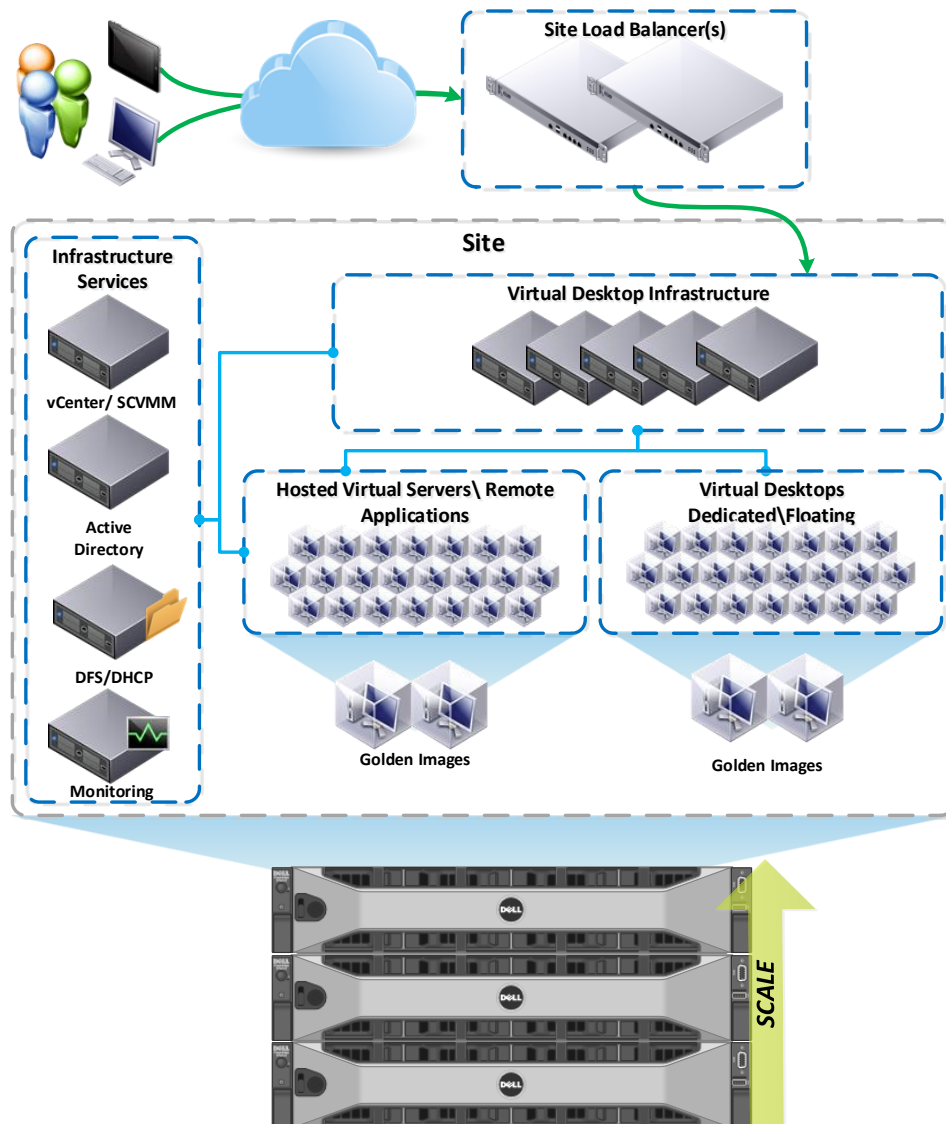
2.3 Dell XC Web-Scale – Solution Pods

The networking layer consists of the 10Gb Dell Networking S4810 utilized to build a world-class leaf/spine architecture with robust 1Gb switching in the S60 for iDRAC connectivity.
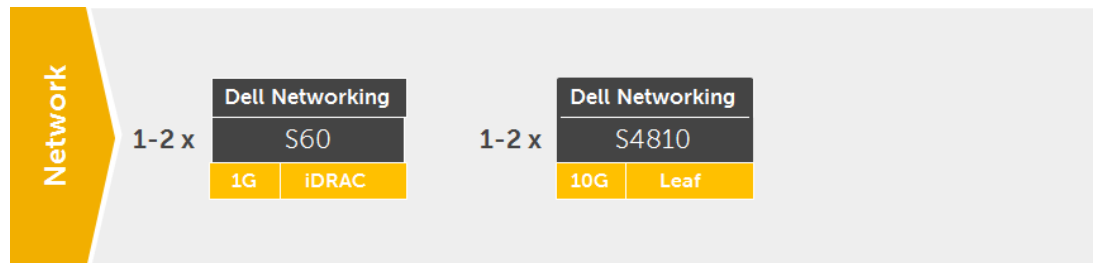
The compute, management and storage layers are 'converged' into a single XC appliance, hosting VMware vSphere hypervisor. A minimum of three nodes per cluster is required. The recommended boundaries of an individual pod follow two schools of thought: hypervisor-based or VDI solution-based. Limits can be established based on the number of nodes supported within a given hypervisor cluster-in this case, 32 nodes for vSphere. A single Nutanix NDFS cluster can span an unlimited number of appliance nodes with several hypervisor clusters contained within.

Dell recommends that the VDI management infrastructure nodes be separated from the compute resources onto their own appliance with a common storage namespace shared between them based on NFS for vSphere. One node for VDI management is required, minimally, and expanded based on size of the pod. The designations ds_rdsh, ds_compute, and ds_mgmt are logical containers used to group VMs of a particular type. Compute hosts can be used for RDSH. Distinct clusters should be built for management and compute host types for HA, respectively, to plan predictable scale and load across the environment. The NFS or SMB namespace can be shared across the two clusters adding disk capacity and performance for each distinct cluster.
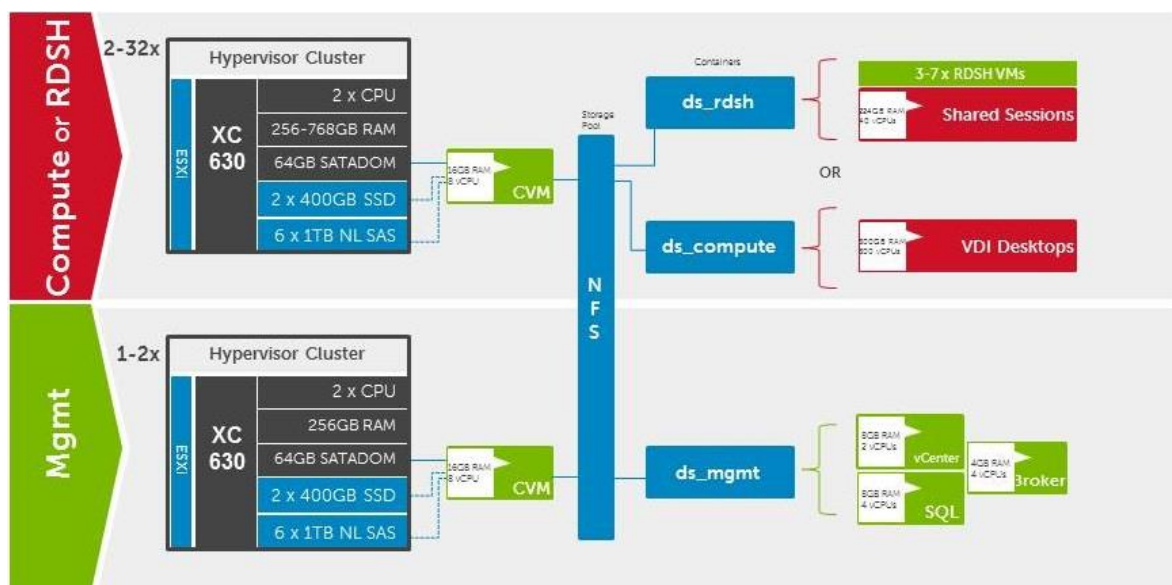
2.3.1 Network Architecture

Designed for true linear scaling, Dell XC series leverages 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.

The following figure shows a design of a scale-out Leaf-Spine network architecture that provides 20Gb active throughput from each node to its Leaf and scalable 80Gb active throughput from each Leaf to Spine switch. This provides scale from 3 XC nodes to thousands without any impact to available bandwidth:
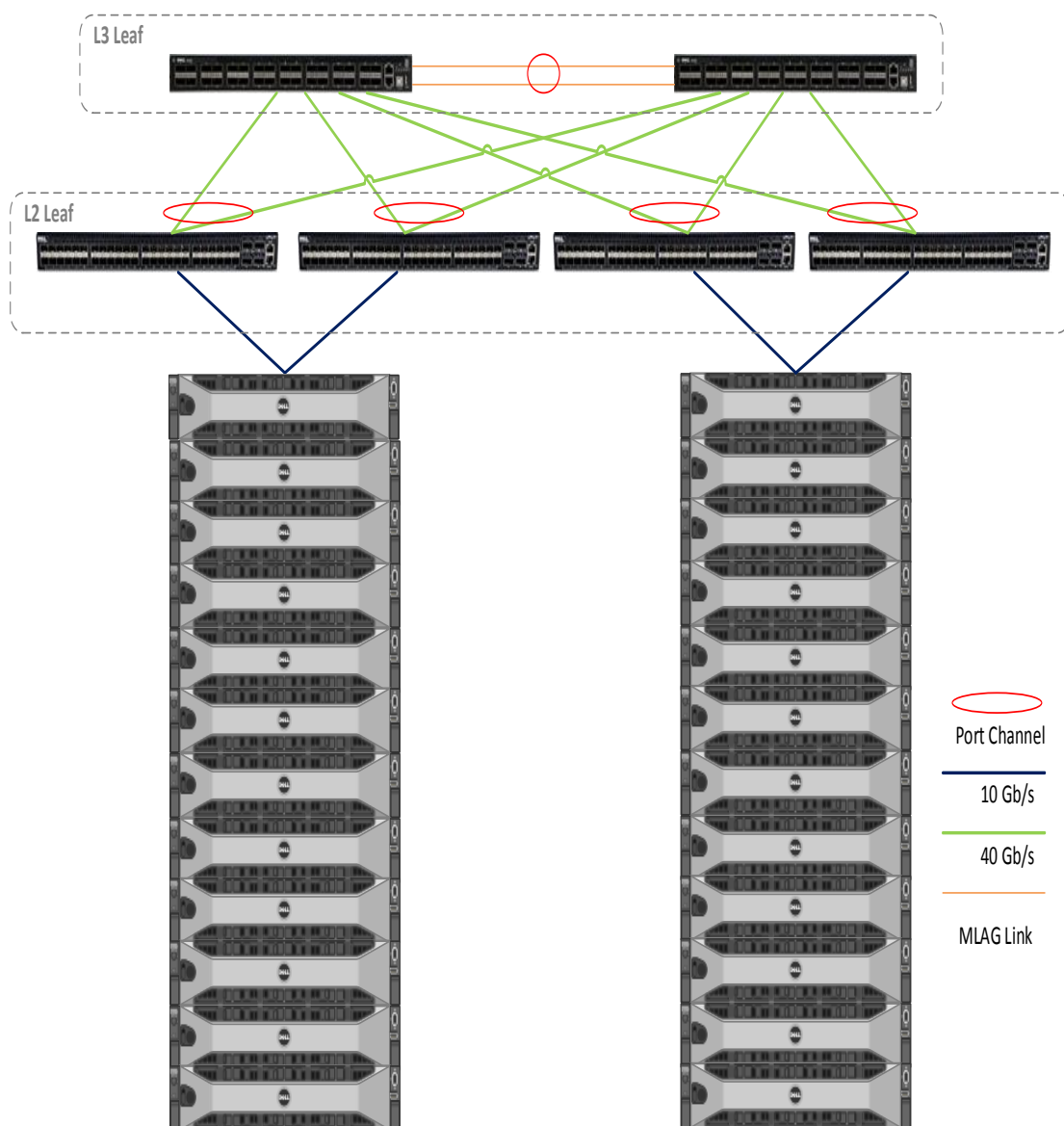
3 Hardware Components

3.1 Network

The following sections contain the core network components for the Dell XC Web-Scale Converged Appliance for VMware View. General uplink cabling guidance to consider in all cases is that TwinAx is very cost effective for short 10Gb runs and for longer runs use fiber with SFPs.

3.1.1 Dell Networking S60 (1Gb ToR Switch)

The Dell Networking S60 is a high-performance 1Gb access switch optimized for lowering operational costs at the network edge and is recommended for iDRAC connectivity. The S60 answers the key challenges related to network congestion in data center ToR (Top-of-Rack) and service provider aggregation deployments. As the use of bursty applications and services continue to increase, huge spikes in network traffic that can cause network congestion and packet loss, also become more common. The S60 is equipped with the industry's largest packet buffer (1.25 GB), enabling it to deliver lower application latency and maintain predictable network performance even when faced with significant spikes in network traffic. Providing 48 line-rate Gb ports and up to four optional 10Gb uplinks in just 1-RU, the S60 conserves valuable rack space. Further, the S60 design delivers unmatched configuration flexibility, high reliability, and power and cooling efficiency to reduce costs.

Model	Features	Options	Uses
Dell Networking S60	44 x BaseT (10/100/1000) + 4 x SFP High performance High Scalability	Redundant PSUs	1Gb connectivity for iDRAC
		4 x 1Gb SFP ports that support copper or fiber	
		12Gb or 24Gb stacking (up to 12 switches)	
		2 x modular slots for 10Gb uplinks or stacking modules	
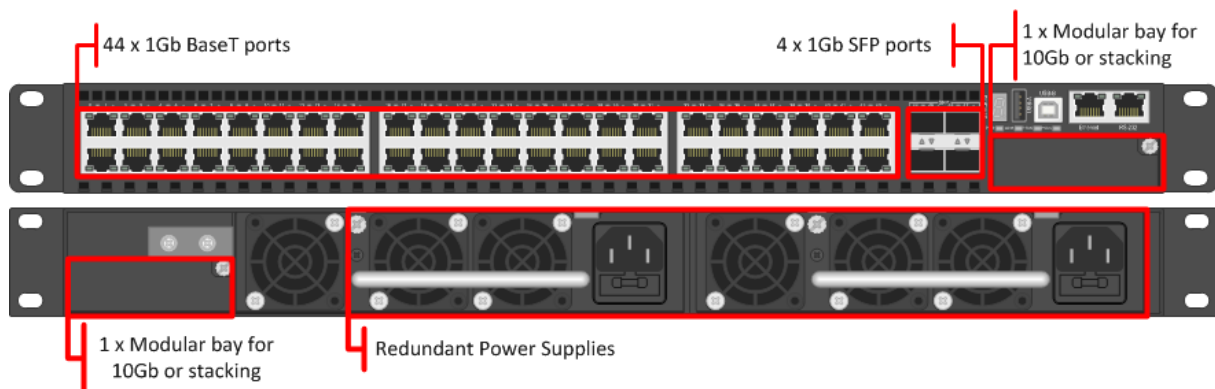
Guidance:

- 10Gb uplinks to a core or distribution switch are the preferred design choice using the rear 10Gb uplink modules. If 10Gb to a core or distribution switch is unavailable the front 4 x 1Gb SFP ports are used.

- The front 4 SFP ports can support copper cabling and are upgraded to optical if a longer run is needed.

- The S60 is appropriate for use in solutions scaling higher than 6000 users.

For more information on the S60 switch and Dell Networking products, please visit: LINK

3.1.1.1 S60 Stacking

The S60 switches are optionally stacked with 2 or more switches, if greater port count or redundancy is desired. Each switch will need a stacking module plugged into a rear bay and connected with a stacking cable. The best practice for switch stacks greater than 2 is to cable in a ring configuration with the last switch in the stack cabled back to the first. Uplinks need to be configured on all switches in the stack back to the core to provide redundancy and failure protection.
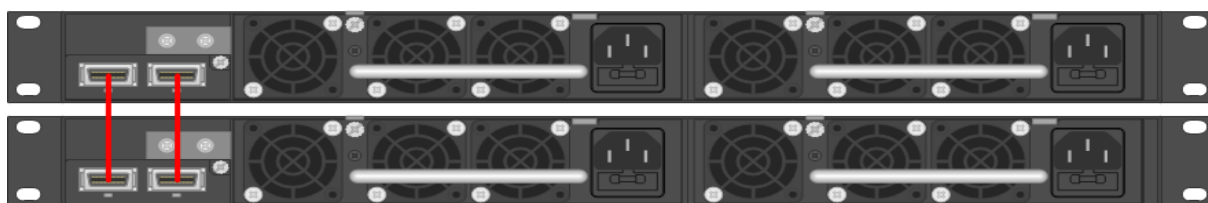
3.1.2 Dell Networking S4810 (10Gb ToR Switch)

The Dell Networking S-Series S4810 is an ultra-low latency 10/40Gb Top-of-Rack (ToR) switch purpose-built for applications in high-performance data center and computing environments. Leveraging a non-blocking, cut-through switching architecture, the S4810 delivers line-rate L2 and L3 forwarding capacity with ultra-low latency to maximize network performance. The compact S4810 design provides industry-leading density of 48 dual-speed 1/10Gb (SFP+) ports as well as four 40Gb QSFP+ uplinks to conserve valuable rack space and simplify the migration to 40Gb in the data center core (Each 40Gb QSFP+ uplink can support four 10Gb ports with a breakout cable). Priority-based Flow Control (PFC), Data Center Bridge Exchange (DCBX), Enhance Transmission Selection (ETS), coupled with ultra-low latency and line rate throughput, make the S4810 ideally suited for converged leaf/spine environments.

Model	Features	Options	Uses
Dell Networking S4810	48 x SFP+ (1Gb/10Gb) + 4 x QSFP+ (40Gb) Redundant Power Supplies	Single-mode/ multimode optics, TwinAx, QSFP+ breakout cables	ToR switch for 10Gb converged connectivity
		Stack up to 6 switches or 2 using VLT, using SFP or QSFP ports	

Guidance:

- The 40Gb QSFP+ ports are split into 4 x 10Gb ports using breakout cables for stand-alone units, if necessary. This is not supported in stacked configurations.

- 10Gb or 40Gb uplinks to a core or distribution switch is the preferred design choice.

- The front 4 SFP ports can support copper cabling and are upgraded to optical if a longer run is needed.

- The S60 is appropriate for use in solutions scaling higher than 6000 users.

For more information on the S4810 switch and Dell Networking, please visit: LINK
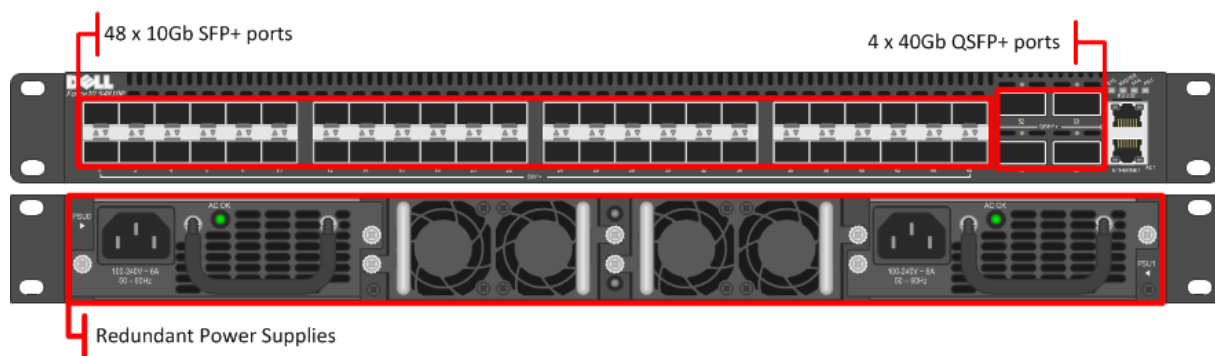
3.1.2.1   S4810 Stacking

The S4810 switches are optionally stacked up to 6 switches or configured to use Virtual Link Trunking (VLT) up to 2 switches. Stacking is supported on either SFP or QSFP ports as long as that port is configured for stacking. The best practice for switch stacks greater than 2 is to cable in a ring configuration with the last switch in the stack cabled back to the first. Uplinks need to be configured on all switches in the stack back to the core to provide redundancy and failure protection. It is recommended that the S4810 be configured for use in the leaf layer connective up to another switch in the spine layer.
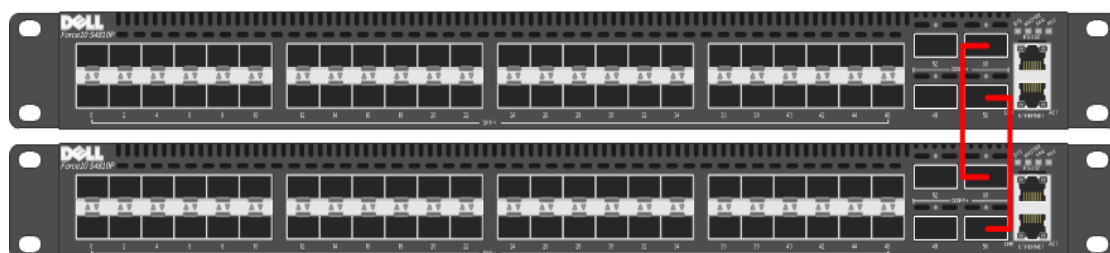
3.2 Appliance

The server platform for this appliance is the Dell XC630 (13G). This dual socket CPU platform runs the fastest Intel Xeon E5-2600v3 family of processors, can host up to 768GB RAM, and supports up to 10 x 2.5" SAS disks (two or four SSDs and four to eight HDDs). The Dell XC630 offers uncompromising performance and scalability in a 1U form factor. For more information please visit: Link
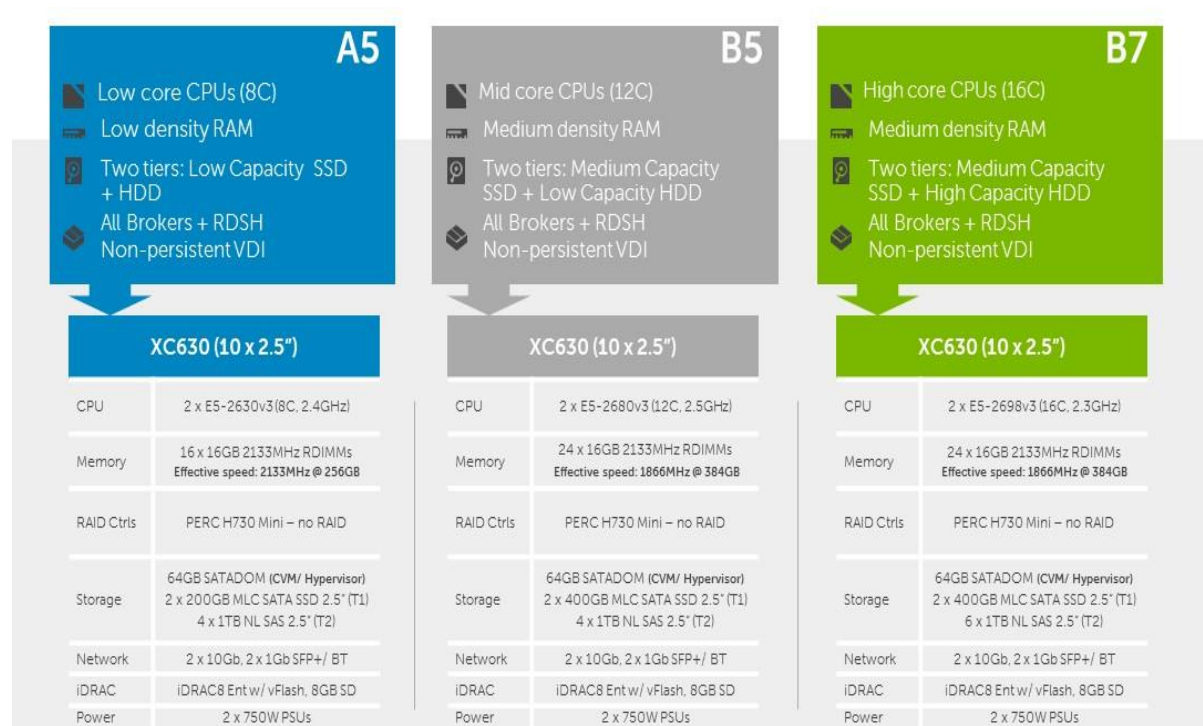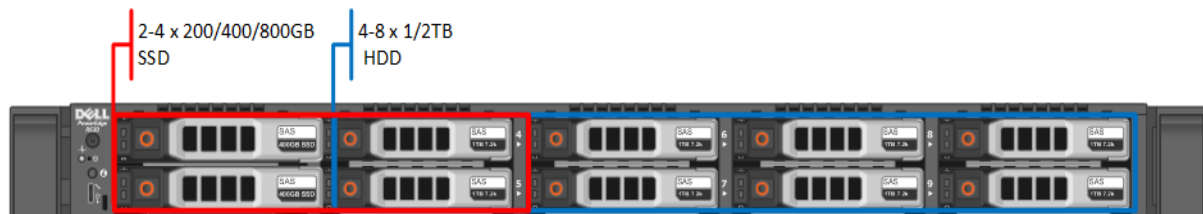
3.2.1 Dell XC A5

The Dell XC A5 platform is perfect for POCs, lighter user workloads, shared sessions or application virtualization. Each appliance comes equipped with dual 8-core CPUs and 256GB of high-performance DDR4 RAM. The 64GB SATADOM module hosts the hypervisor and Nutanix Controller VM while the H730 PERC is configured in pass-through mode to present the SSDs and HDDs to the NDFS cluster. A minimum of six disks come in each host, 2 x 200GB SSD for the performance tier (Tier1) and 4 x 1TB NL SAS disks for the capacity tier (Tier2), each tier can be expanded as required.

These six disks are presented to the Nutanix Controller VM running locally on each host. Each platform can be outfitting with SFP+ or BaseT NICs.

A5 — Dell XC630	
CPU	2 x E5-2630v3 (8C, 2.4GHz)
Memory	16 x 16GB 2133MHz RDIMMs **Effective speed: 2133MHz @ 256GB**
RAID Ctrls	PERC H730 Mini — no RAID
Storage	64GB SATADOM **(CVM/ HV)** 2 x 200GB MLC SATA SSD 2.5" (T1) 4 x 1TB NL SAS 2.5" (T2)
Network	2 x 10Gb, 2 x 1Gb SFP+/ BT
iDRAC	iDRAC8 Ent w/ vFlash, 8GB SD
Power	2 x 750W PSUs

3.2.2   Dell XC B5

The Dell XC B5 platform is perfect for larger POCs, medium user workloads, shared sessions or application virtualization. Each appliance comes equipped with dual 12-core CPUs and 384GB of high-performance RAM. The 64GB SATADOM module hosts the hypervisor and Nutanix Controller VM while the H730 PERC is configured in pass-through mode to present the SSDs and HDDs to the NDFS cluster. A minimum of six disks come in each host, 2 x 400GB SSD for the performance tier (Tier1) and 4 x 1TB NL SAS disks for the capacity tier (Tier2), each tier can be expanded as required. These six disks are presented to the Nutanix Controller VM running locally on each host. Each platform can be outfitting with SFP+ or BaseT NICs.

B5 — Dell XC630	
CPU	2 x E5-2680v3 (12C, 2.5GHz)
Memory	24 x 16GB 2133MHz RDIMMs **Effective speed: 1866 MHz @ 384GB**
RAID Ctrls	PERC H730 Mini — no RAID
Storage	64GB SATADOM **(CVM/ HV)** 2 x 400GB MLC SATA SSD 2.5" (T1) 4 x 1TB NL SAS 2.5" (T2)
Network	2 x 10Gb, 2 x 1Gb SFP+/ BT

iDRAC	iDRAC8 Ent w/ vFlash, 8GB SD
Power	2 x 750W PSUs

3.2.3   Dell XC B7

The Dell XC B7 platform is ideal for high performance requirements, heavy user workloads, shared sessions or application virtualization. Each appliance comes equipped with dual 16-core CPUs and 384GB of high-performance RAM. The 64GB SATADOM module hosts the hypervisor and Nutanix Controller VM while the H730 PERC is configured in pass-through mode to present the SSDs and HDDs to the NDFS cluster. A minimum of six disks come in each host, 2 x 400GB SSD for the performance tier (Tier1) and 6 x 1TB NL SAS disks for the capacity tier (Tier2), each tier can be expanded as required. These six disks are presented to the Nutanix Controller VM running locally on each host. Each platform can be outfitting with SFP+ or BaseT NICs.

B7 – Dell XC630	
CPU	2 x E5-2698v3 (16C, 2.3GHz)
Memory	24 x 16GB 2133MHz RDIMMs **Effective speed: 1866 MHz @ 384GB**
RAID Ctrls	PERC H730 Mini – no RAID
Storage	64GB SATADOM **(CVM/ HV)** 2 x 400GB MLC SATA SSD 2.5" (T1) 6 x 1TB NL SAS 2.5" (T2)
Network	2 x 10Gb, 2 x 1Gb SFP+/ BT
iDRAC	iDRAC8 Ent w/ vFlash, 8GB SD
Power	2 x 750W PSUs

3.3   Dell Wyse Cloud Clients

The following Dell Wyse clients will deliver a superior VMware user experience and are the recommended choices for this solution.

Dell Wyse P25

Uncompromising computing with the benefits of secure, centralized management. The Dell Wyse P25 PCoIP zero client for VMware View is a secure, easily managed zero client that provides outstanding graphics performance for advanced applications such as CAD, 3D solids modeling, video editing and advanced worker-level office productivity applications. Smaller than a typical notebook, this dedicated zero client is designed specifically for VMware View. It features the latest processor technology from Teradici to process the PCoIP protocol in silicon and includes client-side content caching to deliver the highest level of

performance available over 2 HD displays in an extremely compact, energy-efficient form factor. The Dell Wyse P25 delivers a rich user experience while resolving the challenges of provisioning, managing, maintaining and securing enterprise desktops.

Dell Wyse D10DP

The Dell Wyse D10DP is a high-performance and secure ThinOS 8 thin client that is absolutely virus and malware immune. Combining the performance of a dual core AMD G-Series APU with an integrated graphics engine and ThinOS, the D10DP offers exceptional thin client PCoIP processing performance for VMware Horizon View environments that handles demanding multimedia apps with ease and delivers brilliant graphics. Powerful, compact and extremely energy efficient, the D10DP is a great VDI end point for organizations that need high-end performance but face potential budget limitations.

Dell Wyse P45

Uncompromising computing with the benefits of secure, centralized management. The Dell Wyse P45 PCoIP zero client for VMware View is a secure, easily managed zero client that provides outstanding graphics performance for advanced applications such as CAD, 3D solids modeling, video editing and advanced worker-level office productivity applications. About the size of a notebook, this dedicated zero client designed specifically for VMware View. It features the latest processor technology from Teradici to process the PCoIP protocol in silicon and includes client-side content caching to deliver the highest level of display performance available over 4 HD displays in a compact, energy-efficient form factor. The Dell Wyse P45 delivers a rich user experience while resolving the challenges of provisioning, managing, maintaining and securing enterprise desktops.

Dell Wyse Z50D

Designed for power users, the Dell Wyse X50D is the highest performing thin client on the market. Highly secure and ultra-powerful, the X50D combines Dell Wyse-enhanced SUSE Linux Enterprise with dual-core AMD 1.65 GHz processor and a revolutionary unified engine for an unprecedented user experience. The Z50D eliminates performance constraints for high-end, processing-intensive applications like computer-aided design, multimedia, HD video and 3D modelling.

Dell Wyse Z90D

This is super high performance Windows Embedded Standard 7 thin client for virtual desktop environments. Featuring a dual core AMD

processor and a revolutionary unified engine that eliminates performance constraints, the Z90D7 achieves incredible speed and power for the most demanding embedded windows applications, rich graphics and HD video. With touch screen capable displays, the Z90D7 adds the ease of an intuitive multi touch user experience and is an ideal thin client for the most demanding virtual desktop workload applications.

Dell Chromebook 11

The lightweight, easy-to-use Dell Chromebook 11 helps turn education into exploration - without the worries of safety or security. Priced to make 1:1 computing affordable today, Chromebook 11 is backed by Dell support services to make the most of your budget for years to come. The Chrome OS and Chrome browser get students online in an instant and loads web pages in seconds. A high-density battery supported by a 4th Gen Intel® processor provides up to 10 hours of power. Encourage creativity with the Chromebook 11 and its multimedia features that include an 11.6" screen, stereo sound and webcam. Finally, with a fully compliant HTML5 browser; the Dell Chromebook11 is an excellent choice as an endpoint to a HTML5/BLAST connects Horizon View VDI desktop.

4 Software Components

4.1 VMware Horizon View

The solution is based on VMware Horizon View which provides a complete end-to-end solution delivering Microsoft Windows virtual desktops to users on a wide variety of endpoint devices. Virtual desktops are dynamically assembled on demand, providing users with pristine, yet personalized, desktops each time they log on.

VMware Horizon View provides a complete virtual desktop delivery system by integrating several distributed components with advanced configuration tools that simplify the creation and real-time management of the virtual desktop infrastructure. For the complete set of details, please see the Horizon View resources page at http://www.vmware.com/products/horizon-view/resources.html

The core Horizon View components include:

View Connection Server (VCS) – Installed on servers in the data center and brokers client connections, The VCS authenticates users, entitles users by mapping them to desktops and/or pools, establishes secure connections from clients to desktops, support single sign-on, sets and applies policies, acts as a DMZ security server for outside corporate firewall connections and more.

View Client – Installed on endpoints. Is software for creating connections to View desktops that can be run from tablets, Windows, Linux, or Mac PCs or laptops, thin clients and other devices.

View Portal – A web portal to access links for downloading full View clients. With HTML Access Feature enabled enablement for running a View desktop inside a supported browser is enabled.

View Agent – Installed on all VMs, physical machines and Terminal Service servers that are used as a source for View desktops. On VMs the agent is used to communicate with the View client to provide services such as USB redirection, printer support and more.

View Administrator – A web portal that provides admin functions such as deploy and management of View desktops and pools, set and control user authentication and more.

View Composer – This software service can be installed standalone or on the vCenter server and provides enablement to deploy and create linked clone desktop pools (also called non-persistent desktops).

vCenter Server – This is a server that provides centralized management and configuration to entire virtual desktop and host infrastructure. It facilitates configuration, provision, management services. It is installed on a Windows Server 2008 host (can be a VM).

View Transfer Server – Manages data transfers between the data center and the View desktops that are checked out on the end users' desktops in offline mode. This Server is required to support desktops that run the View client with Local Mode options. Replications and syncing are the functions it will perform with offline images.

4.1.1 NUMA Architecture Considerations

Best practices and testing has showed that aligning RDSH design to the physical Non-Uniform Memory Access (NUMA) architecture of the server CPUs results in increased and optimal performance. NUMA ensures that a CPU can access its own directly-connected RAM banks faster than those banks of the other processor which are accessed via the Quick Path Interconnect (QPI). Ensuring that your virtual RDSH servers do not span physical CPUs will ensure the greatest possible performance benefit.

The general guidance for RDSH on the Dell XC appliance is as follows:

4.1.1.1 A5

Eight physical cores per CPU in the A5 platform, assuming 2x oversubscription per core, gives us a total of 32 consumable cores per appliance. The Nutanix CVM will receive its vCPU allotment from the first physical CPU and so configuring the RDSH VMs as shown below will ensure that no NUMA spanning occurs which could lower performance.
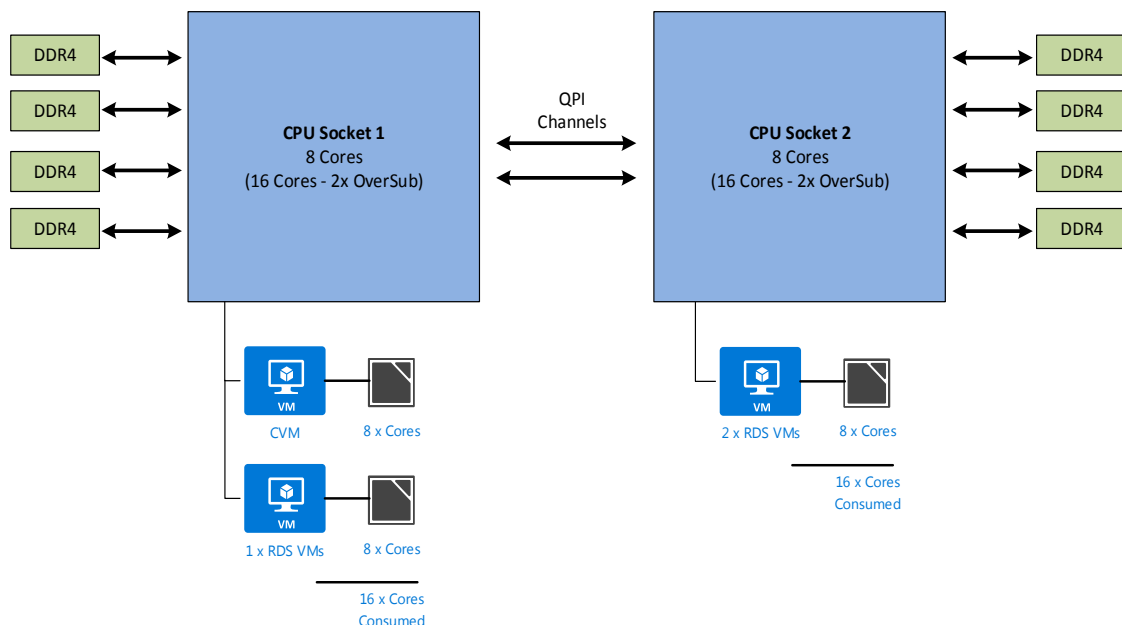
4.1.1.2 B5

12 physical cores per CPU in the B5 platform, assuming 2x oversubscription per core, nets a total of 48 consumable cores per appliance. The Nutanix CVM receives its vCPU allotment from the first physical CPU so configuring the RDSH VMs as shown below will ensure that no NUMA spanning occurs which can lower performance.

4.1.1.3 B7

16 physical cores per CPU in the B7 platform, assuming 2x oversubscription per core, nets a total of 64 consumable cores per appliance. The Nutanix CVM receives its vCPU allotment from the first physical CPU so configuring the RDSH VMs as shown below will ensure that no NUMA spanning occurs which can lower performance.

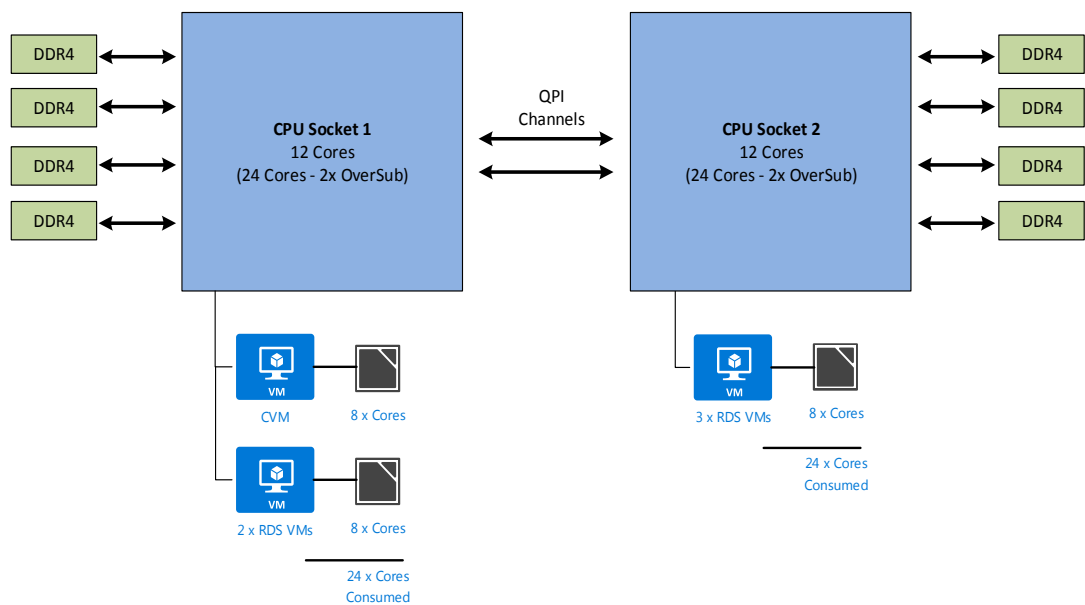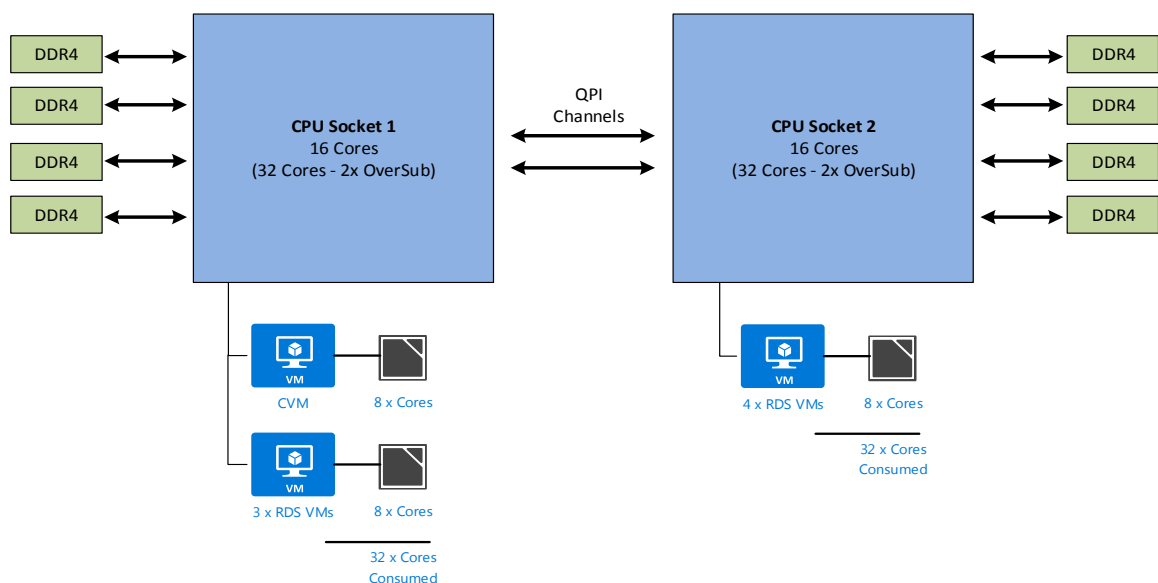## 4.2   VDI hypervisor platform

4.2.1   VMware vSphere 5.5

VMware vSphere 5.5 (currently vSphere 5.5 U2) is a virtualization platform used for building VDI and cloud infrastructures. vSphere 5.5 represents a migration from the ESX architecture to the ESXi architecture.

VMware vSphere 5.5 includes three major layers: Virtualization, Management and Interface. The Virtualization layer includes infrastructure and application services. The Management layer is central for configuring, provisioning and managing virtualized environments. The Interface layer includes the vSphere client and the vSphere web client.

Throughout the Dell Wyse Datacenter solution, all VMware best practices and prerequisites are adhered to (NTP, DNS, Active Directory, etc.). The vCenter 5.5 VM used in the solution will be a single Windows Server 2012 R2 VM (Check for current Windows Server OS compatibility at: http://www.vmware.com/resources/compatibility), residing on a host in the management tier. SQL server is a core component of vCenter and will be hosted on another VM also residing in the management tier. All additional Horizon View components need to be installed in a distributed architecture, 1 role per VM.

For more information on VMware vSphere, visit http://www.vmware.com/products/vsphere

5 Solution Architecture for View

5.1 Management Infrastructure

5.1.1 vSphere

The Management role requirements for the vSphere-based solution are summarized below.

Role	vCPU	vRAM (GB)	NIC	OS vDisk	
				Size (GB)	Location
Nutanix CVM	8*	16	2	-	C:\ (SATADOM)
Connection Server	4	8	1	40	SDS: ds_mgmt
Primary SQL	4	8	1	40 + 200	SDS: ds_mgmt
vCenter Appliance	2	8	1	125	SDS: ds_mgmt
Total	**18 vCPUs**	**40GB**	**5 vNICs**	**405GB**	**-**

5.1.2 SQL Databases

The VMware databases will be hosted by a single dedicated SQL 2012 SP1 Server VM (check DB compatibility at: http://partnerweb.vmware.com/comp_guide2/sim/interop_matrix.php?) in the Management layer. Use caution during database setup to ensure that SQL data, logs and TempDB are properly separated onto their respective volumes. Create all Databases that will be required for:

- View Connection Server
- vCenter

Initial placement of all databases into a single SQL instance is fine unless performance becomes an issue, in which case database need to be separated into separate named instances. Enable auto-growth for each DB.

Best practices defined by VMware are to be adhered to, to ensure optimal database performance.

Align all disks to be used by SQL Server with a 1024K offset and then formatted with a 64K file allocation unit size (data, logs and TempDB).

5.1.3 DNS

DNS plays a crucial role in the environment not only as the basis for Active Directory but will be used to control access to the various VMware software components. All hosts, VMs and consumable software components need to have a presence in DNS, preferably via a dynamic and AD-integrated namespace. Microsoft best practices and organizational requirements are to be adhered to.

Pay consideration for eventual scaling, access to components that may live on one or more servers (SQL databases, VMware services) during the initial deployment. Use CNAMEs and the round robin DNS mechanism to provide a front-end "mask" to the back-end server actually hosting the service or data source.

5.1.3.1 DNS for SQL

To access the SQL data sources, either directly or via ODBC, a connection to the server name\ instance name must be used. To simplify this process, as well as protect for future scaling (HA), instead of connecting to server names directly, alias these connections in the form of DNS CNAMEs. So instead of connecting to SQLServer1\<instance name> for every device that needs access to SQL, the preferred approach would be to connect to <CNAME>\<instance name>.

For example, the CNAME "VDISQL" is created to point to SQLServer1. If a failure scenario was to occur and SQLServer2 would need to start serving data, we would simply change the CNAME in DNS to point to SQLServer2. No infrastructure SQL client connections would need to be touched.

SQLServer1	Host (A)	10.1.1.28
SQLServer2	Host (A)	10.1.1.29
SQLVDI	Alias (CNAME)	SQLServer1.fcs.local

5.2 Storage Architecture Overview

All Dell XC Web Scale appliances come with two tiers of storage by default, SSD for performance and HDD for capacity. A single common software defined namespace is created across the Nutanix cluster and presented as either NFS or SMB to the hypervisor of each host. This constitutes a storage pool and one should be sufficient per cluster. Within this common namespace, logical containers are created to group VM files as well as control the specific storage-related features that are desired to be enabled such as deduplication and compression.

5.2.1 Containers

The following table outlines the recommended containers, their purpose and settings given the use case. Best practices suggest using as few features as possible, only enable what is absolutely required. For example, if you are not experiencing disk capacity pressure then there is no need to enable Capacity Tier Deduplication. Enabling unnecessary services increases the CPU demand on the Controller VMs.

| Container	Purpose	Replication Factor	Perf Tier Deduplication	Capacity Tier Deduplication	Compression
Ds_compute	Desktop VMs	2	Enabled	Disabled	Disabled
Ds_mgmt	Mgmt Infra VMs	2	Enabled	Disabled	Disabled
Ds_rdsh	RDSH Server VMs	2	Enabled	Disabled	Disabled

5.3 Networking

5.3.1 vSphere

The network configuration for the Dell XC Web Scale appliances utilizes a 10Gb converged infrastructure model. All required VLANs will traverse 2 x 10Gb NICs configured in an active/ active team. For larger scaling it is recommended to separate the infrastructure management VMs from the compute VMs to aid in predictable compute host scaling. The following outlines the VLAN requirements for the Compute and Management hosts in this solution model:

- Compute hosts
  - Management VLAN: Configured for hypervisor infrastructure traffic – L3 routed via spine layer
  - Live Migration VLAN: Configured for Live Migration traffic – L2 switched via leaf layer
  - VDI VLAN: Configured for VDI session traffic – L3 routed via spine layer
- Management hosts
  - Management VLAN: Configured for hypervisor Management traffic – L3 routed via spine layer
  - Live Migration VLAN: Configured for Live Migration traffic – L2 switched via leaf layer
  - VDI Management VLAN: Configured for VDI infrastructure traffic – L3 routed via spine layer
- An iDRAC VLAN is configured for all hardware management traffic – L3 routed via spine layer

5.3.1.1 vSphere

The Management host network configuration consists of a standard vSwitch teamed with 2 x 10Gb physical adapters assigned. The CVM connects to a private internal vSwitch as well as the standard external vSwitch. All VMkernel service ports connect to the standard external vSwitch. All VDI infrastructure VMs connect through the primary port group on the external vSwitch.
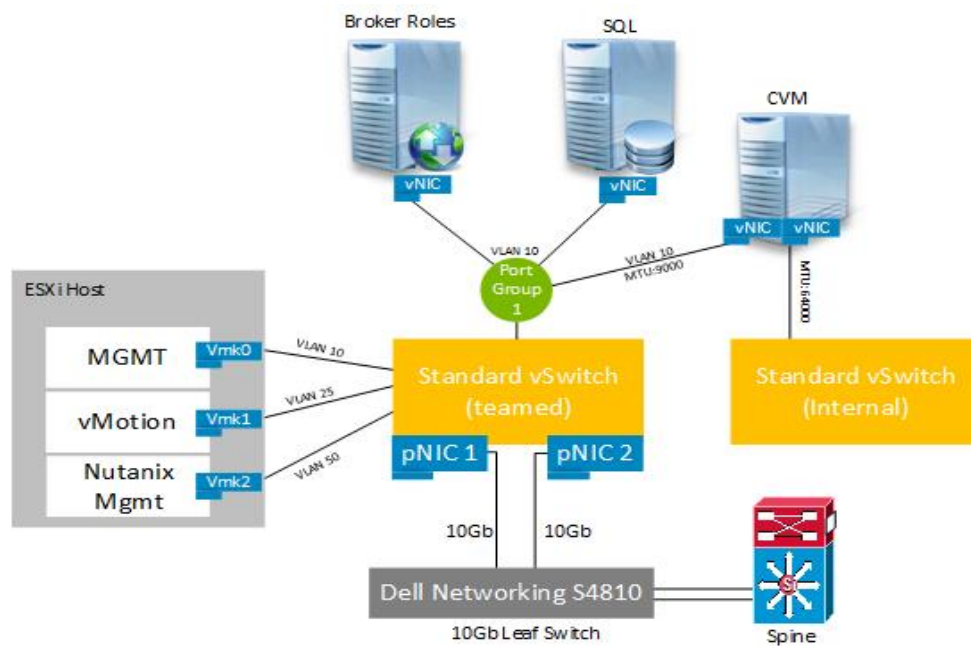
5.4 Scaling Guidance

Each component of the solution architecture scales independently according to the desired number of supported users. Using the new Intel Ivy Bridge CPUs, rack and blade servers now scale equally from a compute perspective.

- The components can be scaled either horizontally (by adding additional physical and virtual servers to the server pools) or vertically (by adding virtual resources to the infrastructure)
- Eliminate bandwidth and performance bottlenecks as much as possible
- Allow future horizontal and vertical scaling with the objective of reducing the future cost of ownership of the infrastructure.

Component	Metric	Horizontal Scalability	Vertical Scalability
Virtual Desktop Host/Compute Servers	VMs per physical host	Additional hosts and clusters added as necessary	Additional RAM or CPU compute power
View Composer	Desktops per instance	Additional physical servers added to the Management cluster to deal with additional management VMs.	Additional RAM or CPU compute power
View Connection Servers	Desktops per instance	Additional physical servers added to the Management cluster to deal with additional management VMs.	Additional VCS Management VMs.
VMware vCenter	VMs per physical host and/or ESX hosts per vCenter instance	Deploy additional servers and use linked mode to optimize management	Additional vCenter Management VMs.
Database Services	Concurrent connections, responsiveness of reads/ writes	Migrate databases to a dedicated SQL server and increase the number of management nodes	Additional RAM and CPU for the management nodes
File Services	Concurrent connections, responsiveness of reads/ writes	Split user profiles and home directories between multiple file servers in the cluster. File services can also be migrated to the optional NAS device to provide high availability.	Additional RAM and CPU for the management nodes

5.5   Solution High Availability

High availability (HA) is offered to protect each layers of the solution architecture, individually if desired. Following the N+1 model, additional ToR switches for LAN are added to the Network layer and stacked to provide redundancy as required, additional compute and management hosts are added to their respective layers, vSphere clustering is introduced in the management layer, SQL is mirrored or clustered, an F5 device can be leveraged for load balancing and a NAS device can be used to host file shares. Storage protocol switch stacks and NAS selection will vary based on chosen solution architecture.

The HA options provides redundancy for all critical components in the stack while improving the performance and efficiency of the solution as a whole.

- An additional switch is added at the network tier which will be configured with the original as a stack and equally spreading each host's network connections across both.

- At the compute tier an additional ESXi host is added to provide N+1 protection provided by vSphere for computer tier protection.
- A number of enhancements occur at the Management tier, the first of which is the addition of another host. The Management hosts will then be configured in an HA cluster. All applicable Horizon View server roles can then be duplicated on the new host where connections to each will be load balanced via the addition of a F5 Load Balancer. SQL will also receive greater protection through the addition and configuration of a SQL mirror with a witness.

5.6 VMware Horizon View Communication Flow
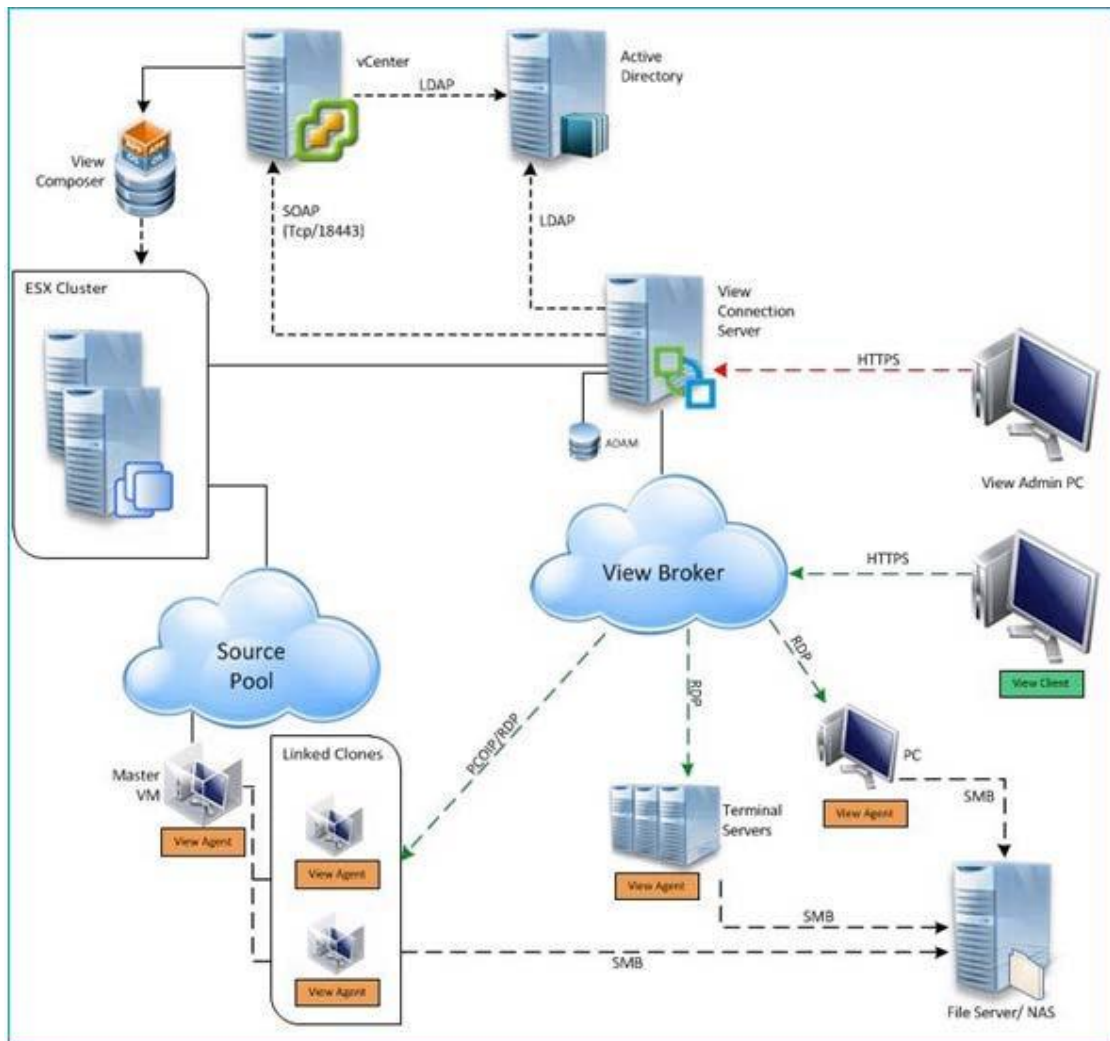
6 Solution Performance and Testing

6.1 Load Generation and Monitoring

### 6.1.1 Login VSI 4 – Login Consultants

Login VSI is the de-facto industry standard tool for testing VDI environments and server-based computing / terminal services environments. It uses launcher systems to connect a specified number of users to available desktops within the environment and then executes a standard collection of desktop application software (e.g. Microsoft Office, Adobe Acrobat Reader etc.) on each VDI desktop/session. Once the user is connected, a logon script configures the user environment and then starts the test workload. Each launcher system can launch connections to multiple 'target' machines (i.e. VDI desktops). A centralized management console is used to configure and manage the Login VSI environment and launcher systems.

6.1.2 VMware vCenter

VMware vCenter was used for VMware vSphere-based solutions to gather key data (CPU, Memory and Network usage) from each of the desktop hosts during each test run. This data was exported to .csv files for each host and then consolidated to show data from all hosts. While the report does not include specific performance metrics for the Management host servers, these servers were monitored during testing and were seen to be performing at an expected performance level.

6.2   Performance Analysis Methodology

In order to ensure the optimal combination of end-user experience (EUE) and cost-per-user, performance analysis and characterization (PAAC) on Dell Wyse Datacenter solutions is carried out using a carefully designed, holistic methodology that monitors both hardware resource utilization parameters and EUE during load-testing. This methodology is based on the three pillars shown below. Login VSI is currently the load-testing tool used during PAAC of Dell Wyse Datacenter solutions; Login VSI is the de-facto industry standard for VDI and server-based computing (SBC) environments and is discussed in more detail below.
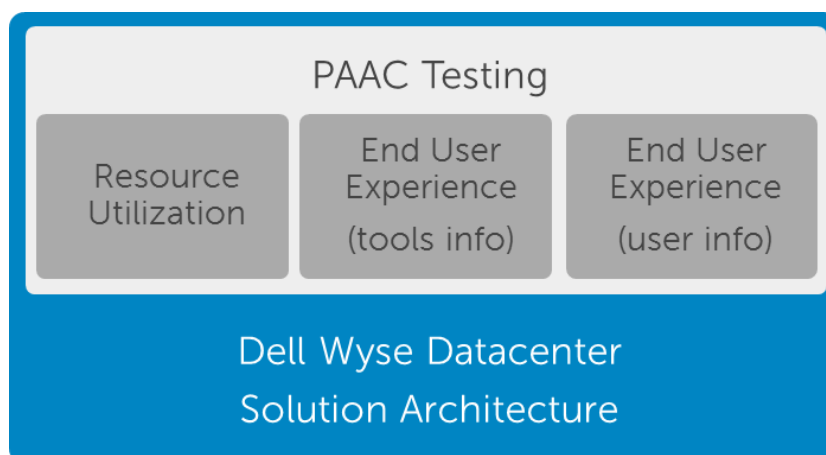
Each of the pillars shown in the above diagram is discussed in more detail below.

6.2.1   Resource Utilization

Poor end-user experience is one of the main risk factors when implementing desktop virtualization but the root cause for poor end-user experience is resource contention – hardware resources at some point in the solution have been exhausted, thus causing the poor end-user experience. In order to ensure that this has not happened (and that it is not close to happening), PAAC on Dell Wyse Datacenter solutions monitors the relevant resource utilization parameters and applies relatively conservative thresholds as shown in the table below. As discussed above, these thresholds are carefully selected to deliver an optimal combination of good end-user experience and cost-per-user, while also providing burst capacity for seasonal / intermittent spikes in usage. These thresholds are used to decide the number of virtual desktops (density) that are hosted by a specific hardware environment (i.e. combination of server, storage and networking) that forms the basis for a Dell Wyse Datacenter RA.

Resource Utilization Thresholds	
Parameter	Pass / Fail Threshold
Physical Host CPU Utilization	85%
Physical Host Memory Utilization	85%
Network Throughput	85%
Storage IO Latency	20ms

6.2.2   Dell Wyse Datacenter Workloads and Profiles

It's important to understand user workloads and profiles when designing a desktop virtualization solution in order to understand the density numbers that the solution can support. At Dell, we use five workload / profile levels, each of which is bound by specific metrics and capabilities. In addition, we use workloads and profiles that are targeted at graphics-intensive use cases. We present more detailed information in relation to these workloads and profiles below but first it is useful to define the terms "workload" and "profile" as they are used in this document.

- **Profile**: This is the configuration of the virtual desktop - number of vCPUs and amount of RAM configured on the desktop (i.e. available to the user).
- **Workload**: This is the set of applications used by performance analysis and characterization (PAAC) of Dell Wyse Datacenter solutions e.g. Microsoft Office applications, PDF Reader, Internet Explorer etc.

6.2.3   Dell Wyse Datacenter Profiles

The table shown below presents the persistent user profiles used during PAAC on this solution. These profiles have been carefully selected to provide the optimal level of resources for common use cases.

User Profile	vCPUs	ESXi Memory Reservation	ESXi memory configured	OS
Standard	1	2GB	2GB	x86
Enhanced	2	3GB	3GB	x86
Professional	2	4GB	4GB	x64

Standard user profile equates to a task worker, Enhanced to a knowledge worker and Professional to a power user.

6.2.4 Dell Wyse Datacenter Workloads

Load-testing on each of the profiles described in the above table is carried out using an appropriate workload that is representative of the relevant use case. In the case of the non-graphics workloads, these workloads are Login VSI workloads and in the case of graphics workloads, these are specially designed workloads that stress the VDI environment to a level that is appropriate for the relevant use case. This information is summarized in the table below:

| Profile Name	Workload	OS Image
Standard	Login VSI Light	Shared
Enhanced	Login VSI Medium	Shared
Professional	Login VSI Heavy	Shared + Profile Virtualization

As noted in the table above, further information for each of the workloads is given below. It is noted that for Login VSI testing, the following login and boot paradigm is used:

- For single-server / single-host testing (typically carried out to determine the virtual desktop capacity of a specific physical server), users are logged in every 30 seconds.
- For multi-host / full solution testing, users are logged in over a period of 1-hour, to replicate the normal login storm in an enterprise environment.
- All desktops are pre-booted in advance of logins commencing.

For all testing, all virtual desktops run an industry-standard anti-virus solution (currently McAfee VirusScan Enterprise) in order to fully replicate a customer environment.

6.2.4.1 Login VSI 3 versus Login VSI 4

PAAC on Dell Wyse Datacenter solutions is currently carried out using Login VSI version 4. However, some previous Dell Wyse Datacenter solutions used Login VSI version 3 for this PAAC work. Login VSI version 3 used a slightly different set of workloads to those used by Login VSI 4 and in order to allow comparison of results obtained using these different Login VSI versions, it is useful to be aware of the information presented in the figure below. This information (for Login VSI medium workload) illustrates higher CPU utilization and lower disk IO for Login VSI 4 (green / upper set of graphs) when compared to Login VSI3 (red / lower set of graphs). The exact variation between these Login VSI versions will vary between environments.
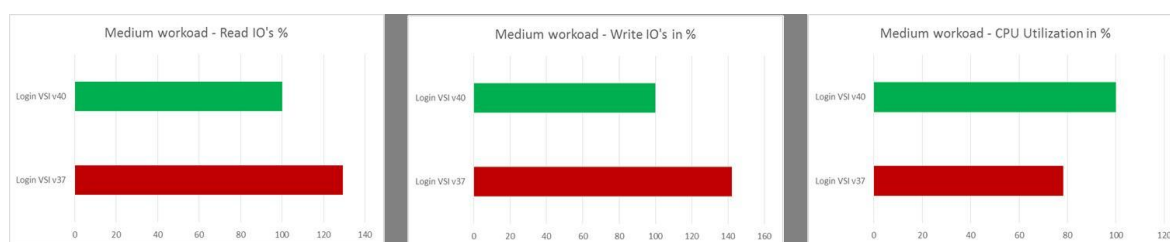
6.2.4.2 Login VSI Light Workload

Compared to the Login VSI medium workload described below, the light workload runs fewer applications (mainly Excel and Internet Explorer with some minimal Word activity) and starts/stops the applications less frequently. This results in lower CPU, memory and disk IO usage.

6.2.4.3 Login VSI Medium Workload

The Login VSI medium workload is designed to run on 2vCPU's per desktop VM. This workload emulates a medium knowledge worker using Office, IE, PDF and Java/FreeMind. The Login VSI medium workload has the following characteristics:

- Once a session has been started the workload will repeat (loop) every 48 minutes.
- The loop is divided in four segments; each consecutive Login VSI user logon will start a different segment. This ensures that all elements in the workload are equally used throughout the test.
- The medium workload opens up to five applications simultaneously.
- The keyboard type rate is 160 ms for each character.
- Approximately two minutes of idle time is included to simulate real-world users.

Each loop opens and uses:

- Outlook, browse messages.
- Internet Explorer, browse different webpages and a YouTube style video (480p movie trailer) is opened three times in every loop.
- Word, one instance to measure response time, one instance to review and edit a document.
- Doro PDF Printer & Acrobat Reader, the Word document is printed and exported to PDF.
- Excel, a very large randomized sheet is opened.
- PowerPoint, a presentation is reviewed and edited.
- FreeMind, a Java based Mind Mapping application.

6.2.4.4 Login VSI Heavy Workload

The heavy workload is based on the medium workload except that the heavy workload:

- Begins by opening four instances of Internet Explorer. These instances stay open throughout the workload loop.
- Begins by opening two instances of Adobe Reader. These instances stay open throughout the workload loop.
- There are more PDF printer actions in the workload.
- Instead of 480p videos a 720p and a 1080p video are watched.
- Increased the time the workload plays a flash game.
- The idle time is reduced to two minutes.

6.3 Testing and Validation

6.3.1 Testing Process

The purpose of the single server testing is to validate the architectural assumptions made around the server stack. Each user load is tested against four runs. First, a pilot run to validate that the infrastructure is functioning and valid data can be captured, and then, three subsequent runs allowing correlation of data.

At different stages of the testing the testing team will complete some manual "User Experience" Testing while the environment is under load. This will involve a team member logging into a session during the run and completing tasks similar to the User Workload description. While this experience will be subjective, it will help provide a better understanding of the end user experience of the desktop sessions, particularly under high load, and ensure that the data gathered is reliable.

For all workloads, the performance analysis scenario will be to launch a user session every 30 seconds.   Once all users have logged in, all will run workload activities at steady-state for 16 minutes and then logoffs will commence.

6.4 Horizon View Test Results

6.4.1 Configuration

Validation for this project was completed for VMWare View on the following platforms.

ESXi 5.5 U2 / VMware View 6.0/ XC630

ESXi 5.5 U2 / RDSH brokered by VMware View 6.0/ XC630

The platform config is shown below and the Login VSI workloads used for load testing on each environment.

Platform	Config	CPU	RAM	RAID Ctlr	HD Config	Network	Workload
XC630	B5	E5-2690v3 (12Core, 2.6GHz)	512GB	H730P	1 X 64GB SATADOM **(CVM/ HV)** 2 x 400GB, Intel S3700, SATA SSD's 2.5" (T1) 8 x 1TB NL SAS 2.5" (T2)	2 x 10Gb, 2 x 1Gb SFP+/ BT	Heavy

1GB networking was used for the deployment of the appliances while all 10GB networking was used for the PAAC testing.

Compute and Management resources were split out with the following configuration across a three node Nutanix cluster and all test runs were completed with this configuration.

- Node 1 – XC630– Dedicated Management (vCenter Appliance, SQL Server, VMware View Connection Server, VMware View Composer.)
- Node 2 – XC630– Dedicated Compute
- Node 3 –XC630 – Dedicated Compute

Please refer to Section 5.1 for the configuration of each management virtual machine.

The virtual machines were non-persistent linked clone desktops each configured on Windows 8.1 aligning with the Login VSI 4.X virtual machine configuration, Office2010 was used with each Virtual Machine sized at 32 GB. User Workload configuration of the load generation virtual machines is shown in the table below.

Workload	vCPUs	Memory	OS	HD Size
Professional User	2	4GB	x64	32 GB

The RDSH environment was configured as shown in the table below for each host according to platform. The RDSH images were built using Windows Server 2012 R2 Standard edition and Microsoft Office 2010 plus the support tools required for Login VSI testing.

Role	Config	VMs per host	vCPUs	RAM (GB)	NIC	Disk Size GB
RDSH VM	B5	5	8	32	1	80

6.4.1.1   Test Results

Validation was performed using CCC standard testing methodology using LoginVSI 4 load generation tool for VDI benchmarking that simulates user workloads.

Each test run adhered to PAAC best practices with a 30 second session logon interval, 16 minutes of steady state after which user sessions would begin logging off.

The following table summarizes the test results for the various workloads and configurations.

Hyper-visor	Provisioning	Workload	Density Per Host	Avg CPU	Avg Memory Consumed	Avg Memory Active	Avg IOPS/User	Avg Net /User
ESXi	View	Professional	80	98%	166 GB	42 GB	8.1	961 Kbps
ESXi	RDSH	Professional	120	98%	97 GB	21 GB	2.8	372 Kbps

CPU Utilization* CPU % for ESX Hosts was adjusted to account for the fact that on Intel E5-2690v3 series processors the ESX host CPU metrics will exceed the rated 100% for the host if Turbo Boost is enabled (by default). The Adjusted CPU % Usage is based on 100% usage and but is not reflected in the charts. The figure shown in the table is the Compute host steady state peak CPU Usage. One Nutanix CVM Controller virtual machine is located on each node and is allocated 10000MHz of CPU reservation.

Memory Utilization. The figure shown in the table above is the average memory **consumed** per Compute host over the recorded test period. Active is the average **active** memory per Compute host over the recorded test period.

The IOPS results are calculated from the average Nutanix **Cluster Disk IOPS** figure over the test period divided by the number of users.

Network Utilization The figure shown in the table is the average **Kbps/User per Compute host** over the recorded test period.

Summary of Tests Run

The 'Number of Users' column indicates the total number of user desktop test sessions run across two Compute hosts.

Config	Number of Users	Desktop Type	Login VSI Workload
B5	160	Linked Clone	Heavy
B5	240	RDSH	Heavy

6.4.1.2   Results summary

**B5 Configuration, ESXi/VMware View, 160 User Linked Clone Test, Heavy Workload**

The architecture for this solution includes a dedicated management host and two compute hosts. Each of the compute hosts was populated with 80 linked clone virtual machines and one Nutanix CVM per host.

The Nutanix CVM's took up approximately 15% of the compute hosts CPU Usage at the start of the test run.

This chart does not include the additional 10% of CPU available from the Turbo boost feature. Without the inclusion there is a total of 62380 MHz available for Desktops, with Turbo boost the total available MHz value is 68168 MHz.

The below graph shows the performance data for 80 user sessions per host on a pair of Compute hosts. The CPU reaches a steady state peak of 98% during the test cycle when approximately 80 users are logged on to each compute host.

The Management host in the cluster runs the vSphere management virtual machines and a Nutanix CVM virtual machine. Its CPU utilization is significantly lower than the compute hosts in the cluster. The CPU utilization for the management host does not exceed 16% at any point in the test cycle.
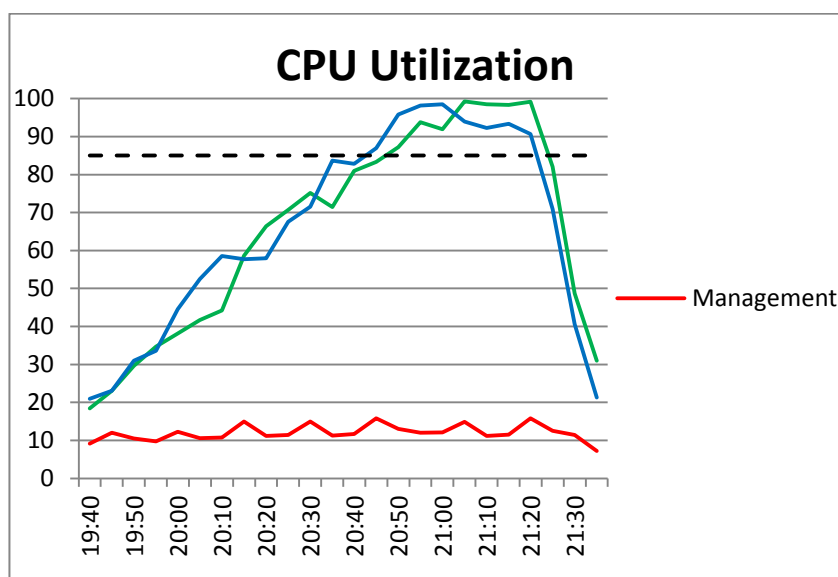
In regards to memory consumption for the cluster, out of a total of 512 GB available memory per node there were no constraints for any of the hosts. The Compute Hosts reached a max memory consumption of 208 GB with active memory usage reaching a max of 63GB. There was no memory ballooning or swapping.
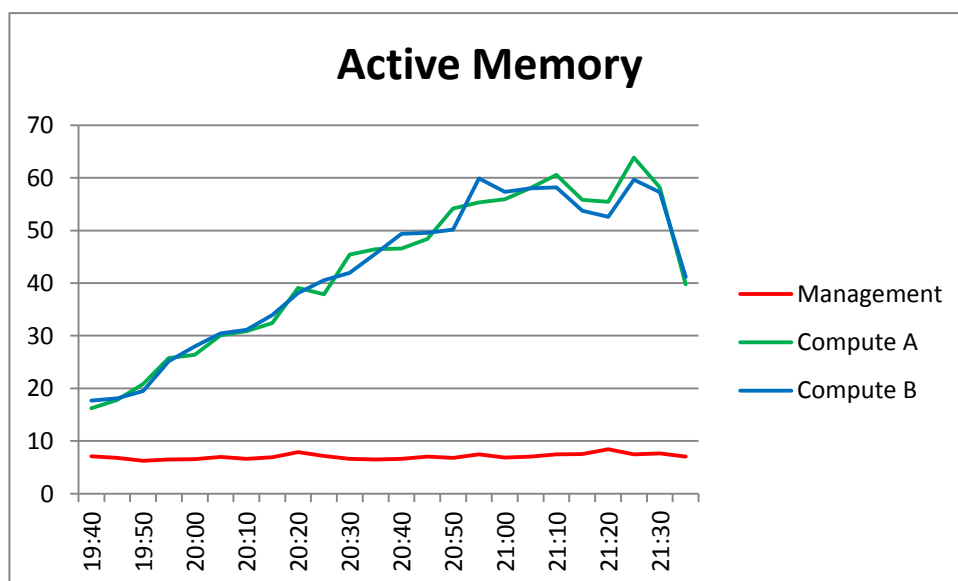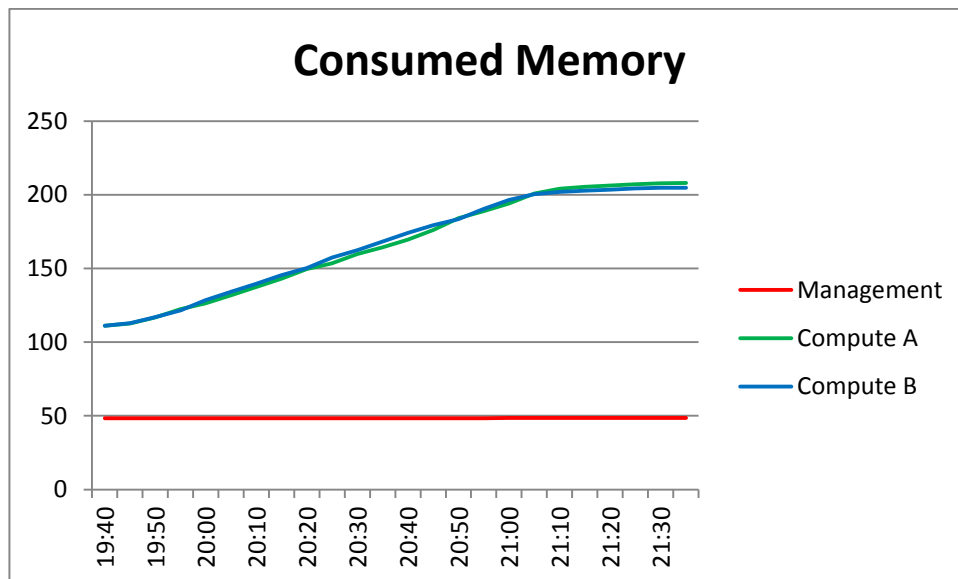
Consumed Memory

Active Memory

Network bandwidth is not an issue on this solution with a steady state peak of approximately 110,000 Kbps and a spike to 190,000 Kbps during the period of desktop session logoffs.
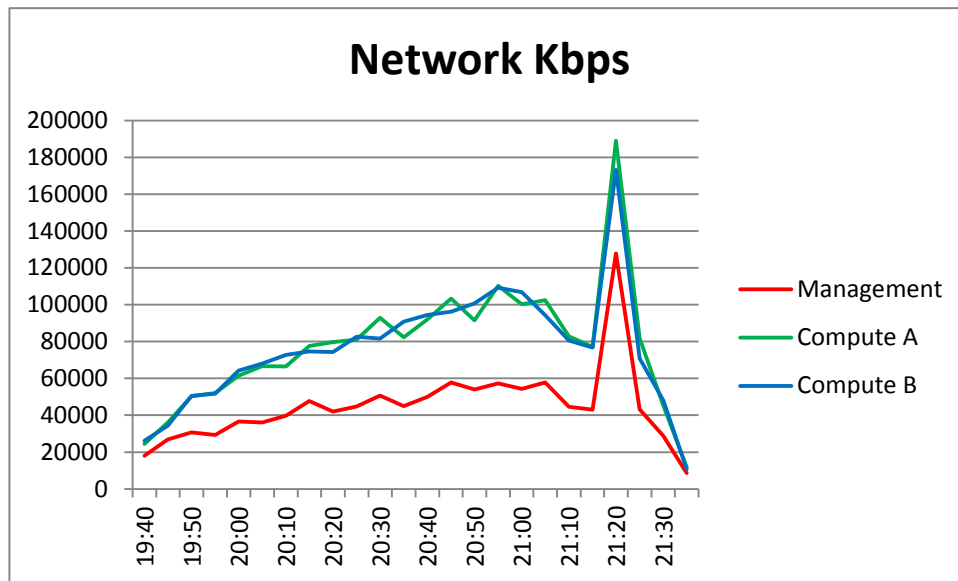
The Nutanix Cluster IOPS figures taken from the Nutanix Web Console showed that the cluster reached a maximum of 2,127 Disk IOPS during the test run with the Management node having the lowest disk IOPS of the three nodes. The activity spike in the graph after steady state peak shows the user logoffs taking place at the end of the test cycle.

The Login VSI Max user experience score for this test indicates that the VSI Max score not reached so there should have been little deterioration of user experience during testing.
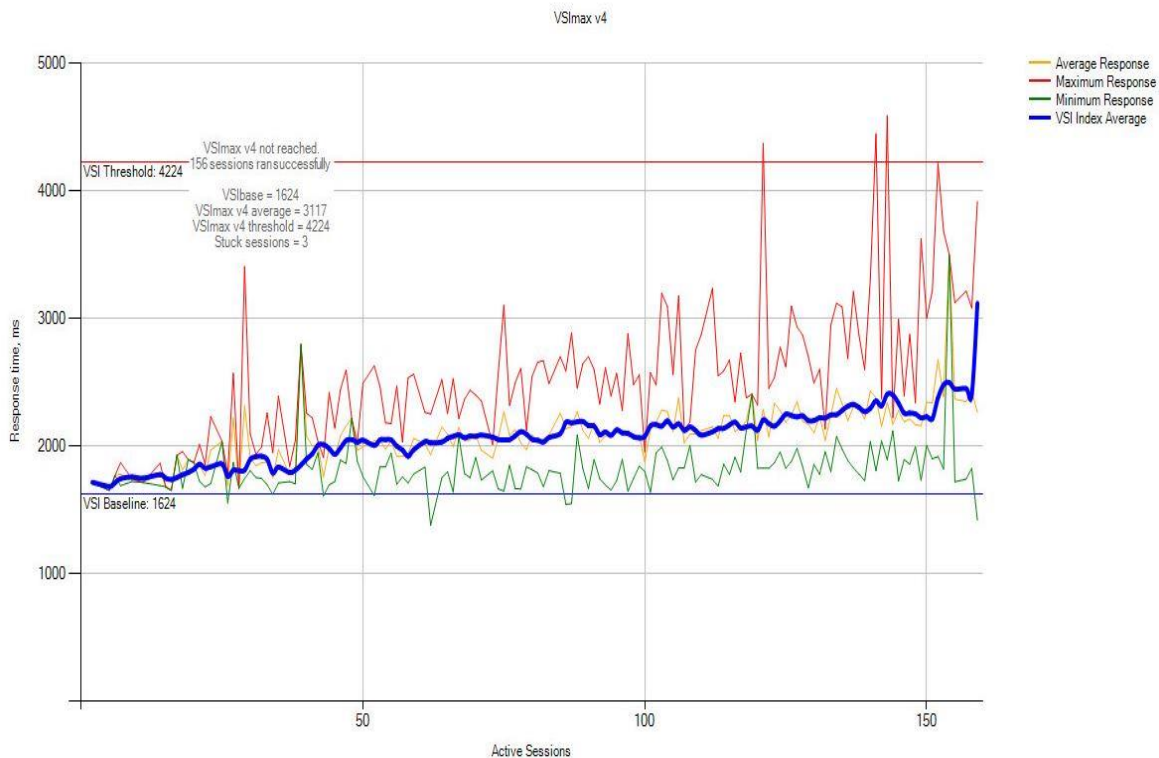
Notes:
- As indicated above, the CPU graphs do not take into account the extra 10% of CPU resources available through the 2690v3's turbo feature.
- Subjective user experience showed mouse movement and window response times when clicking within a running session during steady state were good. Video playback while somewhat slow was still acceptable.
- User login times were consistent up until the final few sessions were launching when a few session took some extra time to login.
- 512 GB of memory installed on each node is more than enough for this configuration and should run equally well with less memory installed.

### B5 Configuration, ESXi/VMware View, 240 User RDSH Test, Heavy Workload

The architecture for this solution includes a dedicated management host and two compute hosts. Each of the compute hosts were populated with 5 RDS session host virtual machines running Windows Server 2012 R2 and one Nutanix CVM per host. Each session host vm was allocated 32GB of RAM and 8vCPU's.

The Nutanix CVM's took up approximately 4% of the compute hosts CPU Usage at the start of the test run.

This chart does not include the additional 10% of CPU available from the Turbo boost feature. Without the inclusion there is a total of 62380 MHz available for Desktops, with Turbo boost the total available MHz value is 68168 MHz.

The below graph shows the performance data for 120 user sessions per host on a pair of Compute hosts. The CPU reaches a steady state peak of 98% during the test cycle when approximately 120 users are logged on to each compute host.

The Management host in the cluster runs the vSphere management virtual machines and a Nutanix CVM virtual machine. Its CPU utilization is significantly lower than the compute hosts in the cluster. The CPU utilization for the management host does not exceed 13% at any point in the test cycle.
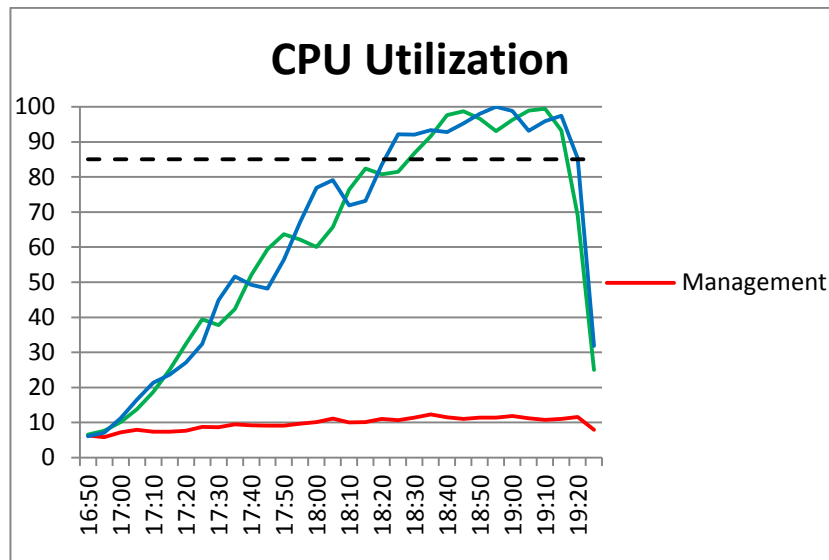

In regards to memory consumption for the cluster, out of a total of 512 GB available memory per node there were no constraints for any of the hosts. The Compute Hosts reached a max memory consumption of 106 GB with active memory usage reaching a max of 49GB. There was no memory ballooning or swapping.
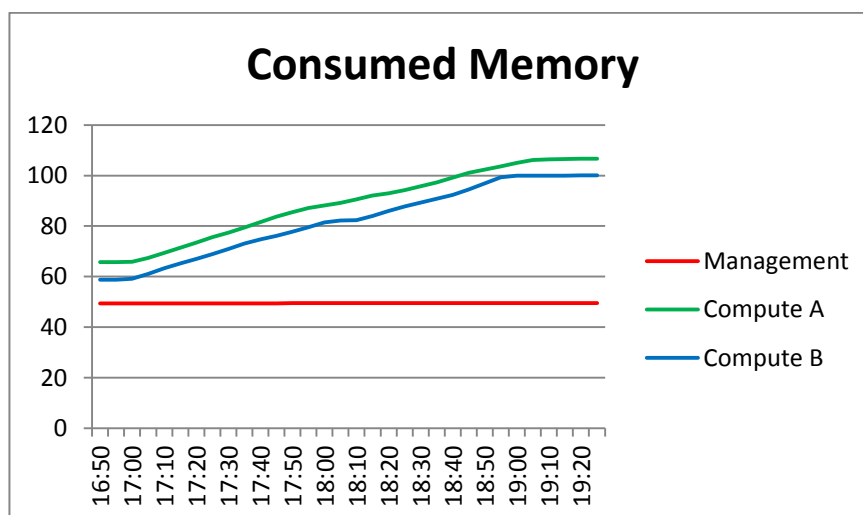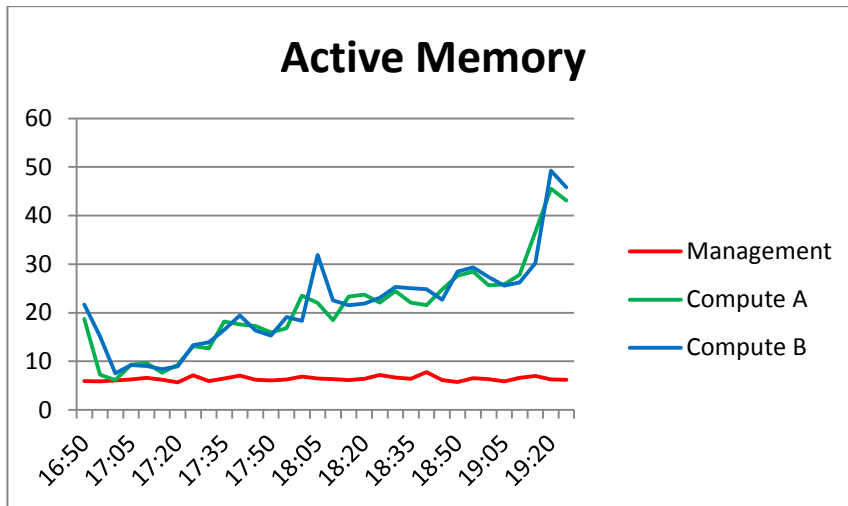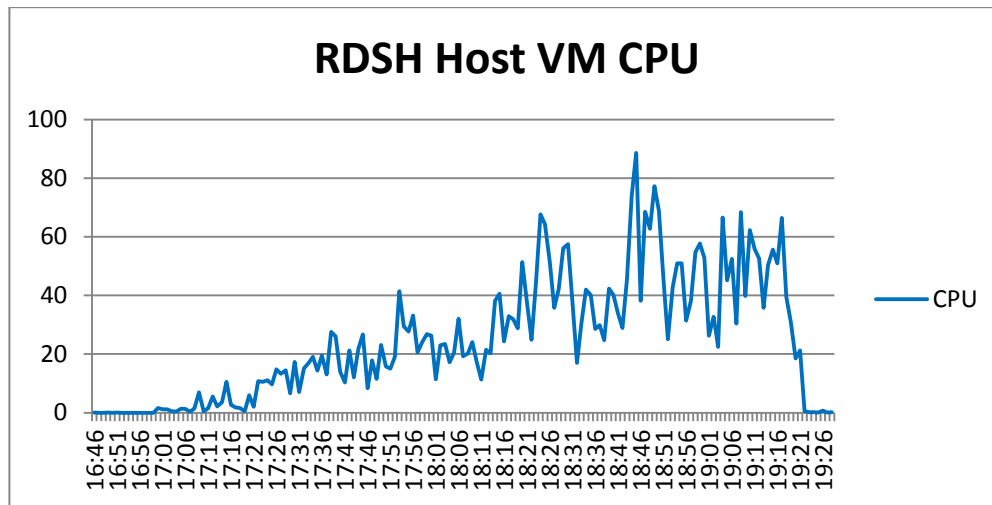
Active Memory

Network bandwidth is not an issue on this solution with a steady state peak of approximately 70,000 Kbps and a spike to 77,000 Kbps during the period of desktop session logoffs.


Network Kbps
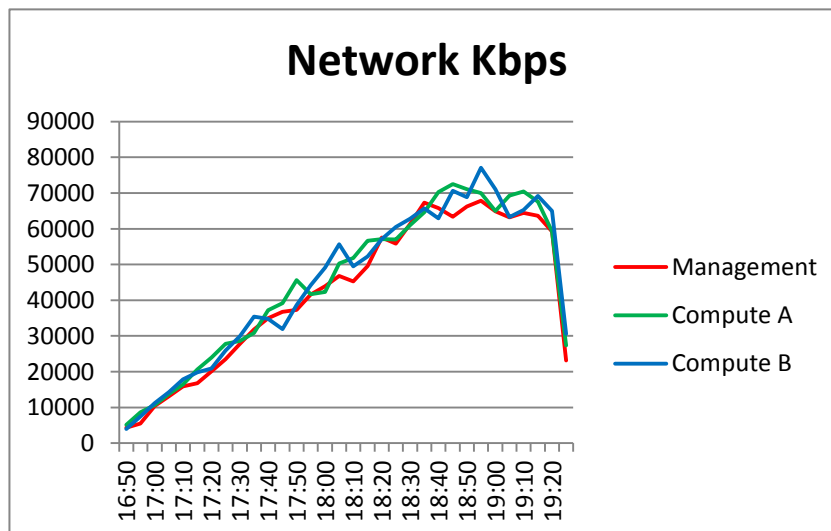
Monitoring of a single, sample RDSH session host showed that its processor reached a maximum of 88% during the test run.

The Nutanix Cluster IOPS figures taken from the Nutanix Web Console showed that the cluster reached a maximum of 1133 Disk IOPS during the test run with the Management node having the lowest disk IOPS of the three nodes.
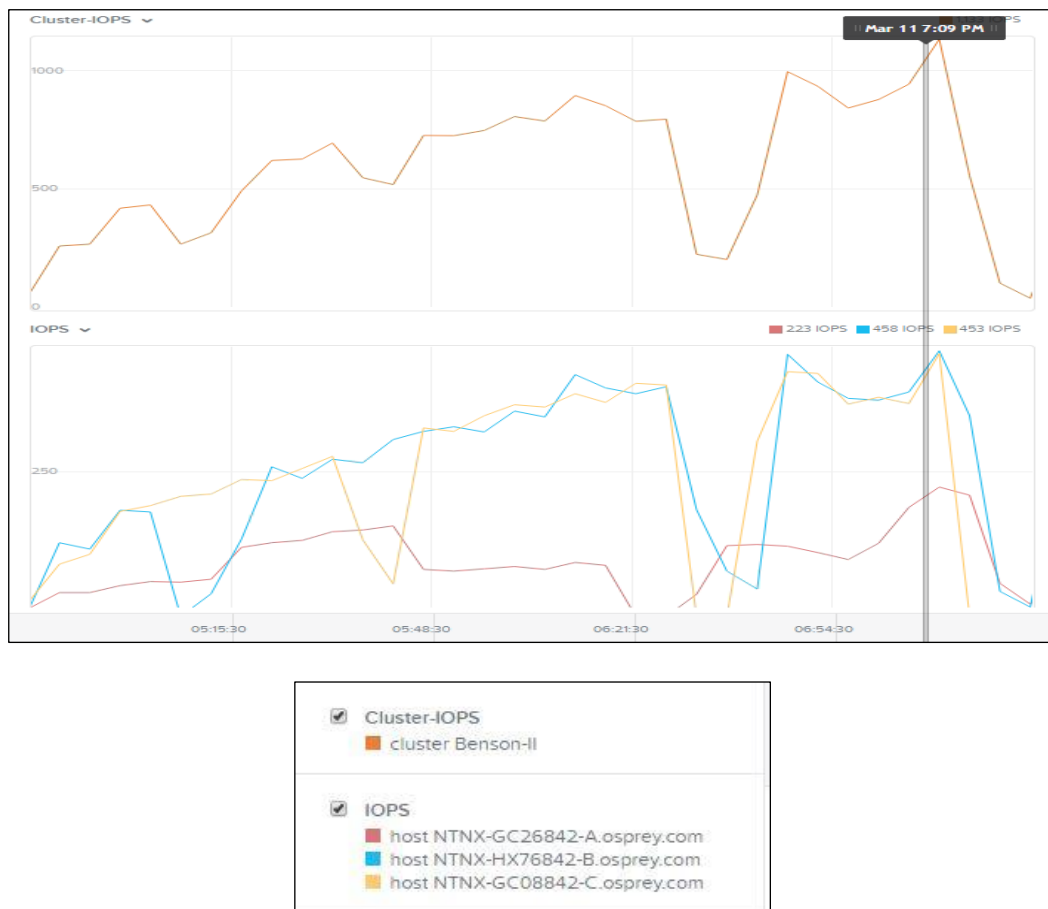
The Login VSI Max user experience score for this test indicates that the VSI Max score not reached so there should have been little deterioration of user experience during testing. There were some erratic login times outside of the normal right from the beginning of test however.
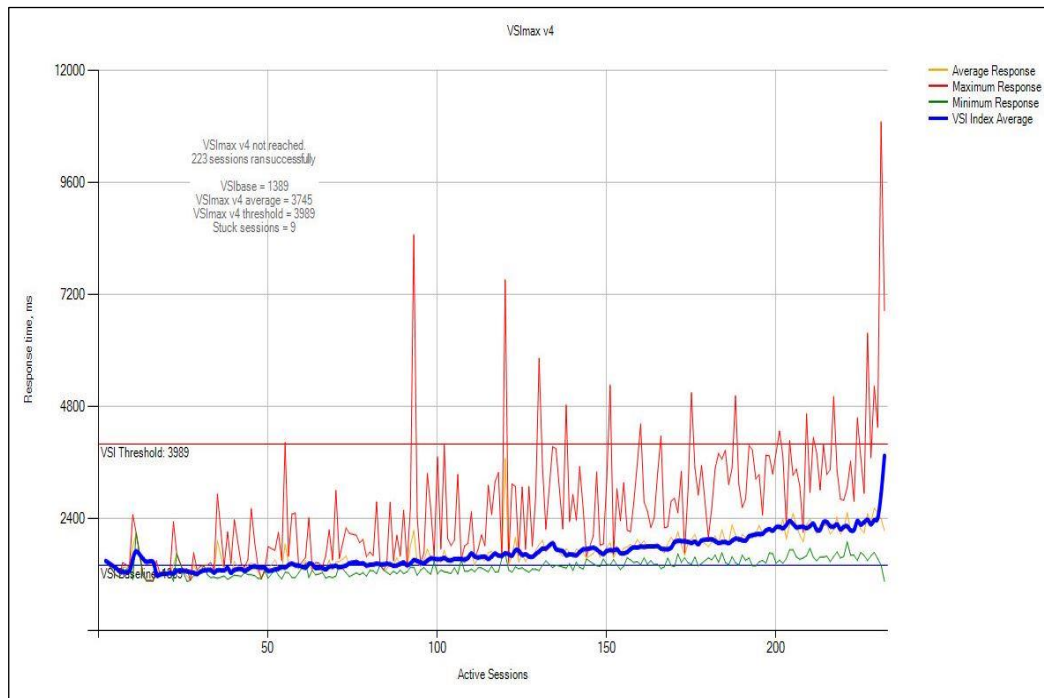
Notes:

- As indicated above, the CPU graphs do not take into account the extra 10% of CPU resources available through the 2690v3's turbo feature.
- Subjective user experience showed mouse movement and window response times when clicking within a running session during steady state were good. Video playback was good quality with no jumpy playback even with the maximum amount of users logged on.
- User login times were a bit erratic from the beginning of the test cycle with some taking extra time to login.
- 512 GB of memory installed on each node is more than enough for this configuration and should run equally well with less memory installed.

Acknowledgements

About the Authors

Peter Fine is the Sr. Principal Solutions Architect for enterprise VDI solutions at Dell. Peter has extensive experience and expertise on the broader Microsoft, Citrix and VMware solutions software stacks as well as in enterprise virtualization, storage, networking and enterprise data center design.

Rick Biedler is the Solutions Development Manager for Datacenter appliances at Dell, managing the development and delivery of Enterprise class Desktop virtualization solutions based on Dell datacenter components and core virtualization platforms.

Paul Wynne is a Principal Engineer within the Cloud Client Computing team at Dell. Paul has over 10 years' experience architecting and supporting VMware, Microsoft & Citrix technologies on a range of Enterprise Storage, Switching and Server technologies.

Andrew Breedy is a staff engineer working in the Dell-Wyse Datacenter engineering group. Andrew has extensive experience in server hardware and on Microsoft and VMWare virtualization solutions as well as a background in engineering test processes and performance analysis and characterization.

Manish Chacko is the Sr. Technical Marketing Advisor for Citrix-based solutions at Dell. Before writing about technology, Manish spent time designing, implementing and supporting technology- in IT, Systems Engineering & Network Performance & Monitoring. Manish was a long-time Dell customer and advocate before becoming a Dell employee.