Wyse Datacenter for Citrix® XenDesktop® Reference Architecture

Dell Engineering
September 2016

A Dell Reference Architecture

Revisions

Date	Description
April 2014	Initial release
March 2016	Document overhaul, new Broadwell CPUs, networking, servers
September 2016	Minor updates to Endpoints section

Contents

1 Introduction

1.1 Purpose

This document addresses several architecture design, configuration and implementation considerations for the key components required to deliver virtual desktops or shared sessions via Citrix® XenDesktop® and XenApp® on Microsoft® Windows Server® Hyper-V® 2012 R2 or VMware® vSphere® 6. Proposed design choices include rack or blade servers, local disks or shared storage. Guidance contained within this document follows a building block methodology enabling the combination of several different components each with their own scaling capabilities.

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 component selection guidance.

1.3 What's New

- Introduce support for Intel Broadwell CPUs (E5-2600v4) and DDR4-2400 RAM
- Introduce support for new Dell Networking options
- Introduce the Dell Appliance for Wyse as base platform for Hyper-V solutions
- Introduce all-flash for Local Tier 1 rack solution

1.4 External Publications & Industry Vertical Information

- Dell Wyse Datacenter for Citrix and Dell PowerEdge VRTX: LINK
- Dell Wyse Datacenter for Citrix and Microsoft Lync 2013: LINK
- Dell Wyse Datacenter for Mobile Clinical Computing: LINK
- Dell Wyse Datacenter for Government: LINK
- Dell Wyse Datacenter for Education: LINK
- Validated 3000 User XenDesktop deployment using Dell Compellent All Flash Array: LINK

2 Solution Architecture Overview

2.1 Introduction

Dell Wyse Datacenter solutions provide a number of deployment options to meet your desktop virtualization requirements. Our solution is able to provide a compelling desktop experience to a range of employees within your organization from task workers to knowledge workers to power users. The deployment options for Dell Wyse Datacenter include:

- Pooled Virtual Desktops (Non-persistent)
- Full Clone Virtual Desktops (Persistent)
- Shared XenApp Sessions (RDSH)

Additionally, our solution includes options for users who require:

- Graphics Acceleration
- Unified Communications

2.2 Physical Architecture Overview

The core Dell Wyse Datacenter architecture consists of two models: Local Tier1 and Shared Tier1. "Tier 1" in the Dell Wyse Datacenter context defines from which disk source the VDI sessions execute. Local Tier1 includes rack servers or blades with SSDs while Shared Tier 1 can include rack or blade servers due to the usage of shared Tier 1 storage. Tier 2 storage is present in both solution architectures and, while having a reduced performance requirement, is utilized for user data and Management VM execution. Management VM execution occurs using Tier 2 storage for all solution models. Dell Wyse Datacenter is a 100% virtualized solution architecture.
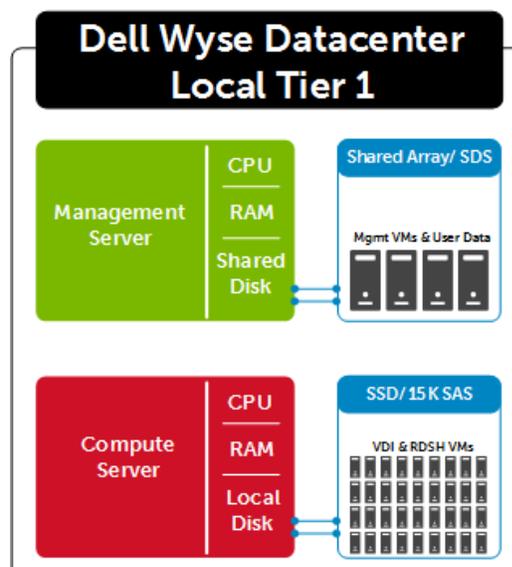
In the Shared Tier 1 solution model, an additional high-performance shared storage array is added to handle the execution of the VDI sessions. All compute and management layer hosts in this model are diskless.
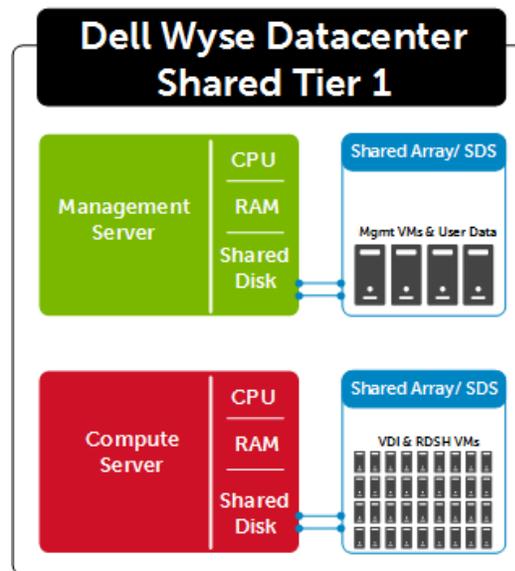
2.3 Solution Layers

The Dell Wyse Datacenter Solution leverages a core set of hardware and software components consisting of five primary layers:

- Networking Layer
- Compute Server Layer
- Management Server Layer
- Storage Layer
- Thin Client Layer (please refer to section 3)

These components have been integrated and tested to provide the optimal balance of high performance and lowest cost per user. The Dell Wyse Datacenter stack is designed to be cost effective allowing IT departments to implement high-performance fully virtualized desktop environments.

2.3.1 Networking

Only a single high performance Dell Networking 48-port switch is required to get started in the network layer. This switch hosts all solution traffic consisting of 1Gb iSCSI and LAN sources for smaller stacks. Above 500 users we recommend that 10Gb LAN be used with iSCSI traffic separated into discrete switching Fabrics. Additional switches are added and stacked as required to provide High Availability for the Network layer.
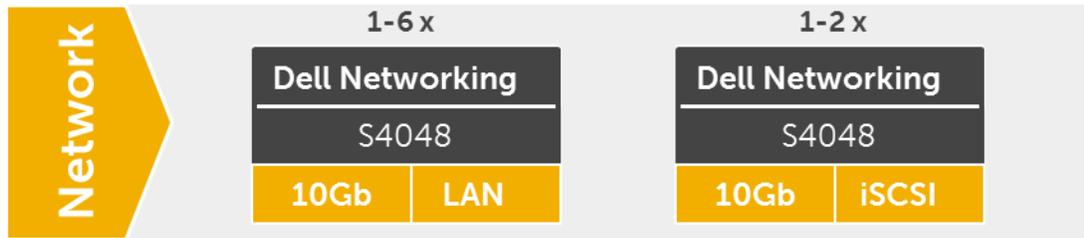
2.3.2   Compute

The compute layer consists of the server resources responsible for hosting the XenDesktop or XenApp user sessions, hosted either via VMware vSphere or Microsoft Hyper-V hypervisors, local or shared Tier 1 solution models (local Tier 1, all-flash, pictured below).
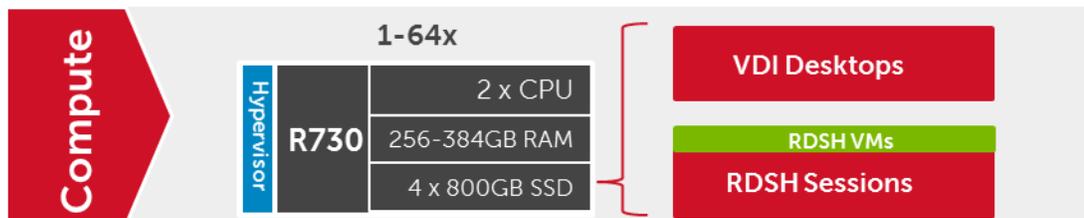
2.3.3   Management

VDI management components are dedicated to their own layer so as to not negatively impact the user sessions running in the compute layer. This physical separation of resources provides clean, linear, and predictable scaling without the need to reconfigure or move resources within the solution as you grow. The management layer will host all the server VMs necessary to support the VDI infrastructure.
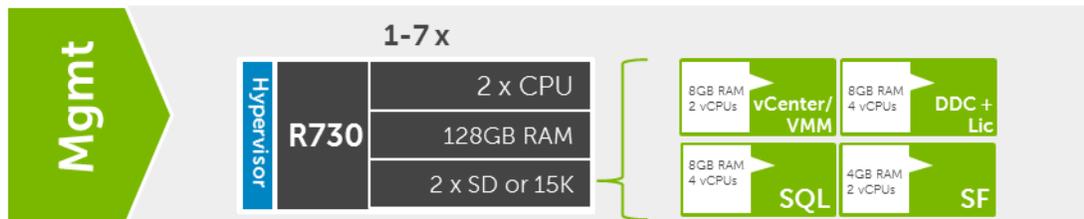
2.3.4   Storage

The storage layer consists of options provided by EqualLogic (EQL) for iSCSI, Compellent (CML) arrays for Fiber Channel or Software Defined to suit your Tier 1 (T1) and Tier 2 (T2) scaling and capacity needs. CML pictured below.
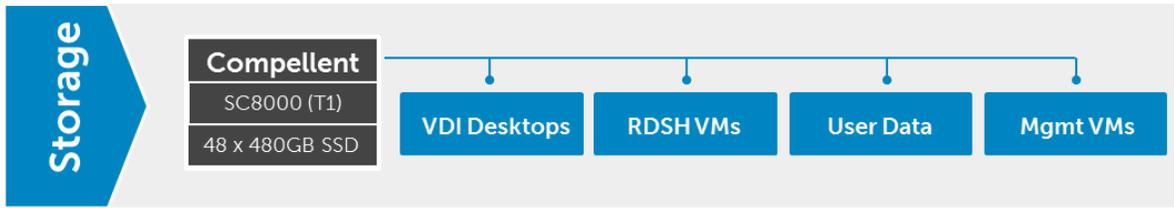
2.4 Local Tier 1

2.4.1 Base Offerings

2.4.1.1 Dell Appliance for Wyse

For pilot or small deployments, the Dell Appliance for Wyse provides an optimized server base with a Quick Start Tool (QST) utility to speed initial deployments. This architecture is non-distributed with all VDI, Management, and storage functions on a single host running Hyper-V. If additional scaling is desired, you can grow into a larger distributed architecture seamlessly. Dell QST will help you configure the host, deploy the management infrastructure and create the desktop pool using Citrix Machine Creation Services (MCS). For more information on the Dell Appliance for Wyse, please visit: LINK.
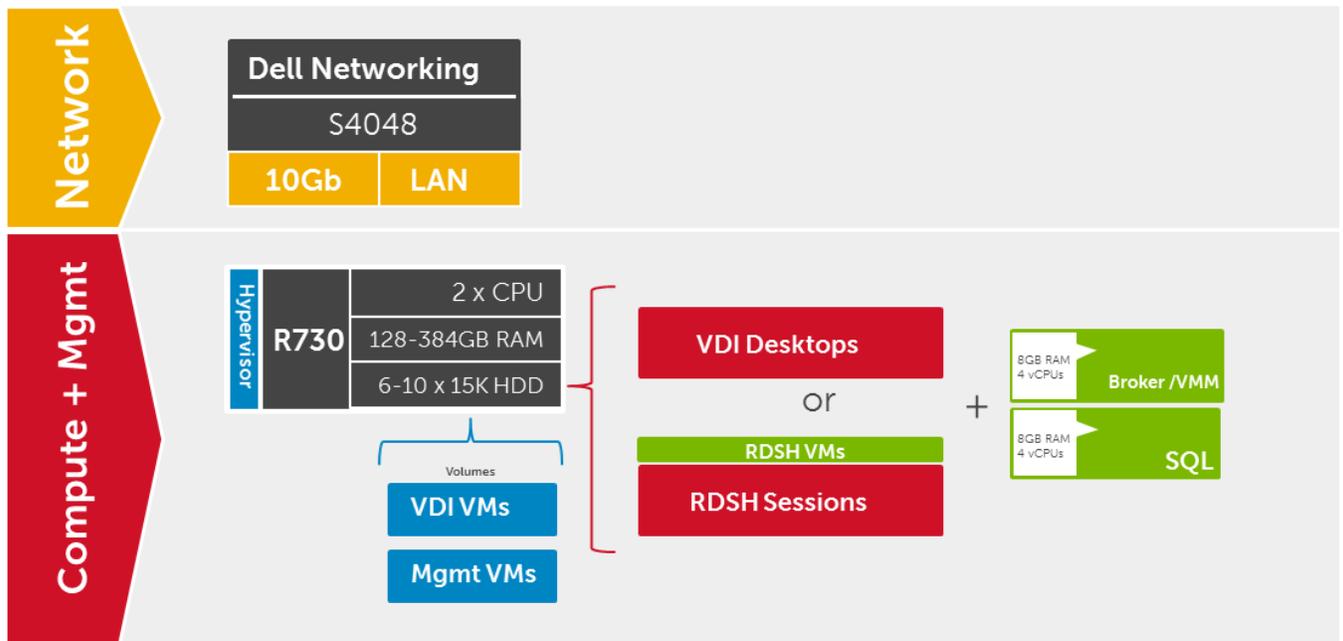
2.4.1.2   Dell Appliance for Wyse (Scale Out)

The Dell Appliance for Wyse can be scaled easily by adding incremental compute nodes for desktop or RDSH VMs and separating the management infrastructure onto its own physical servers protected by external EqualLogic Tier 2 storage. The compute nodes use the large appliance configuration, the mgmt nodes make use of the small appliance configuration.
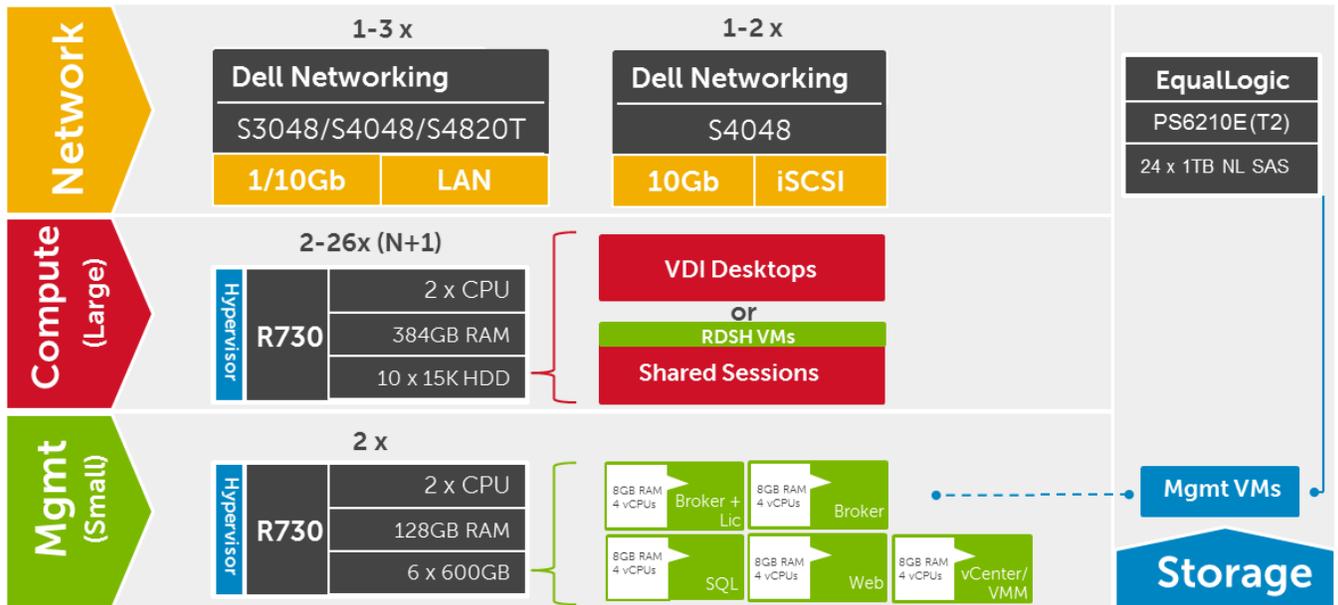
2.4.2   Local Tier 1 for Rack Servers (iSCSI)

The Local Tier 1 solution model provides a high-performance scalable rack-based configuration that hosts user VDI sessions on local SSD or spinning disk in the compute layer. vSphere or Hyper-V based solutions are available and scale based on the chosen hypervisor. All-flash pictured below, if spinning disk is desired substitute the SSDs with 10x 600GB 15K SAS HDDs for vSphere, 12 HDDs for Hyper-V.

Diagram 1:

Network
- 1-6 x — Dell Networking S4048 — 10Gb | LAN
- 1-2 x — Dell Networking S4048 — 10Gb | iSCSI
- EqualLogic — PS6210E (T2) — 24 x 1TB NL SAS

Compute
- 1-64x — Hypervisor — R730 — 2 x CPU | 256-384GB RAM | 4 x 800GB SSD
- VDI Desktops
- RDSH VMs
- RDSH Sessions

Mgmt
- 1-7 x — Hypervisor — R730 — 2 x CPU | 128GB RAM | 2 x SD or 15K
- 8GB RAM 2 vCPUs vCenter/VMM | 8GB RAM 4 vCPUs DDC + Lic
- 8GB RAM 4 vCPUs SQL | 4GB RAM 2 vCPUs SF

Storage
- User Data
- Mgmt VMs

High-performance graphics capabilities compliment the solution and can easily be added at any time to any new or existing deployment. Simply add the appropriate number of graphics-enabled compute hosts to your architecture and provide a superior user experience with NVIDIA GRID vGPU technology.

Diagram 2:

Network
- 1-6 x — Dell Networking S4048 — 10Gb | LAN
- 1-2 x — Dell Networking S4048 — 10Gb | iSCSI
- EqualLogic — PS6210E (T2) — 24 x 1TB NL SAS

Compute
- 1-64x — Hypervisor — R730 — 2 x CPU | 256-384GB RAM | 4 x 800GB SSD
- VDI Desktops
- RDSH VMs
- RDSH Sessions

HDX3D Pro
- 1-64x — Hypervisor — R730 Or 7910 — 2 x CPU | 256GB RAM | 8-12 x 15K HDD | Up to 3 x GPUs
- GFX-enabled VDI Desktops

Mgmt
- 1-7x — Hypervisor — R730 — 2 x CPU | 128GB RAM | 2 x SD or 15K
- 8GB RAM 2 vCPUs vCenter/VMM | 8GB RAM 4 vCPUs DDC + Lic
- 8GB RAM 4 vCPUs SQL | 4GB RAM 2 vCPUs SF

Storage
- User Data
- Mgmt VMs

DELL

2.4.2.1   Local Tier 1 – Network Architecture

In the Local Tier 1 architecture, a single Dell Networking switch is shared among all network connections for both management and compute, up to 500 users. Over 500 users Dell Wyse Solutions Engineering recommends separating the network Fabrics to isolate iSCSI and LAN traffic as well as making each switch stack redundant. Only the management servers connect to iSCSI storage in this model. All Top of Rack (ToR) traffic is layer 2 (L2)/ switched locally, with all layer 3 (L3)/ routable VLANs trunked from a core or distribution switch. The following diagrams illustrate the logical data flow in relation to the core switch.
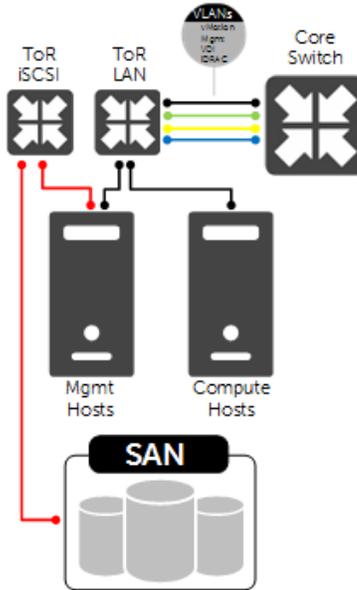
2.4.2.2   Local Tier 1 – Cabling (HA)

The following diagram depicts the LT1 rack solution including optional components for HA:
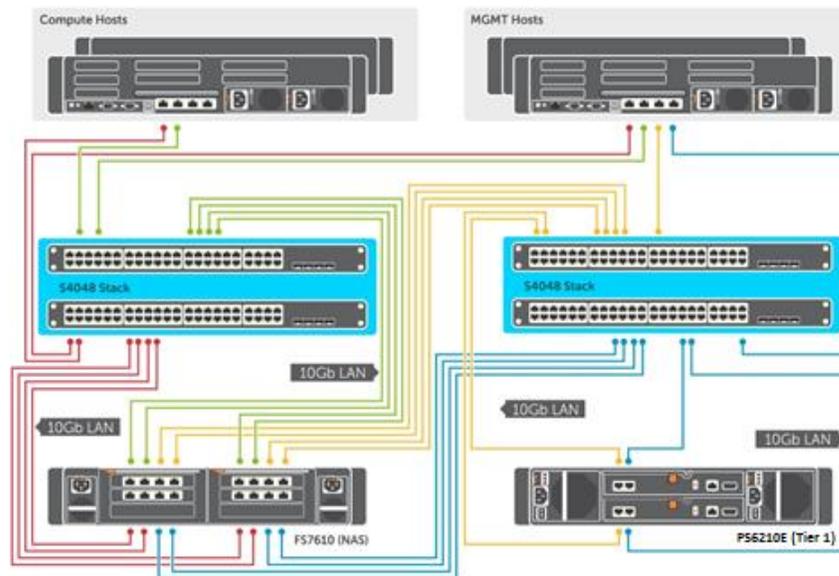
2.4.2.3　Local Tier 1 – Rack Scaling Guidance

Local Tier 1 HW Scaling (iSCSI)				
User Scale	ToR LAN	ToR 10Gb iSCSI	EQL T2	EQL NAS
0-1000	S4048	S4048	6210E	
0-1000 (HA)	S4048	S4048	6210E	FS7610
0-6000+	S4048	S4048	6210E	FS7610

2.4.3　Local Tier 1 for Blade Servers

The Local Tier 1 solution model for blade servers provides a high-performance 800GB SSD configuration that does not require shared storage but Tier 2 is added to provide HA to the management layer infrastructure. User VDI sessions are hosted locally on SSDs in each blade server using Citrix MCS for desktop delivery. A pair of PowerEdge M I/O Aggregator switches are required in the A Fabric. The B Fabric, ToR iSCSI switches and Tier 2 storage are optional for HA or additional bandwidth.
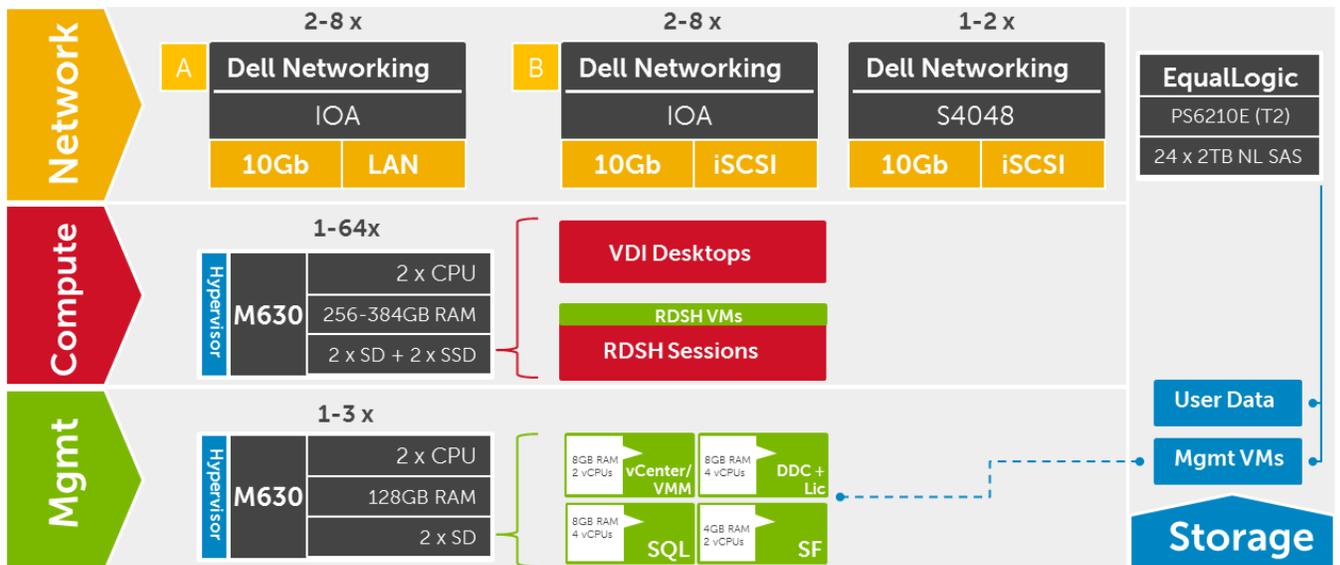
2.4.3.1   Local Tier 1 – Network Architecture

In the Local Tier 1 architecture for blades, there is no need for a ToR switch unless T2 iSCSI is present for HA. The A Fabric IOA switches can connect directly to the core or distribution network layer. Both Management and Compute servers connect to all VLANs in this model via a single vSwitch. The following diagram illustrates the server NIC to chassis switch connections, vSwitch assignments, as well as logical VLAN flow in relation to the core switch.
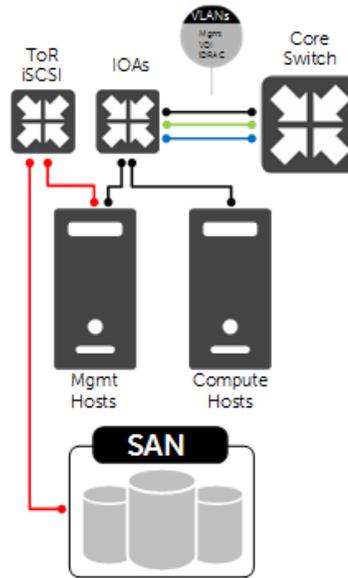
2.4.3.2   Local Tier 1 – Cabling (HA)

The following diagram depicts the LT1 blade solution including optional components for HA. The A Fabric, B Fabric and ToR switches are stacked, respectively.
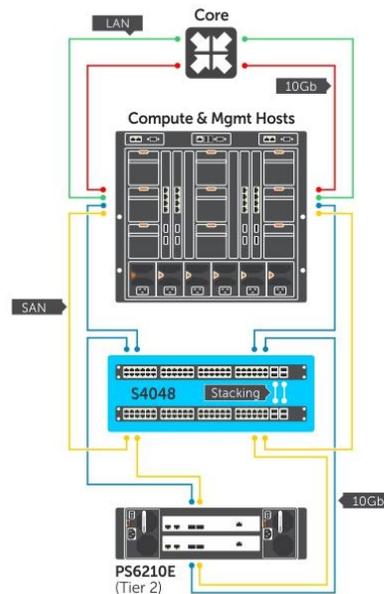
2.4.3.3 Local Tier 1 – Blade Scaling Guidance

Local Tier 1 HW scaling (Blade)				
User Scale	Blade LAN (A Fabric)	Blade iSCSI (B Fabric)	ToR 10Gb iSCSI	EQL T2
0-10000 (no HA)	IOA	-	-	-
0-1000 (HA)	IOA	IOA	S4048	4210E
0-6000+ (HA)	IOA	IOA	S4048	6210E

2.5 Shared Tier 1 Rack

Shared Tier 1 for rack servers incorporates shared Tier 1 storage used for execution of VDI sessions.

2.5.1 Shared Tier 1 for Rack Servers (iSCSI)

For POCs or small deployments, Tier1 and Tier2 can be combined on a single EqualLogic PS6210XS storage array. For over 500 users the storage layers are separated into discrete arrays, as depicted in the figure below. Additional arrays are added for Tier 1 or Tier 2 as the user count grows.
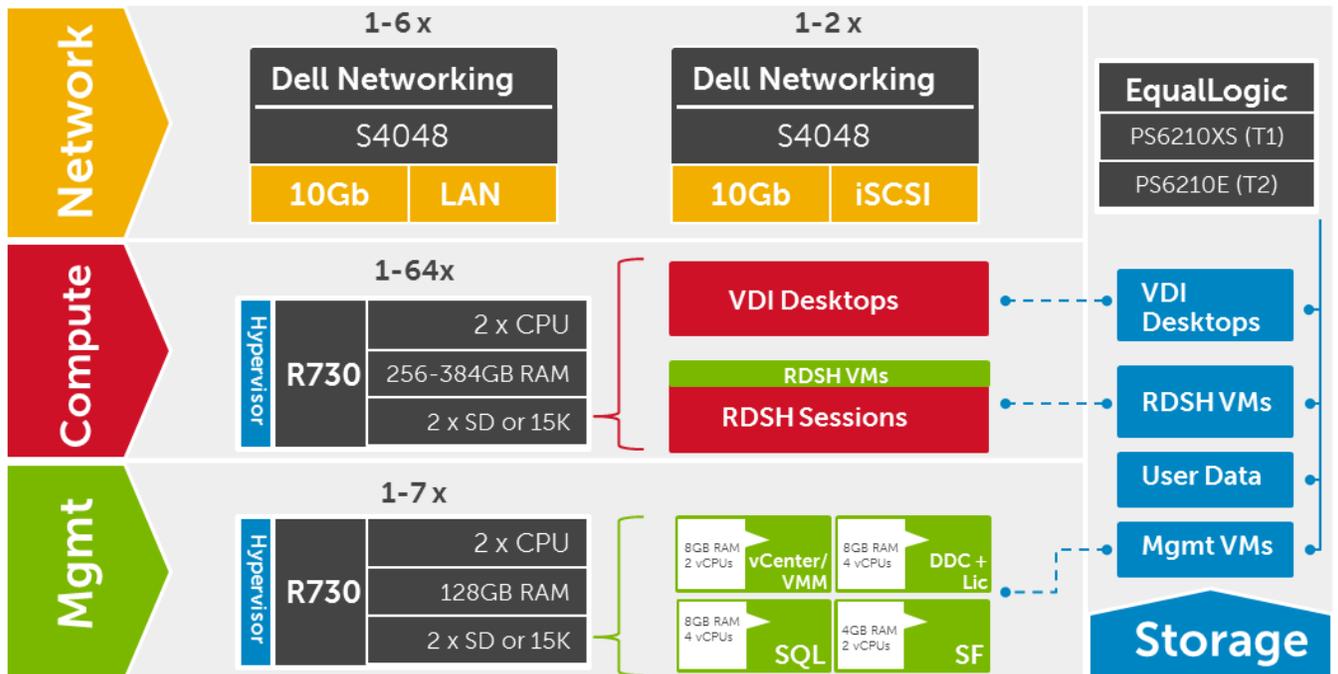
2.5.1.1 Shared Tier 1 – Network Architecture

In the Shared Tier 1 architecture for rack servers, both management and compute servers connect to shared storage. All ToR traffic is layer 2 switched locally, with all layer 3 routable VLANs routed through a core or distribution switch. The following diagrams illustrate the server NIC to ToR switch connections, vSwitch assignments, as well as logical VLAN flow in relation to the core switch.
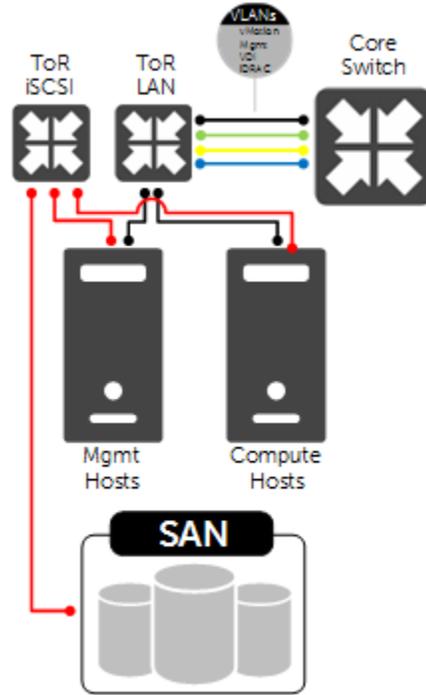
2.5.1.2 Shared Tier 1 – Cabling (HA)
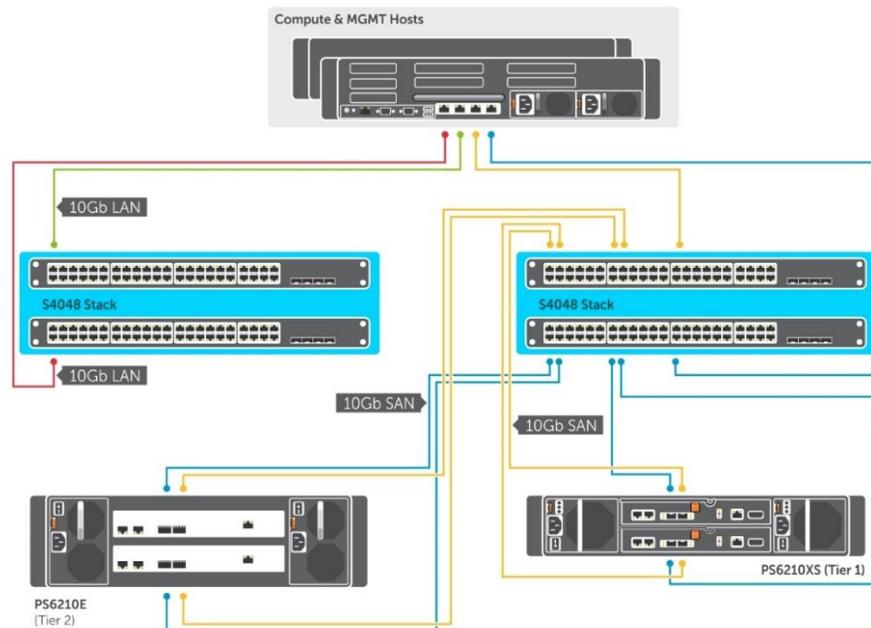
2.5.1.3　Shared Tier 1 – Rack Scaling Guidance

Shared Tier 1 HW scaling (Rack – iSCSI)					
User Scale	ToR LAN	ToR 10Gb iSCSI	EQL T1	EQL T2	EQL NAS
0-500	S4048	S4048	6210XS	-	-
500-1000				6210E	-
0-6000+ (HA)					FS7610

2.5.2　Shared Tier 1 for Rack Servers (FC)

Utilizing Compellent storage for Shared Tier 1 provides a fiber channel (FC) solution where Tier 1 and Tier 2 can optionally be combined in a single array. Tier 2 functions (user data + Mgmt VMs) are removed from the array if the customer has another Tier 2 solution in place or if a Tier 2 Compellent array is used. Scaling this solution is very linear by predictably adding Compellent arrays for every 2000 basic users, on average. The image below depicts a 3000 user array. Please see section 3.3 for more information.
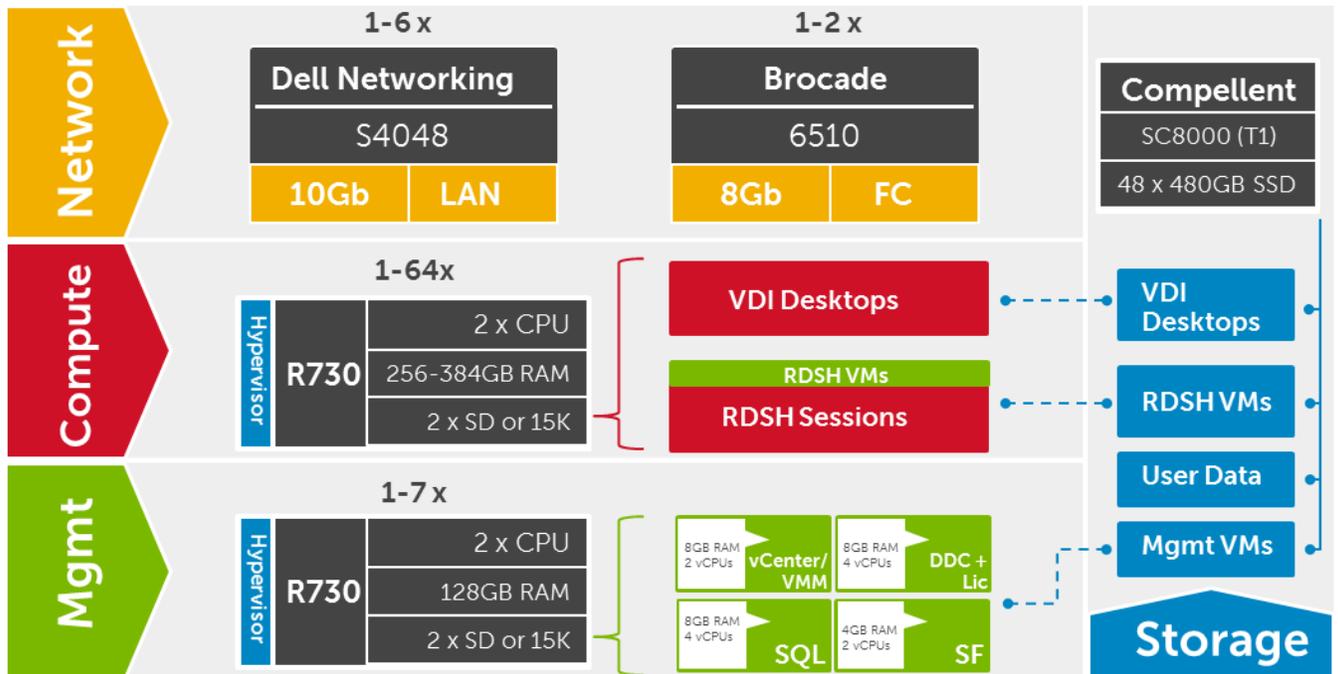
2.5.2.1   Shared Tier 1 – Network Architecture

In the Shared Tier 1 architecture for rack servers using FC, a separate switching infrastructure is required for FC. Management and compute servers both connect to shared storage using FC. Both management and compute servers connect to all network VLANs in this model. All ToR traffic has designed to be layer 2/ switched locally, with all layer 3/ routable VLANs routed through a core or distribution switch. The following diagrams illustrate the server NIC to ToR switch connections, vSwitch assignments, as well as logical VLAN flow in relation to the core switch.
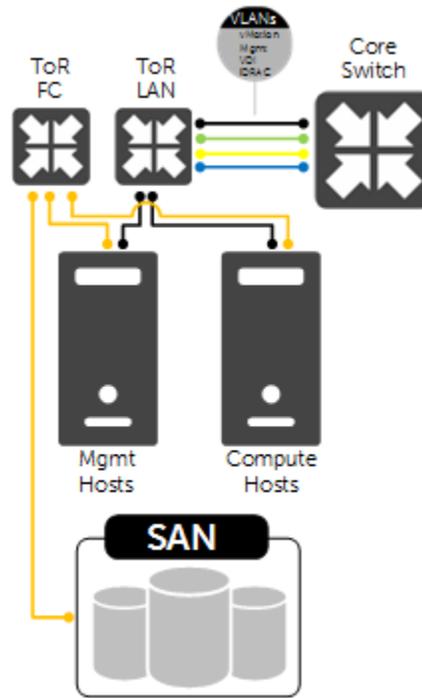
2.5.2.2   Shared Tier 1 – Cabling (HA)
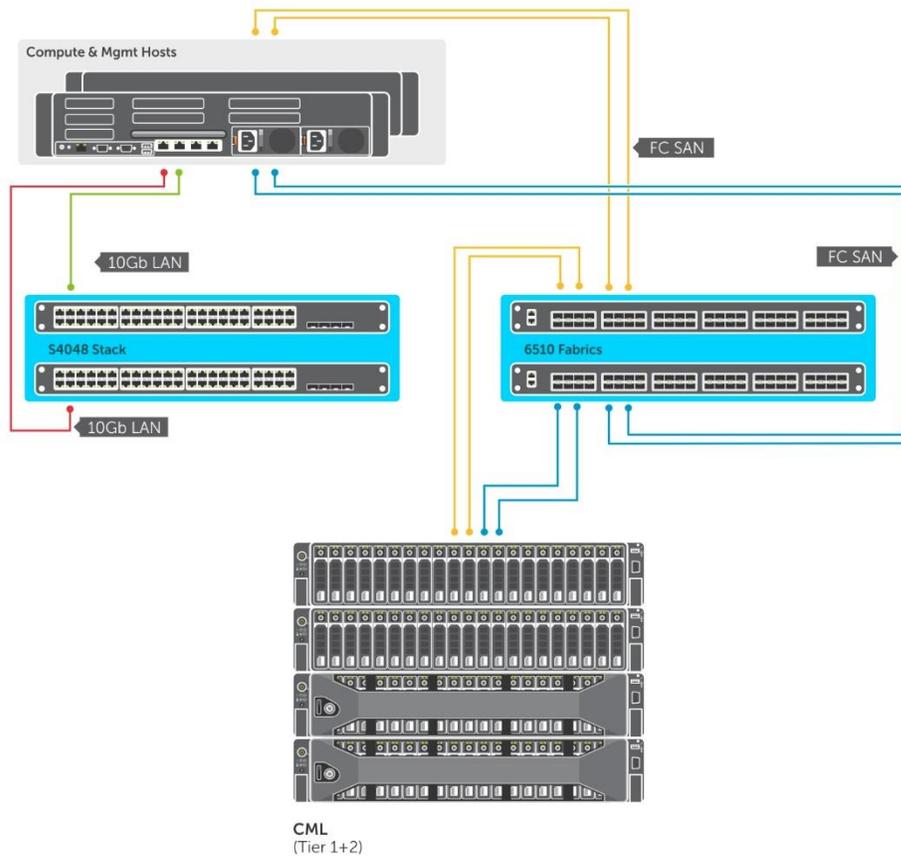
2.5.2.3   Shared Tier 1 – Rack Scaling Guidance

Shared Tier 1 HW scaling (Rack - FC)					
User Scale	LAN Network	FC Network	CML T1	CML T2	CML NAS
0-1000	S4048	6510	SSD	-	-
0-6000+ (HA)	S4048	6510	SSD	SSD or 15K	FS8600

2.6 Shared Tier 1 Blade

2.6.1 Shared Tier 1 for Blade Servers (iSCSI)

As is the case in the Shared Tier 1 model using rack servers, blades can also be used in a 500 user bundle by combing Tier 1 and Tier 2 on a single 6210XS array. Above 500 users, separate Tier 1 and Tier 2 storage into discrete arrays.
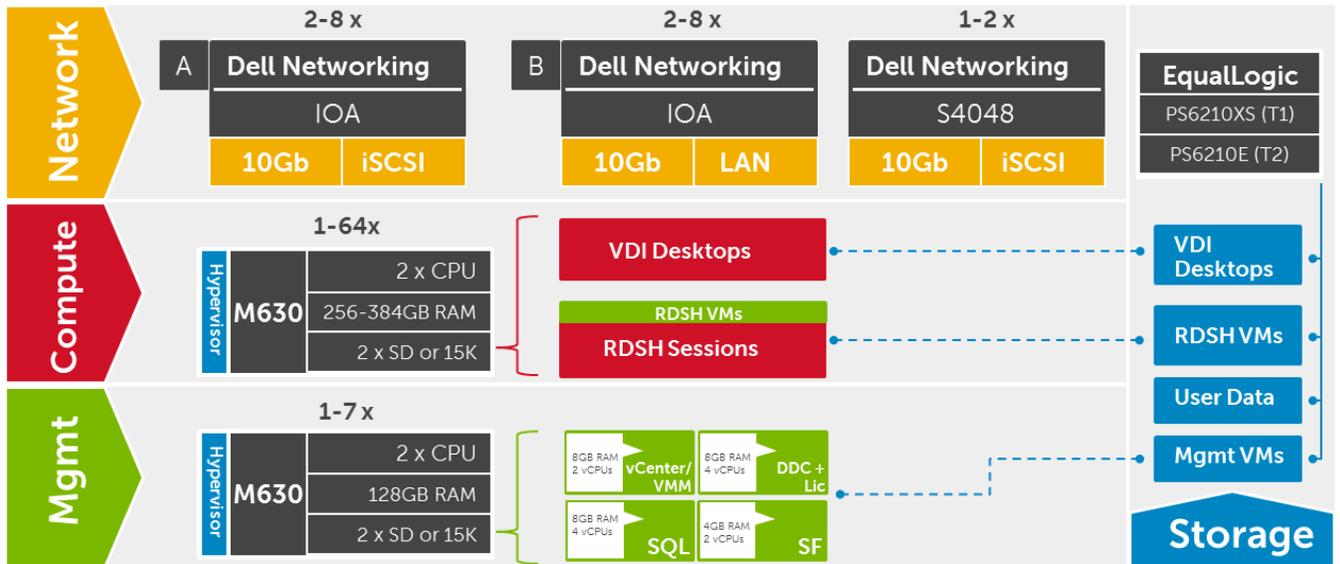
2.6.1.1 Shared Tier 1 – Network Architecture

In the Shared Tier 1 architecture for blades, only iSCSI is switched through ToR. There is no need to switch LAN ToR since the IOAs in the chassis supports LAN to the blades and are uplinked to the core or distribution layers directly. For greater redundancy, a ToR switch is used to support iDRAC used outside of the chassis. Both Management and Compute servers connect to all VLANs in this model. The following diagram illustrates the chassis switch uplink ports to ToR/ core switch connections, vSwitch assignments, as well as logical VLAN flow in relation to the core switch.
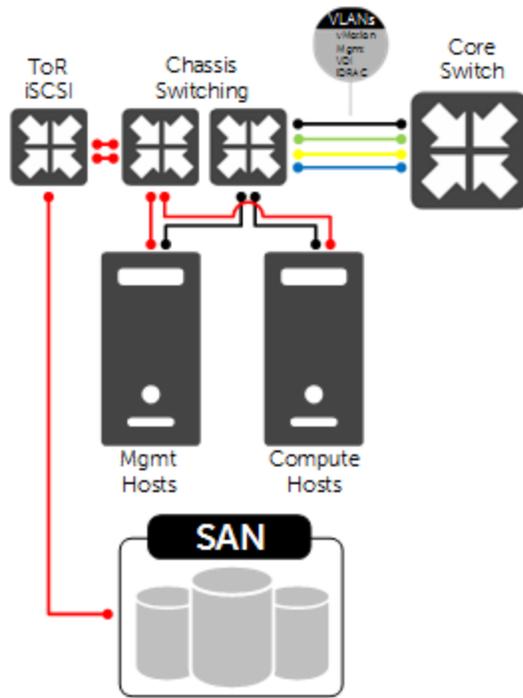
2.6.1.2 Shared Tier 1 – Cabling (HA)
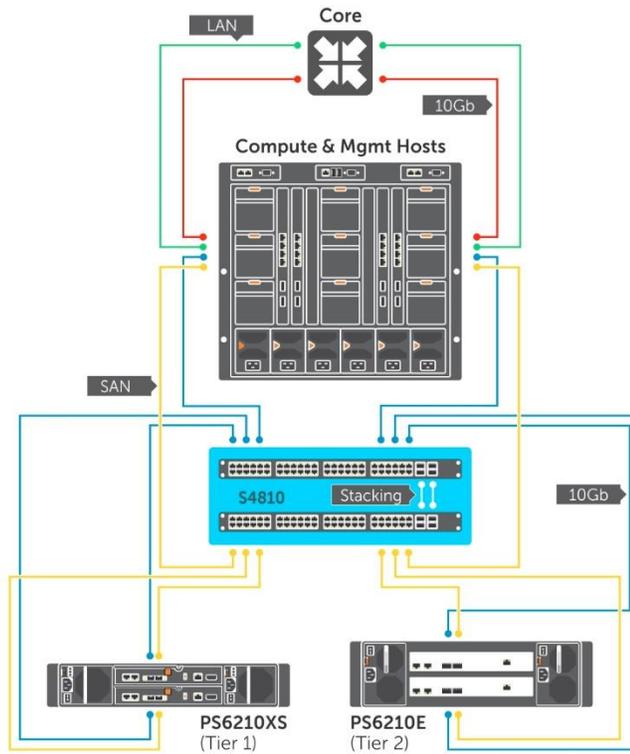
2.6.1.3   Shared Tier 1 – Scaling Guidance

Shared Tier 1 HW scaling (Blade - iSCSI)						
User Scale	Blade LAN (B Fabric)	Blade iSCSI (A Fabric)	ToR 10Gb iSCSI	EQL T1	EQL T2	EQL NAS
0-500	IOA	IOA	S4048	6210XS	-	-
500-1000					6210E	-
0-6000+ (HA)						FS7610

2.6.2   Shared Tier 1 for Blade Servers (FC)

Fiber channel is also an option in the Shared Tier 1 model for blades. There are a few key differences using FC with blades instead of iSCSI: Blade chassis interconnects, FC HBAs in the servers, and FC IO cards in the Compellent arrays. ToR FC switching is optional if a suitable FC infrastructure is already in place. The image below depicts a 4000 user stack.
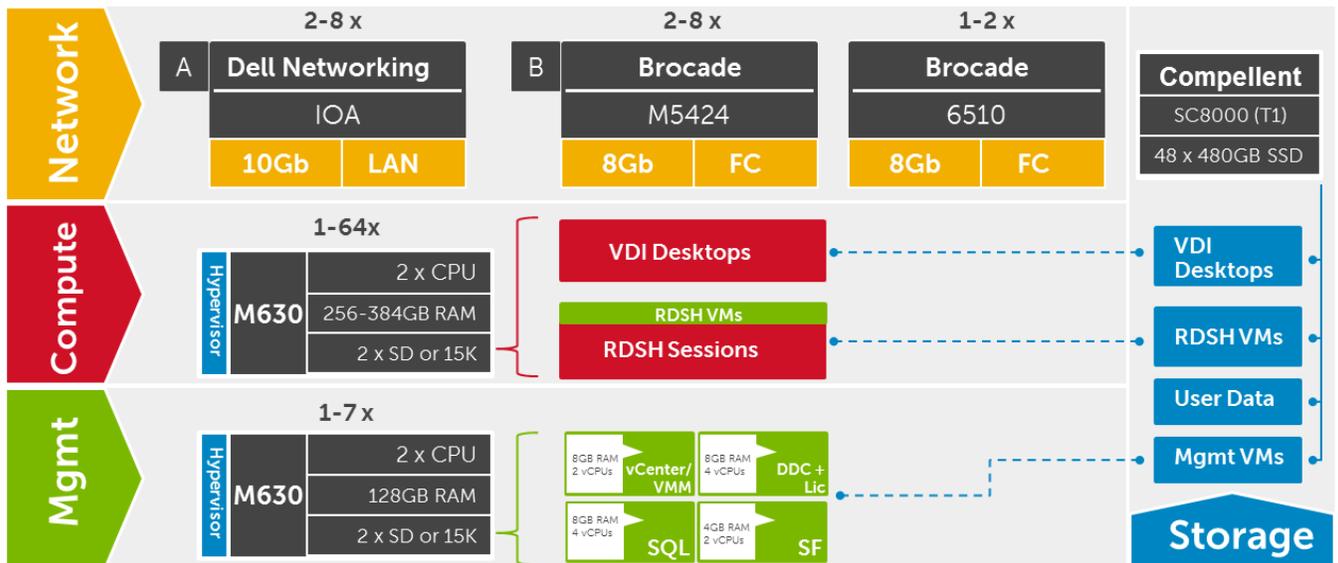

2.6.2.1   Shared Tier 1 – Network Architecture

In the Shared Tier 1 architecture for blade servers using FC, a separate switching infrastructure is required for FC. Management and compute servers both connect to shared storage using FC switched via chassis interconnects. Both management and compute servers connect to all network VLANs in this model. All ToR traffic has designed to be layer 2/ switched locally, with all layer 3/ routable VLANs routed through a core or distribution switch. The following diagrams illustrate the server NIC to ToR switch connections, vSwitch assignments, as well as logical VLAN flow in relation to the core switch.
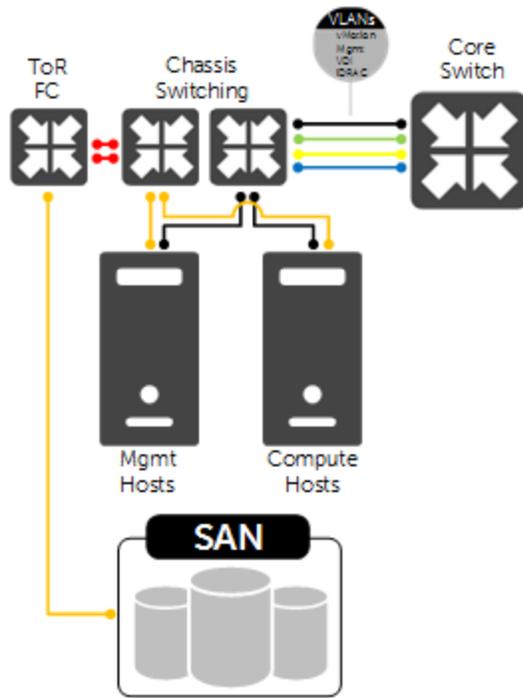
2.6.2.2   Shared Tier 1 – Cabling
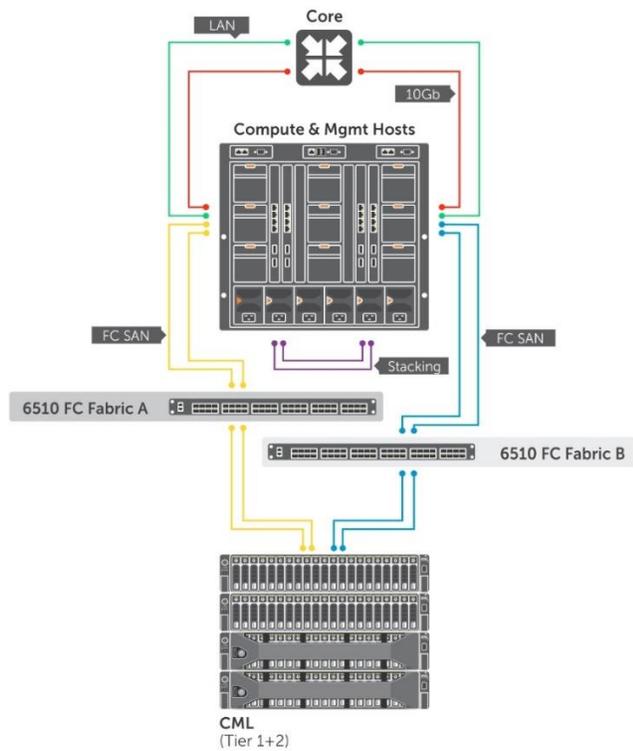
2.6.2.3 Shared Tier 1 – Scaling Guidance

Shared Tier 1 HW scaling (Blade - FC)						
User Scale	Blade LAN (A Fabric)	Blade FC (B Fabric)	ToR FC	CML T1	CML T2	CML NAS
0-500	IOA	5424	6510	SSD	-	-
500-1000					-	-
0-6000+ (HA)					SSD or 15K	FS8600

3 Hardware Components

3.1 Network

The following sections contain the core network components for the Dell Wyse Datacenter solutions. 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 S3048 (1Gb ToR Switch)

Accelerate applications in high-performance environments with a low-latency top-of-rack (ToR) switch that features 48 x 1GbE and 4 x 10GbE ports, a dense 1U design and up to 260Gbps performance. The S3048-ON also supports Open Network Installation Environment (ONIE) for zero-touch installation of alternate network operating systems.

Model	Features	Options	Uses
Dell Networking S3048-ON	48 x 1000BaseT 4 x 10Gb SFP+ Non-blocking, line-rate performance 260Gbps full-duplex bandwidth 131 Mpps forwarding rate	Redundant hot-swap PSUs & fans VRF-lite, Routed VLT, VLT Proxy Gateway User port stacking (up to 6 switches) Open Networking Install Environment (ONIE)	1Gb connectivity

Model	Features	Options	Uses
Dell Networking S3048-ON	48 x 1000BaseT 4 x 10Gb SFP+ Non-blocking, line-rate performance 260Gbps full-duplex bandwidth 131 Mpps forwarding rate	Redundant hot-swap PSUs & fans VRF-lite, Routed VLT, VLT Proxy Gateway User port stacking (up to 6 switches) Open Networking Install Environment (ONIE)	1Gb connectivity

48 x 1GB BaseT ports

4 x SFP+ ports

Redundant Power Supplies

3.1.2 Dell Networking S4048 (10Gb ToR Switch)

Optimize your network for virtualization with a high-density, ultra-low-latency ToR switch that features 48 x 10GbE SFP+ and 6 x 40GbE ports (or 72 x 10GbE ports in breakout mode) and up to 720Gbps performance. The S4048-ON also supports ONIE for zero-touch installation of alternate network operating systems.

Model	Features	Options	Uses
Dell Networking S4048-ON	48 x 10Gb SFP+ 6 x 40Gb QSFP+ Non-blocking, line-rate performance 1.44Tbps bandwidth 720 Gbps forwarding rate VXLAN gateway support	Redundant hot-swap PSUs & fans	10Gb connectivity
		72 x 10Gb SFP+ ports with breakout cables	
		User port stacking (up to 6 switches)	
		Open Networking Install Environment (ONIE)	

48 x 10Gb SFP+ ports   6 x QSFP+ ports

Redundant Power Supplies
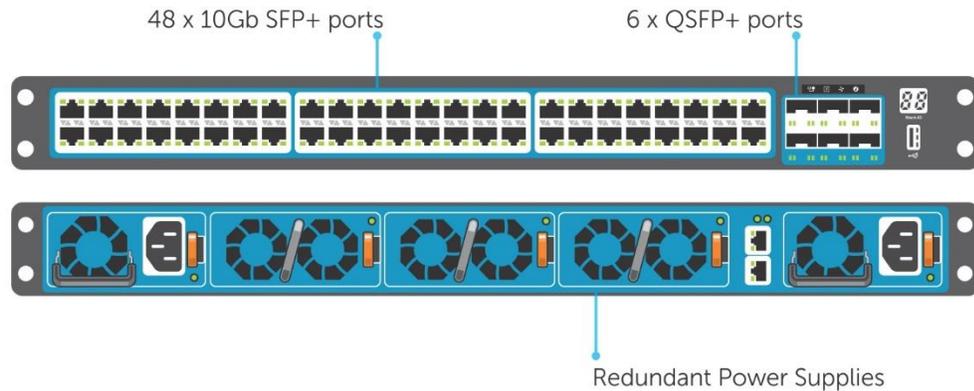
For more information on the S3048, S4048 switches and Dell Networking, please visit: LINK

3.1.3   Brocade 6510 (FC ToR Switch)

The Brocade® 6510 Switch meets the demands of hyper-scale, private cloud storage environments by delivering market-leading speeds up to 16Gb Fibre Channel (FC) technology and capabilities that support highly virtualized environments. Designed to enable maximum flexibility and investment protection, the Brocade 6510 is configurable in 24, 36, or 48 ports and supports 2, 4, 8, or 16Gb speeds in an efficiently designed 1U package. It also provides a simplified deployment process and a point-and-click user interface—making it both powerful and easy to use. The Brocade 6510 offers low-cost access to industry-leading Storage Area Network (SAN) technology while providing "pay-as-you-grow" scalability to meet the needs of an evolving storage environment.

Model	Features	Options	Uses
Brocade 6510	48 x 2/4/8/16Gb Fiber Channel Additional (optional) FlexIO module Up to 24 total ports (internal + external)	Ports on demand from 24, 36, and 48 ports	FC ToR switches for all solutions. Optional for blades

48 x Auto-sensing Ports
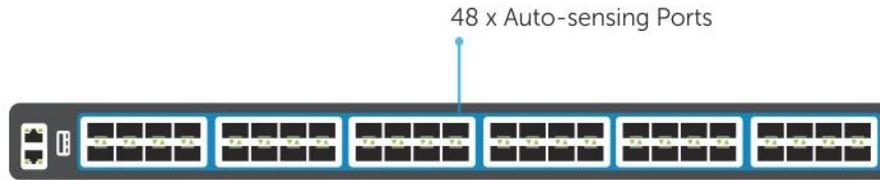
For more information on the Brocade 6510 switch, please visit: LINK

3.1.4   Brocade M5424 (FC Blade Interconnect)

The Brocade® M5424 switches and Dell™ PowerEdge™ M1000e Blade enclosures provide robust solutions for FC SAN deployments. Not only does this offering help simplify and reduce the amount of SAN hardware components required for a deployment, but it also maintains the scalability, performance, interoperability and management of traditional SAN environments. The M5424 can easily integrate FC technology into new or existing storage area network (SAN) environments using the PowerEdge™ M1000e Blade enclosure. The Brocade® M5424 is a flexible platform that delivers advanced functionality, performance, manageability, scalability with up to 16 internal Fabric ports and up to 8 2GB/4GB/8GB auto-sensing uplinks and is ideal for larger storage area networks. Integration of SAN switching capabilities with the M5424 also helps to reduce complexity and increase SAN manageability.

Model	Features	Options	Uses
Brocade M5424	16 x internal Fabric ports Up to 8 2/4/8Gb auto-sensing uplinks	Ports on demand from 12 to 24 ports	Blade switch for FC in Shared Tier 1 model

8 x Auto-sensing Uplinks
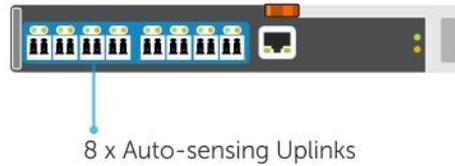
For more information on the Brocade M5424 switch, please visit: LINK

3.1.5   PowerEdge M I/O Aggregator (10Gb Blade Interconnect)

Simplify network management and increase server bandwidth with the PowerEdge™ M I/O Aggregator, enabling easy, plug-and-play data center convergence.

Model	Features	Options	Uses
PowerEdge M I/O Aggregator (IOA)	Up to 32 x 10Gb ports + 4 x external SFP+ 2 x line rate fixed QSFP+ ports 2 optional FlexIO modules	2-port QSFP+ module in 4x10Gb mode	Blade switch for iSCSI in Shared Tier 1 blade solution, LAN + iSCSI in Local Tier 1 blade solution
		4-port SFP+ 10Gb module	
		4-port 10GBASE-T copper module (one per IOA)	
		Stack up to 2 IOAs using QSFP+ ports	

2 x QSFP+ Ports (4 x 10Gb)

FlexIO Slot to Provide additional copper or SFP+ Ports
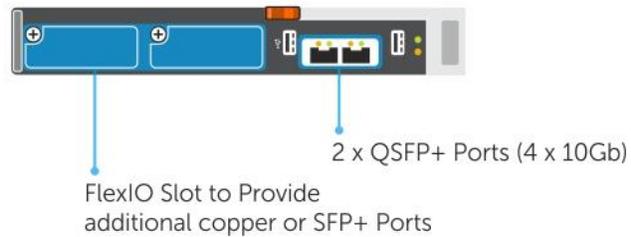
For more information on the Dell IOA switch, please visit: LINK

3.2   Servers

3.2.1   PowerEdge R730

The foundation of the Dell Wyse Datacenter solution portfolio is the best-in-class Dell PowerEdge R730. This dual socket CPU platform runs the fastest Intel Xeon E5-2600 v4 family of CPUs up to 22 cores (Broadwell), can host up to 24 DIMMs of DDR4-2400 RAM, supports up to 16 x 2.5" SAS disks and can be outfitted with 2 double-wide GPU accelerators from AMD or NVIDIA. The Dell PowerEdge R730 offers uncompromising performance and scalability in a 2U form factor.
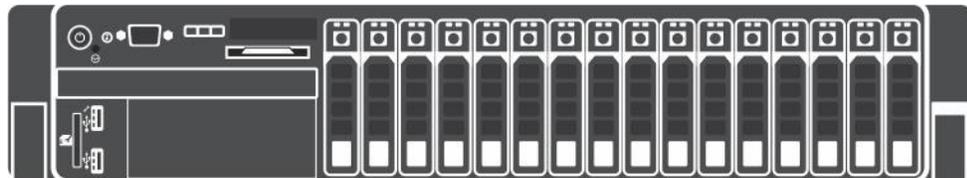
For more information on the R730, please visit: [Link](#)

3.2.1.1   Dell Appliance for Wyse

The Dell Appliance for Wyse can be ordered in two optimized configurations, Large and Small as shown below, or customized as needed. For all-in-one POCs or as a basis for a larger deployment, the Large model provides the foundation of the Local Tier 1 model on Hyper-V. For smaller deployments, POCs using XenApp or distributed architectures intended to scale, the Small platform serves as the basis for the Mgmt stack.

Dell Appliance for Wyse – PowerEdge R730	
Large	**Small**
2 x Intel Xeon E5-2660v4 14C CPU (2Ghz)	2 x Intel Xeon E5-2630v4 10C CPU (2.2Ghz)
384GB Memory (24 x 16GB RDIMMs, (2400MT/s)	128GB Memory (8 x 16GB RDIMMs, 2400MT/s)
Microsoft Hyper-V on 10 x 600GB 15K SAS	**Microsoft Hyper-V** on 6 x 600GB 15K SAS
PERC H730 Integrated RAID Controller – RAID10	PERC H730 Integrated RAID Controller – RAID10
Embedded 2 x 10Gb + 2 x 1Gb NDC	Embedded 2 x 10Gb + 2 x 1Gb NDC
iDRAC8 Enterprise	iDRAC8 Enterprise
2 x 750W PSUs	2 x 750W PSUs

3.2.1.2   Graphics Compute Server

In the Local Tier 1 model, VDI sessions execute on local storage on each compute server. In this model, the management server hosts access shared storage to support the VDI management VMs. Due to the reduced densities on GPU-enabled compute hosts, recommended RAM has been reduced. The management server host does not require local disk to host the management VMs. The recommended management host configuration should follow the standard Local or Shared Tier 1 guidance.

Local or Shared Tier 1 GFX Compute Host – PowerEdge R730	
2 x Intel Xeon E5-2698v4 14C CPU (2Ghz)	2 x Intel Xeon E5-2698v4 14C CPU (2Ghz)
256GB Memory (16 x 16GB RDIMMs, (2400MT/s)	256GB Memory (16 x 16GB RDIMMs, 2400MT/s)
VMware vSphere on internal 8GB Dual SD	**Citrix XenServer** on dual 15K SAS HDDs
4 x 800GB SSD or 10 x 15K SAS	4 x 800GB SSD or 10 x 15K SAS
PERC H730 Integrated RAID Controller – RAID10	PERC H730 Integrated RAID Controller – RAID10
Embedded 2 x 10Gb + 2 x 1Gb NDC	Embedded 2 x 10Gb + 2 x 1Gb NDC
2 x NVIDIA GRID K1/K2 or Tesla M60 GPUs	2 x NVIDIA GRID K1/K2 or Tesla M60 GPUs
iDRAC8 Enterprise	iDRAC8 Enterprise
2 x 1100W PSUs	2 x 1100W PSUs

3.2.1.3   Local Tier 1 Rack

In the Local Tier 1 model, VDI desktops or XenApp sessions execute from local storage on each Compute server. The hypervisors used in this solution are vSphere or Hyper-V. In this model, only the Management server hosts access iSCSI storage to support the solution's Management role VMs using the 1Gb NICs on the embedded Network Daughter Card (NDC). Additional NICs can be added as required for increased bandwidth or resiliency requirements. Refer to section 2.4 for cabling implications. The Management server host has reduced RAM and CPU and does not require local disk to host the management VMs. Any of the compute options below can also be used for XenApp while optionally reducing the amount of RAM. All-flash shown for compute below, the SSDs can be optionally substituted for 10-12 15K SAS.

Local Tier 1 Compute Host – PowerEdge R730	
2 x Intel Xeon E5-2698v4 20C CPU (2.2Ghz)	2 x Intel Xeon E5-2698v4 20C CPU (2.2Ghz)
512GB Memory (16 x 32GB RDIMMs, 2400MT/s)	512GB Memory (16 x 32GB RDIMMs, 2400MT/s)
VMware vSphere on internal 8GB Dual SD	**Microsoft Hyper-V on dual 15K SAS HDDs**
4 x 800GB SSD or 10 x 15K SAS	4 x 800GB SSD or 10 x 15K SAS
PERC H730 Integrated RAID Controller – RAID10	PERC H730 Integrated RAID Controller – RAID10
Embedded 2 x 10Gb + 2 x 1Gb NDC	Embedded 2 x 10Gb + 2 x 1Gb NDC
iDRAC8 Enterprise	iDRAC8 Enterprise
2 x 750W PSUs	2 x 750W PSUs

Local Tier 1 Management Host – PowerEdge R730	
2 x Intel Xeon E5-2660v4 14C CPU (2Ghz)	2 x Intel Xeon E5-2660v4 14C CPU (2Ghz)
128GB Memory (8 x 16GB RDIMMs, 2400MT/s)	128GB Memory (8 x 16GB RDIMMs, 2400MT/s)
VMware vSphere on 2 x internal 8GB Dual SD	**Microsoft Hyper-V** on 2 x 600GB 15K SAS
Embedded 2 x 10Gb + 2 x 1Gb NDC	Embedded 2 x 10Gb + 2 x 1Gb NDC
iDRAC8 Enterprise	iDRAC8 Enterprise
2 x 750W PSUs	2 x 750W PSUs

3.2.1.4   Shared Tier 1 Rack (iSCSI)

In the Shared Tier 1 model, VDI desktops or XenApp sessions execute on shared storage so there is no need for local disk on each server to host VMs. To provide server-level network redundancy using the fewest physical NICs possible, both the Compute and Management servers use a QP NDC: 4 x 10Gb ports 2 x 10Gb ports for iSCSI, 2 x 1Gb ports for LAN. All configuration options (beyond the hypervisor differences noted in the table below) are identical except for CPU and RAM which are reduced on the Management host.

Shared Tier 1 Compute Host PowerEdge R730	
2 x Intel Xeon E5-2698v4 20C CPU (2.2Ghz)	2 x Intel Xeon E5-2698v4 20C CPU (2.2Ghz)
512GB Memory (16 x 32GB RDIMMs, 2400MT/s)	512GB Memory (16 x 32GB RDIMMs, 2400MT/s)
VMware vSphere on internal 8GB Dual SD	**Microsoft Hyper-V** on 2 x 300GB 15K SAS
Embedded 4 x 10Gb NDC	Embedded 4 x 10Gb NDC
iDRAC8 Enterprise	iDRAC8 Enterprise
2 x 750W PSUs	2 x 750W PSUs

Shared Tier 1 Management Host PowerEdge R730	
2 x Intel Xeon E5-2660v4 14C CPU (2Ghz)	2 x Intel Xeon E5-2660v4 14C CPU (2Ghz)
128GB Memory (8 x 16GB RDIMMs, 2400MT/s)	128GB Memory (8 x 16GB RDIMMs, 2400MT/s)
VMware vSphere on internal 8GB Dual SD	**Microsoft Hyper-V** on 2 x 300GB 15K SAS
Embedded 4 x 10Gb NDC	Embedded 4 x 10Gb NDC
iDRAC8 Enterprise	iDRAC8 Enterprise
2 x 750W PSUs	2 x 750W PSUs

3.2.1.4.1 Shared Tier 1 Rack (FC)

Fiber Channel is optionally leveraged as the block storage protocol for Compute and Management hosts with Compellent Tier 1 and Tier 2 storage. Aside from the use of FC HBAs to replace the 10Gb NICs used for iSCSI, the rest of the server configurations are the same.

Shared Tier 1 Compute Host PowerEdge R730	
2 x Intel Xeon E5-2698v4 20C CPU (2.2Ghz)	2 x Intel Xeon E5-2698v4 20C CPU (2.2Ghz)
512GB Memory (16 x 32GB RDIMMs, 2400MT/s)	512GB Memory (16 x 32GB RDIMMs, 2400MT/s)
VMware vSphere on internal 8GB Dual SD	**Microsoft Hyper-V** on 2 x 300GB 15K SAS
Embedded 4 x 10Gb NDC	Embedded 4 x 10Gb NDC
2 x QLogic 2562 8Gb DP FC HBA	2 x QLogic 2562 8Gb DP FC HBA
iDRAC8 Enterprise	iDRAC8 Enterprise
2 x 750W PSUs	2 x 750W PSUs

Shared Tier 1 Management Host PowerEdge R730	
2 x Intel Xeon E5-2660v4 14C CPU (2Ghz)	2 x Intel Xeon E5-2660v4 14C CPU (2Ghz)
128GB Memory (8 x 16GB RDIMMs, 2400MT/s)	128GB Memory (8 x 16GB RDIMMs, 2400MT/s)
VMware vSphere on internal 8GB Dual SD	**Microsoft Hyper-V** on 2 x 300GB 15K SAS
Embedded 4 x 10Gb NDC	Embedded 4 x 10Gb NDC
2 x QLogic 2562 8Gb DP FC HBA	2 x QLogic 2562 8Gb DP FC HBA
iDRAC8 Enterprise	iDRAC8 Enterprise
2 x 750W PSUs	2 x 750W PSUs

3.2.2   PowerEdge M630

The blade server platform recommendation for the Dell Wyse Datacenter solution is the PowerEdge M630. This half-height blade server is a feature-rich, dual-CPU platform that offers a blend of density, performance, efficiency and scalability. The M630 offers remarkable computational density, scaling up to 24 cores, 2 socket Intel Xeon CPUs (Broadwell) and 24 DIMMs (768GB RAM) of DDR4 memory in an extremely compact half-height blade form factor.
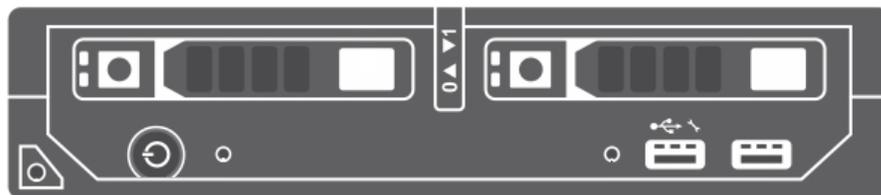
For more information on the PowerEdge M630, please visit: Link

---

3.2.2.1 Local Tier 1 Blade

In the Local Tier 1 model for blades, VDI desktops or XenApp sessions execute on local high-performance SSDs on each compute host. vSphere is the supported hypervisor in this solution due to its ability to run from integrated SD freeing the pair of SSDs for VDI execution only. In this model, shared storage is not required for Tier 2 unless management host-level HA is required. All management and desktop VMs is hosted locally on their respective blades. The recommended provisioning method is MCS with non-persistent desktops. Refer to Section 2.2.3 for solution cabling implications.

Local Tier 1 Compute Host – PowerEdge M630
2 x Intel Xeon E5-2698v4 20C CPU (2.2GHz)
512GB Memory (16 x 32GB RDIMMs, 2400MT/s)
VMware vSphere on 2 x 8GB internal SD
QLogic 57810S-k 10Gb DP KR NDC (LAN)
iDRAC8 Enterprise w/ vFlash, 8GB SD

Local Tier 1 Management Host – PowerEdge M630
2 x Intel Xeon E5-2660v4 14C CPU (2GHz)
128GB Memory (8 x 16GB RDIMMs, 2400MT/s)
VMware vSphere on 2 x 8GB internal SD
QLogic 57810S-k 10Gb DP KR NDC (iSCSI-HA)
iDRAC8 Enterprise w/ vFlash, 8GB SD

3.2.2.2 Shared Tier 1 Blade (iSCSI)

The Shared Tier 1 blade server is configured in line with its rack server equivalent. Two network interconnect Fabrics are configured for the blades: the A-Fabric dedicated to 10Gb iSCSI traffic, the B-Fabric dedicated to 10Gb LAN.

Shared Tier 1 Compute Host – PowerEdge M630		
2 x Intel Xeon E5-2698v4 20C CPU (2.2GHz)		2 x Intel Xeon E5-2698v4 20C CPU (2.2GHz)
512GB Memory (16 x 32GB RDIMMs, 2400MT/s)		512GB Memory (16 x 32GB RDIMMs, 2400MT/s)
VMware vSphere on 2 x 8GB internal SD	Or	**Microsoft Hyper-V** on 2 x 300GB 15K SAS
QLogic 57810S-k 10Gb DP KR NDC (iSCSI)		QLogic 57810S-k 10Gb DP KR NDC (iSCSI)
QLogic 57810-k 10Gb DP KR NDC (LAN)		QLogic 57810-k 10Gb DP KR NDC (LAN)
iDRAC8 Enterprise w/ vFlash, 8GB SD		iDRAC8 Enterprise w/ vFlash, 8GB SD

Shared Tier 1 Management Host – PowerEdge M630		
2 x Intel Xeon E5-2660v4 14C CPU (2GHz)		2 x Intel Xeon E5-2660v4 14C CPU (2GHz)
128GB RAM (8 x 16GB RDIMMs @ 2400MT/s)		128GB RAM (8 x 16GB RDIMMs @ 2400MT/s)
VMware vSphere on 2 x 8GB internal SD	**Or**	**Microsoft Hyper-V** on 2 x 300GB 15K SAS
QLogic 57810S-k 10Gb DP KR NDC (iSCSI)		QLogic 57810S-k 10Gb DP KR NDC (iSCSI)
QLogic 57810-k 10Gb DP KR NDC (LAN)		QLogic 57810-k 10Gb DP KR NDC (LAN)
iDRAC8 Enterprise w/ vFlash, 8GB SD		iDRAC8 Enterprise w/ vFlash, 8GB SD

3.2.2.3   Shared Tier 1 Blade (FC)

Fiber Channel is optionally leveraged as the block storage protocol for Compute and Management hosts with Compellent Tier 1 and Tier 2 storage. Aside from the use of FC HBAs to replace the 10Gb NICs used for iSCSI, the rest of the server configurations are the same. Please note that FC is only currently supported using vSphere.

Shared Tier 1 Compute Host – PowerEdge M630		
2 x Intel Xeon E5-2698v4 20C CPU (2.2GHz)		2 x Intel Xeon E5-2698v4 20C CPU (2.2GHz)
512GB Memory (16 x 32GB RDIMMs, 2400MT/s)		512GB Memory (16 x 32GB RDIMMs, 2400MT/s)
VMware vSphere on 2 x 2GB internal SD	**Or**	**Microsoft Hyper-V** on 2 x 300GB 15K SAS
QLogic 57810S-k 10Gb DP KR NDC (LAN)		QLogic 57810S-k 10Gb DP KR NDC (LAN)
1 x QLogic QME2572 8Gb FC mezz (FC)		1 x QLogic QME2572 8Gb FC mezz (FC)
iDRAC8 Enterprise w/ vFlash, 8GB SD		iDRAC8 Enterprise w/ vFlash, 8GB SD

Shared Tier 1 Management Host – PowerEdge M630		
2 x Intel Xeon E5-2660v4 14C CPU (2GHz)		2 x Intel Xeon E5-2660v4 14C CPU (2GHz)
128GB RAM (8 x 16GB RDIMMs @ 2400MT/s)		128GB RAM (8 x 16GB RDIMMs @ 2400MT/s)
VMware vSphere on 2 x 8GB internal SD	**Or**	**Microsoft Hyper-V** on 2 x 300GB 15K SAS
QLogic 57810S-k 10Gb DP KR NDC (LAN)		QLogic 57810S-k 10Gb DP KR NDC (LAN)
1 x QLogic QME2572 8Gb FC mezz (FC)		1 x QLogic QME2572 8Gb FC mezz (FC)
iDRAC8 Enterprise w/ vFlash, 8GB SD		iDRAC8 Enterprise w/ vFlash, 8GB SD

3.3　Storage

3.3.1　EqualLogic Tier 1 – PS6210XS (iSCSI)

Implement both high-speed, low-latency solid-state disk (SSD) technology and high-capacity HDDs from a single chassis. The PS6210XS 10GbE iSCSI array is a Dell Fluid Data™ solution with a virtualized scale-out architecture that delivers enhanced storage performance and reliability that is easy to manage and scale for future needs. For more information please visit: LINK

Model	Features	Options	Uses
EqualLogic PS6210XS	24 drive hybrid array (SSD + 10K SAS), dual HA controllers, Snaps/Clones, Async replication, SAN HQ, 4 x 10Gb interfaces per controller (2 x SFP + 2 x 10GBT)	13TB – 7 x 400GB SSD + 17 x 600GB 10K SAS	Tier 1 array for Shared Tier 1 solution model (10Gb – iSCSI)
		26TB – 7 x 800GB SSD + 17 x 1.2TB 10K SAS	

7 x SSD + 17 x 10K SAS

10Gb Ethernet Ports　　Mgmt Ports
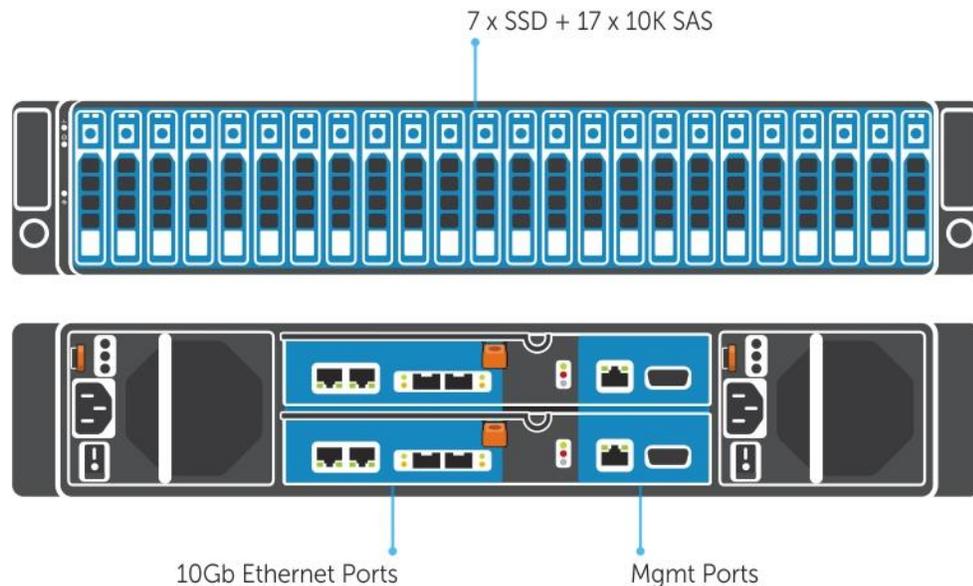
3.3.2　EqualLogic Tier 2 – PS6210E

EqualLogic PS6210 Array Series provide enhanced storage performance for the small-to-medium enterprise with the simplicity of the PS Series product line. Dell EqualLogic PS6210 arrays can drive up to approximately 2GB/sec of throughput per array for sequential, large-block workloads.

Model	Features	Options	Uses
EqualLogic PS6210E	24 drive bays (NL-SAS/ 7.2K RPM), dual HA controllers, Snaps/Clones, Async replication, SAN HQ, 4 x 10Gb ports, 4U chassis	24TB – 24 x 1TB HDs 48TB – 24 x 2TB HDs 72TB – 24 x 3TB HDs 96TB – 24 x 4TB HDs	Tier 2 array for up to 1500 users, per array, in shared Tier 1 solution model (10Gb)

24 x 7.2K NL SAS Drives

10Gb Ethernet Ports   Mgmt Ports
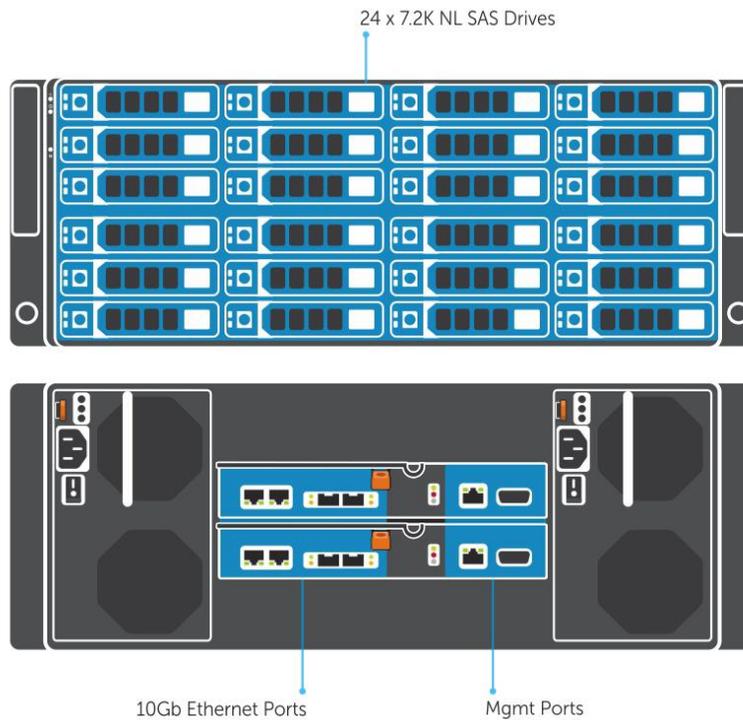
3.3.3   EqualLogic Configuration

Each tier of EqualLogic storage is managed as a separate pool or group to isolate specific workloads. Manage shared Tier 1 arrays used for hosting VDI sessions grouped, while managing shared Tier 2 arrays used for hosting Management server role VMs and user data grouped separately.
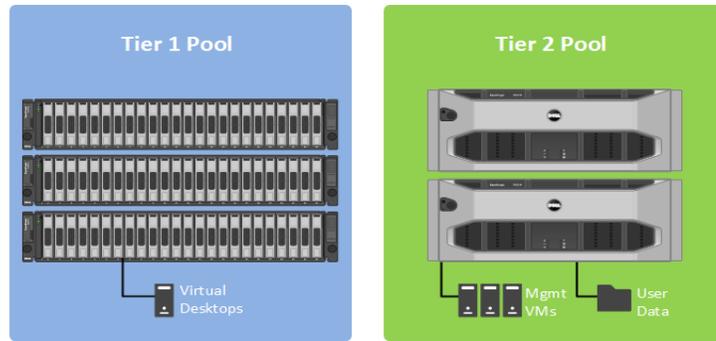
3.3.4 Compellent Tier 1 – SC8000

Experience enhanced storage performance, efficiency and scalability with the next-generation Dell Compellent SC8000 Storage Center Controller. Dell Wyse Solutions Engineering recommends that all Compellent storage arrays be implemented using two controllers in an HA cluster. Fiber Channel is the preferred storage protocol for use with this array, but Compellent is fully capable of supporting iSCSI as well. Key Storage Center applications used strategically to provide increased performance include:
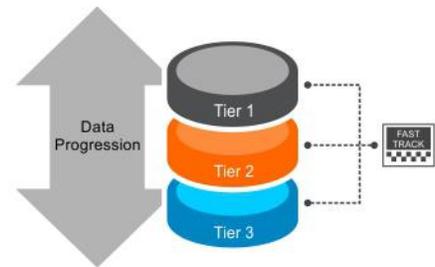
- Simultaneous iSCSI, Fibre Channel (FC) and Fibre Channel over Ethernet (FCoE) front-end interconnects
- Controls SAS and FC drive enclosures, including those with write-intensive and read-intensive SSDs in SLC, MLC and TLC formats
- Scales up to 960 SAS drives in multiple enclosures per dual-controller system and scales-out to many systems across numerous sites, monitored by a single console
- Seamlessly integrates with the FS8600 NAS appliance for fully interoperable block and file storage solutions

Compellent Tier 1 storage consists of a standard dual controller configuration and scales upward by adding disks/ shelves and additional discrete arrays. A single pair of SC8000 controllers will support Tier 1 and Tier 2 for up to 2000 knowledge worker users, as depicted below, utilizing all 15K SAS disks. If Tier 2 is to be separated then an additional 30% of users are added per Tier 1 array. Scaling above this number, additional arrays will need to be implemented. Additional capacity and performance capability is achieved by adding larger disks or shelves, as appropriate, up to the controller's performance limits. Each disk shelf requires 1 hot spare per disk type. RAID is virtualized across all disks in an array (RAID10 or RAID6). Please refer to the test methodology and results for specific workload characteristics in section 7. SSDs are added for use in scenarios where boot storms or provisioning speeds are an issue.

Controller	Front-End IO	Back-End IO	Disk Shelf	Disks	SCOS (min)

2 x SC8000 (16GB)	2 x dual-port 8Gb FC cards (per controller)	2 x quad-port SAS cards (per controller)	2.5" SAS shelf (24 disks each)	2.5" 300GB 15K SAS (~206 IOPS each) or 480GB SSD	6.3

3.3.5   Compellent Tier 2 – SC8000 or SC4020

Compellent Tier 2 storage is optional if a customer wishes to deploy discrete arrays for each Tier. The guidance below is provided for informational purposes. The optional Compellent Tier 2 array consists of a standard dual controller configuration and scales upward by adding disks and shelves. A single pair of SC8000 or SC4020 controllers is able to support Tier 2 for 10,000 basic users. Additional capacity and performance capability is achieved by adding disks and shelves, as appropriate. Each disk shelf requires 1 hot spare per disk type. When designing for Tier 2, capacity requirements will drive higher overall array performance capabilities due to the amount of disk that is on hand. Our base Tier 2 sizing guidance is based on 1 IOPS and 5GB per user.

Controller	Front-End IO	Back-End IO	Disk Shelf	Disks
2 x SC8000 (16GB)	2 x dual-port 8Gb FC cards (per controller)	2 x quad-port SAS cards (per controller)	2.5" SAS shelf (24 disks each)	2.5" 1TB NL SAS (~76 IOPS each)
2 x SC4020 (16GB)	2 x dual-port 8Gb FC cards (per controller)	2 x quad-port SAS cards (per controller)	2.5" SAS shelf (24 disks each)	2.5" 300GB 15K SAS (~206 IOPS each)

Sample Tier 2 Scaling Guidance (based on 1 IOPS per user + 5GB minimum):

Users	Controllers	Disk Shelves	Disks	RAW Capacity
500	2	1	7	7TB
1000	2	1	14	14TB
5000	2	3	66	66TB
10000	2	6	132	132TB

3.3.6 Network Attached Storage (NAS)

3.3.6.1 EqualLogic FS7610

Model	Features	Scaling	Uses
EqualLogic FS7610	Dual active-active controllers, 24GB cache per controller (cache mirroring), SMB & NFS support, AD-integration. Up to 2 FS7610 systems in a NAS cluster (4 controllers). 10Gb iSCSI via 8 x Ethernet ports.	Each controller can support 1500 concurrent users, up to 6000 total in a 2 system NAS cluster.	Scale out NAS for Local Tier 1 to provide file share HA.

3.3.6.2 Compellent FS8600

Model	Features	Options	Uses
Compellent FS8600	Dual active-active controllers, 24GB cache per controller (cache mirroring), SMB & NFS support, AD-integration. Up to 4 FS8600 systems in a NAS cluster (8 controllers). FC only.	Each controller can support 1500 concurrent users, up to 12000 total in a 4 system NAS cluster.	Scale out NAS for Shared Tier 1 on Compellent, to provide file share HA (FC Only).

3.4 GPUs

3.4.1 NVIDIA GRID K1 and K2

NVIDIA GRID™ technology offers the ability to offload graphics processing from the CPU to the GPU in virtualized environments, allowing the data center manager to deliver true PC graphics-rich experiences to more users for the first time. NVIDIA's Kepler™-based GRID K1 and K2 boards are specifically designed to enable rich graphics in virtualized environments.

GPU Virtualization

GRID boards allow hardware virtualization of the GPU. This means multiple users can share a single GPU, improving user density while providing true PC performance and compatibility.

Low-Latency Remote Display

NVIDIA's patented low-latency remote display technology greatly improves the user experience by reducing the lag that users feel when interacting with their virtual machine. With this technology, the virtual desktop screen is pushed directly to the remoting protocol.

Maximum User Density

NVIDIA GRID boards have an optimized multi-GPU design that helps to maximize user density. GRID K1 boards, which include four Kepler-based GPUs and 16GB of memory, are designed to host the maximum number of concurrent users. GRID K2 boards, which include two higher end Kepler GPUs and 8GB of memory, deliver maximum density for users of graphics-intensive applications.

Specs	Grid K1	Grid K2
Number of GPUs	4 x Kepler GPUs (GK107)	2 x high-end Kepler GPUs
Total CUDA cores	768 (192 per GPU)	3072 (1536 per GPU)
Core Clock	850 MT/s	745 MT/s
Total memory size	16 GB DDR3	8 GB GDDR5
Max power	130 W	225 W
Form Factors	Dual slot (4.4" x 10.5")	Dual slot (4.4" x 10.5")
Aux power	6-pin connector	8-pin connector
PCIe	x16 (Gen3)	x16 (Gen3)
Cooling solution	Passive	Passive/ Active

3.4.2 NVIDIA Tesla M60

The NVIDIA® Tesla® M60 is a dual-slot 10.5 inch PCI Express Gen3 graphics card featuring two high-end NVIDIA Maxwell™ GPUs and a total of 16GB GDDR5 memory per card. This card utilizes NVIDIA GPU Boost™ technology which dynamically adjusts the GPU clock to achieve maximum performance. Additionally, the Tesla M60 doubles the number of H.264 encoders over the NVIDIA® Kepler™ GPUs.

Accelerate your most demanding enterprise data center workloads with NVIDIA® Tesla® GPU accelerators. Scientists can now crunch through petabytes of data up to 10x faster than with CPUs in applications ranging from energy exploration to deep learning. Plus, Tesla accelerators deliver the horsepower needed to run bigger simulations faster than ever before. For enterprises deploying VDI, Tesla accelerators are perfect for accelerating virtual desktops.

Specs	Tesla M60
Number of GPUs	2 x NVIDIA Maxwell™ GPUs
Total CUDA cores	4096 (2048 per GPU)
Base Clock	899 MHz (Max: 1178 MHz)
Total memory size	16GB GDDR5 (8GB per GPU)
Max power	300W
Form Factors	Dual slot (4.4" x 10.5")
Aux power	8-pin connector
PCIe	x16 (Gen3)
Cooling solution	Passive/ Active

For more information on NVIDIA Grid, please visit: Link

3.5   Dell Wyse Endpoints

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

3.5.1   Wyse 3020 Thin Client with ThinOS

The Wyse 3020 sets the standard for thin clients. Providing an exceptional user experience, it features the incredibly fast Wyse ThinOS, for environments in which security is critical—there's no attack surface to put your data at risk. Boot up in just seconds and log in securely to almost any network. It delivers a superior Citrix VDI user experience, along with usability and management features found in premium thin clients. The Wyse 3020 delivers outstanding performance based on its dual core system-on-a-chip (SoC) design, and a built-in media CPU delivers smooth multimedia, bi-directional audio and Flash playback. Flexible mounting options let you position it vertically or horizontally on your desk, on the wall or behind your display. Using about 7-watts of power in full operation, the Wyse 3020 creates very little heat for a greener, more comfortable working environment. For more information, please visit: Link

3.5.2   Wyse 3020 Zero Client for Citrix

The Wyse 3020 zero client for Citrix provides an exceptional user experience at a highly affordable price for Citrix XenDesktop and XenApp published desktops. With no attack surface, these zero clients offer protection against network-borne viruses and malware zero attacks. The Wyse 3020 draws less than 7 watts of power in full operation and powers on in just six seconds. The Wyse 3020's extremely compact and efficient design sports gigabit LAN and dual band wireless a/b/g/n, so it's ready for serious tasks whether you're wired in or not. These can be connected to LAN, WAN, or Wi-Fi networks, as these zero clients easily adapt to the latest network security and protocols. For more information, please visit: Link

3.5.3   Wyse 3030 LT Thin Client with ThinOS

The Wyse 3030 LT thin client offers an excellent user experience within a cost-effective offering, and features the virus resistant and extremely efficient Wyse ThinOS, for environments in which security is critical—there's no attack surface to put your data at risk. The 3030 LT delivers outstanding performance based on its dual core processor design, and delivers smooth multimedia, bi-directional audio and flash playback. Boot up in just seconds and log in securely to almost any network. In addition, the Wyse 3030 LT is designed for smooth playback of high bit-rate HD video and graphics within a very compact form factor, with very efficient energy consumption and low heat emissions. Using less than 7 watts of electricity, the Wyse 3030 LT's small size enables discrete mounting options: under desks, to walls, and behind monitors, creating cool workspaces in every respect. For more information, please visit: Link

3.5.4   Wyse 5010 Thin Client with ThinOS

Designed for knowledge workers and power users, the Wyse 5010 is a high-performance thin client based on Wyse ThinOS, the virus-resistant firmware base designed for optimal thin client security, performance, and ease-of-use. Highly secure, compact and powerful, it combines a dual-core AMD 1.4 GHz CPU with a revolutionary unified graphics engine for an outstanding user experience. It addresses the performance challenges of processing-intensive applications like computer-aided design, multimedia, HD video and 3D modeling. Scalable on-premise or cloud-based management provides simple deployment, patching, and updates. Take a unit from box to productivity in minutes with auto configuration. Delivering outstanding processing speed and power, security and display performance, the Wyse 5010 offers a unique combination of performance, efficiency, and affordability. For more information, please visit: Link

3.5.1   Wyse 5010 Zero Client for Citrix

Wyse 5010 is the next-generation zero client for Citrix HDX and Citrix XenDesktop, delivering ultimate performance, security and simplicity. With a powerful dual core AMD G-series CPU, it is faster than competing devices. This additional computing horsepower allows dazzling HD multimedia delivery without overtaxing your server or network. Scalable enterprise-wide management provides simple deployment, patching and updates—your Citrix XenDesktop server configures it out-of-the-box to your preferences for plug-and-play speed and ease of use. Virus and malware resistant, the Wyse 5010 draws under 9 watts of power in full operation—less than any PC on the planet. For more information please visit: Link

3.5.2   Wyse 7010 Thin Client with Linux OS

Designed for power users, the Wyse 7010 is the highest performing thin client on the market. Highly secure and ultra-powerful, it combines Dell Wyse-enhanced SUSE Linux Enterprise with a dual-core AMD 1.65 GHz CPU and a revolutionary unified engine for an unprecedented user experience. It eliminates performance constraints for high-end, processing-intensive applications like computer-aided design, multimedia, HD video and 3D modeling. Scalable enterprise-wide management provides simple deployment, patching and updates. Take a unit from box to productivity in minutes with auto configuration. Delivering unmatched processing speed and power, security and display performance, it's no wonder no other thin client can compare. For more information, please visit: Link

3.5.3   Wyse 7020 Thin Client with WES7

The Wyse 7020 is a super high-performance Windows Embedded Standard 7 thin client for virtual desktop environments. Featuring a quad-core AMD CPU, and an integrated graphics engine that significantly boost performance; it achieves exceptional speed and power for the most demanding VDI and embedded Windows applications, rotational 3D graphics, 3D simulation and modeling, unified communications, and multi-screen HD multimedia. Take a unit from box to productivity in minutes. Just select the desired configuration and it does the rest automatically—no need to reboot. Scale to tens of thousands of endpoints with Wyse Device Manager or leverage your existing Microsoft System Center Configuration Manager platform. The Wyse 7020 is the thin client for power users who need workstation-class performance on their desktop or within a desktop virtualization environment (x86 or x64). For more information, please visit: Link

3.5.4   Wyse 7020 Thin Client with WES8

Wyse 7020 is a super high-performance Windows Embedded 8 Standard thin client for virtual desktop environments. Featuring a quad-core AMD CPU, it offers a vibrant Windows 8 experience and achieves exceptional speed and power for the most demanding embedded Windows applications, rich 3D graphics, and HD multimedia. You can scale to tens of thousands of Wyse 7020 endpoints with Wyse Device Manager (WDM) software or leverage your existing Microsoft System Center Configuration Manager platform. With single-touch or multi-touch capable displays, it adds the ease of an intuitive touch user experience. The Wyse 7020 is an ideal thin client for offering a high-performance Windows 8 experience with the most demanding mix of virtual desktop or cloud applications (x86 or x64). For more information please visit: Link

3.5.5   Wyse 7040 Thin Client with WES7P

The Wyse 7040 is a high-powered, ultra-secure thin client. Equipped with 6th generation Intel i5/i7 processors, it delivers extremely high graphical display performance (up to three displays via display-port daisy-chaining, with 4K resolution available on a single monitor) for seamless access to the most

demanding applications. The Wyse 7040 is compatible with both data center hosted and client-side virtual desktop environments and is compliant with all relevant U.S. Federal security certifications including OPAL compliant hard-drive options, VPAT/Section 508, NIST BIOS, Energy-Star and EPEAT. Wyse enhanced Windows Embedded Standard 7P OS provides additional security features such as BitLocker. The Wyse 7040 offers a high level of connectivity including dual NIC, 6 x USB3.0 ports and an optional second network port, with either copper or fiber SFP interface. Wyse 7040 devices are highly manageable through Intel vPRO, Wyse Device Manager (WDM), Microsoft System Center Configuration Manager (SCCM) and Dell Command Configure (DCC). For more information, please visit: Link

3.5.6   Dell Venue 11 Pro 5000

Meet the ultimate in productivity, connectivity and collaboration. Enjoy full laptop performance in an ultra-portable tablet that has unmatched flexibility for a business in motion. This dual purpose device works as a tablet when you're out in the field but also enables you to work on your desktop in the office thanks to an optional dock. For more information, please visit: Link

3.5.7   Dell Chromebook 13

The lightweight, easy-to-use Dell Chromebook 13 helps turn education into exploration - without the worries of safety or security. Priced to make 1:1 computing affordable today, Chromebook 13 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 5th Gen Intel® CPU provides up to 12 hours of power. Encourage creativity with the Chromebook 13 and its multimedia features that include a 13.3" screen, stereo sound and webcam. For more information, please visit: Link

4 Software Components

4.1 Citrix

4.1.1 Citrix XenDesktop

The solution is based on Citrix XenDesktop which provides a complete end-to-end solution delivering Microsoft Windows virtual desktops or server-based hosted shared sessions 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.

Citrix XenDesktop 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.
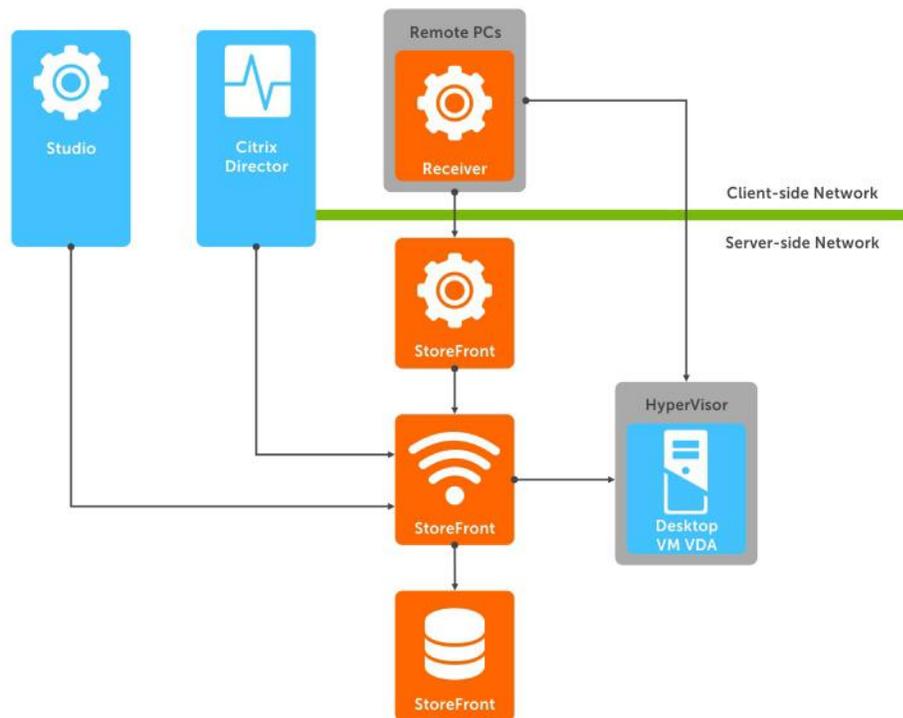
The core XenDesktop components include:

- **Studio**

Studio is the management console that enables you to configure and manage your deployment, eliminating the need for separate management consoles for managing delivery of applications and desktops. Studio provides various wizards to guide you through the process of setting up your environment, creating your workloads to host applications and desktops, and assigning applications and desktops to users.

- **Director**

Director is a web-based tool that enables IT support teams to monitor an environment, troubleshoot issues before they become system-critical, and perform support tasks for end users. You can also view and interact with a user's sessions using Microsoft Remote Assistance.

- **Receiver**

Installed on user devices, Citrix Receiver provides users with quick, secure, self-service access to documents, applications, and desktops from any of the user's devices including smartphones, tablets, and PCs. Receiver provides on-demand access to Windows, Web, and Software as a Service (SaaS) applications.

- **Delivery Controller** (DC)

Installed on servers in the data center, the controller authenticates users, manages the assembly of users' virtual desktop environments, and brokers connections between users and their virtual desktops.

- **StoreFront**

StoreFront authenticates users to sites hosting resources and manages stores of desktops and applications that user's access.

- **License Server**

The Citrix License Server is an essential component at any Citrix-based solution. Every Citrix product environment must have at least one shared or dedicated license server. License servers are computers that are either partly or completely dedicated to storing and managing licenses. Citrix products request licenses from a license server when users attempt to connect.

- **Machine Creation Services** (MCS)

A collection of services that work together to create virtual servers and desktops from a master image on demand; optimizing storage utilization and providing a pristine virtual machine to users every time they log on. Machine Creation Services is fully integrated and administrated in Citrix Studio.

- **Provisioning Services** (PVS)

The Provisioning Services infrastructure is based on software-streaming technology. This technology allows computers to be provisioned and re-provisioned in real-time from a single shared-disk image.

- **Virtual Delivery Agent** (VDA)

The Virtual Desktop Agent is a transparent plugin that is installed on every virtual desktop or XenApp host (RDSH) and enables the direct connection between the virtual desktop and users' endpoint devices.

4.1.2 Machine Creation Services (MCS)

Citrix Machine Creation Services is the native provisioning mechanism within Citrix XenDesktop for virtual desktop image creation and management. Machine Creation Services uses the hypervisor APIs to create, start, stop, and delete virtual desktop images. Desktop images are organized in a Machine Catalog and within that catalog there are a number of options available to create and deploy virtual desktops:

- **Random**: Virtual desktops are assigned randomly as users connect. When they logoff, the desktop is reset to its original state and made free for another user to login and use. Any changes made by the user are discarded at log off.

- **Static**: Virtual desktops are assigned to the same user every time with three options for how to handle changes made to the desktop: Store on local vDisk, Personal vDisk, or discarded on user log off.
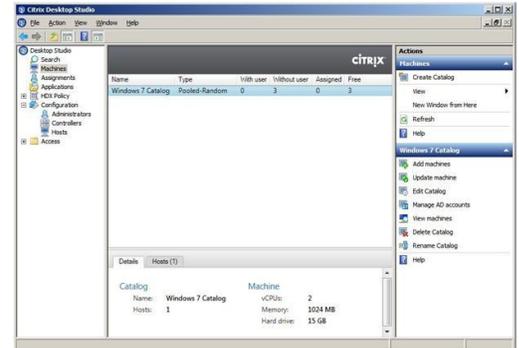
All the desktops in a random or static catalog are based off a master desktop template which is selected during the catalog creation process. MCS then takes snapshots of the master template and layers two additional virtual disks on top: an Identity vDisk and a Difference vDisk. The Identity vDisk includes all the specific desktop identity information such as host names and passwords. The Difference vDisk is where all the writes and changes to the desktop are stored. These Identity and Difference vDisks for each desktop are stored on the same data store as their related clone.
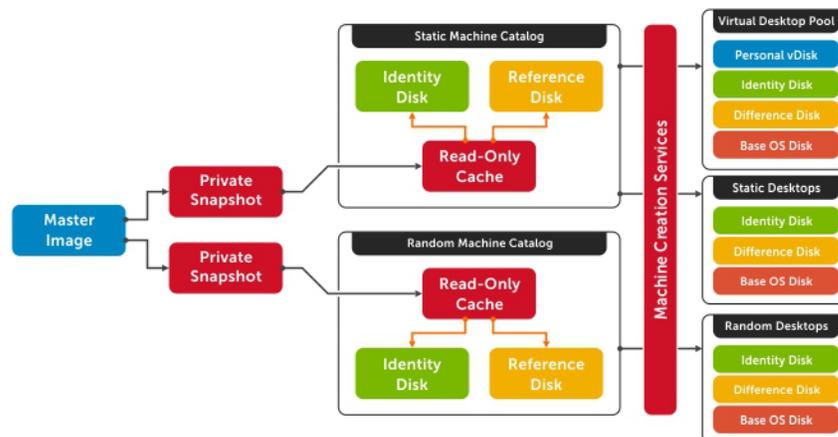
While traditionally used for small to medium sized XenDesktop deployments, MCS can bring along with it some substantial Tier 1 storage cost savings because of the snapshot/identity/difference disk

methodology. The Tier 1 disk space requirements of the identity and difference disks when layered on top of a master image snapshot, is far less than that of a dedicated desktop architecture.

4.1.3   Provisioning Services (PVS)

PVS is an alternative method of image provisioning which uses streaming to share a single base vDisk image instead of copying images to VMs. PVS are used to deliver shared vDisk images to physical or virtual machines. Another potential use is the serial provisioning of XenApp to enable scale-out hosted shared desktop infrastructure. Provisioning Services enables real-time streamed provisioning and re-provisioning which enable administrators to completely eliminate the need to manage and patch individual systems.

Desktop images are organized in a Machine Catalog and within that catalog there are a number of options available to create and deploy virtual or physical desktops:

- **Random**: Virtual or physical desktops are assigned randomly as users connect. When they logoff, the desktop is reset to its original state and made free for another user to login and use. Any changes made by the user are discarded at log off.

- **Static**: Virtual desktops are assigned to the same user every time with user changes stored on a separate Personal vDisk.

Using Provisioning Services, vDisk images are configured in Standard Image mode, read-only, or Private Image mode, read/write. A vDisk in Standard Image mode allows multiple desktops to boot from it simultaneously greatly reducing the number of images that must be maintained and the amount of storage that is otherwise required (non-persistent). Private Image mode vDisks are equivalent to dedicated hard disks and can only be used by one target device at a time (persistent). The Provisioning Server runs on a virtual instance of Windows Server 2012 R2 on the Management Server(s).

4.1.3.1   PVS Write Cache

Citrix Provisioning Services delivery of standard images relies on write-caches to store any writes made by the target OS. The most common write-cache implementation places write-cache on the target machine's storage. Independent of the physical or virtual nature of the target machine, this storage has to be allocated and formatted to be usable.

While there are 4 possible locations for storage of the write cache in PVS, the Dell Wyse Datacenter solution recommends placement of the PVS write cache in the target compute host's RAM with overflow enabled. We recommend using a cache size of 512MB for virtual desktops and 21GB for XenApp VMs delivered via PVS.

4.1.4　Personal vDisk

Citrix Personal vDisk is an enterprise workspace virtualization solution that is built into Citrix XenDesktop. Personal vDisk provides the user customization and personalization benefits of a persistent desktop image with the storage savings and performance of a single/shared image.

Used in conjunction with a static desktop experience, Citrix Personal vDisk allows each user to receive personal storage in the form of a layered vDisk (3GB minimum). This personal vDisk enables users to personalize and persist their desktop environment while providing storage for any user or departmental apps.

Personal vDisk provides the following benefits to XenDesktop:

- Persistent personalization of user profiles, settings and data
- Enables deployment and management of user installed and entitlement based applications
- Fully compatible with Microsoft SCCM and App-V
- 100% persistence with VDI pooled Storage management
- Near Zero management overhead
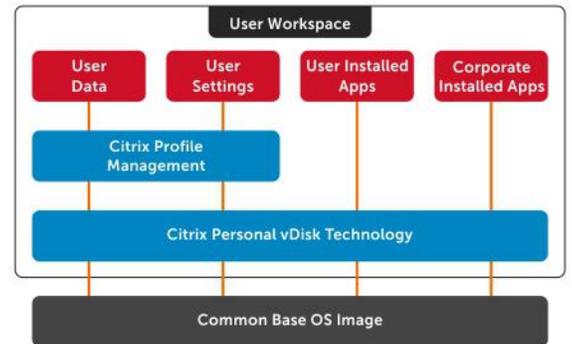
4.1.5　AppDisks

Citrix AppDisk, included in XenDesktop 7.8, provides layering technology to manage departmental applications as an independent storage layer. AppDisk eases the management burden of maintaining multiple departmental images by instantly layering applications onto a single golden image that remains separate and pristine. AppDisks can be associated with either published desktops or published applications via XenApp. AppDisk does not replace the functionality provided by Personal vDisk but currently cannot be used within the same golden image. AppDisks when integrated with AppDNA provides the ability to analyze OS and application performance, compatibility as well as remediation capabilities.
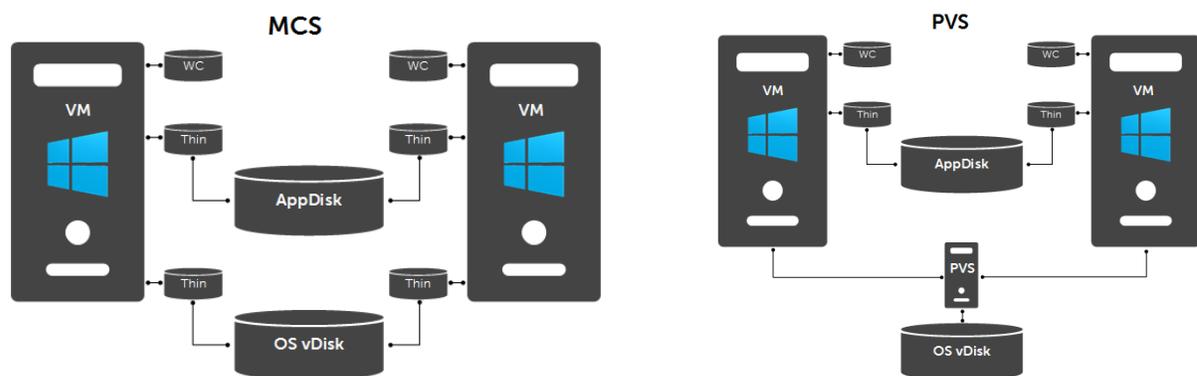
4.1.6   HDX 3D Pro

XenDesktop with HDX 3D Pro is a desktop and app virtualization solution that supports high-end designers and engineers of 3D professional graphics applications and provides cost-effective support to viewers and editors of 3D data. With XenDesktop, you can deliver a persistent user experience and leverage other virtualization benefits such as single-image management and improved data security.

Use HDX 3D Pro technologies with:

- Computer-aided design, manufacturing, and engineering (CAD/CAM/CAE) applications
- Geographical information system (GIS) software
- Picture Archiving Communication System (PACS) workstations for medical imaging
- Latest OpenGL, DirectX, CUDA and CL versions supported
- Latest NVIDIA Grid cards
- Shared or dedicated GPUs or a mix of both on desktop or server OS VMs

HDX 3D Pro provides the best user experience over any bandwidth using Framehawk integration:

- On wide area network (WAN) connections: Deliver an interactive user experience over WAN connections with bandwidths as low as 1.5 Mbps.
- On local area network (LAN) connections: Deliver a user experience equivalent to that of a local desktop on LAN connections with bandwidths of 100 Mbps.

4.1.7   Citrix Profile Manager

Citrix Profile Management is a component of the XenDesktop suite which is used to manage user profiles and minimize many of the issues associated with traditional Windows roaming profiles in an environment where users may have their user profile open on multiple devices at the same time. The profile management toolset has two components: the profile management agent, installed on any device where the user profiles is managed, and a Group Policy Administrative Template, which is imported to a group policy.

In order to further optimize, the profile management folders within the user profile is redirected the users' home drive. The folder redirection is managed via group policy objects within Active Directory. The following folders are redirected:

- Contacts
- Downloads
- Favorites
- Links
- My Documents
- Searches
- Start Menu
- Windows
- My Music
- My Pictures
- My Videos
- Desktop
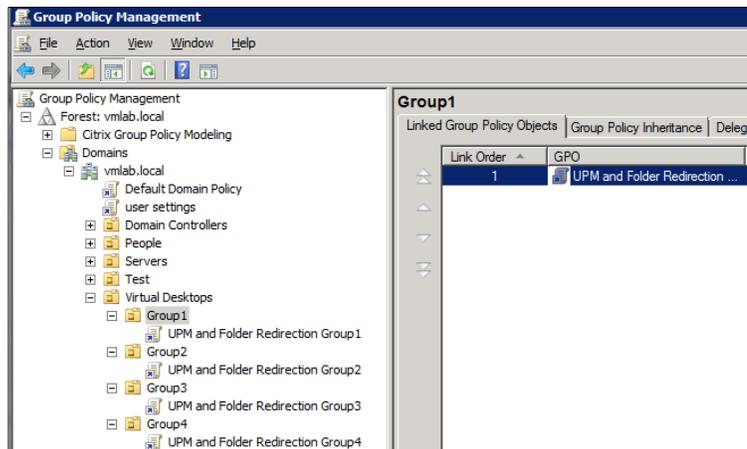
4.1.8   Citrix XenApp

Citrix XenApp 7.6 includes enhancements in the areas of faster access to virtual apps with higher connection resiliency, improved graphics rendering, and new app-usage reporting and monitoring tools.

Citrix XenApp delivers Windows apps as secure mobile services. With XenApp, IT can mobilize the business - increasing user productivity, while reducing costs by centralizing control and security of intellectual property. XenApp delivers high-performance apps to any PC, Mac, laptop, tablet or smartphone that enable the delivery of a native experience that is optimized for the type of device, as well as the network. XenApp is built on a 3rd generation FlexCast Management Architecture (FMA) and is the only hybrid cloud-ready platform that separates the management plane from the workload to enable IT to securely deliver published apps on-premises, and manage workers and mobile workspaces either on-premises or in the cloud.
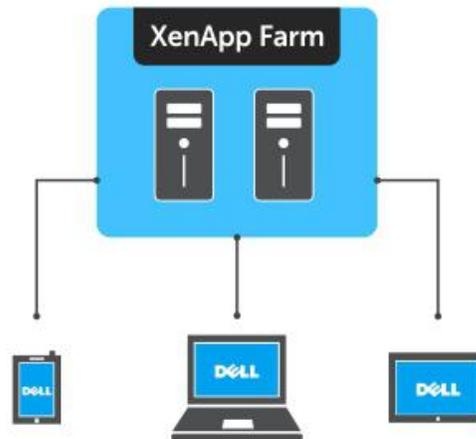
Benefits of hosted desktop sessions and applications:

- Management of applications (single instance)

- Management of simple desktop images (no applications installed)

- PVS to stream XenApp servers as well as user desktops

- Scalability of XenDesktop compute hosts: CPU and IOPS reduction via application offload

- Shared storage scalability: less IOPS = more room to grow

Citrix XenDesktop with XenApp integration can effectively deliver a desktop/application hybrid solution as well. Specifically where a single or small number of shared VDI desktop images are deployed via XenDesktop, each with common shared applications installed within the golden image. A user-specific application set is then deployed and made accessible via the hosted application compute infrastructure, accessible from within the virtual desktop.

User Environment	XenDesktop	XenApp
User-Specific Applications		✔
Profile and User Data	✔	✔
Shared Applications	✔	
Shared Virtual Desktop Image	✔	

Alternatively, XenApp provides a platform for delivering Windows server-based sessions to users who may not need a full desktop VM. Hosted desktops increase infrastructure resource utilization while reducing complexity as all applications and sessions are centrally managed.

User Environment	XenDesktop	XenApp
User-Specific Applications		✔
Profile and User Data		✔
Dedicated Virtual Desktop Image		✔

4.1.8.1 XenApp Integration into Dell Wyse Datacenter Architecture

The XenApp servers can exist as physical or virtualized instances of Windows Server 2012 R2. A minimum of one (1), up to a maximum of eight (7) virtual servers are installed per physical compute host. Since XenApp instances are easily added to an existing XenDesktop stack, the only additional components required are:

- One or more Windows Server OS instances running the Citrix VDA added to the XenDesktop site

The total number of required virtual XenApp servers is dependent on application type, quantity and user load. Deploying XenApp virtually and in a multi-server farm configuration increases overall farm performance, application load balancing as well as farm redundancy and resiliency.

4.1.8.2   XenDesktop with XenApp and Personal vDisk Integration

In a XenDesktop implementation that leverages hosted applications, these execute from a centralized Windows Server and are then accessed via the Citrix Receiver. There are some instances, however, where certain departmental or custom applications cannot run using XenApp. At the same time for organizational policy or certain storage considerations, delivering these applications as a part of a base image is not possible either. In this case, Citrix Personal vDisk technology is the appropriate solution.

With Citrix Personal vDisk, each user of that single shared virtual desktop image also receives a personal layered vDisk, which enables the user to personalize their desktop and receive native application execution within a Windows client OS and not from a server. When leveraging the integration of XenApp within XenDesktop, all profile and user data is seamlessly accessed within both environments.

4.1.8.3   XenApp Integration into Dell Wyse Datacenter Architecture

The XenApp servers can exist as physical or virtualized instances of Windows Server 2012 R2. A minimum of one (1), up to a maximum of eight (8) virtual servers are installed per physical compute host. Since XenApp instances are easily added to an existing XenDesktop stack, the only additional components required are:

- One or more Server OS instances running the Citrix VDA added to the XenDesktop site

The total number of required virtual XenApp servers is dependent on application type, quantity and user load. Deploying XenApp virtually and in a multi-server farm configuration increases overall farm performance, application load balancing as well as farm redundancy and resiliency.

4.1.8.4   PVS Integration with XenApp

One of the many benefits of PVS is the ability to quickly scale the XenApp instances within a farm. Bandwidth is a key consideration and PVS bandwidth utilization is mostly a function of the number of target devices and the portion of the image(s) they utilize. Network impact considerations include:

- PVS streaming is delivered via UDP, yet the application has built-in mechanisms to provide flow control, and retransmission as necessary.

- Data is streamed to each target device only as requested by the OS and applications running on the target device. In most cases, less than 20% of any application is ever transferred.

- PVS relies on a cast of supporting infrastructure services. DNS, DHCP need to be provided on dedicated service infrastructure servers, while TFTP and PXE Boot are functions that may be hosted on PVS servers or elsewhere.

4.1.8.5 NUMA Architecture Considerations

Best practices and testing has showed that aligning XenApp design to the physical Non-Uniform Memory Access (NUMA) architecture of the server CPUs results in increased and optimal performance. NUMA alignment ensures that a CPU can access its own directly-connected RAM banks faster than those banks of the adjacent CPU which are accessed via the Quick Path Interconnect (QPI). The same is true of VMs with large vCPU assignments, best performance will be achieved if your VMs receive their vCPU allotment from a single physical NUMA node. Ensuring that your virtual XenApp servers do not span physical NUMA nodes will ensure the greatest possible performance benefit.

The general guidance for XenApp NUMA-alignment on the Dell Wyse Datacenter solution is as follows:

4.1.8.6 NUMA Alignment

20 physical cores per CPU given the 2698v4 part, 40 logical cores with Hyperthreading active, gives us a total of 80 consumable cores per compute node and falls in line with a 2x oversubscription rate. Configuring the XenApp VMs as shown below (as XA VMs) will ensure that no physical NUMA node spanning occurs which could lower performance for an effected VM.
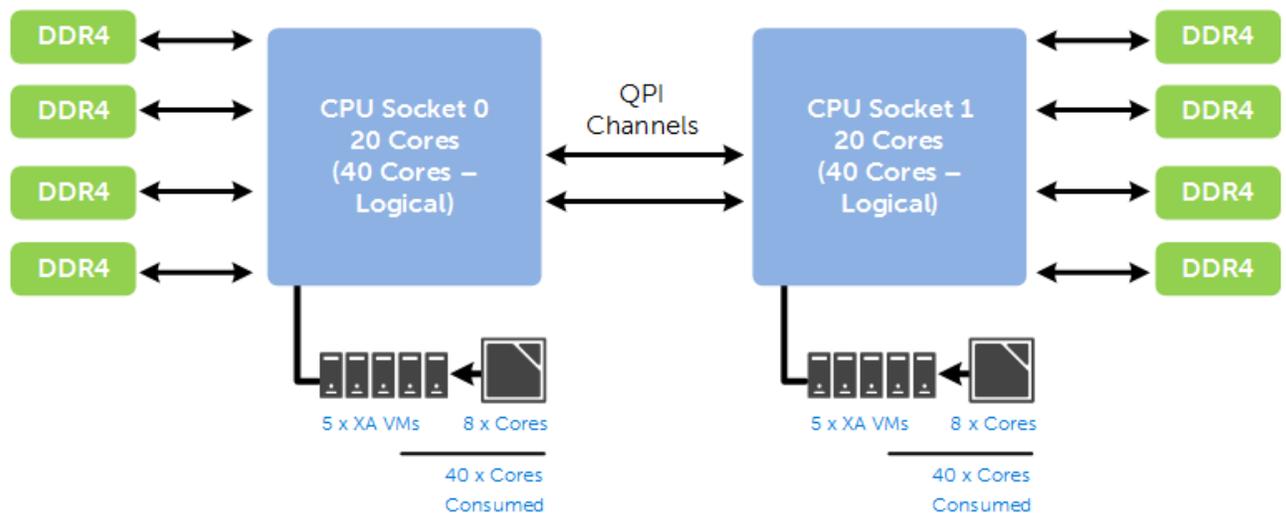
4.1.9 Citrix NetScaler

Citrix NetScaler is an all-in-one web application delivery controller that makes applications run five times better, reduces web application ownership costs, optimizes the user experience, and makes sure that applications are always available by using:

- Proven application acceleration such as HTTP compression and caching
- High application availability through advanced L4-7 load balancer
- Application security with an integrated Application Firewall
- Server offloading to significantly reduce costs and consolidate servers

A NetScaler appliance resides between the clients and the servers, so that client requests and server responses pass through it. In a typical installation, virtual servers (vservers) configured on the NetScaler provide connection points that clients use to access the applications behind the NetScaler. In this case, the NetScaler owns public IP addresses that are associated with its vservers, while the real servers are isolated in a private network. It is also possible to operate the NetScaler in a transparent mode as an L2 bridge or L3 router, or even to combine aspects of these and other modes. NetScaler can also be used to host the StoreFront function eliminating complexity from the environment.
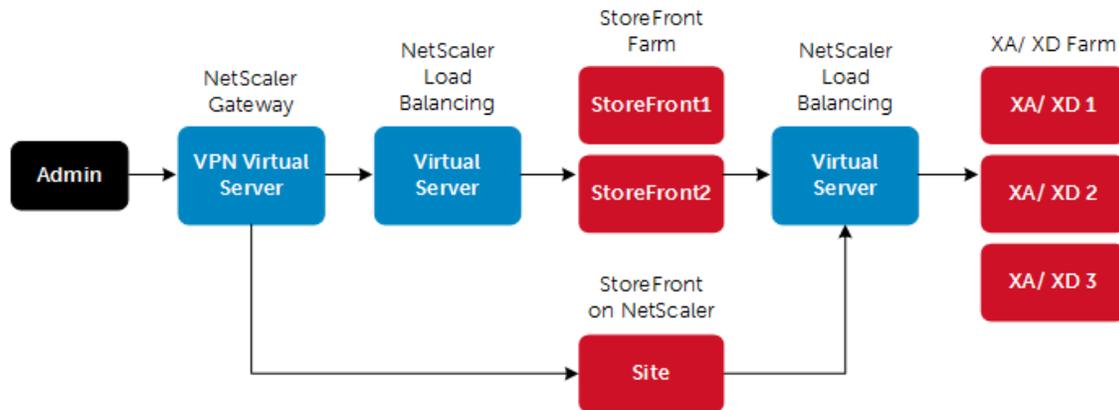
Global Server Load Balancing

GSLB is an industry standard function. It is in widespread use to provide automatic distribution of user requests to an instance of an application hosted in the appropriate data center where multiple processing facilities exist. The intent is to seamlessly redistribute load on an as required basis, transparent to the user community. These distributions are used on a localized or worldwide basis. Many companies use GSLB in its simplest form. They use the technology to automatically redirect traffic to Disaster Recovery (DR) sites on an exception basis. That is, GSLB is configured to simply route user load to the DR site on a temporary basis only in the event of a catastrophic failure or only during extended planned data center maintenance. GSLB is also used to distribute load across data centers on a continuous load balancing basis as part of normal processing.

XenDesktop HA with NetScaler White Paper: Link

Several of the management components of the XenDesktop stack are made highly-available using NetScaler to load balance traffic. The following management components require the use of a load balancer to function in a high availability mode:

- StoreFront Servers
- Licensing Server
- XenDesktop XML Service
- XenDesktop Desktop Director

- Provisioning Services TFTP Service
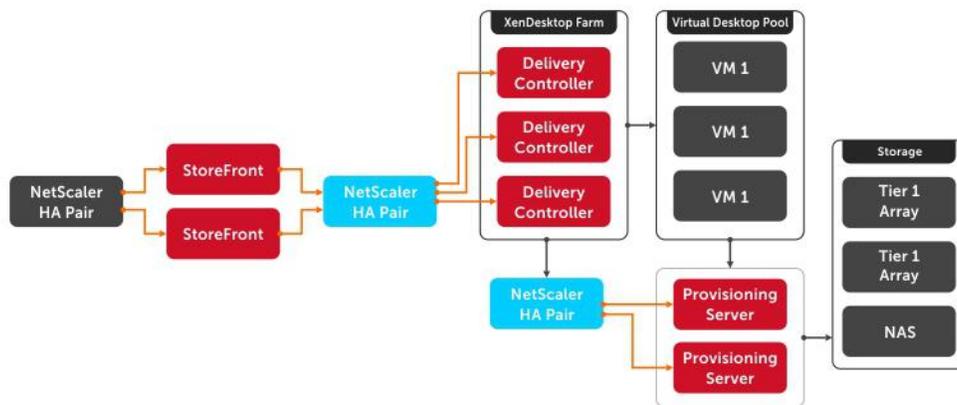
4.2 Hypervisor Platforms

4.2.1 VMware vSphere 6

The vSphere hypervisor also known as ESXi is a bare-metal hypervisor that installs directly on top of your physical server and partitions it into multiple virtual machines. Each virtual machine shares the same physical resources as the other virtual machines and they can all run at the same time. Unlike other hypervisors, all management functionality of vSphere is done through remote management tools. There is no underlying operating system, reducing the install footprint to less than 150MB.
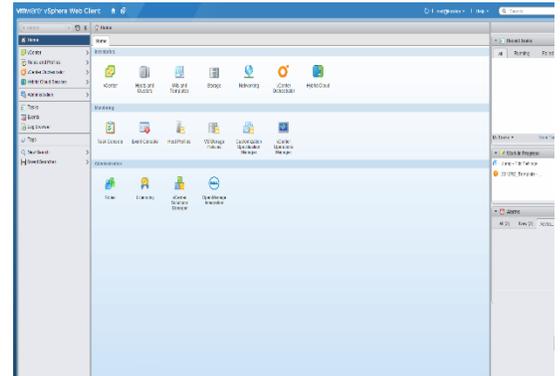
VMware vSphere 6 includes three major layers: Virtualization, Management and Interface. The Virtualization layer includes infrastructure and application services. The Management layer is central for configuring, provisioning and managing virtualized environments. The Interface layer includes the vSphere web client.

Throughout the Dell Wyse Datacenter solution, all VMware and Microsoft best practices and prerequisites for core services are adhered to (NTP, DNS, Active Directory, etc.). The vCenter 6 VM used in the solution is a single Windows Server 2012 R2 VM or vCenter virtual appliance, residing on a host in the management layer. SQL server is a core component of the Windows version of vCenter and is hosted on another VM also residing in the management layer. It is recommended that all additional XenDesktop components be installed in a distributed architecture, one role per server VM.

4.2.2 Microsoft Windows Server 2012 R2 Hyper-V

Windows Server 2012 R2 Hyper-V ™ is a powerful virtualization technology that enables businesses to leverage the benefits of virtualization. Hyper-V reduces costs, increases hardware utilization, optimizes business infrastructure, and improves server availability. Hyper-V works with virtualization-aware hardware to tightly control the resources available to each virtual machine. The latest generation of Dell servers includes virtualization-aware CPUs and network adapters.

From a network management standpoint, virtual machines are much easier to manage than physical computers. To this end, Hyper-V includes many management features designed to make managing virtual machines simple and familiar, while enabling easy access to powerful VM-specific management functions. The primary management platform within a Hyper-V based XenDesktop virtualization environment is Microsoft Systems Center Virtual Machine Manager SP1 (SCVMM).
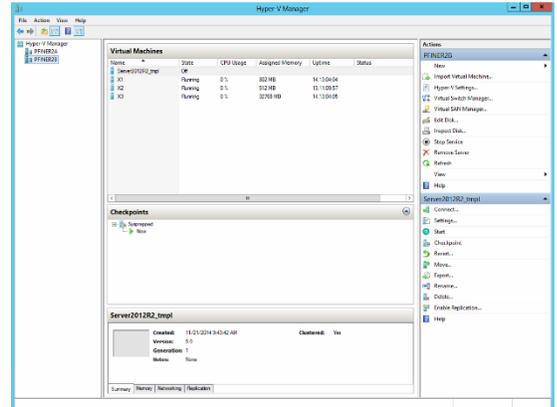
SCVMM provides centralized and powerful management, monitoring, and self-service provisioning for virtual machines. SCVMM host groups are a way to apply policies and to check for problems across several VMs at once. Groups are organized by owner, operating system, or by custom names such as "Development" or "Production". The interface also incorporates Remote Desktop Protocol (RDP); double-click a VM to bring up the console for that VM—live and accessible from the management console.

4.3 NVIDIA GRID vGPU

NVIDIA GRID vGPU™ brings the full benefit of NVIDIA hardware-accelerated graphics to virtualized solutions. This technology provides exceptional graphics performance for virtual desktops equivalent to local PCs when sharing a GPU among multiple users.

GRID vGPU™ is the industry's most advanced technology for sharing true GPU hardware acceleration between multiple virtual desktops—without compromising the graphics experience. Application features and compatibility are exactly the same as they would be at the user's desk.

With GRID vGPU™ technology, the graphics commands of each virtual machine are passed directly to the GPU, without translation by the hypervisor. This allows the GPU hardware to be time-sliced to deliver the ultimate in shared virtualized graphics performance.

VIRTUAL MACHINE

APPLICATIONS

NVIDIA GRAPHICS DRIVER
GUEST OS

NETWORK

NVIDIA GRID™ vGPU™

NVIDIA GRID™ SOFTWARE
HYPERVISOR

CLIENT

NVIDIA GPU

Between 1 and 16 concurrent clients/users per GPU
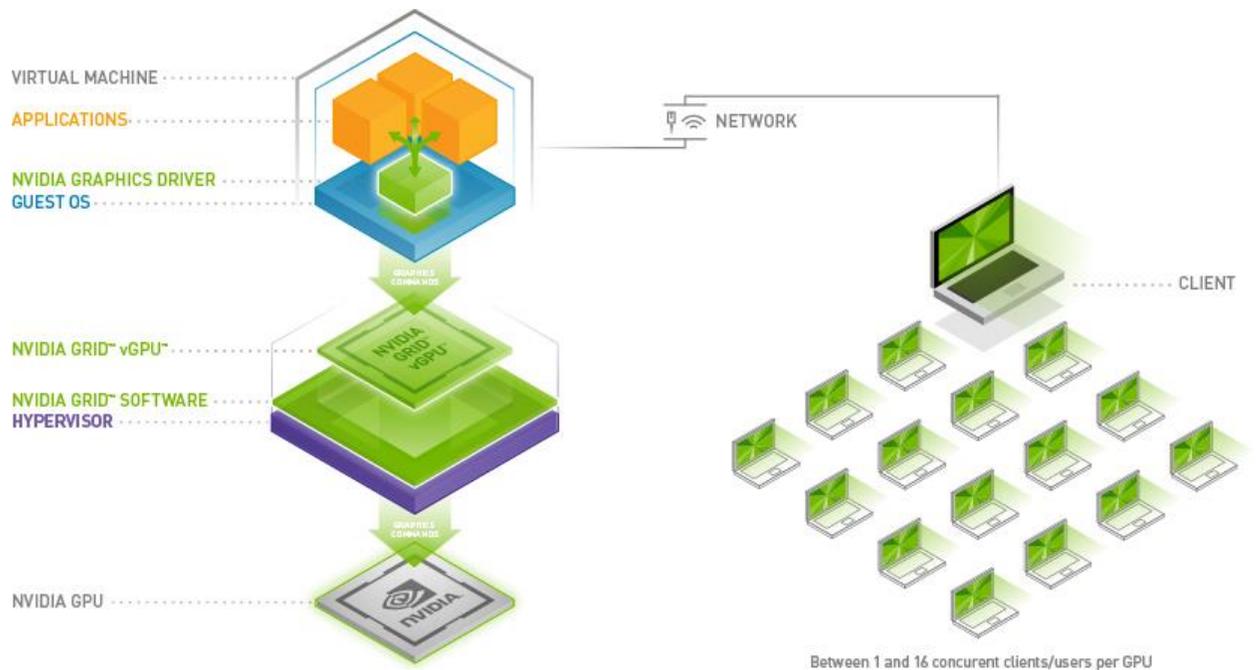
Image provided curtesy of NVIDIA Corporation, Copyright NVIDIA Corporation

4.3.1   vGPU Profiles

Virtual Graphics Processing Unit, or GRID vGPU™, is technology developed by NVIDIA® that enables hardware sharing of graphics processing for virtual desktops. This solution provides a hybrid shared mode allowing the GPU to be virtualized while the virtual machines run the native NVIDIA video drivers for better performance. Thanks to OpenGL support, VMs have access to more graphics applications. When utilizing vGPU, the graphics commands from virtual machines are passed directly to the GPU without any hypervisor translation. All this is done without sacrificing server performance and so is truly cutting edge.

The Dell graphics server platform is the award winning PowerEdge R730 which can accommodate two NVIDIA K1 or K2 graphics cards. The combination of Dell servers, NVIDIA vGPU technology and NVIDIA GRID cards enable high-end graphics users to experience high fidelity graphics quality and performance, for their favorite applications at a reasonable cost.

NOTE: GRID vGPU™ is supported on NVIDIA K1/K2 cards and is available as a licensed feature on NVIDIA Tesla M60 cards.

GRID K1 and K2 vGPU Profiles:

Card	vGPU Profile	Graphics Memory	Virtual Display Heads	Maximum Resolution	Physical GPUs	Maximum vGPUs		Intended User(s)
						Per GPU	Per Card	
GRID K2	K280Q	4GB	4	2560x1600	2	1	2	Designer
	K260Q	2GB	4	2560x1600		2	4	Designer
	K240Q	1GB	2	2560x1600		4	8	Power User
	K220Q	512MB	2	2560x1600		8	16	Power User
GRID K1	K180Q	4GB	4	2560x1600	4	1	4	Entry Designer
	K160Q	2GB	4	2560x1600		2	8	Power User
	K140Q	1GB	2	2560x1600		4	16	Power User
	K120Q	512MB	2	2560x1600		8	32	Power User

NVIDIA® Tesla® M60 GRID vGPU Profiles:

Card	vGPU Profile	Graphics Memory (Frame Buffer)	Virtual Display Heads	Maximum Resolution	Maximum vGPUs		Guest VM OS Support*		GRID License Required
					Per GPU	Per Card	Win	64bit Linux	
Tesla M60	M60-8Q	8GB	4	4096x2160	1	2	●	●	GRID Virtual Workstation
	M60-4Q	4GB	4	4096x2160	2	4	●	●	
	M60-2Q	2GB	4	4096x2160	4	8	●	●	
	M60-1Q	1GB	2	4096x2160	8	16	●	●	
	M60-0Q	512MB	2	2560x1600	16	32	●	●	
	M60-1B	1GB	4	2560x1600	8	16	●		GRID Virtual PC
	M60-0B	512MB	2	2560x1600	16	32	●		
	M60-8A	8GB	1	1280x1024	1	2	●		GRID Virtual Application
	M60-4A	4GB			2	4	●		
	M60-2A	2GB			4	8	●		
	M60-1A	1GB			8	16	●		

*NOTE: Windows guest operating systems supported with all profiles. 64-Bit Linux operating systems supported are listed in the table on the next page and are only supported by vGPU on the M60 cards.

DELL

Windows	Linux
Windows 7 (32/64-bit)	RHEL 6.6 & 7
Windows 8.x (32/64-bit)	CentOS 6.6 & 7
Windows 10 (32/64-bit)	Ubuntu 12.04 & 14.04 LTS
Windows Server 2008 R2	
Windows Server 2012 R2	

## 4.3.1.1 GRID vGPU Licensing and Architecture

NVIDIA GRID vGPU™ is offered as a licensable feature on Tesla M60 GPUs. vGPU can be licensed and entitled using one of the three following software editions.

NVIDIA GRID Virtual Applications	NVIDIA GRID Virtual PC	NVIDIA GRID Virtual Workstation
For organizations deploying XenApp or other RDSH solutions. Designed to deliver Windows applications at full performance.	For users who need a virtual desktop, but also need a great user experience leveraging PC applications, browsers, and high-definition video.	For users who need to use professional graphics applications with full performance on any device, anywhere.
Up to 2 displays supporting virtualized Windows applications	Up to 4 displays supporting Windows desktops, and NVIDIA Quadro features	Up to 4 displays supporting Windows or Linux desktops, NVIDIA Quadro, CUDA*, OpenCL* & GPU pass-through

*CUDA and OpenCL only supported with M60-8Q profile

The GRID vGPU Manager, running on the hypervisor installed via the VIB, controls the vGPUs that can be assigned to guest VMs. A properly configured VM obtains a license from the GRID license server during the boot operation for a specified license level. The NVIDIA graphics driver running on the guest VM provides direct access to the assigned GPU. When the VM is shut down, it releases the license back to the server.
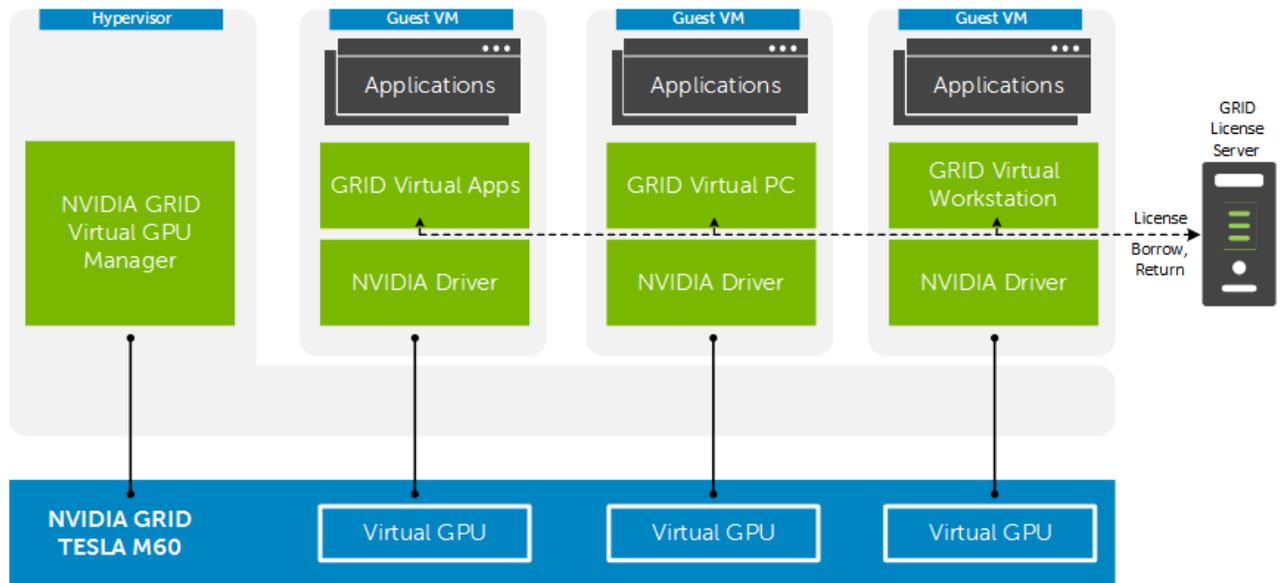
For more information about NVIDIA GRID vGPU, please visit: LINK

5   Solution Architecture for XenDesktop 7

5.1   Management Role Configuration

5.1.1   vSphere

The Management role requirements for the base solution are summarized below and take into account the use of PVS. If MCS provisioning is desired remove this from your build plan. Depending on the scale of the deployment, the number of these services must be adjusted. Use data disks for role-specific application files such as data, logs and IIS web files in the Management volume.

Role	vCPU	RAM (GB)	NIC	OS + Data vDisk (GB)	Tier 2 Volume (GB)
DDC + Lic	4	8	1	40 + 5	-
Storefront	2	4	1	40 + 5	-
vCenter	2	8	1	40 + 5	50 (VMDK)
PVS	4	24	1	40 + 5	100 (VMDK)
SQL Server	4	8	1	40 + 5	200 (VMDK)
File Server	2	4	1	40 + 5	2048 (RDM)
TOTALS	**18**	**56**	**6**	**270**	**2398**

5.1.2   NVIDIA GRID License Server Requirements

When using NVIDIA Tesla M60 cards, graphics enabled VMs must obtain a license from a GRID License server on your network to be entitled for vGPU. To configure, a virtual machine with the following specifications must be added to a management host in addition to the management role VMs.

Role	vCPU	RAM (GB)	NIC	OS + Data vDisk (GB)	Tier 2 Volume (GB)
NVIDIA GRID License Srv	2	4	1	40 + 5	-

GRID License server software can be installed on a system running the following operating systems:

- Windows 7 (x32/x64)
- Windows 8.x (x32/x64)
- Windows 10 x64
- Windows Server 2008 R2
- Windows Server 2012 R2
- Red Hat Enterprise 7.1 x64
- CentOS 7.1 x64

Additional license server requirements:

- A fixed (unchanging) IP address. The IP address may be assigned dynamically via DHCP or statically configured, but must be constant.
- At least one unchanging Ethernet MAC address, to be used as a unique identifier when registering the server and generating licenses in NVIDIA's licensing portal.
- The date/time must be set accurately (all hosts on the same network should be time synchronized).

5.1.3   Hyper-V

| Role | vCPU | Startup RAM (GB) | Dynamic Memory | | | NIC | OS + Data vDisk | |
			Min|Max	Buffer	Weight		Size (GB)	Tier2 Vol
DDC + Lic	4	8	384MB | 10GB	20%	Med	1	40	-
Storefront	2	4	384MB | 6GB	20%	Med	1	40	-
SCVMM	2	8	384MB | 10GB	20%	Med	1	40 + 5	50 (VHDX)
PVS	4	24	384MB | 24GB	20%	Med	1	40 + 5	100 (VHDX)

Primary SQL	4	8	384MB | 10GB	20%	Med	1	40 + 5	200 (VHDX)
File Server	2	4	384MB | 6GB	20%	Med	1	40 + 5	2048 (PTM)
Total	**18**	**56GB**	**2.3GB | 66GB**	**-**	**-**	**6**	**260GB**	**2398GB**

5.1.4   XenApp VM Configuration

The recommended number of XenApp VMs and their configurations on vSphere or Hyper-V are summarized below and take into account proper NUMA balancing assuming the CPU in use is the E5-2698v4. For more information on NUMA please refer to section 4.1.8.

Role	VMs per host	vCPUs per VM	RAM (GB)	NIC	OS vDisk Size (GB)
XenApp VM	10	8	32	1	80

5.1.5   SQL Databases

The Citrix, Microsoft and VMware databases are hosted by a single dedicated SQL 2012 R2 Server VM 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 are required for:

- Citrix XenDesktop
- vCenter or SCVMM

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 Citrix, Microsoft and 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).

5.1.6   DNS

DNS plays a crucial role in the environment not only as the basis for Active Directory but is used to control access to the various Citrix and Microsoft 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, Citrix 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.

5.1.6.1   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 is 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

The Dell Wyse Datacenter solution has a wide variety of Tier 1 and Tier 2 storage options to provide maximum flexibility to suit any use case. Customers have the choice to leverage best-of-breed iSCSI solutions from EqualLogic or Fiber Channel solutions from Dell Compellent while being assured the storage Tiers of the Dell Wyse Datacenter solution will consistently meet or outperform user needs and expectations.

5.2.1   Local Tier 1 Storage

Selecting the local Tier 1 storage model means that the compute host servers use 10 locally installed hard drives to house the user desktop VMs. In this model, Tier 1 storage exists as local hard disks or SSDs on the Compute hosts themselves. To achieve the required performance level, RAID 10 is recommended for use across all local disks. A single volume per local Tier 1 Compute host is sufficient to host the provisioned desktop VMs along with their respective write caches.

5.2.2 Shared Tier 1 Storage

Selecting the Shared Tier 1 model means that the virtualization compute hosts are deployed without Tier 1 local storage and leverage shared storage hosted on a high performance Dell Storage array. In this model, shared storage is leveraged for Tier 1 and used for VDI execution and write cache. Based on the heavy performance requirements of Tier 1 for VDI, it is recommended to use separate arrays for Tier 1 and Tier 2 when possible. We recommend using 500GB LUNs for VDI and running no more than 125 VMs per volume along with their respective write caches. Sizing to 500 basic users will require 4 x 500GB volumes.

Volumes	Size (GB)	Storage Array	Purpose	File System
VDI-1	500	Tier 1	125 x desktop VMs + WC	VMFS or NTFS
VDI-2	500	Tier 1	125 x desktop VMs + WC	VMFS or NTFS
VDI-3	500	Tier 1	125 x desktop VMs + WC	VMFS or NTFS
VDI-4	500	Tier 1	125 x desktop VMs + WC	VMFS or NTFS

5.2.3 Shared Tier 2 Storage

Tier 2 is shared iSCSI or FC storage used to host the Management server VMs and user data. EqualLogic 6210 arrays are used for Shared Tier 2 (Local Tier 1 only), or 65x0 series for larger deployments (up to 16 in a group), or a single CML array scaled up to 10K users. The Compellent Tier 2 array, as specified in section 3.3.2 scales simply by adding disks. The table below outlines the volume requirements for Tier 2. Larger disk sizes are chosen to meet the capacity needs of the customer. The user data are presented either via a file server VM using RDM/ PTD for small scale deployments or via NAS for large scale or HA deployments. The solution as designed presents all SQL disks using VMDK or VHDX formats. RAID 50 are used in smaller deployments but is **not** recommended for critical environments. The recommendation for larger scale and mission critical deployments with higher performance requirements is to use RAID 10 or RAID 6 to maximize performance and recoverability. The following depicts the component volumes required to support a 500 user environment. Additional Management volumes are created as needed along with size adjustments as applicable for user data and profiles.

Volumes	Size (GB)	Storage Array	Purpose	File System
Management	350	Tier 2	vCenter/ SCVMM, XD roles, File & SQL	VMFS/ NTFS
PVS	100	Tier 2	PVS vDisks	VMFS/ NTFS
User Data	2048	Tier 2	File Server/ NAS	RDM/PTD
User Profiles	20	Tier 2	User profiles	VMFS/ NTFS
SQL DATA	100	Tier 2	SQL	VMFS/ NTFS

SQL LOGS	100	Tier 2	SQL	VMFS/ NTFS
TempDB Data	5	Tier 2	SQL	VMFS/ NTFS
TempDB Logs	5	Tier 2	SQL	VMFS/ NTFS
Templates/ISO	200	Tier 2	ISO storage (optional)	VMFS/ NTFS

5.2.4   Storage Networking – EqualLogic iSCSI

Dell's iSCSI technology provides compelling price/performance in a simplified architecture while improving manageability in virtualized environments. Specifically, iSCSI offers virtualized environments simplified deployment, comprehensive storage management and data protection functionality, and seamless VM mobility. Dell iSCSI solutions give customers the "Storage Direct" advantage – the ability to seamlessly integrate virtualization into an overall, optimized storage environment.

If iSCSI is the selected block storage protocol, then the Dell EqualLogic MPIO plugin or MEM (Multipathing Extension Module) is installed on all hosts that connect to iSCSI storage. This module is added via a command line using a Virtual Management Appliance (vMA) from VMware. This plugin allows for easy configuration of iSCSI on each host. The MPIO plugin allows for the creation of new or access to existing data stores and handle IO load balancing. The plugin will also configure the optimal multi-pathing settings for the data stores as well. Some key settings to be used as part of the configuration:

● Specify 2 IP Addresses for iSCSI on each host

● Specify NICs

● Specify Jumbo Frames at 9000 MTU

● Initialize iSCSI initiator

● Specify IP for the EqualLogic Storage group.

## 5.2.5 Storage Networking – Compellent Fiber Channel (FC)

Based on Fluid Data architecture, the Dell Compellent Storage Center SAN provides built-in intelligence and automation to dynamically manage enterprise data throughout its lifecycle. Together, block-level intelligence, storage virtualization, integrated software and modular, platform-independent hardware enable exceptional efficiency, simplicity and security.

Storage Center actively manages data at a block level using real-time intelligence, providing fully virtualized storage at the disk level. Resources are pooled across the entire storage array. All virtual volumes are thin-provisioned. And with sub-LUN tiering, data is automatically moved between Tiers and RAID levels based on actual use.

If Fiber Channel is the selected block storage protocol, then the Compellent Storage Center Integrations for VMware vSphere client plug-in is installed on all hosts. This plugin enables all newly created data stores to be automatically aligned at the recommended 4MB offset. Although a single Fabric are configured to begin with to reduce costs, as a best practice recommendation, the environment is configured with two Fabrics to provide multi-pathing and end-to-end redundancy.

The following QLogic HBA BIOS settings are used:

- Set the "connection options" field to 1 for point to point only

- Set the "login retry count" field to 60 attempts

- Set the "port down retry" count field to 60 attempts

- Set the "link down timeout" field to 30 seconds

- Set the "queue depth" (or "Execution Throttle") field to 255

- This queue depth are set to 255 because the ESXi VMkernel driver module and DSNRO can more conveniently control the queue depth
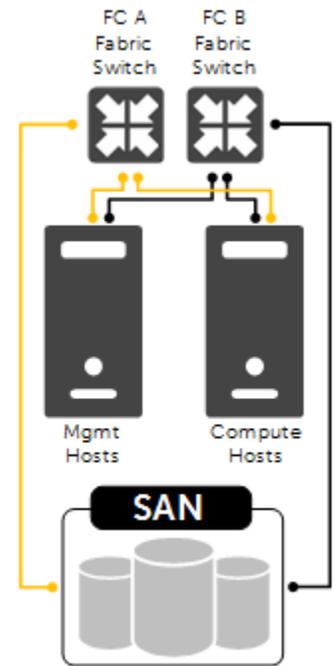
5.2.5.1 FC Zoning

Zone at least one port from each server HBA to communicate with a single Compellent fault domain. The result of this is two distinct FC Fabrics and four redundant paths per server as shown in the diagram below. Round Robin or Fixed Paths are supported. You can leverage Compellent Virtual Ports to minimize port consumption as well as simplify deployment. Zone each controller's front-end virtual ports, within a fault domain, with at least one ESXi initiator per server.

5.3 Virtual Networking

5.3.1 Local Tier 1 – iSCSI

The network configuration in this model will vary between the Compute and Management hosts. The Compute hosts do not need access to iSCSI storage since they are hosting VDI VMs on local disk. Since

the Management VMs are hosted on shared storage, they can take advantage of HA including Live Migration. The following outlines the VLAN requirements for the Compute and Management hosts in this solution model, applicable to rack or blade servers with HA:

- Compute hosts (Local Tier 1)

    o Management VLAN: Configured for hypervisor infrastructure traffic – L3 routed via core switch

    o VDI VLAN: Configured for VDI session traffic – L3 routed via core switch

- Management hosts (Local Tier 1)

    o Management VLAN: Configured for hypervisor Management traffic – L3 routed via core switch

    o Live Migration VLAN: Configured for Live Migration traffic – L2 switched only, trunked from Core (HA only)

    o Failover Cluster VLAN: Configured for Cluster and Cluster Shared Volume traffic – L2 switched only, trunked from core (Hyper-V only)

    o iSCSI VLAN: Configured for iSCSI traffic – L2 switched only via ToR switch

    o VDI Management VLAN: Configured for VDI infrastructure traffic – L3 routed via core switch

- A VLAN for iDRAC is configured for all hardware management traffic – L3 routed via core switch

Following best practices, LAN and block storage traffic is separated in solutions >500 users. This traffic is combined within a single switch in smaller stacks to minimize the initial investment, however, VLANs are required for each traffic type to enable traffic separation. Each Local Tier 1 Compute host will have a quad port NDC which includes both 10Gb and 1Gb interfaces. Configure the LAN traffic from the server to the ToR switch as a LAG.

5.3.1.1   vSphere

Standard vSwitches depicted below; however, dvSwitches should be used as desired for VM traffic especially in larger deployments to ease the management burden across numerous hosts. In the LT1 rack model, only the mgmt. hosts connect to shared storage so require additional VMK ports. Network share values should be configured equally among the VMKernel port groups that share a physical set of network adapters.
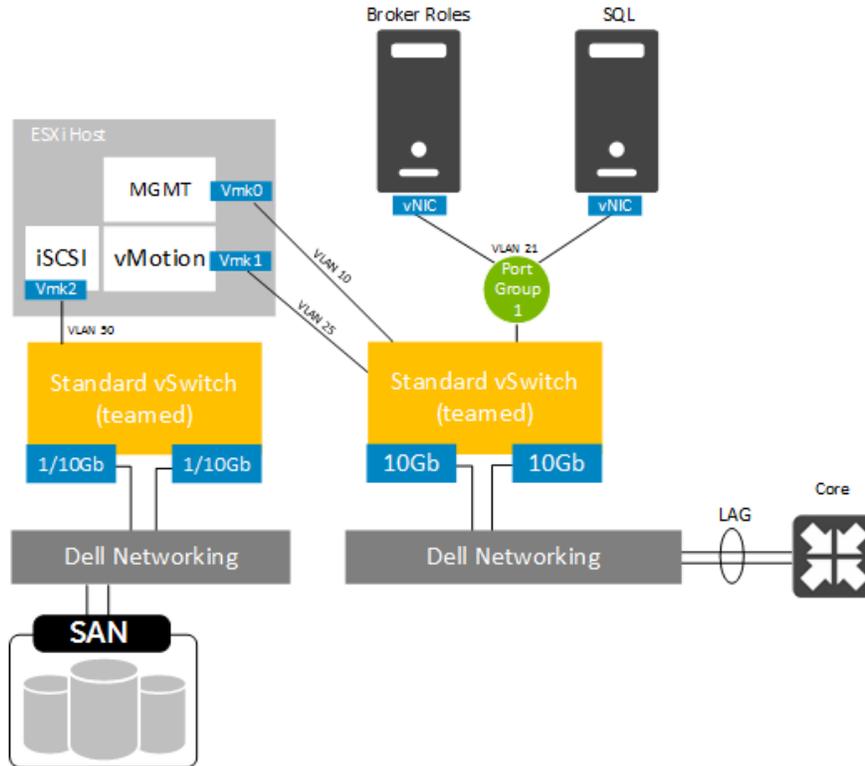
The Compute hosts are configured in the same basic manner, minus the shared storage, with the desktop VMs connecting to the primary port group on the external vSwitch.
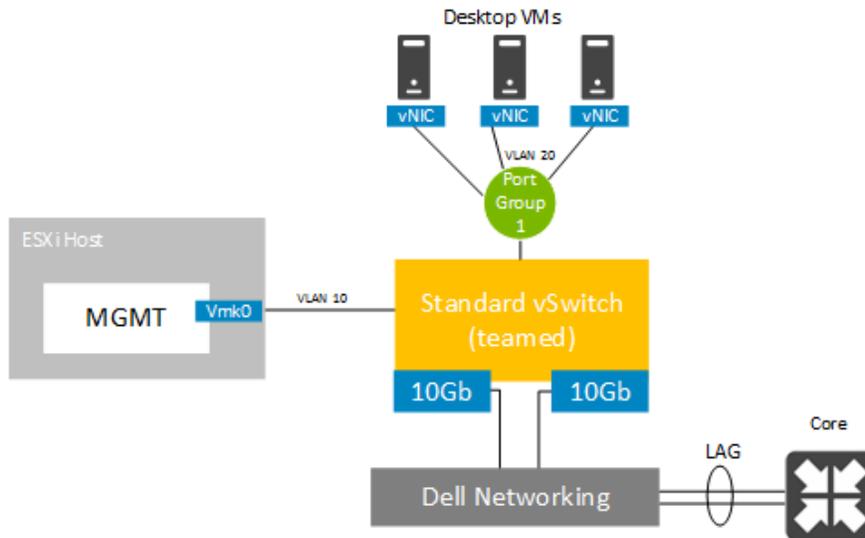

5.3.1.2   Hyper-V

The Hyper-V configuration, while identical in core requirements and hardware, is executed differently due to how Hyper-V and Windows Server 2012 R2 implement networking and virtual switches. As shown in the diagram below, native Windows Server 2012 R2 NIC Teaming is utilized to load balance and provide

resiliency for network connections. For the compute host in this scenario, a single LBFO NIC team is configured to connect to a Hyper-V switch for external traffic. All vNICs associated with the Management OS connect directly to the external Hyper-V switch.
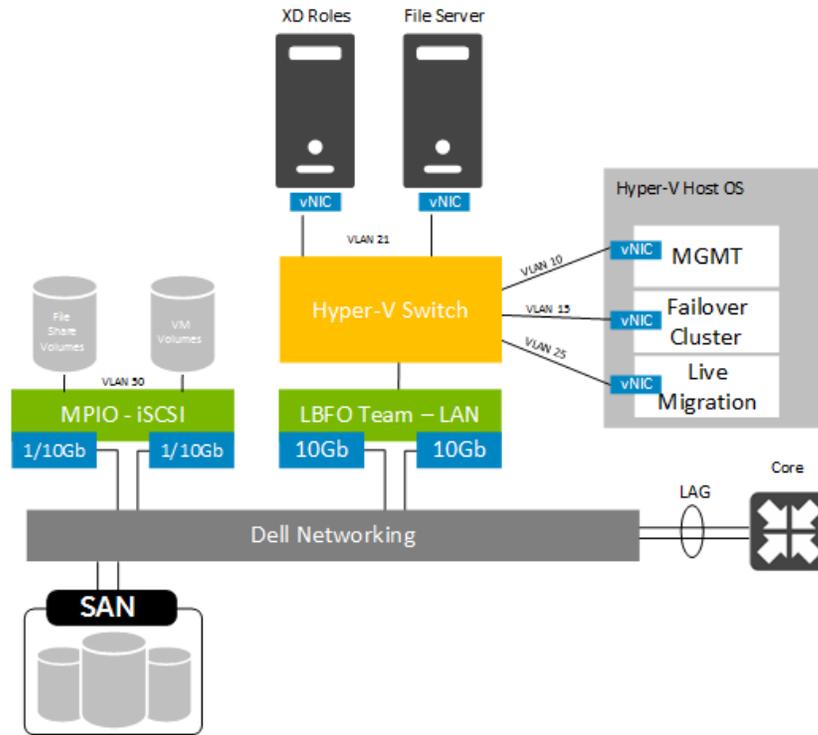
The dedicated compute host configuration is shown in the diagram below and configured similarly to the management host configuration but without the features enabled by shared storage.
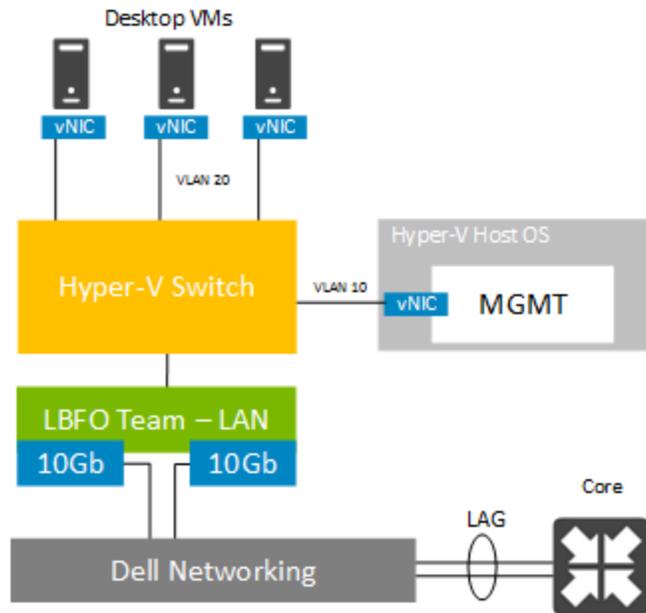
5.3.2 Shared Tier 1 – iSCSI

The network configuration in this model is identical between the Compute and Management hosts since all VMs are hosted on shared storage, whether T1 for desktops or T2 for Mgmt. The benefits of shared storage are available to all hosts such as Live Migration and HA. The following outlines the VLAN requirements for the Compute and Management hosts in this solution model:

- Compute hosts (Shared Tier 1)
    - o Management VLAN: Configured for hypervisor Management traffic – L3 routed via core switch
    - o Live Migration VLAN: Configured for Live Migration traffic – L2 switched only, trunked from Core
    - o Failover Cluster VLAN: Configured for Cluster and Cluster Shared Volume traffic – L2 switched only, trunked from core (Hyper-V only)
    - o iSCSI VLAN: Configured for iSCSI traffic – L2 switched only via ToR switch
    - o VDI VLAN: Configured for VDI session traffic – L3 routed via core switch

- Management hosts (Shared Tier 1)
    - o Management VLAN: Configured for hypervisor Management traffic – L3 routed via core switch
    - o Live Migration VLAN: Configured for Live Migration traffic – L2 switched only, trunked from Core
    - o Failover Cluster VLAN: Configured for Cluster and Cluster Shared Volume traffic – L2 switched only, trunked from core (Hyper-V only)
    - o iSCSI VLAN: Configured for iSCSI traffic – L2 switched only via ToR switch
    - o VDI Management VLAN: Configured for VDI infrastructure traffic – L3 routed via core switch

- An VLAN for iDRAC is configured for all hardware management traffic – L3 routed via core switch

Following best practices, iSCSI and LAN traffic is physically separated into discrete Fabrics. Each Shared Tier 1 Compute and Management host have a quad port NDC with 4 x 10Gb SFP+ ports each. iSCSI is isolated onto its own vSwitch with redundant ports to guarantee storage bandwidth. Configure the LAN traffic from the server to the ToR switch as a LAG.

5.3.2.1 vSphere

The same basic configuration applies to rack or blade servers although the physical NIC and switching components differ. Network share values should be configured equally among the VMkernel port groups that share a physical set of network adapters.
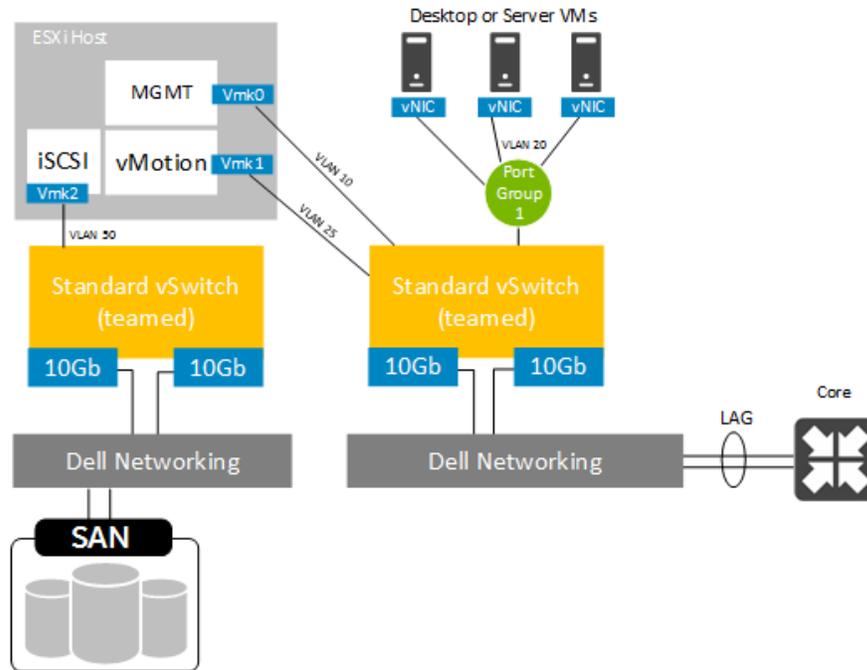
5.3.2.2 Hyper-V

The Hyper-V configuration, while identical in core requirements and hardware, is executed differently due to how Hyper-V and Windows Server 2012 R2 implement networking and virtual switches. As shown in the diagram below, native Windows Server 2012 R2 NIC Teaming is utilized to load balance and provide resiliency for network connections. For the compute or management hosts in this scenario, a single LBFO NIC team is configured to connect to a Hyper-V switch for external traffic. All vNICs associated with the Management OS connect directly to the external Hyper-V switch with MPIO used to connect to shared storage.
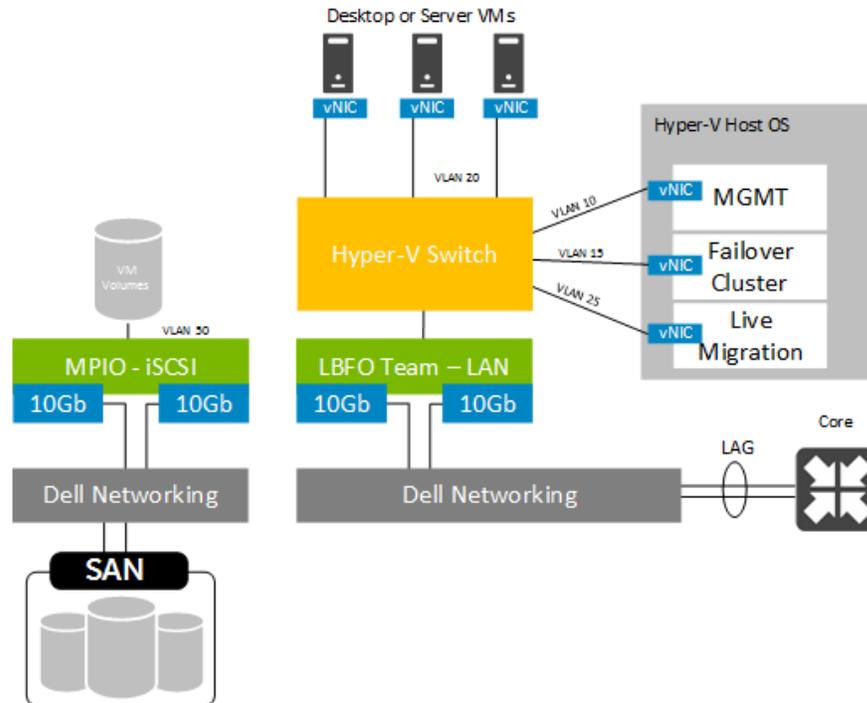
5.3.3 Shared Tier 1 – FC

Using Fiber Channel based storage eliminates the need to build iSCSI into the network stack but requires additional storage fabrics to be built out. The network configuration in this model is identical between the Compute and Management hosts. The benefits of shared storage are available to all hosts such as Live Migration and HA. The following outlines the VLAN requirements for the Compute and Management hosts in this solution model:

- Compute hosts (Shared Tier 1)

    o Management VLAN: Configured for hypervisor Management traffic – L3 routed via core switch
    o Live Migration VLAN: Configured for Live Migration traffic – L2 switched only, trunked from Core
    o Failover Cluster VLAN: Configured for Cluster and Cluster Shared Volume traffic – L2 switched only, trunked from core (Hyper-V only)

    o VDI VLAN: Configured for VDI session traffic – L3 routed via core switch
- Management hosts (Shared Tier 1)

    o Management VLAN: Configured for hypervisor Management traffic – L3 routed via core switch
    o Live Migration VLAN: Configured for Live Migration traffic – L2 switched only, trunked from Core
    o Failover Cluster VLAN: Configured for Cluster and Cluster Shared Volume traffic – L2 switched only, trunked from core (Hyper-V only)

    o VDI Management VLAN: Configured for VDI infrastructure traffic – L3 routed via core switch

- A VLAN for iDRAC is configured for all hardware management traffic – L3 routed via core switch

FC and LAN traffic are physically separated into discrete switching Fabrics. Each Shared Tier 1 Compute and Management host have a quad port NDC (4 x 10Gb) as well as 2 x 8Gb dual port FC HBAs. LAN traffic from the server to the ToR switch is configured as a LAG.

5.3.3.1   vSphere

The same basic configuration applies to rack or blade servers although the physical NIC and switching components differ. Network share values should be configured equally among the VMkernel port groups that share a physical set of network adapters.

5.3.3.2   Hyper-V

The Hyper-V configuration, while identical in core requirements and hardware, is executed differently due to how Hyper-V and Windows Server 2012 R2 implement networking and virtual switches. As shown in the diagram below, native Windows Server 2012 R2 NIC Teaming is utilized to load balance and provide resiliency for network connections. For the compute or management hosts in this scenario, a single LBFO NIC team is configured to connect to a Hyper-V switch for external traffic. All vNICs associated with the Management OS connect directly to the external Hyper-V switch with MPIO used to connect to shared storage.

Desktop or Server VMs

vNIC | vNIC | vNIC

VLAN 20

Hyper-V Host OS

VLAN 10 → vNIC MGMT

VLAN 15 → vNIC Failover Cluster

VLAN 25 → vNIC Live Migration

VM Volumes

Hyper-V Switch

VLAN 50

MPIO - FC

8Gb | 8Gb

LBFO Team – LAN

10Gb | 10Gb

Core

LAG

Brocade

Dell Networking

SAN
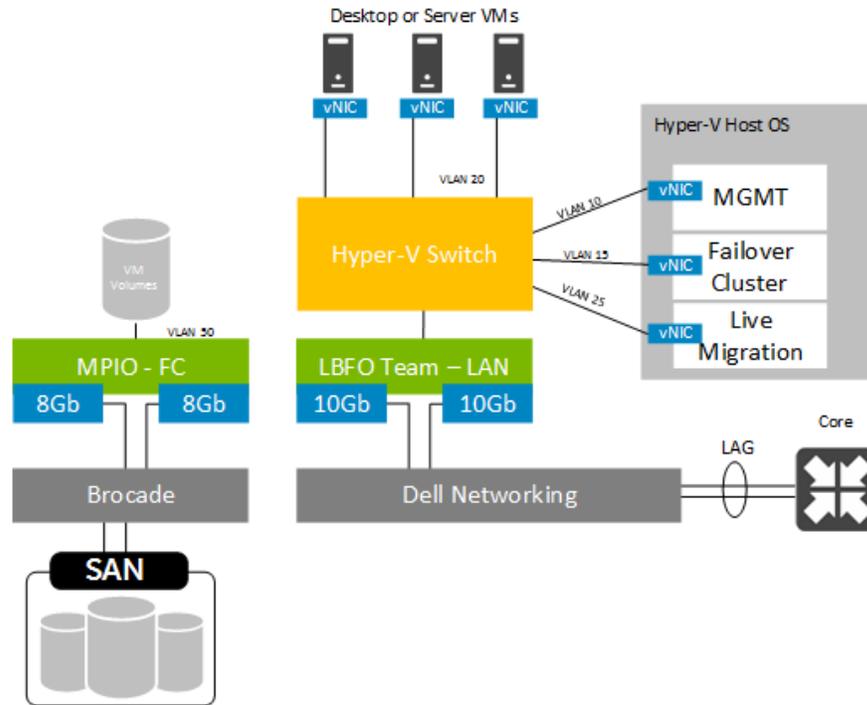
5.4　Scaling Guidance

- The components are 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
Compute Servers	Desktop VMs per physical host based on available CPU	Additional hosts and clusters added as necessary	Additional RAM or CPU compute power
Mgmt Servers	Number of server VMs per host	Add additional hosts	Add RAM or network adapters
Provisioning Servers	Desktops per instance	Additional servers added to the Provisioning Server farm	Additional network and I/O capacity added to the servers
Desktop Delivery Servers	Desktops per instance (dependent on SQL performance as well)	Additional servers added to the XenDesktop Site	Additional virtual machine resources (RAM and CPU)
XenApp Servers	Desktops per instance	Additional virtual servers added to the XenDesktop Site	Additional physical servers to host virtual XenApp servers.
Storefront Servers	Logons/ minute	Additional servers added to the Storefront environment	Additional virtual machine resources (RAM and CPU)
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 architecture solution layer, individually if desired. Following the N+1 model, additional ToR switches 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 or Hyper-V clustering is introduced in both the management and compute layers, SQL is

configured with AlwaysOn and NetScaler is leveraged for load balancing. Storage protocol switch stacks and NAS selection will vary based on chosen solution architecture.

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

- Additional switches added to the existing thereby equally spreading each host's network connections across multiple switches.

- Additional ESXi or Hyper-V hosts added in the compute or mgmt layers to provide N+1 protection.

- Applicable Citrix infrastructure server roles are duplicated and spread amongst mgmt host instances where connections to each are load balanced via the addition of virtual NetScaler appliances.

- Each PVS Server instance is responsible for the Write Cache of each desktop that it is hosting. This Write Cache is readable by the resilient Provisioning Server and as it is held on the target device. In the event of a Provisioning Server failure, all desktops that were hosted will transfer to an alternate provisioning server for that site and users will be unaware of the failure.

- SQL Server databases also are protected through the addition and configuration of an "AlwaysOn" Failover Cluster Instance or Availability Group.

5.5.1   SQL Server High Availability

HA for SQL is provided via AlwaysOn using either Failover Cluster Instances or Availability Groups. This configuration protects all critical data stored within the database from physical server as well as virtual server problems. DNS is used to control access to the primary SQL instance. Place the principal VM that will host the primary copy of the data on the first Management host. Additional replicas of the primary database are placed on subsequent Management hosts.
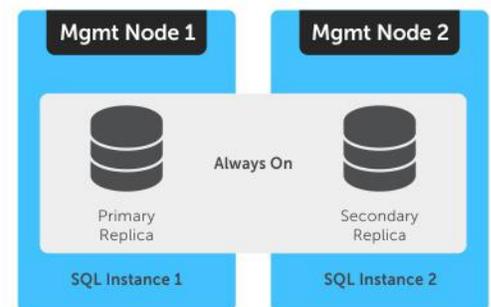
Please refer to these links for more information: LINK1 and LINK2

6   Dell Wyse Datacenter Architecture
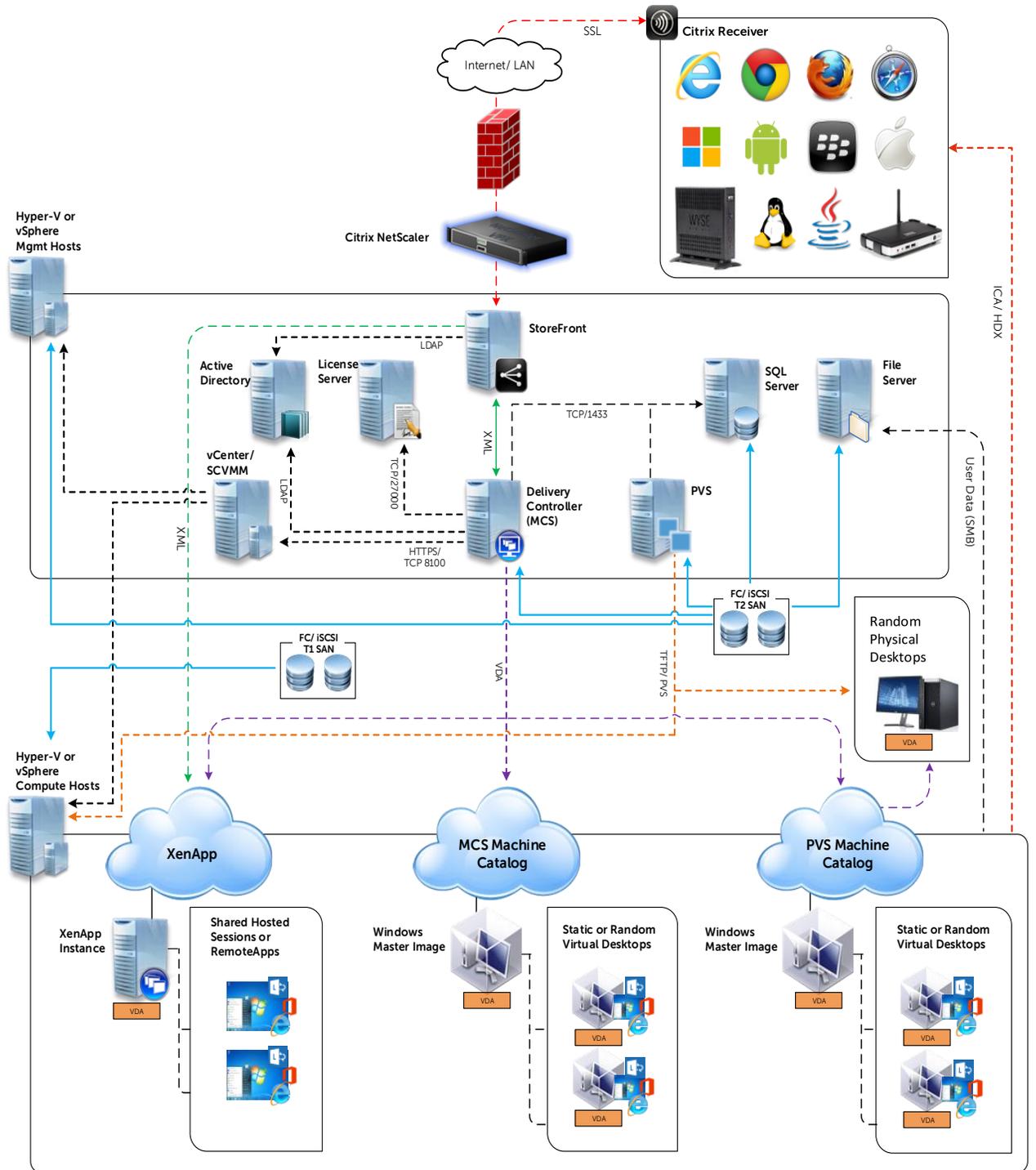
7 Solution Performance and Testing

At the time of publication here are the available appliance density recommendations.

Hypervisor	Provisioning	Workload	Template OS	Config	User Density
Hyper-V	XenDesktop MCS	Standard	Windows 8.1	Large Appliance	245
Hyper-V	XenDesktop MCS	Enhanced	Windows 8.1	Large Appliance	200
Hyper-V	XenDesktop MCS	Professional	Windows 8.1	Large Appliance	150
Hyper-V	XenApp	Enhanced	Server 2012 R2	Large Appliance	340
Hyper-V	XenDesktop MCS	Standard	Windows 8.1	Small Appliance	95
Hyper-V	XenDesktop MCS	Enhanced	Windows 8.1	Small Appliance	75
Hyper-V	XenDesktop MCS	Professional	Windows 8.1	Small Appliance	55
Hyper-V	XenApp	Enhanced	Server 2012 R2	Small Appliance	170
ESXi	Linked Clone	Standard	Windows 8.1	Local Tier 1	330
ESXi	Linked Clone	Enhanced	Windows 8.1	Local Tier 1	240
ESXi	Linked Clone	Professional	Windows 8.1	Local Tier 1	180

For detailed up-to-date validation results and analysis of these reference designs and more, please visit: LINK

Acknowledgements

About the Authors

Peter Fine is the Chief Architect and CTO of EUC Enterprise Engineering at Dell. Peter owns the strategy, architecture and leads the engineering of the datacenter EUC product and solutions portfolio. Follow Peter @ExitTheFastLane or www.ExitTheFastLane.com.

Jerry Van Blaricom is a Lead Architect in the Cloud Client Solutions Engineering Group at Dell. Jerry has extensive experience with the design and implementation of a broad range of enterprise systems and is focused on making Dell's virtualization offerings consistently best in class.

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.