Technical white paper

HP VirtualSystem VS2 for Microsoft

Reference architecture for building a Microsoft Hyper-V environment on HP Converged Infrastructure with integrated management

Table of contents

Executive summary

The HP VirtualSystem VS2 for Microsoft® Reference Architecture (VS2 RA) is designed to conform to the Microsoft Private Cloud Fast Track reference architecture and reduces data center complexity, increases agility, and mitigates the risks that come with deployment of a virtualized environment. The VS2 RA design leverages best-in-class HP Converged Infrastructure and Microsoft Hyper-V – the industry's fastest growing hypervisor and the best choice for virtualizing Microsoft applications and workloads. At its core, the VS2 RA provides a foundation for building a high-performance Microsoft Hyper-V virtualization platform that has been optimized to consolidate and provision hundreds to thousands of workloads while providing extremely high availability at all levels – from the underlying network and storage fabrics up to the virtual machine (VM) layer.

The VS2 RA solution is built on the HP Converged Infrastructure, including the HP BladeSystem architecture, HP Virtual Connect Flex-10, and HP LeftHand 4800 SAN Solution for BladeSystem. By designing on the HP BladeSystem, the VS2 RA can be sized and scaled in a modular fashion, simplifying scaling up and out while reducing the footprint of the solution and environmental requirements through advanced power and thermal capabilities. HP Virtual Connect provides the converged fabric for the VS2 RA and the ability to specifically allocate network ports and associated bandwidth per workload requirements – a very important capability when building a virtualization platform due to the complex networking requirements. Coupling these technologies with the HP LeftHand 4800 SAN Solution for BladeSystem (4800), the VS2 RA provides an extremely dense platform for the deployment of virtualized environments that require high levels of storage performance.

The VS2 RA is designed to give customers the flexibility to implement a solution themselves or to leverage HP's or a partner's services to assemble the hardware and even implement the solution on-site. The combination of an optimized architecture with services implementation makes it possible to speed deployment, quickly provision workloads, simplify management, and ultimately reduce IT costs. This document describes the benefits of the VS2 RA, describes the solution architecture, and describes a reference example for building out a Microsoft Private Cloud reference architecture solution.

Target audience: This document is intended for technical decision-makers and solution architects.

The virtualization reality

Virtualization of IT resources delivers a wide range of benefits. It can help an IT organization increase agility, operate more efficiently, cut costs, simplify application deployment and management, improve availability, and establish a foundation for a cloud environment. Yet despite its many benefits, this now-dominant computing paradigm has also introduced greater complexity into IT environments.

To capture the full benefits of virtualization for Microsoft environments, IT organizations need to do much more than simply create and manage virtual machines. They must go beyond the challenges associated with integrating virtualization into an existing environment and configuring storage area networks (SANs) to work in a virtual realm. And they must find the right tools and approaches for managing virtualized applications and operations, maintaining a secure environment, scaling beyond the initial configuration, and evolving to a private cloud. The ability to deploy new applications faster and keep them up and running more reliably is the central mission of IT and a competitive differentiator. To gain a real edge, you need to go beyond just managing infrastructure. You need to manage applications, deploy them faster, and keep applications up and running more reliably.

With the VS2 RA, HP and Microsoft help organizations break through the integration challenges of virtualization and accelerate the journey to cloud optimized converged infrastructure.

Benefits

There are a number of benefits that the VS2 RA design provides for organizations looking to deploy a virtualized environment. These benefits help to drive down both acquisition and operational costs and reduce your total cost of ownership (TCO).

Time to value

The VS2 RA is based on a pre-tested, validated architecture design, thus reducing the time-consuming tasks associated with designing, testing, and certifying the solution on-site.

Optimized and balanced architecture

When designing virtualized environments with varying workloads it can often be a challenge to ensure there is sufficient I/O capacity in the design to meet the requirements of the workloads while also efficiently utilizing the other server and network resources. One of the significant driving factors for customers looking at virtualization initiatives is that there is often a high percentage of servers in the data center that run at very low CPU utilizations, consuming energy and floor space while performing very little work.

The VS2 RA platform has been designed from the ground up to effectively utilize all of the resources (processing, network, I/O and capacity) required by varying consolidated workloads.

Enhanced efficiency & high availability

By standardizing platforms and system configurations, you can enhance IT efficiency and enable automation. The portability of VMs can enhance disaster recovery and business continuity.

Flexible

Building on the HP Converged Infrastructure, the VS2 RA provides a phased growth design, allowing easy expansion of storage and/or compute nodes to improve I/O and processing power as needed. With HP Virtual Connect Flex-10 technology, the VS2 RA utilizes a single fabric that can be configured to meet the specific requirements for virtualization. It provides the flexibility to define individual networks and allocate bandwidth to those networks according to the utilization and availability requirements, while dramatically reducing the cabling and wiring complexity and associated fabric costs.

The VS2 RA design can easily scale from a 4-blade and single 4800 configuration to a 10-blade design with three 4800s. Furthermore, as extra resources are required, additional racks can be incorporated into the management envelope and grouped together as new resource pools.

The VS2 RA also provides the flexibility to build this solution using your own in-house IT staff or engage with experienced HP consultants to customize and tailor the design to meet the demands of your business.

Simplified management

With HP Insight Control for Microsoft System Center (IC-SC) customers get deeper insight and monitoring control into the hardware from System Center Operations Manager (SCOM). IC-SC provides hardware alerting and event information in a single pane of view within System Center Operations Manager. Additionally, with System Center Virtual Machine Manager (SCVMM), customers can easily deploy, migrate, and manage virtual machines (VM) from a single pane of glass, as well as script multi-system activities.

Foundation for the Cloud

The VS2 RA is an application-centric design that supports Microsoft virtualization and Microsoft Private Cloud. HP Insight Control with Microsoft System Center enables comprehensive monitoring of Microsoft Private Cloud deployments. The VS2 RA allows organizations to consume cloud on their terms providing the choice and flexibility of a hybrid cloud model through common management, virtualization, identity and developer tools.

VS2 RA overview

As virtualized environments may have multiple different workloads, sizing them appropriately can be a challenge. However, HP has carefully considered many general performance needs of solutions with varying workloads, and has designed an innovative, enterprise-class solution that can be used to consolidate and provision multiple different systems and applications. The result is a fault-tolerant solution built on HP Converged Infrastructure.

This document provides specific details about the configuration and bill of materials for the HP VS2 RA based on the HP ProLiant BL460c Gen8 server blade (BL460c) and the HP LeftHand 4800 G2 42TB SAS SAN Solution for BladeSystem (4800) storage. This reference architecture accommodates both I/O-intensive and CPU-intensive workloads and can be easily modified to increase storage or compute capabilities.

VS2 RA components

HP BladeSystem

The BladeSystem c7000 enclosure provides all the power, cooling, and I/O infrastructure needed to support modular server, interconnect, and storage components today and throughout the next several years. The enclosure is 10U high and holds up to 16 server and/or storage blades plus optional redundant network and storage interconnect modules. It includes a shared, 5 terabit per second high-speed NonStop midplane for wire-once connectivity of server blades to network and shared storage.

HP ProLiant servers

The latest generation of HP blade server technology, the HP ProLiant Generation 8 servers, features embedded automation and intelligence that cut lifecycle operations tasks, which in turn reduces overhead and downtime costs. These new features and capabilities help further simplify management of the VS2 RA, help lower compute costs, and increase performance. These include HP SmartMemory, HP Integrated Lights-Out 4 (iLO 4), HP Active Health System, HP Agentless management, and HP Intelligent Provisioning features.

HP Virtual Connect Flex-10

HP Virtual Connect Flex-10 technology creates a dynamically scalable internal network architecture for virtualized deployments. The c7000 enclosure in the VS2 RA contains two Virtual Connect (VC) Flex-10 interconnect modules, and each module connects to a dual port 10Gb Flex-10 adapter in each server. Each Flex-10 adapter has four FlexNICs on each of its dual ports. The multiple FlexNICs can support the iSCSI storage, specialized VM function, management and production networks recommended by HP for Hyper-V deployments.

HP 3Gb SAS BL Switch

HP 3Gb SAS BL Switch for HP BladeSystem c-Class enclosures (SAS BL Switch) allows blades in the enclosure to connect directly and easily to SAS storage. It enables a straight forward external zoned SAS or shared SAS storage solution and is a key component of the VS2 RA. The 3Gb SAS BL Switch is required to use 4800 storage, as it provides the storage blades connectivity to the drive enclosures in a high-performance, highly-available configuration.

HP Networking

HP Networking offers a range of products, from low-cost switches to highly available and modular models. High performance top-of-rack (ToR) switches are a powerful part of the HP FlexNetwork portfolio. With the added benefit of Intelligent Resilient Framework (IRF), which allows multiple switches to appear logically as a single device, these devices provide extra availability and bandwidth.

The VS2 RA takes advantage of two switch options: a low-cost HP 5120-16G SI Switch (5120) for non-critical, non-redundant connections; and the high-performing, highly available HP 5920AF-24XG Switch (5920) for the remaining critical connections with IRF.

HP Storage

The HP LeftHand 4800 SAN for BladeSystem delivers storage that is tightly integrated with HP BladeSystem and offers a dense storage solution that dynamically and linearly scales as the infrastructure expands. Being converged into the HP BladeSystem architecture enables simplified administration, improved security, flexibility, and enhanced performance through network convergence of the virtual server and storage fabric using 10Gb Ethernet and Virtual Connect Flex-10.

Solution design

The VS2 RA was tested and can be configured as a Microsoft Windows® Failover Cluster resource with Microsoft Hyper-V enabled as the virtualization technology. This provides the infrastructure resource pool to provision hundreds of VM instances for consolidation and new VM provisioning.

The solution is composed of storage (4800) and compute (BL460c) building blocks, and consists of two main configurations ("base" and "extended") as summarized in the bulleted lists below. Both configurations have been balanced for roughly 1500 small-block, random IOPS per BL460c. However, as each virtualization environment has varying applications and workloads, storage and compute blocks may be added independently as appropriate, ranging between the base and extended configurations detailed below and shown in figures 1 and 2. If more resources are still needed, additional VS2 RA racks can be added and integrated. For more performance information and some guidance on sizing the environment, please see the Performance, sizing, and scalability section of this paper.

Base configuration

- 4 x HP ProLiant BL460c Gen8 Servers for virtualization (a total of 64 cores and 512 GB or optionally 1 TB RAM)
- 1 x HP 4800 G2 42TB SAS SAN Solution for BladeSystem (42 TB raw storage, 70 spindles)
- 2 x HP ProLiant DL360p Gen 8 Servers for management (a total of 32 cores and 256GB RAM)
- 2 x HP 5920AF-24XG Switches
- 1 x HP 5120-16G SI Switch

Extended configuration

- 10 x HP ProLiant BL460c Gen8 Servers for virtualization (a total of 160 cores and 1.25 TB or optionally 2.5 TB RAM)
- 3 x HP 4800 G2 42TB SAS SAN Solution for BladeSystem (126 TB raw storage, 210 spindles)
- 2 x HP ProLiant DL360p Gen 8 Servers for management (a total of 32 cores and 256GB RAM)
- 2 x HP 5920AF-24XG Switches
- 1 x HP 5120-16G SI Switch

Figure 1. VS2 RA Base Front and Rear Views

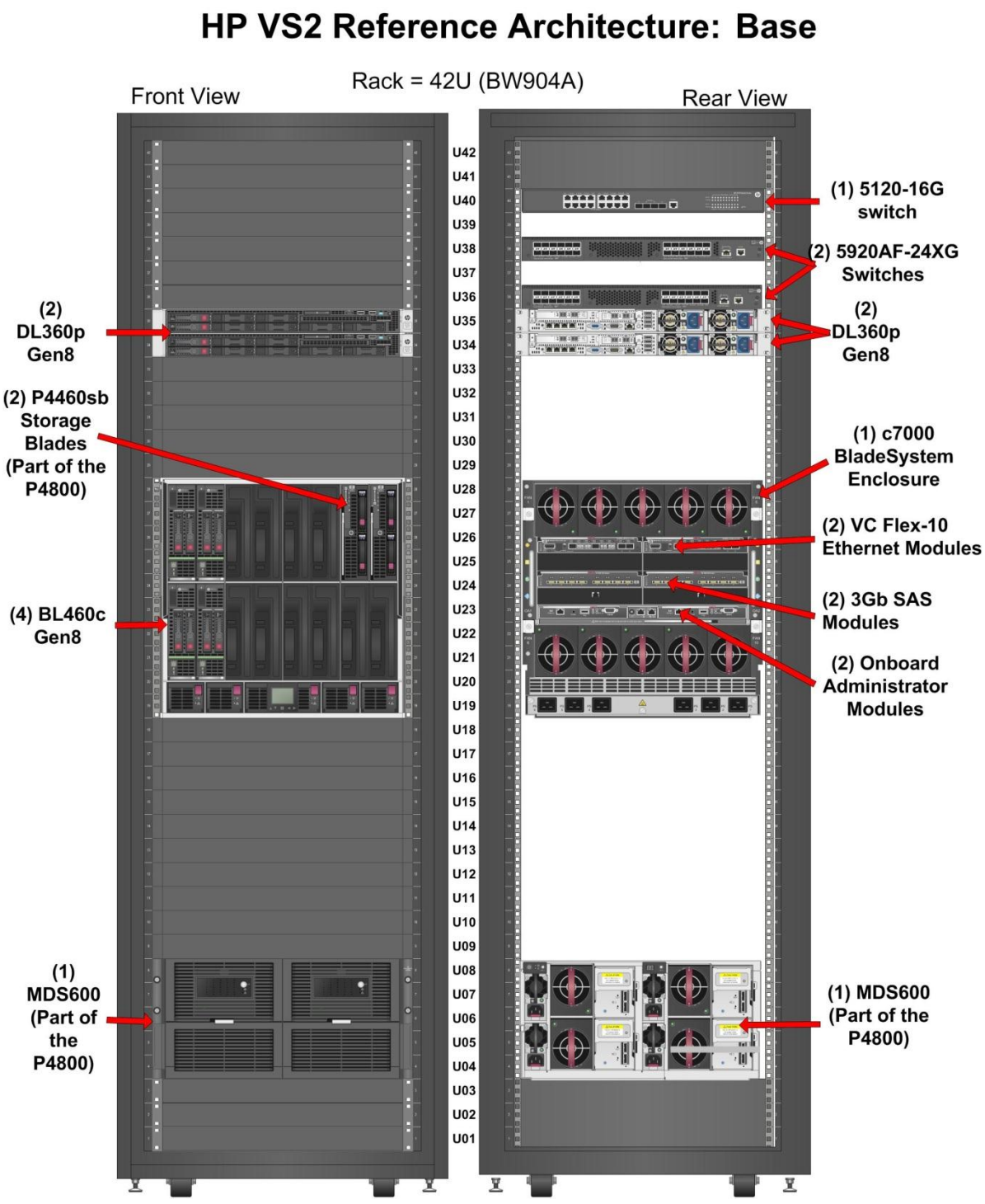
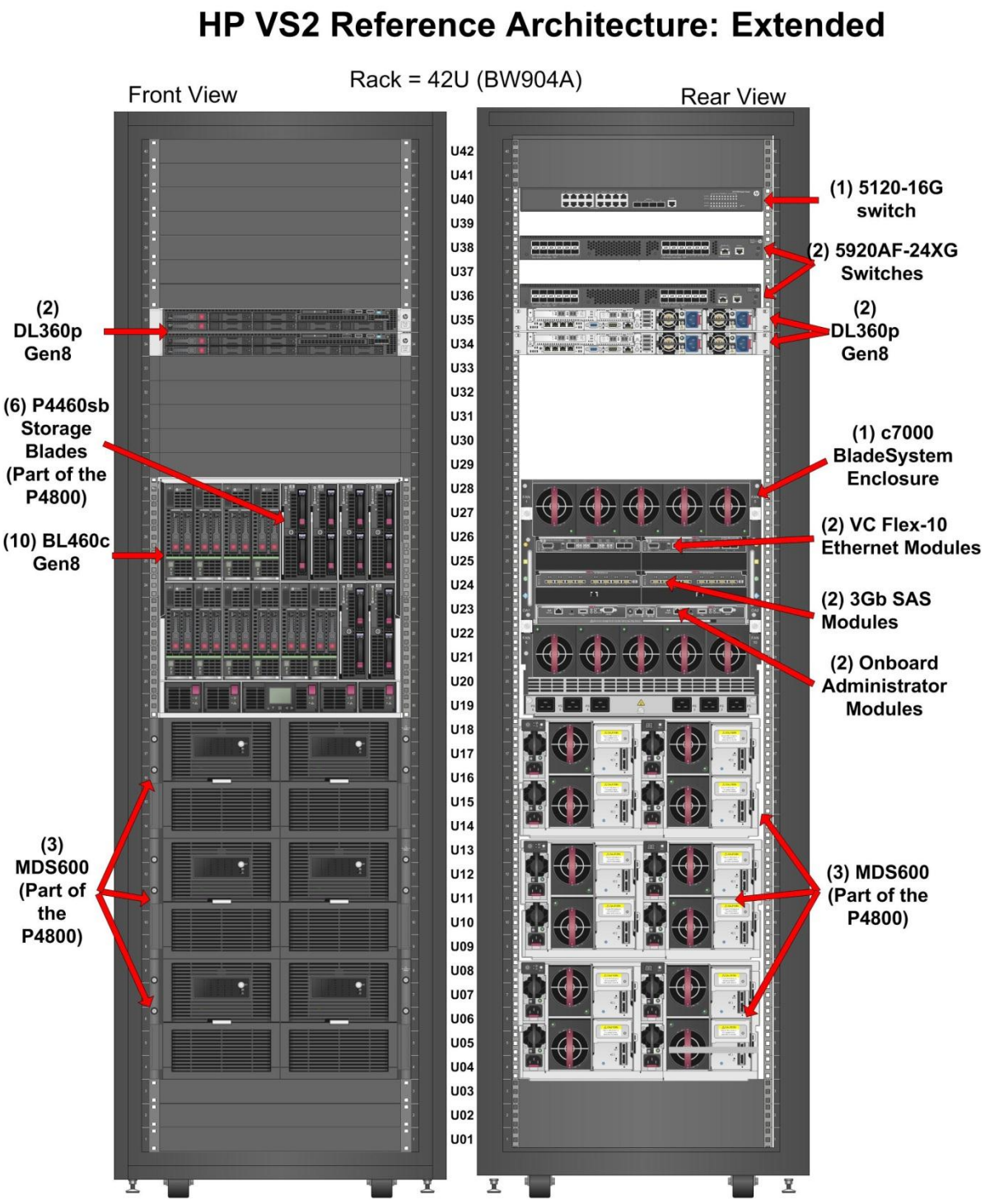


Figure 2. VS2 RA Extended Front and Rear Views



Servers

The server building block consists of HP ProLiant BL460c Gen8 server blades deployed within an HP BladeSystem c7000 enclosure (with a minimum of 4 blades). Each blade has 2 Intel® Xeon® E5-2670 Processors (totaling 32 logical CPU cores when Hyper-Threading is enabled), 128 GB or 256 GB RAM, and two 300 GB hard disk drives (HDDs).

In the VS2 RA testing, each server was running Microsoft Windows Server 2008 R2 SP1 Datacenter edition with Hyper-V enabled. With the Datacenter edition, users have the right to run an unlimited number of VMs on each server. This will be the best option when running the VS2 RA given the anticipated number of VMs that can be consolidated onto this platform.

Along with the Hyper-V role, Failover Clustering is also recommended on each server. Each additional rack then becomes its own Failover Cluster resource that can be managed as additional resources with the System Center Virtual Machine Manager environment.

In addition to the BL460c compute nodes, two HP ProLiant DL360p Gen8 (DL360) servers provide management of the VS2 RA rack. HP also recommends clustering these servers and running Hyper-V with several VMs running the management software used in the environment. Each DL360 has 16 CPU cores, 128GB RAM and two 300GB HDDs.

Storage

Storage is also designed in a building block fashion for the VS2 RA solution, with each core storage building block consisting of an HP 4800 G2 42TB SAS SAN Solution for BladeSystem. The 4800 itself is composed primarily of one HP 600 Modular Disk System (MDS600) enclosure with seventy 600GB 15K SAS disk drives, and two redundant active-active scalable P4460sb G2 Storage Blades for BladeSystem for c-Class Enclosure.

For a base configuration there is one 4800 block and for an extended configuration there are three storage blocks. The storage fabric chosen in the VS2 RA design is 10Gb iSCSI. This specifically provides capability for configuring guest clusters between two VMs (requiring direct access to LUNs on the storage array). This capability enables a higher degree of availability.

The 4800 allows multiple levels of redundancy and high availability. Each half of the MDS600 drive enclosure can be setup in a RAID configuration (RAID5 by default). Additionally, the two halves are mirrored using Network RAID 10 to further protect against drive, power, and controller failure.

Customers can also extend the VS2 RA by connecting to existing iSCSI SAN environments for additional storage capacity if needed.

Networking

When designing Hyper-V host servers, there are a number of distinct networks that are required for the solution to be in accordance with accepted best practices, and also to meet Microsoft Fast Track compliance.

Thus, the following networking design is recommended for each server in the VS2 RA:

- Teamed production (Prod) networks
- Teamed management (Mgt) networks
- Live VM migration (LM) network
- Cluster Shared Volume (CSV) network
- Redundant iSCSI networks for Multipath I/O (MPIO)

The VS2 RA is designed to follow the above recommendations, providing redundant connectivity to each of the major network segments that are used in a Microsoft Hyper-V deployment.

The management and guest (production) networks use NIC teaming for improved redundancy, whereas the Live Migration and CSV networks are set up to fail over to each other in case of a failed connection.

The combination of the HP 4800 SAN solution and HP ProLiant blade servers in the same c7000 enclosure allows iSCSI traffic to traverse the backplane of the enclosure, reducing the impact of iSCSI traffic on the switches at the top of the rack. Each configuration should also use MPIO drivers for the redundant iSCSI networks as recommended above.

The HP Virtual Connect Flex-10 10 Gb Ethernet Module for c-Class BladeSystem (Flex-10 Module) reduces cabling required and provides valuable networking flexibility with profiles that can be designed for the networking needs of each blade. Figures 3 and 4 below show the VS2 RA network design and BL460c connectivity configuration.

Figure 3. VS2 RA Ethernet diagram

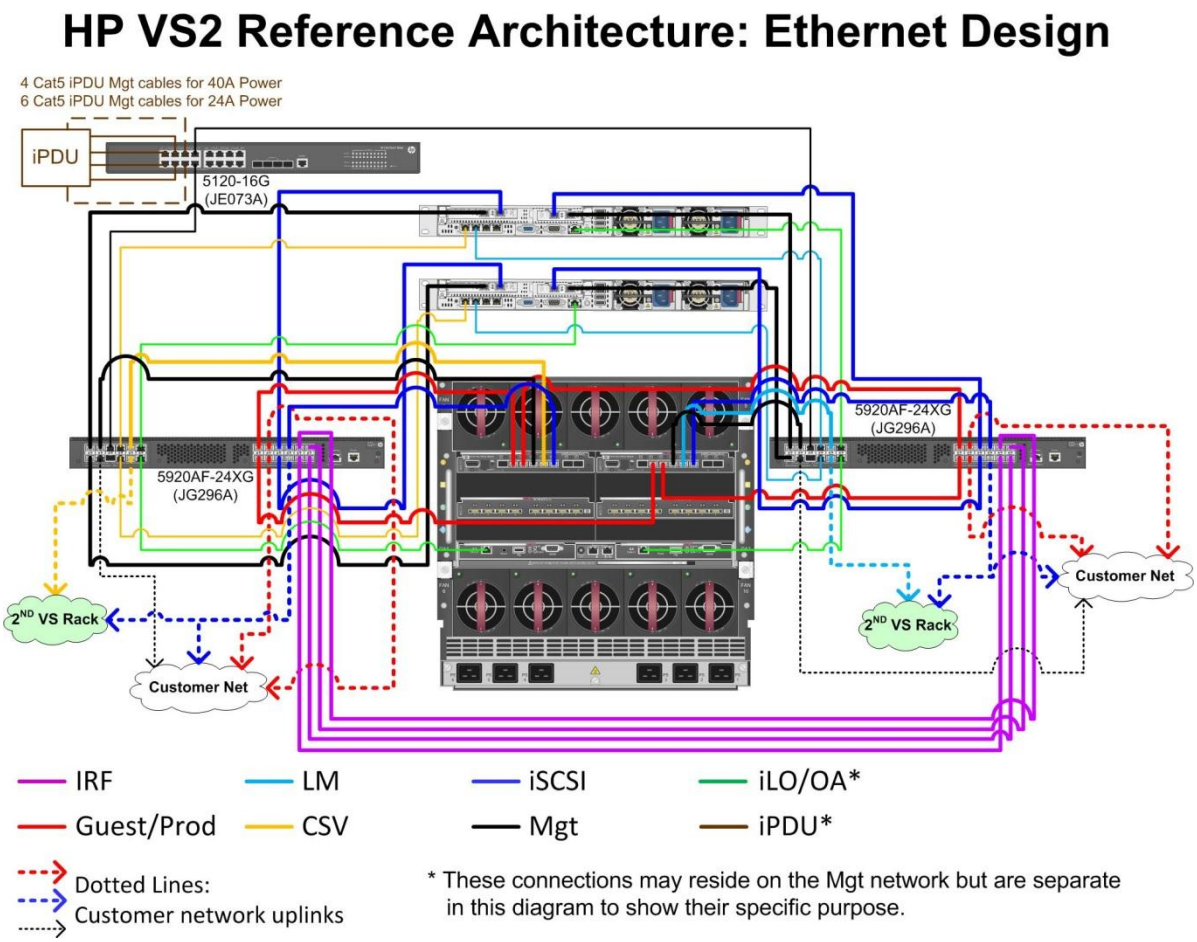
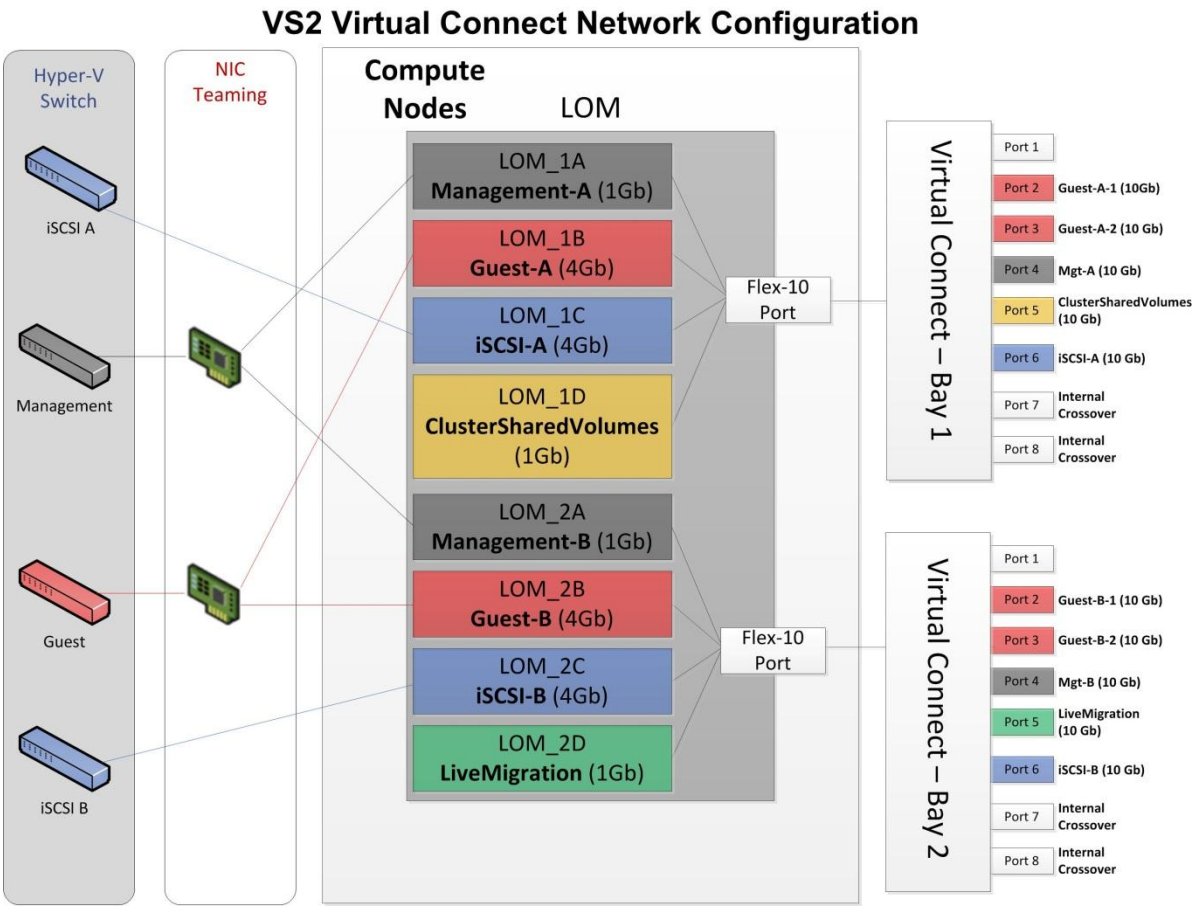
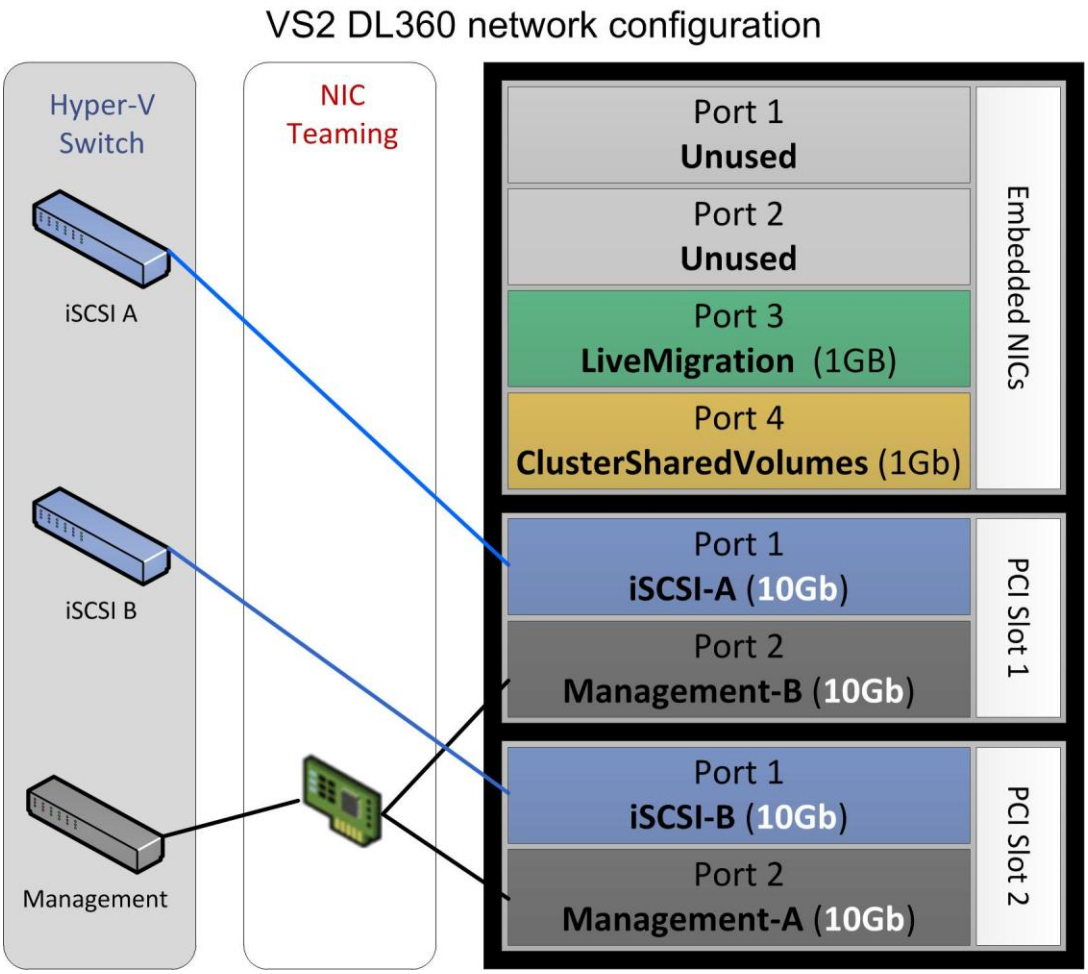


Figure 4. VS2 RA blade network configuration



The DL360 management servers have no need to access the production network, and therefore lack connections for that network, but otherwise use the same networks. Figure 5 shows the DL360 network configuration.

Figure 5. VS2 RA DL360 network configuration



Connectivity within the VS2 RA is achieved via the following switches:

HP 5920AF-24XG Switches (5920) – Redundant 10 Gb switches for the majority of the Ethernet traffic leaving the c7000 enclosure.

The 5920 switches provide added performance and redundancy with the Intelligent Resilient Framework (IRF), which allows the switches to appear logically like a single device. For example, in the VS2 RA, the Flex-10 module in interconnect bay 1 of the c7000 enclosure connects to each 5920 switch. However, with IRF configured between the 5920 switches, they appear to be a single switch to the Flex-10 module, allowing the 2 connections to be joined in an LACP (802.3ad) NIC team.

HP 5120-16G Switch (5120) – low-cost 1Gb switch for intelligent PDU management.

The 5120 switch provides the extra necessary 1Gb ports to manage the HP Intelligent Power Distribution Unit (iPDU) discussed in more detail in the section below. These connections are non-critical and do not require full redundancy, hence the cost-effective switch used for those connections.

Power

The VS2 RA utilizes the HP Intelligent Power Distribution Units (iPDU) which bring an increased level of precision, control, and automation to power distribution. This design allows intelligent power monitoring and management of all servers, including both blades and DL servers. With the HP iPDUs, operations staff have the ability to get exact power consumption at a core, stick or outlet level and get precise reporting of the power utilization and consumption requirements of the HP VS2 RA environment. Additionally the HP iPDUs can be configured to send out SNMP based alerts to monitor the power consumption for the VS2 RA from the SCOM console or from an HP Systems Insight Manager environment for the data center.

For fault tolerance, each VS2 RA rack allows four HP iPDUs for 32A and 40A, or six HP iPDUs for 24A. These are to be connected to the customer's utility supply via multiple feeds to sustain the loss of a single power feed in the data center. Components in the rack are connected in a redundant manner, such that devices with multiple power supplies may survive a power failure in one supply. Due to the many power options used per region, further details on the power configuration are left to the customer to design.

Management software components

The VS2 RA provides the underlying infrastructure for running multiple workloads requiring high levels of performance and availability. To complete the solution, a software stack that can deliver the management, monitoring, operational control, and lifecycle requirements is needed for a virtualized environment. To deliver these capabilities System Center 2012 with HP Insight Control for System Center is recommended for the VS2 RA.

Microsoft System Center 2012

System Center 2012 is a comprehensive management platform consisting of capabilities for infrastructure management, service delivery and automation, and application management and control. The System Center 2012 product includes a number of core components that can be installed including the following:

- App Controller (AC)
- Operations Manager (SCOM)
- Orchestrator (ORC)
- Service Manager (SM)
- Virtual Machine Manager (VMM)
- Data Protection Manager (DPM)
- Endpoint Protection (EP)
- Configuration Manager (SCCM)

The System Center 2012 management stack can be installed on a core set of VMs running on the VS2 RA management servers (DL360s), through which the services are delivered to the business users.

For more information on Microsoft System Center 2012 please see microsoft.com/systemcenter.

HP Insight Control for Microsoft System Center

HP Insight Control for Microsoft System Center provides essential infrastructure management, making it easy to deploy, migrate, monitor, control, and optimize an IT infrastructure from a single, simple management console. Within the VS2 RA, HP Insight Control for Microsoft System Center seamlessly integrates unique HP ProLiant and HP BladeSystem manageability features into System Center Operations Manager, exposing the native capabilities of servers and enclosures to System Center. This increased visibility makes it easier to perform root cause analysis to subsystem and component level.

For the VS2 RA, Gen8 servers and the integration of Insight Control with System Center provide the following capabilities:

- Proactive hardware monitoring, managing and alerting of hardware health and intelligently responding to hardware events on servers running Windows.
- The HP ProLiant Agentless Management Pack manages the health of ProLiant Gen8 servers without the need for loading OS-based SNMP agents or WBEM providers.

- The HP SCVMM 2012 Integration Kit provides HP WinPE and production drivers for ProLiant servers to assist with OS deployment via SCVMM 2012.
- The HP ProLiant Updates Catalog can be used by SCVMM 2012 to provide simplified Windows driver and firmware updates via a rotating, automated workflow for Hyper-V clusters.

In addition, the HP Storage Management Pack for System Center provides seamless integration with SCOM 2012 by integrating predefined discovery and state monitoring policies, event processing rules and tasks. This comprehensive integration solution complements HP Insight Control for Microsoft System Center, and allows administrators to proactively streamline IT operations and increase systems availability by monitoring ProLiant server environments and HP Storage products through a common console.

HP LeftHand Storage Centralized Management Console

The HP LeftHand Storage Centralized Management Console (CMC) provides the ability to manage multiple LeftHand arrays as well as setup advanced features such as multi-site clusters.

The CMC software can be installed on a VM on the management servers. Additionally, if fewer than 3 storage nodes exist – as in the case with the VS2 RA Base configuration – the Failover Manager (FOM) software is installed on a VM.

Built-in high availability

The VS2 RA has been designed to provide extremely high availability at all levels – from the underlying network and storage fabrics to the VM layer. Thus, the VS2 RA can tolerate the failure of any active component, ensuring there is minimal impact to VMs in the event of planned or unplanned downtime. High availability can be extended to the solution level through the live migration of active workloads between VMs, with zero downtime.

In the hardware design, there is full redundancy for 5920 switches and PDUs, while server components such as power supplies, fans, and storage are hot swappable. The VS2 RA also uses NIC teaming and MPIO to maximize throughput and availability. High availability between the VS2 RA and your upstream network infrastructure is achieved through Link Aggregation Control Protocol (LACP) bindings in conjunction with redundant Virtual Connect Flex-10 modules.

In addition, host and guest clustering are leveraged to enhance availability, protecting VMs or specific workloads from planned maintenance and unexpected hardware failures, while enabling load balancing.

Host clustering for VM mobility

To enhance the availability of VMs, virtualization hosts in a VS2 RA should be configured through Windows Failover Clustering as a four-node (Base version) or ten-node (Extended version) cluster.

Installed in the parent partition of each virtualization host, the cluster service allows VMs to move from server to server (though not from one cluster to another in a multi-rack VS2 RA configuration). For example, if downtime is planned, a VM can be moved seamlessly from one node to another server via the Live Migration feature of SCVMM; alternatively, if the downtime is unplanned, VMs undergo a fast restart on a surviving server.

Use cases for host clustering include:

- **Zero-downtime host patching**: Supporting hardware changes or software updates to the parent partition
- **Load balancing**: Migrating a VM to a different server when the original server's resources are being saturated

Guest clustering for workload mobility

Optionally, you can create guest clusters to provide higher levels of availability for specific workloads within a VS2 RA configuration. In this scenario, the cluster service runs within VMs in a two-node cluster and manages the movement of a mission-critical workload from one VM to the other, as required.

Use cases for guest clustering include:

- **Zero-downtime VM patching**: Supporting updates to the OS or application running in the VM
- **Automatic recovery**: Failing over a database instance from one VM to another in the event of a failure

Guest clustering is a powerful option that is most useful for protecting mission-critical workloads. To support VM level guest clustering, an iSCSI storage fabric is part of the core VS2 RA design.

Performance, sizing, and scalability

Sizing any environment requires knowledge of the applied workloads and the hardware resources. However, this is especially difficult in virtualized environments, as multiple different workloads are often applied, and resources support more applications than in the past.

The VS2 RA design allows storage and compute resources to be increased independently and gradually to satisfy both scale-up (within a rack) and scale-out (multiple rack) requirements. As more resources are required, extra BL460c blades can be added to the base configuration, as can 4800 (or other iSCSI) storage, until becoming the extended configuration. At this point, the solution can be scaled-out by adding extra VS2 racks. Additionally, the memory in each blade can be increased from 128GB to 256GB.

Table 1 lists the total resources in VS2 RA configuration. The storage performance in the following tables is based on using RAID 5 on each half of the MDS600 storage shelf and Network RAID 10 between the two halves. Because this configuration mirrors host IOPS between the two halves of the storage, and also uses RAID5, performance is sacrificed somewhat for redundancy. Customers seeking better performance may choose a difference storage configuration on the 4800. Performance will also vary based on the workload applied.

Although only the base and extended configuration sizing information is listed below, remember that components may be added incrementally as needed. Therefore, between 4 and 10 blades may be used, as well as two 4800s (instead of only 1 or 3), or additional external iSCSI storage may be configured.

Note
The storage IOPS and throughput (MB/s) values are based on lab testing with RAID 5 on each half of the storage and Network RAID 10 between the storage halves. Performance will vary depending on the specific workload and application configuration. Also note that the logical CPU core count in Table 1 is twice the number of physical cores, as it assumes Hyper-Threading is enabled.

Table 1. VS2 RA Configuration Sizing

Config	Server Blades	Total Logical CPU Cores	Total RAM (GB)	Ethernet Uplink Bandwidth	IOPS: 8KB-block Random 60/40 R/W	MB/s: 512KB-block Sequential Reads	MB/s: 64KB-block 50% Random, 50% Reads
VS2 Base	4	128	512 / 1024	40 Gbps	~5,000	~800	~160
VS2 Extended	10	320	1280 / 2560	40 Gbps	~15,000	~2,400	~480

Although Table 1 shows resources in the RA, to properly size a customer solution, one must consider memory and CPU overhead of the host operating system and hypervisor. Also, enabling Hyper-Threading makes the number of logical CPU cores appear to double in Windows systems, however, performance testing suggests the actual improvement with Hyper-Threading is only 30%-50%.[1][2]

Therefore, in an attempt to account for overhead and the tested Hyper-Threading performance improvements, and to properly represent available server resources, the memory and CPU values will be adjusted as follows:

1. For each physical server, memory will be reduced by 8 GB RAM for the host OS and hypervisor. Note that this is a rough estimate for memory used for the host and does not necessarily follow published equations for calculating memory required by the host.[3]
2. To determine the number of total logical CPU cores, rather than doubling the number of physical CPU cores, the physical cores will be multiplied by 1.5. To account for the host OS and hypervisor, the number of logical cores is further reduced by one per host.

[1] http://software.intel.com/en-us/articles/intel-hyper-threading-technology-analysis-of-the-ht-effects-on-a-server-transactional-workload/
[2] http://software.intel.com/en-us/articles/performance-insights-to-intel-hyper-threading-technology/
[3] http://msdn.microsoft.com/en-us/library/cc768529(v=bts.10).aspx

Table 2 can be used to help in sizing customer environments.

Table 2. VS2 RA resources adjusted for host OS and hypervisor overhead

Config	Server Blades	Total Logical CPU Cores	Total RAM (GB)	Ethernet Uplink Bandwidth	IOPS: 8KB-block Random 60/40 R/W	MB/s: 512KB-block Sequential Reads	MB/s: 64KB-block 50% Random, 50% Reads
VS2 Base	4	92	480 / 992	40 Gbps	~5,000	~800	~160
VS2 Extended	10	230	1200 / 2480	40 Gbps	~15,000	~2,400	~480

When sizing an environment, it is critical to look at each of the key components and potential bottlenecks, and then size for the limiting component. For example, if I/O is the bottleneck, size appropriately for that, not for the surplus in compute or networking power. To assist in sizing workloads on a VS2 RA configuration, several sample sizing exercises are provided below.

Note:
These scenarios are *only* examples. Each customer environment will vary and should therefore be sized and tested appropriately. HP does not guarantee the VM quantities represented on the VS2 RA below, as they depend heavily on the actual workload applied, nor is HP responsible for improperly sized customer environments.

Scenario 1: Light server load

Assume a rack matching the VS2 RA has multiple VMs, each needing 1 virtual CPU at 50% utilization. Also assume each VM uses Hyper-V's Dynamic Memory, and actively uses 3GB RAM. Finally, assume 30 small-block random IOPS per VM are needed as well as 20Mbps bandwidth.

To estimate the number of VMs of this type that could reside in this solution, calculate anticipated usage values for each component, and see which component bottlenecks first. Table 3 shows expected resource consumption based on the number of VMs described above.

Table 3. Scenario 1 expected resource consumption

VMs	Total Logical Cores Used	Total RAM (GB)	Ethernet Uplink Bandwidth (Mbps)	IOPS: Small-block Random 60/40 R/W
1	1 (at 50%)	3	20	30
160	80	480	3200	4800
170	85	**510** (Upgrade RAM)	3400	**5100** (Add Storage)
180	90	540	3600	5400
190	**95** (Add blades)	570	3800	5700

At around 170 VMs, the blades require more memory (given the assumptions we've made), in which case the memory upgrade option (256 GB per blade) could be applied. Additional storage might also be needed now. By the time 190 VMs are reached, additional blades may also be necessary.

Scenario 2: OLTP-like workload

Assume a particular VS2 rack is to serve primarily as an OLTP database consolidation target such that each consolidated system will be converted to a VM. The average workload applied generates about 80 IOPS per VM, consisting of small block, random I/O, with roughly 60% reads. Each server to be consolidated actively uses 10 GB RAM. Also assume each uses the equivalent of 4 Intel Xeon E5-2670 processors at about 65% utilization. Lastly, assume the network utilization is roughly 100Mbps each. Table 4 shows anticipated resource consumption for scenario 2.

Table 4. Scenario 2 resource consumption

VMs	Total Logical Cores Used	Total RAM (GB)	Ethernet Uplink Bandwidth (Mbps)	IOPS: Small-block Random 60/40 R/W
1	4 (at 65%)	10	100	80
30	78	300	3000	2400
40	**104** (Add blades)	400	4000	3200

Notice that with 40 VMs each getting 4 virtual CPUs (vCPU), that totals 160 vCPUs, but at 65% utilization, that's roughly equivalent to 104 vCPUs, which is above the available resources with 4 blades (96 vCPUs available, as listed in Table 2). Therefore, the demand on the CPU cores is expected to exceed the available resources.

In this case, additional blades could be added to provide the necessary logical CPUs.

Scenario 3: Memory-consuming workload

In this scenario, assume a configuration and workload similar to that in Scenario 2. However, in this scenario, each VM requires 16GB RAM and only uses its 4 vCPUs at about 50%. In the default VS2 Base RA configuration, the memory becomes the expected bottleneck at 40 VMs as seen in orange in Table 5.

To accommodate for this, the 256 GB per-blade option can be used to temporarily relieve that bottleneck. At 50 VMs, however, CPU cores become exhausted, and server blades need to be added. If 2 blades are added (for a total of 6), the next expected bottleneck isn't until 70 VMs are created, at which point the storage becomes the expected bottleneck requiring a second 4800. The progression of this analysis is seen in Table 5.

Table 5. Scenario 3 resource consumption

VMs	Total Logical Cores Used	Total RAM (GB)	Ethernet Uplink Bandwidth (Mbps)	IOPS: Small-block Random 60/40 R/W
1	4 (at 50%)	16	100	80
30	60	480	3000	2400
40	80	**640** (Upgrade RAM)	4000	3200
50	**100** (Add blades)	800	5000	4000
60	120	960	6000	4800
70	140	1120	7000	**5600** (Add storage)

Scenario 4: BI-like workload

In scenario 4, assume the workload is similar to some Business Intelligent (BI) workloads, where large (512KB) sequential reads are pulled from the storage, followed by computation on those blocks. Based on this description, assume that each VM pulls 20 MB/s from the storage and consume roughly 500 Mbps network bandwidth. The workload also uses 50% of 4 vCPUs (doing relatively heavy computation, but only 50% of the time) and 12 GB RAM.

Given these characteristics, Table 6 below shows expected resource demands.

Table 6. Scenario 4 resource consumption

VMs	Total Logical Cores Used	Total RAM (GB)	Ethernet Uplink Bandwidth (Mbps)	MB/s: 512KB Blocks, sequential reads
1	4 (at 50%)	12	500	20
30	60	360	15000	600
40	80	480	20000	800
50	**100** (Add blades)	**600**	25000	**1000** (Add storage)
60	120	720	30000	1200
70	140	**840** (Upgrade RAM)	35000	1400

Beginning with a VS2 Base RA, at around 40 VMs, the storage is expected to reach its maximum bandwidth, suggesting the need to add more storage to exceed 40 VMs. Next, the memory and logical cores required exceed that available, so adding a blade or two will help. Assuming a total of 6 blades, the next anticipated bottleneck appears at 70 VMs, with a lack of memory, which could be resolved by upgrading to 256 GB RAM per blade.

Scenario 5: SharePoint collaboration and document management solution

The initial deployment of SharePoint 2010 is primarily to provide collaboration and document management capabilities to a company and its various groups, teams and users. The design of the solution will vary depending on the workload (function usage and frequency), however a very common approach is to deploy multiple Web Front End (WFE) servers (which facilitate providing content pages to the user, as well as run the majority of available application services), a SQL server to provide required databases and content storage; and optionally one or more dedicated application servers to support frequently used services such as Search. SharePoint is designed as a highly scale-out application, and thus promotes a strategy of deploying different services on separate servers, or VMs.

The example shown in Table 7 depicts 3 x WFE servers, a SQL server and 2 x dedicated Application servers (Search, etc.) and is a very common deployment design. With the VMs configured as shown with respect to the number of cores allocated, memory, etc. this configuration can support between 1000 and 2000 users, depending on workload mix and degree of content modification, etc. The table shows the resource consumption (cores, memory, network bandwidth, and storage I/O and throughput for a typical collaboration workload, including a high degree of document management and content modification. The line “Block x 1” summarizes these resources for the whole solution (all services and VMs), and reflects a typical deployment. Should further capacity be needed, it is typical to scale-out the solution by increasing the number of VMs running the required services. As a single active SQL server can support about three WFEs, it is typical to also scale-out SQL servers to provide the required capacity, and optionally host specific content library databases to equalize the load as may be needed. The subsequent table lines “Block x 2” and “Block x 4” depict this scale-out. Note that “Block x 4” approaches the throughput capacity limit of a single 4800, and thus in this case it would be recommended to add storage to the VS2 solution.

Table 7 shows example resource consumption for the simulated SharePoint workload. For more information on deploying SharePoint 2010 in a virtualized environment, please see the HP white paper titled Virtualizing SharePoint 2010 on HP Converged Infrastructure. This paper also provides details of the test workload, detailed performance results and strategies for scaling-out the solution by running specific application services on multiple VMs.

Table 7. Scenario 5 resource consumption

VMs (1 Block)	Total Logical Cores Used	Total RAM (GB)	Ethernet Uplink Bandwidth (Mbps)	MB/s: 64KB Blocks, 50% Random, 50% reads
3 (WFE)	12 (4 each at 85%)	30 (10GB each)	60 (20Mbps each)	None
1 (SQL)	4 (at 30%)	16	70	40
2 (App)	8 (4 each at 10%)	20 (10GB each)	20 (10Mbps each)	Insignificant
Block x1 (6 VMs)	12.2	66	150	40
Block x2 (12 VMs)	24.4	132	300	80
Block x4 (24 VMs)	48.8	264	600	160 (Add storage)

Scenario 6: Mixed workload

In a VS2 RA, it's likely that multiple different workloads will be combined, unlike in the examples above. However, while more complicated, the concepts of sizing remain the same. One must determine how many VMs with each type of workload will exist, then combine the anticipated resource utilizations from each of those workloads and determine where the bottlenecks are and how to overcome them.

Remember that the flexibility built into the design of VS2 allows customers to increase needed resources in building blocks requiring minimal effort. Memory can be upgraded to 256 GB RAM in each blade by adding extra DIMMs, blades can be added to increase memory and logical CPU cores, and storage can be added to increase IOPS and MB/s. If the solution requires more network bandwidth extra uplink cables can added, or existing VC network configurations can be modified to allow more traffic on certain networks. Finally, if a rack's resources are exhausted, a second VS2 RA rack can be added and integrated with the existing rack to allow multi-rack clusters for improved performance and availability.

Summary

The VS2 RA discussed in this paper simplifies the design process of a scalable, highly-virtualized environment. Properly sizing virtualized environments requires some knowledge of the consolidated workloads applied, and this paper provides some sizing guidance in a building block structure to allow for expansion and optional customization.

Appendix – Bill of materials

This bill of materials is a summary of the key components needed for building a VS2 RA. This does not include cables and other minor components. Due to varying power options, components of the power configuration are also not included in this bill of materials.

VS2 Base Reference Architecture configuration

Table 8. VS2 Base RA configuration components

Quantity	HP Part Number	Description
Rack and Network Infrastructure		
1	BW904A	HP 642 1075mm Shock Intelligent Rack
1	433718-B21	HP BLc7000 10K Rack Ship Brkt Opt Kit
1	BW906A	HP 42U 1075mm Side Panel Kit
1	BW930A	HP Air Flow Optimization Kit
1	BW891A	HP Rack Grounding Kit
2	AF070A	HP 10pk Carbt 1U Universal Filler Panel
1	BW932A	HP 600mm Jb Rack Stabilizer Kit
Virtualization Hosts		
1	507019-B21	HP BLc7000 CTO 3 IN LCD ROHS Encl
4	641016-B21	HP BL460c Gen8 10Gb FLB CTO Blade
4	662064-L21	HP BL460c Gen8 E5-2670 FIO Kit
4	662064-B21	HP BL460c Gen8 E5-2670 Kit
32	672631-B21	HP 16GB 2Rx4 PC3-12800R-11 Kit
8	652564-B21	HP 300GB 6G SAS 10K 2.5in SC ENT HDD
4	684212-B21	HP FlexFabric 10Gb 2P 554FLB FIO Adptr
4	589251-B21	MS WS08 R2 DataCnt2CPU VS FIO Npi en SW
2	AJ864A	HP 3Gb SAS BL-c Switch
1	591973-B21	HP VC Flex-10 Ent Edn for BLc7000 Opt
1	517521-B22	HP 6X 2400W Plat Ht Plg FIO Pwr Sply Kit
1	456204-B21	HP BLc7000 DDR2 Encl Mgmt Option
1	677595-B21	HP BLc 1PH Intelligent Power Mod FIO Opt
6	582765-B21	HP IC ML/DL/BL 1-Svr FIO 24x7 SW
1	517520-B21	HP BLc 6X Active Cool 200 FIO Fan Opt

Quantity	HP Part Number	Description
Management Servers		
2	654081-B21	HP DL360p Gen8 8-SFF CTO Server
2	664011-L21	HP DL360p Gen8 E5-2690 FIO Kit
2	664011-B21	HP DL360p Gen8 E5-2690 Kit
16	672631-B21	HP 16GB 2Rx4 PC3-12800R-11 Kit
4	652564-B21	HP 300GB 6G SAS 10K 2.5in SC ENT HDD
2	652238-B21	HP 9.5mm SATA DVD ROM Jb Kit
2	684208-B21	HP Ethernet 1GbE 4P 331FLR FIO Adptr
2	631679-B21	HP 1GB FBWC for P-Series Smart Array
2	663201-B21	HP 1U SFF BB Gen8 Rail Kit
4	614203-B21	HP NC552SFP 10GbE 2P Svr Adapter
4	656363-B21	HP 750W CS Plat PL Ht Plg Pwr Supply Kit
2	582765-B21	HP IC ML/DL/BL 1-Svr FIO 24x7 SW
2	589251-B21	MS WS08 R2 DataCnt2CPU VS FIO Npi en SW
Networking		
3	592774-B21	HP SL Universal Switch Rail Kit
1	JE073A	HP 5120-16G SI Switch
2	JG296A	HP 5920AF-24XG Switch
4	JC680A	HP A58x0AF 650W AC Power Supply
4	JG297A	HP 5920AF-24XG Bk(pwr)-Frt(prt) Fn Tray
Storage		
1	BV932A	HP 4800 G2 42TB SAS SAN BladeSystem

VS2 Extended Reference Architecture configuration

Table 9. VS2 Extended RA configuration components

Quantity	HP Part Number	Description
Rack and Network Infrastructure		
1	BW904A	HP 642 1075mm Shock Intelligent Rack
1	433718-B21	HP BLc7000 10K Rack Ship Brkt Opt Kit
1	BW906A	HP 42U 1075mm Side Panel Kit
1	BW930A	HP Air Flow Optimization Kit
1	BW891A	HP Rack Grounding Kit
2	AF070A	HP 10pk Carbt 1U Universal Filler Panel
Virtualization Hosts		
1	507019-B21	HP BLc7000 CTO 3 IN LCD ROHS Encl
10	641016-B21	HP BL460c Gen8 10Gb FLB CTO Blade
10	662064-L21	HP BL460c Gen8 E5-2670 FIO Kit
10	662064-B21	HP BL460c Gen8 E5-2670 Kit
80	672631-B21	HP 16GB 2Rx4 PC3-12800R-11 Kit
20	652564-B21	HP 300GB 6G SAS 10K 2.5in SC ENT HDD
10	684212-B21	HP FlexFabric 10Gb 2P 554FLB FIO Adptr
10	589251-B21	MS WS08 R2 DataCnt2CPU VS FIO Npi en SW
2	AJ864A	HP 3Gb SAS BL-c Switch
1	591973-B21	HP VC Flex-10 Ent Edn for BLc7000 Opt
1	517521-B22	HP 6X 2400W Plat Ht Plg FIO Pwr Sply Kit
1	456204-B21	HP BLc7000 DDR2 Encl Mgmt Option
1	677595-B21	HP BLc 1PH Intelligent Power Mod FIO Opt
16	582765-B21	HP IC ML/DL/BL 1-Svr FIO 24x7 SW
1	517520-B21	HP BLc 6X Active Cool 200 FIO Fan Opt

Quantity	HP Part Number	Description
Management Servers		
2	654081-B21	HP DL360p Gen8 8-SFF CTO Server
2	664011-L21	HP DL360p Gen8 E5-2690 FIO Kit
2	664011-B21	HP DL360p Gen8 E5-2690 Kit
16	672631-B21	HP 16GB 2Rx4 PC3-12800R-11 Kit
4	652564-B21	HP 300GB 6G SAS 10K 2.5in SC ENT HDD
2	652238-B21	HP 9.5mm SATA DVD ROM Jb Kit
2	684208-B21	HP Ethernet 1GbE 4P 331FLR FIO Adptr
2	631679-B21	HP 1GB FBWC for P-Series Smart Array
2	663201-B21	HP 1U SFF BB Gen8 Rail Kit
4	614203-B21	HP NC552SFP 10GbE 2P Svr Adapter
4	656363-B21	HP 750W CS Plat PL Ht Plg Pwr Supply Kit
2	582765-B21	HP IC ML/DL/BL 1-Svr FIO 24x7 SW
2	589251-B21	MS WS08 R2 DataCnt2CPU VS FIO Npi en SW
Networking		
3	592774-B21	HP SL Universal Switch Rail Kit
1	JE073A	HP 5120-16G SI Switch
2	JG296A	HP 5920AF-24XG Switch
4	JC680A	HP A58x0AF 650W AC Power Supply
4	JG297A	HP 5920AF-24XG Bk(pwr)-Frt(prt) Fn Tray
Storage		
3	BV932A	HP 4800 G2 42TB SAS SAN BladeSystem

For more information

HP ProLiant Servers	hp.com/go/proliant
HP BladeSystem	hp.com/go/bladesystem
HP Converged Infrastructure	hp.com/go/convergedinfrastructure
HP Insight Control	hp.com/go/insightcontrol
HP Networking	hp.com/go/networking
HP LeftHand Storage	hp.com/go/lefthand
HP VirtualSystem	hp.com/go/virtualsystem
Microsoft System Center	microsoft.com/en-us/server-cloud/system-center/default.aspx
HP CloudSystem	hp.com/go/cloudsystem

To help us improve our documents, please provide feedback at hp.com/solutions/feedback.

Get connected

hp.com/go/getconnected

Current HP driver, support, and security alerts
delivered directly to your desktop

4AA4-3583ENW, Created September 2012