# HP VirtualSystem VS2 for Red Hat

Reference architecture for building a Red Hat Enterprise Virtualization environment on HP Converged Infrastructure with integrated management

## Table of contents

**Executive summary** 3

**Introduction** 3
  - **Benefits** 4

**Solution overview** 4

**Solution components** 5
  - **HP BladeSystem** 5
  - **HP ProLiant servers** 6
  - **HP Virtual Connect Flex-10** 7
  - **HP 3Gb SAS BL Switch** 8
  - **HP Networking** 8
  - **HP Storage** 8

**Performance, sizing, and scalability** 9

**Environment setup** 13
  - **Servers** 16
  - **Storage** 16
  - **Networking** 16
  - **Power** 19
  - **Management software components** 20
  - **Built-in high availability** 20
  - **Management cluster configuration** 21
  - **Virtual Connect profiles** 22
  - **Host node configuration** 23
  - **Final configuration** 23

**Implementing a proof-of-concept** 24

**Appendix A: Bill of Materials (BOM)** 25
  - **VS2 Base Reference Architecture configuration** 25
  - **VS2 Extended Reference Architecture configuration** 27

**Appendix B: Software/Firmware versions** 30
Appendix C: Installing the SIM service 31
For more information 39
Executive summary

The HP VirtualSystem VS2 for Red Hat Reference Architecture (VS2 RA) is intended to reduce data center complexity, increase agility, and mitigate the risks that come with deployment of a virtualized environment. The VS2 RA design leverages best-in-class HP Converged Infrastructure and Red Hat Enterprise Virtualization, the open source choice for virtualizing workloads. At its core, the VS2 RA provides a foundation for building a high-performance Red Hat Enterprise Virtualization platform that has been optimized to consolidate and provision hundreds to thousands of workloads while providing 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 HP Converged Infrastructure, including the HP BladeSystem architecture, HP Virtual Connect Flex-10, and HP LeftHand P4800 SAN Solution for BladeSystem. The HP BladeSystem allows the VS2 RA to be sized and scaled in a modular fashion, simplifying scaling up and out as additional resources are required. Additionally the HP BladeSystem architecture helps to not only reduce the footprint of the solution but also reduce the 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 the VS2 RA requirements – a very important capability when building a virtualization platform due to the complex networking requirements. Coupling these technologies with the HP LeftHand P4800 SAN Solution for BladeSystem (P4800), the VS2 RA provides an extremely dense platform for the deployment of virtualized environments that require high levels of storage performance.

At the core of the VS2 RA are the HP ProLiant BL460c Gen8 blade servers. This latest generation of blade servers provides a number of advancements in intelligence, control, and performance, helping drive down the operational expenses of the VS2 RA platform. Innovations such as HP Integrated Lights-Out 4 (iLO 4), HP Active Health System, HP Smart Update Manager, HP SmartMemory, and HP SmartDrive technologies help customers save substantial administrative time and effort in managing and supporting the VS2 RA platform; enabling customers to reinvest that time savings in other higher value activities.

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, outlines some typical use cases, and describes a reference example for building out a reference solution.

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

Introduction

Most of today’s IT spending is consumed by operating and maintaining the existing infrastructure, leaving little space in the budget for new investments that can add value. However, while it may seem obvious that these ongoing costs need to be reduced, the inefficiency and inflexibility of your infrastructure may provide few opportunities.

Key areas that may need to be addressed in your infrastructure include the following:

- **Server utilization**: Servers are often dedicated to individual applications, leading to massive underutilization of resources. Such servers can typically be virtualized and consolidated.

- **Manageability**: Giving each new application its own system configuration creates a significant management burden. When patches and upgrades have to be applied individually, it becomes difficult to automate such management tasks; thus, in order to simplify the environment, it is important to minimize the number of hardware platforms, operating systems, and system configurations.

- **Inflexibility**: With a proprietary or monolithic infrastructure, it can be difficult to scale applications up or down, or provide any level of fault-tolerance. The inherent inflexibility of such an infrastructure also makes it almost impossible to implement the updates required to accommodate changing business needs.

However, with virtualization, a single HP ProLiant server can run multiple application instances in isolation as virtual machines (VMs). Automated tools can allocate resources (CPU, storage, and networking) to individual VMs, allowing them to scale up and down in tune with the demands of the particular workload. Moreover, since each operating system and application environment is stored on a virtual disk, you can easily copy this disk and create one or many clones of the virtual machine.
VMs are also highly portable and can easily be migrated to a new physical server (host) to support maintenance activities on the original host or to better balance the workload between hosts. This portability also allows you to quickly restart an application on a new host if the original host were to fail.

Additional benefits of HP VirtualSystem are summarized below.

**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 are 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 and high availability**

By standardizing platforms and system configurations, you can enhance IT efficiency and enable automation. The portability of VMs enhances 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 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 to a 10-blade design with up to three P4800s in a single rack. 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**

The VS2 RA solution is built on HP Converged Infrastructure components and provides the infrastructure layer for deploying a private cloud. By deploying Red Hat Enterprise Virtualization Manager to manage the VS2 RA, customers can manage the entire lifecycle of both the virtual machines and the workloads running on them.

With HP Systems Insight Manager (SIM) customers get deeper insight and monitoring control into the hardware. HP SIM provides hardware alerting and event information on a single pane of glass.

**Solution 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 Appendix A: Bill of Materials (BOM) for the HP VS2 RA based on the HP ProLiant BL460c Gen8 server blade (BL460c) and the HP LeftHand P4800 G2 42TB SAS SAN Solution for BladeSystem (P4800) storage. This reference architecture accommodates both I/O-intensive and CPU-intensive workloads and can be easily modified to increase storage or compute capabilities, ranging between the base and extended configurations detailed below and shown in Figure 7 and Figure 8.

The solution discussed in this paper is intended to operate as a fully self-contained virtual hosting solution. All management and storage is contained within a single rack with multiple virtual networks connected out to the local data center via trunked high-speed links. Use of additional external storage or expansion to a secondary rack is included in the design but is not covered in this document.

**Solution components**

Management infrastructure is provided by two HP ProLiant DL360p Gen8 servers operating as a cluster.

4 to 10 HP BL460c Gen8 “compute” nodes provide virtual hosting and CPU power for virtual machines.

2 or 3 P4800 G2 (each consisting of 2 P4460sb G2 servers and a 42TB HP 600 Modular Disk System (MDS600)) provide high-speed iSCSI shared storage.

**HP BladeSystem**

Drive business innovation and eliminate sprawl with HP BladeSystem engineered to maximize every hour, watt, and dollar. With HP BladeSystem, it is possible to create a change ready, power efficient, network optimized, simple to manage, and high performance infrastructure on which to build and scale your virtualized environment. 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 midplane for wire-once connectivity of server blades to network and shared storage. An HP BladeSystem c7000 enclosure populated with HP ProLiant BL460c Gen8 blades is shown in Figure 1.

_Figure 1: HP BladeSystem c7000 Enclosure (pictured with HP ProLiant BL460c Gen8 server blades)_;
**HP ProLiant servers**

The latest generation of HP blade server technology, the HP ProLiant Generation 8 servers, features embedded automation and intelligence that simplify lifecycle operations, 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, iLO 4, HP Active Health System, HP Agentless management, and HP Intelligent Provisioning features.

With HP SmartMemory technology, servers can identify and verify that system memory has passed enhanced levels of qualification and testing. This prevents a common cause of downtime and service interruption due to lower quality memory DIMMs requiring replacement. Additionally, certain HP ProLiant Generation 8 servers with HP SmartMemory can run memory at lower voltages (1.35 vs. 1.5) without impacting performance. This helps lower power consumption up to 20% while still maintaining equal performance.

HP iLO has been a standard component for HP servers for many years, simplifying server setup, health monitoring and remote server management. With the release of the Gen8 server family, iLO 4 has been enhanced to provide greater control and awareness of the system health and management and is the core element enabling features such as the HP Active Health System, Agentless management and HP Intelligent Provisioning.

HP iLO 4 includes embedded storage, replacing the traditional HP SmartStart DVD by including the support and configuration utilities directly on the servers shipping from the factory. This not only includes the common setup utilities but also includes the drivers and firmware required for installing the OS using HP Intelligent Provisioning. A simple setup wizard allows you to download updated drivers and firmware as needed and to link back to HP for additional support through HP Insight Online (described in the following section).

The HP Active Health System is a diagnostic tool which is used to monitor and record all the changes in server hardware and system configuration, changes in temperature and voltage, and alerts. It is used to diagnose problems and resolve system failures. This information can be used by HP support engineers if there is a service issue. And since the system is built on iLO, there is no impact on server performance to gather and collect this information.

Additionally, the iLO architecture enables the HP Agentless management capability which collects core management functions directly from iLO for all internal system components, including health monitoring and alerts.

These HP ProLiant Gen8 technologies coupled with the performance improvements and other enhanced features such as additional thermal sensors to provide more efficient cooling instrumentation yield a platform that can simplify management and lower operational costs in both the deployment and on-going management of the VS2 RA. Figure 2 shows the BL460c Gen8 server blade.

**Figure 2:** HP BL460c Gen8 Server Blade

For more information on the HP ProLiant Gen8 technology please visit [hp.com/go/proliant](http://hp.com/go/proliant).
**Simplified support and lifecycle maintenance**

With the HP ProActive Insight architecture, HP ProLiant Gen8 servers continuously monitor more than a thousand system parameters to optimize application performance and proactively decrease downtime, while providing organizations insight into every aspect of their IT infrastructure. With this continuous monitoring along with HP Insight Online (a comprehensive cloud based management portal for viewing system health, asset, and warranty information), the VS2 RA provides simplified monitoring and support capabilities.

The embedded HP Active Health System collects logs for every action performed on the server. In the event of a service issue, these logs provide valuable time savings when performing root cause analysis. Customers can log into the portal to see the service events on these systems in real-time. Additionally, as part of the HP Insight Online and iLO technology, customers can leverage the embedded remote support feature that provides Phone Home Service Support for the HP servers in the VS2 RA (without any additional software installation).

For server and storage upgrades, HP has established a set of tools to make updating driver and firmware changes to complex solutions a more streamlined and simplified process. With the HP Service Pack for ProLiant (SPP) and HP Smart Update Manager (HP SUM), customers can download tested and validated firmware/driver bundles on a periodic basis from HP and push those updates out to the components in the VS2 RA in a controlled manner. This allows administrators to target updates on a server by server basis and control migration of the workloads to eliminate any workload downtime associated with planned maintenance activities. Additionally, the redundancy built into the VS2 RA design provides multiple data paths for updating infrastructure components, such as the network switches, without needing to take the solution offline.

For more information on the HP SPP and HP SUM tools visit [hp.com/go/spp](http://hp.com/go/spp).

HP Proactive Care Services provides an additional level of support for organizations managing complex IT environments. Geared for converged, virtualized and cloud-based environments, Proactive Care Services features remote and onsite support, proactive scans and reports and regular consultations with HP technology experts. By effectively using the power of industry-leading remote management technology, expert proactive planning, and rapid problem resolution, we can help you realize the full value of your investments.


**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. Each FlexNIC is recognized by the server as a PCI-e physical function device with customizable speeds from 100Mb to 10Gb for each of the iSCSI storage, management and production networks recommended by HP for VS2 deployments.

VC Flex-10 modules and adapters aggregate traffic from multiple networks into a single 10Gb link. Flex-10 technology partitions the 10Gb data stream into multiple (up to four) adjustable bandwidths, preserving routing information for all data classes. These and other features of the VC Flex-10 modules make them an excellent choice for virtualized environments.

**Figure 3:** HP Virtual Connect Flex-10 Ethernet Module
**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 straightforward external zoned SAS or shared SAS storage solution via an HP Smart Array P700m controller installed in the server.

The SAS BL Switch is a key component of the P4800 storage used in the VS2 RA, as it provides the storage blades connectivity to the drive enclosures in a high-performance, highly-available configuration. Figure 4 shows the SAS BL Switch in the VS2 RA.

*Figure 4: HP 3Gb SAS BL Switch*

**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 switches: 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. The HP 5920AF-24XG is a high-density 10 GbE, ultra-deep packet buffering, top-of-rack (ToR) switch. The availability of this switch was further increased by adding a second switch configured with IRF providing active, redundant connections. Figure 5 shows the 24-port 10Gb Ethernet 5920 switch used in the VS2 RA.

*Figure 5: HP 5920AF-24XG Switch*

**HP Storage**

Storage is a critical architectural component of any virtualized environment as it directly impacts performance and availability of VMs in the system, and therefore must be chosen carefully. HP provides storage solutions to meet small to large server virtualization deployments with a range of SAN, NAS, and DAS solutions.
HP LeftHand P4800 SAN for BladeSystem

The HP LeftHand P4800 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 integrated 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.

Simplicity and performance are enhanced with high-speed storage paths, and dense disk spindle counts, all without external storage switching fabric required. Several storage software features are included with P4800 storage, including thin provisioning, snapshots, remote copy, and Network RAID for high availability. Convergence with the HP BladeSystem architecture also provides a significant density and footprint advantage further saving costs.

In the VS2 RA, 2 or 3 HP P4800 units (Figure 6) are configured, each unit consisting of a node pair serving 70 spindles.

Figure 6: HP P4800 SAN for BladeSystem

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 power to be increased independently and gradually to satisfy both scale-up (within a rack) and scale-out (multiple racks) requirements. As more resources are required, extra BL460c blades can be added to the base configuration, as can P4800 storage. The solution can be scaled-out further by adding additional VS2 racks. Additionally, the memory in each blade can be increased.

Table 1 lists the total resources in the 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. This configuration mirrors host IOPS between the two halves of the storage, and utilizes RAID 5 locally on each half for increased protection. Customers seeking better performance may choose a different storage configuration on the P4800. 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.
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. I/O performance information was collected using iozone throughput testing. Performance will vary depending on the specific workload and application configuration. Also note that the logical CPU core count is twice the number of physical cores, as it assumes Hyper-Threading is enabled.

Table 1: VS2 RA Configuration Sizing

<table>
<thead>
<tr>
<th>Config</th>
<th>Server Blades</th>
<th>Total Logical CPU Cores</th>
<th>Total RAM (GB)</th>
<th>Aggregate Ethernet Uplink Bandwidth</th>
<th>IOPS: 8KB-block Random 60/40 R/W</th>
<th>MB/s: 512KB-block Sequential Reads</th>
<th>MB/s: 64KB-block 50% Random, 50% Reads</th>
</tr>
</thead>
<tbody>
<tr>
<td>VS2 Base</td>
<td>4</td>
<td>128</td>
<td>512 / 1024</td>
<td>40 Gbps</td>
<td>~9,000</td>
<td>~1600</td>
<td>~320</td>
</tr>
<tr>
<td>VS2 Extended</td>
<td>10</td>
<td>320</td>
<td>1280 / 2560</td>
<td>40 Gbps</td>
<td>~15,000</td>
<td>~2,400</td>
<td>~480</td>
</tr>
</tbody>
</table>

Although Table 1 shows resources suggested in the RA, to properly size a customer solution, one must consider memory and CPU overhead of host operating systems. Enabling Hyper-Threading makes the number of logical CPU cores appear to double, however, performance testing suggests the actual improvement is generally less than double\(^1\). Therefore, in an attempt to more accurately represent available server resources, to determine the available CPU resources the number of physical cores will be multiplied by 1.5 rather than 2.0. Table 2 can be used to help in sizing environments.

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

<table>
<thead>
<tr>
<th>Config</th>
<th>Server Blades</th>
<th>Total Logical CPU Cores</th>
<th>Total RAM (GB)</th>
<th>Aggregate Ethernet Uplink Bandwidth</th>
<th>IOPS: 8KB-block Random 60/40 R/W</th>
<th>MB/s: 512KB-block Sequential Reads</th>
<th>MB/s: 64KB-block 50% Random, 50% Reads</th>
</tr>
</thead>
<tbody>
<tr>
<td>VS2 Base</td>
<td>4</td>
<td>96</td>
<td>480 / 992</td>
<td>40 Gbps</td>
<td>~9,000</td>
<td>~1600</td>
<td>~320</td>
</tr>
<tr>
<td>VS2 Extended</td>
<td>10</td>
<td>240</td>
<td>1200 / 2480</td>
<td>40 Gbps</td>
<td>~15,000</td>
<td>~2,400</td>
<td>~480</td>
</tr>
</tbody>
</table>

When sizing an environment, it is critical to look at each of the key components and potential bottlenecks, and 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.

\(^1\) http://software.intel.com/en-us/articles/intel-hyper-threading-technology-analysis-of-the-ht-effects-on-a-server-transactional-workload
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 environments.

Scenario 1: 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. Assume the average workload applied generates about 80 IOPS per VM, consisting of small block, random I/O, with roughly 60% reads. Also assume each server to be consolidated actively uses 10 GB RAM. Assume each uses the equivalent of 4 processors at about 65% utilization. Lastly, assume the network utilization is relatively low per VM – roughly 100Mbps each.

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 resource consumption

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

Notice that with 40 VMs each getting 4 virtual CPUs (vCPU), we are using 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 2: Memory-consuming workload

In this scenario, assume a configuration and workload similar to that in Scenario 1. 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 bold in Table 4.

To relieve the memory bottleneck upgrade the memory on each host blade to 256 GB. At 50 VMs, however, CPU cores become exhausted, and server blades need to be added. The progression of this analysis is seen in Table 4.

Table 4: Scenario 2 resource consumption

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

Scenario 3: BI-like workload

In scenario 3, 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 consumes 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 5 below shows expected resource demands.

Table 5: Scenario 3 resource consumption

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

As in the previous scenario, the workload reaches the CPU bottleneck at 50 virtual machines, 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 4: 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 P4800s can be added to increase IOPS and MB/s. If the solution requires more network bandwidth extra uplink cables can be 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.

Environment setup

The VS2 RA was tested with a Red Hat Enterprise Linux cluster providing highly available access to the management environment and highly available virtualization hosting provided by Red Hat Enterprise Virtualization (RHEV). This provides the infrastructure resource pool to provision hundreds of VM instances for consolidation and new VM provisioning.

RHEV running on HP ProLiant servers is a proven enterprise-grade virtualization platform, offering leading performance, scalability, and SELinux hardened security for your virtualized deployment – offering 60–80% cost savings over proprietary alternatives. Built on Linux Kernel-based Virtual Machine (KVM) technology, Red Hat Enterprise Virtualization is a natural choice for Linux workloads and an option for Microsoft® Windows® workloads as well. The entire platform is available through HP – including RHEV subscriptions and support, and implementation and consulting services for larger deployments. For the latest RHEV information and offerings from HP, visit http://h71028.www7.hp.com/enterprise/us/en/partners/redhat-enterprise-virtualization.html

The solution is composed of storage (P4800) and compute (BL460c Gen8) building blocks, and consists of two main configurations (“base” and “extended”) as summarized in the bulleted lists below. As each virtualization environment has varying applications and workloads, storage and compute blocks may be added as appropriate, ranging between the base and extended configurations detailed below and shown in Figure 7 and Figure 8. 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)
- 2 x HP P4800 G2 42TB SAS SAN Solution for BladeSystem (a total of 84 TB raw storage, 140 spindles)
- 2 x HP ProLiant DL360p Gen8 Servers for management
- 2 x HP 5920AF-24XG Switches
- 1 x HP 5120-16G SI Switch

Figure 7: VS2 RA Base Front and Rear Views
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 P4800 G2 42TB SAS SAN Solution for BladeSystem (126 TB raw storage, 210 spindles)
- 2 x HP ProLiant DL360p Gen 8 Servers for management
- 2 x HP 5920AF-24XG Switches
- 1 x HP 5120-16G SI Switch

Figure 8: 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 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) and is running Red Hat Enterprise Virtualization Host 6.2.

Note
With the Red Hat Enterprise Linux unlimited VM entitlement (BC323AAE RHEL Srv, 2 skt, Unltd Guest 24x7 3yr Lic), users have the right to run an unlimited number of desktop or server VMs on each entitled RHEV-H node. As each environment is unique and may not run Red Hat Enterprise Linux exclusively, this entitlement is not included in the parts list for the reference architecture. Licensing of the virtual machines is the responsibility of the customer.

In addition to the BL460c compute nodes, two HP ProLiant DL360p Gen8 (DL360) servers provide management of the VS2 RA rack. Each DL360 has 16 CPU cores, 128GB RAM and two 300GB HDDs. These servers are clustered using Red Hat Enterprise Linux 6.2 with the High Availability Add On to provide highly available management of the solution.

Storage

Storage is also designed in a modular fashion for the VS2 RA solution, with each core storage building block consisting of an HP P4800 G2 42TB SAS SAN Solution for BladeSystem. The P4800 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 are two P4800 blocks and for an extended configuration there are three storage blocks. The storage fabric 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 P4800 allows multiple levels of redundancy and high availability. Each half of the MDS600 drive enclosure was set up in a RAID configuration (RAID 5 by default). Additionally, the two halves are mirrored using Network RAID 10 to further protect against drive, power, and controller failure. Customers can further extend the storage by connecting to existing iSCSI SAN environments for additional storage capacity if needed.


Networking

When designing a virtualization solution, there are a number of distinct networks that are required and should be isolated for best performance. Thus, the following networking design is recommended for each server in the VS2 RA:

- Bonded production network
- Bonded management network
- Bonded iSCSI network

Additionally for the management servers, a dedicated, bonded network should be used to isolate cluster traffic for maximum reliability and performance.

The VS2 RA is designed to follow the above recommendations, providing redundant connectivity to each of the major network segments. All the recommended networks utilize host-level NIC bonding for improved redundancy and throughput.

The combination of the HP P4800 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 bonded network interfaces as recommended above for maximum performance and availability.
The HP Virtual Connect Flex-10 10Gb Ethernet Module for c-Class BladeSystem (Flex-10 Module) reduces the cabling required and provides valuable networking flexibility with profiles designed for the networking needs of each blade. Figure 9 and Figure 10 show the VS2 RA network design and BL460c Gen8 connectivity configuration.

**Figure 9: VS2 RA Ethernet diagram**

The Virtual Connect wiring and network definitions tested as part of this reference architecture were designed to obtain maximum throughput and maximum availability. The design is based on "Scenario 1:3 – Multiple Simple Networks Providing Redundancy and Link Aggregation 802.3ad (LACP) with VLAN Tunneling" from **HP Virtual Connect Ethernet Cookbook: Single and Multi Enclosure Domain (Stacked) Scenarios**

Figure 10: VS2 RA blade network configuration
The DL360 management servers have no need to access the production (guest) network; however they could be configured to access them in the event certain services (NTP, RHN, etc.) are unavailable on the management network. Figure 11 shows the DL360 Gen8 network configuration.

![Figure 11: VS2 RA DL360 network configuration](image)

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

**HP 5920AF-24XG Switches (5920)** – Redundant 10 Gb switches handle the majority of the network 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 configured in a LACP (IEEE 802.3ad) configuration allowing connections to survive failure of an entire switch.

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

The 5120 switch provides the extra necessary 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 iPDU, the operations staff can 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 iPDU are configured to send out SNMP based power events to HP Systems Insight Manager to provide a centralize view.

For fault tolerance, each VS2 RA rack configures four HP iPDU for 32A and 40A, or six HP iPDU 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 delivers the management, monitoring, operational control, and lifecycle requirements is needed for a virtualized environment.

Red Hat Enterprise Virtualization Manager

Red Hat Enterprise Virtualization Manager delivers a centralized management system to administer and control all aspects of a virtualized environment. Red Hat Enterprise Virtualization (RHEV) Manager provides a rich user interface that allows an administrator to manage their virtual infrastructure from a web browser allowing even the most advanced configurations such as network bonding and VLANs to be centrally managed from a graphical console.

In RHEV 3.0, the manager requires a Microsoft Internet Explorer client with .NET 4 runtime for administrative tasks. The user console supports both Firefox and Internet Explorer connections. In either case, the client must connect on the management network using the DNS resolvable name. This reference architecture does not include the Internet Explorer client.

HP Systems Insight Manager

HP Systems Insight Manager (HP SIM) is the foundation for the HP unified server–storage management strategy. HP SIM is a hardware–level management product that supports multiple operating systems on HP ProLiant, Integrity and HP 9000 servers, HP MSA, EVA, XP, and third-party storage arrays. Through a single management view HP SIM provides the basic management features of system discovery and identification, single–event view, inventory data collection, and reporting. The core HP SIM software uses Web Based Enterprise Management (WBEM) to deliver the essential capabilities required to manage all HP server platforms.

HP SIM can be extended to provide systems management with plug-ins for HP client, server, storage, power, and printer products. HP Insight Control and Matrix Operating Environment build on and complement the HP SIM capabilities with deployment, migration, power and performance management, remote monitoring and control, integrated support for virtualization, infrastructure provisioning and optimization, and continuity of services protection. Plug-in applications for workload management, capacity management, virtual machine (VM) management, and partition management using HP Integrity Essentials enable you to choose the value-added software that delivers complete lifecycle management for your hardware assets. With HP Systems Insight Manager and optional Insight software, HP delivers a heterogeneous multi–OS server and storage management solution that enables you to holistically monitor and control your environment which improves operational efficiency and reduces costs.

HP P4000 Centralized Management Console

The HP P4000 Centralized Management Console (CMC) provides the ability to manage multiple P4000 arrays as well as setup advanced features such as multi-site clusters. The CMC software is installed on both nodes of the management cluster.

Built-in high availability

The VS2 RA has been designed to provide high availability at all levels – from the underlying network and storage fabrics to the VM layer. Thus, the VS2 RA tolerates 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 HP 5920 switches and iPDUs, while server components such as power supplies, fans, and storage are hot swappable. The VS2 RA also uses NIC bonding to maximize throughput and availability. High availability between the VS2 RA and your upstream network infrastructure is achieved through Link Aggregation Control Protocol (LACP, IEEE 802.3ad) 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

Virtual machines can be migrated (online or offline) within a Red Hat Enterprise Virtualization cluster from within the virtualization manager. To further enhance the availability of VMs, virtual machines in a VS2 RA can be configured as highly available in Red Hat Enterprise Virtualization Manager. This will restart the virtual machine on another node of the cluster in the event of an unplanned outage of the host. Not all virtual machines need to be configured as highly available, allowing the administrator to control resource usage in the event of hardware downtime.

The availability of online migration of the virtual machine allows for zero downtime of the virtual machine while allowing the hosting server to undergo maintenance. Under Red Hat Enterprise Virtualization Manager clusters can be configured with different policies for different reasons:

- **Load balancing**: Migrating a VM to a different server when the original server's resources are being saturated
- **Power management**: Policies can be configured to locate virtual machines to reduce the number of physical hosts that need to be powered on.

Guest clustering for workload mobility

Optionally, you can create highly available clusters of virtual machines to provide even higher levels of availability for specific workloads. In this scenario, the clustered service runs within the virtual machines in a cluster and manages the movement of a mission-critical workload from one 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 virtual machine.
- **Automatic recovery**: Failing over a database instance from one virtual machine to another in the event of a failure.

Guest clustering is a powerful option that is most useful for protecting mission-critical workloads.

Management cluster configuration

The two-node management cluster was configured using two fencing devices and a quorum disk to ensure proper failure in the event of a 'split-brain' scenario following HP best practices:


iLO 4 power fencing was used as the primary fence device. Fencing using HP iLO 4 requires the use of the ipmilan fence utility and the additional parameter of 'lanplus=1' to work properly. It is recommended that a dedicated iLO user be created with operator level privileges for use by the fence daemon. Using RIBCL (hponcfg) a fence user with the required permissions can be created as follows:

```xml
<RIBCL VERSION="2.0">
  <LOGIN USER_LOGIN="dummy_value" PASSWORD="UsingAutologin">
    <USER_INFO MODE="write">
      <ADD_USER USER_NAME="fence" USER_LOGIN="fence" PASSWORD="password">
        <ADMIN_PRIV value="N"/>
        <REMOTE CONS_PRIV value="Y"/>
        <RESET_SERVER_PRIV value="Y"/>
        <VIRTUAL_MEDIA_PRIV value="Y"/>
        <CONFIG_ILO_PRIV value="N"/>
      </ADD_USER>
    </USER_INFO>
  </LOGIN>
</RIBCL>
```

In the event of a primary fence failure, a secondary fence device was configured using SCSI fencing. This secondary fence will prevent any data corruption on the shared storage volume in the event of a failure of the primary fencing device.

In the event that both cluster nodes are running but unable to communicate with each other (split-brain) a tiebreaker is needed, this third vote is provided by the quorum device. The quorum device uses an iSCSI LUN presented from the P4800. The special two-node configuration of “master_wins” was used to provide the deciding vote in the event of a loss of cluster communications.
All networks on the management nodes (Figure 11) were configured using bonded interfaces – including the directly connected cluster network. The cluster communications network was an active/failover (mode 1) bond on the internal 1Gb NICs. The iSCSI network configuration used adaptive load balancing (mode 6) with an MII link monitoring frequency of 100 milliseconds using a port from each of the PCI-E 10Gb network cards. The management (rhevm) and data center links shared the final network bond using VLAN tags to separate the traffic; access to the data center network is not required so this link could be dedicated to internal management traffic only. The management network bond was configured using IEEE 802.3ad Dynamic link aggregation (LACP/mode 4) to the HP 5920 switches. For details on configuring bonded network interfaces on Red Hat Enterprise Linux refer to Chapter 8 of the Red Hat Enterprise Linux Hypervisor Deployment Guide: [https://access.redhat.com/knowledge/docs/Red_Hat_Enterprise_Linux/](https://access.redhat.com/knowledge/docs/Red_Hat_Enterprise_Linux/).

Cluster services provide highly available access to HP Systems Insight Manager and the Red Hat Enterprise Virtualization Manager. The HP P4000 Centralized Management Console does not rely on any databases; it is simply installed on all management nodes to ensure availability.

Red Hat has documented the process for taking a pair of “Basic Server” installations of Red Hat Enterprise Linux to a cluster running RHEV-M as a highly available cluster service (requires RHN account): [https://access.redhat.com/knowledge/techbriefs/setting-rhev-m-highly-available-cluster-rhev-30](https://access.redhat.com/knowledge/techbriefs/setting-rhev-m-highly-available-cluster-rhev-30)

For details on configuration of HP Systems Insight Manager as a cluster service see Appendix C: Installing the SIM service.

Domain name services (DNS) and network time protocol (NTP) services are provided by the management cluster to keep all systems synchronized. Dynamic host configuration protocol (DHCP) and preboot execution environment (PXE) deployment services, for the management network only, are also hosted on the management nodes.

**Virtual Connect profiles**

The Virtual Connect design is based on “Scenario 1:3 – Multiple Simple Networks Providing Redundancy and Link Aggregation 802.3ad (LACP) with VLAN Tunneling” from HP Virtual Connect Ethernet Cookbook: Single and Multi Enclosure Domain (Stacked) Scenarios [http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01990371/c01990371.pdf](http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01990371/c01990371.pdf)

We define six network connections providing three different networks, one of which will allow VLAN tunneling to provide additional virtual networks to the virtual machines as needed.

```bash
add network DataCenter-1 -quiet NAGs=Default VLanTunnel=Enabled Color=red
add uplinkport enc0:1:X2 Network=DataCenter-1 Speed=Auto
add uplinkport enc0:1:X3 Network=DataCenter-1 Speed=Auto
set network DataCenter-1 SmartLink=Enabled

add network Management-1 -quiet NAGs=Default VLanTunnel=Disabled PrefSpeedType=Custom PrefSpeed=1000 MaxSpeedType=Custom MaxSpeed=1000
add uplinkport enc0:1:X4 Network=Management-1 Speed=Auto
set network Management-1 SmartLink=Enabled

add network iSCSI-1 -quiet NAGs=Default VLanTunnel=Disabled Color=blue
add uplinkport enc0:1:X6 Network=iSCSI-1 Speed=Auto
set network iSCSI-1 SmartLink=Enabled

add network DataCenter-2 -quiet NAGs=Default VLanTunnel=Enabled Color=red
add uplinkport enc0:2:X2 Network=DataCenter-2 Speed=Auto
add uplinkport enc0:2:X3 Network=DataCenter-2 Speed=Auto
set network DataCenter-2 SmartLink=Enabled

add network Management-2 -quiet NAGs=Default VLanTunnel=Disabled PrefSpeedType=Custom PrefSpeed=1000 MaxSpeedType=Custom MaxSpeed=1000
add uplinkport enc0:2:X4 Network=Management-2 Speed=Auto
set network Management-2 SmartLink=Enabled

add network iSCSI-2 -quiet NAGs=Default VLanTunnel=Disabled
add uplinkport enc0:2:X6 Network=iSCSI-2 Speed=Auto
set network iSCSI-2 SmartLink=Enabled
```
Two different VC profiles are required, one for the hosting servers (BL460c Gen8) and a second for the P4460sb components of the P4800 storage. After creating the first instance of each profile they should be copied to create the correct number of profiles for the environment and then assigned to the correct bays.

```plaintext
add profile Compute_01 -NoDefaultEnetConn -NoDefaultFcoeConn
NAG=Default
add enet-connection Compute_01 Network=Management-1 PXE=UseBIOS
add enet-connection Compute_01 Network=Management-2 PXE=UseBIOS
add enet-connection Compute_01 Network=DataCenter-1 PXE=UseBIOS
add enet-connection Compute_01 Network=DataCenter-2 PXE=UseBIOS
add enet-connection Compute_01 Network=iSCSI-1 PXE=UseBIOS
add enet-connection Compute_01 Network=iSCSI-2 PXE=UseBIOS

add profile StorageBlade_01 -NoDefaultEnetConn -NoDefaultFcoeConn -NoDefaultFcConn
NAG=Default SNType=Factory-Default
add enet-connection StorageBlade_01 Network=iSCSI-1 PXE=UseBIOS
add enet-connection StorageBlade_01 Network=iSCSI-2 PXE=UseBIOS
AddressType=Factory-Default
```

**Host node configuration**

Each BL460c Gen8 server is configured as a virtual machine host running Red Hat Enterprise Virtualization Hypervisor 6.2. These servers are configured into a single iSCSI Storage Domain in a single Data Center. Power management is enabled for each host, as with the management cluster, ipmilan is used to communicate with the iLO 4 processors. The power management of these systems utilize the standard fencing daemons so the same additional argument, lanplus=1, is required. Again, HP recommends a special user be created for this purpose. The same RIBCL commands used to create the fence user on the management servers can be used here.

Each node is also configured with three dedicated networks (Figure 10): management (rhevm), iSCSI, and data center. These networks are made highly available by configuring a network bond for each one. A custom configuration of “mode=5 miimon=100” is used to achieve high availability and increased bandwidth beyond the standard active/failover bonding.

The easiest means of installation is to use the virtual media capability of iLO and mount the RHEV-H ISO from the RHEV-M server. To do so it must first be exported via either FTP or HTTP. Red Hat has documented the installation process for RHEV-H in the Red Hat Enterprise Linux Hypervisor Deployment Guide: https://access.redhat.com/knowledge/docs/Red_Hat_Enterprise_Virtualization/

**Final configuration**

The final configuration of the environment relies heavily on the Red Hat Enterprise Virtualization 3.0 Administration Guide: https://access.redhat.com/knowledge/docs/Red_Hat_Enterprise_Virtualization/

Each host needs to be accepted into a data center and cluster before you can configure it. The data center is a loose association of hosts used primary to contain the clusters and storage domain. HP recommends not using the default domain and cluster but configuring new ones based on the policies you want. The process of managing data centers in RHEV-M is covered in Chapter 2 of the Administration Guide.

Once you have the data center defined, configure the individual nodes with power management. In the event of a communication issue, this allows RHEV-M to reboot the node and have it rejoin the cluster. You will not be able to test the power management until you have at least two members in the cluster. Using the IPMI agent configure power management using the IP address of the iLO, the dedicated fence user and ipmilan=1. Details on configuring power management are in Chapter 13 of the Administration Guide.

With power management configured the next step is to configure the networks on each host (RHEV-H) node. These are bonded configurations, all using a custom configuration of “mode=5 miimon=100”. Section 4.1.2.2.2.1 of the Administration guide discusses configuring the bonded interface. HP recommends you name these networks based on their purpose.

- Management network using eth0 and eth1
- Production network using eth2 and eth3
- iSCSI network using eth4 and eth5
Using the HP P4000 CMC, we now need to present storage to the blades for the virtual machine disks. RHEV-M will manage the storage; any LUNs assigned into a storage domain will be joined together and divided up as needed based on the virtual machine definition. The iSCSI software initiator name for each node can be found on the administration console of each node under the “Remote Storage” group. To simplify storage presentation HP recommends creating a ‘cluster’ in the CMC for each cluster you have defined in RHEV-M. Once you have the cluster(s) defined, present LUNs based on your requirements to each of them. Discovering the iSCSI storage and creating the storage domain is covered in Chapter 3 of the Administration Guide.

**Implementing a proof-of-concept**

As a matter of best practice for all deployments, HP recommends implementing a proof-of-concept using a test environment that matches as closely as possible the planned production environment. In this way, appropriate performance and scalability characterizations can be obtained. For help with a proof-of-concept, contact an HP Services representative ([hp.com/large/contact/enterprise/index.html](http://hp.com/large/contact/enterprise/index.html)) or your HP partner.
Appendix A: Bill of Materials (BOM)

This bill of materials is a summary of the key components needed for building a VS2 RA and does not include all required cables and other minor components. Due to varying power options, components of the power configuration are also not included in this bill of materials. Additional cables and transceivers for uplink out of the rack are not included as they are dependent on the end network infrastructure. Ports have been reserved on the IRF’ed 5920 switch pair for this purpose:

- 4 ports for Data Center (Guest/Production) uplink
- 2 ports for iSCSI uplink to connect to external iSCSI
- 2 ports for Management Network.

Additional network cables will also be needed to connect the iPDU into the network infrastructure. Each iPDU will require one (1) CAT5e cable for connection to the 5120 switch.

Because each environment is different support contracts and certain localization options are not included in the BOM. HP recommends a Proactive Care Service solution to correspond with the platform licensing (3 year). Proactive Care Services is available directly from HP and through our global network of authorized ServiceONE Support Partners.

The following BOMs contain electronic license to use (E-LTU) parts. Electronic software license delivery is now available in most countries. HP recommends purchasing electronic products over physical products (when available) for faster delivery and for the convenience of not tracking and managing confidential paper licenses. For more information, please contact your reseller or an HP representative.

VS2 Base Reference Architecture configuration

Table 6: VS2 Base RA configuration components

<table>
<thead>
<tr>
<th>Quantity</th>
<th>HP Part Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Rack Infrastructure</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>BW904A</td>
<td>HP 642 1075mm Shock Intelligent Rack</td>
</tr>
<tr>
<td>1</td>
<td>433718-B21</td>
<td>HP BLc7000 10K Rack Ship Brkt Opt Kit</td>
</tr>
<tr>
<td>1</td>
<td>BW906A</td>
<td>HP 42U 1075mm Side Panel Kit</td>
</tr>
<tr>
<td>1</td>
<td>BW930A</td>
<td>HP Air Flow Optimization Kit</td>
</tr>
<tr>
<td>1</td>
<td>BW891A</td>
<td>HP Rack Grounding Kit</td>
</tr>
<tr>
<td>2</td>
<td>AF070A</td>
<td>HP 1U 10-pack Carbon Universal Filler Panel</td>
</tr>
<tr>
<td>1</td>
<td>BW932A</td>
<td>HP 600mm JackBlack Rack Stabilizer Kit</td>
</tr>
<tr>
<td><strong>Virtualization Hosts</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>5D7019-B21</td>
<td>HP BLc7000 CTO 3 IN LCD ROHS Encl</td>
</tr>
<tr>
<td>4</td>
<td>641016-B21</td>
<td>HP BL460c Gen8 10Gb FLB CTO Blade</td>
</tr>
<tr>
<td>4</td>
<td>662064-L21</td>
<td>HP BL460c Gen8 E5-2670 FIO Kit</td>
</tr>
<tr>
<td>4</td>
<td>662064-B21</td>
<td>HP BL460c Gen8 E5-2670 Kit</td>
</tr>
<tr>
<td>32</td>
<td>672631-B21</td>
<td>HP 16GB 2Rx4 PC3-12800R-11 Kit</td>
</tr>
<tr>
<td>Quantity</td>
<td>HP Part Number</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>------------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Quantity</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>HP Part Number</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Description</strong></td>
</tr>
<tr>
<td>8</td>
<td>652564-B21</td>
<td>HP 300GB 6G SAS 10K 2.5in SC ENT HDD</td>
</tr>
<tr>
<td>4</td>
<td>684212-B21</td>
<td>HP FlexFabric 10Gb 2P 554FLB FIO Adptr</td>
</tr>
<tr>
<td>8</td>
<td>BD503AAE</td>
<td>HP iLO Adv-BL incl 3yr TS U Electronic LTU</td>
</tr>
<tr>
<td>8</td>
<td>A1L28AAE</td>
<td>RHEV 1 Socket 3yr 24x7 Electronic LTU</td>
</tr>
<tr>
<td>2</td>
<td>AJ864A</td>
<td>HP 3Gb SAS BL-c Switch</td>
</tr>
<tr>
<td>1</td>
<td>591973-B21</td>
<td>HP VC Flex-10 Ent Edn for BLc7000 Opt</td>
</tr>
<tr>
<td>1</td>
<td>517521-B22</td>
<td>HP 6X 2400W Plat Ht Plg FIO Power Supply Kit</td>
</tr>
<tr>
<td>1</td>
<td>456204-B21</td>
<td>HP BLc7000 DDR2 Enclosure Mgmt Option</td>
</tr>
<tr>
<td>1</td>
<td>677595-B21</td>
<td>HP BLc 1PH Intelligent Power Mod FIO Opt</td>
</tr>
<tr>
<td>1</td>
<td>517520-B21</td>
<td>HP BLc 6X Active Cool 200 FIO Fan Opt</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Management Servers</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>654081-B21</td>
<td>HP DL360p Gen8 8-SFF CTO Server</td>
</tr>
<tr>
<td>2</td>
<td>664011-L21</td>
<td>HP DL360p Gen8 E5-2690 FIO Kit</td>
</tr>
<tr>
<td>2</td>
<td>664011-B21</td>
<td>HP DL360p Gen8 E5-2690 Kit</td>
</tr>
<tr>
<td>16</td>
<td>672631-B21</td>
<td>HP 16GB 2Rx4 PC3-12800R-11 Kit</td>
</tr>
<tr>
<td>4</td>
<td>652564-B21</td>
<td>HP 300GB 6G SAS 10K 2.5in SC ENT HDD</td>
</tr>
<tr>
<td>2</td>
<td>652238-B21</td>
<td>HP 9.5mm SATA DVD ROM JackBlack Kit</td>
</tr>
<tr>
<td>2</td>
<td>684208-B21</td>
<td>HP Ethernet 1GbE 4P 331FLR FIO Adapter</td>
</tr>
<tr>
<td>2</td>
<td>631679-B21</td>
<td>HP 1GB FBWC for P-Series Smart Array</td>
</tr>
<tr>
<td>2</td>
<td>663201-B21</td>
<td>HP 1U SFF BB Gen8 Rail Kit</td>
</tr>
<tr>
<td>4</td>
<td>614203-B21</td>
<td>HP NC552SFP 10GbE 2P Svr Adapter</td>
</tr>
<tr>
<td>4</td>
<td>656363-B21</td>
<td>HP 750W CS Plat PL Ht Plg Power Supply Kit</td>
</tr>
<tr>
<td>2</td>
<td>BD506AAE</td>
<td>HP iLO Adv incl 3yr TS U Electronic LTU</td>
</tr>
<tr>
<td>2</td>
<td>BC387AAE</td>
<td>Red Hat High-Availability 2 Sockets 3 Year E-LTU</td>
</tr>
<tr>
<td>2</td>
<td>BC321AAE</td>
<td>RHEL Srv, 2 skt, 1 Guest 24x7 3yr Electronic LTU</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Networking</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>592774-B21</td>
<td>HP SL Universal Switch Rail Kit</td>
</tr>
<tr>
<td>1</td>
<td>JE073A</td>
<td>HP 5120-16G S1 Switch</td>
</tr>
<tr>
<td>Quantity</td>
<td>HP Part Number</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>----------------</td>
<td>-------------------------------------------------------------</td>
</tr>
<tr>
<td>2</td>
<td>JG296A</td>
<td>HP 5920AF-24XG Switch</td>
</tr>
<tr>
<td>4</td>
<td>JC680A</td>
<td>HP AJ8x0AF 650W AC Power Supply</td>
</tr>
<tr>
<td>4</td>
<td>JG297A</td>
<td>HP 5920AF-24XG Bk(pwr)-Frt(prt) Fn Tray</td>
</tr>
<tr>
<td>4</td>
<td>C7533A</td>
<td>HP Ethernet CAT5e RJ45 M/M Cable (1.2m)</td>
</tr>
<tr>
<td>4</td>
<td>C7535A</td>
<td>HP Ethernet CAT5e RJ45 M/M Cable (2.1m)</td>
</tr>
<tr>
<td>6</td>
<td>JD089B</td>
<td>HP X120 SFP RJ45 T Transceiver</td>
</tr>
<tr>
<td>4</td>
<td>JD095C</td>
<td>HP X240 SFP+ SFP+ Direct Attach Cable (0.65m)</td>
</tr>
<tr>
<td>16</td>
<td>JD096C</td>
<td>HP X240 SFP+ SFP+ Direct Attach Cable (1.2m)</td>
</tr>
<tr>
<td>0</td>
<td>JD092B</td>
<td>HP X130 SFP+ LC SR Transceiver (optional for uplink)</td>
</tr>
</tbody>
</table>

### Storage

<table>
<thead>
<tr>
<th>Quantity</th>
<th>HP Part Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>BV932A</td>
<td>HP P4800 G2 42TB SAS SAN BladeSystem</td>
</tr>
</tbody>
</table>

**VS2 Extended Reference Architecture configuration**

**Table 7: VS2 Extended RA configuration components**

<table>
<thead>
<tr>
<th>Quantity</th>
<th>HP Part Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td><strong>Rack Infrastructure</strong></td>
</tr>
<tr>
<td>1</td>
<td>BW904A</td>
<td>HP 642 1075mm Shock Intelligent Rack</td>
</tr>
<tr>
<td>1</td>
<td>433718-B21</td>
<td>HP BLc7000 10K Rack Ship Brkt Opt Kit</td>
</tr>
<tr>
<td>1</td>
<td>BW906A</td>
<td>HP 42U 1075mm Side Panel Kit</td>
</tr>
<tr>
<td>1</td>
<td>BW930A</td>
<td>HP Air Flow Optimization Kit</td>
</tr>
<tr>
<td>1</td>
<td>BW891A</td>
<td>HP Rack Grounding Kit</td>
</tr>
<tr>
<td>2</td>
<td>AF070A</td>
<td>HP 1U 10-pack Carbon Universal Filler Panel</td>
</tr>
<tr>
<td>1</td>
<td>BW932A</td>
<td>600mm JackBlack Rack Stabilizer Kit</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Virtualization Hosts</strong></td>
</tr>
<tr>
<td>1</td>
<td>507019-B21</td>
<td>HP BLc7000 CTO 3 IN LCD ROHS Encl</td>
</tr>
<tr>
<td>10</td>
<td>641016-B21</td>
<td>HP BL460c Gen8 10Gb FLB CTO Blade</td>
</tr>
<tr>
<td>10</td>
<td>662064-L21</td>
<td>HP BL460c Gen8 E5-2670 FIO Kit</td>
</tr>
<tr>
<td>Quantity</td>
<td>HP Part Number</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>----------------</td>
<td>-------------</td>
</tr>
<tr>
<td>10</td>
<td>662064-B21</td>
<td>HP BL460c Gen8 E5-2670 Kit</td>
</tr>
<tr>
<td>80</td>
<td>672631-B21</td>
<td>HP 16GB 2Rx4 PC3-12800R-11 Kit</td>
</tr>
<tr>
<td>20</td>
<td>652564-B21</td>
<td>HP 300GB 6G SAS 10K 2.5in SC ENT HDD</td>
</tr>
<tr>
<td>10</td>
<td>684212-B21</td>
<td>HP FlexFabric 10Gb 2P 554FLB FIO Adptr</td>
</tr>
<tr>
<td>16</td>
<td>BD503AAE</td>
<td>HP ILO Adv-Bl incl 3yr TS U Electronic LTU</td>
</tr>
<tr>
<td>20</td>
<td>A1L28AAE</td>
<td>RHEV 1 Sckt 3yr 24x7 Electronic LTU</td>
</tr>
<tr>
<td>2</td>
<td>AJ864A</td>
<td>HP 3Gb SAS BL-c Switch</td>
</tr>
<tr>
<td>1</td>
<td>591973-B21</td>
<td>HP VC Flex-10 Ent Edn for BLc7000 Opt</td>
</tr>
<tr>
<td>1</td>
<td>517521-B22</td>
<td>HP 6X 2400W Plat Ht Plg FIO Pwr Sply Kit</td>
</tr>
<tr>
<td>1</td>
<td>456204-B21</td>
<td>HP BLc7000 DDR2 Encl Mgmt Option</td>
</tr>
<tr>
<td>1</td>
<td>677595-B21</td>
<td>HP BLc 1PH Intelligent Power Mod FIO Opt</td>
</tr>
<tr>
<td>1</td>
<td>517520-B21</td>
<td>HP BLc 6X Active Cool 200 FIO Fan Opt</td>
</tr>
<tr>
<td><strong>Management Servers</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>654081-B21</td>
<td>HP DL360p Gen8 B-SFF CTO Server</td>
</tr>
<tr>
<td>2</td>
<td>664011-L21</td>
<td>HP DL360p Gen8 ES-2690 FIO Kit</td>
</tr>
<tr>
<td>2</td>
<td>664011-B21</td>
<td>HP DL360p Gen8 ES-2690 Kit</td>
</tr>
<tr>
<td>16</td>
<td>672631-B21</td>
<td>HP 16GB 2Rx4 PC3-12800R-11 Kit</td>
</tr>
<tr>
<td>4</td>
<td>652564-B21</td>
<td>HP 300GB 6G SAS 10K 2.5in SC ENT HDD</td>
</tr>
<tr>
<td>2</td>
<td>652238-B21</td>
<td>HP 9.5mm SATA DVD ROM JackBlack Kit</td>
</tr>
<tr>
<td>2</td>
<td>684208-B21</td>
<td>HP Ethernet 1GbE 4P 331FLR FIO Adptr</td>
</tr>
<tr>
<td>2</td>
<td>631679-B21</td>
<td>HP 1GB FBWC for P-Series Smart Array</td>
</tr>
<tr>
<td>2</td>
<td>663201-B21</td>
<td>HP 1U SFF BB Gen8 Rail Kit</td>
</tr>
<tr>
<td>4</td>
<td>614203-B21</td>
<td>HP NC5525FP 10GbE 2P Svr Adapter</td>
</tr>
<tr>
<td>4</td>
<td>656363-B21</td>
<td>HP 750W CS Plat PL Ht Plg Pwr Supply Kit</td>
</tr>
<tr>
<td>2</td>
<td>BD506AAE</td>
<td>HP ILO Adv incl 3yr TS U Electronic LTU</td>
</tr>
<tr>
<td>2</td>
<td>BC387AAE</td>
<td>Red Hat High-Availability 2 Sockets 3 Year E-LTU</td>
</tr>
<tr>
<td>2</td>
<td>BC321AAE</td>
<td>RHEL Srv, 2 skt, 1 Guest 24x7 3yr Electronic LTU</td>
</tr>
<tr>
<td>Quantity</td>
<td>HP Part Number</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>----------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>Networking</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>592774-B21</td>
<td>HP SL Universal Switch Rail Kit</td>
</tr>
<tr>
<td>1</td>
<td>JE073A</td>
<td>HP 5120-16G SI Switch</td>
</tr>
<tr>
<td>2</td>
<td>JG296A</td>
<td>HP 5920AF-24XG Switch</td>
</tr>
<tr>
<td>4</td>
<td>JC680A</td>
<td>HP A58x0AF 650W AC Power Supply</td>
</tr>
<tr>
<td>4</td>
<td>JG297A</td>
<td>HP 5920AF-24XG Bk(pwr)-Frt(prt) Fn Tray</td>
</tr>
<tr>
<td>4</td>
<td>C7533A</td>
<td>HP Ethernet CAT5e RJ45 M/M Cable (1.2m)</td>
</tr>
<tr>
<td>4</td>
<td>C7535A</td>
<td>HP Ethernet CAT5e RJ45 M/M Cable (2.1m)</td>
</tr>
<tr>
<td>6</td>
<td>J0089B</td>
<td>HP X120 SFP RJ45 T Transceiver</td>
</tr>
<tr>
<td>4</td>
<td>J0095C</td>
<td>HP X240 SFP+ SFP+ Direct Attach Cable (0.65m)</td>
</tr>
<tr>
<td>16</td>
<td>J0096C</td>
<td>HP X240 SFP+ SFP+ Direct Attach Cable (1.2m)</td>
</tr>
<tr>
<td>0</td>
<td>J0092B</td>
<td>HP X130 SFP+ LC SR Transceiver (optional for uplink)</td>
</tr>
<tr>
<td>Storage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>BV932A</td>
<td>HP P48000 G2 42TB SAS SAN BladeSystem</td>
</tr>
</tbody>
</table>
## Appendix B: Software/Firmware versions

The reference configuration was created using these versions of software and firmware.

<table>
<thead>
<tr>
<th>Description</th>
<th>Version</th>
</tr>
</thead>
<tbody>
<tr>
<td>HP 5920AF-24XG Switch</td>
<td>Version 7.1.023, Release 2108P02</td>
</tr>
<tr>
<td>HP 5120-16G SI Switch</td>
<td>Version 5.20, Release 1512P05</td>
</tr>
<tr>
<td>ProLiant DL360p Gen8</td>
<td>P71 02/25/2012</td>
</tr>
<tr>
<td>HP Ethernet 1GbE 4P 331FLR FIO Adptr</td>
<td>5719-v1.29</td>
</tr>
<tr>
<td>HP NC522SFP 10GbE 2P Svr Adapter</td>
<td>4.1.402.20</td>
</tr>
<tr>
<td>HP Smart Array P420i Controller 2.14</td>
<td>2.14</td>
</tr>
<tr>
<td>Red Hat Enterprise Linux</td>
<td>6.2</td>
</tr>
<tr>
<td>Red Hat Enterprise Virtualization Manager</td>
<td>3.0</td>
</tr>
<tr>
<td>HP Systems Insight Manager</td>
<td>7.1</td>
</tr>
<tr>
<td>BladeSystem c7000 DDR2 Onboard Administrator with KVM</td>
<td>3.56</td>
</tr>
<tr>
<td>HP VC Flex-10 Ent Edn for Blc7000 Opt</td>
<td>3.60</td>
</tr>
<tr>
<td>HP 3Gb SAS BL-c Switch</td>
<td>2.82</td>
</tr>
<tr>
<td>HP SAS Management Software</td>
<td>2.2.17.0</td>
</tr>
<tr>
<td>ProLiant BL460c Gen8</td>
<td>System ROM131 02/25/2012</td>
</tr>
<tr>
<td>iLO 4</td>
<td>1.05 Apr 19 2012</td>
</tr>
<tr>
<td>HP FlexFabric 10Gb 2P 554FLB FIO Adptr</td>
<td>4.1.402.20</td>
</tr>
<tr>
<td>Red Hat Enterprise Virtualization Hypervisor</td>
<td>6.2</td>
</tr>
<tr>
<td>SAN iQ</td>
<td>9.5.00.1215.0</td>
</tr>
<tr>
<td>HP P4000 Centralized Management Console</td>
<td>Version 9.5.01, Build 0282.0</td>
</tr>
<tr>
<td>P4460sb</td>
<td>System ROM127 05/05/2011</td>
</tr>
<tr>
<td>iLO 3</td>
<td>1.28 Jan 13 2012</td>
</tr>
<tr>
<td>MDS600</td>
<td>3.60</td>
</tr>
</tbody>
</table>
Appendix C: Installing the SIM service

The process included here for installing SIM as a cluster service assumes that you already have a functioning cluster and shared storage configured. We further assume that the storage is managed by HA LVM and not CLVM.

Before we install HP SIM we need to make sure that all the dependencies are met on all management servers. This dependency information is based on HP SIM 7.1 installation for Red Hat Enterprise Linux 6.2 (x86-64) and is subject to change.

<table>
<thead>
<tr>
<th>Package</th>
<th>Architecture</th>
</tr>
</thead>
<tbody>
<tr>
<td>net-snmp</td>
<td>x86-64</td>
</tr>
<tr>
<td>cluster-snmp</td>
<td>x86-64</td>
</tr>
<tr>
<td>glibc</td>
<td>i686</td>
</tr>
<tr>
<td>nss-softokn-freebl</td>
<td>i686</td>
</tr>
<tr>
<td>libgcc</td>
<td>i686</td>
</tr>
<tr>
<td>libstdc++</td>
<td>i686</td>
</tr>
<tr>
<td>zlib</td>
<td>i686</td>
</tr>
<tr>
<td>ncurses-libs</td>
<td>i686</td>
</tr>
<tr>
<td>cyrus-sasl-gssapi</td>
<td>i686</td>
</tr>
<tr>
<td>keyutils-libs</td>
<td>i686</td>
</tr>
<tr>
<td>krb5-libs</td>
<td>i686</td>
</tr>
<tr>
<td>libcom_err</td>
<td>i686</td>
</tr>
<tr>
<td>libselinux</td>
<td>i686</td>
</tr>
<tr>
<td>popt</td>
<td>i686</td>
</tr>
<tr>
<td>pam</td>
<td>i686</td>
</tr>
<tr>
<td>audit-libs</td>
<td>i686</td>
</tr>
<tr>
<td>cracklib</td>
<td>i686</td>
</tr>
<tr>
<td>db4</td>
<td>i686</td>
</tr>
</tbody>
</table>

```
```
The latest version of the HP Systems Insight Manager software can be downloaded from [hp.com/go/hpsim](http://hp.com/go/hpsim). Install HP SIM on all management nodes using the manual installation process as documented in the installation guide, with one exception. Replace the ‘rpm -i’ commands with ‘yum install’ commands. This will avoid future complaints from yum about the database being modified outside of yum. For increased security HP recommends not disabling SELinux or the firewalls. Instructions on adding the necessary rules to the firewall are included below; for Red Hat Enterprise Linux 6.2 no changes to the default (targeted) SELinux policy are required.

http://h20564.www2.hp.com/portal/site/hpsc/public/kb/docDisplay/?docId=emr_na-c03345285

```bash
# cd /tmp
# sh HP_SIM_* .bin --noexec --target SIM
# cd SIM
# yum install -y hpsmdb-rhel6-*.$(uname -m).rpm
# chkconfig hpsmdb on
# service hpsmdb start
# yum install -y hpsim*.rpm
# /opt/mx/bin/mxinitconfig -l
```

If there are errors with the checks resolve them before continuing, see the install guide for troubleshooting. When all checks pass, continue to configure HP SIM and verify the daemons start. The final configuration (`mxinitconfig -a`) should only be completed on one server.

```bash
# /opt/mx/bin/mxinitconfig -a
# ps -elf | grep mx
```

Now we need to stop the services and disable them to ensure they don’t try to start on reboot.

```bash
# service hpsim stop ; chkconfig hpsim off
# service hpsmdb stop ; chkconfig hpsmdb off
```

The installation of SIM as a clustered service will be very similar to the installation of RHEV-M. First, we want to define an IP address for the cluster service. If you are using the local DNS configuration `sim.example.com (172.23.1.4)` has already been defined for this purpose.

```bash
# ccs -h localhost --addresource ip address=172.23.1.4 monitor_link=on sleeptime=10
# ccs -h localhost --addresource ip address=172.23.1.4 monitor_link=on sleeptime=10
# ccs -h localhost --addservice sim domain=prefer_mgmt2 recovery=relocate
# ccs -h localhost --addsubservice sim ip ref=172.23.1.4
# ccs -h localhost --addsubservice sim ip ref=172.23.16.4
# ccs -h localhost --sync --activate
# clustat -s sim
Service Name  Owner (Last)  State
--------  --------  ------
service:sim  mgmt-node2  started
```

On the node with the HP SIM service configured, we now need to add the shared storage.

```bash
# parted -s /dev/disk/by-path/ip-172.23.16.254:*
  /dev/sdb:53.7GB
1049kB 53.7GB primary
```

---

**Table:**

<table>
<thead>
<tr>
<th>Service Name</th>
<th>Owner (Last)</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>service:sim</td>
<td>mgmt-node2</td>
<td>started</td>
</tr>
</tbody>
</table>
Now that the device has been partitioned we need to initialize it for use by LVM and create the volume group.

```bash
pvcreate /dev/disk/by-path/ip-172.23.16.254:*\:mgmt-sim-lun-0-part1
vgcreate SIMVG /dev/disk/by-path/ip-172.23.16.254:*\:mgmt-sim-lun-0-part1
vgchange --addtag @mgmt-node2 SIMVG
lvcreate -n lv_sim -l 100%FREE SI
```

You may have noticed there is an extra step to activate the volume group; this is because we are running the HA LVM. Before returning control to the cluster we will need to remove this tag as well. Make and mount the filesystem at /hpsimlnx.

```bash
mkfs -t ext4 /dev/SIMVG/lv_sim
mkdir -p /hpsimlnx
mount /dev/SIMVG/lv_sim /hpsimlnx
```

Now we need to move the files from the local storage to the shared storage.

```bash
# cd /hpsimlnx
# mkdir -p etc/sysconfig etc/pam.d etc/opt etc/init.d var/log var/opt opt
# mv {,/hpsimlnx}/etc/init.d/hpsim
# mv {,/hpsimlnx}/etc/init.d/hpsmdb
# mv {,/hpsimlnx}/etc/pam.d/hpsmdb
# mv {,/hpsimlnx}/etc/pam.d/mxpamauthrealm
# mv {,/hpsimlnx}/etc/sysconfig/hpsmdb
# mv {,/hpsimlnx}/etc/opt/mx
# mv {,/hpsimlnx}/opt/hpsmdb
# mv {,/hpsimlnx}/opt/hpwebadmin
# mv {,/hpsimlnx}/opt/mx
# mv {,/hpsimlnx}/var/log/hpsmdb
# mv {,/hpsimlnx}/var/opt/mx
# mv {,/hpsimlnx}/var/opt/hpsmdb
# ln -sf {/hpsimlnx,}/etc/init.d/hpsim
# ln -sf {/hpsimlnx,}/etc/init.d/hpsmdb
# ln -sf {/hpsimlnx,}/etc/pam.d/hpsmdb
# ln -sf {/hpsimlnx,}/etc/pam.d/mxpamauthrealm
# ln -sf {/hpsimlnx,}/etc/sysconfig/hpsmdb
# ln -sf {/hpsimlnx,}/etc/opt/mx
# ln -sf {/hpsimlnx,}/opt/hpsmdb
# ln -sf {/hpsimlnx,}/opt/hpwebadmin
# ln -sf {/hpsimlnx,}/opt/mx
# ln -sf {/hpsimlnx,}/var/log/hpsmdb
# ln -sf {/hpsimlnx,}/var/opt/mx
# ln -sf {/hpsimlnx,}/var/opt/hpsmdb
```
On the other node, we need to clean up the install.

```bash
# rm -rf /etc/init.d/hpsim
# rm -rf /etc/init.d/hpsmdb
# rm -rf /etc/pam.d/mxpamauthrealm
# rm -rf /etc/sysconfig/hpsmdb
# rm -rf /etc/opt/mx
# rm -rf /opt/hpsmdb
# rm -rf /opt/hpwebadmin
# rm -rf /opt/mx
# rm -rf /var/log/hpsmdb
# rm -rf /var/opt/mx
# rm -rf /var/opt/hpsmdb

# ln -sf {/hpsimlnx,}/etc/init.d/hpsim
# ln -sf {/hpsimlnx,}/etc/init.d/hpsmdb
# ln -sf {/hpsimlnx,}/etc/pam.d/mxpamauthrealm
# ln -sf {/hpsimlnx,}/etc/sysconfig/hpsmdb
# ln -sf {/hpsimlnx,}/etc/opt/mx
# ln -sf {/hpsimlnx,}/opt/hpsmdb
# ln -sf {/hpsimlnx,}/opt/hpwebadmin
# ln -sf {/hpsimlnx,}/opt/mx
# ln -sf {/hpsimlnx,}/var/log/hpsmdb
# ln -sf {/hpsimlnx,}/var/opt/mx
# ln -sf {/hpsimlnx,}/var/opt/hpsmdb
```
Once we have all the files transferred to the shared storage, there are a few updates we need to make. Because both HP SIM and RHEV-M are running instances of PostgreSQL we need to update the status logic for hpsmdb to reflect this. To do this we create a wrapper, this avoids issues caused by upgrading SIM in the future.

```bash
# cat > /hpsimlnx/etc/init.d/hpsmdb-rhcs <<EOF
#!/bin/sh
# hpsmdb-rhcs Wrapper to cleanup HPSMDB for RHCS use

INITD=${0%/*}

# For SELinux we need to use 'runuser' not 'su'
SU=runuser
if test -x /sbin/runuser ; then
  SU=runuser
fi

# Set defaults for configuration variables
PGENGINE=/hpsimlnx/opt/hpsmdb/bin
PGDATA=/hpsimlnx/var/opt/hpsmdb/data
PGPORT=50006

# See how we were called.
case "$1" in
  start)
    if test -f ${PGDATA}/postmaster.pid ; then
      $0 status
      if test $? -ne 0 ; then
        echo Cleaning up unclean exit
        rm -f ${PGDATA}/postmaster.pid
      fi
    fi
    exec $(INITD)/hpsmdb start
    ;;
  stop)
    exec $(INITD)/hpsmdb stop
    ;;
  status)
    exec $(SU) -l hpsmdb -c "pg_ctl -D '${PGDATA}' status"
    ;;
  restart)
    exec $(INITD)/hpsmdb restart
    ;;
  *)
    echo "$Usage: $0 {start|stop|status|restart}"
    exit 1
esac

exit 1
EOF
# chmod +x /hpsimlnx/etc/init.d/hpsmdb-rhcs
```

The HP SIM init script does not provide a means to query status, to be able to use it as a ‘script’ sub service we need to create a wrapper that does.

```bash
# cat > /hpsimlnx/etc/init.d/hpsim-rhcs <<EOF
#!/bin/sh
# This is the init script for starting up the HPSIM under RHCS
#
# chkconfig: - 99 01
# description: Starts and stops the HPSIM backend daemons
# processname: hpsim-rhcs
# pidfile: /var/run/hpsim-rhcs.pid

# Source function library.
INITD=/etc/rc.d/init.d
. $INITD/functions

# Get function listing for cross-distribution logic.
TYPESET=`typeset -figrep "declare"`
EOF
```
NAME=hpsim-rhcs
PIDFILE=/var/run/hpsim-rhcs.pid

# Check that networking is up.
# Pretty much need it
[ "${NETWORKING}" = "no" ] && & & exit 0

script_result=0

start(){
    echo -n "Starting ${NAME} service: "
    /etc/init.d/hpsim start
    script_result=$?
    pid=$(pidof -s /opt/mx/lbin/mxdomainmgr /opt/mx/lbin/mxdtf)
    echo $pid > ${PIDFILE}
    if [ $script_result -eq 0 ]; then
        success "${NAME}"
        touch /var/lock/subsys/${NAME}
    else
        failure "${NAME} (hpsim)"
        fi
    echo
}

stop(){
    echo -n "$Stopping ${NAME} service: "
    /etc/init.d/hpsim stop
    if [ $? -eq 0 ] ; then
        echo_success
        rm -f /var/lock/subsys/${NAME}
        rm -f ${PIDFILE}
    else
        echo_failure
        script_result=1
        fi
    echo
}

restart(){
    stop && start
    script_result=$?
}

# See how we were called.
case "$1" in
    start)
        start
        ;
    stop)
        stop
        ;
    status)
        status /opt/mx/lbin/mxdomainmgr && status /opt/mx/lbin/mxdtf
        ;
    restart)
        restart
        ;
    *)
        echo "$Usage: $0 {start|stop|status|restart}"
        exit 1
esac

exit $script_result

EOF

chmod +x /hpsimlnx/etc/init.d/hpsim-rhcs

Finally we need to update some configuration files to ensure that HP SIM uses the cluster alias when configuring other systems.

# sed -i 's/localhost/sim/ /hpsimlnx/etc/opt/mx/config/mx.properties
Now that we have all the configuration and scripts in place, we finish defining the cluster service. But first we need to stop everything and unmount and release the LVM volume.

```bash
# cd /
# umount /hpsimlnx
# vgchange -a n -deltag @mgmt-node2 SIMVG
```

With the services stopped and the file systems unmounted we can update the cluster service to include the new sub-services for the file system and SIM.

```bash
# ccs -h localhost --addresource lvm name="SIMVG" self_fence="on" vg_name="SIMVG"
# ccs -h localhost --addresource fs fstype=ext4 device=/dev/SIMVG/lv_sim name=lv_sim mountpoint=/hpsimlnx
# ccs -h localhost --addsubservice sim lvm ref=SIMVG
# ccs -h localhost --addsubservice sim lvm:fs ref=lv_sim
# ccs -h localhost --addresource script name="hpsmdb-rhcs" file="/hpsimlnx/etc/init.d/hpsmdb-rhcs"
# ccs -h localhost --addsubservice sim script ref=hpsmdb-rhcs
# ccs -h localhost --addsubservice sim script:script ref=hpsim-rhcs
# ccs -h localhost --sync --activate
```

Now we need to open the firewall to allow incoming connections on the SIM web interface and SNMP alerts.

```bash
# iptables -N SIM
# iptables -A SIM -m state --state NEW -p tcp --dport http-mgmt -j ACCEPT
# iptables -A SIM -m state --state NEW -p tcp --dport 50000 -j ACCEPT
# iptables -A SIM -m state --state NEW -p tcp --dport snmptrap -j ACCEPT
# iptables -A SIM -m state --state NEW -p udp --dport snmptrap -j ACCEPT
# iptables -A SIM -m state --state NEW -p tcp --dport 50004 -j ACCEPT
# iptables -A Mgmt_Network -p tcp --dest 172.23.1.4 -j SIM
# iptables -A Mgmt_Network -p udp --dest 172.23.1.4 -j SIM
# iptables -A iSCSI_Network -p tcp --dest 172.23.16.4 -j SIM
# iptables -A iSCSI_Network -p udp --dest 172.23.16.4 -j SIM
# service iptables save
```

With the service up and running, we need to replace the certificate that was generated during install with one that reflects the cluster name/IP.

```bash
# mxcert -n CN=sim.example.com ALT=172.23.1.4,172.23.16.4 OU="Red Hat RA" O=HP L=EXAMPLE ST=CA C=US
WARNING: Creating a new server certificate will invalidate any existing trust relationships between the central management server and other systems that rely on the certificate, such as browsers, and including HP web-based managed systems configured to trust this central management server.

Additionally, any applications on the central management server using the shared certificate directory will also be updated with a new server certificate and private key, which may affect their operation.

You must restart HP Systems Insight Manager for this change to properly take effect throughout the application.

Do you wish to continue? [Y=yes,N=no] y
```

Now restart the cluster service and verify the new certificate is in use. It could take several minutes until the SIM web interface is back online.

```bash
# clusvcadm -R service:sim
# sleep 300
# openssl s_client -connect 172.23.1.4:50000 < /dev/null | grep "^subject=/C=US/ST=CA/L=EXAMPLE/O=HP/CN=Red Hat RA/CN=sim.example.com"
```
At this point you can connect to the web interface via your favorite browser and configure HP SIM for your environment. Alternatively, the command line can be used to perform the configuration; for example, defining the default credentials and adding hosts.

http://h20564.www2.hp.com/portal/site/hpsc/public/kb/docDisplay/?docId=emr_na-c03345226

```
# mxnodesecurity -a -p ssh -c admin:Password1234
# mxnodesecurity -a -p ssh -c root:Password1234
Listing all global credentials...

NODENAME  PROTOCOL  USERNAME  PASSWORD
@default1  snmp       public    private
@default1  ssh        admin     ********
@default2  ssh        root      ********
Listing all system credentials...
Could not list specified credential(s)
Credentials may not exist
```

The order in which hosts are added is important. It is recommended that the iLO/OA be discovered before the operating systems.

```
# mxnode -a mgmt1-ilo 172.23.0.1
Node has been added/modified:
Node has been added/modified: mgmt1-ilo
```

Discovery of the operating systems may require some configuration for maximum value, for example opening the firewall or setting the SNMP trap destination. The latter can be done by SIM upon discovery of the system. At a minimum HP recommends opening SSH and WBEM inbound to the systems to be managed; this should be restricted to the IP address of the management cluster nodes.

```
# iptables -N SIM_Discovery
# iptables -A SIM_Discovery -m state --state NEW -p tcp --dport ssh -j ACCEPT
# iptables -A SIM_Discovery -m state --state NEW -m multiport -p tcp --dports 80,280,2301,2381,50000 -j ACCEPT
# iptables -A Mgmt_Network -p tcp --src 172.23.1.1,172.23.1.2 -j SIM_Discovery
# service iptables save
```

SNMP monitoring allows HP SIM to collect performance and thermal information from the system. However, due to its insecure nature many environments do not allow SNMP traffic. To address this concern the ProLiant Gen8 servers offer the HP Agentless Management Service. This service, installed from the HP Software Delivery Repository (http://downloads.linux.hp.com/SDR), provides operating system SNMP information through iLO. This allows for more security by not requiring SNMP access to the server and isolates management traffic to the management network.


A complete list of the ports that can be opened, including those for HP System Management Homepage (SMH) can be found in Managing HP servers through firewalls with Insight Management 7.1 or greater White Paper http://h20564.www2.hp.com/portal/site/hpsc/public/kb/docDisplay/?docId=emr_na-c03345481
For more information

HP ProLiant Servers  hp.com/go/proliant
HP BladeSystem  hp.com/go/bladesystem
HP Systems Insight Manager  hp.com/go/sim
HP Networking  hp.com/go/networking
HP P4000 Storage  hp.com/go/p4000
HP VirtualSystem  hp.com/go/virtualsystem
HP & Red Hat  hp.com/go/redhat

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

© Copyright 2012 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

Microsoft and Windows are U.S. registered trademarks of Microsoft Corporation. Intel and Xeon are trademarks of Intel Corporation in the U.S. and other countries.

4AA4-4310ENW, Created October 2012