Building a Microsoft Windows Server 2008 R2 Hyper-V failover cluster with HP Virtual Connect FlexFabric

Technical white paper

Table of contents

Executive summary .......................................................................................................................... 2
Overview ....................................................................................................................................... 2
   HP Virtual Connect FlexFabric ................................................................................................. 4

Building a Windows Server 2008 R2 Hyper-V failover cluster with HP FlexFabric .................. 5
   Setting up the Virtual Connect FlexFabric infrastructure ....................................................... 6
   Installing and configuring the Hyper-V servers ..................................................................... 16

Summary ....................................................................................................................................... 28

For more information .................................................................................................................. 29
Executive summary

The server edge is one of the most complex areas of today’s data center infrastructure, with the most server to server traffic and congestion bottlenecks. The growing demands for virtual servers and applications and their associated networking needs exacerbates this problem creating a sprawl of wires, adapter cards, switches, and processes that are out of control, with complexity that isn’t manageable, costs that are unsustainable, and an architectural approach that isn’t easily scalable.

By implementing HP Virtual Connect FlexFabric, businesses can eliminate network sprawl at the server edge, increase and control network utilization, eliminate confusion regarding network convergence and begin to use it effectively. They can now benefit economically from the advantages of Converged Infrastructure and converged network technology.

This white paper provides an illustration of how to build a two-node Microsoft® Windows® Server 2008 R2 Hyper-V Failover Cluster with HP ProLiant server blades and HP Virtual Connect (VC) FlexFabric. The focus of the document is on the steps necessary to configure the VC FlexFabric modules based on an example environment. This includes more complex details such as setting up NIC teaming and VLAN configurations correctly for the Hyper-V configuration.

The purpose of this example scenario is to test and validate that live migration support for a test virtual machine (VM) is functioning correctly in this configuration. Using the steps detailed in this paper, a VM was created and both planned (live migration) and unplanned failover tests were successfully completed.

**Target audience:** This white paper is intended for several audiences, including decision makers and architects responsible for designing and planning the data center infrastructure as well as administrators responsible for deploying and configuring the environment. This white paper assumes that the reader has a base level of familiarity with HP Virtual Connect and Windows Server 2008 R2 Hyper-V virtualization and live migration. For links to additional background information on these topics see the “For more information” section at the end of this white paper.

This white paper describes testing performed in May 2010.

Overview

Server virtualization is quickly becoming a key strategic element in many organizations, addressing a number of different issues companies are facing in their data centers. For companies facing significant physical server sprawl and a lack of space, power or cooling resources, virtualization provides a platform for consolidation onto virtual machines (VMs). This not only provides a reduction in the number of physical servers (saving on real estate, power and cooling costs) but also addresses and improves on server utilization rates. But beyond the consolidation benefits, another significant business driver for virtualization is to improve data center flexibility.

A dynamic business requires a dynamic IT infrastructure – a data center that provides the flexibility needed to dynamically add additional server processing power for handling spikes in web traffic and the agility to shift resources from one application to another as resource demands change. And while server virtualization is one of the components of building a dynamic data center, to truly achieve this vision, companies must start thinking beyond just server virtualization and look at pooling and virtualizing resources across the entire datacenter. This rational is why virtualization is one of the five overarching design requirements of HP Converged Infrastructure.

HP Converged Infrastructure is a framework for building a dynamic data center that eliminates costly and rigid IT silos, and unlocks resources for IT innovation rather than IT management. Along with virtualization, the Converged Infrastructure has four other overarching requirements: being resilient, orchestrated, optimized and modular. The HP Converged Infrastructure matches the supply of IT
resources with the demand for business applications. By transitioning away from a product-centric approach to a shared-service management model, organizations can accelerate standardization, reduce operational costs, and accelerate business results.

There are four core components of the HP Converged Infrastructure architecture framework as shown in Figure 1 below.

---

**Figure 1 – HP Converged Infrastructure architecture**

Virtual Connect is a cornerstone of the HP Converged Infrastructure strategy because it simplifies the server edge by consolidating, virtualizing and automating server connections. With Virtual Connect FlexFabric – using Converged Enhanced Ethernet (CEE) and Fibre Channel over Ethernet (FCoE), which is built-in to every ProLiant G7 server blade – HP will lead the next industry transformation toward the simplest converged connections for Converged Infrastructure. The HP Virtual Connect

---

FlexFabric is a virtualized, high-performance, low latency network that consolidates Ethernet and storage networks onto a single fabric to dramatically lower network complexity and cost.

**HP Virtual Connect FlexFabric**

Virtual Connect FlexFabric represents the 3rd generation of HP’s award winning Virtual Connect technology with over 2.6 million ports deployed in data centers today. HP Virtual Connect FlexFabric is a solution to this growing problem that simplifies the server edge by replacing traditional switches and modules with a converged and virtualized way to connect servers to different networks with one device, over one wire, utilizing industry standards. HP Virtual Connect FlexFabric allows you to eliminate separate interconnect modules and mezzanine cards for disparate protocols like Ethernet and FC. This allows you to consolidate all of the traffic onto a single interconnect architecture, simplifying the design and cabling complexity of your environment.

And beyond just these benefits, VC FlexFabric modules also provide significant flexibility in designing the network and storage connectivity requirements for your server blades. All of the new ProLiant G7 blades will include VC FlexFabric capability built-in with integrated VC FlexFabric NICs, and any ProLiant G6 blades can be easily upgraded to support VC FlexFabric with new Virtual Connect FlexFabric Adapter mezzanine cards. These adapters and mezzanine cards have two, 10 Gb FlexFabric ports which can be carved up into as many as 4 physical functions per port. A physical function can be either a FlexNIC, FlexHBA-FCoE, or FlexHBA-iSCSI supporting Ethernet, FCoE, or iSCSI traffic respectively. Each adapter port can have up to 4 FlexNIC physical functions if there are no storage requirements or 3 FlexNICs and 1 FlexHBA physical function for FCoE or iSCSI connectivity. And the bandwidth of these ports can be configured to satisfy the requirements of your environment. Each port is capable of up to 10 Gb which can be allocated as necessary to the physical functions of that port.

This gives you a tremendous amount of flexibility when designing and configuring your network and storage requirements for your environment. This is especially salient for virtualization environments which have a requirement for a lot of different networks that have varying bandwidth and segmentation requirements. In a traditional networking design, this would require a number of additional network cards, cables, and uplink ports which quickly drive up the total cost of the solution. With VC FlexFabric, you have the unique flexibility to allocate and fine-tune network and storage bandwidth for each connection and define each of those physical functions with the specific bandwidth requirements for that network, without having to overprovision bandwidth based on static network speeds.
Building a Windows Server 2008 R2 Hyper-V failover cluster with HP FlexFabric

The purpose of this white paper is to describe the implementation steps that are required to build a Windows Server 2008 R2 Hyper-V failover cluster supporting live migration with HP VC FlexFabric. Specifically the focus is on the steps necessary to setup the underlying VC FlexFabric infrastructure to support the failover cluster configuration. This includes two of the more complex configuration tasks, setting up VLANs and NIC teaming.

This document will focus on an example scenario to implement a two-node Hyper-V failover cluster with live migration support. The example will demonstrate how to design and implement shared SAN storage, NIC teaming for high availability on the hosts, and a multi-network, VLAN scenario on a single VM guest; the type of networking requirements that a multi-tier application such as Microsoft SharePoint 2010, Exchange Server 2010, or SQL Server 2008 R2 might require. Of course the specific networking and storage requirements for your environment will vary significantly depending upon what application workloads are being virtualized. But this can serve as a baseline reference configuration, and the steps described in this document can be used as a model for building larger clusters with more than two nodes and a single VM.

For this example, a single HP BladeSystem c7000 enclosure with two HP ProLiant BL460c G6 server blades will be used to run the two-node Hyper-V cluster. To support the backend storage requirements, an HP StorageWorks B400 Enterprise Virtual Array (EVA8400) FC array is utilized. To provide Ethernet networking and FC storage connectivity to the c7000 enclosure, two HP Virtual Connect FlexFabric modules are installed in bays 3 and 4 and two HP NC551m FlexFabric Adapters are installed in the corresponding mezzanine slot (slot 1) in the blades. The reason that the FlexFabric modules are installed in bays 3 and 4 is because the signal paths for the blade mezzanine cards in slot 1 are connected to those bays. One port is mapped to bay 3 and one port is mapped to bay 4. Thus, installing the two FlexFabric modules provides additional bandwidth and redundancy for the server network and storage connections.

The VC FlexFabric modules form the foundation of the converged networking fabric for this implementation, consolidating both upstream FC and Ethernet connections and downstream Ethernet and Fibre Channel over Ethernet (FCoE) connections into a single pair of modules.

Table 1 lists the networking connectivity requirements for this example. While it’s not mandatory, it’s helpful to have a clear understanding of these requirements up front before configuring the VC FlexFabric modules and defining the server profiles. However, changes can be easily made if required after you perform the initial configuration.

<table>
<thead>
<tr>
<th>Network name</th>
<th>VLAN ID</th>
<th>Host or VM</th>
<th>Bandwidth</th>
<th>NIC Teaming</th>
<th>IP Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Management</td>
<td>NA</td>
<td>Host</td>
<td>1 Gbps</td>
<td>Yes</td>
<td>192.168.1.x</td>
</tr>
<tr>
<td>Cluster</td>
<td>NA</td>
<td>Host</td>
<td>1 Gbps</td>
<td>Yes</td>
<td>200.1.1.x</td>
</tr>
<tr>
<td>Production</td>
<td>NA</td>
<td>Host</td>
<td>4 Gbps</td>
<td>Yes</td>
<td>NA (supports FE and BE VLANs for VMs)</td>
</tr>
<tr>
<td>FE</td>
<td>101</td>
<td>VM</td>
<td>NA</td>
<td>NA</td>
<td>172.1.1.x</td>
</tr>
<tr>
<td>BE</td>
<td>201</td>
<td>VM</td>
<td>NA</td>
<td>NA</td>
<td>10.1.1.x</td>
</tr>
</tbody>
</table>
To satisfy the storage requirements, each host requires redundant ports to the SAN fabric with 4 Gbps bandwidth. This means for each 10 Gb port on the server, 6 Gb will be allocated for networking and 4 Gb for FCoE bringing the total required bandwidth for each host to 20 Gbps when factoring in redundancy.

Setting up the Virtual Connect FlexFabric infrastructure

For this example scenario, we need two insert two VC FlexFabric modules in bays 3 and 4 of the blade enclosure. The modules must be in bays 3 and 4 rather than in bays 1 and 2 because the ProLiant BL460c G6 server blades do not have the appropriate FlexFabric adapters on the motherboard and require an additional Virtual Connect FlexFabric Adapter mezzanine card in mezzanine slot 1 to utilize the new FlexFabric capabilities. If you are using the ProLiant G7 blades which have integrated Virtual Connect FlexFabric NICs on the motherboard then the Virtual Connect FlexFabric modules can be inserted into bays 1 and 2.

After inserting the modules into the proper bays in the enclosure, the next step is to connect the modules to the external LAN and SAN networks. A diagram showing the rear side of the enclosure and applicable connections for this example is shown in Figure 2.

Figure 2 – Rear view of c7000 enclosure showing wiring of Ethernet and FC SAN connections to VC FlexFabric modules

As the figure above shows, there are two ports on each of the Virtual Connect FlexFabric modules (4 total uplinks) which provide multiple, redundant Ethernet network connections from the enclosure to the HP ProCurve network switches. There are also two FC SAN uplinks (1 per module) connecting to
two external SAN switches. This provides redundant connections from the server blades to the storage fabric and ultimately to the backend storage arrays (an EVA8400 not shown in the diagram). And remember that these connections are using the standard FC protocol from the Virtual Connect FlexFabric modules to the upstream switches, not FCoE. The FCoE protocol is only used for communications from the modules downstream to the server blades themselves, allowing you to eliminate the necessity for a dedicated FC HBA mezzanine card in the blade and associated FC interconnect module in the BladeSystem enclosure.

To handle multiple different communication protocols (FC and Ethernet) requires protocol specific transceivers. The FlexFabric modules are designed to provide a great deal of flexibility in the type of traffic that each port can support and different transceivers can be plugged into various ports to support different protocols\(^2\). In this example, 4 Gb FC SFP transceivers were installed in ports 1 on both modules and 1 Gb RJ-45 SFP modules were installed in ports 5 and 8. 1 Gb uplinks were utilized in this testing due to limited 10 Gb networking equipment in the lab environment. However, in production, 10 Gb SFP+ transceivers can provide 10 Gb uplink speeds to supported 10 Gb switches.

**Defining the Virtual Connect FlexFabric networks**

Now that the Virtual Connect FlexFabric modules have been inserted into the enclosure and the uplinks have been connected, the next step is to define the required Ethernet networks, SAN fabrics and server profiles for the Hyper-V host servers. This configuration can be performed using either the VC command line interface (CLI) or using the web-based Virtual Connect Manager (VCM). This white paper will highlight using the VCM to perform these steps.

For our example implementation, we’ve defined a requirement for two SAN fabrics and eight Ethernet networks to satisfy the LAN and SAN connectivity requirements. Figure 3 depicts a representation of the FlexFabric adapter and specific physical function configuration for this environment.

---

\(^2\) For specific configurations rules and guidelines for the SFP+ transceivers please see the associated Virtual Connect FlexFabric configuration and setup guides.
From this diagram you can see that three of the physical functions on each port are configured as FlexNICs to support the networking requirements. The other physical function is configured as a FlexHBA-FCoE connection. One point of interest is that the second physical function is the only physical function that supports both storage connectivity (FCoE or iSCSI) or network connectivity. So if you define an iSCSI or FCoE connection in your server profile, that network will always be mapped to the second physical function of the FlexFabric port.

To define the SAN fabrics within VCM, the “Define -> SAN Fabric” drop-down menu option was selected. Figure 4 depicts the summary information for the two SAN fabrics.
As the blue circle highlights, port 1 on each module was used as the uplink for the SAN fabrics. The port speed was also set to 4 Gb during the creation of the SAN fabrics to match the speed of the upstream SAN fabric in the test environment. The two SAN fabric networks defined in this example provide redundant paths to the backend EVA8400 SAN array for this configuration.

While defining the SAN fabrics is relatively straightforward, defining the Ethernet network requirements is a bit more complex. This is because we have requirements for both standard VC Ethernet networks as well as VC shared uplink sets (to support multiple tagged VLAN networks on the same NIC). Shared uplink sets provide the ability to support multiple tagged VLANs to a single server NIC (or multiple NICs) and minimize the number of uplinks required. In this example we'll define two shared uplink sets that will support the FE and BE Ethernet networks (VLANs 101 and 201 respectively) as were defined in Table 1 above.

In order to utilize the multiple networks capability of Virtual Connect, the “Map VLAN Tags” feature needs to be turned on under the Ethernet Settings/Advanced Settings tab within the Virtual Connect Manager as shown in Figure 5.

Figure 5 – Ethernet Settings, Advanced Settings tab for enabling mapped VLAN tag support

**Ethernet Settings**

<table>
<thead>
<tr>
<th>MAC Addresses</th>
<th>Port Monitoring</th>
<th>Advanced Settings</th>
</tr>
</thead>
</table>

**Server VLAN Tagging Support**

- Tunnel VLAN Tags
- Map VLAN Tags

- Force server connections to use the same VLAN mappings as shared uplink sets

**Multiple Networks Link Speed Settings**

When using mapped VLAN tags (multiple networks over a single link), these settings will be used for the overall link speed control.

- Set a Custom value for Preferred Link Connection Speed
- Set a Custom value for Maximum Link Connection Speed
The “Map VLAN Tags” feature provides the ability to use a Shared Uplink Set to present multiple networks to a single NIC. This way you can select multiple networks when assigning a network to the server profile. This then allows you to configure multiple VLANs on that server NIC. Further discussion on this step is covered in the “Defining the Virtual Connect FlexFabric server profiles” section below.

To define the shared uplink sets and Ethernet networks the “Define – Shared Uplink Set” and “Define – Ethernet Network” drop-down menu options were selected respectively from within VCM. To support the networking requirements for our example implementation, eight Ethernet networks were created in VCM. The figure below (Figure 6) provides a detailed summary of the eight networks.

Figure 6 – Ethernet networks summary

<table>
<thead>
<tr>
<th>Ethernet Networks</th>
<th>Shared Uplink Set (VLAN)</th>
<th>Port Status</th>
<th>Connected To</th>
<th>Uplink Port</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mgt-A</td>
<td>ProVLAN0-A (2M)</td>
<td>Linked-Active</td>
<td>1G6</td>
<td>SFP-RJ45</td>
</tr>
<tr>
<td>Mgt-B</td>
<td>ProVLAN0-B (2M)</td>
<td>Linked-Active</td>
<td>1G6</td>
<td>SFP-RJ45</td>
</tr>
<tr>
<td>Cluster-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cluster-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BE-1</td>
<td>ProVLAN-BE (1M)</td>
<td>Linked-Active</td>
<td>1G6</td>
<td>SFP-RJ45</td>
</tr>
<tr>
<td>BE-2</td>
<td>ProVLAN-BE (1M)</td>
<td>Linked-Active</td>
<td>1G6</td>
<td>SFP-RJ45</td>
</tr>
<tr>
<td>Mgt-A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mgt-B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

There are a couple of interesting aspects to highlight from this summary. First, the two management networks (Mgt-A and Mgt-B) are assigned to port 8 on each module respectively. This enables both networks to be active as the “Port Status” column indicates. An alternate approach would be to create a single management network that contained port 8 on both modules. However, that would leave one port as active and one in standby. While either option is viable, for this example we’ve chosen to define two separate networks (both networks staying as active) to take full advantage of the bandwidth for both ports. Redundancy and load balancing across both of these ports will be handled using NIC teaming at the host level.

Another thing to notice are the lack of Cluster-A and Cluster-B network properties. There are no external uplink ports on the FlexFabric modules associated with these networks. This is because the cluster heartbeat network was defined such that all traffic on this network will stay inside the enclosure.

And finally, the last four networks are associated with the production traffic. These include the BE-1, BE-2, FE-1, and FE-2 networks. To see the specific details of one of these networks you can right-click on the network name and select edit. Figure 7 shows the details and the associated network properties on the edit network page for the BE-1 network.
You can see that ID 201 is assigned as the VLAN ID for this network and that the network is part of the ProdVLANs-A shared uplink set. BE-2 is also assigned the same VLAN ID (201) but assigned to ProdVLANs-B. The front end networks, FE-1 and FE-2, use VLAN ID 101 and are also assigned to ProdVLANs-A and ProdVLANs-B respectively.

The uplink ports for all four of these networks are port 5 on both of the modules. This is the reason why the shared uplink sets have been defined so that the two networks can traverse the same uplink port with appropriate VLAN mappings. To get additional detail on the shared uplink you can right-click on the shared uplink set name and select edit. The details for the ProdVLANs-A shared uplink set are shown in Figure 8.
As you can see, the FE-1 and BE-1 Ethernet networks are both assigned to this shared uplink set and the external uplink is set to port 5 on the module in bay 3. The other shared uplink set, ProdVLANs-B is assigned to port 5 on the module in bay 4 and has the FE-2 and BE-2 networks associated to it so that both shared uplink sets carry traffic for VLANs 101 and 201.

One thing that is important to point out is that the upstream ProCurve switches (or whichever switches are the next hop from the Virtual Connect FlexFabric modules) need to be configured appropriately to handle the VLAN traffic. Please check with the specific switch manufacturer and model documentation for more details on configuring the upstream switches accordingly to handle the VLAN tagged networks.

**Defining the Virtual Connect FlexFabric server profiles**

The next step in the process is to define the server profiles that will be used for the two Hyper-V host servers. To create a server profile, select the “Define->Server profile” option from the drop-down menu. For this example, we need to add six Ethernet network connections and two FCoE connections to each server profile. This will utilize all eight available ports on the blade’s FlexFabric adapter. To help in understanding the steps that were taken to create the server profiles, we’ll examine the details of a complete server profile. Figure 9 depicts a server profile for one of the Hyper-V host servers (the second profile will be identical except for the profile name).
The first thing that stands out in this figure is that rather than the 6 network connections that were mentioned in the previous paragraph, there are actually 12 network connections that have been defined. This is due to the specific hardware configuration in this environment, and the way that the network connections are mapped to specific physical ports on the adapter cards. When using a FlexFabric mezzanine card in a server blade, port mappings are alternated between the integrated Flex-10 adapter on the motherboard of the BL460c G6 and the FlexFabric adapter in mezzanine slot 1. This behavior is easy to see by looking at the “Mappings” column in the figure above. The port mappings go back and forth between the LOM and MEZZ1 cards.

Since we aren’t using any additional modules in bays 1 and 2 (the bays the integrated adapters are mapped to), we need to make sure that we’re assigning the appropriate networks to ports that map to the mezzanine card rather than the integrated adapter ports. To do this is relatively straightforward; you just need to create networks that are left unassigned which map to the integrated adapter ports.

So for this scenario we create 12 Ethernet network connections and leave networks 1, 2 (not shown in the figure), 5, 6, 9, and 10 unassigned. This leaves us with the 6 networks that we need to define for our implementation (3, 4, 7, 8, 11 and 12).
There is no particular order in which the networks should be defined, but the first two network connections were assigned to the management networks, Mgt-A and Mgt-B. By default when a network is assigned the port speed is set to “Preferred” and if left in that state, the network bandwidth would be equally divided amongst all the networks that have been assigned to the server profile. However, there are specific bandwidth allocations that we’d like to set for each network based on the cluster implementation that we’re building. To configure a specific network speed, select “Custom” for the option in the “Port Speed” column. This will pop-up an additional window (Figure 10) that will allow you to dial-in to the exact speed you want to allocated for that network. In this case, we’ve set 1 Gb for the management networks.

Figure 10 – Setting the custom network bandwidth allocation

This same process is repeated for the cluster networks, Cluster-A and Cluster-B. The final two network connections, connections 11 and 12 in Figure 9, have a slightly different process for configuration. Recall in the previous section the main reason we configured shared uplink sets with mapped VLAN tagging was to allow multiple networks to be connected to the same port. To accomplish this during the server profile creation, select the “Multiple Networks” option in the drop-down menu in the “Network Name” column. This will bring up another configuration window which will allow you to either manually add network mappings or to use previously created shared uplink set mappings as shown in Figure 11.
For this network assignment (connection number 11 in the server profile), the multiple networks assignment was set to "ProdVLANs-A". For the last network connection (12), the VLAN mappings were set to the values in the “ProdVLANs-B” shared uplink set. For both of these networks, a port speed of 4 Gbps is set. This provides a total allocation of 6 Gbps for networking requirements, leaving 4 Gbps of bandwidth for each FCoE port.

The two FCoE connections for the server profile will map to the two SAN fabrics (EVA8400-A and EVA8400-B) that were defined. As was mentioned earlier, the two FCoE connections are assigned to the second physical function on the FlexFabric adapter ports since this is the only physical function that supports FCoE or iSCSI storage connections.

Finally we can see that the allocated bandwidth (highlighted by blue circles in Figure 9 above) shows that both ports of the FlexFabric Adapter are utilizing the total 10 Gbps of allotted bandwidth for the adapter. And the flexibility of the Virtual Connect FlexFabric design has allowed us to allocate that bandwidth to meet the exact needs of the implementation as we’ve defined above.

At this point, the server profiles can be assigned to the two server blades appropriately. In this example, we’re using two BL460c G6 blades as the Hyper-V host servers.
Installing and configuring the Hyper-V servers

After creating and assigning the server profiles, the next step is to install the OS and Hyper-V server role on the two server blades. There are a number of mechanisms that can be used to install and configure the OS and server appropriately to support Hyper-V. These range from a manual step-by-step process to something that can be completely automated. This white paper will provide some high level details of each of the major steps in the process, but in-depth details and implementation steps for each of these tasks are outside the scope of this paper.

In this example, a Windows 2008 R2 DVD ISO was attached to each server using the HP Integrated Lights-Out (iLO) virtual media capability and the base OS install was performed manually on each blade. Once the OS has been installed, the latest ProLiant Support Pack³ (PSP) should be installed to ensure that all driver versions are up to date. In this example, PSP 8.40 was installed.

Note, the one item that should not be installed from the PSP at this time is the HP Network Configuration Utility (NCU), which is HP’s NIC teaming software. This must be installed after the Hyper-V server role has been added. This is covered in more detail below.

After the OS and PSP have been installed, there is one final step that must be performed before the Hyper-V server role can be added. In order for the server to support the Hyper-V server role, the following server BIOS modifications are required:

1. Enable the “No-Execute Memory Protection” advanced option.
2. Enable the “Intel® Virtualization Technology” advanced option.

These options are specific for an Intel chipset platform (BL460c G6). For further information on what the BIOS modifications are for an AMD™ platform or for general information on implementing Hyper-V on HP ProLiant servers see the “Implementing Microsoft Windows Server 2008 R2 Hyper-V and Microsoft Hyper-V Server 2008 R2 on HP ProLiant servers” document.

After modifying the server BIOS to support virtualization, the next step is to install the Hyper-V server role. In Windows Server 2008, Hyper-V is a server role that can be easily installed by running the “Add Roles” wizard (from within Server Manager) and selecting the Hyper-V role option. This step is also covered in additional detail in the aforementioned white paper on implementing Hyper-V on ProLiant servers.

Configuring NIC teaming and VLANs with Hyper-V

Once the Hyper-V role has been installed, the HP NCU software can be installed from the PSP or by downloading the installation package separately from HP. In this testing, NCU version 9.90.0.17 was installed. This section will cover the steps that were taken to setup NIC teaming with Hyper-V for this deployment. For more general setup information and details please read the “Using HP ProLiant Network Teaming Software with Microsoft Windows Server 2008 Hyper-V or with Microsoft Windows Server 2008 R2 Hyper-V” document.

For this implementation there are six NICs that are enabled, three NICs per VC FlexFabric module. The NCU utility will allow us to setup three NIC teams with NICs on separate VC FlexFabric modules to prevent the loss of a single module causing a network outage.

³ The latest PSP for each server OS can be downloaded from the following page, http://h20000.www2.hp.com/bizsupport/TechSupport/DriverDownload.jsp?lang=en&cc=us&prodNamId=3716247&taskId=135&prodTypId=18964&prodSeriesId=3716246&lang=en&cc=us.
Figure 12 shows the NCU utility and the six available networks. The best way to make sure that you are teaming the correct NICs together is to validate the MAC addresses from the server profile (figure 9 above) with the MAC addresses of the NICs listed in the NCU. In this case, the actual networks were renamed beforehand to ensure that the correct teams were made within NCU.

Figure 12 – NCU configuration screen before forming NIC teams
Once you’ve determined the correct NICs that will be teamed together, simply highlight the two NICs and select the “Team” option. Repeat this step to form the remaining two teams. For the management and cluster teams, the only modification made was to the name of the teamed network for easy identification. These were called “Mgt Team” and “Cluster” respectively. This can be accomplished by highlighting the newly formed NIC team and selecting the “Properties” option. The third NIC team requires an additional setup step to enable the team to handle multiple VLANs. Highlight the network team, called “prod”, and select the “VLAN(802.1Q)” option as shown in Figure 13.

Figure 13 – NCU configuration screen after NIC teams formed
Selecting this option brings you directly to the VLAN tab of the network team properties page. Here you need to add the VLANs and associated VLAN IDs that have been defined in VCM for the FlexFabric modules. These include the FE and BE VLANs with IDs 101 and 201 respectively as shown in Figure 14.

Figure 14 – VLAN tab in production NIC team properties page
Once the VLAN information has been entered press OK and then OK a second time to close and save the changes for the NCU. At this point, the three network teams have been created. These can be viewed by opening up the Windows network properties page. You can also see the associated bandwidths allocated to each team. If the MAC addresses were matched appropriately, the bandwidth values will match the values set in the server profile. Figure 15 shows the bandwidth allocations for the newly created teams. Note that there are actually four network property pages, rather than three. This is because each VLAN on the “prod” network team has a separate network attribute entry.

Figure 15 – NIC teams and bandwidth
Now that the networks have been defined and the teams created, the final configuration step to perform prior to configuring the Hyper-V virtual networks is to set the IP addresses on the host networks. The following table, Table 2, provides the IP address information used in this example.

**Table 2. NIC team information for host networks**

<table>
<thead>
<tr>
<th>Network name</th>
<th>Server name</th>
<th>Static/ DHCP</th>
<th>IP Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster team</td>
<td>HV-FlexFabric-1</td>
<td>Static</td>
<td>200.1.1.1</td>
</tr>
<tr>
<td>Mgt team</td>
<td>HV-FlexFabric-1</td>
<td>DHCP</td>
<td>192.168.1.19</td>
</tr>
<tr>
<td>Cluster team</td>
<td>HV-FlexFabric-2</td>
<td>Static</td>
<td>200.1.1.2</td>
</tr>
<tr>
<td>Mgt team</td>
<td>HV-FlexFabric-2</td>
<td>DHCP</td>
<td>192.168.1.25</td>
</tr>
</tbody>
</table>

The remaining two VLANs, FE-ProdVLANs and BE-ProdVLANs, will be used to create the Hyper-V virtual networks for guest VM network access as described in the next section. Depending on the applications that will run on the VMs and your specific network design, either static or DHCP IP addresses can be used for these networks.
Creating the Hyper-V virtual networks

Once you have defined the NIC teams (and setup the appropriate VLANs on the NIC team that will handle VM guest networking) you can create the necessary Hyper-V virtual networks. To create a Hyper-V virtual network, open up the Virtual Network Manager from the Hyper-V Manager interface.

In this example there are two external virtual networks that need to be defined to provide isolation for front-end and back-end networking. These will be called FE and BE respectively. These two virtual networks will map to the VLANs that were created on the “prod” NIC team above. To create the FE network, create a new external virtual network and select the “FE:prod” connection from the external drop-down menu box in the “Connection Type” section as shown in Figure 16 below.

Figure 16 – Hyper-V Virtual Network Manager window

This same process is repeated for the “BE” network, with the exception that the connection utilizes the “BE:prod” VLAN connection.
At this juncture the networks have been defined but the specific VLAN ID information for these networks has not been set. This step occurs once the guest VMs have been defined and is covered in the “Implementing Failover Clustering and Live Migration” section below.

**Configuring the storage resources**

Before adding the failover cluster role and configuring the Hyper-V live migration cluster, the final step that we need to perform is to make sure the storage has been presented and configured correctly. These steps will differ depending on which storage subsystem you are using. The backend storage array for this environment is the EVA8400 array and this section will discuss high level details of the steps necessary to present and configure the storage for this array. For more specific step by step instructions on how to setup and configure the EVA array (or any other array) please read the product specific installation and configuration documents.

The first step in getting the storage presented to the two Hyper-V host servers is to add the two hosts to the EVA. To add a host to an EVA array, HP StorageWorks Command View (CV) EVA software can be utilized. The key attributes needed for the host definition are the HBA world wide port names (WWPNs). Like the MAC addresses, these values can be ascertained by looking at the details for the two FCoE ports in the server profile (figure 9). With the correct WWPNs the two hosts can be created on the array. These values are also used to modify the SAN switch zoning information to ensure that the LUNs are visible from the ports on both servers.

After the two hosts have been added to the array, the next step is to configure the LUNs on the array. This scenario is a simple example demonstrating a single VM with live migration support. Thus we’re only using three LUNs for this example as detailed in the following table (Table 3).

<table>
<thead>
<tr>
<th>LUN name</th>
<th>Size</th>
<th>VRAID level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quorum</td>
<td>5 GB</td>
<td>VRAID1</td>
</tr>
<tr>
<td>VM guest OS</td>
<td>50 GB</td>
<td>VRAID1</td>
</tr>
<tr>
<td>VM guest Data</td>
<td>50 GB</td>
<td>VRAID1</td>
</tr>
</tbody>
</table>

For more complex, real-world scenarios, the actual LUN configurations will be based on the specific application requirements and the number of VMs in the cluster. However, this basic process can be easily extrapolated out for the more complex configurations.

Using CV EVA, these three LUNs are created and then presented to both of the hosts. This enables both servers to see all the LUNs which will be required when creating the Windows cluster. However, at this point, there are multiple paths to each LUN that are visible in Disk Manager. This is because we haven’t installed multipathing software. Once the multipathing software has been installed, you will only see a single path to each LUN in Disk Manager. The EVA array has specific multipathing software and in this environment version 4.00.00 of the HP MPIO Full Featured DSM for EVA8400 was installed. Note, the EVA specific MPIO software leverages the base Microsoft MPIO feature in the Windows 2008 R2 OS.

Now that the MPIO software is handling multipathing, the final step in preparing the LUNs for the cluster installation is to format and partition them. Remember that since this is a cluster configuration, the LUNs are visible to both host servers. Thus, we only need to format and partition the LUNs from a single host. And once the failover clustering feature is enabled, the underlying cluster technology will appropriately handle multiple hosts accessing the same LUNs.
Implementing Failover Clustering and Live Migration

At this stage the underlying network and storage infrastructure has been setup and configured to support the two-node Hyper-V cluster. As the general setup and installation of Hyper-V and Failover Clustering is covered in great detail by Microsoft in various TechNet articles (See the “For More Information” section for links), this white paper will not go into step by step details for this process. Rather the remainder of this paper will focus on the high-level details necessary to build out the Hyper-V cluster and support live migration. We’ll also examine the final steps necessary during VM creation to support the VLAN networking requirements for this environment.

To implement the failover cluster with live migration support the following steps were followed:

1. Install failover cluster feature on both servers.
2. Validate the cluster configuration.
3. Create the cluster.
4. Create a VM.
5. Make the VM highly available.
6. Configure the cluster network for live migration.
7. Install the guest VM OS and application(s).
8. Test and validate failover scenarios.

The first step is to install the failover cluster feature on both servers. This feature can be installed by running the “Add Features Wizard” from within Server Manager. Once the failover cluster feature has been installed, the next step is to run the cluster validation wizard. The validation check will confirm that the server, storage, and network configuration meet the specific requirements necessary to install a failover cluster.

To run the validation wizard, open the Failover Cluster Management interface and select the “Validate a Configuration” option from the “Management” section in the center of the console. If everything checks out then you can continue on to the next step of creating the cluster.

Creating the cluster is another wizard based operation that is executed from the Failover Cluster Management interface by selecting the “Create a cluster” option. For this example, a cluster called “Flexcluster.flexfabric.com” was created with the two server nodes and a “Node and Disk Majority” quorum configuration.

With the cluster created you can now go ahead and create the VMs for your environment. In this example we’ll create a single VM and then enable that VM for HA within the cluster to test and validate both planned and unplanned failover scenarios.

To create the VM, you can use the “New Virtual Machine Wizard” option from the Hyper-V Manager interface. Note, it is very important that when running this wizard, that the storage you select to host the VM’s VHD file is located on one of the shared storage LUNs, and not the host server’s local storage. Otherwise you won’t be able to enable HA for the VM and add it to the cluster. For this configuration, the “VM guest OS” LUN was selected. Also during this initial wizard you can leave the networking disconnected for the time being. The networking adapters can be added directly to the VM after the wizard completes as modifications to support the VLAN requirements must be made.
After the wizard completes, highlight the newly created VM and open the settings window. From here the two network adapters for the VM can be added. One of the network adapters will be connected to the “FE” network and the second network adapter connected to the “BE” network. To configure the appropriate VLAN information, check the “Enable virtual LAN identification” option and set the VLAN IDs to 101 and 201 respectively for the “FE” and “BE” networks. Figure 17 shows this configuration step for the “FE” network.

![Figure 17 – Modifying the network properties for the WinVM test VM](image)

After configuring the two network adapters there is one final setting to modify before making the VM part of the cluster. You want to ensure that the cluster service will manage the state of the VM once the VM is part of the cluster. To ensure this will happen correctly you should set the “Automatic Start Action” setting for the VM to none (as pointed out by the blue circle in the lower left hand corner of the figure above).

Now you can save these changes by pressing the OK button on the bottom of the settings page. Note, do not start up the VM at this point as the VM must be powered off to make it highly available.
To make the VM highly available we need to switch back to the Failover Cluster Manager interface. Within this snap-in, click on “Services and Applications” and select the “Configure a Service or Application” option from within the action pane to open up the high availability wizard. Select “Virtual Machine” as the option for the service that you want to make highly available and then select the specific VM. In this case, the “WinVM” VM is the VM that we are making highly available.

Note, once the VM is part of a cluster, any modifications to the settings of the VM should be made from within the Failover Cluster Manager rather than from within Hyper-V Manager. To change the settings of the VM, right-click on the VM resource and select the “Settings…” option as shown in Figure 18.

Figure 18 – WinVM service action options
By default, the networks in the cluster are configured for live migration. However, you do want to make sure that the preference order is set to use the network that is dedicated to live migration. In this example, we have two networks, the management and cluster networks. To make sure that the cluster network is set as the preferred network for live migration, select the “Properties” option as shown in the figure above. This will open up the properties window, click the “Network for live migration” tab and validate that the order of the networks is correct.

At this point, the VM can be powered up (from within the Failover Cluster Manager) and the guest OS, along with any applicable software, installed.

Once the guest OS and software has been installed and configured, the final step in the process is to test and validate the failure scenarios and HA capabilities. To run a planned live migration or quick migration scenario, right-click on the VM within the Failover Cluster Manager and select either the “Live migrate virtual machine to another node” or “Quick migrate virtual machine(s) to another node” option from the available action menu as shown in Figure 19.

Figure 19 – Initiating live migration planned failover test

For this configuration, both live migration and quick migration were tested and completed successfully. An unplanned downtime event was also successfully simulated by powering off the server that currently owned the VM. The VM failed over to the second node in the cluster with minimal downtime.

A quick note about Cluster Shared Volumes

With Windows Server 2008 R2, Microsoft introduced the concept of Cluster Shared Volumes (CSVs) to allow multiple nodes in a cluster to read and write to a volume at the same time. Traditionally, only a single node in a cluster could own and access a volume at one time. This effectively meant that each VM required its own volume. Otherwise, multiple VMs on the same LUN would have to be failed over from node to node together. This also meant that for larger deployments with a lot of VMs in a cluster, the number of LUNs to manage can quickly become unruly. With CSVs, these restrictions are removed.
Now multiple VMs can utilize a single LUN (and be moved/failed to another node in the cluster) without impacting the other VMs on that LUN. This subsequently means that fewer LUNs need to be managed.

This example did not utilize CSVs for the cluster configuration as the example focused on a single VM and the paper was more focused on the underlying VC FlexFabric configuration steps. However, implementing CSVs in this process is fairly straightforward and only requires a few additional steps that were not covered. For more detail on this and step by step guides for configuring CSVs with Live Migration please see the “Hyper-V: Using Live Migration with Cluster Shared Volumes in Windows Server 2008 R2” TechNet article.

Summary

As more and more customers deploy virtualization in their data centers, and rely on Windows Failover Clustering to provide high availability for their VMs, it is critical to understand how to setup and configure the underlying infrastructure appropriately. This white paper provides example implementation details on how to build a Windows Server 2008 R2 Failover Cluster (with Live Migration support) using the HP Virtual Connect FlexFabric architecture.

HP Virtual Connect FlexFabric provides clients an open, standards-based choice to network convergence in an industry that previously had limited options. HP Virtual Connect FlexFabric converges Ethernet, Fibre Channel, and iSCSI traffic into a single common fabric module, helping to not only simplify and reduce the costs of the infrastructure fabric, but also providing a key technology necessary to implement the FlexFabric layer of the HP Converged Infrastructure. This includes the ability to allocate and fine-tune network bandwidth for each connection and dynamically make changes to those connections on the fly.
For more information

The following links provide more information on Microsoft Hyper-V:

HP and Microsoft virtualization solution page – www.hp.com/go/hyper-v

HP ActiveAnswers for Microsoft Hyper-V solutions page –
http://www.hp.com/solutions/activeanswers/Hyper-v

For more information on HP Virtual Connect – www.hp.com/go/virtualconnect

For more information on HP Converged Infrastructure – www.hp.com/go/convergedinfrastructure

For more general information on setting up and installing Hyper-V and Failover Clustering, see the

To help us improve our documents, please provide feedback at