OUTDATED

Top 10 SAN Tips and Lessons Learned From Virtualization

Written by: **Brien M. Posey**

Brien Posey, a freelance technical writer, received Microsoft's MVP award six times for his Exchange Server, Windows Server, IIS, and File Systems Storage work.

Server virtualization is a science in its own rite, but the virtualization infrastructure takes on an additional layer of complexity when SANs are brought into the picture. The key to ensuring the optimal functionality of such a complex environment is to adhere to the best practices for virtualization / SAN connectivity. Of course many of the best practices vary depending on which vendor's products are being used, but there are some best practices that are more or less universal. This whitepaper outlines ten such best practices for using SANs in a virtual datacenter.

1 Verify That Your SAN Hardware is Certified for Your Virtualizatio Solution

One of the first considerations that you must take into account is compatibility. It goes without saying that if you want to connect your virtualization infrastructure to a SAN then the SAN must be compatible with your virtualization platform. However, in many cases it is important to look beyond mere compatibility and instead verify that the hardware is actually certified for use with your virtualization solution.
For example, organizations that use Hyper-V as a virtualization solution often depend on the use of failover clusters to provide fault tolerance and live migration capabilities. However, Microsoft's support policy clearly states that failover clustering solutions are only supported if all of the hardware components are marked as certified for Windows Server 2008 or Windows Server 2008 R2. Additionally, the Failover Cluster Manager in Windows Server contains a Wizard that is designed to validate your failover clustering configuration. Microsoft will not support a failover clustering configuration unless it has passed all of the validation tests.

The concept of using certified components is not unique to Microsoft virtualization solutions. VMware also publishes their own Certified Compatibility Guides (http://www.vmware.com/resources/guides.html). These guides list which systems, storage / SAN, and backup solutions have been certified to be compatible with VMware.

One area in which VMware differs from Microsoft is that while Microsoft requires storage hardware to be certified for use with Windows Server in order to be supported, VMware publishes a list of community supported hardware and software (http://communities.vmware.com/cshwsw.jspa). The items appearing on this list are not officially certified by VMware, but have been tested either by the hardware vendors or by VMware's customers.

2 Be Consistent With Your Hardware

While there is no denying the importance of making sure that your SAN hardware is compatible with (and preferably certified for use with) your virtualization solution that alone might not be enough. You need to be consistent with your hardware.

Suppose for instance that that you have a series of servers that are being used as nodes in a virtualization host server cluster, and that those nodes connect to your SAN through Fibre Channel. In such a case, each server's Fibre Channel adapter should be identical.

Having identical Fibre Channel adapters means more than simply installing the same make and model in each server. It also means verifying that each Fibre Channel adapter is using the same firmware version. Likewise, you should make sure that each adapter uses identical communication settings (speed, duplex mode, flow control, media type, etc.).

3 Use Lots of Spindles

When you are allocating storage to a virtualization host server (or to a host server cluster) it is important to configure the storage to deliver performance that is scaled to match the total workload requirements for the virtual machines that are running on the host. Although there are numerous factors that affect storage performance, you should pay particular attention to the number of spindles that you use in a given situation.

Generally speaking you will receive the best performance by spreading the workload across as many spindles as possible (within reason). Given the importance of virtual machine hosting, it is also important to make your RAID array fault tolerant. RAID 5 will get the job done, but RAID 10 yields much better performance.

Another consideration for those seeking to achieve high performance is to use Solid State Drives (SSDs). SSDs have no moving parts and are not subject to the mechanical limitations of HDDs. As such, they generally offer twenty to thirty percent better performance than regular hard drives. One thing to consider however, is that SSDs have a higher cost per gigabyte and SSDs have a much lower overall capacity.

4 Consider Using Multipath I/O Software

One of the most important considerations for any storage fabric is reliability, and one of the best ways to achieve reliability is through redundancy. Although using a failover clustering solution for your virtualization platform is a good start, clustering alone may not offer adequate protection against hardware failures.

A failover cluster will insulate you against a server failure, but may not protect you against the loss of SAN connectivity. Depending on what clustering solution you are using for example, it is possible that a cluster node could lose SAN connectivity, and yet remain online. In this type of situation the virtual machines running on the node on which the failure occurred would fail, but a failover to another cluster node may never occur because the cluster itself is still able to communicate with the ailing node and assumes that the node is healthy.
The best way to address this problem is through the use of multipath I/O, which comes standard with

modern hypervisors and operating environments. Multipath I/O will allow a server to continue to communicate with the SAN even in the event of a host bus adapter failure.
There are a few important considerations that you should take into account when planning a multipath I/O solution. First, base your multipath I/O solution around the use of two separate host bus adapters. While it is true that you can achieve multipath I/O through a dual port host bus adapter, doing so will only protect you against a cable failure. The adapter itself remains a potential single point of failure. As such, you can only achieve true redundancy by using multiple host bus adapters.

Another consideration is that multipath I/O solutions can be very version sensitive. It is extremely important that your host bus adapters use consistent firmware versions and that you work closely with your hardware vendor to choose hardware that will work in a multipath environment.

Finally, some operating systems can be very picky about multipath I/O software. If you are using Windows Server 2008 R2 for example, then your multipath I/O solution must be based on Microsoft's Multipath I/O component (which is included with the operating system) and a Device Specific Module. You can use a Device Specific Module that was included with the operating system, or you can use one that was provided by a hardware vendor.

5 Follow the Recommended Best Practices for iSCSI

Even though Fibre Channel delivers the best performance for SAN connectivity, many organizations opt to use iSCSI instead because it is far less expensive to implement. iSCSI tends to be very flexible, and this flexibility sometimes leads to less than ideal configurations. As such, there are a few best practices that you should keep in mind with regard to using iSCSI connectivity.

First, iSCSI communications should take place over a dedicated network. This is important not only for performance reasons, but also for security. You should never under any circumstances use the same NIC for both iSCSI traffic and for general network communications.

Second, avoid confusion by using a different subnet for iSCSI communications than what you use for normal network communications. Yes, you can technically get away with using a single subnet for all types of network traffic so long as you use unique IP addresses, but remember that the goal is to limit iSCSI traffic to a dedicated network (dedicated adapters, switches, etc.). Using the same subnet for two separate networks can potentially cause routing problems, so it is best to create a dedicated subnet for iSCSI traffic.

A third consideration has to do with the performance limits that you might encounter. When a network adapter is unable to deliver sufficient bandwidth then a common solution to the problem is to use NIC teaming (which is sometimes referred to as bonding). NIC teaming is a technique through which network communications are distributed across two or more NICs as a way of overcoming the performance limitations of a single NIC. Although NIC teaming works well for general network communications, some flavors of iSCSI do not work with teamed NICs.

6 Don't Fall Into the Bandwidth Trap

When developing a storage architecture, most people have the same basic goal in mind – to get the best possible performance at the lowest possible price. With this goal in mind, consider the choice between 8 gigabit Fibre Channel and 10 gigabit iSCSI connectivity.

On the surface this choice seems obvious. iSCSI is much less expensive than Fibre Channel and 10 gigabit iSCSI would seem to be faster than 8 gigabit Fibre Channel. However, all is not what it might at first seem.

The problem with assuming that 10 gigabit iSCSI is the faster of the two connectivity methods is an adapter's throughput rating does not reflect the amount of overhead involved in the data transfer process.

Every communication protocol has overhead associated with it. This is true for both iSCSI and Fibre Channel. This means that the amount of data that a protocol can transfer over a given medium within a specific length of time is largely affected by the efficiency of the protocol.

Fibre Channel is a far simpler protocol than iSCSI. The iSCSI protocol has a lot of overhead because iSCSI commands are encapsulated in TCP/IP packets. Network adapters with TCP/IP offload engines can help to improve iSCSI's efficiency, but not enough to put it on par with Fibre Channel. Furthermore, iSCSI robs the server of CPU cycles and does not scale as well as Fibre Channel in multi-port configurations.

Benchmark testing of 10 gigabit iSCSI and 8 gigabit Fibre Channel has confirmed that even with its 2 gigabit advantage, 10 gigabit iSCSI simply does not perform as well as 8 gigabit Fibre Channel(Link).

7 Jumbo Frames

Another lesson for SAN users in virtual datacenters is that if you are using iSCSI connectivity you might be able to achieve better overall storage performance by enabling jumbo frames.

The concept of frames was first introduced out of necessity during the infancy of the Internet. The distributed (and sometimes unreliable) nature of the Internet meant that data sometimes had to be re-transmitted. TCP/IP was designed to break the data into a series of chunks (frames). That way, if some of the data was lost or corrupted during transit then the host could retransmit the frames that were lost rather than having to retransmit everything.

In a SAN environment, storage is attached via reliable, high speed connectivity. As such, frame retransmission becomes a moot point. While you can't get around the need for frames, you can increase the size of the frames by enabling jumbo frames. Doing so reduces overhead and results in more efficient use of your bandwidth.

VMware and Hyper-V both support the use of jumbo frames. Jumbo frame support was first introduced by VMware in vSphere 4.0. Microsoft followed suit by enabling jumbo frames in Hyper-V R2 (although the Integration Services are required to use jumbo frames from within a virtual machine).

The method for enabling jumbo frames varies from one product to the next. Furthermore, all of your networking components must support jumbo frames in order to facilitate end to end use.

You can confirm that jumbo frames are enabled and that they are being used end to end by using a simple ping test. The syntax is as follows:

Ping –n 1 –l 8000 –f (host name or IP address)

The N switch tells Ping how many echo requests you want to send. The L switch tells Ping how many packets to send, and the F switch tells Ping not to fragment the packets. The switches are case sensitive.

If jumbo frames are not enabled end to end then Ping will return an error stating that packets need to be fragmented.

8 Configure NIC Ports Appropriately

Host servers contain a limited number of ports in which NICs can be installed. As such, it is important to allocate your NIC ports in a way that maximizes performance and that prevents network bottlenecks.

One thing that you should do in any host server is to use multi-port NIC cards. The goal is to maximize the number of NICs that the server can accommodate.

If you are connecting the host to a Fibre Channel SAN then you will want to reserve at least two of the server's expansion ports for host bus adapters. Using multiple adapters in conjunction with multipath I/O provides enhanced performance and insulation against the failure of a host bus adapter.

If the SAN is iSCSI based then you will want to dedicate at least two NIC ports to iSCSI traffic. Although some flavors of iSCSI do not support NIC teaming, you should be able to achieve multipath I/O for iSCSI without teaming the NICs.

Regardless of which type of SAN storage connectivity you are using, you will need to reserve a NIC port on each host server for heartbeat traffic. Clustered hosts use heartbeats over a dedicated network segment to exchange health status information with other hosts in the cluster.

It is also a good idea to reserve a NIC port for server management. Although this is not an absolute requirement, having a dedicated management port allows you to manage the server without depriving any of the virtual machines of network bandwidth. In clustered environments it is common to create a dedicated backend management network.

Finally, you will need to have some NICs dedicated to servicing the virtual machines. The way in which these NICs should be configured varies depending upon the needs of the guest machines. If you have some virtual machines that run extremely bandwidth intensive applications then you can dedicate physical NIC ports to those virtual machines. However, it is usually preferable to team the NICs that will be servicing virtual machines, because dedicating NICs to specific virtual machines can make moving virtual machines from one cluster node to another difficult.

Regardless of how you choose to configure your NIC ports, it is important to take redundancy into account. For example, if you decide to use multiple ports for iSCSI connections then it is better to use ports on separate physical NICs than to use ports on a single, multi-port NIC. That way the server won't be cut off from the storage if a NIC fails.

9 Keep Things Organized

As you saw in the previous section, a properly configured host cluster makes use of multiple NIC ports and several dedicated network segments. One of the keys to keeping things running smoothly is to adopt a naming convention for NIC ports, and physical and virtual switches. Your naming conventions should be reflected within the hypervisor if possible (by doing things such as renaming your virtual switches) and in your network documentation. If you ever need to replace a dead host server you don't want to have to guess as to what each of the network cables was used for or how the server's NIC ports should be configured.

10 Avoid Using Removable Media

Another best practice that is often overlooked is that removable media should be completely avoided on clustered host servers. To see why this is the case, imagine that you use a DVD to upgrade an application that is installed on a virtual server. When the process completes, you eject the installation media, but you forget to uncapture the DVD drive. If a failure were to occur later on then the failover process probably will not work. Different hypervisors behave differently, but typically you cannot live migrate a virtual machine that has a physical DVD attached.

A better solution is to configure a portion of your SAN to store ISO files for each DVD that you use. When it comes time to perform an action that would normally require a DVD, simply use your SAN based ISO library instead. NIC port on each host server for heartbeat traffic. Clustered hosts use heartbeats over a dedicated network segment to exchange health status information with other hosts in the cluster.

It is also a good idea to reserve a NIC port for server management. Although this is not an absolute requirement, having a dedicated management port allows you to manage the server without depriving any of the virtual machines of network bandwidth. In clustered environments it is common to create a dedicated backend management network.

Finally, you will need to have some NICs dedicated to servicing the virtual machines. The way in which these NICs should be configured varies depending upon the needs of the guest machines. If you have some virtual machines that run extremely bandwidth intensive applications then you can dedicate physical NIC ports to those virtual machines. However, it is usually preferable to team the NICs that will be servicing virtual machines, because dedicating NICs to specific virtual machines can make moving virtual machines from one cluster node to another difficult.

Regardless of how you choose to configure your NIC ports, it is important to take redundancy into account. For example, if you decide to use multiple ports for iSCSI connections then it is better to use ports on separate physical NICs than to use ports on a single, multi-port NIC. That way the server won't be cut off from the storage if a NIC fails.

ABOUT STARWIND
Since 2003, StarWind has been a pioneer in the iSCSI storage industry and has been the solution of choice for thousands of global customers in over 50 countries, from SMBs, to governments, and to Fortune 1000 clients.