Storage Protocol Comparison White Paper

TECHNICAL MARKETING DOCUMENTATION
V 1.0/UPDATED APRIL 2012

vmware®

Table of Contents

Introduction

VMware frequently is asked for guidance regarding the best storage protocol to use with VMware vSphere®. vSphere supports many storage protocols, with no preference given to any one over another. However, many customers still want to know how these protocols stack up against each other and to understand their respective pros and cons.

This white paper looks at common storage protocols from a vSphere perspective. It is not intended to delve into performance comparisons, for the following two reasons:

- The Performance Engineering team at VMware already produces excellent storage performance white papers.
- Storage protocol performance can vary greatly, depending on the storage array vendor. It therefore does not make sense to compare iSCSI and NFS from one vendor, because another vendor might implement one of those protocols far better.

If you are interested in viewing performance comparisons of storage protocols, the "Conclusion" section of this paper includes links to several such documents.

Storage Protocol Comparison Table

	iSCSI	NFS	FIBRE CHANNEL	FCoE
Description	iSCSI presents block devices to a VMware® ESXi™ host. Rather than accessing blocks from a local disk, I/O operations are carried out over a network using a block access protocol. In the case of iSCSI, remote blocks are accessed by encapsulating SCSI commands and data into TCP/IP packets. Support for iSCSI was introduced in VMware® ESX® 3.0 in 2006.	NFS presents file devices over a network to an ESXi host for mounting. The NFS server/array makes its local file systems available to ESXi hosts. ESXi hosts access the metadata and files on the NFS array/server, using an RPC-based protocol. VMware currently implements NFS version 3 over TCP/IP. Support for NFS was introduced in ESX 3.0 in 2006	Fibre Channel (FC) presents block devices similar to iSCSI. Again, I/O operations are carried out over a network, using a block access protocol. In FC, remote blocks are accessed by encapsulating SCSI commands and data into FC frames. FC is commonly deployed in the majority of mission-critical environments. It has been the only one of these four protocols supported on ESX since the beginning.	Fibre Channel over Ethernet (FCoE) also presents block devices, with I/O operations carried out over a network using a block access protocol. In this protocol, SCSI commands and data are encapsulated into Ethernet frames. FCoE has many of the same characteristics as FC, except that the transport is Ethernet. VMware introduced support for hardware FCoE in vSphere 4.x and software FCoE in VMware vSphere® 5.0 in 2011.
Implementation Options	• Network adapter with iSCSI capabilities, using software iSCSI initiator and accessed using a VMkernel (vmknic) port. or: • Dependent hardware iSCSI initiator. or: • Independent hardware iSCSI initiator.	Standard network adapter, accessed using a VMkernel port (vmknic).	Requires a dedicated host bus adapter (HBA) (typically two, for redundancy and multipathing).	• Hardware converged network adapter (CNA). or: • Network adapter with FCoE capabilities, using software FCoE initiator.

	iSCSI	NFS	FIBRE CHANNEL	FCoE
Performance Considerations	iSCSI can run over a 1Gb or a 10Gb TCP/IP network. Multiple connections can be multiplexed into a single session, established between the initiator and target. VMware supports jumbo frames for iSCSI traffic, which can improve performance. Jumbo frames send payloads larger than 1,500. Support for jumbo frames with IP storage was introduced in ESX 4, but not on all initiators. (See VMware knowledge base articles 1007654 and 1009473.) iSCSI can introduce overhead on a host's CPU (encapsulating SCSI data into TCP/IP packets).	NFS can run over 1Gb or 10Gb TCP/IP networks. NFS also supports UDP, but the VMware implementation requires TCP. VMware supports jumbo frames for NFS traffic, which can improve performance in certain situations. Support for jumbo frames with IP storage was introduced in ESX 4. NFS can introduce overhead on a host's CPU (encapsulating file I/O into TCP/IP packets).	FC can run on 1Gb/2Gb/4Gb/8Gb and 16Gb HBAs, but 16Gb HBAs must be throttled to run at 8Gb in vSphere 5.0. Buffer-to-buffer credits and end-to-end credits throttle throughput to ensure a lossless network. This protocol typically affects a host's CPU the least, because HBAs (required for FC) handle most of the processing (encapsulation of SCSI data into FC frames).	This protocol requires 10Gb Ethernet. With FCoE, there is no IP encapsulation of the data as there is with NFS and iSCSI. This reduces some of the overhead/latency. FCoE is SCSI over Ethernet, not IP. This protocol also requires jumbo frames, because FC payloads are 2.2K in size and cannot be fragmented.
Load Balancing	VMware Pluggable Storage Architecture (PSA) provides a round-robin (RR) path selection policy (PSP) that distributes load across multiple paths to an iSCSI target. Better distribution of load with PSP_RR is achieved when multiple LUNs are accessed concurrently.	There is no load balancing per se on the current implementation of NFS, because there is only a single session. Aggregate bandwidth can be configured by creating multiple paths to the NAS array, accessing some datastores via one path and other datastores via another.	VMware Pluggable Storage Architecture (PSA) provides a round-robin (RR) path selection policy (PSP) that distributes load across multiple paths to an FC target. Better distribution of load with PSP_RR is achieved when multiple LUNs are accessed concurrently.	VMware Pluggable Storage Architecture (PSA) provides a round-robin (RR) path selection policy (PSP) that distributes load across multiple paths to an FCoE target. Better distribution of load with PSP_RR is achieved when multiple LUNs are accessed concurrently.

	iSCSI	NFS	FIBRE CHANNEL	FCoE
Resilience	VMware PSA implements failover via its Storage Array Type Plug-in (SATP) for all supported iSCSI arrays. The preferred method to do this for software iSCSI is with iSCSI binding implemented, but it can be achieved by adding multiple targets on different subnets mapped to the iSCSI initiator.	Network adapter teaming can be configured so that if one interface fails, another can take its place. However, this relies on a network failure and might not be able to handle error conditions occurring on the NFS array/server side.	VMware PSA implements failover via its Storage Array Type Plug-in (SATP) for all supported FC arrays.	VMware PSA implements failover via its Storage Array Type Plug-in (SATP) for all supported FCoE arrays.
Error Checking	iSCSI uses TCP, which resends dropped packets.	NFS uses TCP, which resends dropped packets.	FC is implemented as a lossless network. This is achieved by throttling throughput at times of congestion, using B2B and E2E credits.	FCoE requires a lossless network. This is achieved by the implementation of a pause frame mechanism at times of congestion.
Security	iSCSI implements the Challenge Handshake Authentication Protocol (CHAP) to ensure that initiators and targets trust each other. VLANs or private networks are highly recommended, to isolate the iSCSI traffic from other traffic types.	VLANs or private networks are highly recommended, to isolate the NFS traffic from other traffic types.	Some FC switches support the concepts of a VSAN, to isolate parts of the storage infrastructure. VSANs are conceptually similar to VLANs. Zoning between hosts and FC targets also offers a degree of isolation.	Some FCoE switches support the concepts of a VSAN, to isolate parts of the storage infrastructure. Zoning between hosts and FCoE targets also offers a degree of isolation.

	iSCSI	NFS	FIBRE CHANNEL	FCoE
VMware vSphere Storage APIs – Array Integration (VAAI) Primitives	Although VMware vSphere® Storage APIs – Array Integration (VAAI) primitives can vary from array to array, iSCSI devices can benefit from the following full complement of block primitives: • Atomic test/set • Full copy • Block zero • Thin provisioning • UNMAP These primitives are built in to ESXi and require no additional software installed on the host.	Again, these vary from array to array. The following VAAI primitives are available on NFS devices: • Full copy (but only with cold migration—not with VMware vSphere® Storage vMotion®) • Preallocated space (WRITE_ZEROs) • Cloned offload using native snapshots A plug-in from the storage array vendor is required for VAAI NAS.	Although VAAI primitives can vary from array to array, FC devices can benefit from the following full complement of block primitives: • Atomic test/set • Full copy • Block zero • Thin provisioning • UNMAP These primitives are built in to ESXi and require no additional software installed on the host.	Although VAAI primitives can vary from array to array, FCoE devices can benefit from the following full complement of block primitives: • Atomic test/set • Full copy • Block zero • Thin provisioning • UNMAP These primitives are built in to ESXi and require no additional software installed on the host.
ESXi Boot from SAN	Yes	No	Yes	Software FCoE – No Hardware FCoE (CNA) – Yes
RDM Support	Yes	No	Yes	Yes
Maximum Device Size	64TB	Refer to NAS array vendor or NAS server vendor for maximum supported datastore size. Theoretical size is much larger than 64TB but requires NAS vendor to support it.	64TB	64TB
Maximum Number of Devices	256	Default: 8 Maximum: 256	256	256
Protocol Direct to Virtual Machine	Yes, via in-guest iSCSI initiator.	Yes, via in-guest NFS client.	No, but FC devices can be mapped directly to the virtual machine with NPIV. This still requires prior RDM mapping to the virtual machine, and hardware must support NPIV (FC switch, HBA).	No
Storage vMotion Support	Yes	Yes	Yes	Yes

	iSCSI	NFS	FIBRE CHANNEL	FCoE
Storage DRS Support	Yes	Yes	Yes	Yes
Storage I/O Control Support	Yes, since vSphere 4.1.	Yes, since vSphere 5.0.	Yes, since vSphere 4.1.	Yes, since vSphere 4.1.
Virtualized MSCS Support	No. VMware does not support MSCS nodes built on virtual machines residing on iSCSI storage. However, the use of software iSCSI initiators within guest operating systems configured with MSCS, in any configuration supported by Microsoft, is transparent to ESXi hosts. There is no need for explicit support statements from VMware.	No. VMware does not support MSCS nodes built on virtual machines residing on NFS storage.	Yes. VMware supports MSCS nodes built on virtual machines residing on FC storage.	No. VMware does not support MSCS nodes built on virtual machines residing on FCoE storage.
Ease of Configuration	Medium – Setting up the iSCSI initiator requires aptitude and the FDQN or IP address of the target, plus some configuration for initiator maps and LUN presentation on the array side. After the target has been discovered through a scan of the SAN, LUNs are available for datastores or RDMs.	Easy – This requires only the IP or FQDN of the target, plus the mount point. Datastores appear immediately after the host has been granted access from the NFS array/server side.	Difficult – This involves zoning at the FC switch level and LUN masking at the array level after the zoning is complete. It is more complex to configure than IP storage. After the target has been discovered through a scan of the SAN, LUNs are available for datastores or RDMs.	Difficult – This involves zoning at the FCoE switch level and LUN masking at the array level after the zoning is complete. It is more complex to configure than IP storage. After the target has been discovered through a scan of the SAN, LUNs are available for datastores or RDMs.

	iSCSI	NFS	FIBRE CHANNEL	FCoE
Advantages	No additional hardware is necessary. Can use existing networking hardware components and iSCSI driver from VMware, so it's inexpensive to implement. Well-known and well-understood protocol. Quite mature at this stage. Administrators with network skills should be able to implement. Can be troubleshooted with generic network tools such as Wireshark.	No additional hardware is necessary. Can use existing networking hardware components, so it's inexpensive to implement. Well-known and well-understood protocol. It also is very mature. Administrators with network skills should be able to implement. Can be troubleshooted with generic network tools such as Wireshark.	Well-known and well-understood protocol. Very mature and trusted. Found in majority of mission-critical environments.	Enables consolidation of storage and other traffic onto the same network via converged network adapter (CNA). Using Data Center Bridging Exchange (DCBX) protocol, FCoE has been made lossless even though it runs over Ethernet. DCBX does other things, such as enabling different traffic classes to run on the same network, but that is beyond the scope of this discussion.
Disadvantages	Inability to route with iSCSI binding implemented. Possible security issues because there is no built-in encryption, so care must be taken to isolate traffic (e.g., VLANs). Software iSCSI can cause additional CPU overhead on the ESX host. TCP can introduce latency for iSCSI.	Because there is only a single session per connection, configuring for maximum bandwidth across multiple paths requires some care and attention. No PSA multipathing. Same security concerns as with iSCSI, because everything is transferred in clear text, so care must be taken to isolate traffic (e.g., VLANs). NFS is still version 3, which does not have the multipathing or security features of NFS v4 or NFS v4.1. NFS can cause additional CPU overhead on the ESX host. TCP can introduce latency for NFS.	Still runs only at 8Gb, which is slower than other networks (16Gb throttled to run at 8Gb in vSphere 5.0). Requires dedicated HBA, FC switch, and FC-capable storage array, which makes an FC implementation somewhat more expensive. Additional management overhead (e.g., switch zoning) is required. Might prove harder to troubleshoot than other protocols.	Somewhat new and currently not quite as mature as other protocols. Requires a 10Gb lossless network infrastructure, which can be expensive. Cannot route between initiator and targets using native IP routing. Instead, it must use protocols such as FIP (FCoE Initialization Protocol). Might prove complex to troubleshoot/isolate issues, with network and storage traffic using the same pipe.

Conclusion

The objective of this white paper is to provide information on storage protocols and how they interoperate with VMware vSphere and related features. Not all supported storage protocols are discussed. Some notable exceptions are ATA over Ethernet (AoE) and shared/switched SAS. However, the protocols that are included in this paper are the ones that VMware is most frequently asked to compare.

As mentioned in the introduction, we have intentionally avoided comparing the performance of each of the protocols. The following VMware white papers already have examined the performance of the protocols from a vSphere perspective:

- *Achieving a Million I/O Operations per Second from a Single VMware vSphere 5.0 Host*
  http://www.vmware.com/files/pdf/1M-iops-perf-vsphere5.pdf
- *Comparison of Storage Protocol Performance in VMware vSphere 4*
  http://www.vmware.com/files/pdf/perf_vsphere_storage_protocols.pdf
- *Comparison of Storage Protocol Performance*
  http://www.vmware.com/files/pdf/storage_protocol_perf.pdf

In addition, ease of configuration is highly subjective. In this paper, the author simply shares his own experiences in configuring these protocols from a vSphere perspective.

About the Author

Cormac Hogan is a senior technical marketing manager with the Cloud Infrastructure Product Marketing group at VMware. He is responsible for storage in general, with a focus on core VMware vSphere storage technologies and virtual storage, including the VMware vSphere® Storage Appliance. He was one of the first VMware employees at the EMEA headquarters in Cork, Ireland, in April 2005. He spent a number of years as the technical support escalation engineer for storage before moving into a support readiness training role, where he developed training materials and delivered training to technical support and VMware support partners. He has been in technical marketing since 2011.

- Follow Cormac's blogs at http://blogs.vmware.com/vsphere/storage.
- Follow Cormac on Twitter: @VMwareStorage.

vmware®

VMware, Inc. 3401 Hillview Avenue Palo Alto CA 94304 USA Tel 877-486-9273 Fax 650-427-5001 www.vmware.com