NetApp White Paper

Converged Enhanced Ethernet– Good for iSCSI SANs

Graham Smith, BLADE Network Technologies;
Jason Blosil, NetApp; Mike DiMeglio, NetApp
December 2008 | WP-7052-1208

EXECUTIVE SUMMARY

The introduction of new features to enhance the Ethernet protocol will improve the performance characteristics of storage applications. Although commonly associated with an emerging standard, Fibre Channel over Ethernet (FCoE), these new enhancements will contribute to improved performance of iSCSI storage traffic as well.

TABLE OF CONTENTS

1 INTRODUCTION 3

2 INTRODUCING LOSSLESS ETHERNET 4

PRIORITY FLOW CONTROL (IEEE 802.1QBB) 4

CONGESTION NOTIFICATION (IEEE 802.1QAU) 5

SHORTEST PATH BRIDGING (IEEE 802.1AQ) 5

ENHANCED TRANSMISSION SELECTION (802.1QAZ) 6

3 CONCLUSION 7

1 INTRODUCTION

Today, Ethernet is a predominant network choice for interconnecting resources in the data center. It is ubiquitous and well understood by network engineers and developers worldwide, and it has stood the test of time against challengers trying to displace it as the popular option for data center network environments. However, emerging demands by applications require additional capabilities in networking infrastructures, resulting in deployment of multiple, separate, application-specific networks. It is common for enterprise data centers to deploy an Ethernet network for IP traffic, Fibre Channel storage area networks (SANs) for block mode SCSI traffic, and maybe even an InfiniBand fabric for high-performance computing clusters.

The combined capital and operating costs for deployment and management of three distinct network types are high, creating an opportunity for consolidation on a unified fabric. When the three types of networks are evaluated technically, Ethernet has the most promise for meeting most of the requirements of all three network types, but it requires some additional capabilities and features to extend its abilities as a robust network storage fabric. These features offer a number of benefits and are usually associated with an emerging protocol that is currently in standards review, Fibre Channel over Ethernet (FCoE). Although these new Ethernet features are not required to support IP SANs (iSCSI) as they are with FCoE, the benefits can also be extended to IP SANs to deliver a more robust, higher performance iSCSI SAN implementation.

Network consolidation using 10Gb Ethernet is one of the main business benefits of this approach. The ubiquity and increased bandwidth capability of 10 GbE can eliminate the need for a separate physical network infrastructure for storage network traffic. The increased bandwidth can give a boost to existing Ethernet-based storage protocols iSCSI and CIFS/NFS, as well as increase traction for the upcoming Fibre Channel over Ethernet (FCoE) standard. As data volumes increase, bandwidth is likely to become something of a bottleneck, but storage success is based on more than raw throughput. Robustness, cost reduction, and ease of use are key goals for all organizations, and the convergence between SAN and LAN, made possible by storage over Ethernet, and lossless Ethernet will be a major step toward accomplishing these goals.

This paper reviews these new Ethernet features and describes how they improve robustness, management, and performance of iSCSI SANs.

2 INTRODUCING LOSSLESS ETHERNET

Today Ethernet is a best-effort network that may drop packets or deliver packets out of order when the network is busy, resulting in retransmissions and time-outs. The iSCSI protocol allows retransmissions and time-outs because iSCSI leverages the TCP and IP standards. However, reductions in possible network losses can improve iSCSI SAN performance.

New standards are being developed that will create a new, more capable family of Ethernet protocols. These standards, referred to collectively as Converged Enhanced Ethernet (CEE), are being developed by IEEE 802.1 and IETF standards bodies.

- Priority Flow Control (IEEE 802.1Qbb)
- Congestion notification (IEEE 802.1Qau)
- Shortest path bridging (802.1aq)
- Link layer routing protocol (IETF – TRILL)
- Enhanced Transmission Selection (802.1Qaz)

PRIORITY FLOW CONTROL (IEEE 802.1QBB)

Converged Enhanced Ethernet (CEE) capable products will enable lossless Ethernet fabrics by using IEEE 802.1Qbb Priority Flow Control (PFC) to pause traffic based on the priority levels. 802.1Qbb allows eight virtual lanes to be created in an Ethernet link, with each virtual lane assigned a priority level. During periods of heavy congestion, lower priority traffic can be paused, while allowing high-priority and latency-sensitive tasks such as data storage to continue.

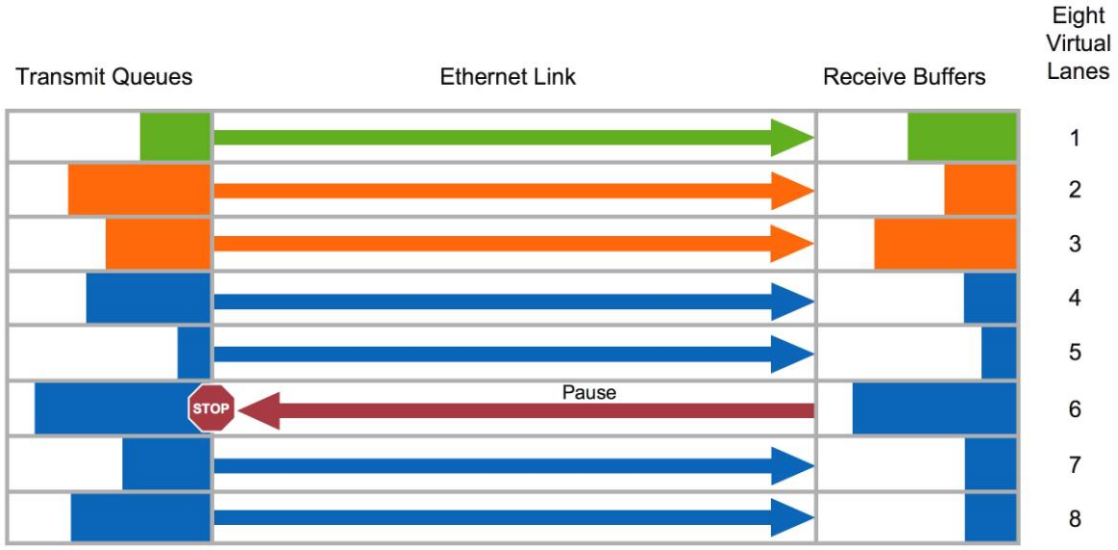

Figure 1) Priority Flow Control allows up to eight prioritized virtual lanes per Ethernet link.

IEEE 802.1Q (Virtual LAN) can be used to partition the physical Ethernet fabric to create high levels of security by isolating traffic types and to enhance quality of service (QoS)by configuring guaranteed bandwidth and latencies per VLAN. Using VLANs and 802.1Qbb flow control, several high-performance lanes of lossless Ethernet can be established on a single 10 Gigabit Ethernet fabric.

Not all data traffic has the same priority. For instance, storage traffic is generally higher priority than other network traffic, such as e-mail or instant messaging. For this reason, among others, it is common to have a dedicated storage network where security, QoS, and performance can be managed independently from the LAN. PFC, coupled with VLANs, allows LANs, SANs, and other application-specific networks to coexist on the same wire, while remaining isolated from each other logically to ensure I/O security. The result is improved iSCSI performance and savings associated with sharing a common wire and switch fabric.

CONGESTION NOTIFICATION (IEEE 802.1QAU)

Ethernet will be further enhanced by IEEE 802.1Qau, which will provide end-to-end flow control by allowing congested points to request that ingress ports limit their transmission when congestion is occurring. When the congestion has eased, the ingress ports are informed that they can increase their transmission again.

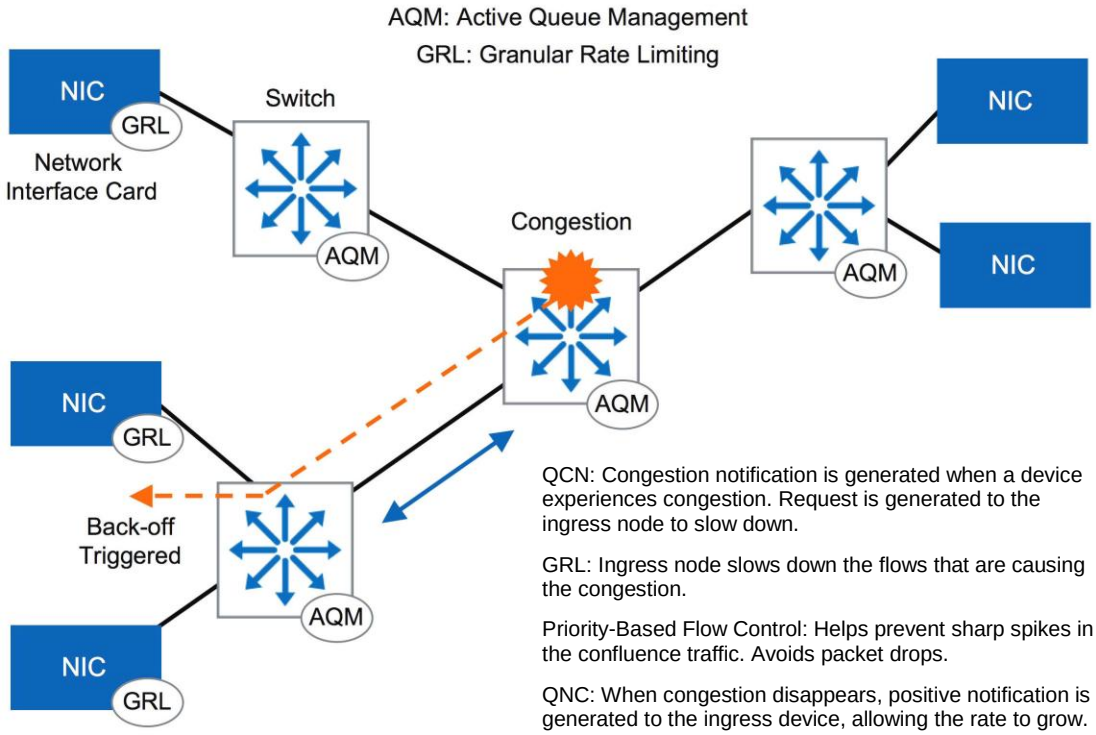

Figure 2) Congestion management throttles down I/O traffic until network congestion points are relieved.

iSCSI performance is enhanced with congestion notification. Instead of experiencing dropped packets, with their associated long timeouts and retransmissions due to traffic congestion, data traffic is automatically throttled down until congestion points are relieved, at which time throughput increases. Although this feature doesn't increase the speed of transmissions, a temporary reduction in transmission rate is preferred over lost packets causing long timeouts and requiring retransmissions.

SHORTEST PATH BRIDGING (IEEE 802.1AQ)

Shortest path bridging is an incremental advance to the Multiple Spanning Tree Protocol (MSTP), which uses the link state protocol (IS-IS) to share learned topologies between switches and to enable the rapid learning of the shortest paths between end points across the Ethernet fabric. This greatly improves performance and efficiency because the fabric dynamically adjusts to changes as links are added or removed.

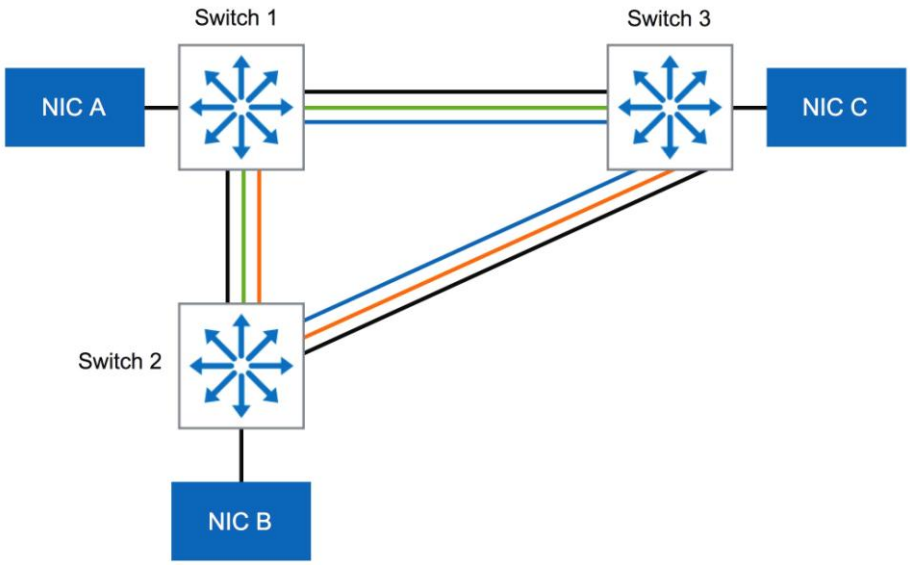

Figure 3) Shortest path bridging discovers optimal network path between source and destination.

The Internet Engineering Task Force (IETF) is also working on a similar solution called TRILL (Transparent Interconnect of a Lot of Links) to enable shortest path frame routing across Ethernet fabrics.

iSCSI networks benefit from shortest path bridging to reduce unnecessary hops along the data path. As network devices learn the shortest paths from node to node, all traffic benefits from more efficient data paths.

ENHANCED TRANSMISSION SELECTION (802.1QAZ)

For effective fabric consolidation, important traffic like storage data can be assigned high priorities and guaranteed bandwidths. To improve the overall network efficiency, 802.1Qaz (ETS) allows lower priority traffic to use unused bandwidth from the high-priority queues and to exceed their own bandwidth guarantees.

In Figure 4, four traffic types have been prioritized and assigned guaranteed minimum bandwidths. Over time, the rate of high-priority traffic like iSCSI varies. When the rate is lower than the guaranteed minimum, the bandwidth is made available for lower priority traffic types, such as e-mail in this example.

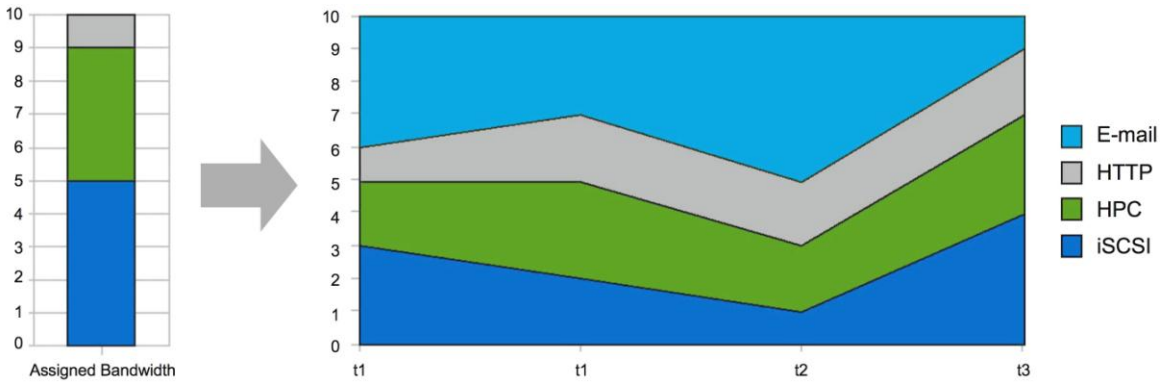

Figure 4) ETS allows lower priority traffic to consume unused bandwidth until required by higher priority traffic.

Although enhanced transmission selection may not improve performance for the highest priority traffic, such as iSCSI storage, it does contribute to overall network efficiency by optimizing network bandwidth, further justifying the move to a converged Ethernet fabric.

3 CONCLUSION

The continuing expansion of 10 Gigabit Ethernet combined with lossless Ethernet supports mixing of traffic types between servers, switched networks, and networked storage targets. Lossless Ethernet extensions (PFC, ETS, Bridging, and Congestion Notification) enable a 10 Gigabit Ethernet connection to support multiple traffic types simultaneously, preserving their respective traffic properties. With these extensions, the same 10 Gigabit Ethernet link can enhance iSCSI performance and robustness and also be able to support Fibre Channel storage traffic by offering a no-drop capable consolidated I/O fabric.

Combining enhancements in consolidated I/O Ethernet Storage with changes in server architectures is influencing the move to unified I/O as well. The adoption of PCI-Express enables servers to overcome the I/O bottleneck at the PCI bus, so that they can use a full 10 Gigabit Ethernet interface. At the same time, servers are using higher density chips, quad cores, and multiprocessor platforms, resulting in more demand for greater bandwidths into and out of the servers. With multiple processors, cores, and virtual machines existing on single servers, 10 Gigabit Ethernet is being widely adopted as a method for managing multiple traffic types simultaneously and meeting the emerging demands of applications.

IT centers of all sizes can benefit from converging LAN and SAN traffic on one consolidated Ethernet wire with all of the QoS, security, and performance associated with a dedicated network. And as other applications, such as Voice over IP, become more widely used, the efficiencies associated with a single unified fabric around Ethernet become that much more attractive.