

10G Ethernet aggregation on uplinks to interface with storage enables significant performance and management advantages for IP-based storage using iSCSI or NAS.

Fabric Convergence: Simplify, Reduce Costs & Power NOW

Today, Ethernet is the dominant, preferred and familiar network choice for connection of server and storage resources in the data center. Ethernet’s worldwide acceptance amongst network engineers and developers results in a much higher degree of continual downward price pressure compared to more specialized technologies. Ethernet has stood up to many challenges over the years, often incorporating the best features of its competitors, and through it all has remained the most popular option for data center network environments.

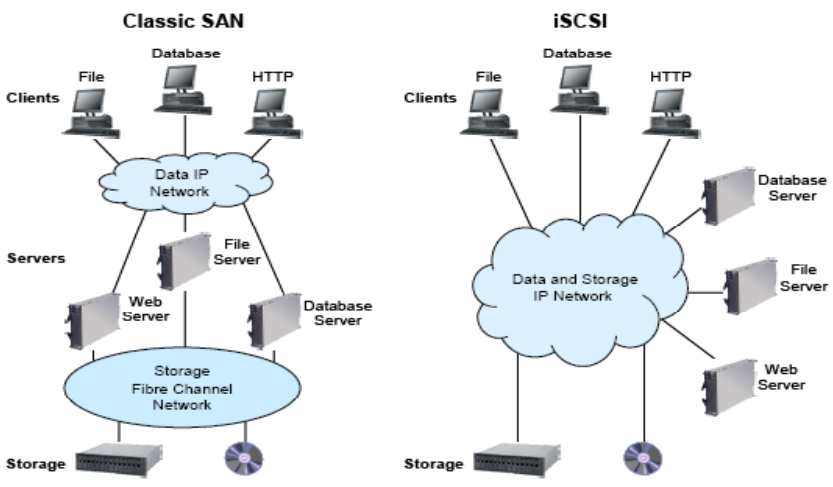
The exponential increase in bandwidth enabled by the market maturity of 10 Gigabit Ethernet (10 GE) has further increased the attractiveness of Ethernet as a data center network fabric. Efforts are underway to establish Ethernet as the sole fabric for convergence of the data and fiber channel storage-area networks (SANs) in the medium- to long-term. But even in the short-term, the move to 10 GE aggregation on uplinks connected to 10 GE interfaces on storage controllers enables a substantial performance increase and favorable infrastructure for implementation of IP-based storage using either iSCSI for block-based access or Networked Attached Storage (NAS) for file-based transfer.

In this paper we explore the potential and positioning of an IP SAN solution using software or hardware iSCSI initiation technology combined with 10 GE Uplink switching modules from BLADE Network Technologies on the IBM BladeCenter platform.

The Evolution of Storage Access

Networks must keep up with the evolving demands of applications. As a result, multiple network fabrics have been deployed to serve specific requirements of these applications. Today, enterprises commonly implement up to three separate network fabrics in their data centers; Ethernet for IP traffic, Fibre Channel storage area networks (SANs) for block mode storage traffic, and perhaps even InfiniBand for high-performance computing clusters.

The iSCSI protocol is a transport for the block-based SCSI protocol over a TCP/IP network. In the past, standard IP protocol infrastructure (for example, Ethernet) could not provide the required high bandwidth



and low-latency needed for communication to storage. Up until now, standard Ethernet has provided a network connection on a best-effort basis that could cause packets to be either dropped or delivered out of order when the network was busy, resulting in time-outs and subsequent retransmissions. Evolving to a 10 GE uplink connection within the data center will alleviate a potential bottleneck on the uplink network, and provide an iSCSI connection with much improved performance

compared to standard 1 GE uplinks. While best-effort communications of IP protocol have been highly effective over unreliable network infrastructure such as for wide-area communications, the risk of retransmissions is less critical in a controlled data center environment serving the connection of server and storage resources over a short distance.

1 Gigabit Ethernet is well established in the data center and is competitive with 1 and 2 Gigabit Fibre Channel. With the recent market transition in Ethernet technology from 1 GE to 10 GE, it is now practical (from a performance perspective) to access high performance storage devices over an IP network. 10 GE has recently become affordable and is competitive with 8 Gigabit Fibre Channel, offering a high performance and low latency network infrastructure that allows customers to implement a single TCP/IP network infrastructure for storage and data traffic. (For detailed performance data comparing 10 GE iSCSI and NAS with 4 Gb Fibre Channel, please refer to the link to the benchmark report at the end of this document.)

A major advantage of iSCSI has over competing transport protocols such as FCP, is that iSCSI can run over standard, off-the-shelf Ethernet network components. A network that incorporates iSCSI SANs need use only a single kind of network infrastructure (Ethernet) for both data and storage traffic, while use of FCP requires a separate type of infrastructure (Fibre Channel) and administration for the storage. Furthermore, iSCSI (TCP) based SANs can extend over arbitrary distances, and are not subject to distance limitations that limit FCP.

Deployments of iSCSI have experienced rapid growth in a variety of applications, including file servers, Microsoft Exchange and Microsoft SQL Server databases, and VMware ESX servers. A variety of iSCSI target systems, including some running at 10 GE, are available from major vendors including Hitachi Data Systems, IBM and NetApp.

Converging the Network Fabric in the Data Center

BLADE Network Technologies has developed optimized solutions for converged SAN and LAN networks for the IBM BladeCenter. The combination of software iSCSI initiators (or hardware initiators if boot from SAN is a requirement) on blade servers, directly connected over the blade chassis backplane to a BLADE Network Technologies 1/10G Ethernet Uplink switch (IBM part 44W4404) provides a proven and consolidated solution for connecting IBM BladeCenter to an iSCSI storage array over Ethernet running at either 1 Gigabit/sec or 10 Gigabits/sec. BLADE Network Technologies' embedded switches for IBM BladeCenter operate at 100% line-rate, ensuring maximum throughput and a reduced risk of re-transmissions within the data center.

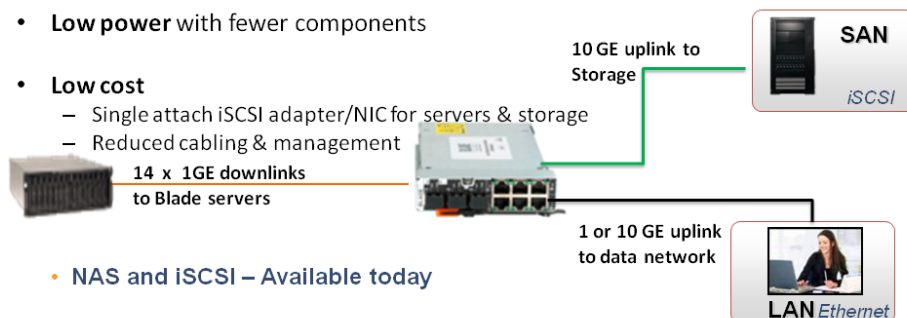
Consolidation of Network Fabrics with the 1/10 GE Uplink Switch

Carry both SAN and LAN traffic on 10G Ethernet

- **Line Rate**
 - Blade Network switches perform at 100% line rate
 - IP SANs - NAS and iSCSI performance improvement
 - VLAN security and QOS functionality for prioritization of storage traffic

- **Low power** with fewer components

- **Low cost**
 - Single attach iSCSI adapter/NIC for servers & storage
 - Reduced cabling & management



- **NAS and iSCSI – Available today**

This line-rate Ethernet solution addresses the mostly likely bottleneck in a data center networking environment for the IBM BladeCenter—the uplink from the consolidation of 14x 1 GE blade servers—while taking advantage of the increasing prevalence of 10 GE interfaces on storage arrays. The BLADE Network Technologies 1/10 GE Uplink module provides a total of 36 Gigabits/sec uplink bandwidth, reduced complexity and lower power usage, all at a price comparable to competitive solutions that have only 1 GE uplinks. Plus clients are also entitled to SmartConnect™ software, which provides innovative capabilities like switch stacking and VMready™ network virtualization. The combination of an iSCSI initiator on the host with the 1/10 GE Uplink module enables a feature rich iSCSI implementation that can grow with the customer's needs. This is a future-proof solution—when the customer decides to move from 1 GE to 10 GE uplinks, no re-investment in new switch hardware is required. SFP+ 10 GE uplink technology allows for both short-range (up to 300 meters) and long-range (up to 10 kms) connectivity. Direct attach cabling enables connectivity for a fraction of the cost of previous 10 GE fiber connectivity solutions.

Converged Enhanced Ethernet

New standards are now being developed that will create a new, more capable family of Ethernet protocols even better suited to serving combined data and storage access. These standards, collectively referred to as Converged Enhanced Ethernet (CEE), are being developed by IEEE 802.1 and IETF standards bodies. CEE will enable a guaranteed “lossless” environment which will further benefit iSCSI transport over a converged network based on 10 GE Ethernet, as well as enabling Fiber Channel SAN traffic to be combined on a single link with TCP/IP data traffic. This converged network, on the basis of Fiber Channel over Ethernet (FCoE) will enable storage elements to be discovered and the storage traffic to be prioritized over other types of traffic.

CEE incorporates features that make it suitable as a universal data center I/O fabric, with the ability to transport storage payloads without the overhead of TCP/IP, and without packet loss.

Conclusion

Combining enhancements in consolidated I/O Ethernet Storage with changes in server architectures is driving the move to a unified I/O fabric in the data center. Servers are now 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. iSCSI connectivity in the IBM BladeCenter using iSCSI hardware initiation and 10 GE uplink connectivity is a solution available now. The next dimension of network fabric convergence of data and SAN traffic is in lossless Ethernet and FCoE. The benefits of a converged network using the 10 GE are available today with an iSCSI SAN and 10 GE uplink in IBM BladeCenter.

iSCSI vs. FCoE		
	iSCSI	FCoE
<u>Payload</u>	<u>SCSI</u>	<u>Fibre Channel</u>
<u>Transport</u>	<i>TCP/IP over Ethernet</i>	<i>Enhanced Ethernet</i>
<u>Network scope</u>	<i>LAN, MAN, WAN</i>	<i>Data center (non-WAN routable)</i>
<u>Server CPU overhead</u>	<i>Medium to Low with Offload</i>	<i>Low</i>
<u>Speed</u>	<i>1 or 10 Gigabit Ethernet</i>	<i>10 Gigabit Ethernet</i>
<u>Works with pre-CEE Ethernet</u>	<i>Yes</i>	<i>No</i>
<u>Typical applications</u>	<i>Small, midsize SANs</i>	<i>Midsize, large SANs</i>
<u>Availability</u>	<i>Now</i>	<i>Later in 2009*2010</i>

Source:
<http://www.informationweek.com/story/showArticle.jhtml?articleID=212901028>

Source: <http://www.informationweek.com/story/showArticle.jhtml?articleID=212901028>

Robustness, cost reduction, and ease of use are key goals for all organizations, and the convergence between SAN and LAN, made possible by carrying storage over Ethernet, as well as lossless Ethernet will be major steps toward accomplishing these goals. BLADE Network Technologies is an innovator in converged SAN and LAN technologies. Expect more offerings in this segment for the future.

Additional Information:

- *Converged Enhanced Ethernet – Good for iSCSI SANs* Graham Smith, BLADE Network; Jason Blasil, NetApp; Mike DiMeglio, NetApp, Oct. 2008
http://www.bladenetwork.net/userfiles/file/PDFs/WP_NetApp_Enhanced_Ethernet.pdf
- *IBM BladeCenter iSCSI SAN Solution* Rufus Credle, Khalid Ansari, Scott Lorditch (BLADE Network), P Chris Mojica <http://www.redbooks.ibm.com/redp4153>
- *Ethernet Technology Grows Up--And Out - Developments on several fronts could give iSCSI the power to muscle in on higher-end Fibre Channel SANs.*
<http://www.informationweek.com/story/showArticle.jhtml?articleID=212901028>
- *Reference to iSCSI Benchmark*
http://www.bladenetwork.net/userfiles/file/PDFs/IBM_10Gb_iSCSI_PerfTestSum.pdf

BNT 1/10 Uplink ESM Switch for IBM BladeCenter

Highlights

- *Supported by all IBM BladeCenter chassis IBM Part Number 44W4404*
 - *Standard Switch form factor - Also supported in MSIM*
- *Ports*
 - *22x 1G copper RJ-45*
 - *6x 1G RJ45 External uplinks*
 - *14x 1G Internal server links*
 - *2x Internal MM ports*
 - *3x 10G SFP+ uplink ports*
 - *SFP+ Transceivers*
 - *IBM 10GBase-SR SFP+ (PN 44W4408)*
 - *BLADE 10GBase-LR SFP+ (BNT PN BN-CKM-SP-LR)*
 - *SFP+ Copper Direct Attach Cables (up to 10m) – no transceiver*
 - *1M cable (BLADE P/N: BN-SP-CBL-1M)*
 - *3M cable (BLADE P/N: BN-SP-CBL-3M)*
 - *7M cable (BLADE P/N: BN-SP-CBL-7M)*
- *Line rate performance with no packet drop*
- *Power ~ 40W*
- *IPV6 Capable*
- *Future upgrades: Stacking, VMready and more*
- *Web Price \$4,999*



©2009 BLADE Network Technologies, Inc. BLADE Network Technologies, RackSwitch, BLADEHarmony, VMready, SmartConnect, ServerMobility, and the BLADE logo are trademarks of BLADE Network Technologies. Other trademarks are owned by their respective companies. All rights reserved. Information in this document is subject to change without notice. BLADE Network Technologies assumes no responsibility for any errors that may appear in this document. All statements regarding BLADE's future direction and intent are subject to change or withdrawal without notice, at BLADE's sole discretion. <http://www.bladenetwork.net>.

MKT090428