

Brocade Advanced Accelerator for FICON—Setting a New Standard for FICON Extension

SETTING THE BENCHMARK IN EXTENDED FICON PERFORMANCE

FICON device emulation and read/write tape pipelining technologies were first available on the Brocade USD-X and Brocade Edge M3000 extension products. These technologies provide for virtually unlimited distance extension of FICON tape and a popular mainframe disk mirroring solution from IBM, called IBM z/OS Global Mirror (zGM), previously known as Extended Remote Copy (XRC), an application that is also licensed and resold by both Hitachi Data Systems and EMC. The Brocade USD-X, Edge M3000, 7500, 7800, and FX8-24 platforms with FICON emulation and pipelining capabilities set the industry standard for FICON distance extension and became the solution of choice for thousands of mainframe enterprises around the world.

Now in its sixth-generation implementation, Brocade has leveraged these technologies to expand the FICON extension capabilities of the Brocade 7840 Extension Switch and the Brocade SX6 Extension Blade, setting yet another industry benchmark for extended FICON performance.

The emulation and pipelining technologies on the Brocade® Extension platforms provide for virtually unlimited distance extension of FICON® tape, IBM z/OS Global Mirror, and Teradata data warehousing systems. This paper details the operation and advantages of these technologies in extended FICON applications.

Introduction

The Brocade 7840 Extension Switch and the Brocade SX6 Extension Blade are the newest members of the Brocade Extension product family and the latest generation of Brocade products for FICON extension beyond traditional distance limitations, providing FICON Inter-Switch Link (ISL) extension over a Fibre Channel over Internet Protocol (FCIP) connection. This solution is very effective in configurations in which the FICON protocols and application requirements can be met, despite network bandwidth and propagation delay constraints, which present challenges for straight ISL extension.

FICON device emulation and read/write tape pipelining technologies available on all Brocade Extension products provide for virtually unlimited distance extension of FICON tape, a popular mainframe DASD mirroring solution from IBM called z/OS Global Mirror (zGM), previously known as Extended Remote Copy (XRC),

and mainframe host connectivity to Teradata data warehousing systems. The Brocade FICON emulation and pipelining capabilities set the industry standard for FICON distance extension and are the solution of choice for thousands of mainframe enterprises around the world.

Now in the sixth-generation implementation, Brocade has leveraged these technologies and taken the next step in increasing FICON device scalability and throughput performance, adding non-disruptive firmware upgrades and integrating advanced Brocade Fabric Vision® technology to expand the FICON extension capabilities of the Brocade 7840 and the Brocade SX6, setting yet another industry benchmark for extended FICON performance.

FICON Protocol Efficiencies Background

Standard FICON protocols are far more efficient than their antecedents in either ESCON or parallel channel block multiplexer protocols. The entire

FICON protocol is designed to reduce or eliminate the multiple end-to-end protocol exchanges traditionally required to support legacy device access methods and I/O drivers. In so doing, FICON delivers much better protocol efficiencies over media that subjects the protocol to increased propagation delays. For many applications, a transparent FICON frame shuttle can be expected to deliver remarkably good overall performance, although not for significantly extended distance applications (that is, hundreds or thousands of kilometers). In addition, FICON offers far greater flexibility of connectivity (up to 32,000 I/O devices per FICON Express16S channel), multiple concurrent data transfers (up to 64 concurrent operations), and full-duplex channel operations (multiple simultaneous reads and writes).

Finally, FICON is mapped over the Fibre Channel FC-2 protocol layer in 1-, 2-, 4-, 8-, and 16-gigabit per second (Gbps) implementations. In the Fibre Channel standard, FICON is defined as a level-4 layer called SB-2/3/4, which is the generic term for the IBM single-byte command architecture for attached I/O devices. FICON and SB-2/3/4 are interchangeable terms—both implemented as connectionless point-to-point or switched connectionless point-to-point Fibre Channel topologies.

The Effect of Long-Distance Networks on FICON Protocols

Information Unit (IU) pacing is an SB-2/3/4 level-4 function that limits the number of Channel Command Words (CCWs), and therefore the number of IUs, that can either transmit (write) or solicit (read) without the need for additional control-unit-generated acknowledgements, which are called “command responses.” IU pacing is not, strictly speaking, a flow control mechanism, since flow control is adequately addressed by the FC-PH level buffer-to-buffer crediting function. Rather, IU pacing is a mechanism intended to prevent I/O operations that might introduce very large data transfers from monopolizing access to Fibre Channel facilities by other concurrent I/O operations.

In essence, IU pacing provides a load-sharing or “fair-access” mechanism for multiple competing channel programs. While this facility yields desirable results, ensuring more predictable I/O response times on heavily loaded channels, it produces less optimal results for very long-distance deployments. In these cases, increased link latencies can introduce dormant periods on the channel and its Wide Area Network (WAN) link. Dormant periods occur when delays in waiting for anticipated command responses increase to the point where the pacing window prohibits

the timely execution of CCWs that might otherwise be executed to insure optimal performance. The nominal IU pacing window for 1-, 2-, 4-, 8-, and 16-Gbps FICON implementations permits no more than 16 IUs to remain uncredited. As discussed in the following section, the one exception is IBM's Extended Distance FICON enhancement for z/OS Global Mirror, where IU pacing is essentially no longer a limiting factor.

IBM z/OS Global Mirror Emulation

The IBM z/OS Global Mirror (zGM) application for mainframe is one of the most comprehensive DASD replication offerings available for z/OS environments. zGM is a coordinated communications and mirroring solution that spans storage controllers and a host-based system data mover. The system data mover controls the mirror and is responsible for reading a copy of all modified data located on primary storage, as well as maintaining a copy on secondary storage.

For Business Continuity and Disaster Recovery (BC/DR), the primary and secondary storage are typically geographically separated. The unique Brocade FICON emulation feature provides capabilities for extending the primary Direct Access Storage Device (DASD) over virtually unlimited distances, giving enterprises maximum flexibility in geographic placement.

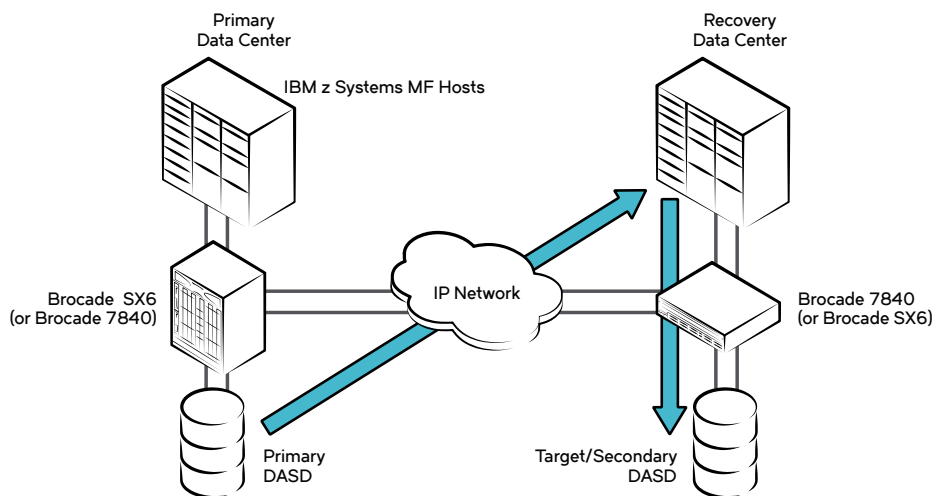


Figure 1: IBM z/OS Global Mirror (zGM) emulation.

zGM operates on two basic types of commands:

- Read Track Sets (RTSs) are used to read entire tracks during a synchronization phase.
- Read Record Sets (RRSs) are used to read records that have been modified on the primary DASD.

In zGM applications, high performance depends on the efficient execution of lengthy chains of the RRS channel programs. Since these channel programs can ultimately chain hundreds of RRS commands together, IU pacing may impose unnecessary performance constraints on long-distance and high-latency WANs. In these applications, the outbound flow of commands or command-data IUs to the device may be interrupted, due to delays in the timely receipt of required command response IUs from the device.

As part of the IBM System z10 and DS8000 announcements, IBM

announced IU pacing enhancements that allow customers to deploy zGM over long distances without a significant impact to performance. This is more commonly known as persistent IU pacing, or Extended Distance FICON. The FC-SB3 standard was amended (FC-SB3/AM1) to incorporate the changes made with persistent IU pacing. At the time of the IBM announcement, this capability was available only on the System z10 coupled with the latest DS8000 firmware. It is supported on the System z10 and later processors running Driver 73 with MCL F85898.003 or Driver 76 with MCL N10948.001. EMC supports persistent IU pacing with the Enginuity 5874 Q4 service release on their VMAX DASD array. HDS also supports persistent IU pacing with their Virtual Storage Platform.

This capability applies only to FICON ISL extension. Environments that need to traverse an IP WAN require FCIP protocol translation and associated channel extension equipment.

To take advantage of persistent IU pacing, the DASD array and control unit microcode must support it, and a System z10 or later processor is required. The Brocade zGM emulation still has a significant positive impact on performance for pre-z10 processors and earlier DASD arrays. In addition, Brocade zGM emulation can be used in conjunction with persistent IU pacing to increase capacity, compared to persistent IU pacing alone. In single System Data Mover (SDM) reader implementations, the combination of Brocade emulation with Extended Distance FICON delivered increased application I/Os per second until reaching limits of the SDM. When multiple SDM readers are deployed, the capacity expansion enabled by Brocade emulation is dramatic over all distances, increasing application I/O rates by up to 25 percent or more.

Brocade Emulation Functions in a zGM Environment

The Brocade zGM emulation process serves only this uniquely formatted channel program. Within this single channel program type, it seeks only to alleviate the dormancy that may be introduced by the effect of IU pacing over long distances and increased latency. Latency is introduced by several mechanisms, including intermediate buffering delays, signal propagation delays, and, finally, bandwidth restrictions imposed by WAN links with lower bandwidth than the Fibre Channel media (one or two gigabits). Additionally, since the IU pacing mechanism has no means of predicting the actual number of data bytes solicited by input (RRS) operations, the resulting small inbound IUs raise the opportunity to increase periods of link dormancy.

Brocade zGM emulation yields performance increases over non-emulating protocols based on the combination of all of the factors already mentioned. It is important to appreciate the impact of these factors, in order to adequately assess the impact of zGM emulation. Methods of developing such information prior to deployment include data reduction of traces of the system data mover's active channel programs with primary volumes. Additionally, information about the average propagation delay and total average bandwidth must be obtained from the WAN link provider.

In order to alleviate the excessive latency introduced by IU pacing across a WAN link, a zGM RRS emulation algorithm is implemented. RRS channel programs are differentiated from other channel programs via a unique prefixed Define Subsystem Operation (DSO) command called Command-Data IU. Information contained in the data buffer associated with this command carries both the number and the buffer size of all subsequent RRS Command IUs. Only channel programs associated with this unique DSO command are subject to the emulation process. All other channel programs and their associated information units are shuttled across the WAN link without any further processing. Thus, the emulation issue addresses only the

performance issues associated with the Channel Command Words (CCWs) that are involved in RRS operations. All other supported functions are subject to the effects of WAN link latency to some degree.

Brocade Emulation Processes for FICON Access to Tapes

Improved FICON protocol efficiencies reduce the number of end-to-end exchanges required to support tape operations, as compared with its antecedent ESCON and parallel channels implementation. However, many legacy access methods generate small channel programs consisting of as little as a single read or write CCW, normally preceded in a chain by an operating system-supplied mode-set command and, in some cases, a terminating no-op command. Thus, small channel programs that support tape operations are still serialized on a device basis by the command-data-status exchanges that typify tape read and write operations.

While these end-to-end exchanges may be considered trivial in native FICON-attached tape implementations, they can become a significant impediment to acceptable I/O access times and bandwidth utilization for WAN-supported FICON configurations. In addition to the command-data-status exchange

required to support tape operations, the effect of IU pacing may also introduce additional undesirable delays in FICON-attached tape devices accessed through WAN facilities. This is true particularly for tape write operations, where outbound data frames play a significant role in the IU pacing algorithm. The Brocade suite of emulation and pipelining functions reduces the undesirable effects of latency on these exchanges and improves overall performance for WAN-extended FICON-attached tape and virtual tape devices.

Tape pipelining refers to the concept of maintaining a series of I/O operations across a host-WAN-device environment, and it should not be confused with the normal FICON streaming of CCWs and data in a single command chain. Normally, tape access methods can be expected to read data sequentially until they reach the end-of-file delimiters (tape marks) or to write data sequentially until either the data set is closed or an end-of-tape condition occurs (multivolume file). The emulation design strategy attempts to optimize performance for sequential reads and writes, while accommodating any other nonconforming conditions in a lower-performance, non-emulating frame shuttle. Since write operations can be expected to comprise the larger percentage of I/O operations for tape devices (for archival purposes), they are addressed first.

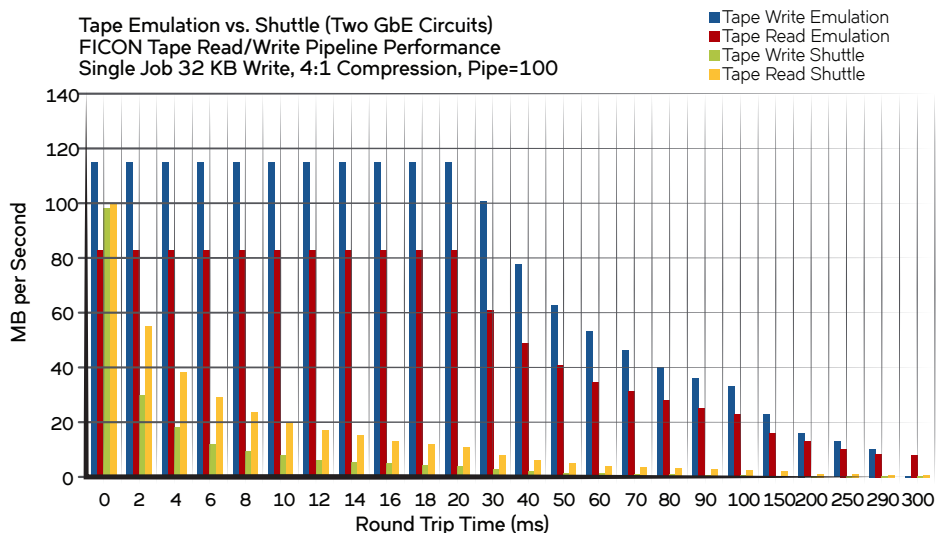


Figure 4: Brocade emulation performance advantage with one tape job.

In the second test, read pipelining performance was measured using eight concurrent tape jobs. The results show that read pipelining improves throughput from 25 percent to a massive 10x improvement over long distances. Even without emulation, better utilization of

the WAN link is achieved when scaling the number of virtual tape engines and concurrent tape jobs. When there is only one (or a few) tape jobs, FICON read emulation is critical, in order to obtain better utilization of the WAN link and higher throughput on the tape job.

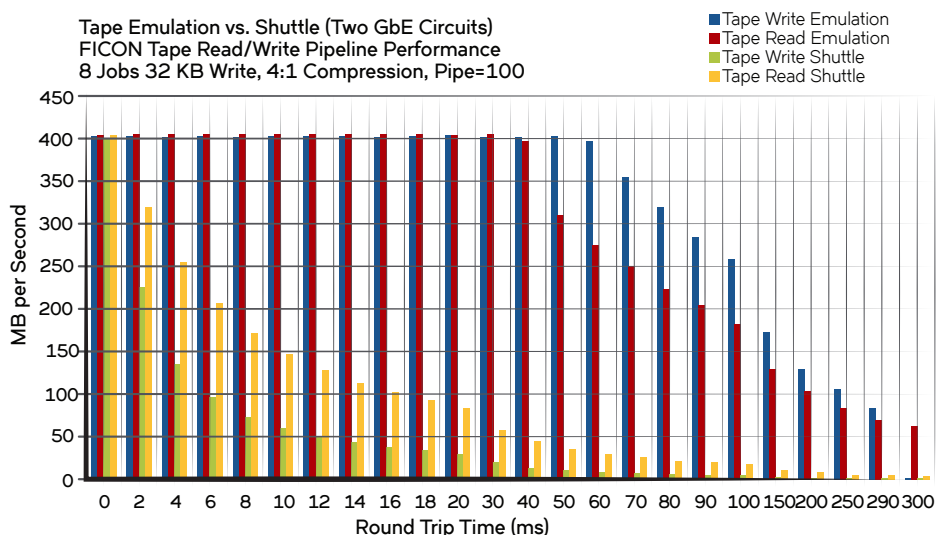


Figure 5: Brocade emulation performance advantage with eight tape jobs.

Brocade Support for Extended FICON Tape

The Brocade FICON extension feature set includes emulation support for the IBM and Oracle virtual tape systems, and Read and Write Pipelining for extended FICON tape operations. A wide range of configurations are supported for flexible disaster recovery implementations, which can be deployed in a variety of switch environments, including FICON Cascade. These configurations include:

- Extension between host and remote tape controller (front-end extension)
- Extension between local tape controller and remote tape library (back-end extension)
- Extension between clustered systems for high availability

Brocade Support for Teradata Data Warehousing Systems

The Brocade 7840 Extension Switch and the Brocade SX6 Extension Blade provide emulation support for FICON-based Teradata data warehousing systems. Emulation support for FICON-based Teradata data warehousing systems was first available on the Brocade 7800 and Brocade FX8-24 with Brocade Fabric OS® (Brocade FOS) v7.0. As with tape and zGM emulation, Brocade emulation for Teradata data warehousing systems accelerates FICON read and write commands over distance using an IP WAN.

FICON Teradata emulation services reduce latency caused by WAN propagation delays and bandwidth restrictions, by processing selected FICON commands and their associated control/data/status responses. For write

commands, control and status frames are generated for the host side of the WAN in order to pipeline write commands over the same or multiple exchanges. Such pipelined write commands and their data are queued at the device side of the WAN for asynchronous transfer to the device. For read operations received by the device side of the WAN, a number of anticipatory read commands are autonomously generated and transferred to the device. The data and status associated with such commands are sent to the host side of the WAN and queued in anticipation of host-generated read commands.

To quantify the performance gains from Brocade emulation in a Teradata environment, performance testing was conducted across a variety of latencies with one, four, and eight Control Processes (CPs). First, FastLoad (initial database loading) was tested with one CP across latencies ranging from 0 to 200 milliseconds (ms).

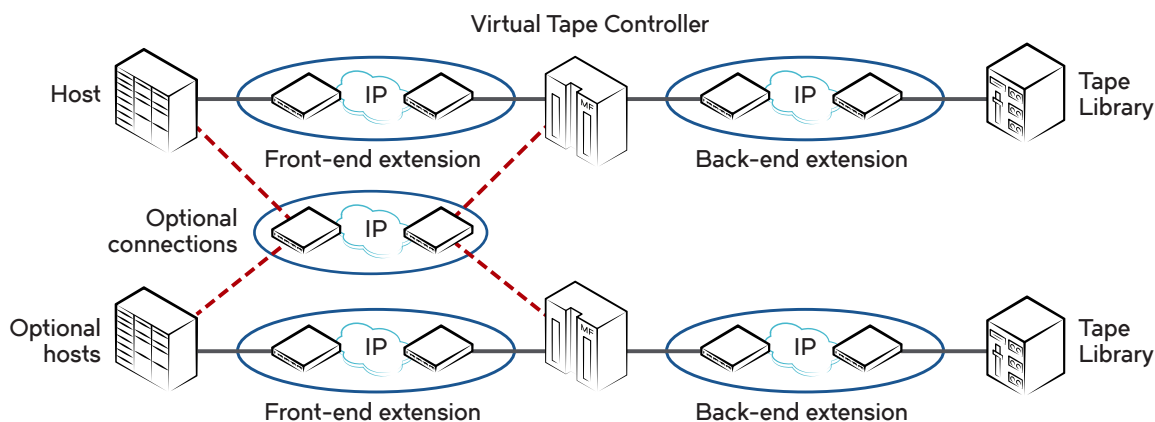


Figure 6: Supported configurations for flexible disaster recovery tape environment.

As shown in Table 1, even at 10-ms round-trip latency, pipelining reduces elapsed start-to-finish time by 80 percent. As latency increases, the performance advantage of pipelining becomes quite apparent. For instance, at 60-ms round-trip latency, pipelining reduces the start-to-finish time by 29 minutes for 1 CP, from 33 minutes down to 4 minutes. At 200 ms, elapsed time is reduced 1 hour and 35 minutes. When running four or eight concurrent control processes, pipelining elapsed times are only slightly higher than when running 1 CP.

Table 1: Brocade emulation dramatically reduces initial database loading time.

Teradata FastLoad Utility (Initial Database Loading)				
Latency (Round Trip)	Pipeline Disabled	Brocade Extension with Pipeline Enabled		
	1 CP	1CP	4CPs	8 CPs
0 ms	0 min 54 sec	0 min 54 sec	1 min 6 sec	1 min 1 sec
10 ms	5 min 51 sec	1 min 7 sec		
30 ms	16 min 40 sec	2 min 27 sec		
60 ms	33 min 48 sec	4 min 30 sec	4 min 51 sec	5 min 7 sec
100 ms	55 min 18 sec	7 min 24 sec		
200 ms	1 hr 49 min 46 sec	14 min 9 sec	15 min 1 sec	

In the second test, (shown in Table 2), write (using the Teradata ARC Restore utility) performance was measured across the same latencies. As shown below, using a Brocade extension product with pipelining dramatically reduces the completion time compared to non-pipelined writes. At 100-ms round-trip latency, pipelining reduces the elapsed time of 1 CP by 55 minutes, an 8x reduction compared to non-pipelined performance. Pipelining performance remains consistent even when running multiple concurrent CPs.

Table 2: Brocade emulation dramatically improves write performance over all distances.

Teradata ARC Restore Utility (Write to Device)				
Latency (Round Trip)	Pipeline Disabled	Brocade Extension with Pipeline Enabled		
	1 CP	1CP	4CPs	8 CPs
0 ms	3 min 22 sec	3 min 3 sec	3 min 40 sec	3 min 52 sec
10 ms	7 min 10 sec	3 min 15 sec		
30 ms	19 min 9 sec	3 min 45 sec		
60 ms	38 min 51 sec	5 min 15 sec	6 min 14 sec	6 min 3 sec
100 ms	1 hr 2 min 55 sec	8 min 18 sec		
200 ms	2 hr 4 min 48 sec	16 min 9 sec		

A third test, (shown in Table 3), was performed to measure read performance (using the Teradata ARC Dump utility). Again, the impact of Brocade pipelining is dramatic. At 30-ms round-trip latency, elapsed time goes from 19 minutes and 30 seconds with pipelining disabled to only 26 seconds with pipelining enabled. The improvement in read performance is equally dramatic across all latencies and with multiple concurrent CPs running.

Table 3: Brocade emulation dramatically improves read performance over all distances.

Teradata ARC Dump Utility (Read Back to Mainframe)				
Latency (Round Trip)	Pipeline Disabled	Brocade Extension with Pipeline Enabled		
	1 CP	1CP	4CPs	8 CPs
0 ms	43 sec	22 sec	24 sec	23 sec
10 ms	6 min 57 sec	24 sec		
30 ms	19 min 30 sec	26 sec		
60 ms	38 min 48 sec	46 sec	53 sec	55 sec
100 ms	1 hour 3 min 33 sec	1 min 15 sec		
200 ms	2 hours 6 min 4 sec	2 min 25 sec	2 min 39 sec	

What About Synchronous DASD Replication?

Clearly, with a z/OS Global Mirror implementation, using Brocade zGM emulation technology is a sound idea for achieving optimal performance, Recovery Point Objective (RPO), and Recovery Time Objective (RTO) requirements. Also, synchronous DASD copy techniques, such as IBM Metro Mirror (aka PPRC) clearly do not benefit from nor require any type of emulation technology. With synchronous DASD copy, performance, RPO, and RTO depend on available bandwidth, number of paths, and distance.

The question then becomes: What benefits does an FCIP implementation bring to a synchronous DASD remote copy configuration in a mainframe environment? If emulation technology

is not needed for synchronous copy, and long distances are not spanned, why choose FCIP instead of native Fibre Channel channel/FICON through cascaded directors or using Dense Wavelength Division Multiplexing (DWDM)?

FCIP offers several advantages over native Fibre Channel and DWDM for mainframe synchronous copy implementations. Specifically, Brocade Extension Trunking can improve the performance, RPO/RTO, availability, and reliability of synchronous copy implementations. Brocade Extension Trunking provides seamless failover, lossless link loss, redundant circuits, and load balancing across the circuits in an FCIP tunnel. All of these can greatly improve a synchronous copy implementation.

Extension Switch or Blade?

Organizations frequently ask whether they should implement the extension switch or the extension blade. The answer depends on several factors.

First, what is the organization's historical preference? Many end users find, in the absence of major technology differences between the two options, that from a management perspective it is simpler to continue using the technology that they are familiar with. Also, many end users prefer to keep the FICON directors limited to host and local storage connectivity, and keep the long-distance replication connectivity in separate devices.

Second, how many ports does the organization have installed/planned for installing in the FICON directors? In the Brocade X6-4 and X6-8 Directors,

installing a Brocade SX6 Extension Blade means using a slot in the director chassis that could be used for a Brocade FC32-48 port blade (giving up the potential 48 ports of FICON connectivity for the extension blade). Would doing so necessitate purchasing an additional Brocade X6-4 or X6-8 director chassis in order to have more ports? If the answer to that is “yes,” it may be more financially attractive to use the Brocade 7840 Extension Switch for distance replication needs.

Third, if an organization is using Brocade DCX® 8510 FICON directors, the Brocade SX6 Extension Blade is supported only in the Brocade X6 family of directors. In this case, an organization would need to use the Brocade 7840 Extension Switch.

Fourth, the Brocade 7840 Fibre Channel ports are limited to 16 Gbps, while the Brocade SX6 Extension Blade supports 32 Gbps Fibre Channel speeds.

Summary

The FICON channel architecture is designed to greatly expand the distance, throughput, capacity, and scale of implementation—all of which are disaster recovery and business continuity challenges faced by IBM z Systems end users. The FICON emulation and read/write pipelining capabilities on the Brocade Extension platforms further extend FICON capabilities, enabling mainframe physical tape, virtual tape, and zGM to be extended over virtually unlimited distances.

This dramatically improves backup and recovery performance, increases scalability, and delivers maximum WAN bandwidth utilization.

The advantages of this Brocade technology are widely deployed and proven across the large, worldwide Brocade mainframe customer base. Now in its sixth generation, FICON emulation and pipelining capabilities, plus additional capabilities such as non-disruptive firmware upgrade and advanced Brocade

Fabric Vision technology, have been implemented on the industry-leading Brocade 7840 Extension Switch and the Brocade SX6 Extension Blade. Organizations can now deploy these solutions for remote tape, Teradata data warehousing systems, and IBM z/OS Global Mirror environments over virtually unlimited distances. Such deployment furthers Brocade’s market leadership in FICON extension solutions and provides unprecedented flexibility, performance, and data protection improvements to Brocade mainframe customers.

About Brocade

Brocade networking solutions help organizations achieve their critical business initiatives as they transition to a world where applications and information reside anywhere. Today, Brocade is extending its proven data center expertise across the entire network with open, virtual, and efficient solutions built for consolidation, virtualization, and cloud computing. Learn more at www.brocade.com.

Corporate Headquarters

San Jose, CA USA
T: +1-408-333-8000
info@brocade.com

European Headquarters

Geneva, Switzerland
T: +41-22-799-56-40
emea-info@brocade.com

Asia Pacific Headquarters

Singapore
T: +65-6538-4700
apac-info@brocade.com



© 2017 Brocade Communications Systems, Inc. All Rights Reserved. 03/17 GA-WP-1909-01

Brocade, the B-wing symbol, and MyBrocade are registered trademarks of Brocade Communications Systems, Inc., in the United States and in other countries. Other brands, product names, or service names mentioned of Brocade Communications Systems, Inc. are listed at www.brocade.com/en/legal/brocade-legal-intellectual-property/brocade-legal-trademarks.html. Other marks may belong to third parties.

Notice: This document is for informational purposes only and does not set forth any warranty, expressed or implied, concerning any equipment, equipment feature, or service offered or to be offered by Brocade. Brocade reserves the right to make changes to this document at any time, without notice, and assumes no responsibility for its use. This informational document describes features that may not be currently available. Contact a Brocade sales office for information on feature and product availability. Export of technical data contained in this document may require an export license from the United States government.

BROCADE 