



MAINFRAME

Setting a New Standard for FICON Extension—Brocade Emulation and Tape Pipelining for FICON

The emulation and pipelining technologies on the Brocade extension platforms provide for virtually unlimited distance extension of FICON tape and the IBM z/OS Global Mirror (formerly eXtended Remote Copy or XRC) application. This paper details the operation and advantages of these technologies in extended FICON applications.

BROCADE

CONTENTS

Introduction.....	3
FICON Protocol Efficiencies Background.....	3
The Effect of Long Distance Networks on FICON Protocols.....	3
IBM eXtended Remote Copy (XRC) Emulation	4
Brocade Emulation Functions in an XRC Environment.....	5
Brocade Emulation Processes for FICON Access to Tapes.....	5
Performance Gains with Device Emulation	7
Brocade Support for Extended FICON Tape	7
Summary	8

INTRODUCTION

The Brocade® 7500 Extension Switch and FR4-18i Extension Blade for the Brocade 48000 Director and the DCX Backbone were the first Brocade products with FICON extension beyond traditional distance limitations, providing FICON Inter-Switch (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, while network bandwidth and propagation delay constraints are still taken into account. The next step forward for Brocade was to add selective FICON device emulation to the product line, enabling FICON extension to expand beyond the standard bandwidth and propagation delay constraints—which presents challenges for straight ISL extension.

FICON device emulation and read/write tape pipelining technologies were first available on the Brocade USD-X and Brocade Edge M3000 extension products (formerly McDATA UltraNet Storage Director eXtended and UltraNet Edge Storage Router). These technologies provide for virtually unlimited distance extension of FICON tape and a popular mainframe disk mirroring solution from IBM, called eXtended Remote Copy (XRC), an application which is also licensed and resold by both Hitachi Data Systems and EMC. The Brocade USD-X and M3000 platforms with 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 their fourth-generation implementation, Brocade has leveraged these technologies to expand the FICON extension capabilities of the Brocade 7500, 7500E, and FR4-18i platforms, 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 and/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, but not for significantly extended distance applications (that is, hundreds or thousands of kilometers). In addition, FICON offers far greater flexibility of connectivity (up to 64,000 devices per channel), multiple concurrent data transfers (up to 32 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 and 8 Gbit/sec implementations. In the Fibre Channel standard, FICON is defined as a level-4 layer called “SB-2/3,” which is the generic term for the IBM single-byte command architecture for attached I/O devices. “FICON” and “SB-2/3” 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 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 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, insuring 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 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 and 8 Gbit/sec FICON implementations permits no more than 16 IUs to remain uncredited. Pacing credits can be adjusted dynamically from these values by control unit requests for specific protocol sequences; however, the channel is not bound to honor control unit requests for larger IU pacing windows.

IBM EXTENDED REMOTE COPY (XRC) EMULATION

The IBM Global Mirror application for mainframe is one of the most comprehensive disk replication offerings available for z/OS environments. Global Mirror (formerly XRC) 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 and also maintaining a copy on secondary storage. For Business Continuity and Disaster Recovery (BC/DR), the primary and secondary storage are typically geographically separated. Brocade's unique 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.

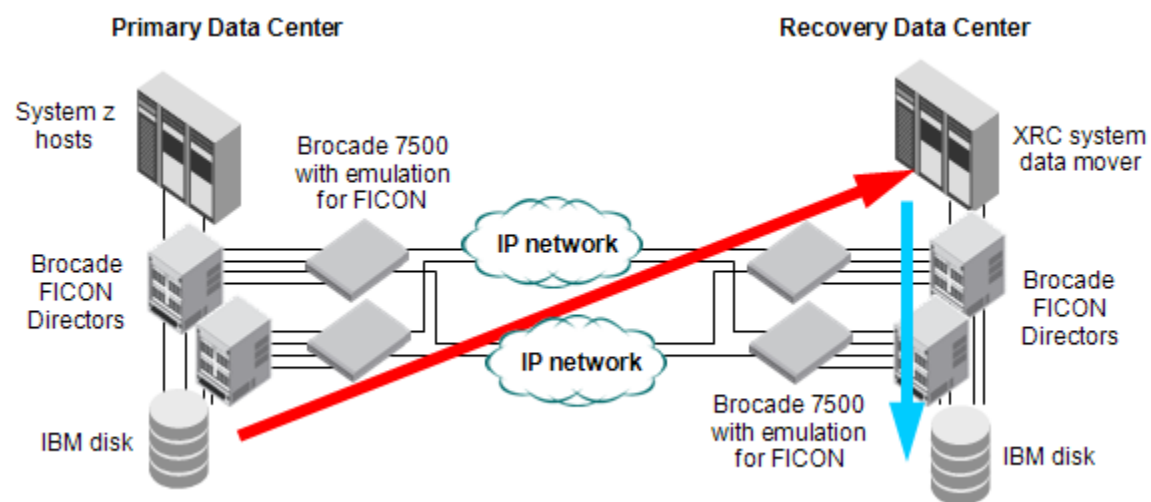


Figure 1. IBM eXtended Remote Copy (XRC) emulation

IBM Global Mirror for System z (XRC) operates on two basic types of commands:

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

In XRC 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/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 new System z10 announcement, IBM announced IU pacing enhancements that will allow customers to deploy XRC over long distances without a significant impact to performance. This new capability is available on the new System z10 coupled with the latest DS8000 firmware, while other systems are still negatively impacted by IU pacing constraints. This new capability also applies only to FICON ISL extension. Environments that need to traverse an IP WAN require FCIP protocol translation and associated channel extension equipment.

NOTE: To take advantage of the IU pacing enhancements, the control unit microcode and a System z10 processor must be updated. The Brocade Global Mirror for System z (XRC) emulation still has a significant positive impact on performance for pre-z10 processors.

BROCADE EMULATION FUNCTIONS IN AN XRC ENVIRONMENT

The Brocade Global Mirror for System z (XRC) emulation process serves only this uniquely formatted channel program and within this single channel program type 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, resultant small inbound IUs raise the opportunity to increase periods of link dormancy.

Brocade's XRC emulation yields performance increase over non-emulating protocols based on the combination of all of the aforementioned factors. In order to adequately assess the impact of XRC emulation, an appreciation of all of the factors involved is needed. 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 excessive latency introduced by IU pacing across a WAN link, an XRC RRS emulation algorithm has been implemented. RRS channel programs are differentiated from other channel programs via a unique prefixed Define Subsystem Operation (DSO) command 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 Channel Command Words (CCW) involved in RRS operations. All other supported functions will be 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 compared with its antecedent ESCON and parallel channels implementations. 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, particularly for tape write operations, where outbound data frames play a significant role in the IU pacing algorithm. Brocade's 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 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 (multi-volume file). The emulation design strategy attempts to optimize performance for sequential reads and writes, while accommodating any other non-conforming 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.

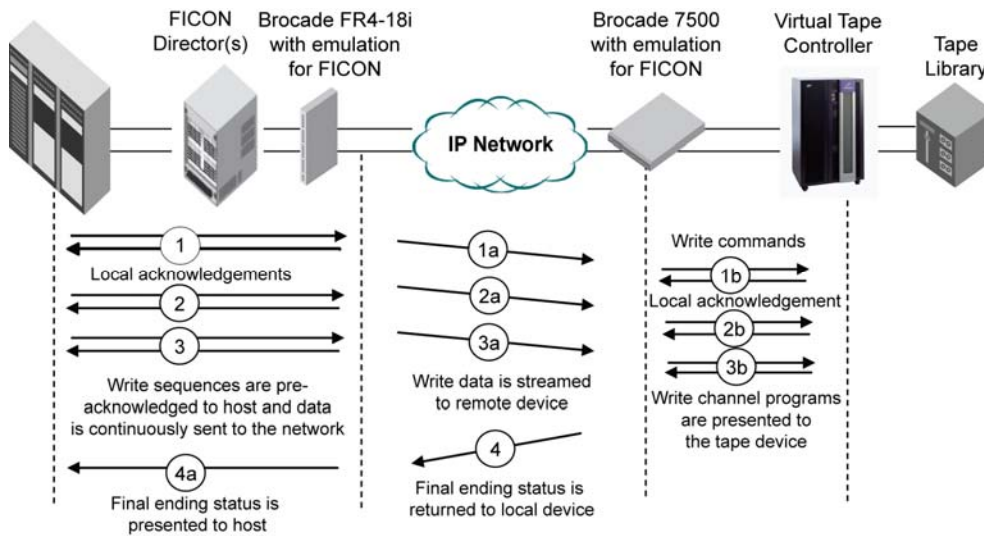


Figure 2. Brocade FICON write tape pipelining

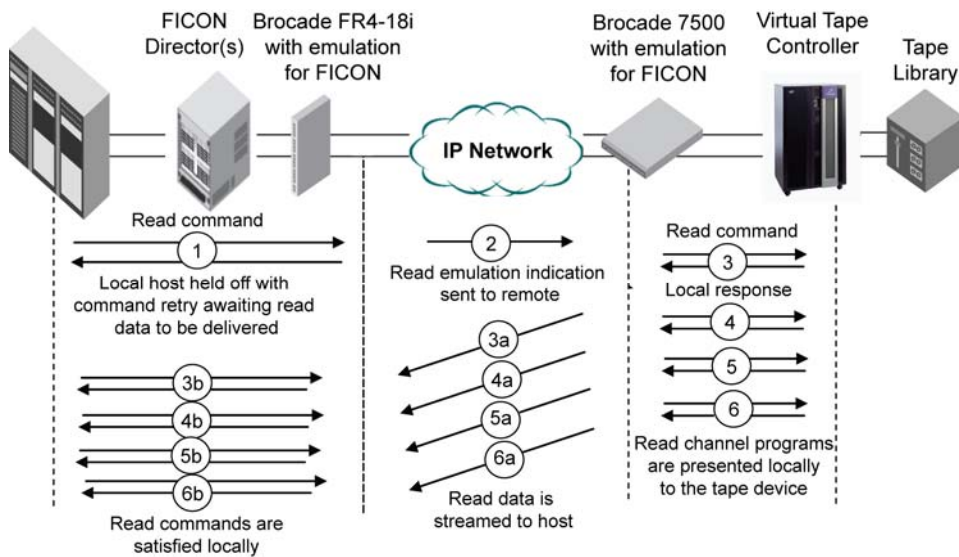


Figure 3. Brocade FICON read tape pipelining

The tape emulation features exist as an extension of the previously developed XRC emulation functions, and much of the existing XRC emulation infrastructure has been exploited. Indeed, it is possible for the code to concurrently support both the XRC emulation functions and the tape emulation functions concurrently, provided that their separate control units are dynamically switched. As with XRC emulation, the tape emulation techniques apply to the Fibre Channel frame level and not to the buffer or command level. Again, as with XRC emulation, the only parts of the tape access to be emulated are those associated with standard writes and reads, and possible accompanying mode-sets or no-ops, which constitute the bulk of the performance path for tape operations. All other control functions are not emulated and the emulation process provides no surrogate control unit/device image to the host and no surrogate host image to the device.

PERFORMANCE GAINS WITH DEVICE EMULATION

When discussing emulation, the term “shuttle mode” is often used for purposes of comparison. Shuttle mode refers to the behavior of extending FICON, in which every operation goes from end to end and performance is directly impacted by the propagation delay of each operation.

In the first FICON tape performance graph below, tests were run against a single extended IBM Virtual Tape Controller (VTC) with eight 32 KB block size tape jobs, using Brocade 7500 extension devices. Read and write pipelining performance versus shuttle mode performance was captured for various link latencies. As shown in Figure 4, read performance is lower than that of write operations and is limited by the VTC. However, the performance improvement FICON emulation provides relative to shuttle mode is pronounced for both reads and writes.

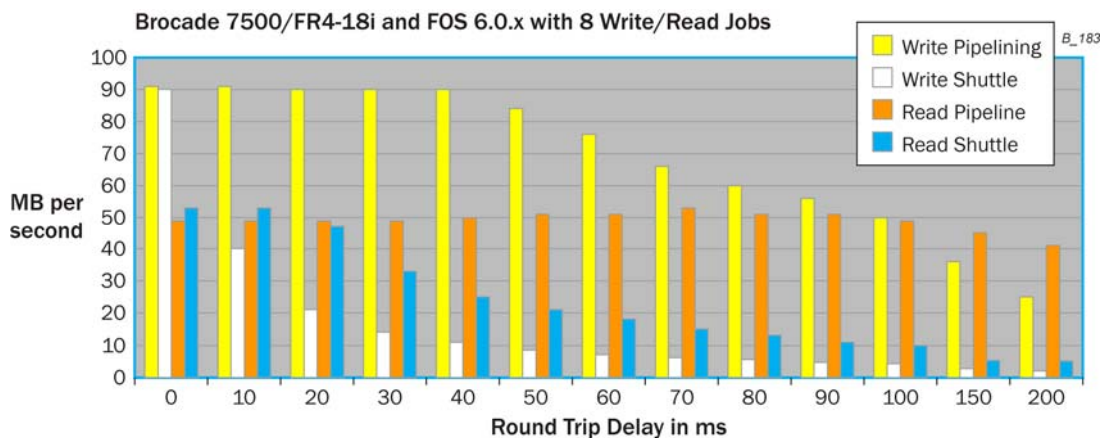


Figure 4. FICON tape read/write pipeline performance for a single VTC and 8 read/write jobs

In the second test, read pipelining performance was measured using 4 VTCs and 16 concurrent tape jobs. As shown in the graph below, read pipelining improves throughput from 25 percent to a massive 15x improvement over long distances. Even without emulation, better utilization of the WAN link is achieved when scaling the number of VTCs and concurrent tape jobs. When there is only one or a few tape jobs, FICON read emulation is critical to obtain better utilization of the WAN link and higher throughput on the tape job.



Figure 5. FICON tape read/write pipeline performance for 4 VTCs and 16 read/write jobs

BROCADE SUPPORT FOR EXTENDED FICON TAPE

The Brocade FICON extension feature set includes emulation support for the IBM Virtual Tape Server (VTS), the Sun Virtual Storage Manager (VSM), 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 (frontend extension)
- Extension between local tape controller and remote tape library (backend extension)
- Extension between clustered systems for high availability.

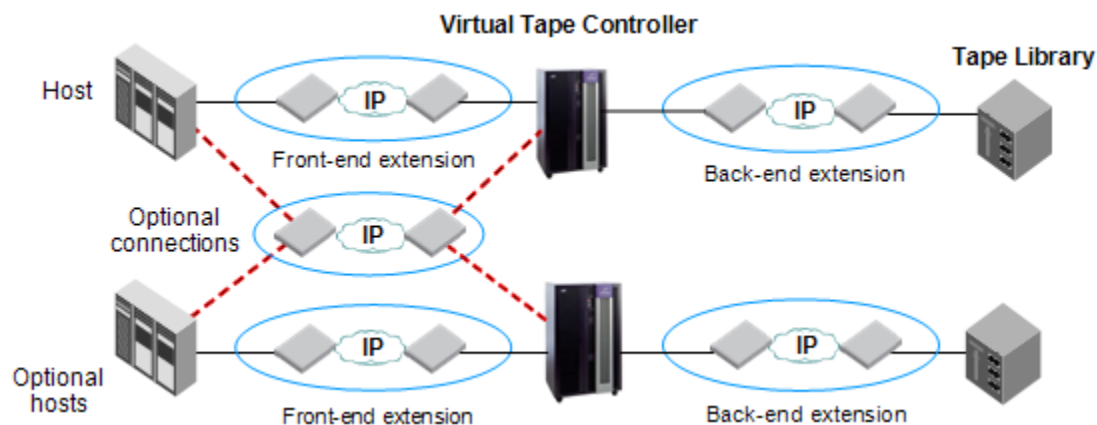


Figure 6. Supported configurations for flexible disaster recovery tape environment

SUMMARY

The FICON channel is architected to greatly expand the distance, throughput capacity, and scale of implementation—all challenges faced by the IBM System z. The FICON emulation and read/write pipelining capabilities on the Brocade USD-X and Edge M3000 extension platforms further extended FICON capabilities, enabling tape and IBM XRC to be extended over virtually unlimited distances, dramatically improving backup and recovery performance, increasing scalability, and delivering maximum WAN bandwidth utilization.

The advantages of this Brocade technology are widely deployed and proven across the Brocade large mainframe customer base. Now in its fourth generation, FICON emulation and pipelining capabilities have been implemented on the industry-leading Brocade 7500 and 7500E Extension Switches and the FR4-18i Extension Blade for the Brocade 48000 Director and DCX Backbone. These solutions can now be deployed for remote tape and XRC environments over virtually unlimited distances, furthering Brocade's market leadership in FICON extension solutions and providing unprecedented flexibility, performance, and data protection improvements to Brocade mainframe customers.

© 2008 Brocade Communications Systems, Inc. All Rights Reserved. 03/08 GA-TP-035-01

Brocade, Fabric OS, File Lifecycle Manager, MyView, and StorageX are registered trademarks and the Brocade B-wing symbol, DCX, and SAN Health are trademarks of Brocade Communications Systems, Inc., in the United States and/or in other countries. All other brands, products, or service names are or may be trademarks or service marks of, and are used to identify, products or services of their respective owners.

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.