The Association of System
Performance Professionals

The **Computer Measurement Group**, commonly called **CMG**, is a not for profit, worldwide organization of data processing professionals committed to the measurement and management of computer systems. CMG members are primarily concerned with performance evaluation of existing systems to maximize performance (eg. response time, throughput, etc.) and with capacity management where planned enhancements to existing systems or the design of new systems are evaluated to find the necessary resources required to provide adequate performance at a reasonable cost.

This paper was originally published in the Proceedings of the Computer Measurement Group's 2005 International Conference.

For more information on CMG please visit www.cmg.org

Copyright 2005 by The Computer Measurement Group, Inc. All Rights Reserved
Published by The Computer Measurement Group, Inc., a non-profit Illinois membership corporation. Permission to reprint in whole or in any part may be granted for educational and scientific purposes upon written application to the Editor, CMG Headquarters, 151 Fries Mill Road, Suite 104, Turnersville, NJ 08012. Permission is hereby granted to CMG members to reproduce this publication in whole or in part solely for internal distribution with the member's organization provided the copyright notice above is set forth in full text on the title page of each item reproduced. The ideas and concepts set forth in this publication are solely those of the respective authors, and not of CMG, and CMG does not endorse, guarantee or otherwise certify any such ideas or concepts in any application or usage. Printed in the United States of America.

Taking FICON to the Next Level: Cascaded High Performance FICON

Stephen Guendert, McDATA Corporation

Prior to the January 2003 introduction of support for cascaded FICON director connectivity on IBM zSeries mainframes, only a single level of FICON directors for connectivity between a processor and peripheral devices was allowed to be used. Cascaded FICON introduced the open systems SAN concept of the Inter-switch link (ISL). IBM now supports the flow of traffic from the processor through two FICON directors that are connected via ISL and on to the peripheral devices such as disk and tape.

Introduction

FICON, like most technological advancements, evolved from the limitations of its predecessor. ESCON (IBM's Enterprise System Connection) is a very successful storage network protocol for mainframe systems and has been considered the father of the modern SAN. Since its introduction in 1999, FICON has seen quite an evolution in its brief history. We started out with FICON bridge mode (FCV) and, in 5 short years, have gone from single director Native FICON (FC) implementations to configurations that intermix FICON and open systems fibre channel protocol (FCP). The most recent addition to the capabilities of FICON is the ability to create cascaded fabrics of FICON directors and switches.

Cascaded FICON allows the end user to have a FICON channel, or FICON CTC connect an IBM zSeries server to another zSeries server or peripheral device such as disk, tape library, or printer via two FICON directors or switches between the connected devices and/or servers. This permits tremendous flexibility and fabric cost savings in the FICON architecture, better utilization of storage resources and higher data availability in the enterprise. It also allows for more robust disaster recovery and business continuity.

This paper will discuss what cascaded FICON is, how it works, complete with an in depth discussion on buffer to buffer credits and their impact on performance in a cascaded FICON environment, and what goes into the planning, design and implementation of a cascaded FICON environment.

A few words on HA/DR/BC

The greater bandwidth of and distance capabilities FICON has over ESCON are starting to make it an essential and cost effective component in high availability/disaster recovery/business continuity (HA/DR/BC) solutions. Since Sept 11, 2001 more and more companies are insourcing DR/BC and those that are doing so are building the mainframe piece of their new DR/BC datacenters using FICON, rather than ESCON. Until IBM announced FICON cascading as being Generally Available (GA), the FICON architecture was limited to a single domain due to the single byte addressing limitations inherited from ESCON. FICON cascading allows the end user to have a greater maximum distance between sites, i.e., up to an unrepeated distance of 36 km at 2 Gb/sec bandwidth. For details please refer to Tables 1 and 2.

September 11, 2001 hammered home how critical it is for an enterprise to be prepared for disaster. This was even more so for large enterprise mainframe customers. A complete paradigm shift has occurred since 9/11/01 when the topic of DR/BC is discussed. Disaster recovery is no longer limited to problems such as fires or a small flood. Companies now need to consider and plan for the possibility of the destruction of their entire data center and, possibly, the people that work in it. A great many articles, books and other publications have discussed the IT "Lessons learned" from September 11, 2001:

1) To manage business continuity it is absolutely critical to maintain geographical separation of facilities and resources. Any resource the enterprise has that cannot be replaced from external sources within the recovery time objective (RTO) should be

available within the enterprise. It is also preferable to have these resources (buildings, hardware, software, data, and staff) in multiple locations. Cascaded FICON gives this geographical separation that post 9/11 business requires; ESCON does not.

2) The most successful DR/BC implementations are oftentimes based on as much automation as possible. 9/11/01 proved that key staff and skills may no longer be present after a disaster strikes.

3) Financial, government, military, and other enterprises now have critical Recovery Time Objectives that are seconds or minutes and not days and hours. For these end users it has become increasingly necessary to implement in in-house (insourced) DR solution. By in-house we mean that the facilities and equipment needed to achieve the HA/DR/BC solution are owned by the enterprise itself. Cascaded FICON allows for considerable cost savings compared with ESCON when insourcing HA/DR/BC.

14 Years of Evolution

In 1990 the ESCON channel architecture was introduced to the world as the way to address the limitations of parallel (bus and tag) architectures. As such, ESCON provided noticeable, measurable improvements in distance capabilities, switching topologies and, most importantly, response time and service time performance. By the end of the 1990s, ESCON's strengths over parallel channels had become its weaknesses. FICON evolved in the late 1990s to address the technical limitations of ESCON in bandwidth, distances and channel/device addressing.

Initially, the FICON (FC-SB-2) architecture did not allow the connection of multiple FICON directors. Of course, neither does ESCON except when static connections of "chained" ESCON directors were used to extend ESCON distances. Both ESCON and FICON defined a single byte for the link address, the link address being the port attached to "this" director. As of 31 January 2003, this changed. Now, it is possible to have two-director configurations, with it also being possible to have separate geographic sites. This is done by adding the domain field of the fibre channel destination ID to the link address in order to specify the exiting director and the link address on that director.

What is Cascaded FICON?

Cascaded FICON refers to an implementation of FICON that involves one or more FICON channel paths to be defined over 2 FICON directors that are connected to each other using an Inter-Switch Link (ISL). [TROW02] The processor interface is connected to one director, while the storage interface is connected to the other. This configuration is supported for both disk and tape, with multiple processors, disk subsystems and tape subsystems sharing the ISLs between the directors. Multiple ISLs between the directors is also supported.

There are some hardware and software requirements specific to cascaded FICON [NEVI05]:

1) The FICON directors themselves must be from the same vendor (i.e., both should be from either Brocade, Cisco, or McDATA)

2) The mainframes must be z Series machines: z800, 890, 900, or 990. Cascaded FICON requires 64-bit architecture to support the two-byte addressing scheme. Cascaded FICON is, therefore, not supported on 9672 G5/G6 mainframes.

3) z/OS version 1.4 or greater, and/or z/OS version 1.3 with required PTFs/MCLs to support 2-byte link addressing (DRV3g and MCL (J11206) or later).

4) The high integrity fabric feature for the FICON director must be installed on all directors involved in the cascaded architecture.

What are the Benefits?

Cascaded FICON delivers many of the same benefits of open systems SANs to the mainframe space. Cascaded FICON allows for simpler infrastructure management, lowered infrastructure cost of ownership, and higher data availability. This higher data availability is important in delivering a more robust enterprise disaster recovery strategy. Further benefits are realized when the ISLs connect directors in two or more locations and/or are extended over long distances. Figure 1 shows a non-cascaded two-site environment.

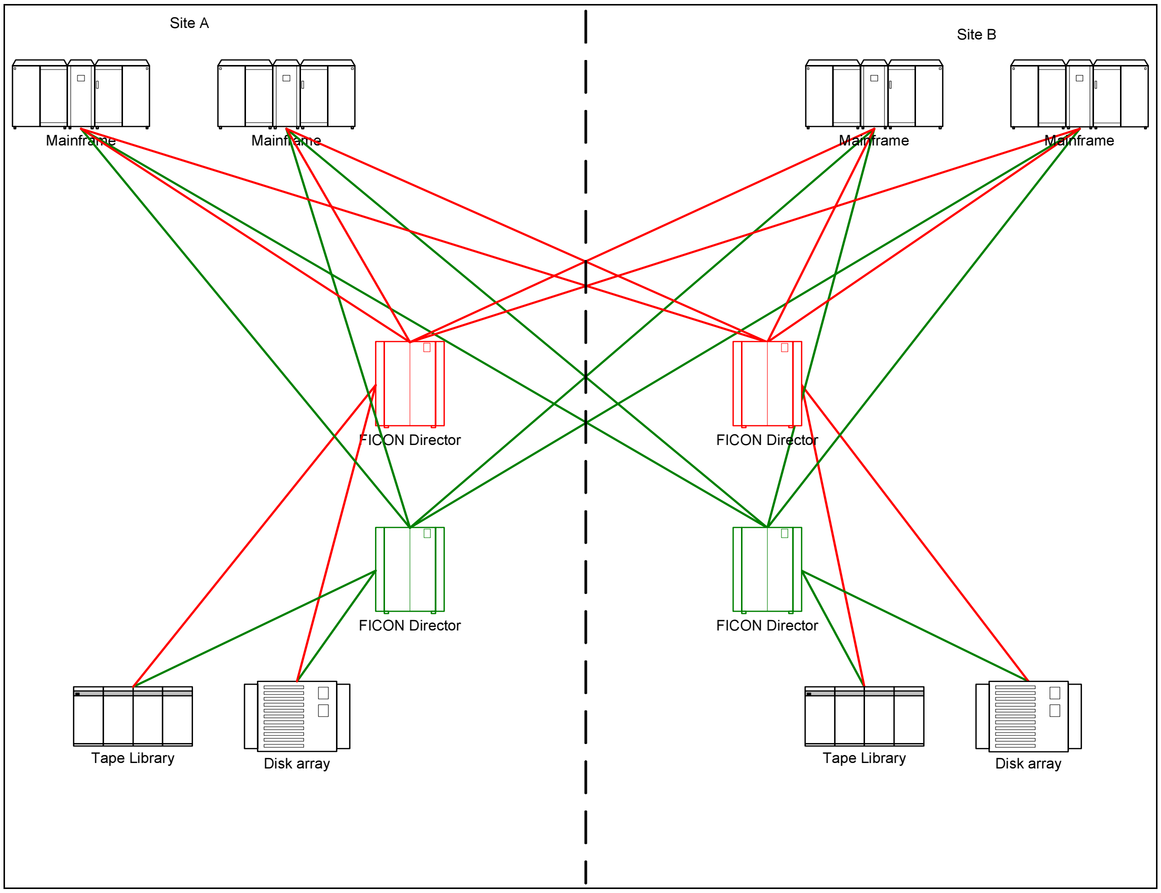

Figure 1 – Two-Site, Non-cascaded FICON environment

In the configuration above, all hosts have access to all of the disk and tape subsystems at both locations. The host channels at one location are extended to the FICON directors at the other location to allow for cross-site storage access. If each line represents two FICON channels, then the above configuration would need a total of sixteen (16) extended links. These links would only be utilized to the extent that the host has activity to the remote devices.

The most obvious benefit when comparing the Figure 1 configuration with one that is cascaded is the reduction in the number of links across the WAN. Figure 2 shows a cascaded, two-site FICON environment.

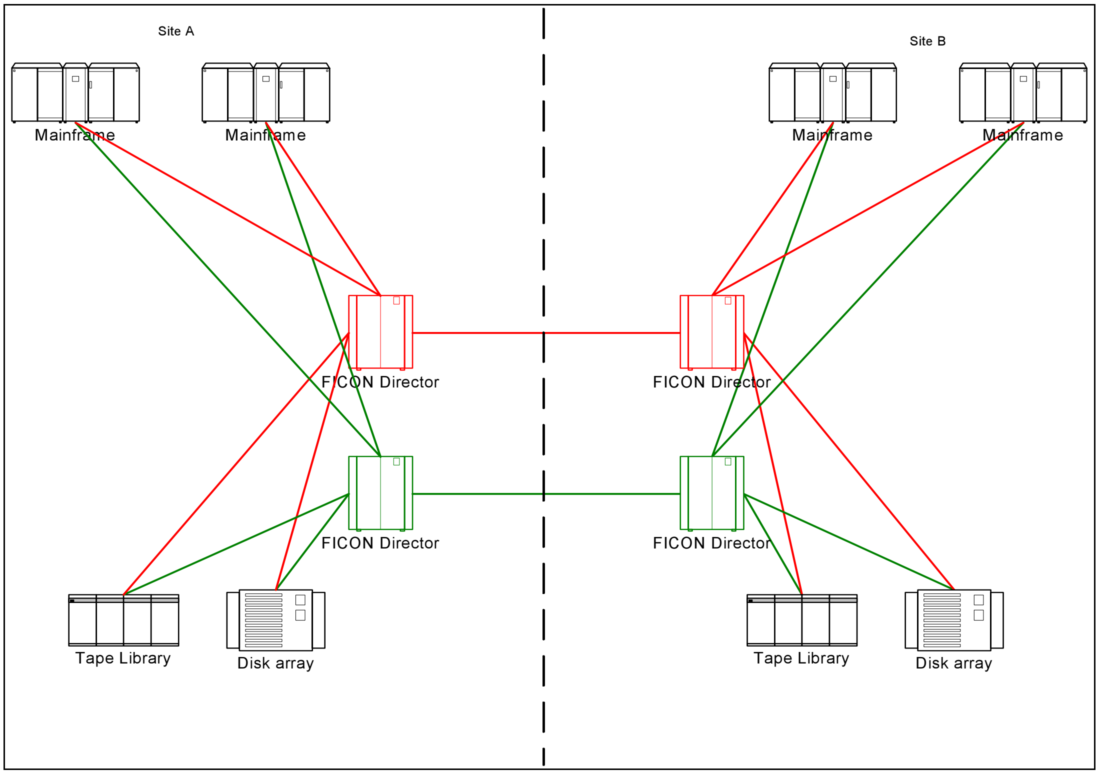

Figure 2 – Two-Site, Cascaded FICON environment

In this configuration, if each line represents two channels, only four (4) extended links are required. Since FICON is a packet-switched protocol (as opposed to the circuit-switched ESCON protocol), multiple devices can share the ISLs, and multiple I/Os can be processed across the ISLs at the same time. This allows for the reduction of links between sites and allows for more efficient utilization of the links in place. In addition, ISLs can be added as the environment grows and traffic patterns dictate.

This is the key way in which a cascaded FICON implementation can reduce the cost of the enterprise architecture. As can be seen in both Figure 1 and Figure 2, the cabling schema for both intersite and intrasite has been simplified. Fewer intrasite cables translate into lower cabling hardware and management costs. It also reduces the number of FICON adapters, director ports, and host channel card ports required, thus lowering the connectivity cost for mainframes and storage devices as well. In Figure 2, the sharing of links between the two sites reduces the number of physical channels going across between sites thereby lowering the cost by consolidating channels and the number of director ports. By the way, the faster the channel speeds are between sites, the better the intersite cost savings from this consolidation will be. So, as 4 Gbps FICON and 10Gbps FICON become available, the more attractive this becomes.

Another benefit to this approach, especially over long distances, is that the FICON director typically has many more buffer-credits per port than do the processor and the disk/tape subsystem cards. More buffer-credits allow for a link to be extended farther distances without significantly impacting response times to the host.

In their March 2003 white paper on Cascaded FICON director performance considerations, [CRON03] Cronin and Bassener listed 7 main factors affecting the performance of a cascaded FICON director configuration:

1. The number of ISLs between the two cascaded FICON directors and the routing of traffic across ISLs.
2. The number of FICON/FICON Express channels whose traffic is being routed across the ISLs.
3. The ISL link speed.
4. Contention for director ports associated with the ISLs.

5. The nature of the I/O workload (I/O rates, block sizes, use of data chaining, and read/write ratio).
6. The distances of the paths between the components of the configuration (the FICON channel links from processor(s) to the first director, the ISLs between directors, and the links from the second director to the storage control unit ports).
7. The number of switch port buffer to buffer credits.

The last point –the number of buffer to buffer credits-- is typically the one examined most carefully, as well as the one that is most often misunderstood. As such, it deserves more detailed attention.

Basics of Fibre Channel Flow Control and Buffer Credit Theory

The traditional goal of flow control is to prevent a transmitter from overrunning a receiver by providing real-time signals back from the receiver to pace the transmitter. Where the receiver is distant from the transmitter, the pacing signals back to the transmitter are so delayed that overrun and low performance can occur. Credit based flow control is a technique used in fibre channel (and therefore in FICON) that can effectively prevent receiver overrun in long distance circuits while still allowing for high transmitter activity.

The basic information carrier in Fibre Channel (and, hence, FICON) is the frame. Other than ordered sets, which are used for communicating low-level link conditions, all information is contained within frames. When discussing the concept of frames, a good analogy to use is that of an envelope: When sending a letter via the United States Postal Service, the letter is encapsulated within an envelope. When sending data via a FICON network, the data is encapsulated within a frame. Fortunately, FICON networks have far better service times than the USPS.

To prevent a target device (either host or storage) from being overwhelmed with frames, the Fibre Channel architecture provides flow control mechanisms based on a system of credits. Each of these credits represents the ability of the device to accept an additional frame(s). If a recipient issues no credits to the sender, no frames can be sent. Pacing the transport of subsequent frames on the basis of this credit system helps prevent the loss of frames and reduces the frequency of entire fibre channel sequences needing to be retransmitted across the link.

Upon arrival at a receiver, a frame goes through several steps. It is received, deserialized, decoded, and stored in a receive buffer where it is processed by the receiving port. If another frame arrives while the receiver is processing the first frame, a second receive buffer is needed to hold this new frame. Unless the receiver is capable of processing frames as fast as the transmitter is capable of sending them, it is possible for all of the receive buffers to fill up with received frames. At this point, if the transmitter should send another frame, the receiver will not have a receive buffer available and the frame will be lost. To prevent this type of error condition, the Fibre Channel architecture provides a two level flow control mechanism that allows the receiver to control when the transmitter may send frames. The receiving port controls the frame transmission by giving the sending port permission to send one or more frames to the receiving port in question. This permission is called a credit. The actual credit(s) are granted during the login process between two ports. The credit value is decremented when a frame is sent and replenished when a response is received. If the available credits for a given port reaches zero, the supply of credits is said to be exhausted. Further transmission of frames with that port is then suspended until the amount of credits can be replenished to a non-zero value. A good analogy would be a pre-paid calling card: there are a certain amount of minutes to use, and one can talk until there is no more time (minutes) on the card.

One of the goals of Fibre Channel (and hence FICON) is to provide reliable delivery of information from sender to receiver. Providing a data link with a low bit-error rate is a good start, but simply minimizing the quantity of bit-level transmission errors is not enough. We need to guarantee/ensure consistent and reliable frame delivery. Flow control is one of the primary mechanisms for providing this reliability. There are two types of flow control mechanisms in FICON/Fibre Channel; that are used. The first is End-to-End Flow Control. The second, which will be the focus for the remainder of this paper, is called Buffer-to-Buffer Flow Control.

End-to-End Flow Control

Transmission credit is initially established when two communicating nodes log in and exchange their respective communication parameters. End-to-End Flow Control, also known as EE Credit, is used by Class 1 and Class 2 service between two end nodes. The nodes monitor end-to-end flow control themselves. Intervening switches or directors do not participate in EE Credit. As data is sent from the sending port to the destination port, the sender subtracts a credit from its pool of end-to-end credits. Next, when the destination port receives the transmission, it sends an acknowledgement control word (ACK) back to the sender indicating that the frame was received. When this acknowledgement is received back at the sending port, it then adds a credit back to its own credit pool. Therefore, end-to-end credits used by the sending port are replenished when it receives the acknowledgement from the destination port. What we need to keep in mind is that End-to-End Flow Control is always managed between a specific pair of node ports. Therefore, an individual node port may have many different end-to-end credit values, each corresponding to a different destination node port.

Buffer to Buffer Flow Control

In contrast, Buffer-to-Buffer Flow Control is flow control between two optically adjacent ports in the I/O path, i.e., transmission control over individual network links. This is true for both N_Port (the port on the host or device) to F_Port (the port on the FICON director connected to an N_Port) and E_Port to E_Port (ports on the FICON directors connected to a port on another FICON director via an interswitch link, or ISL) adjacency. A separate, independent pool of credits is used to manage Buffer-to-Buffer Flow Control. Similar to what was earlier discussed for End-to-End Flow Control, Buffer-to-Buffer Flow Control works by a sending port using its available credit supply and waiting to have the credits replenished by the port on the opposite end of the link. These Buffer-to-Buffer Credits (BB Credits) are used by Class 2 and Class 3 service and rely on the fibre channel receiver-ready (R_RDY) control word to be sent by the receiving link port to the sender. An end node attached to a FICON director will establish its BB Credit during login to the fabric. A communicating partner attached elsewhere on the FICON director will establish its own and most likely different BB Credit value to the director during its login process. This value is used by the transmitter to track the consumption of receive buffers and pace transmissions if necessary. A typical FICON director would track the consumption of BB Credits as follows:

- Before any data frames are sent, the transmitter sets a counter equal to the BB-Credit value.
- For each data frame sent by the transmitter, the counter is decremented by one.
- Upon receipt of a data frame, the receiver sends a status frame (R_RDY) to the transmitter indicating that the data frame was received AND the buffer can receive another data frame.
- For each R_RDY received by the transmitter, the counter is incremented by one.

As long as the transmitter count is a non-zero value, the transmitter is free to continue sending data. This mechanism allows for the transmitter to have a maximum number of data frames in transit equal to the value of BB_Credit, with an inspection of the transmitter counter indicating the number of receive buffers. Hence, BB Credit has no end to end component. The sender decrements the BB Credit by 1 for each R_RDY received. The initial value of BB Credit must be non-zero. The rate of frame transmission is regulated by the receiving port based on the availability of buffers to hold received frames. [NEVI05]

It should be noted that the FC-FS specification allows the transmitter to be initialized at zero, or at the value BB_Credit and either count up or down on frame transmit. Different switch/director vendors may handle this with either method, and the counting would be handled accordingly.

At first glance, it is readily apparent that this system may leave something to be desired in terms of overall performance and efficiency. This is due to the time required for frames to travel from the sending port to the receiving port and responses to return from the receiving port back to the sending port. Now, consider that it takes light approximately 5 nsec to propagate through 1 meter of optical fiber, or 50 microseconds to travel 10km. This behavior becomes even less efficient and more of a performance drag on faster links, longer distance links, or when traveling through complex topologies that contribute significant delivery latencies. So, to achieve the higher performance while preventing the overrun of receive buffers, we need to use BB Credit values>1 and frame streaming. If a sending port is allowed to send more than one frame without having to wait for a response to each, performance can be improved. This is referred to as frame streaming. As more credits are made available, link utilization (and performance) will increase until link utilization reaches 100%. When the link is thus fully utilized, frames can be sent as rapidly as allowed but additional credits will not help matters. So, the key questions become what are the implications to asset deployment, how many frames can be transmitted by the sending port prior to a response to the first frame being received from the receiving port and, most importantly, what is the optimal amount of credits?

Implications to Asset Deployment

There are four implications to asset deployment to consider:

1. For write intensive applications across an ISL (tape and disk replication) the BB_Credit value advertised by the E_Port on the target side gates performance. In other words, the number of BB Credits on the target side cascaded FICON director is the major factor.
2. For read intensive applications across an ISL (regular transactions) the BB_Credit value advertised by the E_Port on the host side gates performance. In other words, the number of BB Credits at the local location is the major factor.
3. Two ports do not negotiate BB_Credit down to the lowest common value. A receiver simply "advertises" BB_Credits to a linked transmitter.
4. The depletion of BB_Credits at any point between an initiator and a target will gate overall throughput.

What is the optimal number of BB Credits?

Kembel states that the optimal amount of credits is determined by the distance (frame delivery time), the processing time at the receiving port; link signaling rate, and the size of the frames being transmitted. He developed a formula to determine the optimal amount of credits: [KEMB02]

Credit=(Round_trip_time+Receivingport_processing _time) /Frame Transmission _time.

In other words, the optimal number of BB credits depends on 3 key parameters:

1) round trip time, i.e., the distance
2) frame processing time
3) frame transmission time*

*As the link speed increases, the frame transmission time is reduced; therefore, as we get fast iterations of FICON such as FICON Express and FICON Express 2, the amount of credits need to be increased to obtain full link utilization, even in a short distance environment!

Consider this in another way. Assuming the speed of light in a fiber optic cable (in a non-vacuum) is 200,000,000 meters/second, it requires 500 microseconds to send a bit of information 100 km. Thus, a 100 km link can contain 62,500 bytes by the time the first bit is read at the other end (or 53,125 bytes when accounting for the FC 8-bit/10-bit encoding at the FC-1 layer). In addition, for each frame that is transferred, the hardware at the other end must acknowledge that the frame has been received before a successful transmission occurs. This requires enough capacity in the hardware to allow continuous transmission of frames on the link, while waiting for the acknowledgement to be sent by the receiver at the other end. Therefore, in order to maintain 100% utilization of a 1 Gb link for 100 km, the sending director hardware must have enough resources (BB credits) to keep 106.250 bytes on the link and the receiving switch hardware must have enough resources to allow the sender to transmit continuously. Recall earlier discussions of the BB Credit mechanism/BB Flow Control. To theoretically achieve 100% utilization of a 1 Gbps link for 100 km, the required BB Credits range from 49 to 1155 depending on the average frame size. When the link speed is increased to 2 Gbps, the required BB Credits range from 98 to 2310. This assumption is that the smallest frame size is 68 bytes (36 bytes of Fibre Channel Header plus 32 bytes of command information for a FICON FC-SB2 command IU) and the largest frame size is 2148 bytes (36 bytes of Fibre Channel Header plus 2112 bytes of data).

Typical environments generally do not sustain one pattern for the transfer of data. Normally, reads and writes are mixed and the data block sizes vary. To get an idea of the resources necessary to support long distances in a FICON environment, the data transfers can be expressed in terms of average block sizes. Tables 1 and 2 characterize the BB Credit requirements for 1 Gbps and 2 Gbps links at 100 km based on different average block sizes for data transfers.

Typical Block Size	Frames per Block	Average Frame Size	Number of Buffers Requried
1024	3	341	291
2048	4	512	198
4096	5	819	126
8192	7	1170	89
16384	11	1489	70
32768	19	1725	61
65536	35	1872	56
131072	67	1956	54
262144	131	2001	52

Table 1-Theoretical 1 Gbps utilization at 100 km (these values are close approximations and are intended to convey the approximate Buffer-to-Buffer Credit requirements).

Typical Block Size	Frames per Block	Average Frame Size	Number of Buffers Requried
1024	3	341	582
2048	4	512	396
4096	5	819	252
8192	7	1170	178
16384	11	1489	140
32768	19	1725	121
65536	35	1872	112
131072	67	1956	107
262144	131	2001	105

Table 2-Theoretical 2 Gbps utilization at 100Km (these values are close approximations and are intended to convey the approximate buffer to buffer credit requirements.)

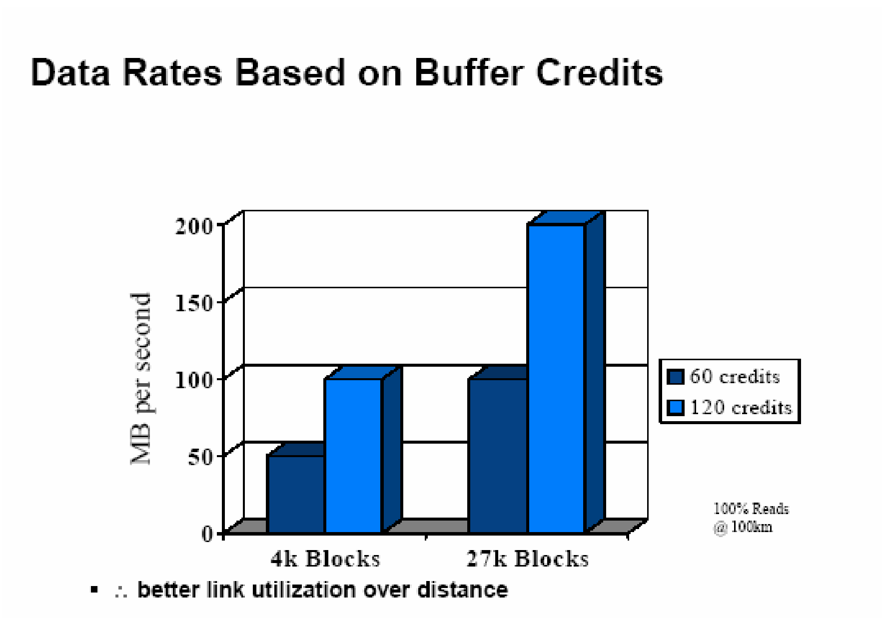

Figure 3. Data rates based on buffer to buffer credits

Figure 3 represents data from an IBM White paper on Cascaded FICON performance considerations written by Cronin and Basener. This graph shows that at 100 km distance between 2 nodes, having double the number of buffer credits yields double the effective data regardless of block size. Figure 4 below shows a similar type of result specifically for cascaded FICON.

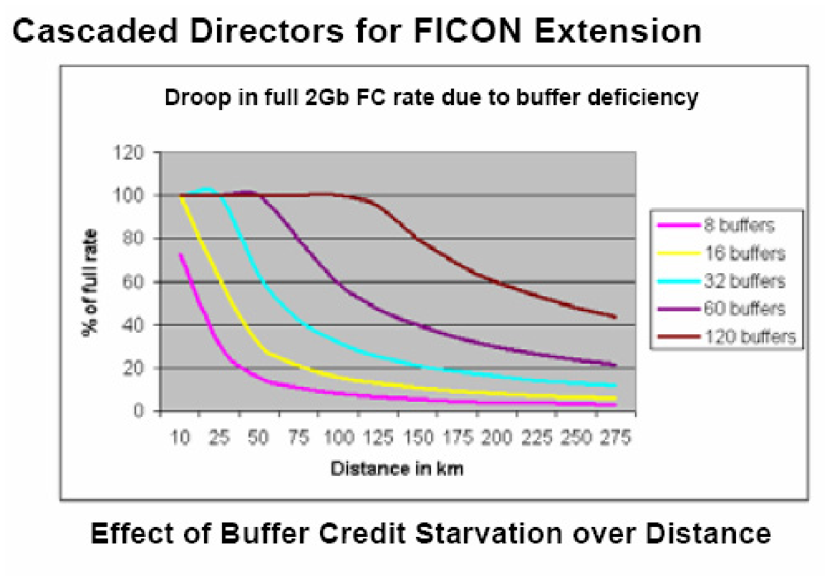

Figure 4

As companies bring DR back in house via cascaded FICON, distance becomes more important and, therefore, the amount of BB Credits present on FICON directors has become much more important in the past two years.

Therefore, there is an optimal number of BB credits for a given distance, data characteristic (block size), and link speed. While at first glance, it may appear that the more BB credits available the better, there is clearly an optimal number that can be defined mathematically using the variables cited above. This is quite analogous to DASD and cache sizing. Studies have been done to mathematically prove that there is an optimal amount of DASD cache, i.e., there is a law of diminishing marginal returns with cache size, to the point of where adding more does not improve performance. While the option will often be provided by a FICON switch/director to use more than this optimal number of BB credits, much like with exceeding optimal cache size, exceeding the optimal number of BB credits does not improve performance and is not optimal to an end-users economic situation: paying for something not truly needed.

In a nutshell it means that cascaded FICON configurations with the optimal number of BB Credits allows for even higher availability, disaster recovery, and business continuity (HA/DR/BC).

Technical Discussion on FICON Cascading

The basics

First, as stated earlier, cascaded FICON is limited to z Series mainframes only. Please review the HW/SW requirements outlined earlier in this paper. For more details on some of the fibre channel terminology that is in this section, please refer to the key references listed in the bibliography.

Referring back to Figure 2, observe that a cascaded FICON director configuration involves at least three fibre channel links. The first link is between the FICON channel card on the mainframe (known as an N_Port) and the FICON director's fibre channel adapter card (which is considered an F_Port). The second link is between the two FICON directors via what are known as E_Ports. The link between E_Ports on the directors is known as an inter-switch link, or ISL. The final link is from the F_Port to a FICON adapter card in the control unit port (N_Port) of the storage device. The physical paths are the actual fibre channel links connected by the FICON directors providing the physical transmission path between a channel and a control unit. Please also note that the links between the cascaded FICON directors may be multiple ISLs, both for redundancy and to ensure adequate I/O bandwidth.

Addressing Support

Single byte addressing refers to the link address definition in the Input-Output Configuration Program (IOCP). Two-byte addressing (cascading) allows IOCP to specify link addresses for any number of domains by including the domain address with the link address. This allows the FICON configuration the capability of creating definitions in IOCP that span more than one director.

Figure 5 shows that the FC-FS 24 bit FC port address identifier is divided into three fields:

1) Domain
2) Area
3) AL Port

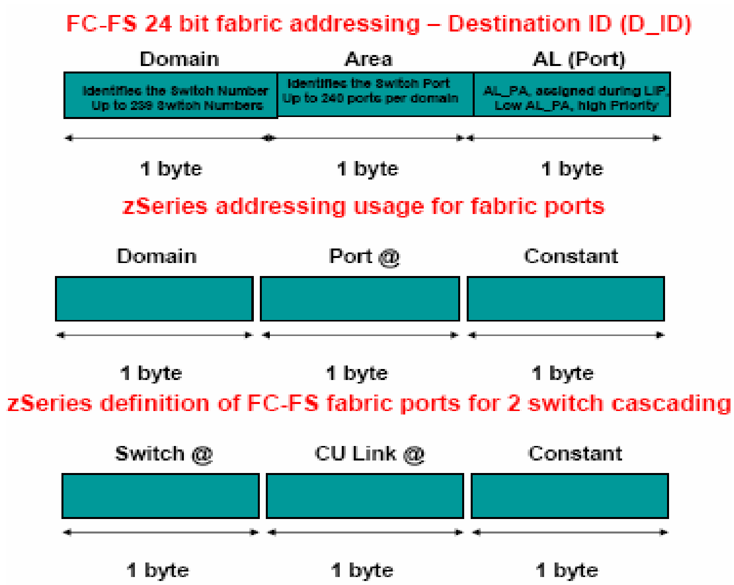

Figure 5. Fabric Addressing Support (1)

In a cascaded FICON environment, 16 bits of the 24-bit address must be defined for the zSeries server to access a FICON control unit. The FICON directors provide the remaining byte used to make up the full 3-byte FC port address of the CU being accessed. The AL_Port (arbitrated loop) value is not used in FICON and will be set to a constant value. The zSeries "domain and "area" fields are referred to as the F_Port's "port address" field. It is a 2-byte value, and when defining access to a CU that is attached to this port using the zSeries Hardware Configuration Definition (HCD) or IOCP, the port address is referred to as the link address. Figure 6 further illustrates this. Figure 7 is an example of a cascaded FICON IOCP gen.

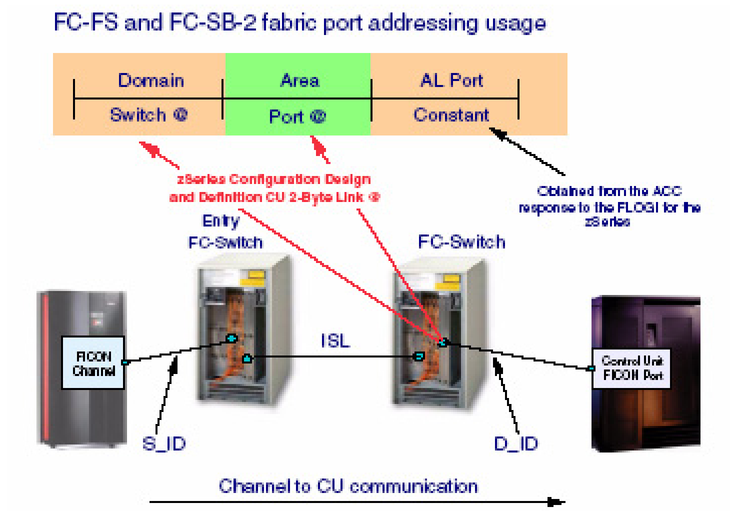

Figure 6. Fabric Addressing Support (2)

The connections between the two directors are established through the exchange of link parameters (ELP). The directors then pause for a device fabric login (FLOGI), and then, assuming the device is another switch they initiate an ELP exchange. This results in the formation of the ISL connection(s).

In a cascaded FICON configuration, three additional steps occur beyond the "normal" FICON switched point-to-point communication initialization. A much more detailed discussion of the entire FICON initialization procedure can be found in Chapter 3 of the IBM Redbook: FICON Native Implementation and Reference Guide, pp 23-43. [TROW02] A hyperlink is given in the references. The 3 steps are:

1) If a 2-byte link address is found in the CU macro in IOCDS, a "Query Security Attribute" (QSA) command is sent by the host to check with the fabric controller on the directors if the directors have the high integrity fabric features installed.
2) The director responds to the QSA.
3) If it is an affirmative response, indicating that a high integrity fabric is present (fabric binding and insistent domain ID) the login continues. If not, login stops and the ISLs are treated as invalid (not a good thing).

```
ID (no change in ID statement for FICON)
RESOURCE PARTITION=((CSS(0),(SYSA,1),(SYSB,2),(SYSC,3)),
      (CSS(1),(SYSD,1),(SYSE,2),(SYSF,3)))

CHPID PATH=(CSS(0),80),SHARED,PART=((CSS(0),(SYSA,SYSB),(=))),TYPE=FC,SWITCH=61,PCHID=160
CHPID PATH=(CSS(0),81),PART=((CSS(0),(SYSB),(=))),TYPE=FC,SWITCH=61,PCHID=1A0
CHPID PATH=(CSS(0),82),PART=((CSS(0),(SYSB),(=))),TYPE=FC,SWITCH=61,PCHID=130
CHPID PATH=(CSS(0,1),83),SHARED,PART=((CSS(0),(SYSC),(=)),(CSS(1),(SYSD),(=))),
      TYPE=FC,SWITCH=61,PCHID=1D0

CNTLUNIT CUNUMBR=8000,PATH=((CSS(0),80,81,82,83),(CSS(1),83)),
      UNITADD=((00,256)),LINK=((CSS(0)6212,6222,6232,6242),(CSS(1),6242)),
      CUADD=0,UNIT=2105
CNTLUNIT CUNUMBR=8100,PATH=((CSS(0),80,81,82,83),(CSS(1),83)),
      UNITADD=((00,256)),LINK=((CSS(0)6212,6222,6232,6242),(CSS(1),6242)),
      CUADD=1,UNIT=2105
        *
        *
CNTLUNIT CUNUMBR=8700,PATH=((CSS(0),80,81,82,83),(CSS(1),83)),
UNITADD=((00,256)),LINK=((CSS(0)6212,6222,6232,6242),(CSS(1),6242)),
      CUADD=7,UNIT=2105

IODEVICE  (no change for FICON)
```

Figure 7. Sample IOCP gen for Cascaded FICON (z890/z990)

High Integrity Enterprise Fabrics

Data integrity is paramount in a mainframe environment. End to end data integrity absolutely must be maintained throughout a cascaded FICON director environment. Why? We must ensure that any changes made to the customer's data stream are always detected and that the data is always delivered to the correct end point. What does high integrity fabric architecture and support entail?

1) Support of Insistent Domain IDs. This means that a FICON director will not be allowed to automatically change its address when a duplicate switch address is added to the enterprise fabric. Intentional manual operator action is required to change a FICON director's address. Insistent Domain Ids prohibit the use of dynamic Domain IDs, thereby ensuring that predictable Domain IDs are being enforced within the fabric. It also makes certain that duplicate Domain IDs are not used within the fabric.

2) Fabric Binding. Fabric binding enables companies to allow only FICON directors that are configured to support high-integrity fabrics to be added to the storage/FICON network. The FICON directors that you wish to connect to the fabric must be added to the fabric membership list of the directors already in the fabric. This membership list is composed of the "acceptable" FICON director's Worldwide Name (WWN) and Domain ID. Using the Domain ID ensures that there will be no address conflicts, i.e. duplicate domain IDs when the fabrics are merged. The two connected FICON directors then exchange their membership list. This membership list is a Switch Fabric Internal Link Service (SW_ILS) function, which ensures a consistent and unified behavior across all potential fabric access points.

Applications for Cascaded FICON directors

Any application or environment that requires cross-site FICON or fibre channel connectivity for multiple channels can benefit from using cascaded directors. Those that come to mind first are Geographically Dispersed Parallel Sysplex (GDPS) and array-based replication. Others include remote tape vaulting, where the backup application writes directly to tape drives at the remote location.

Conclusion

The evolution of FICON to support cascading is a clear-cut example of a protocol that was designed to fill ever-changing requirements. FICON is the basis on which mainframe storage networks will be built well into the future. As an upper level layer in the Fibre Channel standard, FICON will continue to be around, most likely evolving into a shared mainframe/open systems storage network.

Bibliography

[ARTI05] Dr. H. Pat Artis and Robert Ross. *Managing Complex FICON Configurations.* 2005. www.perfassoc.com

[BENE01] Alan F. Bener. *Fibre Channel for SANs.* McGraw-Hill (2001). ISBN 0-07-137413-2.

[CRON03] Catherine Cronin and Richard Basener. *Performance Considerations for a Cascaded FICON Director Environment Version 0.2x (March 2003).* www.ibm.com

[KEMB02] Robert Kembel. *Fibre Channel: A Comprehensive Introduction.* Northwest Learning Associates (2002). ISBN 0-931836-84-0

[NEVI05] Iain Neville, Bill White & Hans-Peter Eckham. *FICON Implementation Guide.* IBM Redbooks (2005). SG24-6497. www.ibm.com

[SCHU04] Greg Schulz. *Resilient Storage Networks.* Elsevier (2004). ISBN 1-55558-311-3.

[SIMI03] Huseyin Simitci. *Storage Network Performance Analysis.* Wiley Publishing (2003). ISBN 0-7645-1685-X.

[TROW02] Ken Trowell, Bill White et.al. *FICON Native Implementation and Reference Guide.* IBM Redbooks (2002). SG24-6266-01