February 2002

# FICON and FICON Express Channel Performance Version 1.0

*Cathy Cronin*
*zSeries I/O Performance*
*Catherine Cronin/Poughkeepsie/IBM@IBMUS*
*8-295-8200 or (845) 435-8200*

# *Table of Contents*

## Introduction

This white paper has been developed to help IBM® field sales specialists and technical representatives understand the performance characteristics of the different versions of FICON™ and the new FICON Express™ channels that are now available to our customers. IBM has made significant improvements to FICON channels since this product was initially shipped in 1999. The following chart depicts some of those improvements:

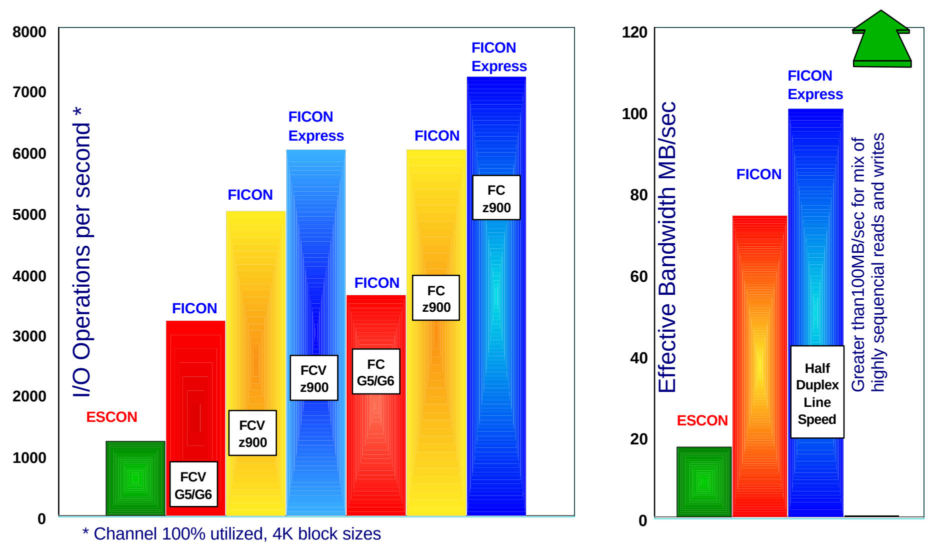


Figure 1

Reflected in the left bar chart is the "best can do" capabilities of each of the FICON channels in both FCV or FICON bridge and FC or native FICON mode measured using an I/O driver measurement program for 4K byte read hits. 4K bytes is the size of most online database I/O operations. These are the maximum possible or 100% channel utilization 4K I/O rates possible for each channel. Normally customers should keep their channels at 50% or less channel utilization to achieve good online transaction response times.

Reflected in the right bar chart is the "best can do" capabilities of each of the FICON channels in FC mode measured using an I/O driver measurement program for 6x27K or 6 half-track reads and writes. This is representative of the type of channel programs used in disk to tape backup jobs or other highly sequential batch jobs.

As you can see, the IBM ~ zSeries™ 900 FICON Express represents a significant improvement in both 4K I/O per second throughput and maximum bandwidth capability compared to ESCON® and our original IBM S/390® G5/G6 FICON offering.

This paper will explain in more detail the benefits of FICON and FICON Express channels in both FC and FCV mode.

## Benefits of FICON and FICON Express

FICON channels have many benefits over the traditional ESCON and Parallel channels. These benefits include:

- *Increased Data Transfer Rate: The data transfer rate across the link is limited to 4.5 MB/sec on a Parallel channel and 17 MB/sec on an ESCON channel. The FICON link data rate during data transfer is 100 MB/sec. The new FICON Express channel can sustain this "line speed" rate for large data transfers doing all reads or all writes and FICON Express can deliver greater than 100 MB/sec for a mix of large read and write data transfers. One must however, keep in mind that the data transfer time for any application is some fraction of the total job elapsed time and therefore the bottom line benefit of this improvement will vary by workload.*
- *Improved Performance: Native FICON attachment offers much higher data transfer rates than ESCON attachments. For critical data-intensive applications, this can translate into significant improvements in overall elapsed time. For example, QSAM workloads (typical of batch) and large non-parallel DB2® queries (typical of data warehousing and data extracts) can run over two times faster with FICON than with ESCON.*
- *Reduced Backup Windows: Working together, IBM's FICON-attached Enterprise Storage Server™ and 3590-A60 Magstar® tape system can slash elapsed times for backup operations by up to half (depending on data compressibility).*
- *Increased Distance: With traditional parallel channels, the distance from a host processor (CEC) to the Control Units (CUs) was limited to 400 feet. This distance limitation was eliminated with the introduction of the ESCON channels in 1990 where the distance was up to 3 km without switches or repeaters and up to 43 km with switches or repeaters although the data rate dropped rapidly after 9km. FICON channels now support distances up to 20 km without repeaters and 100 km with repeaters without data rate droop.*
- *Channel Aggregation: A single FICON channel can replace multiple ESCON channels. This means fewer channels, director ports and control unit ports or a simpler configuration to manage.*
- *Combined CTC and Channel function: With zSeries 900 GA2, a single FICON CHPID can be simultaneously used to talk to another FICON CHPID and any other native FICON controller for example, DASD, TAPE or Printer.*
- *Reduced Infrastructure: The infrastructure costs will be reduced by the aggregation of multiple ESCON channel fibers onto a single FICON fiber. This will be most significant when connecting host systems to remote locations either across the street, across town, or even further.*

- *Increased Addressability: The current addressing limitations for ESCON are 1024 Unit Addresses (UA) per channel. With FICON, the addressability is increased to 16,384 per FICON Channel. This increase can help customers that are currently constrained either by the number of CUs or number of devices attached to their systems.*

### Simpler Configurations with native FICON Channels

FICON running in FC mode with native FICON attachments offers the opportunity to substantially consolidate and simplify path configurations, particularly the complex configurations found in large z/OS™ environments. This is due to the much higher capabilities on attached volume addresses, the multiplexing capability of FICON channels, directors and attachments and the sharply increased throughput capability of FICON and FICON Express as implemented on the IBM ~ zSeries 900 server.

An example helps to clarify the opportunities that exist when configuring storage for FICON attachment. Consider, as a starting point, an existing configuration in which each of two disk storage systems are connected to each of two hosts, both running on separate zSeries 900 servers. Each storage system is attached to ESCON directors via 16 paths; each host, in turn, has 16 paths set aside for accessing each storage system, via the corresponding directors. Figures 2 and 3 illustrate this multi-system ESCON configuration for each ESS.

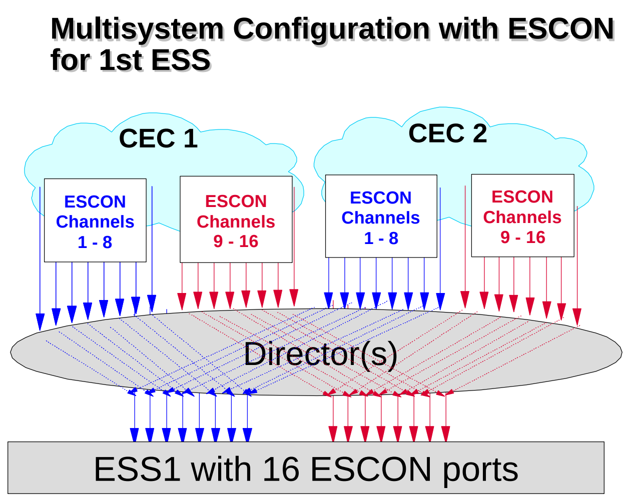


Figure 2

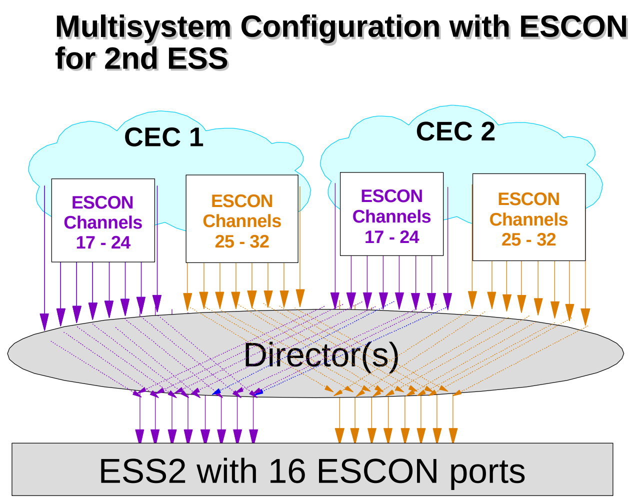


Figure 3

It is important to note that, although each storage system in such a configuration has only 16 paths into it, there are 32 host paths dedicated for access to that system. If perfect balance of the I/O load could be achieved between the two hosts, it would, in principle, be possible to reduce to 8 the number of ESCON paths from each host to the director accessing a given storage system. Since perfect balance cannot generally be achieved however, it is often impractical to reduce ESCON paths in this way. Configurations of the kind just outlined are one reason that ESCON channel path utilizations tends to be low in large z/OS environments. Since the utilization of the ESCON storage subsystem port generally needs to stay below 50% to achieve good response times in an online database transaction environment and this utilization can be approximated by summing up the utilizations of the individual ESCON channels that are connected to it, one can see how this leads to low ESCON channel utilizations on each CEC in this environment. The fact that ESCON directors are circuit-switched and that each ESCON storage subsystem port can only execute one I/O operation at a time also contributes to the over configuration and low utilization of ESCON channels in large z/OS multi-system environments.

Naturally, it is also possible to reproduce the identical scheme with FICON or FICON Express. In this case, each of two ESSs are attached to a FICON director via 4 paths; each host, in turn, has 4 paths to the director set aside for accessing that storage system. This FICON configuration offers a four times reduction in the number of host-to-director and director-to-storage-system paths. With FICON, we can also refine the configuration just outlined, by allowing resources to be shared. Rather than dedicating 4 paths from each host to access each ESS, a better configuration would be to provide 8 paths from each host, all of which access both ESSs. In this way, any host channel can access any storage volume, on either ESS.

Nevertheless, this refined configuration still reflects the same "overkill" seen previously with ESCON. Although the two ESSs, taken together, are accessed via 8 FICON paths, these connect to a total of 16 FICON or FICON Express channels. With FICON, however, there is a practical way to further consolidate some of the 16 channels. It is not necessary to assume perfect balance of the I/O load in order to accomplish such a consolidation on a zSeries 900 server. As discussed later in this paper, the performance capability of this server's FICON Express channel is sufficient so that even with a significant imbalance of I/O load, it is still possible with many workloads to service the 8 FICON storage paths using 8 FICON Express channels (4 from each host). In this way, a total of 64 ESCON channels have been reduced to a total of 8 FICON Express channels, while also simplifying connectivity since every channel connects to all of the accessible storage volumes. This example demonstrates the fact that the N to 1 reduction achieved by the channels is not necessarily the same as the N to 1 reduction achieved by the storage subsystem host adapters when migrating from ESCON to FICON. In this example, the channels were reduced from 64 to 8 or by an 8 to 1 factor while the adapters on each of the storage subsystems were reduced from 16 ESCON to 4 FICON or a 4 to 1 reduction. Figure 4 illustrates this much simpler multi-system FICON configuration to 2 ESSs which is possible for those workloads that meet the consolidation criteria discussed in the next section of this paper.

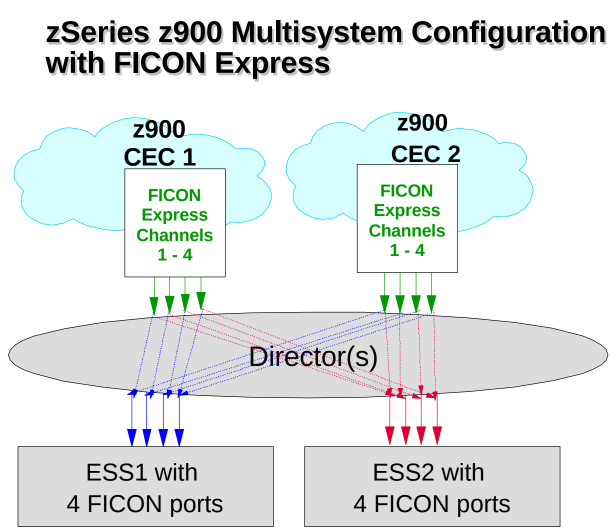


Figure 4

### Channel Consolidation with native FICON attachments

This section discusses the features and capabilities of native FICON (FC) channels relevant to the task of consolidating work from many ESCON channels to fewer FICON or FICON Express channels. As the example of the previous section illustrates, the assessment of how many ESCON channels can be replaced by a single FICON or FICON Express channel in a particular customer configuration should include the following factors: ESCON channel utilizations, I/O activity rates, data transfer sizes, storage system cache hit ratios, the number of channels connected to a single control unit (CU) port, the number of CU ports daisy-chained to a single channel and the distance between the CEC (G5, G6 or zSeries 900 server) and the storage system. The fundamental design characteristics of FICON and ESCON are also important factors in this analysis.

For the last ten or more years, customers have become accustomed to the fact that the capacity of an ESCON CEC channel is effectively the same as the capacity of an ESCON control unit port. This is due to the fact that an ESCON channel can only operate on one channel program at a time and the channel, the link and the control unit port all remain busy during the total time that this one channel program is actively executing. If an ESCON control unit disconnects from an ESCON channel, the ESCON channel will begin work on a new channel program and will not continue processing the first channel program until a reconnection occurs. FICON channels present a different paradigm, however, due to the multiplexing capability of FICON. Multiplexing allows a single FICON channel to be simultaneously working on one or more sets of channel program instructions for one or more control units. At the same time, the FICON adapter on the control unit port interface can be working on a different set of channel program instructions from the same or a completely different channel program, and the microprocessors and device adapters in the control unit can be working on yet another completely different set of channel program instructions. When ESCON channels are connected to a control unit port, there are hundreds of microseconds of time where the channel is idle and yet unable to initiate new work or accept reconnections for previously initiated work while the control unit port adapter is working and vice versa. The multiplexing capability of FICON allows this idle time to be converted to productive processing time on additional channel programs. This is a contributing factor to the increased capacity of FICON and FICON Express compared to ESCON channels.

Multiplexing also means that the capability of a single FICON CEC channel is not the same as the capability of a single FICON interface on a control unit. In other words, there is no longer a 1:1 relationship between the effective capacity of the CEC channel and the effective capacity of the control unit port that is connected to the channel. In the next few years, as we improve both channels and control units, there will be times when the single CU FICON interface throughput is greater than, less than or equal to the single FICON channel throughput. When the capacity of a single FICON CU interface is greater than that of a single FICON channel, as is the case with first generation native FICON control units such as the

ESS models F10 and F20 and the S/390 G5/G6 server FICON channel, higher FICON CU interface throughput can be obtained by connecting multiple channels through a director to the same CU port. Similarly, when the capacity of a single FICON channel is greater than the capacity of a FICON CU interface as is the case with the zSeries 900 FICON and FICON Express channels, higher channel throughput can be achieved by connecting multiple control units through a director to the same FICON or FICON Express channel.

FICON and ESCON channels respond very differently to increased distance. Because of the multiplexing capability of FICON, the maximum throughput capability of a FICON channel is not reduced as the distance between the channel and the control unit is increased. With ESCON, the channel, the link and the control unit port all remain busy during the additional time that it takes to execute a program over many kilometers of distance. This extended busy time with ESCON over distance reduces the total throughput capability of an ESCON channel. The multiplexing capability of FICON allows it to begin or continue working on additional channel programs during the time that any one of its channel programs is elongated due to distance. With FICON, the utilization of the processor and the bus and therefore its total throughput capability is not impacted by distance. Please see the FICON versus ESCON Distance Comparison section later in this document for a more thorough discussion of the topic of distance.

FICON channel utilizations are also not impacted by control unit cache miss activity. With ESCON, when a control unit cache miss occurs, the control unit must disconnect from the channel, service the cache miss and then attempt to reconnect to the channel again. The additional overhead involved in disconnecting and reconnecting reduces the maximum throughput capabilities of an ESCON channel. One must also be careful not to drive ESCON channel utilizations too high since the probability of a successful reconnection decreases as ESCON channel utilizations increase. With FICON channels and control units, disconnect and reconnect protocols are eliminated. When the frames of data for one cache miss are ready to be transferred to the channel, they take their position on the queue along with any other frames that may be ready to transfer from any of the other I/O operations that are concurrently active. When a FICON control unit services a cache miss, it simply keeps track of the amount of time it takes to perform this operation for RMF™ reporting purposes only. This time appears as “disconnect” time on the Device Activity Report even though an actual disconnect operation never occurred, i.e., the FICON channel and the control unit never had any formal communications about this event. Effectively, from a FICON channel perspective, a cache miss is treated the same as a cache hit, except that it takes longer for the data transfer to start. The utilization of the FICON channel processor and bus will be the same for the same activity level regardless of whether the control unit cache is achieving a 30%, 60%, 90% or any other cache hit ratio. For workloads with relatively low control unit cache hit ratios, it is therefore possible to achieve greater than 4 to 1 FICON channel consolidation by connecting a single FICON channel to many different physical control units.

To obtain high levels of throughput (i.e. io/sec) especially on a zSeries z900 FICON or FICON Express channel, it is a good idea to logical daisy-chain one control unit with low cache hits ratios (less than 50%) with other control units that have high cache hit ratios. This is because there is a limit of 32 open exchanges per FICON channel and cache misses will keep an exchange open for much longer ( 10ms or more ) than a cache hit. A combination of laboratory measurements and simulation modeling has shown that with the extreme case of 100% control unit cache misses, a FICON Express channel can only be driven to about 30% of its maximum capability with 4K byte data transfer sizes. However with control unit cache hit ratios of 50% or more, a FICON Express channel can be driven to 50% or more of its maximum capability. Since we recommend that all channels including ESCON, FICON and FICON Express be kept at utilizations of 50% or less for reasonable response times with online database processing in a real production workload environment, and the cache sizes of modern control units are as high as 32GB and should therefore yield high cache hit ratios for the vast majority of customer workloads, the open exchange limit should realistically not be a concern. I mention it here only for those customers that are in the habit of running their own I/O driver measurements as part of their new product evaluation process. In some cases, the open exchange limit could restrict the maximum throughput that you are able to obtain in a test environment with a single FICON control unit configured to a single CEC. Setting an open exchange limit per FICON channel was a design implementation decision we made for the benefit of many of our large customers that connect channels from many different CECs to a single FICON control unit port. Its purpose is to prevent a single FICON control unit port from being overloaded with work and to put a limit on the amount of connect time elongation that could be experienced when many different I/O's are simultaneously active on the same channel and control unit port.

If a single ESS box is connected to a single CEC, IBM's laboratory measurement results that appear in the IBM ESS FICON Performance White Paper made available through IBM reps and Business Partners on August 14, 2001 reported that four ESCON ESS host adapters can be replaced by a single FICON ESS host adapter for I/O workloads similar to the Cache Standard and Cache Friendly measurements which have moderate to high cache hit ratios. Greater than 4 to 1 channel consolidation can be achieved however if multiple ESSs are connected to multiple CECs via a native FICON director or if there are many kilometers of distance between the CEC and the control units. The number (N) of ESCON channels that can be consolidated onto a single FICON channel is higher for those customers that have many terabytes of data spread out across multiple control units with relatively low activity on any one CU port. On the other hand, N is low for those customers that want to do all of their I/O on one physical control unit and have high activity rates per CU port with high control unit cache hit ratios. The majority of customers are between these two extremes.

Current ESCON channel utilization patterns are an important input to the FICON channel consolidation analysis process. On a zSeries 900 processor, up to 400% ESCON channel utilization can be consolidated onto a single native FICON Express channel if two or more FICON CU ports are daisy-chained to the same FICON Express channel using a FICON director. This means that as long as the peak of the sum of all the utilizations of the ESCON channel consolidation candidates stays below the 400% limit, these channels can be consolidated onto a single zSeries 900 FICON Express channel. There is no requirement on the peak of any individual ESCON channel. For example, during the time interval that represents the peak of the sum of all the utilizations, the individual ESCON channel utilizations could be as follows: (100%, 50%, 50%, 30%, 70%, 40%, 60%) or (100%, 100%, 100%, 100%) or (50%, 50%, 50%, 50%, 50%, 50%, 50%, 50%) or (20%, 30%, 40%, 50%, 60%, 20%, 30%, 40%, 50%, 60%) or any other combination that adds up to less than or equal to 400%. This 400% limit is an improvement over the 300% limit on the zSeries 900 FICON channels (if two or more FICON CU ports are connected to the same FICON channel using a FICON director) and the 200% limit on the G5 and G6 FICON channels. The superior aggregation capability of the new FICON Express channel allows more customers to take advantage of the simplified configuration example presented in the previous section.

In most customer production environments, variations in I/O activity naturally occur across different control units throughout the day. For many customers, ESCON channel activity rates (io/sec) peak during the day but ESCON channel utilizations peak during their nighttime batch activity. To assist customers with ESCON to FICON migration planning, IBM can analyze customer-supplied SMF and configuration data to suggest ESCON channel candidates for consolidation onto FICON and FICON Express channels. Factors including current ESCON channel utilization and I/O rate, and projected FICON or FICON Express channel processor and bus utilization will be included in the analysis. Guidelines are provided to ensure that the work consolidated on the FICON channel will remain below the "knee of the response time curve." Contact your IBM Sales Representative or IBM Business Partner to request a FICON Channel Consolidation Analysis from the IBM TechLine Organization.

## Native FICON Directors

Native FICON directors simplify the multi-system configuration process for our large zSeries customers.  Native FICON directors allow multiple systems to perform I/O operations concurrently to the same control unit port. With ESCON technology if one system was connected to a director port, all other systems would receive busy signals if they tried to connect to the same port. One system would have to wait until the other system either finished its entire I/O operation or disconnected before it could attempt to do an I/O to that same control unit port. With native FICON packet directors, frames from multiple channel programs from multiple systems can be intermixed at each director and control unit port. The net effect is to improve the overall throughput of a control unit when multiple systems are concurrently active or reduce the number of channels, director and control unit ports needed to support a large configuration as was illustrated in the Simpler Configurations with Native FICON Channels section of this paper.  One must be careful, however, not to overload a single FICON director or control unit port, since this will result in elongated response times. Just like many other resources, a FICON director or control unit port has a level of utilization, beyond which response times increase dramatically and this point is referred to as the “knee of the curve.”  One should consult the performance white papers for each type of FICON attached control unit to understand the “knee of the curve” response time points for that product.

## FICON Improvements on IBM ~ zSeries 900

The z900 has more total bandwidth available to it than previous processors - largely due to the availability of more FICON channels and the higher bandwith capability of each FICON and FICON Express channel.  zSeries 900 supports up to 96 FICON in contrast to G6 servers which only support up to 36 FICON and G5 servers which only support up to 24 FICON channels. The reason for this improvement is the new I/O cages on zSeries 900, which eliminate the physical card slot limits that exist on the G5 and G6 servers. On G5 and G6, once you are at the 256 CHPID limit, it is necessary to remove an ESCON card with 4 ESCON CHPIDs on it in order to install a FICON card with 1 FICON CHPID on it. Therefore when the maximum number of 36 FICON channels is installed on a G6 server, there are only enough card slots left for 120 ESCON channels.  Similarly on a G5 server when the maximum 24 FICON channels are installed, there are only enough card slots left for 168 ESCON channels.  With zSeries 900, these limitations no longer exist.  When the maximum 96 FICON channels are installed on a z900, there are still plenty of card slots left to plug the full complement of 160 ESCON channels.  The only thing you need to remember is that the total number of ESCON and FICON CHPIDs must add up to less than or equal to 256.  Therefore, once you are at the 256 CHPID limit on the z900, only one ESCON CHPID needs to be removed for each new FICON CHPID that you want to add.

zSeries 900 also has a new Flexible Channel CHPID Assignment feature which disassociates CHPIDs from physical locations and therefore allows the assignment of all 256 CHPIDs to usable channel paths. On G5 and G6, when you install an OSA, Crypto or FICON card with 1 active CHPID, you lose the ability to address the other 3 CHPID numbers that were assigned to that physical card slot position. With zSeries 900, all 256 CHPIDs can be accessed since you can assign the 256 CHPIDs to whatever physical location you choose.

The FICON channel on zSeries 900 has a new 333 MHz POWERPC® processor versus the 166 MHz processor on the G5 and G6 FICON channels. An L2 cache was also added to the FICON channel design on zSeries 900. These design changes resulted in about a 30% to 50% improvement in the number of 4K byte ssch/sec that can be processed on each zSeries 900 FICON channel versus a G5/G6 FICON channel. The FICON Express channel on zSeries 900 introduces a 64bit, 66Mhz internal bus, which is twice the width and twice the speed of the original FICON internal bus. The FICON Express bus is capable of “line speed” or 100MB/sec for all reads or all writes and greater than 100 MB/sec for a mix of large reads and write data transfers.

One way to summarize the total increased bandwidth capability of a z900 vs a G5 or G6 server is to compare the peak instantaneous bandwidth of the sum of all the channels on each server. In all cases, the bottom line effective bandwidth of each server is workload dependent. But, assuming for comparison purposes, a peak instantaneous bandwidth of 20 MB/sec for each half duplex ESCON channel and 200 MB/sec for each full duplex FICON channel, a server with 256 ESCON channels could drive a total of 5.1 GB/sec of peak instantaneous bandwidth. A G5 server with 24 FICON and 168 ESCON could drive about 8.2 GB/sec and a G6 server with 36 FICON and 120 ESCON could drive 9.6 GB/sec of peak instantaneous bandwidth. In contrast, a z900 server with 96 FICON and 160 ESCON can drive a whopping 22.4 GB/sec of peak instantaneous bandwidth or 2.3 to 4.4 times as much as previous servers.

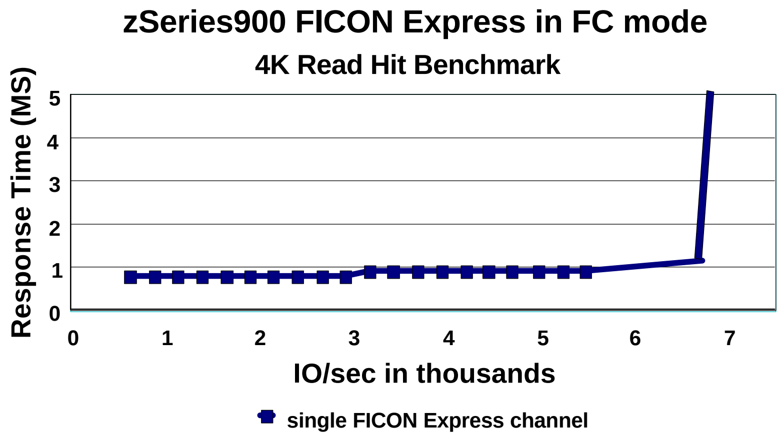


Figure 5

In addition to its bandwidth improvements, the zSeries 900 native FICON Express channel also improves the number of 4K byte operations/sec that can be processed. If a single native FICON Express channel is connected via a native FICON director to two different native FICON Shark CU ports, it can process up to 7200 IO/sec as shown in Figure 5 above. With 96 FICON and 160 ESCON, a z900 could theoretically drive a peak of over 800,000 4K I/O operations per second. This rate, however would not yield reasonable response times. Reasonable response time rates vary by workload. The 4K read hit workload is a very simplistic measurement. Real production workloads are significantly more complicated than this. If you were getting reasonable response times at about 1500 operations/sec on a single G5 or G6 FICON channel or 500 operations/sec on a single ESCON channel, then you could expect to see similar results at about 3000 operations/sec on a zSeries 900 FICON Express channel. Each customer should understand the unique complexities of their workload and adjust the resulting measurements appropriately.

### FICON and FICON Express Channels in FCV mode

The FICON Bridge solution is made up of a FICON or FICON Express channel in FCV mode attached to a FICON Bridge card in a 9032-005 ESCON Director. The FICON Bridge solution allows customers to achieve ESCON channel consolidation by using new high-bandwidth FICON channels with existing ESCON devices.

The following charts show the improvement we have made in the FICON Bridge product since it was originally introduced. These measurements were done with 4 FICON or FICON Express channels running in FCV or FICON Bridge mode connected to 8 ESCON LCU's as shown in Figure 6 below.

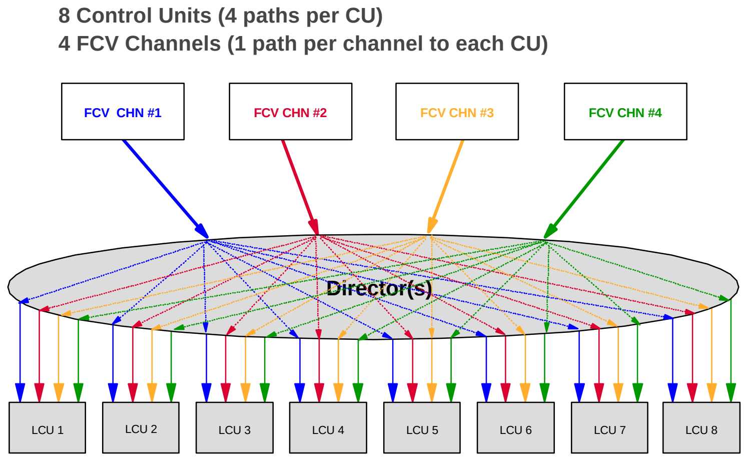


Figure 6

Just as with native mode, FICON Express channels in FCV mode offer substantial improvements in the capability to process both 4K IO operations/sec and high bandwidth applications compared to previous generations of FICON channels. This is illustrated in Figures 7 and 8. Neither the 4K IO operations per second or the large block MB/sec capabilities are as high in FCV mode as they are in native FICON FC mode however, since FCV mode is limited by the bandwidth of the ESCON links and the additional overhead required to do the protocol conversion from ESCON to FICON. A single zSeries 900 FICON Express channel in FCV mode can now support the aggregation of up to 8 or more ESCON channels whose utilizations sum up to less than 350 to 400% utilization. This is a substantial improvement over the 160 to 200% utilization level of aggregation supported by the G5 and G6 server FICON channels in FCV mode. With the much higher aggregation capabilities of the z900 FICON Express channel, one might be tempted to aggregate more than 8 ESCON channels onto a single FICON Express Bridge channel, but in this case one must remember that no more than 8 IO operations can be active concurrently in FCV mode.

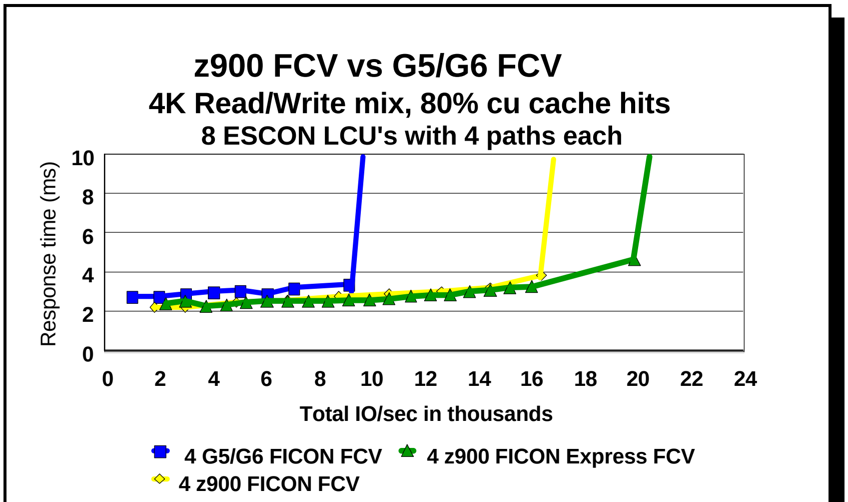


Figure 7

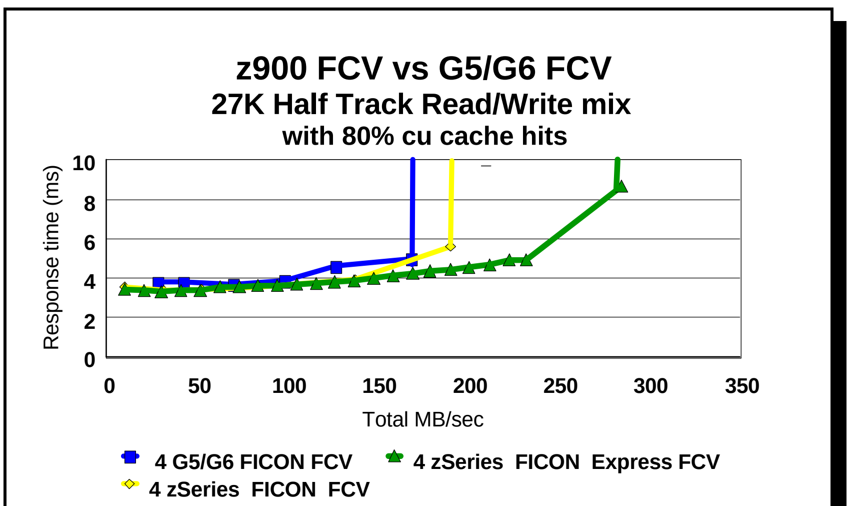


Figure 8

### FICON and FICON Express vs ESCON Distance Comparison

FICON and FICON Express channels are far superior to ESCON channels at increased distances. In particular, ESCON channels experience higher response time elongations than FICON as distances increase and an ESCON channel experiences a much more significant data rate droop at much shorter distances than FICON channels. With ESCON, droop begins at 9 km and at 20 km ESCON delivers less than half the maximum effective bandwidth it provides at 0km. In contrast, FICON channels can maintain a 70 MB/sec effective data rate at 100 km and FICON Express channels can maintain the full 100 MB/sec single direction link rate at 100 km. There are two main reasons for this difference between ESCON and FICON channels: increased buffer sizes in the FICON channel card and improved protocols with fewer interlocks or round trip handshakes across the link.

Droop for both FICON and ESCON is a matter of physics. Droop begins when the link distance reaches the point where the time light takes to make one round trip on the link is equal to the time it takes to transmit the number of bytes that will fit in the receiver's buffer. So the important factors are the speed of light through a fiber, the link data rate and the buffer capacity. The following chart summarizes the droop effect on ESCON, FICON and FICON Express channels.

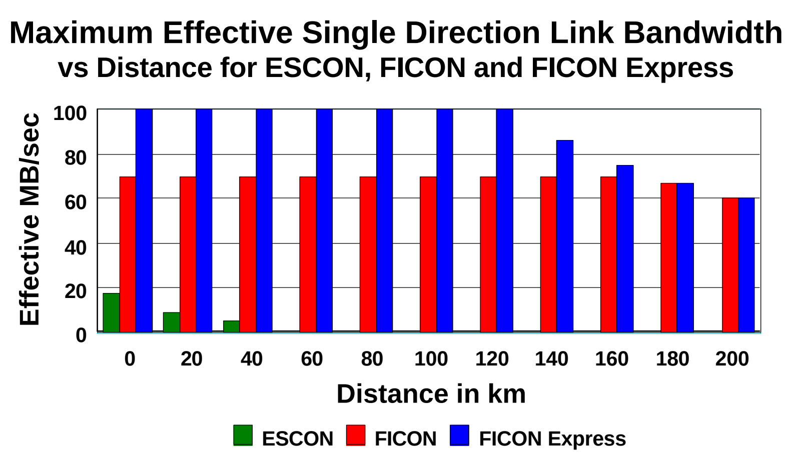


**Figure 9**

With FICON, IBM claims “no droop” out to 100 km. This means that you can maintain the 100 MB/sec link speed out to 100 km. At 150 km, you should theoretically be able to get about 80% of the maximum link capability. However, the original FICON product that we delivered on our G5/G6 and z900 servers had an internal bus limit of about 70 MB/sec depending on the workload you were running. So, it is therefore possible to see less "effective droop" at 150 km with this FICON product than you would see with the new FICON Express product that we announced as part of zSeries 900 GA2 which has a much better internal bus that is capable of delivering the full link speed of 100 MB/sec for highly sequential workloads doing all reads or all writes. Based on the physics of droop both FICON and FICON Express using the same fiber link connected to the same director and control unit should theoretically be able to deliver about the same throughput at 180km or more and the FICON Express product is superior to FICON at distances less than that point. Both FICON and FICON Express are far superior to ESCON at all distances.

In all cases, one needs to remember that the basic speed of light delay will add 0.1ms of increased response time for every 10 km of distance and every interlocked handshake or round-trip per channel program. Native FICON and FICON Express channels simply open a communication exchange at the beginning of the channel program and the control unit closes the exchange at the end for a total of one interlocked handshake or round-trip per channel program. The speed of light propagation delay is 5 microseconds( µs ) per km one way or 10 µs per km round trip. Therefore a simple channel program (4 KB read hit ) with 6 interlocked handshakes on ESCON channels that experienced up to a 60 µs increase in response time every km or 0.6 ms per 10 km will now experience a propagation delay of only about 0.1 ms per 10 km of distance with native FICON or FICON Express channels. When FICON and FICON Express channels are operating in FCV or FICON Bridge mode, there is about a 0.2 ms increase in response time for every 10 km of distance. The following chart summarizes the increase in response time in milliseconds (ms) for ESCON, FICON and FICON Express in FCV and FC mode for simple channel programs.

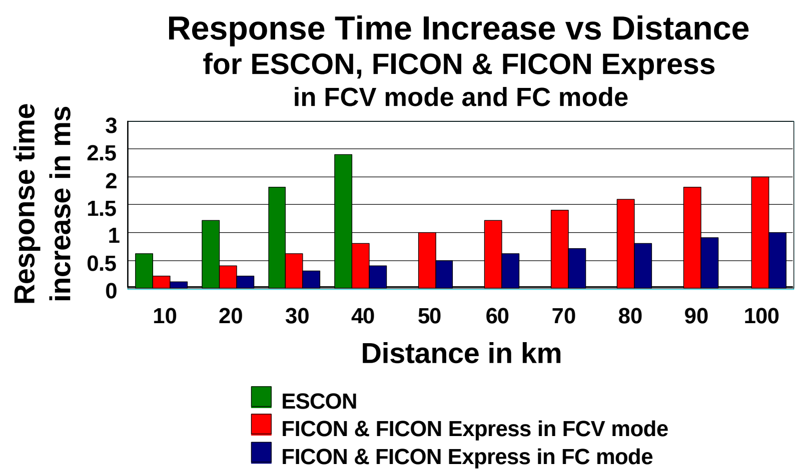


Figure 10

One must also remember that a distance of 100 km was the IBM design point for FICON. This means that the IBM Poughkeepsie test team did a complete set of functional tests for FICON at 100 km and that IBM requested that vendors developing native FICON directors and other FICON attachments provide enough (i.e. at least 50) buffer credits to support FICON 1 Gigabit links at 100km. So, while it is theoretically possible for FICON to maintain high bandwidth at distances greater than 100 km, these distances have not been tested and are only achievable if enough buffer credits exist to support the link speed.

## FICON RMF Information

The primary RMF report of interest for FICON is the Channel Path Activity report.

FICON channels can be identified from the TYPE column; their type begins with FC:

type FC indicates a native FICON channel;
type FC_S indicates a switched native FICON channel;
type FCV indicates a FICON bridge channel which connects to an ESCON control unit via a bridge card in a 9032 model 5 ESCON director.

For a given FICON channel there are three possible entries under UTILIZATION (%):

PART denotes the FICON processor utilization due to this logical partition (this field will be blank in BASIC mode or if EMIF is not installed).
TOTAL denotes the FICON processor utilization for the sum of all the LPARs.
BUS denotes the FICON internal bus utilization for the sum of all the LPARs.

The FICON processor is busy for channel program processing, which includes the processing of each individual channel command word (CCW) in the channel program and some setup activity at the beginning of the channel program and cleanup at the end. The FICON bus is busy for the actual transfer of command and data frames from the FICON channel to the adapter card, which is connected via the FICON link to the director or control unit. The FICON bus is also busy when the FICON processor is polling for work to do. This is why one can see anywhere from 5 to 12% FICON bus utilization on the RMF Channel Activity report during time intervals when there are no I/Os active on that channel. We do not have an accurate way of determining actual FICON processor utilization, i.e. We do not have a way of knowing when the FICON processor is busy versus when it is idle. RMF estimates FICON processor utilization by comparing the actual number of command and data sequences processed with a “maximum” number. The “maximum” number is determined by the number of sequences processed while running an I/O driver program doing 100% 4K read hits on the IBM Poughkeepsie test floor.

The actual FC channel processor and bus utilizations as reported by RMF will vary by workload and by channel type. For the same workload and same activity level, utilizations reported for zSeries 900 FICON Express channels will be different than utilizations reported for zSeries 900 FICON channels and G5/G6 FICON channels. In general, small data transfer sizes will drive both zSeries 900 and G5/G6 FICON channel processor utilizations higher than bus utilizations while large data transfer sizes will drive the FICON channel bus utilization higher than the processor utilization. Figure 11 shows the processor and bus utilizations that are reported on the RMF Channel Activity report for 4KB read hit I/O operations on the zSeries 900 native FICON channel. At the “knee of the response time curve,” for this very simple 4KB read hit program, FICON processor utilizations are about 80% and FICON bus utilizations are about 40%.

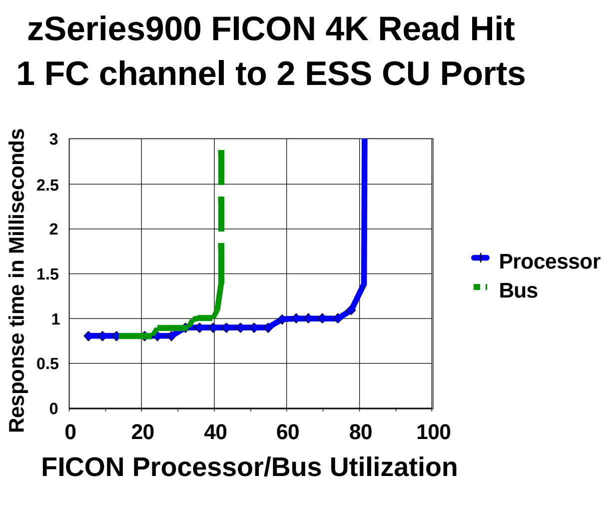


Figure 11

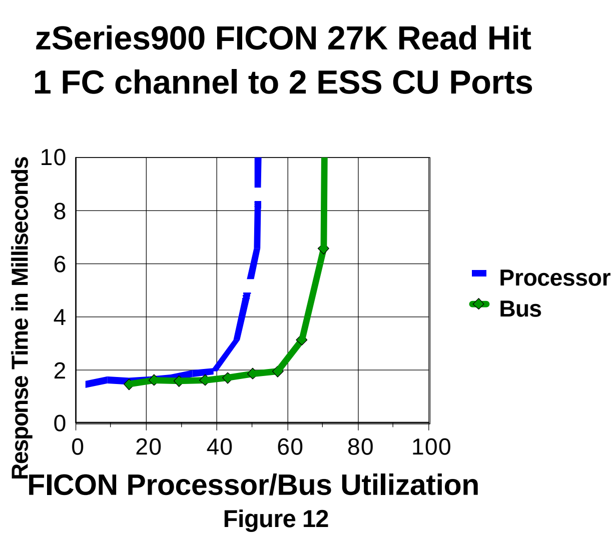


Figure 12

Figure 12 shows the processor and bus utilizations that are reported on the RMF Channel Activity report for 27K or half-track read hit I/O operations on the zSeries 900. For this series of measurements, the "knee of the response time curve" occurs at 60% FICON bus utilization and 40% FICON processor utilization. In general, to achieve good response times with real production workloads, one should keep the maximum of both the FICON processor and FICON bus utilization less than 50%.

It is important to note that the maximum FICON bus utilization achieved in our laboratory measurements was about 80% and not 100% as one might expect. This is due to the fact that the bus utilization represents the number of cycles that the bus was busy compared to the theoretical maximum possible bus busy cycles based on the cycle time of the bus. In the process of transferring command and data frames across the bus, there are gaps that occur between transfers depending on the arrival pattern and the mix of different channel programs running at any one point in time, so it is therefore not possible to achieve 100% FICON bus utilization with normal channel program processing.

Since the bus on the zSeries 900 FICON Express channel has a higher capacity than that of the zSeries 900 and G5/G6 FICON channel, the FICON Express bus utilizations will be lower than the FICON Express processor utilizations for the vast majority of workloads. In the laboratory performance measurements that we have done for the FICON Express channel, we have not seen the bus be the limit for maximum capability. For workloads with small data transfer sizes, the FICON Express processor will be the limit to maximum throughput just as it is with FICON channels. For workloads transferring large amounts of data all in one direction, the 100 MB/sec link will be the limiting factor. In our measurements, the FICON Express bus utilization was less than 50% when the 100 MB/sec link limit was achieved.

It is therefore recommended that you pay more attention to the MB/sec information than the bus utilization on the RMF report for FICON Express channels.

FICON and FICON Express channels provide bandwidth information (MB/SEC) not available for ESCON channels. This is provided separately for READs and WRITEs since the fibre channel link is full duplex, at both the logical partition level (PART) and the entire system level (TOTAL). Figure 13 shows the MB/sec that can be achieved with channel programs using 4 K byte and 27 K byte blocksizes for both FICON and FICON Express channels. With data transfer sizes of only 4K bytes, the maximum attainable bandwidth is in the range of 20 to 30 MB/sec, while data transfer sizes of 27 K or 6x27 K bytes are capable of 60 to 80 MB/sec for FICON and 80 to 120 MB/sec for FICON Express. A bandwidth of just over 50 MB/sec was achieved when 32 4K byte blocks were chained together in a single 32x4 K channel program. We would say that this channel program has a blocksize of 4 K with a total data transfer size of 128 K bytes. As you can see the MB/sec attainable with this channel program is not as high as the MB/sec attainable using channel programs with 27 K byte blocksizes.

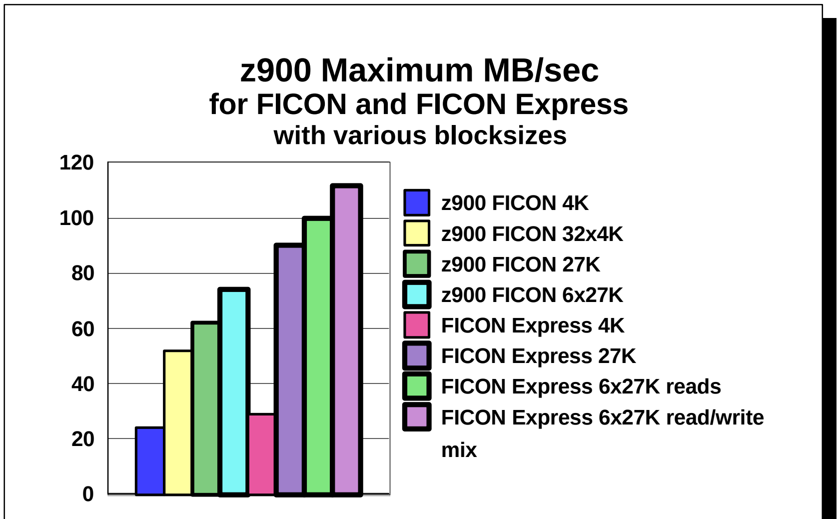


Figure 13

For FICON channels it is possible to estimate the average number of bytes transferred per SSCH by dividing the MB/sec of a FICON channel from the Channel Path Activity report by the total SSCH/sec processed by a FICON channel from the I/O Queuing Activity report. The total SSCH/sec processed by a FICON channel can be determined by adding up all of the "CHPID taken" fields on the I/O Queuing Activity report for each LCU that a single FICON channel is connected to. If the average data transfer sizes of your channel programs are greater than 27 K bytes, then you should compare the MB/sec fields on your RMF Channel Activity Report to graph in Figure 13 to assess the maximum capability of FICON channels for your workload.

For a native FICON ESS, it will also be of interest to consult the Direct Access Device Activity report. Here one can examine the AVG RESP TIME and various response time components for activity to the LCUs attached to the FICON channels. In particular, AVG CONN TIME for large blocksize transfers should be significantly less for native FICON channels than for the same transfer size on ESCON or FICON Bridge channels due to the 100 MB/sec link transfer speeds of native FICON.

With native FICON, customers should continue to analyze their I/O activity by looking at the DASD or TAPE activity reports, just as they did with ESCON channels. If response time is a problem, then the response time components need to be looked at. If disconnect time is a problem, then an increase in CU cache size might help. If IOSQ time is a problem, then ESS with Parallel Access Volumes might help. If pend or connect times are too high, then one can look at the FICON processor and bus utilizations. If either one of these utilizations is above 50% then overuse of the FICON channel could be contributing to additional pend and connect time delays. If, on the other hand, pend and connect times are high and FICON channel utilizations are less than 50%, then overuse of a FICON director port or control unit port could be contributing factors. If FICON channels from multiple CECs are connected to the same director destination port , then one must add up the activity from all the CECs to determine the total destination port activity. This total activity level should be less than the "knee of the curve" points depicted in the measurement results that appear in the white papers for the specific native FICON DASD or TAPE product that is being used.

One of the basic differences between native FICON and ESCON channel performance is the CONNECT time component of response time. Since an ESCON channel is only capable of executing one I/O at a time, the amount of time that it takes to execute the protocol + data transfer components of CONNECT time is relatively constant from one I/O operation to the next with the same exact channel program. With FICON however, CONNECT time can vary from one execution of a channel program to another. This is a side effect of the multiplexing capability of FICON. Since both the channel and the control unit can be concurrently executing multiple I/O operations, the individual data transfer frames of one I/O operation might get queued up behind the data transfer frames of another I/O operation. So, the CONNECT time of an I/O with FICON is dependent upon the number of I/Os that are concurrently active on the same FICON channel, link and control unit connection. Multiplexing also means that the start and end of the CONNECT time for one native FICON I/O operation can overlap the start and end of the CONNECT time for several other native FICON I/O operations. With ESCON, the additional queuing delays caused by having multiple I/Os concurrently active appear in the PEND or DISC time component of response time. If the same workload with the same activity rate and the same level of I/O concurrency is run on native FICON channels instead of ESCON channels, then one could see the PEND and DISC time components of response time decrease and the CONNECT time component increase for small data transfer sizes. For large data transfers, the improved CONNECT time due to the 100 MB/sec link transfer speed will most likely offset any increased CONNECT time due to multiplexing queuing delays. Figure 14 illustrates the improvement in connect time for FICON compared with ESCON for large data transfers when I/Os were active to 512 different volumes across 4 FICON channels or 128 volumes per FICON during a 1 minute RMF interval.

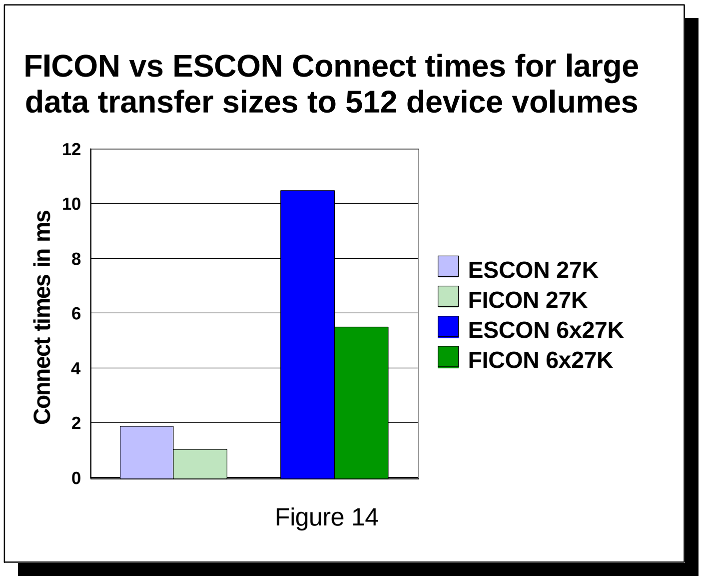


Figure 14

A few other fields on the RMF I/O Queuing activity report will be different for FICON versus ESCON. For example, the %DP BUSY should be 0 for native FICON due to the elimination of destination port busy signals with native FICON directors. %CU BUSY should also be 0 for native FICON in most customer production environments. CU busies will only occur with native FICON when an individual CU port is being overloaded with work from many different FICON or FICON Express channels simultaneously. If this is the case and if no errors were made in the IOCDS then some "tuning" of your configuration is necessary to reduce these CU busies and achieve better response time results. With z/OS V1R2 and RMF Release 12, DP and CU busies will now be summarized for each IOP on the RMF I/O Queuing Activity report. The % ALL CH PATH BUSY field has been eliminated from this release of RMF since it is meaningless for FICON. Instead, channel path busies as well as IOP utilizations will be reported for each IOP. One of the benefits of native FICON is that it makes Saps or IOPs more productive due to the reduction in DP and CU busies. For the same activity rate, one should see less IOP utilization % busy with native FICON and FICON Express channels than with ESCON channels. One must be careful not to misinterpret IOP utilization %'s however. High IOP utilization %'s are usually an indicator of contention especially with ESCON channels, directors and control units. Adding additional IOPs will NOT help reduce channel configuration contention. One must identify the source of the configuration contention and fix it. Migrating from ESCON to native FICON configurations is a natural solution to this problem.

In summary, the basic architecture and design differences between FICON and ESCON results in many changes to the performance data that appear on RMF reports. Additional information in the form of FICON processor and bus utilizations and READ and WRITE MB/sec is provided to help analyze the multiplexing capability of FICON. Since ESCON is only capable of executing one I/O operation at a time, RMF reports the time that the entire

CHPID path is busy for ESCON channel utilization. With FICON, we must consider the individual components of the total CHPID path such as the FICON channel processor and bus, the fibre link, the director destination port and the control unit port adapter microprocessor and bus. The charts and examples provided in this paper should help guide you in assessing the maximum capability of FICON and FICON Express channels for your workload.

## Conclusion

zSeries 900 FICON Express channels offer many benefits over ESCON and our original FICON channels. The increased throughput and bandwidth capabilities of these channels offer the opportunity for improved performance with simpler configurations and reduced infrastructure over longer distances to meet the needs of future datacenter growth including backup and disaster recovery requirements. The total native FICON solution — DASD, TAPE and Printer attachments, directors and the new and improved FICON Express channels — are now available and ready for your installation.

Additional FICON product information is available on the IBM System Sales Web site and the zSeries I/O connectivity Web site at wcs.haw.ibm.com/servers/eserver/zseries/connectivity/

## Acknowledgement

The data presented in this paper is based upon measurements carried out over the last couple of years with a mixture of internal IBM tools and non-IBM I/O driver programs. The measurement results depicted in Figures 5, 7, 8, 11, 12 and 14 were achieved by using Version 10 of the PAI/O Driver® for OS/390® program. The measurement results depicted in Figures 1 and 13 are a mixture of internal IBM tools and the PAI/O Driver® program.

I would like to thank all of the reviewers of this paper for their comments and special thanks goes to the LSPR measurement team and the following people for their continued support on this effort:

| | |
|---|---|
| Rich Basener | Rich Basener/Poughkeepsie/IBM@IBMUS |
| Mario Borelli | Mario Borelli/Poughkeepsie/IBM@IBMUS |
| Larry Blount | Lawrence Blount/Tucson/IBM@IBMUS |

Copyright IBM Corporation 2002

IBM Corporation
Marketing Communications
Route 100
Somers, NY 10589
U.S.A.

Produced in the United States of America
2/02
All Rights Reserved

IBM, IBM logo, e-business logo, DB2, Enterprise Storage Server, ESCON, FICON, FICON Express, Magstar, POWERPC, RMF, S/390, z/OS, and zSeries are trademarks or registered trademarks of International Business Machines Corporation of the United States, other countries or both.

Java and all Java-based trademarks and logos are trademarks of Sun Microsystems, Inc. in the United States, other countries or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Intel is a trademark of Intel Corporation in the United States, other countries or both.

Linux is a registered trademark of Linus Torvalds.

PAI/O Driver is a Trademark of Performance Associates, Inc.

Other company, product and service names may be trademarks or service marks of others.

Information concerning non-IBM products was obtained from the suppliers of their products or their published announcements. Questions on the capabilities of the non-IBM products should be addressed with the suppliers.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

GM13-0120-00