ESCON-to-FICON Bridge Channel Aggregation Guidelines

These guidelines are intended to assist in the selection of ESCON channel candidates for aggregation onto FICON Bridge channels. Like ESCON channels, FICON Bridge channels have a level of utilization beyond which response time increases dramatically - this is referred to as the "knee of the curve". Using these aggregation guidelines should help keep I/O on FICON Bridge channels below the knee of the response time curve.

There are several reasons for moving ESCON channel I/O to a FICON Bridge channel: increase processor I/O bandwidth; increase connectivity distance; provide additional device addressability; and reduce connectivity infrastructure. Except when implementing connectivity at 10 kilometer or more, moving well-tuned ESCON I/O to a FICON Bridge channel will result in some I/O operation response time elongation. If however, significant IOSQ or PEND time delays exist in the current ESCON configuration, then moving to the FICON bridge channel may result in a slight improvement, or no change, to the existing I/O response times.

Response time elongation can vary from 400 microseconds for short records at low activity rates to about one millisecond for very long records or a mix of record sizes at high FICON Bridge channel utilization levels. While there will be some response time elongation, the focus should not be on the elongation per se, but rather on whether or not the elongation will have an effect on users and applications. For the most part, 400 or so microseconds of short-record response time elongation will go unnoticed by on-line transaction processing. The primary consideration is that larger response time elongation may occur with batch job processing of very long records, or at high levels of FICON Bridge channel utilization. For this reason, FICON Bridge channel utilization should be kept at recommended levels, and the effect of response time elongation needs to be estimated for batch job "windows". The formula for response time elongation is (sequentially_dependent I/O count) * (average elongation per I/O) -- for example, a batch job network with one million sequentially_dependent I/O operations that experience an average .9 millisecond elongation each, will result in a total elongation of about 900 seconds, or 15 minutes. On the other hand, some batch job measurements resulted in no job elongation, since improvements in IOSQ time compensated for the slight degradations in PEND and CONN time.

These aggregation guidelines are aimed at configurations with all I/O subsystem connectivity to FICON Bridge channels. Connecting an I/O subsystem to a combination of FICON Bridge channels and ESCON channels should permit higher channel utilization levels than recommended here.

Remember, these are only guidelines, and their use does not guarantee I/O performance on aggregated FICON Bridge channels.

All Subsystem Aggregations

Input for the selection of ESCON channel (actually, ESCON I/O subsystem) candidates for aggregation to FICON channels is provided by MVS measurement data, the number of FICON Bridge channels required to satisfy connectivity, and a "best fit" judgment by the user of the FICON Channel Aggregation tool.

These guidelines assume that RMF data has been provided, and that the FICON Channel Aggregation tool is used to present measured ESCON channel utilization and estimated FICON Bridge channel utilization data over an extended period of time.

These guidelines are applicable to aggregating all ESCON I/O subsystem types onto FICON Bridge channels.

- Take utilization measurements, in the form of SMF type 70-78 records, during periods of highest I/O subsystem (not necessarily processor) utilization - typically month-end, quarter-end, year-end, and seasonal peak periods. Gather data in 15 minute intervals for a full week, including weekends - this is most critical for tape activity measurements, since tape use may not be as predictable as DASD.
- Consider the effect of subsystem aggregation when an application or image is moved to a backup processor or image.
- Consider customer requirements for connectivity symmetry across all images and processors, where any aggregation on one processor may need to be duplicated on all processors.
- Do not exceed estimated FICON Bridge channel processor and bus utilization levels of 50%.
- When aggregating ESCON channels with logically daisy chained I/O subsystems, be aware that these subsystems will no longer be "throttled" by sharing a single channel, but will have full access to the eight concurrent channel paths of the FICON Bridge channel. All I/O subsystems aggregated onto a FICON Bridge channel will "compete" evenly, with all other aggregated subsystems, for access to that channel,.
- Be cautious when aggregating subsystems with zero utilization during the measurement period. Verify the utilization or validate it with measurements from other periods. Also, be aware of future plans for these subsystems.
- Implement only as many FICON Bridge channels as needed to meet current additional connectivity requirements. The strategy should be to use FICON Bridge channels to provide additional ESCON connectivity until native FICON channels and I/O are available later this year.
- Spread FICON channel connectivity to as many I/O subsystems as possible, and retain the maximum possible ESCON channel connectivity to each subsystem. This will "dilute" any response time elongation by spreading I/O activity over both ESCON and FICON channels - at high levels of channel utilization, FICON channel activity "spills over" to ESCON channels.
- Plan sufficient utilization "headroom" for I/O subsystem growth or additional I/O subsystems on aggregated FICON channels.

DASD-only Subsystem Aggregation

These additional guidelines are intended for FICON Bridge channel aggregations when only DASD subsystems are included in the aggregation.

- Do not exceed a combined ESCON channel utilization level of 160% for subsystem I/O with a one millisecond or less connect time, and 200% for subsystem I/O with greater than one millisecond connect time. When multiple I/O subsystems are connected to the same FICON channel, the result will most likely be a mix of I/O with short and long connect times, and using a combined utilization of somewhere between 160% and 200% should keep response times below the knee of the curve.

Tape-only Subsystem Aggregation

These additional guidelines are intended for FICON Bridge channel aggregations when only tape subsystems are included in the aggregation.

Additional care should be taken when selecting tape subsystems for aggregation, because tape device activity has irregular use patterns when compared with DASD, and because tape utilization often occurs in short bursts of near-100% followed by longer periods of zero utilization.

First, it must be determined if elongated tape subsystem response times will affect a batch job completion "window". For planning purposes, I/O operation elongation could be as much as one millisecond when very large blocks of data are being transferred.

- Understand the characteristics of applications that use tape, and overall tape allocation and use patterns.
- Verify tape activity over several days, including a weekend, to determine a tape usage profile - patterns of use and peak activity.
- Do not exceed a combined ESCON channel utilization level of 300%. This will result in aggregating no more than three or four heavily-used (75-100%), concurrently-operating ESCON tape channels onto a single FICON Bridge channel.
- Assess the potential impact of the simultaneous activity of all tape subsystems in an aggregation.
- Be cautious when aggregating tape subsystems with zero channel utilization during the available measurement intervals. It is likely that, when used, these subsystems will operate at near-100% channel utilization.
- Based on the potential for different allocation patterns and channel utilization characteristics, consider aggregating different types of tape subsystems onto the same FICON Bridge channel.

Note: Anticipate that some tape-only FICON Bridge channel aggregation recommendations may not result in an ESCON-equivalent channel savings.

Mixed DASD and Tape Subsystem Aggregation

These additional guidelines are intended for FICON Bridge channel aggregations when both DASD and tape subsystems are included in the aggregation.

There are potential benefits when ESCON DASD and tape subsystems can be aggregated onto the same FICON Bridge channel because these different classes of I/O subsystems may have non-concurrent usage patterns and different data blocksize or read/write characteristics.

- Apply all DASD-only and tape-only subsystem aggregation guidelines.
- Aggregate different DASD and tape device types and application use on the same FICON channel. When different device types, with their different I/O usage patterns and characteristics, support different applications, the result is often a minimum concurrent usage of the aggregated I/O subsystems.
- Separate the aggregation of subsystems used for dump-restore operations - do not group all dump or restore source and target subsystems on the same FICON Bridge channels. Spread an aggregation so that dump and restore operations are spread across all available FICON Bridge channels. Because more I/O subsystems will be connected to one FICON Bridge channel than were connected to a single ESCON channel, more careful allocation of dump and restore devices may be necessary to spread allocation across available FICON Bridge channels.
- Avoid aggregating tapes on all FICON Bridge channels connected to an individual DASD subsystem. For example, if an eight-path DASD subsystem has four ESCON paths and four FICON Bridge paths, avoid aggregating tape subsystems on more than two of the FICON Bridge paths. This is most significant for applications that are sensitive to DASD response times than to tape response times.
- Avoid aggregating tapes with key DASD subsystems together on the same FICON Bridge channels when the tapes are expected to be active during critical DASD response time windows. Instead, aggregate tapes and DASD that are active during different or non-critical times.
- Do not exceed a combined ESCON channel utilization level of 200% on a FICON Bridge channel with concurrent tape and DASD activity if tape is executing long record jobs and DASD is executing short record jobs and you are concerned about DASD response times. A utilization level of 300% is acceptable during periods of tape-only activity.

Aggregating Other Subsystems

There are no specific recommendations for aggregating other I/O subsystems on FICON Bridge channels, but these are some things to evaluate when considering including other I/O subsystems in a FICON Bridge channel aggregation.
.
- Including an I/O subsystem in a FICON Bridge channel aggregation requires that the I/O subsystem be attached through an ESCON Director (with a FICON Bridge card installed.)
- A high I/O rate can cause some elongation of response time. This may be critical for XCF Channel-to-Channel communication traffic, where sub-millisecond response time is the norm.