FICON Multihop

Requirements and Configurations

April 2017 (Updated June 2017)

Pasquale “PJ” Catalano

Dr. Steve Guendert

Document WP102704

Systems Group

Notices, Disclaimer and Trademarks

Table of Contents

FICON Multihop: Requirements and Supported Configurations

Executive Summary:

Through this whitepaper and supporting z Systems Storage Area Network (SAN) qualification letters, IBM announced support for a new set of cascaded FICON configurations for FICON interswitch links (ISLs) called FICON Multihop. FICON Multihop allows support for cascading up to four switches, three hops, expanding the previous restriction of two cascaded switches, one hop. Multihop is only supported using traditional static routing methods, and is not supported in conjunction with FICON Dynamic Routing (FIDR).

z13 servers using Multihop have advantages for simplifying configurations in a single site or across multiple sites with ISL and cascaded FICON directors:

- Support multiple configurations for easy deployment of GDPS
- Improve utilization of FICON directors and switches
- Simpler and easier to manage configurations for SAN availability

This white paper will discuss requirements for FICON Multihop, review the basics of cascaded FICON, and supported configurations for FICON Multihop. The paper will conclude with a discussion of design considerations and the possible use cases where implementing FICON Multihop will be of the most benefit to z Systems end users.

Introduction

IBM originally announced FICON channels for the S/390 9672 G5 processor in May of 1998. Over the ensuing 18 years, FICON has rapidly evolved from 1 Gbit FICON Bridge Mode (FCV) to the 16 Gbit FICON Express16S channels that were announced in January 2015 as part of the IBM z Systems z13 announcement. While the eight generations of FICON channels have each offered increased performance and capacity, from the perspective of this paper the most important enhancements relate to the fibre channel upper level protocol (ULP) employed by FICON. Specifically, the first FICON ULP was called FC-SB-2. This specification supported native FICON storage devices, channel-to-channel connections, FICON directors (switches), as well as a number of other key features. However, the FC-SB-2 ULP did not support switched paths over multiple directors. When IBM introduced the FC-SB-3 ULP in January of 2003, this limitation was addressed. While FC-SB-2 employed a single byte link address that specified just the switch port on the director, FC-SB-3 employs a 2 byte address where the first byte is the address of the destination director and the second byte is the address of the port on the destination director. This is known as cascaded FICON. Since 2003, there have been many changes with how the FICON switching devices send fibre channel frame traffic between each other. The focus of this paper is to illustrate the new configurations that are supported in a Multihop environment.

Cascaded FICON Defined

Cascaded FICON refers to an implementation of FICON that allows storage fabrics to be linked via connections between pairs of FICON directors and/or FICON switches. The connections between these switching devices are known as interswitch links (ISLs). ISLs support processor-to-processor, processor-to-disk or tape subsystem, and subsystem-to-subsystem logical switched connections. Cascaded FICON facilitates the design and implementation of robust disaster recovery/business continuity solutions such as IBM's Geographically Dispersed Parallel Sysplex™ (IBM GDPS®). Cascaded FICON architectures substantially reduce the infrastructure costs and complexities associated with these implementations. Cascaded FICON also permits greater flexibility in the FICON architecture, more effective utilization of fibre links, and higher data availability in the enterprise.

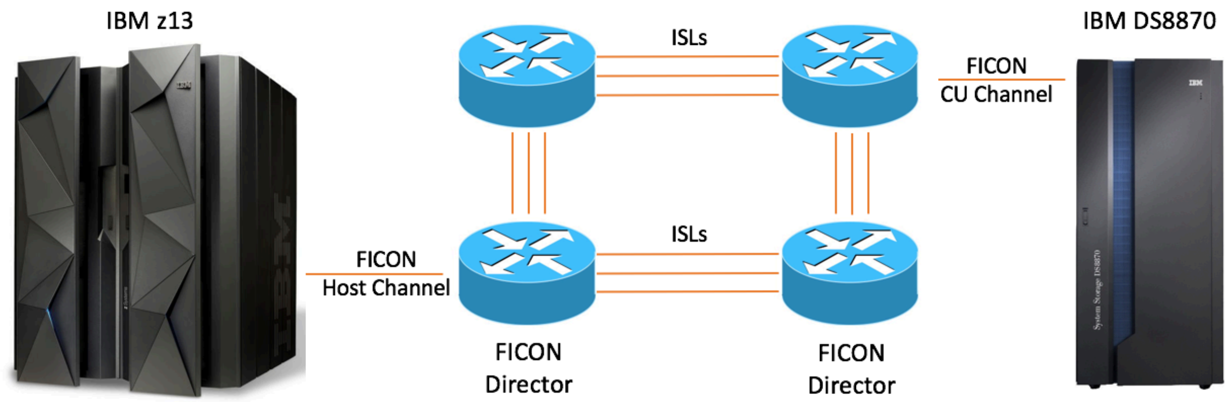

Figure 1: Cascaded Switch Configuration in a Multihop Environment

As is shown in Figure 1, a logical connection over a cascaded FICON director configuration incorporates three end-to-end links. The first link connects the FICON channel N_Port (node port) to an F_Port (fabric port) on the FICON director. The links between directors connect E_Ports (expansion port) on the local director to E_Ports on the remote directors. An ISL is a link between two switches, E_Port-to-E_Port. The ports of the switches automatically come online as E_Ports once the login process finishes successfully. Finally, the third link connects the F_Port on the director to an N_Port on the control unit (CU) subsystem.

The Hardware Configuration Definition (HCD) defines the relationship between channels and directors (by switch ID) and specific switch port (as well as the destination switch ID for cascaded connections) for the switch to director segment of the link. However, HCD does not define the ISL connections. While this may initially appear to be surprising based on a reader's prior experience with the exacting specificity required by HCD, it means that the management of the traffic over ISLs is controlled exclusively by the directors. Hence, additional ISL bandwidth can be added to a configuration without any modification to the environment's HCD definitions. HCD simply assumes that the links are present and requires that 2 byte addressing be employed to define the connectivity to storage subsystems. The first byte is the switch ID and the second byte is the port to which the storage subsystem is connected. During initialization, the switches identify their peers and create a routing table so that frames for remote subsystems may be forwarded to the correct director.

Hardware Requirements for FICON Multihop

FICON Multihop has several hardware requirements that need to be met prior to implementation. There are z Systems, storage, SAN and network/DWDM device requirements.

Z Systems hardware requirements

FICON Multihop requires z13/z13s running Driver 27L with bundle S08a or later with FICON Express 16S or FICON Express 8S channels.

DASD/Storage requirements

FICON Multihop is supported with IBM DS8870 running firmware release 7.5 or later. Multihop will work with both 16Gbps and 8Gbps host adapter cards. Other storage OEMs may support Multihop as well. Clients should check with their DASD vendor to verify support.

SAN Hardware requirements

FICON Multihop is supported on the IBM B type (Brocade) z Systems qualified Gen 5 products SAN768B-2 SAN384B-2, SAN 48B-5, and SAN42B-R (DCX 8510-8, DCX 8510-4, 6510, and 7840) and Gen 6 products SAN512B-6, SAN256B-4, SAN64B-6 (X6-8, X6-4 and G620) running FOS levels 7.4.0a and later qualified versions. Brocade Network Advisor 12.4.1 or later qualified is also recommended.

FICON Multihop is supported on the Cisco z Systems qualified products MDS 9710 and MDS 9706 running NxOS 6.2(11c), and MDS 9250i running NxOS 6.2(11d). Data Center Network Manager 7.2(1) or later qualified version is also recommended.

This document is not meant to determine qualified switch products. To ensure that the planned products to be implemented are qualified, registered users can see the IBM Resource Link® library for current information about qualified switch products and routing modes supported:

https://www-304.ibm.com/servers/resourcelink/lib03020.nsf/pages/switchesAndDirectorsQualifiedForIbmSystemZRFiconRAndFcpChannels?OpenDocument

Network/DWDM requirements

If a network or Dense Wavelength Division Multiplexer (DWDM) device is supported in a z Systems environment, in general it is supported for Multihop FICON. A list of DWDM products qualified for use with z Systems can be found at IBM Resource Link® library:

https://www-304.ibm.com/servers/resourcelink/lib03020.nsf/pages/systemzQualifiedWdmProductsForGdpsSolutions?OpenDocument&pathID=

However, clients should also verify support of network/DWDM devices with their FICON SAN hardware components. Clients should also verify specific support with ISL features such as the use of R_RDY and VC_RDY. Clients should check with both their switch, and DWDM vendors to verify support.

Software/OS requirements for FICON Multihop

FICON Multihop is supported on all operating systems that run on z13 and z13s that allow FICON channels to be defined.

Review of Fabric Shortest Path First (FSPF)

The Fabric Shortest Path First (FSPF) protocol is the standardized routing protocol for Fibre Channel (hence FICON) SAN fabrics. FSPF is a link state path selection protocol that directs traffic along the shortest path between the source and destination based upon the link cost. FSPF detects link failures, determines the shortest route for traffic, updates the routing table, provides fixed routing paths within a fabric, and maintains correct ordering of frames. FSPF also keeps track of the state of the links on all switches in the fabric and associates a cost with each link. The protocol computes paths from a switch to all the other switches in the fabric by adding the cost of all links traversed by the path, and chooses the path that minimizes the costs. This collection of the link states, including costs, of all the switches in the fabric constitutes the topology database or link state database.

FSPF is based on a replicated topology database that is present in every switching device in the FICON SAN fabric. Each switching device uses information in this database to compute paths to its peers via a process known as path selection. The FSPF protocol itself provides the mechanisms to create and maintain this replicated topology database. When the FICON SAN fabric is first initialized, the topology database is created in all operational switches. If a new switching device is added to the fabric, or the state of an interswitch link (ISL) changes, the topology database is updated in all the fabric's switching devices to reflect the new configuration.

For more information regarding how the FSPF Protocol works, please see the FICON Dynamic Routing (FIDR) whitepaper, where the mechanism is discussed in detail:

https://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP102651

FICON Multihop Supported Configurations

The configurations supported for Multihop environments fall into two categories. The first is using native ISLs only. The second, is using a mix of native ISLs with Fiber Channel over Internet Protocol (FCIP) extension networks. In the examples below we will refer to the actual link distance between adjacent switch X and switch Y as $\overline{XY}$.

Native ISL Configurations

Below are the configurations supported for native ISLs. These configurations are comprised of two-hop (three switch) environments and three-hop (four switch environments). For these configurations, we are just using native ISLs to transport FICON traffic, and not using any emulation or encapsulation technologies to extend the base supported distance for FICON links. Links that are greater than 10km require the use of a qualified DWDM solution.

Two-Hop, Three Switch, Configurations

In this set of configurations, we are looking at the combinations of switch topologies, but restricting them to only three switches. These environments are considering two-hop configurations, because at most you may traverse two switches to reach your target.

Two-Hop Array

The first configuration is three switches that are connected linearly, as seen in Figure 3. This is called the Two-Hop Array configuration. For native ISL configurations, we need to make sure we still adhere to the FICON timeout limitations. This means the longest distance a FICON packet can traverse is 300km. Therefore, the maximum distance of $\overline{AB} + \overline{BC} <= 300$km.

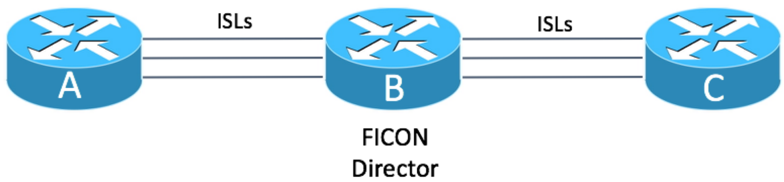

Figure 2: Two-Hop Array Configuration

The Triangle

The next configuration is three switches connected in a triangle configuration, as seen in Figure 3. This is called the Triangle configuration. Again, we must observe the maximum 300km distance for FICON links. To ensure that requirement is still met, we need to examine the worst-case scenarios in the event of path loss. If the shortest path between adjacent switches were lost, then traffic would have to traverse the two longer paths to reach its targets. In the picture below, if we assume $\overline{BC}$ is the shortest distance of the three, then $\overline{AB} + \overline{AC} <= 300$km.

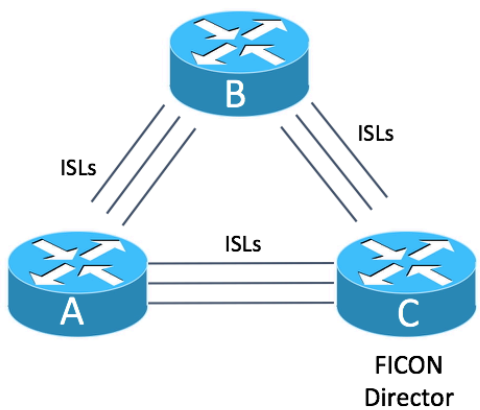

Figure 3: Triangle Configuration

Three Hop, Four Switch, Configurations

Next, we explore the combinations of switching topologies, but restricting them to four switching devices. These environments are considering three-hop configurations, because at most you may traverse three switches and/or directors to reach your target.

Three-Hop Array

In Figure 4, three switches that are connected linearly. This is called the Three-Hop Array configuration. In this configuration, the maximum distance of $\overline{AB} + \overline{BC} + \overline{CD} <= 300$km.

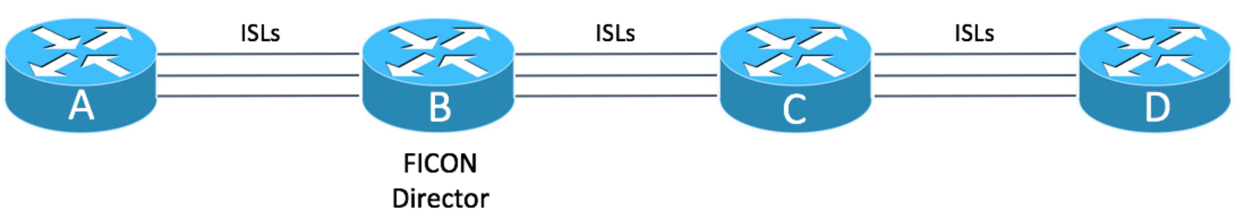

Figure 4: Three-Hop Array Configuration

The Square

The next configuration is four switches connected in a square configuration, as seen in Figure 5. This is called the Square configuration. Again, we must observe the maximum 300km distance for FICON links. To ensure that requirement is still met, we need to examine the worst-case scenarios in the event of path loss. If the shortest path between adjacent switches were lost, then traffic would have to traverse the three longer paths to reach its targets. In the picture below, if we assume $\overline{BC}$ is the shortest distance of the four, then $\overline{AB} + \overline{AD} + \overline{CD} <= 300$km.

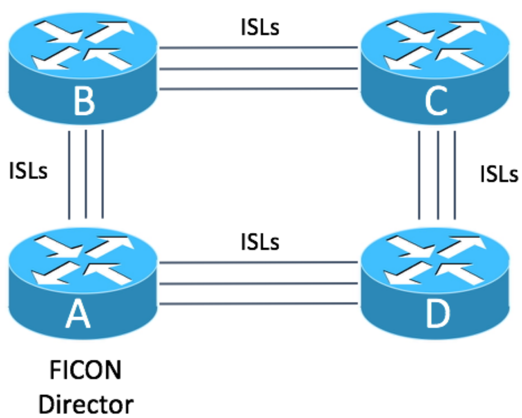

Figure 5: Square Configuration

FCIP Configurations

The next set of configurations use FCIP technologies to increase the distance of FICON links. These environments are mainly used for asynchronous replication applications. This is mainly because of the latency overhead of traveling across an IP network. The maximum distance supported over an FCIP connection for FICON is 3000km. Only one FCIP bridge is allowed per Multihop configuration.

Two-Hop Array with FCIP Bridge

This configuration has two switches connected with native ISLs and a third switch connected over FCIP, as seen in Figure 6. The maximum distance of $\overline{AB}$ is 300km, and the maximum distance of $\overline{BC}$ is 3000km.

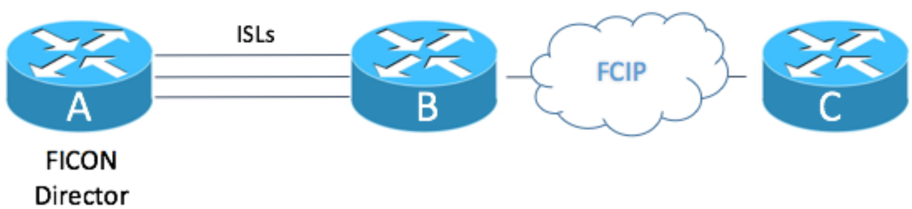

Figure 6: Two-Hop Array with FCIP

Three Hop, Four Switch, Configurations with FCIP

In the four switch configurations, there are three variants. This is due to the location of the FCIP bridge connection and the configuration of the native ISLs.

Three-Hop Array with Middle FCIP Bridge Configuration

In Figure 7, the FCIP bridge is in the middle of the configuration. The maximum distance of the bridge is 3000km. The distance of the native ISLs $\overline{AB} + \overline{CD}$ <= 300km.

Figure 7: Three-Hop Array with Middle FCIP Bridge

Three-Hop Array with Leaf FCIP Bridge Configuration

In Figure 8, the FCIP is position at the last hop. The maximum distance of the bridge is 3000km. The distance of the native ISLs $\overline{AB} + \overline{BC}$ <= 300km.

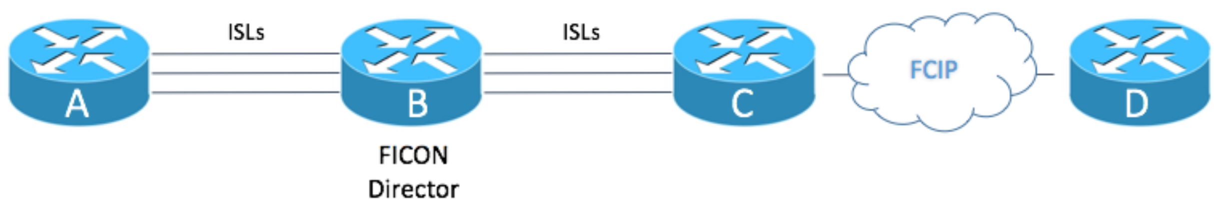

Figure 8: Three-Hop Array with Lead FCIP Bridge

The Triangle with Leaf FCIP Bridge Configuration

The final configuration is a combination of the native ISL Triangle configuration with an FCIP connection to a fourth switch. The maximum distance rules of the native ISL Triangle configuration still apply to the native ISLs in this configuration. The maximum distance of the FCIP bridge is 3000km.

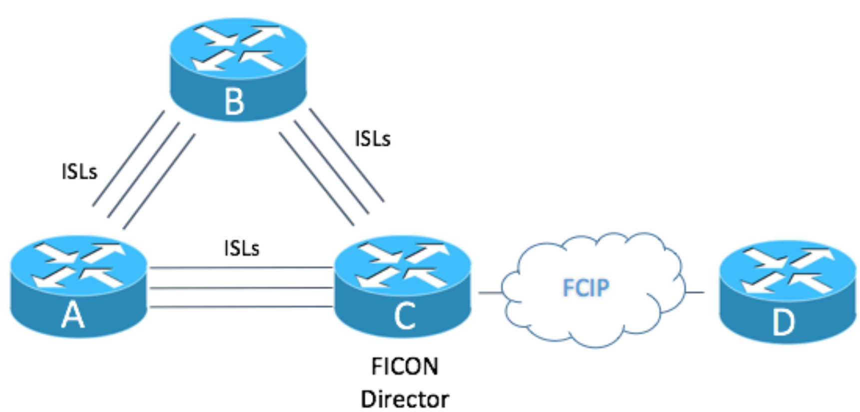

Figure 9: The Triage with Leaf FCIP Bridge

Multihop Benefits, Use Cases and Design Considerations

Multihop Benefits

The characteristics of Multihop environments provide two primary benefits to the end user. Which benefit applies most will depend on your specific circumstances and configuration.

1) Potential configuration consolidation/switch consolidation. The limit of only having two switches/1 hop supported in the past, led to some unnatural FICON switch configurations for clients running multisite configurations. With the new supported Multihop configurations, the management of these configurations is much simpler.

2) Better availability. Using the square and triangle configurations enables the access to higher SAN availability in the event of a total connectivity break between two switches. These configurations now allow traffic to be routed around to ensure you still maintain access to all the devices and hosts in your SAN fabric. Prior to Multi-Hop support, clients with three and/or 4 site architectures experiencing these types of connectivity failures could see themselves being in an unsupported configuration for failover. Multi-hop FICON eliminates this potential issue.

Design Considerations

There are some design implications that need to be considered with the use of FICON Multihop environments. The first is bandwidth planning and allocation for failure scenarios, and the second performance and latency impacts of traffic being re-routed to longer paths.

When planning, and designing your FICON Multihop fabric, you need to keep in mind the additional bandwidth that may be required when there is a total loss of connectivity between two switches. If this event were to occur, then all traffic that was flowing between these two adjacent switches will be rerouted over the non-failed ISL paths. To ensure there is minimal impact to the other traffic in the fabric additional bandwidth should be considered between all routes.

Secondly, if all the traffic is rerouted due to a total connectivity loss between switches, applications using those paths may see a negative impact to performance due to the increased latency of traversing the longer path. Unfortunately, this is the consequence of a high availability design, but there is little that can be done to mitigate this.

Bandwidth Sizing Between Sites, Virtual Fabrics and DWDM Considerations

In any cascaded FICON architecture, it is a best practice to perform a bandwidth sizing study. Such a study can help you determine the bandwidth required for the cascaded links, the number of ISLs required to support that bandwidth requirement, as well as allow you to plan for anticipated storage growth and associated bandwidth growth over a multi-year period. Factors to consider in the bandwidth sizing study include:

1) What type of traffic is going across the cascaded FICON links (ISLs)? DASD, tape, CTC, all the above?
2) Do I have replication traffic going across the ISLs? If so, what type (synchronous, asynchronous, both)?
3) Do I wish to isolate a particular traffic type to its own set of ISLs (by OS, storage type, replication type)?

4) Am I using trunking/port channels in conjunction with Multi-Hop ?
5) Am I using virtual fabrics and how does that effect my ISL allocation?
6) Do I have an SLA that must be met on a specific replication traffic?
7) Does my environment include a GDPS or similar architecture? Will I be performing a HyperSwap?

There are a variety of tools that can be used for this study. Some of the better tools include the Intellimagic family of software.

If you are employing a virtual fabrics architecture, it is recommended that you consult with your switch vendor for specific configuration guidelines on ISL usage with multihop and virtual fabrics.

Finally, if you are employing DWDM devices in a cascaded FICON architecture, check with your switch vendor **AND** DWDM vendor for specifics on Multi-Hop support, as well as support for other cascaded FICON technologies (for example, not every DWDM vendor supports VC_RDY and requires use of R_RDY on ISLs).

Summary

A great deal of technological innovation has occurred over the past twelve years since cascaded FICON was initially introduced by IBM in 2003. These technical innovations have driven the need for further innovations in how the interswitch links manage traffic flows, and how the end user manages the ISLs. FICON Multihop is the latest such management innovation. There are some design implications that need to be considered with the use of FICON Multihop environments, FICON Multihop provides significant configuration flexibility and management improvements over the older FICON single hop policy.

Revision History

4/2017: Initial published version

6/2017: Minor text corrections.

3/2018: Authors update.

