The FICON features support FC and SCSI devices in IBM z/VM, IBM z/VSE, Linux on IBM Z, and the KVM hypervisor, as shown in Figure 3-2.

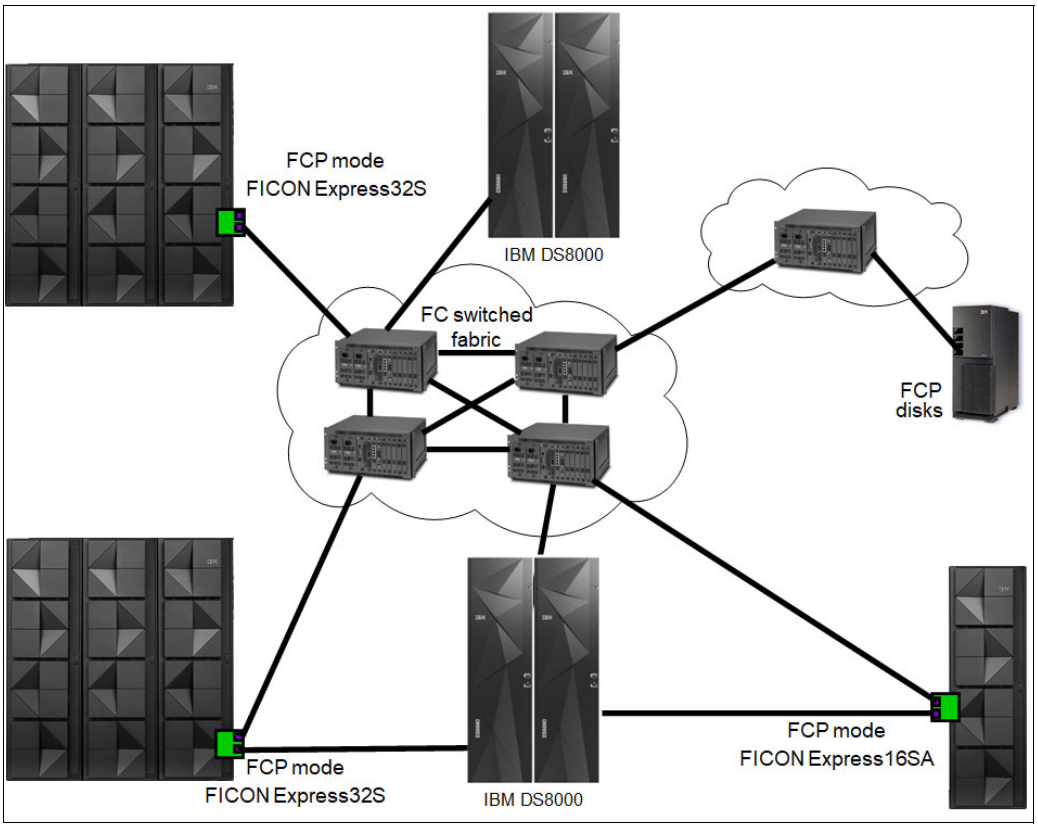

Figure 3-2 IBM Z FCP example topology

With IBM Z, point-to-point connections can be used to access data that is stored on devices without using an FC switch. In addition, an operating system or other stand-alone program can undergo an IPL through a point-to-point connection by using the SCSI IPL feature. N_Port ID Virtualization (NPIV) is not supported by FCP point to point. For more information, see “Worldwide port name tool” on page 39.

The FCP support allows z/VM, Linux on IBM Z and the KVM hypervisor, and z/VSE operating systems to access industry-standard SCSI devices. For disk applications, these FCP storage devices use fixed block sectors rather than the Extended Count Key Data (ECKD) format.

3.1.2 FCP channel

The FC-FCP standard was developed by the International Committee of Information Technology Standards (INCITS) and published as an American National Standards Institute (ANSI) standard. The IBM Z FCP I/O architecture conforms to the FC standards that are specified by the INCITS. For more information about the FC standards, see the INCITS Technical Committee T11 website and their page for SCSI Storage Interfaces (this committee within INCITS is responsible for the FC Interface).

FICON channels in FCP mode provide full fabric attachment of SCSI devices to the operating system images by using the FCP, and point-to-point attachment of SCSI devices. This technique allows z/VM, Linux on IBM Z and the KVM hypervisor, and z/VSE to access industry-standard SCSI storage controllers and devices.

FCP channel full fabric support means that multiple numbers of directors or switches can be placed between the IBM Z platform and the SCSI device. This technique enables many *hops* through a storage area network (SAN) and provides improved use of intersite-connected resources and infrastructure. This expanded ability to attach storage devices provides more choices for storage solutions and the ability to use existing storage devices. This configuration can facilitate the consolidation of UNIX server farms onto the IBM Z platform, which protects investments in SCSI-based storage.

For a list of switches, storage controllers, and devices that are verified to work in an FC network that is attached to FCP channel, and the software requirements to support FCP and SCSI controllers or devices, see the I/O Connectivity website.

FICON channels in FCP mode are based on the FC standards that are defined by INCITS and published as ANSI standards. FCP is an upper-layer FC mapping of SCSI on a common stack of FC physical and logical communication layers.

SCSI is supported by a wide range of controllers and devices, complementing the classical storage attachment capability through FICON channels. FCP is the base for industry-standard FC networks or SANs.

FC networks consist of servers, storage controllers, and devices as end nodes, which are interconnected by FC switches, directors, and hubs. Switches and directors are used to build FC networks or fabrics. Fibre Channel Arbitrated Loops (FC-ALs) can be constructed by using FC hubs. In addition, different types of bridges and routers can be used to connect devices with different interfaces, such as parallel SCSI. All these interconnections can be combined in the same network.

SCSI is implemented by many vendors in many different types of storage controllers and devices. These controllers and devices are widely accepted in the marketplace and have proven to be able to meet the reliability, availability, and serviceability (RAS) requirements of many environments.

FICON channels in FCP mode use the queued direct input/output (QDIO) architecture for communication with the operating system. The QDIO architecture for FCP channels derives from the QDIO architecture, which was defined initially for the OSA-Express features and HiperSockets communications.

FCP channels do not use control devices. Instead, data devices that represent QDIO queue pairs (QPs) are defined, and they consist of a request queue and a response queue. Each QP represents a communication path between an operating system and the FCP channel. A QP allows an operating system to send FCP requests to the FCP channel through the request queue. The FCP channel uses the response queue to pass completion indications and unsolicited status indications to the operating system.

Hardware Configuration Definition (HCD) or Input/Output Configuration Program (IOCP) is used to define the FCP channel type and QDIO data devices. However, there is no definition requirement for the FC storage controllers and devices, or for the FC interconnect units, such as switches, directors, and bridges. The FCP industry standard architecture requires that the FC devices (end nodes) in an FC network are addressed by using worldwide names (WWNs), FC IDs, and logical unit numbers (LUNs).

These addresses are configured on an operating system level and passed to the FCP channel together with the corresponding FC I/O or service request through a logical QDIO device (queue).

Figure 3-3 shows the necessary FCP I/O definitions and compares them to FICON I/O definitions.

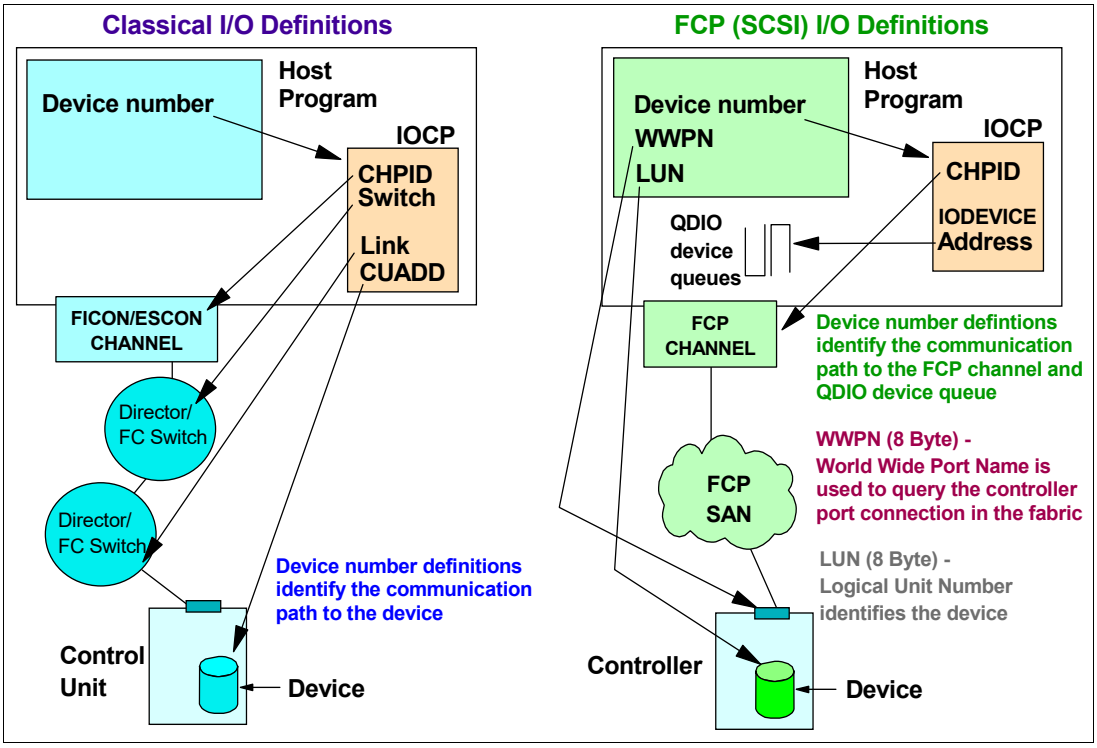

Figure 3-3 I/O definition comparison (FCP to FICON)

Channel and device sharing

An FCP channel can be shared among multiple Linux operating systems, each running in a logical partition (LPAR) or as a guest operating system under z/VM. To access the FCP channel, each operating system needs its own QDIO QP, which is defined as a data device on an FCP channel in the HCD/IOCP.

Each FCP channel can support up to 480 QDIO QPs with the IBM Z platform. This support allows each FCP channel to be shared among 480 operating system instances (with a maximum of 252 guests per LPAR).

Host operating systems that share access to an FCP channel can establish up to 2048 concurrent connections to up to 512 different remote FC ports that are associated with FC controllers. The total number of concurrent connections to end devices, which are identified by LUNs, must not exceed 4096.

Although multiple operating systems can concurrently access the same remote FC port through a single FCP channel, FC devices, which are identified by their LUNs, can be reused only serially. For two or more unique operating system instances to share concurrent access to a single FC or SCSI device (LUN), each of these operating systems must access this device through a different FCP channel.

If two or more unique operating system instances attempt to share concurrent access to a single FC or SCSI device (LUN) over the same FCP channel, a LUN sharing conflict occurs and errors result. A way to alleviate this sharing conflict on the IBM Z platform is to use NPIV.

Worldwide port name tool

The worldwide port name (WWPN) tool assigns WWPNs to each FCP channel or port by using the same WWPN assignment algorithms that a system uses when you are assigning WWPNs for channels that use NPIV. Therefore, the SAN can be set up in advance, allowing operations to proceed faster after the system is installed.

WWPN tool: The WWPN tool is supported in IBM Processor Resource/Systems Manager (PR/SM) mode. DPM uses a different assignment algorithm for generating and assigning NPIV WWPNs.

The WWPN tool can calculate and show WWPNs for both virtual and physical ports ahead of system installation. This feature means that the SAN configuration can be retained rather than altered by assigning a WWPN to physical FCP ports when a FICON feature is replaced.

The WWPN tool takes an adhesive file that contains the FCP-specific I/O device definitions and creates the WWPN assignments that are required to set up the SAN. A binary configuration file that can be imported later by the system is also created. The CSV (`.csv`) file can be either created manually or exported from the HCD or Hardware Configuration Manager (HCM).

The WWPN tool can be downloaded from the Tools section of IBM Resource Link (requires registration).

FCP SCSI IPL feature enabler

This function runs an IPL of an operating system from an FCP channel-attached disk that is either in an LPAR or as a guest operating system under z/VM. SCSI IPL can directly run an IPL of a Linux operating system that was installed previously on a SCSI disk. Therefore, there is no need for a classical, channel-attached device (FICON), such as an ECKD disk CU, to install and a Linux operating system and run an IPL of it.

The IPL device is identified by its SAN address, which consists of the WWPN of the disk controller and the LUN of the IPL device.

Important: If a second-level z/VM system undergoes an IPL from an FCP SCSI LUN, a minimum virtual memory is required, which depends on the model of the processor on which the z/VM system is running. To ensure success on all processor models, you should define at least 768 MB of virtual storage.

SCSI IPL is supported in the following conditions:

- FCP access control
- Point-to-point connections
- NPIV

A *stand-alone-dump program* can also undergo an IPL from an FCP channel that is attached to a SCSI disk. The stand-alone-dump program can also store the generated dumped data on a SCSI disk. z/VM support of SCSI IPL allows Linux and other guest operating systems that support this feature to undergo an IPL from an FCP-attached SCSI disk when z/VM is running on an IBM Z platform. Therefore, Linux guests can be started and run from an FCP channel-attached disk.

FCP multipathing concept

Multipath access to disk subsystems is a basic feature with the channel subsystem (CSS) on the IBM Z platform. FICON connections support multiple hardware paths to any physical disk device. The IBM Z platform handles multipathing invisibly through the operating system. With FICON channels, the operating system is presented with a single device for I/O operations, and multipathing happens under CSS control.

Multipathing over FCP is different. With FCP multipathing on Linux on IBM Z, each path to each LUN appears to the operating system as a separate device. For example, if there are four paths to five LUNs, the Linux kernel defines 20 SCSI devices.

At the time of writing, supported distributions use device-mapper multipathing in the Linux kernel along with multipath-tools in the user space. For more information, see the corresponding distribution documentation and How to use FC-attached SCSI devices.

FCP access control

The ability to control access to nodes and devices is provided as a function in switches and controllers. It is called *LUN masking* and *zoning*, which can be used to prevent systems from accessing storage that they are not permitted to access:

LUN masking A LUN represents a portion of a controller, such as a disk device. With the use of LUNs, a controller can be logically divided into independent elements or groups of elements. Access to these LUNs can be restricted to distinctive WWPNs as part of the controller configuration. This method is known as *LUN masking*.

Zoning Segmentation of a switched fabric is achieved through *zoning*. It should be used to fence off certain portions of the switched fabric, allowing only the members of a zone to communicate within that zone. All others that attempt to access from outside of that zone are rejected.

I/O devices

The IBM Z FCP channel implements the FCP standard as defined by the INCITS Fibre Channel Protocol for SCSI (FC-FCP) and Fibre Channel Protocol for SCSI Second Version (FCP-2), and the relevant protocols for the SCSI-2 and SCSI-3 protocol suites. Theoretically, each device that conforms to these protocols works when attached to an IBM Z FCP channel. However, experience shows that there are small deviations in the implementations of these protocols.

Also, for certain types of FCP and SCSI controllers and devices, specific drivers in the operating system might be required to use all capabilities of the controller or device. The drivers might also be required to cope with unique characteristics or deficiencies of the device.

Note: Do appropriate conformance and interoperability testing to verify that a storage controller or device can be attached to an IBM Z FCP channel in a configuration. For example, test that it can be attached through a type of FC switch, director, or point-to-point connection.

Hardware assists for z/VM guests

A complementary virtualization technology is available for the IBM Z platform. The technology includes these capabilities:

- QDIO Enhanced Buffer-State Management (QEBSM), with two hardware instructions that are designed to eliminate the overhead of hypervisor interception.
- Host Page-Management Assist (HPMA), which is an interface to the z/VM main storage management function that allows the hardware to assign, lock, and unlock page frames without z/VM hypervisor assistance.

These hardware assists allow a cooperating guest operating system to start QDIO operations directly to the applicable channel without interception by z/VM, which provides more performance improvements. This support is integrated into the IBM Z platform. Consult the appropriate Preventive Service Planning (PSP) buckets (3931DEVICE, 8561DEVICE, 3906DEVICE, 3932DEVICE, 8562DEVICE, or 3907DEVICE) before implementation.

Support of T10-Data Integrity Field for enhanced reliability

Because high reliability is important for maintaining the availability of business-critical applications, the IBM Z FCP supports the ANSI T10 Data Integrity Field (DIF) standard. Data integrity protection fields are generated by the operating system and propagated through the SAN. IBM Z helps to provide added end-to-end data protection between the operating system and the storage device.

An extension to the standard that is called Data Integrity Extensions (DIX) provides checksum protection from the application layer through the host bus adapter (HBA), where cyclical redundancy check (CRC) protection is implemented.

T10-DIF support by the FICON features, when defined as CHPID type FCP, is available on the IBM Z platform. Using the T10-DIF standard requires support by the operating system and the storage device.

3.1.3 FCP and FICON mode characteristics

The single largest difference between the FICON channel and FCP channel mode types is the treatment of data access control and security. FICON channels rely on a multiple image facility (MIF) to address concerns about shared channels and devices. MIF provides ultra-high access control and security of data so that one operating system image and its data requests cannot interfere with another operating system's data requests. With the introduction of IBM Z, MIF continues this ultra-high level of access control and security across CSSs.

FCP and MIF

Linux guest operating systems under z/VM can have access to an FCP channel defined to the z/VM operating system. Using MIF, an FCP channel can also be shared between Linux LPARs and z/VM LPARs with Linux guests.

The FCP industry-standard architecture does not use the data access control and security functions of MIF. As a result, FCP has the following limitations:

- Channel sharing

  When NPIV is not implemented, and if multiple Linux images share an FCP channel and all Linux images have connectivity to all devices that are connected to the FCP fabric, all Linux images use the same WWPN. They use this name to enter the fabric, and they are indistinguishable from each other within the fabric. Therefore, the usage of zoning in switches and LUN masking in controllers is not effective in creating appropriate access controls among the Linux images.

  By using NPIV, each operating system that shares an FCP channel is assigned a unique WWPN. The WWPN can be used for *device-level* access control in storage controllers (LUN masking) and in *switch-level* access control (zoning).

- Device sharing

  Without using NPIV, an FCP channel prevents logical units from being opened by more than one Linux image at a time. Access is on a first-come, first-served basis. This system prevents problems with concurrent access from Linux images that share an FCP channel and the same WWPN. This serialization means that one Linux image can block other Linux images from accessing the data on one or more logical units unless the sharing images (z/VM guests) are not in contention.

FICON versus FCP

FICON and FCP have other significant differences. Certain differences are fundamental to the IBM Z family, and others are fundamental to the two channel architectures. Others depend on the operating system and the storage device being attached. Without taking the operating system and the storage device into consideration, I/O connectivity through IBM Z FCP and FICON channels has the following differences:

- Direct connection

  With all the FICON features on the IBM Z platform, storage controllers can be directly connected to the channel by using point-to-point attachment when in FCP mode. There is no need for a director or switch between the FCP channel and storage controllers.

  Note: NPIV is supported in a switched topology, and FCP with NPIV is *not* supported in a point-to-point topology.

- Switch topology

  FCP channels support full fabric connectivity, which means that several directors or switches can be used between a IBM Z platform and the device. With the FICON cascaded director support, the FICON storage network topology is limited to a two-director, single-hop configuration.

- Enterprise fabric

  The usage of cascaded FICON Directors ensures the implementation of a high-integrity fabric. For FCP, a high-integrity fabric solution is not mandatory, although it must be considered. For example, if an FCP Inter-Switch Link (ISL) must be moved, data might potentially be sent to the wrong path without notification. This type of error does not happen on an enterprise fabric with FICON.

- Transmission data checking

  When a transmission is sent through an FCP channel, because of its full fabric capability, FCP checks data for each leg of that transmission. FICON also checks intermediate data.

- Serviceability:
  - Licensed Internal Code (LIC) updates and the IBM Z platform itself allow concurrent FCP fixes. FICON channels, when configured as CHPID type FCP, support concurrent fixes, allowing the application of a LIC without requiring a configuration of off/on. This exclusive FCP availability feature is available with all FICON features.
  - The FICON features have Small Form-factor Pluggable (SFP) optics to permit each channel to be individually serviced during a fiber optic module failure. The traffic on the other channels on the same feature can continue to flow if a channel requires servicing.
- Problem determination:
  - Request Node Identification (RNID)

    RNID assists with the isolation of FICON detected cabling errors. Resolving fiber optic cabling problems can be a challenge in a fiber optic environment with extended distances. To facilitate resolution, the operating system can request the RNID data for each device or CU that is attached to native FICON channels. Then, you can display the RNID data by using an operator command. RNID is available to the IBM Z platform and is supported by all FICON features (CHPID type FC) and IBM z/OS.
  - Link incident reporting

    Link incident reporting is integral to the FICON architecture. When a problem on a link occurs, this mechanism identifies the two connected nodes between which the problem occurred, which leads to faster problem determination and service. For FCP, link incident reporting is not a requirement for the architecture, although it might be offered as an optional switch function. Therefore, important problem determination information might not be available if a problem occurs on an FCP link.

    IBM Z allows z/OS to register for all FICON link incident records. This feature improves your ability to capture data for link incident analysis across multiple systems.
  - Simplified problem determination

    To more quickly detect fiber optic cabling problems in a SAN, all FICON channel error information is forwarded to the Hardware Management Console (HMC). This function facilitates detection and reporting of trends and thresholds for the channels with aggregate views, including data from multiple systems.

    Problem determination can be simplified by using the HMC to pinpoint fiber optic cabling issues in your SAN fabric without involving IBM service personnel.

    All FICON channel error information is forwarded to the HMC. In the HMC, this information is analyzed to detect and report the trends and thresholds for all FICON channels on the IBM Z platform. The Fibre Channel Analyzer task on the HMC can be used to display analyzed information about errors on FICON channels (CHPID type FC) of attached Support Elements (SEs). Data includes information about the physical channel ID (PCHID), CHPID, channel type, source link address, and destination link address where the error occurred. This report shows an aggregate view of the data and can span multiple systems.

    Starting with IBM z13®, similar FICON problem determination tools were implemented for FCP channels. These channel problem determination tools for FCP channels include functions such as analyze channel information, subchannel data, device status, serial link status, and link error statistic block. In addition to the analyze functions, fabric status login and SAN explorer functions are also available. These FCP problem determination tools are accessed from the HMC in the same way as for the FICON channels.

- FICON purge path extended

  The purge path extended function enhances FICON problem determination. The FICON purge path error recovery function is extended so that it transfers error-related data and statistics between the channel and entry switch and the CU and its entry switch to the host operating system. FICON purge path extended use requires a switch or device that supports this function. The purge path extended function for FC channels is available on IBM z16, IBM z15, and IBM z14.

- FICON error recovery

  IBM Z platform, z/OS, and I/O recovery processing are designed to allow the system to detect switch or director fabric problems that might cause FICON links to fail and recover multiple times in a short time.

  This feature allows the system to detect these conditions and keep an affected path offline until an operator action is taken. This process is expected to limit the performance impacts of switch or director fabric problems. The FICON error recovery function is available in z/OS.

Forward Error Correction

Forward Error Correction (FEC) is a technique that is used for controlling errors in data transmission over unreliable or noisy communication channels. By adding redundancy and error-correcting code (ECC) to the transmitted information, the receiver detects and corrects a limited number of errors in the information instead of requesting a retransmission. This process improves the reliability and bandwidth utilization of a connection and reduces retransmissions due to bit errors. This advantage is true especially for connections across long distances, such as an ISL in an IBM Geographically Dispersed Parallel Sysplex (IBM GDPS®) Metro Mirror environment.

FICON Express16SA, FICON Express16S+, and FICON Express16S support FEC coding on top of their 64 b/66 b data encoding for 16 Gbps connections. Their FEC design can correct up to 11 bit errors per 2112 bits that are transmitted. FICON Express32G uses 256b/257b encoding and can correct up to 20 bit errors per 5140 bits that are transmitted. Thus, when connected to devices that support FEC at 16 or 32 Gbps connections, the FEC design allows FICON Express channels to operate at higher speeds over longer distances and with reduced power and higher throughput. Concurrently, the FEC design maintains the same reliability and robustness that FICON channels are known for.

With IBM DS8870 or later, the IBM z16, IBM z15, IBM z14, and IBM z14 ZR1 can extend the usage of FEC to the fabric N_Ports[1] for a completed end-to-end coverage of 16 or 32 Gbps FC links. For more information, see *IBM DS8900F and IBM Z Synergy DS8900F: Release 9.3 and z/OS 2.5*, REDP-5186.

FICON Dynamic Routing

With the IBM z16, IBM z15, IBM z14 and IBM z14 ZR1, FICON channels are no longer restricted to the usage of static SAN routing policies for ISLs for cascaded FICON directors. IBM Z now support dynamic routing in the SAN with the FICON Dynamic Routing (FIDR) feature. It supports the dynamic routing policies that are provided by the FICON director manufacturers, such as Brocade Exchange Based Routing 7 (EBR 7) and Cisco Open Exchange ID Routing (OxID).

[1] Node ports

With FIDR, IBM z16, IBM z15, IBM z14, and IBM z14 ZR1 have advantages for performance and management in configurations with ISL and cascaded FICON directors:

- Support sharing of ISLs between FICON and FCP (Peer-to-Peer Remote Copy (PPRC) or distributed)
- Better balanced I/O traffic between all available ISLs
- Improved utilization of the FICON director and ISL
- Easier management with a predicable and repeatable I/O performance

FIDR can be enabled by defining dynamic routing capable switches and CUs in HCD. Also, z/OS has implemented a health check function for FIDR.

FICON performance

For more information about FICON and FCP performance, see the IBM server connectivity web page.

3.2 FICON elements

FICON enables multiple concurrent I/O operations to occur simultaneously to multiple CUs. FICON channels also permit intermixing of large and small I/O operations on the same link. The data center I/O configuration now has increased flexibility for connectivity because of the increased I/O rate, increased bandwidth, and multiplexing of mixed workloads.

3.2.1 FICON channel

FICON channel architecture is compatible with the following protocols:

- Fibre Channel Physical and Signaling standard (FC-FS)
- Fibre Channel Switch Fabric and Switch Control Requirements (FC-SW)
- Fibre Channel Single-Byte-3 (FC-SB-3) and Fibre Channel Single-Byte-4 (FC-SB-4) standards

Cabling specifications are defined by the Fibre Channel - Physical Interface - 4 (FC-PI-4) standard and used by IBM Z FICON features. Table 3-1 identifies cabling types and link data rates that are supported in the FICON environment, which include their allowable maximum distances and link loss budget. The link loss budget is derived from the channel insertion loss budget that is defined by the FC-PI-4 standard (Revision 8.00).

Table 3-1 Fiber optic cabling for FICON: Maximum distances and link loss budget

FC-PI-4 Fiber core	Cable type	2 Gbps	4 Gbps	8 Gbps	16 Gbps	32 Gbps	10 Gbps ISL[a]
		Distance / Link-loss budget (decibels (dB))	Distance / Link-loss budget (dB)	Distance / Link-loss budget (dB)	Distance / Link-loss budget (dB)	Distance / Link-loss budget (dB)	Distance / Link-loss budget (dB)
9 µm SM	OS1/ OS2	10 km / 7.8	10 km / 7.8	10 km / 6.4	10 km / 6.4	10 km / 6.34	10 km / 6.4
9 µm SM	OS1/ OS2	4 km / 4.8	4 km / 4.8	N/A	N/A	N/A	N/A
50 µm MM	OM4	500 m / 3.31	400 m / 2.95	190 m / 2.19	125 m / 1.95	100 m / 1.86	N/A
50 µm MM	OM3	500 m / 3.31	380 m / 2.88	150 m / 2.04	100 m / 1.86	70 m / 1.75	300 m / 2.6
50 µm MM	OM2	300 m / 2.62	150 m / 2.06	50 m / 1.68	35 m / 1.63	20 m / 1.57	82 m / 2.3
62.5 µm MM	OM1	150 m / 2.1	70 m / 1.78	21 m / 1.58	N/A	N/A	N/A

a. ISL between two FICON Directors.

Note: IBM does not support a mix of 50 µm and 62.5 µm fiber optic cabling in the same physical link.

When an application performs an I/O operation to a device that is represented by a unit control block (UCB), it initiates an I/O request by using macros or a Supervisor Call to the Input/Output Supervisor (IOS). The application or access method also provides the channel program (channel command words (CCWs)) and extra parameters in the operation request block (ORB). This request is queued on the UCB. The IOS services the request from the UCB on a priority basis.

Then, the IOS issues a start subchannel (SSCH) instruction with the subsystem identifier (SSID) that represents the device and the ORB as operands. The CSS is signaled to perform the operation. This flow is shown in Figure 3-4 on page 47.

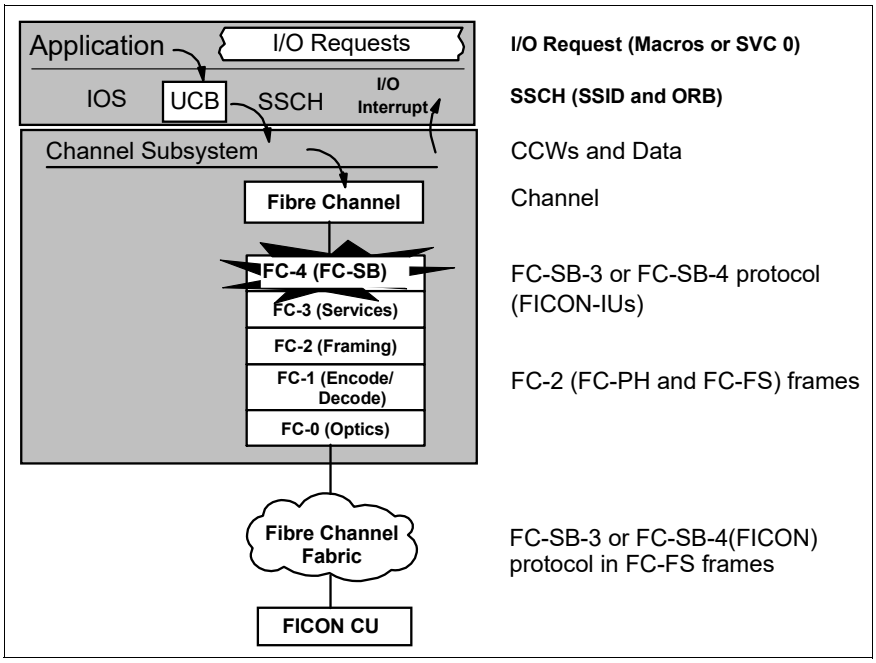

Figure 3-4 FICON channel operation flow

The most appropriate FICON channel is selected by the CSS. The FICON channel fetches channel programs (CCWs) that are prepared by the application, fetches data from memory (write) or stores data into memory (read), and presents the status of the operation to the application (I/O interrupt).

The z/Architecture channel commands, data, and status are packaged by the FICON channel into FC-SB-3 or FC-SB-4 (FC-4 layer) Information Units (IUs). IUs from several different I/O operations to the same or different CUs and devices are multiplexed or demultiplexed by the FC-2 layer (framing). These FC-2 frames with encapsulated FC-SB-3 or FC-SB-4 IUs are encoded or decoded by the FC-1 layer (encode or decode) and sent to or received from the FC-0 fiber optic medium (optics).

On a FICON channel, CCWs are transferred to the CU without waiting for the first command response from the CU or for a CE/DE after each CCW execution. The device presents a logical *end* to the CU after each CCW execution. If the last CCW of the CCW chain has been run by the device, the CU presents CE/DE to the channel. Figure 3-5 shows a CCW operation on a FICON channel that uses CCW chaining.

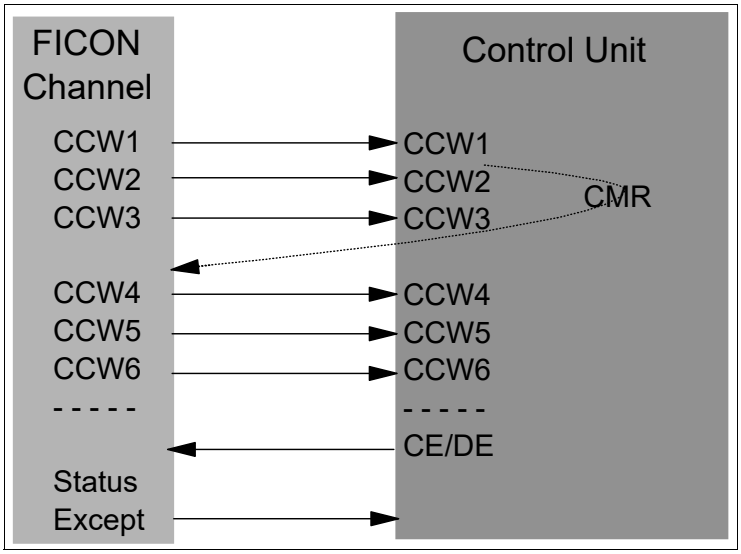

Figure 3-5 CCW chaining

3.2.2 IBM High-Performance FICON for IBM Z

IBM High-Performance FICON for IBM Z (zHPF) is an enhancement of the FICON channel architecture that is compatible with these protocols:

- FC-FS standard
- FC-SW
- FC-SB-4 standard

Using zHPF with the FICON channel, the z/OS operating system, and the CU reduces the FICON channel overhead. This goal is achieved by protocol optimization and reducing the number of IUs processed, which results in more efficient usage of the fiber link.

The FICON Express32S, FICON Express16SA, FICON Express16S+, FICON Express16S, and FICON Express8S features support both the existing FICON architecture and the zHPF architecture. From the z/OS point of view, the existing FICON architecture is called *command mode*, and the zHPF architecture is called *transport mode*. A parameter in the ORB is used to determine whether the FICON channel is running in command or transport mode.

The mode that is used for an I/O operation depends on the CU that is supporting zHPF and the settings in the z/OS operating system. An **`IECIOSxx`** parameter and **`SETIOS`** commands in z/OS can enable or disable zHPF. Support is also added for the **`D IOS, ZHPF`** system command to indicate whether zHPF is enabled, disabled, or not supported on the system.

During link initialization, both the channel and the CU indicate whether they support zHPF. The Process Login (PRLI) support indicator is presented in response to the RNID Extended Link Services (ELS). If PRLI is supported, the channel sends a PRLI ELS. Then, the PRLI response indicates that zHPF is supported by the CU.

Like the existing FICON channel architecture, the application or access method provides the channel program (CCWs) and parameters in the ORB. Bit 13 in word 1 of the ORB specifies how to handle the channel program in either command mode or transport mode.

The way that zHPF transport mode manages CCW operation is different from the CCW operation for the existing FICON architecture command mode, as shown in Figure 3-6. In command mode, each single CCW is sent to the CU for execution. In transport mode, all CCWs are sent over the link in one single frame to the CU. Certain complex CCW chains are not supported by zHPF. Figure 3-6 shows an example of the optimization by a zHPF transport mode read operation.

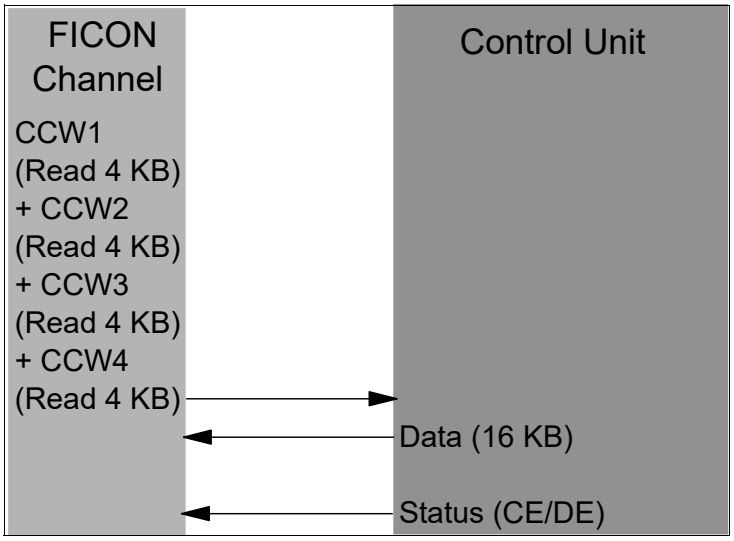

Figure 3-6 High-performance FICON read operation

The channel sends all the required CCWs and read operations of 4 KB of data in one single frame to the CU. The CU transfers the requested data over the link to the channel, followed by a CE/DE if the operation was successful. Less overhead is generated compared with the existing FICON architecture.

Figure 3-7 shows the same reduction of frames and open exchanges for a zHPF transport mode write operation.

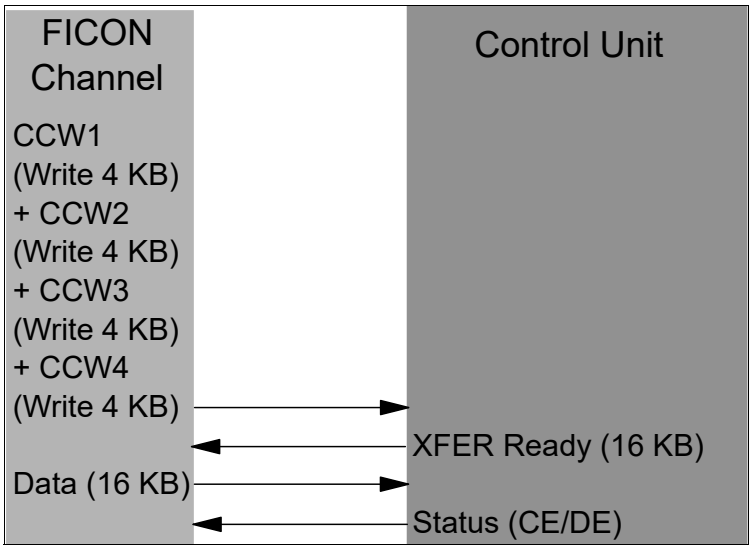

Figure 3-7 High-performance FICON write operation

The channel sends all the required CCWs and write operations of 4 KB of data in one frame to the CU. The CU responds with XFER when it is ready to receive the data. The channel then sends the 16 KB of data to the CU. If the CU successfully receives the data and finishes the write operation, the CE/DE status is sent by the CU to indicate the completion of the write operation.

zHPF supports multi-track operations. It allows the channel to operate at rates that fully use the bandwidth of a FICON Express channel. The zHPF fully supports multiple tracks of data that can be transferred in a single operation.

For more information about the FICON channel protocol and zHPF, see *FICON Planning and Implementation Guide*, SG24-6497.

3.2.3 Platform and name server registration in FICON channel

All FICON features on the IBM Z platform support platform and name server registration to the fabric. That support exists only if the FICON feature is defined as CHPID type FC.

Information about the channels that are connected to a fabric, if they are registered, allow other nodes or SAN managers to query the name server to determine what is connected to the fabric. The following attributes are registered for the IBM Z platform:

- Platform information:
  - Worldwide node name (WWNN). This name is the node name of the platform and it is the same for all channels that belong to the platform.
  - Platform type.
  - Host computer type.
  - Platform name. The platform name includes vendor ID, product ID, and vendor-specific data from the node descriptor.
- Channel information.
- WWPN.

- Port type (N_Port_ID).
- FC-4 types supported.
- Classes of services that are supported by the channel.

The platform and name server registration service are defined in the Fibre Channel - Generic Services 4 (FC-GS-4) standard.

3.2.4 Open exchanges

An *open exchange*, which is part of FICON and FC terminology, represents an I/O operation in progress over the channel. Many I/O operations can be in progress over FICON channels at any one time. For example, a disk I/O operation might temporarily disconnect from the channel when performing a seek operation or while waiting for a disk rotation. During this disconnect time, other I/O operations can be managed as follows:

- Command mode open exchanges

  In command mode, the number of open exchanges is limited by the FICON Express feature. FICON Express32S, FICON Express16SA, FICON Express16S+, FICON Express16S, and FICON Express8S allow up to 64 open exchanges. One open exchange (an exchange pair) in command mode is the same as one I/O operation in progress.

- Transport mode open exchanges

  In transport mode, one exchange is sent from the channel to the CU. Then, the same exchange ID is sent back from the CU to the channel to complete the I/O operation. The maximum number of simultaneous exchanges that the channel can have open with the CU is 750 exchanges. The CU sets the maximum number of exchanges in the status area of the transport mode response IU. The default number is 64, which can be increased or decreased.

In addition, FICON channels can multiplex data transfer for several devices concurrently. This feature also allows workloads with low to moderate CU cache hit ratios to achieve higher levels of activity rates per channel.

If the open exchange limit is reached, more I/O operations are refused by the channel, which can result in queues and retries by the operating system.

Extended distances

Degradation of performance at extended distances can be avoided by implementing an enhancement to the industry standard FICON architecture (FC-SB-3). This enhancement is a protocol for persistent IU pacing. CUs that use the architecture can increase the pace count, which is the number of IUs that are allowed to be underway between the channel and the CU. Extended distance FICON channels retrieve the last pacing information and use this information for later operations. This feature avoids performance degradation at the start of a new operation.

IU pacing helps to optimize the link usage and simplifies the requirements for channel extension equipment because more commands can be in-flight. Extended distance FICON is apparent to the operating systems and it is applicable to all FICON features that are defined with CHPID type FC.

Modified Indirect Data Address Word

On IBM Z, Modified Indirect Data Address Word (MIDAW) provides alternatives to using CCW data chaining in channel programs. The MIDAW facility was added to z/Architecture and can coexist with the current CCW IDAW facility.