Today’s Presenters

Joe Kimpler
ATTO Technologies

Patty Driever
IBM

Howard Johnson
Broadcom
Agenda

• Brief Review of Relevant FICON 101 Concepts
  – Introduction
• Mapping of FICON (FC-SB-x) on to Fibre Channel FC-2 layer
  – The Basics of the ULP Mapping
• Evolution of FICON protocol optimizations
  – Persistent IU Pacing
  – MIDAWs
  – Command Mode vs Transport Mode
• Q&A
Brief Review of FICON 101
Characteristics of the Mainframe

- Integrity
- Security
- Flexibility
- Availability
- Serviceability
- Transactions
- Efficiency
- Reliability

Operating System
Partition
Operating System
Partition
Operating System
Partitions sharing physical resources
Hypervisor
Partition
Hardware Hypervisor
Mainframe Hardware

Mainframe Hardware
System Architecture – The I/O Operation

• Start Subchannel parameters:
  – Subchannel number => target device
  – Pointer to an Operation Request Block (ORB)
    • Pointer to sequence of Channel Command Words (CCWs)
      – Command
      – Data address
      – Quantity of data
      – Control conditions (flags)
        » e.g. command chaining, conditional execution, and the ability to construct a command chain in discontiguous storage
Link Architecture

- Exists on both the System and Device ends
- Translation
  - System/device architecture to link transport layer
- Describes behaviors required to execute I/O operations
  - Both host and storage ends of the link
- Describes:
  - Link Initialization process
  - Unique Link Level functions
  - How individual I/O operations are handled at this layer
  - Link level error detection
  - Link level recovery
Mapping FICON to Fibre Channel
Fibre Channel Single-Byte Command Code Sets

- FC-SB
  - ESCON
- FC-SB-2 (12/2000)
  - Mapping of command mode
- FC-SB-3 (2/2007)
  - Persistent IU pacing
    - Amendment to existing standards
- FC-SB-4 (9/2009)
  - Definition of transport mode
- FC-SB-5 (5/2012)
  - FC4 Link Service to aid in discovery and auto-configuration
  - Bi-directional transport mode operations
- FC-SB-6 (8/2016)
  - Transport mode extended distance support
  - Introduction of extended CRC Offset Block
  - Added conditions for sending Purge Path Extended
Information Unit (IU)

- An IU is defined in FC-FS as an “organized collection of data specified by an upper level to be transferred as a single Sequence by FC-2”

- FC-SB IUs contain FC-SB:
  - device-level commands
  - status
  - data
  - data descriptor or control information, or
  - FC-SB-6 link control information

- FC-SB IUs are sent as FC-4 device-data frames (FC-FS routing control bits set to ‘0000’b)

- FICON (Command Mode) makes use of 5 of the FC-FS Information categories
  - Unsolicited command
  - Solicited control
  - Unsolicited control
  - Solicited data
  - Unsolicited data
**Sequence of an I/O Operation**

**CPU**
- Start Subchannel instruction issued
  - Target device
  - Sequence of commands (CCWs) to be executed and associated data areas

**Channel Subsystem**
- Channel Path Selection

**FICON Channel**
- Fetches command (CCW) sequence
- Builds Fibre Channel IUss containing Device Information Blocks
  - Command
  - Command Data
  - Data
  - Status
  - Control
  - Link Control

- IUs are then passed to the Fibre Channel FC-2 layer where Fibre Channel frames are built
- FC-1/FC-0 layers send the frames on the link if buffer-to-buffer credit is available
Each concurrent FICON I/O operation uses two Fibre Channel Exchanges

- One unidirectional Exchange for IUs from the channel to the storage controller
- A separate unidirectional Exchange for IUs from the storage controller to the channel

The **PAIR** is commonly known as a “FICON Exchange”
Packaging CCWs into IUs

Commands and data are packaged into information units as specified in the FICON Upper Level Protocol (FC-SB)

- A single CCW command is in an IU
- The first IU sent for an I/O operation (CCW sequence) will open an IU exchange between the sender and receiver
- A CCW command and its associated write data can both be in the same IU, up to a maximum of 8K
- As many IUs are built for a single CCW and its associated data as required

Using a technique called ‘pipelining’, multiple CCWs in the sequence can be transferred to the control unit without waiting for a response to each command
Building IUs Into Frames

FC-SB general IU payload structure for command-mode IUs consists of:

- **FC-SB Header (8 bytes)**
  - Provides SB-x addressing info to identify the logical path and device for the exchange

- **IU Header (8 bytes)**
  - Provides SB-x control flags (including DIB type) and necessary info to associate an IU to a specific CCW
  - DIB types = Command/Data/Command-Data/Status/Control/Link Control

- **Device Information Block (DIB)**
  - DIB Header (12 bytes) – includes IU Count and DIB Data byte count
  - LRC (4 bytes)
  - DIB Data (0-8160 bytes)

FC-2 frames are built from the IU request and IU data:

- Maximum frame packet size of 2112 bytes (528 words)
- As many frames will be sent to transfer an IU as it takes to hold the IU data
- FC-2 Header ‘Type’ field is set to ‘1B’ for IUs from a channel and ‘1C’ for IUs from a control unit
Evolution of the FICON Protocol
FICON “Leaps” and “Bounds”

FICON (and, therefore, FC-SB) evolved to improve the “IS FASTER” characteristics:

– The evolution of Persistent IU pacing to SB-5/6 improved “Flexibility, Availability, and Serviceability”

– How the use of IDAWs/MIDAWs/TIDAWs of the S/390 System Architecture affects FICON efficiency on the link

– The transition from Command mode to Transport mode improved “Transactions, Efficiency, and Reliability”

All while maintaining “Integrity and Security”
Persistent IU Pacing
Default IU Pacing

- **Pipelining** = transfer of CCW commands (and associated data) without waiting for responses to prior commands
- **IU pacing credit** = maximum number of IUs that a channel may send on a given outbound exchange before it receives a Command-Response IU
  - default value is 16 IUs
  - CU may request an increase in IU pacing credit
    - *Remains in effect for duration of the exchange*

- Command-Response IU can be requested on any command sent within IU Pacing Count
  - By setting CRR bit in the command flags

- CU/Device responds with a CMR:
  - at the start of the 1st CCW execution
  - when it sees one has been requested

- On receipt of the pacing CMR the channel will send more IUs (if there are more IUs to send)
Persistent IU Pacing

- Increased pacing counts help performance at distance

- Distance between a host and storage controller for an individual logical path is a constant

- **Persistent IU pacing**
  - CU requests modified IU pacing count
  - Takes effect on next channel program on the logical path
    - *Persist until changed by CU or system reset is performed on the logical path*

- Support for concurrent enablement of persistent IU pacing is indicated in the Node Descriptor

Can be requested in a Command Response IU or when sending in status.
IDAWs to MIDAWs to TIDAWs
Data Transfer

- I/O operations (SSCH instructions) are used to transfer data to or from an I/O device, to or from host main storage.
- The ‘channel program’ (CCW chain) guides this data transfer.

- Start at the CCW data address location
- End when the CCW count is zero
Data Transfer

- A single CCW operation is very efficient when the data is in contiguous real storage, but......
  - Device volumes have grown in capacity over time
  - Generally applications do not get contiguous real storage greater than 4k in length
    - Larger amounts of contiguous virtual storage can be obtained, but this will not be in contiguous real
  - Channels always use real storage addressing
  - For a single I/O operation (CCW data transfer), data is sent contiguously, a single source data address was insufficient
  - The most efficient means to transfer data is to send it in the minimum number of IUs
Indirect Data Address Words (IDAWs)

- IDAWs provide a scatter-gather list type of function, BUT they carry strict limitations on data addresses
  - The data pointed to by the first IDAW may start anywhere within a 2k/4k area of storage, but
  - Subsequent IDAW data areas must start and end on a 2k or 4k boundary (except last)

When the data location virtual storage addresses meet address boundary requirements they can be translated to real storage addresses by the application or an Operating System function and IDAWs can be used for this single CCW data transfer command I/O operation

A single channel CCW operation is very efficient when using IDAWs
Chain Data (CD)

- Some application or subsystem single CCW I/O operations require that data that is fetched or stored is in scattered storage locations.
- These scatter/gather storage address locations and lengths do not meet the strict requirements of IDAWs, so IDAWs cannot be used.
- A Chain Data CCW operation can be used, but they are inefficient as pseudo CCW channel commands are sent for each chain data CCW.

<table>
<thead>
<tr>
<th>Host virtual storage addressing</th>
<th>Host real storage addressing</th>
</tr>
</thead>
<tbody>
<tr>
<td>2K / 4K pages</td>
<td>2K / 4K pages</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

Non-contiguous 2K / 4K pages (in virtual storage) with non-contiguous real addressing

Transfer data to or from storage device. A pseudo CCW command is sent for each CD CCW (1 IU per CCW)

A single channel CCW operation using Chain Data is very inefficient
Multiple Indirect Data Address Word (MIDAW)

- For the scenario where data comes from scattered storage locations and boundaries, use of MIDAWs has the channel efficiency of an IDAW operation and the flexibility of Chain Data storage location addressing.
- Use of MIDAWs is essentially exploiting a scatter/gather list of addresses and counts with relaxed boundary rules.

Host virtual storage addressing

4K pages

1
3
4
2
4
6

Non-contiguous 2K / 4K pages (in virtual storage) with non-contiguous real addressing

4K page

1

4K page

5

4K page

3

4K page

2

4

6

Channel

Transfer data to or from storage device.

A single channel CCW operation using MIDAW is very efficient—data is packaged into a minimal number of IUs.

FCIA
First Channel Industry Association
TRANSPORT MODE
Command Mode

Start Subchannel (SSCH)

ORB = Operation Request Block

ORB
CCW
CCW
CCW
IDAW / MIDAW
IDAW / MIDAW
IDAW / MIDAW
Data
Data
Data

CCWs can point directly to where the data resides (on a Write) or will reside (on a Read), or they can use an IDAW/MIDAW, which is a scatter/gather list of addresses and counts.

Each CCW (or Chain Data ‘pseudo CCW’) is sent on the link individually.
FICON – Better Performance?

Two aspects of FICON impacting performance:

• Amount of processing overhead in both the channel and storage controller

• Each command (and/or associated data block) sent individually

• Unique nature of the protocol did not exploit performance assists available to largest distributed systems FC ULP (i.e. SCSI attachment)

Newer protocols (like FC-NVMe) were designed with this in mind!

Although they can be streamed!

• Hardware assists built into Fibre Channel host bus adapter chips
• Firmware accelerations designed for Fibre Channel Protocol (FCP)
FICON – Transport Mode

The host communicates directly with the control unit

- The channel is acting as a conduit
- No individual commands or state tracking
- The entire channel program sequence is sent to the control unit in one descriptor
- Uses the Fibre Channel FCP link protocol. A FICON channel provides concurrent support for both the new and old protocols
- Transport Mode provides increased performance for both small block transfers and large block bandwidth
- Complex channel programs that are not easily converted to the new protocol still execute with the existing FICON Command Mode protocol
- Devices are accessible using both old and new protocols
Transport Mode

- ORB
- TCW
- TSB
- Data
- TIDAW
- Addressing & Control Information
- Device Command Word
- Device Command Word
- Device Command Word
- Device Command Word
- Checking Block
- Count

Replaces CCW

Same front-end as Command Mode

Replaces IDAWs/MIDAWs

ORB

TSB
Transport Status Block

TCCB
Transport Command & Control Block

Checking Block
Count

Word
0
1
2
3
4
5
6
7

00000000000000000000000000010000
LPM 00000000

Key

Channel-Program Address

CSS Priority
Reserved
Rsv. for Pgm.
Reserved
Reserved
Reserved
Reserved

Same front-end as Command Mode

Replaces CCW

ORB

Specifies Transport Mode

FCIA
Transport Command IU

A TC_IU is a direct mapping of an FCP_CMND IU with a variable length CDB.

### Addressing Info
- **FCP_LUN (8 bytes)**
- **Task Description & Control Fields (4 bytes)**
  - **FCP CDB**
  - Any additional FCP CDB
- **FCP_DL (data length)**
- **FCP_Bidirectional Read DL**

### Description of work to be performed
- **FC-SB-6 Header (8 bytes)**
  - **Transport Command Header (TCH) (4 bytes)**
  - **Transport Command Control Block (TCCB) (variable length)**
  - **Transport Count (data length)**
  - **Bidirectional Read DL**

### Notes
- **FC-NVMe places Connection Identifier in this field**
- **Identifies host and CU image IDs and target device**
- **Control info (including Additional CDB length)**
- **Contains the list of individual command words, control data, LRC for the entire TC_IU, and byte count of the total amount of data transferred**
- **FC-NVMe places Submission Queue Element (SQE) in this field**
**Link View of 4 Reads**

**Command Mode**
Prefix Cmd + 64 bytes → Read (4k) → Cmd Resp → Read (4k) → Data (4k) + CRC → Read (4k) → Data (4k) + CRC → Read (4k) → Data (4k) + CRC → Status Accept

**Transport Mode**
Prefix + 64 bytes of prefix data + Read (4k) + Read (4k) + Read (4k) + Read (4k) → Data (16k) + CRC → Status

<table>
<thead>
<tr>
<th></th>
<th>Channel to Control Unit</th>
<th>Control Unit to Channel</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Command Mode</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commands</td>
<td>5</td>
<td>N/A</td>
</tr>
<tr>
<td>Exchanges</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Sequences</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Frames</td>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td>CRC Generate / Check</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Channel to Control Unit</th>
<th>Control Unit to Channel</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Transport Mode</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commands</td>
<td>5</td>
<td>N/A</td>
</tr>
<tr>
<td>Exchanges</td>
<td>1</td>
<td>1 (same)</td>
</tr>
<tr>
<td>Sequences</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Frames</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>CRC Generate / Check</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
FICON 201 Themes

• FICON has evolved from its inception

• Integrity, Transactions, and Efficiency are the forces behind the evolution

• “IS FASTER” drives innovation!
Protocol Analysis 201
For High-Speed Fibre Channel
April 11, 2019
10:00 am PT/1:00 pm ET
Register at:
https://www.brighttalk.com/webcast/14967/349710
After this Webcast

• Please rate this event – we value your feedback
• We will post a Q&A blog at http://fibrechannel.org/ with answers to all the questions we received today
• Follow us on Twitter @FCIAnews for updates on future FCIA webcasts
• Visit our library of FCIA on-demand webcasts at http://fibrechannel.org/webcasts/ to learn about:
  – Fibre Channel Fundamentals
  – FC-NVMe
  – Long Distance Fibre Channel
  – Fibre Channel Speedmap
  – FCIP (Extension): Data Protection and Business Continuity
  – Fibre Channel Performance
  – Fibre Channel Cabling
  – 64GFC
Thank You