Fibre Channel Resilience for Enterprise Storage

Roger Hathorn
STSM – IBM Systems
Vice Chairman of T11.3 Fibre Channel Interconnect Schemes
Chairman FC-SB-6 Ad Hoc Working Group

FIBRE CHANNEL INDUSTRY ASSOCIATION

Fibre Channel Resilience for Enterprise Storage

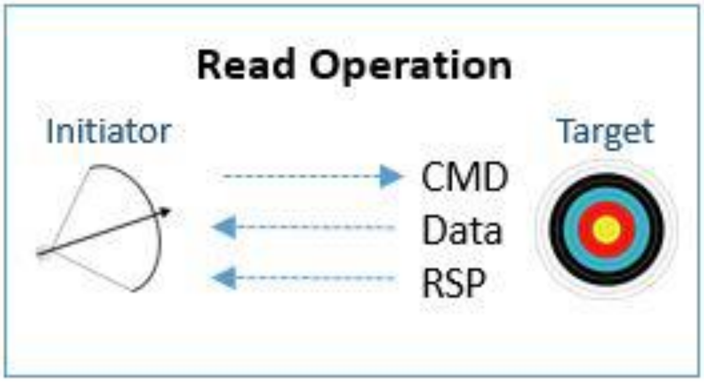
Fibre Channel is the choice for resilient storage networks due to its reliability and robust set of management capabilities. Mainframe computers that run the world's financial systems, high speed transactions processing, fraud protection, and reservation systems rely on Fibre Channel because it provides performance and reliability these systems demand for 24/7 availability.

Mainframe systems use FICON and High Performance FICON protocol to carry Single Byte commands (SBCCS) over Fibre Channel to communicate with storage. FICON and High performance FICON are defined in the T11 Fibre Channel standards as FC-SB-6 Command Mode and Transport mode, respectively. FC-SB-6 (FICON's 6th generation) is an FC-4 Upper Layer Protocol (ULP) that runs on top of the Fibre Channel FC-2 layer (defined by FC-FS-4) just like the FCP protocol (T10's FCP-4) is used to carry SCSI commands over Fibre Channel.

It's a common misconception that FCP is Fibre Channel and Fibre Channel is FCP and it is all about SCSI commands. FCP is merely an FC-4 layer command and data transport protocol that was initially defined to carry SCSI commands over Fibre Channel Layer 2. However, realizing that the FCP-4 transport provides the lowest overhead and streamlined operations, High Performance FICON uses the same FCP-4 transport for commands and data as SCSI FCP does.

In fact, the new NVMe over Fibre Channel Fabrics standard (FC-NVMe) also chose to make use of the FCP transport for commands and data because it provides for the lowest latency to carry NVMe operations over Fibre Channel. The steps of an FCP operation very much parallel the steps of an RDMA operation, which is touted in the technologies outside of Fibre Channel as a necessary component for low latency I/O. Figure 1 shows the sequences of an FCP Read operation.

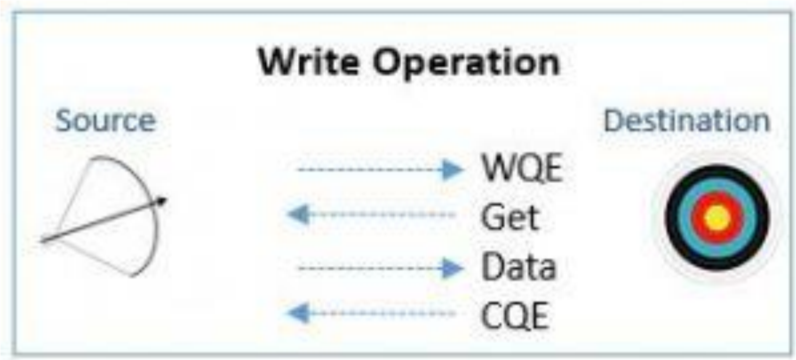
Figure 1 - FCP Read Operation



Similarly, Figure 2 shows an FCP write operation.

Figure 2 - FCP Write Operation

The host (initiator) transmits a command to the storage device (target) in a Command IU (CMD). The target controls the movement of the data to or from the host by sending the Data IU for a read or requesting the data for a write (XFER_RDY), and then the target completes the operation with a Response IU (RSP). This operation is performed over a Fibre Channel Exchange. The initiator of the exchange associates the host memory address for the data to the exchange prior to sending the operation to the target. When the target sends the data or requests the data, the data is automatically stored or retrieved from host memory. The difference from RDMA is that the host memory addresses are not pre-allocated or shared across the fabric.



With some implementations that support the use of first burst capability on writes, such as High Performance FICON (zHPF), the Transfer Ready (or Get) operation is eliminated which can eliminate an additional 5us/km (the speed of light) in each direction which is 10us/km for a round trip. Implementations use the new standard FC-SB-6 First Transfer Buffer Credits to allow an entire write sequence to be transmitted as a first burst of data, eliminating all roundtrip delays.

Fibre Channel is extendable over IP networks using FC-IP capable switches providing for any distance required to reach around the world. Use of such extension is not uncommon in disaster recovery solutions and this first burst capability is essential for performance at distances in the 100's to 1000's of kilometers.

Imagine standing at the ATM or in line at the checkout from your favorite store waiting for approval of your transaction. Every transaction, while started over the internet or telephone lines, likely winds up on a backend system powered by the Mainframe and Fibre Channel to perform the fraud protection and verification required to complete the transaction. If the transaction cannot be authorized, the store or bank must either reject transaction or take on the liability of accepting it. This is why enterprise applications demand high performance, low latency storage networks that can provide for guaranteed response times on such transactions. Fibre Channel lossless credit-based protocol was built from the ground up to provide such reliability.

Not only does Fibre Channel provide the reliability and performance necessary for these demanding applications, the community also realizes that the infrastructure needs to be manageable. This is why industry experts are bringing new technology to the standards to help vendors provide solutions to help manage storage networks. One such technology is the Read Diagnostic Parameters (RDP) Extended Link Service, now part of the FC-LS-4 standard. This link service provides the ability for any entity in the SAN including servers, storage, or switches to provide visibility to the operational port and link characteristics of any other port in the SAN.

Examples of the information available include;

a) optical transceiver power information including transmit and receive power levels, module voltage, current, and temperature so that link issues can be diagnosed without sending someone out to the floor with power meters, potentially causing further disruption or introducing further issues on the links;
b) Buffer to Buffer credit levels provided by the ports to help diagnose performance bottlenecks and appropriate values for the distances involved.
c) Current operating link speed as well as the capable operating speeds of the port (remember that Fibre Channel ports can auto-negotiate between at least 3 link-speed generations), allowing identification of links that are not operating at optimal speed;
d) a standard set of link error statistics to identify excessive bit errors, uncorrectable errors, CRC errors and link failures, etc.; and
e) World Wide node and port names of the port and the attached port to identify both ends of any link.

The list goes on and keeps growing as this technology is quickly being adopted by venders in the community and further uses are discovered. This information, when provided for each end of a link or each port in the path from server to device through the SAN, gives visibility to the health of the SAN which can be monitored for current failing links and transceivers as well as provide the information that, with cognitive analysis over time, can predict when failures may occur and allow for proactive maintenance before issues occur.

Conclusion

Fibre Channel has provided a regular cadence of upgrades in speeds and feeds since the mid 1990s as well as continuous functional protocol enhancements and manageability improvements. With over 30 participating companies, the INCITS T11 Fibre Channel standards committee has a clear roadmap to the future with the next generation of 64GFC and 256GFC (parallel operation) just around the corner (1H 2018 standard ratification expected) and 128GFC (and 512GFC) projects already being proposed. Systems and storage vendors continue to use Fibre Channel to meet the demands of enterprise storage networks and continue to drive valuable enhancements into T11 standards for continued interoperability. Fibre Channel has and always will provide a paramount level of performance and reliability the future of business demands.