The Emulex LightPulse™ Fibre Channel host bus adapter architecture is designed and optimized for high performance enterprise systems. With a sophisticated ASIC controller, a powerful RISC processor, and specialized storage area network (SAN) hardware, Emulex host bus adapters provide the performance and flexibility needed in complex SANs.

This Technology Brief focuses on the unique capabilities of Emulex’s large data buffer, which is specialized SAN hardware that provides support for 64 buffer credits, providing improved performance and greater flexibility. The benefits of a large data buffer are particularly evident in long distance applications, when operating at 2Gb/s data rates, or in systems with a heavily loaded PCI bus.

**Background**

Applications such as remote backup, restore, campus-wide SANs, movement of high bandwidth data (images, video, large files, etc), and FICON are just a few of the applications that create the need to move data extended distances over Fibre Channel links. Long distance applications are becoming more common as SAN deployment grows and the use of repeaters becomes more prevalent. While improving the utility of the Fibre Channel network, long distance connections create challenges for high performance data movement that can be addressed with the right host bus adapter (HBA). As will be shown in this white paper, a key enabler to the performance of a long distance connection is the amount of buffer credit that a host bus adapter supports.

The move to higher data rates also is driving the need for increased data buffering on the Fibre Channel host bus adapter. This white paper will demonstrate how higher data rates, including the emerging 2Gb/s data rate, place a greater burden on the HBA data buffering to maintain performance even at connections of moderate distances.

Systems with heavily loaded PCI busses also benefit from Fibre Channel HBAs with large data buffers. If congestion on the PCI bus occurs, the HBA data buffer can serve as a data reservoir, allowing the Fibre Channel fabric to continue delivering data to the HBA and maintain link performance. Without a large data buffer, a heavily loaded PCI bus could force the Fibre Channel link to stop delivering data until the congestion clears. This ability to maintain SAN fabric data flow, even in the case of a heavily loaded PCI bus, enables PCI scalability, allowing increased I/O performance through the addition of Fibre Channel host bus adapters to the system.
Buffer credit overview

Fibre Channel data is organized in frames that can be concatenated into sequences to create large block transfers. The frame size is dependent upon the host bus adapter and the target, and can be in the range of 512 bytes to 2KB. Multiple sequences can be combined into a single exchange, permitting up to 128MB of data to be transferred with a single I/O command. The maximum amount of frame data that can be in flight at any given time is governed by buffer credits, which if not sufficient for the link distance and speed will severely limit performance.

Buffer credit is a mechanism defined by the Fibre Channel standard that establishes the maximum amount of data that can be sent at any one time. This scheme is used to allow Fibre Channel to be self-throttling, thereby allowing it to establish a reliable connection without the need to accommodate dropped frames due to congestion. The use of buffer credit ensures guaranteed delivery using the hardware mechanism provided by the Fibre Channel protocol, rather than software error correction, delivering significantly higher performance and lower server CPU overheads compared to other networking technologies.

Buffer credit limits, between each device and the fabric, are communicated at the time of fabric login. One buffer credit allows a device to send one frame of data (typically 1KB or 2KB) before a “receiver-ready” acknowledgement is received. In order to send more data than a single frame, more buffer memory must be available. The size of the data buffer dictates the amount of buffer credit that can be extended by the device, and in turn limits the amount of data that can be in flight at any given time.

Calculating buffer credit requirements

The amount of buffer credit required to sustain maximum bandwidth over extended distances can be easily calculated. For example, light travelling down an optic fiber has a latency of 5nsec per meter, so the round trip propagation time of data traveling through a 10km fiber optic cable and a “receiver-ready” acknowledgement is 100µsec. When transmitted through this cable, the transfer of a 2KB frame on a 1Gb/s link has a transmit time, to get fully on the link, of 20µsec. Therefore, for single buffer credit transaction, only 20% of the link bandwidth is available for a connection with a single buffer credit. Multiple frames must be in flight in order to more fully utilize the link bandwidth, and seven 2KB credits being the minimum required to fill a 10km 1Gb/s pipe. With similarly, a 100km cable requires a minimum of 61 buffer credits to maintain maximum link bandwidth.

When operating at 2Gb/s data rate, twice as many buffer credits are needed to achieve the same link utilization for a given distance, since the data transmit time is cut in half. The proposed 4Gb/s and 10Gb/s data rates will depend even more heavily on large data buffers to sustain high bandwidth utilization.

Emulex buffered data architecture

Emulex host bus adapters feature a buffered architecture that accommodates up to 64 buffer credits of 2KB each. Emulex HBAs have 256KB of dual ported buffer memory on board that can receive data at full link speed, while simultaneously moving data out to the PCI bus to the host server. This architecture enables high sustained throughput in any conceivable SAN application, and is unique in its ability to sustain full bandwidth utilization of a 1Gb/s or 2Gb/s link over long distance. In addition, Emulex HBAs are uniquely capable of supporting the 20-100km distance required by the FICON specification. Therefore, Emulex’s HBAs are ideally suited for the industry transition to higher data rates, including 2Gb/s, 4Gb/s, and even 10Gb/s.

Typical Fibre Channel architectures

Typical competitor architectures use a small amount of integrated memory in the controller ASIC. This memory, typically enough for four frames of 1K data or two frames at 2K, is all that is available to extend buffer credit to the SAN fabric. As can be calculated from the examples above, this amount of buffer credit is capable of utilizing less than 40 percent of the available link bandwidth at 10km. Although sufficient perhaps for small SANs, this architecture is limited to environments requiring less than 500m of cabling, severely limiting the flexibility of these solutions. This architecture will become even more limiting as data rates increase from 2Gb/s to 4Gb/s, or even 10Gb/s. Additionally, it does not address PCI congestion issues related to adding multiple HBAs per system.
Performance benefits of a large data buffer

As described, HBAs featuring large data buffers can provide significant performance advantages. The performance chart to the left shows that Emulex host adapters maintain high throughput over extended distances, while the leading competitor’s performance is dramatically impacted due to its limited number of frame buffers. As this test shows, a single Emulex HBA provides more than twice the performance of the leading competitor in a 10km test. In addition to higher performance, this architecture is inherently more scalable and provides greater flexibility in configuring a SAN.

Summary

Host bus adapters featuring a large data buffer provide compelling performance advantages in Fibre Channel SANs. In addition to enabling higher bandwidth performance over extended distances, a large data buffer provides the ability for performance to scale linearly with increases in link data rate, and by adding more host connections, per system, to the SAN. HBAs featuring a large data buffer also provide unmatched flexibility in implementing SANs. Emulex’s Fibre Channel HBAs are designed with the largest data buffer offering in the industry, providing dramatic performance benefits in extended distance tests. Competing host adapters, by contrast, utilize a small integrated frame buffer that significantly limits performance in today’s SANs, and does not allow for scaling to higher data rates.

This report is the property of Emulex Corporation and may not be duplicated without permission from the Company. | April, 2001