EMULEX®

White Paper

Advancing Storage Systems into the Future

AT A GLANCE

This paper reviews the advancement of the primary Fibre Channel embedded storage architectures used in the past, present and future to offer the reader a way to understand the benefits of each embedded storage architecture.

PRODUCTS

InSpeed SOCs and Root Switches, FibreSpy SOCs and Root Switches

APPLICATIONS

SBODs, Root Switch, Fully-Switched, Tiered Storage, and SuperScalar Architectures

This paper reviews the advancement of the primary Fibre Channel embedded storage architectures used in storage systems in the past, present and future (Figure 1) to offer the reader a way to understand the benefits each embedded storage system architecture offers.

Within a storage system, embedded storage architectures provide the connection between a storage controller to one or more disk drive enclosures, and to the disk drives. A storage controller is a printed circuit card that typically provides the connectivity to the storage system from the outside world and also delivers the RAID capability to manage the physical and logical disk drives within a storage system. When this paper refers to a "storage controller" it is referring to a disk RAID based controller, a tape library controller, and a NAS head.

Externally attached enterprise and mid-range storage systems have significantly evolved over the last decade from systems based on shared loop Fibre Channel technology. These shared-loop based systems, while mostly providing adequate levels of performance, were limited in their ability to scale and offered only the most basic levels of reliability, availability and serviceability (RAS).

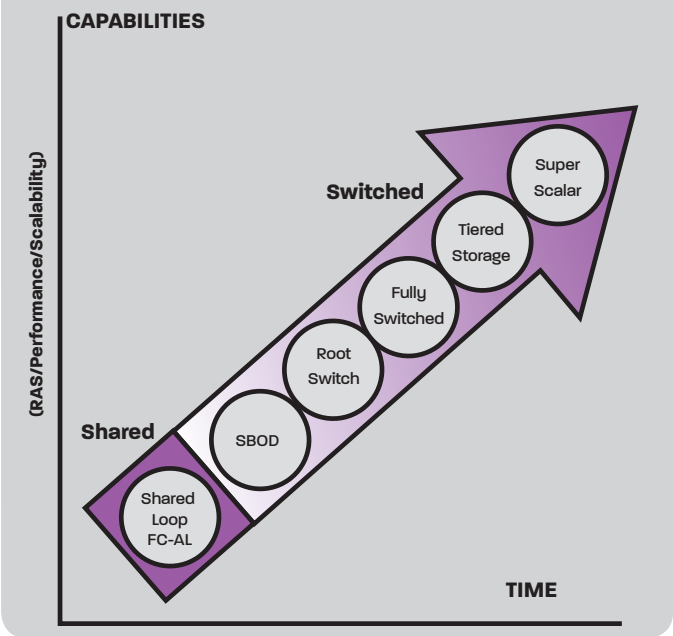
"There is no intelligent argument that says having a non-switched back-end is better than switched, so it's merely a matter of money. . .(Emulex) has pretty much eliminated that concern." Duplessie said everyone who doesn't switch the back-end of their storage array will have to justify it to customers. "And there won't be any good answers."

Steve Duplessie,
President and Senior Analyst
Enterprise Strategy Group

However, Fibre Channel based storage systems have greatly advanced over the last four years due in great part to Emulex's efforts to implement switching into the back-end of storage systems. Emulex has continually worked to enhance the fundamentals of Fibre Channel to leverage existing designs and investments forward into the future to meet new market requirements as the storage system industry evolves.

These new technologies offer significantly higher levels of RAS and now maintain performance while scaling, while also offering ways to lower total storage system costs — all accomplished while preserving previous technology and development investments made by storage system providers. The ability to leverage past investments by offering backward compatibility with field-proven, existing designs has been a key requirement for today's storage system providers to quickly deliver trusted and robust storage systems to market, and to leverage the systems already installed in IT departments.

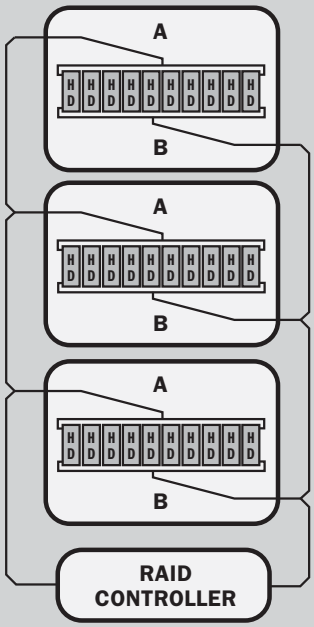
Figure 1 - Fibre Channel Embedded Storage Infrastructure Evolution



Just a Bunch Of Disks (JBOD) Architecture

Just like its name, a JBOD (Figure 2) is just a bunch of disk drives interconnected by a Port Bypass Circuit (PBC). Fairly much extinct at the time of this paper, the PBC is a basic electronic component used within JBOD systems and Fibre Channel hubs. PBC's operate on a shared-loop topology and allow disk drives to be powered down and removed without interrupting traffic to other disk drives or impacting data integrity on the loop. PBC's also by-pass and electronically remove improperly behaving disk drives from the shared loop. However, much like a string of holiday lights, the PBC is not capable of detecting or isolating a specific misbehaving disk drive — all it can do is remove a bad disk from the loop. Therefore, PBC's do not generally provide the level of robustness that today's embedded storage systems require. Furthermore, shared-loop environments can easily get saturated and impact performance when as few as 30 disk drives are installed within a storage system. This performance impact becomes increasingly worse as additional drives are added to the storage system since more and more devices are fighting over the shared loop line to the storage controller.

Figure 2 – JBOD Architecture

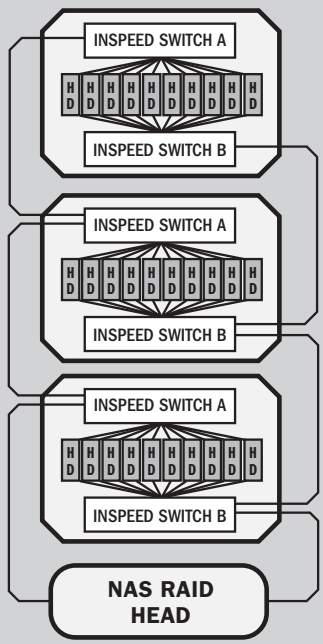


Switched Bunch of Disks (SBOD) Architecture

The SBOD® is one of the most commonly switched architectures deployed by storage system providers today. It was developed by Emulex to be fully compatible with Fibre Channel designs implemented for PBCs, yet eliminate their inherent drawbacks and issues. As shown in Figure 3, an SBOD is created by replacing the PBC with an Switch-on-a-Chip (SOC) that provides multiple dynamic direct connections between the storage controller and the disk drives. Emulex's InSpeed® Embedded Storage Switch is the most widely shipped and deployed SOC on the market today. Because of its direct point-to-point connections, InSpeed SOC-based SBODs deliver over 2-times higher performance than shared loop PBC-based JBODs when using the SPC 1.0 benchmark. SBODs provide the following benefits for storage systems:

- Intelligent Device Monitoring – SBODs introduce intelligent per device monitoring, providing extensive and critically needed improvements over PBCs. An InSpeed-based SBOD increases RAS during every phase of a storage systems operation. Even the most advanced PBCs fail to help identify the source of errors or provide the level of diagnostics offered by an Embedded Storage Switch, like InSpeed.
- Trend Monitoring – SBODs provide access to several metrics and diagnostics tools that can be used to help monitor trends over time or during an active troubleshooting session. SOCs in particular are in a unique position to constantly monitor disk drive health as disk drives are the primary source of storage system failures.
- Performance Improvements – The performance of an SBOD is greatly improved over a JBOD. SBOD performance increases are dependent on both the topology and the number of disk drives in an enclosure. The performance boost introduced by an SOC can be over 2-times. Furthermore, technical analysis has shown that as the storage industry moves to 4Gb/s, it can be expected that a 2Gb/s SBOD will exceed the performance of a 4Gb/s JBOD.

Figure 3 –SBOD Architecture



Root Switch Architecture

Root Switches are also widely deployed today in disk-based storage systems, and especially in tape libraries. Root Switch Architectures (Figure 4) provide a truly drop-in, fastest time-to-market, compatible solution for placing switched connectivity between the storage controller and the disk drive enclosures within the storage system. Like the SBOD implementation, Root Switch implementations bring increased levels of fault isolation and dynamic recovery to the storage system, enable new levels of diagnostics and performance while scaling for adding disk drives. Root Switches provide 3-times greater performance than shared-loop PBC implementations using the SPC 1.0 benchmark, while providing RAS down to the disk drive enclosure-level.

Fully-Switched Architecture

In a Fully-Switched Architecture (Figure 5), both Root Switch and SBOD Architectures are combined to deliver the highest performance, and ultra-reliable solution – from the storage controller down to each individual disk drive. By having a maximum of two connections (hops) between the storage controller and the disk drives, overall latency greatly decreases to deliver up to 4-times performance benefit over shared-loop PBC-based architectures using the SPC 1.0 benchmark. Fully-Switched Architectures deliver the maximum reliability, availability, and serviceability (RAS) characteristics a storage system can offer because service actions can be performed at the disk drive and enclosure-levels.

Figure 4 – Root Switch Architecture

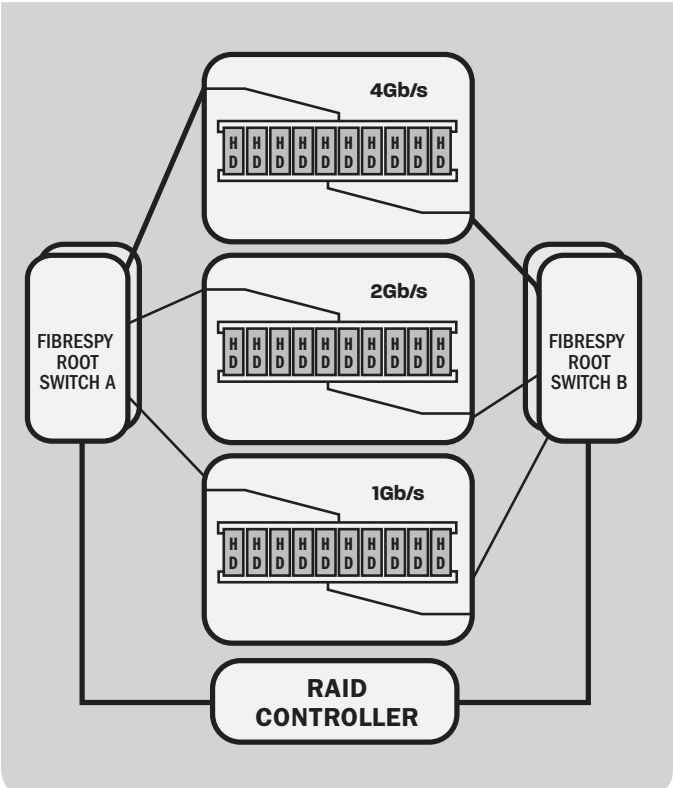
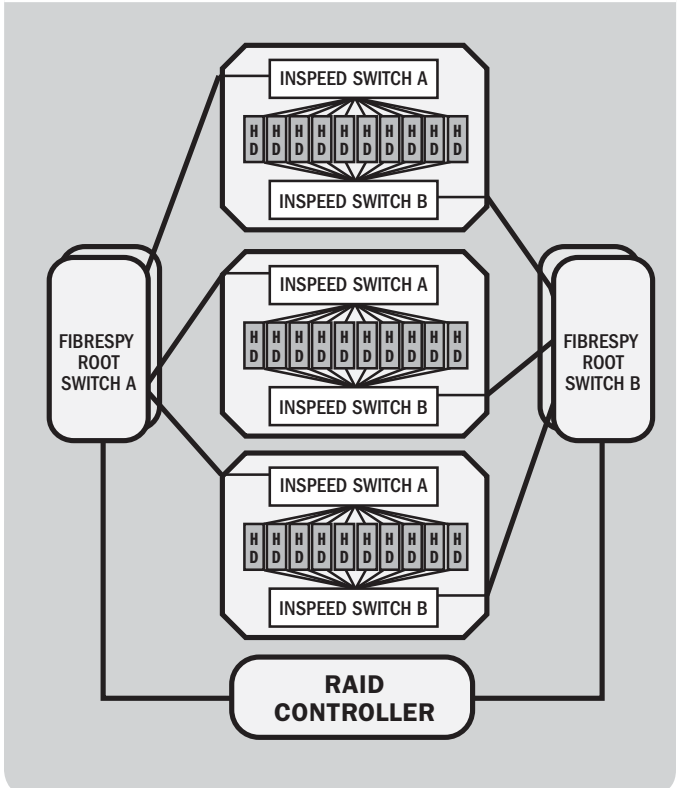


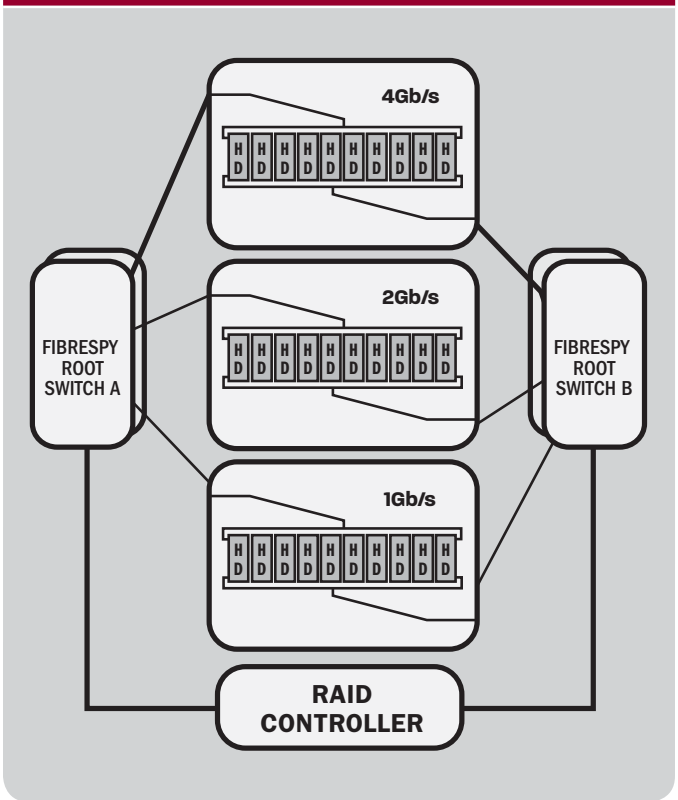
Figure 5 – Fully-Switched Architecture



Tiered Storage Architecture

Tiered Storage Architectures (Figure 6) represent a new breed of capabilities for storage systems by providing several types of data storage for a variety of applications within a single storage system to enable Information Lifecycle Management (ILM) capabilities. This new approach to storage, which is enabled by Emulex's FibreSpy® family of products, allows IT managers to allocate disk drives with differing capabilities to applications that may also have differing RAS or performance needs. Tiered Storage can be implemented within an SBOD, Root Switch or Fully-Switched Architecture. In any case, the SBOD or Root Switch typically must offer the ability to connect different speeds of disk drives or disk drive enclosures within a single storage system backend. Today these speeds are 4, 2 or 1Gb/s.

Figure 6 – Tiered Storage in a Root Switch Architecture



A new emerging technology to watch is FC-SATA. FC-SATA provides the ability to natively connect Serial ATA (SATA) disk drives into Fibre Channel embedded storage architectures, thus providing investment protection to storage system providers who want to lower their storage system costs and quickly meet Tiered Storage market needs using SATA disk drives. Also developed by Emulex, FC-SATA is now being offered to the ANSI T11 standards body to enhance the Fibre Channel industry standard so all component suppliers can offer complete FC-SATA solutions to advance Fibre Channel into the future.

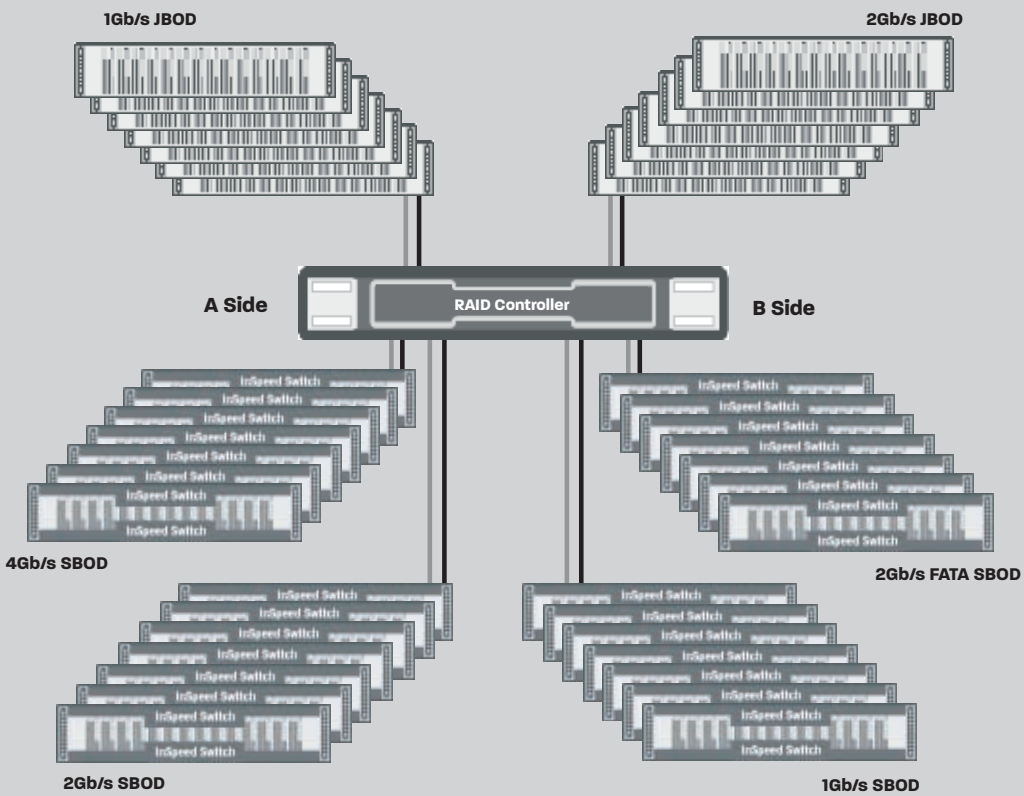
The primary advantages of implementing Tiered Storage within a storage system are:

1. Allows IT managers to use a single common software tool to allocate storage. This removes the requirement to have to learn the software utility tools of several different products, vendors and operating systems, making IT managers significantly more productive.
2. The versatility of allocating the right storage to the right application allows for a reduction in storage costs for applications that can get by with less costly disk drives. These changes can be made through Tiered Storage without compromising application performance or availability.
3. Tiered Storage systems make possible a true pay-as-you-grow storage provisioning strategy since they can — at a granular level — be used to add disk drives behind the most expensive part of the storage system, the storage controller. This capability frees IT managers to utilize their storage pools most effectively and minimize wasted disk space.

SuperScalar Architecture

All of the previous architectures discussed so far are capable of connecting a total maximum of 126 disk drives and storage controllers into an embedded storage network. This 126-device limitation is the maximum number of device addresses offered by the Fibre Channel Arbitrated Loop (FC-AL) protocol. SuperScalar Architectures (Figure 7) are an emerging embedded storage system architecture developed to break through this 126-device limitation in the back-end of Fibre Channel storage systems. Enabled by Emulex's FibreSpy family of products, SuperScalar Architectures give storage system providers the ability to now connect thousands of disk drives behind the storage controller(s). This new capability helps lower the total cost of storage systems because the storage controller, the most costly part of the storage system, can now be amortized across many more terabytes of storage. Like Tiered Storage Architectures, SuperScalar Architectures can be implemented using any combination of SBOD, Root Switch or Fully-Switched Architectures. And, SuperScalar implementations can be implemented to be Tiered Storage capable as well, as shown in Figure 7.

Figure 7 – SuperScalar Tiered Storage Architecture



Conclusion

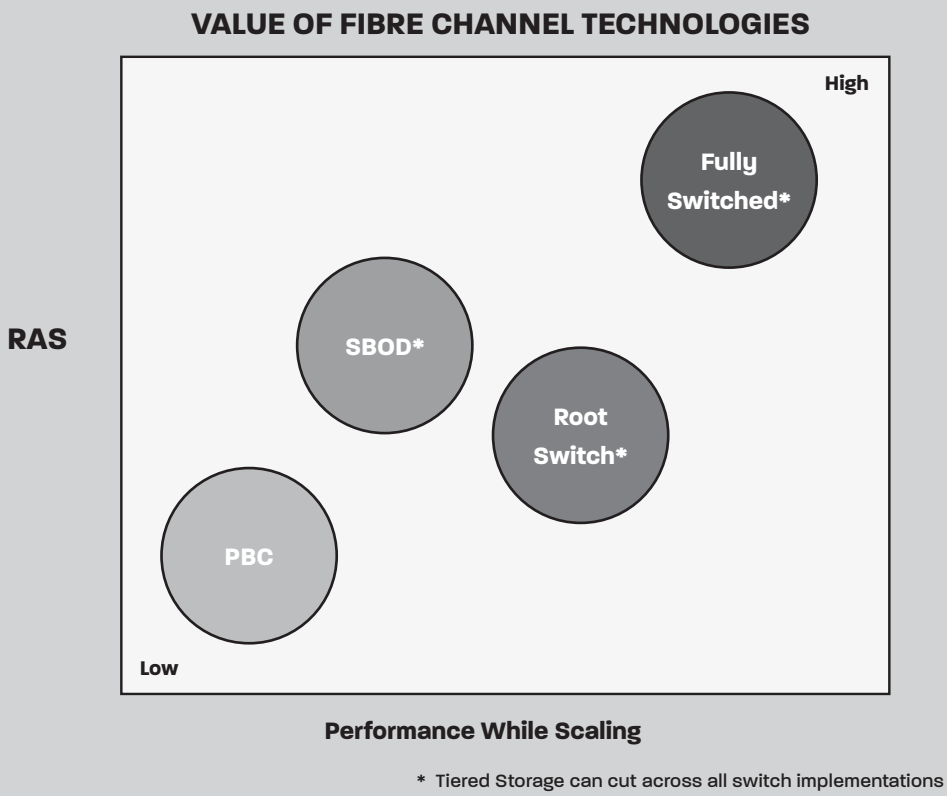
There are several different approaches that can be implemented in the back-end of storage systems depending on the storage system providers' target market, cost objectives, level of performance while scaling, RAS, and advanced capabilities required, as shown in Figure 8. Within a Fibre Channel infrastructure, each architecture leverages existing Fibre Channel designs forward into the future while preserving existing investments. What becomes critical, however, is as the number of disk drives increase and volume of data passing through the embedded network increases, it becomes much more essential to ensure that performance and RAS are maximized. Any loss of connectivity means loss of access to an application and/or its data, and translates into business downtime.

By configuring storage systems with SBOD, Root Switched, Fully-Switched, Tiered Storage, or SuperScalar Architectures, storage system providers can be certain to meet their desired level of performance while scaling, RAS and cost-effectiveness objectives.

Emulex's InSpeed products are field-proven with over 15-million installed ports implementing SBOD, Root Switched, and Fully-Switched storage configurations.
And, Emulex's FibreSpy product offerings deliver new Tiered Storage and SuperScalar capabilities for the very first time.

Additional information about Emulex's InSpeed and FibreSpy products can be found at www.emulex.com.

Figure 8 - Value Comparison of Embedded Storage System Architectures



STORAGE NETWORKING DEFINITIONS	
Port Bypass Circuits (PBC)	A circuit that automatically opens and closes a shared Fibre Channel arbitrated loop (FC-AL) so that nodes can be added to or removed from the loop with minimal disruption of operations. PBC's are typically limited in their ability to diagnose problems to a specific device level, and have decreasing performance characteristics as disk drives are added to the shared loop. PBC's are typically found in shared loop Fibre Channel hubs and JBOD disk enclosures where performance and RAS are not critical considerations.
SBOD (Switched Bunch of Disks)	Switched JBOD. An SBOD is a JBOD that uses an Embedded Storage Switch to improve performance while scaling and RAS characteristics. An SBOD is an Emulex trademark term describing a mechanism for accessing hard drives in an enclosure through dynamically switching initiators and targets such that multiple simultaneous conversations can occur. SBODs enable performance acceleration while scaling to the full AL_PA address space, and dramatically improve reliability, availability and serviceability (RAS) in storage systems. SBODs also support the ability to be serially connected either through a single cascade, or through trunking, where both bandwidth multiplication and failover are introduced.
Root Switch	Root Switch implementations place dynamic switched connectivity between the storage controller (or NAS head) and the drive shelves within a storage array, and can also be similarly deployed into tape libraries. Root Switches are defined as being positioned in the back-end of a storage system. Root Switch applications bring increased levels of fault isolation and dynamic recovery to the storage system, enable higher levels of diagnostics, and the scaling of drive counts to the full range of an AL_PA space. Root Switch configurations bring the highest level of performance acceleration to a storage system, and deliver the simplest mechanisms for removing and replacing drive shelves, resulting in dramatically improved serviceability.
Fully-Switched Architecture	A Fully-Switched Architecture combines a Root Switch with SBODs to provide switching at all (the disk drive and disk drive enclosure) levels. Each drive can be accessed deterministically through two hops from an initiator, providing maximum performance, reliability, availability and serviceability. A Fully-Switched Architecture also supports trunking to attached SBODs ensuring both bandwidth multiplication and failover for even higher levels of performance and availability.
Tiered Storage	Provides the ability to mix and match different classes of disk or tape storage behind a storage systems storage controller. These different classes of disk can be high performance 4Gb/s Fibre Channel disk drives, legacy 2Gb/s Fibre Channel disk drives, and/or emerging Fibre Channel nearline disk drives. Tiered Storage provides the ability to place data on the most cost effective disk storage possible for a particular application. The proposed FC-SATA standard will allow SATA disk drives to be natively and cost effectively attached into the same storage system as Fibre Channel disk drives.
SuperScalar	SuperScalar addresses the needs of creating highly scalable storage architectures by allowing thousands of disk drives to be connected — breaking the 126 disk drive limitation found in storage systems prior to now. This limitation is due to the number of device addresses provided by the FC-AL protocol typically used in storage systems. Emulex's FibreSpy technology uses the FC-FLA protocol to break the FC-AL AL_PA space barrier and enable storage system OEMs to place thousands of disk drives behind a single storage controller. SuperScalar Architectures help lower the total cost of storage systems because the storage controller can now be amortized across a larger number of targets while using different classes of drive technology.

Corporate HQ: 3333 Susan Street, Costa Mesa, CA 92626 (714) 662-5600
Wokingham, U.K. (44) 118-977-2929 | **Paris, France** (33) 41-91 19-90
Beijing, China (86 10) 6849-9547

07-315 11/06