

SANBlaze: VirtualLUN

FCoE Technology primer

Inside contents

Overview	1
Why FCoE	1
Key FCoE features	2
FCoE Storage Hierarchy	3
FCoE Test Challenges	4
Summary	5



Overview

FCoE (Fibre Channel over Ethernet) is a new technology that that will allow Ethernet networks to carry Fibre Channel storage I/O data frames. Fibre channel (FC) is both a protocol and a specialized high-speed optical network, developed over a 20-year period, most often used for storage networking and is the backbone of most modern data centers. Because data center “up-time” is crucial to so many corporate operations, corporations often deploy a separate FC network because it is fast, reliable (loss-less) , and boasts very low latency. If FCoE lives up to its promise, data centers and blade server backplanes will soon have the option to consolidate future infrastructure investments and operate FC protocol on 10-Gb Ethernet equipment.

Why FCoE ?

Before diving into the FCoE test challenges, it might be helpful to understand why the industry would consider another Ethernet based storage protocol when we already have iSCSI, FCIP, NFS and CIFS. By their very nature, all storage protocols require guaranteed packet delivery. When a document is saved, there can be no ambiguity about whether the data arrives at the disk. Yet Ethernet is by definition a best-effort transport. It attempts to deliver all packets, but Ethernet may drop packets whenever the network becomes very busy.

To tolerate Ethernet’s intrinsic lossy nature, many applications employ TCP. TCP is effectively a “tally system” an application can use to guarantee data delivery between sender and receiver. With TCP, every packet is numbered and tracked as it travels through the IP network. TCP protocol detects dropped or lost Ethernet packets, and retransmits whenever necessary. TCP software runs on both the transmitter and receiver, and can sometimes consume a large percentage of the host computers CPU bandwidth.

FCoE abandons IP protocol and IP addresses (e.g. 192.168.1.xx), instead employing a new 802.3 Ethertype. The approach eliminates TCP altogether, replacing it with a hardware flow control scheme which guarantees packet delivery and can match performance metrics found on native Fibre channel networks. Moreover, FCoE uses the same protocols as native FC, which allows inexpensive gateways to shuttle traffic between new FCoE equipment and legacy FC equipment. Consolidation means fewer adapters, transceivers, and cables (cost), less labor to manage a single infrastructure (cost), and less energy and cooling (cost). Consolidation can be especially useful in ATCA blade server environments, which often feature Ethernet XAUI signaling as the only means to conduct high-speed blade-to-blade I/O.

Key features for high-performance FCoE

To match the performance metrics established by native Fibre channel technology, users should look for Ethernet controllers with two key characteristics. First, look for controllers with PAUSE frame management implemented in hardware. As mentioned, FCoE is the only major Ethernet storage protocol that will not utilize TCP. To guarantee packet delivery, FCoE utilizes a flow control scheme defined in IEEE 802.3X called PAUSE commands. PAUSE will exist at the data link layer, and allows sender and receiver to temporarily pause traffic. Optionally, some Ethernet controllers will offer a priority pause scheme, which allows suspension of FCoE frames, but permits non-FCoE traffic to pass through the controller. The second key feature is jumbo frame support, already common on most commercial Ethernet controllers. Fibre channel allows data payloads that are 2KB in size. Jumbo frame support will allow the FCoE network to carry the maximum size, un-altered native FC frames.

Understanding FCoE Storage Hierarchy

The equipment or tools selected to validate FCoE must have the ability to understand block storage hierarchy. The foundation of FCoE storage is still the disk drive, constructed of platters that hold magnetic charges representing our data. Conceptually, disk firmware divides each platter into concentric circular tracks- much like an old LP record, (for those old enough to remember). The disk divides each track into sectors, and each sector contains some number of blocks. A block is the smallest addressable chunk of data on a disk; generally 512Bytes. (It is the “block” in what we call “block I/O”). The disk linearly numbers each block from 0 to its max capacity. The index is called a logical block address (LBA).

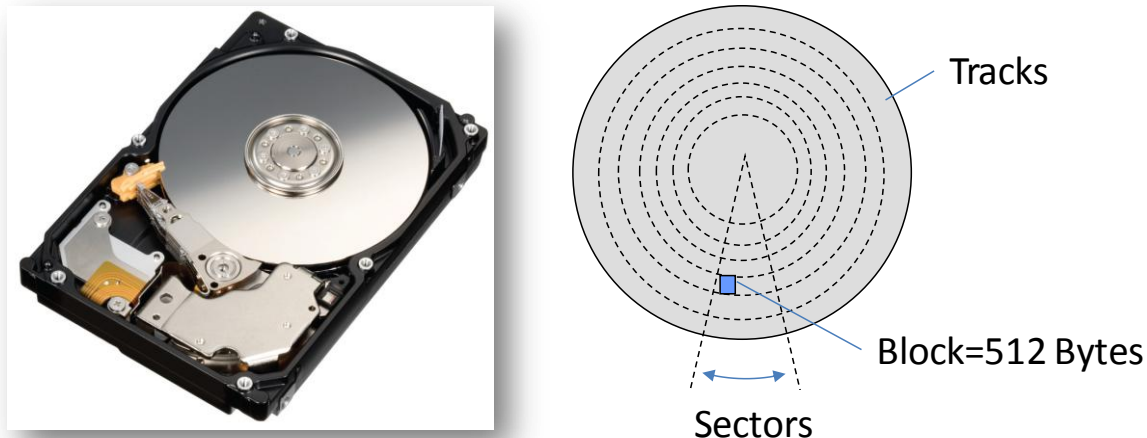


Fig 2. Disk platter storage terminology.

If you walk into a data center, you will see 100's of disks installed into storage devices that provide virtualization and RAID protection services and ultimately makes the storage visible to a host. Several vendors announced FCoE storage products in 2008, and users can expect product announcements to accelerate in 2009. Let us briefly examine what is happening inside an FCoE storage array:

- 1) Real physical disks are installed into the storage product.
- 2) Software in the storage product may apply RAID services to these disks
- 3) Software in the storage product is used to create a virtual volume, and virtual disks
- 4) Software assigns virtual disks to hosts using access control lists (ACL)
- 5) Hosts access virtual disks using FC protocol
- 6) FC commands are encapsulated in Ethernet packets (FCoE) for transmission

The final piece of FCoE storage hierarchy is the FCoE protocol. A Russian doll is a great analogy to convey what is happening. The outer “shell” is the Ethernet 802.3 packet. Ethernet Switches use the address information to deliver the packet. The Ethernet frame carries a new Frame type designated for FCoE traffic. The FCoE frame carries Fibre Channel frames. Within the FC frame, is the FC protocol (FCP), which carries a SCSI command. SCSI commands are the “Block I/O language” that hosts use to communicate with storage products. The SCSI command set includes SCSI read, SCSI write, and dozens for disk management that all exist to facilitate block I/O storage activities.

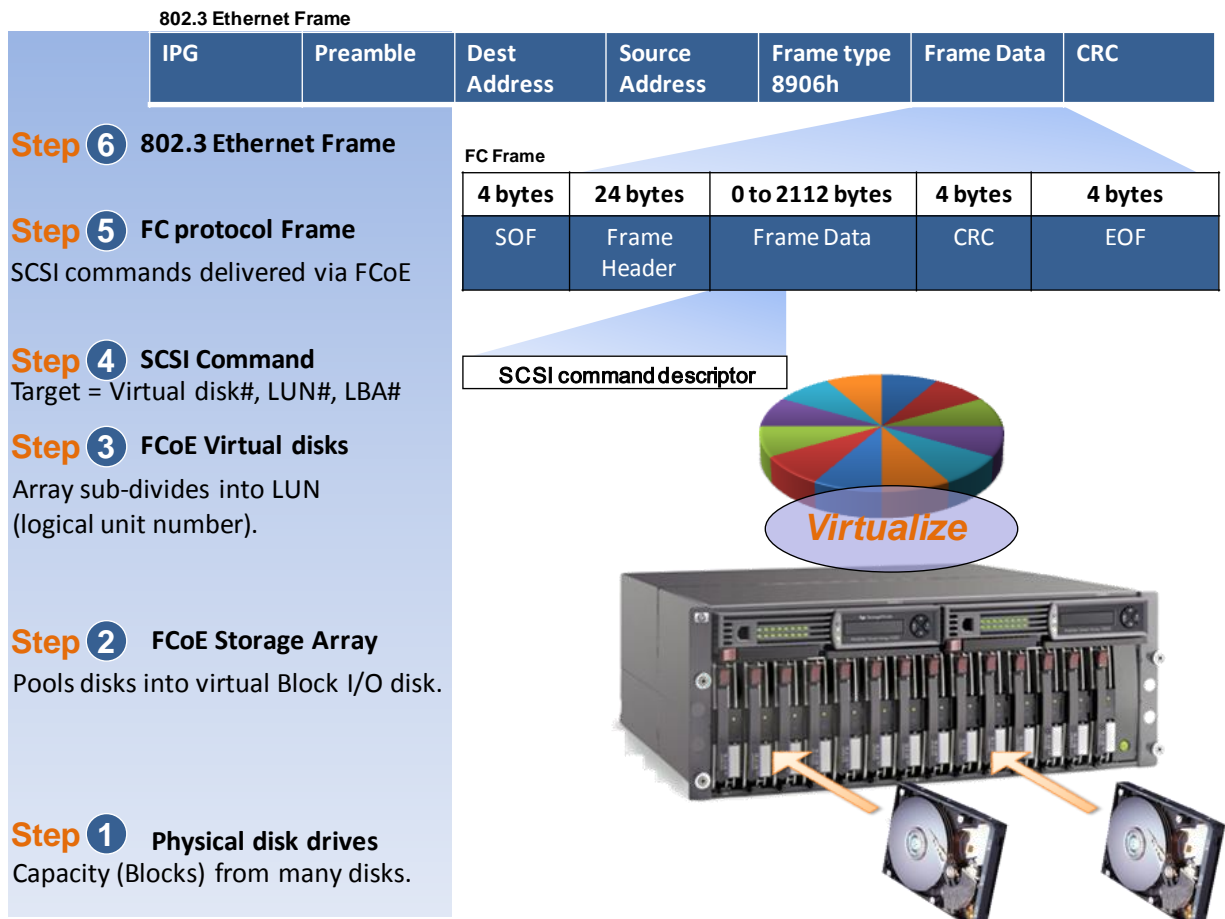


Fig 3. FCoE Protocol Hierarchy.

FCoE test challenges

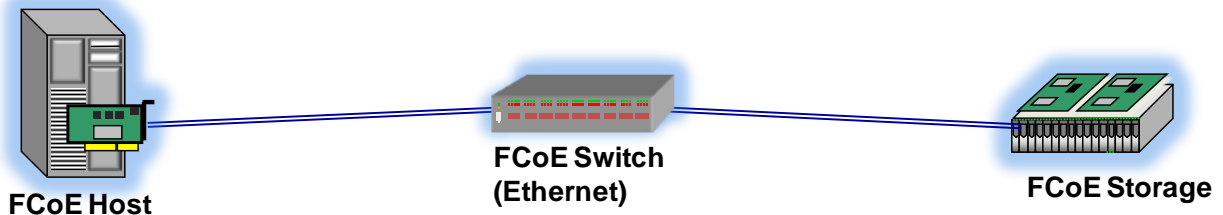


Fig 4. FCoE ecosystem: Initiator, switch, and targets.

Whether FCoE deploys in the data center or the backplane storage choice in a chassis, there will be servers, switches and storage products. Each product category presents a slightly different test challenge, which requires validation of functionality, performance and error path responses. As the minimum criteria for consideration, FCoE test equipment must support this varied list of functional FCoE protocol exercise capability:

FCoE Hosts (Initiators)	FCoE Switch / Infrastructure	FCoE Storage/Array (Targets)
<ul style="list-style-type: none"> • Validate different CPU/OS • OS boot (various) • Disk/array discovery • Driver installation testing • Address testing (nexus) • Max number of targets • Max number of LUNs • Max number of I/O • Performance • Data integrity (patterns) • Sleep/hibernate • Tape/medium changers • Error handling 	<ul style="list-style-type: none"> • Link speed negotiation • Endpoint registrations • Zoning, ACL, • Flow control, priority pause • E-Port (ISL) switch cascade • VLAN • Link up/down reset • Statistics • FC Fabric services • Latency • Error handling 	<ul style="list-style-type: none"> • Max Targets capability • Max LUN capability • Max I/O (queue full) • Max capacity (volumes) • All CDB (norm, vendor unique) • Persistent Reservations • Reserve and Release • Variable Block Size (512/520) • Error Handling

This list is far from exhaustive. To validate error responses, the test product must have the ability to emulate 100's of initiators and targets and provide the ability to inject errors within any level of the protocol hierarchy. The SANBlaze VirtualLUN offers all of this plus command customization capability to create vendor unique commands, adjust latency and the ability to author, save and retrieve custom profiles.

Testing FCoE Storage Protocol

SANBlaze Technology, Inc.

5 Clock Tower Place
Suite 100
Maynard, MA. 01754

Phone
(978) 897-1888

Fax
(978) 897-3171

E-mail
info@sanblaze.com

FCoE Test Challenges (cont.)

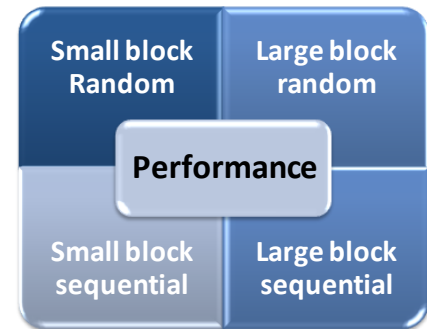
The best FCoE equipment will also include a four-corner I/O performance measurement and traffic generation.

Small block random- Typical for database (OLTP), One single disk yield ~150 I/O per sec. Arrays, have dozens of disk and Caches

Small block sequential- Typical of operating systems 1000's of I/O per second per disk, Test OS boot, different versions updates, RH, SLES, windows, updates, patches etc.

Large block Random- Typical of VOD, IPTV playback (multiple streams).

Large block sequential- Typical of Logging, journaling, data acquisition. These applications generate the highest bandwidth in system, virtually no idle time.



Summary

Optimal FCoE test equipment will feature both FCoE and native FC test ports for gateway validation. They will feature scriptable test execution to facilitate deterministic regression testing. SANBlaze's VirtualLUN emulation systems deliver all the features discussed in this paper. SANBlaze systems will assist product developers, Q/A, test and manufacturing professionals in delivering products with improved time to market, product quality and reduced product development costs. With the right test equipment, the transition to FCoE- Ethernet based storage can be made confidently and quickly.

For more information please visit the SANBlaze web site at: www.sanblaze.com or send email info@sanblaze.com



SANBlaze Technology, Inc. is a pioneer in SAN Emulation technologies and a leading provider of storage solutions for embedded systems. SANBlaze emulation products provide storage engineers, test, QA and manufacturing teams with scalable, high performance and configurable emulated environments for Fibre Channel, SAS and FCOE targets and initiators. SANBlaze emulation systems are installed at most major storage hardware and software vendors worldwide.