Using StoreEngine™ and StorePak™ for Embedded RAID Direct Attach Storage (DAS)

Abstract

StoreEngine and StorePak provide embedded RAID Direct Attached Storage (DAS) capability for block-level access to onboard SATA storage via standard PCIe interfaces, and via optional Fibre Channel and 10Gb Ethernet interfaces. Embedded RAID is ideal for bandwidth driven applications such as the real-time acquisition of wide-band sensor, radar and video data streams. StoreEngine's embedded RAID can support up to 750 MBytes/s of sustained performance for a single blade and can be scaled to higher levels through the use of additional StoreEngines and/or StorePaks. StoreEngine (and optional StorePaks) provide up to 3TB of file storage per blade using either fixed or easily removable SSD storage media.

Using StorePak & StoreEngine as Embedded RAID Direct Attached Storage

StoreEngine/StorePak Overview

StoreEngine and StorePak are flexible storage building blocks that are used to implement flexible, scalable data storage systems. **StoreEngine** is an ultra-high performance *Storage Controller* blade that can also host up to 1.5 TB of non-removable on-board SSD storage. The StoreEngine single slot blade can simultaneously serve block data (like a disk drive or RAID system) as well as NAS file sharing (like a NFS/CIFS file server). **StorePak** is a *Storage* expansion blade that can host up to 3 TB of easily removable and hot swappable SSD storage.

StoreEngine, with optional StorePaks, provides unmatched storage capability, ultra high performance and high capacity all within a small size, weight, and power (SWaP) footprint. StoreEngine is ideal for high bandwidth embedded data recording, NAS file serving, and general purpose RAID applications. StoreEngine is easily scalable in capacity and performance by simply adding additional StoreEngine and/or StorePak blades.

Using StoreEngine/StorePak for Embedded RAID (DAS) Operation

StoreEngine supports Direct Attached Storage (DAS, aka embedded RAID) applications using PCIe, Fibre Channel, or 1/10 Gb Ethernet connections. Up to six PCIe x4 backplane connections, as well as four ports of 1Gb Ethernet are built in to StoreEngine. Fibre Channel and 10GbE connectivity is available using an optional Rear Transition Module (RTM). StoreEngine also provides options for both fixed and removable SSD storage media. The fixed media version is based on the StoreEngine blade, each hosting three fixed SSD drives. The removable media version uses a combination of StoreEngine VPX blades, along with StorePak VPX removable drive assemblies. Each removable StorePak driver carrier unit hosts six SSD drives.

StoreEngine and StorePak(s) can be utilized in three different ways to provide embedded RAID storage capabilities. The first of these is shown in figure 1 below, where StoreEngine is used stand-alone to provide up to 1.5 TB of fully managed RAID 0/5 storage. This method also supports simultaneous operation of embedded RAID along with StoreEngines NAS and Recorder modes of operation.

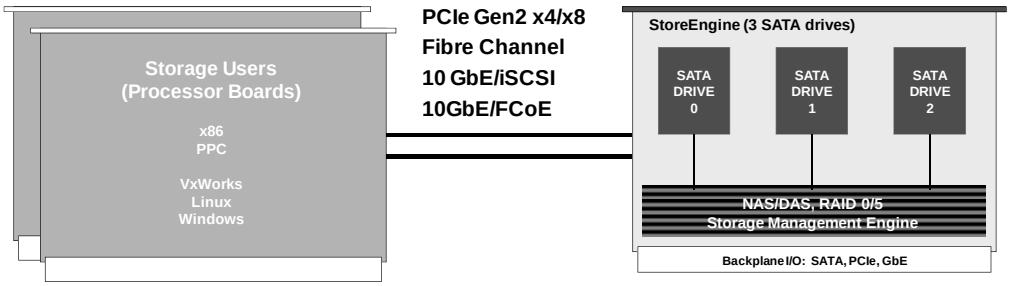

Figure 1. StoreEngine embedded RAID 0/5 -- 1.5TB *managed* non-removable SSD storage

Method 2, shown in figure 2 below, implements the raw SSD storage using one or more removable StorePaks, rather than embedding SSD the storage on StoreEngine. This provides the same capabilities (simultaneous support for embedded RAID 0/5, NAS, and Recorder) as was described for method 1, but with the added benefit of easily removable, hot-swappable StorePaks.

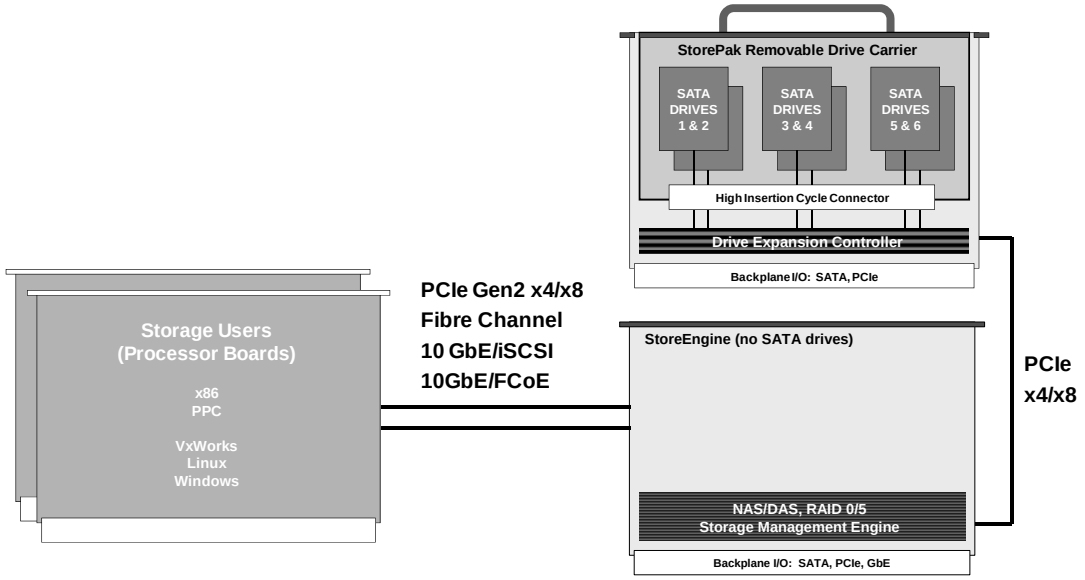

Figure 2. StoreEngine embedded RAID 0/5 -- 1.5TB managed *removable* SSD storage

And method 3, shown in figure 3 below, utilizes StorePaks(s) only to provide a more limited unmanaged embedded RAID 0/1 only solution. This approach allows a processor board to connect directly to StorePak(s) using PCIe. The NAS and Recorder modes of operation are not supported in this simplified usage model.

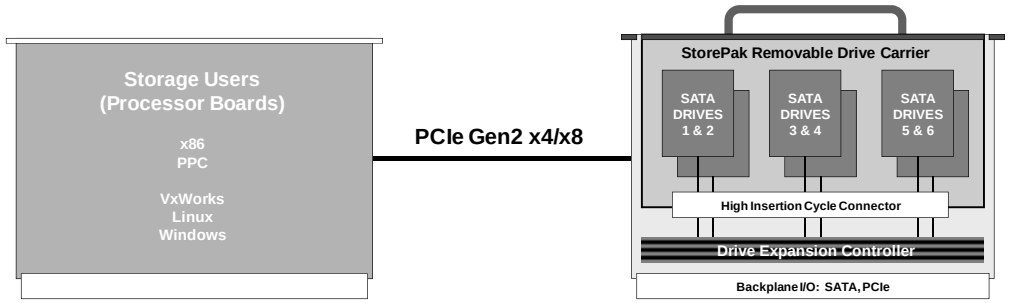

Figure 3. StoreEngine embedded RAID 0/1 -- 1.5TB *unmanaged* removable SSD storage

Data Storage Models

Storage usage models are divided into two main categories: Direct Attached Storage (DAS), which provides ***block level*** storage access (including RAID), and Network Attached Storage (NAS), which provides ***file level*** storage access. The difference between these two lies largely in where the file system is hosted, as illustrated in figure 4 below. For DAS storage systems, the storage client hosts a *local file system*, while for a NAS system, a *server file system* is hosted within the NAS storage device itself.

NAS storage is always accessed by clients via an Ethernet network interface using protocols such as NFS, CIFS/SMB, and FTP. DAS storage may be accessed using a variety of interfaces/protocols, including PCIe, Fibre Channel, Ethernet/iSCSI, and Ethernet/FCoE.

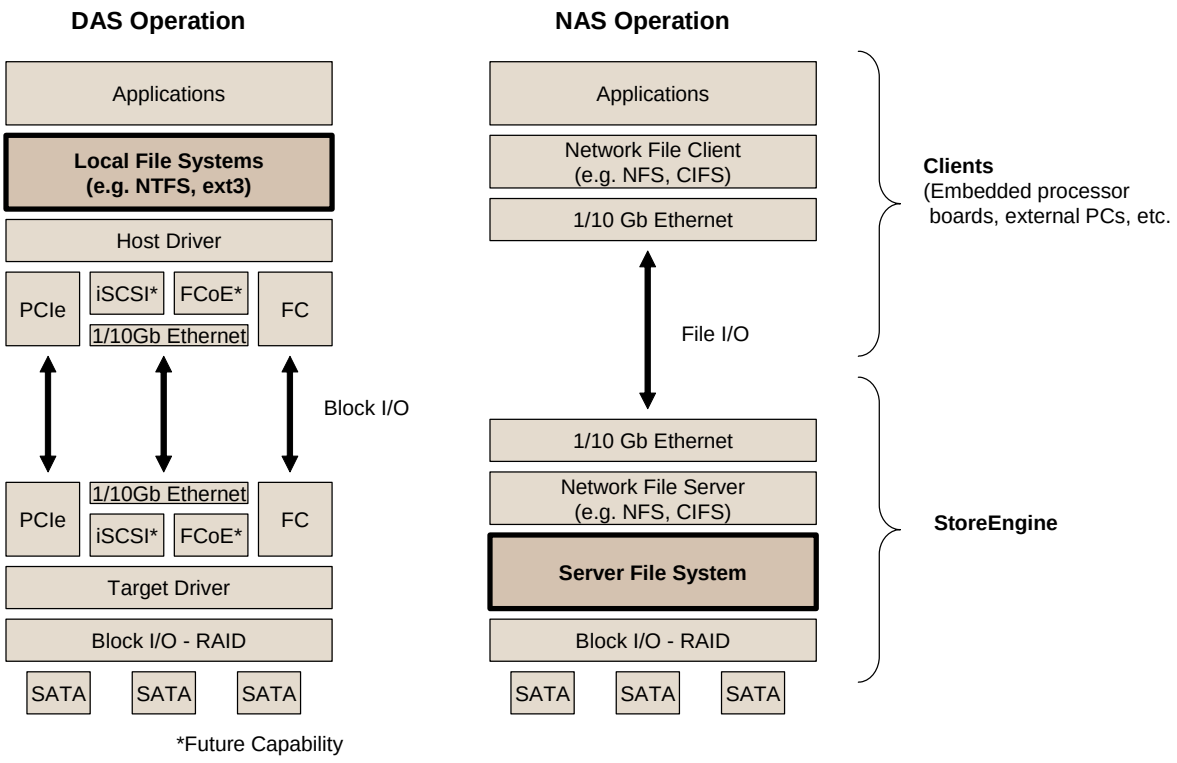

Figure 4. StoreEngine DAS operation contrasted with NAS operation

StoreEngine Embedded RAID (DAS) Operation

StoreEngine's embedded RAID Direct Attached Storage (DAS) capability provides block-level access to StoreEngines onboard storage for bandwidth driven applications. Access to storage is provided via a standard backplane PCIe fabric, as well as support for multiple optional interfaces such as Fibre Channel, Ethernet/iSCSI, and Ethernet/FCoE. StoreEngine's embedded RAID can support up to 750 MByte/s of sustained performance for a single blade and can be scaled to higher levels through the use of additional StoreEngines and/or StorePaks.

In DAS applications, the client processor hosts a file system that is provided by the client's operating system. The client thus utilizes StoreEngine storage in a low level block mode. The allocation and use of these low level storage blocks is controlled completely by the client file system, thus (unlike NAS storage) DAS stored data cannot typically be shared between multiple clients. Data transfer rates for DAS storage are typically very high, up to 750 MB/s for a standard StoreEngine configuration.

DAS Protocols

DAS clients utilize block level storage access protocols to store or access data. StoreEngine supports several commonly used DAS block level protocols, including PCIe Direct Connect, Fibre Channel, Ethernet/iSCSI*, and Ethernet/FCoE*.

PCIe Direct Connect

PCIe Direct Connect is an extremely efficient method of adding high performance RAID storage to a PCIe enabled embedded processor board. A lightweight StoreEngine PCIe client driver runs on the client board, supporting sustained data transfer rates to and from StoreEngine of up to 750 MByte/s. StoreEngine client drivers are available for both VxWorks and Linux, for both x86 and PowerPC architectures.

Fibre Channel

Fibre Channel (FC) is a multi-gigabit storage networking network technology which has become the standard connection type for storage area networks (SAN). Fibre Channel networks use the Fibre Channel Protocol (FCP) as a transport protocol to transport SCSI commands over the network. Client applications leverage a Fibre Channel client driver that communicates with StoreEngine.

iSCSI (*future capability)

iSCSI is an IP based storage networking standard which transfers SCSI storage commands over IP (Ethernet) networks. iSCSI can be used to transmit data over Ethernet networks for storage and retrieval by providing a network transport for clients to send SCSI commands and data to/from StoreEngine.

FCoE (*future capability)

Fibre Channel over Ethernet (FCoE) provides encapsulation of Fibre Channel frames over Ethernet networks. This allows Fibre Channel to use Ethernet networks as a transport while preserving the basic Fibre Channel protocol. FCoE maps Fibre Channel natively over Ethernet by replacing the FC0 and FC1 layers of the Fibre Channel stack with Ethernet. By retaining the native Fibre Channel constructs, FCoE also allows integration with existing Fibre Channel networks. In contrast to iSCSI which runs on top of TCP and IP, FCoE operates directly above Ethernet in the network protocol stack.

StoreEngine DAS Connectivity Options

PCIe Direct Connect DAS Connection

PCIe Direct Connect applications leverage standard PCIe fabrics for client connectivity to StoreEngine. The PCIe connections are typically made using embedded backplane high speed fabrics (figure 5), though in some situations the connections may also be made using optional RTM PCIe access (figure 6).

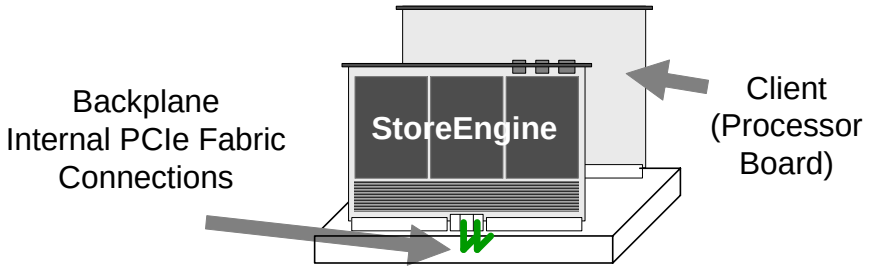

Figure 5. PCIe connection through internal backplane high speed fabric

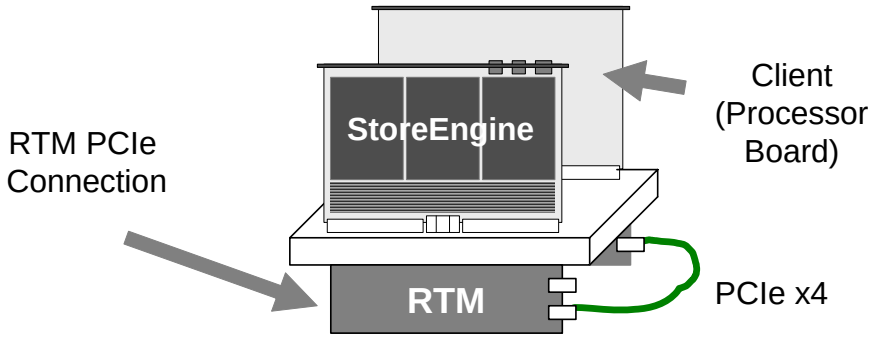

Figure 6. PCIe connection using an optional RTM

Fibre Channel DAS Connection

Fibre Channel requires the use of the optional Fibre Channel Rear Transition Module in conjunction with StoreEngine, as shown in figure 7. Client boards may use standard Fibre Channel HBA adapters to access StoreEngine using these protocols, along with use of standard Fibre Channel initiator drivers.

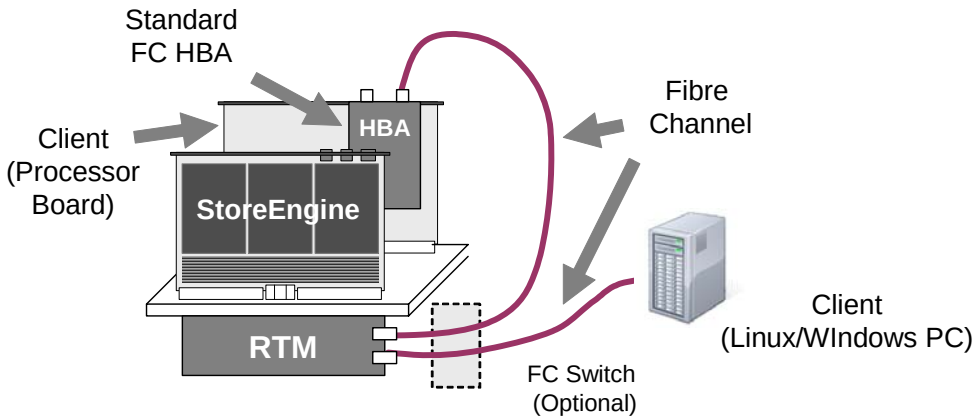

Figure 7. Fibre Channel DAS connection using an optional RTM

10 Gb Ethernet (iSCSI* or FCoE*) DAS Connection

FCoE and iSCSI operation over 10Gb Ethernet requires the use of the optional 10GbE Rear Transition Module in conjunction with StoreEngine (figure 8). Client boards use standard or TOE 10GbE adapter boards to access StoreEngine using iSCSI, and FCoE protocols, along with use of standard software or TOE based iSCSI or FCoE initiator drivers.

**iSCSI and FCoE are planned features*

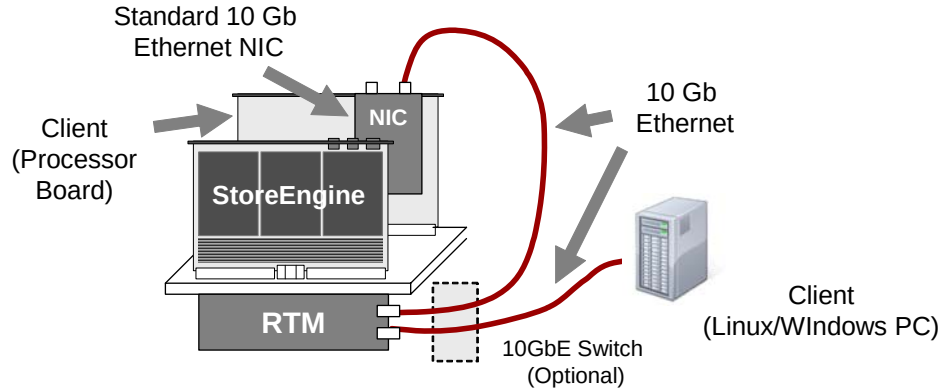

Figure 8. Ethernet (iSCSI, FCoE) DAS connection using an optional RTM

StoreEngine DAS - Optional Capabilities

RAID Options

StoreEngine storage can be configured to operate as either RAID 0 or RAID 5. Both RAID 0 and RAID 5 modes aggregate the storage of all of the StoreEngine drives into one or more "logical drives". RAID 0 and RAID 5 differ greatly in levels of write performance and data protection.

RAID 0 stripes data across all of the StoreEngine drives, which provides the highest possible read and write performance, as well as the highest usable capacity. RAID 5 also stripes data across drives, but in addition creates a parity block for each data stripe which provides data protection and continued operation (at a lower performance level) in the event of a drive failure. Since one "drives worth" of performance and capacity is reserved for parity in RAID 5, the capacity and write performance when using RAID 5 is approximately 2/3 that of RAID 0 for a standard three drive StoreEngine configuration.

Logical Volumes

StoreEngine supports the optional use of Logical Volumes. Logical volumes provide a method for dividing the available storage into a number of separate volumes. For example, a single RAID 5 array can be created that aggregates all storage. This array can then be divided up into several logical volumes, which may be used for different purposes. (i.e., one volume exported as a block RAID device to a DAS client, while another volume is exported for FTP use). The use of logical volumes also offers the advantage of being able to grow or shrink volumes dynamically.

Encryption

StoreEngine provides an option to encrypt all data as it is written to disk. Note that while StoreEngine has hardware accelerated encryption, the use of encryption may still affect performance in some situations.

Secure Erase

Store Engine provides a Secure Erase operation which may be invoked via the Web Management interface. Secure Erase fully erases all data on the selected drive, restoring the drive to an unused condition. Note that Secure Erase can take several minutes to performance on a Solid State drive, and up to several hours to perform on a rotating hard drive.

Quick Config

StoreEngine provides two "quick configuration" options. The first method provides a set of pre-defined configurations that can, with a single mouse click, fully configure StoreEngine to one of six commonly used configurations. The second method allows users to create and store "configuration snapshots". These configuration snapshots can be reapplied with a mouse click to restore a StoreEngine to a known configuration. These configuration snapshots can also be transported to other StoreEngines, allowing users an easy way to "clone" configurations onto additional StoreEngine units.

DAS Performance Optimization

With all data transfer and storage devices, the specifics of how the device is used will affect performance. While there are numerous tunable parameters in DAS applications, there are three basic keys to maximizing performance: 1) maximize the data transfer sizes, and 2) use sequential access to files to the maximum extent possible.

Maximize Data Transfer Size

Data transfer sizes should be optimized on two levels. First, applications should perform reads and writes in the largest block sizes possible. For example, a single 256K data read from StoreEngine will be much more efficient that sixteen 16K reads.

Maximize Sequential Access

Sequential data access is always more efficient that random access. There are two reasons for this. Sequential access allows multiple accesses to be gathered in to fewer, larger accesses using read-ahead and write-behind techniques. And second, solid state drives and especially rotating hard drives both perform better in sequential access scenarios.

StoreEngine Web Management Interface

StoreEngine DAS/RAID functionality is enabled and managed via a simple web based management interface, which supports monitoring and configuration of all StoreEngine interfaces, operating modes, and storage options. Specific web management capabilities include:

- BIT status (self test, voltages, currents, temperatures)
- Storage status (available/used capacity, status, errors)
- Network statistics
- Interface Status (link status, errors)
- Performance
- Enable/Disable protocols & options
- Manage NAS and DAS exports
- Security and permissions
- Secure Erase

An example of a StoreEngine management page (this specific one is used to create a PCIe DAS export or "target") is shown in figure 9. The other main management categories are shown in the menu on the left side of the page.

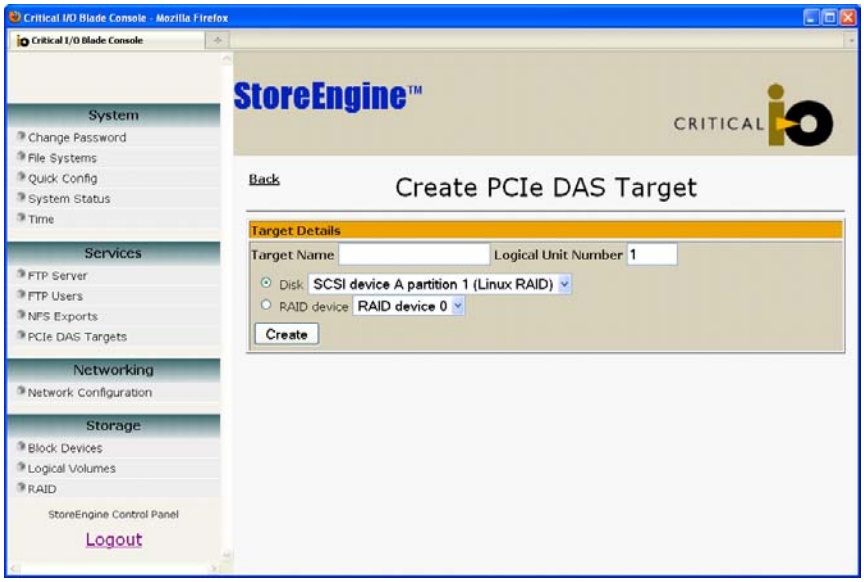

Figure 9. StoreEngine web management interface

StoreEngine Hardware Architecture

The StoreEngine hardware architecture is shown in figure 10. The key components are:

- Storage Management Processor – The highly power efficient PowerPC based storage management processor provides overall control of the StoreEngine, implementing storage management functionality as well as implementing the NAS file sharing protocols.

- Hardware Accelerated RAID – The RAID hardware function provides acceleration of RAID-5 operations.

- PCIe Switch – An 8 port PCIe Gen 2 switch provides internal and external connectivity for StoreEngine. It connects all of the internal resources, and provides up to four backplane PCIe ports (depending on the specific StoreEngine model) for external PCIe backplane or RTM connections.

- Gb Ethernet NICs– StoreEngine has four on-board Gb Ethernet interfaces. Two interfaces provide TCP Assist Hardware capabilities.

- SATA Controller – The SATA controller manages the three internal SATA drives. It operates under the control of the Storage Management Processor.

- SATA Drives – Three internal SATA drives provide data storage. Up to four additional external SATA drives may also be connected using the RTM SATA connection. The SATA drives are used only for user data; no StoreEngine management software or configuration information is stored on the SATA drives.

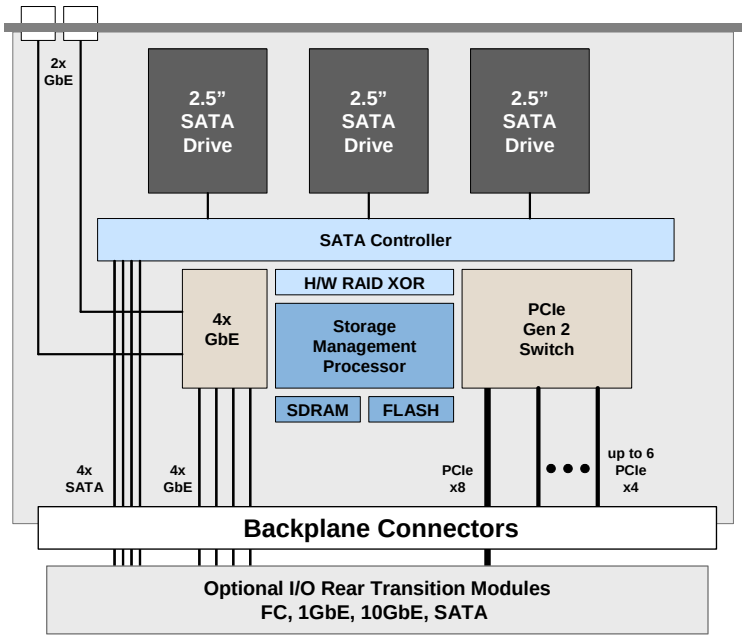

Figure 10. StoreEngine hardware architecture.

StoreEngine Software Architecture

The StoreEngine internal software architecture is shown in figure 11. The key layers in the architecture are 1) the Transport layer, which provides the implementation of the various storage access protocols, 2) the Storage Management Layer, which manages the file and block level storage resources, and 3) the Storage Layer, which provides the interface to the three physical on-board SATA SSD or HDD storage devices.

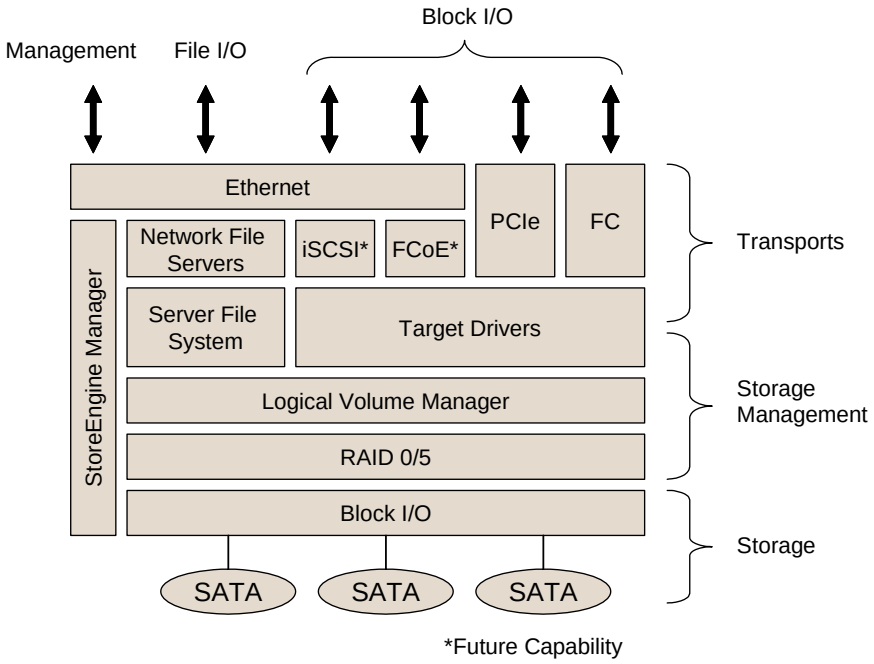

Figure 11. StoreEngine Software Architecture