Configurable High Performance Storage using StoreRack™

Abstract

StoreRack leverages Critical I/O's StoreEngine and StorePack storage blades to provide highly configurable rack mount storage solutions targeted to real-time embedded systems. This paper describes the configurable aspects of StoreRack, as well as the storage modes that StoreRack supports, including block level DAS, file sharing NAS, and scalable high speed recording.

Revision 3/20/13

CRITICAL IO

1. Introduction

StoreRack™ is a rack mount highly configurable storage system that utilizes rugged, removable and hot swappable SSD modules. Designed for use in real-time embedded applications, StoreRack provides high performance and high capacity rugged solid-state storage within a small size, weight, and power (SWaP) footprint. StoreRack is designed for high bandwidth data recording, file serving, and general-purpose RAID storage applications, where high performance and capacity, combined with a small SWaP footprint are critical.

StoreRack leverages Critical I/O's StorePak™ and StoreEngine™ modules. StorePak modules are specifically designed for easy removal and replacement of SSD storage. StorePak modules can be used stand-alone for Direct Attached Storage (DAS) operation, or they can be used in conjunction with StoreEngine storage manager modules to provide an expanded set of powerful storage capabilities. As there are no cables to unplug or tools needed to remove StorePak modules, StoreRack is ideal for applications where removal and replacement of onboard storage must be accomplished quickly and easily.

Figure 1. Five slot and three slot StoreRack chassis

StoreRack is comprised of a 3 or 5 slot OpenVPX chassis populated with combinations of StoreEngine, StorePak, and RTM I/O modules. The StoreRack chassis are 19"rack mount, air-cooled, with integral power supplies. An internal OpenVPX backplane provides PCIe mesh "fat-pipe" connectivity between all board positions, facilitating PCIe communications between board slots for scalable storage architectures. A rear transition module (RTM) capability is also provided at each board position, supporting external interface options including PCIe, 1 or 10GbE, and Fibre Channel.

Many aspects of StoreRack are highly configurable and scalable, including:

- Storage capacity: 1.5 TB to 15 TB
- Data rate: 100 MB/s to 7.5 GB/s (sustained)
- Interface type/number: PCIe, 1/10 GbE, Fibre Channel
- RAID levels: 0, 1, or 5
- Chassis size: 3 slot or 5 slot
- Modes of operation: NAS, DAS, Recording
- Recording architectures: Direct/buffered, multi-channel

StoreRack can provide simultaneous support of three different data storage architectures:

- Direct Attached Storage (aka block storage, RAID)
- High Speed Data Recording
- Network Attached Storage (aka File Server)

These modes of operation are discussed in more detail later in this paper.

2. StoreRack Building Blocks

StoreRack systems are comprised of a set of common hardware building blocks as described in table 1 below. The key hardware components are StoreEngine and StorePak; flexible storage building blocks that can be used to implement a wide range of data storage systems. Both StoreEngine and StorePak provide rich PCIe connectivity, with up to eight x4 PCIe backplane ports per VPX board. These ports are used for connections to the user's data sources as well as implementing the PCIe connection network between all StoreEngines and StorePaks within a StoreRack.

StoreEngine hosts Critical I/O's multi-mode storage management software which supports a rich set of storage interfaces, management capabilities, and modes of operation. A comprehensive web-based storage management user interface allows the configuration, control, and management of these capabilities.

StorePaks can be used alone to provide a raw PCIe connected SSD storage capability. An example is a PCIe connection from StoreRack from another rack containing one or more SBCs. All other modes of StoreRack operation, including NAS, recording, and Fibre Channel DAS, require the inclusion of one or more StoreEngine modules in the StoreRack. StoreEngine's embedded storage management processor hosts powerful and flexible storage management software to support these additional modes.

StoreRack external I/O interfaces are normally provided via a complement of Rear Transition Modules (RTMs), but specialized interface options can also be accommodated using optional VPX XMC/PMC carrier cards.

Table 1. StoreRack building blocks.

StorePak and StorePak Carrier	StorePak is a two-part VPX storage expansion blade that can host up to 3TB of removable SSD storage. The StorePak™ removable SSD module is designed for easy removal and replacement, while the StorePak carrier module remains plugged into the VPX backplane, providing hot-swap capability while avoiding wear and tear on the VPX backplane connectors.
StoreEngine	StoreEngine is an ultra-high performance VPX storage controller blade) that hosts up to 1.5TB of non-removable on-board SSD storage. StoreEngine hosts powerful storage management software that simultaneously provides high performance recording functionality; provides block data (aka RAID/DAS); and provides NAS file sharing (aka NFS/CIFS file server).

Mezzanine RTM	The RTM-VPX-MC1 rear transition module provides several rear I/O connectivity options for StoreRack applications. The RTM provides a PMC/XMC site that can host mezzanine I/O cards for additional interface options in lab and evaluation applications, as well as providing standard optical and copper Gb Ethernet connectivity, and SATA expansion connectivity.
PCIe RTM	The RTM-VPX-PCIe rear transition module provides Gen1 or Gen2 PCIe rear I/O cable connectivity options for StoreRack applications. The RTM provides one x8 and two x4 PCIe iPass connectors for external PCIe Gen2 connectivity, as well as providing standard optical and copper Gb Ethernet connectivity.
FC RTM	The RTM-VPX-8GFC rear transition module provides a dual 8Gb Fibre Channel interface capability. Critical I/O Fibre Channel products employ an architecture which completely offloads all protocol processing into dedicated silicon for high sustained throughput and low latency.
10GbE RTM	The RTM-VPX-10GBE rear transition module provides a dual 10 Gb Ethernet interface option for StoreEngine. Critical I/O Ethernet products employ an architecture which offloads much of the protocol processing into dedicated silicon for high sustained throughput and low latency.
VPX-Mezz	VPX-Mezz is a VPX dual XMC/PMC carrier board that be used to expand StoreRack front panel I/O connectivity through the use of standard XMC/PMC modules.

3. Standard and Custom StoreRack Configurations

StoreRack is highly configurable; there are four standard configurations, along with a multitude of optional custom configurations. The standard StoreRack configurations are shown in the table below, but virtually any combination of StorePaks, StoreEngines, and RTMs can be supported. Several examples of custom configuration variations are shown later in this paper.

Table 2. Standard StoreRack configurations

Model	Slots	Module Set	Modes	Standard Interfaces	Capacity
SR19R6U3	3	3x StorePak	DAS (RAID 0/1)	up to 3x PCIe x8	up to 9 TB
SR19RC6U3	3	2x StorePak 1x StoreEngine	DAS (RAID 0/1/5) NAS Recording	up to 2x PCIe x8 up to 4x GbE	up to 7.5 TB
SR19R6U5	5	5x StorePak	DAS (RAID 0/1)	up to 5x PCIe x8	up to 15 TB
SR19RC6U5	5	4x StorePak 1x StoreEngine	DAS (RAID 0/1/5) NAS Recording	up to 4x PCIe x8 up to 4x GbE	up to 13.5 TB

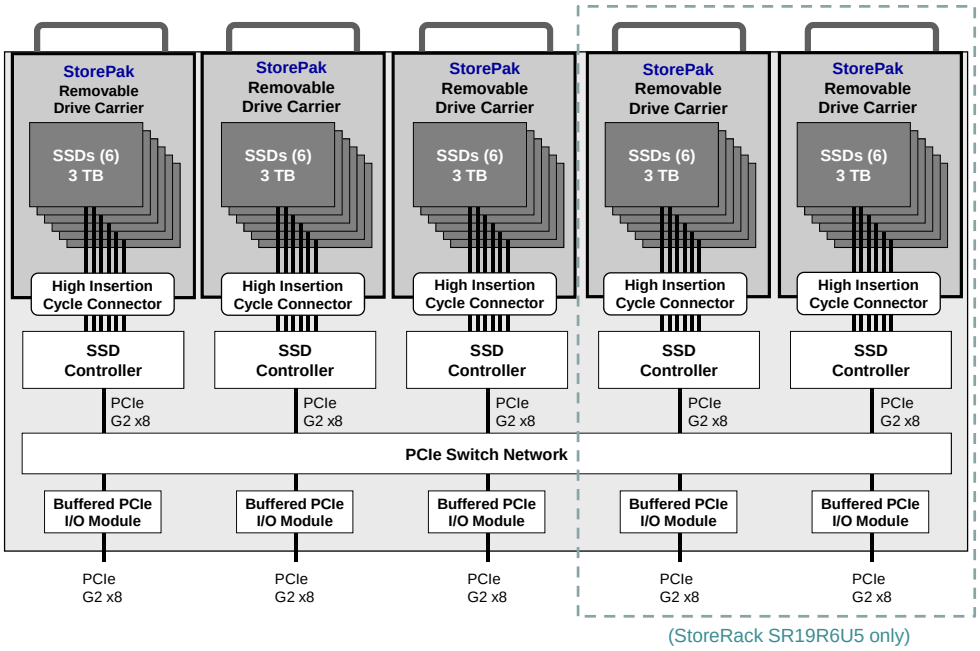

Figure 2. StoreRack standard configurations for unmanaged PCIe DAS operation

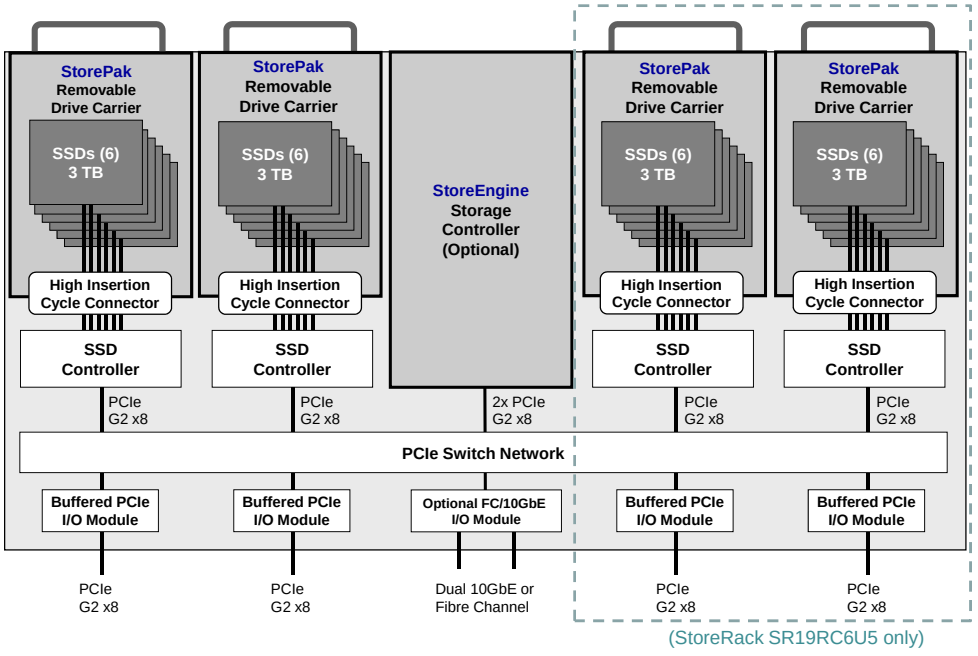

Figure 3. StoreRack standard configurations for managed DAS, NAS, and recording operation

4. StoreRack Modes of Operation

StoreRack can provide concurrent support for Direct Attached Storage (DAS, aka RAID, block storage), Network Attached Storage (aka NAS, file server) and high performance recording. Note that at least one StoreEngine storage manager module must be included in a StoreRack for managed DAS, NAS and Recorder modes.

Table 3. Characteristics of StoreRack operating modes

StoreRack Mode	StoreEngine Needed?	Typical Performance	Characteristics
DAS (RAID/block storage) *unmanaged*	No	500-1500 MB/s	Low level storage (i.e. disk blocks) are presented directly to the storage users (hosts). Hosts manage storage directly, typically through the use of standard file systems. Supports RAID 0/1.
DAS (RAID/block storage) *managed*	Yes	500-1000 MB/s	As above, but low level storage is first aggregated and virtualized prior to presentation to storage users (hosts). Supports RAID 0/1/5.
NAS (file server)	Yes	50-200 MB/s	Low level storage is managed by StoreEngine(s). Data is accessed by storage users via file sharing protocols such as NFS or CIFs/SMB.
Recorder	Yes	1-6 GB/s	Low level storage is managed by StoreEngine(s) through a local high performance distributed recording file system. Protocol free data streams are transferred directly between the recording data sources and StoreRack.

Note that the performance numbers mentioned in this paper represent the sustained performance levels that can be achieved in a properly architected system that is used in an optimum manner. The performance level that is actually achieved in any given application depends on the details of the system architecture and the details of how the application accesses StoreRack storage.

Direct Attached Storage Modes

StoreRack supports two modes of Direct Attached Storage (DAS, aka RAID/block storage) operation. The *unmanaged* DAS mode uses only StorePak modules, which are PCIe connected to the block storage users (hosts), typically processing boards. RAID 0 and RAID 1 are supported, with storage capacity of up to 15 TB. Each StorePak is presented to the block storage user (host) as one or more “drives”. The host provides all storage management capability, typically through the use of a standard file system. PCIe is connections are used for this mode, supporting data rates of up to 1.5 GB/s per PCIe x8 connection.

StoreRack's *managed* DAS mode utilizes a combination of StoreEngines and StorePaks. StoreEngines run storage management software, which aggregates and virtualizes all of the storage that is available on the designated StorePaks. The management software allows the user to define multiple RAIDs, logical volumes, and multiple exports to Fibre Channel and/or PCIe connected hosts. This mode supports RAID 0, 1, and 5, with data rates of up to 1.0 GB/s per PCIe x8 interface, and capacity of up to 13.5 TB.

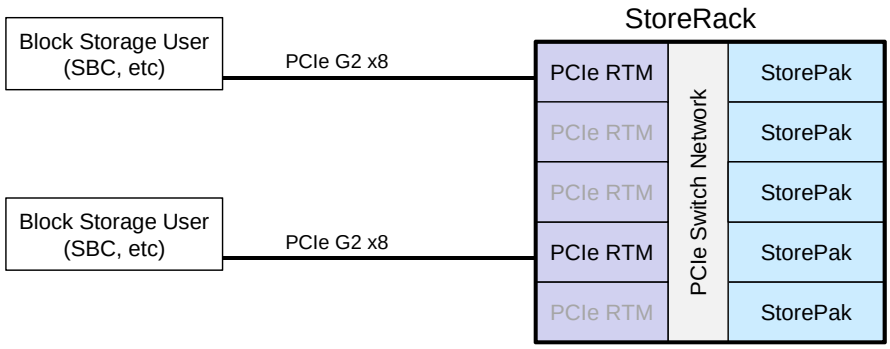

Figure 4. Unmanaged PCIe connected DAS

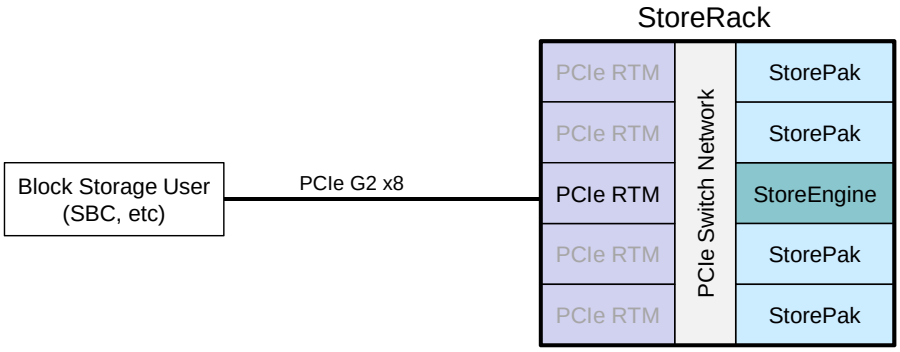

Figure 5. Managed PCIe connected DAS

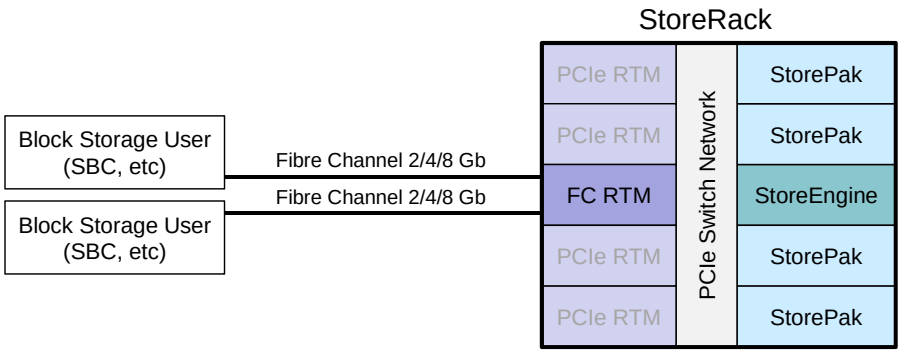

Figure 6. Managed Fibre Channel connected DAS

Network Attached Storage (File Server) Mode

StoreEngine's NAS (Network Attached Storage) provides file-level access to its onboard RAID volumes to a networked environment. This allows storage to be accessed and shared via standard file access protocols (such as NFS and CIFS) with PCs, servers, SBCs, etc., that are connected via an Ethernet network. StoreEngine supports NAS data rates of up to 200 MB/s. NAS is ideal for applications that require file-level shared access to data among multiple clients at low to moderate data rates.

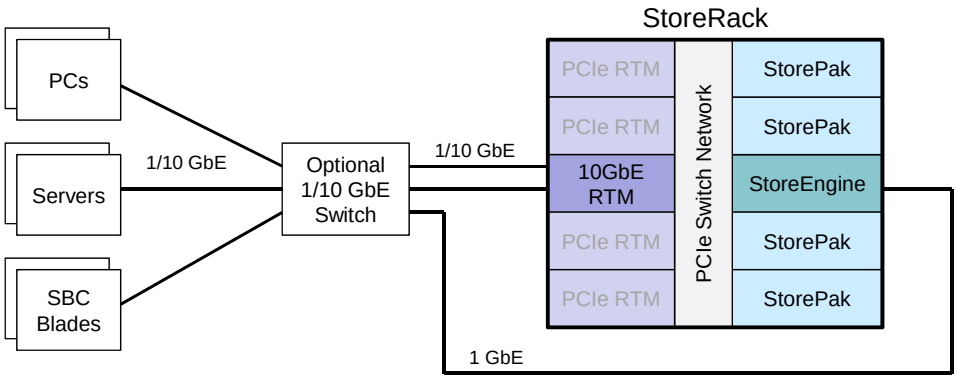

Figure 7. 1 or 10 GbE connected NAS file server

Scalable Data Recorder Mode

When configured with one or more StoreEngine modules, StoreRack provides a turn-key, flexible and scalable recording platform that can record high bandwidth data streams from ADCs, FPGAs, video streams, and other sources. By aggregating up to multiple StoreEngines and StorePaks, a recorder can support sustained recording at rates of up to 6 GB/s with a capacity of up to 13.5 TB.

Options for the recorded data source range from a data stream from a simple PCIe connected ADCs, to more intelligent PCIe connected devices such as FPGA processors, or a PCIe connected standard CPU board (SBC). The source may also consist of 10GbE network connection to record a UDP/IP data stream.

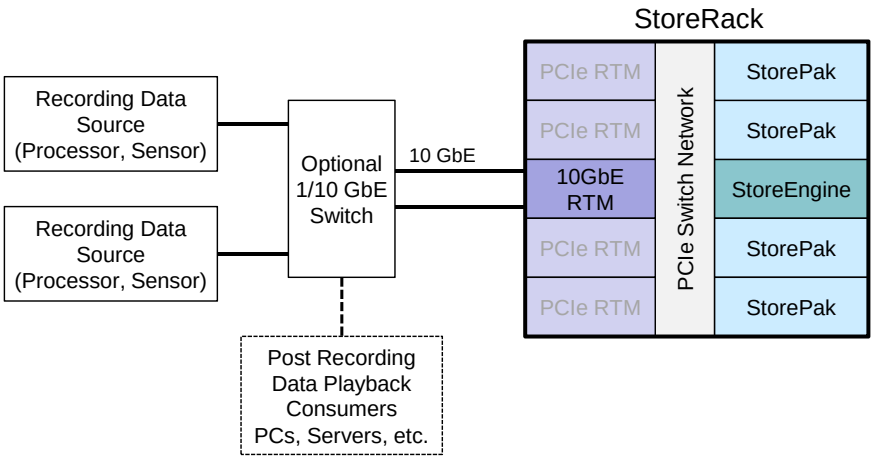

Figure 8. 10 GbE recorder

For PCIe connected data sources, a StoreRack recorder can operate in either a direct mode or a buffered mode. Direct mode results in data being transferred directly from PCIe data sources to/from the SSDs hosted on StorePak modules, without first being buffered in a StoreEngine. This contrasts with buffered mode, where data is buffered in a StoreEngine prior to being moved to the SSD storage.

Direct mode has the advantage of higher performance and parallelization, as data does not need to be first buffered in a StoreEngine. Direct mode has a further advantage of increased storage density, as multiple StorePak modules can be connected directly to separate data sources. Direct mode requires that the StorePaks be able to read or "pull" data directly from PCIe accessible memory on the PCIe data sources. A requirement of direct mode is that the PCIe data sources must provide very high performance PCIe target intefaces, and further must be able to buffer a fairly large amount of data to ensure smooth high recording rates.

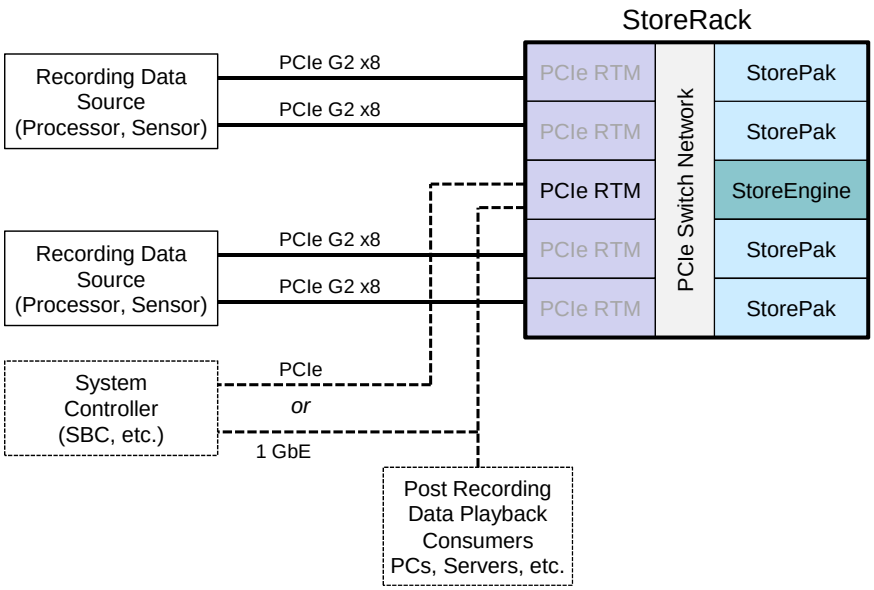

Figure 9. Four channel 6GB/s PCIe direct mode recorder.

In contrast, buffered mode operation allows data to first be buffered in a StoreEngine RAM prior to being written to SSD storage. Buffered mode has the advantage of flexibility. PCIe data sources may "push" data into data buffers on the StoreEngine(s). The timing of these data transfers is thus more flexible, and can be controlled by the data source. StoreEngine then transfers buffered data to the SSD storage on a decoupled timeline. RAID 0 and RAID 5 are supported in buffered mode. Buffered mode also supports a "distributed" variation, where the performance of multiple StoreEngine may be aggregated to record data from a single ultra-high bandwidth stream of > 2GB/s

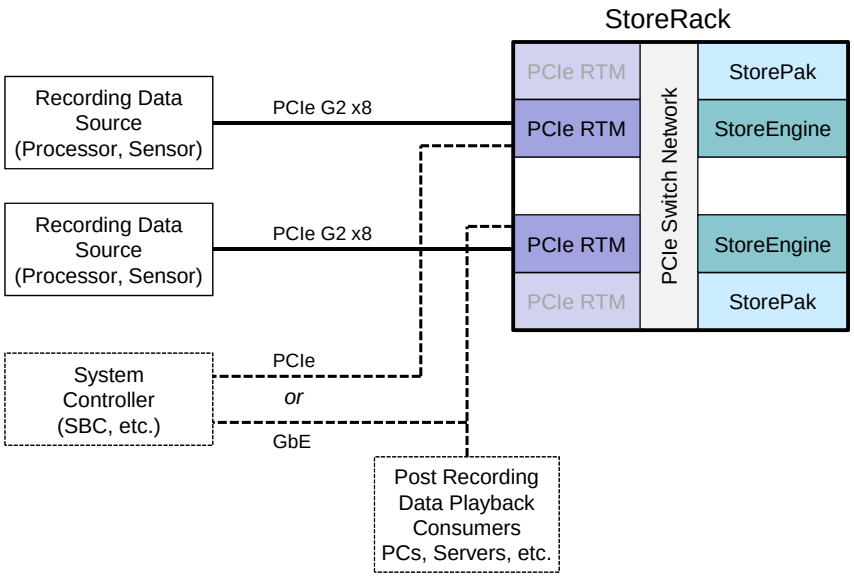

Figure 10. Two channel 2GB/s buffered mode recorder

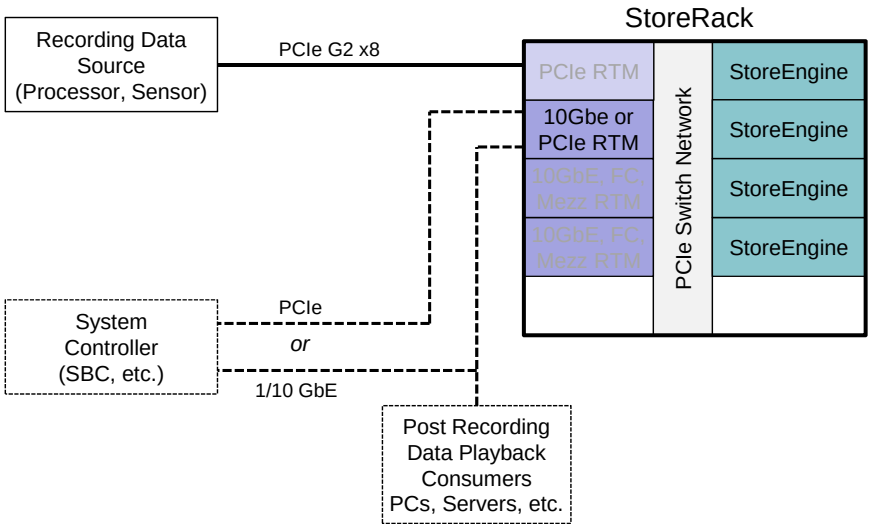

Figure 11. High bandwidth single channel 2 GB/s distributed mode recorder

Data that is recorded on a StoreRack recorder may be played back via the same path that was used for the recording source, or alternatively via any other supported recording interface, or via NAS file access. For example, data that is recorded using a PCIe connected ADC data source may be retrieved via a PCIe connected processor board, or via an Ethernet connected PC using NFS. Critical I/O also provides a specialized 10Gb Ethernet file transfer protocol that allows recorded data to be offloaded from StoreRack to standard PCs or servers at sustained rates of >1 GB/s

All of the functionality to control both the recording of data onto multiple StoreEngine/StorePak blades, and to play back the aggregated data from these multiple blades, is built into the StoreEngine recording software. All recorder functionality is managed through a single StoreEngine board, using either the

built-in web based management interface, or through the use of a Recording Driver that is hosted on a user's System Controller SBC.

5. External I/O Connectivity

Physical I/O connections to StoreRack are made using Rear Transition Modules. The interface options included are specific to the StoreEngine/StorePak/RTM complement of the StoreRack.

For each StoreEngine (and its associated RTM):

- 1x 1GbE mini-DB9 front panel
- 1x Serial mini-DB9 front panel
- 2x 1GbE (via optical LC and copper RJ45 on the PCIe or Mezz RTMs)
- 1x PCIe Gen2 x8 (via a x8 iPass copper connector on the PCIe or Mezz RTMs)
- 1x Serial (via a DB9 connector on the PCIe or Mezz RTMs)
- 2x 10GbE (via optical LC on the 10GbE RTM)
- 2x 2/4/8 GB Fibre Channel (via optical LC on the Fibre Channel RTM)

For each StorePak (and its associated RTM):

- 2x PCIe Gen1/Gen2 x4 (via x4 iPass copper connectors on the PCIe RTMs)
- 1x PCIe Gen1/Gen2 x8 (via a x8 iPass copper connector on the PCIe RTMs)

All cables are routed from the StoreRack RTM through and access panel located at the rear of the StoreRack.

As noted above, external PCIe cable connections are made via standard iPass connectors, following the conventions of the PCIe Cable Specification. This allows the use of off the shelf PCIe cables from suppliers such as Molex. Copper PCIe cable lengths should be limited to a maximum of two meters for Gen 2 operation. Alternatively, PCIe optical cables are also available that are compatible with the standard StoreRack iPass connectors. These cables feature built-in optical transceivers, which allow for significantly longer cable lengths. Use of these cables requires that iPass connectors on the users PCIe interfaces supply the power connections as defined in the PCIe cable specification.

6. Optional XMC/PMC Carrier VPX Boards

StoreRack can be configured with optional VPX dual XMC/PMC carrier boards. This allows StoreRack to accommodate external interface options outside of the standard StoreRack interfaces. In most cases, specialized interface options will require some modification to StoreEngine management software.

7. Optional User Supplied OpenVPX Boards

As StoreRack is based on a standard OpenVPX chassis, StoreRack configurations with unused slots can host user-supplied OpenVPX boards. Examples might include specialized interface boards, DSP/FPGA processor boards, or standard SBCs.