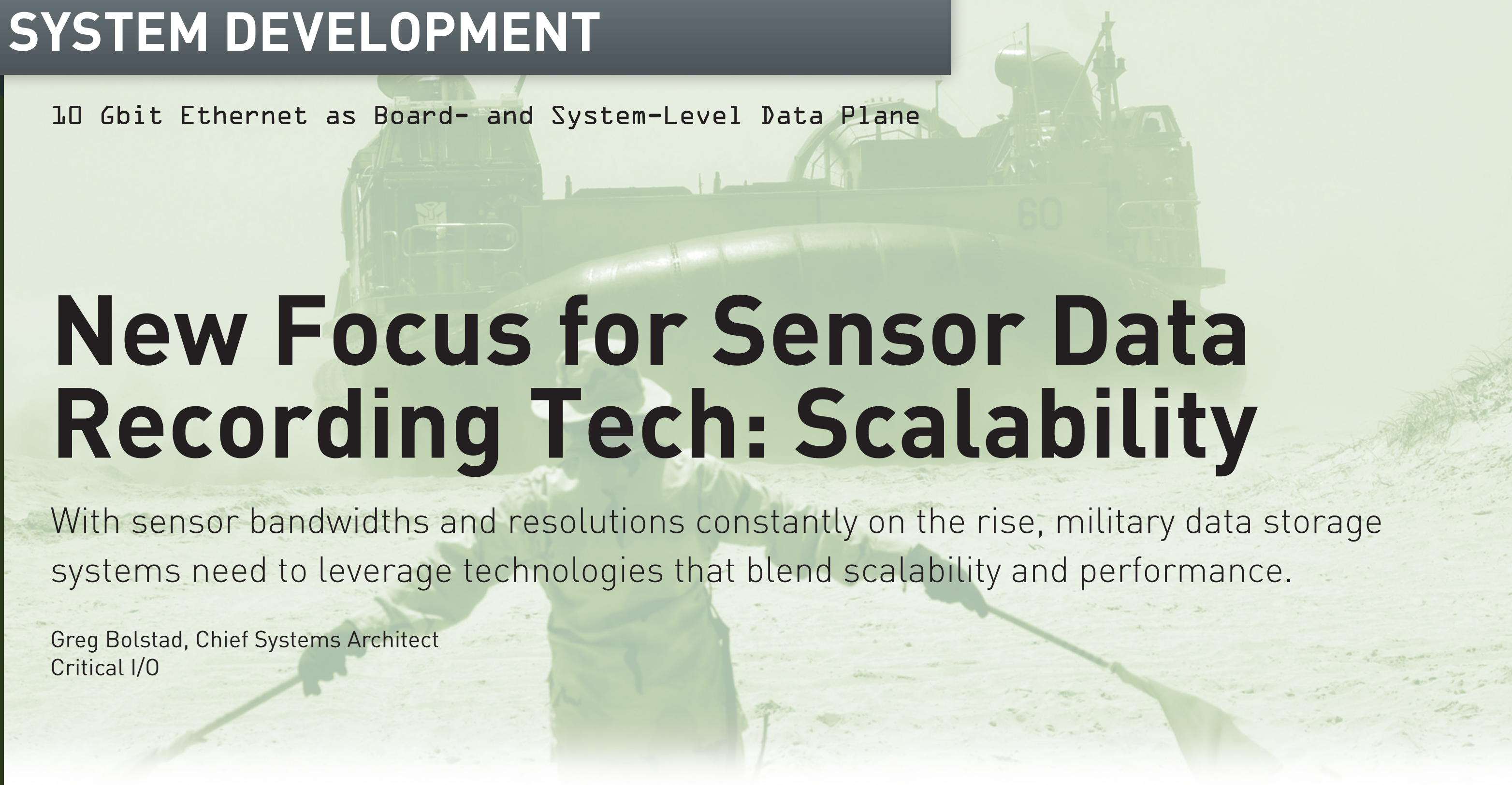

New Focus for Sensor Data Recording Tech: Scalability

With sensor bandwidths and resolutions constantly on the rise, military data storage systems need to leverage technologies that blend scalability and performance.

Greg Bolstad, Chief Systems Architect
Critical I/O

The defense and intelligence communities face an increasing reliance on high-resolution and high-bandwidth sensors of varying types. Raw data rates in these systems are increasing—due to higher bandwidth and higher resolution sensors—as are mission durations. The majority of these sensor systems require that some or all of the raw or partially processed data be recorded for later analysis, post-processing or archiving. The high data rate and capacity requirements place a huge burden on sensor data recording subsystems. This burden is often compounded by severe SWaP constraints and the need for high reliability in sometimes demanding environments.

Because sensor bandwidths and resolutions are increasing so rapidly, the storage capacity and bandwidth requirements of the sensor system are also escalating rapidly. Thus, it is essential that the recording storage solution be able to scale so that these increasing system recording storage requirements can be accommodated without the need to discard the recording subsystem architecture and start over again.

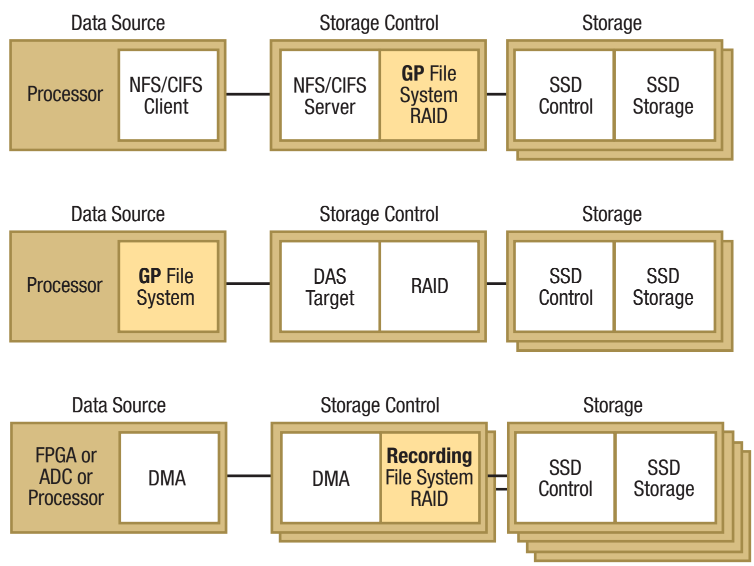

Figure 1

These three data recording system modes each offer varying degrees of performance and scalability.

Three Main Components

A recording storage system is comprised of the following three main components: storage controllers, storage modules and the interfaces. Storage controllers implement the intelligence needed to manage storage resources. Storage modules provide the raw data storage—often SATA SSD based. The interfaces comprise not just the physical recording/playback interfaces, but also the associated protocols. A scalable recording storage system must allow for growth—in both capacity and bandwidth—and flexibility in all three of these areas.

Storage technology, especially SSDs and controllers, is evolving rapidly. A recording building block approach that leverages the latest in commercial and enterprise storage technologies can meet the data recording needs of current and future sensor platforms. In particular, blade-based recorder building blocks, including storage controller blades and storage blades, can be used to build a wide variety of military sensor recording systems. These systems can support scalability in different dimensions, allowing them to meet the current and future needs of a wide range of sensor platforms.

Sensor Data Recorder Usage Models

The majority of sensor data recording applications are mission-oriented, with data typically captured for a large part of each mission. Generally after mission completion, the full set of acquired data must be transported to an analysis/post-processing/archiving facility. In the mission-oriented model, the recording system data storage modules start out empty and the complete set of recorded data is generally saved (not overwritten). As a result, the recorder storage is filled only once per mission. The total storage capacity generally has a direct impact on supportable mission durations. Storage capacities ranging from a few Terabytes to over 100 Terabytes are common in these systems, with aggregate recording bandwidths from a few hundred Mbytes/s to over 10 Gbytes/s.

In this mission model, the recording system is configured with freshly erased storage prior to the start of a mission. Upon the completion of the mission, recorded data is extracted from the platform, generally using one of two methods. The first is to offload data to some other storage device through the use of a high speed playback interface. After offload, the data storage modules can be erased and made ready for the next mission. The second method is to physically remove the data storage modules, replacing them with fresh storage modules. The removed modules are then physically transported to the analysis/post-processing/archive facility. Thus the platform is made immediately available for the next mission.

Other usage models include a continuous recording model, where data is recorded continuously for long durations. Older data is continuously overwritten with newer data, and only "data regions of interest" are retained for further processing or analysis. This model can result in recorder storage being overwritten many times during the course of the mission, which can place additional requirements on the specific types of SSD storage that can be used.

Hybrid usage models may also combine the above two models, as well as partitioning parts of the storage system for use in non-recording applications. For example, the storage sub-system may use 50 percent of its total available storage for the mission oriented write-once model, another 25 percent for continuous "region

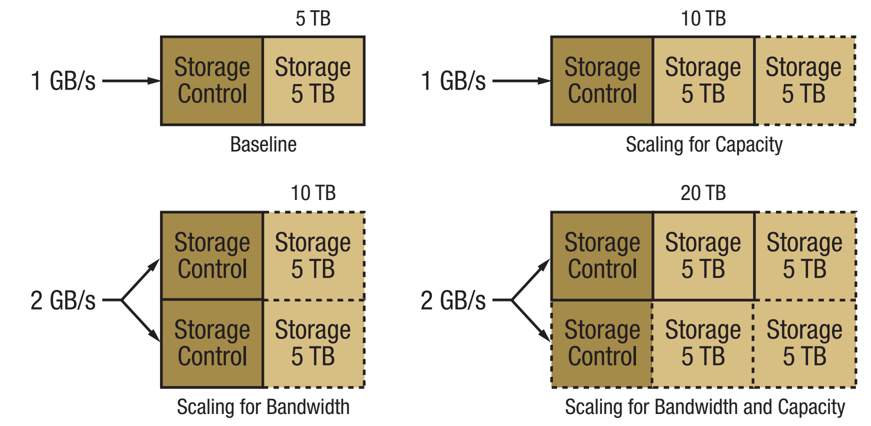

Figure 2

A building block approach allows bandwidth and capacity to be scaled independently.

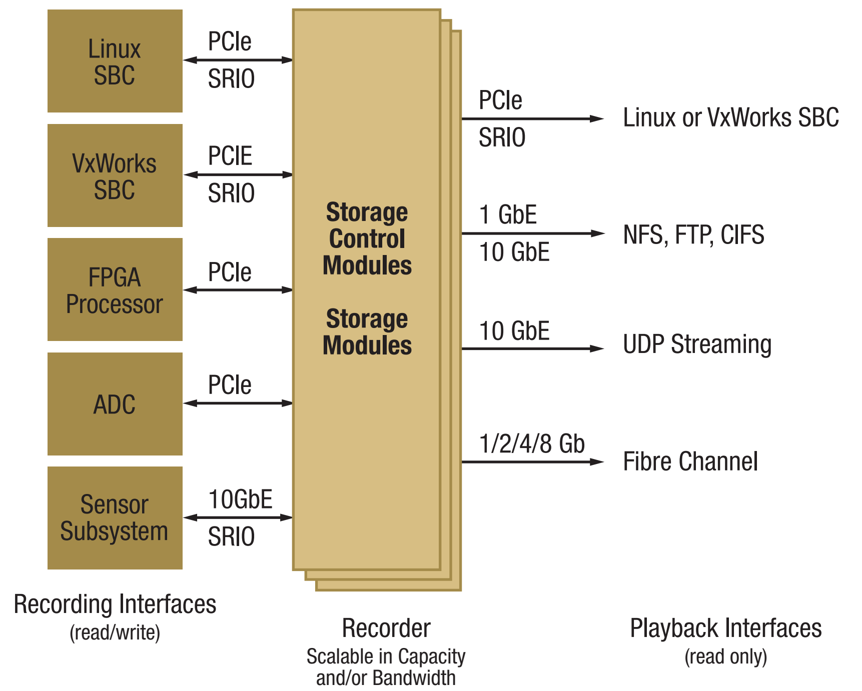

Figure 3

Options for the physical interfaces and interface protocols of the sensor recorder subsystems.

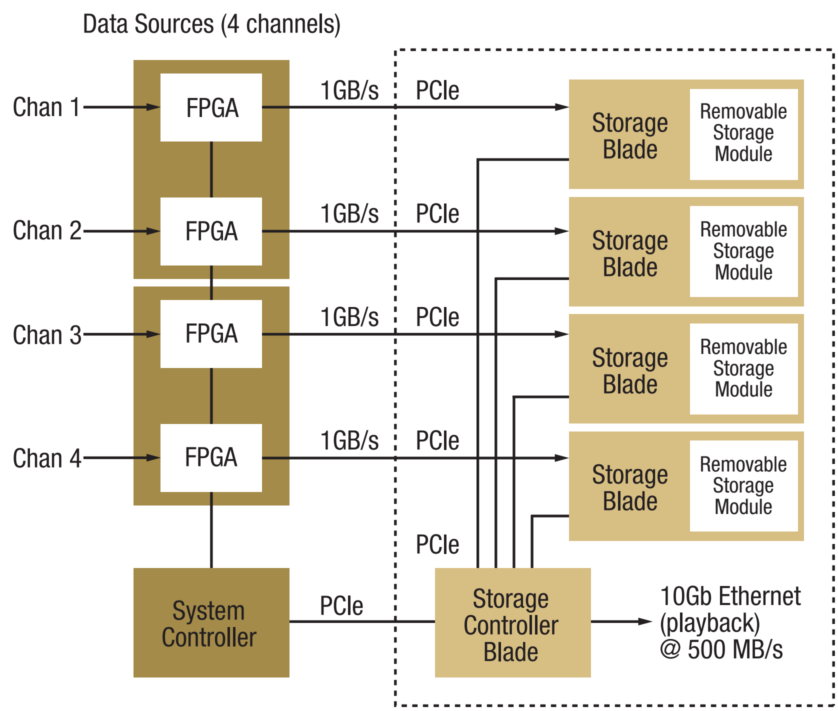

Figure 4

This real-world example application uses a single storage control blade along with multiple storage blades to form a four channel data recorder that can record data at a sustained rate of 4.0 Gbytes/s.

of interest" recording, and the remainder for general purpose network-based file sharing. That general purpose data is not typically removed or erased.

Choices in Recorder Modes

As illustrated in Figure 1, there are several data recording system modes that can be used, each with varying degrees of performance and scalability. The first two, NAS and DAS modes, require the involvement of a host processor board, which can limit performance and scalability. The third mode, Recorder mode, does not require that data pass through an intervening host processor board; data may be captured directly from the raw data source.

Network Attached Storage (NAS) Mode: In NAS mode, the recorder resource is configured to act like a network file server. A host processor board writes data to the storage system using NFS, FTP (for Unix, Linux file sharing) and CIFS (for Windows file sharing). Playback of recorded data is typically via the same interface as recording. Performance is lower when using NAS mode as compared to DAS or Recorder modes—typically about 200 Mbytes/s maximum per NAS recorder.

Direct Attached Storage (DAS) Mode: In DAS modes, raw storage is aggregated and presented to a host processor board as one or more large disk drives (actually, RAID 0/5 arrays). DAS host use a PCIe or Fibre Channel (FC) connection to the host CPU board. The host CPU fully controls the allocation and use of the 'virtual' storage blocks presented by the DAS recorder resource, generally through the use of a general purpose file system on the host. Playback of recorded data is typically via the same interface as recording. Typical maximum performance for DAS mode is moderate—about 500 Mbytes/s.

Recorder Mode: In Recorder mode, one or more Storage Controllers implement a high performance recording file system that fully manages the raw storage resources. It can accept a stream of data directly from either a "dumb" data source, such as an ADC or FPGA; a "smart" data source, such as a CPU or DSP board; or from a network data source like an Ethernet or UDP/TCP stream. In all cases, data is transferred from the data source to the recorder using highly efficient DMAs of data blocks directly to recording storage system. The recording file system may allow striping of data from a single data source across multiple storage blades for increased capacity and performance. Playback of recorded data may be via a PCIe/SRIO connection or via Ethernet (NFS, FTP, CIFS, etc.) Typical recording performance is about 1 Gbyte/s per blade, with possible aggregation of multiple blades for higher rates.

Scalability: Multiple Dimensions

There are several dimensions of scalability in sensor recording systems. First, there is scaling in the number of data channels. There is also scaling in the bandwidth that can be handled for each channel. And finally, there is scaling in the total storage capacity per channel. Note that scaling for increased bandwidth per channel or number of data channels often also results in increased capacity as well. Figure 2 illustrates scaling for bandwidth and capacity in a blade based recording architecture.

Capacity Scalability: This is the most basic form of recording system scaling. It simply involves adding more storage—storage blades—associated with an existing storage controller blade.

Single Channel Bandwidth Scalability: Each storage entity has a limited bandwidth capability. "Entity" here means either storage controller blade or storage blade or both. The only way to attain higher bandwidths is to "stripe" data from a signal channel across multiple storage entities, while still retaining a common "single point" recording and data playback model. That is, the multiple storage entities are combined for increase bandwidth, but to the user they still appear as a single entity.

Multi-Channel Scalability: A multichannel recorder supplies multiple data storage entities, each dealing with its own

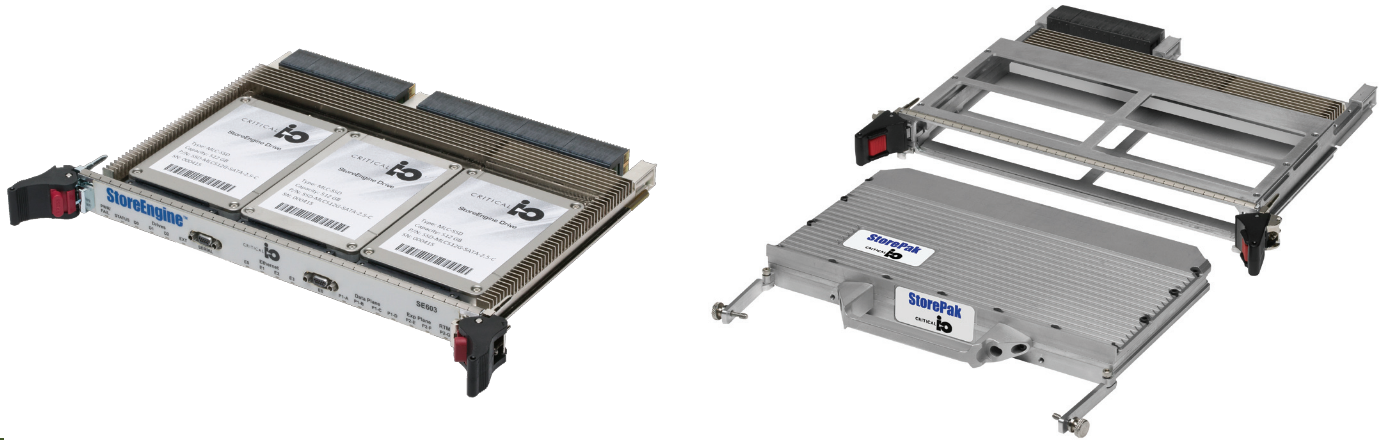

Figure 5

The StoreEngine (left) is a VPX blade that can also host up to 3 Terabytes of non-removable onboard SSD storage. StorePak (right) is a Storage VPX blade that can host up to 6 Terabytes of hot swappable SSD storage.

streams of data. It is still highly desirable to retain a common single point control and playback paradigm. A system can scale from one channel to three channels by adding only storage blades. The three-channel system in this case still only requires a single storage controller blade.

Interfaces and Protocols

Sensor data recorders must also support a diverse set of record and playback interfaces. Recorders are somewhat unique in that the record interface used to capture data from the sensor is not necessarily the same as the playback interface used to transfer the recorded data to some consumer.

As shown in Figure 3, there's a variety of options for the physical interfaces and interface protocols of the sensor recorder sub-systems. These are a key element of system flexibility, as different sensor systems have a diverse set of "preferred" interfaces. And in addition, for airborne systems in particular, it is highly desirable to be able to rapidly extract recorded data after a mission. The interfaces, and especially the protocols that are most effective to use for data playback, may be different than those that are most effectively used for data recording during the mission.

There are a number of record and playback data interfaces from which to choose. PCI Express and Serial RapidIO both provide multi-lane high speed serial. Gen 1 of both technologies runs at 2.5 Gbps per lane, while Gen 2 runs at 5.0 Gbps per lane (optionally 6.25 for SRIO). Both can scale to configurations of up to 32 lanes wide for highest performance applications. A typical G1 x8 or G2 x4 configuration supports data rates of up to 2 Gbytes/s. While these two technologies are significantly different in terms of addressing and routing, they can be viewed as roughly equivalent with respect to sensor recording interface usage.

Ethernet Simplicity

Ethernet, meanwhile, is seeing increasing usage as a sensor interface. Common protocols include "vanilla" TCP or UDP, optimized stream oriented implementations of UDP, as well as specialized protocols such as GigE Vision(1) that layer on top of UDP. Recording data rates of up to 120 Mbytes/s to 1200 Mbytes/s can be achieved using 1 GbE and 10 GbE respectively. Ports can be aggregated for higher recording data rates.

Ethernet is probably the most commonly used playback interface, especially where full-speed playback is not required. Depending on performance requirements, normal network file sharing protocols such as NFS, FTP or CIFS can be used, or for highest performance "vanilla" TCP or UDP, optimized stream oriented implementations of UDP can be used. File sharing protocols typically provide playback performance of several hundred MB/s, while optimized UDP stream protocols can easily support playback at full 10 GbE line rate of 1200 MB/s.

Fibre Channel (1/2/4/8 Gb FC) is a multi-gigabit storage networking technology that is commonly used as the standard for direct attached storage applications. Not generally used as a recording interface, but sometimes used for playback. Playback rates of up to 8 Gbytes/s can be achieved with 8Gb FC.

Example Storage Blade Application

A real-world example application that uses a single storage control blade along with multiple storage blades is shown in Figure 4. It forms a four-channel data recorder that can record data at a sustained rate of 4.0 Gbytes/s. The system provides a total storage capacity of 12 Terabytes, expandable to 24 Terabytes by simply adding more storage blades. This system usage generally follows the "mission-oriented" usage model, where the storage is filled during the course of a mission, and then removed and replaced with fresh storage prior to the next mission. During the course of the mission, however, certain sections of recorded days must be played back for "on-the-fly" analysis. This capability is implemented using a 10GbE interface utilizing a UDP stream protocol for efficient data access at greater than 500 Mbytes/s.

Two examples of storage building blocks along those lines are Critical I/O's

StoreEngine and StorePak, shown in Figure 5. StoreEngine is an ultra-high performance storage controller VPX blade that can also host up to 3 Terabytes of non-removable on-board SSD storage. StorePak is a Storage VPX blade that can host up to 6 Terabytes of easily removable and hot swappable SSD storage. Various combinations of multiple StoreEngines and multiple StorePaks can be interconnected using PCIe backplane connections to support a huge variety of easily scalable sensor data recording systems.

Leveraging Commercial Developments

StoreEngine and StorePak leverage best-of-breed commercial storage technologies, and build on these technologies to adapt them to the needs of military recording systems. In addition to clear cost advantages, this approach allows system designers to take advantage of the latest commercial developments while still meeting SWaP, performance and environmental requirements of their systems. StoreEngine and StorePak support a wide variety of recording and playback interfaces, including PCIe, 10GbE, and Fibre Channel, and supports recording in NAS, DAS and Recorder modes.

In Recorder mode, StoreEngine completely manages its storage resources using a recording file system that runs on the StoreEngine. This file system allows striping data from a single data stream across multiple StoreEngines for increased capacity and performance as well as aggregating the additional removable storage provided by multiple StorePaks. Time stamps are added to each data block, and a provision for user extensible meta-data is provided.

Modular, scalable, blade-based recording systems provide a mechanism to help meet the ever-increasing storage capacity and bandwidth needs of high resolution sensor systems. Critical I/O's StoreEngine and StorePak VPX blades are examples of blade-based storage building blocks that leverage the latest in COTS commercial and enterprise storage technologies to meet these rapidly increasing requirements.

Critical I/O
Irvine, CA.
(949) 553-2200.
[www.criticalio.com].