Omneon Spectrum Media Server Architecture

Whitepaper

OMNEON
VIDEONETWORKS
www.omneon.com

Omneon Spectrum Media Server Architecture

Don Craig, Chief Technology Officer

Paul Turner, VP Product Marketing

Omneon Video Networks, Inc.

May 2004

Table Of Contents:

Background

Omneon Video Networks develops and markets an advanced media server system that consists of specialized hardware and software components connected by various networking technologies. Omneon deploys a building block approach in the design of its servers. This approach ensures that a system will exactly match a customers needs in terms of real-time channel count, media format types, IP connectivity, storage capacity, overall bandwidth and redundancy. Customers can purchase only what they need initially, confident that they can add capacity and formats as their needs grow. This building block approach is not only applied to the physical hardware, but also encompasses the system software. This paper describes the individual components and how they relate to one another.

System Hardware Components

An Omneon Spectrum media server is comprised of 3 main hardware components: MediaPorts/ MultiPorts, which are the physical media I/O devices; MediaServers, which manage access to the hard drives, along with providing IP based connectivity; and MediaStores, which are the disk subsystems onto which data is stored. (See *Figure 1* below.) The most basic Omneon server is comprised of one of each of these, but the system can be expanded up to dozens of channels, and many Terabytes of storage, simply by adding more "blocks" to the system.

Interconnection between the MediaPorts/MultiPorts and the MediaServers is via an industry standard High Speed Serial Bus, and interconnection between the MediaServers and the MediaStores is via Multiple 2Gb/s Fibre Channel loops.

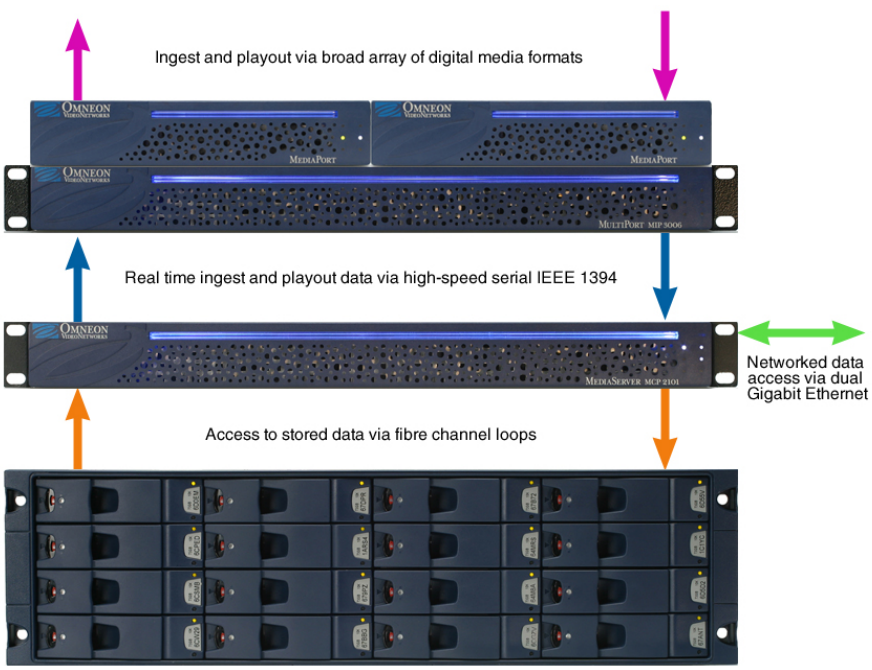

Figure 1: The basic building blocks of the Omneon Spectrum media server system showing the MediaStore storage array (bottom), the MediaServer system unit (middle), and MultiPort and MediaPort I/O components. Each element of the system can be scaled independently for additional channels, storage capacity, bandwidth or redundancy as needed.

MediaPort I/O Components

MediaPorts are the physical I/O devices used to input or output media to an Omneon server. Each MediaPort is responsible for a single channel, and processes video, audio and timecode. (See *Figure 2* below.) They are all bi-directional, offering record or playback capability (though not simultaneously – a MediaPort is either in record mode, or play mode at any one point in time). This is an important architectural point – since MediaPorts are bi-directional, they can be used to record material, then reconfigured via software to play that same material out, which minimizes the number of MediaPorts needed for any application. There are multiple variants of MediaPorts, three of which deal with input and output of baseband (serial digital) video (including VITC), AES audio (if required), and LTC. All 3 take in this baseband media and compress it into either DV or MPEG2 files, which are then sent to the MediaServer for storage. The difference in these MediaPorts merely defines the kind of MPEG2 files encoded – Iframe only, Iframe and Long GOP, or Iframe/Long GOP/IMX. All 3 can play back all DV and MPEG2 types, so the decision on which a customer should use is purely dependent on the record format.

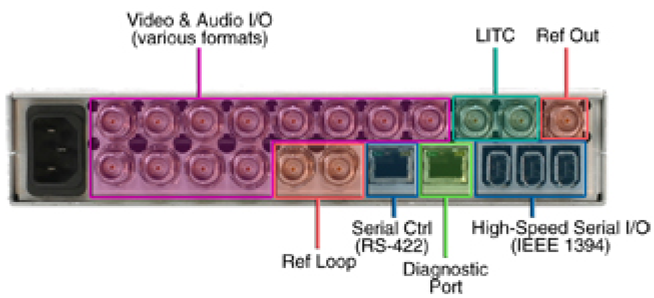

Figure 2: Rear panel view of a MediaPort I/O component.

Another MediaPort variant is also capable of inputting serial digital video, but passes this down on to the MediaServer uncompressed, yielding the highest possible picture quality. This MediaPort is also capable of inputting video over SDTI (Serial Data Transport Interface). Currently, this is the method used to input DV or DVCPRO material at faster than real time, to input High Definition material compressed in HDCAM format, or to input High Definition material compressed in DVCPRO HD format.

A final MediaPort variant inputs material via DVB/ASI. This video format is used for transportation of customized MPEG video streams (i.e. the video has already been compressed before input to the MediaPort). Most common uses for ASI are as the physical transport mechanism for MPTS (Multi-Program Transport Stream), where 3 or 4 video signals are compressed and multiplexed onto a single BNC cable or for transport of High Definition material which has been MPEG2 compressed in an external encoder.

The use of these last two variants of MediaPorts enables Omneon servers to store HD material in any of the available industry standard formats. The building block approach to hardware (and system software) means that these formats can be simultaneously stored on the same server (in the same file system), and these High Definition signals can coexist with standard definition video. This is a very important differentiating feature of Omneon servers, allowing users to mix and match compression techniques, or even change formats without the need to replace the entire server – you simply add the format you need, when you need it. No other server offers this level of flexibility.

Omneon media servers find significant use in playout applications. These applications have their own specific needs – in general, they require significantly more outputs than inputs, and the high channel counts force constraints both in terms of cost/channel and rack space/channel. For these applications, Omneon offer another type of I/O device, called a MultiPort.

MultiPort I/O Components

MultiPorts are I/O devices which offer significantly improved cost/channel and rack space/channel by limiting the ports to output only and putting many outputs into a single rack unit enclosure. (See *Figure 3* below.) They are offered in 3 channel and 6 channel variants and operate on MPEG based material only (the vast majority of playout servers are MPEG based). By coupling MultiPorts for playout with a small number of MediaPorts for ingest, customers can again design systems that are tailored to their exact needs, while ensuring expandability at any time.

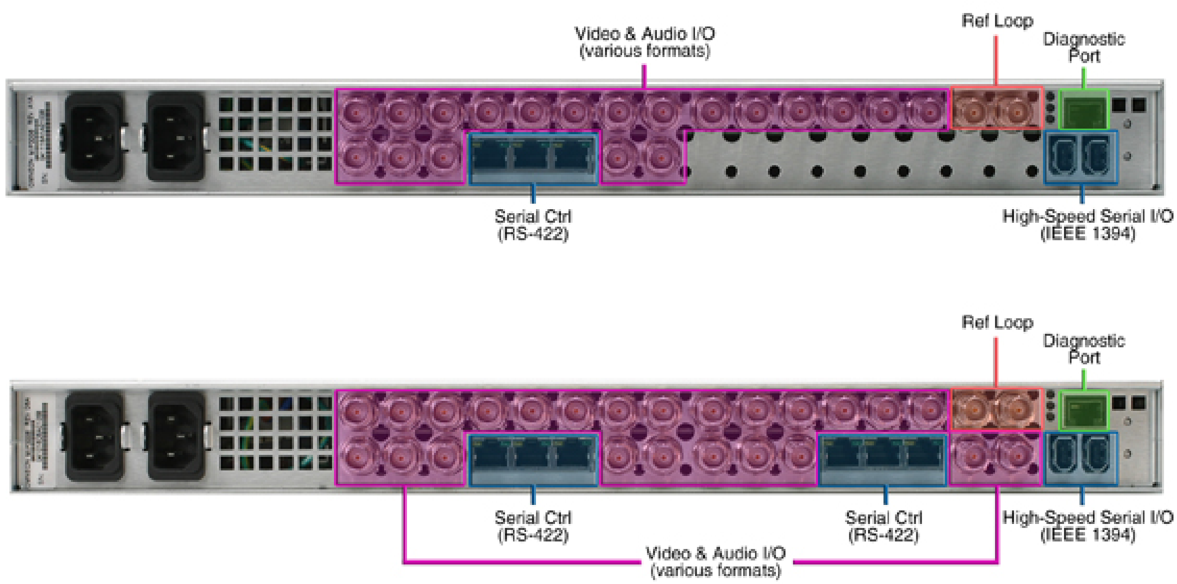

Figure 3: Rear panel views of the MultiPort MIP 3003 (top) and MultiPort MIP 3006 (bottom).

All MediaPorts and MultiPorts transfer their data to and from the MediaServer component via an industry-standard high speed serial interface. The use of this external interconnect enables the expandability of the system, while simultaneously maximizing the serviceability of the system. The interconnect offers "hot plug" capability, which means that MediaPorts/MultiPorts can be removed, replaced or added to the system without the need to take the system off-line. This is especially important in the serviceability issue: Most customers with mission-critical applications will run in an N+1 or mirrored configuration. This configuration ensures that in the event of a failure of an I/O device, a spare channel is available to take on the load and ensure uninterrupted playout. Sooner or later, though, that failed I/O must be replaced. The use of this high speed serial interface in an Omneon server means that this can be done while the server remains on air. All competing products require that the server be taken off-line in order to replace an I/O, requiring the use of back up equipment. This "hot plug" capability also means that additional I/Os may be added to the system without requiring that the system be taken off-line for the upgrade.

MediaServer System Units

MediaServers are one of the core components of an Omneon server. (See Figure 4 below.) Each MediaServer is a 1RU chassis which interfaces to the MediaPorts/MultiPorts via the High Speed Serial Interface (IEEE 1394). This interconnect is used to configure the operation of the I/O devices, and to receive media packets from them when in they are in record mode and transmit packets to them when they are in playback mode. The high speed interface offers isochronous (deterministic) transfer of data, guaranteeing video delivery.

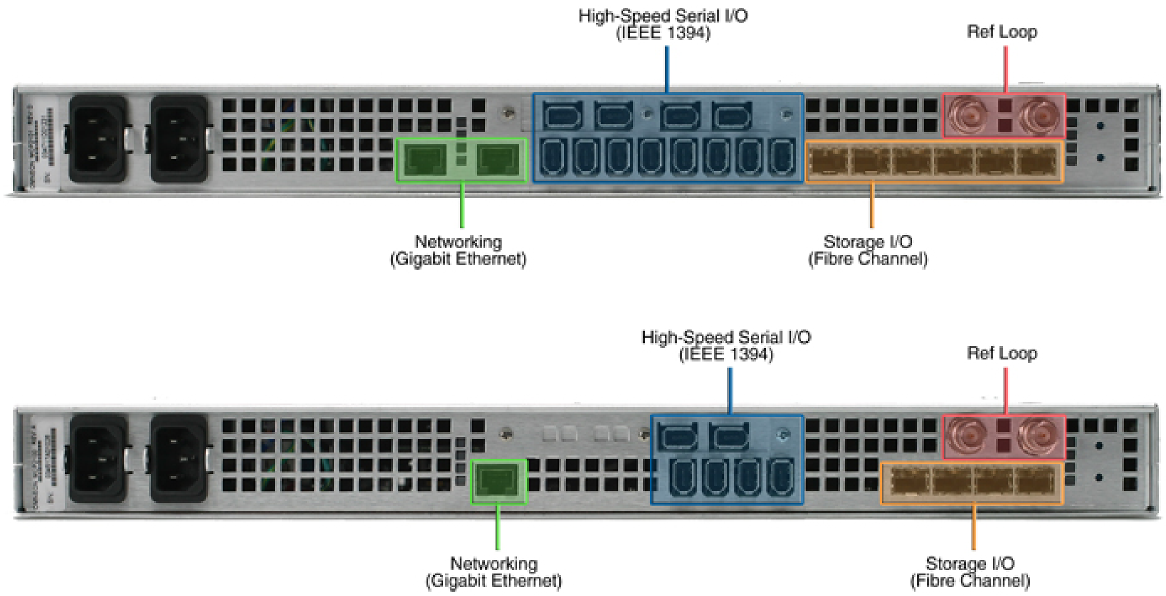

Figure 4: Rear panel views of the MediaServer MCP 2101 (top) and MediaServer MCP 2100 (bottom).

Today's playout architectures involve significant network connectivity between servers (and between servers and external clients such as Non-Linear Editors). As a second function, MediaStores offer IP connectivity to the network via built in Gigabit Ethernet ports. The MediaServer manages IP bandwidth to ensure that the real-time record and playback requirements of a video server are never compromised by high volume network traffic (such as when a complete clip is being transferred over the network to the server. In such cases where significant real-time operation is required, the MediaServer will reduce the transfer rate over the network, to ensure that real time operation is always maintained.

Unique to the Omneon MediaServer is the ability to mount the entire server as a shared drive via Samba (for Windows) or Apple File Protocol (for Mac). This significantly improves workflows which involve external computing platforms, allowing them to browse content via standard explorer or Finder functionality. Competing products offer the user FTP connectivity, but none of the sophisticated "drag and drop" capabilities offered by an Omneon server (of course, Omneon servers also offer FTP transfers as another choice).

A third function of a MediaServer is to act as the RAID controller for the MediaStores. Omneon servers use MediaStores which are configured as JBOD (Just a Bunch of Disks) or SBOD (Switched Bunch of Disks). The difference between these two configurations is contained fully in the disk enclosure itself. Obviously, for playout, the media needs to be RAID protected. The MediaServers "stripe" the information onto the drives in a RAID3 configuration. This is done entirely in software (rather than in hardware), as the flexibility of a software RAID system is inherently greater than that of a hardware solution (more on that in the discussion about RAID control later in this paper).

The final function of the MediaServer is to act as the File System controller for the MediaStores. The Omneon file system is highly optimized for use in a video server, combining optimizing disk usage while simultaneously maximizing available storage capacity. The Omneon file system is also natively SAN capable. This means that multiple MediaServers can share storage in a SAN based manner, without the need for external file system controllers. The file system ensures that all MediaServers have synchronized views into the storage at all times. This synchronization is maintained even if a MediaServer has to be replaced – file system information is stored onto the disk subsystem along with the actual media, and uploaded into the replacement MediaServer upon startup.

This synchronization becomes more and more important as Non-Linear Editors (NLEs) and other clients transfer data to and from the playout server. Many of the files used by an NLE (such as EDLs) are significantly smaller than the media files themselves.

MediaStore Storage Arrays

A MediaStore is a 3U or 1U full-width rack-mount chassis with a pair of fibre channel interfaces containing JBOD or SBOD disk arrays. All active sub-assemblies are hot swappable, and each chassis contains up to 16 (3U chassis) or 4 (1U chassis) dual port high performance fibre channel disk drives. (See Figure 5 below.)

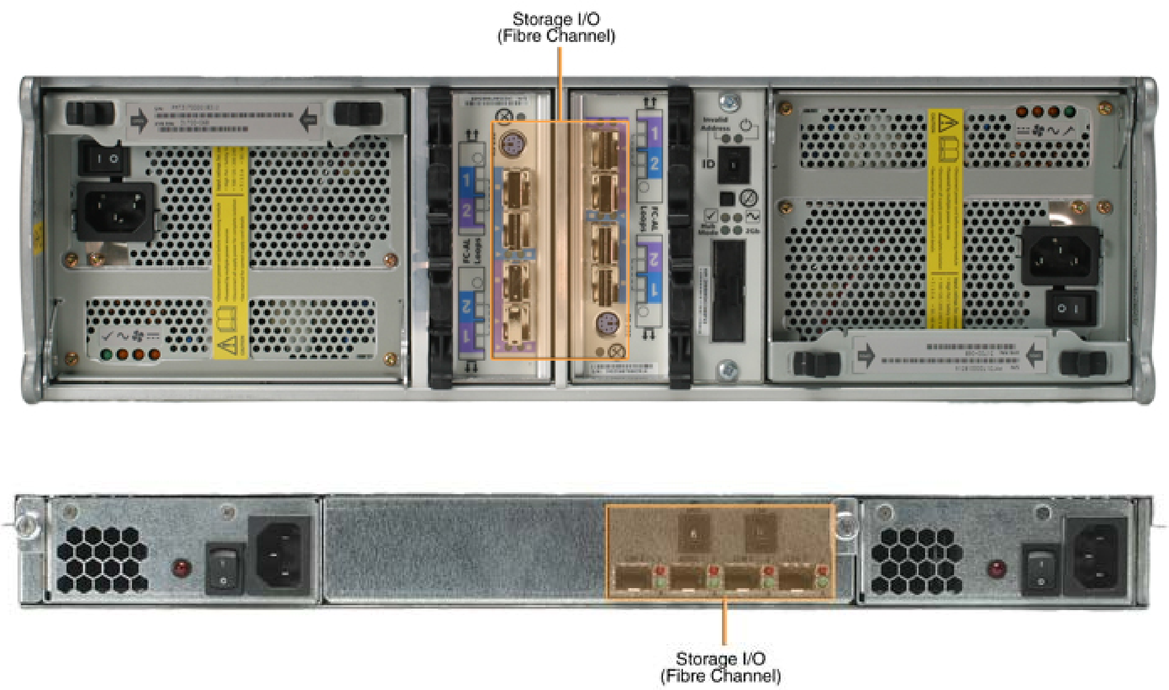

Figure 5: MediaStore MSC 4000 (top) and MediaStore MSC 4400 (bottom) rear panel views.

MediaStore disks accessed through the fibre channel ports are organized into Raidsets (Groups of 4, 7 or 8 disks), which are in turn organized into the filesystem. MediaServers write to the disks a Raidset at a time, and an important property of the system is ensuring Raidsets are wide enough (have enough disks), and that the filesystem is wide enough (have enough Raidsets) to support the bandwidth requirements of the total number of MediaPorts and Ethernet interfaces connected to the system. A single MediaServer supports up to 16 MediaPorts operating at about 30 Mb/s or less, but at higher data rates a single MediaServer will be limited (for example, a single

MediaServer will be limited to 4 simultaneous streams of uncompressed 270 Mb/s SDI each with an associated 4 tracks of 1.8 Mb/s AES/EBU stereo audio).

One additional system component which deserves mention is the Omneon Networked System Manager. This is a self contained control console, based on an x86 platform. It is used for system configuration and diagnostics, and offers a web-based GUI, which can be accessed either directly on the platform itself, or remotely over the network (subject to network security rules, of course). Although an important engineering tool, it is not needed for system operation. Provision is made for a backup System Manager to provide redundancy if so required.

System Software Components

QuickTime and MXF wrapper formats are both available on Omneon media servers. Omneon servers are unique in the methodology of storing the “clips” on the MediaStore. A “wrapper” is used to encapsulate the actual media, for a variety of reasons. At present, Omneon servers use a QuickTime wrapper to encapsulate the media. To be specific, a “Reference” QuickTime wrapper is used. In this case, the media is not physically part of the QuickTime “movie”, but is pointed to by the .mov file. This offers great advantages in a media server, as now, audio and video can be edited separately (see *Figure 6* below), or additional audio tracks can by added simply by importing the new .wav file into the system and modifying the clip’s QuickTime wrapper to point to the new audio in addition to the existing audio. Other tracks and file types can also be accommodated, as technology advances, without the need for a major change in the operation of the system. This approach is extended to MXF wrapped material also: MXF (as a wrapper format) is actually very close to QuickTime, which has meant that the transition to MXF support has been extremely simple. In fact, both MXF-wrapped material and QuickTime-wrapped material can coexist on the same file system, and MXF and QuickTime clips can be played “back to back” without the need to reconfigure, or any black or corrupted frames – again offering the customer the freedom to expand or change the system without major disruption.

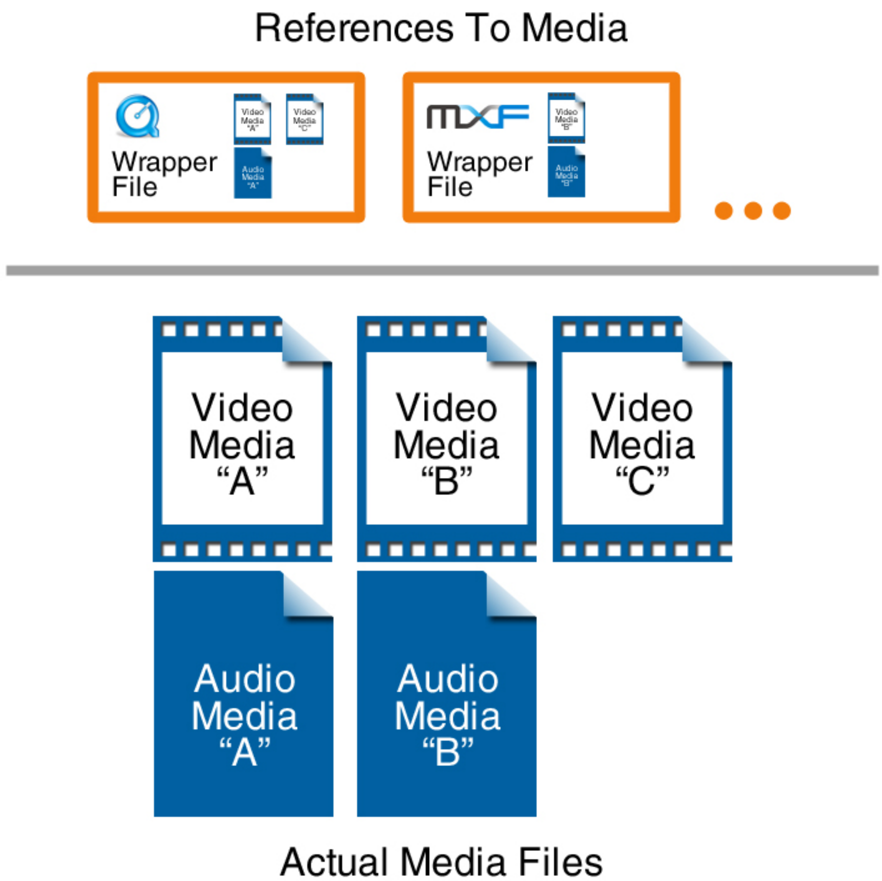

Figure 6: In the Omneon file system, media files of different formats are stored independently from the wrapped references to the media. Wrapped references are exposed to automation and editing systems. The complexities of multiple file formats and separate or linked audio are eliminated for the user.

SAN Properties

In addition to the usual SAN functionality, the Omneon file system offers intelligent striping of the data. As mentioned earlier, protection of storage is critical to the system's reliability. Manufacturers have many options when choosing a RAID scheme for their systems, and each solution offers its own unique benefits while also having a direct impact on the amount of storage available for media. In most cases, manufacturers tend to choose RAID3 or RAID5. Both technologies employ the idea of data drives plus parity drives for redundancy purposes. The difference between these two schemes is simply whether there is a dedicated drive for parity, or parity is striped across the disks along with the data.

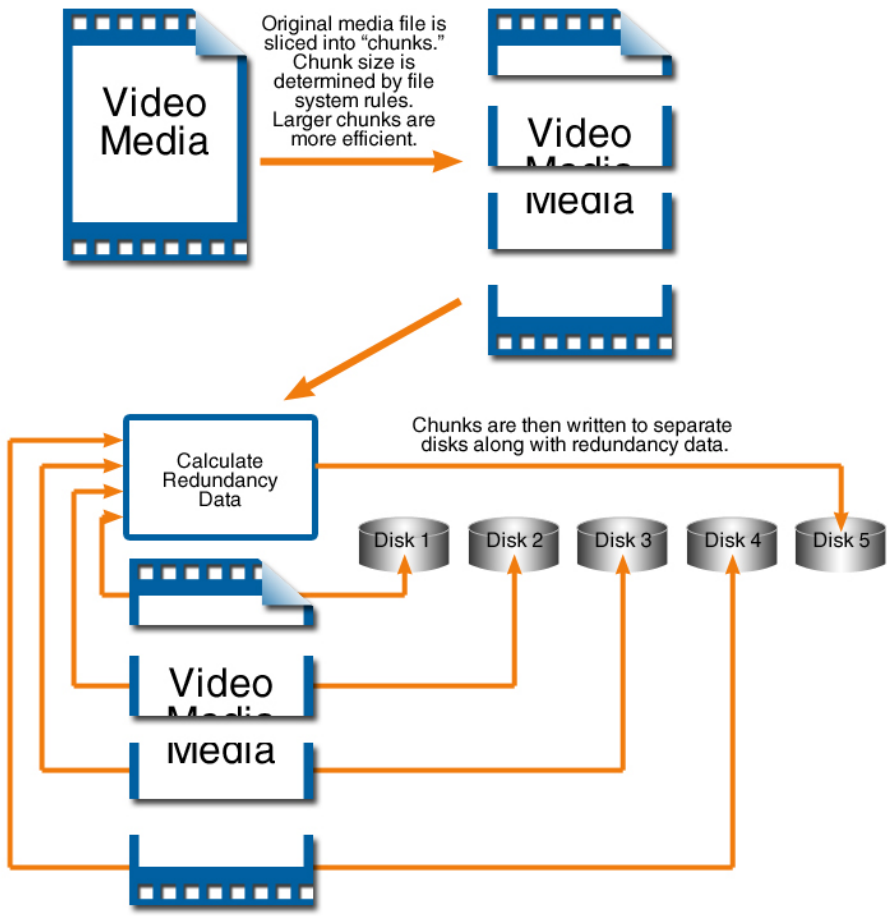

Figure 7: Typical RAID 3 "workflow."

In recognizing that the Broadcast industry is moving away from a tape based infrastructure to file based storage and management, Omneon recognized that simply using a fixed size RAID stripe for storage would not be the most intelligent and cost effective solution, and designed an advanced file system to manage the storage, dynamically matching file sizes with RAID topologies. This maximizes both performance and storage capacity, regardless of physical file size. An example may make this clearer. Most playout servers use a RAID stripe that is configured for storage of large files. The general rule is to make the smallest "chunk" that can be written to disk as large as possible. By doing so, you can read more data per read operation, simplifying buffering issues further down the line. This is reasonable, as media files are generally very large. A consequence of this decision is that small files can take up substantially more storage space than what the size of the file would indicate. The crux of the issue here is that given a large "chunk" size, a small file is "striped" across many disks, with only a fraction of each chunk being filled with real data. The rest remains unused, and the file takes up much more space than is absolutely necessary. This issue is illustrated in fig. 1. The chunk size is set by the file system for the most efficient storage of large files, and is regularly 64KB or larger. If we assume a chunk size of 64KB, and a small file of (say) 64KB, then in the RAID system shown, the original file would be split up into 4 chunks (i.e. 16KB of data), plus another 16KB chunk of redundancy data. The system can only write in chunks of 64KB to each drive, so it "zero fills" each chunk of the file to make it 64KB, and then writes the chunks to each disk. So your 64KB file now takes up 320KB of storage!!! An Omneon file system dynamically changes the stripe methodology, in order to use a more appropriate striping system for each file size, thus maximizing return on investment. Intelligent use of the correct RAID technology is the only way of maximizing storage efficiency for varying file sizes. (By the way, there's absolutely no reason why you can't have more than one RAID striping configuration on an array – you just have to manage them correctly!). This issue is of increasing importance as servers start to embrace the idea of standard networking topologies, with the resulting increase in the number of smaller files stored on the server.

The Omneon SAN and its components have the following properties:

- Between 2 and 8 data disks are clustered together to create a Raidset:
    - Recording media streams are buffered until all disks in a Raidset can be written at once (writing a stripe). This approach maximizes throughput to the drives, ensuring writes are extremely efficient (an important factor in SAN based systems).
    - The system continues uninterrupted if any drive in a protected Raidset fails.
    - A replacement for a failed drive can be rebuilt at an operator controllable bandwidth; only allocated storage is rebuilt.
    - Raidsets can be spread across multiple disk chassis.
- A filesystem is made up of 1 or more Raidsets:
    - Filesystem writes can be striped across Raidsets, or the filesystem can be organized to fill a Raidset before passing on to the next, or both (X Raidsets Wide, Y Raidsets deep).
    - New Raidsets can be added to running filesystems:
        - Expansion without rebuild: If the expansion is simply adding capacity, the system will complete filling pre-existing Raidsets before continuing on to the next. There is no need to rebuild the array.
        - Dynamic resizing: If the number of Raidsets across which the filesystem is being striped is being increased, any resulting increase in filesystem bandwidth will only affect files written after the change. There is no need to rebuild the array.
        - Re-striping: Within the Omneon system is a utility that, using a specified amount of system bandwidth, will read every pre-existing file in a filesystem and re-write it. In this way the peak I/O channel count of an existing filesystem can be dynamically increased.
- All MediaServers share a common view of all the filesystems in a SAN, and synchronize writes.
- Each MediaServer may have 1 or 2 paths to each disk drive, and dynamically selects the path with the least delay.

- In the event of a path failure, the system automatically reconfigures to use only the functioning path.

NAS

As mentioned earlier, the server may be mounted as a shared resource ("Network drive") over the network. Networked Attached Storage (NAS) views of the filesystem on the SAN are available via its Ethernet ports. If additional IP bandwidth is required, multiple MediaServers can be connected to the network in order to increase NAS bandwidth.

Conclusion

This has been a brief description of some of the more important aspects in the design of the Omneon Spectrum media server system. In addition to its duties as a server for the playout market, the system functions extremely well in the Production market. Its ability to store multiple file types and multiple resolutions, coupled with its advanced architecture, which couples SAN and NAS technologies into a highly sophisticated file system puts it head and shoulders above the competition, both protecting customers' investments and offering unprecedented expansion, reliability and serviceability. It is truly a mixture of the best practices of the Video industry and the IT industry, bridging the gap between technologies often considered to be completely at odds with each other in a way that enhances the server workflow, rather than limiting it.