grass valley
A BELDEN BRAND

K2 Shared Storage Systems

Building Systems with Flexible Storage Technology

Adolfo Rodriguez, March 2016

APPLICATION NOTE

K2 Shared Storage Systems Strategy

By optimizing the best off-the-shelf IT products for Grass Valley K2 systems infrastructure, users benefit from capacity increases and performance gains as technology progresses. Grass Valley, a Belden Brand, handles these changes by making sure there are upgrade paths for all components. Investment is protected by designing new systems that can be used with multiple generations of technology.

The key element of any K2 shared storage system is bandwidth, and how much bandwidth needs to be supplied depends on the application.

K2 Media Server

K2 clients with record/play channels, editing and production tools are all consumers of bandwidth. Storage arrays are the suppliers of bandwidth. In between, there needs to be an arbiter which functions as data bridge and data manager between consumers and suppliers. This is the function of the K2 media server.

These servers can be configured to manage 100, 200, 300, 400, 600, or 800 MB/s of data, depending on the system requirements. A very large system might require more than 800 MB/s of bandwidth. *Unlike other systems, K2 can add additional media servers and aggregate the bandwidth.* The multiple servers connect on the client side through a 10 GigE switch and on the storage side through an 8 Gb Fibre Channel switch. *While other systems are limited to as few as two servers, K2 can scale up to 20-30 servers.*

File transfers also consume bandwidth in shared storage systems. It is critical to overall system performance to manage this type of bandwidth requirement as well. For smaller systems, the same K2 media server can supply up to 95 Mb/s of continuous bandwidth for this task at all times — regardless of whatever activities are occurring. For larger systems, additional file transfer bandwidth can be added.

FTP Servers

A K2 shared storage system can incorporate special file transfer (FTP) servers. Similar to the K2 media servers, these file transfer servers connect on the client side through a 10 GigE switch and on the storage side through an 8 Gb Fibre Channel switch. Each of these servers can manage up to 600 MB/s of file transfer bandwidth. Like the K2 media servers, multiple file transfer servers can be added to scale up to 10 GB/s or more of constant file transfer bandwidth. The file transfer bandwidth is independent of bandwidth needed for other operations, such as play and record, but unlike other systems, the specified file transfer amount for the system is always fully available, no matter what load is placed on the system. Like all other types of bandwidth requirements, file transfer bandwidth must be supplied by a certain number of drives. K2 storage components can be precisely configured to match file transfer bandwidth requirements.

Managing Assets

Another task of the K2 media server is managing the assets stored on the system. The media server hosts the K2 file system and database. A K2 shared storage system can have up to 50,000 clips with maximized access performance. The system can manage up to 200,000 clips with only slightly slower access times. For every asset imported or recorded, a K2 system generates a unique ID, so even clips with the same name are not an issue. What gets written to storage is an accessible container file containing elementary streams comprising video, audio, timecode, ancillary data, and XML-based metadata. K2 can unwrap and wrap these container files using standard formats such MXF and MOV. For example, a QuickTime file could be imported into K2, used for some operations, and then exported out again as MXF.

The file system keeps track of the changes that will be made in a journal before committing them to the main file system. In the event of a system failure, the file system is quicker to bring back online and less likely to become corrupted. Metadata and journaling for the file system is contained in three drives in the first LUN of the primary RAID storage chassis. Two drives are mirrored and the third one is used as a hot spare. For high-bandwidth systems, solid-state drives can be used. The rest of the RAID storage chassis will have media drives for actual storage.

K2 Storage Components

The storage implementation for K2 systems is not only flexible, but easily scalable as well. The principal storage component is the primary RAID chassis. This chassis contains a highly reliable, fast failover hardware RAID controller and can be configured with dual controllers for redundancy.

There are various permutations of primary RAID chassis, depending on what kind of system is being deployed. One variant is used for any non-redundant K2 system. A second variant is used for smaller redundant systems with only a single primary chassis/controller. A third variant is for larger redundant systems with multiple primary chassis/controllers.

It is this capability of having multiple primary chassis/controllers that distinguishes K2 shared storage systems from other vendors' offerings. To scale for capacity and bandwidth, multiple primary chassis/controllers can be linked using an 8 Gb/s Fibre Channel switch. Behind each primary chassis will be multiple expansion chassis. *In this way, all the capacity and bandwidth can be aggregated into one unified pool of storage that can scale much greater than competitive systems.* Capacity can go up to 6.8 PB of raw storage using 6 TB production drives and up to 2.1 PB of raw storage using 1.8 TB online drives. The bandwidth of a K2 system can scale up to 24.4 GB/s.

Redundancy

Systems can be configured to stripe the drives using RAID-5 or RAID-6. Storage drives are all of the 6 Gb/s SAS type. The latest generation of online storage uses 2.5 inch 600 GB, 1.2 TB and 1.8 TB 10,000 RPM drives, and production storage uses 3.5-inch 4 TB and 6 TB 7,200 RPM drives.

Storage configurations can span a wide performance range, depending on read vs. write, bit rate of material, access method, etc., but as a general rule of thumb, production storage will have 12 drives per chassis representing 150 MB/s of bandwidth. Online storage will have 24 drives per chassis, representing 600 MB/s of bandwidth.

K2 ships with higher-performance 8 Gb Fibre Channel-based RAID systems and servers. These systems were designed to be backwards-compatible with the previous two generations of storage. This application note documents how to expand systems with the newest storage offerings.

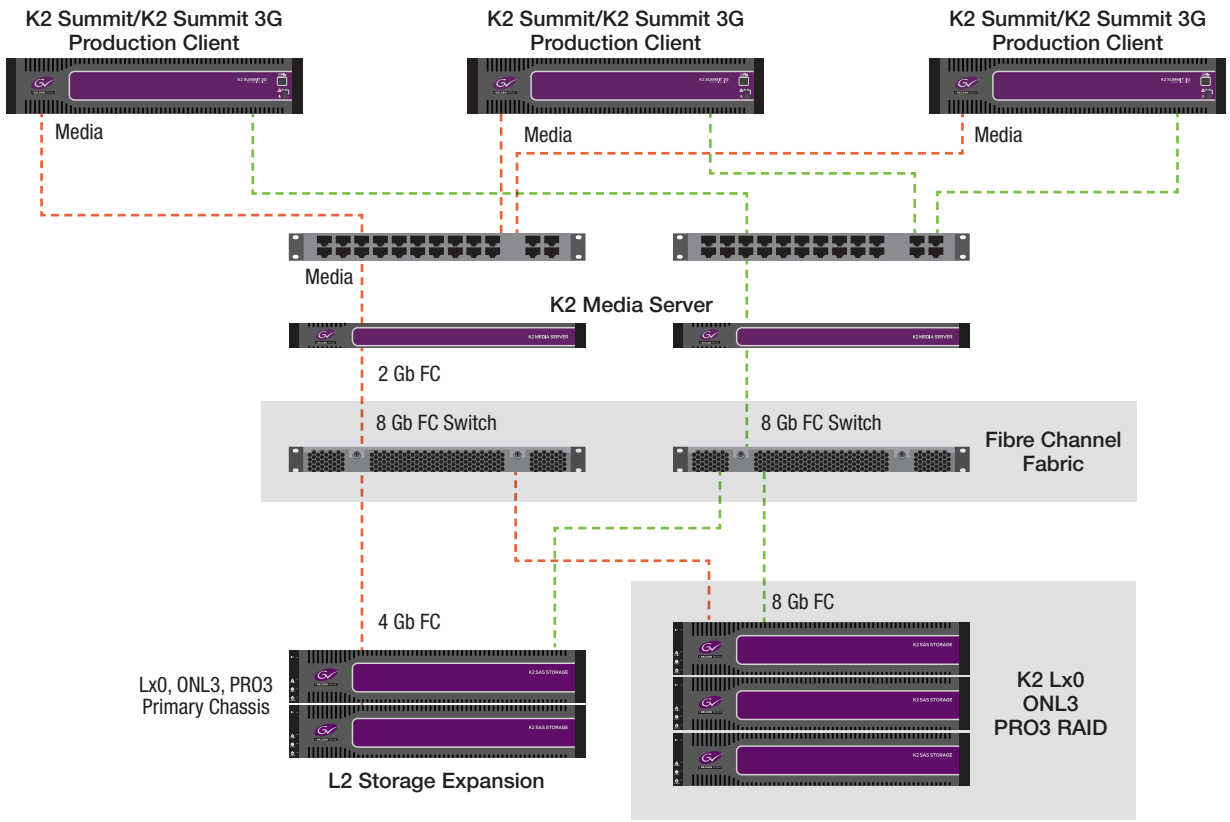

Figure 1 – The general concept for expanding K2 systems is to keep existing storage and to expand with additional RAID systems by adding a Fibre Channel fabric.

General Storage Expansion Concepts

In understanding storage expansion, there are some general concepts and terms that are important to understand. There are three reasons why storage might be added:

1. For additional capacity
2. For additional bandwidth
3. To replace a legacy storage system

For K2 systems, as they grow with additional channels, more bandwidth will probably be needed as well as capacity.

Terminology

LUN/RANK – A set of drives that are treated as one logical unit. The term RANK is also used in place of LUN in K2 documentation.

K2 ONL/ONL2/ONL3 and K2 PRO/PRO2/PRO3 systems have 6-drive LUNs for RAID-5 or RAID-6 configurations.

Stripe Group – This is a logical grouping of LUNs in which data is written across all drives. With a stripe group of 3 LUNs, data is written across all 3 LUNs on each write. Within the 6-drive LUN, data is also written across all drives.

Multiple Stripe Groups – K2 systems can have multiple stripe groups in which the system will write data across one or the other (but not both) stripe groups. This process is transparent to the user, as only one file system is seen. It is a low-level file system operation only.

Bandwidth – The amount of data that can be read from storage. The bigger the stripe group, the more bandwidth is available. Generally after 5-6 LUNs, additional LUNs offer minimal if any additional bandwidth (but will add capacity).

I/O Operations – In addition to bandwidth (how fast data can transfer), the number of I/O operations (data requests) the file system can handle is critical. For example, a simple request for data is only a few I/Os where as managing a database requires many I/Os. Large amounts of I/Os can reduce the overall bandwidth of a system.

Capacity – The total number of hours stored. Amount of hours is determined by the bit rate of the video file. The calculation is more complex when there is a mix of compression rates, or when both SD and HD files are on the same system.

Mixing Drive Sizes – When a drive size is no longer available from the manufacturer, the general strategy in the industry is to replace drives in a LUN with a larger drive as long as the technology is similar. The usable capacity of the new drive is the capacity of the smaller drive, so if a 300 GB drive is replaced with a 600 GB drive, only half of the drive is used.

Mixing drives in a system can be done with multiple stripe groups as seen in the Expansion by Capacity diagram (Figure 3).

Bandwidth-Constrained Systems – In such systems, the number of LUNs is determined by the bandwidth required by real-time video/audio I/O needs and non-real-time needs such as FTP and CIFS.

Capacity-Constrained Systems – In such systems, the number of LUNs is determined by the number of hours required. They will have excess bandwidth available for adding more channels or increasing the video bit rate. When the system is originally quoted, you will know if it is bandwidth or capacity constrained.

Storage Expansion

The file system can be expanded for more bandwidth, or more capacity, or both. It is generally advisable to expand for bandwidth to get the advantage of both capacity and bandwidth. These upgrades are non-destructive, which means that all existing content is preserved.

When adding a storage to a 4 Gb/s Lx0 system please consult with Grass Valley technical staff to check on issues around software versions, switches, and server types.

Expansion by Bandwidth with Restripe

This is the preferred method of expansion. This method adds more LUNs to an existing stripe group. Bandwidth is increased by adding additional LUNs and, in many cases, additional RAID controllers to the initial stripe group.

- Additional drives and servers are connected to the switch
- The system needs a quick reboot just to see the new storage components
- A background task will start a copy of all material on the system. This restripe task will read each file and then write it back to the file system. The new file is verified before the old file is deleted for complete security
- When the copy task is finished, the additional bandwidth will be available
- Time to complete the restripe is dependent on available bandwidth and number of clips so it's not possible to estimate the time — but it could be in the range of 24 hours or so for a large system with tens of thousands of clips
- Note: The system can remain operational while the restripe process is occurring, but the additional bandwidth (for new clients or codecs) cannot be used until that process finished

The new drives must be the same size or larger. However, if using larger drives only the amount of space on the smaller drives is used.

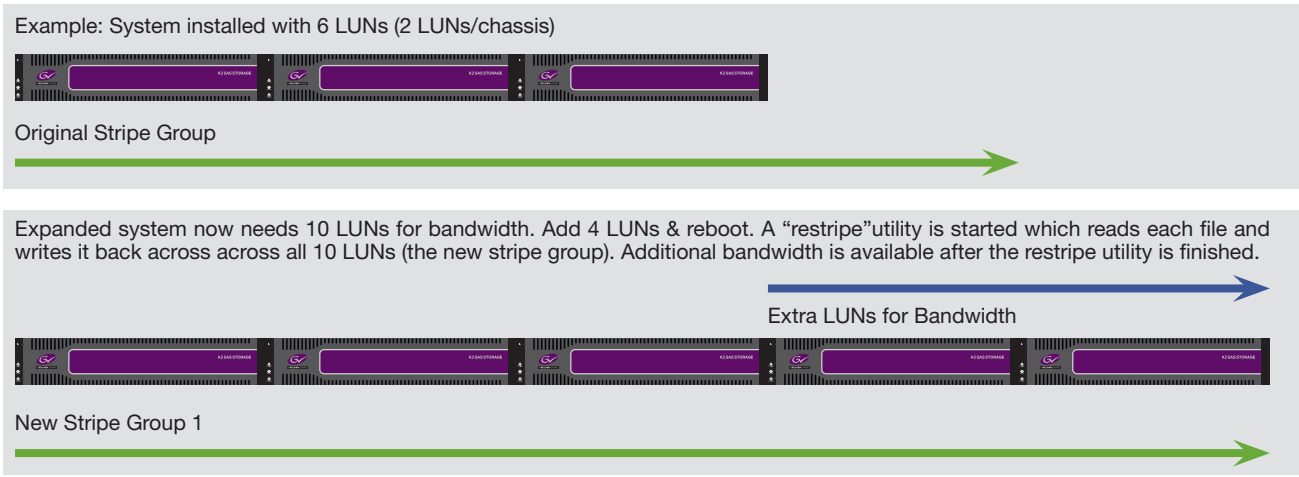

Figure 2 – Example of expansion by bandwidth with restripe. This upgrade is non-destructive and preserves all existing content.

Expansion by Bandwidth with No Restripe

In a system with small expansion (in the range of 20%) and a high churn rate (where new material is constantly being ingested), the re-stripe happens automatically as new content is recorded and the old content is deleted. In this case, it may be decided to not perform the restripe step.

For example, for a user who has less than 3,000 hours of content that is mostly replaced every eight days, there is no need to restripe the file system.

Expansion by Capacity

The file system can have multiple stripe groups, but this is transparent to the user, who will only ever see one file system.

One advantage of this approach is that it permits expansion with a larger capacity drive. Consult with a technical specialist before expanding by capacity.

Notes:

- The bandwidth of the system is the bandwidth of the smallest stripe group
- When adding a second stripe group, it needs to be the same number of LUNs as the first stripe group
- This upgrade is non-destructive and preserves existing content

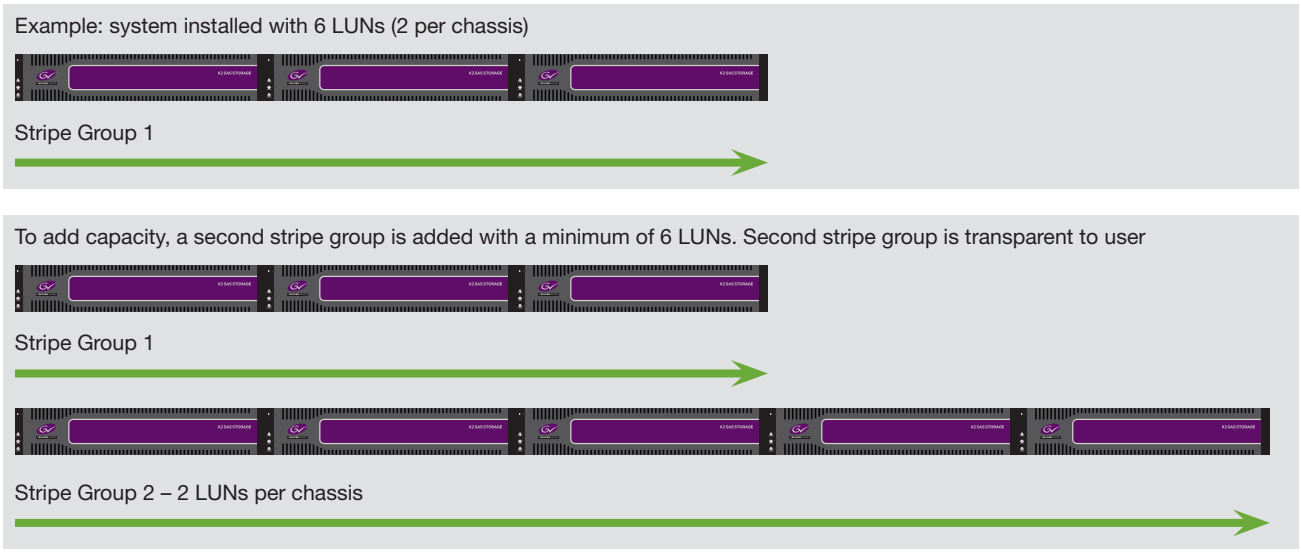

Figure 3 – Example of expansion by capacity. Adding a second stripe group is another expansion option. This will add additional storage but not additional bandwidth. One advantage of this approach is it permits mixing different size drives. This upgrade is non-destructive and preserves all existing content.

Fibre Channel Fabric Notes

K2 Lx0 systems are based on 4 Gb Fibre Channel RAID controllers, whereas K2 ONL/ONL2/ONL3, PRO/PRO2/PRO3 and NL2/NL3 systems use 8 Gb RAID controllers. When expanding systems, the Fibre Channel fabric will operate at 4 Gb. An existing 4 Gb Fibre Channel switch with open ports can be used. If adding a 8 Gb Fibre Channel switch, it will operate in 4 Gb mode.

Storage Calculations

With a single compression format, storage calculations are straightforward and based on the tables below.

When mixing compression formats (for instance SD and HD), the GVConfig Tool can used for estimation.

Bandwidth Calculations

The slowest RAID controller and slowest LUN determine the performance of the file system. For example, if the K2 Lx0 system has 3 LUNs and 3 LUNs of ONL or PRO storage is added, the bandwidth will only double even though the newer storage is of much higher performance.

For a major expansion of a bandwidth-constrained system, it may be less expensive to discard the original storage system. A Grass Valley sales specialist can calculate both options for the least expensive solution.

Summary

System storage and expansion options should be examined very carefully and reviewed with a specialist, and installation should always be performed by a service engineer.

It is important to consider:

- Total bandwidth needs
- Total capacity needs
- Using the GVConfig Tool to calculate bandwidth and capacity calculations
- Compatibility of all applications with the upgraded version of K2 system software
- Time needed to upgrade the system

Expanding current systems with new storage can add many more years to the life of a K2 system as the core video server will generally last much longer than the storage system drives.

K2 Online Storage in Hours

Times are for video with four 16-bit audio channels, no ancillary data (for 1 LUN of 6 drives). Stated hours are for 95% of total usable capacity.

RAID-5	8 Mb/s	12 Mb/s	15 Mb/s	25 Mb/s	50 Mb/s	80 Mb/s	100 Mb/s	DV25	DV50
450 GB	450	300	240	144	72	46	37	149	78
500 GB	500	333	267	160	80	51	41	166	87
600 GB	606	404	323	194	97	61	49	199	105
900 GB	906	604	483	290	145	93	74	299	157
1 TB	1006	671	537	322	161	103	82	332	174
1.2 TB	1206	804	643	386	193	124	99	398	209
1.8 TB	1813	1208	967	580	290	185	148	597	314
2 TB	2013	1342	1073	644	322	206	165	663	348
3 TB	3019	2013	1610	966	483	309	247	995	523
4 TB	4019	2679	2143	1286	643	411	329	1327	697
6 TB	6031	4021	3217	1930	965	618	494	1990	1045
RAID-6	8 Mb/s	12 Mb/s	15 Mb/s	25 Mb/s	50 Mb/s	80 Mb/s	100 Mb/s	DV25	DV50
450 GB	363	242	193	116	58	38	30	119	63
500 GB	400	267	213	128	64	41	33	133	70
600 GB	481	321	257	154	77	50	40	159	84
900 GB	725	483	387	232	116	74	59	239	125
1 TB	806	538	430	258	129	83	66	265	139
1.2 TB	963	642	513	308	154	99	79	318	167
1.8 TB	1450	967	773	464	232	149	119	478	251
2 TB	1606	1071	857	514	257	165	132	531	279
3 TB	2413	1608	1287	772	386	248	198	796	418
4 TB	3219	2146	1717	1030	515	329	263	1061	557
6 TB	4825	3217	2573	1544	772	494	395	1592	836

GVB-1-0544B-EN-AN