Aggregation is Key to Optimizing IP Network Design

Chuck Meyer, CTO, Grass Valley, a Belden Brand

August 2017

By using aggregation, today's networks can be protected against near-term obsolescence without necessarily adding the most expensive IP switch available

The world of IP technology moves rapidly. The speed with which data moves through an Ethernet connection and the amount of bandwidth that connection can accommodate increases at a rate that makes network planning challenging. What makes it even more challenging? Usually, higher bandwidth interconnects, such as active optical connectors, are initially more expensive and less available due to their rapid adoption by data center and telecommunications installations. Supply eventually catches up to the demand, but costs sometimes make early adoption prohibitive.

In the world of broadcast and media distribution, IP networks are being deployed to exploit the benefits that this technology presents: scalability; simplified geographic distribution; workflow efficiency; and the promise of future bandwidth increases. They are also being deployed in an effort to provide faster adoption of new media distribution services and formats as the role of traditional broadcast evolves into a more consumer choice-driven business model.

One approach to meeting this goal, in the presence of consistently-changing bandwidth, is to buy what is newest. However, this creates an even greater potential problem; that which is newest today is more expensive, and guaranteed to be slower tomorrow. Even worse, it may not even be fast enough to accommodate the next format to come along.

In order to design IP networks using Commercial Off the Shelf (COTS) solutions to accommodate for physical expansion and future bandwidth increases, it is constructive to analyze non-blocking, multicast switch architectures and the equipment located at the edge of the switch fabric.

Table 1 – Cost and Bandwidth Efficiency

This table shows the simple story of cost by comparing 10 GbE, 40 GbE (provided by four 10 Gbps links), 25 GbE and 50 GbE (coming soon). The video format is 4K, 10 bit 4:2:2 with the frame rate indicated in each column.

	10 GbE x 2		25 GbE	40 GbE	50 GbE
	50 Hz	60 Hz	60 Hz	60 Hz	60 Hz
Number of 4K Channels	2	1	2	3	4
Relative Cost*	2.00	2.00	4.25	4.00	6.00
Cost per Channel	1.000	2.000	2.125	1.333	1.500
% BW Used	0.87	0.52	0.84	0.79	0.84

* The Relative Cost is normalized per Gbps, and based on street pricing for SFP+, QSFP, SFP28 and QSFP28 optical components. For example, if a SFP+ is $60.00USD, then the 40 GbE QSFP is $240.00USD.

Bandwidth per link is only one aspect of system design, but it is essential. Table 1 demonstrates that 10/40 GbE can cost less than 25 GbE, which, in turn, may be more expensive than 50 GbE. The reason is channel efficiency. The more channels that fit in the link, the lower the system cost. High resolution video signals, for example, use lots of bandwidth and there's obvious incentive to economize the carriage of these signals. If 3 signals can fit, rather than 2, the broadcaster realizes a 50% improvement.

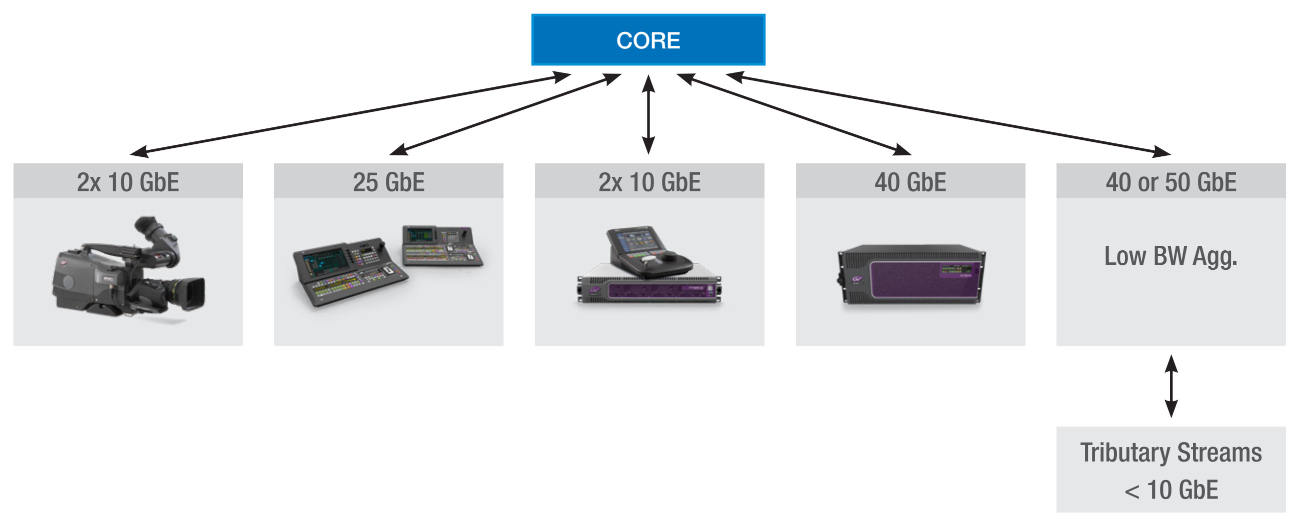
In a more practical example, let's talk about a signal emanating from one of today's popular 4K UHD cameras. In a network that is operating inefficiently, you may find that a 25 GbE connection is being used for this camera when it only requires 12 Gbps. The other 13 Gbps is completely unutilized – and the cost of wasting this bandwidth adds up fast.

The Case for Aggregation

IP switches are designed to exploit aggregation. Groups of lower bandwidth signals are gathered together and placed into a single, higher bandwidth signal. Aggregation is essentially a time division multiplexing process where tributary streams, or signals, are multiplexed into a trunk. Trunks can then be aggregated into larger trunks, or super trunks, to gain even greater bandwidth efficiency.

IP switches are designed to switch the signals within each trunk, meaning they manage the disassembly, routing and re-assembly of multiple tributary streams, carried within multiple trunks, and even super trunks. In leaf and spine topologies, trunks are used to connect leaves with spines. The combination of leaves and spines creates a switch that is equivalent to a core. When a modular, core IP routing switch is analyzed, it is fair to consider port cards as leaves and the fabric cards as spines.

Figure 1 – Modular Core with Aggregation



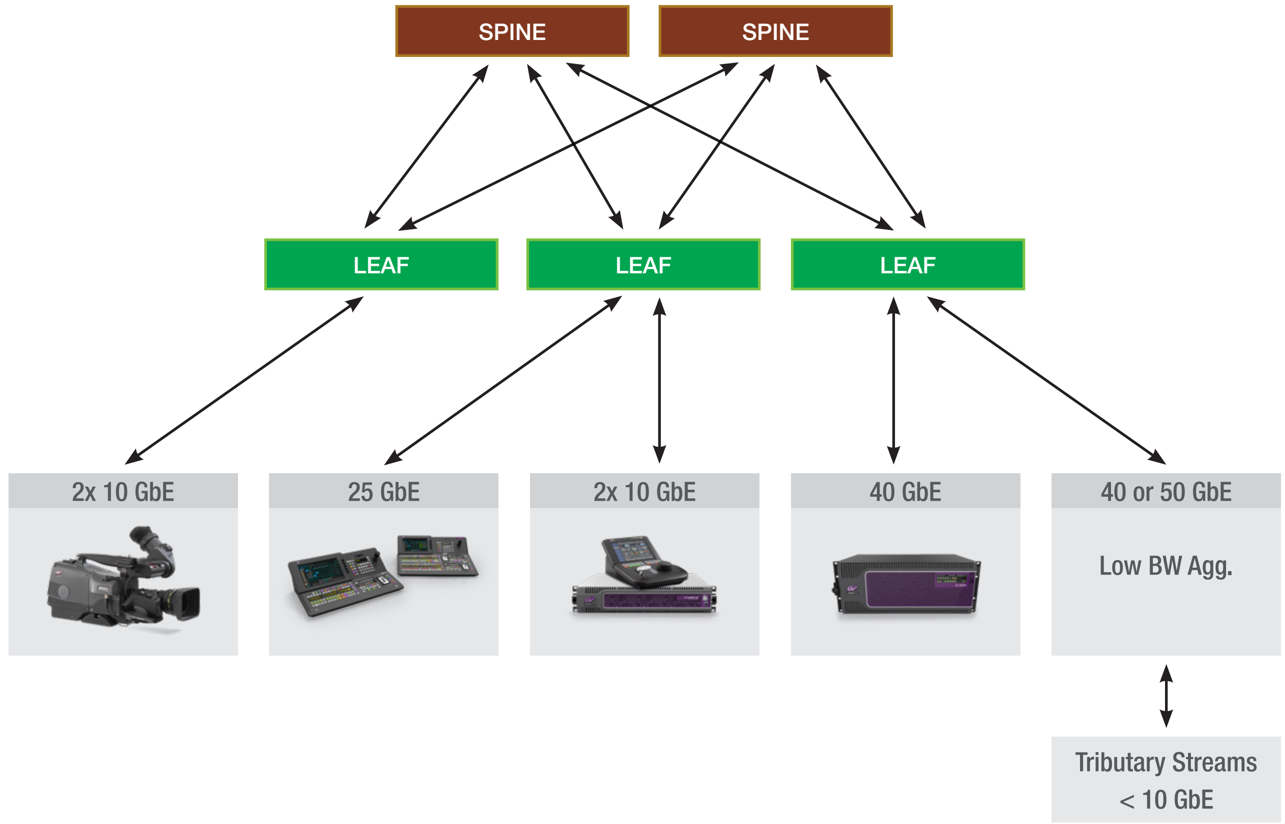
Spine switches typically contain very high bandwidth ports. For example, 36 ports of 100 GbE trunks can be connected to a number of leaf switches. Leaf switches tend to have fewer trunk ports, and more edge ports with lower bandwidth. Leaf switches are aggregators, combining signals into trunks. Typical edge port bandwidth is 10/40 and 25/50 GbE. The uplink ports, connected to a spine, are typically 100 and 200 Gbps. I/O modules in a core switch are typically used for tributary inputs with 10/40 and 25/50 Gbps bandwidth and additionally, there are I/O modules designed for 100/200/400 GbE trunks. These are most often used to connect between core switch.

In some cases, such as audio tributary streams, even 10 Gbps is too coarse, so a third layer of routing can be used to aggregate 1 GbE tributary signals into 10/40 GbE, as shown in Figure 1. These smaller trunks can then be fed to a leaf, or coreI/O module.

Due to broadcast workflows requiring multiple, granular data rates, sizes ranging from 1.5G to 12G for video and as small as 1.5Mbps for AES67 audio, purpose-built media aggregators can add significant efficiencies to a system's design. GV NODE is an edge routing and processing platform manufactured by Grass Valley, a Belden Brand, and provides an example of the type of aggregation functionality that can be used to a network operator's advantage. GV Node acts as a bandwidth aggregator with built-in conversion between legacy audio/video signals (AES/SDI) and IP, while also providing video processing functions. This approach optimizes bandwidth utilization and equipment density, both of which contribute to an overall lower cost system.

For example, (see Figure 2) a system designed to handle uncompressed 4K can be more cost-effective when this approach is applied. By leveraging a pair of 10 GbE SFP+ connections for uncompressed 4K video, the overall cost per channel and bandwidth efficiency compares favorably to a single 25 GbE connection, as shown earlier in Table 1.

Figure 2 – Leaf and Spine with Aggregation

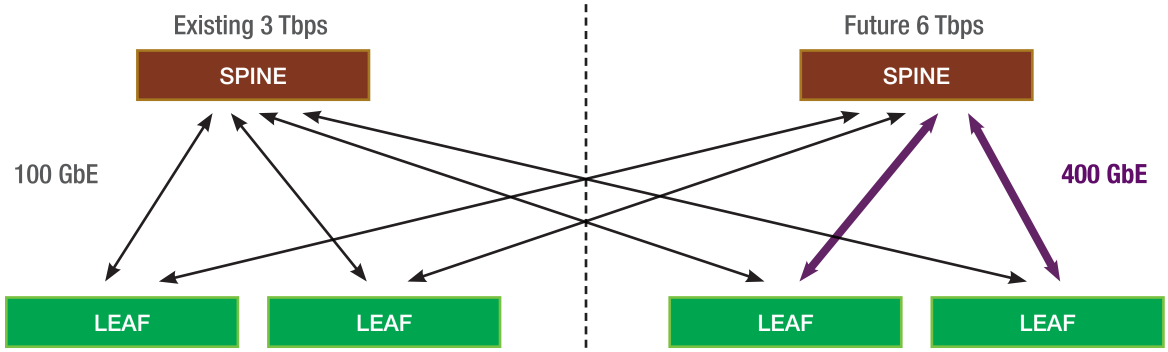


Planning for Future Needs

In the world of broadcast, future system expansion is a near certainty, so it is essential to consider future expansion and upgrades when looking at system design. With broadcast SDI routers, or a modular core IP switch, extra card slots may be left unused to expand in the future. SDI routers were typically designed for a fixed, peak bandwidth: a new format led to the need for a new router. However, in the case of a modular core IP switch, the ability to adjust bandwidth enables the adoption of a new format without installing a new router, but there's still the matter of peak bandwidth capacity to consider. Otherwise, the switch's capacity could be exceeded.

A leaf and spine topology provides the flexibility to mitigate this bandwidth capacity problem. Leaf and spine architectures can allow for open ports to be reserved on the leaf and spine elements, rather than leave open card slots. So, in addition to expanding the network by adding another leaf or spine, each element can also be expanded to enable a bandwidth increase. For example, Figure 3 shows a new/future spine that is 400 GbE capable – and in practice will likely support 100/200 GbE, as well. These lower bandwidth uplinks from the existing/old spine can now tie into a higher bandwidth new/future spine. Thus, the resultant network topology has more overall bandwidth.

Figure 3 – Future Network Expansion



This provides for important design choices as broadcasters look to create efficient, dynamic networks that are protected against near-term obsolescence. **As the strategies described earlier illustrate, it's not necessary to upgrade, or install, the latest, greatest, most expensive IP switch in order to future-proof your network.** In fact, today's 10G/40G network is a very suitable option. This technology is far from obsolete and offers many different switch sizes and price points that can be used as aggregators at the leaf. When combined with 25G/50G connections for essential areas where trunks have already been constructed, both bandwidth and cost are effectively optimized.

Of course, if additional budget is available, one can choose to start at the highest bandwidth possible, provided the next generation of necessary bandwidth is anticipated. As pointed out earlier, in some cases, 40 GbE and 50 GbE can both be more cost effective than 25 GbE, based on optical I/O. In this way, a network can be designed that meets your system needs today – and saves money – while remaining prepared to easily upgrade in the future to increase overall network size and bandwidth when requirements and economics align.

Many factors determine the final system price and topology. However, designing based on optimizing bandwidth and channel count can lead to lower cost systems. 10/40 GbE is not going to be obsolete any time soon so don't be too hasty to overlook the financial advantages of using the optimal combination of 10/40 and 25/50 GbE.

GVB-1-0650C-EN-WP