SMPTE Meeting Presentation

Large Scale Deployment of SMPTE 2110: The IP Live Production Facility

Steve Sneddon
NBCUniversal, steve.sneddon@nbcuni.com
Chris Swisher
NBCUniversal, christopher.swisher@nbcuni.com
Jeff Mayzurk
NBCUniversal, jeff.mayzurk@nbcuni.com

Written for presentation at the
SMPTE 2019 Annual Technical Conference & Exhibition

Abstract. *In 2016, NBCUniversal began the project to design and build the new global headquarters for Telemundo Enterprises in Miami Florida. The facility that became known as Telemundo Center would feature 13 production studios and seven control rooms supporting scripted episodic content, daily live news and sports programming, beginning with FIFA World Cup 2018. To support the scale and flexibility required for a facility of this magnitude, the key technical design consideration was the use of a software-defined video network infrastructure. At the time of launch in spring of 2018, Telemundo Center was home to the largest SMPTE ST 2110 environment in the world, consisting of over 12,000 unique HD sources and 150,000 multicast streams across audio and video.*

This paper will explore the major considerations and challenges in building such a large scale, all-IP broadcast production facility. We will demonstrate design factors around switching of video flows, redundancy, control and orchestration, PTP master clock systems and handoffs to multi-manufacturer SMPTE ST 2110 devices as well as non-IP enabled devices. This paper will also discuss our experience and lessons learned with utilizing a Software Defined Network (SDN) control plane and routing commands that abstract the underlying physical and link-level connectivity.

Some key topics include approaches to pooled resources and management of centralized operations; gaps in existing standards, with strategies to overcome limitations; the promise and the peril of differing ergonomic and performance characteristics of SMPTE ST 2110 endpoints – support for redundancy, clean switching, and audio/video synchronization

We will propose a reference architecture for supporting a GPS-sourced, large-scale PTP distribution to over 500 end points and explore some of the limitations and corresponding solutions encountered in PTP distribution at scale. Finally, we will demonstrate new software-defined infrastructure concepts such as virtual sources and virtual destinations which replaced legacy physical design patterns in this build.

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.

Contents

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.

Introduction

In 2016, NBCUniversal broke ground on a project to design and build Telemundo Center, the new global headquarters for Telemundo Enterprises in Miami, Florida. This new facility would bring together a combination of offices and studios for Telemundo Network, Telemundo Studios, Telemundo International, and Universo Network, as well as being the home of NBCUniversal International's Latin America offices. Prior to the opening of Telemundo Center, the staff of Telemundo Enterprises had been located at many older facilities around the Miami metro area. Telemundo Center allowed for all groups within the Telemundo Enterprises umbrella to come together under one roof in a modern facility. Apart from bringing business units together, one of the many goals of the project was to make the facility as technically future-proof and flexible as possible to be able to best serve Telemundo's needs in an evolving media landscape.

> "Telemundo Center is the manifestation of our commitment to the Hispanic market and a representation of our core values of innovation, collaboration and transparency," said Cesar Conde, Chairman, NBCUniversal Telemundo Enterprises and NBCUniversal International Group. "Latinos are a growing cultural, political, and economic force influencing every aspect of our country. Telemundo Center is the only facility that can fuel the preferences and demands of this dynamic audience, while driving unlimited growth and opportunity for our company, our employees and our community for years to come."
> *(http://www.nbcuniversal.com/article/nbcuniversal-telemundo-enterprises-celebrates-new-global-headquarters)*

Telemundo Center opened its doors in mid-2018 with the premier event being the 31 days of coverage of the FIFA World Cup 2018. The facility is now a hub of content creation delivering daily live news, entertainment shows, sports programming, and scripted episodic content across multiple media platforms including broadcast, cable and digital.

One of the major technologies we deployed to future-proof Telemundo center was video over IP using SMPTE ST 2110. At the time of launch, Telemundo Center was home to the largest ST 2110 environment in the world, consisting of over 12,000 unique HD sources and 150,000 multicast streams across audio and video. This paper will explore the major considerations and challenges in building such a large scale, all-IP broadcast production facility. We will demonstrate design factors around switching of video flows, redundancy, control and orchestration, PTP master clock systems and handoffs to multi-manufacturer SMPTE 2110 devices as well as non-IP enabled devices. This paper will also discuss our experience and lessons learned from designing, building and launching a large IP-only facility from the ground up.

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.

Facility Overview

Telemundo Center is approximately a 500,000 square-foot facility located on 21 acres in Miami, Florida. In addition to office space with a capacity to house 1,500 employees, the building features full production facilities to enable news, sports and scripted Entertainment for broadcast and digital outlets.

In support of those productions, the building features the following:

- 13 production studios in various sizes up to 8000 SF
- 5 live production control rooms
- 72 edit seats – approximately half of which are desktop edit and half are edit rooms
- 60 graphics creation seats
- A central video playback area
- A central graphics playback area
- A central camera shading area
- A transmission operations center

A central equipment room supports the above operational areas. At the heart of the central equipment room (CER) is a redundant set of core IP video routers using the SMPTE ST 2110. The CER itself is designed as a collection of 12 pods of 14-28 racks each, total of 290 racks. The CER also houses the fiber core—over 10,000 strand presentations of mostly single-mode fiber for plant-wide cross patching. Fiber core frames feed out to intermediate distribution frame (IDF) rooms throughout the facility for secondary cross patching to local endpoints. To feed studios, fiber bundles of up to 144 strands terminate in studio support rooms serving the IDF and production demark function for each production studio.

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.

Network Architecture and Core Technologies

The timeline of the build of Telemundo Center coincided with many technology shifts in the broadcast industry. The use of SMPTE ST 2110, and video over IP in general, was just emerging as a viable solution for a facility of the size of Telemundo Center and there were not many reference designs upon which to base the technology architecture. As part of the build process, we did an analysis of available technologies related to network architecture. The following sections detail some of the architectures that were considered and explain what we ultimately selected in each area.

Leaf-spine and single tier architectures

Leaf-spine is a network topology consisting of a small number of large core switches (spines) with a large number of smaller switches (leaves) which aggerate end-point devices. Generally, in leaf-spine topologies, end-point devices are only connected to the leaf switches and never to the spine switches.

With the goal of having a non-blocking network, leaf switches have low speed ports (10 GbE or 25 GbE) for endpoint connections and higher bandwidth (40 GbE or 100 GbE) uplink ports which equal in the aggregate the traffic from the endpoint devices to the leaf switches.

While a leaf-spine topology in a commodity datacenter environment for traditional IT applications may oversubscribe the uplink ports, in an uncompressed video environment it is often desirable for the available bandwidth of the uplink channels to equal or exceed that of the combined endpoints connected to the low speed ports. Note that the uplink ports do not need necessarily need to offer the combined *maximum available* bandwidth of the endpoint facing ports, only the *actual* combined bandwidth of all connected endpoint devices. A device, for example, might only use 1.5 Gbit/s of a 10 GbE port.

The use of leaf-spine topology in this manner offers an advantage in the ability to aggregate lower bandwidth endpoint devices into higher bandwidth links to the spine switches. Many endpoint devices may only produce or consume a relatively lower amount of bandwidth in comparison to the bandwidth available on the switchport they are consuming. For example, a camera may only produce a single 1080i video stream at 1.5 Gbit/s. However, as that device may occupy a 10 GbE port, the remaining bandwidth of that port is rendered locked and unavailable for use by other devices. If such “port waste” is inevitable due to endpoint designs, it is advantageous to absorb that waste at leaf switches rather than on the spine switch to maximize overall available network bandwidth. For example, a leaf switch of 10-count of 10 GbE ports fully subscribed would require a 100 Gbit/s of uplink bandwidth, but the same switch with only 1.5 Gbit/s utilized on each port would require less than 25 Gbit/s uplink bandwidth.

One of the major advantages of a leaf-spine topology is the ability to distribute leaf switches throughout a facility and locate them close to the endpoint devices to which are connected. This advantage helps make the use of top of rack switches attractive. All endpoint devices within a rack may be easily connected to leaf switches located within that rack. The “top of rack” model limits the need for copper or fiber cabling to traverse an equipment room or a facility and instead

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.

stay in a relatively short area and thereby limit the cost and complexity of physical integration by minimizing the total length of fiber in the build. It may also enable the use of lower-cost cabling such as multimode vs. single-mode fiber and, in some cases, the use of distance-limited copper ethernet or direct attach cable (DAC) solutions. The most expensive cabling, at the greatest length, would carry the greatest amount of bandwidth.

An alternative topology to leaf-spine is a single tier topology. In a single tier topology, there are very large core switches just like in a leaf-spine network. However, unlike a leaf-spine network, endpoint devices are connected directly to switch ports on the core switches. This direct connection of endpoints means that most of the ports on the core switches are endpoint-facing and therefore lower bandwidth. In a single tier topology servicing many endpoint devices it would not be uncommon to have several thousand lower bandwidth ports (10 GbE or 25 GbE) as opposed to the several hundred higher bandwidth ports (40 GbE or 100GbE) uplink ports typically seen in a leaf-spine network. Core switches in a single tier topology may still include some number of higher bandwidth ports (40Gb or 100Gb) for various specialty or bulk transport purposes, including high-bandwidth endpoints, facility interlink or connection between core switches.

One of the advantages of a single tier network, with core switches of very high port counts, is that no consideration needs to be given to the engineering or sizing of inter-switch bandwidth as required with a leaf-spine network. Generally, a single core of a modern data-center grade core switch is non-blocking at bandwidth that is equal to the sum of the available bandwidth of all ports. Though there may be scenarios where a core of blocking bandwidth could be considered for cost reasons. Switch specifications should be reviewed to confirm the expected performance.

The other advantage of a single tier topology is hardware deployment and infrastructure simplicity. In a single tier topology, there is simply less networking hardware. In both leaf-spine and single tier topologies, there are large core switches, but in a leaf-spine deployment, there is could be a substantial amount of leaf switches, which add an additional layer of cost, complexity, maintenance and wiring above that of a single tier deployment. Maintenance items such as periodic software and firmware updates can become time consuming when the number of switches in the network runs into the hundreds.

As part of the evaluation of technology leading up to the design and build phases of Telemundo Center, we considered a leaf-spine architecture for all the reasons listed above. Leaf-spine is a proven model with strong backing from vendors in the video space as well as many deployments in adjacent spaces. Leaf-spine is also more in line with current trends in network design for non-broadcast applications and may therefore be considered the more common or standard model.

However, as we embarked on the design phase of the project, we ultimately pivoted to a single tier topology for several reasons. First, as we designed the facility, we recognized that Telemundo Center was unusually large when compared with video over IP deployments to date. The leaf-spine network would have been very large to accommodate the production needs of Telemundo Center. While, larger leaf-spine networks have been routinely deployed in non-video data-center provider industries, the non-blocking or near non-blocking requirements of uncompressed video routing meant that leaf-spine topology was going to be extremely complex. The leaf-spine network would have to consist of multiple core switches and hundreds of leaf switches to support the production needs of the facility.

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.

Cost, complexity and maintenance became key concerns of installing and operating such a large, hardware-intensive network. Leaf-spine architecture requires network links both from endpoint to leaf, and from leaf to core. As such the cost of fiber and optics is greater than the equivalent connectivity for a single tier core.

This is not to say that a leaf-spine topology would be a bad choice for other facilities – but given the considerations and technology available at the time, a single tier network provided the most advantages for Telemundo Center.

We eventually landed on two single tier networks serving functional areas of the plant. These two functional areas are “Production” and “Acquisition”. Each of Production and Acquisition is serviced by a large core switch with over two thousand 10 GbE ports for connection of endpoint devices.

Division of endpoints between Production and Acquisition is chosen due to operational affinities, a functional requirement for each group of endpoints to primarily consume video within its group. On the Production core, everything related to studios and control rooms is connected, such as cameras, graphics devices, production switchers, playback servers. The Acquisition core supports incoming remote feeds, outgoing distribution, disk recorders and post-production.

The Production and Acquisition cores are connected to each other through a relatively small number of “tieline” ports, 3-count of 120 Gbit/s physical ports each supporting 12 lanes of 10 GbE, for a total 360 Gbit/s bandwidth. These cross-core connections are directly analogous to tielines that would have been used to connect two SDI routers together. The tieline connections are logically channelized such that they can service several hundred video flows traversing between the acquisition and production cores in both directions. The total switching capability between the cores does not need to approach non-blocking. These cross-core links simply need to support enough channels of video flows to support any functional requirement for video traversing from one functional area to another. Since, by design, most video flows stay within their functional areas, only very specific use cases need to be accounted for on these cross-core links. Key examples of cross-core video traffic are live video remote feeds from Acquisition to Production, and live program release feeds from Production to Acquisition.

SDN and Hardware Controlled Network

The video network at Telemundo Center is a Software-defined network (SDN), meaning that there is a software controller that instructs the control plane of the core switches how to route video flows. This software controller understands the physical network topology, with ingress ports and stream information for each source flow, egress ports and host information for each endpoint consuming video flows. The software controller provides a user interface to issue route requests, then instructs core switches to direct flows from ingress to egress ports. This control also includes the construction of multicast replication where appropriate if multiple endpoints are consuming the same source flow.

The alternative to a software-defined network would be a more traditional hardware-controlled network, where packet forwarding decisions are individually made in the hardware control plane of each node of the network. With no central controller, each network switch operates

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.

autonomously forwarding packets based on a set of predefined rules. Route requests may be issued directly to network nodes by endpoints via Internet Group Management Protocol (IGMP). While this type of network control scheme is the most popular by far for commodity data networks, we felt that it was not well-suited to routing video flows in a large live production plant. Hardware-controlled networks may not perform well at channelizing links with small counts of these constant high bandwidth streams and may not create multicast replication trees in the most efficient manner.

In summary, a software-controlled network is the solution we deemed best suited to provide functionality that would most nearly replicate the experience of an SDI video router. The "SDI-like" experience includes intelligent link provisioning to support the non-blocking and tieline performance expectations, as well as video production industry user interfaces, such as router control panels.

Video Network Redundancy

The previous section in the paper discussed the overall architecture of the video transport network. However, what was not mentioned in the previous section was the video transport network redundancy model. The model for redundancy in the Telemundo Center build was based on the SMPTE ST 2022-7 standard for "Seamless Protection Switching".

In an ST 2022-7 environment, there are two video transport networks that are always actively transporting video. These two active networks can be thought of as X/Y networks where every endpoint device is simultaneously connected and active on both the X and Y networks using double the amount of physical network interfaces that would be required in a non-redundant network. Half of the network interfaces on the endpoint device are connect to the X network and half the network interfaces are connected to the Y network. Having double the amount of network interfaces allow for all the bandwidth that the endpoint requires to be used on each the X and Y networks simultaneously. For example, a device sending or receiving 6-count of 1.5 Gbit/s streams would feature 2-count of 10 GbE ports – one each for the X and Y network.

Endpoints transmitting flows send packets to both X and Y networks simultaneously. Endpoints receiving flows receive and process packets from both and Y networks and perform a "hitless merge" of a single stream based on packets from either network.

In more detail: SMPTE ST 2022-7 specifies that Seamless Protection Switching senders will construct redundant X/Y packets with identical payloads, marked with identical RTP time stamp and sequence numbers. SMPTE ST 2022-7 receivers will receive the redundant packets from both X and Y networks. The receiver will identify redundant pairs based on the RTP time stamp and sequence numbers. If a receiver detects redundant packets from both X and Y networks, it then reviews packets for errors and the preferable packet is selected for further processing, display or de-encapsulation. If either X or Y streams are missing a packet, then the existing packet is preferable and selected for further processing. The net effect of this X/Y packet selection process is the "hitless merge" of redundant network streams. X and Y are considered to be an active/active redundant pair, and the merged stream may be constructed of payload data from either X or Y on a per-packet basis.

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.

One of the important factors to consider, when designing a video network using SMPTE ST 2022-7, is the time delta between the redundant networks. Because the receiving endpoint device must analyze and compare the RTP sequence number of packets coming from redundant networks, it must be allowed time to buffer incoming packets that arrive earlier from one network so that it can compare those packets to ones received from a redundant network. If the path that the X and Y networks take have very similar topologies and distances, the time delta between receiving packets on the primary and redundant networks will be minimal, and therefore the buffer needed for comparison can be very small. However, if the X and Y networks take different paths and/or the topologies of those networks are very different, the buffer needed by the endpoint device might be significant. If the receive buffer is adjusted too low for the network conditions, the receiver will not be able to effectively use the packets it is receiving from the lagging network and video dropout may occur in the event of a packet receive failure on the leading network. Conversely, if the receive buffer is adjusted to be too high, a perceivable video delay will be apparent to the end users. Therefore, care must be taken to implement the minimum buffer so as not to needlessly insert video delay into a production plant, yet still support the timing delta between the redundant video transport networks.

In the case of a WAN transport, it may be desirable to have very different network topologies which may have different delay characteristics. For example, when transporting video between two geographically different regions, one might want to take care that the redundant transport links do not share any common infrastructure. In long haul applications, this may mean that a flow on one will take significantly longer to arrive at the endpoint than its redundant pair on another link. In this scenario, the need for diverse infrastructure may force an increase in the ST 2022-7 receive buffer to move effectively utilize both paths for redundancy. Increased delay in exchange for better redundancy protection may be worthwhile tradeoff.

In the case of intra-facility video transport, it is typically not desirable to have any noticeable video delay. Therefore, care should be taken such that the X and Y video networks have very similar topologies and delay characteristics, such that the buffers on the receiving endpoint devices can be set very low, thus minimizing video delay within the facility. Software controlled network architecture may help with this, by ensuring parity in the flow path between redundant networks.

In the case of the Telemundo Center facility, we designed the X and Y networks to be as practically identical as possible. As explained earlier, the Production and Acquisition network are both based on single tier cores. To implement ST 2022-7 redundancy, we commissioned a redundant pair of such networks for each of Production and Acquisition. We have Production X and Production Y networks, with each production endpoint connected to each, and Acquisition X and Acquisition Y networks, with each acquisition endpoint connected to each.

Each of the X cores were located on physically separate power/UPS and cooling systems from the Y cores. While Telemundo Center is a large building, it is not large when taking into consideration the speed of light, with most fiber cable runs being well under 1 KM in length. Therefore, not a lot of care was needed to assure that cable runs were of similar lengths between the X and Y cores, as the delta between cable runs was not significant enough to affect the timing buffers of the receiving endpoint devices. We configured all endpoints at or near the minimum allowable buffer delay. The net timing impact of this is a configuration in which the overall delay though a receiver is not more than one video frame – inclusive of buffer delay and all other processing.

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.

An important thing to understand about ST 2022-7 Seamless Protection Switching is that the protection model offered does have limits. One important limit is that ST 2022-7 cannot protect against the full failure of an endpoint. A transmitting endpoint is responsible for creating both redundant video flows. So, if a transmitting endpoint fails completely, because of a power issue for example, no video flows will be transmitted to either redundant network and therefore video will not be received by any endpoint on either network. Similarly, if a receiving endpoint device completely fails it will not be able to process or display video regardless of the redundancy of the network. Because of this limit, highly important video should be also backed up using additional redundancy models – for example those redundancy models that might be utilized in an SDI plant. A completely redundant model for the most critical feeds would employ discrete source and destination feeds across diverse hardware, each transported redundantly with ST 2022-7.

An additional limit of the ST 2022-7 Seamless Protection Switching model is when there are multiple network path failures which are on the X and Y networks. For example, consider a transmitting endpoint that is dual fiber connected to an X and Y core and a receiving network device that is dual fiber connected from the X and Y cores. Under normal operation all is well, with full redundancy in place. In the event of a fiber failure on the X path from the transmitting device to X core, all is still well as video can be received by the receive endpoint via the Y path and Y core. However, if second failure occurs, this time on the Y link between the Y core and receive endpoint, video is now lost between the transmitting endpoint and the receiving endpoint – even though each endpoint still has a one good link up. Even more interestingly, other endpoints in the system will still be able to send and receive flows with these endpoints because they still have both of their links up. This situation can lead to a confusing troubleshooting which defies traditional broadcast source/destination testing logic – where a source is available to all but one destination, and a destination is available to all but one source. The problem may not "move" in a way that certain troubleshooting logic may suggest.

SMPTE ST 2022-7 Seamless Protection Switching is one of those most powerful new tools offered by IP video as compared with SDI. Redundancy is especially valuable in the case of very large IP networks with a failure block potentially equal to the entire video environment. But it is critical to understand the nature of the redundancy model and its limitations. The use of ST 2022-7 does not alone convey a "bulletproof" property to the video network and certainly not to the facility overall.

The use of a robust system health monitoring and alerting toolset is recommended to keep support teams informed of actual or imminent failures. Seamless redundancy may have the effect of masking critical system faults, and in the event of link failure or other outage protected by ST 2022-7 all care should be taken to repair the impacted leg and restore ST 2022-7 protection.

PTP and Reference Systems

In a traditional broadcast plant, a reference signal, commonly referred to as blackburst or genlock, is distributed to every piece of equipment the produces or processes video. Black burst itself is an analog video signal used as a common phase reference to synchronize the video generated throughout the plant, allowing for every source to be vertically in time with every other.

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.

In a video over IP plant, a newer method of synchronization is used: Precision Time Protocol (PTP). PTP is not a video specific technology; it has uses for providing highly accurate clock information across all kinds of computer networks. Major users of PTP outside of the broadcast industry are cellular telephone networks and financial networks where accurate time data is important. SMPTE ST 2110-10 defines the use of PTP as the synchronization method within an ST 2110 video over IP deployment. When designing an IP production plant, it is important to understand the mechanics of how a PTP clock system interacts with end-point devices.

Traditional black burst distribution is a one-way clock signal. Endpoint devices receive a clock pulse from a master sync generator, but never communicate back to that master clock. Black burst can, therefore, be duplicated and distributed through the plant using analog video distribution amplifiers. In contrast, PTP is two-way communication protocol in which endpoints both receive timing information and communicate back to the clock. And there are a variety of clock types that must be installed and maintained within a PTP network.

As a bidirectional protocol, PTP can be understood like a client-server relationship. Unlike blackburst, which can be distributed to an effectively unlimited number of endpoints, PTP clocks have a limit to the number of endpoints with which they can interact. The clock server can only support so many clients simultaneously.

It is outside the scope of this document to present a detailed technical analysis of how PTP works. However, we will touch on a few items that were relevant to the overall design at Telemundo Center. To help with that overview, it is important to define some of the components of a PTP generation and distribution system:

PTP Grandmaster – This is the ultimate PTP generator that sits at the top of a network. It can take time and phase data from an external source such as GPS and generate the master PTP signal on a network.

Boundary Clock – A boundary clock acts as a sub master clock on PTP network. It connects to a PTP grandmaster to obtain PTP clock data and then acts as a master to downstream devices, including endpoint devices or other boundary clocks. Boundaries serve in this way to segment the PTP network into smaller zones. Boundary clocks are important because, as stated above, any particular PTP master can service only a finite number of endpoint devices and it is important not to oversubscribe any PTP master clock. Adding boundary clocks in parallel to the existing boundary clocks is the proper way to scale a PTP distribution network as endpoint devices are added.

In a typical IT or Datacenter centric installation, primary and backup PTP grandmasters would be purpose-built devices that serve that are connected to the core or leaf switches on a network, but the boundary clocks may be integrated as features into the network switches themselves. This model limits the overall count of PTP-specific devices on the network, however it also may result in non-deterministic relationships between PTP devices.

As part of the Telemundo Center project we looked at a number of ways to build a PTP generation and distribution plant. We eventually settled on an architecture which allows for a more deterministic approach to PTP propagation than one relying directly on network nodes (switches) to act as boundaries. In our architecture, a set of four redundant PTP grandmasters are synchronized to GPS. Each of these is cross-connected to a redundant pair of 1 Gbit/s “Non-PTP-Aware” switches acting as a PTP distribution tier. Also connected to this PTP distribution

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.

network are a set of purpose-built boundary clocks configured to synchronize with the grandmasters. Each of these boundaries, in turn, is connected to the 10 Gbit/s ST 2110 video network to act as master clocks to endpoint devices. No switches or other network nodes are configured to act as PTP masters. The IP video network is configured to segment PTP traffic between these boundary clocks and endpoints so as to specify exactly which endpoints are locking to which boundaries and thereby ensure that we do not oversubscribe the boundary clocks. This solution serves overall to limit PTP complexity and reduce the possibility of cross-vendor issues in PTP support by effectively eliminating certain “automatic” features of PTP in favor of a more deterministic approach.

PTP is not the only reference system we commissioned for Telemundo Center. Telemundo, like many current IP installations, features a significant amount of SDI hardware requiring legacy black burst reference signals. Our PTP distribution solution also suited this need well; the boundary clocks, installed as pairs per network segment, are also configurable to serve as generators for black burst and other legacy analog reference signals. One of the several pairs of boundary clocks is wired to a changeover switch for analog signals and feeds a traditional black burst distribution system based on analog video DAs for SDI equipment.

As a final note on PTP, we found it important to take care in properly configuring the system to suit our needs. One of the foundational concepts of PTP is the “best master clock algorithm,” or the BMC algorithm. The BMC algorithm allows clock devices on the network to perform a “voting” procedure to elect one of the several available masters as the one to which they will synchronize. Several factors play into the BMC algorithm voting procedure, but chief among them is a tiered priority setting configured on the clock itself. The BMC algorithm also considers more nuanced factors such as quality of signal and the source of the master’s upstream synchronization. But priority is the principal tool that system administrators have to control the voting behavior. In an improperly configured PTP environment, the BMC algorithm may result in clocks assuming the role of master in contradiction of the system administrator’s intention.

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.

SMPTE ST 2110

SMPTE ST 2110 Professional Media Over Managed IP Networks is a suite of standards for use in professional content production which describe the mechanism for using Internet Protocol to transport video, audio and metadata streams. The roots of SMPTE ST 2110 come from the Video Services Forum (VSF) Technical Recommendation for Transport of Uncompressed Elementary Stream Media Over IP (TR-03).

- SMPTE ST 2110-10/-20/-30 - Addresses system concerns and uncompressed video and audio streams
- SMPTE ST 2110-21 - Specifies traffic shaping and delivery timing of the uncompressed video
- SMPTE ST 2110-31 - Specifies the real-time, RTP-based transport of AES3 signals over IP networks, referenced to a network reference clock
- SMPTE ST 2110-40 - Maps ancillary data packets (as defined in SMPTE ST 291-1) into Real-Time Transport Protocol (RTP) packets that are transported via User Data Protocol/Internet Protocol (UDP/IP) and enables those packets to be moved synchronously with associated video and audio essence streams

  (https://www.smpte.org/smpte-st-2110-faq)

One of the key advancements of SMPTE ST 2110 is that video, audio and metadata are all transmitted as separate IP multicast data flows. Having separate elementary essence streams over IP allows for a wide variety of content creation scenarios that would not be easily achievable if audio, video, and metadata were more tightly bundled together as in SDI.

As SMPTE and others provide countless resources to understand this standard, we will not dive further into the technical details of ST 2110 here. However, this standard was an essential component of the Telemundo Center build, and we exploited features of ST 2110 to solve workflow requirements at Telemundo.

At the time of the Telemundo Center build, ST 2110 was only newly available and not yet widely adopted. We did consider alternatives to ST 2110, which may have provided some short-term benefits in simplicity and supportability of the build. However, it was commonly expected in the industry overall that ST 2110 would mature to become the de factor industry standard. Telemundo Center needed to be forward-looking to ensure supportability well into the future. Ultimately, this consideration meant that ST 2110 was the only viable option.

As the project progressed, vendors rapidly moved to release ST 2110 based products to meet out timeline, and in the two years since we began the build ST 2110 has in fact matured to become the de facto industry standard for uncompressed IP video.

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.

Design Considerations and Implementation

Video Standard – HD, 3G, 4K

One of the often-advertised advantages of moving to an all IP infrastructure is the promise of being “format agnostic”. SDI and baseband technologies had a tight coupling between the bandwidth demands of their formats and their underlying transport medium. This model served SDI well, as the data rates required for SD (270 Mbit/s), HD (1.5 Gbit/s) and 1080p (3 Gbit/s) exceeded those provided by commodity Ethernet throughout the 1990’s and early 2000’s. In short, SDI baseband networks supported significantly higher data rates than were cost-effective for IP at the time. However, by now commodity Ethernet has far exceeded the capabilities that could be developed economically for baseband video transport. IP transport technologies have a well-defined separation between their transport bandwidth abilities and their payloads, meaning that payloads transported over an IP network can be as large as the underlying link allows. As commodity Ethernet bandwidths increase, IP will be able to transport video formats with ever-increasing bandwidth demands.

Telemundo Center is largely produced in 1080i 59.94. However, we sized all infrastructure with the assumption that each stream could run at up to 1080p. For example, 6 video streams at 1080i could be carried over a single 10 Gbit/s link. We would provision bandwidth for those 6 streams as 2x10 Gbit/s to support stream growth to 3G. The net effect of this is that our system overall has approximately 50% reserved bandwidth for future growth *per link*. There remains, of course, additional remaining available bandwidth from unused ports on the network.

We provisioned these links for 1080p for two main reasons. First, there would always be some small subset of content running at 1080p. Most notably, this includes multiviewer displays for control room monitor walls. These outputs have a much better perceived resolution when operating at 1080p vs. 1080i. Our 1080p bandwidth reservation ensures that these multiviewer mosaic displays can be routed anywhere in the plant and displayed. There are other examples of the need for 1080p, including studio monitor feeds and some content for post-production.

Second, and more importantly, sizing for 1080p supports a path to 4K UHD. Many devices in our production chain support 4K operation modes wherein a set of 4 video ports typically run discretely at 1080i can be run at 1080p in groups of four 3 Gbit/s signals for quad-link 4K. The most prominent example of this is the production switcher. By sizing IP bandwidth to 3 Gbit/s reservations for each stream, we enabled future support for this kind of 4K operation mode.

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.

Network Aggregation and Bandwidth Efficiency

One critical design consideration for IP video, especially in designs using a single tier network topology, is the efficiency of port utilization on the core switches. Multiple IP streams flow across each network link, and each port can be subscribed up its maximum available bandwidth. A 10 GbE port, for example, can transport up to 6-count of 1080i (1.5 Gbit/s) streams in each direction. Though that totals only 9 Gbit/s utilization, since the 10 Gbit/s cannot support an additional 1.5 Gbit/s stream it would be considered fully subscribed. The remaining 1 Gbit/s will be partially utilized for supporting traffic such as control and PTP.

A 10 GbE port would be considered significantly undersubscribed if it were transporting, for example, only one or two streams of 1080i 1.5 Gbit/s. There is no technical problem with undersubscribing ports in this way, but it is an inefficient use of overall network capacity. The remaining bandwidth on the undersubscribed port is unavailable for any other use, so it is considered "wasted." A significant undersubscription of many ports on a network can result in a significant amount of such wasted bandwidth, under-utilization of expensive infrastructure, and is recommended to be avoided wherever possible. This undersubscription problem works bi-directionally, as well. Even if the port is fully subscribed in the network ingress direction, network egress is also bound in that link.

One way to get around the issue of high port counts occupied with low bandwidth devices would be to install smaller switches in areas of the plant to aggregate bandwidth more effectively and thereby free up higher bandwidth ports on the core switches. This method of aggregation is exactly the advantage of a leaf/spine network, but it can also be developed in a more targeted fashion in an otherwise single tier network. For example, 6 devices each outputting a single 1.5 Gbit/s stream can be connected to an aggregation switch occupying a single 10 GbE port on the core. Add another 6 devices each receiving only a single 1.5 Gbit/s stream to fully subscribe the 10G core port bidirectionally.

Aggregation of this sort can help unlock the full capacity of the network. The downside is that such aggregation switches cost money and add complexity. Broadcasters need to strike a careful balance between avoidable port waste and excessive aggregation. With too much aggregation, some of the architectural simplicity advantages of a single tier topology may be lost.

For the Telemundo Center build, we considered all levels of aggregation, from a very aggressive model where we would try to conserve as many core ports as possible, to using no aggregation at all. We eventually landed on a model of using limited or light aggregation. We accepted some portion of port waste on the core but provisioned some aggregation switches supporting banks of similar low stream-count devices. This model left us with more than enough ports in the core switches for future growth and allowed us to still have a simple core switch design.

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.

Audio Transport Considerations

As we discussed earlier, an important feature of SMPTE ST 2110 is that video, audio, and ancillary data are transported as separate multicast streams which can easily be routed to different destination endpoints independently. Audio streams encapsulated as AES67 may be routed separately from video streams to an audio-only endpoint such as an audio mixer. Unlike in an SDI environment with embedded audio, no multiplexing or de-multiplexing equipment is needed to separate or combine audio streams to their related video streams. This means that there is no need to waste link bandwidth on transporting video multicasts to audio-only devices. Discrete multicast routing also means that we can virtualize audio embedders (multiplexers) by routing video and audio streams from separate devices (for example, production switcher and audio mixer) to the same destination endpoint.

For Telemundo Center we had initially planned to use AES67 audio networking for everything - including streams that were part of SMPTE ST 2110 video sources and also independently generated audio sources such as the output of studio microphone pre-amplifiers. The goal here was to have one media network for all video and audio.

As the project progressed we found that this truly unified audio and video network environment was not yet ready for market; large scale AES67 deployments were not ready to fully interoperate with ST 2110. With that limitation in mind, we landed on design to bridge the SMPTE ST 2110 video/audio transport environment with a more traditional audio-only router environment connecting production audio consoles to studio microphone pre-amps and IFBs. The bridge between these two networks worlds is a bank of bidirectional ST 2110 to MADI converters. These convertors serve as sources and destinations on their respective network to pass audio between environments.

Another key audio consideration is that there remains limited standardization in the packaging of mono audio channels within AES67 multicasts. Some vendors have chosen a method of packaging one mono audio per multicast, while others have chosen 16 and some have chosen four. This non-standardization of packaging of multicasts has led to some difficulties interfacing audio streams between vendors. This potential incompatibly was another reason we chose to keep the audio/video world and the audio only world separate, only connecting the two with bridges that could reformate audio packages in the way we needed.

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.

Pooled Resource Management & Operational Presentation

This section shifts focus away from IP video standard and technical considerations, toward exploring some of the ramifications of a large-scale IP build for production operations. First, we will consider traditional paradigms of video source presentation in an SDI video architecture. Imagine, for this example, an SDI broadcast plant featuring several Production Control Rooms each with its local router and a Core router for shared resources. There are three main ways in which sources in this environment are presented to destinations within the PCR local router.

- **Local Sources** – These are source directly wired to the PCR local router, with a typically non-blocking capability to route to any destination on the local router. These sources would have "local" naming, which would not need to specify the PCR to which they belong. For example, a camera CCU wired as a source on the control room local router can be named "CAM-1." And any number of PCR local routers can have a separate source, each called "CAM-1." There is no inherent conflict with this. There is no functional benefit to giving these sources any kind of globally unique name, either specifying the PCR to which they belong ("PCR 1 CAM-1," "PCR 2 CAM-1") or counting they within a group of similar sources across the plant ("CCU 19," "CCU 43").
- **Core Sources via Managed Tielines** – These are sources wired directly to the Core router and routed into the PCR local router via automatic tielines managed by the router control system. The switching capacity of these sources to PCR local router destinations is blocking – it has an upper limit to count of sources equal to the number of available managed tielines. These sources would be required to have a globally unique naming convention, as they are shared with and available to all PCRs. For example, the shared pool of all camera CCUs for the plant would have names like "CAM-1" through "CAM-25," without a separate "CAM-1" for each of PCR 1 and 2 as seen in the local source example. Core sources routed to local destinations via managed tielines would retain that globally unique name. Therefore, a destination in the PCR local router would see the globally unique name, e.g. "CAM-25," not localized for that PCR. Tallies and UMDs, similarly, would see these globally unique names.
- **Core Sources via Manual Callups** – These sources are wired directly to the Core router but routed into the PCR via manual destination routing on the Core. This core destination might would be wired directly to a source on the PCR local router. Then the PCR local router source behaves like a locally wired source device on that router. Routing capacity would be blocking from Core to Local, but the callup source on the local router would have non-blocking capacity to any destination on the local router. That source, therefore, would behave from the perspective of the local router exactly like a source device wired directly to the local router. The key functional difference between manual callups and managed tielines is that with manual callups the Core router source will not retain its globally unique name through to local router destinations. The manual callup represents a break point where the core source with its globally unique name can be "localized" to the PCR and assume only a locally unique name. For example, core router source "CCU 25" can be manually routed to core destination "PCR 1 CAM 1" and enter the PCR local router as a fully localized "CAM-1." And each PCR can have its own "CAM-1" to provide a localized presentation of core shared resource.

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.

Each of these paradigms may be appropriate for various workflows within such an SDI plant, and each has advantages and disadvantages:

> **Local Sources** are port-efficient, requiring only a single local router source to present to a local router destination. Their usage is clearly indicated, operational routing is straightforward and easy to understand. The disadvantage is that these resources are essentially “locked” to the PCR local router, not inherently available for sharing with other PCRs connected to the core. It is easy but inflexible.
>
> **Core Sources via Managed Tielines** are port-inefficient, requiring a core router source and destination, as well as a local router source, to present to a local router destination. The operational routing experience is also straightforward, with a single “take” delivering a core source all the way to local destination. Managed tielines also carries several disadvantages. First, it requires an intelligent tieline management system, which increases system complexity and support overhead. Second, it does not allow for a localized presentation of the core resource. So rather than an operations-friendly “CAM-1,” the PCR would interact directly with the globally unique source “CCU 25.” Directors may find it confusing to “Ready CCU 25” today and “ready CCU 43” tomorrow. However, such globally unique naming may be acceptable and appropriate for certain uses. For example, a patchable router input presented at a studio broadcast service panel (BSP) would always be considered a globally unique source outside of the PCR.
>
> **Core Source via Manual Callups** are similarly port-inefficient, again requiring a core router source and destination, as well as a local router source, to present to a local router destination. Manual callups require two separate “takes” – first, to route core source to core destination, then a second to route local source to local destination. However, the key benefit of manual callups over managed tielines is that they allow for localized presentation of the core source. Therefore, directors will always have a local source to refer to as “CAM-1” in every PCR irrespective of which global CCU is routed into it. Additionally, local source-destination routing may be “pre-loaded,” for example to ensure that the localized “CAM-1” is always routed to the appropriate local destinations (e.g. switcher inputs or monitors). So even though there are two “take” events required, they do not necessarily need to occur at the same time. This can ease the burden of operational use of this paradigm.

We recognize that this model of multiple “Local” vs” “Core” routers is a foreign concept to many broadcasters, where smaller facilities may require only a single router to support the entire plant infrastructure. For a facility at the large scale and complexity of Telemundo Center, this type of complex multi-router interconnect would have been a reality if we designed the plant in SDI. However, while not all projects may have the same scale challenge as Telemundo, the solutions discussed below are applicable to any infrastructure based on a single flat router environment where resources must be shared between control rooms.

For Telemundo Center, a key design strategy was the elimination of the split between “Core” and “Local” routers for production video – specifically, sources and destinations pertaining to studios and control rooms. All of these I/Os are connected to a single non-blocking IP network. This includes camera CCUs, GFX, DDR and other playback systems, production switchers, audio mixers, processing gear such as color correctors, studio BSP video ports, multiviewers and displays.

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.

The key challenge for operational presentation in the build was how to simulate each of the SDI-era paradigms for operational presentation within a single large, flat non-blocking IP environment. All physical sources and destinations are required to have globally unique names, and there is no physical "break point" or interconnect to create a localized presentation based on manual callup.

Without care to manage operational presentation in this environment, the default operational experience would most nearly mimic the Managed Tielines paradigm in SDI – but in a fully non-blocking capacity. Every route would be performed from a globally unique source to a globally unique destination. Tally and UMD information would always present the globally unique naming, and there would be no mechanism to "pre-load" or pre-route PCR setups without foreknowledge of the global resources to be used. A major component in the functional design of Telemundo Center was a development of strategies to make a flat matrix operate like a legacy environment in terms of localized presentation of sources. These strategies include virtual loopback routing and an extensive use of router I/O namesets.

First, we will examine virtual loopback routing. Of course, in an IP build there is an option for true "non-virtual" loopback routing. A destination or router output may be wired directly to another source on the same router – either strictly within the IP domain or via a SDI gateways. This would allow for a for a workflow simulating the Manual Callup paradigm, with all the inherent benefits, however it comes with the same port-inefficiency downside of callups between SDI routers. It is wasteful of actual ports and bandwidth on the IP network.

We made extensive use of virtual loopbacks within the router control system to simulate Manual Callups in a more efficient manner than physical loopbacks. These virtual loopbacks do not require any physical hardware, the use of ports or extra bandwidth on the IP network. They are created as a virtual object within the control system and present themselves as both a destination and a source. And they can be seen and controlled from user interfaces (e.g. router control panels) just like any physical source or destination. In a single virtualized object, they simulate an SDI core router destination and its direct connection to an SDI local router source. We created many hundreds of these for each PCR and, in the aggregate, they form a completely virtualized PCR local router.

The workflow for virtual loopbacks is to route a global physical source, e.g., "CCU 19" to a local PCR loopback destination, e.g., "PCR 1 CAM 1." That virtualized local "CAM-1," can be used both for operationally-friendly local naming (Director calls "CAM-1" irrespective of the physical CCU assigned to it). It also allows for a break point to pre-route all the localized virtual resources within the PCR. So local "CAM-1" can be pre-routed to various destinations (switcher inputs or monitors), and those route relationships are persistent and flow from whatever physical source is routed to the local virtual loopback. The localized loopbacks also offer non-blocking routing capacity to any destination through the plant, either inside or outside the PCR. We thereby simulate within the IP environment an SDI Manual Callup paradigm, but with the key advantage that it is non-blocking across the IP production network. Since these localized virtual loopbacks consume no bandwidth and no physical ports, there is no inherent limitation on the number of them that can be created. In a facility with 100-count of physical, global "CCU 1" through "CCU 100," we can build a matching set of localized PCR "CAM-1" through "CAM-100" to provide total operational flexibility in assigning the pooled CCU resources to the control room. There is also, of course, no requirement that the global CCU and local CAM numbers match 1-to-1. So, while "CCU-1" may be used as "PCR 1 CAM-1," we may use "CCU-19" as "PCR 2 CAM-1." The solution

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.

allows for non-blocking resource sharing between PCRs with the benefit of locally friendly naming, pre-routing of local resources, and no impact on actual network bandwidth or port capacity.

Next, we will examine the use of namesets. Without a layer of nameset management, all sources on the network, whether physical or virtual, would be required to have a globally unique name. This is clearly the case for true global physical resources, but what about the local PCR virtuals? In the above explanation, these virtual loopbacks were sometimes referred to with globally unique name such as "PCR 1 CAM-1" and sometimes with a truly localized name such as "CAM-1." Our use of namesets allows us to have it both ways. Different columns within the nameset table for all these virtual I/Os present either globally unique or PCR-localized naming, with different namesets presented to different PCRs and other functional areas as appropriate. The below table provides an example of this:

Type	Global Name	PCR 1 Name	PCR 2 Name
Physical	CCU 19	CCU 19	CCU 19
Physical	CCU 20	CCU 20	CCU 20
Virtual	PCR 1 CAM 1	**CAM-1**	PCR 1 CAM 1
Virtual	PCR 1 CAM 2	**CAM-2**	PCR 1 CAM 2
Virtual	PCR 2 CAM 1	PCR 2 CAM-1	**CAM-1**
Virtual	PCR 2 CAM 2	PCR 2 CAM-2	**CAM-2**

Note the pattern, which has the following properties:

- Every source has a globally unique name
- Physical sources present their global name across all namesets
- Virtual sources present their global name within "foreign" (another PCR) namesets, but present a localized name, stripped of the PCR specification, within their local PCR nameset.

Within a given PCR, UMD labeling for virtuals does not the specify the PCR. The implication to operators in that room is that the source is "my" local CAM-1. To make that resource available in other rooms, we specify its PCR locality. Within PCR 2, we have a local "my" CAM-1, as well as a source explicitly labeled as "PCR 1's" CAM-1 to distinguish remote from local.

These namesets are exposed as appropriate to each PCR and operational area so that users in those spaces see the simplest version of the source name as it applies to them. Shared service areas outside of any PCR will always see the Global Name, since they have no necessary affinity to any PCR.

Namesets may also be used in this way to localize any physical sources which are permanently assigned to the PCR and which do not require routing through the virtual loopback infrastructure. This mechanism is used to simulate the paradigm of SDI local router sources. It can be explained via the below table.

Type	Global Name	PCR 1 Name	PCR 2 Name
Physical	DDR 19	**DDR 1**	PCR 1 DDR 1
Physical	DDR 20	**DDR 2**	PCR 1 DDR 2

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.

Physical	DDR 21	PCR 2 DDR 1	**DDR 1**
Physical	DDR 22	PCR 2 DDR 2	**DDR 2**

Note the pattern, which has the following properties:

- Every source has a globally unique name
- Unlike the pooled resources with local presentation via virtual routing, we use only namesets to directly localize the physical source to the PCR
- These sources present a name to "foreign" PCRs which is derived from the fully-localized name, but with PCR or zone specified.
- In this example, the globally unique nameset is shown as distinct from the "foreign" localized name in each PCR nameset, but that is not required. Either is acceptable and may be used interchangeably.

At Telemundo Center, we sparingly employed this method for localizing physical sources without the virtual loopback layer. This method limits flexibility to dynamically reassign resources to PCRs but carries the benefit that the local assignment is baked into the naming and thus doesn't require any operational management moving forward. We employed this only for devices which for practical reasons could not be shared between control rooms, or where there was no operational benefit to such sharing. A key example would be physical video monitoring destinations within the PCR. Those are inherently bound to the PCR itself, so they require no virtualization layer to localize to the PCR – only the local friendly naming via namesets. This model would also be appropriate for the limited use of actual, non-virtualized physical loopbacks we briefly mentioned above – where we have switchable/assignable resources based on physical rather than virtual loopbacks.

Finally, note that since all physical devices are connected directly to the production IP network in any case, there is only a software configuration difference between the use of virtual loopbacks and localized assignment of direct physicals. Any source can be ported from one to the other paradigm with a control system configuration change and would not require any wire work or hardware installation.

Top Down and Bottom Up

At a macro level, there are broadly two techniques to manage the localization of resources in a large pooled environment. We will term these the "Top Down" and "Bottom Up" paradigms. There exist control systems solutions in the market to provide Top Down management of pooled resources. These solutions may offer advanced intelligent functionality such as scheduling and automatic/managed assignments of resources from the pool. However, they may also be a "black box" performing assignments that cannot be traced directly to any action within the underlying router control system. While such a "black box" may provide an opportunity for a managed user presentation layer, it would tend to be limited to those functions and interfaces specifically built for user interaction. Developing these functions and interfaces may be complex and time consuming, and such a system architecture may be such that there is no available or convenient "back door" option to work around it either to perform ad-hoc assignments outside the scope of pre-built functions, or to operate the facility in a DR capacity if the management system is in a non-functioning state (system crash, etc.).

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.

Due to our concerns about these limitations of true Top Down management, we elected for Telemundo Center to develop the virtual loopback solution as a Bottom Up alternative. In the Bottom Up approach, all resource assignments can be expressed as a route event within the underlying control system. Assigning global "CCU 25" to local "PCR 1 CAM-1" can be performed as a Source-Destination route. This model imposes no inherent limitation on operations to work only within a particular managed environment. Assignments can be performed manually, one by one. They can be performed in groups via salvos to describe standard setups and common assignments. Additionally, since these assignments are route events within the control system, they can also be performed via remote automation interface using common video router control protocols. Access via automation leaves us the option to employ any one or a variety of external control systems to perform advanced functions available in a Top Down system (scheduling or automatic/managed allocations), and seamlessly move back and forth between assignments managed by automation and those managed manually. In fact, we did employ at Telemundo a Top Down management system to perform certain complex assignment management tasks, but since that system acts as an automation interface rather than a self-contained "black box," we have options available to work around it where necessary.

In summary, by building a resource management solution out of basic building blocks with multiple standardized control points, we have developed a user-friendly experience allowing for production flexibility and the potential for adding or changing external automation solutions as future needs require.

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.

IP to the Edge – How Far to Go?

Some consideration should be applied to the question of how for to the edge should an IP infrastructure extend. In the case of a content creation facility this question applies largely to "in front of the camera" display technologies such as on-set monitors and LED walls. These devices are typically not IP natively enabled and the overall environment is subject to physical strain and possible damage through constant production movement. From a network infrastructure simplicity perspective, an ideal situation would be for Ethernet to extend all the way to an endpoint, even if that endpoint is an on-camera monitor. Practically though, extending IP connectivity all the way to such an end point may introduce a level of risk that is unacceptable. On-camera set elements may be moved, disconnected, and reconfigured often. While a fiber-based Ethernet link can be ruggedized, it will probably never be as comparable to a coax-based BNC connection that SDI uses from a reliability standpoint when plugged and unplugged often. Additionally, since an SDI connection needs very little configuration at the endpoint, it is more ideal for use in an environment used to that sort of reconfigurability.

When building the studios at Telemundo Center, we looked a wide variety of options around where to make the transition from IP to SDI to feed elements in the studios. One possibility was installing SDI gateways in a central location to feed all the studios. However, because of the size of Telemundo Center and the size of the studios, we would have quickly exceeded the allowable cable length for SDI over coax. In the other direction, an ideal landing point for IP to SDI conversion would have been inside the studio broadcast service panels (BSPs). This location would have been the best practical place because it would have brought IP right into the studios and then allowed for SDI over coax for "last mile" connectivity to the endpoints on the sets. However, at the time of the build there was very little IP gateway vendor equipment that was both low profile and had sufficiently quiet fans to be usable in a studio BSP enclosure. We eventually landed on placing the IP gateways for each studio in an IDF closet that was co-located with each studio which also housed corporate network equipment serving the studio. This solution was a good compromise, though since the time of the Telemundo Center build, vendors now have a larger variety of IP gateway equipment that would be acceptable to be located within a quiet studio.

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.

Conclusion

Missing Components – What do we need?

Adoption of a common stream connection management protocol across vendors. This need is directly addressed by the AMWA NMOS specifications, and a variety of other proprietary and niche options are available as well. But limitations on interoperability between endpoints and control systems remains a challenge for IP systems design. At the time of the Telemundo Center build, we often stayed within a single vendor family, or opted to use the SDI versions of end-point equipment with IP Gateway devices, to avoid cross-vendor connection management integration. A fully compatible and broadly adopted stream connection management system is key to the success of larger installations that intend to use products from multiple manufacturers.

Adoption of common stream switching mechanics across vendors. This need is distinct from the stream connection management addressed by NMOS. Endpoints, control systems and network switches support a variety of different mechanisms for stream switching at route time. The route time event requires three distinct steps – Teardown of the existing flow; Sending a new flow; Instructing the endpoint to subscribe to the new flow. Available methods include:

- **Break Before Make** - The existing flow is first disconnected, then the new flow is sent along with the connection instruction. The method has two distinct advantages. It is bandwidth efficient for links, and it does not require precise synchronization of the routing steps. The key disadvantage is that it is the slow and will produce a visual artifact at the time of the route switch – such as black, "freeze frame," or a glitch.
- **Make Before Break** - Endpoints undersubscribe their link, with up 50% reserved to accommodate stream switching. The new flow is sent along with the connection instruction. After the endpoint connects to the new stream, the prior existing flow is disconnected. This method has the advantage that it can provide a visually seamless switch. It also requires no more precise event synchronization than "Break Before Make." The key disadvantage is the undersubscription of bandwidth, which at scale can mean a significant amount of waste in network infrastructure.
- **Synchronous Switching** - The existing flow teardown, send of the new flow and connection instruction are precisely coordinated in time to occur during the SMPTE RP 168 vertical interval switching point. Where available, synchronous switching can provide the most "SDI-like" route experience, without the downsides of "Break Before Make" or "Make Before Break." However, since this method involves precise timing coordination of both network and endpoint components, it has specific support requirements and is generally unavailable outside of single-vendor implementations.

Each of these switching mechanics may be considered acceptable for various uses, however there remains little commonality amongst vendors in adoption of any mechanic. Optimally, endpoints and network infrastructures would support a multitude of switching mechanics for the widest possible compatibility. But broad adoption of synchronous switching mechanics would provide the best possible user experience.

A related concept to the stream switching mechanics is the viability of various methods within software and hardware-controlled network infrastructures. A hardware-controlled infrastructure would tend to support a "bottom-up" method of route initiation, such as one based on IGMP

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.

signaling. Here, an endpoint would issue a stream request directly to its local network node, and the route event request would be propagated through the network to achieve existing flow teardown and new flow delivery. This method is not well suited to support the kind of precise synchronization required for truly synchronous switching and is therefore more applicable to a “Break Before Make” model. Our requirement at Telemundo Center for the most “SDI-Like” route switching experience is a key component of our choice to deploy a software-controlled network environment.

The adoption of advanced SMPTE ST 2110-40 ancillary data stream processing. Specifically, the ability of endpoint receiving devices to subscribe to multiple ST 2110-40 ancillary data streams simultaneously and utilize them in a combined fashion. The functionality of stacked ancillary streams would be the equivalent of SDI passthrough data inserters wired in series. An example use case is the scenario in which some set of endpoints requires just closed captioning data, while another set requires both closed captioning and ANC triggers. The SDI model for this would be a forked path with an upstream passthrough caption encoder and a downstream ANC trigger inserter. The serial nature of such an SDI path imposes restrictions on recombination and elimination of ancillary data – for example, both forks would necessarily include any other data inserted upstream of the caption encoder. In a fully-realized IP solution, passthrough data inserters would be replaced with ST 2110-40 data senders, outputting ANC-only streams to the network. Receivers could then subscribe to any number of these atomic ANC streams and combine them in de-encapsulation. This multiple subscriptions would be analogous to the way in which ST 2110 receivers today can receive multiple ST 2110-30 audio streams, along with video streams, and combine them in de-encapsulation. While the ST 2110-40 standard should allow for this scenario, it seems that no vendor has yet to implement such functionality. Indeed, there is limited existing product for IP-native ANC processing, even in a passthrough capacity. For Telemundo center, all ANC encoding services, including closed captioning, are provided with SDI passthrough devices connected to the network via SDI gateways.

One potential interim step to a “stackable/atomic” ANC workflow directly to generic endpoints in ST 2110 would be a discrete subsystem for receiving multiple ANC streams for processing and explicit recombination. An ST 2110-40 ANC combiner could receive a stacked set of ANC streams and output a single stream representing the combined payload. This concept of “pre-grooming” ancillary data would alleviate the need for new stacked ANC processing features in generic endpoints while still allowing for a dynamic recombination workflow. Similar subsystems for audio purposes are common in ST 2110 environments, where audio grooming solutions receive multiple audio streams, then process and repackage them to suit the requirements of various endpoints – to support audio channel shuffling, limitations on audio multicast receive counts, or diverse multicast channel count requirements.

Audio stream packaging standardization and improved flexibility. The ST 2110 standard specifies that devices support modes of between 1 and 8 audio streams per multicast. In practice, this standard has not been widely adopted, and vendors today will typically offer only a single option for audio multicast channel counts. In an ST 2110 environment assumed to support 16 channels of audio per endpoint, there is a tradeoff to be considered in implementing channel count specifications. Low stream counts (e.g. 1) offer flexibility and improved dynamic channel control, while increasing network overhead, configuration management complexity processing requirements for endpoints. High stream counts (e.g. 16) offer reduced flexibility with limited dynamic channel control, while minimizing network overhead, configuration management and

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.

processing requirements. In practice, most endpoints support just a single channel count specification somewhere between 1 and 16 as a way of balancing the advantages and disadvantage of both extremes.

There are two key problems with the current state of audio multicast support: first, multi-vendor interoperability, where channel counts will generally match for senders and receivers within single-vendor product families, such agreement is not at all guaranteed between different product families in a multi-vendor solution. Second, single-vendor dynamic control, where even if send and receive channel counts per multicast agree, there may be addition limitations imposed on multicast receive counts, due to processing capacity or similar.

For example, consider a receiver with a limited audio multicast receive count and a locked channel count, supporting 4-count multicast audio streams with 4 channels each. All 4 multicast receivers must be engaged for a total 16 channel audio payload. While such an endpoint may be able to perform channel shuffling within those 4x4 receive channels, it would be unable to replace any one of those 16 channels with even a single channel from a 5th multicast stream. An audio grooming subsystem would then be required to subscribe to the 5th multicast and package it inside a new 4-count of 4 channel multicast streams. In fact, this exact scenario exists in the endpoint solution deployed at Telemundo Center. For this reason, we have chosen to limit the overall expected audio payload (as a general plant specification) to 12 channels, down from the 16 channels expected in SDI. This 12-channel count leaves one audio multicast subscription available at each endpoint to support shuffling in up to 4 additional channels beyond the baseline 3x4 audio multicast standard.

Adoption across vendors of common methods for IP-based trigger and tally data. In the legacy SDI environment, tally and triggers were communicated via a variety of physical connections and protocols – including general purpose input/output (GPIO), serial connections, and IP-based solutions. While such device control considerations are out of the scope of the ST 2110 standard, it would be desirable in IP builds to deprecate wherever possible the need for GPIO and serial connectivity and replace those functions with IP based solutions. One major driver for this need is the increasing use of virtual computing solutions in place of legacy hardware appliances. For such virtualized devices, non-IP (Serial/GPIO) interfaces tend to be either impossible or impractical to implement.

A variety of possible solutions exist to address this problem, including specifications address in NMOS IS-07 and preexisting proprietary protocols. At Telemundo Center, for example, GPI triggering from production switcher is accomplished via a solution consisting of GPIO-to-IP interfaces, virtual GPIs and custom control drivers for implementing device-specific APIs over IP. Broad support for IP-native triggering and tally protocols would dramatically simplify builds from the perspective of hardware, wiring installation, and dynamic functionality.

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.

Lessons Learned

Telemundo Center represented a unique opportunity to build a large IP plant as a greenfield project. Over the course of design, installation, testing and operational commissioning we development some key learnings and recommendations that we hope will be helpful to other broadcasters seeking to implement IP solutions.

Try to avoid legacy SDI coax wiring practices where possible. This includes such components as jackfields and distribution amplifiers to feed test points, in-rack QC monitoring and other similar uses. Even with the extensive use of SDI-IP gateway devices at Telemundo center, we eliminated the use of jackfields and DAs. While eliminating jackfields and DAs may increase complexity in maintenance troubleshooting, it results in significantly streamlined physical builds in terms of time and wiring complexity. Infrastructural simplification also enables a significant compression of required rack space, and a reduction of passive gear in equipment rooms – which tends to result in “orphaned” power and cooling capacity in modern datacenters.

Develop a strong, well-considered fiberoptic cabling plan. While fiber is not new or unique to IP video builds, IP will require a greatly increased use of fiber relative to SDI builds. Implement stringent standards around fiber cleaning and general fiber cabling management, including training programs for integration teams. All fiber connections should be cleaned, inspected, then cleaned again prior to insertion in devices or bulkheads. Many IP builds suffer from improperly cleaned fiber connections, which may result in data loss and force a re-cleaning process after the nominal conclusion of physical integration. Save time in the build by touching only each fiber connection once and cleaning it properly.

Additionally, consider the strategy for fiber distribution around the plant. As a general rule, fiber distribution should be as simple as the required functionality allows. There should be as few physical fiber connection points as possible between any two linked devices. Perform a cost-benefit analysis on centralized fiber cross-patching relative to direct connections. For large-count strand bundles, consider options around termination and breakout. Unterminated fiber bundles may be easier to pull through conduit, but field termination at scale is a costly, time-consuming and error-prone. Bundles pre-terminated with high-density MPO connectors of 12-or-24 strands may be convenient to break out using cassette-type solutions, however that model introduces an additional connection point contributing to power loss. Consider pre-terminating with simplex or duplex connectors instead for connection to pass-through bulkheads. Complex fiber distribution systems with multiple bulkhead and cross-patch links between devices have numerous points of failure (including human error) and are difficult to troubleshoot. While many of the strategies we discuss here can contribute to plant flexibility and reusable cabling infrastructure, that benefit may not be great enough to overcome the complexity both in build and support.

Complexity in installation phase vs. configuration phase of a plant build. In a typical SDI build, every source is wired through a jackfield, to a DA, then back through the jackfield before landing at its destination. It is typical, then to have 4 cables in path for a single link – with potentially hundreds of such links per rack. Each of these must be engineered, cut to length, terminated, labeled, installed and dressed. In the aggregate, this means that SDI builds are extremely complex to wire. The advantage of this is that every cable carries a single unidirectional stream and point-to-point connection requirements can be decided at the time of design. As a result, commissioning is fairly straightforward. If devices and cabling function as designed, video paths are pre-set, and the overall system powers up and “just works.” Documentation is also

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.

straightforward to generate and to read. Signals on a diagram read left to right and the intended functionality of the system can be understood by reading a drawing.

Consider, by contrast, a typical IP build. In a typical IP build, every device is wired directly to the IP network. As we discussed above, there is no need for jackfields and DAs. And in a well-designed fiber plan, connections are as close as possible to being directly "point-to-point." As a result, and in combination with a good fiber cleaning strategy, physical integration is fairly straightforward. The disadvantage of this is the intended functionality of the system may be difficult or impossible to understand by reading a physical wiring diagram. Wiring will not imply any relationship between endpoint devices, as these relationships in IP move from a serial unidirectional model to a hubbed (via the network) bidirectional model. Device relationships and functional requirements are all defined in software, including signal routing at show time. There is an increased burden on developing control systems, user interfaces, managing IP addresses and device naming. There is no physical connectivity diagram that can show which device output is intended to feed a given device input.

In summary, SDI wiring is complex, but commissioning, troubleshooting and documentation are simple. IP wiring is simple, but commissioning, troubleshooting and documentation are complex. A key takeaway from this observation is that, relative to SDI, IP projects should include an increased timeline between the end of physical integration and the start of production readiness. Another takeaway is that project teams should attempt to begin developing control models, functional use cases, IP addressing and naming schemes, as early as possible in the overall project. This will help assure that workflow intentions are well understood and ready to be implemented promptly during system commissioning.

Finally, IP builds require new kinds of functional documentation to augment wiring documentation – documents explaining not just how the system is wired, but how it is meant to be used. For Telemundo Center, this kind of documentation was enabled by the virtual loopback routing solution we discussed above. These virtual loopbacks are virtual objects in software, however they act like devices with inputs and outputs. That means they can be included on a "virtual signal line" diagram to indicate software function as an analog of physical connectivity. The use of such virtual path diagrams at Telemundo Center has helped significantly to document the plant both for operations and engineering audiences.

The perfect IP network is not a minimum requirement. Broadcasters should not feel they have to wait to implement IP until they can provide total interoperability in native IP on a single non-blocking redundant network for all media types. While total unification may be an industry goal in the long term, for the foreseeable future it is perfectly acceptable to make concessions to an idealized view of IP.

First, vendor-agnostic "native IP" solutions are nearly impossible to achieve currently. Even where products share support for ST 2110 itself, there is no common control standard adopted, either for stream subscription or switching mechanics. This means that broadcasters will likely have three choices available for integrating "third party" devices into their IP plant for the foreseeable future. These options are SDI gateways, NAT solutions for static endpoint stream subscription, and device-specific or proprietary endpoint control APIs. Each of these options comes with its own advantages and disadvantages. For Telemundo Center, we selected an architecture in which most third-party devices are connected with SDI gateways. Gateways provided for us the most frictionless install and commissioning experience, without compromising

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.

any of the functional benefits of IP. As control standards and IP product support progress, we have the option to abandon gateways in favor of direct native connectivity.

Second, the use of diverse IP media networks may be advisable for different use cases and media types. Such diversity may include audio-only networks for audio-only applications as well as separate networks for compressed and uncompressed video. A well-designed and completely flexible IP plant solution may involve a number of different networks without any overall functional loss relative to a perfectly unified network. Consider the example of studio microphone sources – these will not ever need to be routed to production switcher inputs or multiviewer displays, so that capability should not be a requirement of any IP system build. Limitations to a diverse network can be overcome with media conversion gear at network interconnect points, and by wiring devices to multiple networks where they need to send or consume a variety of media types.

Finally, the use of “split” networks, even for a given media type such as ST 2110, may be cost effective and functionally seamless relative to a truly non-blocking network. There may be technical limitations to non-blocking scale for a given network solution, and other considerations (including cost) may make a large non-blocking network impractical. Broadcasters should not assume that non-blocking scale limitations set a hard ceiling on the overall scale of an ST 2110 environment. A key benefit of IP video is that large scale affords elastic production and flexible sharing of resources, however there are real operational affinities to consider in network design. Even in a build as large and ambitious as Telemundo Center, we recognized that there was little benefit to a non-blocking network for the entire ST 2110 environment. As such, we deployed the ST 2110 environment as a pair of non-blocking networks. Transmission, Ingest and Post systems are connected to one network, where non-blocking performance is required. Production systems, including components associated with studios and control rooms, are connected to a separate network. The Acquisition and Production networks are interconnected with a set of managed IP tielines of enough capacity that they are effectively transparent. While the overall system is not truly non-blocking, the functional requirement is well within the blocking capability.

Every IP De-Encapsulation is a unique event. And not all endpoints are guaranteed to perform de-encapsulation in an identical way. Areas where performance may vary are Audio/Video sync, route switching characteristics (Seamless or not), and ST 2022-7 redundancy (Hitless or not). An SDI video stream consists of three key components – Audio, Video and Ancillary Data. These components are interleaved in the serial data stream and can be expected to remain united and synchronized through various devices and infrastructure, including through an SDI router. In ST 2110, these three components are split into multiple different multicast streams. Even for signals that entered the IP network as video and audio combined, they are still demultiplexed in the network and remultiplexed separately at each receiving device. Receiving endpoints typically use PTP time stamps to time-align the streams. However, one device, due either to intrinsic limitations or configuration error, may perform this synchronization differently than another. For a given source, we may see lip sync issues at one output but not another. Such a scenario does not typically occur for multiplexed streams on different outputs of an SDI router.

It is possible to use virtual loopback routing in a “2x1” configuration to create a virtual 2x1 switch. In an SDI router, all downstream consumers of a switched router output would typically perceive the switch event the same way, whether glitchy or seamless. In an IP network, a virtual 2x1 switch does not produce a single switched stream. Instead, it produces a set of instructions for each listening endpoint to disconnect from one stream and reconnect to another. Not all endpoints may perceive the same result of this switch; some may be seamless, some glitchy.

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.

Finally, since different endpoints have their own internal mechanisms for genlock of de-encapsulated streams (such as in an IP to SDI gateway), reference timing for a given source may be measured differently between two different endpoints.

One takeaway from these observations is the need to be mindful about the value of QC in an IP environment. Any encoded data stream—SDI included—requires decoding and interpretation by a receiving device. While IP video adds a new level of capabilities in a broadcast plant, it remains an immature technology in many ways, as evidenced by the variation in de-encapsulation performance across endpoints. Understanding this variation is key to troubleshooting to maximize performance in an IP video environment.

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.

Final Thoughts

As demonstrated by the implementation at Telemundo Center, large scale SMPTE ST 2110 deployments are not only possible, but also provide a level of flexibility and scale unattainable with a traditional SDI broadcast plant. However, as with the adoption of most new technologies, SMPTE ST 2010 raises a number of considerations, namely the fundamental shift from hard-wired connectivity to a system defined by software configuration. As more broadcast engineering teams move toward SMTPE ST 2010, we expect the industry to evolve, filling many of the gaps we identified in this inaugural installation.

Authorized licensed use limited to: SMPTE ALL INCLUSIVE. Downloaded on October 29,2020 at 16:35:55 UTC from IEEE Xplore. Restrictions apply.