Telestream
Whitepaper

An Introduction to IP Video and Precision Time Protocol (PTP)

The broadcast industry is going through tremendous technology changes. Some changes such as 8K/4K/UHD, High Dynamic Range and High Frame Rate video could be regarded as evolutionary, but the transition to an all IP video workflow is regarded by many as a revolutionary and disruptive technology change that will demand entirely new skillsets and infrastructure.

The migration to IP will impact everyone across the broadcast chain to varying degrees, including content producers, broadcasters, content providers, content distributors and equipment manufacturers. However, possibly the biggest impact will be to live production workflows.

Although many see IP Video as new technology, in fact in video distribution workflows, the transition from ASI to IP began over 15 years ago and IT technology began to enable the transition to file-based workflows more than 10 years ago. In fact, production is the last remaining stronghold for SDI and that is set to change. It is reasonable to ask why SDI remains in use in live production workflows. The answer is that the technology works very well, giving outstanding image quality, with extremely low levels of jitter and latency as well as offering an extremely "thin" unidirectional protocol that is easy to deploy and makes frame accurate switching inherently simple. In addition, SDI is an open, non-proprietary and universally supported standard.

So why would we want to move to using IP? The most commonly quoted reasoning is the ability to use Commercially Off-The-Shelf (COTS) IT-based infrastructure, which takes advantage of the economies of scale of the IT industry when compared with the relatively small broadcast industry. In addition, it offers advantages of reducing cabling cost and weight. All this certainly true, but probably the biggest advantage is the much greater routing flexibility offered along with enabling new workflows such as downstream/centralized

production. These new workflows in turn are likely to lead to new types of content and with it new sources of revenue. One aspect of using IP for transporting video that is often overlooked is that scalability is no longer a function of port density, but instead is merely a function of bandwidth.

Having said this, IP does bring some challenges, including jitter; latency; the risk of dropped packets, an inherent lack of synchronicity along with asymmetry which results in different path delays upstream and downstream. However, all the above are surmountable, but it does not change the fact that IP is a complex set of bi-directional protocols requiring a knowledge of both the source and destination before deployment. It is often thought that transporting uncompressed or lightly compressed video is the most difficult application for IP, but in fact it could be argued that trading floors, where time is money demand even greater levels of performance.
Switches intended for trading room applications typically offer latencies less than 250 ns, which offers more than enough performance for IP video applications, where timing accuracy is typically around 1 μs.

The Application of Standards

In general, when we refer to video over IP in the context of any video production workflow, we are referring to the distribution of either baseband or lightly compressed video over Real Time Protocol, commonly referred to as RTP. The advantage of using RTP as opposed to Universal Datagram Protocol (UDP) for the transport layer is twofold. RTP packets are time-stamped making the measurement of packet delay variation easier, but critically the packets also carry a sequence number, making the detection of dropped or out-of-order packets relatively straight forward.

In addition to carrying Video over IP, in a live production environment it is critical to consider synchronization and timing. The asynchronous nature of IP has the advantage that many different traffic types can be carried across a network without having to be concerned with synchronization, but this presents a challenge in the production environment where synchronization is critical to enable frame-accurate switching as well as synchronous video processing. To provide the necessary “genlock”, a precise timing standard is required, for both IP and Ethernet networks that is standardized in IEEE 1588-2008 Precision Time Protocol, commonly referred to as PTP version 2. This is also the basis of a SMPTE PTP standard, specifically intended for the timing and synchronization of video transmitted over RTP networks – defined in two parts SMPTE 2059-1 and 2059-2.

Likewise, there is an AES67-2018 PTP profile for use with audio transmitted over RTP using the AES67 format. The first part of the SMPTE ST 2059 standard refers to “the generation and alignment of interface signals to the SMPTE Epoch” (Date 1970-01-01 Time 00:00:00 TAI) and the second part refers to the definition of a “SMPTE profile for use of IEEE 1588 Precision Time Protocol in professional broadcast applications”. It should be noted that while PTP provides a mechanism to synchronize the real-time clocks of devices on an Ethernet-based network to the same time, it does not make the network itself synchronous.

Coming back to the carriage of Video over IP, there are a number of specific industry standards and proprietary methods for its distribution. Firstly SMPTE ST 2022-6 is a standard designed to transport uncompressed SDI video (Figure 2), embedded audio and metadata over RTP/UDP. This SDI encapsulation process into IP packets has the advantage that video, audio and data will remain synchronized across the network and is ideal for distribution of content. However, it limits the flexibility in having separate streams for just video, audio and data that is needed within a live production workflow. Initially audio was encapsulated into IP packets using AES67 allowing mono, stereo or multi-channel audio to be carried as separate streams and there is a need for separate flows of video and data as well. Hence SMPTE ST 2110 has a suite of standards (Figure 3) that allow for video images to be packetized in (ST2110-20), audio is based on AES67 (ST2110-30) and data (ST2110-40) is based on RFC8331. Each type of video, audio and data stream can be sent separately across the network and allows for easier mixing of the different streams into the final production output. However, this means that timing becomes a critical component of ensuring synchronization across the network of each of these streams using PTP.

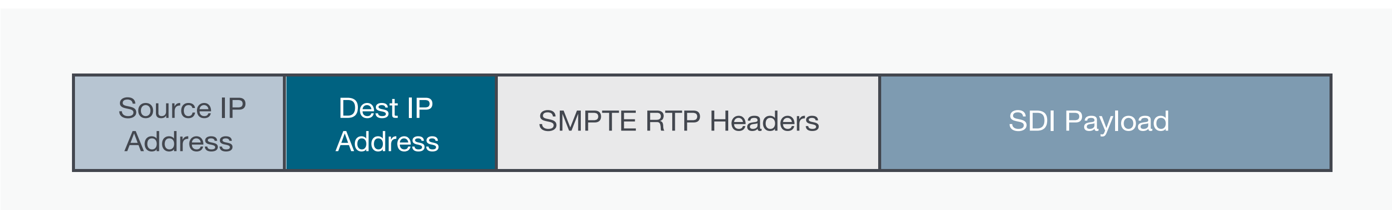


FIGURE 2. SMPTE ST 2022-6 IP Packet Format (A Single IP Flow Can Carry Video, Audio and Metadata encapsulated from SDI).

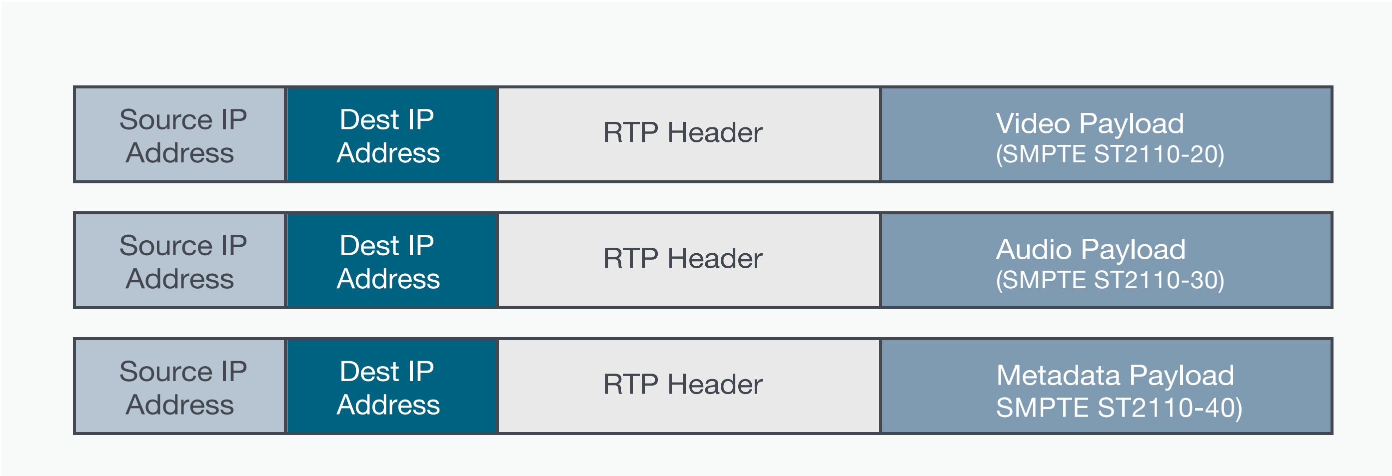


FIGURE 3. SMPTE ST 2110 suite of standards that provides encapsulation of video, audio and data in separate streams.

Within a broadcast network redundancy is a critical part of the infrastructure to ensure transmission of the network. SMPTE 2022-7 Seamless IP Protection Switching is enabled by IGMP multicasts and provides a method to clean switch RTP packets using frame numbers. It can tolerate the complete failure of one network path. With seamless (otherwise known as "hitless") failover, the receiver selects packets from the main or backup streams in order to produce an error-free output, at the cost of doubling required network bandwidth.

SMPTE 2022-7 can be used for both ST2022-6 and ST2110 streams to provide redundancy within the network. The example shown below (Figure 4) shows an error-free output even though the mainstream has suffered a total network failure.

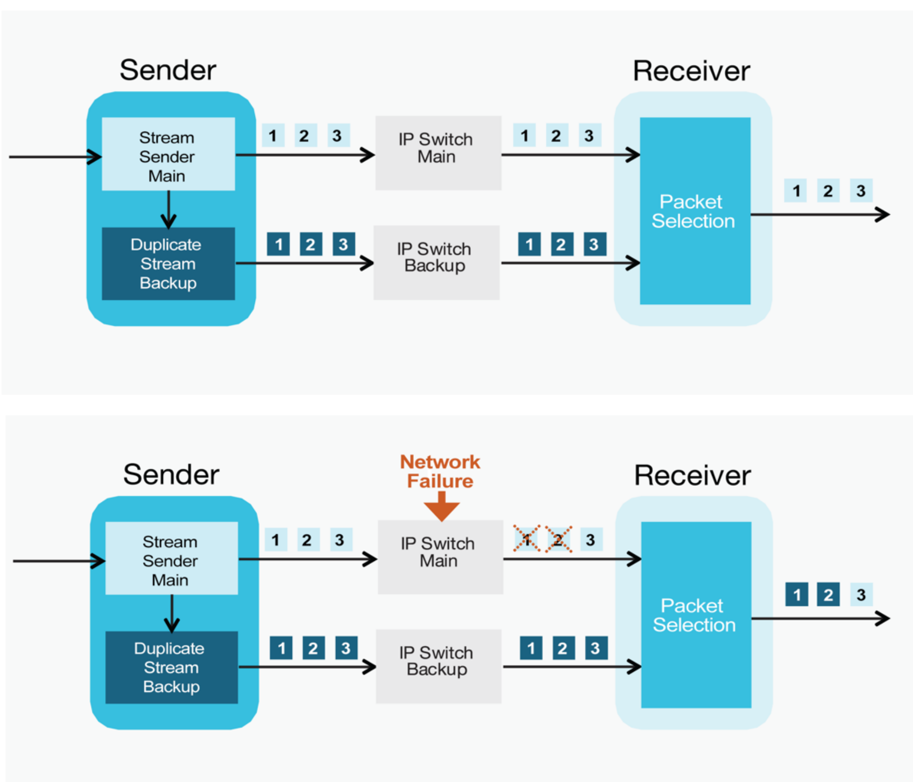


FIGURE 4. SMPTE ST 2022-7 Seamless IP Protection Switching.

To Compress or Not to Compress that is the Question

The practical and affordable deployment of 8K/4K/UHD is likely to lead to the question of the need to compress the content or increase the bandwidth of the pipe. The scalability of IP networks allows for 10G, 25G, 40G, 100G or 400G networks, but at a price and the decision becomes how scalable do you need your network to be or can you lightly compress the content to save on bandwidth. The evolution of SDI from HD to 3G to Quad Link lead to the development of 12G SDI to allow for the production of 4K/UHD content and now 8K will lead to quad link 12G initially. The extensibility of IP media networks means you can fit multiple HD or 3G streams within a 10G infrastructure and you can just fit 3840x2160p50 when using ST2110-20. However, 3840x2160p59/60 will not fit so the next logical extension is to use a 25G infrastructure within the network for these streams.

While 10-bit High Dynamic Range (HDR) has minimal impact on bitrate, the adoption of 12-bit HDR results in an approximately 20% increase in required bandwidth. It is perhaps obvious that High Frame Rate (100/120 fps) requires light compression to be used within the IP media network.

All these new technologies conspire to drive adoption of light compression methods in order to fit ever more data into the IP media pipe. All compression methods are a trade-off between latency, compression ratio and picture quality. In live production applications only low levels of compression (typically 4:1) are required, whilst conversely, latency needs to be low and the picture quality needs to be of the highest order. The compression methods complexity and ease of implementation within standard building blocks needs to be considered. Block transform codecs (MPEG-2, H.264, HEVC etc.) deliver high levels of compression at the expense of high levels of complexity and latency. The wavelet-based codecs deliver lower levels of compression for high quality or lossless applications, but with much lower levels of complexity and associated latency. There are several methods commonly proposed for use in production applications and all are relatively simple and light weight wavelet compression algorithms. They are the Sony Low Latency Video CODEC – LLVC (submitted to SMPTE as RDD 34); VC-2 (Figure 5) – also known as Dirac Pro, which developed by BBC research and development and is standardized as SMPTE ST 2042; the Intopix Tiny CODEC – TICO (submitted to SMPTE as RDD 35); and JPEG-XS (ISO/IEC 21122). All these wavelet CODECs are intra-coded and are designed to deliver extremely high-quality video at low levels of compression and with low latency.

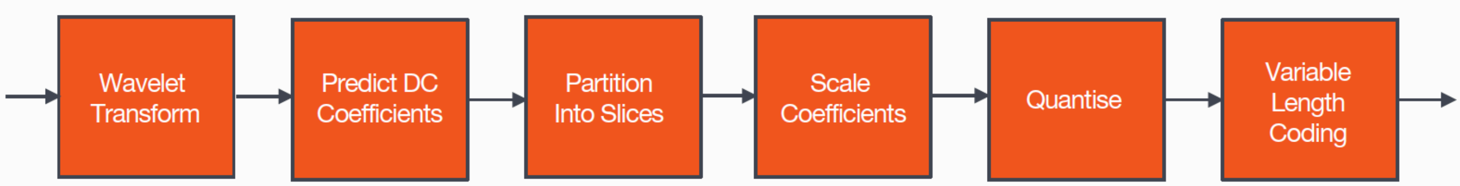


FIGURE 5. VC-2 Signal Processing Chain.

Keeping an eye on IP streams

IP media networks can either use ST2022-6 where the SDI data is encapsulated into IP packets and means that video, embedded audio and data are synchronized together. This makes ST2022-6 ideal for distribution and allows a facility to replace their SDI switching infrastructure with IP switches and a series of SDI to IP gateway cards to make the conversion between SDI to ST2022-6 streams or vice versa. This approach allows simple adoption of an IP infrastructure and a hybrid architecture between SDI and IP. Synchronization relies on primarily on analog references to the gateway cards or the use of PTP within the IP infrastructure.

However, ST2022-6 is not the most efficient use of the bandwidth since it requires the image and blanking to be encapsulated within the IP packets. SMPTE ST2110 suite of standards addresses this by only encapsulating the video image (ST2110-20) within the IP packets and separate streams are used for audio using AES67 (ST2110-30) and ancillary data using RFC8331 (ST2110-40). Carrying separate streams for video, audio and data provides greater flexibility in combining a variety of sources together in a live IP workflow. With this greater flexibility comes complexity in ensuring synchronization of the video, audio and data streams with the reliance on PTP to ensure each stream is referenced to a common clock.

Monitoring of the various streams becomes a critical component of your IP workflow, since you need to monitor traffic flowing across the network ensuring an “error-free” network. To ensure that all senders provide all their packets to the appropriate receiver in a timely manner.

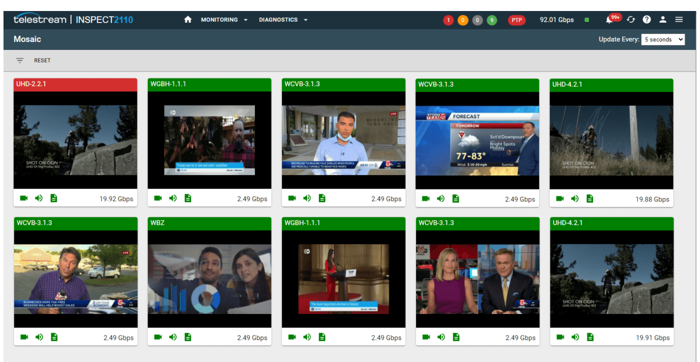


FIGURE 6. IP Video Monitoring using Inspect 2110.

Monitoring can be as simple as decoding the picture and listening to the audio. However, to be able to monitor multiple programs at scale, a different approach is required by monitoring by exception. By focusing on each program to make sure they are present and healthy using Telestream’s Inspect 2110 (Figure 6). Notification can be triggered if the format of the video, audio or data has errors. Allowing an operator or engineer to investigate further the problem using the ‘click to view in PRISM’ the Telestream PRISM waveform monitor can be automatically configured to investigate the issue and can provide a deeper analysis of an individual stream.

When needing to analyze issues PRISM can monitor the RTP sequence and check for lost or out of order packets indicating a traffic problem across the network for a stream.

Timing is critical within an IP media network and PRISM can monitor PTP and provide trend analysis over minutes or hours. Additionally, check that the source is locked to PTP can be done within the Timing display and analyzed within the Stream Timing display to compare video to PTP or RTP offsets and for audio the video to audio or audio to RTP offset can be monitored.

FIGURE 7. PRISM monitoring 2110 streams showing Stream Timing, IP Status and TIming displays.

Keeping PTP Simple

The adoption of video over IP along with the use of PTP to synchronize the real-time clocks of different network nodes infers that any such network requires a network time server, in order to provide the PTP genlock functionality equivalent to that delivered by a Sync Pulse Generator (SPG) in SDI networks. Any logical grouping of clocks that are synchronized together are referred to as a PTP domain. It should be noted that clock in one domain may not be synchronized to clocks in another domain.

This PTP network time server is generally referred to as a PTP Grandmaster, with a device that derives its timing synchronization from PTP being referred to as a PTP Follower. A Leader clock is a device that provides the time in a given PTP domain and a Follower is a device that synchronizes to a Leader. A Grandmaster is a Leader that is providing the ultimate source of clock synchronization in a network. In the context of broadcast applications, PTP Grandmasters are usually synchronized to GPS, GLONASS or both, in order to derive accurate timecode relative to the 1970 Epoch. It should be noted that PTP Grandmasters always use the 1970 Epoch. To enable legacy equipment support, the Tektronix SPG8000A hybrid PTP Grandmaster and SDI SPG is able to phase its baseband timing outputs relative to either the 1970 or 1958 Epoch dates.

Within any PTP domain there are a number of message types used to establish time within that network. Announce messages are used to establish the synchronization hierarchy and provide the clock status and clock criteria used to determine which clock becomes the Grandmaster. Sync and Follow-up messages are transmitted by the Grandmaster and are used by Followers to derive the time.

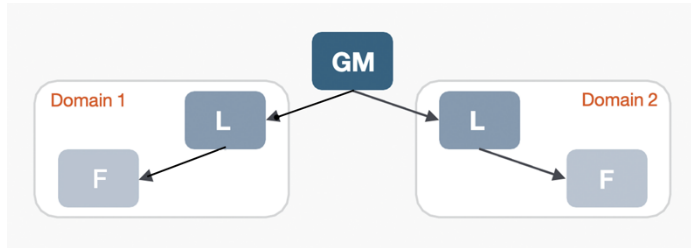
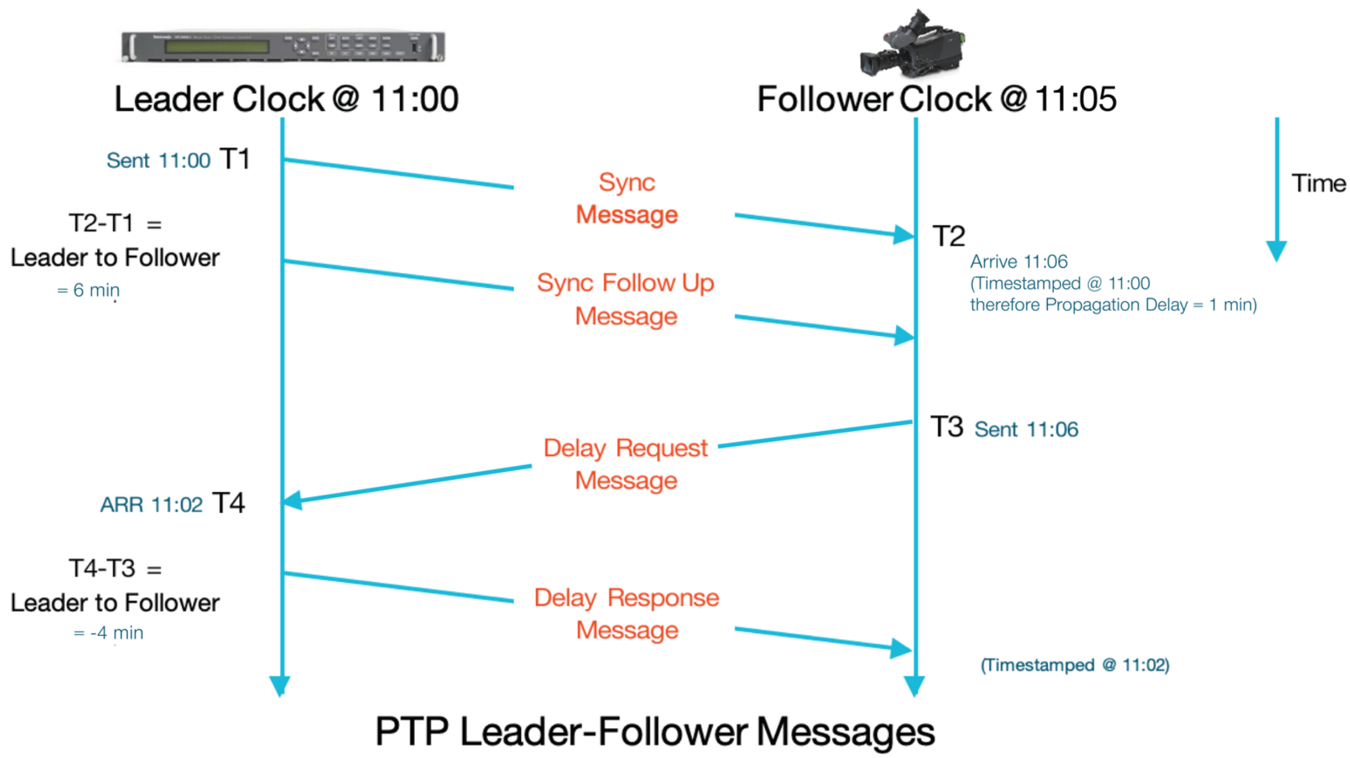


FIGURE 8. PTP Domains Synchronized to a Common Grandmaster.

Delay Request messages are a request for timing information and are sent from the Follower to the Grandmaster in order to determine the reverse path propagation delay between the Follower and the Grandmaster. A Delay Response message is sent by the Grandmaster and contains the time of receipt of the Delay Request message by the Grandmaster.

As defined, PTP is a method for distributing time over a network, with a single Grandmaster providing the source of time, to synchronize one or more Followers. The Grandmaster periodically transmits Sync and Follow-up messages, which the Followers use to derive the time. In an ideal World the network delay could be programmed into each Follower which could then be offset to the time in the received packet to derive the correct time. Such symmetry can only be relied upon in point-to-point IP links. Unfortunately, the delay in switched / routed IP networks is both variable and asymmetric, so the Follower devices must periodically send Delay Request messages to the Grandmaster. The Grandmaster accurately time stamps these messages on receipt and the time of receipt is sent back to the Follower in a Delay Response message.



Offset = (Leader to Follower $_t$ – Follower to Leader $_t$)/2 **= 5 Mins**
Oneway Delay = (Leader to Follower$_t$ + Follower to Leader$_t$)/2 **= 1 Min**

FIGURE 9. Deriving the Correct Time in a PTP Network.

Using the diagram (Figure 9) above as a reference, the Follower is now able to calculate the difference between its own clock and that of the Grandmaster using the Leader-to-Follower sync packet delay (T2- T1) and Follower-to-Master delay request packet-delay (T4-T3). The Offset (Follower Time – Leader Time) = [(T2-T1)-(T4-T3)]/2 and the Oneway delay = [(T2-T1)+(T4-T3)]/2. For the Follower time to be now correct, the propagation delay in both directions must be equal.

If the propagation delay in both directions is in fact different, then the Follower is offset to “correct” for this by adjusting its clock to a value of half the asymmetry. The clock’s control loop adjusts the Follower time to make the Leader-to-Follower and Follower-to-Leader propagation delays appear to be equal. That is, the control loop adjusts the Follower time such that T2-T1 = T4-T3.

Accuracy and Reliability is Key – The BMCA

One reason for PTP's suitability to broadcast applications is the resilience provided by the use of the Best Master (Leader) Clock Algorithm (BMCA). The BMCA allows the most accurate Leader to automatically take over the duties of Grandmaster when the previous Grandmaster loses its GPS lock, gets disconnected from the network, or is unable to act as Grandmaster for any reason.

The BMC Algorithm runs on all clocks in a network and uses a number of criteria to determine which Leader should be Grandmaster including the following in priority order:

1. User Definable Priority 1 Field (the lowest value <= 128 wins)
2. Clock Class (e.g. GPS vs free running)
3. Clock Accuracy
4. Clock Variance (jitter and wander)
5. User Definable Priority 2 Field (the lowest value <= 128 wins)
6. Clock Source Port ID (usually the Ethernet MAC Address)

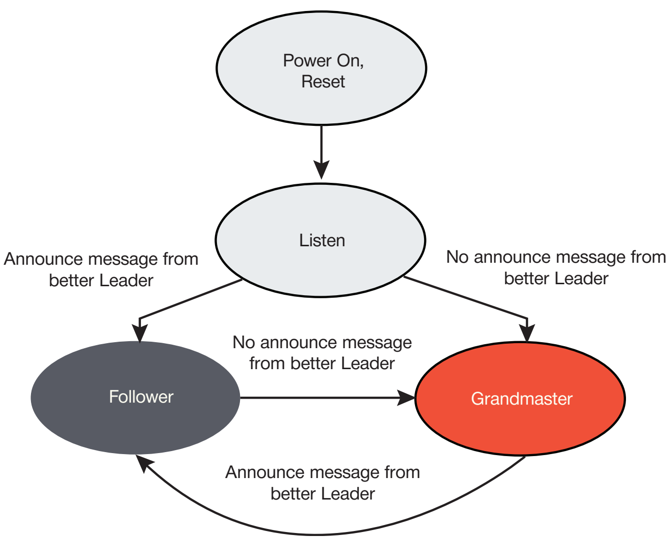


FIGURE 10. Determining Leader/Follower Clock State.

Grandmaster Failover

In order to establish an automatic main and backup Grandmaster fail over the Priority 2 field is used to identify main and backup clocks between two or more otherwise identical redundant Grandmasters as follows:

- Main Grandmaster
  (Priority Field 1 = 1; Priority Field 2 = 10)
- Backup Grandmaster
  (Priority Field 1 =1; Priority Field 2 = 11)

If both identical Leader clocks are locked to GPS, they will have the same clock quality, so the lowest Priority Two Field value will select which is the Grandmaster. If the Main clock loses GPS lock, then the Backup clock becomes the Better Leader and will take over as Grandmaster.

It is worth noting that if any GPS synchronized Leader clock loses GPS lock, it will of-course itself become free running and will be reliant upon its own internal local oscillator. However good this oscillator is, over an extended period of time it will drift, even if slightly relative to the GPS clock. Once GPS lock is re-acquired, unless the Leader's clock local oscillator phase-lock loop (PLL) is driven slowly to re-synchronize with the GPS clock, then the system can suffer from what is known as "Sync Shock" when the Leader's clock frequency suddenly changes. Whilst this may be acceptable in some IT applications, this is of course highly undesirable in a video production application. In the case of the SPG8000A, the "Stay Genlock" feature is designed specifically to avoid the problem of Sync Shock through careful control of the PLL.

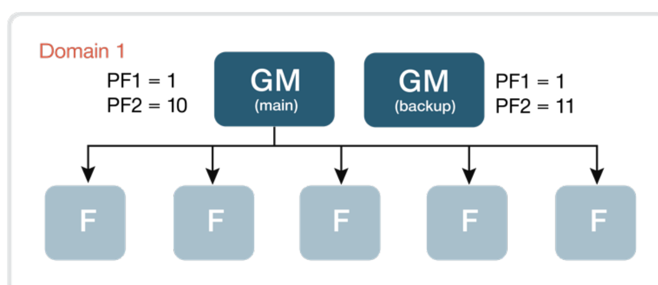


FIGURE 11. Configuration of Main/Backup Grandmasters for Automatic Failover.

Although in theory it is possible to use a Leader with software-based time-stamping, in the case of live video production applications, it is highly unlikely that such a device could be devised with the necessary clock accuracy required for synchronous video processing. A hardware time-stamped Grandmaster device such as the SPG8000A (Figure 12) is locked to GPS (or GLONASS or both to provide greater constellation resilience), with the Grandmaster's local oscillator being phase-locked to the GPS reference. This local oscillator is the reference clock used with dedicated hardware for the precise timestamp of the incoming PTP messages and PTP sync packets. A dedicated hardware approach is unaffected by operating system behavior or network traffic latency.

PTP Clock Types

Ordinary Clocks are those devices that are at either end of a network and are not switches or routers. A Follower Only clock never acts as a Leader, whereas a Leader/Follower clock can act as either and a Preferred Grandmaster is configured to never become Follower.

It is vital that switches and routers in any IP video network that relies upon PTP for synchronization are "PTP Aware". That is they are able to account for their own queuing delay, to ensure downstream timing accuracy. This can be achieved in one of two ways. The first is by the switch acting as a Transparent Clock which hardware time stamps Sync and Delay Request messages on arrival and departure and adds the difference to a correction field in the message.

The second way for a switch or router to account for its own queuing delay is to act as a Boundary Clock, which receives time from a Leader on one Follower port and provides one or more Leader (not Grandmaster) ports to downstream Followers in a PTP Domain and in doing so, removes the effect of its own queue.

Transport	Type
Reserved	Version
Length	
Domain	
Reserved	
Flags	
Correction Field	
Reserved	
Source Port	
Sequence ID	
Control	
Log	
Time Stamp	

FIGURE 12. SPG8000A Master Sync and Clock Reference Generator.

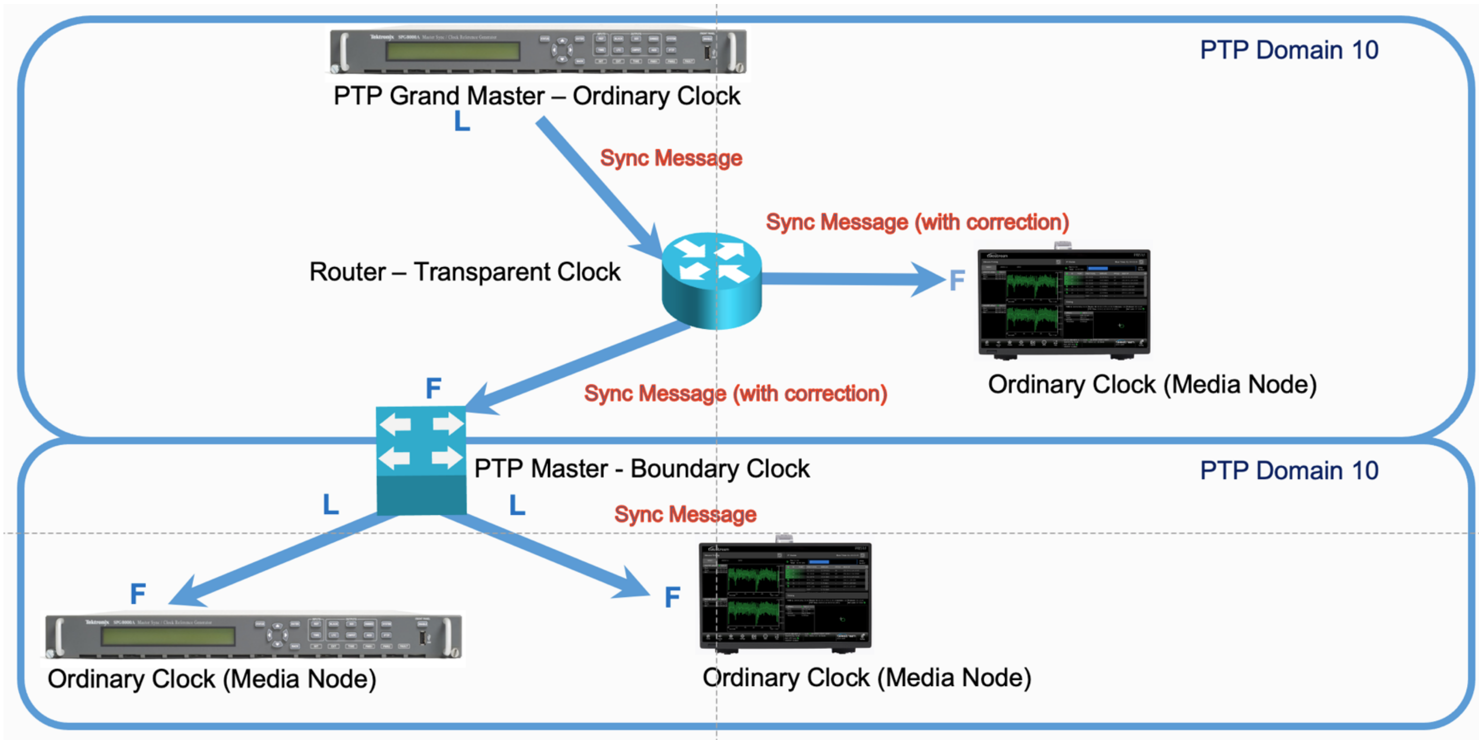


FIGURE 13. PTP Clock Types in a Network.

Summary

Although the transition to IP is seen by some as inevitable, not all equipment is available with IP interfaces. It must also be considered that the investment in SDI equipment has been so considerable, that the use of hybrid IP/SDI networks is likely for the foreseeable future. For broadcast applications, it is essential that the PTP Grandmaster such as the Telestream SPG8000A provides support for the application specific video and audio PTP profiles, such as SMPTE 2059 and AES67, as well traditional SPG features including black burst, tri-level and SDI out. All the above protocols must be referenced to the same GPS clock, or such a hybrid IP/SDI network would be inoperable. It must also be considered that a broadcast live production network is entirely reliant on a stable reference and any timing and synchronization devices “must work”.

Although the concept of carrying uncompressed (or lightly compressed) video over IP is perceived as being very new, and indeed revolutionary, the precedent for the broad adoption of IT infrastructure for live production facilities has in fact been in place for many years. As was mentioned at the beginning of this paper, IT infrastructure began to be adopted over fifteen years ago for compressed video distribution using MPEG-2 transport streams over IP. IT infrastructure is also in industry-wide use as the distribution and control component for file-based workflows. In both these cases, Telestream provided technology support for early adopters; with TS over IP test equipment as well as with the first file-based QC tool on the market.

As such, Telestream was involved with the earliest adopters of compressed video over IP and file-based QC and continues to be closely engaged with these latest developments with baseband video over IP. This has led to the development of Inspect 2110 to monitor multiple programs within an IP media network that can work in conjunction with the Telestream PRISM for deeper analysis of video waveform, audio, data and PTP.

We are at the beginning of a long-term transition to IT-based infrastructure and those involved in the production and facility side of video have little experience with the new technology, but conversely are extremely experienced using SDI and all the issues associated with its use. This coupled with a huge investment in existing technology and workflows implies that the transition will take place gradually, making it likely that hybrid SDI/IP infrastructure will be in place for some years. Such production facilities will require equipment that is able to operate seamlessly and reliably in such a hybrid environment. The companies best placed to provide equipment that meets those requirements are those who have experience both of the challenges of the live production environment, as well as extensive experience of the challenges associated with the distribution of video over IP networks.

telestream

www.telestream.net | info@telestream.net | tel +1 530 470 1300