White Paper

Intel® Media Transport Library

Open Source Library Enables Real-Time Media over IP Networks

Built upon the SMPTE ST 2110 standard, Intel® Media Transport Library supports the industry transition from proprietary to open architectures, delivering real-time media over Internet Protocol (IP) connections with high performance, reliability and efficiency.

Authors

Intel
Ping Yu,
Jianxin Du,
Ming Li,
Jing Wang,
Raul Diaz

With growth in bandwidth for both local area networks and internet gateways, media is increasingly carried on IP networks for applications that include video production, video communications, visual analytics, AR/VR, immersive 360 video, digital signage, industrial usages and the emerging metaverse. Compressed video and audio have been used for media distribution for quite some time, but other application areas have either been entirely unable to work with compressed media or able to use only modest levels of compression to preserve video quality for both human viewing and AI analysis.

Live video production, industrial applications and digital signage (among others) require strict time references and media synchronization between multiple streams. The Intel Media Transport Library supports those requirements for media communications between server nodes while providing the following benefits:

- High throughput.
- Low latency.
- High reliability.
- Low compute overhead.
- Network traffic management.
- Synchronized, real-time media transport.

To achieve strict requirements around flexibility, scalability and cost, the open source Intel Media Transport Library implements the protocols in the Society of Motion Picture and Television Engineers (SMPTE) ST 2110 suite of standards.

Starting in 1989, SMPTE introduced a family of standards collectively referred to as Serial Digital Interface (SDI), a synchronous networking protocol designed to carry media in which each stream consists of precisely timed samples of interleaved audio, video and ancillary data conforming to specific video resolutions. SDI uses a fixed clock frequency for each datum placed on the network, so it is referred to as an isochronous network, from the Greek ισο (iso) "equal" + χρόνος (chronos) "time." SDI has dedicated video format synchronization packets known as the Timing Reference Signal (TRS) with flags in those packets marking the start and end of each video line and video frame. TRS allows multiple SDI streams to be precisely synchronized.

In 2007, SMPTE published a family of standards called SMPTE ST 2022 that define how to send digital video over an IP network. ST 2022-6 mimicked the SDI format of interleaved audio, video and ancillary data streams using the Real-time Transport Protocol (RTP), which is transported by the User Datagram Protocol (UDP), which in turn is encapsulated in the IP datagram. In 2017, SMPTE

Table of Contents

significantly expanded the capabilities and flexibility of digital video transmission over IP with a new standard called ST 2110. While both ST 2022 and ST 2110 support the transport of synchronized audio, video and ancillary data, ST 2110 separates audio, video and ancillary data into separate "essence" streams in which each stream of packets only carries one type of media. For example, a video essence stream contains only video data, and a separate audio essence stream would contain the corresponding audio. The two essence streams can be transmitted and processed separately, even in separate servers, greatly simplifying processing and enabling scalability. Since IP networks are asynchronous, the IEEE 1588 standard entitled Precision Time Protocol (PTP) is used to timestamp every ST 2110 packet. The PTP timestamps are used to synchronize the multiple essence streams as they travel through the network and through servers.

In addition to PTP for timestamps, the ST 2110-21 standard enforces a precise packet pacing that not only prevents network bottlenecks that can lead to packet loss, but also simulates the TRS in SDI, allowing video equipment that requires precise audio and video timing to interface cleanly with IP networked servers.

IP networks are called "best effort" networks since they do not guarantee the delivery of packets. Normally, this is not a significant problem because packets that are lost during transmission can be detected, and a request can be made to re-send those missing packets. However, in many applications including live video production, industrial, AR/VR and others, there is not sufficient time to request missing packets because the requests themselves would cause delays (latency) that would prevent the media from being delivered in real-time. Because of this, SMPTE created a standard called ST 2022-7, "Seamless Protection Switching of SMPTE ST 2022 IP Datagrams." ST 2110 also uses ST 2022-7, which allows an essence stream to be duplicated and sent through separate network switches and Ethernet network adapters, creating redundant paths. If the primary stream should lose one or more packets or have them delayed, then the secondary stream can be used to recover the packets.

The Intel Media Transport Library Implementation

Intel Media Transport Library has been designed to facilitate ST 2110 network traffic that flows between servers based on Intel® Xeon® Scalable processors using Intel® Ethernet Network Adapters. The library efficiently transfers ST 2110 media and data between the processor and a network adapter such as the Intel® Ethernet Network Adapter E810, enabling many ST 2110 sessions to maintain consistent, simultaneous high throughput without interruption.

To maintain the high packet processing throughput required by these applications, the Intel Media Transport Library relies on an efficient asynchronous tasklet architecture that allows multiple ST 2110 sessions to run in parallel without impacting each other. The tasklet architecture manages the media and data buffer transfers between the processor and the NIC. The Intel Media Transport Library implements the following SMPTE standards:

- ST 2110-10 System and Timing
- ST 2110-20 Uncompressed Video
- ST 2110-21 Video Stream Packet Pacing
- ST 2110-22 Compressed Video
- ST 2110-30 AES67 Uncompressed Audio
- ST 2110-31 AES3 Compressed Audio
- ST 2110-40 Ancillary Data
- ST 2022-7 Seamless Packet Switching.
- ST 2059 Synchronize Video Equipment over an IP Network

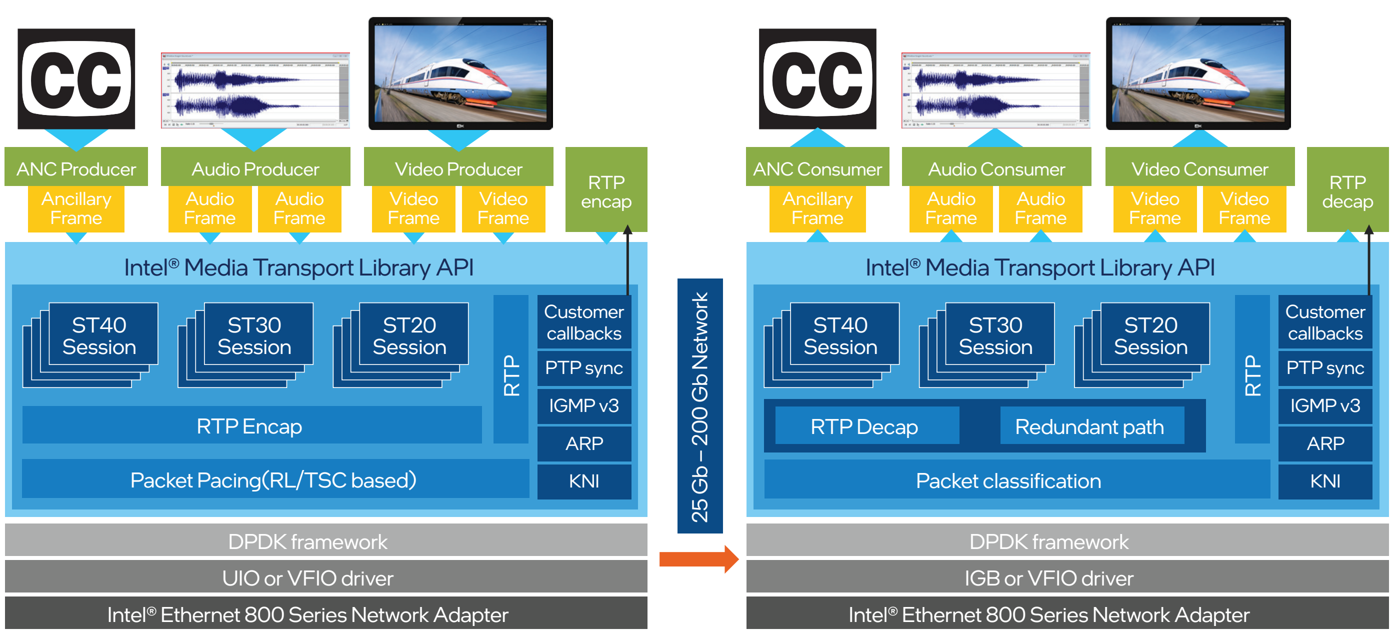

Figure 1. Intel® Media Transport Library architecture.

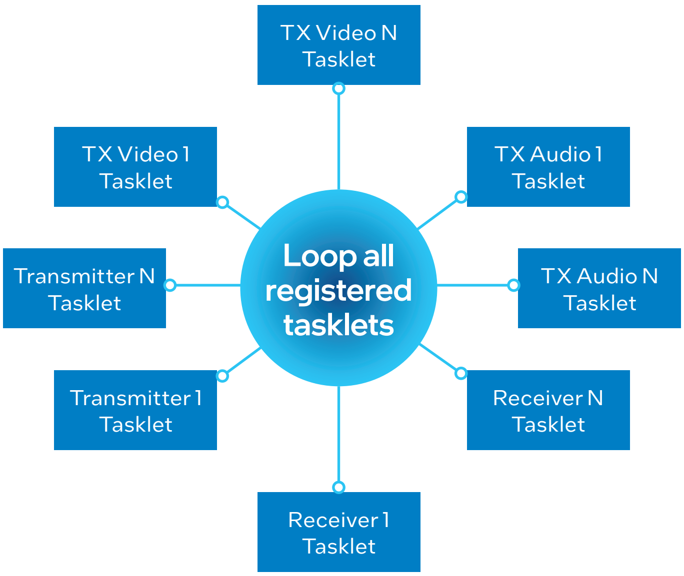

Figure 2. Intel® Media Transport Library tasklets.

The ST 2110 standards refer to many other SMPTE and non-SMPTE standards. Refer to the Appendix for a list of the standards used or referenced in ST 2110. The Intel Media Transport Library runs on Linux and Windows with support for bare metal, containers and virtual machines (VMs). The library implements a sophisticated architecture using specialized optimizations to conform with ST 2110's rigorous requirements.

In Figure 1, the media and data flow are shown with a transmitter on the left and a receiver on the right. Each session for each media and data type is a tasklet that manages the transfer from the application buffers to the network adapter and vice versa.

Tasklets include transmitter (Tx) video sessions, audio sessions, receiver (Rx) sessions, NIC transmitters, NIC receivers and others. Tasklets are managed in a pool and cycled in a loop through all the registered tasklets as shown in Figure 2.

The Data Plane Development Kit (DPDK)

To support the high bandwidth and low latency needed by the ST 2110 network traffic, the Intel Media Transport Library is built using DPDK, which is a set of user-space libraries and drivers that accelerate packet processing.

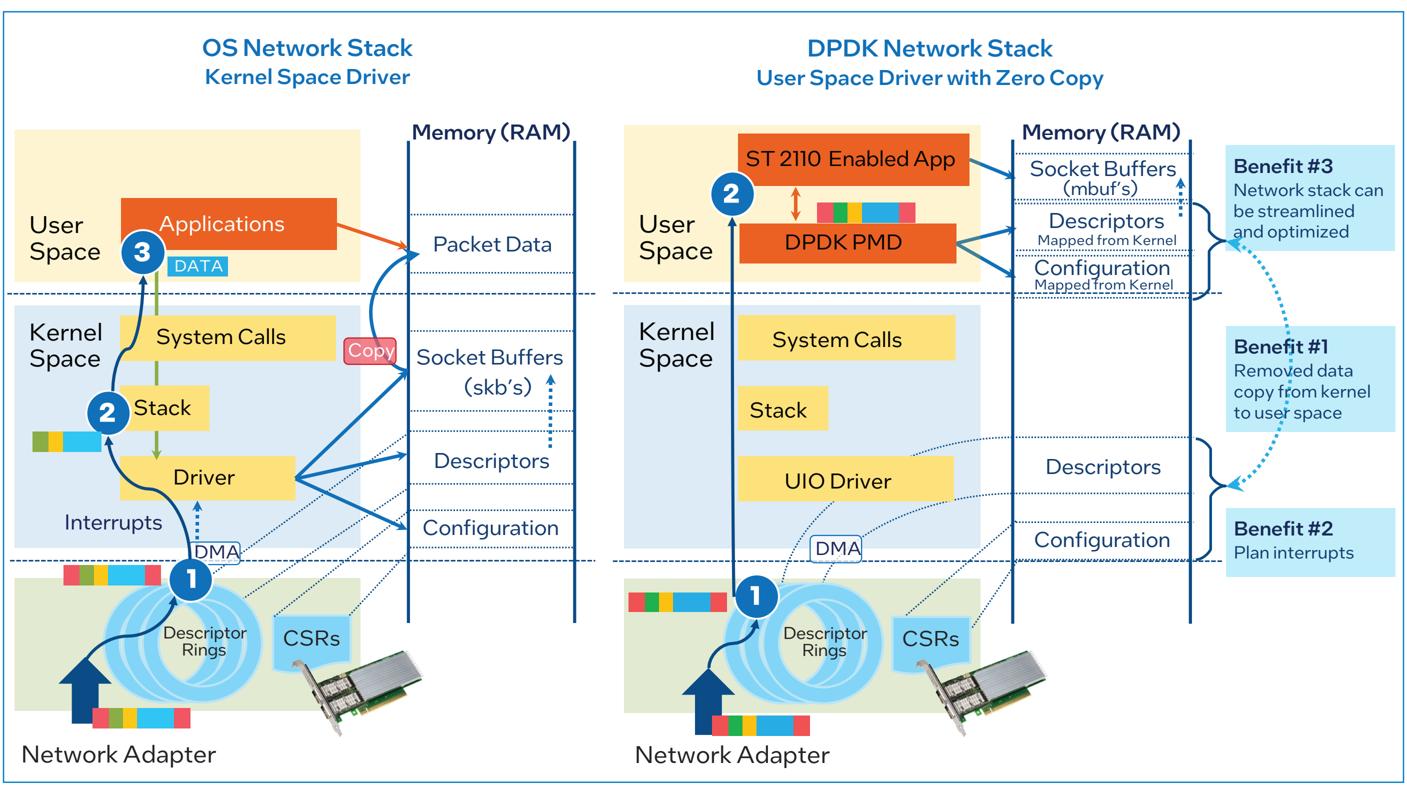

Figure 3. DPDK transfers payloads directly from user space to network adapter.

By bypassing the kernel-space network stack, DPDK allows the Intel Media Transport Library to reduce or eliminate memory copies of the ST 2110 data buffers as they are sent to or received from the network adapter. In Figure 3, the left side shows how data is sent to the network adapter using the operating system's kernel-space driver. The payload data for the packets is placed into a buffer which is copied from the user space to the kernel space, at which point that data is passed to the kernel. On the right of the figure, the DPDK stack bypasses the copy of the buffer into the kernel-space driver. The kernel-space driver is used solely to configure the handoff of the user-space data to the network adapter.

ST 2110-10 System and Timing

ST 2110-10 defines the configuration, packet format and system timing model for ST 2110 networking. The Intel Media Transport Library implements both unicast and multicast networking modes with support for IGMP2 and IGMP3.

IEEE 1588 Precision Time Protocol (PTP) plays a critical role to synchronize ST 2110 essence streams. Each device in the network maintains its own internal clock. These clocks are synchronized across the network by sending PTP packets. A grandmaster (GM) clock is normally used as the reference clock generating the PTP packets for all other devices. Intel platforms beginning with 2nd Gen Intel Xeon Scalable processors support hardware synchronization with the Intel Ethernet Network Adapter's internal clock. The improved accuracy of this synchronization enables conformance with ST 2110-10 requirements for packet timestamp accuracy.

ST 2110-20 Uncompressed Video

Uncompressed video is used extensively for video production applications and increasingly in other application areas that require pixel-perfect video transmission and processing. For example, since color keys are used extensively in video production, uncompressed video enables pixel-perfect image compositing which is required to avoid overlay and transition artifacts.

The huge bandwidth needed to transport uncompressed video presents significant challenges to IP networks. Consider the common example of a single 1080p HD video stream with a resolution of 1080 lines and 1920 columns, high dynamic range (HDR) 10-bit-color pixels, a nominal 60 video frames per second and a 4:2:2 chroma video sub-sampling format. After being formatted into ST 2110 RTP packets, this video stream consumes nearly 2.5 gigabits per second (Gbps). The network adapter can transmit and receive 100 Gbps simultaneously, so it could transmit and receive up to 40 of these HD video streams, although in practice, other non-ST 2110 network traffic and certain system and application limitations generally limit this theoretical maximum.

ST 2110-21 Video Stream Packet Pacing

For large numbers of video streams to flow successfully in a cluster of servers and to send traffic through IP switches, the high-bandwidth video traffic must be controlled so that bursts do not overwhelm the switch buffers and cause packet loss. ST 20110-21 defines transmission traffic shaping models based on a leaky bucket algorithm. The network compatibility model defines the rate at which the sender shall fill the leaky buffer, and the virtual receiver buffer model defines the rate at which the receiver shall empty the leaky buffer. Implementing ST 2011-21's strict pacing models generally requires hardware mechanisms to maintain the precise timing of the packets.

For ST 2110-21 packet pacing, the Intel Media Transport Library relies on the network adapter's rate-limiter technology using a special software optimization in the library to implement ST 2110's traffic pacing.

ST 2110-22 Compressed Video

Although some video workflows require uncompressed video, many others benefit from video compression to reduce network traffic and increase the number of streams that can be processed and transmitted. For example, video transmitted over the internet, video from specialized compressed video cameras and playout servers that contain compressed video streams that are scheduled to be broadcast or streamed all require support for the transport of compressed video.

The Intel Media Transport Library has implemented ST 2110-22 to support compressed video streams. Any video codec format can be supported, including AVC H.264, HEVC H.265, AV1 and the new low-latency, high-quality JPEG XS codec, which is commonly used in video production.

Several different pipelines can be configured using a combination of ST 2110-20 and ST 2110-22, as shown in Figure 4 using the JPEG XS codec as an example. The "Media Processing" block represents the applications that are processing the video essence streams.

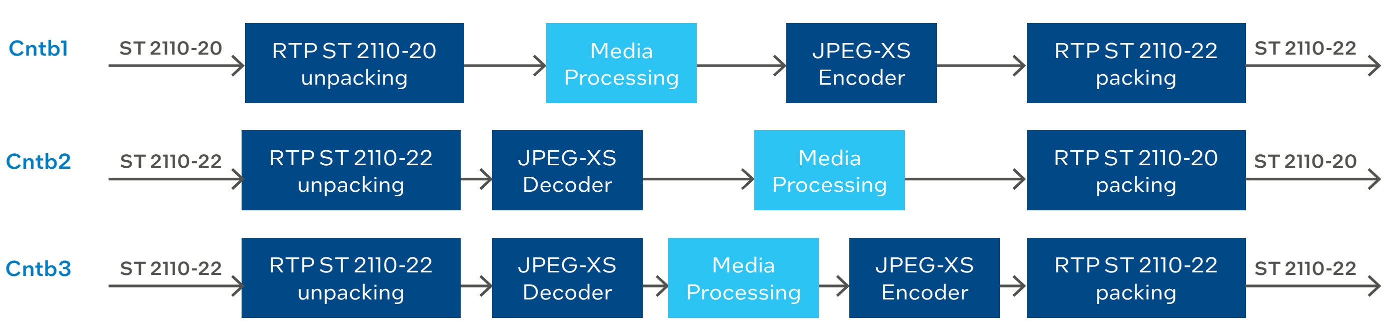

Figure 4. Example contribution transcoding pipelines for ST 2110 workloads.

ST 2110-30 Uncompressed Audio and ST 2110-40 Ancillary Data

Audio and ancillary data generally requires much lower bandwidth than video. However, the ST 2110-30 audio standard has strict timing to make sure that audio and video remain synchronized. ST 2110 specifies that the PTP standard be used to timestamp packets and a grandmaster clock be used as a reference to synchronize time across the network devices. The library uses the dedicated hardware registers in Intel Xeon Scalable processors and in the Intel Ethernet 800 Series Network Adapters to maintain tight synchronization between the server internal clock and the network adapters's internal clock, which is automatically synchronized with the grandmaster clock.

The ST 2110-40 ancillary data packets are similarly timestamped so that data such as closed captions stays in sync with the audio and the video.

Intel Media Transport Library performance testing

Testing measures the number of cores required to maintain 200 Gbps network bandwidth using 54 streams of 1080p video content at 60 fps on 4th Gen Intel Xeon Scalable processors, as shown in Table 1. Intel® Data Streaming Accelerator (Intel® DSA) is a high-performance data copy and transformation accelerator that is native to this hardware platform. It reduces the number of cores required by the library to process ST 2110 traffic.

Intel DSA optimizes streaming data movement and transformation operations common with applications for high-performance storage, networking, persistent memory and various data processing applications. For more details on Intel Media Transport Library on 4th Gen Intel Xeon Scalable processors, see https://www.intel.com/content/www/us/en/products/docs/processors/xeon-accelerated/network/dsa-solution-brief.html.

Open Source

The Intel Media Transport Library has been open sourced to simplify integration into customer applications. The APIs for media transport are clearly defined, and their implementations are available should developers wish to add specific functionality. For example, a developer might want to parse the header buffer on receive or transmit to modify timestamps, or they may wish to add support for custom video formats and layouts. The library repository is located at: https://github.com/OpenVisualCloud/Media-Transport-Library.

Conclusion

The Intel Media Transport Library has been designed to support a broad range of media applications that need to transport and synchronize media and other data from multiple sources such as cameras and microphones, and across multiple servers. The library uses DPDK to enable high-throughput, low-latency transport and relies on hardware support in Intel Xeon Scalable processors and Intel Ethernet 800 Series Network Adapters to implement the strict timing and packet pacing required by the ST 2110 standard. The Intel Media Transport Library is designed to provide the fundamental ingredients to support varied media industries through an open and generic solution enabling media transport, including meeting the demands of increased video resolution with 4K and 8K.

More Information

Intel Media Transport Library: github.com/OpenVisualCloud/Media-Transport-Library

Intel Xeon Scalable processors: intel.com/xeonscalable

Intel Ethernet Technology: intel.com/content/www/us/en/architecture-and-technology/ethernet.html

Data Plane Development Kit (DPDK): dpdk.org

Society of Motion Picture and Television Engineers (SMPTE): smpte.org

IEEE 1588 Precision Time Protocol (PTP) standard: standards.ieee.org/ieee/1588/6825

Table 1. Intel® Media Transport Library Rx Performance on 4th Gen Intel® Xeon® Scalable processors.

	Number of Network Adapters	Number of Cores per Network Adapter	1080p@60 # of Sessions (max/core)	Total Max Sessions
4th Generation Intel® Xeon® Scalable Processor	2	1	27	54(2*27)

Appendix

List of SMPTE ST 2110 standards and standards referenced:

1. SMPTE
    1.1. ST 2110
        1.1.1. ST 2110-10 2017 Professional Media Over Managed IP Networks - System Timing and Definitions ST 2110-10-2017
        1.1.2. ST 2110-20-2017 Professional Media Over Managed IP Networks - Uncompressed Active Video
        1.1.3. ST 2110-21 2017 Professional Media Over Managed IP Networks Traffic Shaping and Delivery Timing for Video
        1.1.4. ST 2110-22 2019 Professional Media over managed IP Networks Constant Bit Rate Video ST 2110-22-2019
        1.1.5. ST 2110-30-2017 Professional Media Over Managed IP Networks - PCM Digital Audio
        1.1.6. ST 2110-31 2018 Professional Media Over Managed IP Networks AES3 Transparent Transport ST 2110-31-2018
        1.1.7. ST 2110-40 2018 Professional Media Over Managed IP Networks SMPTE ST 291-1 Ancillary Data ST 2110-40-2018
    1.2. ST 2022
        1.2.1. ST 2022-6-2012 Transport of High Bit Rate Media Signals over IP Networks
        1.2.2. ST 2022-7-2013 Seamless Protection Switching of SMPTE ST 2022 IP Datagrams
        1.2.3. ST 2022-8-2019 Professional Media Over Managed IP Networks: Timing of ST 2022-6 Streams in ST 2110-10 Systems
    1.3. ST 2059
        1.3.1. ST 2059-1-2015 Generation and Alignment of Interface Signals
        1.3.2. ST 2059-2-2015 SMPTE Profile for Use of IEEE-1588 Precision Time Protocol in Professional Broadcast Applications
    1.4. SMPTE Audio and Ancillary Data
        1.4.1. ST 272:2004 Formatting AES Audio and Auxiliary Data into Digital Video Ancillary Data Space
        1.4.2. ST 291-1 Ancillary Data ST 2110-40-2018
        1.4.3. ST 299-1:2009 24-Bit Digital Audio Format for SMPTE 292 Bit-Serial Interface
        1.4.4. ST 299-2:2010 Extension of the 24-Bit Digital Audio Format to 32 Channels for 3 Gb/s Bit-Serial Interfaces
    1.5. Miscellaneous
        1.5.1. RP 157:2012 Key and Alpha Signals
        1.5.2. ST 428-1:2006 D-Cinema Distribution Master — Image Characteristics
        1.5.3. ST 2065-1:2012 Academy Color Encoding Specification (ACES)
        1.5.4. ST 2065-3:2012 Academy Density Exchange Encoding (ADX) — Encoding Academy Printing Density (APD) Values
        1.5.5. RP 2077:2013 Full-Range Image Mapping
2. AES67 (Audio)
    2.1. AES67:2015, AES standard for audio applications of networks - High-performance streaming audio-over-IP interoperability
3. IEEE (Clock Synchronization)
    3.1. IEEE 1588_2008-2002 IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems
4. IETF (SDP, IP, UDP, RTP, Timing)
    4.1. RFC 768 User Datagram Protocol
    4.2. RFC 791 Internet Protocol
    4.3. RFC 2460 Internet Protocol, Version 6 (IPv6)
    4.4. RFC 3376 Internet Group Management Protocol

    4.5. RFC 4566 SDP: Session Description Protocol
    4.6. RFC 3550 2003 RTP - A Transport Protocol for Real-Time Applications
    4.7. RFC 3551 RTP Profile for Audio and Video Conferences with Minimal Control
    4.8. RFC 4175 RTP Payload Format for Uncompressed Video
    4.9. RFC 4566 SDP: Session Description Protocol
    4.10. RFC 5285 A General Mechanism for RTP Header Extension
    4.11. RFC 7104 Duplication Grouping Semantics in the Session Description Protocol
    4.12. RFC 7273 RTP Clock Source Signaling

5. ISO
    5.1. ISO 11664-1:2007 Colorimetry — Part 1: CIE standard colorimetric observers

6. ITU-R BT Series (Broadcasting Service - Television)
    6.1. ITU-R BT.601-7 Studio encoding parameters of digital television for standard 4:3 and wide screen 16:9 aspect ratios
    6.2. ITU-R BT.656-5-200712
    6.3. ITU-R BT.709-6-201506 Parameter values for the HDTV standards for production and international programme exchange
    6.4. ITU-R BT.1543-1-201506
    6.5. ITU-R BT.1847-1-201506
    6.6. ITU-R BT.1886 Reference electro-optical transfer function for flat panel displays used in HDTV studio production
    6.7. ITU-R BT.2020-2-201510 Parameter values for ultra-high definition television systems for production and international programme exchange
    6.8. ITU-R BT.2100-0 (07/2016) Image Parameter Values for High Dynamic Range Television for use in Production and International Programme Exchange

0723/DL/MESH/349349-001US