IEEE 1722 Media on AVB Networks

Presentation to the AVnu Alliance Broadcast Advisory Council

Rob Silfvast
Principal System Architect, Avid
rob.silfvast@avid.com

6-June-2013

Outline

- Brief overview of IEEE 1722 AV Transport Protocol
- Description of AVTP Professional Video Format
- Comparison and co-existence with ST-2022
- Clock Reference Streams on AVB
- Q&A

AVB Terminology

AVB Cloud (or AVB Domain)

A collection of interconnected network nodes, and the links between them, all of which support the 802.1 AVB standards

Stream

A regular flow of packets in an AVB cloud that contains one or more channels of media (or other real-time) data. A stream has a unique ID and is reserved using 802.1Q-2011 Stream Reservation Protocol.

Talker

An entity in the AVB cloud that can send a stream

Listener

An entity in the AVB cloud that can receive a stream

Controller

An entity on the network which configures and connects Talkers and Listeners in an AVB network

Media Synchronization across a shared LAN

- Maintain A/V lip-sync between rendering devices
  - No loss of sync due to separate transport flows across network
  - Separate streams can have different delivery times
- Maintain tight audio sync between multiple transducers
  - Need < 2 usec to support phased arrays (mics or loudspeakers)
  - Tight video sync also useful for multi-display (compound) screens
- Convey media clocks from master nodes to slave nodes
  - Across a fundamentally asynchronous network infrastructure (!)
  - Support multiple, independent clock domains on same network

Audio Video Transport Protocol (AVTP)

IEEE Std 1722™-2011

IEEE Standard for Layer 2 Transport Protocol for Time-Sensitive Applications in Bridged Local Area Networks

IEEE Computer Society

Sponsored by the
Microprocessor Standards Committee

IEEE 1722 Basics

5.1.1 AVTP network requirements

All devices that send, receive, or forward AVTP data are required to support the services provided by gPTP, SRP, and FQTSS. AVTP relies on these services being available to function properly. The behavior when a Talker, Listener, or bridge on the network does not support these AVB protocols is out of the scope of this standard.

AVTP makes use of gPTP to provide a network wide time base that can be used to convey timing information from a Talker to Listener(s).

AVTP makes use of SRP and FQTSS to provide reliable network delivery with bounded network latency for transporting audio/video data from Talker to Listener.

AVTP data can be sent from Talker to Listener either directly or forwarded by AVB bridges to the Listener.

IEEE 1722 Basics

- Ethernet frame format
- Carried at the Data-Link layer (OSI Layer 2)
- Ethertype = 0x22F0
- Max Frame size = 1542 Bytes (on the wire)
- Priority Tagging used to differentiate from non-AVB traffic
- VLAN Tagging supported
- Destination Address is (typically) a Multicast Address
- Unique ID per Stream

802.3 Ethernet frame structure

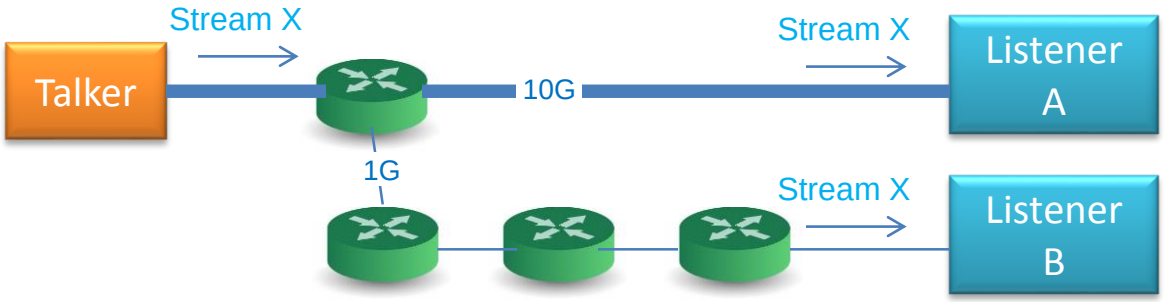
Preamble	Start of frame delimiter	MAC destination	MAC source	802.1Q tag (optional)	Ethertype (Ethernet II) or length (IEEE 802.3)	Payload	Frame check sequence (32-bit CRC)	Interframe gap
7 octets	1 octet	6 octets	6 octets	(4 octets)	2 octets	42[note 2]–1500 octets	4 octets	12 octets
	← 64–1522 octets →							
← 72–1530 octets →								
← 84–1542 octets →								

IEEE 1722-2011 Data Format

- Based on IEC 61883
- Used for media transport on IEEE 1394 (FireWire)
- Supports myriad audio and video formats
- Even supports timecode and MIDI
- Has some “legacy baggage”

IEEE 1722 Presentation Timestamp

- Tells Listener the exact time to present media samples
  - Using the common 802.1AS clock as a “measuring stick”
  - De-couples media play-out time from network transit time



Both nodes render Stream X payload at the same time

IEEE P1722a – new features for AVTP

- AVTP Video Formats
  - **“APVF” = AVTP Professional Video Format**
  - RTP Video encapsulated in an AVB stream
- AVTP Audio Formats
  - Lose the baggage from IEC 61883 formatting
- Clock Reference Streams
  - Convey media clocks across the network
  - Leverages Presentation Time concept of 1722

APVF (AVTP Pro Video Format)

- Designed to carry SDI-encoded bitstreams, including all HANC and VANC data
  - 270 Mbit (SMPTE 259)
  - 1.5 Gbit (SMPTE 292)
  - 3 Gbit (SMPTE 425)
- Packetization scheme correlates to raster scanning
  - Start of video line always starts a new packet
  - AVTP timestamp for every horizontal blanking event
- Marker flags to indicate key events in the stream
  - Vertical blanking
  - Interlaced: field 0 vs field 1
  - metadata frames

SMPTE Standards for Video over IP

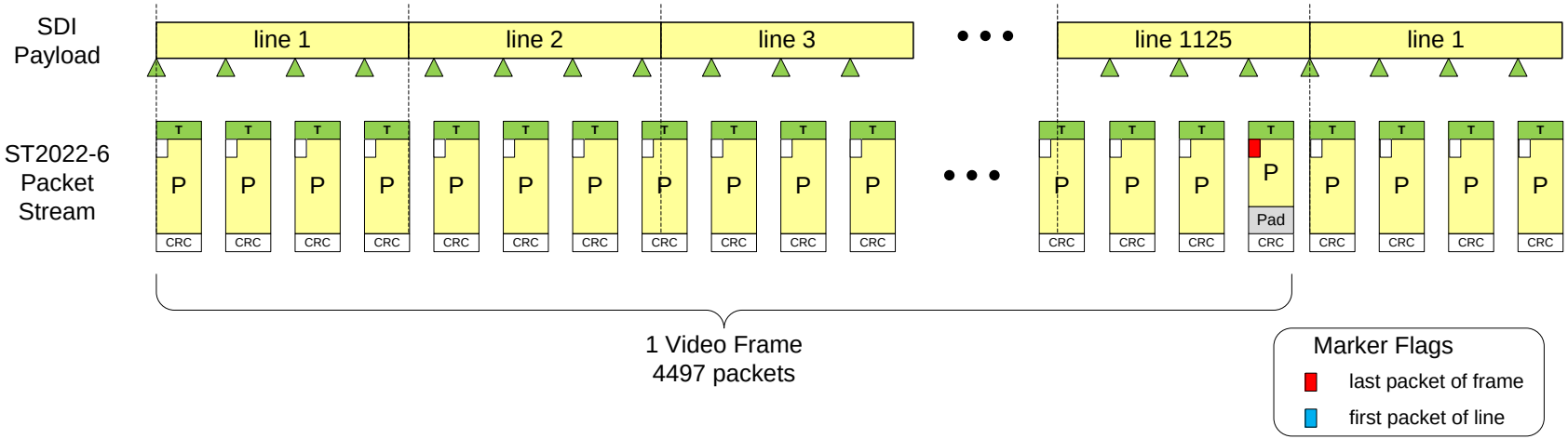
- ST-2022-1 and ST-2022-2
  - MPEG-2 Transport Stream (TS) in Constant Bit Rate (CBR) with Forward Error Correction (FEC)
- ST-2022-3 and ST-2022-4
  - As above but with Variable Bit Rate (VBR)
- ST-2022-5 and ST-2022-6
  - Uncompressed video over RTP/UDP IP from 270 Mb/s to 3 Gb/s
  - Payload encoding uses SDI formats (ST-259/292/495)
  - Optional FEC (ST-2022-5)

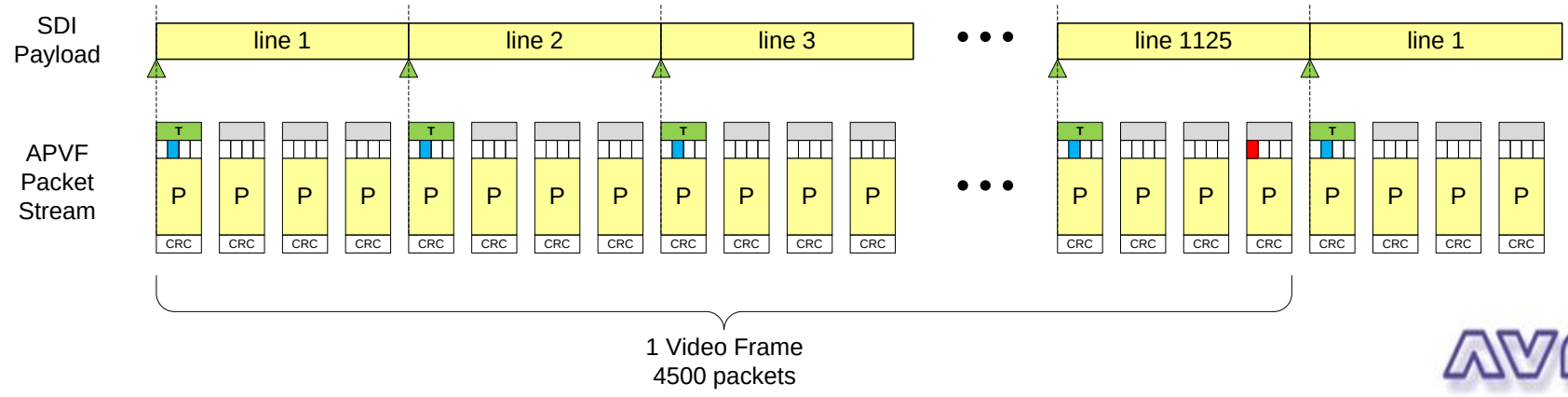
Comparison of ST-2022 and AVB

Attribute	SMPTE 2022 (over IP)	AVB Pro Video Format *
Compression	MPEG, H.264, J2K, none	None
FEC	Optional	None
Link types	WAN, MAN (Layer-3 datagrams)	LAN (Layer-2 datagrams)
End-to-end Sync	External to the standard	802.1AS and 1722 presentation time
QoS	External to the standard	802.1Q SRP and FQTSS
Expected Latency (based on encoding)	Medium (~1 video frame)	Low (~1 video line)
Best Application	Long haul, back haul, contribution links, primary distribution	SDI-like transport of live media within a facility

*IEEE P1722A Draft

Packetization schema: ST-2022-6 vs IEEE P1722A APVF





Timestamps in ST2022-6 vs 1722 APVF

- ST-2022-6 has two timestamp fields per packet
  - RTP timestamp correlates to “first byte in the RTP datagram”
    - This first byte is the media payload header
    - Unrelated to the actual media
  - Video timestamp
    - Is optional
    - Correlates to first sample in media payload to a video clock at the sender
    - Most often this is not a particular point in the raster
    - No mandate for sender’s video clock to be replicated at all nodes
- APVF has one timestamp per video line (every Nth packet)
  - Presentation time of first pixel in the line

Why Operate at Layer 2 ?

- Limiting the domain size can be a good thing
  - Don’t try to “boil the ocean”
    - L3 (IP) datagrams can go anywhere
    - An all-IP solution would need different modes for LAN vs WAN routing
  - Expectations of QoS and Latency more deterministic on a LAN
  - Low-latency maps to Local Area (Long haul is tolerant to latency)
  - Tighter user control => more secure
- Media packets handled lower in the network stack
  - Less header/overhead processing (3G SDI: 360,000 packets/sec!)
  - Filter at L2 => avoid implementing a HW-based IP parser
- L2-only is lower cost than L3+L2

ST-2022 and AVB together

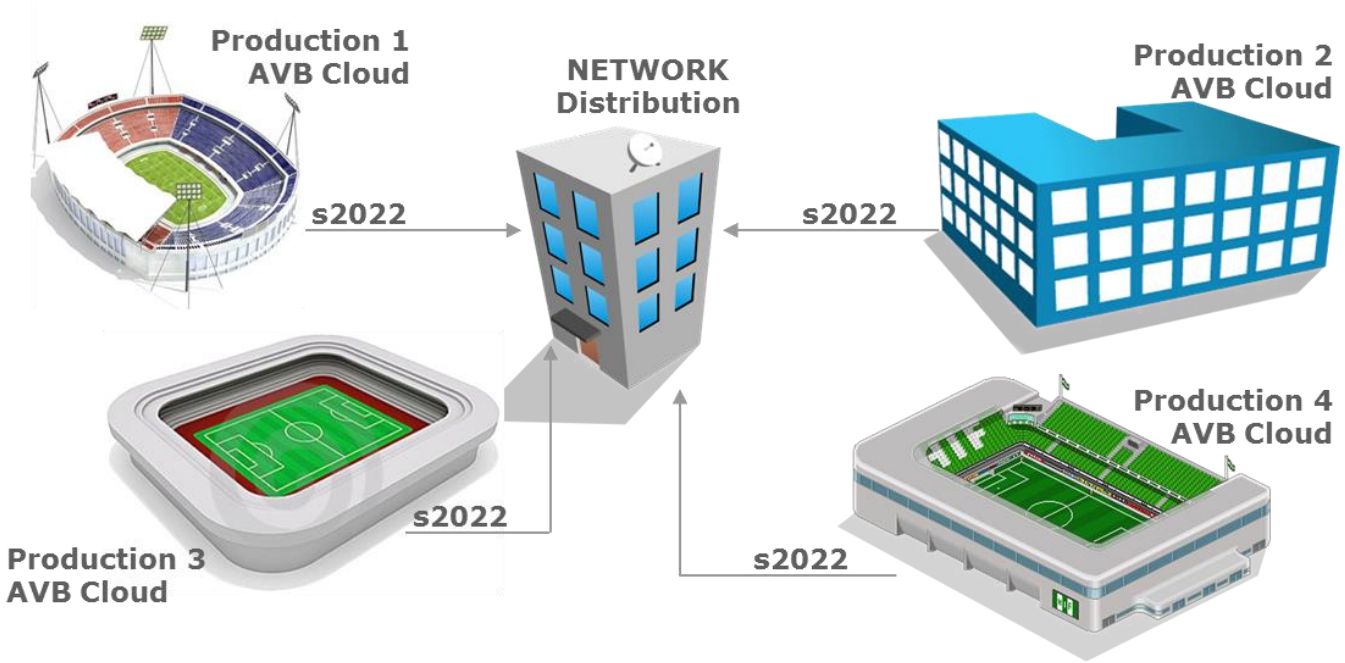


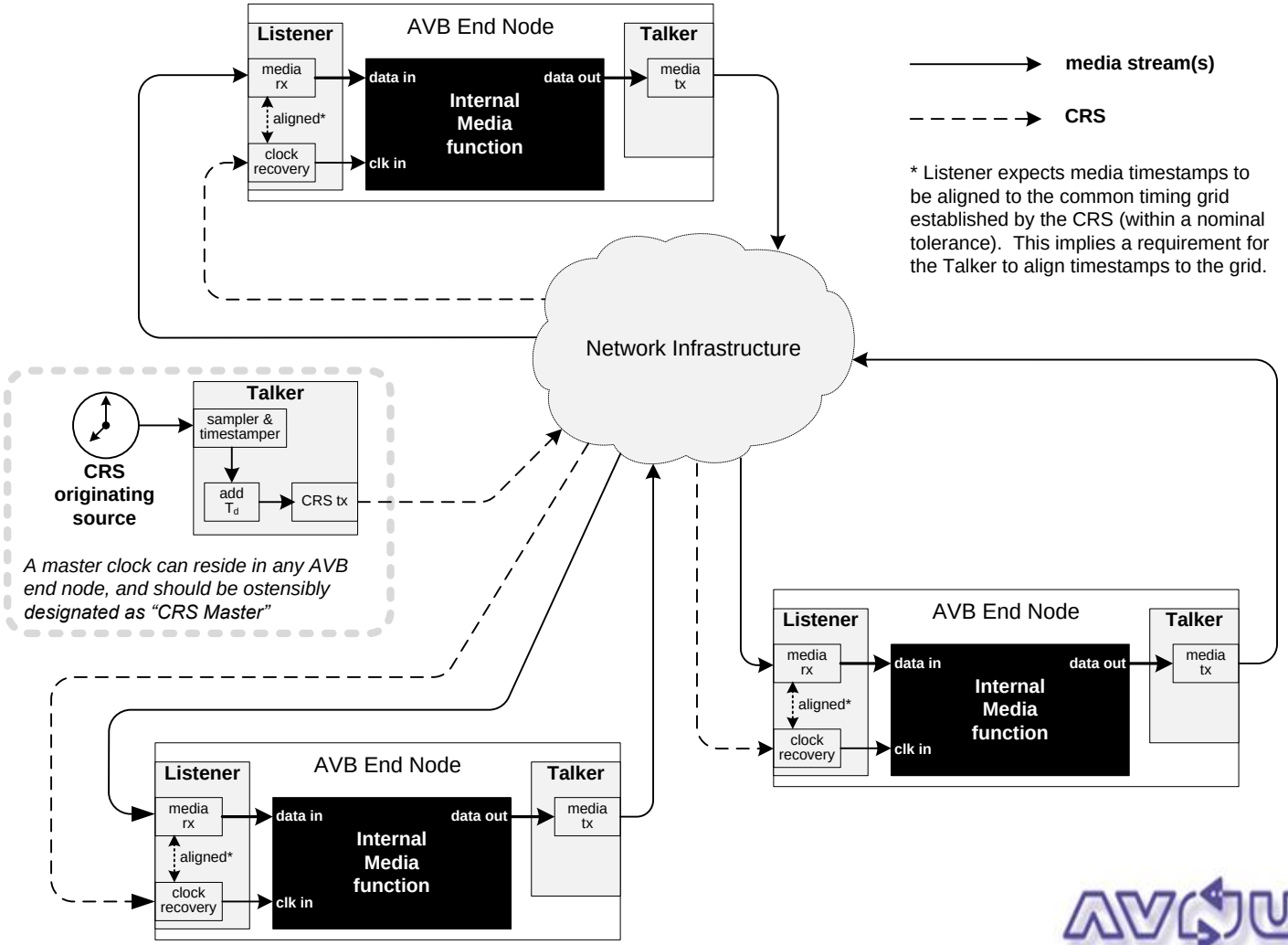
Diagram by
Axon Digital

"We think s2022 and AVB will work hand in hand: AVB for a closed, layer-2 production environment, and s2022 for connecting venues with the network as an alternative to satellite uplinks." -- Peter Schut, CTO Axon Digital

AVTP Clock Reference Streams

- Convey timing of “clock tick” events across an AVB LAN
- Establish common timing grid for all participating nodes
  - Same concept as House Sync
- No media payload => low bandwidth usage
- Talker = Clock Master (CRS Master)
- Multicast to all listeners in same clock domain
- Supports multiple clock domains on same LAN
  - CRS (and media) timing is decoupled from 802.1AS clock
  - Any number of CRS can be carried on a LAN

System Diagram with CRS



AVTP CRS details

- Common timing grid is typically the line clock for video or the sampling clock for audio
- ***CRS payload = timestamps***
  - corresponding to every Nth tick on the common grid
- Listeners can up-sample the timestamps to re-establish the grid clock
  - Use a PLL when very low jitter is required
- Simple logic at listener: compare CRS timestamps to locally running 802.1AS clock

Early Adoption of IEEE 1722a Pro Video Format

Axon Digital CEO Jan Eveleens at NAB show, April 2013

Demonstrating a working prototype of 3G SDI running bi-directionally across 10G Ethernet/AVB, through an Extreme Networks switch

Thank You

(Q&A)