AES67 & SMPTE ST 2110: Transport & Synchronization

Wed, November 10, 2021
17:00 h (CET)

Andreas Hildebrand, ALC NetworX
Topics:

• Transport protocols
  – RTP / UDP / IP
  – Multicasting
  – QoS

• RTP packet setup
  – Media encoding
  – Packet time

• Session description

• Timing & Synchronization
  – PTP
  – Media clocking
  – Alignment
  – Latency
Preliminary remark:

RAVENNA, AES67 & SMPTE ST 2110

share the same principles and

utilize the same protocols and methods for

synchronization and transport of real-time media.
Network Transport
7-layer OSI Model
7-layer OSI Model
7-layer OSI Model
TCP vs. UDP (Layer 4): TCP features

- Secured transport protocol between networked hosts:
  - applications on hosts create a connection one to another
  - guarantees reliable in-order delivery of sender to receiver
  - sequence numbers for ordering received TCP segments and detecting duplicate data
  - checksums for segment error detection
  - acknowledgements and timers for detecting and adjusting to loss or delay
  - retransmission and timeout mechanisms for error control
  - unpredictable delay characteristics

⇒ not suitable for real-time communication
TCP vs. UDP (Layer 4): UDP features

- simple “unreliable” datagram transport service:
  - does not provide reliability and ordering guarantees
  - datagrams may arrive out of order or go missing without notice
  - checksum for detection of packages with bit errors
  - faster and more efficient for lightweight or time-sensitive purposes

- obvious choice for real-time audio and/or video transmission
  - RFC 768, August 1980
7-layer OSI Model
RTP Packets (Layer 5)

- Consist of RTP header, optional payload headers and the payload itself
- RTP header (overhead) = **12 Bytes**

<table>
<thead>
<tr>
<th>0</th>
<th>4</th>
<th>8</th>
<th>16</th>
<th>31</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ver</td>
<td>P</td>
<td>X</td>
<td>CC</td>
<td>M</td>
</tr>
<tr>
<td>Payload Type</td>
<td>Sequence Number</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Timestamp</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Source Synchronization Identifier (SSRC)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Options + Padding (optional)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Audio Data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>or</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Video Data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
RTP Packets (Layer 5)

- Consist of RTP header, optional payload headers and the payload itself
- RTP header (overhead) = **12 Bytes**
- IP + UDP + RTP overhead = 20 + 8 + 12 = **40 Bytes**
- MTU (maximum transmission unit, largest size of a packet that can be transmitted without being split) 1500 Bytes in an IP/Ethernet LAN: in principle **0 to 1460 bytes** available for RTP payload data per packet
Layered Packet Encapsulation

Layer 2 (Link Layer)

Layer 3 (Network Layer)

Layer 4 (Transport Layer)

Layer 5 (Session Layer)

Ethernet Header

IP Header

UDP Header

RTP Header

RTP Payload (PCM Modulated Data)

Ethernet Trailer

Bytes

14/18

20

8

12

1460

4

1518 / 1522
RTP - Layered Packet Encapsulation
Multicast
Multicast vs. unicast

- **Unicast:**
  - Point-to-point connection between sender and receiver
  - Every additional receiver adds another individual connection
  - Network traffic increases with every additional unicast stream
Multicast vs. unicast

- **Unicast:**
  - Point-to-point connection between sender and receiver.
  - Every additional receiver adds another individual connection.
  - Network traffic increases with every additional unicast stream.

![Diagram of multicast vs. unicast network connections]
Multicast vs. unicast

- **Unicast:**
  - Point-to-point connection between sender and receiver
  - Every additional receiver adds another individual connection
  - Network traffic increases with every additional unicast stream

⇒ Scenario: individual stream between two devices (e.g. multi-channel stream between two remote locations)
Multicast vs. unicast

- **Multicast (one-to-many):**
  - Only one connection per stream on transmitter side
  - Switches “know” participants (receivers) of any multicast and forward packets only to registered receivers (IGMP)
Multicast vs. unicast

- Multicast (one-to-many):
  - Only one connection per stream on transmitter side
  - Switches “know” participants (receivers) of any multicast and forward packets only to registered receivers (IGMP)
Multicast vs. unicast

- **Multicast (one-to-many):**
  - Only one connection per stream on transmitter side
  - Switches “know” participants (receivers) of any multicast and forward packets only to registered receivers (IGMP)

![Diagram showing multicast and unicast traffic flow](image-url)
Multicast vs. unicast

- Multicast (one-to-many):
  - Only one connection per stream on transmitter side
  - Switches “know” participants (receivers) of any multicast and forward packets only to registered receivers (IGMP)
  - Switches copy packets as required
Multicast vs. unicast

- **Multicast (one-to-many):**
  - Only one connection per stream on transmitter side
  - Switches “know” participants (receivers) of any multicast and forward packets only to registered receivers (IGMP)
  - Switches copy packets as required
Multicast vs. unicast

- **Multicast (one-to-many):**
  - Only one connection per stream on transmitter side
  - Switches “know” participants (receivers) of any multicast and forward packets only to registered receivers (IGMP)
  - Switches copy packets as required
Multicast vs. unicast

- Multicast (one-to-many):
  - Only one connection per stream on transmitter side
  - Switches “know” participants (receivers) of any multicast and forward packets only to registered receivers (IGMP)
  - Switches copy packets as required
  - Minimizes network load - traffic increases on last (closest to receiver) segment(s) of network path only
Multicast vs. unicast

- **Multicast (one-to-many):**
  - Only one connection per stream on transmitter side.
  - Switches “know” participants (receivers) of any multicast and forward packets only to registered receivers (IGMP).
  - Switches copy packets as required.
  - Minimizes network load - traffic increases on last (close to receiver) segment(s) of network path only.
Multicast vs. unicast

- **Multicast (one-to-many):**
  - Only one connection per stream on transmitter side
  - Switches “know” participants (receivers) of any multicast and forward packets only to registered receivers (IGMP)
  - Switches copy packets as required
  - Minimizes network load - traffic increases on last (closest to receiver) segment(s) of network only
Multicast vs. unicast

- **Multicast (one-to-many):**
  - Only one connection per stream on transmitter side
  - Switches “know” participants (receivers) of any multicast and forward packets only to registered receivers (IGMP)
  - Switches copy packets as required
  - Minimizes network load - traffic increases on last (closest to receiver) segment(s) of network path only

⇒ *Scenario:* distribution of a single source to many potential recipients (e.g. program stream to journalist desktops)
Packet Setup
Encoding / packet setup (AES67):

- **Media format**: 16 / 24 bit linear PCM
- **Channel count**: 1..8
- **Sample rates**: 48 (44.1 / 96) kHz
- **Packet time**: 1 ms = 48 samples/channel (6 / 12 / 16 / 192)
- **max. payload size (MTU)**: 1440 bytes (taking IPv6 constraints into consideration)

⇒ Examples:

- #1: 16 bit PCM, 2 channels, 48 samples (1 ms @ 48 kHz): 192 bytes
- #2: 24 bit PCM, 2 channels, 48 samples (1 ms @ 48 kHz): 288 bytes
- #3: 24 bit PCM, 8 channels, 48 samples (1 ms @ 48 kHz): 1152 bytes
- #4: 16 bit PCM, 2 channels, 192 samples (4 ms @ 48 kHz): 768 bytes
**RTP Packets (RFC 3550)**

- Consist of RTP header, optional payload headers and the payload itself
- RTP header (overhead) = **12 Bytes**

<table>
<thead>
<tr>
<th>0</th>
<th>4</th>
<th>8</th>
<th>16</th>
<th>31</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ver</td>
<td>P</td>
<td>X</td>
<td>CC</td>
<td>M</td>
</tr>
</tbody>
</table>

- Sequence Number
- Timestamp
- Source Synchronization Identifier (SSRC)
- Options + Padding (optional)
- Audio Data
  - or
  - Video Data
RTP payload types (PT)

- static types: index 0 - 95, none applicable for broadcast audio quality (except L16 / 44100Hz)
- Dynamic payload types: index 96-127
  - L16, L24 defined (RFC 3551/3190)
  - transparent AES3 via AM824 format (not officially registered with IANA, but defined as audio subtype in SMPTE ST 2110-31)
- Dynamic formats need to be signaled: ⇒ SDP (Session Description Protocol)
Session Description
AES67 & SMPTE ST 2110: Transport & Synchronization

AES67 - SDP Session Description Protocol (RFC 4556)

• Required to describe stream formatting, synchronization and connection information
• Provided by a sender (or management instance) for each stream
• Human-readable text:

```
v=0
o=1 0 IN IP4 192.168.1.100
s=RAVENNA demo stream
t=0 0
a=ts-refclk:ptp=IEEE1588-2008:00-60-6e-ff-fe-7c-23-0f:0
a=mediакl:direct=0
m=audio 5004 RTP/AVP 98
a=rtpmap:98 L24/48000/2
a=c=IN IP4 239.3.14.142
a=recvonly
a=ptime:1
```
Quality of Service
QoS – Differentiated Services (DiffServ)
QoS – Differentiated Services (DiffServ)

• Defined in RFC 2474
• Defines up to 64 traffic classes (i.e. EF, AFx, CSx, BE etc.)
• Packets are tagged with DSCP value (0 – 63)
QoS – Differentiated Services (DiffServ)

ToS = DSCP (6 bits)
QoS – Differentiated Services (DiffServ)

Flow 1: BE
Flow 2: EF, EF
Flow 3: CS6, CS6
Flow 4: BE
Flow 5: 
Flow 6: BE
Flow 7: AF, AF
Flow 8: BE

CS6 = q1
EF = q2
AF = q2
BE = q3
QoS – Differentiated Services (DiffServ)

- Defined in RFC 2474
- Defines up to 64 traffic classes (i.e. EF, AFx, CSx, BE etc.)
- Packets are tagged with DSCP value (0 – 63)
- Switches store packets in different priority queues (requires proper configuration)
QoS – Differentiated Services (DiffServ)

Flow 1
Flow 2
Flow 3
Flow 4
Flow 5
Flow 6
Flow 7
Flow 8

q1
q2
q3

CS6
EF
AF
BE

Highest Priority
Middle Priority
Lowest Priority

CS6
EF
AF
BE

q1
q2
q3
**QoS – Differentiated Services (DiffServ)**

- Defined in RFC 2474
- Defines up to 64 traffic classes (i.e. EF, AFx, CSx, BE etc.)
- Packets are tagged with DSCP value (0 – 63)
- Switches store packets in different priority queues (requires proper configuration)
- Egress scheduler forwards packets from higher prioritized queues first (strict priority / weighted round robin / guaranteed minimum bandwidth …)
QoS – Differentiated Services (DiffServ)

Flow 1: BE
Flow 2: EF, EF
Flow 3: CS6, CS6
Flow 4: BE
Flow 5: BE
Flow 6: BE
Flow 7: AF, AF
Flow 8: BE

Classifier:
- Highest Priority: CS6 = q1
- Middle Priority: EF = q2, AF = q2
- Lowest Priority: BE = q3

Scheduler:
- Port: CS6
QoS – Differentiated Services (DiffServ)

Flow 1: BE
Flow 2: EF, EF
Flow 3: CS6, CS6
Flow 4: BE
Flow 5: BE
Flow 6: BE
Flow 7: AF, AF
Flow 8: BE

Classifier:
- CS6 = q1
- EF = q2
- AF = q2
- BE = q3

Scheduler:
- Highest Priority
- Middle Priority: EF, AF, EF
- Lowest Priority: BE, BE, BE, BE

Port:
AF
QoS – Differentiated Services (DiffServ)

- Defined in RFC 2474
- Defines up to 64 traffic classes (i.e. EF, AFx, CSx, BE etc.)
- Packets are tagged with DSCP value (0 – 63)
- Switches store packets in different priority queues (requires proper configuration)
- Egress scheduler forwards packets from higher prioritized queues first (strict priority, weighted round robin, guaranteed minimum bandwidth)
- Needs to be supported along full path from the transmitting to the receiving end
- No admission control → congestion / packet dropping possible when bandwidth is exceeded
Timing & Synchronization
Synchronization & Media Clocks

- All nodes are running local clocks
- Local clocks are precisely synchronized to a common wall clock via PTP

  *PTPv1 standardized by IEEE in 2002 (IEEE 1588-2002)*
  *PTPv2 followed in 2008 (IEEE1588-2008)*
  *PTPv1 and PTPv2 are not compatible!*
Synchronization & Media Clocks

• All nodes are running local clocks
• Local clocks are precisely synchronized to a common wall clock via PTP
• Media clocks are generated locally from synchronized local clock
Synchronization & Media Clocks

Master Clock

Slave Clocks (nodes)

Media Clocks

GPS

PTP
Synchronization & Media Clocks

• All nodes are running local clocks
• Local clocks are precisely synchronized to a common wall clock via PTP
• Media clocks are generated locally from synchronized local clock
• Generation of any desired media clock (sample rate) possible
• Concurrent operation of different media clocks possible
• Phase accuracy of AES 11 ($\pm$ 5% of sample period) achievable by deployment of PTP-aware switches (BC or TC)
• Synchronization across facilities possible by reference to absolute time (TAI / GPS)
• Essence data (audio samples or video frames) is related to the media clock upon intake - essentially receiving a generation “time stamp” with respect to the media clock
RTP Packets (Layer 5)

- Consist of RTP header, optional payload headers and the payload itself
- **RTP header** (overhead) = 12 bytes, **RTP payload** (linear audio data) = up to 1440 bytes
- **RTP Timestamp** = media clock counter (for linear PCM audio) = 32 bits (4 bytes)
  - rollover will occur roughly once per day (~ 1d, 51m, 19s)
**Synchronization & Media Clocks**

- Offset $R$ is established on stream start-up
- $R$ may be random to defeat crypto-text attacks
- This offset will be constant throughout the stream’s lifetime
- The offset ($R$) will be conveyed via SDP (a=mediaclk:direct=<offset>) – must be “0” in ST2110
Synchronization & Media Clocks

- All nodes are running local clocks
- Local clocks are precisely synchronized to a common wall clock via PTP
- Media clocks are generated locally from synchronized local clock
- Generation of any desired media clock (sample rate) possible
- Concurrent operation of different media clocks possible
- Phase accuracy of AES 11 (± 5% of sample period) achievable by deployment of PTP-aware switches (BC or TC)
- Synchronization across facilities possible by reference to absolute time (TAI / GPS)
- Essence data (audio samples or video frames) is related to the media clock upon intake - essentially receiving a generation “time stamp” with respect to the media clock
- Fixed / determinable latency by configuring a suitable link offset (“playout delay”)

AES67 & SMPTE ST 2110: Transport & Synchronization
AES67 synchronization - link offset (latency)

RTP timestamp of (first) sample (in packet) + Link offset - RTP offset
SDP (a= mediack: direct=<offset>)

Desired playout time for sample

Ingress time reference point

IEEE 1588 measurement planes

IP network

Sender network system

Network stack and controller

Sender packet buffer

Media clock

Receiver network system

Network stack and controller

Receiver packet buffer

Media clock

ADC

DAC
Synchronization & Media Clocks

- All nodes are running local clocks
- Local clocks are precisely synchronized to a common wall clock via PTP
- Media clocks are generated locally from synchronized local clock
- Generation of any desired media clock (sample rate) possible
- Concurrent operation of different media clocks possible
- Phase accuracy of AES 11 (± 5% of sample period) achievable by deployment of PTP-aware switches (BC or TC)
- Synchronization across facilities possible by reference to absolute time (TAI / GPS)
- Essence data (audio samples or video frames) is related to the media clock upon intake - essentially receiving a generation “time stamp” with respect to the media clock
- Fixed / determinable latency by configuring a suitable link offset (“playout delay”)
- Inter-stream alignment by comparing and relating the time stamps of individual essence data
Production Workflow Timing

Image courtesy of Andy Rayner (Nevion)
Production Workflow Timing

Image courtesy of Andy Rayner (Nevion)
Latency
Latency – determining factors:

- Underlying network technology (Fast Ethernet / Gigabit Ethernet / …)
- Network topology (number of hops, distances)
- Network jitter (switch performance, hops, bandwidth utilization, competing traffic, QoS)
- Stream / packet configuration (samples per packet, sampling rate)
Latency – dependency on frames per packet & sampling rate:

<table>
<thead>
<tr>
<th>Bytes per sample</th>
<th>3</th>
<th>MTU-RTP 1460</th>
<th>Freq 48000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frames per packet</td>
<td>1</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>Latency ms</td>
<td>0.02</td>
<td>0.17</td>
<td>0.25</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Channels per frame</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>24</td>
<td>36</td>
<td>96</td>
<td>144</td>
<td>384</td>
<td>768</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>48</td>
<td>72</td>
<td>192</td>
<td>288</td>
<td>768</td>
<td>1536</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>72</td>
<td>108</td>
<td>288</td>
<td>432</td>
<td>1152</td>
<td>2304</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>96</td>
<td>144</td>
<td>384</td>
<td>576</td>
<td>1536</td>
<td>3072</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>15</td>
<td>120</td>
<td>180</td>
<td>480</td>
<td>720</td>
<td>1920</td>
<td>3840</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>18</td>
<td>144</td>
<td>216</td>
<td>576</td>
<td>864</td>
<td>2304</td>
<td>4608</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>21</td>
<td>168</td>
<td>252</td>
<td>672</td>
<td>1008</td>
<td>2688</td>
<td>5376</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>30</td>
<td>240</td>
<td>360</td>
<td>960</td>
<td>1440</td>
<td>3840</td>
<td>7680</td>
<td></td>
</tr>
</tbody>
</table>
Latency – determining factors:

- Underlying network technology (Fast Ethernet / Gigabit Ethernet / ...)
- Network topology (number of hops, distances)
- Network jitter (switch performance, hops, bandwidth utilization, competing traffic, QoS)
- Stream / packet configuration (samples per packet, sampling rate)

⇒ Base line: sub-milliseconds latency achievable (even with layer 3 packet-switching technology)
Questions?
Next RAVENNA webinar:

**Status update on standards & industry activities**

17th November 2021 - 17:00 h CET

Presenters: Andreas Hildebrand (ALC NetworX)

In this webinar, Andreas will provide a status update on relevant standards and activities relating to IP-based real-time media transport. The webinar includes reports on latest developments in AES, SMPTE, AMWA NMOS and AIMS, with particular focus on AES67 and ST 2110. A quick status update on iPMX related activities will round-up this webinar.

Register

www.ravenna-network.com
Contact information:

Andreas Hildebrand
ALC NetworX GmbH
ravenna@alcnetworx.de

www.ravenna-network.com