Synchronization of ST 2110 Audio

- Andreas Hildebrand –
  RAVENNA Technology Evangelist
  ALC NetworX, Munich
Andreas Hildebrand, RAVENNA Technology Evangelist
• more than 25 years in the professional audio / broadcasting industry
• graduate diploma in computer science
• R&D, project & product management experience
• member of AES67 TG and ST2110 DG

ALC NetworX GmbH, Munich / Germany
• established 2008
• R&D center
• developing & promoting RAVENNA
• Partnerships with > 40 manufacturers

RAVENNA
• IP media networking technology
• designed to meet requirements of professional audio / broadcasting applications
• open technology approach, license-free
• fully AES67-compliant (built-in)
Timing & Synchronization – General Requirements

- Media bit-transparency
  - no sample rate conversion
  - streams need to run on same media clock
- Concurrent operation of different sample rates on same network
- Determinable (low) end-to-end latency
- Time alignment between media streams
- Replacement for “house clock” distribution (word clock, black burst etc.)

⇒ Clock reassembly from stream data not appropriate
⇒ Distribution of master clock beats not sufficient
⇒ Common understanding of absolute time required (“wall clock”)
Timing & Synchronization – Accuracy Requirements

• Audio applications have highest time accuracy & precision demands:

  ➲ Sample accurate alignment of streams (± ½ sample)
    – @ 48 kHz: ± 10 µs
    – @ 96 kHz: ± 5 µs
    – @ 192 kHz: ± 2.5 µs

  ➲ “Distribution” of word clock reference
     (AES11 calls for ± 5% max jitter / wander):
    – @ 48 kHz: ± 1 µs
    – @ 96 kHz: ± 500 ns
    – @ 192 kHz: ± 250 ns
Synchronization & Media Clocks

- All nodes are running local clocks
- Local clocks are precisely synchronized to a common wall clock via IEEE 1588-2008 (PTPv2)
  - *PTPv1 standardized by IEEE in 2002 (IEEE 1588-2002)*
  - *PTPv2 followed in 2008 (IEEE1588-2008)*
  - *PTPv1 and PTPv2 are not compatible!*
How PTPv2 works

- Nodes are organized in a master/slave hierarchy
  *The grandmaster is at the top, it is elected according to clock quality.*

- Grandmaster multicasts periodic sync messages
  *Clients learn the grandmaster time, and correct their own time.*

- Transmission delay is measured with a delay_request / delay_response message pair
  *Measured delay is used to correct the time extracted from the sync message. Delay measurement can be very accurate with support from switches (BC or TC).*

- Received grandmaster time is used to drift-compensate local clock
  *Local clock can be a disciplined oscillator (VCO or VCXO), or it can be a free-running clock with digital correction (more common).*

- Local clock in each node is used to timestamp PTP messages
  *Highest precision requires hardware timestamping support in a node, either in the PHY, or in the MAC, or in-between (the closer to the wire the better).*
Grandmaster Selection

- The standard defines a common Best Master Clock Algorithm (BMCA)
  *Every node follows the same algorithm → all arrive at the same result.*

- Every node holds a data set describing the qualities of its own clock
  *There are several different quality criteria which are considered.*

- Data sets are distributed in the network with *Announce* messages
  *All nodes know data sets of any other node and can compare against their own sets.*

- BMC Algorithm is re-run when current grandmaster disappears
  *There is a period of time without sync messages until the new grandmaster takes over. The clients must be able to bridge the gap.*

- Grandmaster does not need to be a “dedicated” GM device
  *Master capability can be a function of an ordinary node.*
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 (± 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 (network clock)
Synchronization & media clocks

- 3 type of clocks in the system:
  - Wall clock (reference clock) - provided by Grandmaster
    - local copy of the wall clock in each node
  - Media clock – derived from the local clock (i.e. 48 kHz for audio, 90 kHz for video)
  - RTP clock (stream clock) – derived from the media clock
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=mediackl:direct=<offset>) – must be “0” in ST2110
RTP Packets (Layer 5)

- Consist of RTP header, optional payload headers and the payload itself
- RTP header (overhead) = 12 bytes, payload (linear audio data) = up to 1440 bytes
- RTP Timestamp = media clock counter (for linear PCM audio) = 32 bits (4 bytes)
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”)
Synchronization & Media Clocks - Link offset

Link offset
IEEE 1588 measurement planes

ADC → Sender packet buffer → Network stack and controller → Network clock → Sender network system

Media clock → Sender network system

IP network

Network stack and controller
Network clock
Receiver network system
Media clock

Receiver packet buffer → DAC

Egress time reference point

Ingress time reference point
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)
How to synchronize streams across various processing stages

• Problem:
  – Any stream leaving a (processing) device is a new stream
  – New alignment of (processed) essence to wall clock (reference) time
  – Alignment of original essence is lost

• Possible solutions:
  – Use of original time alignment for new stream (RTP timestamps adjusted to those of original essence)
    o Offset increases, might be too large for downstream Rx buffer
    o Which timestamps serve as reference when mixing essence?
    o How does the (processing) host know the exact relationship between ingress / and egress essence?
  – Carry origin timestamps as in-band meta data
    o Requires new payload format (audio essence data + audio meta data), or
    o Needs to make use of RTP header extensions mechanism
      (which in turn may result in variable / decreased audio payload segments)
  – Carry origin timestamps as out-of-band meta data
    o Requires new standard (in the works → AES X242, ST2110-41/-42, NMOS)
How to synchronize streams across various processing stages

- Problem:
  - Any stream leaving a (processing) device is a new stream
  - New alignment of (processed) essence to wall clock time
  - Alignment of original essence is lost

- Intermediate (?) / current solution:
  - Leave alignment task to management layer (i.e. Broadcast Controller)
    - Devices report processing delays to BC (or have fixed / configurable delays)
    - BC configures required Rx delay for subsequent stages (playout delay)
Production Workflow Timing
Thank you for your attention!

Contact information:

Andreas Hildebrand  
Technology Evangelist  
ravenna@alcnetworx.de

ALC NetworX GmbH  
Am Loferfeld 58  
81249 Munich  
Germany

www.ravenna-network.com