Audio Video Bridging (AVB) Assumptions

IEEE 802.1 AVB Plenary

May 2008 – Eilat, Israel

Green Text = Agreed to at a Plenary (was Blue)

Blue Text = Newly Agreed to (was Red at last Face 2 Face)

Black Text = Not Decided

Changes Marked with Red from last version

Don Pannell

Marvell

dpannell@marvell.com

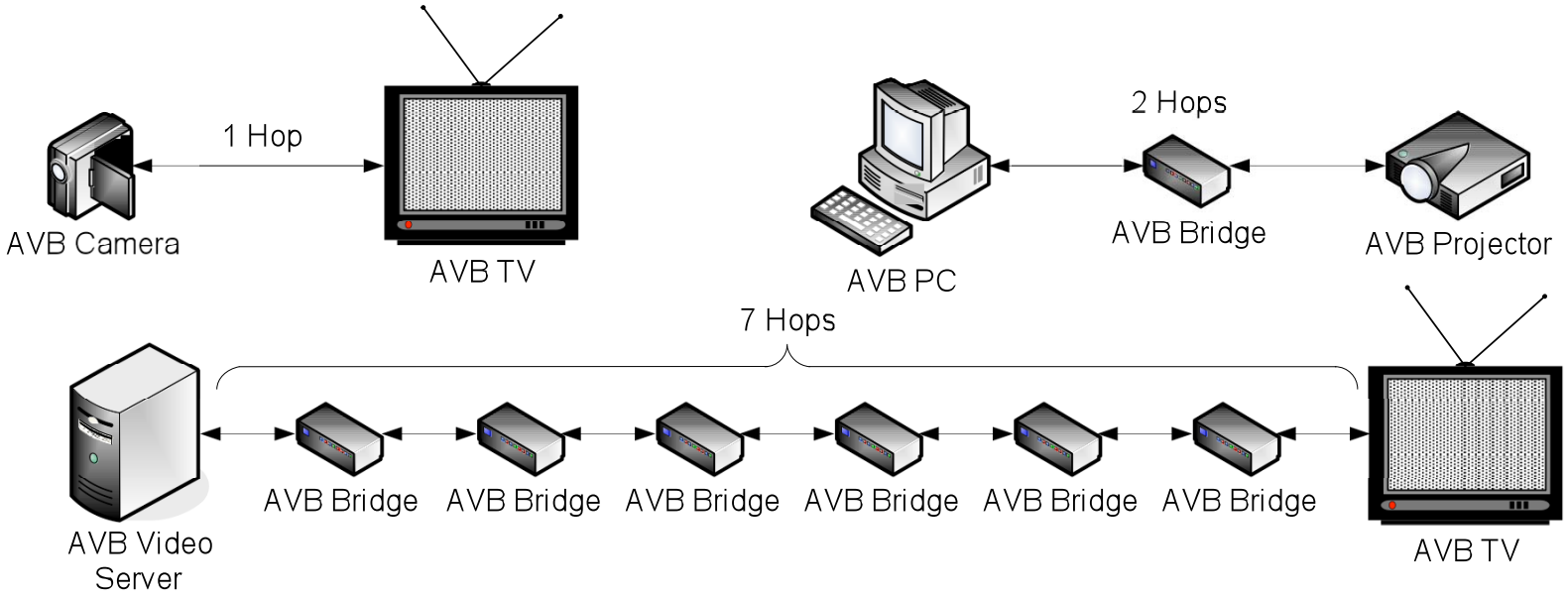
(updates May 12-15, 2008 by MJT)

Revision History

- Avb-pannell-assumptions-0508-v15: May 12-15, 2008 Eilat, Isael
- Avb-pannell-assumptions-0508-v14: May 2008 – calls
- Avb-pannell-assumptions-0408-v13: calls
- Avb-pannell-assumptions-0308-v12: 03/20/08 Orlando, FL
- Avb-pannell-assumptions-0308-v11: 03/05/08 Call
- Avb-pannell-assumptions-1107-v10: 01/28/08 Los Gatos
- Avb-pannell-assumptions-1107-v9: After many calls – 11/13/07 Atlanta
- Avb-pannell-assumptions-0907-v8: After many calls – 9/04/07 Stockholm
- Avb-pannell-assumptions-0707-v7: After many calls – 7/17/07 San Francisco
- Avb-pannell-assumptions-0607-v6: After Geneva Interim & after 5/23/07 call
- Avb-pannell-assumptions-0507-v5: After many calls
- Avb-pannell-assumptions-0407-v4: After 4/11/07 call
- Avb-pannell-assumptions-0407-v3: After 4/04/07 call
- Avb-pannell-assumptions-0307-v2: After 3/28/07 call
- Avb-pannell-assumptions-0307-v1: Before 3/28/07 call

Performance Goals

- Max Latency (802.1Qav – MAC to MAC, not analog source to sink eg., microphone to speaker or hard drive to display)
  - 802.3: AVB Class A: Less than 2 mSec over 7 Hops (this is based on a total trip budget of 10 - 15 mSec for lip to ear synchronization – Ref 4, Section 5.1)
    - This is affected by port speed, # ports on the bridges, and the max frame size? – Ref 5
  - 802.3: AVB Class B: Less than 20 mSec over 7 Hops?
  - 802.11: ??
- End node latency needs to be known and communicated, but it's out of scope
- Hops are defined in 802.1 terms, i.e., 1 LAN = 1 Hop



Performance Goals / Names

- Latency Variation over 7 Hops
  - 802.3: Bounded only by the maximum latency for each Class?
- Maximum Interference Amount per Hop
  - Class A: 1 Max size frame + Sum of the Maximum size of the Class A frames on each of its other ports – Ref 5
  - Class B: 1 Max size frame + 1 Max size Class A burst (based on max Class A BW allocation) + Amount of other Class B frames on each of its other ports
- Need to define a Traffic Class Measurement Period for bandwidth calculations
  - This results in Talkers using smaller frames on its per flow shapers – Use 125 uSec for Class A? 250 uSec for Class B?
- AVB Class A together with AVB Class B cannot use more than 75% of a link's bandwidth (i.e., not to exceed 75%)
  - The Remaining 25% (or more) is used for Legacy (non-AVB) streams
- Functional Device Type Names
  - AVB will use Talkers, Listeners & Bridges
    - Talker is the source of a stream, Listener is a receiver of a stream
    - A Bridge is an 802.1 Bridge
    - Any physical device could be any combination of these

802.3 Link Requirements

- Link Speed
  - 802.3: 100 Mbit/sec or faster (i.e., no 10 Mbit support)
  - Energy Efficient Ethernet issues need to be considered when we know its impact
    - We THINK that something like 10usec delays are OK?
- Link Duplex
  - 802.3: Full Duplex only (i.e., no half duplex support)
- AVB behavior over Link Aggregation (802.3 clause 43, now 802.1AX) links will not be specified
  - Problems with AS and Qat and Qav
  - i.e., Link Aggregation support is dependent on contributions from those that need it and may need to become part of a separate PAR
- Maximum Frame Size
  - Defined as Layer 2 bytes – or start of DA through end of CRC
  - 802.3: 1088 bytes for AVB Class A for 100 Mbit links (assumes 125 uSec Bandwidth Measurement Interval)

    For Everything else: 2000 bytes (i.e., Max Legal frame size) – actually this is the original 1500 bytes at the MSDU (Layer 3 & above bytes)!!! (i.e., no Jumbo frame support)
- Flow Control
  - 802.3x is not supported and cannot be used on AVB links
  - 802.1Qbb (Per Priority Pause) is not addressed by AVB but is not allowed for any AVB traffic

AVB Frame Format

- 802.1 Q Tagging
  - All AVB Streams will be Q Tagged
    - Talkers that are not VLAN aware must use a VID of 0x000 in these frames
    - Bridges transmitting to Listeners must not strip the Q Tag on AVB classes
  - All PTP frames (for 802.1AS) will NOT be Q Tagged
  - All SRP frames (for 802.1Qat) will be Q Tagged
  - All Q Tagged frames must be single Tagged at a Talker and a Listener
- VLANs
  - The VID is a VLAN and not a Stream Identifier
  - Stream Identifiers must be unique per VID
- Ether types
  - The type field of a frame is not a Stream Identifier
- Priorities
  - AVB Class A Streams will use a default Q Tag priority of 5 (PCP)
  - AVB Class B Streams will use a default Q Tag priority of 4 (PCP)

Stream Identification - SRP

- Stream Identification (for Queue Usage in Bridges & Listeners – Ref 1)
  - An Stream Reserved (SR) Frame is any Frame with a designated SR priority entering a SR port
  - A SR Port is a port mode used to differentiate the port from being a Edge port (SR ports are part of the SR domain, Edge ports are at the edge of the SR domain, not connected to an SR-aware device)
    - (need to find a way to describe this on a per class basis)
    - (also need to differentiate between .1BA and .1Qat/Qav ports, is it enough to say .1BA port is the same as a port that supports both .1Qat/av and .1AS?)
- Stream Policy (for Queue Usage in Bridges – Ref 1)
  - Only those frames that meet the Stream Identification (above) can be placed into the Stream Reserved (SR) Egress Queues
  - All other frames are placed into the non-SR Egress Queues
  - Frames that contain a designated SR priority entering a non-SR port must have their ~~Q Tag~~ priority re-mapped

SRP Assumptions

- Stream Identification (for Filtering in Bridges)
  - Different Multicast Streams must use Unique Multicast Addresses (within the same VLAN)
  - The standard will support Unicast (Individual) & Multicast (Group) DA for SR streams
  - Different Unicast streams may use the same Unicast address
  - Filtering is done on frame's DA + VID
- Each stream can use only one ~~PCP Q Tag~~ priority value
- A Reservation Identifier for SRP (i.e., Talker MAC Address along with a 16-bit higher level identifier) is for one stream only
- SRP will ~~currently~~ allocate a Static bandwidth for a stream (i.e., Dynamic, sometimes called Variable Bit Rate, support is dependent on contributions from those that need it and may need to become part of a separate PAR)
  - Therefore, Variable Bit Rate streams need to reserve the peak bandwidth that they need
- Reconfiguration of a stream's bandwidth is allowed

SRP Assumptions

- SR Frame Priorities are changeable
  - These are changeable for ‘engineered’ networks
  - For Home networks the set of specified defaults simply must work!
- ~~The mechanism for changing~~ the association of an SR Class to its frame Priority can be changed by management
- Hook to existing higher layer protocols without imposing any new requirements on the higher layer protocols
- Detailed diagnostics about the SR domain and its capabilities are not a part of SRP
  - The intent is to use the work being done in 802.1ag
- A Listener is ready to accept a stream at the time it makes a reservation request

SRP Assumptions

- SRP will Respond to changes in the network topology (e.g., Spanning Tree change) by re-converging after the topology change
- SRP will respond to changes in usable bandwidth (e.g., wireless and/or coordinated shared networks) by recalculating (i.e., by dropping lower ranking streams)
- Need to report error reason(s) to affected Nodes when a 'No' is given to an SRP stream request, e.g.,:
  - Can't store current DA (i.e., bridge is out of DA resources)
  - Use alternate DA (i.e., DA is already in use or Hash problem)
  - Don't have enough bandwidth on a link (report what bandwidth is left and which link? it is)
  - Out of MMRP database resources
  - Requested Priority is not an AVB Class in this device
    - Do we report all these at once (vector), or prioritize the error
    - Only report the first error on a path from talker to listener

SRP Assumptions

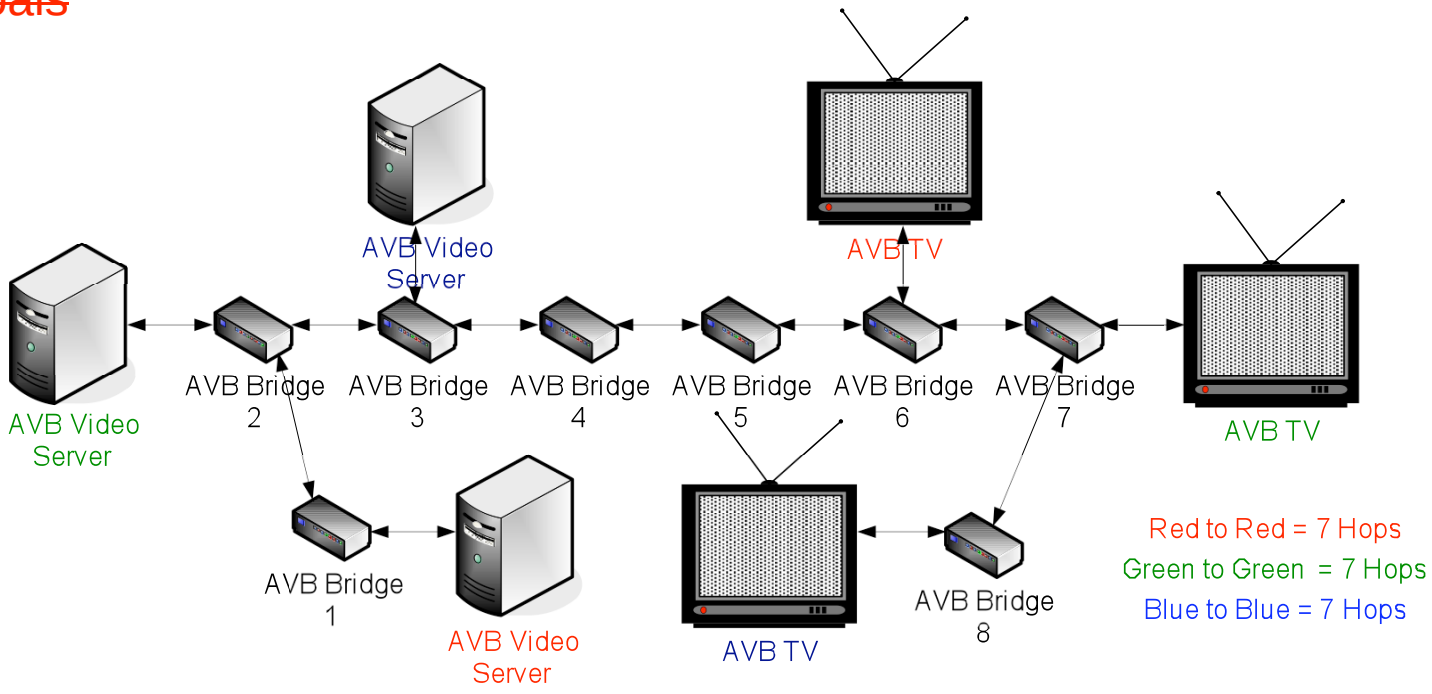
- MAC DA's are Allocated by the Talker:
  - And the Talker gets it from a higher layer ~~if its IP or From ??? if its not IP (IEEE 1722 need this)~~
    - 1722 is developing a MAC Address Acquisition Protocol (Annex C in 1722)
  - ~~It is in scope of SRP to allocate DA's, but not sure its necessary for SRP~~
- SRP v1 will not support dynamic bandwidth changes.
  - A later version will ~~Devices will Respond to changes in bandwidth request based on user requests (eg., when the user requests an increase or decrease in bandwidth)~~ by a make before break method
    - This is done by requesting the same Stream ID with a new T-Spec
    - To do this the T-Spec parameters (bandwidth & packets per second) can not change in opposite directions
- SRP v1 will not support multiple talkers using the same Stream ID.
  - A later version will support ~~If you have~~ multiple Talkers using the same Stream ID and they have different T-Specs, one or more of the Talker registrations will be rejected such that any remaining successful registrations, if any, will have the same T-Spec (multiple cameras and select just one case)
- The maximum time to make or break an SRP reservation is:
  - This defined per hop assuming a max of 7 hops
    - for consumer remote control applications this must not exceed 100ms
    - for professional video applications this must not exceed 20ms

SRP Assumptions

- It should be possible for SRP to indicate the presence/absence of protocol activity from adjacent Bridges (receipt of Talker registrations, etc.) and that can be used directly as a means of determining where the boundary lies.
  - This may make the boundary somewhat mobile, as it may move depending on whether or not a Bridge is issuing registration; however, if is isn't, then it also isn't a source of streams, so may as well be outside the region as in it.
  - Hence there is no need to use AB (LLDP) as a means of detecting the boundary
  - For Ethernet, we need to detect ports connected to Buffered Repeaters
    - To set these ports as non-SR ports use 802.1AS's path delay
    - How will this be connected?
    - This can be overridden for managed networks
- Talker reservations are propagated everywhere (by default) but they can be limited to only the ports where the destination address has been registered by MMRP. This is called Talker Pruning - which removes unnecessary Talker registrations in the rest of the network at the expense of slower re-convergence.

SRP Assumptions

- The figure below is a valid SR Cloud
- ~~Therefore,~~ SRP Needs to ~~Count Hops and~~ report the total worst case latency on a path ~~or Port Speed reported from all the Hops~~
  - a Listener can determine if it wants to listen to the stream as it may not meet the stream latency requirements ~~AVB Latency and/or Clock Quality Goals~~



SRP Questions

- When there is a change in the network:
  - Will not transmit them at best effort, nor oversubscribe the AVB queues
  - This means the flows will be interrupted
- MMRP or IGMP could be used on a 'Listener Proxy' port on an AVB Bridge. Specifying this proxy is highly desirable for 802.1BA
  - MMRP can be used for talker pruning in very large networks and will be specified by SRP, however it's important to note that talkers do not need to wait for MMRP registrations … other registration mechanisms may be available.
- Talkers and listeners must be MSRP aware
  - Bridges can proxy as a Listener
- ~~Does the SRP ACK's need to report the upper and lower bounds of what a path can support (to support 802.11)? Need a contribution to progress.~~
- Need to specify a function and associated parameters in Qav to respond to T-Spec requests (e.g., is 802.1AE enabled,…
- Need to make sure that there are margins built into the BW requested vs the Qav parameters to make sure that the +/- 100ppm clock difference between adjacent devices does not cause overflow

Domain Questions

- What is an SRP domain & its forwarding rules?
  - An SRP class domain consists of connected ports that are configured to use the same priority for the stream class
  - In the default/non-engineered case 802.1AS is needed to detect 'buffered repeater' however so this is a requirement of 802.1BA
- What is an AVB (or 802.1BA) domain?
  - Assume it to be a subset of an 802.1AS domain
  - It is the intersection of the 802.1AS domain and an SRP Class domain
- What is an 802.1AS domain?
  - A set of contiguous bridges, end stations and links that meet the requirement of 802.1AS
- AVB (802.1BA) streams must reside entirely inside an AVB Cloud? Yes
  - But SRP allows streams to be mapped to the edge of an SRP domain to proxy for a non-AVB Listener (like a PC)
    - In this case the AVB bridge snoops IGMP join requests from the Listener
  - SRP Guarantees are made only within the Talker's AVB Cloud

SRP Questions

- T-Spec Parameters?
  - Traffic Class Measurement Period done per stream at the Talker before the Talker's Class shaper: Class A: 125 uSec Class B: 250 uSec?
  - The traffic class, which implies a TSPEC measurement period depending on its Class (either Class A or Class B)?
  - The bandwidth, which is measured in total bytes transmitted in mac_service_data_units during a traffic class measurement period?
  - The maximum number of packets transmitted during a traffic class measurement period?
  - Minimum of 42 mac_service_data_unit bytes in min. size frames?
- The maximum packet size will be determined by bandwidth, traffic class and layer 2 technology (among other things)
- Fan-in is the main component of latency

PTP Assumptions

- PTP Clock Quality (802.1AS clock, Ref 2)
  - Application Jitter & Wander in Ref 3 are met by passing the .1AS clock through the reference (measurement) filter (filter to be defined in .1AS)
  - PTP time is based upon TAI which does NOT experience step changes at daylight savings time changes nor does it change when there is a leap second, nor does it change across time zones
  - +/- 100ppm or better from a free running ≥ 25 MHz clock
  - Less than 4 ppm per second drift (crystal drift) References needed! Spectrum (or Alan variance or TDEV) spec needed.
  - End point time synchronization accuracy for steady-state (up to 7 Hops) <= 1 uSec (i.e., any 2 PTP clocks separated by at most 7 Hops differ by no more than 1 uSec i.e., +/- 500ns from the grandmaster)
  - Endpoint media output synchronization accuracy within 1us after the 802.1AS to AES3 filter?
  - Network settling time at startup: < 2 Sec
    - Would like to see < 200 mSec? – need a contribution
  - Applications above 802.1AS will be provided with the PPM offset of the new grandmaster relative to the previous grandmaster - the grandmaster that won knows the previous grandmaster

PTP Assumptions

- PTP Clock Quality (802.1AS clock, Ref 2) Continued:
  - If a new grandmaster that is not traceable to TAI takes over, it takes over using the time of the previous grandmaster, if any (i.e., any new grandmaster must 1st sync to the current grandmaster before taking over), but provides no bound on the PPM change (other than specified above)
  - If the new grandmaster is traceable to TAI (or if any grandmaster clock is changed by, for example, the user), PTP time may experience an arbitrarily large step change in time
    - The new grandmaster will communicate the step change in time
  - 802.1AS provides 'GM Disruption Count' information such that when a stream is mapped and de-mapped, it can be done such that both ends know they're using the same grand master time. This information is important when a grandmaster experiences a step change in time.
  - Assuming no spanning tree reconfiguration, on grandmaster changeover, the time between the last Sync of the old grandmaster and the time of the 1st Sync of the new grandmaster (holdover time) plus the time from the 1st Sync of the new grandmaster until the clocks in the domain have settled, shall be less than 2 Sec. Goal < 200ms? – need a contribution
    - Note: Spanning tree may cause additional settling time of several seconds

PTP Assumptions

- PTP Clock Quality (802.1AS clock, Ref 2) Continued:
  - When a bridge stops receiving Sync messages, it quickly stops sending Sync messages while the new grandmaster is selected, allowing stations to enter holdover quickly
  - A minimum Sync rate of TBD Syncs (e.g. 10 for 802.3, 3 for 802.11) per second is sufficiently frequent to meet PTP clock quality requirements (to be verified – we know 100/sec works)
  - Various links within a single 802.1AS domain may have differing SYNC intervals (e.g. 802.11 vs. 802.3) but we will specify the minimum interval per medium so that processing load will be understood.
    - Fix the values in 802.1BA for different profiles
  - A bridge will send Syncs at accelerated TBD rate for TBD sync messages when a link comes up or when a node becomes the new grandmaster?
    - guidance needed in the future – should be a 802.1BA profile attribute?
  - A bridge doesn't implement Best Master Clock, but a device may include both a bridge and station functionality
  - 802.3 PHY latency variation is as specified in (Ref 2)
    - No more than 86 ns per Hop or 43 ns per PHY (i.e., portion that is not known and corrected for)

PTP-BMCA assumptions

- No pre-master state
- Every bridge will support BMCA and will create an AS spanning tree
- BMCA will be described using RSTP formalisms, with a possible annex using 1588 formalisms
- Path cost using sum of squares or simple sum?
- Does it help significantly to run synchs and announces at a higher rate temporarily when a new GM goes active
- Need default BMC priorities to keep 802.11 STA from becoming GM
- Loss of GM is detected by a lack of synch msgs, no announce receipt timeout
- Announce msgs are sent immediately on loss of GM with no randomization
- No 2-msg qualification of announce msgs from potential masters

PTP Assumptions

- Spanning Tree
  - A spanning tree (not part of AS?) will eliminate routing loops in an AVB cloud for PTP
  - Sync messages are passed along a spanning tree
  - Only on 802.3 PathDelay messages are transmitted/received even on ports blocked by spanning tree
    - AS must drop messages it receives that came from itself
    - If multiple responses are received on a port from a single request, the link is considered to be outside the AS Domain
    - If the measured link delay exceeds a per link TBD value then the link is considered to be outside the AS Domain (buffered repeater issue - a 802.1BA issue)
    - Until a response is received the link is considered to be outside the AS Domain
    - Hence there is no need to use AB (LLDP) as a means of detecting the boundary
  - PathDelay information is then provided to RSTP to establish the link cost (which may cause a change in the Spanning Tree to avoid buffered repeaters)
    - exact thresholds and preferences are set by an 802.1BA profile

PTP Assumptions

- 802.1AS **does not modify** frames on-the-fly (i.e. **all clocks are two-step clocks**)
- An 802.1AS domain is a set of bridges/end-stations that meet the 802.1AS requirements that are directly connected or associated via links that are operating as 802.1AS links
- The 802.1AS protocol packets are not propagated beyond the 802.1AS domain, except for determining the edge of the 802.1AS domain
- The timing reference plane is defined as the interface between the PHY and the physical medium
- The point in the message corresponding to the beginning of the first symbol following the start of frame delimiter (SFD) is a distinguished point termed the Message Time Stamp point

PTP Assumptions

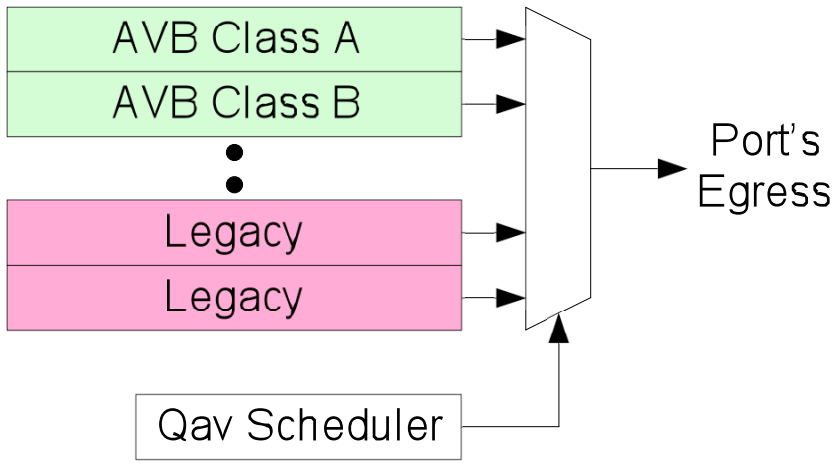
- Those clock components associated with both an end-station and bridge-station are the same
- All potential grandmasters must have a Clock Target function – even those grandmaster's with a clock traceable to TAI
  - A Clock Source-only device is disallowed, because it introduces possible time discontinuities when the grandmaster changes, and because a clock target is needed to measure the time and frequency offsets

PTP Questions

- Bridges will introduce less than TBD Residence Time?
  - probably a profile-specified value

PTP/MGT/Qav Priority

- All stream frames that are ready are transmitted ahead of all other frames (even MGT/PTP … e.g. BPDUs)



Qav Assumptions

- Bridges do per class shaping
- Talkers need to do both per stream and per class shaping
  - Streams shall be shaped prior to reaching the Talker's MAC
  - In the absence of congestion (i.e., interfering traffic), Talkers will not burst AVB stream frames (i.e., they will transmit 1 frame then an extended IFG to even out the frame spacing)
- Policing at ingress is not required and will not be specified
- Talkers are required to follow the shaping algorithm that is specified for Bridges (i.e., that which is defined in the 802.1Qav drafts)
- Burst Credits do not accumulate in the absence of congestion
- The queue draining algorithm does not use 802.1AS's (or 1588) time or rate; therefore the Scope of the project is being revised

Qav Questions

- How will Qav work with Qaz (i.e., enhanced transmission selection)?
  - We will work with Qaz to make sure we don't conflict in terms of "bandwidth available" (managed Class A/B limits)
- How will Qav's domain be determined?
  - There is no Qav domain!

BA Assumptions

- Talkers are required to do this…
  - All Talker capable devices shall be Grandmaster capable
- Bridges are required to do this…
  - May be Grandmaster capable
- Listeners are required to do this…
  - May be Grandmaster capable
- Profiles
  - Consumer plug-and-play - Kevin
  - Professional A/V - Gael
  - Industrial – Franz-Josef
- The edge of an AVB cloud is detected with 802.1AS mechanisms
  - For Ethernet, the path delay portion of PTP can be run without a Grandmaster thus buffered repeaters can be found
  - For other layer 2 technologies, there are layer 2 specific mechanisms
- An AVB cloud must have exactly one AS domain
  - Don't want an AVB cloud using different time services

CE-BA Assumptions

- Bridges must snoop IGMP on ports that go to a Listener and then generate MMRP requests and also proxy for the listener ready?

Professional-BA assumptions

- MMRP required to support talker pruning
- Very high performance Qat operation
  - 20ms through 7 hops of Ethernet
- Do we need to specify a mechanism to allow a Listener to move from one Talker to another Talker without using a pre-reservation (since a pre-reservation cannot be done)? Do we need to synchronize the change in the Network?

BA Questions

- How will MSTP select an SRP path over a CM (Congestion Management) path or a non-SRP/non-CM path using 'out of the box' defaults?
  - For AVB with non-AVB devices: Use MSTP with at minimum one spanning tree instance and set AVB to AVB path costs low (match terms in capability vectors) and playing with root costs using MSTP's priority vector?
  - May need a new PAR to define how these vectors are defined?
- A TPMR needs to support 802.1AS between its ports if it wishes to be an AVB link (or it will look like a buffered repeater)
- Do we generalize the stream MC address acquisition (1722 is looking at using this by requesting a block of addrs), or use a new layer 2 DHCP type of server, or allocate bridges a unique ID to use in a MC OUI?
  - Start out with a request & claim, and then you get a response which could be a defend or it's an answer from a server that will indicate what address(es) to use. A bridge that is also a server can prevent your request/claim from being distributed throughout the network.
    - Use an MRP block address for this request protocol?

802.11 Requirements

- Link Speed
  - 802.11: 100 Mbits or faster?
- Link Duplex
  - 802.11: Half Duplex (no choice)
- Maximum Frame Size (Layer 2 bytes – DA through CRC)
  - 802.11: 1088 bytes for AVB Class A (depending on the speed of the link)? 1522 bytes? Or 2000 bytes for Legacy frames? (i.e., no Jumbo frame support)
- Link Level Flow Control
  - 802.11: Won't be allowed
- Precise Timing Protocol
  - 802.11v location/time measurement is required
- AVB 802.11 devices must support 802.1 bridging

Other Issues

- 802.1AE Environments
  - Any AVB Streams and PTP & SRP frames can be AE Tagged
  - Implementation detail … but we already want 802.3 to give us timing information!
- PONs are currently not specifically supported?
  - i.e., PON support is dependent on contributions from those that need it and may need to become part of a separate PAR
- Provider Networks are currently not specifically supported?
  - i.e., Provider Network support is dependent on contributions from those that need it and may need to become part of a separate PAR
- Other Assumptions …
  - (this is a growing work in process)

References

- Ref 1: at-pannell-policies-0707-v04
  - Title: 802.1 Qat Policies Proposals
- Ref 2: as-garner-assumptions-for-error-sources-time-synch-0507-v03
  - Title: Assumptions for Sources of Time Synchronization Error in IEEE 802.1AS
- Ref 3: avb-garner-requirements-summary-r4-060217
  - Title: Summary of AVB Bridging Network Requirements
- Ref 4: avb-dvj-white-paper-060306
  - Title: Residential Ethernet (RE) (a working paper)
- Ref 5: av-mjt-max-delay-0308-v1
  - Title: Max Delay in Ethernet Networks