Audio Video Bridging (AVB) Assumptions
IEEE 802.1 AVB Plenary
April 2008 - Calls

Green Text = Agreed to on Various AVB Calls/Meetings

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

Black Text = Not Decided

Changes Marked with Red from last version
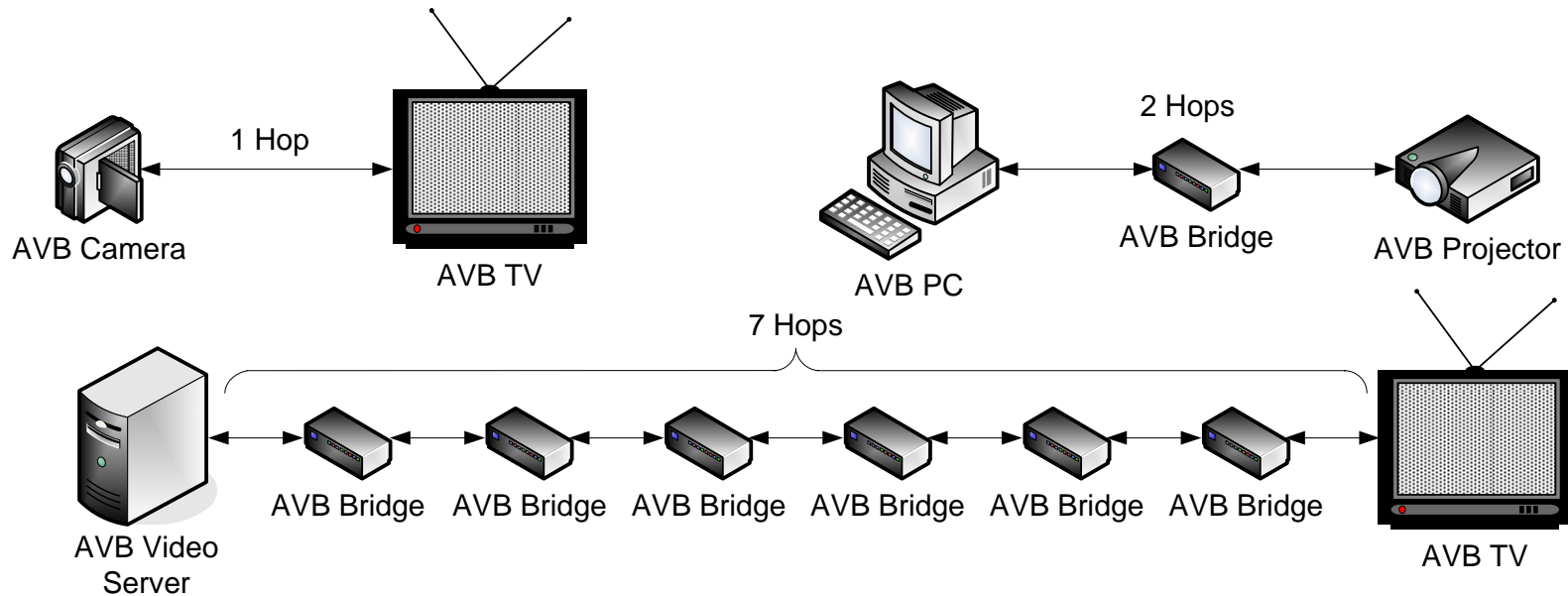
Don Pannell

Marvell

dpannell@marvell.com

Revision History

- 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

1 Hop

AVB Camera

AVB TV

2 Hops

AVB Bridge

AVB PC

AVB Projector

7 Hops

AVB Video Server

AVB Bridge   AVB Bridge   AVB Bridge   AVB Bridge   AVB Bridge   AVB Bridge

AVB TV

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 + Amount of other Class B frames (at Class B allocated data rate)
- Need to define a Traffic Class Measurement Period for bandwidth calculations
  - This results in Talkers using smaller frames – 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
  - ~~Need a knob to limit the % of link bandwidth that can be reserved for Class A?~~
- 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
- 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) cannot be used on AVB traffic or protocols

AVB Frame Format

- ## 802.1 Q Tagging
  - All AVB Streams will be Q Tagged
  - 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 Ether type 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 – Ref 1)
  - An AVB Stream Frame is any Tagged Frame with a designated AVB priority entering an AVB port
  - An AVB Port is a port mode used to differentiate the port from being a Legacy port (AVB ports are part of the AVB Cloud, Legacy ports are at the edge of the AVB Cloud, not connected to an AVB aware device)

- Stream Policy (for Queue Usage in Bridges – Ref 1)
  - Only those frames that meet the Stream Identification (above) can be placed into the AVB Egress Queues
  - All other frames are placed into the Legacy Egress Queues
  - Frames that contain a designated AVB priority entering a non-AVB port (i.e., a Legacy 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 AVB 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
- 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

- AVB 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 AVB Class to its frame Priority is in the MIB
  - How is this going to be signaled? LLDP? What is needed here?

- Hook to existing higher layer protocols without imposing any new requirements on the higher layer protocols

- Detailed diagnostics about the AVB cloud 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)

  - Can't increase the bandwidth when requested (report remaining bandwidth)

  - Out of MMRP database resources

  - Requested Priority is not an AVB Class in this device

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 Acquition Protocol (Annex C in 1722)
  - It is in scope of SRP to allocate DA's, but not sure its necessary
- 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 – but the feature may not make the 1st iteration of the spec
  - 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
- 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) – but the feature may not make the 1st iteration of the spec
- The maximum time to make or break an SRP reservation is: ???
  - Is this defined per hop assuming a max of 7 hops?

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

    – Need to detect ports connected to Buffered Repeaters?

        - To set these ports as SRP edge ports (i.e., non-AVB ports)
        - Use 802.1AS's delay? How will this be connected?
        - Do we need a MIB to enable this or not?

SRP Assumptions

- The figure below is a valid AVB Cloud
- Therefore, SRP Needs to Count Hops and report the Max Latency or Port Speed reported from all the Hops
  - So a Listener can determine if it wants to listen to the stream as it may not meet the AVB Latency and/or Clock Quality Goals

Red to Red = 7 Hops
Green to Green = 7 Hops
Blue to Blue = 7 Hops
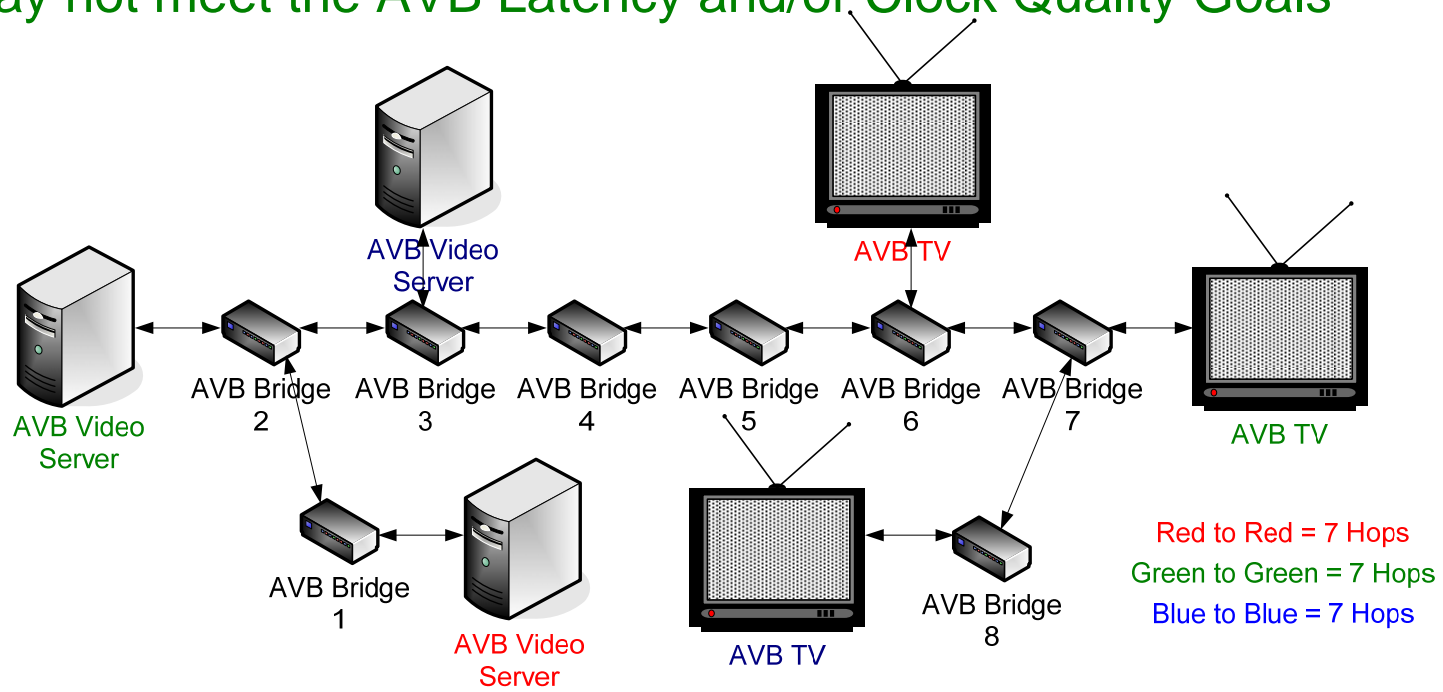
SRP Questions

- Are Talker reservations propagated everywhere or only on ports where the destination address has been registered by MMRP? Doing so will allow faster re-convergence at the expense of unnecessary Talker registrations in the rest of the network.

- When there is a change in the network, it is better to interrupt the flow for longer or transmit them at best effort, or oversubscribe the AVB queues?

- How do we handle the case where a Bridge or Listener is out of Address Resources?
  - Or it can't handle the current address but it can support others

- Do Listeners need to run MMRP? Maybe no, because IGMP is there instead? But this puts the listener outside the AVB cloud and there is a problem if the 'last' link has a buffered repeater (transparent switch) in the path. Or get the stream as far as we can to a 'Listener Proxy' port on an AVB Bridge?

- Talkers must be MSRP aware? Yes..
  - Listeners?

- Does the SRP ACK's need to report the upper and lower bounds of what a path can support (to support 802.11)?

SRP Questions

- Need to specify a sublayer interface in Qav to respond to T-Spec requests (e.g., is 802.1AE enabled,…)?
- What is an SRP domain & its forwarding rules?
  - Is 802.1AS required if timing is not needed?
  - In the default/non-engineered case 802.1AS is needed to detect 'transparent bridges' however
- What is an AVB domain?
  - Assume it to be a subset of an 802.1AS domain
- What is an 802.1AS domain?
- AVB streams must reside entirely inside an AVB Cloud? No…
  - SRP PDUs are confined within an AVB Cloud
  - Guarantees are made only within the Talker's AVB Cloud
  - Stream data can go anywhere

SRP Questions

- What is the Class B measurement period?
- Is the bandwidth measured over a continuously moving window ~~equal to the traffic class interval~~, or is it measured using a fixed window (one that has fixed boundaries in time, although the "phase" of those boundaries might be arbitrary)? Needed for certification.
- Similarly, what do we mean by "max packets in an interval" is this the maximum number of packet starts in an interval, or what?
- T-Spec Parameters?
  - Traffic Class Measurement Period is: 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?

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 $\geq$ 25 MHz clock
  - Less than 4 ppm per Second drift (crystal drift)
  - 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 <= +/- 500ns after the 802.1AS to AES3 filter? (either this or the previous 500ns must change)
  - 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 < 100ms? – 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 and maximum per medium?
    - Fix the values in 802.1BA?
  - 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
  - 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 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 the spanning tree
  - 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 TBD then the link is considered to be outside the AS Domain (buffered repeater issue - is this an 802.1BA issue?)
    - Until a response is received the link is considered to be outside the AS Domain
  - PathDelay information is then provided to spanning tree to establish the link cost (which may cause a change in spanning tree to avoid buffered repeaters)

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 contiguous set of bridges/end-stations that meet the 802.1AS requirements

- The 802.1AS protocol packets are not propagated beyond the 802.1AS domain, except for determining the edge of the 802.1AS domain

- The 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

- Frames received that are smaller than the expected size based on the fields expected to be contained, are discarded
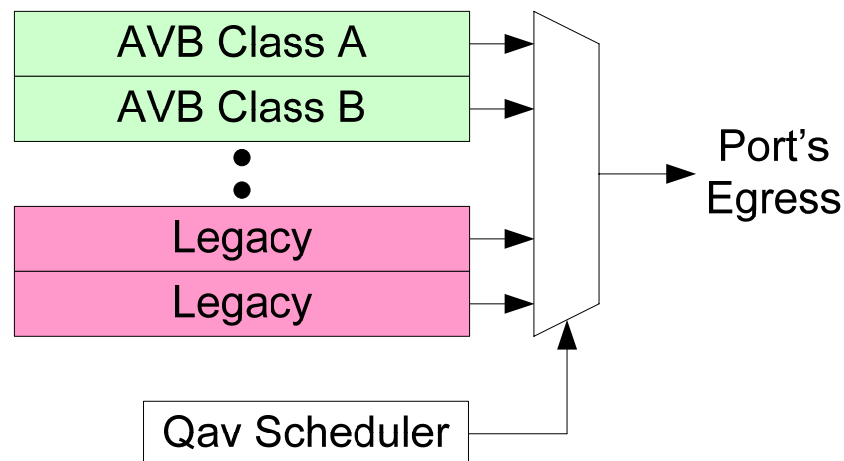
PTP Assumptions

- 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 Assumptions

- 802.1AS can detect the edge of an AS region by means of protocol activity – as it will send PDUs to measure path delay. Lack of response to these, or an excessive delay measurement value, would both be indication of the presence of a boundary.

  - Hence there is no need to use AB (LLDP) as a means of detecting the boundary

  - The excessive delay measurement value for an AS boundary will be defined in 802.1BA

- Bridges will introduce less than TBD Residence Time?

PTP/Qav Priority

- Priority Models (do we need to spec this or just spec the latency)?
  - Concern is where to put PTP and/or Management (BPDU type) frames
  - Assume PTP is a MGMT (Management) frame & all MGMT are given the same scheduling priority?
    - At what point do MGMT frames get MUX'ed to the Port's Egress?

Qav Assumptions

- Bridges do per class shaping
- Talkers need to do both per stream and per class shaping (i.e., what's in the Talker's MAC)
  - Streams shall be shaped prior to reaching the Talker's MAC
- Policing at ingress is not required and will not be specified
  - Per Flow Ingress policing is not possible at layer 2 for unicast flows
- 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)
- 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., pacing for congestion notification)? Do we need to worry?
- How will Qav's Auto SRP bit get set (to determine the edge of the Qav cloud)? Qat requests + AS results (to find the buffered repeaters)?
- How will Qav's domain be determined?
  - Is it different from Qat's domain? Probably not

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
  - Class A & B capable devices or ports
  - Class B only capable devices or ports
- The edge of an AVB cloud is detected with 802.1AS path delay measurements
  - The path delay portion of PTP can be run without a Grandmaster thus buffered repeaters can be found
- The AVB cloud is the intersection of exactly one AS domain and exactly one Qat domain
  - Don't want an AVB cloud using different time services

BA Assumptions

- Talkers and Listeners (along with Bridges) must support MMRP?
  - Necessary for IEEE 1722 as IGMP can't work there due to its being Layer 2
- A Bridge must support MSRP Pruning but its default is off (i.e., no pruning)?

BA Assumptions for Existing Stuff

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

BA Questions

- MSTP is the best place to determine like minded devices (for AVB and CN) and this should be handled in the BPDUs?

- 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)

- 4/2/08 call - Specify STP for 802.11 so that it doesn't suffer from the time delays normally encountered with 'dumb yellow coax shared media' so reconfiguration is faster than 2 to 6 sec?

- 4/2/08 call - Use Bonjour for 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?

  - 4/2/08 call - 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

Other Issues

- 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?

- How is the edge of the 'stream' cloud determined (i.e., the cloud that uses Qat and Qav)? Fill this out and add to BA Assumptions

- Use 802.1AS to detect 'transparent bridges' and then indicate that the port it is connected to is not 'AVB Capable'?
  - Where does this fit into the building of STP?

- 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? Is the same issue as on Slide 11 2nd bullet?

Other Issues

- BA part: To establish an AVB Cloud? Need to define.

- 802.1AE Environments
  - Any AVB Streams and PTP & SRP frames can be AE Tagged
  - Need to understand implications of MacSEC on PTP Timestamping
    - Will Need Help with 802.3 on this

- 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