Audio Video Bridging (AVB) Assumptions

IEEE 802.1 AVB Plenary

Jan 28 & 29, 2008 – Los Gatos

Green Text = Agreed to on Various AVB Calls/Meetings

Black Text = Not Decided

Changes Marked with Red from last version

Don Pannell

Marvell

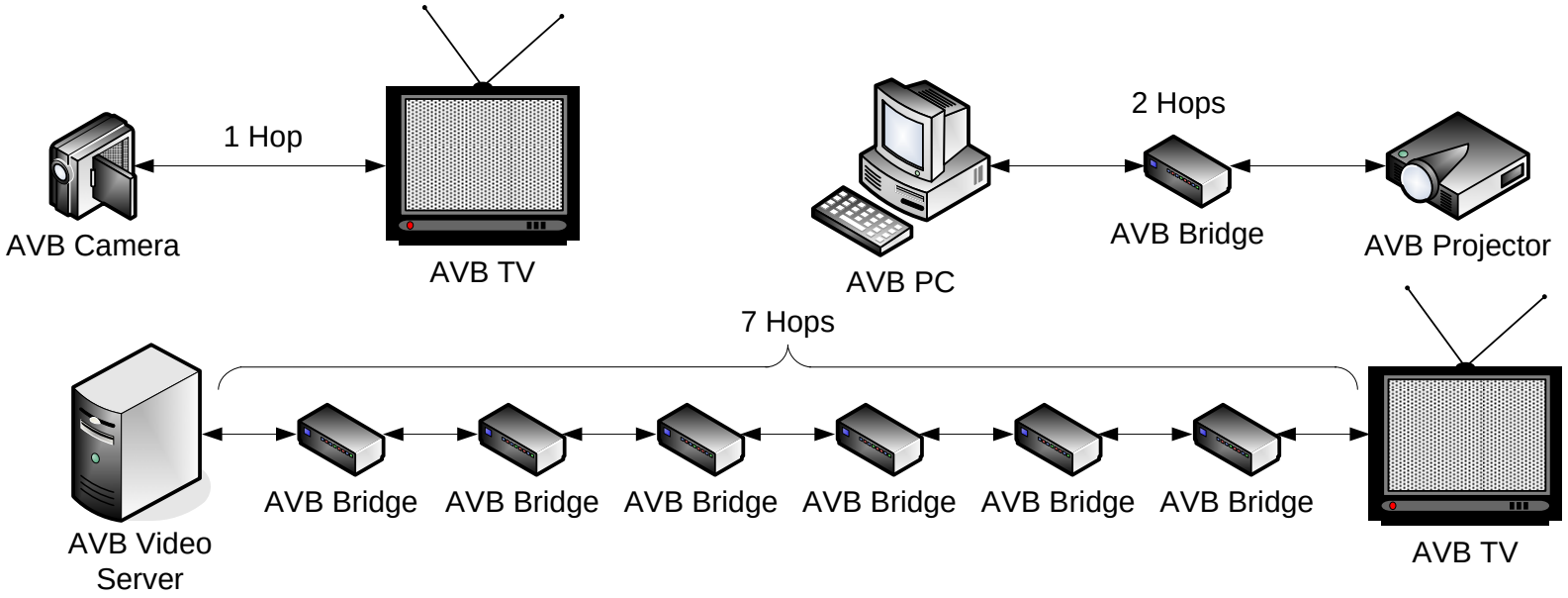
dpannell@marvell.com

Revision History

- Avb-pannell-assumptions-0108-v11: 01/28/08 Los Gatos
- Avb-pannell-assumptions-1107-v10: After many calls – 11/13/07 Atlanta
- 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)
  - 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: Class A: 0 to 2 mSec – Class B: 0 to 20 mSec?
- Maximum Interference Amount per Hop
  - Class A: 1 Max size frame
  - Class B: 1 Max size frame + 1 Max size Class A burst
- 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
- Link Duplex
  - 802.3: Full Duplex only (i.e., no half duplex support)
  - Works with clause 43 Link Aggregation (data and PTP)
- 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

    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

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)
    - These Q Tag PRI defaults may change – voice your wish!

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 priority
- A Reservation Identifier for SRP (i.e., Talker MAC Address along with 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)
- Variable Bit Rate streams need to reserve the peak bandwidth that they need
- Reconfiguration of a stream's bandwidth is allowed

SRP Assumptions

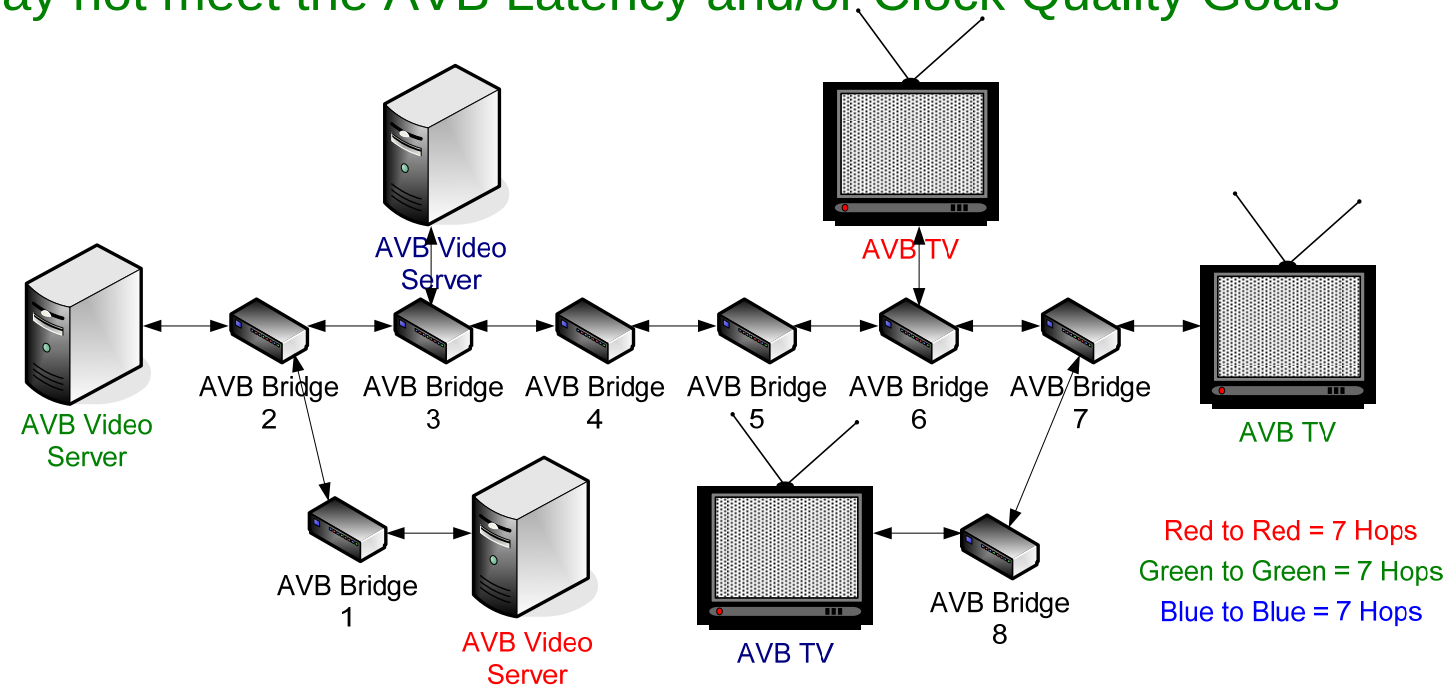
- AVB Frame Priorities are changeable
  - These are changeable for 'engineered' networks
  - For Home networks the set of specified defaults simply must work!
- The mechanism for changing the association of an AVB Class to its frame Priority is in the MIB
- Hook to existing higher layer protocols without 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 piggy back on 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 (eg., Spanning Tree change) by re-converging after the topology change
- SRP will respond to changes in usable bandwidth (eg., wireless and/or coordinated shared networks) by recalculating
- Need to report error reason(s) to affected Listeners 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

- 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



SRP Questions

- 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
- How are Stream MAC DA's going to be Allocated?
  - By the Talker, and how the Talker gets it is:
    - From a higher layer if its IP, From ??? if its not IP (IEEE 1722 need this)
  - It is in scope of SRP, but we are not sure its necessary
- How will SRP Respond to changes in bandwidth request based on user requests (eg., when the user requests an increase in bandwidth)? By a make before break method by making a request for the new rate
- Need to specify a sublayer interface in Qav to respond to T-Spec requests (e.g., is 802.1AE enabled,…)?
- IEEE 1722 needs a Null Stream ID, would like all zeros?
  - IEEE 1722 will look for a bit to indicate a valid Stream ID or not

SRP Questions

- 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
- 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?

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: ?
  - 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 filter.
  - 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)
  - 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 <= 1 uSec
  - Network settling time at startup: < 2 Sec
  - If a new grandmaster (which is not traceable to TAI) takes over, it takes over using the time of the previous grandmaster, but provides no bound on the PPM change (other than that specified above).
    - Even if it was possible to include it in the Announce, the new grandmaster will not communicate its estimate of the PPM step

PTP Assumptions

- PTP Clock Quality (802.1AS clock, Ref 2) Continued:
  - 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 an estimate of the step change in time
  - Every time a grandmaster introduces a time change, it increments a time-base ID which uniquely identifies the time step. Perhaps the initial time-base ID is a random number
  - 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 2s. ?Goal < 100ms?
    - Note: this doesn't work if a bridge sends Sync even if it stops receiving Sync.
    - Note: Spanning tree may cause additional settling time of several seconds
  - 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

PTP Assumptions

- A minimum Sync rate of TBD (e.g. 10 for 802.3, 3 for 802.11) Syncs per second is sufficiently frequent to meet PTP clock quality requirements (to be verified)
- ?Question: Should a PTP device be allowed to send Sync at accelerated rates for a little while when a links comes up, or when a node becomes the new Grand Master?
- Only master ports initiate Sync or equivalent (need to change the “LocationRequest” direction on 802.11)
- No bridge is required to have a best master selection function (but any bridge but may)
- 802.3 PHY Latency Jitter/Wander 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)
- Various links within a single PTP domain may have differing SYNC intervals (e.g. 802.11 vs. 802.3) but **we** will specify the rate per medium?

PTP Assumptions

- Spanning Tree
  - A spanning Tree (or equivalent) will eliminate routing loops in an AVB cloud for PTP?
    - Should 802.1AS construct its OWN spanning tree, used only for time distribution or should we use an MSTI?
    - NOTE: This is bigger than an 802.1AS problem – also impacts streams
  - Sync messages are passed along the spanning tree.
  - PathDelay messages are transmitted/received even on ports blocked by spanning tree
  - PathDelay measurements are used by spanning tree to establish the link cost
- PTP Clock Discontinuities
  - PTP slave clocks can see discontinuities in Global Time (i.e. step changes in epoch) during media streaming operation
  - Discontinuities may originate from selection of new Grandmaster, or may occur within a clock while it is acting as Grandmaster
  - A step change in epoch may be any value (e.g. usec to years)
  - PTP will inform the application upon change in Grandmaster or other clock discontinuities (note: this will require extra flag in some message from the grand master)

PTP Assumptions

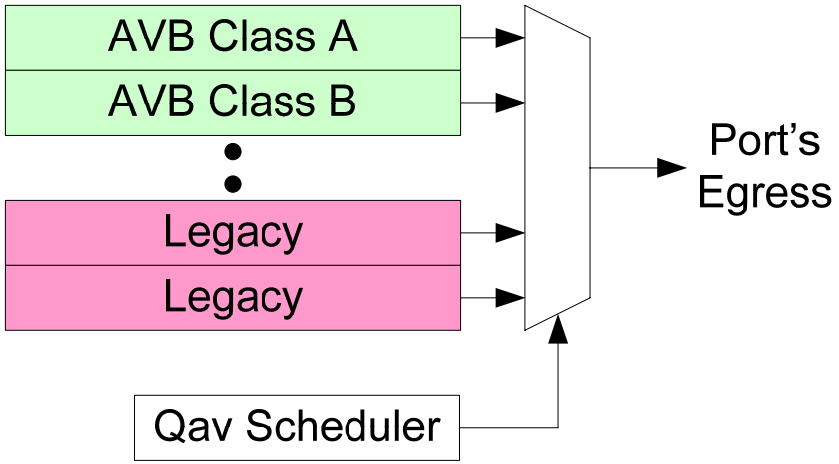
- 802.1AS **does not modify** frames on-the-fly (i.e. **all clocks are two-step clocks**)
- An 802.1AS domain is a contiguous set of bridges/end-stations that meet the 802.1AS requirements
  - 802.1AS domain may be larger than an AVB (SRP) Cloud (move to .1BA)
- The 802.1AS protocol packets are not propagated beyond the 802.1AS domain, except possibly for determining the edge of the 802.1AS domain
- The reference point 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

- The clock components associated with an end-station and bridge-station are the same?
- The end-station clock components can include a clockTarget, traceable clockSource, or clockSource & clockTarget components?
- The functional distinction between traceable and nontraceable sources is the behavior when not currently the grandMaster:
  - A traceable ClockSource behavior does not change.
  - A nontraceable ClockSource sync's to network-supplied time.
- A nontraceable ClockSource-only device is disallowed, because it introduces possible time discontinuities when the GM changes?

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?



PTP Questions

- All Talker capable devices must be Grandmaster capable?
- **MOVE TO 802.1BA:** How do we detect the edge of an AVB cloud?
  - The path delay portion of PTP can be run without a Grandmaster thus ‘transparent bridges’ can always be found?
- MOVE TO 802.1BA section
  - An AVB cloud is contained inside an 802.1AS domain
    - Don’t want an AVB Cloud using different time services

Qav Assumptions

- Bridges do per class shaping
- Talkers do per class shaping (i.e., what's in the Talker's MAC)
  - Streams shall be shaped prior to reaching the Talker's MAC – but this is out of scope for AVB
- Policing at ingress is not required and will not be specified
- 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)

Qav Questions

- How will Qav work with Qaz (i.e., pacing for congestion management)? Do we need to worry?
- How will Qav's Auto SRP bit get set? LLDP + AS results (to find a 'transparent bridge')?

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

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: 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)?
- 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?

Other Issues

- 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)