Summary of A/V Bridging Network Requirements

Geoffrey M. Garner (Consultant)
Gmgarner@comcast.net

SAMSUNG Electronics

IEEE 802.1 AVB TG / IEEE 1588 Design Meeting
Gaithersburg, MD
2006.02.21

Acknowledgment

❑The author would like to thank David V. James for having supplied slides 3 and 4 (i.e., the two slides that follow the current slide).

Requirements

- Inexpensive (insignificant incremental bridge costs):
  - Standard bridge crystals and uP (e.g., 8051) are sufficient
  - Snapshot times are sent within the next frame
  - Same snapshot HW for 1588+ and 802.1-RB
- Efficient: 64-byte frames are preferred (must be less than 72 bytes)
- Accurate:
  - 1us maximum wander from the grand-master station
  - Not response-time dependent
  - Limited only by HW snapshot accuracies
- Robust:
  - Endpoint accuracies are quickly attained (≤ 2 seconds)
  - Concurrent grand-master selection and synchronization
  - No administration required (plug-and-play)
  - Minimal grand-master handover transients
  - Reasonable rogue-frame lifetimes (≤ 2 seconds)
  - Grand-master precedence is higher-level selectable
  - Seconds never overflow and are 1588 compatible
- Sampling times:
  - ~10 ms (100ms ≥ t ≥ 1ms) offset adjustments (sufficient firmware execution time)
  - ~100 ms rate adjustments (rate difference errors are insignificant), work in progress

Assumptions

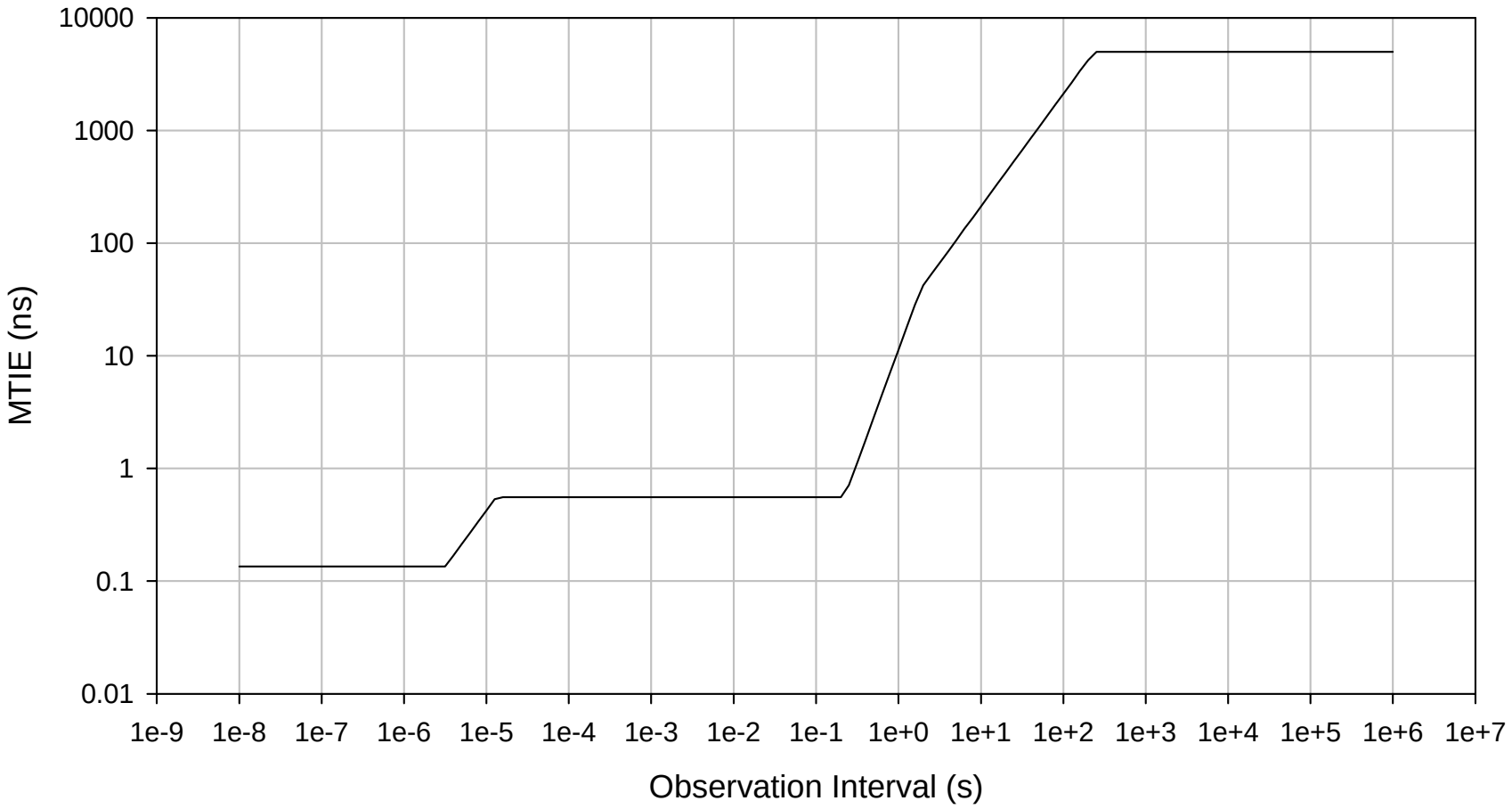
- Bridges have ±100 PPM quartz crystals
- Bridge cascade depth is limited to 7
- The range of 1588 seconds is 80 bits
- Ethernet media:
  - Full-duplex
  - PAUSE is never used
  - 100 Mbs and higher (not 10 Mbs)
- Bridges and endpoints have uPs
  - 10 ms is “doable” by the simplest of uPs
  - 10 ms is “tolerable” by the fastest of PCs

Requirements

- Jitter and wander of time difference between the synchronization signals at any two nodes must within a budget allocation of the MTIE mask on the next two slides
  - This mask is the lower envelope of the jitter/wander accumulation MTIE masks for time-sensitive applications (uncompressed and compressed digital video; digital audio) at the network egress (see backup slides for mask for each application)
  - Note that the AVB Network synchronization gets only an allocation of this requirement
    - for applications that are delivered to the residence via service provider network(s), those networks will also have jitter and wander
    - There will be jitter and wander due to mapping/demapping the application into and out of the AVB network, i.e., due to the finite granularity of the network timing when creating application time stamps
- Time difference from Grandmaster not to exceed 5 μs (application requirement)
- Cost of A/V Bridges should be in same ballpark as cost of present consumer-grade Ethernet switches, routers, or wireless access points (or products that combine these functions) [5]

End-to-End Application Jitter and Wander Requirement

Lower Envelope of Network Interface MTIE Masks for Digital Video and Audio Signals



End-to-End Application Jitter and Wander Requirement

$$\mathrm{MTIE}(S) = \begin{cases} 0.1347\ \mathrm{ns} & S < 3.183\times10^{-6}\ \mathrm{s} \\ 4.231\times10^{4} S\ \mathrm{ns} & 3.183\times10^{-6}\ \mathrm{s} < S \le 1.31316\times10^{-5}\ \mathrm{s} \\ 0.5556\ \mathrm{ns} & 1.31316\times10^{-5}\ \mathrm{s} < S \le 0.22196\ \mathrm{s} \\ 11.27745 S^{2}\ \mathrm{ns} & 0.22196\ \mathrm{s} < S \le 1.87631\ \mathrm{s} \\ 21.16 S\ \mathrm{ns} & 1.87631\ \mathrm{s} < S \le 47.26\ \mathrm{s} \\ 5000\ \mathrm{ns} & S > 236.29\ \mathrm{s} \end{cases}$$

This implies:

- High-band jitter ≤ 0.1347 ns (100 kHz high-pass measurement filter)
- Wide-band jitter ≤ 0.5556 ns (1.434 Hz high-pass measurement filter)
- Frequency Drift ≤ 0.0226 ppm/s for observation intervals > 0.222 s
- Frequency offset ≤ 0.021 ppm for observation intervals > 1.876 s
- Maximum peak-to-peak wander of 5000 ns (5 μs)

Assumptions – Additional Detail

- AVB will have inexpensive Ethernet clock/oscillator
  - 25 MHz (40 ns granularity)
  - Will not be OCXO and extremely likely not TCXO; may be possible to bound noise generation (but bound will be loose)
- AVB will have inexpensive processor, for which timing/synch functions will be a small subset of all its functions
- Low cost requirement implies it will likely not be feasible to have special hardware at the PHY to allow accurate on-the-fly time stamp measurement
  - Time stamp measurement will very likely need to be sent in subsequent frame
  - Time stamp measurement accuracy will be no better than 40 ns
- Low cost requirement implies that a solution should allow any expensive filtering to be done at end device (and therefore have cost associated with the application that needs it); expensive filtering should not be required in the bridges

Tradeoffs

- Tradeoff between bridge oscillator phase error (due noise generation, temperature changes, granularity, and time stamp measurement error), sync interval, and endpoint filter bandwidth and gain peaking
  - Narrower bandwidth for endpoint filter implies that (1) oscillator must be more stable to obtain the same noise generation, and (2) transient response will be longer or more expensive design with fast-acquisition mode will be needed
- Different compensation schemes have different requirements on information that must be exchanged (e.g., both free-running and frequency corrected phases versus just frequency-corrected phases; cumulative and differential information versus only one or the other)

References

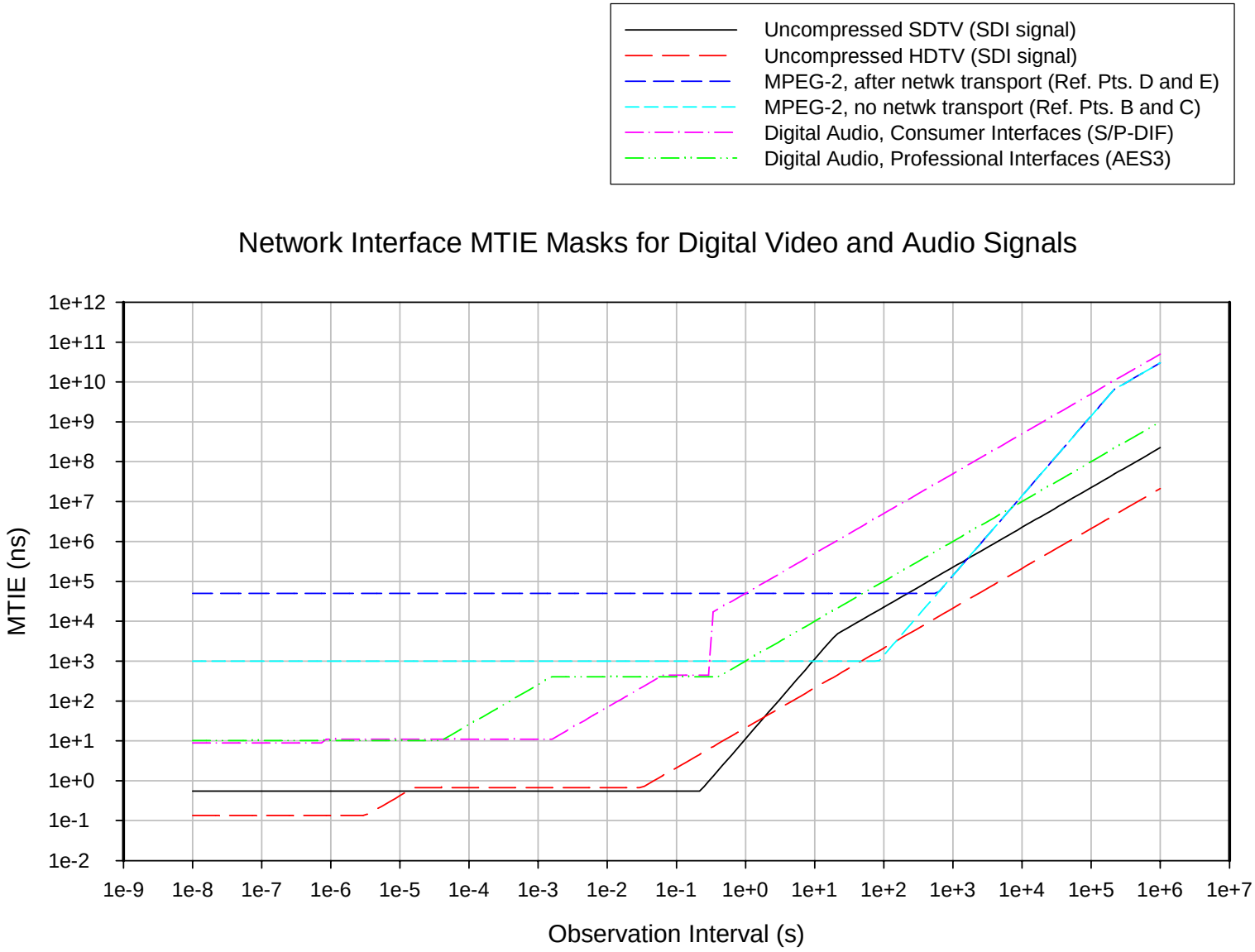
1. Geoffrey M. Garner, *Description of ResE Video Applications and Requirements*, Samsung presentation at May, 2005 IEEE 802.3 ResE meeting, Austin, TX, May 16, 2005.
2. Geoffrey M. Garner, *Description of ResE Audio Applications and Requirements*, Samsung presentation at May, 2005 IEEE 802.3 ResE meeting, Austin, TX, May 16, 2005.
3. Geoffrey M. Garner, *End-to-End Jitter and Wander Requirements for ResE Applications*, Samsung presentation at May, 2005 IEEE 802.3 ResE meeting, Austin, TX, May 16, 2005.
4. Michael Johas Teener, *Residential Ethernet Study Group Closing Plenary Report IEEE 802.3*, San Francisco, CA, July 21, 2005
5. *Timing and Synchronization for Time-Sensitive Applications in Bridged Local Area Networks*, Draft 5 Criteria, November 16, 2005.
6. *ATSC Implementation Subcommittee Finding: Relative Timing of Sound and Vision for Broadcast Operations*, ATSC, Doc. IS-191, 26 June, 2003.

Backup

Detail of End-to-End Jitter Application and Wander Requirements

Requirement	Uncompressed SDTV	Uncompressed HDTV	MPEG-2, with network transport	MPEG-2, no network transport	Digital audio, consumer interface	Digital audio, professional interface
Wide-band jitter (UIpp)	0.2	1.0	50 µs peak-to-peak phase variation requirement (no measurement filter specified)	1000 ns peak-to-peak phase variation requirement (no measurement filter specified)	0.25	0.25
Wide-band jitter meas filt (Hz)	10	10			200	8000
High-band jitter (UIpp)	0.2	0.2			0.2	No requirement
High-band jitter meas filt (kHz)	1	100			400 (approx)	No requirement
Frequency offset (ppm)	±2.79365 (NTSC) ±0.225549 (PAL)	±10	±30	±30	±50 (Level 1) ±1000 (Level 2)	±1 (Grade 1) ±10 (Grade 2)
Frequency drift rate (ppm/s)	0.027937 (NTSC) 0.0225549 (PAL)	No requirement	0.000278	0.000278	No requirement	No requirement

Detail of End-to-End Jitter Application and Wander Requirements



Inter-Stream Synchronization Requirements

- Time synchronization requirements of different audio/video streams for acceptable QoS, for several applications (see [3] and Reference [42] cited in [3]; for Lip-synch, see [6])
  - Tightly coupled audio (e.g., audio streams delivered to multiple speakers)
    - ±10 μs
      - This is a real requirement for certain digital speaker applications for which there is a precise phase relation among the acoustic signals from the various speakers (including the effects of walls; the phase relation is maintained adaptively)
      - In any application involving multiple digital speakers, the difference in synchronization between multiple speakers must be stable within ±10 μs
  - Lip-synch [6]
    - Sound should not lead video by more than 15 ms
    - Sound should not lag video by more than 45 ms
  - Tightly coupled video with accompanying audio (e.g., music with displayed notes, for tutoring purposes)
    - ±5 ms
  - Other examples, and detailed description of experiments, given in Reference [42] cited in [3]