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
2006.02.21
Summary of Requirements

- Jitter and wander accumulation for time-sensitive applications (uncompressed and compressed digital video; digital audio) at the network egress must be within the respective MTIE masks on slide 3 (derived from the requirements on slide 4; see [3], and background in [1] and [2])
  - Note that the AVB Network gets only an allocation of these requirements for applications that are delivered to the residence via service provider network(s)
  - For applications that contain multiple streams, time synch of the streams (inter-stream synch requirements) must be on the order of several ms, and possibly as stringent as 10 µs (but no more stringent than this); see slide 5 for details [3]

- Requirements must be met for an application whose streams traverse up to 7 hops [4]
  - This is an assumption on the maximum expected network diameter; the total number of bridges in the network may be larger

- Possible maximum latency requirement of 2 ms [4]
  - may be relaxed by several ms depending on implications for bridges and applications (there was discussion of this at the September, 2005 interim meeting)

- Cost of Audio/Video 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]

- Minimal or no administration required by users; bridges should be “plug and play” and self-configure (including GM selection)
End-to-End Jitter and Wander Requirements

Network Interface MTIE Masks for Digital Video and Audio Signals

- Uncompressed SDTV (SDI signal)
- Uncompressed HDTV (SDI signal)
- MPEG-2, after network transport (Ref. Pts. D and E)
- MPEG-2, no network transport (Ref. Pts. B and C)
- Digital Audio, Consumer Interfaces (S/P-DIF)
- Digital Audio, Professional Interfaces (AES3)
## End-to-End Jitter and Wander Requirements

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

Digital audio, professional interface
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 (note: it was indicated in 1/25 AVB call that 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). It needs to be decided how precisely to describe this situation in this VG.)
  - Lip-synch [6]
    - Sound should not lead video by more than 15 ms
    - Sound should not lag video by more than 45 ms
  - Video animation with accompanying audio
    - ±80 ms
  - Other examples, and detailed description of experiments, given in Reference [42] cited in [3]
Additional Assumptions

- 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 improve time stamp measurement accuracy

- 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 to noise generation, temperature changes, granularity, and time stamp measurement error), sync interval, and endpoint filter bandwidth and gain peaking.

- **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


