Possible re-write of Clause 5.5 (now Clause 5.3 in 1722a) on Timing and Synchronization

Note: This is a work in progress, but is offered to the working group to stimulate discussion. In the process of working on text for the Common Timing Grid to support new diagrams, I realized that some of the concepts in 1722-2011 are either confusing or inconsistent with the CTG concepts

5.5 Timing and synchronization

5.5.1 General

Because AVTPDUs are carried on asynchronous MAC layers (such as IEEE 802.3 Ethernets with IEEE 802.1 bridges, and IEEE 802.11 "Wi-Fi" networks), any AVTPDU is subject to a variable and indeterminate transit time between Talker and Listener. The quality-of-service mechanisms provided by <IEEE 802.1Q-2011 SRP and FQTSS> allow the transit times of AVTPDUs to be minimized, and—importantly—bounded for a given network size and topology. A key principle <feature?> of time-sensitive networking is to account for indeterminate transit times and facilitate re-synchronization of each AVTPDU's payload data at the Listener.

5.5.2 AVTP presentation time

AVTP defines the concept of presentation time to account for indeterminate transit times <across the network> and to support re-synchronization of payload data at the Listener.

The presentation time represents the gPTP time at which a designated[1] media sample or event within the AVTPDU payload is to be rendered, or otherwise consumed, by the time-sensitive application within each Listener. When multiple Listeners receive copies of the same AVTPDU, presentation time is the mechanism by which they all can render the media payload at the same time.

The AVTP presentation time is conveyed in the avtp_timestamp field of AVTPDUs. The AVTP presentation time is not required in all AVTPDUs; for example, an AVTPDU carrying a MAAP message <cite ref> does not require precise timing or synchronization.

While time-sensitive streams generally rely on presentation time for proper synchronization, the presentation time might not be valid in every AVTPDU within a given stream. If an AVTPDU contains a valid presentation time, then the tv (avtp_timestamp_valid) bit shall be set to one (1).

NOTE: A Listener can generally keep itself synchronized by processing presentation timestamps less frequently than it receives stream PDUs. Also, when a gPTP grandmaster change occurs on the network, all Talkers should clear their avtp_timestamp valid bits (set to zero (0)) until the changeover has completed, i.e. the new gPTP clock has stabilized across the network.

<5.x.x> Calculation of presentation time

[1] In most cases the designated sample is the first sample in the payload; however, with some formats such as IEC 61883-6 this might not be true (see <ref>) <<Rob note: Is this correct?>>

Robert Silfvast 2/24/14 7:57 PM
Deleted: AVTP defines a presentation time to achieve timing synchronization between a Talker and Listener(s). The presentation time represents in nanoseconds the gPTP time when the data contained in the AVTPDU are to be available to the AVTP Listener(s). ... [1]

Robert Silfvast 2/24/14 8:10 PM
Deleted: contained

Robert Silfvast 2/24/14 8:10 PM
Deleted: stream data

Robert Silfvast 2/24/14 8:11 PM
Deleted: may

Robert Silfvast 2/24/14 8:10 PM
Deleted: timestamp

Robert Silfvast 2/24/14 8:32 PM
Deleted: The AVTP presentation time represents the timestamp of when the media sample was presented to AVTP at the Talker plus a constant, Max Transit Time, to compensate for network latency. Network latency is dependent on the network configuration and speed. Max Transit Time represents the worst-case network latency assumed for a given configuration.

At the instant when a time sensitive application within a Talker generates a media sample or event, the corresponding gPTP time is captured (or otherwise known) within the Talker. This time is denoted as $\mathbf{t_G}$. The sample or event is then inserted into an AVTPDU, along with a presentation time $\mathbf{t_P}$, and queued for transmission[2]. In order to guarantee that the sample or event is ready for consumption within the Listener(s) at time t_P, the difference $t_P - t_G$ must be greater than or equal to the known and bounded maximum transit time between the Talker and all of its Listeners. This difference value is defined as "Max Transit Time" ($\mathbf{t_{MT}}$), and it must account for two components, t_{GQ} and t_{WCL}, as defined below:

t_{GQ} = the time within the Talker between generating the sample (or event) and queuing it for transmission[3]

t_{WCL} = the worst-case latency of the network path between the Talker device's transmission queue and any Listener application receiving the given stream. This includes any time required to transfer the payload data from the network interface to the application buffer at the Listener.

Requirement: $t_{MT} = t_P - t_G \geq t_{GQ} + t_{WCL}$

For a given Talker-Listener pair, it is possible to determine t_{GQ} and t_{WCL} based on knowledge of the implementations within both the Talker and the Listener and the path across the network. The means by which t_{WCL} is calculated depends on network link and bridge parameters, and is outside the scope of this standard[4].

For AVTPUs containing media stream data, it is furthermore recommended that Max Transit Time be set to an integer multiple of the nominal media clock period $\mathbf{t_C}$ for the given stream. The reasoning behind this is based on important concepts described in Section <5.x.x> Common Timing Grid.

Recommended: t_{MT} / t_C = an integer

In summary:

$t_P = t_G + t_{MT}$

Unless otherwise determined, the value for t_{MT} shall be computed as follows:

Table 5.3—AVTP default values for Max Transit Time

Class	Max Transit Time (t_{MT})
A	2.0 msec, rounded up to the next integer multiple of the media clock period
B	50.0 msec, rounded up to the next integer multiple of the media clock period

Robert Silfvast 2/24/14 9:34 PM
Comment [1]: This is the same thing as "max timing uncertainty" in the original 1722-2011 standard. I think this is much less confusing.

Robert Silfvast 2/25/14 9:27 AM
Comment [2]: It might make sense to further break this into two components, one being essentially the "MSRP accumulated latency" and other being the latency from rx queue to Listener's application buffer

Robert Silfvast 2/24/14 8:44 PM
Deleted: It is possible for a Talker and a Listener

Robert Silfvast 2/24/14 9:46 PM
Deleted: the Max Transit Time value to use in a given stream. The mechanism for a Talker and a Listener to negotiate Max Transit Time

Robert Silfvast 2/24/14 11:25 PM
Deleted: Unless otherwise negotiated between the Talker and the Listener, the Max Transit Time to calculate the AVTP presentation time is defined by the Default Max Transit Time in Table 5.3

Robert Silfvast 2/24/14 11:16 PM
Deleted: .

Robert Silfvast 2/24/14 10:05 PM
Deleted: ms

[2] In many of the defined AVTP formats, an AVTPDU encapsulates not just one, but multiple, media samples or events for the sake of improving payload/header efficiency. The waiting-time delay imposed by collecting and buffering multiple samples, prior to forming a packet, is outside the scope of this section and should simply be treated as an additional delay in the data path, above and beyond the Max Transit Time described herein.

[3] Note that either t_{GQ} or t_G (but not both!) should account for any buffering delays imposed by collecting multiple samples if the AVTP format uses such a scheme.

[4] The interested reader should consult <IEEE 802.1Q-2011 SRP and FQTSS etc>

5.5.3 timestamp_uncertain

Although the gPTP time is intended to be stable, it is possible for discontinuities to occur in the gPTP time. These can be caused by events such as changes in the identity of the gPTP Grandmaster clock, changes in the timing source of the Grandmaster clock, or other events encountered by a Talker.

When a Talker detects a discontinuity in gPTP time, it should set the timestamp_uncertain field to one (1). This indicates to the Listener(s) of the stream that the AVTP presentation times contained in the avtp_timestamp field may, for a limited period of time, not correspond to the gPTP time. This flag suggests to the Listener that it should take appropriate action. Once the Talker has determined that the gPTP clock has returned to a normal state, the timestamp_uncertain field should be reset to zero (0).

When a Listener detects that the timestamp_uncertain field has been set to 1 or detects a discontinuity, either from a gPTP indication or simple observation, it should stop attempting to correlate AVTP presentation time to gPTP time until the source of the discontinuity has ceased and the Listener has allowed the appropriate time for the gPTP clock to stabilize.

The time required for the gPTP clock to stabilize is implementation dependent and therefore outside the scope of this standard.

5.5.4 Presentation time measurement points

For the purpose of improving clarity, this clause from IEEE 1722-2011 has been removed and replaced by the descriptions and diagrams in Sections <5.x.x> through <5.x.y> above. Despite these changes to the document, the general principles remain the same and backward compatibility remains intact, except for the default values of Max Timing Uncertainty in Table <5.x.x> which did not include the round-up component in IEEE 1722-2011[5].

[5] Talkers not using the rounded-up default values for Max Transit Time might suffer incompatibilities with <p1722a> compatible implementations only if a Common Timing Grid is utilized in the system.

Robert Silfvast 2/24/14 9:54 PM
Deleted: NOTE—The Max Transit Time's numbers are a function of the speed, size (hops), and configuration of the IEEE 802.1 AVB Cloud. The Max Transit Time can be determined by the maximum latency per hop calculations as defined in SRP (SRP, subclause 35.2.2.8.6) and is reported by SRP to the Listener as Accumulated Latency. The Default Max Transit Time is set to support up to seven Ethernet (IEEE Std 802.3-2008) hops for SR class A, and up to two WiFi (IEEE Std 802.11-2007) hops plus six Ethernet (IEEE Std 802.3-2008) hops for SR class B.

Robert Silfvast 2/24/14 10:15 PM
Deleted:

Robert Silfvast 2/24/14 9:54 PM
Deleted: there to be

Robert Silfvast 2/24/14 9:54 PM
Deleted: could

Robert Silfvast 2/24/14 9:56 PM
Comment [3]: Shall?

Robert Silfvast 2/24/14 11:26 PM
Deleted: , either from an gPTP indication or simple observation

Robert Silfvast 2/24/14 9:57 PM
Deleted: and the

Robert Silfvast 2/24/14 9:57 PM
Deleted: needs to

Robert Silfvast 2/24/14 9:58 PM
Deleted: n

Robert Silfvast 2/24/14 10:02 PM
Deleted: This subclause defines the measurement points to be used when calculating the AVTP presentation time. In general, the presentation time marks the time that the IEEE 1722 layer transfers the timestamped data to the next processing layer in the stack. ... [2]