Contents

3.8. Audio Requirements

3.8.1. References

RTCA/DO-214A: Audio Systems Characteristics and Minimum Operational Performance Standards for Aircraft Audio Systems and Equipment

EUROCAE ED112a: Minimum Operational Performance Specifications for Crash Protected Airborne Recorder Systems

IETF RFC 3550 RTP: A Transport Protocol for Real-Time Applications

IEEE 1588-2008 Precision Time Protocol (PTPv2)
The AES67 audio networking interoperability standard includes a PTP profile compatible with SMPTE ST2059-2

3.8.2. Bandwidth and Dynamic Range Considerations

The relationship of frequency bandwidth to dynamic range for communications/navigation and interphone interfaces is shown in RTCA DO-214A. It is noted in RTCA DO-214A that full speech bandwidth extends from below 100 Hz to over 8 kHz for an adult male. However, the octave with the greatest energy (30%) is between 300 Hz to 600 Hz for an adult male, and 550 Hz to 1100 Hz for an adult female. The majority of intelligibility is found between 1 kHz and 5 kHz. Reduction in bandwidth from 6 kHz is stated to reduce intelligibility in the presence of noise. Each of these factors contributes to the overall response of 300 Hz to 6 kHz for the interphone interface. Radios generally support audio response between 300 Hz and 3 kHz.

Although RTCA DO-214A recommends no more than a -3dB reduction in power at the pass-band edges, the overall response must remain within a 3dB envelope. The use of extended frequency response elements in the system is encouraged where practical. This minimizes the cumulative effect of cascaded components (microphone, audio system, headphones) on total system intelligibility. For example, a microphone, an audio system, and headphone, each having a 300 Hz to 6 kHz (-3dB) response, may produce a total system response of 600 Hz to 3 kHz (-3dB) given a gradual roll-off in each component. The net effect is a reduction by four in the speech signal bandwidth and its energy.

Implementation of digital filtering can extend the bandwidth as recommended by RTCA DO-214A. Sample rate selection can balance the digital filter complexity versus consumption of network bandwidth. Typical sampling rates for 3 kHz radio paths and 6 kHz interphone paths are 8 kHz and 16 kHz, respectively. Maintaining dynamic range (minimum signal to maximum signal) is critical for maximizing intelligibility. Typical wiring practices in aircraft maintain noise at levels 40dB below the desired speech level with the dynamic range of VHF and UHF audio Absolute and Differential Delay varying from 6dB to greater than 40dB. RTCA DO-214A requires the audio system to have greater than 50dB to 60dB of dynamic range, measured as ((S+N) / N) depending on the environmental category tested to. To ensure system performance requirements are met, this minimum dynamic range must be maintained.

Choosing the appropriate quantization to balance the dynamic range versus consumption of network bandwidth is also important. The greater the dynamic range, the lower the audible background hiss when the system is idle (noise floor). The dynamic range requirement of 60dB (suggested by RTCA DO-214A) may be achieved by using 11 bits or more per sample (using 2's complement). Recommended audio quantization of 16 bits per sample is specified, yielding 90dB of dynamic range.

3.8.3. Absolute Delay

Absolute (Envelope) delays are specified in RTCA DO-214A. An absolute delay value is important to define because of the potential for adverse system affects like halted speech between operators. It is noted in RTCA DO-214A that the person speaking can become confused when the sidetone is delayed given sufficient amplitude. RTCA DO-214A minimizes the effects of latent sidetone by requiring less than 10ms of absolute delay one way. Audio systems with greater latency in their digital audio transmission may use a local (internal) sidetone to meet the specification. However, careful consideration must be given to the microphone to radio/interphone delay and radio/interphone to headphone delay so as to not impede two-way communication. The maximum value for absolute delay (any microphone to any headset) using digital audio is 20ms round trip for digital audio distribution. This value is based on industry studies of the acceptable tolerance of this delay by the speaker.

ED-112A places further restrictions on absolute delay:

The delay in recording the flight crew audio signals from the time of reception at the microphones to the time of recording on the protected recording medium shall not exceed 250 milliseconds.

3.8.4. Differential Delay

RTCA DO-214A mentions the possible case of two radios tuned to the same frequency and summed to the headphone. The resultant signal will be 6dB louder if the signals have identical delay. However, a difference in delay may cause frequency nulls in the spectrum of the summed signal. This may reduce speech intelligibility due to the loss of formants. It may also result in the loss of pure tones such as 1020Hz. It is recommended to force the first null outside the system bandwidth by minimizing the differential delay.

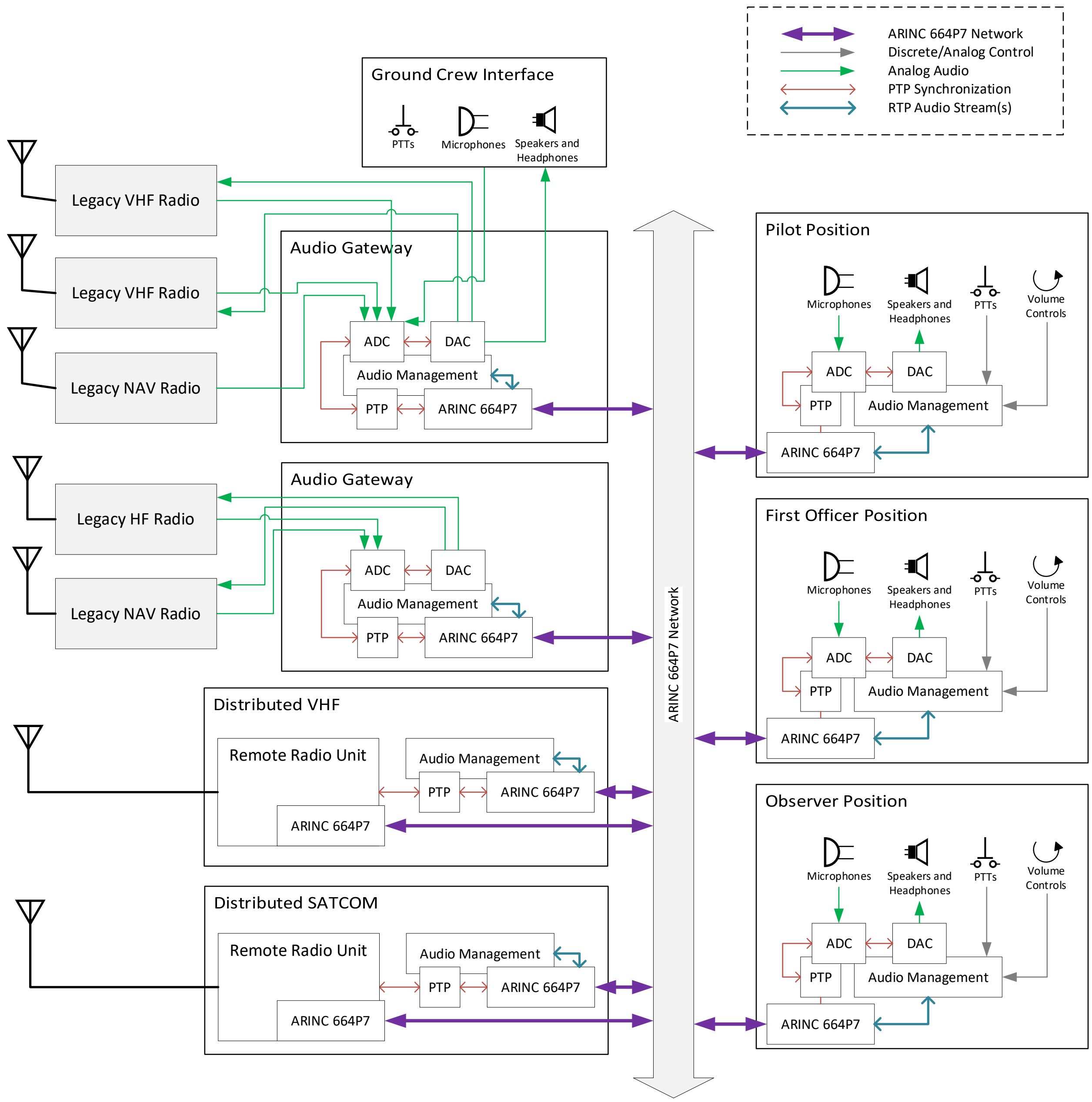
RTCA DO-214A requires less than 9dB of attenuation throughout the frequency bandwidth of the system when two identical signals are summed to the same output. The minimum differential delay to produce 9dB of attenuation can be derived by analysis of the highest frequency in the pass-band. Only the 3 kHz pass-band edge for radio to headphone paths shall be considered since interphone to headphone paths have unique sources all the time. Of the 9dB attenuation allowed, 3dB is consumed by the -3dB pass-band attenuation at 3 kHz, leaving 6dB of attenuation due to differential delay. It requires only one third of a wavelength in time (period) of differential delay to achieve 6dB of attenuation in the summed signal, or 114µs at 3 kHz.

Constant differences in the sampling intervals between asynchronous equipment due to sampling clock "drift" results in timing errors. Timing errors cause performance problems when the samples are too late or too early to properly replicate the audio signal. Accurate sample clocks are required to ensure adequate sampling performance.

3.8.5. Architecture

There are several potential architecture variations, with a few key implementation aspects. This document assumes an ARINC 664P7 network as the digital backbone that carries all Audio and control traffic. The diagram below shows a mixed scenario with two Audio Gateway functions that bridge between the digital audio domain and legacy CNS radio assets. Two digital only examples, a Distributed VHF and a Distributed SATCOM function, are shown below as examples where there is no analog to digital or digital to analog transfer taking place across the function boundary. There is the further potential example where the paired Distributed SATCOM function providing PTP synchronization and Audio Management is connected via the same ARINC664P7 network to the Remote Radio Unit.

On the right side of the diagram, the traditional three Flight Deck user positions is shown, highlighting the digital to analog and analog to digital conversion at each station, synchronized via PTP with all other Audio system assets.



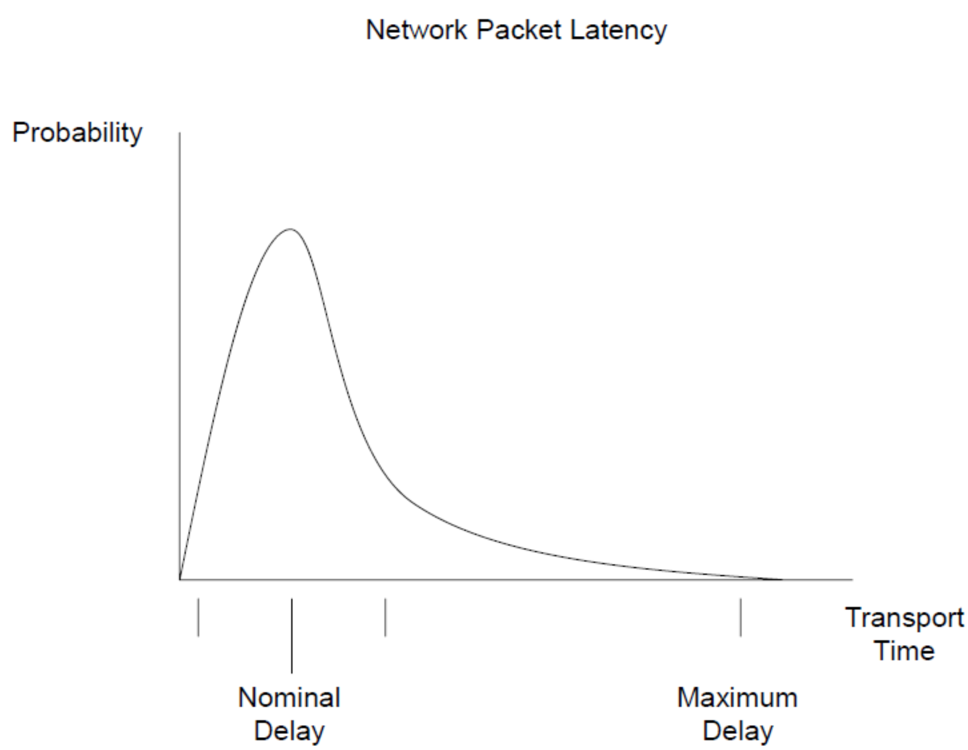
3.8.6. Transport Considerations

Digital audio places additional requirements on the audio distribution method due to the discrete nature of the digital audio sampling process. Digital transmission has the potential to affect audio quality in the event of timing errors. The relationship of these errors to the basic requirements above is fundamental to establishing a digital audio distribution standard for the flight deck.

Real time distribution of digital audio is often referred to as an audio stream. Ideally the digital audio stream is a continuous and steady flow of audio samples like the analog audio signal.

When samples are grouped together to create a network message payload, variable network transport delays may result in timing errors. These timing errors can result in payloads arriving too late to ensure adequate samples are available for replication. Proper buffering is required to ensure adequate streaming performance. These buffers are known as jitter buffers.

Network performance that can bound the nominal transport delay and maximum transport delay will provide predictable digital audio distribution performance. Bounding the maximum network delay allows the jitter buffer to be sized. The combination of the payload time interval, nominal network delay, and jitter buffer size is the audio distribution latency.



Additionally, loss of a group of samples in a network message requires special handling. Without special handling, loss of a single message may result in an audible “pop” or “click”. At the destination, each message must be in the same sequence as it was at the source. If the messages are out of order, the network interface must reorder them. A highly deterministic network is required to minimize the problems due to loss of samples or out-of-order samples.

This document recommends using RTP protocol less QoS, SIP, RTCP (these services are not required as the ARINC 664 P7 network is deterministic.)

Audio streams are multicast when a stream from the publisher needs to be distributed to more than one subscriber.

The Low Fidelity Audio type (RTP Payload Type 101) shall consist of the RTP header with 4 ms of uncompressed 16 bit, 8 kHz samples.

The High Fidelity Audio type (RTP Payload Type 102) shall consist of the RTP header with 4 ms of uncompressed 16 bit, 16 kHz samples.

The Slow Low Fidelity Audio type (RTP Payload Type 103) shall consist of the RTP header with 20 ms of uncompressed 16 bit, 8 kHz samples.

The Audio System shall use a jitter buffer that can accommodate up to 4 ms of network jitter at the subscribing end system.

Maximum network latency shall be less than 2ms (this ensures meeting the 10ms one way differential delay requirement with 4ms sampled by publisher and subscriber.)

The Audio System shall output null audio streams when no audio signal is present.

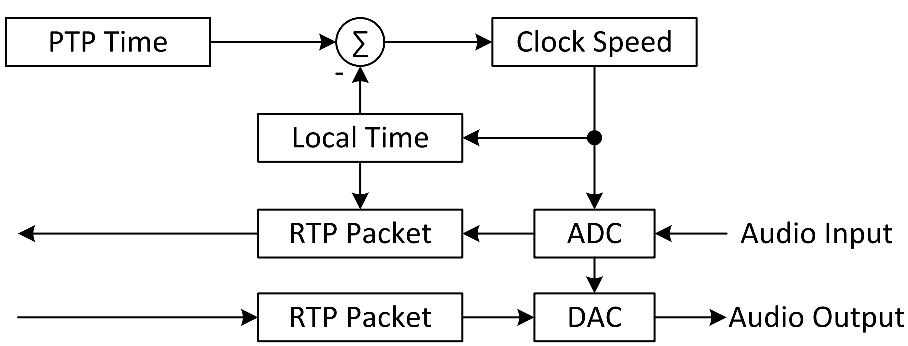
The Audio System shall output a KEY discrete in the same CDN message with the mic audio stream for each mic audio destination to route the mic audio to the appropriate destination.

3.8.7. Synchronization Considerations

Synchronization between real time publishers and subscribers is required to minimize audio noise due to dropped or added samples and to meet the RTCA DO-214A differential delay requirement.

The Audio System shall use PTP per IEEE 1588-2008.

Real Time publishers and subscribers should synchronize their system and sample clocks to within 100µs of the master system time. This provides compliance with the RTCA DO-214A differential delay requirement, and minimizes noise due to dropped/added samples which would otherwise occur due to buffer over/underruns in real time.



This document recommends that the Flight Deck user positions provide the master PTP time service to the network, and be able to negotiate which Flight Deck user position is the single master. This provides redundancy in the case of the loss of an position. An independent PTP network may be used, however redundancy is still recommended.

The Audio System shall use a method to compensate for clock differences between the publishers and subscribers of audio streams when the clock sync between audio equipment is lost.

The Audio System shall provide a means to determine when clock sync between audio equipment is lost for the publishing audio equipment.

ED-112A places additional requirements on channel synchronization: “The recordings for separate channels shall be made such that, when replayed, the relative time between channels can be deduced to better than 4 milliseconds irrespective of recording delay.” However, meeting differential delay requirement is much more restrictive (in order to meet the system must be synchronized to within 114 us.) Once synchronized, a system meeting the differential delay requirement will also meet the ED-112A requirement.

3.8.8. Sidetone Considerations

Systems using local sidetone shall ensure that the HF tune tone is still heard in the local sidetone audio when present.