Audio Engineering Society
Convention Paper
Presented at the 114th Convention
2003 March 22–25 Amsterdam, The Netherlands



# Latency in Audio Ethernet Networks

Nuno Fonseca[1, 2] and Edmundo Monteiro[2]
nfonseca@estg.iplei.pt edmundo@dei.uc.pt

[1] Polytechnic Institute of Leiria, ESTG, Alto do Lena, 2400 Leiria, Portugal

[2] University of Coimbra, Pólo II, DEI, 3030-290 Coimbra, Portugal

## ABSTRACT

In a time when several audio Ethernet networking solutions are being studied and developed, the analysis of the latency introduced by theses networks is fundamental. This analysis is the subject of the present paper, and is necessary not only to enable the identification of the factors that can be optimised, but also to support the decision about the possibility or not of the inclusion of in-band synchronism signalling.

## 1. INTRODUTION

The need for an audio networking solution increases every day, and Ethernet can be one possible solution [1]. This paper will analyse the latency in this type of networks from the audio point of view.

In this study we make the following assumptions. First of all we assume a star topology (or hybrid star topology [5]), meaning that every device is connected to a central point. Although a daisy-chain could offer lower latency, that topology brings several disadvantages regarding connecting, managing, troubleshooting, and bandwidth issues when compared with star.

Secondly, we assume that the central point(s) is an Ethernet switch working on full duplex mode. Nowadays, the price of this kind of equipment is almost similar to Hubs, with much more performance.

Thirdly, we assume that network devices are non-blocking, or that the total amount of traffic is below the blocking state, i.e., the results don't take into account the latency produced by blocking situations [6].

## 2. LATENCY

Let's consider the latency of one Ethernet frame from the moment the first bit is sent, to the moment the last bit is received. There are three types of delays involved in this situation: the propagation delay (the delay of the propagation of one bit between two points of the network); the delay introduced by the network devices (switches); and the frame transmission delay (the delay between the transmission of the first bit and the transmission of the last bit).

### 2.1. Propagation Delay

The propagation delay is due to the propagation of the signal (electrical or optical) between the sender and the receiver. The propagation speed depends on the cable properties, and is measured by comparison to light speed, we can consider, as shown bellow:

$$t = \frac{d}{(p * c)}$$

where:

*d* – cable distance between endpoints (m)

*c* – light speed (m/s)

*p* – cable propagation factor ([0, 1]of light speed)

The graphic of Figure 1 plots propagation delays in a cable with a propagation factor of 0.6.

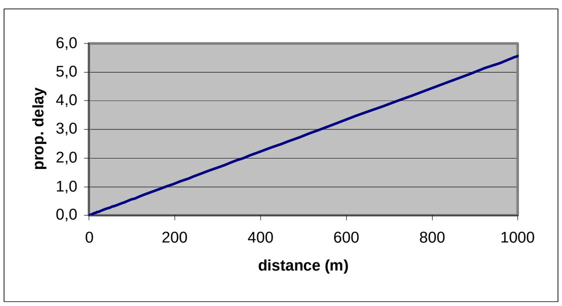


**Figure 1 –** Propagation delay (µs) in a 0.6c cable

### 2.2. Frame Delay

Other type of delay that is added to the total latency is the frame transmission delay [4], i.e., the latency between the transmission of the first bit of the frame and the transmission of the last bit of the frame, which depends on the size of the frame and the bandwidth:

$$t = \frac{\#bits}{bandwidth} = \frac{\#bytes * 8}{bandwidth}$$

Figure 2 shows plots transmission delays against frame sizes in Gigabit Ethernet link and in a Fast Ethernet link.

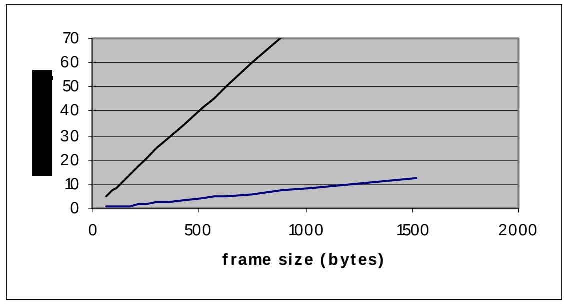


**Figure 2 –** Frame transmission delay (µs) versus frame size in a Gigabit (lower line) and Fast Ethernet link (upper line).

### 2.3. Switch Latency

Every switch adds two types of latency, the buffering latency and the forwarding latency.

The buffering latency is the time the switch takes to receive and store part or the total of the frame, before starting transmitting it again. Some switches use a pure store-and-forwarding architecture, which means that the switch has to receive and store the entire frame before sending it out. Other switches use a fragment free method (a kind of intelligent cut-through), where the switch buffers only the initial 64 bytes, before starting the transmission.

The forwarding delay is a switch specification parameter that represents the processing latency of the switch.

The sum of these two latencies gives the total switch latency between the reception of the first bit, and the transmitting of the first bit:

$$t = \frac{buf}{band} + t_{fw}$$

where:

*buf* – initial buffering (bits)

= frame size, in store-and-forward switches

= 64*8, in fragment free switches

*band* – bandwidth (bits/s)

$t_{fw}$ – forwarding latency (s)

### 2.4. Total Network Latency

So putting all together, we know the latency from the moment the first bit of the frame is transmitted by one endpoint, to the moment the last bit is received by the other endpoint, we will have:

$$t = \frac{d}{p*c} + \sum_{\#Sw}\left[\frac{buf}{band} + t_{fw}\right] + \frac{F_s}{band}$$

where:
*t* – total latency (s)
*d* – distance between endpoints (m)
*p* – cable propagation factor ([0, 1]of light speed)
c – light speed (m/s)
*#Sw* – number of switches
*Fs* – frame size (bits)
*buf* – initial buffering (bits)
= *Fs*, in store-and-forward switches
= 512 (64*8), in fragment free switches
*band* – bandwidth (bits/s)
$t_{fw}$ – forwarding latency of the switch (s)

The expression above represents the total latency introduced by the network but, the total latency of audio delivery, depends also of two other factors. The first is the Audio Buffering delay – due to the fact that the transmitter may not generate an Ethernet packet for each audio sample. The second factor includes the delays that are introduced by the transmission, transport and reception of other frames in the network.

### 2.5. Audio Buffering Delay

If the audio device sends a Ethernet frame for every audio sample, this means that a 64 bytes frame (frame minimum size [2]) is needed to send only two or three bytes of audio (considering that each sample have 16 or 24 bit resolution), which decrease the bandwidth for data, decreasing the number of available audio channels.
One way to solve this problem is to do some buffering at the transmitter and send one Ethernet frame only when more than one sample is available. This will increase the latency by the factor:

$$t_{buf} = \frac{buffersize - 1}{F}$$

where:
$t_{buf}$ – buffering latency
*buffersize* – number of samples in one frame.
*F* – sampling rate (Hz)

Figure 3 shows the delay due to audio buffering versus de number of samples per frames, at a 96 KHz sampling rate.

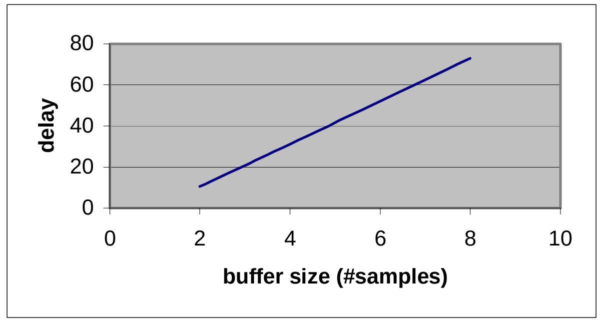


Figure 3 **– Delay (μs) in audio buffering at 96 Khz.**

### 2.6. Delay due to Other Frames

The traffic due to other frames in the network will also add latency to the system.
If the audio device needs to send packets for more than one destination, the transmission of one frame will be delayed when the audio device is busy transmitting frames to other destinations over the same Ethernet link. This means that the transmission of some frame could have an extra delay of:

$$t = \frac{(n-1)*Fs}{band}$$

where:
*n* – number of destinations
*Fs* – frame size (bits)
*band* – bandwidth (bit/s)

Figure 4 plots the delay due to other frames in the network against the number of different destinations on the network.

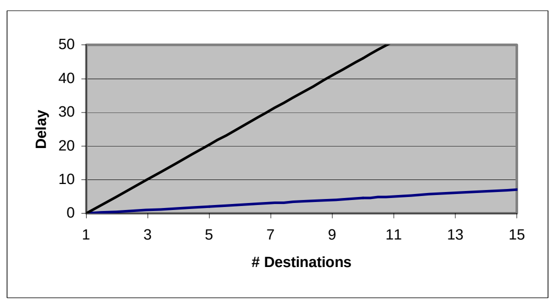


**Figure 4** – Delay (μs) transmitting 64 byte frames on Gigabit link (lower line) and on a Fast Ethernet Link

The traffic due to other frames in the network will also affect the reception of frames. Some audio device could be receiving at the same time, frames from several devices, on the same link, which means that an additional delay will be add:

$$t = \frac{(n-1)*Fs}{band}$$

where:
*n* – number of sources
*Fs* – Frame size (bits)
*band* – bandwidth (bit/s)

This kind of delay will also occur in the communication between switches, because the switch may need to use the same port to send several frames (even frames that don't have the same source and destination).

## 3. ANALYSIS

To evaluate the overall effect of the latency factors discussed in the previous section, let us consider a Fast Ethernet network over copper cable (with a propagation speed of 0.6c), on a network with a maximal distance between endpoints of 50m, with the communication passing on 2 switches with a forwarding latency of 7 us, using a 512 byte frame.

In the scenario described, and not considering the audio buffering latency and the latency added by other frames, we reach a value of T = 137 µs for the total latency introduced. If we were using an audio stream of 44.1 KHz, that latency would represent 6 samples of difference. If we wanted to have a sample sinchronisation, this value should be lower than 0,25 sample [3].

Considering the problem in another way, let us find what network we should have to support a sync signal.

For a 192 KHz audio stream (to take into account future needs in audio systems), a sample 25% is equivalent to 1.3 µs delay.
In Fast Ethernet the total frame latency for a 64 byte frame (which represents the minimal frame size) is 5 µs.
In Gigabit Ethernet, with a 50 m link and using only 2 switches, the total frame latency for a 64 bytes frame will be T = 1.8 µs (even considering forwarding delays of 0).

So, we can conclude that an in-band sync signal is no possible in Ethernet networks. Does this mean that we should discard Ethernet as an Audio Networking Solution? Of course not!

## 4. CONCLUSIONS

This paper presents theoretical analysis about audio latency in today Ethernet networks. The analysis shows that Ethernet audio networks presents good latencies values for audio transmission (even with large networks we can get latencies lower than 1 ms), but not for in-band sync signals (maybe like any other audio networking solution).

Regarding the transportation of sync signals, the solution may pass, not by using sync packets, but by the creation of other ways to inform the remote devices of the sync of the signals.

This kind of problems are, of course, much easier to solve on "point-to-point" solutions, or even, on "daisy-chain" solutions, but to achieve a true networking solution (even for audio transportation in real-time) the solution must be asynchronous (from the transmission point of view) and switched, to allow full flexibility.

Ethernet continues to be a very good solution for audio networking, but there is a need for a networking sync solution (physically integrated with Ethernet or not).

## 5. ACKNOWLEDGMENT

We would like to acknowledge our institutions and the Foundation for Science and Technology from the Portuguese Ministry of Science and High Level Education.

## 6. REFERENCES

[1] Nuno Fonseca, "Foundations for Sound over Ethernet," presented at the AES112th convention, Munich, Germany, 2002 May.

[2] IEEE, "IEEE 802.3 Standard 2002", *IEEE*, 2nd Edition 2002.

[3] Michael Talbot-Smith, "Audio Engineer's Reference Book", *Focal Press*, 1999, pp 6.57-7.1

[4] John L Hennessy, David A Patterson, "Computer Architecture, a Quantitative Approach", *Morgan Kaufmann*, 2nd Edition 1996, pp. 562-632.

[5] James Chellis, Charles Perkins, Matthew Strebe, "MCSE: Networking essentials", *Sybex*, 2nd Edition 1998.

[6] Karen Webb, "Building Cisco Multilayer Switched Networks", *Cisco Press*, 2001, pp 405 – 419