Proceedings of the Institute of Acoustics

Wha wha what wa wa was tha tha that
- or -
Latency in Multichannel Digital Audio Transmission Systems and Its Practical Effects

Sam Wise Arup Acoustics, Winchester, Hampshire, UK

1 INTRODUCTION

According to The Concise Oxford Dictionary, latency is derived from the word latent. Latent is an adjective that means concealed or dormant. Alternatively it is used to describe something that exists but is not yet developed or manifest. Latency is the noun form of the word latent.

Latency (signal delays hidden in the process) in multichannel digital audio transmission systems is a latent (concealed) problem that can produce unexpected effects in the acoustic domain. This paper is a simple tutorial to help prevent users of these new systems from falling into unexpected traps. It is helpful to understand how latency arises, the symptoms that it can produce and ways to plan for it.

1.1 Latency in Acoustic Systems

A good place to start a description of latency is by the use of examples from acoustics. As we know, sound takes a finite time to travel from one place to another. This travel time is not fixed, but varies with distance and with temperature and humidity. Under many circumstances this is of little concern to the acoustician, but in the context of room acoustics or sound system engineering, it becomes very important. Reflected sound arriving from a source within a given time interval can produce an apparent increase in loudness compared to direct sound only. Extending the arrival time (or latency) of this reflected sound a bit will cause it to start to interfere with the direct sound, reducing clarity or intelligibility. A further delay in arrival of the reflection will separate the sounds completely, producing an echo. Thus acoustic latency could be defined as the delay between the emission of the initial sound and the arrival of the reflected sound. A lot of care is taken of this effect in the design of concert halls and similar spaces.

At a fixed distance from a source, the arrival time only varies with atmospheric conditions. In most rooms, this change has only minor significance. However, in very large spaces, such as arenas and stadia, variations of arrival time can become significant enough to require adjustments to prevent echoes, or to improve the imaging to the desired location. Even very small changes in this arrival time can introduce changes to system performance when arrays are built from multiple loudspeakers.

One could consider these signal arrival delays as "latency" caused by acoustic effects. Under some conditions, we don't care. Under other conditions they are of major significance. Competent practitioners will take notice of these delays and devise methods for accounting for their effects.

1.2 Latency in Electronic Systems

Until about 20 years ago, latency in sound system design was mostly acoustic. We have had to keep track of the delay of sound travelling though air in many applications in order to produce a good design. There has always been delay in the electronics as well, but in the analogue domain one had to try very hard to introduce any substantial delay or latency compared to the acoustic delay. In the interim, there were mechanical means for producing substantial delays and for the production of multiple reflections such as spring and plate reverb devices. Analogue bucket brigade type delays appeared in the 70's as a first step toward sampled system delays. With the introduction of high quality electronic reverb, chorus and flanging effects, variable electronic delay or "latency" began to be used to simulate the acoustic delay effects found in real rooms. This then was further extended to produce sounds never heard naturally. But, in most cases, this was done on purpose, *for effect,* and therefore did not often give rise to the unexpected.

My first serious encounter with ill effects caused by unexpected signal processing latency came prior to a concert at which I was bass player, but also responsible for the sound system, although I did not perform the mix. The first sign of a problem followed from plugging up the keyboard. At the mix position this sounded fine, but on stage the piano sounded as though it was down a cave. Ten minutes into rehearsal I could stand it no more. A few head scratches, knob twists and plug patches later, it emerged that left and right outputs from the electric piano were patched to left and right on the house sound, producing a very spacious keyboard sound. However, left and right were both mixed together onto the stage monitor facing me. The resulting unacceptable sound was caused by a delay built-in to the keyboard, intended to produce the spacious stereo effect for the audience. A latent side effect arose when these two outputs were mixed back together and sent to a single speaker, producing a horrible hollow sound. Given the right (or wrong) circumstances it is possible to produce a similar effect from a single loudspeaker cluster using a DSP based signal processor and/or digital signal transmission system. Other effects can arise from the interaction between a stage monitor and a house loudspeaker when they are located near each other.

2 LATENCY – WHAT CAUSES IT

2.1 Latency in Analogue Systems

All physical systems that affect energy will introduce some delay or latency in the energy flow. In analogue audio electronics we usually express this as a phase response variation with frequency. Ie in most analogue systems the latency is somewhat variable with frequency, but remains shorter in time than a single cycle at the highest frequency of interest. If we are looking for a 20 kHz bandwidth, this means that any delays are likely to be substantially less than 50 µs. When trying to achieve matching between two loudspeakers operating together at high frequencies, even this is too much. This is why acceptable phase shift in professional equipment is ordinarily limited to a fraction of a wave cycle, say 45°, equivalent to a signal latency of about 6 µs at 20 kHz.

2.2 Latency in the Digital Conversion Process

When we move to digital audio systems, immediately consideration of the conversion process is required, even if there is no other internal signal processing. Where systems sample at 48 kHz, a delay arises equal to a minimum of one sample between the input and the internal digital representation of the signal. This delay is approximately 20 µs. The delay is not as precise as it sounds, but also has a variation in its timing, which is conventionally referred to as jitter. Jitter is in fact a variation in latency. Others have documented the ill effects that can arise in high quality audio due to jitter.

Most audio signal delays used in professional applications use a 48 kHz sampling rate and are therefore also limited to adjustment steps of 20 µs between any two signals flowing through or derived from that system. Increasing the sampling rate is one way to reduce both the latency and the step size between delayed outputs. Another way is to internally double clock the signals to generate shorter sampling intervals for the start of output conversion. One manufacturer produces a product having step sizes of 10 µs, which is helpful when trying to control signal interaction effects at the highest audio frequencies.

2.3 Latency in Digital Signal Processing

If we now introduce some digital signal processing (DSP) into the system, we immediately also introduce additional delay or latency. This is because DSP algorithms (or programs) are formed by delaying and adding signals together, usually by adding a signal to other, delayed versions of itself. The more complex the process, or the more processes that are set up in series, the longer that delay will be. A typical multi-purpose digital signal processor may introduce total processing delays of 5 ms to 15 ms for moderately complex systems. *This is in addition to any delay purposefully introduced!* As the latency is dependent upon the signal processing used, there may be different delays in different signal processing pathways. Some systems notify the user when this arises and introduce automatic correction if requested. Other systems introduce delay compensation automatically within each processing box, which the user can override. This helps to ensure that all signals departing from one processing unit remain synchronised, but does not remove all risk of latency problems.

2.4 Latency in Digital Signal Transmission

If we now add the further complication of digital signal transmission, we may introduce yet more latency into the system. While the first two causes may be somewhat obvious, this latter issue is not so obvious and may vary according to the way a complex system is patched externally. This means that while a system might be designed to inform the user of the delay arising for internally patched and processed signals, it may not be able to do so when it performs a hidden function such as signal transmission. Current Cobranet based products have a total transmission latency of about 5.3 mS from analogue input *to any output*. This sounds acceptable, but as we will see, an output arises when the signal is transferred directly (hidden) into a compatible manufacturer's DSP system. Another manufacturer's non-Cobranet based product adds minimal delay for transmission, but requires more cabling for medium and large scale systems.

As transmission systems increase in the number of channels or in distance, various hubs, media converters and other items will be required in the system. The signal also passes through these and they too can extend the latency of the signal. The important thing to remember is that transmission latency *adds* to the total signal latency overhead, whether or not signal processing is included and system selection needs to consider potential latency issues as well as other design factors.

2.5 Variations in Latency

Finally, one must consider not only the fixed latency in a system, but also any variability in latency. Leaving aside jitter, which is usually outside system designer control, how can this arise?

All digital systems that connect more than just a pair of units on the ends of one cable must use some method to allow those units to arbitrate between themselves. Which unit has control and is allowed to talk (send its signal)? In administrative type activity, some delay in response is acceptable, as most computer data transfer does not require precise timing. More important here is that the data received is an exact copy of the data sent. Corrupted data may be more harmful than just wrong information, it may cause the whole system to fail. Data integrity is king, along with the ability to send the data just about anywhere.

When transmitting audio or video, continuity and timing becomes king, for the most serious errors arise where the stream is interrupted, introducing dropouts and glitches. This is the worst kind of latency variation. Many of us are observing the development of systems for the transmission of audio and video over the Internet. We experience these defects ourselves. Elimination of this problem can be arranged by buffering the whole item prior to allowing playback. Thus we can exchange terrible signal failure due to short-term variations in latency, for almost no signal failure and very high latency.

In professional audio, neither of these is acceptable, nor is the quality reduction that arises from most bitrate reduction processes that are used to limit the total time we have to wait. Thus manufacturers have had to devise methods for signal transmission that give good promises on the variation in latency (ie none, except jitter), while keeping the fixed latency within acceptable limits.

Another source of latency variation that should be considered in audio systems can arise when DSP processors allow different "programs" to be loaded. One manufacturer suggests that if DSP power runs short in his system, a means of overcoming this would be to create multiple setups, each of which used less total DSP resources, and to load these as required for the particular use of the event. Designers should take care in doing this to ensure that the total processing delay through the DSP engine does not change, introducing signal alignment variations that were not intended.

2.6 Products for Multichannel Digital Audio Transmission

CobraNet (licensed by several manufacturers) is based on the standard Ethernet and TCP/IP protocols that are used within our office computer network systems. Imposed on this is a proprietary method that ensures orderly action by the various CobraNet compatible devices on the network so that transmission delays are bounded and failure does not arise. CobraNet is compatible not only with standard CAT series and fiber-optic cabling, but also with many computer industry standard devices. Up to 64 channels can be transmitted on one cable. The cost of this compatibility and expandability in today's CobraNet products is a fixed latency for transmission of 5.3 ms. It is understood that work is underway to reduce this. It is recommended that standard computer "terminal equipment" (computers, servers or printers) using a standard Ethernet interface does not share the network with CobraNet compatible products.

BSS Audio's Soundweb product range uses a proprietary method of transmission of eight channel blocks between their units. Point-to-point interconnection between units is used. The hub used to interconnect and construct larger systems also contains audio DSP capability. Transmission arbitration therefore is not handled at the network cable interface, but is dealt with inside the DSP portion of each individual processing unit. Soundweb is compatible with standard CAT and fiber-optic cabling, but compatibility ends there and proprietary devices must be used throughout. One benefit is that transmission latency is low.

IEEE-1394 (Firewire, iLink) based AV networks are in development, but will initially be confined to the smaller domestic or studio contexts. Other systems are available, but these are generally limited to either single unit-to-unit transmission, or include latencies that are not acceptable for real-time use.

3 LATENCY EFFECTS IN REAL SYSTEMS

3.1 DSP Based Loudspeaker Controllers

Most (possibly all) audio DSP processors provide a means to ensure that the total delay through the system remains constant for all signal paths. This ensures that relative alignment between outputs does not change, even though the total latency of the system may. Providing this alignment has a cost in terms of DSP power used that can be quite significant. On many system designs, activating this signal alignment adds cost by demanding the addition of further DSP processing devices, no matter which manufacturer's products are used.

3.2 Array Alignment

When DSP or digital audio transmission systems are used to process and/or send signals to an array, the precision of arrival time can be a very significant factor in the quality of sound received across the audience. Mis-alignment produces well known effects, such as comb filtering and the generation of side lobes that take energy into unexpected parts of the room. This is the main instance in sound reinforcement where a detailed look at the signal alignment arising between outputs in the system is essential.

3.3 Digital Signal Transmission from Mixer to Loudspeaker Controller

If DSP processing is used, it is often in the context of adjusting the relationships between analogue inputs and analogue outputs. However, adding digital transmission from a remote location (the mix position) to the loudspeakers can introduce additional fixed delays into the system, ranging up to 5.3 ms. When added to typical DSP processing delays, the total can approach 20 ms.

If such a system is driving a loudspeaker array, then the whole array will be delayed by the total latency, causing the array sound to arrive after the live sound coming from stage. With loudspeakers on stage, this may not be a problem, since this delay would possibly only improve the imaging of the sound toward the performers themselves. However, if the array is flown above the stage, then the digital processing latency must be added to the considerable sound transit time delay.

As an example, consider an array at a typical height of 10 m above the stage. The acoustic transit time from the array to the front seating area would be about 35 ms – or just acceptable to prevent confusion in the sound at those seats. Adding in a total digital processing latency of 20 ms would increase this to 55 ms, which is definitely in the danger zone and may be considered unworkable.

3.4 Including the Microphone Distribution in the System

Now consider the goal that might be achieved today. There is a stage box on the stage into which microphones are connected. This contains microphone pre-amplifiers that can be remote controlled from the mix position. The microphone signals are transmitted in digital form over a single link to the mix position. As the mixer is likely to remain analogue at present, the outputs are taken from the distribution system in analogue form. A transmission delay of up to 5.3 ms may have now been added to the total system latency. In the system above, the flown array might now be 60 ms behind the direct sound in the front seating area.

The advantage to this system is that the microphone signals are virtually totally impervious to noise once in digital form. Stage Monitor, broadcast and recording feeds can be split off in the digital domain, removing the need for analogue microphone splitters and any signal degradation these introduce. The disadvantage is also clear.

3.5 Interference Effects Between Closely Located Loudspeakers

Within an array, it is relatively easy to consider and deal with the variations that might arise in the DSP processing channel. This is especially true as the DSP loudspeaker processor is most commonly located at the amplifier rack, ensuring that all outputs are derived from one system processor.

But consider what might happen if a house sound system uses some loudspeakers stacked on stage, and driven from a similar processor connected to the digital audio transmission system. Let's assume that the equipment is of good quality, so the touring show uses the house sound system. However, there is a strong preference by the monitor mix engineer for the use of the band's own stage monitor system, including analogue compressor/limiters and graphic equalisers. Helpfully, we give the monitor engineer one of our digital signal transmission output boxes from which he takes his microphone feeds. This is nice and tidy, requiring only a single CAT cable to his input system. The delay on this feed direct from the stage is potentially only 5.3 ms, or even less with some systems, not enough to cause significant side effects.

Then, assume a stage monitor loudspeaker or two are placed toward the front of the stage as side fills, directly adjacent to the house loudspeaker onstage stack. As no DSP processing is used on the monitor system, the time difference between sounds emitted by the monitor and house systems might be between 5 ms and 20 ms. In acoustic terms this is equivalent to half a wavelength at 25 to 100 Hz. Below about 250 Hz, most loudspeakers are nearly omni-directional. The result of this interaction is that the outputs of the house and monitor loudspeakers might differ in phase by 180° at 25 Hz to 100 Hz, producing a cancellation at that frequency. Nulls would be repeated at a frequency intervals of double the first cancellation frequency - the well-known comb filter effect.

As the effects are wavelength dependent, differences in the latency (or unplanned delay) between any two closely located loudspeakers can result in unexpected interference at unexpected frequencies. In the range from 1 ms to 10 ms, this difference can have a profound effect on expected system response, even if the loudspeakers are not pointed in the same direction. This *really is* what was called a "latent" problem in the introduction to this paper. The cause is hidden and unexpected. Troubleshooting methods are unclear. In fact it might take some time to devise a solution and in many instances the cause would never be appreciated.

Note that this is a result of DSP processing time, rather than anything to do with digital transmission. It is probably already a routine problem on larger systems, but largely unrecognised. If analogue and digital processing is somehow mixed in a system, or different digital processing is used, then this could occur. No manufacturer's system is immune unless DSP latency is reduced to the µs range, or house technical staff become aware of the problem so that they can recognise and deal with it.

3.6 Large Distributed Systems

When systems get large, digital signal transmission really comes into its own. Cable costs are so substantially reduced and flexibility improved, that the systems themselves can be considered almost free of cost. .

Examples of such uses would be stadia, theme parks, major broadcast and recording complexes, conference and exhibition centres, cultural complexes and even the education industry within audio/visual signal distribution systems.

Here such an integrated system might effectively become the invisible helper, transporting signals between what look like conventional patch facilities. Even where the full flexibility of the system is used to control signal routing via a computer screen, the underlying structure of the hardware would normally be made invisible.

In these circumstances, even where most recipients of distributed signals are not in hearing range of the source, great care may be needed to ensure that routes cannot be accidentally enabled that produce effects like those outlined above.

3.7 Connecting to an Outside Broadcast Vehicle

Feeds to an outside broadcast vehicle could easily comprise:

- a direct digital distribution of the stage microphones
- pre-mixed feeds from the house sound system
- commentators feeds
- a few broadcaster-owned slung microphones wired directly via analogue multicore
- other direct sound pickup from intercom systems or camera-mounted microphones.

It is easy to see that care in planning will be essential to obtain the desired result.

4 SUMMARY

Digital audio processing and transmission systems are extending themselves ever deeper into the sound production process. Recording systems are now largely completely in the digital domain, with satisfactory results. But recording is not real-time and therefore is not so sensitive to delays in the process (except for foldback and monitoring issues).

As digital audio becomes more prominent in the live environment, designers and users need to be aware of the potential effects of delays or latency in the signal and take these into careful consideration when designing and operating the systems.

5 REFERENCES

Davis and Davis, *Sound System Engineering,* Butterworth & Heinemann, 1986

Wise, S., Digital Audio Tranmission Systems, *Installation Europe*, 1999

Murphy, D., Audio quality and capacity issues in network design, *Moving Audio,* Audio Engineering Society, 2000

Moses, B., Audio applications of the IEEE 1394 high performance serial bus, *Moving Audio,* Audio Engineering Society, 2000

Coulon, J., Audio over USB, *Moving Audio,* Audio Engineering Society, 2000

Dunn, J., IEEE 1394 and Sampling Jitter, *Moving Audio,* Audio Engineering Society, 2000

Rave Application Guide, QSC Audio Products, Inc. 1998

Zwieble, R. and Gross, K., *CobraNet Technology Papers*, Peak Audio Website, 2000

Harshbarger, B., *A Networking for Audio Primer, Parts I to IV*, Sound and Communications Magazine, 1999