Clarifying Processing Delays Specific to a Riedel MediorNet IP System

by Renaud Lavoie, Francois Gagnon, Daniel Tremblay

Introduction

The goal of this white paper is to provide some insight into the delays[1] occurring inside the basic blocks of an ST 2110 system. While this document mainly focuses on Riedel products (because not all gateways and processing devices act the same in ST 2110 systems), it also provides a general rule of thumb for delays in switches and in traditional SDI systems.

In this paper, we address one of the most frequent questions from customers who are still at the beginning of their transition from SDI to IP: What is the latency in a MediorNet IP ST 2110 system and how does it differ from an SDI system?

To this end, we divide the system into its basic blocks: the encapsulators and de-encapsulators, the switches, the processing devices, and the encoders and decoders.

1. Basic 2110 Encapsulator Delays

The encapsulator is also referred to as the sender gateway. Encap in NMOS and control systems is composed of the SDI / HDMI input and the output to the network. Riedel's MediorNet IP product line is equipped with frame synchronizers and audio sample rate converters (except the J2K encoders, which do not need such functionality as they do not need PTP reference) to synchronize free run signals to the PTP flow in ST 2110. Of course, the frame synchronizer will add delay for the signal not in sync. For synchronous signals (PTP locked), this function can be disabled to minimize the encapsulator delay (eDelay). The following figure shows the encap blocks:
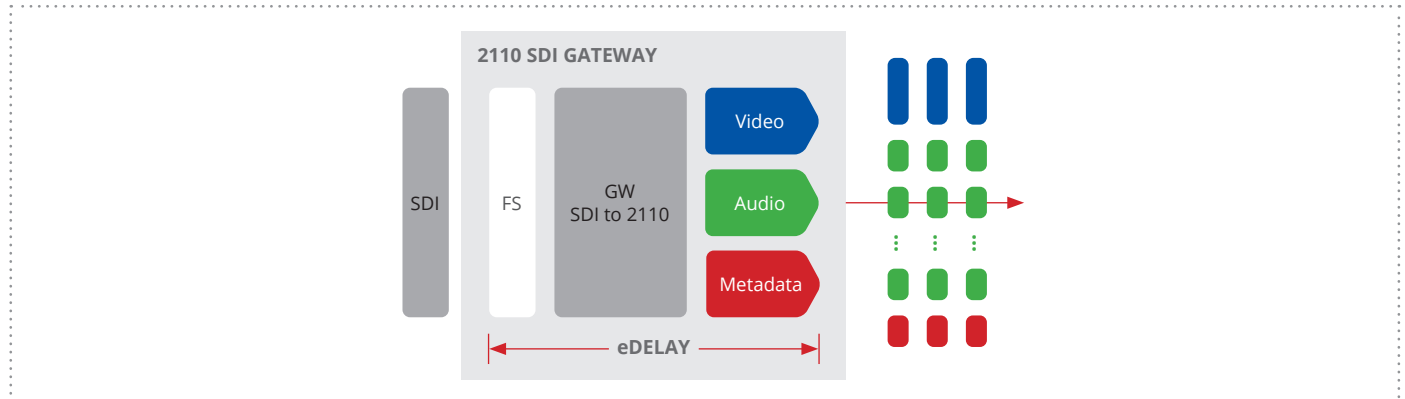

Figure 1. ST 2110 SDI encapsulator delays (eDelay) with/without ST 2022-7[2]

The eDelay depends on the frame synchronizer (FS). Accordingly, the following table is divided into eDelay with and without FS:

Essence	Async eDelay with FS	Sync eDelay without FS
Video	Up to 1 frame	½ lines
Audio	Up to 1 frame	1 packet: Dependent of packet time (pTime; max 1msec)
ANC	Up to 1 frame	Mode 1 (M-bit): 1 frame (1 packet per frame) Mode 2 (MTU mode): Fill MTU (1500 bytes of data) Mode 3 (Quick): Send packet per ANC (packet time at 10G or 25G)

Table 1. ST 2110 SDI encapsulator delays (eDelay)

---

[1] *Delay* and *latency* are often used interchangeably, but they are not always the same. While there is no universally valid distinction between the two, *latency* is often associated with the round-trip time (the time it takes for a packet to be sent plus the time it takes for it to return back). This paper, howver, investigates the different parts that make up the "overall latency" in a system, like the *delays* inside the senders or receivers.
[2] Only de-encapsulators are affected by 2022-7.

Note on Audio: *The audio delay depends on the packet time (pTime). For a 1ms pTime, the delay will match this 1 ms delay + transmission time, which is neglectable.*

Note on ANC: *There are different ANC modes in ST 2110 encapsulators.*
In **Mode 1***, also referred to as the M-bit mode, the ANC packets are kept in the encapsulator until the M-bit arrives.*
Mode 2 *is MTU mode, where packets are filled with ANC before sending. MTU is 1500 bytes per payload. The total MTU is 1526 bytes with the Ethernet header. Jumbo Frame is not supported in ST2110.*
Mode 3 *is the quickest mode: each time an ANC in blanking is completed, the packet is sent over the network.*

The following image summarizes the 3 modes (the number of ANC packets is arbitrary and does not reflect reality)
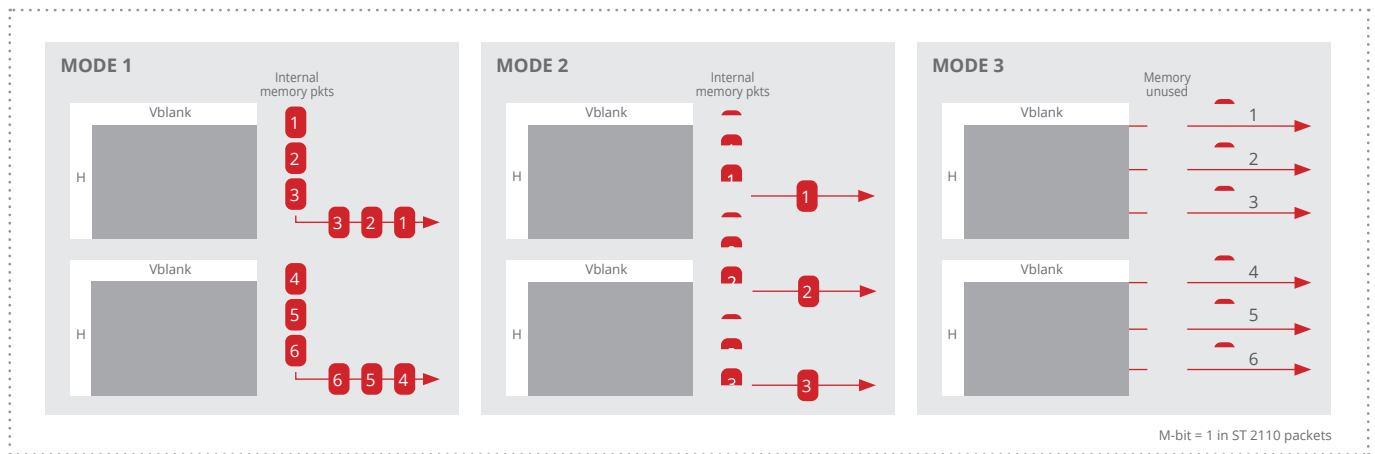

Figure 2. ANC modes in ST 2110 encapsulators

2. Basic 2110 De-encapsulator Delays

The de-encapsulator is also referred to as the receiver gateway. The De-encapsulator receives an ST 2110 set of flows and provides an SDI / HDMI output. Riedel MediorNet IP de-encapsulators are equipped with ST 2022-7 buffers for network redundancy (please note that encapsulators provide ST 2022-7 packet duplication without requiring additional buffer.
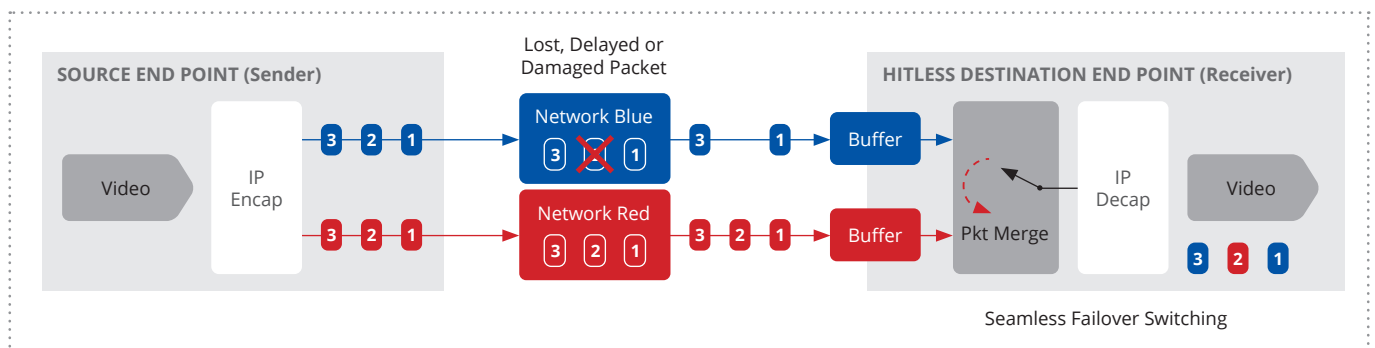

Figure 3. ST 2022-7 explained

The buffer produces the biggest part of the delay inside the receiver, but fortunately, it is configurable to be adapted to your needs. Different types of ST 2022-7 modes are specified as follows:

Class	SBR	HDBR
A	10 ms	10 ms
B	50 ms	50 ms
C	450 ms	150 ms
D[3]	150 µs	150 µs

Table 2. ST 2022-7 delay table

The high bit rate (HBR) is defined as a data rate from 270Mbps, HD, 3G up to UHD and more. The SBR is a lower bit rate defined as lower than 270Mbps.

The following picture represents a de-encapsulator. The table below shows the delay the device will exhibit with and without ST 2022-7.
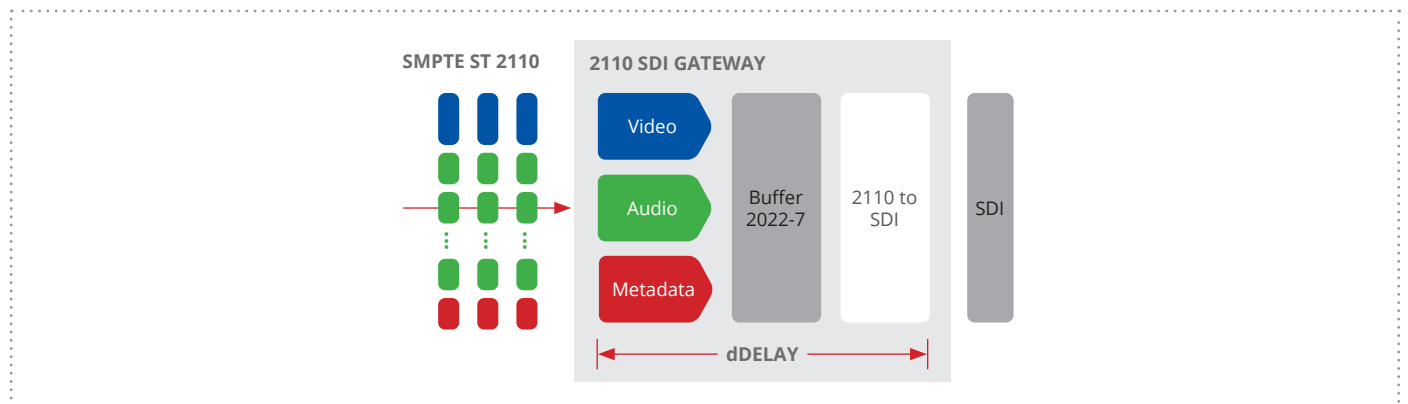

Figure 4. ST 2110 SDI de-encapsulator delays (dDelay)

As you can see from the following table, in order to support all ST 2110 in the market, the receiver will add 1 frame of delay on the ANC in M-bit mode.

Essence	Sync (dDelay) with 2022-7	Sync (dDelay) without 2022-7
Video	2022-7 D: 150 µs	7 lines HD 14 lines in 3G 56 lines in UHD mode
Audio	Audio is realigned with video 2022-7 D: 150 µs	Audio is realigned with video. 7 lines HD 14 lines in 3G 56 lines in UHD mode
ANC	Always 1 frame (mode M-bit to support all 2110 senders)	Always 1 frame (mode M-bit to support all 2110 senders)

Table 3. ST 2110 SDI de-encapsulator delays (dDelay)

---

[3] ST 2022-7 Class D with short delays was added for live production, room to room or studio applications.

3. A Rule of Thumb for Network Switch Delays

Today's network switches are more predictable than old ones. The following rule of thumb assumes that the switch is well-configured (IGMP enabled, ports are not oversubscribed, QoS is present or not, no deep buffers active, packet size MTU 1500, cut through config or store and forward mode, etc.)

With these assumptions, it is fair and realistic to say that the delay per switch is 5 µs. The following image shows a spine-leaf configuration, but we can assume that any configuration (spine-leaf, monolith core), each switch will have the same 5 µs delay.

The total delay on an encapsulator to an de-encapsulator path is 15 µs: 5µs for the TOR switch connected with the encapsulator + 5 µs for the core switch + 5 µs for the TOR switch connected with the de-encapsulator.
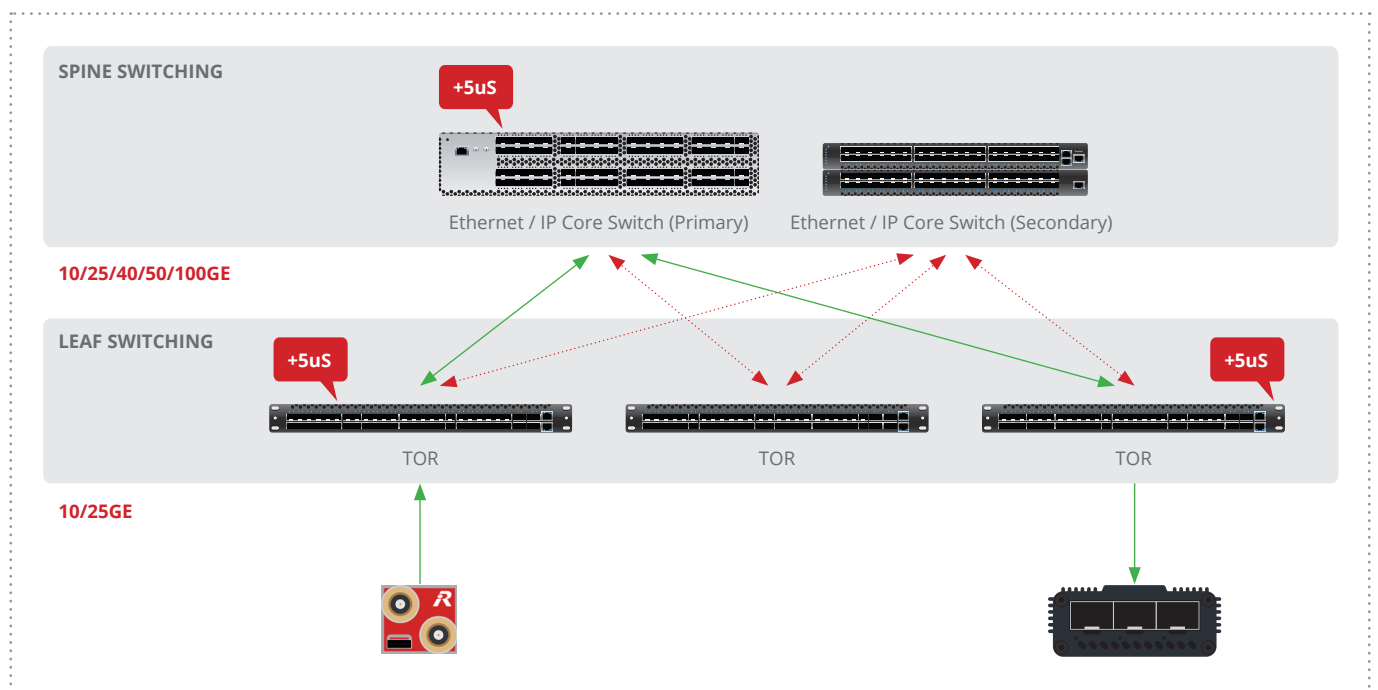

Figure 5. Switch delays in a spine/leaf architecture

4. JPEG2000 (J2K) Delays

The J2K encoders and decoders described here are the ones used for SDI to J2K compression and J2K to SDI decompression. To calculate the delay for this type of compression, all you have to do is combine the J2K delay with the ST 2110 delay. Just be sure not to add the ST 2022-7 block twice, as there's only a single ST 2022-7 delay period present in this path.

The following images show the J2K encoder and J2K decoder without any blocks of automatic repeat request (ARQ). Attached with the image you can see the delay of the J2K encoders and decoders. In addition, ultra-low latency (ULL) is not enabled in the following delay.
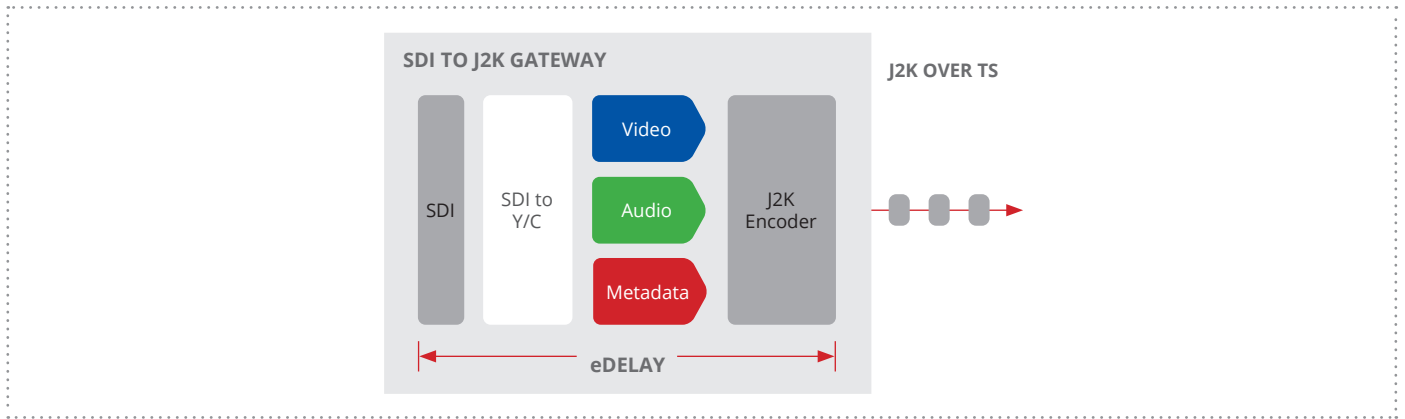
Figure 6. J2K encoder delays (eDelay)

Essence	(eDelay)
A	2 frames
B	2 frames
C	2 frames

Table 4. J2K encoder (eDelay) with/without ST2022-7

The J2K decoder block is composed of a buffer for ST 2022-7 and network jitter. The buffer delay is either 2022-7 (A, B, C or D, please see Table 5.) or a manually configured WAN jitter buffer (when 2022-7 is not used). This buffer has a maximum jitter tolerance of 120 ms as recommended by the ST 2022-2 specification[4].

When ARQ is used, the buffer size will be adjusted based on the number of round trips needed to keep the time for retries long enough to avoid glitches in the essences.
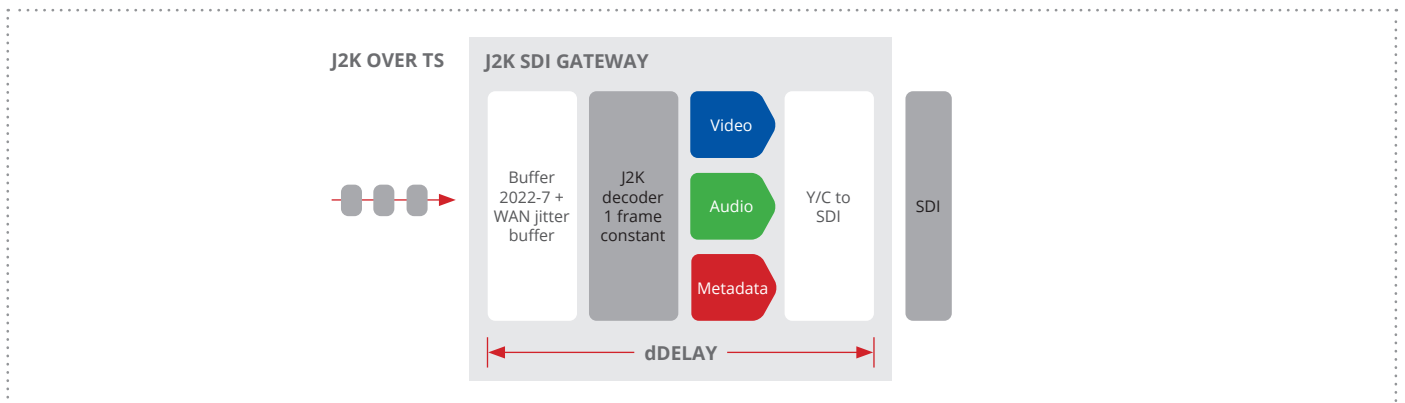

Figure 7. J2K decoder delays (dDelay)

Note: *J2K encode with ULL profile active could reduce the entire eDelay to a minimum of 1/16 video frame processing delay. J2K decoder with ULL profile active can reduce the dDelay to Network buffer + 1/16 frames processing delay.*

---

[4] Riedel MediorNet JPEG-2000 solutions also support 10ms.

Essence	Sync (dDelay) with 2022-7	Sync (dDelay) without 2022-7
Video Audio ANC	1 frame + buffer size (2022-7A: 1frame + 10ms) (2022-7B: 1frame + 50ms) (2022-7C: 1frame + 450ms) (2022-7D: 1frame + 150µs)	1 frame + buffer size (100µs granularity)

Table 5. J2K decoder (dDelay) with/without ST 2022-7

Note: *These delays do not consider wide area networks. Please contact Riedel system consulting to calculate these delays.*

5. JPEG-XS Delays

The JPEG-XS encoders and decoders described here are the ones used from SDI to JPEG-XS and vice versa. To calculate the delay for this type of compression, all you have to do is combine the JPEG-XS delay with the ST 2110 delay. Just be sure not to add the ST 2022-7 block twice, as there's only a single ST 2022-7 delay period present in this path.

The following images show the JPEG-XS encoder and JPEG-XS decoder without any blocks of automatic repeat request (ARQ). Attached with the image you can see the delay of the JPEG-XS encoders and decoders.
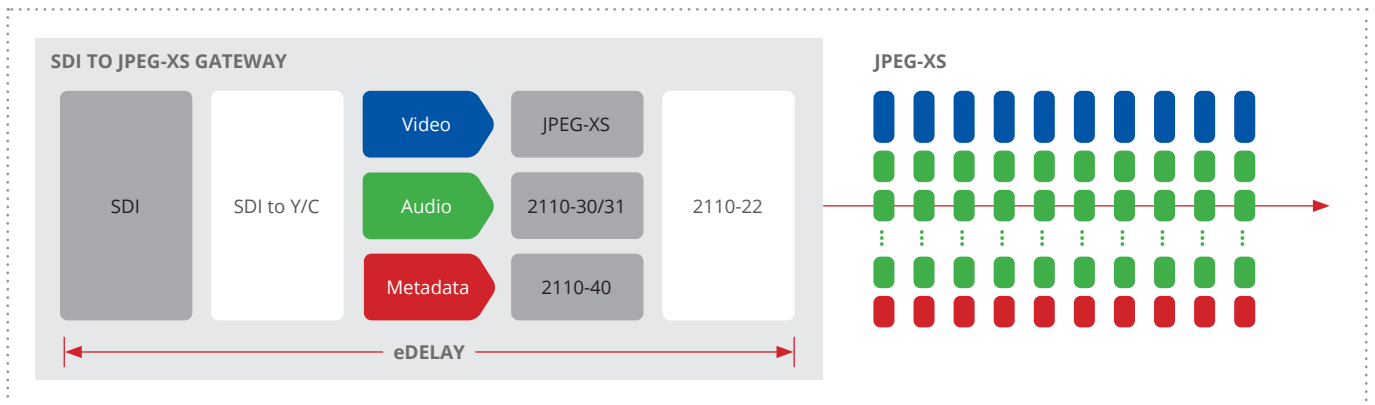

Figure 8. JPEG-XS encoder delays (eDelay)

Essence	Async (Frame Sync on)	Sync (dDelay) without FS
Video	10 lines + 1 frame	10 lines
Audio	Up to 1 frame	1 packet: dependent of pTime (max 1ms)
ANC	Up to 1 frame	Mode 1 (Mbit): 1 frame (1 packet per frame) Mode 2 (MTU mode) : Fill MTU (1500 bytes of data) Mode 3 (Quick): send packet per ANC (packet time at 10G or 25G)

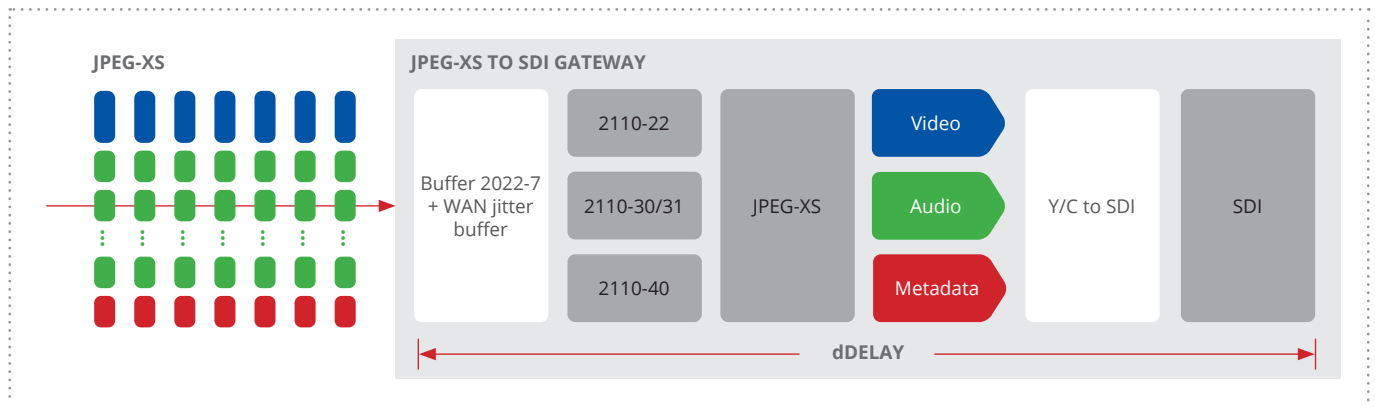*Table 6. JPEG-XS encoder delays (eDelay)*

Figure 9. JPEG-XS decoder delays (dDelay)

Essence	With 2022-7	Without 2022-7
Video	10 lines + 2022-7 delay	10 lines
Audio	Audio is realigned with video 2022-7 delay	Audio is realigned with video. 10 lines
ANC	Always 1 frame (mode M-bit to support all 2110 senders)	Always 1 frame (mode M-bit to support all 2110 senders)

Table 7. JPEG-XS decoder delays (dDelay)

Note: *These delays do not consider wide area networks. Please contact Riedel system consulting to calculate the delay.*

6. 2110 Encapsulator to 2110 De-encapsulator Delays

The following figures give some estimates on delays from ST 2110 encapsulators to ST 2110 de-encapsulators.
Figures 10 and 11 show an encapsulator to de-encapsulator **without frame sync** (input is already synchronized) **with & without ST 2022-7 redundancy** (please note that only 1 network is shown). Figures 12 and 13 show an encapsulator to de-encapsulator **with frame sync** (input is asynchronous / freerun) **with & without ST 2022-7 redundancy** (please note that only 1 network is shown).
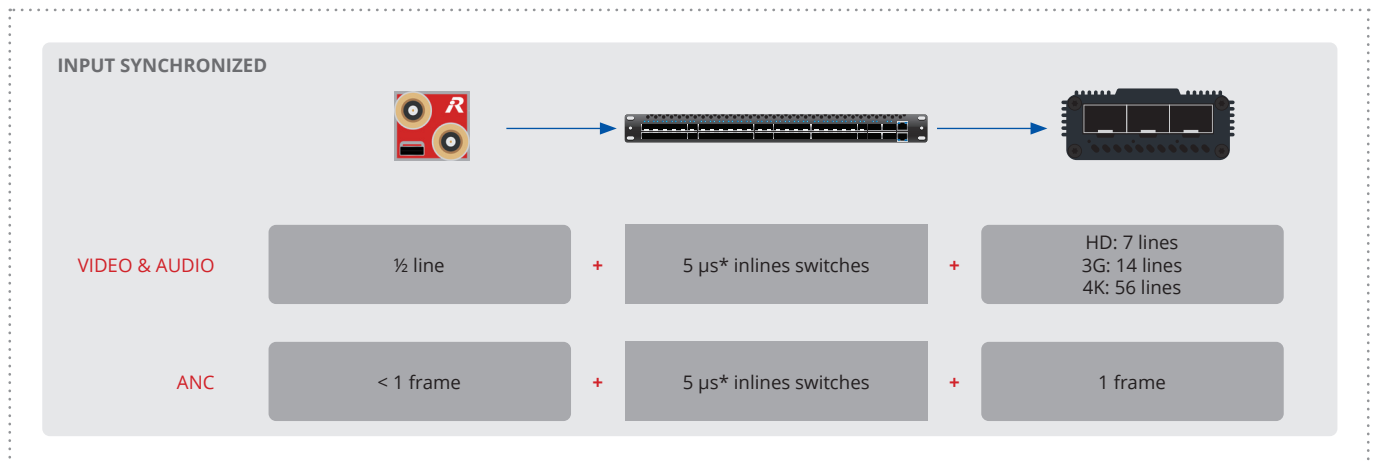

Figure 10. ST 2110 delay without frame sync and without ST 2022-7 redundancy
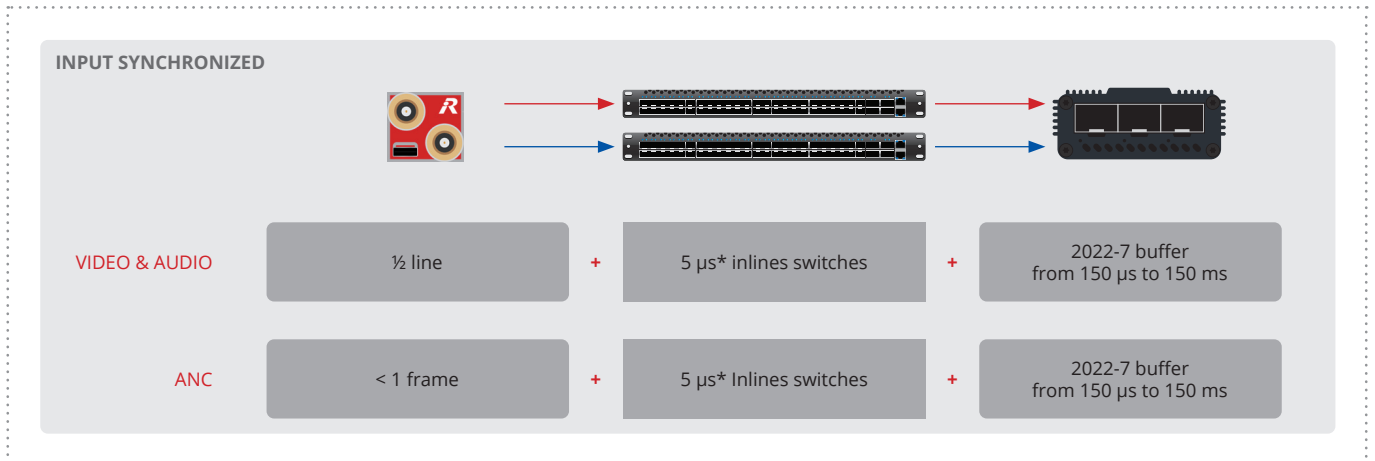
INPUT SYNCHRONIZED

VIDEO & AUDIO	½ line	+	5 µs* inlines switches	+	2022-7 buffer from 150 µs to 150 ms
ANC	< 1 frame	+	5 µs* Inlines switches	+	2022-7 buffer from 150 µs to 150 ms

Figure 11. ST 2110 delay without frame sync and with ST 2022-7 redundancy

ASYNCHRONOUS INPUT

VIDEO & AUDIO	1 frame	+	5 µs* inlines switches	+	HD: 7 lines 3G: 14 lines 4K: 56 lines
ANC	1 frame	+	5 µs* inlines switches	+	1 frame
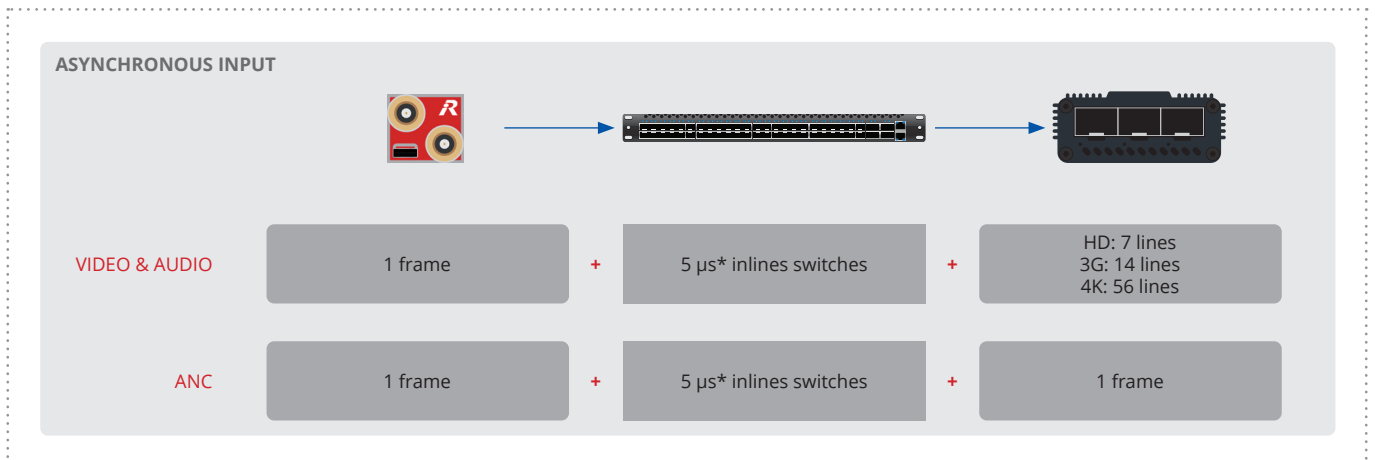
Figure 12. ST 2110 delay with frame sync and without ST 2022-7 redundancy

ASYNCHRONOUS INPUT

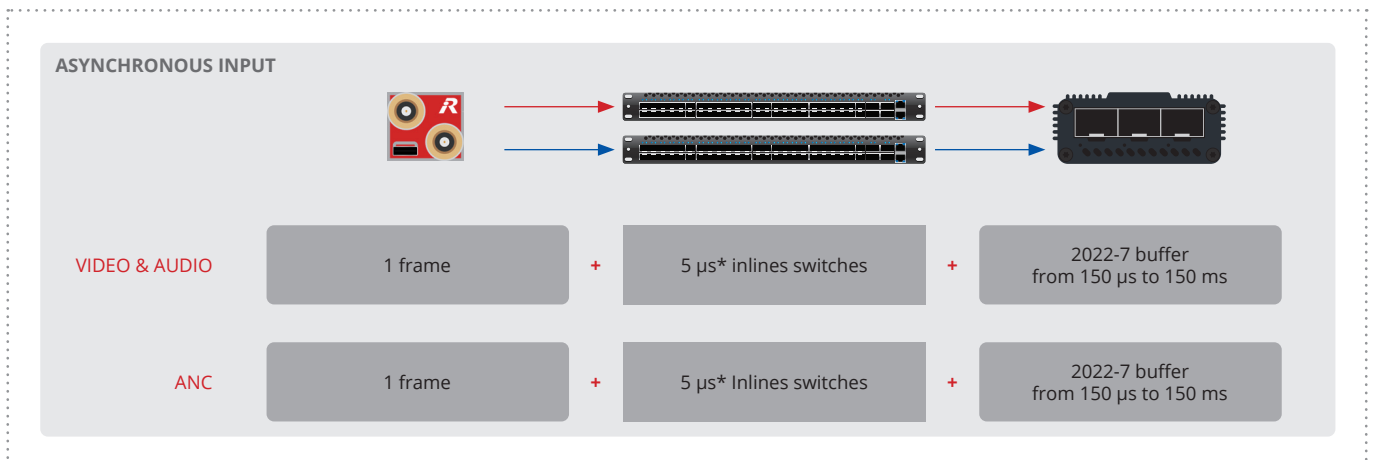| VIDEO & AUDIO	1 frame	+	5 µs* inlines switches	+	2022-7 buffer from 150 µs to 150 ms
ANC	1 frame	+	5 µs* Inlines switches	+	2022-7 buffer from 150 µs to 150 ms

Figure 13. ST 2110 delay with frame sync and with ST 2022-7 redundancy

7. Audio Delay and Lip Sync

The de-encapsulator, or ST 2110 receiver, realigns the RTP timestamp per essence, but does not directly realign the audio on the video timestamp. One important aspect is that the base frequency of the timestamp is different in audio and video: 48Khz for ST 2110-30/31 and 90Khz for ST 2110-20.

This difference is not impossible to overcome. It does, however, make the process less trivial. More importantly, an 'automatic' alignment makes the audio move in time, which creates pops and glitches! To avoid such problems, you can configure the MediorNet IP receiver manually to compensate the delay – and therefore the lip sync. After this compensation, the devices will keep this in memory and no more lip sync errors will be observed.

Please note that having a PTP in ST 2110 does not ensure lip sync, because it is the receiver devices that align the essences.

8. Conclusion

As stated in the introduction, this white paper aimed to shed some light on the various delays inside encapsulators, de-encapsulators, and switches, and explain the various buffers (ST2022-7 and network buffer) and their impact on delay.

All in all, the video over IP path does not exhibit a large delay compared to traditional routers, and the large delay in the encapsulator (eDelay) is present in the same blocks as in SDI (frame synchronizer). On the de-encapsulator or receiver, the delay is adjustable depending on the system and the network's expected delay. For a studio, the redundancy could be minimized to make the dDelay imperceptible.

So you might ask yourself: If I replace my SDI router with an IP system, what's the delay I should expect? Well, there's no straightforward answer to this question, because the delays in IP systems are determined by a range of different factors. Just to name a few: Are your inputs synchronized with PTP? Are they compressed and have to be decompressed or decoded (and then encapsulated in the network)? How many switches do your signals have to travel through? There's no easy equation, no clear-cut rule of thumb, because it all depends on the properties of your network.

If you are looking for an accurate calculation of the delays in your future system, we recommend you contact a Riedel system consultant, who will be happy to answer all your questions and help you design the network to suit your needs!

To make a long story short: In a studio with a thoughtfully designed network, the total delay is comparable to traditional SDI.