A Cloud-Capable Synchronized Transport Architecture for FM and HD Radio Broadcasting

Philipp Schmid
Nautel Limited
Halifax, Nova Scotia, Canada
pschmid@nautel.com

Abstract – Radio broadcast infrastructure and deployment architecture is continually evolving to provide a reliable and secure on-air broadcast service embracing state-of-the-art technology in the broadcast plant to deliver new FM and HD Radio services efficiently. Can a software-based air-chain deployment be coupled with a broadcast transmitter? Can we leverage server and virtualization technology developed for cloud applications? Can we make use of IT standard resiliency and security improvements to better manage centralized air-chains back to the studio and content generation? Today's broadcast architecture using purpose-built embedded devices often deployed at the transmitter site does not interface well to a mission-critical data center lacking a transport stream with sufficient synchronization and security provisions.

An integrated FM and HD Radio transport architecture that allows for a location- and platform-agnostic air-chain architecture bonding all broadcast services into a single transport stream that ends at the transmitter exciter and originates from software-based air-chains (SAC) is presented and demonstrated. Software based air-chain services (ACS) can be cloned, virtualized, or containerized allowing for a common implementation regardless of the underlying host platform, whether this is a physical on-site server, an on-premises data center, or even cloud services like Amazon Web Services (AWS). To maintain consistent timing across the entire air-chain, all air-chain services are synchronized to the high-quality crystal in the broadcast transmitter all the way back to the air-chain ingest in the cloud, on-premises, or local server without the requirement for GPS synchronization.

This architecture definitively solves the FM-HD1 time alignment problem challenging the industry today. Critical to the architecture is a network protocol, based on established MPX and E2X industry standards, along with a synchronous buffer management algorithm that provides a superior level of redundancy and resiliency via multiple concurrent synchronous air-chains that always keep the transmitter on-air. Maximum network reliability through multiple TCP paths can be achieved. Adopting an IT standard client-server paradigm allows broadcasters to adopt state of the art security practices and authentication between transmitter and air-chain services. This shows that this architecture is not only cloud-capable but also suitable for on-premises deployments leading to a flexible, simpler, more secure, and more efficient conversion to HD Radio broadcasting on a modern digital capable broadcast transmitter adopting this architecture.

Why Software Air-Chains? Why the Cloud?

"What makes servers so special?" asked Greg Shay, CTO at The Telos Alliance [1]. His answer was succinctly "The fact that they are not special." If we can define a platform-agnostic architecture based on commodity servers that can prove to be reliable, secure and meets all industry requirements that

This paper is excerpted from the Proceedings of the 2022 NAB Broadcast Engineering and Information Technology (BEIT) Conference, © 2022, National Association of Broadcasters, 1 M Street S.E., Washington, DC 20003 USA.

works in the most challenging use case, the cloud, a milestone proof-point has been achieved. Certainly, the architecture will also work in a more controlled on-premises use case across a studio transmitter link (STL) and it most certainly will handle today's use case with much of the air-chain purpose-built equipment residing at the transmitter site. It does not mean use of the cloud has to be embraced today or even tomorrow but allows for recognition and planning for future possibilities in the cloud today. This new broadcast architecture can be realized today because it is built on industry standard protocols like digital composite MPX for the FM and E2X for the HD Radio broadcast. All air-chain components and building blocks, like the Xperi-supplied Gen4 encoder and exgine modulator, are field-proven and are available today without modification and continue to guarantee HD Radio receiver interoperability through a common air interface implementation; the blocks just have to be recombined under a new architecture paradigm that is location- and platform-agnostic so "non-special servers" may be utilized, whether physical or virtual.

Increasingly, broadcasters are centralizing studios and content distribution across their fleet of stations or affiliated partner stations. The goal is to effectively manage their network by simplifying broadcast transmission sites and keeping complicated system components subject to change in a centrally managed location. Ideally, the implementation is based on non-special industry standard servers or virtualization environments, such that an organization's IT department can service this mission critical infrastructure. Many organizations consider the cloud the ultimate centralized location no matter if it's a corporately managed data center or 3rd party service provider like Amazon Web Services (AWS), Microsoft Azure, or others. Even smaller broadcasters can benefit from shared centralization through simplified management and access to a pool of IT and technical experts with easy and flexible upgrades to HD Radio[1] broadcasting. Software-based air-chain services managed on physical or virtual servers are easily installed, cloned, scaled and provisioned. Nautel's vision is to create a broadcast transmitter that connects to one or more redundant, software-based air-chains that keep the transmitter on-air all the time.

The security concerns mentioned in this paper are of utmost importance for any network paths across the public internet and to the cloud. Demonstrating a solid defense strategy on the public internet also leads to best practices for systems entirely on a corporate wide area network (WAN) that allows you to protect the inner keep of your broadcast defense, the broadcast transmitter. Multiple software air-chains provide an additional layer of security via their expendability through cloning and backup with platform-agnostic deployments across standard server platforms. Cloud and on-premises based studio disaster recovery solutions exist today that are complementary to the multiple software air-chain solution presented here. On the other hand, should your broadcast transmitter or control interface become compromised, an emergency trip to turn off the transmitter with potentially length downtime or backup audio may be required. A backup transmitter may not help if both are sharing the same network resources. Nautel's goal, again, is to use the cloud as a proof-point with wider applicability for all broadcast use cases through best practices.

Fully implementing this vision today is challenging given the lack of an integrated FM and HD Radio transport stream. Bonding all broadcast content into a single data stream at point of ingest at the centralized studio or creation node enables backup air-chains for both FM and HD Radio. Integrated analog FM to HD Radio time alignment is a key aspect required for the broadcast plant to ensure a seamless listening experience when the receiver blends from analog FM to the digital main audio simulcast. The architecture presented herein addresses this requirement without the need of an off-air monitor that reactively adjusts audio delays across the two audio streams as is commonly used today. It is demonstrated that this new architecture never requires any audio delay adjustment over time or any air-chain interruptions including changeovers to hot standby air-chain connections. This important

[1] HD Radio™ is a proprietary trademark of DTS, an Xperi affiliated company. The author is not affiliated with, and this text is not endorsed by DTS.

industry challenge is intrinsically solved in this architecture without the need for tight GPS-based 10 MHz timing tolerances that are challenging to maintain in a cloud-based environment. Seamless time-aligned transmitter cloud connectivity can be achieved no matter the global location of the serving data center.

HD Radio Broadcast Systems Architecture Evolution

To fully understand the limitations of today's HD Radio broadcast architecture, we should review its evolution due to technology advancements and changing requirements since the FCC authorized the use of In-Band On Channel (IBOC) for HD Radio in 2002.

2nd Generation: Early IBOC Adoption

The first commercially available HD Radio architecture topology dubbed Gen2 following an earlier proof of concept 1st generation was made available to early adopters in 2003 for AM and 2004 for FM stations. A simplified architecture topology is shown in Figure 1 depicting functional components in blue rectangles that may be grouped into purpose-built boxes indicated by the dotted black lines. The main audio originates at the studio and is delivered to the transmitter site via a studio-to-transmitter Link (STL), which could be analog audio/MPX, digital audio/MPX, or a switched network connection. This audio is delivered to an audio processor at the transmitter site that often is configured to provide the broadcast transmitter with an analog MPX signal, but other options are possible.

Without an HD Radio signal, the MPX would drive a typical FM exciter and transmitter. In the early days, HD Radio was implemented with the least impact on the FM air-chain. The audio was split at the transmitter site minimizing STL changes, and a second audio processor was used for the HD1 audio that was then fed into an external audio synchronization unit (EASU) and then into the In-Band On Channel (IBOC) exciter. In many cases the IBOC exciter drove a separate digital-only transmitter to either be space or high-level combined with an FM transmitter. Shown in Figure 1 is a low-level combined transmitter that amplifies both the FM and IBOC signal.

The standalone IBOC exciter box was responsible for two functions; encoding the HD1 audio in what later is termed the exporter and modulating the IBOC signal that would later become the exciter engine (exgine). The IBOC exciter required a companion box that is the heartbeat of the entire synchronous

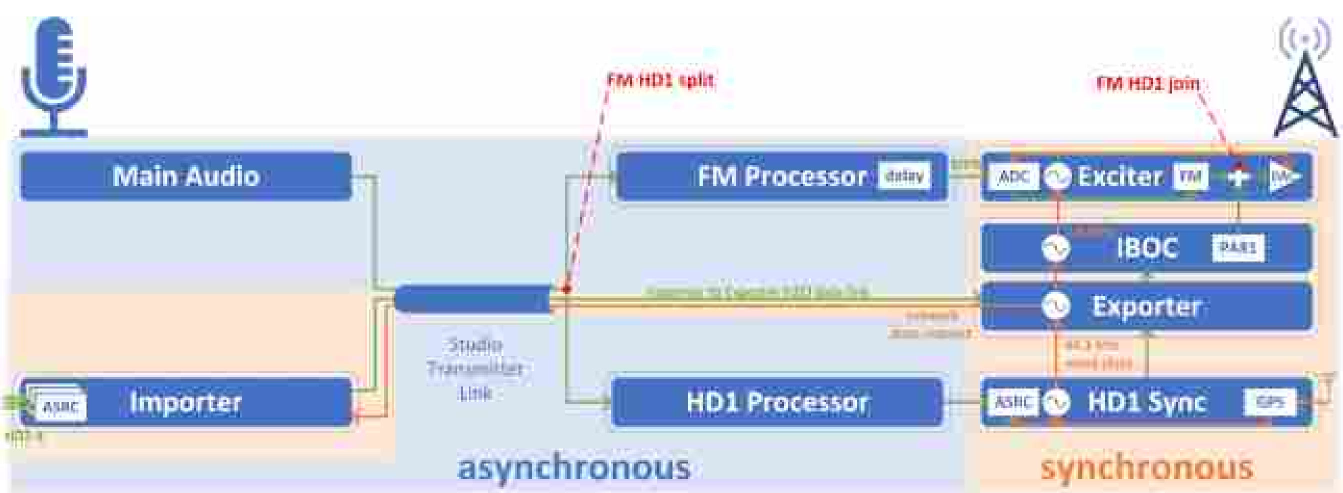

FIGURE 1: 2ND GENERATION BROADCAST ARCHITECTURE PLACES AN IBOC-ONLY EXCITER MODULATOR WITH THE TRANSMITTER. FM MODULATION IS INDEPENDENT AND THE IMPORTER WAS INTRODUCED LATER FOR SUB-CHANNELS.

air-chain of this generation. Based on a GPS reference, it resamples the incoming HD1 audio to a perfect 44.1 kHz sample rate.

In the context of this discussion a synchronous air-chain is any portion of the overall air-chain where all digital samples are fractionally related to each other. In this architecture the GPS module is the root of the clock distribution tree branching out via orange lines shown Figure 1. A synchronous air-chain can span a single or multiple components, provided synchronization exists between all components. For example, the 44.1 kHz word clock drives the exporter encoding rate shown as an oscillator. This oscillator may be a hardware oscillator or a software numerically controlled oscillator (NCO). In either case, operating sample rates are locked in a synchronous air-chain clock tree as shown in Table 1. Note other exciter sample rates are also possible; the ones shown are only exemplary from an early exciter generation.

TABLE 1: IBOC SYNCHRONOUS SAMPLE RATES

	Ratio	Sample Rate	
GPS produces a 1 pulse per second		1	Hz
This drives a 10 MHz voltage-controlled crystal	$\frac{10e6\ cycles}{1\ PPS}$	10	MHz
This drives a 44.1 kHz word clock	$\frac{44100\ samples}{10e6\ cycles}$	44.1	kHz
This generates L1 frames (65536 audio samples)	$\frac{1\ L1\ frame}{65536\ samples}$	0.67291	Hz
This generates 512 IBOC symbols of 2160 IQ samples	$\frac{1105920\ IQ\ Samples}{1\ L1\ frame}$	744187.5	Hz
The 10 MHz drives the exciter MPX ADC sample rate	$\frac{1\ MPX\ Samples}{2\ IQ\ Samples}$	372093.75	Hz
The 10 MHz drives a direct-to-channel exciter DAC	$\frac{640\ DAC\ Samples}{1\ IQ\ Sample}$	476.28	MHz

The key aspect of a synchronous air-chain is that one can obtain any of the intermediate sample rates through precise fractions. Because of this tight fractional relationship, all parts of the air-chain are akin to a fine watchmaker and precision gears and crystals for exact timing. This architecture following the audio flow pushes synchronization from the 44.1 kHz GPS disciplined audio rate out to the transmitter exciter, which ultimately drives the transmitter frequency. This synchronous air-chain can also control delays within the chain by managing the input and output rates of the air-chain.

However, the problem emerges once we recognize the FM and HD1 audio is split outside the synchronous air-chain, which means that portion of the air-chain cannot guarantee fixed throughput delay. This can lead to on-air time differences in the FM and HD1 simulcast as highlighted in later sections. In this architecture the asynchronous branches are constrained to the transmitter site and as long as the audio processors can provide consistent delay the FM-HD1 diversity delay alignment was not a big issue in this architecture. HD Radio has also not achieved critical receiver penetration at that time potentially masking this problem.

The importer was introduced to handle the encoding of HD2 through HD4, and it demonstrated that it is in fact possible to extend a synchronous audio chain all the way back to the studio using the importer 2 exporter (I2E) protocol. Network data requests instruct the encoders to provide additional audio data at precisely the rate required by the exporter data multiplexer and since no time alignment exists for the sub-channels, additional buffering was introduced to ensure a reliable connection to the transmitter site. This extension of the audio chain requires bi-directional communication across the STL and prevents the fan-out of sub-channel to multiple transmitters as they must run at precisely the rate of the exporter encoder.

3rd Generation: Exgine Modulator

Available as early as 2006, the main improvement of the 3rd generation broadcast architecture was the introduction of the exgine IBOC modulator that could now be integrated into the broadcast exciter directly, offering better synchronization internally either through a 10 MHz signal or other clock rates (see Figure 3). While the exporter can still be setup at the transmitter site beside the exgine, this separation from the exporter encoder allows the exporter now to be moved across an IP enabled STL to the studio. A further improvement was the integration of the synchronization unit right into the exporter aiding synchronization at the studio side.

A big driver for this architecture was a) to perform all audio encoding at the studio and reduce STL bandwidth and b) keep mission-critical equipment at the studio for easier maintenance; only an embedded exgine needed to remain at the transmitter site. To facilitate this separation, a new network exporter 2 exciter (E2X) protocol was introduced that could feed the exgine with either UDP or TCP/IP packets allowing for unidirectional or bidirectional IP STLs. A key advantage of this topology is that with the exporter being the root of the clock tree, multiple transmitter branches could be synchronized for a translator or single frequency network (SFN) provided identical content is to be transmitted on-air.

The exciter can be synchronized either through a GPS unit or through a clock packet stream embedded within the E2X packet stream. While the HD portion of the air-chain is synchronous, the analog FM audio now traverses a potential lengthy asynchronous path creating the potential for FM to HD1 misalignment. Early on it became apparent that the network jitter imposed on the E2X clock packet stream posed a significant challenge to exciter synchronization and a reliable HD transport (RHDT) protocol was devised to create a constant bandwidth E2X stream to allow E2X clock packets to traverse the network with reduced network jitter over bandwidth constraint STLs [2]. Even with this improvement, it is challenging to achieve the tight FM to HD1 time alignment specifications required by NRSC-5 of 68 µs [3].

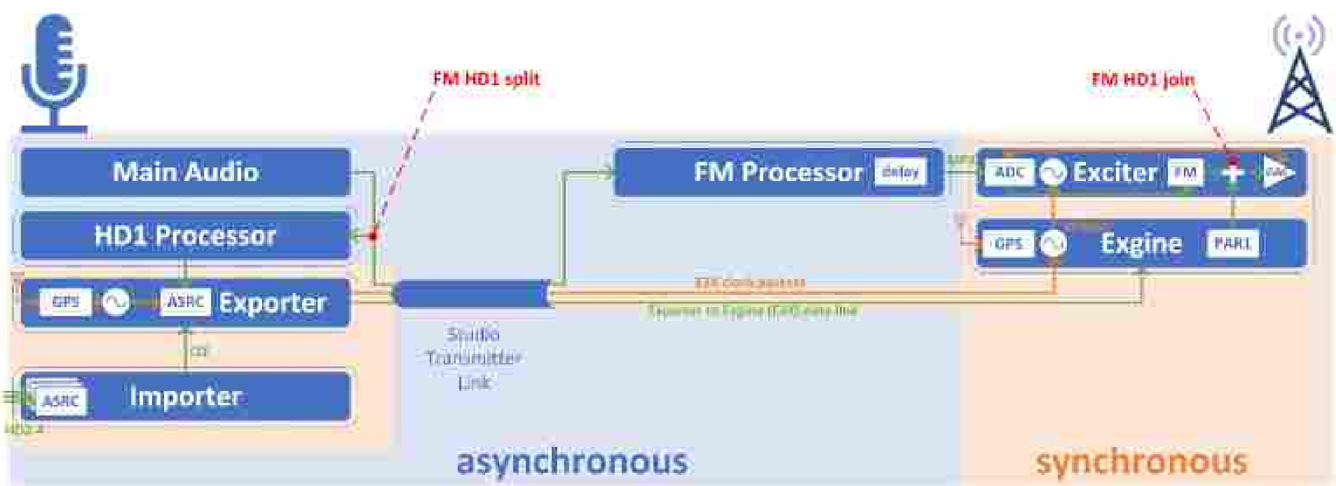

FIGURE 2: THE 3RD GENERATION BROADCAST ARCHITECTURE EMBEDS THE EXGINE MODULATOR RIGHT INTO THE TRANSMITTER EXCITER AND ALLOWS THE EXPORTER TO BE MOVED TO THE STUDIO USING A NEW E2X PROTOCOL. THIS WIDENS THE DIFFERING FM AND HD1 PATHS ACROSS AN STL CREATING FM-HD1 ALIGNMENT CHALLENGES.

4th Generation: Hybrid Peak-to-average Power Reduction

This generation is defined by several key improvements as shown in Figure 3: Hybrid PAR, a combined importer/exporter, dual output FM+HD audio processors, and delay correction receivers.

Hybrid peak-to-average power reduction (PAR) in the exciter was introduced that takes into account the vector addition of FM and IBOC [4] when reducing signal peaks to allow the transmitter to produce a higher output power. Two implementations of hybrid PAR are available today: HD PowerBoost that maintains PAR within the exciter core implementation and PAR2 coupled with the exgine which now also requires the FM signal brought back into the exgine card. A new synchronization challenge emerged in that a batch of IBOC symbols had to be precisely aligned with the corresponding FM rather than sample by sample.

A combined importer/exporter solution allows collapsing two boxes into one, but functionally is equivalent to the two separate boxes. The I2E protocol is no longer required and the synchronization between the encoded HD1 and sub channels is much tighter since the synchronous air-chain distance has been greatly reduced. Subchannels can be encoded right within the Gen4 exporter, or alternatively, the encoder can be placed at a remote location via remote capture clients (RCC) that can be distributed either as software modules or discrete purpose-built encoder boxes are on the market today. The RCC encodes the audio precisely to the data rate required for the on-air broadcast and serves as an effective transport solution across an STL either to the studio or a 3rd party leasing a sub channel. Note that the sub-channel audio needs to be processed prior to audio encoding not shown in Figure 3 for the sake of clarity.

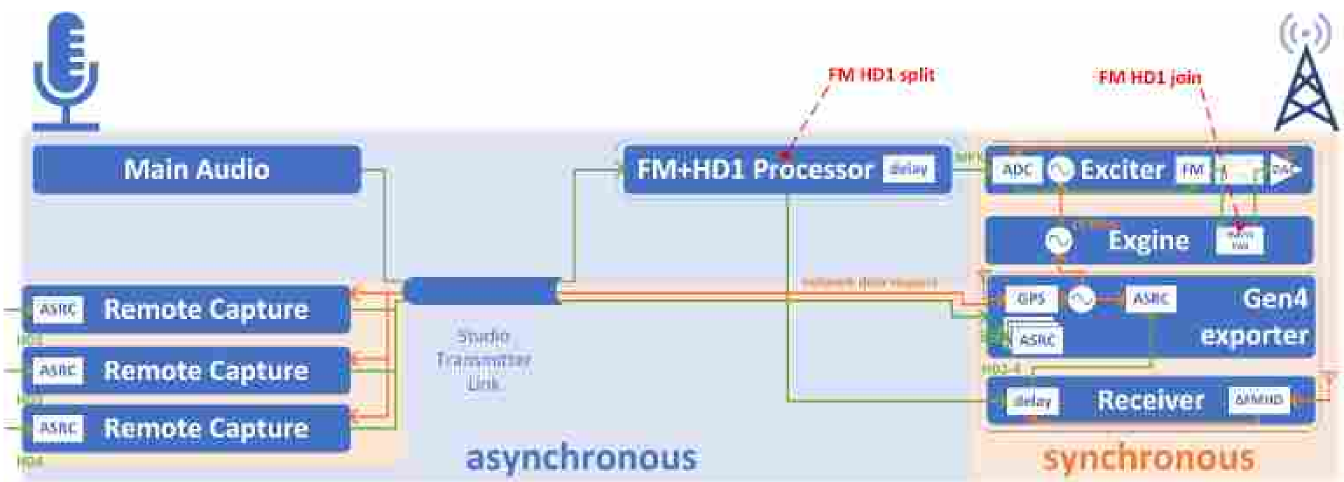

FIGURE 3: THE 4TH GENERATION BROADCAST ARCHITECTURE RECOMMENDS A COMBINED IMPORTER AND EXPORTER TO BE PLACED AT THE TRANSMITTER SITE UNDER GPS SYNCHRONIZATION TO MINIMIZE FM-HD1 PATHS FROM SPLIT IN AN EXTERNAL AUDIO PROCESSOR TO JOIN IN THE TRANSMITTER. A DELAY CORRECTING RECEIVER IS NEEDED TO CORRECT FOR START-UP DELAYS AND DRIFT. REMOTE CAPTURE CLIENTS REDUCE BANDWIDTH ACROSS THE STL.

A dual-output audio processor applies the same base audio processing to both the FM and HD1. It then applies FM-specific processing, like stereo generation, pre-emphasis and composite clipping and limiting to the composite MPX signal only. A dual-output processor has been found necessary to achieve the best possible FM to HD1 blend experience for the listener. Since many commercial stations want to maintain best possible audio processing and loudness control, often the main audio processor is at the transmitter site right beside the transmitter to be able to feed it either with analog or digital MPX. This now necessitates the Gen4 exporter also to be after the audio processor in the air chain at the transmitter site. Having the main FM+HD1 audio processor beside the transmitter exciter means the FM HD1 split to join has been minimized and can be more effectively controlled through common external clocking across all components.

However, all HD Radio generations have maintained the same feed-forward synchronization architecture from exporter to transmitter, splitting the FM and HD1 audio paths over synchronous and asynchronous air-chain portions leading to potential FM to HD1 time misalignment. To address this shortcoming, reactive feedback receivers were developed that measure the FM-HD1 delay and control a delay line to compensate. This delay line can be placed on either the analog or digital audio path provided the unit is told which path it is on to apply the correct delay direction. Often the delay is placed on the digital path by adding additional buffering on the FM to balance the two paths. This way we can continue to feed the transmitter exciter with MPX that is not delay corrected. However, care must be taken should one decide to control the MPX via a dynamic delay buffer since any delay affects the 19 kHz MPX pilot phase and poses a risk to analog FM receivers to switch back to mono reception if lock of the pilot is lost. Until the impact of MPX delay variation is better understood, Nautel recommends employing static MPX delays only.

The Problem of FM and HD1 Time Alignment

As many HD Radio listeners are aware, once an HD Radio receiver is tuned to a station it quickly acquires the analog FM carrier then blends to the HD1 simulcast audio. The blend function ramps down one set of audio then ramps the other audio up over a period of 1 to 2 seconds. If the two audio streams are grossly misaligned whole words may either be dropped or repeated during a transition from FM to HD1, which can become excessively annoying when listening to a station on the fringe of coverage. Listeners have brought back cars to the dealership thinking their radio was broken when in fact the station was misaligned.

Rough alignment can be achieved audibly by tuning some receiver models into “split mode” with the FM audio on speaker and HD1 audio on the other speaker. Human perception of the delay difference typically is limited to around 10 ms, which can address gross misalignment, but falls short of the stringent 3 sample NRSC-5 spec [3] which translates to 68 µs at a 44.1 kHz sample rate. The justification for this stringent specification lies in the audio comb filtering effect during the blend transition. We can model this effect as a Finite Impulse Response (FIR) filter as shown Equation (1), where x[n] represents the analog audio that is ramping down and x’[n] represents the almost identical audio on the HD1 that is ramping up. The z parameter represents an integer sample delay for any delay offset between the FM and HD1 audio. We can now compute the filter’s frequency response shown in Figure 4 based on a variable delay difference (3, 12, 50 and 300 samples are shown) given where in the blend we are. The effect is most pronounced at the midpoint of the blend where both copies of the audio have similar amplitude.

$$y[n] = R_{down} * x[n] + R_{up} * \ x'[n]\ z^{-delay} \qquad (1)$$

From Figure 4 we can see that the filter introduces deep notches that are dependent on the delay factor. While within the NRSC-5 specification of 3 samples, we only see a single notch centered at 7.35 kHz. Should the blend transition occur with the main audio content at that time solely around 7 kHz, the listener would observe a temporary drop in audio level. Typical audio content is well below this frequency; only higher frequencies are impacted briefly.

As the delay increases, lower audio frequencies are impacted with 12 samples centered around 1.8 kHz with a notch bandwidth of 2 kHz. As frequencies are not equally impacted, instead of an audio drop, the listener may experience a wide range of audio artifacts during the blend transition. The process may even make typically inaudible watermarking audible during this time, exacerbating the observed artifacts. As the delay increases to 50 samples, more and more frequencies are notched out. By the time 300 samples (or 6.8 ms) is reached, close to the human ear's perception limit, the entire audio bandwidth is impacted. However, one observation to note is that the notches become increasingly frequency-selective, meaning that only crisp tones within the audio content are affected. While a misaligned blend of that magnitude does affect the blend transition, the effect is that it is less likely to affect large portions of the audio leading to a moderate degree of audio artifacts.

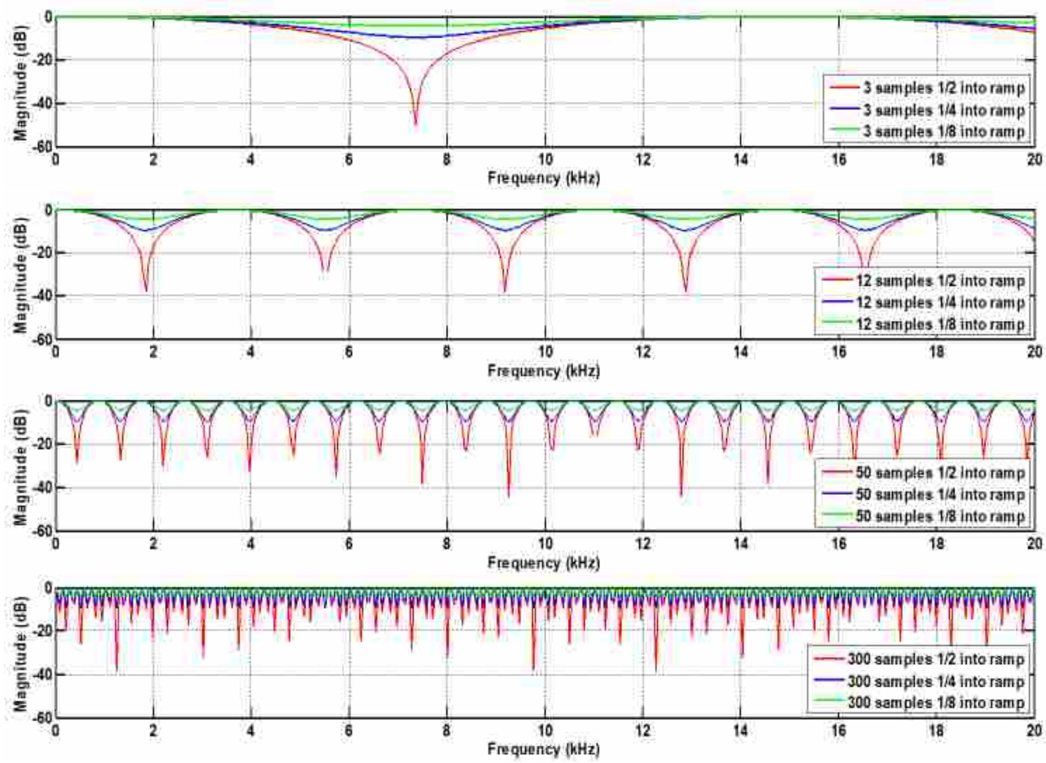

Figure 4: Comb filter effect on Audio during FM/HD1 blend under Increasing Delay. Less than 3 samples presents reasonable audio pass band, 12 samples impacts key audio frequency ranges, 50 and 300 samples impact the entire audio bandwidth but notches are more Frequency selective for subjectively lower blend impairment.

Typical station implementations often experience a delay drift rather than static delay offsets due to a complex broadcast plant. Various audio or network paths get switched in and out or reference clocks are not adequately synchronized leaving audio synchronizers undisciplined. Figure 5 shows typical time alignment drifts observed a 3rd generation topology across an STL with dual parallel FM and HD1 paths. Without a 10 MHz reference (see top plot of Figure 5) the alignment can vary wildly over the course of a few hours traversing the ideal ±3 sample window, entering a region of large artifacts and crossing into regions of moderate artifacts. While offering a large improvement, even with a solid 10 MHz reference at both the studio and the transmitter (see bottom plot of Figure 5), it is challenging to achieve the ±3 sample window across diverse STL paths.

The industry has developed off-air receivers capable of reactively measuring the FM-HD1 delay error and matching the delay across the two audio streams in variable delay buffers; Figure 5 was captured using a Justin 808 receiver by Inovonics Broadcast. HD Radio exporters exist today that can reactively correct for the diversity delay using off-air receivers and consistently keep the delay on-track. Reactive receivers are great at correcting any audio offsets that occur when re-starting any part of the air-chain. However, an underlying delay slip will be compensated by continually adjusting your broadcast audio,

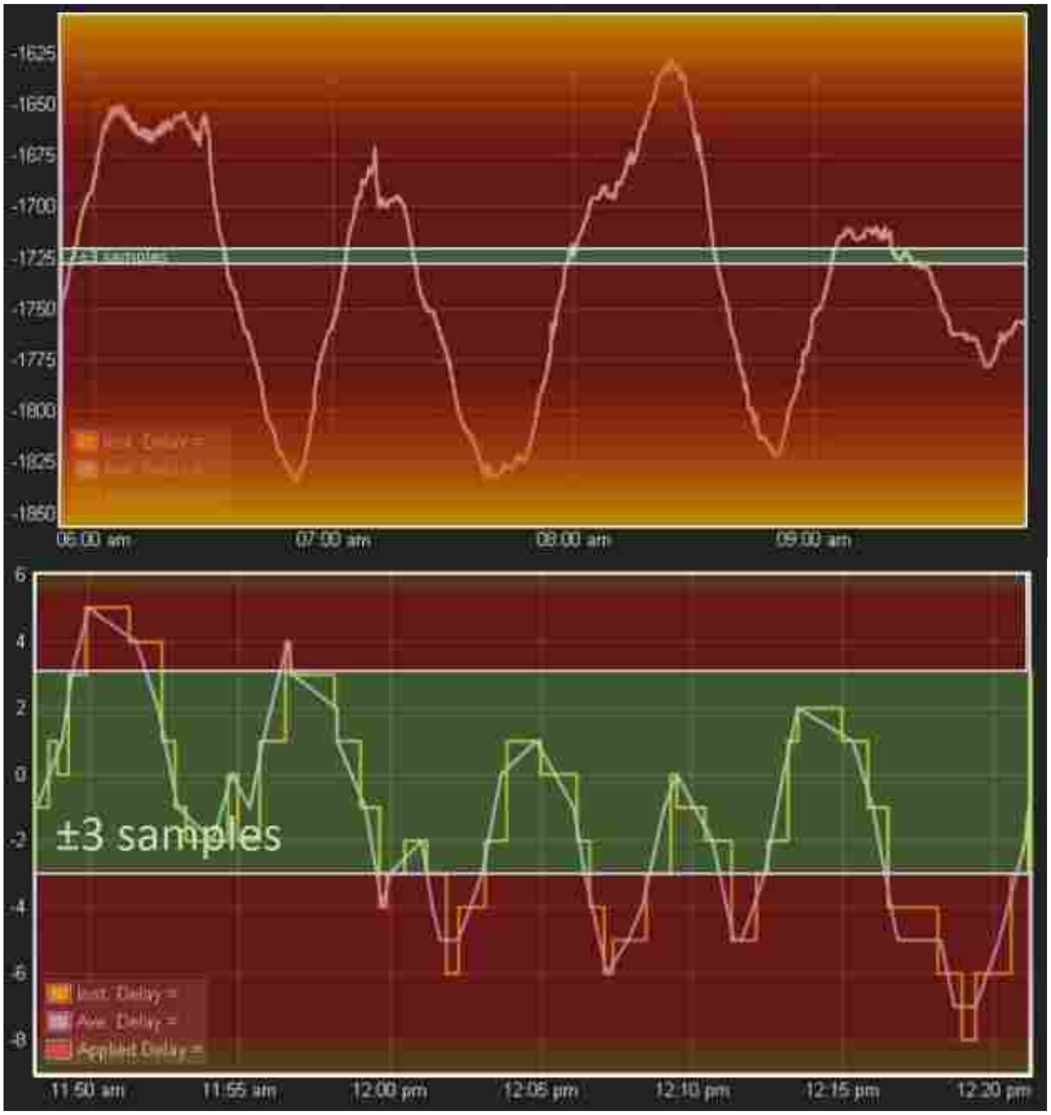

FIGURE 5: TOP: TYPICAL TIME ALIGNMENT ERROR ACROSS AN STL WITHOUT GPS.
BOTTOM: WITH GPS SYNCHRONIZATION [5]. TARGET RANGE ADDED FOR BOTH.
CO-LOCATING LEGACY BROADCAST EQUIPMENT AT THE TRANSMITTER SITE IS RECOMMENDED FOR BETTER TIME SYNCHRONIZATION.

impacting audio quality and possibly hiding underlying systemic issues that should be addressed. Of course, reactive receivers are location limited to the geographical coverage of the broadcast signal.

Caution must be exercised when using reactive receivers in a broadcast plant, as an increasing number of automotive receivers now have built-in reactive delay compensation. Effectively, this cascades two reactive delay adjustment loops along with a dynamically changing broadcast plant. This can lead to instabilities of obtaining a solid diversity delay correction within the automotive receiver forcing the receiver to turn off HD Radio entirely. If a broadcast plant delay is not fixed and stable, there is a risk of the HD transmission being rejected by an increasing number of receivers. The NRSC is formulating a set of specifications that limit the maximum rate adjustment a reactive receiver can apply to a broadcast to minimally impact automotive receivers. The conclusion that NRSC's time and level alignment guide [5] draws is the only solution to avoid reactive correction given today's 3rd and 4th generation equipment is to locate all the crucial HD air-chain components including the main audio processor at the transmitter site.

All HD Radio generations have maintained the same feed forward synchronization architecture from exporter to transmitter splitting the FM and HD1 audio paths over synchronous and asynchronous air chain portions leading to potential FM to HD1 time misalignment. A synchronous air chain from split to join is the definitive solution.

The solution: Synchronous Air-Chain from Split to Join

The lengthy discussion regarding previous generations was necessary to understand the key aspects of this new architecture. For all previous generations, the FM-HD1 path, from split prior to audio processing to join in the exciter, crosses synchronous and asynchronous portions of the audio chain. The only way to guarantee FM-HD1 delay is to encompass the entire path from FM-HD1 split to join in a synchronous air-chain as shown in Figure 6, where only the external audio feeds remain asynchronous.

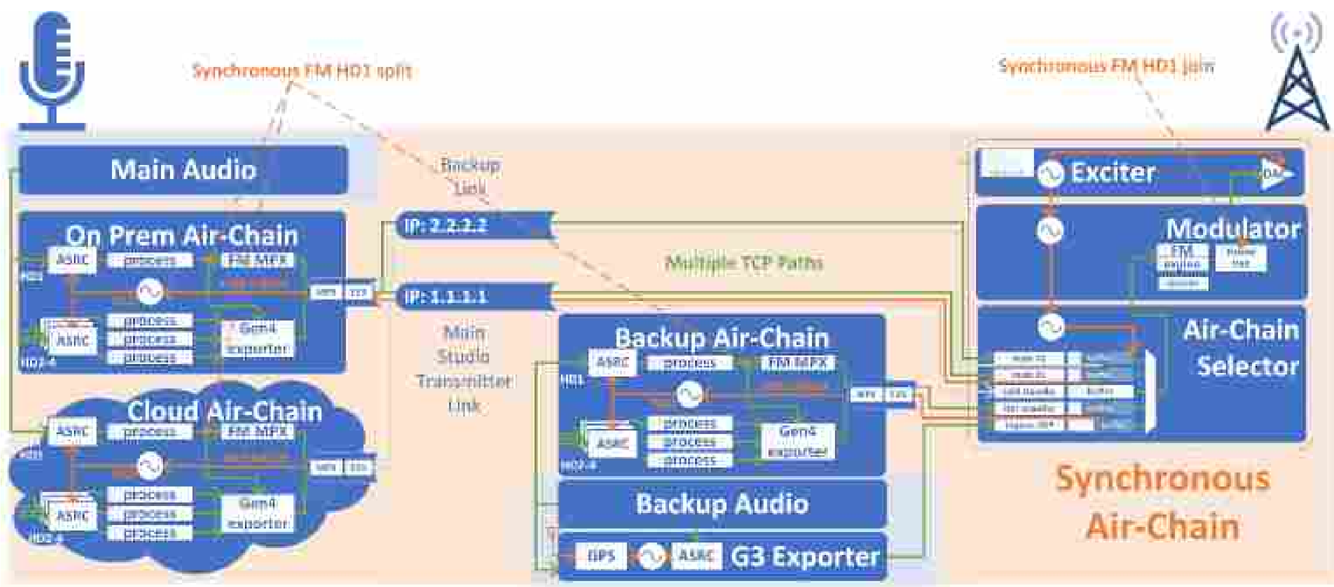

FIGURE 6: SYNCHRONOUS AIR-CHAIN FROM SPLIT TO JOIN WITHOUT ANY ASYNCHRONOUS PATHS. ALL AIR-CHAINS ARE SYNCHRONIZED TO THE HIGH-QUALITY TRANSMITTER CRYSTAL OPERATING ON AN IDENTICAL TIME BASE BY CONTROLLING ALL MAIN AND SECONDARY AUDIO INPUT ASYNCHRONOUS SAMPLE RATE CONVERTERS. THE ACTIVE AIR-CHAIN IS SELECTED BY THE AIR-CHAIN SELECTOR.

This new architecture introduces a further evolution of the Gen4 exporter by coupling it with synchronous audio processing for all HD channels, as well as, analog FM including stereo generation, RDS injection, and composite limiting all in hardware agnostic software modules, termed the Air-chain Services (ACS). The audio processor has several audio input options at arbitrary sample rates, some of which are highly asynchronous, like Internet Icecast connections to an external server with deep buffering and synchronization requirements. It is immaterial to this architecture how the audio arrives at the processor; the key is that the processor sample rate converts the incoming FM-HD1 main audio prior to splitting the audio paths after a common processing stage.

The processor and the exporter are operating in lock step on 2,048 audio sample batch sizes at a sample rate of 44.1 kHz. In turn, 4,096 audio samples produce precisely 1 E2X packet every ~92.8 ms.

In parallel, the audio processor produces MPX samples at a sample rate of 4x44.1 kHz = 176.4 kHz, producing a batch of 16,384 MPX samples for every 4,096 audio samples and every coupled E2X data packet. This allows for definition of a synchronous transport protocol that transmits both E2X and MPX interleaved through a transport pipe like TCP/IP, where samples are strictly locked and cannot pass each other in transit; the same can be achieved using sequence numbers embedded into a protocol but TCP/IP already handles that. Standardizing upon this integrated E2X plus MPX transport protocol will lead to an interoperable ecosystem of products both on the encoding side as well as the transmitter exciter/modulation side, as it is now a single, integrated content flow that carries the entire broadcast signal.

The location of the exgine has not changed, it is still integrated with the transmitter as before. However, the FM audio chain has changed significantly, as now the entire FM content is also carried via the modified E2X protocol that also includes the MPX signal. This allows the flexibility to move the FM modulation process from the transmitter exciter to the exgine itself. The FM modulation process has been greatly simplified as all the MPX generation has been pushed upstream to the audio processor. Now the FM modulation corresponding to a batch of IBOC symbols modulated within the Xperi standard exgine core can be computed on a parallel path in conjunction with the IBOC modulation in a mini synchronous air-chain flow that follows the same principles taught so far on a smaller scale. A massive advantage of the proposed architecture is that it allows a higher degree of sophistication for a variety of main and backup inputs described below. The air-chain selector now provides multiple inputs with synchronized buffers across multiple air-chain inputs.

Table 2 provides an overview of the presented architectures in terms of impact on the studio, the studio transmitter link and transmitter site.

TABLE 2: HD RADIO BROADCAST ARCHITECTURE COMPARISON

	Gen2	*Gen3*	*Gen4*	*Sync Air-Chain*
Transmitter Site Equipment	5 boxes	2 or 3 with external GPS	4 including delay receiver	1 transmitter plus optional backup audio
STL Bandwidth	**1.6 – 4.6 Mbps**	**1.7 Mbps**	**1.6 – 4.6 Mbps**	**0.8-3.0 Mbps**
FM Bandwidth	1.0 - 1.5 Mbps	1.0-1.5 Mbps	1.0 - 1.5 Mbps	2.8 Mbps 600 kbps (compressed)
HD Bandwidth	3-4.5 Mbps (discrete audio) <60 kbps (importer at studio)	~200 kbps	3-4.5 Mbps (discrete audio) <60 kbps (remote capture client)	~200 kbps
Studio Impact	low	high	medium	high
Content flow	forward	forward	forward	forward
Session connection flow	forward	forward	forward	reverse
Sync flow	forward	forward	forward	reverse
Authentication flow	none	none	none	reverse

Buffer Synchronization with Air-Chain Feedback

Unlike previous generations, this architecture supports multiple concurrent air-chains and relies on a set of synchronized input buffers; one per air-chain. This means all air-chains must be clock synchronous. This architecture allows the transmitters high quality crystal to form the root of the entire clock tree rather than a potentially remote exporter clock that could introduce transmitter frequency offsets. The transmitter is already required to have a high-quality crystal to guarantee correct carrier frequency and modulation parameters as defined in 47CFR73.1545 and 47CFR73.1570. By keeping the transmitter crystal constant in an open loop configuration, we make the digital broadcast more robust as receivers are not subjected to swings in the transmitter frequency and OFDM constellation rotations that arises from sudden changes in the transmitter frequency control [6]. Additional GPS synchronization can still be used to control the transmitter crystal but is no longer required to keep the FMHD1 delay aligned.

As stated previously, in a synchronous air-chain the input and output sample rates need to be controlled to be lockstep, but intermediate processing stages enjoy a great degree of flexibility. All air-chains have the same output, the transmitter crystal driving the direct-to-channel DAC. All air-chains may have different inputs over possibly large geographic separation. The key innovation in this air-chain is that the common main audio ASRC located in each of the air-chains can be controlled to govern the chain's throughput delay and align buffer levels to within IBOC L1 frame boundaries. We can now synchronize all air chains together onto a common transmitter time base.

The buffer levels of any air-chain not currently on air can safely be adjusted without any negative impact to the on-air broadcast or affecting the FM-HD1 delay, provided one adjusts the FM and HD1 equally. This means any air-chain that is brought online can pre-charge its buffers prior to becoming active and start reasonably close to a desired target level (see Figure 7). The buffer levels are then continually monitored thereafter.

The air-chains implement a numerically controlled oscillator (NCO) to govern the processing rate of the audio processor. In simplified terms, the air-chain starts a stopwatch based on the underlying platform's time base and counts the number of input samples from that starting point in time. Given the core processing rate is at 44.1 kHz the air-chain compares how many samples have been processed against the number of samples that should have been processed. If the actual number is less than the target number, additional samples will be processed. The initial estimate based on the platform time base will most likely be wrong at first. Even if the time base is GPS, IEEE1588 or NTP synchronized, there is no guarantee the transmitter crystal, while very precise, is perfectly time synchronized.

There is no transmit buffering at the transmit side and all packets are immediately transferred to the exgine, the buffer levels of all air-chains can be monitored at the exgine. The buffer levels will change erratically depending on network jitter and can also be subject to slow-downs due to network congestion. The buffer is averaged over minutes before reporting the average buffer level. The error between the averaged buffer level and the target level becomes an input to a proportional, integral, derivative (PID) control loop commonly used in control theory.

It turns out that only the P and I terms are required for an effective control loop. The P term allows the control loop to react to large errors, such as an initial offset to the target level. The I term provides a long-term average estimate of the air-chain platform's frequency error which can be displayed in terms of parts per million (PPM). This frequency error is only valid after hours of steady-state operation. The PID loop produces as an output a correction of the air-chain platform's time base in terms of additional (or fewer) 44.1 kHz audio samples that are transmitted over the network. Figure 7 shows how the control loop can obtain a desired target buffer level of 24 packets (1.5 IBOC L1 frames) from an initial starting point of 32 packets that were initially pre-charged too high to demonstrate the control loop's ability to govern the buffer level.

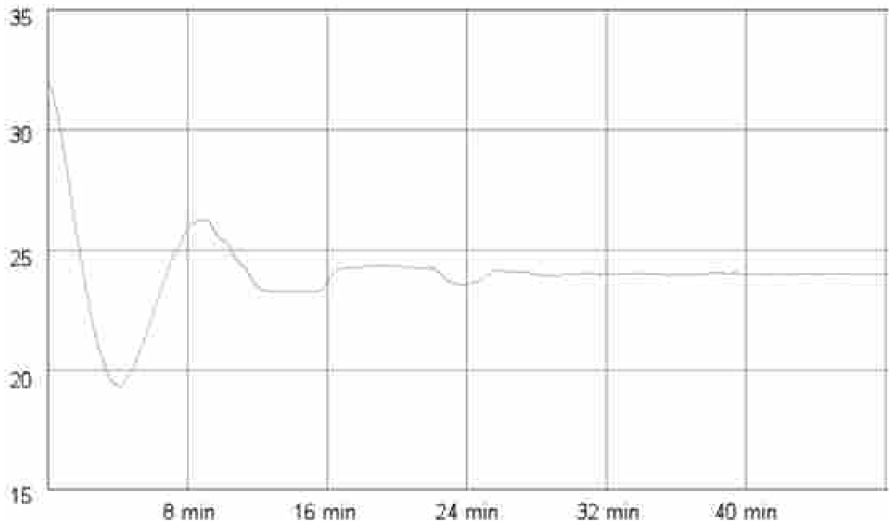

FIGURE 7: AIR CHAIN SELECTOR BUFFER LEVELS SYNCHRONIZED VIA PI FEEDBACK LOOP TO SETTLE ON A TARGET BUFFER LEVEL OF 24 PACKETS STARTING FROM AN INITIAL HIGH PRE-CHARGE OF 32 PACKETS. STEADY STATE AIR CHAIN PLATFORM TIME FREQUENCY OFFSET CAN BE ESTIMATED WITH RESIDUAL ADJUSTMENTS LONG TERM.

A typical implementation would pre-charge close to the target to avoid excessive control loop variation. The natural frequency of the control loop shown is around 10 minutes but may be subject to further tweaking and adjustment given a specific networking environment. Note the step function centered around integer packet numbers when a predominant number of buffer level reads center around the same integer buffer level; the I term will eventually push it to the correct level.

Figure 8 presents the results of an experiment, where 2 air-chains were configured, one on a local HDMultiCast+ server and another on AWS server in the Ohio data center with the transmitter located at Nautel. The on-premises HDMultiCast+ air-chain was pre-charged to a buffer level of 14, while the AWS air-chain was pre-charged with a buffer level of 32. The air-chain integrated Omnia 9 Enterprise OS audio processing module also provides a rate estimation of what it considers its input rate.

On startup, both the transmitter-based estimation and the audio processor estimation exhibit large errors that are induced by the control loop slewing to the desired target buffer level; no significant audio artefacts were noted on either the FM or HD audio. After about 30 minutes, the buffer levels had stabilized producing a frequency offset error of around 35 ppm for the on-premises server and close to 0 ppm for the AWS air-chain after approximately 10 hours of run time. Carrying a reliable, short-term accurate, time base into a virtual environment is a challenge as experienced using VirtualBox on a desktop platform. It appears AWS has effectively solved this challenge; other data centers may not have solved this challenge yet. It is interesting to note that the 1.6 Hz frequency offset reported by the audio processor represents a 36.2 ppm frequency error, very close to the 35.3 ppm transmitter-based estimate. However, both transmitter-based ppm readings fluctuated +-5 ppm due to residual network jitter. The transmitter crystal was running open loop without any GPS discipline demonstrating that GPS is now optional at the transmitter site.

All buffers are pre-charged such that the buffer read operation perfectly aligns with IBOC L1 frame boundaries and are kept in lockstep thereafter. By making the buffer a minimum of 1 L1 frame consisting of 16 packets we can perform a hitless transfer from one air-chain to another air-chain depending on the content switched to. Today the layer 5 audio framing within separate Gen4 exporters is not guaranteed to be perfectly aligned on L1 boundaries, so a 2-3 L1 frame IBOC drop is forced on a changeover causing a 3 to 4s silence on the secondary channels. This drop may be eliminated in the future if layer 5 frames can be aligned in the exporter.

FIGURE 8: INITIAL STARTUP (TOP) WITH ON PREMISES AIR-CHAIN STARTING AT 14 PACKETS AND THE ACTIVE CLOUD BASED AIR-CHAIN STARTING AT 28 PACKETS (LEFT). IMPERCEPTIBLE AUDIO PROCESSOR ASRC ADJUSTMENT (RIGHT) PULLS BOTH BUFFERS TO A TARGET OF 24 PACKETS. ONLY THE LONG-TERM FREQUENCY OFFSET REMAINS WITH THE TRANSMITTER ESTIMATE MATCHING THE PROCESSOR OFFSET (BOTTOM).

A New Transport Protocol Building on Existing Standards

The application layer protocol, from the session management, transport layer and network layer, must be specified; options exist at each level. An initial framework is presented here for consideration of standardization to allow for eventual vendor interoperability.

Application Layer links MPX and E2X

For the presented synchronization architecture to work, all that the application layer is required to provide is a linking between the analog MPX packet content and the E2X packet content that corresponds to the same 4,096 input audio samples at 44.1 kHz. The packet and sample stream must be ordered and transferred without errored or lost packets, a responsibility of the transport layer.

An initial version of the transfer protocol interleaves an E2X packet with 16,384 16-bit MPX samples at 4 x 44.1 kHz = 176.4 kHz sample rate across a TCP/IP pipe for the forward data direction. A new packet for buffer level feedback control is also introduced to adjust the air-chain NCO by a specific number of audio samples. In the future, multiple profiles can be defined to allow enabling MPX or E2X and the substitution of either component with alternate definitions. For example, MPX could be carried as an encoded MPX stream to reduce bandwidth in another profile. The proprietary E2X protocol defined initially by iBiquity, now Xperi, may be substituted by conveying the IBOC constellations occupying the same time slice in a carrier bitmap if an open protocol definition based on the NRSC-5 [7] specification is required.

Session and Presentation Layer Provides Authentication Options

Beyond the MPX and E2X sample transfer, it is recommended that profiles be added for authentication to the protocol. One profile can be based on a private and public key exchange between the transmitter and the air-chain service. The private key resides on the transmitter behind a solid firewall implementation and is never shared publicly and can be uniquely node-locked to the transmitter hardware. The public key needs to be installed on each instance of the air-chain service and can either be floating re-using the same key across all air-chains or can also be uniquely locked to the underlying air-chain platform. In this way, the air-chain service software module can be made available for download publicly or even directly from the transmitter. The private and public key generation will be performed by a trusted body and installed on the air-chain service separately from the installation. One can consider this a basic 2 factor authentication method, where the software module is installed from and linked to the transmitter while the keys are installed by an alternate means. In this way, a level of trust is established between the transmitter and air-chain service.

Only authentication is required by the application, encryption is not required considering the data content will be broadcast publicly on-air within 10 seconds. However, a man-in-the-middle attack must be prevented, where an intermediary node may hi-jack the current session and become an imposter. Encryption methods, like SSL, can solve this problem as the encryption and decryption requires synchronized state on both the sender and transmitter. To avoid the overhead of encoding the entire data flow, which is not required, a running checksum can be injected based on a hashing function like SHA-512 that takes both the private and public key into consideration. A malicious actor would now have to become a man-in-the-middle and learn the running state of the hashing function requiring the air-chain platform to become compromised. Should an air-chain platform or session become compromised by a malicious actor, it can be shut down at the transmitter, then address the attack vector, switch to an uncompromised backup air-chain, re-download the air-chain service or revert to a safe restore point, update the keys and bring the main air-chain back online.

Note that public key infrastructure suffers from a certificate management challenge of how to generate and manage the issued keys. Alternate authentication methods are emerging based on certificateless blockchain technology [8] that may be considered for a future authentication profile.

TCP Transport Defends against Layer 4 Denial of Service Attacks

Broadcasters are in a feed-forward data flow mindset that follows the broadcast paradigm often with a low latency and multicast requirement reaching multiple targets; sometimes even across uni-directional links. It comes as no surprise that previous HD Radio generations placed the TCP server at the downstream device following the data flow. The legacy TCP/IP E2X mode has the downstream exgine listening for an upstream connection from the exporter, the TCP/IP I2E link has the downstream exporter listening for the upstream importer, the downstream importer is listening for connections from the upstream remote capture clients. The Legacy User Datagram Protocol (UDP) protocol also creates a listening UDP socket for incoming connections. Other broadcast streaming protocols like AES67 for PCM audio, Real Time Protocol (RTP), and SMPTE2110 for TV applications, are all based on UDP to allow for IP multicasting. They all introduce open listening ports. It is tempting to add transmitter site firewall port re-directs to allow remote content delivery or transmitter control … don't.

A public listening network socket is a serious attack vector, even if or especially when, re-directing through a firewall:

- **Listening UDP sockets** are vulnerable to **UDP flood attacks**, which "is a type of denial-of-service attack in which a large number of User Datagram Protocol (UDP) packets are sent to a targeted server with the aim of overwhelming that device's ability to process and respond. The firewall protecting the targeted server can also become exhausted as a result of UDP flooding, resulting in a denial-of-service to legitimate traffic." [9]

- **Listening TCP sockets** are vulnerable to **SYN flood attacks**, which "is a type of denial-of-service (DDoS) attack which aims to make a server unavailable to legitimate traffic by consuming all available server resources. By repeatedly sending initial connection request (SYN) packets, the attacker is able to overwhelm all available ports on a targeted server machine, causing the targeted device to respond to legitimate traffic sluggishly or not at all." [10]

It is proposed that the client-server architecture be flipped by making the upstream air-chain service the TCP/IP server and never establishing any listening socket shielding the transmitter from attack. By reaching out to an air-chain server, the transmitter can easily traverse a firewall, without compromising the firewall by opening port redirects or dropping other security guards. Using Network Address Translation (NAT) and IP Masquerading the IP address of the transmitter is never revealed outside the firewall eliminating another potential attack vector. Not only does it solve the firewall traversal problem, one can now reach out to a known location even across the public Internet; session management will authenticate the remote server identity.

A key security aspect of this new architecture is that by making the transmitter the root of your broadcast tree, your transmitter can survive an attack even if you lose an air chain branch. If an air chain is compromised or denied service, drop it and switch to another air chain. Your broadcast continues to be safe.

Lower layer attacks like layer 3 ICMP flood attacks or Smurf attacks must be handled by the firewall and network implementation or the ISP using industry standard practices. Of course, flipping the client-server makes it the responsibility of the air-chain service to guard against denial-of-service attacks. When implementing a cloud solution, this needs to be considered. Fortunately, leading cloud service providers offer DDOS solutions, like AWS Shield or Azure DDOS protection. Load balancers may also be used to limit access to your main air-chain.

Main and Backup Air-chains for Resilient Broadcast Routing

The objective of the network protocol is to transfer audio broadcast content from ingest to transmission. Considering all audio paths from the main air-chain to the various backup air-chains part of our network topology, several different connection types are supported:

Multiple TCP Paths

Last-mile STLs are often the most unreliable link in the broadcast chain. Multiple TCP paths allows for simultaneous utilization of two or more STL paths for improved resiliency to network failure with hitless failover. Depending on the configured IP routing, the entire path from transmitter to ACS can be made redundant.

The connections labelled main #1 and main #2 in Figure 6 originate from the same On Prem Air-chain and terminate at the same transmitter and carry identical content. Using the described synchronous buffer architecture that aligns packet sequence numbers on read, these two connections may be treated as multiple TCP paths providing hitless failover from one STL link to another on complete link loss. This link diversity is like bit splicing in other protocols. Multiple TCP paths have to be accounted for in the air-chain service synchronization as only one buffer can control the air-chain service. The air-chain service can simply ignore all synchronization requests except for the first or main connection, this will automatically synchronize the buffer depth of the other TCP paths, as well.

Multiple Path TCP is accomplished by assigning multiple IP addresses to the air-chain server. In Figure 6 both IP addresses 1.1.1.1 and 2.2.2.2 are assigned to the air-chain server and could either be assigned to two physical Ethernet adapters on separate subnets or both could be assigned to a single Ethernet address. Using the two IP addresses appropriate TCP/IP routing can be established using standard IP practices to have both TCP connections go across separate STLs for all or just part of the entire data path.

The Internet Engineering Task Force (IETF) has approved RFC8684 [11] that shares many design concepts to this implementation using multihomed and multiaddressed hosts:

> TCP/IP communication is currently restricted to a single path per connection, yet multiple paths often exist between peers. The **simultaneous use of these multiple paths for a TCP/IP session** would improve resource usage within the network and thus improve user experience through higher throughput and **improved resilience to network failure.**
>
> Multipath TCP provides the ability to simultaneously use multiple paths between peers. This document presents a set of extensions to traditional TCP to support multipath operation. The protocol offers the same type of service to applications as TCP (i.e., a reliable byte stream), and it provides the components necessary to establish and use **multiple TCP flows across potentially disjoint paths.**

The current Nautel implementation uses two separate TCP/IP sessions and sequences the application stream outside of the connection. However, as Multipath TCP matures, adopting this approach and recommending it for standardization of this architecture will be considered if appropriate.

Hot Standby Air-Chain

Hot standby air-chains are fully connected to a backup air-chain and can be placed on air without the need of connection initiation minimizing down time of the on-air broadcast. By the definition of multiple TCP, the same source and same destination are required for all paths which means the air-chain itself is not redundant. A hot standby air-chain differs from multiple TCP in that it connects to independent air-chains as shown on Figure 6. Simultaneous TCP/IP connections are made to both the on-premises and the local backup air-chain. Each connection fills its respective buffer to the target fill level independently using the buffer control algorithm on each connection. The read side of all buffers is synchronized with the FM and exgine modulators such that on an air-chain changeover initiated by the air-chain selector the exgine modulator produces a continuous stream of IBOC symbols aiding in receiver lock acquisition at the symbol level even if a few IBOC symbols are muted on changeover. A synchronous hot standby changeover maintains L1 frame and IBOC symbol timing for fast receiver acquisition of and fast HD lock on the new air chain. The blend function in an HD Radio receiver produces a pleasant cross-fade to the new audio content.

Furthermore, the buffer read side is synchronized such as to dequeue packets from all connections within the same L1 frame sequence number. For example, if the on-premises air-chain is currently on sequence number 45,641 and L1 frames are cyclical with modulus 16 it is currently on the 9th packet within that L1 frame. The local backup air-chain buffer will be advanced until its read packet sequence number is also the 9th packet, for example 11,705. At least 16 packets are maintained in the buffer, such that even on a complete link failure a graceful changeover may be performed on L1 frame boundaries on the first packet within a frame.

While the exgine modulator is kept synchronous, the modulator is reset to flush its internal state on a changeover. It is possible to perform a changeover without loss of HD lock and a seamless cut-in of the HD audio provided all layer 5 audio frames are synchronized across the two air-chains. Until L5 frame alignment can be guaranteed up to two L1 frames may be dropped on a changeover. The FM content is switched in seamlessly while the receiver is still locked to the old HD stream. Once the HD lock is lost, the blend function performs a pleasant cross-fade to the new analog FM audio content on the main channel. Secondary HD channels will exhibit a 3 to 4.5s silence before producing the new audio content but only if the sub-channel exists within the new air-chain. Most receivers will indicate the channel is no longer available if it does not exist in the new air-chain.

Hot standbys are a great way to perform air-chain software updates. Create a new air-chain, update it, configure it, test it to your satisfaction. Bring it up as a hot standby, changeover, retire the old air-chain all with minimal downtime. Of course, both multiple TCP paths or hot standby double the overall bandwidth requirements unless one can either locally source an air-chain or configure TCP/IP routing across different STLs.

Cold Standby Air-Chain

When STL bandwidth is a concern, a cold standby can be configured to quickly serve up a new connection to a backup air-chain but since the buffer is not prefilled as shown in Figure 6 there is a possibility of exhausting the currently active air-chain on a link failure prior to a new connection being established. Unlike an exciter changeover, even on an empty buffer we can maintain an unmodulated RF carrier at the configured power level eliminating any power ramp up delays.

Cold standbys can take on many forms. It can be a clone of the active air-chain without an active connection. However, more sophisticated topologies emerge. Over a fleet of transmitters, a broadcaster may dedicate a few air-chains for emergency use. As the primary air-chain for a given transmitter fails, it connects to a load balancer system that redirects and connects an emergency air-chain to this failing transmitter air-chain. This may involve waking up a virtual air-chain or container and connecting the appropriate ingest streams automatically. All these steps may take some time. It is possible to increase the transmitter's buffer levels to account for all these steps to minimize downtime at the expense of added throughput delay. The proposed client/server architecture that reaches out to a backup air chain allows us to use standard IT components, like load balancers, to handle many disaster recovery use cases via cold standby changeovers.

Legacy UDP

Legacy UDP used by the E2X protocol today in legacy equipment can be supported, since this architecture is based on established standards. However, since the legacy UDP does not support the clock architecture and buffer management described in this paper, over time the legacy UDP buffer will either underflow, if the exporter rate is too slow or overflow, if it is too fast causing HD dropouts. These can be mitigated by GPS synchronizing both the transmitter and exporter, but then what results is essentially a 4^{th} generation architecture with all its limitations.

As described in the preceding sections, earlier FM HD Radio generations do not handle the analog FM component. In the same manner this legacy UDP implementation also only carries the digital IBOC component and must rely on reactive off-air receivers to guarantee FM-HD1 time alignment. It is however possible to decode the HD1 at the exgine level back into a PCM audio stream to be modulated into FM. While the general audio processing that may have been applied prior to the exporter will be preserved, any FM-specific processing will not necessarily be applied (like composite limiting). This method can, in theory, produce good FM-HD1 alignment but because of the sub-standard FM processing may only be appropriate for a last-defense backup audio stream.

Since legacy UDP relies on a downstream listening port, it is not recommended to use this connection method across public networks. This applies both to this architecture and all HD Radio generations.

Constant FM-HD1 Delay with Synchronous Changeover

To illustrate the architecture's superior changeover behavior, two air-chains in AWS were set up. The transmitter output was monitored with an Inovonics Justin 808 to measure the FM-HD1 alignment and an Inovonics Sofia streaming receiver set to split FM/HD mode to be able to plot audio levels on a graph. The Justin output is shown in the top half of Figure 9, the Sofia output is shown in the bottom half. Air-chain #1 was set to process the audio at a loud level, while air-chain #2 was set to bypass at a lower audio level to be able to distinguish the two sources on the audio level graph.

FIGURE 9: CONSISTENT FM-HD1 ALIGNMENT ACROSS CHANGEOVER. A COLD STANDBY CHANGEOVER MAY CAUSE SHORT SILENCE WITHOUT RF CARRIER DROP, A HOT STANDBY CHANGEOVER BLENDS GRACEFULLY TO THE FM CONTENT, A MULTIPLE TCP CHANGEOVER TO THE SAME AIR-CHAIN IS HITLESS.

The transmitter was also outfitted with dual exciters. To establish a baseline, the impact on the broadcast service was measured using standard exciter changeover with legacy UDP. Since the RF carrier is affected in an exciter changeover, the FM service will produce static depending on the received signal level based on the distance to the transmit antenna until the transmitter has completed ramping to full power; typically, 90% of power is reached within 10s. Long periods of HD re-acquisition were noted after an exciter changeover of up to 70s on several receivers. It is assumed that receivers attempt to decode the old IBOC signal until a timeout is hit forcing a new signal acquisition. This timeout may be receiver dependent. This causes an extended HD Radio outage for the station. It may be possible to SFN synchronize both exciters to solve this problem at added cost and complexity.

In contrast, a cold standby, hot standby and multiple TCP changeover all maintain the RF carrier and transmitter power output level. Figure 9 starts with a cold switch initiated manually at the 60s mark. Since there was no other active buffer at the time the transmitter runs out of modulation content silencing the FM audio shown in the red trace. An HD Radio receiver will hang on to the broadcast for another 4.5 seconds, the depth of the IBOC interleaver of 3 L1 frames, before muting its output by blending back to the silent FM. The second air-chain was brought back online using a manual button push simulating the length of connection establishment. If a connection is established prior to buffer depletion and the interleaver an HD Radio receiver may not experience any down time.

A hot standby changeover will be an instant switch over on the FM and a few seconds and the HD Radio receiver will blend to the new analog FM using a pleasant cross fade. When the receiver blends back to the digital simulcast, the experience will be smooth as the audio is perfectly aligned as shown by the time alignment graph in Figure 9. Switching multiple TCP paths is hitless without any impact on the underlying broadcast. Table 3 provides an overview of the various changeover methods.

TABLE 3: CHANGEOVER COMPARISON CHART

	Dual Exciter	*Cold Standby*	*Hot Standby*	*Multiple Path TCP*
Backup Air-Chain	Yes duplicate equipment	Yes	Yes	No
RF	1-10s to RF on	Not affected	Not affected	Not affected
FM	Initial static	some silence	Instant switch	Hitless
HD1	FM blend 10s-70s	FM blend (short silence)	3-5s FM blend (blends to new)	Hitless
HD2-4	10s-70s silence	<10s silence	3-5s silence	Hitless
FM-HD1 Alignment	No	Yes	Yes	Yes
Supported	All generations	HD Air-chain G3/4 Manual	HD Air-chain	HD Air-chain Bit splicing

Changeover methods are not mutually exclusive and can be combined to best suit the station's needs and requirements based on available STL paths. A multiple path TCP implementation can have a concurrent hot standby air-chain, as well, should the main air-chain and not the STL fail itself. A set of two hot standbys can be maintained from a pool of cold standbys to always have a quick changeover.

Conclusion

This paper presents and demonstrated an integrated FM and HD Radio transport architecture that allows for a location- and platform-agnostic air-chain architecture bonding all broadcast services into a single transport stream that ends at the transmitter exciter and originates from software-based air-chains. Software-based air-chain services can be virtualized or containerized allowing for a common implementation regardless of the underlying host platform, whether this is a physical on-site server, an on-premises data center, or even cloud services like AWS.

To maintain consistent timing across the entire air-chain all air-chain services are synchronized to the high-quality crystal in the broadcast transmitter, all the way back to the air-chain ingest in the cloud without the need for GPS synchronization. This makes the main FMHD audio processor a key component of the air-chain. Critical to the architecture also is a network protocol, based on MPX and E2X industry standards, and a buffer management algorithm that allows for multiple concurrent synchronous backup air-chains that always keep the transmitter on-air with cold standby, hot standby and multiple TCP connections. A standardization effort on the network protocol between transmitter and air-chain services will lead to an eco-system of interoperable products from multiple vendors. Future work on integrating MPX compression codecs can widen the applicability and scalability of this architecture.

The cloud connectivity presented in this paper is a proof point that this new architecture can be realized and fielded today because it is built on industry standard protocols like digital composite MPX for the FM and E2X for the HD Radio broadcast. All air-chain components and building blocks, like the Xperi supplied Gen4 encoder and exgine modulator, are available today without modification and continue to guarantee HD Radio receiver interoperability through a common air interface implementation; the blocks just have to be recombined under this new architecture paradigm that is location and platform agnostic. The discussed synchronous buffer management algorithm allows for multiple concurrent synchronous air-chains that always keep the transmitter on-air provides maximum reliability through multiple TCP paths. Adopting an IT standard client-server paradigm not used in today's HD Radio broadcast architecture allows the industry to adopt state-of-the-art security practices such as solid firewall rules and private/public key authentication between transmitter and air-chain services a major leap forward compared to today's feed-forward purpose-built embedded broadcast devices.

Acknowledgements

The author would like to thank the team at The Telos Alliance for their expertise regarding audio processors and the virtual studio, the team at Xperi for their support and expertise on HD Radio system components, Inovonics Broadcast for the measurement equipment used in this paper, and my friends and colleagues at Nautel helping me create this architecture and review this paper.

References

[1] G. Shay, "The Virtual Radio Station in the Cloud," in *IEEE BTS Pulse*, 2022.

[2] P. Schmid, "E2X Bandwidth and Bit Error Requirements for Ethernet Synchronization," in *NAB Broadcast Engineering COnference*, Las Vegas, 2007.

[3] National Radio Systems Committee, "HD Radio™ FM Transmission System Specifications," 2016.

[4] P. Schmid, "An Improved Method of Peak-to-Average Power Ratio Reduction for FM+IBOC Broadcast Transmission," Halifax, 2009.

[5] National Radio Systems Committee, *NRSC-G203: IBOC Time and Level Alignment Guide,* 2017.

[6] National Radio Systems Committee (NRSC), "Transmission Signal Quality for IBOC Signals," 2011.

[7] National Radio Systems Committee, *NRSC-5-D: In-band/on-channel Digital Radio Broadcasting Standard,* 2017.

[8] Y. Zhang, C. Xu, X. Lin and X. Shen, "Blockchain-Based Public Integrity Verification for Cloud Storage against Procrastinating Auditors," *IEEE Transactions on Cloud Computing,* pp. 923-937, July 2021.

[9] "UDP flood attack," Cloudflare, 13 March 2022. [Online]. Available: https://www.cloudflare.com/en-ca/learning/ddos/udp-flood-ddos-attack/.

[10] "SYN flood attack," Cloudflare, 13 March 2022. [Online]. Available: https://www.cloudflare.com/learning/ddos/syn-flood-ddos-attack/.

[11] Internet Engineering Task Force (IETF), "RFC8684: TCP Extensions for Multipath Operation with Multiple Addresses," 2020.

[12] P. Schmid, "Radio Air-Chain Innovation Webinar," 2020.