White Paper
Using Livewire+AES67 AoIP Over WAN

Gregory F Shay, CSO, The Telos Alliance
February 19, 2016

AoIP Over the WAN

The intention of this document is to explain and describe the issues, terminology, and conditions that affect the use of AoIP over the WAN. This is done to make the reader aware of what this technology can do, explain some of the tradeoffs, and set the context for further planning and decision-making.

AoIP over the WAN can be implemented with multiple methods. The 'best' method is determined by an evaluation of the tradeoffs and constraints, and the scope of the goal is to take advantage of technologies that lead into the future.

The past established method of audio over the WAN is to use point-to-point data compression codecs. The most up-to-date method is using the features of the new industry standard AES67.

AES67 takes advantage of the hindsight and lessons derived from many years of experience of the vendor proprietary audio network technologies that went into the creation of the standard, plus the inclusion of best-practice technology coming from the VOIP WAN.

For context, the different choices affecting AoIP over the WAN will be listed and explained in some detail.

- Network Bandwidth and Quality
- Multicast vs Unicast
- Wide Area Synchronization
- Packet Format and Data Rates
- Discovery
- Routing Control and Status
- Handoff Between Networks

Network Bandwidth and Quality

The primary economic decision driver is the availability and cost of high-quality bandwidth over the WAN.

When bandwidth is limited, codecs specialize in delivering the best audio quality for the least network bandwidth. When low cost and high-quality managed WAN bandwidth is available, it is more practical and economical to connect AoIP directly over the WAN, and not use codecs .

The crossover point can be considered when the WAN bandwidth cost comes down to the point that the cost of the codec devices and their maintenance, and the audio quality reduction of even the best codecs, tips the balance to send full-quality, linear, uncompressed full-fidelity audio directly over the WAN. The seamless integration and interoperation of the local and remote audio facility, as of one large facility, becomes a desirable advantage.

Solutions:
If the WAN network bandwidth is limited, and codecs need to be used, a codec such as the Telos iPort is a high-channel-count solution, and effectively uses AoIP for the high-channel-count local studio interfaces. When the WAN is limited bandwidth, *and* poorly behaved or controlled (i.e. the internet), a codec such as the Telos Z/IP One handles and adapts to the variations in the network. Using codecs that offer audio, codec control, and routing control via the same network brings integration advantages and hardware cable interconnect reduction.

When AoIP is used to make the local connection to codecs, a migration path is open. When the WAN bandwidth permits in the future, AoIP interfaced codecs could be removed, and the interfacing AoIP equipment then would connect directly over the WAN.

When the WAN bandwidth and quality is sufficient, the full range of Telos and Axia Livewire+AES67 can be used to transport AoIP directly over the WAN, with the highest of audio quality, full control, and integration.

Years ago when the Telos Alliance invented Livewire AoIP, we knew that someday the WAN bandwidth would become sufficient so that the efficiency and utility of AoIP used on a LAN inside the facility would then expand outside and between facilities. We have been expecting this future to happen.

A note about the meaning of "Livewire+AES67":
Livewire has always been a solution for building complete "facilities over IP," meaning not only audio streams, but also device control, GPIO (contact closures), discovery and advertising of audio devices, and the routing of audio and GPIO together, as a set of features and functions.

AES67 is a standard for audio streams and audio connections. At the Telos Alliance, we have implemented AES67 by adding the AES67 industry standard audio stream types into the existing extended set of Livewire features. So "Livewire+AES67" means the best of both worlds, AES67 standard audio AoIP format and operation, plus the rest of the "facility over IP" solution of Livewire, as mentioned above.

Multicast vs Unicast

Traditionally, inside a facility a LAN is completely under the control of the engineers and administrators of that facility. However, the WAN network is often under the control of a different group with other responsibilities, a different business group, a different organization, or simply a third-party vendor.

The practical difficulties of network configuration and management of multicast on the WAN have made it the exception rather than the rule. It is often taken for granted that multicast on the WAN is not possible.

Indeed, it can be technically possible, and in this specific case, if the WAN to be used is under the control of the broadcaster, it may be possible to simply configure the WAN as one large multicast network, and Livewire or multicast AES67 AoIP would simply function as one large extended studio facility. The multicast WAN is uncommon, but a possibility.

AES67 recognizes that, in many or most cases, unicast must be used over the WAN. Vendors of AES67 equipment, to be fully compliant with the standard, are required to implement both unicast and multicast modes, so that the customer can configure equipment to connect over the LAN or WAN freely as the use case of the particular audio stream requires, without needing different equipment, or without requiring preplanning of the local/remote nature of the individual audio streams.

It should be noted that unicast mode can perfectly well be used over the LAN, allowing a total unification of operating mode in either the LAN or WAN case. The downside of unicast on the LAN, is that the property of multicast routing that gives a ‘virtual distribution amplifier on every audio channel’, that all audio channels are available everywhere simultaneously is lost.

Solutions:
Axia Livewire was always multicast-only operation. In cases where WAN multicasting was allowed, multicasting Livewire over the WAN has been used, with great success.

Livewire+AES67 equipment has the ability to use AES67 unicast mode, and is suitable for use when WAN multicast is not possible or not desired.

Wide Area Synchronization

Codecs

Codecs such as the Telos iPort and Z/IP One synchronize as point-to-point links, and bring out audio interfaces with internal sample-rate conversion to decouple sample rates and avoid (or hide) the need for wide area synchronization. The Telos Alliance xNode also contains a special 'Unicast Link' mode that self synchronizes point-to-point AoIP audio connections, and decouples the sample rates by using analog or AES/EBU audio interfaces.

Direct AoIP

With direct AoIP operation over the WAN, the number of channels and the desire for seamless wide area operation make it inconvenient to continue to pass every audio stream through sample-rate conversion devices in order to continue to allow each site to have its own master timebase. Instead, the system integration design simplicity, and relatively low cost of synchronizing all sites to one common timebase is practical.

IEEE-1588 PTP industry standard

AES67 uses IEEE-1588 Network Precision Time Protocol for synchronization. IEEE-1588 has at least three methods for wide area synchronization:

- Independent IEEE-1588 grandmasters at each site, each with a GPS receiver, providing globally synchronized timebase on each of the local networks at each of the remote facilities.
- IEEE-1588 over unicast. Each site requires an IEEE-1588 boundary clock device to translate from the unicast connection over the WAN, to multicast over the local site LAN.
- If using a multicast WAN, IEEE-1588 operates just as on the LAN. A grandmaster is elected for the whole system, and boundary clock devices at each site can be used to offload PTP clock packet traffic from the many devices present on a large scale system.

The identity of any single point of failure is different in these cases. Use of IEEE-1588 unicast makes the grandmaster at the main location drive the whole system. As such, it can be made redundant, as IEEE-1588 provides for redundant grandmaster automatic switchover. The multiple independent GPS receivers make each of the sites operate correctly with each other with any partial loss of the network or any of the sites, but has reliance on the GPS system itself.

Some or all of the overhead cost of independent GPS receivers may be able to be absorbed. The design of a modern facility usually always includes a master sync generator, and the incremental cost of using a GPS referenced master sync generator is rapidly coming down, as many industries beyond audio embrace the use of IEEE-1588 PTP.

Time of Flight Compensation

With all remote locations synchronized to the globally common GPS timebase, AES67 permits audio streams to be carried over an arbitrary time of flight path, and used at the remote site with precise audio timing and phase control.

There is a mechanism inside AES67, a setting named the '*link offset*,' which is a per-stream setting determined by the user. Link offset determines the time offset from the time or origination of the audio stream, to when it is used (mixed) or output, at the receiving destination. This offset takes into account the time of flight, the buffering size needed to cover the network jitter according to the characteristics of the WAN network, and finally any time compensating factor for phase-aligning different audio streams coming from different sources (if any).

This AES67 link offset mechanism solves a number of problems, and enables the user to be in control of the synchronization of the audio, given the reality of speed of light time of flight, and the desired audio parallel program synchronization. For instance, if a production was being created, with a music source at one location, a voice at another location, and the resulting two (or more) programs combined at a third location, the link offset mechanism of AES67 could be used to not only overcome the jitter of the WAN networks in use, but to retime and phase-align the arrival of the two programs at the third location for time-aligned mixing. (Please note: There are limits to what can be corrected for when working with WAN delays. Audio path 'loops' cannot have delay subtracted. Two sites cannot simultaneously hear each other in sync. Time alignment is accomplished by always adding compensation delay, never by subtracting delay that is already there.)

Packet Format and Data Rates

AoIP carries audio data in network-formatted packets. There is the choice of how much audio data to carry in each packet. This represents a tradeoff between the network bandwidth used, and the audio latency (or input to output time delay). Many short packets give low latency audio over the network (for example 250us per packet) but use roughly double the network bandwidth due to more frequent packet headers. Fewer maximum-sized packets give moderate audio latency (for example 4ms per packet) and are most network-bandwidth-efficient.

AES67 defines one required 1ms audio packet format, to guarantee that all vendors can interoperate with at least this format. 1ms audio packets are 'in between' the extremes of lowest latency and maximum packet network efficiency.

AES67 also allows other packet sizes to be used, so it can be optimized as desired (as long as the vendor implements in addition the standard 1ms stream format).

Livewire+AES67 equipment provides three packet latency formats: the 1ms required by AES67, and two additional, which are the prior Livewire formats, *Livestreams* at 250us, and *Standard Streams* at 4ms. Note that all three stream formats are fully compliant AES67 formats.

(Side note: AES67 is actually so close to what Livewire has always been, that it was only necessary to add the use of the source timestamp in the RTP header to Livewire stream format to make it into AES67. Furthermore, legacy Livewire devices *ignore* this source timestamp if present, so AES67 streams can be received by legacy Livewire equipment, provided a sync bridge from 1588 to Livewire sync is made using any one xNode device.)

Discovery

Discovery means the process by which you can find, and browse, audio devices that are presently attached to the network. Automatic discovery is done by having each device advertise itself on the network (typically multicast), and then each listener builds a list of all devices present. It is a convenience feature for locating, listing, and managing devices, as well as for programming or setting up audio connections.

Discovery on the LAN

The state of the art of discovery protocols for most audio networking products make the assumption of everything being at one location, and advertising and gathering one master list. Discovery of *all* devices on the network is a handy feature to have, but as the list of devices grows, the convenience of having a single list can diminish.

Discovery on the WAN

Particularly in the case of widespread systems over the whole WAN, it may not be most desirable that a list of all devices in the entire system, every audio endpoint, is presented to the operator.

In the WAN case, there are many more audio signals that are local to each site, and except for perhaps maintenance-checking reasons, would not want to be made visible to other locations. Operationally, a subset of specifically 'exported' audio signals from each site, is all that is likely needed and desired. This 'export list' is determined when the system is designed, and is not a simple function of automatically listing all audio devices. The export list would be compiled by hand, and entered into a customer inter-site routing management software control system.

Interaction of Discovery on the LAN and WAN

The ideal combination of discovery on the LAN and WAN would be in three parts:

1. Discover all audio device channels in the facility, on the local LAN.
2. Designate a subset of these audio channels that need to be 'exported' over the WAN. Make these known to some agent, human, or software management program.
3. At each of the local sites, introduce the lists of exported channels from the other sites onto each of the local sites, via the discovery protocol in use on the local LAN. This would be a software program that acts as an agent or proxy, transmitting the information about the remote sites channels using the local discovery protocol.

When Discovery Is Not Needed

Keep in mind, that if the audio stream network addresses are managed and known ahead of time, there is no requirement nor need to use a discovery protocol. Audio switching and routing control is always done using the address information. This is analogous to typing a web address that you know into a web browser versus using a search engine to find what you are looking for, when you don't know where it is.

Scalable discovery on the WAN is a recognized need. There are some vendors that are studying how to bring solutions to market, and also proposed standards in development. However, because the use of audio channels between sites of a large multi-site operator is typically very closely tied to overall system operation and workflow. The desire for *careful*, *deliberate,* and *failsafe*, means the task of controlling audio switching and routing is often handled by custom-designed or configured control system and software. The quick and easy usefulness of discovery, choosing an audio endpoint from a live drop-down list, can make for a higher risk of operator error.

Product Capabilities

Livewire+AES67 devices today contain multiple type of multicast based advertising and discovery.

1. Legacy Livewire advertising and discovery mechanism
2. Bonjour advertising and discovery
3. Planned to be added: SAP advertising and discovery. Implementation of the use of SAP has been just very recently embraced by the AES67 vendor community.

All of these protocols are directly useful for the case of the local facility, and can be interfaced with by some custom higher level inter-site 'audio device directory management' software. All Livewire+AES67 devices have built-in control protocol over the network, which allows simple and seamless integration of routing under the command of local or wide area routing decisions.

Routing Control and Status

Use Routable Networking Protocols

The fundamental building block for implementing control and status monitoring that can be extended over the WAN, is to have each piece of equipment capable of network control and status, using TCP/IP with its own IP layer 3 unicast address.

Equipment that tries to maintain an older style of control via serial port, or by using network communication via raw ethernet, multicast, or some other protocol less than TCP/IP, will create barriers in the extended network system. Look closely at the specifications to verify TCP/IP is used, not a more primitive (and less extensible) network protocol.

Web Browser Interface

A built in HTTP server allows quick and efficient configuration, control, and monitoring by personnel using a standard web browser. The web browser interface itself is over TCP/IP, so also can be routed and reached remotely. As the scale of a system grows, and the number of devices in use grows, the use of an HTTP web browser interface isolates the control of different versions of the product software from the user.

Prefer Not to *Require* Control Utilities

If the studio is designed with equipment that has specially written configuration or control utility programs that must be installed and run on customer PCs, and is the only way to configure, manage or control audio equipment. This method brings with it the need to always have that PC present, and the proper version of the utility loaded for the version of the OS and for the version of the equipment. It may be quite difficult to enforce that all equipment has the same software version to communicate with one central control computer.

Open Control Protocol

To be sure, software control programs and utilities are powerful and valuable, but they should interface to the equipment via the routable network protocols, using an openly published control protocol if an industry standard control protocol is not available.

SNMP is a commonly used control protocol for monitoring and some limited control. The newly published AES70 control protocol for audio equipment is a strong candidate for future implementations, but is too new and only at the very beginning of any vendor adoption.

It has proven to be a large value for the end customer to be able to write small custom control programs, *as needed,* to fine tune workflow and user interaction with the individual device and the overall system. An open, published, and easy-to-read and understand protocol simplifies doing this. A complex, binary, closed protocol coming from a closed, required control utility, does not make this simple and often presents a barrier.

Scalable and Remote Control

As the system grows (over the WAN), it naturally becomes more complex, and a hierarchy of design is often desired—higher level 'commands' that dispatch multiple low-level commands to individual pieces of equipment.

Doing this remotely, controlling operations at remote sites over the WAN from a central location, or in an emergency allowing control from other locations than *only* the central core, is the reason for the importance of equipment to use routable TCP/IP control, as stated first in this section. Lesser control protocols that are not as scalable and routable as the network itself create barriers. And workaround solutions to add network routing communications to equipment by adding additional hardware, is less reliable, increases complexity and opportunities for things to go wrong.

High Level System Control

AoIP equipment with built-in network routable control allows the use of high-level master control software. The use of a routable network protocol allows the high-level master control system to be centrally located, and to remote control equipment over the WAN.

Unification of Audio and GPIO Over IP

Along with audio over IP, there are still many interface uses for contact closures in / out, called GPIO (General Purpose I/O). Carrying GPIO over the IP network is another step closer to the "Facility over IP," concept mentioned earlier. There is a large benefit for seamless integration by routing audio together with its associated GPIO. The GPIO bits do not have to be actually embedded in the IP audio stream (they are not part of AES67, for example), but being on the same IP network, and routed simultaneously side by side, the IP network performs the job of the multiplex.

Product Capabilities

Livewire+AES67 equipment has all of the above mentioned, scalable network control in over TCP/IP using simple human readable ASCII text-based protocol named *LWCP (Livewire Control Protocol)*, as well as SNMP for basic system monitoring. Built in HTTP servers allow configuration, control and monitoring using a standard web browser. In the xNodes, every audio channel is metered and brought out via the *LWRP (Livewire Routing protocol)*, allowing built-in monitoring of audio levels and implementation of external silence detection alarms and audio levels verification. LWRP is also used to carry and route contact closure GPIO alongside audio routes.

For high-level system control, the Axia **Pathfinder** line of software and hardware provides feature-complete solutions that are easily user-customizable, and that control the Livewire+AES67 equipment over the network via the TCP/IP protocols. High-level 'salvos', called 'stack events' (or 'logic flows' in the second-generation Pathfinder Core PRO routing appliance) can be easily defined to customize the desired system reaction to any high-level signal (via GPIO), button press, or automatic reaction to monitored audio levels (silence detection, etc.).

Handoff of AoIP Between Networks

The design of a large-scale system spread over a wide area, using AoIP and control over the WAN, has to take into consideration the human factors for large personnel teams working efficiently together. Even if a technology may hold the ability to be completely centrally organized and managed, it may be found practical and expedient to split up responsibilities over parts of the system, to different persons. Often, the WAN network itself is not co-owned by the end facility owners, and a point of demarcation with a defined service level agreement contract forms the interface between the organizations. For large systems, a divide-and-conquer subsystem decomposition must be used to keep the system manageable.

AoIP Demarcation Points

With prior hardware-oriented audio technology, a suitable set of physical audio connectors was typically chosen, and designated the point of demarcation between organizations. Now with AoIP, there may be physical *network* connectors present, but a physical network connector does not represent a handoff interface of the audio streams. Audio streams flow *through* any given network jack, but the AoIP stream originates, is consumed, is controlled, and is monitored physically elsewhere, not at that network connection jack.

In other words, a network physical handoff point is not a demarcation point of audio control, for which to answer such questions as:

1. Is the AoIP stream present and correct from the sender, at the demarcation point?
2. Can I confirm the consumer of the stream is taking the audio from the demarcation point I am supplying it to?
3. Is the audio at the demarcation point silent or not?
4. If I want to switch the audio source, which is feeding one of my demarcation point inputs, can I do this without taking control of the other organizations equipment?
5. If audio is in error, on which side of the demarcation point is the problem?

The solution to this issue for AoIP comes from using well-established techniques from SIP telephony, which solved the exact same management and organizational issues more than 15 years ago, by the use of *SIP Registration, SIP Trunking,* and *SIP Back-to-Back User Agents.* This is indeed one of the strongest reasons that the use of SIP was chosen for AES67 unicast mode. The intention is for AES67 to grow in use and practice to be a kind of 'studio quality' phone call, with all of the advanced features and benefits that we have come to expect from how the modern telephone system seamlessly makes connections; routes to where the endpoints are identified (by SIP address); hands off through multiple independent global networks; and keeps track of quality, maintenance, usage and resources consumed (and even bills by the minute!).

AES67 AoIP brings with it more than LAN audio network connectivity, it holds the ability for global WAN connectivity. It is important to note that this capability is not present in the legacy AoIP systems, Livewire, Ravenna, Dante, Cobranet and AVB. This capability in AES67, the use of SIP for unicast connections, is an innovation designed into AES67, coming from the benefit of hindsight for the developers of the legacy AoIP technologies.

Now to briefly overview how these SIP features serve as the AoIP WAN demarcation:

- *SIP Registration* is the process whereby equipment wanting to make an audio connection visible outside of the local facility, *registers* the addresses of those audio channels with an entity called an *SIP registrar*. Typically the SIP registrar is owned by the WAN network provider, has authenticated access, and can permit or deny registration access (for instance if you have not paid your bill, in the commercial case.) Audio channels registered are then exported and put on a list available between sites. This mechanism is also closely related to the WAN discovery function mentioned in section 5 of this document.
- *SIP Trunking* is a kind of a gateway through which you reach remote endpoints by addressing the near end of a virtual trunk channel. If an audio output channel from a local facility is to be exported, that piece of equipment at the local facility doesn't need to be programmed to send its AoIP stream directly to the far end piece of equipment. That may be too much detail, too dynamic, too cumbersome, and unnecessary to know. With trunking, all the local equipment has to be concerned with is pointing to the near end of the trunk, which is fixed and does not change. The rest of the routing is controlled by a separate system to get the signal to go where it is desired.
- *SIP Back-to-Back User Agent* is the specific mechanism that is used to interface between two audio network channels, in the middle of the network, or for instance at the end of the SIP trunk. A SIP back-to-back agent is a small program that is running on a server located at some agreed-on location. This agent program has two SIP interfaces, one facing 'inward' to the LAN and one facing 'outward' to the WAN. The agent acts in all ways as a normal audio device; it simply connects the audio between the two interfaces and provides monitoring features as desired, for the internal audio. In this way, the back-to-back agent has all of the desired properties of a demarcation point, independent access on each side, independent addressability, accountability, isolation of control to each side, and the ability to monitor and verify the presence and quality of audio passing the boundary in each direction.

Product Capabilities

Livewire+AES67 devices contain the AES67 unicast SIP mode, and therefore are capable of operating in the above-outlined methods, provided the SIP system organization and management infrastructure is in place. Livewire+AES67 devices do not themselves *contain* SIP Trunks, SIP Registrar, or SIP Back-to-Back User Agents. These are layers of SIP functionality on top of, and in addition to, the endpoint devices.

It is true that constructing large, geographically spread out, large-scale systems using SIP for AES67 audio is a brand new field, the cutting-edge of modern audio system and facility design. But it draws on more than a decade of VoIP telecom experience, and the right partners will bring over that experience to make solutions for AoIP over the WAN.

THE TELOS ALLIANCE®

THE TELOS ALLIANCE | 1241 SUPERIOR AVENUE | CLEVELAND, OH 44114 USA
MAIN: +1.216.241.7225 | **24/7 SUPPORT:** +1.216.622.0247 | **TELOSALLIANCE.COM**