CCS Audio over IP Tools

White paper

Apr 09

MUSICAM USA

Table of Contents

INTRODUCTION

CCS/MUSICAM USA is one of the pioneering manufacturers of professional IP audio codecs for the broadcast market. We launched our line of next generation IP codecs – starting with the Suprima in 2007. Our expertise in this field has grown along with the various hardware and software solutions that we have developed in conjunction with our customers and other professional broadcast companies, helping them find viable solutions to facilitate their migration to this new IP technology.

This practical guide is aimed at providing a fundamental understanding of all the relevant concepts of audio streaming over IP networks, as well as how to use the different tools provided by CCS/MUSICAM USA to deal with them. Awareness of the issues and the way to deal with them is essential for the broadcaster to achieve a reliable audio connection over IP.

1.- AUDIO OVER IP

Most of the digital audio communications over the years have been achieved on synchronous networks where a common clock is provided at both ends of the connection, so that it is possible to synchronize units with a simple mechanism of clock extraction. Thus, data arrives at the remote end at exactly the same rate and in the same order as it was sent from the transmitter.

In addition to that, the data flow was consistent and so was the delay. Moreover, the type of call tended to be point-to-point and bidirectional.

This is completely different on IP networks, where the flexibility for the type of physical connection is much wider. In addition to that, there is a lack of embedded or separate clock, so a clock synchronization algorithm must be achieved by the IP audio codec to compensate for that.

The data flow is split up in to packets, where each packet in a stream might be treated in a different way from the rest, so that the path from the sender to the receiver might vary depending on several different factors. This might introduce different delays in the arrival time of each of the audio packets, in other words **jitter** in the IP communications.

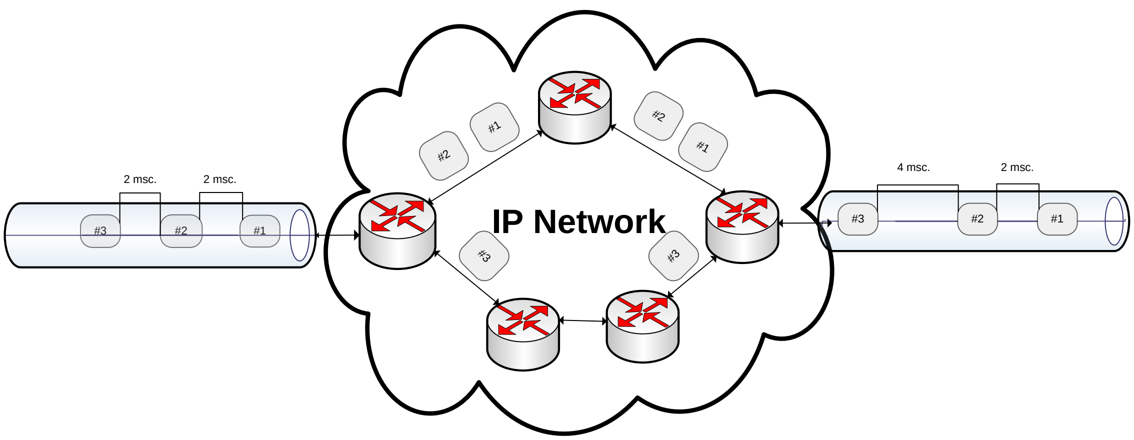

Figure 1: Jitter.

The effect of jitter on the audio communication is potentially a buffer underflow on the decoder side. This data starvation will cause an audio interruption.

This behavior of the network might introduce other important issues, which need to be taken into account, such as lost and disordered packets.

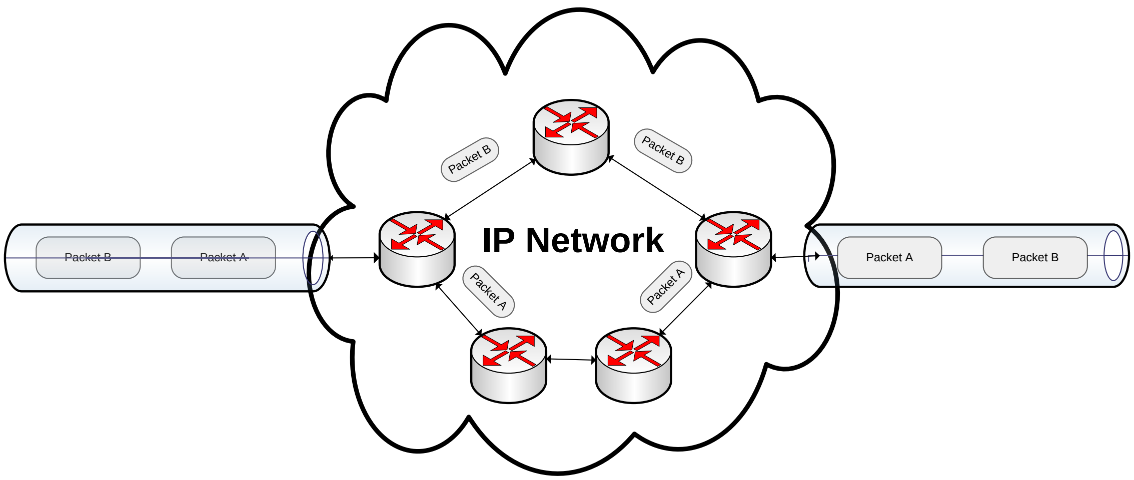

Figure 2: Different paths and packet disorder.

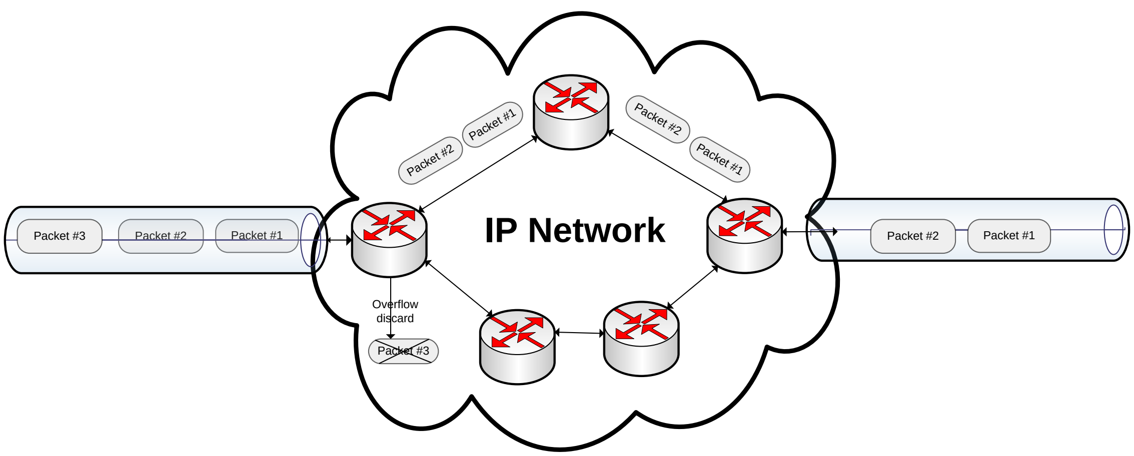

Figure 3: Lost packets.

Usually, one packet is considered as a lost packet when it did not arrive to its destination due to network congestion or any other problem related to the data flow process within the path from the source to the destination.

However, when talking about audio streaming, some packets might not arrive or might arrive too late to be processed. In audio streaming the result in both cases is the same: the packet is considered as a **lost packet** and there is a resulting loss of audio.

This loss of audio might result in an audible glitch depending on whether '**error concealment**' techniques can be applied or not.

To continue the comparison with other digital audio links, IP links need some extra information which is sent over the media along with the actual audio data, in form of headers, wrapping the audio data, which increase the data bandwidth needed to achieve the audio streaming. The size of the audio packets (frames) will affect the overhead percentage so that the smaller the frame, the larger the overhead in relation to the audio data, but the smaller the delay. Therefore, the configuration of this parameter is a trade-off between delay and bandwidth.

Figure 4: Overhead.

That's why the IP audio codec must allow **flexible packet sizes**, so that the user can optimize the unit to fit in with their requirement regarding delay and bandwidth. CCS/MUSICAM USA IP codecs allow the user to configure the packet size for all low delay compression algorithms. We also provide the user with a very helpful information window, which informs them about the final delay and overhead depending on the different packet sizes for the selected compression mode.

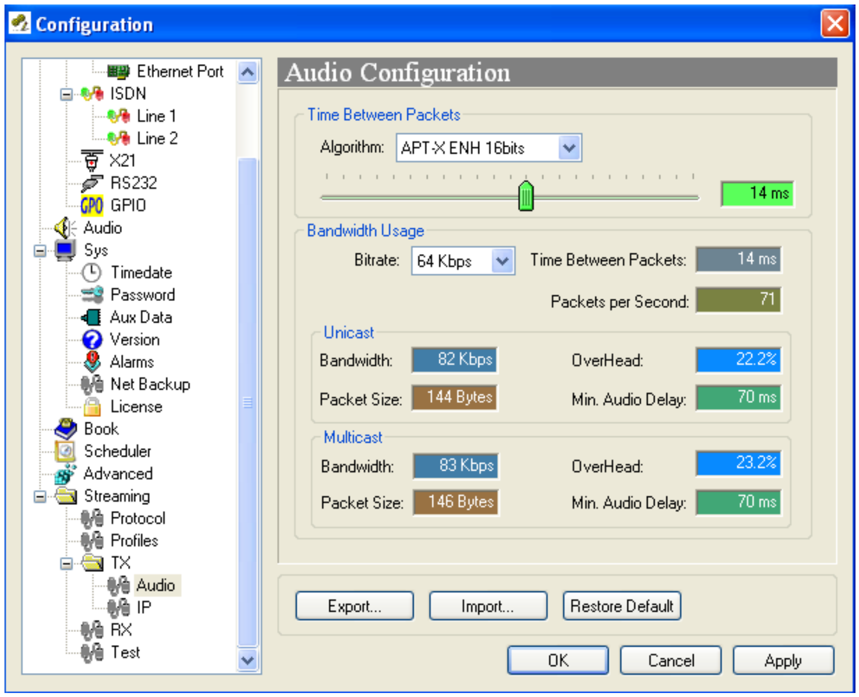

Figure 5: Web information page

And finally, we also provide a **very powerful and easy to use diagnostic tool** that allows **real time** display during the audio connection of critical parameters that are essential for fine tuning of the IP settings. This tool will provide the user with all the information in real time about all the parameters detailed above: jitter, lost packets, disordered packets and buffer occupation. The last chapter explains how this works with some screenshots.

2.- TYPES OF IP CONNECTIONS

There are two different technologies used to establish audio connections over IP: Unicast and Multicast connections. For those networks which do not support Multicast technology, including the Internet, CCS/MUSICAM USA have developed a special protocol: Multi-Unicast that is a combination of these two types of connection and is available on all CCS/MUSICAM USA IP audio codecs.

2.1.- UNICAST

Unicast connections are point to point connections, pretty similar to those ones over ISDN links, for example. Unicast connections can be unidirectional or bi-directional. In addition, unidirectional connections can be outgoing or incoming data connections: Unicast Tx or Unicast Rx connections respectively.

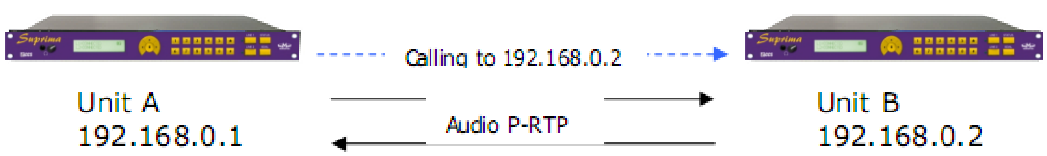

Figure 6: Unicast connection

2.2.- MULTICAST

Multicast connections are point to multi-point unidirectional connections. With MULTICAST the calls must be made from both ends. Both the sender of the data and all the receivers of the data must call to establish a connection to the multicast group. The multicast operation can be seen in the following diagram:

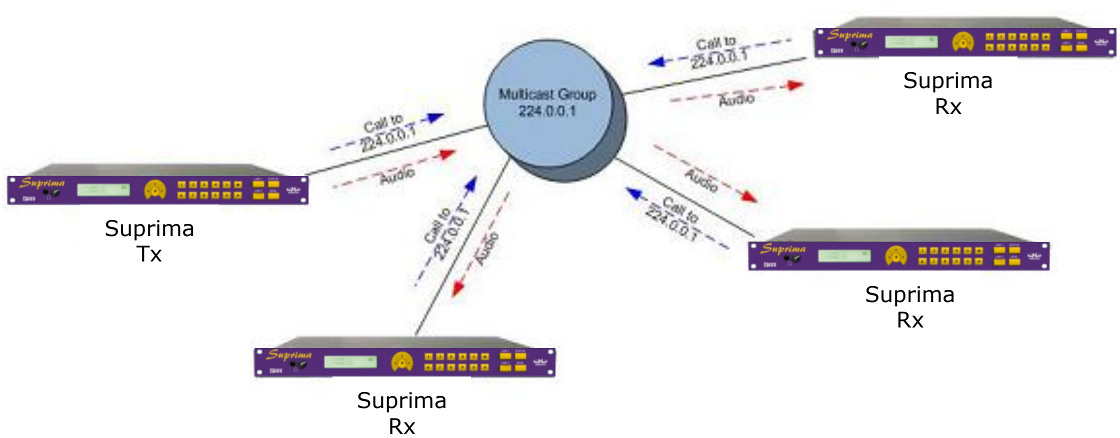

Figure 7: Multicast connection

The big advantage of multicast is that the required bandwidth remains the same regardless of the number of destination, be it 1 or N. This is the most appropriate mode for audio distribution from one source to several destinations at the same time and with the same audio.

The only drawback is that not all IP networks support Multicast. In fact, wide area networks such as the Internet do not support this mode.

2.3.- MULTI-UNICAST

With the inclusion of Multi-Unicast, the user can simulate a Multicast point to multi-point connection over IP by establishing as many unicast connections as the number of receivers. This allows the user to send the same audio to up to 10 different destinations over those IP networks which do not support Multicast.

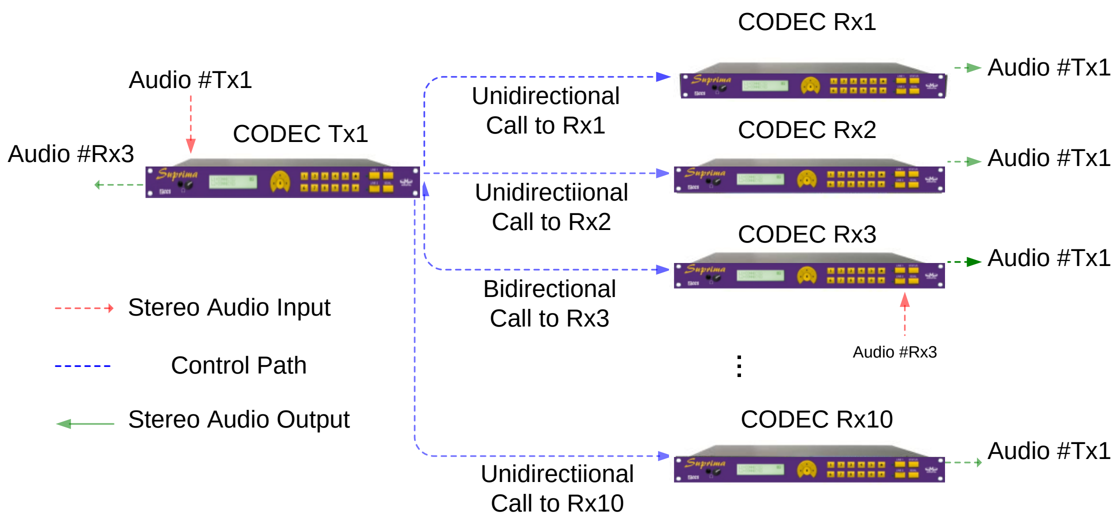

Figure 8. Multi-Unicast

The only drawback of this mode is that the required bandwidth at the transmitter side will multiply by the number of receivers.

3.- PROTOCOLS

The type of connections over IP mentioned above, do only define a means to carry data from one sender to one or multiple receivers, but they do not define the way to establish, control and tear down the connection, as well as the mechanism to negotiate the compression mode for the audio connections and the way to embed other auxiliary data information along with the audio within the audio stream.

CCS/MUSICAM USA has developed hardware and software that supports both proprietary and international standards based protocols (EBU TECH 3326 standard for audio contribution over IP).

3.1.- THE EBU TECH 3326 STANDARD FOR AUDIO CONTRIBUTION OVER IP

The European Broadcast Union (EBU) is promoting the interoperability of audio codecs from different manufacturers. For this purpose, it has been proposed to use and deploy a subset of the existing Internet Protocols: SIP/SAP/RTP and SDP these are the main protocols, although the term ‘SIP’ is commonly used to refer to the whole standard.

This cooperation will allow users to set-up and use audio streaming communications between different vendors equipment in a consistent easy to use way.

CCS/MUSICAM USA is committed to work alongside the EBU NACIP group in the definition and implementation of this new standard for audio over IP. From the beginning CCS/MUSICAM USA have kept at the forefront of these developments and ensuring that all latest protocols stated as 'mandatory' in the standard specification are implemented.

The first version of this standard was released in 2007 as EBU Tech 3326 standard.

3.2.- CCS PROPRIETARY ALGORITHMS

Due to the lack of standards in the early days, CCS started developing its own proprietary protocols to enable IP streaming connections,

- ☑ CCS Real Time Control Protocol (P-RTCP): This is a protocol based on TCP that allows for the establishment and termination of a connection as well as for the negotiation of the codec mode (automatic audio synchronization in all modes).

- ☑ CCS Real Time Protocol (P-RTP): This is a protocol based on UDP used for the transmission of audio.

Even though there is a standard defined by the EBU group in this regard, at the time of writing this document there are still some features which have not been addressed by it, but they are covered by our CCS Proprietary algorithms, such as:

- Asymmetric audio compression: With CCS Proprietary Protocols, it is possible to encode and decode in different compression modes.

- Auxiliary data: SIP does not define a common way of embedding auxiliary data along with the audio data, whilst CCS Proprietary Protocols support auxiliary data over IP for all supported compression modes.

- Back-up procedures: N+1 CCS Redundancy Protocol allows one backup IP codec to provide resiliency for up to N active IP codecs.

  This protocol has been enhanced so that if one of the caller devices fails, the Back-up device takes over and establishes the connection, copying the configuration settings from the failed device.

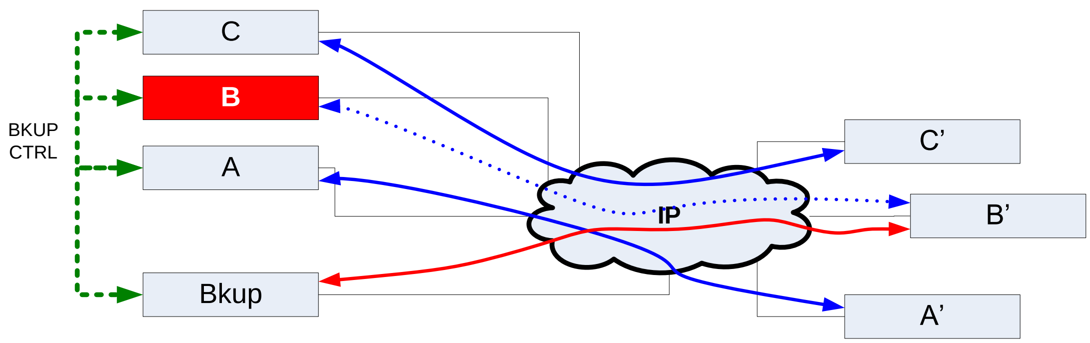

Figure 9. N+1 Redundancy Protocol

The N+1 redundancy protocol runs over IP, and any CCS/MUSICAM USA IP codec can act as a Back-up device for any other CCS/MUSICAM USA codec. This makes it possible to control one unit with another different one (independent of hardware), which does not have to be in the same place necessarily, as long as the corresponding audio signals are available at both sites.

In addition to that, ISDN can be used as a Back-up for the IP & X21 interfaces in case of failure.

None of these procedures, Device Redundancy and Network Redundancy, are covered by the NACIP standard.

- Overhead: Our proprietary protocols add less overhead to the audio connection.

- Control protocol: Using the CCS Proprietary Protocols, it is possible to enable/disable the Control Protocol for the connection. This part is responsible for the establishment, negotiation, monitoring and ending the audio connection. This Control Protocol is similar to the SIP protocol in this regard.

With this control part, it is possible to establish the connection from either end without needing any intervention at the remote end. In addition, as this control protocol continuously checks whether the status of the remote peer is ok, it is possible to know whether the connection is up and ready at any given moment without having to monitor the codec at the remote site.

In some instances, there are some applications where the underlying media for the IP connection is unidirectional so, in this case, a Control Protocol like this would fail, given that it needs a bidirectional link to work. For these scenarios, it is possible to disable the 'Control Protocol' for the audio connection. The drawback is that there won't be any control of the remote end status, and the call will have to be established from both ends.

- NAT Traversal mechanisms: CCS Proprietary Protocol v3 introduces NAT Traversal techniques to allow automatic, configuration free connection through routers. To help the user understand this feature, we will explain this concept in the following sub-chapter.

3.2.1.- Network Address Translation

Any IP device connected to the Internet needs a unique IP address. The available range of IP addresses for the IPv4 standard is limited and it is not large enough to cover all IP devices around the world.

To deal with this limitation, NAT enables multiple hosts on a private network to access the Internet using a single public IP address. Therefore, two different ranges of IP address are defined: internal or private and external or public.

Once an internal host has initiated a connection to an external IP device through the NAT enabled (router), the reverse path on the router is enabled, but any other unsolicited incoming traffic is rejected (firewall). Depending on the type of NAT, this reverse path will be available to any application on any external host, or only to the host and application which was accessed from the internal host. This last type is called Symmetric

NAT, one of the most commonly implemented types in commercial routers.

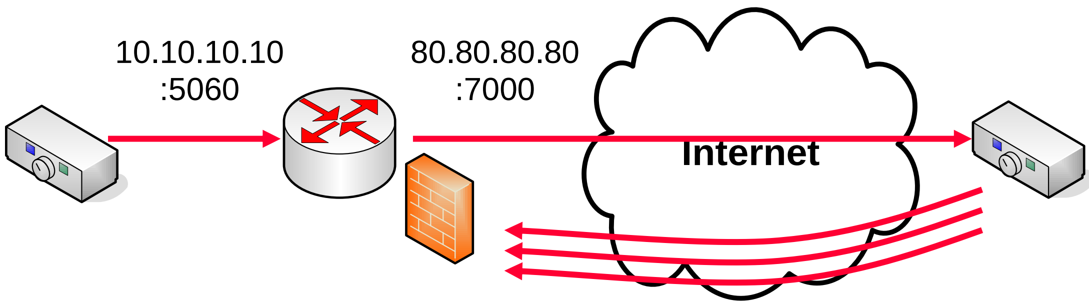

Figure 10. Unsolicited incoming traffic blocked (firewall) by the NAT router.

Some tools to deal with NAT rely on the supposition that this NAT router won't be of the symmetric NAT type, offering an external 'Server' which helps the internal codec connect to the external one by opening a path in the routers at both ends of the connection. STUN is a protocol design to achieve this solution in a more standard way.

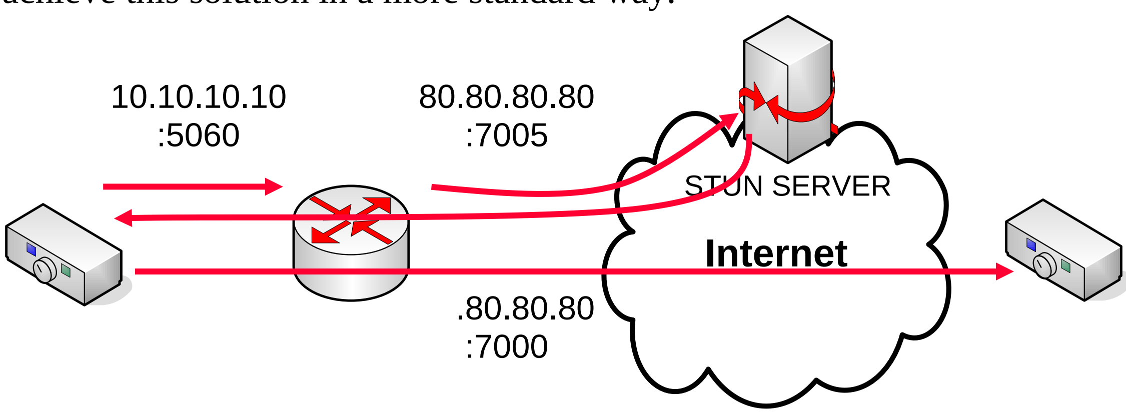

Figure 11. Unsolicited incoming traffic is allowed by the NAT router (not symmetrical NAT) via the reverse path opened with the help of the STUN server.

However, none of these tools work well when the NAT at either end is of symmetric NAT type.

To solve this problem, CCS introduced Proprietary IP Protocol v3, which allows the bidirectional data flow to get through any type of NAT router at the caller side[1], be it symmetric NAT or not, with zero configuration.

4.- CCS PLL ADJUSTING MECHANISM

Due to the lack of a clock to provide a common way of synchronization over the IP networks, a mechanism to compensate for that must be built into the IP codec. Otherwise, the jitter (slightly different clock frequencies) inherent to the AD/DA clocks in the converters at both ends of the connection would lead to inconsistent delay and audio interruptions.

The CCS PLL algorithm controls the DA data rate by fine adjustment of the PLL clock in the receiver, so that when the buffer occupation goes down, this algorithm will slow the PLL down, so that the buffer consumption is decreased and the buffer can recover its original status. The other way around, it will speed up the DA data rate in case of buffer overflow.

This mechanism assures reliable and consistent audio flow even when the underlying transmission media lacks any form of common synchronization clock, as it is the case with all IP networks.

The ‘Streaming bar’ on the web control page of our IP audio codecs informs the user in real time about the operation of the PLL adjusting algorithm. Thus, when this bar goes to the left, it means that the buffer occupation is decreased and the PLL mechanism has slowed down the DA clock, and vice versa.

[1] It is needed to configure the NAT or Port Forwarding capability on the router at the called side, but this side is usually the Studio/Transmitter end, where the router should be accesible.

Figure 12. 'Streaming bar'

5.- JITTER

Jitter is the difference between the fastest proper arrival time for one packet, and the slowest one.

The effect of the jitter in the audio communication is a buffer underflow on the decoder side. This data starvation on the decoder side will cause an audio interruption.

In order to compensate for that, the unit must provide the user with the possibility to set a **jitter buffer (Rx buffer)**. This jitter buffer accounts for the different arrival times of the audio packets, resulting in a steady output stream to feed the audio decoder. The longer the buffer is, the more likely the decoder is to receive data properly. However, increasing the buffer will also increase overall delay.

CCS IP codecs offer the user the possibility to configure a jitter buffer of up to 10 seconds.

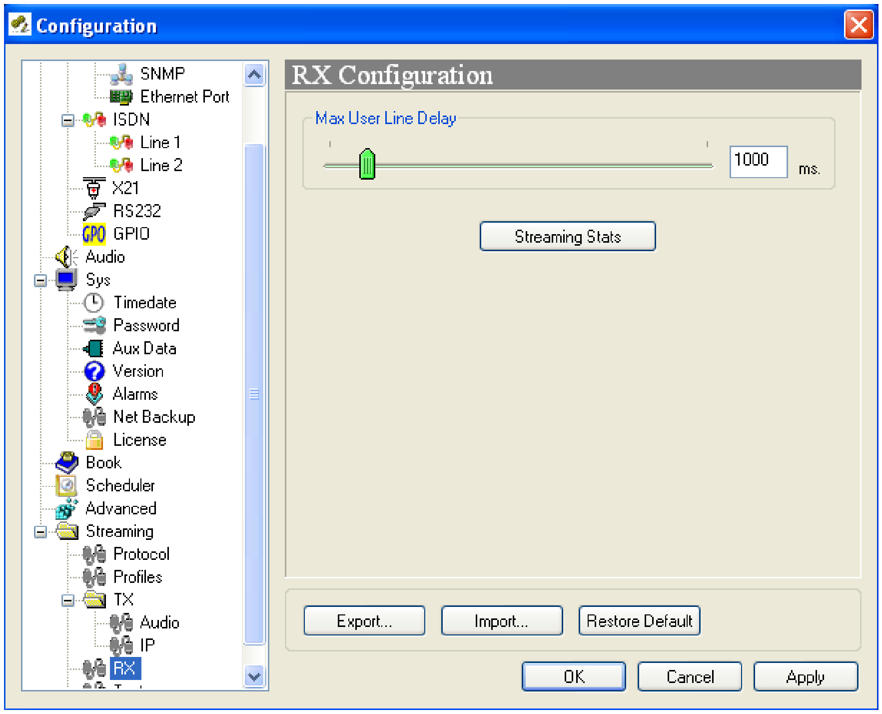

Figure 13. Jitter Configuration Window

6.- FRAME SIZE

As mentioned above, IP links need some extra information (overhead) which is added to the actual audio data, requiring more bandwidth than that of the 'raw' audio bit rate.

Hence, as the header information is of constant size, smaller audio data frames have a larger percentage of added overhead, but the smaller the delay, and vice versa. So, for the end user to achieve an optimum match to their network bandwidth performance and required delay constraints, CCS IP codecs allows the user to configure all low delay compression modes from a very small packet size of 2mS up to a large packet size of 20mS. In addition, a very intuitive, dynamic interactive display shows the user the actual delay and overhead as they change the packet size on the web control page.

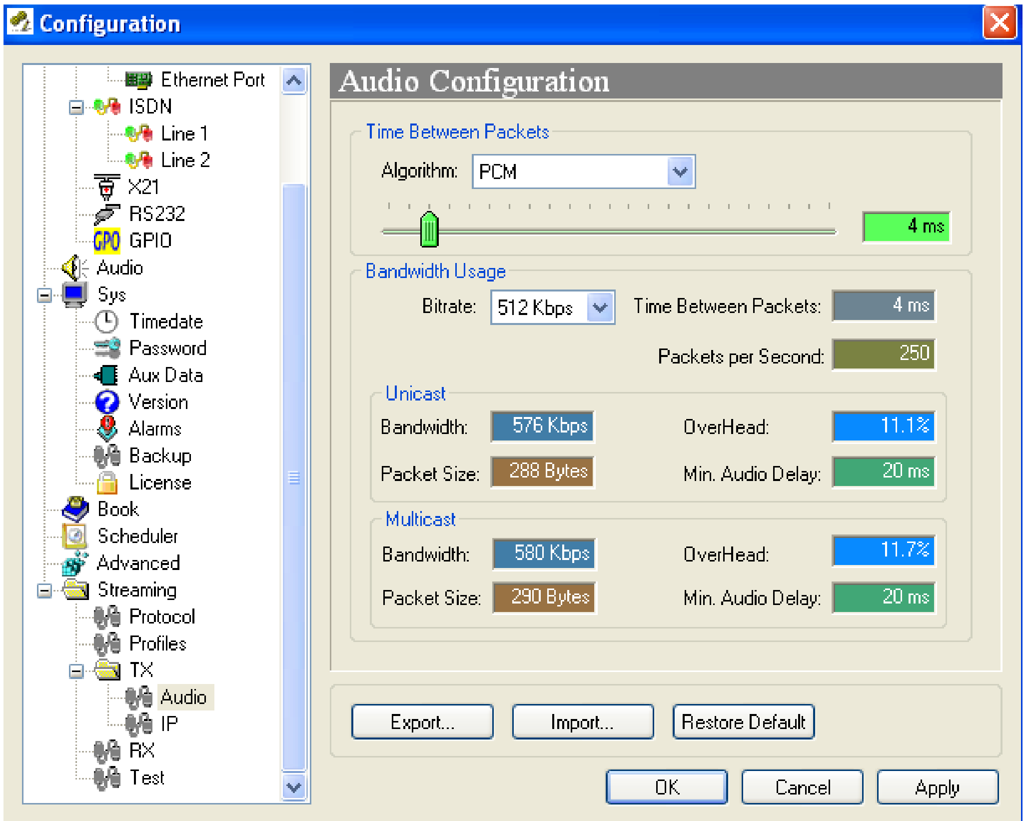

Figure 14. Time Between Packets configuration

7.- LOST PACKETS

The probability of getting lost packets will depend on several factors, the main ones being the quality of the network, the available bandwidth and the frame size used for the audio streaming.

Depending on the circumstances, FEC techniques cannot necessarily totally compensate for loss of data, when this happens the result will be an audio glitch. The audio interruption may be loud or quiet depending on the type of audio, the frame size and type of coding algorithm. In order to make this audio interruption less audible, '**error concealment**' techniques can also be applied.

7.1.- FEC (FORWARD ERROR CORRECTION)

In case of a lost packet, there will be an audio glitch, how audible this is will, to a certain extent depend on the type of audio, the frame size and type of coding algorithm.

Forward Error Correction mechanisms are able to compensate for a small amount of lost packets. FEC mechanisms send special ‘parity’ packets that allow the information (data) to be able to be reconstructed in the case of packet losses based on the redundant information. However these techniques will add more delay and more bandwidth to the audio connection.

CCS IP audio codecs allow the user to set different schemas for FEC.

1. One FEC packet per audio packet (100% overhead).
2. One FEC packet per 2 audio packets (50% overhead).
3. One FEC packet per 3 audio packets (33.3% overhead).
…
24. One FEC packet per 24 audio packets (4.16% overhead).

The second one will increase the overhead by 50%, and will be able to reconstruct isolated lost packets, whilst 100% FEC will double the bandwidth requirements but it will be able to reconstruct up to two consecutive lost audio packets.

It is also possible to select lower overhead schemas, reducing the overhead down to 1 FEC packet per 24 audio packets. This does mean that internal buffers would need to hold at least 24 audio packets, thus increasing delay. Therefore: The longer the audio packet window per FEC packet, the longer the delay, but the smaller the FEC overhead. Consequentially, the longer the audio packet window, the less robust the algorithm, as it can reconstruct less packets (in the simplest scheme, 1 FEC packet per 24 audio packets, a maximum of one packet loss in a cluster of 24 audio packets is supported).

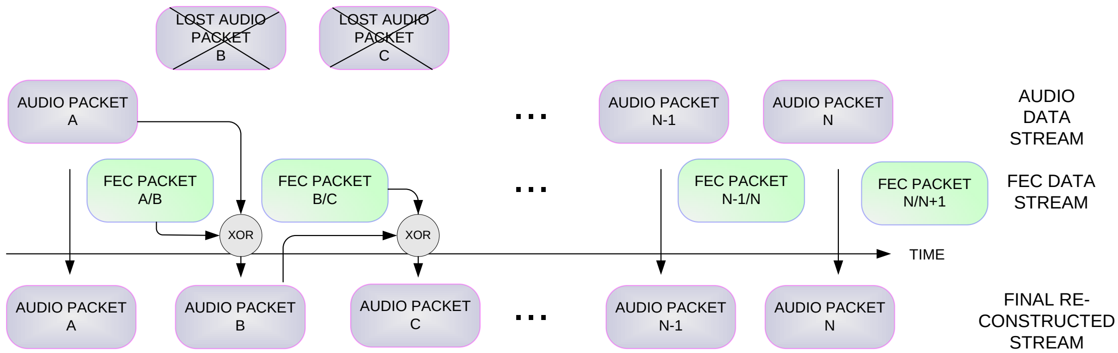

Figure 15: 100% Overhead Scheme: FEC packet per audio packet: it can reconstruct several consecutive lost audio packets

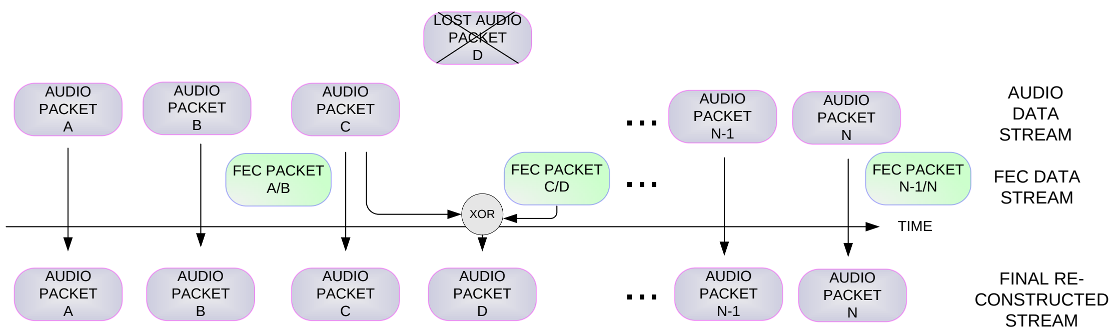

Figure 16: 50% Overhead Scheme: FEC packet per 2 audio packets: it it will be able to reconstruct isolated lost packets

The FEC stream is sent via a separate port to maintain compatibility with not FEC capable codecs.

The CCS FEC mechanism has been developed as per the RFC 2733 standard, keeping its commitment for the interoperability.

7.2.- ERROR CONCEALMENT

There are some scenarios where FEC mechanisms cannot be used due to bandwidth or delay restrictions. For those cases, a method to minimize the impact of packet losses in form of artifacts on the decoded audio might be used.

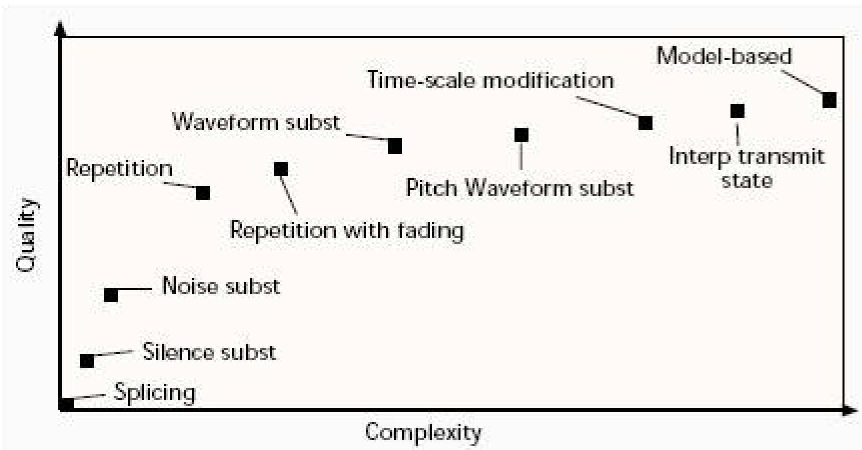

Figure 17. Error Concealment

This figure shows the trade-off between quality and complexity for different 'error concealment' techniques. There are several ways of doing this, and several RFC standards that can be followed.

After carefully analyzing the different techniques, CCS has decided to implement the optimum one in terms of quality vs. complexity: Repetition with fading.

These techniques are part of the MPEG audio standards, and the most effective way to conceal audible artifacts in the audio reproduction caused by lost information.

8.- DISORDERED PACKETS

Given that the data flow is split up in to packets, and each packet in a stream might be routed in a different way from the rest, the path from the sender to the receiver might vary depending on several different factors. This might lead to packet arriving out of sequence. To deal with it, CCS IP Codecs support packet reordering if a receiver (Rx jitter) buffer has been configured for the audio connection. This feature is supported on both Proprietary (CCS v4) and SIP/RTP protocols.

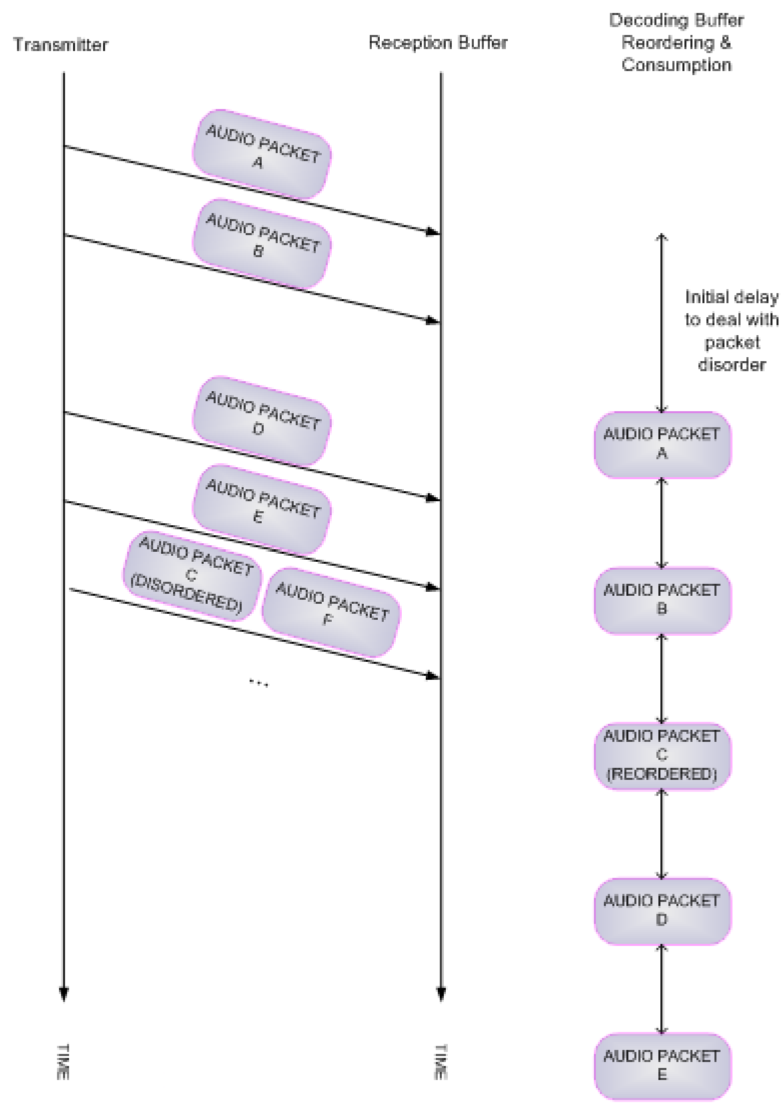

Figure 18. Packet Reordering.

9.- ACTIVE RECOVERY

Active Recovery is based on Retransmission techniques. Unlike FEC (passive recovery), no redundant information is added to the original stream to help the receiver reconstruct the data flow in case of packet losses.

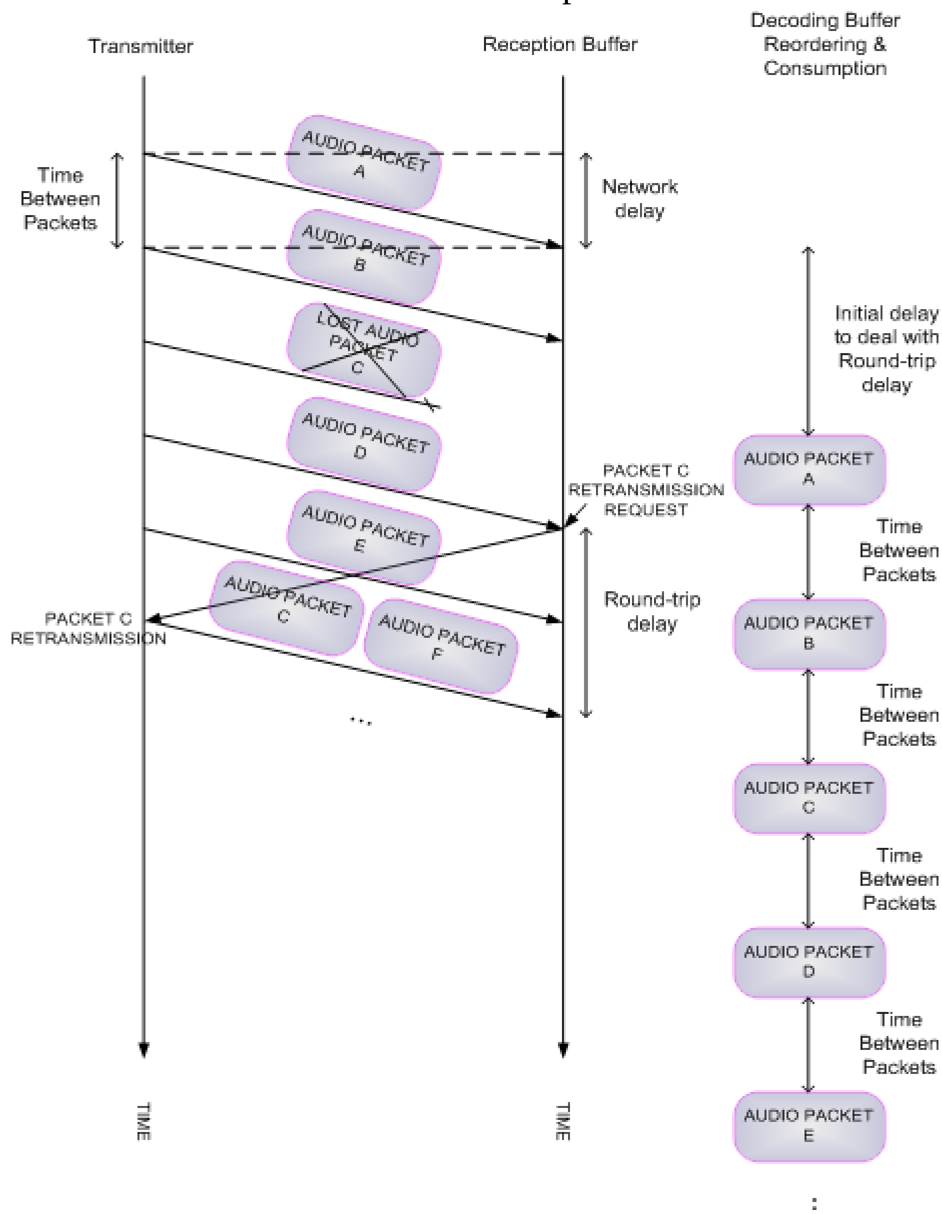

Figure 19. Active Recovery Diagram for a zero jitter network.

Active Recovery does only send additional information when it is actually needed, thus keeping additional network overhead to a minimum. As can be seen in figure 18, Audio Packet C is the only lost packet, so this is the only redundant information being sent from the transmitter.

The disadvantage of this technique is longer end-to-end latency, due to the round-trip delay of the IP link. As can be seen in figure 18, the decoder must wait for an initial period of time (buffering) before starting to process data, to deal with the additional delay caused by the round-trip delay of a packet retransmission in case of a packet loss. Otherwise, there would be an audio interruption when Audio Packet C is lost in figure 18. If this additional delay caused by retransmission were interpreted as jitter in the connection, 'active recovery' buffering would be the jitter buffer.

Take into account that jitter is a different concept, nothing no to do with active recovery techniques and retransmission, and that the jitter buffer would have to be added to the 'active recovery' buffer in case of jitter in the link.

This method is reasonably appropriate for unidirectional audio communications where latency is not an issue.

10.- MEASUREMENT TOOLS FROM CCS

One of the most innovative features on CCS/MUSICAM USA IP family of codecs compared to other manufacturers is the monitoring tools for audio streaming and network diagnostics. The latest of these tools is called the ‘Real time Network Analyzer’ which runs concurrently and in real time with the audio connection.

This tool offers information about the performance of the audio streaming in real time, and it stores the information for 24 hours or up to 7 days if the web control page remains opened for that time.

All the available information can be exported to a file, and it is organized by call, so it’s possible to recover the data from a specific call and any given time within a call. This tool works on the receiver side (decoder side) and it provides the user with the following information:

1. Rx Buffer occupation
2. Jitter
3. Lost packets & Disordered packets

This information shows the quality of the audio streaming operation over the IP network, as well as helping the user fit their network implementation.

The statistics will highlight periods of time when the decoder is not able to reassemble the output audio due lack of the network performance, corresponding to the ‘Framed lost’ and ‘Decoder not framed’ alarms.

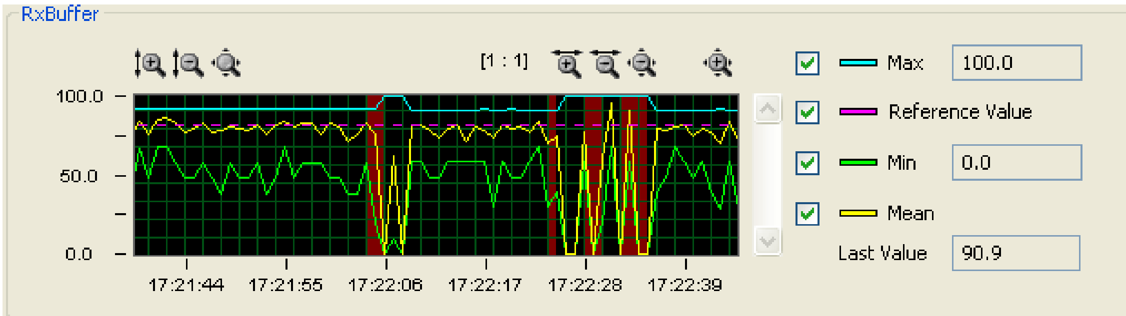

Figure 20. Rx Buffer occupation with decoder alarm notifications

In addition, it runs as part of the web control page, so it is not necessary to install any additional software.

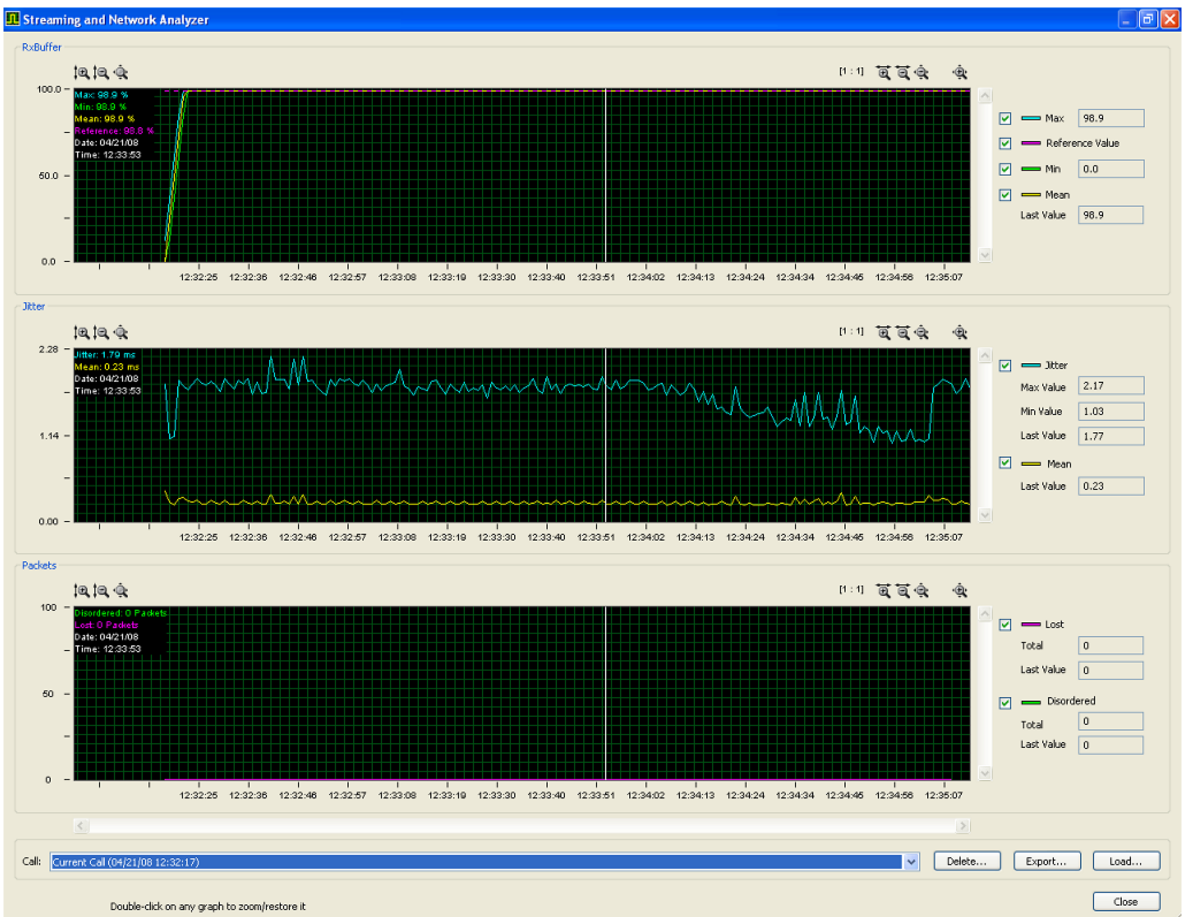

Figure 21. Real Time Analyzer main window

Should the user want to check the available bandwidth, the delay and/or the jitter before establishing a connection, CCS/MUSICAM USA offers their customers another tool, which runs off-line, and allows the user to measure the upload and download bandwidth in the IP link between two CCS/MUSICAM USA units. As well as the round-trip delay, it is also possible to simulate a real audio connection in terms of bandwidth and packet rate to check whether the IP link supports the necessary audio bandwidth requirements. If this option is NOT selected, the tool will try to use as much bandwidth as possible, measuring the total bandwidth available for that link.

Warning - this tool will take **ALL** of the available bandwidth, so should be used with care on an in-service network.

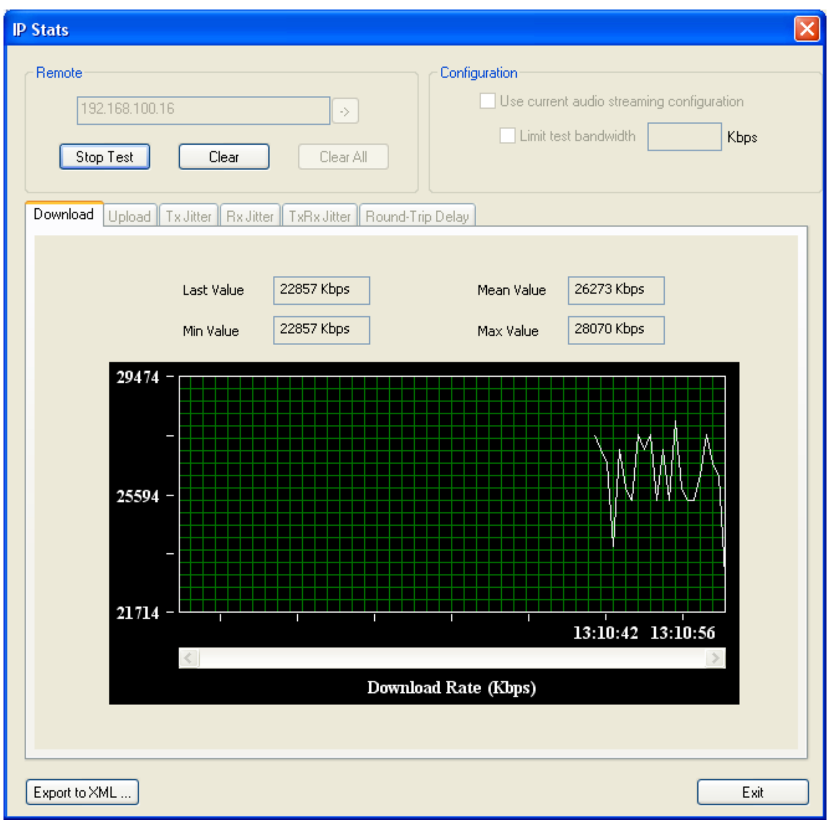

Figure 22. Test Streaming Tool

11.- MONITORING TOOLS: SNMP & ALARM SYSTEM

CCS/MUSICAM USA IP devices support SNMP protocol for control/notification of different configuration settings and events. In addition, the alarm system allows the user to select which type of events the unit will check and notify. Notifications via SNMP traps and/or emails are supported.

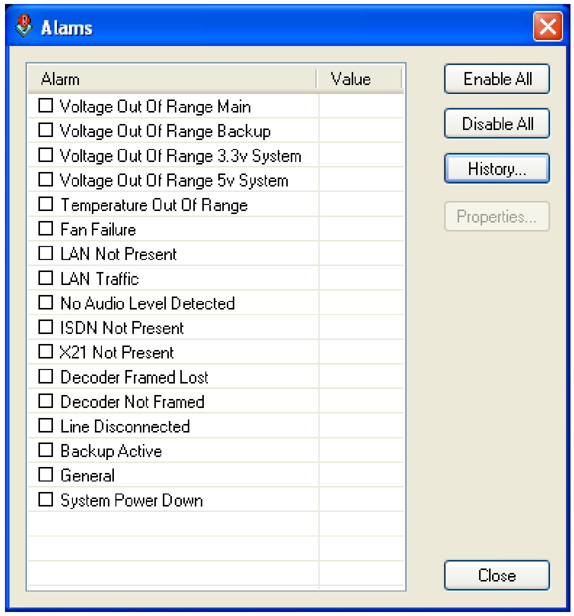

Figure 23. Alarm window

Alarms are saved in the non-volatile memory of the codec, so it is possible to access the alarm information even if the unit has been restarted.

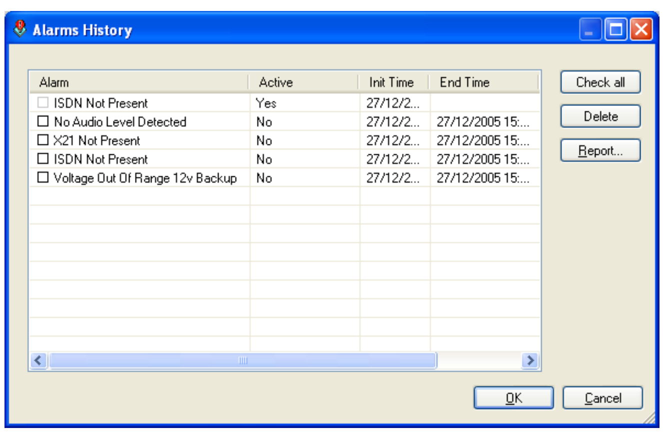

Figure 24. Alarm history window

12.–DOUBLE LAN INTERFACE

On some IP network installations it is desirable to separate the audio network from the control or management network. This may be to prevent other IP devices, such as desktop computers, from putting extra traffic on the audio network, or to prevent external devices and computers accessing a corporate LAN.

If in this case separation cannot be achieved by means of an external IP device (router) for example, CCS/MUSICAM USA already offers the option to include a second Ethernet/IP interface on their units.

This second Ethernet/IP interface has its own IP address, acting as a completely independent interface for control/management purposes.

The main Ethernet interface acts only as the audio interface when the second interface is enabled in the unit and, at this time, the main Ethernet interface will not accept any management.