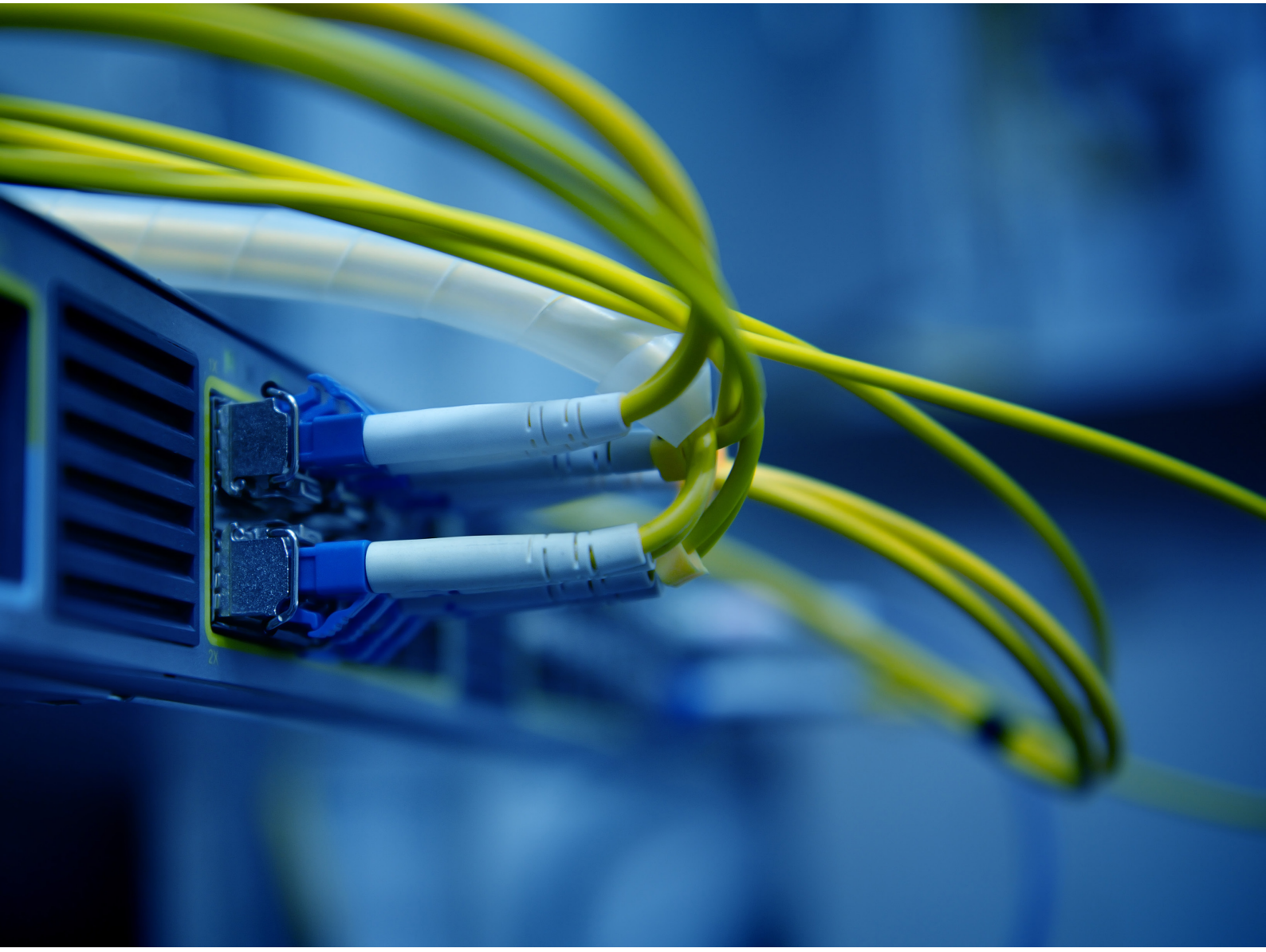
CALREC

IP Resilience

How to build extra resilience into your IP broadcast networks

a **Calrec** whitepaper



IP Resilience

How to build extra resilience into your IP broadcast networks

When it comes to building resilience into an IP environment there are many areas to consider.

Some of these areas aren't new; things like ensuring devices are powered from two independent power sources, or ensuring redundant cable runs are physically routed separately. These physical problems will never go away.

Some challenges *are* new, such as deploying a second network fabric to accommodate a SMPTE ST2022-7 pair of media streams or ensuring that IGMP querier election is redundant and deterministic.

In a free, certified IP training course available at **calrec.com**, we have already discussed Network Design best practices, multiple network topologies and how network design decisions can drastically change the resilience of an IP network.

You can find more information on this here: https://calrec.com/courses/aoip-network-design/

This paper looks at how to build safe, resilient and robust IP network infrastructures and why Calrec is doubling down to providing fully hardware redundant IP solutions.

It is beneficial to networking, technology and architect teams who require a thorough understanding of the operation and performance of Calrec's hardware redundancy model.

CALREC

Contents

Resilience and redundancy are similar terms that are used a lot when describing IP networks that are secured against component failure. Resilience is generally achieved by ensuring adequate redundancy. There is some ambiguity with the term redundancy as it has different meanings across different sectors and technologies.

In a broadcast environment, redundancy is the presence of a secondary technology installation that either works in line with the primary or is available to take over in order to recover a failed primary.

In an IT environment, the term redundancy may refer to underutilised or unused technology that may result in unnecessary operational costs to a business. For the sake of this whitepaper, we refer to redundancy in a broadcast context where redundancy is a provider of resilience.

In live-to-air facilities, a fundamental design principle of highly resilient architectures is the elimination of single points of failure. Let's look at a typical IP deployment; where are the weaknesses?

The nomenclature used to identify each distinct network varies between the IT and broadcast industries. Red/Blue, A/B, Amber/Blue, Primary/Secondary are all regularly used. In this whitepaper we will use Amber/Blue in order to correlate to TR1001-1 published by the JT-NM.

Neither network has priority over the other and both should be configured the same as each other to provide equal performance. Usually there are two networks when ST2022-7 is deployed and perhaps two PTP Grand Master Clocks per network. A separate management network or networks can be installed to mitigate risk to media traffic and allow access to off-site servers via private WAN or the internet. Essential applications may be deployed in off-premises public or private clouds. On-premises servers are often clustered or doubled up, sometimes behind load balancers. Baseband audio I/O can be split across multiple IP devices. Backup mixing consoles can be deployed to ensure whatever happens, the show will go on. Power is redundant, transmission paths are redundant, there is at least two of everything wherever possible.

Figure 1 shows what a resilient architecture might look like.

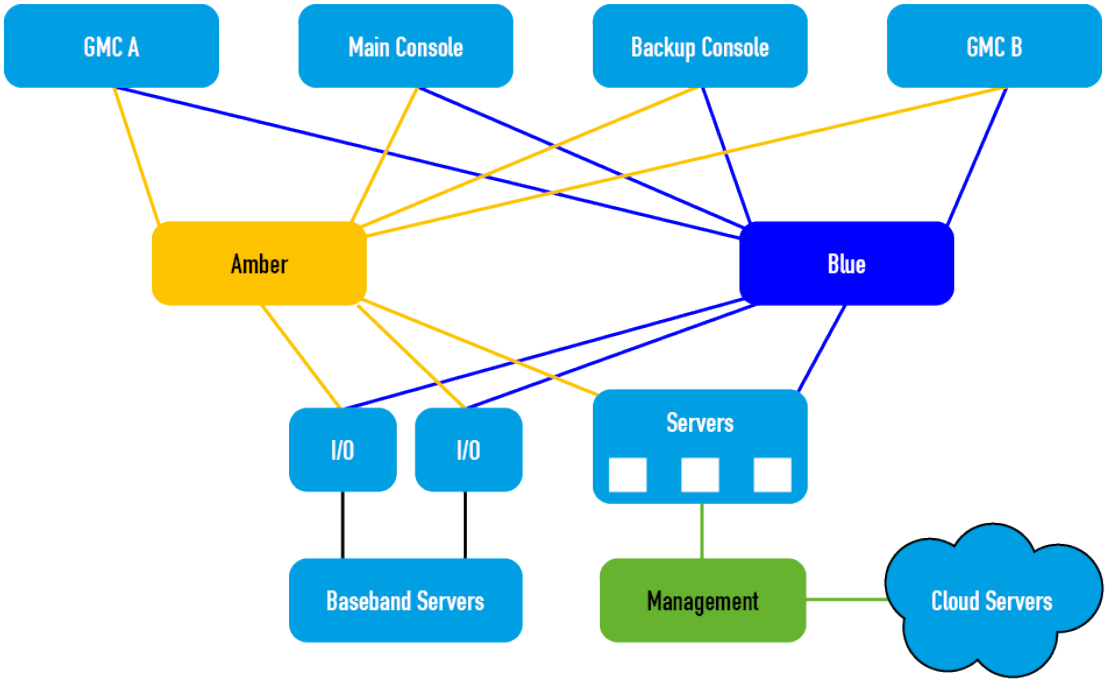


Figure 1: Resilient network architecture

This architecture protects against the majority or failures that could be expected when working in an IP environment; component failures such as PSUs or damaged cabling or accidental unplugging of devices. With this level of resilience, you can expect hitless audio performance. ST2022-7 (which we look at in the next section) ensures that there are always two copies of media on a network at any given time.

But what happens if there are multiple points of failure? We all hope that this never arises , but history has shown us that even the most farfetched disaster scenarios are entirely plausible! A solid disaster recovery strategy is the best insurance against the risk of loss of live content.

This is why Calrec has developed hardware redundancy technologies that allows swift and controlled recovery of mission critical audio mixing functions if the unthinkable was to happen.

Let's take the hypothetical disaster of a fire within a data centre or machine room. The DSP core of the main mixing console has been rendered out of action. In reality, there would be many further safeguards against this, but for the sake of discussion let's assume they have also failed to protect the live equipment.

The facility deploys a backup console, but this now needs manual and procedural effort to be flipped over to. This invariably leads to downtime as the backup system is commissioned for live use.

Protecting your content to air in this scenario may be impossible without significant interruption. Establishing an expected and deterministic time to recovery can be difficult whenever manual intervention is required, so deploying technology that automates media recovery times is essential.

The rest of this whitepaper will be dedicated to describing how Calrec mitigates risk by offering redundant hardware that automatically and deterministically keeps the show on air, even in the event of complete service outage at a geographical location.

SMPTE ST2022-7

The first area of resilience with all Calrec AoIP products is the implementation of SMPTE ST2022-7.

ST2022-7 mandates that a device should produce more than one single media stream for any given media source.

All Calrec Cores and IO Boxes provide a set of two NICS with the expectation that each NIC is connected to a different independent network. Streams with identical content are played out onto each network.

Although the content of the RTP Payload is identical, the transport addresses and ports are encouraged to be different, especially when deployed in a routed multicast environment.

ST2022-7 is an Active/Active redundancy model as IP devices are producing and consuming IP traffic on both networks at the same time. If a packet is missing from either side of the network, the receiver will simply use the packet from the other side.

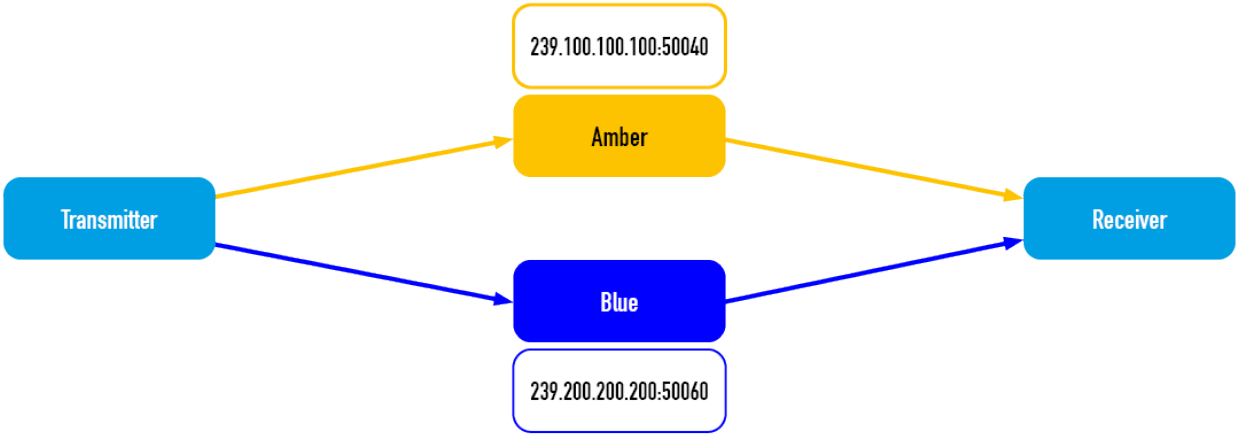


Figure 2: ST2022-7 hitless packet merging

This mechanism, illustrated in *figure 2*, is known as hitless packet merging and ensures clean audio reproduction when packets from either side of the network are missing.

Core Redundancy

Calrec builds on ST2022-7 by adding a second Calrec IP core which can be geographically separate from its twin. Pairs of Calrec's IP cores work in an active/inactive model; this means the inactive core is not responsible for processing audio unless a failure is detected on the active core.

In normal working conditions, the primary core is active and the secondary core is inactive; only one core can be active at a time. In the event of a failure of the primary core, the secondary core will be marked as active and the primary core will be marked as offline. The secondary core is a full duplicate of the primary core offering full hardware redundancy to all components, including the ST2022-7 pairs and dual PSUs. As both cores deploy ST2022-7 NICs, both cores are meshed to the Amber and Blue Networks (*figure 3*).

Both IP cores can be connected by multiple IP links known as twin links. Data is replicated from the active core to the inactive core over the twin links. Each core has its own unique IP addresses on all interfaces, so all components across both cores can be accessed independently over a network.

There are no restrictions on what IP addresses can be allocated, only that all applications and NICs require unique IP addresses, so you can design Layer3 resilience into all areas of both cores giving critical isolation between Layer2 segments.

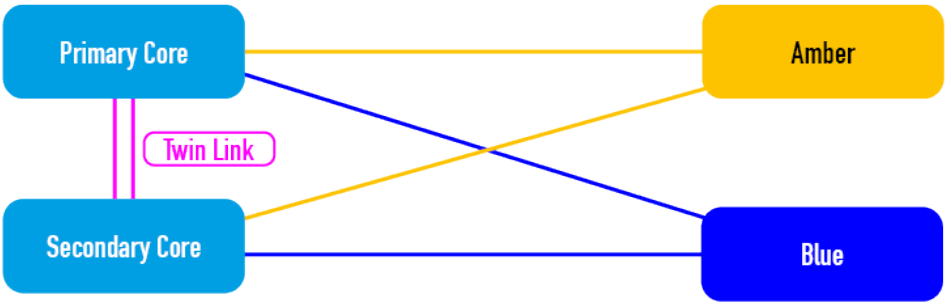


Figure 3: Calrec hardware redundancy media connectivity

With hardware redundancy in mind, there are additional challenges that require innovative solutions. Our goals when developing our hardware redundancy model are:

- Ensure automatic and deterministic failover
- Minimal downtime
- No strict network design requirements or prerequisites other than existing best practices
- Agnostic to any COTS network hardware or software
- No compromise on interoperability to or from other devices on the network
- Ensure hardware failover is as invisible as possible to other devices on the network
- Encourage redundant hardware to be geographically separate

The following sections discuss what Calrec's hardware redundancy model looks like from a network perspective in various operating and failure conditions.

Calrec redundant cores as a AES67/ST2110-30 sender

Calrec only produces streams from the active core (ie. the primary core in normal working conditions). Packets are not played from the inactive core. The reason is simple: if both cores generated the same multicast streams, then any receiving device on the network would receive duplicate packets.

This would double the inbound bandwidth consumption for a receiver's link. We would also be unable to account for a receiver's ability to throw away any unnecessary packets; at the time of writing, there is no standardisation on how an end device should handle a duplicate packet received on a single NIC. This makes it difficult to adopt an active/active redundancy model in an interoperable AES67/ST2110 environment. This is illustrated in *figure 4*.

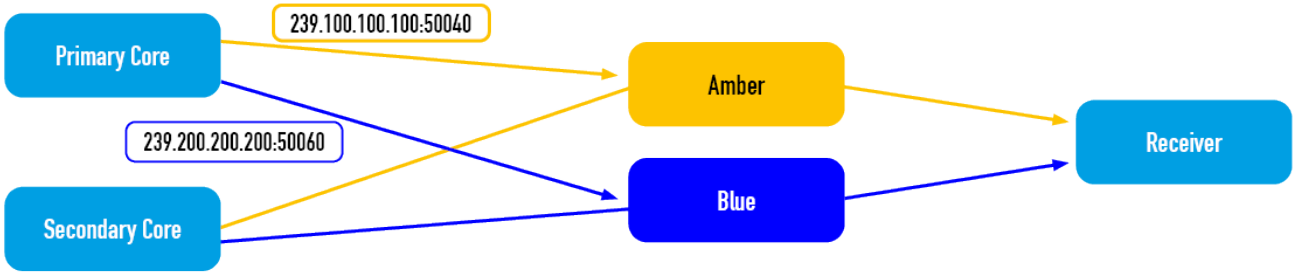


Figure 4: Calrec redundant cores only transmit from the active core

Although packets are not produced from the secondary core, all the data required to activate the stream is configured and ready to go. The active core provides the inactive core with all of the instructions it needs to be ready to send packets if it is promoted. If streams are altered by the user in connect, or via an NMOS controller, these changes are replicated to the inactive core.

In the event of a hardware failover (from primary to secondary), the following actions occur (*figure 5*):

1. All senders on the failed primary core are muted and packet production stops.
2. All senders are activated on the freshly promoted secondary core.
3. The IP packets that are now being produced by the secondary core use the same transport IP addresses and UDP ports. This ensures that any receivers that were tuned to the primary core's streams now receive the secondary core's streams.
4. The source IP address of the stream is changed to the IP address of the NIC it is produced from on the secondary core. This adds some complications in the event that the network and/or the receiving device are implementing source specific multicast (SSM). We will discuss these challenges later.

Calrec redundant cores as a AES67/ST2110-30 receiver

The inactive core receivers operate differently. An inactive core is just an audio device that receives streams from the network, exactly the same as any other IP device on the network.

Therefore, the inactive core can send IGMP Membership Reports for the same multicast addresses as the active core and both can receive the same streams in parallel and the network takes care of packet replication and delivery to the end devices.

As the inactive core is not responsible for processing audio, the packets are simply dropped on arrival by the

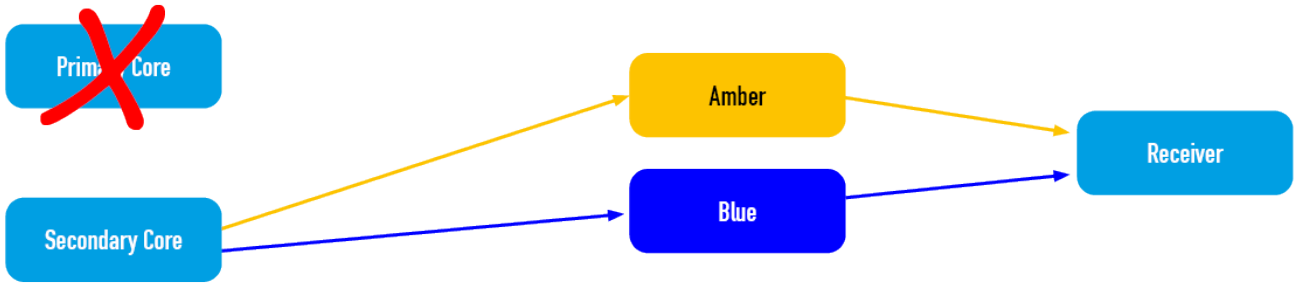


Figure 5: Redundant cores activate streams when they are promoted to active

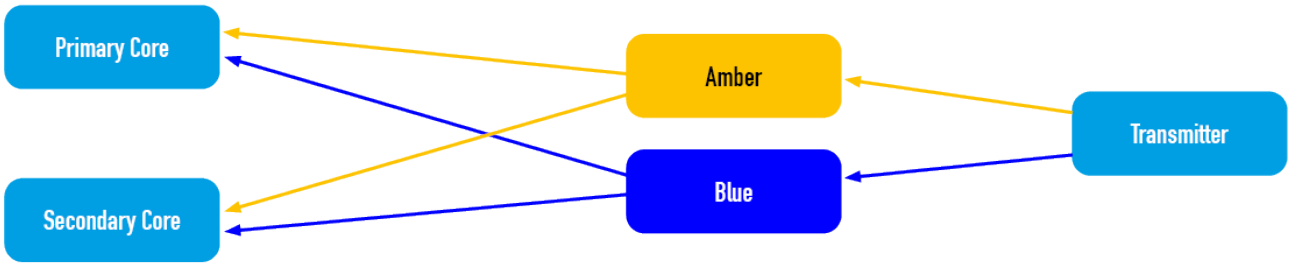


Figure 6: Redundant hardware as a ST2110 receivers

inactive core's routing logic.

In the event of a failover, the network IGMP tables on both amber and blue networks are converged and ready to go.

From a multicast point of view, the network does not need to react to a core failure; it is already directing multicast to the inactive core's receivers (*figure 6*).

Redundant core control

There are a number of methods, applications and protocols that can control Calrec cores. Keeping control of all of these entities in the event of a hardware failover is essential. The following sections will discuss the different ways in which the control of the active and inactive cores remain deterministic in times of crisis.

Calrec offers engineering and operational functionality through multiple software applications on cores:

- Connect – A stream manager which provides a clear overview of IP devices and streams on a network and gives the user a simple way to manage them
- Assist* – A soft control surface that can generate offline console snapshots and memories as well as provide live mixing services from a UI on a laptop or computer.
- Configure – A comprehensive configuration and diagnostics gathering application.
- Software Updater – An application which updates the software of all system components.

Each application is designed with resilience in mind. Due to subtle differences in workflows, the failover mechanism isn't consistent across all the applications; the specifics for each application are discussed in the following sections.

Calrec runs a proprietary protocol called Calrec Control Protocol (CCP) that is used to discover and control I/O devices on the network.

CCP was designed to handle hardware redundancy from the ground up and the background processes are invisible to the user.

Calrec IP cores and I/O boxes also have full support for NMOS IS-04 discovery and IS-05 connection APIs. NMOS does not specify a direction or operation of the APIs for a hardware redundancy model like Calrec's. Information on how Calrec includes full NMOS redundancy is presented in the following sections.

Control application redundancy

The active core will always replicate data to the inactive core. Therefore, some applications will have limited connectivity when they are inactive; this ensures that new control data from a client isn't overwritten by older data from the active core.

The behaviour of different applications that run on a core is determined partially by their requirements to replicate data between cores or not.

Data is typically replicated every five seconds, so you can expect automatic point-in-time recovery of five seconds or less in the event of a hardware failover.

Calrec Connect

Calrec Connect is a standalone server that doesn't have to run on Calrec hardware. This platform agnostic strategy allows flexible deployment and means it can be accommodated into any disaster recovery strategy.

When Connect is running on Calrec hardware, it is only launched on the active core. This prevents client control of the inactive core and ensures a single point of entry for all control data, which avoids duplicate data due to intra-core data replication.

In the event that an inactive core is promoted to be active, then instances of the application are instantly spun up with the latest data that has been replicated. Connect presents data on network devices and will always present a live view of the network. Therefore some data is not subject to the five second point of recovery, and will always present live data.

The passing of audio between devices has no dependencies on Calrec Connect. The availability of Calrec Connect on the network has no bearing on audio performance. Calrec Connect is a control application only.

Calrec Connect Virtual Machine

Calrec also provides a Calrec Connect virtual machine (VM). The virtual machine is provided as a .ova and can be deployed on most popular type 1 and type 2 hypervisors.

To clarify, a hypervisor is a software package that can manage virtual machines, including functions like allocating hardware resources of the host to a guest virtual machine. A .ova is a package that contains a virtual machine in an open and standardised format which allows a .ova to be spun up on any hypervisor that supports the format.

This allows Calrec customers to deploy the application in a flexible and elastic way to meet the needs of their infrastructure. This also opens opportunities for programmatic deployment of the application based on any number of operational circumstances. The Connect VM also allows 'offline' configuration of I/O devices before they are placed into an IP production environment.

Calrec Assist

Calrec Assist* will run on both cores but is only operable from the active core. In the event of a failure, any active sessions will be redirected to the newly promoted active core. If there are no sessions running at the time of a failover, then the client would need to simply log into the newly promoted active core using the application's unique IP address.

Calrec Configure

Calrec configure is used for management and diagnostics. Both cores can be configured and managed for each core's application. For total redundancy, it is recommended that both applications should be made available to the network, but this isn't a strict requirement.

Calrec Software Updater

Software updater is used to load software updates onto the core. Each core has its own application so that each core can have its software updated independently.

How software updater is deployed on an IP network should be considered carefully. Both cores should have a interface that software updater is exposed to, but that doesn't have to be permanent. Software updater is only intended for use during an adequate maintenance window, and not making it available to the network permanently would be good security best practice. The interface that Software Updater is exposed to can be managed from Calrec Configure.

Control application addressing

Each application is addressed uniquely to offer full deployment flexibility. Applications can be in the same subnet or different subnets; they can be attached to the same physical network or separate physical networks. All applications can be freely assigned to a different physical interface on the primary and/or secondary cores.

The way control applications are presented to the network is fully definable using Calrec Configure. It is a good idea when configuring control PCs that there are shortcuts or quick links to both the primary and secondary applications.

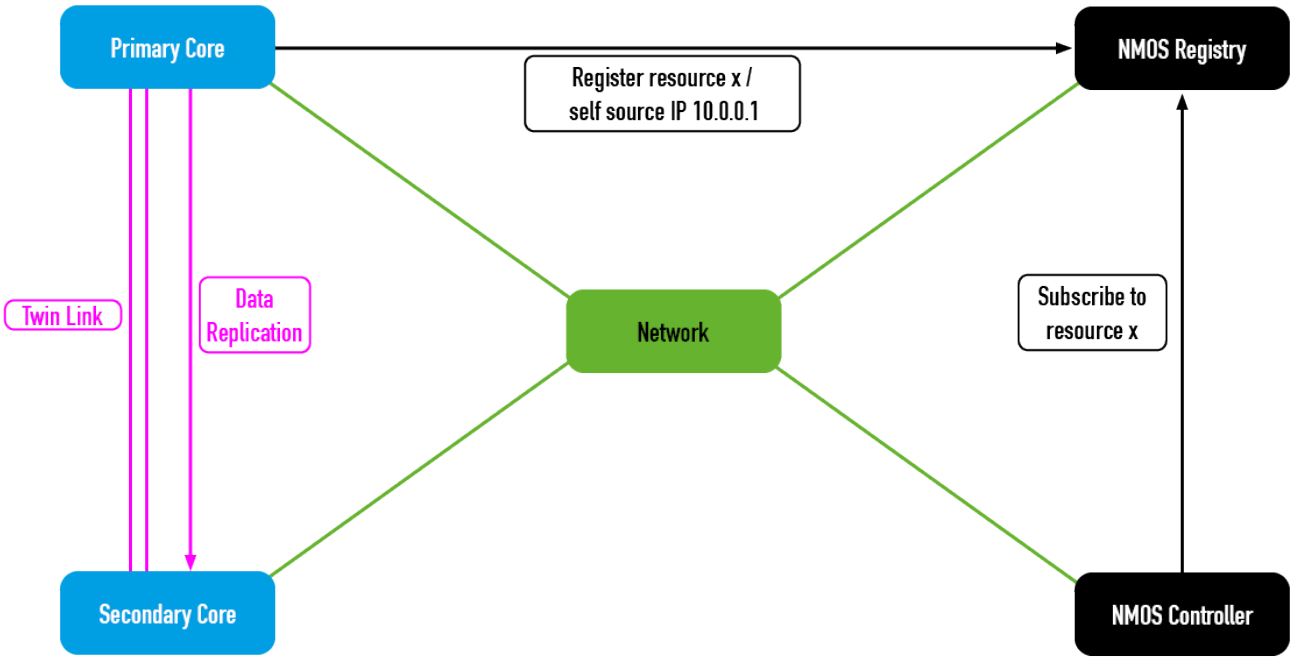


Figure 7: IS04 connectivity under normal working conditions

Calrec Control Protocol (CCP)

CCP is the protocol that Calrec AoIP devices use to talk to and control each other. We developed CCP in order to provide services to our users that are not currently covered by industry standards. As well as hardware redundancy, services such as port protection, software updates and sophisticated port control are important features that we were not willing to compromise on.

No special considerations need to be given to CPP with regards to hardware redundancy. When a hardware failover occurs, CCP devices communicate with each other to update all of the necessary information about active senders and the addresses and ports to use for control.

The networking requirements are identical to that of media streams. There must be at least a single multicast and unicast path between devices.

On large AoIP networks with many CCP devices, you can expect bandwidth consumption to be no greater than 20Mbps peak. Realistically, the bandwidth consumed by CCP is < 1Mbps, but may peak higher during failure conditions or if devices are being added or removed from the network.

NMOS

Calrec is AWMA NMOS IS-04 and IS-05 compliant across redundant cores. NMOS IS-04 is an API for device discovery across an IP network. NMOS IS-04 nodes register themselves and their resources to a registration server. NMOS devices such as other NMOS nodes or NMOS Controllers can communicate with the registry to get connection and stream information for any NMOS node on the network. NMOS IS-05 is a connection API used to allow NMOS controllers to control media flows between NMOS nodes. For more information on NMOS, visit https://www.amwa.tv/nmos.

The NMOS standards do not directly reference operation of the APIs for any specific hardware redundancy model. Calrec have taken an innovative approach to implementing NMOS on physically separate but logically identical hardware devices. The following sections describe the behaviour of NMOS on our redundant core model under normal and failure conditions.

Calrec redundant hardware as an IS-04 Node

The IS-04 Node API only run on the active core. The active core

registers its resources with the NMOS registry. NMOS controllers can query and control the node on the active core through the normal NMOS mechanisms.

Any data that is changed on the active core's node API is replicated to the inactive core. The inactive does not run the node API and does not present the replicated data to the network (*figure 7*).

In the event of an active core failure, the following happens (*figure 8*):

1. The Node API is stopped on the failed core.

2. The Node API is then spun up on the newly promoted active core.

3. The Node API re-registers with the NMOS registry.

   a. The Node API on the newly promoted core inherits all the resource UUIDs from the primary core.

   b. The Node API on the newly promoted core tells the registry about its IP and MAC addresses so devices know that the location of the resources has physically changed. This includes the endpoints and href in /self and the connection API hrefs in /devices.

One of the key benefits to inheriting the resource UUIDs on a failover is that an NMOS device or controller that was previously subscribed to any of the resources on the failed core remain active, but the registry now automatically directs queries to the newly promoted core.

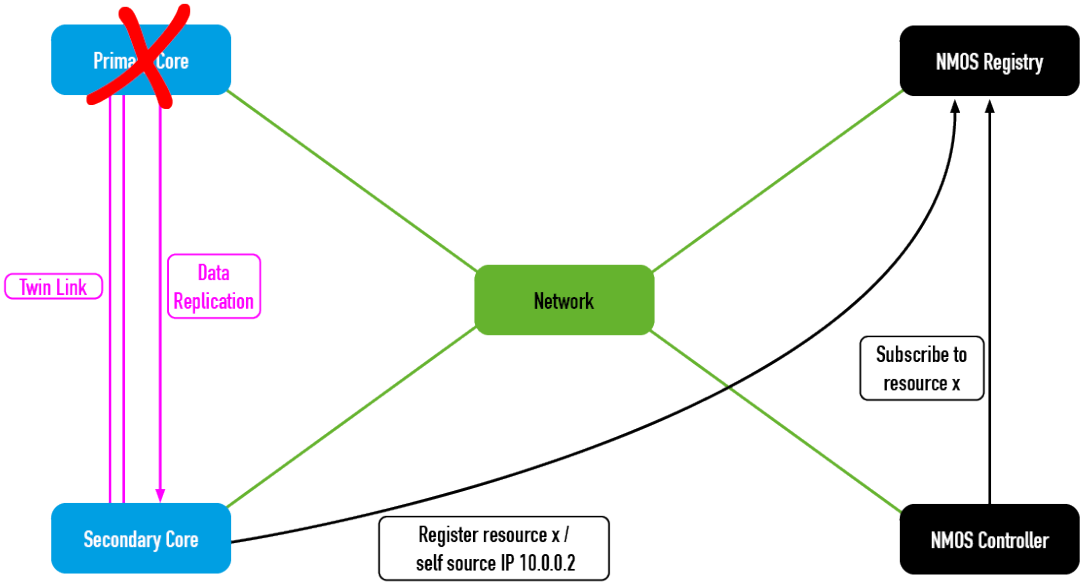
NMOS registries commonly use a 12 second default timeout value. When a core failover occurs, the NMOS APIs are instantly spun up to the newly promoted core. This means that often, the NMOS registry is unaware that the device has gone missing; it simply has its IP addresses and MAC addresses updated to reference the new active core.

Designing NMOS failover in such a way means that the failover is invisible to NMOS controllers and no action is required in order to continue controlling the Calrec core nodes. This helps work towards Calrec's goal of making hardware failovers invisible to other devices on the network.

Controlling Calrec redundant hardware via IS-05

The IS-05 connection API only runs on the active core, which is the same behaviour as the IS-04 Node API. While an NMOS controller is free to change the transport parameters of streams, POST transport files and activate connections to the API running

Figure 8: Resources on the newly promoted core inherit the identities from the failed core



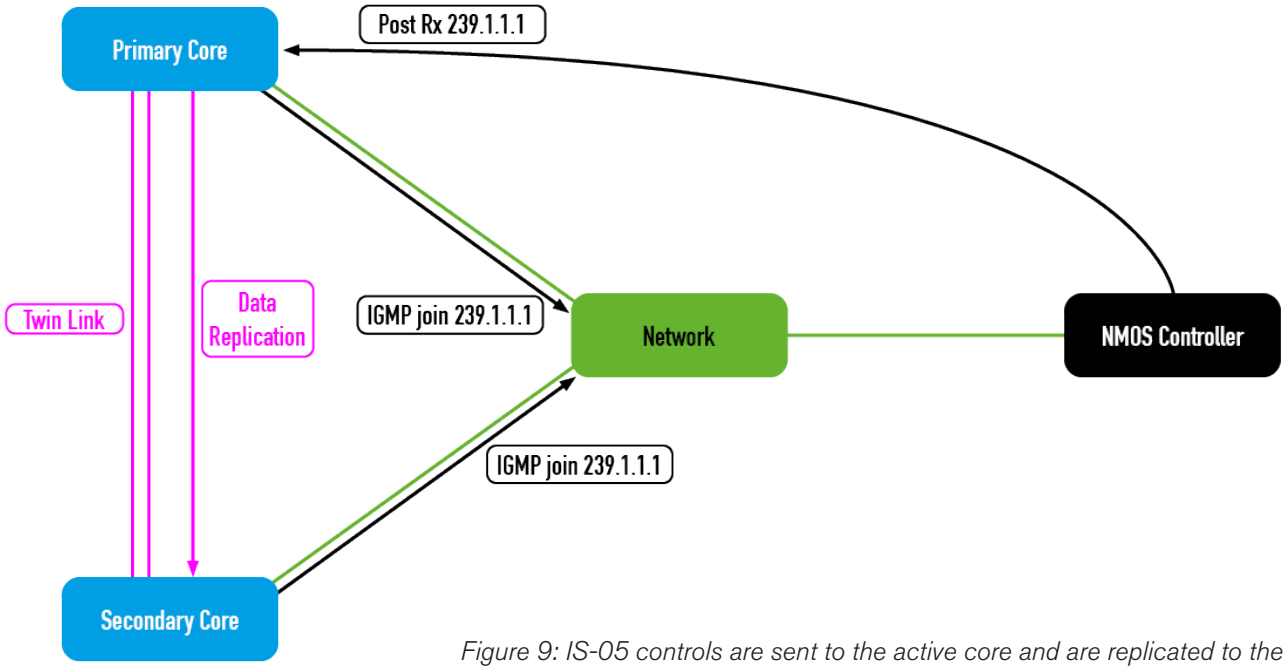


Figure 9: IS-05 controls are sent to the active core and are replicated to the inactive

on the active core, when a change occurs the data is replicated to the inactive core. However, the data is stored outside of the NMOS domain as there is no NMOS API running on the inactive core.

For Calrec receivers, the data received via the active core's IS-05 connection API is used to configure the non-active core's receivers. The receivers will send IGMP Membership reports based on the data that is replicated from the active core. In a fully converged state, both active and inactive cores are receiving multicast streams having already sent the necessary IGMP Membership reports based upon the IS-05 connection on the active core (*figure 9*).

In the event of an active core failure, the following happens:

1. The Connection API is stopped on the failed core.
2. The Connection API is transferred to the newly promoted core.
3. As the discovery information has been updated in the registry through IS-04, NMOS devices and Controllers can now interact directly with the Connection API on the newly promoted core.

PTP

PTP is connected to ImPulse cores using dedicated connections. This means that it is possible to isolate PTP traffic all the way to the end device, bypassing the need to manage traffic priorities between PTP and media traffic on the link to the core.

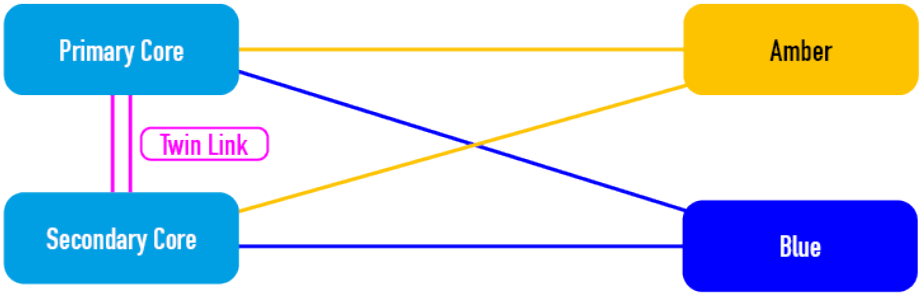


Figure 10: PTP connectivity for hardware redundant cores

Each ImPulse core has two dedicated PTP connections with the expectation that one is fed from the amber network and the other from the blue network.

A redundant ImPulse system requires 4 x dedicated PTP links for full PTP redundancy (*figure 10*).

All other Calrec devices require PTP on the media NICs. In this case, it is encouraged that a suitable QoS strategy is deployed to protect PTP packets through the network.

Performance

Calrec's hardware redundancy model is built on top of ST2022-7, not instead of. That means that ST2022-7 still provides full hitless performance for applicable failures within a facility. Hardware redundancy is an automatic and deterministic method of disaster recovery where the function of an entire area may be lost. The recovery time is determined by multiple elements:

- The number of senders that need to be activated on failover.
- The number of streams in the NMOS/CCP data models.
- The number of Calrec I/O boxes on the network.
- The amount of internal DSP and router resources that need to be activated on failover.
- The size of the current show file.
- External network conditions.

As a rough guide, a relatively small system with a moderate number of streams and a few I/O boxes will be expected to recover in a few seconds.

Conversely, a heavily weighted ImPulse system with the maximum amount of capacity used across 4 x routers (this is a total of 8192 multicast streams between both cores) and a further 30 x I/O boxes on the network will take longer to recover all audio streams.

NMOS service recovery time is also variable depending on system load. Expected recovery time is less than five seconds, subject to NMOS registry configuration.

Data is replicated from the active core to the inactive core every 5 seconds, so Calrec customers can expect recovery performance of 5 seconds or less.

Source Specific Multicast (SSM)

SSM is a technology that requires receiving devices to declare what sources they wish to receive multicast from. The inner workings of SSM are outside the scope of this whitepaper, although RFC 4607 and RFC 3376 are good places to start for further reading.

Whilst you do not need to use SSM to build a resilient network, there are

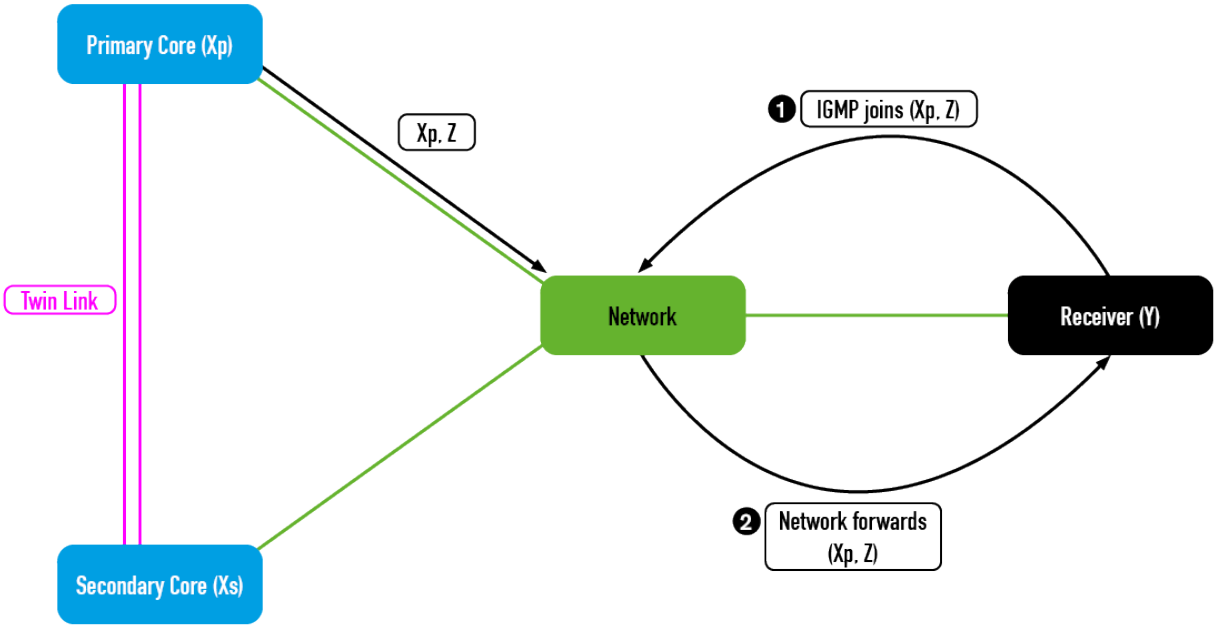


Figure 11: Receiver requests (Xp, Z), the network creates the path

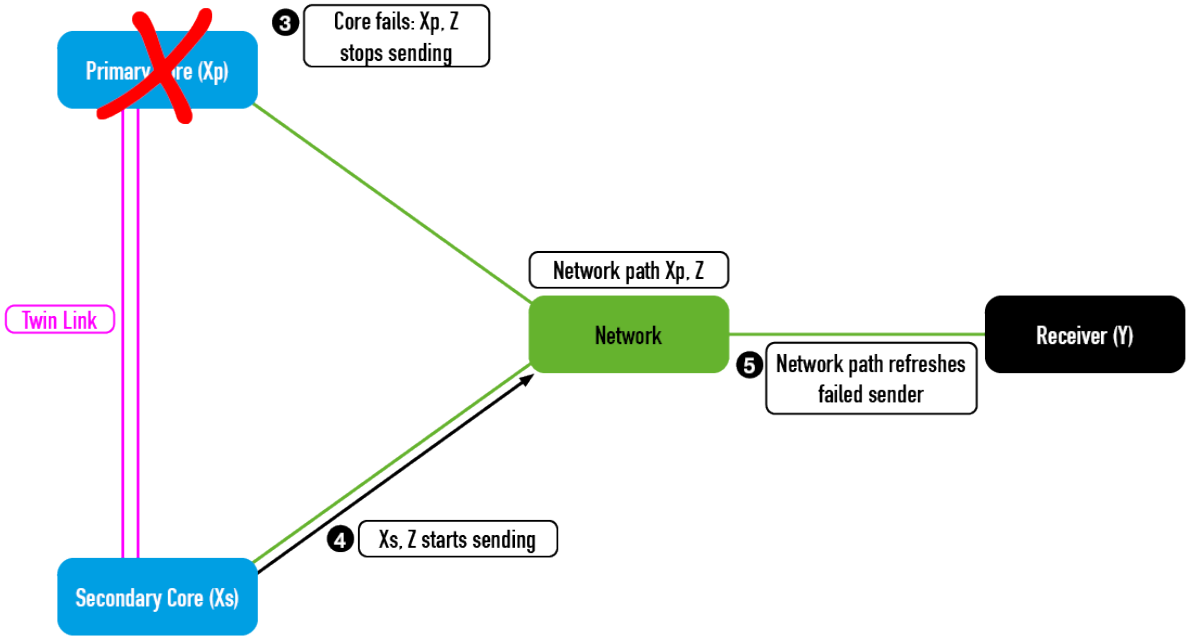


Figure 12: After a failure, the network path does not automatically update to the new source (Xs)

some potential benefits to using this technology. Using SSM can allow a receiver to only receive multicast from trusted senders which vastly improves security over ASM (Any Source Multicast).

A second benefit is ease of deployment through PIM-SSM (Protocol Independent Multicast - SSM). The source data from IGMPv3 request messages are used to find a reverse unicast path to the source to construct the multicast tree. This deployment method reduces the complexity of implementing multicast over a routed network by removing the requirement for a rendezvous point.

As SSM is managed within the network, it does present challenges in the mission to make hardware failover agnostic and invisible to the wider IP environment.

Let's take the following situation (figure 11): Receiver Y wants to receive stream Z from sender Xp. Sender Xp is a Calrec primary core. The manufacturer of receiver Y is irrelevant to this example, but we will assume that it is a non Calrec device. IGMPv3 is used to send an SSM request to the network; the notation is written as (S, G) where S = Source and G = Group. In this example the IGMPv3 Membership Report looks like this: (Xp, Z).

The receiver requests with SSM the multicast Z from source Xp (1). This passes through the network and the network creates the necessary forwarding rules to facilitate this (2). The initial connection that triggers this event could be through NMOS or statically through the Calrec Connect application. In this condition, the connection is established, the receiver receives packets and processes the audio appropriately.

In the event of a hardware failover (figure 12), the processes as described in the previous sections occur and the stream now becomes available from Xs (3). At this point, the stream produced is (Xs, Z) (4) however the network is still forwarding (Xp, Z) (5).

A number of potential approaches can be utilised to resolve this challenge. By its very nature, SSM relies on network interactions rather than the failover behaviour of individual devices, so a network-wide approach is required with cooperation across the network between individual devices and network orchestration.

Summary

This whitepaper has explained the absolute necessity of full hardware redundancy in mission critical broadcast environments and has presented Calrec's solutions in an open and transparent way.

Calrec is proud to engage with all areas of our industry in improving, building and scaling highly resilient IP architectures.

References:

SMPTE ST2110: **https://www.smpte.org/smpte-st-2110-faq**

SMPTE ST2022-7: **https://ieeexplore.ieee.org/document/8716822**

JT-NM TR1001: **https://static.jt-nm.org/documents/JT-NM_TR-1001-1_2020_v1.1.pdf**

AMWA NMOS IS-04: **https://specs.amwa.tv/is-04/**

AMWA NMOS IS-05: **https://specs.amwa.tv/is-05/**

AES67: **https://www.aes.org/publications/standards/search.cfm?docID=96**

RFC 4607 – Source-specific multicast for IP: **https://tools.ietf.org/rfc/rfc4607.txt**

RFC 3376 – Internet Group Management Protocol, Version 3: **https://tools.ietf.org/rfc/rfc3376.txt**

* Assist available on Type R

For more information on how **Calrec** enables its customers to deploy AoIP in resilient and scalable manners, please feel free to contact us at **demo@calrec.com.**

Our Network Specialist Dave Sampson discusses best practices and robust solutions to network design in our series of webinars covering all things AoIP. For more information, visit: ***https://calrec.com/calrec-sound-institute/***

Calrec Audio Ltd
Nutclough Mill
Victoria Road
Hebden Bridge
West Yorkshire
HX7 8EZ
England UK

Tel +44 (0)1422 842159
Email enquiries@calrec.com

CALREC