The state of networked video and integrated system design

Introduction

The convergence of Pro AV and Information Technology.

In this paper we discuss the state of networked video and explore various technologies in use. We also discuss other system components and features necessary to design a complete and secure integrated system rather than a collection of disparate network video endpoints.

Contents

Why networked video?

There are a lot of benefits that can be achieved with network video if the right technology is selected.

Cost

One of the great promises of networked AV is the ability to reduce the overall cost of an installation. Re-using existing network infrastructure and cabling can reduce significant overhead. Any room with a network drop can now support AV. However, based on the type of infrastructure and system design, there may not actually be a significant cost advantage.

Limitless distance

Once video is on the network it can be easily sent anywhere in the building or campus. Open that up to the internet and now it has access to the world.

Boundless creation and consumption points

There is a rich ecosystem of networked products that support AV. AV can be streamed to and from BYOD devices such as laptops, phones, and tablets. There are many network streaming cameras and recording devices on the market at all price points and feature sets. Server-based video streaming solutions such as video reflectors and content delivery networks (CDNs) open up interesting opportunities for new solutions in the pro AV space.

Architectural flexibility

Due to the bidirectional nature of Ethernet networks, switch ports are no longer “inputs” or “outputs.” Even some video endpoints have the capability to both encode and decode video. This makes distribution architectures more flexible, meaning you can reuse existing network infrastructure that was designed with no knowledge of video distribution. And now adding devices to the video distribution system is really easy: just drop an encoder or decoder anywhere on the network.

Network video streaming vs. ethernet-as-infrastructure

There are many different video coding technologies, or codecs, used to distribute video over the network, most of which involve some amount of video compression of the latest video formats.

While all codecs used in the pro AV space offer excellent video quality, they all strike a different balance between bandwidth and video latency. How each codec strikes this balance affects whether content can be streamed on the corporate network, or whether dedicated Ethernet infrastructure must be designed and built. The implications of this tradeoff are discussed below.

Latency

Latency is delay introduced by the process of encoding and decoding the video. Even video distributed over a “perfect” network with “zero latency” will still have some latency introduced by any video codecs.

Tolerance for latency depends on the application. For common in-room applications such as PowerPoint® presentation, up to several hundred milliseconds of latency will go un-noticed. If a presenter is directly interacting with on-screen content via mouse or keyboard, latencies up to 200ms are acceptable, albeit noticeable. As a reference, remote desktop applications typically have latencies in this range.

For applications such as presentations in overflow rooms, in which an audience is observing a presenter via video streamed from another room, even extreme latencies of several seconds will not be noticeable. This is because the viewer has no reference point for what constitutes “live video,” much like a television viewer of a live sports event is unaware of latencies introduced by cable or satellite distribution. For a two-way question-and-answer session the maximum acceptable latency should be limited to a few hundred milliseconds.

Latency can quickly become problematic when using the network for video distribution and analog cabling for audio distribution. Video will be delayed as compared to the audio, and thus out of sync. The human brain is very sensitive to audio arriving ahead of video by even a few dozen milliseconds. However, this effect can be easily corrected by adding the appropriate amount of delay to the analog audio path.

Encoding video means to reformat the data to be suitable for transmission on an Ethernet network. This almost always involves some amount of data compression to fit high-bandwidth video onto lower-bandwidth networks.

Decoding makes the video again suitable for uncompressed baseband transmission technologies such as HDMI® or DisplayPort. In the networked AV world, transmitter and receiver products are often referred to as encoders and decoders.

A codec is an encoding/decoding protocol. It is a combination of “coder” and “decoder.”

Bandwidth

Bandwidth concerns the amount of network traffic used by a given codec, and determines what kind of network is required to support its use. The most bandwidth-efficient codecs typically use around 10 Mbps and as such can be used on existing 1 gigabit corporate infrastructure alongside normal user email and web traffic without much concern. The least bandwidth-efficient codecs can use up to 10 Gbps to transmit 4K content and will immediately monopolize all the available bandwidth of a dedicated 10 gigabit network link, dragging down network operations in the process.

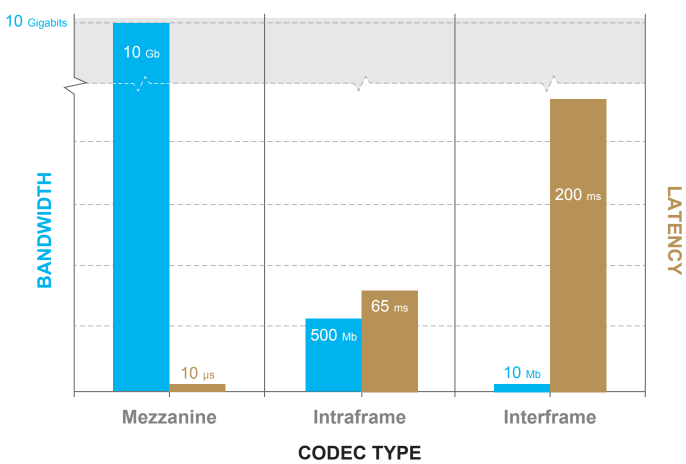
Perhaps the most surprising fact learned by those new to network streaming is that all but the most bandwidth-efficient codecs are inappropriate for use on most corporate networks.

Many devices that use Ethernet as the distribution infrastructure employ high-bandwidth codecs. In these cases, the devices require a dedicated network, and therefore the cost savings associated with leveraging the existing corporate network are not realized. Depending on the application and system design, the cost can even be more expensive than traditional HDBaseT® systems.

Since these systems can't stream over the internet, the rich ecosystem of interoperable devices and services doesn't exist. Users can't send high-bandwidth video to a mobile device or laptop, and it can't be stored by content management systems.

Compression technologies

If implemented correctly and used in the right applications, most modern standard codecs offer excellent image quality. As discussed, each strikes a different balance between latency and bandwidth.



Mezzanine, intra-frame, and inter-frame codecs

Intra-frame codecs treat each frame as an individual entity for encoding. No information from other frames is used to aid in the encode process. Some so-called "mezzanine compression" codecs look only within a small number of video lines. By contrast, an inter-frame codec will look for similarities across different frames to achieve far better compression ratios. The amount of video data used during encoding weighs heavily on the overall latency of the system, since all that data must be buffered to be analyzed as a single entity.

Mezzanine codecs: TICO and DSC

TICO and DSC are competing compression codecs often referred to as "lightweight" compression. They are primarily used when 2:1 or 4:1 compression is needed to squeeze video under a specific bandwidth threshold with minimal latency impact.

While latency is effectively nonexistent, the lightweight nature of these codecs requires the most bandwidth of any of the compression techniques discussed. 4K video will saturate a 10G network link and 1080p video will saturate a 1G network link. These technologies can't be used on existing corporate network infrastructure because they'd erase much of the assumed cost benefit associated with streaming technologies. Video signals at this level are difficult to move around a building without expensive and carefully designed networks, and cannot realistically be distributed between campuses or over the internet.

The TICO Alliance is gaining traction and interoperability in the broadcast video space, but products relevant to the pro AV space are lacking. DSC is employed in the DisplayPort specification, and is widely expected to be employed by future HDMI specifications, as well, but it's not yet readily available for network video products.

Standard video codecs are well understood by both academics and major industry players in the field of video.

Pros, cons, and proper applications for each codec are well researched, well documented and field tested.

The same is not true of proprietary codecs, which are unproven in the industry and may perform poorly in applications not highlighted in product demos and marketing literature.

Intra-frame codecs: JPEG2000 and VC-2

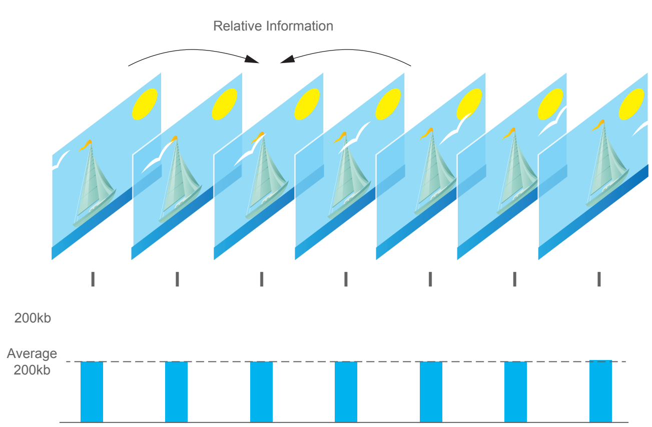
JPEG2000 and VC-2 offer significantly better compression ratios than DSC or TICO. For instance, 1080p and 4K video will compress to several hundred megabytes with these solutions as opposed to the several gigabytes required by mezzanine codecs. These bandwidths may fit on a dedicated 1G network link, but are too high to co-exist on the corporate network with applications such as email and web usage.

Most intra-frame codec installations will require a dedicated network involving significant design effort to properly balance cost and performance. 1G links are fine if they will only ever carry a single video channel, but once two or more channels must be supported 10G links or higher are required. Inter-building links are manageable if planned for appropriately but these codecs still can't realistically be streamed over the internet. And there is no developed third-party ecosystem for these codecs; they are closed solutions. You can't stream to a laptop or mobile device.

At 10-65ms, latency is significantly longer than that of mezzanine solutions, but is still appropriate for most pro AV applications. However, in latency-sensitive applications care must be taken when combining codec latencies with other latency-inducing devices such as video scalers and video teleconferencing codecs (VTCs).

Inter-frame codecs: H.264 and H.265

Inter-frame codecs go a step further to achieve bandwidth reductions that are an order of magnitude better than even the intra-frame codecs. These codecs look at video frames both before and after the frame that's being encoded to take advantage of similarities across the frames.



By way of example, consider a typical meeting in which a presenter is showing a slide deck, and her laptop is set to output video at 60 frames per second. An intra-frame codec such as JPEG2000 encodes every frame the computer generates as a unique entity. An inter-frame codec will determine that many frames are showing the exact same content (the presentation slide) and the bandwidth will essentially drop to zero for those repeated frames. Even for live video, the content usually does not change much from frame to frame, so bandwidth remains low and spikes only upon scene changes.

Since these codecs optimize for bandwidth the latency increases. Latencies are typically in the 200ms range in the best case, with the option of artificially increasing buffer size (which translates to latency) to accommodate video transmission over unstable links or the internet.

Keep in mind that latencies for devices in a system are additive. For instance, video teleconferencing systems (VTCs) use inter-frame codecs and incur their own latency penalty. Using a network video streaming solution to send video to the codec will double the latency penalty, degrading the user experience. For these sorts of applications, low-cost zero-latency HDBaseT links from the laptop to the VTC make the most sense.

H.264/AVC

The Advanced Video Coding (AVC) standard, also known as H.264, is the most common inter-frame codec in use today. H.264 is used in cable and satellite TV distribution, Blu-ray® discs and much of the streaming video used on the internet from providers such as Netflix® and HBO®. H.264 is also commonly used by remote workers in virtual desktop applications and in video teleconferencing systems.

Live encoding devices can typically achieve bandwidths as low as 500Kbps with 1080p video. (That's a 900:1 compression ratio!) However, at these bandwidths video quality is compromised.. Typically, a 10Mbps stream is recommended for high quality video applications, which is still well below the bandwidth threshold for using a 1Gb corporate network.

One important thing to understand is that for particularly bandwidth-sensitive applications, the system designer can elect to optimize for bandwidth rather than image quality. Most codecs will allow the designer to strike a different bandwidth vs. quality balance, but H.264 is the only codec discussed that's suitable for truly bandwidth-sensitive applications. For instance, if streaming to a mobile device over the comparatively expensive cellular network, a lower-bandwidth stream is appropriate.

One downside to H.264 is its lack of 4K support. Due to the extra buffering and compression requirements, it is realistically limited to 1080p for pro AV applications.

H.265/HEVC

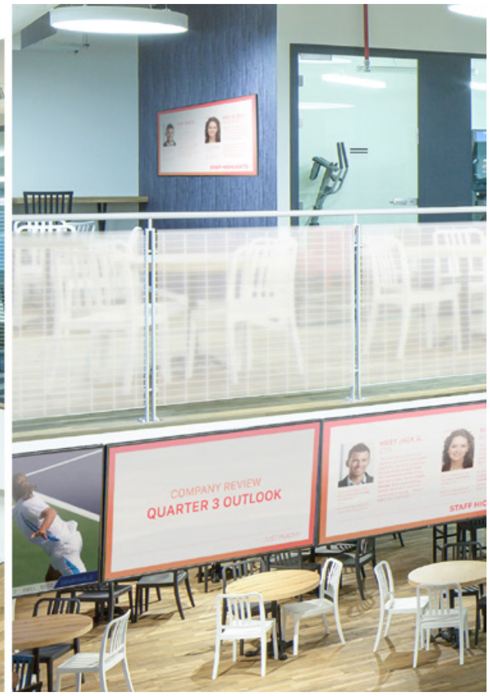
The High Efficiency Video Coding (HEVC) standard, also known as H.265, is the next generation of H.264. H.265 uses similar but more advanced techniques to achieve compression ratios that are about 40% greater than its predecessor. These bandwidth savings will become important as higher-resolution 4K content moves onto the corporate network. However, H.265 encoding is about 10 times as computationally complex as H.264, so 4K encoders currently exist primarily in very expensive devices aimed at the broadcast space. Some high-end video conferencing systems also support H.265 encoding, but only at 1080p. H.265 decoders are significantly simpler than the encoders and will soon be readily available at reasonable prices.

?

Having trouble keeping track of inter-frame vs. intra-frame codecs?

Just keep in mind that intra means within and inter means between or among.

So intra-frame codecs look within each frame while inter-frame codecs look at data across many frames.



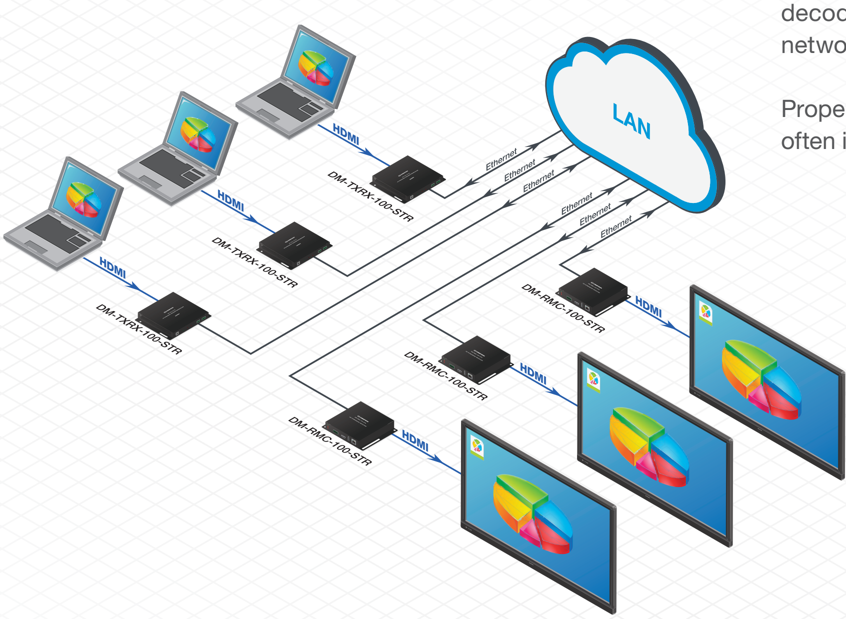
Low-bandwidth network video applications

The low-bandwidth H.264 codec is required to take advantage of existing networks, to enable video transmission at unlimited distances, and to interoperate with the third-party ecosystem. Here we take a look at the various use cases where H.264 makes sense.

Simple matrix switching on the corporate network

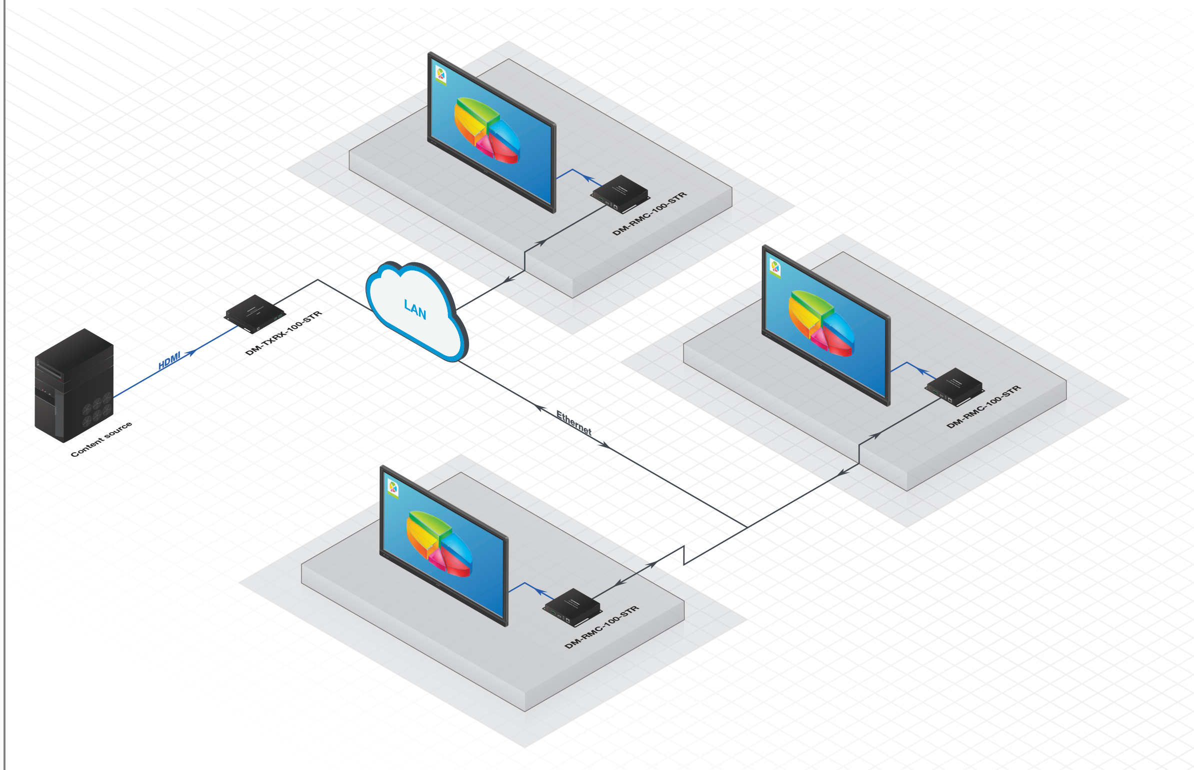
The most conceptually simple way to take advantage of streaming on the network is to use H.264 encoders and decoders to create a large virtual matrix switch, with the network as the switching infrastructure.

Proper HDCP support is critical for these applications as they often involve sources such as Blu-ray players and cable boxes.



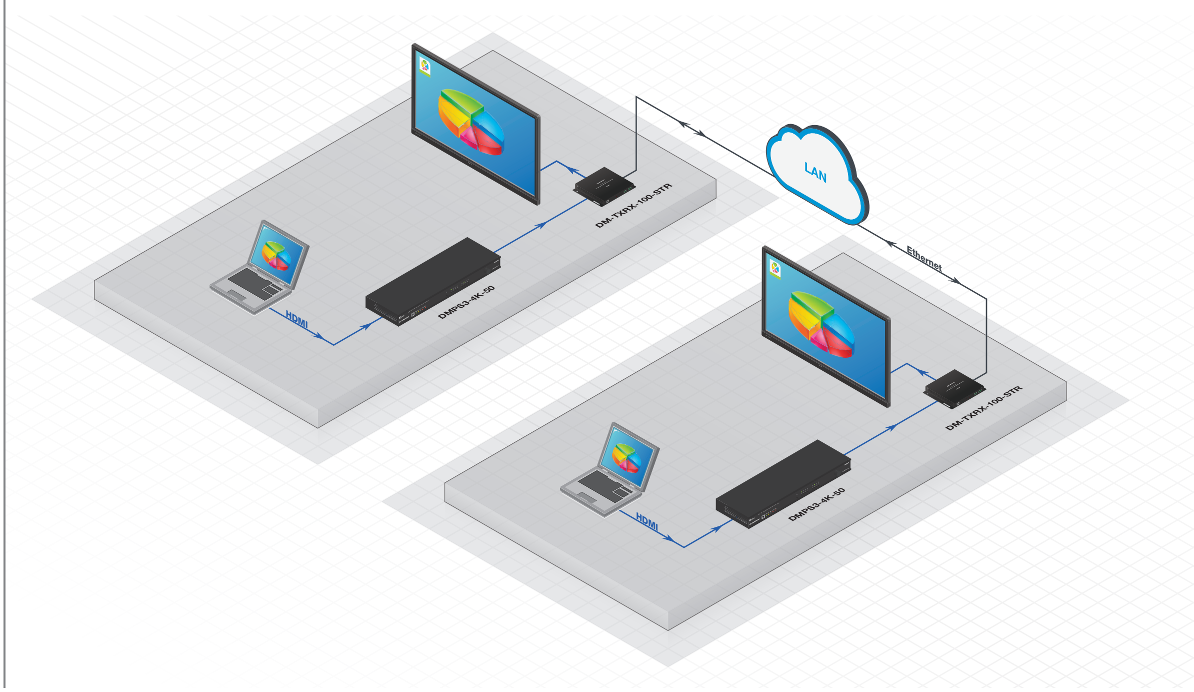
One-to-many distribution

Using multicast streaming, one content source can easily be sent to many displays without using lots of outputs on a matrix switch. This application is common for overflow rooms, lobbies, cafeterias, and digital signage.



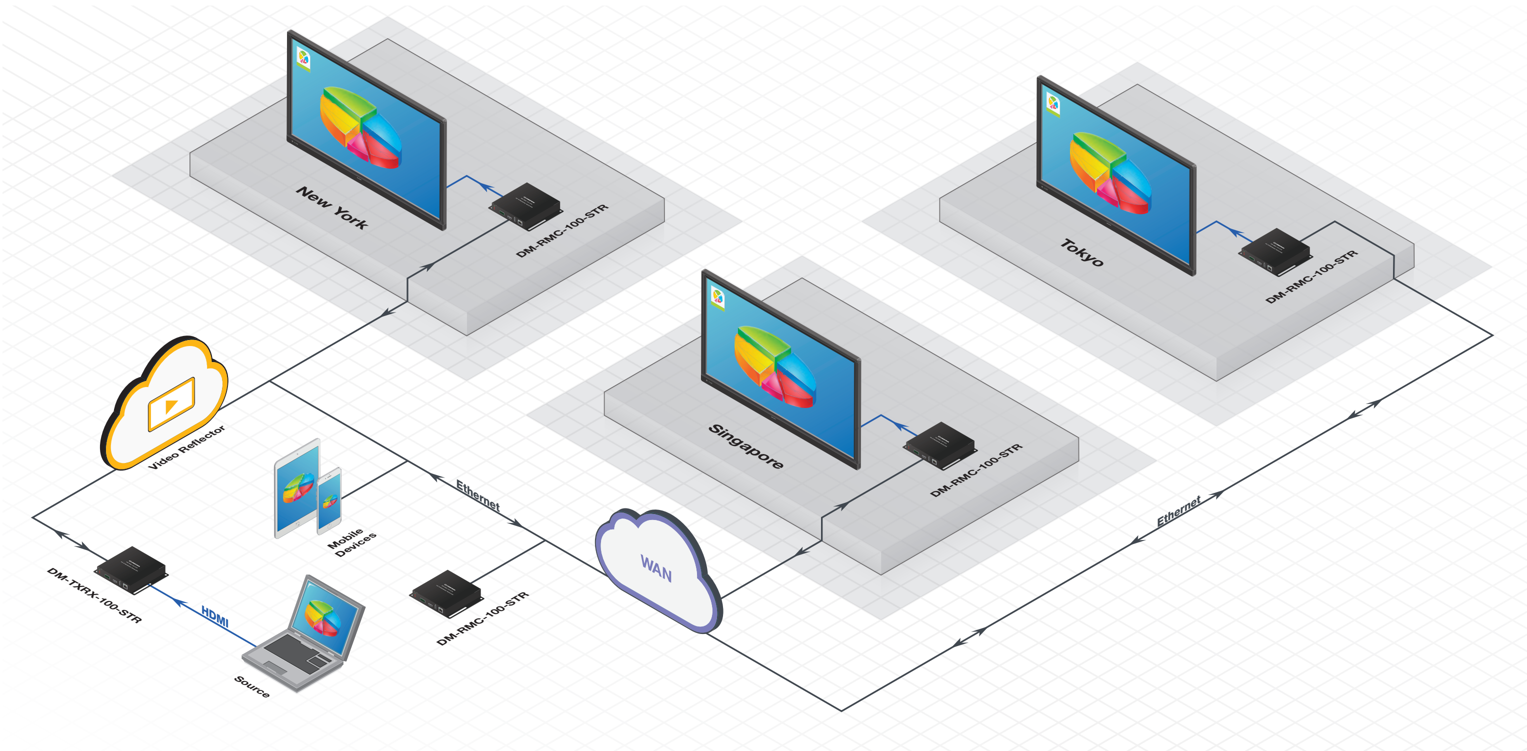
Interconnected rooms in disperse locations

Rooms using HDMI and HDBaseT can be interconnected using streaming encoders and decoders. Any room can become a presentation room and any room can receive the content, even if it's halfway around the world.



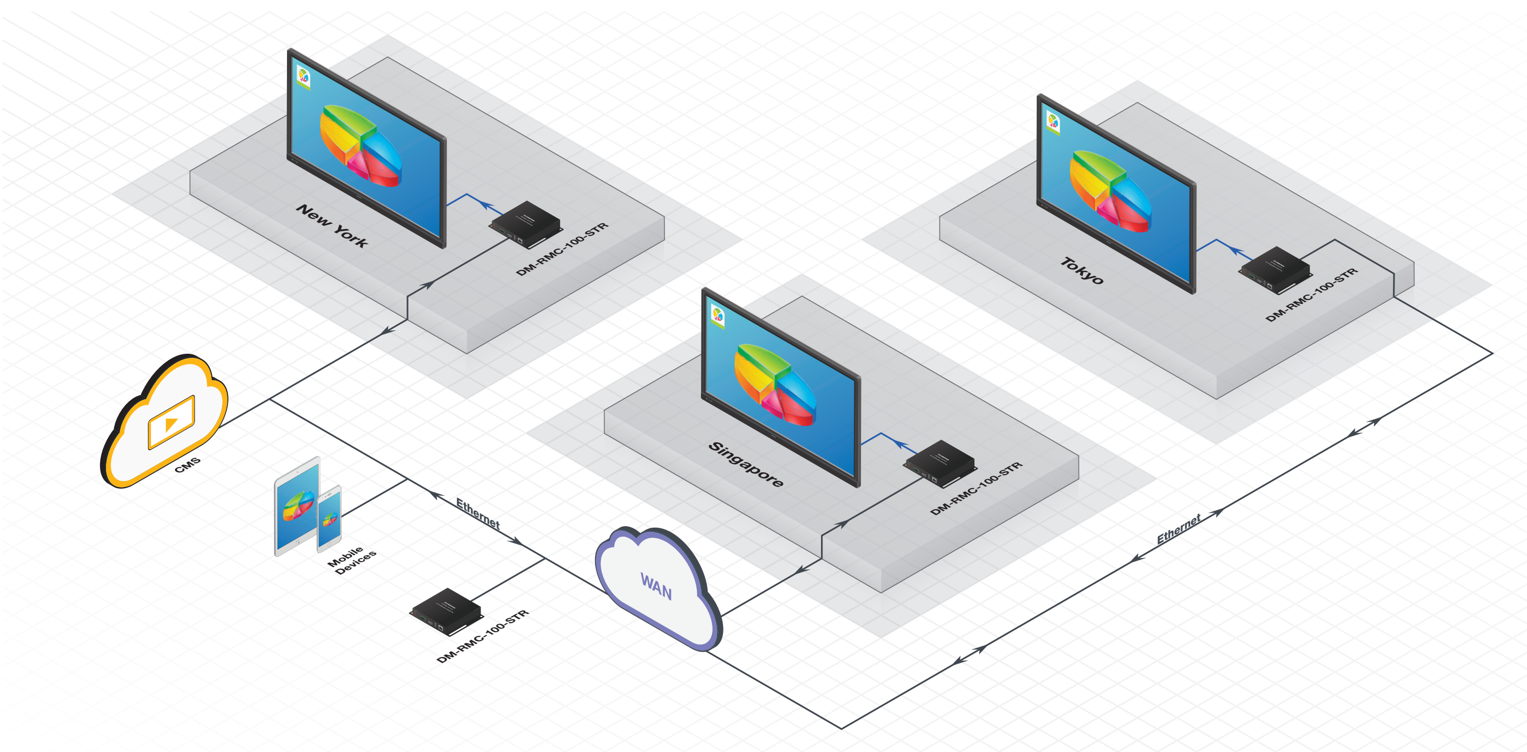
Live video reflectors

Live executive presentations and “town hall” events may be streamed to a reflector service. The reflector transcodes the video to the various formats required by any receiving devices such as laptops, mobile devices, and in-room decoders.



Corporate video library

Many corporations and universities maintain extensive video libraries in content managements systems (CMS) such as Kaltura® or Panopto®. You can stream directly from these systems to endpoint devices such as laptops, mobile devices and in-room decoders.



Beyond codecs: the integrated system

A truly integrated system relies on more than just selection of the right video codec. Irrespective of codec type selected, the devices must be easily and securely managed and controlled.

Integration and control

Video distribution doesn't exist in a vacuum. Routes must be set, video sources and displays must be controlled, and interfaces must be provided to react to user input. A complete solution provides all the necessary pieces for easy integration.

Network management

A modern video distribution system requires monitoring and maintenance. Users may experience difficulties, disconnected cables or power from devices, and devices may occasionally fail, especially in larger systems. Additionally, tracking usage data can help AV and IT departments make more intelligent decisions about future technology investments.

Enterprise network management software enables companies to better support their users and understand how they're using rooms. All components of an integrated system must be manageable, including network video endpoints.

Security

Once video devices are on the network they must adhere to standard IT security policies to safeguard against unauthorized access to content from private meetings. Think about it: while the CEO is presenting her strategy to the corporate board, the content is now moving across the network, which is accessible from any network jack in the building. Properly securing this content is mission critical.

When designing a system, there are two key aspects to consider from a content security perspective: protection of the content and protection of the control interface. Additionally, many corporate IT departments now impose security requirements on all networked devices to protect the network itself.

Content protection

Protecting the content is an obvious need. Luckily, standard network encryption technologies are up to the task. The Advanced Encryption Standard (AES) forms the underpinnings of SSL and TLS, and as such, is used on the open internet to protect all manner of secure online activities, from banking transactions to Facebook® logins. AES is also ideally suited for protecting video, and is the core encryption technology used in HDCP 2.2, the standard content protection protocol in the professional AV industry.

That said, don't rely on HDCP to secure content. Many devices (including most laptops) don't support HDCP, so relying on HDCP to protect the content is not sufficient. If the laptop doesn't turn HDCP on, the content will remain unprotected. It's important to be able to control AES beyond the automatic HDCP control mechanisms.

HDCP also doesn't authenticate the user or the specific receiver device. It only authenticates that the device is a licensed HDCP device. A nefarious actor may bring in a new device from outside the organization, and as long as it's licensed by HDCP, it will happily decrypt the content. An additional authentication layer, such as a content-specific username and password, should be required for the most sensitive content.

HDBaseT is not as susceptible to network-based attacks that put video content at risk. For the most secure applications, HDBaseT remains the right approach.

Control interface protection

Perhaps even more important than protecting the content is protecting the control interface. Content protection is irrelevant if someone can connect to the encoder control interface and disable the encryption! And once access is gained to the configuration interface, video may easily be routed to an additional nefarious device. Web-based configuration interfaces must be protected by HTTPS and the control system interface must also be protected by an encryption mechanism, such as SSL.

Network protection

Corporate IT departments are very careful about which devices have access to their networks. An increasing number of companies now require all networked devices to support an authentication protocol known as 802.1x. This protocol is used to ensure that only devices explicitly authorized by the IT department have access to the network.

802.1x has specific hardware requirements and is not supported by all devices. Be sure this requirement is considered before designing a streaming system or you may find that the IT department won't allow it on their network.

Streaming ecosystem compatibility

Unfortunately, even when two network AV devices use the same video codec, they may not be compatible. Different protocols may be used to transport the video data and communicate system capabilities. Furthermore, different codecs may be used to encode the audio data. It's important to support all industry-relevant protocols, especially to ensure compatibility with devices in the H.264 ecosystem.

HLS and DASH

HTTP Live Streaming (HLS) and Dynamic and Adaptive Streaming over HTTP (DASH) are two of the more common modern web-based video distribution mechanisms. They separate video packets into small HTTP downloads, allowing the video data to easily pass through firewalls and corporate proxy servers. Modern CDNs such as Kaltura and Panopto employ these protocols for content transmission, so receiving devices must support these protocols for compatibility.

RTSP

The Real-Time Streaming Protocol (RTSP) is used to negotiate video format support between encoding and decoding devices. This protocol is also used to relay playback commands such as "Play," "Stop," and "Rewind." RTSP is popular with streaming cameras.

Blind multicast

With blind multicast transmission multiple receivers are subscribed to the same transmission stream. Both sides must be configured with the same streaming parameters (e.g., resolution, bitrate) to ensure interoperability. This format is common in broadcast and digital signage applications.

Audio codecs

While most of this white paper discussion has revolved around video codecs, audio codecs are important as well. The ITU G.711, Advanced Audio Coding (AAC) and pulse code modulation (PCM) standards are both common in the streaming video space.

Networked video & HDBaseT

Network streaming is great for a lot of applications and enables exciting new use cases. However, traditional circuit-switched HDMI and HDBaseT solutions still make sense in many instances, especially compared to high-bandwidth solutions. Given the cost and density advantages of network switching infrastructures, how can this be?

Cost

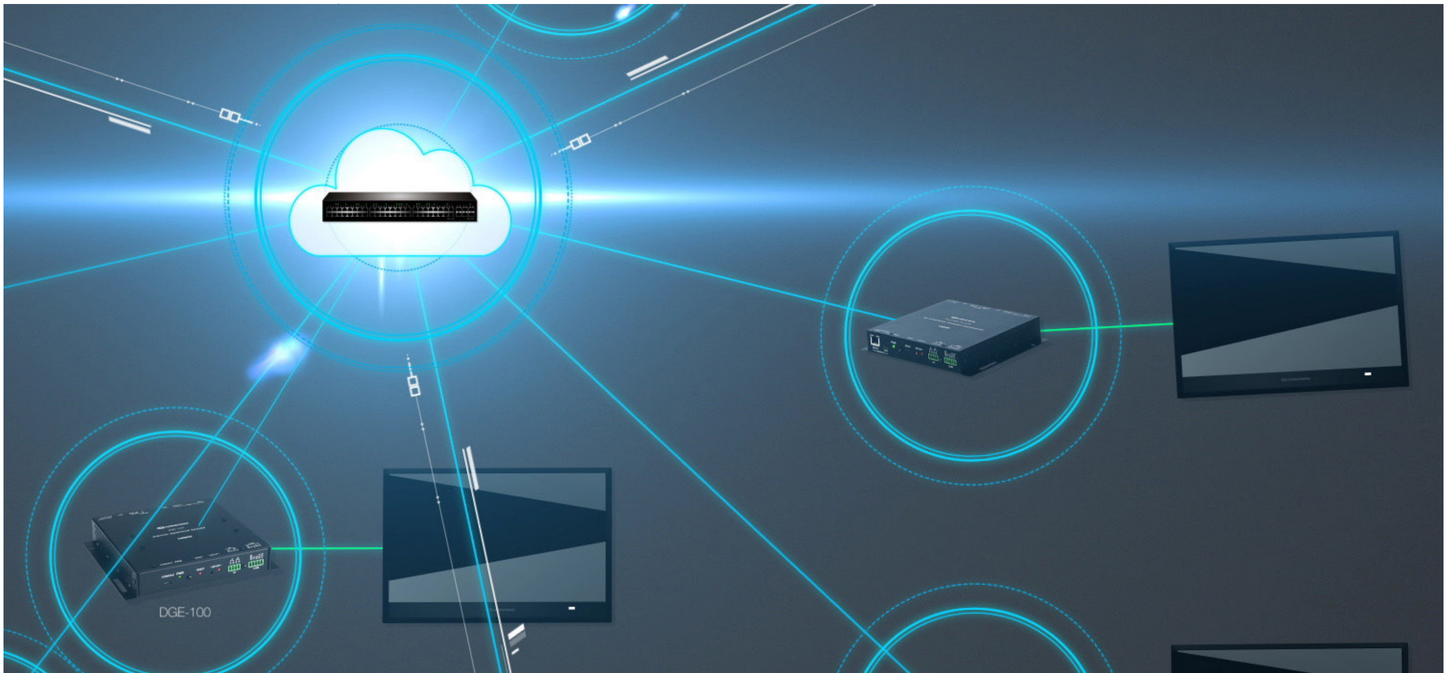
Ethernet ports on a network switch are significantly cheaper than HDBaseT ports on a matrix switch. But high-bandwidth networked encoders and decoders are still significantly more expensive than comparably-equipped HDBaseT transmitters and receivers. So for simple point-to-point video extension HDBaseT remains the more cost-effective option.

The same remains true for many larger installations. In a networked video installation, every source and receiver device requires an expensive encoder or decoder, regardless of whether it's installed in a rack or remotely. In a circuit-switched installation, in-rack sources and receivers can take advantage of relatively low-cost HDMI inputs and outputs native to the chassis.

Density

Ethernet switches clearly beat HDBaseT switches for pure density. A 48-port Ethernet switch can easily be had in a single rack unit. But again, the local in-rack HDMI encoders and decoders erase much of this benefit. In-rack HDMI inputs can be had in significantly higher densities on circuit switches than on network encoders and decoders, quickly eclipsing the density advantage of the network switch.

Installed HDBaseT systems will be used for many years, and new HDBaseT systems will be designed and installed for some time as well. As network video gains popularity, it will be added to existing HDMI and HDBaseT systems, thus you'll need effective and elegant hybrid solutions to manage that transition seamlessly and reliably.



Conclusion

Networked video technologies are not created equal. Low-bandwidth H.264 video is the most widely-available codec that allows integrators to take full advantage of the benefits of networked video: low cost re-use of existing infrastructure, limitless distance, and the well-developed ecosystem of interoperable devices and services. Higher-bandwidth options optimize for latency but require the build-out of dedicated network infrastructure, can't be streamed over the internet, and don't offer interoperability with other devices.

However, it takes more than the proper codec selection to build a complete video distribution system. Systems must securely integrate with control systems for routing as well as control of video sources and displays. And the system must be maintainable, which means it must be able to be monitored on the network.

For latency-sensitive in-room applications, HDBaseT will remain the dominant technology for some time, significantly influencing system design and management even as network AV becomes increasingly popular.

Most modern successful system designs will involve a mix of HDBaseT and network AV, allowing the designer to take advantage of each when appropriate. Be sure to consider how to integrate copper, fiber, and Ethernet infrastructures during system design to ensure a successful installation.