WHITE PAPER

Personal Telepresence: The Next Generation of Video Communication

www.vidyo.com
1.866.99.VIDYO

WP_0609_PT_US

What's Wrong with Legacy Video Conferencing?

Many things do not work in legacy video conferencing. In previous years, the costs of computing for encoding and decoding in video conferencing were high. Hence, groups of people gathered in a room to share a video conferencing system. They sat around a conference table with a video camera and a screen and were connected to a similarly equipped remote conference room. The room system camera was typically set to capture everyone in the room and the meeting participants in the same room frequently faced each other instead of the camera and receiving display, leaving remote participants feeling disconnected from the conversation. For reliable transmission in such legacy systems, expensive dedicated networks or quality of service enhanced bandwidth has to be used.

If the meeting called for a third endpoint to join, a costly Multipoint Control Unit (MCU) was required to perform transcoding for all connected parties. Transcoding at the MCU not only degraded the image quality of the video, but also introduced considerable latency in the video stream, resulting in unnatural communication experiences. Overall, the experience of the users of these legacy video conferencing systems was such that they often decided to just stick with audio calls, leaving the expensive video conferencing equipment commonly unused.

Some of the above problems have been solved by Telepresence systems. Telepresence systems were designed to provide natural, reliable, high quality visual communication with very low latency. The architects of these systems realized that the transcoding MCU had to be eliminated. They replaced the MCU with high speed video switches. Additionally, to provide more natural communication, there is a one to one participant to camera ratio such that the participant is positioned closer to, and facing, the video capture device and receiving display.

Evolution of Video Conferencing

Legacy MCU based video conferencing systems are expensive, require costly QoS networks, are cumbersome to use and provide experiences that fall short of high quality, low latency, natural communication.

Telepresence systems work great and provide high quality, low latency visual communication experiences, but they are very expensive, still rely upon costly QoS bandwidth and constrain users to fixed, dedicated rooms.

Personal Telepresence VidyoRouter™ based systems combine the quality of telepresence with the convenience and universal access of web conferencing, providing high quality, low latency communication over general purpose IP networks, like the Internet, at price points that make it possible to deploy to every desktop in an organization.

However, telepresence systems are very expensive, require dedicated lines and most of all, you cannot take them with you when you travel. At the current price points for these systems, even the largest organizations will not be able to achieve universal access throughout the organization. The video switches that replace the MCU rely upon expensive, nearly lossless, networks to provide the quality and reliability users demand. The fixed location of the endpoints requires a participant to travel to the specific destination to participate.

The Solution: Vidyo Personal Telepresence

Today, the picture has changed. Everybody has the computation power needed for encoding and decoding real-time video conferencing on their desktop making cost effective telepresence feasible. Hence, the Telepresence system can be put onto the screen of every computer, making telepresence personal and universally deployable. However, personal telepresence has the following technical challenges:

- Conference participants that gathered in a conference room in legacy video conferencing are now sitting in front of their desktops. Hence, multi-point calls are much more likely to occur.
- The computational resources available on desktop computers vary greatly. Hence, the compute power required to decode multipoint calls must be adjustable.
- The diversity of location of the participants in the conference make it much more likely to have unreliable connections with varying throughput, e.g. in a hotel room. Hence, a robust video conferencing solution is required.

Legacy systems cannot solve these technical challenges. Transcoding MCUs do not scale to arbitrary numbers of calls and arbitrary numbers of participants in each call. Providing all of the participants with dedicated lines is neither practical nor cost effective.

The industry's first Personal Telepresence system, architected by Vidyo, solves all of the above problems by delivering:

- the high quality, low latency visual communication of telepresence systems
- point and click simplicity on personal devices that travel with users to their natural work environments
- low cost, ubiquitous access, only requiring general-purpose IP networks, such as the Internet

At the core of the architecture is the patented VidyoRouter™, which intelligently routes video streams to both VidyoDesktop™ and VidyoRoom™ endpoints based upon bandwidth availability, endpoint processing power and display resolution capability. The video streams are represented in the scalable video format that is standardized in the SVC (Scalable Video Coding) extension of H.264/AVC. SVC efficiently represents multiple spatial and temporal video resolutions corresponding to different bit rates. This combination provides the solution to the technical challenges of personal telepresence.

The VidyoRouter passes on exactly those packets of the SVC video streams in order to

- provide the video resolutions in multi-point calls allowing continuous presence and personal layout
- limit the number of video data streams to be decoded to account for the hybrid computational resources of the various desktop computers and room systems in the conferences
- matches the number of bits to be transmitted for each transmission link, avoiding delays and congestion

The VidyoRouter together with the SVC decoder in the VidyoDesktop or VidyoRoom provide the error resilience that is needed for personal telepresence. The VidyoRouter provides stronger protection against packet losses. But only the lower video resolution is protected so that it can be done with very little overhead. If a packet of the higher resolution part is lost, the SVC decoder can conceal the missing part. This is only effective with spatial or quality scalability.

Unlike the MCU, the VidyoRouter introduces virtually no latency to the bit streams carrying the audio and video information since all encoding and decoding is done only at the end points. Vidyo's approach to distributed processing for encoding/ decoding enables a single VidyoRouter to support up to 100 simultaneous HD connections and scales easily to thousands of end points via additional VidyoRouters. A transcoding MCU based system is not able to achieve Personal Telepresence as latency increases with higher volumes of multipoint calls and the expense ratchets quickly as more costly, low port density MCUs are added. Additionally, the use of SVC in all components of the network has made Vidyo's architecture the most error resilient in the industry. VidyoConferencing™ systems are able to perform in environments with more than 20% packet loss. By way of comparison, MCU-based architectures start to fail at just 2% packet loss.

Why a VidyoRouter based solution is required for Personal Telepresence

Personal Telepresence requires high quality, low latency, highly error resilient video communication on personal devices at prices that enable organization wide deployment. Here's why only Vidyo can deliver Personal Telepresence today.

MCU-Based

- MCU transcoding degrades image quality & adds latency
- Dependence on QoS bandwidth adds expense and limits mobility
- Cost prohibitive for universal deployment

Telepresence Switches

- High quality image with low latency
- Dependence on QoS bandwidth adds expense and limits mobility
- Cost prohibitive for universal deployment

VidyoRouter-Based

- High quality image with low latency
- High error resiliency enables use of low cost, existing IP networks, and the ubiquitous Internet.
- Cost effective for large scale deployment using existing personal computers

Gateway from the Past to the Future

For all of the reasons cited above, no legacy system with a transcoding MCU is capable of delivering Personal Telepresence, rendering such systems obsolete in a world where communication needs to take place easily, reliably, from anywhere at any time and universally throughout an organization. But what happens to organizations that have made non-trivial investments in legacy endpoints?

Should they decide to cut their losses and throw out tens to hundreds of thousands of dollars of equipment, or should they continue to invest in the legacy technology, hoping to derive some value from the dollars previously spent? In most cases, the answer is *neither*. Keeping with the philosophy that every endpoint in the network should receive what it is capable of receiving for its own optimal performance, Vidyo designed the VidyoGateway™ to connect legacy endpoints to Vidyo's native SVC endpoints. This way the endpoints which require transcoding get it, and those that don't suffer no performance impact and continue to receive non-transcoded SVC bit streams.

Forcing Vidyo's personal device endpoints to natively connect directly to transcoding MCUs would have been a dysfunctional approach to bridging the two technologies. Why? Doing so would negate all of the benefits of an SVC-enabled endpoint, making it as susceptible to degraded image quality, latency, and error intolerance as MCU dependent endpoints. Since the personal telepresence endpoints can conceivably be deployed to every desktop in an organization, while legacy endpoints are typically limited to a significantly smaller number of conference rooms, it makes sense to enable the majority of users to perform in an optimized setting, via the VidyoRouter, while providing the transcoded bit streams needed to legacy endpoints via an edge of the network device.

No Room System Left Behind

A typical organizational deployment of a video conferencing solution may have 50 room systems, but more than 1000 desktop users. In mixed environments, Vidyo enables the larger community of SVC enabled users to experience optimal performance while providing transcoding only to legacy endpoints that require it.

Not All SVC Implementations Are Equal

You know that SVC is a key ingredient in making personal telepresence a reality, but did you know that it is possible to be SVC conformant but not gain any benefit beyond what is available in H.264/AVC? There are three components to scalability, temporal, spatial, and quality. For the purpose of this discussion we will focus on temporal and spatial. Temporal scalability is available in H.264/AVC and provides some benefit in terms of preserving video quality in lossy environments by enabling an endpoint to skip some of the pictures in the video stream that have not been received in full or have been lost. Spatial scalability is only available in SVC and adds an enhancement layer that enables the endpoint to skip some resolution details that are either lost or not required by the receiver, hence maintaining a higher frame rate in the presence of errors, but just as importantly, allowing for rate matching and layout without transcoding in an MCU. The combination of temporal and spatial scalability, properly implemented in all elements throughout the network is required to achieve the error resiliency demonstrated by the VidyoRouter based architecture. Some MCU based manufacturers claim that they support SVC, but have only implemented temporal scaling. While this makes their implementation a nice marketing checklist item, relative to full SVC implementation, it provides only partial performance enhancement for error resilience in point to point calls, but little benefit to the target application, multi-point conferencing.

SVC is Always Better than H.264/AVC Baseline in Video Conferencing

The benefits of real-time video communication with SVC providing native resolution transrating in lieu of transcoding are clear by now. But what happens if the transmission channel is perfect and no transrating is required as in the case of a point to point call? Even in this case, SVC is always better than H.264/AVC. Thomas Wiegand, an editor of the H.264/AVC standard and its SVC extension, as well as one of the chairmen of the Joint Video Team (JVT) standardization committee that created the standard says, *"An error-resilient SVC video stream is composed of an H.264/AVC Baseline Profile part and the SVC part. When the transmission channel is perfect, and no transrating is required, the H.264/AVC Baseline Profile part can be dropped from the SVC video stream. What remains is the SVC part that has the coding efficiency of the H.264/AVC High Profile and typically requires 10% less bit rate than Baseline Profile for the same video quality."* To put that in context he adds, *"To summarize, even in scenarios where SVC specific coding tools are not required, Vidyo's SVC based solution provides a 10% improvement. In all other scenarios, the Vidyo system is dramatically better than anything else."*

> ***Reduced Processing with SVC***
>
> *Vidyo's implementation of SVC doesn't simply redistribute the processing burden of the transcoding MCU, but actually reduces the overall system processing required. The small incremental processing required to encode and decode SVC as compared to AVC is a tiny fraction of what is saved by eliminating the transcoding MCU from the network.*

Summary

Personal Telepresence is the next generation in video communication. It delivers affordable, universal, and reliable video communication. The VidyoConferencing system provides the solution to the business and technical challenges of personal telepresence.

Seeing Is Believing

As is the case with any disruptive technology that threatens the seated incumbents, there is a lot of noise in the industry regarding the topics addressed in this document. You will undoubtedly hear things that contradict statements made here and you will want to know, "Is this VidyoRouter based architecture really that much better than the legacy MCU based architecture?" See for yourself at www.vidyo.com/comparison_video and you decide.

References:

1. A. Eleftheriadis, R. Civanlar, and O. Shapiro, "Multipoint Videoconferencing with Scalable Video Coding," Journal of Zhejiang University SCIENCE A, Vol. 7, Nr. 5, April 2006, pp. 696-705. (Papers from the Packet Video 2006 Workshop.)

2. M. Noble, "Vidyo Comparison Video," www.vidyo.com/comparison_video, June 4, 2009.