Vidyo®
Personal Telepresence

TECHNICAL NOTE

"Natural Continuous Presence"™ in VidyoConferencing™

www.vidyo.com
1.866.99.VIDYO

TN_1209_LATENCY_US

"Natural Continuous Presence"™ in VidyoConferencing™

The Importance of Low Latency to Natural Video Communication

Michael Horowitz

Compared to what was available just a few years ago, the video quality produced by most high definition (HD) video conferencing systems has significantly improved today's video communication experience. Among the factors contributing to this improvement are more capable hardware platforms, better video coding technology (i.e., H.264), and the availability of the bandwidth required to support HD video at a reasonable cost. While increased resolution is an important part of *natural* video communication, other attributes need to be considered as well, including video quality at lower data rates, resilience to network transmission errors, and – key to high quality natural video communication – low latency.

Latency is largely transparent to users when the one-way or *end-to-end* delay is kept below 150 milliseconds (ms)[1] . Natural conversation begins to break down (e.g., people talk over each other) when end-to-end delay exceeds 250 ms. Today, most video conferencing endpoints in a 30 frame per second (fps) *point-to-point* conference call have an end-to-end delay of approximately 200 ms excluding network transport time. Once you add the typical end-to-end network transport time of 20 ms to 50 ms within the continental United States (more overseas), the important 250 ms threshold is met or exceeded.

Only recently have 60 fps video conferencing endpoints become available, with the launch of VidyoConferencing products and technology in early 2008. The more natural rendition of motion on these systems (as compared to 30 fps systems) is just one benefit resulting from the increased frame rate. Arguably the most important benefit is a significantly lower end-to-end delay. This fact comes as a surprise to most.

To understand why higher frame rate leads to lower latency, consider the following question: Why does the end-to-end delay of a typical video conferencing system exceed 250 ms? That is, where does all the latency come from? The answer is that it originates in many different places throughout the video processing and transportation pipeline including the video encoding and decoding processes and internal buffering of data, commonly referred to as jitter buffering. Jitter buffers are used to ensure the video preserves the natural smoothness of motion and to protect video integrity in adverse network conditions. Many delay sources contribute to the overall delay in a way that is directly proportional to the frame time – a point key to the understanding of the relationship between frame rate and latency.

[1] ITU-T Recommendation G.114, One-way Transmission Time, Section 4, page 2, 05/2003.

For example, if one uses a 3-frame data buffer to ensure natural motion smoothness, at 30 fps the buffer will add a delay of 3 x 33.33 ms = 100 ms. In the same design, at 60 fps the added delay is halved to 3 x 16.67 ms = 50 ms.

In addition, while a 3-frame buffer as described above may be sufficient for ensuring natural motion smoothness when network conditions are excellent, it may not offer adequate protection against the large variation in packet delays often associated with IP and wireless networks. To protect against such network conditions, most systems are forced to use deeper buffers that add yet more delay even at higher frame rates. Vidyo's unparalleled error resilience allows it to protect the integrity of the video stream without increasing the end-to-end delay.

At Vidyo we believe that high video quality *and* low-latency are the cornerstones to natural video communications. In describing his personal experience in a recent Wainhouse Research Bulletin[2] , Video Conferencing Industry Analyst, Andrew Davis said this about the VidyoConferencing: "… Ira and I did our Wainhouse Research proprietary high technology test for latency and found the delay to be almost imperceptible."

In addition to the high level of latency associated with point-to-point calls using traditional video conferencing systems, the end-to-end latency often *doubles* for multi-point calls. Multi-point video is most commonly presented using one of two methods. Using the first method -- "video switching" -- video is displayed from one remote source at a time (typically that of the loudest speaker). The end-to-end latency associated with this method is essentially equivalent to that of the point-to-point system.

Business grade systems have two essential requirements that preclude video switching: rate matching and continuous presence ("CP"). The former is the need to support different line speeds for different users. The latter is the requirement to display video from multiple participants simultaneously. Using traditional techniques, both rate matching and continuous presence require brute-force transcoding of video. In addition to requiring expensive hardware and lowering video quality[3], transcoding requires precious time. In fact, it is common for the end-to-end latency in a continuous presence video conference to exceed 400 ms, well above the point where natural conversation begins to break down.

---

[2] Wainhouse Research Bulletin, Page 2, Volume 9, Issue #25, Sep-15-08. (http://www.wainhouse.com/bulletin)

[3] H.264 video encoding is inherently lossy. Consequently, the re-encoding component of the transcoding process lowers video quality.

Tapping the power of the new H.264 scalable video coding (SVC) standard, Vidyo has developed the VidyoRouter ™ that produces continuous presence as well as rate matching without transcoding. This revolutionary technology eliminates the need for expensive hardware and loss of video quality. Most importantly, the latency is essentially the same as that in Vidyo's point-to-point system – almost imperceptible. Ultra-low-latency continuous presence adds a dimension of naturalness to multi-point video communications never before seen. We invite you to experience Vidyo's Natural Continuous Presence™ and see for yourself.

* * * * *

Dr. Michael J. Horowitz, a leading architect of low-delay video codecs, heads the Algorithms Team at Vidyo, Inc. The team develops H.264 Scalable Video Coding (SVC) and other standards based codecs. Earlier in the decade at Polycom he led the team that developed the first commercially available H.264 Advanced Video Codec (AVC). Dr. Horowitz has 16 patents and patents pending and has served as chair of the ad-hoc group for H.264 Complexity Reduction in the ITU-T's Video Coding Experts Group and the Joint Video Team (JVT).