Vidyo Disaggregates Multipoint Control Unit (MCU) to Scale Videoconferencing

Executive Summary
Camera sensors and video-capable displays are following Moore’s Law: declining cost and increasing quality curves. Video hardware is now ubiquitous in mobile computing devices (smartphones, tablets, and notebooks). As camera module costs continue to decline, a growing range of low-cost consumer and commercial devices will be video-enabled, including many types of Industrial Internet of Things (IIoT) endpoints.

Multipoint control unit (MCU) economics are not keeping pace. They are legacy artifacts of a decade when high-resolution cameras and displays were expensive, the compute power to run video codecs was not available in PCs, and even “good enough” video quality required expensive dedicated communications bandwidth. MCUs became a dominant design for video conferencing, because there was no alternative at the time.

About eight years ago, Vidyo invented a Scalable Video Coding (SVC) based architecture for real-time video communication. Since then, they have promoted their SVC know-how to drive improved video bandwidth efficiency in today’s widely used video compression and transmission standards. Vidyo is also contributing heavily to upcoming standards, notably Google’s WebRTC and H.265 (HEVC).

Leveraging their SVC insight, Vidyo then disaggregated the legacy videoconferencing MCU. They created a simple, virtualizable, low-latency video streaming router technology that they brought to market as their VidyoRouter product line. It is the heart of their VidyoConferencing product portfolio and platform. Vidyo enables third parties developing web, appliance, or IIoT video products to tap into their client and VidyoRouter server control interfaces via their VidyoWorks Client and Web Services APIs. And, if integration partners and customers choose, they may ingress networked video through the VidyoReplay API for content management.

Vidyo does not treat media conferences as scarce resources. The VidyoConferencing platform enables secure, reservationless meeting invitations that support growing mobile and remote work habits, and it connects legacy room-based deployments to desktop and mobile conferencing. Audio-only endpoints, both phones and room-based systems, can also join VidyoConferencing sessions so that anyone can join a Vidyo videoconference from any audio or video communications system.

Moor Insights & Strategy believes the path to widespread deployment of low-cost video streaming IIoT endpoints, and actual use of videoconferencing on existing personal endpoints, is that video streams must adapt fluidly to instantaneous network conditions.
Vidyo’s technology suite and product portfolio are uniquely designed to continuously deliver the highest possible video quality over whatever bandwidth is available for the duration of a videoconference or video stream. And it does so with off-the-shelf server and endpoint hardware.

**Should I Stay or Should I Go?**

Face-to-face (F2F) interaction is mandatory for the smooth functioning of any organization. While virtual meetings are unlikely to displace all physical F2F meetings completely, the high costs of travel (direct expenses, opportunity costs, impact on personal life, *etc.*) is causing an increasingly mobile workforce to look for videoconferencing solutions that better replicate the physically-immediate experience of in-person F2F interaction.

*Cisco’s Global Business Service Adoption Trends* calls out that “*desktop video conferencing will be the fastest growing service across all business categories (data & mobile) from 2012 to 2017*. Cisco is predicting a 51.7% CAGR for desktop videoconferencing over those years, and the installed base of these systems will grow from 26.6 million users in 2012 to 213.5 million users in 2017.

The costs of deploying this new videoconferencing infrastructure are not always direct. Organizations have to weigh the impact of introducing videoconferencing solutions on their networking infrastructure and also weigh the impact of their network investments on the quality of their videoconferencing system’s user experience.

In the business world there is no such thing as a “free lunch”. *Cisco’s Visual Networking Index Services Adoption Forecast* says that increased adoption of “advanced video communications” in the enterprise segment will triple enterprise network traffic between 2012 and 2017. Upgrading network bandwidth is expensive. It might involve running new wire or deploying new wireless access points, upgrading to higher bandwidth switches and routers, adding additional tiers of switches and routers to add bandwidth (or changing their logical network topology), and/or buying lower latency switching and routing solutions (and perhaps increasing their bandwidth while doing so).

This is why Cisco is so interested in videoconferencing. A large portion of their future revenue will be derived from the impact of increased video communications usage over enterprise networks. “*For every dollar we sell of video endpoints we typically will also sell between $3-$5 of networking.*” (*Guido Jouret*, Cisco’s GM of Internet of Things)

Cisco’s interest—and their data—are derived from a traditional “desktop” view of videoconferencing within a firewalled enterprise network. However, new mobile and work from home (or coffee shop, *etc.*) usage models are opening up videoconference participation to low-cost, variable quality endpoints joining videoconferences from unmanaged networks across the open Internet.
When there is no guarantee for quality of endpoint audio and video nor their network latency, bandwidth, and packet integrity (how many packets are lost or damaged in transit)...when all of that is completely outside of an organization’s IT control, then it is time to consider different solutions. Whomever can deliver low latency video with “good enough” visual quality over existing network infrastructure will have the market’s attention.

Convenience Impact on Frequency of Business Interactions
Workflow integration and personal frustration also play a large part in the adoption of videoconferencing technologies into everyday business life.

We have written about the history of computing and the evolution from mainframes and minicomputers to microcomputers and personal computing. There is a close parallel here. High-quality videoconferencing rooms with dedicated MCUs and network bandwidth, guarded by reservation systems and stern administrative staff, remind us of glass-room datacenters. There are a mid-tier of video-enabled conference rooms with merely expensive videoconferencing systems; they are more accessible, but kind of like desktop PCs.

The market is clearly heading toward usage models more in tune with mobility and distributed workforce demographics. The future is not about bringing your own device to work but instead bringing your work to any device.

This means that videoconferencing must be embedded seamlessly in an employee’s most-used tools and workflow processes. It cannot be an impediment to workflow. Microsoft embedded Lync into their Office suite for this reason (if there is one thing Microsoft is indisputably good at, it is personal productivity).

This shift means that videoconferences will no longer be just two or three reserved rooms collaborating within a corporate firewall. The current dominant usage model of two or three groups of people meeting via dedicated rooms is giving way to a model where most participants will access a videoconference through their own personal devices. The result is that videoconferencing systems with expensive, limited capacity multipoint control units (MCUs) are now bottlenecks, not enablers. MCUs were designed for fixed infrastructure in closed and protected environments. MCUs are not appropriate for fluid, scalable video communications.

If videoconferencing is to attain widespread use on a daily basis, as part of normal workflows, then it must be enabled through frictionless use of common scheduling tools and by the devices people use to do their work.
VidyoConferencing Platform

Vidyo labels its entire solution suite under the name of the server-side application that runs its video conference sessions: VidyoConferencing. See Figure 1 for a breakdown of the major components of the VidyoConferencing platform. Not shown are higher-level software products layered above VidyoWorks and VidyoPortal.

Figure 1 is not a network diagram. It shows that every nearly every product in Vidyo’s portfolio either talks directly to a VidyoRouter or talks to a gateway (either a dedicated appliance or a software application running on a server) that talks to a VidyoRouter. The VidyoConferencing platform is, practically speaking, a disaggregated MCU.

Figure 1: VidyoConferencing Platform Components

The notable exception in Vidyo’s portfolio is their recently announced VidyoH2O for Google+ Hangouts. A bit of a stand-alone chimera, it is a point product aimed at bridging legacy H.323-based videoconferencing room systems, corporate SIP IP-based PBX dial-in audio calls, and as the name suggests, Google+ Hangouts.

The VidyoConferencing platform is an extensible architecture that allows their videoconferencing solution to be integrated with major videoconferencing-enabled workflow tools like Microsoft Lync, Mitel MiCollab, IBM Sametime, and Adobe Connect.

Vidyo has also created a set of bridges between different conferencing protocol domains under the name VidyoGateway. These bridges can take the form of either a discrete physical or a virtual intermediate translation appliance.
We divide Vidyo’s technology and products into three clusters:

1. **Endpoint technologies**, including compression coding and communications protocols,
2. **Adaptive real-time routing** of audio and video based on endpoint capabilities and real-time network performance analysis, and
3. **Workflow and back-end solution integration**.

**Endpoint Technologies**

Vidyo’s home turf is, as mentioned, scalable video coding (SVC). SVC was added to the H.264/MPEG-4 advanced video codec (AVC) standard in 2008. Vidyo has parlayed their early insights into SVC techniques into a strong future position in H.265 high efficiency video coding (HEVC) and Google’s open-sourced WebRTC. (Vidyo has been granted 47 patents worldwide, as of January, 2014.) Their insight into maintaining the highest quality video at each individual endpoint is this: they do not allow endpoints with a weaker network connection or compute resources to degrade the videoconferencing experience for other endpoints. Vidyo also creates the highest possible image at each endpoint, enabling devices with good connections and high resolution displays (stationary or mobile) to achieve better video resolution than many legacy videoconferencing rooms.

In addition to using available endpoint resources to perform media codec functions efficiently and communicate with VidyoRouter, these endpoint apps perform a couple of other high value functions. Endpoints help VidyoRouter determine the quality of a connection between an endpoint and VidyoRouter from moment to moment. Vidyo endpoint apps speak directly to VidyoRouter, such as VidyoDesktop (PC, Mac, and Linux), VidyoMobile (iOS & Android phones and tablets) and VidyoWeb (plug-ins for common browsers).

Vidyo’s endpoint apps also offload a lot of what would have happened at an expensive MCU onto the video endpoints. They send VidyoRouter metadata describing the endpoint’s display resolution as well as its graphics and CPU processing capabilities. Not only do endpoints all encode and decode one standard format, endpoints composite multiple streams of incoming video themselves using their local graphics capability.

**Real-Time Routing Through Disaggregation**

The real thought behind the VidyoConferencing platform architecture becomes apparent when we examine VidyoRouter. Vidyo pried the protocol conversion out of an expensive videoconferencing MCU and put it in special purpose outboard appliances or embedded their native protocol in endpoints. They then assumed that every endpoint will use SVC coding (H.264 annex G, in the future other formats) or that they would build gateways for devices using different video formats (high correlation between needing to convert video formats and needing a different protocol), and they pulled video transcoding out of the MCU. Then Vidyo assumed that Moore’s Law would continue to cram better graphics and display management capabilities into endpoints over time (a safe
assumption) and offloaded graphics and display management functions from the MCU to the endpoints.

What’s left? Routing. Vidyo has disaggregated an expensive MCU into a commodity smart router. No transcoding and no protocol conversion means that VidyoRouter can concentrate on being a really good, low-latency application level router. Vidyo disaggregated the legacy videoconferencing MCU to create their efficient, extensible, and low-cost VidyoConferencing architecture. This flexible architecture allows conferences to flow across multiple VidyoRouters with very little latency added at each router. This lets VidyoRouters analyze and localize videoconference media streams to a single router for better bandwidth use. It also allows single videoconferences to span multiple VidyoRouters and therefore scale to very large participant counts.

While CPUs are not very efficient at converting media formats (endpoint GPUs are), CPUs are very good at high-level logic and making real-time decisions about where to send which packets. CPUs are so good that Vidyo sells a virtual appliance version of VidyoRouter that runs in VMware virtual machines. Customers don’t even need dedicated hardware to run VidyoRouter. Any virtualized or converged server with spare capacity can run VidyoRouter VE software in a VMware ESXi instance (and in an average enterprise datacenter most are underutilized). The baseline VidyoRouter appliance and VidyoRouter VE software instance have equivalent video stream handing capabilities of 100 concurrent high definition connections per VidyoRouter instance.

**Workflow and Solution Integration**

VidyoConferencing integration into partner products and end user workflow starts with the ability to embed hyperlinks or other deeper level APIs. This integration streamlines inviting guests who are using a standard web browsers, web-enabled productivity tools (Microsoft Lync, Mitel MiCollab, IBM Sametime, and Adobe Connect), instant messaging clients (Google and Microsoft including Skype), and audio-only endpoints (SIP IP-based PBX dial-in and smartphone audio calls), into a videoconferencing session without a complex registration process on their part. This software can also be used in IIoT solutions to form video streaming links to embedded video devices, such as connecting hospital beds to nursing stations for remote patient monitoring.

For developer partners who would like to embed VidyoConferencing as a part of their own solutions, VidyoWorks APIs provide access to Vidyo’s physical and virtual appliances. While APIs are not in themselves transformative, Vidyo’s disaggregated architecture is. Vidyo is working with partners to improve the video quality and flexibility of a range of IIoT real-time video streaming applications while reducing costs.

Completing the portfolio, where would workflow and solution integration be without data retention and compliance interfaces? VidyoReplay can be used for archiving live video communications sessions. VidyoReplay has a control interface to record and play back stored conference sessions. Vidyo provides API access to VidyoReplay’s internal video database. At a high level, this API provides software developer hooks for recording control, content management and webcasting. Kaltura has already implemented a data
ingress solution to pull video from VidyoReplay and store it in Kaltura’s larger content management system (CMS).

Security is critical for organizational workflow, personal use of conferencing solutions, and for Internet of Things applications. Vidyo encrypts all point-to-point communications links in the VidyoConferencing platform by default – media is encrypted using SRTP, signaling is encrypted using TLS, and all other transactions are encrypted using HTTPS/SSL. Because encryption is point-to-point, media sessions are not encrypted within VidyoRouter instances, so organizations may choose to control physical access to their VidyoRouter instances behind their own firewall or they can run hosted instances on remote hardware infrastructure. Access to conference sessions is secured end-to-end, including control of directory access and functions, which controls access to VidyoConferencing sessions. The VidyoReplay file system is currently not encrypted, so customers who implement VidyoReplay but wish to maintain file system security may choose to locate VidyoReplay appliances behind their firewall.

**Video and the Industrial Internet of Things**

We have written about the Internet of Things (IoT) and the Industrial IoT (IIoT) from a number of perspectives. Industrial artifacts that stream video in real-time have a diverse set of communications requirements that all summarize into a common observation. There is a high probability that latency, bandwidth, and packet integrity will be negatively affected by constraints on cost, power, mobility, materials, and other design choices made to account for environmental constraints, like long-distance wireless connections, noise (sound and vibration, electromagnetic interference, *etc.*), and harsh environments (thermal, chemical, radiological, biological, *etc.*).

The net effect of a lot (think millions to possibly many billions) of things streaming real-time video with poor packet integrity is that current videoconferencing and web-based video entertainment technologies are inadequate to turn the packets that are received into a million (or billion) low latency, good enough video streams.

The primary goal for most of these systems is to determine that something “interesting” has happened in the field of view of a video sensor. Here again, low latency and good enough video quality are enabling system attributes. Once something interesting happens in a video stream, then human operators or cloud-based analytics systems can decide on a course of action (and maybe send more video-enabled things into the fray to get a better look).

The remote monitoring and control systems for video-emitting IIoT things will require video stream ingress and routing capabilities far beyond the requirements of enterprise real-time video communications. The IIoT will be driven by scale and governed by cost.
Call to Action

Vidyo's disaggregated VidyoRouter products enable a low latency, real-time, multi-stream communications user experience at the highest quality supported by each individual endpoint in a video conversation. The VidyoConferencing system takes into account each endpoint's capabilities and moment-to-moment network performance. The result is hard to describe in words. Individual participants’ experiences differ based on their endpoint and connection, but the experience is remarkably good for each participant—even those on poor connections.

VidyoRouter is designed to leverage widely available, industry standard server technology and is available as an optimized appliance or in a virtual edition for VMware platform. VidyoRouter's primary endpoints are Vidyo apps running on popular consumer software ecosystems and commodity consumer hardware. For endpoints that do not require a VidyoGateway (typically legacy room-based conferencing systems), it is hard to envision a way to achieve lower hardware cost of acquisition and operation.

Enterprises interested in expanding use of videoconferencing beyond room-based systems should evaluate VidyoDesktop (PC, Mac and Linux), VidyoMobile (phones and tablets), VidyoSlate (tablets), and VidyoWeb (browser plug-ins)—Vidyo's front-end interfaces to VidyoRouter. These tools integrate with organizational workflows and with the devices people want to use.

For anyone evaluating video systems for IIoT deployments, we suggest contacting Vidyo to get a deeper understanding of their technologies, as well as how their current and future product portfolio can contribute to real-time decisions given poor quality and low bandwidth network and communications links.
publicly release the results of any revision to these forecasts and forward-looking statements in light of new information or future events.

©2014 Moor Insights & Strategy. Company and product names are used for informational purposes only and may be trademarks of their respective owners.