

Cloud and Traditional Videoconferencing Technology for Telemedicine and Distance Learning

Wei-Li Liu, MBA, Kai Zhang, MS, Craig Locatis, PhD,
and Michael Ackerman, PhD

Office of High Performance Computing and Communications,
National Library of Medicine, Bethesda, Maryland.

Abstract

Introduction: Cloud-based videoconferencing versus traditional systems are described for possible use in telemedicine and distance learning. **Materials and Methods:** Differences between traditional and cloud-based videoconferencing systems are examined, and the methods for identifying and testing systems are explained. Findings are presented characterizing the cloud conferencing genre and its attributes versus traditional H.323 conferencing. **Results:** Because the technology is rapidly evolving and needs to be evaluated in reference to local needs, it is strongly recommended that this or other reviews not be considered substitutes for personal hands-on experience. **Conclusions:** This review identifies key attributes of the technology that can be used to appraise the relevance of cloud conferencing technology and to determine whether migration from traditional technology to a cloud environment is warranted. An evaluation template is provided for assessing systems appropriateness.

Key words: technology, cloud computing, telecommunications, telemedicine, distance learning

Introduction

The generic, distinguishing technology characteristics of cloud-based videoconferencing identified in this review provide a framework for assessing traditional versus cloud videoconferencing in relation to local needs. The review begins by providing background information about cloud computing, the contrasting approaches to communication by cloud and traditional videoconferencing systems, and concepts related to cloud conferencing. Next, criteria used to identify the characteristics of cloud videoconferencing systems and the methodology used in this review are described. Finally, common attributes of cloud videoconferencing systems are identified based on the review that can be used (1) to document and appraise specific products using an evaluation template and (2) to determine if migration from traditional to cloud systems is justified.

BACKGROUND

The National Institute of Standards and Technology defines cloud computing as "...a model for enabling ubiquitous, convenient, on-

demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction."¹ Cloud computing represents a means of delivering computing services via a distributed network like the Internet. Although there are local computers and software, the applications performing tasks (such as word processing or database management) and the data they generate reside on computers somewhere on the network. Applications are accessed on the network as needed, rather than installing duplicate programs locally on individual workstations, and the data they generate are also stored on the network, which facilitates sharing. In cloud videoconferencing, local software and hardware (a computer with a camera) are still needed, but the videoconferencing applications managing conferences and the data they generate reside somewhere on the network (i.e., in the "cloud").

Terms associated with cloud conferencing are Web real-time communication (WebRTC), unified communication, and video as a service. WebRTC refers to an open source research and development effort aimed at incorporating videoconferencing directly into browsers.² Because browsers share video across the Internet, WebRTC involves cloud computing, but only by incorporating video into browser architecture. The term unified communications refers to integrating real-time communication (telephony and videoconferencing) with other network data resources, such as interactive whiteboards, and non-real-time communication, such as e-mail and voicemail, so that voicemail might be accessed as e-mail, or vice versa.³ Cloud conferencing technologies attempt to integrate videoconferencing with other applications, at least those that are real time, and subscribe to the idea of unifying communication to some extent. Video as a service is a term describing accessing network videoconferencing services located in the cloud, usually paid for by subscription.⁴

Cloud videoconferencing technology is best understood, in contrast to "traditional" videoconferencing. Cloud technology's most distinguishing feature is that client conferencing software installed on local computing devices accesses videoconferencing software on servers managing communication. Servers take advantage of the camera and audio resources that are built into or added onto the client devices, such as a laptop's built-in camera and microphone or desktop's external USB camera and microphone.

Traditional videoconferencing is accomplished through the use of appliances usually permanently installed in a room or placed on moveable carts.⁵ A common communication standard (H.323) ensures that end-point appliances interoperate to exchange audio and

SYSTEM	DEVELOPER
SeeVough	SeeVough
Vidyo	Vidyo
Jabber	Cisco
Scopia	Avaya
Fuze	Fuze
Real Presence	Polycom

video at transfer rates from 128 kilobits/s (Kbps) to 4 megabits/s (Mbps), with the latter providing 1920 × 1080 pixels of high-definition progressive video incorporating the H.264 video standard. Typical appliances include pan, tilt, and zoom (PTZ) cameras that can be locally or remotely controlled to show the entire room or persons within it and omnidirectional microphones that can detect audio from long distances, with built-in echo cancellation so that sounds picked up from the speakers do not produce feedback. Every end point usually requires an appliance costing about \$4,000, although occasionally a vendor might offer inexpensive software versions of its products for use with Webcams and computers.

Materials and Methods

Cloud conferencing systems reviewed are listed in Table 1. All were part of the Internet2 research and education network’s Test Drive Program, allowing Internet2 member institutions to try out products with Internet2 staff. The systems should be considered representative, but not inclusive, of those available. As commercial company members of Internet2, they represent some of the major corporate developers and companies having a large footprint of the higher education market, and several exhibit at telemedicine meetings as well. Because the software at Internet2 lagged behind the latest development cycle, tests ultimately were carried out with products hosted directly on developer servers.

Tests were done to determine (1) functional differences between cloud and traditional videoconferencing, (2) whether there were common core features and significant differences in how they were implemented, and (3) what circumstances, if any, might warrant migration to cloud conferencing technology. The cloud-based products reviewed were tested from the last half of 2012 through the first quarter of 2014. Each system’s major components as represented on its main menu, toolbar, or tabs were tested at least twice: once in a point-to-point and once in a multipoint conference. These point-to-point and multipoint tests were repeated several months apart with different users. Four or five end points participated in multipoint tests. Their common features were exercised directly, with the exception of a Webcasting capability, unique to just two systems (Vidyo

[Hackensack, NJ] and Avaya [Santa Clara, CA] Scopia®). An assessment protocol (Fig. 1) was used to document system features.

Results

The review identified core cloud system attributes including video and audio encoding, multipoint conferencing, operating system and computing platform requirements, interoperability, security, content sharing, user interfaces, archiving, and Webcasting. Differences between traditional and cloud systems for each of these attributes are summarized in Table 2.

VIDEO ENCODING

All cloud systems support H.264/MPEG-4 Advanced Video Coding (AVC),⁶ providing the same video quality as MPEG-2 at less than half of the bandwidth, delivering video at rates from 40 Kbps to 10 Mbps for resolutions ranging from 176 × 144 to 1920 × 1080 pixels. Some systems support up to 1080 pixel resolution, whereas all others only support up to 720 pixels. Some support H.264 Scalable Video Coding⁷, an extension of AVC to provide even better quality video. Video quality varies depending upon the cloud system, camera, and type of

1. **Video.** Does the system provide sufficient video resolution at bitrates appropriate for the anticipated telemedicine or distance learning application?
2. **Audio.** Will the built-in audio capabilities of the system suffice or will additional echo cancellation, microphone, or other audio components be required?
3. **Multipoint conferencing and scaling.** To what extent is multipoint and/or simultaneous conferencing needed? Is scalability an issue? Are additional conferences and user capacity needed? Do conferences need to be deployed on multiple devices?
4. **Platform and hardware.** Is the client software available for the operating systems one has deployed? Is pan tilt and zoom camera control required? Does the system support it?
5. **Interoperability.** Is it important for the software to interoperate with more traditional H.323 hardware systems? Are apps available for the kinds of mobile devices one may anticipate using? Is phone bridging needed?
6. **Collaboration and content.** Will less than full application sharing suffice for the system’s anticipated teleconsultation or distance learning application? Will the mark up capabilities of a whiteboard be needed or only simple screen sharing? Are chat capabilities desirable?
7. **Deployment and security.** Does the system provide sufficient encryption, authentication, and password protection for one’s anticipated application and are port settings and firewall requirements realizable in one’s network environment?
8. **Interface.** Is the client software intuitive and easy to use? Does it offer affordances at a level of complexity that are appropriate for its end users? Are the anticipated users the general population or someone trained to help others?
9. **Archiving and webcasting.** Are archiving or simultaneous streaming of conferences needed and, if so, are special file formats and players acceptable or must the files be in more common video formats? Is the quality of the video archived acceptable?
10. **Cost and licensing.** Are the prices and terms reasonable and justifiable considering the system’s planned deployment and application? Does the overall system offer appropriate capabilities at costs justifying replacement of more traditional conferencing technology?

Fig. 1. Cloud conferencing evaluation template.

Table 2. Attribute Comparison

ATTRIBUTES	TRADITIONAL SYSTEMS	CLOUD SYSTEMS
Video quality	Built-in video resolution	Potential windowing/resolution issues
Audio quality	Built-in audio quality control	Potential echo/feedback and loudness issues
Multipoint scalability	Difficult, may need special devices	Built-in multipoint capability, theoretically unlimited
OS/hardware	Built-in OS and PTZ camera control	Need to appraise computer capabilities compatibility, work around limited camera control
Interoperability	High, follows H.323 standard	Low, not standardized
Content sharing	Built-in but limited	Built-in with multiple sharing features
Security/network management	Network management/coordination required	Little network management needed, none for some systems
Interface	Difficult, assumes trained users	Easy, intuitive graphical interfaces intended for anyone
Archiving and Webcasting	Limited	Built-in/easy for some systems
Cost	Expensive, especially for end-point appliances	Expensive for servers, cheaper for clients

OS, operating system; PTZ, pan, tilt, and zoom..

computer display, and because the video is displayed on the on the computer's screen, the video window's size varies depending on the number of end points connected. Some systems allow a conference administrator to control window layout, whereas others do not, but such control does not guarantee identical display at end points because local computers might be configured for different resolutions (e.g., 800 × 600 or 1024 × 768 pixels).

Although cloud conferencing systems use the same H.264 video codec as the latest generation high-definition traditional conferencing systems (720 pixels or 1080 pixels), variations in display resolution, window sizes, and camera quality may produce video inferior to traditional systems that include high-quality cameras and display video on television monitors, especially when conferences involve multiple end points. This limitation can be overcome by connecting additional computer monitors increasing display space, but adds expense. Acceptable image quality depends on specific applications and may be a greater issue in teledermatology than in telepsychiatry, for example. Window size becomes more important if classrooms rather than individuals are participating in distance learning. Theoretically, cloud systems have video quality equal to traditional systems. In practice, they may not.

AUDIO ENCODING

Audio codecs in cloud systems include the G.7xx series (G.711, G.721, G.726), speedx, and MPEG-4 AAC audio standards. All systems allow volume adjustment and muting and have built-in echo cancellation. Echo cancellation built into cloud software still may be insufficient, and an external hardware echo cancellation device or headsets may be needed. Appropriate audio quality also will depend on specific telemedicine and distant learning applications, for example, whether heart sounds, breathing, or speech patterns will need analysis, or whether students will be located in groups in large rooms.

Although traditional H.323 videoconferencing systems use the same audio codecs, they output audio to external speakers and use multidirectional external microphones and more powerful built-in echo cancellers, providing inherently more audio quality than cloud systems, which use the audio inputs and outputs of their computer hosts. This cloud-based default audio may be adequate for many applications, especially if conferences only involve individuals, but additional external speakers, microphones, and echo cancellation devices might needed for applications involving groups, thus increasing costs.

MULTIPOINT CONFERENCING AND SCALING

Servers manage all cloud conferencing, even if only two end points communicate. The servers are designed to accommodate multiple simultaneous conferences and are theoretically limited only by computer capability and network capacity. Traditional videoconferencing systems can call each other directly if there are only two end points, but a device called a multipoint control unit (MCU) is required to bridge multiple connections, an extra expense that escalates depending on the number of connections. MCUs accommodating four or eight connections can be added into a traditional appliance, but a separate MCU device may be required if more than eight connections are needed. Cloud systems are intended for multipoint conferences and large deployments; traditional systems are geared for point-to-point communication with multipoint conferences as an option.

OPERATING SYSTEM AND HARDWARE

Cloud conferencing systems mainly run under Microsoft (Redmond, WA) Windows System 7 and Mac OS (Apple, Cupertino, CA), whereas Linux freeware is supported by just a few. All have applications for mobile platforms. Although they utilize cameras and microphones built into devices, they can be configured to accommodate external audio and video inputs and outputs. Most systems

reviewed allow control of remote external PTZ cameras when interoperating with traditional H.323 systems, a feature that may be particularly important in remote patient examinations, especially for telemedicine programs already using more traditional technology. None of the systems reviewed natively controlled PTZ cameras attached to cloud-based clients, although there are other cloud systems offering some control. Autotracking PTZ cameras that follow an instructor's voice or movement may suffice for some distance learning applications, but add expense, and may be inadequate for telemedicine when examining a specific area of a patient is needed. The only practical option for cloud systems lacking remote camera control may be instructing patients how to position themselves. Because traditional systems use self-contained appliances with all the necessary hardware and software installed, operating system and device requirements are seldom issues.

INTEROPERABILITY

Most cloud systems use SIP, the session initiation protocol (and sometimes additional H.323 protocols), to communicate, but they do not interoperate with other cloud systems.⁸ The H.323 standard ensures traditional systems interoperate, not only to exchange audio and video, but other features such as PTZ camera control and content sharing. Although all cloud systems tested have applications for mobile devices having Google (Mountain View, CA) Android™ or Apple iOS operating systems, only a few provide phone bridging allowing audio-only participation by phone.

COLLABORATION AND CONTENT SHARING

Chat, whiteboard, and screen sharing are three common cloud collaboration features. Chat allows text communication, whereas whiteboard allows free-hand writing and drawing and sharing imported digital images that can be pointed to or marked up. Screen sharing allows conference participants to view a given remote site's computer screen (i.e., desktop and its slide, browser, and other applications) but is not true application sharing, where all participants can access a remote computer's actual applications to generate or edit content. Participant video windows shrink to ensure content legibility, but at a cost of making it more difficult for participants to see each other. Some systems provide more flexibility, allowing users to display content in a separate window and manipulate video and content window size.

Traditional H.323 videoconferencing systems have more limited presentation capabilities and lack chat and whiteboards. There is, however, a sub-standard (H.239) that allows users to connect computers to appliances and transmit content as video. One common work-around is to establish two independent connections at each point: one for sharing audio and video between appliances and another between computers for sharing content. Cloud systems, being computer-based, are superior for content sharing, but managing display real estate can be a problem.

SECURITY AND NETWORK MANAGEMENT

Some cloud systems use plug-in software to display video within a browser, but most use separate client software. Products having their

own clients usually require more network port management and more coordination with network administrators, whereas browser-based systems use the standard browser port 80 usually open on most networks.⁹ All systems require authentication to communicate with servers and provide encryption and password protection, with the latter being the option of the conference initiator. Although cloud and traditional systems share many of the same security mechanisms for authentication, encryption, and password protection, traditional systems require more port utilization than browser-based cloud systems, making them harder to make them comply with institutional network security policies.

USER INTERFACE

All cloud systems have very intuitive, graphical user interfaces, but some are easier to use than others. One, for example, defaults to a very basic interface with minimal tools for video, audio, and content sharing while allowing access to a toolbox revealing additional features only if needed. Browser-based cloud systems also seem easier to use because they draw on features of the standard Web interface. Traditional systems require using the appliance remote control to page through menus and enter alphanumeric data to make calls and configure systems. Their design assumes greater technical competence, and the remote controls are clumsy for entering text. Cloud systems are designed for general users, whereas traditional systems are geared for trained operators.

ARCHIVING AND WEBCASTING

Several cloud systems provide conferencing archive/recording capability for later viewing on demand, whereas only a few support live conference Webcasting, allowing nonconference participants to view the interaction in real time. Archives are more compressed and viewed in smaller windows than the original conferences, so quality tends to be poorer. Some systems require users to download developer-supplied viewer or player software to view archived files, whereas others record conferences in common video formats such as Apple's Quicktime or Windows Media.

Traditional H.323 systems lack built-in archiving and streaming capabilities. Users have to run system video and audio outputs to a computer or other device that is configured to accept these inputs and digitally record them. If content is presented by establishing second connections instead of sending it as video, then additional software is required to capture screen content and synchronize it with the audio and video. An entirely separate system is needed for Webcasting.

COSTS AND LICENSING

Although a few cloud systems have flexible pricing accommodating a small number of end points, most are priced for more enterprise-wide, large deployments. Cloud server software can minimally cost upwards of \$20,000 and have annual service and maintenance fees of several thousand dollars. Many developers assume software will be used for education and base their prices on the number of "seats" or end-user clients issued. These costs are reasonable on a per seat basis considering they allow hosting several simultaneous conferences involving many users. Moreover, traditional hardware-based systems with similar capabilities are even more expensive. Still, costs are harder

to justify for users not needing so much capability. If a low-capacity MCU can suffice, for example, one accommodating four end points, traditional MCU costs can be comparable to cloud servers, although end-point appliances are more expensive. The cost advantages for cloud end points is mitigated, however, if additional hardware, such as echo cancellation, PTZ cameras, and monitors, is required.

Discussion and Conclusions

Cloud conferencing systems represent a newer, alternative technology and differ significantly from traditional conferencing systems. Software implementations of videoconferencing have several theoretical benefits. First, client software can be more widely and immediately deployed because it is installed on computers and other devices already in use, provided, of course, the machines have sufficient computing power to run the client software, have the operating system for which the software was designed, and have video and audio capabilities. Second, videoconferencing can be made increasingly mobile because client software can be installed on laptops, tablet computers, smartphones, and other devices equipped with cameras and one or more forms of wireless technology. Although traditional videoconferencing appliances can be connected to wireless antennae and moved about, their inherent size and that of cameras and monitors to which they connect limit mobility. Third, cloud technology is more scalable, limited only by the capabilities of the computers on which the server software is installed, the capacity of the networks used for communication, and licensing costs. Because traditional videoconferencing relies on MCUs for multipoint conferencing and various models have an upper limit connection capacity (e.g., 4, 8, 16, etc.), each of which is priced at multiples of single units, scalability is an issue. Finally, the use of client software on existing computing platforms introduces economies of scale and the potential to reach more end users directly. The intended target user population is anyone working anywhere. In contrast, traditional videoconferencing units are costly, intended for institutional use in exam rooms, conference rooms, or classrooms, and usually require trained users and technical support.

Cloud conferencing technologies are improved with each new software release, and certain technology limitations identified here may be rectified in the future. For example, code can be added for remote control of PTZ cameras or to drive other telemedicine devices. Still, theoretical cloud advantages must be balanced by practical concerns about the current quality of cloud system video and audio, computing requirements, possible need for additional hardware, possible limited remote camera control, lack of interoperability, display restrictions, and whether archiving provided by some systems is needed. Because cloud systems are usually priced for enterprise deployment involving many end points, they may not be cost-effective for modest applications except, perhaps, if used as a service. They may be more appropriate for large-scale deployments such as monitoring patients at many sites or providing education to different locations. The pricing terms and features of the cloud systems reviewed (e.g., screen sharing) suggests they are currently most suited

for education, especially by individuals sitting at their own desktops or laptops rather than for classes in auditoriums or rooms.

Acknowledgments

This review was supported by the National Institutes of Health/National Library of Medicine intramural research program. The authors acknowledge the contributions of Willis Nguyen, a student intern who helped install and test the programs.

Disclosure Statement

No competing financial interests exist.

REFERENCES

1. Mell P, Grance T. *The NIST definition of cloud computing*. Version 15.10.07. Gaithersburg, MD: Information Technology Laboratory, National Institute of Standards and Technology, 2009.
2. Rodriguez P, Cervifio J, Trajkovska I, Salvachua J. Advanced videoconferencing services based on webrtc. *IADIS International Conferences Web Based Communities and Social Media 2012 and Collaborative Technologies*. Prague, Czech Republic: IADIS, 2012;180–184.
3. Evans D. An introduction to unified communication: Challenges and opportunities. *Aslib Proc New Inf Perspect* 2004;56:308–314.
4. Rodriguez P, Gallego D, Cervino J, Escribano F, Quemada J, Salvachua J. Vaas: Videoconference as a service. *5th International Conference on Collaborative Computing: Networking, Applications and Worksharing, 2009*. CollaborateCom. Washington, DC: IEEE Xplore, 2009;1–11.
5. Liu W, Zhang K, Locatis C, Ackerman M. Internet-based videoconferencing coder/decoders and tools for telemedicine. *Telemed J E Health* 2011;17:358–362.
6. Wiegand T, Sullivan GJ, Bjontegaard G, Luthra A. Overview of the H.264/AVC video coding standard. *IEEE Trans Circuits Syst Video Technol* 2003;13:560–576.
7. Hewage C, Karim H, Worrall S, Dogan S, Kondoz A. Comparison of stereo video coding support in mpeg-4 mac, h.264/avc and h.264/svc. *Proceedings of VIE2007*. London: Institution of Engineering and Technology, 2007.
8. Ho J, Hu J, Steenkiste P. A conference gateway supporting interoperability between SIP and H.323. *Proceedings of the Ninth ACM International Conference on Multimedia*. New York: ACM, 2001;421–430.
9. Mahoney MV, Chan PK. Learning rules for anomaly detection of hostile network traffic. *Third IEEE International Conference on Data Mining, 2003*. ICDM 2003. New York: IEEE, 2003;601–604.

Address correspondence to:

Wei-Li Liu, MBA

Office of High Performance Computing and Communications

National Library of Medicine

8600 Rockville Pike

Bethesda, MD 20894

E-mail: wliu@mail.nih.gov

Received: June 17, 2014

Revised: July 7, 2014

Accepted: July 23, 2014