



WHITE PAPER

An Introduction to the Basics of Video Conferencing

Introduction

In the next few years we shall see explosive growth in the use of video conferencing (VC) as a fundamental tool for businesses to enhance communication and collaboration between employees, partners and customers. The technology has developed considerably from early adopters to its current form of mass market roll-out. It's anticipated that nearly half of information workers will have some type of personal video solution in 2016, up from just 15% today*. With VC becoming a core component of IT infrastructure that enables communication and collaboration, businesses will be looking to providers of telephony, business applications and network infrastructure services to include this capability as part of their offering.

This report will examine the basic components of the technology, considerations for deploying VC solutions, and will introduce the Polycom® RealPresence™ software platform to readers.

What is Video Conferencing and how does it work?

To set the foundations for future elaboration, at the simplest level, a video conference is an online meeting (or a meeting over distance) that takes place between two parties, where each participant can see an image of the other, and where both parties are able to speak and listen to the other participants in real time. The components necessary to make this happen include:-

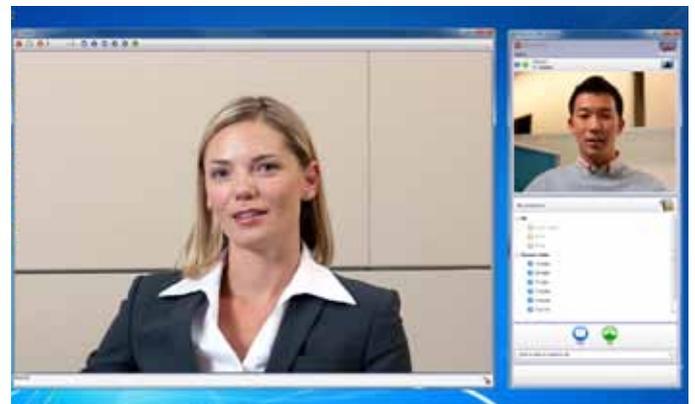
- a microphone, webcam and speakers
- a display
- a software program that captures the voice stream from the microphone, encodes it, transmits to the other participant, and simultaneously decodes the digital voice stream being received from the remote participant in the video conference (most commonly referred to as a "Codec".)
- a software program that bridges both parties together across a digital connection, managing the exchange of voice and video between participants. At either end of the connection, the video and voice traffic is combined and delivered to each participant in the form of a real-time video image and audio stream
- an optional management tool for the scheduling of VC sessions

At a slightly more advanced level, it is also possible to provide the ability to share content from a device during a video call. The quality and type of content that can be shared depends on the rate of data exchange during the call.

Terminology used by VC users to describe the process of dialling into and participating in a virtual meeting is known as "joining a bridge". Different virtual meeting rooms are assigned unique "bridge numbers", and users join a video call by "dialling a bridge number".

Point-to-Point Video Conferencing

Video-enabled meetings happen in two distinct ways: either point-to-point or with multi-point. In point-to-point, the simplest scenario is where one person or group is connected to another. The physical components (i.e. microphone and camera) that enable the meeting to take place are often integrated in to desktop computing solutions like a laptop or tablet, or can be combined into dedicated, room-based hardware solutions.



An example of point-to-point video conferencing

Where desktop solutions tend to be used by individuals, room-based solutions utilise dedicated VC technology where groups of people can be seen, heard and can naturally participate in the meeting.

Multi-point Video Conferencing

In multi-point video calls, three or more locations are connected together, where all participants can see and hear each other, as well as see any content being shared during the meeting.



A use-case scenario of multi-point video conferencing

Source: *Forrester - Preparing for Uneven Corporate Adoption of Video Communications, May 9, 2011

In this scenario, digital information streams of voice, video and content are processed by a central, independent software program. Combining the individual participant's video and voice traffic, the program re-sends a collective data stream back to meeting participants in the form of real-time audio and video imagery.

Individuals can participate in a meeting in an "audio only" mode, or combine audio with video images of the meeting on screen. Depending upon the technical capability of the VC system being used, images seen by participants are either classified as "Active Speaker" or "Continuous Presence".

In "Active Speaker" mode, the screen only provides an image of the person that is speaking at any point in time. In more advanced solutions with "Continuous Presence" mode, the bridge divides the image on the screen into a number of different areas. The person speaking at any point in time is presented in a large central area, and other meeting participants are shown displayed around the central image. The "Continuous Presence" mode thus allows meeting participants to view and interact with all meeting participants in a 'virtual meeting room'.

The software program which creates the "virtual meeting room" and the digital processing hardware on which it resides, is often called a Video Bridge, or "bridge", for short. Another term for a bridge which is often used is a VC "multi-point control unit" or "MCU".

Whereas point-to-point VC is relatively simple, the creation and management of multi-point video conferences can be complex. An MCU must be able to create, control and facilitate multiple simultaneous live VC meetings. A further complexity is added when different locations may connect to the meeting over digital or analogue streams at different speeds, with different data transport and signalling protocols employed to facilitate the communication.

To link these users into a common, virtual meeting, the MCU must therefore be able to understand and translate between several different protocols (i.e. H.264 for communication over IP, and H.263 for ISDN). The MCU will also allow those joining the video bridge to do so at the highest speed and the best possible quality that their individual system can support. Although there are two separate processes taking place here, this is often jointly referred to as "Transcoding".

It is important to note that not all bridges provide such transcoding capability, and failure to do this can seriously impact the quality and experience of video calls. When transcoding is not provided and users dial into a bridge over a range of different connection speeds, it is possible that the bridge may only be able to support the video meeting by establishing the connections at the lowest common denominator. To illustrate the negative effect of this, consider a meeting that takes place with most users joining the bridge from the high-speed corporate network, but where one or two individuals dial into the meeting from home on low-bandwidth DSL or ISDN. In this case the experience of the many corporate users is downgraded to the lowest common denominator of the home-users, potentially making the video call ineffective. Where effective transcoding is supported by the MCU, those on the corporate network will continue to enjoy HD video quality, while remote users receive quality commensurate with their connection speeds.

In summary, when an MCU is designed well, integrating easily with multiple vendors and allowing users to call in at the data rate and resolution they want or need to – the result is an easy, seamless experience for all users, allowing people to focus on the meeting, not the technology.

The Language of Video Conferencing

As VC technology has evolved, two main protocols have emerged to provide the signalling control for the establishment, control and termination of VC calls: SIP (Session Initiation Protocol) and H.323.

For the encoding and decoding of visual information, the industry is moving towards the industry standard known as H.264, which was developed to provide high-quality video at lower bandwidth over a wide range of networks and systems. An extension to the H.264 protocol is Scalable Video Coding (SVC), which is established to facilitate the enablement of VC on a wider range of devices, such as tablets and mobile phones.

Bridging Architecture and Functionality

As described above, the combination of software and the hardware that creates the virtual meeting rooms is called a "Video Bridge". Virtual meeting rooms are identified by their "bridge numbers". With multiple calls taking place simultaneously, software analyses all the different data streams coming into the bridge processors, and assigns data streams accordingly.

At the simplest level, the processing workload for bridges is dependent upon four factors:

- the number of locations that dial into each bridge
- the number of conferencing calls that each bridge must handle simultaneously
- the amount of data that is being received on each digital stream: higher resolutions of images and sound (i.e. High Definition) generate more data that needs to be processed.
- the degree of transcoding that the bridge must perform while handling calls being received at different connection speeds and utilising different protocols.

As the workload increases, each bridge must process more data. Performance can therefore be improved by increasing the number of Digital Signalling Processors (DSPs) utilised to decode and encode the digital streams entering and leaving MCUs. If the bridging function becomes overloaded, video and voice information may be lost, causing latency to be introduced into calls, both of which can degrade the video meeting experience.

Extra processing resource can be provided for the bridging function by either utilising a more powerful bridge (with a greater number of DSPs) or through a virtual software approach, where the software that controls the signalling function can operate independently of the physical hardware.

A conference call with an assigned conference number does not have to take place, or be processed by a dedicated piece of hardware. The call can be “virtualised”, and assigned to whatever physical bridge has the correct resource or capacity to handle the call. A virtualisation manager oversees which physical bridge has the capacity, and assigns incoming calls accordingly.

In extreme, but rare circumstances, the virtualisation manager may assign resources for a call across several different physical bridges that work in tandem together. Known as “auto-cascading”, the resources within the physical bridge can be instructed by the software to operate in a “Parent-Child” arrangement, with one bridge “owning” the conference call, and the others sharing the workload.

In the continuous presence mode of presentation, the bridge will automatically provide the screen templates in which the viewers will see the other meeting participants. The bridge can

also provide some administrative functionality for the call such as; assigning passwords to enter each meeting, and providing Interactive Voice Response (IVR) functionality, where call participants can be greeted and instructed by customised voice greetings.

Although most participants will actively dial into a VC meeting, the bridge can be programmed to automatically dial out to participating locations and automatically connect them in to a meeting. For example, the bridge could automatically wake up the cameras in remote meeting rooms, and link those meeting rooms into a prescheduled call. Participants of such a meeting would simply have to walk into the video room at the correct time, and join the meeting.

Video Call Management and Protocol Conversion

In order to build an architecture that scales, the software platform must be able to provide call signalling functionality, and dynamically manage the set-up and maintenance of a large number of video calls. The software architecture has to be capable of reconfiguring itself and its resources in real-time, so that these resources are used to their best ability. In addition, the software architecture has to understand the bandwidth requirements of each call being placed, the policy that is associated with each call (the prioritisation and importance of a call), and where the participants of a call are geographically located. By understanding this, the software platform can utilise local resources instead of redirecting data streams & call signalling to resources that are far away, an approach which would eat up large amounts of bandwidth on WAN links that are very costly.

The software platform should also be able to instantly detect any failure of hardware resources or loss of communication across infrastructure links, so that it can re-direct traffic and re-establish calls utilising alternative resources, without overly impacting video calls or their quality.

When systems on different customer premises try to join the same video call using devices which run different protocols (i.e. H.323, RTV or SIP), the VC platform must first perform protocol conversion to a common language so the infrastructure can understand and process information correctly. In other words, the software platform should provide intrinsic gateway functionality between devices that talk different languages.

The Polycom RealPresence Virtualisation Manager sits in front of the bridges, and interfaces between the outside world and the bridging resources. This optimises how incoming video calls are handled by virtual resources at its disposal. The Virtualisation Manager can apply business rules that help it place incoming meetings on bridges that make the most sense, either for capacity, geography, or other priority rules.

Let us consider three examples of this approach and see how it simplifies the process:

<p>Example A</p>	<p>Customer A in California wants to meet with Customer B in New York, Customer C in London and Customer D in Paris. The Customer has a video bridge in Denver and a video bridge in Paris and a virtualisation manager on a server in London. In this situation, the virtualisation management software would identify that two participants wanted to join the call from the U.S., and may, for example purposes, direct them to the resources on the Denver bridge. Likewise, the European participants may be directed to the Paris bridge, with overall control of the call being given to the Master Denver bridge. Under this scheme, large amounts of video data are not shipped across a transatlantic WAN, thereby potentially providing cost savings.</p>
<p>Example B</p>	<p>In the above example, the U.S. customers are using an H.264 based system, and in Europe they are using Microsoft® Lync™ enabled video conferencing based upon RTV. In this scenario, the virtualisation management software on the London server acts as a Gateway between Microsoft, and the U.S. Video resources, converts the Microsoft signalling, and establishes the whole call using the bridges in the U.S. and Paris.</p>
<p>Example C</p>	<p>In this example, the call is proceeding but the bridge in Denver suddenly stops functioning due to a fire in the data centre. The Virtualisation Manager in London detects this, and redirects the video traffic across the WAN link to the Paris bridge. Users connecting via H.323 simply redial to re-join the call, with the administration and management being performed seamlessly in the background. However, for SIP based calls there is an added advantage: the Polycom platform will detect the problem and reconnect the participants back into the call automatically, hopefully before the user has even noticed that there was a problem.</p>

Device Management

To enable large-scale deployment and management of VC solutions, the software platform provides for the management and maintenance of hardware infrastructure components through a separate functional area: The Device Manager.

The Device Manager can help dynamically provision devices and components of the VC infrastructure. Once component hardware is deployed within the network and its infrastructure, the Device Manager will monitor and help troubleshoot problems with these devices. When software updates are required, the Device Manager will help deploy them.

A significant contributing factor to the rise in demand for VC is because of the ease of use by which calls can be established by users. The scheduling and management of calls has become easy, through the creation of user-friendly scheduling portals, or via integration into Microsoft® Outlook™.

The Device Manager will also provide reporting, and comprehensive details of video calls, processing the information to evaluate current system usage, and expansion plans for the video network.

Security

Many organisations who have invested in VC will inevitably need to be able to assist mobile or home workers wanting to dial into their company network, and participate in video calls with colleagues. The software platform must therefore provide the capability to enable, and manage this.

Likewise, VC-enabled organisations will also want to use the technology to communicate with their partners and customers. This will only be possible if video traffic is able to securely traverse the firewalls from one customer to another. Firewall traversal is a particular challenge to video, as the data firewalls try to re-organise data packets. The implementation of a video firewall such as the VBP (H323) can eliminate this issue.

Content Management

Historically, the primary motivating factor for most companies has been to use VC as a way of saving business travel costs. Recently, organisations are beginning to understand that the benefits of VC can impact many different parts of an organisation including; training, marketing, education, compliance, internal communications, advertising, PR, to name just a few.

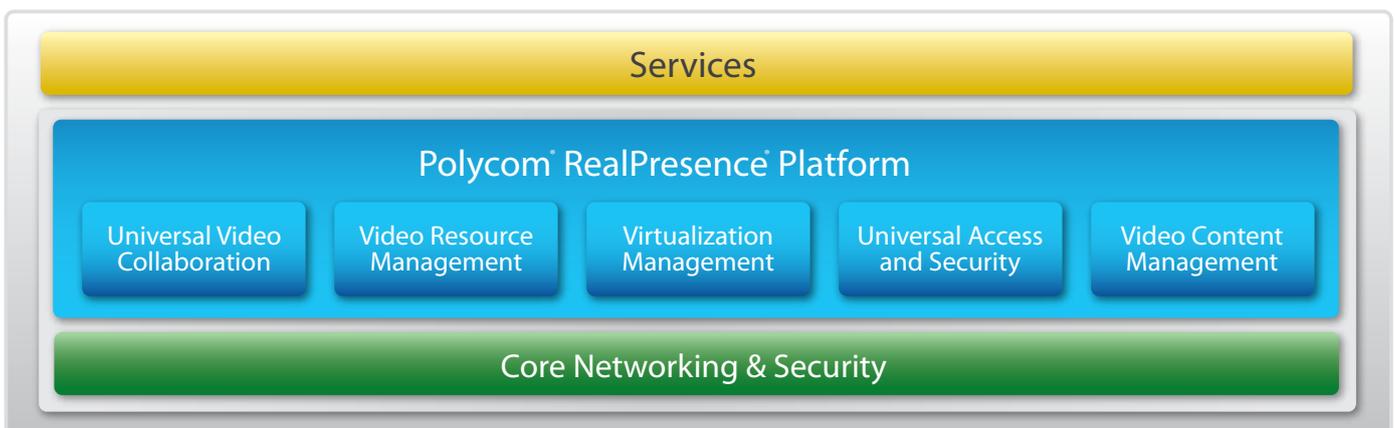
As the usage of VC in these fields has begun to grow, customers have discovered the potential to not only use VC to communicate in real-time, but also to uncover the possibilities that exist for re-using digital recordings of past events and communications.

Moving beyond “meetings”, the same technology is being used to create digitally encapsulated rich media, which can then be edited, enhanced, archived, and broadcast across multiple media. These assets can be made available to target audiences on-demand.

For example:

- **Live Event Multicasting:** The software platform enables the streaming of recorded webcasts, and supports both the push and pull of video to the streaming servers.
- **Video-On-Demand:** The software platform automates the creation of archived versions of any live event webcast so that customers can replay them on demand, as desired.
- **Media Management:** The software platform can be used to control how video content will be aggregated, approved, categorised, edited and published.
- **Storage and Archiving:** The software platform establishes rules for the lifecycle of storage for bandwidth-intensive video content: customers can determine how the content will be retained, transcoded and stored in the Cloud, or across corporate resources without daily, hands-on maintenance.

The Polycom® RealPresence™ Platform



In the previous section, we explained the five basic functional areas that constitute the software platform Polycom has developed to enable scalable, reliable, and cost-efficient VC solutions. The Polycom RealPresence platform breaks down the core infrastructure for enabling VC into five main areas:

Universal Video Collaboration: Providing the bridging capability at the core of video conferencing, this provides the software for multipoint video, voice and content collaboration that connects the most people at highest quality and lowest cost.

Virtualization Management: Providing the call management and protocol conversion that allows the bridging resources to be virtualized, this provides the software that enables multi-tenancy and massive scale, redundancy and resiliency.

Video Resource Management: Providing the device and software management of endpoints and infrastructure, enabling central management, monitoring and the delivery of video collaboration across organisations.

Universal Access and Security: The Software that easily and securely connects video participants in and outside a customer firewall and optimizes for a best collaboration experience.

Video Content Management: Software that enables organisations to support their business customers for secure video capture, content management, administration and delivery.

Summary

With demand for VC growing rapidly, solutions need to provide a robust, resilient, scalable, and manageable infrastructure that delivers unparalleled quality of experience.

The Polycom RealPresence Platform is the software infrastructure that underlies Polycom's VC, and binds it together by integrating with core networking and security infrastructure. This enables secure collaboration for any number of users across heterogeneous networks, with up to 50% less bandwidth consumption than other competitor solutions. By delivering an architecture based upon open standards, Polycom has created a flexible, state-of-the-art VC platform that provides superior investment protection, at the lowest total cost of ownership.

About Polycom

Polycom is the global leader in open standards-based unified communications (UC) solutions for telepresence, video, and voice powered by the Polycom® RealPresence® Platform. The RealPresence Platform interoperates with the broadest range of business, mobile, and social applications and devices. More than 400,000 organizations trust Polycom solutions to collaborate and meet face-to-face from any location for more productive and effective engagement with colleagues, partners, customers, specialists, and prospects. Polycom, together with its broad partner ecosystem, provides customers with the best total cost of ownership, interoperability, scalability, and security for video collaboration, whether on-premises, hosted, or cloud-delivered.

Visit www.polycom.com or connect with Polycom on Twitter, Facebook, and LinkedIn.

Polycom EMEA Headquarters

270 Bath Road, Slough, Berkshire SL1 4DX, UK

T +44 (0)1753 723282 | F +44 (0) 1753 723010

www.polycom.co.uk

