Telepresence: High-Performance Video Conferencing

ITU-T Technology Watch Briefing Report Series, No. 2 (November 2007)

1. W(h)ither videotelephony?

Do you remember your last video conference? Blurry faces on tiny screens, with sound that doesn't quite synchronize with the stilted movement of the lips. After the laborious setup of cameras and microphones, you seem to spend more time worrying about technical problems than talking about the topic at hand, with repeated loss of connection. As frustration grows, and attention wanders, it is difficult to avoid the feeling that you should have arranged a face-to-face meeting instead.

ITU Technology Watch Briefing Reports are intended to evaluate the potential of emerging technologies, in a manner that is accessible to non-experts, with a view to:

- Identifying candidate technologies for standardization work within ITU.
- Assessing their implications for ITU Membership, especially developing countries.

Other reports in the series include:

#1 Intelligent Transport System and CALM

#3 ICTs and climate change

See: http://www.itu.int/ITU-T/techwatch

The technology of videotelephony made an uncertain start at the 1964 New York World's Fair, when AT&T tested its *Picturephone* service on members of the public. The response was not very positive.[1] Although videotelephony continues to occupy prime space at many TELECOM exhibitions, because of its visual impact, successive attempts to create a commercial market for mass-market videotelephony services have generally failed. This is partly as a result of high costs and lack of bandwidth, but also because of consumer resistance to being seen on camera. As a result, the videotelephony market has developed in two quite different directions:

- At the low end, for residential consumers, many PCs now come equipped with webcams, and mobile phones have in-built cameras, which can be used for adding tiny, slow-refresh images to applications like instant messaging or video ringtones. Generally speaking, this is an *application* rather than a *service*, and the level of commercialization is based mainly on sales of equipment and bandwidth rather than of minutes of use. But, with the phenomenal success of user-recorded short videos (e.g. YouTube), and the rapid increase in residential broadband speeds, there is an expectation that wider use of real-time video will follow.
- At the high end, for business users, studio-based video conferencing has grown as a means of encouraging collaborative work among offices spread around the globe and as a substitute for travel. The aim is to give users the illusion of sitting on the opposite side of conference table from one or more remote parties (See Figure 1). High-definition (HD) video images and audio are transmitted via a packet-based Next-Generation Network (NGN), connecting multiple conference rooms around the world, and covering thousands of kilometres with virtually zero latency.

This report, the second in the Technology Watch Briefing Report series, looks in more detail at the second of these trends, namely the development of high-performance, studio-based video conferencing or "**Telepresence".** Such systems are already available on the market, and vendors have identified the technology as a potential billion US dollar market[2].

Figure 1: A sample Telepresence studio
Source: Cisco Telepresence 3000 virtual conference table[1]

2. Fields of application

The main market opportunity for high-performance video conferencing lies in the business sector, especially for in-company usage. However, other opportunities lie in the distance education, telemedicine and entertainment markets, as well as other important fields of application (see Box 1).

In the business world, telepresence applications include executive meetings, remote interviewing for recruitment, local presence for remote assistants or receptionists, remote expert consultation in product development processes, face-to-face customer support, outsourcing and business conferences.

Distance learning offers access to increased educational resources to students, regardless of their location. Scientific experiments and demonstrations carried out by teachers can be viewed remotely, with real-time interaction. Classes for hearing impaired students can also be offered. Documents, slides, spreadsheets, website and other resources can be displayed in addition to voice and video.

Box 1: Sample applications of telepresence and video conferencing

Telepresence and high-quality video conferencing solutions help to increase productivity and to save time by offering distance collaboration. Applications include:

- **Outsourcing:** the multinational information technology services company *Infosys* has deployed one of Asia's largest video conferencing facilities in its headquarters "Infosys City" in Bangalore, India. This room has the capacity to simultaneously video-conference to 24 locations, in order to collaborate with customers and its global offices across the globe.
- **Enhancing customer relationships:** In the initial rollout of its telepresence product series, *Cisco* deployed 110 telepresence endpoints in selected offices worldwide. The project resulted in more communication with customers while traveling less. Customer relationships have been enhanced by giving more opportunities to meet Cisco experts with "same room" experience.
- **Education & Training:** The *Singapore-MIT Alliance* is an innovative education and research partnership involving some of the top engineering research universities in the world. Videoconferencing systems are used every day as an integral part of these programs, providing students with access to world-class teaching resources and the possibility to share knowledge with fellow students abroad.
- **Telemedicine:** Ten French hospitals have installed a video conferencing network that links their emergency rooms to the prestigious stroke centre at *Bichat Hospital* in Paris. With better, faster treatment, stroke victims are far less likely to die or suffer permanent impairment.

Adapted from various sources, including Tandberg's customer overview, see http://tandberg.com/ind_focus/index.jsp.

In medicine, telepresence through teleconferencing can be used for remote diagnosis and therapy. Transmitted information may include medical images, multi-point audio and video conferences, a patient's medical records and output data from medical devices. The quality of the image transmitted is essential for the doctor while, for the patient, the psychological reassurance provided by telepresence is comforting. One of the most specialized and demanding applications is remote surgery, or the ability for a surgeon to perform on a patient even though they are not physically in the same location.[3].

3. From video conference to telepresence

Although commercial video conferencing dates from the early 1960s, in practice it was not until the early 1980s when Integrated Services Digital Networks (ISDN) standards allowed digital signals, such as compressed video and audio, to be transmitted over long distances, that the equipment market for video conference products began to take off[4]. Despite the high initial costs and usage charges, the benefits of saving money and time by collaborating without traveling quickly became evident. Then, as now, vendors promised increasing productivity and profitability by using video conferencing. Decisions can be made faster and travel expenses reduced. Nowadays, marketing literature also stresses the positive impact on environment. Setting up a multi-user video conference implies a significant reduction in carbon dioxide emissions compared with flying each participant to a central conference venue. A study conducted by the European Telecommunication Network Operators' association (ETNO) and the World Wide Fund for nature (WWF) showed that, by replacing of 20 per cent of business travel in the EU-25 Countries by non-travel solutions (e.g. video conferencing), it would be possible to avoid some 22 million tonnes of CO_2 emission per year[5].

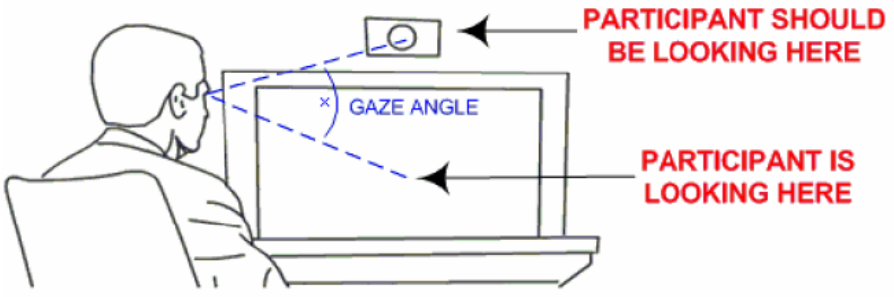

Figure 2: Camera-Screen displacement causes the loss of eye-contact.
"You are basically looking at the other person's forehead." (Brian Riggs, principal analyst for enterprise telephony at Current Analysis)

Source: "Cisco sets sail for teleconferencing" [2]
Illustration from "Emerging Technologies for Teleconferencing and Telepresence" (Wainhouse Research 2005), see http://www.wainhouse.com/whitepapers/.

Nevertheless, the last thirty years have showed hardly any impact of video conferencing in slowing the rise in business travel and video conferencing has never really taken off as a standalone market. The drawbacks lie in the difficult setup, insufficient quality and poor reliability of the video and audio transmission. In addition to these technical defects, users noted socio-cultural aspects such as lack of "true" eye-contact (see Figure 2) as well as the self-consciousness of appearing on a television screen.

4. Telepresence characteristics

In order to improve the video-conference experience of users, it is necessary to make significant advances in all three areas – network technologies, conference hardware, conference software – in order to provide the user with the experience of "being there without going there"[6].

In today's telepresence offerings, participants may now appear "life size", on one metre plus HD plasma monitors or LCD displays. Live video resolutions can go up to 1080p at 30 frames per second, where 1080 represents the number of lines of vertical resolution and p the progressive, non-interlaced mode of scanning. Those parameter values implement ITU-R Recommendation BT.709 which defines HDTV standards. The first demonstrations of HDTV in Europe and North America date back 25 years, and future technologies will be based upon ITU-R and ITU-T Recommendations on LSDI (Large Screen Digital Imagery). This set of Recommendations defines how "super HDTV" images – up to four times the quality of standard HDTV – can be delivered to cinema-like venues, bypassing traditional distribution methods.

In telepresence, spatial CD quality audio is directed to the conversation partner simulating the acoustical feeling of face-to-face talks. To improve eye-contact between users, multiple HD cameras are deployed closely above the screens in order to obtain a small angle between camera, eyes and screen. Concealing the camera in the centre of an immersive screen would help to achieve "true" eye-contact[7] [8]. Conferences become more life-like, and allow for much more interactive communication, including the use of body language.

Figure 1 shows a typical telepresence solution. This virtual conference room is a three-panel, 165 cm plasma screen system complete with a specially designed table that seats six participants per side or a "virtual table" for twelve. To give the illusion of debating in the same room, conference rooms are equipped with similar decoration.

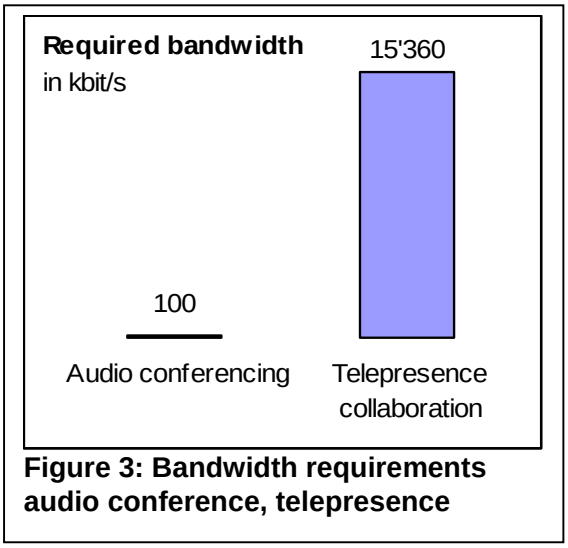

Figure 3: Bandwidth requirements audio conference, telepresence

Conference software now focuses on usability, simplicity and interoperability, allowing the user to set up easily conferences between two or more offices (multipoint). Presentations, documents and files can be shared between conference rooms and instantly be made available on an additional display, improving collaboration and interactivity.

Transmitting video and audio in HD quality demands high-bandwidth connections. To achieve life-like experience, potential delays must be negligible for human eye and ear. Bandwidth requirements for 1080p conferences are specified as 15 Mbit/s (Megabits per second)[9]. More bandwidth would be required to connect additional offices to a conference; whilst using a lower, but HD resolution of 720p would decrease bandwidth demands[10]. In relative terms, telepresence requires around 150 times more bandwidth than traditional voice conference calls, as illustrated in Figure 3.

While the demand for bandwidth rockets upwards, so too does the level of service available. By the end of 2006, there were around 280 million broadband Internet subscribers in 166 countries, representing two-thirds of the total number of Internet subscribers[11]. Furthermore, available bandwidth has been increasing by 66 per cent per year while median price has been falling by 41 per cent per year since 2003, which is faster than Moore's Law for semiconductor price-performance[12]. This suggests that high performance video conferencing is now becoming more available and affordable for domestic users.

Providing the telepresence experience, however, may still require dedicated networks (see Section 5), as it results in more challenges for network service providers (NSP) than merely offering higher bandwidth. The guaranteed availability of bandwidth on demand is essential as rescheduling meetings due to network unavailability is not an acceptable option for business customers or for remote surgery. Before starting a session, users should have the possibility to reserve bandwidth (via call admission control, CAC). Telepresence traffic should be detected automatically by network operators and be given high priority in return for a higher price and to comply with strict service-level agreements (SLA) to safeguard QoS. For applications such as remote surgery, sticking to SLA can be "vital". For business conferences it is essential that end-to-end security is also assured and that the NSP can protect their networks from distributed-denial-of-service (DDoS) attacks or unauthorized access.

5. Telepresence market

The key players in the field involve both, telepresence solution vendors, like Cisco, HP, Polycom and Tandberg, and NSPs like NTT and Verizon, which are already members of ITU and actively involved in standardization activities. However, new providers with different service models will emerge. Telepresence studios can either be purchased (current prices for fully furnished rooms can cost between US$250'000 and US$500'000) or leased on a monthly rate.

As a vendor of standalone telepresence applications, Cisco expects its solutions, which run over the customers' own network provided it meets the necessary network requirements, to generate US$1 billion annually in revenue from hardware sales by 2013.[13]. According to research done by Cisco, telepresence network services from the full range of providers will represent a US$4 billion opportunity for NSPs by 2010[9]. HP, another leading vendor, offers its customers both a telepresence studio, called Halo, and a private, high bandwidth (45 Mbit/s plus), full duplex, worldwide fibre optic network – called Halo Video Exchange Network (HVEN) – connecting via private leased lines that are dedicated exclusively for the use in video conferencing. In addition to the cost of a studio, additional monthly fees are payable for network and operation costs[14].

A different service model focuses on building networks of telepresence studios in important business locations around the world and renting the studios on an hourly or daily rate to companies without their own video conferencing facilities. They offer concierge-level services around the actual conference, including call scheduling and suite reservation services, call management, remote monitoring and monthly reporting. Other major providers of telepresence solutions, who are not yet ITU members, include Digital Video Enterprises, Telanetix and Teliris.

Customers may include companies, organizations and states that recently have implemented policies to become carbon neutral, which is increasingly seen as good corporate or state responsibility. A growing list of corporations (e.g. *PepsiCo, Google, Yahoo!, Dell*) and territories have already announced dates for when they intend to become fully carbon neutral. Video conferencing and remote collaboration play key roles in those policies to reduce travel, principally flights.

6. Implications for ITU-T

Telepresence has implications for ITU both as a potential user organization and as the leading global standards development organization in the ICT field.

ITU-T is currently experimenting with remote collaboration tools, including *GoToMeeting* and *WebEx*, as a way of facilitating remote participation in its meetings, especially from developing countries. A first trial is being carried out with a link between ITU's headquarters in Geneva and the Cairo regional office, during the workshop on human exposure to electromagnetic fields (EMF)[15] on 20 November 2007. Such remote collaboration tools that may include additive video transmission and do not require more than a web browser and a conventional Internet connection.

Box 2: Working definition of Next Generation Network
A Next Generation Network (NGN) is a packet-based network able to provide services including Telecommunication Services and able to make use of multiple broadband, QoS-enabled transport technologies and in which service-related functions are independent from underlying transport-related technologies. It offers unrestricted access by users to different service providers. It supports generalized mobility which will allow consistent and ubiquitous provision of services to users.
Source: ITU-T Recommendation Y.2001, see http://www.itu.int/rec/T-REC-Y.2001/en.

In contrast, telepresence requirements cannot be met on today's public Internet, but are implicit in the specifications for next generation networks (NGN) (see Box 2), which is a major focus of ITU-T standards-making. The rollout of NGN will usher in a new era of multimedia communications and bring with it a need to consider updating or replacing the currently used multimedia protocols, such as H.323 (developed in ITU-T Study Group 16) and Session Initiation Protocol (SIP) (developed by the Internet Engineering Task Force SIP Working Group). Interoperability is a key requirement of telepresence to ensure broad connectivity with traditional and emerging video environments. Today, most of the available products support both, H.323 and SIP. A workshop in held in May 2006, jointly organized by ITU-T and IMTC (International Multimedia Telecommunications Consortium), identified strong and weak points in both protocols (see Box 3), and proposed to migrate H.323 and SIP into a new generation of multimedia protocols, called H.325 or Advanced Multimedia Systems (AMS), that takes into consideration special aspects of security, flexibility and QoS.

Work on AMS also addresses the current lack of multimedia support for mobile systems. Portable devices, as well as PCs and IP desk phones will become more powerful, hence bringing the two segments of the video conferencing market (low-end residential use and high-performance, studio-based business use) closer together. The video functionality in instant messaging applications has become very popular among PC users. Standards require low complexity codices– for mobile use – and have to focus on low power consumption as well as interoperability among devices and different systems. Today's standards for video compression, like ITU-T H.264, are very appropriate for high-motion video content. Nevertheless, in order to obtain quality beyond HD, existing standards have to be enhanced in matters of resolution, frame rate, colour accuracy and efficiency.

Box 3: H.323, SIP: is H.325 next?

ITU-T and the IMTC jointly organized an ITU-T workshop and the IMTC Forum 2006 in San Diego, California, USA, from 9 to 11 May 2006, on the topic "H.323, SIP: is H.325 next?" H.323 describes terminals and other entities that provide multimedia communications services over Packet-Based Networks (PBN) which *may not* provide a guaranteed QoS. SIP is an application-layer control (signaling) protocol for creating, modifying, and terminating sessions with one or more participants. These sessions include Internet telephone calls, multimedia distribution, and multimedia conferences. SIP is characterized by its proponents as having roots in the IP community rather than the telecommunications industry[16]. While SIP originally had a goal of simplicity, in its current state it has become just as complex as H.323[17].

The conclusions and recommendations from the Workshop include:

- Bandwidth is getting cheaper quickly, therefore:
  - Compression is still important for video, but less than it used to be.
  - Compression for audio is already adequate; the focus now is on features, quality, etc.
  - Flexibility and interoperability are key issues. Security, rate adaptation, complexity/power and error robustness, etc are likely to be more important in the future.

Participants identified limitations in existing protocols, such as

- Poor or complex capability exchange;
- Poor error handling and fault management;
- Multiple interoperability issues;
- SIP and H.323 are problematic for mobile systems; operators have adopted H.324M;
- Little consideration is given to NAT/FW and other IP network issues;
- Important aspects, such as QoS, security, lawful interception, emergency services, provisioning and management were only considered at a late stage in standards development, resulting in less-than-ideal solutions.

Those challenges result in an opportunity for ITU & IMTC to take a lead in addressing some of these issues, notably:

- Quality of service/experience;
- Availability of content;
- Interoperability;
- Mobility.

Source: Adapted from ITU-workshop site, see http://www.itu.int/ITU-T/worksem/h325/200605/index.html.

7. Implications for developing countries

For developing countries, the success of video conferencing in general and telepresence in particular is tightly linked to the deployment of NGN infrastructure and the higher bandwidth required for high-performance services. Applications in education, medicine (see Box 4) and business promise great benefits for developing countries, but depend on the availability and reliability of more powerful networks. ICT vendors and service providers with global operations may establish branches and research centres in emerging economies, like India, and use telepresence to collaborate with their head office or other research units. Universities and institutions of higher education in developing countries have also been co-operating with universities in the developed countries to share knowledge via distance learning, and to make it available in remote regions[18]. Telepresence

Box 4: Telemedicine in Mozambique

The Government of Mozambique, in cooperation with ITU, has established a telemedicine link between the central hospitals of Maputo, the capital, and Beira, the country's second largest city some 1'000 km away from the capital. The link allows the hospitals to exchange messages regarding laboratory results and treatments as well as X-Rays. As a result, doctors in Beira can refer cases to the central hospital in Maputo for primary or secondary opinions and send medical records to the capital so that experts there can determine whether patients facing more serious problems can be treated locally or should be transferred to Maputo. The project was especially important for the hospital in Beira since it had no radiologist when the telemedicine link was established.

For developing countries, such telemedicine projects tend to be relatively expensive to implement. The approximate cost in hooking up Maputo and Beira was US$50'000, with the main cost being the digitization of the X-ray images. Mozambique's Government was so happy with the results that its Prime Minister wrote to the ITU to ask for its help in establishing additional telemedicine links with a hospital in Nampula, the country's third-largest city, with part of the cost to be covered by the government. Similar telemedicine projects with which ITU is involved are currently underway in Senegal, Uganda and Ukraine.

Adapted from "Internet and Health: Is there a doctor online?"
http://www.itu.int/newsarchive/wtdc2002/Internet_Health.html.

will help to enhance the degree of interactivity and collaboration between students and educators. In addition to communication in high definition, personal video communication on mobile devices will also play a major role in developing countries, once the infrastructure is provided, as this is likely to be more affordable and more available, given that the number of mobile users in developing countries is sometimes more than ten times greater than the number of fixed line connections.

8. Conclusion

With the help of high-definition displays, NGN and improvements in usability, telepresence should be able to resolve some of the disadvantages associated with today's video conferencing. Telepresence users still comprise a small elite, but globalization, the increasing need for international collaboration and a desire to reduce carbon emissions will drive demand, as will frustrations and higher costs (due to rising oil prices) associated with long-distance travel. HD video communication will enhance interactivity and productivity in business, as well as in applications in the fields of education and medicine. In the future, high performance video conferencing can be expected also to trickle down to personal and mobile video communications, and therefore standards have to be adapted for the use on a wide range of devices to guarantee interoperability.

Glossary of abbreviations and acronyms used in the document

AMS	Advanced Multimedia Systems
CAC	Call Admission Control
CD	Compact Disc
CO_2	Carbon dioxide
DDoS	Distributed Denial of Service Attack
ETNO	European Telecommunication Network Operators' association
FW	Firewall
HD	High-Definition
HDTV	High-Definition Television
HVEN	Halo Video Exchange Network
ICTs	Information and Communication Technologies
IETF	Internet Engineering Task Force
IMTC	International Multimedia Telecommunications Consortium
IP	Internet Protocol
ISDN	Integrated Services Digital Networks
ITU	International Telecommunication Union
ITU-R	ITU Radiocommunication sector
ITU-T	ITU Telecommunication standardization sector
LCD	Liquid Crystal Display
LSDI	Large Screen Digital Imagery
Mbit/s	Megabits per second
MIT	Massachusetts Institute of Technology
NAT	Network Address Translation
NGN	Next-Generation Network
NSP	Network Service Provider
PBN	Packet-Based Network
PC	Personal Computer
QoS	Quality of Service
SIP	Session Initiation Protocol
SLA	Service-Level Agreement
VoIP	Voice over IP
WG	Working Group
WSIS	World Summit on the Information Society
WWF	World Wide Fund for nature

Notes, sources and further reading

[1] See, for instance, a brief history of videophones on the "Connected Earth" website at: http://www.connected-earth.com/Galleries/Frombuttonstobytes/Digitaltelephony/Videofromthephone/.

[2] Report on Cisco Systems CEO John Chamber's opening keynote at the Networkers 2006 conference in Las Vegas, see http://www.vnunet.com/vnunet/news/2158699/cisco-sets-sail.

[3] One of the earliest remote surgery operations was conducted on 7 September 2001, with a surgeon performing a gallbladder operation on a patient 6'230km away. See http://en.wikipedia.org/wiki/Lindbergh_Operation.

[4] For a detailed history of video conferencing see http://www.wiredred.com/video conferencing-history.html or http://myhome.hanafos.com/~soonjp/vchx.html.

[5] See Chapter 3 of the joint WWF-ETNO road-map "Saving the Climate @ the speed of light", which discusses the opportunities for ICT services to reduce CO_2 emissions. It is online at http://www.etno.be/Portals/34/ETNO%20Documents/Sustainability/Climate%20Change%20Road%20Map.pdf. This topic is further discussed in ITU-T Technology Watch Briefing Report #3 (October 2007), on "ICTs and Climate Change".

[6] This is the slogan for HP's telepresence solution, "Halo". See http://www.hp.com/halo.

[7] Digital Video Enterprises offers telepresence solutions using embedded cameras to obtain true eye contact between conversation partners, see http://www.dvetelepresence.com/products/eyeContactSilhouette.asp.

[8] In January 2006, Apple Computer patented a new kind of video monitor with tiny cameras between the pixels, see http://appft1.uspto.gov/netacgi/nph-Parser?Sect1=PTO1&Sect2=HITOFF&d=PG01&p=1&u=%2Fnetahtml%2FPTO%2Fsrchnum.html&r=1&f=G&l=50&s1=%2220060007222%22.PGNR.&OS=DN/20060007222&RS=DN/20060007222.

[9] White Paper "Delivering a Cisco TelePresence Network Connection Service: The Value of an IP Next-Generation Network", Cisco Systems, see http://www.cisco.com/en/US/netsol/ns672/networking_solutions_white_paper0900aecd805bbda0.shtml.

[10] Tandberg's telepresence solution is called "Experia". Using a resolution of 720p, an 8 Mbit/s capacity is required, see http://tandberg.com/collateral/product_brochures/TANDBERG_Experia.pdf.

[11] Source: ITU World Telecommunication Indicators Database, at: http://www.itu.int/ITU-D/ICTEYE/Indicators/Indicators.aspx.

[12] Global average estimates, based on ITU's annual survey of broadband services and prices, as reported in ITU/UNCTAD "World Information Society Report 2007: Beyond WSIS", available at www.itu.int/wisr.

[13] VoIP-News article "Getting started with Telepresence", see http://www.voip-news.com/feature/getting-started-with-telepresence-040407/.

[14] In October 2006, HP released Halo Collaboration Studio in Asia Pacific. A press statement includes pricing information, see http://www.hp.com/halo/pdf/Halo_China_Computer_World_Press_English.pdf.

[15] See the event website at: http://www.itu.int/ITU-T/worksem/emc-emf/200711/index.html.

[16] For a comparison of H.323 and SIP see http://www.packetizer.com/voip/h323_vs_sip/.

[17] "SIP's Future: Complicated and Competitive" (Article by John G. Waclawsky, Business Communication Review 2/2007), see http://www.packetizer.com/labs/documents/BCR-Feb2007.pdf.

[18] Distance learning ventures include the African Virtual University or MIT Open Courseware