

Telepresence: Beautiful And Expensive

John Bartlett

The quality is stunning, but so is the price—and the impact on your network

John Bartlett is a consultant and VP with NetForecast (www.netforecast.com), specializing in data and real-time application performance on enterprise networks and the Internet. He can be reached at john@netforecast.com.

Immersive large-format face-to-face videoconferencing systems have been around for a number of years now, but the new telepresence systems from Hewlett Packard and Cisco, introduced last year, are really giving this technology a boost in visibility. With these two market powerhouses pushing the technology, telepresence has suddenly become a conversation topic in boardrooms around the world.

Those who have tried the new systems find the experience startlingly good, and a big step up from older videoconferencing. The idea of high-quality, face-to-face meetings conducted at a distance over the network is compelling, especially in these times of high fuel costs, increasingly difficult travel, and a growing awareness of our travel's impact on the environment.

But this kind of good news always comes at a cost. The systems are quite expensive, and the enterprise network must be prepared for the demands of this new application. It's no exaggeration to say that telepresence, which depends on very low packet loss, very low latency and minimal jitter (variation in latency), will be the toughest application your network has ever had to support. That top-level execs will literally see the results of your efforts in their telepresence sessions ups the ante even further.

Enterprises need to take a hard look at how this new application will be integrated with the existing network and with the applications currently running on it, both to ensure the quality of the telepresence experience and to prevent impacts from telepresence on the performance of existing business-critical applications.

Isn't This Just High-end Videoconferencing?

Many of us have had the experience of using a standard videoconferencing system and being

TABLE 1 Telepresence Product Comparison

Vendor	Screen Configuration	Number of Participants per Room	Number of Connected Rooms Max (no image switching, full size images)	Shared Data Connection
Cisco TelePresence	Models come in 1 or 3 screen configs, with a projector for data sharing or document camera	Up to 6 for a System 3000, 2 for a System 1000	2 for 1-screen system, 4 for 3-screen system	Yes, VGA input and document camera
HP Halo	3 Screens plus 4th for data sharing, image camera or H.323 interconnect	Up to 6	4, camera concentrates on center 2 or center 4 participants	Yes, VGA input and document camera
Polycom RPX	2 screens or 4 screens, each a 4-foot x 42-inch cinematic screen, LCD display in table for every 2 participants for collaboration	Up to 28 (up to 2 tiered rows of seats behind conference table seats)	3 for 2-screen system, 5 for 4-screen system	Yes, VGA input and document camera
Tandberg Experia	3 screens plus 1 below image screen for collaboration	Up to 6	4	Yes, VGA input and document camera
Telirus VirtualLive	5 or more	One more than the number of screens	Customized as needed	Yes, laptop VGA, document camera, smart board, other input as needed

frustrated with the complexity of the setup. With such a small screen, it's hard to tell which person at the remote conference room table is which and it's hard to distinguish who is speaking. We often wonder if it is any better than just having a good telephone conference. But to answer the question above: No, telepresence is not just high-end video-conferencing.

Have you ever been to an iMax theater, where you are surrounded by the movie screen, and felt like you were falling or leaning as the image shifts around you? When there is that much screen you forget that you are sitting in a theater seat, and start to believe you are on that little biplane flying through the Grand Canyon.

Now let me set your expectations down a bit: Telepresence isn't an iMax theater, but the idea is the same. The goal of telepresence is to provide an immersive experience—that is, one with sufficiently powerful audio and visual effects so that you feel you are together in the same room.

Telepresence vendors Hewlett Packard, Cisco and Polycom told me about how they had involved cinematographers and/or psychologists to help them design these systems, and the results are impressive:

- The full-sized images of the remote participants are very lifelike.
- The audio is very clear, and provided on multiple channels, so that a person's voice comes from the direction of their screen image.
- The lighting is set up to make you look good without makeup, and to light faces so there are no deep shadows.
- Attention is paid to eye contact, so that when

you look at a remote participant's image, they see you looking (nearly) at them, and vice versa.

All these factors enhance the communications, and make you feel like you really met with those remote folks, looked them in the eye, and had a good discussion. The accumulation of these cues (visual, auditory, eye contact, etc.) takes us across a threshold of perception that is more than an incremental improvement. It allows us to believe at some level that we are meeting together.

But They Cost A Small Fortune!

Indeed they do. Systems range from a low of \$60,000 up to nearly \$700,000 per room (Table 1), and may carry a management contract which adds another \$18,000 per month to the budget.

These prices definitely put telepresence in the category of a strategic investment that must return value to the company in a significant way. How can these costs be justified?

There are three levels of justification to consider. Level 1 is the traditional rationale that video-conferencing will replace (at least some) travel, so travel costs will fall. While travel savings probably return the least amount of dollars, compared to productivity improvements and business acceleration (discussed below), the numbers for travel are well understood, and the financial team can count them easily.

Unfortunately, even when the numbers look good, many CFOs just don't believe that travel really will be reduced, and often they are right. The question to ask is: Can we really reduce the travel budget if we install this equipment? Some companies are tackling this the other way around

The goal is to make you feel as though you are all in the same room

Connect to H.323 Video Conferencing	Connect to Audio-only participant	Bridging (accomodating more sites through switching or multiple small images per screen)	Equipment List Price (per room)	Managed Service Costs (per month)
Announced, to be supported later this year	Yes	Up to 36, uses switching, each site always viewed on a single screen with life-size image	\$79,000 for 1-screen system, \$299,000 for 3-screen system	\$3,211 for 1-screen, \$5,900 for 3-screen, no network
Yes, through a single screen, part of Tandberg alliance	Yes	Future- part of Tandberg alliance	\$329,000	\$18,000 per month per room includes network
Yes	Yes	Yes, both bridging and switching modes available	\$299,000 to \$695,000	\$6,700 - \$12,700 per room per month, no network
Yes	Yes	Yes	\$199,000 and up	Managed services available through Tandberg partners
Yes	Yes	No	\$60,000 to \$250,000	\$8,000 per month per room, includes network

**Telepresence
saves travel time
and expense, and
it can accelerate
the pace of
business**

and starting with a travel budget cut, then forcing employees and work groups to find alternative communications methods. This approach often also leads to productivity enhancements, the level 2 cost justification.

Here's an example of how productivity increases: I am based in Boston. It will cost me 2 days to have an important face-to-face meeting in California with a client, partner or colleague. I have to leave home in the early afternoon to catch a long flight to California, then spend the night, have my important meeting in the morning, catch a return flight and get home well after dinner.

Even using mobile laptops, PDAs and cell-phones, much of that time is not productive, and I have missed two evenings with my family. If I and my California colleague had telepresence systems, we could meet for an hour or two on the first day, and the task would be done. If I had three

meetings to do in three different cities around the country or the world, I still could get all three done in a day. In short, I get more done in less time—and I don't have to miss my wife's cooking or my daughter's ballet recital.

These level 2 justifications are obvious to the participants, but they can be a bit difficult to quantify. They also will have a correlated effect, which brings us to level 3, the acceleration of business. If the use of telepresence can shorten a product design cycle, close a key contract sooner, establish a partnering arrangement and make it productive or solve a problem for an important customer in hours instead of days, weeks or months, then the pace of business accelerates. Shortening the cycle for each of these business components provides a handsome return and quickly justifies the cost of a telepresence investment.

Substantially the same cost justifications were

FIGURE 1 Point-to-Point Connections

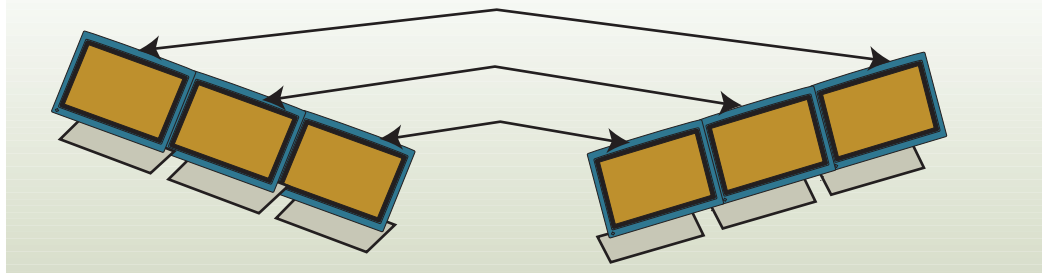
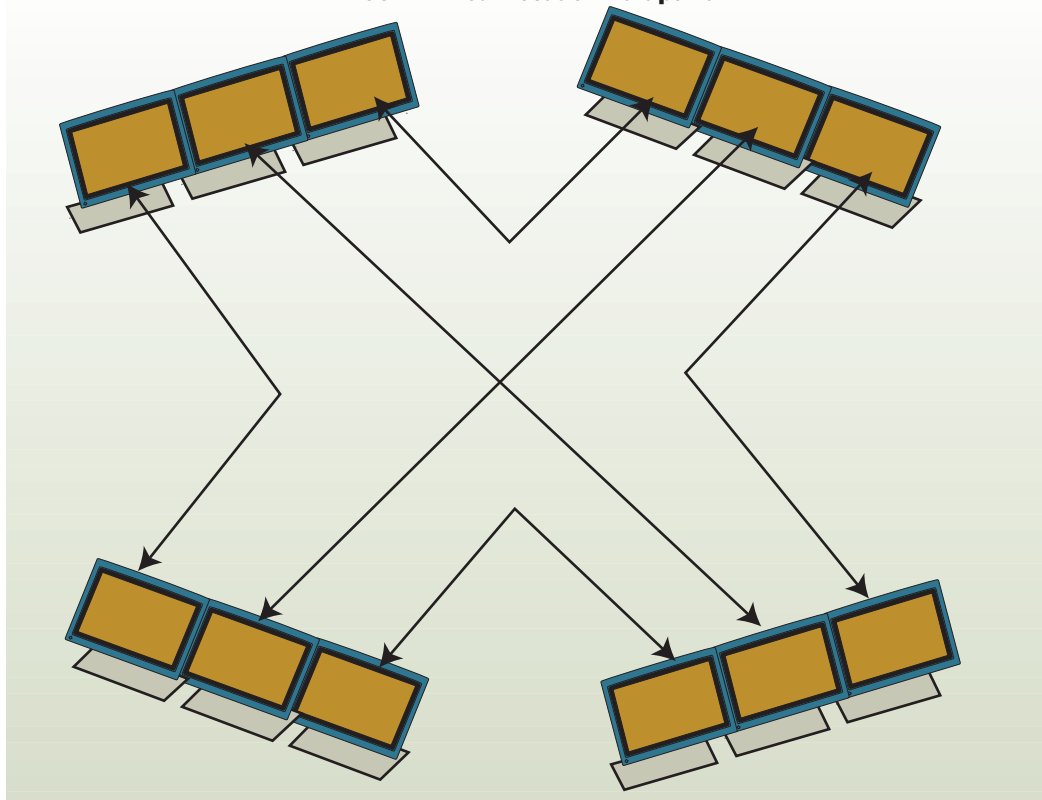


FIGURE 2 Four-Location Multipoint



put forward, with mixed success, in the early '90s for rollabout room videoconferencing systems, and later in the '90s when desktop videoconferencing was being hyped, but these new telepresence systems could have better luck. They are so much better and so much easier to use, and the vendors' pitch is being aimed at executives instead of network/IT people. If the execs want telepresence, they will certainly find it is easier to justify than, say, a corporate jet.

Isn't Videoconferencing Complicated And Unreliable?

Well, yes, it can be. Those who have been most successful with traditional videoconferencing have put a concerted effort into managing the videoconferencing environment for their users. High-level meetings are scheduled in advance and they are completely managed. The systems are started and verified 10 minutes before the meeting starts, and a technician often remotely monitors the conference to ensure that systems are alive and that quality is being maintained.

Telepresence vendors are targeting high-level executives as their first customers because these users can most quickly justify the expenditures and benefit from the systems. Simplicity, quality and reliability are paramount to gaining acceptance with C-level users, so most telepresence vendors are providing a managed service offering with their equipment (Table 1). These managed services handle all aspects of operation including scheduling, call setup and quality monitoring. Some of these services also include the global high-speed IP network needed to interconnect telepresence sites.

For example, Polycom and Cisco sell a managed telepresence solution without a network, while HP and Teliris include the network in their managed offerings. Cisco automates scheduling for the user via integration with the calendaring functions of Microsoft's Outlook and integration with their VOIP telephony system.

The new systems differ from traditional videoconferencing in that most do not require a multi-point control unit (MCU or bridge) for conferences in which the other locations can be seen on one of the available screens. Thus a 3-screen system can support up to 4 sites with no MCU. Tandberg, Polycom and Cisco offer bridging to support additional sites when needed. Tandberg and Polycom bridges provide both switching and continuous presence, while Cisco provides only a switching mode of operation. Switching means that remote sites consume a whole screen, and different sites appear or disappear based on who is currently speaking. Continuous presence means that a screen is broken up into smaller screens, and each remote site can be seen all the time, but in a smaller form factor.

Some vendors say that bridging detracts from the telepresence experience. Polycom acknowl-

edges that this is true, but notes that the importance of including those additional locations and people in the conference often offsets the desire to maintain a perfect telepresence environment. All the vendors support a telephone connection, which means that one additional exec who is on his way to the airport can at least participate in the audio portion of the conference. (For more product comparisons, see Table 1.)

What Is The Impact On My Network?

Significant. Telepresence has all the real-time requirements of voice over IP (VOIP) such as very low packet loss, low jitter and low end-to-end latency, and it carries the additional requirement of very high bandwidth—up to 20 Mbps per room.

In fact, telepresence can be the most performance-challenging application your network will carry. You will have to allocate sufficient bandwidth, get the QOS settings right, and make sure there aren't any conflicts with the performance parameters of other apps.

These difficult requirements present another argument for the telepresence vendors' managed service offerings. When you buy an HP system, for example, they bring their network into your building, all the way to the telepresence room. The telepresence system traffic never flows over any of the enterprise LAN or WAN components. HP manages the traffic end-to-end in order to guarantee a high-quality experience. The HP team says this reduces the risk for the customer in terms of quality, schedule and unexpected costs.

Teliris also provides a managed network as a part of their service, but only brings its network to a demarc in the building. The customer must provide connectivity from the demarc to the conference room, but high-bandwidth LAN equipment is relatively inexpensive and this is not too difficult. Even if the customer doesn't use a dedicated LAN, and chooses instead to run the telepresence system on a converged LAN, it is not hard to use a separate VLAN and configure it appropriately.

Cisco, Polycom and Tandberg offer the telepresence endpoints, and they expect the enterprise will carry the traffic on their existing IP networks, or work out a network solution with a third-party WAN service provider. If the enterprise plans to use the management service offered by Polycom, the enterprise must also handle the connectivity back to the Polycom's Video Network Operations Center (VNOC).

If you decide to use your own network for telepresence, homework is required to ensure that the network can properly support this traffic, which is no easy task. Let's take a look at the issues involved in providing transport for these telepresence rooms:

■ **Bandwidth**—Telepresence systems consume about 5 Mbps per screen of full-duplex network bandwidth. Each screen needs about 4 Mbps of real data transfer, and then the overhead of the



Managed equipment and network services are part of some telepresence offerings

Packet loss, latency and jitter can break the illusion of being in the same room

RTP, UDP, IP and Ethernet headers adds about 20 percent, bringing the total close to 5 Mbps per screen. Cisco claims that much lower bandwidth is used most of the time because there is not a lot of movement or rapid gesturing in most business meetings, but that the full bandwidth should be provisioned for best quality.

If we use three screens in each of two telepresence rooms, and connect them point-to-point, then there are three full-duplex 5-Mbps streams flowing in each direction, as shown in Figure 1.

Now consider a multipoint telepresence meeting with four sites, each with three screens. In this scenario, most vendors make point-to-point connections, tying one screen in each telepresence room to one in each of the other three sites (Figure 2). Participants in each room must scrunch together a bit so that they all can be seen on a single screen. In this configuration we have a full-duplex, 5-Mbps stream flowing between each site and each of the three other sites.

■ **Packet Loss**—Real-time voice and video streams use the UDP protocol instead of TCP, because there is no point in going back and getting a missing voice or video packet, as TCP does for other traffic flows. Real-time streams must stay on time and in sync so the sounds and images appear natural at the far end.

To overcome occasional packet loss, manufacturers have designed clever ways of masking this missing information, to minimize its impact. But the information really is missing, so each of these algorithms is just an approximation of the original content. Needless to say, the best quality is achieved when the data is not lost.

The more highly compressed the original information, the greater the impact on quality when data is lost. Video compression algorithms often use an incremental change approach, sending only the changes in an image from the previous frame (rather than sending the full frame each time). This is very effective for a conference-room environment, where most of the background of the image is static, and movement of individuals against that background is the “new” information that needs to be updated frame by frame.

But now consider the impact of a lost packet. If new frames are based on incremental changes from older frames, then the lost information not only causes a quality problem for the frame to which it belongs, but also to subsequent frames, since they are built on information which never reached the receiving end. The results can be pops and breaks in the audio, blurry blocks in the video, and portions of the image freezing for a half second or so until the system recovers.

High-quality telepresence meetings require consistently low packet loss on a high-speed, real-time connection for hours at a time. Cisco recommends that the network packet loss be less than 0.05 percent. Guaranteeing such a tight tolerance requires careful network design. Yet this still rep-

resents about 20 lost packets per screen per minute. The effect of those lost packets will be small if they are well-distributed in time, and if they don’t represent critical parts of the image, but if they occur in bursts they will still be noticed.

■ **Latency**—Latency is the delay imposed by the network and the telepresence system between one end and the other. If a listener on one end of the telepresence connection raises a finger in protest, how long will it be before the speaker on the other end sees that raised finger and notes the objection? This latency value is important to keeping the natural feel of a face-to-face meeting.

Most humans don’t notice audio delays of less than 150 ms, so this is the well-accepted one-way maximum latency in the voice environment. This value is very difficult to achieve in current video-conferencing environments, and, because the audio has to sync with the video, whichever is slower will govern the user’s experience.

Cisco says their system creates less than 200 ms of latency between two adjacent systems, with no network in between. Most humans notice delays above 250 ms, so this leaves only about 50 milliseconds of latency budget for the networks connecting the rooms.

Unfortunately, 50 ms barely gets you across the continental U.S. Table 2 shows the network latency between New York and various cities, calculated with a conservative model I created a few years ago based on a large dataset from an international network service provider. By conservative, I mean that some improvement to these latencies is possible through careful network design, although the speed of light will always contribute a significant portion of these values.

Network latency does not affect the quality of the video images or the sound, but it does affect the interactive nature of the conference. People can learn to adjust to these delays, as they did with satellite-based long distance calls, but it takes practice, it’s annoying, it’s tiring and it breaks the illusion of being in the same room.

It’s clear in Table 2 that delays will be noticeable on telepresence sessions between New York and most U.S. cities west of the Mississippi. Even lengthier delays will occur between New York and

TABLE 2 Latency From New York (ms)

City	Network Latency	User to User Latency
Houston	40	240
Denver	43	243
San Francisco	55	255
London	67	267
Tokyo	108	308
Hong Kong	125	325
Bangalore	128	328
Sydney	149	349

the cities in the Asia Pacific region. Telepresence vendors need to improve the quality of experience with innovations that reduce the latency added by the video equipment.

■ **Jitter**—Jitter is the variation in latency. If a packet normally takes 100 ms to traverse a network path, but due to congestion experienced along the path it actually takes 150 ms, it has incurred 50 ms of jitter.

Jitter is important because packets arriving late can miss their scheduled play window. The video image is being reproduced in real time, and thus the data must arrive on schedule so that it can be used as a part of the video construction. A packet arriving late cannot be used.

Most video and audio systems have a jitter buffer, which delays the audio or video reconstruction by some amount of time, like starting an event 10 minutes late to allow tardy patrons the opportunity to arrive and be seated. Typical jitter buffers are 50 milliseconds long, allowing packets to be up to 50 ms late and still be used.

Of course, the jitter buffer delay contributes to the latency of the system and to the overall latency of the end-to-end connection. System designers must trade off the quality of the network (e.g., its ability to deliver packets with low jitter) against the total latency of the video system.

The jitter buffers of traditional videoconferencing systems range from 50 to 100 milliseconds. VOIP phones often have dynamic jitter buffers which contract when the network quality is good and expand as the quality of the network degrades. Some can expand to 100 or 200 milliseconds.

In contrast, the newer telepresence systems, with their emphasis on providing as close to a “real” conferencing environment as possible, are designed with smaller jitter buffers and thus require the network to cause less jitter. Cisco recommends that the network jitter remain below 10 milliseconds for telepresence system support. This is a significant challenge in a converged, global network.

Whose Network To Use?

Before you consider adding telepresence traffic to your corporate WAN, you need to be able to answer “yes” to the following questions:

■ Do you have QOS deployed on your network? Is it operating both at layers 2 (Ethernet) and 3 (IP), on the LAN as well as on the WAN?

■ Have you fully implemented voice over IP on your network?

■ Do you have network resiliency supporting less than 50 ms failover times for your network links?

■ Are all your links at least 45 Mbps?

■ Do you have real-time end-to-end network testing tools to measure loss, latency and jitter?

If you have to say “no” to any of these questions, but your execs have already decided that you are implementing telepresence, then you will probably want to start with a managed telepres-

ence service, or at least engage a service provider with a focus on telepresence to provide the links.

All the answers are not yet in on how to run this high-demand traffic across a converged network and keep all the applications happy. And remember, any performance problems you encounter will be displayed, as they occur, on 50-inch plasma screens to your top management.

Conclusion

The telepresence market is still very young, but all the vendors are reporting high interest, and high utilization of the systems that have been installed at customer sites and within their own companies.

Many of the usability and quality concerns that dogged traditional room-based videoconferencing systems have been overcome in telepresence. The new systems have bigger screens, lower latency, higher quality spatial audio, better attention to room details (color, lighting, etc.) and a focus on ease of use that should please many users—especially those who know first-hand how hard traditional videoconferencing can be. And there is this hard-to-describe immersive quality, which really makes the experience work.

Unfortunately, we also have the usual early market problem of proprietary systems, notably those from Cisco, HP and Teliris, that don’t interoperate. Polycom and Tandberg have standards-based systems that will talk to each other, and Tandberg is working with HP to make their systems interconnect, albeit not fully at this time.

Despite these obstacles, the vendors tout their interest in business-to-business telepresence, and how they are working to support it (at least between their own customers). Today, for example, customers of Teliris can quickly arrange a business-to-business call. The Teliris network is designed to open a secure MPLS tunnel between the two companies for the duration of a scheduled call, and then close it up again to maintain security once the call completes.

The vendors will have to make major business and technical efforts if they want to offer an interoperable business-to-business telepresence capability. Expensive travel, environmental woes, terrorist threats and other worries would suggest there is a growing market for such systems. However, even if the vendors do not move in this direction, telepresence is already a great step forward from room-based videoconferencing □

Any performance problems will be displayed, as they occur, on 50-inch plasma screens to your top executives

Companies Mentioned In This Article

Cisco (www.cisco.com)

HP (www.hp.com)

Polycom (www.polycom.com)

Tandberg (www.tandberg.com)

Teliris (www.teliris.com)