

Web Conferencing, Instant Messaging & Streaming Media
Prepared for the International Communications Industries Association
To support Infocomm 2002

By Andy Nilssen & Marc Beattie
Senior Analysts and Partners, [Wainhouse Research](#), Brookline, MA, USA

.....

Marc F. Beattie has authored public and private reports on product strategies, distribution structures, emerging technologies and industry applications. Marc is a co-author in Wainhouse Research's *Conferencing Markets & Strategies* series. He is the principle instructor for the *Implementing Streaming Media* seminar. Prior to joining Wainhouse Research Marc was an early team member at PictureTel and Polycom – holding his last position at Polycom as Director of Business Development; he has spent 15 years working within the industry. Marc has been an independent analyst for the last three years. Mr. Beattie attended Gordon College where he earned a Bachelor of Arts degree in Economics.

Andrew H. Nilssen is a co-author of Wainhouse Research's *Conferencing Markets & Strategies* series, and consults for industry leaders, Fortune 500 companies, and venture capitalists about the multimedia communications industry. Andy has over 20 years of experience in bringing high-technology products to market. Prior to Wainhouse Research, he was Director of Marketing, New Business Development at PictureTel, Inc. where he defined strategies and identified partners to expand business by leveraging broadband and streaming technologies. Andy earned his MBA and BSEE degrees from the University of New Hampshire and holds two ease-of-use related patents.

.....

Web conferencing, instant messaging (IM) and streaming media are three emerging “rich media” communication technologies that have been made possible through the combination of computer-generated information and the widespread growth of the Internet. All three technologies have experienced rapid growth due to the impact of their powerful yet inexpensive capabilities. Each has also spawned dedicated companies that offer the component technology in product form or bundle the products with integrated services to present a turnkey rich-media conferencing solution for the user.

Web Conferencing

Web conferencing is built on the premise that graphical data represents a large portion

of the content in a conference call, and the new reality that there is access to a network-connected web browser virtually everywhere a telephone is located. Web conferencing sells itself on providing a way to augment a typical audio conference with powerful graphical information exchange.

Most web conferencing systems use client-server architecture to enable simultaneous exchange of audio and graphical content. As shown in the figure above, a web conference typically consists of two parallel conferences – an audio conference using telephones on the PSTN overlaid with a separate data conference using web browsers on a network for data. The degree of audio integration in a web conference can vary greatly, ranging from requiring the user to set up a totally separate audio call, to

displaying the conference number to call in to via the web browser, to automatically calling out to all participants (scheduled or ad-hoc).

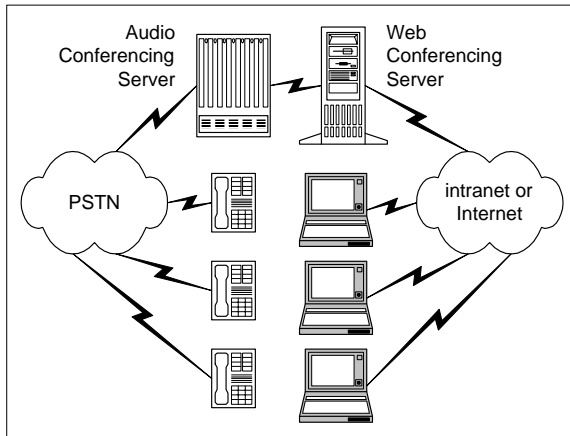


Figure 1 - Web Conference Architecture

In a web conference, a presenter can present PowerPoint slides, demonstrate software applications, use an interactive whiteboard, and take polls to gauge the audience's reaction. In some specialized customer support applications, the presenter can actually take control of a viewer's PC to help walk them through a problem.

Every web conferencing system requires the use of a local software client on the participant's PC to add web conferencing features to a standard web browser. This usually takes the form of a Java applet, which is downloaded "on the fly" and usually saved locally in the browser's cache so that it does not need to be downloaded again for subsequent web conferences. The size, features, and efficiency of this local Java client are different from one web conferencing service provider to another. The method for capturing and transferring the graphical image from the presenter to the participant's web browser can vary. In most web conferences, uploading a PowerPoint presentation to the web conferencing server is the norm – though some systems (ex. WebEx) actually process and distribute the presentation from the client itself. Basic

web conferencing systems will simply translate the PowerPoint slides to static bitmapped GIF or JPEG images, thereby losing slide transitions and animations, and also locking the slide resolution into whatever translation is used. More sophisticated systems use either a proprietary object descriptor language or DHTML to translate PowerPoint slides into graphic objects, and to preserve high resolution and possibly transitions and animations as well. Graphical object-based systems also allow slide resolution to scale to take full advantage of the resolution of an individual participant's browser, and can be more bandwidth efficient. Application sharing and "web tours" have similar considerations – the basic systems just "peel & send" the presenter's screen, while more sophisticated systems support true multi user application sharing – some are even built on top of the T.120 (specifically T.128) application sharing engine which is also used in products like Microsoft NetMeeting.

Web conferencing systems need to get along with firewalls – or more precisely, need to get through firewalls reliably. If the web conference involves participants located on different corporate LANs, and the web conferencing service is not designed to understand a particular firewall, the audience behind that firewall will not be able to participate. Penetrating firewalls is not a trivial issue. The more advanced web conferencing technologies will actually try different mechanisms on the fly to figure out which policy is optimal for getting through a particular firewall. The web browsing protocol, HTTP, uses port 80 – which is probably left open by most firewall administrators, but it is not the most efficient. More efficient than HTTP is TCP, which is also used for streaming media, file transfers, and terminal emulators; however, most administrators will use a firewall to block ports to unsolicited TCP traffic. Finally, IP ports, which act like a sub-address to an IP address, may also be blocked or restricted by the firewall. If a firewall restricts port 80, web browsing will

be blocked. Because web conferencing systems involve a local client, a good system will be able to try each successive method to get through the firewall in the preferred order – TCP, HTTP with dynamic ports, and HTTP via port 80 – in order to establish the most efficient web conferencing connection through a firewall.

Two last considerations are the browser itself and standards. Web conferencing vendors spend considerable effort debugging their clients to run with each new revision from major browser vendors to be sure a particular web conferencing service works reliably with most popular browsers. For most web conferencing implementations, the vendors' attitude towards standards is often "who cares?" since server communications are web conference specific and the clients are lightweight on-the-fly Java downloads. Most vendors believe there just isn't too much investment or interoperability to be concerned with. However, this will start becoming a little muddy as IP audio and video become part of a web conference and as users start using many different web conferencing services and multiple web conferencing Java clients accumulate on user's PCs.

Instant Messaging

Instant Messaging (IM) is becoming an influential factor in the rich media conferencing services space for two reasons:

- 1) IM systems uniquely use *presence* to display personal contacts along with their real-time availability status. Thus the probability of successfully connecting with someone is known before trying to make the actual connection. No more telephone tag, no need for voice mail – IM is bringing a welcome change to the age-old hit-and-miss paradigm for calling someone.
- 2) IM systems are increasingly becoming much more feature rich,

and are quickly moving to augment basic text-based chat with file transfer, data sharing, PC remote control, and even real-time IP-based voice and video.

As with other forms of conferencing, IM can be implemented by purchasing an IM server or by using an IM service. The big three commercial IM service providers – AOL/ICQ, Yahoo IM, and Microsoft MSN – are not compatible, meaning a user on one service cannot communicate with a user on another. "Commercial-grade" IM software server packages are available for organizations and service providers to purchase to host their own IM sessions for security or cost reasons (ex Jabber, Bantu, Microsoft Exchange Server). The presence features of IM are also beginning to merge with web conferencing in products such as the Polycom WebOffice and Lotus Sametime.

The introduction of Microsoft Messenger for Windows XP marks a major move by Microsoft to expand their MSN IM service into a full conferencing service. Messenger for XP is IM-based, but features integrated data conferencing (file transfer, whiteboard, application and desktop sharing), and taunts users with "Add Voice" and "Add Camera" icons. Wainhouse Research has found that Messenger for XP's wide-band audio can be excellent and the software-based video compression using a USB camera is adequate for effective desktop videoconferencing. Messenger for XP can rendezvous with presence information only when connected to either Microsoft's .NET services (which requires each user to get a Microsoft Passport account) or to an enterprise-hosted Microsoft Exchange server – so clearly Microsoft is playing on both the client and the server side. The service also requires the network to support a new network protocol called "Universal Plug-and-Play", which may take many years to enable.

Streaming Media

Streaming media technology is a relatively new vehicle for extending multimedia conferences to any user with a network-connected PC and web browser.

Streaming technology involves creating, serving, and playing a mix of audio, video, and/or presentation data over a network without having to wait to receive the entire multimedia file of the event. Because IP networks introduce variable delay, streaming requires the receiver to accumulate pieces of data off the network at the top of a buffer while simultaneously emptying the bottom of the buffer for real-time playback of audio and video. Since receiving data from the Internet in particular can be unpredictable, the entire streaming process can introduce a fair amount of delay, easily measured in 10's of seconds; thus streaming technology is uniquely applicable only to receiving one-way live or stored presentations, where real-time, two-way interaction is not a priority. The audio and video stream can also be accompanied by synchronized presentation data in the form of still images or Flash-encoded visuals.

There are many different ways to look at the streaming industry. Wainhouse Research has divided the industry into four segments, each of which includes hardware, software, services and components, each of which represents a distinct application set. Most vendors operate in just one of these segments, although some of the vendors and service providers are expanding their offerings to cover multiple pieces of the puzzle.

- Content Creation. Products and services in this sector enable customers to create content that is optimized for streaming media delivery. We include here all the steps taken to capture, edit, and encode content for streaming delivery.
- Content Management. This is also known as digital asset management. These products and services make libraries of digital assets more valuable and more accessible. One of the key elements of this market segment is the area of audio and video indexing, which enables intelligent searching on streaming media files. Another major segment is represented by systems which enable owners of digital content to control access, manage libraries, track usage, and generate detailed reports and billing statements.
- Content Delivery Networks. If a user is delivering streaming media on their own LAN, chances are they will have ample bandwidth and control. But if the stream is to be delivered to far-flung locations, or to sites not on the LAN, or if the event is to be accessed by the public at large, then some type of delivery network or systems architecture should be used. A variety of content delivery networking schemes exist; most use the public Internet for part of their solution, some use satellites, some use intelligently placed servers to move content closer to the end users, thereby minimizing the traffic that has to move over the Internet backbone.
- Client-Server Products. There is no recognized standard for streaming media today. The major client-server choices today include Apple's Quicktime, Microsoft's Windows Media Technologies, and Real Network's RealSystem. Some client software systems are designed to be used as stand-alone player applications; others are designed to reside inside a standard web browser. Some streaming products use no client at all; instead a Java applet is loaded automatically when needed to play the streaming file.

Other important elements in the client-server segment are the data compression/decompression engines or codecs, which are usually implemented in software though some codecs use special hardware to increase their speed. Some of the most popular streaming codecs are Sorenson Broadcaster, RealVideo, MPEG-1, MPEG-2, MPEG-4, and H.263. Most of the client-server products on the market support multiple codecs.

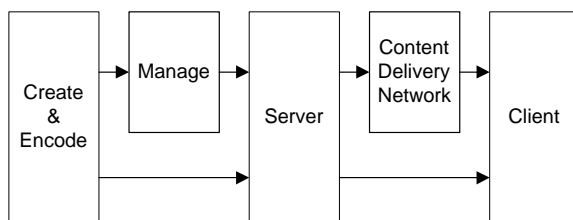


Figure 2 The Four Streaming Functional Blocks

Figure 2 illustrates how these four functional blocks fit together. All streaming applications start with a content creation and encoding process. Sometimes, the content is indexed, managed, or subject to access control, although this is not always the case. All streaming files eventually are placed on a streaming server from where they are delivered to the users over some type of IP delivery network, corporate LAN, extranet, or Internet connection. The content is eventually viewed (or heard) with some type of client software that must be compatible with the protocols, file formats, and compression techniques that are supported by the server.

Streaming media services fall into two broad segments: consumer streaming services, which uses the Internet for delivery, and professional corporate streaming services, which uses both private intranets as well as the Internet.

The streaming hardware and software servers available to streaming service providers vary according to a wide variety of user needs within these two categories:

- Number of simultaneous streams supported – this can range from single digit numbers to several thousand
- Number of files available – primarily a function of storage space
- Bandwidth supported – most of the consumer/Internet users today are connected at 28.8 kbps, but corporate LAN users are in a position to take advantage of the much higher data rates possible. Higher data rates translate to much higher audio and video quality.
- Security features - authentication features ensure site security or enable pay per view for direct revenue generation. A server interface may connect to Web-based backend credit card and transaction processing systems.
- User Accounting – provides the hosting service the ability to know who watched what video when.
- Interfaces to live video feeds such as news, weather, etc. Since these feeds are generally NTSC analog video, the server also may contain an encoding solution which digitizes the video and then compresses it to the streaming format required. The encoding solution may be software residing on the host computer or it may be a codec hardware accessory on an add-in board. Industrial-strength streaming applications often use hardware codecs which enables a single PC to encode multiple streams simultaneously.
- Management systems that present to users an organized list of what is available for viewing.

For conferencing applications, streaming servers are specifically configured to work with videoconferencing systems and real-

time input. The first generation streaming conferencing servers use the video and audio signals directly from the videoconferencing system as the input to the streaming subsystem. Future conferencing servers are likely to have unified user interfaces and digital trans-coding to make the conferencing to streaming connection seamless.

Streaming media clients, which are designed to run as both browser plug-ins and stand-alone applications, can be quite large and are thus downloaded and installed on the receiver's PC prior to a streaming media event.

Streaming video can span a wide range of bit rates from 20 kbps up to MPEG1 (typically 1.5 Mbps) and occasionally MPEG2 rates (more than 8 Mbps). Determining the optimal bit rate requires a trade-off between the network bandwidth available versus the fidelity/quality needed. Network bandwidth can be segmented into three categories: Internet/modem speeds (< 56 kbps), ISDN speeds (< 256 kbps), and intranet speeds (> 256 kbps).

Audio- and video-on-demand technology enables service providers to complement real-time streaming offerings with conference storage and retrieval. Training videos, briefings, audio conferences, earnings calls, employee benefits programs,

training sessions, or recordings of live feeds can be stored on a server, to be reviewed on demand by users from any browser-enabled desktop. Only a single copy of the original event is maintained on the server, which simplifies updating, access management, and distribution control.

Match the Media to the Message

While many of these conferencing technologies are new and often more exciting to work with than a standard audio call, Wainhouse Research encourages clients to evaluate the message they are trying relay and apply the appropriate medium to match the message.

Figure 3 presents two curves that one should keep in mind while choosing the rich media conferencing vehicle for a particular event. The first curve illustrates the trade-off between Cost and "Connected-ness" – for example, if cost is not an object, then why not use videoconferencing for everything? Videoconferencing provides maximum visual feedback with your audience, which is important if the meeting requires persuasion, selling, or consensus building on a controversial topic. If this level of connected-ness is not required, then a lower cost conferencing vehicle could be used. The second curve illustrates the trade-off between Interactivity and Reach, which fortunately is consistent with a natural

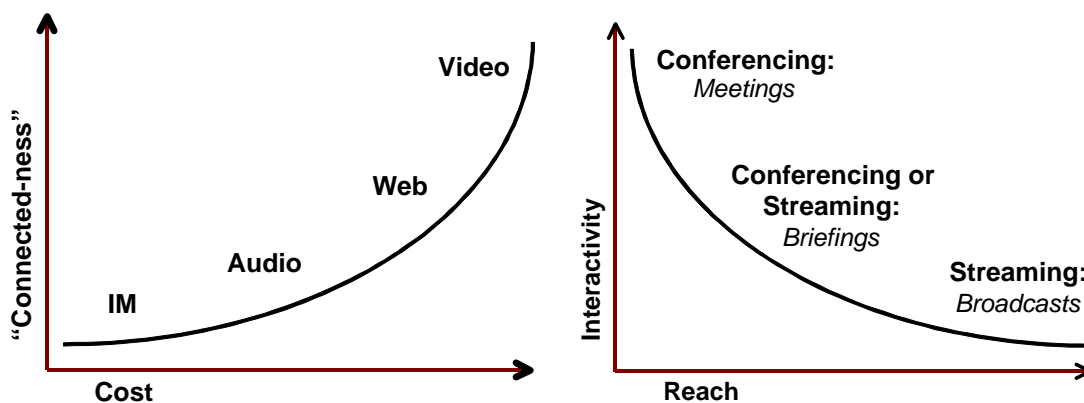


Figure 3 Match the Media to the Message

meeting dynamic: the more people involved in an event (“reach”), the less interaction there tends to be. Hence an interactive Meeting should use a real-time conferencing technology, while an informational broadcast can use streaming. A briefing could use either technology depending on the degree of real-time interaction required to cover the content.

It may be that a simple audio conference achieves 90% of the intent. If the sharing data is important many easy-to-use web conferencing services are available to add to your audio calls. While streaming can provide impact and reach, knowing what your target audience is capable of receiving - PC clients on a modem are not often good candidates for streaming - will determine if they ever get your message. If the message involves critical persuasion, the added “connected-ness” benefit provided by a true two-way videoconference may be most effective.

If the web conference is a presentation where one site is doing the majority of the talking, audio streaming of the presenter’s voice is becoming a popular option. In this case, audience feedback to the presenter and to other sites is limited to text chat, polling, and “mood” indicators – which, for structured presentations, can be quite effective. Very high-quality human voice

audio can be delivered at low bandwidth using standard streaming platforms such as Real Networks G2 or Windows Media Technologies. Each individual stream will typically use less than 8kb/sec of network bandwidth per user. Audio over 28k modem connections was considered little more than a toy as few as 24 months ago. Today effective streaming audio can co-exist with graphical data updates and still be delivered very well over modem speed connections. If the network supports multicasting, it can be used to minimize network utilization for the one-way audio stream as well.

While each of these technologies has its independent capabilities, they can also be used together. By way of example, IM can be used to determine the “presence” and status of colleagues. If your IM client indicates their availability, you can request a web conference to share a document or object on your PC. While engaged in that web conference participants can IM one another on the subject being discussed while not disturbing the conversation of the larger group – this would be similar to leaning over to someone in a conference room and having a sidebar conversation. If some team members were not available for the web conference the meeting could be captured – both the audio and data - and archived for viewing at a later time via streaming.