Will Your Next Video Bridge Be Software-Based?

Examining a Next-Generation Software-Based Video Conference Server
Will Your Next Video Bridge Be Software-Based?

Examining a Next-Generation Software-Based Video Conference Server
Executive Summary

A conference server is the software and related hardware implementation that enable video bridging along with significant ancillary functions such as browser based facilities to control conferences, manage data, route information, and sometimes serve web- or T.120-based data flows to conference participants. Many current-generation video bridges may be classified as conference servers because of the significant amount of supplementary server-like functionality they provide.

Due to unique new software architectures and more powerful CPUs, video bridging can in many instances be successfully accomplished using a software-based conference server running on an inexpensive PC platform instead of requiring an expensive hardware-based video bridge. Hardware-based bridges are still required in certain computationally intensive conferences, such as when transcoding between different audio or video compression technologies or when rate matching between endpoints is needed; however, for many conferences, particularly those internal to a single enterprise over the IP (Internet protocol) network, a software-based conferencing server may now be a viable alternative.

Software-based conference servers running on commonly deployed CPUs using standard operating systems are less expensive to own and to deploy than hardware-based servers. Upgrades are easy: adding ports is done by entering a new license key either on the same or a new PC while new or updated server functionality can be obtained through a simple software installation. This contrasts markedly to a hardware-based conference server upgrade where adding ports often involves cracking the case to add additional blades to a hardware backplane and adding new functionality typically requires more deployment time and perhaps replacement of an entire main processor blade.

Because software-based conference servers run on standard PC operating systems, many ancillary functions can be easily added. For example, some software conference servers natively support web conferencing capability seamlessly integrated with audio and video bridging functionality. Furthermore, new methods for handling video through advanced server cascading functionality are capable of significantly reducing bandwidth when using video over the wide area network.

Given the recent advances in software-based conferencing servers, enterprises bringing web conferencing and IP video bridging in-house should carefully consider whether an integrated software-based conference server will meet their need for functionality and cost versus a hardware-based video bridge and a second web conferencing server.

From Video Bridges to Conference Servers

Early video bridges supported the H.320 standard, which enabled multipoint videoconferencing over ISDN connections. ISDN created great opportunities for widespread videoconferencing but it imposed constraints on the architecture and design of the video bridges required for multipoint conferencing. Meetings were often held where participants used different ISDN connection parameters; as a consequence, powerful digital signal processors (DSPs) were required on each video port to enable rate matching between the endpoints. Higher processing capability also enabled continuous presence videoconferences where three, four, or more individuals were able to be viewed simultaneously in small panes on the screen. Until recently, video bridges were large rack mount units sold primarily to large corporations and service providers.

With the advent of H.323 for conferencing over packet-switched networks, a new breed of multipoint bridging units designed for small to large enterprises has evolved. These bridges are smaller, cheaper, and often contain functionality designed to enable more widespread use of video within the enterprise. One of the most notable improvements has been the development of web-based interfaces to these devices. Today, most video bridges can be configured and diagnosed through a simple point and click
interface; furthermore, many bridges now have added capabilities to control individual conferences, call out to video endpoints, disconnect endpoints, and put the conference in lecture mode.

Enter the Age of the Conference Server

With the significant increase in integrated functionality added to video bridges in recent years, these devices are evolving into sophisticated conference servers that do much more than just mix and retransmit audio and video. There is a growing trend toward integrated client-server conferencing environments where the conference server enables a whole host of conferencing services including audio, video, and web conferencing, as well as instant messaging, streaming and archiving.

The entire conferencing industry is moving toward integrated collaboration solutions because there is demand for increased conferencing flexibility. Video bridging manufacturers are developing, either organically or with partners, integrated audio, web and videoconferencing solutions. These solutions are often based on expensive single-purpose audio and video bridging devices employing proprietary hardware loosely linked to an in-house or third party web conferencing server.

These DSP based hardware servers are still the right solution when a conference needs to handle computationally intensive processes such as transcoding between different audio and video compression algorithms, rate matching, and laying out one or more continuous presence views where significant video image manipulation is required. These types of devices are well suited for video bridging service companies and for very large organizations that regularly conference using ISDN videoconferencing devices of varying capability and bandwidth.

The Case for a Software-Centric Conference Server Solution

A handful of companies including, First Virtual Communications (FVC - www.fvc.com), are taking a different approach to video bridging and conferencing. Rather than investing resources in proprietary hardware-based servers for specialized video applications, these companies are developing highly sophisticated software-based conferencing servers that run on off the shelf PC platforms. FVC is clearly the market leader in developing a standards-based, non-proprietary software-based conference server solution.
Recognizing that most IP videoconferences can be negotiated to use similar conferencing parameters, software-based conference servers avoid processing huge amounts of video data by functioning under the following conditions:

1. Transcoding is not required. This means that each endpoint enters a particular conference using the same video compression codecs. This condition is still fulfilled when a gateway is involved, because the gateway does any necessary transcoding between endpoints running H.320 and H.323.
2. Rate matching is not required. This simply means that each endpoint connects to the conference at the same bit rate, a condition that can be met in the majority of IP videoconferences.
3. High resolution continuous presence images are not required. Either the conference can be done using voice activated switching, or continuous presence multipoint video images are deemed of satisfactory quality by combining four QCIF images (176 x 144 pixels) into a single CIF sized image (352 x 288 pixels), which is not computationally expensive.

Under these conditions, software conference servers avoid the computationally intensive video manipulation steps and can provide satisfactory performance.

Performance Data

One may justifiably question whether or not an off the shelf PC running conference server software is capable of adequately performing video bridging, and whether such a solution is scalable. To give an example of the CPU utilization video bridging has on the PC CPU, FVC conducted benchmarking tests and provided the data used in the tables below.¹ The first table shows CPU utilization under four different voice activated switching scenarios.

<table>
<thead>
<tr>
<th>Conference Parameters</th>
<th>CPU</th>
<th>Average CPU Utilization</th>
<th>Maximum CPU Utilization</th>
</tr>
</thead>
<tbody>
<tr>
<td>H.323: H.261 CIF video; G.711 Audio</td>
<td>Single 1.5 GHz Xeon</td>
<td>28%</td>
<td>61%</td>
</tr>
<tr>
<td>H.323: H.261 CIF video; G.711 Audio</td>
<td>Dual 1.5 GHz Xeon</td>
<td>20%</td>
<td>31%</td>
</tr>
<tr>
<td>H.323: H.261 CIF video; G.722 Audio</td>
<td>Single 1.5 GHz Xeon</td>
<td>53%</td>
<td>64%</td>
</tr>
<tr>
<td>H.323: H.261 CIF video; G.722 Audio</td>
<td>Dual 1.5 GHz Xeon</td>
<td>18%</td>
<td>32%</td>
</tr>
</tbody>
</table>

Table 1: FVC Conference Server CPU utilization statistics.

These data show that in addition to serving multipoint audio and video, there is significant CPU capability remaining to serve complementary conferencing applications, such as web conferencing or instant messaging.

¹ These tables show representative performance data for FVC’s software-based Conference Server version 7.0.
The next table shows data from scalability testing, again using voice activated switching. What makes these tests especially compelling is the number of simultaneous conferences that can occur on a software-based platform.

<table>
<thead>
<tr>
<th>Bit Rate</th>
<th>Codec Parameters</th>
<th>Endpoints/Conference</th>
<th>Total Conferences</th>
<th>Total Users</th>
</tr>
</thead>
<tbody>
<tr>
<td>384 kbps</td>
<td>H.323 H.261 CIF video G.711(\mu) audio</td>
<td>4</td>
<td>20</td>
<td>80</td>
</tr>
<tr>
<td>768 kbps</td>
<td>H.323 H.261 CIF video G.711(\mu) audio</td>
<td>4</td>
<td>13</td>
<td>52</td>
</tr>
<tr>
<td>256 kbps</td>
<td>CUSeeMe(^2) H.263 QCIF video G.711(\mu) audio</td>
<td>10</td>
<td>20</td>
<td>120</td>
</tr>
<tr>
<td>512 kbps</td>
<td>H.323 H.261 CIF video G.711(\mu) audio</td>
<td>4</td>
<td>50</td>
<td>200</td>
</tr>
</tbody>
</table>

Table 2: FVC Software-Based Conference Server scalability performance parameters

These performance data suggest that software-based conference servers running on off the shelf processors are capable of handling significant numbers of conferences. They also show that high scalability can be accomplished in two ways: using faster CPUs and employing multiple CPUs. Faster CPUs obviously allow the various processes involved in serving the audio and video to run more rapidly, providing scalability by allowing more streams to be served in the same amount of time. Multiple processor machines effectively allow processes and threads within processes to execute more rapidly because pieces of the calculations are distributed over more than one processor.

One may naively suppose that scalability would be infinite simply by continuing to add additional processors. In fact, tests show that as the number of processors increases, the amount of overhead required to manage the tasks running on the individual processors also increases until a point is reached where additional scalability is not possible due to multiprocessor overhead. However, network bandwidth saturation occurs due to the number of simultaneous endpoint connections long before multiprocessor overhead limitations are encountered.

The list price for the 200 port 512 kbps basic videoconferencing bridging solution referenced above including the Quad Xeon 1.5 GHz is under $175,000. This compares favorably with a hardware conference server deployment which could list at around $500,000\(^3\), depending upon the vendor and the options selected. Consequently, if an enterprise is able to meet the conditions for which a software-based conference server is designed (no transcoding, no rate matching, and sending QCIF image size when using continuous presence), then significant capital expenditure reductions can be realized investing in a software-based conferencing server.

Wainhouse Research is aware of at least one IP video network service provider using FVC’s software-based conference server to provide video bridging capability to their customers running any H.323 compliant endpoint. The solution is sufficiently cost effective that the provider offers unlimited video bridging as part of the standard service agreement.

---

\(^2\) CUSeeMe is a proprietary protocol used in FVC’s Click to Meet web-based software endpoint.

\(^3\) For the additional money, the hardware device would be able to do transcoding, rate matching, and sophisticated continuous presence in addition to basic video bridging which are not always required in enterprise videoconferencing over IP.
Distributed Servers and Intelligent Linking

When conferencing across an enterprise wide area network, it may be desirable to host a conference on more than one conference server. Because conference servers are licensed based upon the total number of simultaneous ports utilized, enterprises can install copies of the software on CPUs in remote locations creating distributed conferencing servers. Distributed servers provide significant flexibility and convenience when designing a network topology to enable integrated IM, web, and video conferencing.

There are three primary reasons why one may consider installing conference servers at different locations within the network: 1) to add media control points, and 2) to minimize wide area network bandwidth, and 3) to better facilitate large conferences. Enterprises will typically want to install a conference server at strategic locations where gatekeeper functionality is needed in order to control media access to the network and local area network bandwidth utilization. This is accomplished using the gatekeeper functionality included within each conference server. The gatekeeper does not consume conferencing ports and the distributed use of gatekeepers gives network administrators multiple conferencing control points.

FVC’s conference server minimizes the bandwidth consumed in a multipoint call by using a technology called intelligent linking. With intelligent linking, the video image from each call participant is sent to the local conference server. By communicating with other conference servers, each conference server is aware of which video streams are being viewed by those participants connected to that server. If no one happens to be viewing a particular participant’s video image, the stream is not served to the network, thus conserving bandwidth across the wide area connection.

![Figure 2: Conference Server intelligent linking conserves bandwidth across wide area networks by serving only the video streams being viewed by other users.](image-url)
Intelligent linking over distributed conference servers can also better facilitate larger video conferences while simultaneously minimizing bandwidth. H.323 supports the concept of simple cascading which consists of starting multipoint conferences on two or more video bridges, then connecting the conferences together by having the video bridges call one another. Each bridge is each viewed by the other bridge(s) in the call as if it were a single endpoint. In voice activated switching multipoint calls, this is usually not a problem because only the video image of the last person speaking is served to other bridges, which in turn serve the image to the participants connected to that bridge. However, in a continuous presence call, where there are multiple callers connected to each bridge, the multiple images from one MCU appear as miniature images within a single pane on the other, as illustrated below.

![Figure 3: Cascaded continuous presence calls without Intelligent Linking illustrating miniature panes.](image)

FVC’s intelligent linking eliminates these miniature tiled panes because each server is aware of the video images to be displayed. Rather than sending pre-mixed images to other conferencing servers, intelligent linking allows each conference server to send and receive four separate QCIF streams and mix them locally on each conference server. Intelligent linking works even when standard H.323 endpoints are the only endpoints in the conference: each endpoint will see four individual QCIF sized images stitched together to form a single CIF sized image instead of the mini-panes from conventional bridges.

Simple H.323 cascading on some video bridges can also lead to incorrect audio switching when individuals on different bridges are speaking. When a mixed audio stream received from a remote bridge contains multiple speakers, the increased amplitude can make it difficult to determine which video image should be served to each endpoint. Intelligent linking overcomes this difficulty because unmixed audio is received by each conference server, then mixed locally before being sent to endpoints attached to that particular server. Because audio is handled in this way, each conference server knows which image(s) to display.

Confidential Servers Enable Rich Media Web Conferencing

Video bridging is one example of the capability of a conference server. Software-based conference servers are also able to provide access to a variety of integrated rich media conferencing tools. Besides video bridging, FVC’s conference server provides a foundation for instant messaging, web conferencing and collaboration capability, desktop videoconferencing, and streaming through the company’s Click to Meet family of solutions. Click to Meet provides instant messaging capability, a full web conferencing application as well as a fully functional software videoconferencing endpoint that uses the same audio and video codecs as those found in H.323, but encapsulated within a proprietary FVC wrapper. The Click to Meet Application Family can also invite standards based H.323, SIP (i.e. Microsoft Messenger for Windows XP and SIP telephones), and T.120 clients into conferences using the feature set provided by the underlying conference server.
Although video is not required to participate in a conference, Click to Meet’s video capability works with inexpensive web cameras, H.323 desktop videoconferencing cameras, or with set top box video conferencing units from Polycom, Tandberg, Sony, and others.

Using FVC’s Click to Meet software endpoint has some significant advantages because each user is allowed to individually control how many or how few video images they view. This is in contrast to a conventional multipoint call with standard desktop or set top box videoconferencing units, where each individual receives only what the organizer of the call pre-established with the bridge or arranges through a bridge’s dynamic continuous presence options. By communicating with the conference server, the Click to Meet endpoint allows participants to choose who and what they see and when they see it.

Mixing interactive participants and view-only participants in a single meeting is possible using the conference server’s streaming capability. FVC’s Conference Server supports streaming using Microsoft’s Windows Media, Real Network’s streaming technologies, or Apple’s QuickTime.

**Conclusion**

Software-based conference servers offer an attractive alternative to single purpose dedicated hardware video bridges from both a price and a functionality standpoint. Under the conditions many enterprises will encounter in their IP conferencing environments, software conferencing servers can successfully bridge video calls. Among software-based conference servers, FVC’s Conference Server coupled with the Click to Meet family of solutions provides one of the most compelling implementations of a tightly integrated collaboration environment by seamlessly providing instant messaging, web conferencing, audio and video conferencing within a single application. FVC’s intelligent linking capability has the ability to conserve bandwidth when calling over wide area connections, and it can actually enhance large multipoint conferences by assuring that the video images are properly displayed in cascaded conferences. Whereas software-based bridging has often been discounted in the past, it may be time for serious conferencing users to take a careful look into software-based conference server capabilities.
About First Virtual Communications

First Virtual Communications is a premier provider of next generation web conferencing and collaboration solutions, delivering integrated communications solutions for businesses that need to remotely collaborate, train, demonstrate or sell. The Company’s solutions integrate seamlessly with existing tools and methodologies, such as email and web browsing, while extending the advantages of instant messaging and collaboration environments, such as MSN Messenger and Microsoft Exchange. FVC’s innovative solutions are deployed in over 1,200 customer sites worldwide, including Fortune 500 companies, government agencies and service providers. Additional information about First Virtual Communications can be found on the Web at http://www.fvc.com.

About Wainhouse Research

Wainhouse Research (http://www.wainhouse.com) is an independent market research firm that focuses on critical issues in rich media communications, videoconferencing, teleconferencing, and streaming media. The company conducts multi-client and custom research studies, consults with end users on key implementation issues, publishes white papers and market statistics, and delivers public and private seminars as well as speaker presentations at industry group meetings. Wainhouse Research publishes a number of reports detailing the current market trends and major vendor strategies in the multimedia networking infrastructure, endpoints, and services markets, as well as the segment reports Comparing IP Video Network Service Providers Versus the Naked Internet and Surviving in the Conferencing Reseller Channel and the free newsletter, The Wainhouse Research Bulletin.

About the Author

E. Brent Kelly is a Senior Analyst and Consultant at Wainhouse Research. Brent was formerly VP of marketing at Sorenson Vision where he launched the company’s live streaming and IP videoconferencing products. He has authored articles on IP conferencing and has developed seminars on implementing IP-based Rich Media Conferencing. As an executive in a manufacturing firm, he developed and implemented a marketing and channel strategy that helped land national accounts at major retailers. Brent has significant high tech product management and development experience, working on the team that built the devices that Intel uses to test their microprocessors. He has also led teams developing real-time data acquisition and control systems and adaptive intelligent design systems for Schlumberger. He has worked for several other multinational companies including Conoco, and Monsanto. Mr. Kelly has a Ph.D. in engineering from Texas A&M and a B.S. in engineering from Brigham Young University. He can be reached at bkelly@wainhouse.com.