A Web Conference System Architecture for the Broadband Era

Yoichi Kato
General Manager, Web Conference Systems
NTT-IT Corp.
“Web Conference System/Service” is gaining recognition in Japan. The term means a system with the following aspects:

- All software based (clients running on PC/PDA).
- Using IP networks.
- Multi-user.
- Audio.
- Video.
- Document Sharing.

“MeetingPlaza”: a web conference system widely used in Japan
A Typical Application: Meetings with branch offices and the head quarter

Purpose: Regular Meetings, Information Sharing, etc.
Number of locations: 3 – 30+  ~100
Number of participants/location: 1 - 10

Other Applications: Meetings among the members of an association
Learning for medical representatives
Typical Terminal settings: Personal

PC, Headset, Web Camera
Typical Terminal Settings: Group of 2-4

PC, In-expensive Mic/Speaker with EC, Web Camera
Typical Terminal Setting: Group of 5+

PC, Microphones, Mixer, Speaker, EC, Camera, Projector
Requirements on Networks

- Various Speed (20Kbps ~ 10Mbps)
  - 20Kbps~ Modem/ISDN/Mobile
  - 200Kbps~ ADSL/Cable
  - 2Mbps~ LAN/Fiber Optics

- Various Configurations
  - NAT/IP Masquerade
  - HTTP Proxy
  - F/W (port filtering)
    - Virus scanners...
What is important (user’s voices)?

- Most important medium: Audio
  - Without this, no communication established...
- Next: Document Sharing
  - In contemporary meetings, “document sharing” is essential.
- ....
- ....
- Necessity of video is less...
  - But comparing to Audio/Document conferences, users recognize it is more comfortable feeling “other people” in distance conferences through video.
  - To “feel” other people, small/slow image is enough in many applications.
Audio (most important medium)

- Robustness against:
  - Packet loss
    - ~10% packet loss observed in Japan-China connections these days...
  - Narrow bandwidth
    - 20Kbps ~

- Simultaneous talking (~ 4 persons)
  - Bandwidth usage needs to be low

- Audio Bandwidth (~ sampling freq.)
  - 8KHz -- Enough for most applications
  - 11KHz -- Foreign language lessons need at least this
  - 22KHz -- Special cases such as distance piano lesson (one example exists)
“Server-Client” vs. “Peer to Peer”

| connectivity in various network configurations | GOOD using HTTP Tunneling/Client-leading session initiation | No connectivity when both clients are under F/W |
| possible bandwidth bottleneck? | All traffic concentrate at the server | Receiving bandwidth increase as the number of client increase |
| administration (user management, usage logging, etc.) | Easy | Not Easy |

Hybrid architecture may be possible (further study needed)
Server mixing or Client mixing?

**Client Mixing**

- Audio stream A
- Audio stream B
- Audio stream C

Decode and mixed in client

**Server Mixing**

- Audio stream A
- Audio stream B
- Audio stream C

Decode/mix/re-encode at Server

Receiving plural audio stream at the same time (not usable with modem/ISDN/mobile)

Receiving only one audio stream always (Usable with Modem/ISDN/mobile)
Audio transmission architecture

- Server-Mixing ("N-1" mixing)
  - Avoiding to hear my talk with delay

Diagram showing a server mixing audio streams from multiple sources (A, B, C, D) into a single output stream.
**Video Transmission Requirements**

- Optimized for:
  - Various network speed (20Kbps ~ 10Mbps)
  - Heterogeneous network configurations
    - Terminals with high speed connections can transmit larger/fast video each other
    - Terminals with different network speed can communicate at small/slow video
    - The above two must be realized at the same time (see next slide)

- Robustness against:
  - Packet loss
  - Rapid bandwidth changing
    - Audio and document sharing data must have priority over video communication
Video transmission for multi-user communications

Video Coding Algorithms used in *MeetingPlaza*
- Frame drop
- Pyramid coding
- Conditional replenishment
- DCT/Variable length coding

Joint ITU-T Workshop and IMTC Forum 2006 “H.323, SIP: is H.325 next?”
San Diego, 9-11 May 2006
Document Sharing Functions

- Web Co-Browsing
  - URL/Input area/Scroll synchronized
- File Sharing
  - Slide page flip synchronization (PPT)
  - Scroll synchronization (text, images)
- Virtual Printer Image Sharing
  - Documents of any application with “print” menu can be shared
  - Scroll/Scaling/Page synchronization
- AP Sharing (display image sharing)
- Shared Whiteboard
- Annotation
  - Drawing on shared documents
- Cursor sharing

See the demonstration
Demonstration

- Live
  - Server: NY

- Replay pre-recorded session
  - Video performance
What’s the most crucial change from H.32x?
It’s: “All Software Based”

- The “Platform” of the system is:
  - Generic Personal Computer/PDA
  - The Internet or IP-based private networks
  - Server (may not be used always)

- PC or PDA works as “terminal”.

- “Software” is available when needed.
  - “Downloadable”, for example.

Terminals can communicate each other
(because they run the same software)

No protocol level standardization needed...
What can be “standardized”?

- Definition of typical conceptual system structure
  - Example:
    - Category 1 Web Conference System: peer to peer based
    - Category 2 Web Conference System: server-client
    - Category 3 Web Conference System: mixed

- Definition of capabilities
  - Like xx@yy in MPEG

- ...

- ...
Further Information

- **Web:**
  - [www.meetingplaza.com](http://www.meetingplaza.com)

- **Contact:**
  - Yoichi Kato (yoichi@ntt-it.co.jp)
  - NTT-IT Corp.

  - Phone: +81 45 651 7555
  - Fax: +81 45 224 6799