Videoconferencing: standards update and industry trends

Roberto Flaiani
AETHRA
June 3rd, 2004

ITU standards for videoconferencing

- **H.320 ('90)**
  - → **ISDN, V.4 approved 2004**
- **H.324 ('96)**
  - → **POTS, ISDN, mobile networks; V.3 approved 2002**
- **H.323 ('96)**
  - → **Packet networks, V.5 approved 2003**
- **H.310, H.321**
  - → **ATM networks**
  - → **Not in use any more**

H.350, H.239

- **H.350 - Directory Services Architecture for Multimedia Conferencing ('03)**
  - ➞ **H.350.1 - DSA for H.323**
  - ➞ **H.350.2 - DSA for H.235**
  - ➞ **H.350.3 - DSA for H.320**
  - ➞ **H.350.4 - DSA for SIP**
  - ➞ **H.350.5 – DSA for non-standard protocols**
- **H.239 – "Role management and additional media channels for H.300-series terminals" ('03)**
  - ➞ **Powerful technology which will likely replace data conferencing in most room environments**
    - **T.120 is complex and problematic in large conferences**
    - **WEB conferencing is plagued by lack of standardization**

Latest ITU work

- **H.320 V.4 (Jan ’04)**
  - ➞ **Support for H.264, H.239 (“dual video”)**
  - ➞ **Audio MPEG4 AAC**
  - ➞ **Enhanced security: AES (H.233, H.234)**
- **H.350.6 - DSA for Call Forwarding and Preferences”**
- **H.qosarch (QoS Architecture)**
- **H.460 series Recs:**
  - ➞ **H.460.9 Annex B (Extended reporting)**
  - ➞ **H.460.10 (Call Party Category)**
  - ➞ **H.460.11 (Delayed Call Establishment)**
  - ➞ **H.460.12 (Glare Control Indication)**
  - ➞ **H.460.13 (Called User Release Control)**
  - ➞ **H.460.14 (Multilevel Precedence and Preemption)**
  - ➞ **H.460.15 (TCP Channel Suspension)**

Coding standards

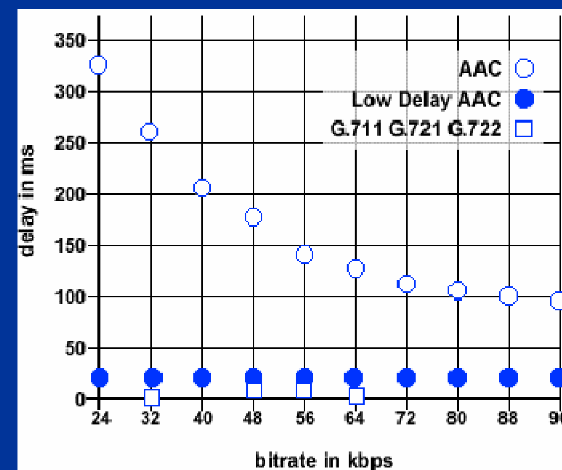
- **Narrowband audio (4 kHz bw)**
  - → **G.711** 64 kbit/s (56)
  - → **G.728** 16 kbit/s; 9.6, 12.8 kbit/s
  - → **G.729** 6.4, 8, 11.8 kbit/s
  - → **G.723.1** 5.3, 6.3 kbit/s
- **Wideband audio (7 kHz bw)**
  - → **G.722** 48, 56, 64 kbit/s
  - → **G722.1** 16, 24, 32 kbit/s
  - → **G.722.2** 6.6, 8.85, 12.65, 14.25, 15.85, 18.25, 19.85, 23.05 and 23.85 kbit/s
- **High quality audio (15 kHz bw)**
  - → **MPEG4 AAC LD** 64 kbit/s stereo 7kHz, 128 kbit/s stereo 15 kHz

G.722.2

- **Supported only in H.323 systems**
- **Aligned with 3GPP AMR-WB codec**
- **Built in VAD (Voice Activity Detection) and CNG (Comfort Noise Generation)**
- **-For clean speech signals, G.722.2 is**
  - ➞ **At 14.55 kbit/s, equal or better than G.722 at 64 kbit/s**
  - ➞ **At 12.65 kbit/s, at least equal to G.722 at 56 kbit/s**
  - ➞ **At 8.85 kbit/s, equivalent to G.722 at 48 kbit/s**
- **For music signals, G.722.2 at 23.85 kbit/s is equivalent to G.722 at 48 kbit/s**

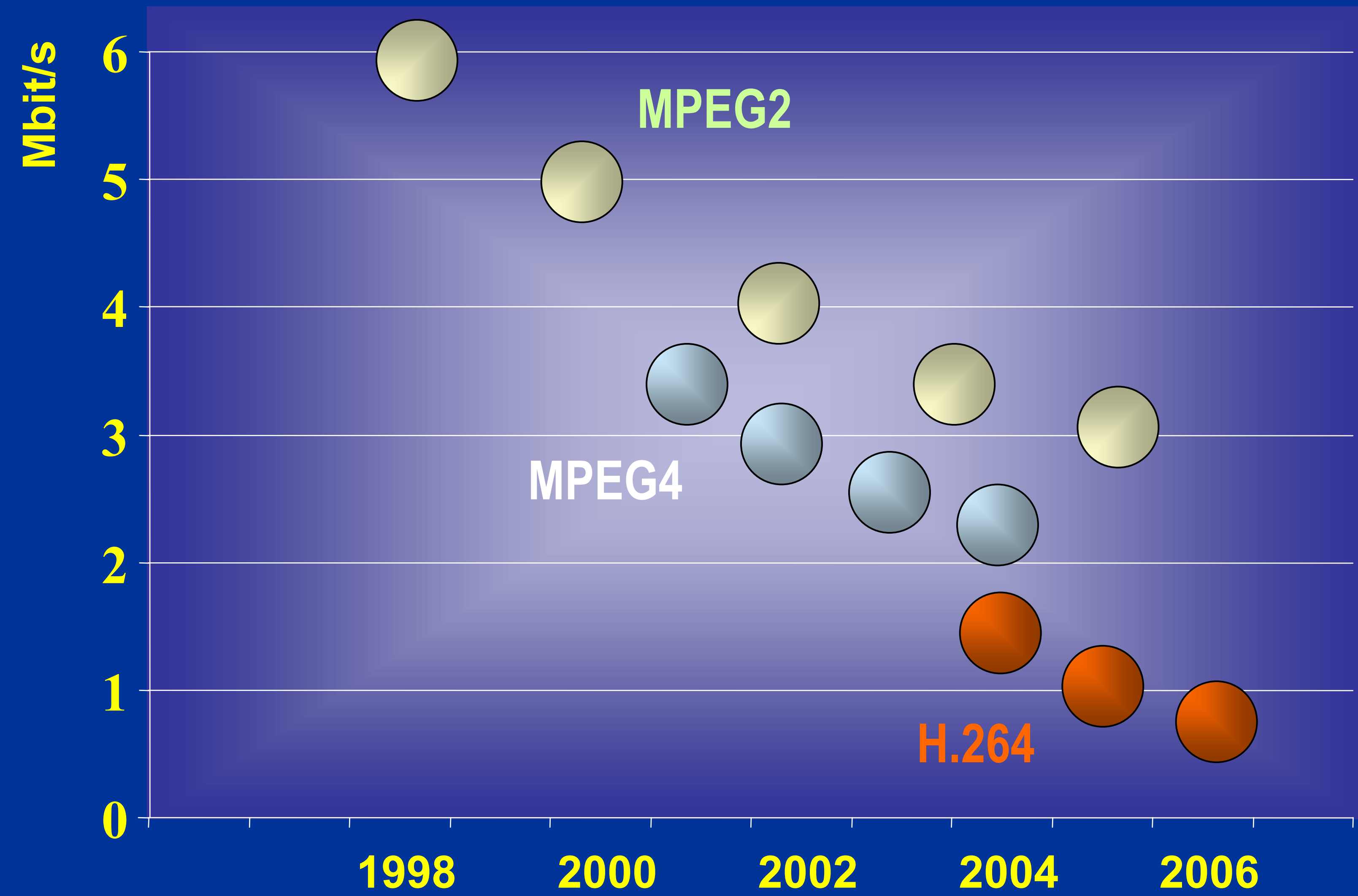
MPEG4 AAC-LD codec

- **Derived from MPEG4 AAC**
  - ➞ **AAC achieves "indistinguishable quality" for stereo signals at 128 kbit/s**
  - ➞ **"CD quality" (stereo) at 96 kbit/s**
  - ➞ **Delay > 100 ms**

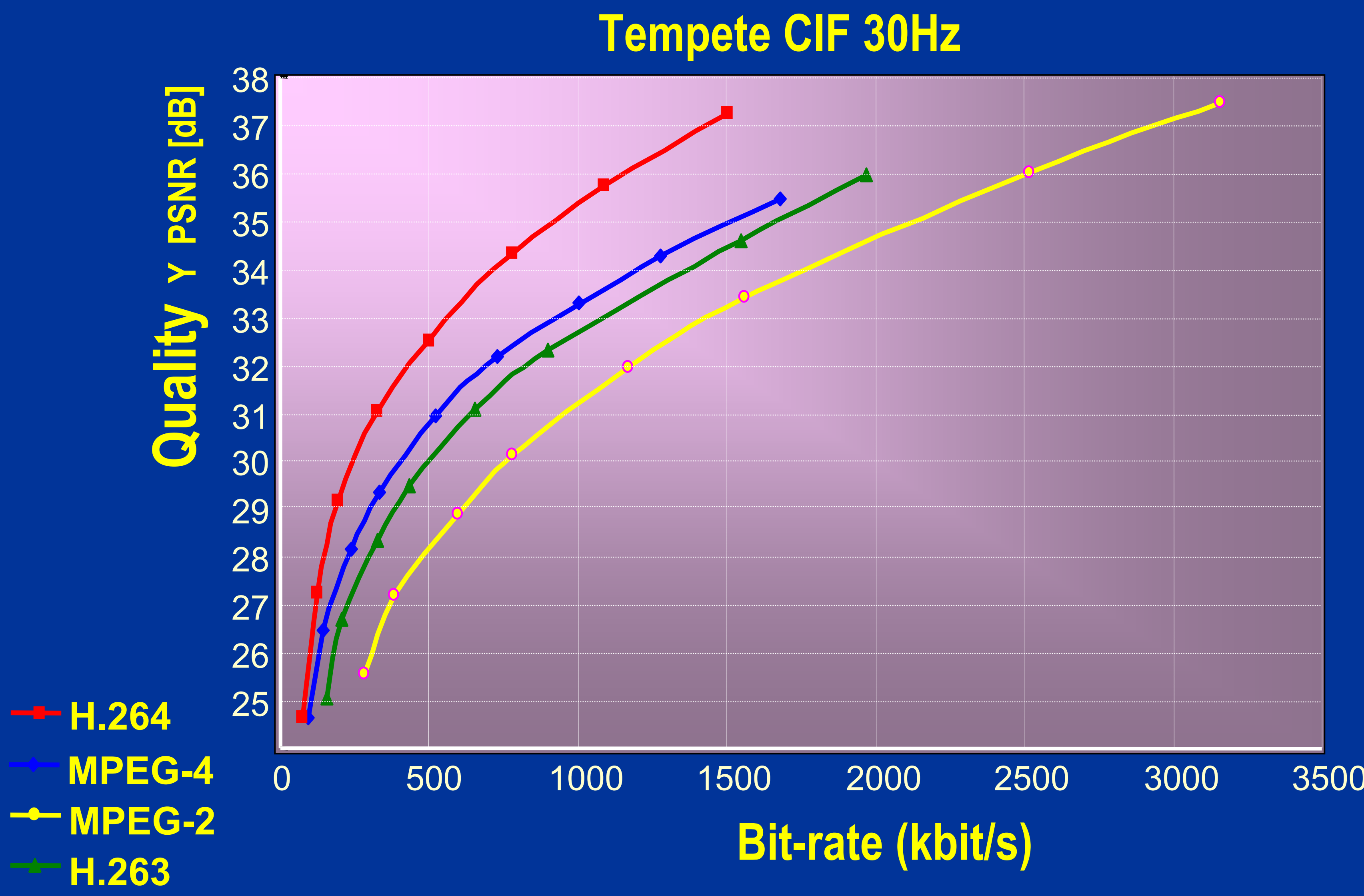


- **AAC-LD**
  - ➞ **Algorithmic delay: 20 ms**
  - ➞ **"Real world delay"; as low as 30 ms**
  - ➞ **At 32 kbit/s, same quality of AAC at 24 kbit/s**
  - ➞ **Always equal or better than MP3 at the same rate**

TV quality bandwidth



Codec comparison



H.264

- **H.264/MEG4 AVC gaining momentum**
  - ➞ **Selected by 3GPP, DVD forum, Jap. Broadcasters**
- **It may bridge wired and wireless networks, television and videoconferencing**
- **Complexity**
  - ➞ **Encoder ~ 6 times more complex than MPEG4 SP**
    - **~ 4 times more complex than MPEG2**
  - ➞ **Decoder 2 ÷ 3 times more complex than MPEG4 SP**
  - ➞ **Real complexity depends on features implemented**
- **Interoperability tests in the initial stage**

H.264 implementors

- Aethra
- Ahead Software / ATEME
- Amphion
- Apple Computer
- British Telecom
- Broadcom / Sand Video
- Conexant
- Deutsche Telekom
- DG2L
- Dicas
- DSP Research
- Emblaze Group
- Envivio
- Equator
- FastVDO
- France Telecom
- Hantro
- Harmonic
- HHI
- i3 Micro Technology
- iVast
- Intel
- KDDI R&D Labs
- Ligos
- LSI Logic / Videolocus
- Mainconcept
- Mcubeworks
- Media Excel
- Mobile Video Imaging
- Mobilygen
- Modulus Video
- Moonlight Cordless
- Motorola
- Neomagic
- Nokia
- Oki Electric
- Optibase
- Packetvideo
- PixelTools
- PixSil Technology
- Polycom
- Prodys
- Radvision
- Richcore
- Samsung
- Scientific Atlanta
- Setabox
- SkyStream Networks
- Sony
- ST Micro
- Tandberg
- TandbergTV
- Tektronix
- Techno Mathematical
- Telesuite
- thin multimedia
- Thomson
- TI
- Toshiba
- Tuxia
- UB Video
- Videosoft / Vanguard
- VideoTele.com
- VCON
- Vqual

New directions in audio video

- **Video**
  - → **H.265** **(4 to 6 years away)**
  - → **“3D” video**
    - **Panoramic (from one viewpoint in every direction)**
    - **Interactive stereo (one view for each eye)**
    - **Free viewpoint (N cameras; N small or large)**
    - **3D synthetic video**
- **Audio**
  - → **Lossless** **(scalable) coding**
  - → **VBR** **coding**

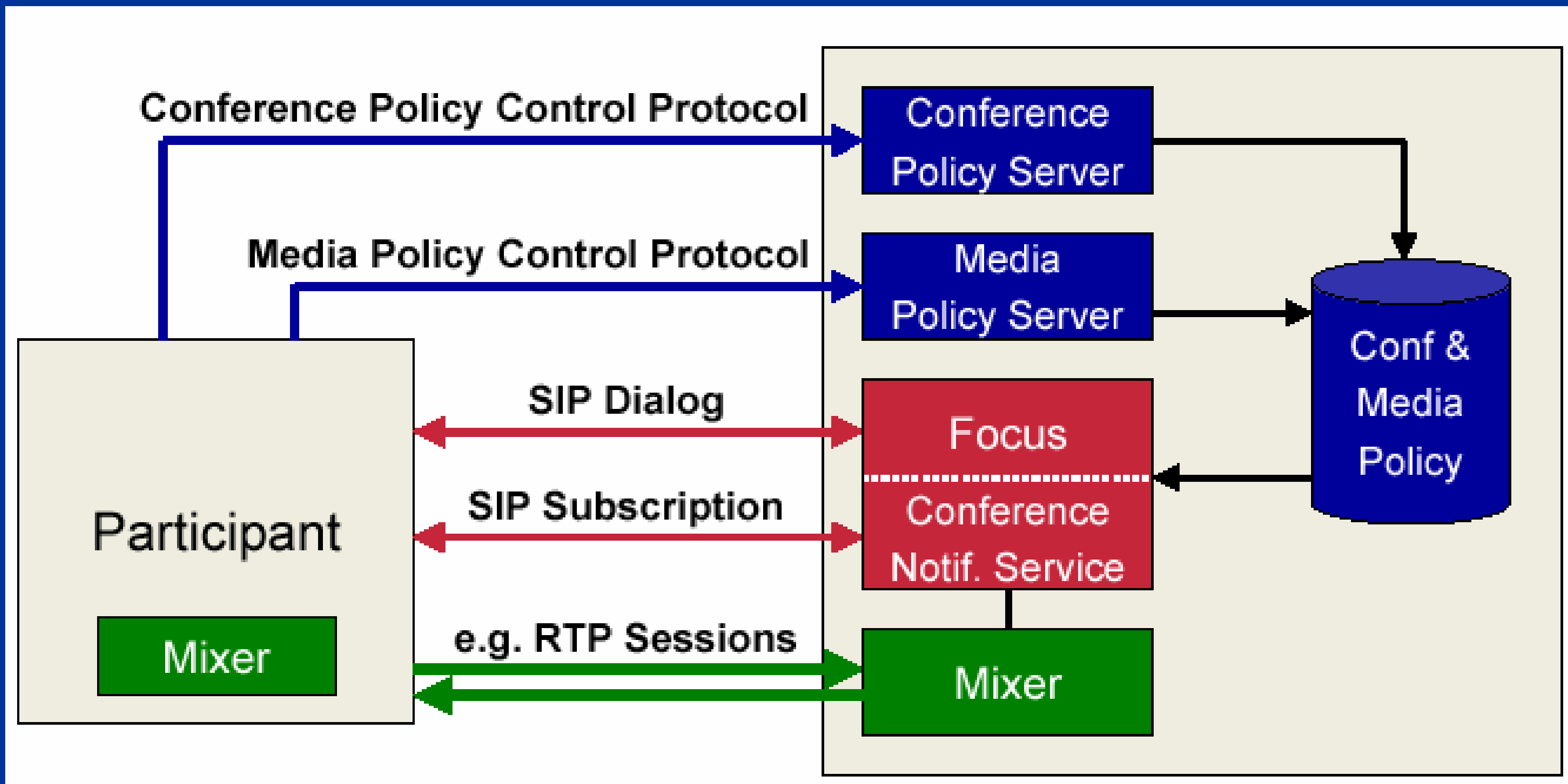
SIP for videoconferencing

- **The SIP community is addressing how to use SIP in tightly-coupled multiparty conferencing, including "advanced" features such as sidebars and media policy manipulation**
  - → **Some mechanisms could be used by non-SIP appl.**
- **SDP, originally designed for multicast session announcements (rather than unicast interactive negotiations), is still problematic when trying to express capabilities:**
  - → **No way to indicate allowed parameter range or constraints (e.g. simultaneous codec supported)**
  - → **Difficult to express video codecs resolution, frame rate, bit rates supported, options (annexes) available**

SIP conferencing docs.

- **draft-ietf-sipping-conferencing-requirements-00**
- **draft-ietf-sipping-conferencing-framework-01**
- **draft-ietf-sipping-conference-package-02**
  - ➞ **Allows users to subscribe and to be notified about changes in the conference (membership, status of users, sidebars).**
- **draft-ietf-sipping-cc-conferencing-03**
  - ➞ **Explains how to create a conference using SIP only methods, allowing participation to conference aware or unaware UAs.**
  - ➞ **Defines the functionality of a "conference factory" enabling the automatic creation of ad-hoc conferences**

XCON model



XCON is working on conference and media control

XCON WG

- **Documents produced so far:**
  - **draft-ietf-xcon-conference-scenarios-00.txt**
  - **draft-ietf-xcon-cpcp-reqs-03.txt**
  - **draft-ietf-xcon-floor-control-req-00.txt**
  - **draft-ietf-xcon-cpcp-xcap-00.txt**
- **Design team active on floor control protocol**
- **XCON/SIPPING interim meeting last week in Boston**

SIP in the video community

- **SIP popularity has been boosted by adoption in 3GPP standards and MS Messenger/LCS**
  - ➞ **Even though videoconferencing in 3GPP is mostly done in circuit mode with H.324M**
- **Better integration with "presence" and IM may be a driver for a shift to SIP, starting from the desktop**
- **SIP is currently less complete and mature than H.323/H450.x for sofisticated suppl. services**
  - ➞ **But in applications such as "enterprise video PBX" or "video call centers" SIP is attracting developers for its simplicity and low implementation cost**

SIP in the video community

- **SIP has enjoyed limited acceptance so far in the video community, but this is going to change soon**

- **Unfortunately, adoption of SIP in CPE and network equipment is not going to improve interoperability:**
  - **➞SIP as a protocol is less mature than H.323**
  - **➞H.323 is more precisely defined**
  - **➞SIP-H.323 gateways will add complexity**

NAT – Firewall traversal

- **It is widely recognized as a big problem**
- **It affects both signalling and media traversal**
- **Various solutions have been proposed, each with its own drawbacks/limitations:**
  - → **Application awareness in middleboxes: ALGs**
  - → **Relay servers: proxies, reflectors (STUN, TURN)**
  - → **Middlebox configuration (MIDCOM, UPnP…)**
  - → **Middlebox traversal by tunneling**

NAT – Firewall traversal

- **The IETF has tried to solve the problem**
  - **➞MIDCOM failed to fulfill expectations**
    - **Only STUN promoted to RFC**
- **Something new is being developed in NSIS (“Next Steps in Signalling”):**
  - **➞IETF draft: “A NAT/Firewall NSIS Signaling Layer Protocol (NSLP)”**
  - **➞Difficult to say if it will fly at this stage**

Short term study items in ITU

- **Signalling traversal**
  - **→RAS over TCP**
    - **The assumption is that once a connection is established through the NAT then messages from the outgoing destination address can flow back**
    - **If RAS and H.225.0 share the same (persistent) connection, after registration an EP can be reached from outside**
- **Media traversal**
  - **→Support of ICE**
- **ITU has opened a new Question (Q4/16) to study NAT/firewall traversal related issues**

Endpoints state of the art today

- H.323 and H.320, >2 Mbit/s
- H.264, H.263, H.261
- Video res. QCIF, CIF, 4CIF, XGA
- MCU 7 sites + Gateway
  - → Audio and video transcoding
- H.239 (2 video flows)
- XGA input, XGA output
- Dual screen support
- Embedded PowerPoint pres.
- Voice tracking

- AES encryption
- WiFi
- Streaming

What are we going to see next?

- **Multistandard terminals:**
  - → **H.320, H.323, SIP**
- **TV quality is the target for video**
  - → **With H.264, achievable at around 1.5 Mbit/s**
  - → **But latency problems still unaddressed by H.264**
  - → **Transcoding and cascading get the problem worse**
- **Advanced audio codecs, stereo**
  - → **MPEG 4 AAC-LD, stereo echo cancellation**
- **Improved security**
- **More telephony features (support of advanced PBX functionality)**

What are we going to see next?

- **Tighter integration with presence, IM and directory services**
- **MCUs:**
  - → **Dual video**
  - → **Advanced control**
  - → **Personalized layouts**
- **Next generation protocol not on the horizon yet**
- **What about the consumer side of the story?**
  - → **Endpoint at an attractive price point possible now**
  - → **Interest from providers of triple play services**
  - → **What will the consumer reaction be? VC on mobile phones could set the stage.**