Recent developments in QoE aspects of videoconferencing services

Patrick Luthi, TANDBERG
Rapporteur for Q.1/16
(multimedia systems, terminals and data-conferencing)

QoE aspects of videoconferencing

- Recent developments in SG16 significantly enhanced the videoconferencing user's quality of experience (QoE).
- Standards defining new features or codecs such as:
  - dual video streams (H.239)
  - video coding (H.264)
  - audio coding (AAC-LD, G.722.1)
  - NAT/firewall traversal methods (H.460.18,19)
  - H.324m call setup acceleration techniques (H.324 Annex K)
- Improved perception of quality of a video call even if end to end QoS is not guaranteed when using the Internet as a transport network.

ITU-T

Dual video streams

- Best feature of the decade! (IMTC 2002 Spring Forum)
- H.239 - Role management and additional media channels (AMC) for H.300-series terminals
  - Use of more than one video channel in H.320-based systems
  - Indication of video bitrate supported for AMC and in-call bitrate management
  - Labelling of individual channels with a "role“ applicable to H.320 and H.245 signalling-based systems
- Collaborative standardization effort
  - ITU identified the need to standardize a way to transport multiple media streams in H.320 in February 2002
  - First approved in July 2003, revised in Sept. 2005
  - For the future: multiple audio channels in H.239 for multi-lingual and simultaneous translation applications

ITU-T

Video coding – H.264

- H.264 considerably increased coding gain
- Significantly better than H.263 and MPEG-2 - same video quality at half of the bitrate
- Enhanced error and packet loss resilience
- HD ready
  - Maximum picture size: 4096 x 2048 (variable aspect ratio)
  - bandwidth requirement of 1 Mbit/s is sufficient if the picture has little motion
  - Example: HD MPEG-2 content (1920x1080) traditionally runs at 12-20 Mbps. H.264 can deliver 1920x1080 content at 7-8 Mbps at the same or better quality.
- Computational efficiency with new Annex B/H.241 which specifies a reduced-complexity decoding process to be applied to H.264 bitstreams

ITU-T

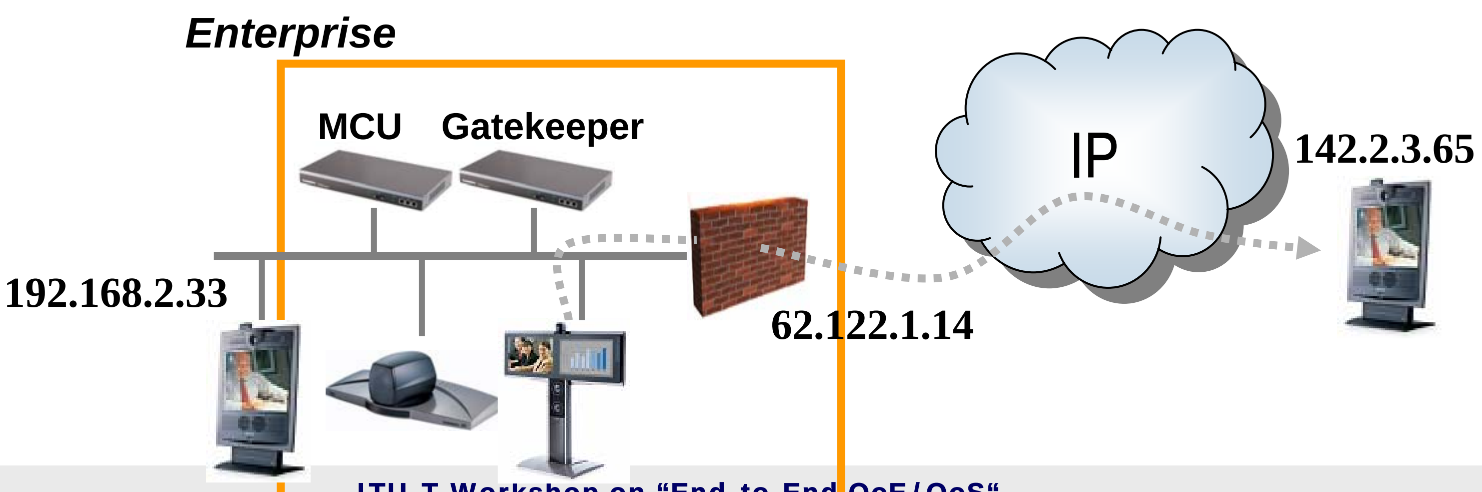
Audio Coding – "super" wideband

- Improved intelligibility and transparency
- Better music and sound effect quality
- Users experience less fatigue during video meetings
- Low Delay Advanced Audio Coding (AAC-LD)
  - high quality low delay audio coding standard of MPEG-4.
  - closely derived from MPEG-2 Advanced Audio Coding (AAC).
  - algorithmic delay of only 20 ms.
  - supports e.g. 20 kHz at 64 kbit/s or stereo 20 kHz at 128 kbit/s
- G.722.1 Annex C
  - extension to G.722.1 (7 kHz audio).
  - low complexity
  - algorithmic delay of 40 ms.
  - supports 14 kHz at 24, 32 and 48 kbit/s.

Network Address Translation: the Challenge

ITU-T

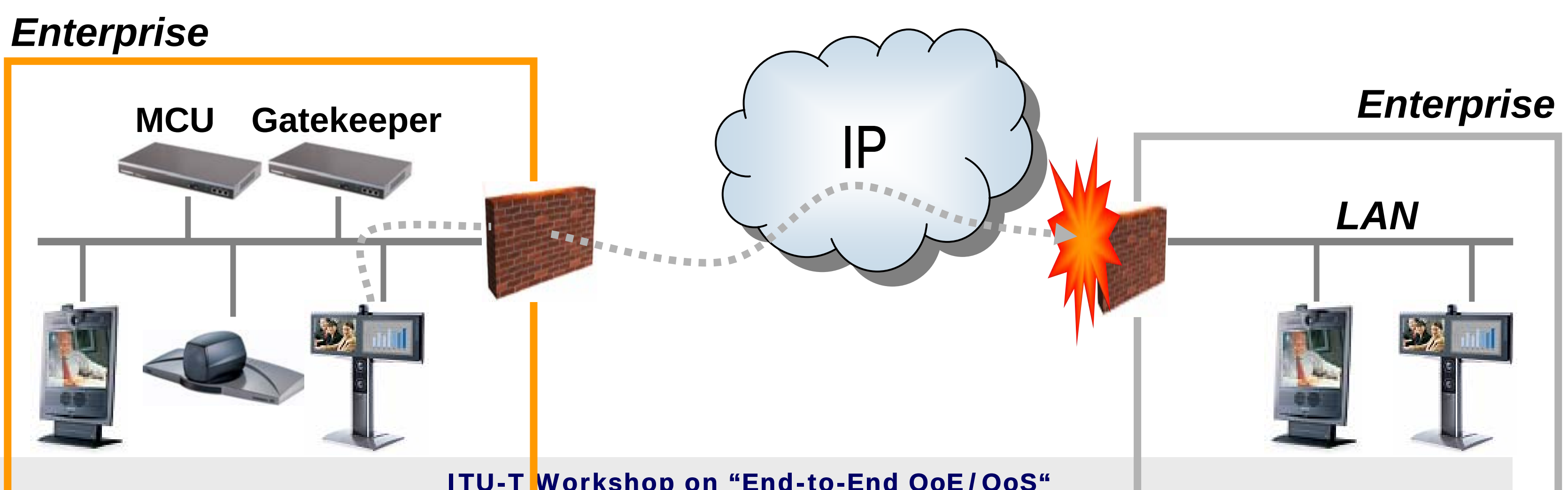
- Organizations use private addressing schemes and share a public address
- Outside terminals cannot access to those private addresses
- Translation function is widely exploited as a security feature
- Address translation typically applies to packet headers, but not to the protocol within the packets (H.245, etc.)



ITU-T

The Firewall Problem

- Rules that allow everyone to connect to everyone else are unusual and unwelcome to the security administrator.
- Unsolicited incoming connections are typically not allowed
- A firewall can be “opened” for video calls, but results in either loss of features (such as encryption) or reduced security
- IP communication protocols use a wide range of network ports



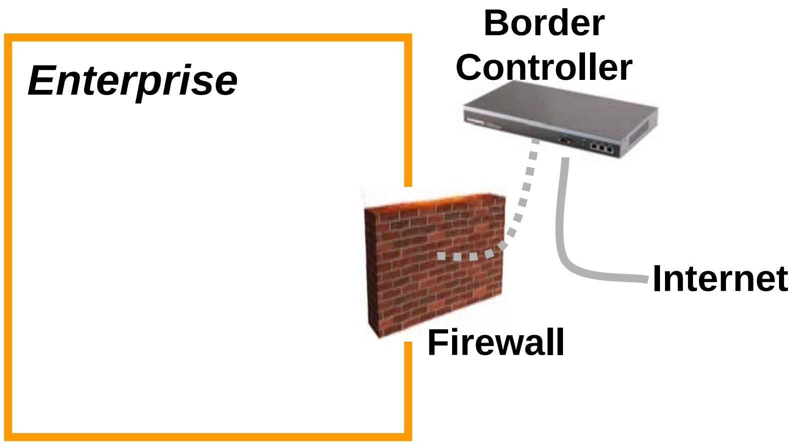
What solutions?

- No firewall, no NAT (public IP address)
- VPN to connect separate locations
- ISDN Gateway
- Proprietary protocols
- Stand-alone or Gatekeeper Proxy
- Application Level Gateway
- MIDCOM, a protocol to let an outside box issue commands to open and close ports on the firewall
  - Complex and unproven standard still in development
- ICE (Interactive Connectivity Establishment), a methodology for NAT traversal
  - Still in development, makes use of existing protocols (STUN, TURN, RSIP)
  - Works only for SIP

The Solution: A Border Controller

- Secure tunnelling of H.323 calls through any firewall
- No features are lost – works with H.264, MPEG-4, AES, H.239, etc.
- Border controller provides traversal for ALL other H.323 endpoints and MCUs

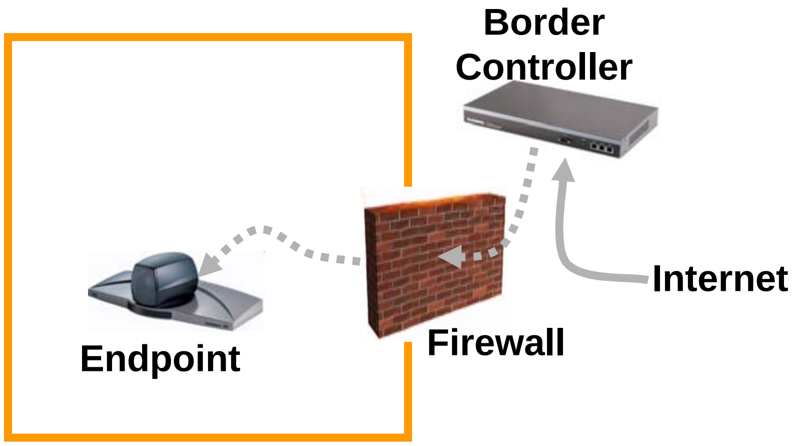
- The Border Controller allows secure traversal through firewall
- The Border Controller might be hosted by a service provider, or hosted in an enterprise DMZ along with the enterprise mail and web proxies.



The Solution: Enhanced Endpoints

- Embedded in endpoints
- Allows secure traversal of ANY firewall
- The firewall only needs to allow connections between the solution components
- The Border Controller and endpoints are designed to use a very small number of registered ports.

- Enhanced Endpoint and Border Controller create a route through the firewall
- The outside world calls through the Border Controller



ITU-T

The standardized solution: H.460.18 and 19

- Collaborative standardization effort
  - Approved in September 2005
  - What took so long? IETF kept promising, but couldn't deliver because of lots of "religious" disagreements
  - Key decision for ITU 2005-08 Study Period: Focus narrowly on H.323 solution
  - Based on TANDBERG's Expressway technology

The standardized solution: H.460.18 and 19

- H.460.18 - Traversal of H.323 signalling across network address translators and firewalls
  - H.323 signalling traversal & call setup
  - Is mandatory
  - All traffic originates inside NAT/FW boundary
  - Port symmetry lets response to pass through NAT/FW
  - This opens bi-directional “pinhole” through NAT/FW
  - Keep-alive packets sent periodically to keep pinholes open
- H.460.19 - Traversal of H.323 media across Network Address Translators and Firewalls
  - H.323 media traversal
  - Depends on signaling traversal mechanism such as H.460.18
  - Uses similar principles as H.460.18 (pinholes and keep-alive packets to maintain path)

Call setup acceleration in H.324m (Annex K)

ITU-T

- Motivation comes from mobile world (3g-324m)
- A 3G video call can take a long time…
  - ~8 seconds to set up bearer (pre-ringing)
  - Ringing and answering time
  - 4 to 6 seconds for “call setup” (H.324)
    - Setting up multiplex level (H.223)
    - Exchange caps, configure mux, open audio/video channels (H.245)
    - Send / Receive / Render initial Audio & Video

ITU-T

Call setup acceleration – The MONA solution

- Annex K/H.324 – Media oriented negotiation acceleration (MONA)
- Exchange of “fast call setup” capabilities and preferences (“MONA Preference” Messages)
- Quick set up of audio and video channels
- Maintains full compatibility with legacy terminals (using regular H.245 – not accelerated)
- Expected performance (MONA-to-MONA case)
  - Time to receive media ranges 0.5 to 1.5 RT
  - “Typical” RT = 800 ms
  - So call setup may range 400ms to 1.2 sec

ITU-T

Call setup acceleration contributing techniques

- Media Preconfigured Channels **(MPC)**
  - Small table of commonly used codec + mux configurations
  - Early-bearer may be used to send media
- Signaling Preconfigured Channel **(SPC)**
  - Early-bearer exchange of capabilities/prefs + inference model
  - Preserves full flexibility of H.245 channel establishment
- Accelerated H.245 Procedures **(A2P)**
  - Media can be sent without waiting for OLC and MES exchanges
  - Implemented as minor changes to existing H.245 procedures

FM

FSS

ACN

Mapping to original proposals

Thank you!