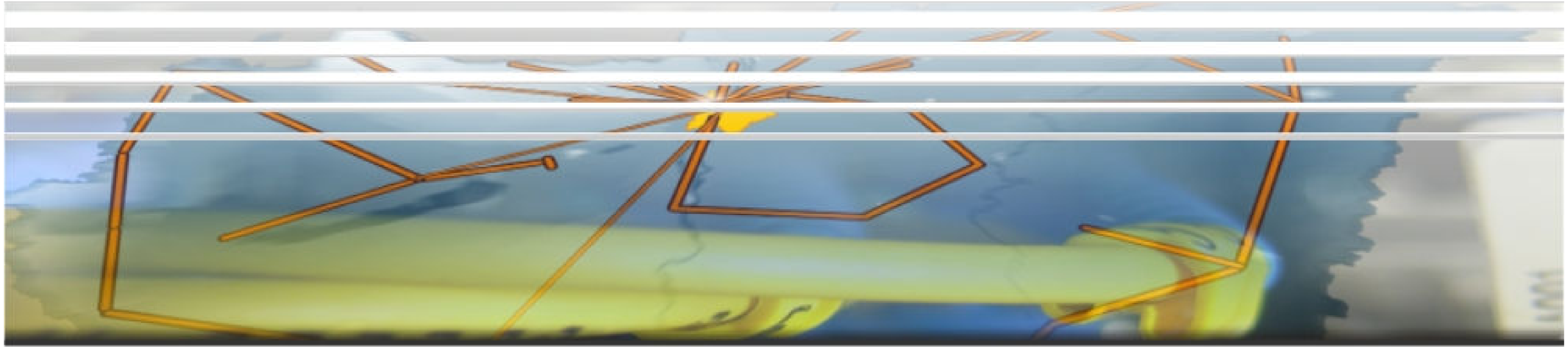
Protocol primer: H.323 & SIP



András Kovács

NIIF/HUNGARNET

<akov@niif.hu>

VICES Project partner visit

Budapest

27 May 2009

Content

- H.323 protocol overview
  - Protocol basics
  - Network building blocks
  - Gatekeepers & addressing
  - MCUs
  - Firewall considerations
- SIP overview
  - Protocol basics
  - Network building blocks
  - Addressing (ENUM)

Videoconference history

- 1956: AT&T Picturephone
- 1982: CCITT H.120 (2Mbit/sec codec)
- 1984: PictureTel VC terminal ($80.000)
- 1990: CCITT H.320 ISDN VC & H.261 codec
- 1992: RTP/RTCP v1 protocols (IETF)
- 1996: ITU-T H.323 v1 & H.263 codec
- 1997: VRVS (Caltech-CERN)
- 1999: SIP → IETF Proposed Standard
- 2000: ITU-T H.323 v4
- 2001: NTT DoCoMo 3G WCDM videophone ($570)
- 2003: ITU-T H.264 (MPEG4) codec
- 2004: H.239 *GA graphics transmission (VGA, XGA, etc.)
- 2004: SIP videoconference implementations

H.323

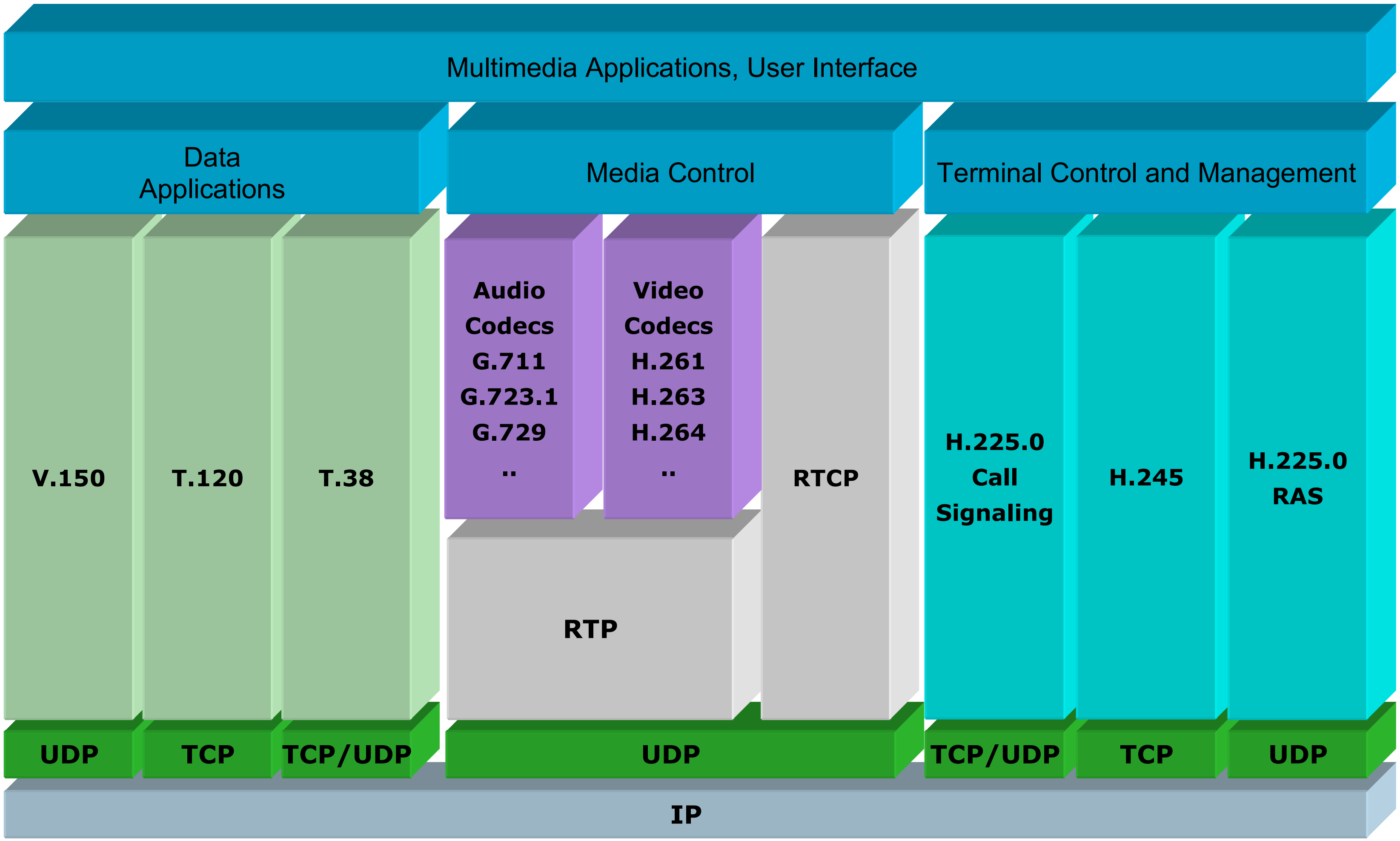
H.323 basics

- ITU-T recommendation (1996)
- Definition: multimedia conferencing protocol for packet switched networks allowing multipoint audio, video and data conferencing
- Family of H.* communication protocols:
  - H.320: N-ISDN
  - H.321: B-ISDN (ATM)
  - H.324: PSTN
- H.32x: “umbrella” recommendations:
  - Refers to various number of protocols
  - H.32x version
  - Annex, Appendix

H.323 & related standards

- H.323: foundational recommendation
  - Architectural principles
  - Reference to related standards
- H.225.0:
  - Call Signaling (Q.931) → ISDN
  - RAS (Registration, Admission & Status)
- H.245: multimedia control protocol
  - Terminal capability negotiation (codecs, datarate, etc.)
  - Alteration of call parameters (in-call)
- All protocol messages: ASN.1 encoded
- +10x related protocol/codec recommendations
- IETF: IP, UDP, TCP, RTP/RTCP

H.323 – protocol stack



H.323 – additional standards

- T.120 family
  - T.12x → Whiteboard, Chat, File Transfer, ...
  - Died out, very complex configuration
- H.239 Dual Stream – VGA graphics transmission in parallel to audio/video
  - Approved: 2004
  - Transmission of a presentation
  - VGA, SVGA, XGA, SXGA resolutions
- T.38 – Fax over IP
- V.150 – Modem over IP
- H.235 – Security framework (rarely used)
- H.323 Annex Q - Far End Camera Control

Audio codecs

- G.711 – PCM 3.1 KHz @ 64Kbps (normal telephony)
- G.722 - ADPCM 7KHz @ 48, 56, and 64Kbps
- G.722.1 - 7kHz @ 24 and 32Kbps
- G.723.1 - Dual rate speech codec @ 5.3 and 6.3Kbps
- G.728 – LD-CELP speech codec @ 16Kbps
- G.729 – CS-ACELP speech codec @ 8Kbps
- Vendor specifics:
  - Polycom Siren7: 7KHz @ 24 & 32Kbps → G.722.1
  - Polycom Siren14: 14KHz @ 24, 32 & 48Kbps → G.722.1 Annex C
  - Polycom Siren22: 22KHz mono/stereo @ 64, 96 & 128Kbps → G.719
  - ISO: MPEG-4 AAC-LD (Low Delay audio coding)

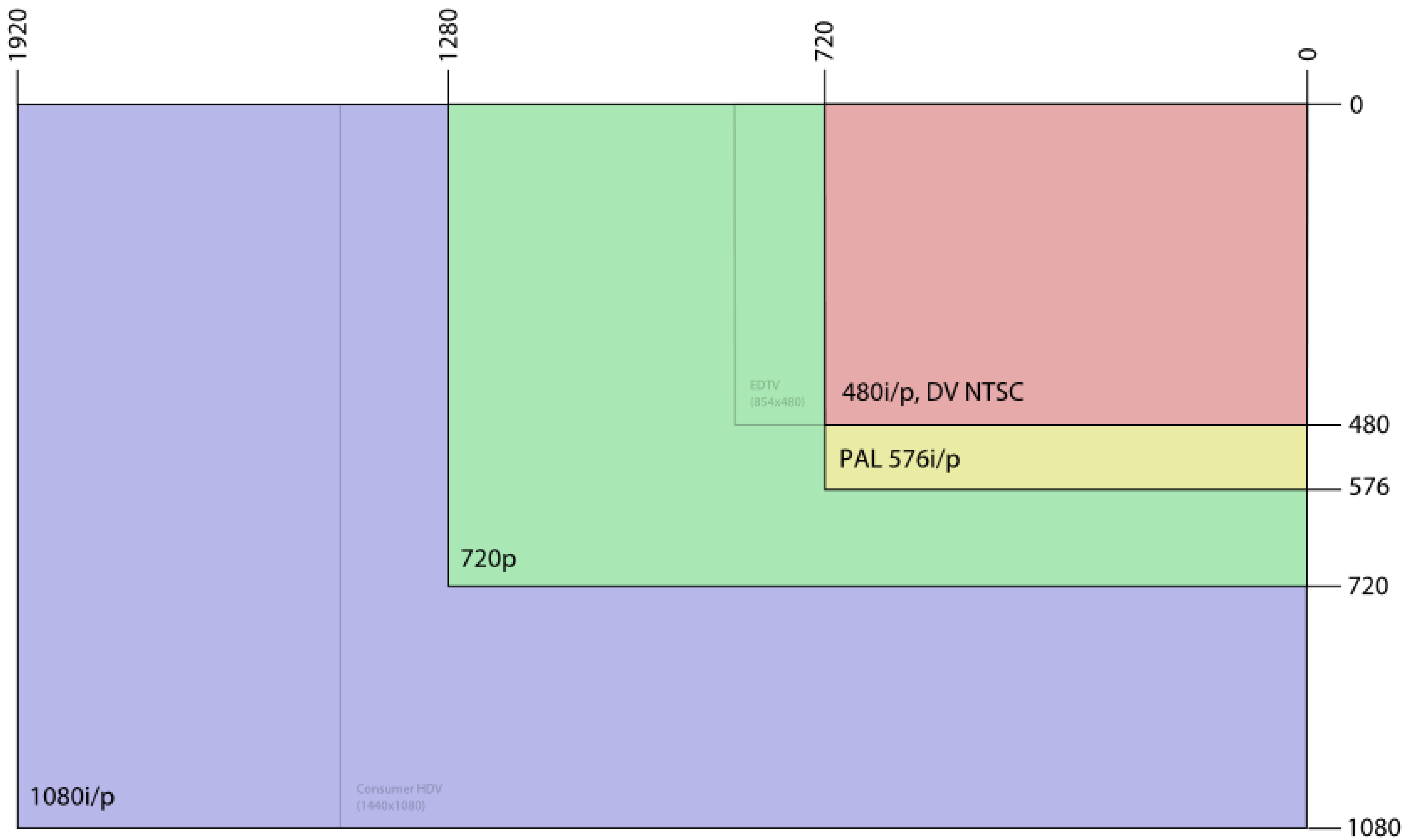
Video codecs

- **H.261** - Video codec for audiovisual services at Nx64 Kbps
  - MPEG1 based
  - Mandatory in all H.323 video terminals
  - N = 1…M → ISDN B channels
- **H.263** - Video Coding for Low Bitrate Communication
  - MPEG2 based
  - Half bandwidth, same quality as H.261
  - 2x computing complexity to H.261
  - Resolutions up to SD
  - Additional annexes: v2 (=H.263+), v3 (=H.263++)
- **H.264** – Very low speed, good quality
  - MPEG4 based
  - Half bandwidth, same quality as H.263
  - 4x computing complexity to H.263
  - Resolutions:
    - Up to HD 1080i/p
    - Any resolution is allowed

Typical videoconference resolutions

Resolution	Name	Codec
176 x 144	QCIF	H.261 & H.263 mandatory
128 x 96	SQCIF	H.263 mandatory
352 x 240	SIF (Source Input Format)	
352 x 288	CIF (Common Intermediate Format)	H.261, H.263 optional, approx. VHS quality
704 x 480	4SIF	H.263 optional (NTSC)
704 x 576	4CIF	H.263 optional (PAL)
1280 x 720p	720p	H.264
1920 x 1080i/p	1080i/p	H.264
512 x 288	w288p	*Tandberg*
576 x 448	448p	*Tandberg*
768 x 448	w448p	*Tandberg*

Evolution to HD



(source: wikipedia)

H.323 network elements - Terminals

- Telepresence
- Phone/Videophone
- IVR systems
- Voice Mail
- Softphone (pl. NetMeeting/EKIGA)

H.323 network elements - MCUs

- Multipoint Control Unit (MCU)

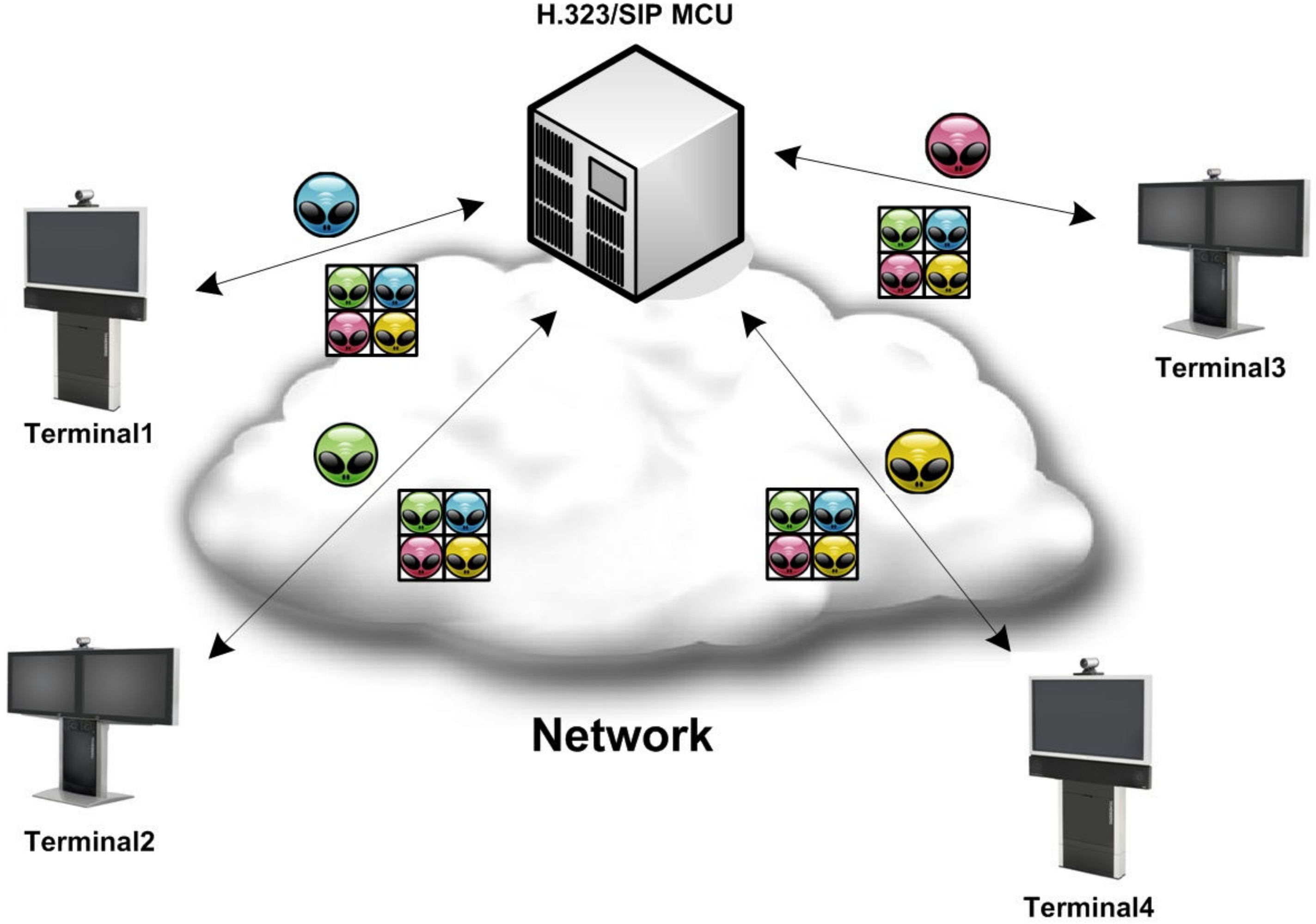
  - Multipoint conferences (audio/video)
  - Terminating several p2p (terminal-MCU) calls
  - Parallel conferences
  - Mixing and switching media (audio, video & data)
  - Bridge terminal capability differences (transcoding)

H.323 network elements - MCUs

- (contd.)
  - Real time video/audio decode → compose → encode
    - High capacity & quality needs HW (DSP)
    - SW MCU: low capacity & quality
  - Conference types:
    - Voice Switching: Based on volume
    - Continuous Presence: terminals' video composed on one screen
  - Other functions (e.g. IVR, webconferencing, streaming, recording, etc.)

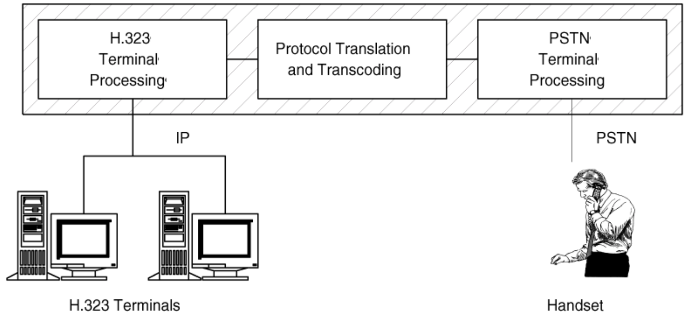
H.323 network elements - MCUs

- Example: CP multipoint composition



H.323 network elements – Gateways & Gatekeepers

- Gateway:
  - Traversal to other networks (e.g. ISDN, SIP, etc.)
  - Can be SW or HW,
    MCU component, etc.



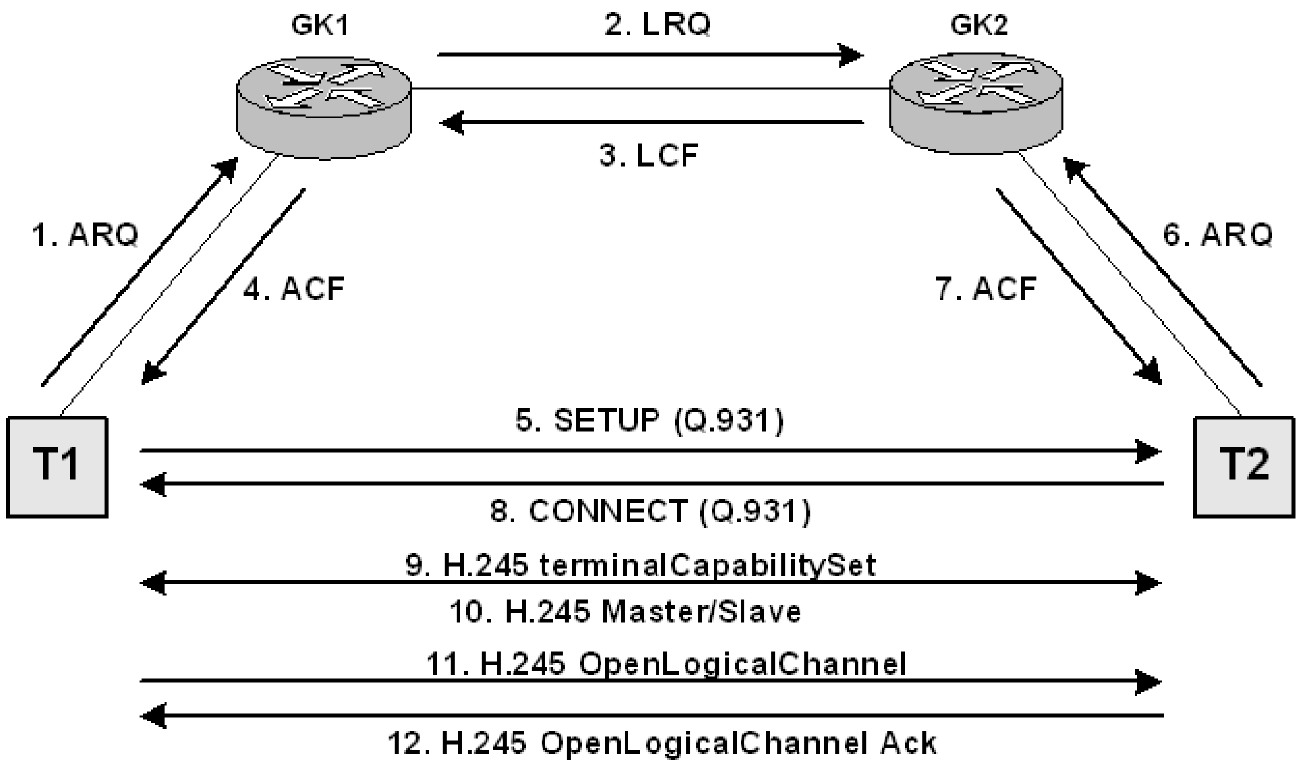
- Gatekeeper: optional but important
  - Address resolution (e.g. e.164, name → IP)
  - Call Routing/Call Admission/Call Authorization
  - RAS (Registration, Admission & Status) → Maintaining terminals (between GK and terminal)
  - Accounting and statistics (e.g. Call Detail Records)
  - Large scale deployments: GK hierarchy

H.323 Addressing (Gatekeeper)

- IP address/domain name (GK is not needed)
- H.323 ID (GK)
  - Max. 256 unicode string (e.g. NIIF1)
  - Inconvenient with remote control
- e.164 ID (GK)
  - Phone numbers (recommended)
  - Max. 128 digits (0-9) és #, * (pl. 00361001234)
- E-mail (GK)
  - user@domain format
  - vsfx1@vidkonf.niif.hu
- URL (GK)
  - Max. 512 chars (e.g. ras://vsfx1.vidkonf.niif.hu)
  - Very unusual, inconvenient

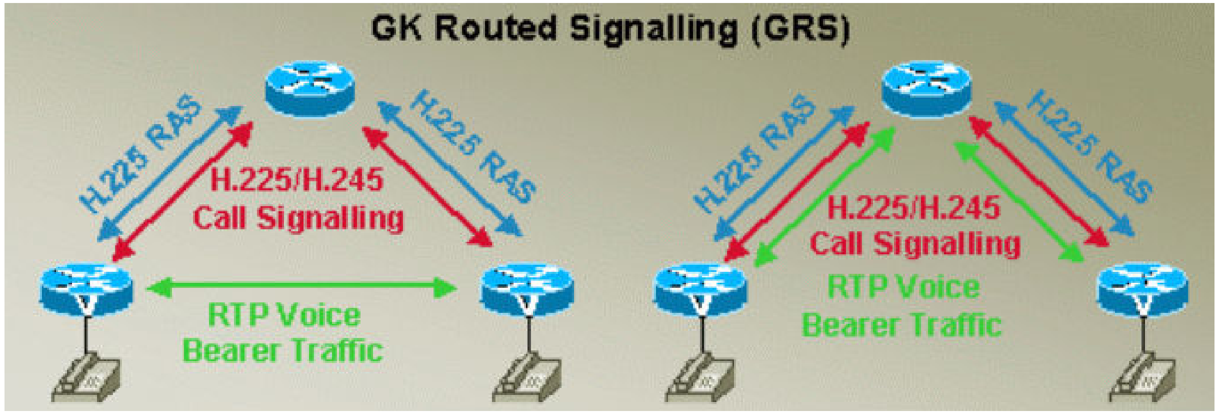
H.323 Signalling example

- Call setup
  - H.225.0 = 1-4, 6, 7 (RAS) ill. 5, 8 (Q.931)
  - H.245 = 9-12
  - Audio/video communication: RTP/RTCP



H.323 & Firewalls

- Ports used:
  - H.225.0 RAS: UDP 1719
  - H.225.0 Q.931: TCP 1720
  - H.245: TCP >1024
  - RTP/RTCP: UDP > 1024
- Type of call signaling:
  - Direct Signaling
  - Routed Signaling (all signals through GK)
  - Proxy (RTP/RTCP through GK) → Firewall



H.323 & Firewalls

- Difficulties:
  - Many TCP/UDP ports are opened
  - Random ports for H.245 & RTP/RTCP
  - ASN.1 encoded messages (demanding to decode/encode)
  - H.323 aware firewalls are rare
- Possible solutions:
  - Basic:
    - Restriction on used random ports (endpoints, gatekeepers)
    - Open holes in firewall
  - Proxy gatekeeper (see previous slide):
    - Allow proxy IP to send and receive
    - Proxy load?
  - H.460.17: TCP 1720 tunneled signaling (H.225, H.245)
  - H.460.18 & 19: Traversal server based solutions

Session Initiation Protocol

Session Initiation Protocol

- SIP = Session Initiation Protocol
  - IETF defined RFC 3261
  - Goal: establish sessions over the Internet
  - Simple textual (HTTP-like) protocol (no ASN.1)
    - Allows simple devices (from HW to browsers, softphones)
  - Internet approach instead of telco perpective
  - Why SIP?
    - All-IP
    - Audio, video & data integration
    - Open to other applications (messaging, presence, etc.)
    - Opens competition
    - Textual protocol allows any extension
  - SIP vs. H.323 battle:
    - Over (years ago), SIP has won
    - VC vendors still H.323 biased…

SIP network elements

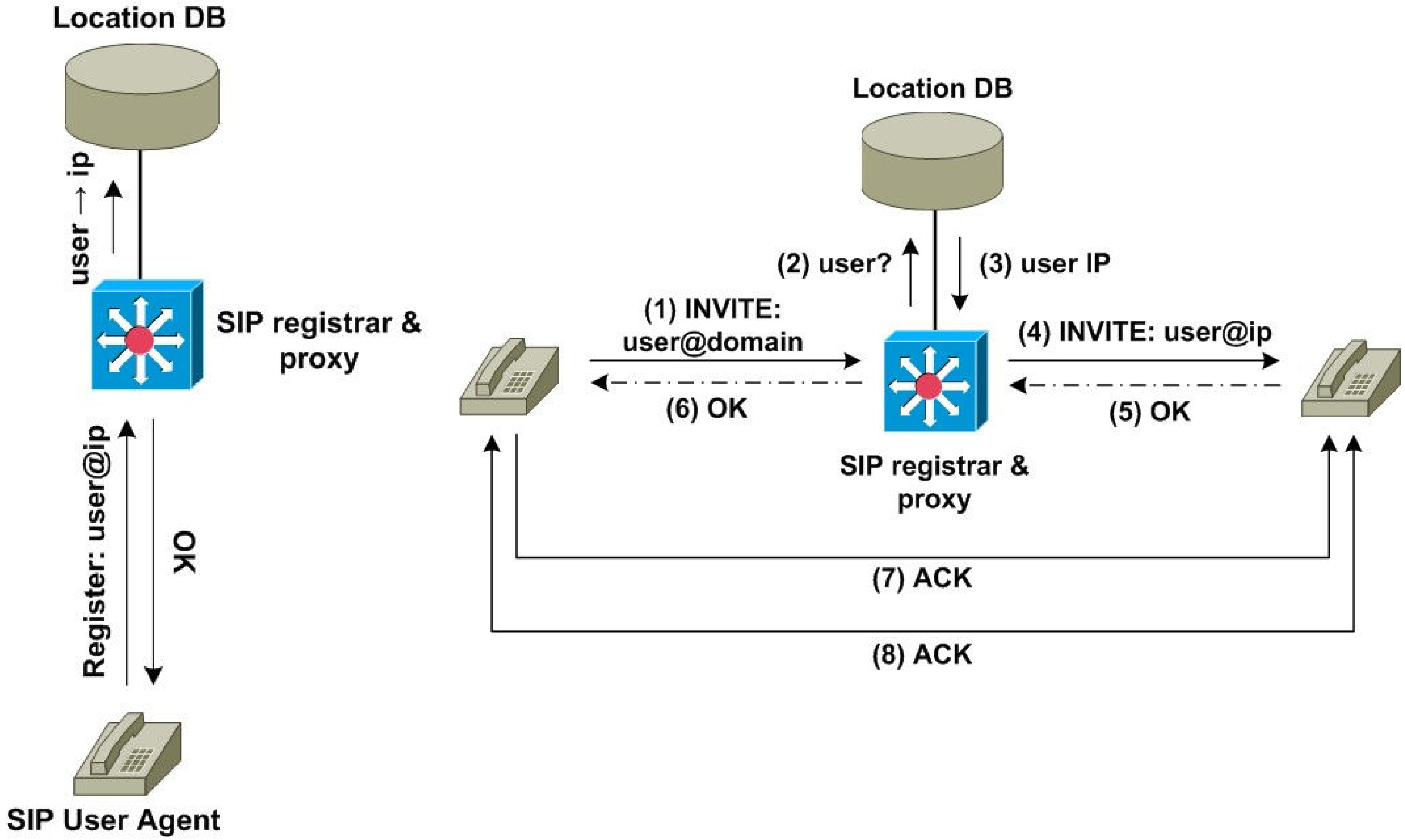
- Terminal → User Agent (UA)
  - Client (phone, VC device)
  - Server (IVR service, voice mail, etc.)
- SIP registrar:
  - UA logs in when started
  - Keeps track of UA location (IP)
- SIP proxy:
  - Relays call signaling
  - Forwards UA request to next-hop
  - Based on location database and routing rules
  - Acts as both client and server
- SIP redirect server:
  - Redirection to other servers/services
- Typically: all three are part of a server
- Other components:
  - Gateways (e.g. SIP-PSTN, SIP-H.323, etc.)
  - Application servers (e.g. MCU)

SIP – related protocols and mechanisms

- TCP, UDP
- RTP/RTCP for audio/video transport
- DNS for domain name record lookup
- ENUM
- STUN (NAT traversal)
- Codecs: H.323 details apply here
- Session Description Protocol (SDP) for media stream description

SIP communication example

- Registration & simple call



SIP call routing

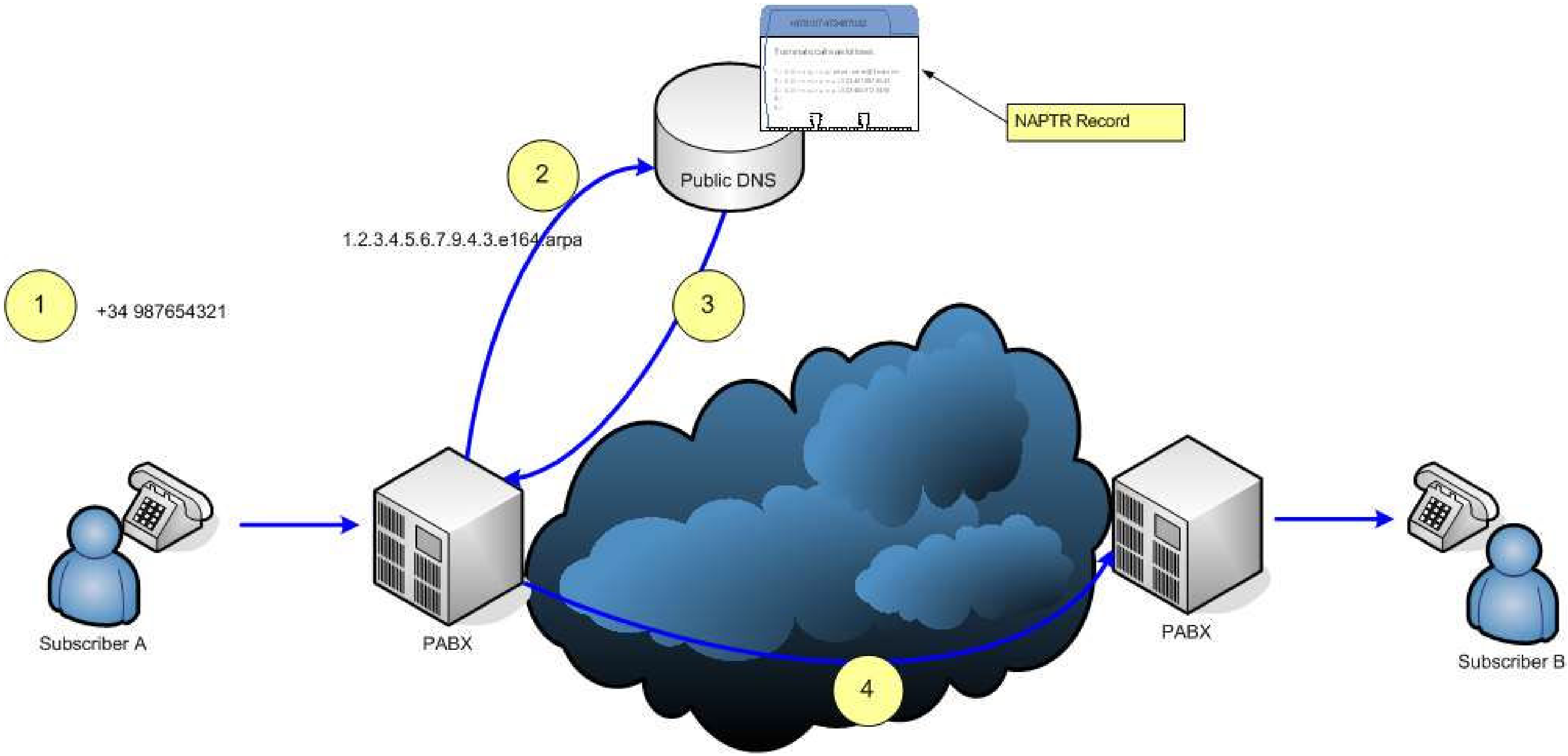
- Inside domain: location DB
- Between domains:
  - Based on DNS
    - SRV records
    - ENUM
  - Static routing information
  - Dynamic routing information

ENUM

- ENUM = E.164 number mapping
  - IETF developed (RFC 3761)
  - Linking a single ID to 1…N services. Usage:
    - Personal IDs
    - Linking telephony and IP worlds
  - DNS based mapping of E.164 to URI/IP
- 1 x E.164 to many services:
  - e-mail addresses
  - fax number
  - personal website
  - VoIP number
  - mobile telephone numbers
  - voice mail systems
  - IP-telephony addresses
  - GPS coordinates
  - etc.

ENUM example

- ENUM example



(Source: Wikipedia)

ENUM NAPTR example

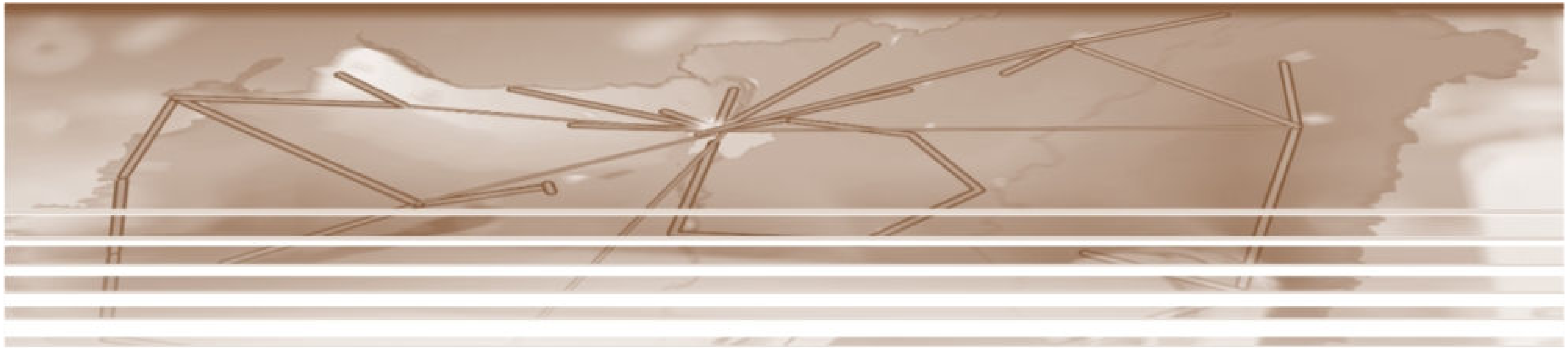
- NAPTR:
  - Naming Authority Pointer Resource
  - RFC 3403 (previously RFC 2915)
- Example:
  - E.164 number: +36 1 666 1111
  - Remove everything: 3616661111
  - Reverse it and use dots: 1.1.1.1.6.6.6.1.6.3
  - Add e164.arpa suffix: 1.1.1.1.6.6.6.1.6.3.e164.arpa
  - Add NAPTR record to this zone:

```
IN NAPTR 100 1 "u" “E2U+sip" "!^.*$!sip:akov@niif.hu!”
```

SIP and firewalls

- SIP ports:
  - Clients use TCP or UDP 5060 and/or 5061
  - Media transport uses random ports…
- Firewall traversal required

Thank You!



Kovács András

<akov@niif.hu>