Communication Networks

VITMAB06

VoIP, SIP

Gusztáv Adamis

BME TMIT

2024

TMiT

©

Voice over IP

VoIP

Voice transmission networks

- Voice transmission in chronological order :
  - PSTN (Public Switched Telephony Network)
    - From 1876
  - ISDN (Integrated Services Digital Network - digital landline telephone network)
    - From the end of 1980s, only moderate success
  - Mobile systems
    - From the beginning of 1980s
  - VoIP (Voice over IP)
    - „Voice” is “speech”, not “sound” or “voice in general”
    - From appr. 2010s has really spread
    - Both on fixed and mobile networks

Why VoIP is good?

- Basic idea: it is unnecessary to maintain two (three) networks
  - fixed telephone network and fixed IP network
- Voice traffic requires a very small bandwidth from an IP point of view
  - 6...64 kb/s a voice channel
  - appr. 200 Mb/s was the national telecommunications backbone network
- There will also be fewer wires in the flat/office
- Costs can be reduced
- Not only voice transmission, but also integrated data and image transmission
  - E.g. sending a URL in a voice call,
  - Look at the referred web page
  - Web-based phone book

Challenges of VoIP

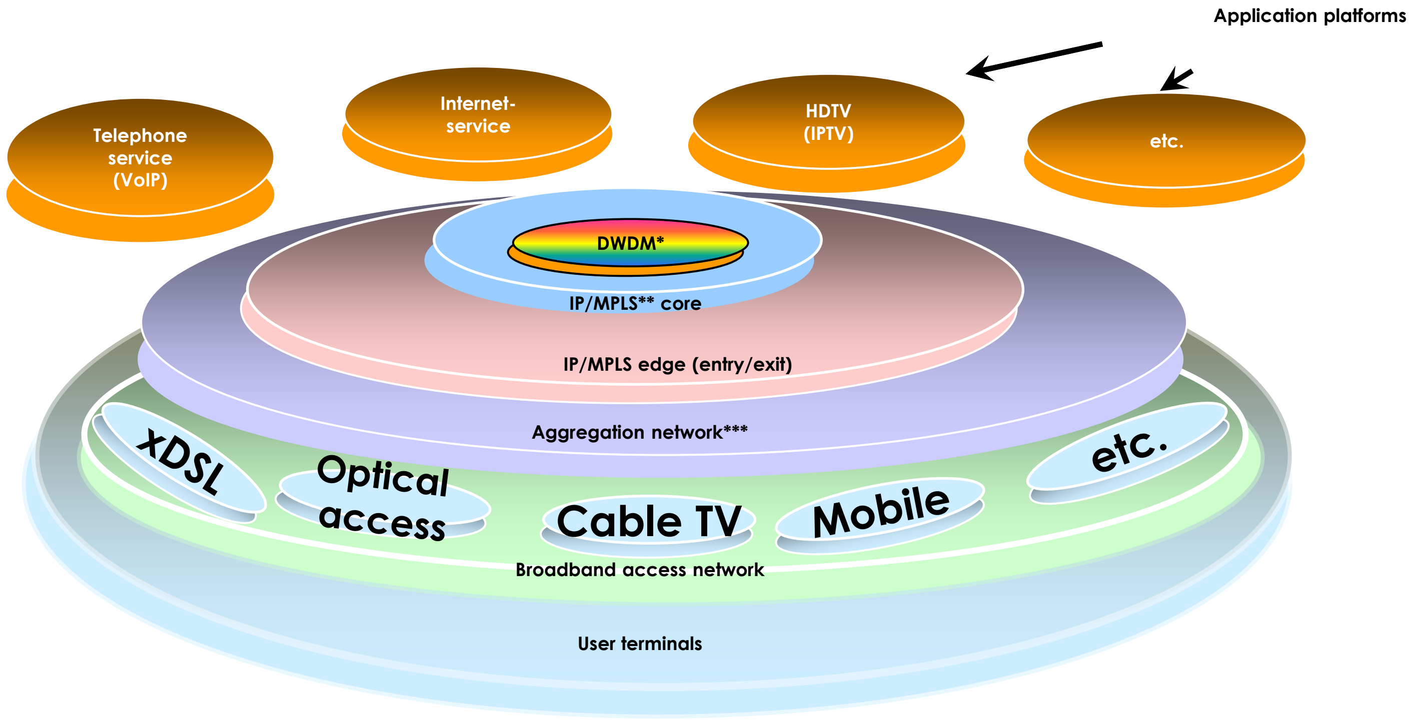
- PSTN/ISDN/mobile (e.g. GSM/UMTS) networks were designed to be "bombproof".
  - High availability
    - High reliability devices
    - Redundancy
  - Thoroughly tested protocols
  - Closed network (intrusion protection)
  - Many, many, many years of experience
- Guaranteed service quality
  - Thanks to circuit switching
- Additional services
  - E.g. the emergency call with a harmonized number
- All this had to be created for VoIP-based networks

VoIP architectures

- At first glance, it seems like an IP application layer problem
  - True on some level. However, there are:
  - dedicated protocols
    - for different tasks: data transmission, connection establishment
  - dedicated hardware
    - end devices, network nodes

VoIP architectures

- VoIP is a general term. Question: which part of the network is IP?
  - Initial solution: replacing trunks with IP, switches remain TDM
  - By now: switches are also IP-based networks.
  - Reminder: Next Generation Networks (NGN)

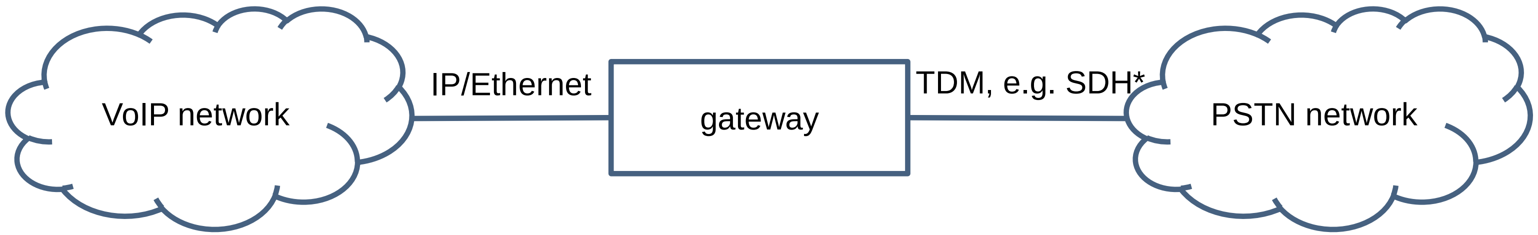


VoIP architectures

- … and the end devices? (=telephone devices)
  - At home subscribers, this often remains wired and analog
    - They are connected to IP devices, but the user does not need to know this
      - E.g. IP telephone exchange, cable modem, ADSL Home Gateway
    - the phone number remains unchanged
    - analog telephones can also have push-button, display, etc.
      - but the voice on wire pair coming out of it is analog
  - In a corporate environment, this is also often IP-based: VoIP terminal equipment
    - looks similar to a "traditional" phone (see e.g. VoIP measurement)
    - With IP address, Ethernet connector
    - Even with extra services (e.g. web-based phone book)
  - Softphone = VoIP software
    - e.g. Skype, Viber, Facebook Messenger, WhatsApp, Google Hangouts, Discord, TeamSpeak, Ekiga, Linphone, etc.
    - Can run on a PC, laptop, tablet, mobile phone, …

VoIP architectures

- Where there is also a "traditional" telephone network...
  - (= PSTN, Public Switched Telephone Network)
  - ... there needs to be a gateway at the border of the VoIP/PSTN networks
  - even if only the end device is analog



*TDM = Time Division Multiplexing
SDH = Synchronous Digital Hierarchy

VoIP functions

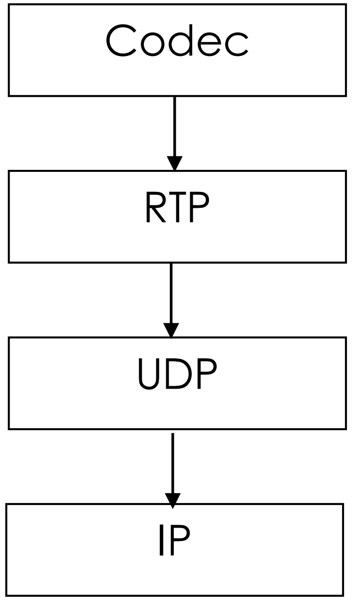
- Four function groups
  1. Voice encoding and decoding
  2. Transmission of voice packets
  3. Signaling (control)
  4. Interworking with other VoIP/PSTN networks (gateway functions)

1. Voice encoding and decoding
   - Codecs
   - Discussed previously
   - The main point now: their output is a bit stream with a speed of about 5-64 kb/s

VoIP functions

2. Transmission of voice packets

– Typically, in an RTP (Real-time Transport Protocol) packet embedded in a UDP packet (see IPTV presentation later)



IP header (20 byte)	UDP header (8 byte)	RTP header (12 byte)	Voice information (4-100 byte)

- RTP: timestamp
  - (Re-)ordering of packets
  - Can be played periodically, though they arrived with different delay
- Why NOT TCP used? **Too large delay of ACK/retransmission**
- Why UDP used? **ports**
- How large voice info shall be in an IP frame?
  - Larger IP packet:
    - Smaller overhead
    - Larger delay
  - it is recommended to keep the total one-way delay ("mouth to ear") below 150ms, but never above 400ms

Delay and Quality

- One-way, mouth-to-ear delay
- From: ITU G.114

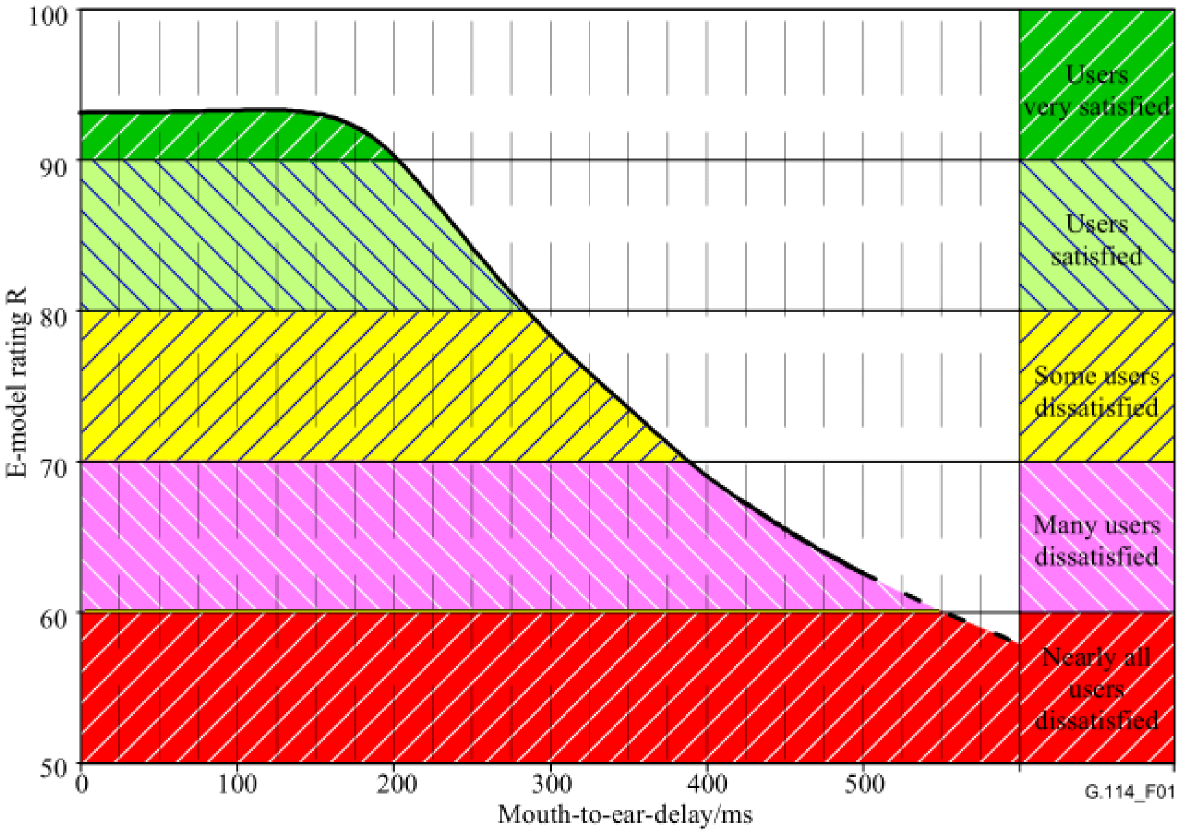


Figure 1/G.114 – Determination of the effects of absolute delay by the E-model

VoIP functions

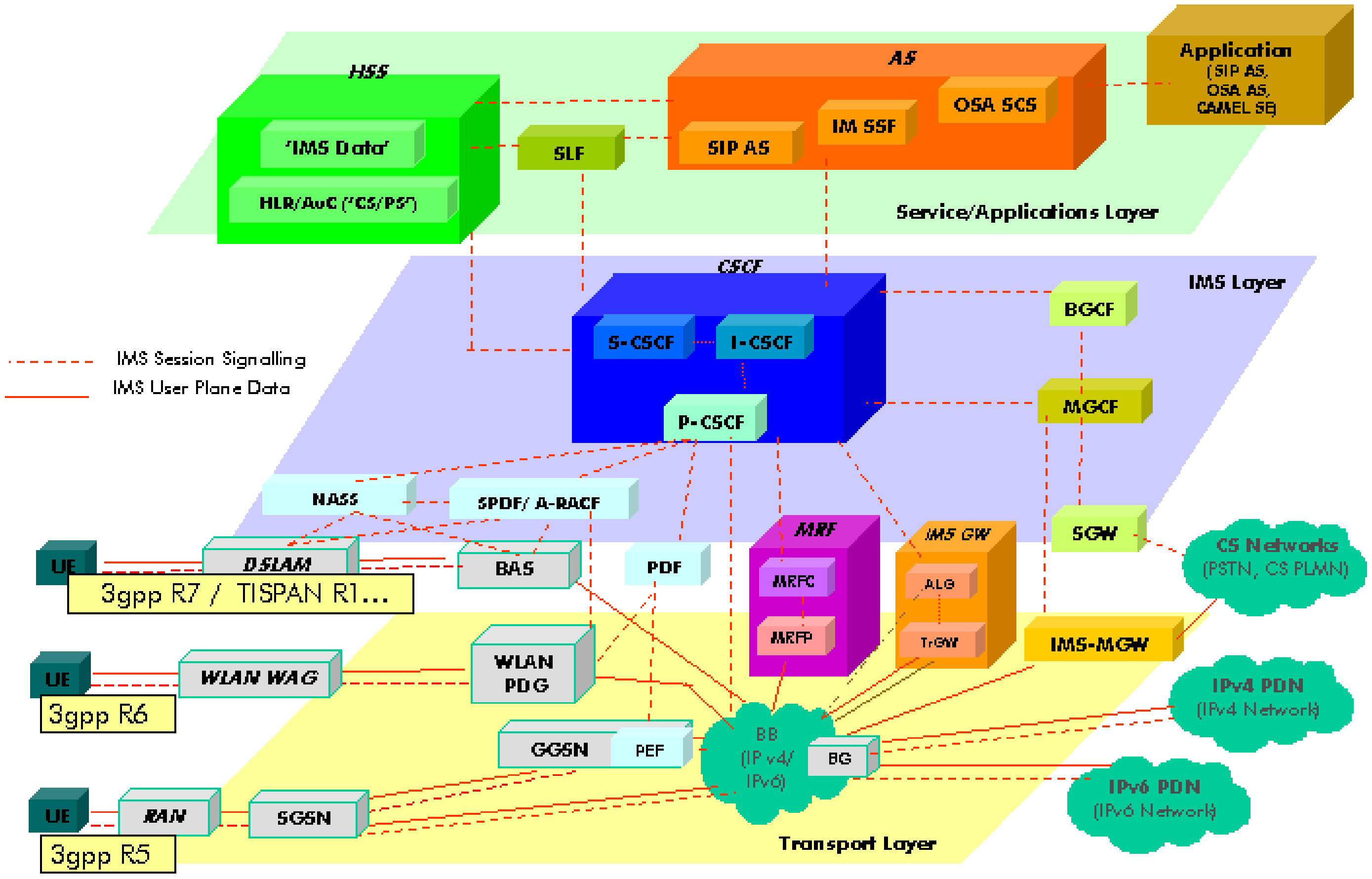
3. Signaling
   - Most important: establishment and release of a connection
   - Lots of signal system recommendations. The two most common:
     - H.323 (MegaCo) (ITU -- International Telecommunication Union) (older)
     - SIP (IETF -- Internet Engineering Task Force)
4. Interworking with other VoIP/PSTN networks (gateway functions)
   - A Gateway is needed that speaks the language of both PSTN and VoIP networks
     - according to all three aspects above, *for example*:
       - Voice encoding: PCM ↔ G.729 (it is a VoIP codec)
       - Transmission: SDH (which is a TDM transmission system) ↔ IP/UDP/RTP
       - PSTN signaling system (e.g. SS7) ↔ H.323 (VoIP signaling system)

IMS

- IMS = IP Multimedia Subsystem
  - It will also be mentioned in connection with VoLTE
  - Architecture for the implementation of the IP-based backbone of wired and mobile multimedia networks
  - The data is transmitted in IP packets, via routers
  - The data format and signals must also be converted to other networks
    - there are dedicated servers for these
  - Dedicated servers for managing signals
  - Application servers that implement certain functions
    - e.g. establishment of a conference call

IMS

Details won't be asked on exam, only the basics



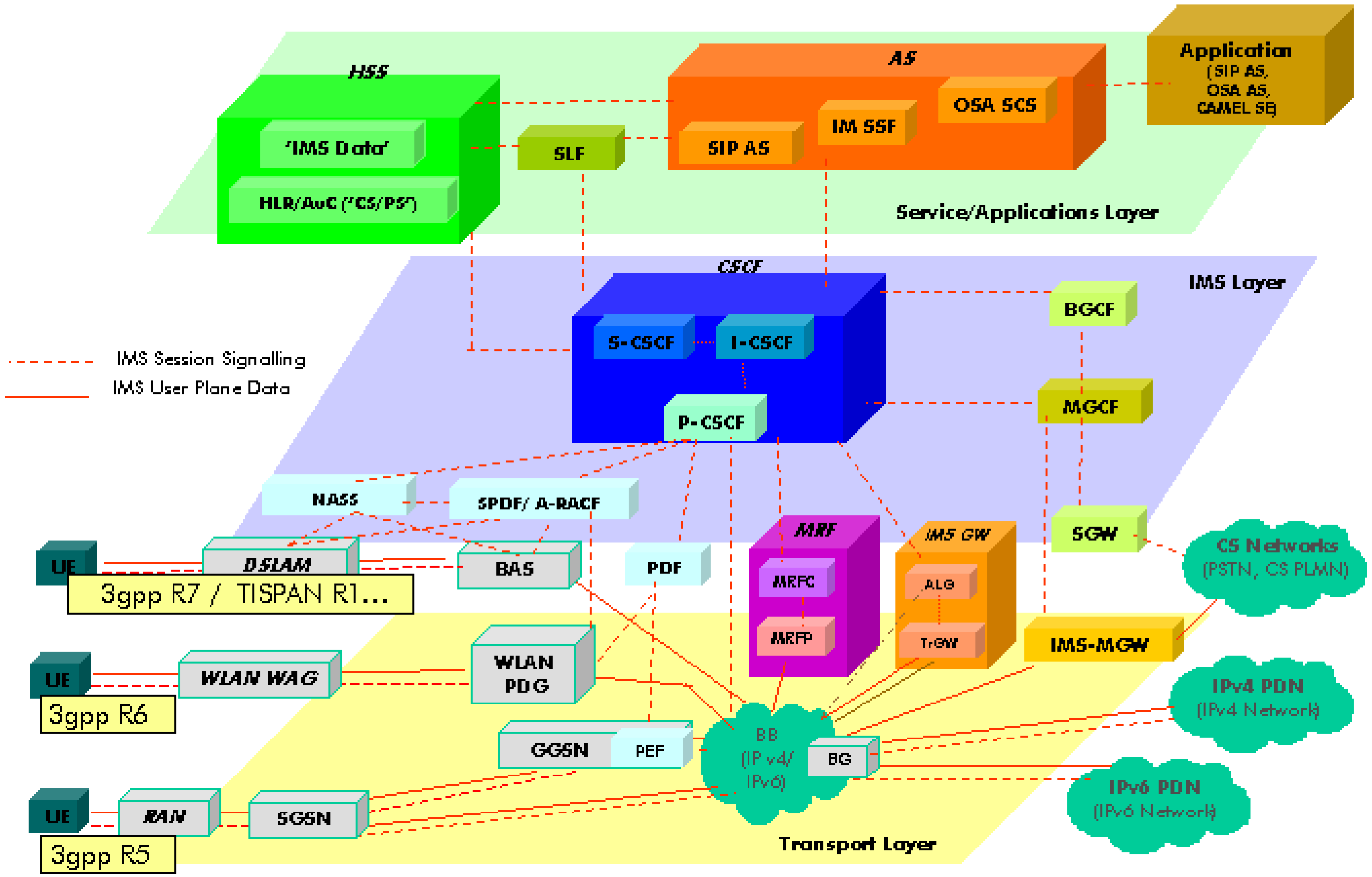
CSCF – Call Session Control Function

S-CSCF (Serving) – serves the user, downloads user profile from HSS to know which Application Server to be used

I-CSCF (Interrogating) – asks HSS to decide which S-CSCF to be used

P-CSCF (Proxy) SIP proxy (routing of registration, session requests; informs S-CSCF on access ntw; info for charging; maintains secure connection to user

IMS



- IMS: IP Multimedia Subsystem
- IP core-network architectural framework for media (e.g. VoIP) in mobile and fix (wired) networks
- Defined by 3GPP (3rd Generation Partnership Project) as industrial forum

- Some main functions:
  - IP packet transport
  - servers for signaling (SIP) and charging
  - servers for media-flow control
  - servers for gateway functions
  - application servers
    - e.g. for setting-up conference-calls

Session Initiation Protocol

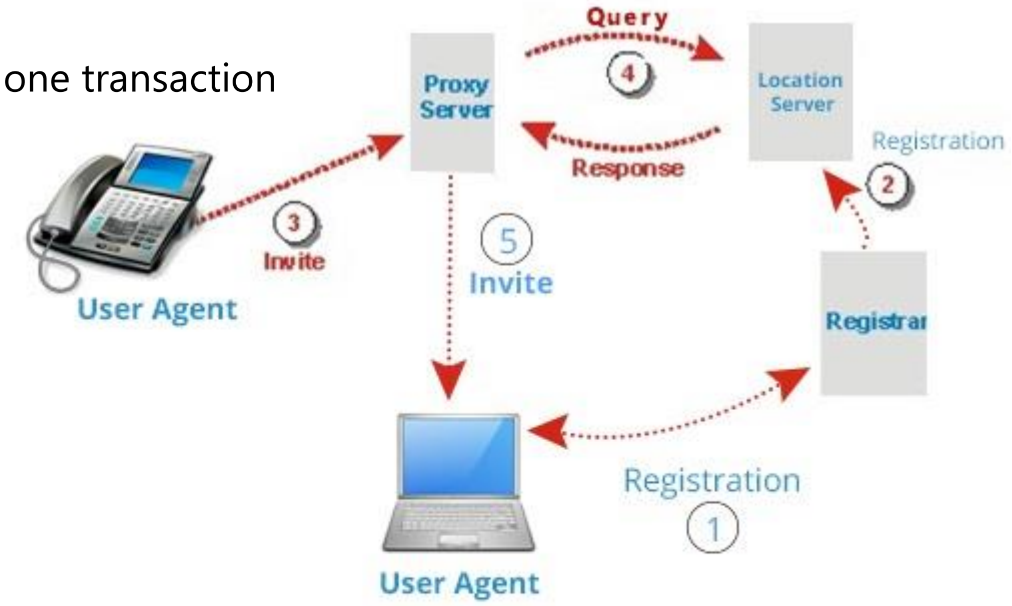
SIP

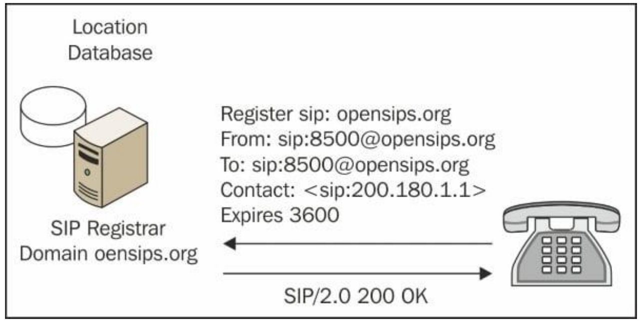
Session Initiation Protocol (SIP)

- IETF RFC 3261 and connecting RFCs (Internet Engineering Task Force, Request for Comments) https://tools.ietf.org/html/rfc3261
- It was originally published in 1999 and since it has been updated
- *Signaling* protocol for creating, managing, and terminating real-time connections (sessions).
- E.g. speech, video call, instant messaging
  - Between two or more parties
- Text-based, HTTP-like
- The *media stream* goes *separately*, not via SIP (mainly RTP - Real-time Transport Protocol)
- Over UDP/TCP/SCTP (Stream Control Transmission Protocol)

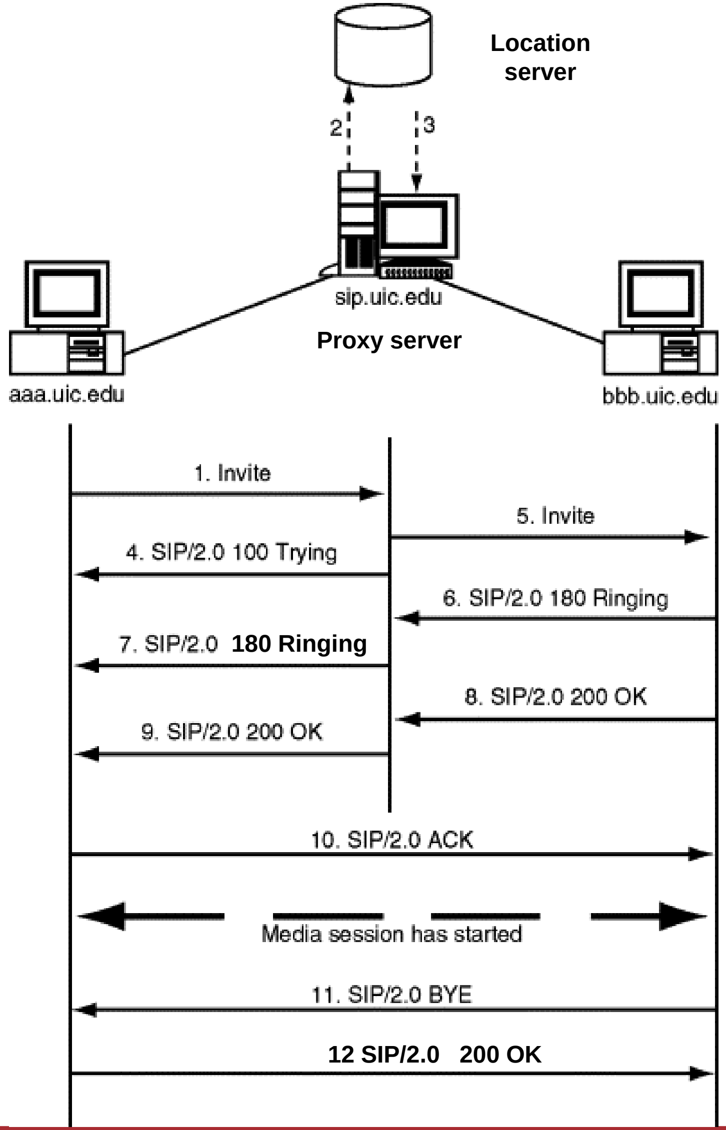
SIP network components

- User agent, UA:
  - UAS: UA Server, sends a request
  - UAC: UA Client, receives a request
  - The roles may change and are only permanent for the duration of one transaction
- Proxy server
  - Its main task is to route the call
  - Basically, the role of a proxy server is much like a router
  - It has some intelligence to understand a SIP request and send it ahead with the help of URI
- Registrar (Registration server)
  - A UA registers to it
  - It stores the URI and the location of users
- Redirect server
  - Redirects the requester to another endpoint (as in HTTP)
- Location server
  - The location server provides information about a caller's possible locations to the redirect and proxy servers.
  - Only a proxy server or a redirect server can contact a location server.
- Gateway
  - Connection to other networks, e.g. PSTN
    - PSTN = Public Switched Telephony Network, there aren't many like that anymore





SIP Call Flow



- SIP Requests
  - Each request has a method and a URI
- Identifier: URI (Uniform Resource Identifier)
  - E.g. sip:adamisgusztav@tmit.bme.hu
- The possible methods:
  - REGISTER Registration with a given IP address
  - INVITE Request to establish a connection
  - ACK Used to acknowledge the final response to an INVITE method
  - BYE Releasing a connection
  - CANCEL Cancellation of a pending request
  - MESSAGE Sending a text message

  Won't be asked on exam

  - Other methods: UPDATE, REFER, PRACK, SUBSCRIBE, NOTIFY, PUBLISH, INFO, OPTIONS

Session Description Protocol (SDP)

- RFC 8866 (2021, originally RFC2327, 1998)
- Description of the multimedia connection before the connection is established (in INVITE)
  - May be modified in ACK
- Important content :
  - Type of media (voice, video, etc.)
  - Transmission protocol used (e.g. RTP)
  - Format of media (e.g. codec type)
  - Accessibility of media (e.g. IP address, port)

Like:

```
v=0
o=- 3823730069 3823730069 IN IP4
192.168.0.33
s=pjmedia
b=AS:84
t=0 0
a=X-nat:0
m=audio 4062 RTP/AVP 8 101
c=IN IP4 192.168.0.33
b=TIAS:64000
a=rtcp:4063 IN IP4 192.168.0.33
a=sendrecv
a=rtpmap:8 PCMA/8000
a=rtpmap:101 telephone-event/8000
a=fmtp:101 0-16
a=ssrc:2070035071
cname:26517c500b201b33
```

Won't be asked on exam

- v = (protocol version)
- o = (owner/creator and session identifier)
- s = (session name)
- u =* (URI of description)
- e =* (email address)
- p =* (phone number)
- c =* (connection information - not required if included in all media)
- t= field contains the start time and stop time of the session. t=start-time stop-time (if both 0: permanent connection)
- z =* (time zone adjustments)
- a =* (zero or more session attribute lines)

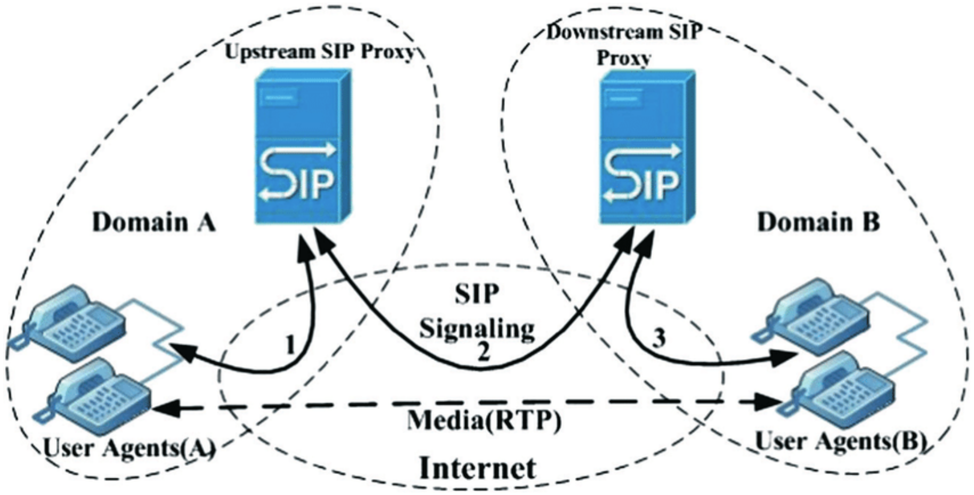
- m= (media name and transport address)
  - i=* (media title or information field)
  - c=* (connection information — optional if included at session level)
  - b=* (zero or more bandwidth information lines)
  - k=* (encryption key)
  - a=* (zero or more media attribute lines — overriding the Session attribute lines)

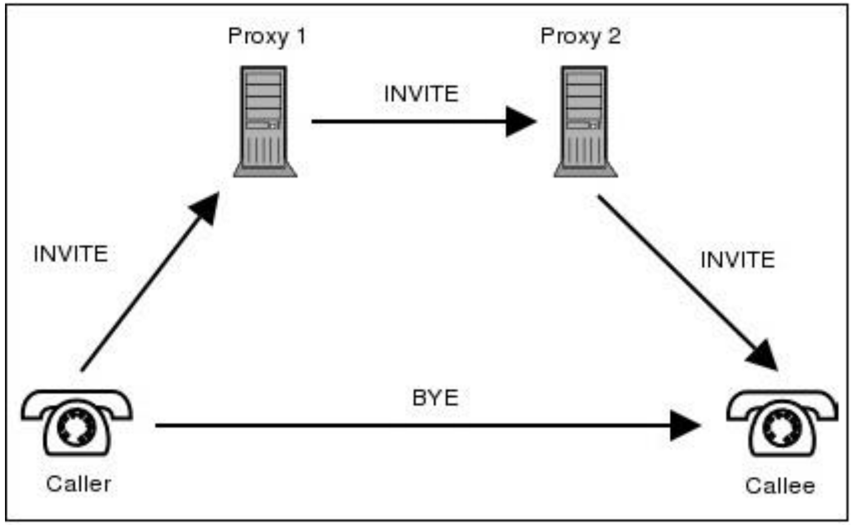
* denotes optional

SIP Responses

- 1xx: Provisional Responses (to serve the request may take longer tine)
  - E.g. 180 Ringing: waiting for the called party to answer
  - 181 Call is being forwarded
- 2xx: Successful Responses
  - E.g. 200 OK
- 3xx: Redirection Responses
  - E.g. 301 Moved Permanently: The requested URI no longer exists, but there is a new one that is returned
  - 302 Moved Temporarily: A new address will be given in the contact header field, which can be retried by the requesting client. This address should not be saved for future invite requests.
- 4xx: Client Failure Responses
  - E.g. 404 Not found: no such URI
- 5xx: Server Failure Responses
  - E.g. 500 Internal Server Error: The server could not fulfill the request due to some unexpected condition
- 6xx: Global Failure Responses: none of the servers can fulfill the request
  - E.g. 600 Busy Everywhere: A called party was contacted successfully but the called is busy
  - 603 Decline: the called party rejected the call

SIP Call Flow





- A caller initiates a call,
  - INVITE sent to the proxy server
  - The proxy server attempts to resolve the address of the callee with the help of the DNS server
- After getting the next route,
  - caller's proxy server (Proxy 1, outbound proxy server) forwards the INVITE to the callee's proxy server which acts as an inbound proxy server (Proxy 2) for the callee
- The inbound proxy server contacts the location server
  - to get information about the callee's address where the user registered
- After getting information from the location server,
  - Proxy 2 forwards the call to its destination
- Once the user agents get to know their address,
  - they can start the call (conversation) directly
  - media not SIP, but (typically) RTP

SIP message exchanges

- In the VoIP measurement, you can see some up close
- Much more complicated situations are also possible
- But all this is just a small introduction to SIP and VoIP