Voice over IP: Security and Mobility

Prof. Dr. Gerald Q. Maguire Jr.

<maguire@it.kth.se>

http://www.it.kth.se/~maguire

Wireless@KTH
&
Inst. for Micorelectronics and Information Technology (IMIT)
Royal Institute of Technology (KTH), Stockholm, Sweden

KTH Microelectronics and Information Technology

IVA Syd
Lund, Sweden

3 March 2004

Last modified: 04.03.03:10:42

Gerald Q. Maguire Jr.

July 1994 to present : Professor, Computer Communication (Datorkommunikation), formally the chair is described in the following Swedish:

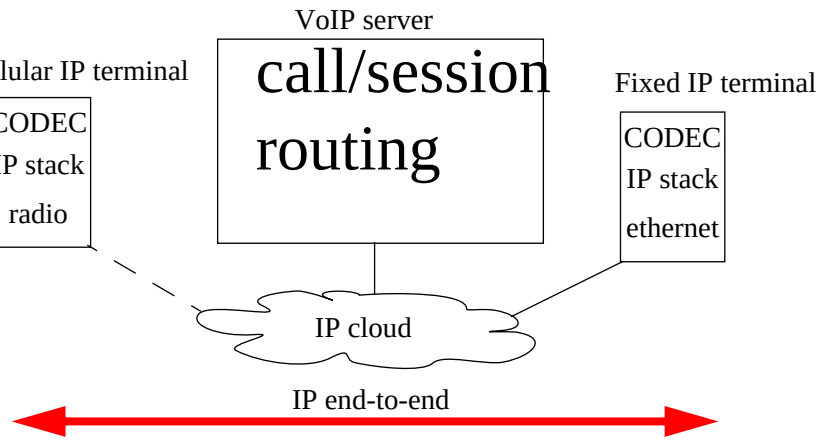
> Ämnet omfattar integrerade dator- och kommunikationssystem för personlig databehandling och personalkomunikation, speciellt trådlös kommunikation och effektiva implementerings strategier. (Ref. nr 473/93).

Jan. 1983 .. Dec. 1993 - on the faculty of Columbia University in CS

- Doctoral students (completed): 4 (KTH) and 4 (Columbia University); 8 licentiate students (KTH), Member of 34 doctoral dissertation committees
- MS & Civ. Ing. main advisor: 6 (KTH), 1 (Columbia); examiner: >220 (KTH)
- Teaching:
  - Voice Over IP, Sensor Networks, Wireless and Mobile Network Architectures, Internet-working, Mobile Personal Communication, Access Points, Location/Context Aware Personal Communication, Telecommunication and Data Communication (basic courses); total > 3400 students at KTH
  - VLSI design, Operating Systems, Artificial Intellegence, Programming Languages and Translators, Computer Organization, Symbolic Computation, Performance Analysis, Computer Networks, … ; total > 700 (Columbia)
- Education: Ph. D. 1983 and M.S. 1981 - both in Computer Science, University of Utah; B.A., magna cum laude, Physics, Indiana University of Pennsylvania, 1975

Voice over IP (VoIP)

VoIP is an End-to-End Architecture which exploits *processing in the end points*.



Unlike the traditional Public Switch Telephony Network - where processing is done **inside the network**.

Network Convergence:

In the past, many different networks - *each optimized for a specific use*: POTS, data networks (such as X.25), broadcast radio and television, … and each of these in turn often had specific national, regional, or proprietary implementations)

$\Rightarrow$

(Now) we think about a *converged* network which is a *global* network

VoIP, largely *independent* of the underlying network, i.e., LAN, Cellular, WLAN, PAN, Ad hoc, … .

Three major alternatives for VoIP

Concept	Implementation
Use *signalling* concepts from the traditional telephony industry	H.323
Use *control* concepts from the traditional telephony industry	Softswitches
Use an internet-centric *protocol*	Session Initiation Protocol (SIP)

SIP ⇒ a change from telephony's "calls" between handsets controlled by the network to "sessions" which can be between **processes** on **any** platform **anywhere** in the Internet and with both **control** and **media content** in *digital* form and hence can be easily manipulated.

- thus a separate voice network is **not** necessary
- open and distributed nature enables lots of innovation
  - since **both** *control* and *media* can be manipulated and
  - "events" are no longer restricted to start and end of calls

Traditional Telecom vs. Datacom

Circuit-switched	Packet-switched
standardized **interfaces**	standardized **protocols** and **packet formats**
lots of internal state (i.e., each switch & other network nodes)	very limited internal state • caches and other state are soft-state and dynamically built based on traffic • no session state in the network
long setup times - since the route (with QoS) has to be set up from end-to-end before there is any further traffic	End-to-End Argument ⇒ integrity of communications is the reponsibility of the end node, **not** the network
services: built **into** the network ⇒ hard to add new services • **operators** decide what services users can have • all elements of the net have to support the service before it can be introduced • Application programming interfaces (APIs) are often vendor specific or even proprietary	Services can be added by anyone • since they can be provided by **any** node *attached* to the network • users control their choice of services
centralized control	no central control ⇒ no one can easily turn it off
“carrier class” equipment and specifications • target: very high availability 99.999% (5 min./year of unavailability) • all equipment, links, etc. must operate with very high availability	a mix of “carrier class”, business, & consumer equip. • backbone target: high availability >99.99% (50 min./year unavailability) • local networks: availability >99% (several days/year of unavailability) • In aggregate - there is extremely high availability because most of the network elements are **independent**
long tradition of slow changes • PBXs > ~10 years; public exchanges ~30yrs	short tradition of very fast change • Moore’s Law doublings at 18 or 9 months!
clear operator role (well enshrined in *public law*)	unclear what the role of operators is (or even **who** is an operator)

VoIP vs. traditional telephony

Henning Schulzrinne in a slide entitled “Why should carriers worry?”[1] nicely states the threats to traditional operators:

- Evolution from application-specific infrastructure ⇒ **Content-neutral** bandwidth delivery mechanism - takes away the large margins which the operators are used to (and **want**!):
  - “GPRS: $4-10/MB, SMS: >$62.50/MB, voice (mobile and landline): $1.70/MB”
- Only operators can offer services ⇒ Anybody can offer phone services
- SIP only needs to handle signaling, not media traffic
- High barriers to entry ⇒ No regulatory hurdles[2]

In addition to this we can add:

- Only vendors can create services ⇒ anybody can create a service

---

1. Henning Schulzrinne, “When will the telephone network disappear?”, as part of Intensive Graduate Course "Internet Multimedia", University of Oulu, 3-6 June 2002.

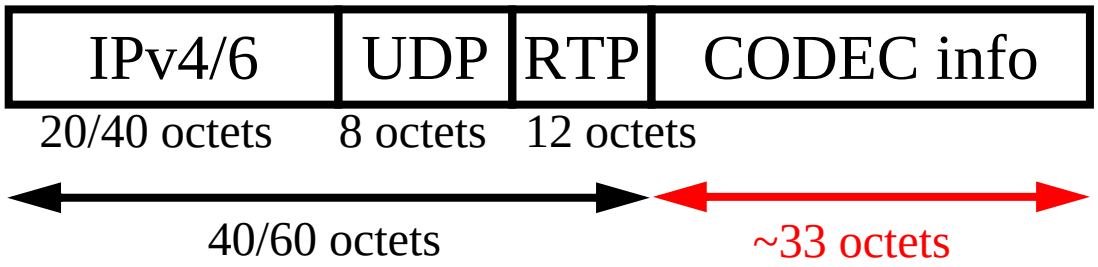
2. J. Sununu said “VoIP providers should be free from state regulation, free from the complexity of FCC regulations, free to develop new solutions to address social needs, and free to amaze consumers.” E-BUSINESS: New Hampshire Senator Readies, "Hands-Off VoIP" Bill, Internetweek.com, January 12, 2004 *http://www.internetweek.com/e-business/showArticle.jhtml?articleID=17300570*

VOIP Modes of Operation

- PC to PC
- PC-to-Telephone calls
- Telephone-to-PC calls
- Telephone-to-Telephone calls via the Internet
- Premises to Premises
  - use IP to tunnel from one PBX/Exchange to another
- Premises to Network
  - use IP to tunnel from one PBX/Exchange to a gateway of an operator
- Network to Network
  - from one operator to another or from one operator's regional/national network to the same operator in another region or nation

VoIP details: Protocols and Packets

Carry the speech frame inside an Real-Time Transport Protocol (RTP) packet



Typical packetization time of 10-20ms per audio frame.

A real-time application ⇒ data must be delivered with the same time relationship as it was created (all be it with a delay). Two aspects of real-time delivery:

Concept		**Implementation**
Order	data should be played in the same order as it was created	sequence number
Time	the receiver must know when to play the packets, in order to reproduce the same signal as was input	time stamp

The timestamp *granularity* (i.e., the units) are determined by the payload type {often based on the sampling rate}.

VoIP need not be “toll quality”

Public Switched Telephony System (PSTN) uses a **fixed** sampling rate, typically 8kHz and coding to 8 bits, this results in 64 kbps voice coding

However, VoIP is *not* limited to using this coding and could have **higher** or **lower** data rates depending on the CODEC(s) used, the available bandwidth between the end points, and the user’s preference(s).

One of the interesting possibilities which VoIP offers is quality which is:

- better that “toll grade” telephony or
- worse than “toll grade” telephony (but perhaps still acceptable)

This is unlike the *fixed* quality of traditional phone systems.

Audio Encodings: G.711 A-law, G.711 mu–law, 8 bit linear, 16 bit linear, Linear Predictive Coding (LPC), code-excited linear prediction (CELP) -- Federal Standard FED-STD 1016, GSM 06.10: RPE/LTP (residual pulse excitation/long term prediction) coding at a rate of 13 kb/s , Interactive Multimedia Association’s DVI ADPCM Wave Type, ITU’ G.721, ITU’s G.722: 7 kHz audio-coding within 64 kbit/s, ITU’s G.728: Coding of speech at 16 kbit/s using low-delay code excited linear prediction, MPEG-I or MPEG-II audio encapsulated as elementary streams, defined in ISO standards ISO/IEC 11172-3 and 13818-3, variable-rate version of DVI4, …

Delay and delay variance (jitter)

The end-to-end delay (from mouth to ear - for audio), includes the encoding, packetization, (transmission, propagation, switching/routing, receiving,)$^{+}$ dejittering, decoding, playing

To hide the jitter we generally use playout buffer **only** in the final receiver. Note: This playout buffer **adds additional delay** in order to *hide* the delay variations (this is called: **delayed playback**), playback delay > delay variance

There are very nice studies of the effects of delay on peceived voice quality, see R. G. Cole and J. H. Rosenbluth, "Voice over IP Performance Monitoring", Computer Communications Review, Vol. 21, Number 2, April, 2001, pp. 9-24.

- the delay impairment has roughly two *linear* behaviors, thus

$$I_d = 0.024d + 0.11(d - 177.3)H(d - 177.3)$$

$$d = \text{one-way delay in ms}$$

$$H(x) = 0 \quad \text{if}(x < 0) \quad \text{else} \quad H(x) = 1 \quad \text{when} \quad x \geq 0$$

- for delays less than 177ms, conversation is very natural, while above this it become more strained (eventually breaking down ⇒ simplex)

Playout delay

- Playout delay should track the network delay as it **varies** *during* a session
- This delay is computed for each talk spurt based on *observed* average delay and deviation from this average delay -- this computation is similar to estimates of RTT and deviation in TCP
- Beginning of a talk spurt is identified by examining the timestamps and/or sequence numbers (if silence detection is being done at the source)
- The intervals between talk spurts give you a chance to catch-up
  - without this, if the sender's clock were slightly faster than the receiver's clock the queue would build without limit! This is important as the 8kHz sampling in PC's codecs is rarely exactly 8kHz.

When to play

The actual playout time is **not** a function of the arrival time, only of the end-to-end delay which can be calculated as shown below:

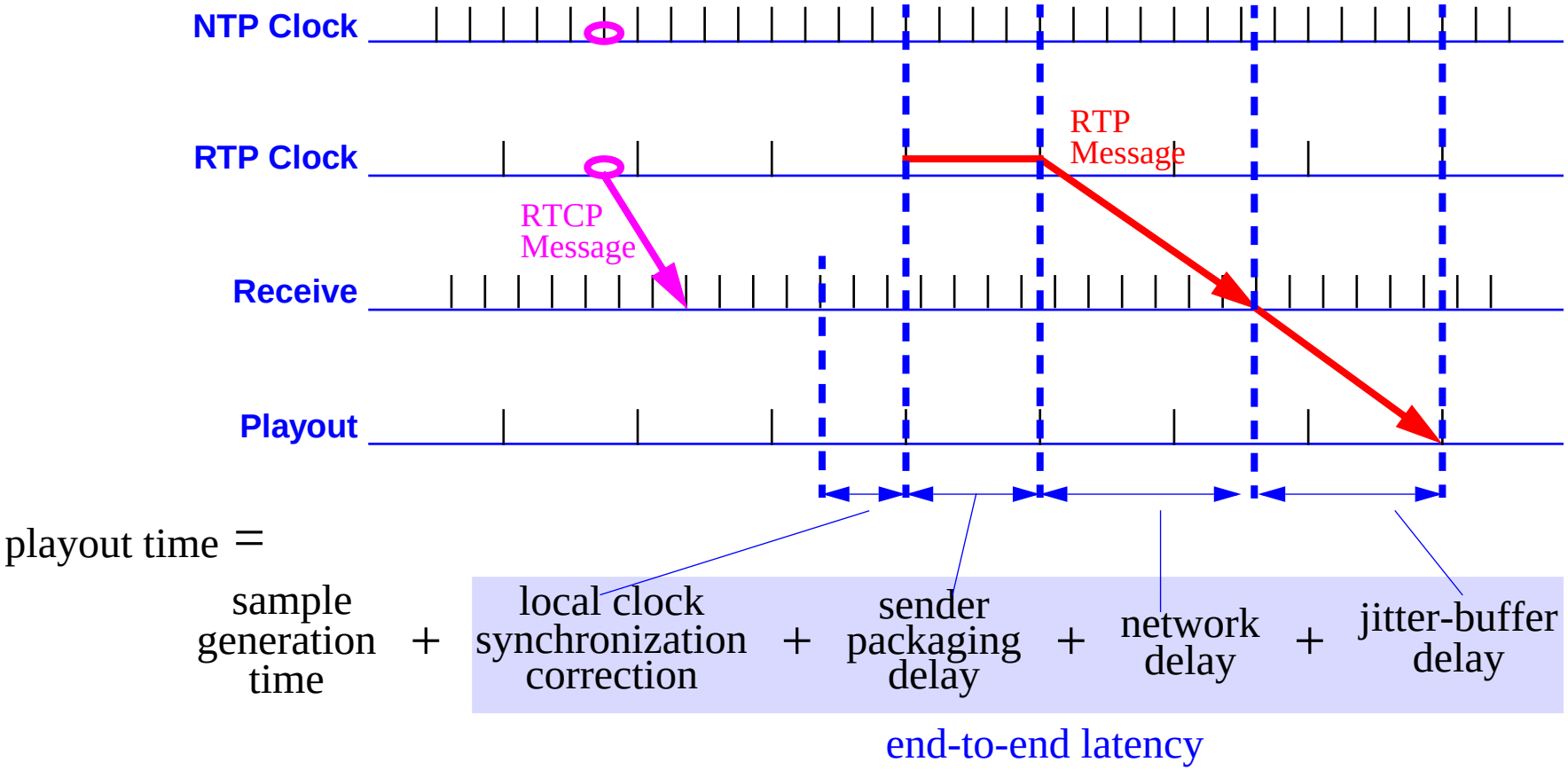


Figure adapted from slide 11 on page 6 of Kevin Jeffay, “Lecture 9: Networking Performance of Multimedia Delivery on the Internet Today”, Lecture notes for COMP 249: Advanced Distributed Systems Multimedia, Dept. of CS, Univ. of North Carolina at Chapel Hill, November 9, 1999. http://www.cs.odu.edu/~cs778/jeffay/Lecture9.pdf

Retransmission, Loss, and Recovery

For interactive real-time media we generally don't have time to request the source to retransmit a packet and to receive the new copy ⇒ **live without it** or *recover it using Forward Error Correction (FEC)*, i.e., send sufficient redundant data to enable recovery.

For **non**-interactive media we can use retransmission at the cost of a longer delay before starting playout

If you do have to generate output, but don't have any samples to play:

- audio
  - Comfort noise: play **white nosie** or play noise like in the last samples {as humans get uncomfortable with complete silence, they think the connection is broken!}
  - if you are using highly encoded audio even a BER of 10^{-5} will produce very noticable errors

Various techniques for loss concealment (i.e., hiding losses), such as those used in the UCL's Robust Audio Tool (RAT) http://www-mice.cs.ucl.ac.uk/multimedia/software/rat/

Session Initiation Protocol (SIP)

- Defined in RFC 3261 *http://www.ietf.org/rfc/rfc3261.txt*
- provides application layer signaling
  - Used to **establish**, **modify**, and **terminate** multimedia *sessions*
- can utilize UDP, TCP, TLS, SCTP, … for underlying transport
- HTTP-like
  - uses **textual** rather than binary (ala H.323) messages (⇒ humans can read them)
  - uses Uniform Resource Indicators (URIs) to designate calling and called parties
- target applications : voice, video, gaming, instant messaging, presence, call control[1], …

SIP **only** covers **signaling** parts of H.323. Does not use RTP itself, but sessions can use RTP.

- SIP provides ability to **discover** remote users and **establish** interactive **sessions**
- Does **not** ensure QoS or deliver large quantities of data

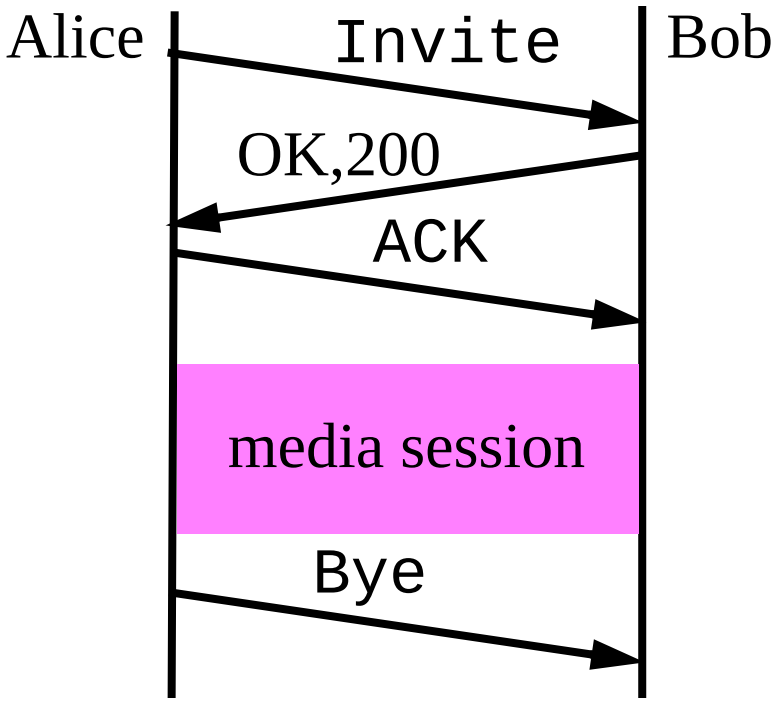
SIP uses SDP (Session Description Protocol) to provide information about a call, such as, the media encoding, protocol port number, multicast addresses, etc.

1. Largely taken from Advanced Intelligent Network (AIN).

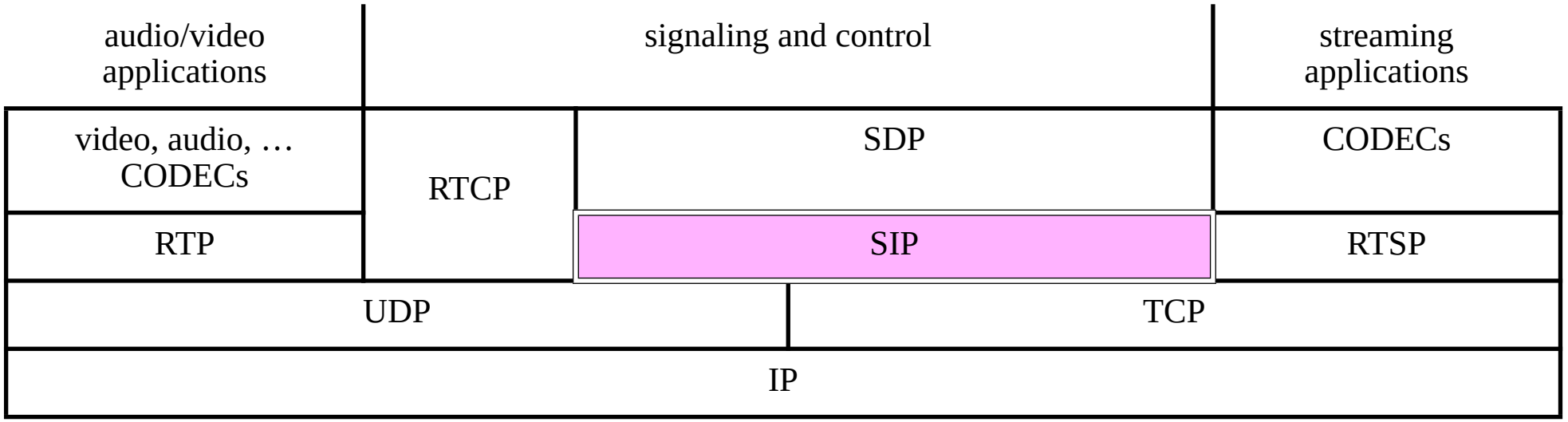
Is SIP simple?

- 25 RFCs (for SIP and SDP) - total of 823 pages
- RFC3261 is longest RFC ever (based on byte count)
- There are claims that one can still build a simple user agent in a (long) evening, but there is **substantial** work required with respect to security (due to TLS, S/MIME, AAA, Denial of Service issues, …)

SIP timeline - showing a simple version of Alice invites Bob to a SIP session:



SIP, RTP, and RTSP



RTP Control Protocol (RTCP)

[upward] enables endpoints to provide meta-information to the source - this enables the sources to be adaptive to the endpoints. For example, by using an adaptive coding algorithm the source can accommodate the actually data rate of packets arriving at the endpoint.

[downward] enables sources to send the endpoints information about a session

Real Time Streaming Protocol (RTSP), defined in RFC 2326 - remote media playback control (think in terms of controlling a remote VCR/DVD/CD player).

SIP actors

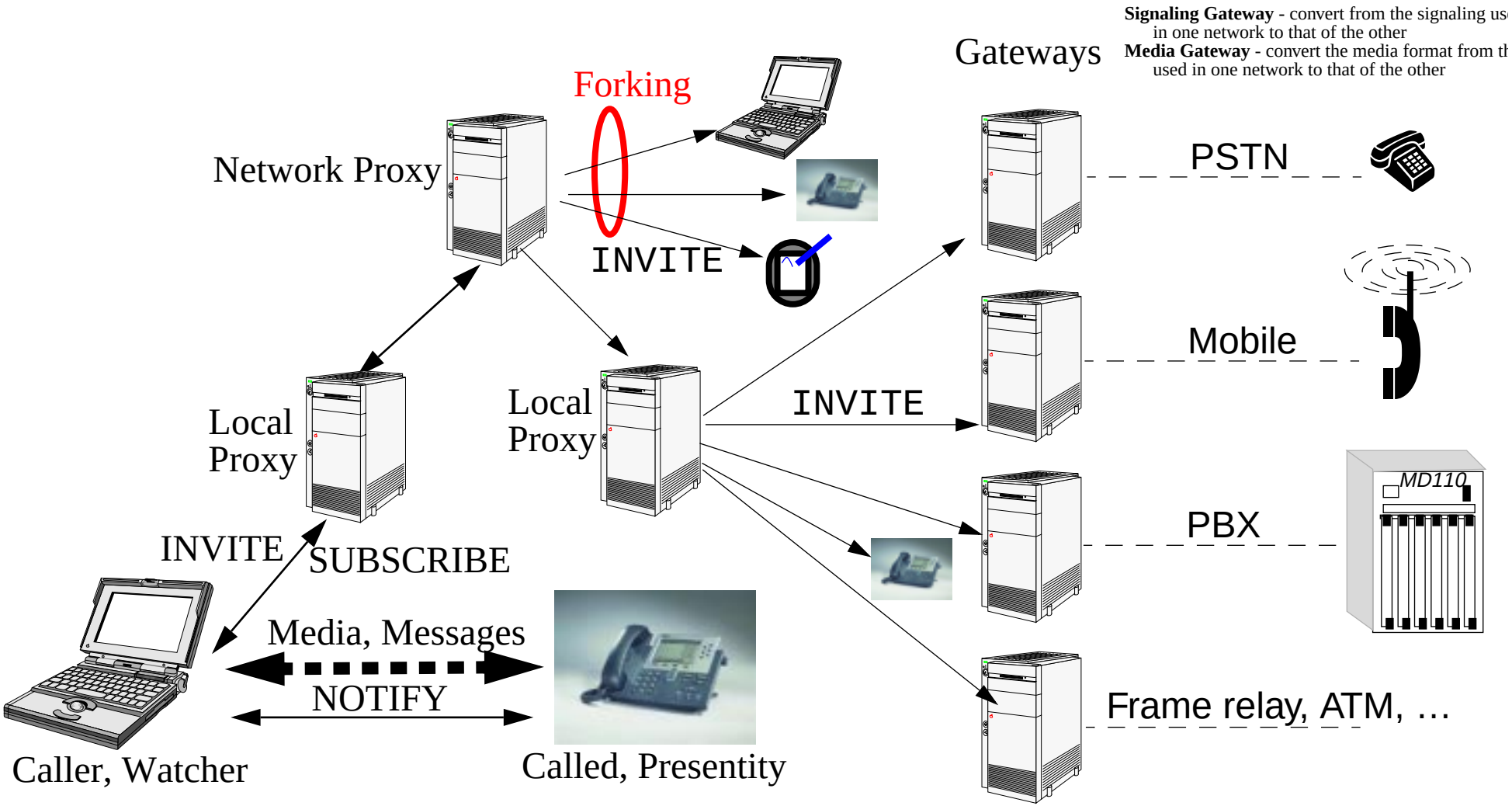


Figure 1: SIP Actors

SIP Methods and Status Codes

Method	Purpose
INVITE	Invites a user to join a call.
ACK	Confirms that a client has received a final response to an INVITE.
BYE	Terminates the call between two of the users on a call.
OPTIONS	Requests information on the capabilities of a server.
CANCEL	Ends a pending request, but does not end the call.
REGISTER	Provides the map for address resolution, this lets a server know the location of a user.

At least 8 additional methods have been defined.

SIP Status codes - patterned on and simular to HTTP's status codes:

Code	Meaning
1xx	Informational or Provisional - request received, continuing to process the request
2xx	Final - the action was successfully received, understood, and accepted
3xx	Redirection - further action needs to be taken in order to complete the request
4xx	Client Error - the request contains bad syntax or cannot be fulfilled at this server
5xx	Server Error - server failed to fulfill an apparently valid request (Try another server!)
6xx	Global Failure - the request cannot be fulfilled at any server (Give up!)

SIP Uniform Resource Indicators (URIs)

URI's have the same basic form as e-mail addresses: user@domain

Two URI schemes:

- SIP URI - introduced in RFC 2543
  - example: `sip:maguire@kth.se`
- Secure SIP URI - introduced in RFC 3261
  - example: `sips:maguire@kth.se`
  - Requires TLS over TCP as transport for security

Two types of SIP URIs:

- Address of Record (AOR) (identifies a user)
  - example: `sip:maguire@kth.se`
  - Need DNS SRV records to locate SIP Servers for kth.se domain
- Fully Qualified Domain Name (FQDN) (identifies a specific device)
  - examples: `sip:maguire@130.237.212.2` or `sip:maguire@chipsphone.it.kth.se`
  - `sip:+46-8-790-6000@kth.se; user=phone` the main KTH phone number in E.164 format via a gateway; note that the visual separators in a phone number (dashes, dots, etc.) are ignored by the protocol

SIP Invite (method/URI/version)

```
INVITE sip:bob@biloxi.com SIP/2.0
Via: SIP/2.0/UDP pc33.atlanta.com:5060;branch=z9hG4bK776asdhds
To: Bob <sip:bob@biloxi.com>
From: Alice <sip:alice@atlanta.com>;tag=1928301774
Call-ID: a84b4c76e66710
CSeq: 314159 INVITE
Contact: <sip:alice@pc33.atlanta.com>
Content-Type: application/sdp
Content-Length: 142

(Alice's SDP not shown)
```

Start Line is the first line of a SIP message which contains:

- method or Request type: `INVITE`
- Request-URI which indicates who the request is for: `sip:bob@biloxi.com`
- SIP version number: `SIP/2.0`

Issues to be considered

- Address Resolution - first step in routing the SIP request is to compute the **mapping** between the **URI** and *a specific user* at *a specific host/address*.
  - This is a very general process and the source of much of SIP's power.
  - Provides support for mobility and portability
  - Can utilize: DNS SRV lookup, ENUM, Location Server lookup
- Session Setup
- Media Negotiation
- Session Modification
- Session Termination
- Session Cancellation
- Mid-call Signaling
- Call Control
- QoS Call setup

Several types of SIP Servers

- User agent server runs on a SIP terminal (could be a SIP phone, a PDA, laptop, …) - it consists of two parts:
  - User Agent Client (UAC): initiates requests
  - User Agent Server (UAS): responds to requests
- SIP proxy - interprets (if necessary, rewrites specific parts of a SIP request message) before forwarding it to a server closer to the destination:
  - SIP **stateful** proxy server - remembers its queries and answer; can also forward several queries in parallel (can be **Transaction Stateful** or **Call Stateful**).
  - SIP **stateless** proxy server
  - They ignore SDP and don't handle any media (content)
  - Outgoing proxy: used by a user agent to route an outgoing request
  - Incoming proxy: proxy server which supports a domain (receives incoming requests)
- SIP redirect server - directes the client to contact an alternate URI
- Registrar server - receives SIP REGISTER requests updates LS
- Location server (LS) - knows the current binding and queried by proxies to do their routing
  - SIP can also use DNS SRV (Service) Records used to locate (inbound) proxy.
  - note in RFC 2543: a location server is a generic term for a **database**

SIP Trapezoid[1]

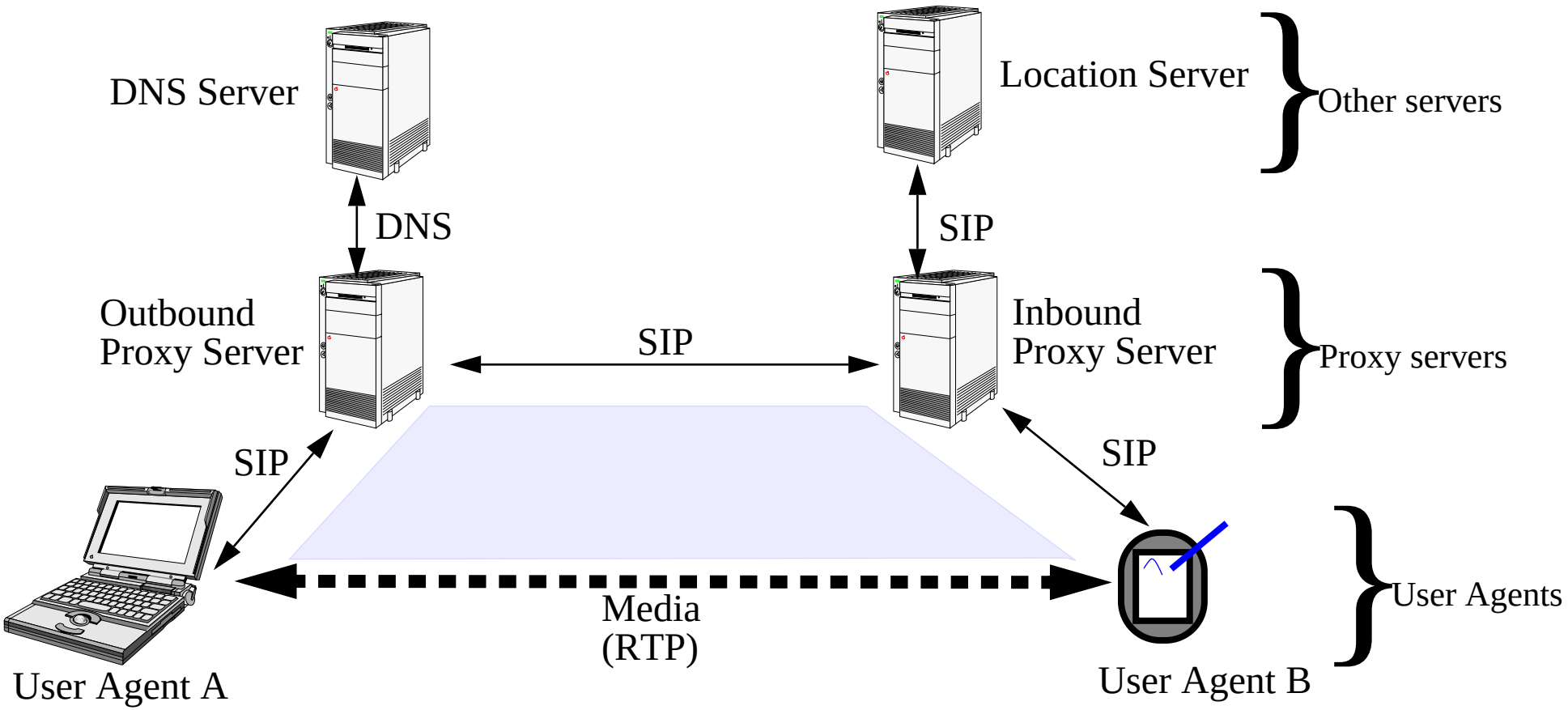


Figure 2: SIP Trapezoid

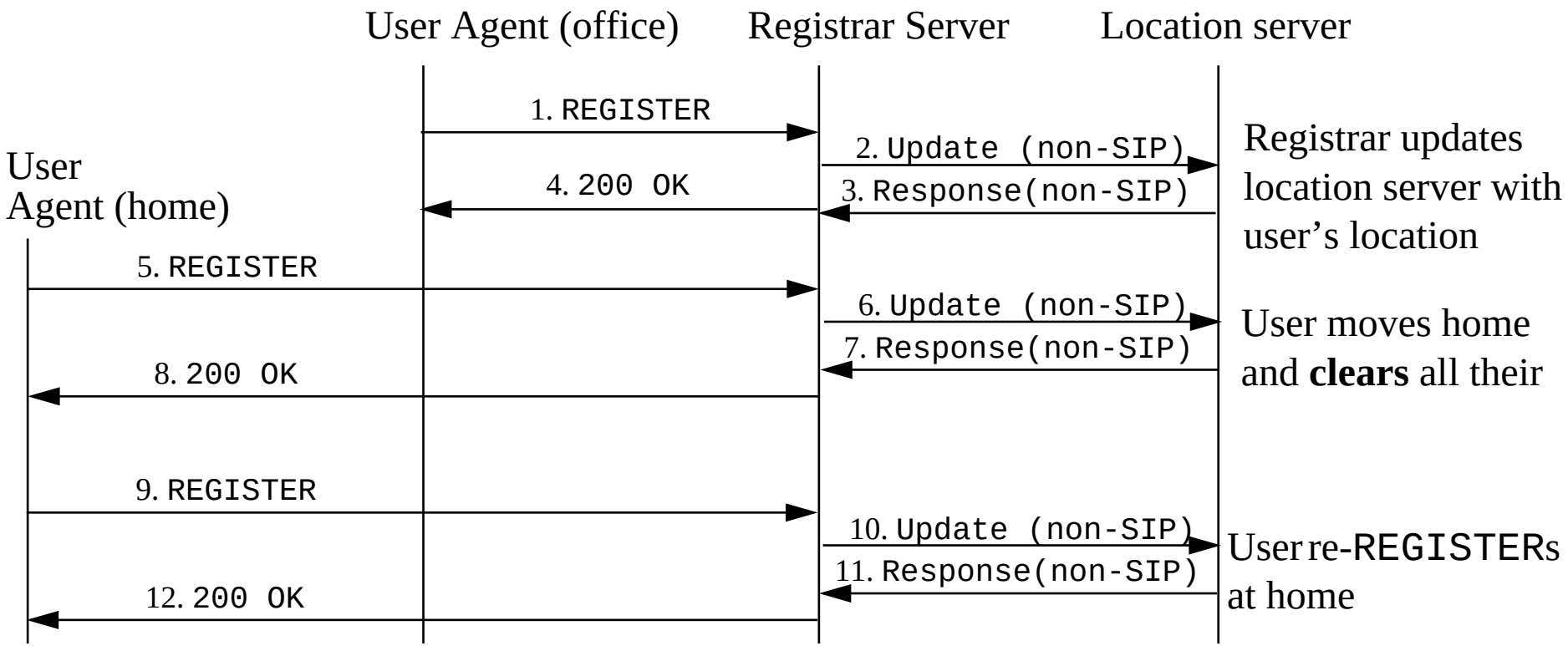
1. From the lecture notes "SIP Tutorial: Introduction to SIP" by Henry Sinnreich and Alan Johnston, *http://smuhandouts.com/8393/SIPTutorial.pdf*

Purpose of registration

User B registers in order to establish their current device and location

- Only *their* location server need know
- Therefore this scales well
- The location server need not disclose this location to "just anyone", but can apply various polices to decide who can learn of it

REGISTERing



REGISTER request includes one or more Contact headers:

```
Contact: <sip:UserA@4.3.2.1>;class=personal
Contact: <sip:UserA-msg-depot@voicemail.provider.com>;feature=voicemail
Contact: <sip:+13145551212@gateway.com;user=phone>;class=business
Contact: <sip:+13145553333@cellphone.com;user=phone>;mobility=mobile
Contact: <tel:+13145551212>
Contact: <mailto:UserA@hotmailer.com>
```

Details at: Sinnreich & Johnston, pp. 78-79 and **User Preferences** on page 242.

SIP Call Setup Attempt[1]

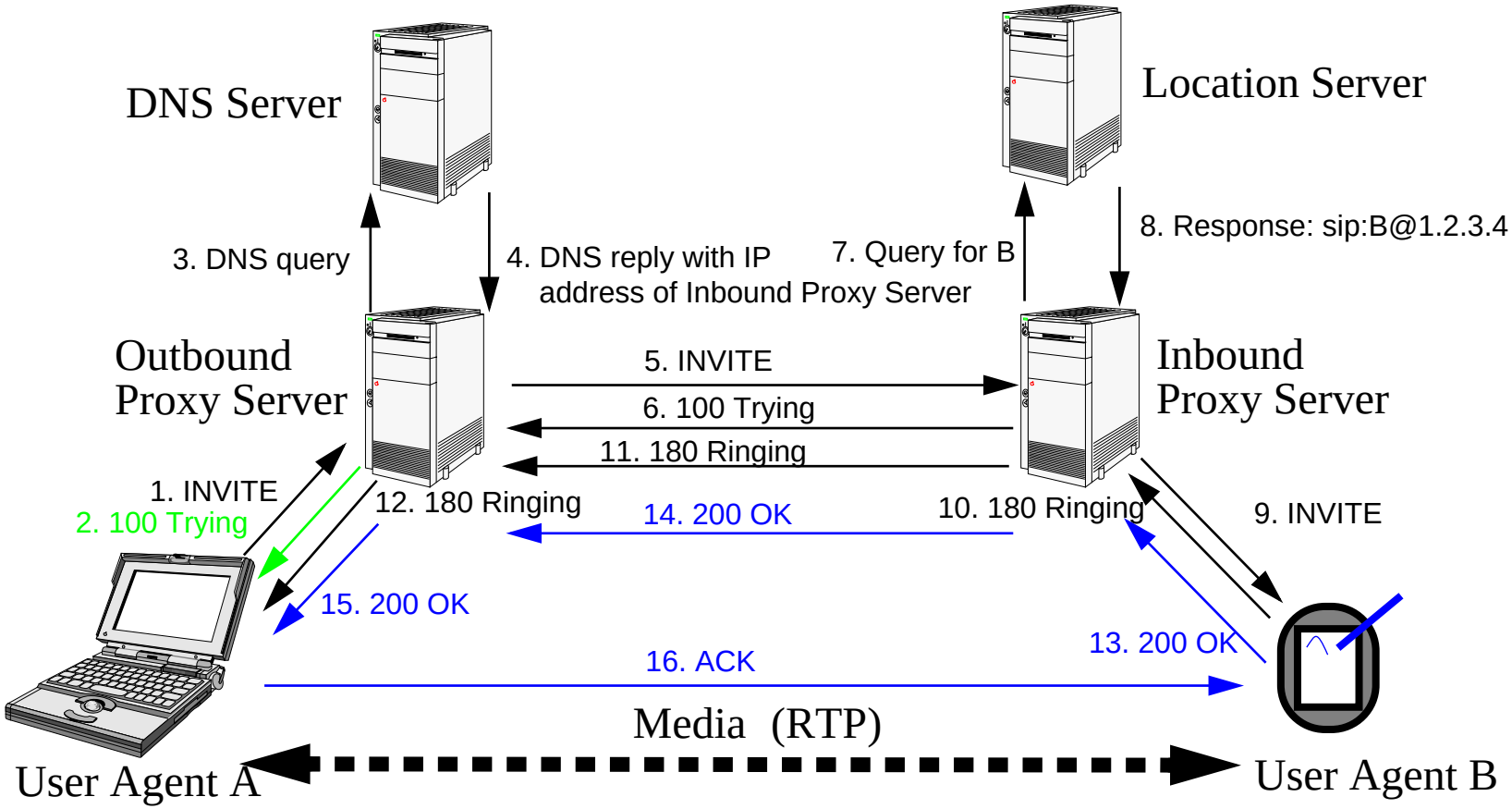
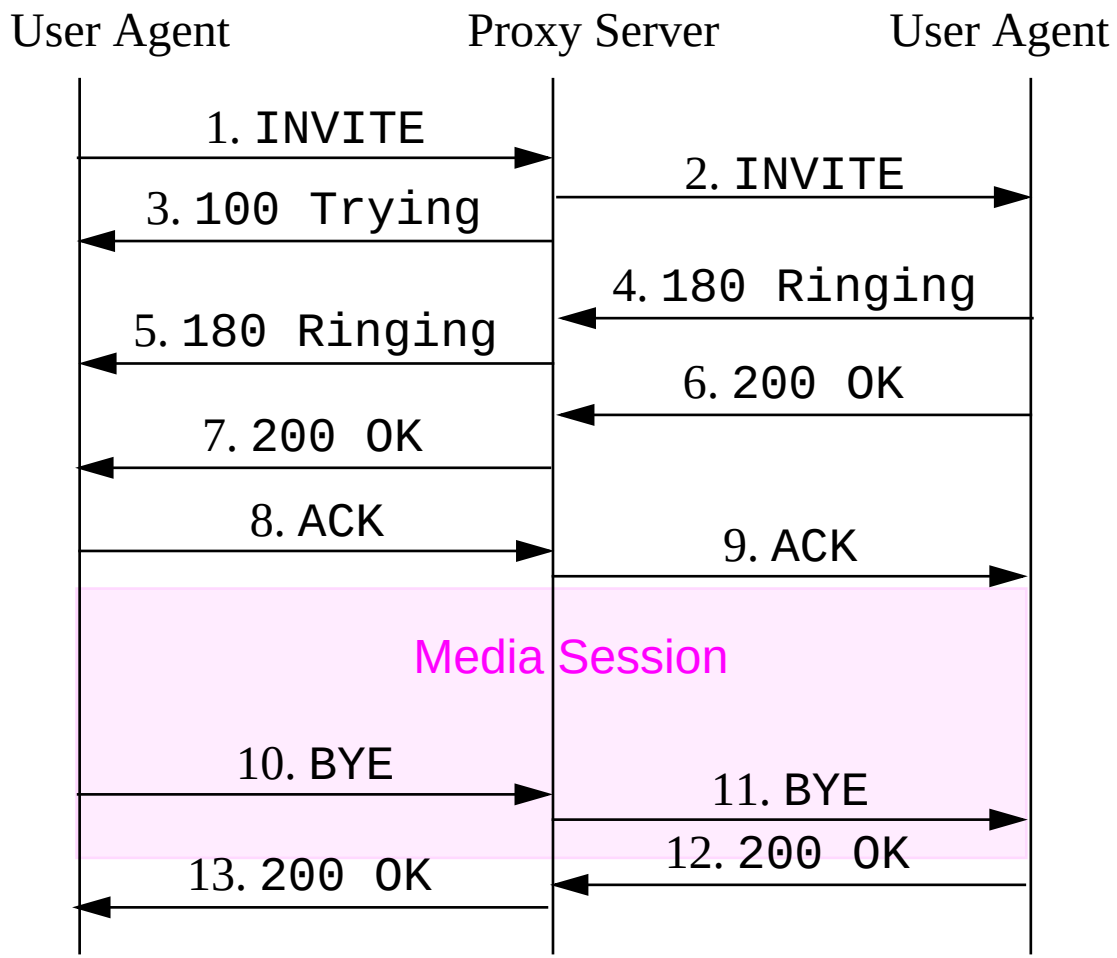


Figure 3: SIP Call Setup Attempt - when B has registered

1. Adapted from the lecture notes "SIP Tutorial: Introduction to SIP" by Henry Sinnreich and Alan Johnston, *http://smuhandouts.com/8393/SIPTutorial.pdf*

SIP Session - terminates with BYE



BYE causes the media session to be torn down.

Note: BYE like `INVITE` is an **end-to-end** method.

Authentication

Builds upon authentication schemes developed for HTTP (see RFC 2716), for example challenge/response, digest, …

Two forms:

- user agent-to-user agent
  - `401 Unauthorized` ⇒ Authentication Required
- user agent-to-server
  - `407 Proxy Authentication Required` ⇒ Authentication Required (reponse sent by a proxy/server)

Note: Any SIP request can be challenged for authentication.

Note: There is **no** *integrity* protection, for additional information see **SIP Security** on page 254.

SIP Extensions and Features

- Method Extensions
  - Unknown methods rejected by User Agent using `405` or `501` response
  - Listed in `Allow` header field
  - Proxies treat unknown methods as a non-`INVITE`
- Header Field Extensions
  - Unknown header fields are ignored by user agents and proxies
  - Some have feature tags registered, these can be declared in a `Supported` or `Require` header field
- Message Body Extensions
  - Unknown message body types are rejected with a `406` response
  - Supported types can be declared with an `Accept` header field
  - `Content-Disposition` indicates what to do with it
- Extension must define failback to base SIP specification.

$\Rightarrow$ No Profiling is needed

- unlike for example, Bluetooth!

SIP Method Extensions are defined in a number of RFCs.

RFC 3261 - New Services

- Customized ringing
  - A trusted proxy can insert an `Alert-Info` header field into an `INVITE`
- Screen Pops
  - A trusted proxy can insert an `Call-Info` header field into an `INVITE`
  - URI can be HTTP and can contain call control “soft keys”
- Callback
  - Reply-to and In-Reply-To header - to assist in returning calls
- Annoucement handling
  - UAS or proxy need not make a decision about playing an early media annoucement
    - Error response contains new `Error-Info` header field which contains the URI of the annoucement
  - UAC makes a decision based on the user’s interface

Beyond the PSTN: Presence & Instant Messaging

- **Presence**, i.e., Who is available?
- **Location**, i.e., Where are they?: office, home, traveling, …
- **Call state**: Are they busy (in a call) or not?
- **Willingness**: Are they available or not?
- **Preferred medium**: text message, e-mail, voice, video, …
- **Preferences** (*caller* and *callee* preferences)

See Sinnreich and Johnston's Chapter 11 (Presence and Instant Communications) & course *2G5565 Mobile Presence: Architectures, Protocols, and Applications*.

- Reuters has deployed a SIP-based instant-messaging platform for the financial services industry that has 50,000 users each week.
- IBM's NotesBuddy application for ~315k employees - an *experimental* messaging client that integrates instant messaging (IM), email, voice, and other communication.

Presence-Enabled Services

- Complex call screening
  - Location-based: home vs. work
  - Caller-based: personal friend or business colleague
  - Time-based: during my “working hours” or during my “personal time”
- Join an existing call ⇒ Instant Conferencing, group chat sessions, …
- Creating a conference when a specific group of people are all *available* and *willing* to be called
- New services that have **yet** to be invented!
- SIP Messaging and Presence Leveraging Extensions (SIMPLE) Working Group *http://www.ietf.org/html.charters/simple-charter.html*

Significance

- In July 2002, 3GPP adopted SIP for their signalling protocol (Release5)
- 3GPP adops SIMPLE as instant messaging/presence mechanism (Release6)

While there are some differences between the 3GPP and IETF points of view

From Henning Schulzrinne, "SIP - growing up", SIP 2003, Paris, January 2003, slide 5.

3GPP	IETF
Network does not trust the user	User only partially trusts the network
layer 1 and layer 2 specific	generic
walled garden	open access

Not suprisingly the 3GPP system for using SIP is rather complex with a number of new components: Proxy Call Session Control Function (P-CSFC), Interrogating Call Session Control Function (I-CSFC), Serving Call Session Control Function (S-CSFC), Home Subscriber Server (HSS), Application Server (AS), Subscription Locator Function (SLF), Breakout Gateway Control Function (BGCF), Media Gateway Control Function (MGCF), and Media Gateway (MGW)

Programmable “phone”

Programming environments

- Symbian
- Java
- Linux
- …

Avoids lock-in driven by operators and telecom equipment vendors

Greatly increases numbers of developers

⇒ more (new) services

⇒ more security problems

Erik Eliasson's miniSIP - as used in the test

miniSIP supports pluggable CODECs:

- each RTP packet says which codec was used
- SDP can specify multiple codecs each with different properties (including better than toll quality)
- tests used PCM ⇒ sending 50 packets of 160 byte RTP payload length (packet size is 176 bytes) per second (i.e. 64 Kbps), i.e., 20 ms between packets
  - time to transmit/receive a packet ~55-60 µs
  - Laptop ASUS 1300B with Pentium III processor, 700 MHz
  - 112 MB RAM (no swapping)
  - Operating System: SuSE Linux 7.1 Personal Edition
  - Security Services: confidentiality and message authentication (with Replay Protection)
  - Cryptographic Algorithms: AES in Counter Mode for the confidentiality and HMAC SHA1 for the message authentication
  - Lengths: master key: 16 bytes; salting key: 14 bytes; authentication key: 16 bytes; encryption key: 16 bytes; block: 128 bytes

Secure Real Time Protocol (SRTP) for securing the media data transport

Israel M. Abad Caballero, *Secure Mobile Voice over IP*, M.Sc. Thesis, June 2003.

ftp://ftp.it.kth.se/Reports/DEGREE-PROJECT-REPORTS/030626-Israel_Abad_Caballero-final-report.pdf

Sender behavior	Receiver behavior
	Read the SRTP packet from the socket.
Determine cryptographic context to use	Determine the cryptographic context to be used
Derive session keys from master key (via MIKEY)	Determine the session keys from master key (via MIKEY)
	If message authentication and replay protection are provided, check for possible replay and verify the authentication tag
Encrypt the RTP payload	Decrypt the Encrypted Portion of the packet
If message authentication required, compute authentication tag and append	If present, remove authentication tag
Send the SRTP packet to the socket	Pass the RTP packet up the stack

- AES CM (Rijndael) or Null Cipher for encryption (using libcrypto)
- HMAC or, Null authenticator for message authentication
- SRTP packet is 176 bytes (RTP + 4 for the authentication tag if message authentication is to be provided)
- Packet creation: RTP 3-5 μs; RTP+SRTP 76-80 μs (throughput 20Mbps)
  - ~1% of the time there are packets which take as long as 240 μs

Multimedia Internet KEYing (MIKEY) as the key management protocol

Johan Bilien, *Key Agreement for Secure Voice over IP*, M.Sc. Thesis, Dec. 2003.

ftp://ftp.it.kth.se/Reports/DEGREE-PROJECT-REPORTS/031215-Johan-Bilien-report-final-with-cover.pdf

Extends earlier thesis - Runs on a Laptop or iPAQ under linux

Secure Call Setup[1]

Total delay (in ms)	Calling Delay	Answering Delay
No security	19.5	9.5
MIKEY, shared key	20.9	10.5
MIKEY, Diffie-Hellman	52.5 (UDP)	47.6 (UDP)
	58.9 TCP)	48.9 (TCP)

- name-servers (BIND 8.2 on Linux 2.4, 500 MHz Pentium 3 laptops)
- root name-server ns.lab manages the delegation of minisip.com and ssvl.kth.se to their respective name server
- two routers (1.1 GHz Celeron desktops) perform static routing, and each router also runs a SIP server, SIP Express Router (SER v0.8.11))
- Alice and Bob use minisip, running on 1.4 GHz Pentium 4 laptops,running Linux 2.4

1. Johan Bilien, Erik Eliasson, and Jon-Olov Vatn, “Call establishment Delay for secure VoIP”, WiOpt’04: Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks, University of Cambridge,UK, 24-26 March, 2004

Communications and Privacy

- Encryption as the norm - even onetime pads are feasible
  - Since all speech and other media content will be in digital form, it will be trivial to provide encryption and authentication of all communication (if the participants want to)
  - traditional public telephony **less secure** than using: VPNs, SRTP, MIKEY, …
  - For WLANs: IEEE 802.11i security features along with 128-bit Advanced Encryption Standard (AES) encryption, …
- Identity hiding
  - Authentication when you mutually want to
- Mobile presence has to be done carefully
- Anonymous network access
- Location hiding & Privacy
  - Alberto Escudero-Pascual, http://www.it.kth.se/~aep
    - *Anonymous and Untraceable Communications - Location privacy in mobile internetworking*, Licentiate Thesis, June 2001
    - *Privacy in the Next generation Internet: Data Protection in the context of the European Union Policy*, Dissertation, Dec. 2002
- Location mis-direction ⇒ End of Sovereignty
- Traffic pattern hiding
- Traffic hiding

If WLANs are widely available

- How many different places to do you frequently spend time?
- What would happen if you had WLAN access in X% of these places? (perhaps with X > 90%)
- What if you also had VoIP service in all the places you have WLAN access?
  - For example, via a Cisco Wireless IP Phone 7920

⇒ Is there a business case for 3G in urban areas?

⇒ Is there a business case for 3G anywhere?

Handoffs for real-time media: J-O Vatn's upcoming dissertation

Future work

To combine:

- **Mobility**: WLAN + GPRS (via a private complete GSM/GPRS system *http://csd.ssvl.kth.se/monaco/main.htm*)
- **Security**: IPsec, TLS, SRTP+MIKEY, … + SIP ⇒ secure VoIP
- **Context and location awareness**: minimizing manual (re-)configuration as users move about and facilitating their interaction with each other & the things around them - Adaptive and Context-Aware Services (ACAS)[1]

⇒ **New services**: such as audio services - managing a 3D (or 4D) audio environment, *automatic* call diversion, …

In a challenging environment of **socially correlated** user movements (i.e., classes, meetings, etc.)

Questions: What services do students *want*? Which services do they *need*? How will this *change interactions* with other students, teachers, staff, … .

1. *http://psi.verkstad.net/acas/* (part of AWSI *http://www.wireless.kth.se/AWSI/*)

- **Why audio?** Because users can utilize audio interaction **while** *on the move*
- **Why PDAs?** Because they support both computing and communication in a small form-factor, it is possible to have multiple wireless interfaces, audio is good enough quality to use for entertainment (MP3files, streaming audio, voice interaction, and interactive voice), and we can have **enough** devices **which people will use on the move** to start to understand the effects of **correlation** and the **demands on the underlying infrastructure**[1].

1. HP grant “Applied Mobile Tech. Solutions in Learning Environments”

¿Questions?

For details, references, … see my lectures notes for:

- 2G1325/2G5564 Practical Voice Over IP (VoIP): SIP and related protocols *http://www.imit.kth.se/courses/2G1325/*
- Magnus Sjöstedt and Oskar Bergquist, VoIP regulatory issues, M.Sc. Thesis, June 2003 *ftp://ftp.it.kth.se/Reports/DEGREE-PROJECT-REPORTS/030627-Magnus_Sjostedt-and-Oskar_Bergquist-Report.pdf*

Books

- Luan Dang, Cullen Jennings, and David Kelly, *Practical VoIP: Using VOCAL*, O'Reilly, 2002, ISBN 0-596-00078-2
- Henry Sinnreich and Alan B. Johnston, *Internet Communications Using SIP: Delivering VoIP and Multimedia Services with Session Initiation Protocol*, Wiley, 2001, ISBN: 0-471-41399-2.

Software

- Vovida Open Communication Application Library (VOCAL) system - source code available from vovida.org
- SIP Express Router (SER) - http://www.iptel.org/ser/
- ...