2G1325/2G5564 Practical Voice Over IP (VoIP): SIP and related protocols
Spring 2007, Period 4

Lecture notes of G. Q. Maguire Jr.

For use in conjunction with:

VoIP a major market................................................................. 38
Handsets......................................................................................... 39
VoIP Chipsets .............................................................................. 40
Deregulation ⇒ New operators.................................................. 41
Deregulation ⇒ New Suppliers................................................... 42
Let them fail fast!.......................................................................... 43
Latency .......................................................................................... 44
VOIP Modes of Operation.......................................................... 45
IP based data+voice infrastructure .......................................... 46
Voice Gateway.............................................................................. 47
Voice over IP (VOIP) Gateways ............................................... 48
Voice representation ............................................................... 48
Signaling ....................................................................................... 49
Fax Support ................................................................................... 49
Management ................................................................................. 49
Compatibility .............................................................................. 50
Cisco’s Voice Over IP .................................................................. 51
Intranet Telephone System....................................................... 54
Wireless LANs.............................................................................. 55
<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Telia’s HomeRun</td>
<td>56</td>
</tr>
<tr>
<td>Ericsson’s &quot;GSM on the Net&quot;</td>
<td>57</td>
</tr>
<tr>
<td>VOIP vs. traditional telephony</td>
<td>58</td>
</tr>
<tr>
<td>Economics</td>
<td>59</td>
</tr>
<tr>
<td>VoIP vs. traditional telephony</td>
<td>60</td>
</tr>
<tr>
<td>Patents</td>
<td>61</td>
</tr>
<tr>
<td>Deregulation → Trends</td>
<td>63</td>
</tr>
<tr>
<td>Carriers offering VOIP</td>
<td>64</td>
</tr>
<tr>
<td>MCI (formerly WorldCom) Connection</td>
<td>65</td>
</tr>
<tr>
<td>Previously</td>
<td>65</td>
</tr>
<tr>
<td>Today</td>
<td>65</td>
</tr>
<tr>
<td>Level 3 Communications Inc.</td>
<td>66</td>
</tr>
<tr>
<td>TeliaSonera Bredbandstelefoni</td>
<td>67</td>
</tr>
<tr>
<td>Emulating the PSTN</td>
<td>68</td>
</tr>
<tr>
<td>Calling and Called Features</td>
<td>70</td>
</tr>
<tr>
<td>Beyond the PSTN: Presence &amp; Instant Messaging</td>
<td>71</td>
</tr>
<tr>
<td>Presence-Enabled Services</td>
<td>72</td>
</tr>
<tr>
<td>Three major alternatives for VoIP</td>
<td>73</td>
</tr>
<tr>
<td>...</td>
<td></td>
</tr>
</tbody>
</table>
Negatives ................................................................................................ 74
Deregulation ➞ New Regulations ................................................................. 75
Regulations in Sweden ............................................................................... 76
Programmable “phone” ............................................................................. 77
Conferences ................................................................................................. 78
Not with out problems .............................................................................. 79
References and Further Reading ................................................................. 80
Acknowledgements ..................................................................................... 84
Module 2: VoIP details ............................................................................... 85
Traditional Telecom vs. Datacom ................................................................. 86
VoIP details: Protocols and Packets ............................................................ 87
RTP and H.323 for IP Telephony ................................................................. 88
RTP, RTCP, and RTSP ................................................................................ 89
Real-Time Delivery ..................................................................................... 90
Packet delay ................................................................................................. 91
Dealing with Delay jitter ........................................................................... 92
Delay and delay variance (jitter) ................................................................. 93
Playout delay .................................................................................................................. 94
When to play .................................................................................................................... 95
Retransmission, Loss, and Recovery ............................................................................... 96
Patterns of Loss ............................................................................................................... 97
Loss concealment .............................................................................................................. 98
VoIP need not be “toll quality” .................................................................................... 99
RTP: Real-Time Transport Protocol .............................................................................. 100
Payload types .................................................................................................................. 101
Audio Encodings .............................................................................................................. 103
Timestamps ....................................................................................................................... 105
Stream translation and mixing ....................................................................................... 106
RTP Control Protocol (RTCP) ....................................................................................... 107
Compound Reports ......................................................................................................... 108
Proposed RTCP Reporting Extensions .......................................................................... 109
RTP translators/mixers .................................................................................................... 112
Synchronizing Multiple Streams .................................................................................... 113
RTP Transport and Many-to-many Transmission ............................................................ 114
<table>
<thead>
<tr>
<th>Issue</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Issues to be considered</td>
<td>136</td>
</tr>
<tr>
<td>Address Resolution</td>
<td>137</td>
</tr>
<tr>
<td>SIP timeline</td>
<td>138</td>
</tr>
<tr>
<td>SIP Invite</td>
<td>139</td>
</tr>
<tr>
<td>Bob’s response to Alice’s INVITE</td>
<td>140</td>
</tr>
<tr>
<td>ACK</td>
<td>141</td>
</tr>
<tr>
<td>SIP Invite (method/URI/version)</td>
<td>142</td>
</tr>
<tr>
<td>SIP Via</td>
<td>143</td>
</tr>
<tr>
<td>Dialog (Call leg) Information</td>
<td>144</td>
</tr>
<tr>
<td>SIP CSeq</td>
<td>145</td>
</tr>
<tr>
<td>SIP Contact</td>
<td>146</td>
</tr>
<tr>
<td>SIP Content Type and Length</td>
<td>147</td>
</tr>
<tr>
<td>SIP Max-Forwards</td>
<td>148</td>
</tr>
<tr>
<td>Other header fields</td>
<td>149</td>
</tr>
<tr>
<td>Several types of SIP Servers</td>
<td>150</td>
</tr>
<tr>
<td>SIP Trapezoid</td>
<td>151</td>
</tr>
<tr>
<td>SIP Call Setup</td>
<td>152</td>
</tr>
<tr>
<td>Topic</td>
<td>Page</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>SIP Call Setup Attempt</td>
<td>153</td>
</tr>
<tr>
<td>SIP Call Setup Attempt</td>
<td>154</td>
</tr>
<tr>
<td>SIP Presence</td>
<td>155</td>
</tr>
<tr>
<td>SIP B not Present</td>
<td>156</td>
</tr>
<tr>
<td>SIP Registration Example</td>
<td>157</td>
</tr>
<tr>
<td>Purpose of registration</td>
<td>158</td>
</tr>
<tr>
<td>REGISTERing</td>
<td>159</td>
</tr>
<tr>
<td>SIP Call Setup Attempt</td>
<td>160</td>
</tr>
<tr>
<td>SIP Session Termination using BYE</td>
<td>161</td>
</tr>
<tr>
<td>SIP Session Termination using CANCEL</td>
<td>162</td>
</tr>
<tr>
<td>CANCEL and OPTIONS</td>
<td>163</td>
</tr>
<tr>
<td>CANCEL</td>
<td>163</td>
</tr>
<tr>
<td>OPTIONS</td>
<td>163</td>
</tr>
<tr>
<td>Unsuccessful final responses are hop-by-hop</td>
<td>164</td>
</tr>
<tr>
<td>Authentication</td>
<td>165</td>
</tr>
<tr>
<td>SIP Method Extensions in other RFCs</td>
<td>166</td>
</tr>
<tr>
<td>SIP Extensions and Features</td>
<td>167</td>
</tr>
<tr>
<td>SIP Presence - Signed In</td>
<td>168</td>
</tr>
<tr>
<td>Topic</td>
<td>Page</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Gateways</td>
<td>192</td>
</tr>
<tr>
<td>Significance</td>
<td>193</td>
</tr>
<tr>
<td>References and Further Reading</td>
<td>194</td>
</tr>
<tr>
<td>SIP</td>
<td>194</td>
</tr>
<tr>
<td>ITU Services CS-1 and CS-2</td>
<td>198</td>
</tr>
<tr>
<td>Module 4: Session Announcement Protocol (SAP)</td>
<td>200</td>
</tr>
<tr>
<td>Session Announcement Protocol (SAP)</td>
<td>201</td>
</tr>
<tr>
<td>References and Further Reading</td>
<td>202</td>
</tr>
<tr>
<td>SAP</td>
<td>202</td>
</tr>
<tr>
<td>Module 5: Session Description Protocol (SDP)</td>
<td>203</td>
</tr>
<tr>
<td>Session Description Protocol (SDP)</td>
<td>204</td>
</tr>
<tr>
<td>Session Description Protocol (SDP)</td>
<td>205</td>
</tr>
<tr>
<td>SDP Message Details</td>
<td>208</td>
</tr>
<tr>
<td>Session description</td>
<td>209</td>
</tr>
<tr>
<td>SDP Offer/Response Example</td>
<td>210</td>
</tr>
<tr>
<td>SDP Response Example</td>
<td>211</td>
</tr>
<tr>
<td>Session Modification</td>
<td>212</td>
</tr>
<tr>
<td>Session modification (continued)</td>
<td>213</td>
</tr>
</tbody>
</table>
EUNM Timeline ................................................................. 235
Sweden’s ENUM Mapping ...................................................... 237
VISIONng Association ............................................................... 238
SIP goes beyond ENUM .......................................................... 239
References and Further Reading ............................................... 240
DNS ........................................................................ 240
ENUM ........................................................................ 240
Module 7: SIP Mobility ........................................................... 245
SIP Mobility ........................................................................ 246
Local Number Portability ......................................................... 247
References and Further Reading ............................................... 248
SIP Mobility ........................................................................ 248
Service Mobility .................................................................... 248
Module 8: SIP Service Creation ............................................... 249
SIP Service Creation ............................................................... 250
Services implemented by x ...................................................... 251
Services implemented by Extensions .......................................... 252
SIP Service Logic .................................................................. 253
Call Processing Language (CPL) .............................................. 254
SIP Common Gateway Interface (CGI) ................................................ 255
SIP Java Servlets .................................................................................. 256
JAIN APIs.................................................................................................. 257
US National Institute of Standards and Technology - SIP and Jain .... 260
Parlay ....................................................................................................... 261
SIP Request-URIs for Service Control .................................................. 262
Reason Header ..................................................................................... 263
Voice eXtensible Markup Language (VoiceXML ™).............................. 264
References and Further Reading ......................................................... 265
SIP Service Creation ............................................................................ 265
JAIN ........................................................................................................ 266
Parley ..................................................................................................... 266
SIP Request URI .................................................................................. 266
Reason Header ..................................................................................... 266
VoiceXML ............................................................................................... 267
Module 9: User Preferences ................................................................. 268
User Preferences .................................................................................. 269
Contact parameters ............................................................................. 270
Contact header example ...................................................................... 271
NAT traversal methods....................................................................................................... 297
STUN (Simple Traversal of UDP through NATs (Network Address Translation)) ........ 299
STUN steps....................................................................................................................... 300
UDP and TCP Firewall Traversal problems..................................................................... 301
UDP and TCP NAT Traversal problems............................................................................ 302
SIP Application Level Gateway (ALG) for Firewall Traversal .................................. 303
Middlebox communications (MIDCOM) ........................................................................ 304
Application aware Middlebox ........................................................................................ 305
Security flaws in Abstract Syntax Notation One (ASN.1) ............................................ 306
Swedish Electronic Communications Act ....................................................................... 308
Recording of Call Contents ........................................................................................... 309
Privacy & Lawful Intercept (LI)..................................................................................... 310
Reasonably Available Information.................................................................................... 311
EU privacy and Lawful Intercept (LI) ............................................................................ 312
Intercept architecture ....................................................................................................... 313
Voice over IP Security Alliance ....................................................................................... 314
<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>VoIP quality over IEEE 802.11b</td>
<td>374</td>
</tr>
<tr>
<td>Application Policy Server (APS)</td>
<td>375</td>
</tr>
<tr>
<td>References and Further Reading</td>
<td>376</td>
</tr>
<tr>
<td>Module 15: SIP Applications</td>
<td>382</td>
</tr>
<tr>
<td>Session Initiation Protocol Project INvestiGation (SIPPING)</td>
<td>383</td>
</tr>
<tr>
<td>Application Service Components</td>
<td>385</td>
</tr>
<tr>
<td>Advantages</td>
<td>386</td>
</tr>
<tr>
<td>Collecting DTMF digits for use within a service</td>
<td>387</td>
</tr>
<tr>
<td>Repsonse “3. 200 OK” looks like:</td>
<td>388</td>
</tr>
<tr>
<td>Controller issues a “re-Invite” at 11 which looks like:</td>
<td>389</td>
</tr>
<tr>
<td>Voice Portal Service using Interactive Voice Response (IVR)</td>
<td>390</td>
</tr>
<tr>
<td>Managing Services</td>
<td>391</td>
</tr>
<tr>
<td>Lots more services</td>
<td>392</td>
</tr>
<tr>
<td>References and Further Reading</td>
<td>393</td>
</tr>
<tr>
<td>SIPPING</td>
<td>393</td>
</tr>
<tr>
<td>Module 16: More than Voice</td>
<td>394</td>
</tr>
<tr>
<td>Non-voice Services and IP Phones</td>
<td>395</td>
</tr>
<tr>
<td>XML</td>
<td>396</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Invoking RTP streams</td>
<td>397</td>
</tr>
<tr>
<td>More details</td>
<td>398</td>
</tr>
<tr>
<td>Services for sale - building a market</td>
<td>399</td>
</tr>
<tr>
<td>Network Appliances</td>
<td>400</td>
</tr>
<tr>
<td>Proposed Extension of SIP</td>
<td>401</td>
</tr>
<tr>
<td>Service Location Protocol (SLP) URL</td>
<td>402</td>
</tr>
<tr>
<td>Example service</td>
<td>403</td>
</tr>
<tr>
<td>Example of service portability</td>
<td>405</td>
</tr>
<tr>
<td>Text</td>
<td>408</td>
</tr>
<tr>
<td>Interleaved text</td>
<td>408</td>
</tr>
<tr>
<td>Timed Text</td>
<td>408</td>
</tr>
<tr>
<td>References and Further Reading</td>
<td>409</td>
</tr>
<tr>
<td>Phone Services</td>
<td>409</td>
</tr>
<tr>
<td>Network Appliances</td>
<td>409</td>
</tr>
<tr>
<td>Module 17: VOCAL</td>
<td>412</td>
</tr>
<tr>
<td>VOCAL System Overview</td>
<td>413</td>
</tr>
<tr>
<td>VOCAL Servers</td>
<td>414</td>
</tr>
<tr>
<td>Scaling of a VOCAL system</td>
<td>415</td>
</tr>
<tr>
<td>For comparison with a PBX</td>
<td>416</td>
</tr>
</tbody>
</table>
Module 1: Introduction

Lecture notes of G. Q. Maguire Jr.

For use in conjunction with:

Welcome to the course!

The course should be fun.

We will dig deeper into Voice over IP - with a focus on SIP and related protocols, but may also examine some of the other protocols which are used.

Information about the course is available from the course web page

http://www.imit.kth.se/courses/2G1325/

Note that the above URL will change - due to the reorganization of KTH to:

http://www.cos.ict.kth.se/education/msc/ccs/courses/2G1325/
Staff Associated with the Course

Instructor (Kursansvarig)

prof. Gerald Q. Maguire Jr.  <maguire@it.kth.se>

Administrative Assistant: recording of grades, registration, etc.

Irina Radulescu <irina.radulescu at wireless.kth.se>
Goals, Scope and Method

Goals of the Course

• To understand what Voice over IP (VoIP) systems are, their basic architectures, and the underlying protocols
• To be able to read and understand the literature.
• To provide a basis for your own research and development in this area.

Scope and Method

• You are encouraged to examine the Vovida Open Communication Application Library (VOCAL) system\(^1\), SIP Express Router\(^2\), Minisip\(^3\), …
  • to understand both the details of the system(s) and
  • to abstract from these details some architectural features and examine some places where it can be extended (thus using it as a platform on which you can explore).
• You will demonstrate your knowledge by writing a written report and giving an oral presentation describing your project.

---

1. The source code is available from [vovida.org](http://vovida.org)
2. The source code is available from [http://www.iptel.org/ser/](http://www.iptel.org/ser/)
3. The source code is available from [http://www.minisip.org/](http://www.minisip.org/)
Prerequisites

- Internetwork (2G1305) or
- Equivalent knowledge in Computer Communications (this requires permission of the instructor)
The focus of the course is on what Voice over IP (VoIP) systems are, their basic architectures, and the underlying protocols. We will primarily focus on the Session Initiation Protocol (SIP) and related protocols.

The course consists of 10 hours of lectures and a project of ~50 hours effort.
Topics

• Session Initiation Protocol (SIP)
• Real-time Transport Protocol (RTP)
• Real-time Streaming Protocol (RTSP)
• Common Open Policy Server (COPS)
• SIP User Agents
• Location Server, Redirect Server, SIP Proxy Server, Registrar Server, ...
  , Provisioning Server, Feature Server
• Call Processing Language (CPL)
Examination requirements

- Written and Oral project reports

Grades: U, 3, 4, 5
Project

Goals: to gain analytical or practical experience and to show that you have mastered some knowledge in this area and to encourage you to find a topic which interests you (since this will motivate you to really understand the material)

- Can be done in a group of 1 to 3 students (formed by yourself). Each student must contribute to the final written and oral reports.
- Discuss your ideas about topics with the instructor before starting.
Assignment Registration and Report

• Registration: Mon. 9-April 2007, to <maguire@it.kth.se>, subject=2G1325 topic
  • Group members, leader
  • Topic selected

• Written report
  • Length of the final report should be 10 pages (roughly 5,000 words) for each student.
  • Report may be in the form of a collections of papers, with each paper suitable for submission to a conference or journal
  • Contribution by each member of the group - must be clear (in the case where the report is a collection of papers - the role of each member of the group can be explained in the overall introduction to the papers.
  • The report should clearly describe: 1) what you have done; 2) who did what; if you have done some implementation and measurements you should describe the methods and tools used, along with the test or implementation results, and your analysis.


oral presentations scheduled Friday 25-May 2007 from 08:00-18:00

• Send e-mail with URL link to <maguire@it.kth.se>

Note that it is OK to start working well in advance of the deadlines!

1. Alternative dates can be scheduled with the instructor’s permission.
The course will mainly be based on the book:


We will refer to other books, articles, and RFCs as necessary. A list of interesting literature will be available on the course web page and in the references and further reading section of each lecture module.

In addition, you will be searching & reading the literature in conjunction with your projects. Please make sure that you properly reference your sources in your report. If you use someone else’s words - they must be clearly indicated as a quotation. Note also that individual figures have their own copyrights, so if you are going to use a figure/picture/… from some other source, you need to both cite this source and have the copyright owner’s permission to use it.
Lecture Plan

• Introduction
  • Course arrangement
  • Set the context of VoIP, both technically and economically

• VoIP details
  • Session Initiation Protocol (SIP)
  • Session Description Protocol (SDP)
  • DNS and ENUM

• Mobility

• Service Creation

• User preferences

• Security, NATs, and Firewalls

• SIP Telephony

• Conferencing

• Mixed Internet - PSTN services

• AAA and QoS

• More than just voice!
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)  

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

- We will focus on VoIP, largely *independently* of the underlying network, i.e., LAN, Cellular, WLAN, PAN, Ad hoc, ….
Internetworking

Internetworking is

• based on the interconnection (concatenation) of multiple networks
• accommodates multiple underlying hardware technologies by providing a way to interconnect heterogeneous networks and makes them inter-operate.

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 endpoints, and the user’s preference(s).

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

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

This is unlike the fixed quality of traditional phone systems.
VoIP a major market

Voice over IP has developed as a major market - which began with H.323 and has now moved to SIP. There are increasing numbers of users and a large variety of VoIP hardware and software on the market. With increasing numbers of vendors, the competition is heating up - is it a maturing market?

“Cisco began selling its VoIP gear to corporations around 1997, but until the past year, sales were slow. Cisco notes that it took more than three years to sell its first 1 million VoIP phones, but the next 1 million took only 12 months.”

Ben Charny , “Is VoIP pioneer Cisco losing momentum?”, CNET News.com, September 17, 2003, 4:00 AM PT

As of their fiscal year 2005 (ending July 30, 2005), they had shipped their 6 millionth IP phone[10]. (This is 3 million more than one year earlier.)
Handsets

There are now lots of USB attached VoIP handsets:

- http://www.clarisys.com/ Claritel-i750, Plantronics DA60, Net2Phone, Linksys, D-Link, IPEVO, ...

WLAN Handsets:

- starting with Symbol Technologies’s NetVision® Data Phone
  - runs speech recognition software in a network attached server
  - unfortunately it uses a proprietary protocol between the handset and their server, but I expect others will make similar devices which will **not** have this mis-feature.
- Toshiba announced two PDAs (models e800 and e805) with VoIP and Wi-Fi; as have NTT’s DoCoMo and Sony Corp.

VoIP cellular handsets, e.g. TTPCom’s combined GSM & 802.11 ("GSM.11"), Nokia, Qtek, …
VoIP Chipsets


Two ICs:

- **T8302 IPT_ARM (Advanced RISC Machine)**
  - Up to 57.6 MHz general-purpose processor
  - controls the system I/O: two 10/100Base-T Ethernets, USB, IrDA, SPI, 16 programmable I/O pins (some could be used to interface to an LCD module), …
  - general telephone control features: 7 row outputs and 8 column inputs/outputs to control up to 56 LEDs and scan up to 56 keys, 6 different flash rates, …

- **T8301 IPT_DSP (digital signal processor)**
  - Based on Agere Systems DSP1627 digital signal processor core running at 80 MIPS
  - single-cycle multiply accumulate instruction supports voice compression/decompression and echo cancellation algorithms
  - Includes two 16-bit digital-to-analog (D/A), one 16-bit analog-to-digital converters (A/D), low-pass filters, audio amplifier, lots of buffers (for input and output)

A special feature is **acoustic** echo cancellation to enable high quality speakerphone. See also [4].
Deregulation ⇒ New operators

Lots of new actors appeared as operators:

- **MCI (formerly Worldcom)** - [http://www.mci.com/](http://www.mci.com/)
- **Qwest** - [http://www.qwest.com/](http://www.qwest.com/)
- **Level3** [http://www.level3.net/](http://www.level3.net/)
  - (3)Voice, an IP based long distance service using Softswitch technology
- **Vonage** - [http://www.vonage.com](http://www.vonage.com)
  - 2.2 million subscriber lines as of 2006.12.31
- **Skype™ Technologies** [http://www.skype.com/](http://www.skype.com/)
  - “Skype is free Internet telephony that just works.”
  - 522,932,765 downloads as of 2007.03.27 at 01:15:01; 200 million downloads as of 2005.11.08 (67,430,762 downloads as of 2005.02.10)
    - in 2005: ~1 Million downloads/day, downloads at peak are ~0.5 Gbit/sec
  - 5,512,395 simultaneous users (as of 2007.03.27); 9 million on 2007.01.29; 8 million on 2006.11.09; > 5 million simultaneous users on 2006.01.23 [14]
  - statistics as an RSS feed at: [http://share.skype.com/stats_rss.xml](http://share.skype.com/stats_rss.xml) updated every few minutes

- …
Deregulation ⇒ New Suppliers

Lots of new actors as equipment suppliers:

- Cisco, 3Com, Nortel Networks, ...

Traditional telecom equipment vendors buying datacom vendors.

Lots of mergers and acquisitions among datacom vendors.

As of Fall 2002, many of these vendors (similar to operators) were reorganizing, selling off divisions, reducing staffing, … -- due to the Telecom meltdown! However, some have survived (or been reborn).

For a list of SIP products see: http://www.pulver.com/products/sip/
Let them fail fast!

We hold that the primary cause of current telecom troubles is that Internet-based end-to-end data networking has subsumed (and will subsume) the value that was formerly embodied in other communications networks. This, in turn, is causing the immediate obsolescence of the vertically integrated, circuit-based telephony industry of 127 years vintage.

Izumi Aizu, Jay Batson, Robert J. Berger, et al.,

The extent of this transformation is well described in their complete letter which recommends that the FCC:
• “Resist at all costs the telephone industry’s calls for bailouts. The policy should be one of "fast failure."
• Acknowledge that non-Internet communications equipment, while not yet extinct, is economically obsolete and forbear from actions that would artificially prolong its use.
• Discourage attempts by incumbent telephone companies to thwart municipal, publicly-owned and other communications initiatives that don’t fit the telephone company business model.
• Accelerate FCC exploration of innovative spectrum use and aggressively expand unlicensed spectrum allocation.”
For example:

**Round-trip times from dumburken.it.kth.se**

(as of 2004.03.26)

<table>
<thead>
<tr>
<th>Destination</th>
<th>min (ms)</th>
<th>avg (ms)</th>
<th>max (ms)</th>
<th>hops</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local LANs (<a href="http://www.imit.kth.se">www.imit.kth.se</a>)</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>to northern Sweden (cdt-lisa.cdt.luth.se)</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>12</td>
</tr>
<tr>
<td>to Austria (<a href="http://www.tu-graz.ac.at">www.tu-graz.ac.at</a>)</td>
<td>38</td>
<td>39</td>
<td>44</td>
<td>14</td>
</tr>
<tr>
<td>To my machine in eastern US (via an SDSL link)</td>
<td>122</td>
<td>124</td>
<td>135</td>
<td>15</td>
</tr>
<tr>
<td>To US west coast (<a href="http://www.stanford.edu">www.stanford.edu</a>)</td>
<td>199</td>
<td>199</td>
<td>200</td>
<td>23</td>
</tr>
<tr>
<td>To Australia (<a href="http://www.uow.edu.au">www.uow.edu.au</a>) {via the US west coast}</td>
<td>350</td>
<td>350</td>
<td>351</td>
<td>20</td>
</tr>
</tbody>
</table>

---

**Figure 2:** Usability of a voice circuit as a function of end-to-end delay (adapted from a drawing by Cisco\(^{a}\))

\(a\) (this was at http://www.packeteer.com/solutions/voip/sld006.htm)
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
  - see Time Warner’s “Telecom One Solution”
- 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
IP based data+voice infrastructure

- Public cells
  - Handset
- Exchange
- E-1 to PSTN
- Gateway
- Workstation
- Home
- FW/Switch
- Voice Gateway
- Workstation
- Office
- In-building WLAN system
  - AP
  - Handset
- AP
- Internet
- Mobile
- Router

Maguire
maguire@it.kth.se

IP based data+voice infrastructure
2007.03.27

Module 1: 46 of 84
Practical Voice Over IP (VoIP): SIP and related protocols
Use access servers filled with digital modems (currently (formerly?) used for current analog modem pools) as voice gateways or special purpose gateways such as that of Li Wei [5].
Voice over IP (VOIP) Gateways

Gateways not only provide basic telephony and fax services, but can also enable lots of value-added services, e.g., call-centers, integrated messaging, least-cost routing, … .

Such gateways provide three basic functions:

- Interface between the PSTN network and the Internet
  Terminate incoming synchronous voice calls, compress the voice, encapsulate it into packets, and send it as IP packets. Incoming IP voice packets are unpacked, decompressed, buffered, and then sent out as synchronous voice to the PSTN connection.

- Global directory mapping
  Translate between the names and IP addresses of the Internet world and the E.164 telephone numbering scheme of the PSTN network.

- Authentication and billing

**Voice representation**

Commonly: ITU G.723.1 algorithm for voice encoding/decoding or G.729 (CS-ACELP voice compression).
Signaling

Based on the H.323 standard on the LAN and conventional signaling will be used on telephone networks.

NB: In conventional telephony networks signalling only happens at the beginning and end of a call. See Theo Kanter’s dissertation for what can be enabled via SIP so that you can react to other events.

Fax Support

Both store-and-forward and real-time fax modes.

• In store-and-forward the system records the entire FAX before transmission.

Management

Full SNMP management capabilities via MIBs (Management Information Base)

• provided to control all functions of the Gateway
• Extensive statistical data will be collected on dropped calls, lost/resent packets, and network delays.
Compatibility

De jure standards:
- ITU G 723.1/G.729 and H.323
- VoIP Forum IA 1.0

De facto standards:
- Netscape’s Cooltalk
- Microsoft’s NetMeeting (formerly H.323, now SIP)

Session Initiation Protocol (SIP) [RFC 2543] is much simpler than H.323
Cisco’s Voice Over IP

Enables Cisco 3600 series routers to carry live voice traffic (e.g., telephone calls and faxes) over an IP network.

They state that this could be used for:

- Toll bypass
- Remote PBX presence over WANs
- Unified voice/data trunking
- POTS-Internet telephony gateways

Uses Real-Time Transport Protocol (RTP) for carrying packetized audio and video traffic over an IP network.

Cisco 3600 supports a selection of CODECs:

- G.711 A-Law 64,000 bits per second (bps)
- G.711 u-Law 64,000 bps
- G.729 8000 bps
Cisco 3800 supports even more CODECs:

- ITU G.726 standard, 32k rate
- ITU G.726 standard, 24k rate
- ITU G.726 standard, 16k rate
- ITU G.728 standard, 16k rate (default)
- ITU G.729 standard, 8k rate

By using Voice Activity Detection (VAD) - you only need to send traffic if there is something to send {Note: telecom operators like this because it enables even higher levels of statistical multiplexing}.

An interesting aspect is that users worry when they hear absolute silence, so to help make them comfortable it is useful to play noise when there is nothing useful to output. Cisco provide a “comfort-noise command to generate background noise to fill silent gaps during calls if VAD is activated”.

Cisco 3600 series router can be used as the voice gateway with software such as Microsoft NetMeeting.
Cisco 3800 also supports “fax-relay” - at various rates either current voice rate or 2,400/4,800/7,200/9,600/14,400 bps fax rates.

For further information see

http://www.cisco.com/univercd/cc/td/doc/product/software/ios113ed/113t/113t_1/voip/config.htm
Intranet Telephone System

On January 19, 1998, Symbol Technologies and Cisco Systems announced that they had combined the Symbol Technologies’ NetVision™ wireless LAN handset and Cisco 3600 to provide a complete wireless local area network telephone system based on Voice-Over-IP technology.

The handset uses a wireless LAN (IEEE 802.11) infrastructure and a voice gateway via Cisco 3600 voice/ fax modules. The system conforms to H.323.

"I believe that this is the first wireless local area network telephone based on this technology" -- Jeff Pulver

Seamless roaming via Symbol’s pre-emptive roaming algorithm with load balancing.

Claims each cell can accommodate ~25 simultaneous, full-duplex phone calls.

Wireless LANs

“The wireless workplace will soon be upon us”\(^1\)

Telia has strengthened its position within the area of radio-based data solutions through the acquisition of Global Cast Internetworking. The company will primarily enhance Telia Mobile’s offering in wireless LANs and develop solutions that will lead to the introduction of the wireless office. A number of different alternatives to fixed data connections are currently under development and, *later wireless IP telephony will also be introduced.*

…

The acquisition means that Telia Mobile has secured the resources it needs to maintain its continued expansion and product development within the field of radio-based LAN solutions. *Radio LANs are particularly suitable for use by small and medium-sized companies as well as by operators of public buildings such as airports and railway stations.*

Today’s radio-LAN technology is based on *inexpensive products that do not require frequency certification.* They are *easy to install* and are often used to replace cabled data networks in, for example, large buildings.

…”

---

Telia’s HomeRun

http://www.homerun.telia.com/

A subscription based service to link you to your corporate network from airports, train stations, ferry terminals, hotels, conference centers, etc. via WLAN.

Look for Telia’s HomeRun logo:
Ericsson’s "GSM on the Net"

- Provide communication services over an integrated GSM-IP (Internet Protocol) network
- Support local and global mobility
- Support multimedia capabilities and IP-based applications
- Uses small radio base stations to add local-area GSM coverage to office LANs
- Provides computer-telephony integration: applications include web-initiated telephony, directory-assisted dialing, unified messaging and advanced conferencing and application-sharing using voice datacoms and video.
VOIP vs. traditional telephony

As of 2003 approx. 14% of International traffic to/from the US is via VoIP, based on 24 billions minutes vs. 170.7 billion minutes via PSTN [11] (the article cites the source of data as TeleGeography Research Group/Primetrica Inc.) For further statistics see: http://www.telegeography.com/products/tg/index.php


There is a move for traditional operators to replace their exchanges with IP telephony, see Niels Herbert and Göte Andersson, “Telia ersätter all AXE med IP-telefoni”, Elektronik Tidningen, #3, 4 March 2005, page 4.

For information about the development of the AXE switches see [12].
Economics


"What is the reality in the battle over packet-versus-circuit telephony, and what is hype?

Looking at the potential savings by cost element, it is clear that in 1998, access arbitrage is the major economic driver behind VOIP. By 2003, we anticipate that switched-access arbitrage will diminish in importance, as the ESP exemption disappears and/or access rates drop to true underlying cost.

However, we believe that the convergence between voice and data via packetized networks will offset the disappearance of a gap in switched access costs. As a result, VOIP will continue to enjoy a substantial advantage over circuit-switched voice. Indeed, as voice/data convergence occurs, we see standalone circuit-switched voice becoming economically nonviable."

Note: Enhanced Service Provider (ESP) exemption means that ISPs do not pay access charges to local phone companies {since the ISP just receives calls from users}
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

**NB. These new services can be far broader than traditional telephony services.**

---

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

\(^2\) see “Regulations in Sweden” on page 76
Mixing voice and data in the LAN goes back to at least this patent:

US 4581735 : Local area network packet protocol for combined voice and data transmission
INVENTORS: Lois E. Flamm and John O. Limb
ASSIGNEES: AT&T Bell Laboratories, Murray Hill, NJ
ISSUED: Apr. 8, 1986
FILED: May 31, 1983

ABSTRACT: In order to control the transfer of packets of information among a plurality of stations, the instant communications system, station and protocol contemplate first and second oppositely directed signal paths. At least two stations are coupled to both the first and the second signal paths. A station reads one signal from a path and writes another signal on the path. The one signal is read by an arrangement which electrically precedes the arrangement for writing the other signal. Packets are transmitted in a regular, cyclic sequence. A head station on a forward path writes a start cycle code for enabling each station to transmit one or more packets. If a station has a packet to transmit, it can read the bus field of a packet on the forward path. Responsive thereto, a logical interpretation may
be made as to whether the forward path is busy or is not busy. If the path is not busy, the packet may be written on the path by overwriting any signal thereon including the busy field. If the path is busy, the station may defer the writing until the path is detected as not busy. In order to accommodate different types of traffic, the head station may write different start cycle codes. For example, a start-of-voice code may enable stations to transmit voice packets; a start-of-data code may enable stations to transmit data packets, etc. for the different types of traffic. Further, the start cycle codes may be written in a regular, e.g., periodic, fashion to mitigate deleterious effects, such as speech clipping. Still further, the last station on the forward path may write end cycle codes in packets on a reverse path for communicating control information to the head station. Responsive to the control information, the head station may modify the cycle to permit the respective stations to, for example, transmit more than one packet per cycle or to vary the number of packet time slots, which are allocated to each of the different types of traffic.
Deregulation ⇒ Trends

- replacing multiplexors with Routers/Switches/… << 1/10 circuit swi. cost
- Standard telco interfaces being replaced by datacom interfaces
- New Alliances:
  - HP/AT&T Alliance, 3Com/Siemens, Bay/Ericsson, Cabletron/Nortel, Alcatel integrating Cisco IOS software technology, Ericsson Radio Systems & Cisco Systems collaborate wireless Internet services, …
- future developments building on VOIP
  ♦ Fax broadcast, Improved quality of service, Multipoint audio bridging, Text-to-speech conversion and Speech-to-Text conversion, Voice response systems, …
  ♦ Replacing the wireless voice network’s infrastructure with IP:
    U. C. Berkeley’s ICEBERG: Internet-based core for CEllular networks BEyond the thiRd Generation

See the Univ. of California at Berkeley ICEBERG project report:
http://iceberg.cs.berkeley.edu/release/

⇒ Telecom (only) operators have no future
⇒ Telecom (only) companies have no future
Carriers offering VOIP

“Equant, a network services provider, will announce tomorrow that it is introducing voice-over-frame relay service in 40 countries, ... The company says customers can save 20% to 40% or more by sending voice traffic over its frame relay network. "This is the nearest you’re going to get to free voice," says Laurence Huntley, executive VP of marketing for Equant Network Service. ... Equant isn’t alone in its pursuit to send voice traffic over data networks. Most of the major carriers are testing services that would send voice over data networks. ... .”¹

• October 2002:
  • Verizon offering managed IP telephony via **IPT Watch** for US$3-4/month
  • WorldCom offering SIP based VoIP for DSL customers for US$50-60/month for unlimited local, domestic long distance, and data support {price does **not** include equipment at US$200-300 per phone and DSL/Frame relay/ATM connection} The Service Level Agreement (SLA) specifies >99.9% network availability, <55ms round trip latency, and >99.5% packet delivery.

• December 2004:
  • Verizon offering VoiceWing - with unlimited calling within the US for US$34.95/month
  • “As we see the industry fundamentals continue to shift, the future will be about the convergence of computing and telecommunications. And where these two worlds meet is where MCI will be.” -- Michael D. Capellas, MCI CEO ²

---

¹. Mary E. Thyfault, Equant To Roll Out Voice-Over-Frame Relay Service, InformationWeek Daily, **10/21/98**.

². [http://global.mci.com/about/publicpolicy/voip/](http://global.mci.com/about/publicpolicy/voip/)
MCI (formerly WorldCom) Connection

Previously
- 3 or more separate networks (often each had its own staff!)
- Duration/geography-based pricing
- Expensive moves, adds, and changes (typically 1+ move/person/year)
- Standalone applications - generally expensive
- Closed PBX architecture

Today
- via gateway to the PSTN, service expands beyond the LAN to the WAN
- centralized intelligence is offered; customers utilize a Web browser to control and manage their network
- MCI incurs the costs of buying major equipment, thus limiting customer’s risk and capital investment
- **One** source for all services
- Easy mobility
- Choice of vendors for Customer Premises Equipment (CPE)
Introduced (3) VoIP Toll Free service: “a toll-free calling service across the United States, rounding out its local and long distance voice over Internet protocol offerings.”


Level 3 sells services to carriers, who then offer VoIP and data services to their customers.

Uses **softswitch** networking technology to convert voice signals from the PSTN to IP packets and conversely converts packets to voice signals when a call is routed to the public switched network. (>30 x 10^9 minutes of calls per month - as of January 13, 2005)
February 5th, 2004 TeliaSonera announce their *residential* broadband telephony service using server and client products from Hotsip AB (www.hotsip.com). In addition to telephony, the service includes: video calls, presence, and instant messaging.[7]

- The startup cost is 250 kr and the monthly cost 80 kr.
- Calls to the fixed PSTN network are the same price as if you called from a fixed telephone in their traditional network.
- Customers get a telephone number from the “area/city” code 075 (i.e., +46 75-15xxxxxxx)
- They do **not** support calls to “betalsamtal” (0900-numbers)
Emulating the PSTN

Many people feel that VoIP will really only “take off” when it can really emulate all the functions which users are used to in the PSTN:

• Integration with the web via: Click-to-connect
• “Dialing” an e-mail address or URL {digits vs. strings}
• Intelligent network (IN) services:
  • Call forward, busy
  • Call forward, no ans.
  • Call forward, uncond.
  • Call hold
  • Call park
  • Call pick-up
  • Call waiting
  • Consultation hold
  • Do not disturb
  • Find-me
  • Incoming call screen/Outgoing call screen
  • Secondary number in/Secondary number out
  • Three-way conference
  • Unattended transfer
• additional PBX features (which in Sweden means providing functions such as “I’m on vacation and will not return until 31 August 2005”)
• Computer-Telephony Integration (CTI), including Desktop call management, integration with various databases, etc.
• PSTN availability and reliability (thus the increasing use of Power over Ethernet for ethernet attached IP phones - so the wall outlet does not have to provide power for the phone to work)
• Roaming - both personal and device mobility
• Phone number portability
• E911 service {How do you handle geographic location of the station?}
Calling and Called Features

- **Calling** feature - activated when placing a call
  - e.g., Call Blocking and Call Return
- **Called** feature - activated when this entity would be the target of a call
  - Call Screening and Call Forward
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! (This is a good area for projects in 2G5565 Mobile Presence: Architectures, Protocols, and Applications)

- SIP Messaging and Presence Leveraging Extensions (SIMPLE) Working Group was formed in March 2001

http://www.ietf.org/html.charters/simple-charter.html
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
Negatives

Although VoIP equipment costs less than PBXs:

- the technology is new and thus upgrades are frequent (this takes time and effort)
- PBXs generally last ~10 years and public exchanges ~30yrs; while VoIP equipment is mostly computer equipment with a ~3 year amortization
“I am preparing legislation to preserve the free regulatory framework that has allowed VoIP applications to reach mainstream consumers,” Sununu, Republican from New Hampshire, said in a statement. “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.”

http://www.internetweek.com/e-business/showArticle.jhtml?articleID=17300570
Regulations in Sweden

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

Conferences

Voice on the Net (VON) http://www.von.com/

Interoperability testing:

• SIP development community’s interoperability testing event is called SIPit http://www.sipit.net/\(^1\). Note: The SIPit event is **closed** to the public and press, and no information is released about which products **fail** to comply with the standard.
  • Why have it closed? So that the testing can be done without risk of public embarrassment.
  • Interoperability is one of the most important aspects of wide deployment using multiple vendors products[6].
  • Proper handing of server failover is considered by some to be the most critical interoperability issue at present[6].

---

Not with out problems

It is not necessary a smooth transition to VoIP. Numerous organizations have faced problems [15] and there remain vast areas where further work is needed.

Potential for Spam over Internet Telephony (SPIT), Denial of Service, …
References and Further Reading

- IP Telephony (iptel)
- SIP Forum [http://www.sipforum.org](http://www.sipforum.org)

A great set of references compiled by prof. Raj Jain is available at: [http://www.cis.ohio-state.edu/~jain/refs/ref_voip.htm](http://www.cis.ohio-state.edu/~jain/refs/ref_voip.htm)


Acknowledgements

I would like to thank the following people and organizations for their permission to use pictures, icons, …

- Ulf Strömgren <ustromgr@cisco.com> for sending the Cisco 7960 picture on 2002.10.30
- Henry Sinnreich and Alan Johnston, both of WorldCom, for the wonderful SIP tutorial which Henry sent on 2002.10.30
# Traditional Telecom vs. Datacom

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

Carry the speech frame inside an RTP packet

<table>
<thead>
<tr>
<th></th>
<th>IPv4/6</th>
<th>UDP</th>
<th>RTP</th>
<th>CODEC info</th>
</tr>
</thead>
<tbody>
<tr>
<td>Octets</td>
<td>20/40</td>
<td>8</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>~33 octets</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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


This should be compared to the durations relevant to speech phenomena:

- “10 µs: smallest difference detectable by auditory system (localization),
- 3 ms: shortest phoneme (plosive burst),
- 10 ms: glottal pulse period,
- 100 ms: average phoneme duration,
- 4 s: exhale period during speech.” (from slide titled ‘What is a “short” window of time?’[31])
RTP and H.323 for IP Telephony

<table>
<thead>
<tr>
<th>audio/video applications</th>
<th>signaling and control</th>
<th>data applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>video codec</td>
<td>RTCP</td>
<td>H.225 registration</td>
</tr>
<tr>
<td>audio codec</td>
<td>H.225 Signaling</td>
<td>H.245 Control</td>
</tr>
<tr>
<td>RTP</td>
<td></td>
<td>T.120</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>UDP</th>
<th>TCP</th>
<th>IP</th>
</tr>
</thead>
</table>

H.323 framework of a group protocols for IP telephony (from ITU)
H.225 Signaling used to establish a call
H.245 Control and feedback during the call
T.120 Exchange of data associated with a call
RTP Real-time data transfer
RTCP Real-time Control Protocol

We will not examine H.323 in much detail, but will examine RTP and RTCP.
RTP, RTCP, and RTSP

- Audio/video applications
- Video, audio, … CODECs
- RTP
- UDP
- TCP
- IP

- Signaling and control
- SDP
- SIP

- Streaming applications
- CODECs
- RTSP
- UDP
Real-Time Delivery

In a real-time application \(\Rightarrow\) data must be delivered with the same time relationship as it was created (but with some delay)

Two aspects of real-time delivery (for protocols):

<table>
<thead>
<tr>
<th>Order</th>
<th>data should be played in the same order as it was created</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>the receiver must know when to play the packets, in order to reproduce the same signal as was input</td>
</tr>
</tbody>
</table>

We keep these separate by using a sequence number for order and a time stamp for timing.

Consider an application which transmits audio by sending datagrams every 20ms, but does silence detection and avoids sending packets of only silence. Thus the receiver may see that the time stamp advances by more than the usual 20ms, but the sequence number will be the expected next sequence number. Therefore we can tell the difference between missing packets and silence.
Packet delay

A stream of sampled audio packets are transmitted from the source \( (s_n) \), received at the destination \( (r_n) \), and played \( (p_n) \), thus each packet experiences a delay before playout \( (d_n) \).

If a packet arrives too late \( (r_3 \text{ arrives after we should have started to play at } p_3) \), then there is a problem (for some or all of the third packet’s audio).
Dealing with Delay jitter

Unless packets are lost, if we wait **long enough** they will come, but then the total delay may exceed the threshold required for interactive speech! (~180ms)
Delay and delay variance (jitter)

The end-to-end delay (from mouth to ear - for audio), includes:

- 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 perceived voice quality, see R. G. Cole and J. H. Rosenbluth, “Voice over IP Performance Monitoring”[18].

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

\[
I_d = 0.024d + 0.11(d - 177.3)H(d - 177.3)
\]

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

\[
H(x) = \begin{cases} 
0 & \text{if } x < 0 \\
1 & \text{when } x \geq 0
\end{cases}
\]

- 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 [23][24]
- 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 spurs 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.

---

1. Average silence duration (~596 ms) combine with the average talk-spurt duration (227ms) ⇒ a long-term speech activity factor of 27.6% [32].
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:

![Diagram showing the relationship between different clock and message synchronization processes](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.

However, 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 noise or play noise like in the last samples {as humans get uncomfortable with complete silence, they think the connection is broken!} [19]
  - if you are using highly encoded audio even a BER of $10^{-5}$ will produce very noticeable errors

- video
  - show the same (complete) video frame again
  - you can drop every 100th frame (for a BER of $10^{-2}$), but the user will not notice! [20]

There may also be compression applied to RTP see [36].
Patterns of Loss

With simple FEC you could lose *every other* packet and still not be missing content, but if pairs of packets are lost then you lose content.

To understand temporal patterns of speech, various models have been developed, see for example [33].
Loss concealment

There are various techniques for loss concealment (i.e., hiding losses), such as those used in the Robust Audio Tool (RAT):

  


- UCL’s Robust Audio Tool (RAT) page:
  
  http://www-mice.cs.ucl.ac.uk/multimedia/software/rat/

See also [222] and [223].
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** than “toll grade” telephony or
- **worse** than “toll grade” telephony (but perhaps still acceptable)

This is unlike the **fixed** quality of traditional phone systems.
RTP: Real-Time Transport Protocol

- First defined by RFC 1889, now defined by RFC 3550 [27]
- Designed to carry a variety of real-time data: audio and video.
- Provides two key facilities:
  - Sequence number for order of delivery (initial value chosen randomly)
  - Timestamp (of first sample) - used for control of playback

Provides no mechanisms to ensure timely delivery.

```
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
+-------------+----------------+---------------------+
<table>
<thead>
<tr>
<th>VER</th>
<th>P</th>
<th>X</th>
<th>CC</th>
<th>M</th>
<th>PTYPE</th>
<th>Sequence number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Timestamp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Synchronization source identifier</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contributing source ID …</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

- VER - version number (currently 2)
- P - whether zero padding follows the payload
- X - whether extension or not
- M - marker for beginning of each frame (or talk spurt if doing silence detection)
- PTYPE - Type of payload - first defined as Profiles in RFC 1890 now defined in RFC 3551

We will address the other fields later.
# Payload types

Payload types (PT) for standard audio and video encodings (Adapted from Tables 4 and 5 of RFC3551 [28])

<table>
<thead>
<tr>
<th>PT</th>
<th>encoding name</th>
<th>audio clock rate (Hz)</th>
<th>audio channels (audio)</th>
<th>PT</th>
<th>encoding name</th>
<th>video clock rate (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>PCM U</td>
<td>8,000</td>
<td>1</td>
<td>24</td>
<td>unassigned</td>
<td>V</td>
</tr>
<tr>
<td>1</td>
<td>reserved</td>
<td>8,000</td>
<td>1</td>
<td>25</td>
<td>CelB</td>
<td>V 90,000</td>
</tr>
<tr>
<td>2</td>
<td>reserved</td>
<td>8,000</td>
<td>1</td>
<td>26</td>
<td>JPEG</td>
<td>V 90,000</td>
</tr>
<tr>
<td>3</td>
<td>GSM</td>
<td>8,000</td>
<td>1</td>
<td>27</td>
<td>unassigned</td>
<td>V</td>
</tr>
<tr>
<td>4</td>
<td>G723</td>
<td>8,000</td>
<td>1</td>
<td>28</td>
<td>nv</td>
<td>V 90,000</td>
</tr>
<tr>
<td>5</td>
<td>DVI4</td>
<td>8,000</td>
<td>1</td>
<td>29</td>
<td>unassigned</td>
<td>V</td>
</tr>
<tr>
<td>6</td>
<td>DVI4</td>
<td>16,000</td>
<td>1</td>
<td>30</td>
<td>unassigned</td>
<td>V</td>
</tr>
<tr>
<td>7</td>
<td>LPC</td>
<td>8,000</td>
<td>1</td>
<td>31</td>
<td>H.261</td>
<td>V 90,000</td>
</tr>
<tr>
<td>8</td>
<td>PCMA</td>
<td>8,000</td>
<td>1</td>
<td>32</td>
<td>MPV</td>
<td>V 90,000</td>
</tr>
<tr>
<td>9</td>
<td>G722</td>
<td>8,000</td>
<td>1</td>
<td>33</td>
<td>MP2T</td>
<td>AV 90,000</td>
</tr>
<tr>
<td>10</td>
<td>L16</td>
<td>44,10</td>
<td>2</td>
<td>34</td>
<td>unassigned</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>L16 A</td>
<td>44,10</td>
<td>1</td>
<td>72</td>
<td>reserved</td>
<td>N/A</td>
</tr>
<tr>
<td>12</td>
<td>QCELP</td>
<td>8,000</td>
<td>1</td>
<td>77</td>
<td>unassigned</td>
<td></td>
</tr>
</tbody>
</table>

N/A (N/A = Not Applicable)
Payload types (PT) for standard audio and video encodings (Adapted from Tables 4 and 5 of RFC3551 [28])

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>CN</td>
<td>A</td>
<td>8,000</td>
<td>1</td>
<td>96..1</td>
<td>dynamic</td>
</tr>
<tr>
<td>14</td>
<td>MPA</td>
<td>A</td>
<td>90,00</td>
<td>see RFC 0</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>G728</td>
<td>A</td>
<td>8,000</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>DVI4</td>
<td>A</td>
<td>11,02</td>
<td>1</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>DVI4</td>
<td>A</td>
<td>22,05</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>G729</td>
<td>A</td>
<td>8,000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>reserved</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20..</td>
<td>unassigned</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Dynamic assignment of mapping between a payload type and an encoding is defined by SDP or H.323/H.245 mechanisms; these start with 96 - but can use lower numbers, if more than 32 encodings are needed - see RFC3551 [28].

RFC3551 says no new static assignments are to be made.
# Audio Encodings

Properties of Audio Encodings (adapted from Table 1 of RFC1990 and updated by RFC3551 [28])

<table>
<thead>
<tr>
<th>encoding</th>
<th>encoding</th>
<th>sample/frame</th>
<th>bits/sample</th>
<th>ms/frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>DVI4</td>
<td>Interactive Multimedia Association’s DVI ADPCM Wave Type</td>
<td>sample</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>G722</td>
<td>ITU’s G.722: 7 kHz audio-coding within 64 kbit/s</td>
<td>sample</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>G723</td>
<td>ITU’s G.723: Dual-rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s</td>
<td>frame</td>
<td>N/A</td>
<td>30.0</td>
</tr>
<tr>
<td>G726</td>
<td>ITU’s G.726</td>
<td>frame</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>G728</td>
<td>ITU’s G.728: 16 kbit/s using low-delay CELP</td>
<td>frame</td>
<td>N/A</td>
<td>2.5</td>
</tr>
<tr>
<td>G729</td>
<td>ITU’s G.729: 8 kbit/s using conjugate structure-algebraic code excited linear prediction (CS-ACELP)</td>
<td>frame</td>
<td>N/A</td>
<td>10.0</td>
</tr>
<tr>
<td>GSM</td>
<td>GSM 06.10: RPE/LTP (residual pulse excitation/long term prediction) coding at a rate of 13 kb/s</td>
<td>frame</td>
<td>N/A</td>
<td>20.0</td>
</tr>
<tr>
<td>L8</td>
<td>8 bit linear</td>
<td>sample</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>L16</td>
<td>16 bit linear</td>
<td>sample</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>LPC</td>
<td>Linear Predictive Coding</td>
<td>frame</td>
<td>N/A</td>
<td>20.0</td>
</tr>
<tr>
<td>MPA</td>
<td>MPEG-I or MPEG-II audio encapsulated as elementary streams, from ISO standards ISO/IEC 11172-3 &amp; 13818-3</td>
<td>frame</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>PCMA</td>
<td>G.711 A-law</td>
<td>sample</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>PCMU</td>
<td>G.711 mu–law</td>
<td>sample</td>
<td>8</td>
<td></td>
</tr>
</tbody>
</table>
Properties of Audio Encodings (adapted from Table 1 of RFC1990 and updated by RFC3551 [28])

<table>
<thead>
<tr>
<th>encoding</th>
<th>encoding</th>
<th>sample/frame</th>
<th>bits/sample</th>
<th>ms/frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>QCLEP</td>
<td></td>
<td>frame</td>
<td>variable</td>
<td>20.0</td>
</tr>
<tr>
<td>VDVI</td>
<td>variable-rate version of DVI4</td>
<td>sample</td>
<td>variable</td>
<td></td>
</tr>
</tbody>
</table>

See also internet Low Bitrate Codec (iLBC) [34].

There is also a lot of work in wideband CODECs, such as Extended Adaptive Multi-Rate Wideband (AMR-WB+) Audio Codec [37], [38]
**Timestamps**

The *initial* timestamp is to be chosen *randomly* (just as the initial sequence number is selected randomly):

- to avoid replays
- to increase security (this assumes that the intruder does not have access to all the packets flowing to the destination)

The timestamp *granularity* (i.e., the units) are determined by the payload type often based on the sampling rate
Stream translation and mixing

<table>
<thead>
<tr>
<th>mixing</th>
<th>combining several RTP streams to produce a single stream</th>
</tr>
</thead>
<tbody>
<tr>
<td>translation</td>
<td>converting from one encoding to another (also know as <strong>transcoding</strong>)</td>
</tr>
</tbody>
</table>

Each source has a unique 32 bit **Synchronization Source Identifier**.

When several sources are mixed the new stream gets its own unique **Synchronization Source Identifier** and the IDs of the contributing sources are included as **Contributing Source IDs**, the number of which is indicated in the 4-bit **CC** field of the header.
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

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sender Report</td>
<td>SR</td>
<td>Time information for each synchronization source and a count of data octets sent</td>
</tr>
<tr>
<td>Receiver Report</td>
<td>RR</td>
<td>Report of packet loss and jitter, information for timing and round-trip estimation</td>
</tr>
<tr>
<td>Source Description</td>
<td>SDES</td>
<td>Description of who owns the source</td>
</tr>
<tr>
<td>Goodbye</td>
<td>BYE</td>
<td>Receiver leaving the session</td>
</tr>
<tr>
<td>Application</td>
<td>APP</td>
<td>Application-specific report</td>
</tr>
</tbody>
</table>

Name Type Meaning
Sender Report SR 200 Time information for each synchronization source and a count of data octets sent
Receiver Report RR 201 Report of packet loss and jitter, information for timing and round-trip estimation
Source Description SDES 202 Description of who owns the source
Goodbye BYE 203 Receiver leaving the session
Application APP 204 Application-specific report

1. RTCP uses compound packets with multiple RTCP messages in a single packet.
Compound Reports

If and only if (IFF) the compound packet is to be encrypted: it is prefixed by a random 32-bit quantity selected for each compound packet transmitted.

The first RTCP packet in the compound packet must always be a report packet (either RR or SR). Followed by up to 30 more report packets (as RC is only 5 bits).

This is followed by an SDES packet containing a CNAME item (other information such as NAME, EMAIL, PHONE, LOC {geographic location}, TOOL, NOTE, and PRIV {private extension to SDES} are optional).

BYE should be the last packet sent with a given SSRC/CSRC.
Proposed RTCP Reporting Extensions

See RFC 3611 RTP Control Protocol Extended Reports (RTCP XR)[35]

VoIP Metrics Report Block - provides metrics for monitoring VoIP calls.

<table>
<thead>
<tr>
<th></th>
<th>8</th>
<th>16</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>BT=64</td>
<td>reserved</td>
<td>length=7</td>
<td></td>
</tr>
<tr>
<td>loss rate</td>
<td>discard rate</td>
<td>burst duration</td>
<td></td>
</tr>
<tr>
<td>burst density</td>
<td>gap duration</td>
<td>gap density</td>
<td></td>
</tr>
<tr>
<td>round trip delay</td>
<td>end system delay</td>
<td></td>
<td></td>
</tr>
<tr>
<td>signal power</td>
<td>doubletalk</td>
<td>noise level</td>
<td>Gmin</td>
</tr>
<tr>
<td>R factor</td>
<td>ext. R factor</td>
<td>MOS-LQ</td>
<td>MOS-CQ</td>
</tr>
<tr>
<td>RX Config</td>
<td>JB Nominal</td>
<td>JB Maximum</td>
<td>JB Abs Max</td>
</tr>
</tbody>
</table>
block type (BT) the constant $64 = 0x40$
reserved 8 bits - MUST be set to zero unless otherwise defined.
length length of this report block in 32-bit words minus one, including the header; constant 6.
loss rate fraction of RTP data packets from the source lost since the beginning of reception, as a fixed point number with the binary point at the left edge of the field\(^a\)
discard rate fraction of RTP data packets from the source that have been discarded since the beginning of reception, due to late or early arrival, under-run or overflow at the receiving jitter buffer, in binary fixed point
burst duration mean duration of the burst\(^b\) intervals, in milliseconds
burst density fraction of RTP data packets within burst intervals since the beginning of reception that were either lost or discarded, in binary fixed point
gap duration mean duration, expressed in milliseconds, of the gap intervals that have occurred
gap density fraction of RTP data packets within inter-burst gaps since the beginning of reception that were either lost or discarded, in binary fixed point
round trip delay most recently calculated round trip time between RTP interfaces, in milliseconds
end system delay most recently estimated end system delay, in milliseconds
signal level voice signal relative level is defined as the ratio of the signal level to overflow signal level, expressed in decibels as a signed integer in two’s complement form
doubletalk level defined as the proportion of voice frame intervals during which speech energy was present in both sending and receiving directions
noise level defined as the ratio of the silent period back ground noise level to overflow signal power, expressed in decibels as a signed integer in two’s complement form
R factor  a voice quality metric describing the segment of the call that is carried over this RTP session, expressed as an integer in the range 0 to 100, with a value of 94 corresponding to "toll quality" and values of 50 or less regarded as unusable; consistent with ITU-T G.107 and ETSI TS 101 329-5

ext. R factor a voice quality metric describing the segment of the call that is carried over an external network segment, for example a cellular network

MOS-LQ estimated mean opinion score for listening quality (MOS-LQ) is a voice quality metric on a scale from 1 to 5, in which 5 represents excellent and 1 represents unacceptable

MOS-CQ estimated mean opinion score for conversational quality (MOS-CQ) defined as including the effects of delay and other effects that would affect conversational quality

Gmin gap threshold, the value used for this report block to determine if a gap exists

RX Config PLC - packet loss concealment: Standard (11)/enhanced(10)/disabled (01)/ unspecified(00); JBA - Jitter Buffer Adaptive: Adaptive (11) / non-adaptive (10) / reserved (01)/ unknown (00). Jitter Buffer is adaptive then its size is being dynamically adjusted to deal with varying levels of jitter;JB Rate - Jitter Buffer Rate (0-15)

Jitter Buffer nominal size in frames (8 bit)

Jitter Buffer Maximum size in frames (8 bit)

Jitter Buffer Absolute size in frames

Maximun

a. Here after simply referred to as a binary fixed point number.
b. A burst is defined as a longest sequence of packets bounded by lost or discarded packets with the constraint that within a burst the number of successive packets that were received, and not discarded due to delay variation, is less than some value Gmin.
RTP translators/mixers

Translator  changes transport (e.g., IPv4 to IPv6) or changes media coding (i.e., transcoding)
Mixer  combines multiple streams to form a combined stream

Connect two or more transport-level “clouds”, each cloud is defined by a common network and transport protocol (e.g., IP/UDP), multicast address or pair of unicast addresses, and transport level destination port.

To avoid creating a loop the following rules must be observed:

• “Each of the clouds connected by translators and mixers participating in one RTP session either must be distinct from all the others in at least one of these parameters (protocol, address, port), or must be isolated at the network level from the others.

• A derivative of the first rule is that there must not be multiple translators or mixers connected in parallel unless by some arrangement they partition the set of sources to be forwarded.”

From §7.1 General Description of RFC 1889
Synchronizing Multiple Streams

One of the interesting things which RTP supports is synchronization of multiple streams (e.g., audio with a video stream)

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>8</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>VER</td>
<td>P</td>
<td>RC</td>
<td>PTYP</td>
<td>Length</td>
<td></td>
</tr>
<tr>
<td>Sender's Synchronization Source ID</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NTP Time Stamp (most significant 32 bits)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NTP Time Stamp (least significant 32 bits)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RTP Timestamp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sender's Packet Count</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sender's Octet Count</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First Synchronization Source</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction Lost</td>
<td>Total Packets Lost</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extended Highest Sequence Received</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inter-arrival Jitter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Last Sender Report</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delay Since Last Sender Report</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Unfortunately since the time stamps of each stream started at a random number we need some other method to synchronize them!
- Thus use Network Time Protocol (NTP) based time stamps \( \Rightarrow \) an absolute timestamp
- Since we now include the stream timestamps we can correlate these to absolute time (and hence from one stream to another)
RTP Transport and Many-to-many Transmission

RTP uses a connectionless transport (usually UDP):

- Retransmission is undesirable (generally it would be too late)
- Since RTP handles flow control and sequencing we don’t need this from the transport protocol
- RTP is packet oriented
- Enables us to easily use multicast (when there are many endpoints that want the same source stream)
  - multicast identified a group
  - these multicast groups can be dynamic
## Sessions, Streams, Protocol Port, and Demultiplexing

<table>
<thead>
<tr>
<th>Session</th>
<th>All traffic that is sent to a given IP address, port</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stream</td>
<td>a sequence of RTP packets that are from a single synchronization source</td>
</tr>
</tbody>
</table>

### Demultiplexing:

<table>
<thead>
<tr>
<th>session demultiplexing</th>
<th>occurs at the transport layer based on the port number</th>
</tr>
</thead>
<tbody>
<tr>
<td>stream demultiplexing</td>
<td>occurs once the packet is passed to the RTP software, based on the synchronization source identifier - then the sequence number and timestamp are used to order the packet at a suitable time for playback</td>
</tr>
</tbody>
</table>
Further details of RTP and RTCP


Note that an important aspect of RTCP is the rate of sending reports.

• “It is RECOMMENDED that the fraction of the session bandwidth added for RTCP be fixed at 5%.” [27]

• “It is also RECOMMENDED that 1/4 of the RTCP bandwidth be dedicated to participants that are sending data so that in sessions with a large number of receivers but a small number of senders, newly joining participants will more quickly receive the CNAME for the sending sites.” [27]

• Senders can be divided into two groups “… the RECOMMENDED default values for these two parameters would be 1.25% [active senders] and 3.75% [in-active senders] …”. [28]
  • ⇒ in-active sender ≡ receivers should generate at a rate of ~3.75% of the session traffic
  • of course: receivers on receive only links can not generate any reports
Real Time Streaming Protocol (RTSP)


- remote media playback control (think in terms of controlling a remote VCR/DVD/CD player)
- similar to HTTP/1.1, but
  - introduces new methods
  - RTSP servers maintain state
  - data carried out of band (i.e., in RTP packets)
- can use UDP or TCP
- Uses Web security methods (see [40])

See also the internet draft:
- Real Time Streaming Protocol (RTSP)
RTSP session description

<title>Twister</title>
<session>
  <group language=en lipsync>
    <switch>
      <track type=audio e="PCMU/8000/1" src="rtsp://audio.example.com/twister/audio.en/lofi">
      <track type=audio e="DVI4/16000/2" pt="90 DVI4/8000/1" src="rtsp://audio.example.com/twister/audio.en/hifi">
    </switch>
    <track type="video/jpeg" src="rtsp://video.example.com/twister/video">
  </group>
</session>

From figure 6: “Sample RTSP session description” of Henning Schulzrinne, “A comprehensive multimedia control architecture for the Internet”

References and Further Reading


http://www.ietf.org/html.charters/mmusic-charter.html

Also important are the measures of delay, delay jitter, throughput, packet loss, etc. IP Performance Metrics (ippm) is attempting to specify how to measure and exchange information about measurements of these quantities.


[23] Miroslaw Narbutt and Liam Murphy, “Adaptive Playout Buffering for Audio/Video Transmission over the Internet”, Department of Computer Science, University College Dublin, Dublin, Ireland,


RTP and RTCP


http://www.ietf.org/rfc/rfc2833.txt


http://www.cnel.ufl.edu/~markskow/papers/windows.ppt

http://starlet.deltatel.ru/ccitt/1988/ascii/5_1_06.txt {A later version of the standard is ITU-T Recommendation P.80, Methods for Subjective Determination of Transmission Quality, March, 1993}
[33] M. Y. Kim and W. B. Kleijn, “Rate-Distortion comparisons between FEC and MDC based on Gilbert channel model”, in Proc. IEEE Int. Conf. on Networks (ICON), 2003, pp. 495 - 500, Sydney, Australia.


RTSP

http://www.ietf.org/rfc/rfc2326.txt

Session Initiation Protocol (SIP)

Developed by the IETF Multiparty Multimedia Session Control (MMUSIC) working group and since September 1999 in the IETF SIP working group.

SIP is a text-based protocol, similar to HTTP and SMTP, for initiating interactive communication sessions between users. Sessions include: voice, video, chat, interactive games, and virtual reality.

SIP working groups charter “… to maintain the basic model and architecture defined by SIP. In particular:

1. Services and features are provided end-to-end whenever possible.\(^1\)
2. Extensions and new features must be generally applicable, and not applicable only to a specific set of session types.
3. Simplicity is key.
4. Reuse of existing IP protocols and architectures, and integrating with other IP applications, is crucial.

---

\(^1\) The use of end-to-end control is the exact opposite of the centralized control in traditional telecommunication networks.
SIP WG’s deliverables

• SIP specification
• callcontrol: call control specifications, which enables multiparty services, e.g., transfer and bridged sessions
• callerpref: caller preferences extensions, enables intelligent call routing services
• mib: a MIB for SIP nodes
• precon: extensions needed to assure satisfaction of external preconditions, e.g., QoS establishment
• state: extensions needed to manage state within signaling, aka SIP "cookies"
• priv: extensions for security and privacy
• security: security and privacy mechanisms and requirements
• provrel: extensions needed for reliability of provisional messages
• servfeat: extensions needed for negotiation of server features
• sesstimer: Session Timer extension
• **events**: Events extensions (Subscribe/Notify)
• **natfriend**: Extensions for making SIP a NAT-friendly protocol
Related working groups

- Session Initiation Proposal Investigation (sipping) WG - to analyze the requirements for application of SIP to several different tasks
- SIMPLE WG - using SIP for messaging and presence
  - SIP for Presence & IMPP to define payload
- IP telephony (IPTEL) WG
  - Call Processing Language (CPL), Telephony Routing over IP (TRIP)
- SPIRITS - SIP as ‘transport’ mechanism
- Distributed Call Signaling (DCS) Group of the PacketCable Consortium (http://www.packetcable.com/) for distributed telephony services
- 3rd Generation Partnership Project (3GPP), 3rd Generation Partnership Project 2 (3GPP2), -- 3rd generation wireless network efforts
- Peer-to-Peer Session Initiation Protocol (p2psip) WG
- Speech Services Control (speechsc) WG

Historic
- There was a PSTN and Internet Internetworking working group (PINT)
  - origin of SUBSCRIBE/NOTIFY
Session Initiation Protocol (SIP)

- Defined in RFC 3261 [46], updated by RFC 3853[62] & RFC 4320[63]
- 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 control¹, …

SIP is an alternative to H.323 proposed by IETF. 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 (SIP alone: 269 pages)
- RFC3261 was longest RFC ever (based on byte count; 663,043 bytes)
- 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:

Alice | Invite | Bob
--- | --- | ---
Alice | OK,200 | Bob
Bob | ACK | Alice
Bob | media session
Bob | Bye

Maguire
maguire@it.kth.se
2007.03.27
Practical Voice Over IP (VoIP): SIP and related protocols
SIP actors

Figure 3: SIP Actors (presence entity names in blue)
SIP Methods and Status Codes

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

At least 8 additional methods have been defined see SIP Method Extensions in other RFCs on page 166.

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

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

Two URI schemes - similar to the form of an e-mail addresses: user@domain

- **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

Three 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
- **Globally Routable UA URIs (GRUU)** (identifies a instance of a user at a given UA, for the duration of the registration of the UA to which it is bound)[59]
Issues to be considered

• Address Resolution
• Session Setup
• Media Negotiation
• Session Modification
• Session Termination
• Session Cancellation
• Mid-call Signaling
• Call Control
• QoS Call setup
Address Resolution

The 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.

- providing support for mobility and portability
- Can utilize:
  - DNS SRV lookup
  - ENUM
  - Location Server lookup

We will look at this in detail (see **DNS and ENUM** on page 225), but for now will assume a simple DNS lookup based on the URI.
Simple version of Alice invites Bob to a SIP session:

We begin by examining the details of session setup. For lots of example of basic call flows see [56].
SIP Invite

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)

SIP is a text-based protocol and uses ISO 10646 character set in UTF-8 encoding (RFC 2279). The message body uses MIME and can use S/MIME for security.

The generic form of a message is:

```
generic-message = start-line
message-header*
CRLF
[ message-body ]
```

---

1. Example adapted from draft-ietf-sip-rfc2543bis-06.ps
Bob’s response to Alice’s INVITE

SIP/2.0 200 OK
Via: SIP/2.0/UDP pc33.atlanta.com:5060;branch=z9hG4bK776asdhds
Via: SIP/2.0/UDP bigbox3.site3.atlanta.com:5060;branch=z9hG4bK77ef4c2312983.1
Via: SIP/2.0/UDP pc33.atlanta.com:5060;branch=z9hG4bKnashds8
To: Bob <sip:bob@biloxi.com>;tag=a6c85cf
From: Alice <sip:alice@atlanta.com>;tag=1928301774
Call-ID: a84b4c76e66710
CSeq: 314159 INVITE
Contact: <sip:bob@192.0.2.8>
Content-Type: application/sdp
Content-Length: 131

{Bob’s SDP not shown}

---
1. Example adapted from draft-ietf-sip-rfc2543bis-06.ps
ACK

ACK 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 ACK
Content-Length: 0

A successful set-up sequence was: INVITE/200/ACK

A set-up failure would be a sequence such as: INVITE/4xx¹/ACK

NB: INVITE is the only method in SIP that involves a 3-way handshake with ACK

The further setup of the call can proceed directly between Alice and Bob, based on the information (especially that in SDP) which they have exchanged.

Now we will examine the details of these initial SIP messages!

¹ or 5xx or 6xx
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
INVITE sip:bob@biloxi.com SIP/2.0
Via: SIP/2.0/UDP proxy.stockholm.se:5060;branch=82.1
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)

- **Via** headers show the path the request has taken in the SIP network
  - A Via header is inserted by the User Agent which initiated the request (this will be last in the list of Via headers)
  - Via headers are inserted above this by proxies in the path (i.e., this details the path taken by the request)
- **Via** headers are used to route responses back the same way the request came
  - this allows stateful proxies to see both the requests and responses
  - each such proxy adds the protocol, hostname/IP address, and port number
- The “branch” parameter is used to detect loops
Dialog (Call leg) Information

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)

• Dialog¹ (formerly “call leg”) information is in headers:
  • To tag, From tag, and Call-ID All requests and responses in this call will use this same Dialog information.
  • “To” specifies the logical recipient of the message, “From” the logical sender
    – the string “Bob” is called a “display name”

• Call-ID is unique identifier
  • Call-ID number is arbitrary, but it uniquely identifies this call (i.e., session), hence all future references to this session refer to this Call-ID
  • usually composed of pseudo-random string @ hostname or IP Address

---
¹. A Dialog formally begins upon receipt of a response containing a tag. It is called an “Early dialog” when the response was a 18x provisional response.
SIP CSeq

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)

• Command Sequence (CSeq) Number
  • Initialized at start of call (1 in this example)
  • Incremented for each subsequent request
  • Used to distinguish a retransmission from a new request
• Followed by the request type (i.e., SIP method)
SIP Contact

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)

- **Contact** header contains a SIP URL for direct communication between User Agents
  - If Proxies do not Record-Route\(^1\), they can be bypassed
- **Contact** header is also present in 200 OK response

---

1. Note that the Record-Route and Route headers approach of RFC 2543 was found not to work.
SIP Content Type and Length

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)

• **Content-Type** indicates the type of message body attachment (others could be text/plain, application/cpl+xml, etc.)
  • Here “application/sdp” indicates that it is SDP

• **Content-Length** indicates length of the message body in octets (bytes)
  • 0 indicates that there is no message body.
SIP Max-Forwards

INVITE sip:bob@biloxi.com SIP/2.0
Via: SIP/2.0/UDP pc33.atlanta.com:5060;branch=z9hG4bK776asdhds
Max-Forwards: 30
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)

- **Max-Forwards** is decremented by each proxy that forwards the request.
- When count goes to zero, request is **discarded** and **483 Too Many Hops** response is sent.
- Used for stateless loop detection.
Other header fields

- **Content-Encoding**: 
- **Allow**: 
- **Expires**: 
- **In-Reply-To**: 
- **Priority**: indicated priority of displaying a message to a user
  - Normal 
  - Urgent 
  - Non-Urgent 
  - Emergency 
- **Require**: contains a list of options which the server is expected to support in order to process a request 
- **Retry after**: number of seconds after which a requestor should try again 
- **Supported**: enumerates all the extensions supported the sender (NB: this differs from a “Require” which requires that a destination supports the given extension)
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** - directs 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
User Agent A

DNS Server

Outbound Proxy Server

Location Server

Other servers

Proxy servers

User Agents

Figure 4: SIP Trapezoid

Figure 5: SIP Call Setup - when B has not registered

SIP Call Setup Attempt


Figure 6: SIP Call Setup Attempt - when B has not registered
Figure 7: SIP Call Setup Attempt - when B has not registered (continued)

SIP Presence

Figure 8: SIP Presence: A asks to be told when B registers

Figure 9: NOTIFY A that B has <Not Signed In>

SIP Registration Example

Purpose of registration

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

- Only *their* location server need know
  - The location server need not disclose this location to "just anyone", but can apply various polices to decide who can learn of it, i.e., their location server can decide **who** can ask for B’s location and **when** they can ask (perhaps even limiting it to **where** they can ask from).
  - This has significant privacy implications.

- This scales well - as B only has to update their location server, rather than having to inform all *possible* callers.
Registering 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=voice
- 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 268.
SIP Call Setup Attempt⁴

Figure 11: SIP Call Setup Attempt - when B has registered

SIP Session Termination using BYE

BYE causes the media session to be torn down.

Note: BYE like INVITE is an end-to-end method.
SIP Session Termination using CANCEL

CANCEL causes the session to be cancel. Note: If a reply is 481 Transaction Unknown, then the user agent may need to send a BYE since the CANCEL was received after the final response was sent (there was a race condition).

No media session was established.
CANCEL and OPTIONS

CANCEL

- In addition to canceling a pending session
- CANCEL can also be sent by a proxy or user agent
  - for example, when a parallel fork has been done, once you have a successful match, then you can cancel the others

OPTIONS

- Used to query a server or user agent for its capabilities
- sometimes used for very simple presence information
Unsuccessful final responses are hop-by-hop

Unsuccessful final responses (3xx, 4xx, 5xx, 6xx) are always acknowledged on a hop-by-hop basis.

Only 200 OK is end-to-end.
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 (response 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, NATs, and Firewalls on page 280.
SIP Method Extensions in other RFCs

See “Guidelines for Authors of Extensions to the Session Initiation Protocol (SIP)”[51]

- **INFO** - Call signaling information during a call

- **PRACK** - Reliable ACK
  - RFC 3262: Reliability of Provisional Responses in Session Initiation Protocol (SIP), June 2002

- **SUBSCRIBE/NOTIFY**

- **REFER**
  - RFC 3515: The Session Initiation Protocol (SIP) Refer Method, April 2003 [52]

- **MESSAGE**
  - RFC 3428: Session Initiation Protocol Extension for Instant Messaging, December 2002 [54]

- **UPDATE** - Early media and preconditions
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.

⇒ No Profiling is needed
  • unlike for example, Bluetooth!
SIP Presence - Signed In

Figure 12: NOTIFY A that B has \(<\text{Signed In}\>\)

If user B’s agent does not wish to provide user A’s agent with a notification it sends a 603 Decline response.

Media Session
SIP Instant Messaging Example


Figure 13: A sends a message to B
SIP Instant Messaging Example (continued)¹

Figure 14: B sends a message to A

A simple Instant Message (IM) as SIP:

MESSAGE im:UserB@there.com SIP/2.0
Via: SIP/2.0/UDP 4.3.2.1
To: User B <im:UserB@there.com>
From: User A <im:UserA@here.com>
Call-ID: a5-32-43-12@4.3.2.1
CSeq: 1 MESSAGE
Content-type: text/plain
Content-Length: 16

Hi, How are you?

The response will be a 200 OK from B.

Note: the example uses IM URIs instead of SIP URIs.

A MESSAGE request can be sent at anytime (even without a session).

For further information about the work of the IETF working group on Instant Messaging and Presence Protocol (impp) see

Midcall signalling

Midcall signalling used when the session parameters don’t change, to exchange information between two user agents via the body of an INFO message. If the session parameters did change then you would use a re-INVITE.

Note in the above figure the ISUP messages: IAM (Initial address message), INM (Answer message), and USR (user-to-user message).
Call Control

SIP is peer-to-peer -- thus a proxy can’t issue a BYE, only end devices (UAs) can.

To methods for third party call control:

- A proxy passes an invite on, but **stays in the signaling path**
- Use **REFER** to initial third party control (the third party is no longer in the signaling path).

Useful for:

- click-to-call
- Automatic Call Distribution (ACD)
- web call center
- …
Example of using **REFER**

Third party call control, by User A to set up a session between Users B and C.

Note: the use by A of an **INVITE** with a **Refer-to** header and the user by B of an **INVITE** with a **Referred-By** header.
QoS and Call Setup

The path which SIP takes may be different that the media path, thus new extensions were added to enable more handshaking:

- **Early Media** - by allowing SDP to be included in the 183 Session Progress response (allows establishment of QoS requirements before call is answered) - may also enable one-way RTP {hence the name “early media”}, formally: “media during early dialog”

- **Reliable Provisional Responses** extension allows detection of a lost 183 Session Progress response based on using Provisional Response Acknowledgement (PRACK)

- **UAs can use** preCOnditions MET (COMET) method to indicate that the QoS requirements can be met and that the user can be alerted by ringing the phone.

SDP in the INVITE contains an attribute-value pair: "a=qos:mandatory".

For further details see: RFC3312 [47] and RFC3262 [48]; more about SDP in the next lecture module.
1. INVITE
   3. 100 Trying

2. INVITE
   5. 100 Trying

4. INVITE
   6. 100 Trying

User Agent A  Proxy Server  Proxy Server  User Agent B

5. 100 Trying

6. 100 Trying

7. 183 Session Progress

8. 183 Session Progress

9. 183 Session Progress

10. PRACK

11. PRACK

12. PRACK

13. 200 OK

14. 200 OK

15. 200 OK

QoS Setup

16. COMET

21. 200 OK

22. 180 Ringing

17. COMET

20. 200 OK

23. 180 Ringing

18. COMET

19. 200 OK

24. 180 Ringing

27. 200 OK

25. 200 OK

28. ACK

26. 200 OK

29. ACK

30. ACK

RTP media session with QoS
# SIP Message retransmission

<table>
<thead>
<tr>
<th>Timer</th>
<th>default</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>500ms</td>
<td>Set when SIP request is first sent</td>
</tr>
<tr>
<td>T2</td>
<td>4 sec.</td>
<td>Longer timeout when a provisional response has been received</td>
</tr>
</tbody>
</table>

If a request is lost, then timeout T1 will generate a retransmission of the request.

If a request is received and a provisional response is received, then sender switches to timeout T2 (to wait for the final response).

**INVITE** is different:
- receiving a provisional response stops all re-transmissions of the **INVITE**;
- however, the sender of the provisional response starts a T1 timer when it sends its final response and if it does not get an ACK in time it retransmits the final response.

If you want/need acknowledgement of provisional responses use **PRACK**. {For some problems with timeouts for non-INVITE transactions see [60][61].}
RFC 3261 - Routing Changes

- Introduced “loose routing” vs. RFC 3543’s “strict routing”
  - Examples:
    - Pre-loaded (initial INVITE) Route header can be used instead of the default outbound proxy (DOP)
    - Pre-loaded Route header can be used to invoke “home proxy” services (when you are roaming)
    - Additional proxies can be added as needed (for example, adding routing during a call)
- All elements must insert branch parameter as a transaction ID in Via header fields
- Contact header required in all requests that establish a dialog
- From and To tags are now mandatory
- Recommend users of Fully Qualified Domain Name (FQDN) instead of IP addresses
- Via loop detection no longer required of proxies
  - Use of Max-Forwards is now mandatory
- Via hiding is deprecated (i.e., should no longer be used)
  - because it turned out not to be secure or useful
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
- Announcement handling
  - UAS or proxy need not make a decision about playing an early media announcement
    - Error response contains new `Error-Info` header field which contains the URI of the announcement
  - UAC makes a decision based on the user’s interface
Compression of SIP

As textual protocols, some might thing that SIP and SDP are too verbose, hence RFC 3486 [58] describes how SIP and SDP can be compressed. RFC 3485 [57] describes a static dictionary which can be used with Signaling Compression (SigComp) to achieve even higher efficiency.
Intelligent Network service using SIP

ITU has defined a set of service features (think of them as primitives which can be use to construct more complex services). These are divided into two sets:

- Capability Set 1: Service Features
- Capability Set 2

J. Lennox, H. Schulzrinne, and T. F. La Porta, “Implementing Intelligent Network Service with the Session Initiation Procol” [65] addresses Capability Set 1:

Abbreviated Dialing (ABD)  
Attendant (ATT)  
Authentication (AUTC)  
Authorization code (AUTZ)  
Automatic callback (ACB)  
Call distribution (CD)  
Call forwarding (CF)  
Call forwarding on busy/don’t answer (CFC)  
Call gapping (GAP)  
Call hold with announcement (CHA)  
Call limiter (LIM)  
Call logging (LOG)  
Call queueing (QUE)  
Call transfer (TRA)  
Call waiting (CW)  
Closed usergroup (CUG)  
Consultation calling (COC)  
Customer profile management (CPM)  
Customer recorded announcement (CRA)  
Customized ringing (CRG)  
Destinating user prompter (DUP)  
Follow-me diversion (FMD)  
Mass calling (MAS)  
Meet-me conference (MMC)  
Multi-way calling (MWC)  
Off-net calling (ONC)  
One number (ONE)  
Origin dependent routing (ODR)  
Originating call screening (OCS)  
Originating user prompter (OUP)  
Personal numbering (PN)  
Premium charging (PRMC)  
Private numbering plan (PNP)  
Reverse charging (REVC)  
Split charging (SPLC)  
Terminating call screening (TCS)  
Time dependent routing (TDR)
## Capability Set 1: Services

<table>
<thead>
<tr>
<th>Abbreviated dialling (ABD)</th>
<th>Originating call screening (OCS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Account card calling (ACC)</td>
<td>Premium rate (PRM)</td>
</tr>
<tr>
<td>Automatic alternative billing (AAB)</td>
<td>Security screening (SEC)</td>
</tr>
<tr>
<td>Call distribution (CD)</td>
<td>Selective call forwarding on busy/don’t answer (SCF)</td>
</tr>
<tr>
<td>Call forwarding (CF)</td>
<td>Selective call forwarding</td>
</tr>
<tr>
<td>Call rerouting distribution (CRD)</td>
<td>Call forwarding on busy</td>
</tr>
<tr>
<td>Completion of calls to busy subscriber (CCBS)</td>
<td>Call forwarding on don’t answer (no reply)</td>
</tr>
<tr>
<td>Conference calling (CON)</td>
<td>Split charging (SPL)</td>
</tr>
<tr>
<td>Credit card calling (CCC)</td>
<td>Televoting (VOT)</td>
</tr>
<tr>
<td>Destination call routing (DCR)</td>
<td>Terminating call screening (TCS)</td>
</tr>
<tr>
<td>Follow-me diversion (FMD)</td>
<td>Universal access number (UAN)</td>
</tr>
<tr>
<td>Freephone (FPH)</td>
<td>Universal personal telecommunications (UPT)</td>
</tr>
<tr>
<td>Malicious call identification (MCI)</td>
<td>User-defined routing (UDR)</td>
</tr>
<tr>
<td>Mass calling (MAS)</td>
<td>Virtual private network (VPN)</td>
</tr>
</tbody>
</table>
Capability Set 2

Wireless services
Inter-network services
Multimedia
Call pick-up
Calling name delivery
## Features

List of features adopted from http://www.miercom.com/survey - augmented with my own notes with respect to SIP supporting this feature:

<table>
<thead>
<tr>
<th>SIP</th>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>✔️</td>
<td>911/E-911 support</td>
<td>Emergency services</td>
</tr>
<tr>
<td>✔️</td>
<td>Audible message waiting</td>
<td>An audible indicator when there is a new message</td>
</tr>
<tr>
<td>✔️</td>
<td>Automated attendant</td>
<td>Answers and routes calls automatically based on caller responses; e.g., via Interactive Voice Response (IVR) or DTMF prompts</td>
</tr>
<tr>
<td>✔️</td>
<td>Automatic alternate routing</td>
<td>Routes calls automatically based on user-defined routing parameters, priorities, and failover/availability decisions.</td>
</tr>
<tr>
<td>✔️</td>
<td>Automatic call back</td>
<td>Calls an extension back automatically when a busy signal or no answer is encountered. Also known as Camp on.</td>
</tr>
<tr>
<td>✔️</td>
<td>Bridged call appearance</td>
<td>Allows the same phone number to appear and be answered on multiple phone sets.</td>
</tr>
<tr>
<td>✔️</td>
<td>Call blocking</td>
<td>Selectively blocks calls from user-defined origins</td>
</tr>
<tr>
<td>Feature</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>-------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Call conference</td>
<td>An audio path for multiple parties on a single call, established via user keystrokes and no outside intervention.</td>
<td></td>
</tr>
<tr>
<td>Call drop</td>
<td>Terminates a call without hanging up the receiver.</td>
<td></td>
</tr>
<tr>
<td>Call forward all</td>
<td>Redirects all calls to another station or location.</td>
<td></td>
</tr>
<tr>
<td>Call forward on busy</td>
<td>Redirects all calls to another station or location <em>when the user's is busy</em>.</td>
<td></td>
</tr>
<tr>
<td>Call forward on no answer</td>
<td>Redirects all calls to another station or location <em>after a specified number of rings</em>.</td>
<td></td>
</tr>
<tr>
<td>Call hold</td>
<td>Places an incoming call on hold or retrieves a call placed on hold.</td>
<td></td>
</tr>
<tr>
<td>Call pick-up</td>
<td>Allows a user to place a call on hold, then resume it from another phone in the system.</td>
<td></td>
</tr>
<tr>
<td>Call return</td>
<td>Calls back the last incoming number.</td>
<td></td>
</tr>
<tr>
<td>Call transfer</td>
<td>Redirects an answered call to another user.</td>
<td></td>
</tr>
<tr>
<td>Call waiting</td>
<td>An audible indicator heard when there is another call pending.</td>
<td></td>
</tr>
<tr>
<td>Caller ID</td>
<td>Displays the name and/or number of the calling party.</td>
<td></td>
</tr>
<tr>
<td>Feature</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Call Detail Recording (CDR)</td>
<td>Records call data on a specific extension or group of extensions.</td>
<td></td>
</tr>
<tr>
<td>Class of service</td>
<td>Restricts access to features based upon users' privilege level(s).</td>
<td></td>
</tr>
<tr>
<td>Direct inward system access</td>
<td>Dial-in system station appearance.</td>
<td></td>
</tr>
<tr>
<td>Direct transfer to voicemail</td>
<td>Automatically redirects all calls to users' voicemail at the push of a button.</td>
<td></td>
</tr>
<tr>
<td>Directory lookup</td>
<td>Allows users to look up an extension from the corporate LDAP directory.</td>
<td></td>
</tr>
<tr>
<td>Distinctive ringing</td>
<td>Uses a different ringtone for different call characteristics, for example, internal vs external calls.</td>
<td></td>
</tr>
<tr>
<td>Do not disturb</td>
<td>Makes the phone appear to be out of service.</td>
<td></td>
</tr>
<tr>
<td>Follow me</td>
<td>Rings multiple, disparate phones simultaneously when one extension is dialed.</td>
<td></td>
</tr>
<tr>
<td>Free seating/Hoteling</td>
<td>Allows a user to move from one location to another, accessing all calls, features, button mappings, etc.</td>
<td></td>
</tr>
<tr>
<td>Hot line</td>
<td>Private line automatic ring-down connection between two phones.</td>
<td></td>
</tr>
<tr>
<td>Hunt groups</td>
<td>Diverts calls to busy extensions to any extension in a pre-defined group.</td>
<td></td>
</tr>
<tr>
<td>Feature</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>✤ Intercom - phone-to-phone</td>
<td>An internal intercom that initiates calls within a predefined group or department.</td>
<td></td>
</tr>
<tr>
<td>✤ Intercom - phone-to-multi-phone</td>
<td>An internal intercom that initiates voice paging through the speakers of multiple phone systems.</td>
<td></td>
</tr>
<tr>
<td>✤ Intrude</td>
<td>Allows specific users to intrude on calls already in progress.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>See R. Mahy and D. Petrie, &quot;The Session Initiation Protocol (SIP) 'Join' Header&quot; [69] - a new header for use with SIP multi-party applications and call control; to logically join an existing SIP dialog, for: ‘Barge-In’, ‘Message Screening’, ‘and Call CenterMonitoring’</td>
<td></td>
</tr>
<tr>
<td>✤ Last number redial</td>
<td>Redials the last outgoing call.</td>
<td></td>
</tr>
<tr>
<td>✔ Least-cost routing</td>
<td>Routes outbound calls to the least expensive alternative, based on user-defined prioritization.</td>
<td></td>
</tr>
<tr>
<td>Leave word calling</td>
<td>Allows internal users to leave short, pre-programmed messages for other internal users.</td>
<td></td>
</tr>
<tr>
<td>Malicious call trace</td>
<td>Allows users to initiate a call trace.</td>
<td></td>
</tr>
<tr>
<td>Message waiting indicator</td>
<td>Visibly indicates when new voicemail arrives, often via a blinking light.</td>
<td></td>
</tr>
<tr>
<td>✤ Missed call indicator</td>
<td>Lists missed calls.</td>
<td></td>
</tr>
<tr>
<td>Feature</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Multiple call appearance</td>
<td>Allows a single phone to have multiple, repeated instances of a single phone extension.</td>
<td></td>
</tr>
<tr>
<td>✫ Multiple ring styles</td>
<td>Changes the ringtone based on user preference.</td>
<td></td>
</tr>
<tr>
<td>✫ Music on hold</td>
<td>Plays music for the caller when placed on hold.</td>
<td></td>
</tr>
<tr>
<td>✫ Mute</td>
<td>Disables the microphone. (This is really just a feature of the client.)</td>
<td></td>
</tr>
<tr>
<td>✫ Night service</td>
<td>Changes call coverage based on the time of day, for example, plays a common recording for all calls at night.</td>
<td></td>
</tr>
<tr>
<td>✫ One-button send all calls</td>
<td>Automatically redirects all calls to someone else who provides coverage with a single button.</td>
<td></td>
</tr>
<tr>
<td>✫ One-button speed dial</td>
<td>Dials a predefined number with a single button.</td>
<td></td>
</tr>
<tr>
<td>✔ Personal call routing</td>
<td>Defines routing parameters</td>
<td></td>
</tr>
<tr>
<td>Priority ringing</td>
<td>Uses a different ringtone for specified numbers.</td>
<td></td>
</tr>
<tr>
<td>Recorded announcements</td>
<td>Provides predefined announcements to certain calls, for example, “Your call cannot be completed as dialed”.</td>
<td></td>
</tr>
<tr>
<td>✔ System speed dialing</td>
<td>Dials frequently-called numbers using an abbreviated access code.</td>
<td></td>
</tr>
<tr>
<td>✫ User directory</td>
<td>Allows any system endpoint to browser a database of names, extensions, etc.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Feature</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td></td>
<td>Volume control</td>
<td>Changes the volume individually for the speaker, handset, and ringer.</td>
</tr>
<tr>
<td>✷</td>
<td>Whisper page</td>
<td>Allows someone else (such as an assistant) to bridge into a call, allowing only the local party to hear.</td>
</tr>
<tr>
<td>✔</td>
<td>Supported by SIP</td>
<td></td>
</tr>
<tr>
<td>✷</td>
<td>Supported by SIP + additional methods</td>
<td></td>
</tr>
</tbody>
</table>
SIP development, evolution, ...

In traditional IETF fashion is based on **running code**

- So in your projects you should make sure that what you propose is really feasible by implementing it!
  - should have at least **2 interoperable implementations** for each feature
- See the SIP mailing list (**listen until** you have sufficient knowledge to contribute)
- See the SIP Working Group for what is being worked on by others
- See “Guidelines for Authors of Extensions to the Session Initiation Protocol (SIP)” [51]
Gateways

- **Gateway Location Protocol** (GLP) - a protocol used between Location Server (LSs) {similar to BGP}
- **Signaling Gateway** - to convert from the signaling used in one network to that of the other
- **Media Gateway** - to convert the media format from that used in one network to that of the other
Significance

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

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


<table>
<thead>
<tr>
<th>3GPP</th>
<th>IETF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network does not trust the user</td>
<td>User only partially trusts the network</td>
</tr>
<tr>
<td>layer 1 and layer 2 specific</td>
<td>generic</td>
</tr>
<tr>
<td>walled garden</td>
<td>open access</td>
</tr>
</tbody>
</table>

Not surprisingly 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)
References and Further Reading


http://www.ietf.org/html.charters/mmusic-charter.html

[43] Session Initiation Protocol (sip) Working Group
http://www.ietf.org/html.charters/sip-charter.html

Also important are the measures of delay, delay jitter, throughput, packet loss, etc. IP Performance Metrics (ippm) is attempting to specify how to measure and exchange information about measurements of these quantities.

SIP

[44] Henning Schulzrinne’s Session Initiation Protocol (SIP) web page
http://www.cs.columbia.edu/sip/
http://www.ietf.org/rfc/rfc2543.txt


http://www.ietf.org/rfc/rfc3853.txt

[63] R. Sparks, Actions Addressing Identified Issues with the Session Initiation Protocol’s (SIP) Non-INVITE Transaction, IETF, RFC 4320, January 2006
http://www.ietf.org/rfc/rfc4320.txt


ITU Services CS-1 and CS-2


[66] Study Group 11 of the International Telecommunications Union Telecommunications Standards Sector (ITU-T), ITU-T recommendation
Q.1211: Introduction to Intelligent Network Capability Set 1, Annex B


   http://www.ietf.cnri.reston.va.us/rfc/rfc2871.txt

Module 4: Session Announcement Protocol (SAP)

Lecture notes of G. Q. Maguire Jr.
### Session Announcement Protocol (SAP)

Defined in RFC 2974[70]

Primarily for **multicast** session announcement. It provides the session setup information to **prospective** participants.

Each SAP announcer periodically multicasts an announcement:

- to a well known multicast address on port 9875
  - IPv4 global scope sessions use multicast addresses in the range 224.2.128.0 - 224.2.255.255 - their SAP announcements are sent to 224.2.127.254
  - IPv4 administrative scope sessions using administratively scoped IP multicast are defined in [x], the multicast address to be used for announcements is the highest multicast address in the relevant administrative scope zone, e.g., if the scope range is 239.16.32.0 - 239.16.33.255, then SAP announcements use 239.16.33.255
  - IPv6 sessions are announced on the address FF0X:0:0:0:0:1234:5678 where X is the 4-bit scope value, e.g., an announcement for a link-local session assigned the address FF02:0:0:0:1234:5678, is advertised on SAP address FF02:0:0:0:0:2:7FFE
- has same scope as the session it is announcing (the use of TTL scoping for multicast is discouraged)
- IP time-to-live of 255
References and Further Reading

SAP

http://www.ietf.org/rfc/rfc2974.txt
Module 5: Session Description Protocol (SDP)

Lecture notes of G. Q. Maguire Jr.
Session Description Protocol (SDP)

- audio/video applications
- signaling and control
- streaming applications

- video, audio, …
- CODECs
- RTP
- RTCP
- SDP
- SIP
- CODECs
- RTSP
- UDP
- TCP
- IP
Session Description Protocol (SDP)


RFC 4566: SDP: Session Description Protocol

- describes media session
- a text-based protocol
- carried in MIME as a message body in SIP messages
- uses RTP/AVP Profiles for common media types [76]

Note: It is more a session description **format** than a **protocol**.

- RFC 3108: Conventions for the use of the Session Description Protocol for ATM Bearer Connections
- RFC 3264: An Offer/Answer Model with the Session Description Protocol (SDP)
- RFC 3266: Support for IPv6 in Session Description Protocol
- RFC 3388: Grouping of Media Lines in the Session Description Protocol
- RFC 3407: Session Description Protocol Simple Capability Declaration
• RFC 3485: The Session Initiation Protocol and Session Description Protocol Static Dictionary for Signaling Compression (SigComp)
• RFC 3556: Session Description Protocol Bandwidth Modifiers for RTP Control Protocol (RTCP) Bandwidth [82]
• RFC 3605: Real Time Control Protocol (RTCP) attribute in Session Description Protocol
• RFC 3890: A Transport Independent Bandwidth Modifier for the Session Description Protocol
• RFC 4092: Usage of the Session Description Protocol: Alternative Network Address Types (ANAT) Semantics in the Session Initiation Protocol (SIP)
• RFC 4145: TCP-Based Media Transport in the Session Description Protocol
• RFC 4317: Session Description Protocol Offer/Answer Examples [83]
• RFC 4567: Key Management Extensions for Session Description Protocol and Real Time Streaming Protocol (RTSP).
• RFC 4568: Session Description Protocol Security Descriptions for Media Streams [86]
• RFC 4570: Session Description Protocol Source Filters [81]
• RFC 4572: Connection-Oriented Media Transport over the Transport Layer Security (TLS) Protocol in the Session Description Protocol
• RFC 4574: The Session Description Protocol Label Attribute [84]
• RFC 4579: Session Initiation Protocol Call Control - Conferencing for User Agents
• RFC 4583: Session Description Protocol Format for Binary Floor Control Protocol (BFCP) Streams
• RFC 4796: The Session Description Protocol Content Attribute [87]

From the list above you can see that SDP is a changing area.
SDP Message Details

v=0
o=Tesla 289084526 28904526 IN IP4 lab.high-voltage.org
s=--
c=IN IP4 100.101.102.103
t=0 0
m=audio 49170 RTP/AVP 0
a=rtpmap:0 PCMU/8000

- **Version number** (ignored by SIP)
- **Origin** (not used by SIP)
- **Subject** (ignored by SIP)
- **Connection Data**
  - connection: network (IN == Internet), Address type (IPv4), and Address
- **Time** (ignored by SIP): start stop
- **Media** (type, port, RTP/AVP Profile)
- **Attribute** (profile, CODEC, sampling rate)
Session description

v= protocol version
o= owner/creator and session identifier
s= session name
[i= session information] { [xx] ⇒ xx is optional}
[u= URI of description]
[e= email address]
[p= phone number]
[c= connection information- not required if included in all media]
[b= bandwidth information]

<Time description>+ { <xx>+ ⇒ one or more times}
[z= time zone adjustments]
[k= encryption key]
[a= zero or more session attribute lines]* { <xx>* ⇒ zero or more times}

<Media descriptions>*

Time description

t= time the session is active
[r= zero or more repeat times]*

Media description

m= media name and transport address
[i= media title]
[c= connection information-optional if included at session-level]
[b= bandwidth information]
[k= encryption key]
[a= zero or more media attribute lines]*
SDP Offer/Response Example

\[\text{v=0} \]
\[\text{o=} \]
\[\text{c=IN IP4 130.237.212.6} \]
\[\text{t=} \]
\[\text{m=video 4004 RTP/AVP 14 26} \]
\[\text{a=rtpmap:14 MPA/90000} \]
\[\text{a=rtpmap:26 JBEG/90000} \]
\[\text{m=audio 4006 RTP/AVP 0 4} \]
\[\text{a=rtpmap:0 PCMU/8000} \]
\[\text{a=rtpmap:4 GSM/8000} \]

Version of SDP (0)
Origin - not use by SIP
Connection INternet, IPv4, address=130.237.212.6
Time - not use by SIP
Media Video, port=4004, type=RTP/AVP profile, profiles: 14 and 26
Attribute for profile 14, codec=MPA, sampling rate=90000
Attribute for profile 26, codec=JBEG, sampling rate=90000
Media Audio, port=4006, type=RTP/AVP profile, profiles: 0 and 4
Attribute for profile 0, codec=PCMU (PCM \mu\text{law}), sampling rate=8000
Attribute for profile 4, codec=GSM, sampling rate=8000

If the RTCP port is not the next port number, then an rtcp-attribute can be specified in the form [77] (this might be useful in conjunction with a NAT):

"a=rtcp:" port [nettype addrtype connection-address] <CRLF>
SDP Response Example

v=0

Version of SDP (0)

o=

Origin - not use by SIP

c=IN IP4 130.237.21.87

Connection INternet, IPv4, address=130.237.21.87

t=

Time - not use by SIP

m=video 0 RTP/AVP 14

Media Video, port=0, type=RTP/AVP profile, profiles: 14
Receiver declines the video, indicated by port = 0

m=audio 6002 RTP/AVP 4

Media Audio, port=6002, type=RTP/AVP profile, profiles: 4
Receiver declines the PCM coded audio and selects the GSM coded audio

a=rtpmap:4 GSM/8000

Attribute for profile 4, codec=GSM, sampling rate=8000
Session Modification

① Alice invites Bob to a session with the parameters in sdp1.

② Bob’s modified this in his response sdp2.

They communicate.

③ Bob proposes a change in the session (sdp2’), Alice does not accept this change.

④ Bob tries with a new proposal (sdp2”).

⑤ Alice accepts with the session description sdp1’.

They communicate with the new spec.
Session modification (continued)

• The re-INVITE could have been done by either party - it uses the same To, From, and Call-ID as the original INVITE.

• Note that the re-INVITEs do not cause a 180 Ringing or other provisional messages, since communication between Alice and Bob is already underway.

• Note that the first media session continues despite the SIP signalling, until a new agreement has been reached - at which time the new media session replaces the former session.

• The re-INVITE can propose changes of any of the media characteristics, including adding or dropping a particular media stream.
  • this adding or dropping may be because the user has moved from one wireless cell to another, from one network to another, from one interface to another, from one device to another, …
Start and Stop Times

Enable the user to join a broadcast sessions during the broadcast.
Grouping of Media Lines in the Session Description Protocol (SDP)[74]

Defines two SDP attributes:

- "group" and
- "mid" - media stream identification

Allows grouping several media ("m") lines together. This is to support:

- Lip Synchronization (LS) and
- Flow Identification (FID) - a single flow (with several media streams) that are encoded in different formats (and may be received on different ports and host interfaces)
  - Changing between codecs (for example based on current error rate of a wireless channel)

Note FID does not cover the following (but SDP can -- see [74]):

- Parallel encoding using different codecs
- Layered coding
Lip Synchronization

Example adapted from section 6.1 of [74].

A session description of a conference that is being multicast. First and the second media streams MUST be synchronized.

```
v=0
o=Laura 289083124 289083124 IN IP4 one.example.com
t=0 0
c=IN IP4 224.2.17.12/127
a=group:LS 1 2
m=audio 30000 RTP/AVP 0
i=voice of the speaker who speaks in English
a=mid:1
m=video 30002 RTP/AVP 31
i=video component
a=mid:2
m=audio 30004 RTP/AVP 0
i=This media stream contains the Spanish translation
a=mid:3
```
Next generation of SDP (SDPng)

- Designed to address SDP’s ‘flaws’:
  - Limited expressiveness
    - For individual media and combinations of media
    - Often only very basic media descriptions available -- desire for more complex media
  - No real negotiation functionality - as SDP today is a “take it or leave it” proposal
  - Limited extensibility (not nearly as easy to extend as SIP)
  - No semantics for media sessions! Sessions are only implicit.
- SDPng should avoid "second system syndrome"
  - Hence it should be simple, easy to parse, extensible, and have limited scope
  - Session Description and Capability Negotiation

Drafts:
- Flemming Andreasen, “SDP Capability Negotiation: Requirements and Review of Existing Work”, Internet-draft, 6-Mar-07,
  \texttt{http://draft-ietf-mmusic-sdp-capability-negotiation-reqts-01.txt}
- Flemming Andreasen, “SDP Capability Negotiation”, Internet-draft, 6-Mar-07,
  \texttt{http://draft-ietf-mmusic-sdp-capability-negotiation-05.txt}
- Flemming Andreasen,, “SDP media capabilities Negotiation”, Internet-draft, 21-Feb-07,
  \texttt{http://draft-ietf-mmusic-sdp-media-capabilities-01.txt}
SDPng structure

Uses XML syntax - example adapted from Appendix C in [80]:

```xml
<?xml version="1.0" encoding="UTF-8"?>
<sdpng xmlns="http://www.iana.org/sdpng"
       xmlns:audio="http://www.iana.org/sdpng/audio"
       xmlns:video="http://www.iana.org/sdpng/video"
       xmlns:rtp="http://www.iana.org/sdpng/rtp"
       xsi:schemaLocation="http://www.iana.org/sdpng sdpng-base.xsd"
         http://www.iana.org/sdpng/audio sdpng-audio-pkg.xsd"
           http://www.iana.org/sdpng/video sdpng-video-pkg.xsd"
             http://www.iana.org/sdpng/rtp sdpng-rtp-pkg.xsd" >
  <cap><audio:codec name="avp:dvi4">
    <audio:encoding>DVI4</audio:encoding>
    <audio:channels>1</audio:channels>
    <audio:sampling>8000 11025 16000 22050</audio:sampling></audio:codec>
  <audio:codec name="avp:116">
    <audio:encoding>L16</audio:encoding>
    <audio:channels>1 2</audio:channels>
    <audio:sampling>44100</audio:sampling></audio:codec>
  <video:codec name="avp:celb">
    <video:encoding>CelB</video:encoding>
    <video:framerate>4 6 8 12 16 20 24 30</video:framerate></video:codec>
  <rtp:udp name="rtpudpip6">
    <rtp:network>IP6</rtp:network></rtp:udp></cap>
  <def><rtp:udp name="rtp-cfg1" ref="rtpudpip6">
    <rtp:ip-addr>::$1</rtp:ip-addr>
    <rtp:port>9546</rtp:port></rtp:udp></def>
  <cfg><component>
    <alt>
      <audio:codec ref="avp:116"/>
      <rtp:udp ref="rtp-cfg1"></rtp:pt></rtp:udp></alt></component></cfg></sdpng>
```

For details see appendices A.1 “SDPng Base DTD” and A.2 “SDPng XML-Schema Specification” in [80].
Why XML?

If, at this date and time, you want to not use XML, then you need an extremely strong case. XML is well understood, there are many support tools, and many more are in development. The W3C is producing a schema description language which is considered adequate for many business applications, many of which are way more complex than SDP.

The talks about ASN.1 are just that -- talks. The only possible advantage of ASN.1 is the size of the messages, but even that is debatable. On the other hand, the cost is very well known: you need specialized parsers and libraries, you cannot easily use text tools for debugging or monitoring purposes, and the syntax is hard to understand and a pain to extend. Most of the proponents of ASN.1 actually propose some variation of it, which is even worse, since it would require even more specific tools.

The main inconvenient of XML is that it can be bulky. I am not convinced that this is an actual problem: SDP is used for describing multimedia sessions, that normally last a few minutes and carry at a minimum several tens of kilobytes of media; the media stream dwarfs the signaling stream by orders of magnitude. If it is an actual problem, then we can indeed use compression. In fact, we can safely assume that other applications will be hurt before us, and that we will get generic XML compression tools sooner or later. All in all, that should not be a big problem.

Let’s not be silly. Just pick XML.

-- Christian Huitema

http://bmrc.berkeley.edu/mhonarc/openmash-developers/msg00315.html
QoS and SDP

James Polk, “Quality of Service (QoS) Mechanism Selection in the Session Description Protocol (SDP)”, Internet-draft, 5-Mar-07,

http://draft-polk-mmusic-qos-mechanism-identification-03.txt
SDP

[71] SDP: Session Description Protocol (RFC 2327)
    http://www.ietf.org/rfc/rfc2327.txt


[74] G. Camarillo, G. Eriksson, J. Holler, and H. Schulzrinne, Grouping of Media
    Lines in the Session Description Protocol (SDP), IETF RFC 3388, December

[75] F. Andreasen, Session Description Protocol (SDP) Simple Capability
    Declaration, IETF RFC 3407, October 2002.
    http://www.ietf.org/rfc/rfc3407.txt
[76] S. Casner and P. Hoschka, "‘MIME Type Registration of RTP Payload Formats", IETF RFC3555, July 2003

[77] C. Huitema, "Real Time Control Protocol (RTCP) attribute in Session Description Protocol (SDP)", IETF RFC 3605, October 2003
http://www.ietf.org/rfc/rfc3605.txt

http://www.ietf.org/rfc/rfc3524.txt


[80] Dirk Kutscher, Jörg Ott, and Carsten Bormann, “Session Description and Capability Negotiation”, IETF Internet-Draft, February 20, 2005, Expires:
August 21, 2005

http://www.ietf.org/internet-drafts/draft-ietf-mmusic-sdpng-08.txt


Module 6: DNS and ENUM
Telephony URL and Phone-Context

SIP URIs include Telephony URLs [101].

A Telephony URL looks like:

```
tel: +358-555-1234567  a telephone terminal
fax: +358-555-1234567  a fax machine
```

Digit separators of "-" or "." are ignored.

A Phone-Context sets the conditions under which the number can be used, e.g.

```
tel: 1-800-555-1234;phone-content:+1 972
```

- a phone number that can only valid within North America (+1) and within the 972 exchange
- the absence of the "+" in the telephone number indicates that this is a local number, rather than a global number -- but the interpretation of these local numbers is problematic (i.e., there is no assured geographic area nor can one depend on 7 digit numbers being local to a Class 5 exchange {the traditional case in North America}) \( \Rightarrow \) a proposal to deprecate the use of unqualified local digit strings see [94].
**SIP URL**

used in SIP messages to indicate: originator (From), current destination (Request-URI), final destination (To), and redirection address (Contact)

Examples:

<table>
<thead>
<tr>
<th>SIP URL</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>sip:<a href="mailto:firstname.lastname@company.com">firstname.lastname@company.com</a></td>
<td>simple example</td>
</tr>
<tr>
<td>sip:<a href="mailto:+1-212-555-1212@gateway.com">+1-212-555-1212@gateway.com</a>;user=phone</td>
<td>a call from the Internet to the PSTN E.164 phone number (user=phone is not necessary, but just a hint to parsers that it is a numeric phone number)</td>
</tr>
<tr>
<td>sips:<a href="mailto:+1-212-555-1212@gateway.com">+1-212-555-1212@gateway.com</a>;user=phone</td>
<td>a call from the Internet to the PSTN E.164 phone number - the SIP messages should be passed via TLS</td>
</tr>
<tr>
<td>sip:<a href="mailto:+1-212-555-1212@proxy.gateway.com">+1-212-555-1212@proxy.gateway.com</a>;user=phone</td>
<td>proxy server determines gateway and forwards the request</td>
</tr>
<tr>
<td>sip:<a href="mailto:firstname.lastname@registrar.com">firstname.lastname@registrar.com</a>;method=register</td>
<td>to register a user at a SIP registrar</td>
</tr>
</tbody>
</table>
IETF’s E.164 Number Mapping standard uses Domain Name Server (DNS) to map standard International Telecommunication Union (ITU-T) international public telecommunications numbering plan (E.164) telephone numbers to a list of Universal Resource Locators (URL). SIP uses these URL’s to initiate sessions.

For example, ENUM DNS [89] converts a telephone number in E.164 format, e.g. +46812345, and returns e.g., a Universal Resource Identifier (URI) SIP:olle.svenson@telia.se

Thus a SIP client makes a connection to the SIP gateway telia.se passing the local part olle.svenson.

ENUM can return a wide variety of URI types.

The draft “The E.164 to URI DDDS Application (ENUM)”[90] updates the ENUM specification to be compatible with the Dynamic Delegation Discovery System (DDDS) Application specification in RFC 3401.
For details of Dial Sequences and Global Switched Telephone Network (GSTN) see [92]. {Dial Sequences include pauses and other signalling in addition to the phone number}

Note that ENUM maintains the nation-state “ownership” of E.164 numbers.

Why bother? {see [96]}

- In order for PSTN/IDSN user to call VoIP users, there must be a way of translating an E.164 number to some way of reach the VoIP user.
  - Since the PSTN user only has a telephone dialing pad - this limits what they can enter (for example ‘+’ entered as ‘*’).
  - However, due to ITU-T Rec. E.105 [99] -- this means that VoIP become a part of the global public telephony service -- hence this translation has to follow at least some of the ITU rules
  - Which gateway should be used?

- For VoIP users to call a PSTN/ISDN user, caller needs to do an ENUM lookup and utilize a VoIP to PSTN/ISDN gateway
  - Which gateway?
  - Can the called user opt-in or opt-out of having calls from the Internet?

- VoIP caller to VoIP callee when the caller dials an E.164 number
  - Does it get routed to the PSTN and back? {I.e., going through two VoIP gateways!}
• Use of Geographic numbers for fixed VoIP terminals
  • easily enables 911 like services for their terminals too
• (Global | National) [non-geographic] personal numbers
  • A personal or global or national number - which can be your single number
• ...

One problem is that IP communications is not IP Telephony, it is VoIP + Chat + Instant Messaging + Video + … .
DNS

Scales well (due to caching)

ENUM typically uses a 3 layer hierarchy

- **Tier 0: ENUM Root Level**
  - Top level domain for telephone numbers is: `e164.arpa`
  - DNS look up to find the country for a specific E.164-Country Code (CC)
  - Manager: IAB; Registry: RIPE NCC; Registrar: ITU TSB .e164.arpa

- **Tier 1: ENUM CC Level - DNS look up to find the ENUM subscribers**
  - Manager: ITU Member State; Registry: choice of Manager; ENUM Registrar: national choice
  - Swedian example: 6.4.e164.arpa - registry: NIC-SE (as of 13 Dec. 2002)

- **Tier 2: ENUM E.164 Number Level**
  - DNS stores a list over different internet based addresses (URIs) in NAPTR records
  - Thus a look up $\Rightarrow$ a list over different internet based addresses associated with each E.164-number
  - Manager: E.164-subscriber; DNS Service Provider: choice of Manager

For details see RFC 2916[89] and RFC 2915[88].
> set querytype=NAPTR
> e164.arpa

Authoritative answers can be found from: e164.arpa
origin = ns.ripe.net
mail addr = e164-contacts.ripe.net
serial = 2002100901
refresh = 14400 (4H)
retry = 3600 (1H)
expire = 2419200 (4W)
minimum ttl = 14400 (4H)
To find the DNS names for a specific E.164 number

Procedure is:

• Write the E.164 number in its full form, including the countrycode IDDD. Example: +46-8-9761234
• Remove all non-digit characters with the exception of the leading ‘+’. Example: +4689761234
• Remove all characters with the exception of the digits. Example: 4689761234
• Put dots ("." ) between each digit. Example: 4.6.8.9.7.6.1.2.3.4
• Reverse the order of the digits. Example: 4.3.2.1.6.7.9.8.6.4
• Append the string ".e164.arpa" to the end. Example: 4.3.2.1.6.7.9.8.6.4.e164.arpa
• Ask the DNS it returns:
  • mailto: foo@kth.se
  • sip: foo@kth.se
  • ...
ENUM Services

- www.netnumber.com
- Neustar (www.neustar.com) has their ENUM activity at http://www.enum.org/
- EnumWorld (a venture between Network Solutions & Telcordia) www.enumworld.com

The ITU-T “List of ITU-T Recommendation E.164 Assigned Country Codes” as of 1 February 2004 can be found at: http://www.itu.int/itudoc/itu-t/ob-lists/icc/e164_763.html

The RIPE list of e-mail concerning the European assignment of ENUMs can be found at http://www.ripe.net/enum/request-archives/

For a summary of the status of ENUM deployment see [91].

For a summary of the IANA assignments for ENUM services see [109].
EUNM Timeline

Sept. 2000  IETF ENUM WG produced RFC2916

2001  Various Workshops (ITU-T, Europe, US, …) to spread the idea
Swedish PTS releases first ENUM report in April 2001

2002  ITU-T Interim Procedures (IAB, RIPE-NCC)
ETSI SPAN11 TS “ENUM Administration in Europe”

2003  ETSI SPAN11 TS “Minimum Requirements for Interoperability of European ENUM Trials”
IETF RFC2916bis
National and international ENUM Trials using:
  ♦ different scenarios and numbering resources
  ♦ different ENUM-enabled products
  ♦ Swedish PTS releases their ENUM report on 31 July 2003; trial to continue until May 2004, final report due 30 June 2004 (see [102] and [103])

2004  ENUM considered ready for production ⇒ commercial deployments

The IAB instructions regarding ENUM to the RIPE NCC (to whom they had delegated e164.arpa) can be found at:
http://www.ripe.net/enum/instructions.html

Sweden ENUM status is described at [95].

Interesting open questions (as described in [95]):

- Should the state have a permanent **operational** role (as opposed to simply an administrative role)
  - important that the subscriber with a given E.164 number also control the associated ENUM domain name {Who is responsible for maintaining this synchronization and validating changes?}

- Who finances the Tier 1 registry?
- Need for regulations? Self-regulation? …
- Privacy: need E.164 subscriber’s permission to list them in the DNS
- Are there business opportunities?
- Will ENUM be successful?
- …
Sweden’s ENUM Mapping

Approved by ITU TSB on Fri, 29 Nov 2002 12:03:02 +0100

Domain Object

domain: 6.4.e164.arpa
descr: Swedish ENUM Mapping
admin-c: PTSE46-RIPE
tech-c: SE194-RIPE
zone-c: SE194-RIPE
nservers:
  a.ns.6.4.e164.arpa
  b.ns.6.4.e164.arpa
  c.ns.6.4.e164.arpa
  d.ns.6.4.e164.arpa
...

Administrative Contact

role: ENUM Tier 1 Manager
address: National Post and Telecom Agency
address: Box 5398
address: SE-102 49 Stockholm
address: Sweden
phone: +46 8 678 55 69
fax-no: +46 8 678 55 05
e-mail: pts-enum-admin@localhost
trouble: enum-test-admin@localhost
nic-hdl: PTSE46-RIPE
...

Maguire
maguire@it.kth.se
VISIONng Association

Mission of VISIONng (http://www.visionng.org/index.htm): “to provide a framework for the deployment of worldwide inter-domain and multi-vendor IP Communications”

ITU-T has assigned part of the country code for Universal Personal Telecommunication (UPT) to VISIONng for deployment of a UPT Service: +878 10

As of May 2002 VISIONng received ITU-TSB permission and an ENUM Delegation from RIPE NCC; BearingPoint Inc. acting as Tier 1 Manager, Telekom Austria acting as Tier 2 DNS.

These E.164 numbers can be used for both: IP-IP and PSTN-IP.

See also [100].
SIP goes beyond ENUM

by offering additional features:

- User preferences
- Personal/Service/… mobility
- Easy and secure updating of information by the end-user

A given User Agent need not directly implement call routing, LDAP lookup, …, but can instead utilize a default SIP outgoing proxy (which in turn does the work).

Call Processing Language (CPL) can be used to support rapid changes in user preferences (see Call Processing Language (CPL) on page 254)

---

1. See SIP Mobility on page 245.
References and Further Reading

**DNS**


**ENUM**


http://www.ietf.org/rfc/rfc3761.txt


http://www.ripe.net/ripencc/about/presentations/ipv6-enum-paris-20031204/

[92] C. Allocchio, “Text String Notation for Dial Sequences and Global Switched

http://www.ietf.org/rfc/rfc3824.txt


http://www.ietf.org/rfc/rfc3762.txt

http://www.ietf.org/rfc/rfc3764.txt


[100] Richard Stastny, "Status of ENUM Trials", SG2 Plenary, Florianopolis, Brazil, October 2003
http://enum.nic.at/documents/AETP/Presentations/Austria/0025-2003-10_SG2_ENUM.ppt

[101] H. Schulzrinne, ‘The tel URI for Telephone Numbers”, IETF Internet-Draft,
March 20, 2004, Expires: September 18, 2004


   http://www.pts.se/Dokument/dokument.asp?ItemId=3232


   http://www.itu.int/osg/spu/presentations/2004/enum-country-experiences-ftra-uganda-rs.pdf&e=10053


Protocol”, Internet-Draft, December 1, 2004, Expires: June 1, 2005


Module 7: SIP Mobility

Lecture notes of G. Q. Maguire Jr.
SIP Mobility

- **Terminal mobility**\(^1\) ⇒ the terminal moves between subnets
  - Note: Mobile IP supports this at the network layer, while SIP supports this at the application layer (*without* requiring Mobile IP be underneath)

- **Personal Mobility** ⇒ the person moves between terminals

- **Service mobility** ⇒ the person has access to the same services *despite* their movement between terminals and/or networks
  - note: the service may be reduced in quality or capabilities subject to the current network’s capabilities -- but it is the same service
  - this implies that personalization of services must be distributed to the various terminals that the user wishes to use - see the dissertation of Roch Glitho [113]

- **Session mobility** ⇒ the same session is maintained despite the user changing from one device to another

---

\(^1\) Also known as network-level mobility.
Local Number Portability

In the PSTN this means a complex set of lookups for the number, since the number is no longer tied to an exchange.

In SIP the portability occurs because of the lookup of name@domain, which can be mapped to wherever the user wants this mapped to! (i.e., fully qualified domain names are unique, but are not tied to an underlying network address -- it is the name to address mapping which establishes this mapping and it is always dynamic).

For some considerations of tel URIs and number portability see [111] and [112].
References and Further Reading

SIP Mobility


Service Mobility

Module 8: SIP Service Creation

Lecture notes of G. Q. Maguire Jr.
SIP Service Creation

It is the increased opportunities for the exchange of signaling information via SIP which enables many new features and services.
Services implemented by x

Where x is:

- proxy server,
- called user agent,
- calling user agent, or
- Back-to-Back User Agent (B2BUA)

See examples of call-forward, no-answer service in chapter 6 of Sinnreich and Johnston[2].
Services implemented by Extensions

i.e., new methods and headers

See the activities of the IETF SIP, SIPPING, and SIMPLE working groups

Proxy servers - simply treat unknown methods as an OPTION request, unless there is a Proxy-Require header.

User agents return:

- **405 Method Not Allowed** if the method is recognized, but not supported
- **500 Bad Request** if it does not recognize the method
- **420 Bad Extension** if the UAS does not support the requested feature

- **All SIP extensions which use the Require or Supported header** must be documented as an RFC - to prevent interoperability problems
- **All standardized SIP extensions** must document how the extension interacts with elements that don’t understand this extension

---

1. See Other header fields on page 149
SIP Service Logic

- Call Processing Language (CPL)
- SIP Common Gateway Interface (CGI)
- SIP Java Servlets
RFC 2824: Call Processing Language (CPL) [114] and [115]

An XML-based scripting language for describing and controlling call services.

CPL is a very simple language without variables, loops, or the ability to run external programs! {Hence non-trusted end users can upload services to their SIP server} However, it has **primitives for making decisions** and **acting based on call properties** (e.g., time of day, caller, called party, …).

There is a Document Type Definition (DTD) “cpl.dtd” and strict parsing\(^1\) is done based on this DTD.

See also Chapter 13 of *Practical VoIP: Using VOCAL*[1], this includes an example of developing a feature in CPL

See also the dynamic loading of CPL in [118].

---

1. Thus **any** discrepancies between the script and the scheme are errors.
RFC 3050: Common Gateway Interface for SIP [116]

Similar to HTML CGI, a SIP CGI script resides on the server and passes message parameters via environment variables to a separate process. This process sends instructions back to the server through its standard output file descriptor.

Scripts can be written in Perl, Tcl, C, C++, Java, … Of course these scripts (being based on general purpose programming languages) do not have the limitations of CPL and hence only trusted users can be allowed to provide such scripts.

CGI scripts have access to both the request headers and the body and can therefore do general computations based on all this information.
SIP Java Servlets

Extends functionality of SIP client by passing messages to the SIP servelets.

Servlets are similar to the CGI concept, but instead of using a separate process, the messages are passed to a class that runs within a Java Virtual Machine (JVM) inside the server.

Servlets are portable between servers and operating systems, due to the portability of the Java code.


http://www.cs.columbia.edu/sip/drafts/draft-peterbauer-sip-servlet-ext-00.txt

SIP Servlets were defined in A. Kristensen and A. Byttner, “The SIP Servlet API”, IETF Draft, September 1999,

http://www.cs.columbia.edu/sip/drafts/draft-kristensen-sip-servlet-00.txt

• Unfortunately this draft expired and was not carried forward, but is referenced (and large parts included) in subsequent work. See also [117].
JAIN APIs

Providing a level of abstraction for service creation across circuit switched and packet networks, i.e., bridging IP and IN protocols. Goal is provisioning of telecom services by:

- **Service Portability**: - Write Once, Run Anywhere. (via Java portability)
- **Network Convergence**: (Integrated Networks) - Any Network
- **Service Provider Access** - By Anyone!
  - to allow services direct access to network resources and devices

SIP APIs - especially those within the JAIN™ initiative ([http://java.sun.com/products/jain/index.jsp](http://java.sun.com/products/jain/index.jsp)):

- **JAIN SIP** (JSR-000032) - a low level API that maps directly to RFC 2543 - [http://jcp.org/en/jsr/detail?id=32](http://jcp.org/en/jsr/detail?id=32)

- **JAIN SIP Lite** (JSR-000125)- a high-level API, to allow application developers to create applications that have SIP as their underlying protocol **without** needing extensive knowledge of SIP - [http://jcp.org/en/jsr/detail?id=125](http://jcp.org/en/jsr/detail?id=125)
• SDP API (JSR-000141) - to enable users to manipulate SDP messages


• SIMPLE related APIs
  • JAIN SIMPLE Instant Messaging (JSR-000165) - to exchange messages between SIMPLE clients
  • JAIN Instant Messaging (JSR-000187) - to control, manage and manipulate instant messages between clients through the use of presence servers
  • JAIN SIMPLE Presence (JSR-000164) - to manipulate presence information between a SIMPLE client (watcher) and a presence server (presence agent)
  • JAIN Presence and Availability Management (PAM) API (JSR-000123) -
  • JAIN Presence (JSR-000186) - to control, manage and manipulate Presence information between Presence clients and servers

• JAIN Service Provider APIs (SPA) - Java implementation of Parlay APIs
  • JAIN SPA Common API (JSR-000145) common across the JAIN SPA JSRs
  • JAIN SPA Integrity Management and Event Notification API (JSR-000119)

• Regarding Location
• JAIN User Location and Status (ULS) (JSR-000194) - to interrogate the location and status of a user’s mobile device  
  http://jcp.org/en/jsr/detail?id=194

• JAIN OAM API Specification v2.0 (JSR-000132) -  

• JAIN ENUM API Specification (JSR-000161) - API to query and provision E.164 telephone numbers and their service-specific Uniform Resource Identifiers (URI)  
  http://jcp.org/en/jsr/detail?id=161

• JAIN 3G MAP Specification (JSR-000137) - to enable mobile applications in the 3G domain to talk to each other  

The full list of JAIN related specification can be found at:  
http://www-x.antd.nist.gov/proj/iptel/

- NIST-SIP 1.2
- JAIN-SIP Proxy
- JAIN-SIP Instant Messaging Client
- JsPhone - a JAIN-SIP Video Phone
- NIST-SIP traces viewer
- JAIN-SIP gateway
- JAIN-SIP Third Party Call Controller
Parlay

Parlay Group formed (1998) to specify and promote open APIs that “intimately link IT applications with the capabilities of the communications world”.

Goal: to allow applications to access the functionality of the telecoms network in a secure way.

Parlay APIs:

• Service interfaces - provide access to network capabilities and information

• Framework interfaces provide the underlying supporting necessary for the service interfaces to be secure and manageable.

The APIs are defined in Universal Modeling Language (UML).

For further info see: [http://www.parlay.org/](http://www.parlay.org/) and [122].
B. Campbell and R. Sparks, “Control of Service Context using SIP Request-URI”, IETF RFC 3087, April 2001 [123] - proposes a mechanism to communicate context information\(^1\) to an application (via the use of a distinctive Request-URI).

Using different URIs to provide both state information and the information about lead to this state transition (for example, you were forwarded to the voicemail system because the user did not answer vs. being forwarded to the voicemail system because the user is busy with another call).

\(^1\) Call state information, such as the calling party, called party, reason for forward, etc.
Reason Header

Since it is (often) useful to know why a Session Initiation Protocol (SIP) request was issued, the Reason header was introduced. It encapsulates a final status code in a provisional response.

This functionality was needed to resolve the "Heterogeneous Error Response Forking Problem" (HERFP).

For details see [124].
Voice eXtensible Markup Language (VoiceXML ™)

VoiceXML designed for creating audio dialogs (i.e., audio in and out) that feature: synthesized speech, digitized audio, recognition of spoken and DTMF key input, recording of spoken input, telephony, and mixed-initiative conversations.

Goal: To bring the advantages of web-based development and content delivery to interactive voice response applications.

For details see: [http://www.w3.org/TR/voicexml][125]


VoiceXML is designed to go beyond Interactive Voice Response (IVR) systems.
References and Further Reading

SIP Service Creation


JAIN


[121] JAIN SIP 1.0 API specification

Parley


SIP Request URI


Reason Header

the Session Initiation Protocol (SIP)”, IETF RFC 3326, December 2002

VoiceXML

http://www.w3.org/TR/2000/NOTE-voicexml-20000505
Module 9: User Preferences

Lecture notes of G. Q. Maguire Jr.
User Preferences

• Caller preference
  • allows caller to specify how a call should be handled
  • to specify media types: audio, video, whiteboard, …
  • to specify languages (of the callee -- consider for example a help desk call where you want to get help in your choice of language)
  • do you want to reach the callee at home or only at work?, via a landline or on their mobile phone? …
  • examples: should the call be forked or recurse, do you want to use a proxy or redirect, do you want to CANCEL 200 messages or not,

• Called party preference
  • accepting or rejecting calls: based on time of day, day of week, location of called party, from unlisted numbers, …

Caller/callee different
  • Callee is passive, caller is active
    – Thus callee’s preferences must be defined ahead of time (for example by CPL)
    – However, caller’s preferences can be in request
  • Services (usually) run on callee server
  • A given caller might contact any of a large number of number of servers (each of which will have to decide how to process this caller’s request)

Conclusion: Include caller preferences in request
## Contact parameters

Values are either pre-set or indicated when a user REGISTER’s:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>example(s)</th>
<th>Explanation of example(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>class</td>
<td>personal</td>
<td>class=personal</td>
<td>Call should go the &quot;home&quot; not the office.</td>
</tr>
<tr>
<td></td>
<td>business</td>
<td></td>
<td></td>
</tr>
<tr>
<td>duplex</td>
<td>full</td>
<td>duplex=full</td>
<td>should be a full duplex call</td>
</tr>
<tr>
<td></td>
<td>half</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>send-only</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>receive-only</td>
<td></td>
<td></td>
</tr>
<tr>
<td>feature</td>
<td>voicemail</td>
<td>feature=voicemail</td>
<td>Caller wants to be connected to voicemail server</td>
</tr>
<tr>
<td></td>
<td>attendant</td>
<td></td>
<td></td>
</tr>
<tr>
<td>language</td>
<td>language tag</td>
<td>language=&quot;en,de,se,!fi&quot;</td>
<td>Connect caller to someone who speaks English, German, Swedish, not Finnish</td>
</tr>
<tr>
<td>media</td>
<td>MIME types</td>
<td>media=&quot;text/html&quot;</td>
<td>use HTML as the media type</td>
</tr>
<tr>
<td>mobility</td>
<td>fixed</td>
<td>mobility=fixed</td>
<td>connect to the callee’s fixed rather than mobile terminal</td>
</tr>
<tr>
<td></td>
<td>mobile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>priority</td>
<td>urgent</td>
<td>priority=urgent</td>
<td>call is urgent (as seen by the caller).</td>
</tr>
<tr>
<td></td>
<td>emergency</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>non-urgent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>service</td>
<td>fax</td>
<td>IP</td>
<td>ISDN</td>
</tr>
</tbody>
</table>
Contact header example

Contact: maguire <sip:maguire@it.kth.se> ;language="en,de,se,!es"
    ;media="audio,video,application/chat"
    ;duplex="full"
    ;priority="urgent"
Accept/Reject-Contact header(s)

SIP request contains Accept-Contact and Reject-Contact headers

Reject-Contact indicates URI’s not acceptable

Accept-Contact indicates ordered list of acceptable URI’s

Indication by means of rules
  • set intersection and non-intersection of parameters
  • string match of URIs

Example:

Accept-Contact: sip:sales@acme.com ;q=0,
   ;media="!video" ;q=0.1,
   ;mobility="fixed" ;q=0.6,
   ;mobility="!fixed" ;q=0.4

In the second example, the caller does not want to talk to sales@acme.com, but has a preference for video and somewhat prefers the user’s fixed to non-fixed (i.e., mobile) terminal.
Callee (i.e., called party) Parameter processing

- Proxy obtains list of URI’s and the parameters for each, for callee
- Those that match a rule in Reject-Contact are discarded
- Matching set of URI’s determined
- q parameters merged
- Result split into sets of q-equivalency classes
- Parallel search of highest preference q-equivalence class
Accept-Contact Example

example from http://www.ietf.org/proceedings/99nov/I-D/draft-ietf-mmusic-sip-caller-00.txt

1. Note that the internet draft [106] was replaced by [107].

```
sip:mjh@aciri.org;language=en;media=audio,video;q=.8
sip:m.handley@acm.org;class=business;q=0.3
```

INVITE
sip:mark.handley@company.co

PROXY

From Location Service

INVITE sip:markh@aol.com
SIP/2.0

---

Maguire
maguire@it.kth.se

Accept-Contact Example
2007.03.27

Module 9: 274 of 279
Practical Voice Over IP (VoIP): SIP and related protocols
# Request-Disposition

Defines services desired from proxy servers

<table>
<thead>
<tr>
<th>Feature values</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>proxy</td>
<td>whether to proxy or redirect</td>
</tr>
<tr>
<td>redirect</td>
<td></td>
</tr>
<tr>
<td>cancel</td>
<td>whether to return just the first 200-class response, or all 2xx responses</td>
</tr>
<tr>
<td>no-cancel</td>
<td></td>
</tr>
<tr>
<td>fork</td>
<td>whether to fork or not (i.e., proxy to only a single address)</td>
</tr>
<tr>
<td>no-fork</td>
<td></td>
</tr>
<tr>
<td>recurse</td>
<td>whether a proxy server upon receiving a 3xx-class response should recurse (i.e., send requests to the addresses listed in the response) or not (i.e., simply forward the list of addresses upstream towards the caller)</td>
</tr>
<tr>
<td>no-recurse</td>
<td></td>
</tr>
<tr>
<td>parallel</td>
<td>For a forking proxy server, should it send the request to all known addresses at once (parallel), or go through them sequentially, i.e., contacting the next address only after receiving a non-2xx or non-6xx final response.</td>
</tr>
<tr>
<td>sequential</td>
<td></td>
</tr>
<tr>
<td>queue</td>
<td>If called party is temporarily unreachable, caller can indicate that it wants to enqueue rather than be rejected immediately.</td>
</tr>
<tr>
<td>no-queue</td>
<td>Pending call be terminated by a SIP CANCEL or BYE request.</td>
</tr>
</tbody>
</table>

Based on a list of keywords

- example: Request-Disposition: fork, parallel
### SIP Service Examples

Some examples of SIP Services are listed below (from [128])

<table>
<thead>
<tr>
<th>Feature</th>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Call Hold</td>
<td>Single Line Extension</td>
</tr>
<tr>
<td>Consultation Hold</td>
<td>Find-Me</td>
</tr>
<tr>
<td>Music On Hold</td>
<td>Call Management (Incoming Call Screening)</td>
</tr>
<tr>
<td>Unattended Transfer</td>
<td>Call Management (Outgoing Call Screening)</td>
</tr>
<tr>
<td>Attended Transfer</td>
<td>Call Park</td>
</tr>
<tr>
<td>Call Forwarding Unconditional</td>
<td>Call Pickup</td>
</tr>
<tr>
<td>Call Forwarding - Busy</td>
<td>Automatic Redial</td>
</tr>
<tr>
<td>Call Forwarding - No Answer</td>
<td></td>
</tr>
<tr>
<td>3-way Conference - Third Party is Added</td>
<td></td>
</tr>
<tr>
<td>3-way Conference - Third Party Joins</td>
<td></td>
</tr>
</tbody>
</table>

You should compare these to the list we saw earlier: **Features** on page 185
Privacy-Conscious Personalization

Bell Labs’ has developed software designed to give cell phone users greater control over the disclosure of their location[130].

Preferences could depend on:
• who is requesting the location data,
• what time of day it is,
• or the callers’ activities,
• … .

Requests for location are then filtered through these preferences, and are permitted or blocked accordingly.

Operators might provide users with a selection of “preference palettes” to start with, the user could then customize their preferences over time.
User Preferences


http://www.ietf.org/rfc/rfc3841.txt

http://www.ietf.org/internet-drafts/draft-ietf-sipping-service-examples-08.txt

Module 10: SIP Security, NATs, and Firewalls

Lecture notes of G. Q. Maguire Jr.
SIP Security

SIP Security - RFC 3261 [131]

If you want to secure both the SIP and RTP traffic, then you should probably be using an IPSec VPN.

SIP’s rich signalling means that the traffic reveals:

- caller and called parties IP addresses
- contact lists
- traffic patterns

For further details concerning how complex it is to protect such personal information see the dissertation by Alberto Escudero-Pascual, “Privacy in the next generation Internet, Data Protection in the context of European Union Data Protection Policy” [159].

For an example of a call anonymizer service -- using a back-to-back user agent (B2BUA), see figure 8.6 on page 121 of Sinnreich and Johnston.
SIP Digest Authentication

Built upon HTTP’s challenge/response mechanism

Challenges:
- 401 Authentication Required or
- 407 Proxy Authorization Required

Header fields:

Digest
username="A"
realm="sip:proxy.com"
nonce="e288df84f1cecc341ade6e5a359"
opaque="63632f41"
uri="sip:UserB@there.com"
response="1d19580cd833064324a787ecc"

the schema name
The user name as specified in the credentials
realm - copied from the challenge
realm indicates the domain for the authentication
nonce - copied from the challenge
a unique string - typically generated from a timestamp (and possibly a seed), then encrypted with the user’s private key
opaque string which should be returned unchanged to be matched against the challenge (allows for a stateless system)
URI from the Request-URI
message digest computed using user’s credentials and the nonce
RFC 3261 describes the use of Secure MIME (S/MIME) message bodies:

- SIP header fields can be encrypted in an S/MIME message body
- see RFC 2633 [132]

Provides:

- **Message integrity**
  - Allows detection of any modification of message contents
- **Message privacy**
  - Private headers protected by S/MIME
- **Identity**
  - Certificates can be verified to validate identity
SDP & RTP security

As noted earlier SDP enables you to say that you will encrypt the media stream which is sent via RTP - such as DES in CBC Mode (DES-CBC) or AES in f8-mode [139].

This is done via adding to the SDP for each media description:

\[ k = \text{encryption key} \]

1. All encryption capable RTP clients must support this as their default algorithm. In addition, to prevent known plain text attacks, RTCP headers have a 32 bit random prefix.
User identity

J. Peterson and C. Jennings in an IETF draft [133] define mechanisms and practices to assure the identity of the end user that originates a SIP request (does not cover identity for responses).

Their identity mechanism derives from the following principle:

If you can prove you are eligible to register in a domain under a particular address-of-record (AoR), then you are also proving that you are capable of receiving requests for that AoR.

∴ when you place that AoR in the From header field of a SIP request other than a registration (e.g., INVITE), you are providing a 'return address' where you can legitimately be reached.

adapted from [133]

Introduces:
(a) authentication service (at either a user agent or a proxy server) and
(b) two new SIP headers, Identity & Identity-Info headers
Identity header example

from [133]

```
INVITE sip:bob@biloxi.example.org SIP/2.0
Via: SIP/2.0/TLS pc33.atlanta.example.com;branch=z9hG4bKnashds8
To: Bob <sip:bob@biloxi.example.org>
From: Alice <sip:alice@atlanta.example.com>;tag=1928301774
Call-ID: a84b4c76e66710
CSeq: 314159 INVITE
Max-Forwards: 70
Date: Thu, 21 Feb 2002 13:02:03 GMT
Contact: <sip:alice@pc33.atlanta.example.com>

Identity:
"CyI4+nAkHrH3ntmaxgr01TMxTmtjP7MASwliNRdupR1vpkXRvZXx1ja9k0nB2sN3W+v1PDsy32MaqZi0M5WfEkXxbgTnPvYW0jIoK8HMyY1VT7egt0kk4XrKFCHYWGC1sM9CG4hq+YJZTMaSR00MUBhikVijnQ8ykeD6UXNOyfI="
Identity-Info: https://atlanta.example.com/cert
Content-Type: application/sdp
Content-Length: 147

v=0
o=UserA 2890844526 2890844526 IN IP4 pc33.atlanta.example.com
s=Session SDP
c=IN IP4 pc33.atlanta.example.com
t=0 0
m=audio 49172 RTP/AVP 0
a=rtpmap:0 PCMU/8000
```
Saying **BYE** also needs to be authenticated!

```
BYE sip:alice@pc33.atlanta.example.com SIP/2.0
Via: SIP/2.0/TLS 192.0.2.4;branch=z9hG4bKnashds10
Max-Forwards: 70
From: Bob <sip:bob@biloxi.example.org>;tag=a6c85cf
To: Alice <sip:alice@atlanta.example.com>;tag=1928301774
Date: Thu, 21 Feb 2002 14:19:51 GMT
Call-ID: a84b4c76e66710
CSeq: 231 BYE
Identity: 
"A5oh1tSWpbmXTyXJDhaCiHjT2xR2PAwBroi5Y8tdJ+CL3ziY72N3Y+1P8eoiXlrZ0uwb0DicF9GGxA5vw2mCTUxc0XG0KJOhpBnzoXnuPNAZdcZEWsVOQAKj/ERsYR9BfxNPazWmJZjGmDoFDbUNamJRjiEPOKn13uAZIcuf9zM=
Identity-Info: https://biloxi.example.org/cert
Content-Length: 0
```
Erik Eliasson’s miniSIP

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
- Configuration used in the test described next:
  - 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)

Described in IETF RFC 3711 [142], provides confidentiality, message authentication, and replay protection for RTP and RTCP traffic.

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


- 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) [144] as the key management protocol


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

**Secure Call Setup [137]**

<table>
<thead>
<tr>
<th>Total delay (in ms)</th>
<th>Calling Delay</th>
<th>Answering Delay</th>
</tr>
</thead>
<tbody>
<tr>
<td>No security</td>
<td>19.5</td>
<td>9.5</td>
</tr>
<tr>
<td>MIKEY, shared key</td>
<td>20.9</td>
<td>10.5</td>
</tr>
<tr>
<td>MIKEY, Diffie-Hellman</td>
<td>52.5 (UDP)</td>
<td>47.6 (UDP)</td>
</tr>
<tr>
<td></td>
<td>58.9 TCP</td>
<td>48.9 (TCP)</td>
</tr>
</tbody>
</table>

- 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

Joachim Orrblad in his thesis, “Alternatives to MIKEY/SRTP to secure VoIP” [136], examines the use of MIKEY together with IPSec.
Timed
Efficient Stream Loss-tolerant Authentication (TESLA)

SRTP TESLA [145] was designed to provide efficient data origin authentication for multicast and broadcast session.

This is needed since we don’t want to create all possible pairwise authentications for the participants in a conference.
NATs and Firewalls

Because Network Address Translation (NAT) devices change addresses and sometimes port numbers and because addresses and port numbers are inside both SIP and SDP there can be a problem!

Fredrik Thernelius, “SIP, NAT, and Firewalls”, looked at this in detail in his M.Sc. thesis[146].

See also the other documents at

Note: CNAME’s in RTCP may need to be updated by the Network Address Translation (NAT) to hide private network addresses.

To protocols being developed to help deal with NATs:
• Simple Traversal of User Datagram Protocol Through Network Address Translators (STUN)
• Globally Routable User Agent Universal (GRUU) Resource Indicator[152]
• a URI which can be used by anyone on the Internet to route a call to a specific UA instance

See also pages 237-239 of *Practical VoIP: Using VOCAL*[1]; particularly the example of using a Cisco ATA (Analog Telephone Adaptor) behind a Linksys firewall (which configures the firewall to pass incoming traffic on port 5060, 4000, and 4001 to the Cisco ATA) - which also refers to [http://www.dyndns.org/](http://www.dyndns.org/)

See the internet drafts:

• Interactive Connectivity Establishment (ICE): A Methodology for Network Address Translator (NAT) Traversal for Multimedia Session Establishment Protocols [158]

• How to Enable Real-Time Streaming Protocol (RTSP) traverse Network Address Translators (NAT) and interact with Firewalls (<draft-ietf-mmusic-rtsp-nat-03> has expired)
### Types of NAT

**Source NAT**
All callers look like they come from the same IP address

**Destination NAT**
Which internal address should traffic to a given port be forwarded to?

## Four types of NATs [153] and

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Cone</td>
<td>maps a specific internal IP address and port number to a given external IP address and port number&lt;br&gt;This is the only type of NAT that allows an external host to contact an internal host (i.e., behind the NAT) <strong>without</strong> having previously received packets from this internal host.</td>
</tr>
<tr>
<td>Restricted Cone</td>
<td>external hosts must have the IP address of an internal host <strong>prior</strong> to communicating with this internal host</td>
</tr>
<tr>
<td>Port Restricted Cone</td>
<td>external hosts must have the IP address and port number of an internal host <strong>prior</strong> to communicating with this internal host</td>
</tr>
<tr>
<td>Symmetric</td>
<td>assigns unique internal IP address and port numbers based on the specific internal destination</td>
</tr>
</tbody>
</table>
Cone vs. Symmetric NAT

Figure 15: (a) Cone NAT vs. (b) Symmetric NAT - figure inspired by figures 1 and 2 of [157]
NAT traversal methods

- Symmetric media streams
- STUN protocol
  - also: Extended STUN for Symmetric NAT
- rport SIP extension
  - See RFC 3581[154] - defines a new parameter for the Via header field, called "rport", this "allows a client to request that the server send the response back to the source IP address and port from which the request originated."
- OPTIONS request registration refresh
  - Causes the UA to send traffic out - thus refreshing the NAT bindings
- Outgoing INVITE transaction refresh
- Traversal using Relay NAT (TURN)
  - insert a server in the media and signalling path (to deal with Symmetric NATs)
- Application Layer Gateway (ALG)
  - Here the NAT knows about SIP and "does the right thing"
- Universal Plug and Play (UPnP)
  - Use UPnP to control the NAT to open a specific “pinhole” in the firewall
- Manual Configuration
  - manually configure a set of addresses and ports for SIP to use
• Tunnel
  • Tunnel the traffic - inside IPsec, HTTP (i.e., act like HTTP), …

A NAT support “hairpinning” if it can route packets coming from the private network addressed to a public IP address back into the private network. For example, a mobile user might actually be connected to the private network - thus packets to this user don’t actually need to be sent out and then sent back into the private network!
STUN (Simple Traversal of UDP through NATs (Network Address Translation))

STUN, defined in RFC 3489 [150], assists devices behind a NAT firewall or router with their packet routing.

• enables a device to find out its public IP address and the type of NAT service its sitting behind
  • By querying a STUN server with a known public address, the STUN client learns the public IP and port address that were allocated by this client’s) NAT.
• operates on TCP and UDP port 3478
• uses DNS SRV records to find STUN servers attached to a domain. The service name is _stun._udp or _stun._tcp
• Unfortunately, it is not (yet) widely supported by VOIP devices

Note: The STUN RFC states: This protocol is not a cure-all for the problems associated with NAT.

Open source STUN servers from vovida.org, larry.gloo.net, stun.fwdnet.net, stun.softjoys.com, and others.
STUN steps

1. Client queries a STUN server for a shared secret username and password.
2. Server responds with a unique username/password combination for this client.
3. Client sends a binding request using this username/password to the server via UDP.
4. Server copies the source IP and port number into a binding response, and sends this response back to the client.
5. Client compares the IP address and port number received from the server with its local IP address and port number. If they do not match, then the client is behind some type of NAT.

- A full flowchart to find each of the potential situations is shown as Figure 14 “Flow Chart: Determining NAT type” in [153].
UDP and TCP Firewall Traversal problems

1. INVITE/UDP
   User Agent A
   (outside) Firewall (inside)
   2. 180 Ringing/UDP
   3. 200 OK/UDP
   RTP packets/UDP
   4. INVITE/TCP
   5. 180 Ringing/TCP
   6. 200 OK/TCP
   7. ACK/TCP
   RTP packets/UDP

Using UDP all of B’s responses and packets are filtered out by the firewall and there is no session!

Using TCP for SIP enables the session to be setup, but B’s RTP packets are still filtered out by the firewall!
UDP and TCP NAT Traversal problems

1. INVITE
   via:10.1.2.3  
   Contact: A@10.1.2.3  
   SDP 10.1.2.3  

2. 180 Ringing
   via:10.1.2.3; received=1.2.3.4  
   Contact: B@130.101.102.103

3. 200 OK
   via:10.1.2.3; received=1.2.3.4  
   Contact: B@130.101.102.103  
   SDP 130.101.102.103  

4. ACK

SIP can negotiate the NAT, but A’s SDP contains a private address

B’s RTP packets are directed to a private address and hence can not be routed; similarly B’s requests also fail

UDP and TCP NAT Traversal problems

Maguire  
maguire@it.kth.se

UDP and TCP NAT Traversal problems  
2007.03.27  

Module 10: 302 of 325  
Practical Voice Over IP (VoIP): SIP and related protocols
SIP Application Level Gateway (ALG) for Firewall Traversal

Use a proxy within the (possibly private) network:

1. INVITE
2. INVITE sdp ALG
3. 100 Trying

User Agent A  SIP ALG Proxy  Firewall/NAT  User Agent B

4. 180 Ringing
5. 180 Ringing
6. 200 OK sdp B
7. 200 OK sdp ALG
8. ACK
9. ACK

RTP Media Session

10. BYE
11. BYE
12. 200 OK
13. 200 OK

Firewall permits SIP and RTP traffic to/from the Application Level Gateway (ALG) proxy.
The generic problem of enabling complex applications through the middleboxes is being addressed by the Middlebox communications (MIDCOM) Working Group, they do so via MIDCOM agents which perform ALG functions, logically external to a middlebox [148].
Application aware Middlebox

Newport Networks’ Automatic Channel Mapping™ (ACM) [156]:

- SignallingProxy™ acts as a high-performance B2BUA (Back to Back User Agent)
- MediaProxy™ provides a transit point for RTP and RTCP media streams between User Agents
Security flaws in Abstract Syntax Notation One (ASN.1)

Note that the vulnerability was discovered in June 2002!

The United Kingdom National Infrastructure Security Co-Ordination Centre revealed in Jan. 2004, “that it had discovered security flaws that affect the products of dozens of vendors. The flaws were found in software that support a variety of applications and technologies, including voice over IP, videoconferencing, text messaging, Session Initiation Protocol, devices and hardware, and critical networking equipment such as routers and firewalls.” …

“CIOs need to be aware that voice over IP creates exposure to vulnerabilities, says David Fraley, a principal analyst at Gartner Dataquest. "While there are very real and neat opportunities with VoIP, as convergence increases, the risks to attacks to these systems are going to increase," he says.


Risks range from denial-of-service attacks to allowing access to malicious code. according to the

see http://www.cert.org/advisories/CA-2004-01.html#vendors
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

- Location mis-direction $\Rightarrow$ End of Sovereignty

- Traffic pattern hiding

- Traffic hiding
Swedish Electronic Communications Act

Swedish Electronic Communications Act (SFS 2003:389) [160] (see also http://www.pts.se/Sidor/sida.asp?SectionID=1340 ) provides the regulatory framework for electronic communications networks and services. It is based on EU directives and became effective on July 25th, 2003. It defines what/who an operator is and what their obligations are. (note: it replaces the earlier swedish definition of “teleoperator”).

It is relevant to publically available telephone services in 3 major areas:
• emergency calls (Chapter 5, section 7)[160]
• number portability (Chapter 5, section 9)[160], and
• legal intercept (Chapter 6, section 19)[160]
The lawful “use of electronic recording equipment” - when can you make a recording of a call’s contents (i.e., wiretapping and eavesdropping)?

The US Federal government (18 U.S.C. Sec 2511,) and many states have “one-party consent” statutes, i.e., if you are a party to the conversation you can record it. However, note that not all states permit this (some have an “all-party” rule)! Note that these rules often apply to in-person recordings, radio/telecommunication, … , all “electronic communications”.

There are additional rules concerning Broadcasters - who must inform the person that the recording may be subsequently broadcast before the recording begins.

A summary of the rules for the US can be found at:
http://www.rcfp.org/taping/index.html

In addition, there are also laws concerning “employee privacy” which may also be relevant.
There is a proposal that Communications Assistance for Law Enforcement Act (CALEA) {47 U.S.C. § 1001 et seq. [161]} should be applied to VoIP services (and other data services) to "conduct lawful electronic surveillance":


Types of surveillance [164]:

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;pen register&quot;</td>
<td>records call-identifying information for calls originated by a subject</td>
</tr>
<tr>
<td>&quot;trap and trace&quot;</td>
<td>records call-identifying information for calls received by a subject, and</td>
</tr>
<tr>
<td>&quot;interception&quot;</td>
<td>records the conversations of the subject, as well as call identifying information</td>
</tr>
</tbody>
</table>

There is a great variety of proposals for LI [173].
Reasonably Available Information

Operators are only required to provide information to law enforcement if it is reasonably available. For example, “call-identifying information is reasonably available to a carrier if it is present at an intercept access point and can be made available without the carrier being unduly burdened with network modifications”.

The EU statute is similar in identifying when information is technically feasible and economically feasible available.

Thus **Call Forwarding Information** might not always be reasonably available in a SIP environment - since the call forwarding could happen outside the control of a given operator.

Similarly **Dialed-Digit Extraction** might not be available in a SIP environment since the actual IP address of the source and destination might be inside encrypted SDP.
EU privacy and Lawful Intercept (LI)

EU Directive 95/46/EC - Data Protection Directive,
EU Directive 97/66/EC - Telecommunications Data Protection, and


A good summary of the EU situation can be found at [163].

ETSI is defining a standard LI architecture see [166] and [167]. For a list of the LI standards as collected by the Global LI Industry Forum, Inc. [169] see [170].
Intercept architecture

- The existence of Intercepts should be transparent to both the subject and other LEAs!
- The dotted links (probably SNMPv3) must be secured to prevent Unauthorized Creation and Detection of intercepts - while solid red links must be secured to protect intercept related information (IRI) [165]
- Intercept [Access] Point (IAP): router, PSTN gateway, SIP proxy, RADIUS server, ...

Figure 16: Interfaces in RED should be standard to allow interoperability; HI_n = Handover Interface_n
Voice over IP Security Alliance

The Voice over IP Security Alliance (http://www.voipsa.org/) was formed February 7, 2005

They have a moderated mailing list: VOIPSEC
Spam over Internet Telephony (SPIT)

There is rising concern that misconfigured voice gateways, … will lead to increased IP telephony SPAM.

One solution is using speaker recognition and then checking to see if this speaker is on:

- a white list (automatic accept),
- a black list (automatic reject), or
- unknown (message could be recorded and the user listens to it later and then adds the user to their white or black lists).

See for example [171].

Issues of SIP and SPAM and solutions in addition to the above are discussed in [172].
References and Further Reading

SIP Security


RTP encryption


http://www.ietf.org/rfc/rfc1423.txt

<draft-bloom-rtp-encrypt-00.txt>


<draft-ietf-avt-srtp-05.txt>
http://www.ietf.org/rfc/rfc3830.txt

http://www.rfc-editor.org/rfc/rfc4383.txt

NATs and Firewalls


[147] List of sources about SIP and Firewalls

[148] P. Srisuresh, J. Kuthan, J. Rosenberg, A. Molitor, and A. Rayhan,
“Middlebox Communication Architecture and framework”, IETF RFC 3303, August 2002
http://www.ietf.org/rfc/rfc3303.txt

http://www.ietf.org/rfc/rfc3304.txt

http://www.ietf.org/rfc/rfc3489.txt


21, 2005, expires August 22, 2005

http://www.ietf.org/internet-drafts/draft-ietf-sip-gruu-03.txt


http://www.ietf.org/rfc/rfc3581.txt


Privacy


http://www.imit.kth.se/~aep/PhD/docs/escuderoa-PhD-20021030.pdf


http://www.techlawjournal.com/agencies/calea/47uscl001.htm


[163] Jaya, Baloo, Lawful Interception of IP Lawful Interception of IP Traffic, Draft 1,

http://www.blackhat.com/presentations/bh-europe-03/bh-europe-03-baloo.pdf


http://www.ietf.org/rfc/rfc3924.txt

[166] ETSI TS 101 331, Telecommunications security; Lawful Interception (LI); Requirements of law enforcement agencies, V1.1.1, August 2001.


SIP Telephony

SIP Telephony (SIP-T) -- for details see RFC 3204 [180].

Gateway between the SIP world and the PSTN world looks like a SIP user agents to other SIP entities and like a terminating telephone switch to the PSTN.

Advantages Provides ISUP transparency (by carrying ISUP message as multipart MIME messages in the SIP messages between SIP-T gateways)

Disadvantages Does not interwork with SIP
Perpetuates ISUP!

For example of call flows between SIP and PSTN see [181].

Stream Control Transmission Protocol (SCTP) can be used to carry telephony signalling [185].
Telephony Routing over IP (TRIP)

- TRIP[182] is a gateway to Location Server (LS) protocol
- Designed for an interdomain gateway
- Allows the gateway to advertise what PSTN number range it is a gateway for

For within a domain there is a version for between a gateway and a proxy: TRIP-lite

A Location Server is responsible for a Internet Telephony Administrative Domain (ITAD).

See also: **Telephony Routing over IP (TRIP)** on page 354 and **Telephony Gateway REgistration Protocol (TGREP)** [184].
Call Control Services

Generally include advanced telephony services such as:

- Call Transfer, both Attended and Unattended
- Call Park/Un-Park
- Multistage Dialling
- Operator Services
- Conference Call Management
- Call Mobility
- Call Pickup

See the slides starting on *Intelligent Network service using SIP* on page 182.
Call Center Redesign using SIP

- Replace the call center switch via VoIP
- Interactive Voice Response (IVR) - using a media server (for pre-recorded clips) and SIP signalling
- Automatic Call Distribution (ACD) - replace with scripts using Call Processing Language (CPL)
- Agent Workstation - a PC with a SIP client
- The agent has access via Web and various databases to information, which can be indexed by the agent using information from the SIP request.
Additional SIP Telephony services

• SIP for the Hearing Impaired
• Emergency Services
• Precedence signalling (military, government, emergency services, …)
  • RFC 3487 [174] gives the requirements for resource priority mechanisms for SIP
• Message Waiting, Voice Mail, and Unified Messaging
  • See for example Interactive Intelligence’s Communité® (“ka-mune-i-tay”)
• Call Waiting
• SIP *continuing* presence service
  • The I-Am-Alive (IAA) database [179] is a distributed database system that users can query after-the-event to determine the status of a person - it does not require the session properties of SIP
  • Is there a SIP corollary - for *continuing* presence?
Telephony Signaling when used in Internet-based telephony services in addition to the general requirements specified in [177] needs to support a number of additional requirements [178]:

- Telephony signaling applications (used with Internet-based telephony) **must** be able to carry labels.
- The labels **must** be extensible
  - to support various types and numbers of labels.
- These labels **should** have a mapping to the various emergency related labels/markings used in other telephony based networks, e.g., PSTN
  - To ensure that a call placed over a hybrid infrastructure (i.e., PSTN+Internet) can carry the labels end-to-end with appropriate translation at PSTN/Internet boundaries.
  - Only authorized users or operators **should** be able to create non-ordinary Labels (i.e., labels that may alter the default best effort service).
  - Labels **should** be associated with mechanisms to providing strong end-to-end integrity
  - Operators **should** have the capability of authenticating the label
• Application layer IP telephony capabilities **must not** preclude the ability to do application layer accounting.

• Application layer mechanisms in gateways and stateful proxies that are specifically in place to recognize ETS type labels **must** be able to support “best available” service (i.e., better than “best effort”).
Emergency Services (E911)

We need to support 3 things[175]:

• There must exist an emergency address (similar to 911, 112, help, …)
• find Public Safety Answering Point (PSAP)
  • outbound proxy -- only if there is a well bounded geographic area servered by this proxy
  • use DNS where the user or device enters a relevant name: e.g., pittsburgh.pa.911.arpa
  • SLP - but scope not likely to coincide with ESR
• call volume:
  – Sweden: SOSAlarm.se has 20 call centers distributed around Sweden with ~18 million calls/year with ~20% of them calls to 112 the rest are automatic alarms;
  – US: National Emergency Number Association (NENA) reports >500,000 calls/day or 190 million a year (more than 80% are not emergencies ⇒ 311 non-emergency number)
• obtain caller’s identity and geographical address
  • this is done to minimize prank calls
  • caller provides in request
    – Geographic position: N 59° 24.220' E017° 57.029' +/- 77m and/or
    – Geographic Location: "5th floor, Isafjordsgatan 22, Kista, Stockholm, Sweden"
  • or PSAP queries caller
  • or PSAP queries third party based on caller identity

note: Enhanced 911 (E911) - mandated by FCC for cellular phones in US
Public Safety Answering Point (PSAP)

For example MapInfo has an E911 database called “PSAP Pro” which contains the following PSAP information for the U.S.:

- 10-digit emergency numbers
- Administrative phone number
- Contact person
- Jurisdictional boundaries
- Address information
- Fax number
- Latitude and longitude

~4,400 records: both primary PSAPs and sheriff’s departments and offices in areas not served by a PSAP.


So finding the nearest one can be done based on geography, but is it the most relevant or useful one? In Sweden SOS Alarm works with the digital maps of CoordCom.

Location Interoperability Forum became part of Open Mobile Alliance (OMA) and no longer exists separately: http://www.openmobilealliance.org/tech/affiliates/lif/lifindex.html
Vonage 911 service


• User must pre-designate the physical location of their Vonage line and update Vonage when the user moves

• 911 dialing is not automatically a feature of having a line
  • users must pre-activate 911 dialing
  • user may decline 911 dialing

• A 911 dialed call will be connected to a general access line at the Public Safety Answering Point (PSAP)
  • thus they will not know you phone number or location

• Service may not be available due to
  • a local power failure (your IP phone needs power)
  • you local ISP not being able to offer service
  • one of the transit networks not being able to offer service
  • the voice gateway to the PSTN not being in service
  • ...
Vonage equips PSAs with VoIP

Vonage Equips Over 100 New Counties and 400 Calling Centers With E911 in Just One Month, Vinage Press Release, March 7, 2006

- "Nearly 65 Percent of Vonage Customers Now Have E911"
- "In February alone, Vonage equipped an additional 400 calling centers in over 100 new counties with E911 -- bringing the total number of calling centers across the nation with E911 service to over 3400, which is more than half of the nation’s calling centers. While it took Vonage less than a year to turn on E911 in more than one-half of the nation’s PSAP’s, it took the wireless industry 10 years to accomplish the same feat."
- "In the event Vonage is unable to connect to the 911 system or for customers who are using mobile devices such as wifi phones or soft clients, Vonage offers a national emergency call center which enables customers to get local help when they need it."
GEOPRIV (http://www.ietf.org/html.charters/geopriv-charter.html) an IETF working group tasked with establishing a means of disseminating geographic data that is subject to the same sorts of privacy controls as presence is today.

Jon Peterson, “A Presence-based GEOPRIV Location Object Format”, IETF draft (original version 14-Jan-04 - current version September 9, 2004), based on earlier work done in formulating the basic requirements for presence data -- the Presence Information Data Format (PIDF).

Emergency services


[176] Europe’s 112 web site: http://www.sos112.info/


**SIP Telephony**


**TRIP**


http://www.rfc-editor.org/rfc/rfc4166.txt

Module 12: SIP Conferencing

Lecture notes of G. Q. Maguire Jr.
Conferencing

- Multimedia conferencing
  - Synchronized Multimedia Integration Language (SMIL) to enable other media (e.g., text, graphics and URLs) to be added to audio/video streams for synchronized display[194]
  - SMIL documents are XML 1.0 documents

- Multipoint conferencing
  - can exploit multicast where available

- Call control for conferencing

- Floor control [190]
  - this a particular focus of Push-to-talk service [191]
  - see also Florian Maurer’s push-to-talk service for minisip
    http://push2talk.floHweb.ch

- RFC 4245: High-Level Requirements for Tightly Coupled SIP
Conferencing defines the media types for the languages of the W3C Speech Interface Framework [189]:
  - Voice Extensible Markup Language (VoiceXML),
  - Speech Synthesis Markup Language (SSML),
  - Speech Recognition Grammar Specification (SRGS),
  - CallControl XML (CCXML), and
  - Pronunciation Lexicon Specification (PLS).
## Conferencing Models [187]

<table>
<thead>
<tr>
<th>Type of Conference</th>
<th>Description</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endpoint mixing</td>
<td>One end point acts as a mixer for all the other end points</td>
<td>small</td>
</tr>
<tr>
<td>SIP Server and distributed media</td>
<td>Central SIP server establishes a full mesh between all participants - each participant does their own mixing</td>
<td>medium</td>
</tr>
<tr>
<td>Dial-in conference</td>
<td>All participants connect to a conference bridge which does the mixing for each participant</td>
<td>medium</td>
</tr>
<tr>
<td>Ad hoc centralized conference</td>
<td>Two users transition to a multiparty conference, by one of them using third-party signaling to move the call to a conference bridge</td>
<td>medium</td>
</tr>
<tr>
<td>Large multicast conference</td>
<td>user join the multicast based on the multicast address (which they got via:</td>
<td>small to</td>
</tr>
<tr>
<td></td>
<td>• announcement on the web</td>
<td>very large</td>
</tr>
<tr>
<td></td>
<td>• e-mail</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Session Annoucement Protocol (SAP) [192]</td>
<td></td>
</tr>
</tbody>
</table>

Commercial conference bridge authenticate the users joining the conference.
RFC 4353 [188] defines SIP procedures for the following common operations:

- Creating Conferences
- Adding Participants
- Removing Participants
- Destroying Conferences
- Obtaining Membership Information
- Adding and Removing Media
- Conference Announcements and Recordings

How to realize conferences

- Centralized Server
- Endpoint Server
- Media Server Component
- Distributed Mixing
- Cascaded Mixers
Speaker recognition in a conference

Abstract:

A system and method for identifying a participant during a conference call include the capability to receive a packet containing data that represents audible sounds spoken by one of a plurality of participants in a conference call and to determine a speaker of the audible sounds using voice profile information of the participants. The system and method further include the capability to provide identification information of the speaker to the other participants in the conference call contemporaneously with providing audible sounds based on the data to those participants.

Shmuel Shaffer and Michael E. Knappe, US patent 6,853,716 [195]
References and Further Reading

SIP Conferencing


http://www.rfc-editor.org/rfc/rfc4353.txt


http://www.ietf.org/internet-drafts/draft-ietf-sipping-conferencing-requirements-01.txt


http://www.rfc-editor.org/rfc/rfc4376.txt


**Session Announcement Protocol**


http://www.ietf.org/rfc/rfc2974.txt


**SMIL**


**Speaker recognition in a conference**

Mixed Internet-PSTN Services

- PSTN and Internetworking (PINT)
- Servers in the PSTN Initiating Requests to Internet Servers (SPIRITS)
- Telephony Routing over IP (TRIP)
PSTN and Internetworking (PINT)

PSTN and Internetworking (PINT)[196] - action from the internet invokes a PSTN service (note: this is one way invocation), examples:

- Request to Call ⇒ “Click to Connect” from a web page
- Request to Fax Content ⇒ “Click to FAX”
- Request to Speak/Send/Play Content
- ...

Based on SIP extensions (SIPext), which in actuality are SDP extensions (i.e., the body of SIP messages). Redefines some methods (INVITE, REGISTER, and BYE) and introduces three new methods:

- **Subscribe** - request completion status of a request
- **Notify** - receive status updates
- **Unsubscribe** - cancel subscriptions

PINT extensions to SDP: Network type (TN) and Address type: RFC2543 (SIP)
SPIRITS protocol [199] - implementing a family of IN services via internet server (rather than in the PSTN)

For example, internet call waiting (ICW) - calling a busy phone in the PSTN network could pop up a call waiting panel on the client that is using this telephone line, this replaces earlier solutions such as:

- for example, Ericsson’s PhoneDoubler, Ericsson Review, No. 04, 1997
  http://www.ericsson.com/about/publications/review/1997_04/article55.shtml
- PDF of the entire article:

SPIRITS unlike PINT allows two way interaction between Internet and PSTN.

See also [207].
Telephony Routing over IP (TRIP)

Telephony Routing over IP (TRIP) [205] Finding a route from the Internet to a gateway nearest to where the call should be terminated

Telephony Routing Protocol is modeled after the Border Gateway Protocol (BGP)
References and Further Reading

PINT


SPIRITS

[199] V. Gurbani (Editor), A. Brusilovsky, I. Faynberg, J. Gato, H. Lu, and M. Unmehopa, “The SPIRITS (Services in PSTN requesting Internet Services)
Protocol”, IETF RFC 3910, October 2004 [200]

I. Faynberg, H. Lu, and L. Slutman, “Toward Definition of the Protocol for PSTN-initiated Services Supported by PSTN/Internet Internetworking”, IETF Internet draft, March 2000, work in progress {expired}


[204]IETF Service in the PSTN/IN Requesting InTernet Service working group
http://www.ietf.org/html.charters/spirits-charter.html

TRIP


ISUP

Module 14: AAA and QoS for SIP

Lecture notes of G. Q. Maguire Jr.
Authentication, Authorization, Accounting (AAA)

This become a major issue especially in conjunction with QoS since for better than best effort service, someone probably has to pay for this high QoS - AAA is necessary to decide who you are, if you are allowed to ask for this service, and how much you should be charged. See [213] and “Authentication, Authorization and Accounting Requirements for the Session Initiation Protocol”[208].
SIP Accounting

For definition of terms see RFC 2975

Purposes:
• controlling resource usage (for example, gateways to PSTN from which someone could place a very expensive international call)
• real-time
  • fraud detection
  • pre-paid subscriptions
• off-line
  • monthly/quarterly billing
  • deriving usage patterns ⇒ planning upgrades (resource dimensioning), input for fraud detection, …

Resources to account for:
• resources used by SIP itself
• resource consumed once initiated by SIP
• services initiated and controlled by SIP {voice mail, media translation/transcoding, …}
Open Settlement Protocol (OSP)

(mostly) off-line settlement between operators based on Call Detail Records

Open Settlement Protocol developed as part of ETSI project TIPHON (Telecommunications and Internet Protocol Harmonization Over Networks) [211]

Based on exchange of Extensible Markup Language (XML) messages via HTTP

```xml
<!DOCTYPE Message [ 
  <!ELEMENT Message(( PricingIndication | PricingConfirmation | AuthorisationRequest | AuthorisationResponse | AuthorisationIndication | AuthorisationConfirmation | UsageIndication | UsageConfirmation | ReauthorisationRequest | ReauthorisationResponse )+ ) >

... ]>
```
Achieving QoS

- Over provision!
  - Simplest approach

- If this fails, then use TOS field or Diffserv
  - Much of the problem is on the access network - hence TOS or Diffserv even only on these links may be enough

- If this fails, then use RSVP
  - Much more complex - especially when done over several operator’s domains
Some measured delays

Actual performance of SIP phone to SIP phone and software applications over a LAN, shows that the performance of SIP phones is well within acceptable delay.

Measurements of mouth to ear one-way delay, from “Aside: SIP phone QoS” slide 15 of [50]

<table>
<thead>
<tr>
<th>end-point A</th>
<th>end-point B</th>
<th>A⇒B</th>
<th>B⇒A</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSM</td>
<td>PSTN</td>
<td>115 ms</td>
<td>109 ms</td>
</tr>
<tr>
<td>3Com</td>
<td>Cisco</td>
<td>51 ms</td>
<td>63 ms</td>
</tr>
<tr>
<td>NetMeeting</td>
<td>NetMeeting</td>
<td>401 ms</td>
<td>421 ms</td>
</tr>
<tr>
<td>Messanger XP</td>
<td>Messanger XP</td>
<td>109 ms</td>
<td>120 ms</td>
</tr>
</tbody>
</table>
Some statistics from Qwest for POP to POP measurements

### Table 1: February Monthly Averages

<table>
<thead>
<tr>
<th>City</th>
<th>Loss (%)</th>
<th>Latency (ms)</th>
<th>Jitter (ms)</th>
<th>Loss (%)</th>
<th>Latency (ms)</th>
<th>Jitter (ms)</th>
<th>Loss (%)</th>
<th>Latency (ms)</th>
<th>Jitter (ms)</th>
<th>Loss (%)</th>
<th>Latency (ms)</th>
<th>Jitter (ms)</th>
<th>Loss (%)</th>
<th>Latency (ms)</th>
<th>Jitter (ms)</th>
<th>Loss (%)</th>
<th>Latency (ms)</th>
<th>Jitter (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atlanta</td>
<td>0.00</td>
<td>39.64</td>
<td>0.05</td>
<td>0.00</td>
<td>24.13</td>
<td>0.05</td>
<td>0.00</td>
<td>45.21</td>
<td>0.05</td>
<td>0.00</td>
<td>52.10</td>
<td>0.08</td>
<td>0.00</td>
<td>20.35</td>
<td>0.00</td>
<td>0.00</td>
<td>61.10</td>
<td>0.14</td>
</tr>
<tr>
<td>Chicago</td>
<td>0.00</td>
<td>39.46</td>
<td>0.09</td>
<td>0.00</td>
<td>24.10</td>
<td>0.08</td>
<td>0.00</td>
<td>23.32</td>
<td>0.07</td>
<td>0.00</td>
<td>56.13</td>
<td>0.33</td>
<td>0.00</td>
<td>20.17</td>
<td>0.01</td>
<td>0.00</td>
<td>48.10</td>
<td>0.09</td>
</tr>
<tr>
<td>Dallas</td>
<td>0.00</td>
<td>24.13</td>
<td>0.05</td>
<td>0.00</td>
<td>24.12</td>
<td>0.05</td>
<td>0.00</td>
<td>21.21</td>
<td>0.07</td>
<td>0.00</td>
<td>40.77</td>
<td>0.07</td>
<td>0.00</td>
<td>44.24</td>
<td>0.02</td>
<td>0.00</td>
<td>46.14</td>
<td>0.10</td>
</tr>
<tr>
<td>Denver</td>
<td>0.00</td>
<td>45.16</td>
<td>0.09</td>
<td>0.00</td>
<td>23.32</td>
<td>0.05</td>
<td>0.00</td>
<td>21.23</td>
<td>0.08</td>
<td>0.00</td>
<td>32.16</td>
<td>0.06</td>
<td>0.00</td>
<td>44.13</td>
<td>0.00</td>
<td>0.00</td>
<td>25.08</td>
<td>0.06</td>
</tr>
<tr>
<td>LA</td>
<td>0.00</td>
<td>52.07</td>
<td>0.06</td>
<td>0.00</td>
<td>56.09</td>
<td>0.20</td>
<td>0.00</td>
<td>40.68</td>
<td>0.07</td>
<td>0.00</td>
<td>32.22</td>
<td>0.06</td>
<td>0.01</td>
<td>76.09</td>
<td>0.02</td>
<td>0.00</td>
<td>8.01</td>
<td>0.07</td>
</tr>
<tr>
<td>NY</td>
<td>0.00</td>
<td>20.36</td>
<td>0.00</td>
<td>0.00</td>
<td>20.21</td>
<td>0.01</td>
<td>0.00</td>
<td>44.23</td>
<td>0.00</td>
<td>0.00</td>
<td>44.24</td>
<td>0.00</td>
<td>0.01</td>
<td>76.17</td>
<td>0.00</td>
<td>0.00</td>
<td>68.19</td>
<td>0.00</td>
</tr>
<tr>
<td>Sunnyvale</td>
<td>0.02</td>
<td>61.14</td>
<td>0.09</td>
<td>0.01</td>
<td>48.20</td>
<td>0.08</td>
<td>0.00</td>
<td>46.17</td>
<td>0.10</td>
<td>0.01</td>
<td>25.12</td>
<td>0.08</td>
<td>0.00</td>
<td>8.14</td>
<td>0.09</td>
<td>0.00</td>
<td>68.24</td>
<td>0.00</td>
</tr>
</tbody>
</table>

1. Numbers takes from http://209.3.158.116/statqwest/statistics.jsp
Voice Quality

Some major tests:

- **Mean Opinion Score (MOS)** - defined in ITU-T P.800
  - ITU test based on using 40 or more people from different ethnic or language backgrounds listening to audio samples of several seconds each
  - **Human listeners** rating the quality from 1 to 5; 5 being perfect, 4 “toll-quality”, …

- **Perceptual Speech Quality Measurement (PSQM)** - ITU-T P.861
  - A computer algorithm - so it is easy to automate
  - scale of 0 to 6.5, with 0 being perfect
  - Designed for testing codecs
  - test tools from Agilent[215], QEmpirix, Finisar, … - cost US$50k and up

- **PSQM+**
  - Developed by Opticom
  - for VoIP testing

- **PESQ (Perceptual Evaluation of Speech Quality)**
  - submitted to ITU-T by Psytechnics, Opticom, and SwissQual
  - 0.95 correlation with human listeners

- **Perceptual Analysis Measurement System (PAMS)**
  - Developed by British Telecommunications ~1998
• ITU-T’s P.563
  • Passive monitoring
  • 0.85 to 0.9 correlation with human listeners
  • ITU standard May 2004
• Psytechnics algorithm: psyvoip
  • passive listening
  • uses RTP statistics
• "E Model" - ITU-T G.107
  • passive monitoring
Rating voice quality in practice

One approach is to occasionally ask IP phone users to indicate how the quality of their call was at the end of the call ⇒ MOS scoring!

Another is exemplified by Susan Knott, global network architecture for PricewaterhoursCoopers:

“But I’ve found that if my vice president of finance can talk to my CIO [over a VoIP connection], and they both say the quality of the connection is OK, then I say that’s good enough.”

QoS Proprietary vs. Standards based

Past

Agere Systems, Inc. VoIP “Phone-On-A-Chip” used a proprietary voice packet prioritization scheme called Ethernet Quality of Service using BlackBurst (EQuB), an algorithm (implemented in hardware) ensures that voice packets are given the highest priority in their collision domain.

2002

Their Phone-On-A-Chip solution now implements a software-based IEEE 802.1q tagging protocol (i.e. Virtual local area network (VLAN) tagging) for outgoing Ethernet frames.¹

QoS for SIP

SDP can be used to convey conditions which must be met:

- direction for QoS support: send, receive, or bidirectional
- along with a “strength” parameter: optional or mandatory

If conditions can be met then a COMET is sent.

See also [228].
VoIP traffic and Congestion Control

RFC 3714: IAB Concerns Regarding Congestion Control for Voice Traffic in the Internet [220] - describes the concerns of the IAB due to the persistence of VoIP clients which continue to send RTP streams **despite** high packet loss rates WRT\(^1\):

- the risks of congestion collapse (along the end-to-end route) and
- fairness for congestion-controlled TCP traffic sharing the links.

When a steady-state packet drop rate >> a specified drop rate the flow should be terminated or suspended. Thus:

- RFC3551: RTP Profile for Audio and Video Conferences with Minimal Control - should be changed to say:
  - “… RTP receivers **SHOULD** MUST monitor packet loss to ensure that the packet loss rate is within acceptable parameters.” and hence “**MUST** detect and respond to a persistent high loss rate”
- CODECs - should adapt so as to reduce congestion

Suggested heuristic: VoIP applications should suspend or terminate when:

- RTCP reported loss rate is greater than 30%, or
- N back-to-back RTCP reports are missing

---

1. With Respect To
Delay and Packet Loss effects

Effect of delay and packet loss on VoIP when using FEC has been studied by many researchers [222], [223], [224], [225].

A rule of thumb: When the packet loss rate exceeds 20%, the audio quality of VoIP is degraded beyond usefulness (cited as [S03] in [220]).

Normally in telephony, when the quality falls below a certain level users give up (i.e., they hang up). Does this occur in the absence of a cost associated with not hanging up?

∴ according to [220]:

\[ \text{if loss rate is persistently unacceptably high relative to the current sending rate} \]
\[ \text{& the best-effort application is unable to lower its sending rate:} \]

⇒ flow must discontinue:

• multicast session ⇒ receiver withdraws from the multicast group
• unicast session ⇒ unicast connection termination
When to continue (try again)

Probabilistic Congestion Control (PCC) [226] based on:

- calculating a probability for the two possible states (on/off) so that the expected average rate of the flow is TCP-friendly
- to perform a random experiment that succeeds with the above probability to determine the new state of the non-adaptable flow, and
- repeat the previous steps frequently to account for changes in network conditions.

The off periods need to be fairly distributed among users and the on period need to be long enough to be useful.

When to try again is determined by: **Probing the network while in the off** state (the authors of [226] have not implemented this yet).

Note that PCC only applies when there is a significant level of statistical multiplexing on the link (otherwise the use of statistics is not meaningful).

Other examples of probe based measurements are described at [227].
More about congestion

D. Willis and B. Campbell in “Session Initiation Protocol Extension to Assure Congestion Safety”, and Internet-Draft, October 13, 2003 examine:

• UAC may require that any proxy processing its requests must transmit those requests over a transport protocol providing congestion management
  • with a "Proxy-Require: congestion-management" header field

• In turn the UAS receiving these requests can be required to respond in similar fashion

• If a proxy finds that it has no route supporting congestion management it may reject the request with a 514 response (“No available route with congestion management”)

• If the request would be fragmented, the proxy can reject it with a 516 response ("Proxying of request would induce fragmentation")

• If the originating request did not require congestion-managed transport, then a UAS may reject a request that would result in a response that requires congestion-managed transport.
VoIP quality over IEEE 802.11b

Two exjobb reports:

Juan Carlos Martín Severiano, “IEEE 802.11b MAC layer’s influence on VoIP quality: Measurements and Analysis”[230]

Victor Yuri Diogo Nunes, “VoIP quality aspects in 802.11b networks” [231]
Gross, et al. proposed the use of an Application Policy Server (APS) [214].

IETF Integrated Services over Specific Lower Layers (issll) Working group ([http://www.ietf.org/html.charters/issll-charter.html](http://www.ietf.org/html.charters/issll-charter.html)) is defining protocols to control the link layer.
References and Further Reading


[211] Telecommunications and Internet Protocol Harmonization Over Networks (TIPHON): Inter-domain pricing, authorisation, and usage exchange; ETSI DTS/TIPHON-03004 V1.4.0 (1998-09).

SIP-based Distributed Call Control Mechanisms”, IETF Internet Draft, June 6, 2002

http://www.ietf.org/internet-drafts/draft-dcsgroup-sipping-arch-00.txt


[217] netIQ’s Vivinet Manager Suite
   http://www.netiq.com/products/vm/default.asp

[218] Cisco’s “Monitoring Voice over IP Quality of Service”

   (IQ-VW)”, Project Research Paper, for CS522 Computer Communications,
   University of Colorado at Colorado Springs, December 2002.
   http://cs.uccs.edu/~cs522/projF2002/msoliman/doc/QoS%20of%20VoIP%20over%20WLAN.doc

   Control for Voice Traffic in the Internet”, IETF, RFC 3714, Network

[221] Sally Floyd and Kevin Fall, “Promoting the use of end-to-end congestion

[222] Wenyu Jiang and Henning Schulzrinne, “Modeling of Packet Loss and
Delay and Their Effect on Real-Time Multimedia Service Quality”, NOSSDAV, 2000. [http://citeseer.nj.nec.com/jiang00modeling.html](http://citeseer.nj.nec.com/jiang00modeling.html)


Module 15: SIP Applications

Lecture notes of G. Q. Maguire Jr.
Documents the use of SIP for several applications related to telephony and multimedia, and develops requirements for any extensions to SIP needed for them.

One of the significant features of using SIP for building applications is that it is much easier to build open, distributed, and scalable services that the traditional method of Intelligent Networks (IN); thus putting services into the hands of user!

The specific tasks for SIPPING will be:

1. PSTN and/or 3G telephony-equivalent applications that need a standardized approach
   - informational guide to common call flows
   - support for T.38 fax
   - requirements from 3GPP for SIP usage
   - framework of SIP for telephony (SIP-T)
   - call transfer and call forwarding
   - AAA application in SIP telephony
• mapping between SIP and ISUP

2 Messaging-like applications of SIP
• support for hearing-/speech-impaired calling
• User Requirements for the Session Initiation Protocol (SIP) in Support of Deaf, Hard of Hearing and Speech-impaired individuals (RFC 3351)
  http://www.ietf.org/rfc/rfc3351.txt
• development of usage guidelines for subscribe-notify (RFC 2848, SIP events) to ensure commonality among applications using them, including SIMPLE WG’s instant messaging.

3 Multi-party applications of SIP

4 SIP calling to media servers
• develop a requirements draft for an approach to SIP interaction with media servers, e.g., whether a voicemail server is just a box that a caller can send an INVITE to.
Application Service Components

Gateways

Application Service Providers
AAA and Policy Server
Interactive Response Voice (IVR)
Web
SIP Server
Service Controller
Presence Server
Transaction Server
Credit card Verification

SIP clients

IP Network, SIP and HTTP Messages, RTP for Media

Text to Speech
Directory and ENUM
Unified Messaging Voicemail storage
Conf Schedule
Media Mixer
Streaming Content Server
DTMF Collector

SIP clients

PSTN

PBX

AAA and Policy Server
Interactive Response Voice (IVR)
Web
SIP Server
Service Controller
Presence Server
Transaction Server
Credit card Verification

SIP clients

Gateways

Application Service Providers
AAA and Policy Server
Interactive Response Voice (IVR)
Web
SIP Server
Service Controller
Presence Server
Transaction Server
Credit card Verification

SIP clients

PSTN

PBX
Advantages

• Decomposition
  • No complex APIs, just HTTP and SIP ⇒ rapid development
  • User can provide input to the service controller via Web servers, DTMF digit collector, voice portal (via VoiceXML), DTMF input, ... ⇒ just about any internet attached device can be used to provide input.
  • Easy to scale
  • New services can combine the "best of the best" (thus allowing developers to specialize)
  • Servers and services can be located anywhere on the internet and operated by anyone

• Decoupling
  • Loosely coupled and distributed
    – Flexible location of servers
    – if properly designed, implemented, and operated ⇒ higher reliability and resilience
  • Separation of businesses (leads to a rich variety of outsourcing, reseller, ... models)
  • Since the functions are highly independent ⇒ rapid development

• Anyone can introduce a new service

However, if you want to use service components of others, then you may need to work out a suitable agreement (which will probably include an agreement about authorization) ⇒ security can be more complex.
Collecting DTMF digits for use within a service

Call flow diagram:

1. INVITE
2. INVITE
3. 200 OK
4. ACK
5. INVITE
6. 200 OK
7. 200 OK
8. ACK
9. SIP ACK
10. INVITE
11. 200 OK
13. ACK
14. RTP DTMF digits
15. HTTP GET
16. HTTP OK

Legend:
- RTP media session
- DTMF collector
Response “3. 200 OK” looks like:

SIP/2.0 200 OK
Via: SIP/2.0/UDP 100.101.102.103
To: User A <sip:UserA@here.com>
From: UserB <sip:UserB@there.com>
Call-ID: a84b4c76e66710100.101.102.103
CSeq: 1 INVITE
Contact: <sip:UserB@there.com>
Content-Type: application/sdp
Content-Length: ...

v=0
o=UserA 289375749 289375749 IN IP5 110.111.112.113
S=-
c=IN IP4 110.111.112.113
t=0 0
m=audio 5004 RTP/AVP 0
Controller issues a “re-Invite” at 11 which looks like:

```
INVITE sip:UserB@there.com SIP/2.0
Via: SIP/2.0/UDP 100.101.102.103
To: UserB <sip:UserB@there.com>
From: User A <sip:UserA@here.com>
Call-ID: a84b4c76e66710100.101.102.103
CSeq: 1 INVITE
Contact: <sip:UserB@there.com>
Content-Type: application/sdp
Content-Length: ...
```

```
v=0
o=UserA 289375749 289375749 IN IP5 100.101.102.103
S=
  c=IN IP4 100.101.102.103
  t=0 0
m=audio 5004 RTP/AVP 0
m=audio 53000 RTP/AVP 0
  c=IN IP4 200.201.202.203
a=rtpmap:96 telephone-event
```

Note the 2nd “m=audio” line in the SDP (see Sinnreich Johnston page 257), this second connection is the RTP connection to the DTMF digit collector.
The service controller proxies the caller to the IVR system.
Managing Services

Avgeropoulos Konstantinos in “Service Policy Management for User-Centric Services in Heterogeneous Mobile Networks”[234] proposes the use of SIP as signaling protocol for policy based management of a user’s multiple UAs. He proposes a new SIP entity, called the SIP Service Manager (SSM).
Lots more services

<< more to be added here - as time permits >>
References and Further Reading

SIPPING


   http://www.ietf.org/rfc/rfc3263.txt

Module 16: More than Voice
Non-voice Services and IP Phones

Phone Services: built using scripts which the IP phone executes to acquire information and display it

For example, some of the Cisco IP telephones (7940 and 7960) have a web browser which understands XML and a 133x65 pixel-based LCD display to display output.

Sample services:
- Conference room scheduler
- E-mail and voice-mail messages list
- Daily and weekly schedule and appointments
- Personal address book entries (⇒ any phone can become “your” phone)
- Weather reports, Stock information, Company news, Flight status, Transit schedules, …
- Viewing images from remote camera (for security, for a remote receptionist, …)
XML objects include: CiscoIPPhoneMenu, CiscoIPPhoneText, CiscoIPPhoneInput, CiscoIPPhoneDirectory, CiscoIPPhoneImage, CiscoIPPhoneGraphicMenu, CiscoIPPhoneIconMenu, CiscoIPPhoneExecute, CiscoIPPhoneError and CiscoIPPhoneResponse.

Cisco IP Phone Services Software Developer’s Kit:
Invoking RTP streams

On the Cisco phones it is possible to invoke RTP streaming (transmit or receive) via URIs in above services. RTP information for the stream types must be of the form:

<table>
<thead>
<tr>
<th>CODEC</th>
<th>G.711 mu-Law</th>
</tr>
</thead>
<tbody>
<tr>
<td>Packet size</td>
<td>20 ms</td>
</tr>
</tbody>
</table>
see ‘Thinking Outside the “Talk” Box: Building Productivity-Boosting Applications for Your Cisco IP Phones’ by Anne Smith, Cisco Packet, Third Quarter, 2002, pp. 21-23


The book includes a CD which has a CallManager Simulator - so you can write applications with just a web server and a Cisco IP phone.

You can download the SDK, etc. from:

http://cisco.com/warp/public/570/avvid/voice_ip/cm_xml/cm_xmldown.shtml
Services for sale - building a market

Purchase existing services or contract for new third party XML services or support for Cisco’s IP Telephony products: HotDispatch

http://www.hotdispatch.com/cisco-ip-telephony

They have 91 Existing products as of 20 October 2002

HotDispatch has partnered with Cisco to provide IP Telephony marketplace
Figure 17: Using SIP for Service Portability (adapted from figure 2 of Moyer, Maples, Tsang, and Ghosh)

Proposed Extension of SIP

Add **DO** message type

Adding a new optional header: History-Info

- provides information as to how and why a call arrives at a specific application or user [245]

Build upon Event extensions (specifically **SUBSCRIBE** and **NOTIFY**)

- For example, you can subscribe to know when a user is invited to a session or there is a change in a state of an INVITE initiated dialog [244]

Add a new payload type via the new MIME type: **Device Message Protocol** (DMP) -- this payload is translated into device specific payload at the SIP User Agent.

Note that you could also send **SOAP** payload either separately or as part of DMP.
Service Location Protocol (SLP) URL

To: [SLP:/d=lamp, r=office, u=maguire]@it.kth.se

Note that the information inside the [ ] can be encoded in BASE-64 and encrypted, those making it opaque to entities outside the domain.

See [241].
Example service

This example is adapted from the above article, and the service is a network-based alarm clock service:

• delivers user specific information (latest news, weather, etc.)
• at a user selected time
• to the user’s “alarm clock” network appliance

Specifically:

1. REGISTER register@home.net
   To: [sdp:/d=alarmclock, r=bedroom, u=maguire]@ua.chips.home.net
   From: [sdp:/d=alarmclock, r=bedroom, u=maguire]@ua.chips.home.net
   Content-type: application/ddp
   [Device address]

2. INVITE sip:[sdp:/d=alarmclock, r=bedroom, u=maguire]@home.net SIP/2.0
   From: sip:announcement@alarmclock.net
   To: [sdp:/d=alarmclock, r=bedroom, u=maguire]@ua.chips.home.net
   Via: alarmclock.net
   Content-type: application/sdp
   [SDP for uni-directional RTP stream]
3. INVITE sip:[slp:/d=alarmclock, r=bedroom, u=maguire]@home.net SIP/2.0
   From: sip:announcement@alarmclock.net
   To: [slp:/d=alarmclock, r=bedroom, u=maguire]@ua.chips.home.net
   Via: home.net
   Via: alarmclock.net
   Content-type: application/sdp
   [SDP for uni-directional RTP stream]

4. INVITE sip:[slp:/d=alarmclock, r=bedroom, u=maguire]@home.net SIP/2.0
   From: sip:announcement@alarmclock.net
   To: [slp:/d=alarmclock, r=bedroom, u=maguire]@ua.chips.home.net
   Via: chips.home.net
   Via: home.net
   Via: alarmclock.net
   Content-type: application/sdp
   [SDP for uni-directional RTP stream]

5. The alarm clock responds with its RTP parameters and the RTP session plays the announcement to the user via the “alarm clock” network appliance
Example of service portability

This example is adapted from the above article, Chip visits his friend Mark:

- delivers user specific information (latest news, weather, etc.)
- at a user selected time
- to the user’s “alarm clock” network appliance
- But the service now has to be delivered to the correct “alarm clock”
  - Either Chip takes his alarm clock with him or
  - Utilizes Mark’s guest alarm clock as his alarm clock

1. REGISTER register@home.net
   To: [slp:/d=alarmclock, r=bedroom, u=maguire]@ua.chips.home.net
   From: [slp:/d=alarmclock, r=bedroom, u=maguire]@ua.chips.home.net
   Contact: *; expires=0

   The above cancels the service to Chip’s home alarm clock

2. REGISTER register@home.net
   To: [slp:/d=alarmclock, r=bedroom, u=maguire]@ua.chips.home.net
   From: [slp:/d=alarmclock, r=bedroom, u=maguire]@ua.chips.home.net
   Contact: sip:[slp:/d=alarmclock, r=guest_bedroom, u=maguire]@ua.marks.home.net
   Content-type: application/ddp
   [Device description (including address)]
3. INVITE sip: [slp:/d=alarmclock, r=bedroom, u=maguire]@home.net SIP/2.0
   From: sip:announcement@alarmclock.net
   To: [slp:/d=alarmclock, r=bedroom, u=maguire]@us.chips.home.net
   Via: alarmclock.net
   Content-type: application/sdp
   [SDP for uni-directional RTP stream]

   Now the SIP Proxy at home.net looks up
   [slp:/d=alarmclock, r=bedroom, u=maguire]@home.net and determines that it is
   [slp:/d=alarmclock, r=guest_bedroom, u=maguire]@ua.marks.home.net] so it
   forwards the messages to the SIP proxy at marks.home.net

4. INVITE sip: [slp:/d=alarmclock, r=guest_bedroom,
   u=maguire]@ua.marks.home.net] SIP/2.0
   From: sip:announcement@alarmclock.net
   To: [slp:/d=alarmclock, r=bedroom, u=maguire]@ua.chips.home.net
   Via: home.net
   Via: alarmclock.net
   Content-type: application/sdp
   [SDP for uni-directional RTP stream]

5. INVITE sip: [slp:/d=alarmclock, r=guest_bedroom,
   u=maguire]@ua.marks.home.net] SIP/2.0
   From: sip:announcement@alarmclock.net
   To: [slp:/d=alarmclock, r=bedroom, u=maguire]@ua.chips.home.net
   Via: marks.home.net
   Via: home.net
   Via: alarmclock.net
Content-type: application/sdp
[ SDP for uni-directional RTP stream ]

6. Mark’s guest bedroom alarm clock responds with its RTP parameters and the RTP session plays the announcement to the user via the “alarm clock” network appliance.
RTP can be used to carry real-time text conversations, the contents use ITU-T Recommendation T.140.[243]

Timed Text

The 3rd Generation Partnership Project (3GPP) has defined "timed text" as “time-lined, decorated text media format with defined storage in a 3GP file”[242].

“Timed Text can be synchronized with audio/video contents and used in applications such as captioning, titling, and multimedia presentations.”[242]
References and Further Reading

Phone Services


Network Appliances


[238] Open Services Gateway Initiative (OSGi), [http://www.osgi.org](http://www.osgi.org)


[245] M. Barnes (Editor), An Extension to the Session Initiation Protocol (SIP) for Request History Information, IETF, RFC 4244, November 2005

Module 17: VOCAL

Lecture notes of G. Q. Maguire Jr.

Figure 18: VOCAL: Simplified overview

MGCP = Media Gateway Control Protocol
VOCAL Servers

- Marshal server (MS)
  - User Agent (UA) Marshal server
    - interface to/from IP phones connected to this network
    - can do different types of authentication on a per-user basis
  - (PSTN) Gateway Marshal servers
    - provides interworking with PSTN
  - Internet Marshal server
    - interface to/from a SIP proxy server on another IP network
    - authenticate calls via Open Settlement Protocol (OSP)
    - can request QoS via Common Open Policy Service (COPS)
  - Conference Bridge Marshal server
    - interface to/from third party conference servers

- Feature server (FS) - to provide advanced telephony services
- Redirect server (RS) - keep track of registered users and provide routing to/from them
- Provisioning server (PS) - for configuration
- Call Detail Record (CDR) server - stores start/end information about calls for billing and other purposes
Scaling of a VOCAL system

From table 3-1 of *Practical VoIP: Using VOCAL*

<table>
<thead>
<tr>
<th>Server types</th>
<th>6-host system</th>
<th>14-host system</th>
<th>26-host system</th>
</tr>
</thead>
<tbody>
<tr>
<td>Redirect servers</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Feature servers</td>
<td>1</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Marshal servers</td>
<td>2</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>Call Detail Record servers</td>
<td>1/2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Provisioning servers</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Policy servers</td>
<td>1/2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Total number of hosts</td>
<td>6</td>
<td>14</td>
<td>26</td>
</tr>
<tr>
<td>Capacity in calls per second</td>
<td>35</td>
<td>70</td>
<td>175</td>
</tr>
<tr>
<td>Capacity in busy-hour call attempts (BHCA)</td>
<td>125,000</td>
<td>250,000</td>
<td>630,000</td>
</tr>
</tbody>
</table>

*Each host is a 700MHz Pentium III with 512 MB or RAM.*

- Note that unlike a PBX or Public Exchange, the capacity in calls per second (or BHCA) is independent of the call durations, since the call traffic is carried directly between the endpoints via RTP and does not use the VOCAL system!
For comparison with a PBX

- NEC’s PBX: EAX2400 IMX - Integrated Multimedia eXchange, model ICS IMGdxh uses a Pentium control process and the claimed BHCA is 25,600.
- Tekelec’s softswitch "VXiTM Media Gateway Controller" claims a capacity which scales from 250,000 to over 1 million BHCA - a Class 5 exchange.
- Lucent’s 5E-XC™ Switch High Capacity Switch - supports 4 million BHCA, 250K trunks, and 99.9999% availability [249]
- Frank D. Ohrtman Jr. says that a Class 4 Softswitch should handle 800,000 BHCA, support 100,000 DS0s (i.e., 100K 64 bps channels), with a reliability of 99.999%, and MOS of 4.0 (i.e., high quality voice)[246].
  - His pricing data shows that softswitches are about 1/4 the price per DS0 of Class 4 exchanges (e.g., Nortel DMS250 and Lucent 4ESS vs. Convergent Networks’s ICS2000 and SONUS GSX9000) -- additionally the softswitches are physically much smaller.
  - Many claim that softswitch and VoIP reliability already exceeds that of central office exchanges; because with VoIP it is cheaper to implement redundancy and easier to build physically distributed systems; plus more features {sooner}, while also providing potentially better quality (i.e., better than "toll" quality)!

Radcom’s MegaSIP test software generates 3,500,000 BHCA calls per server.

---

1. Was available from http://www.stfi.com/STF_part3e.html
2. "A softswitch is the intelligence in a network that coordinates call control, signaling, and features that make a call across a network or multiple networks possible."[246]
Marshal server (MS)

A SIP proxy server which provides:

• authentication of users
• generates call detail records (CDRs)
• provides an entry point for SIP messages into the VOCAL system
  • thus the other elements of the VOCAL system don’t need to authenticate each message
• monitor heart beats - can use this for load balancing across RSs
• SIP transaction stateful, but not call (dialog) stateful

Allows better scaling, since these servers can be replicated as needed; while allowing the redirect server to focus just on keeping registration information.
Redirect Server (RS)

- receives SIP `REGISTER` messages from User Agents (UAs)
- keeps track of registered users and their locations (i.e., registrations)
- provides routing information for SIP `INVITE` messages
  - based on caller, callee, and registration information (for either or both parties)
  - based on where the `INVITE` message has already been
- Supports redundancy
  - Utilizes multicast heartbeat
    - starts by listening for 2s for another RS
    - if found, then it synchronizes with this RS and will act as a redundant backup RS (following synchronization)
    - if not found, then it starts transmitting its own heartbeat
  - a given RS must mirror `REGISTER` messages (received from the MS) to the other RSs
Feature Server (FS)

- Implements Call Forward, Call Screening, Call Blocking
  - The “Core Features” are implemented “within the network”
    - for example, you can’t implement features in a phone which is not there!
    - you can’t give an end system the caller’s ID, but guarantee that they don’t display it, …
- Execute arbitrary Call Processing Language (CPL) scripts written by users
  - CPL is parsed into eXtensible Markup Language (XML) document object model (DOM) trees, these are then turned into state machines (in C++), then executed.
Residential Gateway (RG)

A residential gateway (RG) provides “… Internet access throughout the home and remote management of common household appliances such as lights, security systems, utility meters, air conditioners, and entertainment systems.”

Open Services Gateway Initiative (OSGi™) Alliance http://www.osgi.org/ is attempting to define a standard framework and API for network delivery of managed services to local networks and devices.

An alternative to using a residential gateway to attach analog phones are devices such as the Cisco Analog Telephone Adaptor (ATA) 186 [250].

In VOCAL: “SIP Residential Gateway is an IP Telephony gateway based on SIP which allows a SIP user agent to make/receive SIP call to/from the Public Switched Telephone Network (PSTN).”

1. from http://www.national.com/appinfo/solutions/0,2062,974,00.html - “National Semiconductor signed a definitive agreement in August 2003 to sell its Information Appliance (IA) business unit, consisting primarily of the Geode™ family of microprocessor products, to Advanced Micro Devices (AMD)”

References and Further Reading


[248] Abacus2 Test equipment, Spirent Communications, Calabasas, CA, USA, www.spirentcom.com - generates and switches more than 20 million calls per hour


SIP Express Router (SER)

http://www.iptel.org/ser/

An open-source implementation which can act as SIP registrar, proxy or redirect server. SER features:

- an application-server interface,
- presence support,
- SMS gateway,
- SIMPLE2Jabber gateway,
- RADIUS/syslog accounting and authorization,
- server status monitoring,
- Firewall Communication Protocol (FCP)\(^1\) security, …
- Web-based user provisioning (serweb)

For configuration help see: http://www.mit.edu/afs/athena/project/sip/sip.edu/ser.shtml

\(^1\) http://www.iptel.org/fcp/
**SipFoundry**

http://www.sipfoundry.org/

Formed on March 29, 2004 - goal improving and adopting open source projects related to SIP

Pingtel Corp. contributed their sipX family of projects (distributed under the LGPL). This includes:

<table>
<thead>
<tr>
<th>Project</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>sipXphone</td>
<td>SIP soft phone</td>
</tr>
<tr>
<td>sipXproxy</td>
<td>pair of applications that together form a configurable SIP router.</td>
</tr>
<tr>
<td>sipXregistry</td>
<td>SIP Registry/Redirect server</td>
</tr>
<tr>
<td>sipXpublisher</td>
<td>server to handle SIP SUBSCRIBE/NOTIFY handling + flexible plugin architecture for different event types.</td>
</tr>
<tr>
<td>sipXvxml</td>
<td>VXML scripting engine supporting creation of IVR and other VXML applications (including auto-attendant and voice mail)</td>
</tr>
<tr>
<td>sipXconfig</td>
<td>SIP configuration server</td>
</tr>
<tr>
<td>sipXpbx</td>
<td>full PBX solution; combining sipXproxy, sipXregistry, sipXpublisher, sipXvxml, and sipXconfig</td>
</tr>
<tr>
<td>sipXtest</td>
<td>testing tools and frameworks</td>
</tr>
</tbody>
</table>
Other SIP Proxies

- **JAIN-SIP Proxy**
  - JAIN-SIP proxy, JAIN-SIP IM client, SIP communicator, SIP trace viewer, JAIN-SIP gateway, JAIN-SIP 3PCC, ...

- **SaRP SIP and RTP Proxy**
  - written in Perl

- **sipd SIP Proxy**

- **Siproxd SIP and RTP Proxy**
  - [http://sourceforge.net/projects/siproxd/](http://sourceforge.net/projects/siproxd/)
  - an proxy/masquerading daemon for the SIP protocol

- **partysip**
  - [http://www.nongnu.org/partysip/partysip.html](http://www.nongnu.org/partysip/partysip.html)
  - Has a plugin to enable use of SCTP transport

- **Yxa: Written in the Erlang programming language**
  - [http://www.stacken.kth.se/projekt/yxa/](http://www.stacken.kth.se/projekt/yxa/)

...
SIP Tools

• Callflow
  - http://callflow.sourceforge.net/
  - Generates SIP call flow diagrams based on an ethereal capture file

• SIPbomber
  - A SIP proxy testing tool for server implementations (i.e., proxies, user agent servers, redirect servers, and registrars)

• Sipsak
  - http://sipsak.berlios.de/
  - Sipsak a command line tool for developers and administrators of SIP applications

• PROTOS Test-Suite
  - http://www.ee.oulu.fi/research/ouspg/protos/testing/c07/sip/
  - SIP Testing tools from the "PROTOS - Security Testing of Protocol Implementations" project
SIP Clients

- **kphone**
  - [http://www.wirlab.net/kphone/](http://www.wirlab.net/kphone/)
  - IPv4 and IPv6 UA for Linux, also supports Presence and Instant Messaging
  - UA for Linux - for KDE

- **Linphone**
  - [http://www.linphone.org/?lang=us&rubrique=1](http://www.linphone.org/?lang=us&rubrique=1)
  - UA for Linux - for GNOME

- **Xten’s X-Lite - free demo version for Windows and Linux**
  - [http://www.xten.net/](http://www.xten.net/)
  - The also have a "Business-class SIP Softphone" called X-PRO

- **Xten’s ineen** [http://www.ineen.com/index.html](http://www.ineen.com/index.html)
  - Instant Messenger & Buddy List
  - Audio Conferencing [10 People]
  - Distributed Conferencing [Infinite]
  - Video Conferencing [4 People]
  - Call Recording [Audio and Video]
  - Speakerphone Mode
  - Call Transfer
  - Call Logs & Missed Call Indicator
  - Automatic Answer
  - Automatic Conference
• Do Not Disturb
• Open Standards Compliant

• Zultys Technologies SIP Soft Phone
  • http://www.lipz4.com/lipz4.htm
  • for linux

• …
CPL and Ontology extensions to SER

See:


References and Further Reading


[252] Amos Nungu, VoIP Service Provider (Internet Telephony Service Provider using SIP Protocol), Masters thesis, School of Information and Communication Technology, Royal Institute of Technology (KTH), April 2005
Skype

Skype™ Technologies http://www.skype.com/

- “Skype is free Internet telephony that just works.”
- 67,430,762 downloads as of 2005.02.10
- 4,707,596,653 minutes served as of 2005.02.10