



现代IP技术丛书

MODERN IP TECHNOLOGY

# 下一代 网络核心控制协议 ——SIP及其应用

黄永峰 李建庆 等 编著

 人民邮电出版社  
POSTS & TELECOM PRESS

现代 IP 技术丛书

# 下一代网络核心控制协议 ——SIP 及其应用

黄永峰 李建庆 等 编著

人民邮电出版社  
北京

## 图书在版编目 (C I P) 数据

下一代网络核心控制协议: SIP 及其应用 / 黄永峰等编  
著. —北京: 人民邮电出版社, 2009. 1  
(现代IP技术丛书)  
ISBN 978-7-115-18965-3

I. 下… II. 黄… III. 计算机网络—通信协议 IV.  
TN915.04

中国版本图书馆CIP数据核字 (2008) 第155708号

## 内 容 提 要

本书以 SIP 的基本标准 RFC 3261 为蓝本, 根据作者多年来对 SIP 的研究经验和积累, 系统介绍了 SIP 的基本概念、协议结构以及 SIP 网络实体的行为规则, 详细论述了常见 SIP 协议栈和应用软件, 包括 SIP 协议的事务和传输、SIP 系统的安全、SIP 协议栈的分析以及 SIP 扩展技术, 最后还介绍了 SIP 在 VoIP 中的应用、SIP 和 P2P 融合技术等最新研究动态。

本书内容翔实, 贴近实际开发, 可供从事互联网多媒体通信的工程技术人员和研究开发人员阅读, 也可作为信息专业的本科生、研究生和教师的教学参考书。

现代 IP 技术丛书

### 下一代网络核心控制协议——SIP 及其应用

- 
- ◆ 编 著 黄永峰 李建庆 等  
责任编辑 陈万寿
  - ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街 14 号  
邮编 100061 电子函件 315@ptpress.com.cn  
网址 <http://www.ptpress.com.cn>  
北京楠萍印刷有限公司印刷
  - ◆ 开本: 787×1092 1/16  
印张: 9.5  
字数: 231 千字  
印数: 1—3 000 册
- 2009 年 1 月第 1 版  
2009 年 1 月北京第 1 次印刷

---

ISBN 978-7-115-18965-3/TN

定价: 25.00 元

读者服务热线: (010)67120142 印装质量热线: (010)67129223  
反盗版热线: (010)67171154

# 前 言

会话初始协议 (Session Initiation Protocol, SIP) 是 2001 年推出的 IETF 标准 (RFC 3261), 用于 IP 网络上建立、改变和结束多媒体会话。随着 SIP 标准的扩展, SIP 可以实现多种业务, 如 IP 电话、多媒体会议、远程白板、即时消息等, 并可以简化对 IPsec VPN 的连接。SIP 还可以方便地将多媒体服务与其他类型的服务组合在一起, 这些服务组如目录信息、Web 浏览、定位和列席服务、寻找和跟踪等。由于 SIP 是应用层协议, 它是相对独立的, 可以连接 IP 网络上任何类型的参与者。例如, SIP 可以提供固定和移动网络之间的无缝服务功能, 这是实现固定与移动业务融合的关键要素, 而且, 这种功能将随着 3G 移动网络在世界各地的部署而变得越来越重要。

正因为 SIP 应用具有巨大潜力, SIP 的研究与开发得到了极大的重视, 基于 SIP 的应用产品如雨后春笋。国际上许多研究单位、大学以及一些著名公司。例如 MIT、Stanford、Cisco、Nortel、Siemens、Alcatel、Broadsoft、DynamicSoft 和 UTStarcom 等, 投入大量的资金和人力进行 SIP 开发和研究。目前出现了几家国际上具有相当影响的基于 SIP 的提供信息传输服务的运营商, 如 Delta 3 等, 还出现了一些风靡全球的 SIP 应用软件, 如 SIP Phone 等。在我国, 由于各种因素, SIP 应用的发展要落后于 H.323 协议。但是, 目前国内专业人士也意识到 SIP 在全球的发展浪潮以及 SIP 应用的巨大商机, 正加紧对 SIP 的研究和开发, 近年来形成了国内的 SIP 研究热。我国也正在制定自己的 SIP 标准。

SIP 已被公认是下一代网络的核心控制协议, 其功能得到不断扩展, 在 SIP 基本标准文档 RFC 3261 的基础上, IETF 推出了许多 SIP 扩展标准, 例如 RFC 3262 (规定了临时响应的可靠性)、RFC 3263 (确定了 Proxy 的定位规则)、RFC 3264 (提出了提议和应答模型)、RFC 3265 (提出了事件通知规则) 以及 RFC 3248 (提出了即时消息规则) 等。人们正在研究 SIP 在下一代网络中的进一步应用, 例如, 网络家电、IMS、多人游戏, 等等。

因此, MP3 的创始者 Michael Robertson 认为, SIP 技术出现的意义不亚于 MP3 的发明, MP3 引发了音乐的数字化, 而 SIP 所带来的就是音、视频信息数字化。

因此, 作者根据自己多年来对 SIP 的研究和开发经验, 历经两年的思索, 撰写了本书。作者在组织该书时, 注意突出如下特点。

(1) 所覆盖的内容尽量广泛。本书除了包含 SIP 最基本的内容 (RFC 3261 规范) 外, 还将结合 SIP 最新的研究成果, 如 SIP P2P 技术等。

(2) 面向开发应用。本书将从开发的角度, 结合一些实际案例来剖析 SIP 技术的精髓, 主要目标是指导 SIP 的应用开发, 因此, 实用性更强。

(3) 内容组织方式层次化。将按照 SIP 技术的知识层次结构, 结合实际案例来编写, 可读性增强。

本书的编写分工如下：黄永峰编著第 1、6 章，陈桂生编著第 4、5 章，李建庆编著第 3、7、8 章，冯文峰编著第 9 章，田晖编著第 2 章和附录；另外，全书由李建庆和黄永峰负责组织和统稿。

本书是澳门基金会 MF-MUST-2007-52 科研项目的主要成果。另外，本书部分成果也源自于国家“973”重大基础理论科研项目（2007CB310806）。在此一并表示感谢。

由于时间仓促，作者水平有限，本书一定有不尽人意之处，欢迎同行批评指正。

作者

# 目 录

<b>第 1 章 概述</b> ..... 1	<b>3.6 proxy 路由过程总结</b> ..... 34
1.1 H.323 和 SIP 的比较..... 1	3.6.1 基本的 SIP 四边传送路由 实例..... 34
1.2 SIP 简介..... 2	3.6.2 穿越严格路由 proxy 的路由 实例..... 35
1.2.1 SIP 基本功能概述..... 3	<b>3.7 SIP 会话建立过程实例分析</b> ..... 36
1.2.2 SIP 的基本功能实体..... 4	<b>第 4 章 SIP 协议的事务和传输</b> ..... 42
1.2.3 SIP 呼叫建立的基本过程..... 5	4.1 客户端事务..... 42
<b>第 2 章 SIP 的基本概念及协议结构</b> ..... 8	4.1.1 INVITE 客户事务..... 43
2.1 SIP 中的基本概念..... 8	4.1.2 构造 ACK 请求..... 44
2.1.1 SIP 中定义的网络实体..... 8	4.1.3 非 INVITE 客户端事务..... 45
2.1.2 SIP 消息中定义的概念..... 9	4.2 服务器事务..... 47
2.2 SIP 的协议结构..... 11	4.2.1 INVITE 服务器事务..... 47
2.3 SIP 消息类型和结构..... 12	4.2.2 非 INVITE 服务器事务..... 48
2.3.1 请求..... 12	4.2.3 服务器事务匹配请求..... 49
2.3.2 应答..... 13	4.3 传输 (transport) ..... 50
2.3.3 头域..... 13	4.3.1 客户端的传输处理..... 50
2.3.4 消息体..... 15	4.3.2 服务器的传输处理..... 51
<b>第 3 章 SIP 网络实体的行为规则</b> ..... 17	<b>第 5 章 SIP 系统的安全</b> ..... 53
3.1 用户代理的行为规则 ..... 17	5.1 SIP 系统的安全威胁模式..... 53
3.1.1 UAC 行为..... 17	5.2 SIP 的安全机制及实现..... 56
3.1.2 UAS 行为..... 21	5.2.1 传输层和网络层的安全机制..... 56
3.2 重定向服务器的行为规则 ..... 24	5.2.2 SIPS URI 方案..... 57
3.3 注册服务器的行为规则 ..... 25	5.2.3 HTTP 认证和 S/MIME 的 消息加密..... 57
3.4 有状态代理服务器的行为规则 ..... 27	5.2.4 SIP 安全机制的实现..... 58
3.4.1 有状态 proxy 的事务..... 28	5.3 SIP 的认证机制..... 61
3.4.2 验证请求..... 28	5.3.1 HTTP Digest 认证..... 61
3.4.3 路由信息的处理..... 29	5.3.2 S/MIME 认证..... 64
3.4.4 请求转发..... 30	5.3.3 SIP 隧道实现头部隐私完整性..... 67
3.4.5 应答的处理..... 32	
3.5 无状态 proxy 的行为规则..... 33	

5.4 SIP 安全机制的应用实例 .....	71		
5.4.1 注册 .....	72		
5.4.2 建立会话 .....	72		
<b>第 6 章 SIP 协议栈的分析</b> .....	78		
6.1 常见 SIP 协议栈简介 .....	78		
6.2 oSIP 协议栈分析 .....	80		
6.2.1 oSIP 的结构 .....	80		
6.2.2 oSIP 的应用结构 .....	84		
6.2.3 oSIP 的安全性问题 .....	85		
6.3 SER 服务器分析 .....	87		
6.3.1 SER 的配置文件 .....	88		
6.3.2 SER 服务器的模块分析 .....	88		
6.3.3 SER 服务器的启动 .....	93		
<b>第 7 章 SIP 扩展</b> .....	96		
7.1 传送 ISUP 信令的扩展 .....	96		
7.1.1 SIP-T 介绍 .....	96		
7.1.2 SIP-I 介绍 .....	96		
7.2 穿越防火墙 .....	97		
7.2.1 穿越防火墙的解决方案 .....	97		
7.2.2 穿越 NAT 的解决方案 .....	98		
7.3 INFO 扩展 .....	99		
7.3.1 INFO 方法 .....	100		
7.3.2 INFO 消息体 .....	100		
7.3.3 利用 INFO 扩展的指导方针 .....	100		
7.4 SIMPLE .....	101		
7.4.1 即时消息 .....	101		
7.4.2 呈现服务 .....	101		
7.5 第三方呼叫控制及会话传递 .....	102		
7.6 SIP 与资源管理 .....	103		
7.7 SIP 会议系统 .....	103		
		<b>第 8 章 SIP 在 VoIP 中的应用</b> .....	105
		8.1 SIP VoIP 系统的实现 .....	105
		8.1.1 VoIP 基本原理 .....	105
		8.1.2 VoIP 关键技术 .....	105
		8.1.3 SIP 的 VoIP 系统及其 业务环境 .....	107
		8.2 软交换与 SIP VoIP 系统互通 .....	108
		8.2.1 互通方式 .....	109
		8.2.2 呼叫控制流程 .....	109
		8.2.3 基于 SIP 的软交换设备间 互通 .....	113
		<b>第 9 章 SIP 和 P2P 融合技术</b> .....	116
		9.1 P2P-SIP 的概念结构 .....	117
		9.1.1 P2P-SIP 覆盖网络概述 .....	108
		9.1.2 P2P-SIP 覆盖网络的概念 定义 .....	108
		9.1.3 P2P-SIP 覆盖网络例子 .....	120
		9.1.4 P2P-SIP 覆盖网络的分布式数 据库功能 .....	121
		9.1.5 如何使用分布式数据库 功能 .....	122
		9.1.6 P2P-SIP 覆盖网络中的 NAT 穿越问题 .....	123
		9.2 P2P-SIP 的两种体系结构 .....	124
		<b>附录 A SIP 消息的头域列表</b> .....	128
		<b>附录 B SIP 的应答代码</b> .....	139
		<b>参考文献</b> .....	145

# 第 1 章 概述

在传统电话系统中，一次通话从建立连接到拆除连接都需要一定的信令配合完成。信令在通信网中是个很重要的概念。所谓信令就是一种机制，通过这种机制，构成通信网的用户终端以及各个业务节点可以互相交换各自的状态信息以及提出对其他设备的接续要求，从而使网络作为一个整体运行。信令协议是有关信令传递和处理的标准协议和规范。

目前，电话通信系统的信令协议可以分为两类：随路信令和公共信道信令。其中，随路信令是指信令和话音在同一条话路中传送的信令方式。目前我国采用的随路信令称为中国一号信令系统。公共信道信令是以时分方式在一条高速数据链路上传送一群话路信令的信令方式，通常用于局间。目前我国采用的公共信道信令就是中国七号信令。七号信令的特点是：信令速度快，具有提供大量信令的潜力，具有改变和增加信令的灵活性，便于开放新业务，在通话时可以随意处理信令，成本低。七号信令方式采用分组交换原理，将交换机之间的信令表示成消息的形式，把每一个信令消息都作为一个分组（消息信号单元）在信令点之间传送。为保证传送的可靠性，分组中除包含消息本身外，还包括传送控制字段和检错校验字段。因此，可以将通信网络中的信令系统比作通信网的神经系统，它是通信网必不可少的组成部分。

近年来，随着互联网技术飞速发展和应用的日益普及，在 IP 上传送语音（Voice over IP, VoIP）应运而生，并得到了普遍应用。VoIP 作为语音通信网络，和传统的电话网络一样，同样需要信令系统的支撑。国际上相关标准组织，先后制定了一些基于 IP 网络的多媒体通信信令协议，目前比较有影响力的标准有 ITU-T 提出的 H.323 协议和 IETF 提出的 SIP 等。

## 1.1 H.323 和 SIP 的比较

H.323 和 SIP 分别是电信领域与互联网领域两大阵营推出的标准。因此，这两个标准协议充分体现 ITU-T 和 IETF 两个不同标准组织之间完全不同的设计理念，两者的设计风格截然不同。下面对它们异同之处进行一些初略分析。

(1) H.323 试图将 VoIP 设计为已经众所周知的传统电话使用模式，与传统电话相比只是传输方式发生了改变，由电路交换变成了分组交换。而 SIP 侧重于将 VoIP 作为互联网上的一个应用，和其他应用（如 FTP、E-mail 等）相比只是增加了信令和 QoS 的要求。但是，H.323 和 SIP 支持的业务集基本相同，都是利用 RTP 作为媒体传输协议。

(2) H.323 采用基于 ASN.1 和压缩编码规则的二进制方法表示其信令消息。ASN.1 通常需要特殊的代码生成器来进行词法和语法分析，因此 H.323 是一个相对复杂的协议。而 SIP



是基于文本的协议，类似于超文本传输协议（Hypertext Transfer Protocol, HTTP）。基于文本的编码意味着头域的含义是一目了然的，如 From、To、Subject 等域名。这种分布式、几乎不需要复杂文档说明的标准规范风格，其优越性在过去的实践中得到了充分的证明（现在广为流行的邮件协议 SMTP 就是一个很好的例子）。SIP 的消息体部分采用会话描述协议（Session Description Protocol, SDP）进行描述，SDP 中的每一项的格式为<type>=<value>，也比较简单。

(3) 在支持会议电话方面，H.323 由多点控制单元（Multipoint Control Unit, MCU）集中执行会议控制功能，所有参加会议的终端都向 MCU 发送控制消息。MCU 可能会成为瓶颈，特别是对于具有附加特性的大型会议。H.323 不支持信令的组播功能，其单播功能限制了可扩展性，降低了可靠性。而 SIP 设计采用分布式的呼叫模型，具有分布式的组播功能，其组播功能不仅便于会议控制，而且简化了用户定位、群组邀请等，能节约带宽。H.323 的集中控制便于计费，对带宽的管理也比较简单、有效。

(4) H.323 中定义了专门的协议用于补充业务，如 H.450.1、H.450.2 和 H.450.3 等。SIP 并未定义专门的协议用于此目的，但它能很方便地支持补充业务或智能业务。只要充分利用 SIP 已定义的头域（如 Contact 头域），并对 SIP 进行简单的扩展（如增加几个域），就可以实现这些业务。例如对于呼叫转移，只要在 BYE 请求消息中添加 Contact 头域，加入意欲转至的第三方地址就可以实现此业务。对于通过扩展头域较难实现的一些智能业务，可在体系结构中增加业务代理，提供一些补充服务或智能网设备的接口。

(5) 在 H.323 中，呼叫建立过程涉及 3 条信令信道：RAS 信令信道、呼叫信令信道和 H.245 控制信道。通过这 3 条信道的协调，H.323 的呼叫才得以进行，呼叫建立时间很长。在 SIP 中，会话请求过程和媒体协商过程等一起进行。尽管 H.323v2 已对呼叫建立过程作了改进，但与之相比，SIP 只需要 1.5 个回路时延来建立呼叫。

(6) H.323 的呼叫信令信道和 H.245 控制信道需要可靠的传输协议。而 SIP 独立于低层协议，一般使用 UDP 等无连接的协议，用其应用层的可靠性机制来保证消息的可靠传输。

总之，H.323 沿用的是传统的实现电话信令的模式，比较成熟，已经出现了不少 H.323 产品。H.323 符合通信领域传统的设计思想，进行集中、层次式控制，采用 H.323 协议便于与传统的电话网相连。SIP 借鉴了互联网的标准和协议的设计思想，在风格上遵循互联网一贯坚持的简练、开放、兼容和可扩展等原则，比较简单。因此，SIP 被公认是下一代互联网的核心控制协议，其应用领域也在不断地扩展，基于 SIP 的应用前景非常看好。

## 1.2 SIP 简介

互联网的许多应用都需要建立和管理一个会话，会话在这里是指在参与者之间的数据交换。由于考虑到参与者的实际情况，这些应用的实现往往是很复杂的：参与者可能是在代理间移动，它们可能可以有多个名字，它们中间的通信可能是基于不同的媒介（例如文本、视频、音频等），甚至多种媒介同时交互。目前有多种通信协议应用于实时的多媒体通信，SIP 就是其中之一。SIP 是 2001 年推出的 IETF 标准（RFC 3261），用于在 IP 数据网络上建立、改变和结束多媒体会话。SIP 能使用互联网端点（用户代理）来发现参与者，并且协商它们

共享会话的特性。为了能够定位会话参与者和为了其他的用途，SIP 创建了一个网络主机（代理服务器）的架构，通过它允许终端用户注册、发出会话邀请，或者发出其他请求。SIP 独立运行于底层的通信协议之上，并且不依赖于所建立的会话类型。

### 1.2.1 SIP 基本功能概述

SIP 是一个应用层的控制协议，可以用来建立、修改和终止多媒体会话（或者会议）例如互联网电话。SIP 也可以邀请参与者参加已经存在的会话，例如多方会议。它可以很方便地在一个已经存在的会话中增加（或者删除）媒体。SIP 透明地支持名字映射和重定向服务，这些特性用于支持个人移动业务，也就是说用户可以使用一个唯一的外部标识，而无须考虑它们所在的实际网络位置。SIP 在建立和终止多媒体通信方面，支持如下 5 个特性。

- (1) 用户定位：检测用于通信的端系统位置。
- (2) 用户有效性：检测参与通信被叫方的意愿。
- (3) 用户能力：检查所采用的媒体和媒体参数。
- (4) 建立会话：“ringing”，在主叫和被叫双方之间建立会话参数。
- (5) 会话管理：包括发送和终止会话、修改会话参数和激活服务等。

SIP 不是一个垂直集成的通信系统。SIP 可能更适合称为一个部件，它可以用作其他 IETF 协议的一个部分，用来构造完整的多媒体架构。例如，该架构将会包含实时数据传输协议（Realtime Transport Protocol, RTP）（RFC 1889）用来传输实时的数据，并且提供 QoS 反馈，同时也包含实时流协议（Real Time Streaming Protocol, RTSP）（RFC 2326）用于控制流媒体的传输，以及媒体网关控制协议（Media Gateway Control Protocol, MEGACO）（RFC 3015）用来控制到公众电话交换网（Public Switched Telephone Network, PSTN）的网关，还有 SDP（RFC 2327）用于描述多媒体会话。因此，SIP 应该和其他的协议一起工作，才能对用户提供的完整的服务。然而，SIP 的基本功能和操作并不依赖于这些协议。一个 SIP 构建一个 VoIP 系统所需要的典型协议栈结构如图 1-1 所示，说明如下。

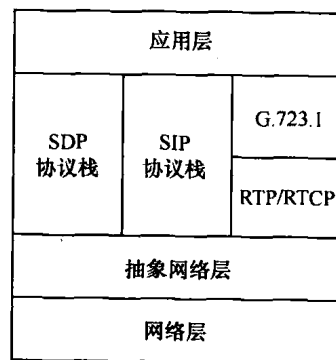


图 1-1 基于 SIP 的 VoIP 系统的协议栈结构

**SIP (RFC 2543)**：该协议是一个应用层控制协议，是多媒体通信的应用层控制（信令）协议栈，负责呼叫的建立与中止，并且也负责通话双方之间媒体流信息的交互。

**SDP (RFC 2327)**：该协议用来描述会话，负责“媒体流描述信息”的生成与分析，从而可以完成会话的公告、建立或者更改。SDP 是一个与通信无关的协议，它被封装在 SIP 消息体中传输。因此，该协议栈的主要任务是生成要发送的 SIP 消息体，并分析收到的 SIP 消息体。

**应用程序层**：它的主要功能是通过调用协议栈提供的 API，完成与用户的交互。

**G.723.1**：语音压缩编码标准，码率为 5.3/7.3kbit/s。

**RTP/RTCP**：实时传输协议和实时传输控制协议，主要负责媒体流的实时传输。

**抽象网络层**：主要任务是对底层函数进行封装，编写一个适合于 SIP Stack 和 SDP Stack 调用抽象网络编程的 API。

**网络层**：负责按 IP 完成数据的转发。

从上述实例可知，SIP 本身并不提供服务。但是，SIP 提供了一些可以用来实现不同服务的原语。例如，SIP 可以用来定位用户，也可以传输一个不透明的对象到对方当前位置。如果利用此原语传输 SDP 会话描述，就可以使端点之间协商会话参数。如果利用此原语来传输主叫方的照片和会话描述 (SESSION DESCRIPTION SD)，就可以很容易地实现“主叫 ID”服务。通过上述简单的例子说明，一个 SIP 原语可以用来提供很多不同的服务。

SIP 并不提供会议控制服务 (例如主席控制或者投票管理)，也没有描述会议应该如何管理。但可以通过在 SIP 上建立其他的会议控制协议来发起一个会议。由于 SIP 的消息和建立的会话可以跨越不同的异构网络，所以 SIP 没有也不打算提供任何形式的网络资源预留管理功能。

安全对于提供的服务来说特别重要。为了达到理想的安全程度，SIP 提供了一套安全服务，包括防止拒绝服务、认证服务 (用户到用户，代理到用户)、完整性保证、加密和隐私服务等。

SIP 可以基于 IPv4，也可以基于 IPv6。

### 1.2.2 SIP 的基本功能实体

SIP 系统采用互联网中常用的客户端/服务器 (Client/Server, C/S) 模型，定义了若干种不同的服务器和用户代理，通过与服务器之间的请求和响应完成呼叫和传送层的控制。用户代理是指为了向服务器发送请求而与服务器建立连接的应用程序。SIP 系统的端系统称为用户代理 (User Agent, UA)。因为 UA 既要能发出呼叫，又要能接收呼叫，所以一个 UA 包含一个用户代理服务器 (User Agent Client; UAC) 和一个用户代理客户端 (User Agent Server, UAS)。UAC 负责呼叫的发出，而 UAS 负责呼叫的接收。服务器是用于向客户端发出的请求提供服务并回送应答的应用程序。共有 4 类基本服务器。

(1) 代理服务器：SIP 请求可以经由多个代理服务器，每个服务器接收请求后将其转发给下一跳服务器。下一跳服务器可能是另一个代理服务器，也可能是最终的用户代理服务器。代理服务器代表其他客户端发起请求，是既充当服务器又充当客户端的媒体程序。在转发请求之前，它可能改写原请求消息中的内容。

(2) 重定向服务器：一种 SIP 服务器，用来从 UAC 接收请求，并将该请求中的 SIP URL 映射到零或多个下一跳服务器的地址，然后将这些地址以响应消息的方式告诉 UAC。UAC 根据收到的新地址，重新向下一跳服务器发送请求消息。值得注意的是：重定向服务器不能发送任何请求，同时也不能接收通话请求。它仅接收 SIP 请求，并把请求中的原地址映射成零个或多个新地址，返回给客户端。

(3) 定位服务器：是互联网中的一种公共服务器，不在 SIP 系统的范畴之内。其查询可采用多种协议，这些协议通常都比较成熟，例如 Finger (RFC 1288)、RWhois (RFC 2167)、LDAP (RFC 1777) 等。

(4) 注册服务器：它接收客户端的注册请求，完成用户地址的注册。

用户终端程序往往需要包括用户代理客户端和用户代理服务器。代理服务器、重定向服务器和注册服务器可以看作是公众性的网络服务器。

SIP 共规定了 6 种方法：INVITE、ACK、CANCEL、OPTIONS、BYE 和 REGISTER。其中 INVITE 和 ACK 用于建立呼叫，完成 3 次握手，或者用于呼叫建立以后改变会话属性；BYE 用于结束会话；OPTIONS 用于查询服务器能力；CANCEL 用于取消已经发出但未最终

结束的请求；REGISTER 用于客户端向注册服务器注册用户位置等信息。SIP 基本实体之间的关系如图 1-2 所示。

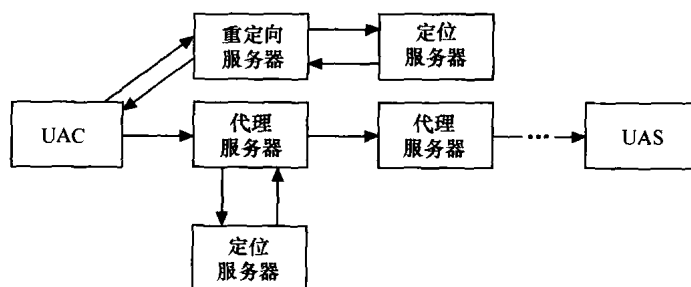


图 1-2 SIP 基本实体之间的关系

### 1.2.3 SIP 呼叫建立的基本过程

SIP 支持 3 种呼叫方式：由 UAC 向 UAS 直接呼叫，由 UAC 在重定向服务器的辅助下进行重定向呼叫和由代理服务器代表 UAC 向被叫发起呼叫。

#### 1. UAC 和 UAS 直接建立或者终止会话过程

在 UAC 和 UAS 之间直接建立或者终止会话的流程如图 1-3 所示。

(1) 主叫方向被叫方发出 INVITE 类 SIP 请求消息，表示主叫方邀请被叫方加入该会话。其中，消息体中包含了该会话的媒体流描述，该描述包含“主叫方能够接收的媒体格式与参数”和“主叫方准备发送的媒体格式与参数”。

(2) 被叫方发回状态码为 200 的响应消息，表示被叫方同意加入该会话，即呼叫被接纳。其中，消息体中也包含了该会话的媒体流描述，该描述必须包含“被叫方能够接收的媒体格式与参数”，有时也可能包含“被叫方准备发送的媒体格式与参数”。

(3) 主叫方向被叫方发出 ACK 类 SIP 请求消息，表示客户端已经收到了服务器对 INVITE 类 SIP 请求消息的最终响应，即已经接收到第 2 步中状态码为 200 的 SIP 响应消息。该 ACK 请求消息中可能还包含媒体流描述信息，被叫方应该按照该描述中的格式和参数发送或者接收媒体流；如果该 ACK 请求消息中不包含媒体流描述，则以最初的 INVITE 请求消息中的描述为准。

(4) 通话完毕后，主叫方首先挂机。主叫方向被叫方发出 BYE 类 SIP 请求消息，表示主叫方希望结束该会话。被叫方收到该消息后，应立即停止向主叫方发送媒体流。

(5) 被叫方发回状态码为 200 的响应消息，表示会话结束。其中，消息体为空。

#### 2. 基于代理服务器的呼叫建立过程

基于代理服务器（Proxy Server）参与的呼叫建立过程如图 1-4 所示。

(1) 主叫方向被叫方发出 INVITE 类 SIP 请求消息，而被叫方注册的域中有一个代理服务器，该请求消息被这个代理服务器接收。

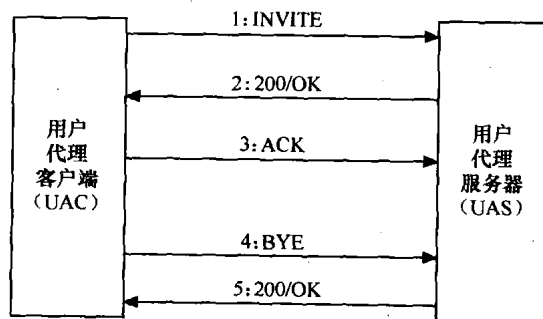


图 1-3 UAC 和 UAS 直接建立或者终止会话的流程

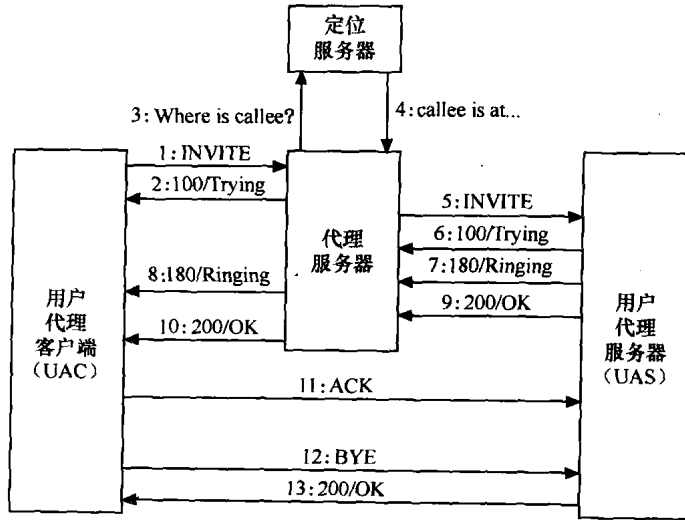


图 1-4 基于代理服务器的呼叫建立过程

(2) 代理服务器收到 INVITE 类请求消息后，立即向主叫方发回状态码为 100 的响应消息。该响应消息表示请求已收到，并且代理服务器已经产生了某种动作处理它，但被叫方的精确地址还没有找到。

(3) 代理服务器向定位服务器查询被叫的当前精确地址，其中代理服务器不属于 SIP 服务器的范畴。

(4) 定位服务器向代理服务器返回被叫当前可能的精确地址。

(5) 代理服务器“代理”主叫方向被叫方发出 INVITE 类 SIP 请求消息，该请求消息是根据第 1 步接收到的 INVITE 请求消息产生的新 INVITE 类请求消息。

(6) 被叫方向代理服务器发回状态码为 100 的响应消息，表示请求正在处理中。

(7) 被叫方向代理服务器发回状态码为 180 的响应消息，表示已经找到了一个被叫方最近注册过的地址（被叫方可能在此，也可能不在此），并且试图振铃提示被叫方。

(8) 代理服务器向主叫方发回状态码为 180 的响应消息，意义与第 7 步相同。

(9) 被叫方向代理服务器发回状态码为 200 的响应消息，表示被叫方同意加入该会话，即呼叫被接纳。

(10) 代理服务器向主叫方发回状态码为 200 的响应消息，表示被叫方同意加入该会话，即呼叫被接纳。

(11) 主叫方向被叫方发出 ACK 类 SIP 请求消息，表示客户端已经收到了服务器对 INVITE 类 SIP 请求消息的最终响应，即已经接收到第 10 步中状态码为 200 的 SIP 响应消息。

注：其中第 1、5、9、10、11 步中消息的消息体均可能包含用 SDP 描述的媒体流格式与参数。

(12) 通话完毕后，主叫方首先挂机。主叫方向被叫方发出 BYE 类 SIP 请求消息，表示主叫方希望结束该会话。被叫方收到该消息后，应立即停止向主叫方发送媒体流。

(13) 被叫方发回状态码为 200 的响应消息，表示会话结束。其中，消息体为空。

### 3. 重定向过程

基于“重定向过程”的呼叫建立过程如图 1-5 所示。

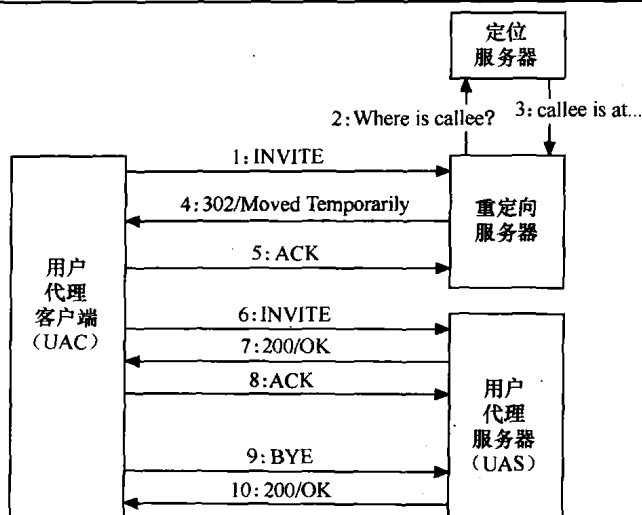


图 1-5 基于“重定向过程”的呼叫建立过程

(1) 主叫方向被叫方发出 INVITE 类 SIP 请求消息，这个请求被被叫方所在域中的重定向服务器所接收。

(2) 重定向服务器向定位服务器查询被叫方当前的地址。

(3) 定位服务器向重定向服务器返回被叫方当前的地址。

(4) 重定向服务器向主叫方发回状态码为 302 的 SIP 响应消息，指出被叫方当前暂时移动到了某个新地址。这个新地址包含在该响应消息中。

(5) 主叫方向重定向服务器发出 ACK 类 SIP 消息，表示已经收到新地址。

(6) 主叫方向新地址发出 INVITE 类 SIP 请求消息，要求被叫方加入该会话。

(7) 被叫方发回状态码为 200 的响应，表示被叫方同意加入该会话，即呼叫被接纳。

(8) 主叫方向被叫方发出 ACK 类 SIP 请求消息，表示客户端已经收到了服务器对 INVITE 类 SIP 请求消息的最终响应，即已经接收到第 10 步中状态码为 200 的 SIP 响应消息。

注：其中第 1、6、7、8 步中消息的消息体均可能包含用 SDP 描述的媒体流格式与参数。

(9) 通话完毕后，主叫方首先挂机。主叫方向被叫方发出 BYE 类 SIP 请求消息，表示主叫方希望结束该会话。被叫方收到该消息后，应立即停止向主叫方发送媒体流。

(10) 被叫方发回状态码为 200 的响应，表示会话结束。其中，消息体为空。

## 第 2 章 SIP 的基本概念及协议结构

SIP 是一个灵活而又简单的互联网应用层的控制协议，其设计原理完全符合互联网的基本理念。它具有端到端、可扩展、分布式等特征。要真正理解 SIP 的思想和内涵，首先需要了解 SIP 中的一些基本概念以及协议的基本结构。下面就以 IETF RFC 3261 为依据，介绍 SIP 中一些重要概念。

### 2.1 SIP 中的基本概念

从上一章 SIP 简介可以了解到，SIP 是一个基于客户端/服务器模式的多媒体通信控制协议。多媒体通信是一个复杂的过程，为了更清楚地阐述该协议过程，SIP 中定义了一些基本概念。在了解 SIP 之前，首先分类介绍这些概念。

#### 2.1.1 SIP 中定义的网络实体

在 SIP 网络中，定义了多种服务器和用户代理等网络实体。需要说明的是，下面介绍这些网络实体只是逻辑实体，在一个具体实现中，可能将它们组合到一个物理实体上实现，当然也可以分开在不同网络实体上实现。

**服务器 (Server):** SIP 系统中，服务器是指能接收客户请求、处理请求，并且发送回应给请求方的网络实体。SIP 系统中典型服务器有代理服务器、用户代理服务器、重定向服务器和注册服务器。

**代理服务器 (Proxy Server):** 它是组成 SIP 系统的中间实体。它本身既作为客户端，又作为服务器，为其他客户端提供请求的转发服务。代理服务器首先提供的是路由服务，也就是说保证请求被发到更加“靠近”目标用户的地方。代理服务器可以强制执行某些策略（例如，确认一个用户是否允许建立一个呼叫等）。代理服务器可以在转发前重写请求消息。

**有状态的代理服务器 (Stateful Proxy):** 在逻辑上，有状态的代理服务器就是处理请求过程中，维护本规范所定义的客户端和服务器的状态机，也称作事务状态代理服务器 (Transaction Stateful Proxy)。状态代理服务器分为事务状态代理服务器和会话状态代理服务器。

**无状态的代理服务器 (Stateless Proxy):** 在逻辑上，无状态的代理服务器在处理请求中，并不维护客户端和服务器的状态机。无状态的代理服务器直接转发每一个接收到的请求和应答。

**重定向服务器 (Redirect Server):** 重定向服务器是产生 3xx 应答的 UAS 服务器，指示

客户端连接别的 URI。

**注册服务器 (Registrar):** 注册服务器是接收 REGISTER 请求的服务器。它把请求的信息放到定位服务器中, 这样可以定位服务器很方便地查找位置信息。

**客户端 (Client):** 客户端是能发出 SIP 请求和接收 SIP 应答的网络实体, 它不一定直接与人交互。客户端又可分为用户代理客户端和代理服务器。

**用户代理 (User Agent, UA):** UA 是一个逻辑实体的概念, 它包含用户代理客户端和用户代理服务器。

**用户代理客户端 (User Agent Client, UAC):** UAC 是一个逻辑的概念, 它创建一个新请求, 并且用客户端事务状态机发送这个请求。UAC 角色只在事务中存在。换句话说, UAC 就是一小段代码初始化一个请求, 并且在事务中遵循 UAC 的规则。如果它接下来收到一个请求, 那么在那个事务中, 它则作为 UAS 来处理请求。

**用户代理服务器 (User Agent Server, UAS):** UAS 是一个逻辑的实体, 对 SIP 请求作出应答、接收、拒绝或者转发对应的请求。UAS 角色只存在于一个事务期间。换句话说, 作为对请求进行应答的一段程序, 在某一个事务期间作为 UAS 存在。而如果它接着发出请求, 那么它就在事务中作为 UAC 存在。

**背靠背的用户代理 (Back-to-Back User Agent, B2BUA):** B2BUA 是一个逻辑实体。该逻辑实体一方面具有 UAS 功能, 能接收和处理请求; 另外还需要具有 UAC 功能, 为了确定该如何应答一个请求, 它必须能转发请求。B2BUA 需要维持对话状态, 并且参与已经建立对话中的每一个请求。因此, B2BUA 可以理解为 UAC 和 UAS 的直接串连。

**呼出代理服务器 (Outbound Proxy):** 它指不是由请求消息中 Request URI 所指定, 而是由 UA 手工配置或者通过自动配置协议所配置的、能接收客户端请求的代理服务器。

**核心 (Core):** 核心定义了 SIP 实体的特定类别。根据核心不同, 可以将网络实体定义为有状态和无状态的代理服务器、用户代理或者注册服务器等。例如, UAC 核心 (UAC Core) 是指在事务层和传输层之上的 UAC 实现的功能集合; 而 UAS 核心 (UAS Core) 是指在事务层和传输层之上的 UAS 实现的功能集合。

### 2.1.2 SIP 消息中定义的概念

SIP 除了定义一些重要的网络实体如服务器的概念外, 还定义了构建、实现消息时所用到的的一些概念。这些概念对于理解 SIP 具有重要的定义。

**消息 (Message):** SIP 元素之间传送的协议数据就是消息。SIP 消息既可以是请求也可以是应答。

**邀请者 (Initiator):** 用 INVITE 消息发起一个会话 (和对话) 建立的一方, 又称为主叫方或主叫 (Calling Party, Caller)。一个 Caller 从发出 INVITE 请求建立对话开始, 到对话终止都一直是这个角色。

**受邀者 (Invitee):** 收到 INVITE 请求并且建立会话的一方, 又称为被叫方或被叫 (Called Party, Callee)。一个 Callee 从收到 INVITE 请求起, 到终止 INVITE 建立的对话结束, 都一直是这个角色。

**方法 (Method):** 方法是组成请求消息的一部分, 是请求服务器进行处理的主要功能。典型的方法就是 INVITE 和 BYE。



**请求 (Request):** 由客户端发到服务器的 SIP 消息, 用于执行特定的功能。

**应答 (Response):** 由服务器发到客户端的 SIP 消息, 用来标识从客户端发往服务器的请求被处理的情况。

**临时应答 (Provisional Response):** 服务器用来标识自己正在处理的应答, 但是本应答并不结束一个 SIP 事务。1xx 应答是临时的, 其他应答标志着事务的结束。

**最终应答 (Final Response):** 一个终结 SIP 事务的应答, 和事务中间的临时应答相反。所有的 2xx、3xx、4xx、5xx 和 6xx 应答都是最终应答。

**头域 (Header):** 头域是在 SIP 消息头部用来描述 SIP 消息信息的部分。它由一堆头域字段组成。

**头域字段 (Header Field):** 头域字段是在 SIP 消息头域的字段。它可以由多个头域字段行组成。一个头域字段由一个头域名和零个或多个头域值组成。多个头域值用逗号分割。某些头域字段只能有单个值, 例如结果 (result) 域就只能有一个值。

**头域值 (Header Field Value):** 头域值是一个单个的值, 一个头域字段可以有零个或者多个头域值。

**会话 (Session):** 多媒体会话是一个由多媒体发送方和接收方组成的集合, 它还包括在发送方和接收方之间的数据流。多媒体会议是一个典型的多媒体会话。会话由 SDP 来描述, SDP 规定一个会话可以由 SDP 用户名、会话标识、网络类型、地址类型和地址元素等信息来表示。

**呼叫 (Call):** 呼叫是指在端点之间的一个通信行为, 通常用于建立多媒体对话。

**呼叫状态 (Call Stateful):** 它指一个对话从 INVITE 消息开始到 BYE 消息终结的整个通话过程。如果一个代理服务器 (proxy) 保存呼叫状态, 那么这个代理服务器就是请求有状态的。一个请求有状态的代理服务器也一定是事务有状态的, 但是事务有状态的代理服务器不一定是请求有状态的。

**对话 (Dialog):** 对话是指持续一段时间的两个 UA 之间的端到端的 SIP 关系。一个对话由 SIP 消息建立。对话可以采用呼叫标识号 (call identifier)、本地标签 (local tag)、对方标签 (remote tag) 来标识。

**会议 (Conference):** 它指一个包含多个参与方的多媒体会话。

**路由集 (Route Set):** 路由集是一个有顺序的 SIP 或者 SIPS URI 列表。这些 URI 描述了传递一个请求所必须经历的代理服务器列表。由于包头中包含了记录路由 (Record-Route), 因此一个路由集可以是自适应的, 也可以是根据配置得到的。

**环路 (Loop):** 当一个请求抵达代理服务器, 代理服务器将转发这个请求, 如果这个请求再次来到同一个代理服务器, 就称之为环路。当第二次抵达时, Request URI 中包含了上次抵达的信息, 并且由于对转发策略没有进行改变, 这样就导致这个请求还会再次被转发回来。环路请求是错误的路由, 处理程序需要检测和防止协议中出现环路请求。

**松散路由 (Loose Routing):** 代理服务器在处理路由头域时按照下列过程来处理, 就称该过程为松散路由。该过程将请求的目的地地址 (由请求 URI 表示) 从沿途需要访问的代理服务器的集合中分离出来处理。如果一个代理服务器与该机制兼容, 则被称为松散路由器。

**下行流 (Downstream):** 它表示事务中的消息传递方向。它特指从 UAC 到 UAS 的请求流方向。