

GUILHERME CÉSAR PILON

SIP - O PROTOCOLO PARA CONVERGÊNCIA NAS REDES DE NOVA GERAÇÃO

Trabalho de Conclusão de Curso apresentado à
Escola de Engenharia de São Carlos, da
Universidade de São Carlos.

Curso de Engenharia de Computação

ORIENTADORA: Mônica de Lacerda Rocha

São Carlos

2014

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTE TRABALHO,
POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS
DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

P643s Pilon, Guilherme César
SIP - O PROTOCOLO PARA CONVERGÊNCIA NAS REDES DE
NOVA GERAÇÃO / Guilherme César Pilon; orientadora
Mônica de Lacerda Rocha. São Carlos, .

Monografia (Graduação em Engenharia de Computação)
-- Escola de Engenharia de São Carlos da Universidade
de São Paulo, .

1. SIP. 2. NGN. 3. VoIP. 4. IMS. 5. Protocolo. I.
Título.

FOLHA DE APROVAÇÃO

Nome: Guilherme Cesar Pilon

Título: “SIP - O protocolo para convergência nas redes de nova geração”

Trabalho de Conclusão de Curso defendido em 20 / 11 / 2014.

Comissão Julgadora:

Resultado:

Profa. Dra. Mônica de Lacerda Rocha
(Orientadora) - SEL/EESC/USP

Aprovada

Prof. Dr. João Paulo Pereira do Carmo
SEL/EESC/USP

Aprovado

Mestre Tania Regina Tronco
Doutoranda - SEL/EESC/USP

Aprovado

Coordenador do Curso Interunidades - Engenharia de Computação:

Prof. Associado Evandro Luís Linhari Rodrigues

Agradecimentos

Agradeço a Deus, por ter aberto as portas necessárias para que eu chegasse até aqui;

A minha família, por ter me dado apoio e ter acreditado em mim durante esta caminhada;

À Professora Mônica, por ter me guiado no desenvolvimento deste trabalho;

A todos os professores que transmitiram seus conhecimentos ao longo deste curso;

A todos os amigos, que proporcionaram anos inesquecíveis durante minha vida universitária;

Aos meus colegas da Trópico Telecomunicações, que abriram as portas e despertaram o interesse pela área de Telecomunicações.

“Be like water making its way through cracks. Do not be assertive, but adjust to the object, and you shall find a way around or through it. If nothing within you stays rigid, outward things will disclose themselves.”

(Bruce Lee)

Lista de Figura

Figura 1 - Sistema de Transmissão na rede legada

Figura 2 - Quantidade linhas fixas telefônicas (Milhões por Ano)

Figura 3 - Modelo OSI

Figura 4 - Interação entre SIP *User Agent*, Servidor SIP e o Serviço de localização

Figura 5 - Estabelecimento de uma chamada SIP simples

Figura 6 - Mobilidade SIP

Figura 7 - Troca de mensagens para a autenticação de uma sessão

Figura 8 - Troca de mensagens H.323 para se estabelecer um *streaming* de mídia

Figura 9 - Quantidade linhas móveis telefônicas (Milhões por Ano)

Figura 10 - Número de pessoas por país usando a Internet (Porcentagem por Ano)

Figura 11 - Arquitetura em camadas do IMS

Lista de Tabelas

Tabela 1 - Divisões do Sistema Telefônico

Tabela 2 - Descrição das camadas do modelo OSI

Tabela 3 - Exemplo de mensagem SIP

Tabela 4 - Descrição dos principais métodos do protocolo SIP

Tabela 5 - Descrição das respostas às solicitações do protocolo SIP

Tabela 6 - Comparação entre SIP e H.323

Lista de Abreviações

3GPP - 3rd Generation Partnership Project

ASCII - American Standard Code for Information Interchange

ASN.1 - Abstract Syntax Notation One

B2BUA - Back to Back User Agent

BGCF - Breakout Gateway Control Function

CIC - Circuit Identification Code

CSCF - Call Session Control Function

DNS - Domain Name System

GPRS - General Packet Radio Service

GSM - Global System for Mobile

HSS - Home Subscriber Server

HTTP - Hypertext Transfer Protocol

IETF - Internet Engineering Task Force

IM - Instant Message

IMS - IP Multimedia Subsystem

IP - Internet Protocol

IPSec - Internet Protocol Security

IPTV - Internet Protocol Television

ISDN - Integrated Services Digital Network

ISUP - ISDN User Part

ITU - International Telecommunication Union

ITU-T - ITU Telecommunication Standardization Sector

I-CSCF - Interrogating Call Session Control Function

LAN - Local Area Network

LTE - Long-Term Evolution

MG - Media Gateway

MGCF - Media Gateway Control Function

MMUSIC - Multiparty MULTimedia Session Control

MRF - Media Resource Function

MRFC - Media Resource Function Controller

MRFP - Media Resource Function Processor

NGN - Next Generation Network

OSI - Open Systems Interconnection

PBX - Private Branch Exchange

PC - Personal Computer

PCM - Pulse Code Modulation

PSTN - Public Switched Telephone Network

P-CSCF - Proxy-Call Session Control Function

QoS - Quality of Service

RFC - Request for Comment

RPC - Remote Procedure Call

RTP - Real-time Transport Protocol

SDP - Session Description Protocol

SIP - Session Initiation Protocol

SLF - Subscriber Location Function

SMTP - Simple Mail Transfer Protocol

SS7 - Signalling System Number 7

S/MIME - Secure/Multipurpose Internet Mail Extensions

S-CSCF - Serving Call Session Control Function

TCP - Transmission Control Protocol

TDM - Time-division multiplexing

TLS - Transport Layer Security

UA - User Agent

UAC - User Agent Client

UAS - User Agent Server

UDP - User Datagram Protocol

UE - User Equipment

URI - Uniform Resource Identifier

VoIP - Voice over IP

Resumo

O mundo da telefonia vem passando por mudanças através da evolução dos meios de comunicação, dispositivos móveis cada vez mais “inteligentes” e a Internet se tornando algo que faz parte do cotidiano das pessoas,. Hoje em dia, conversar com outra pessoa que está a quilômetros de distância já não é o suficiente para satisfazer as necessidades do usuário, sendo que este, atualmente, quer além de conversar, ver a outra pessoa, saber onde ela está em tempo real e ainda compartilhar conteúdo multimídia durante a chamada. Estes recursos são possíveis de serem implementados através de uma infraestrutura concebida para o mundo telefônico graças a uma tecnologia conhecida como NGN (*Next Generation Network*), passando a utilizar a rede IP, ao invés de redes privadas de comutação de circuito, sendo possível a utilização de *softwares* para o controle e processamento das chamadas.

O protocolo SIP (*Session Initiation Protocol*) , que será descrito nesta monografia, foi então desenvolvido pelo IETF (*Internet Engineering Task Force*) para criar, controlar e encerrar as sessões de comunicação entre dois usuários através da NGN. Com ele, é possível não apenas fazer uma chamada telefônica, através do VoIP (*Voice of IP*), mas o protocolo também proporciona a criação de qualquer sessão multimídia entre os usuários, podendo ela ser voz, vídeo, texto e permite até mesmo que os usuários duelem em um jogo *online*.

A utilização do protocolo SIP também ajuda na criação de uma arquitetura de rede para a NGN que possibilita a convergência das redes móveis e fixa facilitando a operação de chamadas entre prestadores de serviços (operadoras) e reduzindo os custos para o usuário final.

Esta monografia tem como objetivo descrever o protocolo SIP e como ele é utilizado hoje em dia no núcleo das redes das operadoras de telefonia.

Palavras Chave: NGN, SIP, Protocolo de Sinalização, VoIP, IMS (*IP Multimedia Subsystem*).

Abstract

With the evolution of media, increasingly "smart" mobile devices and the Internet becoming part of everyday life, the world of telephony is undergoing changes. Today, a conversation with another person who is kilometers away is no longer enough to satisfy the user's will. Now this user wants to chat, while seeing the other person and knowing where the other person is in real time and still share multimedia content during the call. These features are possible to be implemented using an infrastructure designed for telephone world thanks to a technology known as NGN (Next Generation Network), that uses the IP network to carry data, instead of circuit switching used by private telephony networks. With the NGN it is possible to use software for controlling and processing the calls.

The (Session Initiation Protocol) SIP protocol, which is described in this monograph, was then developed by the IETF (Internet Engineering Task Force) to create, manage and terminate communication sessions between two users via the NGN. With it, you can not just make a phone call through VoIP (Voice of IP), but the protocol also provides resources for the creation of any multimedia session between users, being voice, video, text and even online games.

The use of the SIP protocol also helps creating a network architecture for NGN that enables convergence of mobile and wired networks facilitating the call operation between service providers (carriers) and reducing the cost to the end user.

This paper aims to describe the SIP protocol and how it is used today in the core network of telephone operators.

Key Words: NGN, SIP, Signaling Protocol, VoIP, IMS (IP Multimedia Subsystem).

Sumário

Introdução.....	12
1 A telefonia na rede legada	14
1.1 Limitações deste sistema	16
1.2 A consolidação da Internet	17
1.3 A rede NGN	18
2 O protocolo SIP	20
2.1 Descrição	20
2.1.1 SIP User Agents	23
2.1.2 Servidores	23
2.1.3 Serviços de localização	24
2.2 Funcionamento do protocolo	24
2.3 Recursos do protocolo	26
2.3.1 Resolução de Endereço	26
2.3.2 Configuração da Sessão	27
2.3.3 Negociação de Mídia	28
2.3.4 Modificação da Sessão	29
2.3.5 Encerramento e Cancelamento da Sessão	29
2.3.6 Controle da Chamada	30
2.3.7 Mobilidade	31
2.3.8 Transporte de Mensagem de texto	31
2.3.9 Notificação de Evento	32
2.3.10 Registro de Presença	33
2.3.11 Autenticação	33
2.3.12 Extensibilidade	34
2.4 Limitações do SIP	34
2.5 Comparação com outros protocolos NGN	35
2.5.1 Introdução ao H.323	35
2.5.2 Comparação	36
3 Aplicação do SIP na NGN - IMS	40
3.1 Descrição do IMS	40
3.2 O SIP no IMS	43
3.2.1 Servidor CSCF (S-CSCF)	43
3.2.2 <i>Proxy</i> CSCF (P-CSCF)	44
3.2.3 <i>Interrogating</i> CSCF (I-CSCF)	44
3.2.4 Servidores de Aplicação	44
3.2.5 <i>Gateways</i>	45
3.2.6 BGCF (<i>Breakout Gateway Controller Function</i>)	45
3.2.7 MRF (<i>Media Resource Function</i>)	45
Conclusão	47
Referencias Bibliográficas	48

Introdução

Este trabalho tem como objetivo estudar o protocolo SIP, descrevendo suas principais características, funcionalidades e suas aplicações no mundo das telecomunicações.

Quando o SIP foi desenvolvido, era esperado que a NGN substituiria toda a rede legada de telefonia e que cada usuário teria em suas casas e escritórios telefones que suportassem o protocolo SIP, com este fazendo a sinalização de toda a chamada, do aparelho do usuário chamador, até o aparelho do usuário chamado, porém não é isso que vemos ainda hoje. Apesar de existirem aparelhos de telefone que suportam o protocolo SIP, estes são muito caros, não despertando o interesse do usuário final em investir na troca do seu aparelho. Mas, ao contrário do que ocorre na ponta da rede, com o usuário final, para as operadoras é vantajoso a troca das centrais telefônicas da rede legada para novas centrais da rede NGN, que nada mais são que computadores, facilitando assim sua manutenção. Com isso, o SIP ganhou espaço no núcleo da rede de telefonia e hoje faz a sinalização das chamadas entre as centrais e operadoras e através de *gateways*, esta sinalização é traduzida para trafegar na rede legada que chega até o usuário final.

A monografia está dividida em 4 partes: 'Introdução', 'Descrição e Funcionalidades do Protocolo', 'Aplicação do Protocolo' e 'Conclusão'. Na Introdução, será discutido como era feita a sinalização telefônica antes do surgimento da NGN, assim como seus problemas e limitações e como o desenvolvimento da Internet impactou no mundo das telecomunicações. Em seguida, na seção Descrição e Funcionalidades do Protocolo, o SIP será apresentado e seu comportamento como um dos principais protocolos de sinalização nas redes NGN, explicado. Já na parte final, uma comparação com o H.323, será feita com o intuito de ressaltar as características do SIP. Na seção Aplicação do Protocolo, uma introdução ao IMS será feita e como o SIP é usado nesta arquitetura, que visa à convergência de todas as redes de telecomunicações,

facilitando a integração dos serviços. Por fim, a Conclusão apresenta as considerações finais deste trabalho.

Buscando ter uma visão mais ampla do protocolo SIP, suas funcionalidades e sua implicação nas redes de telecomunicações, este trabalho não entra em quesitos estritamente técnicos do protocolo, ao invés, busca uma visão em mais alto nível do protocolo para que dela possa ser apreendidas as vantagens de sua utilização.

1 A telefonia na rede legada

A rede telefônica fixa é o sistema básico de telecomunicações que compõe os aparelhos utilizados pelos usuários do sistema - e de um vasto conjunto de acessórios, tudo isto com o objetivo de prover a interligação dos usuários do sistema de telefonia (assinantes) à central telefônica e as várias centrais entre si, a **Figura 1** ilustra este conceito.

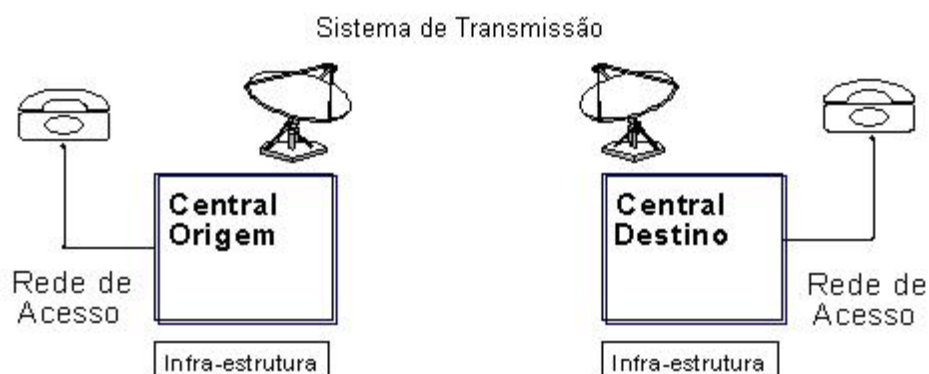


Figura 1 - Sistema de Transmissão na rede legada.

Outro termo clássico utilizado é sistema telefônico, que pode ser definido como o sistema que permite a comunicação de dois assinantes através do telefone. Esse sistema divide-se em subsistemas, descritos na **Tabela 1**, que interagem operacionalmente para formar a rede de telefonia como conhecemos: Rede de Comutação, Rede de Acesso e Rede de Transmissão [2].

Além desses subsistemas, existe o subsistema de sinalização por canal comum nº 7, responsável pela inteligência de comunicação entre os elementos da rede de telecomunicações.

O ISDN (*Integrated Services Digital Network*) *User Part* ou ISUP é parte do sistema de sinalização nº 7 (SS7), utilizado para estabelecer chamadas telefônicas na rede telefônica pública comutada (PSTN - *Public Switched Telephone Network*). Ele é especificado pelo órgão de padronização internacional ITU-T (*ITU Telecommunication Standardization Sector*) como parte da série Q.76x [1].

Tabela 1 - Divisões do Sistema Telefônico [2]

Subsistemas	Descrição
Rede de Comutação	Equipamentos necessários à seleção do caminho que possibilita a comunicação entre os usuários
Rede de Acesso	Suporte físico necessário para a comunicação dos telefones às centrais de comutação
Rede de Transmissão	Suporte que permite a propagação da informação entre as centrais de comutação

Quando uma chamada telefônica é efetuada a partir de um assinante para outro, várias centrais telefônicas podem estar envolvidas, possivelmente cruzando fronteiras internacionais. Para permitir que uma chamada seja efetuada corretamente, dentro da rede PSTN (*Public Switched Telephone Network*), um comutador vai sinalizar informações referentes à chamada, como o número chamado para o próximo comutador na rede usando mensagens ISUP.

As centrais telefônicas podem ser conectadas via troncos E1 (padrão utilizado pelo Brasil) ou T1 (padrão utilizado pelos Estados Unidos) que transportam os dados das chamadas. Esses troncos são divididos em intervalos de tempo de 64 Kbit/s, sendo que cada intervalo de tempo transporta os dados referentes a uma chamada, resultando em uma taxa agregada de 2 Mbit/s no padrão brasileiro e 1,544 Mbit/s no caso do padrão americano [2]. Cada circuito entre dois comutadores é unicamente identificado por um código de identificação do circuito (CIC), que é incluso nas mensagens ISUP [1]. O CIC é usado, juntamente com a informação de sinalização recebida (especialmente o número chamado) para determinar quais os circuitos de entrada e saída devem ser ligados entre si para fornecer um caminho ponta-a-ponta para que a chamada aconteça [7].

Além de transportar informações relacionadas à chamada, o ISUP também é usado para a troca de informações de *status* e gerenciamento dos circuitos disponíveis. No caso de nenhum circuito de saída estar disponível em uma determinada rota, uma mensagem de liberação é enviada de volta aos comutadores precedentes à cadeia, até chegar ao originador da chamada.

A sinalização em sistemas de telefonia é o principal mecanismo pelo qual as chamadas telefônicas são criadas, modificadas e encerradas. Por exemplo, a sinalização de um telefone empresarial diz ao PBX (*Private Branch eXchange*) para encaminhar a chamada para outro telefone. Na rede telefônica pública, a sinalização telefônica instrui os sistemas de comutação a enviar uma ligação 0800 para um *call center* específico onde um atendente irá atender a chamada.

Um exemplo da importância de sinalização é a diferença existente entre uma ligação normal entre residências e uma chamada 0800 para um Serviço de atendimento ao cliente. Tais chamadas são tarifadas de forma diferente. Para o usuário final, ambos os telefonemas possibilitaram o mesmo, a conversa entre duas pessoas, exceto que a sinalização permitiu a adição de valor comercial para a chamada 0800.

A sinalização permite ser aplicado diferentes serviços nas chamadas telefônicas, tais como, chamadas ponto-a-ponto, conferência multiponto, texto, voz e vídeo, entre outros.

1.1 Limitações deste sistema

Apesar da sinalização nº 7 proporcionar estes serviços no mundo telefônico, também conhecido como TDM (*Time-division multiplexing*), este chegou ao limite de sua evolução. Como a PSTN utiliza uma rede própria para seu funcionamento, o aumento do número de linhas telefônicas a partir da década de 1970, como podemos ver no gráfico da **Figura 2**, implicou também na necessidade do aumento da rede PSTN para suportar a demanda, e com isso, o custo das operadoras para ampliar e dar manutenção em suas redes, criando assim um problema de escalabilidade deste sistema.

Porém, com os avanços nos meios de comunicação, com o desenvolvimento e viabilidade do uso comercial das fibras óptica e a criação de protocolos que garantissem uma boa qualidade em serviços que necessitam da transmissão de conteúdo em tempo real, como o protocolo SIP (*Session Initiation Protocol*) e o RTP (*Real-time Transport Protocol*), as redes IP (*Internet Protocol*) passaram a ter um papel importante no mundo da telecomunicações.

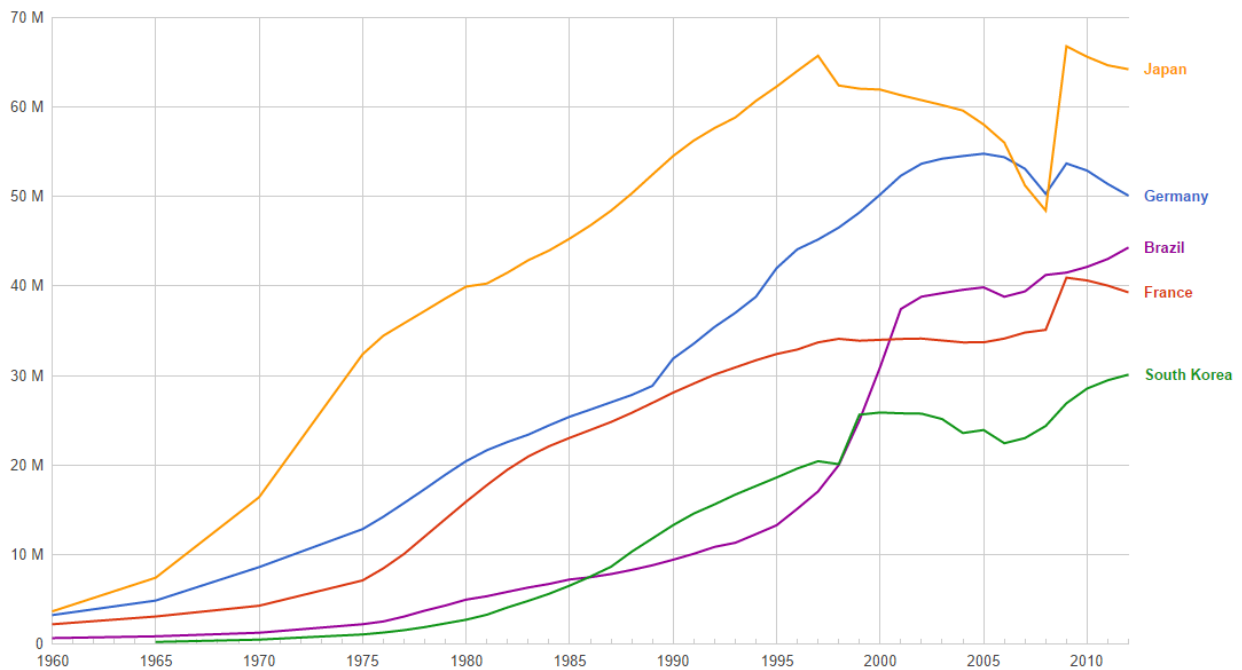


Figura 2 - Quantidade linhas fixas telefônicas (Milhões por Ano) [8]

1.2 A consolidação da Internet

A Internet se beneficia de diferentes fundamentos em comparação aos das redes legadas de telefonia, como, por exemplo, o enorme progresso na rede IP e dos protocolos que a definem. Este avanço rápido pode ser atribuído a pesquisas acadêmicas e às comunidades de engenharia cuja dedicação e colaboração aberta em uma base global ultrapassaram a pressão comercial usual e o sigilo competitivo empresarial, que eram a base para o desenvolvimento da rede legada de telefonia.

O resultado é uma Internet que utiliza protocolos consistentes com um padrão global, e é igualmente robusta para o transporte de dados, transações e comunicações em tempo real, nomeadamente mensagens instantâneas (IM - *Instant Message*), voz, vídeo e conferência.

Na verdade, a Internet pode ser considerada uma rede transparente, projetada para qualquer aplicação, mesmo aquelas que ainda não foram inventadas.

Essa transparência para qualquer aplicação é devido à organização em camadas padronizada pelo modelo OSI.

Este modelo é dividido em camadas hierárquicas, sendo que cada uma dessas camadas é servida pela sua camada anterior e serve a sua camada superior,

implementando uma comunicação horizontal, onde cada camada se comunica apenas com a sua camada correspondente da outra máquina [3].

As camadas são empilhadas de acordo com a **Figura 3** e são descritas na **Tabela 2**:



Figura 3 - Modelo OSI

1.3 A rede NGN

NGN (*Next Generation Network*) é o nome dado às mudanças arquiteturais das redes centrais de telecomunicações e de acesso. A ideia geral da NGN é que uma única rede transporte todas as informações e serviços (voz, dados, e todos os tipos de mídia, como vídeo), através de pacotes, pela Internet. A NGN é construída utilizando os protocolos da Internet, e, portanto, o termo “tudo sobre IP” também pode ser usado para descrever a transformação das redes de telefonia, anteriormente centradas, em direção à NGN.

A NGN é uma rede baseada em pacotes que podem fornecer serviços de telecomunicações e é capaz de fazer uso da banda larga, provendo QoS (Quality of Service) no transporte destes serviços e fazendo com que as funções relacionadas com o serviço sejam independentes das tecnologias de transporte utilizados. Ela oferece acesso irrestrito dos usuários à diferentes prestadores de serviços (operadoras). [15]

Tabela 2 - Descrição das camadas do modelo OSI [3]

Camada	Descrição
Física	Define especificações elétricas e físicas dos dispositivos
Enlace	Detecta e corrige erros que possam acontecer no nível físico
Rede	Fornece os meios funcionais e de procedimento de transferência de dados da origem para o destino em uma rede. Cria os pacotes.
Transporte	Responsável por receber os dados enviados pela camada de sessão e segmentá-los para que sejam enviados a camada de rede
Sessão	Permite que duas aplicações em computadores diferentes estabeleçam uma comunicação
Apresentação	Converte os dados em um formato entendido pelo protocolo usado
Aplicação	Corresponde às aplicações que serão utilizadas para promover a interação entre a máquina e o usuário

Na NGN, há uma separação mais definida entre a parte de transporte (conectividade) da rede e os serviços que são executados (aplicações) sobre o que é transportado. Isto significa que sempre que uma operadora queira habilitar um novo serviço, ela pode defini-lo diretamente na camada de serviço, sem considerar a camada de transporte - ou seja, os serviços são independentes de detalhes de transporte. Com isso, as aplicações, incluindo voz, tendem a ser independente da rede de acesso e residem mais em dispositivos de usuário final (por exemplo, telefone, PC, *set-top box*).

2 O protocolo SIP

2.1 Descrição

O SIP é um protocolo codificado em texto com base em elementos da *HyperText Transport Protocol* (HTTP), que é usado para navegação na Internet, e também no *Simple Mail Transport Protocol* (SMTP), que é usado para o envio e recebimento de e-mail na Internet [10]. A **Tabela 3** ilustra a codificação em texto do protocolo SIP.

Tabela 3 - Exemplo de mensagem SIP.

```
INVITE sip:bob@biloxi.com SIP/2.0
Via: SIP/2.0/UDP pc33.atlanta.com;branch=z9hG4bK776asdhds
Max-Forwards: 70
To: Bob <sip:bob@biloxi.com>
From: Alice <sip:alice@atlanta.com>;tag=1928301774
Call-ID: a84b4c76e66710@pc33.atlanta.com
CSeq: 314159 INVITE
Contact: <sip:alice@pc33.atlanta.com>
Content-Type: application/sdp
Content-Length: 142

v=0
o=CiscoSystemsCCM-SIP 811669 1 IN IP4 10.105.40.14
s=SIP Call
c=IN IP4 10.133.92.102
t=0 0
m=audio 25268 RTP/AVP 18 101
```

O protocolo foi desenvolvido pela *Multiparty Multimedia Session Control IETF* (MMUSIC) *Working Group* como parte da arquitetura da *Internet Multimedia Conferencing* [9], mas, desde então, ganhou sua própria *SIP Working Group* dentro do IETF. Como o nome indica, a função primária do SIP é a iniciação de uma sessão, mas também tem outros usos e funções importantes, como a notificação de presença e de mensagens instantâneas. O SIP é usado para comunicação

ponto-a-ponto, isto é, aqueles em que ambas as partes na chamada são considerados iguais, no entanto, o SIP utiliza um modelo cliente-servidor de transações semelhante ao HTTP, como descrito na próxima seção. Um cliente SIP gera uma solicitação SIP. Um servidor SIP responde ao pedido, gerando uma resposta.

O conjunto principal de solicitações SIP (conhecidas como métodos) são mostrados na **Tabela 4**. As seis primeiras são definidas na RFC (*Request For Comment*) 3261 [11], a especificação base do protocolo. As demais são extensões do SIP e são definidas em RFCs separadas. Novos métodos estão sendo continuamente propostos para adicionar funcionalidades adicionais para o protocolo.

As respostas às solicitações SIP são numéricas. Muitos códigos de resposta são baseados no protocolo HTTP, bem como os novos criados. Códigos de resposta SIP são divididos em seis classes, identificadas pelo primeiro dígito do código, como mostra a **Tabela 5**.

Tabela 4 - Descrição dos principais métodos do protocolo SIP

Método	Descrição
INVITE	Inicia uma sessão SIP
ACK	Confirmação da resposta ao Invite
BYE	Termina uma sessão SIP
CANCEL	Cancela uma sessão pendente
REGISTER	Registro de URI do usuário
OPTIONS	Consulta de opções e capacidades
INFO	Transporte sinalização durante uma sessão
PRACK	Confirmação à uma resposta provisória
UPDATE	Atualiza informações da sessão
REFER	Transfere o usuário para uma URI
SUBSCRIBE	Solicitar notificação de um evento
NOTIFY	Transporte de notificação de eventos
MESSAGE	Transporte de uma mensagem instantânea
PUBLISH	Publica estado de presença para um servidor

Tabela 5 - Descrição das respostas às solicitações do protocolo SIP

Classe	Descrição
1xx	Provisória ou Informativa - Solicitação está em progresso, mas ainda não está completa.
2xx	Sucesso - Solicitação foi concluída com êxito.
3xx	Redirecionamento - Solicitação deve ser tratada em outro local.
4xx	Erro no cliente - Solicitação não foi concluída por causa de um erro na origem, pode ser repetida quando corrigido.
5xx	Erro no servidor - Solicitação não foi concluída por causa de um erro no destinatário, pode ser repetido em outro local.
6xx	Falha Global - Solicitação falhou

Há três elementos principais em uma rede SIP: *SIP User Agents* (os agentes), servidores e serviços de localização, que estão dispostos como mostra a **Figura 4**.

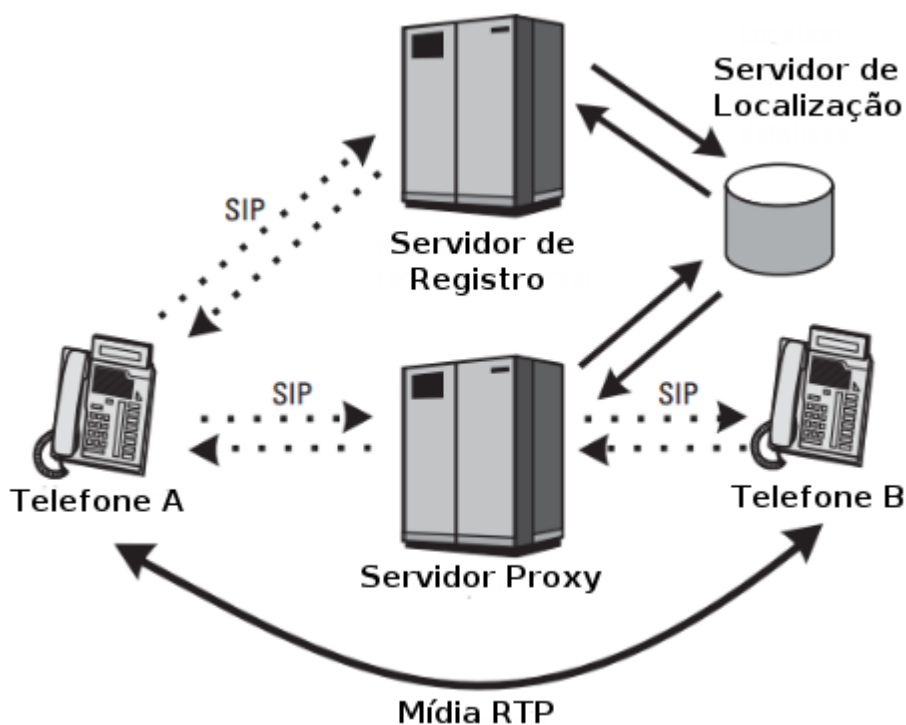


Figura 4 - Interação entre SIP User Agent, Servidor SIP e o Serviço de localização [9]

2.1.1 SIP User Agents

Um dispositivo que suporta o protocolo SIP é chamado de SIP *user agent* [1], podendo este ser um telefone SIP ou um software executado em um computador ou *smartphone*. Um SIP *User Agent* pode ser uma porta de entrada para uma outra rede, como um *gateway* para a PSTN, que permite que um telefone SIP receba e faça chamadas para a PSTN.

Uma das finalidades do SIP é permitir que as sessões sejam estabelecidas entre os *user agents* (UAs). Um UA é comandado por um usuário e atua como um agente em seu nome para estabelecer e terminar sessões de mídia com outros UAs. Na maioria dos casos, o usuário será uma pessoa, mas também pode ser um outro protocolo, tal como no caso de um *gateway*.

Um SIP UA se comporta como um cliente e um servidor. Durante uma sessão, o agente irá geralmente operar tanto como um *user agent client* (UAC) e um *user agent server* (UAS).

O UAC inicia solicitações enquanto a UAS responde a essas solicitações. Um SIP UA também deve suportar o protocolo SDP (*Session Description Protocol*) para descrição da mídia que a sessão a ser iniciada usará. Outros tipos de protocolos para descrição de mídia podem ser usados em conjunto com o SIP, contudo, o suporte ao SDP pelo SIP UA é obrigatório [11].

2.1.2 Servidores

Servidores SIP são entidades que aceitam solicitações e respondem a elas. Um servidor SIP não deve ser confundido com um UAS ou a natureza cliente-servidor do protocolo, que descrevem as operações em termos de clientes (criadores das solicitações) e servidores (originadores de respostas às solicitações). Como os servidores fornecem serviços e recursos para os agentes, eles devem suportar os protocolos de transporte TCP (*Transmission Control Protocol*), TLS (*Transport Layer Security*) e UDP (*User Datagram Protocol*). A **Figura 4** mostra a

interação entre os agentes, servidores e um serviço de localização. Note-se que o protocolo utilizado entre um servidor e o serviço de localização (ou banco de dados) não é necessariamente o SIP.

Existem 3 tipos básicos de servidores SIP:

- *SIP proxys* recebem solicitações SIP de um UA ou outro *proxy* e encaminham esta solicitação para outro local;
- *Redirect servers* recebem uma solicitação de um UA ou de um *SIP proxy* e retorna uma resposta de redirecionamento (3xx), indicando para que local a solicitação deve ser tentada novamente;
- *Registrar servers* recebem pedidos de registro de um SIP UA e atualizam as informações do usuário em um serviço de localização ou outro banco de dados.

SIP proxy, *Redirect server* e *Registrar servers* são somente elementos de sinalização, eles não têm capacidade de manipular media, sendo também passivos, isto é, não iniciam nenhuma solicitação, a não ser que o UA tenha solicitado.

2.1.3 Serviços de localização

Serviço de localização (do inglês, *location services*) é um termo geral utilizado na RFC 3261 [11] para se referir a um banco de dados. O banco de dados pode conter informações sobre os usuários, tais como URIs, endereços IP, *scripts*, características distintas, e outras preferências. Ele também pode conter informações de roteamento sobre a rede, incluindo os locais de *proxys*, *gateways* e outros serviços de localização. Os UA geralmente não interagem diretamente com um serviço de localização, mas recorrem a um *SIP proxy*, *Redirect server* ou *Registrar server* para obter alguma informação sobre outro UA.

2.2 Funcionamento do protocolo

A **Figura 5** mostra um exemplo de troca de mensagens SIP entre dois dispositivos habilitados para SIP.

Os dois dispositivos podem ser telefones SIP, computadores, *palmtops*, ou telefones celulares conectados a uma rede IP (como a Internet) e devem conhecer o endereço IP de cada um, pois não há a participação de um SIP *proxy*, configurando assim uma ligação SIP ponto-a-ponto.

A parte que inicia a sessão, o UA 10.20.30.100, começa a troca de mensagens através do envio de um *INVITE* para a parte chamada (o UA 10.30.100.200). Essa mensagem contém os detalhes do tipo de sessão ou chamada que é solicitado, podendo ser uma simples chamada de voz, uma sessão multimídia (por exemplo, uma videoconferência), ou uma sessão de jogo.

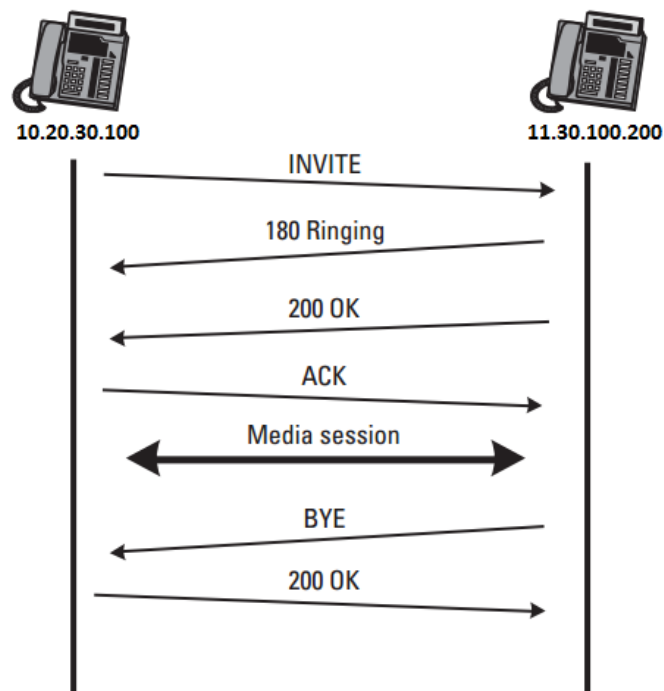


Figura 5 - Estabelecimento de uma chamada SIP simples [9]

O protocolo SIP separa o estabelecimento da sessão de sua descrição de mídia. Esta separação deve ser mantida sempre que novas extensões forem adicionados ao protocolo, a fim de assegurar que o protocolo continuará genérico, podendo estabelecer qualquer tipo de sessão, independente do tipo de mídia que será utilizado.

2.3 Recursos do protocolo

À seguir serão descritos os recursos básicos que o protocolo SIP possibilita dentro da rede NGN.

2.3.1 Resolução de Endereço

Resolução de endereço é uma das funções mais importantes do protocolo SIP. O processo de resolução de endereço no protocolo SIP geralmente começa com um URI e termina com um *username* em um endereço IP. Esta resolução de um nome genérico para um usuário pertencente a um *host* é extremamente poderosa, pois possibilita facilmente a implementação de vários serviços úteis para a telefonia, como mobilidade e portabilidade numérica. A resolução de endereço pode ser realizada tanto pelo UA quanto pelos servidores. A resolução de endereço é obtida através da utilização de DNS para a conversão.

Embora seja possível que um SIP UA tenha acesso a um serviço de localização, esta pesquisa é geralmente realizada por um *proxy* ou *Redirect server* ao invés de um SIP UA.

Em geral, o processo envolve a resolução de endereços e saltos múltiplos das mensagens SIP. Isso permite que UAs e *proxys* realizem o roteamento das mensagens SIP, *hop-by-hop*, isto é, cada elemento consulta seu servidor DNS ou uma tabela de roteamento, em seguida, encaminha a solicitação para o próximo salto. Este processo continua até que a mensagem seja enviada para o destino. Ao contrário das requisições, as respostas em SIP não envolvem a resolução de endereço; todas as respostas são roteadas de volta à origem através do mesmo conjunto de *proxys* que a mensagem de requisição trafegou. Isso é possível por causa do cabeçalho VIA, presente nas mensagens SIP, que empilha os URI dos *proxys* por onde a mensagem de requisição passou.

2.3.2 Configuração da Sessão

A configuração inicial de uma sessão entre duas partes é a função principal do protocolo SIP.

O SIP usa uma mensagem *INVITE* para iniciar a configuração de uma sessão entre dois UAs. O corpo da mensagem do *INVITE* geralmente contém uma descrição do tipo de sessão que o UA deseja estabelecer.

Um UA SIP que deseja iniciar uma sessão com outro UA, ativa os campos da mensagem *INVITE* da seguinte maneira:

- *header To*: insere a URI do destino (podendo ser o destino final ou de um servidor que fará o roteamento da mensagem);
- *header From*: insere a URI do originador da mensagem e uma *tag* gerada aleatoriamente;
- *header Call-ID*: parâmetro gerado aleatoriamente para que identifique a chamada.

O UA que gera a resposta para esta mensagem (uma mensagem 200 *OK*, por exemplo) adiciona uma *tag* ao *header To*. A combinação da *tag* do *header To*, da *tag* do *header From*, e do *header Call-ID* é usada para identificar unicamente esta sessão, criando a identificação de um diálogo em SIP. Estes cabeçalhos nunca são modificados durante uma sessão. Essas informações, assim como os elementos de mídia exigidos, descritos no corpo da mensagem SIP por um protocolo de descrição de mídia, geralmente o SDP, representam a quantidade mínima de informações da chamada que um UA deve manter.

A configuração da sessão SIP ocorre através de um *three-way handshake-INVITE / 200 OK / ACK* para uma configuração bem sucedida e *INVITE / 4xx ou 5xx ou 6xx / ACK* para uma configuração com falha. A **Figura 5** mostra uma configuração de sessão bem sucedida entre dois telefones SIP envolvendo um *INVITE*, uma respostas provisórias (180 *Ringin*g) e uma resposta final (200 *OK*), que recebe um *ACK*.

Uma vez estabelecida, uma sessão de mídia continua indefinidamente sem a necessidade de mais trocas de mensagens de sinalização SIP. Um *timer* de sessão SIP pode, no entanto, ser usado para terminar sessões SIP excessivamente longos

[11]. Se um dos UA's participantes da sessão deseja modificar ou encerrar a sessão, uma nova troca de mensagens de sinalização SIP segue.

Este *three way handshake* permite uma bifurcação, que é uma pesquisa paralela iniciado por um *proxy*, em que várias respostas podem ser devolvidas para um único *INVITE*, isso possibilita que, se estiver configurado, mais de um UE seja notificado da chamada, dando a possibilidade ao usuário escolher onde quer atendê-la, seja no seu *smartphone*, computador ou outro dispositivo que suporte chamadas SIP.

2.3.3 Negociação de Mídia

A negociação de mídia faz parte da sequencia inicial de troca de mensagens *INVITE / 200 / ACK*. O próprio SIP não fornece os meios para a negociação de mídia, mas permite que essa negociação ocorra entre os UA's usando o *Session Description Protocol* (SDP). O SDP é uma linguagem de descrição textual, o qual é definido pelo RFC 2327[13].

A negociação de mídia é um modelo *offer-answer* (oferta-aceitação, traduzindo literalmente) definido pela RFC 3264 [14], em que um UA propõe um ou mais tipos de mídia, e outro UA aceita ou recusa cada sessão de mídia em uma resposta.

Normalmente, a oferta da mídia é feita no *INVITE* inicial pelo UA que solicita a chamada, e a resposta é realizada nos 200 *OK* pelo UA chamado. No entanto, o que chama pode permitir que o UA chamado selecione o tipo de sessão de mídia, não enviando o SDP no convite. Neste caso, o UA chamado faz a oferta da mídia no 200 *OK* (ou em uma resposta provisória confiável), e o 'chamador' responde aceitando a oferta da mídia no *ACK*. No SDP ligado à mensagem SIP, os UA's especificam o tipo de mídia, *codec*, o endereço IP e a porta por onde a mídia negociada será transportada. Mais de um *codec* pode ser especificado para cada tipo de mídia. Uma vez que um *codec* foi negociado, os UA's devem estar preparados para receber a mídia com esse *codec* por toda a duração da sessão.

Se por algum motivo um dos UA's participantes da sessão desejarem mudar o *codec* da mídia previamente negociado, isso pode ser feito através re-*INVITES*, que

são mensagens de *INVITE* SIP enviadas depois que a sessão está estabelecida, como será descrito na próxima sessão.

2.3.4 Modificação da Sessão

Uma vez que uma sessão foi estabelecida usando a sequencia *INVITE / 200 / ACK*, ele pode ser modificado por outra sequencia *INVITE / 200 / ACK*, dentro do mesmo diálogo, o que caracteriza um *re-INVITE*. Como só pode haver um pedido SIP pendente por vez em um diálogo, um *re-INVITE* não pode ser enviado até que a negociação do *INVITE* inicial tenha sido concluída com seu respectivo *ACK*. O *re-INVITE* pode ser feito por qualquer uma das partes e usa o mesmos *headers To, From* (incluindo as *tags*), e *Call-Id* do *INVITE* inicial. No entanto, o SDP contido no *re-INVITE* é assumido diferente do SDP contido no *INVITE* inicial, já que esta troca de mensagens tem como objetivo alterar a mídia negociada previamente. Se o *re-INVITE* falhar ou for recusado, o SDP e a sessão de mídia original continuarão até um *BYE* ser enviado por qualquer uma das partes.

Essa modificação da sessão durante uma chamada é um recurso que torna o protocolo SIP muito poderoso. Através da modificação da sessão durante a chamada, é possível que uma pessoa que inicialmente iniciou uma chamada de voz com outra pessoa, modifique a sessão para que ocorra a inclusão de vídeo, fazendo com que ambas as pessoas se vejam, além de se comunicarem. Isso subentende que não se terminou a chamada.

2.3.5 Encerramento e Cancelamento da Sessão

Encerramento e cancelamento da sessão são duas operações distintas em SIP, que podem ser facilmente confundidas. Encerramento da sessão ocorre quando um UA envia um *BYE* fazendo referência a um trecho de chamada existente (isto é, uma sessão estabelecida com sucesso usando a troca de mensagens *INVITE / 200 / ACK*). Já o cancelamento da sessão ocorre quando um UA termina uma chamada antes do fim da configuração e estabelecimento da chamada. Neste cenário, um UA que enviou um *INVITE*, mas ainda não recebeu uma resposta final

(2xx, 3xx, 4xx, 5xx ou 6xx), envia uma mensagem SIP *CANCEL*. Esta mensagem pode também ser originada por um SIP *Proxy* para cancelar pernas individuais que responderam com uma mensagem de erro uma bifurcação ou uma pesquisa paralela.

No mundo da telefonia, o encerramento da sessão ocorre quando, por exemplo, o assinante A liga para o assinante B e após terminarem de conversar, o assinante A se despede do assinante B e coloca o seu telefone no gancho, gerando uma mensagem SIP *BYE* para o assinante B. Já no caso do cancelamento de uma sessão, o assinante A efetua uma ligação para o assinante B, porém ainda enquanto o telefone de B está tocando, o assinante A resolve cancelar a ligação e põe seu telefone no gancho, gerando assim uma mensagem SIP *CANCEL* para o assinante B, cujo telefone para de tocar.

2.3.6 Controle da Chamada

A arquitetura SIP foi concebida de modo que a comunicação ocorra ponto-a-ponto, ao passo que o controle da chamada seja feito somente pelas UA extremos da chamada. Por exemplo, um *proxy* não pode emitir uma mensagem de *BYE* encerrando uma chamada. Isso só pode ser feito por um dos UA (dispositivos finais) participantes da chamada.

No entanto, há possibilidade de um terceiro elemento controlar uma chamada entre dois UAs, esta funcionalidade possibilitada pelo SIP pode ser extremamente útil em várias implementações de serviço para a telefonia. Por exemplo, um controlador de chamadas pode ser implementado para fazer a tradução numérica entre o número discado por um UA para um número totalmente diferente, o que já é feito hoje em dia no caso da portabilidade numérica entre operadoras. Ou um serviço de *call-center* pode ser implementado, onde uma pessoa pode discar um único número referente ao serviço de atendimento e o controlador de chamadas fazer o roteamento para os atendentes seguindo uma série de fatores, como disponibilidade, especialização do atendente, hora do dia, entre outros.

2.3.7 Mobilidade

Em uma mensagem de registro, um usuário envia para um *Registrar Server* a URI para o qual deseja receber chamadas. Este suporte à mobilidade é uma característica extremamente útil do protocolo SIP e é um dos mais citados benefícios do protocolo sobre os outros protocolos usados para sinalização na NGN. É também este o suporte à mobilidade que levou o protocolo a ser utilizado em muitas novas aplicações, incluindo o controle de chamadas em redes sem fio de terceira geração (3G).

A mensagem SIP *REGISTER* é utilizada para realizar esta função. O pedido contém o *header Contact*, contendo as URIs a serem registradas pelo usuário.

Por exemplo, na **Figura 6**, um usuário inicialmente registra seu telefone de trabalho enviando uma mensagem *REGISTER* para o *Registrar Server*, que atualiza o registro do usuário no *Location Server* que retorna uma mensagem 200 OK, confirmando o registro. No final do dia, o usuário sai do trabalho e vai para sua casa, onde ele cancela sua inscrição do telefone do escritório e registra seu telefone de casa. As chamadas destinadas ao URI do usuário serão agora encaminhadas para o endereço IP do telefone SIP de sua casa.

Observe que o telefone da casa do usuário não tem a necessidade de suportar o protocolo SIP para esta mobilidade, o usuário também pode registrar um telefone da PSTN usando o acesso à *Web*, e-mail, e telefone celular, tendo a possibilidade de atender uma chamada em um desses dispositivos.

2.3.8 Transporte de Mensagem de texto

O protocolo SIP também permite o envio de mensagens de texto. Esta função é feita pelo método SIP *MESSAGE* que transporta a mensagem de texto no corpo da mensagem SIP, ao invés do SDP. O UA de destino responde com 200 OK quando recebe a mensagem.

Esse método se diferencia dos outros, pois não há a necessidade de estabelecer uma sessão (*INVITE / 200 / ACK*) para que ocorra o envio da mensagem.

Os outros métodos em SIP para suportar comunicações instantâneas são subscrição de eventos e notificação de presença.

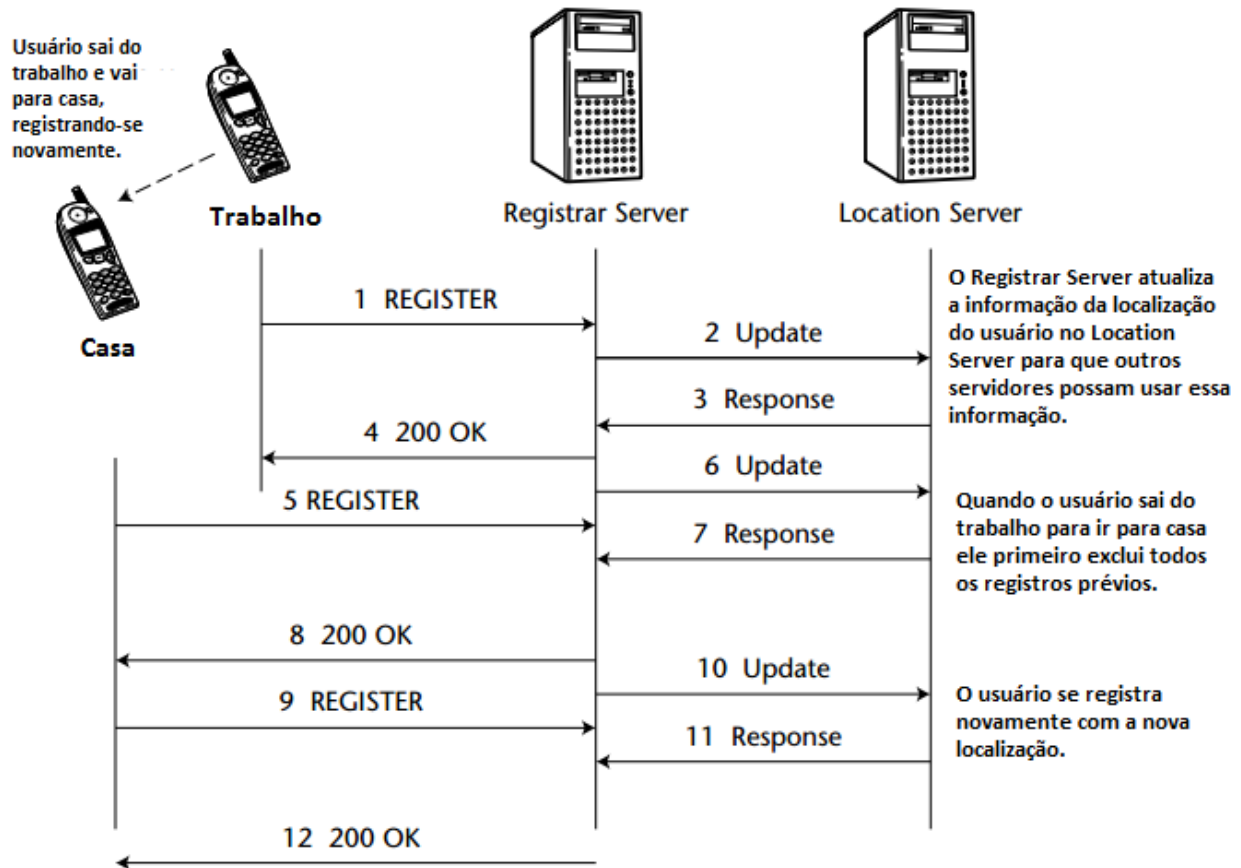


Figura 6 - Mobilidade SIP [9]

2.3.9 Notificação de Evento

O protocolo SIP proporciona ao UA a capacidade de solicitar e receber uma notificação quando um determinado evento ocorre, através dos métodos *SUBSCRIBE* e *NOTIFY*. Pode ser usado, como exemplo desta funcionalidade, a característica de retorno automático de uma chamada telefônica quando o UA chamado está ocupado (fora do gancho) e o UA chamador deseja ser notificado assim que o UA chamado estiver disponível novamente. Isto é feito da seguinte maneira, o usuário A envia um *INVITE* e recebe uma mensagem 486 *Busy Here* como resposta do usuário B. O usuário A então envia uma mensagem *SUBSCRIBE* para o usuário B solicitando uma notificação quando este estiver disponível para

estabelecer uma sessão. Quando o usuário B esteve novamente disponível para estabelecer uma sessão, ele envia uma mensagem *NOTIFY* indicando este fato. O usuário A então estabelece imediatamente a sessão.

Se o usuário B não estiver disposto a fornecer a notificação deste evento, uma mensagem 603 *Decline* pode ser enviada como resposta à mensagem *SUBSCRIBE* enviada pelo usuário A.

2.3.10 Registro de Presença

O método SIP *PUBLISH* permite que um UA que publique ou faça *upload* de informações de presença informação para um *Location Server*, como a URI em que o usuário está aceitando chamadas no momento. O servidor de presença pode então distribuir estas informações para outros elementos da rede SIP, dando suporte para mobilidade em SIP.

2.3.11 Autenticação

O protocolo SIP suporta dois tipos de autenticação: UA para UA, e UA para servidor. Ele não suporta atualmente servidor para servidor, embora isso possa ser feito usando um outro protocolo, como IPSec. O protocolo SIP também usa a título de empréstimo o esquema de autenticação do protocolo HTTP. A autenticação SIP se baseia em um desafio / resposta e um segredo compartilhado entre o UA solicitante e o do servidor que necessite de autenticação. Qualquer pedido SIP pode ser desafiado para se autenticar.

O segredo compartilhado normalmente será um nome de usuário e senha criptografada. A típica de troca de mensagens de autenticação SIP entre os UAs tem a seguinte sequência de mensagens SIP, *INVITE / 401 Authentication Required / ACK* quando o UA descobre que o pedido requer autenticação e simultaneamente descobre a natureza do desafio de autenticação do servidor, presente na mensagem 401. Nesta situação, uma nova mensagem *INVITE* contendo um *header Authorization* é reenviada. Se este *header* contém as credenciais corretas, a

chamada continuará como normal. Caso contrário, uma outra mensagem 401 será recebida. A **Figura 7** ilustra este comportamento.

2.3.12 Extensibilidade

O protocolo SIP foi concebido para ser extensível. Como consequência, o protocolo foi projetado para que os UAs possam implementar novas extensões usando novos *headers* e corpos de mensagens sem necessidade de servidores intermediários, tais como *proxys*. Além disso, permite também suportar essas novas extensões. Por padrão, um *proxy* encaminha inalterado os *headers* desconhecidos presentes nas mensagens SIP. O uso do header *Supported* permite a um UA solicitante informar a rede e outros UAs e quais extensões e recursos ele suporta, permitindo-lhes a opção de usar o recurso. Se existe a necessidade da função ser suportada pelo UA solicitado, o *header Require* informa ao UA solicitado quais as extensões ou funcionalidades necessárias para completar a chamada. Um UA que receber tal mensagem contendo o *header Require* deve retornar um erro se não suportar a extensão requerida.

2.4 Limitações do SIP

O protocolo SIP foi projetado para iniciar, modificar e terminar sessões. Ele não é um protocolo de controle de dispositivo ou chamadas de procedimento remoto (RPC).

Apesar de transportar pequenas mensagens de texto através do método *MESSAGE*, o SIP não é um protocolo de transporte para grandes fluxos de dados. Ele também não é um protocolo para reserva de recursos da rede, uma vez que o caminho de mensagens SIP geralmente não é o caminho do mídia resultantes. O SIP não é um substituto para protocolos utilizados na PSTN - sua abordagem é muito diferente de modelos de chamadas e sinalização em telecomunicações. SIP pode até interagir com a PSTN através de *gateways*, mas esta não é a função principal do SIP.

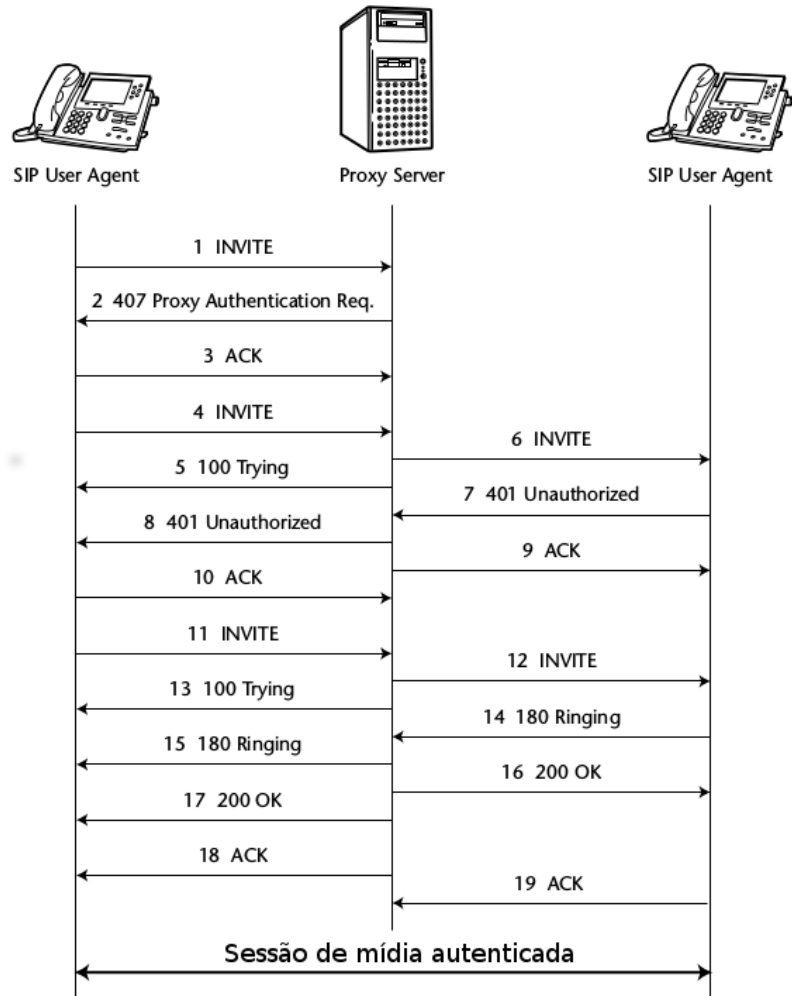


Figura 7 - Troca de mensagens para a autenticação de uma sessão. [9]

O SIP também não é um protocolo VoIP, embora VoIP seja um serviço possível a implementar em uma rede compatível com SIP. Ele é puramente um protocolo de sinalização e não faz nenhuma especificação sobre os tipos de mídia, descrições, serviços e etc.

2.5 Comparação com outros protocolos NGN

2.5.1 Introdução ao H.323

O protocolo H.323 é definido pela ITU *Telecommunication Standardization Sector* (ITU-T) [4], que define os protocolos para fornecer sessões de comunicação audiovisuais em qualquer rede de pacotes. Este protocolo padroniza sinalização e

controle de chamadas, transporte e controle de conteúdo multimídia, e controle de largura de banda para conferências ponto-a-ponto e multiponto.

É usado dentro de várias aplicações da Internet em tempo real, tais como *GnuGk* e *NetMeeting* e é amplamente implantado em todo o mundo pelos prestadores de serviços e empresas de serviços de voz e vídeo através de redes IP.

É uma parte da série de ITU-T H.32x de protocolos, que também abordam a comunicação multimídia sobre ISDN, o PSTN ou SS7, e redes móveis 3G.

A sinalização de chamada H.323 é baseada na Recomendação Q.931 [12] do ITU-T e é adequado para a transmissão de chamadas entre redes que utilizam uma mistura de IP, PSTN, ISDN, e QSIG em ISDN. Um modelo de chamada, semelhante ao modelo chamada ISDN, facilita a introdução da telefonia IP nas redes existentes de sistemas de PBX baseados em ISDN, incluindo transições para PBXs baseados em IP.

Dentro do contexto de H.323, um PBX baseado em IP pode ser um elemento de controlo de *gateway* ou outra ligação, que fornece um serviço de telefones ou chamada de vídeo. Tal dispositivo pode fornecer ou facilitar ambos os serviços básicos e serviços complementares, tais como transferência de chamadas e *hold*.

2.5.2 Comparação

Os protocolos SIP e H.323 foram desenvolvidos para diferentes propósitos por organismos de normalização com requisitos muito diferentes. H.323 foi desenvolvido pela ITU. A sua concepção e implementação refletem seu histórico voltado para a PSTN e rede legada, utilizando codificação binária e reutilização da sinalização utilizada na ISDN. O protocolo SIP, por outro lado, foi desenvolvido pelo IETF com uma perspectiva voltada para a Internet, concebido para ser mais escalável e de trabalhar de uma forma Inter domínio utilizando o conjunto completo de funções na Internet. A **Figura 8** ilustra a troca de mensagens que deve ocorrer entre os usuários usando o protocolo H.323 para que um *streaming* de mídia seja estabelecido entre dois usuários. Se compararmos com a **Figura 5**, vemos que com menos mensagens do protocolo SIP podemos estabelecer o mesmo *streaming*, evitando congestionamento da rede.

Enquanto o protocolo H.323 foi implantado no início da utilização do VoIP e videoconferência sobre IP, quando a rede legada ainda era muito utilizada, o protocolo SIP, com sua arquitetura voltada para Internet está ganhando força e está emergindo como o futuro padrão de sinalização para comunicação IP.

A primeira diferença fundamental está no esquema de codificação usado pelos protocolos. O SIP é um protocolo baseado em texto como HTTP e SMTP, enquanto o H.323 usa uma codificação binária padrão ASN.1 [5]. Por um lado, a codificação binária do H.323 pode resultar em uma mensagem menor para ser transportada pela rede, porém aumenta a complexidade de implementações e análise do tráfego. Um protocolo baseado em texto como o SIP pode ser facilmente manipulado e não requer ferramentas complexas para monitorar e interpretar suas mensagens. Um simples analisador de pacotes, como o *Wireshark*, pode ser utilizado para capturar e analisar as mensagens ASCII trocadas entre os UA's SIP a partir de uma LAN.

Outra diferença importante é que, enquanto o H.323 é exclusivamente um protocolo de sinalização, o SIP possui funções como informação de presença do usuário e capacidade de envio de mensagem instantânea. Isso faz do SIP um protocolo poderoso que permite a um usuário com vários dispositivos móvel suportando SIP localizar e se comunicar com outro usuário de diferentes maneiras, nomeadamente através de áudio, vídeo, mensagens instantâneas [5]. É por esta razão que os novos serviços de telefonia e/ou comunicação estão sendo desenvolvidos utilizando SIP ao invés do H.323.

O protocolo SIP também foi adotado por operadoras de telefonia móvel como a sinalização de chamada e protocolo de mensagem instantânea para suas redes de terceira geração (3G). Esta rede possui milhões de dispositivos que usam o SIP. Este número tende a aumentar cada vez mais com a chegada das redes LTE (4G) e o crescimento do IMS, que será discutido na próxima sessão, aumentando ainda mais a integração entre esses dispositivos móveis.

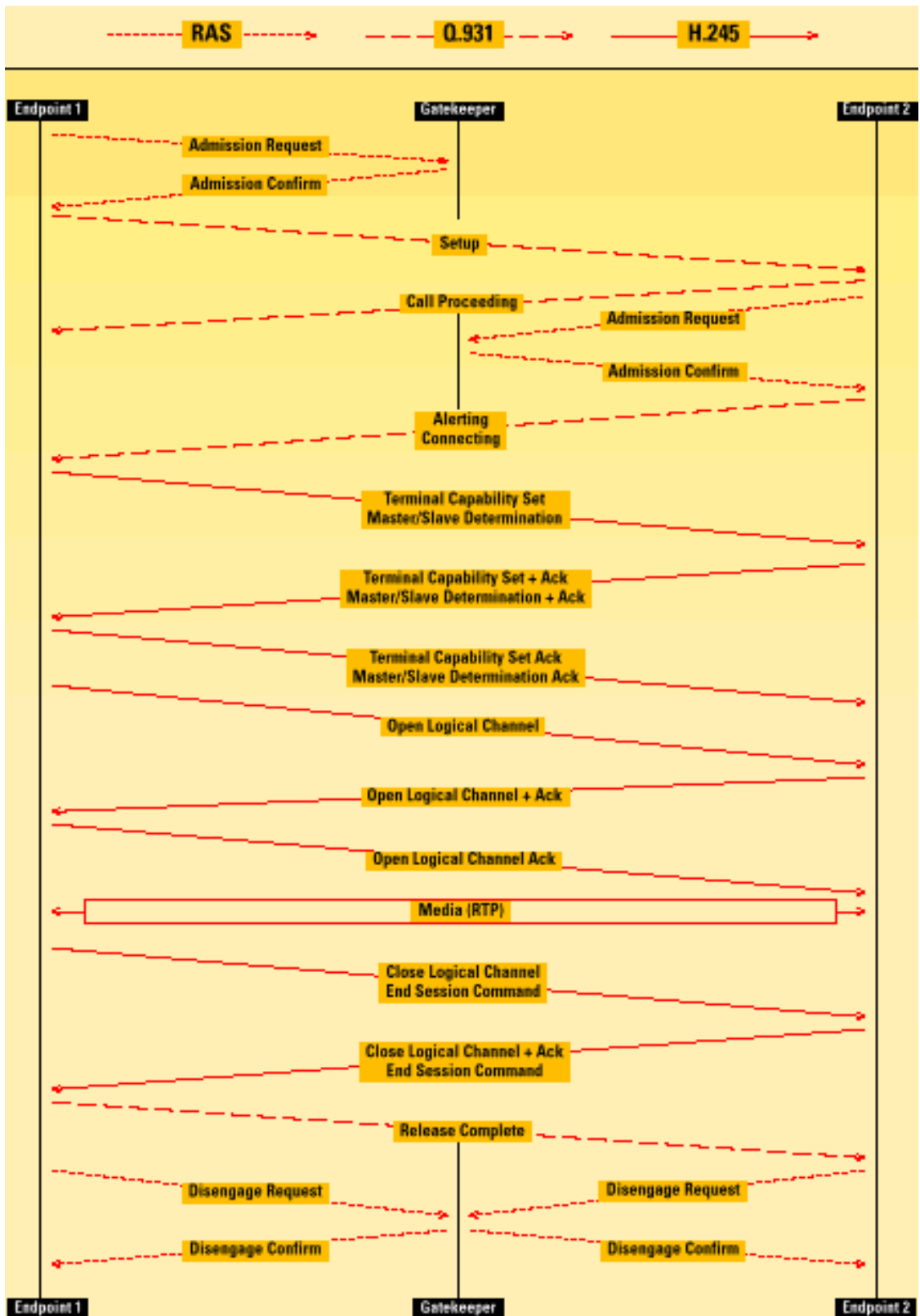


Figura 8 - Troca de mensagens H.323 para se estabelecer um *streaming* de mídia.

Outra diferença importante entre os dois protocolos é o nível de segurança. SIP como definido na RFC 3261[11] tem mecanismos de segurança muito robustos para fornecer encriptação, autenticação usando certificados e garantia de integridade da mensagem de ponta-a-ponta, mesmo na presença de servidores intermediários não confiáveis. Não foi preciso desenvolver estes recursos de segurança no protocolo SIP; em vez disso, uma vez que ele é um protocolo da Internet, o SIP pôde herdar o rico conjunto de protocolos de segurança presentes na Internet, como TLS e S/MIME [5]. Por exemplo, os mesmos mecanismos de segurança que tornam possível digitar informações de um cartão de crédito em um formulário de uma página da Web seguro, é o mesmo que permite ao SIP fornecer uma sinalização segura entre os servidores.

A principal vantagem do SIP é que ele é mais simples e fácil de manipular, em comparação com outros protocolos que são usados na NGN, pois foi desenvolvido já no mundo da Internet como um protocolo de rede.

A **Tabela 6** resume as principais diferenças entre os dois protocolos.

Tabela 6 - Comparação entre SIP e H.323

SIP	H.323
Codificação baseado em texto. <ul style="list-style-type: none"> • Maior • Fácil de manipular 	Codificação binária padrão ASN.1. <ul style="list-style-type: none"> • Menor • Difícil de manipular
Mais recente. <ul style="list-style-type: none"> • Totalmente voltado para Internet 	Desenvolvido no início do VoIP. <ul style="list-style-type: none"> • Voltado para a PSTN
Além de sinalização, possui outras funções. <ul style="list-style-type: none"> • Informação de Presença • Autenticação • Mensagem de texto 	Protocolo exclusivo para sinalização. <ul style="list-style-type: none"> • Menor flexibilidade
Maior nível de segurança, pois utiliza protocolos já robustos da Internet. <ul style="list-style-type: none"> • TLS • S/MIME 	Menor nível de segurança.

3 Aplicação do SIP na NGN - IMS

3.1 Descrição do IMS

No momento estamos experimentando a rápida convergência de mundos telefônicos, fixos e móveis, já que o número de dispositivos móveis está aumentando exponencialmente, como pode ser visto no gráfico da **Figura 9**. Adicionalmente esses dois dispositivos devem se comunicar de uma maneira transparente para o usuário. Estes dispositivos móveis possuem cada vez mais sensores, monitores, câmeras de alta resolução, entre outros recursos para que possam ser usados por diversas aplicações. Com o aumento do uso da Internet (ver **Figura 10**), esses dispositivos móveis estão sempre conectados, redefinindo como essas aplicações são desenvolvidas. Diz-se redefinindo-as, pois agora não são mais entidades isoladas que trocam informações apenas com o usuário, podendo se comunicarem. Discar um número e falar com outra pessoa em breve será visto como um subconjunto da rede IP. Este novo paradigma da comunicação vai muito além das capacidades do antigo sistema de telecomunicações.

Para se comunicar, os aplicativos devem ter mecanismos para se conectar à Internet. A rede telefônica fornece atualmente esta tarefa crítica de estabelecimento da conexão. Ao ligar para o ponto, a rede telefônica pode estabelecer uma conexão *ad hoc* entre dois terminais através da rede IP. Esta capacidade de conectividade IP é oferecida apenas em ambientes de provedores isolados onde a inter-operabilidade entre provedores de diferentes operadoras é uma característica indesejável. Portanto, um sistema global - o *IP Multimedia Subsystem (IMS)* - é de extrema importância para o desenvolvimento de novos serviços que possam beneficiar o usuário. O IMS permite que os aplicativos em dispositivos conectados a rede IP possam estabelecer conexões ponto-a-ponto ou ponto-a-serviço facilmente e com segurança.

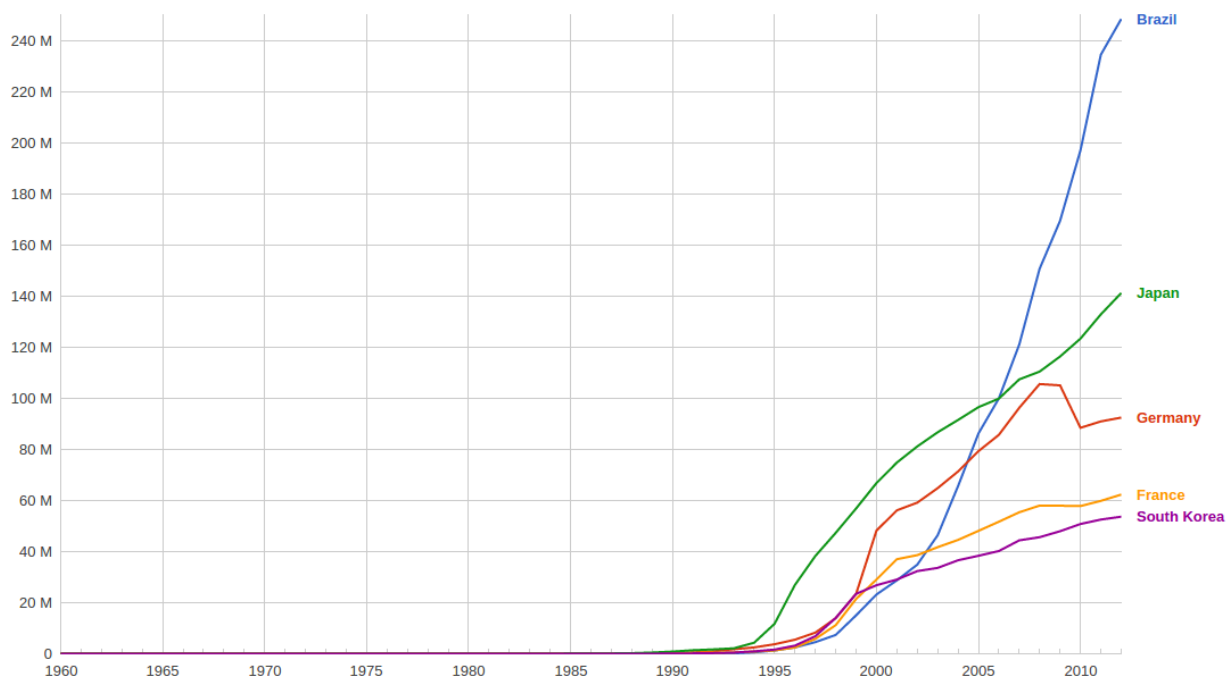


Figura 9 - Quantidade linhas móveis telefônicas (Milhões por Ano) [8]

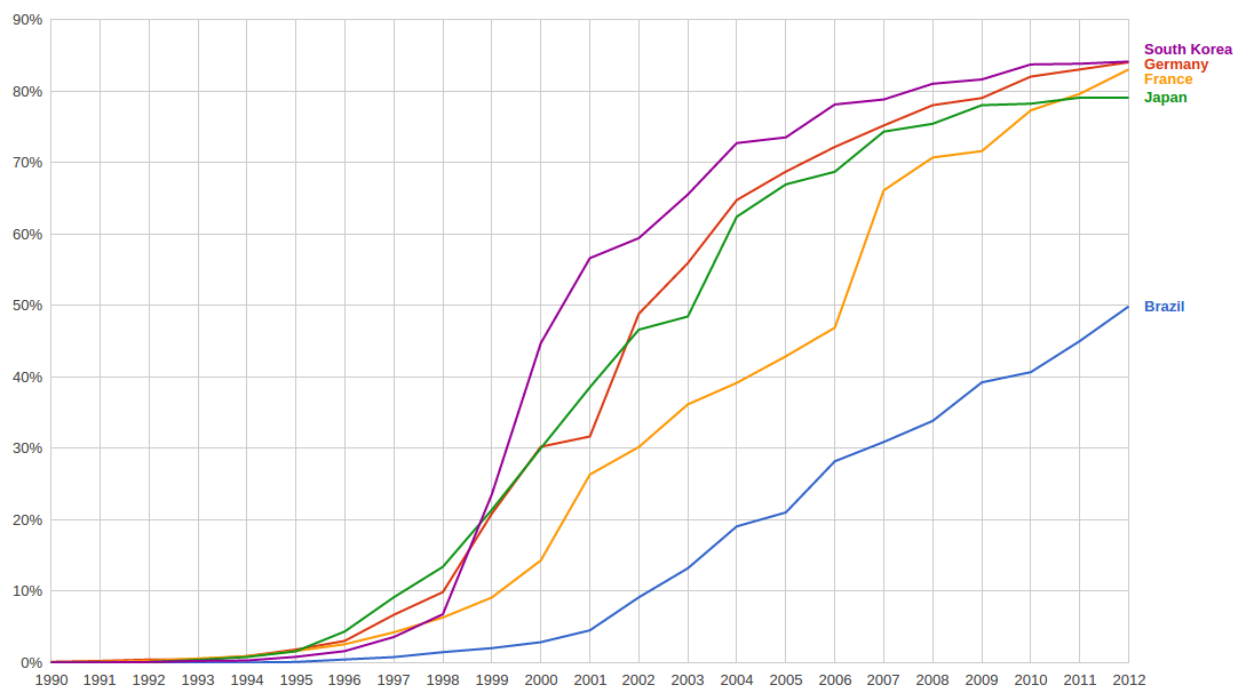


Figura 10 - Número de pessoas por país usando a Internet (Porcentagem por Ano) [8]

A verdadeira integração de serviços de voz e de dados aumenta a produtividade e eficácia geral, enquanto o desenvolvimento de aplicações inovadoras integrando voz, dados e multimídia vai criar demandas por novos

serviços, como bate-papo multimídia, *push to talk* e conferência. A habilidade de combinar a mobilidade e a rede IP é crucial para o serviço de telefonia ter sucesso no futuro.

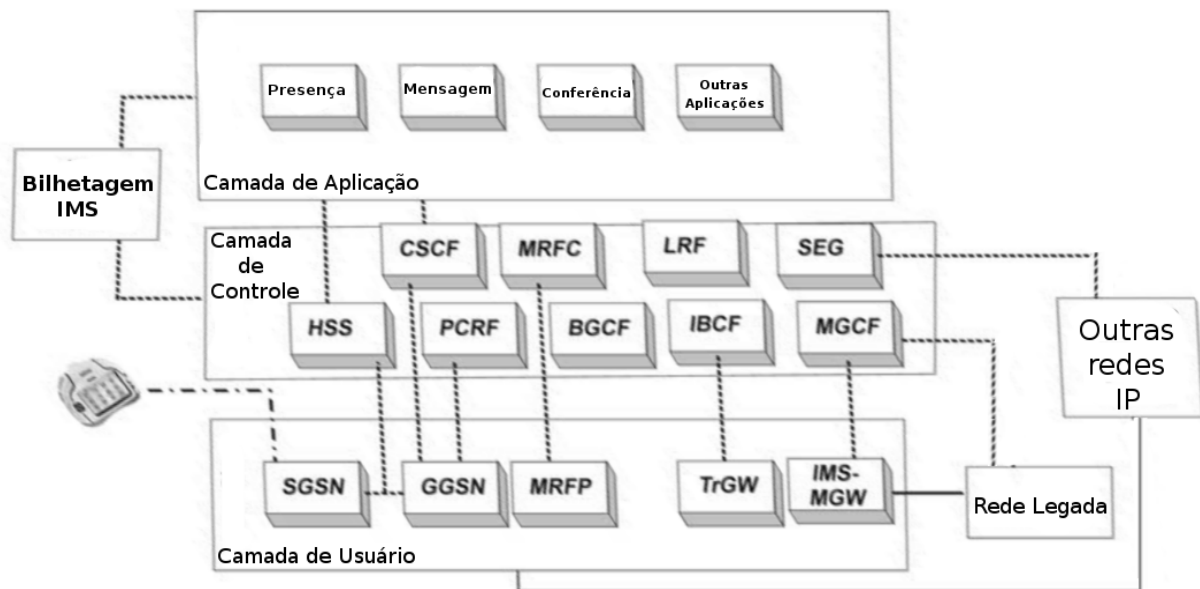


Figura 11 - Arquitetura em camadas do IMS

A convergência das redes móvel e fixa simplifica a experiência do usuário final e dissolve as barreiras e complexidades que separam as ilhas de rede de hoje, assim, os mesmos serviços estão disponíveis em todas as redes, fazendo uso destas fácil, transparente e intuitivo, beneficiando assim o usuário final.

É o IMS que introduz controle de sessão multimídia na rede IP e ao mesmo tempo proporciona a funcionalidade de comutação de circuitos na mesma. O IMS é uma tecnologia chave para essa consolidação da rede IP no mundo da telefonia [6].

O IMS foi originalmente concebido pela 3GPP (*3rd Generation Partnership Project*), visando a evolução das redes móveis GSM (*Global System for Mobile*). Sua formulação original (3GPP Rel-5) representou uma abordagem para oferecer "serviços de Internet" através de GPRS (*General Packet Radio Service*). Esta visão foi posteriormente atualizado para que o IMS também suportasse outras redes além do GPRS, como Wireless LAN, CDMA2000 e linhas fixas.

3.2 O SIP no IMS

O protocolo SIP emprega um papel importante na arquitetura IMS, já que é ele o responsável por cuidar da comunicação entre os elementos que compõe a rede, sobretudo os servidores que compõem a camada de controle do IMS, (conforme ilustrado **Figura 11**).

Estes servidores são também conhecidos como CSCF (*Call / Session Control Function*). Eles executam funções de controle de sessão para sessões de IMS. CSCFs podem ser categorizados em três grupos com base na funcionalidade [16]:

3.2.1 Servidor CSCF (S-CSCF)

O S-CSCF é o centro de todas as funções de sinalização em uma rede IMS. Além de gerenciamento de sessão, um S-CSCF também desempenha o papel de um SIP *Registrar Server* em uma rede IMS. Existe uma interface *Diameter*, protocolo este usado para comunicação com banco de dados, entre S-CSCF e HSS / SLF, que nada mais são do que as entidades do IMS que guardam os dados dos usuários para importar dados de autenticação e perfil de usuário. Todas as mensagens enviadas/recebidas de/para um UA passa por um S-CSCF que fiscaliza essas mensagens para as ações necessárias a tomar (por exemplo, para autorizar um usuário para o uso de um determinado serviço, com base no perfil de utilizador).

Com base na mensagem que recebe, o S-CSCF executa funções de roteamento. Esta função não precisa envolver apenas o roteamento de mensagens para outro servidor SIP, mas pode envolver servidores da camada de aplicação do IMS também. O perfil de usuário (baixado pelo S-CSCF do HSS) é que instrui o S-CSCF se a mensagem de sinalização SIP deve ser encaminhada para um ou mais servidores de aplicativos antes de ser encaminhada para o destino final.

3.2.2 Proxy CSCF (P-CSCF)

O P-CSCF desempenha o papel de um SIP *Proxy Server* para mensagens de entrada e de saída de um terminal IMS. Uma vez que um UA regista-se na rede, é atribuído um único P-CSCF para a duração deste registo.

As tarefas executadas por um P-CSCF são semelhantes as de um *SIP Proxy Server*, com que diz respeito à autenticação, segurança e validação de mensagens SIP, além de gerar dados de tarifação para o UA registrado.

3.2.3 Interrogating CSCF (I-CSCF)

Estritamente falando, um I-CSCF é também um SIP Proxy Server. No entanto, a sua função é específica na rede. Ele está localizado na borda de uma domínio administrativo de uma rede. Quando um P-CSCF quer encontrar o próximo elemento para rotear uma mensagem SIP, ele primeiro obtém o endereço do I-CSCF da rede de destino. Então, o I-CSCF usa sua interface Diameter com o HSS / SLF para encontrar o S-CSCF atribuído ao UA de destino. Posteriormente a mensagem SIP é encaminhada para o S-CSCF apropriado.

3.2.4 Servidores de Aplicação

Os servidores de aplicativos essencialmente executam serviços para os usuários e exercem a função de Servidores de Aplicação SIP. Dependendo do serviço a ser aplicado para aquela sessão, o servidor de aplicações pode operar em um dos seguintes modos [16]:

- Modo SIP *Proxy*;
- Modo SIP *User Agent*;
- Modo SIP *Redirect Server*;
- SIP B2BUA (*Back to Back User Agent* - concatenação de dois UA).0

Os servidores de aplicativos também possuem uma interface com o HSS para fazer *upload* e *download* de dados do usuário, para saber se o serviço pode ou não ser aplicado.

3.2.5 Gateways

Vários tipos de *gateways* são compatíveis com a arquitetura IMS. Por exemplo, a arquitetura inclui *gateways* para a conversão de sinais de comutação de pacotes da Rede IMS para sinais de comutação de circuito da PSTN ou vice-versa.

- Os *gateways* de sinalização fazem a conversão de um protocolo para outro nas camadas mais baixas de uma rede; por exemplo, a conversão do SIP na rede de comutação de pacotes para o H.248/MEGACO, na rede de comutação de circuitos [16].
- Os *gateways* de converção de dados de mídia - *Media Gateway* (MG) e *Media Gateway Controller Function* (MGCF). A MG faz a interface de mídia entre duas redes. Assim, ele converte as mídias que usam o protocolo RTP (na rede IMS) para o PCM (*Pulse Code Modulation*) que geralmente é o protocolo usado para o transporte de mídia do lado da PSTN [16].

3.2.6 BGCF (*Breakout Gateway Controller Function*)

O BGCF também é um servidor SIP que executa funções de roteamento quando a chamada é dirigida a uma rede de comutação de circuitos, como a PSTN. Ele localiza o *gateway* apropriado na rede de destino para que a chamada seja roteada [16].

3.2.7 MRF (*Media Resource Function*)

Um MRF executa várias funções de mídia para a rede SIP, tais como misturar fluxos de mídia (em uma conferência), funções de transcodificação, fazendo anúncios durante um *streaming* de mídia. Um MRF pode ser dividido em MRF *Controller* (MRFC) e MRF *Processor* (MRFP). O MRFC atua essencialmente como um SIP UA conversando com o S-CSCF, enquanto o MRFP executa todas as funções de mídia acima referidos [16].

Todos esses elementos da arquitetura IMS fazem uso do protocolo SIP para desempenharem suas funções dentro da rede.

Conclusão

O protocolo SIP não só possibilitou que a infraestrutura usada na rede de telefonia fosse trazida para a rede IP - diminuindo gastos com novas redes e criação de mais centrais dedicadas, trazendo todos os serviços e funcionalidades anteriormente oferecidos pelas operadoras - mas também trouxe a possibilidade da criação (pelas operadoras telefônicas) de novas funcionalidades, como portabilidade numérica, interceptação legal de chamadas e novos serviços, como IPTV (IP *Television*), através do tratamento de uma única sessão SIP.

Essa mudança do mundo telefônico para as redes IP também possibilitou o uso de computadores conectados à Internet para processar as chamadas telefônicas, substituindo as centrais telefônicas, reduzindo o espaço físico utilizado pelo *hardware*, e facilitando o desenvolvimento de novos serviços pelas operadoras, já que agora estes serviços nada mais são do que mais um *software* sendo executado no computador. Isso acrescentou inteligência aos serviços prestados pelas operadoras, gerando uma maior agilidade e qualidade no atendimento ao consumidor.

Para concluir, refira-se que esta monografia faz parte do trabalho de conclusão de curso de Engenharia de Computação, apresentando uma visão geral do protocolo SIP e seu funcionamento, bem como as facilidades que ele proporciona para o mundo da telefonia.

Referencias Bibliográficas

- [1] ITU-T, Recommendation Q.76 (02/95), General Recommendation on Telephone Switching and Signalling Functions and Information Flows for Services in the ISND, 1995.
- [2] FERRARI, A.M. Telecomunicações - evolução e revolução. São Paulo: Editora Érica, 1997, cap. 3.
- [3] Tanenbaum, D. e J. Wetherall. *Redes de Computadores*. 5ª ed., Pearson, 2011, p. 28.
- [4] ITU-T, Recommendation H.323 (12/2009), Packet-Based Multimedia Communications Systems, 2010.
- [5] Rosenberg, J., e H. Schulzrinne, "A Comparison of SIP and H.323 for Internet Telephony," Network and Operating System Support for Digital Audio and Video (NOSSDAV), Cambridge, England, 1998.
- [6] Heath, M. e Brydon, A., Delivering Strategic Benefits with IP Multimedia Subsystem (IMS), Analysys Research Limited, Cambridge, 2005.
- [7] ITU-T, Recommendation Q.784.1 (07/1996), ISUP basic call test specification: Validation and compatibility for ISUP'92 and Q.767 protocols, 1999.
- [8] ITU, International Telecommunication Union - BDT, <<http://www.itu.int/net4/itu-d/icteye/>>, Acessado em: 03/10/2014.

- [9] Sinnreich, H. e Johnston, A. B., Internet Communications Using SIP: Delivering VoIP and Multimedia Services with Session Initiation Protocol, 2^a ed., Wiley, 2006.
- [10] Johnston, A. B., Understanding the SIP, 2^a ed., Artech House, 2004.
- [11] IETF, RFC 3261, SIP: Session Initiation Protocol, 2002.
- [12] ITU-T, Recommendation Q.931 (05/1998), ISDN user-network interface layer 3 specification for basic call control, 1999.
- [13] IETF, RFC 2327, SDP: Session Description Protocol, 1998.
- [14] IETF, RFC 3264, An Offer/Answer Model with the Session Description Protocol (SDP), 2002.
- [15] ITU, NGN Working Definition, <http://www.itu.int/ITU-T/studygroups/com13/ngn2004/working_definition.html>, Acessado em: 05/10/2014.
- [16] Poikselka, M. e Mayer, G., THE IMS - IP MULTIMEDIA CONCEPTS AND SERVICES, 3^a ed., Wiley, 2009.