

Introdução ao protocolo SIP*

1. SIP (Session Initiation Protocol)

Pode-se dizer que SIP trata-se de um protocolo de controle referente à camada de aplicações do Modelo de Referência OSI (*Open System Interconnection*), que é usado para iniciar, modificar ou terminar sessões ou chamadas multimídia entre usuários. Dentre suas funcionalidades tem-se a localização de usuários, o estabelecimento de chamadas, o suporte a *unicast* ou *multicast*, administração na participação de chamadas (transferências, conferência, entre outros) e possibilidade de participação de um usuário em terminal H.323, via *gateway*. É um protocolo cliente-servidor similar ao HTML no tocante à sintaxe e semântica das estruturas empregadas, com campos explicitamente descritos.

Esta solução de videoconferência, estabelece, modifica e termina sessões multimídia e/ou ligações. Estas sessões podem ser conferências multimídia, aulas pela Internet, telefonia sobre Internet, entre outras. O protocolo SIP é baseado no HTTP e, assim como este, suporta o transporte de qualquer tipo de carga em seus pacotes, pelo uso de *Mime-Types* (*Multipurpose Internet Mail Extensions*). Por utilizar uma arquitetura cliente/servidor, suas operações envolvem apenas métodos de requisição e respostas, como observado também no HTTP e no RTSP.

Como mencionado anteriormente, o fato do SIP tratar-se de um protocolo cliente-servidor (o cliente realiza chamadas que são atendidas pelo servidor), em alguns casos, uma chamada pode envolver diversos servidores e clientes. Um protocolo como o SIP, deve oferecer funções básicas como:

- Conversão de nomes e localização de usuários:

Envolve o mapeamento entre nomes de diferentes tipos de abstração, tais como nomes de um domínio e o nome de um usuário em servidor *Internet*, isto é necessário para que um determinado usuário, que possui um nome qualquer, possa ser convertido em um endereço IP, de modo que possa ser localizado a qualquer momento da ligação;

- Negociação de configuração:

* Adaptado do documento:

TAROUCO, L.; FABRE, M.; GRANVILLE, L.; TAMUSIUNAS, F. Videoconferência, RNP – Rede Nacional de Pesquisa

Permite que um grupo de usuários defina que tipo de informação será trocada e seus respectivos parâmetros. O conjunto e o tipo dos dados que estão sendo enviados não precisam ser uniformes dentro de uma chamada, como diferentes conexões ponto a ponto pode envolver diferentes tipos e parâmetros de dados. Muitos codecs são capazes de receber diferentes tipos de codificações, sendo restritos ao enviar apenas um tipo de dado em cada fluxo;

- Alteração de configuração:

Torna possível a alteração, de maneira dinâmica, ou seja, durante a utilização de uma conexão por seus usuários, dos parâmetros definidos no momento do estabelecimento da conexão.

1.1. Componentes da arquitetura

SIP foi concebido na Universidade de Columbia e depois submetido para aprovação do IETF (*Internet Engineering Task Force*). Foi aprovado como RFC (*Request For Comment*), sendo publicado como RFC 2543, em março de 1999. O IETF define um conjunto de componentes na sua arquitetura operando numa rede IP, este conjunto é definido como “rede” SIP. Estes componentes são apresentados na figura 1.1:

* Adaptado do documento:

TAROUCO, L.; FABRE, M.; GRANVILLE, L.; TAMUSIUNAS, F. Videoconferência, RNP – Rede Nacional de Pesquisa

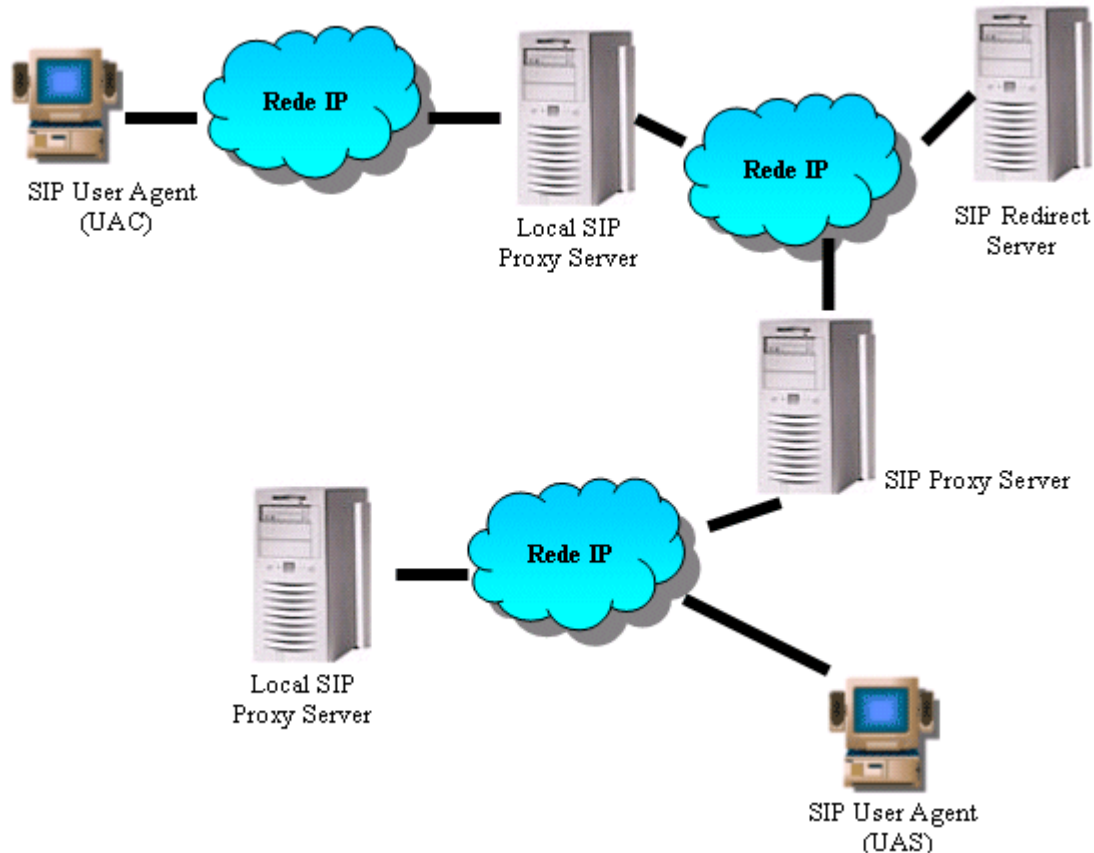


Figura 1.1 - Componentes da arquitetura SIP

- SIP User Agent

Cliente da arquitetura, ou o ponto final da comunicação multimídia que interage com o usuário. Um *user agent* tem dois componentes, um *user agent client* (UAC) e um *user agent server* (UAS). O UAC é responsável por iniciar as chamadas enviando requisições, e o UAS é responsável por responder às chamadas, enviando respostas. Uma aplicação de telefonia Internet por exemplo, contém ambos UAC e UAS. Pode-se destacar a diferença com um *Web Browser* que contém apenas um cliente.

- SIP Proxy Server

* Adaptado do documento:

TAROUCO, L.; FABRE, M.; GRANVILLE, L.; TAMUSIUNAS, F. Videoconferência, RNP – Rede Nacional de Pesquisa

Servidor de redirecionamento de requisições e respostas SIP. Passa a realizar a sinalização como se fosse o originador da chamada, e quando a resposta lhes é enviada, ela é redirecionada para o originador real. O servidor de *proxy* também é conhecido como *next-hop*. O servidor *next-hop* pode ser outro servidor de *proxy*, um UAS ou um servidor de redirecionamento.

Um exemplo da utilização do *proxy* pode ser verificado na Figura 1.2:

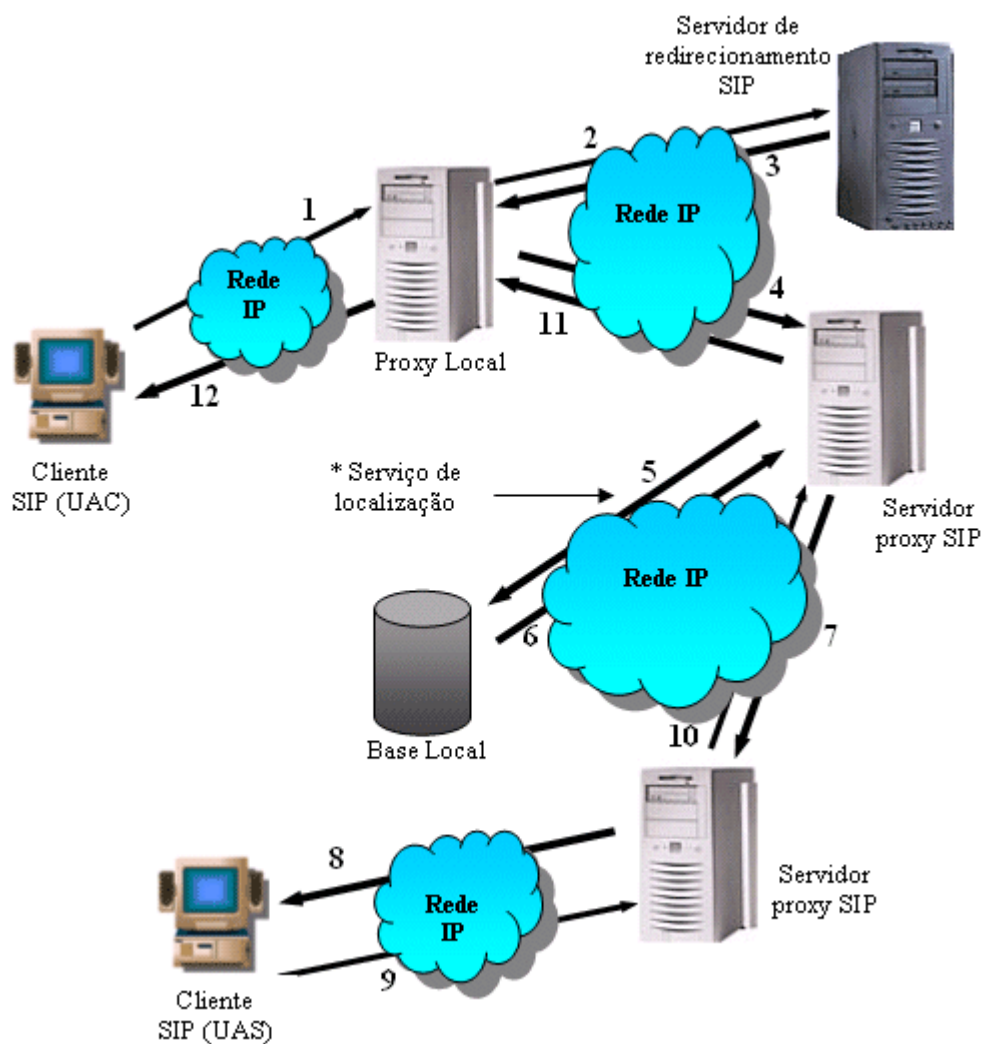


Figura 1.2 - Servidor de *proxy*

O primeiro passo que acontece na figura anterior, é quando o usuário encaminha um *invite request* para um *proxy* local, tendo como destino

* Adaptado do documento:

TAROUCO, L.; FABRE, M.; GRANVILLE, L.; TAMUSIUNAS, F. Videoconferência, RNP – Rede Nacional de Pesquisa

paulo@empresa.com (1). O *proxy* procura empresa.com no DNS (*Domain Name System*) e obtém o endereço IP do servidor que trata requisições SIP deste domínio. O *proxy* então encaminha o *request* para esse servidor (2). O servidor de empresa.com conhece o usuário paulo, mas ele está “logado” no momento como n.aluno@universidade.edu. Então, o servidor redireciona o *proxy* (3) para tentar este endereço. Também com auxílio do DNS, o *proxy* obtém o endereço IP do servidor responsável por universidade.edu e envia o *request* para esse servidor (4). O servidor consulta uma base local (5) e indica que n.aluno@universidade.edu é localmente identificado como paulo.silva@inf.universidade.edu (6). O servidor principal de universidade.edu envia o *request* para o servidor da informática (7), que por sua vez, sabendo o endereço IP do usuário, envia o *request* para ele (8). O usuário aceita a chamada e responde retornando pela cadeia de *proxy* (8), (10), (11) e (12).

- SIP Redirect Server

Recebe requisições e determina um servidor *next-hop*. Todavia, ao contrário de reenviar neste caso, ele retorna o endereço do servidor *next-hop* para o cliente. A função primária dos servidores *proxy* e de redirecionamento são roteamento de chamadas, ou seja a determinação do conjunto de servidores a serem usados no caminho para completar a chamada. Um servidor de *proxy* ou um servidor de redirecionamento pode usar qualquer meio para determinar o servidor *next-hop*, incluindo executar programas e consultar banco de dados. Um servidor *proxy* pode também duplicar uma requisição, enviando cópias para múltiplos servidores *next-hop* de uma vez. Isto permite que uma requisição de início de chamada tente diversas localizações diferentes ao mesmo tempo. A primeira localização que responder é conectada com o cliente chamador.

Em suma, redireciona requisições e respostas, enviando uma mensagem para os clientes como novo endereço SIP procurado, e não fazendo o papel de continuar a chamada.

A arquitetura pode apresentar ainda os seguintes componentes:

- SIP Register Server

Servidor SIP que suporta requisições REGISTER, usadas para registrar as informações dos usuários em algum Servidor de Localização.

* Adaptado do documento:

TAROUCO, L.; FABRE, M.; GRANVILLE, L.; TAMUSIUNAS, F. Videoconferência, RNP – Rede Nacional de Pesquisa

- Servidor de Localização

Possui apenas as funcionalidades de armazenamento e consulta de registros de usuários SIP neste servidor são descritas, ficando a critério do implementador da solução SIP a escolha da melhor tecnologia para esta finalidade.

1.2. Sinalização no SIP

O protocolo SIP é baseado no HTTP e, assim como este, suporta o transporte de qualquer tipo de carga em seus pacotes, pelo uso de Mime-Types (Multipurpose Internet Mail Extensions). O SIP funciona numa arquitetura cliente/servidor, e suas operações envolvem apenas métodos de requisição e respostas, como observado também no HTTP e no RTSP. Os métodos de requisição do SIP são os seguintes:

- INVITE: Indica que o usuário está sendo convidado a participar de uma sessão multimídia. O corpo da mensagem pode conter uma descrição da sessão, utilizando-se o protocolo de descrição de sessão SDP (*Session Description Protocol*).
- ACK: Mensagem recebida como resposta final a um INVITE. A requisição ACK pode conter o SDP de descrição da sessão negociada entre ambos os clientes. Se não contiver o SDP, o usuário chamado pode assumir a descrição dada pelo primeiro INVITE, se houver.
- OPTIONS: Faz uma pergunta sobre quais métodos e extensões são suportados pelo servidor e pelo usuário descrito no campo de cabeçalho <To:>. O servidor pode responder a esta pergunta com o conjunto de métodos e extensões suportado pelo usuário e por ele mesmo.
- BYE: Usado para liberar os recursos associados a uma ligação e forçar a desconexão da mesma.
- CANCEL: Cancela uma requisição que ainda esteja pendente, ou seja, em andamento. Uma requisição é considerada pendente, se e somente se, ela não foi atendida com uma resposta final.

* Adaptado do documento:

TAROUCO, L.; FABRE, M.; GRANVILLE, L.; TAMUSIUNAS, F. Videoconferência, RNP – Rede Nacional de Pesquisa

- REGISTER: Um cliente usa este método para registrar o "alias" (apelido) do seu endereço em algum servidor SIP, que, por aceitar registro de usuários, chamamos de serviço REGISTRAR.

Para cada requisição ou resposta, temos um grupo de cabeçalhos, divididos em: cabeçalhos gerais, com informações importantes sobre a chamada; cabeçalhos de entidade, com meta-informação sobre o corpo da mensagem; e os cabeçalhos específicos, que permitem passar informações adicionais, que não couberam na linha de status da requisição ou da resposta.

As entidades SIP se comunicam usando “transações”. Quando requisições (transações) são atendidas, as respostas enviadas são identificadas por números, que significam a classe da resposta. Pode-se enviar diversas mensagens provisórias antes de se enviar uma resposta definitiva. Existem seis classes possíveis de resposta: Classe 1XX, respostas temporárias ou informativas; Classe 2XX, resposta final de sucesso; Classe 3XX, redirecionamento da requisição; Classe 4XX, erros no cliente; Classe 5XX, erros do servidor; e Classe 6XX, erros globais na rede.

O iniciador de um pedido SIP é chamado de cliente SIP e a entidade que responde é chamada de servidor SIP. As mensagens trocadas durante uma transação compartilham um número Cseq (campo de cabeçalho) comum, com uma exceção: a mensagem ACK usa o mesmo Cseq da transação à qual ela se aplica, mas é uma transação diferente.

A primeira etapa consiste em abrir uma conexão de sinalização entre os pontos de origem e de destino da chamada. Pontos finais SIP podem usar sinalização UDP ou TCP, a sintaxe das mensagens é independente do protocolo de transporte usado.

Quando se usa o TCP, a mesma conexão pode ser usada para todos os pedidos e respostas SIP (não par os dados de mídia) ou uma nova conexão TCP pode ser usada para cada transação. Se o UDP for usado, o endereço e a porta a serem usados para as respostas a pedidos SIP estarão contidos no parâmetro de cabeçalho ‘via’ do pedido SIP. As respostas não devem ser enviadas para o endereço IP do cliente. Se nenhuma porta for especificada no endereço SIP, a conexão é feita com a porta 5060 tanto para o TCP quanto para o UDP.

Um cliente SIP chama um outro ponto final SIP enviando uma mensagem de pedido *invite* (Figura 1.1). A mensagem *invite* normalmente contém informações suficientes para permitir que o terminal de destino estabeleça imediatamente a conexão de mídia solicitada com o ponto de origem da chamada. Essas informações incluem as capacidades de mídia que o ponto de origem da chamada pode receber (e enviar: considera-se que a capacidade do codificador estão enviando ou recebendo no SIP, a menos que os parâmetros SDP *sendonly* ou *reconly* sejam usados) e o endereço de transporte onde o ponto de origem da

* Adaptado do documento:

TAROUCO, L.; FABRE, M.; GRANVILLE, L.; TAMUSIUNAS, F. Videoconferência, RNP – Rede Nacional de Pesquisa

chamada espera que o ponto de destino da chamada envie esses dados de mídia. A maioria dos pontos finais serão capazes de receber muitas codificações diferentes para cada tipo de mídia. A codificação em particular escolhida pelo remetente aparece como parte do cabeçalho RTP.

O ponto de destino da chamada precisa indicar que ele está aceitando o pedido. Esse é o objetivo da mensagem de resposta OK. Uma vez que o pedido foi um convite, a resposta OK também contém as capacidades de mídia do ponto de destino da chamada e onde ele está esperando receber os dados de mídia. O originador da chamada precisa confirmar que recebeu de maneira adequada a resposta do ponto de destino da chamada (lembre-se que podemos estar usando UDP) com a mensagem ACK.

A partir desse exemplo, pode-se notar que o SIP é eficiente, pois o canal de mídia do receptor para o originador da chamada pode ser estabelecido com uma ida e uma volta de mensagens e o canal de mídia do originador para o receptor pode ser estabelecido com uma ida, uma volta e outra ida de mensagens. Isto é muito melhor que as várias idas-e-voltas exigidas pelo H.323 v1.

* Adaptado do documento:

TAROUCO, L.; FABRE, M.; GRANVILLE, L.; TAMUSIUNAS, F. Videoconferência, RNP – Rede Nacional de Pesquisa

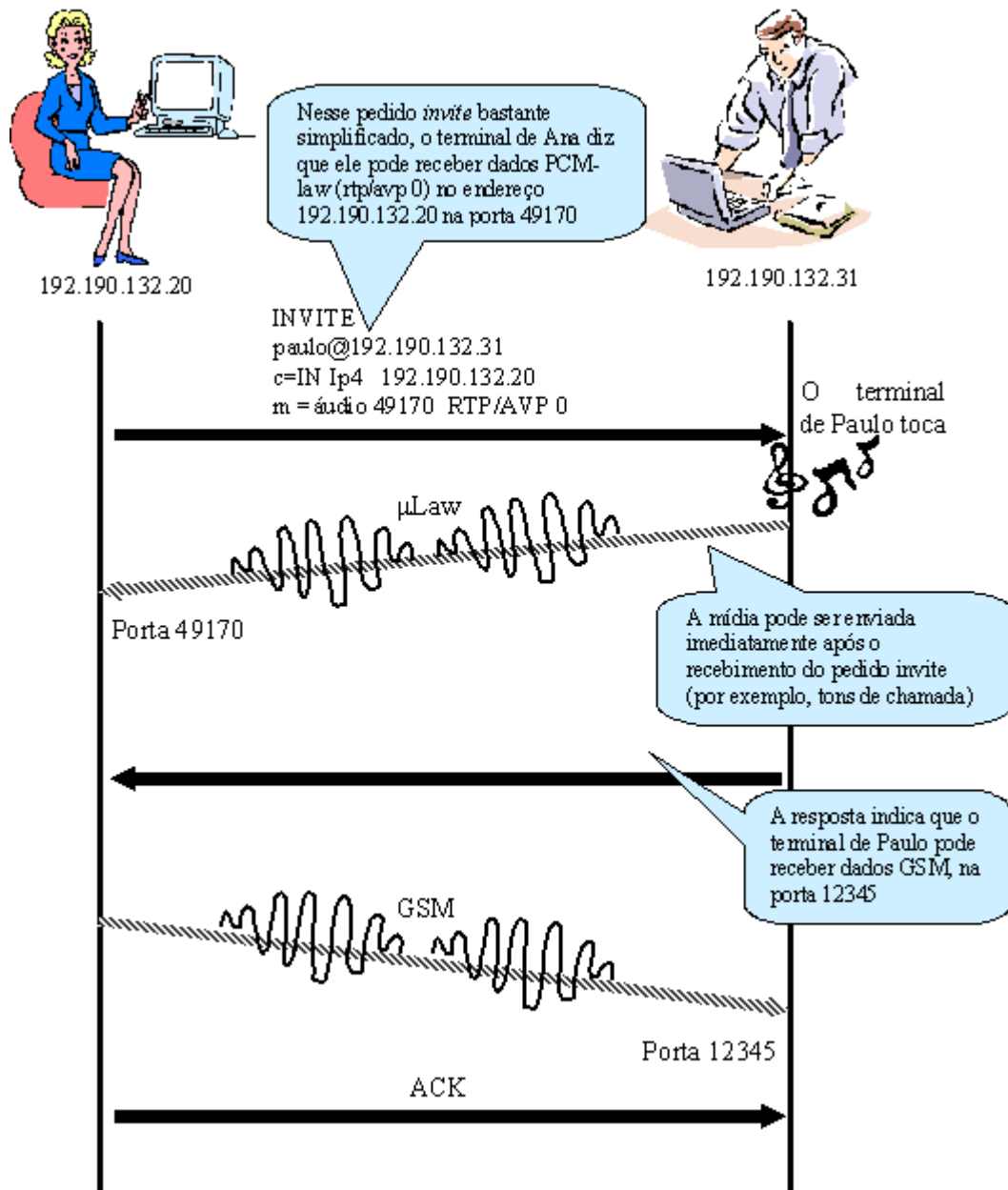


Figura 1.3 - Chamada para um endereço IP conhecido com SIP

Na chamada anterior, o terminal de Ana ofereceu-se para receber um canal de áudio codificado em PCM. Isto pode não ser aceitável para o terminal de Paulo, ou porque Paulo não dispõe de largura de banda suficiente (o PCM exige 64Kbit/S, mais o *overhead* RTP/UDP/IP/PPP) ou porque o terminal de Paulo não possui um codificador PCM m-law. Nesse caso, o terminal de Paulo responderá com um *606 Not Acceptable* e finalmente relacionará o conjunto de codificadores que ele pode usar. Com essa informação, o terminal

* Adaptado do documento:

TAROUCO, L.; FABRE, M.; GRANVILLE, L.; TAMUSIUNAS, F. Videoconferência, RNP – Rede Nacional de Pesquisa

de Ana pode enviar um novo pedido de *invite*, com o mesmo identificador de chamada, anunciando o código apropriado (mas se ele tivesse essa capacidade, poderia tê-la enviado como uma escolha no primeiro pedido *invite*) ou então reiniciar uma chamada por meio de um *proxy* de transcodificação (Figura 1.3).

* Adaptado do documento:

TAROUCO, L.; FABRE, M.; GRANVILLE, L.; TAMUSIUNAS, F. Videoconferência, RNP – Rede Nacional de Pesquisa

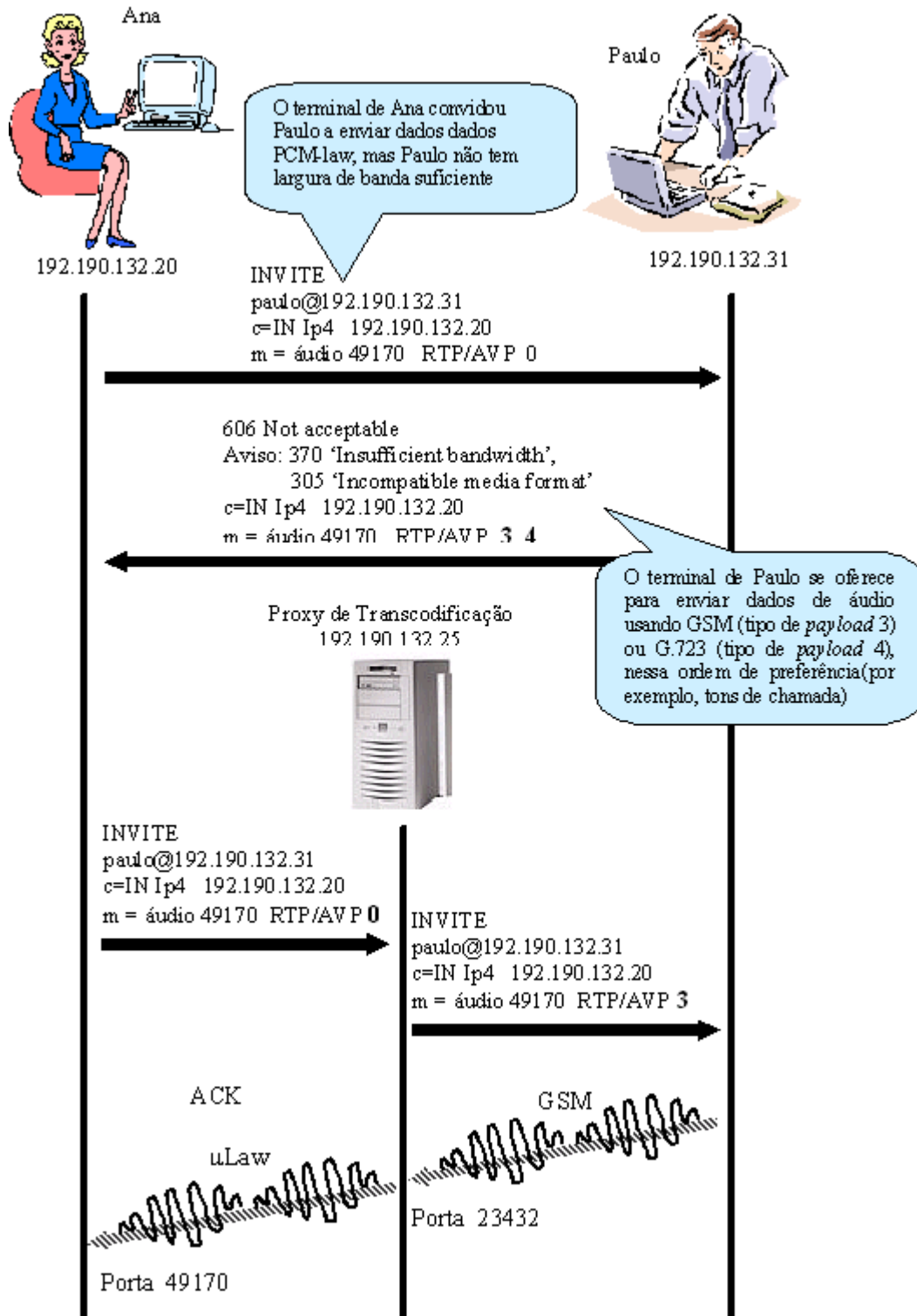


Figura 1.4 - Negociando o estabelecimento da chamada

* Adaptado do documento:

TAROUCO, L.; FABRE, M.; GRANVILLE, L.; TAMUSIUNAS, F. Videoconferência, RNP – Rede Nacional de Pesquisa

O IETF concluiu sobre um codificador padrão a ser usado no caso de voz a baixa taxa de bits, o G.723.1, de maneira a manter mínima a probabilidade desse tipo de incompatibilidade no H.323. Nenhuma recomendação desse tipo existe para o SIP, mas a maioria dos terminais parece ser capaz de receber PCM A-law e m-law e também GSM.

O SIP não tem uma noção de canais lógicos (como definido no H.323). Quando um cliente se oferece para receber vários tipos de mídias em várias portas UDP ou TCP, ele tem de estar preparado para receber mídia imediatamente em qualquer um dessas portas. No entanto, o terminal de destino pode decidir enviar dados em apenas algumas portas (por exemplo, ele não possui capacidades de vídeo e envia apenas áudio). Nada na sinalização informa ao cliente se uma determinada porta estará ativa ou não. Em geral, isso não é realmente um problema, uma vez que a maioria dos pontos finais podem ficar “escutando” as portas não usadas sem um impacto significativo no desempenho.

Em alguns casos, no entanto, as entidades SIP precisam manter vários canais de mídia e têm de reutilizar as portas UDP da maneira mais eficiente possível, este é o caso de grandes *gateways* ou recursos de mídia centralizados. Essas entidades podem ter de multiplexar vários canais de mídia em uma única porta ou fechar portas inativas com base em heurísticas, por exemplo, após um período sem nenhuma atividade.

* Adaptado do documento:

TAROUCO, L.; FABRE, M.; GRANVILLE, L.; TAMUSIUNAS, F. Videoconferência, RNP – Rede Nacional de Pesquisa