Voice over IP (VoIP) & Session Initiation Protocol (SIP)

Redes de Comunicações II

Licenciatura em Engenharia de Computadores e Informática

Prof. Amaro de Sousa (asou@ua.pt)

DETI-UA, 2024/2025

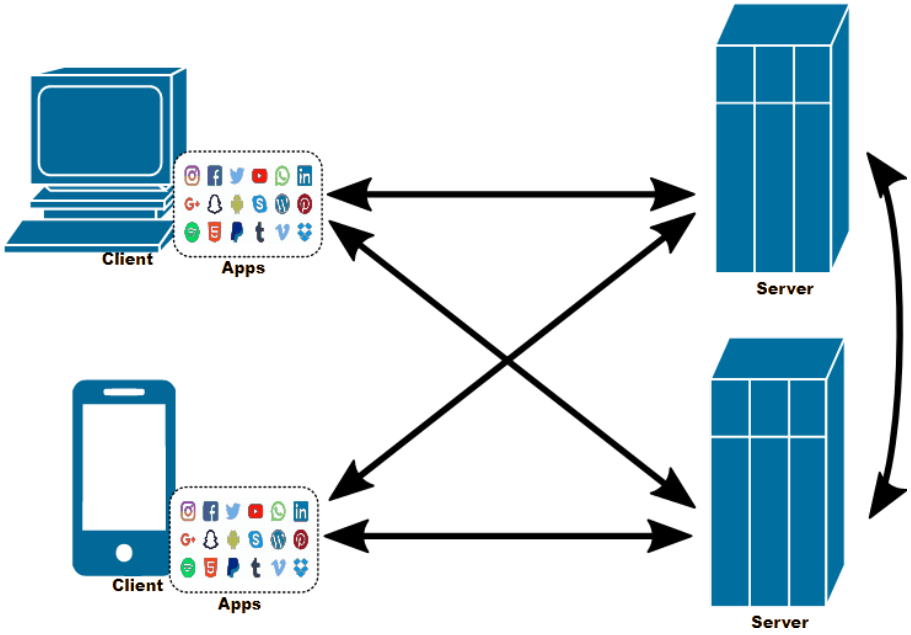
Voice over IP (VoIP)

- Voice over IP (VoIP) is a set of technologies used for voice communication sessions over IP networks.
- How VoIP works:
  1. Analog to Digital: VoIP voice is converted in the transmitter device (a computer or a VoIP phone) into a digital stream (using a codec) which is broken down into a sequence of small data packets.
  2. Transmission: the digital packets are transmitted over the IP network.
  3. Digital to Analog Reassembly: The receiver device reassembles the digital stream converting it back to the analogue voice (using the same codec of the transmitter).
- Typically, VoIP technologies detect silence periods at the transmitter and do not generate data packets during these periods (avoiding transmission of "noise")
- VoIP requires session establishment (for example, to agree on the codec to be used)

Quality of Service (QoS) issues with VoIP

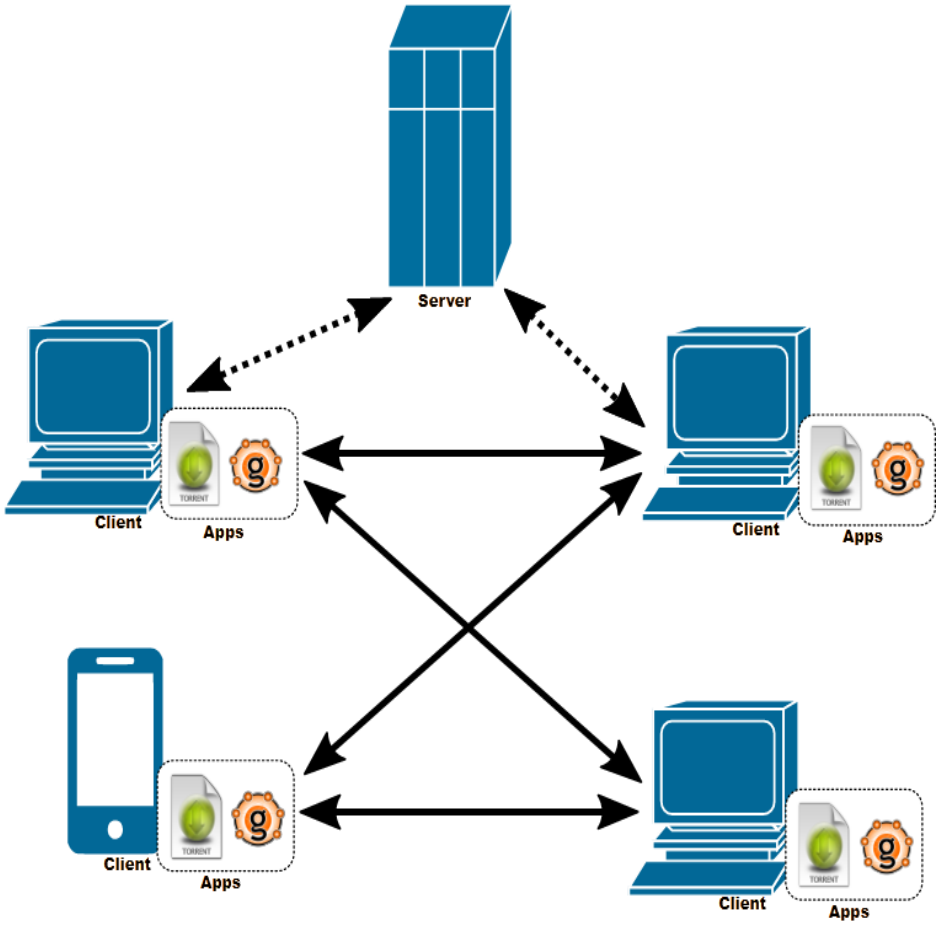
- When transmitted over an IP network, a stream of IP packets of a VoIP session suffer:
  - Packet Delays:
    - due transmission and propagation delays on links,
    - due to switching and queueing delays on routers
  - Packet Loss:
    - due to network congestion (router buffer overflow),
    - due to too late arrival of packets to the receiver for playout.
- Packet Delay tolerance:
  - a one-way delay (mouth-to-ear) should ideally be up to 150 milliseconds (ms) for good quality
- Packet Loss tolerance:
  - depending on the codec, packet loss rates between 0.1% and 1% can be tolerated

Client-Server Model



- Servers:
  - Always ON (i.e. always available to accept communications)
  - IP address is always known (usually through a DNS name)
  - May communicate between them (i.e., may act as clients between them).
- Clients:
  - Communicate with servers (i.e., they initiate the communications with servers).
  - Not always ON (i.e., are ON only when in operation).
  - May have dynamic addresses.
  - They do not communicate between themselves.

Peer-to-Peer (P2P) Model



- Clients:
  - Communicate between themselves.
  - Can be ON only when in operation.
  - May have dynamic addresses.
  - Peer discovery may be done within the P2P network or using central servers.
- Servers:
  - Usually they exist to support the operation of the P2P service.

Session Initiation Protocol (SIP)

- SIP is defined by RFC 3261, June 2002.
- SIP was designed for creating, modifying and terminating sessions between two or more participants.
  - Not limited to VoIP calls (can be multimedia, video-conference).
- SIP can be transported over UDP or TCP protocols.
- It is an alternative to the complex H.323 family of protocols.
  - Due to its simplicity, SIP is much more popular than H.323.
- It provides a peer-to-peer service between endpoints. Peers in a session are called user agents (UAs):
  - User-agent client (UAC): a client application that initiates the request of a SIP session.
  - User-agent server (UAS): a server application that receives and reacts to SIP session requests.
  - A SIP endpoint works both as a UAC and a UAS.

SIP functionalities

- SIP supports five facets of establishing and terminating multimedia communications:
  - User location - determination of the end system to be used for communication
  - User availability - determination of the willingness of the called party to engage in communications
  - User capabilities - determination of the media (audio/video) and media parameters to be used
  - Session setup - "ringing", establishment of session parameters at both called and calling party
  - Session management - including transfer and termination of sessions, modifying session parameters, and invoking services

SIP as a component of a multimedia architecture

- SIP does not provide an integrated communications system.
- SIP is a component to be used with other protocols (to set up a complete multimedia architecture), such as:
  - Real-time Transport Protocol (RTP) for transporting real-time data (audio/video) and providing QoS feedback
  - Real-Time Streaming Protocol (RTSP) for controlling the delivery of streaming media
  - Media Gateway Control Protocol (MEGACO) for controlling gateways to the Public Switched Telephone Network (PSTN)
  - Session Description Protocol (SDP) for describing multimedia sessions.
- SIP is to be used in combination with other protocols to provide complete services to users, but the basic functionalities and operation of SIP do not depend on any of the other protocols.

SIP clients and servers

- SIP Clients
  - Phones (software based or hardware) and Gateways
  - User Agents, that can act as a
    - Client when they initiate a request (UAC),
    - Server when they receive and react to a request (UAS).
- SIP Servers
  - Proxy server
    - Receives SIP requests from a client and forwards them on the client's behalf.
    - Receives SIP messages and forward them to the next SIP server in the network.
    - Provides functions such as authentication, authorization, network access control, routing, reliable request retransmission, and security.
  - Redirect server
    - Provides the client with information about the next hop or hops that a message should take and then the client contacts the next-hop server or UAS directly.
  - Registrar server
    - Processes requests from UACs for registration of their current location.
    - Registrar servers are often co-located with a redirect or proxy server.

SIP messages

- SIP is used for Peer-to-Peer Communications, though it uses a Client-Server model.
- SIP is a text-based protocol (similar to HTTP) and uses the UTF-8 charset.
- A SIP message is either a **Request** from a client to a server, or a **Response** from a server to a client.
  - A Request message consists of a **Request-Line**, one or more Header Fields, an empty line (indicating the end of the Header Fields), and an optional message-body;
  - A Response message consists of a **Status-Line**, one or more Header Fields, an empty line (indicating the end of the Header Fields), and an optional message-body.
  - Similarly to HTTP, all lines (including empty ones) are terminated by a "carriage-return" "line-feed" character sequence (CRLF).

SIP Request messages

- SIP uses SIP Uniform Resource Indicator (URI) to indicate the requested user or service. The general form of a SIP Request-URI is:
  - `sip:user:password@host:port;uri-parameters`
    - `sip:John@doe.com`
    - `sip:+14085551212:pass1234@195.2.43.29:5060`
    - `sip:alice@atlanta.com;maddr=239.255.255.1;ttl=15`
- A Request message contains one of a possible set of methods:
  - RFC 3261 defines six methods: INVITE, ACK, OPTIONS, BYE, CANCEL, and REGISTER.
  - SIP extensions provide additional methods: SUBSCRIBE, NOTIFY, PUBLISH, MESSAGE, ...
- The Request-Line of a Request message includes a Method, a Request-URI, and SIP-Version separated by a single space:
  - Example: `INVITE sip:John@doe.com SIP/2.0`

Methods in SIP Request messages

Basic Methods:

- INVITE – to establish a session
- ACK – to confirm an INVITE request
- BYE – to end a session
- CANCEL – to cancel establishing a session
- REGISTER – to communicate user location (host name, IP)
- OPTIONS – to communicate information about the capabilities of the calling and receiving SIP terminals

Some Extended Methods:

- SUBSCRIBE – to subscribe for notification from the server
- NOTIFY – to notify the subscriber of a new event
- PUBLISH – to publish an event to the server
- MESSAGE – to transport Instant Messages

SIP Response messages

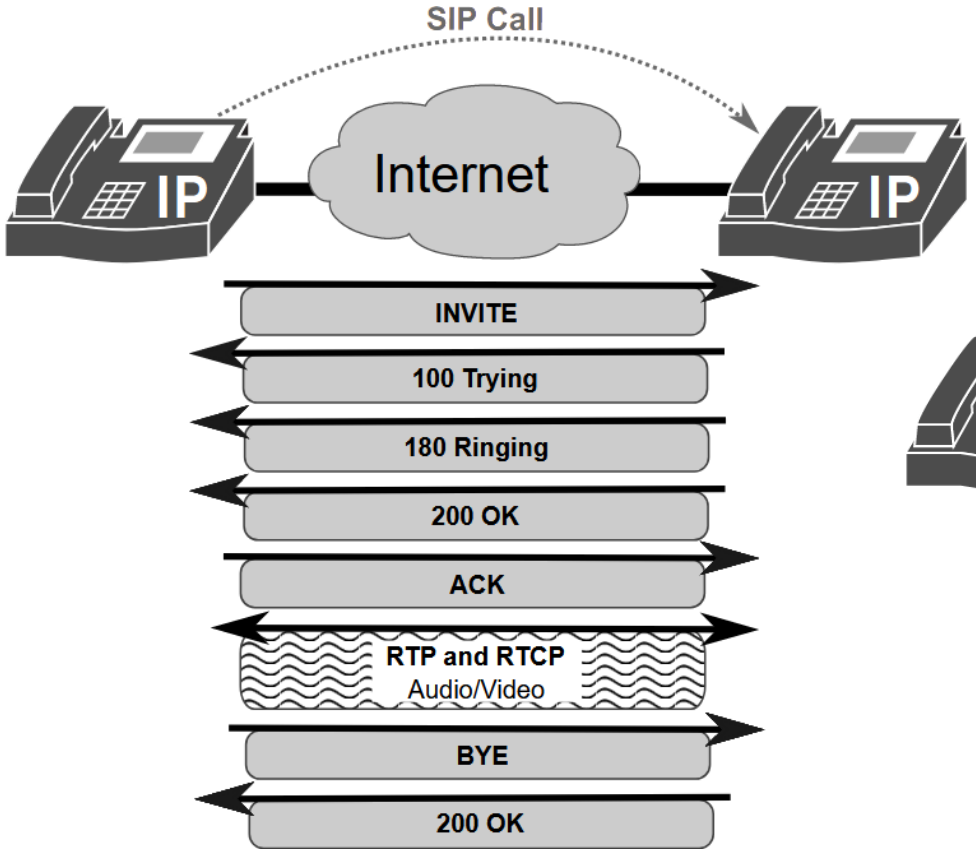
SIP Request messages are replied with SIP Response messages, whose Status-Line is a code of 3 digits. There are six classes of codes:

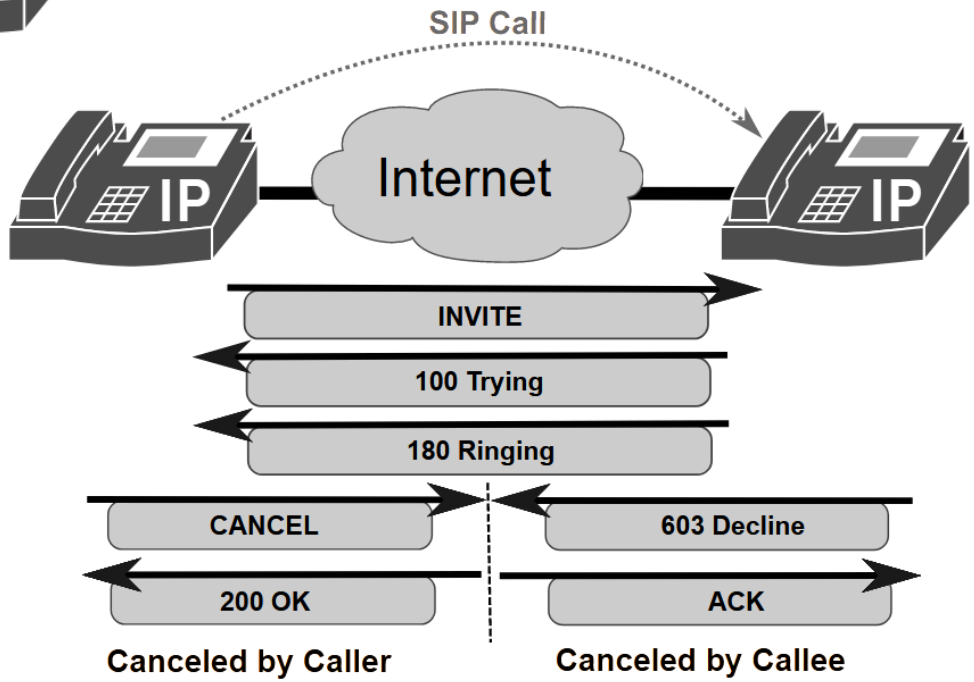
- 1xx – Informational responses, examples:
  - `100 Trying`
  - `180 Ringing`
- 2xx – Success responses, example:
  - `200 OK`
- 3xx – Redirection responses
- 4xx – Request failures, examples:
  - `401 Unauthorized`
  - `403 Forbidden`
- 5xx – Server errors
- 6xx – Global failures

Session Description Protocol (SDP)

- SIP carries (encapsulates) SDP messages.
- When initiating SIP sessions (multimedia teleconferences, VoIP calls, video streaming, or other sessions), participants must know from each other information about media details, transport addresses, and other session description metadata.
- SDP provides a standard representation of such information, irrespective of how the information is transported between participants.
  - SDP is only a format for session description.
  - SDP is intended to be general purpose so that it can be used in a wide range of network environments and applications.
  - SDP does not support negotiation of session content or media encodings.
- SDP is defined by RFC 4566, July 2006.

SIP signalling in direct calls

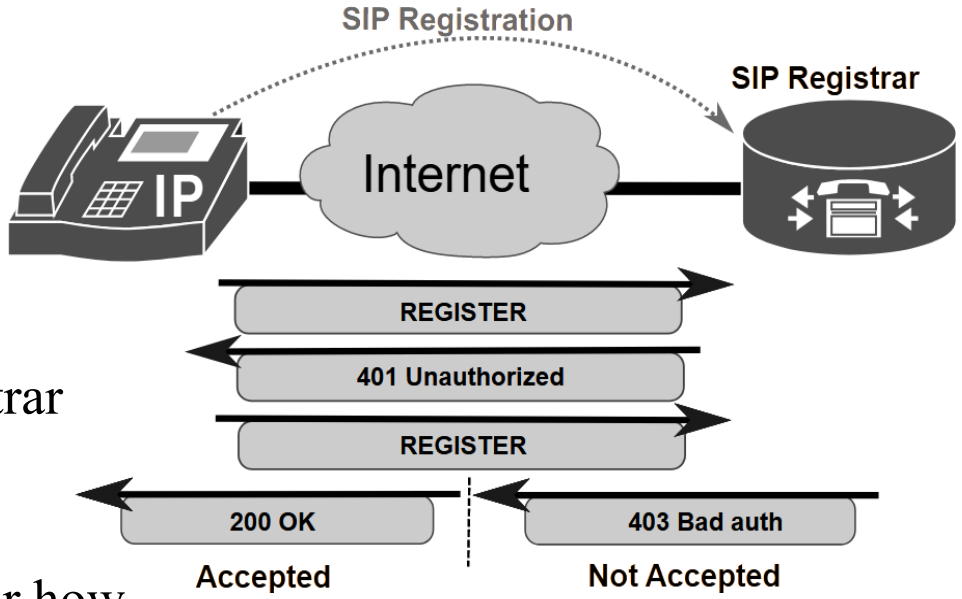




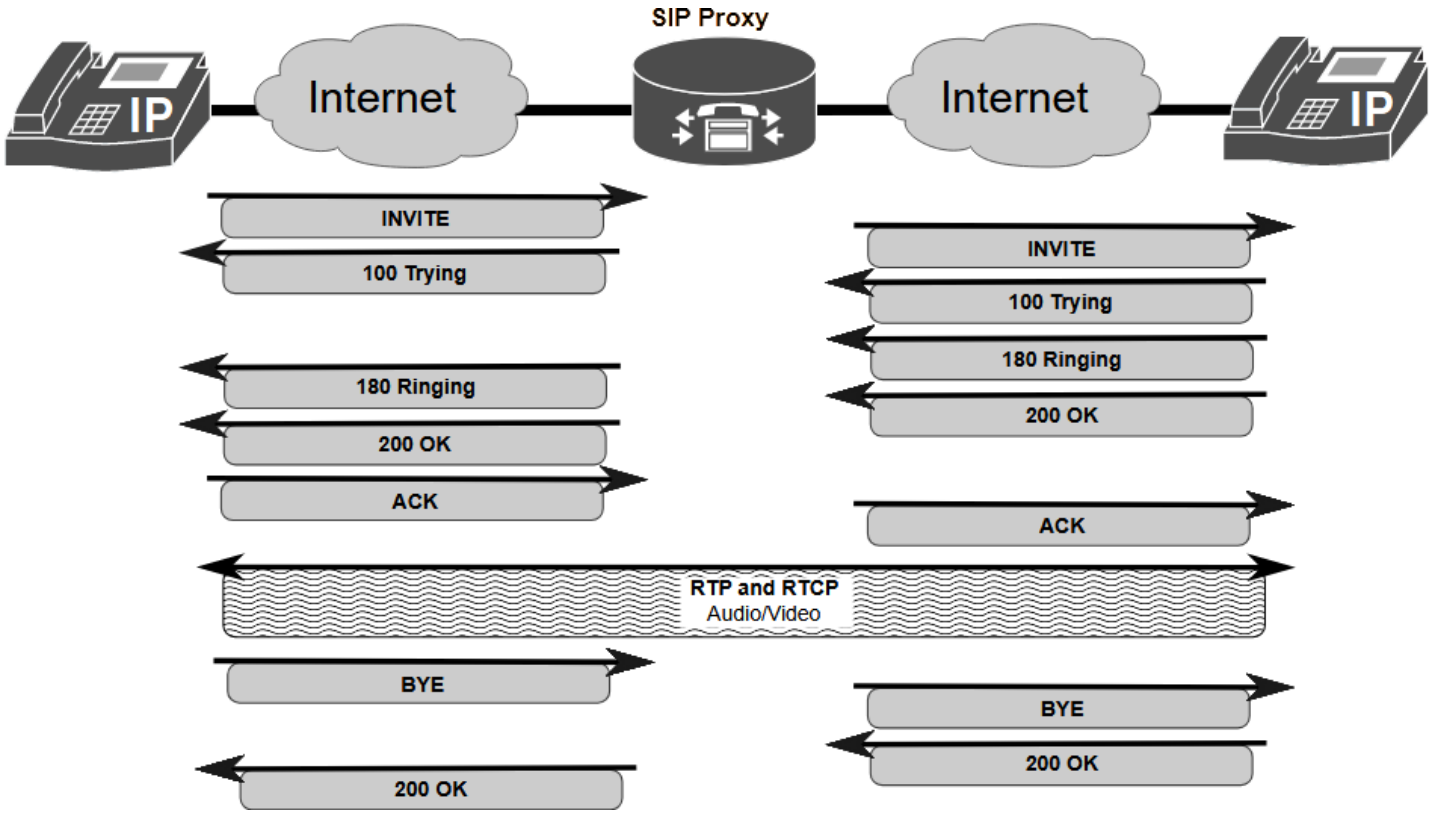
- Issues:
  - The caller must know the IP address of the callee
  - It fails if the callee has a private IP address (i.e., behind a NAT gateway)

SIP Registrar server

- SIP Registrar servers store the location of SIP endpoints.
- A user has an account created which allows them to REGISTER contacts with a particular server.
- SIP Proxy servers query SIP Registrar servers for routing information.
- The REGISTER message has the header field `Expires:` to define for how much time the registration should be valid.
  - A REGISTER message with the header field `Expires:0` deregisters the user.
- If REGISTER has no authentication credentials, the SIP Registrar server responds with `401 Unauthorized.`
  - End-point resends REGISTER with a header field `Authorization` containing the encrypted user password
  - Server accepts registration with the response `200 OK`, or rejects the registration with the response `403 Forbidden`



SIP signalling using Proxy server



- In this case, the users of both SIP endpoints have previously registered in the same domain:
  - For example: from `John.Smith@doe.com` to `Ed.Andersson@doe.com`
- and, so, the Proxy server of the domain can obtain from the Registrar server the IP addresses of the SIP endpoints.

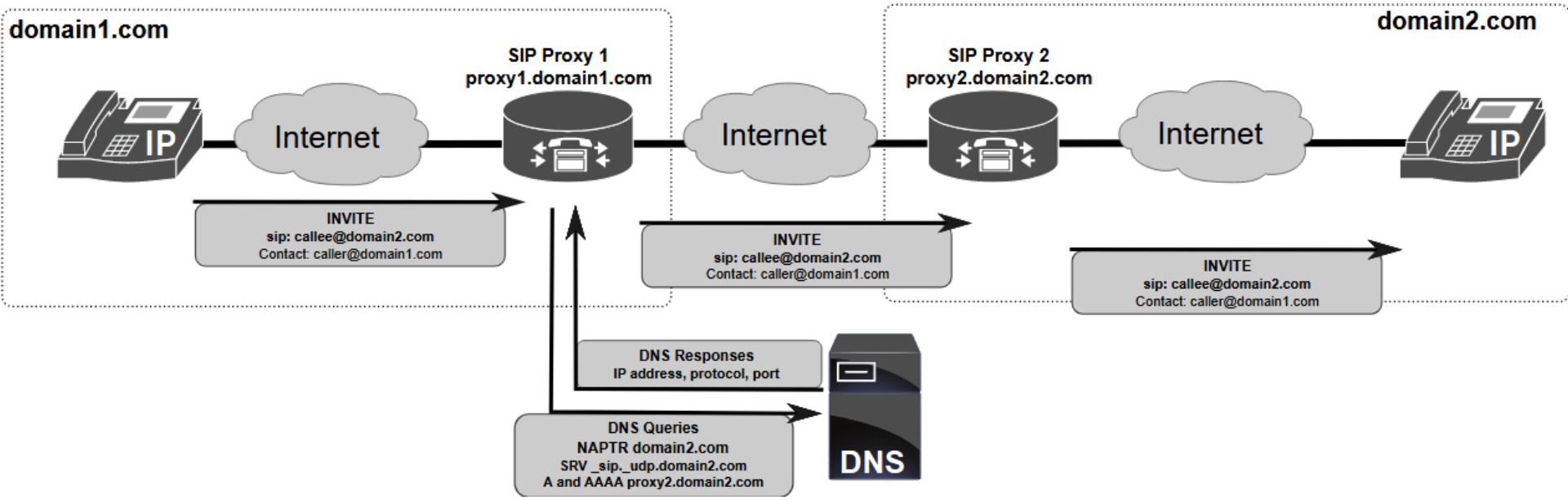
Locating SIP servers

- RFC 3263 (June 2002) defines the DNS procedures to locate SIP Servers.
- When the callee of a SIP session (identified by a SIP URI) is not in the same domain of the caller, the caller's SIP server must identify the callee's SIP server.
  - The caller's SIP server need to know the Transport protocol, IP address and Port number to be used to forward SIP messages to the callee server.
    - If the URI specifies any of them, then they should be used.
    - Otherwise, this information must be retrieved from DNS using the **Service (SRV)** and **Name Authority Pointer (NAPTR)** DNS records.
  - NAPTR DNS records provide a mapping from a domain name to an SRV record and a specific transport protocol.
  - SRV DNS records provide the name(s) of the responsible domain server(s).

Locating SIP servers (example)

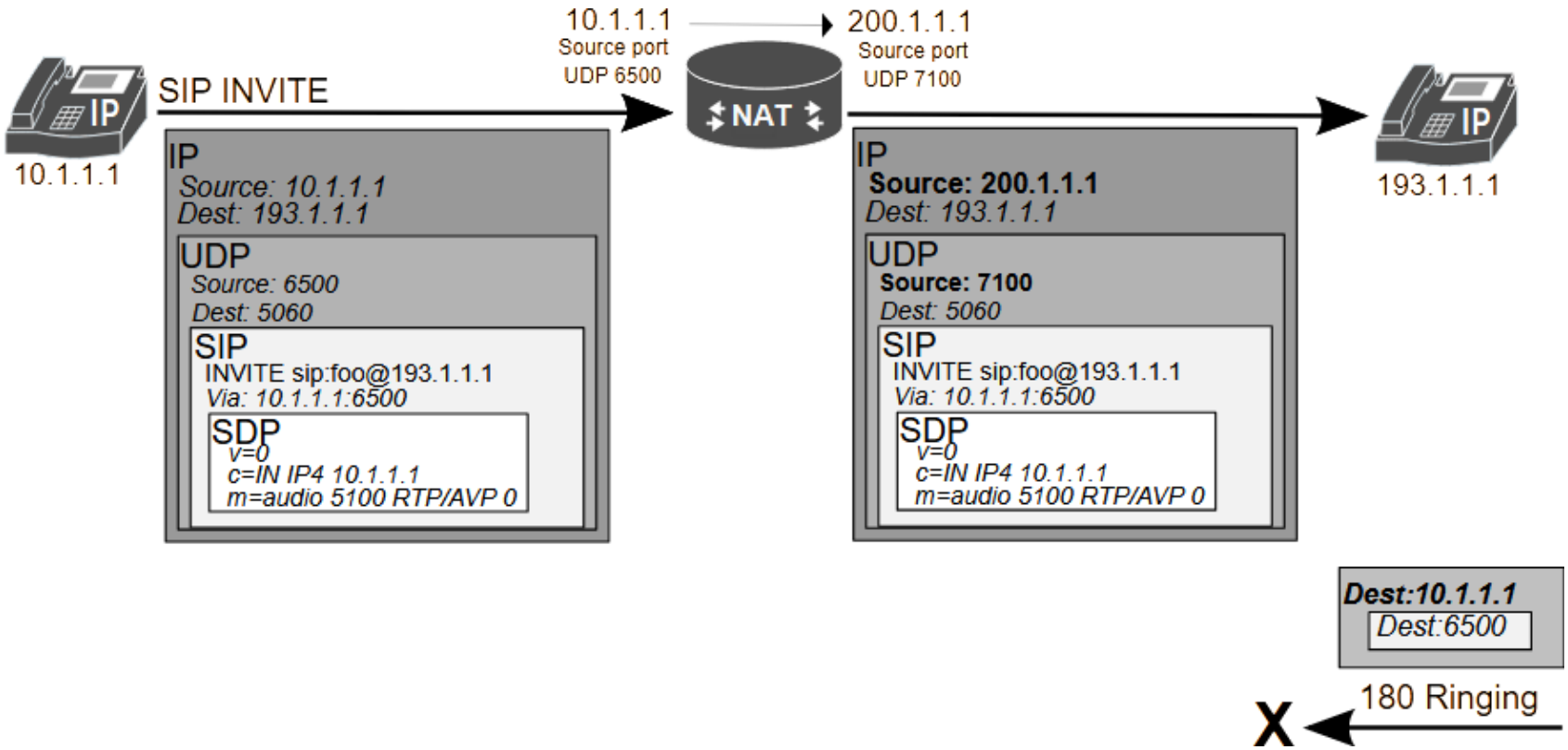
- A client/server wishes to resolve `sip:user@example.com`
- First, it sends a NAPTR DNS query for domain "`example.com`" and the DNS reply is:
  - `IN NAPTR 100 50 "s" "SIP+D2U" "" _sip._udp.example.com`
- it obtains UDP as the transport protocol to be used.
- Then, it sends an SRV DNS query for "`_sip._udp.example.com`" and the DNS reply is:
  - `IN SRV 0 1 5060 server1.example.com`
  - `IN SRV 0 2 5060 server2.example.com`
- it obtains two possible server names and the Port number used by each of them (in this case, 5060 for both).
- Finally, it sends A and/or AAAA queries for the chosen server name to obtain its IPv4 and/or IPv6 addresses.

SIP Proxy forwarding illustration



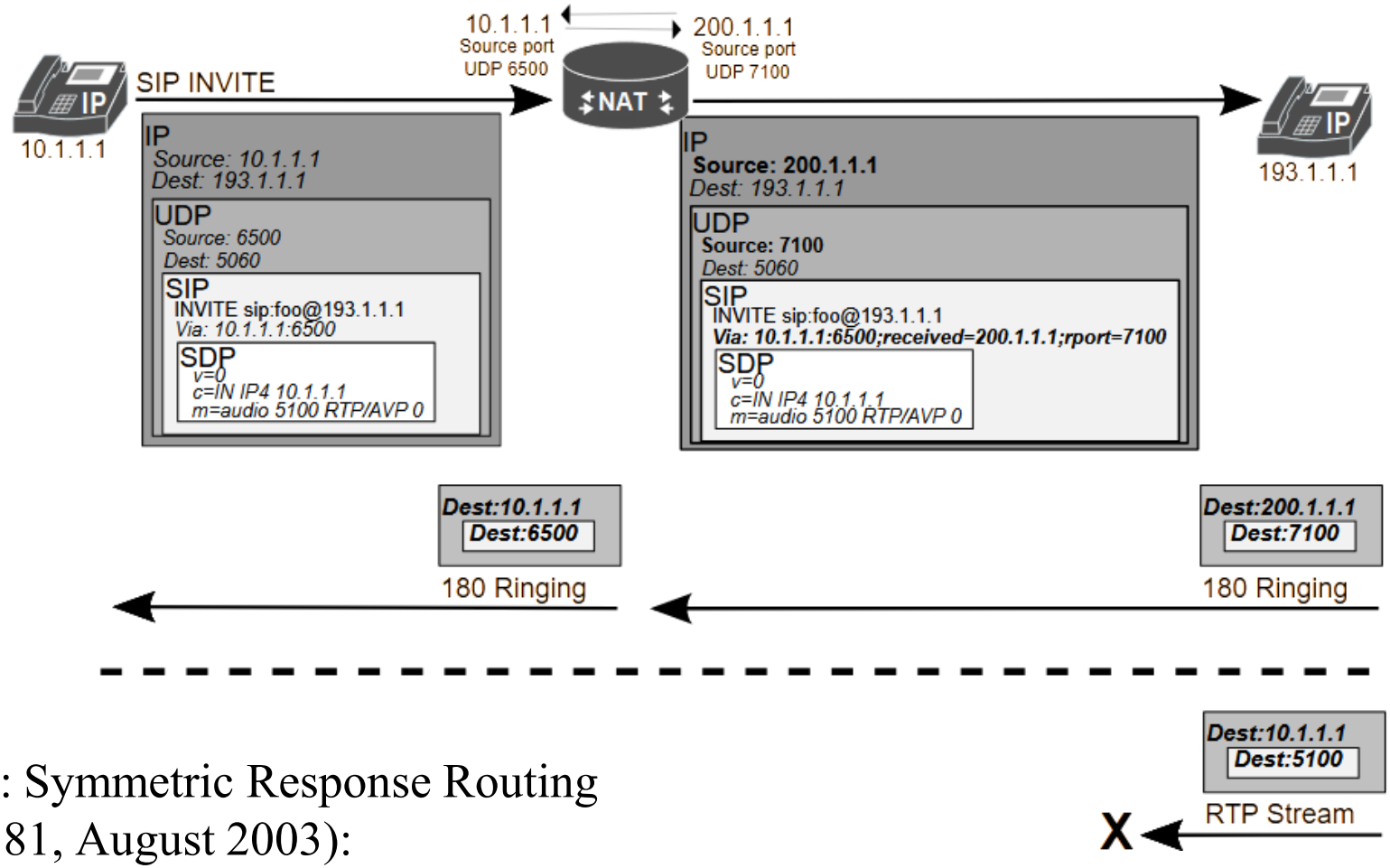
- In this case, the user of each SIP endpoint has previously registered on its domain.
- When the INVITE reaches SIP Proxy 2, it obtains from its Registrar server the IP address of its SIP endpoint.

SIP and NAPT



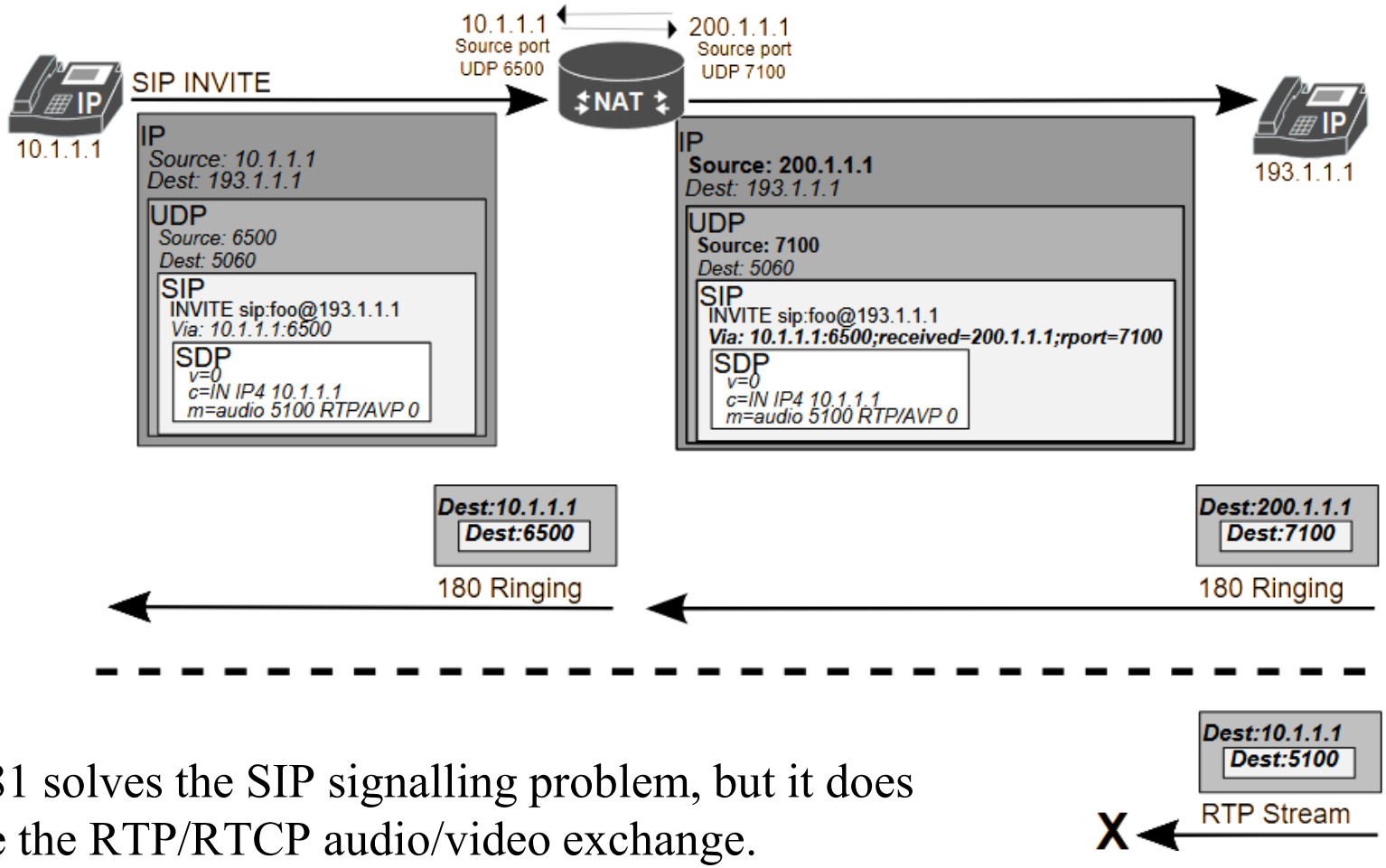
- By default, if the caller is in a private IP network, the NAT translation does not allow successful SIP signalling
  - the caller information in the Header Field "`Via:`" of the SIP Request message is used to reply with the SIP Response message of the callee.

SIP and NAPT: solution for the SIP signalling



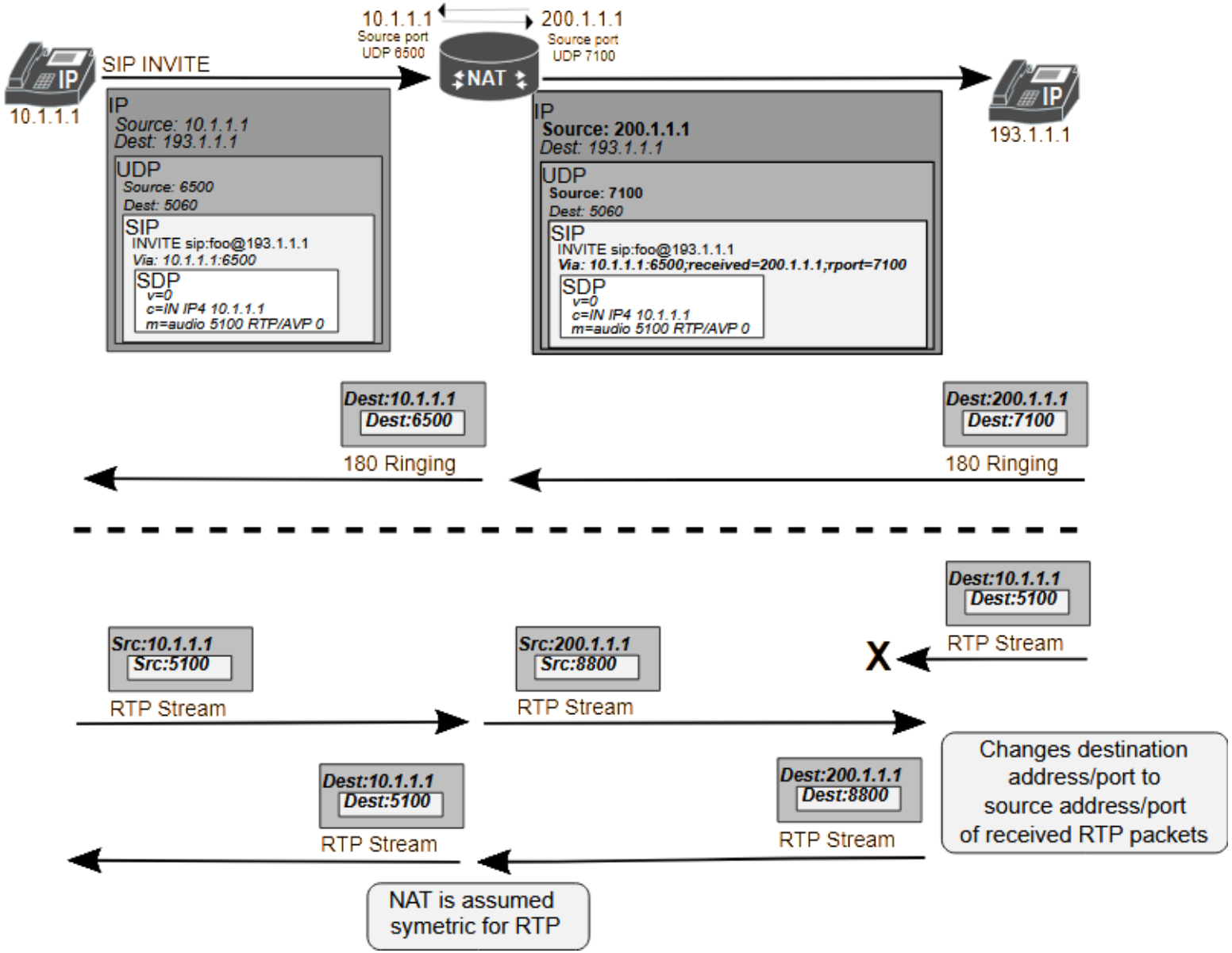
- Solution: Symmetric Response Routing (RFC 3581, August 2003):
  - SIP payload is also "translated" by NAT gateway, adding a **received** and **rport** fields with public address/port.

SIP and NAPT: solution for the SIP signalling

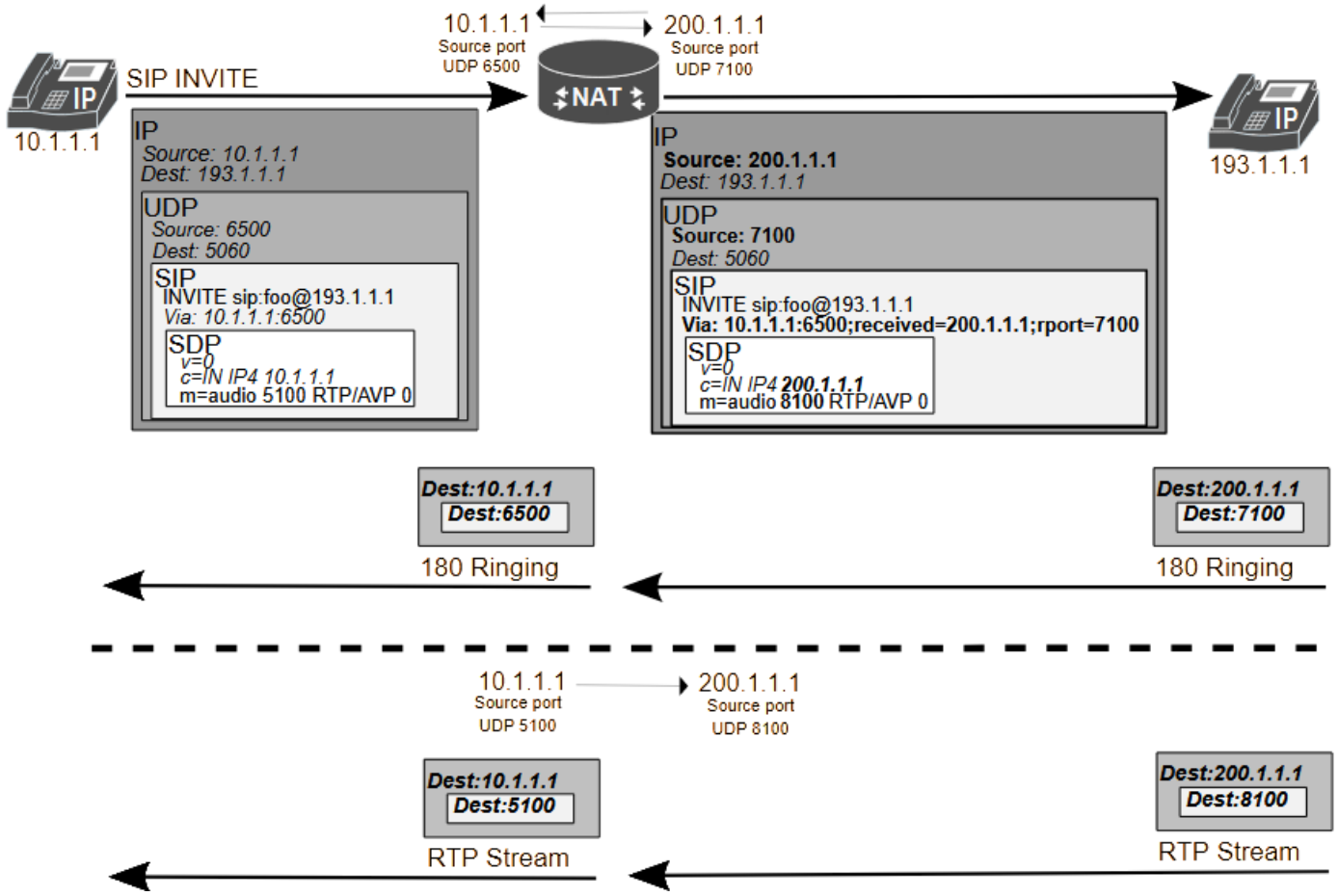


- RFC 3581 solves the SIP signalling problem, but it does not solve the RTP/RTCP audio/video exchange.
  Possible solutions (details in the next 2 slides):
  - Symmetric (RTP/RTCP) NAT, defined in RFC 4961 (July 2007)
  - NAT SIP Application Layer Gateway (ALG)

Symmetric (RTP/RTCP) NAT



NAT SIP Application Layer Gateway (ALG)



- Requires NAT to translate RTP/RTCP packets.
- It imposes an heavy computational burden on NAT gateway.