Master Course
Computer Networks
IN2097

Prof. Dr.-Ing. Georg Carle
Christian Grothoff, Ph.D.

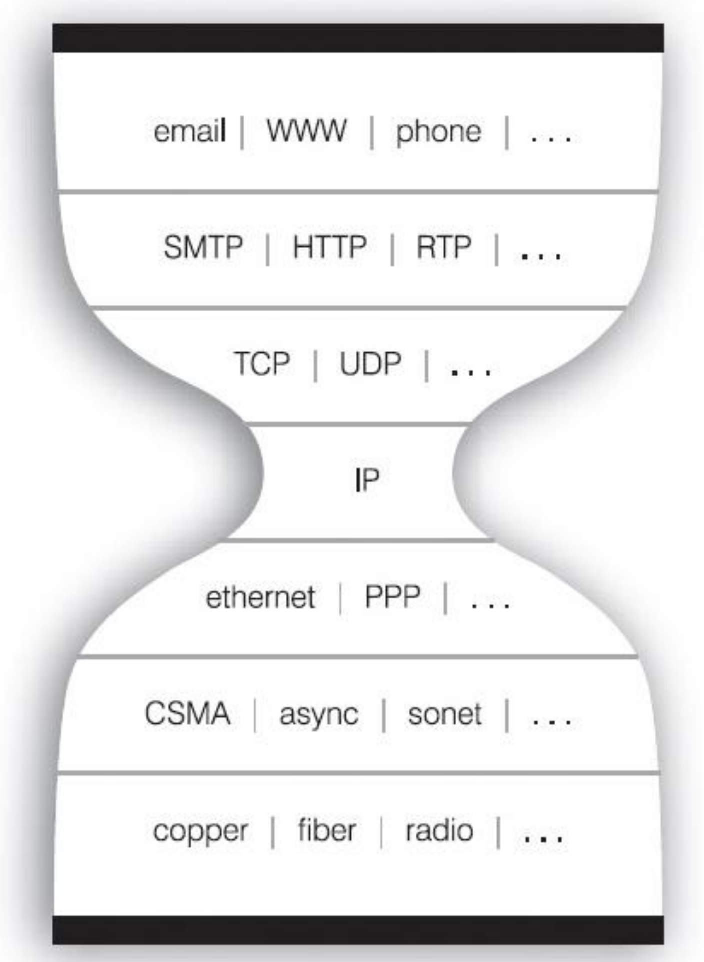
Chair for Network Architectures and Services
Institut für Informatik
Technische Universität München
http://www.net.in.tum.de

Overview

- Introduction to Network Address Translation
- Behavior of NAT
- The NAT Traversal problem
- Solutions to the problem
- Large Scale NATs

Problem

- More and more devices connect to the Internet
  - PCs
  - Cell phones
  - Internet radios
  - TVs
  - Home appliances
  - Future: sensors, cars...



- IP addresses need to be globally unique
  - IPv4 provides a 32bit field
  - Many addresses not usable because of classful allocation

→ We are running out of IP addresses

Address Space

- IP addresses are assigned by the Internet Assigned Numbers Authority (IANA)

- RFC 1918 (published in 1996) directs IANA to reserve the following IPv4 address ranges for private networks
  - 10.0.0.0 – 10.255.255.255
  - 172.16.0.0 – 172.31.255.255
  - 192.168.0.0 – 192.168.255.255

- The addresses may be used and reused by everyone
  - Not routed in the public Internet
  - Therefore a mechanism for translating addresses is needed

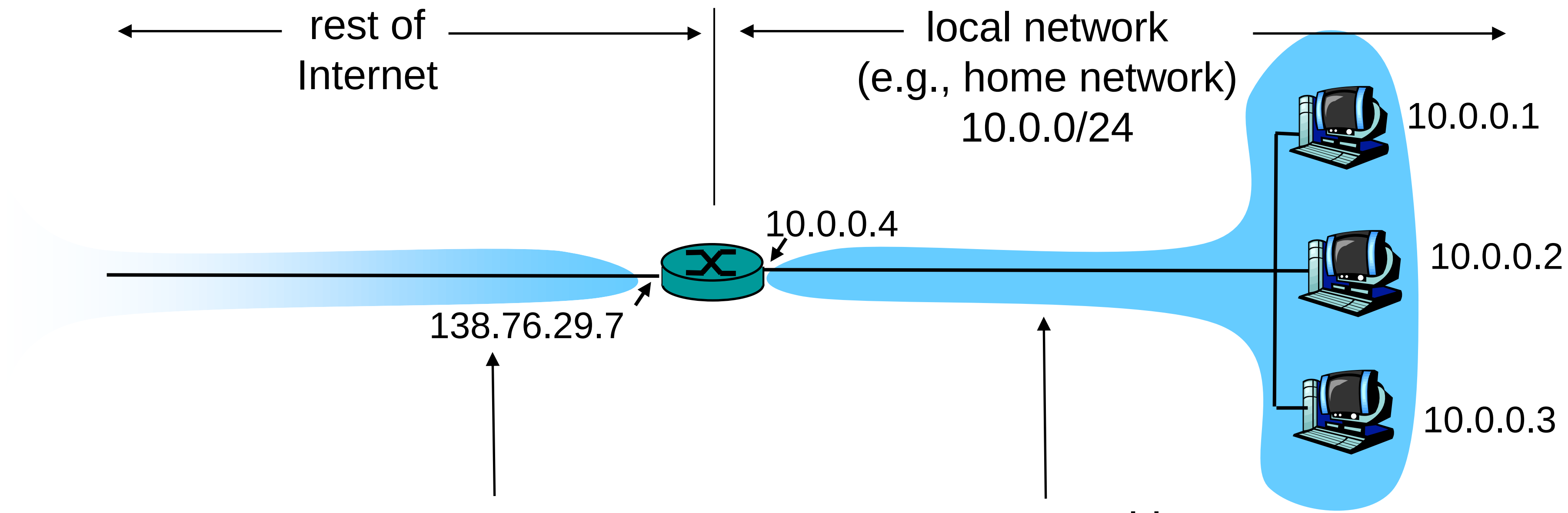
First approach – Network Address Translation

- Idea: only hosts communicating with the public Internet need a public address
  - Once a host connects to the Internet we need to allocate one
  - Communication inside the local network is not affected

- A small number of public addresses may be enough for a large number of private clients

- Only a subset of the private hosts can connect at the same time
  - not realistic anymore (always on)
  - we still need more than one public IP address

NAPT: Network Address and Port Translation



All datagrams *leaving* local network have same single source NAT IP address: 138.76.29.7, different source port numbers

Datagrams with source or destination in this network have 10.0.0/24 address for source, destination as usual

NAT: Network Address Translation

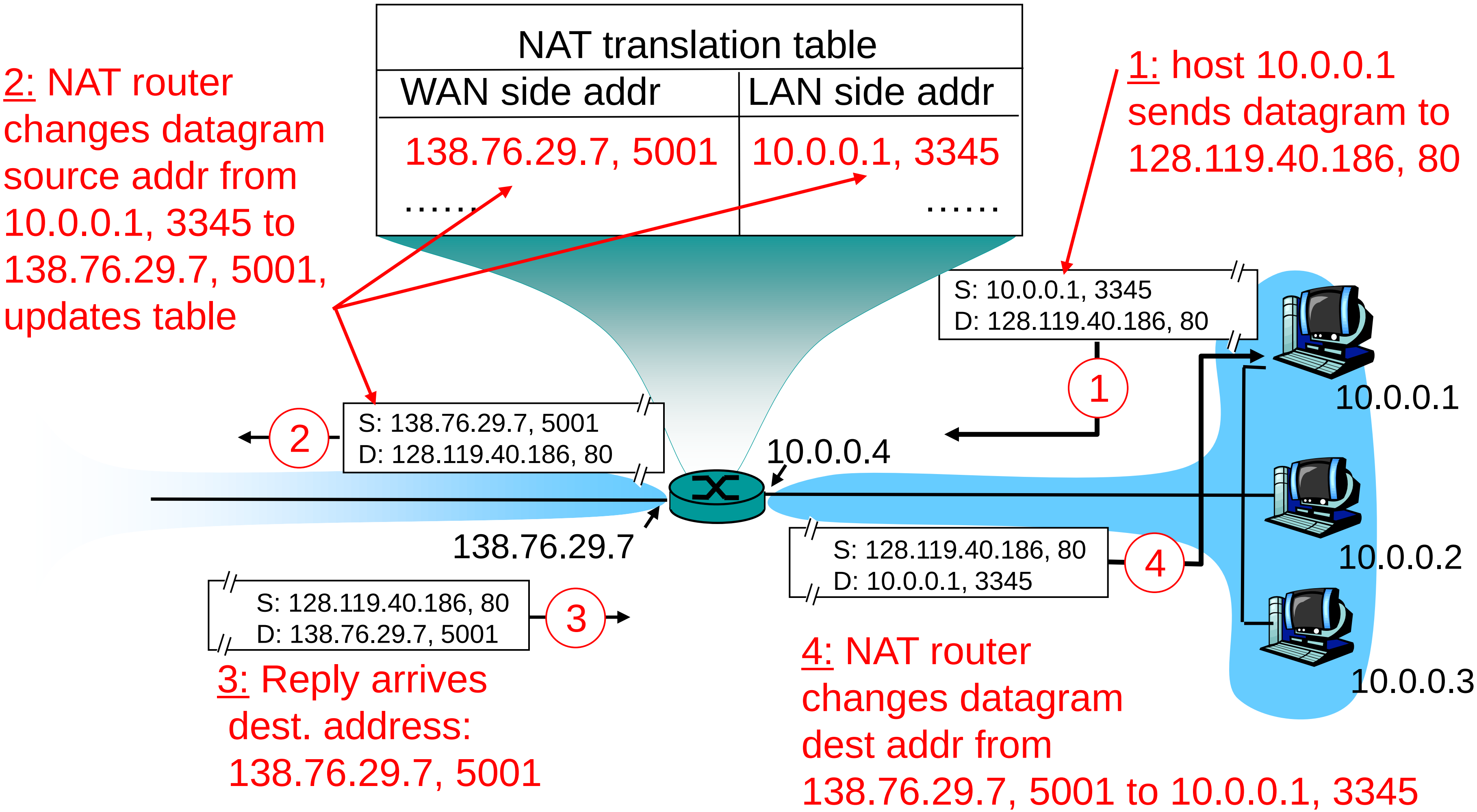
Implementation: NAT router must:

- *On outgoing datagrams: replace* (source IP address, port #) of every outgoing datagram to (NAT IP address, new port #)
  . . . remote clients/servers will respond using (NAT IP address, new port #) as destination addr.

- *remember (in NAT translation table)* every (source IP address, port #) to (NAT IP address, new port #) translation pair
  -> we have to maintain a state in the NAT

- *incoming datagrams: replace* (NAT IP address, new port #) in dest fields of every incoming datagram with corresponding (source IP address, port #) stored in NAT table

NAT: Network Address Translation



NAT: Network Address Translation

- NAPT:
  - ~65000 simultaneous connections with a single LAN-side address!
  - helps against the IP shortage
  - More advantages:
    - we can change addresses of devices in local network without notifying outside world
    - we can change ISP without changing local addresses
    - devices inside local net not explicitly addressable/visible by the outside world (a security plus)

- NAT is controversial:
  - routers should only process up to layer 3
  - violates end-to-end argument

NAT Behavior and Implementation

- Implementation not standardized
  - thought as a temporary solution

- implementation differs from model to model
  - if an application works with one NAT does not imply that is always works in a NATed environment

- NAT behavior
  - Binding (which external mapping is allocated)
    - NAT binding
    - Port binding
  - Endpoint filtering (who is allowed to access the mapping)

Binding

- When creating a new state, the NAT has to assign a new source port and IP address to the connection

- **Port binding** describes the strategy a NAT uses for the assignment of a new external source port
  - Port Preservation (if possible)
  - Some algorithm (e.g. +1)
  - Random

NAT binding

- **NAT binding** describes the behavior of the NAT regarding the reuse of an existing binding
  - two consecutive connections from the same transport address (combination of IP address and port)
  - 2 different bindings?
  - If the binding is the same → Port prediction possible

- Endpoint Independent
  - the external port is only dependent on the source transport address
  - both connections have the same IP address and port

- Endpoint Dependent
  - a new port is assigned for every connection
  - strategy could be random, but also something more predictable
  - Port prediction is hard

Endpoint filtering

- Filtering describes
  - how existing mappings can be used by external hosts
  - How a NAT handles incoming connections

- Independent-Filtering:
  - All inbound connections are allowed
  - Independent on source address
  - As long as a packet matches a state it is forwarded
  - No security

- Address Restricted Filtering:
  - packets coming from the same host (matching IP-Address) the initial packet was sent to are forwarded

- Address and Port Restricted Filtering:
  - IP address and port must match

NAT Types

- With Binding and Filtering 4 NAT types can be defined (RFC 3489)

- Full Cone NAT
  - Endpoint independent
  - Independent filtering

- Address Restricted NAT
  - Endpoint independent binding
  - Address restricted filtering

- Port Address Restricted NAT
  - Endpoint independent binding
  - Port address restricted filtering

- Symmetric NAT
  - Endpoint dependent binding
  - Port address restricted filtering

NAT Types

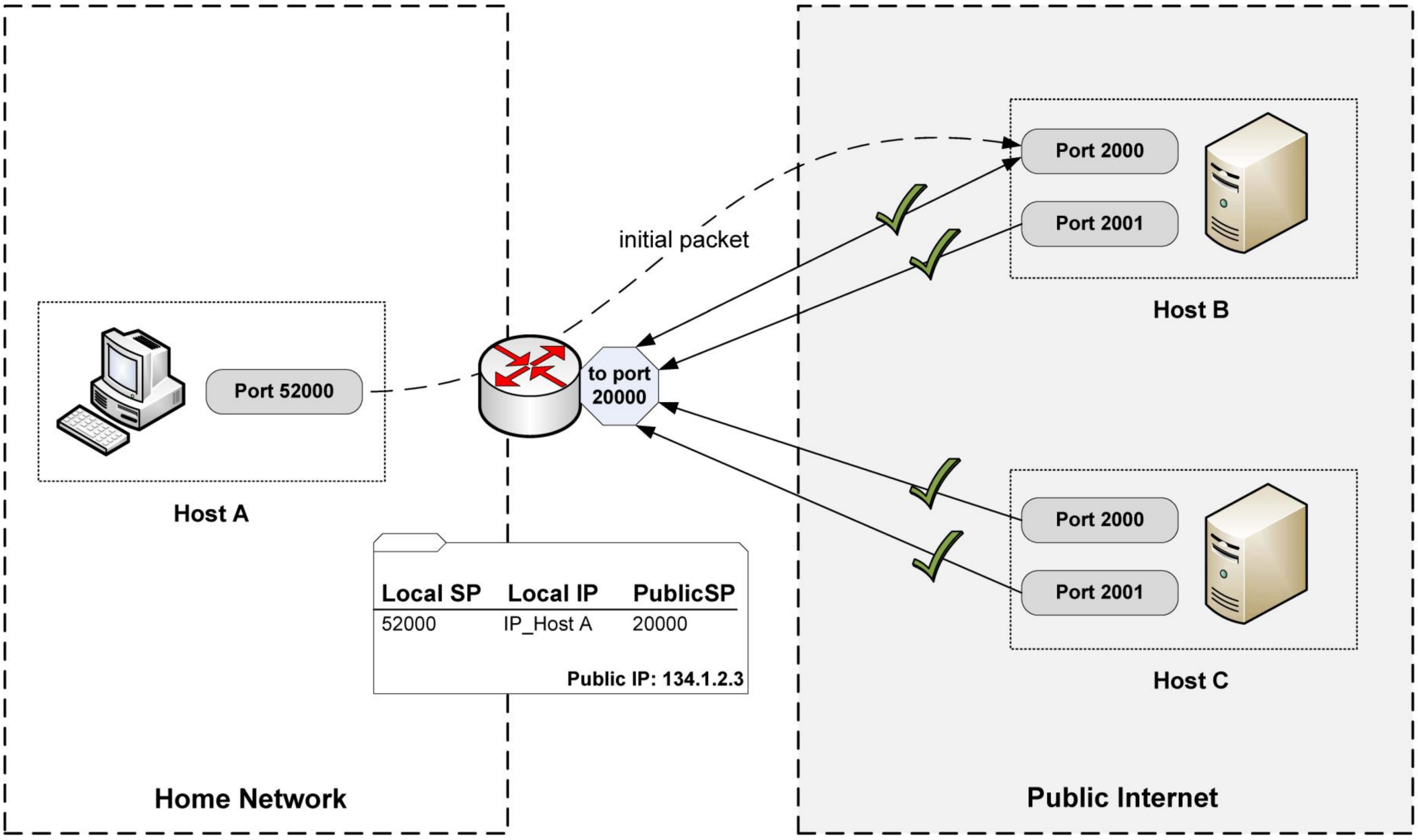
- With Binding and Filtering 4 NAT types can be defined (RFC 3489)

- **Full Cone NAT**
  - **Endpoint independent**
  - **Independent filtering**

- Address Restricted NAT
  - Endpoint independent binding
  - Address restricted filtering

- Port Address Restricted NAT
  - Endpoint independent binding
  - Port address restricted filtering

- Symmetric NAT
  - Endpoint dependent binding
  - Port address restricted filtering


Port 2000
Port 2001
Host B
initial packet
Port 52000
to port 20000
Host A
Local SP
Local IP
PublicSP
52000
IP_Host A
20000
Public IP: 134.1.2.3
Port 2000
Port 2001
Host C
Home Network
Public Internet

NAT Types

- With Binding and Filtering 4 NAT types can be defined (RFC 3489)

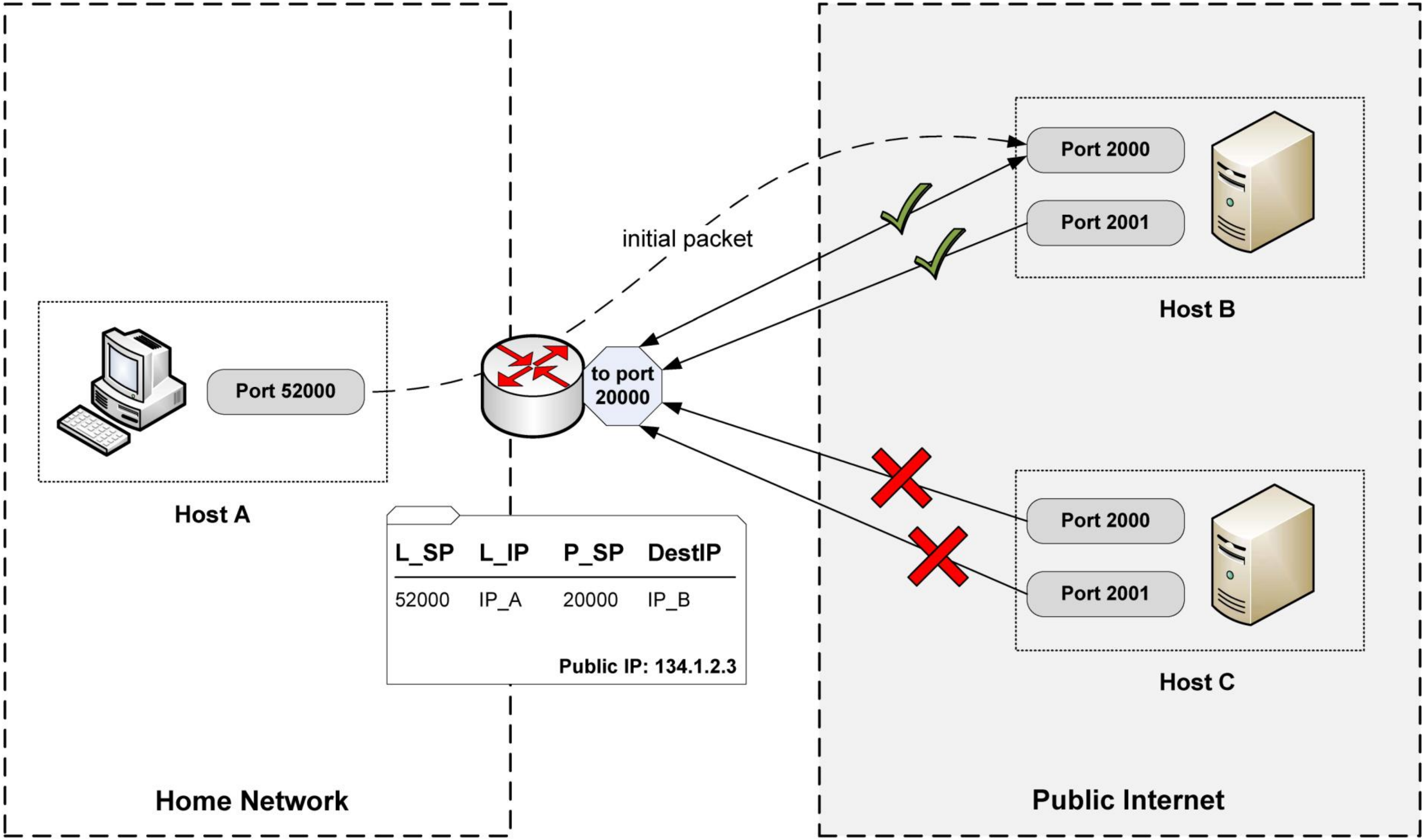
- Full Cone NAT
  - Endpoint independent
  - Independent filtering

- **Address Restricted NAT**
  - **Endpoint independent binding**
  - **Address restricted filtering**

- Port Address Restricted NAT
  - Endpoint independent binding
  - Port address restricted filtering

- Symmetric NAT
  - Endpoint dependent binding
  - Port address restricted filtering

Address Restricted Cone NAT



NAT Types

- With Binding and Filtering 4 NAT types can be defined (RFC 3489)

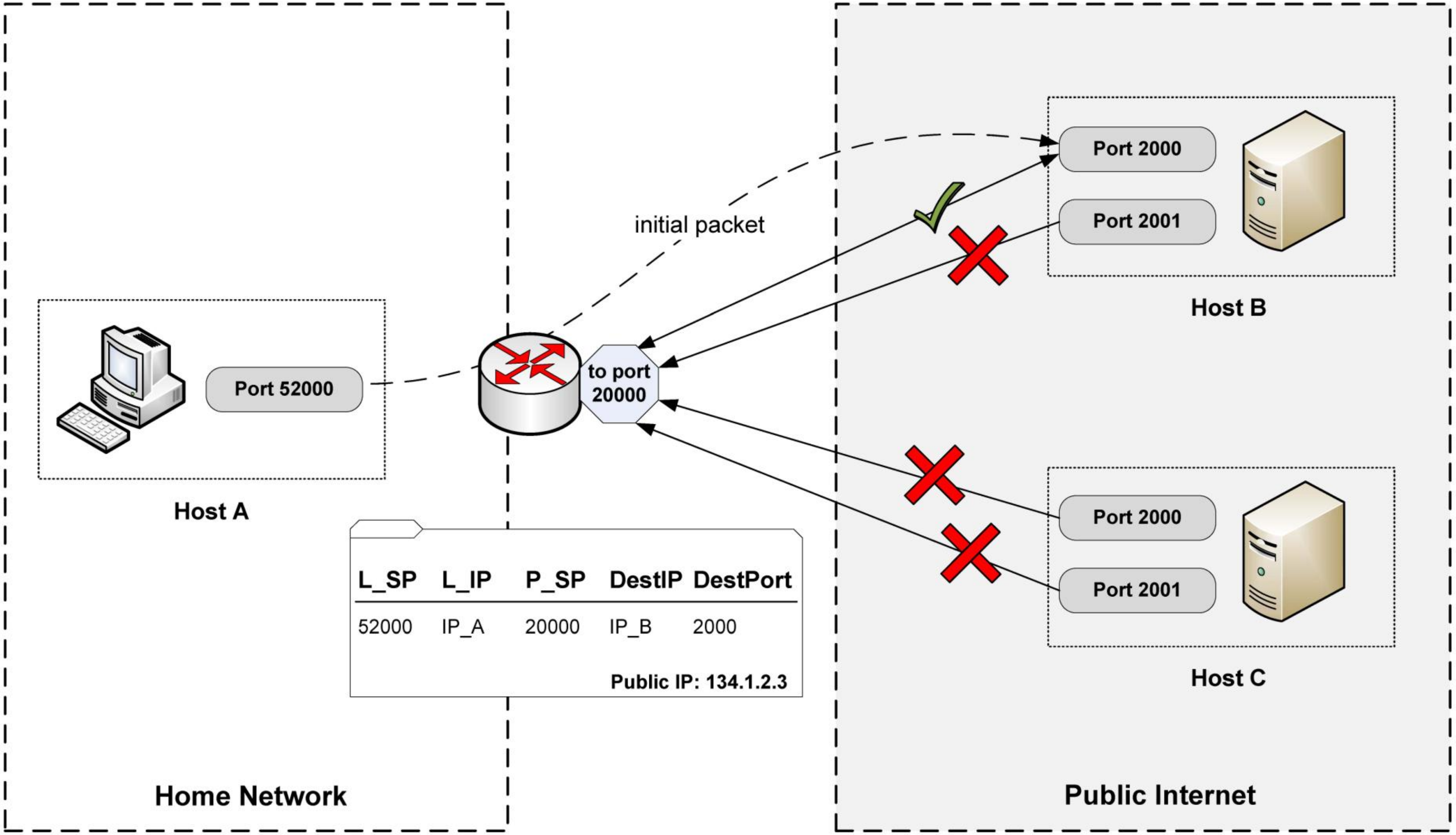
- Full Cone NAT
  - Endpoint independent
  - Independent filtering

- Address Restricted NAT
  - Endpoint independent binding
  - Address restricted filtering

- **Port Address Restricted NAT**
  - **Endpoint independent binding**
  - **Port address restricted filtering**

- Symmetric NAT
  - Endpoint dependent binding
  - Port address restricted filtering

Port Address Restricted Cone NAT



NAT Types

- With Binding and Filtering 4 NAT types can be defined (RFC 3489)

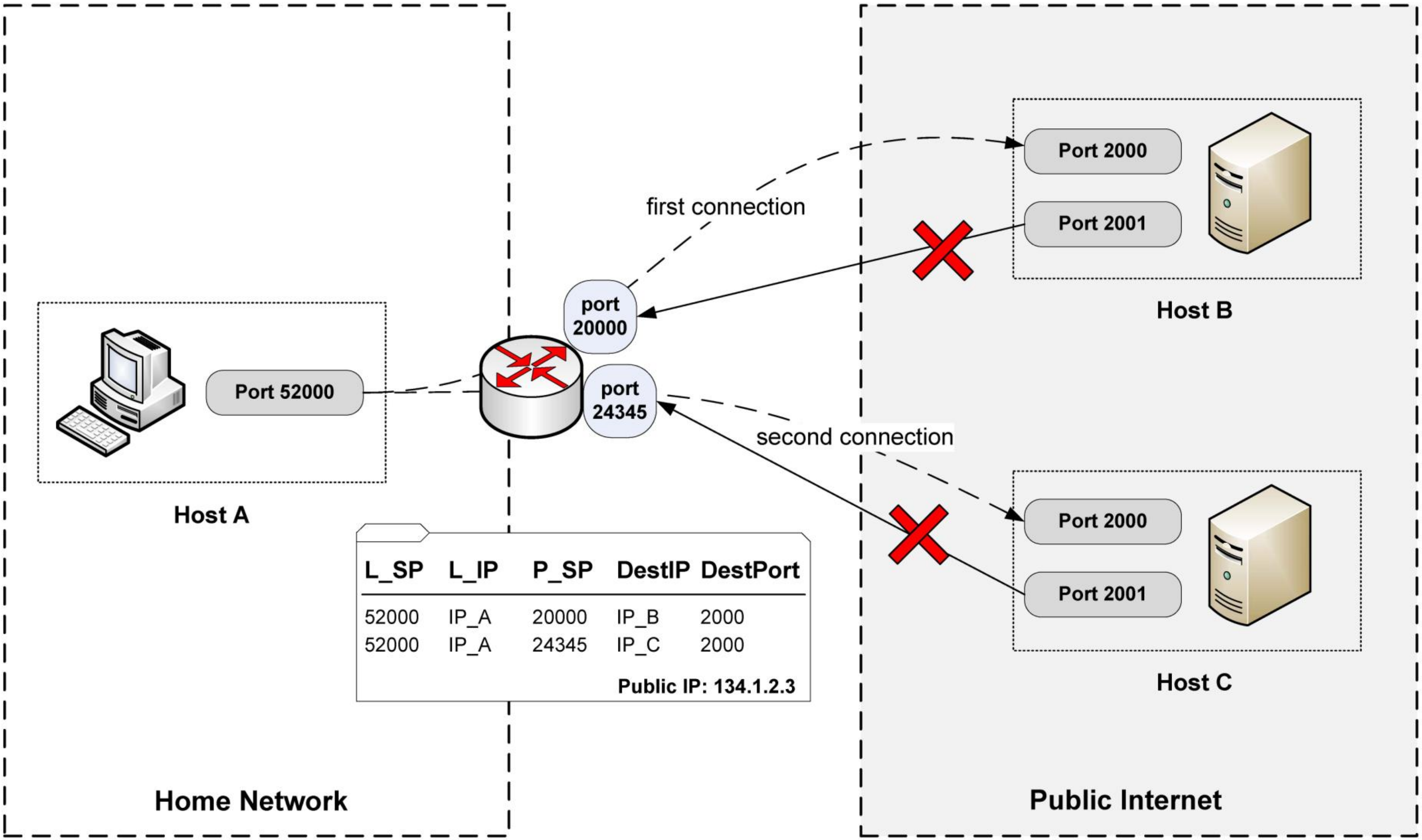
- Full Cone NAT
  - Endpoint independent
  - Independent filtering

- Address Restricted NAT
  - Endpoint independent binding
  - Address restricted filtering

- Port Address Restricted NAT
  - Endpoint independent binding
  - Port address restricted filtering

- **Symmetric NAT**
  - **Endpoint dependent binding**
  - **Port address restricted filtering**

Symmetric NAT



And where is the problem?

- NAT was designed for the client-server paradigm

- Nowadays the internet consists of applications such as
  - P2P networks
  - Voice over IP
  - Multimedia Streams

- Protocols are getting more and more complex
  - Multiple layer 4 connections (data and control session)
  - Realm specific addresses in layer 7

- Connectivity requirements have changed
  - P2P is becoming more and more important
    - Especially for future home and services
  - Direct connections between hosts is necessary

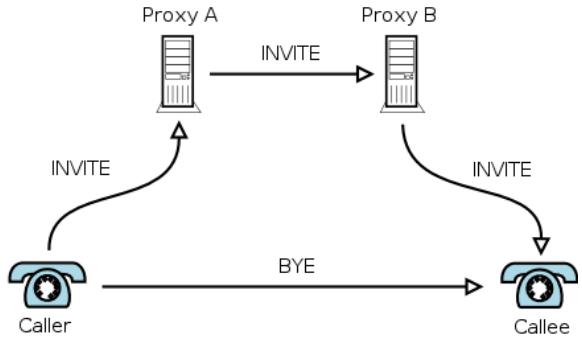
- NATs break the end-to-end connectivity model of the internet
  - Inbound packets can only be forwarded if an appropriate mapping exists
  - Mappings are only created on outbound packets

NAT-Traversal Problem

- Divided into four categories: (derived from IETF-RFC 3027)
  - **Realm-Specific IP-Addresses in the Payload**
    - *Session Initiation Protocol (SIP)*
  - **Peer-to-Peer Applications**
    - *Any service behind a NAT*
  - **Bundled Session Applications (Inband Signaling)**
    - *FTP*
    - *Real time streaming protocol (RTSP)*
    - *SIP together with SDP (Session Description Protocol)*
  - **Unsupported Protocols**
    - *SCTP (Stream Control Transmission Protocol)*
    - *IPSec*

Example: Session Initiation Protocol (SIP)

- Realm Specific IP addresses in the payload (SIP)
- Bundled Session Application (RTP)



Request/Respone Line

INVITE sip:Callee@200.3.4.5 SIP/2.0

Message-Header

Via: SIP/2.0/UDP **192.168.1.5:5060**
From: < sip:Caller@**192.168.1.5** >
To: <sip:Callee@200.3.4.5>
CSeq: 1 INVITE
Contact: <sip:Caller@192.168.1.5:5060>
Content-Type: application/sdp

Message-Body (optional)

v=0
o=Alice 214365879 214365879 IN IP4 **192.168.1. 5**
c=IN IP4 **192.168.1.5**
t= 0 0
m=audio 5200 RTP/AVP 0 9 7 3
a=rtpmap:8 PCMU/8000
a=rtpmap:3 GSM/8000

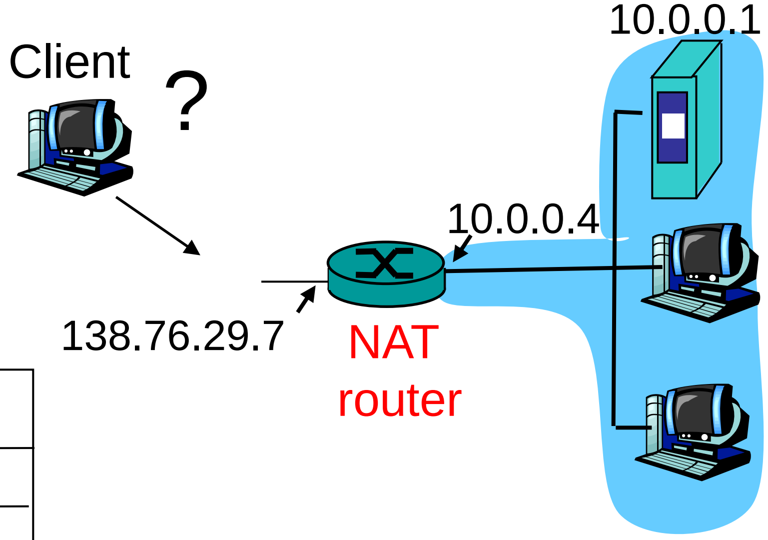
RTP-Session Specification (for 2nd channel)

Media description for 2nd channel

SDP

Example: P2P applications

- Client wants to connect to server with address 10.0.0.1
  - server address 10.0.0.1 local to LAN
    (client can't use it as destination addr)
  - only one externally visible NATted address: 138.76.29.7
  - NAT does not have any idea where to forward packets to



NAT translation table

WAN side addr	LAN side addr
138.76.29.7, 80	10.0.0.1, 80
……	……

Existing Solutions to the NAT-Traversal Problem

- Individual solutions
  - Explicit support by the NAT
    - Static port forwarding, ALG, UPnP, NAT-PMP
  - NAT-behavior based approaches
    - dependent on knowledge about the NAT
    - Hole Punching using STUN (IETF - RFC 3489)
  - External Data-Relay
    - TURN (IETF - Draft)

- Frameworks integrating several techniques
  - framework selects a working technique
  - ICE as the most promising for VoIP (IETF - Draft)

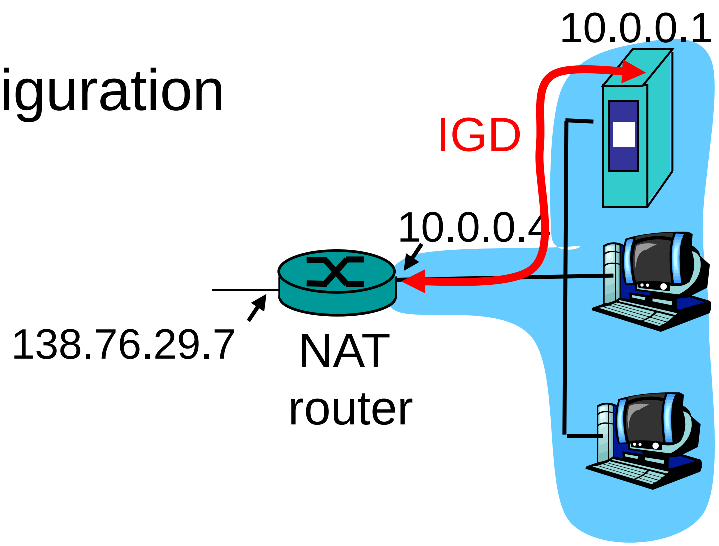
Explicit support by the NAT (1)

- Application Layer Gateway (ALG)
  - implemented on the NAT device and operates on layer 7
  - supports Layer 7 protocols that carry realm specific addresses in their payload
    - SIP, FTP
- Advantages
  - transparent for the application
  - no configuration necessary
- Drawbacks
  - protocol dependent (e.g. ALG for SIP, ALG for FTP...)
  - may or may not be available on the NAT device

Explicit support by the NAT (2)

- Universal Plug and Play (UPnP)
  - Automatic discovery of services (via Multicast)
  - Internet Gateway Device (IGD) for NAT-Traversal

- IGD allows NATed host to
  - automate static NAT port map configuration
  - learn public IP address (138.76.29.7)
  - add/remove port mappings (with lease times)



- Drawbacks
  - no security, evil applications can establish port forwarding entries
  - doesn‘t work with cascaded NATs

Behavior based (1): STUN

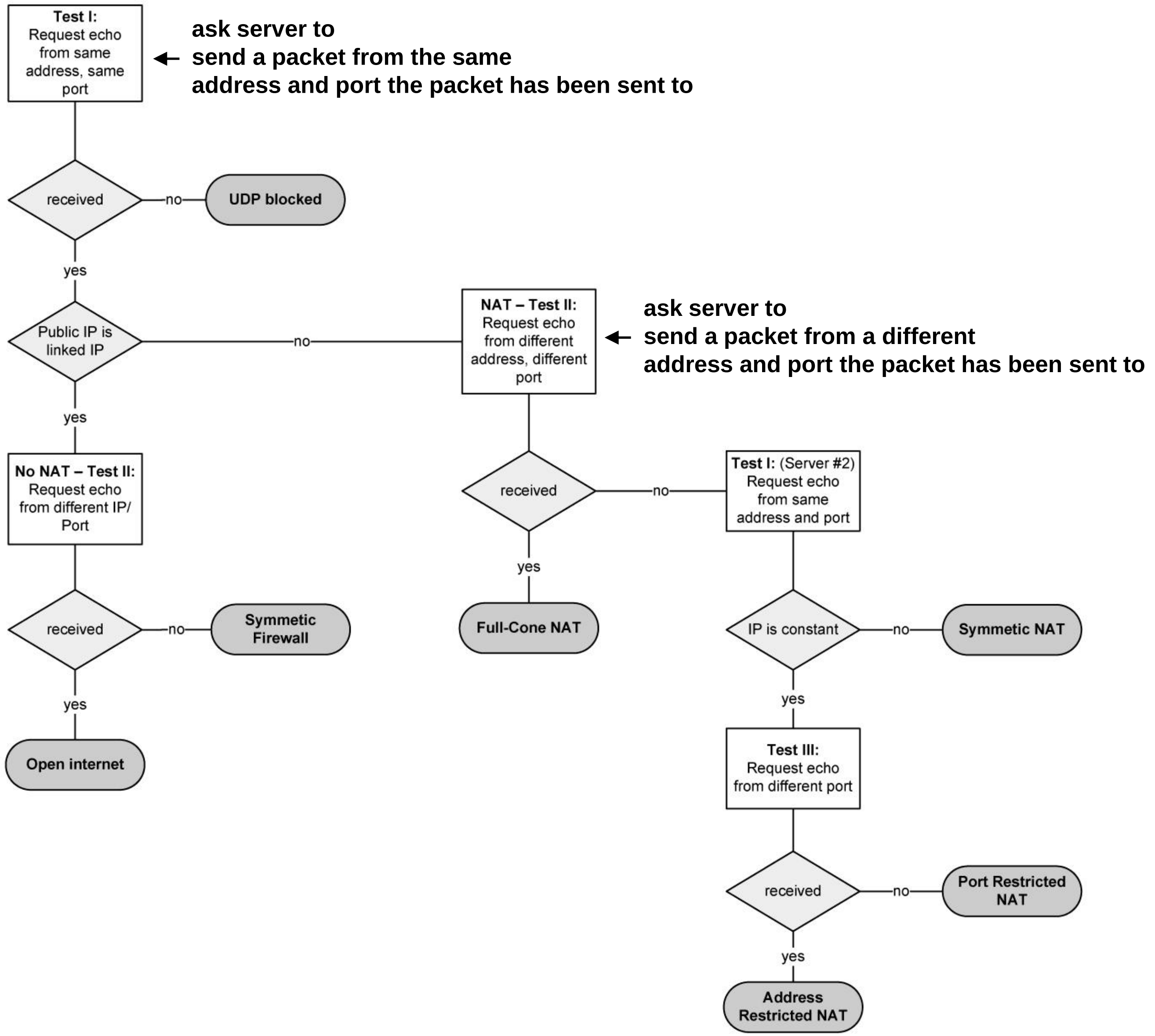
- Simple traversal of UDP through NAT (old) (RFC 3489)
  - Session Traversal Utilities for NAT (new) (RFC 5389)

- Lightweight client-server protocol
  - queries and responses via UDP (optional TCP or TCP/TLS)

- Helps to determine the external transport address (IP address and port) of a client.
  - e.g. query from 192.168.1.1:5060 results in 131.1.2.3:20000

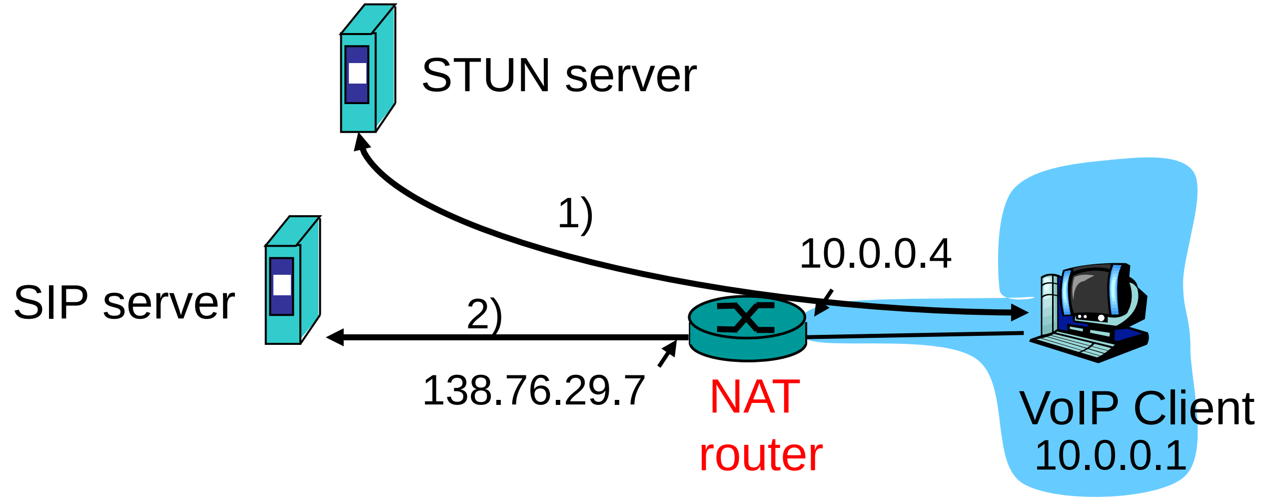
- Algorithm to discover NAT type
  - server needs 2 public IP addresses

STUN Algorithm



Example: STUN and SIP

- VoIP client queries STUN server
  - learns its public transport address
  - can be used in SIP packets



Request/Respone Line	INVITE sip:Callee@200.3.4.5 SIP/2.0
Message-Header	Via: SIP/2.0/UDP **138.76.29.7:5060** From: < sip:Caller@**138.76.29.7** > To: <sip:Callee@200.3.4.5> CSeq: 1 INVITE Contact: <sip:Caller@**138.76.29.7:5060**> Content-Type: application/sdp

Limitations of STUN

- STUN only works if
  - the NAT assigns the external port (and IP address) only based on the source transport address
  - Endpoint independent NAT binding
    - Full Cone NAT
    - Address Restricted Cone NAT
    - Port Address restricted cone NAT
  - Not with symmetric NAT!

- Why?
  - Since we first query the STUN server (different IP and port) and then the actual server
  - The external endpoint must only be dependent on the source transport address

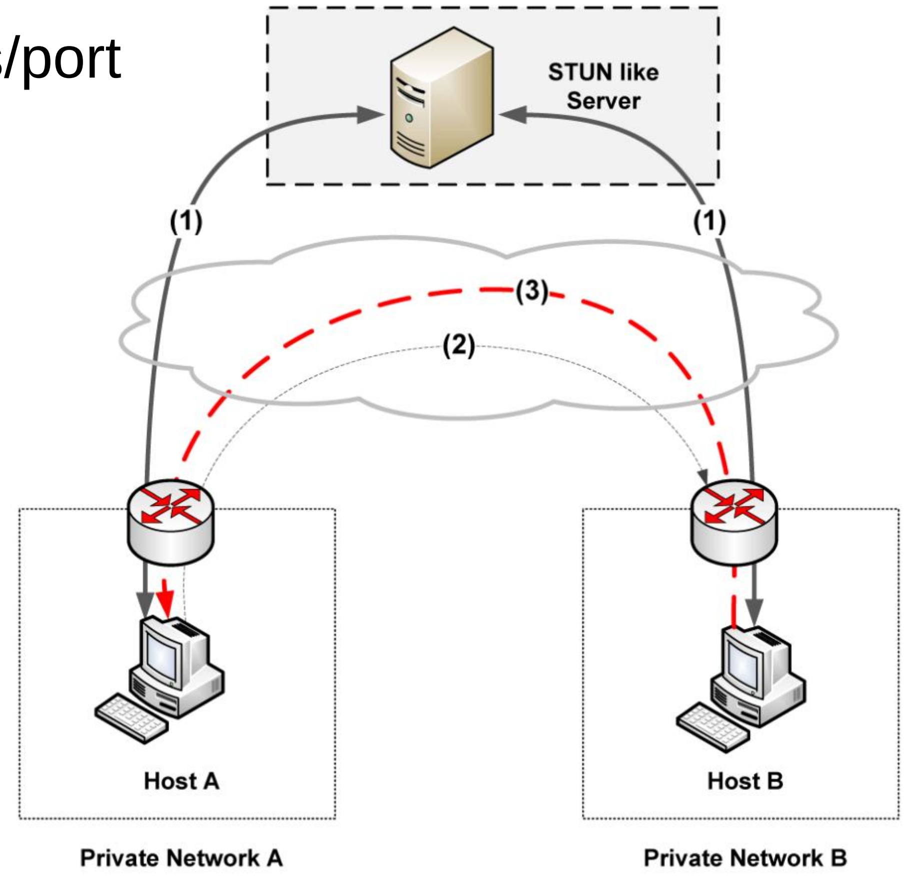
STUN and Hole Punching

- STUN not only helps if we need IP addresses in the payload
  - also for establishing a direct connection between two peers

1) determine external IP address/port and exchange it through Rendezvous Point

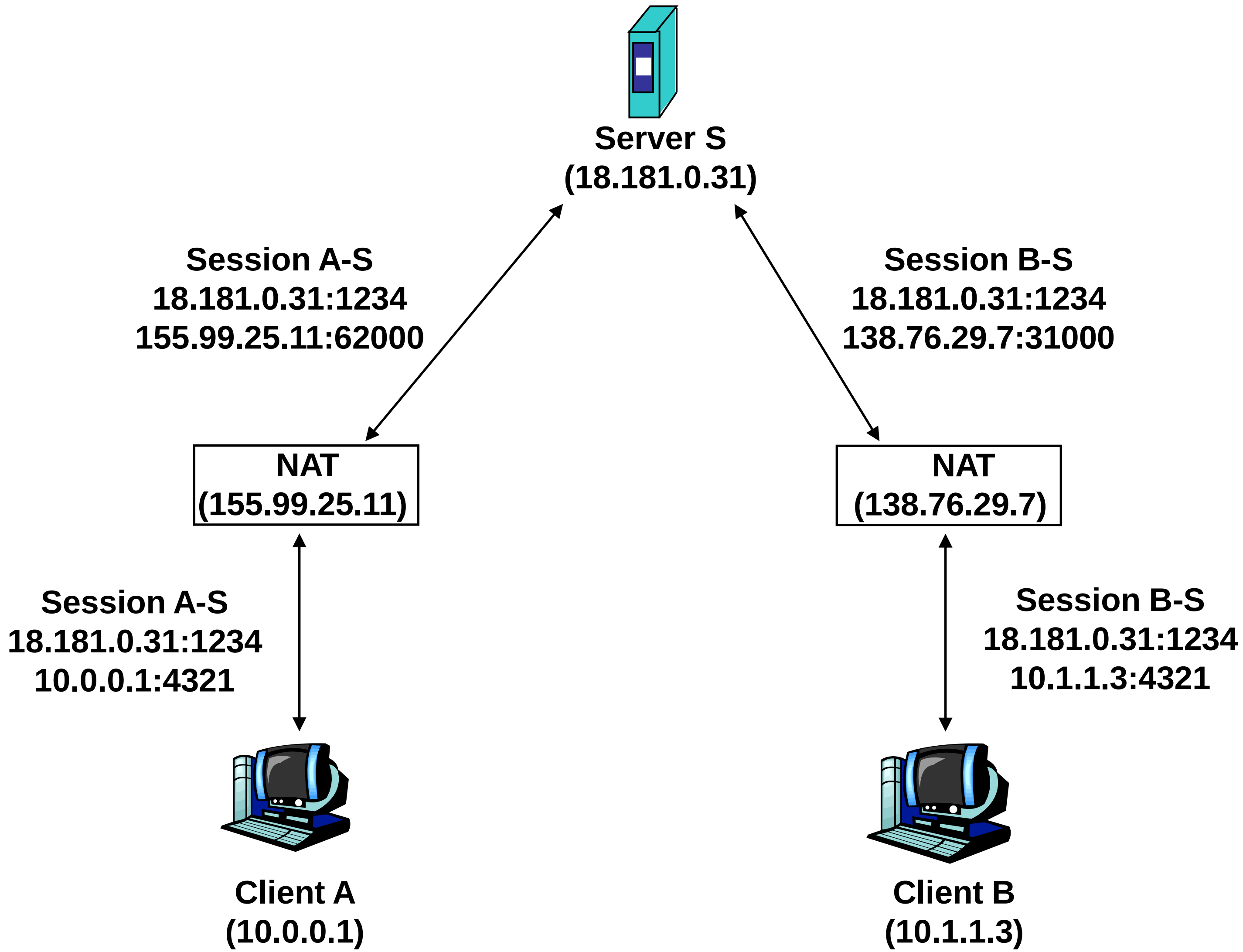
2) both hosts send packets towards the other host outgoing packet creates hole

3) establish connection. hole is created by first packet



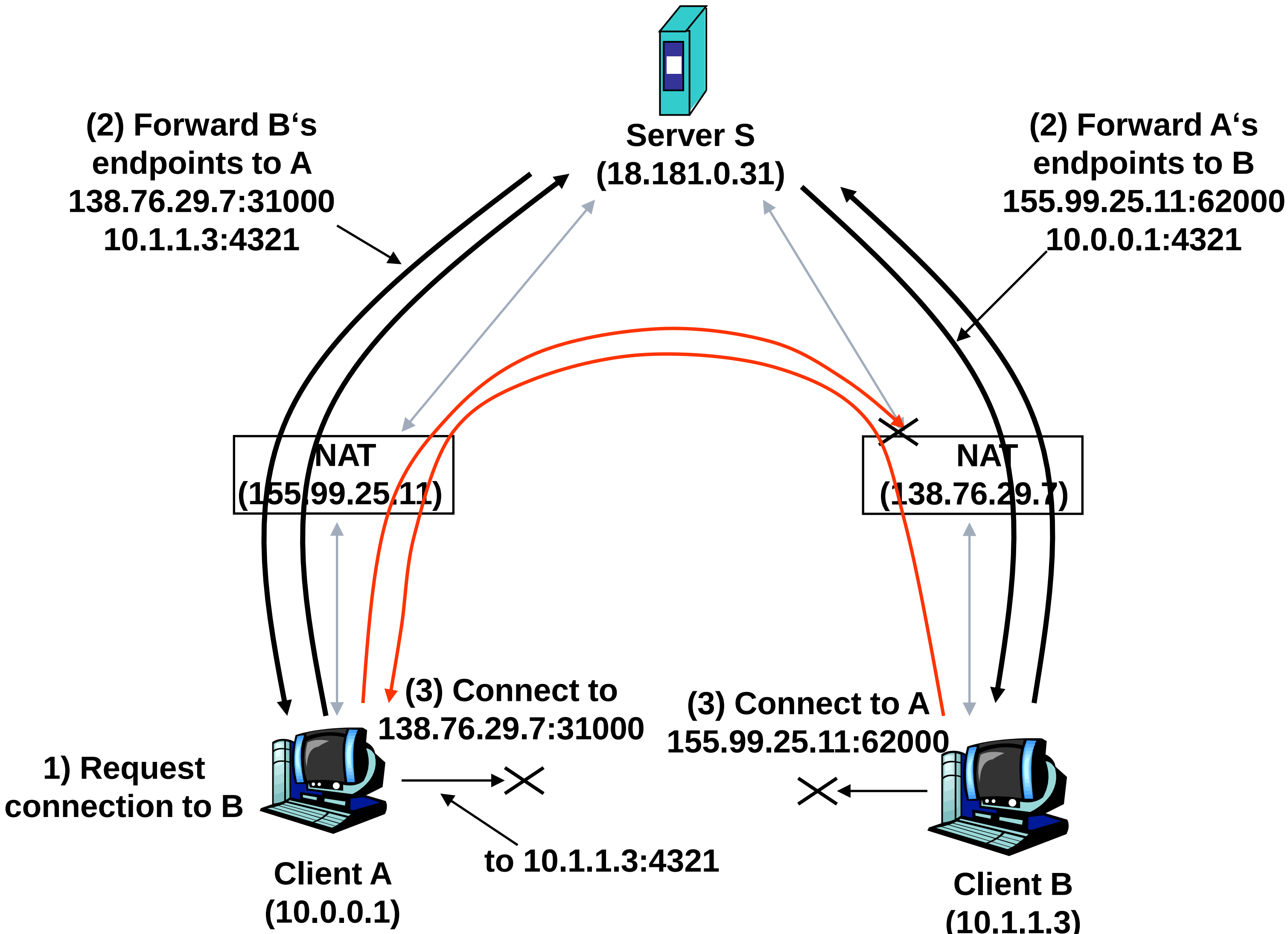
Hole Punching in detail

- Before hole punching



Hole Punching in detail

- Hole punching



DIY Hole Punching: practical example

- You need 2 hosts
  - One in the public internet (client)
  - One behind a NAT (server)

- Firstly start a UDP listener on UDP port 20000 on the “server” console behind the NAT/firewall
  - server/1# nc -u -l -p 20000

- An external computer “client" then attempts to contact it
  - client# echo "hello" | nc -p 5000 -u serverIP 20000
  - Note: 5000 is the source port of the connection

- as expected nothing is received because the NAT has no state

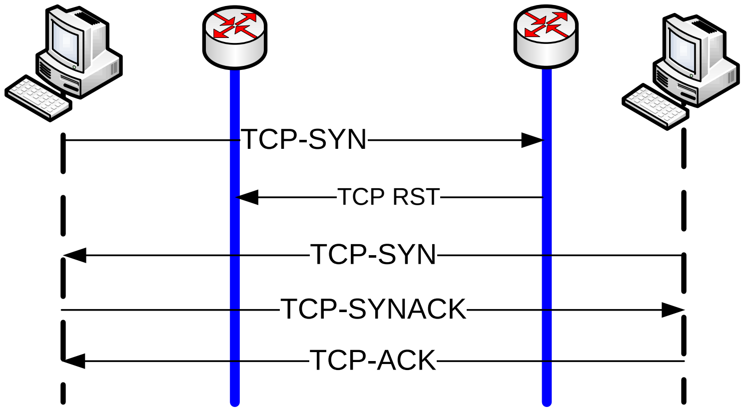
- Now on a second console, server/2, we punch a hole
  - Server/2# hping2 -c 1 -2 -s 20000 -p 5000 clientIP

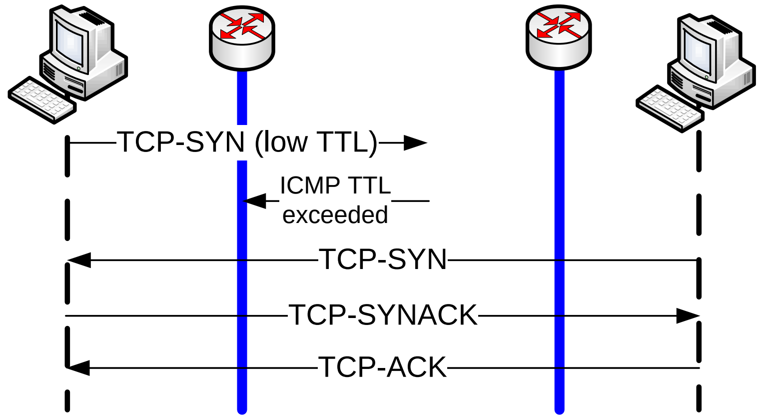
- On the second attempt we connect to the created hole
  - client# echo "hello" | nc -p 5000 -u serverIP 20000

TCP Hole Punching

- Hole Punching not straight forward due to stateful design of TCP
  - 3-way handshake
  - Sequence numbers
  - ICMP packets may trigger RST packets

- Low/high TTL(Layer 3) of Hole-Punching packet
  - As implemented in STUNT (Cornell University)

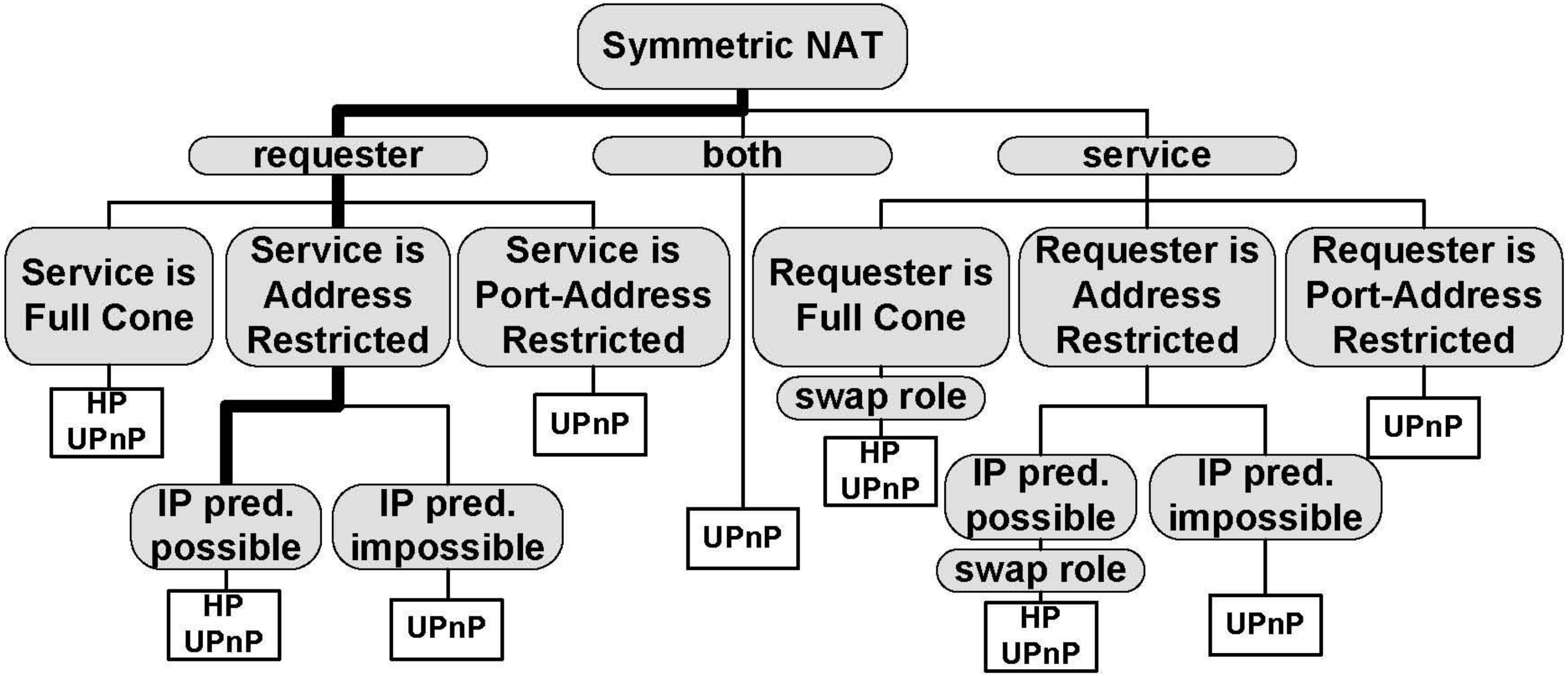




- Bottom line: NAT is not standardized

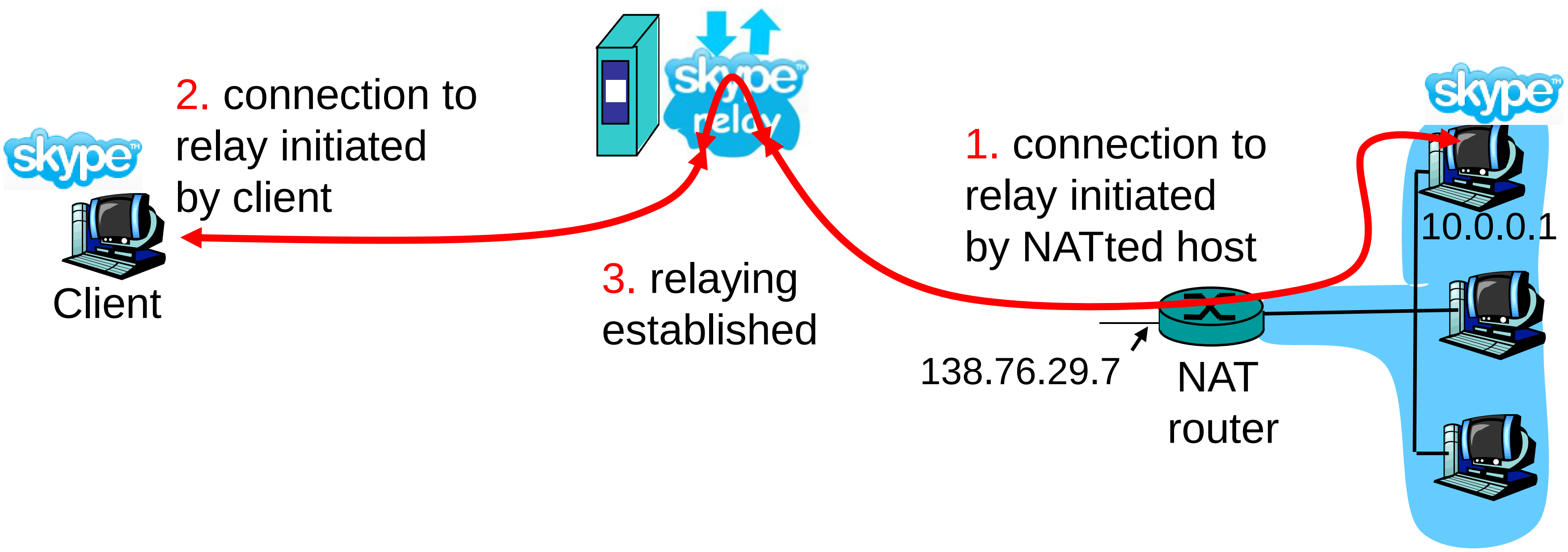
Symmetric NATs

- How can we traverse symmetric NATs
  - Endpoint dependent binding
    - hole punching in general only if port prediction is possible
  - Address and port restricted filtering



Data Relay

- relaying (used in Skype)
  - NATed client establishes connection to relay
  - External client connects to relay
  - relay bridges packets between to connections
  - Traversal using Relay NAT (TURN) as IETF draft



Frameworks

- Interactive Connectivity Establishment (ICE)
  - IETF draft
  - mainly developed for VoIP
  - signaling messages embedded in SIP/SDP

- All possible endpoints are collected and exchanged during call setup
  - local addresses
  - STUN determined
  - TURN determined

- All endpoints are „paired“ and tested (via STUN)
  - best one is determined and used for VoIP session

- Advantages
  - high sucess rate
  - integrated in application

- Drawbacks
  - overhead
  - latency dependent on number of endpoints (pairing)