IPv4 addressing, NAT

http://xkcd.com/195/

Computer Networking: A Top Down Approach
6th edition
Jim Kurose, Keith Ross
Addison-Wesley

Chapter 4: outline

4.1 Introduction

4.2 Virtual circuit and datagram networks

4.3 What's inside a router

4.4 IP: Internet Protocol

- Datagram format
- IPv4 addressing
- Network Address Translation (NAT)
- DHCP
- ICMP
- IPv6
- IPsec

4.5 Routing algorithms

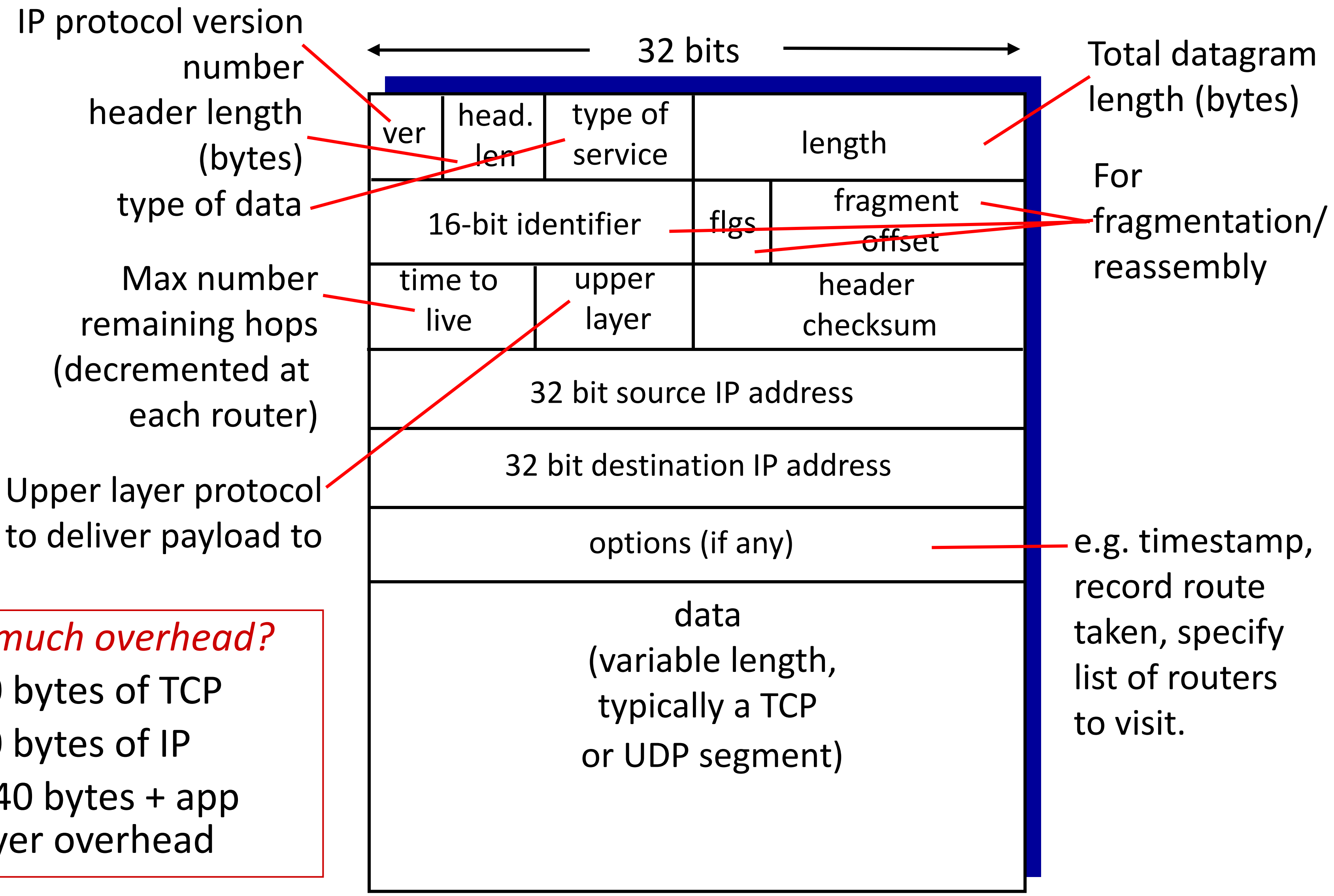
- Link state
- Distance vector
- Hierarchical routing

4.6 Routing in the Internet

- RIP
- OSPF
- BGP

4.7 Broadcast and multicast routing

IP datagram format



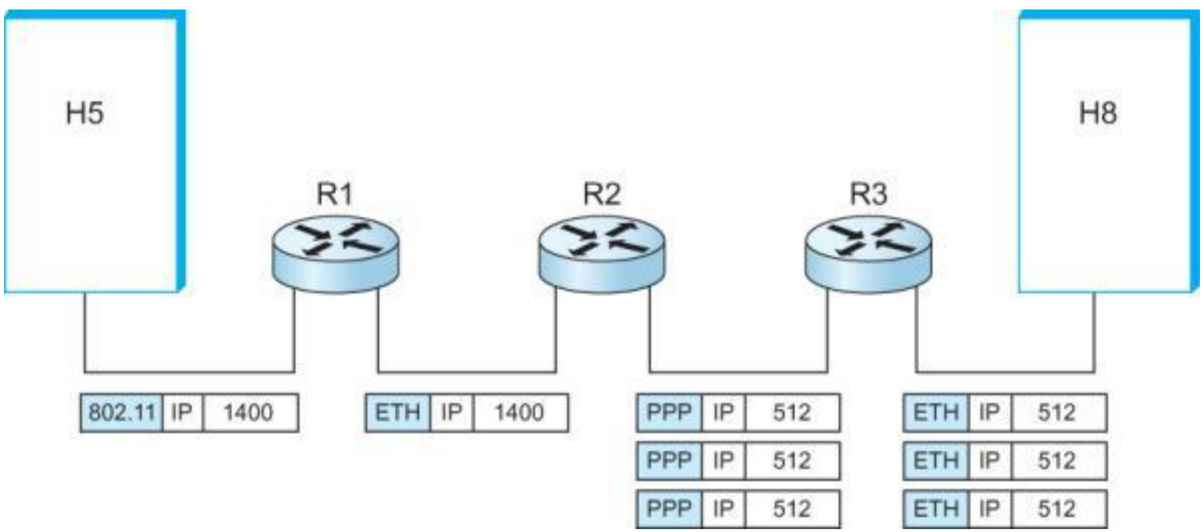
How much overhead?

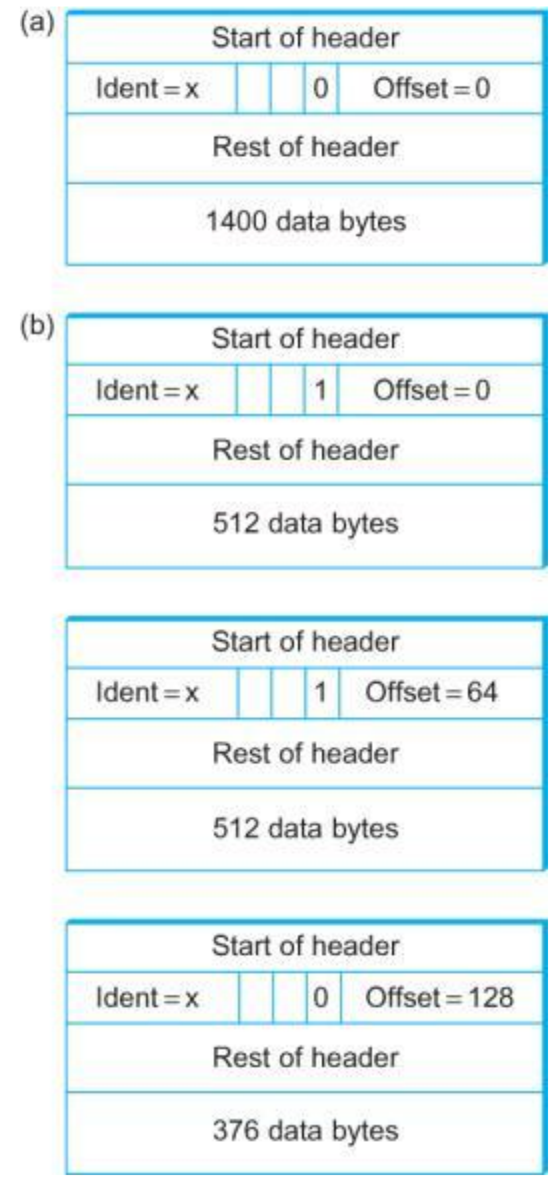
- 20 bytes of TCP
- 20 bytes of IP
- = 40 bytes + app layer overhead

Fragmentation

- Different networks support different packet sizes
  - Maximum Transmission Unit (MTU)
  - e.g. Ethernet 1500 bytes, 802.11 2272 bytes
- Occurs at router
  - If inbound datagram > MTU of outbound network
  - Split into fragments
    - All fragments have same Ident field
    - Each is self-contained datagram

Fragmentation and reassembly



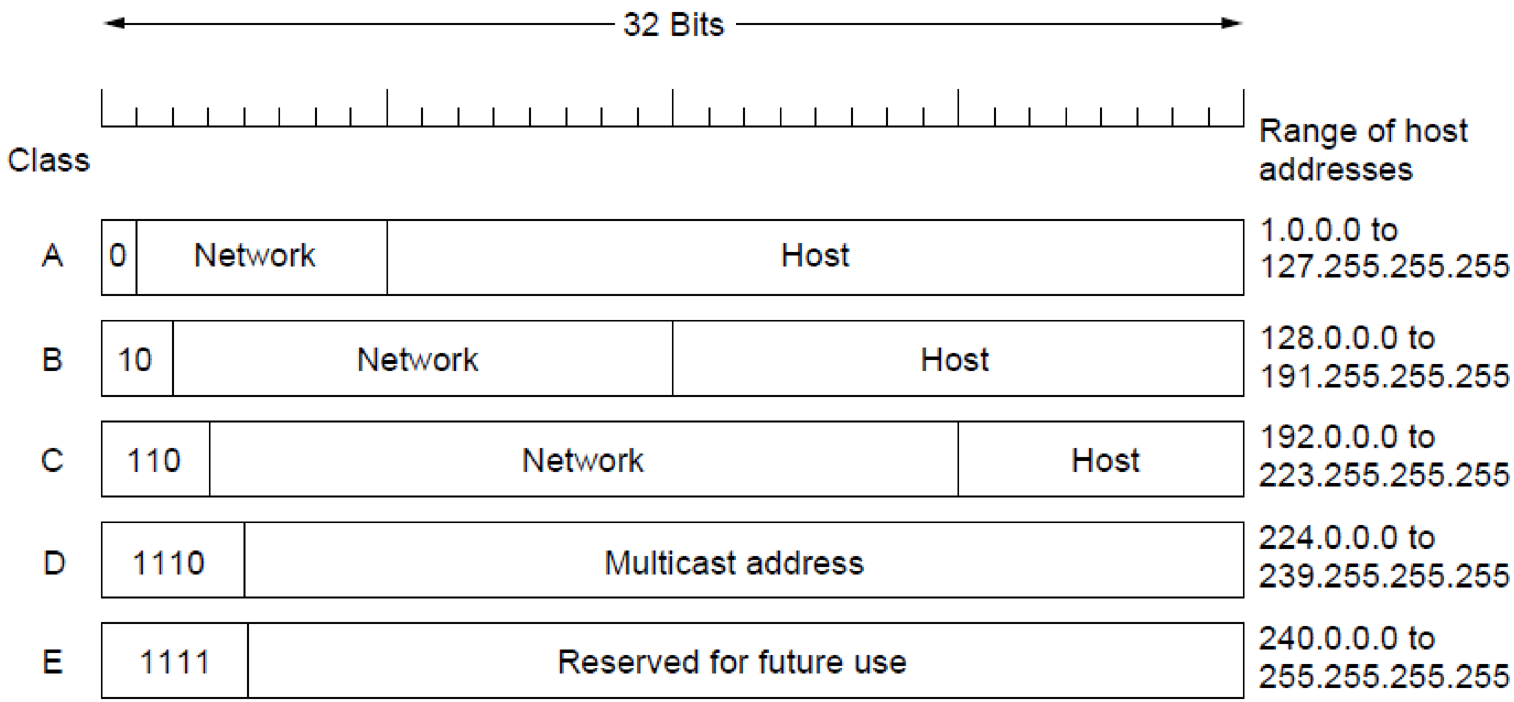


- Reassembly can be done independent of order of arrival
- Fragments may also be fragmented
- No attempt to recover if fragment missing
- Hosts can do MTU discovery
  - Probe message to determine max packet size

Global addressing

- IP service model
  - Assumes global addresses
- Why not 48-bit MAC address?
  - flat structure, no hierarchy
  - e.g. 01:23:45:67:89:ab
- IP addresses
  - IPv4 32 bits
  - network part
  - host part
  - e.g. 10.33.73.165

IPv4 address format



- Classful addressing (before 1993):
  - Class A: 128 networks with 16 million hosts
  - Class B: 16,384 networks with 65,536 hosts
  - Class C: 2 million networks, 256 hosts

Special IP addresses

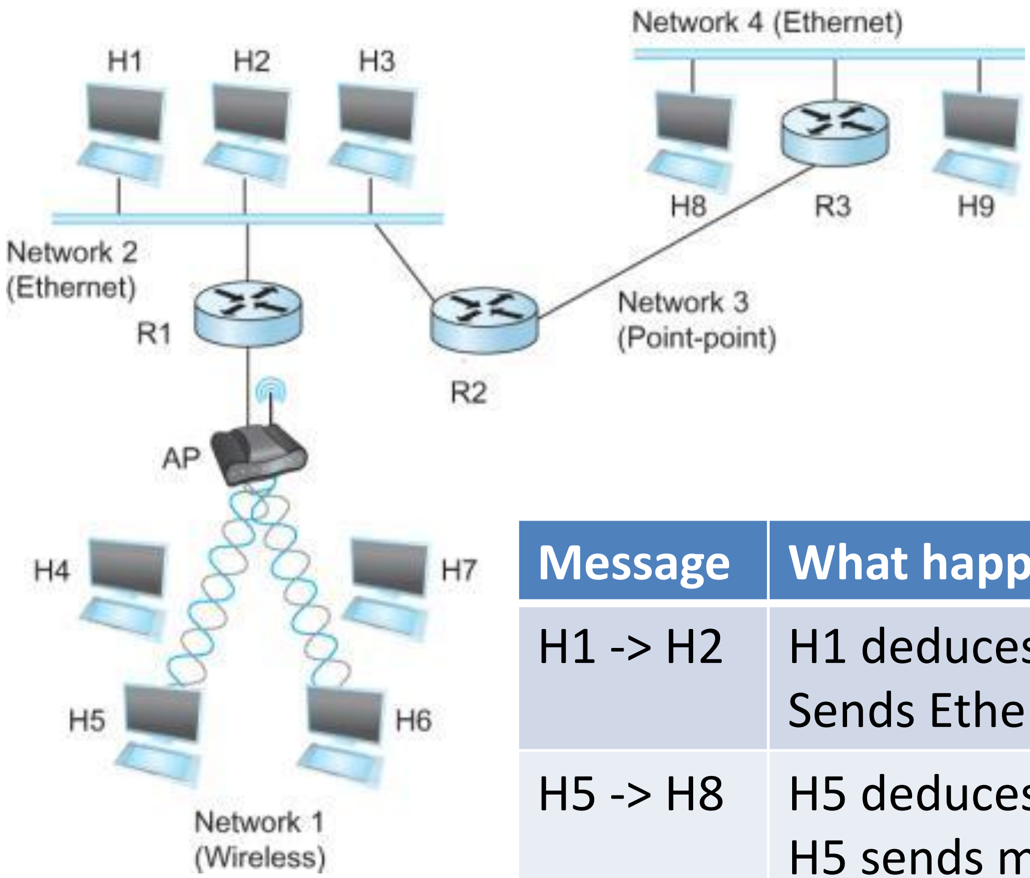
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	This host
0 0 . . . 0 0 | Host	A host on this network
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Broadcast on the local network
Network | 1 1 1 1 . . . 1 1 1 1	Broadcast on a distant network
127 | (Anything)	Loopback

- Dot notation, each byte converted to decimal
  - 1000 0000 1101 0000 0000 0010 1001 0111
  - 128.208.2.151

IP datagram forwarding

- Why (network + host) address help?
  - Routers have a forwarding table
    - Network number -> next hop
  - Without hierarchy:
    - Tables in routers would be huge
    - Machines on same network wouldn't know it
  - Default router
    - Where to send things if not in your table

Routing example



Message	What happens
H1 -> H2	H1 deduces on same network as H2 Sends Ethernet packet directly
H5 -> H8	H5 deduces H8 not on same network H5 sends message to default router R1 R1 can't delivery directly, send to its default router R2 R2 has a forwarding table showing H8 available from R3, sends to R3 R3 delivery to network 4.

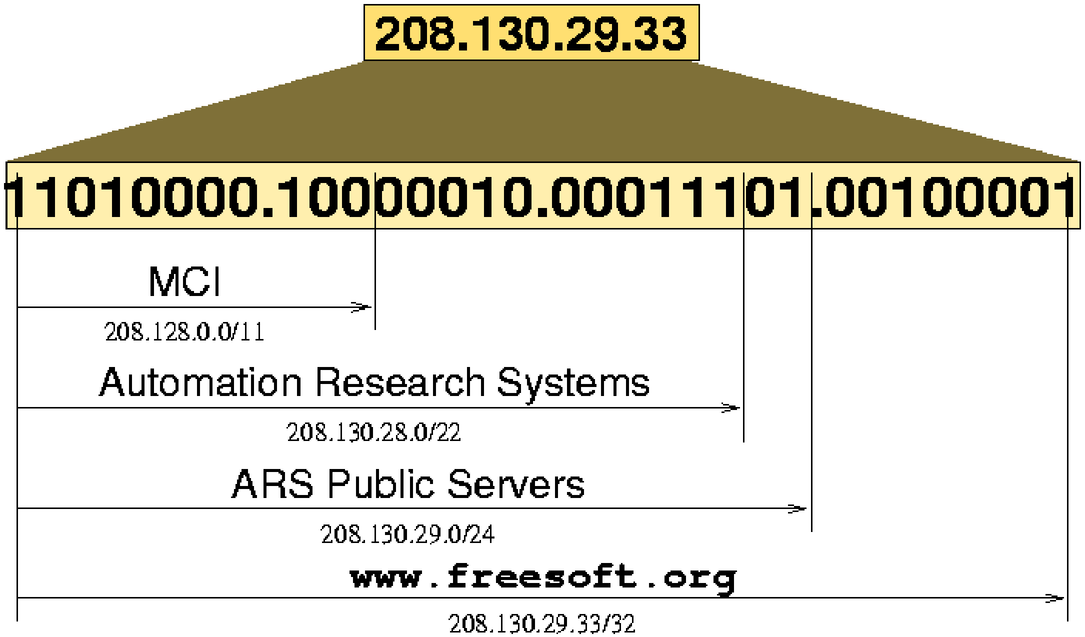
The three bears problem

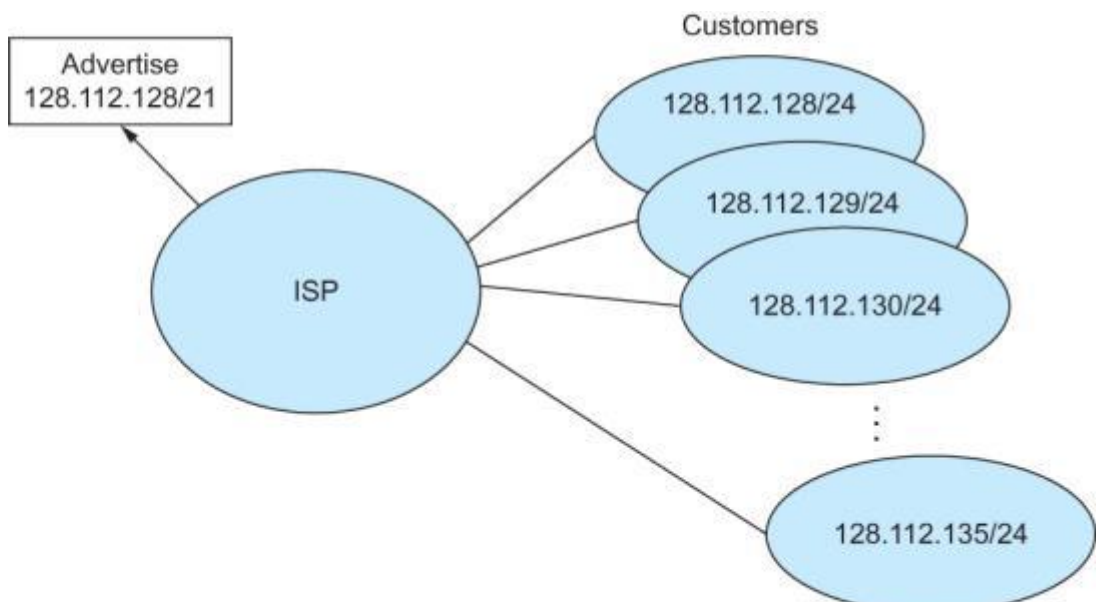
- For most organizations:
  - Class A network too big
  - Class C network too small
  - Class B network... just right
- Actually class B too large for most
  - Half of all class B holders had 50 or fewer hosts
  - 16,384 class B not enough for widespread popularity of the interpipes

Classless addressing

- Classless Interdomain Routing (CIDR)
  - We want:
    - Efficient address allocation
    - Small and fast forwarding tables
- Compromise:
  - Aggregate contiguous blocks of IP addresses
  - /X notation
    - Specify how many prefix bits are network number

CIDR examples





IP forwarding with CIDR

- CIDR prefixes 2-32 bits
  - May have overlapping prefixes in forwarding table
  - Example:
    - Forwarding table: 171.69 (16-bit prefix)
    - Forwarding table: 171.69.10 (24-bit)
    - Destination: 171.69.10.5, matches both
  - Router uses longest match

Private IP addresses

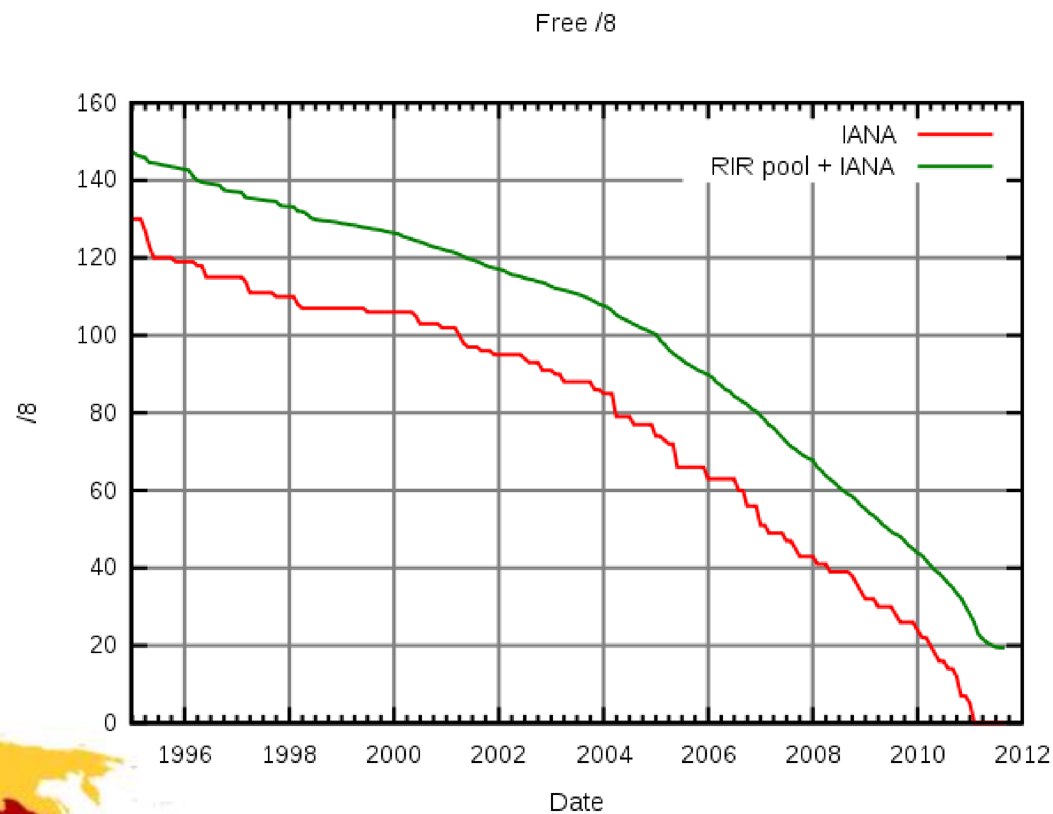
- Private networks (home networks, etc.)
  - Use specified part of IP address space
  - Not globally routable

IP address range	number of addresses	*classful* description	largest CIDR block (subnet mask)	host id size
10.0.0.0 – 10.255.255.255	16,777,216	single class A	10.0.0.0/8 (255.0.0.0)	24 bits
172.16.0.0 – 172.31.255.255	1,048,576	16 contiguous class Bs	172.16.0.0/12 (255.240.0.0)	20 bits
192.168.0.0 – 192.168.255.255	65,536	256 contiguous class Cs	192.168.0.0/16 (255.255.0.0)	16 bits

http://xkcd.com/742/

IPv4 address exhaustion

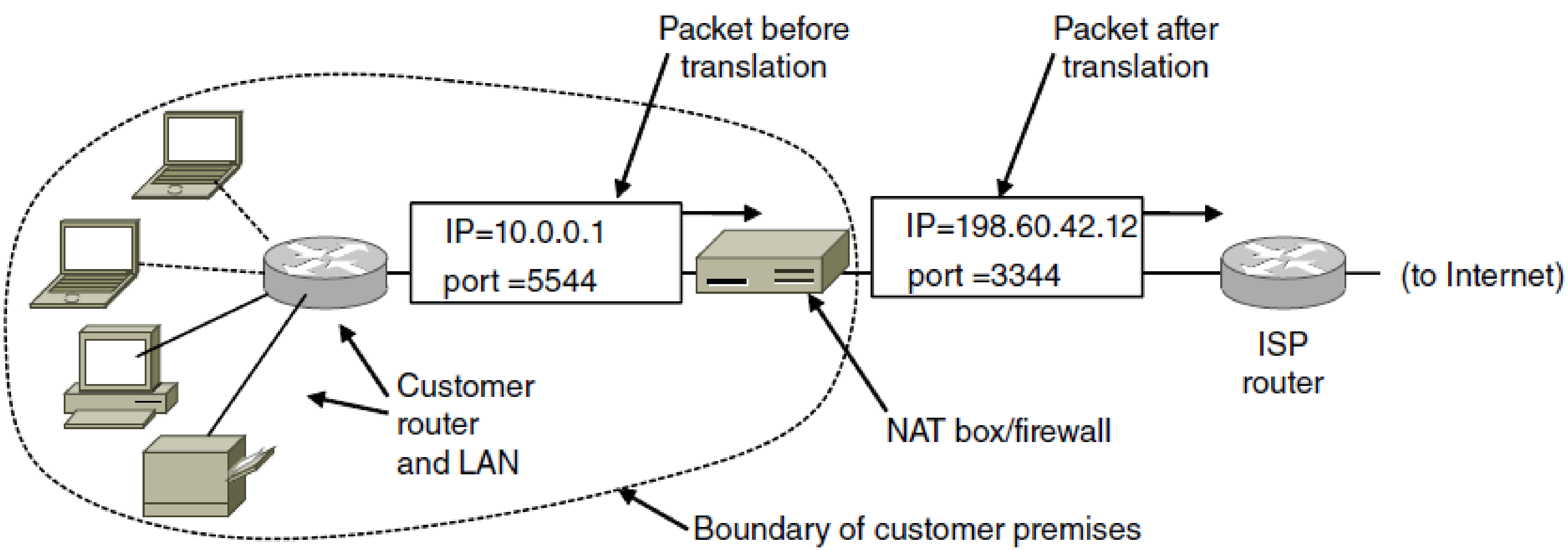
- Jan 31, 2011
  - Last unreserved IANA /8 blocks allocated
  - 5 remaining blocks allocated to each of 5 Regional Internet registries (RIR)



http://www.youtube.com/watch?v=y8WqJum_Gfg

NAT

- Network address translation (NAT)
  - Quick fix to address scarcity
  - Home/business gets one public IP
    - Private IP addresses for all hosts inside network
  - NAT box translates at boundary to public IP



NAT design

- Problem: Where to route reply from remote server?
  - NAT designers observed:
    - Most IP traffic over TCP/UDP
    - TCP/UDP have a 16-bit integer port #
      - Source port and destination port (e.g. 80 for web)
- Solution: Use source port as an index into a translation table

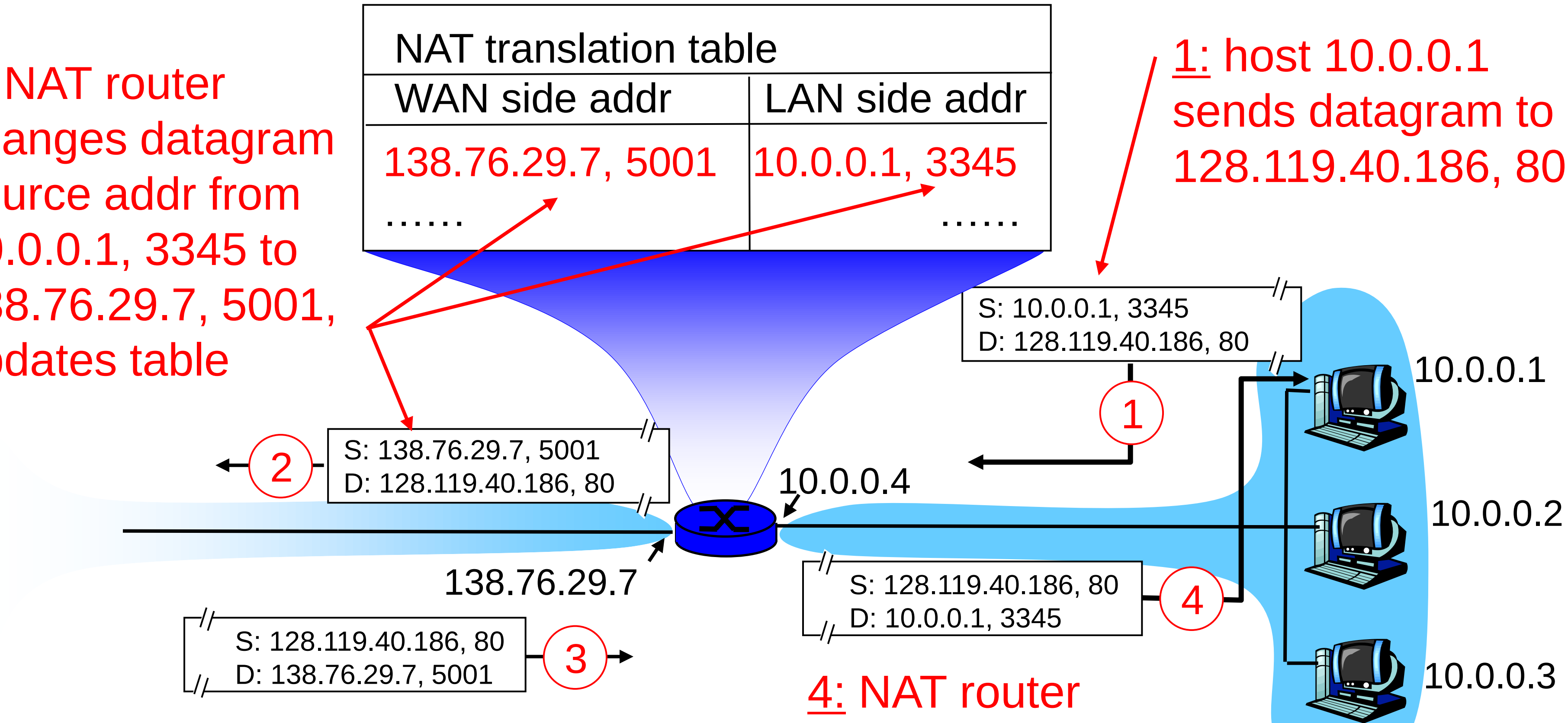
NAT translation

- Map outgoing packets
  - Replace src addr → NAT box addr (public IP)
  - Replace src port # → new port #
- Maintain a translation table
  - (src addr, port #) → (NAT addr, new port #)
  - Free up entry after timeout (frees up port #)
- Incoming packets
  - Consult translation table
  - Rewrite packet and send to local host

NAT example

2: NAT router changes datagram source addr from 10.0.0.1, 3345 to 138.76.29.7, 5001, updates table

1: host 10.0.0.1 sends datagram to 128.119.40.186, 80



3: Reply arrives dest. address: 138.76.29.7, 5001

4: NAT router changes datagram dest addr from 138.76.29.7, 5001 to 10.0.0.1, 3345

Where is NAT implemented?

- Home router
  - Integrates router, DHCP server, NAT, firewall, etc.
  - Single IP address on WAN side from service provider
- Campus or corporate network
  - NAT box at Internet connection point
  - Share a collection of public IP addresses
    - Allows many hosts inside network

NAT advantages

- Helps converse IPv4 addresses
- Easy to switch Internet providers
  - All your devices are using private IPs via DHCP
- Provides a measure of security
  - Outside computers cannot initiate connections
  - However, doesn't protect against:
    - Connections initiated from behind the NAT box to bad places
    - Attacks from hosts inside network

NAT an abomination?

1) Violates the IP model
   – Every host should have unique identifier
2) Breaks end-to-end connectivity model
   – Any host should be able to send a packet to any other host at any time
3) Not connectionless
   – NAT box has state, effectively circuit switching
   – Single point of failure
4) Network layers are not independent
   – NAT looks into the payload

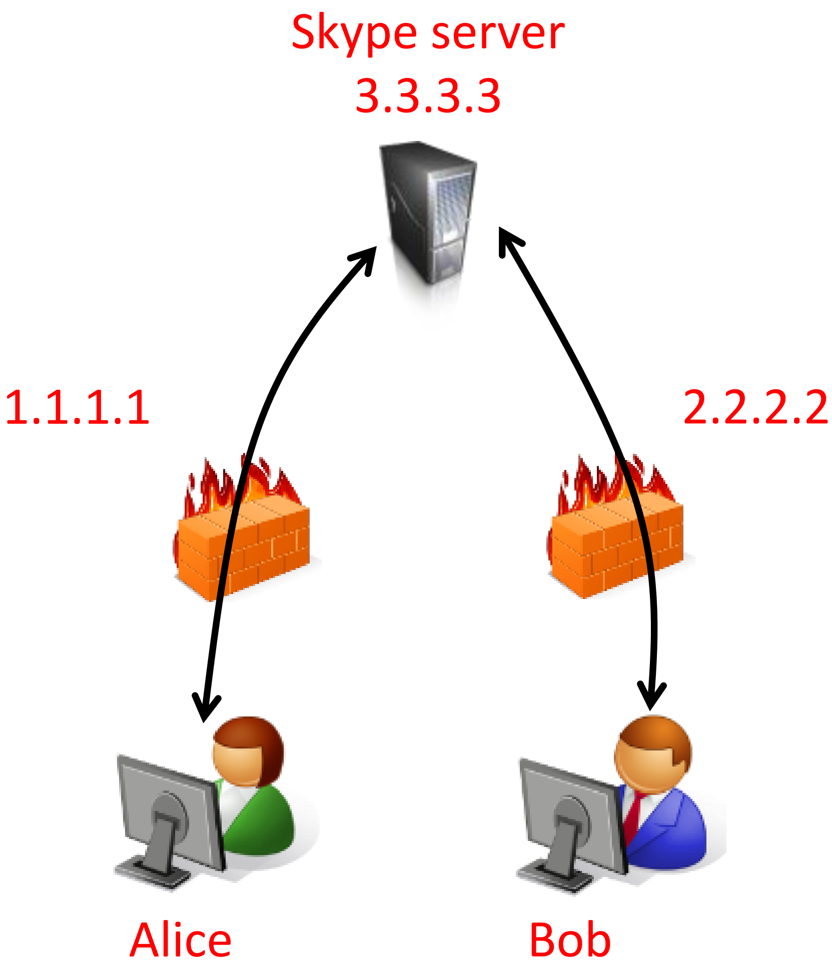
NAT an abomination?

5) Forces use of TCP/UDP protocols
   - Anything else, NAT fails to find TCP Source port
6) Breaks if multiple TCP/IP or UDP ports
   - e.g. FTP and H.323 Internet telephony
7) Limited number of hosts on NAT box
   - Only 16-bits in TCP Source port
   - Can't have > 64K machines on a single IP

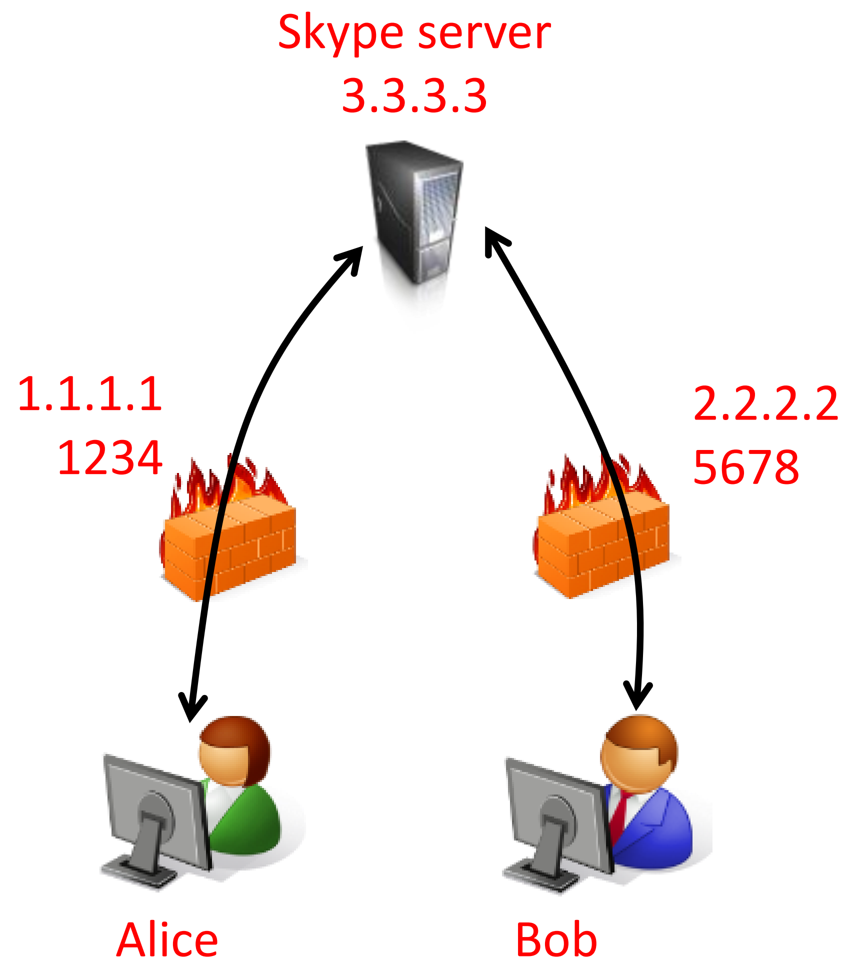
NAT traversal

- Make connections through NAT boxes
  - Client-to-client apps:
    - Voice over IP, video conference, file sharing, gaming
  - One option: UDP hole punching
    - Goal: establish UDP connection between clients
    - Approach: Use central server with public IP to coordinate. Establish direct UDP connections between clients.

UDP hole punching

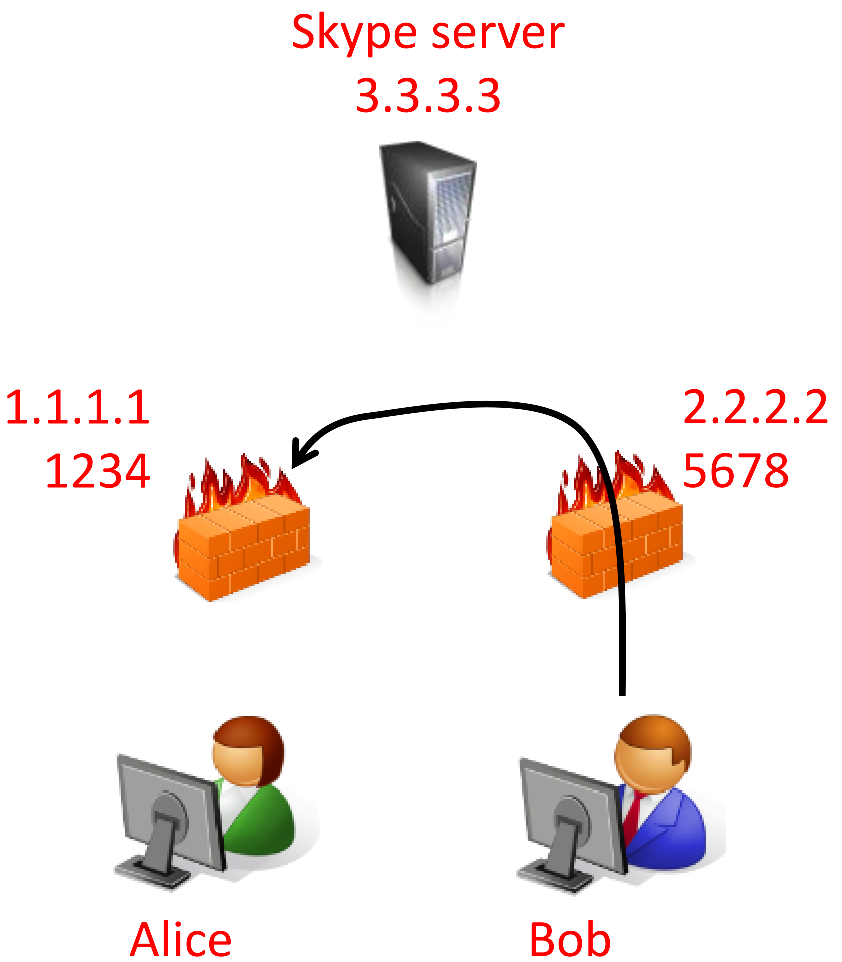


1. Permanent TCP connections to public central server.

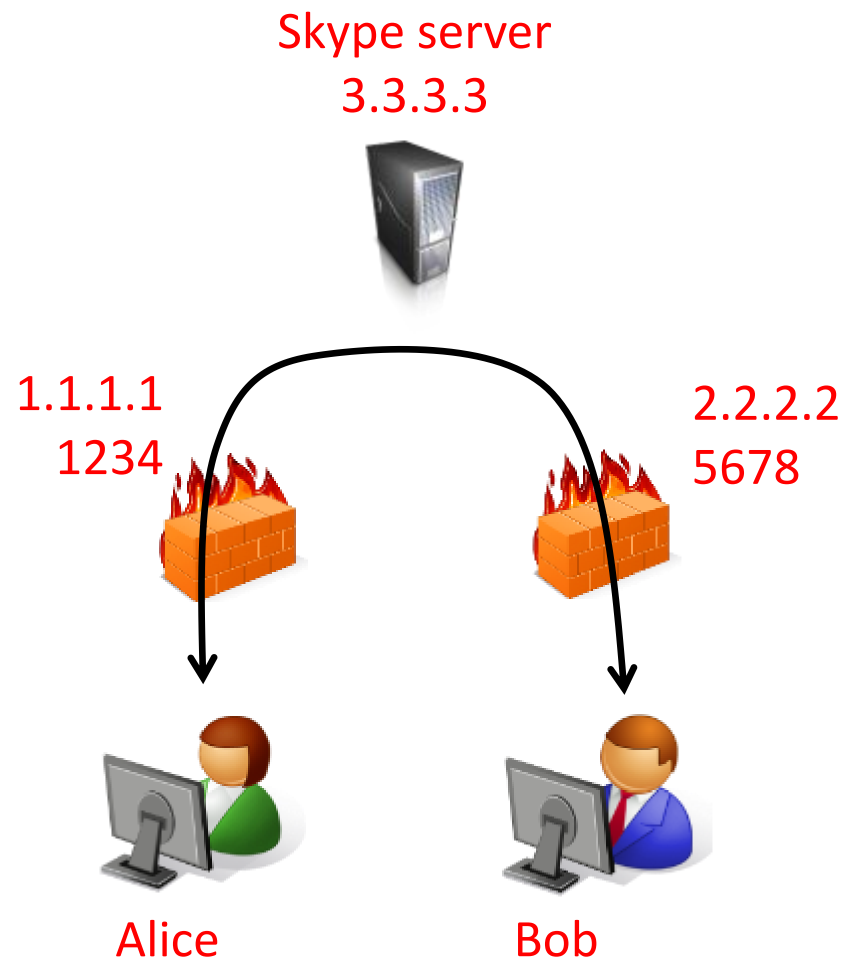


2. Tests reveals UDP port Alice and Bob use to send voice data.

UDP hole punching



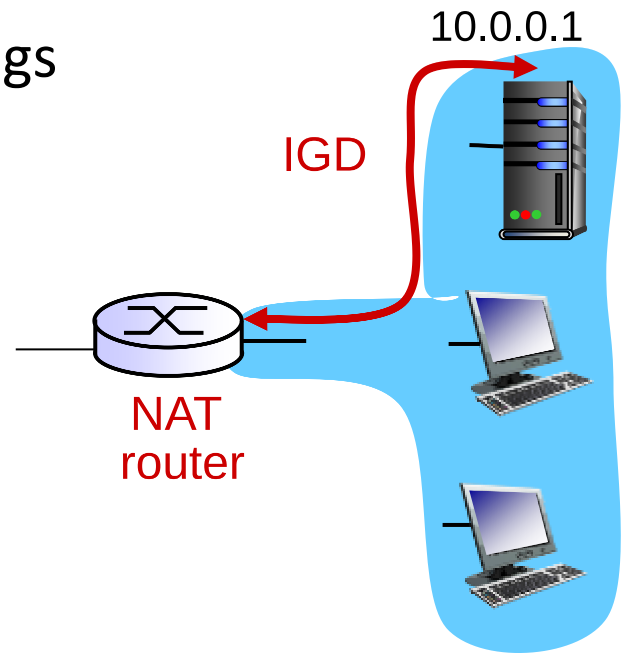
3. Bob sends Alice UDP packet on port 1234. Alice's firewall drops.



4. Alice sends Bob UDP packet on port 5678. Bob's firewall thinks it is a response to his blocked initial packet.

NAT traversal

- Make connections through NAT boxes
  - Another option: Universal Plug and Play (UPnP) Internet Gateway Device (IGD) protocol
  - Allows NAT'd hosts to:
    - Learn public IP address of WAN side of NAT box
    - Add/remove port mappings (with lease times)
    - Enumerate existing port mappings



Summary

- Internet Protocol fragmentation
  - Split at router if IP packet too big for link-layer
- IP addressing
  - Global hierarchical name
  - IPv4, original version, 2^{32} addresses
  - CIDR address, specifies range of IP space
- Network Address Translation (NAT)
  - Helps conserve IPv4 addresses
  - Provides some measure of security
  - UPnP, allows host to configure NAT