

Internet Network Protocols

IPv4/ IPv6

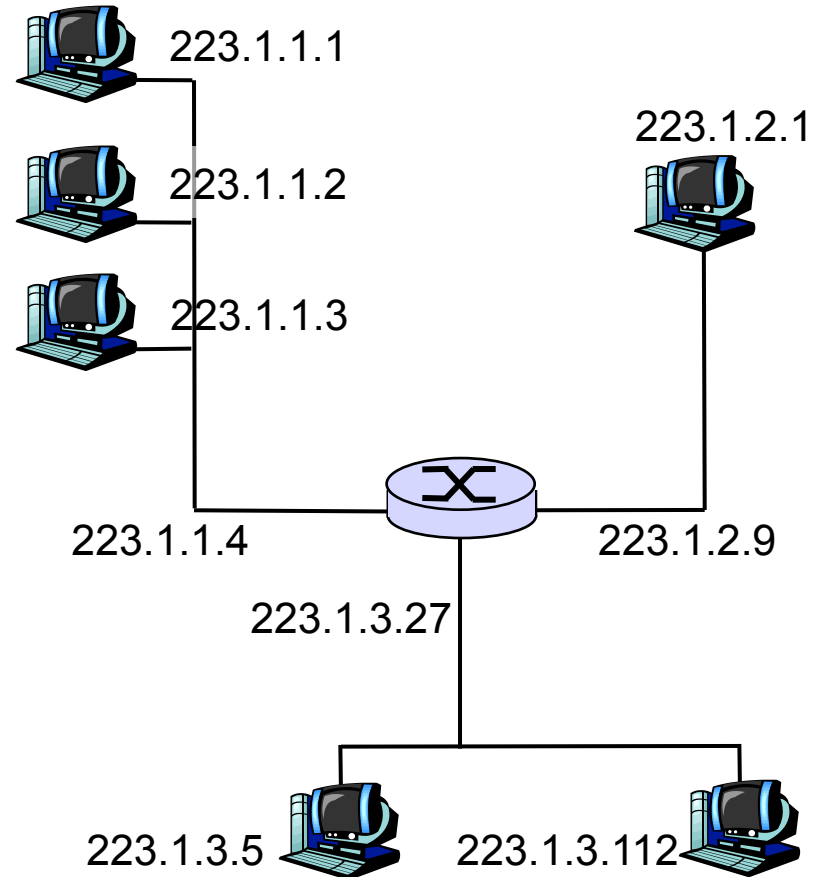
Prof. Anja Feldmann, Ph.D.

anja@inet.tu-berlin.de

TCP/IP Illustrated, Volume 1, W. Richard Stevens
<http://www.kohala.com/start>

IP Interfaces

- ❑ **IP address:** identifier for host or router *interface*
 - IPv4: 32 bit long
 - IPv6: 128 bit long
- ❑ **Interface:** connects a host or router to a physical link
 - Routers typically have multiple interfaces
 - Host may have multiple interfaces
 - IP addresses are associated with interfaces, not hosts or routers

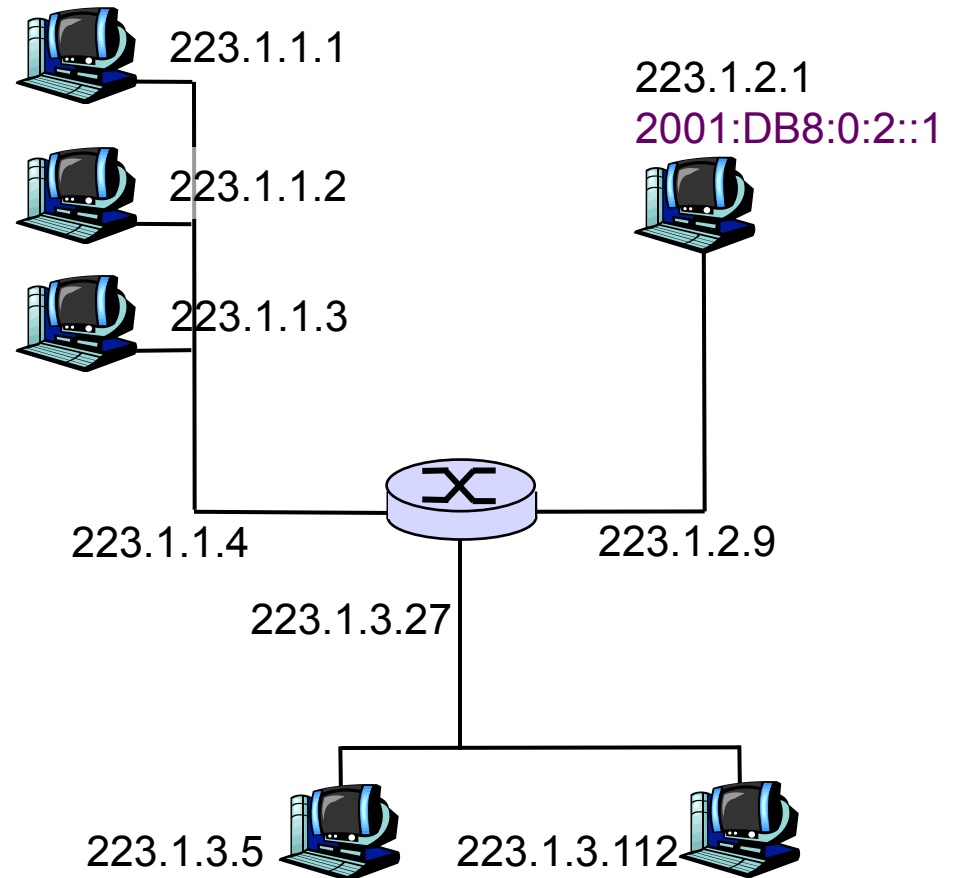


IPv4 Addressing

❑ IP address: Identifier for host or router *interface*

- ❑ IPv4 address:
 - 32 bit
 - written as 4x 8bit in decimal

223.1.1.1 = $\underbrace{11011111}_{223} \underbrace{00000001}_{1} \underbrace{00000001}_{1} \underbrace{00000001}_{1}$



IPv6 Motivation and History

- ❑ IPv4 address space is 32 bit
 - quite limited
- ❑ IPv4 was designed in the 1970ies
 - some requirements changed

Timeline:

- 1992** IETF begins discussion about IPv4 successor
- 1995** First IPv6 RFCs published
- 2000** 50% of IPv4 address space assigned
- 2007** All major OS have IPv6 enabled by default
- 2011** IANA IPv4 assigned last IPv4 block
World IPv6 Day – Major sites test IPv6 for a day
- 2012** World IPv6 Launch Day – Major sites enable IPv6

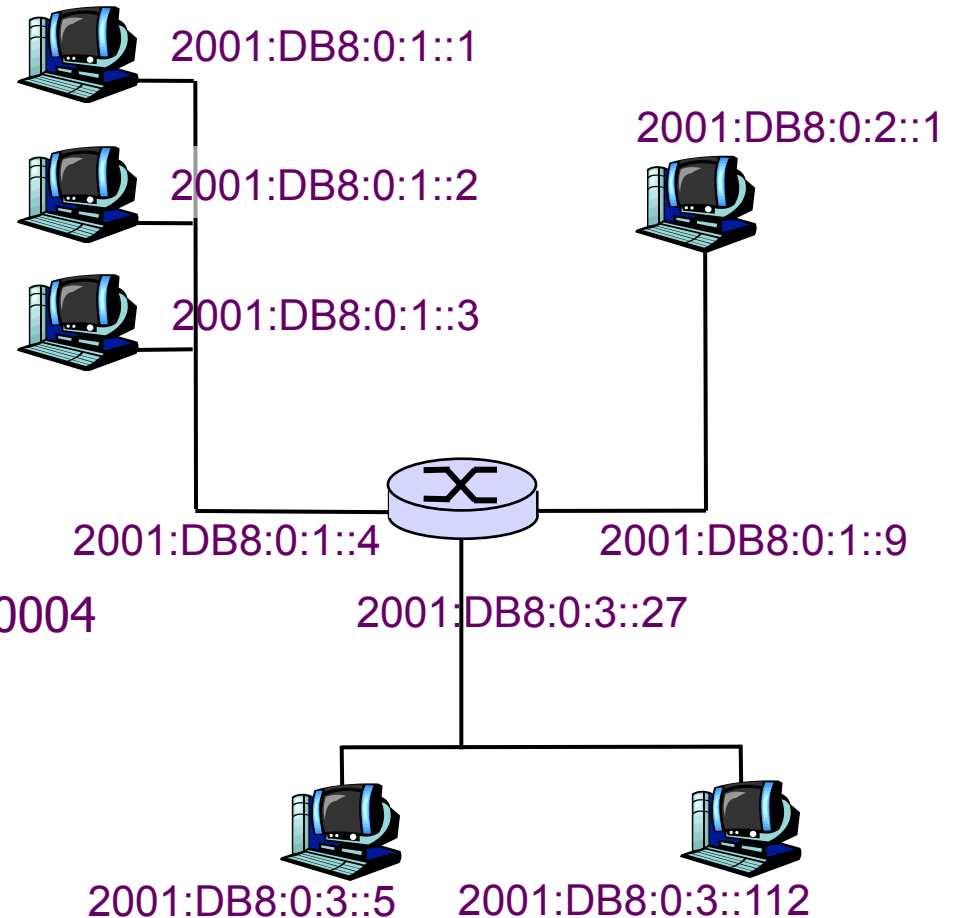
IPv6 Addressing

- ❑ IP address: Identifier for host or router *interface*
- ❑ IPv6 address: 128bit written as 8x 16bit in hex
 - Hextets are separated by colons

2001:0DB8:0000:0001:0000:0000:0000:0004

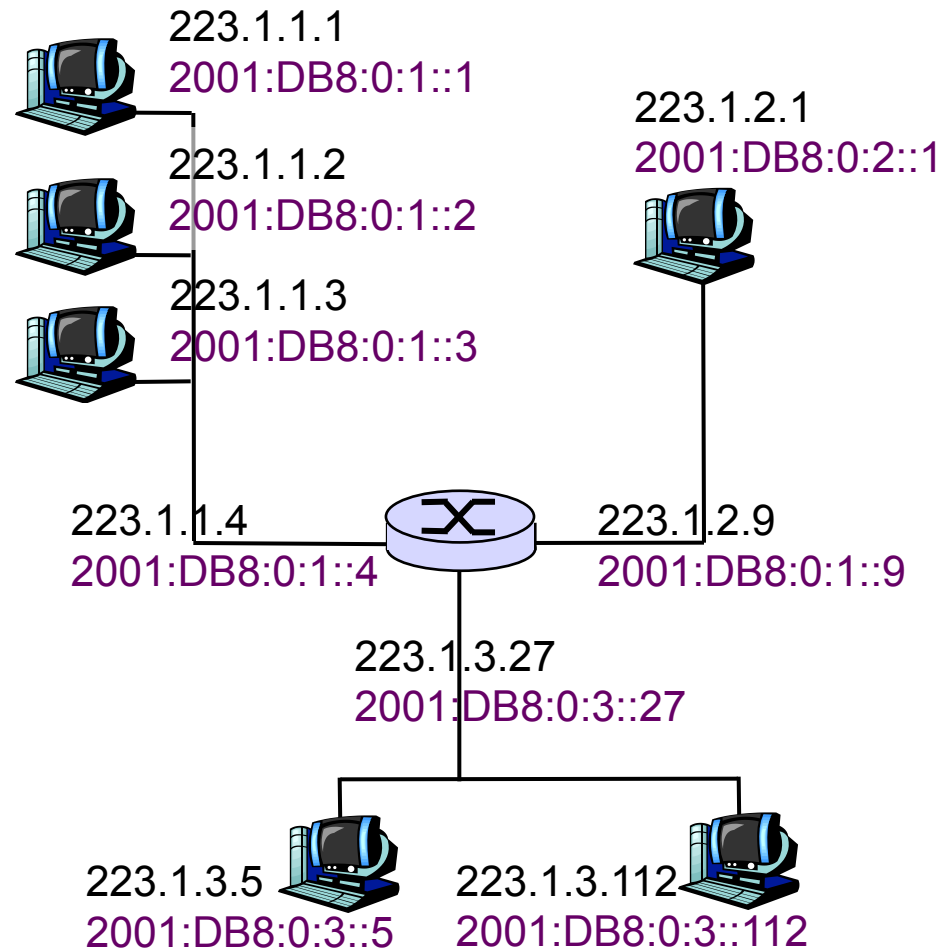
- Leading zeros can be left out
 - Multiple "empty" (all zero) hextets can be abbreviated by a double-colon at one position
- 2001:DB8:0:1::4

shortening multiple times would lead to ambiguous addresses



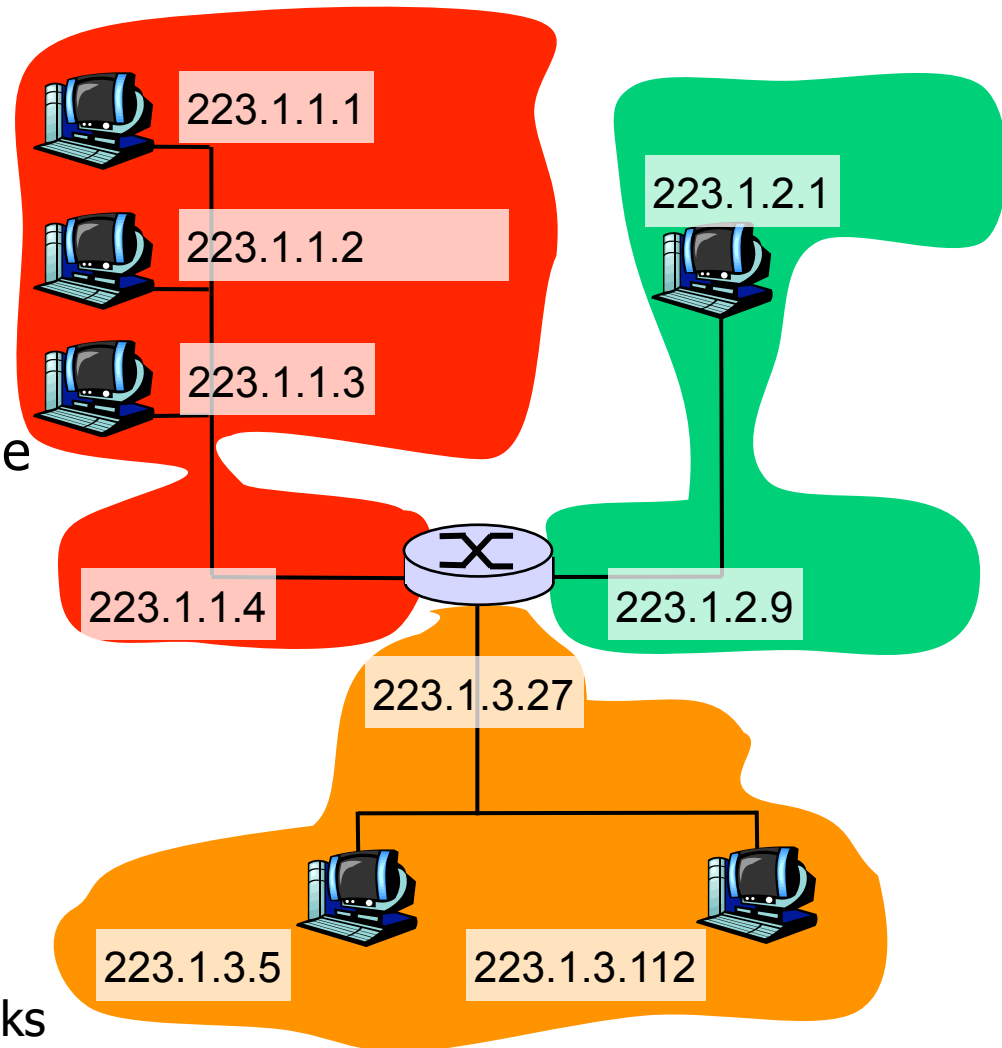
Dual Stack Addressing

- ❑ Multiple identifiers for host or router interface
 - IPv4: 32 bit long
 - IPv6: 128 bit long
- ❑ Host decides whether to use IPv4 or IPv6
 - Routers may support both
 - IPv4 traffic stays IPv4
 - IPv6 traffic stays IPv6
 - IPv4/v6 addresses on the same host are not linked at all at the network layer



IP Network

- ❑ *What's a network?*
(from IP address perspective)
 - Can physically reach each other **without intervening router**
 - Device interfaces with same high order bits of their IP address



Network consisting of 3 IP networks
(for IP addresses starting with 223,
first 24 bits are network address)

IP Network (IPv6)

❑ *What's a network?*

(from IP address perspective)

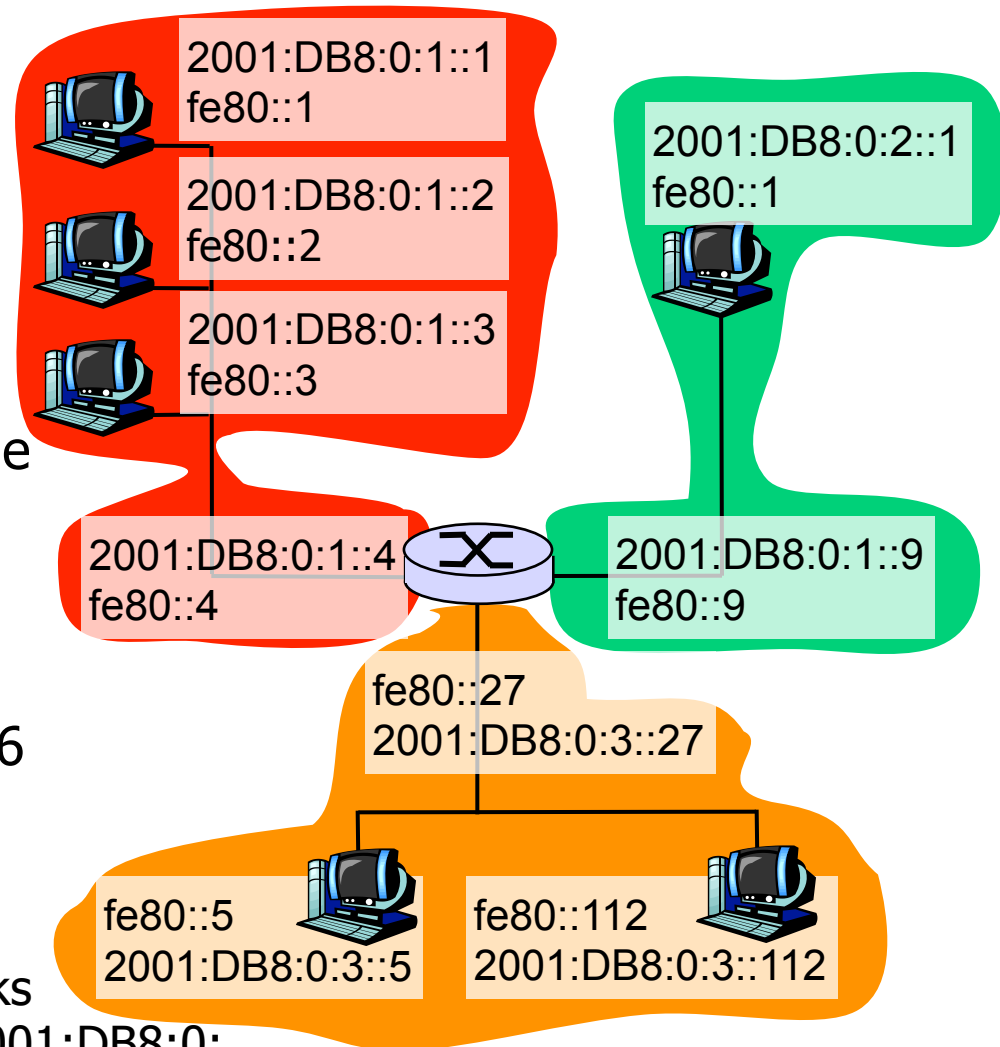
- Can physically reach each other **without intervening router**
- Device interfaces with same high order bits of their IP address

❑ *What's different in IPv6?*

- Usually more than one IPv6 address per host
- Special link-local network

Network consisting of 3 IP networks for IPv6 addresses starting with 2001:DB8:0:

the first 64 bits are network address and link-local-addresses starting with :fe80

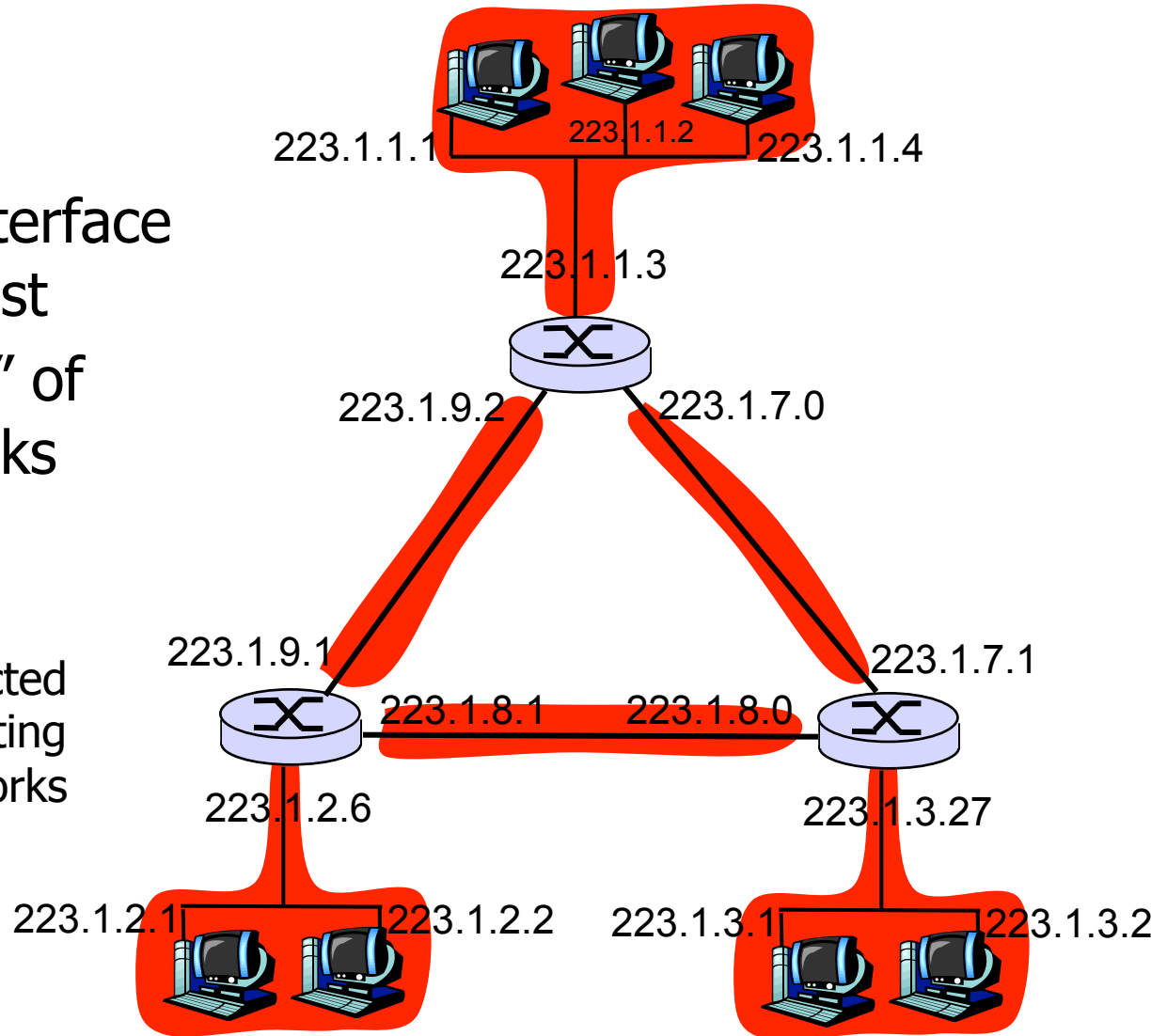


IP Networks (top-down)

How to find the networks?

- ❑ Detach each interface from router, host
- ❑ Create "islands" of isolated networks

Interconnected system consisting of six networks



IP Subnetting

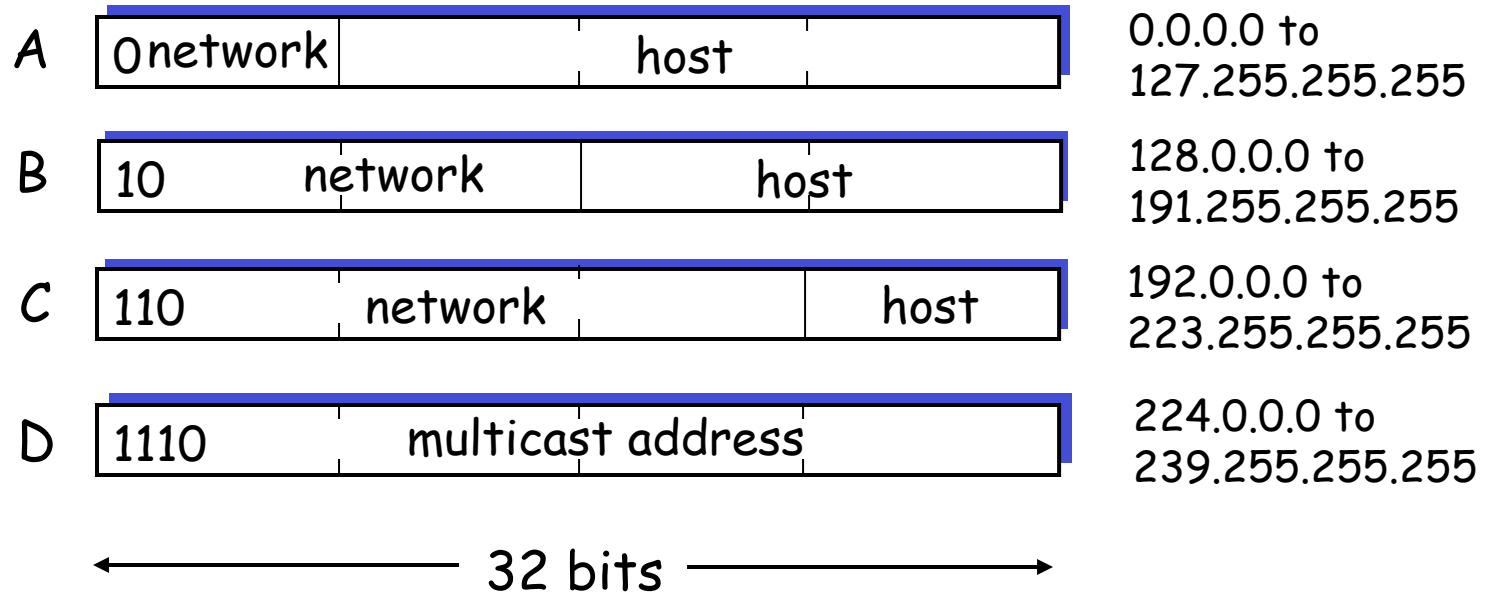
- ❑ Subnetting divides address space in
 - Network part referred to as **prefix**
 - Host address
- ❑ Address format (CIDR)
 - IPv4: a.b.c.d/m 200.23.16.0/24
 - IPv6: x:y:z::/m 2001:DB8:0:3::0/64
- ❑ m: Subnet portion of the address in # of bits; referred to as **prefix length** (bit representation == netmask)



200.23.16.0/24

IPv4 Classful Subnetting (deprecated)

class



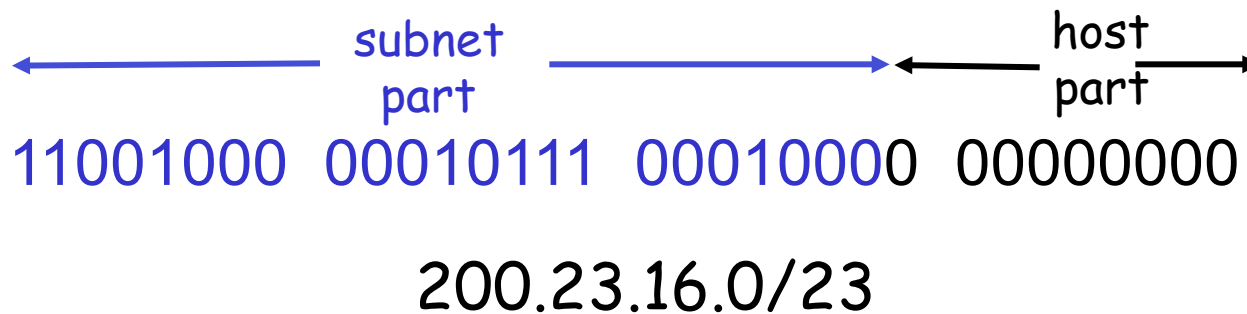
Problem: Wastes IP address space

If you need more addresses than a class C network,
e.g. 256, you need to get at least a class B network (65536)

CIDR (current norm)

CIDR: Classless InterDomain Routing

- ❑ Subnet portion of address of arbitrary length
- ❑ Address format: **a.b.c.d/x**, where x is # bits in subnet portion of address



CIDR

□ Motivation

- Classful class A is too large, classful class C is too small
- Everyone had a Class B network!!!

=>Running out of networks

□ Solution CIDR

- Flexible network masks
 - Sites are given contiguous blocks of class-C addresses and a mask or
 - Parts of former class A/B networks.

Special IP Address Ranges: IPv4

❑ Loopback 127.0.0.0/8

❑ Multicast 224.0.0.0/4

class



224.0.0.0 to
239.255.255.255

← 32 bits →

❑ Private ranges 10.0.0.0/8

172.16.0.0/12

192.168.0.0/16

❑ Link-Local 169.254.0.1/16

Special IP Address Ranges: IPv6

- ❑ Loopback ::1/128
- ❑ **Global Unicast** **2000::/3**
- ❑ Unique Local FC00::/7
- ❑ Multicast FE00::/8
- ❑ Link-Local Unicast FE80::/10

Addresses for use in the Internet are Global Unicast and parts of Multicast.

Link-Local addresses are limited to a physical link (RFC3513).

Multicast Addresses

Addresses a group of hosts at once

- ❑ Useful for streaming and conferencing applications
- ❑ Heavily used in IPv6 for signaling

Only certain ranges usable as multicast

- ❑ IPv6: FE00::/8
- ❑ IPv4: 224.0.0.0/16

Link Local Addresses

Non-routable addresses

- ❑ Can only be used within a network
- ❑ Addresses not unique (!)
- ❑ Heavily used in IPv6 for local signaling

Address ranges used:

- ❑ 169.254.0.1/16 RFC 3927
- ❑ FE80::/10 RFC 4291

Private IP addresses

For local use only - not routable in the Internet

Private IPv4 addresses RFC 1918

- 10/8
- 172.16/12
- 192.168/16

Unique Local IPv6 Unicast addresses RFC 4193

- FC00::/7

How does a host get an IP addresses?

- ❑ Hard-coded by system admin

- ❑ **DHCP / DHCPv6**
Dynamic **H**ost **C**onfiguration **P**rotocol
 - Request the address from a server

- ❑ **IPv6 SLAAC**
Stateless **A**ddress **A**uto-**C**onfiguration
 - Router advertise the IPv6 prefix
 - Hosts add an Interface Identifier as Host-Part

IP Addresses Allocation Process

1. **ICANN** (**I**nternet **C**orporation for **A**ssigned **N**ames and **N**umbers) gives IP address blocks to RIRs
 2. **RIRs** (**R**egional **I**nternet **R**egistries),
i.e. RIPE, ARIN, APNIC, LACNIC, AfriNIC
assign addresses to LIRs
 3. **LIRs** (**L**ocal **I**nternet **R**egistries)
are larger Providers that assign addresses or
address blocks to their customers
- ❑ IPv4 address space
 - None left in the ICANN pool since January 31st 2011,
 - Small blocks at a subset of the RIRs.
 - ❑ IPv6 address space
 - Typical allocation for an LIR: /32
 - Typical allocation for a site: /48 – /56

IP Addresses Allocation Process

Q: What do I do if I don't have a public address?

A1: Recall private IP addresses

- 10/8 RFC 1918
- 172.16/12
- 192.168/16
- FC00::/7 RFC 4193

□ Private use only – not routable in the Internet

A2: Recall link local addresses

- 169.254.0.1/16 RFC 3927
- FE80::/10 RFC 4291

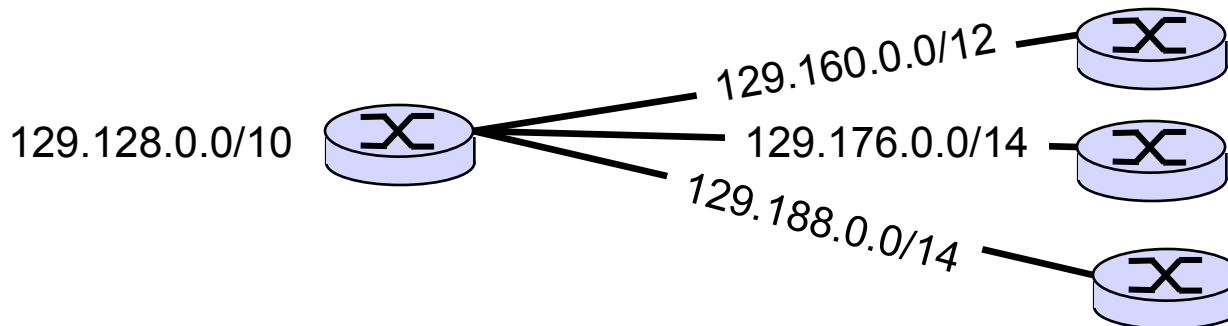
□ Local / single network use only – not routable

Hierarchical address structure

□ Recall: CIDR

128.119.48.12/18 = $\overbrace{10000000\ 01110111\ 00110000\ 00001100}^{18\ \text{relevant bits}}$

- High order bits form the **prefix**
- Once inside the network, can **subnet**: divide remaining bits
- Subnet example:



Note: Picture shows prefix masks, not interface addrs!

□ Forwarding decision: Longest prefix match

Forwarding vs. Routing

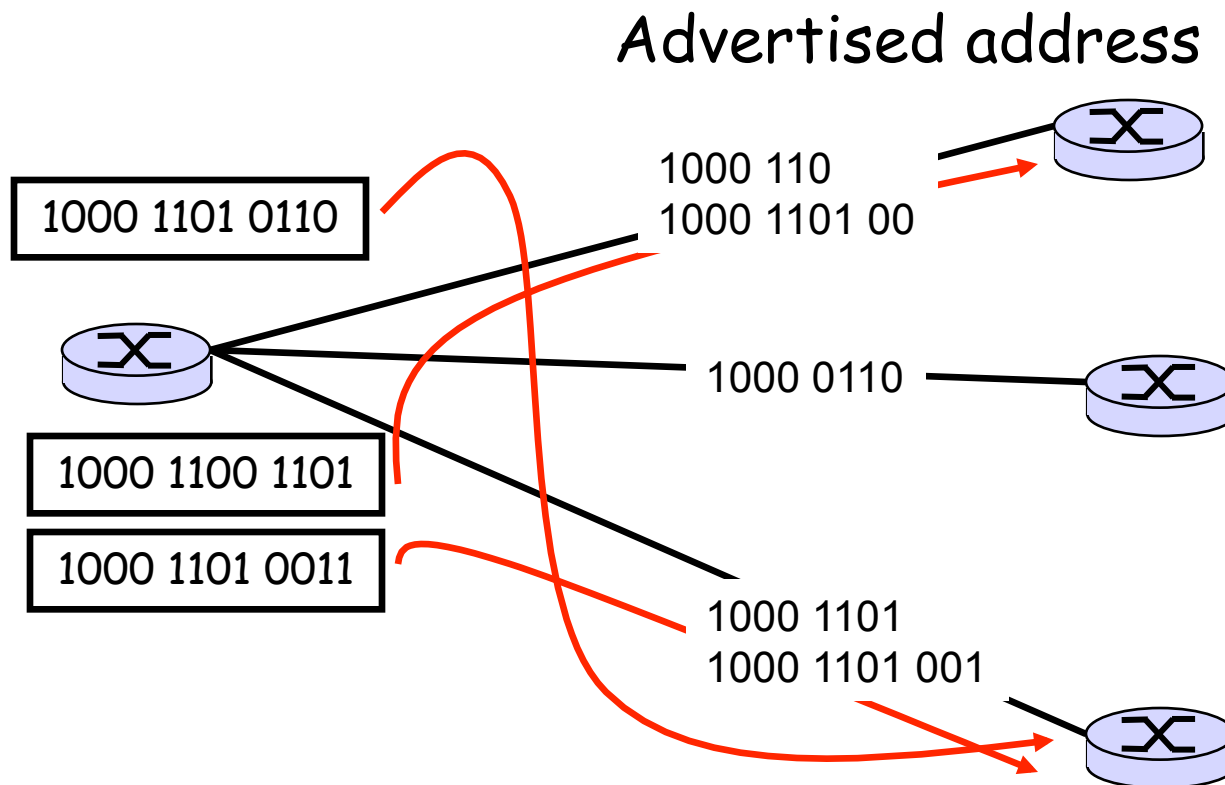
- ❑ **Forwarding:** Process of moving packets from input to output
 - The forwarding table
 - Information in the packet

- ❑ **Routing:** Process by which the forwarding table is built and maintained
 - One or more routing protocols
 - Procedures (algorithms) to convert routing info to forwarding table.

(More later ...)

Forwarding with CIDR

- ❑ Packet should be sent toward the interface with the **longest matching prefix**



Lookup: Longest Prefix Match

- ❑ Forwarding table:
<Network>/<mask> <next-hop>
- ❑ IP Packets: destination IP address
 - Find next-hop via longest prefix match
- ❑ Example (IPv4):

Forwarding table

134.96.252.0/24	A
134.96.0.0/16	C
134.96.240.0/20	B
134.96.252.192/28	B
134.96.252.128/28	A

Packets

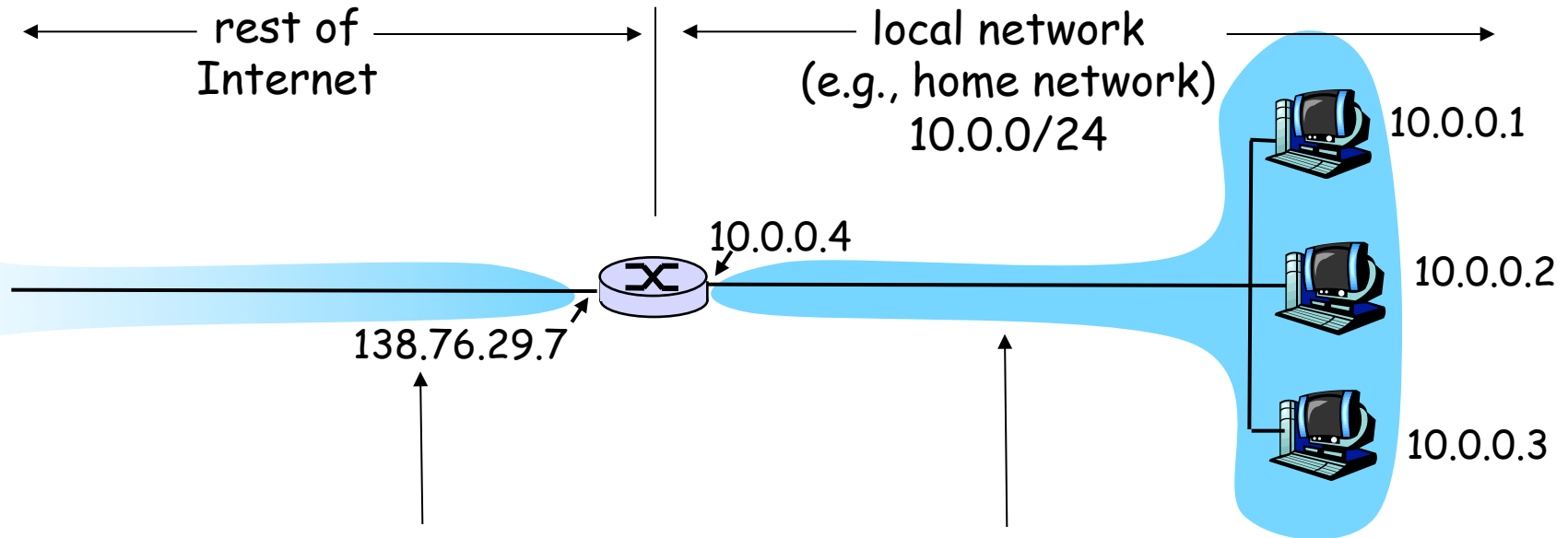
134.96.252.200
134.96.254.2
134.96.239.200
134.97.239.200
134.96.252.191

NAT: Network address translation

Motivation: Local network uses just one IP address as far as outside world is concerned:

- Just one IP address for all devices
- Not needed range of addresses from ISP
- Work around for IPv4 exhaustion (carrier-grade NAT)

NAT: Network address translation (2.)



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 (3.)

Motivation: Local network uses just one IP address as far as outside world is concerned:

- Range of addresses not needed from ISP:
just one IP address for all devices
- Can change addresses of devices in local network without notifying outside world
- Can change ISP without changing addresses of devices in local network
- Devices inside local net not explicitly addressable, visible by outside world.

NAT: Network address translation (4.)

Implementation: NAT router must:

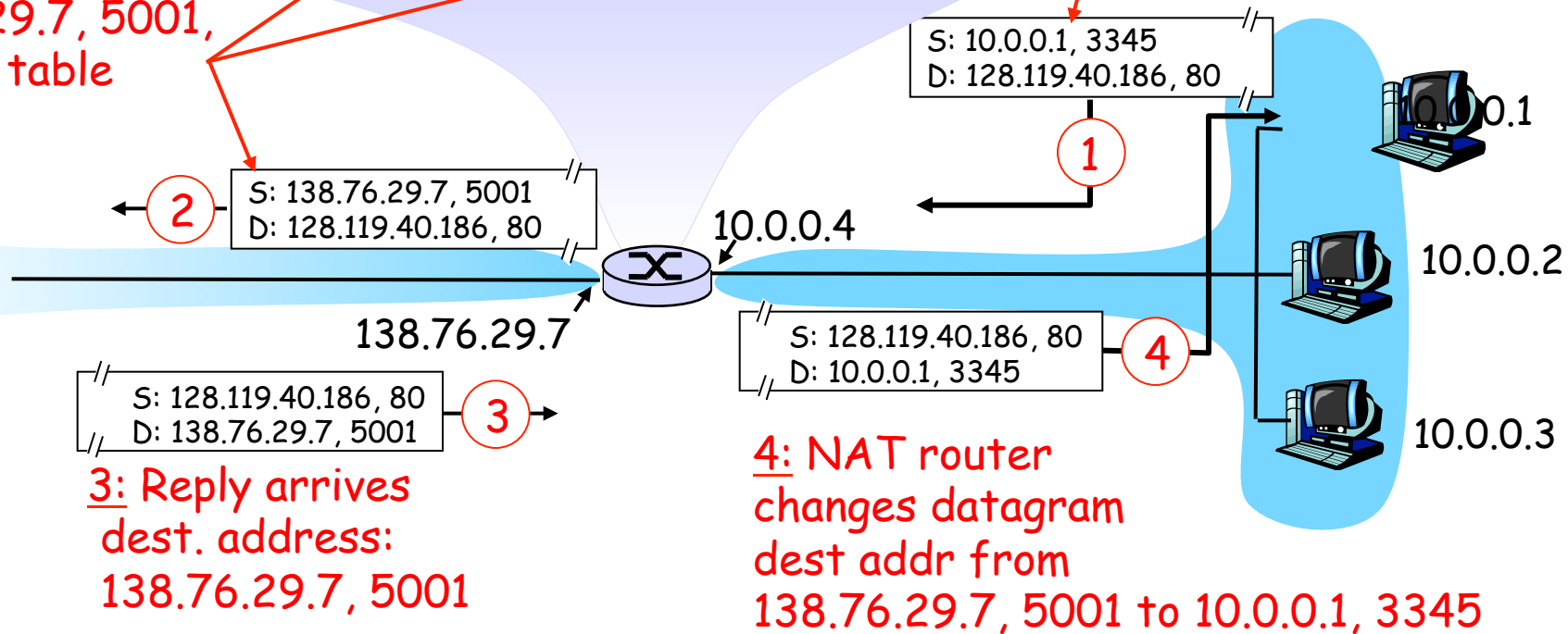
- *Outgoing datagrams: Replace* (source IP address, port #) of every outgoing datagram to (NAT IP address, new port #)
 - . . . remote clients/servers will respond using (NAT IP address, new port #) as destination addr.
- *Remember (in NAT translation table)* every (source IP address, port #) to (NAT IP address, new port #) translation pair
- *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 (5.)

NAT translation table	
WAN side addr	LAN side addr
138.76.29.7, 5001	10.0.0.1, 3345
.....

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



NAT: Network address translation (6.)

- ❑ 16-bit port-number field
 - 60,000 simultaneous connections with a single LAN-side address!
- ❑ NAT is controversial
 - Routers should only process up to layer 3
 - Violates end-to-end argument
 - NAT possibility must be taken into account by app designers, e.g., P2P applications
 - Address shortage should instead be solved by IPv6