



Curtin University

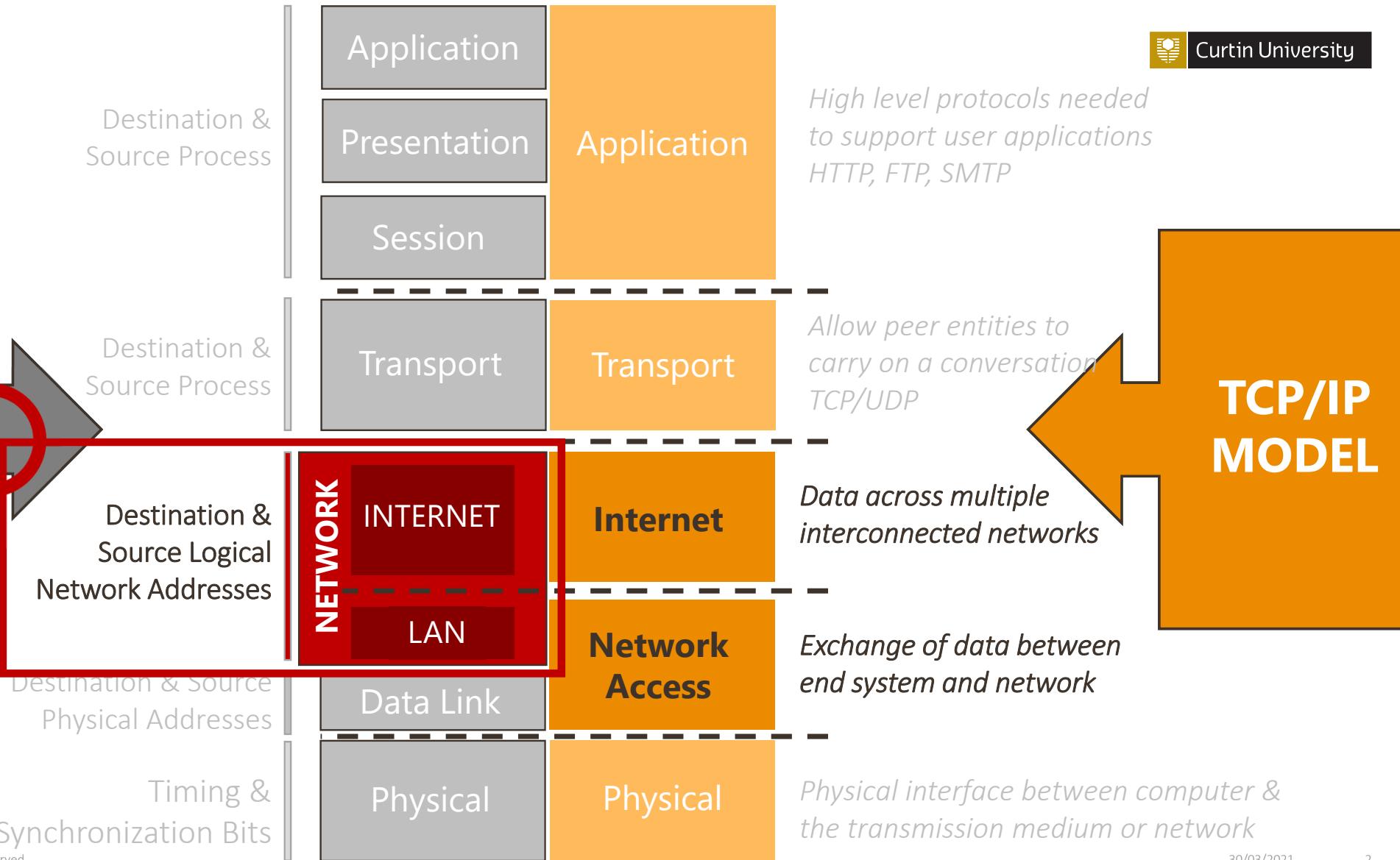
Network Layer I

Prof. Ling Li | Dr. Nadith Pathirage | Lecture 05

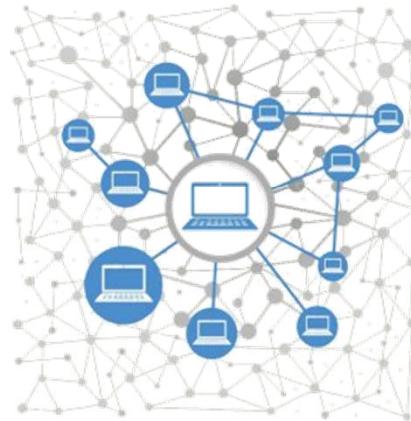
Semester 1, 2021

A GLOBAL UNIVERSITY

WESTERN AUSTRALIA | DUBAI | MALAYSIA | MAURITIUS | SINGAPORE



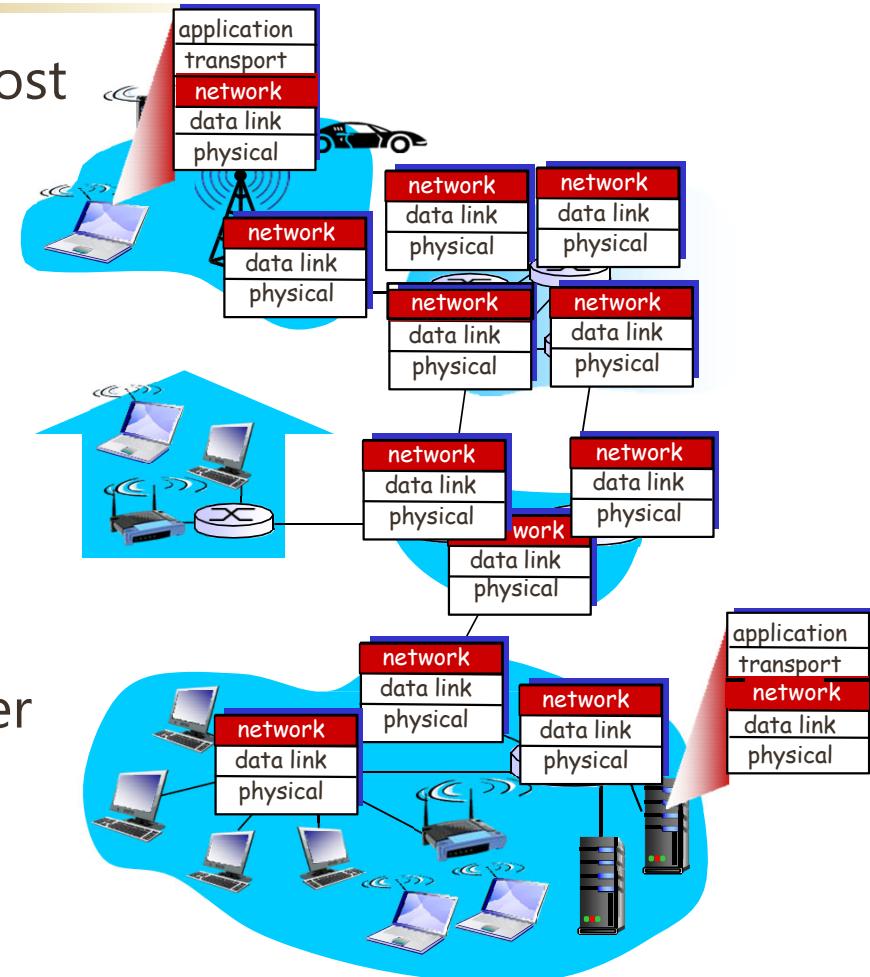
Network Layer



Provides the means of transferring packets from a **source to a destination** host via one or more networks

Network Layer

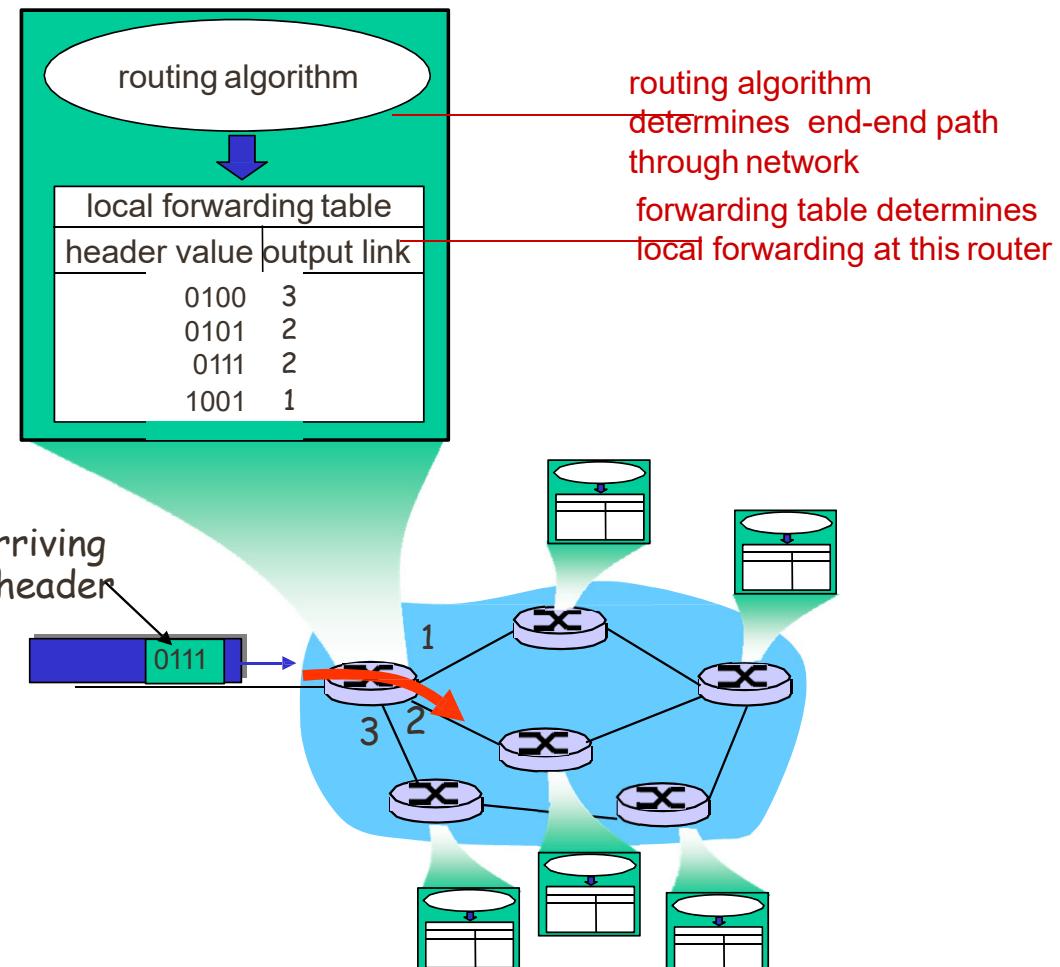
- Transport segment from sending to receiving host
- Sending Side:
 - Encapsulates segments into datagrams
- Receiving Side:
 - Delivers segments to transport layer
- Network layer protocols in every host and router
(layer-3 devices)



Key Functions

1. Store-and-forward

- Move packets from router's input -> appropriate router output
- **Analogy:** process of getting through single interchange



2. Routing

- Determine route taken by packets from source to destination
- **Analogy:** process of planning trip from source to destination (via [routing algorithms](#))

Key Functions – cont.

3. Connection Setup

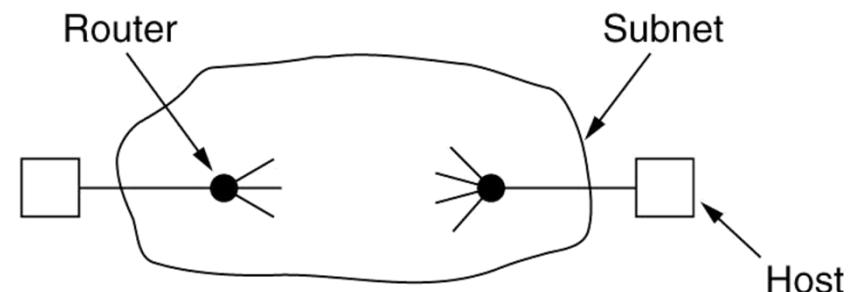
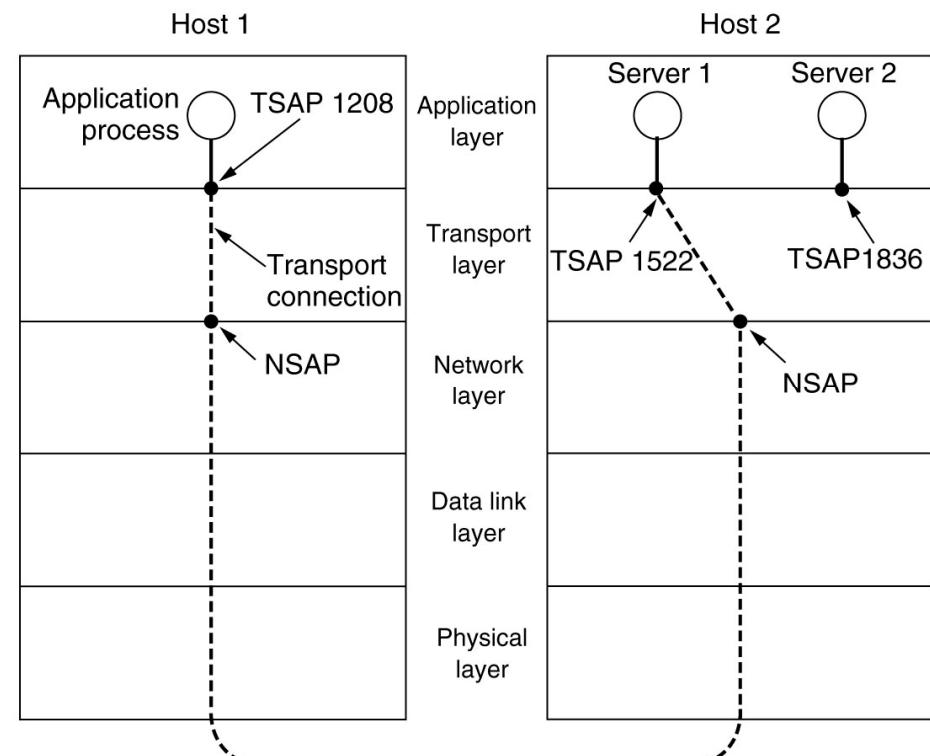
- Only in some network architectures

(i.e. ATM, frame relay, X.25)

- Before datagrams flow: Two end hosts and intervening routers establish *virtual-circuit*

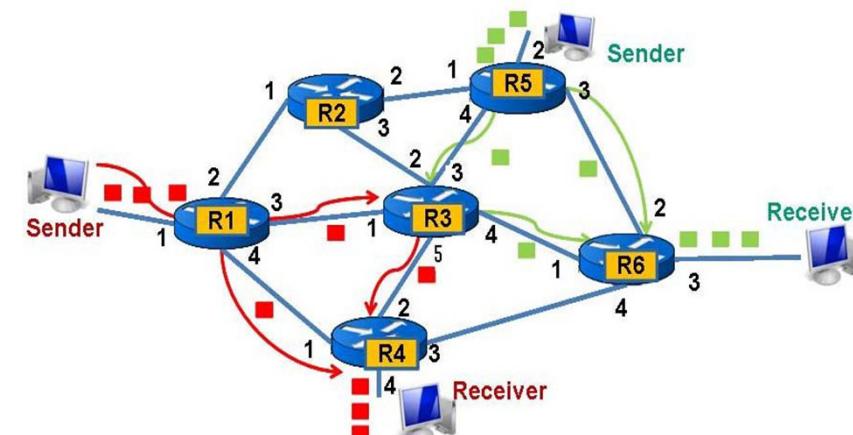
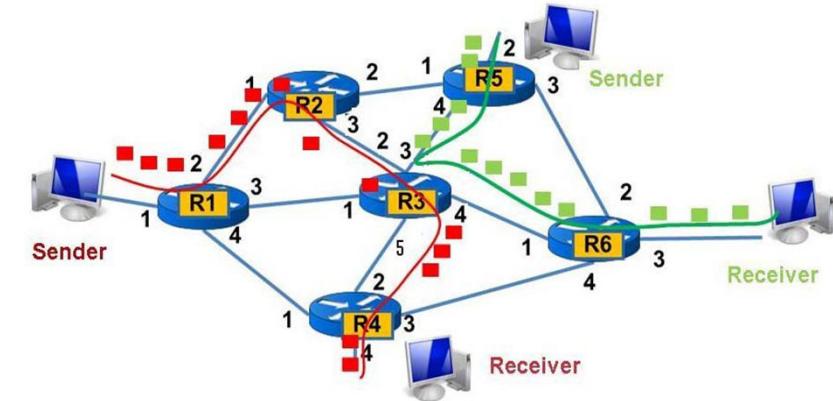
▪ Network vs Transport Layer connection service:

- **Network:** between two hosts (may also involve intervening routers in case of VCs)
- **Transport:** between two processes

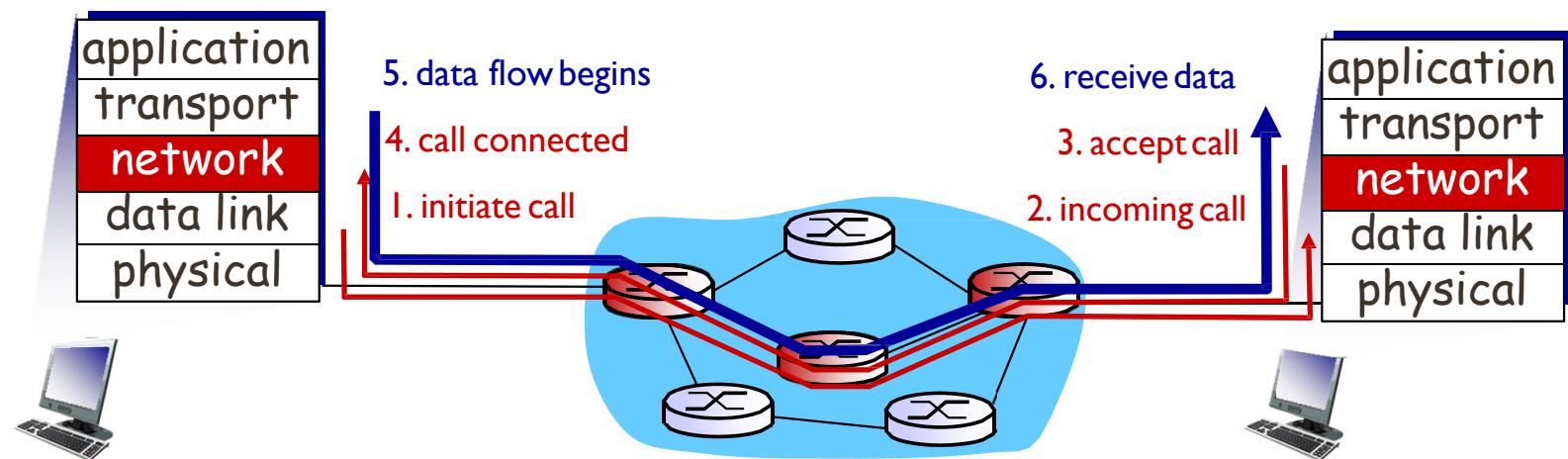


Connection/Connectionless Service

- **Virtual-circuit (PS) Network** provides network-layer **connection service**
- **Datagram (PS) Network** provides network-layer **connectionless service**
- Analogous to **TCP/UDP**, but:
 - **service:** host-to-host
 - **no choice:** network provides one or the other
 - **implementation:** in network core



Virtual Circuit Network

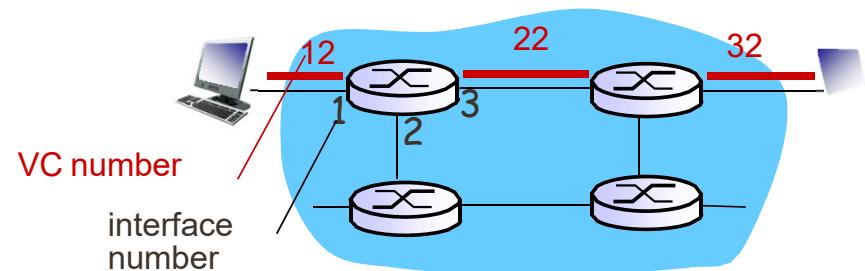


▪ Steps:

- Call setup, teardown for each call before data can flow
- Each packet carries **VC identifier** (*not dest. host address*)
- Every router on source-destination path maintains “state” for each passing connection
- Link, router resources (bandwidth, buffers) may be allocated to VC (dedicated resources = predictable service)

“source-to-destination path behaves much like telephone circuit”

VC Forwarding Table



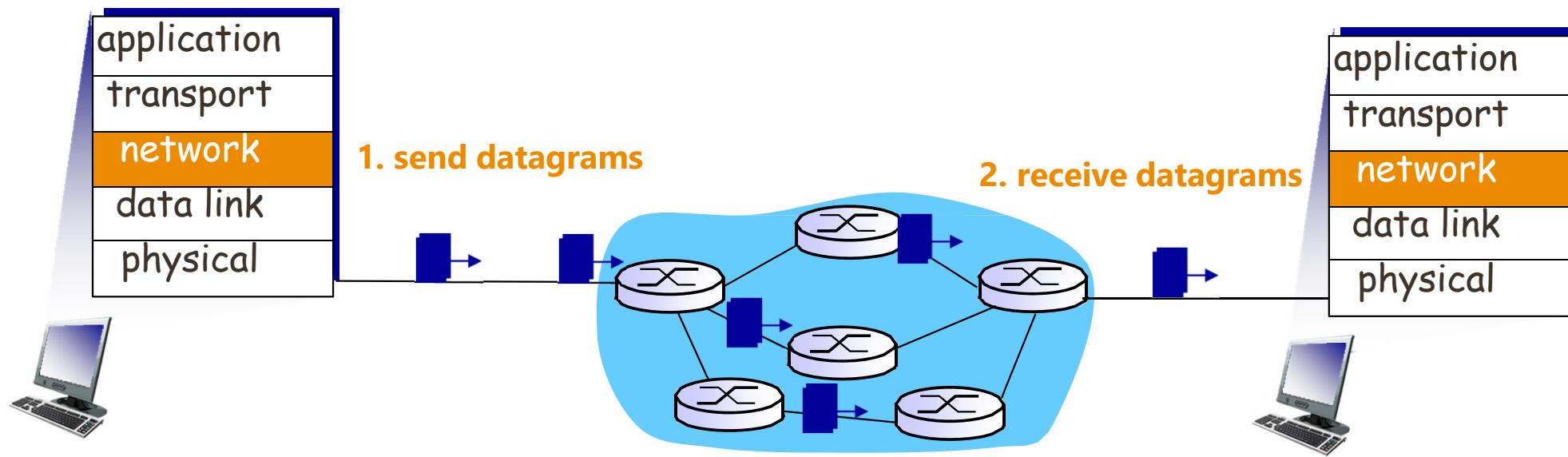
Forwarding table in northwest router:

Incoming interface	Incoming VC #	Outgoing interface	Outgoing VC #
1	12	3	22
2	63	1	18
3	7	2	17
1	97	3	87
...

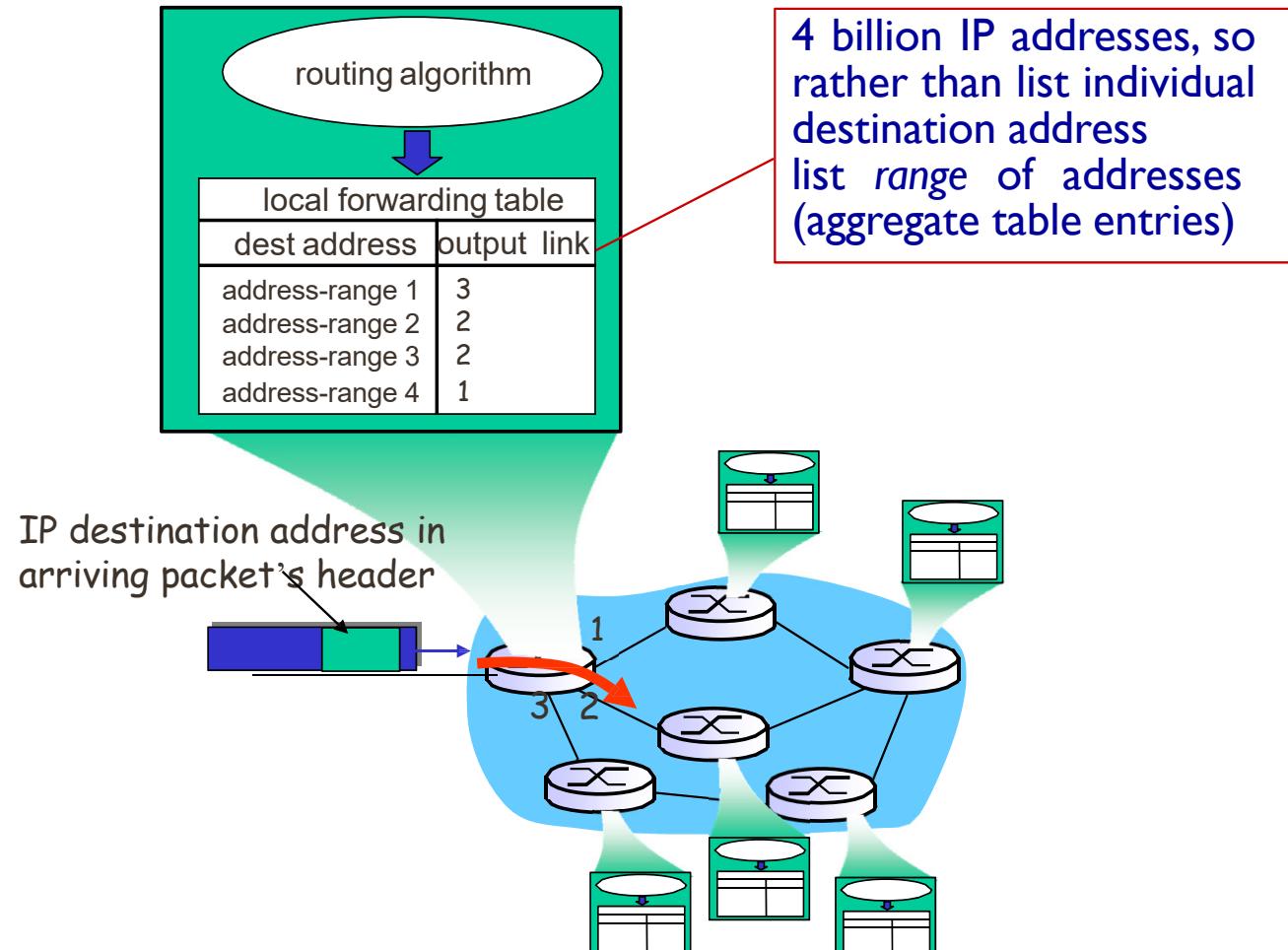
VC routers maintain connection state information!

Datagram Network

- **No call setup** at network layer
- **Routers:** no state about end-to-end connections
- **Packets Forwarding** via destination host address



Datagram Forwarding Table



Datagram Network or VC: why?

▪ Internet (datagram)

- ✓ data exchange among computers
 - “elastic” service, no strict timing req.

▪ many link types

- ✓ different characteristics
- ✓ uniform service difficult

▪ “smart” end systems

(computers)

- ✓ can adapt, perform control,
- ✓ error recovery

**Simple inside network,
complexity at “edge”**

▪ ATM (VC)

- ✓ Evolved from telephony
- ✓ Human conversation
 - Strict timing, reliability requirements
 - Need for guaranteed service

▪ “dumb” end systems

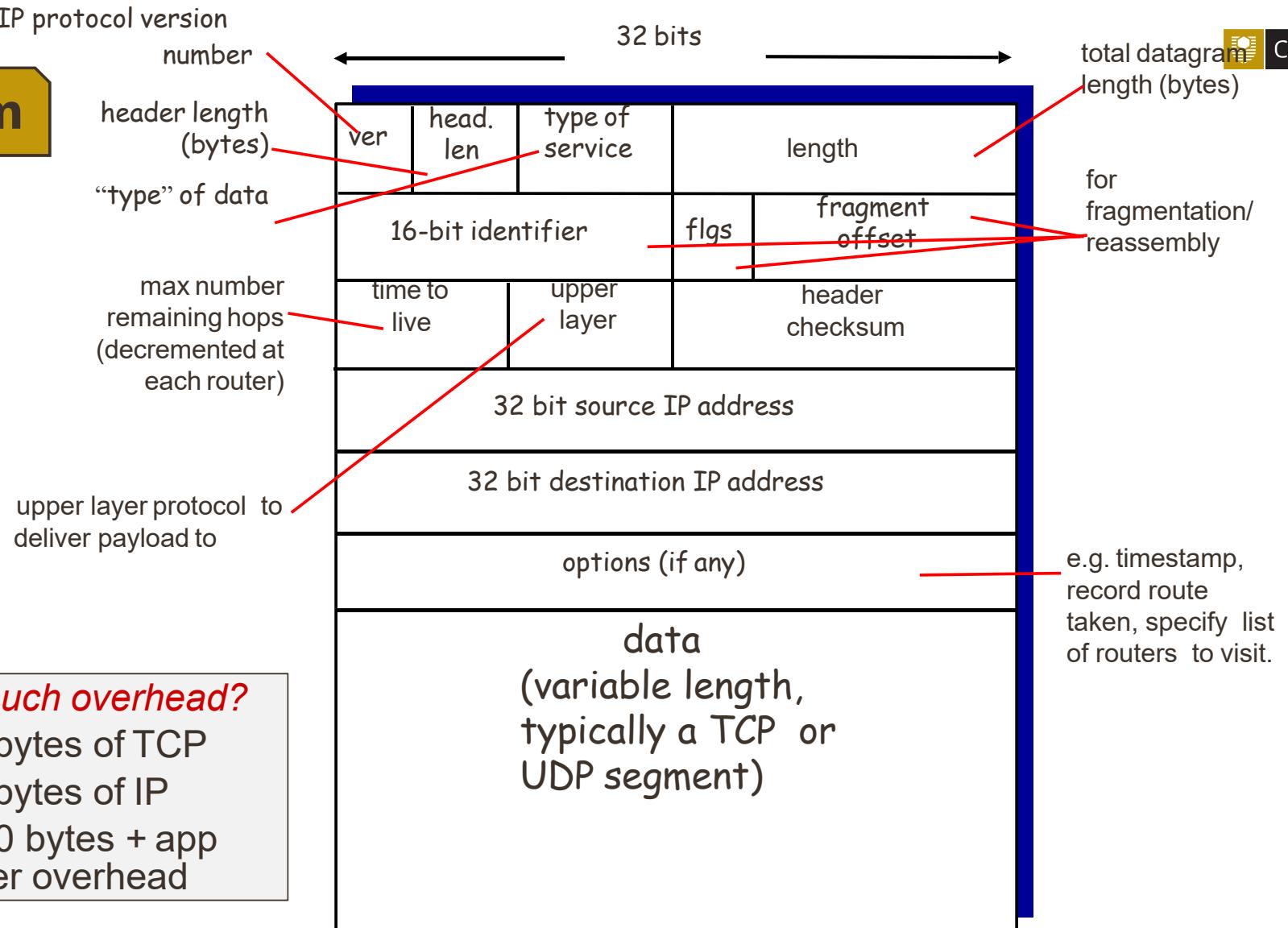
(computers)

- ✓ can adapt, perform control,
- ✓ error recovery

Complexity inside network

IP Datagram

Header

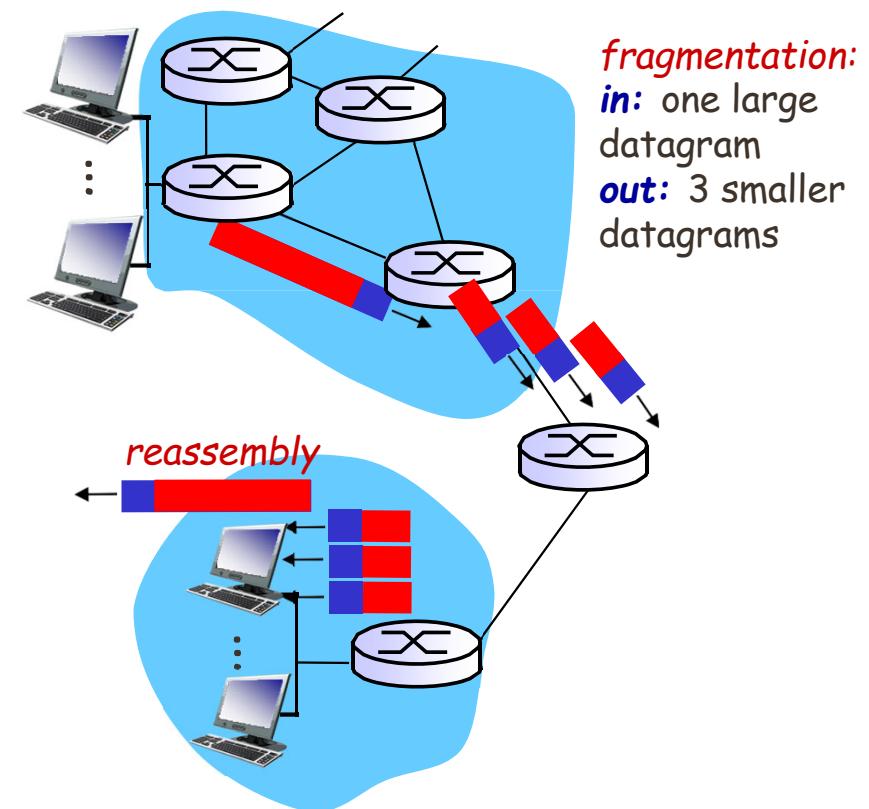


how much overhead?

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

IP Fragmentation, Reassembly

- Network links have **MTU** (Max Transfer Size)
 - Largest possible link-level frame
 - Different link types, different MTUs
- **Fragmentation:** Large IP datagram divided within network
 - One datagram becomes several datagrams
- **Reassembly:** Only at final destination
 - IP header bits used to identify & order related fragments





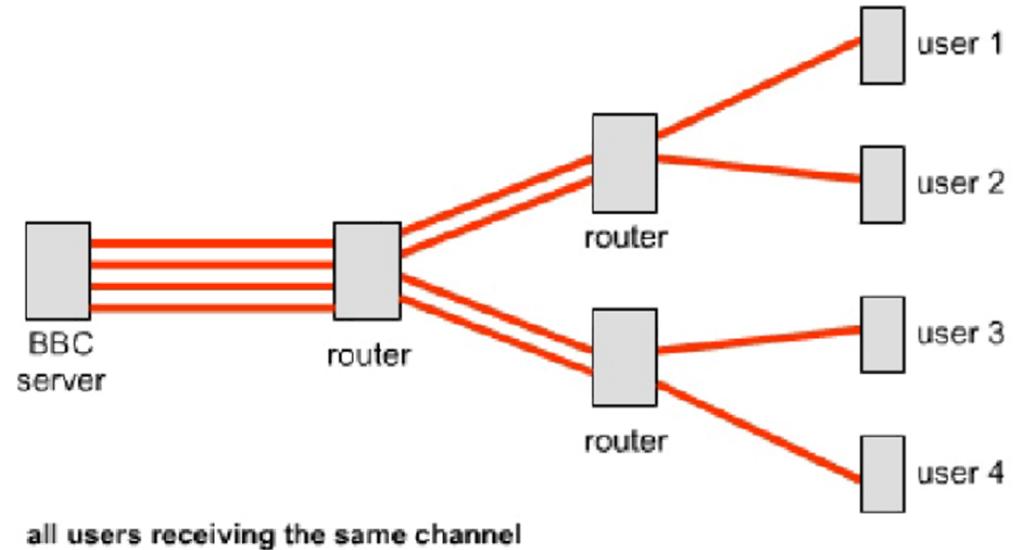
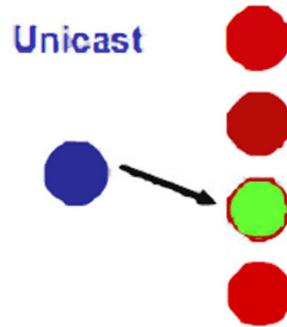
Address Types

- Unicast
- Multicast
- Broadcast
- Anycast
- Geocast

Unicast

- Use IP delivery methods (session-based protocols) such as:

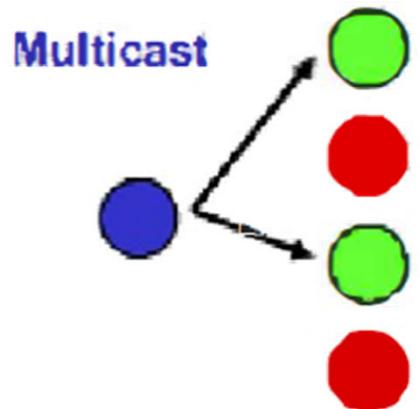
- ✓ Transmission Control Protocol (TCP)
- ✓ User Datagram Protocol (UDP)



- Each unicast client that connects to the server takes up additional bandwidth
- **Client** has a **direct relationship** to the **server**

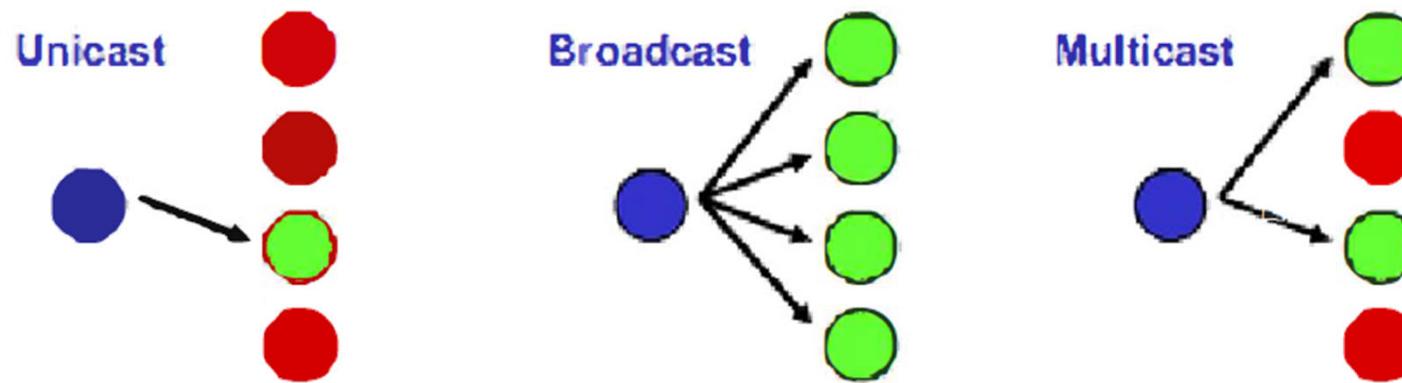
Multicast

- A multicasting is associated with a group of **interested receivers**
- **No direct relationship** between the clients and server
- Clients are connected to the multicast address, **no additional overhead** on the server.
 - i.e. server sends out only one stream per multicast station. same load is experienced on the server whether only one client or 1,000 clients are listening
- Can be used across a **WAN**



Multicast / Broadcast

- **One-to-many** communications



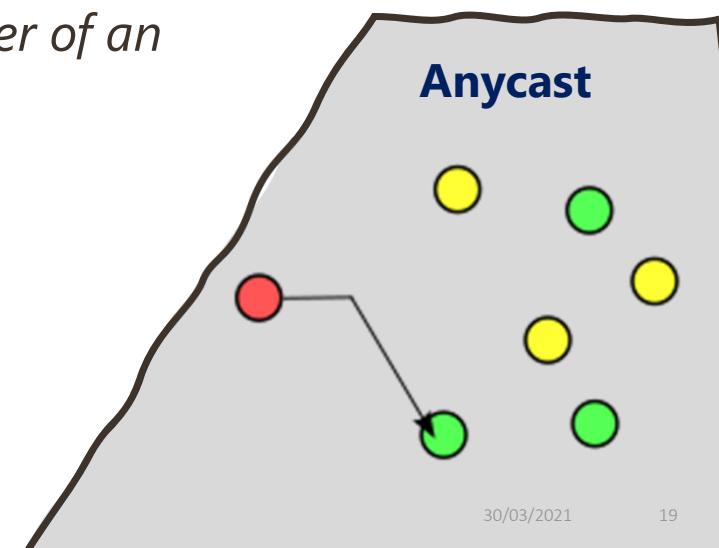
- IP Multicasting refers to implementation of multicast communication in the internet

- **Multicast is driven by receivers:**

- Receivers indicate interest in receiving data

Anycast [RFC: 1546]

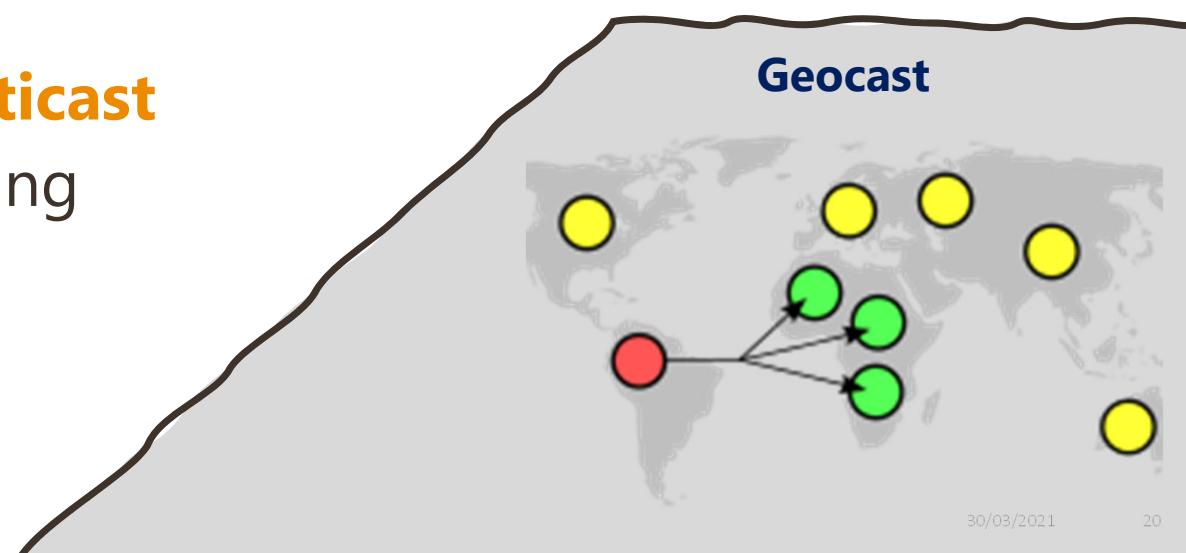
- Anycast addressing is a **one-to-one-of-many** association where datagrams are routed to any single member of a group of potential receivers **that are all identified by the same destination address**
- The routing algorithm selects the single receiver from the group **based on least-expensive routing metric**
 - *i.e. the packets are routed to the topologically-nearest member of an anycast group*
- **Widely used** for Content Delivery Network (**CDN**) products to bring their content closer to the end user



Geocast

- Geocast refers to the delivery of information to a group of **destinations in a network identified by their geographical locations**

- It is a **specialized form of multicast addressing** used by some routing protocols for mobile ad-hoc networks.





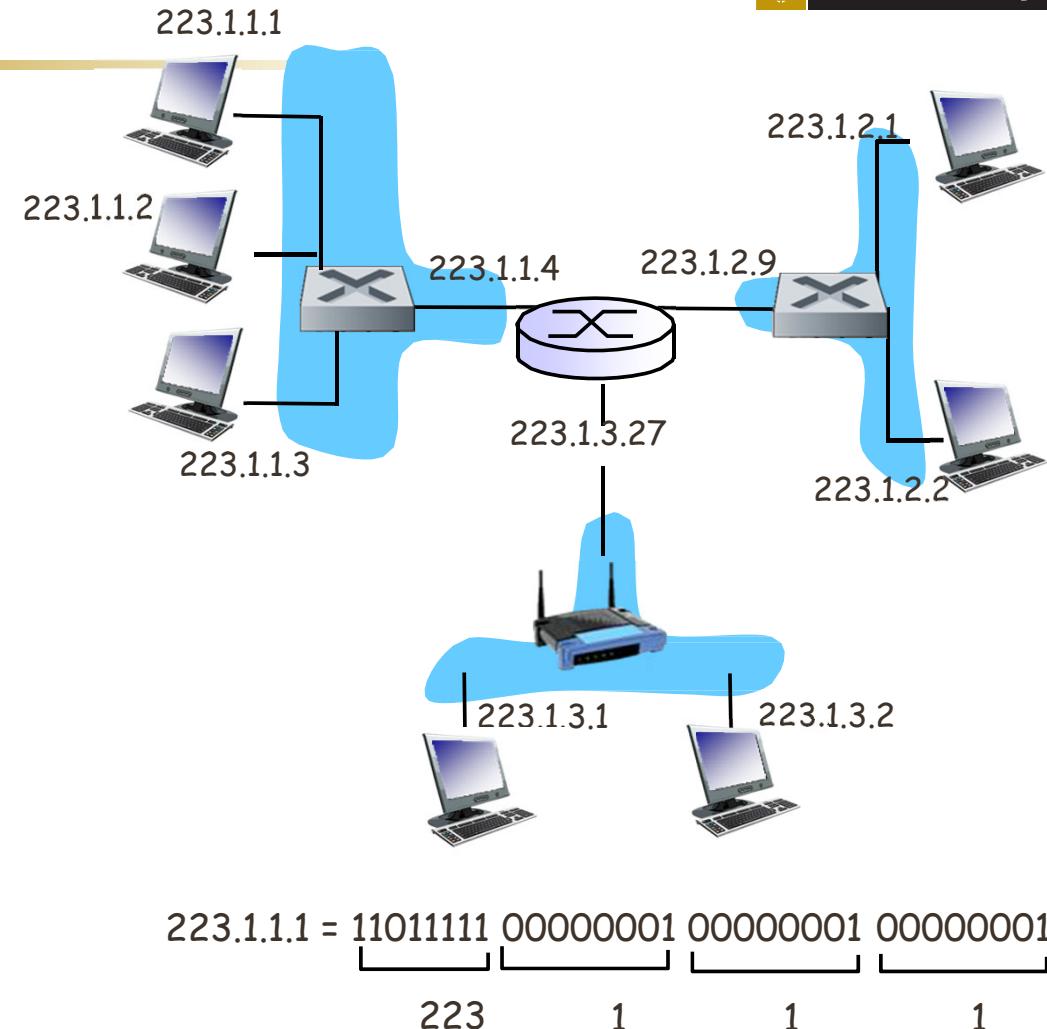
IPv4 Addressing

- Fundamentals
- Subnets
- Classful, Classless (CIDR) IP
- Static vs Dynamic IP
- Obtaining a Global IP Address
 - ISP Address Allocation
 - Hierarchical Addressing
 - ICANN
- Special IPv4 Addresses

IP Addressing

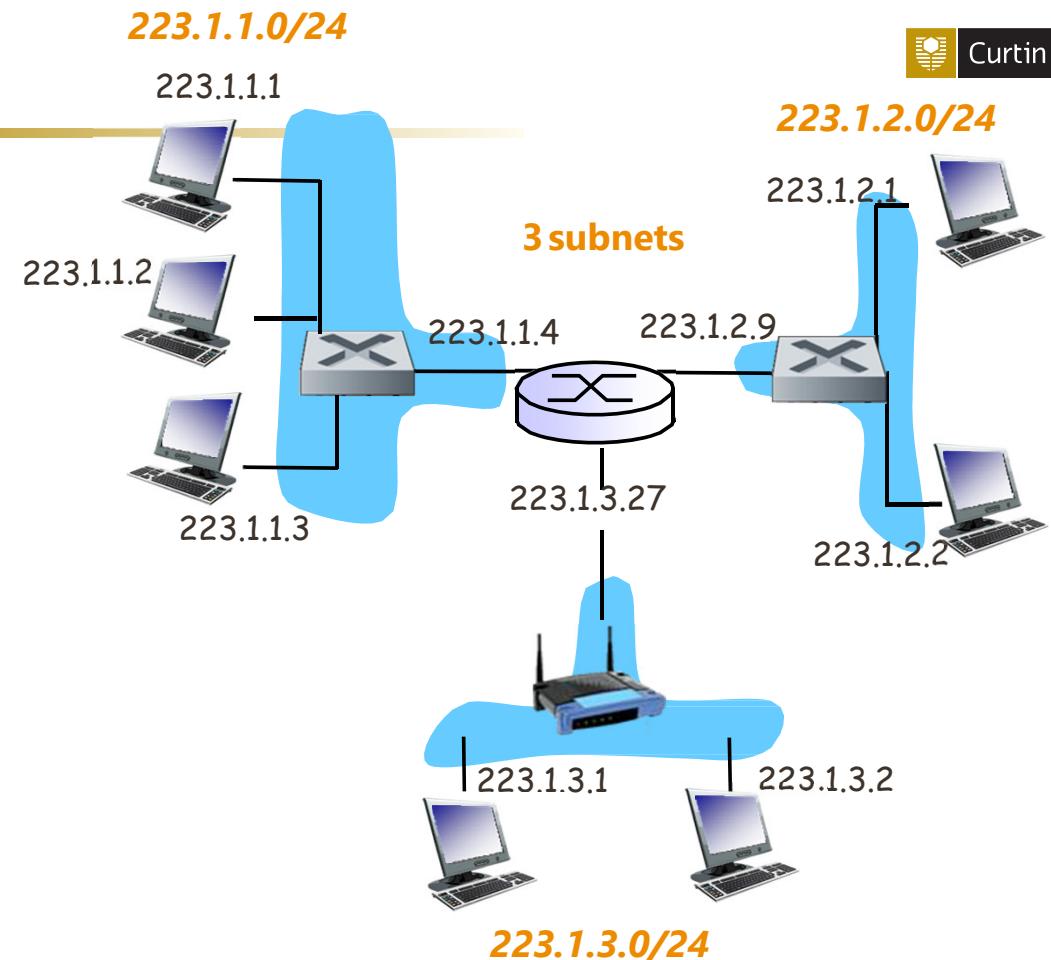
- **IP address:** 32-bit identifier for host and router interface
- **Interface:** connection between host/router and physical link
 - **Routers:** multiple interfaces
 - **Host:** one or two interfaces
(e.g., wired Ethernet, wireless 802.11)

"IP addresses for each interface"



Subnets

- A **subnetwork** or **subnet** is a logical subdivision of an IP network
- Isolated networks with **Subnet ID**
 - Can physically reach each other in same sub-network without intervening router
- **IP address:**
 - **Subnet ID** - high order bits
 - Host ID - low order bits



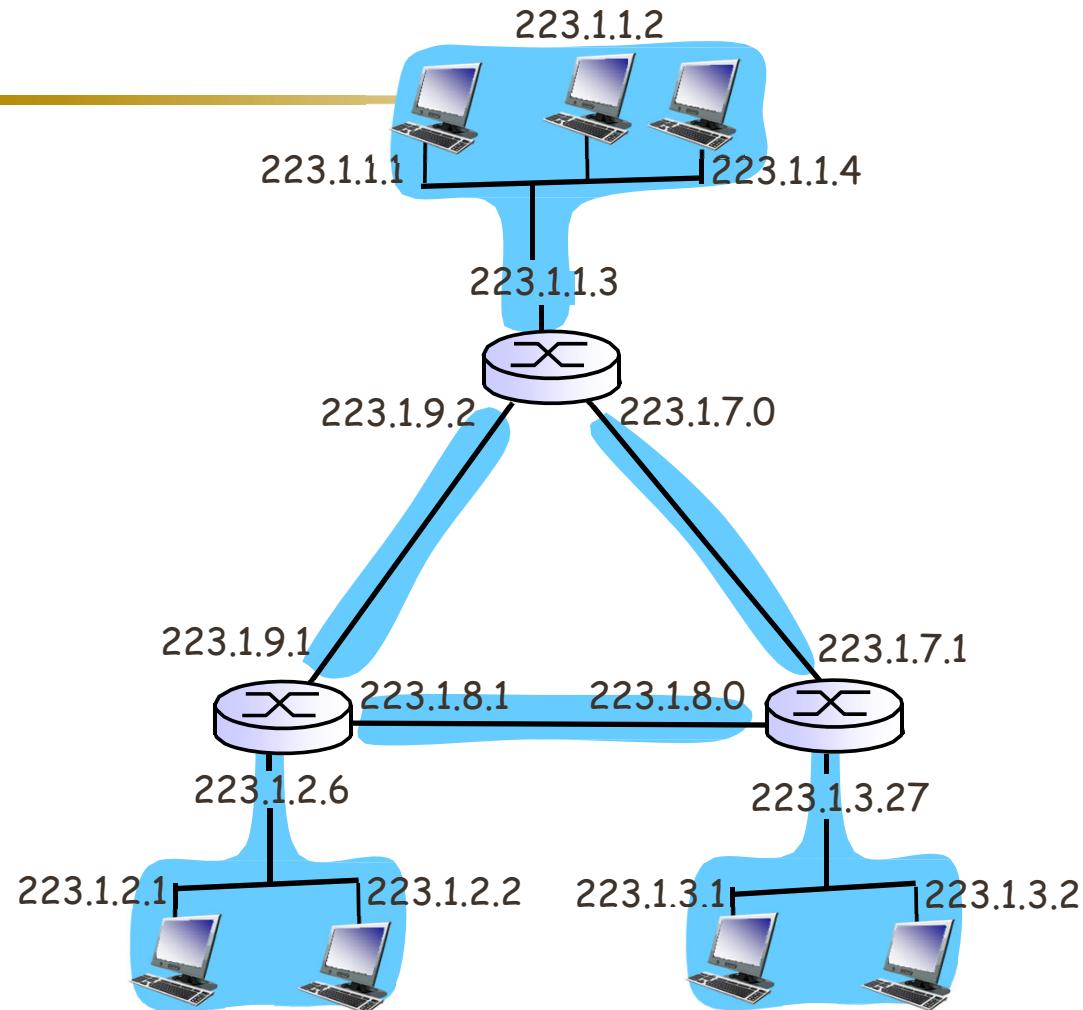
$223.1.1.1/24 = \underline{11011111} \underline{00000001} \underline{00000001} \underline{00000001}$

223 1 1 1

Subnet Mask: /24

Subnets

- How many hosts?
- Each subnet has **equal number of IPv4 addresses**



IP Address Types

- Classful

Class	Start bits	No. of netwk bits	No. of host bits	Range	
A	0	8	24	1.0.0.1 to 126.255.255.254	Supports 16 million hosts on each of 127 networks.
B	10	16	16	128.1.0.1 to 191.255.255.254	Supports 65,000 hosts on each of 16,000 networks.
C	110	24	8	192.0.1.1 to 223.255.254.254	Supports 254 hosts on each of 2 million networks.
D	1110	undef	undef	224.0.0.0 to 239.255.255.255	Reserved for multicast groups.
E	1111	undef	undef	240.0.0.0 to 254.255.255.254	Reserved for future use, or research and development purposes.

IP Address Types

▪ 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
- Possible to split host part further to create additional subnetting



200.23.16.0/23
Subnet mask is /23 or
255.255.254.0

Subnet Mask example

▪ IP address 192.168.40.28/21 - what's the network ID?

- Subnet Mask /21: 255.255.248.0
- 40 = 00101000 with mask 11111000
- Network ID is 192.168.40.0 /21

▪ 192.168.45.52 /21 - what's the network ID?

- Subnet Mask /21: 255.255.248.0
- 45 = 00101101 with mask 11111000
- Network ID is 192.168.40.0 /21

Variable Length Subnet Masking (**VLSM**)

- Use more than 1 subnet masks in the same network to create subnets with **unequal number of IPv4 addresses (hosts)**.
- **E.g. Divide 192.168.10.0 (a Class C network) into**
 - Subnet 1 : 126 IPv4 Addresses (7 bits required)
 - Subnet 2 : 62 IPv4 Addresses (6 bits required)
 - Subnet 3 : 30 IPv4 Addresses (5 bits required)
 - Subnet 4 : 30 IPv4 Addresses (5 bits required)
- Hence:
 - 192.168.10.0 /25 (255.255.255.128): 126 (128-2) usable IPv4 addresses
 - 192.168.10.128 /26 (255.255.255.192): 62 (64-2) usable IPv4 addresses
 - 192.168.10.192 /27 (255.255.255.224): 30 (32-2) usable IPv4 addresses
 - 192.168.10.224 / 27 (255.255.255.224): 30 (32-2) usable IPv4 addresses

Static vs Dynamic IP Addresses

▪ Statically Assigned

- Hard-coded by system admin in a file
 - **Windows:** control-panel->network->configuration->tcp/ip->properties
 - **UNIX:** /etc/rc.config

▪ Dynamically Assigned

- **DHCP:** Dynamic Host Configuration Protocol
- Dynamically get address from a server (or router)
 - Temporarily assigned, or "leased"
 - After a period of time, this lease "expires,"
 - Renews your old address or assigns new IP address

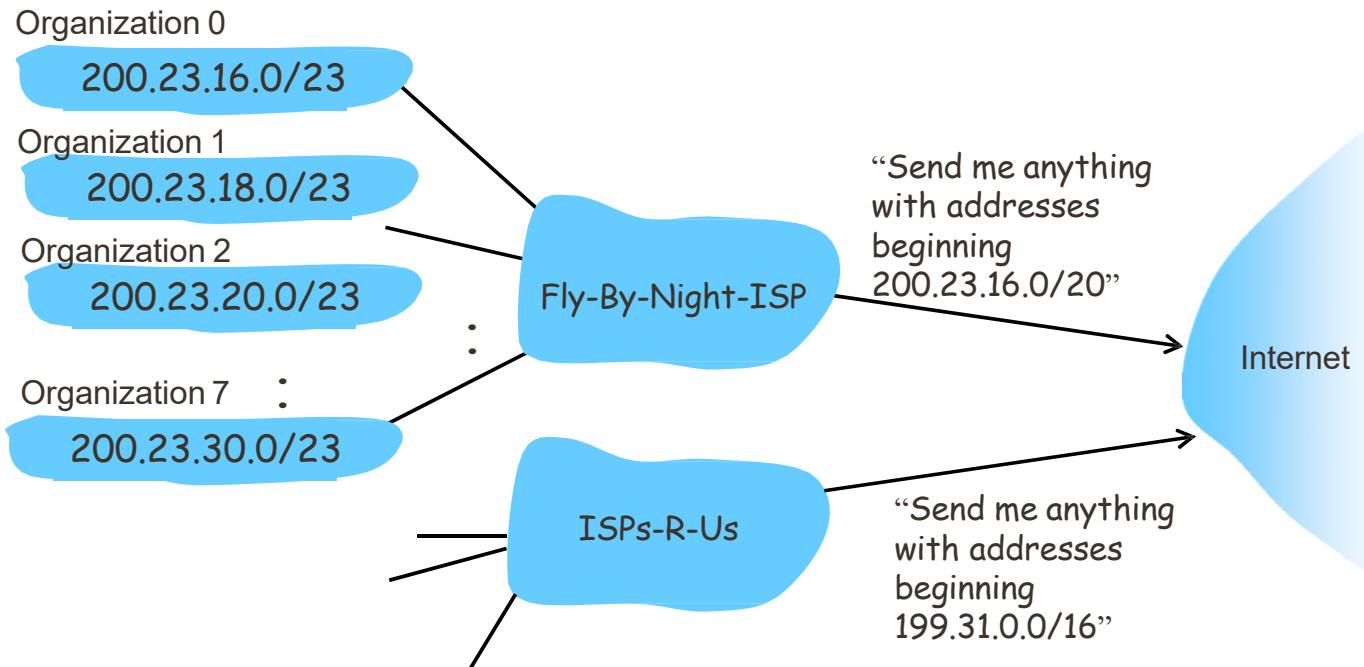
ISP, IP Address Allocation

- ISP (Internet Service Provider), IP address space division

ISP's block	<u>11001000 00010111 00010000 00000000</u>	200.23.16.0/20
Organization 0	<u>11001000 00010111 00010000 00000000</u>	200.23.16.0/23
Organization 1	<u>11001000 00010111 00010010 00000000</u>	200.23.18.0/23
Organization 2	<u>11001000 00010111 00010100 00000000</u>	200.23.20.0/23
...
Organization 7	<u>11001000 00010111 00011110 00000000</u>	200.23.30.0/23

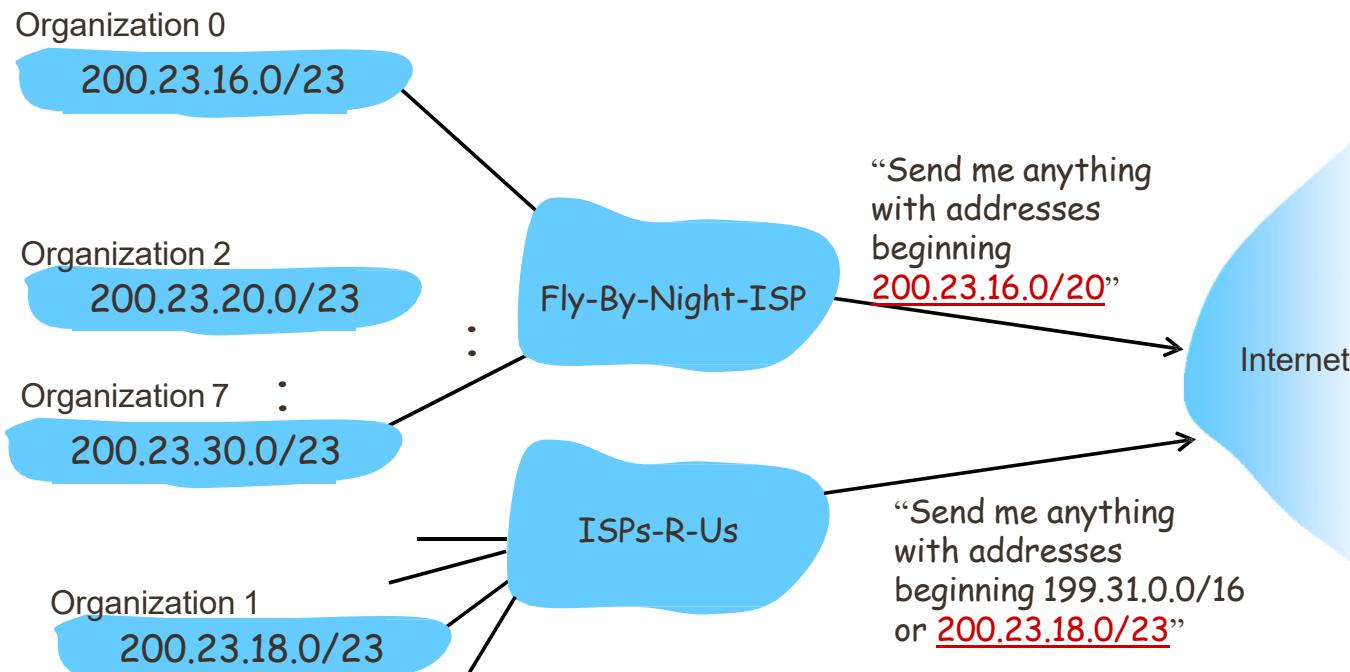
Hierarchical Addressing

- Allows efficient advertisement of routing information:



Hierarchical Addressing – cont.

- ISPs-R-Us has a more **specific route** to Organization 1



ICANN

- How does an ISP get block of addresses?
- **ICANN:** Internet Corporation for Assigned Names and Numbers
<http://www.icann.org/>
 - Allocates addresses
 - Manages DNS
 - Assigns domain names, resolves disputes

Why IP and MAC both?

■ MAC Address Only

- MAC/Physical address is a **globally unique ID** for your device
- Useful and efficient for **local communications (LAN)**
- Broadcast domain – the **internet is not possible**
- MAC address is burnt with hardware NIC card – loss of hardware, loss of all existing connections

tells **who** you are

00:1C:A2:01:A3:45
00:A0:C9:14:C8:29
00:1B:44:11:3A:B7
00:06:5B:BC:7A:C7

■ IP Address Only

- They can **group and organize** different networks (hierarchical)
- They are much like your mailing address
- They are **flexible and changeable** – making a device mobile

tells **where** you are



123.456.789.12

Special IP Addresses

Not assignable to a device or interface

▪ **0.0.0.0 (unspecified address)**

- The address of “this host.” (the primary IP address of the machine executing the instruction)
- In DHCP, when a unique address has not yet been determined, 0.0.0.0 used as the Source IP (i.e. DHCP Discover Packet)
- In the context of a Router: 0.0.0.0 represents Default route
- 0.0.0.0 is not assignable to an interface or used as a destination address

▪ **127.0.0.0/8 (or 127.0.0.0 – 127.255.255.255)**

- Send packets back to source. Used to test the protocol stack locally.
- Typically 127.0.0.1

Special IP Address examples

- **0.0.0.18**

The host with the host address 18 on “this local network”

- **255.255.255.255**

Broadcast on this local network

- **161.115.255.255**

Broadcast on target network 161.115.0.0/16

Link-Local IP Address

- **169.254.0.0/16 (or 169.254.0.0 - 169.254.255.255)**
“**valid only for** communications on a **local network segment**”
- **Self-generated automatically**
 - When a host cannot find a DHCP server
 - When communication problems occur between a host and a DHCP server
- **A link local address means**
 - The host cannot access the internet
 - Link local address is not routable
 - The host can communicate with other devices on the same LAN

Routers do not forward packets with link-local addresses

Private IP Addressing

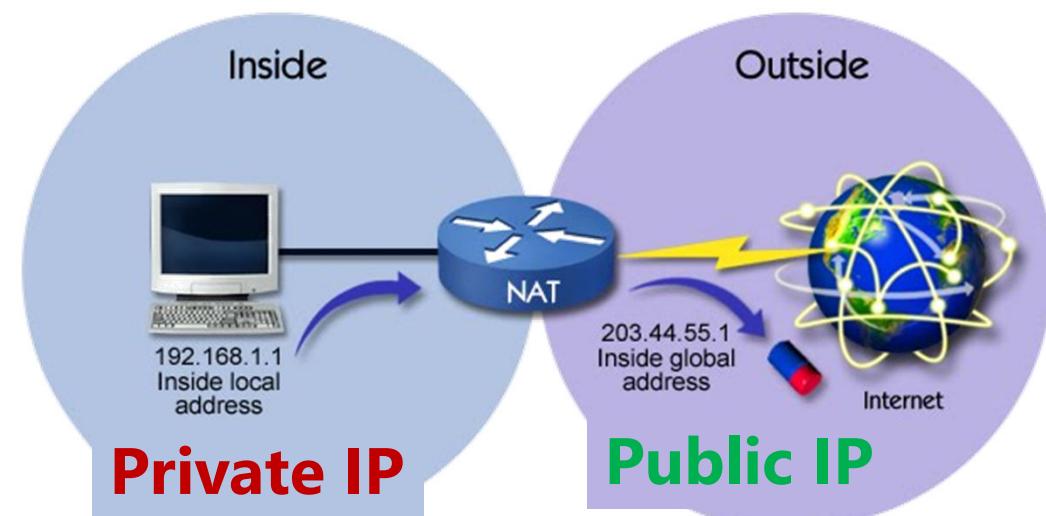
- **What are private IPs?**
 - **Non-routable** addresses **in internet**

- *Routers would not deliver packets with private IP addresses*

- **Free to use** without anyone's permission

▪ Why do we need them?

- There are only 4.3 billion IPv4 addresses
- They have extended IPv4's life



Private IP Addressing – IPv4

Class A	10.0.0.0 to 10.255.255.255	10.0.0.0/8
Class B	172. <u>16</u> .0.0 to 172. <u>31</u> .255.255	172. <u>16</u> .0.0/12
Class C	192.168.0.0 to 192.168.255.255	192.168.0.0/16

Only a portion

Devices with **private IP** address **cannot connect directly to the Internet**

Instead, **access** to the **Internet** must be brokered by a **router** that **supports Network Address Translation (NAT)**

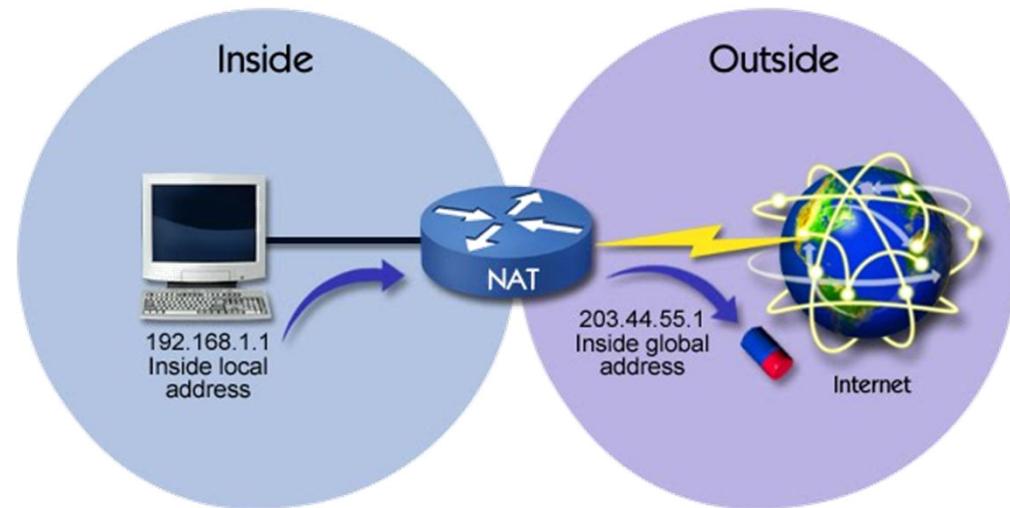


Network Address Translation

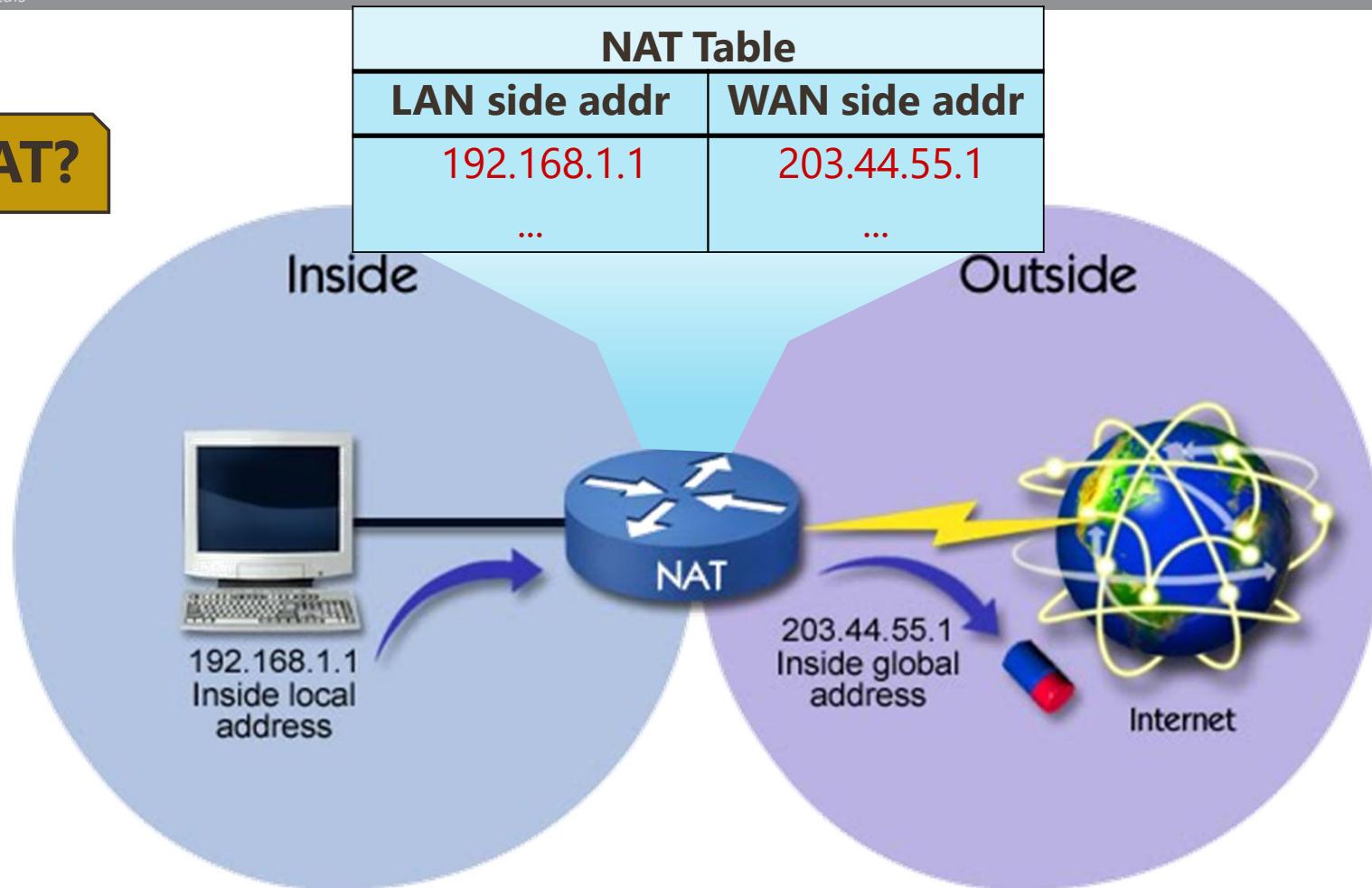
- SNAT
- DNAT
- PAT
- Port Forwarding

Why Network Address Translation?

- **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
- **Secure:** devices inside local network are not explicitly addressable, visible by outside world



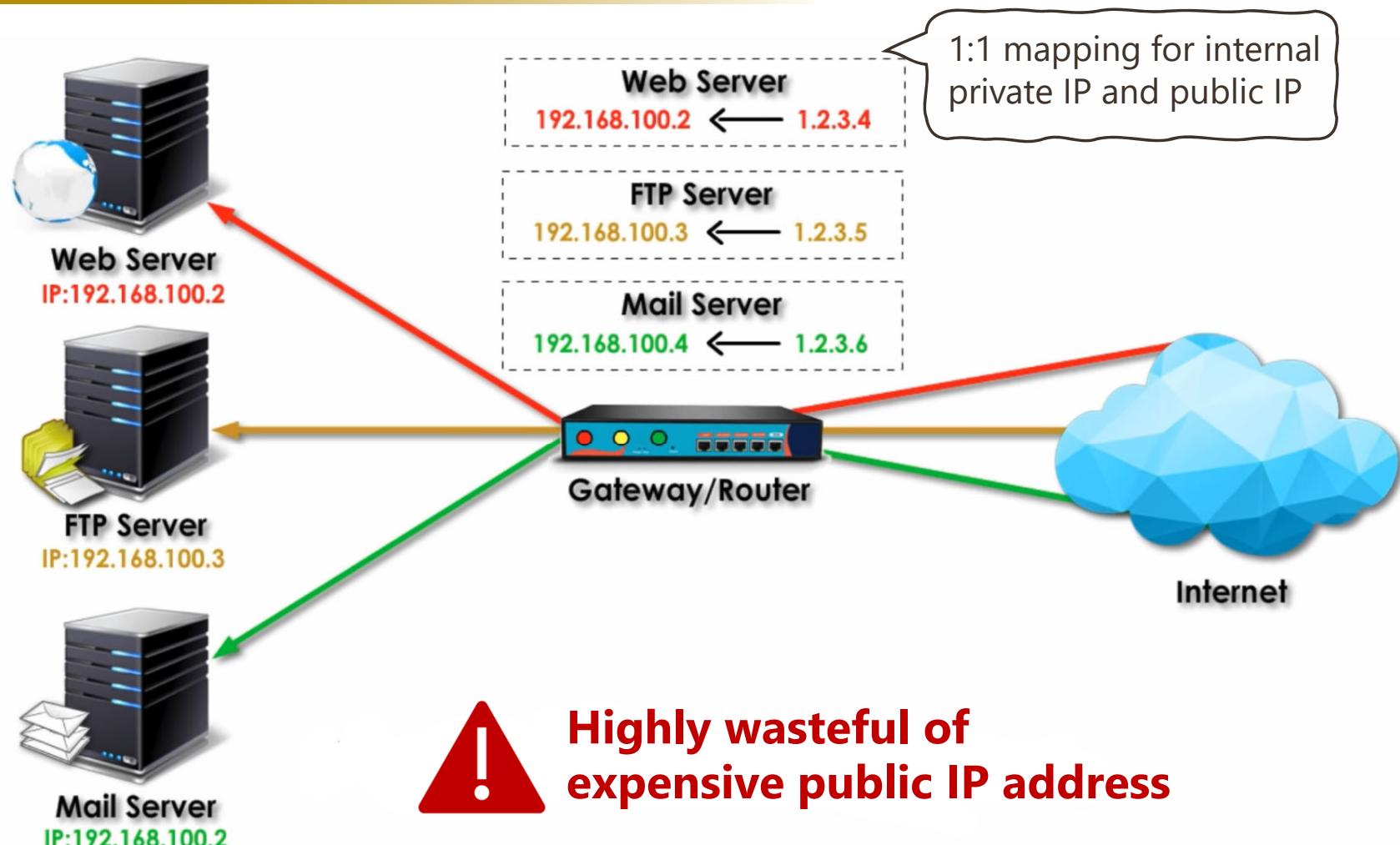
What is NAT?



A router **translate an internal host's private IP address into its public IP address** for outgoing/incoming traffic and vice-versa

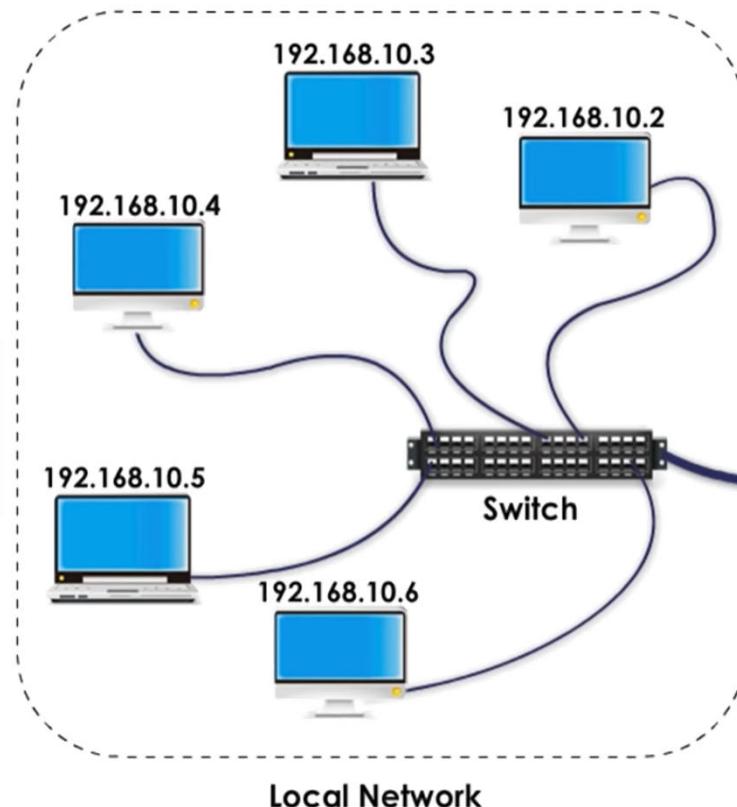
SNAT: Static NAT

deals more with incoming traffic

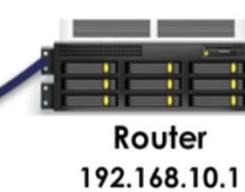


DNAT: Dynamic NAT

deals more with outgoing traffic



Replace internal IP with an IP address of the public IP address pool



56.4.2.1
56.4.2.2
56.4.2.3
56.4.2.4
.....

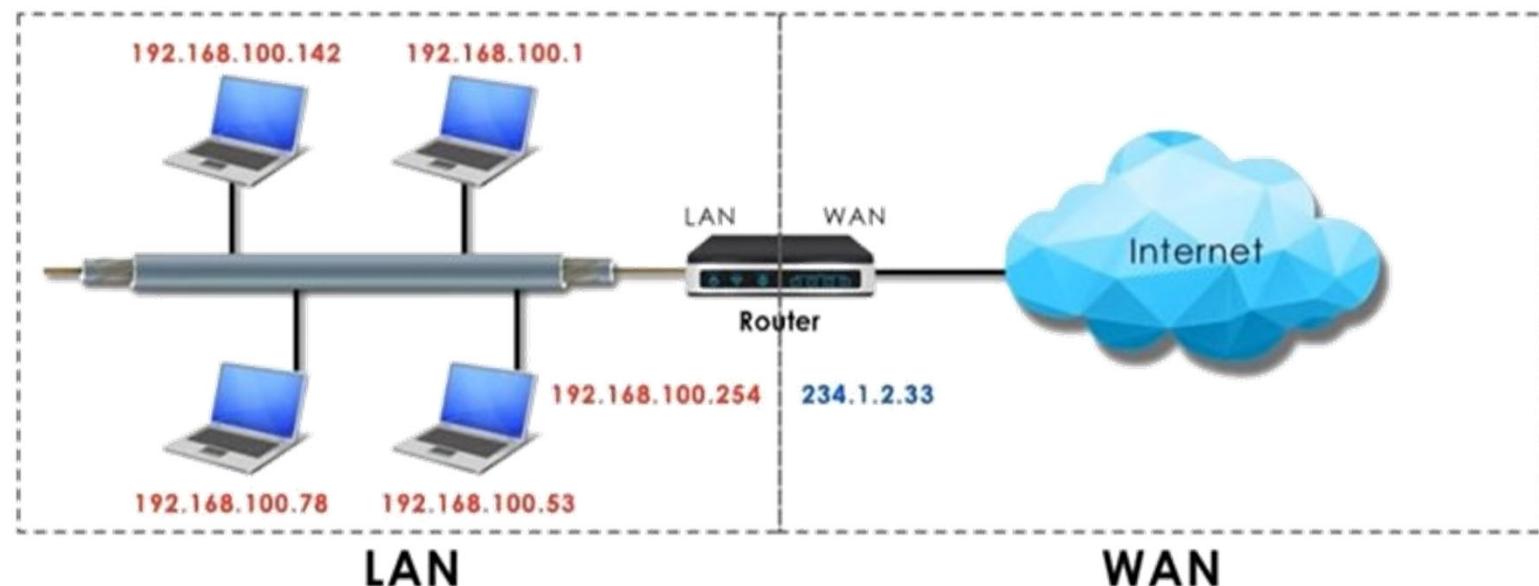
Pool of public IP address

PAT: Port Address Translation

mapping

$192.168.100.1:80 = 234.1.2.33:8000$

$192.168.100.78:80 = 234.1.2.33:8001$

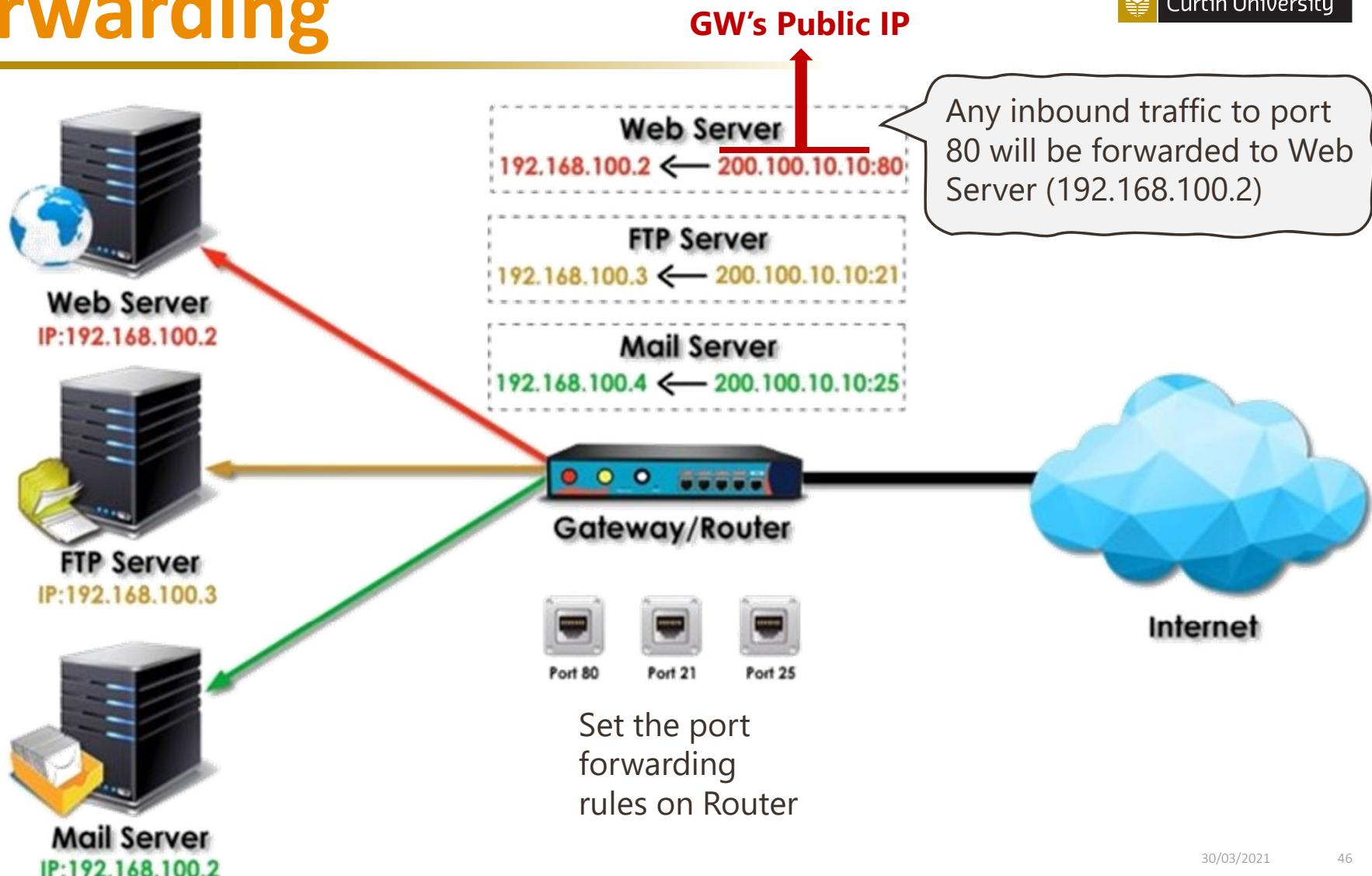


$192.168.100.53:80 = 234.1.2.33:8002$

$192.168.100.142:80 = 234.1.2.33:8003$

Port Forwarding

Port Forwarding
deals more with
incoming traffic





IPv6

- Fundamentals
- Address Space
- IPv6 Datagram
- IPv6 Address Simplification
- Multicast addressing
 - Solicited node address
- Unicast addressing
 - Global Unicast Address

Introduction to IPv6

Initial Motivation:

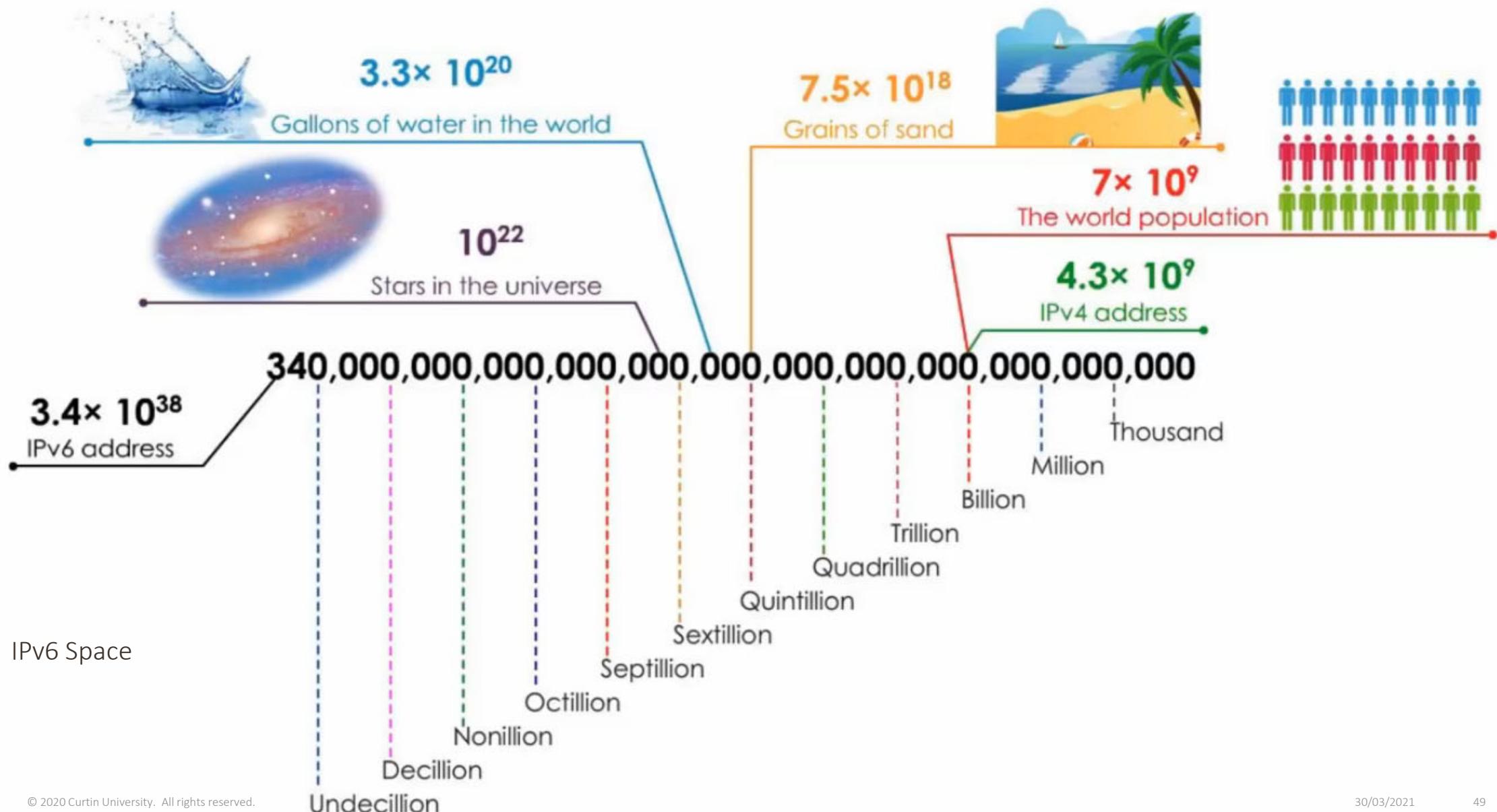
32-bit address space soon to be completely allocated

■ Why IPv6?

- ✓ Supports $2^{128} \sim= 3.4 \times 10^{38}$ addresses

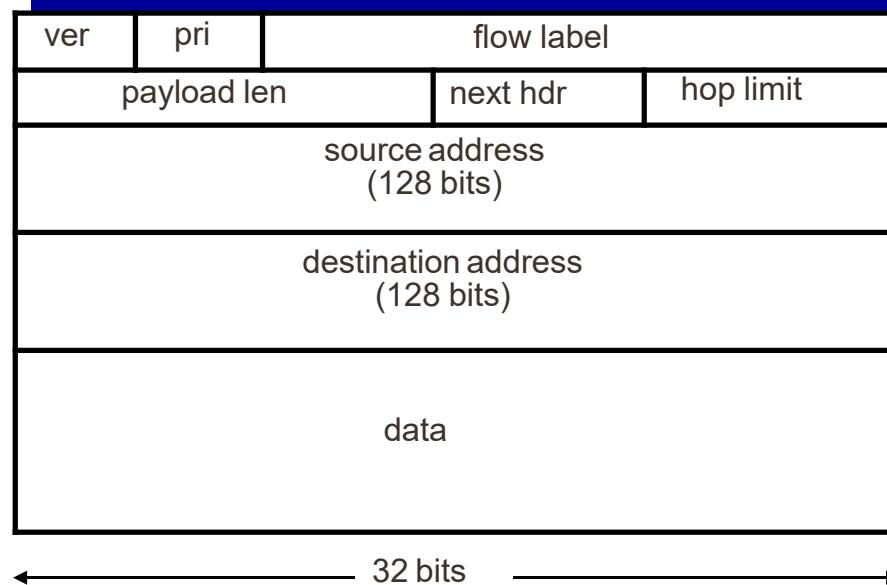
i.e. 2001:000:000:000:0080:ACDE:02CE:1234

- ✓ header format helps speed processing/forwarding
- ✓ header changes to facilitate **QoS**



IPv6 Datagram Format

- **Priority:** identify priority among datagrams in flow
- **Flow Label:** identify datagrams in same “flow.”
(concept of “flow” not well defined)
- **Next Header:** identify upper layer protocol for data



IPv6 Address Simplification

2001:0000:0000:0000:0080:ACDE:02CE:1234

- Leading 0's can be dropped from any group

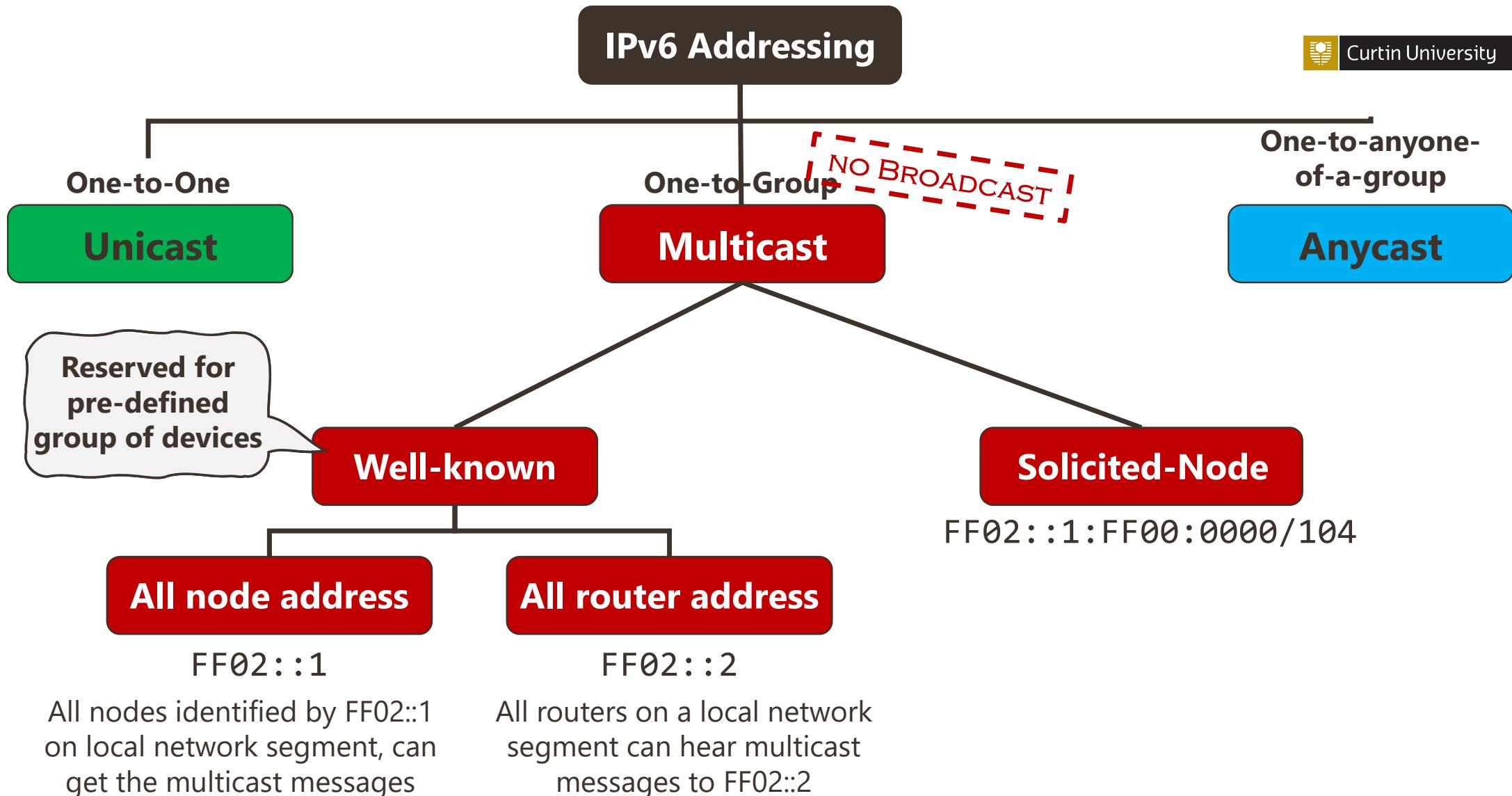
2001:0:0:0:080:ACDE:02CE:1234 → 2001:0:0:0:80:ACDE:2CE:1234

- Using a pair of colons (::) to represent a string of consecutive groups of 0s

2001:0:0:0:80:ACDE:2CE:1234 → 2001~~0:0:0~~::80:ACDE:02CE:1234

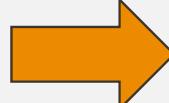
- **Can not use more than one set of colon pairs**

2001:0:0:0:CF:0:0:1234 → 2001~~0:0:0:CF:0:0:1234~~ :: CF:0:0:1234 OR
2001:0:0:0:CF~~0:0:1234~~ :: 1234



Solicited Node Address

- Created automatically by prepending multicast prefix **FF02::1:FF00:0000/104** to last 24 bits of unicast or anycast address

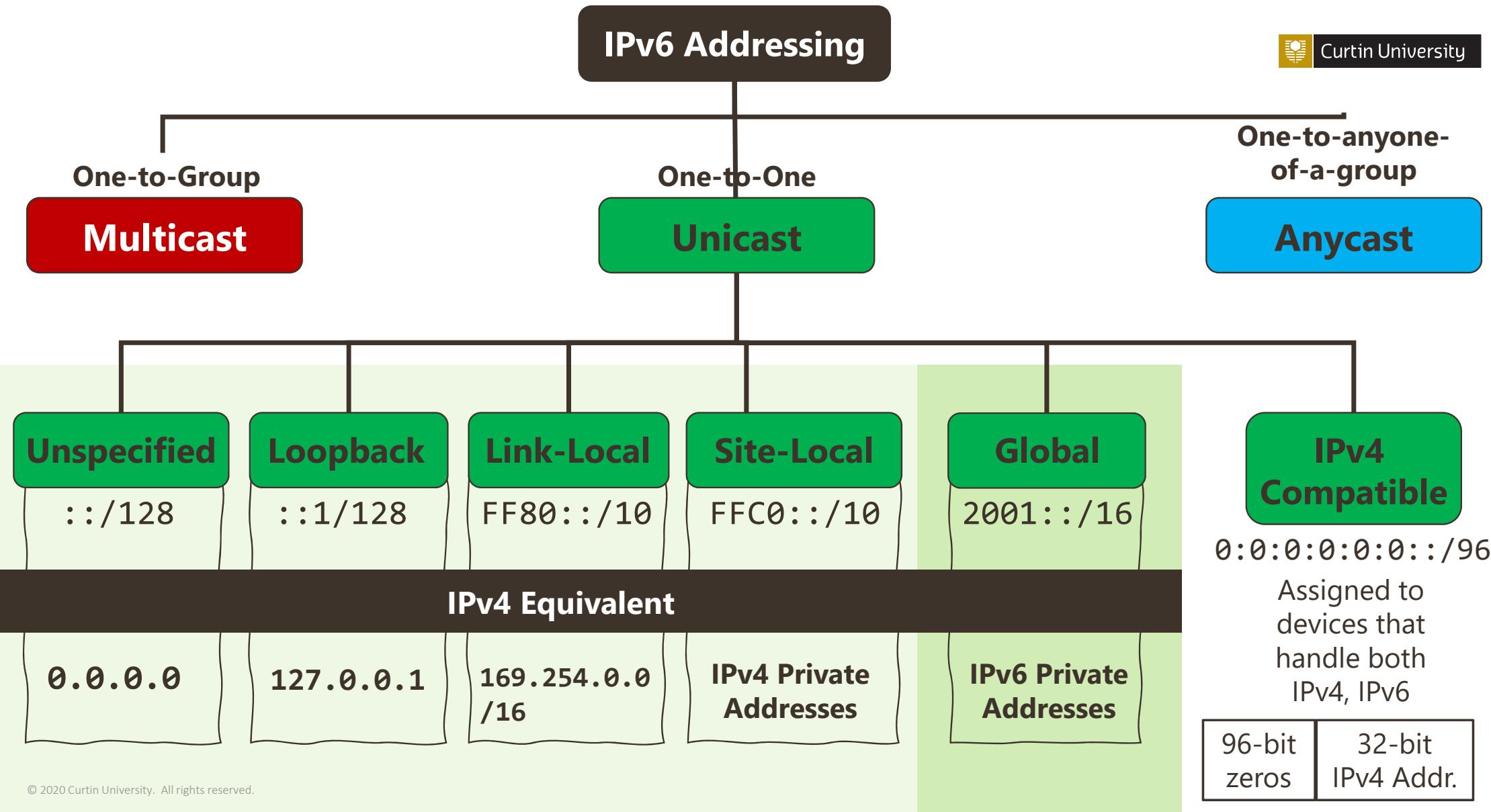
i.e. UnicastIP=2001::01:800:20**0E:8C6C**  FF02::1:**FF0E:8C6C**

- A host is required to join a **solicited-node multicast** group for each of its configured unicast or anycast addresses

Why Solicited Node Address?

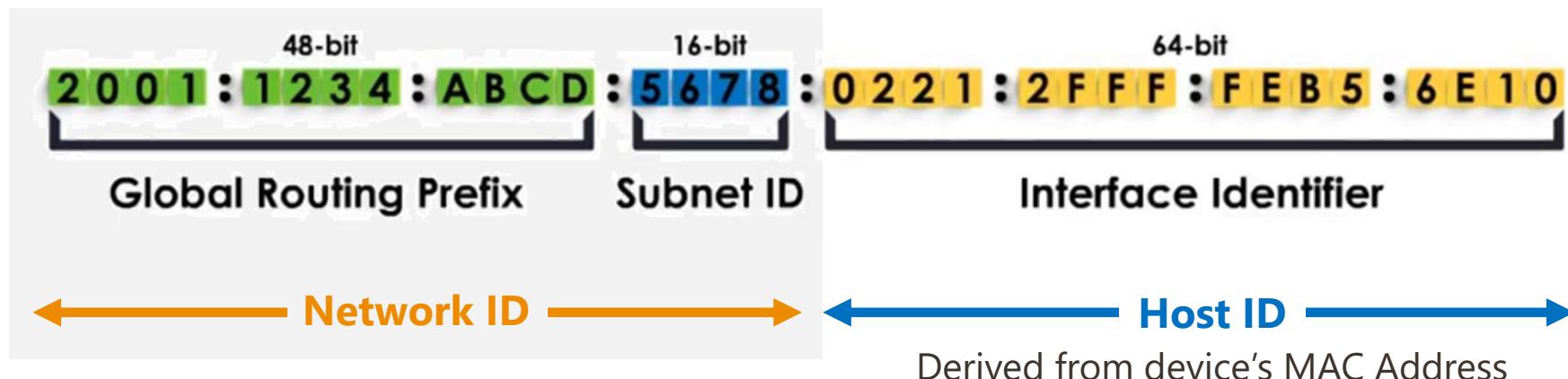
- Useful in link layer address resolution with **Neighbor Discovery Protocol** (NDP) on the link without disturbing all nodes on the local network

IP to MAC Resolution
IPv4 uses ARP,
IPv6 uses **NDP**



Global Unicast

- Similar to IPv4 Public IP Addresses
- Global unicast IPv6 prefixes that are currently allocated by IANA are 2000::/3
 - ✓ They all start with 001 <- binary
 - ✓ At this point, unique global unicast IPv6 address starts with 2001
 - ✓ E.g.



Anyone downstairs will get
2001/16 as routing prefix.

Anyone downstairs will get
2001:1234/32 as routing
prefix.

Anyone downstairs will get
2001:1234:abcd/48 as
routing prefix.

Anyone downstairs will get
2001:1234:abcd:5678/64
as routing prefix.



2001/16



2001:1234/32



2001:1234:abcd/48



2001:1234:abcd:5678/64



Network ID Generation

IANA

01

Your regional ISF

02

Your ISP

03

Your company's
default gateway

04

48-bit

16-bit

2001:1234:ABCD:5678:

Global Routing Prefix

Subnet ID



IPv4 to IPv6 Address Transition

- Dual Stack
- Tunnelling
- NAT64 Translation

IPv4 to IPv6 Transition

- IPv4 and IPv6 will coexist for a long time
- *Techniques*

1. Dual Stack

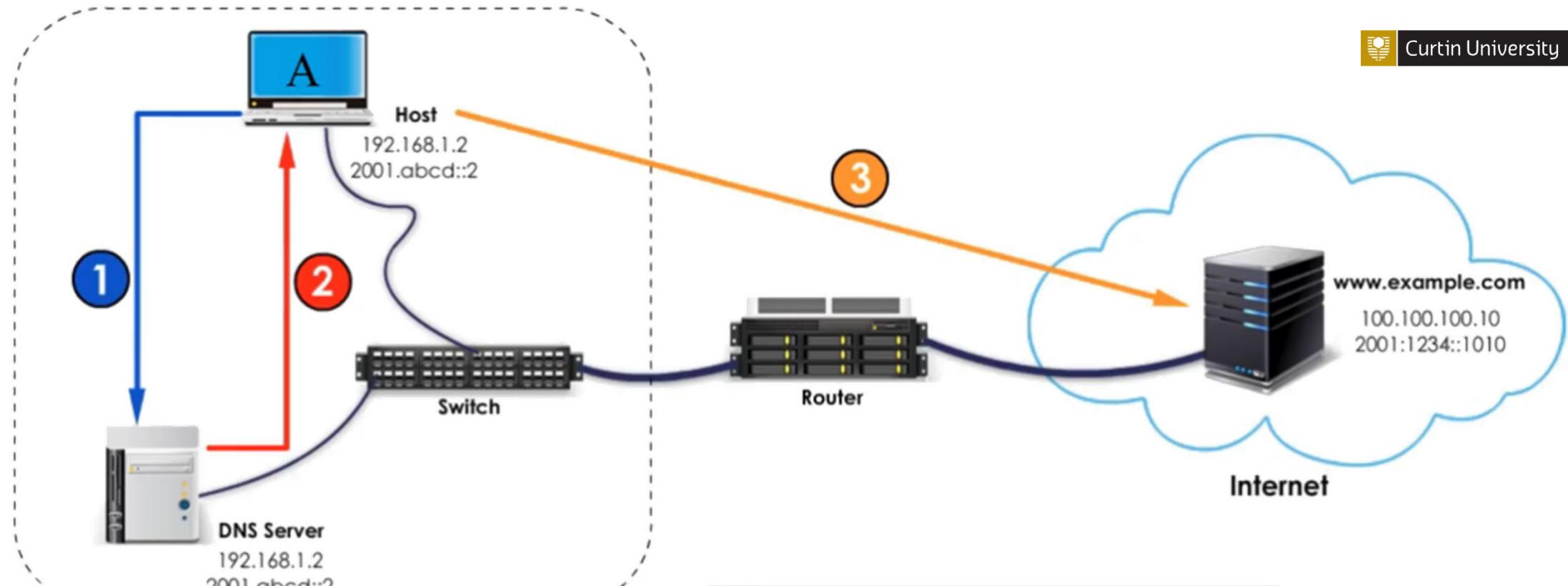
A dual stacked device (PC, Server, Router) supports both IPv4 and IPv6

2. Tunneling

1. Manual Tunneling
2. 6to4
3. ISATAP

3. Translation (NAT64)

Similar to address translation

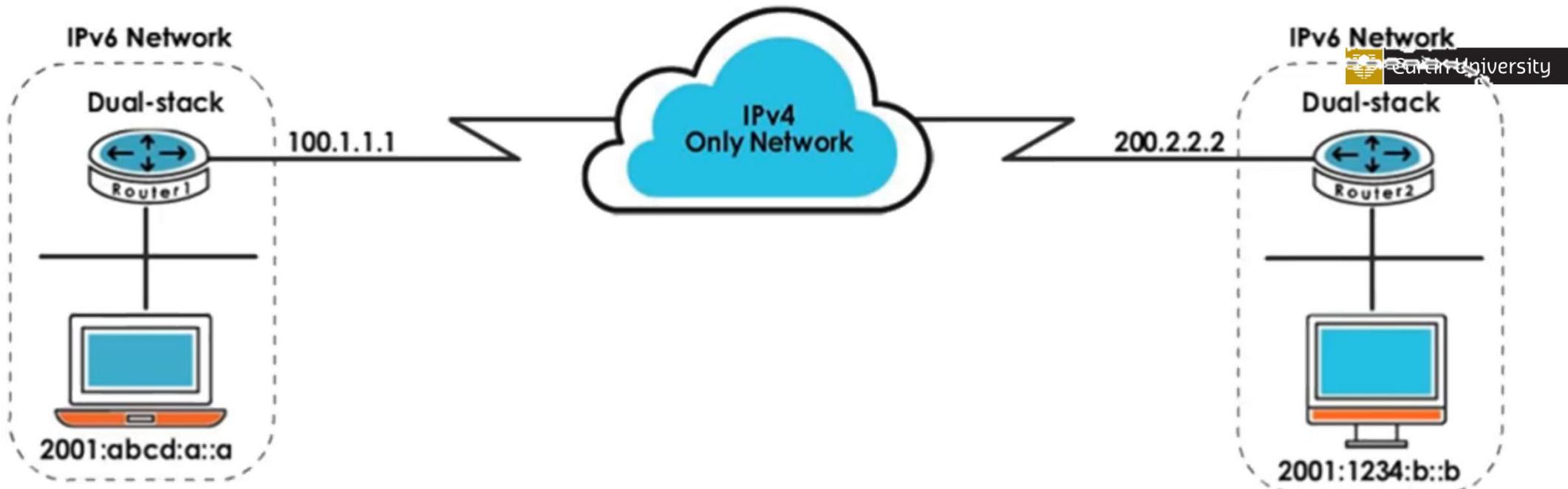


IPv4 and IPv6 Network

Dual Stack

- 1 I need www.example.com IP address
- 2 Type AAAA record: 2001:1234::1010
Type A record: 100.100.100.10
- 3 IPv6 Session with 2001:1234::1010

If failed, IPv4 Session will be established



IPv6 Packet

IPv6 SRC:	IPv6 DST:
2001:abcd:a::a	2001:1234:b::b

Manual Tunnel



IPv6 Packet

IPv6 SRC:	IPv6 DST:
2001:abcd:a::a	2001:1234:b::b

Tunneling

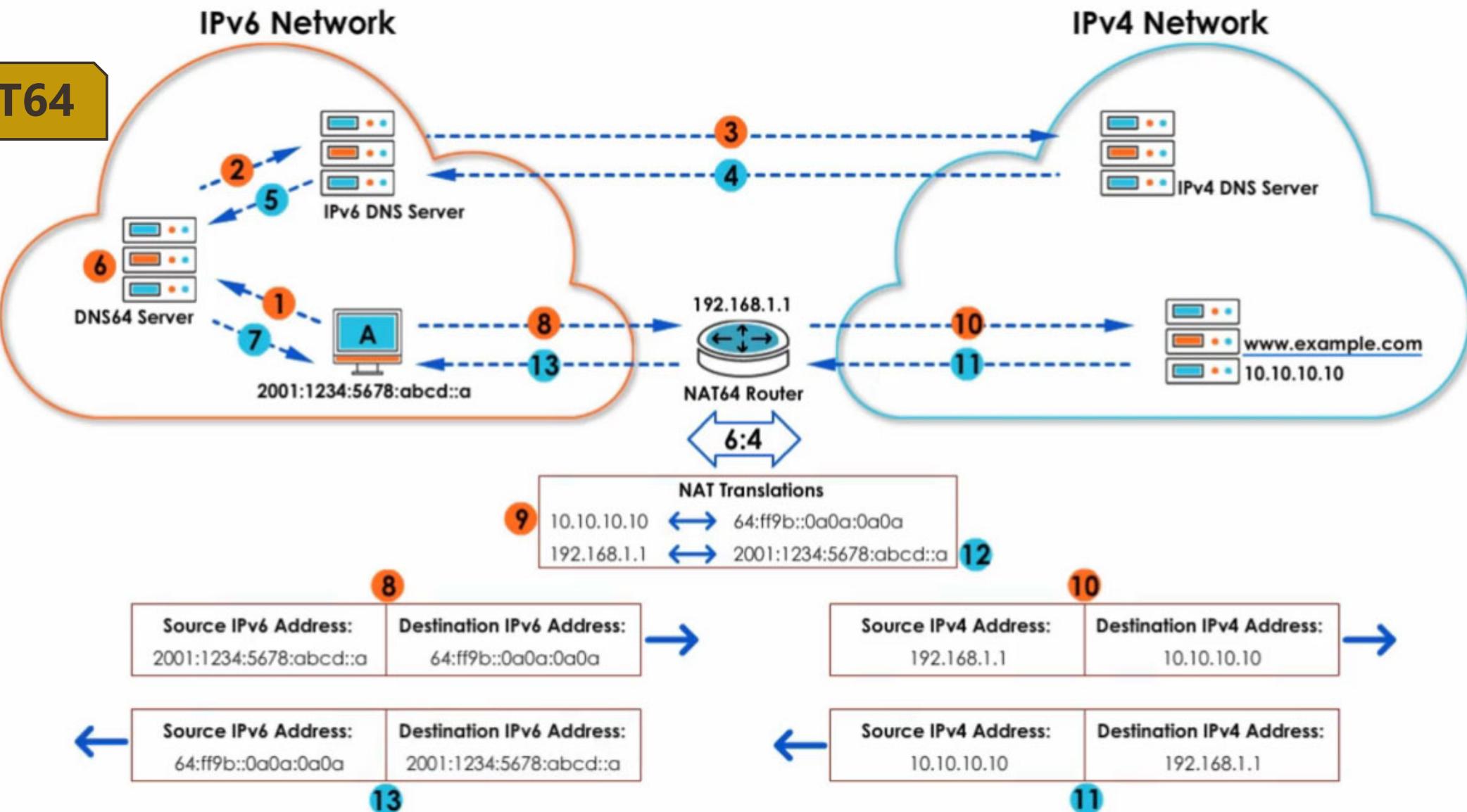
IPv4 SRC: 100.1.1.1
IPv4 DST: 200.2.2.2

41

IPv6 SRC: 2001:abcd:a::a
IPv6 DST: 2001:1234:b::b

Data

Payload – IPv6 Packet



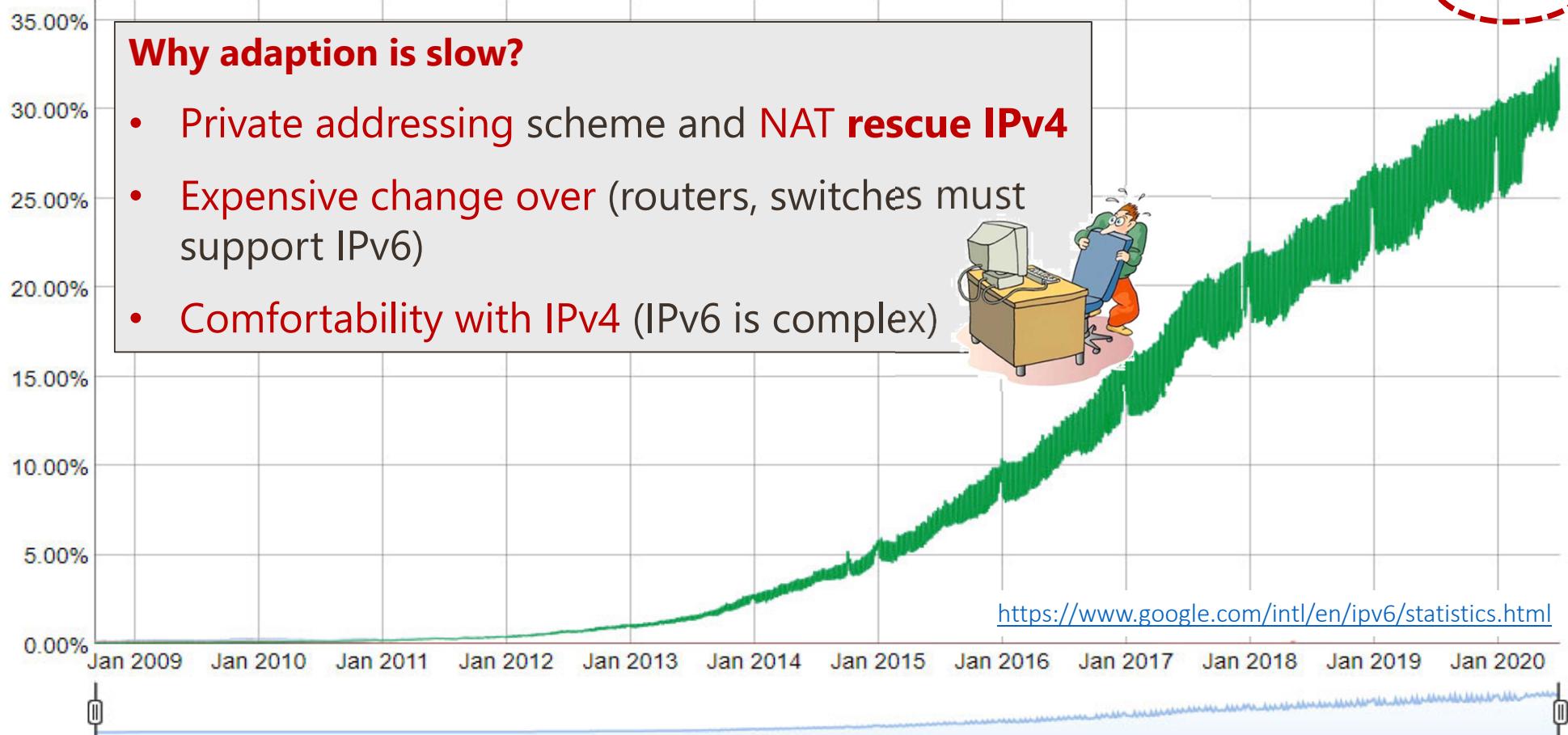
IPv6 Adoption

Per-Country IPv6 adoption

IPv6 Adoption

We are continuously measuring the availability of IPv6 connectivity among Google users. The graph shows the percentage of users that access Google over IPv6.

Native: 31.49% 6to4/Teredo: 0.01% Total IPv6: 31.50% | Jul 3, 2020



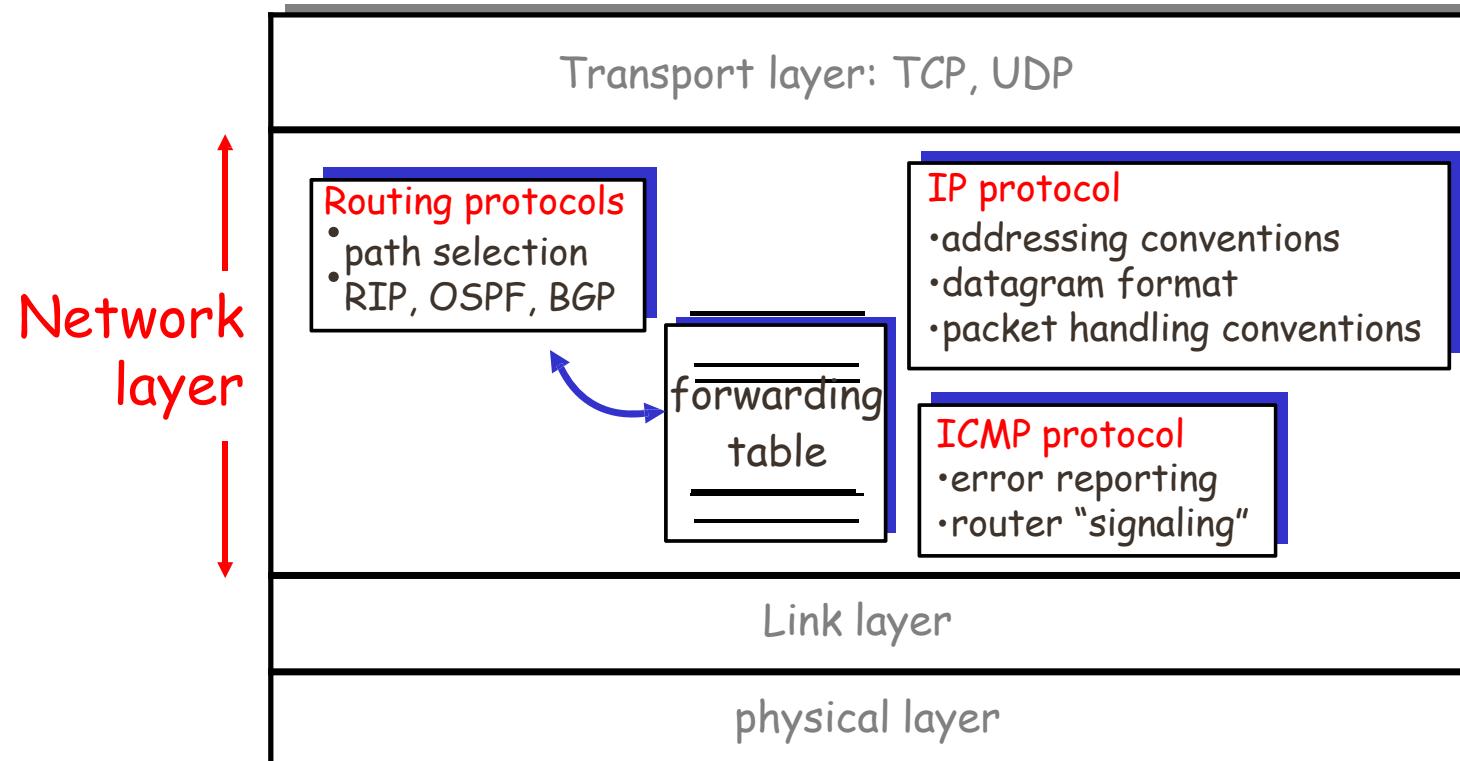
**IPv6
Last
Words**



Network Layer Protocols

- ICMP
- Traceroute Utility
- ARP

Network Layer Protocols



ICMP: Internet Control Message Protocol

- Used by hosts & routers to communicate network-level information

- **error reporting:** unreachable host, network, port, protocol
- echo request/reply (used by **ping**)

▪ ICMP message

- type, code plus first 8 bytes of IP datagram causing error

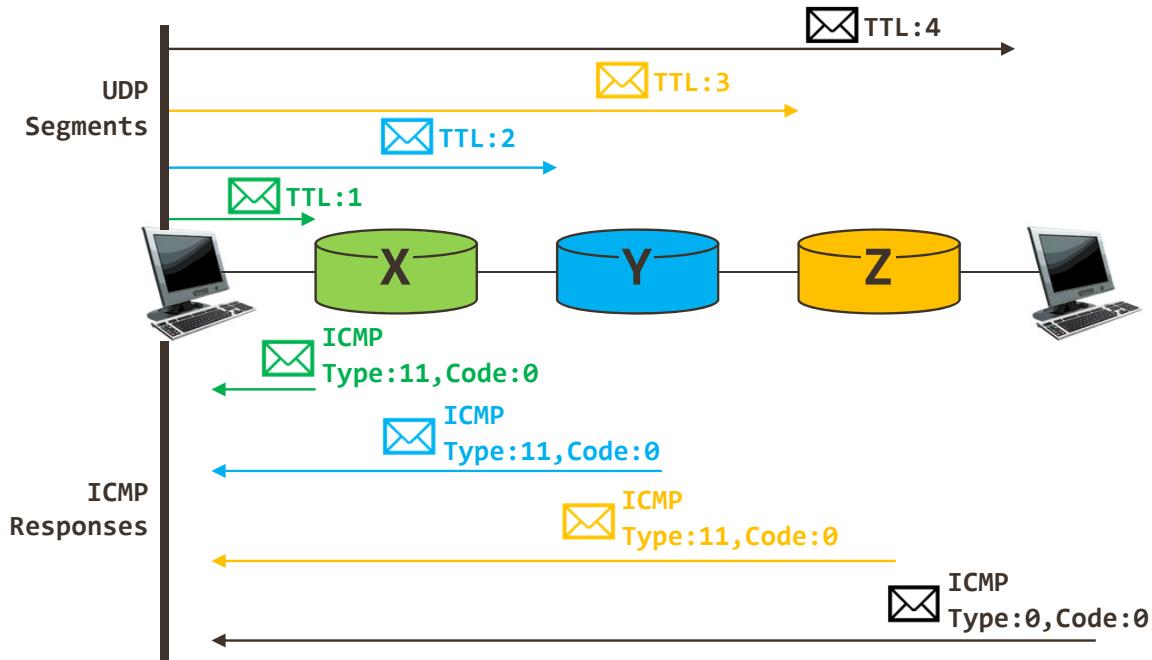
▪ ICMPv6: (IPv6)

- Additional message types, e.g. "Packet Too Big"
- Multicast group management functions

Type	Code	<u>description</u>
0	0	echo reply (ping)
3	0	dest. network unreachable
3	1	dest host unreachable
3	2	dest protocol unreachable
3	3	dest port unreachable
3	6	dest network unknown
3	7	dest host unknown
4	0	source quench (congestion control - not used)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header

ICMP & traceroute utility

- Source sends series of **UDP segments** to destination
- When **TTL** expires on a router
 - Router discards datagram
 - Sends to source an ICMP message (type 11, code 0)
 - Message includes name of router & IP address
- When **ICMP** message arrives, **source** calculates **RTT**
- **Stopping criterion**
 - UDP segment arrives at destination host
 - Destination returns ICMP "host unreachable" (type 3, code 3)



ARP: Address Resolution Protocol

- Maps an **IP Address** to a physical **MAC Address on a LAN**
- A computer uses **ARP program** to find another computer's MAC address based on its IP address
- Why MAC address?
 - **Communication within LAN:** MAC address is used
 - **Communication between LANs:** IP address is used
- **ARP cache:**
 - A table of IP address with their corresponding MAC addresses

LAYER 2
PROTOCOL



ARP Cache

C:\>arp -a		
Internet Address	Physical Address	Type
10.172.112.1	cc-4e-24-1a-d7-00	dynamic
10.172.117.6	0c-84-dc-8e-5a-6b	dynamic
10.172.117.189	4c-0b-be-2c-f7-02	dynamic
10.172.119.255	ff-ff-ff-ff-ff-ff	static
224.0.0.22	01-00-5e-00-00-16	static
224.0.0.251	01-00-5e-00-00-fb	static
224.0.0.252	01-00-5e-00-00-fc	static
239.255.255.250	01-00-5e-7f-ff-fa	static
255.255.255.255	ff-ff-ff-ff-ff-ff	static

Static ARP Entries

- Manually entered

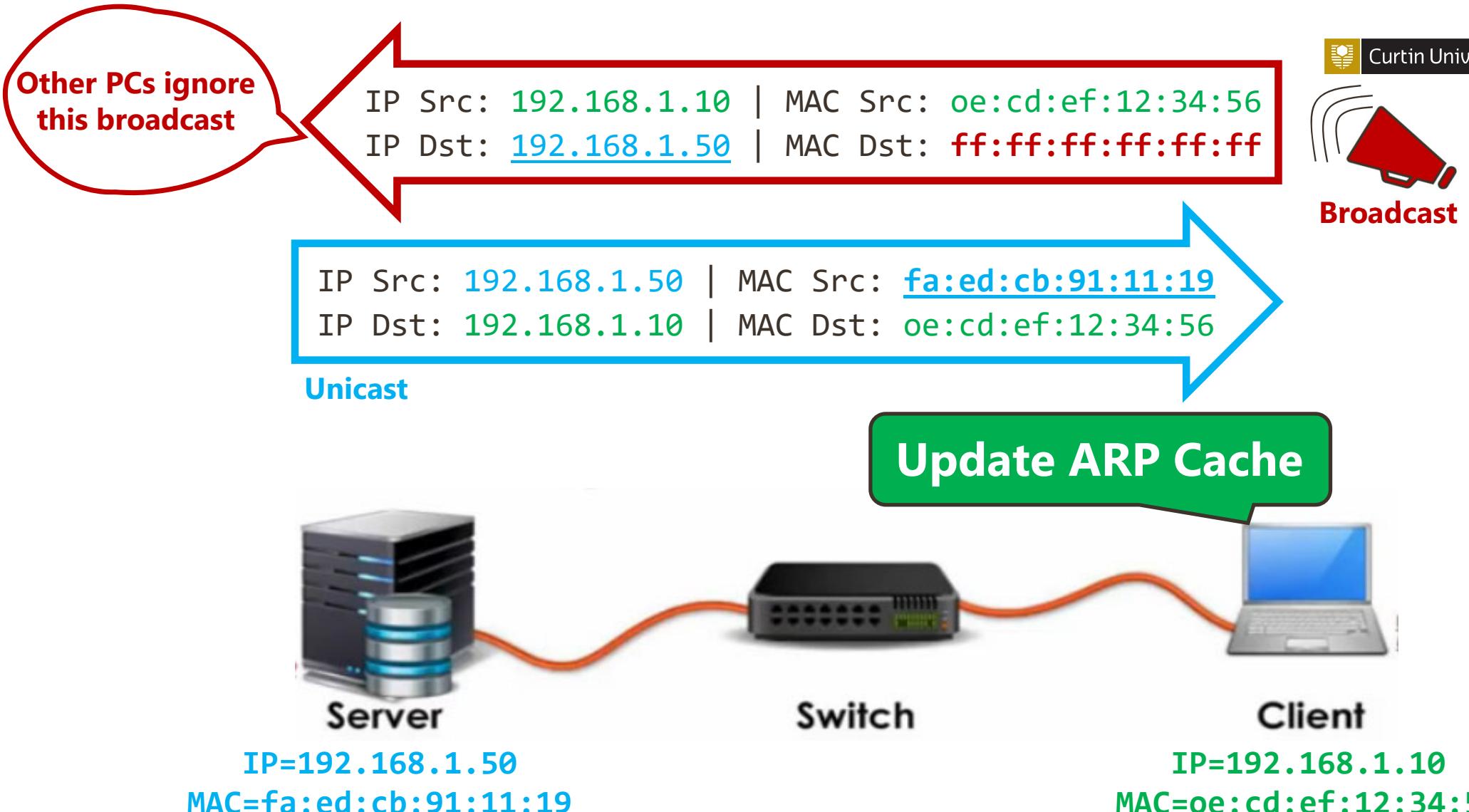
Dynamic ARP Entries

- Via ARP program



ARP Cache Timeout

Timeout for ARP entries





■ Network Layer

- Key Functions
- Services
- Virtual Circuit Networks
- Datagram Networks
 - IP datagram
 - Fragmentation

■ Address Types

- Unicast
- Multicast
- Broadcast
- Anycast
- Geocast

■ IPv4 Addressing

- Classful / Classless Addressing
- VLSM
- Static vs Dynamic IP Addresses
- Obtaining a Global IP Address
 - ISP Address Allocation
 - Hierarchical Addressing
- Special IPv4 Addresses
 - 0.0.0.0, loopback, broadcast, target broadcast
 - Link-local IP address
 - Private IP address

■ Address Translation

- SNAT
- DNAT
- PAT
- Port Forwarding

■ IPv6

- Address Space
- IPv6 Datagram
- IPv6 Address Simplification
- Multicast addressing
 - Solicited node address
- Unicast addressing
 - Global Unicast Address
- IPv6-IPv4 Transition

■ Network Layer Protocols

- ICMP
- Traceroute Utility
- ARP



Curtin University

THANK YOU

Make tomorrow better.