COMPUTER NETWORKS AND INTERNET PROTOCOLS

IP Routing – Introduction [Routing Table]

SOUMYA K GHOSH
COMPUTER SCIENCE AND ENGINEERING
IIT KHARAGPUR

SANDIP CHAKRABORTY
COMPUTER SCIENCE AND ENGINEERING
IIT KHARAGPUR

Discussion Points

- IP addresses and Address allocation
- Packet routing/ forwarding
  - Routing/ Forwarding tables
  - Longest-prefix match forwarding

Ref: *Computer Networking: A Top-Down Approach; Computer Networks; Lecture materials of Jennifer Rexford, Princeton, USA; TCP/IP Tutorials and Technical Overview, IBM Redbooks*

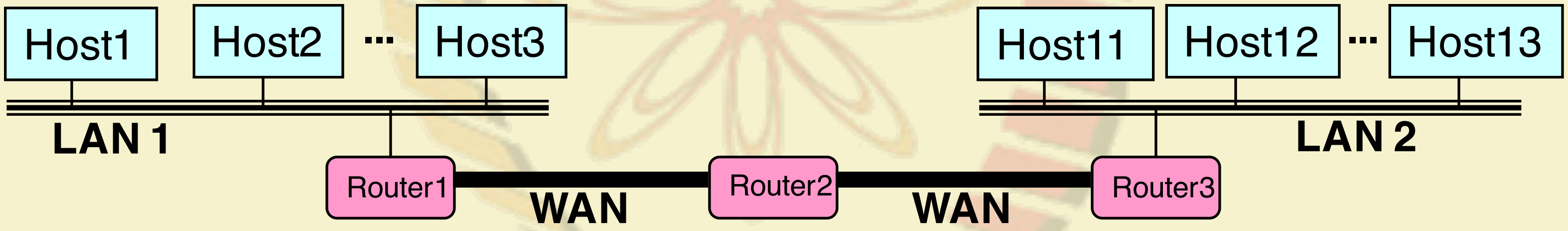
IP Address (IPv4)

- A unique 32-bit number
- Identifies an interface (on a host, on a router, …)
- Represented in dotted-quad notation

14	35	158	15
00001110	00100011	10011110	00001111

Inter-Networking

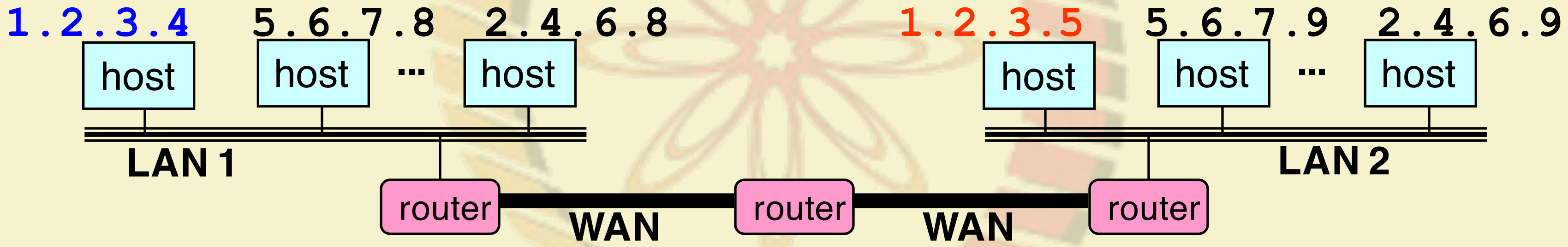
- The Internet is an “inter-network”
  – Used to connect *networks* together, not *hosts*
  – Needs a way to address a network (i.e., group of hosts)



LAN = Local Area Network
WAN = Wide Area Network

Scalability Issue

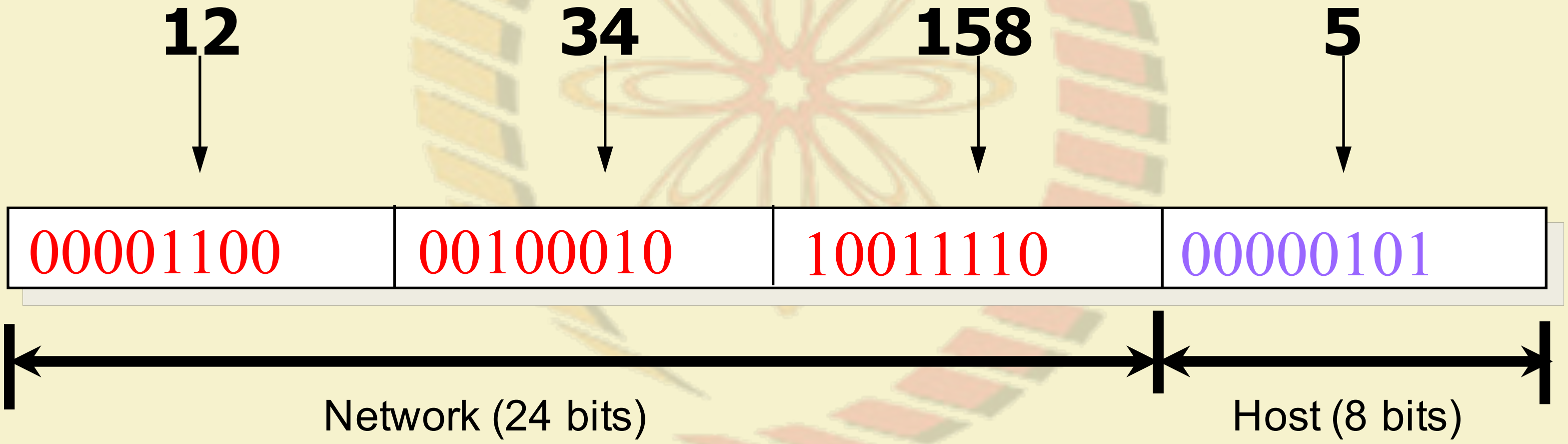
- Suppose hosts had arbitrary addresses, then
  - Every router would need a lot of information
  - To know how to direct packets toward the host



Hierarchical Addressing: IP Prefixes

- Divided into network and host portions (left and right)
- 12.34.158.0/24 is a 24-bit prefix with 2^8 addresses

Analogy with our Postal mail system ?

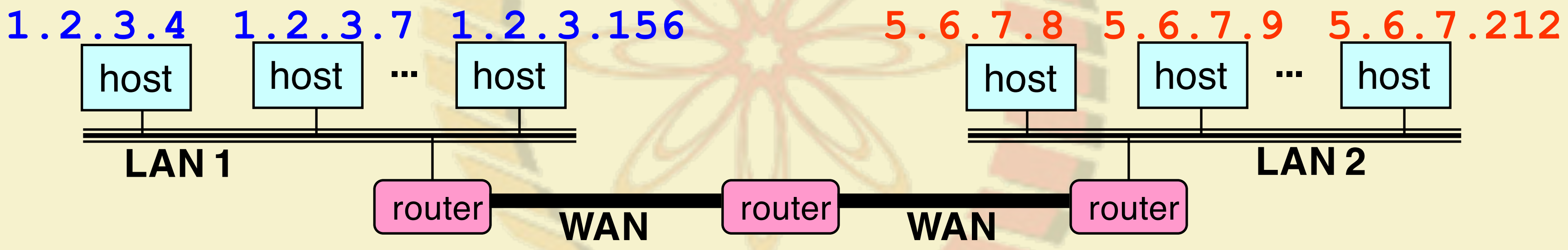


IP Address and a 24-bit Subnet Mask

IP Address	12	34	158	5
	00001100	00100010	10011110	00000101
	11111111	11111111	11111111	00000000
Mask	255	255	255	0

Scalability Improved

- Number related hosts from a common subnet
  - 1.2.3.0/24 on the left LAN
  - 5.6.7.0/24 on the right LAN

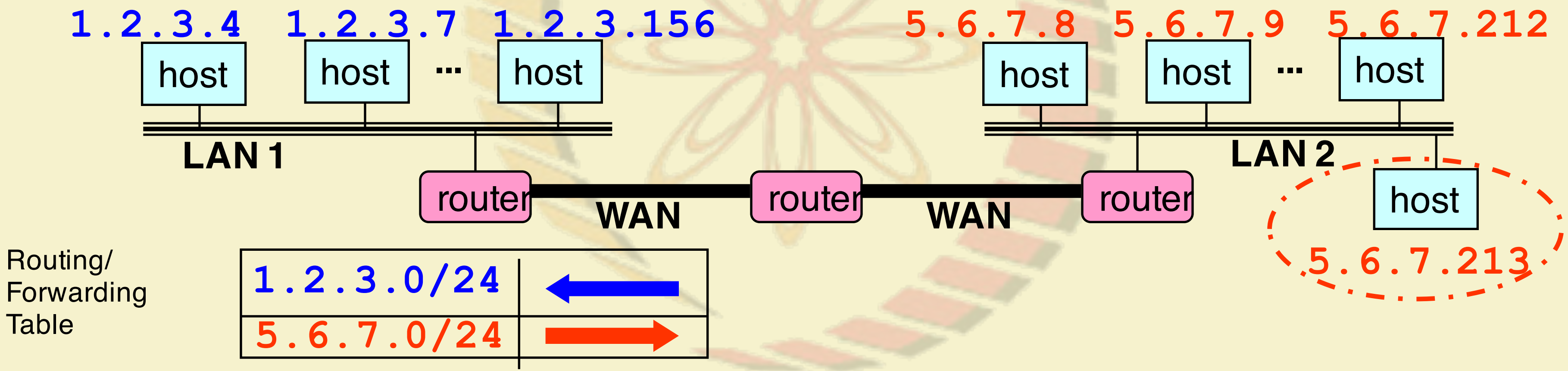


Routing/
Forwarding
Table

1.2.3.0/24	←
5.6.7.0/24	→

Easy to Add New Hosts

- No need to update the routers
  - E.g., adding a new host 5.6.7.213 on the right
  - Doesn't require adding a new forwarding-table entry



IP Address Allocation

Classful Addressing

- Fixed allocation sizes
  - Class A: 0*
    - Very large /8 blocks (e.g., Org. A has 20.0.0.0/8)
  - Class B: 10*
    - Large /16 blocks (e.g,. Org. B has 128.40.0.0/16)
  - Class C: 110*
    - Small /24 blocks (e.g., Org. C has 192.60.225.0/24)
  - Class D: 1110*
    - Multicast groups
  - Class E: 11110*
    - Reserved for future use

Classless Inter-Domain Routing (CIDR)

Use two 32-bit numbers to represent a network.
Network number = IP address + Mask

IP Address : 12.4.0.0 IP Mask: 255.254.0.0

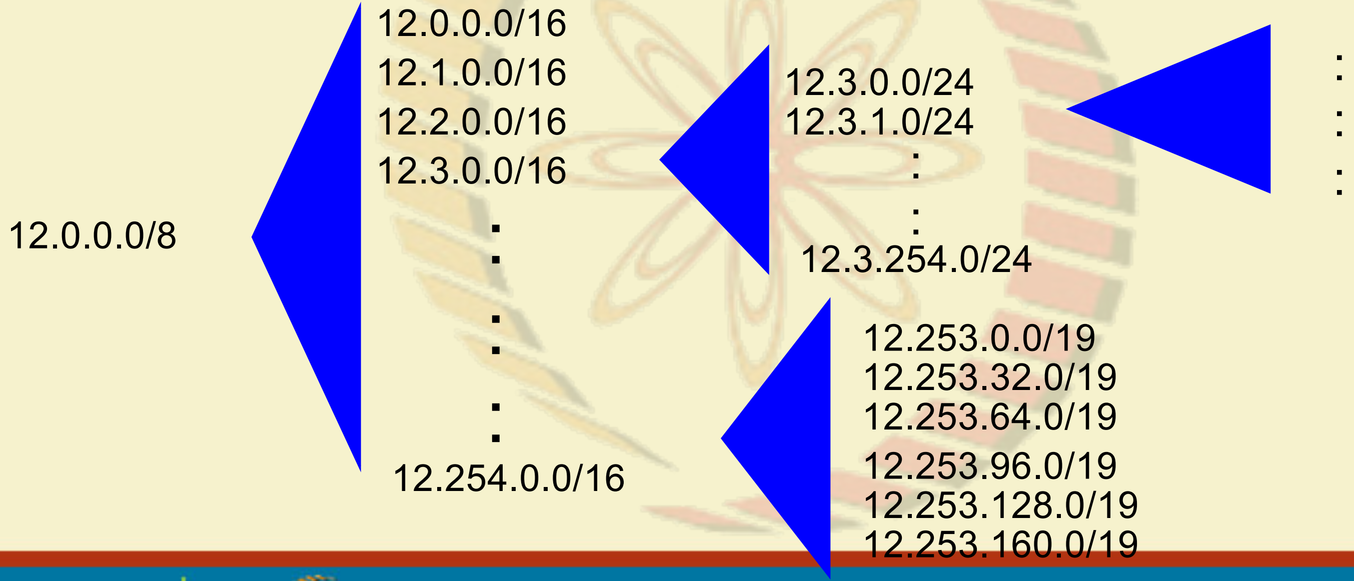
Address	00001100	00000100	00000000	00000000
Mask	11111111	11111110	00000000	00000000

Network Prefix — for hosts

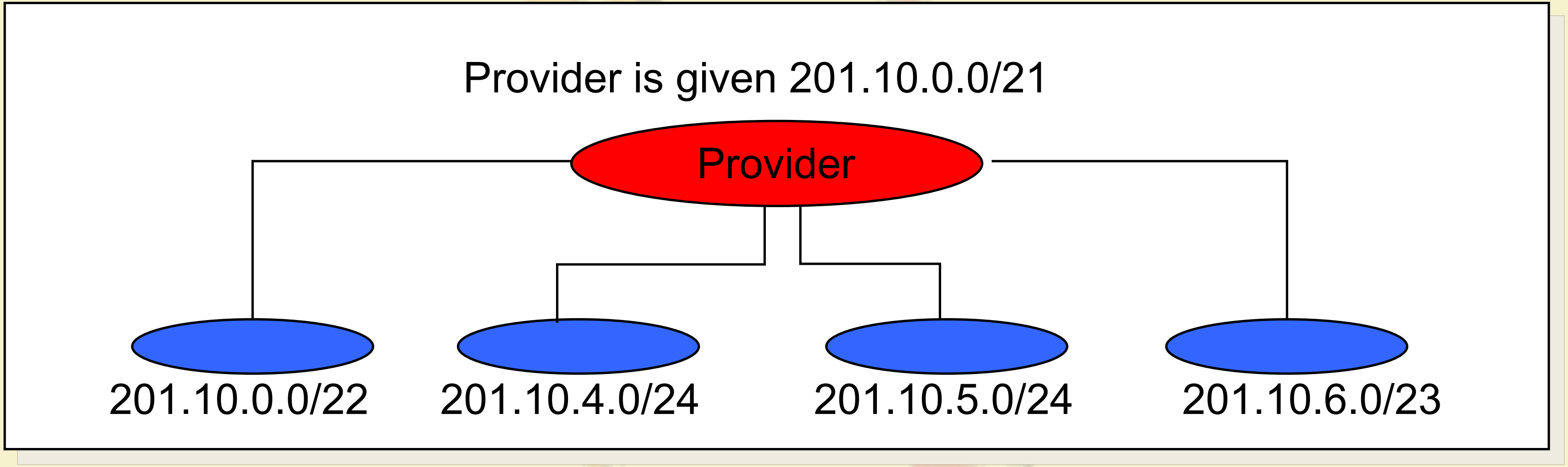
Represented as: 12.4.0.0/15

CIDR: Hierarchal Address Allocation

- Prefixes are key to Internet scalability
  - Address allocated in contiguous chunks (prefixes)
  - Routing protocols and packet forwarding based on prefixes

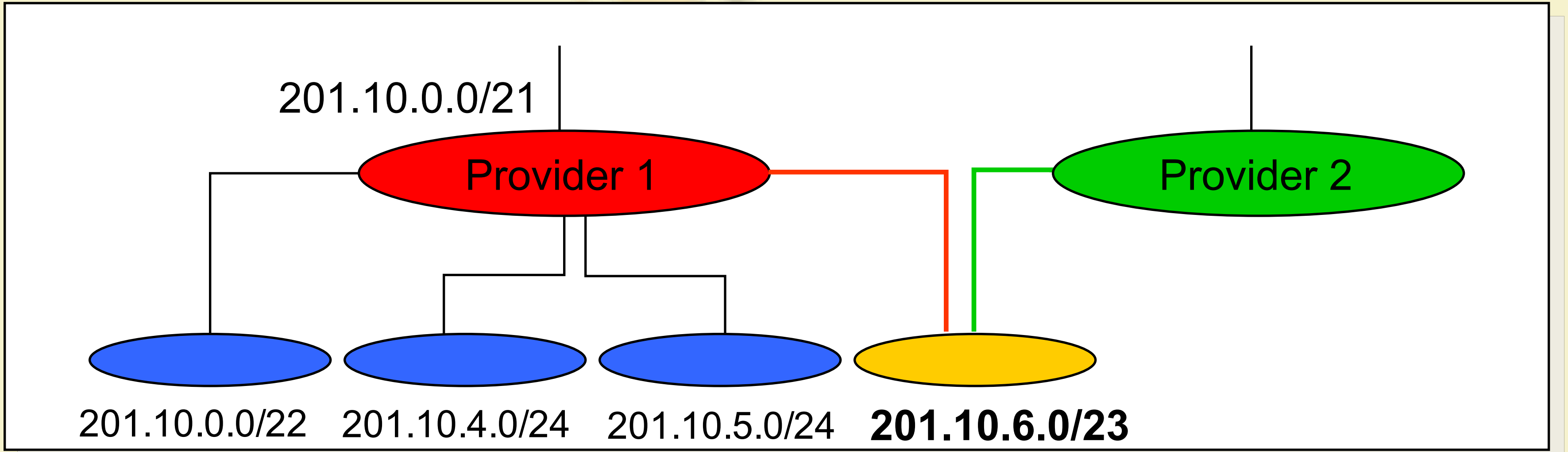


Scalability: Address Aggregation



Routers in the rest of the Internet just need to know how to reach 201.10.0.0/21. The provider can direct the IP packets to the appropriate customer.

Challenges in Aggregation



Multi-homed customer with 201.10.6.0/23 has two providers. Other parts of the Internet need to know how to reach these destinations through *both* providers.

Scalability Through Hierarchy

- Hierarchical addressing
  - Critical for scalable system
  - Don't require everyone to know everyone else
  - Reduces amount of updating when something changes
- Non-uniform hierarchy
  - Useful for heterogeneous networks of different sizes
  - Initial class-based addressing was far too coarse
  - Classless Inter-Domain Routing (CIDR) helps

Obtaining a Block of Addresses

- Separation of control
  - Prefix: assigned *to* an institution
  - Addresses: assigned *by* the institution to their nodes
- Who assigns prefixes?
  - Internet Corporation for Assigned Names and Numbers (ICANN)
    - Allocates large address blocks to Regional Internet Registries
  - Regional Internet Registries (RIRs)
    - E.g., ARIN (American Registry for Internet Numbers)
    - Allocates address blocks within their regions
    - Allocated to Internet Service Providers and large institutions
  - Internet Service Providers (ISPs)
    - Allocate address blocks to their customers
    - Who may, in turn, allocate to their customers…

Ownership of an Address

- Address registries
  - Public record of address allocations
  - Internet Service Providers (ISPs) should update when giving addresses to customers
  - However, records are notoriously out-of-date
- Ways to query
  - UNIX: “whois –h whois.arin.net 128.112.136.35”
  - http://www.arin.net/whois/
  - http://whois.inregistry.in
  - …

Are 32-bit Addresses Enough?

- Not all that many unique addresses
  - $2^{32} = 4,294,967,296$ (just over four billion)
  - Plus, some are reserved for special purposes
  - And, addresses are allocated in larger blocks
- Many devices need IP addresses
  - Computers, PDAs, routers, tanks, toasters, …
- Long-term solution: a larger address space
  - IPv6 has 128-bit addresses ($2^{128} = 3.403 \times 10^{38}$)
- Short-term solutions: with IPv4 address space
  - Private addresses
  - Network address translation (NAT)
  - Dynamically-assigned addresses (DHCP)

Hard Policy Questions

- How much address space per geographic region?
  - Equal amount per country?
  - Proportional to the population?
  - What about addresses already allocated?
- Address space portability?
  - Keep your address block when you change providers?
  - Pro: avoid having to renumber your equipment
  - Con: reduces the effectiveness of address aggregation
- Keeping the address registries up to date?
  - What about mergers and acquisitions?
  - Delegation of address blocks to customers?
  - As a result, the registries are horribly out of date

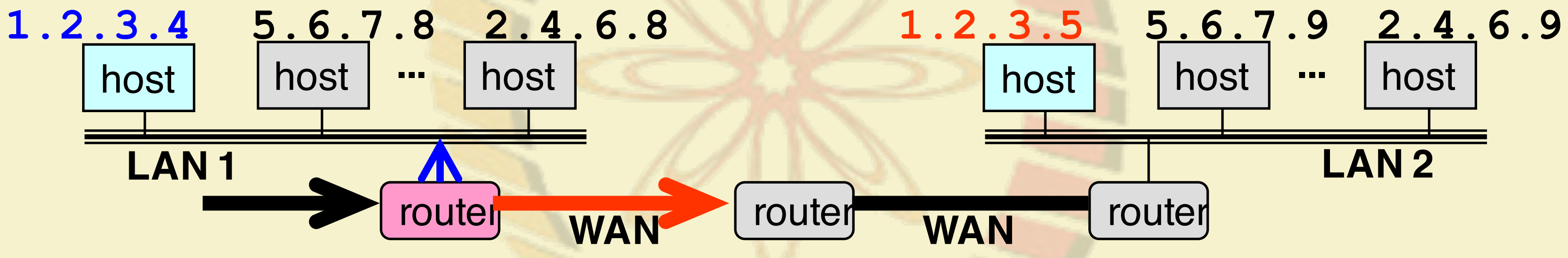
Packet Forwarding

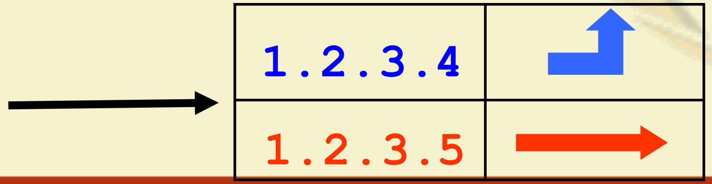
Hop-by-Hop Packet Forwarding

- Each router has a forwarding table
  - Maps destination addresses…
  - … to outgoing interfaces
- Upon receiving a packet
  - Inspect the destination IP address in the header
  - Index into the table
  - Determine the outgoing interface
  - Forward the packet out that interface
- Then, the next router in the path repeats
  - And the packet travels along the path to the destination

Separate Table Entries Per Address

- If a router had a forwarding entry per IP address
  - Match *destination address* of incoming packet
  - ... to the *forwarding-table entry*
  - ... to determine the *outgoing interface*

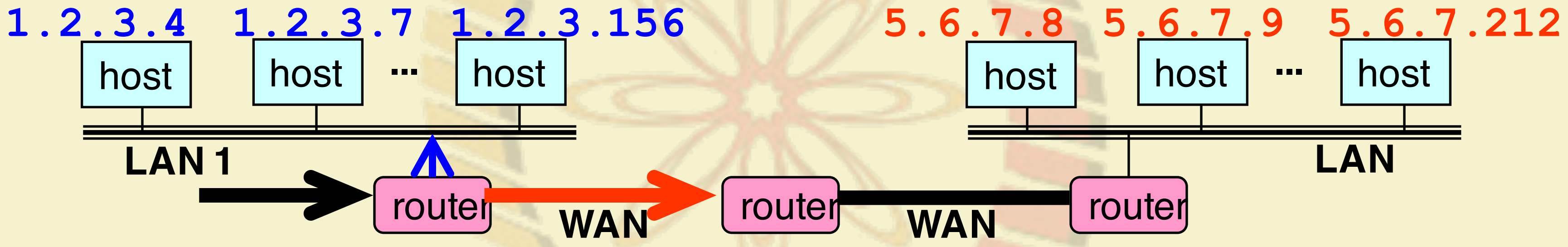




Routing/ Forwarding table

Separate Entry Per 24-bit Prefix

- If the router had an entry per 24-bit prefix
  - Look only at the top 24 bits of the destination address
  - Index into the table to determine the next-hop interface



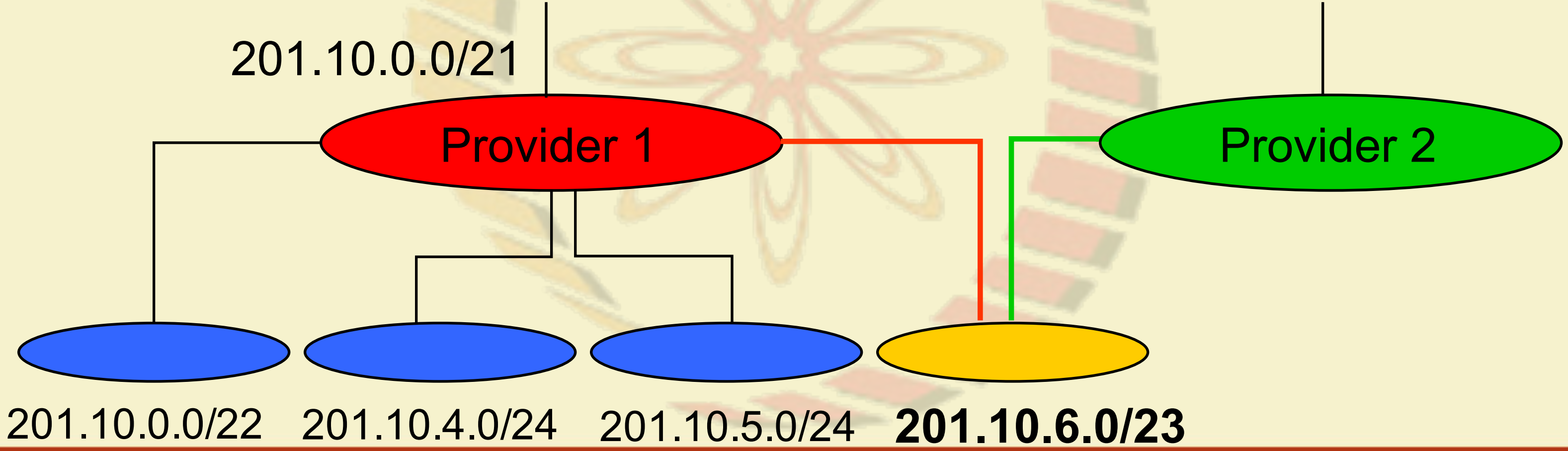
1.2.3.0/24	
5.6.7.0/24	

Separate Entry Classful Address

- If the router had an entry per classful prefix
  - Mixture of Class A, B, and C addresses
  - Depends on the first couple of bits of the destination
- Identify the mask automatically from the address
  - First bit of 0: class A address (/8)
  - First two bits of 10: class B address (/16)
  - First three bits of 110: class C address (/24)
- Then, look in the forwarding table for the match
  - E.g., 1.2.3.4 maps to 1.2.3.0/24
  - Then, look up the entry for 1.2.3.0/24
  - … to identify the outgoing interface

CIDR complicates Packet Forwarding

- CIDR – packet forwarding
  - CIDR allows efficient use of the limited address space
  - But, CIDR makes packet forwarding much harder
- Forwarding table may have many matches
  - E.g., table entries for 201.10.0.0/21 and 201.10.6.0/23
  - The IP address 201.10.6.17 would match *both*!



Longest Prefix Match (LPM) Forwarding

- Forwarding tables in IP routers
  - Maps each IP prefix to next-hop link(s)
- Destination-based forwarding
  - Packet has a destination address
  - Router identifies longest-matching prefix
  - Cute algorithmic problem: very fast lookups

Routing/ forwarding table

destination

201.10.6.17 →

4.0.0.0/8
4.83.128.0/17
201.10.0.0/21
201.10.6.0/23
126.255.103.0/24

→ outgoing link

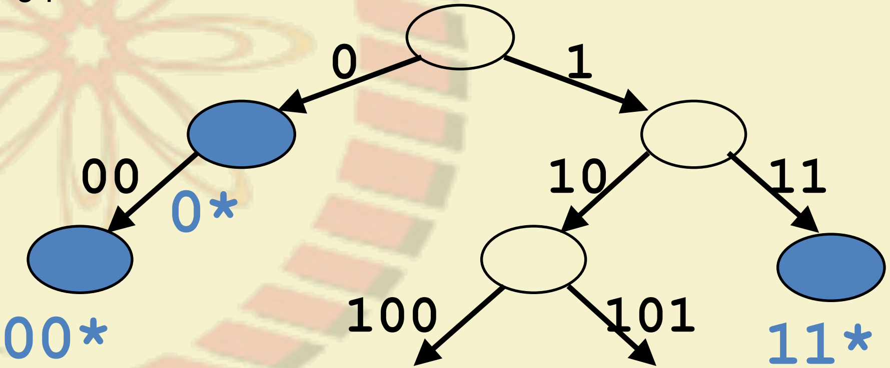
Serial0/0.1

LPM Algorithm

- Scan the forwarding table one entry at a time
  - See if the destination matches the entry
  - If so, check the size of the mask for the prefix
  - Keep track of the entry with longest-matching prefix
- Overhead is linear in size of the forwarding table
  - Table entries are quite large, 200000+
  - The router may have just a few nanoseconds before the next packet is arriving
- Need greater efficiency to keep up with *line rate*
  - Better algorithms
  - Hardware implementations

Patricia Tree

- Store the prefixes as a tree
  - One bit for each level of the tree
  - Some nodes correspond to valid prefixes which have next-hop interfaces in a table
- When a packet arrives
  - Traverse the tree based on the destination address
  - Stop upon reaching the longest matching prefix



Faster Lookups

- Patricia tree is faster than linear scan
  - Proportional to number of bits in the address
- Patricia tree can be made faster
  - Can make a k-ary tree
    - E.g., 4-ary tree with four children (00, 01, 10, and 11)
  - Faster lookup, though requires more space
- Can use special hardware
  - Content Addressable Memories (CAMs)
  - Allows look-ups on a key rather than flat address

Where do Forwarding Tables Come From?

- Routers have forwarding tables
  - Map prefix to outgoing link(s)
- Entries can be statically configured
  - E.g., “map 12.34.158.0/24 to Serial0/0.1”
- But, it is not adadtive
  - to failures
  - to new devices
  - to load balancing
- That is where other technologies come in…
  - Routing protocols, DHCP, and ARP (later in course)

Packet Forwarding by End Devices/ Hosts

- End host with single network interface
  - PC with an Ethernet link
  - Laptop with a wireless link
- Don't need to run a routing protocol
  - Packets to the host itself (e.g., 1.2.3.4/32)
    - Delivered locally
  - Packets to other hosts on the LAN (e.g., 1.2.3.0/24)
    - Sent out the interface
  - Packets to external hosts (e.g., 0.0.0.0/0)
    - Sent out interface to local gateway
- How this information is learned
  - Static setting of address, subnet mask, and gateway
  - Dynamic Host Configuration Protocol (DHCP)

Packet Delivery to the End Hosts?

- How does the last router reach the destination?
- Each interface has a persistent, global identifier
  - MAC (Media Access Control) address
  - Burned in to the adaptors Read-Only Memory (ROM)
  - Flat address structure (i.e., no hierarchy)
- Constructing an address resolution table
  - Mapping MAC address to/from IP address
  - Address Resolution Protocol (ARP)

Summary

- IP address
  - A 32-bit number
  - Allocated in prefixes
  - Non-uniform hierarchy for scalability and flexibility
- Packet forwarding
  - Based on IP prefixes
  - Longest-prefix-match forwarding

MIME Headers

- Located between the Email Header and Body
  - MIME-Version: 1.1
  - Content-Type: type/subtype
  - Content-Transfer-Encoding: encoding type
  - Content-Id: message id
  - Content-Description: textual explanation of non-textual contents

thank you!