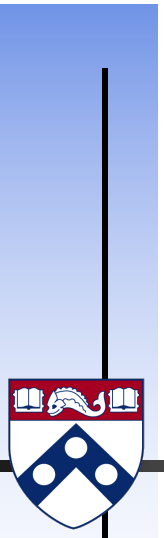


CIS 553: Networked Systems

IP Addressing

February 10, 2020

Agenda

- Link Layer (Part II)
  - Medium access control
  - Switching NEXT
- Network Layer
  - Internet Protocol v4
  - IPv6

Flooding Can Lead to Loops

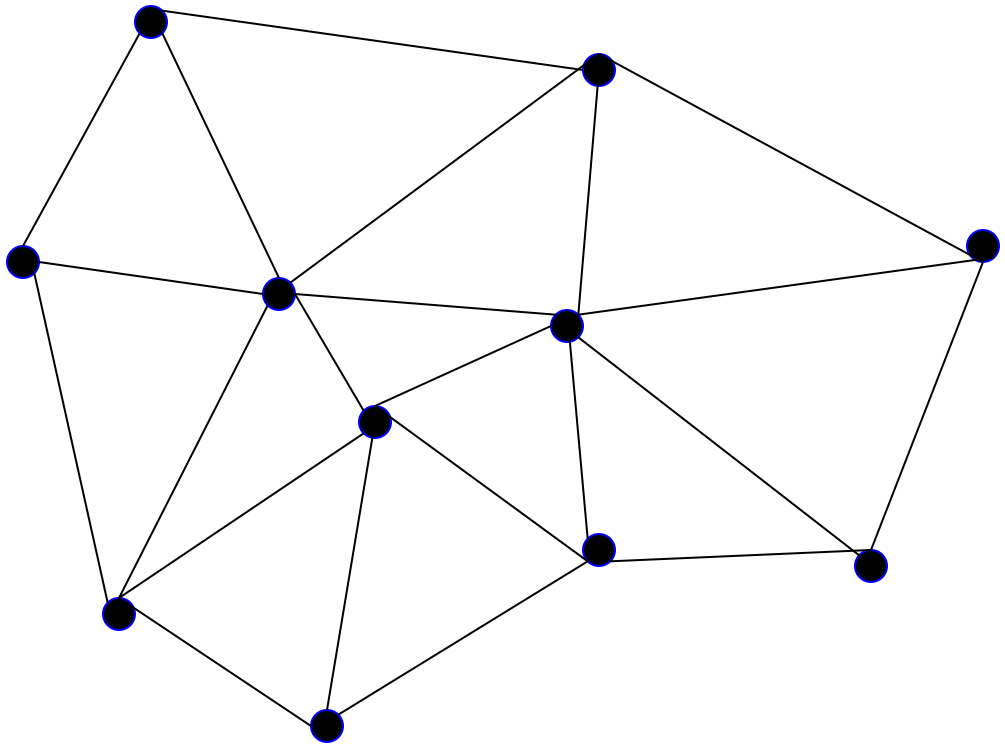
- Flooding can lead to forwarding loops
  - E.g., if the network contains a cycle of switches
  - Either accidentally, or by design for higher reliability

How do we prevent this?

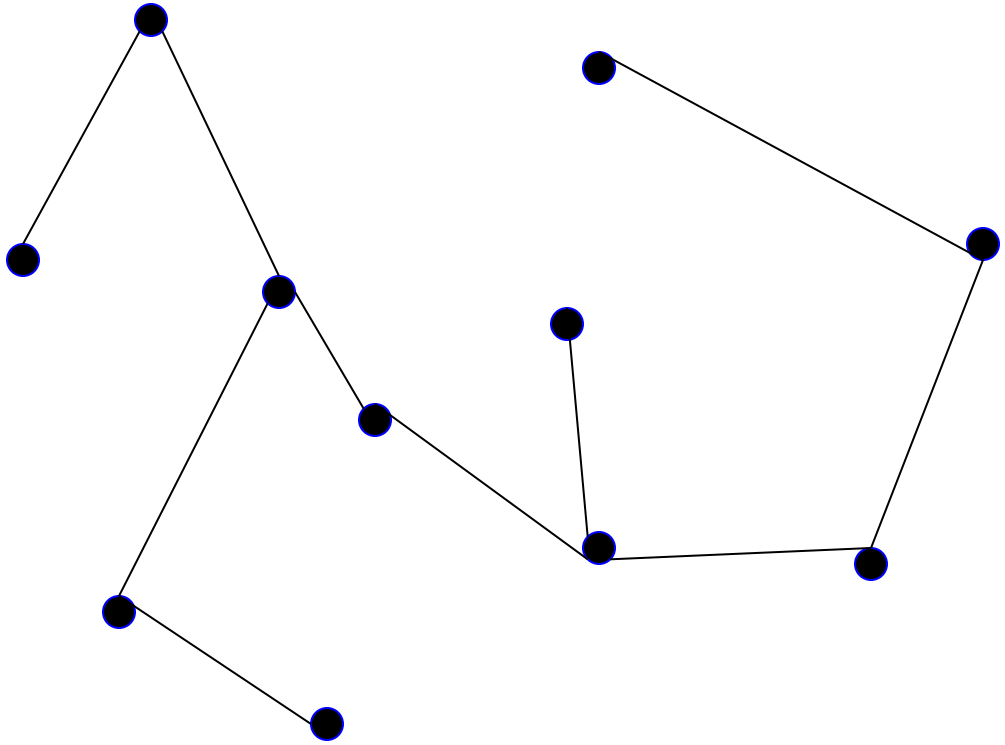
Easiest way to avoid loops

- Use a topology where loops are impossible!
  - This is why Homework 1 didn't have this problem
- Take arbitrary topology and build a spanning tree
  - Sub-graph that includes all vertices but contains no cycles
  - Links not in the spanning tree are not used to forward frames

Consider a graph

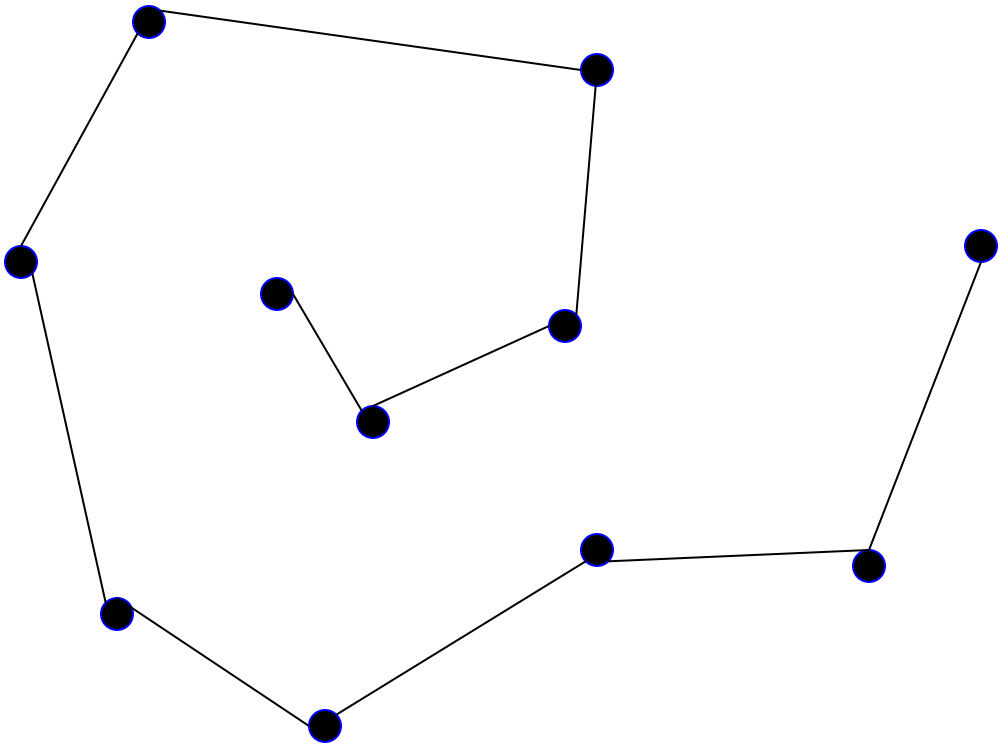


The Spanning Tree

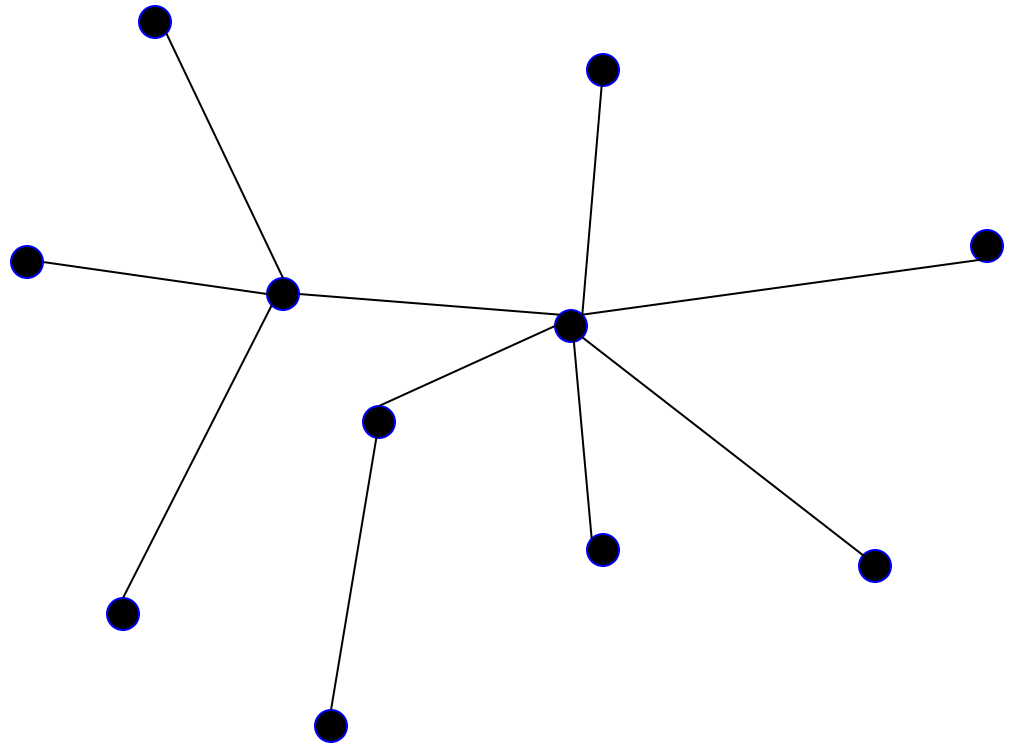


Is this the only possible spanning tree?

Another spanning tree

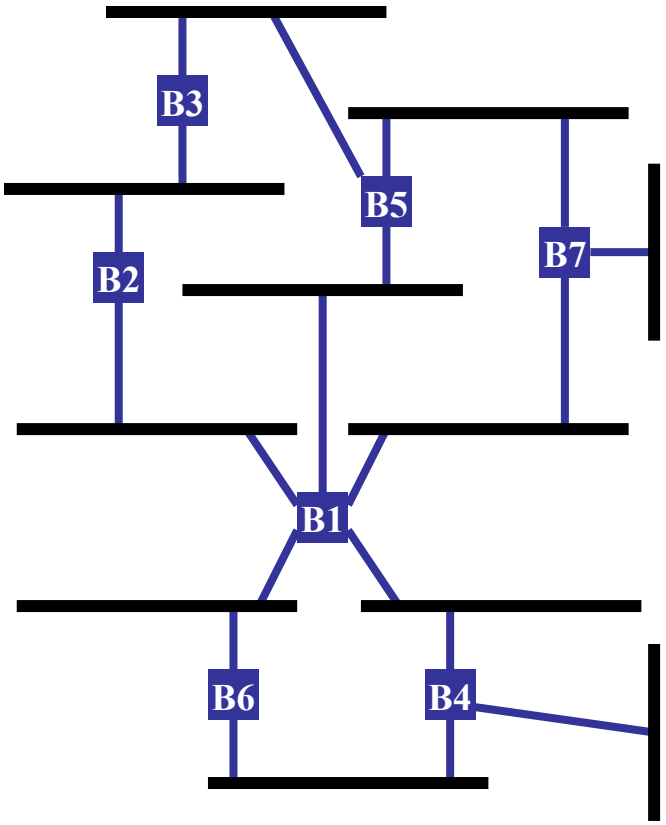


Yet another spanning tree



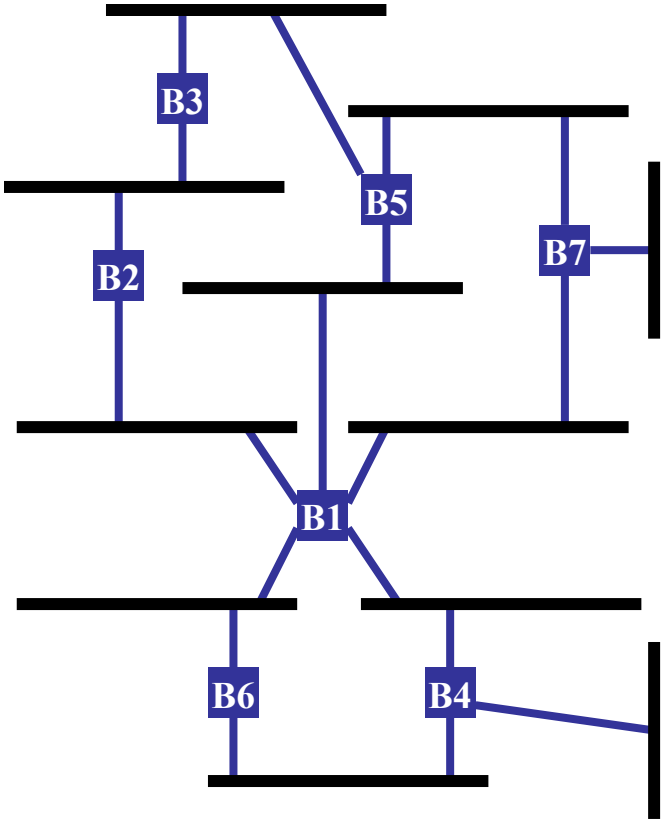
Spanning Tree Algorithm

- Root of the spanning tree is elected first → the bridge with the lowest identifier.
  - All ports are part of tree
- Each bridge finds shortest path to the root.
  - Remembers port that is on the shortest path
  - Used to forward packets
- Select for each LAN a designated bridge that will forward frames to root
  - Has the shortest path to the root.
  - Identifier as tie-breaker



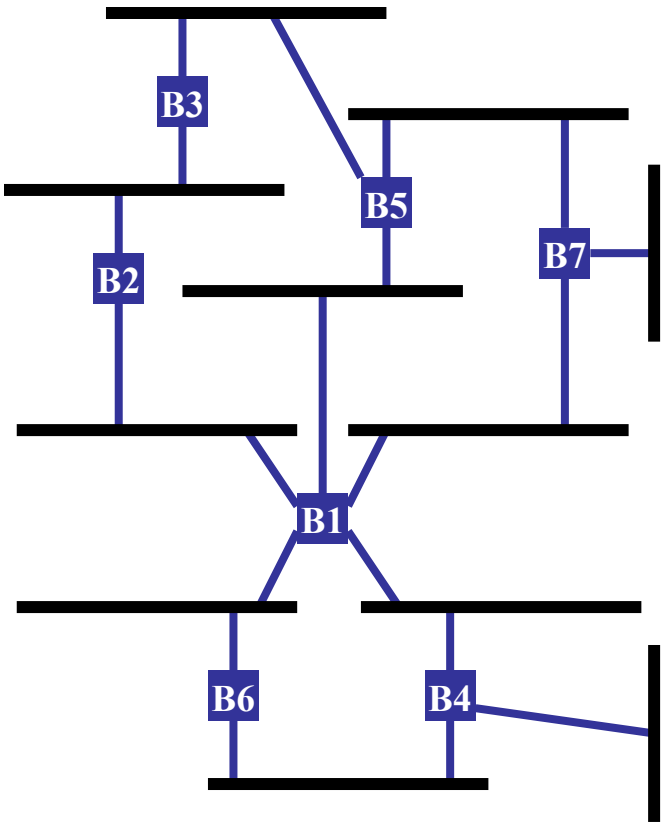
Spanning Tree Algorithm

- Each node sends config message to all neighbors
  - Identifier of the sender
  - Id of the presumed root
  - Distance to the presumed root
- Initially each bridge thinks it is the root
  - B5 sends (B5, B5, 0)
- When B receive a message, it decide whether the solution is better than their local solution
  - A root with a lower identifier?
  - Same root but lower distance?
  - Same root, distance but sender has lower identifier?
- Message from bridge with smaller root ID
  - Not root; stop generating configs, but can forward
- Message from bridge closer to root
  - Not designated bridge; stop sending any config messages on the port



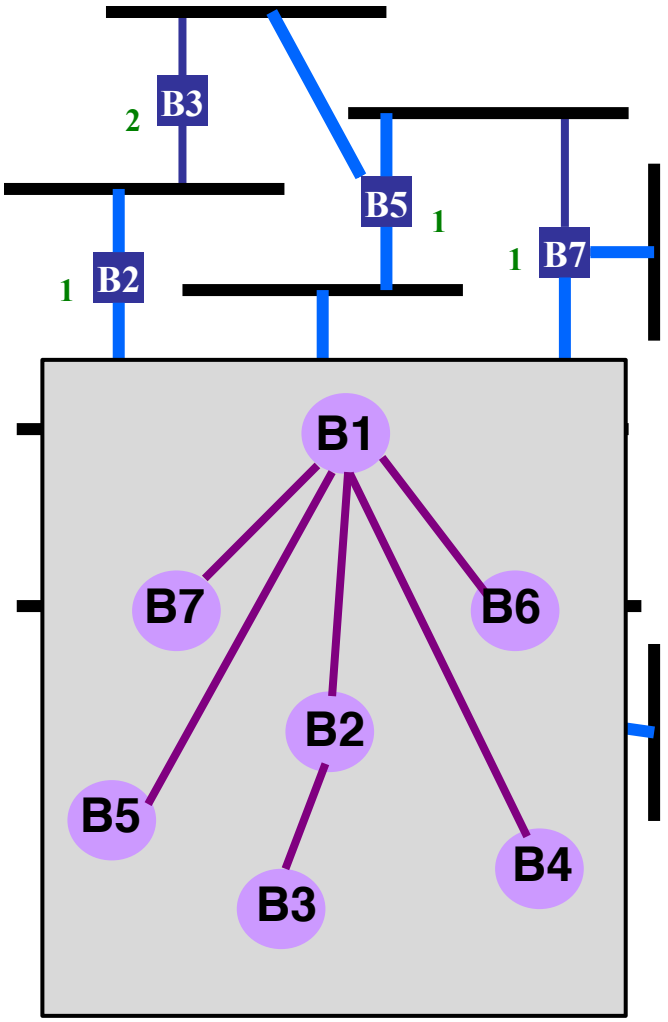
Spanning Tree Algorithm

- Each bridge B can now select which of its ports make up the spanning tree:
  - B's root port
  - All ports for which B is the designated bridge on the LAN
- States for ports on bridges
  - *Forward state* or *blocked state*, depending on whether the port is part of the spanning tree
- Root periodically sends configuration messages and bridges forward them over LANs they are responsible for
- Any bridge failure => Start over



Spanning Tree Algorithm Example

- B3 receives (B2,B2,0)
  - Since 2<3 B3 accepts B2 as a root
- B3 adds one to the distance advertised by B2(0) and thus sends (B3,B2,1) toward B5
- Meanwhile B2 accepts B1 as the root and sends (B2,B1,1)
- B5 accepts B1 as the root and sends (B5,B1,1)
- B3 accepts B1 as the root and figures that B1 and B2 are closer to the root. So stops forwarding on both interfaces.



Robust Spanning Tree Algorithm

- Algorithm must react to failures
  - Failure of the root node
    - Need to elect a new root, with the next lowest identifier
  - Failure of other switches and links
    - Need to recompute the spanning tree
- Key technique: soft-state
  - All entries have timeout, require periodic refreshes
- Root switch continues sending messages
  - Periodically re-announcing itself as the root (1, 0, 1)
  - Other switches continue forwarding messages
  - Stable state: only root send messages
- If switch stops hearing from the root
  - Eventually times out and claims to be the root

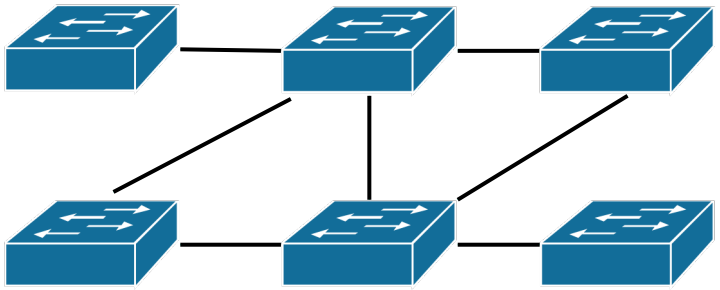
Algorhyme (Radia Perlman, 1985)

I think that I shall never see
A graph more lovely than a tree.
A tree whose crucial property
Is loop-free connectivity.
A tree that must be sure to span
So packets can reach every LAN.
First, the root must be selected.
By ID, it is elected.
Least-cost paths from root are traced.
In the tree, these paths are placed.
A mesh is made by folks like me,
Then bridges find a spanning tree.

Song: https://youtu.be/iE_AbM8ZykI

Layer-2 Summary

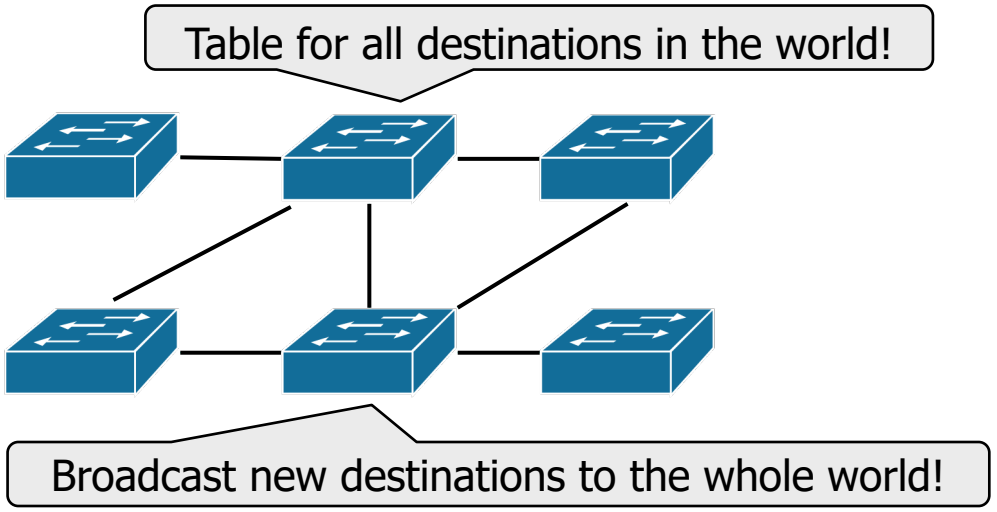
- Switched layer-2 allows:
  - Sending traffic across a network
  - Isolating collision domains



Is this enough to implement the Internet?

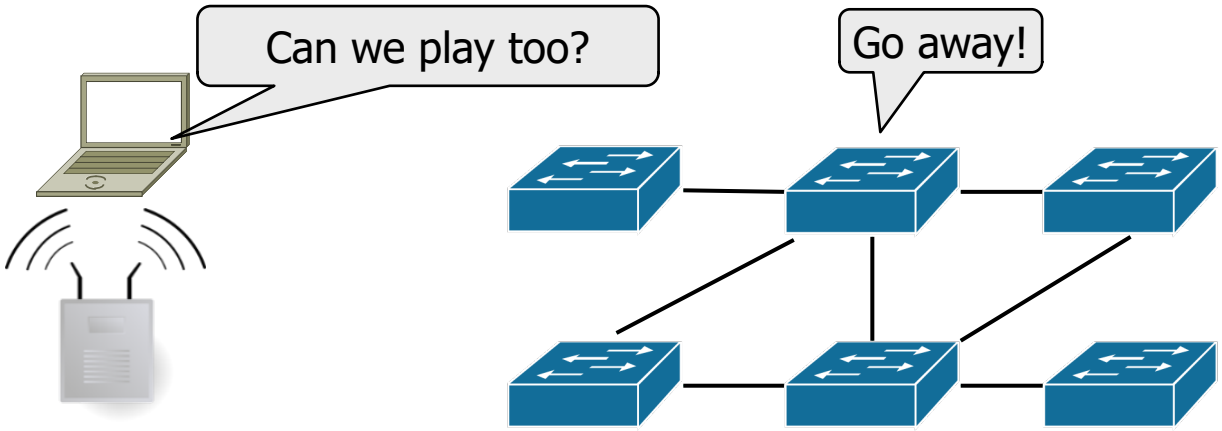
Shortcomings of Switches

1. Don’t scale to large networks
   - Blow up of routing table, broadcast



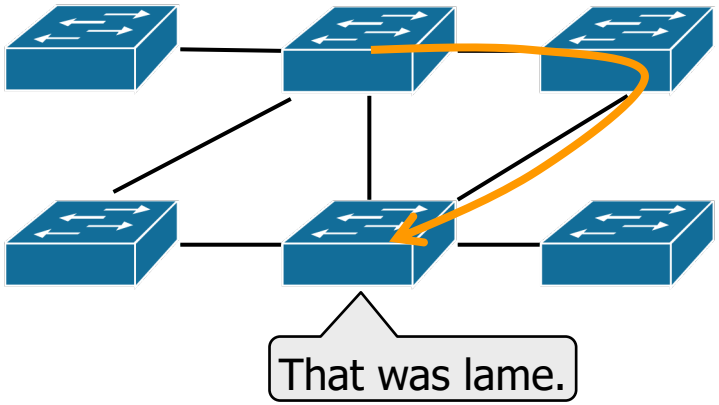
Shortcomings of Switches

1. Don’t scale to large networks
   - Blow up of routing table, broadcast
2. Don’t work across more than one L2 technology
   - Hosts on Ethernet + 3G + 802.11 …



Shortcomings of Switches

1. Don’t scale to large networks
   - Blow up of routing table, broadcast
2. Don’t work across more than one L2 technology
   - Hosts on Ethernet + 3G + 802.11 …
3. Don’t give much traffic control
   - Want to plan routes / bandwidth



Agenda

- Link Layer (Part II)
  - Medium access control
  - Switching
- Network Layer
  - Internet Protocol v4 NEXT
  - IPv6

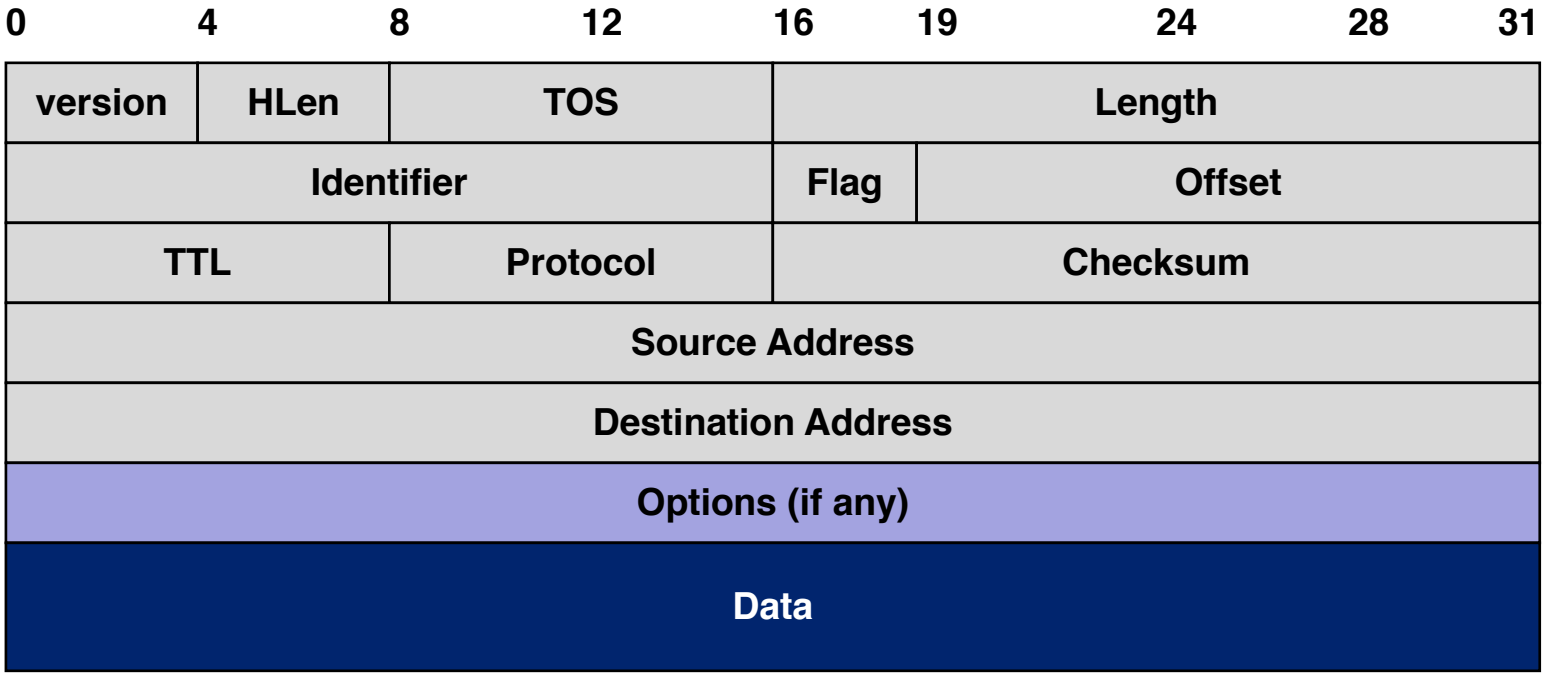
Layer 2: Data link layer

- Connects: Physical interfaces
  - Possibly multiple
  - Usually all the same physical layer protocol
- Name of network: Local Area Network (LAN)
- Name of message: Frame
- Upper interface: A packet of data
- Lower interface: Stream of bits

Layer 3: Network layer

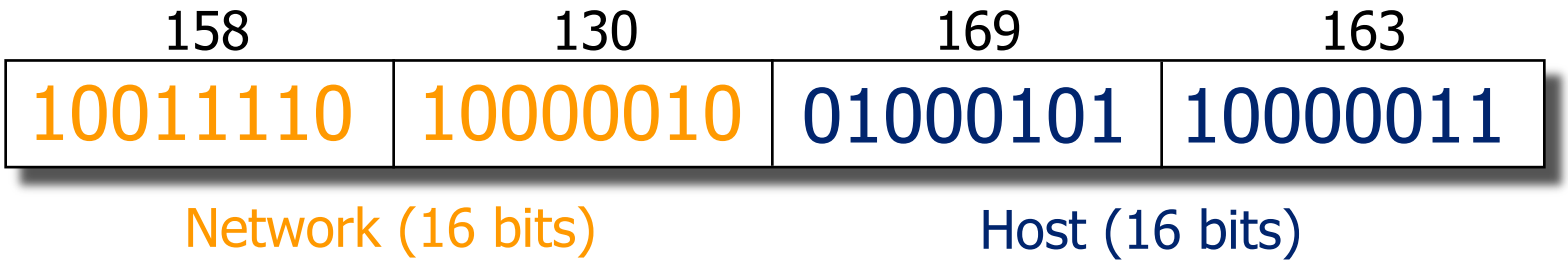
- Connects: LANs
  - All managed by different organizations
- Name of network: Internet
- Name of message: Packet
- Upper interface: A segment of data
- Lower interface: A packet of data
  - Same thing, just has an extra header

IPv4 Packet


0
4
8
12
16
19
24
28
31
version
HLen
TOS
Length
Identifier
Flag
Offset
TTL
Protocol
Checksum
Source Address
Destination Address
Options (if any)
Data

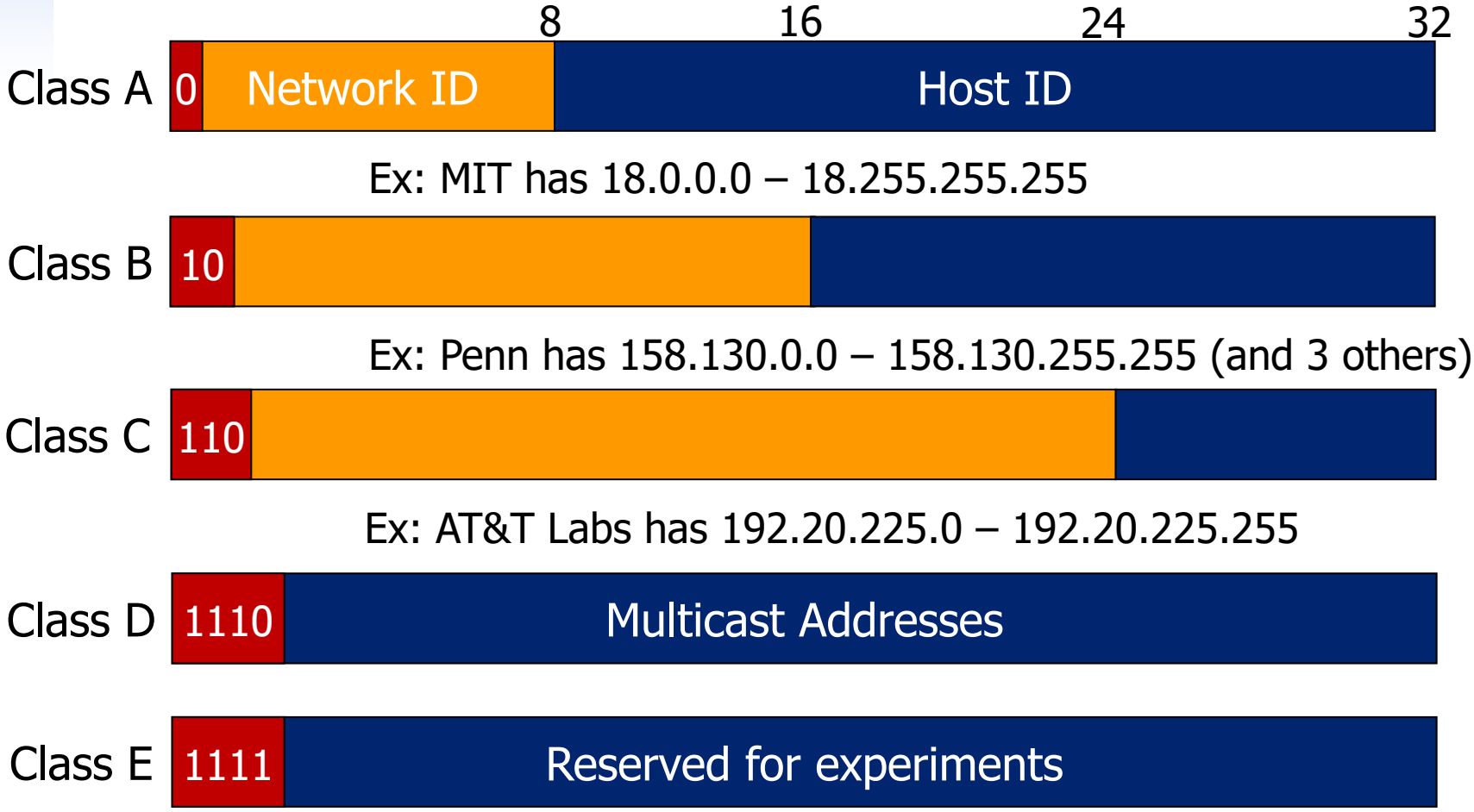
IPv4 Addresses: Networks of Networks

- Location-dependent addressing
- 32-bit number in “dotted-quad” notation
  - cis.upenn.edu --- 158.130.69.163



- Problem: 2^{32} addresses is a lot of table entries
- Solution: Routing based on network and host

Pre-1994: Classful Addressing



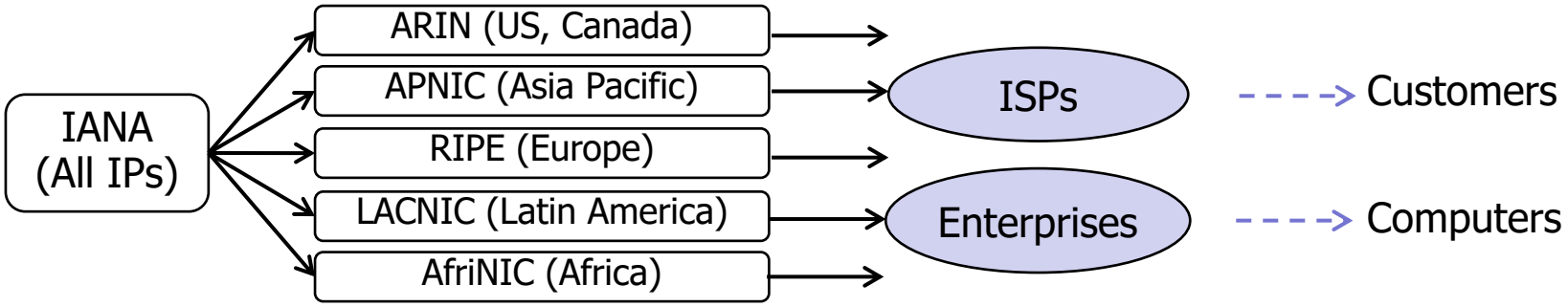
Simple Forwarding: Address range specifies network ID length

Private IP Addresses

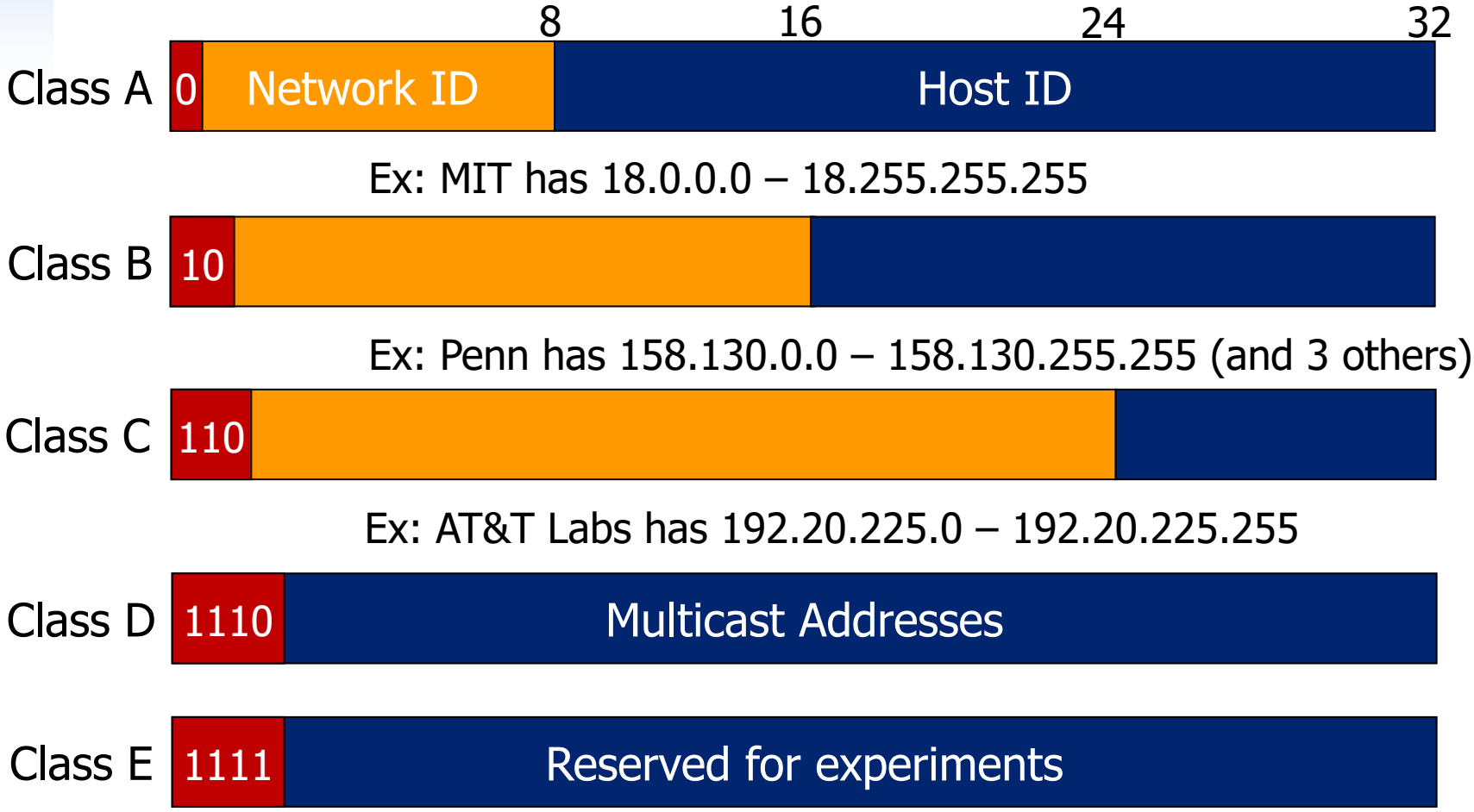
- Can be used freely within private networks (home, small company)
  - 10.0.0.0 – 10.255.255.255
  - 192.168.0.0 – 192.168.255.255
- Need public IP address(es) and NAT to connect to global Internet

Allocating Public IP Addresses

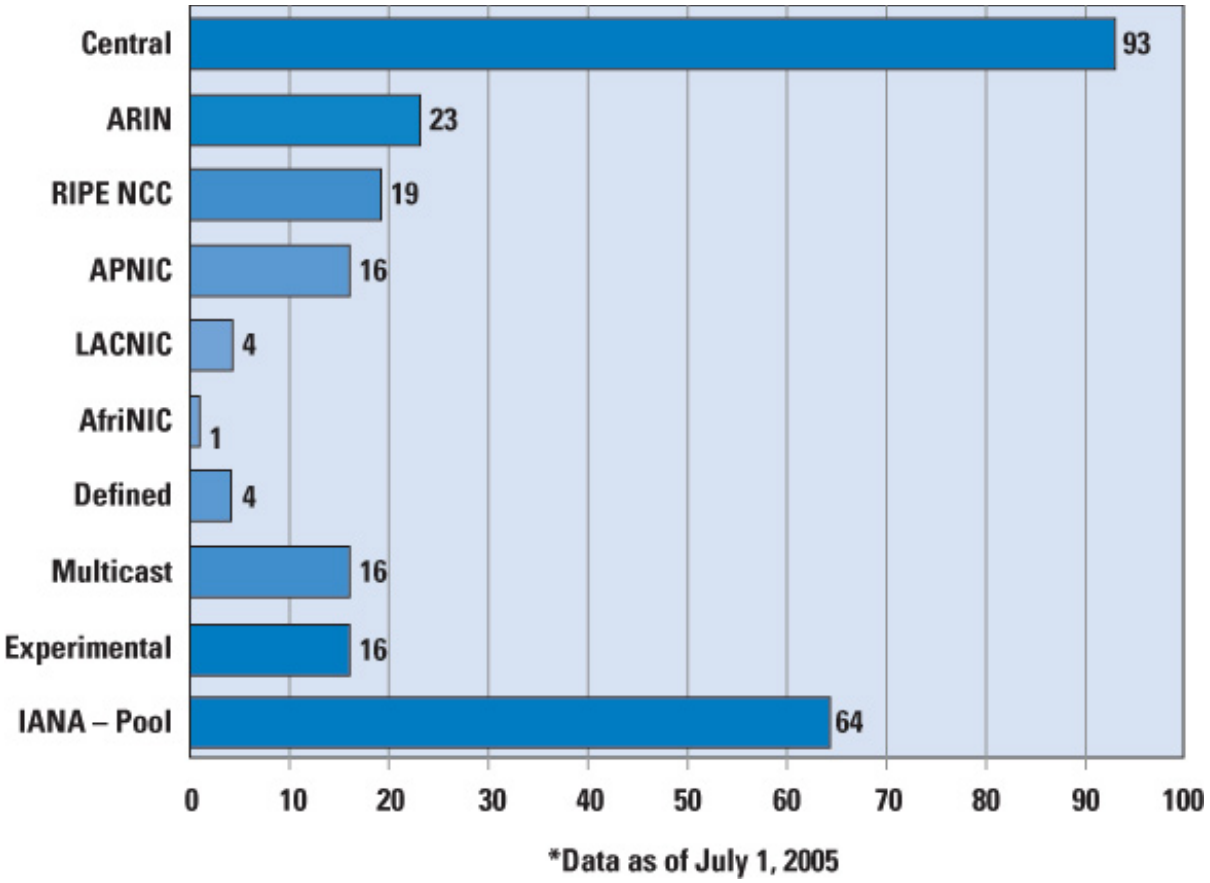
- Follows a hierarchical process
  - IANA delegates to regional bodies (RIRs)
  - RIRs delegate to companies in their region
  - Companies assign to their customers/computers (later, DHCP)



Issues With Classful Addressing



Problem: Fairness and History



- MIT, Ford, Halliburton, Boeing, Merck
- Reclaiming space is difficult. A Class A is a bargaining chip!

Problem: Granularity

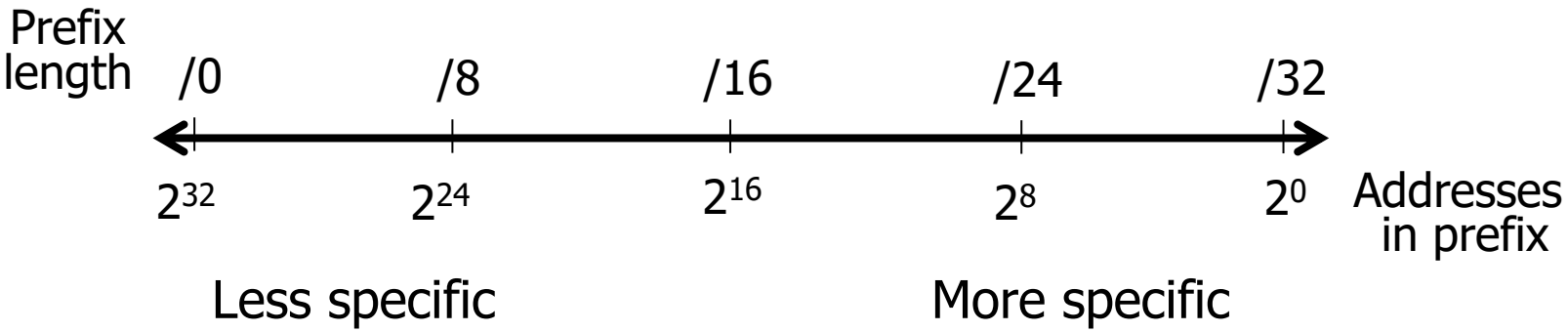
- Example: an organization needs 500 addresses.
- A single class C address not enough
  - 254 hosts
- Instead a class B address is allocated
  - ~65K hosts
- That's overkill, a huge waste!

Classless Interdomain Routing (CIDR)

- Solution: Networks assigned on arbitrary bit boundaries
- Suppose fifty computers in a network are assigned IP addresses 128.23.9.0 - 128.23.9.49
- Range: 01111111 00001111 00001001 00000000 to
  01111111 00001111 00001001 00110001
  - Prefix: 01111111 00001111 00001001 00XX XXXX
- Convention: 128.23.9.0/26
  - There are 32-26=6 bits for the 50 computers
  - 2^6 = 64 addresses

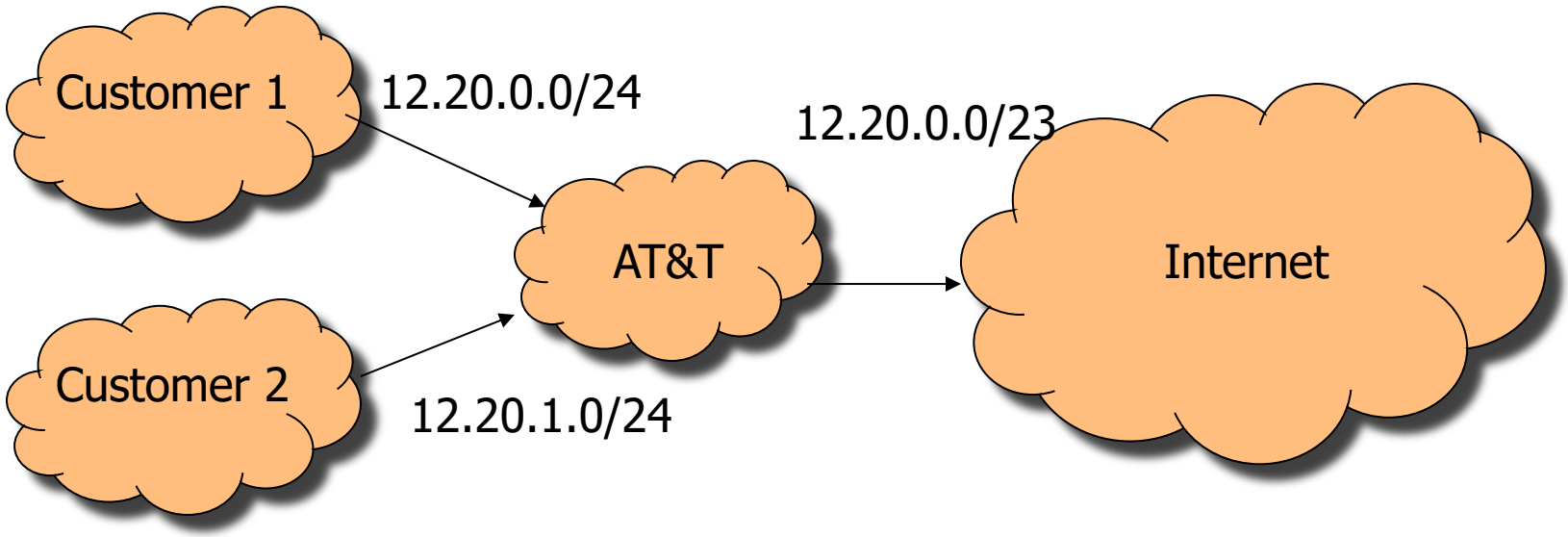
IP Prefixes

- More specific prefix
  - Has longer prefix, hence a smaller number of IP addresses
- Less specific prefix
  - Has shorter prefix, hence a larger number of IP addresses



Benefits of CIDR

- Efficiency: Can allocate blocks of prefixes on a finer granularity
- Hierarchy: Prefixes can be *aggregated* into supernets
  - Typically not done for security…



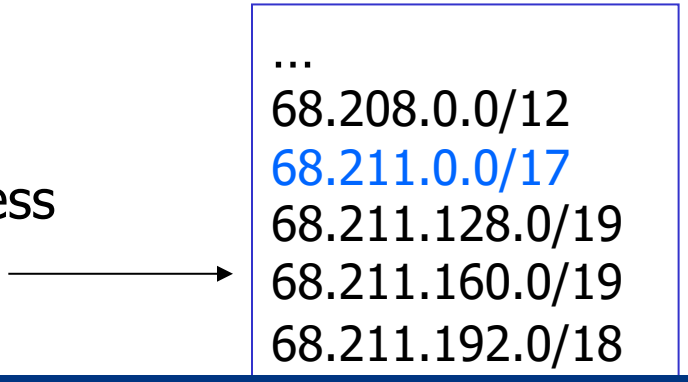
Forwarding: Longest Prefix Match

- Forwarding tables in IP routers
  - Maps each IP prefix to next-hop link(s)
- Entries can overlap!

forwarding table

destination address

68.211.6.120

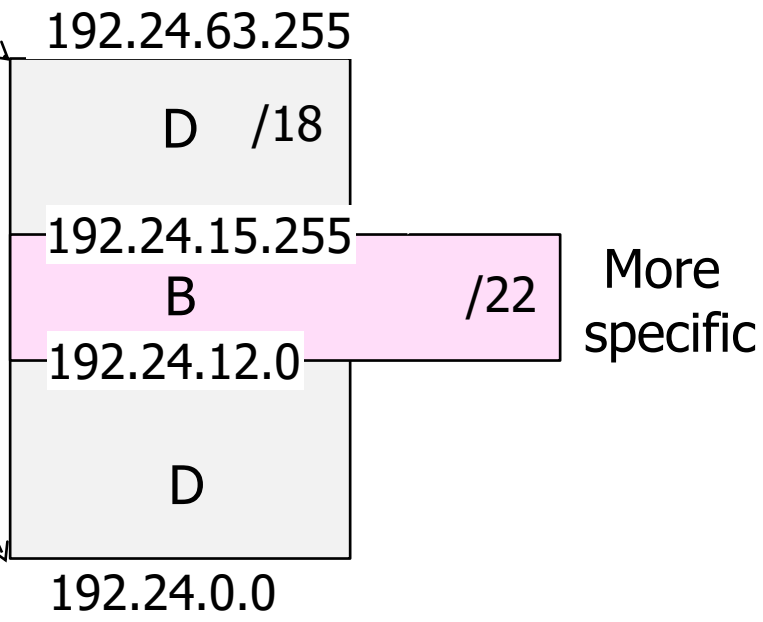


Routing lookup: Find the longest matching prefix (aka the most specific route) among all prefixes that match the destination address.

IP Address Lookup

Prefix	Next Hop
192.24.0.0/18	D
192.24.12.0/22	B

192.24.6.0 →
192.24.14.32 →
192.24.54.0 →



Benefits of Longest Prefix Match

- Can provide default behavior, with less specific prefixes
  - To send traffic going outside an organization to a border router
- Can special case behavior, with more specific prefixes
  - For performance, economics, security, …