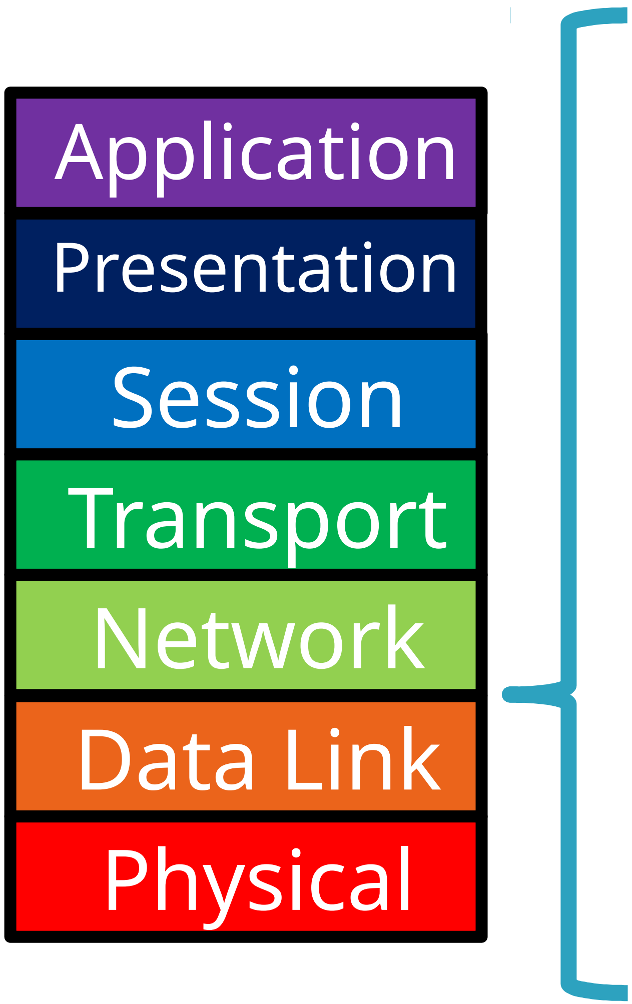
Computer Networks

Lecture 6a: Interconnecting LANs

Based on slides from D. Choffnes Northeastern U. and P. Gill from StonyBrook University
Revised Autumn 2015 by S. Laki

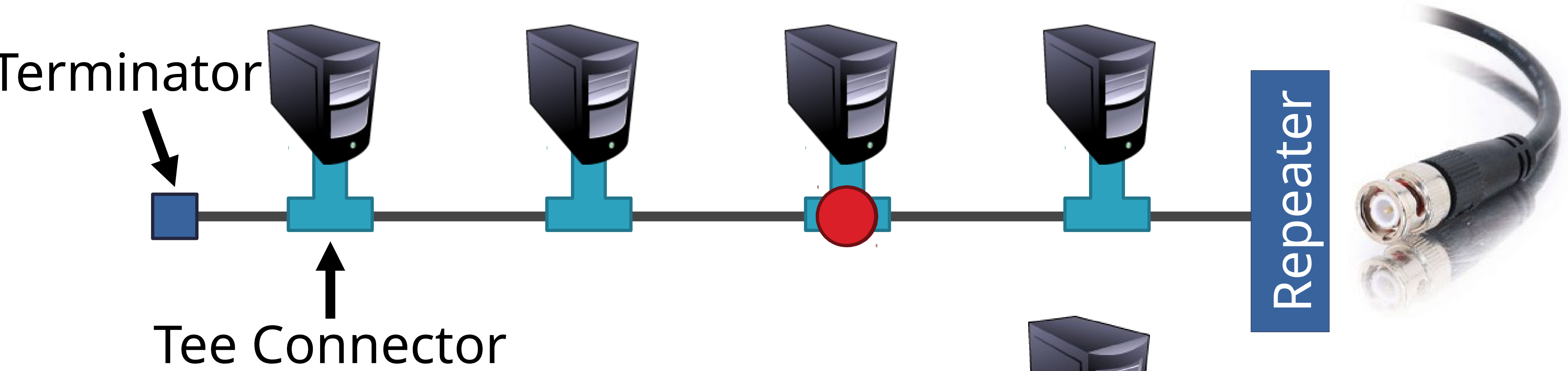
Just Above the Data Link Layer

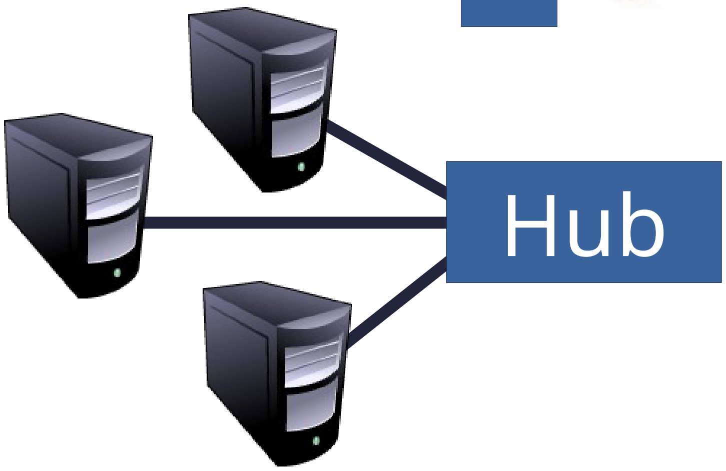


- Bridging
  - How do we connect LANs?
- Function:
  - Route packets between LANs
- Key challenges:
  - Plug-and-play, self configuration
  - How to resolve loops

Recap

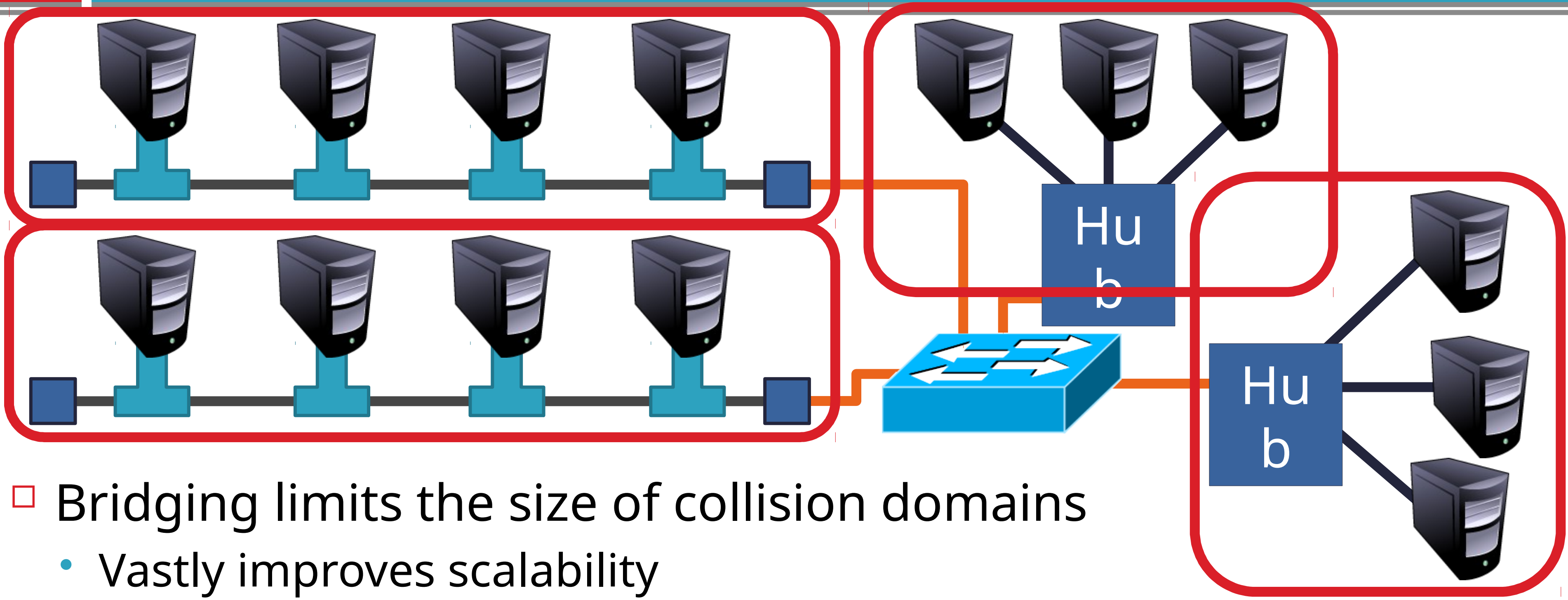
- Originally, Ethernet was a broadcast technology





- Pros: Simplicity
  - Hardware is stupid and cheap
- Cons: No scalability
  - More hosts = more collisions = pandemonium

Bridging the LANs



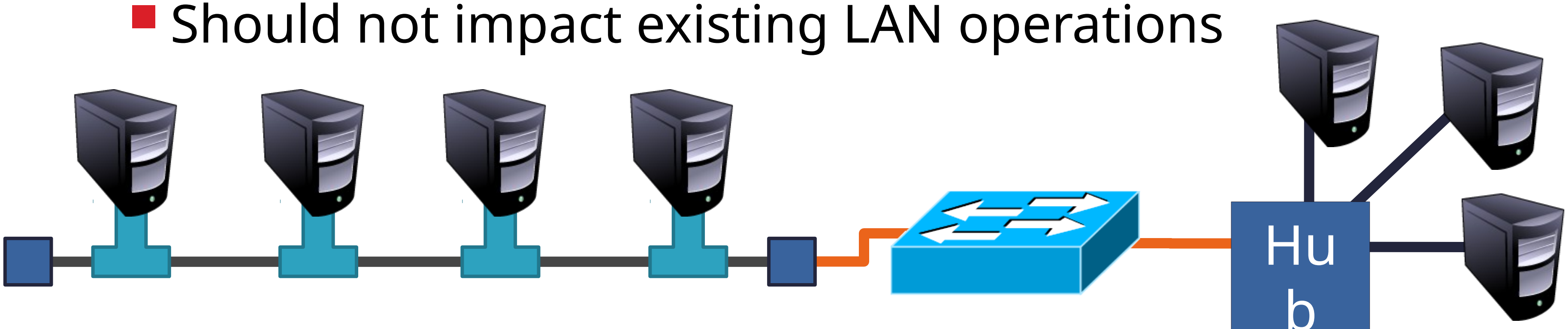
- Bridging limits the size of collision domains
  - Vastly improves scalability
  - Question: could the whole Internet be one bridging domain?
- Tradeoff: bridges are more complex than hubs
  - Physical layer device vs. data link layer device
  - Need memory buffers, packet processing hardware, routing tables

Bridges

- Original form of Ethernet switch
- Connect multiple IEEE 802 LANs at layer 2
-

1. Forwarding of frames
2. Learning of (MAC) Addresses
3. Spanning Tree Algorithm (to handle loops)

  - No hardware of software changes on hosts/hubs
  - Should not impact existing LAN operations



Frame Forwarding Tables

- Each bridge maintains a forwarding table

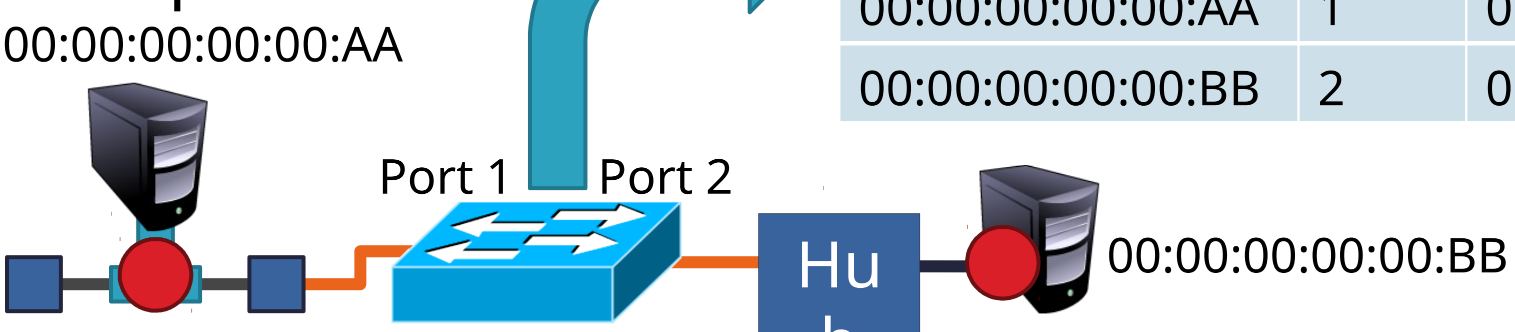
MAC Address	Port	Age
00:00:00:00:00:AA	1	1 minute
00:00:00:00:00:BB	2	7 minutes
00:00:00:00:00:CC	3	2 seconds
00:00:00:00:00:DD	1	3 minutes

Learning Addresses

- Manual configuration is possible, but…
  - Time consuming
  - Error Prone
  - Not adaptable (hosts may get added or removed)
- Instead, learn addresses using a heuristic
  - Look at the source of frames that arrive on each port

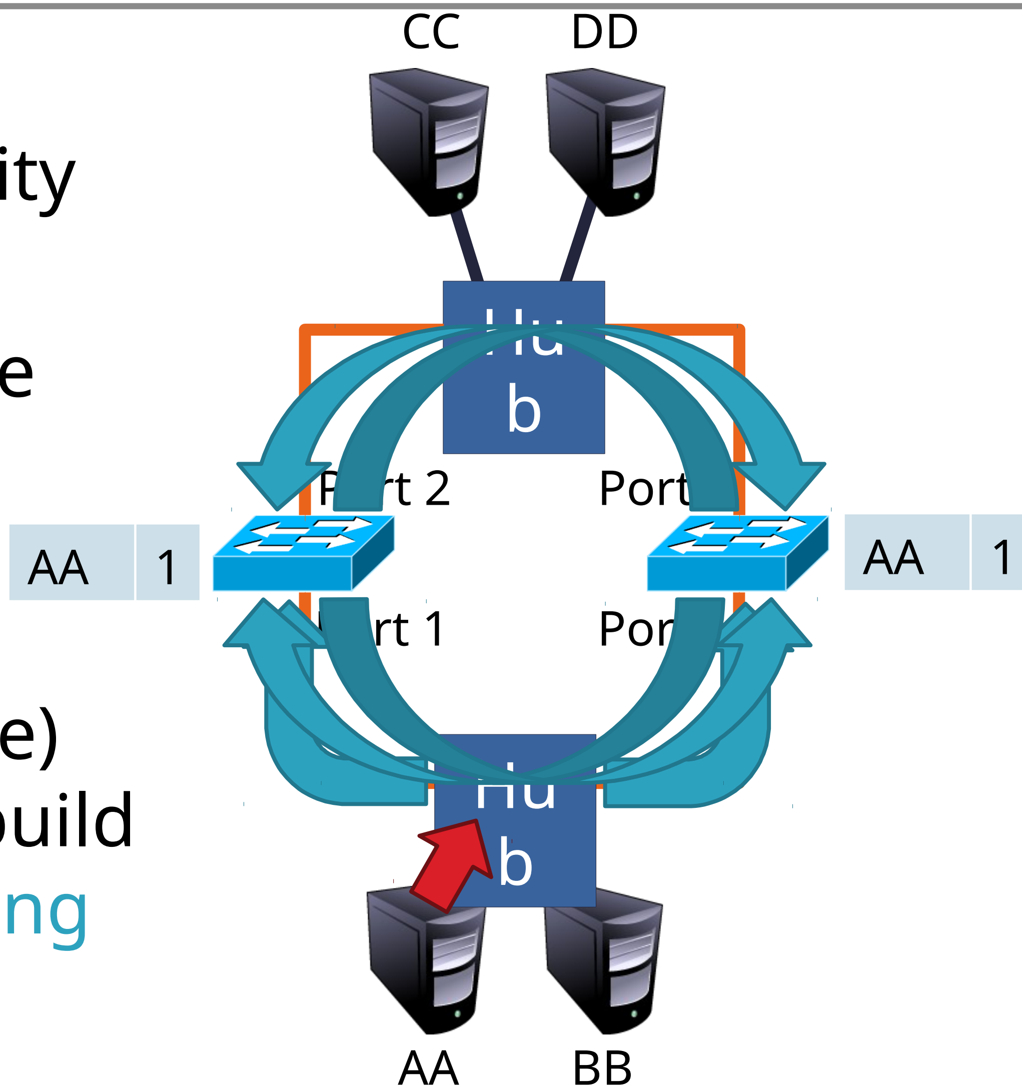
Delete old entries after a timeout

MAC Address	Port	Age
00:00:00:00:00:AA	1	0 minutes
00:00:00:00:00:BB	2	0 minutes



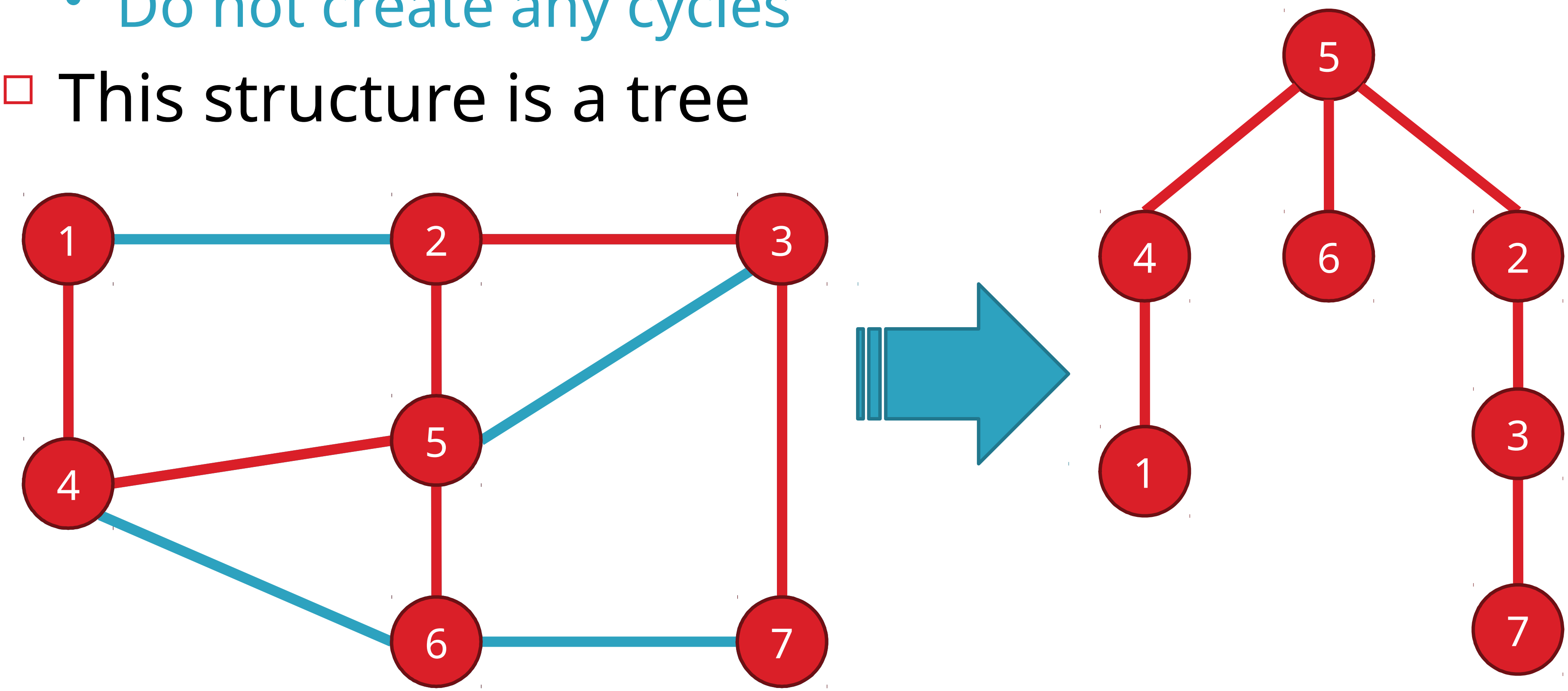
The Danger of Loops

- <Src=AA, Dest=DD>
- This continues to infinity
  - How do we stop this?
- Remove loops from the topology
  - Without physically unplugging cables
- 802.1 (LAN architecture) uses an algorithm to build and maintain a spanning tree for routing



Spanning Tree Definition

- A subset of edges in a graph that:
  - Span all nodes
  - Do not create any cycles
- This structure is a tree



802.1 Spanning Tree Approach

1. Elect a bridge to be the root of the tree
2. Every bridge finds shortest path to the root
3. Union of these paths becomes the spanning tree

- Bridges exchange Configuration Bridge Protocol Data Units (BPDUs) to build the tree
  - Used to elect the root bridge
  - Calculate shortest paths
  - Locate the next hop closest to the root, and its port
  - Select ports to be included in the spanning trees

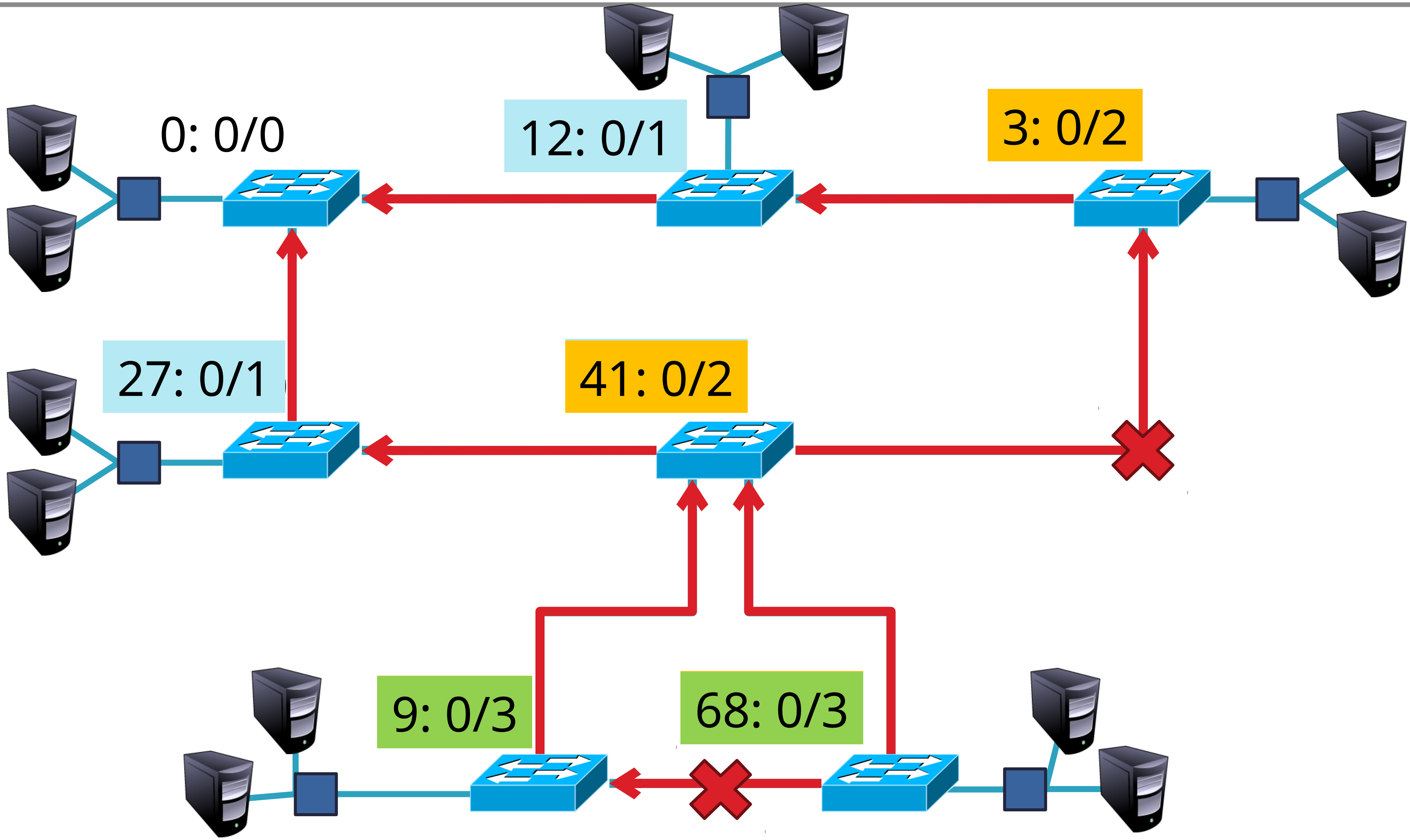
Determining the Root

- Initially, all hosts assume they are the root
- Bridges broadcast BPDUs:

Bridge ID	Root ID	Path Cost to Root

- Based on received BPDUs, each switch chooses:
  - A new root (smallest known Root ID)
  - A new root port (what interface goes towards the root)
  - A new designated bridge (who is the next hop to root)

Spanning Tree Construction



Bridges vs. Switches

- Bridges make it possible to increase LAN capacity
  - Reduces the amount of broadcast packets
  - No loops
- Switch is a special case of a bridge
  - Each port is connected to a single host
    - Either a client machine
    - Or another switch
  - Links are full duplex
  - Simplified hardware: no need for CSMA/CD!
  - Can have different speeds on each port

Switching the Internet

- Capabilities of switches:
  - Network-wide routing based on MAC addresses
  - Learn routes to new hosts automatically
  - Resolve loops
- Could the whole Internet be one switching domain?

NO

Limitations of MAC Routing

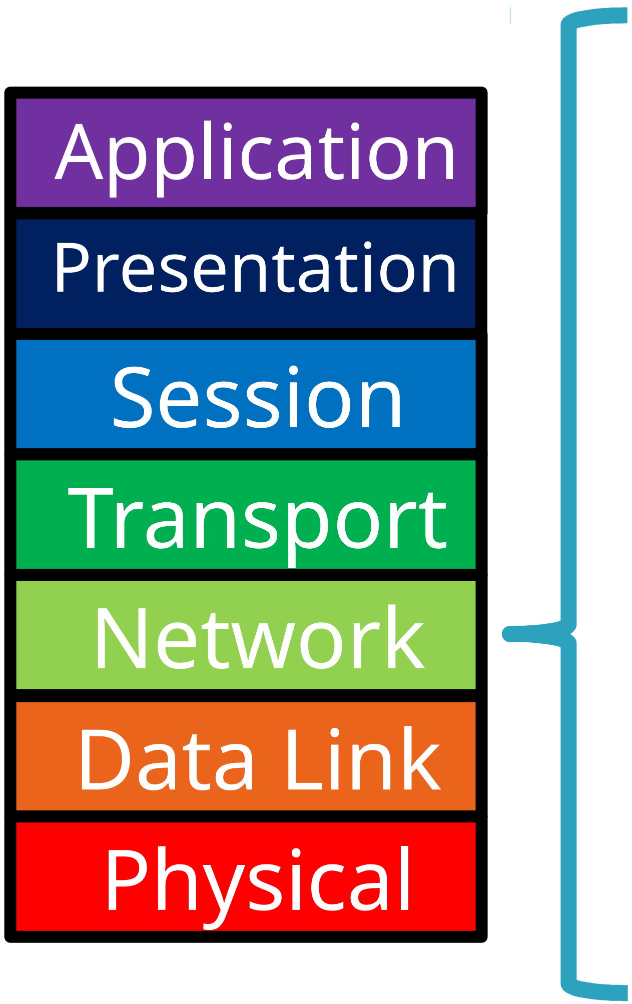
- Inefficient
  - Flooding packets to locate unknown hosts
- Poor Performance
  - Spanning tree does not balance load
  - Hot spots
- Extremely Poor Scalability
  - Every switch needs every MAC address on the Internet in its routing table!
- IP addresses these problems (next …)

Computer Networks

Lecture 6b: Network Layer

Based on slides from D. Choffnes Northeastern U. and P. Gill from StonyBrook University
Revised Autumn 2015 by S. Laki

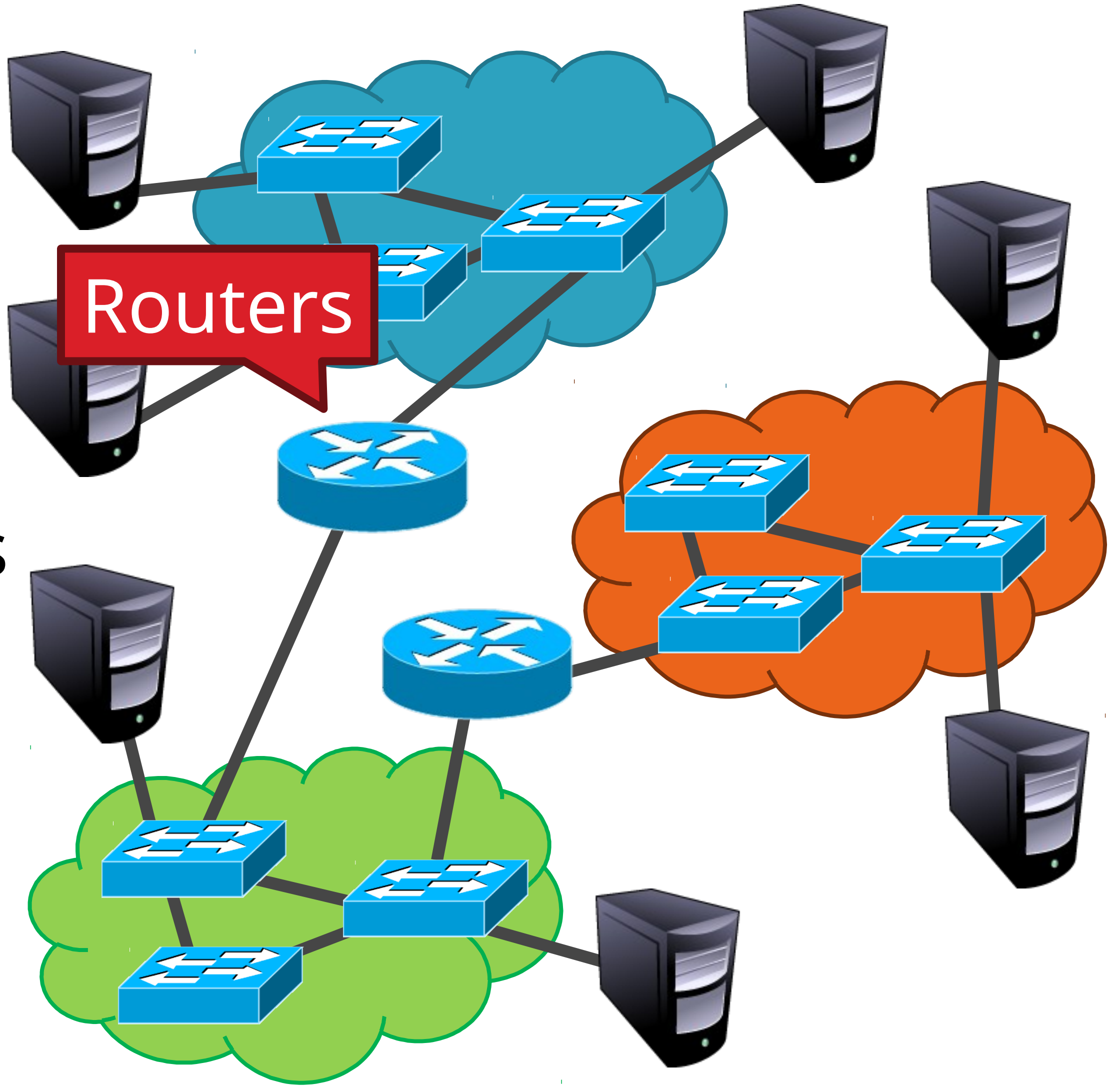
Network Layer



- Function:
  - Route packets end-to-end on a network, through multiple hops
- Key challenge:
  - How to represent addresses
  - How to route packets
    - Scalability
    - Convergence

Routers, Revisited

- How to connect multiple LANs?
- LANs may be incompatible
  - Ethernet, Wifi, etc…
- Connected networks form an internetwork
  - The Internet is the best known example



Int

Internet Service Model

- Best-effort (i.e. things may break)
- Store-and-forward datagram network

Lowest common denominator

- N
  -
- R
  -
  work)
- Service Model
  - What gets sent?
  - How fast will it go?
  - What happens if there are failures?
  - Must deal with heterogeneity
    - Remember, every network is different

Outline

- Addressing
  - Class-based
  - CIDR
- IPv4 Protocol Details
  - Packed Header
  - Fragmentation
- IPv6

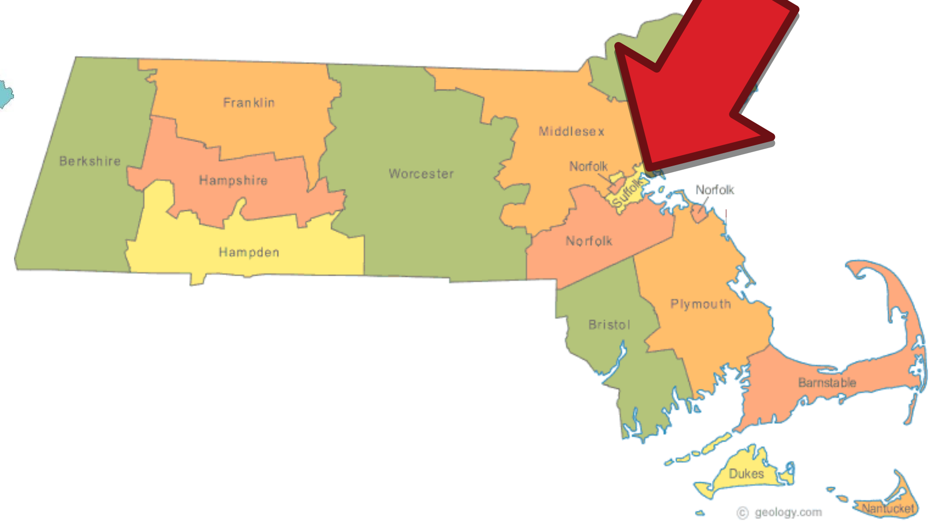
Possible Addressing Schemes

- Flat
  - e.g. each host is identified by a 48-bit MAC address
  - Router needs an entry for every host in the world
    - Too big
    - Too hard to maintain (hosts come and go all the time)
    - Too slow (more later)
- Hierarchy
  - Addresses broken down into segments
  - Each segment has a different level of specificity

Example: Telephone Numbers

1-617-373- 3278

Very General



Northeastern University

West Village G
Room 1234

West Village H
Room 256

Updates are Local

Very Specific

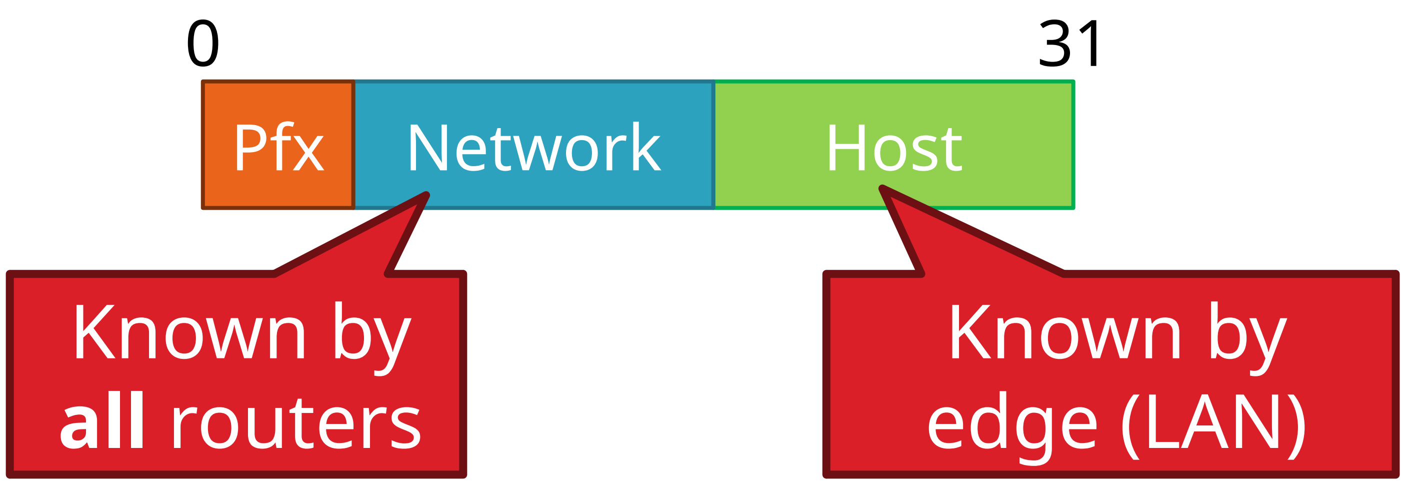
IP Addressing

- IPv4: 32-bit addresses
  - Usually written in dotted notation, e.g. 192.168.21.76
  - Each number is a byte

	0	8	16	24 (to 31)
Decimal	192	168	21	76
Hex	C0	A8	15	4C
Binary	11000000	10101000	00010101	01001100

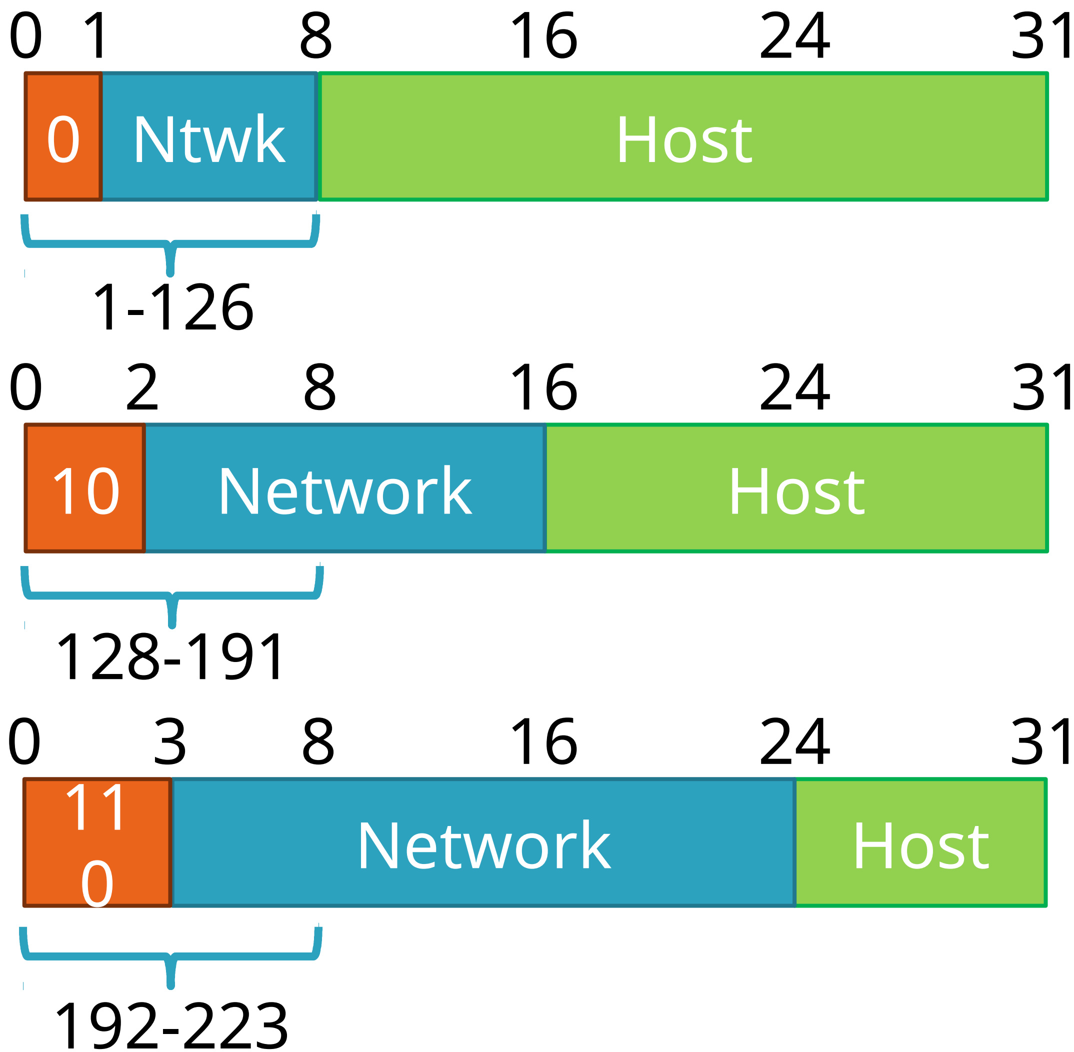
IP Addressing and Forwarding

- Routing Table Requirements
  - For every possible IP, give the next hop
  - But for 32-bit addresses, 2^{32} possibilities!
- Hierarchical address scheme
  - Separate the address into a network and a host



Classes of IP Addresses

- Class A
- Class B
- Class C



Example: MIT
18.*.*.*

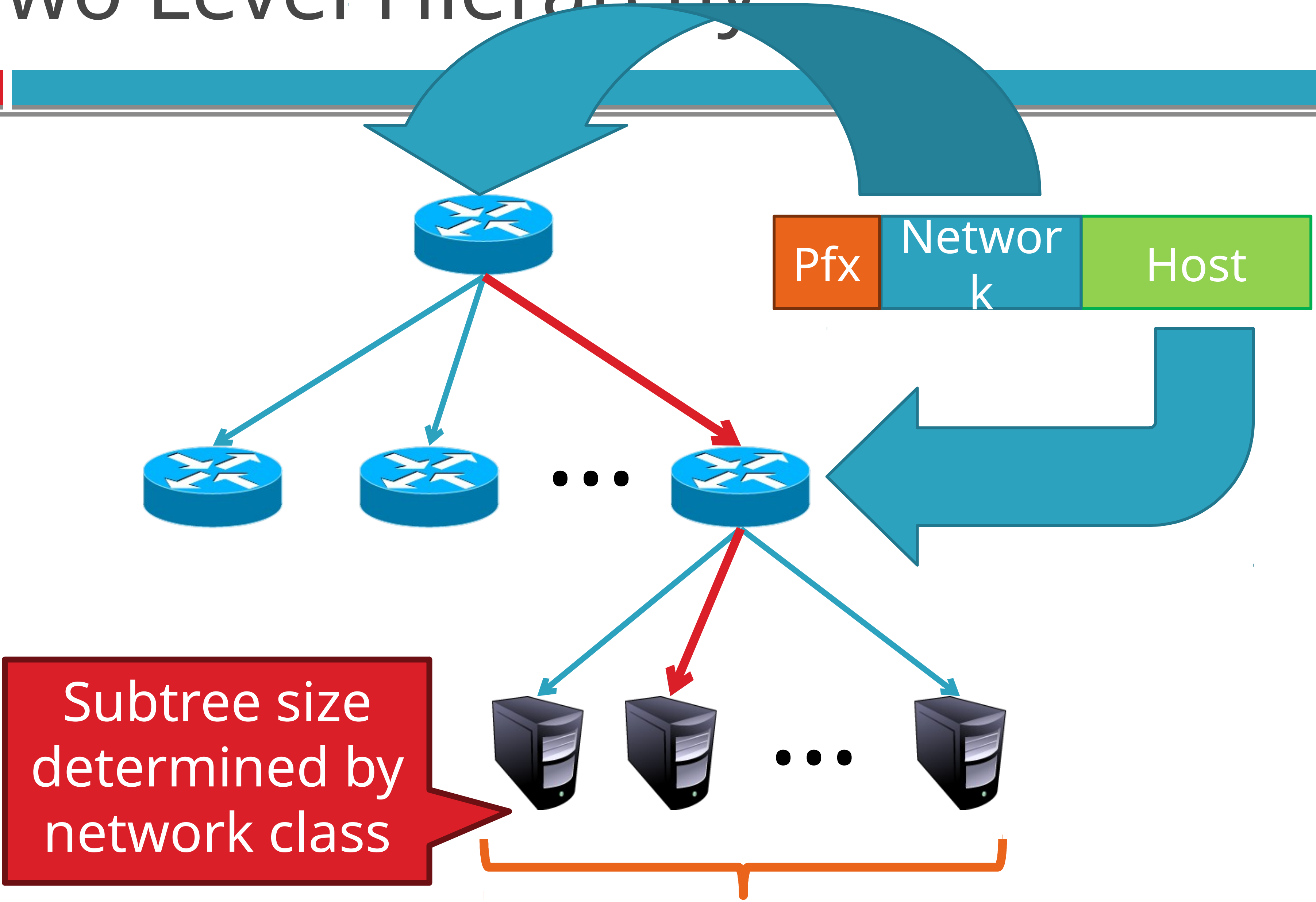
Example: NEU
129.10.*.*

Example:
216.63.78.*

How Do You Get IPs?

- IP address ranges controlled by IANA
  - Internet Assigned Number Authority
  - Roots go back to 1972, ARPANET, UCLA
  - Today, part of ICANN
- IANA grants IPs to regional authorities
  - ARIN (American Registry of Internet Numbers) may grant you a range of IPs
  - You may then advertise routes to your new IP range
  - There are now secondary markets, auctions, …

Two Level Hierarchy



Class Sizes

Way too big

Class	Prefix Bits	Network Bits	Number of Classes	Hosts per Class
A	1	7	$2^7 - 2 = 126$ (0 and 127 are reserved)	$2^{24} - 2 = 16,777,214$ (All 0 and all 1 are reserved)
B	2	14	$2^{14} - 16,398$	$2^{16} - 2 = 65,534$ (All 0 and all 1 are reserved)
C	3	21	$2^{21} = 2,097,512$	$2^8 - 2 =$ 4 (All 0 all 1 are reserved)
			036	

Too many network IDs

Too small to be useful

Subnets

- Problem: need to break up large A and B classes
- Solution: add another layer to the hierarchy
  - From the outside, appears to be a single network
    - Only 1 entry in routing tables
  - Internally, manage multiple subnetworks
    - Split the address range using a subnet mask

Subnet Mask:11111111 11111111 11000000 00000000

Subnet Example

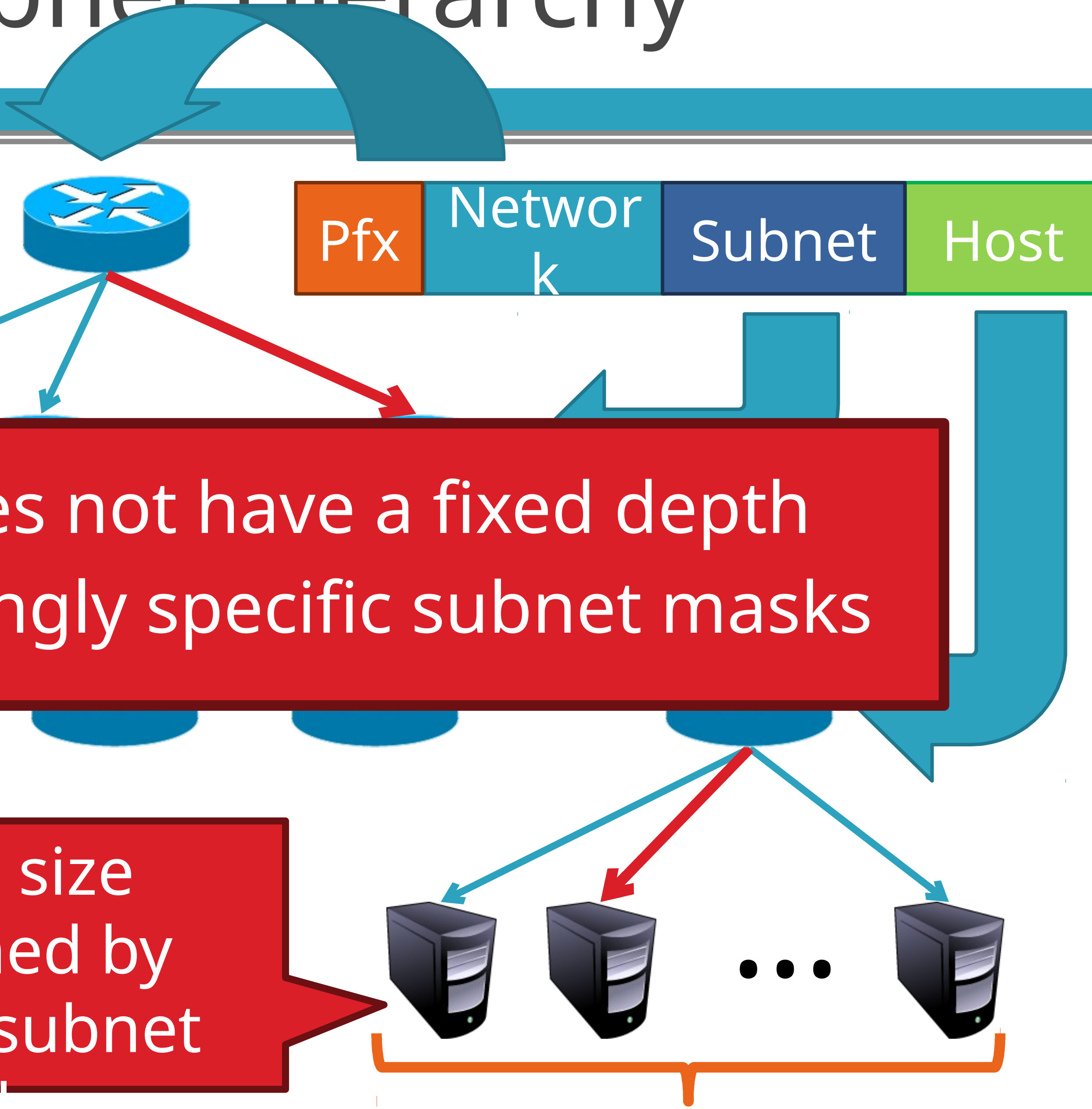
- Extract network:

IP Address:	10110101 11011101 01010100 01110010
Subnet Mask:	& 11111111 11111111 11000000 00000000
Result:	10110101 11011101 01000000 00000000

- Extract host:

IP Address:	10110101 11011101 01010100 01110010
Subnet Mask:	& ~(11111111 11111111 11000000 00000000)
Result:	00000000 00000000 00010100 01110010

N-Level Subnet Hierarchy



- Tree does not have a fixed depth
- Increasingly specific subnet masks

Subtree size determined by length of subnet

Example Routing Table

Address Pattern	Subnet Mask	Destination Router
0.0.0.0	0.0.0.0	Router 4
18.0.0.0	255.0.0.0	Router 2
128.42.0.0	255.255.0.0	Router 3
128.42.128.0	255.255.128.0	Router 5
128.42.222.0	255.255.255.0	Router 1

- Question: 128.42.222.198 matches four rows
  - Which router do we forward to?
- Longest prefix matching
  - Use the row with the longest number of 1's in the mask
  - This is the most specific match

Subnetting Revisited

- Question: does subnetting solve all the problems of class-based routing?

NO

- Classes are still too coarse
  - Class A can be subnetted, but only 126 available
  - Class C is too small
  - Class B is nice, but there are only 16,398 available
- Routing tables are still too big
  - 2.1 million entries per router

Classless Inter Domain Routing

- CIDR, pronounced ‘cider’
- Key ideas:
  - Get rid of IP classes
  - Use bitmasks for all levels of routing
  - **Aggregation** to minimize FIB (forwarding information base)
- Arbitrary split between network and host
  - Specified as a bitmask or prefix length
  - Example: Stony Brook
    - 130.245.0.0 with netmask 255.255.0.0
    - 130.245.0.0 / 16

Aggregation with CIDR

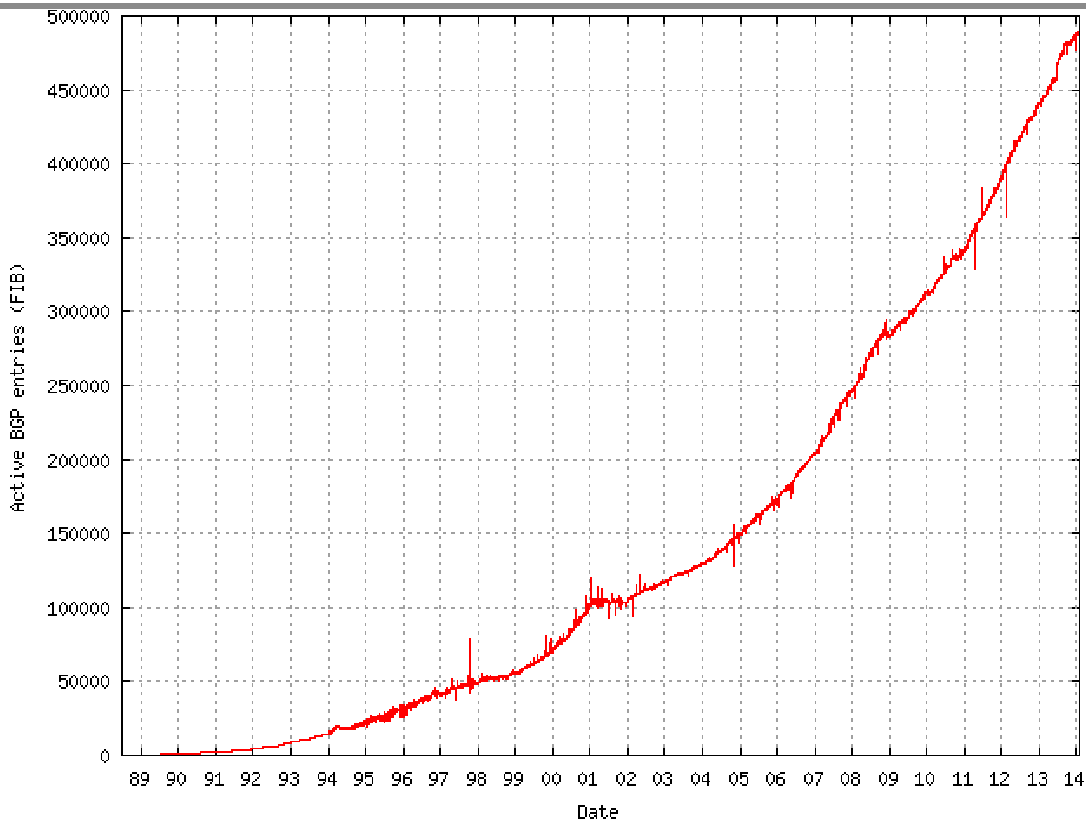
- Original use: aggregating class C ranges
- One organization given contiguous class C ranges
  - Example: Microsoft, 207.46.192.* – 207.46.255.*
  - Represents 2^6 = 64 class C ranges
  - Specified as CIDR address 207.46.192.0/18

	0	8	16	24 31
Decimal	207	46	192	0
Hex	CF	2E	C0	00
Binary	11001111	00101110	11xxxxxx	xxxxxxxx

18 Bits Frozen By Netmask

14 Arbitrary Bits

Size of CIDR Routing Tables



- From www.cidr-report.org
- CIDR has kept IP routing table sizes in check
  - Currently ~500,000 entries for a complete IP routing table
  - Only required by backbone routers

We had a special day in summer 2014!

- 512K day – August 12, 2014
- Default threshold size for IPv4 route data in older Cisco routers 512K routes
  - Some routers failed over to slower memory
    - RAM vs. CAM (content addressable memory)
  - Some routes dropped
- Cisco issues update in May anticipating this issue
  - Reallocated some IPv6 space for IPv4 routes
- Part of the cause
  - Growth in emerging markets
- http://cacm.acm.org/news/178293-internet-routing-failures-bring-architecture-changes-back-to-the-table/fulltext

Takeaways

- Hierarchical addressing is critical for scalability
  - Not all routers need all information
  - Limited number of routers need to know about changes
- Non-uniform hierarchy useful for heterogeneous networks
  - Class-based addressing is too course
  - CIDR improves scalability and granularity
- Implementation challenges
  - Longest prefix matching is more difficult than schemes with no ambiguity