Computer Networks

EDA387/DIT663

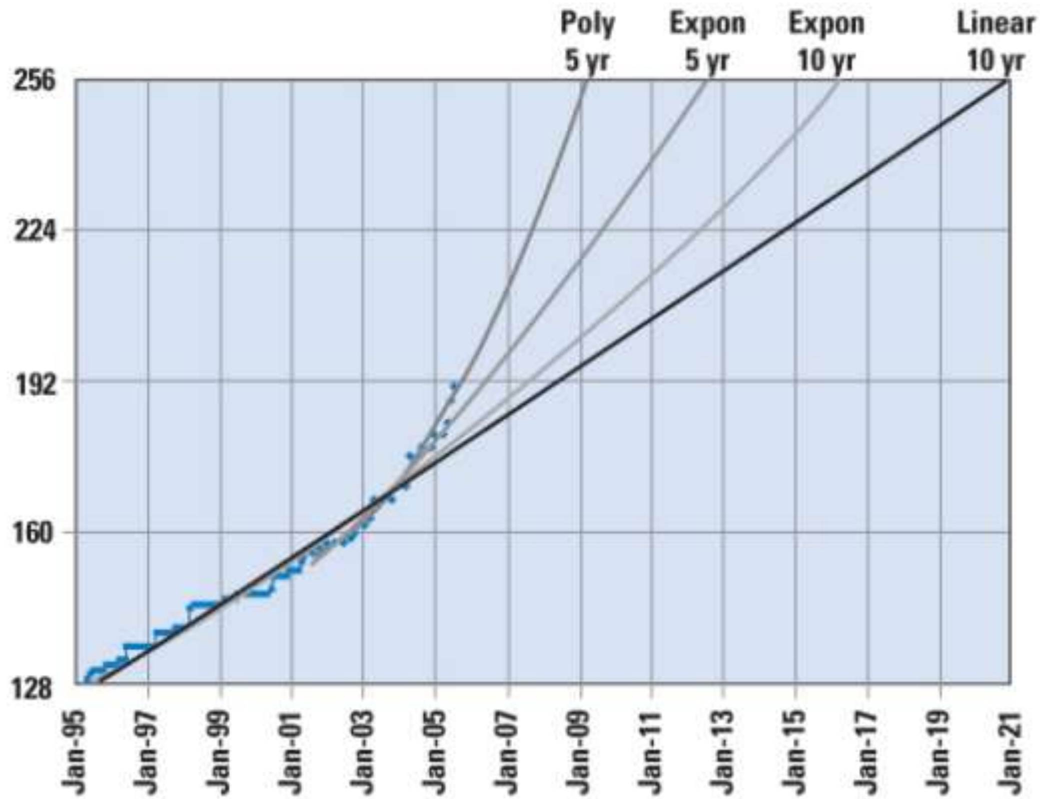
IP variations: IPv6, multicast, anycast

Many slides are barrowed from Philip Levis and David Mazieres and some slides are barrowed from Ali Salehson

Outline

- Next generation IP: IPv6
- IP multicast
- IP anycast
- Practical considerations throughout

I think we have a problem



- Projected use of /8 blocks
- From “A Pragmatic Report on IPv4 Address Space Consumption,” Tony Main, Cisco Systems.

IPv6

- Work started in 1994
- Basic protocol published in 1998 (RFC 2460)
- Brief lull, the progress in 2003-6
- Hard push within IETF today for adoption

FEATURES —

The night the IETF turned off IPv4

At a meeting this week, the Internet Engineering Task Force decided to kill ...

ILJITSCH VAN BEIJNUM - 3/14/2008, 2:05 AM

Taking a detour from IPv4

After working on the new Internet Protocol version 6 (IPv6) for a decade and a half, the Internet Engineering Task Force decided it was time to turn off the old protocol (IPv4 or just IP). So this is what they did for an hour on the network used at the IETF meeting in Philadelphia this week. Network traffic plummeted from some 30Mbps to around 3Mbps as the meeting attendees who had IPv6 enabled could now only get at IPv6-reachable destinations on the Internet. Leslie Daigle, chief Internet technology officer for the Internet Society, who coordinated the IPv4 outage, considers the outage a success.

"The plenary outage provided a focal point for a number of people to dive into and work with IPv6—whether on their own notebooks, or their home web sites (e.g., Paul Hoffman's www.vpnc.org), or their products (ISC's BIND), or public services, such as Google," said Daigle. "From that perspective, the experiment was a success before we even turned off IPv4 access in the plenary session. I hope, though I don't know, that the plenary event also provided a first hand experience to more IETF engineers that they can take into Working Group discussions."

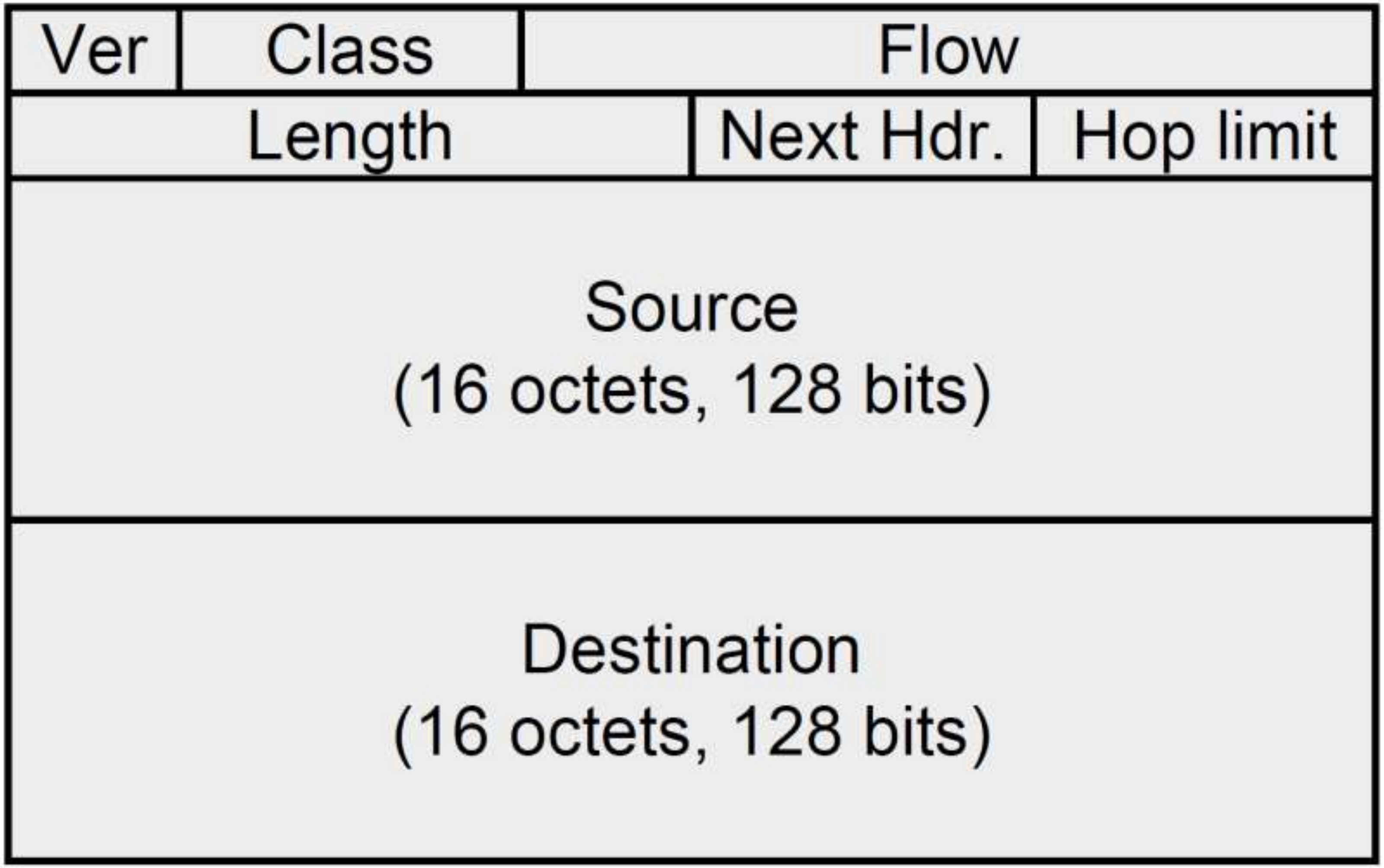
Shortly after the fall meeting, the IETF leadership decided to create an "IPv4 outage" during this week's spring meeting. The triannual IETF meetings attracted some 1,200 Internet engineers from

https://arstechnica.com/features/2008/03/ietf-ipv6-switchoff/

IPv6 Key Features

- 128 bit addresses
  - Autoconfiguration
- Simplifies basic packet format through extension headers
  - 40 byte “base” header
  - Make uncommonly used fields optional
- Security and authentication

IPv6 Header


Ver
Class
Flow
Length
Next Hdr.
Hop limit
Source
(16 octets, 128 bits)
Destination
(16 octets, 128 bits)

IPv6 Header Fields

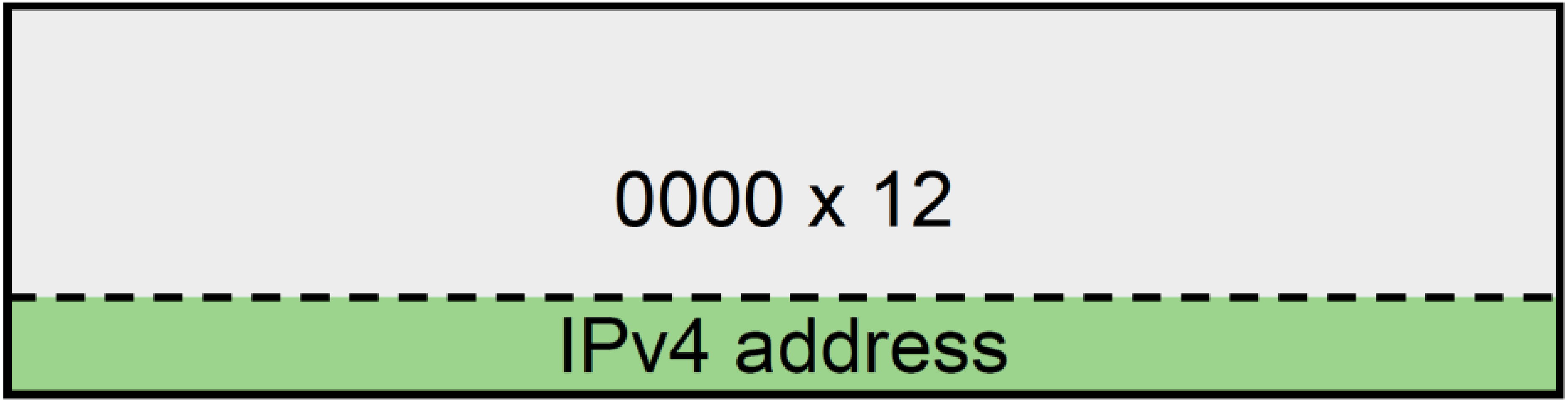
- Version, 4 bits: 6 for IPv6
- Class: 8 bits: like type of service in IPv4
- Flow, 20 bits: identifies a flow
- Length, 16 bits: datagram length
- Next header, 8 bits: more later
- Hop limit, 8 bits: like TTL in IPv4
- Addresses: 128 bits

IPv6 Addresses

- Simplify DHCP and autoconfiguration
- Break 128 bits into 80-bit network and 48-bit interface
- Many link layers have unique interface addresses (more on this later in quarter)
  - E.g., Ethernet is 48 bits
  - Use of 48-bit ID ensures no address collisions, makes DHCP stateless

v4 Interoperability

- RFC 4291
- Every IPv4 address has an associated IPv6 address
- Simply prefix 32-bit IPv4 address with 96 bits of 0



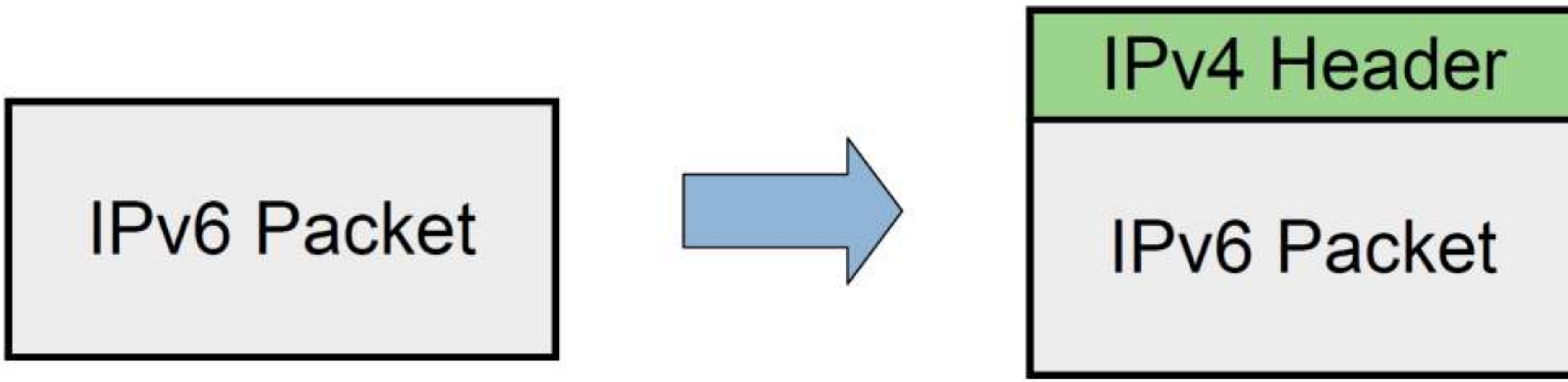
v4 Interoperability, continued

- Two IPv6 endpoints must both have IPv6 stacks
- What about transit network?
- v6 - v6 - v6 (no problem)
- v4 - v4 - v4 (no problem)
- v4 - v6 - v4 (no problem)
- v6 - v4 - v6 (uh-oh)

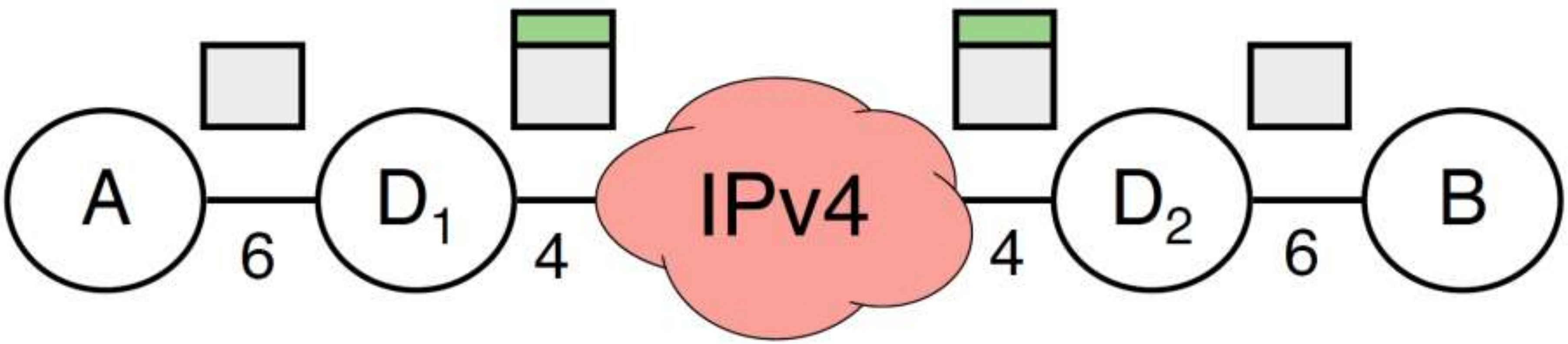
6-4-6 Example

IP Tunneling

- Encapsulate an IP packet inside another IP packet
- Makes an end-to-end path look like a single IP hop



6-4-6 Example, Revisited



Other Tunneling Use: VPN

- Virtual Private Networks
- Use case: two distance corporate offices
  - Want to access each other’s internal networks
  - Make it looks like they’re the actually one network
- Set up an encrypted TCP stream between one host at each network
- Route packets to other office through this host
- If addresses are all private, network is private

MTU Requirement

- IPv4 requires a 576-byte link maximum transmission unit
- IPv6 requires 1280-byte MTU
- If link MTU is smaller, then it MUST support sub-IP fragmentation and assembly to provide a 1280-byte MTU
- It SHOULD provide a 1500-byte MTU; nodes MUST receive 1500 byte packets

Fragmentation Revisited

- High-loss links (e.g., wireless) can be a problem 10-hop route, each link has a 10% drop rate (90% success rate)
  - Probability one fragment arrives is $0.9^{10} \approx 35\%$
  - Each fragment is transmitted

    $1+0.9+0.9^2+0.9^3\ldots0.9^9\approx6.5$ times along the route
  - 100% chance on first hop, 90% on second hop, 81% on third hop, etc.

Fragmentation Revisited, Continued

- Delivery probability of 4 fragment packets: $0{:}35^4 \approx 1.4\%$
- Total transmissions/delivery = $1/0.014 \cdot \Sigma^{9}_{i=0}\ 0.9^{i}$
- Total transmissions/delivery = $65 \cdot 6.5 = 423$
- Fragmentation header in IPv6 is a destination header
  - Fragmentation is possible, but must be done at the source

Link-layer reliability

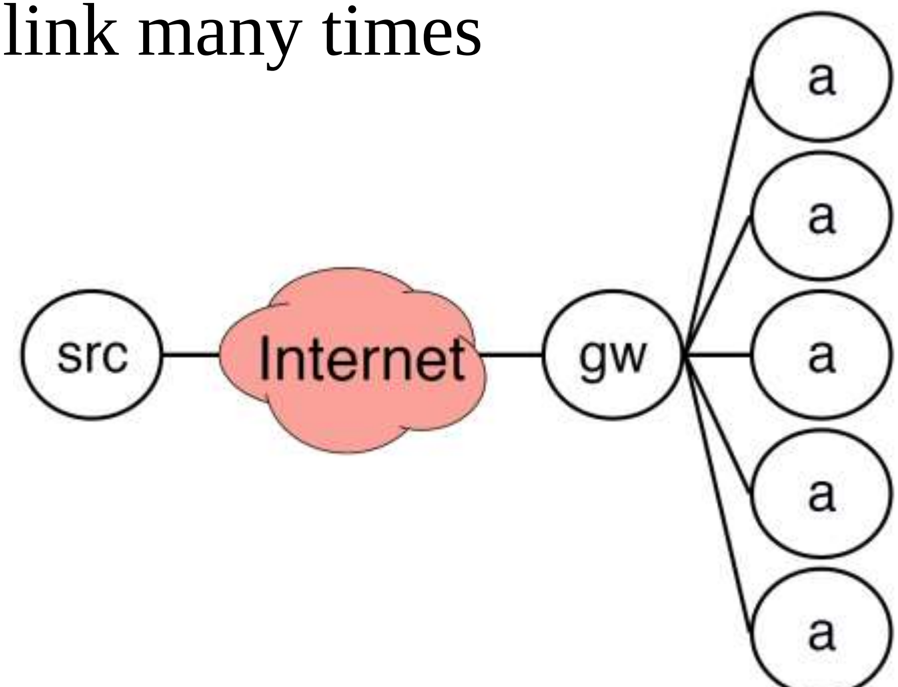
- High-loss link layers usually have single-hop acks and retransmissions
  - End-to-end argument: when can layer 2 reliability fail end-to-end?
- 10-hop route, each link has a 10% drop rate
  - Expect 1/0.9 ≈1.1 transmissions/link
  - 10 links, 11 transmissions
  - 44 transmissions/delivery

Flaw in the Argument

- Original IPv6 motivation was “IPv4 addresses will run out”
- Addresses are a resource; they have a value (you don’t run out of land)
- NATs allow multiple nodes to share an IPv4 address
- IPv6 will become the default when IPv4 addresses are so expensive that it’s cheaper to deploy IPv6
- IETF T-shirt: 32 + 16 > 128

Multicast

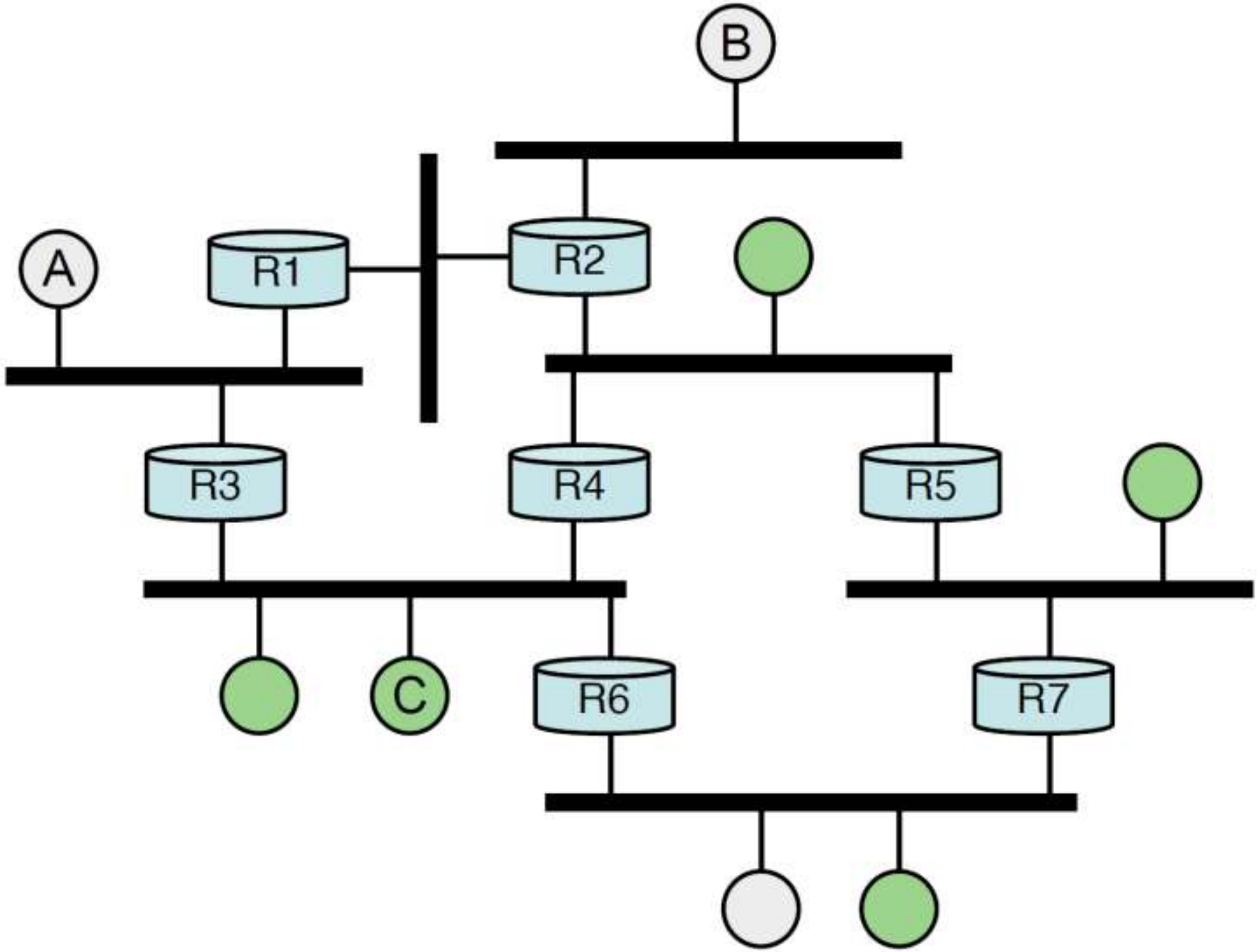
- Problem: want to send a packet to many nodes
  - Examples: IP-TV, large audio stream
- Using *n* unicast packets means the same packet can traverse a single link many times



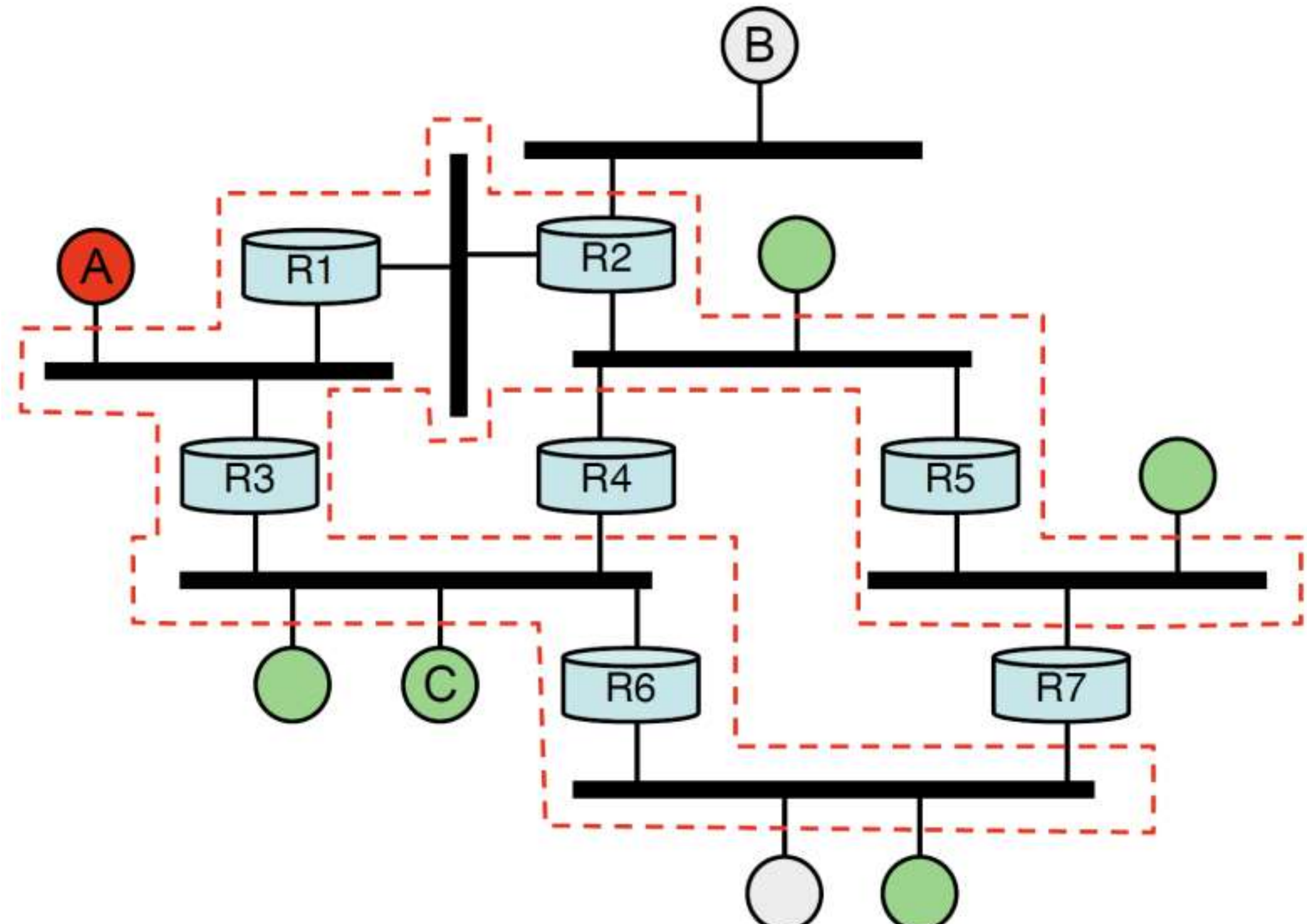
Multicast Approach

- Nodes can join a multicast group
- Denoted by a multicast IP address
- Routers build a routing topology
  - Link state vs. distance vector
- IGMP: Internet Group Management Protocol
  - Protocol for hosts to manage membership in multicast groups
  - Hosts talk to local multicast routers

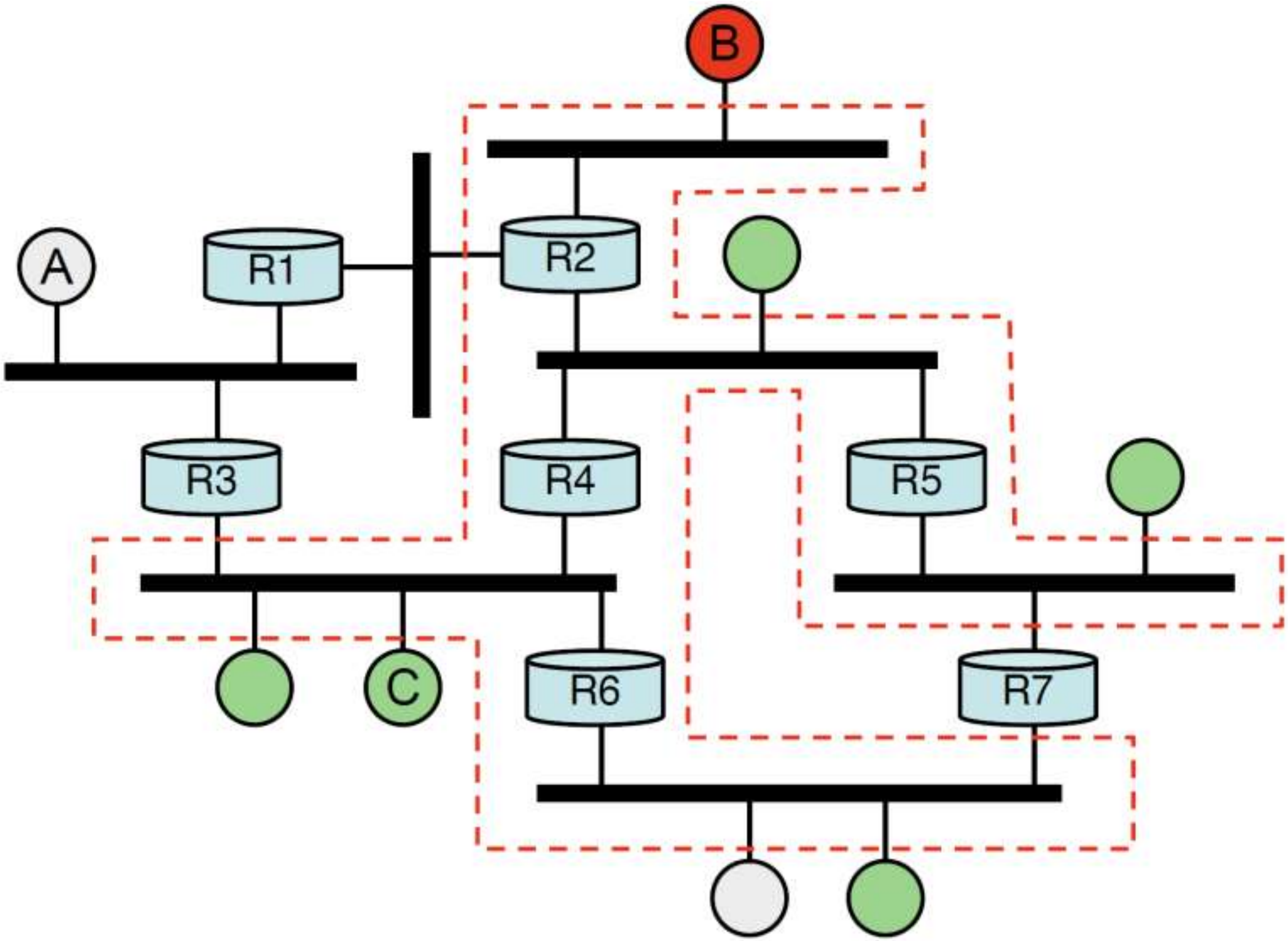
Network Topology



Tree for A as Multicast Source



Tree for B as Multicast Source



Practical considerations

- Multicast protocols end up being very complex
- Introduce a lot of router state
- Turned off on most routers
- Used within a domain, not between domains
- How does one handle congestion control?

Anycast

- Communicate with “any” one of a set of nodes
- We’ve seen this with DNS

```
\$ dig www.google.com
...
;; ANSWER SECTION:
www.google.com. 604799 IN CNAME www.l.google.com.
www.l.google.com. 300 IN A 74.125.19.103
www.l.google.com. 300 IN A 74.125.19.104
www.l.google.com. 300 IN A 74.125.19.147
www.l.google.com. 300 IN A 74.125.19.99
```

Anycast at IP layer

- DNS allows anycast through name → address mappings
- Sometimes we need it at layer 3 itself
  - Single IP address refers to multiple hosts
  - Need to talk to any one of them
- Example: DNS root servers
  - Would like to scale number of root servers with Internet
  - Can't use DNS (remember root servers hard-coded)
  - Want to query closest root server

Anycast in Forwarding Tables

- Remember, forwarding is longest-prefix-match
- An anycast address is a /32 address
- A single router may have multiple entries for the address
- Anycast best used in services where separate packets might go to different destinations

The Cost

- A /32 routing entry!
- Multiple /32 routing entries!

Further Advantages

- Geographic scoping
- Distributed Denial of Service (DDoS)
  - Since anycast is at IP layer, load from DDoS is distributed across many anycast nodes
- F root server made anycast in 2002, now 12 locations

Summary

- Next generation IP: IPv6
- IP multicast
- IP anycast
- Practical considerations throughout
- Next lecture: review
- After midterm: DCCP and NATs