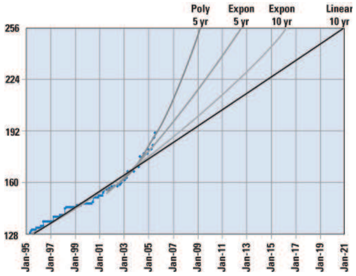
Lecture 16: IP variations: IPv6, multicast, anycast

Overview

- Next generation IP: IPv6
- 6lowpan and the Internet of Things
- IP multicast
- IP anycast
- Practical considerations throughout

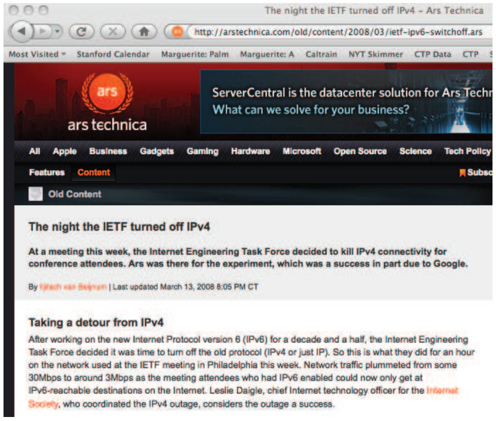
I think we have a problem



- 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



The night the IETF turned off IPv4

At a meeting this week, the Internet Engineering Task Force decided to kill IPv4 connectivity for conference attendees. Ars was there for the experiment, which was a success in part due to Google.

By [illegible] | Last updated March 13, 2008 8:05 PM CT

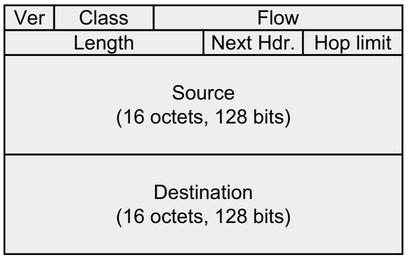
Taking a detour from IPv4

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.

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



IPv6 Header Fields

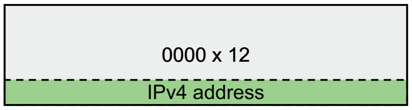
- Version, 4 bits: 6 for IPv6
- Class: 8 bits: like TOSS in IPv4
- Flow, 20 bits: identifies a flow
- Length, 16 bits: datagram length
- Next header, 8 bits: more later
- Hop limit, 8 bits: like TTL in IPv4
- Addresses: 128 bits

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

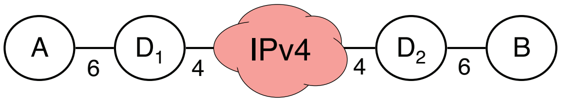
- RFC 4291
- Every IPv4 address has an associated IPv6 address
- Simply prefix 32-bit IPv4 address with 96 bits of 0



v4 Interoperability, continued

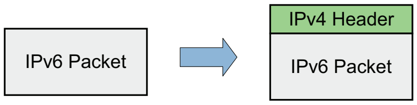
- Two IPv6 endpoints must both have IPv6 stacks
- What about transit network?
- v6 - v6 - v6 (no problem)
- v4 - v4 - v4 (no problem)
- v4 - v6 - v4 (no problem)
- v6 - v4 - v6 (uh-oh)

6-4-6 Example

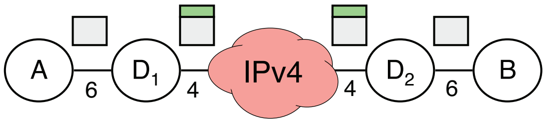


IP Tunneling

- 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

Extension Headers

- Two types: destination and hop-by-hop
- Both have a next header byte
- Destination headers: intended for IP endpoint
  - Fragment header
  - Routing header (loose source routing)
- Hop-by-hop headers: processed by each node
  - Jumbogram: packet is up to 2^{32} bytes long

Example Next Header Values

- 0: Hop-by-hop header
- 1: ICMPv4
- 4: IPv4
- 6: TCP
- 17: UDP
- 41: IPv6
- 43: Routing header
- 44: Fragmentation header
- 58: ICMPv6

MTU Requirement

- IPv4 requires a 576-byte link MTU
- 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 $1 + 0.9 + 0.9^2 + 0.9^3 ... 0.9^9 \approx 6.5$ times along the route
  - 100% chance on first hop, 90% on second hop, 81% on third hop, etc.

Fragmentation Revisited, Continued

- **If a packet has four fragments, delivery probability is** $0.35^4 \approx 1.4\%$
- **Total transmissions/delivery =** $\frac{1}{0.014} \cdot \sum_{i=0}^{9} 0.9^i$
- **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 $\frac{1}{0.9} \approx 1.1$ transmissions/link
  - 10 links, 11 transmissions
  - 44 transmissions/delivery

Practical Considerations

- **IPv6 is only partially deployed**
- **No "killer app"**
- **It's *really* expensive to replace everything!**
- **No switch day: IPv4 will always live on**
- **Multihoming and address fragmentation is causing routing tables to grow very large**
  - IPv6 will make this much, much worse...

IPv4 Status

RIRs

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:** $32 + 16 > 128$

A Market in Addresses

- **A market won't solve the problem, though**
- **IPv4 addresses can't be legally owned in some regions**
- **Trading will further fragment the address space**

6lowpan and the Internet of Things

Internet of Things

- **Increasing connectivity: wireless controllers, light switches, etc.**
- **Home area networks, personal area networks**
- **Today: vertically integrated, separate technologies**
- **Goal: connect them with IP**
- **Imagine every light has an IP address...**

6lowpan

- **IETF working group on IPv6 for low-power personal area networks (PANs)**
- **Tiny, energy constrained, wireless devices: smart homes, ubiquitous computing**
- **Link layers have tiny MTUs: (802.15.4 is 127 bytes)**
- **RFC 4944**

6lowpan Header Compression

- **6lowpan tries to compress common cases: TCP, UDP, etc.**
- **Example: address compression**
  - 6lowpan must allow full 128-bit addresses
  - Address fields alone are 32 bytes!
  - But often they can be shortened...

6lowpan Header Compression

- **draft-ietf-6lowpan-hc-13 (updates RFC 4944)**

```
  0   1   2   3   4   5   6   7   8   9   0   1   2   3   4   5
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
| 0 | 1 | 1 |  TF   |NH | HLIM  |CID|SAC|  SAM  | M |DAC|  DAM  |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
```

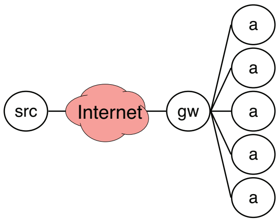
6lowpan Compression Flags

- **SAC: Source address (stateful?)**
- **DAC: Destination address (stateful?)**
- **SAM/DAM: compression scheme used, for stateless:**
  - 00: Full 128 bit address
  - 01: 64-bit address, other 64 are link-local prefix padded with zeros
  - 10: 16-bit address, other 112 are as above
  - 00: 0-bit address, 64-bit link local prefix + 64-bit link layer address

2-minute stretch

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**



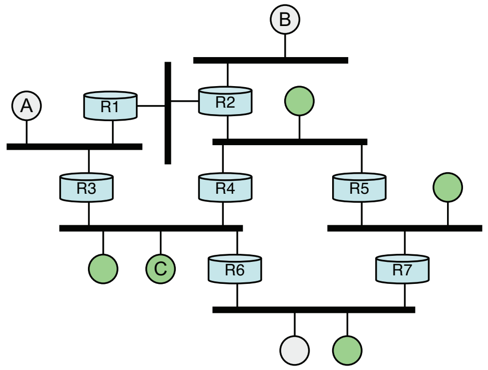
Multicast Approach

- **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

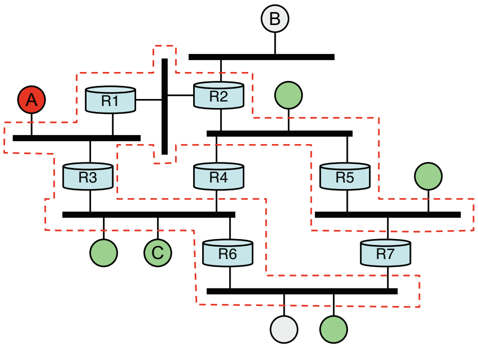
Example: Link State Tree

- **Routers exchange link state**
- **Node advertise presence in group**
- **Routers compute shortest-path multicast tree**
- **Very expensive!**

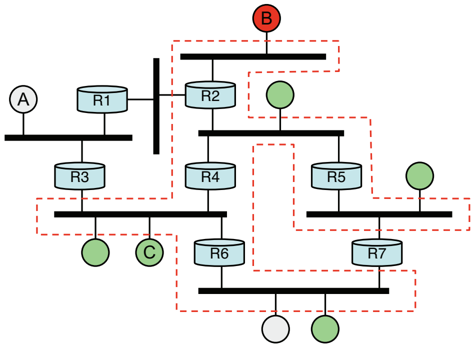
Network Topology



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?

IP Anycast

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 Tablse

- 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

Overview

- Next generation IP: IPv6
- 6lowpan and the Internet of Things
- IP multicast
- IP anycast
- Practical considerations throughout