E6998-02: Internet Routing

Lecture 2
Addressing

John Ioannidis
AT&T Labs – Research
ji+ir@cs.columbia.edu

Announcements

Email: <ji+ir@cs.columbia.edu>

- Mail to anything else will not be answered.

Class web page: http://www.cs.columbia.edu/~ji/F03/

- Check frequently!
- Slides will be available there.
- As will additional reading material (papers, RFCs, source code, man pages, etc.).

Class BBoard: coms6998-002-033@columbia.edu (to post), or https://www1.columbia.edu/sec/bboard/033/coms6998-002/

Office hours: MW 15:00-16:00 in 464 CSC.

TA(s): Angelos Stavrou <angel@cs.columbia.edu>

TA office hours: TR 13:00-14:00

Names, Routes, Addresses

- Common concepts, frequently confused.
- Bit strings referring to nodes, interfaces, or other network elements.
- **Name**: location-independent.
  - **Identifier**.
- **Address**: depends on its location.
  - Shows place in the topology.
- **Route**: depends on both its own and others' location.
  - Directions on how to get there.

- Same entity may have different function around the stack.

- Protocols exist to map between them.

LAN Addresses

- MAC-layer addresses.
- Do not depend on the location in the topology.

  but
- They are used to deliver packets.

- LAN layer does not need address structure.

- Historical reasons for LAN addresses.
- (afterthought) homogeneity with NBMA networks.

Properties of Network Addresses

- Carried in every packet?

- Structure reflecting topology?
  - The need for hierarchical routing.
  - The need for structured addresses.
- Fixed or variable hierarchy boundaries?
- Fixed or variable size addresses?

- Refer to nodes or interfaces?
- Unique?
- One per addressed entity?

- Configuration/acquisition?
- Mapping to/from lower- and higher- layer "addresses".

Structure and Topology

- Flat addressing works for subnet (LAN) addressing.
  - Each node recognizes its own address and picks up packet.
  - <*node*>
- Flat networks don't scale.
  - Break network up into subnets, link them.
  - Designate one node as router.
  - If packet not for subnet, router picks it up (how?).
  - <*subnet*, *node-within-subnet*>
  - Flat routing between subnets.
- If number of subnets becomes large, recurse:
  - <*level-2-subnet*, *level-1-subnet*, *node-within-subnet*>.
- …

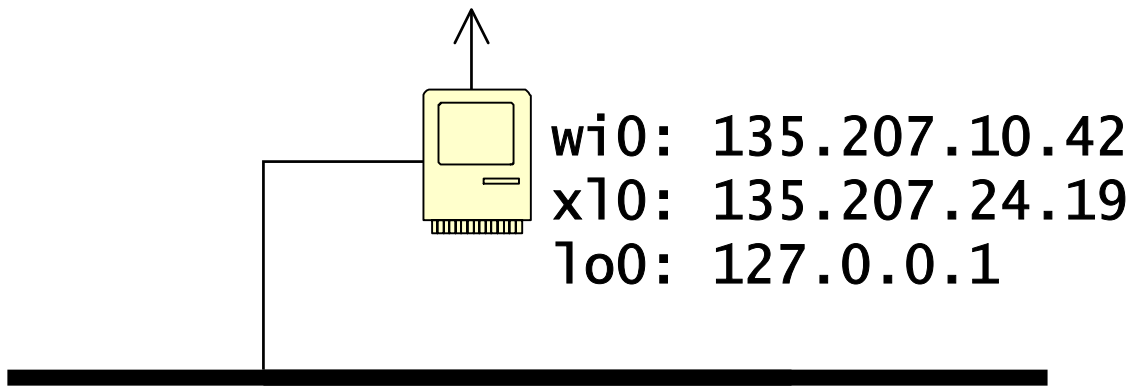
- Hierarchy boundaries can be fixed
  - Faster to process.
  - More wasteful of address bits.
- Address size can be fixed.
  - Faster to process.
  - Will run out!

- Address aggregation.

What does an address label?

- Node or interface?
- Multiple addresses for the same interface?
- Implications for routing.
- Implications for source address selection.
- Implications for packet acceptance.

- Virtual/loopback interfaces?



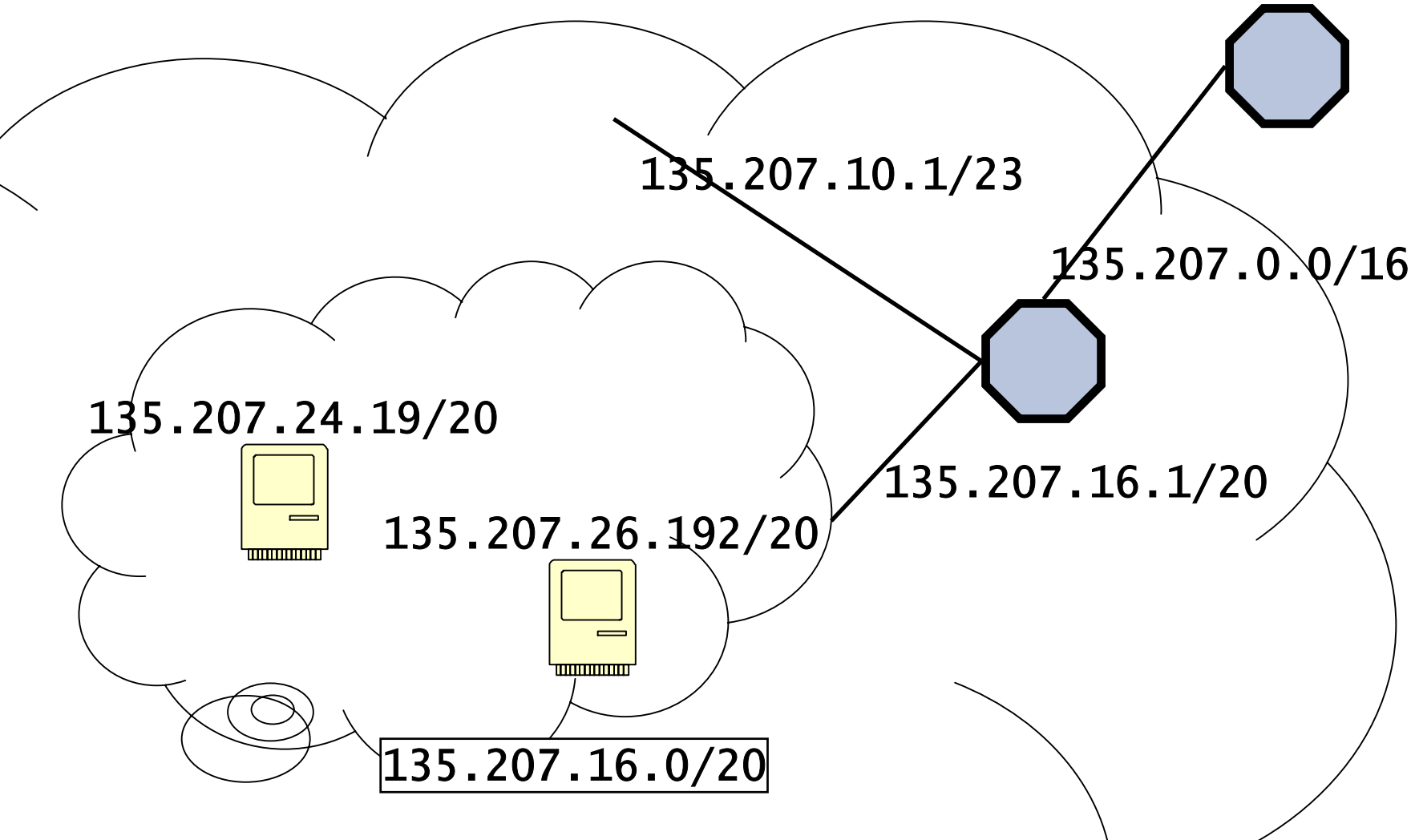
Hierarchical Routing in the Postal Service

John Ioannidis
AT&T Labs – Research
180 Park Avenue, Building 103, Room A-283
Florham Park, NJ 07932
USA

IP Addresses

- Fixed-size, 32-bit addresses.
- Variable hierarchy boundaries: <*subnet-n*, …, *subnet-1*, *node*>
- At each level, it's <*network*, *host*>
  - Only one boundary "visible".
  - Boundary defined by netmask (prefix length).
  - 135.207.24.19/255.255.240.0 (or /20)
  - Network part: 135.207.16.0
  - Host part: 0.0.8.19
  - All hosts in 135.207.16.0/20 designate a router.
- Addresses name interfaces.
- Each interface can have multiple addresses.

- Special addresses ({any,multi,broad}cast) may refer to multiple nodes.


135.207.10.1/23
135.207.0.0/16
135.207.24.19/20
135.207.16.1/20
135.207.26.192/20
135.207.16.0/20

Address Classes (old stuff)

- First few bits determine "address class" and net-host boundaries:
  - 0… (0.0.0.0-127.255.255.255): Class A, 8-bit nets.
  - 10… (128.0.0.0-191.255.255.255): Class B, 16-bit nets.
  - 110… (192.0.0.0-223.255.255.255): Class C, 24-bit nets.
  - 1110… (224.0.0.0-239.255.255.255): Class D, multicast.
  - 1111… (240.0.0.0-255.255.255.255): Class E, reserved.
- Classes A/B/C:
  - Wasteful allocation of address space (not enough Class A/Bs).
  - Routing table explosion (too many "Class Cs").

Old Address Allocation Plan

- An organization gets a Class A, B, or C depending on its expected size.
  - ARPA=10, ATT=12, MIT=18, UCB=32, Stanford=36…
  - Columbia=128.59, Bellcore=128.96
  - GIP-Altaïr=192.27.52
- Turns out Class Cs were too small, Class B and A way too large.
- ~1992, no one was getting any more Class As, and we were afraid we were soon going to run out of Class Bs.
- People hated Class Cs because they had to get several of them, and that increased the size of the routing tables.

Subnetting

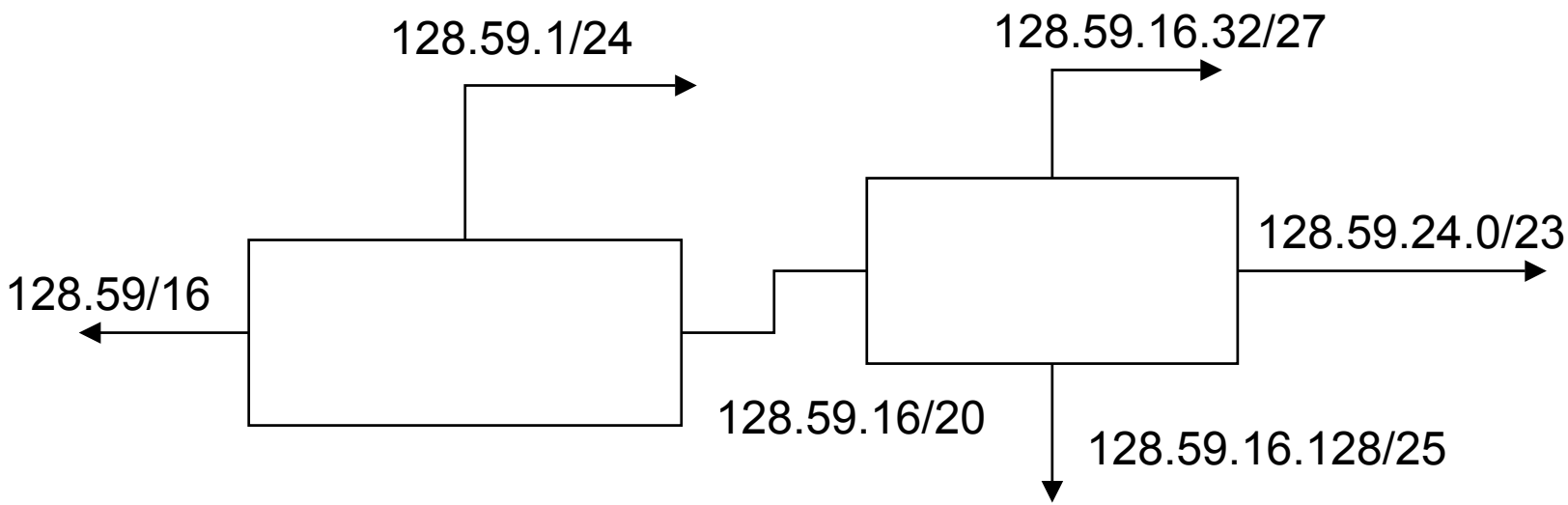
- Breaking up the "host" part of the address into subnet-host.
  - 2 levels of hierarchy good, more are better!
- Bridging too expensive even within a site.
- Originally, along 8-bit boundaries.
- Netmask: shows which part of address is "net":
  - 128.59.16.0/255.255.240.0
  - 10000000 00111011 00010000 00000000
  - 11111111 11111111 11110000 00000000
  - 128.59.16.0/20
- Allows for address aggregation within a site:
  - Access router handles 128.59.0.0/16.
  - Various routers around campus handle subnets.
  - Originally, all subnetted interfaces had to have the same netmask.

Subnetting, Cont'd

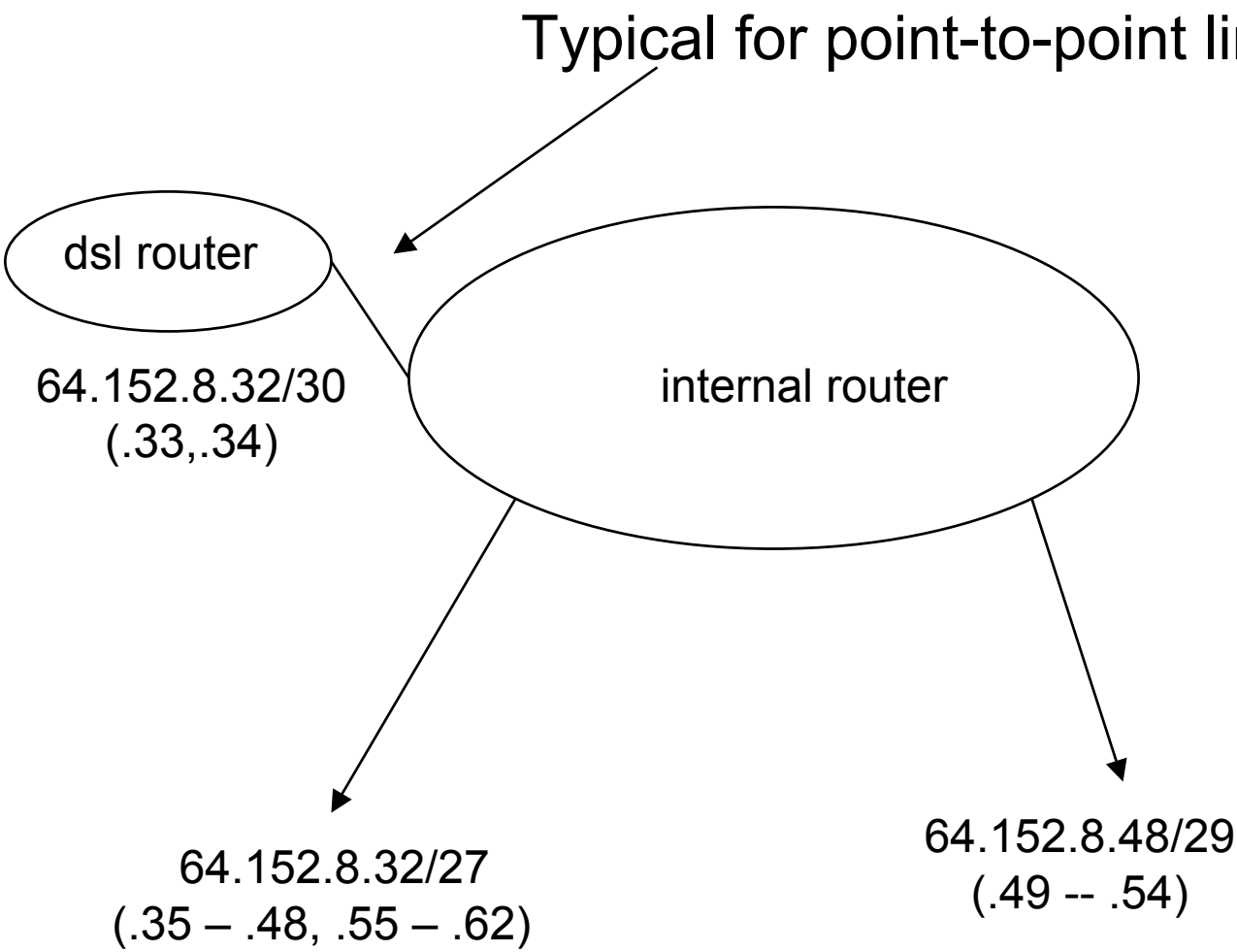
- Host-within-subnet part all zeros → anycast.
  - 135.207.4.64/26
- Host-within-subnet part all ones → directed broadcast.
  - 192.4.13.127/25
- Whether an address is anycast/broadcast or not depends on the receiving router.
  - Router for 135.207.0.0/16 just routes 135.207.4.64.
  - As does router for 135.207.0.0/20.
  - Router for 135.207.4.64/26 considers it an anycast.

VLSM

- Obvious thing to do.
- Different subnets have different size requirements.
- No need to subnet at the longest prefix.
- Forward according to "longest-match".



VLSM can have overlapping allocations



Classless Interdomain Routing (CIDR)

- "Supernetting" (opposite of subnetting).
- Get rid of classes A/B/C.
- Give addresses in terms of prefixes.
- Netmask MUST have contiguous 1s, then contiguous 0s.
- Allows sites to be allocated the proper size of network.
- Allows ISPs to aggregate addresses of clients.
  - Reducing routing table size.
- "CIDR block" or "CIDR prefix".

CIDR Address Allocation

- Pre-CIDR allocations still routed, of course.
- ARIN/RIPE/APNIC have large allocations (/8s) to hand out.
- ISPs get addresses in large blocks from the Registries.
- Allocate chunks of these blocks to customers.
  - "Non-portable" address space: change ISP, change addresses.
  - Aggregation of addresses within ISP.