Part 1: Design Principles

Goals:

- identify, study common architectural principles, protocol mechanisms
- *synthesis:* big picture

Overview:

- Signaling: data/control separation, soft/hard state
- indirection
- randomization
- multiplexing
- virtualization
- design for scale
- end-to-end principle
- implementation techniques

7. Designs for Scale

How to deal with large numbers (millions) of entities in a system?

- IP devices in the internet (0.5 billion)
- Users in P2P network (millions)

Example: routing!

scale: with 200 million destinations:

- can't store all dest's in routing tables!
- routing table exchange would swamp links!

administrative autonomy

- internet = network of networks
- each network admin may want to control routing in its own network

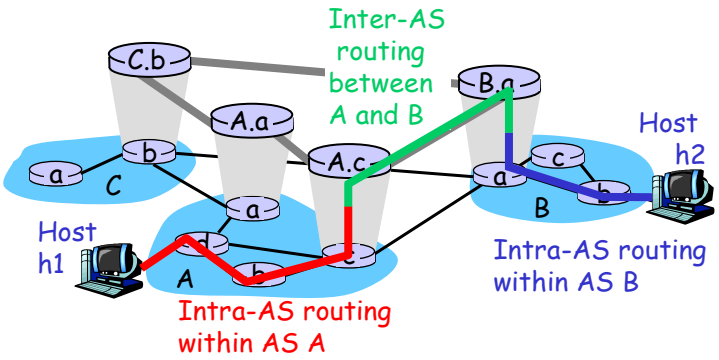
Hierarchical Routing

- aggregate routers into regions, "autonomous systems" (AS)
- routers in same AS run same routing protocol
  - "intra-AS" routing protocol
  - routers in different AS can run different intra-AS routing protocol

gateway routers

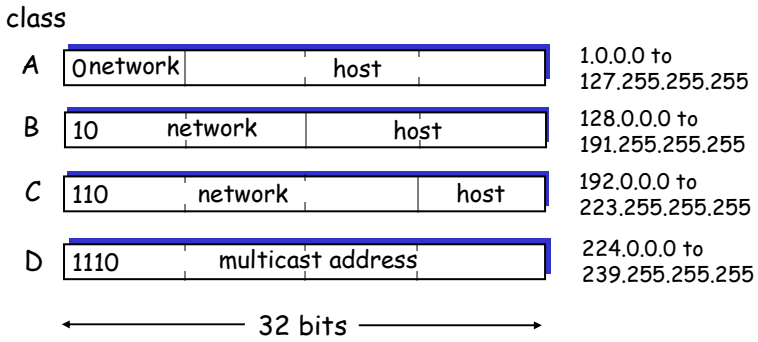
- special routers in AS
- run intra-AS routing protocol with all other routers in AS
- *also* responsible for routing to destinations outside AS
  - run *inter-AS routing* protocol with other gateway routers

Intra-AS and Inter-AS routing



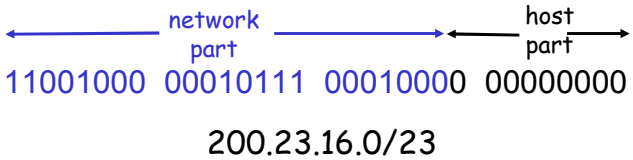
Dealing with scale: addressing

Old-fashioned "class-full" addressing:



IP addressing: CIDR

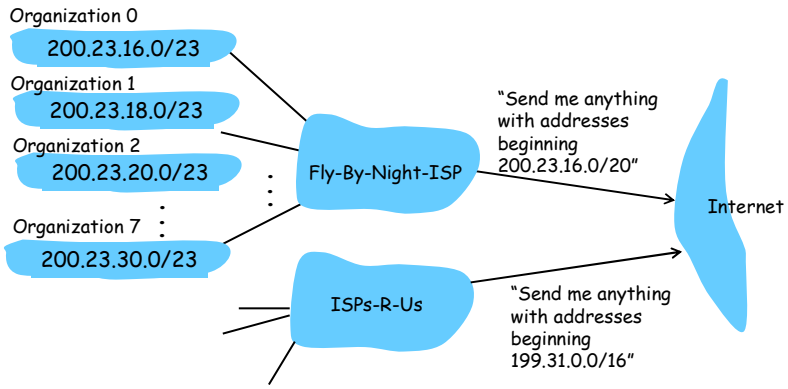
- Classful addressing:
  - inefficient use of address space, address space exhaustion
  - e.g., class B net allocated enough addresses for 65K hosts, even if only 2K hosts in that network
- CIDR: Classless InterDomain Routing
  - network portion of address of arbitrary length
  - address format: a.b.c.d/x, where x is # bits in network portion of address



200.23.16.0/23

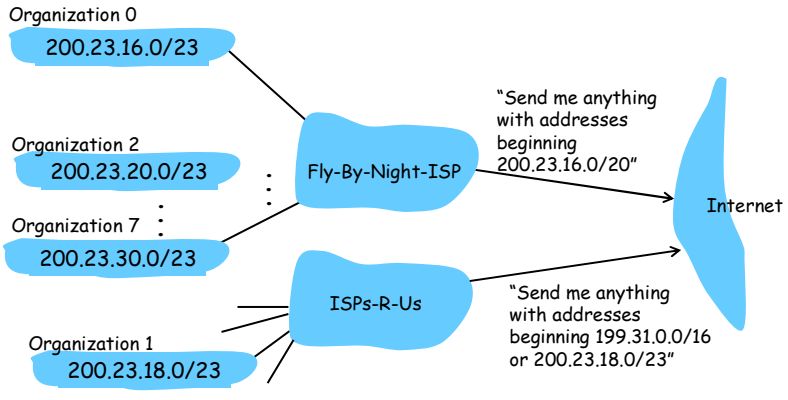
Hierarchical addressing: route aggregation

Hierarchical addressing allows efficient advertisement of routing information:



Hierarchical addressing: more specific routes

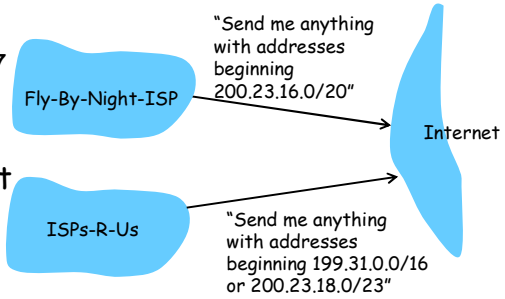
ISPs-R-Us has a more specific route to Organization 1



Hierarchical IP addressing: more specific routes

- Multiple advertised routes could hold destination
  - 200.23.16.0/20
  - 200.23.18.0/23

  both hold 200.23.18.7
- always route to *more specific* destination (longest prefix match)



Dealing with Scale: Other Examples

- Ethernet doesn't scale to very large numbers of users
- routing: fully-connected network only possible when small; use hierarchy to keep elements in same level fully connected (e.g., telephone network)
- p2p networks (like gnutella) scale using hierarchy
- DNS – use hierarchy to scale
- Use of multicast to scale 1-many communication: N-way 1-1 communication OK when small, move to 1-N as N gets large. What about reliability: explicit receiver-to-sender ACK/NAK OK when small, but implosion as N gets big. SRM – uses local recovery
- Human management doesn't scale: DHCP as auto-configuration. Autonomic computing – zero configure. Self-healing.