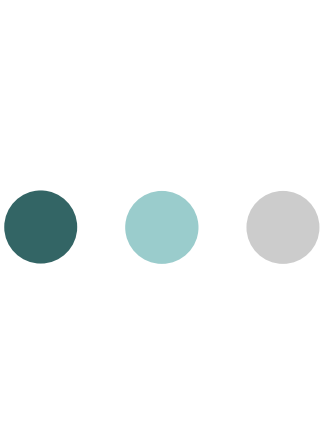


CS419: Computer Networks

Lecture 4, Feb 14, 2004

IP Forwarding Table

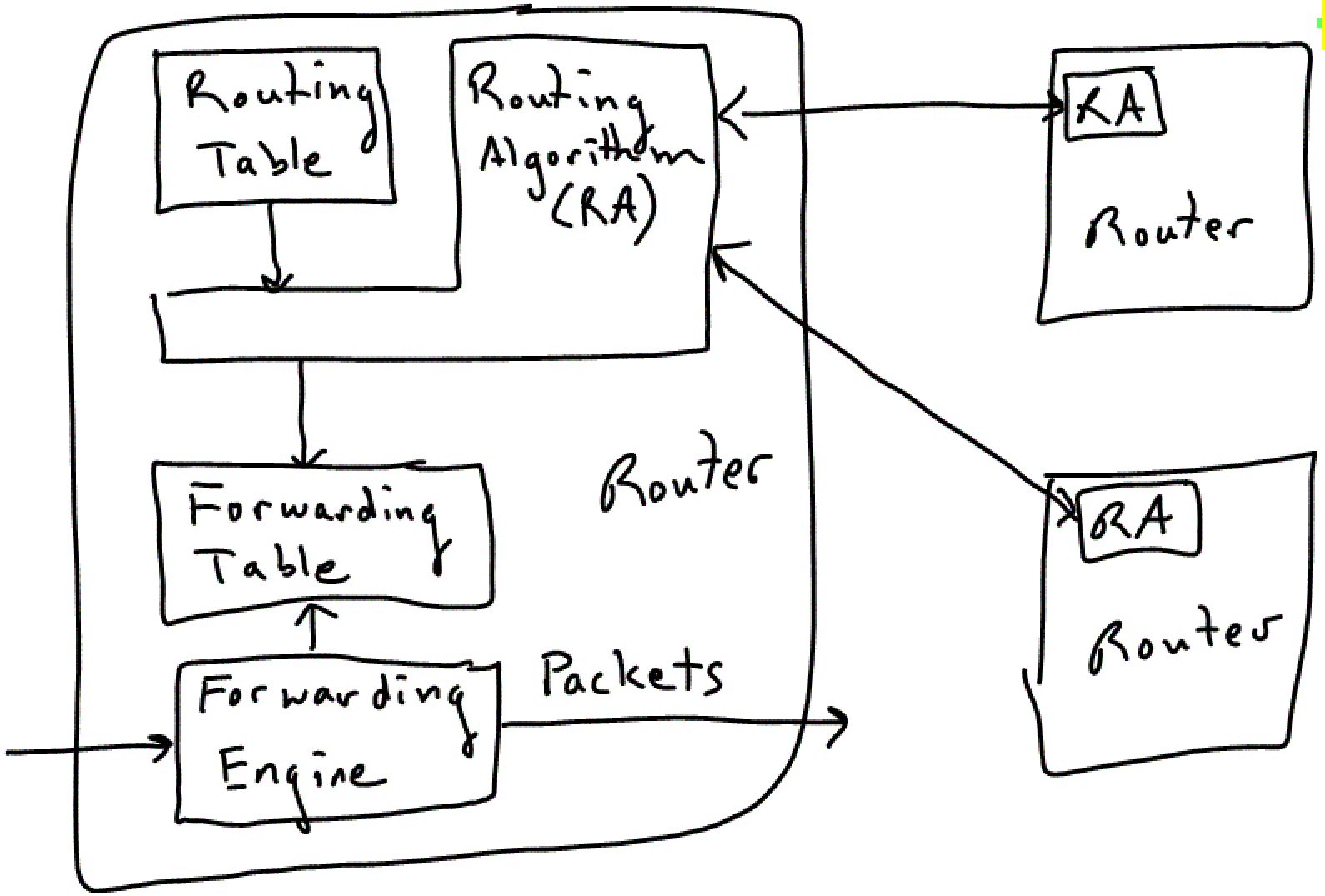
Routing and Forwarding Revisited

CS419

- We separate notion of “routing” and “forwarding”
- Routing algorithm is what a router does in the “background” to figure out where each prefix should be forwarded
  - Address prefixes, next hops, link costs, distances, etc.
- Forwarding is what a router does when a packet arrives
  - Address prefixes, next hops, interface, subnet address

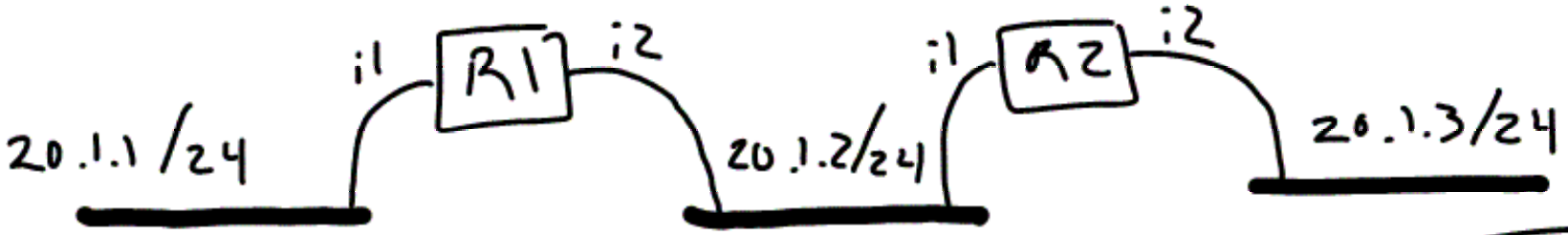
Routing and Forwarding Revisited

CS419



A simple example

CS419



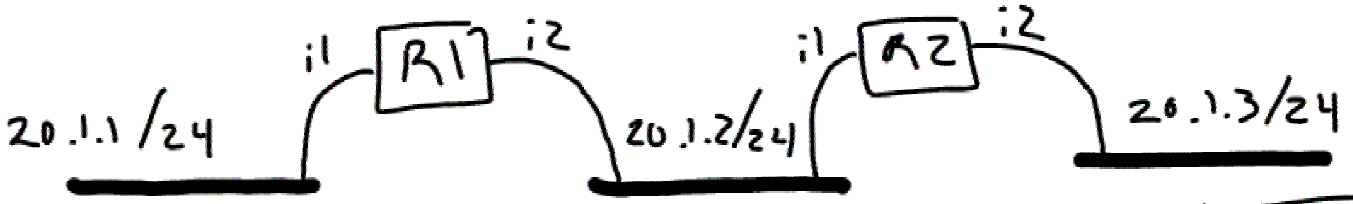
addr	mask	Next Hop	iface	subnet addr
20.1.1.0	255.255.255.0	local	i1	X
20.1.2.0	255.255.255.0	local	i2	X
20.1.3.0	255.255.255.0	R2	i2	1:2:3:4:5:6

R1 forwarding table (FIB)

Simple (naïve) forwarding rule

CS419

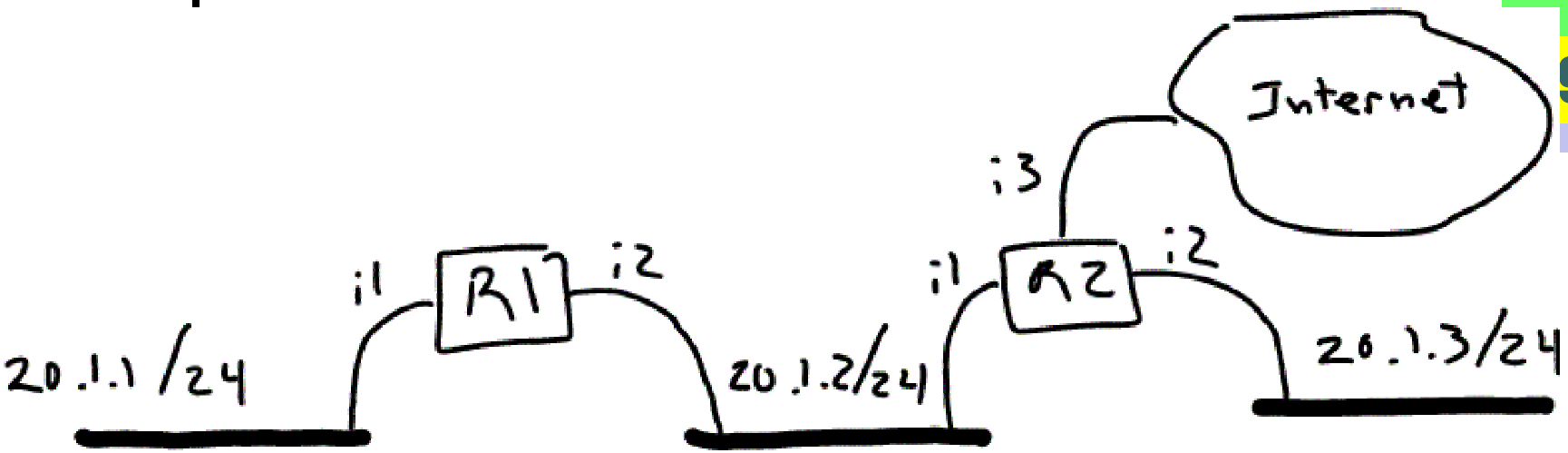
- Step through table from top to bottom
- At each step, apply mask to FIB address and packet address. If results match, then use FIB entry to forward packet
  - If (FIB-addr && FIB-mask) ==
  - (PK-addr && FIB-mask)
  - then use entry
- FIB = Forwarding Information Base
  - i.e. Forwarding Table
  - Routing Table also called RIB



addr	mask	Next Hop	iface	subnet addr
20.1.1.0	255.255.255.0	local	i1	x
20.1.2.0	255.255.255.0	local	i2	x
20.1.3.0	255.255.255.0	R2	i2	1:2:3:4:5:6

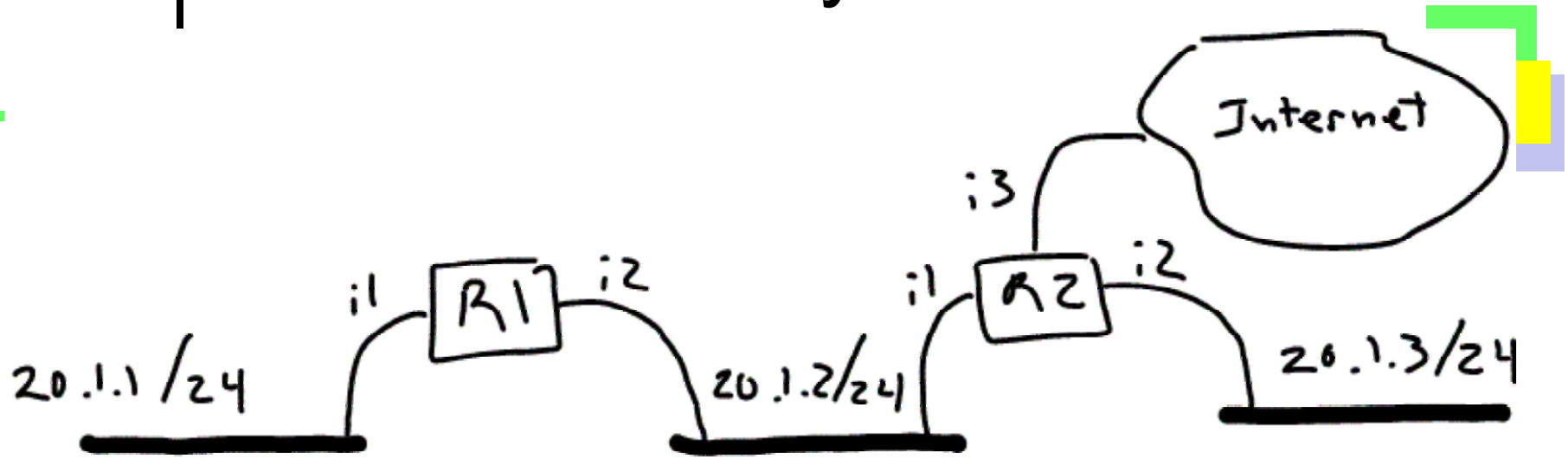
CS419

Simple example with default



20.1.1.0	255.255.255.0	:1
20.1.2.0	255.255.255.0	:2
20.1.3.0	255.255.255.0	R2
0.0.0.0	0.0.0.0	R2

But default entry must be last!

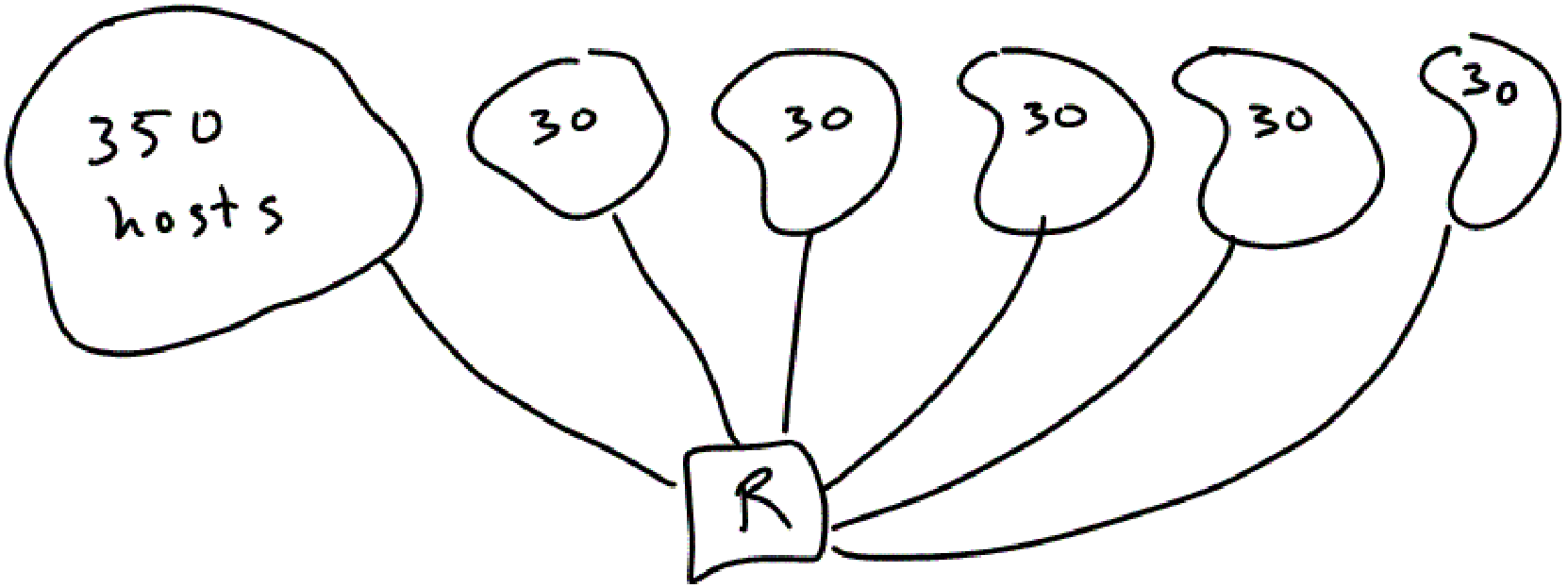


0.0.0.0	0.0.0.0	R2
20.1.1.0	255.255.255.0	:1
20.1.2.0	255.255.255.0	:2
20.1.3.0	255.255.255.0	R2

A more complex example (a site with 500 hosts)

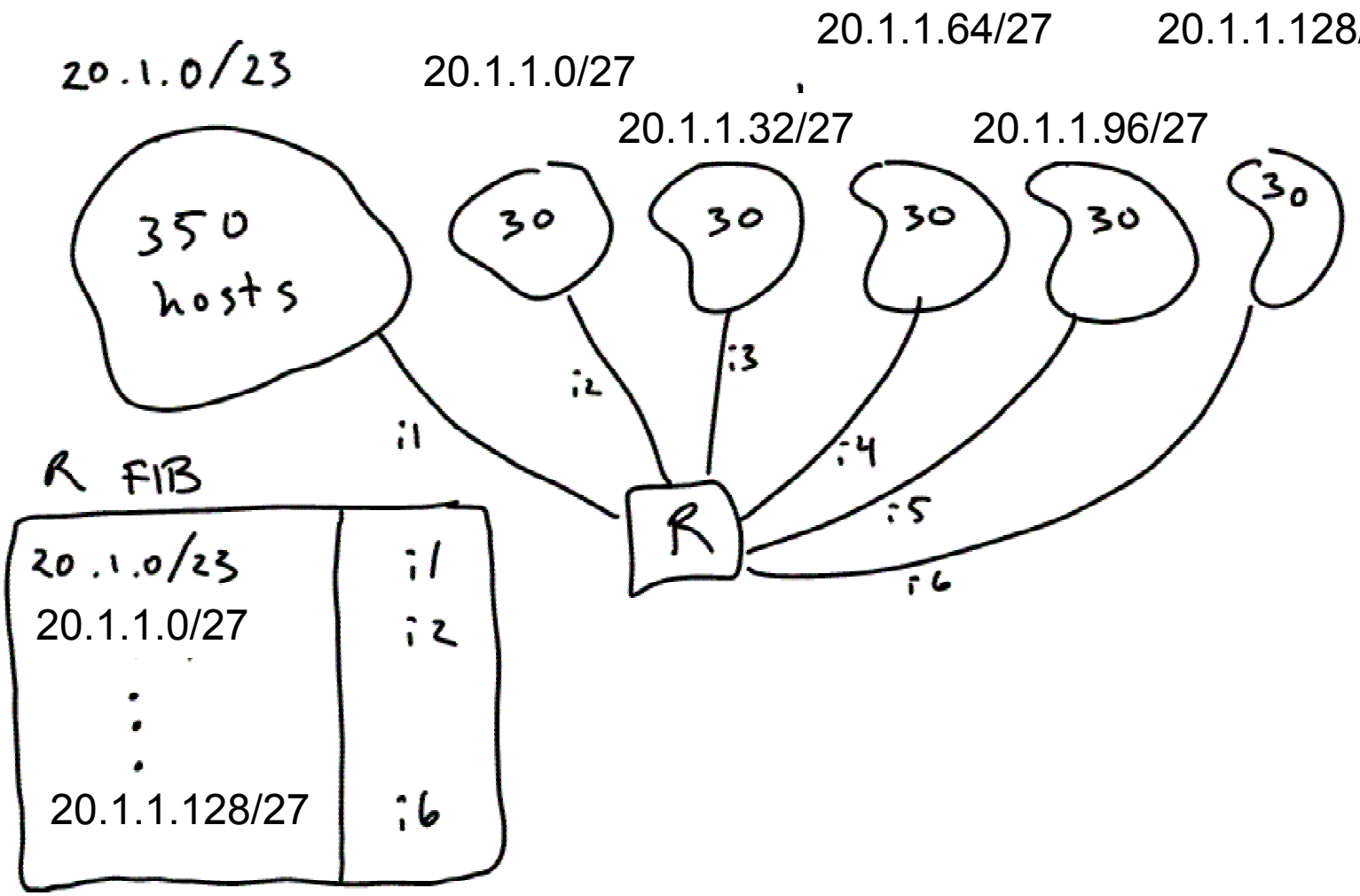
CS419

- How do we assign prefixes (addr and mask) in this case???



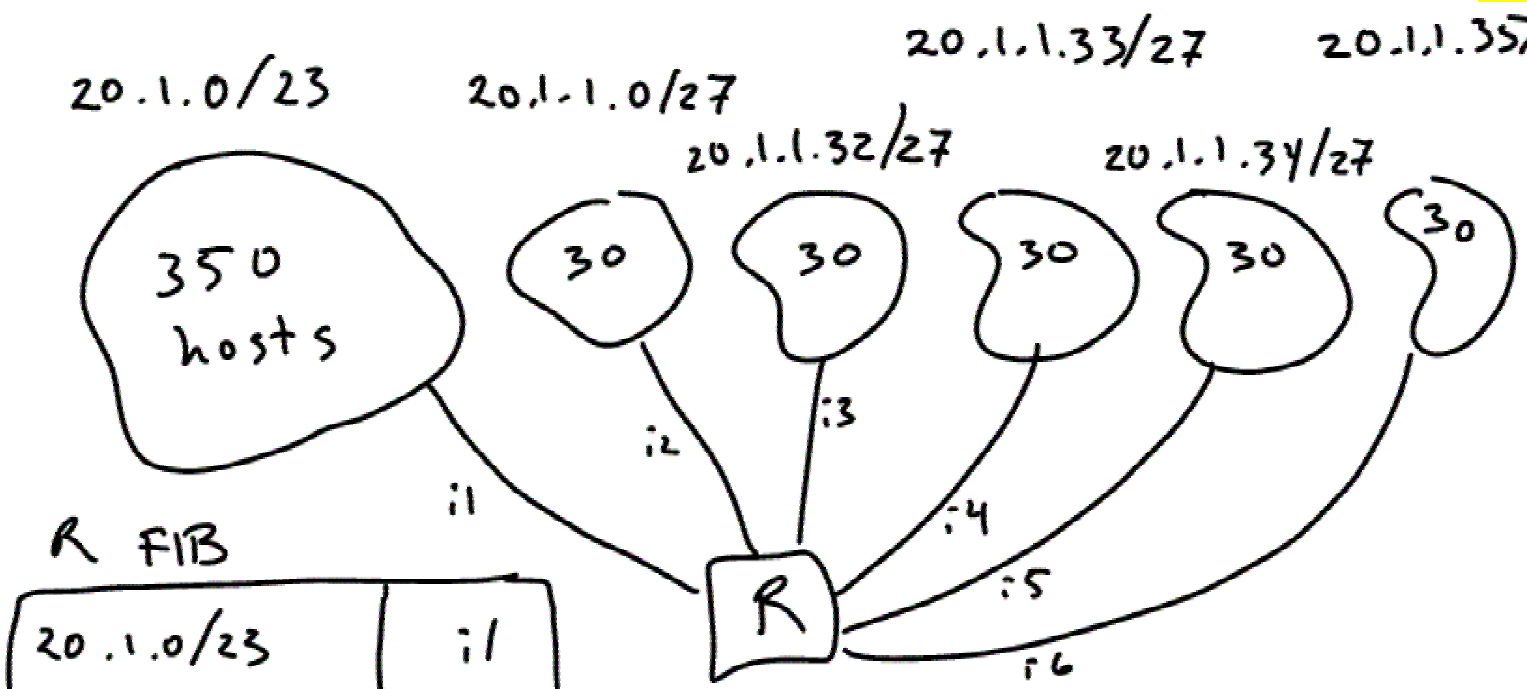
One way to assign prefixes…

CS419



My mistake last year…..

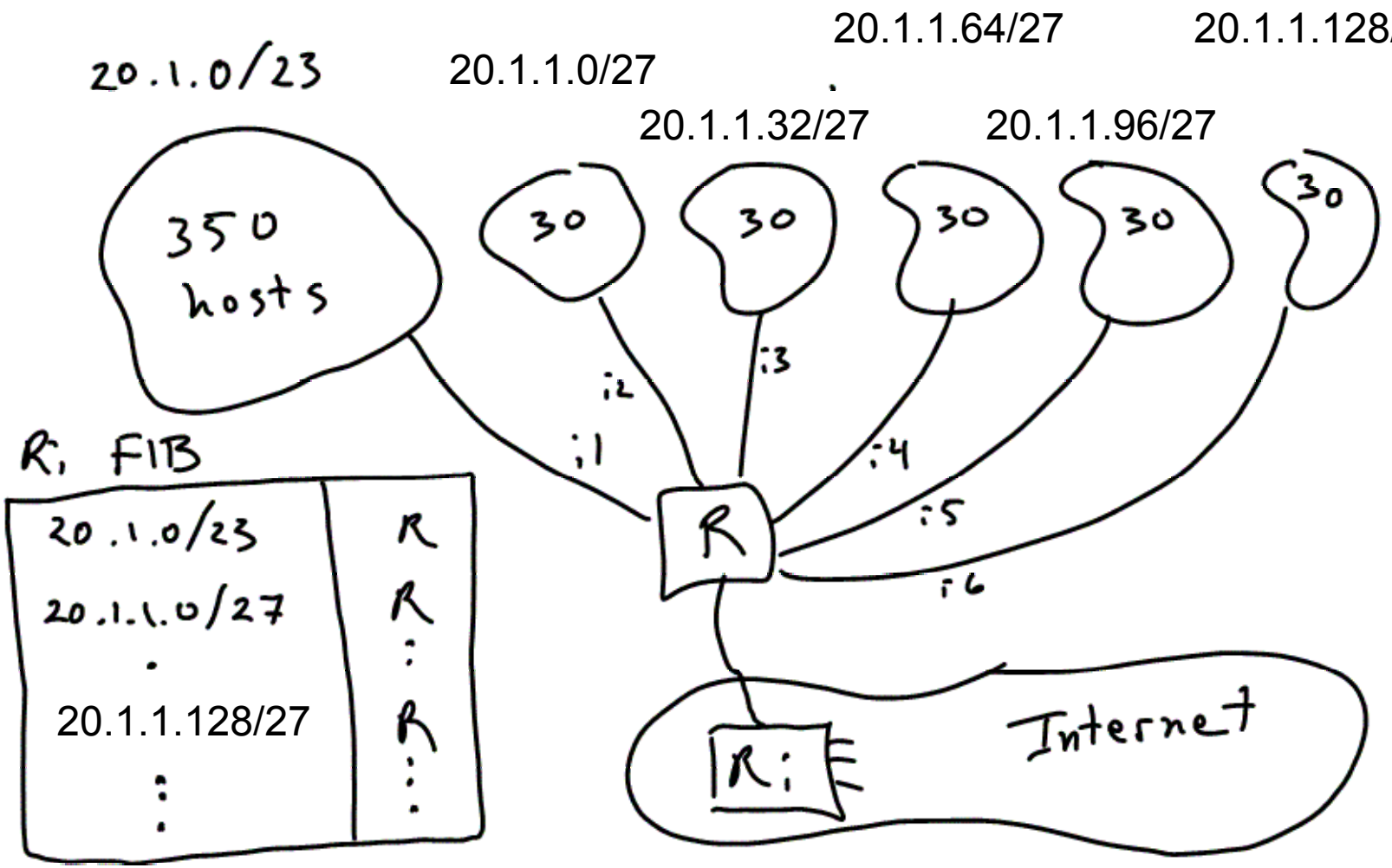
CS419



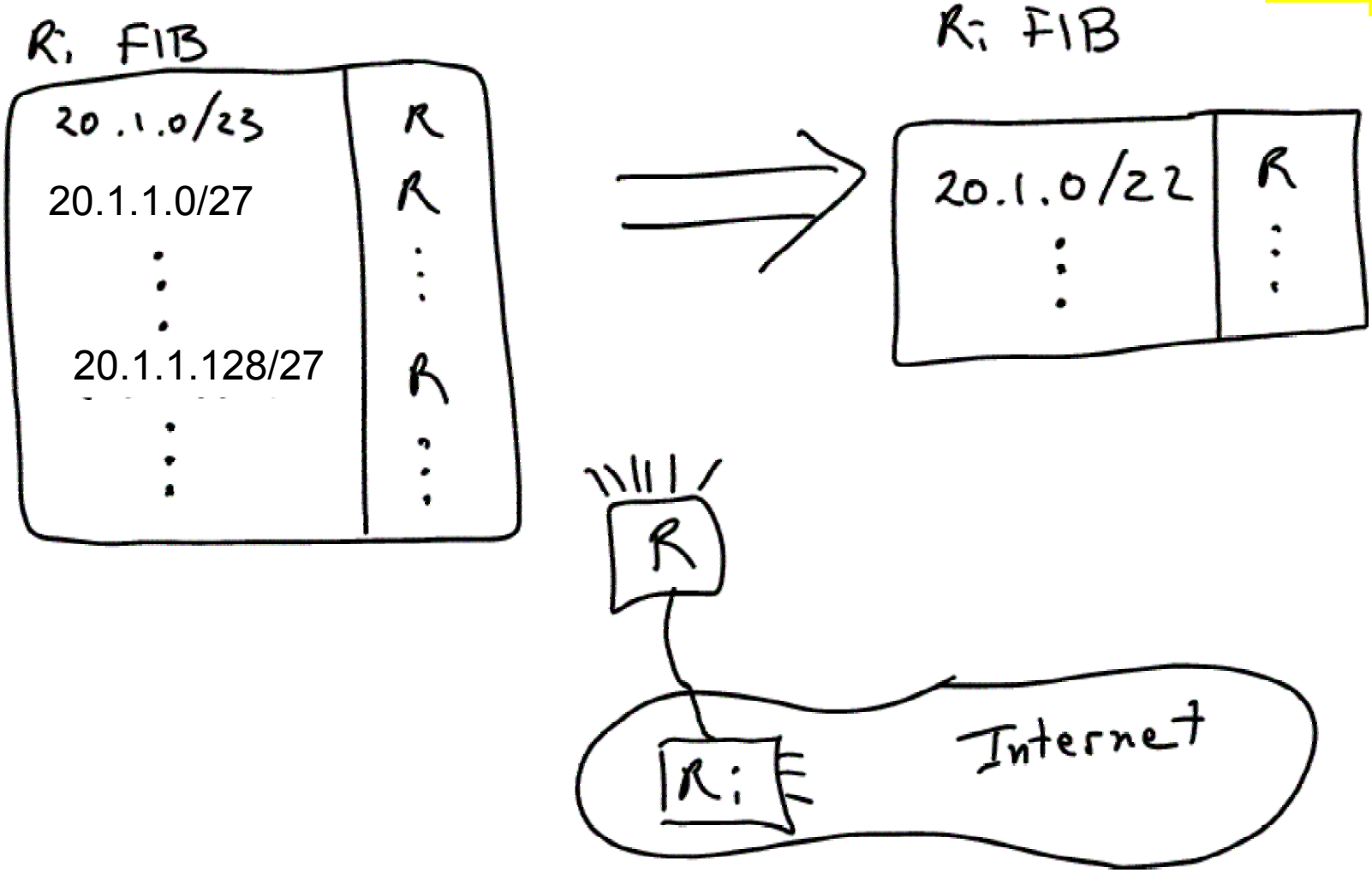
R FIB

20.1.0/23	:1
20.1.1.0/27	:2
⋮	
20.1.1.35/27	:6

The view from the global Internet: 6 FIB entries!

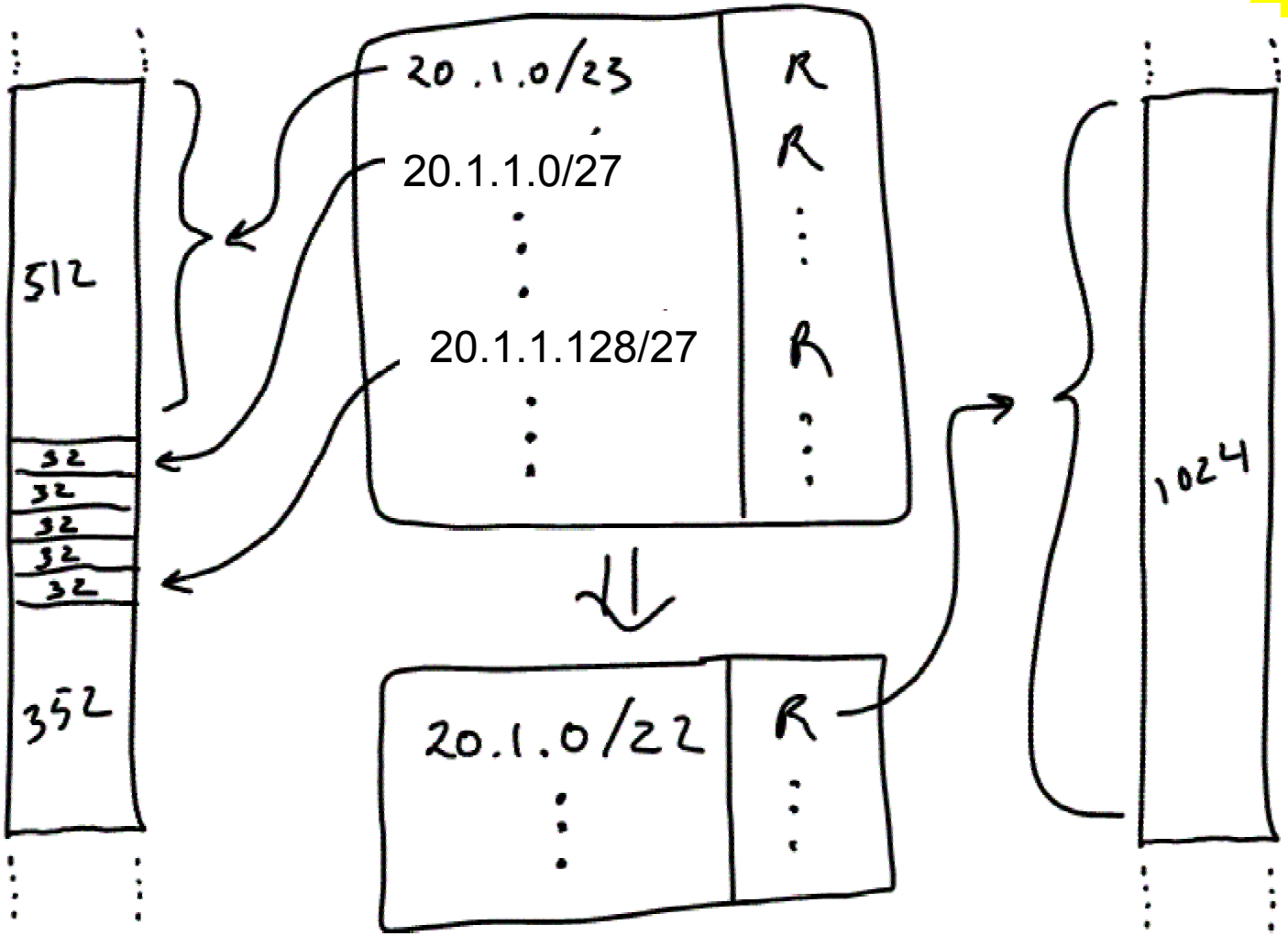


We can shrink that to one FIB entry!



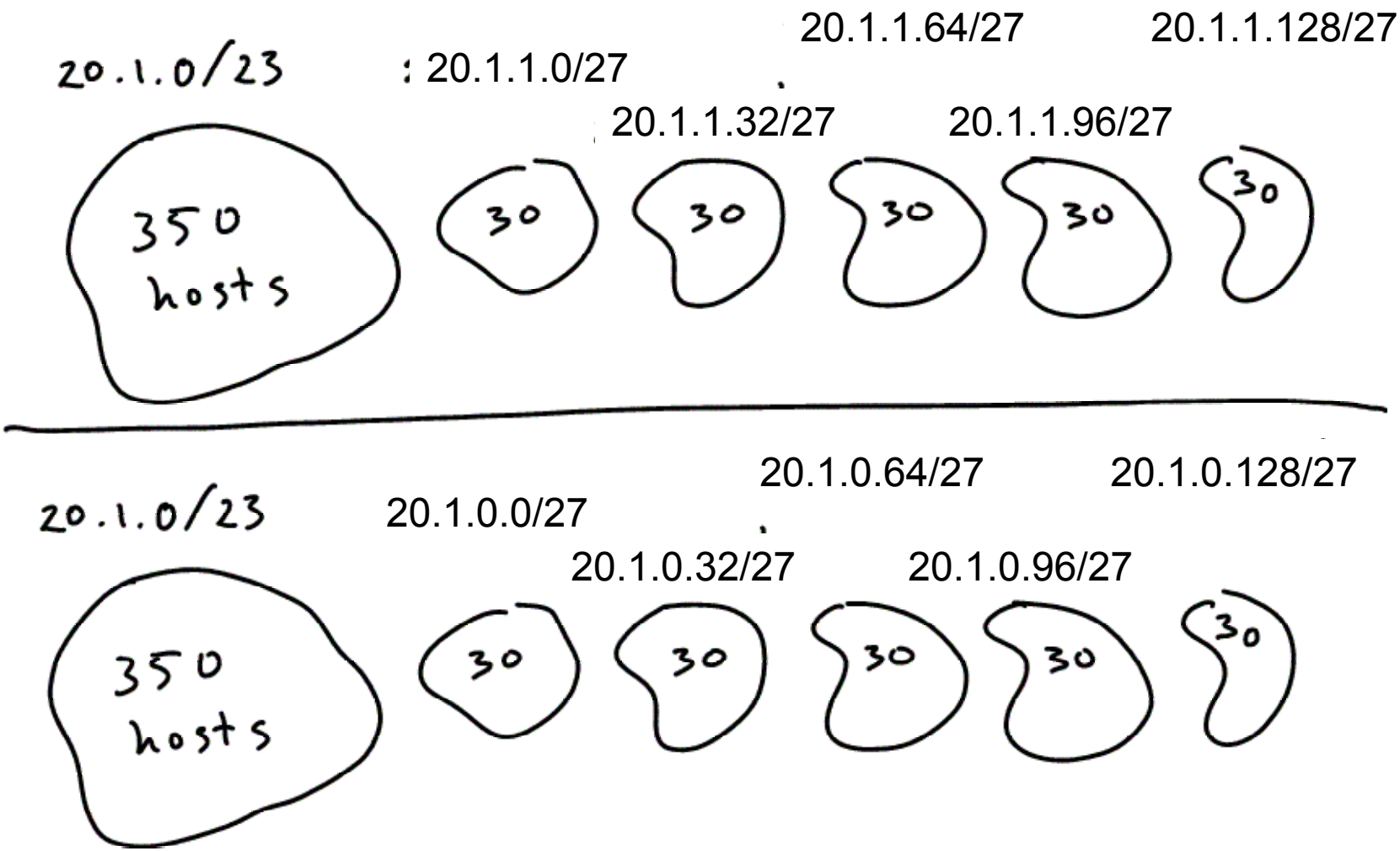
1024 addresses to address 500 hosts! What a waste…

CS419

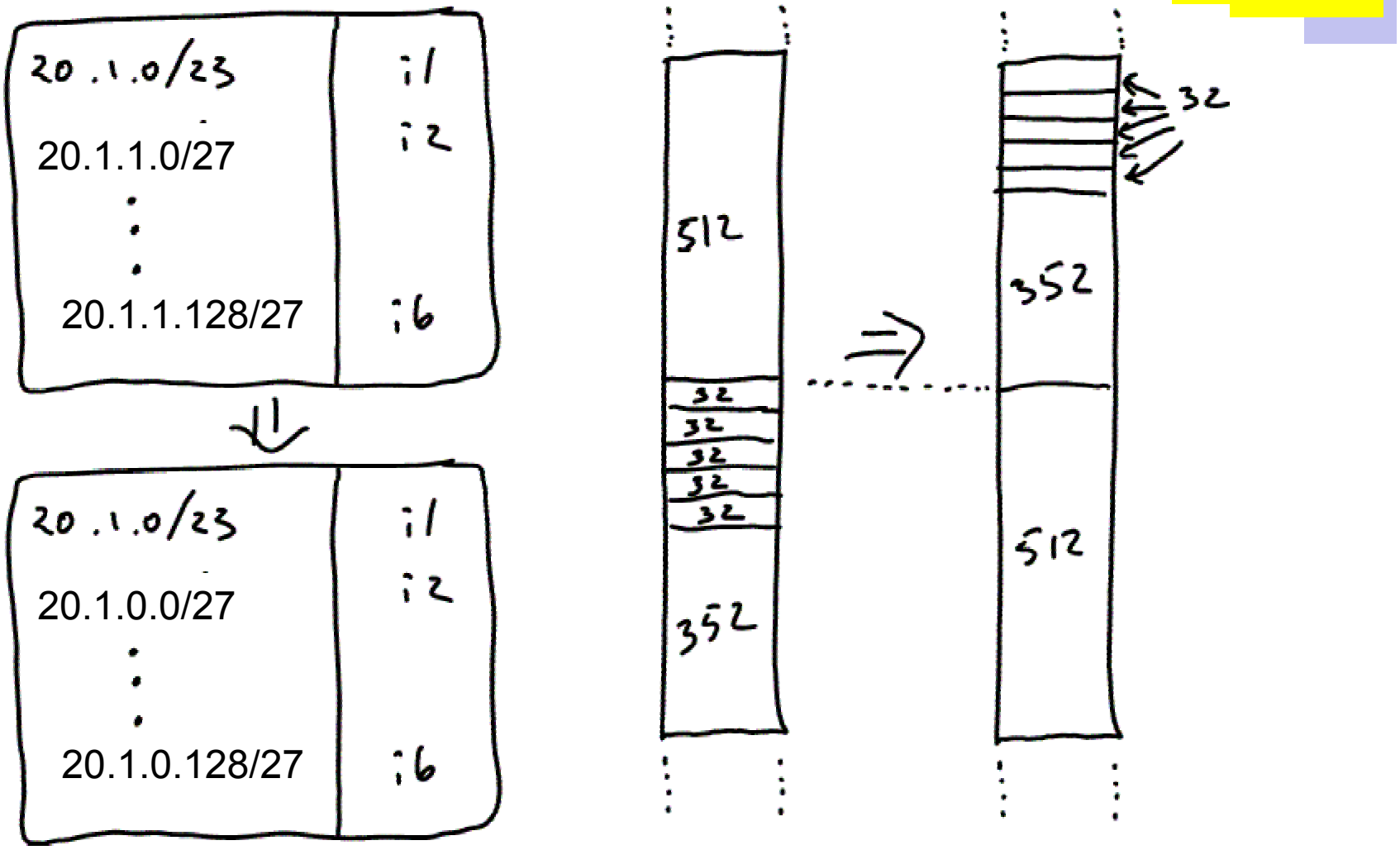


What about this prefix assignment approach instead?

CS419

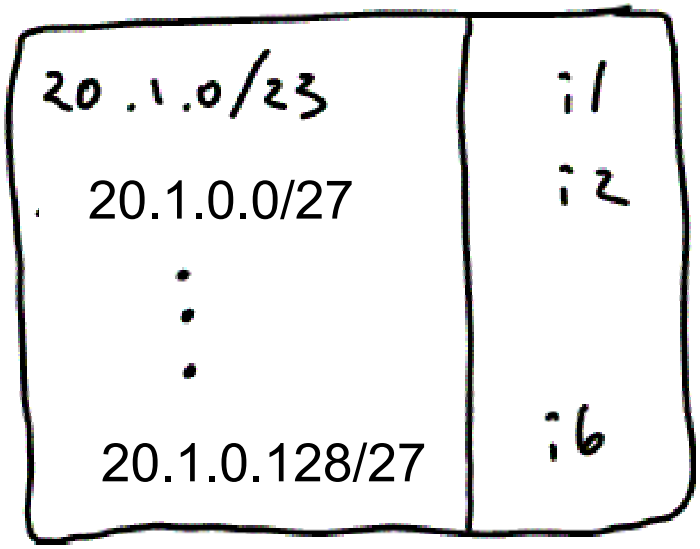


Now 500 addresses fit into a 512 address block!



But now our forwarding rules fail (like with the default)

CS419



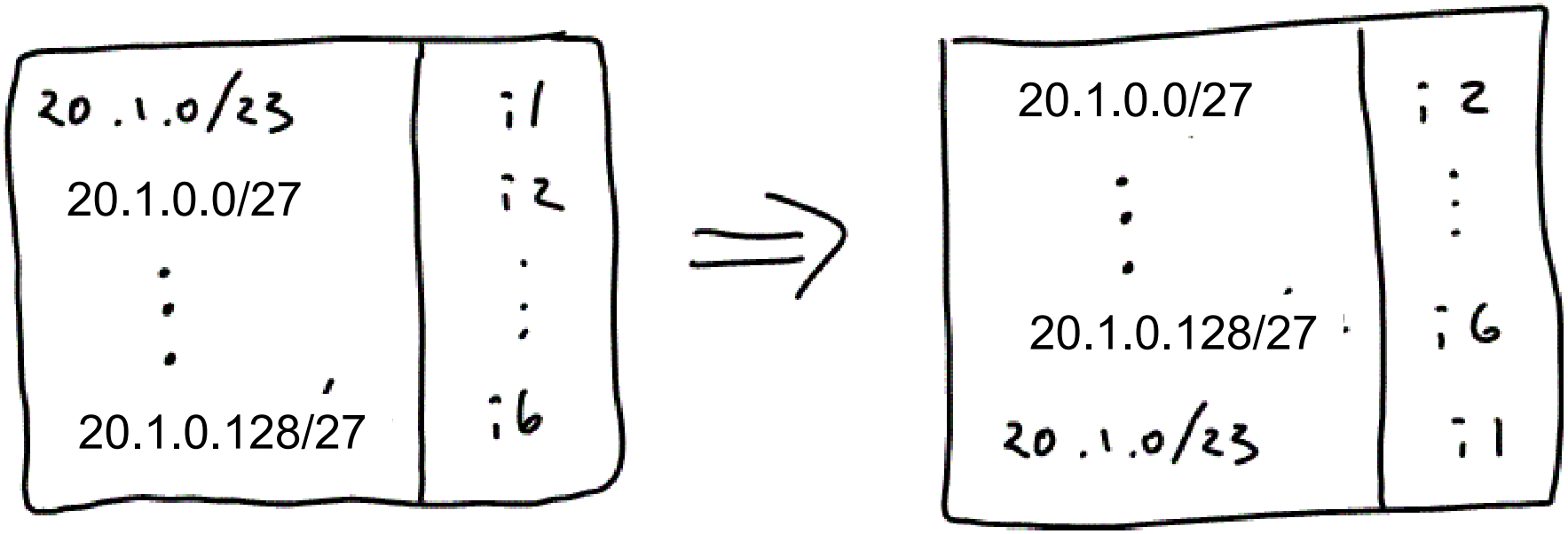
Longest-prefix match

CS419

- Since multiple entries may match, we prefer the entry with the longest mask (prefix)
- Two ways:
  1. Go through the whole FIB, remembering the matching entriey with the longest prefix
  2. Sort FIB in order of longest prefix first, and select first match

First-match Longest-prefix

CS419



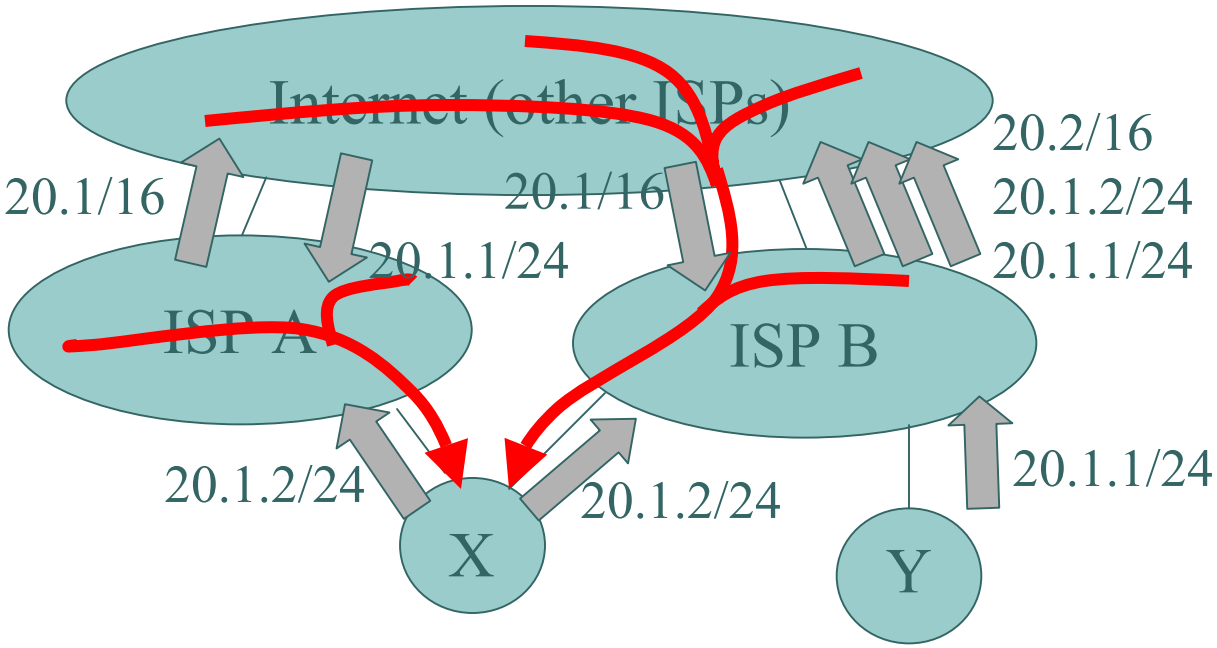
Best-match rules revisited

CS419

- Select matching FIB entry with longest prefix
- If multiple matching FIB entries have the same prefix size, then any may be used
  - Even simultaneously---path splitting for load balancing
  - But try to maintain source affinity (i.e. send different flows along different paths, but don't split a given flow)

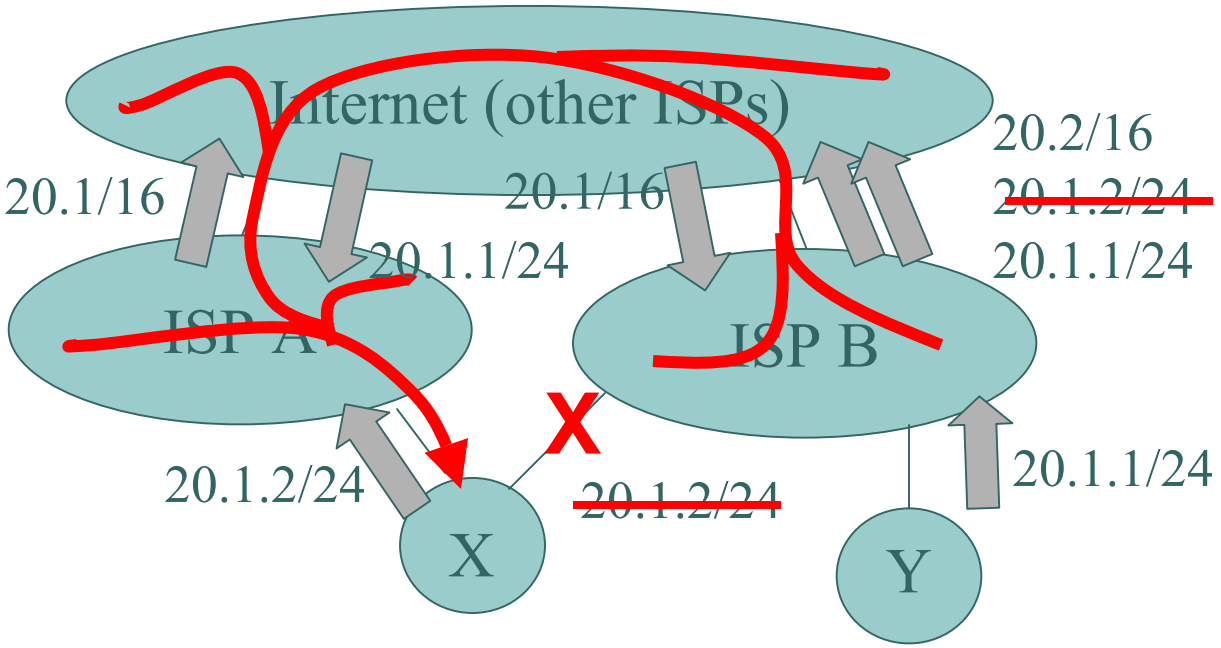
Paths to multi-homed site X

CS419



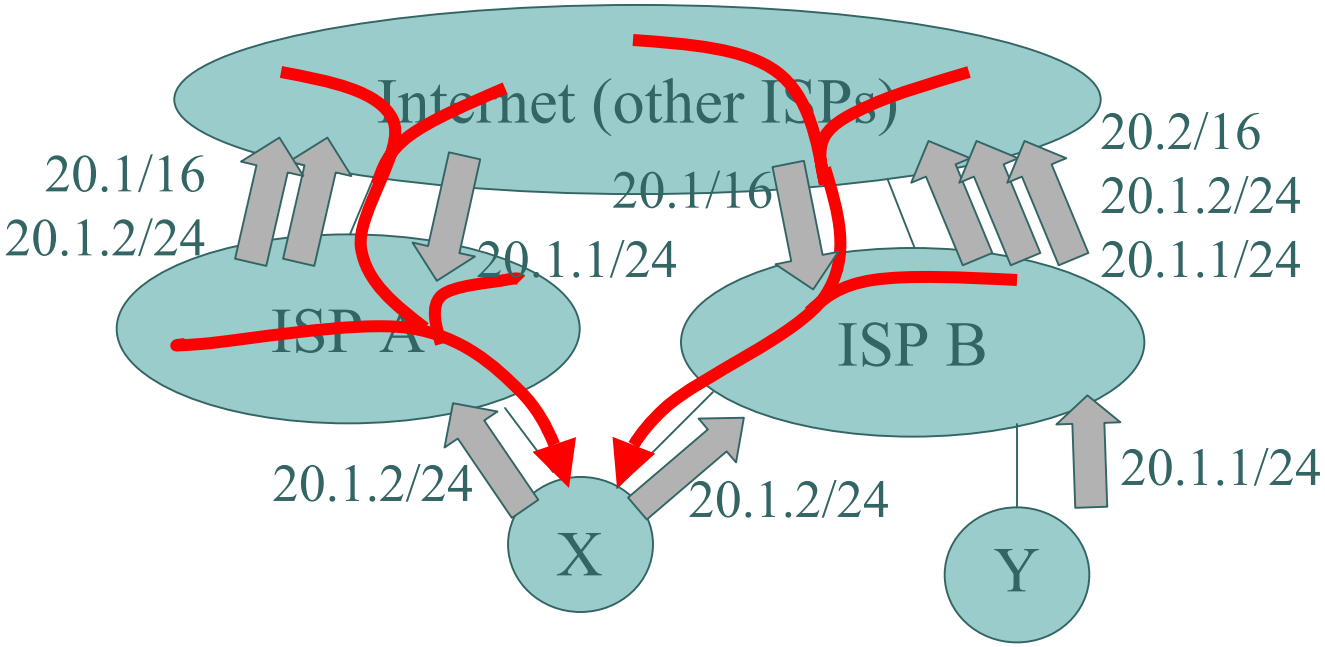
Paths to Site X after X-B link failure

CS419



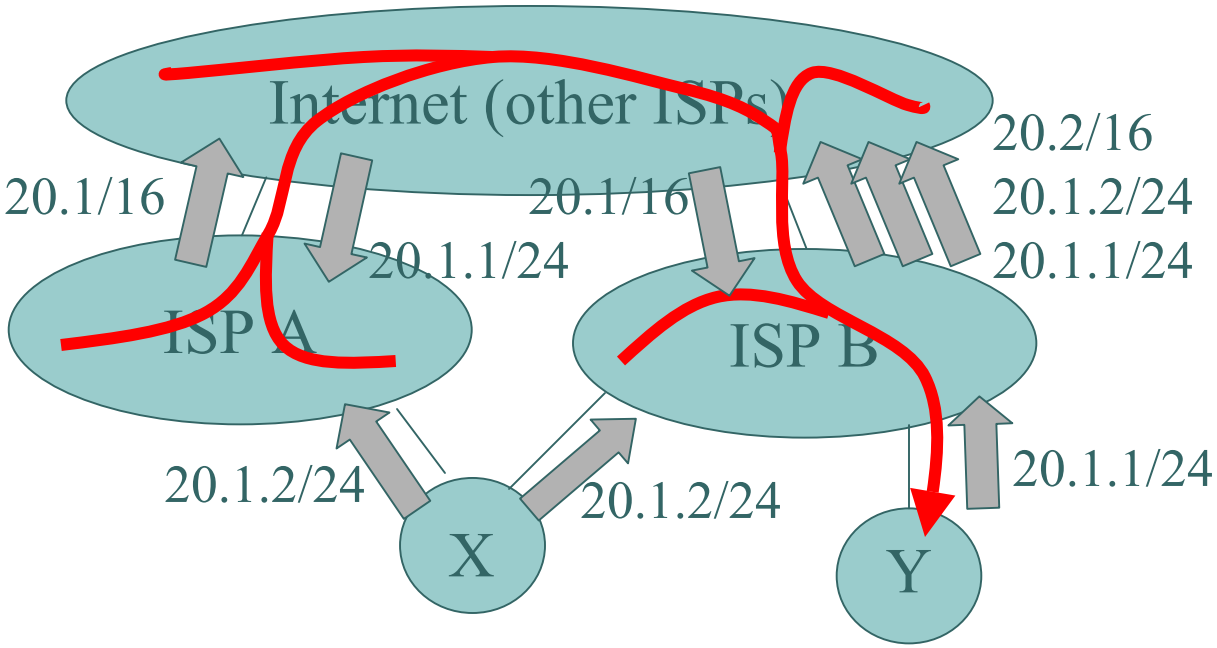
Better load balance (without increasing FIB size)

CS419



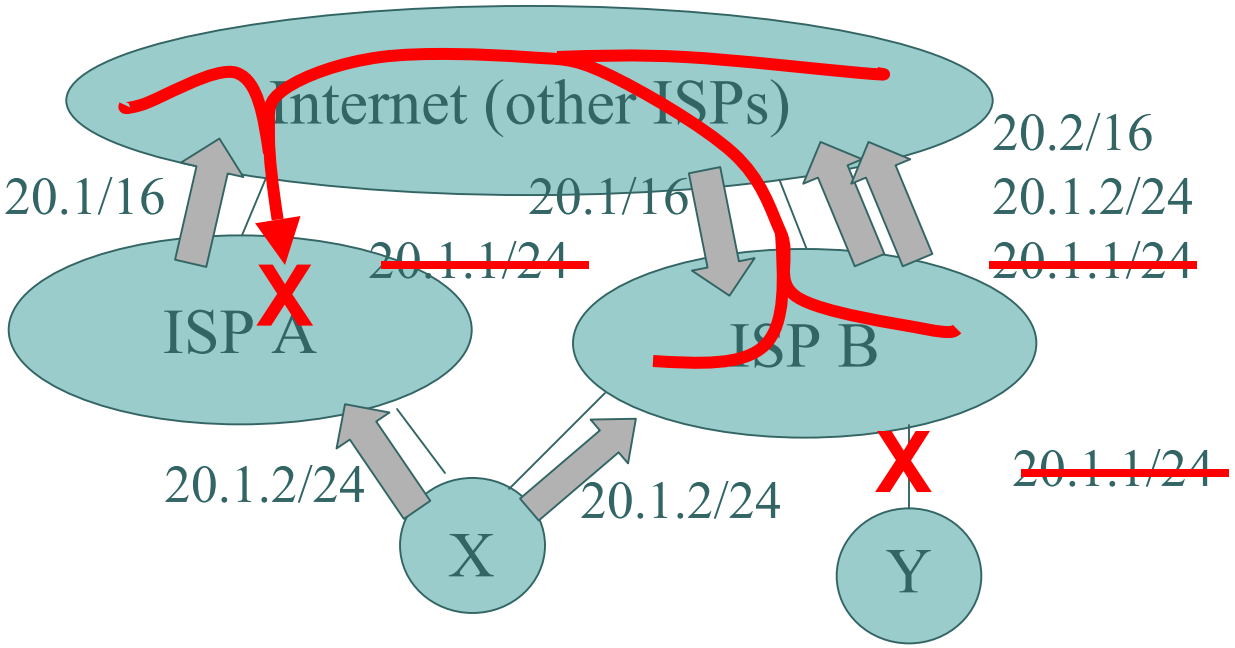
Paths to Site Y

CS419



Paths “to” Site Y after Y-B link failure

CS419



Implementing the forwarding table

CS419

- First-match style ok for small forwarding tables
  - Scales poorly with the number of entries
- Hash structures work for flat addresses, but not hierarchical (masked) addresses
  - “Bridged Ethernets”
- High-end routers implement forwarding table in hardware
  - Combination of a fancy tree structure and CAMs (Content Addressable Memory)
- Otherwise, some kind of tree-like data structure is typically used
  - We’ll look at this later in the course

Other types of forwarding

CS419

- What we looked at so far is *hop-by-hop* forwarding with *hierarchical addresses*
- Hop-by-hop means that every switch in the path makes an “independent” forwarding decision
- But we can also have *source routing*
  - The entire path is listed in the packet
  - IP has a (never used) option for this

Hop-by-hop versus source routing

CS419

- Source routing is (kindof) what you do when you print out directions from mapquest
  - I.e., you carry you path with you
- Hop-by-hop routing is often (kindof) how you find your way around Wal-Mart
  - “where is kids clothing?”, “where are socks?”

Hop-by-hop versus source routing

CS419

- Hop-by-hop is what is used in the Internet
  - Though many people have proposed source routing
- With the exception of routing through a switch fabric within a router
  - But we'll look at router/switch architecture later