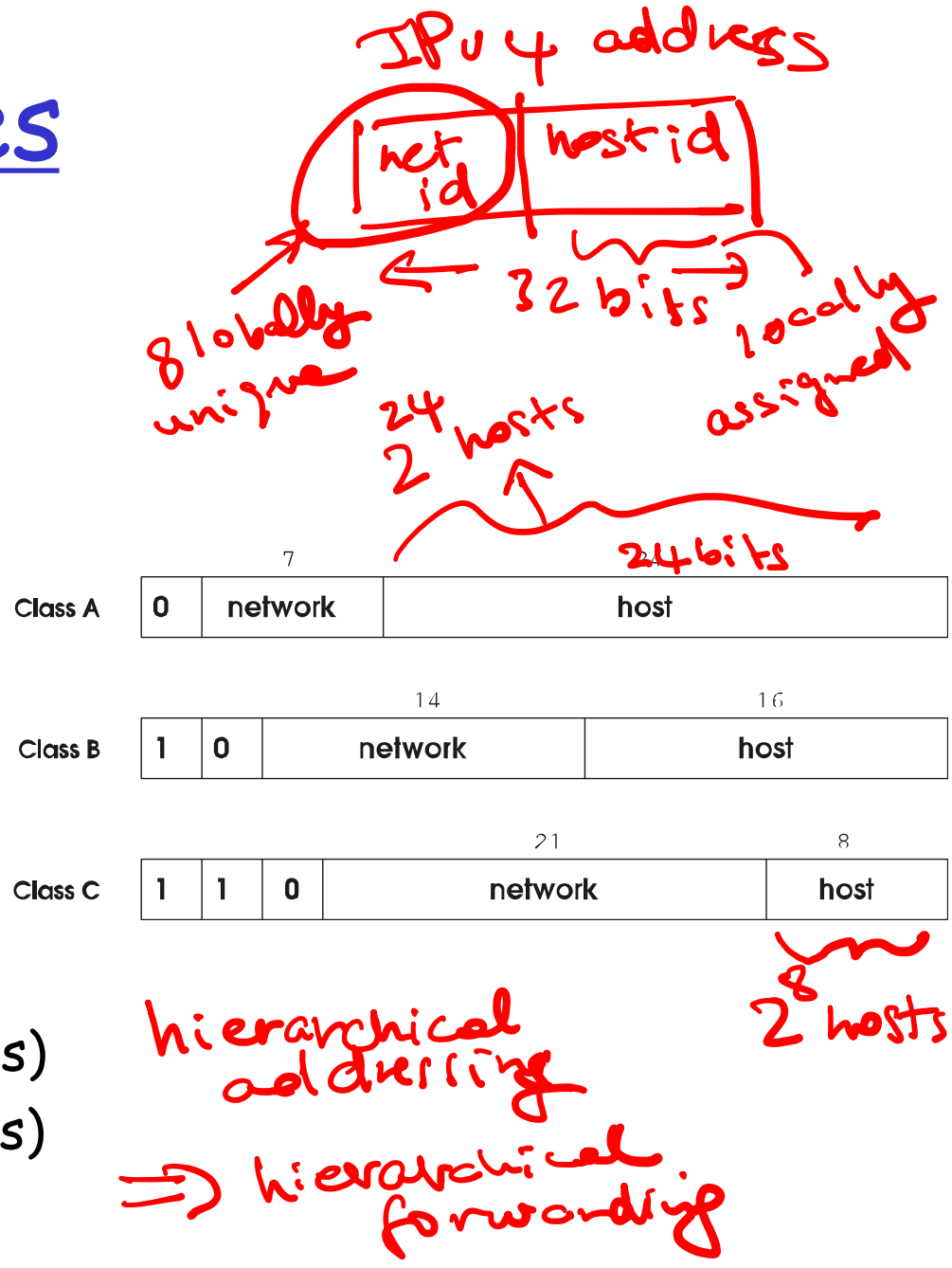
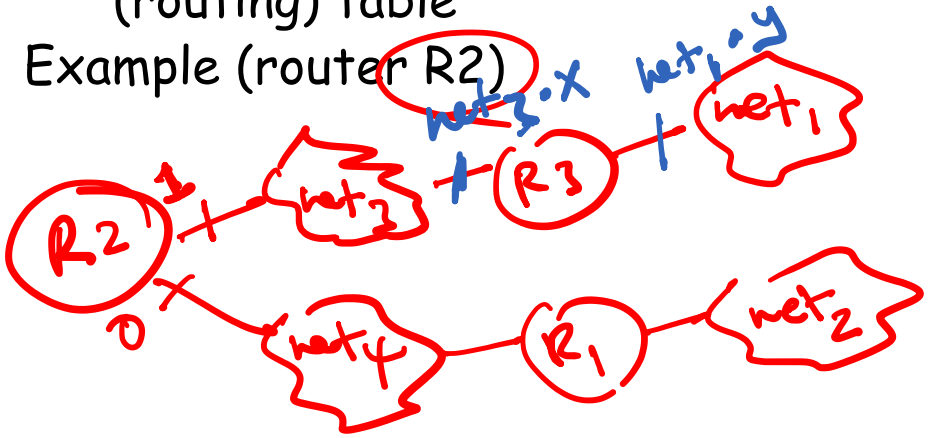
Global Addresses

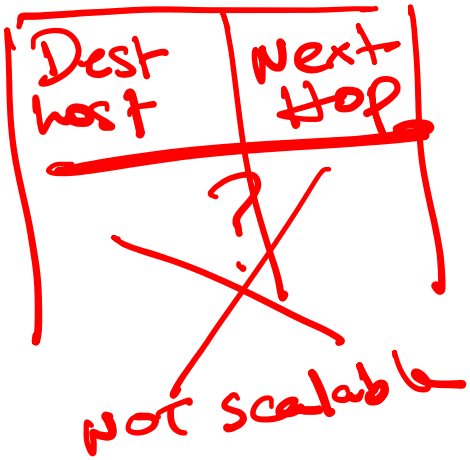
- Properties
  - globally unique
  - hierarchical: network + host
  - network part assigned by ICAAN
  - a router has one IP address for each interface
- Format
- Dot notation
  - 10.3.2.4 (class A address)
  - 128.96.33.81 (class B address)
  - 192.12.69.77 (class C address)



Datagram Forwarding

- Strategy
  - every datagram contains destination's address
  - if directly connected to destination network, then forward to host
  - if not directly connected to destination network, then forward to some router
  - forwarding table maps network number into next hop
  - each host has a default router
  - each router maintains a forwarding (routing) table
- Example (router R2)

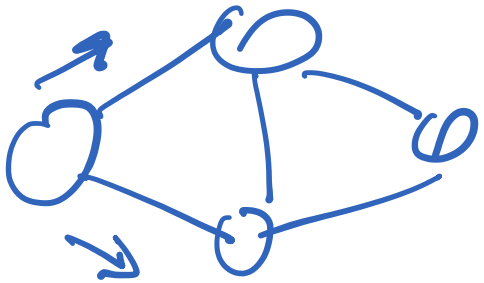


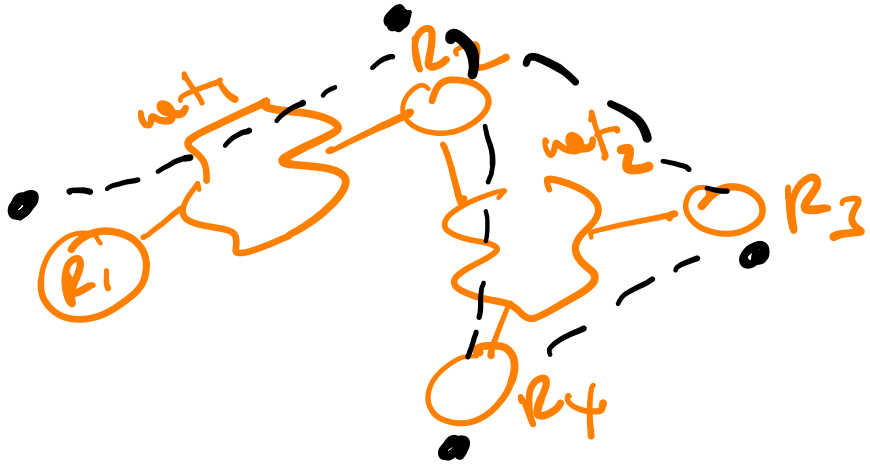


Net Number	Next Hop
1	R3 net3 · x
2	R1
3	Interface 1
4	Interface 0

How do IP Routers Maintain Routing Tables?

- Routing table contains {address of destination network, address of next-hop}
- In distance-vector routing, *neighbor* routers exchange their distances to destination networks
- In link-state routing, each router broadcasts (to *all* other routers) the status of attached networks





Scalability

1) Reduce RT size
2) Conserve IP addresses

IP ``hides'' hosts in address hierarchy, but...

- Inefficient use of address space
  - class C network with 2 hosts (2/255 = 0.78% efficient)
  - class B network with 256 hosts (256/65535 = 0.39% efficient)

want 1.
257 hosts

class B: 2B + 2B hostid

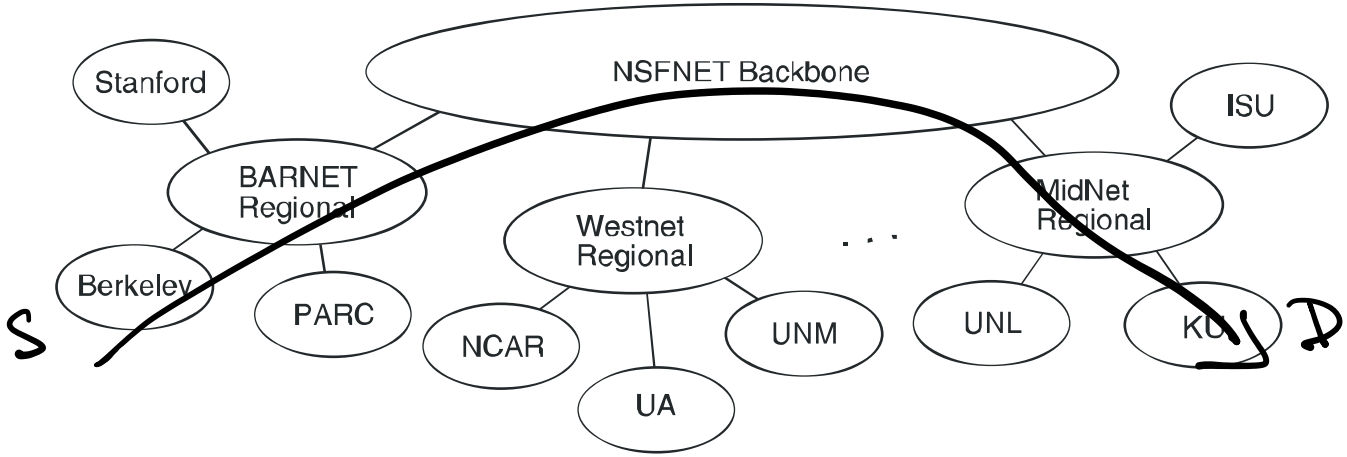
utilization = $\frac{257}{64K} \approx 0.4\%$

$\Rightarrow 2^{16} = 64K$ host addresses

- Too many networks
  - today's Internet has hundreds of thousands of networks
  - routing tables do not scale
  - route propagation protocols do not scale

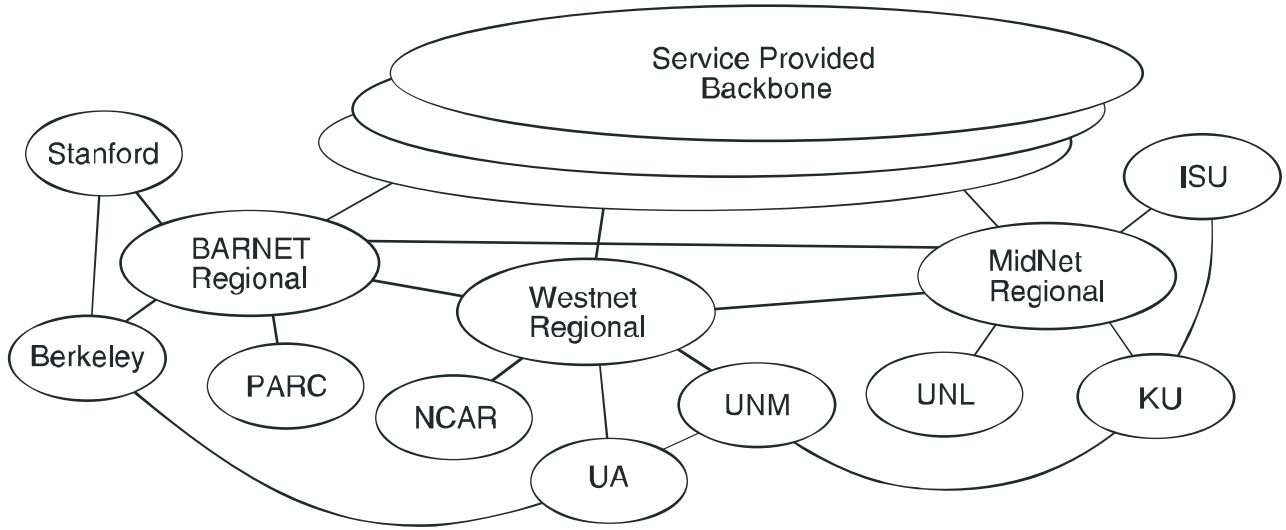
Internet Structure

- Before

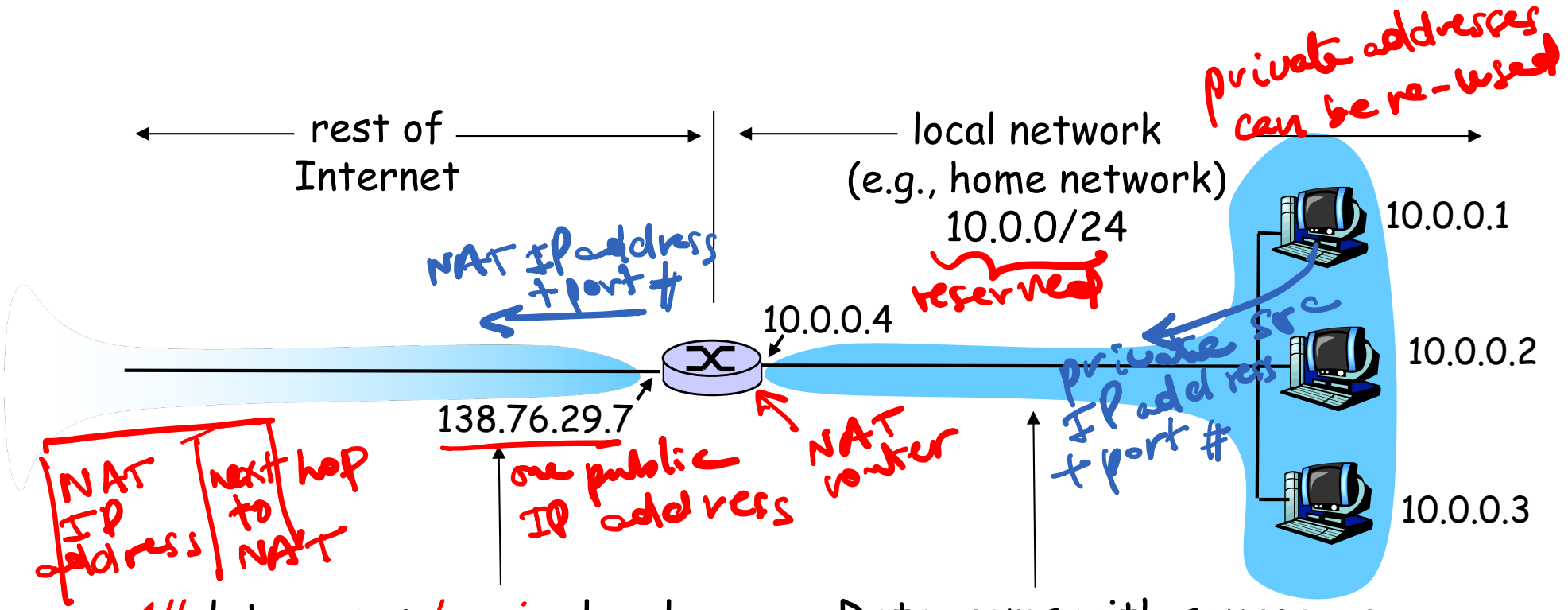


- After

constrained SP problem



NAT: Network Address Translation



All datagrams *leaving* local network have same single source NAT IP address: 138.76.29.7, different source port numbers

Datagrams with source or destination in this network have 10.0.0/24 address for source, destination (as usual)

NAT: Network Address Translation

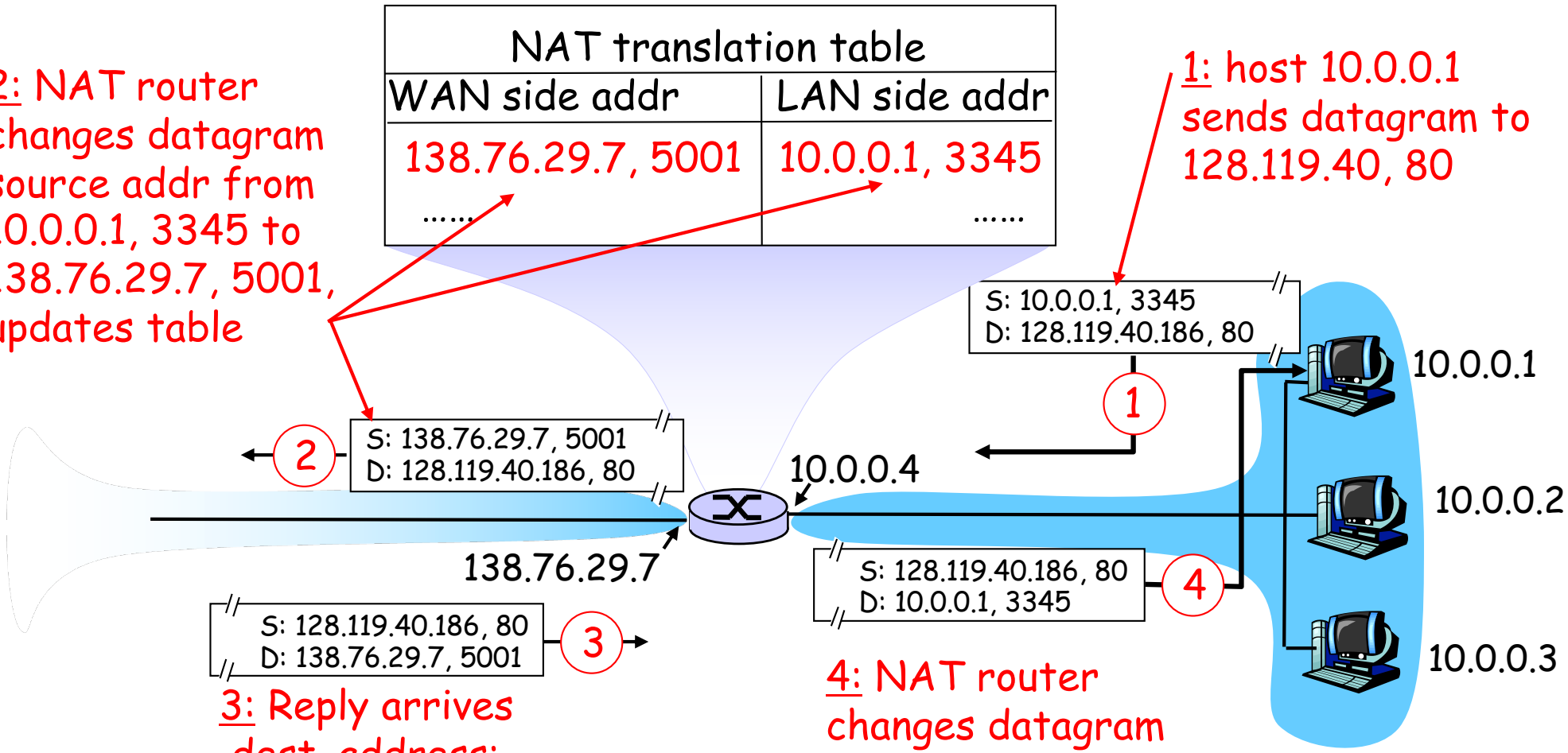
- Motivation: local network uses just one IP address as far as outside word is concerned:
  - no need to be allocated range of addresses from ISP: - just one IP address is used for all devices
  - can change addresses of devices in local network without notifying outside world
  - can change ISP without changing addresses of devices in local network
  - devices inside local net not explicitly addressable, visible by outside world (a security plus)

NAT: Network Address Translation

Implementation: NAT router must:

- *outgoing datagrams: replace* (source IP address, port #) of every outgoing datagram to (NAT IP address, new port #)
  - . . . remote clients/servers will respond using (NAT IP address, new port #) as destination addr.
- *remember (in NAT translation table)* every (source IP address, port #) to (NAT IP address, new port #) translation pair
- *incoming datagrams: replace* (NAT IP address, new port #) in dest fields of every incoming datagram with corresponding (source IP address, port #) stored in NAT table

NAT: Network Address Translation

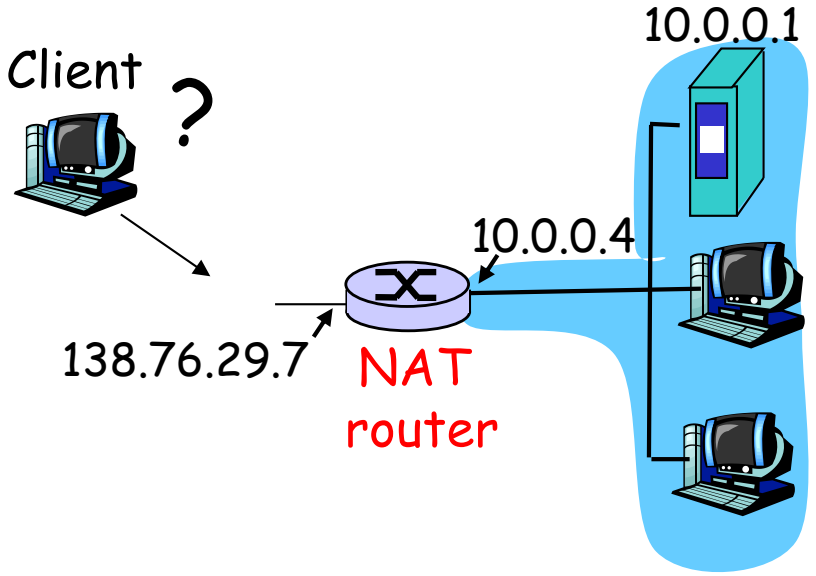


NAT: Network Address Translation

- 16-bit port-number field:
  - Up to 64K simultaneous connections with a single LAN-side address!
- NAT is controversial:
  - routers should only process up to layer 3
  - violates end-to-end argument
    - NAT possibility must be taken into account by app designers, eg, P2P applications
  - address shortage should instead be solved by IPv6 ?

NAT traversal problem

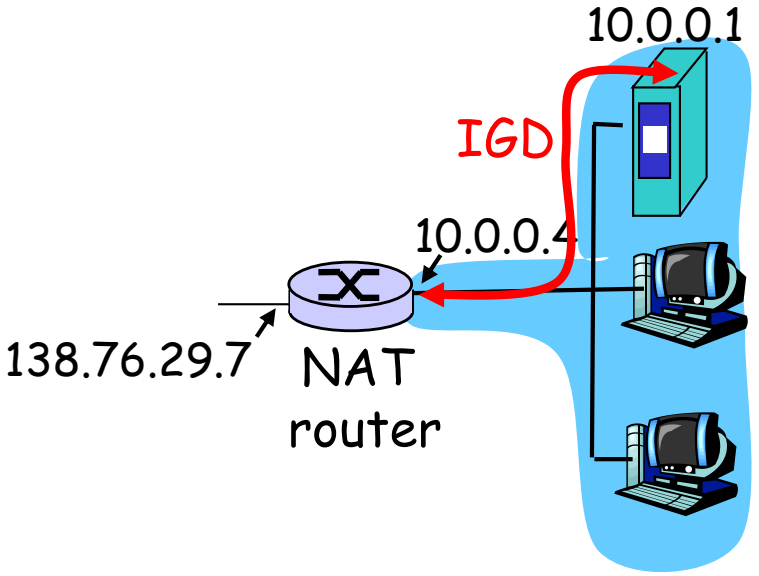
- client wants to connect to server with address 10.0.0.1
  - server address 10.0.0.1 local to LAN (client can't use it as destination addr)
  - only one externally visible NATted address: 138.76.29.7
- solution 1: statically configure NAT to forward incoming connection requests at given port to server
  - e.g., (138.76.29.7, port 2500) always forwarded to 10.0.0.1 port 25000



NAT traversal problem

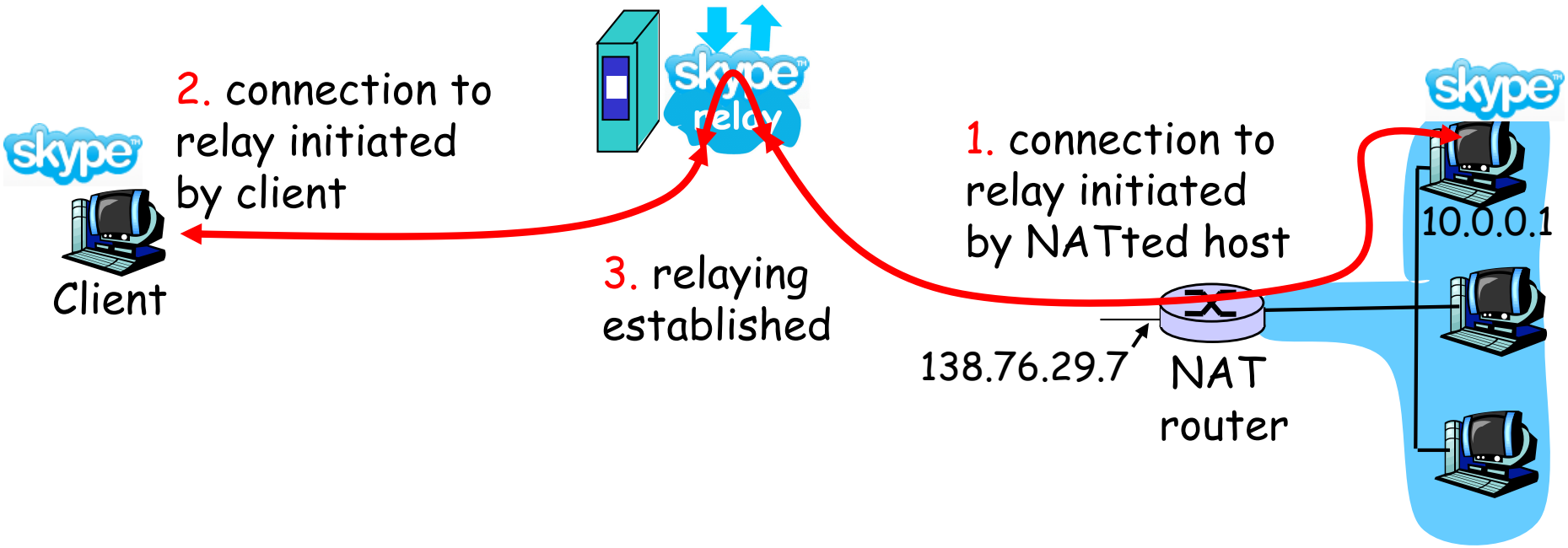
- solution 2: Universal Plug and Play (UPnP) Internet Gateway Device (IGD) Protocol. Allows NATted host to:
  - learn public IP address (138.76.29.7)
  - add/remove port mappings (with lease times)

  i.e., automate static NAT port map configuration

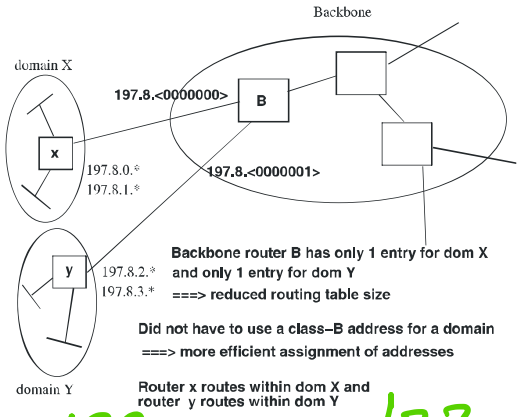


NAT traversal problem

- solution 3: relaying (used in Skype)
  - NATed client establishes connection to relay
  - External client connects to relay
  - relay bridges packets between two connections



Supernetting



4 class-C addresses

198.25.0.*
198.25.1.*
198.25.2.*
198.25.3.*

X: 198.25.0.0/23
256*4 = 1024 hosts

Y: 198.25.2.0/23

Routing Table

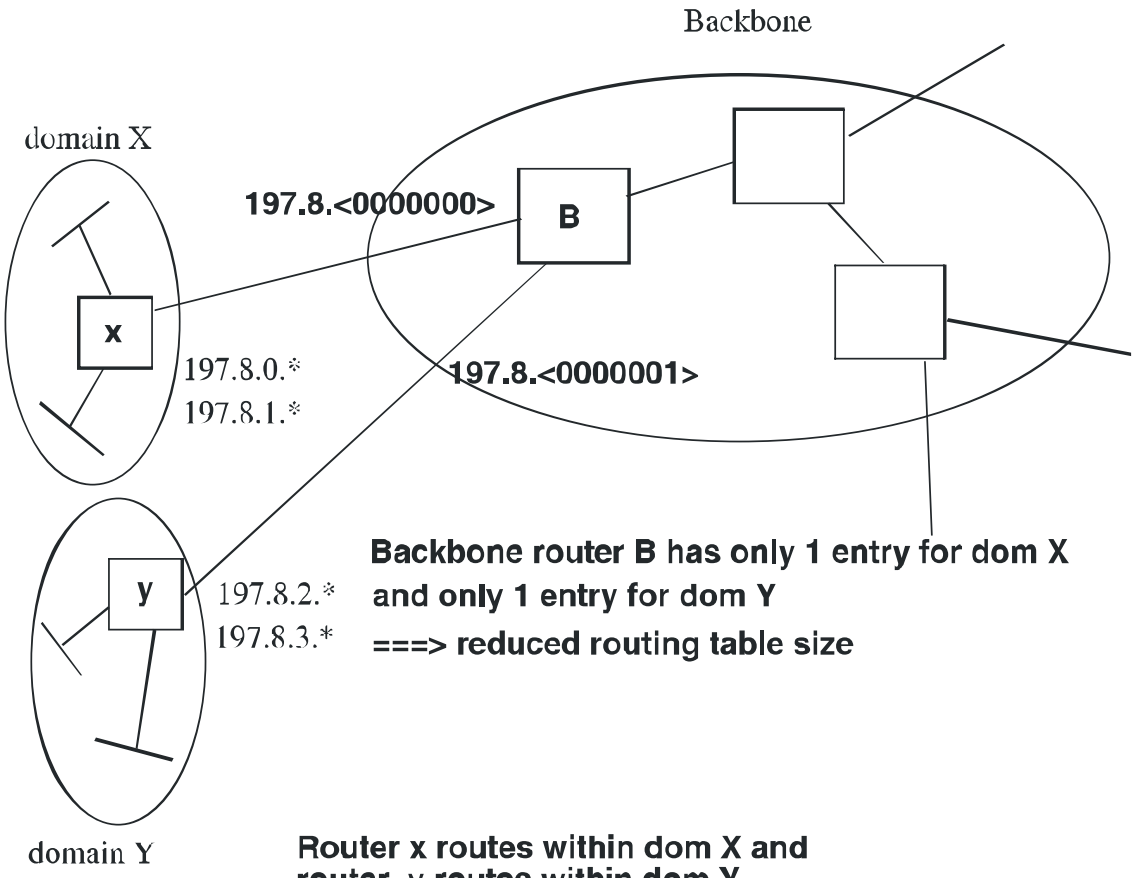
198.25.0.*	to R
198.25.1.*	to R
198.25.2.*	to R
198.25.3.*	to R
198.25.0.0/22	

198.25.2.0/23 By LPM (Longest Prefix Match)

Contiguous block
common prefix of 22 bits
198.25.0.0/22 ← mask
classless routing "CIDR"

Supernetting

- Assign block of contiguous network numbers to near-by networks
- Called CIDR: Classless Inter-Domain Routing
- Represent blocks with a single pair <first_network_address, common prefix>
- Use a bit mask (CIDR mask) to identify block
- All routers must understand CIDR addressing
- Hierarchical addressing and routing
- Longest Prefix Matching when looking up routing table



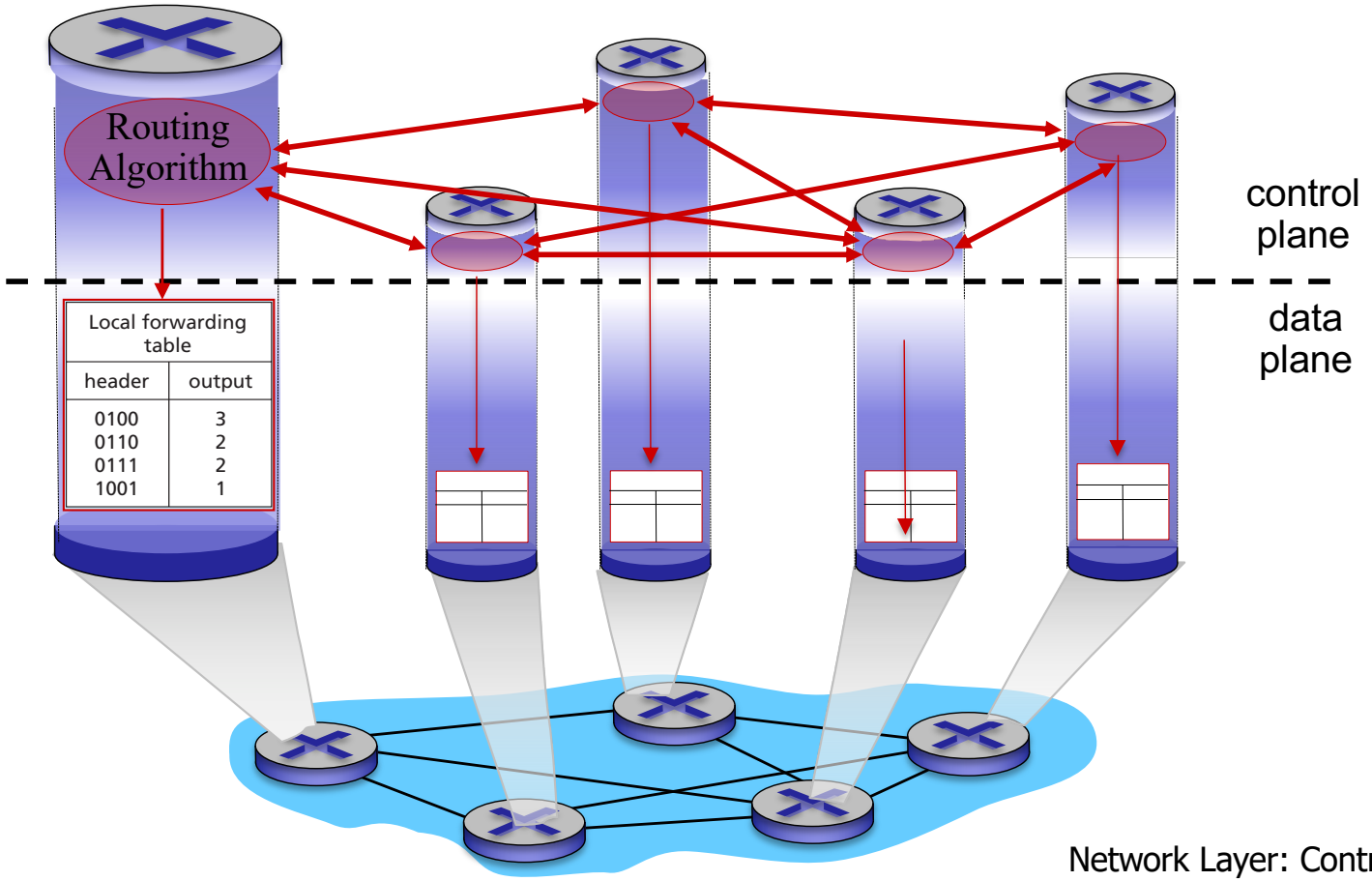
Software defined networking (SDN)

- Internet network layer: historically has been implemented via distributed, per-router approach
  - *monolithic* router contains switching hardware, runs proprietary implementation of Internet standard protocols (IP, RIP, IS-IS, OSPF, BGP) in proprietary router OS (e.g., Cisco IOS)
  - different "middleboxes" for different network layer functions: firewalls, NAT boxes, ..

- ~2005: renewed interest in rethinking network control plane

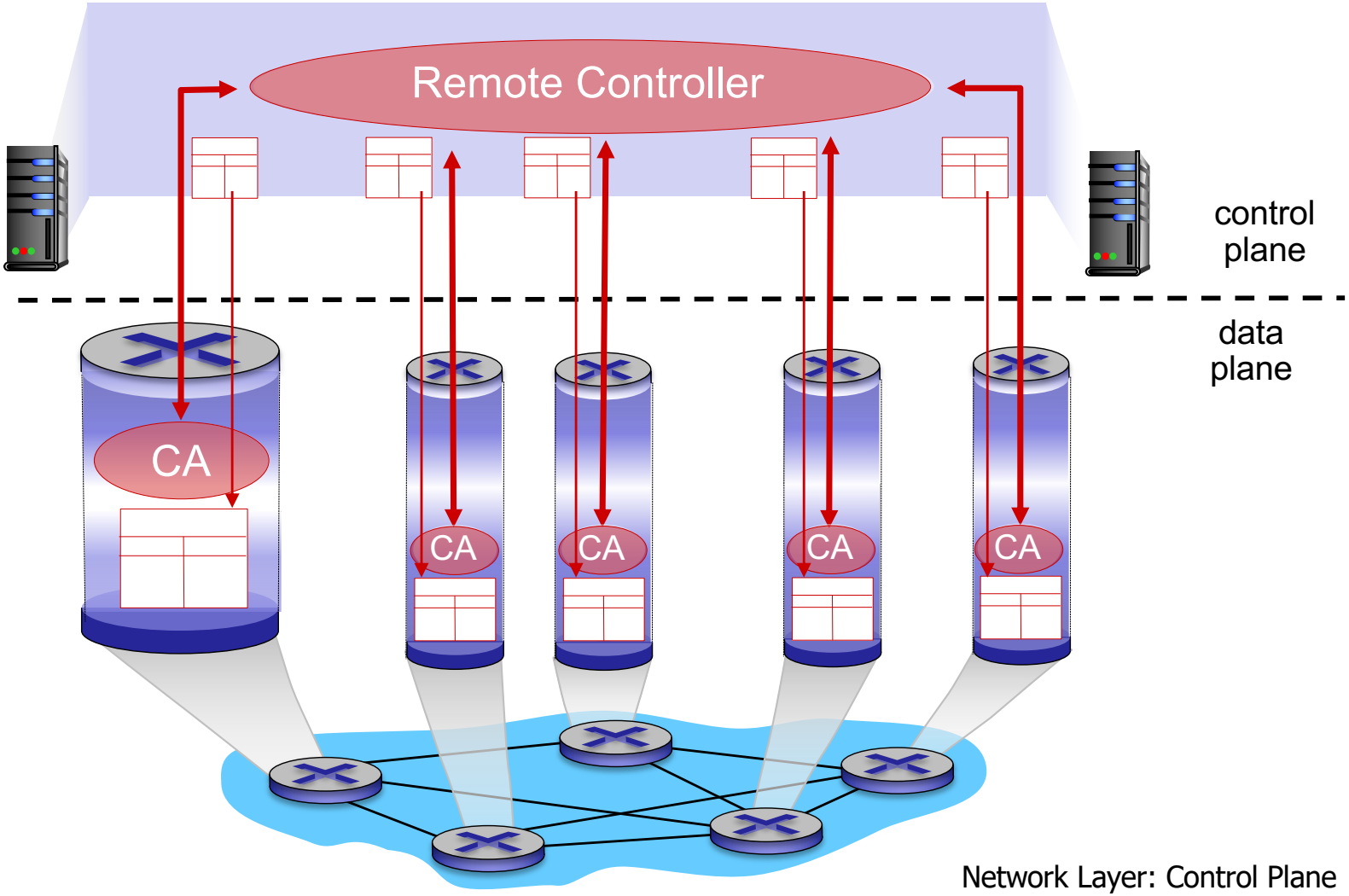
Recall: per-router control plane

Individual routing algorithm components *in each and every router* interact with each other in control plane to compute forwarding tables



Generalized Forwarding and SDN

A distinct (typically remote) controller interacts with local control agents (CAs) in routers to compute forwarding tables

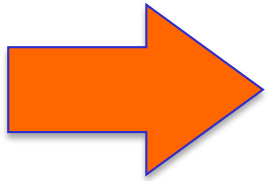


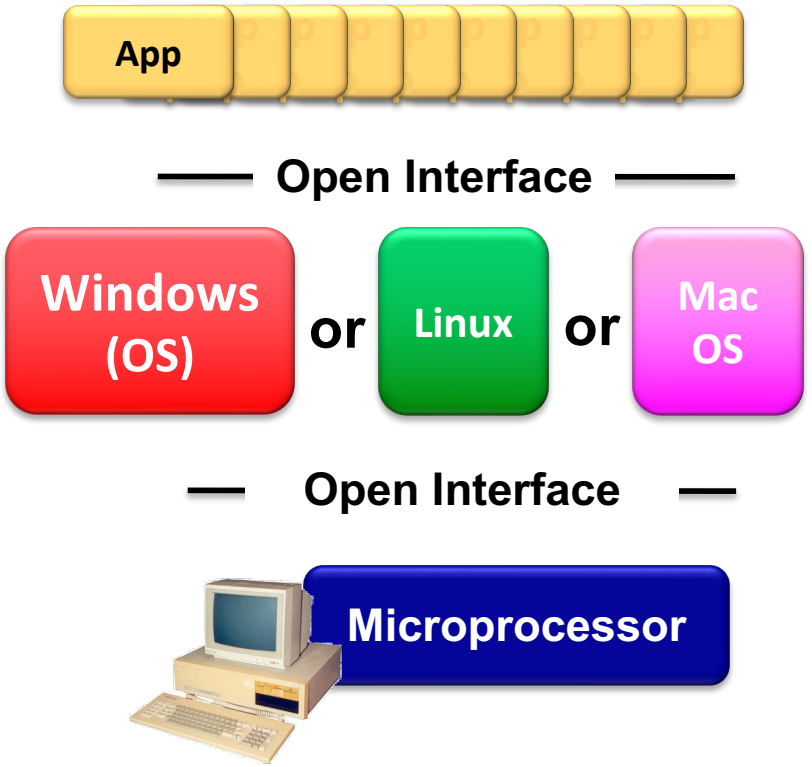
Software defined networking (SDN)

Why a *logically centralized* control plane?

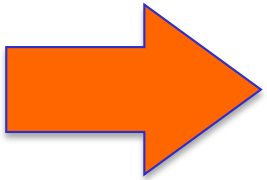
- easier network management: avoid router misconfigurations, greater flexibility of traffic flows
- centralized "programming" easier: compute tables centrally and distribute
- open (non-proprietary) implementation of control plane

Analogy: mainframe to PC evolution*





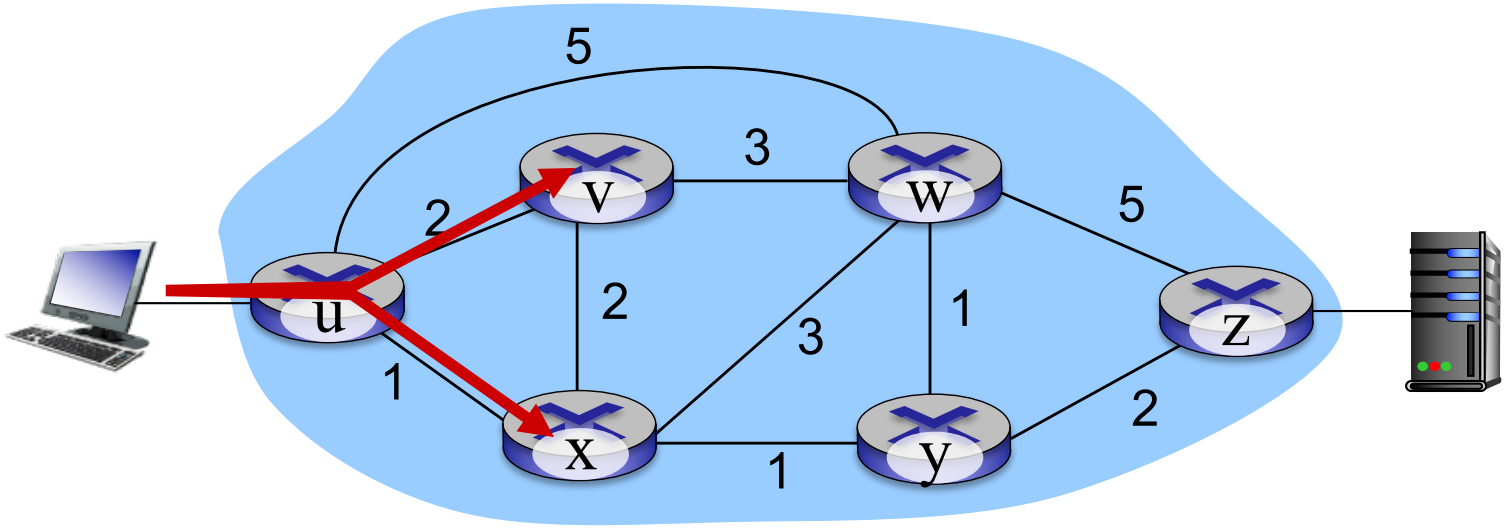
Vertically integrated
Closed, proprietary
Slow innovation
Small industry



Horizontal
Open interfaces
Rapid innovation
Huge industry

* Slide courtesy: N. McKeown

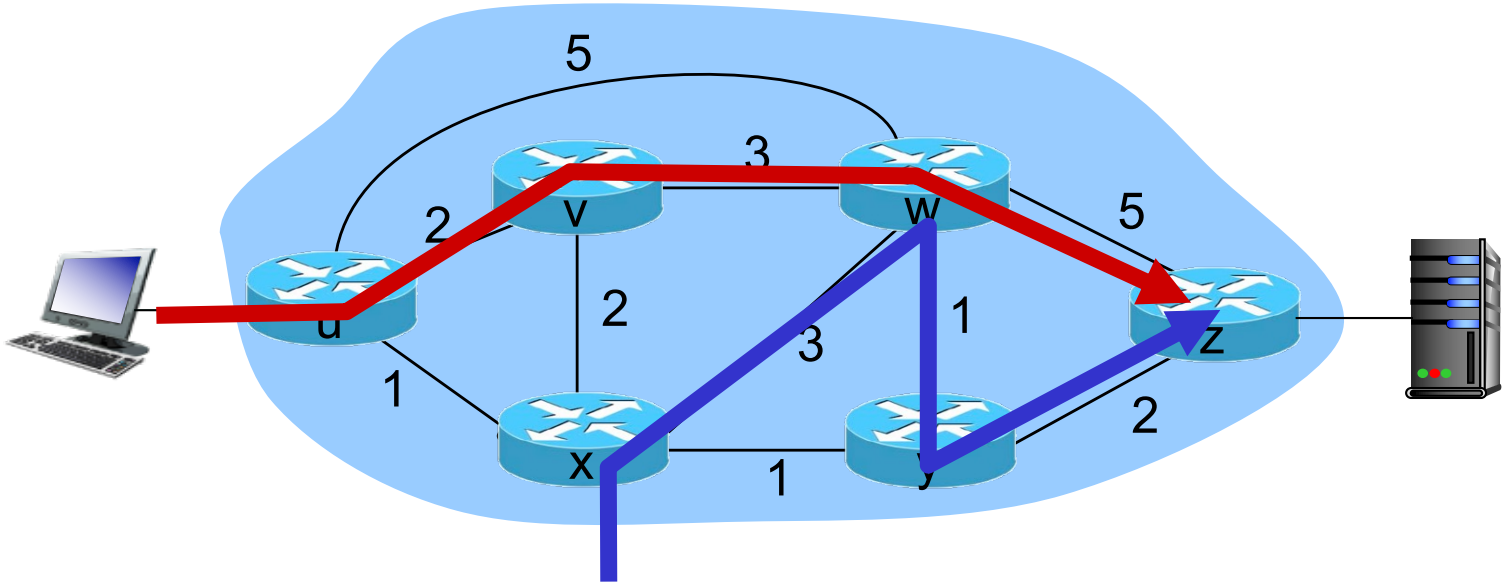
Traffic engineering: difficult



Q: what if network operator wants to split u-to-z traffic along uvwz *and* uxyz (load balancing)?

A: can't do it (or need a new routing algorithm)

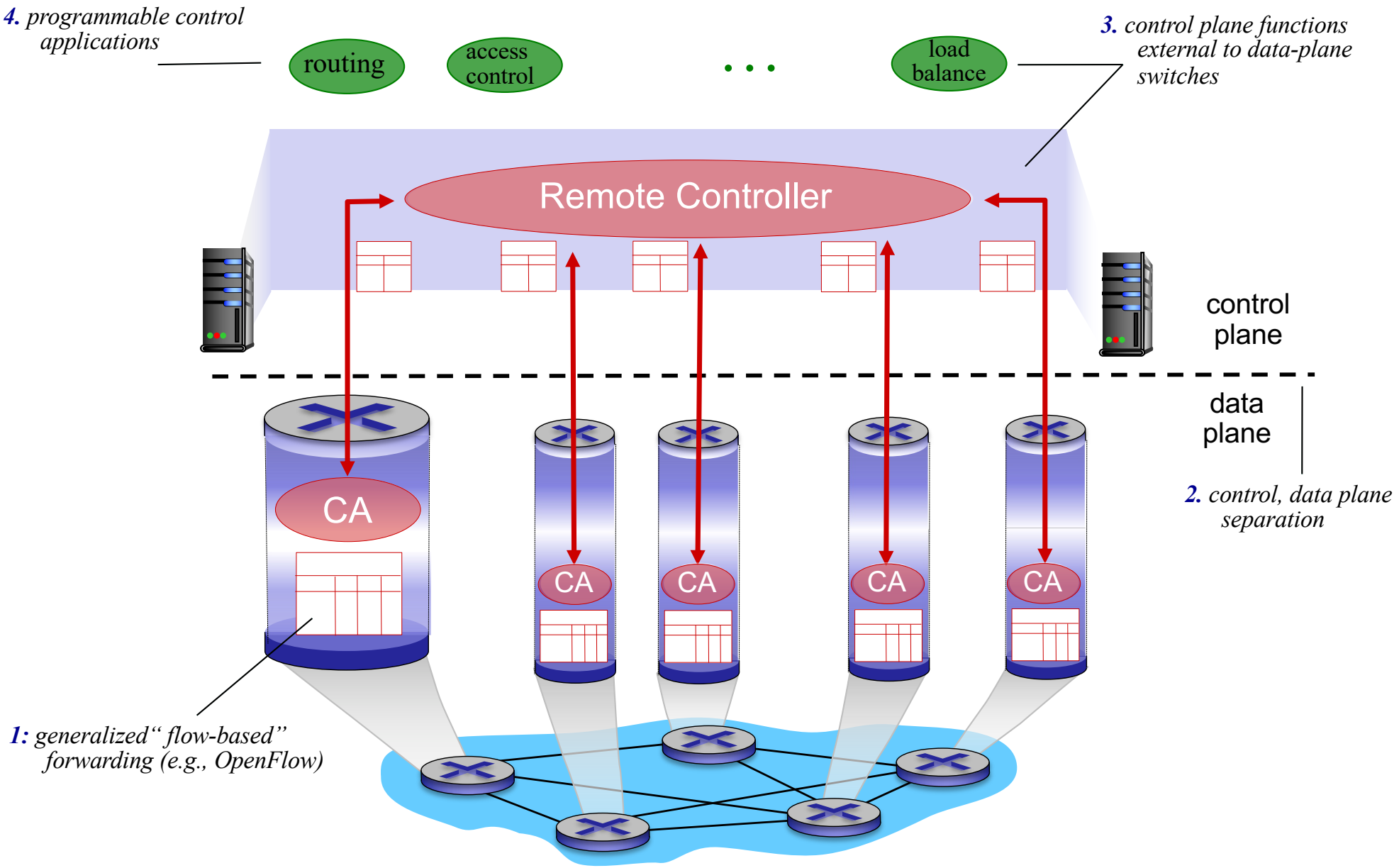
Traffic engineering: difficult



Q: what if w wants to route blue and red traffic differently?

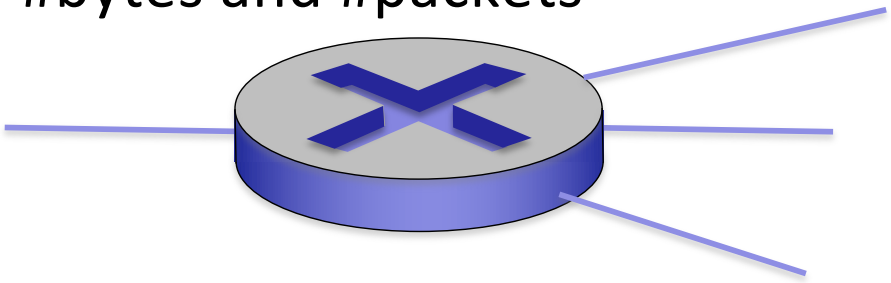
A: can't do it (with destination based forwarding, and LS, DV routing)

Software defined networking (SDN)



OpenFlow data plane abstraction

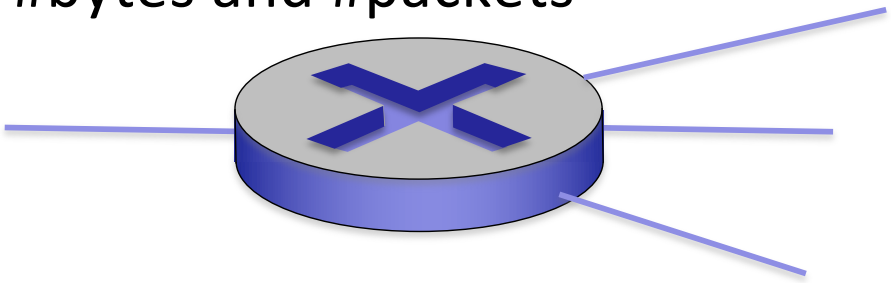
- *flow*: defined by header fields
- generalized forwarding: simple packet-handling rules
  - *Pattern:* match values in packet header fields
  - *Actions: for matched packet:* drop, forward, modify, matched packet or send matched packet to controller
  - *Priority*: disambiguate overlapping patterns
  - *Counters:* #bytes and #packets



Flow table in a router (computed and distributed by controller) define router's match+action rules

OpenFlow data plane abstraction

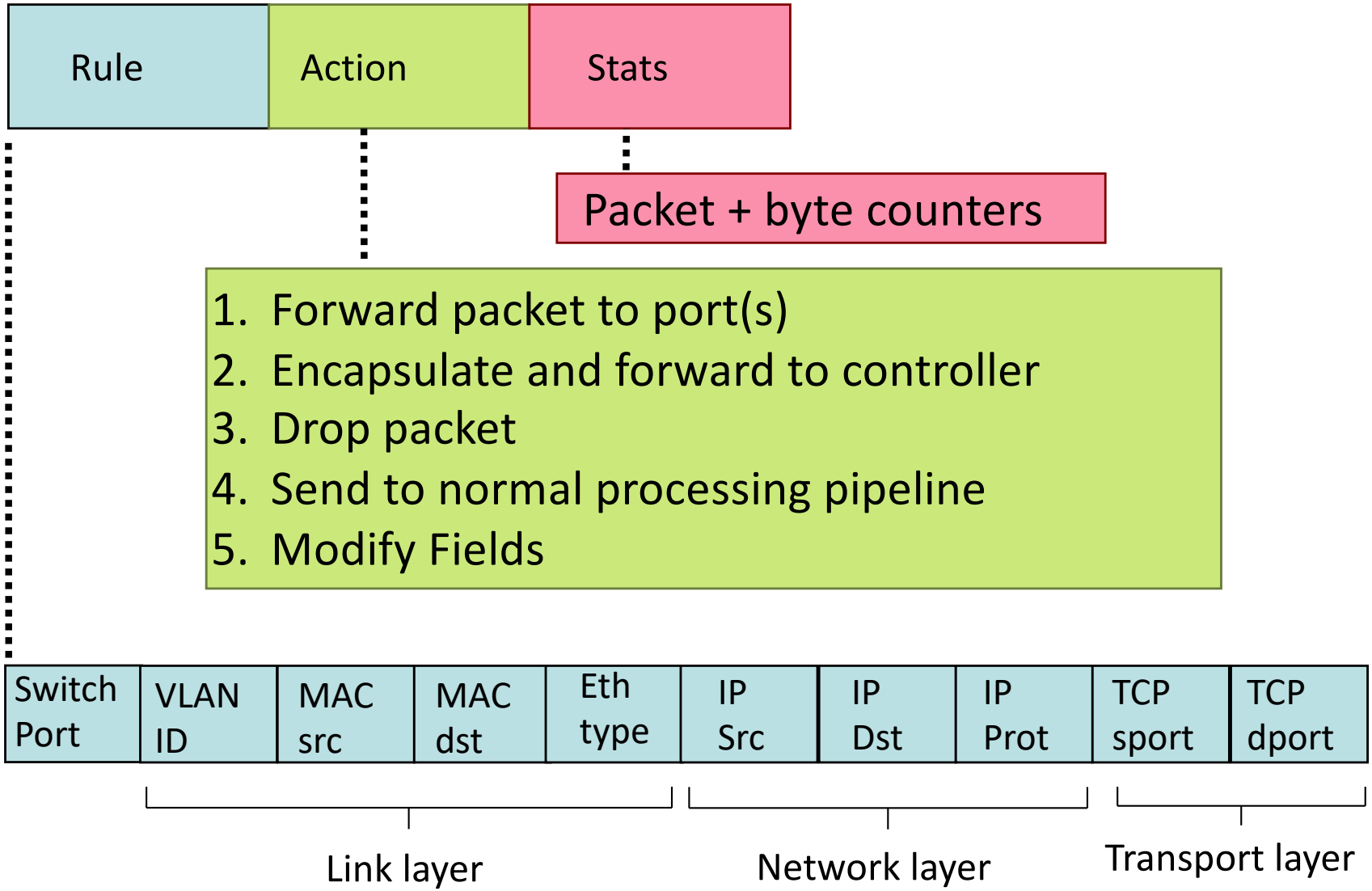
- *flow*: defined by header fields
- generalized forwarding: simple packet-handling rules
  - *Pattern:* match values in packet header fields
  - *Actions: for matched packet:* drop, forward, modify, matched packet or send matched packet to controller
  - *Priority*: disambiguate overlapping patterns
  - *Counters:* #bytes and #packets



* : wildcard

1. src=1.2.*.*, dest=3.4.5.* → drop
2. src = *.*.*.*, dest=3.4.*.* → forward(2)
3. src=10.1.2.3, dest=*.*.*.* → send to controller

OpenFlow: Flow Table Entries



Examples

Destination-based forwarding:

Switch Port	MAC src	MAC dst	Eth type	VLAN ID	IP Src	IP Dst	IP Prot	TCP sport	TCP dport	Action
*	*	*	*	*	*	51.6.0.8	*	*	*	port6

IP datagrams destined to IP address 51.6.0.8 should be forwarded to router output port 6

Firewall:

Switch Port	MAC src	MAC dst	Eth type	VLAN ID	IP Src	IP Dst	IP Prot	TCP sport	TCP dport	Forward
*	*	*	*	*	*	*	*	*	22	drop

do not forward (block) all datagrams destined to TCP port 22

Switch Port	MAC src	MAC dst	Eth type	VLAN ID	IP Src	IP Dst	IP Prot	TCP sport	TCP dport	Forward
*	*	*	*	*	128.119.1.1	*	*	*	*	drop

do not forward (block) all datagrams sent by host 128.119.1.1

OpenFlow abstraction

- *match+action:* unifies different kinds of devices

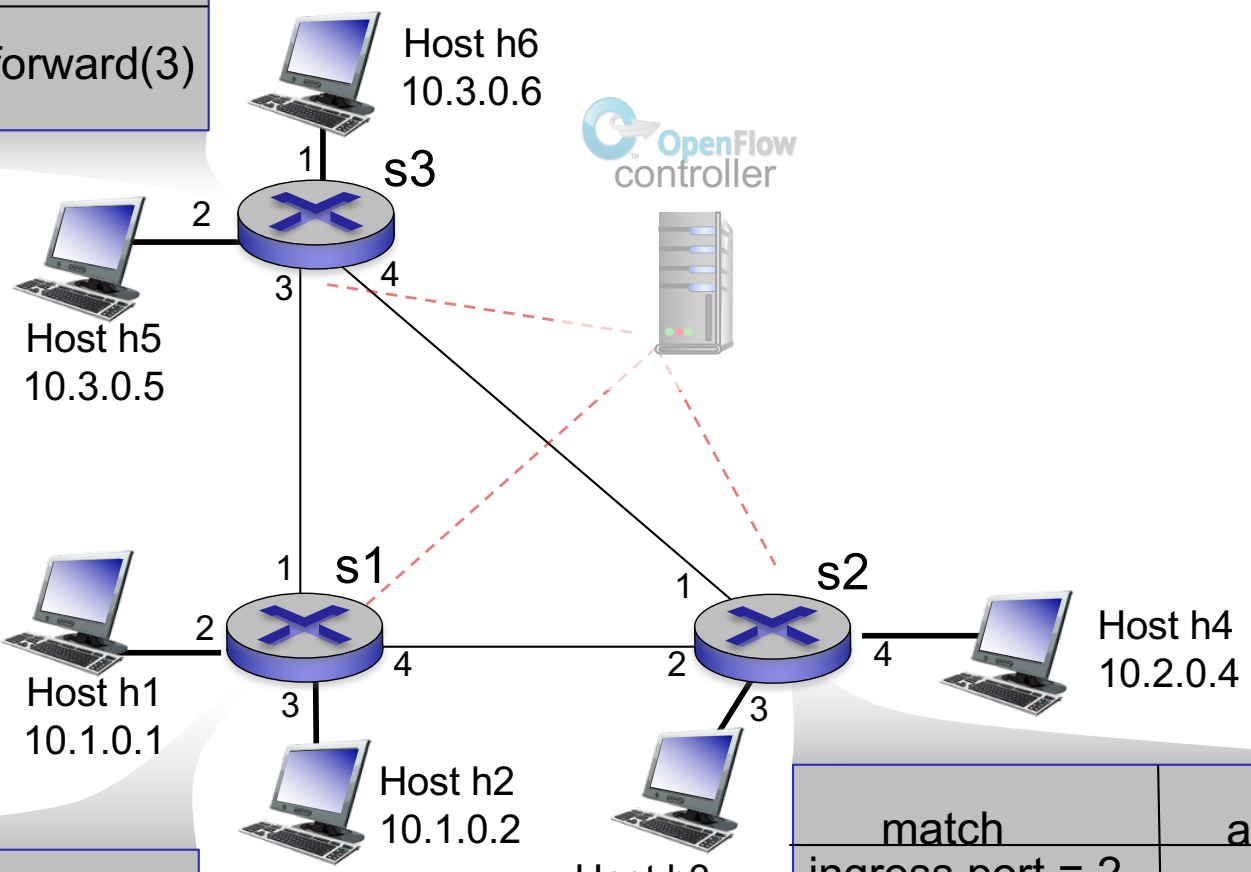
- Router
  - *match:* longest destination IP prefix
  - *action:* forward out a link
- Switch
  - *match:* destination MAC address
  - *action:* forward or flood

- Firewall
  - *match*: IP addresses and TCP/UDP port numbers
  - *action:* permit or deny
- NAT
  - *match:* IP address and port
  - *action:* rewrite address and port

OpenFlow example

Example: datagrams from hosts h5 and h6 should be sent to h3 or h4, via s1 and from there to s2

match	action
IP Src = 10.3.*.* IP Dst = 10.2.*.*	forward(3)



match	action
ingress port = 1 IP Src = 10.3.*.* IP Dst = 10.2.*.*	forward(4)

match	action
ingress port = 2 IP Src = 10.3.*.* IP Dst = 10.2.0.3	forward(3)
ingress port = 2 IP Src = 10.3.*.* IP Dst = 10.2.0.4	forward(4)