# Routing Basics

## Campus Network Design & Operations Workshop

These materials are licensed under the Creative Commons Attribution-NonCommercial 4.0 International license (http://creativecommons.org/licenses/by-nc/4.0/)

UNIVERSITY OF OREGON

Last updated 25th October 2022

NSRC Network Startup Resource Center

# What is a Router?



- A router is a layer 3 device
- Interconnects two or more networks
- Therefore, a router generally has at least two interfaces
  - With VLANs a router could have only one physical interface (known as “router on a stick”) but multiple logical interfaces
- Router looks at the destination address in the IP datagram, and decides how to forward it
- Sometimes also called a "gateway"

# The Forwarding Table

- Each router has a *forwarding table*, indicating the path for a given destination host or network
- The router tries to match the destination address of each datagram against entries in the forwarding table
- If there is a match, the router forwards it to the next-hop gateway router, or directly to the destination host

# The Forwarding Table

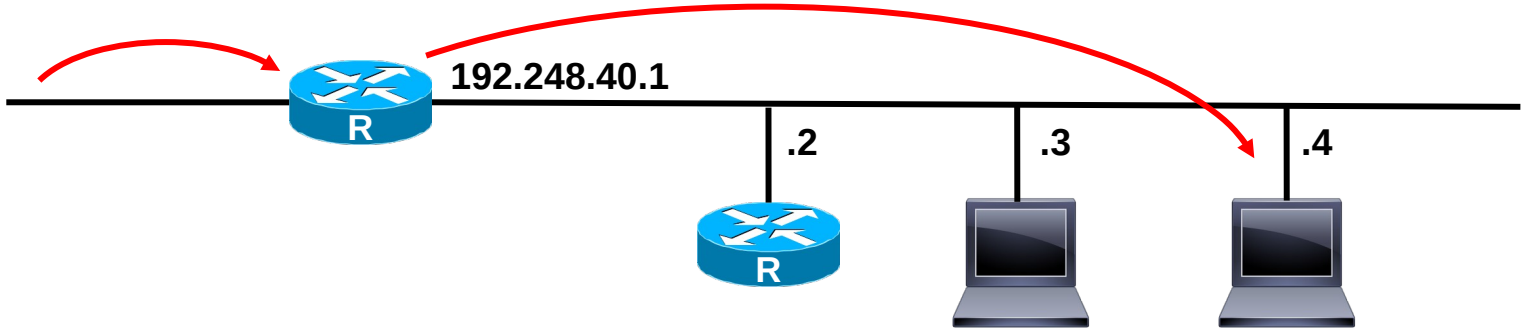
| Destination | Next-Hop | Interface |
|---|---|---|
| 10.40.0.0/16 | 192.248.40.60 | Ethernet0 |
| 192.248.0.140/30 | Directly connected | Serial1 |
| 192.248.40.0/26 | Directly connected | Ethernet0 |
| 192.248.0.0/17 | 192.248.0.141 | Serial1 |
| 203.94.73.202/32 | 192.248.40.2 | Ethernet0 |
| 203.115.6.132/30 | Directly connected | Serial0 |
| Default | 203.115.6.133 | Serial0 |

Typical forwarding table on a simple edge router

# Directly-connected route

- If the destination address is on the same subnet as one of the router's own interfaces, the router can send it directly there

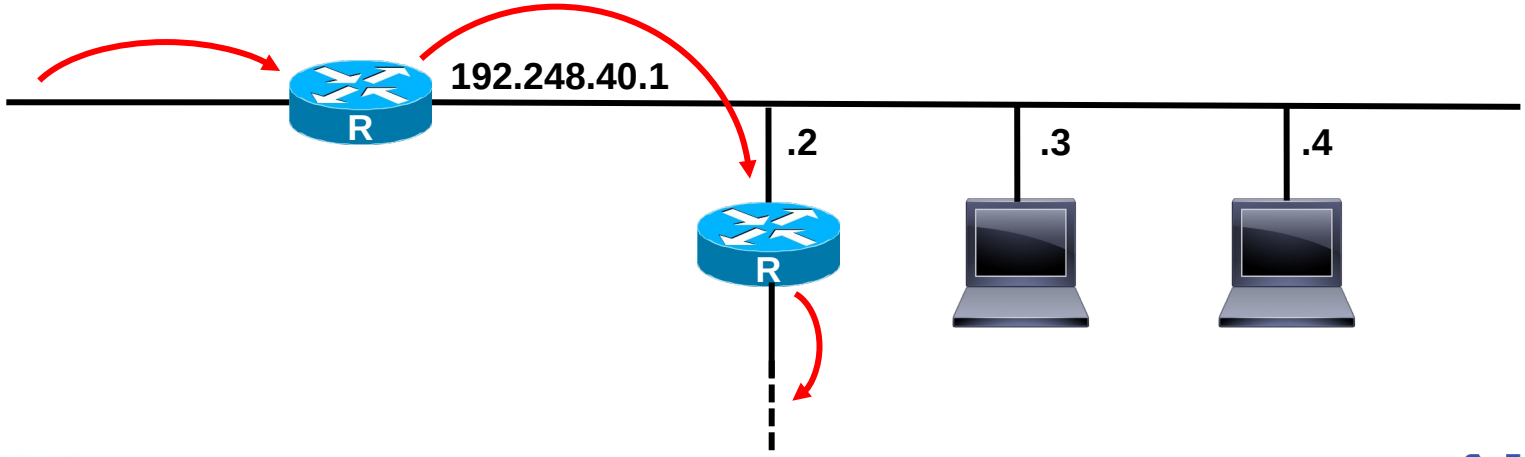
*e.g. datagram with destination address 192.248.40.4*



# Forwarding via next hop

- Otherwise, it sends to a "next hop" router
- The next hop must be on a directly connected subnet

*e.g. destination address 203.94.73.202, next hop 192.248.40.2*



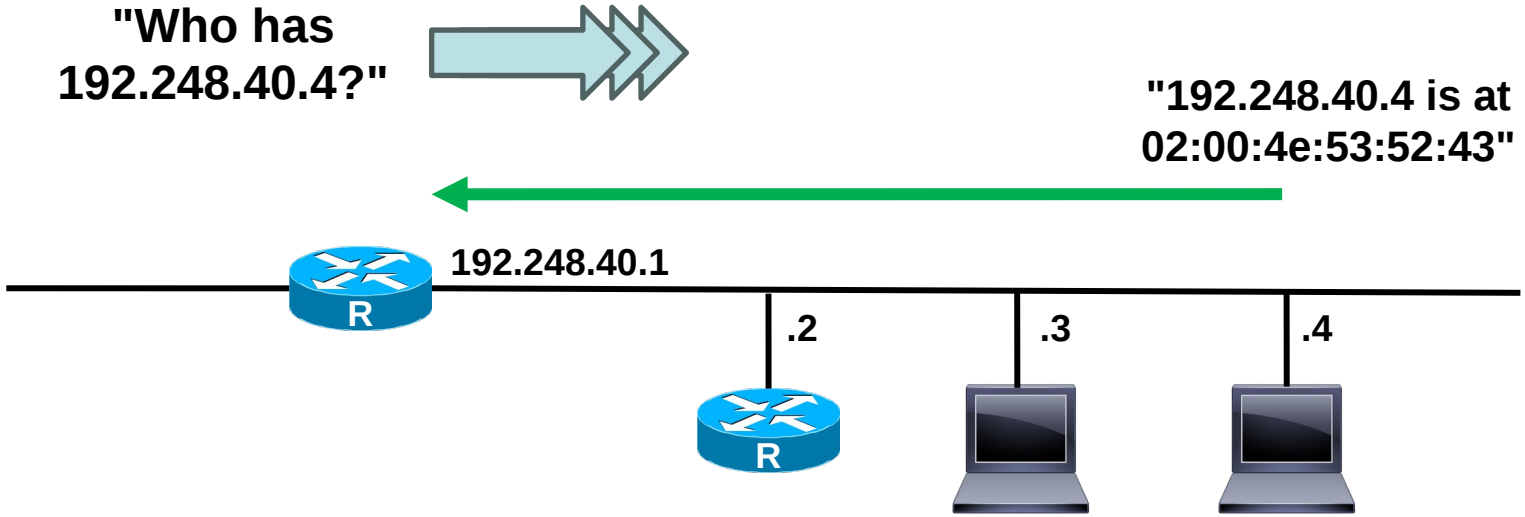
# Encapsulation

- To forward the packet over a shared medium (e.g. ethernet) the router must wrap it in an ethernet frame
  - Source MAC address = the router's interface MAC address
  - Destination MAC address = ???

# Address Resolution

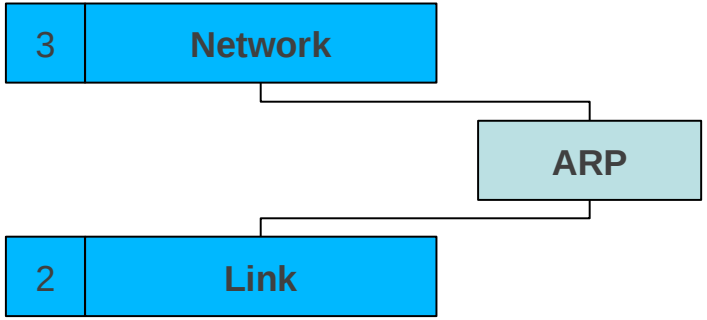
- The next-hop's MAC address must be discovered
  - IPv4: ARP (Address Resolution Protocol)
  - IPv6: NDP (Neighbor Discovery Protocol)
- Send a query for address owner in a broadcast/multicast frame; the owner of the address responds
  - The result is cached for subsequent use
  - Usually for a few minutes, although Cisco routers default to 4 hours

# Address Resolution



# Where does ARP sit in the OSI model?

- Carried inside layer 2 frames
- Provides a service to layer 3
- Is not itself a layer 3 protocol (is not routed)






*(NDP is different: it uses ICMPv6)*

# What about end hosts?

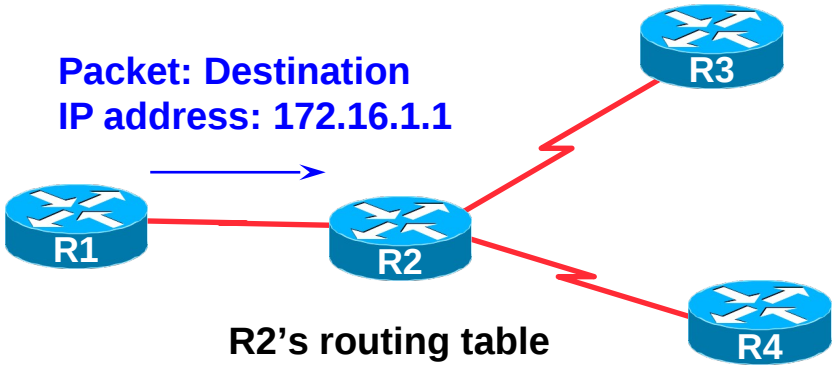
- End hosts also have a forwarding table and ARP/NDP caches
- Usually only have connected routes + default route
- Only one interface, unless "multi-homed"
  - On hosts, IP forwarding should be *disabled*

# IP route lookup

- Matches a prefix of destination IP address (first N bits)
- “Longest match” wins
  - More specific prefix preferred over less specific prefix
  - **Example**: packet with destination of 172.16.1.1 follows the route for 172.16.1.0/24 rather than the one for 172.16.0.0/16.

# IP route lookup

- Based on destination IP address

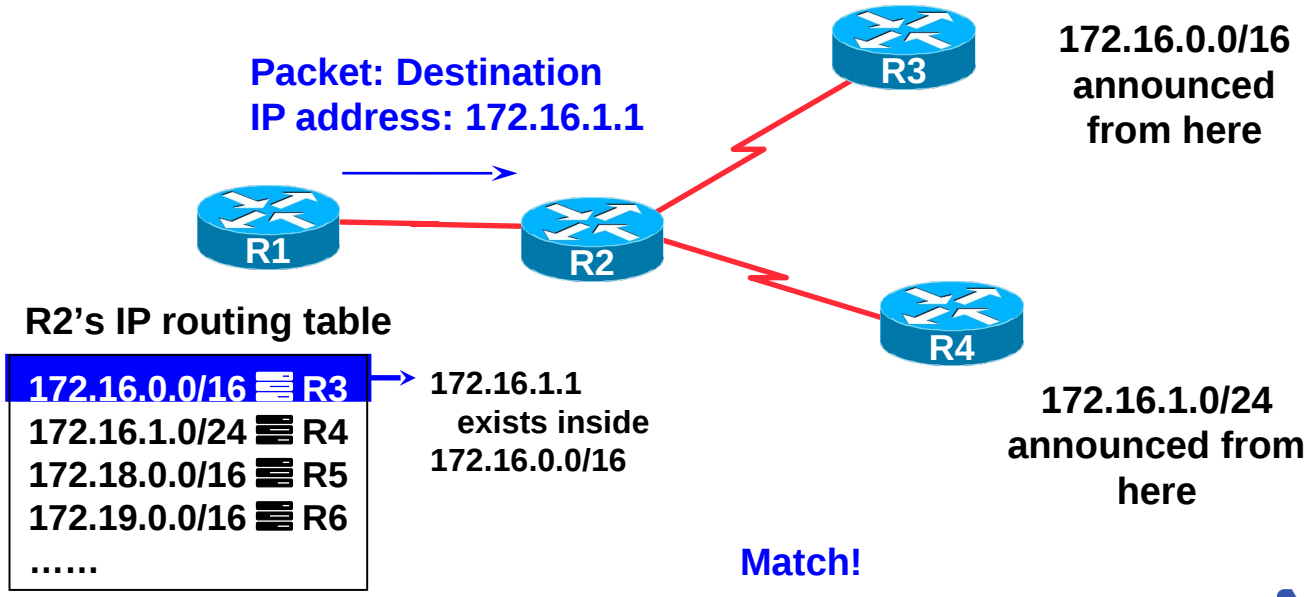



**R2's routing table**

| Destination | NextHop |
|---|---|
| 172.16.0.0/16 | R3 |
| 172.16.1.0/24 | R4 |
| 172.18.0.0/16 | R5 |
| 172.19.0.0/16 | R6 |
| …… | |

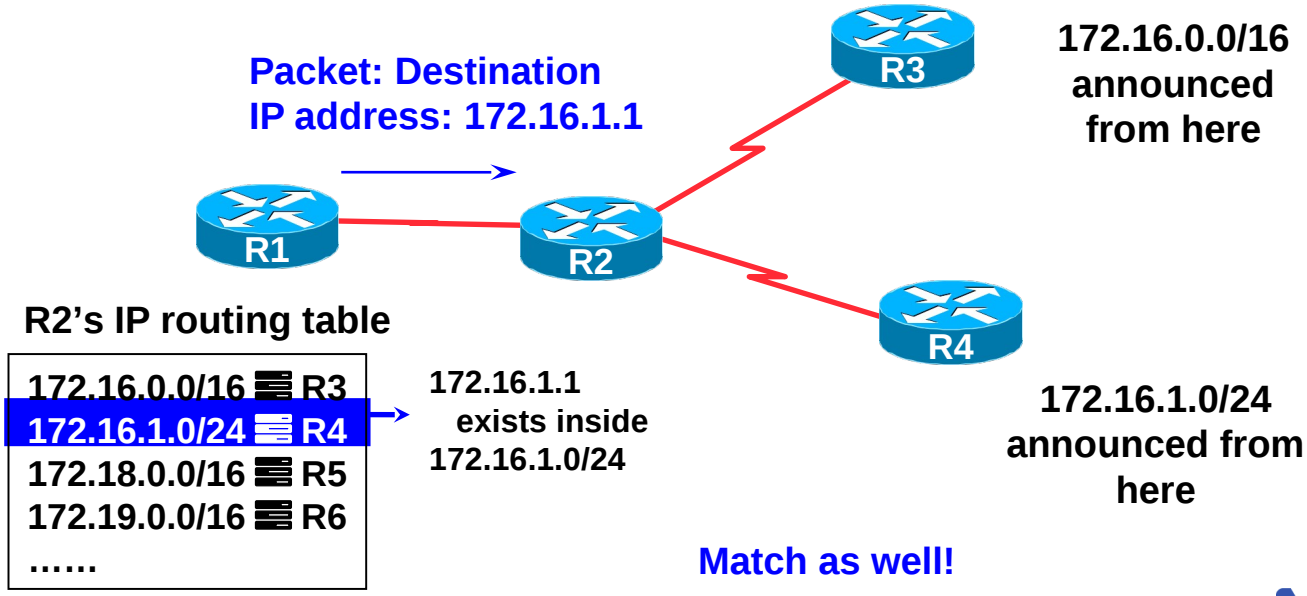
# IP route lookup: Longest match routing

- Based on destination IP address



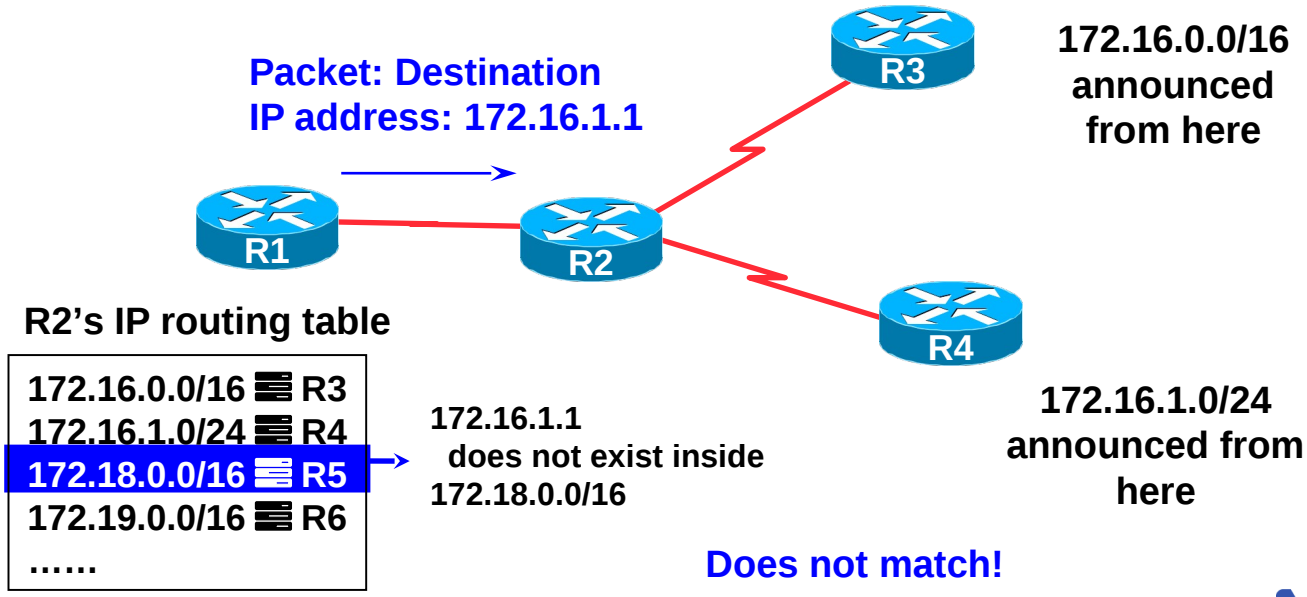
# IP route lookup: Longest match routing

- Based on destination IP address



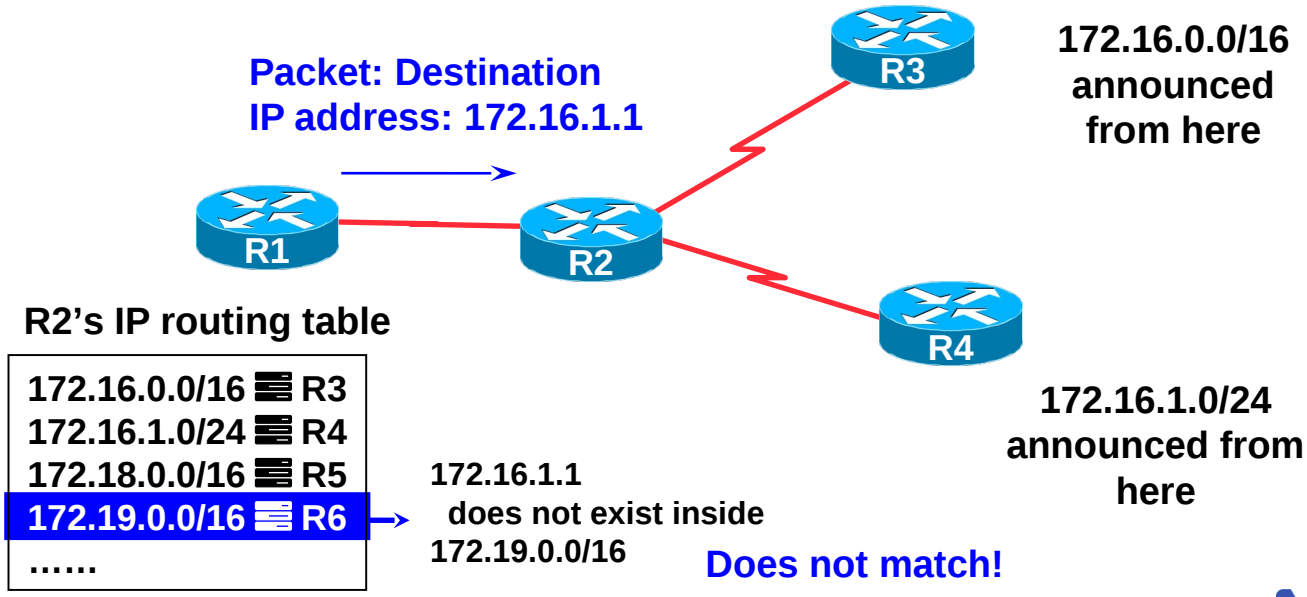
# IP route lookup: Longest match routing

- Based on destination IP address



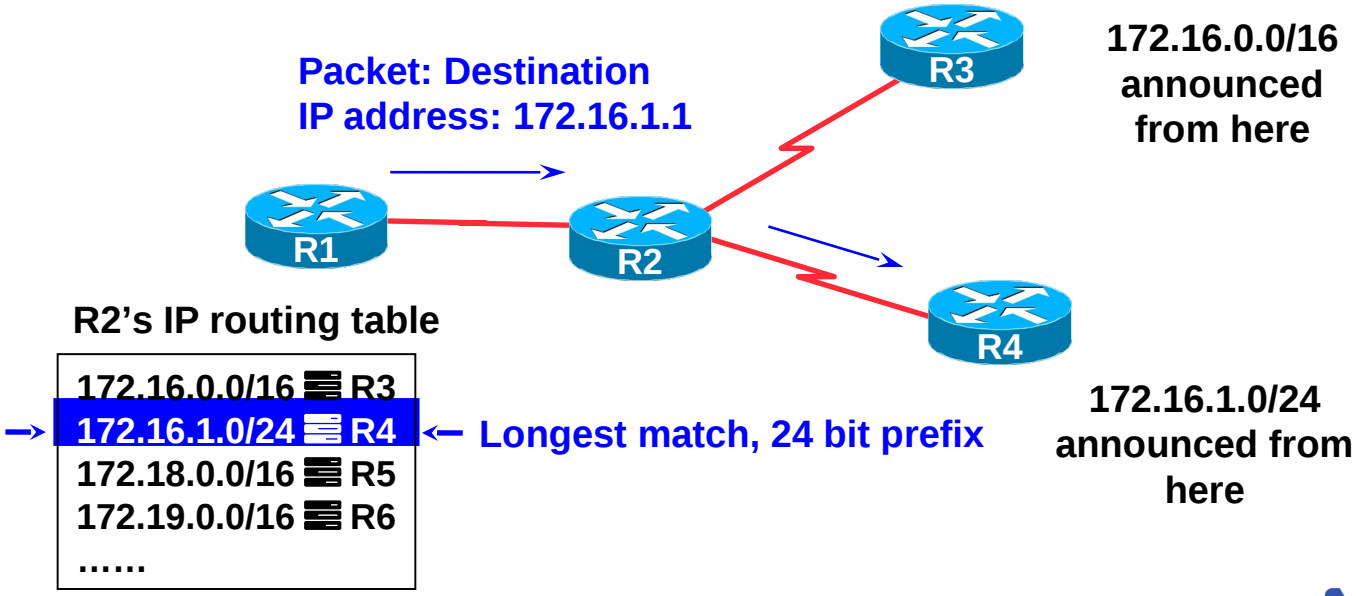
# IP route lookup: Longest match routing

- Based on destination IP address



# IP route lookup: Longest match routing

- Based on destination IP address



# Default route

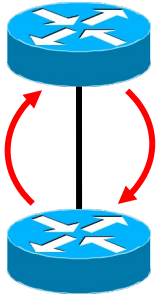
- Default route has a prefix length of zero
  - IPv4 0.0.0.0/0
  - IPv6 ::/0
- The shortest possible; only ever matches if no other route matches
- Sometimes called the "gateway of last resort"

# Forwarding is based on destination address

- Normal forwarding looks at the destination address only
- It is *possible* to configure forwarding which considers the source address as well, but please don't do it
  - This is called "policy-based routing" and is a nightmare to manage
- Access control lists (ACLs) which look at both source and destination IP addresses and/or ports are fine

# Loop prevention and Time-to-Live

- The router decrements a field in the IP header called "TTL"
  - Also updates the header checksum accordingly
- If the TTL drops to zero, the packet is discarded
  - And an ICMP "TTL exceeded" message is sent to the source address
- Avoids packets being forwarded forever with bad configurations



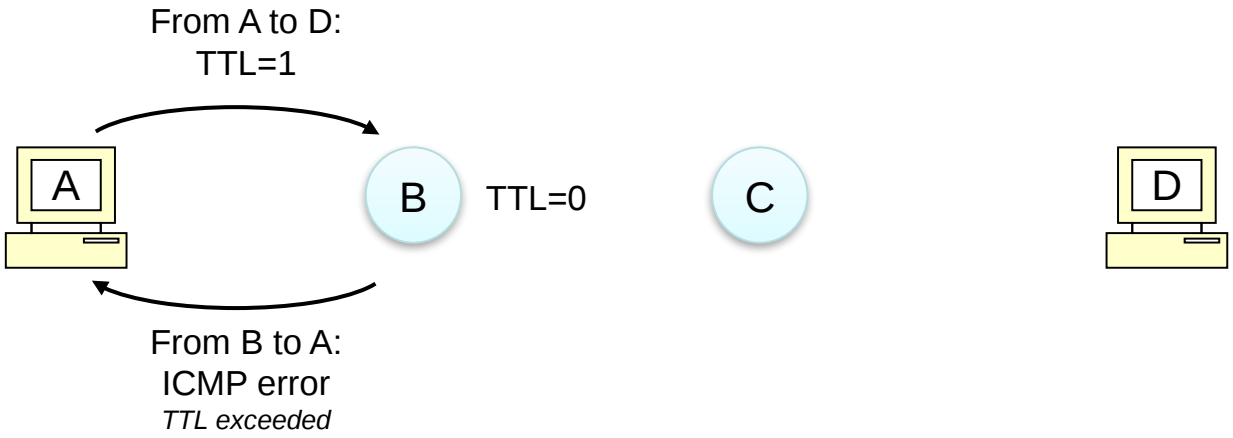
# A use of TTL: traceroute

- Tools like traceroute and mtr are able to show you the routers on the path to a destination IP address
- They do this using a trick
  - Send test packets with TTL=1
  - Then send test packets with TTL=2
  - Repeat until destination reached
- The source addresses of the ICMP "time-to-live exceeded" messages give the routers at each hop

# traceroute: first hop



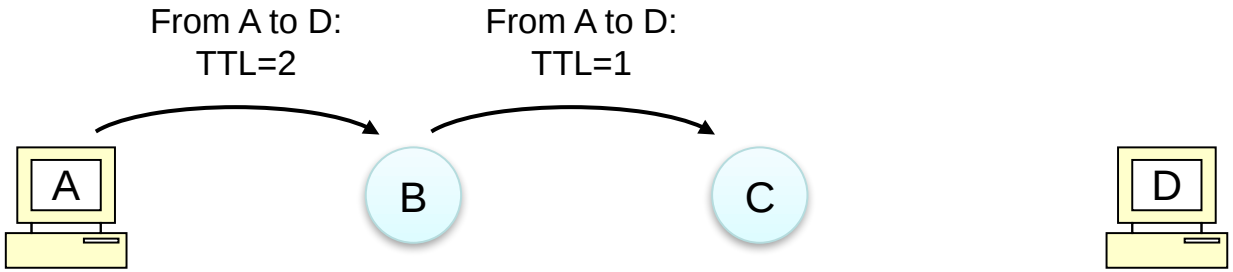
# traceroute: first hop



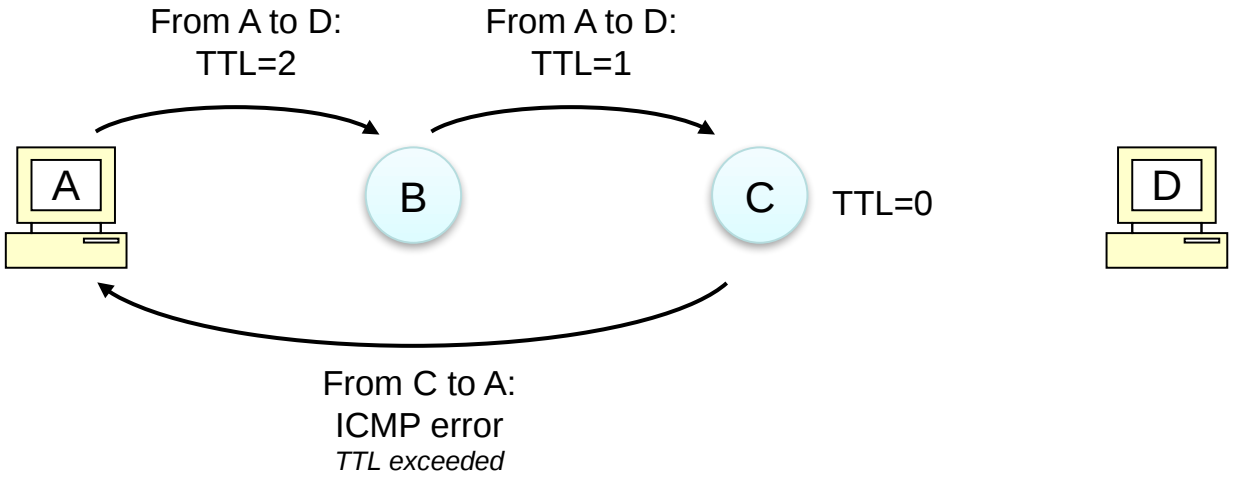
# traceroute: second hop



# traceroute: second hop



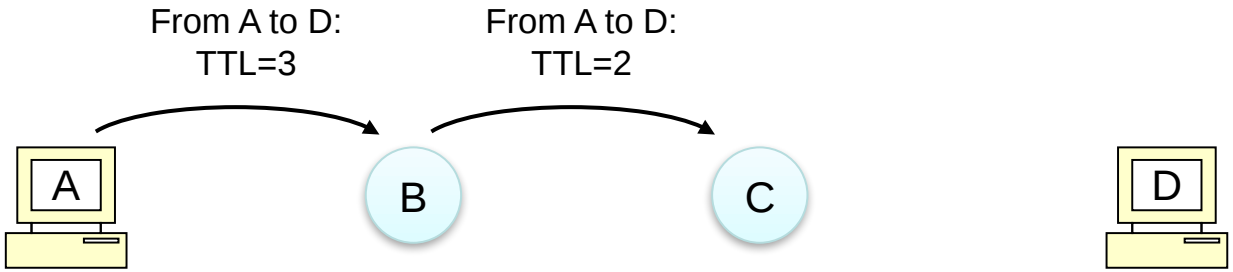
# traceroute: second hop



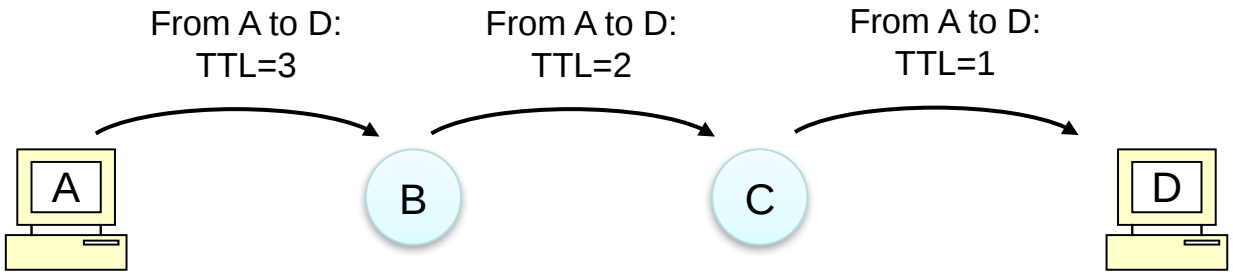
# traceroute: third hop



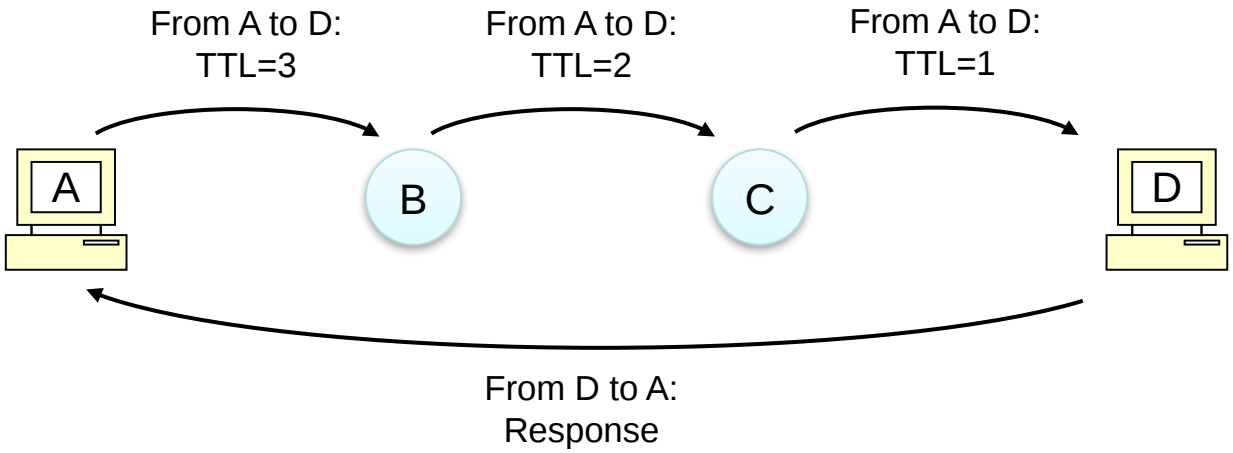
# traceroute: third hop



# traceroute: third hop



# traceroute: third hop



# traceroute notes

- The outbound path could vary from packet to packet
  - e.g. because of load-balancing or route changes
  - By default, traceroute sends 3 test packets at each TTL, so you can see if it's consistent
- You don't learn anything about the return path
  - Asymmetric routing is commonplace
  - You'd need to run a traceroute from the other end (if you can)

# Handling over-sized packets

- MTU = Maximum Transmission Unit = the largest size packet that a link can carry
- If a packet is received on an incoming link which is larger than the MTU of the outgoing link, then either:
  - the router discards it (and sends an ICMP error to source); or
  - the router fragments it into multiple smaller datagrams
  - In IPv4, controlled by "Do not Fragment" (DF) bit in header
- Fragmentation usually involves the CPU and has a very bad impact on router performance

# Questions on forwarding?

# Routing versus Forwarding

- Routing = building maps and giving directions
- Forwarding = moving packets between interfaces according to the “directions”





**Control plane**

# Routing

- Manages tables
- Low volume control traffic (e.g. routing protocols)
- Uses software (CPU)

**Data plane**

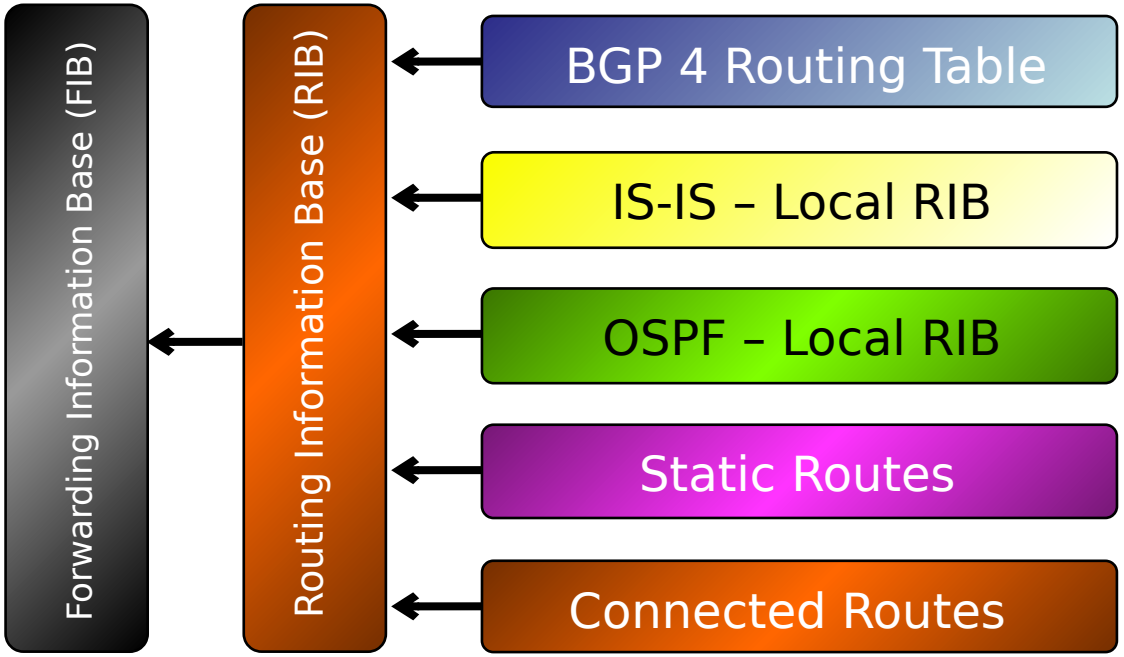
# Forwarding

- Forwards packets according to the tables
- High volume user traffic
- Uses hardware in high-end routers

# Routing – calculating the path

- Routing table entries are created by the administrator (static) and/or learned from routing protocols (dynamic)
- More than one routing protocol may run on a router
  - Each routing protocol builds its own routing table (Local RIB)
- These routes are populated into the router's Global RIB
  - If the same prefix is in multiple Local RIBs, the "administrative distance" controls which one to use – see addendum
- Dynamic routing tables are updated periodically or as topology changes (event driven)

# Routing Tables Feed the Forwarding Table



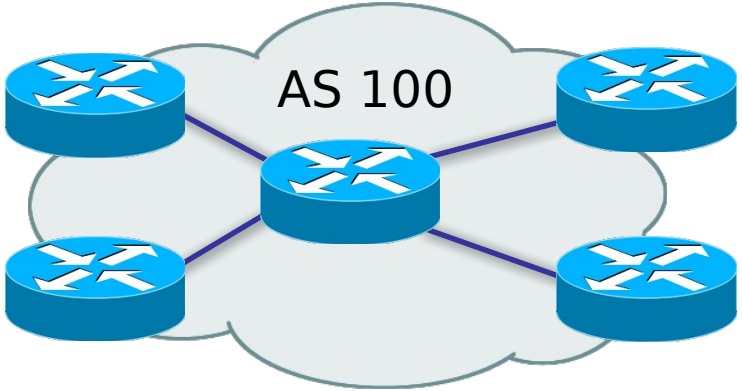
# The FIB

- FIB is the Forwarding Table
  - It contains destinations, the interfaces and the next-hops to get to those destinations (and copied to line cards on high-end routers)
  - It is built from the router's Global RIB
  - Used by the router to figure out where to send each packet
  - Cisco IOS: `show ip cef`

# The Global RIB

- The Global RIB is the Routing Table
  - Built from the routing tables/RIBs of the routing protocols and static routes on the router
  - It contains all the known destinations and the next-hops used to get to those destinations
  - One destination can have lots of possible next-hops – only the best next-hop goes into the Global RIB
  - The Global RIB is used to build the FIB
  - Cisco IOS: `show ip route`

# Autonomous System (AS)

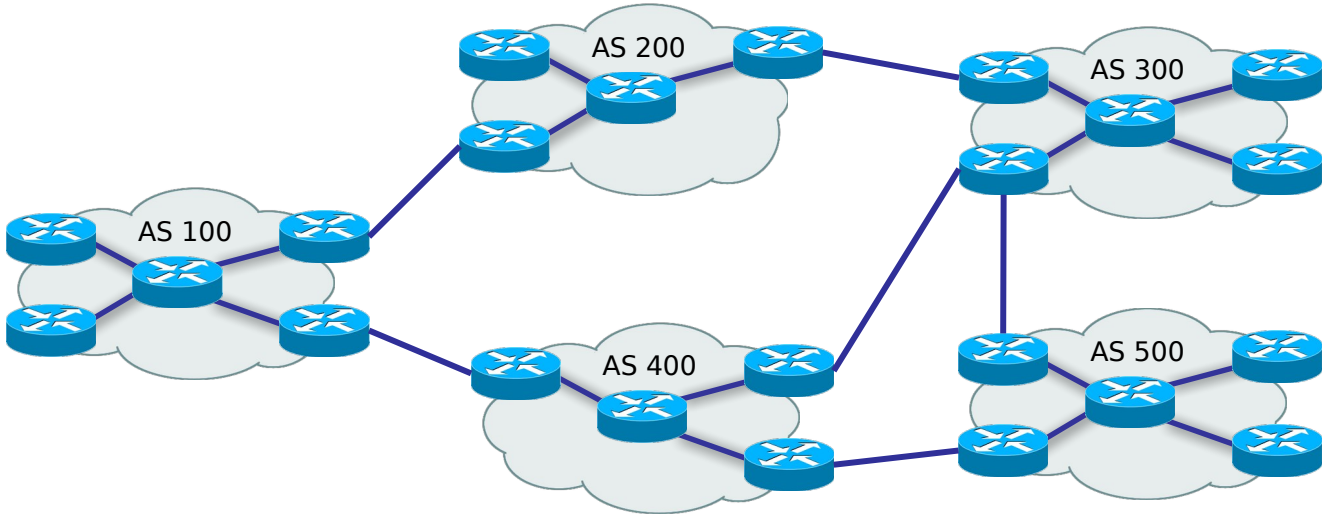



- Collection of routers with same routing policy*
- Usually under single ownership, trust and administrative control
- Examples: a campus; an NREN; an ISP

# Dynamic routing within an AS

- Use an Interior Gateway Protocol (IGP)
  - OSPF (IPv4), OSPFv3 (IPv6)
  - IS-IS (multi-protocol)
- Once IGP is properly configured, routers discover other routers and links automatically
  - Each router calculates the best path to every other destination
  - Topology changes (e.g. link up/down) automatically update

# The Internet consists of interconnected ASes



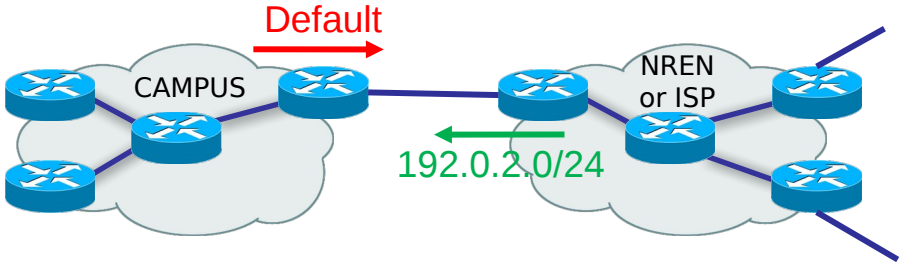

(lots of them!)

# Internet-scale routing

- IGPs do not scale to large numbers of routes, and are not good at route filtering or policy
- Internet routing uses the Border Gateway Protocol (BGP)
  - the only example of an Exterior Gateway Protocol (EGP)
- BGP is configured explicitly between neighboring ASes (eBGP)
- Use BGP inside your AS (iBGP) to distribute Internet routes and your own customer routes
- BGP only understands topology at the level of entire AS's, so you still need an IGP to manage routing within your own network

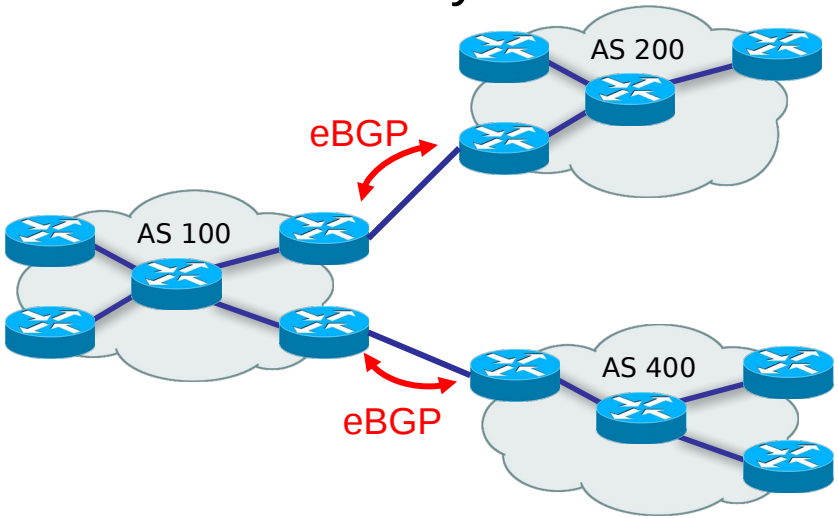
# Do you need BGP?

- If there is only one link out of your network, then a static default route is all you need
- Your ISP configures a static route for your address block



# Do you need BGP?

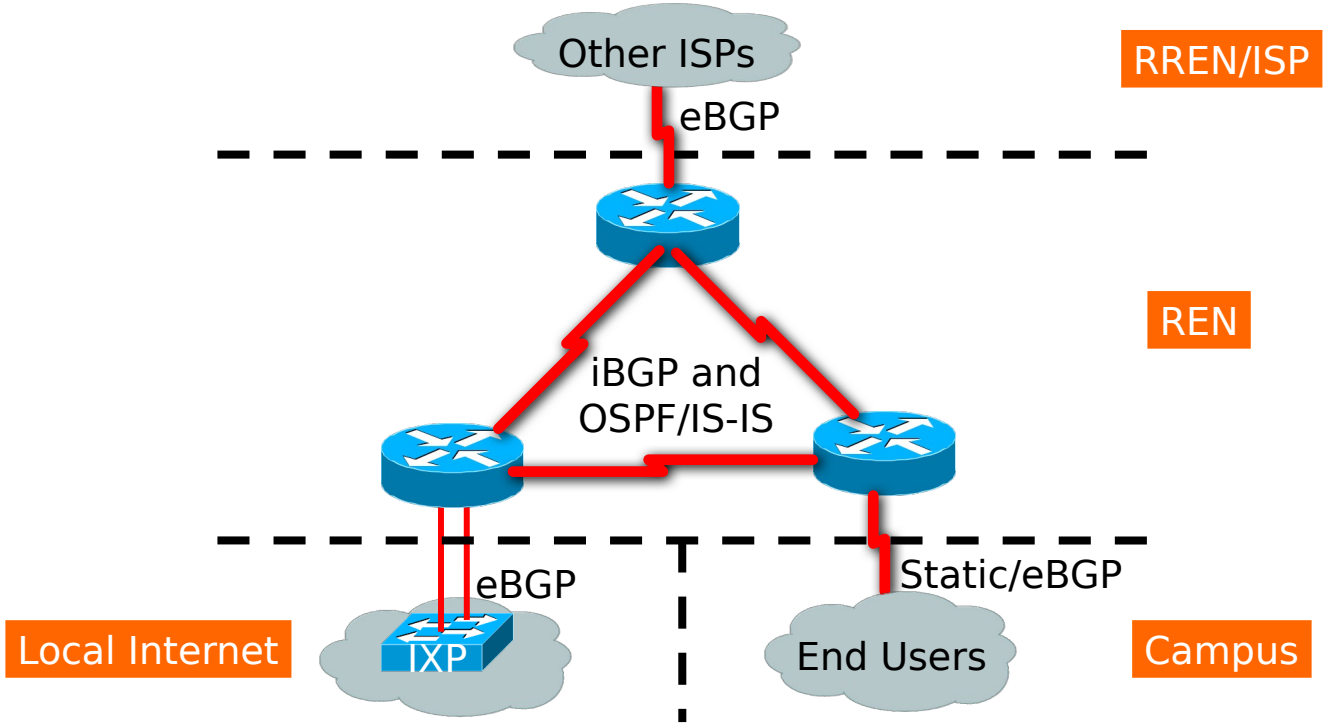
- If you have two or more external links then you need BGP
- You announce your network via BGP, and your providers announce Internet routes to you



# Considerations for BGP

- You can choose to receive just a default route; a full Internet routing table; or default route plus some subset of routes (e.g. only R&E routes, or only in-country routes)
  - The more routes you receive, the more control you have over your outbound traffic flows, but the more RAM and CPU is required
  - A full Internet table requires *a lot* of RAM and CPU!
- If you are running an ISP or NREN network then you are part of the "default free zone" and you will need to carry the full table

# Hierarchy of Routing Protocols



# Protecting the control plane

- All packets where the destination address is one of the router's own interface addresses are diverted to the CPU
  - e.g. management traffic (ssh, snmp); routing protocols; pings to router IP
- The router's CPU can be attacked or overloaded (DoS)
- It should be protected, e.g. using ACLs
- Router will handle such traffic at lower priority than forwarding
  - you may see long response times at certain hops in traceroute; this is normal

# Questions?

# FYI: Default Administrative Distances

| Route Source | Cisco | Juniper | Huawei | Dell | Nokia | Mikrotik |
|---|---|---|---|---|---|---|
| Connected Interface | 0 | 0 | 0 | 0 | 0 | 0 |
| Static Route | 1 | 5 | 60 | 1 | 1 | 1 |
| EIGRP Summary Route | 5 | N/A | ? | N/A | N/A | N/A |
| External BGP | 20 | 170 | 255 | 20 | 170 | 20 |
| Internal EIGRP Route | 90 | N/A | ? | N/A | N/A | N/A |
| IGRP | 100 | N/A | ? | N/A | N/A | N/A |
| OSPF | 110 | 10 | 10 | 110 | 10 | 110 |
| IS-IS | 115 | 18 | 15 | 115 | 18 | 115 |
| RIP | 120 | 100 | 100 | 120 | 100 | 120 |
| EGP | 140 | N/A | N/A | N/A | N/A | N/A |
| External EIGRP | 170 | N/A | ? | N/A | N/A | N/A |
| Internal BGP | 200 | 170 | 255 | 200 | 130 | 200 |
| Unknown | 255 | 255 | ? | 255 | ? | |

