# Routing Basics

ISP Workshops



Last updated 2nd August 2019

# Acknowledgements

- This material originated from the Cisco ISP/IXP Workshop Programme developed by Philip Smith & Barry Greene

- Use of these materials is encouraged as long as the source is fully acknowledged and this notice remains in place

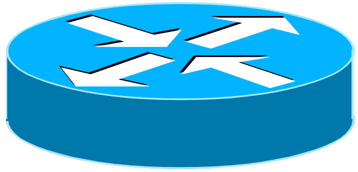
- Bug fixes and improvements are welcomed
  - Please email *workshop (at) bgp4all.com*

Philip Smith

# Routing Concepts

- Routers
- Routing
- Forwarding
- Some definitions
- Policy options
- Routing Protocols

# What is a Router?



- A router is a layer 3 device
- Used for interconnecting networks at layer 3
- A router generally has at least two interfaces
  - With VLANs a router can have only one interface (known as "router on a stick")
- A router looks at the destination address in the IP packet, and decides how to forward it

# The Routing Table

- Each router/host has a *routing table*, indicating the path or the next hop for a given destination host or a network
- The router/host tries to match the destination address of a packet against entries in the routing table
- If there is a match, the router forwards it to the corresponding gateway router or directly to the destination host
- Default route is taken if no other entry matches the destination address

# The Routing 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 | 192.248.40.3 | Ethernet0 |
| 203.115.6.132/30 | Directly connected | Serial0 |
| Default | 203.115.6.133 | Serial0 |

Typical routing table on a simple edge router

# IP Routing – finding the path

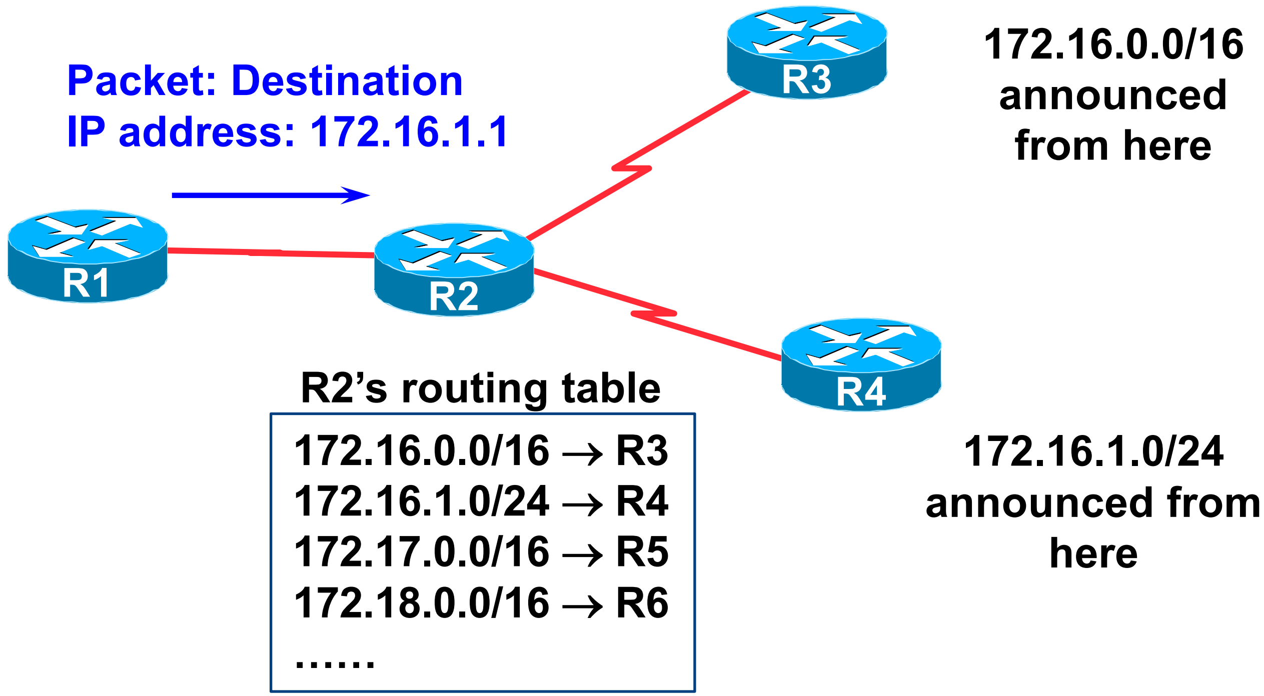
- Routing table entry (the path) is created by the administrator (static) or received from a routing protocol (dynamic)
- More than one routing protocol may run on a router
  - Each routing protocol builds its own routing table (Local RIB)
- Several alternative paths may exist
  - Best path selected for the router's Global routing table (RIB)
- Decisions are updated periodically or as topology changes (event driven)
- Decisions are based on:
  - Topology, policies and metrics (hop count, filtering, delay, bandwidth, etc.)

# IP route lookup

- Based on destination IP address
- "longest match" routing
  - More specific prefix preferred over less specific prefix
  - **Example**: packet with destination of
  - 172.16.1.1/32 is sent to the router announcing 172.16.1.0/24 rather than the router announcing 172.16.0.0/16.

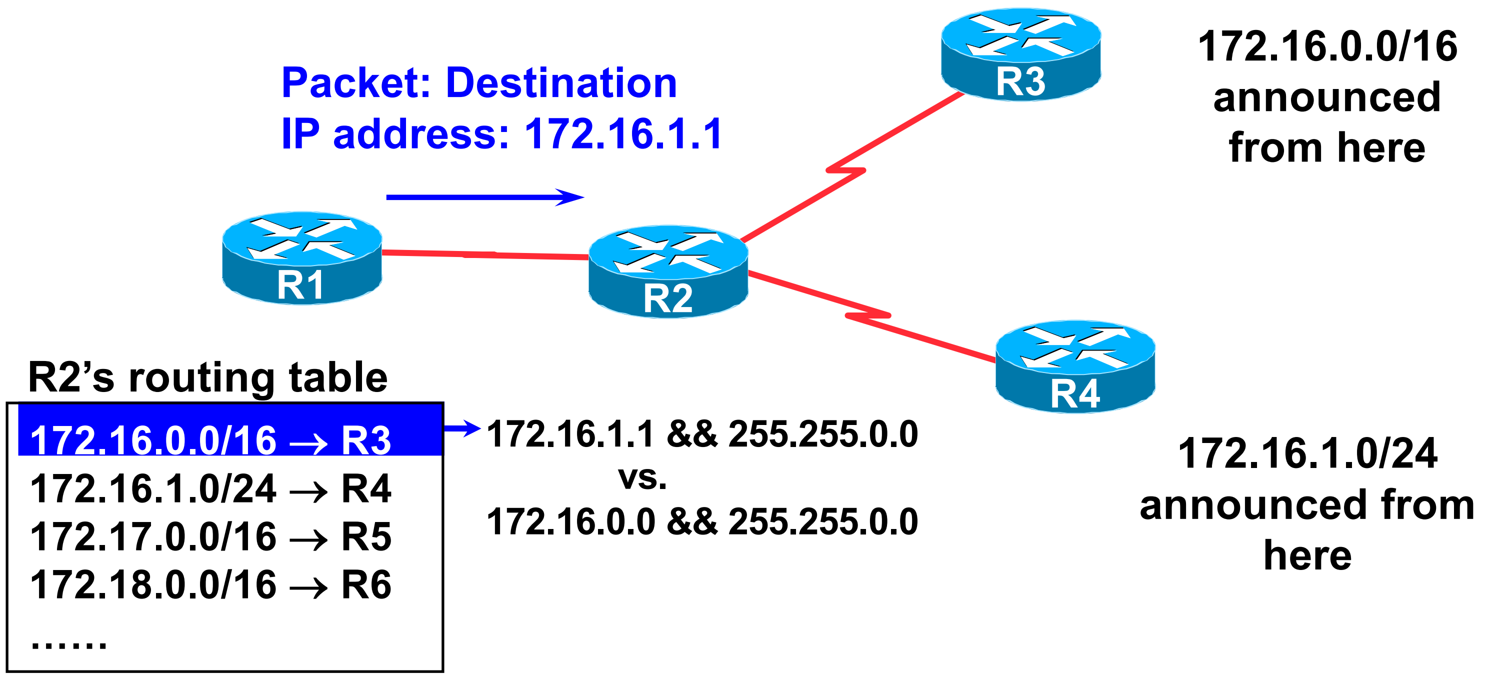
# IP route lookup

- Based on destination IP address



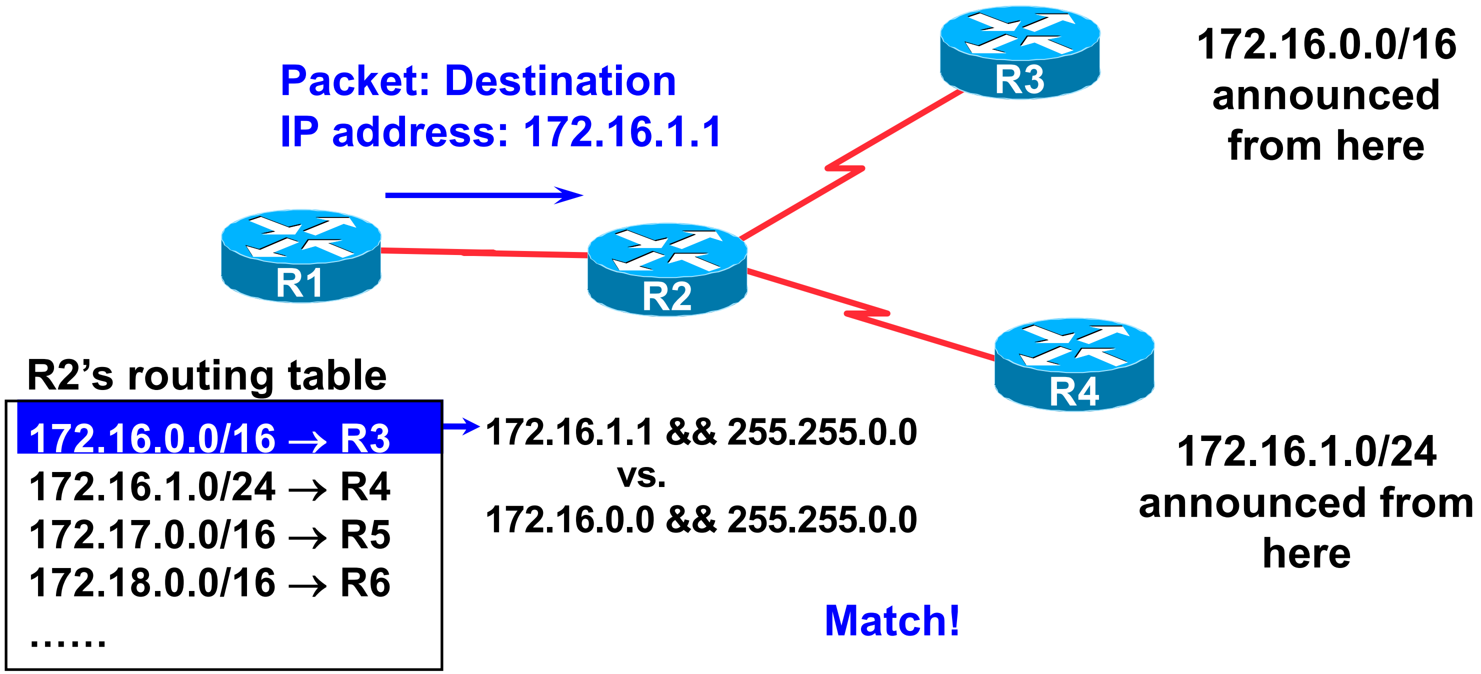
# IP route lookup: Longest match routing

- Based on destination IP address



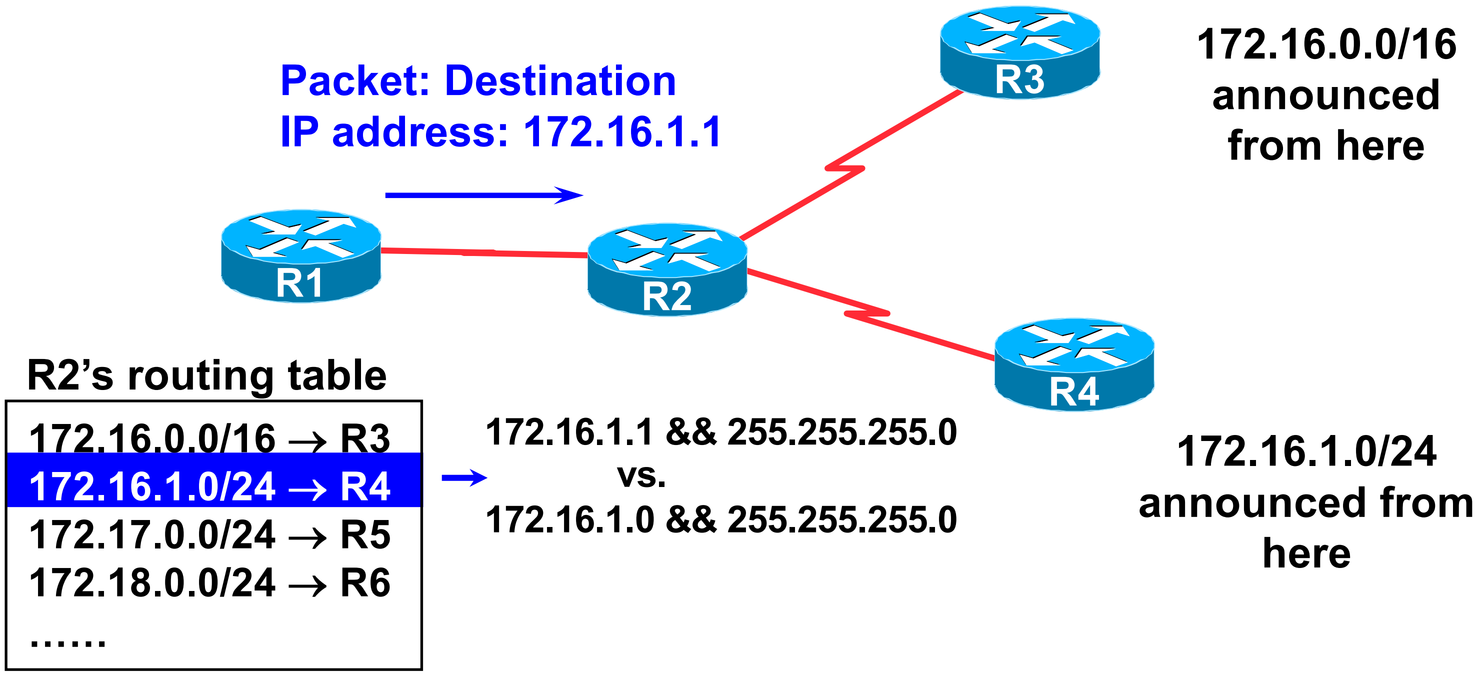
# IP route lookup: Longest match routing

- Based on destination IP address



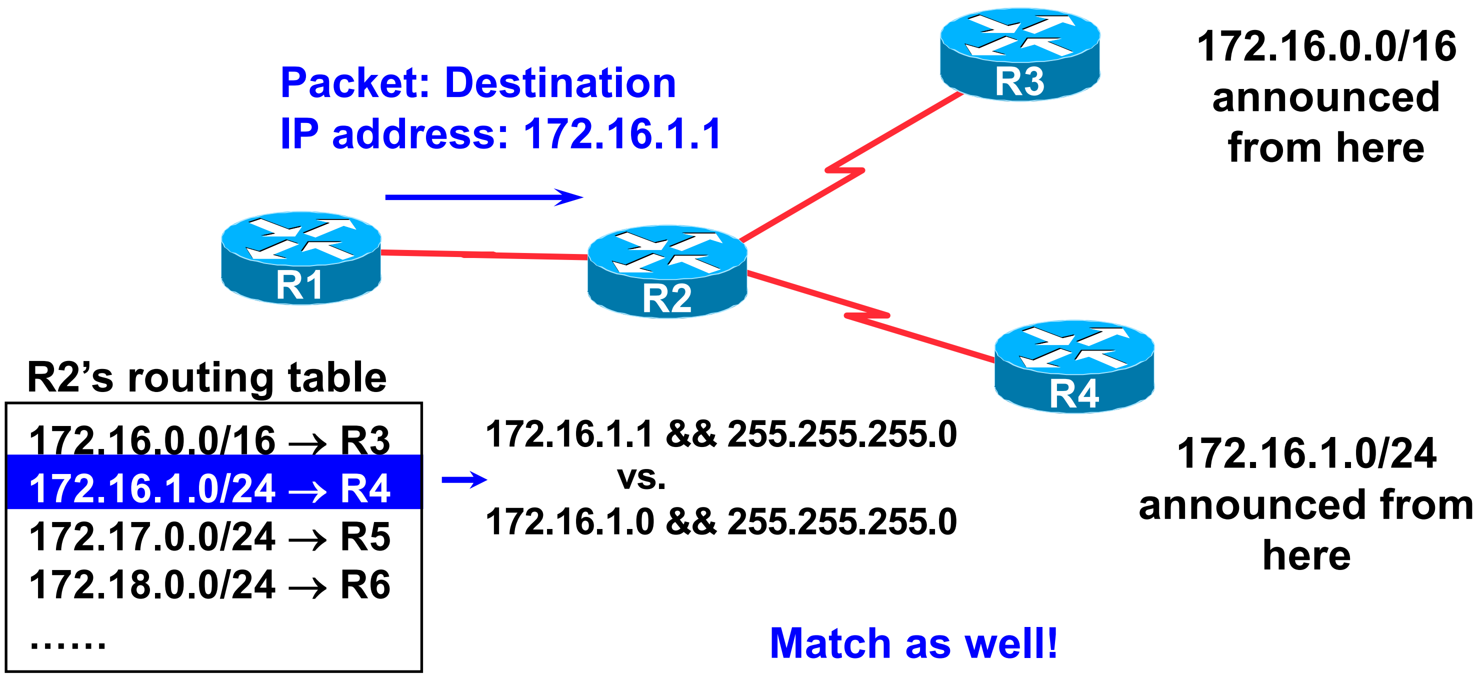
# IP route lookup: Longest match routing

- Based on destination IP address



# IP route lookup: Longest match routing

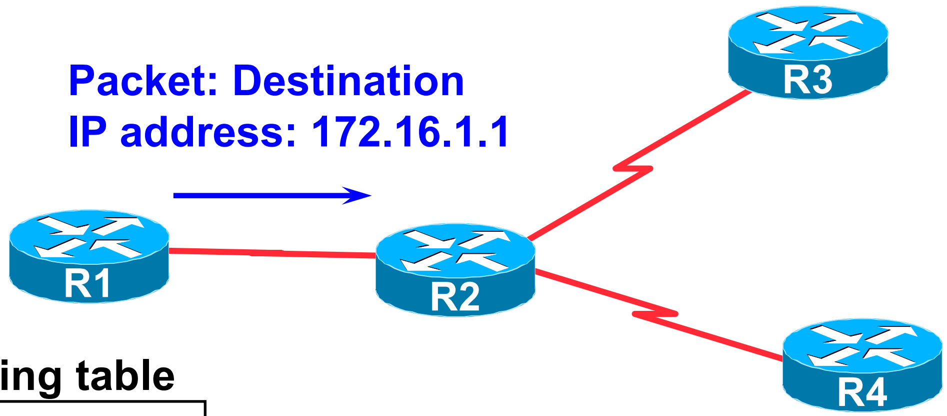
- Based on destination IP address



# IP route lookup: Longest match routing

- Based on destination IP address

**Packet: Destination IP address: 172.16.1.1**




**172.16.0.0/16 announced from here**

**172.16.1.0/24 announced from here**

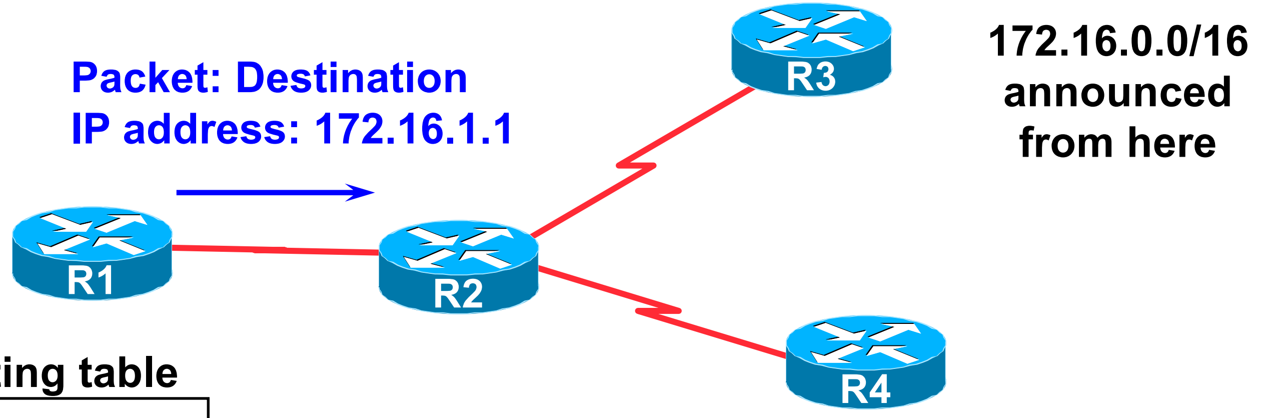
**R2's routing table**

| |
|---|
| 172.16.0.0/16 → R3 |
| 172.16.1.0/24 → R4 |
| **172.17.0.0/16 → R5** |
| 172.18.0.0/16 → R6 |
| …… |

**172.16.1.1 && 255.255.0.0**
**vs.**
**172.17.0.0 && 255.255.0.0**

# IP route lookup: Longest match routing

- Based on destination IP address




172.16.1.0/24
announced from
here

**R2's routing table**

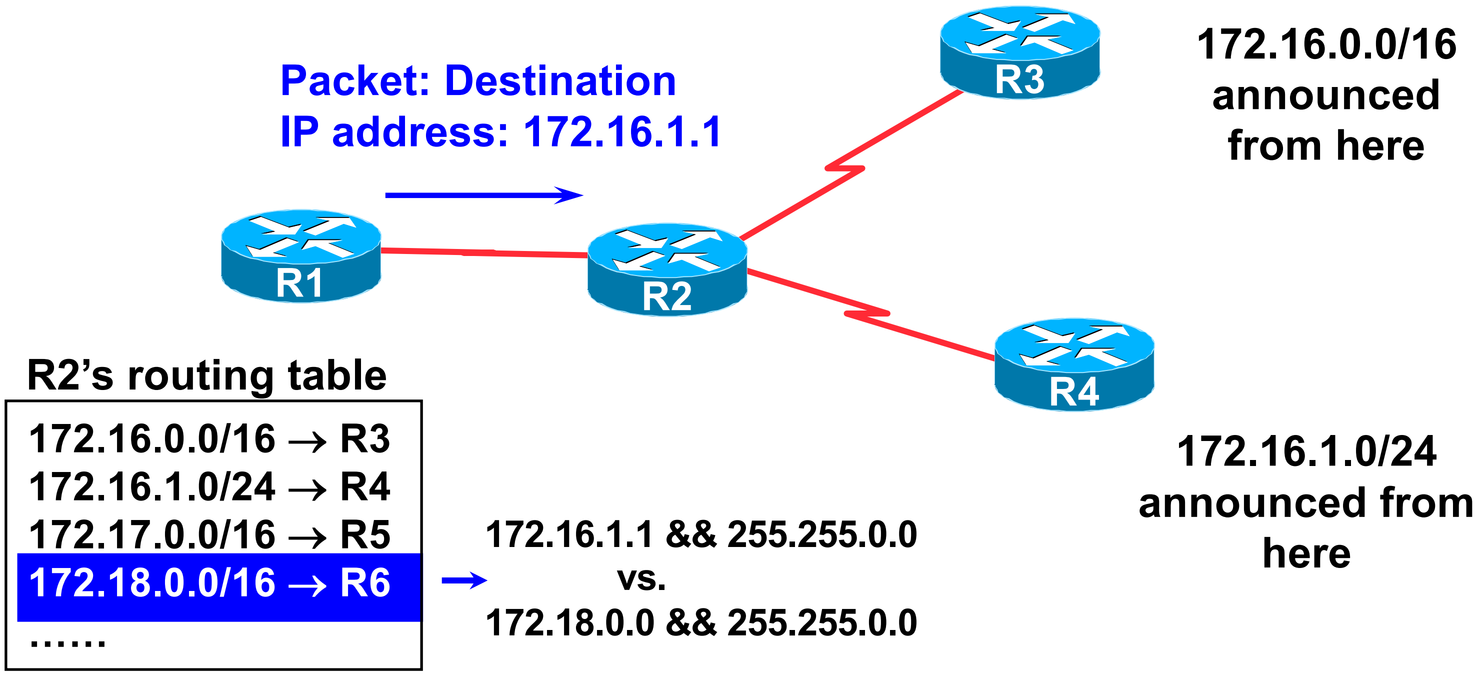
| |
|---|
| 172.16.0.0/16 → R3 |
| 172.16.1.0/24 → R4 |
| 172.17.0.0/16 → R5 |
| 172.18.0.0/16 → R6 |
| …… |

→ 172.16.1.1 && 255.255.0.0
vs.
172.17.0.0 && 255.255.0.0

**Does not match!**

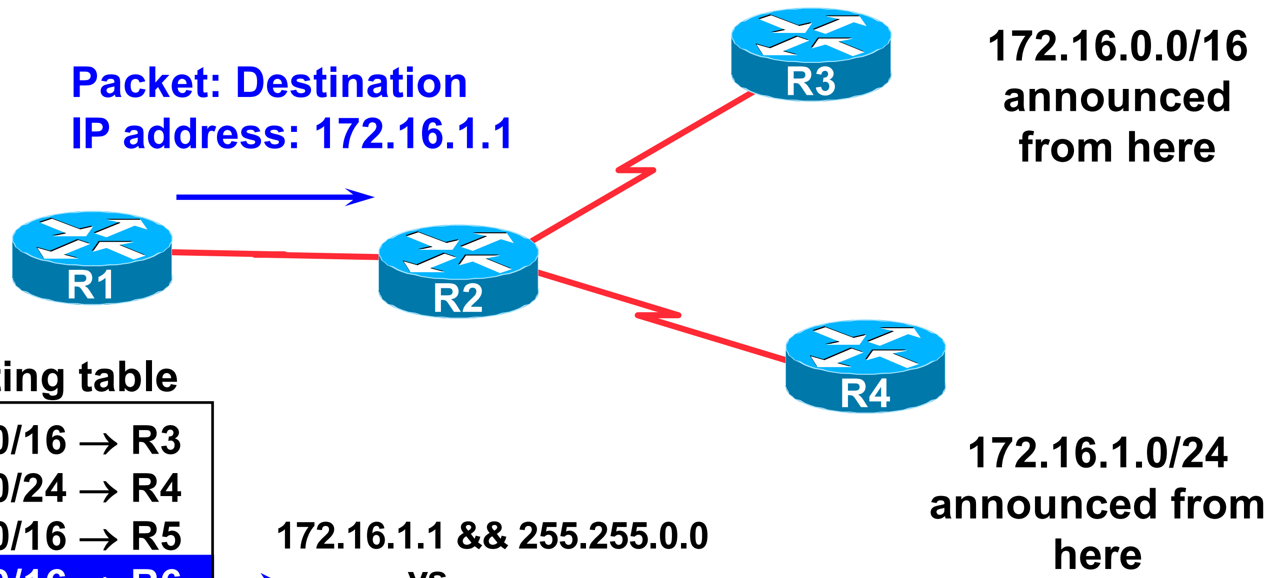
# IP route lookup: Longest match routing

- Based on destination IP address



# IP route lookup: Longest match routing

- Based on destination IP address




**R2's routing table**

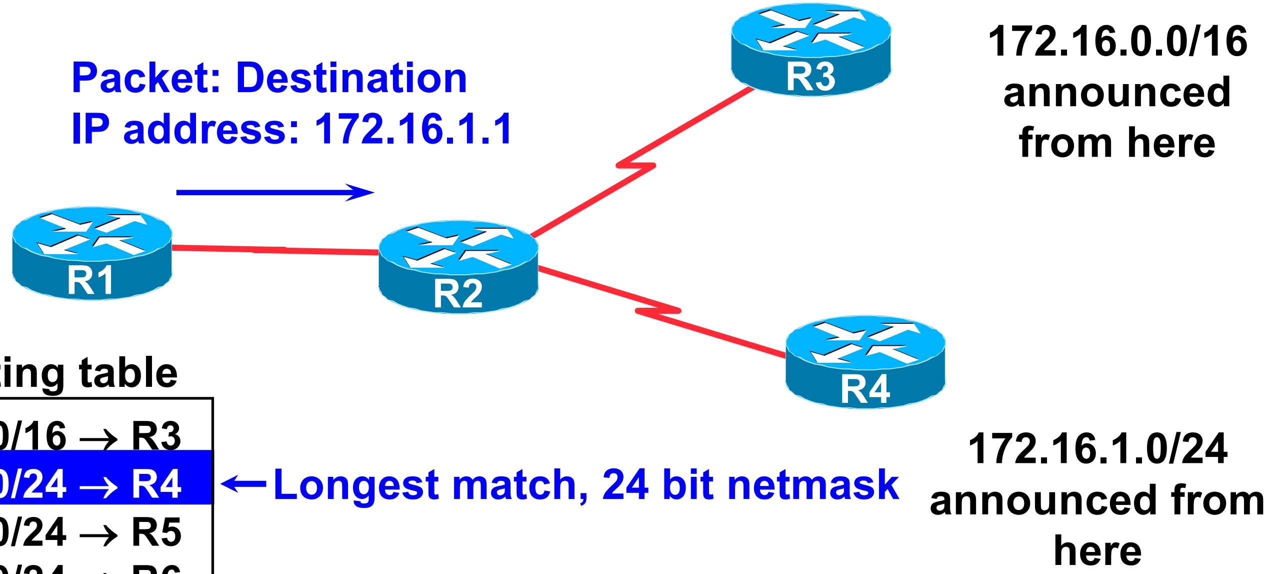
| |
|---|
| 172.16.0.0/16 → R3 |
| 172.16.1.0/24 → R4 |
| 172.16.0.0/16 → R5 |
| **172.18.0.0/16 → R6** |
| …… |

→ 172.16.1.1 && 255.255.0.0
vs.
172.18.0.0 && 255.255.0.0

**Does not match!**

# IP route lookup: Longest match routing

- Based on destination IP address

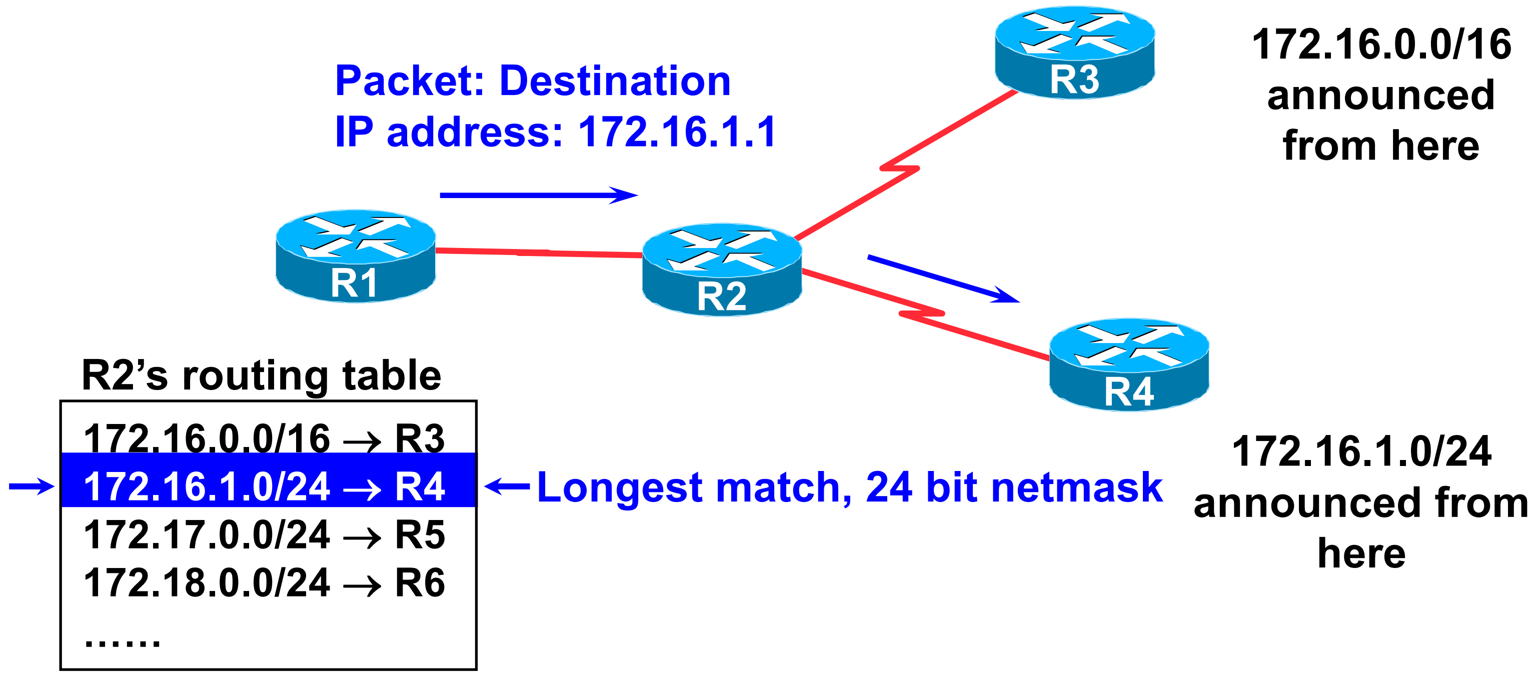



**R2's routing table**

| | |
|---|---|
| 172.16.0.0/16 | → R3 |
| **172.16.1.0/24** | **→ R4** |
| 172.17.0.0/24 | → R5 |
| 172.18.0.0/24 | → R6 |
| …… | |

# IP route lookup: Longest match routing

- Based on destination IP address



# Routing versus Forwarding

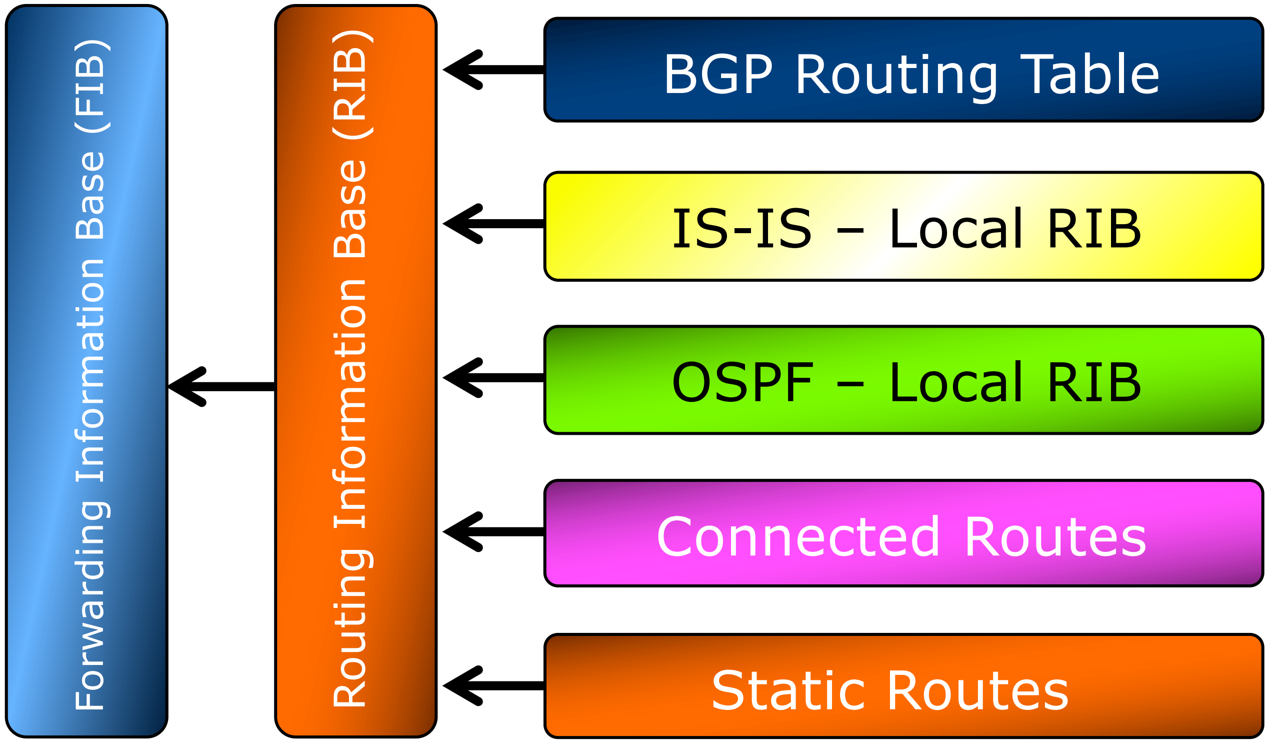
- Routing = building maps and giving directions

- Forwarding = moving packets between interfaces according to the "directions"

# IP Forwarding

- Router decides which interface a packet is sent to
- Forwarding table populated by routing process
- Forwarding decisions:
  - destination address
  - class of service (fair queuing, precedence, others)
  - local requirements (packet filtering)
- Forwarding is usually aided by special hardware

# 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
  - It is built from the router’s Global RIB
  - Used by the router to figure out where to send the 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
    - Routing protocol priority varies per vendor – see addendum
  - 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”

# Explicit versus Default Routing

- Default:
  - Simple, cheap (CPU, memory, bandwidth)
  - No overhead
  - Low granularity (metric games)
- Explicit: (default free zone)
  - Complex, expensive (CPU, memory, bandwidth)
  - High overhead
  - High granularity (every destination known)
- Hybrid:
  - Minimise overhead
  - Provide useful granularity
  - Requires some filtering knowledge

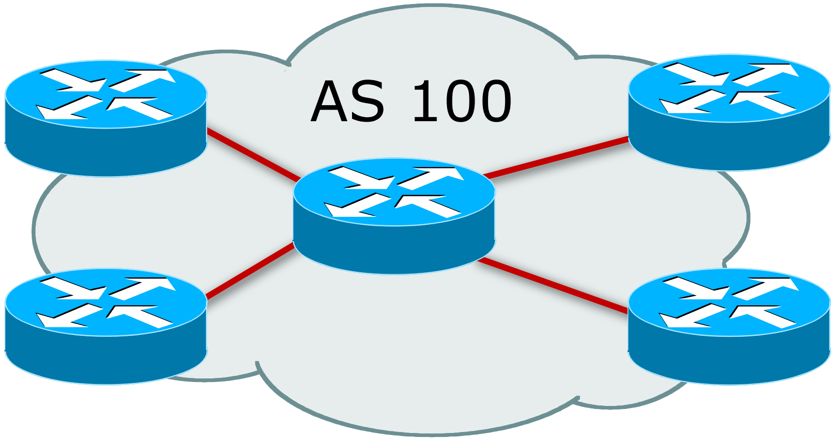
# Egress Traffic

- How packets leave your network
- Egress traffic depends on:
  - Route availability (what others send you)
  - Route acceptance (what you accept from others)
  - Policy and tuning (what you do with routes from others)
  - Peering and transit agreements

# Ingress Traffic

- How packets get to your network and your customers' networks
- Ingress traffic depends on:
  - What information you send and to whom
  - Based on your addressing and AS's
  - Based on others' policy (what they accept from you and what they do with it)

# Autonomous System (AS)

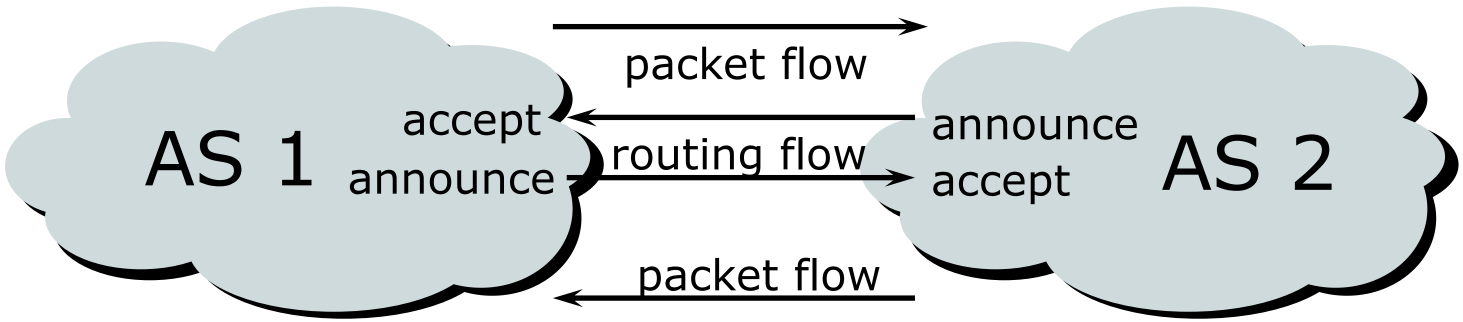



- Collection of networks with same routing policy
- Single routing protocol
- Usually under single ownership, trust and administrative control

# Definition of terms

- Neighbours
  - AS's which directly exchange routing information
  - Routers which exchange routing information
- Announce
  - send routing information to a neighbour
- Accept
  - receive and use routing information sent by a neighbour
- Originate
  - insert routing information into external announcements (usually as a result of the IGP)
- Peers
  - routers in neighbouring AS's or within one AS which exchange routing and policy information

# Routing flow and packet flow




For networks in AS1 and AS2 to communicate:

AS1 must announce to AS2

AS2 must accept from AS1

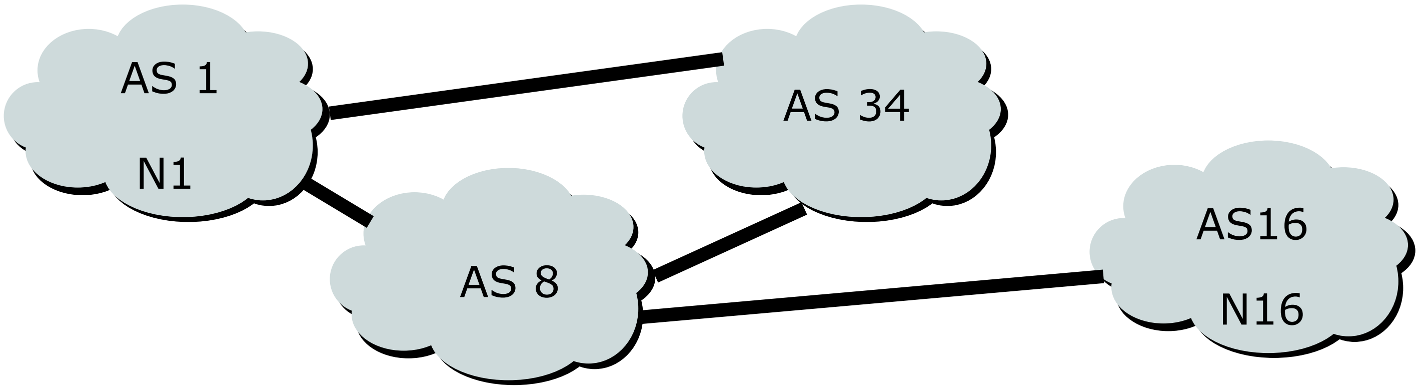
AS2 must announce to AS1

AS1 must accept from AS2

# Routing flow and Traffic flow

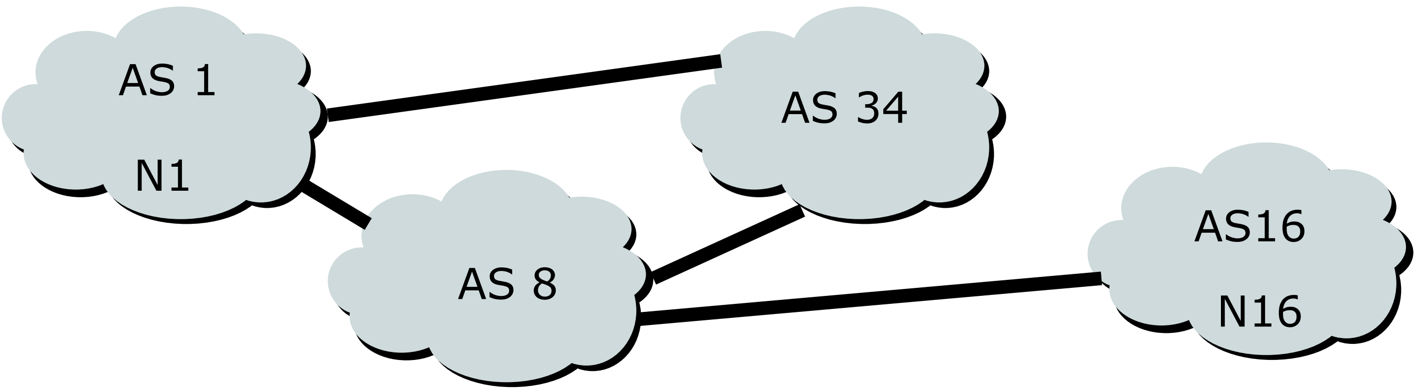
- Traffic flow is always in the opposite direction of the flow of Routing information
  - Filtering outgoing routing information inhibits traffic flow inbound
  - Filtering inbound routing information inhibits traffic flow outbound

# Routing Flow/Packet Flow: With multiple ASes




- For net N1 in AS1 to send traffic to net N16 in AS16:
  - AS16 must originate and announce N16 to AS8.
  - AS8 must accept N16 from AS16.
  - AS8 must announce N16 to AS1 or AS34.
  - AS1 must accept N16 from AS8 or AS34.
- For two-way packet flow, similar policies must exist for N1

# Routing Flow/Packet Flow: With multiple ASes

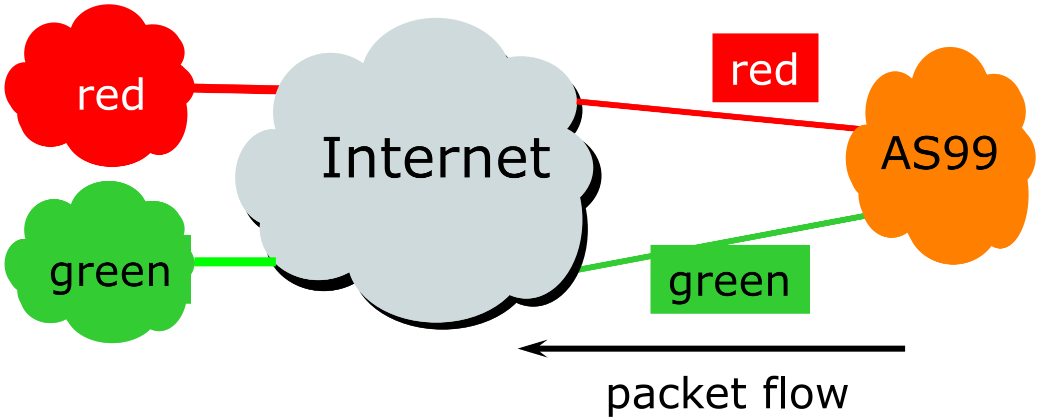



- As more and more paths are implemented between sites it is easy to see how policies can become quite complex.

# Routing Policy

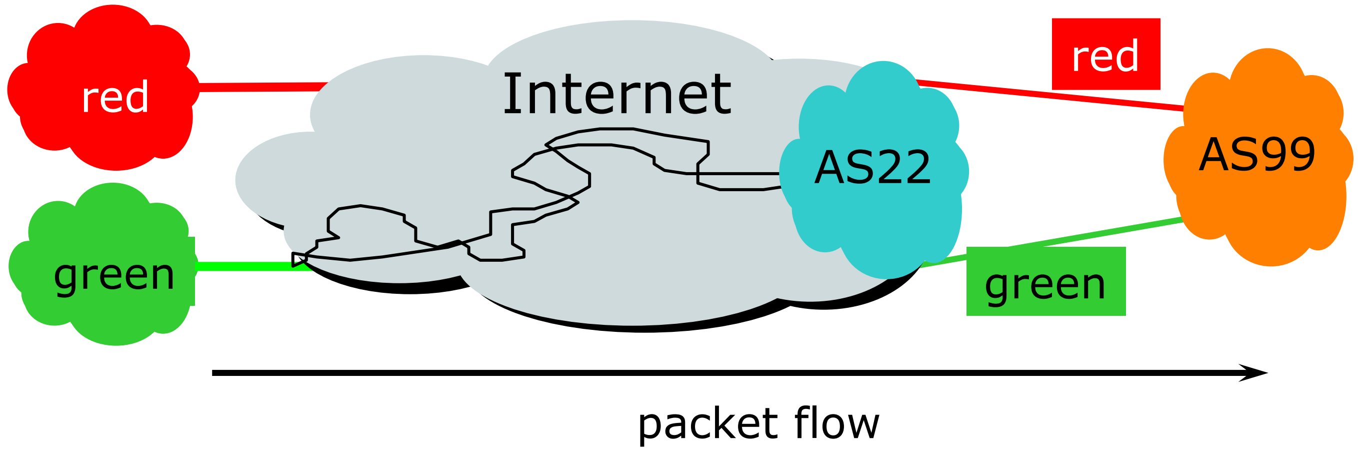
- Used to control traffic flow in and out of a network
- The network operator makes decisions on what routing information to accept and discard from its neighbours
  - Individual routes
  - Routes originated by specific ASes
  - Routes traversing specific ASes
  - Routes belonging to other groupings
    - Groupings which you define as you see fit

# Routing Policy Limitations




- AS99 uses red link for traffic to the red AS and the green link for remaining traffic
- To implement this policy, AS99 has to:
  - Accept routes originating from the red AS on the red link
  - Accept all other routes on the green link

# Routing Policy Limitations




- AS99 would like packets coming from the green AS to use the green link.
- But unless AS22 cooperates in pushing traffic from the green AS down the green link, there is very little that AS99 can do to achieve this aim

# Routing Policy Issues

- July 2019:
  - 67000 IPv6 prefixes & 763000 IPv4 prefixes
    - Not realistic to set policy on all of them individually
  - 65000 origin AS’s
    - Too many to try and create individual policies for
- Routes tied to a specific AS or path may be unstable regardless of connectivity
- Solution: Groups of AS’s are a natural abstraction for filtering purposes

# Routing Protocols

We now know what routing means…

…but what do the routers get up to?

And why are we doing this anyway?

# 1: How Does Routing Work?

- Internet is made up of the Network Operators who connect to each other's networks
- How does an operator in Kenya tell an operator in Japan what end-site customers they have?
- And how does that operator send data packets to the customers of the Japanese operator, and get responses back
  - After all, as on a local ethernet, two way packet flow is needed for communication between two devices

# 2: How Does Routing Work?

- The operator in Kenya could buy a direct connection to the operator in Japan
  - But this doesn't scale – there are thousands of distinct networks, would need thousands of connections, and cost would be astronomical
- Instead, the operator in Kenya tells his neighbouring operators what end-sites they have
  - And the neighbouring operators pass this information on to their neighbours, and so on
  - This process repeats until the information reaches the operator in Japan

# 3: How Does Routing Work?

- This process is called “Routing”
- The mechanisms used are called “Routing Protocols”
- Routing and Routing Protocols ensures that
  - The Internet can scale
  - Thousands of network operators can provide connectivity to each other
  - We have the Internet we see today

# 4: How Does Routing Work?

- The Network Operator in Kenya doesn’t actually tell its neighbouring operators the names of the end-sites
  - (network equipment does not understand names)
- Instead, it has received an IP address block as a member of the Regional Internet Registry serving Kenya
  - Its customers have received address space from this address block as part of their “Internet service”
  - And it announces this address block to its neighbouring operators – this is called announcing a “route”

# Routing Protocols

- Routers use "routing protocols" to exchange routing information with each other
  - IGP is used to refer to the process running on routers inside an ISP's network
  - EGP is used to refer to the process running between routers bordering directly connected ISP networks

# What Is an IGP?

- Interior Gateway Protocol
- Within an Autonomous System
- Carries information about internal infrastructure prefixes
- Two widely used IGPs:
  - OSPF
  - IS-IS

# Why Do We Need an IGP?

- Network Operator backbone scaling
  - Hierarchy
  - Limiting scope of failure
  - Only used for operator's infrastructure addresses, not customers or anything else
  - Design goal is to minimise number of prefixes in IGP to aid scalability and rapid convergence

# What Is an EGP?

- Exterior Gateway Protocol
- Used to convey routing information between Autonomous Systems
- De-coupled from the IGP
- Current EGP is BGP

# Why Do We Need an EGP?

- Scaling to large network
  - Hierarchy
  - Limit scope of failure
- Define Administrative Boundary
- Policy
  - Control reachability of prefixes
  - Merge separate organisations
  - Connect multiple IGPs

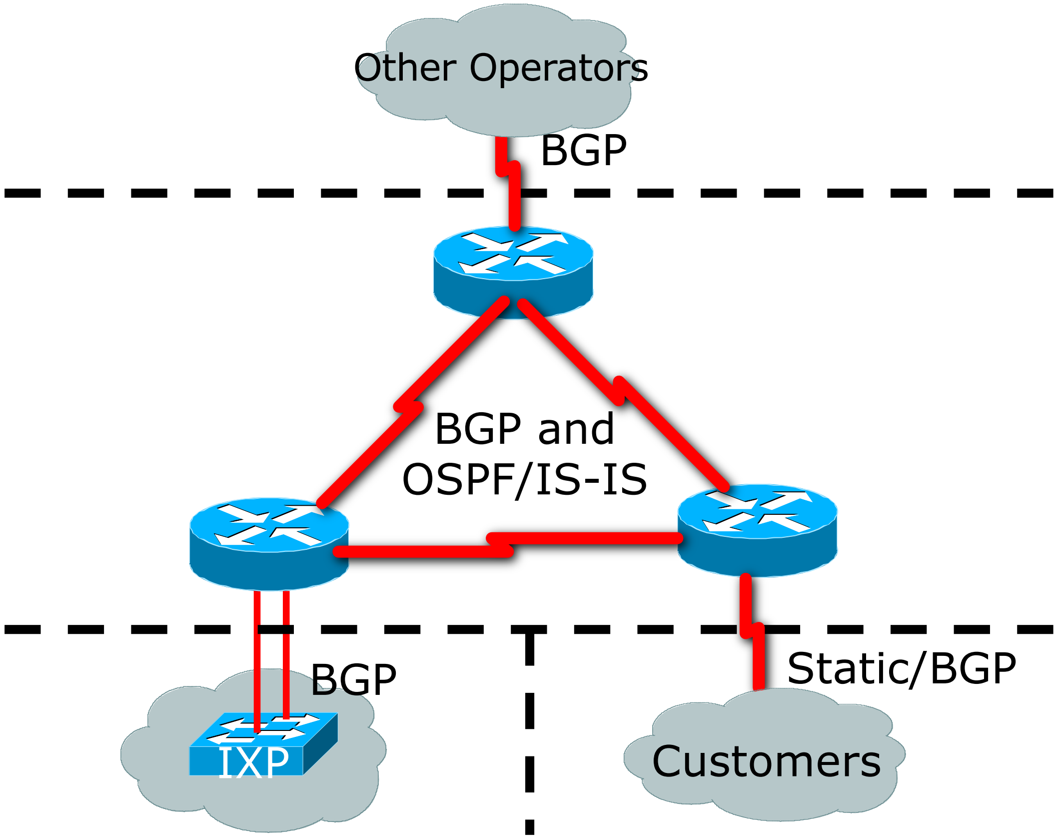
# Interior versus Exterior Routing Protocols

- Interior
  - Automatic neighbour discovery
  - Generally trust your IGP routers
  - Prefixes go to all IGP routers
  - Binds routers in one AS together
- Exterior
  - Specifically configured peers
  - Connecting with outside networks
  - Set administrative boundaries
  - Binds AS's together

# Interior versus Exterior Routing Protocols

- Interior
  - Carries network infrastructure addresses only
  - Network operators aim to keep the IGP small for efficiency and scalability
- Exterior
  - Carries customer prefixes
  - Carries Internet prefixes
  - EGPs are independent of the operator's network topology

# Hierarchy of Routing Protocols



# FYI: Default Administrative Distances

| Route Source | Cisco | Juniper | Huawei | Brocade | Nokia/ALU | 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** | N/A |
| 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 | ? | |

# Routing Basics

ISP Workshops