Campus Network Design Workshop

Routing Basics

UNIVERSITY OF OREGON

Last updated 17th October 2016

NSRC
Network Startup Resource Center

Routing Concepts

- IPv4 & IPv6
- Routing
- Forwarding
- Some definitions
- Policy options
- Routing Protocols

IPv4

- Internet still uses IPv4
  - (considered a legacy protocol)
  - Addresses are 32 bits long
  - Addresses are written as decimal with each 8-bit range separated by a "."
  - Range from 1.0.0.0 to 223.255.255.255
  - 0.0.0.0 to 0.255.255.255 and 224.0.0.0 to 255.255.255.255 have special uses
- IPv4 address has a network portion and a host portion

IPv6

- Internet is starting to use IPv6
  - Addresses are 128 bits long
  - Addresses are written as hexadecimal with each 16-bit range separated by ":"
  - Internet addresses range from 2000::/16 to 3FFF::/16
  - The remaining IPv6 range is reserved or has "special" uses
- IPv6 address has a network portion and a host portion

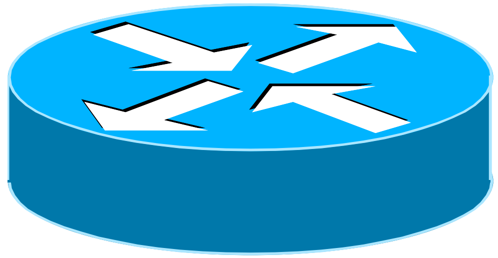
IP address format

- Address and subnet mask
  - IPv4 written as
    - 12.34.56.78 255.255.255.0 or
    - 12.34.56.78/24
  - IPv6 written as
    - 2001:db8::1/126
  - mask represents the number of network bits in the address
    - Usually referred to as the subnet size
  - The remaining bits are the host bits

IP subnets

- IPv4 example – 12.34.56.78/24
  - 32 bits in an IPv4 address
    - 24 bits for the network portion
    - Leaves 8 bits for the host portion
    - 8 bits means there are 2^8 possible hosts on this subnet
- IPv6 example – 2001:db8::1/126
  - 128 bits in an IPv6 address
    - 126 bits for the network portion
    - Leaves 2 bits for the host portion
    - 2 bits means there are 2^2 possible hosts on this subnet

What does a router do?



A day in a life of a router

find path

forward packet, forward packet, forward packet, forward packet...

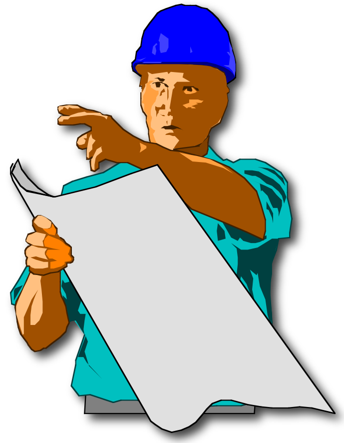
find alternate path

forward packet, forward packet, forward packet, forward packet…

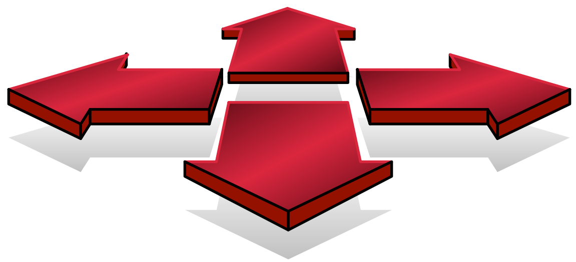
repeat until powered off

Routing versus Forwarding

- Routing = building maps and giving directions



- Forwarding = moving packets between interfaces according to the “directions”



IP Routing – finding the path

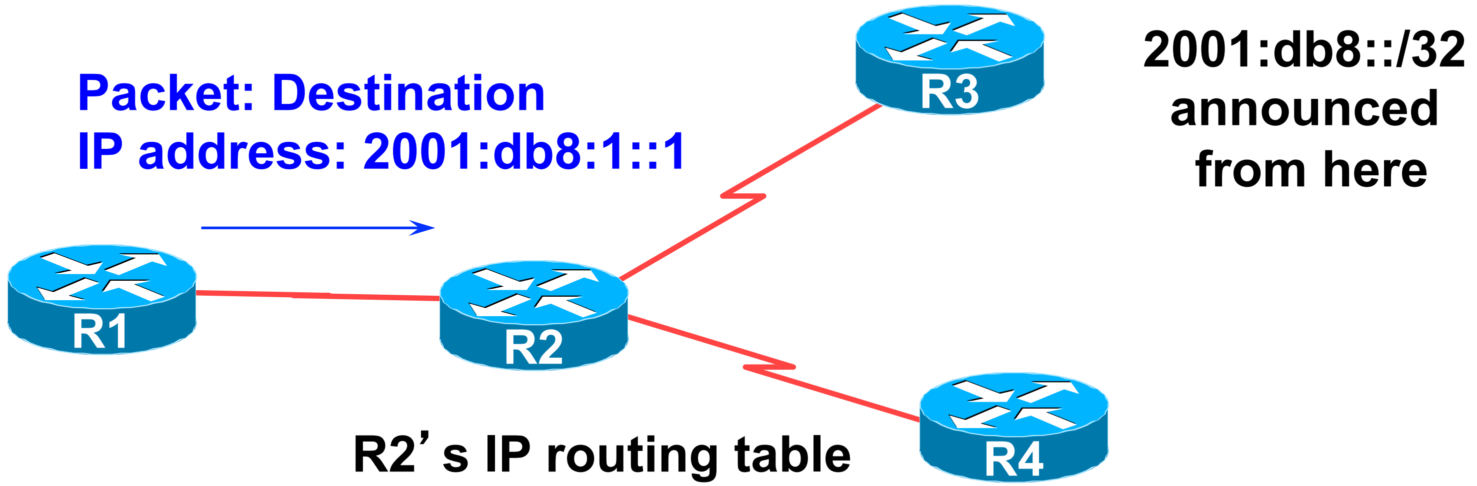
- Path derived from information received from a routing protocol
- Several alternative paths may exist
  - Best path stored in **forwarding** table
- Decisions are updated periodically or as topology changes (event driven)
- Decisions are based on:
  - Topology, policies and metrics (hop count, filtering, 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 2001:db8:1::1/128 is sent to the router announcing 2001:db8:1::/48 rather than the router announcing 2001:db8::/32.

IP route lookup

- Based on destination IP address



R2's IP routing table

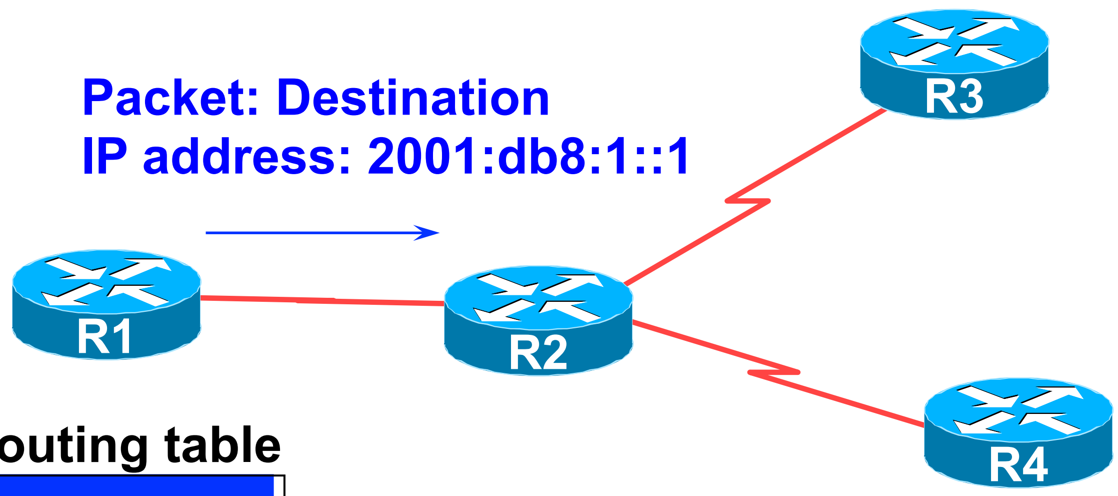
2001:db8::/32 → R3
2001:db8:1::/48→ R4
2001:db9::/32 → R5
2001:dba::/32 → R6
……

2001:db8:1::/48 announced from here

IP route lookup: Longest match routing

- Based on destination IP address

Packet: Destination
IP address: 2001:db8:1::1



2001:db8::/32 announced from here

2001:db8:1::/48 announced from here

R2's IP routing table

2001:db8::/32 → R3
2001:db8:1::/48 → R4
2001:db9::/32 → R5
2001:dba::/32 → R6
……

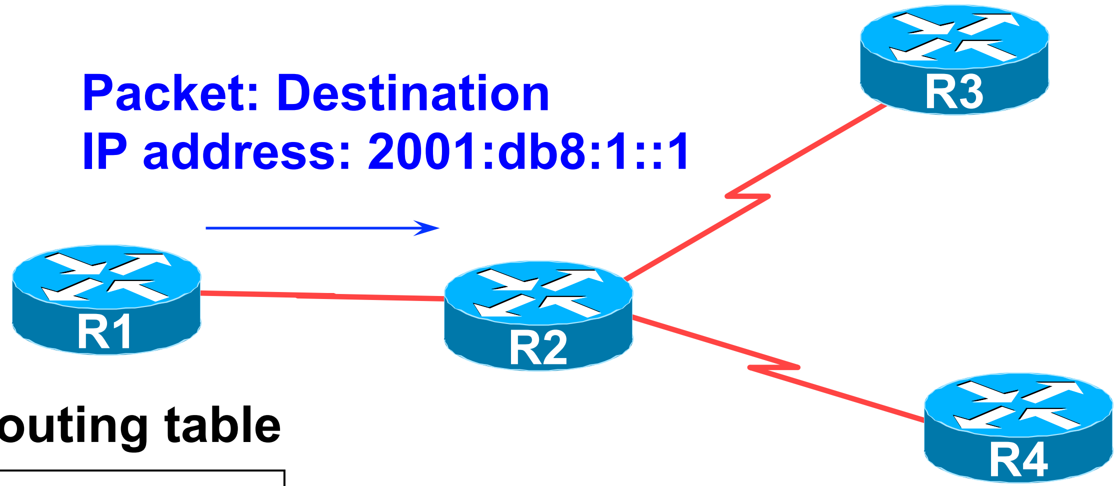
2001:db8:1::1 && ffff:ffff::
vs.
2001:db8:: && ffff:ffff::

Match!

IP route lookup:
Longest match routing

- Based on destination IP address

Packet: Destination
IP address: 2001:db8:1::1



2001:db8::/32 announced from here

2001:db8:1::/16 announced from here

R2's IP routing table

2001:db8::/32 → R3
2001:db8:1::/48 → R4
2001:db9::/32 → R5
2001:dba::/32 → R6
……

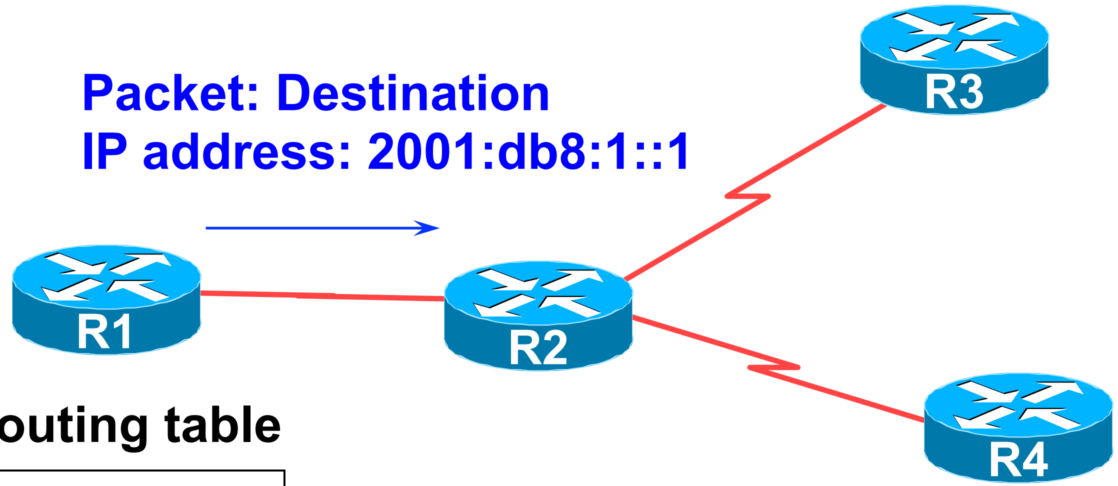
2001:db8:1::1 && ffff:ffff:ffff::
vs.
2001:db8:1:: && ffff:ffff:ffff::

Match as well!

IP route lookup: Longest match routing

- Based on destination IP address

Packet: Destination IP address: 2001:db8:1::1



2001:db8::/32 announced from here

2001:db8:1::/48 announced from here

R2's IP routing table

2001:db8::/32 → R3
2001:db8:1::/48 → R4
2001:db9::/32 → R5
2001:dba::/32 → R6
……

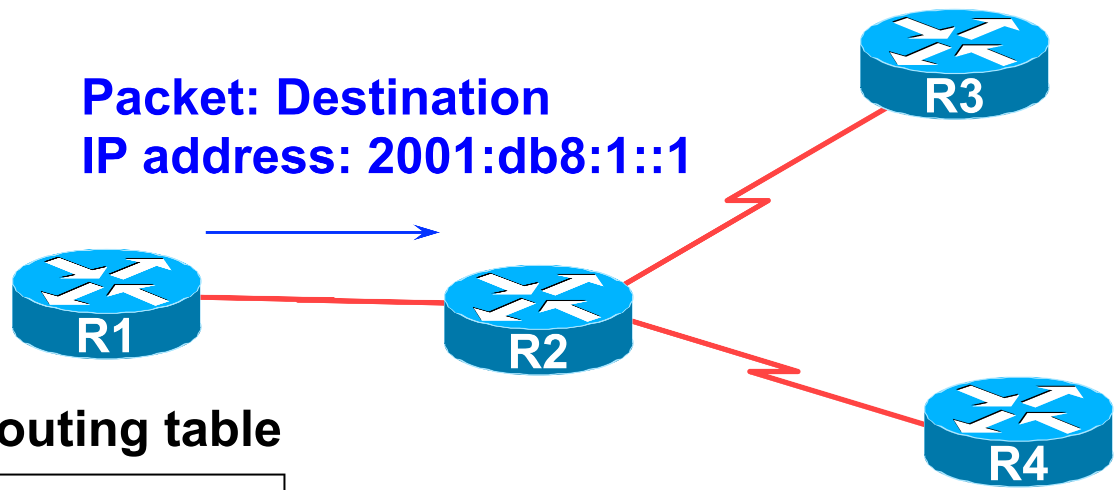
2001:db8:1::1 && ffff:ffff::
vs.
2001:db9:: && ffff:ffff::

Does not match!

IP route lookup: Longest match routing

- Based on destination IP address

Packet: Destination
IP address: 2001:db8:1::1



2001:db8::/32 announced from here

2001:db8:1::/48 announced from here

R2's IP routing table

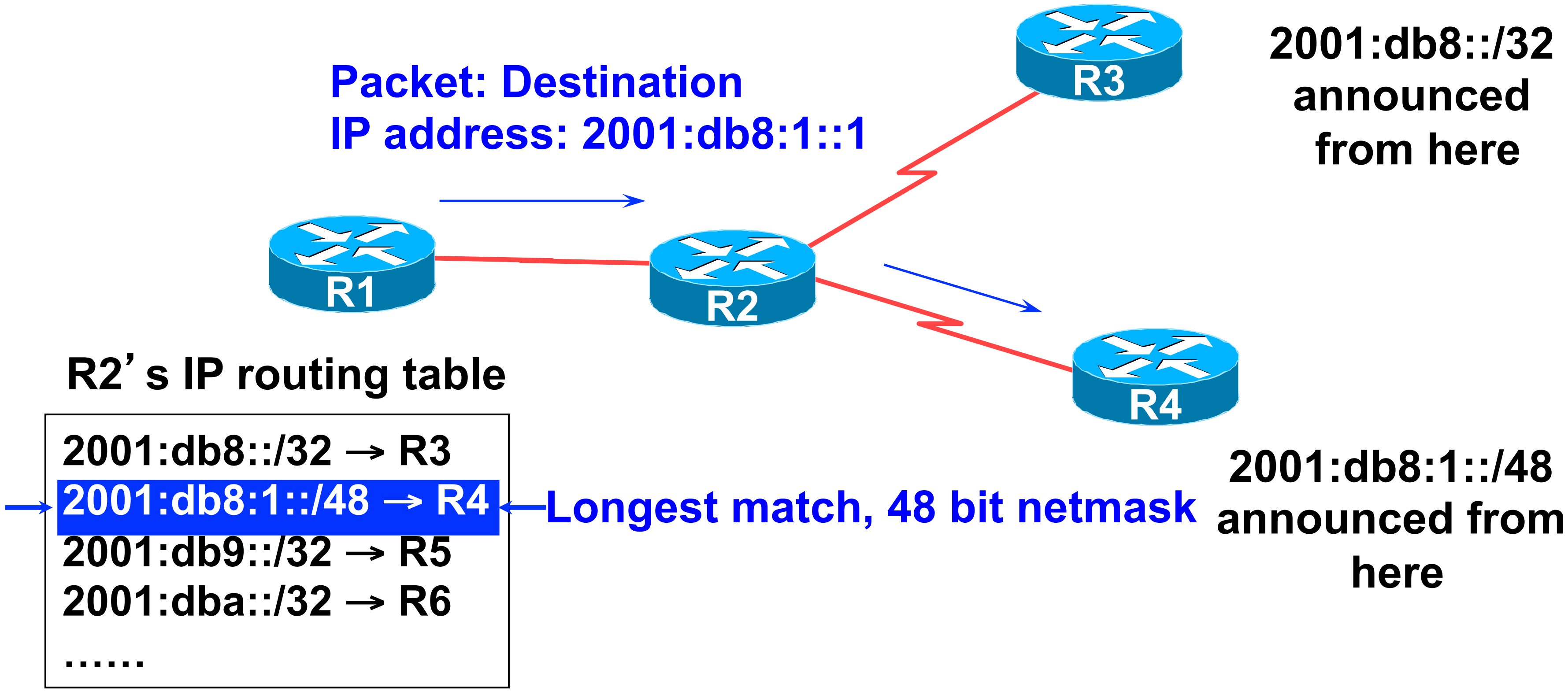
2001:db8::/32 → R3
2001:db8:1::/48 → R4
2001:db9::/32 → R5
2001:dba::/32 → R6
……

2001:db8:1::1 && ffff:ffff::
vs.
2001:dba:: && ffff:ffff::

Does not match!

IP route lookup: Longest match routing

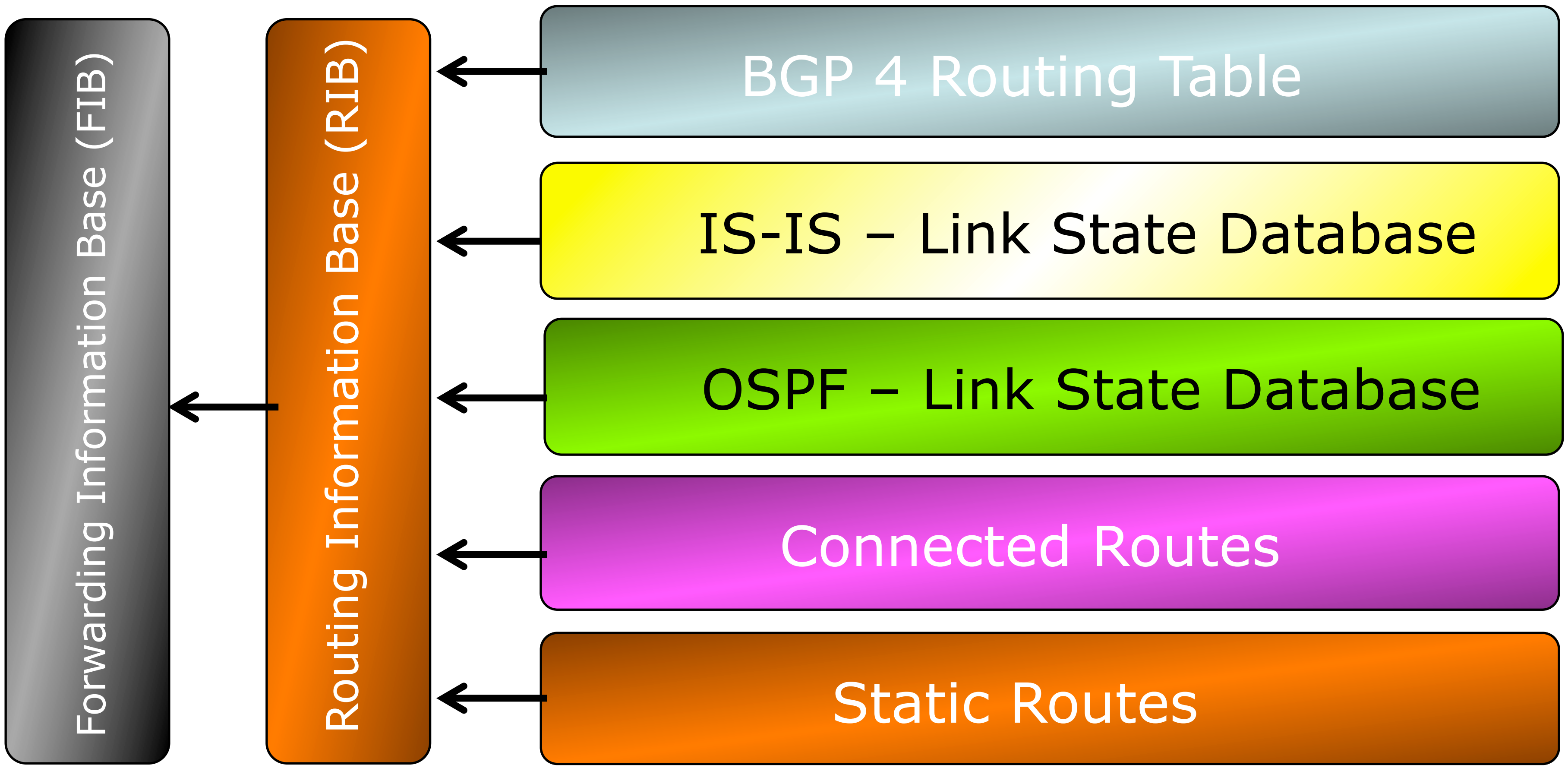
- Based on destination IP address



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 and the interfaces to get to those destinations
  - Used by the router to figure out where to send the packet
  - Careful! Some people still call this a route!
  - Cisco IOS: “show ip cef”

The RIB

- RIB is the Routing Table
  - It contains a list of all the destinations and the various next hops used to get to those destinations – and lots of other information too!
  - One destination can have lots of possible next-hops – only the best next-hop goes into 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

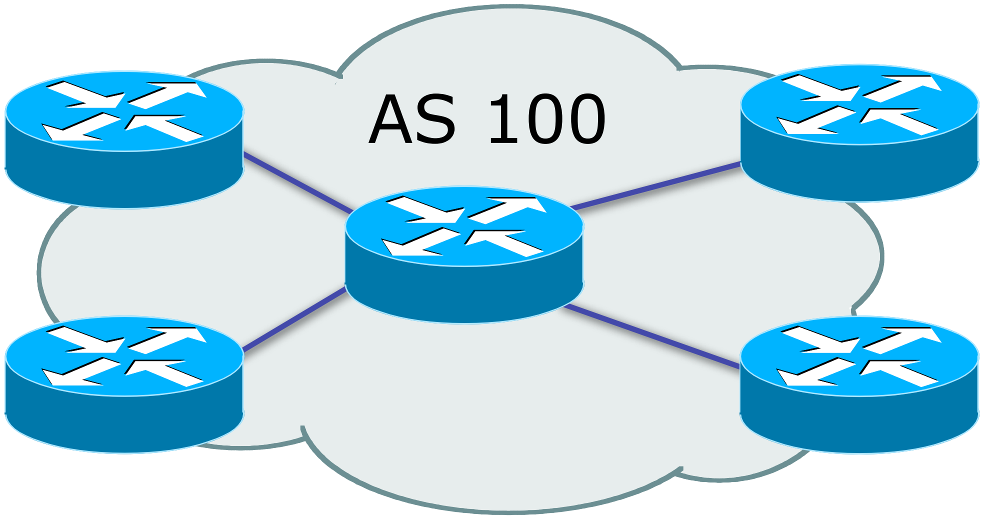
Egress Traffic

- 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)

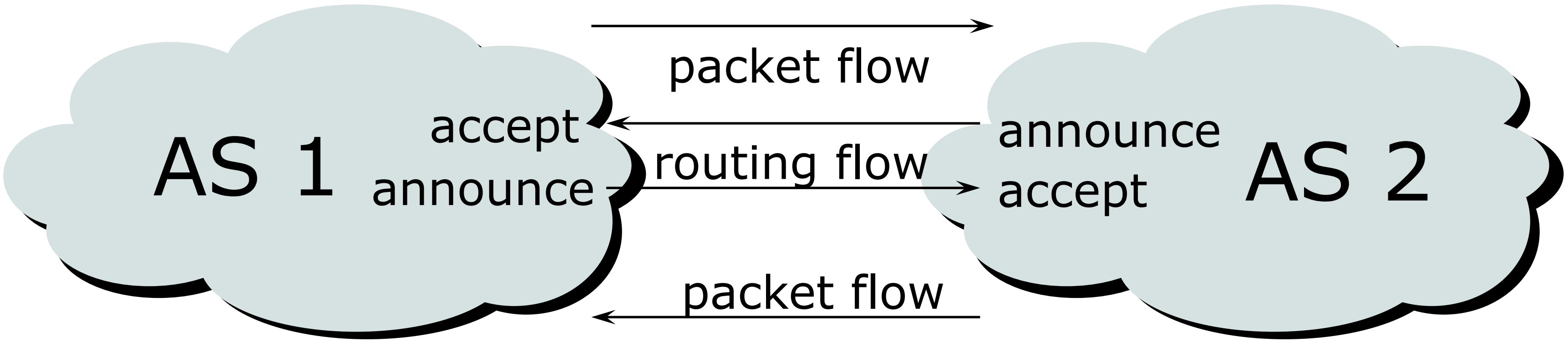


- Collection of networks with same routing policy
- Single routing protocol
- Usually under single ownership, trust and administrative control

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

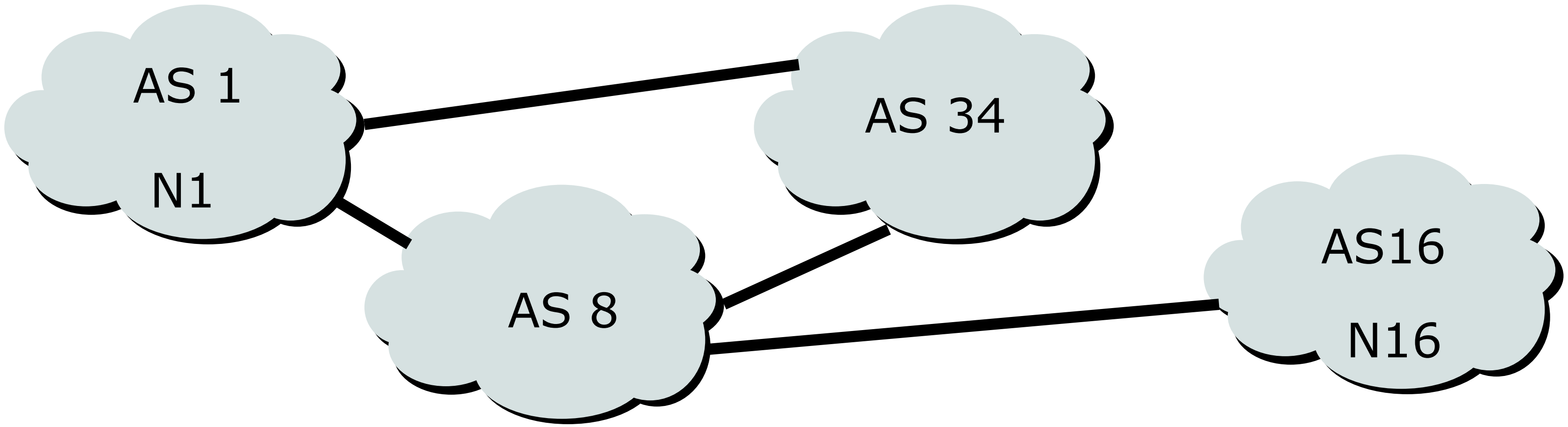
UNIVERSITY OF OREGON

NSRC
Network Startup Resource Center

Routing flow and Traffic flow

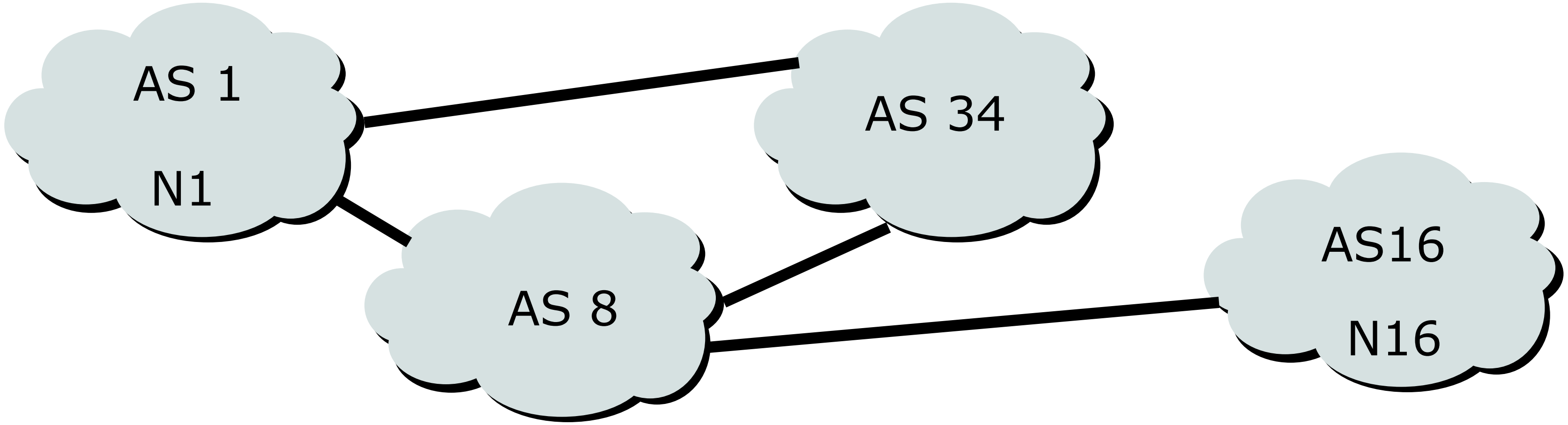
- Traffic flow is always in the opposite direction of the flow of Routing information
  - Filtering outgoing routing information inhibits traffic flow inbound
  - Filtering inbound routing information inhibits traffic flow outbound

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

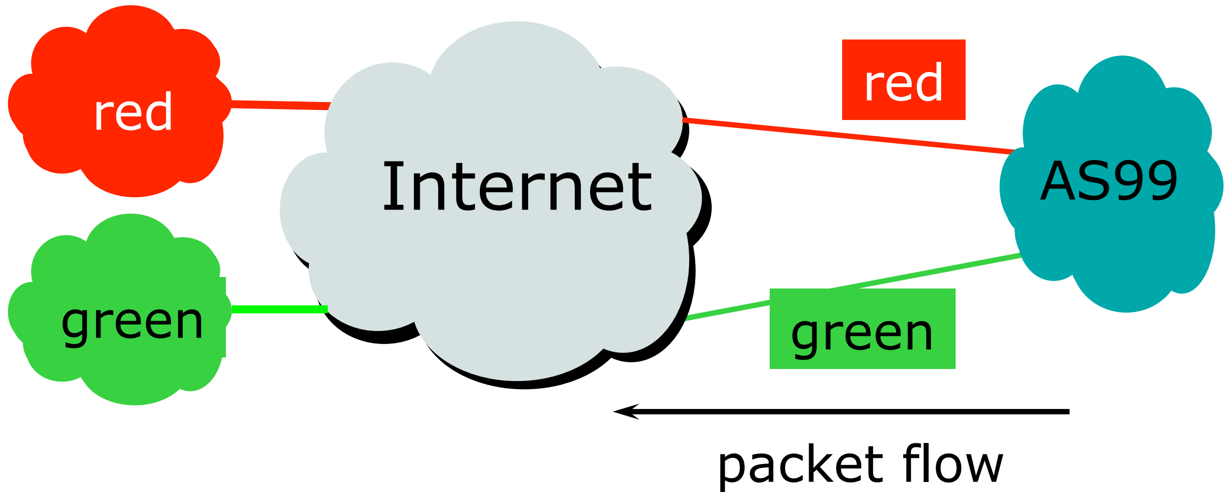


- As multiple paths between sites are implemented it is easy to see how policies can become quite complex.

Routing Policy

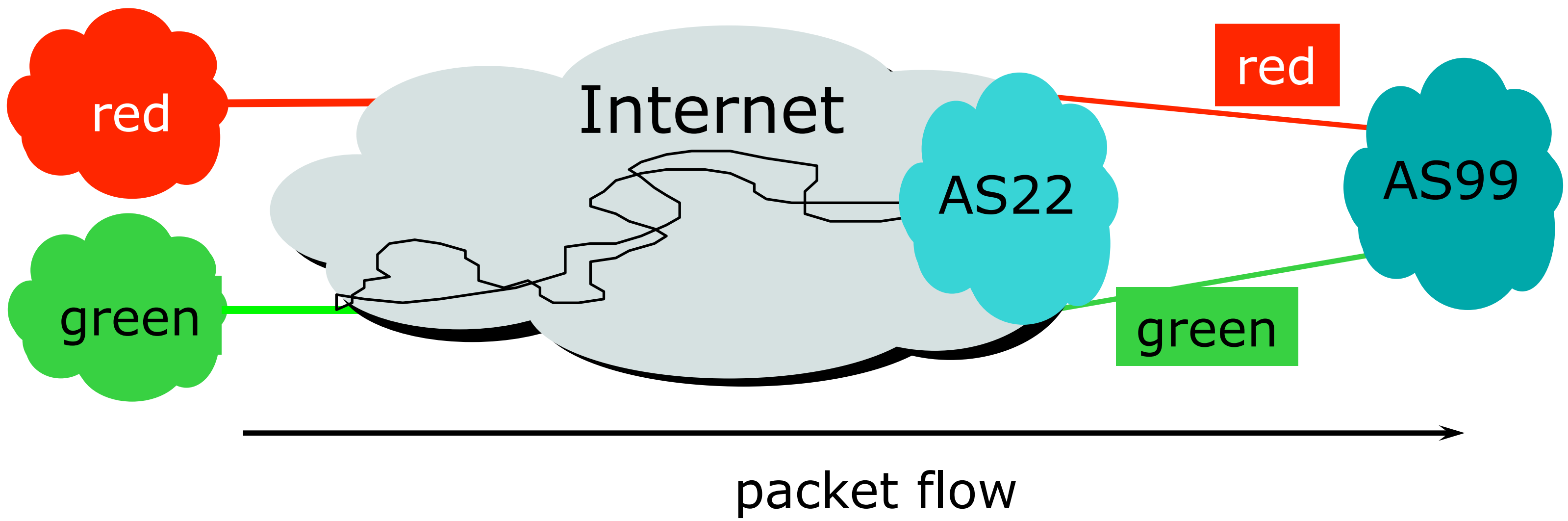
- Used to control traffic flow in and out of an ISP network
- ISP makes decisions on what routing information to accept and discard from its neighbours
  - Individual routes
  - Routes originated by specific ASes
  - Routes traversing specific ASes
  - Routes belonging to other groupings
    - Groupings 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

- September 2016:
  - 31300 IPv6 prefixes & 608000 IPv4 prefixes
    - Not realistic to set policy on all of them individually
  - 57400 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 ISPs who connect to each other's networks
- How does an ISP in Kenya tell an ISP in Japan what customers they have?
- And how does that ISP send data packets to the customers of the ISP in Japan, 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?

- ISP in Kenya could buy a direct connection to the ISP in Japan
  - But this doesn't scale – thousands of ISPs, would need thousands of connections, and cost would be astronomical
- Instead, ISP in Kenya tells his neighbouring ISPs what customers he has
  - And the neighbouring ISPs pass this information on to their neighbours, and so on
  - This process repeats until the information reaches the ISP 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 ISPs can provide connectivity to each other
  - We have the Internet we see today

4: How Does Routing Work?

- ISP in Kenya doesn't actually tell his neighbouring ISPs the names of the customers
  - (network equipment does not understand names)
- Instead, he has received an IP address block as a member of the Regional Internet Registry serving Kenya
  - His customers have received address space from this address block as part of their "Internet service"
  - And he announces this address block to his neighbouring ISPs – 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

UNIVERSITY OF OREGON

NSRC
Network Startup Resource Center

What Is an IGP?

- **I**nterior **G**ateway **P**rotocol
- Within an Autonomous System
- Carries information about internal infrastructure prefixes
- Two widely used IGPs:
  - OSPF
  - IS-IS

Why Do We Need an IGP?

- ISP backbone scaling
  - Hierarchy
  - Limiting scope of failure
  - Only used for ISP'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?

- **E**xterior **G**ateway **P**rotocol
- 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

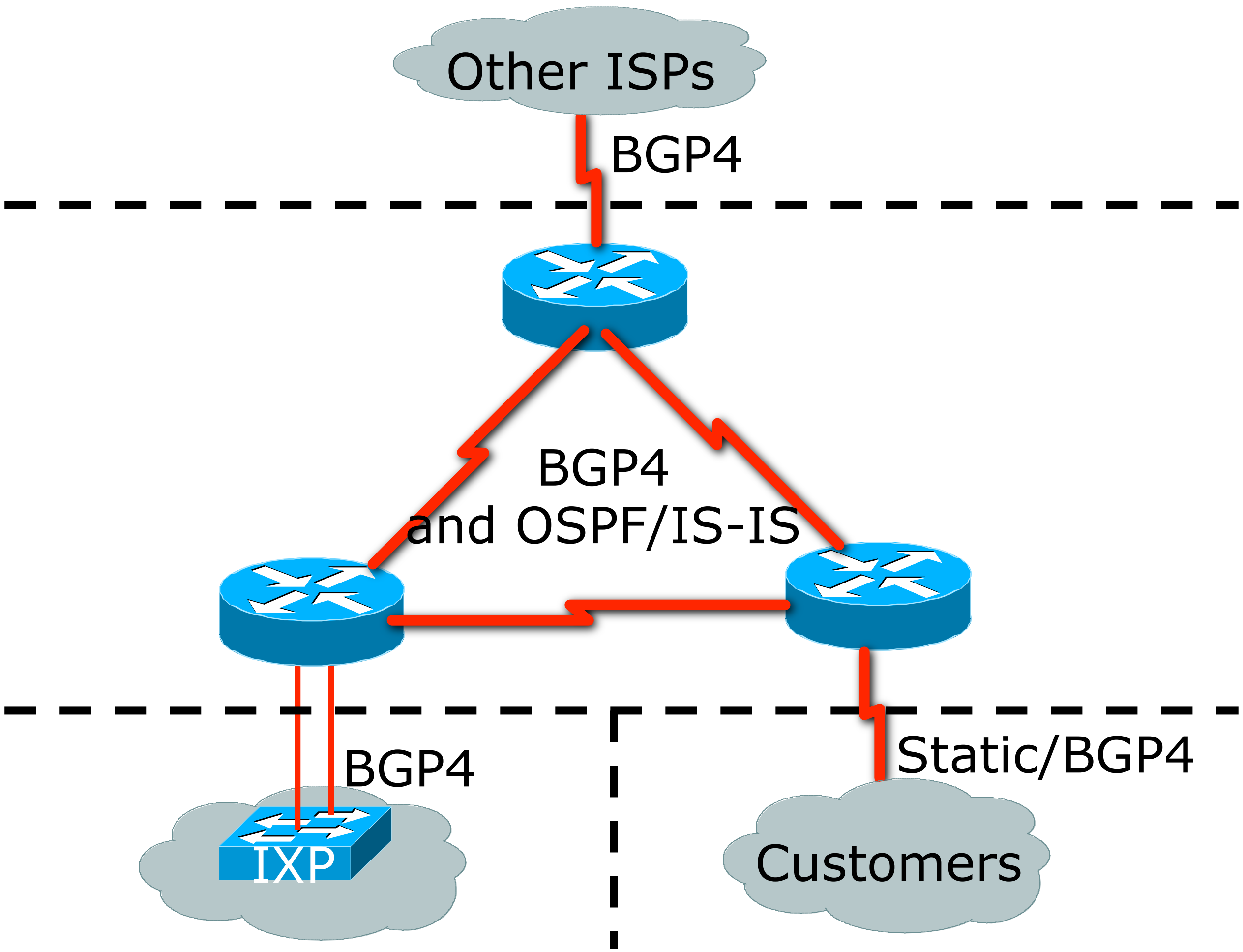
Interior versus Exterior Routing Protocols

- 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 ISP infrastructure addresses only
  - ISPs aim to keep the IGP small for efficiency and scalability
- Exterior
  - Carries customer prefixes
  - Carries Internet prefixes
  - EGPs are independent of ISP network topology

Hierarchy of Routing Protocols



FYI: Default Administrative Distances

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

Questions?

UNIVERSITY OF OREGON

NSRC
Network Startup Resource Center