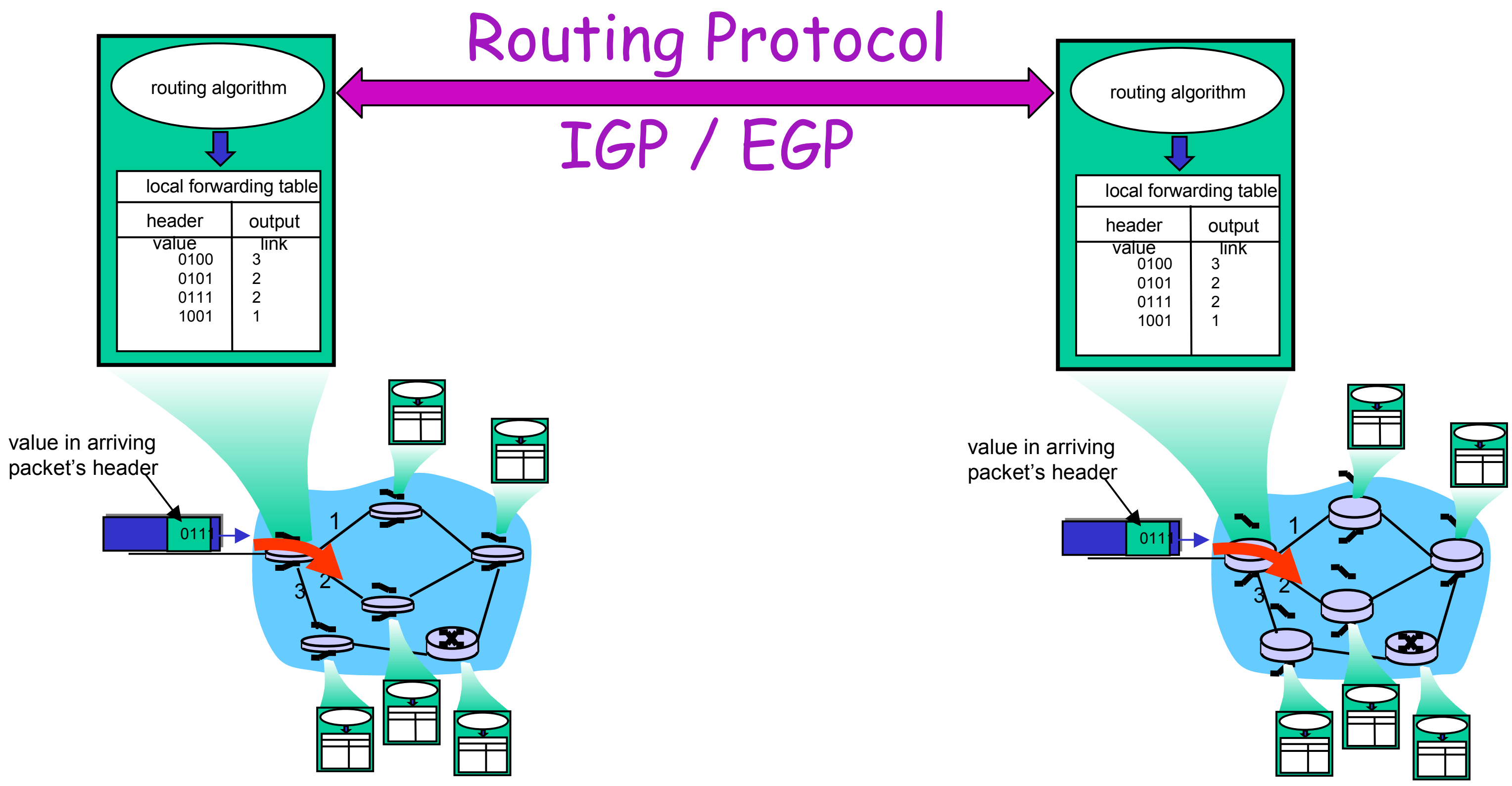
Dynamic Routing

Overview

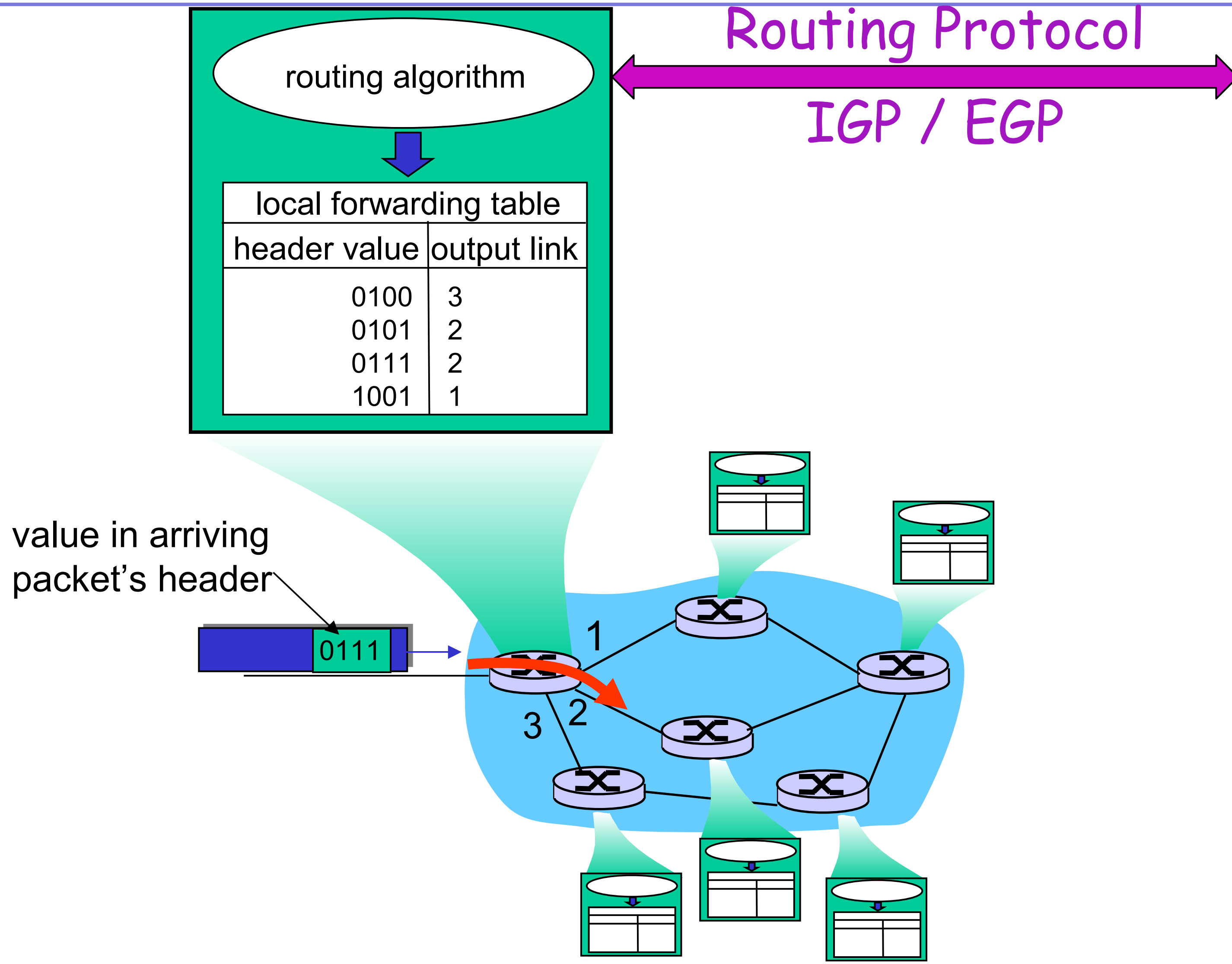
Desirable Characteristics of Dynamic Routing

- Automatically detect and adapt to topology changes
- Provide optimal routing
- Scalability
- Robustness
- Simplicity
- Rapid convergence
- Some control of routing choices
  - e.g., which links we prefer to use

Routers Talk Routing Protocols



Interplay between routing & forwarding



IP Routing – finding the path

- Path is derived from information received from the routing protocol
- Several alternative paths may exist
  - best next hop 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 Forwarding

- Router makes decision on 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)

Convergence – why do I care?

- Convergence is when all the routers have a stable view of the network
- When a network is not converged there is network downtime
  - Packets don’t get to where they are supposed to go
    - Black holes (packets “disappear”)
    - Routing Loops (packets go back and forth between the same devices)
  - Occurs when there is a change in state of router or the links

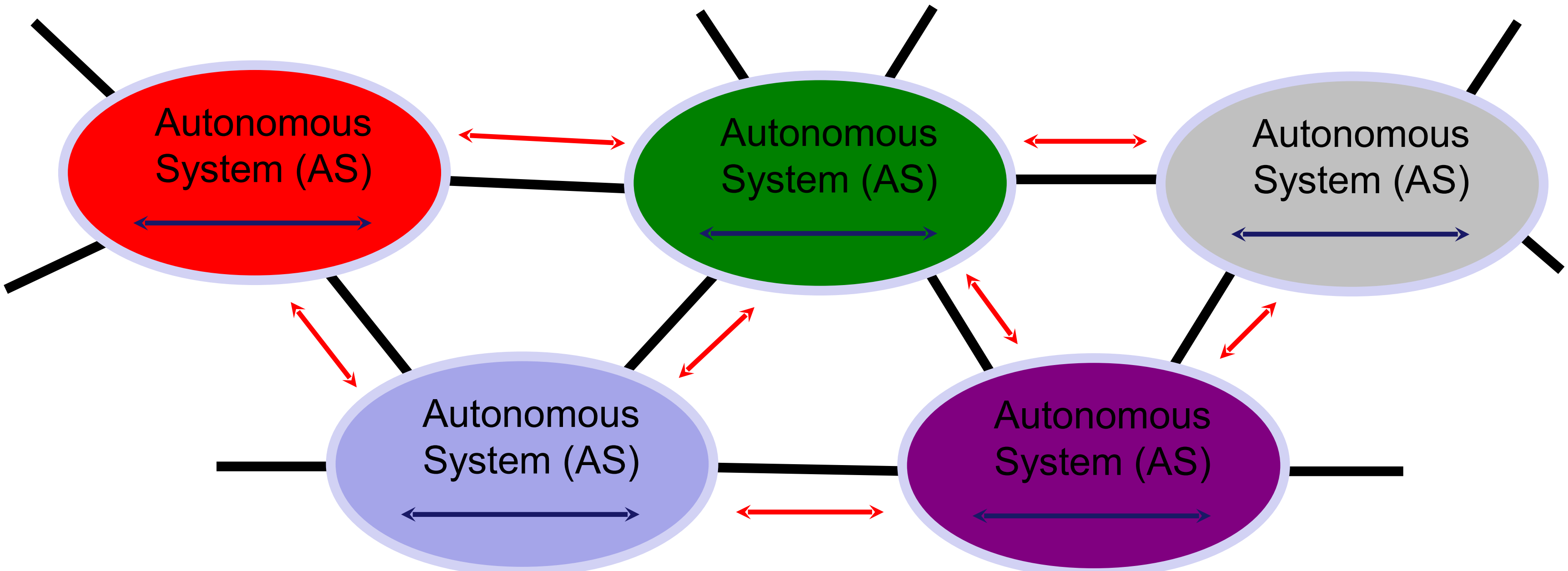
Internet Routing Hierarchy

- The Internet is composed of Autonomous Systems
- Each Autonomous System is an administrative entity that
  - Uses Interior Gateway Protocols (IGPs) to determine routing within the Autonomous System
  - Uses Exterior Gateway Protocols (EGPs) to interact with other Autonomous Systems

IGPs and EGPs

- IGPs provide routing information within your network (LAN, backbone links,etc)
- EGPs consider other networks outside your AS as a black box.

Internet Routing Architecture



Autonomous System: A collection of IP subnets and routers under the same administrative authority.

Interior Routing Protocol

Exterior Routing Protocol

Interior Gateway Protocols

- Four well known IGPs today
  - RIP
  - EIGRP
  - OSPF
  - ISIS

Exterior Gateway Protocols

- One single de-facto standard:
  - BGP

Routing's 3 Aspects

- Acquisition of information about the IP subnets that are reachable through an internet
  - static routing configuration information
  - dynamic routing information protocols (e.g., BGP4, OSPF, RIP, ISIS)
  - each mechanism/protocol constructs a Routing Information Base (RIB)

Routing Aspect #2

- Construction of a Forwarding Table
  - synthesis of a single table from all the Routing Information Bases (RIBs)
  - information about a destination subnet may be acquired multiple ways
  - a precedence is defined among the RIBs to arbitrate conflicts on the same subnet
  - Also called a Forwarding Information Base (FIB)

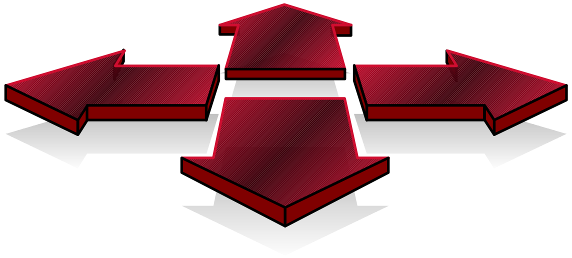
Routing #3

- Use of a Forwarding Table to forward individual packets
  - selection of the next-hop router and interface
  - hop-by-hop, each router makes an independent decision

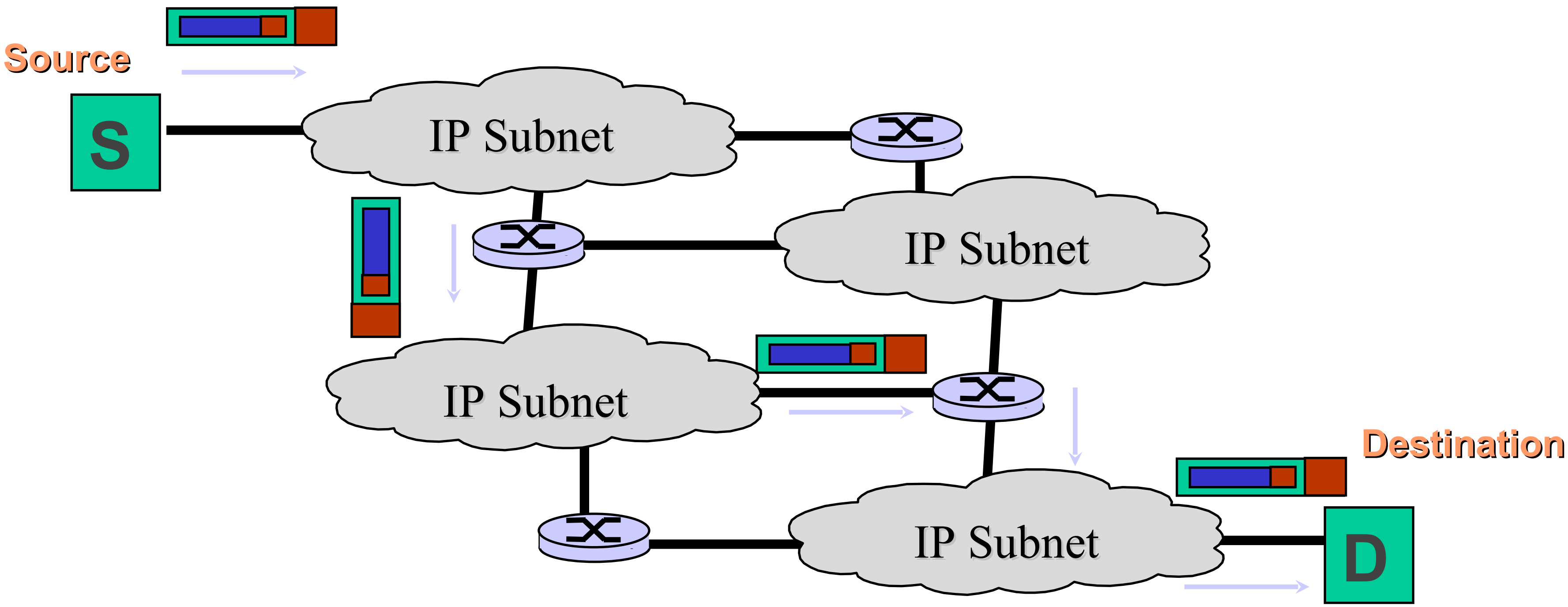
Routing versus Forwarding

- Routing = building maps and giving directions

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

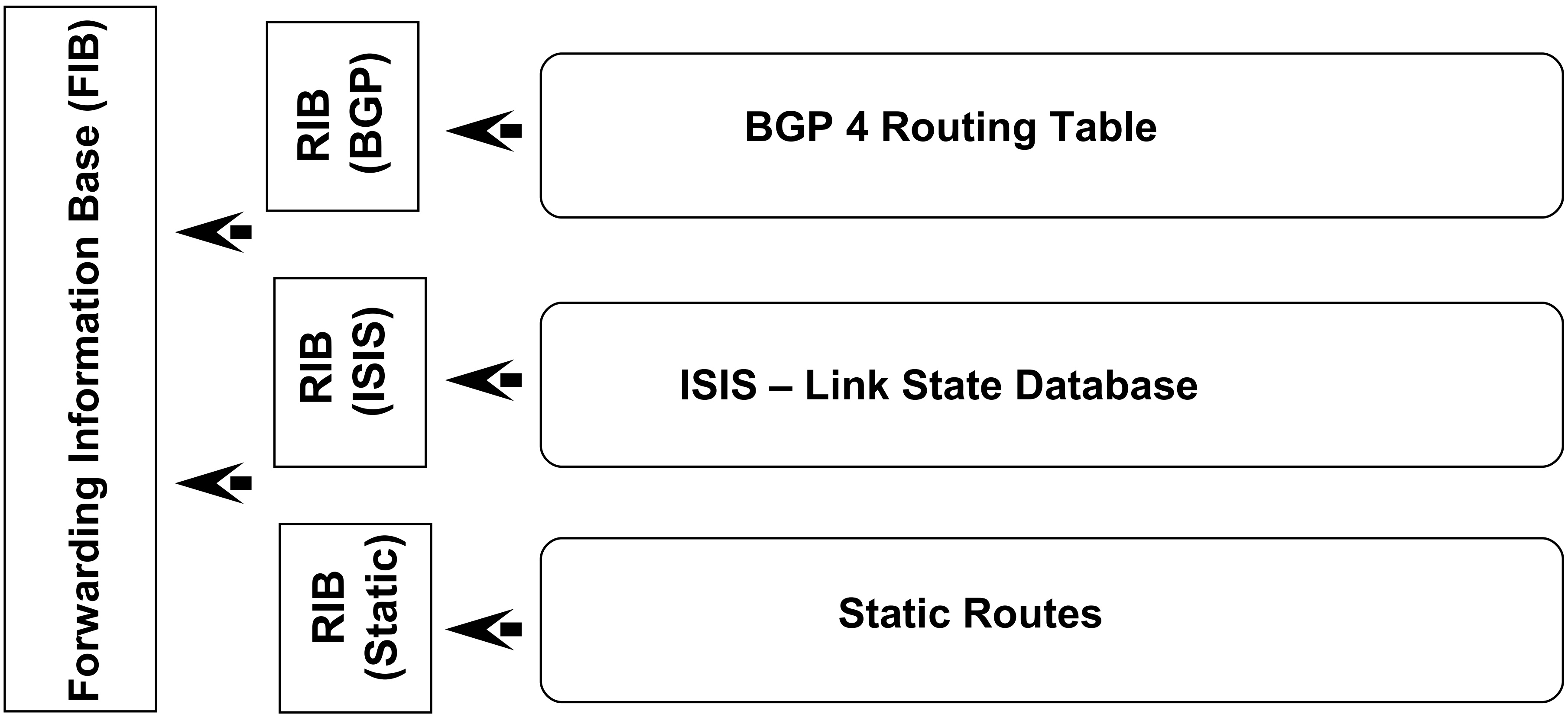


IP Forwarding



- Forwarding decisions:
  - **<u>Destination address</u>**
  - class of service (fair queuing, precedence, others)
  - local requirements (packet filtering)

Routing Tables Feed the Forwarding Table



RIB Construction

- Each routing protocol builds its own Routing Information Base (RIB)
- Each protocol handles route “costs” in its own way.

FIB Construction

- There is only ONE forwarding table!
- An algorithm is used to choose one next-hop toward each IP destination known by any routing protocol
  - the set of IP destinations present in any RIB are collected
  - if a particular IP destination is present in only one RIB, that RIB determines the next hop forwarding path for that destination

FIB Construction

- Choosing FIB entries, cont..
  - if a particular IP destination is present in multiple RIBs, then a precedence is defined to select which RIB entry determines the next hop forwarding path for that destination
  - This process normally chooses exactly one next-hop toward a given destination
- There are no standards for this; it is an implementation (vendor) decision

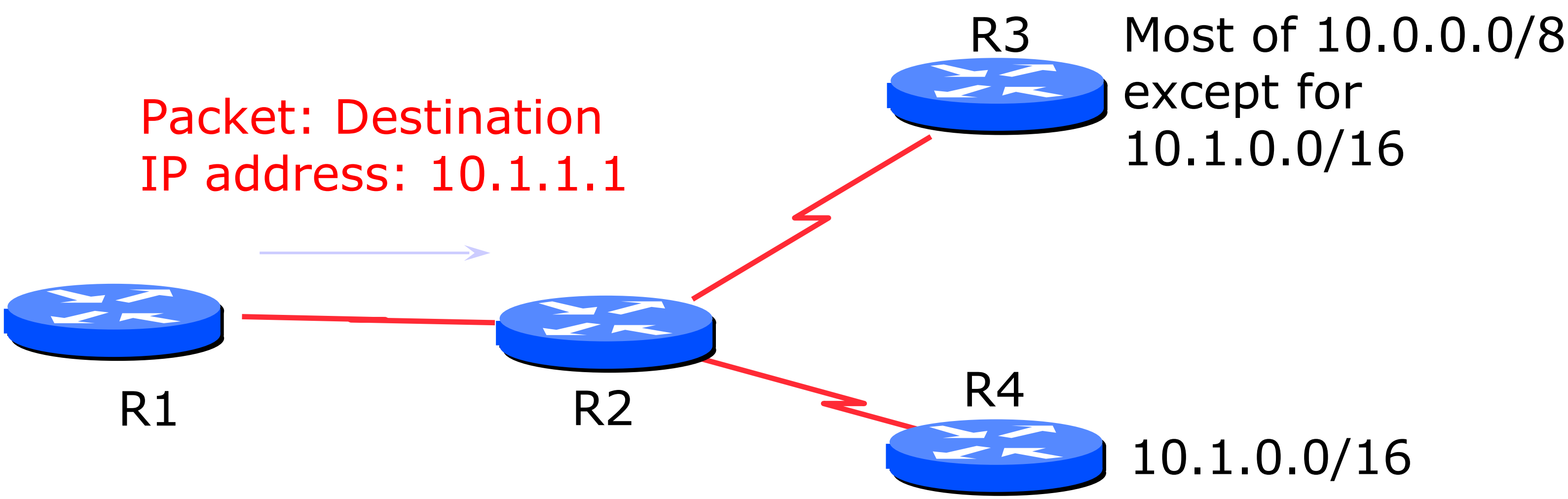
FIB Contents

- IP subnet and mask (or length) of destinations
  - can be the “default” IP subnet
- IP address of the “next hop” toward that IP subnet
- Interface id of the subnet associated with the next hop
- Optional: cost metric associated with this entry in the forwarding table

IP routing

- Default route
  - where to send packets if there is no entry for the destination in the routing table
  - most machines have a single default route
  - often referred to as a default gateway
  - 0.0.0.0/0
    - matches all possible destinations, but is usually not the longest match

IP route lookup: Longest match routing

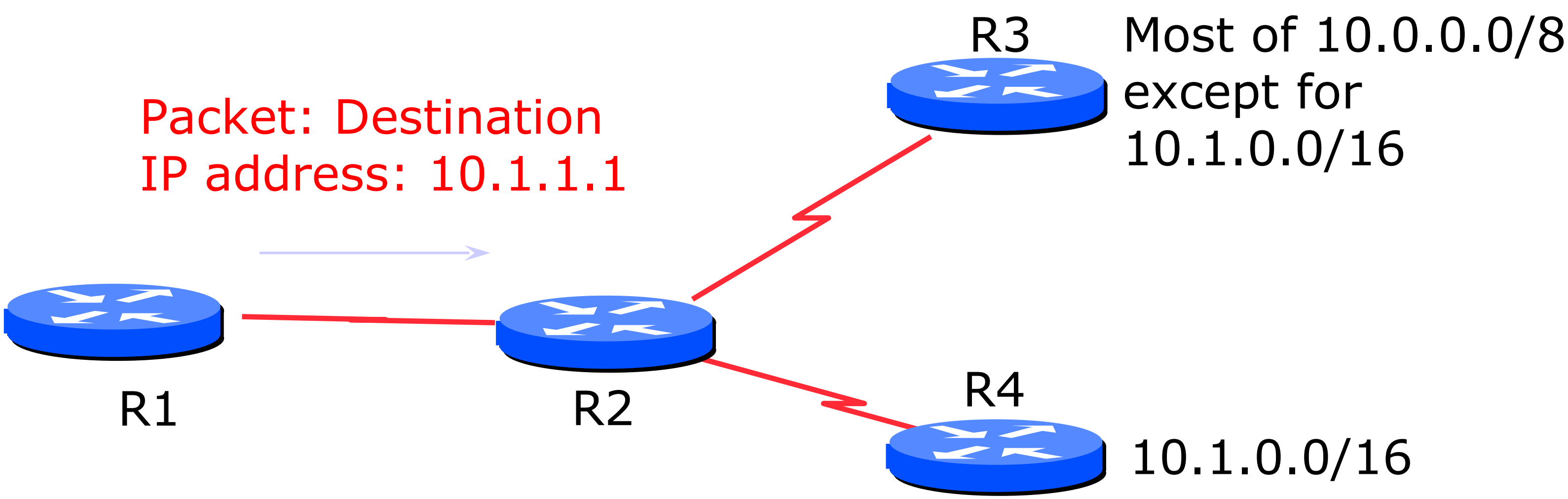


Based on destination IP address

R2's IP forwarding table

10.0.0.0/8 → R3
10.1.0.0/16 → R4
20.0.0.0/8 → R5
0.0.0.0/0 → R1

IP route lookup: Longest match routing

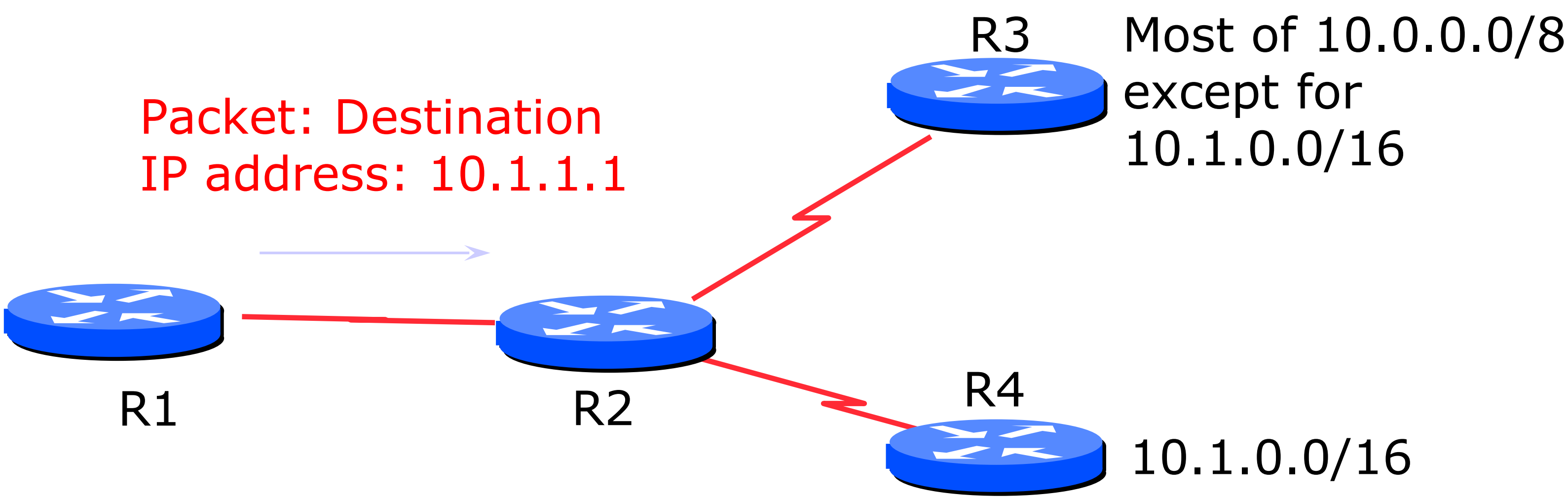


Based on destination IP address

R2's IP forwarding table

10.0.0.0/8 → R3	10.1.1.1 & FF.00.00.00
10.1.0.0/16 → R4	vs.
20.0.0.0/8 → R5	10.0.0.0 & FF.00.00.00
0.0.0.0/0 → R1	Match! (length 8)

IP route lookup: Longest match routing



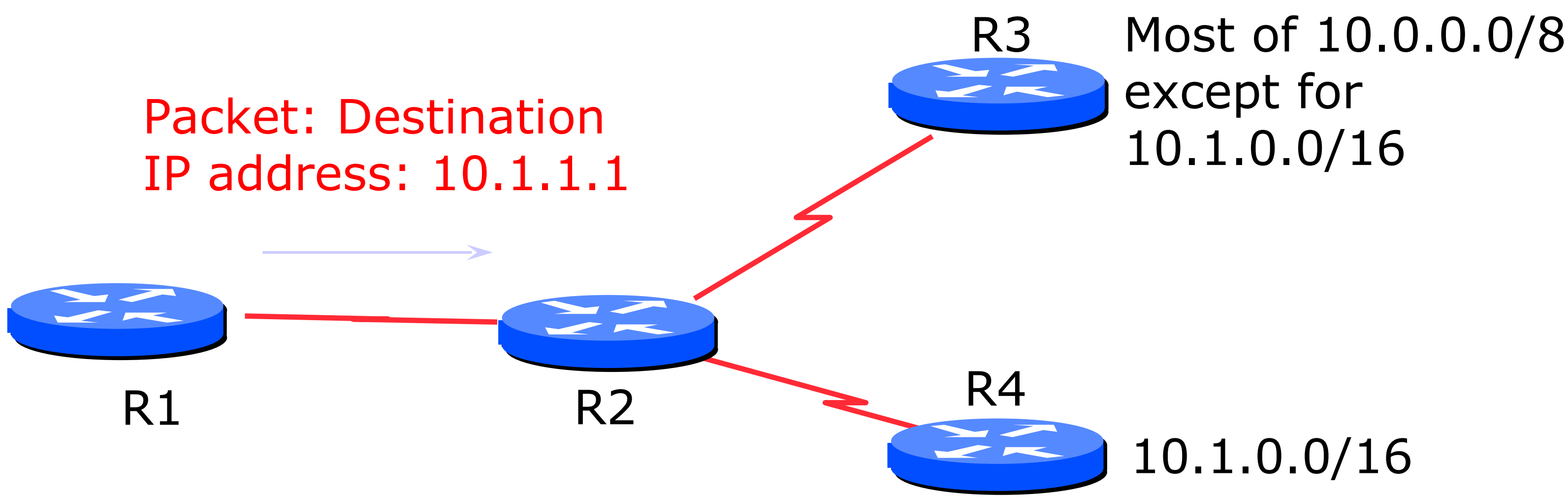
Based on destination IP address

R2's IP forwarding table

10.0.0.0/8 → R3
10.1.0.0/16 → R4
20.0.0.0/8 → R5
0.0.0.0/0 → R1

10.1.1.1 & FF.FF.00.00
vs.
10.1.0.0 & FF.FF.00.00
Match! (length 16)

IP route lookup: Longest match routing



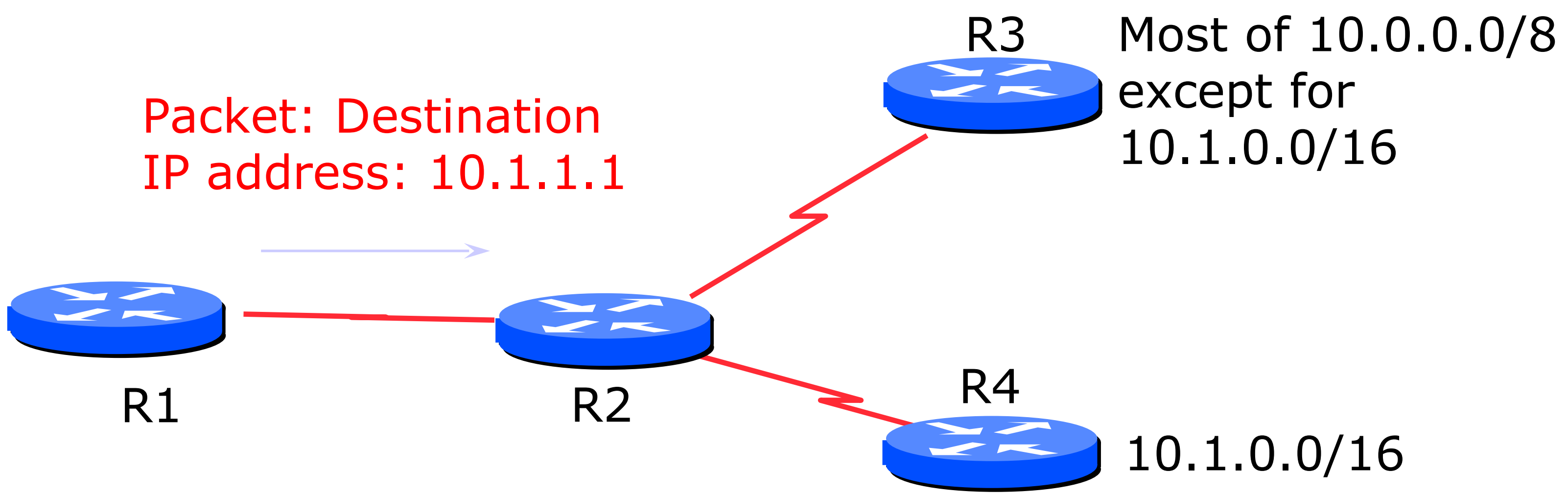
Based on destination IP address

R2's IP forwarding table

10.0.0.0/8 → R3
10.1.0.0/16 → R4
20.0.0.0/8 → R5
0.0.0.0/0 → R1

10.1.1.1 & FF.00.00.00
vs.
20.0.0.0 & FF.00.00.00
No Match!

IP route lookup: Longest match routing



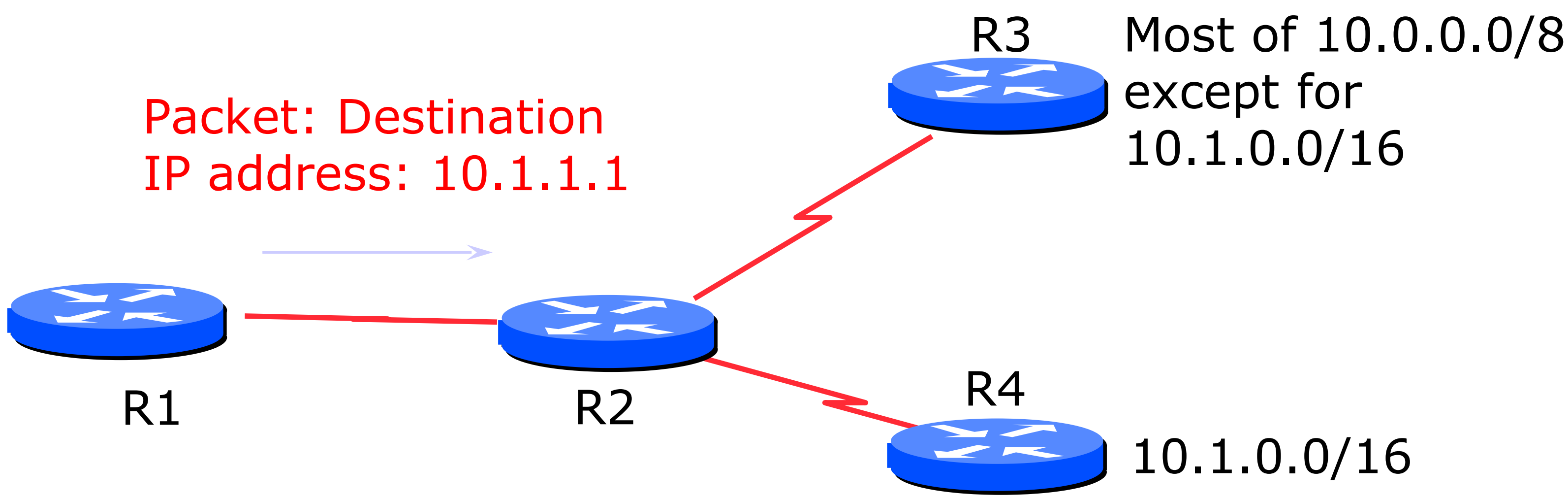
Based on destination IP address

R2's IP forwarding table

10.0.0.0/8 → R3
10.1.0.0/16 → R4
20.0.0.0/8 → R5
0.0.0.0/0 → R1

10.1.1.1 & 00.00.00.00
vs.
0.0.0.0 & 00.00.00.00
Match! (length 0)

IP route lookup: Longest match routing



Based on destination IP address

R2's IP forwarding table

10.0.0.0/8 → R3
10.1.0.0/16 → R4
20.0.0.0/8 → R5
0.0.0.0/0 → R1

This is the longest matching prefix (length 16). "R2" will send the packet to "R4".

IP route lookup: Longest match routing

- Most specific/longest match always wins!!
  - Many people forget this, even experienced ISP engineers
- Default route is 0.0.0.0/0
  - Can handle it using the normal longest match algorithm
  - Matches everything. Always the shortest match.

Distance Vector and Link State

- Distance Vector
  - Accumulates a metric hop-by-hop as the protocol messages traverse the subnets
- Link State
  - Builds a network topology database
  - Computes best path routes from current node to all destinations based on the topology

Distance Vector Protocols

- Each router only advertises to its neighbors, its “distance” to various IP subnets
- Each router computes its next-hop routing table based on least cost determined from information received from its neighbors and the cost to those neighbors

Why not use RIP?

- RIP is a Distance Vector Algorithm
  - Listen to neighbouring routes
  - Install all routes in routing table
    - Lowest hop count wins
  - Advertise all routes in table
    - Very simple, very stupid
- Only metric is hop count
- Network is max 16 hops (not large enough)
- Slow convergence (routing loops)
- Poor robustness

EIGRP

- “Enhanced Interior Gateway Routing Protocol”
- Predecessor was IGRP which was classfull
  - IGRP developed by Cisco in mid 1980s to overcome scalability problems with RIP
- Cisco proprietary routing protocol
- Distance Vector Routing Protocol
  - Has very good metric control
- Still maybe used in some enterprise networks?
  - Multi-protocol (supports more than IP)
  - Exhibits good scalability and rapid convergence
  - Supports unequal cost load balancing

Link State Protocols

Link State Protocols

- Each router “multicasts” to all the routers in the network the state of its locally attached links and IP subnets
- Each router constructs a complete topology view of the entire network based on these link state updates and computes its next-hop routing table based on this topology view

Link State Protocols

- Attempts to minimize convergence times and eliminate non-transient packet looping at the expense of higher messaging overhead, memory, and processing requirements
- Allows multiple metrics/costs to be used

IS-IS

- "Intermediate System to Intermediate System"
- Selected in 1987 by ANSI as OSI intradomain routing protocol (CLNP – connectionless network protocol)
  - Based on work by DEC for DECnet/OSI (DECnet Phase V)
- Extensions for IP developed in 1988
  - NSFnet deployed its IGP based on early ISIS-IP draft

IS-IS (cont)

- Adopted as ISO proposed standard in 1989
  - Integrated ISIS supports IP and CLNP
- Debate between benefits of ISIS and OSPF
  - Several ISPs chose ISIS over OSPF for a number of reasons.
- 1994-date: deployed by several larger ISPs
- Developments continuing in IETF in parallel with OSPF

OSPF

- Open Shortest Path First
  - “Open” means it is public domain
  - Uses “Shortest Path First” algorithm – sometimes called “the Dijkstra algorithm”
- IETF Working Group formed in 1988 to design an IGP for IP
- OSPF v1 published in 1989 – RFC1131
- OSPF v2 published in 1991 – RFC1247
- Developments continued through the 90s and today
  - OSPFv3 based on OSPFv2 designed to support IPv6

Link State Algorithm

- Each router contains a database containing a map of the whole topology
  - Links
  - Their state (including cost)
- All routers have the same information
- All routers calculate the best path to every destination
- Any link state changes are flooded across the network
  - "Global spread of local knowledge"

Summary

- Now know:
  - Difference between static routes, RIP, OSPF and IS-IS.
  - Difference between Routing and Forwarding
  - A Dynamic Routing Protocol should be used in any ISP network
  - Static routes don’t scale
  - RIP doesn’t scale (and is obsolete)