Network Routing

Computer Networks

Faculty of Information Technology
Hanoi University

1. Definition (Network routing)

What is network routing?

- **Network Routing** is an umbrella term for the set of **protocols** that determine the path that data follows in order to travel across multiple networks from its source to its destination.
- Data is routed from its source to its destination through a series of routers, and across multiple networks.

Forwarding vs Routing

- Forwarding
  - Process of selecting an output port based on destination address and routing/switching table
  - Simple and well-defined, performed locally at a node
- Routing
  - Process by which a routing table is built
    - How the network topology information is acquired
    - Determining the best path through the network
  - Depends on complex distributed algorithms

Forwarding Table vs Routing Table

- Forwarding Table
  - Maps a MAC UID to an outgoing interface and/or MAC address of the next hop
  - Optimized for the lookup of data
  - Usually implemented in hardware
- Routing Table
  - Built by the routing algorithm
  - Maps a Network UID to next hop, cost, forwarding policy etc.
  - Optimized for calculating/managing changes in topology
  - Usually implemented in software

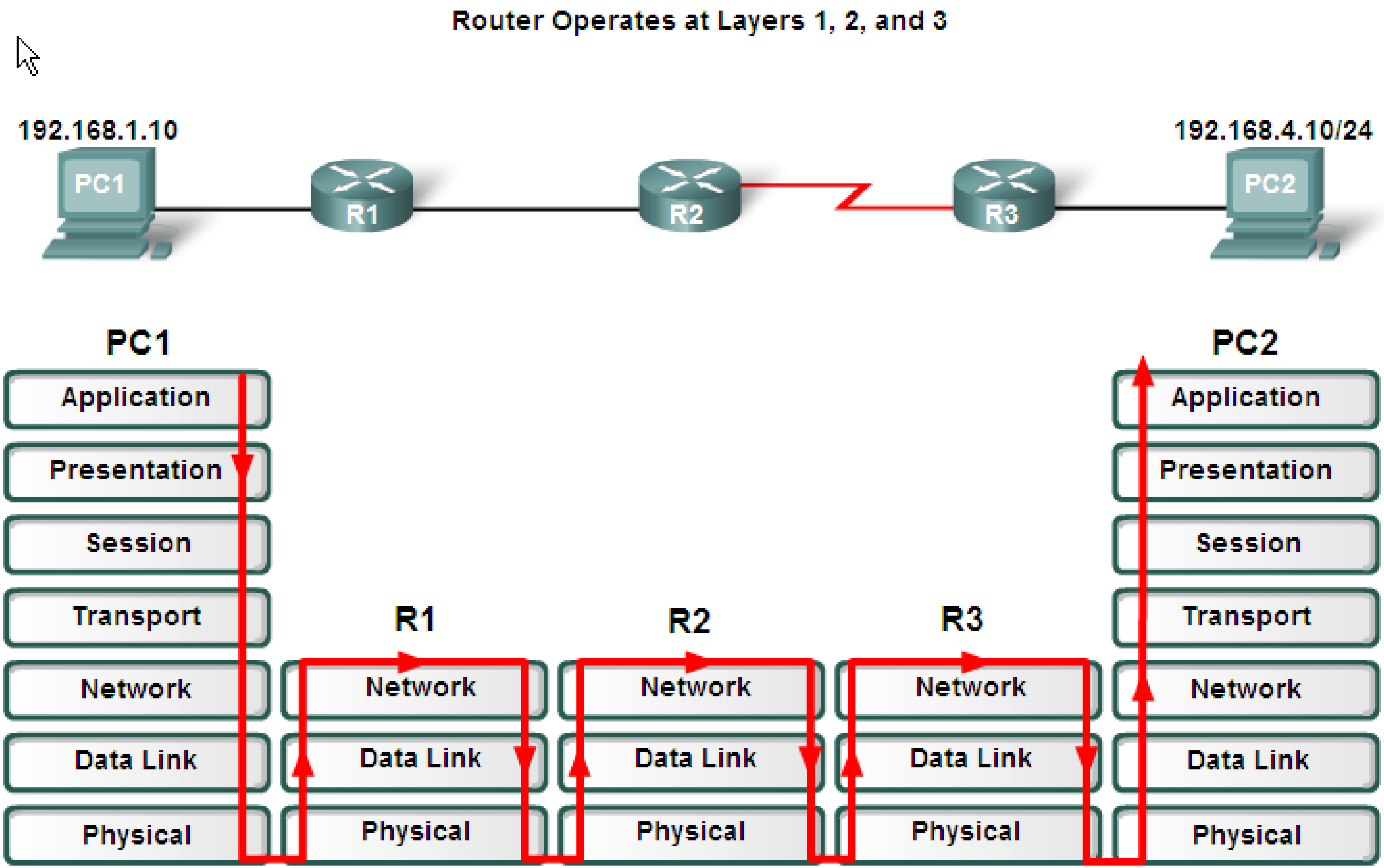
Example: Forwarding table

Host MAC Address	Port
00 00 80 45 FE 21	5
00 00 80 45 DA 47	3
00 40 00 80 45 FE	2
00 40 80 10 AA 21	1
00 00 80 00 FF AB	5

Router

- Router is a small computer (hardware and software as in a computer).
  - CPU
  - RAM
  - NVRAM
  - ROM
  - Flash memory
  - IN/OUT Ports
  - IOS

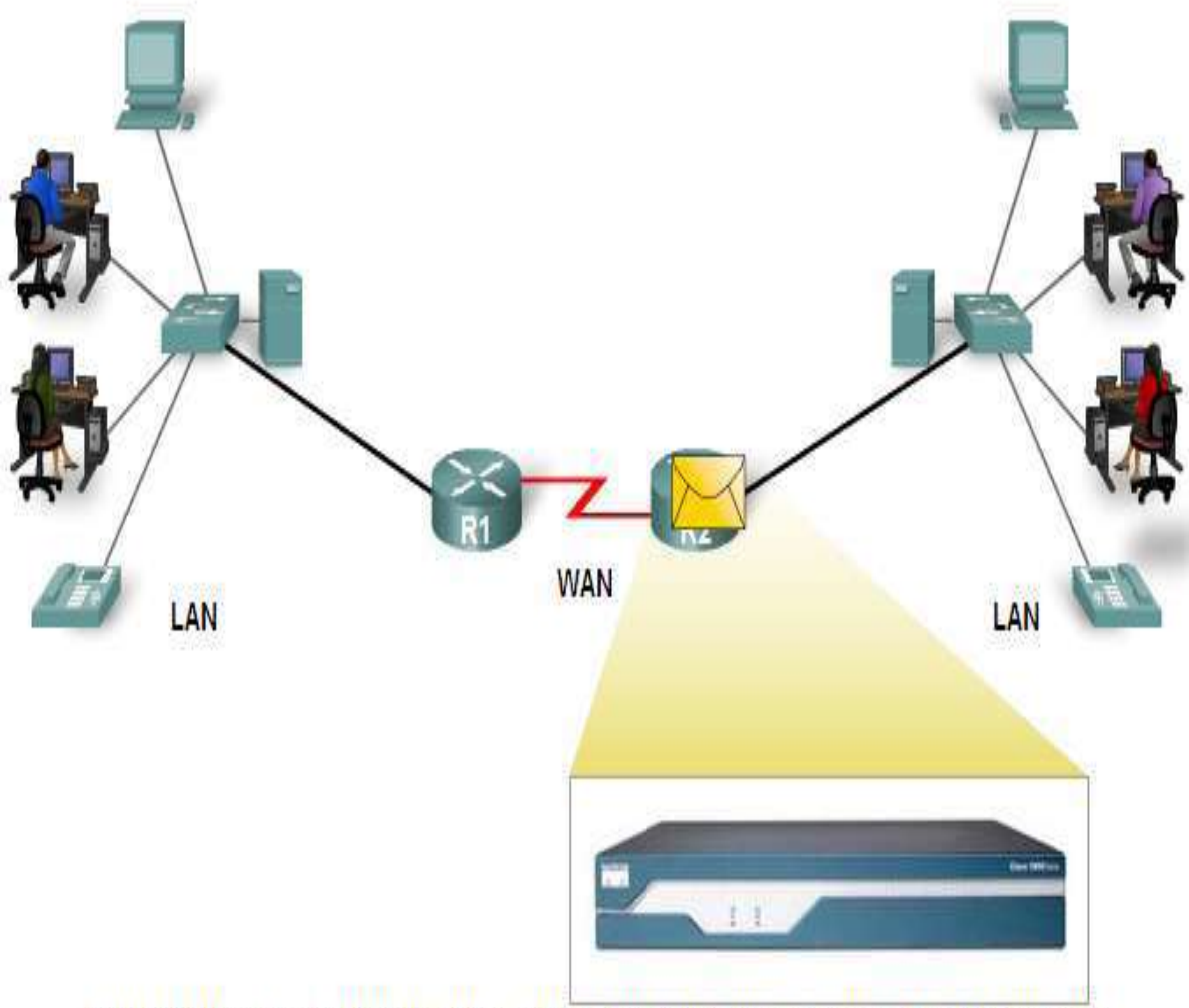
Router



Red arrows indicate flow through the OSI layers.

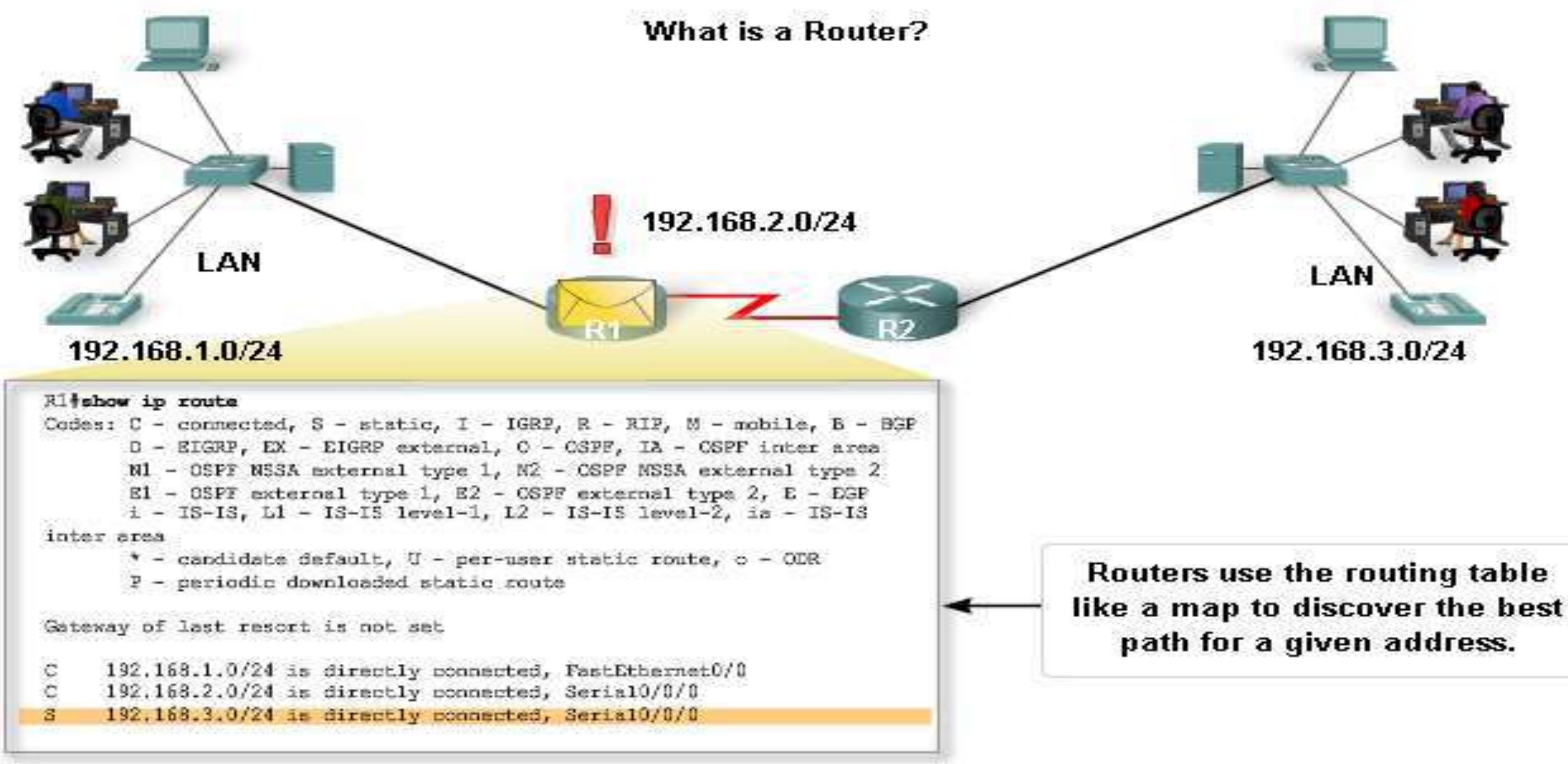
Router

- Router is centre of networks. It is used to connect networks (LAN, WAN).
- Router has some ports, each port belongs to different networks.



Routers direct packets to their proper destination. Routers connect different media.

Router


What is a Router?
192.168.2.0/24
LAN
R1
R2
LAN
192.168.1.0/24
192.168.3.0/24
R1#show ip route
Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B - BGP
D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
E1 - OSPF external type 1, E2 - OSPF external type 2, E - EGP
i - IS-IS, L1 - IS-IS level-1, L2 - IS-IS level-2, ia - IS-IS
inter area
* - candidate default, U - per-user static route, o - ODR
P - periodic downloaded static route
Gateway of last resort is not set
C 192.168.1.0/24 is directly connected, FastEthernet0/0
C 192.168.2.0/24 is directly connected, Serial0/0/0
S 192.168.3.0/24 is directly connected, Serial0/0/0
Routers use the routing table like a map to discover the best path for a given address.

Example: Routing table

```
C:\WINNT\System32\cmd.exe
Microsoft(R) Windows NT(TM)
(C) Copyright 1985-1996 Microsoft Corp.

C:\>route print

Active Routes:

  Network Address          Netmask  Gateway Address        Interface  Metric
          0.0.0.0          0.0.0.0     199.98.126.2    199.98.126.16       1
      38.208.233.0    255.255.255.0     199.98.126.2    199.98.126.16       1
        127.0.0.0        255.0.0.0        127.0.0.1        127.0.0.1       1
      199.98.126.0    255.255.255.0    199.98.126.16    199.98.126.16       1
     199.98.126.16  255.255.255.255        127.0.0.1        127.0.0.1       1
    199.98.126.255  255.255.255.255    199.98.126.16    199.98.126.16       1
        224.0.0.0        224.0.0.0    199.98.126.16    199.98.126.16       1
  255.255.255.255  255.255.255.255    199.98.126.16    199.98.126.16       1

C:\>
```

2. Routing Protocols, Algorithms and Metric

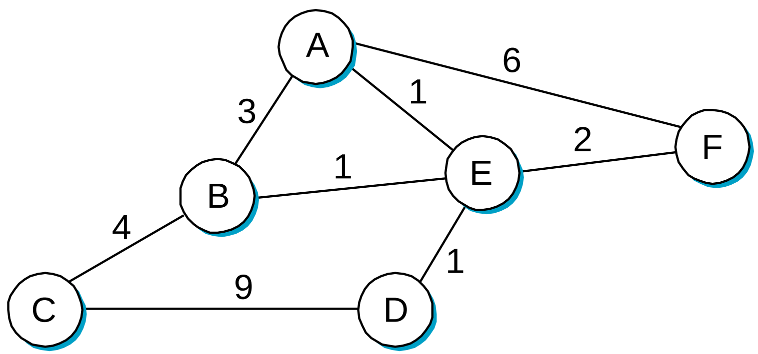
Routing protocol at Network layer

- A major component of the network layer is routing protocol
- Routing protocols use routing algorithms
- Job of a routing algorithm: Given a set of routers with links connecting the routers, find a "good" path from the source to the destination

Modeling a Network

- A network can be modeled by a graph
  - Routers are represented by nodes
  - Physical links between routers are represented by edges
  - Attached computers are ignored
  - Each edge is assigned a weight representing the “cost” of sending a packet across that link
  - The total cost of a path is the sum of the costs of the edges

Network as a Graph



- Problem: Find lowest cost path between two nodes
- Factors
  - Static/Dynamic topology
  - Cost
    - Path length
    - Reliability
    - Delay
    - Bandwidth
    - Load
    - Communication cost

Routing Algorithms

- Routing algorithms that solve a routing problem are based on shortest-path algorithms
- Two types of routing algorithms
  - **Non-Adaptive**
    - form of routing that occurs when a router uses a manually-configured routing entry (by a network administrator)
    - Also called **static routing**
  - **Adaptive**
    - Dynamic information like current topology, load, delay, etc. to select routes is used
    - Also called **dynamic routing**

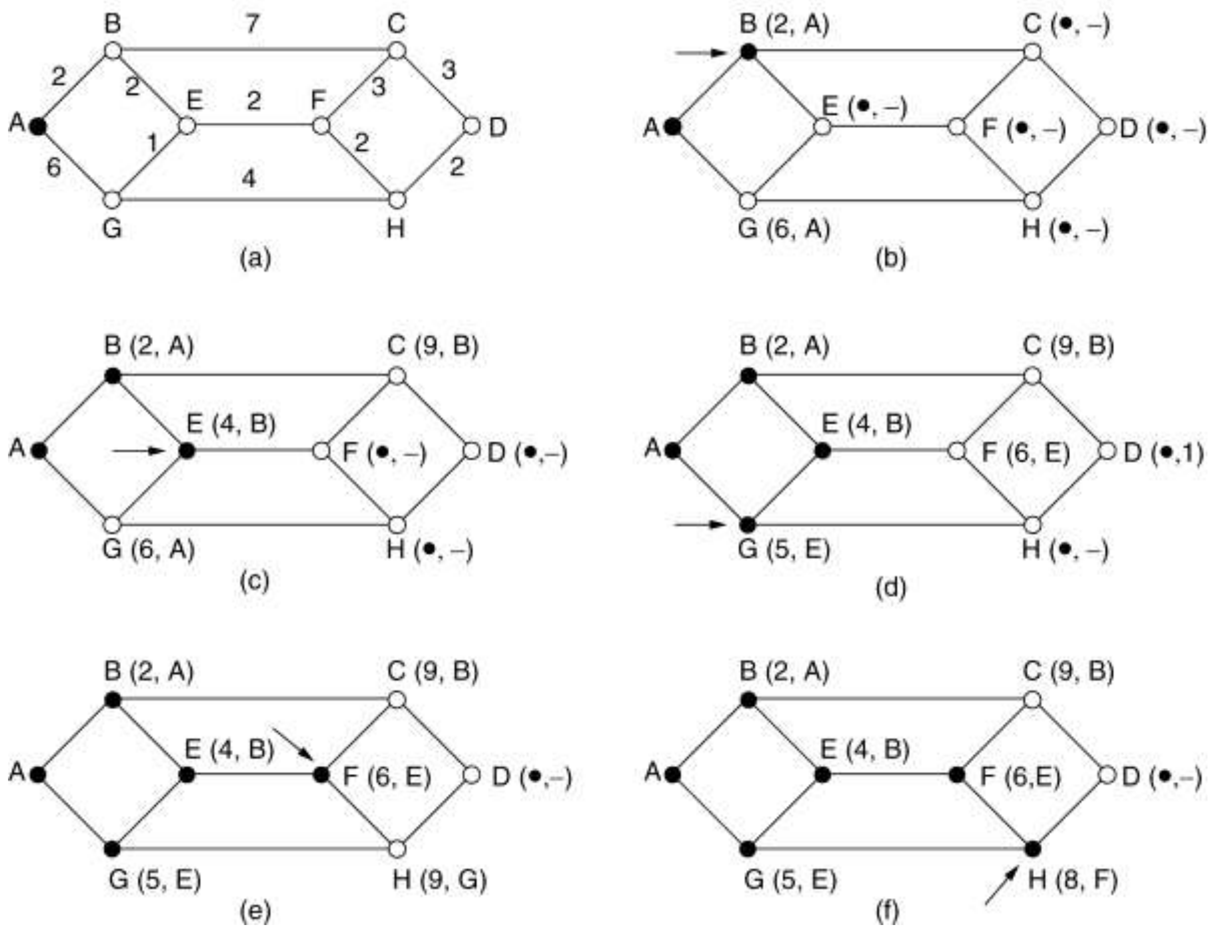
Routing Algorithms

- Correctness, simplicity and minimality
- Robustness
  - During years of continuous operation, should handle all kinds of hardware and software failures, changes in topology and traffic patterns etc. etc.
- Stability
- Fairness and Optimality
- Scalability
  - For routing inside autonomous systems, the internal routing protocols are used e.g., RIP (Interior Gateway Protocols - IGP)
  - For routing between autonomous systems, external routing protocols like EGP (External Gateway Protocol), or BGP (Border Gateway Protocol) are used (Exterior Gateway Protocols - EGP)

Shortest Path Routing

- If we know the topology of the network we can compute the shortest path using Graph theory
  - *Shortest* in terms of which metric
    - No. of hops, geographical distance, queuing and transmission delay, bandwidth, average traffic, communication cost, mean queue length, etc. etc. etc.
  - Once the graph is complete, use Dijkstra's algorithm

Shortest Path Routing



- Basic Idea: During each step, select a newly reachable node at the lowest cost, and add the edge to that node, to the graph built so far

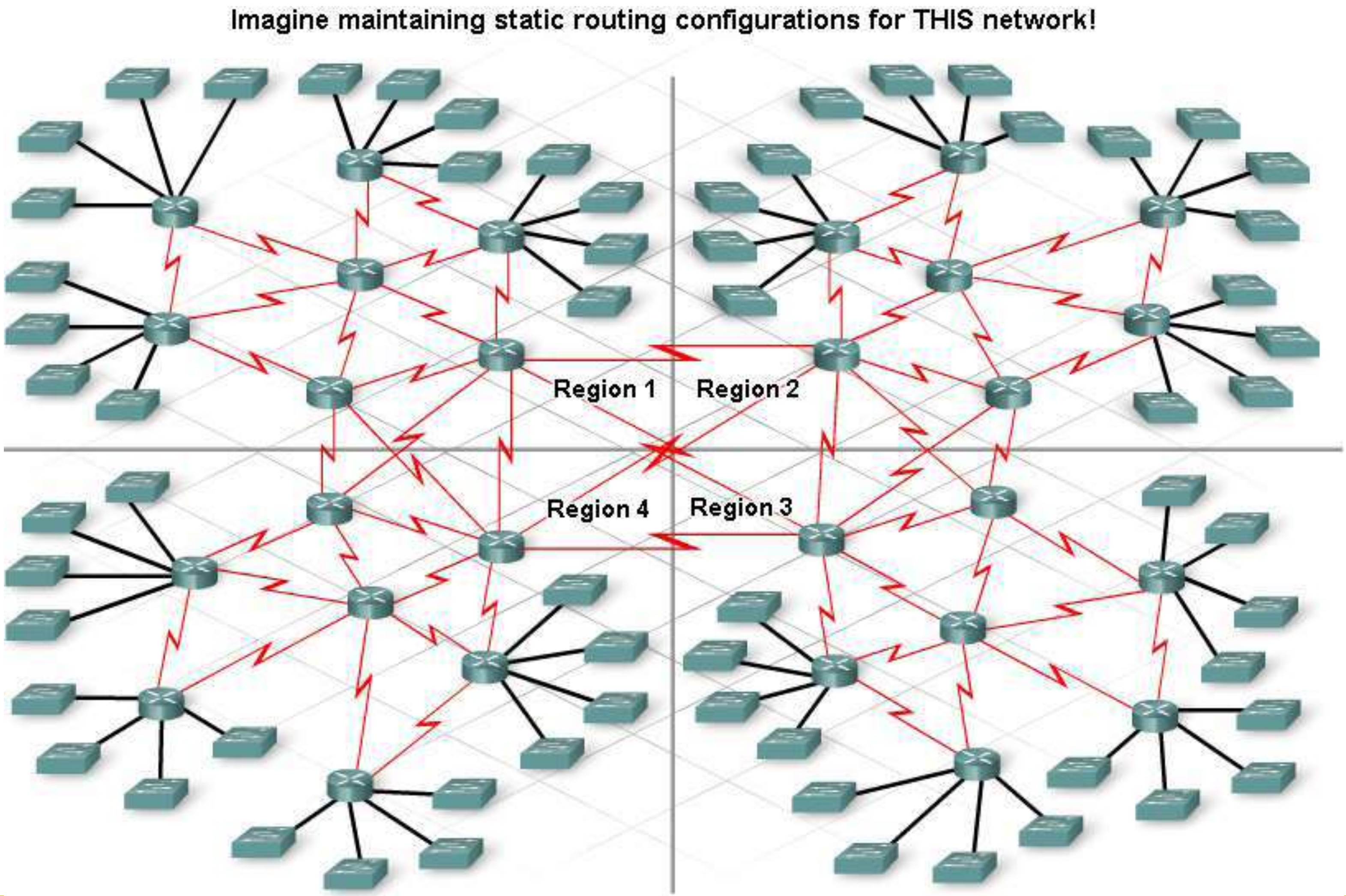
Non-Adaptive Algorithms (static routing)

Non-Adaptive Algorithms

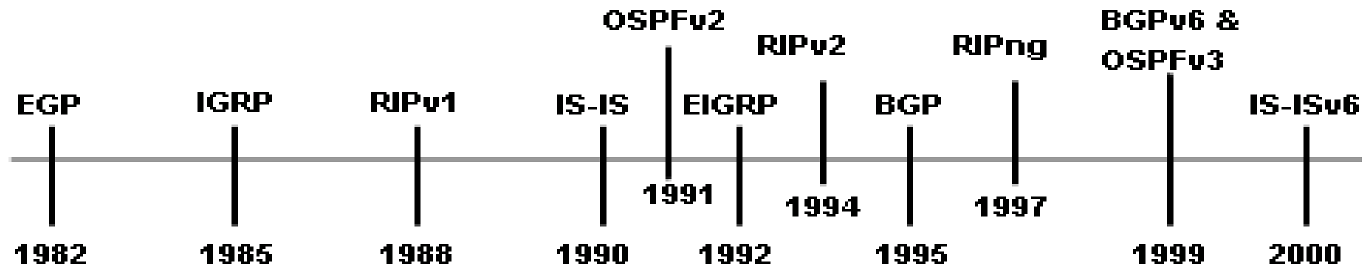
- Non Adaptive routing algorithm is also known as a static routing algorithm.
- When booting up the network, the routing information stores to the routers.
- Non Adaptive routing algorithms do not take the routing decision based on the network topology or network traffic.

Adaptive Algorithms (dynamic routing)

Problem of non-adaptive algorithm

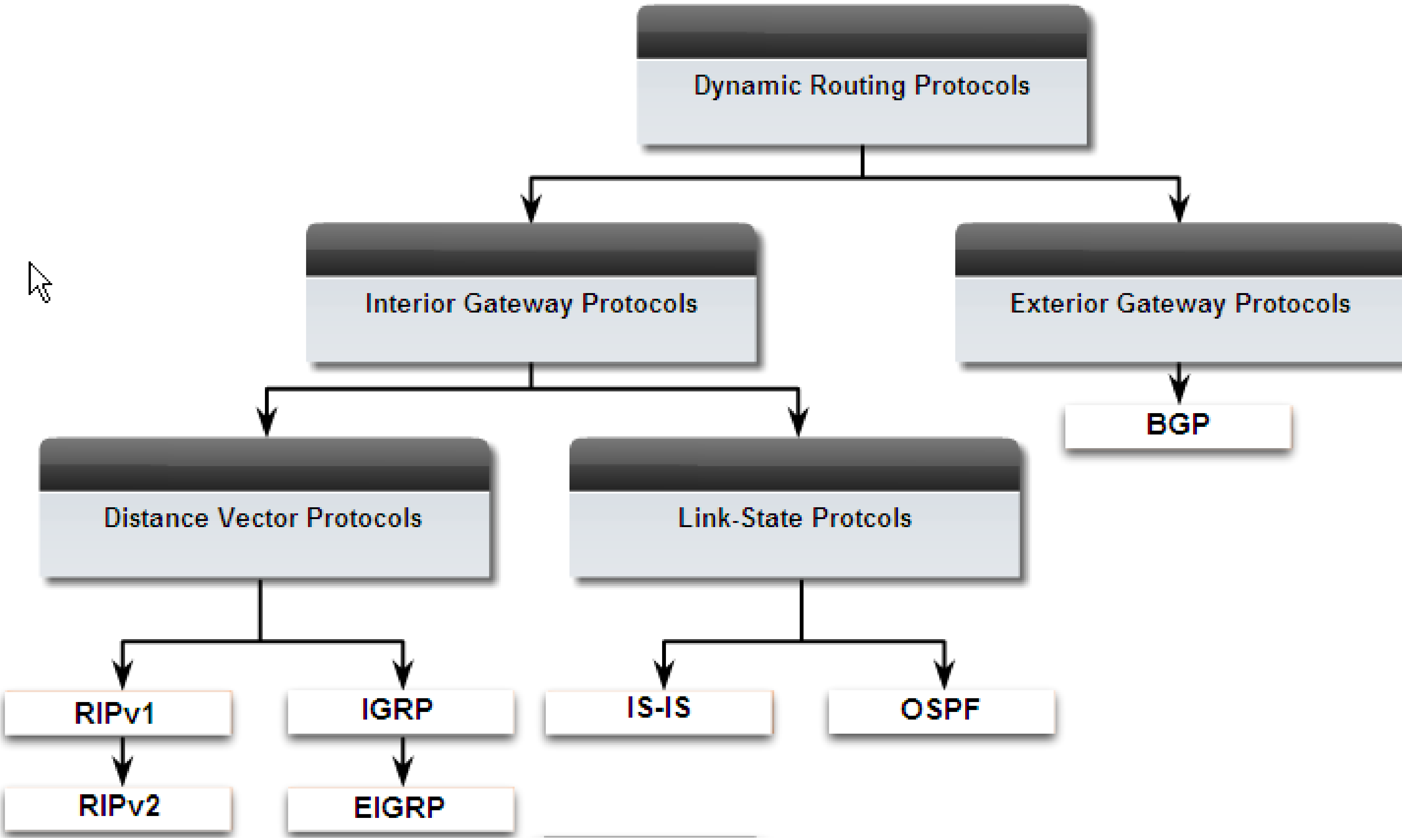


Routing Protocols Evolution and Classification



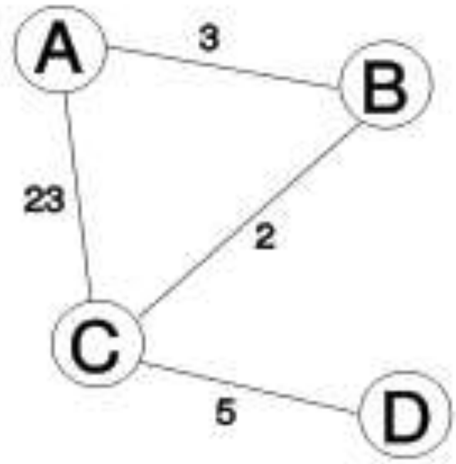
	Interior Gateway Protocols				Exterior Gateway Protocols
	Distance Vector Routing Protocols		Link State Routing Protocols		Path Vector
Classful	RIP	IGRP			EGP
Classless	RIPv2	EIGRP	OSPFv2	IS-IS	BGPv4
IPv6	RIPng	EIGRP for IPv6	OSPFv3	IS-IS for IPv6	BGPv4 for IPv6

Highlighted routing protocols are the focus of this course.


Dynamic Routing Protocols
Interior Gateway Protocols
Exterior Gateway Protocols
BGP
Distance Vector Protocols
Link-State Protcols
RIPv1
IGRP
IS-IS
OSPF
RIPv2
EIGRP

a. Distance Vector Routing Protocol (DVRP)

- A type of Interior Gateway Protocols (IGP)
- Uses the Bellman-Ford algorithm to calculate paths
  1. Each router calculates the distances between itself and all neighboring routers within the AS and stores this information as a table
  2. Each router periodically sends its table to all neighboring routers
  3. When a router receives distance tables from its neighbors, it calculates the shortest routes to all other routers and updates its own table to reflect any changes
     - Upon receipt of an update, for each destination in its table, a router:
       - Compares the metric in its local table with the metric in the neighbor's table plus the cost of reaching that neighbor
       - If the path via the neighbor has a lower cost, the router updates its local table



T=3

from A	via A	via B	via C	via D
to A				
to B		3	25	
to C		5	23	
to D		10	28	

from B	via A	via B	via C	via D
to A	3		7	
to B				
to C	8		2	
to D	13		7	

from C	via A	via B	via C	via D
to A	23	5		15
to B	26	2		12
to C				
to D	33	9		5

from D	via A	via B	via C	via D
to A			10	
to B			7	
to C			5	
to D				

T=0

from A	via A	via B	via C	via D
to A				
to B		3		
to C			23	
to D				

from B	via A	via B	via C	via D
to A	3			
to B				
to C			2	
to D				

from C	via A	via B	via C	via D
to A	23			
to B		2		
to C				
to D				5

from D	via A	via B	via C	via D
to A				
to B				
to C			5	
to D				

T=1

from A	via A	via B	via C	via D
to A				
to B		3	25	
to C		5	23	
to D			28	

from B	via A	via B	via C	via D
to A	3		25	
to B				
to C	26		2	
to D			7	

from C	via A	via B	via C	via D
to A	23	5		
to B	26	2		
to C				
to D				5

from D	via A	via B	via C	via D
to A			28	
to B			7	
to C			5	
to D				

T=2

from A	via A	via B	via C	via D
to A				
to B		3	25	
to C		5	23	
to D		10	28	

from B	via A	via B	via C	via D
to A	3		7	
to B				
to C	8		2	
to D	31		7	

from C	via A	via B	via C	via D
to A	23	5		33
to B	26	2		12
to C				
to D	51	9		5

from D	via A	via B	via C	via D
to A			10	
to B			7	
to C			5	
to D				

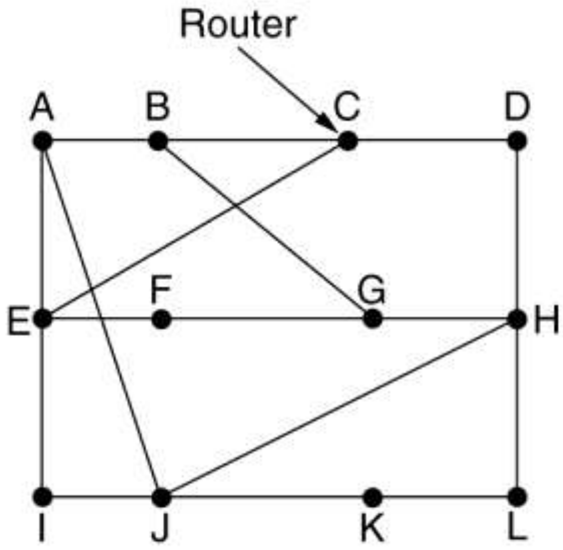
T=3

from A	via A	via B	via C	via D
to A				
to B		3	25	
to C		5	23	
to D		10	28	

from B	via A	via B	via C	via D
to A	3		7	
to B				
to C	8		2	
to D	13		7	

from C	via A	via B	via C	via D
to A	23	5		15
to B	26	2		12
to C				
to D	33	9		5

from D	via A	via B	via C	via D
to A			10	
to B			7	
to C			5	
to D				



(a)

To	A	I	H	K	New estimated delay from J	Line
A	0	24	20	21	8	A
B	12	36	31	28		A
C	25	18	19	36		I
D	40	27	8	24		H
E	14	7	30	22		I
F	23	20	19	40		I
G	18	31	6	31		H
H	17	20	0	19	12	H
I	21	0	14	22	10	I
J	9	11	7	10	0	–
K	24	22	22	0	6	K
L	29	33	9	9		K
	JA delay is 8	JI delay is 10	JH delay is 12	JK delay is 6	New routing table for J	

Vectors received from J's four neighbors

(b)

Input from A, I, H, K, and the new routing table for J

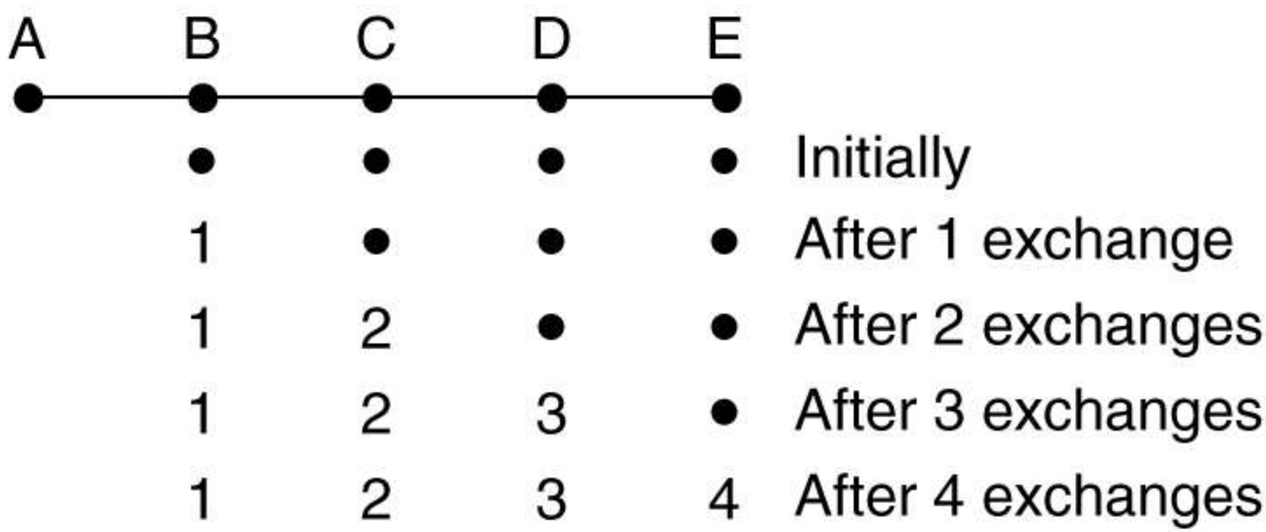
Qs: How to estimate route from J to G?

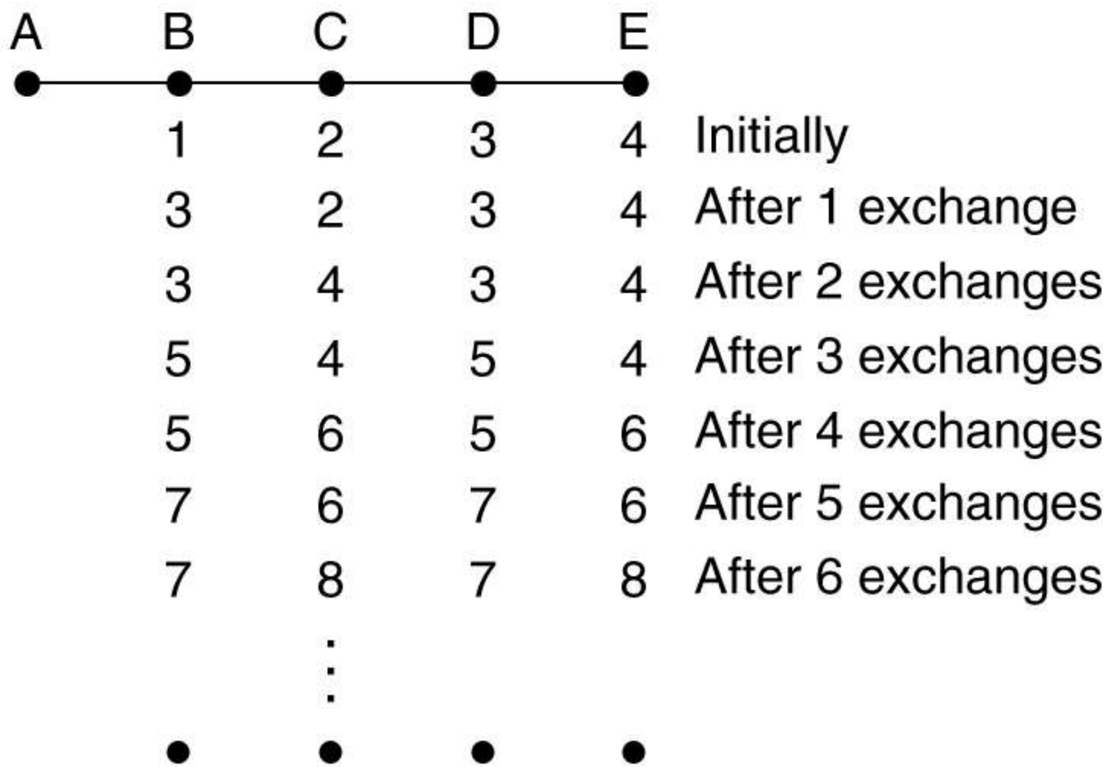
Critique of DVRP

- Advantages
  - Simple to handle and configure
  - Less computational complexity & message overhead
- Disadvantages
  - Does not scale well
    - Slow convergence
      - Convergence is the process of getting consistent routing information to all nodes
  - Count-to-Infinity Problem ('Routing by Rumour')

The count-to-infinity problem

- Assume A was down and just came back online
  - Its neighbors will update their tables with this new information hop-by-hop





- **Basic problem:** when X tells Y it has a path to Z, Y has no way of knowing whether it itself lies on that path

Implementations of DVRP

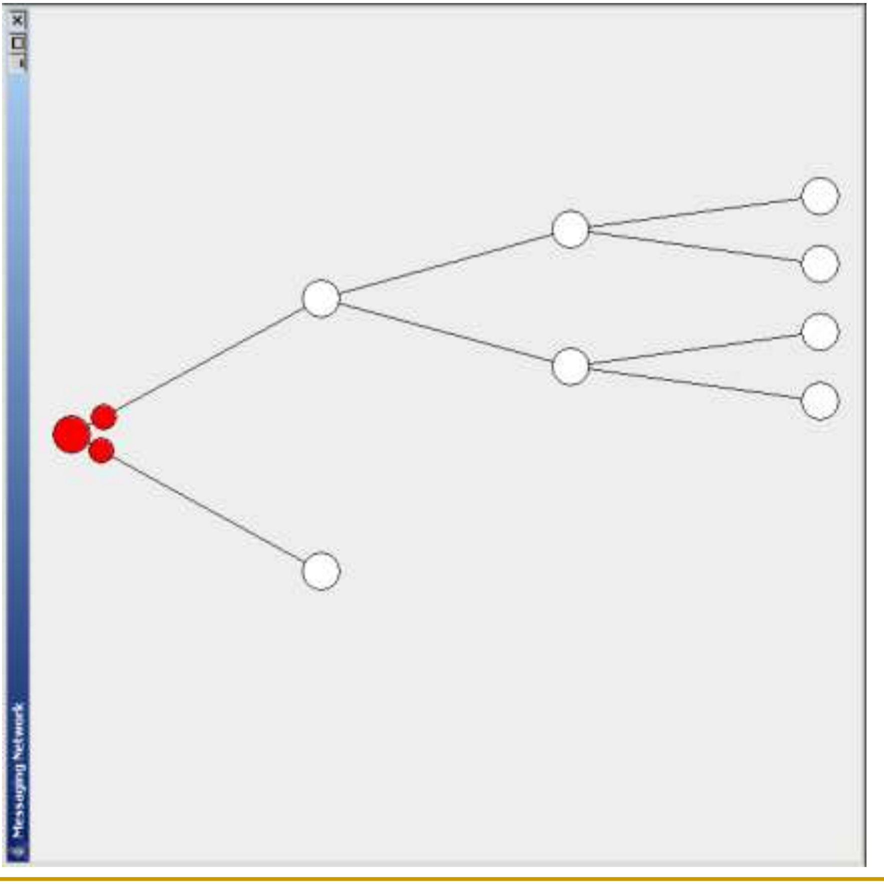
- Examples of real world implementation of distance-vector routing algorithm are:
  - Routing Information Protocol (RIP)
  - Interior Gateway Routing Protocol (IGRP)

b. Link State Routing Protocol

- The network topology and all link costs are known.
- Algorithm used: Dijkstra's Algorithm
- When a network link changes state, a notification, called a *link state advertisement* (LSA) is *flooded* throughout the network

Flooding

- **Basic idea:** Forward an incoming packet to every outgoing line, except the one it came through



Flooding

- **Basic problem:** How to avoid “drowning by packets”?
  - Use a hop counter: after a packet has been forwarded across *N* routers, it is discarded
    - Difficult find the right hop count
  - Be sure to flood a packet only once (i.e. avoid directed cycles)
    - Add a sequence number to each packet's header
    - Each router maintains a private sequence number. When it sends a new packet, it copies the sequence number into the packet, and increments its private sequence number
    - For each source router S, a router:
      - Keeps track of the highest sequence number seen from S
      - Whenever it receives a packet from S containing a sequence number lower than the one stored in its table, it discards the packet
      - Otherwise, it updates the entry for S and floods the packet on
  - **Selective Flooding**: only in the direction that makes sense

When to use Flooding

- Flooding makes sense only when robustness is needed
  - In military applications, the network must remain robust in the face of (extreme) hostility
  - Updating routing tables or distributed databases concurrently

Link State Routing Protocol

- All the routers note the change, and recompute their routes accordingly
  - It is done using only the local copy of the map, and without communicating in any other way with any other node, and calculating a sink tree to the other routers
  - OSPF (Open Shortest Path First) and IS-IS (Intermediate System to Intermediate System) are implementations of Link State routing algorithm

Link State Routing

Each router must do the following:

1. Discover its neighbors, learn their network address
2. Measure the delay or cost to each of its neighbors
3. Construct a packet telling all it has just learned
4. Send this packet to all other routers
5. Compute the shortest path to every other router

More details:

Learning about the Neighbors

- Periodically send HELLO packets on each line
  - Each router replies with its unique ID

Measuring Line Cost

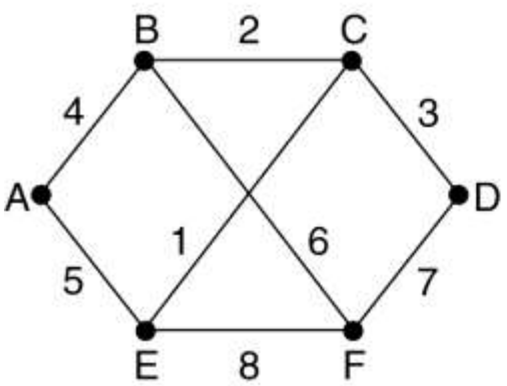
- Just send an ECHO packet through each interface, and measure the round-trip delay

Building Link State Packets

- Each node periodically makes up a short message, the *link-state advertisement* or the *link-state packet*, which:
  - Identifies the node which is producing it
  - Identifies all the other nodes to which it is directly connected
  - Includes a *sequence number*, which increases every time the source node makes up a new version of the message

Building Link State Packets

(a) A subnet (b) The link state packets for this subnet



(a)

Link State Packets

A		B		C		D		E		F	
Seq.		Seq.		Seq.		Seq.		Seq.		Seq.	
Age		Age		Age		Age		Age		Age	
B	4	A	4	B	2	C	3	A	5	B	6
E	5	C	2	D	3	F	7	C	1	D	7
		F	6	E	1			F	8	E	8

(b)

- When to build LSPs
  - Periodically
  - Based on events

Distributing the Link State Packets

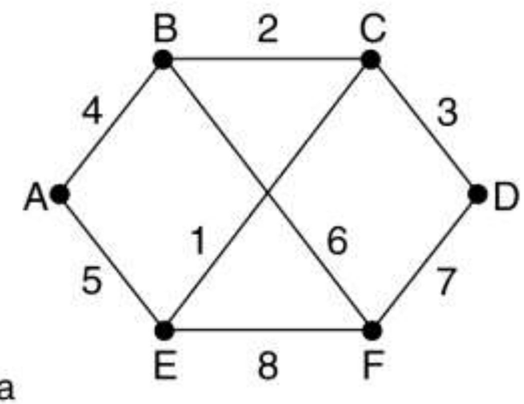
- Problem: Distributing LSPs with integrity and reliability
  - Inefficient distribution leads to different versions of topology → inconsistencies, loops, unreachable hosts, etc. etc.
- Distribution algorithm
  - Flooding
    - Flood is controlled by remembering the (source router, seq no.) pair
      - Duplicates and old seq. nos. are discarded

Distributing the Link State Packets

- Problems with Flooding
  1. If the sequence no. wraps around
  2. If a router crashes, it will lose track of its sequence nos.
  3. An error in seq. no during transmission
- Solution
  - Use 32-bit Sequence numbers
  - Include *Age* in each packet and decrement by 1 every second and every time when its forwarded to a router
  - To increase reliability, all LSPs are acknowledged and checksumed

Distributing the Link State Packets

The packet buffer for router B



			Send flags			ACK flags			
Source	Seq.	Age	A	C	F	A	C	F	Data
A	21	60	0	1	1	1	0	0	
F	21	60	1	1	0	0	0	1	
E	21	59	0	1	0	1	0	1	
C	20	60	1	0	1	0	1	0	
D	21	59	1	0	0	0	1	1	

Avoids the count to infinity problem since all routers get each other router's information

Compute the shortest path to every other router

- Each router constructs the entire subnet graph based on the LSP information
- No link is considered to have been correctly reported unless the two ends agree
    - If one node reports that it is connected to another, but the other node does not report that it is connected to the first, there is a problem, and the link is not included on the map
- Each router uses Dijkstra's algorithm to calculate shortest paths based on the current values in its database
- Because each router makes its calculation using the same information, better routing decisions are made

LS vs DV Routing

- Pros
  - more reliable,
  - more scalable,
  - more convergent,
  - less bandwidth-intensive
    - The link-state packets that are sent over the network are smaller than the packets used in distance-vector routing
      - DV requires a node's entire routing table to be transmitted, while in link-state routing only information about the node's immediate neighbors are transmitted

LS vs DV Routing

- Cons
  - More complex
  - More compute intensive
  - More memory-intensive
    - One LSP for every other node in the network

c. Border Gateway Protocol

- An Exterior Gateway Protocol

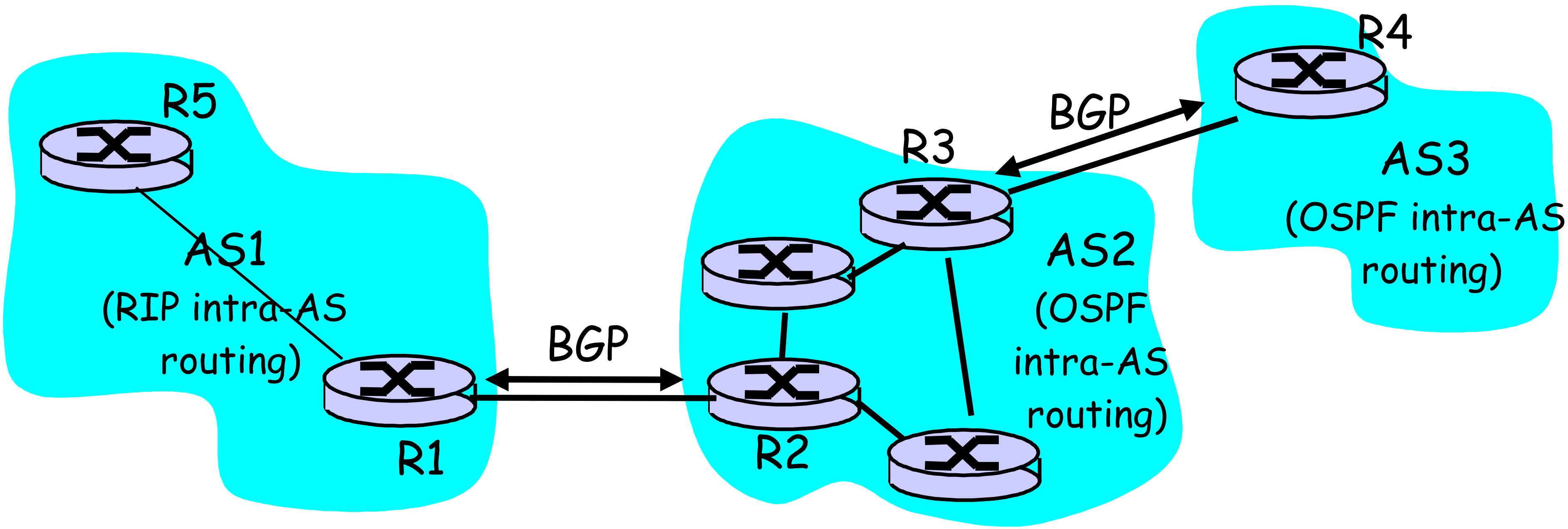
Intro to BGP

- Core routing protocol of the Internet
- Designates network reachability among autonomous systems (AS)
  - Uses Path Vector algorithm
- Current version of BGP is 4 (RFC 4271)

Intra and Inter-AS routing

- Performance vs Policy
  - Intra-AS (IGP-Interior Gat)
    - Move packets as efficiently as possible from source to destination within a single admin domain (Focus is on performance)
  - Inter-AS (EGP)
    - Admin wants control over how traffic is routed (Focus is on policy)
- Scalability
  - Hierarchical routing saves table size, reduced update traffic

Inter-AS routing in the Internet: BGP

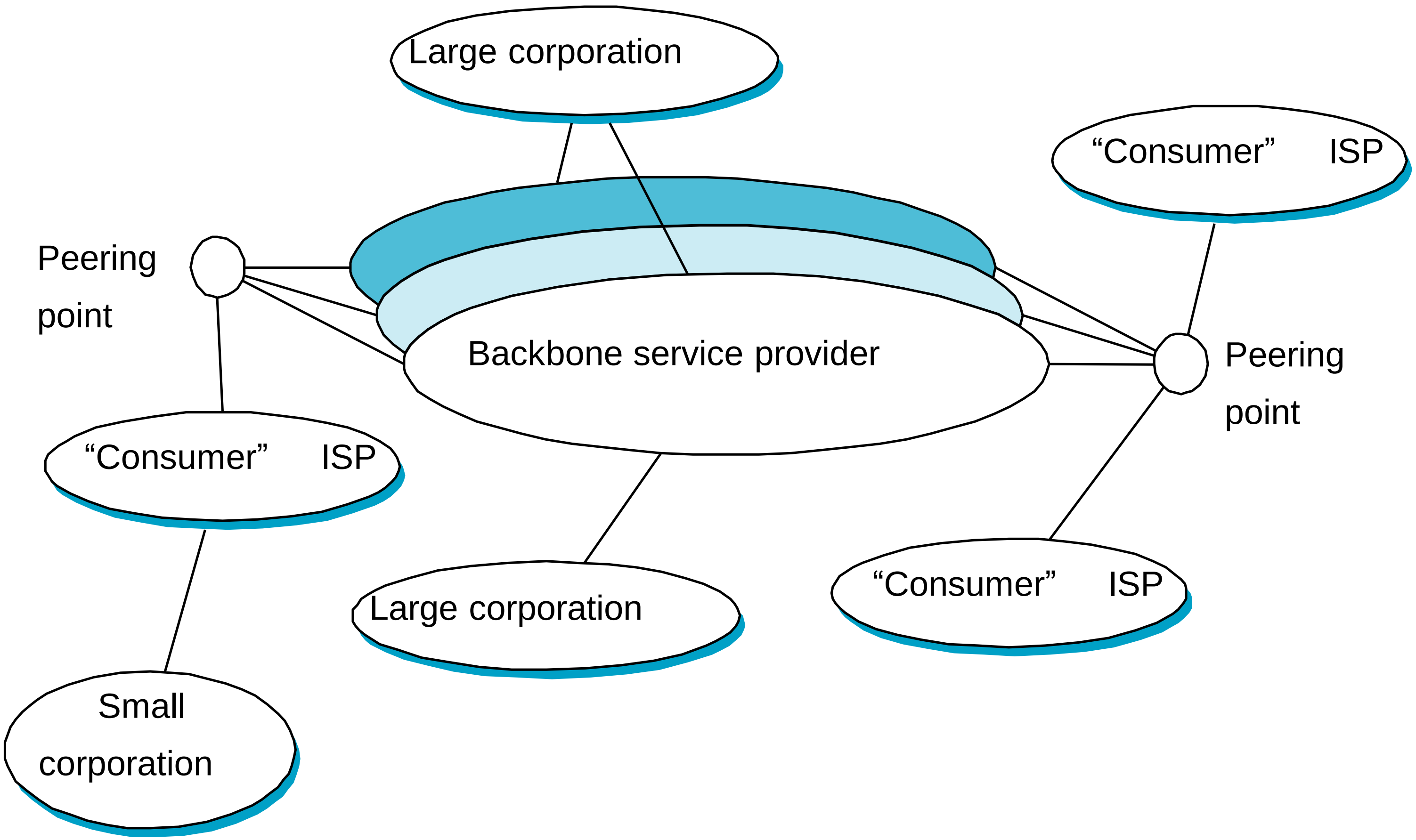


- Local Traffic
  - Traffic that originates at or terminates on nodes within an AS
- Transit Traffic
  - Traffic that passes through an AS

Internet through BGP's eyes

- The Internet consists of three types of Autonomous Systems (AS) interconnected with each other
  - Stub AS
    - An AS that has only a single connection to one other AS
    - Will carry only local traffic
  - Multihomed AS
    - An AS that has connections to more than one other AS but refuses to carry transit traffic
  - Transit AS
    - An AS that has connections to more than one other AS and can carry both local and transit traffic

Internet through BGP's eyes

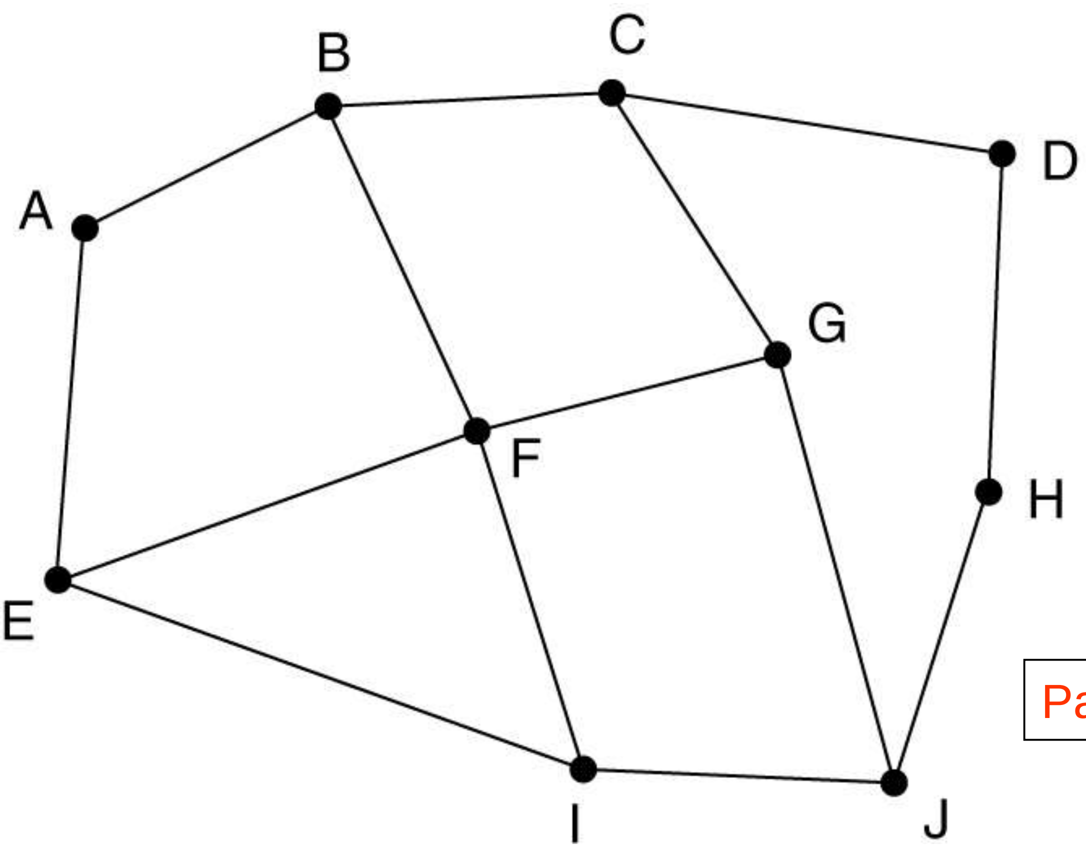


Goals of BGP

- To find *any* path to destination that is loop-free
  - i.e. we are more concerned with reachability than optimality
- Support complex routing policies involving political, security or economic issues

Path Vector Algorithm

- A variant of DV algorithm
  - Instead of maintaining and exchanging *costs* to each destination, BGP router keeps track of *path*



Information F receives
from its neighbors about D

From B: "I use BCD"
From G: "I use GCD"
From I: "I use IFGCD"
From E: "I use EFGCD"

Paths with loops

THANK YOU!

Questions & Comments