Network Layer

Routing Concept

Routing is forwarding of packets from one network to another network choosing the best path from the routing table.

Routing table consist of only the best routes for every destinations.

Types of Routing:

There are three ways a router learn a route to destination, these are:

Static	• It is configure by Administrator **manually** • Need for **destination network ID** • It is **secure and fast** • Used for **small organization** which have network of **10-15** routers
Dynamic	• Means **automatically** routing • Dynamic routes means that the router **learns** of paths of destinations by receiving **periodic updates** from other routers • Is **automatically** choose the best shortest path • Can be done by using **routing protocol**
Default	• Is configured for **unknown** destination • When there is **no entry** for the destination network in a routing table, the router will forward the packet to its **default router**. • It is **last preferred routing**

Two Key Network-Layer Functions

1. **Forwarding:** move a packet from router's input to appropriate router output.
2. **Routing:** determine route taken by packets from source to destination.

Routing Algorithm

Routing Algorithms

- The main function of the network layer is **routing packets from source to destination**.
- The **routing algorithm** is that part of the network layer software responsible for **deciding** which output line an incoming packet should be transmitted on.

- Routing algorithms can be grouped into **two** major classes: Static (**non-adaptive)** and dynamic (**adaptive)**.

Non adaptive algorithms (static routing)	Adaptive algorithms (dynamic routing)
• **Do not** base their routing decisions on measurements or estimates of the current traffic and topology. • It is called **static** algorithm.	• **Change** their routing decisions to reflect changes in the topology, and usually the traffic as well. • It is called **dynamic**.

Static Algorithm

1. Flooding routing.
2. Shortest path routing.

Flooding	• A simple local technique, where each router must make decisions based on local knowledge, **not the complete picture of the network**. • Is a simple algorithm to send a packet **along all paths** (Every incoming packet is sent out **on every outgoing** line **except** the one it arrived on). • Generates **infinite number of duplicate packets** unless some measures are taken to damp the process. • One such measure is to have a **hop counter** in the header of each packet, which is **decremented** at each hop, with the packet being discarded when the counter reaches zero
Shortest Path Routing	• Shortest path routing first developed by **Dijkstra** algorithm. • Find the shortest path from a **specified source to all other destinations** in the network. • In the general case, the labels on the lines could be computed **as a function** of the distance, bandwidth, average traffic, communication cost, measured delay, and other factors.

Q/ Find the shortest path from router A to router H?

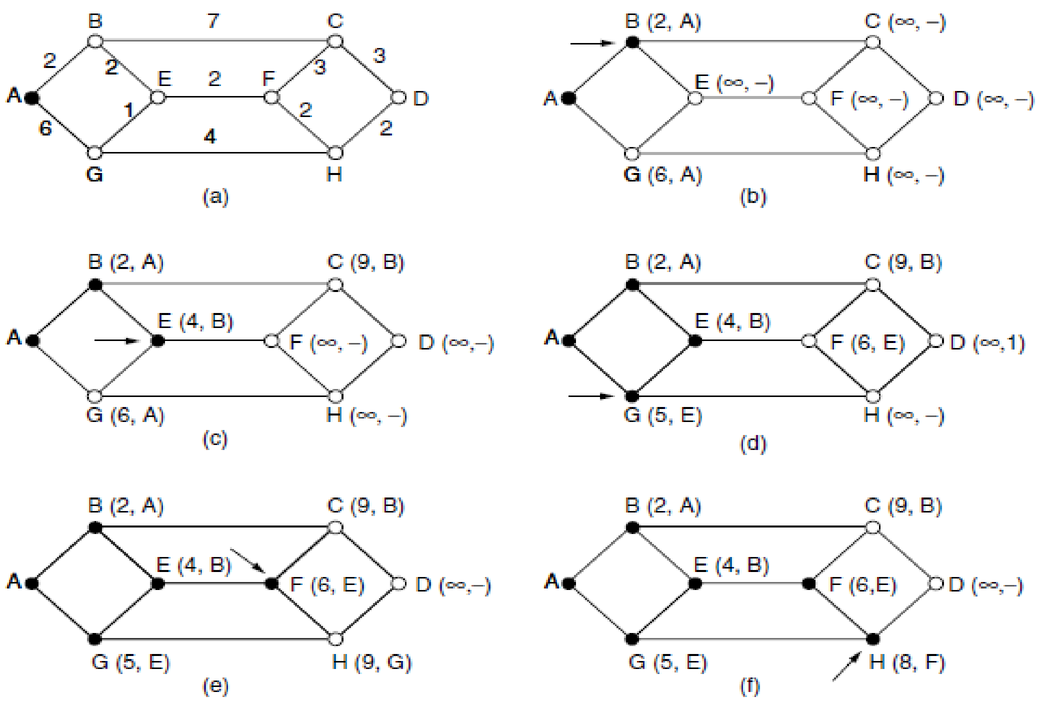


Figure The first six steps used in computing the shortest path from *A* to *D*. The arrows indicate the working node.

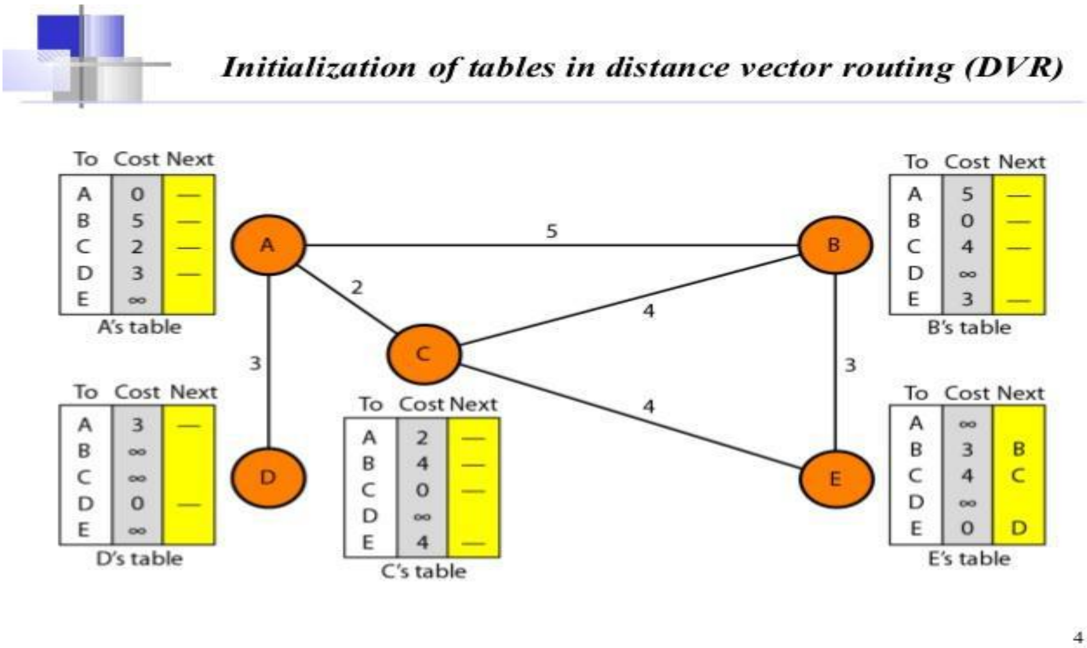
<u>Dynamic Routing Algorithm</u>

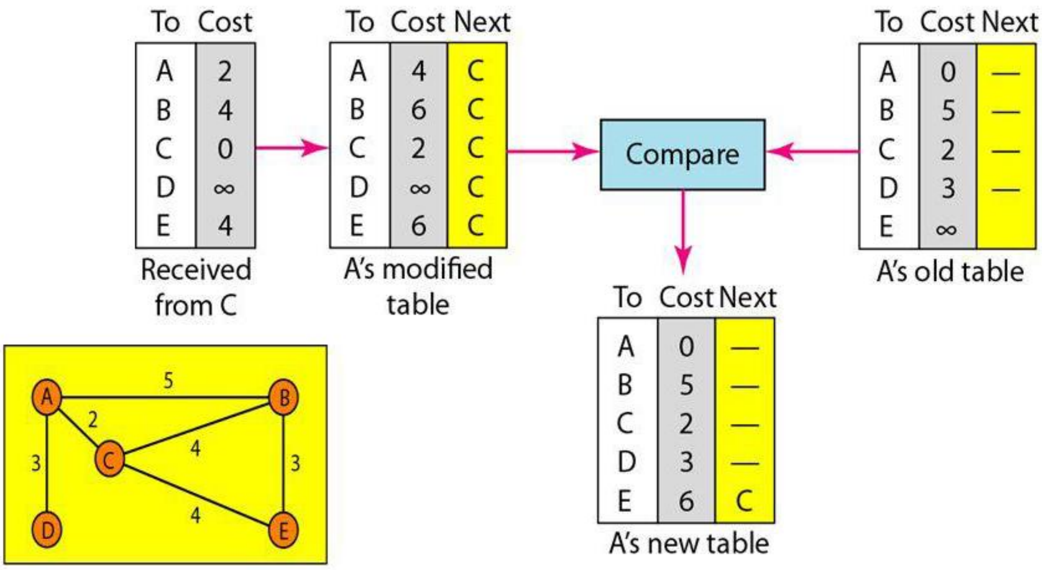
1. Distance Vector Routing.
2. Link state routing.

Distance Vector Routing

- Distance Vector routing is **intra-domain** protocols, **inside** Autonomous system, but not between Autonomous system.
- Distance-vector routing is based on the **least-cost** goal.
- Distance Vector developed by **Bellman-Ford** algorithm.
- Bellman **equation** is used to find the **least** cost (**shortest distance**) between a source and destination.
- A **distance vector routing algorithm** operates by having each router **maintain a table** (i.e., a **vector**) giving the best known distance to each destination.
- These tables are **updated by exchanging** information with the neighbor's router. Every router knows the **best link** to reach each destination.
- Distance Vector router **tells ONLY neighbors** about ALL routes
- **RIP** based on distance vector routing, each router **shares**, **at regular intervals**, its knowledge about entire AS with its neighbor.
- It is so **slow** and does not take **Bandwidth** into consideration when choose the root.

Q/Update Router A using Distance vector algorithm.





Distance vector algorithm

Bellman-Ford equation (dynamic programming)

let

$d_x(y)$:= cost of least-cost path from x to y

Then

$$d_x(y) = \min_v \{c(x,v) + d_v(y)\}$$

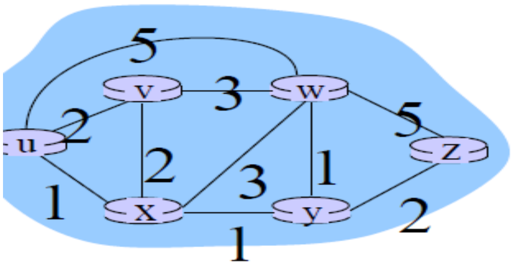
cost from neighbor v to destination y

cost to neighbor v

min taken over all neighbors v of x

- $D_x(y)$ = estimate of least cost from x to y
  x maintains distance vector $\mathbf{D}_x = [D_x(y): y \in N]$
- node x:
  - knows cost to each neighbor v: $c(x,v)$
  - maintains its neighbors' distance vectors. For each neighbor v, x maintains $\mathbf{D}_v = [D_v(y): y \in N]$

Bellman-Ford example



clearly, $d_v(z) = 5$, $d_x(z) = 3$, $d_w(z) = 3$

B-F equation says:

$$\begin{aligned} d_u(z) &= \min \{ c(u,v) + d_v(z), \\ &\qquad c(u,x) + d_x(z), \\ &\qquad c(u,w) + d_w(z) \} \\ &= \min \{2 + 5, \\ &\qquad 1 + 3, \\ &\qquad 5 + 3\} = 4 \end{aligned}$$

node achieving minimum is next
hop in shortest path, used in forwarding table

Link State Routing

The **primary problem in distance vector** that the algorithm often took **too long to converge** after the network topology changed (due to the count-to-infinity **problem**). Consequently, it was **replaced by** a new algorithm, now called **link state** routing.

The idea behind link state routing is simple and can be stated as five parts. Each router must:

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

Note that the Link-State router tells **ALL** other routers about **ONLY** its neighbors and links.

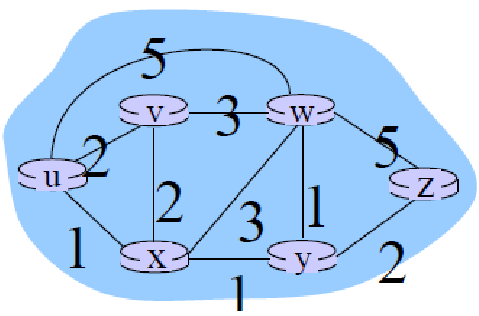
Compared to distance vector routing, link state routing requires **more memory and computation**. Also, the **computation time** grows faster

Nevertheless, in many practical situations, link state routing works **well because it does not suffer from slow convergence problems**.

Q/Consider the following network, define OSPF protocol & construct shortest path tree using dijkstra algorithms:

Dijkstra's algorithm: example

Step	N'	D(v),p(v)	D(w),p(w)	D(x),p(x)	D(y),p(y)	D(z),p(z)
0	u	2,u	5,u	1,u	∞	∞
1	ux	2,u	4,x		2,x	∞
2	uxy	2,u	3,y			4,y
3	uxyv		3,y			4,y
4	uxyvw					4,y
5	uxyvwz					



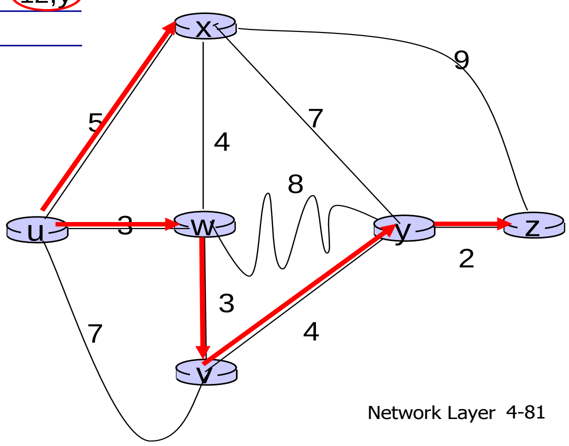
Q/Consider the following network, define OSPF protocol & construct shortest path tree using dijkstra algorithms:

Dijkstra's algorithm: example

Step	N'	D(v) p(v)	D(w) p(w)	D(x) p(x)	D(y) p(y)	D(z) p(z)
0	u	7,u	3,u	5,u	∞	∞
1	uw	6,w		5,u	11,w	∞
2	uwx	6,w			11,w	14,x
3	uwxv				10,v	14,x
4	uwxvy					12,y
5	uwxvyz					

notes:

- construct shortest path tree by tracing predecessor nodes
- ties can exist (can be broken arbitrarily)



Network Layer 4-81

Comparison of LS and DV algorithms

message complexity

- *LS:* with n nodes, E links, O(nE) msgs sent
- *DV:* exchange between neighbors only
  - convergence time varies

speed of convergence

- *LS:* $O(n^2)$ algorithm requires O(nE) msgs
  - may have oscillations
- *DV:* convergence time varies
  - may be routing loops
  - count-to-infinity problem

robustness: what happens if router malfunctions?

LS:

- node can advertise incorrect *link* cost
- each node computes only its *own* table

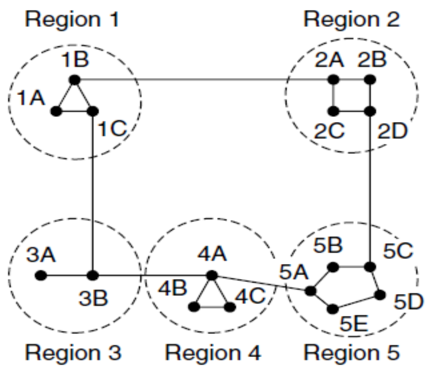
DV:

- DV node can advertise incorrect *path* cost
- each node's table used by others
  - error propagate thru network

Distance Vector	Link State
• Entire routing table is sent as an update	• Updates are incremental & entire routing table is not sent as update.
• Distance vector protocol send periodic update at every 30 or 90 second	• Updates are triggered not periodic.
• Updates are sent to directly connected neighbor only	• Update are sent to entire network & to just directly connected neighbor.
• Routers don't have end to end visibility of entire network.	• Routers have visibility of entire network of that area only.
• Suffer from count to infinity problem	• No routing loops
. **Slow convergence** **after network topology changed due to the count to infinty problem**	• Convergence is fast because of triggered updates.
• Examples: RIP, BGP	• Examples: OSPF, IS-IS
. **Easy to configure**	. **More difficult to configure**

Hierarchical Routing

- As networks **grow in size**, the **router routing tables grow proportionally**.
- Not only is router **memory** consumed by ever-increasing tables, but more **CPU time** is needed to scan them and **more bandwidth** is needed to send status reports about them.
- So router cannot have table about the entire network.
- **When hierarchical routing is used**, the routers are divided into what we will call **regions.**
- Each router knows all the details about how to route packets to destinations within its **own region** but knows nothing about the internal structure of other regions.



(a)

Full table for 1A

Dest.	Line	Hops
1A	–	–
1B	1B	1
1C	1C	1
2A	1B	2
2B	1B	3
2C	1B	3
2D	1B	4
3A	1C	3
3B	1C	2
4A	1C	3
4B	1C	4
4C	1C	4
5A	1C	4
5B	1C	5
5C	1B	5
5D	1C	6
5E	1C	5

(b)

Hierarchical table for 1A

Dest.	Line	Hops
1A	–	–
1B	1B	1
1C	1C	1
2	1B	2
3	1C	2
4	1C	3
5	1C	4

(c)

Broadcast Routing

- For some applications, hosts need to send messages to **many or all other hosts**. **Broadcast routing** is used for that purpose.
- The source should send the packet to **all** the necessary destinations. **One of the problems of this method** is that the source has to have the complete list of destinations.
- We have already seen a better broadcast routing technique: **flooding.**

Multicast Routing

- Sending a message to such a **group** is called **multicasting**, and the routing algorithm used is called **multicast routing**.
- All multicasting schemes require some way to **create and destroy groups and to identify which routers are members of group**.

Network Service Model

Q: What service model for "channel" transporting datagrams from sender to receiver?

individual datagrams	***flow of datagrams***
1. guaranteed delivery 2. guaranteed delivery with less than 40 msec delay	• in-order datagram delivery • Guaranteed minimum bandwidth to flow.

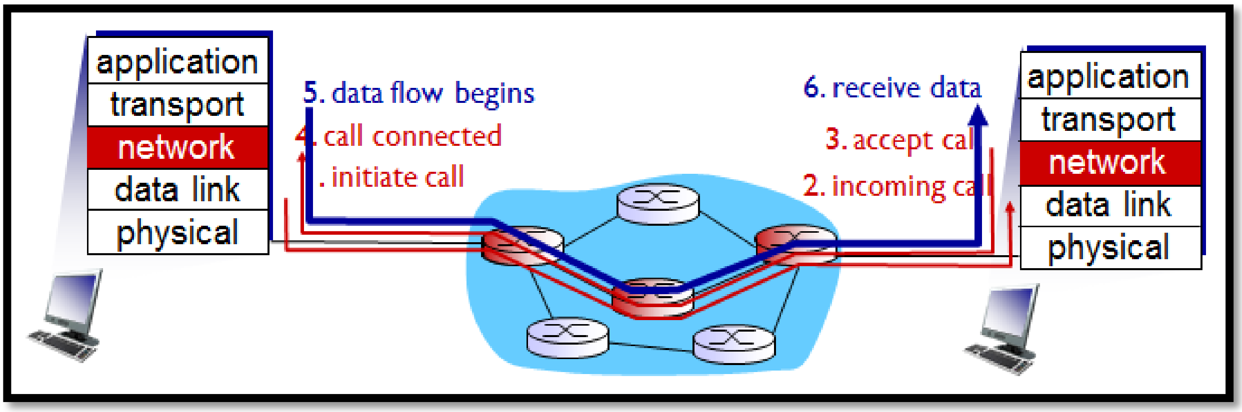
Virtual Circuits and Datagrams

From the network layer points of view, it has to make sure the packets received are in correct order. There are a lot of models existed to help address this problem, among them, two conceptual models namely

1. **Virtual circuit:** the network layer provides the transport layer with a perfect source-to-destination path" behaves much like "telephone circuit" and all packets delivered in order.
   - **Network provides network-layer connection oriented service.**
   - **call setup, teardown for each call *before* data can flow**
   - **each packet carries VC identifier (not destination host address)**
2. **Datagram dominate**: network provides network-layer **connectionless service.**

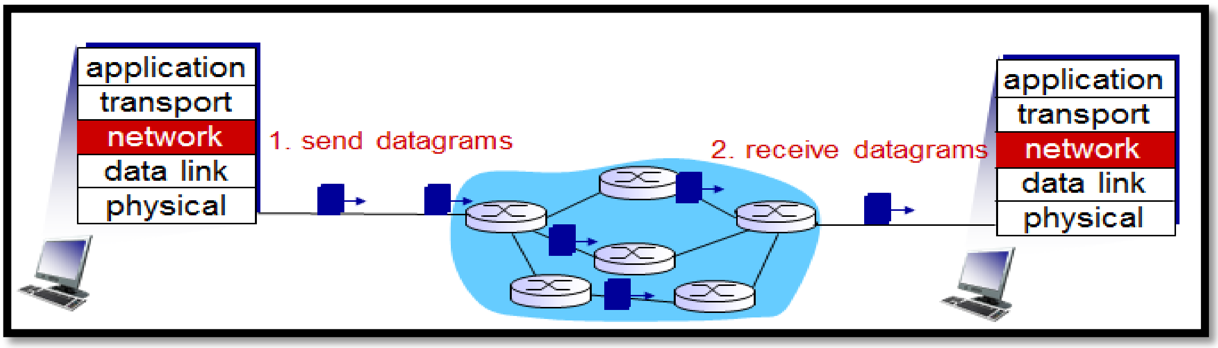
Virtual circuits: signaling protocols

- used to setup, maintain, teardown VC
- used in ATM, frame-relay, X.25
- not used in today's Internet



Datagram Networks

- **No call setup** at network layer
- Routers: **no state about end-to-end connections**
  - no network-level concept of "connection"
- Packets forwarded using destination **host address**



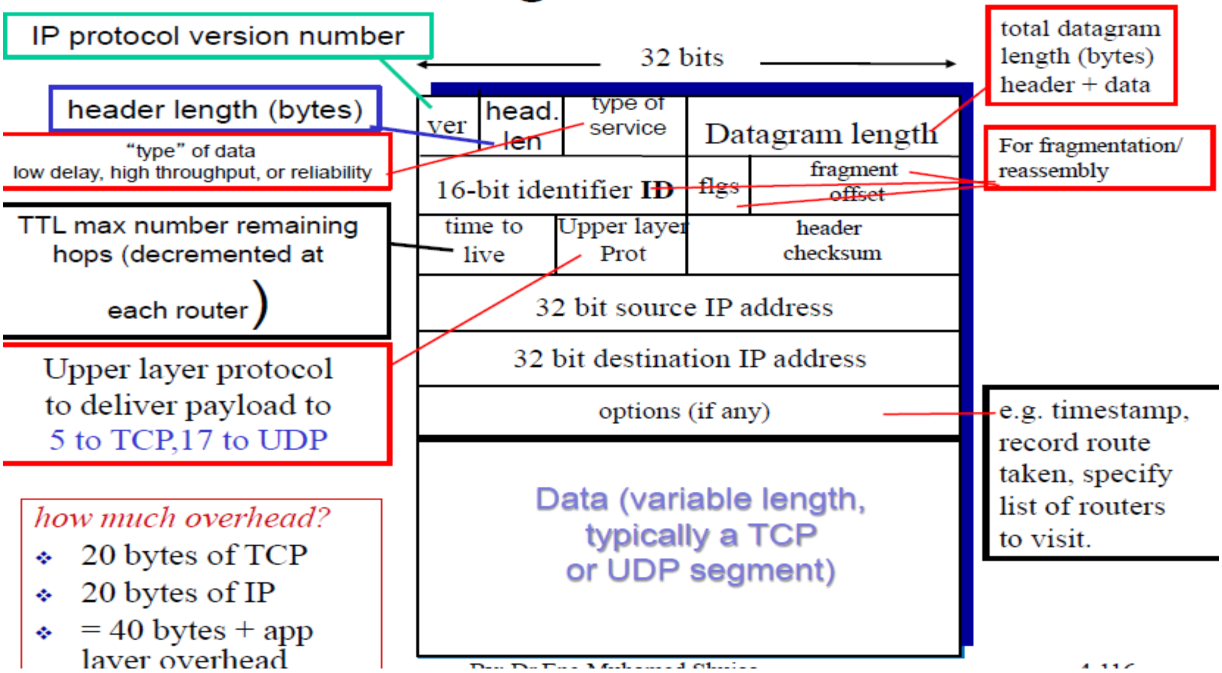
Q/ Datagram or VC network: why?

Internet (datagram)	*ATM (VC) virtual circuit*
1. Data exchange among computers ❖ "Elastic" service, no strict timing req. **2.** Many Link Types ❖ different characteristics ❖ uniform service difficult **3.** "smart" end systems (computers) ❖ can adapt, perform control, error recovery ❖ ***simple inside network, complexity at "edge"***	**1.** evolved from telephony **2.** human conversation: ❖ strict timing, reliability requirements ❖ need for guaranteed service **3.** "dumb" end systems ❖ telephones ❖ ***complexity inside network***

Differences between Virtual Circuit and Datagram Network

Virtual Circuits	Datagram Networks
1. It is connection-oriented simply meaning that there is a reservation of resources like buffers, CPU, bandwidth, etc. 2. Since data follows a particular dedicated path, packets reach in-order to the destination. 3. **Call setup, teardown for each call before data can flow.** 4. **Each packet carries VC identifier (not destination host address)** 5. Virtual Circuits are highly reliable means of transfer. 6. its costly to implement Virtual Circuits. Since each time a new connection has to be setup with reservation of resources and extra information handling at routers. 7. used in ATM, frame-relay, X.25. not used in today's Internet.	1. It is connectionless service. There is no need of reservation of resources as there is no dedicated path for a connection session. 2. The connectionless property makes data packets reach destination in any order. 3. **No call setup** at network layer. 4. Packets forwarded using destination **host address**. 5. Datagram networks are not reliable as Virtual Circuits. 6. But it is always easy and cost efficient to implement datagram networks as there is no extra headache of reserving resources and making a dedicated each time an application has to communicate. 7. used in today's Internet.

IP datagram format



Routing Protocols

- **A routing protocol** is a combination of **rules and a procedure that lets** routers in the internet inform each other of **changes.**
- **Routing metric**: a method by which routing algorithms determines that one route is better than another route, Metric may be (hop count, bandwidth, delay, and load).

Interior versus Exterior Routing Protocols

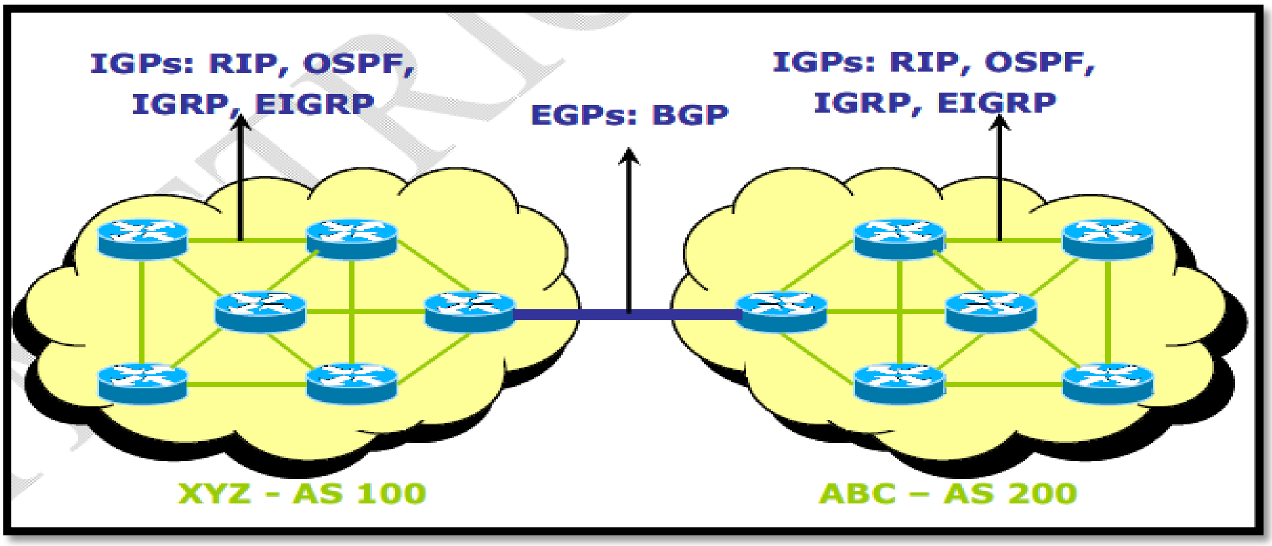
An autonomous system (AS), also known as a domain, is a **collection of routers** that are **under a common administration**, such as a company's internal network or an Internet service provider's (ISP's) network. Because the Internet is based on the AS concept, two types of routing protocols are required:

1. **Interior Gateway Protocols (IGPs):** handle routing **within an Autonomous System**. IGP's figure out how to get from place to place between the routers you own.

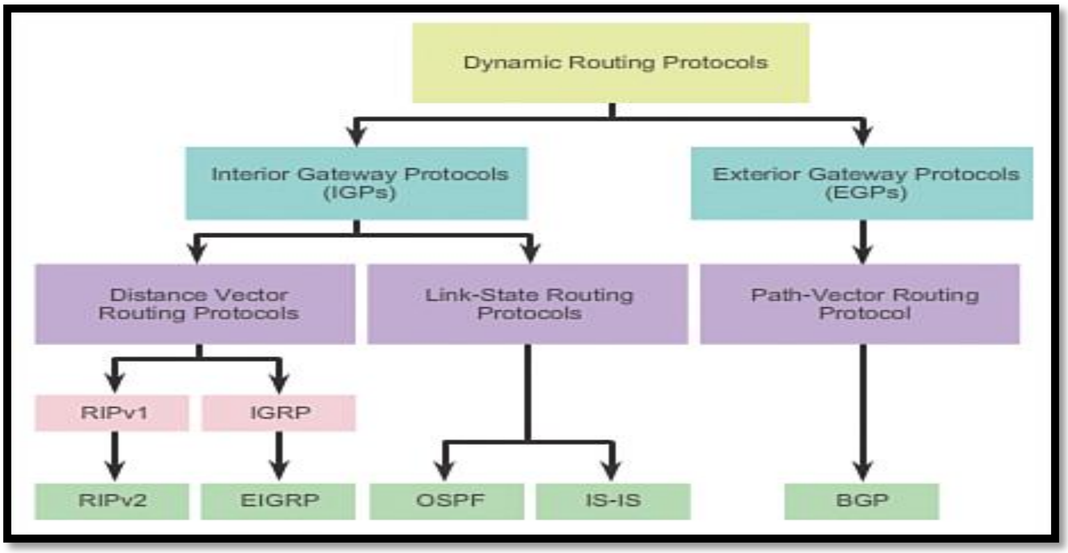
2. **Exterior Gateway Protocols (EGP):** handle routing **outside an Autonomous System** and get you from **your network**, through your Internet provider's network and onto any other network.

Interior gateway protocols (IGP)	**Exterior gateway protocols (EGP)**
Intra-AS (**inside an AS**) routing protocols.	Inter-AS (between autonomous systems) routing protocols.
Examples of IGPs include: • Routing Information Protocol (**RIP**). RIP version 2 (RIPv2), • Open Shortest Path First (**OSPF**), and	Border Gateway Protocol (BGP) is the only widely used EGP protocol on the Internet. BGP version 4 (BGP-4) is considered the acceptable version of BGP on the Internet.

Enhanced Interior Gateway Routing Protocol (**EIGRP**).	
use less-complicated metrics to ease configuration and speed up the decisions about best routing paths for faster convergence.	**Slower** to converge and more complex to configure.



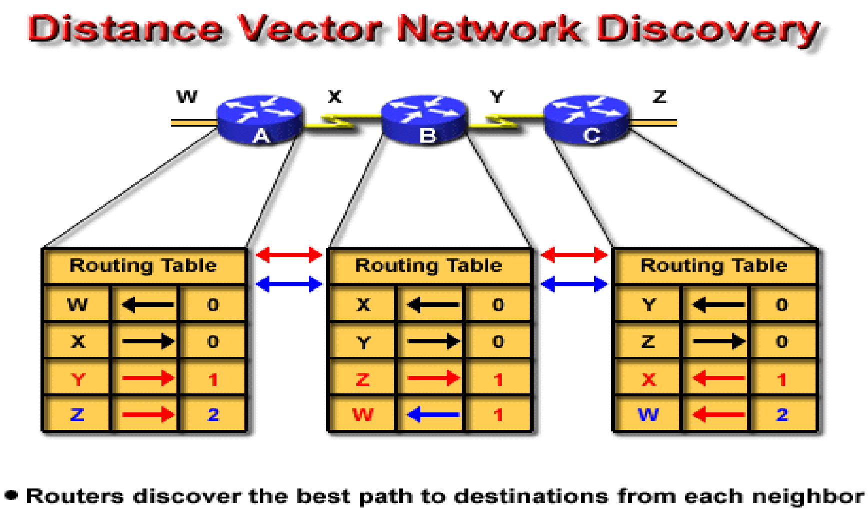
Popular Routing Protocols



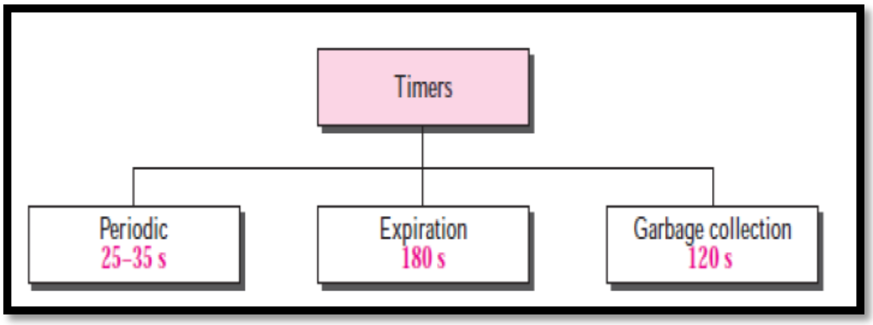
Routing Information Protocol (RIP)

- RIP is a routing protocol **for exchanging routing table information between routers**.
- It is a very simple protocol based on **distance vector routing**.
- Simple **intra-domain** protocol.
- Uses **hop count** as a path selection metric. (RIP **prevents routing loops** by implementing a **limit on the number of hops** allowed in a path from the source to a destination. The

maximum number of hops allowed for RIP is 15. This hop limit, however, also **limits the size of networks that RIP can support**, a hop count of 16 are considered an infinite distance, in other words the route is considered unreachable).



- There are **three types of timers** in RIP routing protocol, these are:



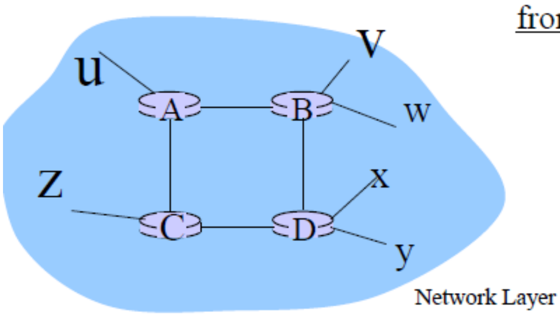
1. **Routing-update timer (periodic timer):** The periodic timer controls the advertising of regular update messages. Routers send **updates every 30** seconds.

2. **Route timeout (*Expiration Timer):*** Expiration Timer if there is a problem on an internet and no update is received within the allotted 180 s, the route is considered expired and the hop count of the route is set to 16, which means the destination is unreachable.(Because 15 hop max).

3. **Route-flush timer (Garbage Collection Timer)**: After the route **timeout expires**, the route-flush timer eventually expires, **deleting** the route from the table.

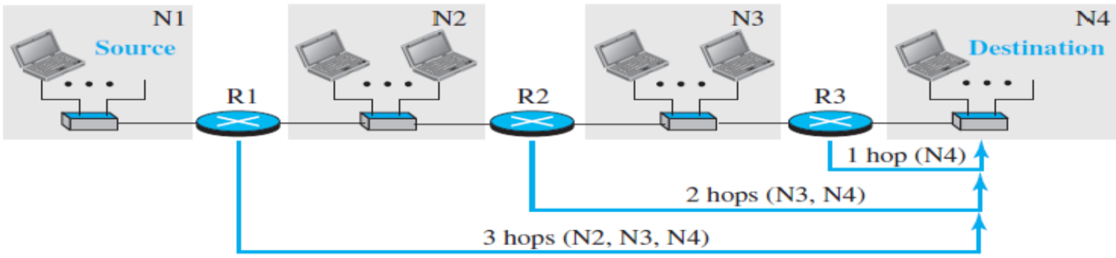
Drawbacks of RIP

- RIP has **slow convergence** and count to infinity problems
- The **hop count cannot exceed 15**, or routes will be dropped.



from router A to destination *subnets:*

subnet	hops
u	1
v	2
w	2
x	3
y	3
z	2

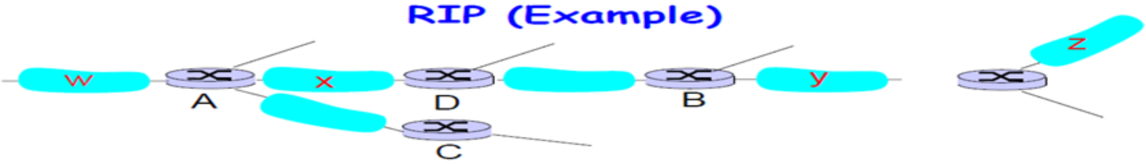


Forwarding table for R1

Destination network	Next router	Cost in hops
N1	——	1
N2	——	1
N3	R2	2
N4	R2	3

Note: After 15 hop the destination unreachable

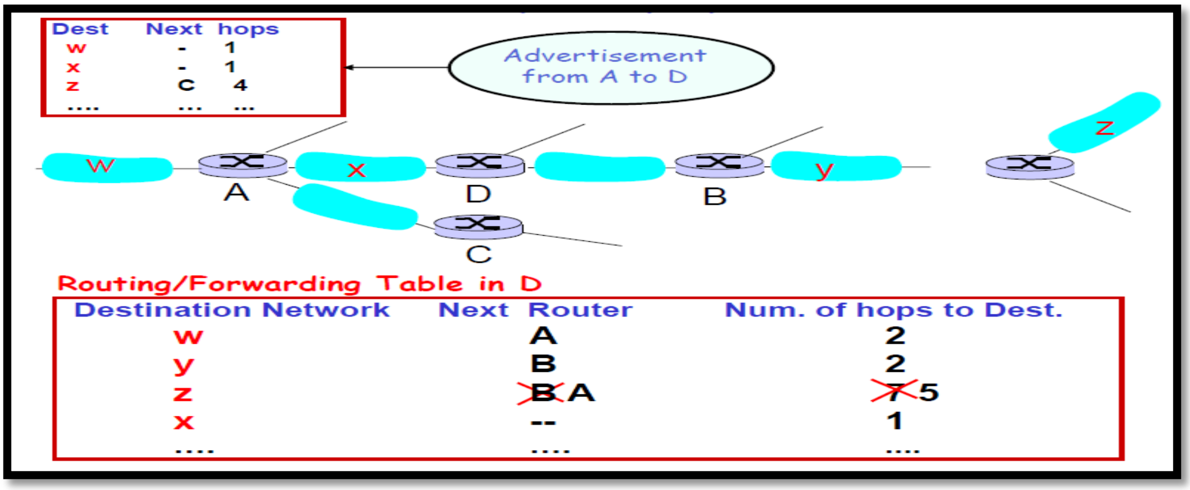
RIP (Example)



Routing/Forwarding Table in D

Destination Network	Next Router	Num. of hops to Dest.
w	A	2
y	B	2
z	B	7
x	--	1
....	....	....

Note that D can reach network z through Router B in 7 hops. See next slide for what would happen if some other router advertised a shorter route to network z.



Open Shortest Path First (OSPF)

The most important features of OSPF are as follows:

1. This protocol is **open**, which means that anyone can implement it without **paying license fees**.
2. OSPF uses a **link state routing algorithm** and falls into the group of interior routing protocols, operating within a single autonomous system (AS). OSPF is perhaps the most widely used interior gateway protocol (IGP) in **large enterprise networks**.
3. **Topology map** at each node.
4. Route computation using Dijkstra’s algorithm.
5. **Advertisements flooded to entire AS.**
6. The OSPF **metric calculation is based on cost**. The Cost is an indication of the **overhead required to send packets across a certain interface**.
7. **"security"**: all OSPF messages authenticated (to **prevent malicious intrusion**)
8. Hierarchical OSPF in large domains.

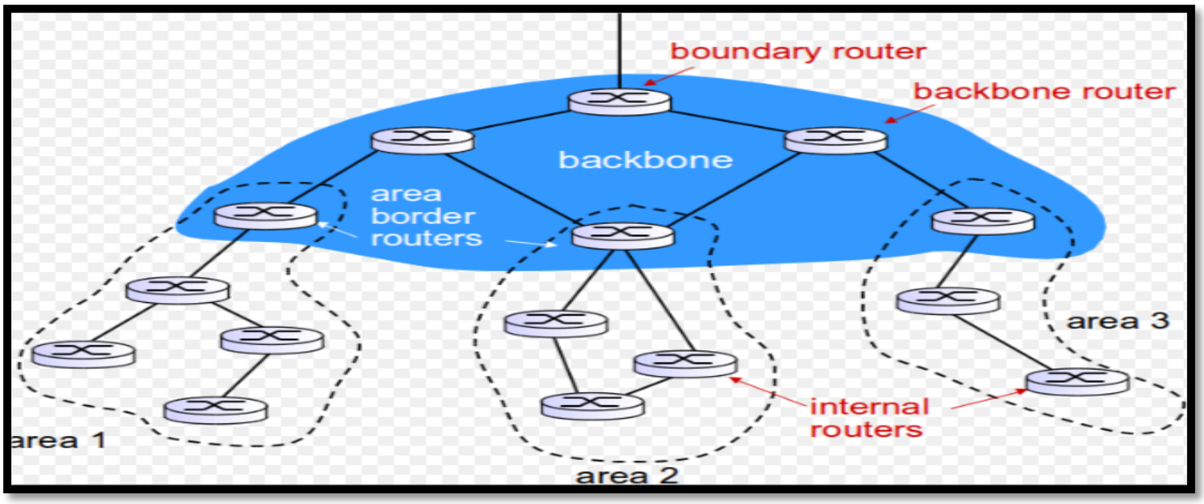
Hierarchical OSPF

To **minimize processing** and **memory requirements**, OSPF can **divide** the routing topology of an Autonomous System into **two-layer hierarchy** called **local area (Standard)** and **backbone area (Referred to as Area 0 or Transit Area).** The main advantages of this division are:

1. It **localizes** impact of a **topology change within an area**.
2. Detailed link state advertisement (LSA) **flooding only in area** (stops at the area boundary).
3. Each node (router) within an area has **detailed area topology**; and only knows direction (shortest path) to networks in other areas.

OSPF Routers May Function As Either:

- **Internal router**: work inside area.
- **Backbone router**: Run OSPF routing limited to backbone.
- **Area Border Router (ABR):** Joining local areas to backbone area. It "summarize" distances to networks in own area; advertise to other Area Border routers.
- **Autonomous System Boundary Router (ASBR):** connect to other AS's.



<u>OSPF is superior to RIP in all aspects, including the following:</u>

- It converges much faster.
- It supports hierarchical structures.
- It has improved metric calculation for best path selection.
- It does not have hop-count limitations
- At its inception, OSPF supported the largest networks.
- Compare to RIP, OSPF has no limitation due to hops (RIP has a limit of 15 hops so any network with more than 15 hops cannot be achieved by RIP.
-

<u>OSPF Messages</u>

Message type	Description
Hello	Used to discover who the neighbors are
Link state update	Provides the sender's costs to its neighbors
Link state ack	Acknowledges link state update
Database description	Announces which updates the sender has
Link state request	Requests information from the partner

EIGRP

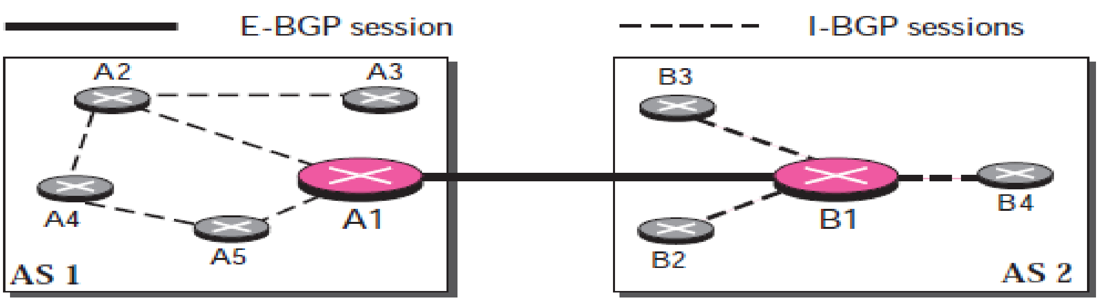
- EIGRP is a protocol for routing IPv4 and IPv6.
- EIGRP, however, is a hybrid routing protocol—**it is a distance vector protocol with additional link-state protocol features.**
- Uses triggered updates (EIGRP has **no periodic** updates).
- Provide **fast convergence to minimize network traffic.**
- Uses the **minimum bandwidth** on the path of the destination network, and calculate a route from the total delay metrics.

Border Gateway Protocol (BGP)

- **Inter-domain routing protocol** for routing between autonomous systems (holds the Internet together)
- BGP is **neither a link state, nor a distance vector protocol**. Routing messages in BGP contain complete routes.
- Network administrators can specify routing policies (**BGP supports flexibility --** paths could be chosen by a provider based on a policy.
- Network administrators can specify routing policies.
- BGP's goal is to find **any path** (not an optimal one).

Internal and external BGP sessions

iBGP	eBGP
• used to connect different routers have **same** AS(same company) • Propagate reach ability information **to all AS-internal routers.**	• used to connect different routers have **different** AS(different company) • Obtain subnet reach ability information **from neighboring ASs**.



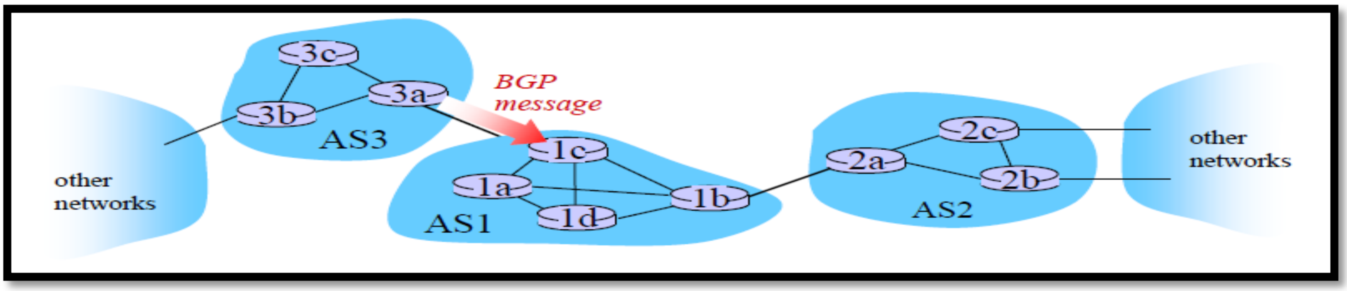
BGP Basics

1. BGP session: two BGP routers ("peers") exchange BGP messages:

- **Advertising paths** to different destination network prefixes ("path vector" protocol).
- Exchanged over semi-permanent **TCP connections**.

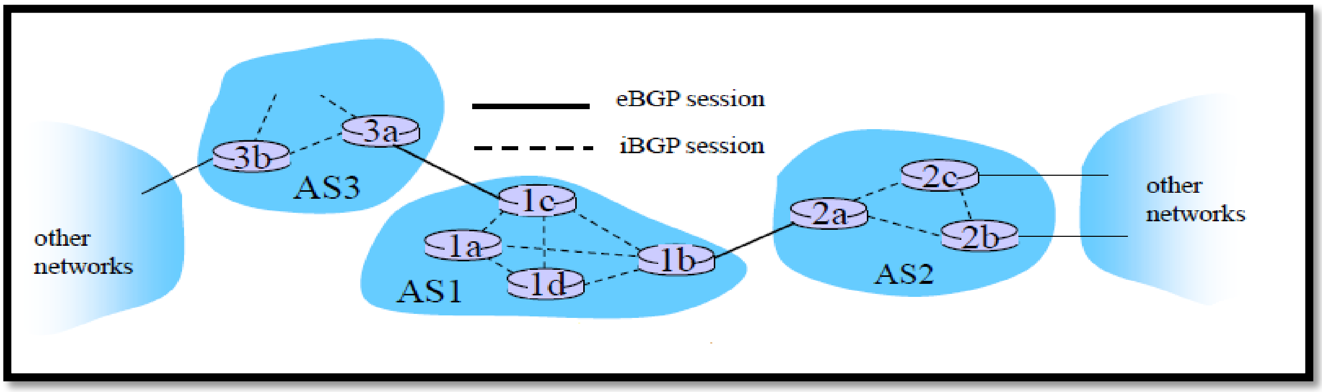
Example: when AS3 advertises a **prefix (How do routers advertise sets of IP addresses)** to AS1:

- AS3 *promises* it will forward datagrams towards that prefix **(Sets of IP addresses)**.
- AS3 can aggregate prefixes in its advertisement.



2. Distributing Path (reachability) Information

- Using **eBGP** session between **3a** and **1c**, AS3 sends prefix reachability info to AS1.
- **1c** can then use **iBGP** do **distribute new** prefix info to all routers in AS1
- **1b** can then **re-advertise** new reachability info to AS2 over **1b-to-2a** eBGP session
- When router learns of new prefix, it creates **entry for prefix** in its forwarding table.



BGP Route Selection

Router may learn about more than 1 route to destination AS, **selects route based on**:

1. Local preference value attribute: **policy decision**
2. **Shortest AS-PATH**
3. **Closest NEXT-HOP router**.
4. **Additional criteria**

BGP Messages

BGP messages exchanged between peers over TCP connection, **BGP has four types of messages**

- **OPEN:** Establish a connection with a BGP peer
- **UPDATE --** advertise or withdraw routes to a destination
- **KEEPALIVE:** Inform a peer that the sender is still alive but has no information to send.
- **NOTIFICATION:** Notify that errors are detected, also used to close connection.