ROUTING PROTOCOL

Routing Protocols are used to find out the lowest or optimum cost path from source to destination. Hence Routing Protocols are used to create routing table at Host & Router side.

Different routing protocols are there based on the type of routing. They are-

(a) Unicast Routing
(b) Multicast Routing.

Unicast Routing Protocol-

To evaluate optimum cost path there are several approach. These are,

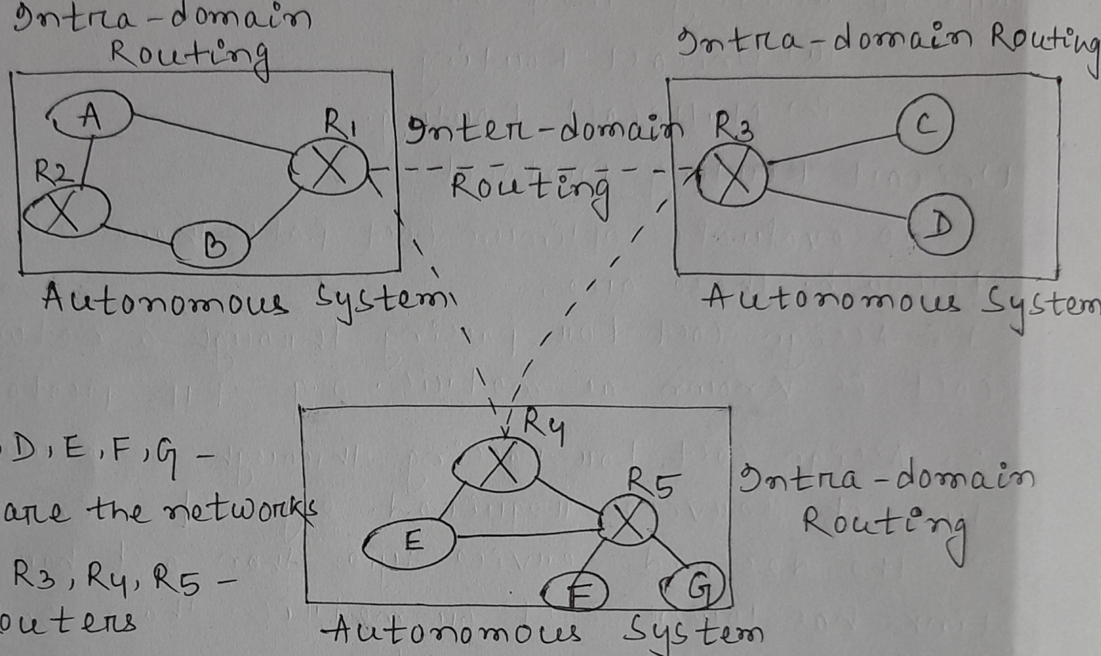
1. Assign cost for passing through a network. This cost is known as metric. e.g- RIP protocol.
2. Assign cost for passing through a network based on type of services required. e.g- If minimum delay is the desired service type then a fiber optic link has lower cost or metric than satellite link. e.g- OSPF protocol.
3. Assign cost based on policy set by network administrator. e.g- BGP protocol.

Routing
→ Intradomain Routing
→ Interdomain Routing

Intradomain Routing	Interdomain Routing
→ An Internet is divided into autonomous system. An autonomous system is group of network & routers under the authority of a single administration. Routing inside an autonomous system is known as intra-domain routing	→ Here is also same concept of autonomous system. Routing between autonomous system is known as inter-domain routing. → Different autonomous system can use different inter-domain routing. → e.g- BGP Routing Protocol.

→ Different autonomous system choose different routing protocols of intra-domain inside an autonomous system.

→ e.g - RIP, OSPF.



A, B, C, D, E, F, G - These are the networks

R1, R2, R3, R4, R5 - Routers

Routing Protocols

- Intra-domain Routing
  - Distance Vector Routing (DVR) → RIP
  - Link State Routing (LSR) → OSPF
- Inter-domain Routing
  - → Path vector Routing (PVR) → BGP

RIP - Routing Information Protocol

OSPF - Open Shortest path First

BGP - Border Gateway Protocol.

1. Distance Vector Routing (DVR) -

→ Each node maintain a vector table or routing table of minimum distance to every node.

→ DVR calculate & choose minimum or least cost route or path between any two nodes.

→ Different phases of DVR are,

(a) Initialization phase.
(b) Sharing phase.
(c) Updating phase.

(a) Initialization phase -

→ During initialization phase, each node knows only the distance between itself and its neighbour node, those that are directly connected to it.

→ In this phase initial cost is calculated by sending a message to the immediate neighbour node.

→ The distance or cost for an entry that is not neighbour is marked as ∞.

(b) Sharing phase -

→ During this phase routing informations are shared between neighbour nodes periodically.

→ Each node send its entire routing table to the neighbour & let the neighbour decide which part is to use & which part is to be discarded.

→ When the neighbour receives routing table, third column is to be replaced with sender's name. If any of the rows can be used, the next node is the sender of routing table. Therefore, a node can send only first two column of its routing table to any neighbour.

(c) Updating phase -

→ During this phase updation of routing table is performed.

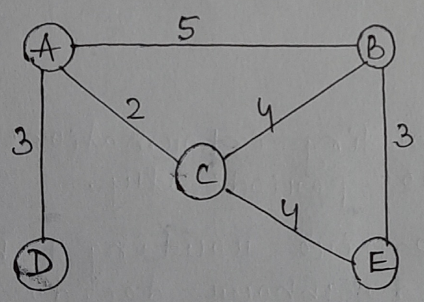
→ Receiving node add the cost between itself and the sending node to each value in the second column of receiving routing table.

→ Receiving node add name of the sending node to each rows as third column, if the receiving node uses information from any row.

→ Receiving node compare each row of its ~~old~~ old routing table with the corresponding row of modified routing table.

- If next node entry is different, then the receiving node chooses row with smaller value of cost. But if there is a tie then keep the older one.
- If next node entry is same, then the receiving node chooses the new one to update the table entry.

e.g- of DVR



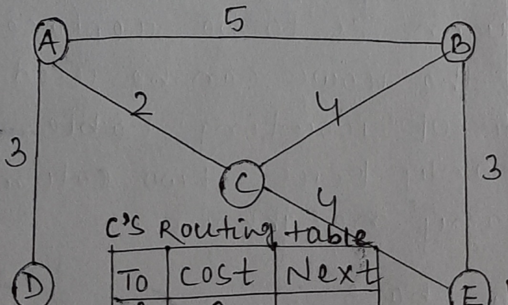
Initialization Phase

A's Routing Table

To	Cost	Next
A	0	—
B	5	—
C	2	—
D	3	—
E	∞	—

B's Routing Table

To	Cost	Next
A	5	—
B	0	—
C	4	—
D	∞	—
E	3	—



D's Routing Table

To	Cost	Next
A	3	—
B	∞	—
C	∞	—
D	0	—
E	∞	—

C's Routing table

To	cost	Next
A	2	—
B	4	—
C	0	—
D	∞	—
E	4	—

E's Routing Table

To	cost	Next
A	∞	—
B	3	—
C	4	—
D	∞	—
E	0	—

Sharing Phase -

Node C share its routing table to node A. Now, node A has two routing table - its own routing table & routing table of C.

Updation phase -

Node A create a new routing table from its old routing table & modified routing table by making comparison.

To	Cost
A	2
B	4
C	0
D	∞
E	4

Routing Table of node 'C'

To	Cost	Next
A	4	C
B	6	C
C	2	C
D	∞	C
E	6	C

A's Modified Routing Table

COMPARE

To	Cost	Next
A	0	—
B	5	—
C	2	—
D	3	—
E	∞	—

A's Old Routing Table

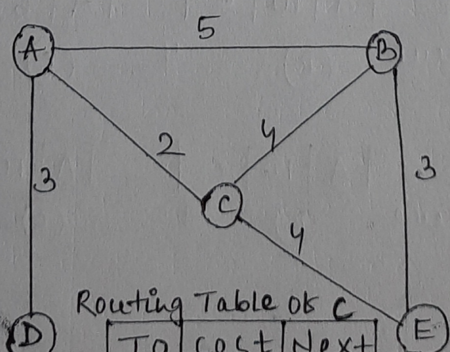
To	Cost	Next
A	0	—
B	5	—
C	2	—
D	3	—
E	6	C

A's New Routing Table

By doing above procedure the routing table of each node by using DVR is,

To	Cost	Next
A	0	—
B	5	—
C	2	—
D	3	—
E	6	C

Routing Table of A



To	Cost	Next
A	5	—
B	0	—
C	4	—
D	8	A
E	3	—

Routing Table of B

To	Cost	Next
A	3	—
B	8	A
C	5	A
D	0	—
E	9	A

Routing Table of D

Routing Table of C

To	Cost	Next
A	2	—
B	4	—
C	0	—
D	5	A
E	4	—

Routing Table of E

To	Cost	Next
A	6	C
B	3	—
C	4	—
D	9	C
E	0	—

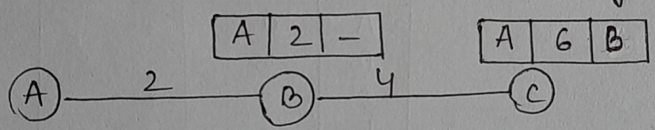
→ A node send its routing table to its neighbour when there is a change. Updating routing table is done by either one of two method. These are → Periodic update
→ Triggered update

Periodic Update	Triggered Update
A node send its routing table periodically after every 30 sec. This is known as periodic update	A node send its 1st & 2nd column of routing table to its neighbour anytime there is a change in its routing table. This is known as triggered update.

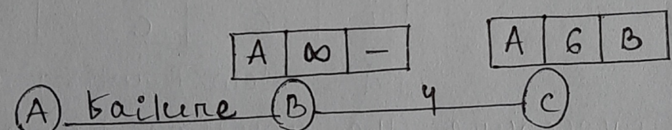
→ Problems associated with DVR is,

1. Two-node instability problem.
2. Three-node instability problem.

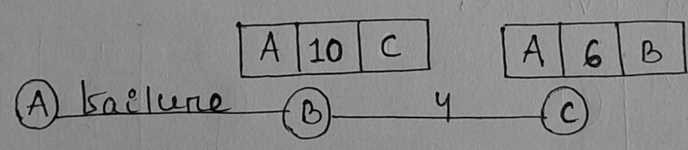
1. Two-node instability problem – Two node instability makes the network unstable. e.g-



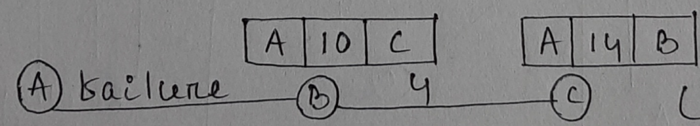
(Before failure)



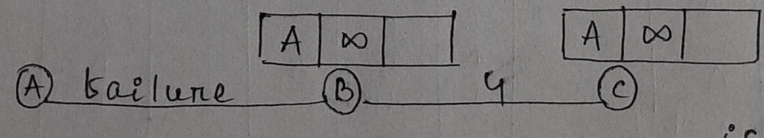
(After failure)



(After A receives update from B)



(After B receives updates from A)



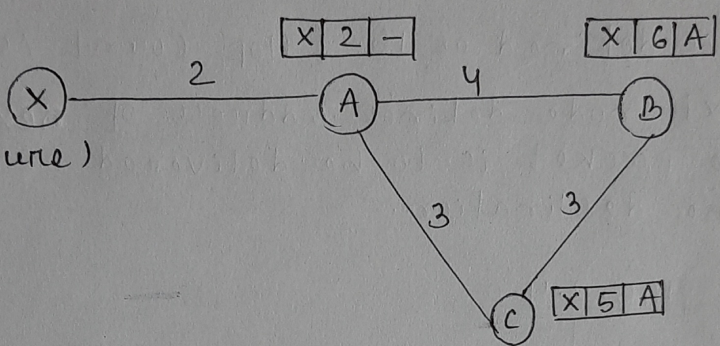
All routing table cost = ∞ means network is unstable.

Remedy to two node instability problem -

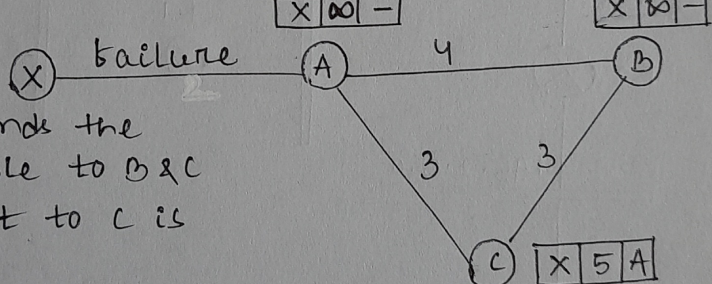
1. Defining ∞ - Redefine ∞ to a smaller value but it is not applicable to large system.

2. Split Horizon & Poison reverse - Replace the distance with ∞ is a warning that "donot use this value & what I know about this route, comes from you."

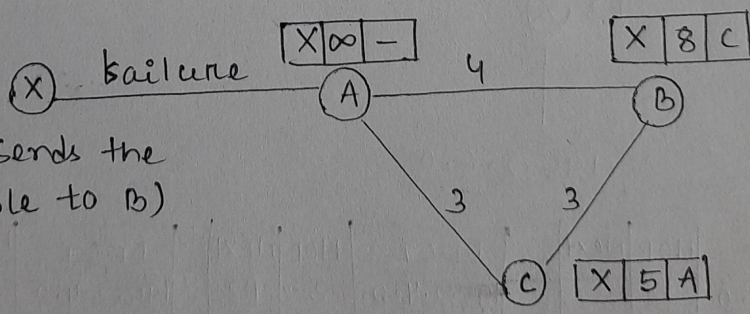
2. Three-node instability problem - If instability arise between three nodes then the network become unstable e.g -



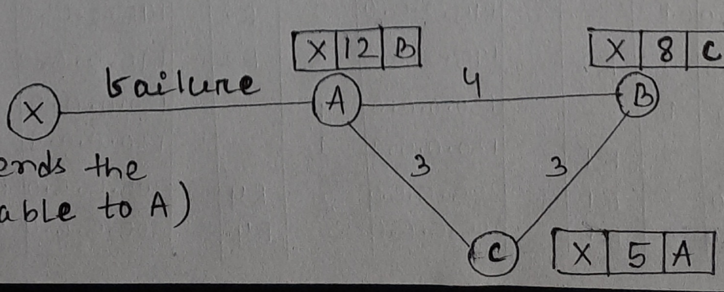
(Before failure)



(After A sends the routing table to B & C but packet to C is lost)



(After C sends the routing table to B)



(After B sends the routing table to A)

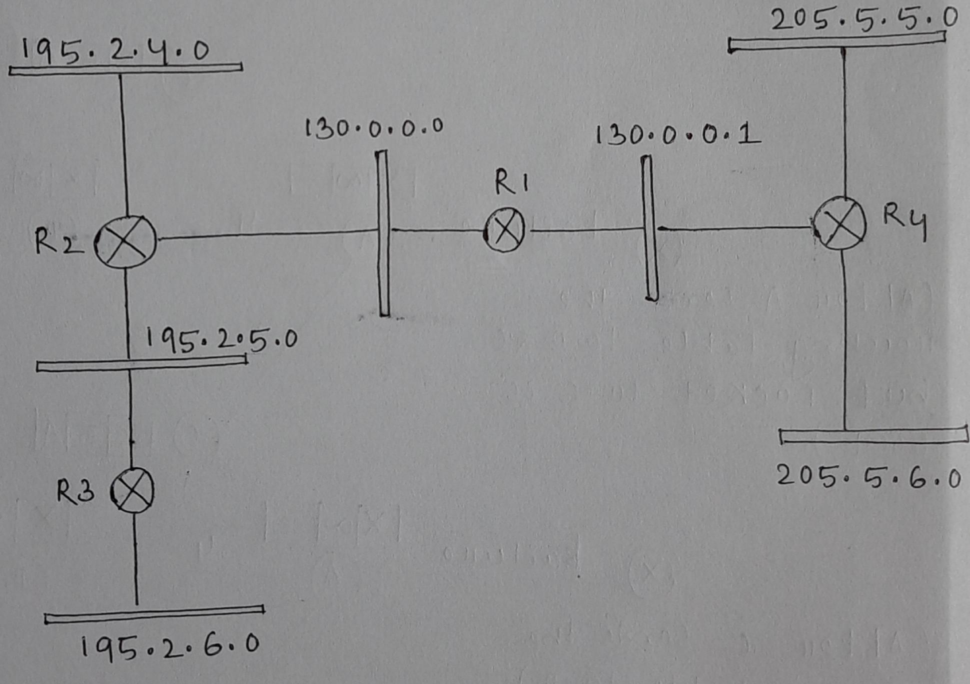
Routing Information Protocol (RIP) -

It is an intra-domain routing protocol based on DVR. RIP is implemented as follows -

Implementation -

1. In RIP we deal with routers & networks of an autonomous system.
2. Destination or 1st column in routing table is network.
3. Metric in RIP is called Hop count which is evaluated by no. of links or networks to reach the destination.
4. ∞ is defined as 16. (Hop count is limited to 15)
5. Next node defines address of next router to which the packet is to be delivered or forwarded to reach the destination.

e.g -



DEST	HOP	Next
195.2.4.0	1	—
195.2.5.0	1	—
195.2.6.0	2	R3
130.0.0.0	1	—
130.0.0.1	2	R1
205.5.5.0	3	R1
205.5.6.0	3	R1

DEST	HOP	NEXT
195.2.4.0	2	R2
195.2.5.0	2	R2
195.2.6.0	3	R2
130.0.0.0	1	—
130.0.0.1	1	—
205.5.5.0	2	R4
205.5.6.0	2	R4

DEST	HOP	NEXT
195.2.4.0	2	R2
195.2.5.0	1	—
195.2.6.0	1	—
130.0.0.0	2	R2
130.0.0.1	3	R2
205.5.5.0	4	R2
205.5.6.0	4	R2

DEST	Hop	NEXT
195.2.4.0	3	R1
195.2.5.0	3	R1
195.2.6.0	4	R1
130.0.0.0	2	R1
130.0.0.1	1	—
205.5.5.0	1	—
205.5.6.0	1	—

Link state Routing - (LSR)

→ In LSR each node use the same topology to create a routing table. But the routing table of each node is unique because calculation are based on different interpretation of topology. i:e: to reach a station different nodes takes different routes based on cost.

→ Topology must be dynamic, representing latest state of each node & each link. If there is changes in any point in the network, the topology must be updated for each node.

→ Concept of LSR is that although the global knowledge about the topology is not clear but each node has partial knowledge about topology.

→ The whole topology can be compiled from partial knowledge of each node. In LSR if each node has the entire topology information i:e: list of nodes, list of links, how nodes are connected, cost & condition of link (up or down) then nodes can use dijkstra's algorithm.

How to build routing table in LSR-

1. Creation of states of link by each node called as link state packet (LSP)
2. Dissemination of link state packet (LSP) to every other nodes or routers called flooding.
3. Formation of shortest path tree
4. Calculation of cost & formation of routing table based on shortest path tree.

Creation of LSP

A Link state packet or LSP carry information about the following-

(a) Node identity
(b) List of links
(c) Sequence no.
(d) Age.

→ Node identity & List of Links are required to make a topology.
→ Sequence no. distinguishes new LSPs from old one.
→ Age prevents old LSPs from remaining in domain for a long time.

LSP are generated in 2 cases -

1. When there is a change in topology of the domain. This is done to update topology information in every node.

2. LSPs are created on a periodic basis. This is done to make sure that old information are removed from domain.

Flooding of LSPs

Flooding process is based on the following -

1. LSP creating node send a copy of that LSPs to each interface.

2. A receiving node compare receiving LSPs with its old one. Then receiving node check sequence no. of the newly arrived LSP. If sequence no. of the newly arrived LSP is older than that of old one it has, then receiving node discard old LSPs & always keep new LSPs.

3. Now, the receiving node send out a copy of the new LSP through each interface, except the one from which the LSP packet has arrived.

4. Flooding stops when there is a node which has only one interface.

Formation of shortest path tree by using Dijkstra's algorithm

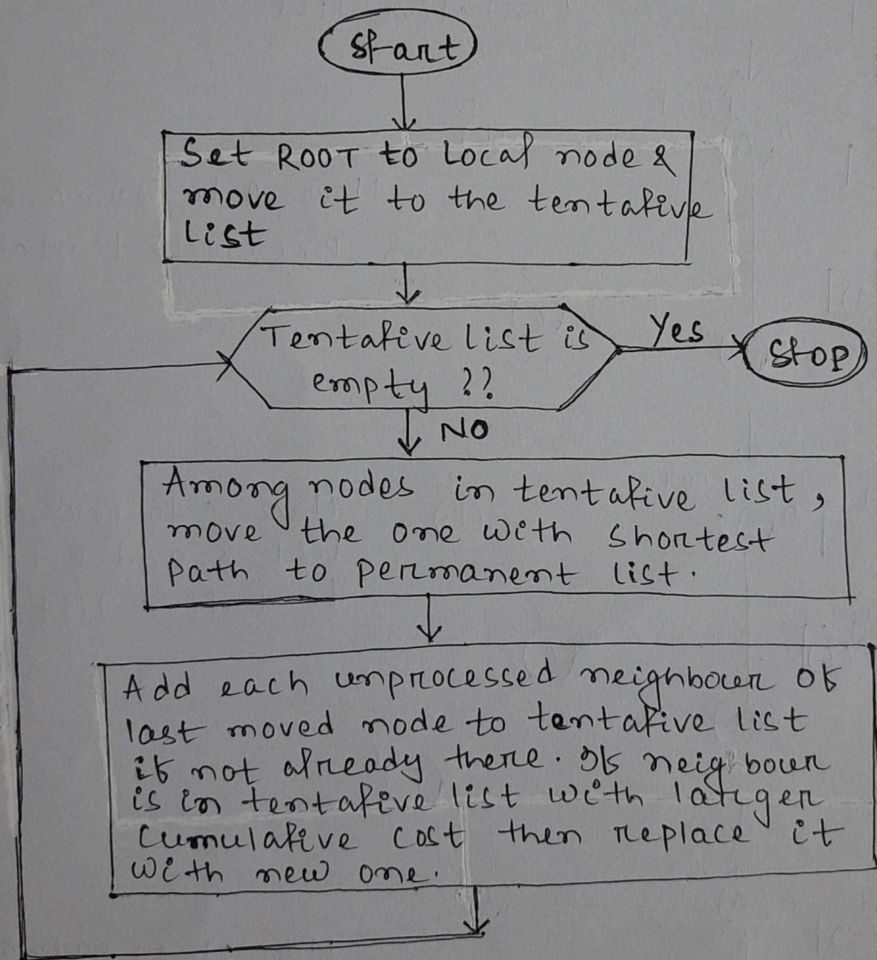
Dijkstra algorithm create a shortest path tree from a graph. A tree is a graph of nodes & links in which one of the node is called as Root. All the other nodes can be reached from the root through only one single root.

→ A Shortest path tree is a tree in which the path between the root & every other node is shortest.

→ Dijkstra algorithm divides the nodes into two types sets. They are - Permanent
Tentative

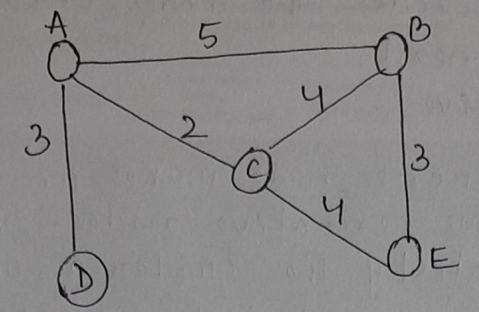
→ Neighbours of a current node makes an entry tentative list & becomes tentative nodes. If these tentative nodes satisfy the criteria of algorithm then they moved to permanent list & becomes a permanent node.

→ Flowchart of Dijkstra algorithm -



<u>Creation of Routing Table from Shortest path tree</u> -

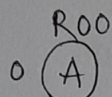
Routing table of LSR gives cost of reaching each node from ROOT. Routing table of LSR consist of 3 column - Destination (Nodes), Cost (metric), Next - Hop (Router next to destination node) e.g.– Create shortest path tree for the following network & form the routing table by using LSR.



Formation of shortest path tree –

Permanent: Empty

Temporary: A(0)

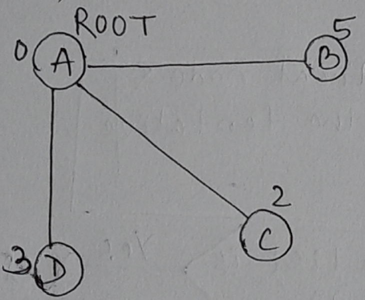


Set A as ROOT node & move it to tentative list.

↓

Permanent: A(0)

Temporary: B(5) C(2) D(3)

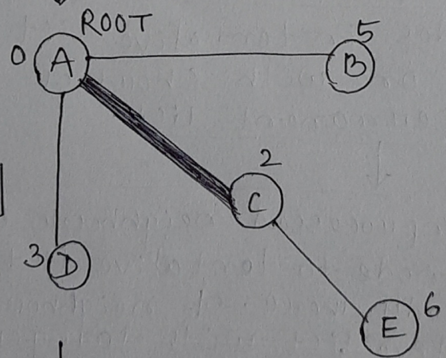


Move A to permanent list & B, C, D to tentative list

↓

Permanent: A(0) C(2)

Temporary: B(5) D(3) E(6)

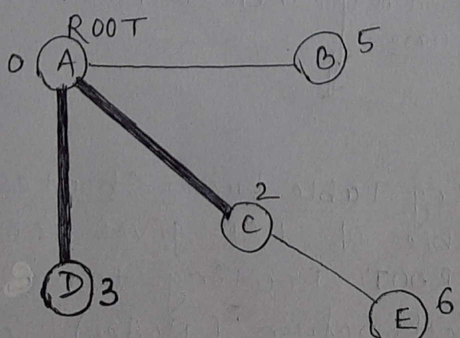


Move C to permanent list & add E to the tentative list.

↓

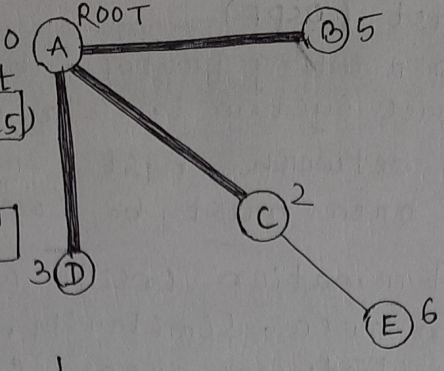
Permanent: A(0) C(2) D(3)

Temporary: B(5) E(6)



Move D to permanent list.

↓



Permanent

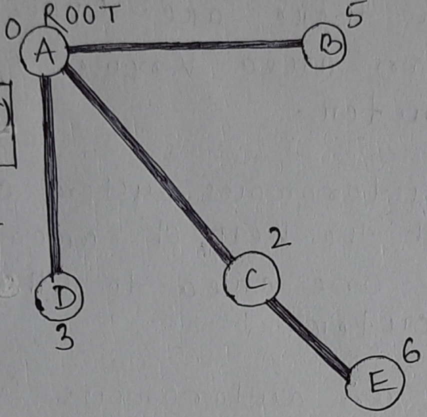
A(0) C(2) D(3) B(5)

Temporary

E (6)

Move B to permanent list

↓



Permanent

A(0) C(2) D(3) B(5) E(6)

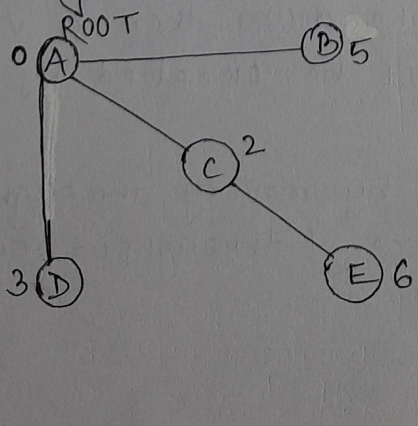
Temporary

Empty

Move E to permanent list Now, tentative list is empty. Stop the Processing here.

In the above example when B moved to permanent list, its neighbour is E which is already in the tentative list. So now we can see that previously cumulative cost to E from root A is 6, but when B moves to permanent list, its neighbour E has cumulative cost from root A is 8. Then we compare E(6) & E(8) and the lowest cost is E(6). Therefore, we keep previous record E(6) & a path exist between C & E. (Not B & E).

Routing table for the derived shortest path tree,



NODE	Cost	Next-HOP
A	0	—
B	5	—
C	2	—
D	3	—
E	6	C

Routing table for node A.

Open Shortest path First (OSPF)-

→ It is an intradomain routing protocol based on LSR. OSPF divides autonomous system into areas. And an area is collection of network, host & routers. All the networks inside an area must be connected.

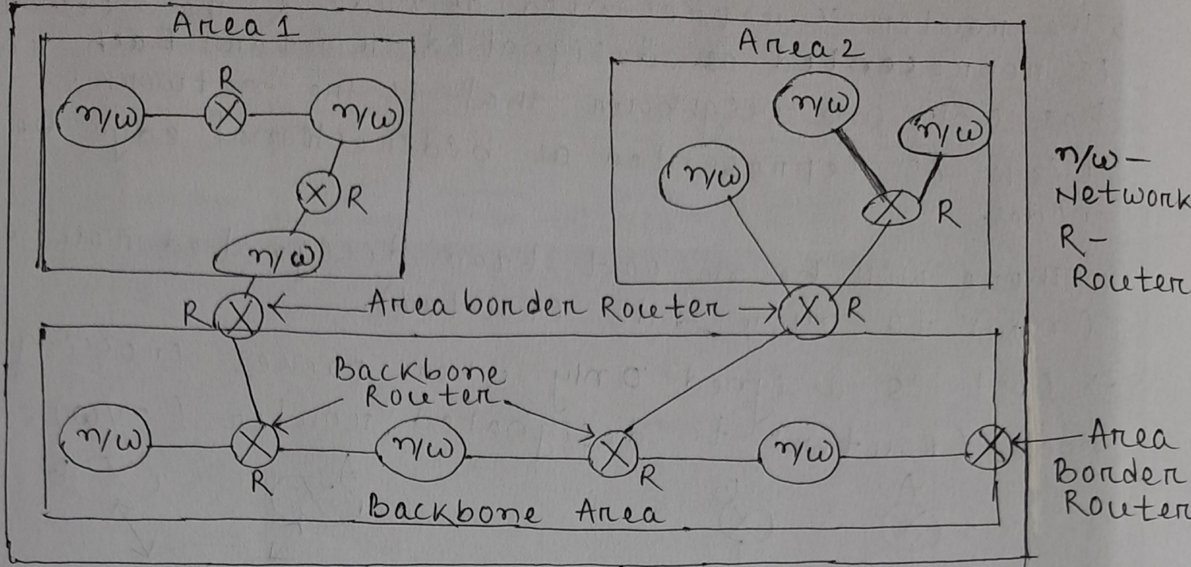
→ To flood routing information inside an area there are routers inside an area. Similarily, to flood the routing information between areas i.e. from one area to other area there are special routers at the border of an area known as border router or area border router.

→ Areas inside an autonomous system are connected with each other with the help of area border router or packets pass from one area to other area through this area border router.

→ Among area inside an autonomous system, there is a special area known as backbone area. A backbone area serve as primary area & all other area serve as secondary area. All the secondary areas must be connected with the primary area inside an autonomous system. And routers present inside backbone area are known as backbone routers. A backbone router can also be act as a area border router.

→ In an autonomous system there must be connectivity between an area border router with backbone router. If due to any reason link between backbone router & area border router broke down then a virtual link between routers must be created by the network administrator.

→ Cost in OSPF is assigned by moving through a network with type of service (throughput or minimum delay etc).



Autonomous System

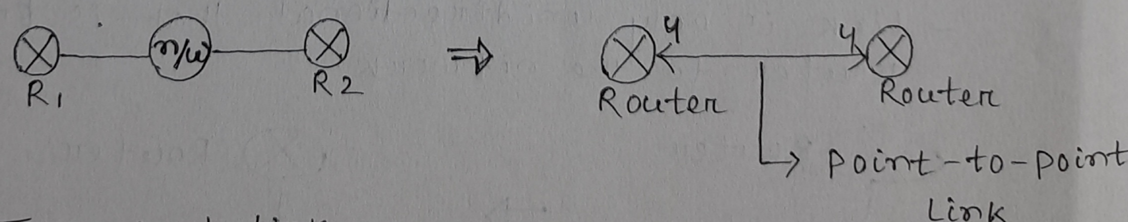
Types of links in autonomous system →

(a) Point-to-point link
(b) Transient link
(c) Stub link
(d) Virtual link

(a) Point-to-point link –

→ This link connects two routers with each other and there is only one link on network between routers.

→ when point-to-point link is graphically represented then each router is represented by a node & link (n/w) is represented by a bidirectional edge with cost or metric in each direction. (cost is same in each direction).



(b) Transient link –

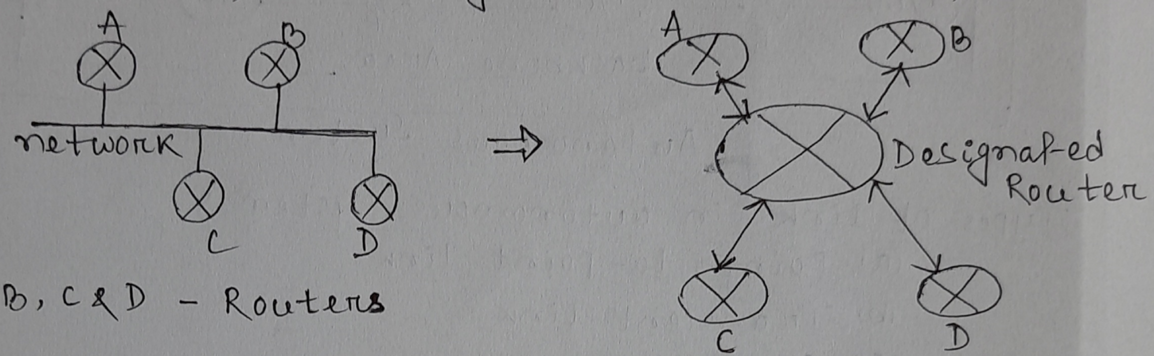
→ A transient link is a network with several routers connected to it. i:e: data packet can enter through any router & leave from any router.

→ To show that each router is connected through one single network to every other router.

→ The router is represented as node & the network is represented as designated router. Each router has only one neighbour that is the network. Link is represented as bidirectional edge between nodes.

→ There will be no cost from designated router (n/w) to other node (Router).

→ Cost is assigned only when packet crossing from node (Router) to designated router (n/w).



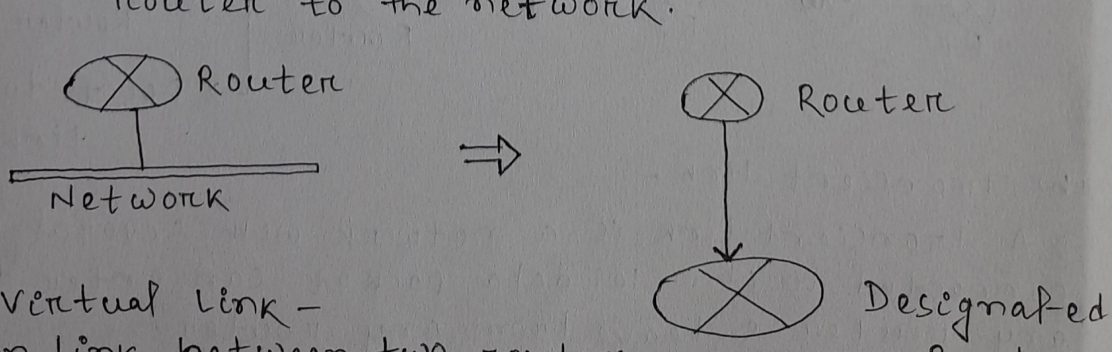
A, B, C & D - Routers

(c) Stub link -

→ A stub link is a network that connected with only one router that is data packets enter to the network through this single router & also data packet leave from the network through this single router.

→ When stub link is represented by graphical representation,

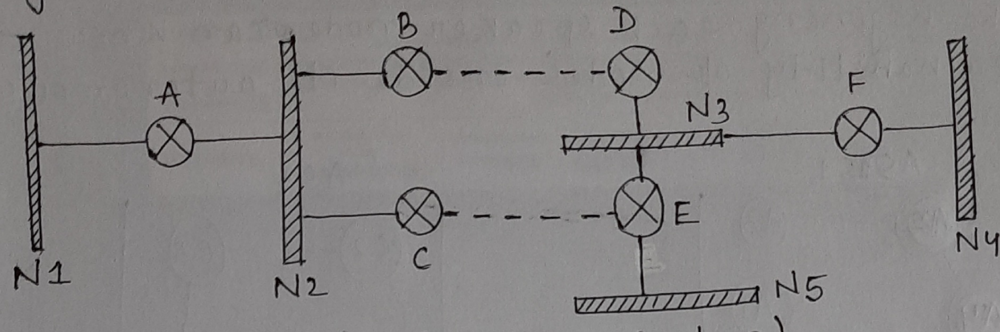
- a router is represented as a node.
- Network is represented by designated router
- link is only unidirectional that is from the router to the network.



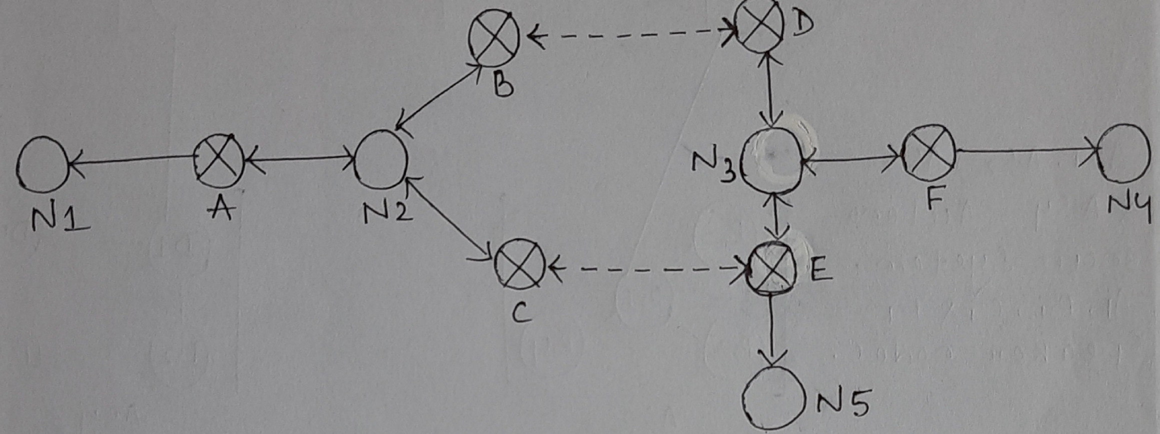
(d) Virtual Link -

When link between two routers is broken, then the network administrator may create a virtual link between them.

e.g of OSPF -



(Autonomous System)



(Graphical Representation of autonomous system)

Path vector Routing (PVR)

→ Path vector routing is used for interdomain routing
In PVR there is one node in each autonomous system that acts on behalf of the entire autonomous system. This node acts as a speaker node for an autonomous system or this node is known as speaker node.

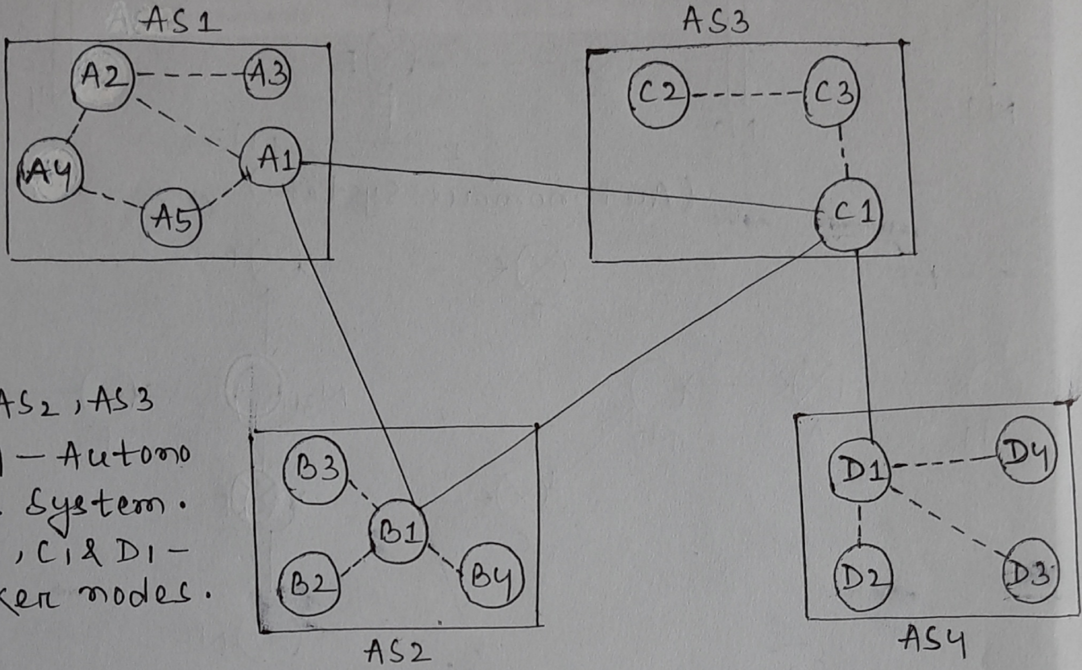
→ Speaker node in autonomous system creates routing table & then advertise it to neighbouring speaker node in neighbouring autonomous system.

→ Speaker node of different autonomous system can communicate with each other.

→ A speaker node advertises the path not the metric or cost to reach the destination. There are 3 steps of PVR process. They are -
(a) Initialization
(b) Sharing
(c) Updating

(a) Initialization -

At the begining each speaker node can know only the reachability of nodes inside its autonomous system.



AS1, AS2, AS3 & AS4 – Autonomous system.
A1, B1, C1 & D1 – Speaker nodes.

Dest	Path
A1	AS1
A2	AS1
A3	AS1
A4	AS1
A5	AS1

A1 Table

Dest	Path
B1	AS2
B2	AS2
B3	AS2
B4	AS2

B1 Table

Dest	Path
C1	AS3
C2	AS3
C3	AS3

C1 Table

Dest	Path
D1	AS4
D2	AS4
D3	AS4
D4	AS4

D1 Table

→ In the above figure A1 is the speaker node of AS1, B1 is the speaker node of AS2, C1 is the speaker node of AS3 & D1 is the speaker node of AS4.

→ Node A1 initializes routing table that shows A1 to A5 are located inside AS1 & can be reached through speaker node A1. Now speaker node A1 advertises this to other speaker nodes of neighbour autonomous system. This initialization & advertisement is done for all the speaker nodes.

(b) Sharing -

→ A speaker node communicates with its immediate neighbour in an autonomous system. For sharing its routing table.

→ speaker node A1 share its routing table with the speaker nodes of neighbour autonomous system. i.e: with B1 & C1. Then C1 also share its routing tables with its immediate neighbour speaker nodes B1 & D1. B1 share its routing table to A1 & C1. Similarily D1 also share its routing table to C1 because C1 is immediate neighbour of D1.

(C) Updation phase -

→ when a speaker nodes receives a two-column routing table from its neighbour speaker node, the receiving speakernode update its own routing table by adding the nodes that are not present in its own routing table. The second column of routing table contain AS (autonomous system) of receiving node with AS (autonomous system) of sending speaker node.

→ All the other node except speaker node when wants to share some data packet to other nodes of other autonomous system then that data packet must be forwarded through the speaker nodes of sending & receiving autonomous system.

→ e.g. $A_2 \xrightarrow[\text{Packet}]{\text{Data}} C_3$

$A_2 \xrightarrow[\text{Packet}]{\text{Data}} A1 \xrightarrow[\text{Packet}]{\text{Data}} C1 \xrightarrow[\text{Packet}]{\text{Data}} C3$

After updation phase the final routing table of each speaker node in each autonomous system becomes,

Dest	Path
A1	AS1
A2	AS1
A3	AS1
A4	AS1
A5	AS1
B1	AS1 - AS2
B2	AS1 - AS2
B3	AS1 - AS2
B4	AS1 - AS2
C1	AS1 - AS3
C2	AS1 - AS3
C3	AS1 - AS3
D1	AS1 - AS3 - AS4
D2	AS1 - AS3 - AS4
D3	AS1 - AS3 - AS4
D4	AS1 - AS3 - AS4

A1 Routing Table

Dest	Path
A1	AS2 - AS1
A2	AS2 - AS1
A3	AS2 - AS1
A4	AS2 - AS1
A5	AS2 - AS1
B1	AS2
B2	AS2
B3	AS2
B4	AS2
C1	AS2 - AS3
C2	AS2 - AS3
C3	AS2 - AS3
D1	AS2 - AS3 - AS4
D2	AS2 - AS3 - AS4
D3	AS2 - AS3 - AS4
D4	AS2 - AS3 - AS4

B1 Routing Table

Dest	Path
A1	AS3 - AS1
A2	AS3 - AS1
A3	AS3 - AS1
A4	AS3 - AS1
A5	AS3 - AS1
B1	AS3 - AS2
B2	AS3 - AS2
B3	AS3 - AS2
B4	AS3 - AS2
C1	AS3
C2	AS3
C3	AS3
D1	AS3 - AS4
D2	AS3 - AS4
D3	AS3 - AS4
D4	AS3 - AS4

C1 Routing Table

Dest	Path
A1	AS4 - AS3 - AS1
A2	AS4 - AS3 - AS1
A3	AS4 - AS3 - AS1
A4	AS4 - AS3 - AS1
A5	AS4 - AS3 - AS1
B1	AS4 - AS3 - AS2
B2	AS4 - AS3 - AS2
B3	AS4 - AS3 - AS2
B4	AS4 - AS3 - AS2
C1	AS4 - AS3
C2	AS4 - AS3
C3	AS4 - AS3
D1	AS4
D2	AS4
D3	AS4
D4	AS4

D1 Routing Table

BGP (BORDER GATEWAY PROTOCOL)

→ It is an interdomain routing protocol. It is based on PVR. The autonomous is categorized into the following types-

(a) Stub AS

(b) Multihomed AS

(c) Transient AS

a) Stub AS -

→ A stub AS has only one connection to another AS. Host in AS can send data traffic to other AS. Hosts in AS can receive data coming from hosts in another AS.

→ Data traffic can not pass through a stub AS. A stub AS can be either a source or sink. e.g- small corporation or Local ISP.

b) Multihomed AS -

→ A multihomed AS can have more than one connection to other AS. It can receive or send data traffic from more than one AS.

→ It does not allow data traffic coming from one AS & going to another AS to pass through it.

→ It is also source or sink of data traffic. e.g - large corporation that connects more than one reginal AS or national AS.

c) Transient AS -

→ Transient AS is one kind of multihomed AS that allows transient data traffic. i.e: data traffic can pass through it. e.g - national or international ISP.

Path Attributes of BGP Protocol →

Path attributes gives information about path. The list of path attributes helps the receiving router to give more information for ~~x~~ deciding or choosing the policy.

Attributes divided or catagorized into 2 different types. They are -

a) Well-Known attributes
- → Well-Known mandatory attributes.
- → Well-Known discretionary attributes

b) Optional attributes
- → Optional transitive attributes.
- → Optinal non-transitive attributes.

a) Well-Known attributes →

Well-Known attributes are those that every BGP router must recognize.

i) Mandatory attributes - Mandatory attributes are those that must appear in the description of a route. e.g - Origin, Next-hop, AS (speaker node) etc

ii) Discretionary attributes - These attributes must be recognized by each & every router but is not required to be included in every update message. e.g - cost or metric.

b) Optional attributes →

These attributes need not be recognized by every router.

i) Transitive attributes - These attributes must be passed to next router by the router that has not implemented this attributes.

ii) Non-transitive attributes - These attributes must be discarded if the receiving router has not implemented it.

BGP session →

Exchange of routing information between two routers using BGP takes place in a session called BGP session.

→ BGP uses services of TCP. When a TCP connection is created for BGP then it can last for a long time until some unusual happens. So for this reason BGP session is called semipermanent connection.

→ Types of BGP session -

a) External BGP session (E-BGP)

b) Internal BGP session (I-BGP)

E-BGP Session

→ E-BGP session is used to exchange information between two speaker nodes belonging to two different autonomous system.

→ Two speaker nodes of two different autonomous system exchange information after collecting information from other routers in an autonomous system which is done by I-BGP is an E-BGP session.
e.g- session established between AS1 & AS2 is an E-BGP session.

I-BGP Session

→ I-BGP session is used to exchange information between two routers inside an autonomous system.

→ Collects information from other routers in autonomous system is done by using the I-BGP session.