BCS502 – COMPUTER NETWORKS

Faculty:

Prof. Ashok Herur

ashok. herur @ eastpoint. ac. in

Module 3

Network Layer

Main topics in Module 3 – Network layer

- Network Layer Services
- Packet Switching
- IPv4 addressing, IPv4 datagram, IPv6 datagram
- Introduction to Routing algorithms
- Unicast Routing algorithms
- Multicast Routing algorithms

Network layer services

Network layer services

- **Functions of the Network Layer:**
- The Network layer makes sure that the packets are sent all the way from the source to the destination, in as efficient a manner as possible.
  - Data link layer just moved the frames from one node to the adjacent one.
- To be able to do this, the Network layer must:
  - Know the topology of the network (the set of all routers and links);
  - Balance the load to avoid overloading some links while leaving others under-utilised.

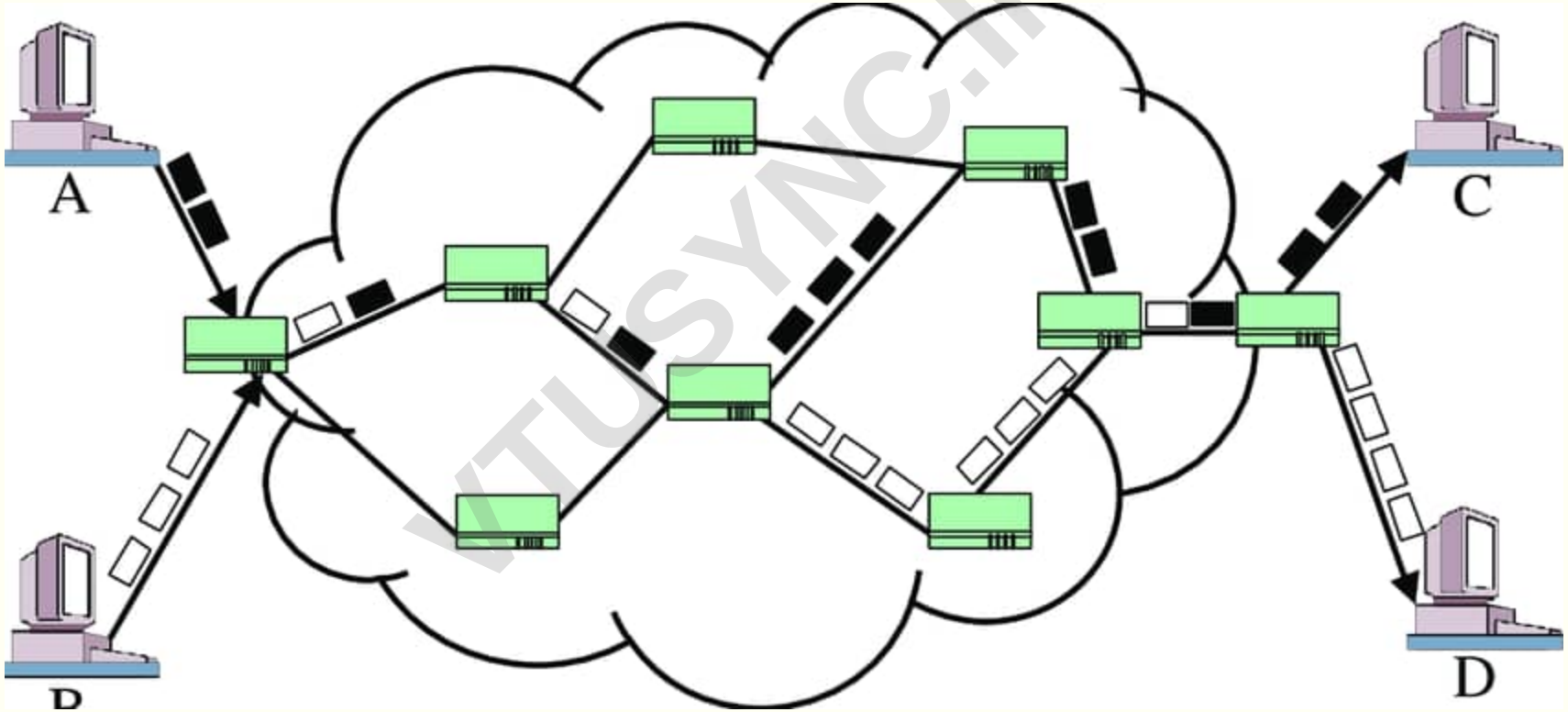
Review of Packet switching.

- Every message is segmented by the Transport layer to comply with the packet size limit imposed by the networks.
- Each packet has information of the message to which it belongs, as well as the destination address.
- There is no dedicated, end-to-end connection.
- The packets are then received, buffered, processed and forwarded by the numerous routers along the way.
- Hence, there is a certain delay encountered at each router.

Review of Packet switching

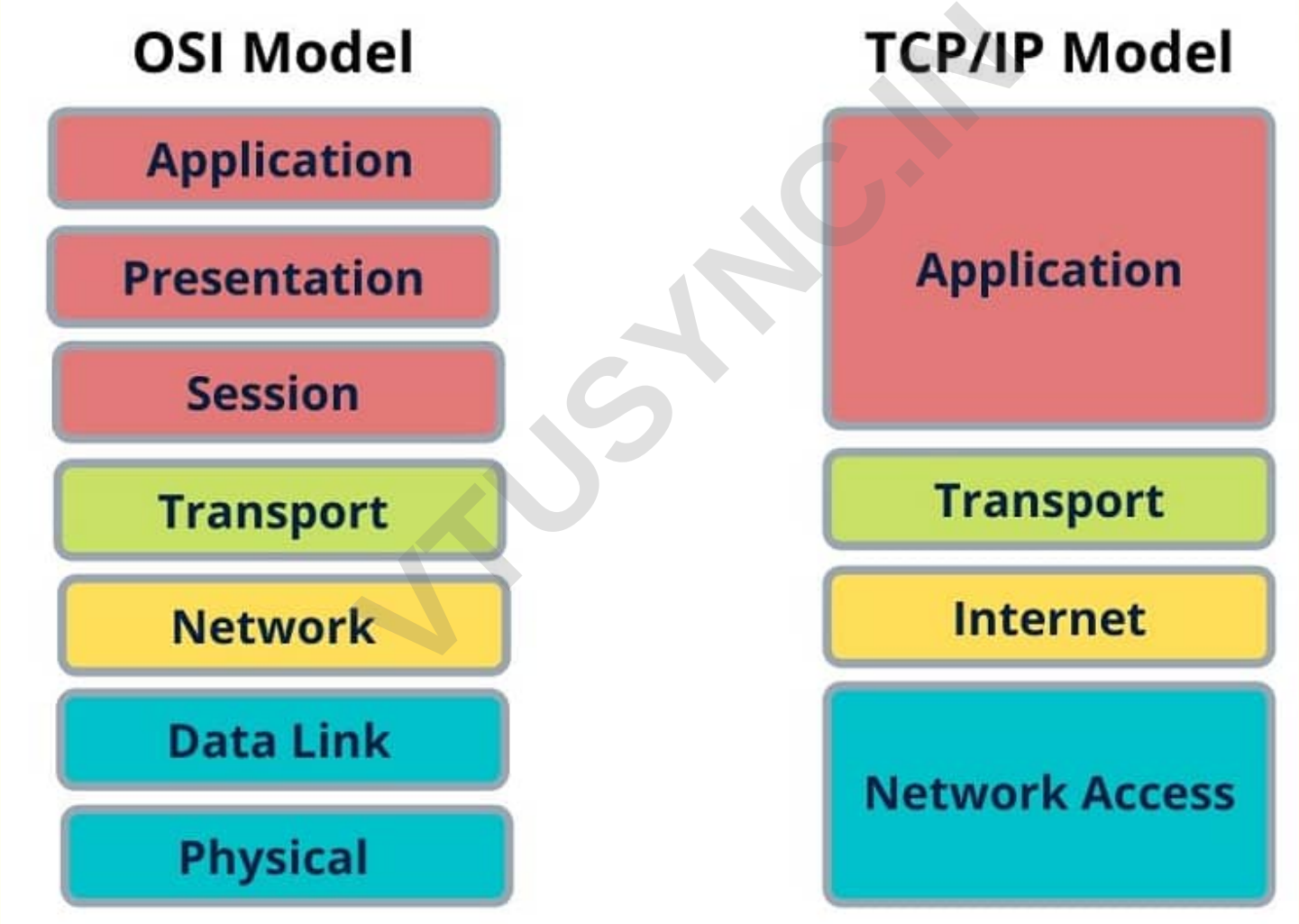
- Advantages:
  - Since the bandwidth of a link is not dedicated (but shared with packets from numerous hosts), the utilisation of the bandwidth is much higher.
  - No set-up time.
- Two methods of Packet switching :
  - Connectionless (Datagram Packet switching);
  - Connection-oriented (Virtual Circuit Packet switching).

Packet switching – Hosts sharing the bandwidth



IPv4 and IPv6

Network layer is called INTERNET layer in TCP/IP model



Internet Protocol (IP)

- IPv4 and IPv6 are the most popular protocols used on the Internet, with IPv6 succeeding IPv4.
- They define the logical addressing of the hosts as well as the datagram (packet) format.
- The hosts as well as the routers understand this protocol.
- The logical address is used to route the datagram from router to router, till they are delivered to the destination host.

IPv4 addresses

- Every device on the WAN is identified by an unique address.
- All IPv4 addresses are 32 bits long.
- They are represented as four octets in dotted decimal format.
    - Example: **132.10.68.73**
- The IP address has two components:
    - The Network ID – Allotted by a global authority
    - The Host ID – Allotted locally by the network administrator

IPv4 address classes

- IP addresses are divided into 5 classes: A,B, C, D, E.
- Class D addresses are used for multicasting while class E is not used
- The Network ID is 1, 2 and 3 bytes long in Class A, B and C respectively, with the remaining part being the Host ID.
- Therefore Class A networks are very large and can accommodate a huge number of hosts within them, while Class C networks are very small.

IPv4 address classes

Network & Host Representation By IP Address Class				
Class	*Octet1*	*Octet2*	*Octet3*	*Octet4*
Class A	0 Network	Host	Host	Host
Class B	1 0 Network	Network	Host	Host
Class C	1 1 0 Network	Network	Network	Host

IPv4 address classes

IP Address Class Assignments	
Class	*First Octet Value*
Class A	0 ~ 127
Class B	128 ~ 191
Class C	192 ~ 223
Class D	224 ~ 239
Class E	240 ~ 255

IPv4 address classes

Characteristics of the IP Address Classes						
Class	Address Range	Identify Bits (binary value)	Bits in Network ID	Number of Networks	Bits in Host ID	Number of Hosts/ Network
A	0 ~ 127	1 (0)	7	126	24	16,777,214
B	128~191	2 (10)	14	16,382	16	65,534
C	192~223	3 (110)	21	2,097,150	8	254

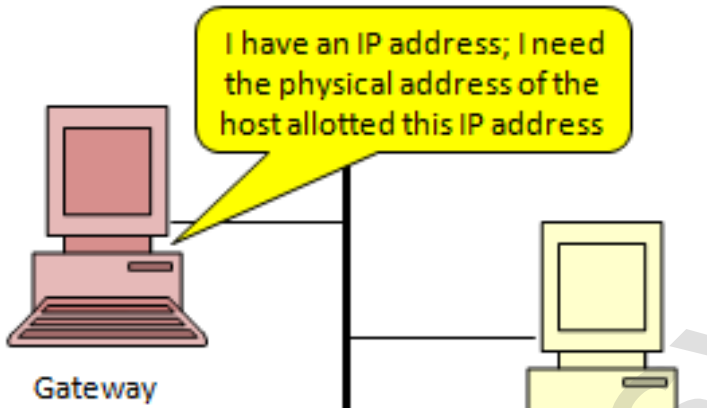
Private IP addresses

- Private IP addresses are valid only within an organization
  - The same addresses can be used in multiple organizations
- When leave the network of the organization, the private addresses are converted to public addresses by the proxy server, and vice-versa.
- Private addresses:
  - Class A: 10.0.0.0 to 10.255.255.255 --- 1 network
  - Class B: 172.16.0.0 to 172.31.255.255 --- 16 networks
  - Class C: 192.168.0.0 to 192.168.255.255 --- 256 networks

Address Resolution Protocol (ARP)

- Routers use the network portion of an IP address to route the packet to the network containing the destination host.
- When the packet reaches the ‘gate’ of the network (belonging to an organization or an ISP), the Gateway device uses the host portion of the IP address to find out where the particular host is, and then deliver the packet.
- Since hosts are uniquely and permanently identified by their physical (MAC) addresses, the IP addresses should be mapped to the physical addresses for eventual delivery of the packet.
    - ARP helps in this mapping.

Address Resolution Protocol (ARP)



- The Gateway broadcasts the IP address and waits for a response.
- Upon receiving the Physical address, it forms an Ethernet frame with that address and sends it on the Bus.

Subnetting

- Network administrators divide large networks, into smaller ones called *sub-networks* (or just subnets), to **improve network performance.**

- Another reason for having subnets is to **reduce the size of a broadcast domain**, when broadcast traffic begins to consume too much of the available bandwidth.
  - Broadcasts are sent to all hosts on a network or a subnetwork.

- Having subnets also leads to **efficient routing** of packets within the network.
- The subnet ID is created by borrowing bits from the host portion of the IP address

Subnetting – Example 1

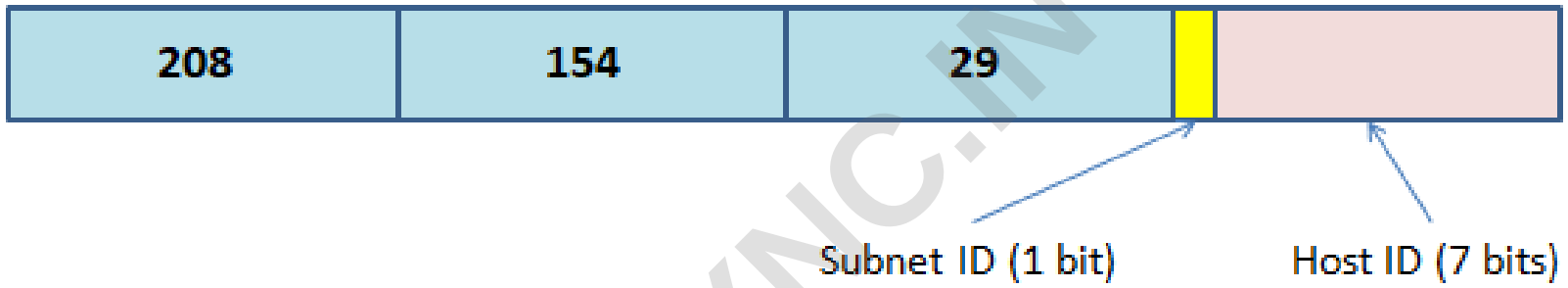
The network 208.154.29.0 needs to be divided into 2 subnets;

Subnet 1 has to support 100 hosts while Subnet 2 has to support 80 hosts.

Design and calculate the range of addresses to be used in both the subnets.

- The network 208.154.29.0 is a Class C network
  - The host ID consists of the last 8 bits of the IP address
  - It can support a total of 254 hosts (when not sub-netted)
- How many bits are required for the subnet ID? ---- Only 1 bit in this case
- How many bits remain for the host ID portion? ---- 7 bits
- How many hosts can each subnet support ? ---- 126 hosts each
- So, the stated requirement of 100 hosts and 80 hosts can be met.

Subnetting – Example 1



- Subnet 1 would have a Subnet ID of '0' while Subnet 2 would have a Subnet ID of '1'
- Host IDs in both the subnets can range from 000 0001 to 111 1110
- Therefore, the fourth byte for subnet 1 can be: 1 to 126, and for subnet 2 can be 129 to 254
- IP address range for Subnet 1 is: 208.154.29.1 to 208.154.29.100
- IP address range for Subnet 2 is:: 208.154.29.129 to 208.154.29.208

Subnetting – Example 2

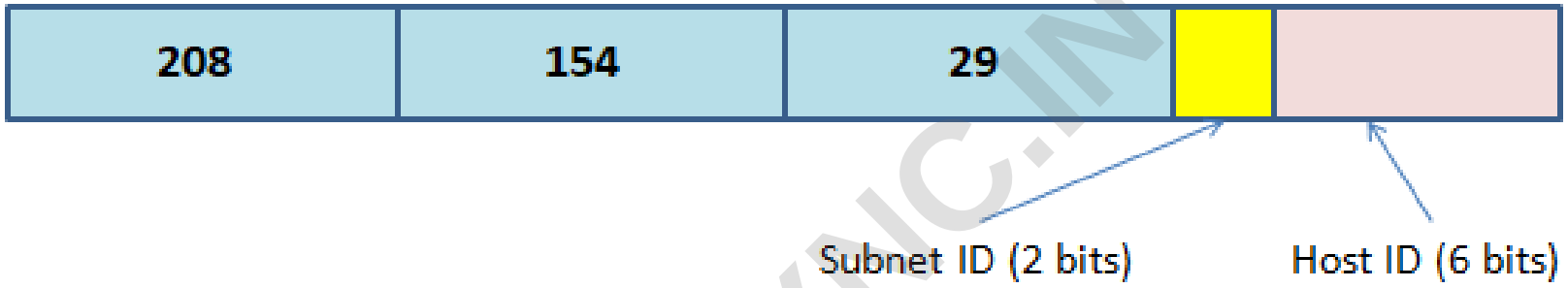
The network 208.154.29.0 needs to be divided into 3 subnets;

The three Subnets have to support a maximum of 50 hosts each.

Design and calculate the range of addresses to be used in the 3 subnets.

- The network 208.154.29.0 is a Class C network
  - The host ID consists of the last 8 bits of the IP address
  - It can support a total of 254 hosts (when not sub-netted)
- How many bits are required for the subnet ID? ---- 2 bits
- How many bits remain for the host ID portion? ---- 6 bits
- How many hosts can each subnet support ? ---- 62 hosts each
- So, the stated requirement of 50 hosts each can be met.

Subnetting – Example 2



- Subnet 1 would have a Subnet ID of '00' while Subnet 2 would have a Subnet ID of '01' and Subnet 3 would have an ID of '10'
- Host IDs in all the subnets can range from 000 0001 to 111 1110
- IP address range for Subnet 1 is: 208.154.29.1 to 208.154.29.62
- IP address range for Subnet 2 is:: 208.154.29.65 to 208.154.29.126
- IP address range for Subnet 3 is:: 208.154.29.129 to 208.154.29.190

Masking

- A ‘mask’ is used to mask out the host portion of an IP address and get only the network (and sub-network) ID
- A mask is a string of 1’s followed by a string of 0’s
  - The number of 1’s equal the number of bits allotted for the network and sub-network ID
- The masking operation is performed by a logic AND operation of the given IP address and the mask for that network

Masking

- The mask for a Class A network with no subnets is 1111111100…….0, which is 255.0.0.0; This is called as Default mask for Class A networks.
  - The Default mask for a Class B network is 255.255.0.0

- The mask for a Class A network with 4 subnets will have 10 ones followed by 22 zeros: 11111111 11000000 00000000 00000000, which is 255.192.0.0
  - The mask for a Class A network with 256 subnets will 255.255.0.0 (same as the default mask for Class B)

Masking

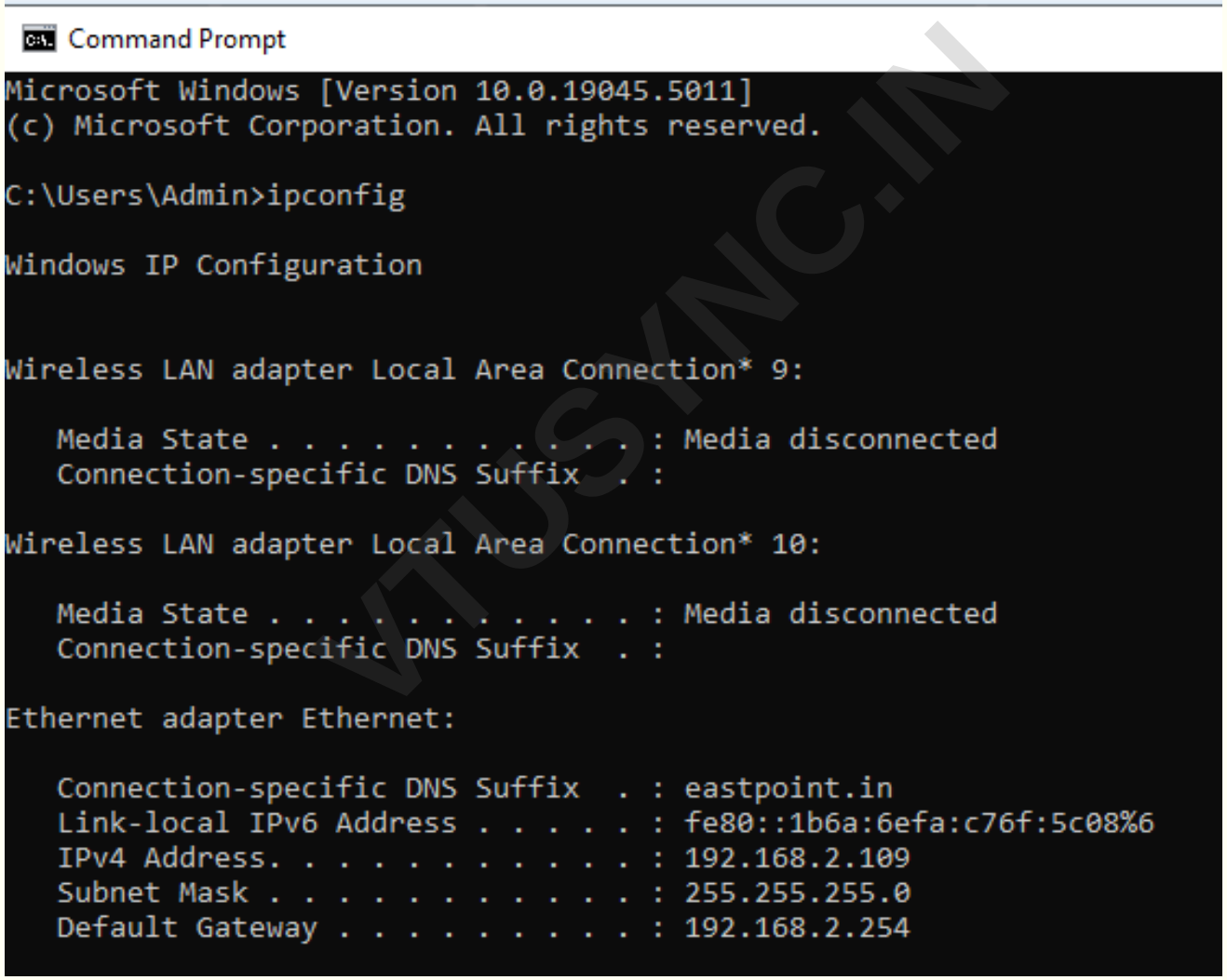
- Find out the Network ID and the Sub-network ID (if any) for the host with an IP address 140.37.163.25 if the mask is 255.255.224.0

- 140.37.163.25 – 10001100. 00100101. 10100011. 00011001

  255.255.224.0 – 11111111. 11111111. 11100000. 00000000

- The AND operation of these 2 would give us the Class B Network ID as 140.37.0.0 and the Sub-network ID bits as 101

Masking



Classful addressing

- Classful network design for IPv4 defined the network address as one or more 8-bit groups, resulting in the blocks of Class A, B, or C addresses.

- Thus, the smallest allocation and routing block contained only 256 addresses - too small for most enterprises, and the next larger block contained 65536 addresses - too large to be used efficiently by even large organizations.

- This led to inefficiencies in
  - Address use: Large number of unused addresses in Class B and A
  - Routing: The large number of allocated small (class-C) networks created large routing tables and heavy demand on routing equipment.

Classless addressing

- **Classless Inter-Domain Routing (CIDR)** allocates address space to Internet Service Providers (ISPs) and end users on any address bit boundary, instead of on 8-bit segments.

- CIDR is based on *variable-length subnet masking* (VLSM), which allows a network to be divided into subnets of various sizes.

- In the CIDR notation for an IP address, the network address (or routing prefix) is written with a suffix that indicates the number of bits of the prefix, such as 192.168.2.0/24.

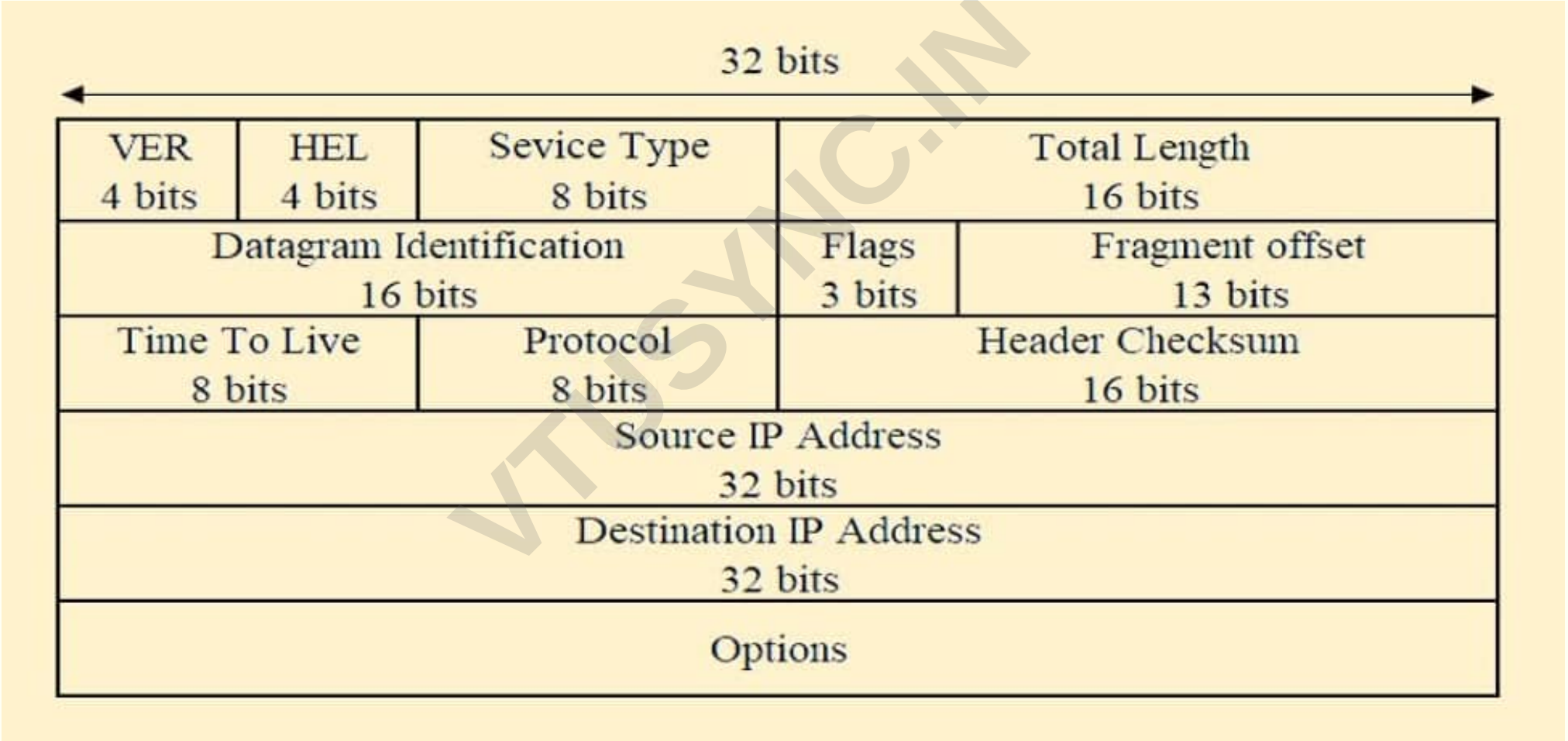
Classless addressing

- The /24 indicates that the network and subnetwork address is 24 bits long
  - Or that the host address is (32-24) = 8 bits in length

- 192.168.2.0/24 is the same as the older notation which mentioned the network ID and the Mask: 192.168.2.0/255.255.255.0
- The IPv4 block 182.168.100.0/22 represents 2^{10} = 1024 addresses from 192.168.100.0 to 192.168.103.255

IPv4 datagram header

IPv4 Datagram Header



IPv4 Header

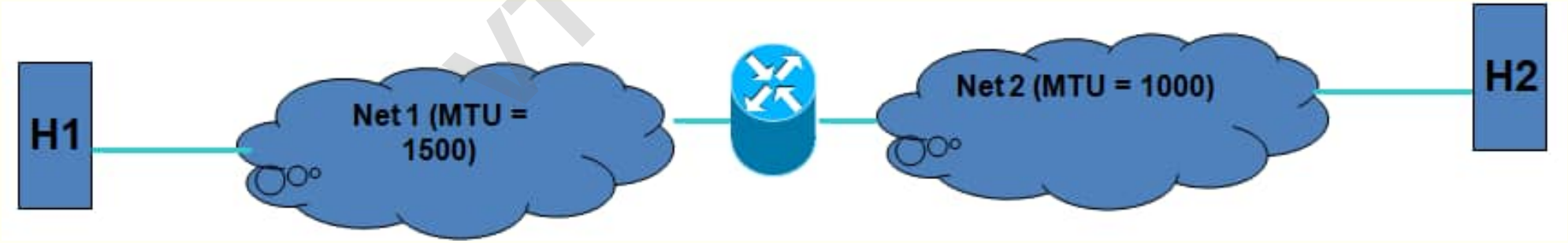
- The IP header is built up in blocks of 32 bits, and is always an integral number of 32-bit words.
- The IP header is divided up into various fields, each having a defined function.
- The 4-bit *version field* indicates the release version of the IP that is used in this datagram; IPv4 would have it as '0100'
- The 4-bit *header length* field identifies the length of the IP header in multiples of 32 bits (or 4 bytes)
- The minimum value for a valid header is 5 (means 5*32 bits = 20 Bytes), Maximum is 15 (means 15*32 bits = 60 Bytes).

IPv4 Header

- The *Service type* field specifies the parameters for the type of service requested by the packet from the network.
- This field specifies things like:
  - The priority to be given to the datagram;
  - Delay to be minimised;
  - Cost to be minimised;
  - Throughput to be maximised;
  - Reliability to be maximised.
- The networks use these parameters to define the handling of the datagram during its transport.

MTU, and Fragmentation

- Every network specifies the maximum size of the data field in a frame. This limit is called the maximum transmission unit (MTU) of that network.
- Any datagram encapsulated in a frame must therefore be smaller than the MTU for that network.
- What happens in the scenario depicted below?



IPv4 Header

- The *Total length* field indicates the length of the entire packet, in bytes.
  - With a 16-bit field, it allows packets as long as 65,535 bytes.
- When a packet is divided into smaller fragments to comply with the MTU restriction, each fragment of the packet will have the same *Identification* field as the parent packet.
- Of the three bits in the *Flag* field, the first one is not currently defined/ used.
- The second flag indicates whether fragmentation is allowed. It is called the *Don't fragment (DF)* flag. If the flag is set to 1, fragmentation is not allowed, and if it is set to 0 fragmentation is allowed.

IPv4 Header

- If the DF flag is set to 1, a datagram will be lost if it has to cross a network that cannot handle its size.
- The third flag is called the *More fragments (MF)* bit. It is used to indicate that there are more fragments to follow, so if a datagram is fragmented this is set on all except on the last fragment.
- The *Fragment offset* field is 13 bits long, and is used to indicate the position of the data in a fragment relative to the beginning of data in the original datagram.

IPv4 Header

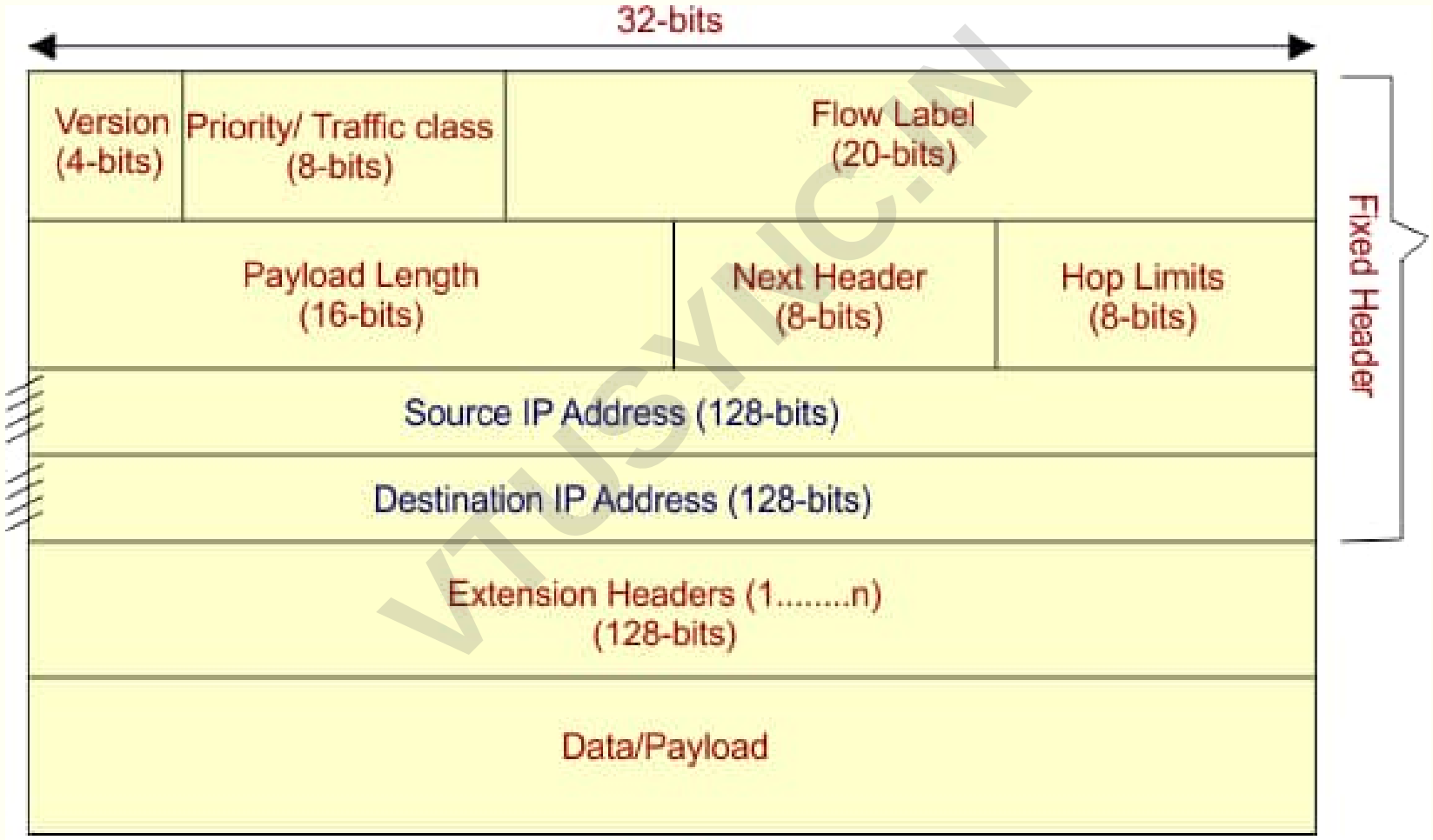
- The *TTL (Time-To-Live)* field contains a initial integer when a packet is sent out by the source.
- This number is decremented by 1 by every router which forwards it.
- When the TTL value becomes zero, routers discard them.
- The TTL field thereby puts a limit to the number of hops (life) a packet can go through and deals with stray packets which keep looping in the network.

IPv4 Header

- The *Header Checksum* field contains the checksum computed on all the bits of the IP header (not the Data)
- This checksum is verified by every router before forwarding the packet.
- If the checksum doesn't match, it indicates an error(s) in the header portion, and the router therefore discards the packet
- It does not attempt to inform the source that the packet is being dropped.

IPv6 datagram header

IPv6 Datagram Header



IPV6 Header

IPv6 Header

- Version field is the same as in IPv4 (of course, the value is now 6 (0110).
- Traffic class is same as “Type of Service” in IPv4.
- Payload length is same as “Total length” in IPv4.
- Hop Limits is same as TTL in IPv4.
- IP address length increased from 32 bits to 128 bits.
- Flow Label field is used by a source to label the packets belonging to the source and destination, in order to request special handling by intermediate IPv6 routers.
  - Between a source and destination, multiple flows may exist because many processes might be running at the same time.

IPv6 address classes

- IPv6 does not have classes, but its address space is divided into different types based on how they are used. Some types of IPv6 addresses include:
  - **IPv6 multicast address**: A major address class in IPv6, with the prefix FF00: allocated to all IPv6 multicast addresses.
  - **IPv6 unicast address**: Used to identify a single network interface. Each node on an IPv6 network has at least one unique unicast address.
  - **IPv6 anycast address**: Used to address multiple interfaces on a single multicast address.

Routing algorithms

Routing algorithms

- A routing algorithm helps decide which output line, an incoming packet should be transmitted on.
- In Datagram Packet switching, this decision must be made anew for every arriving packet since the 'best route' may have changed since the last packet.
- In VC Packet switching, these decisions (at every router along the way) are made when the VC is being set up. Thereafter, the packets are forwarded on the pre-decided route.
- The output of the routing algorithm is stored as a Routing (look-up) Table, one for each router.

Issues in Routing

- Finding the ‘best’ route is not always easy.
- The topology keeps changing: Lines may be added, may go down and would be brought back later.
- A packet may have to travel through heterogeneous networks.
- Delay variability is high.
- Have to deal with Congestion in the network.
- Multicasting and broadcasting is difficult.

Routing algorithms

- Desirable features of a routing algorithm:
  - Correctness
  - Simplicity
  - Robustness (be able to cope with changing topology and traffic)
  - Stability (quickly reach an equilibrium {route} and stay there)
  - Fairness (to all users)
  - Efficiency (at a global level)
- Often, Fairness and Efficiency are contradictory goals.

Fairness versus Efficiency

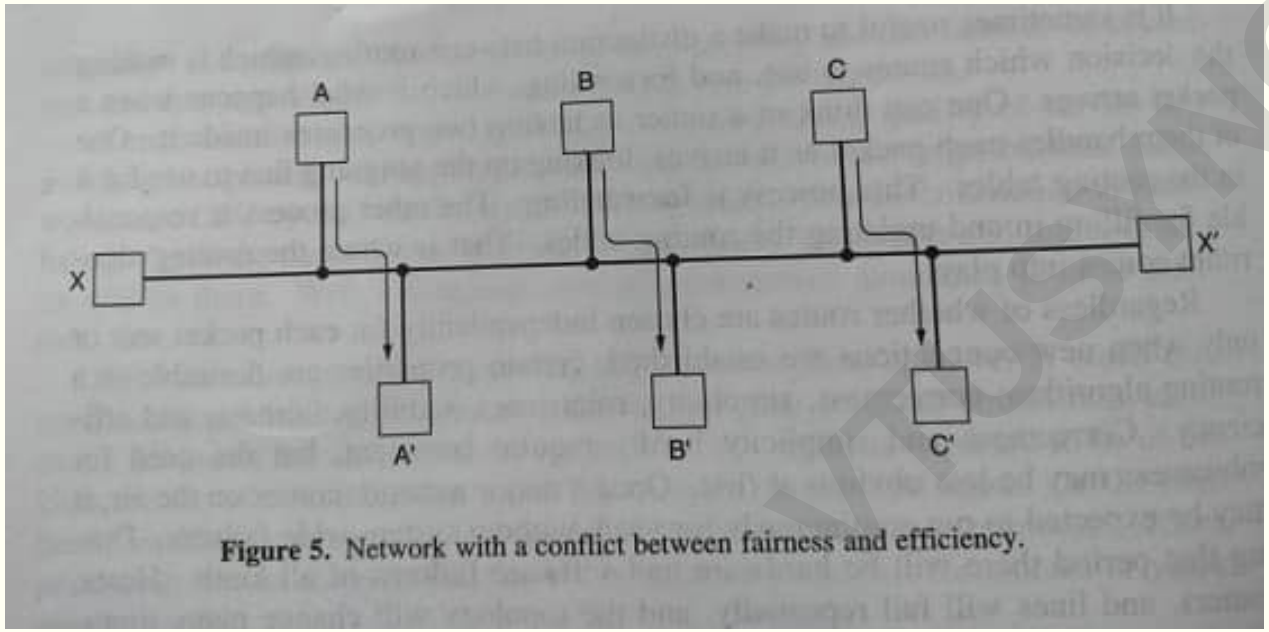


Figure 5. Network with a conflict between fairness and efficiency.

- Let's say that there is enough traffic between A and A', B and B', C and C' to saturate the horizontal links.
- To maximise the total flow, X to X' traffic should be cut off since it would 'eat up' the bandwidth on 3 links.
- Globally efficient but not fair to X and X'

Routing algorithms

- Routing algorithms try to:
  - Minimise the packet delay (mainly depends on the number of hops).
  - Reduce the physical transmission distance, as far as possible.
  - Reduce the bandwidth consumed per packet.
  - Improve the overall network throughput.
- Routing algorithms can be classified as:
  - Non-adaptive (Static) algorithms
  - Adaptive (Dynamic) algorithms

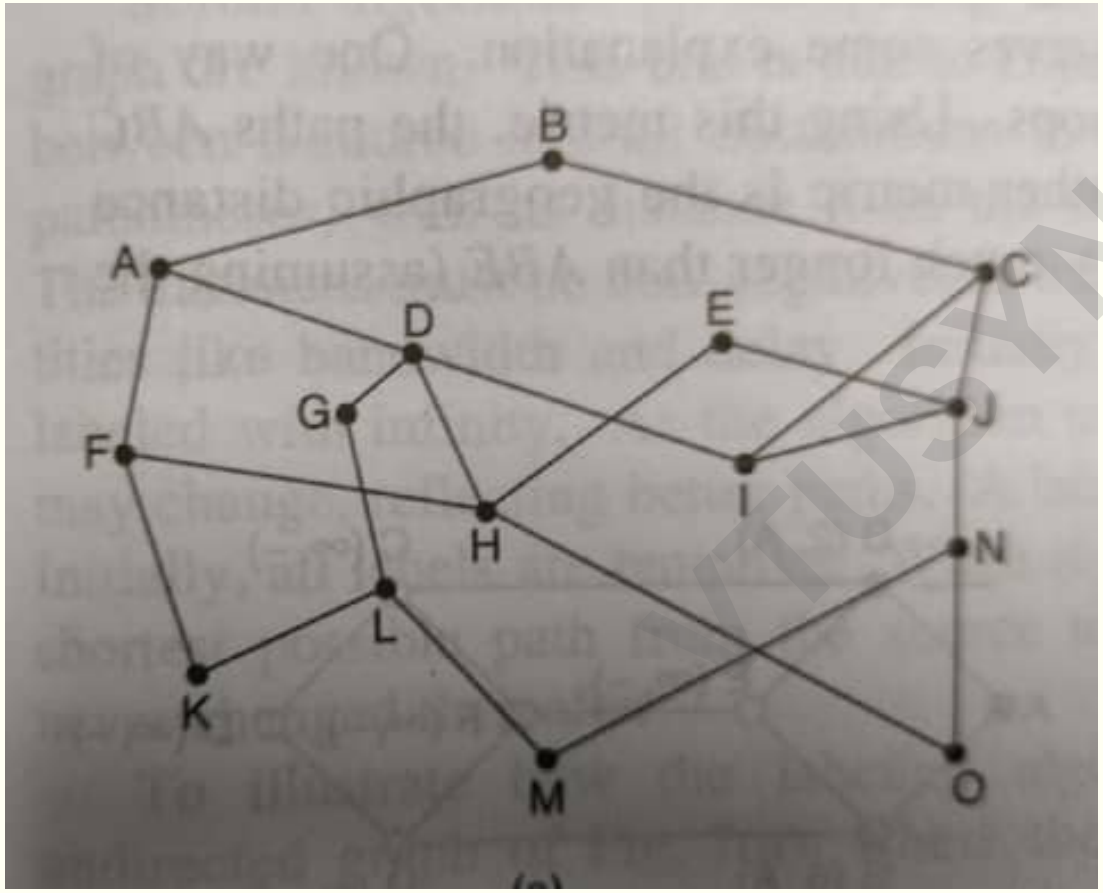
Non-adaptive routing algorithms

- Non-adaptive routing algorithms compute the route from node (router) I to node J, in advance and download that to all the routers along the way.
- The chosen routes remain static for a long period of time (could be days or weeks too).
- They do not base their routing decisions on estimates of the current traffic or short-term changes in topology.
- Since it does not respond to changes or link failures, it is only used in situations where the routing choice is clear (no choice) or has very few options.

Adaptive routing algorithms

- Adaptive routing algorithms base their routing decisions on estimates of the current traffic and changes in the topology.
- These algorithms differ from each other in:
  - Where they get their information (eg., from adjacent routers or from all routers in their sub-net)
  - When they change the routes (eg., when the topology changes or every ΔT seconds as the load changes)
  - What metric is used for optimization (eg., distance, number of hops, or estimated transit time).

The Optimality principle



- It states that if router J is on the optimal path from router I to router K, then the optimal path from J to K also falls along the same route.
- If there was a better route from J to K (say, r1) then the optimal path from I to K would have been (I - J - r1).

Routing algorithms

- Shortest Path algorithm (also called Dijkstra's algorithm)
- Flooding
- Distance Vector routing algorithm
- Link State routing algorithm
- Hierarchical routing algorithm
- Broadcast routing algorithm
- Multicast routing algorithm

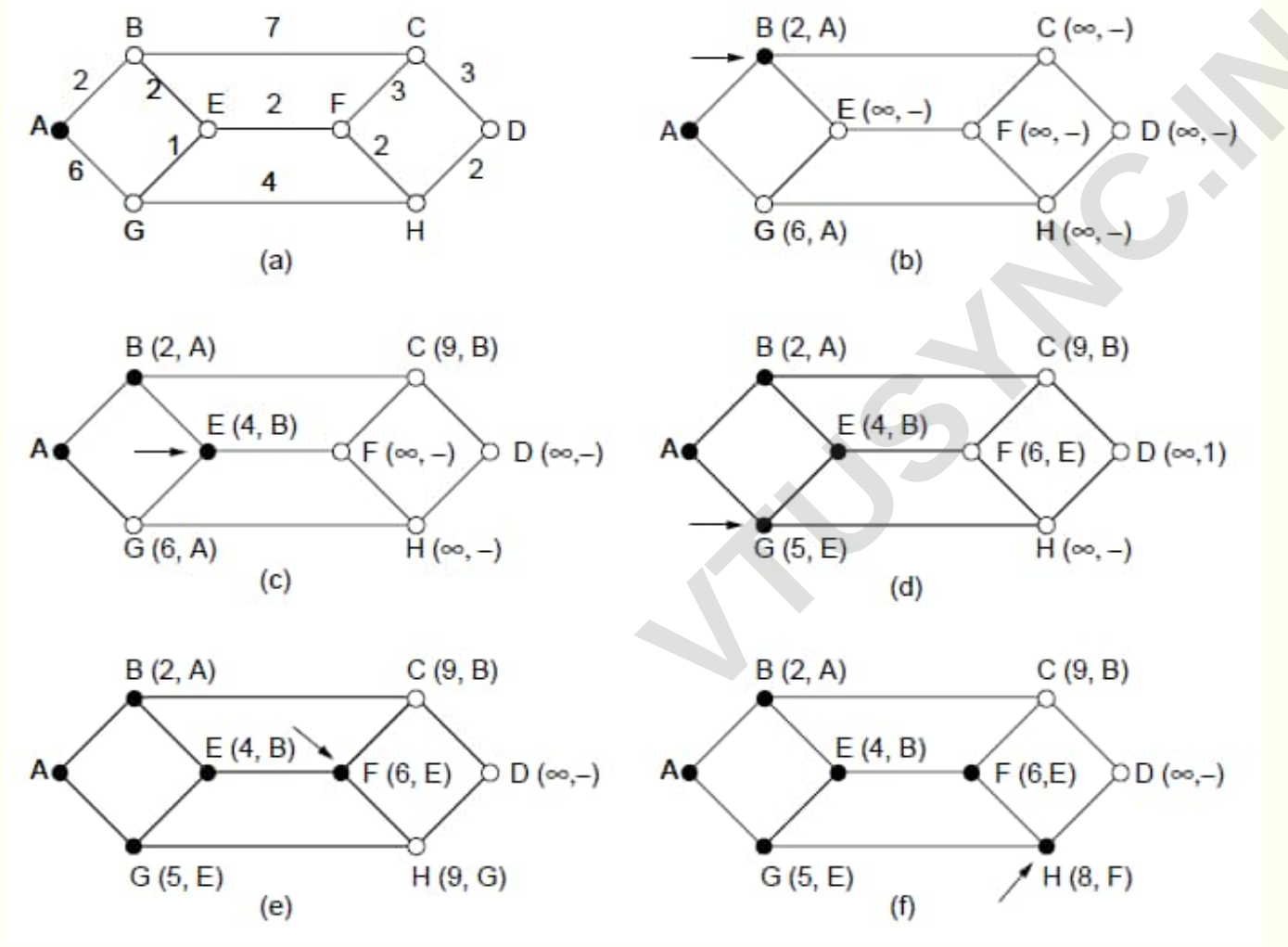
Shortest Path (Dijstra's) routing algorithm

- It is a non-adaptive (static) routing algorithm.
- It is also called as SPF (Shortest Path First) algorithm.
- It helps identify the "shortest" path between a given Source Node and a Destination Node.
- **Procedure:**
  - Start with the Source Node.
  - Mark it as permanent ('seen' node), and designate it as the working node.
  - Set the tentative distance from the Source node to all other nodes to "Infinity" (since, currently, we do not know how 'far' they are from the Source node).

Shortest Path (Dijstra's) routing algorithm

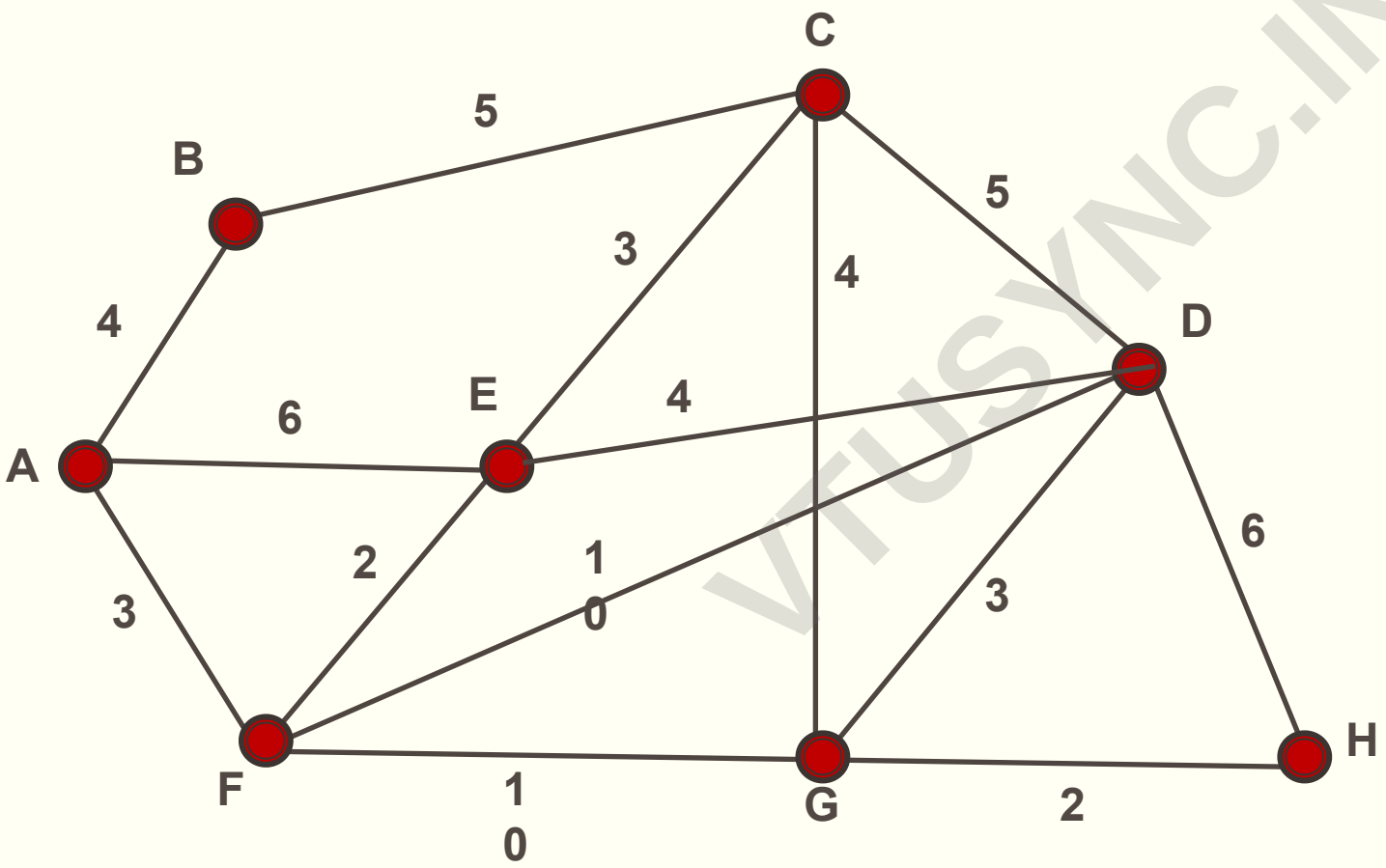
- While some nodes are still not marked permanent, do the following:
  - Compute the tentative distance from the source to all nodes adjacent to the working node.
  - If this is shorter than the current tentative distance, then replace the tentative distance and record the label of the working node there.
  - Select the node with the smallest value and make it the new working node. Designate the node permanent.

Shortest Path (Dijstra's) routing algorithm - Example



- Find the shortest path between node A and H.
- Using the labels, backtrack from H to A, and note down the chosen path.
- Path: A – B – E – F – H
- Distance metric = 8 units

Shortest Path (Dijstra's) routing algorithm - Example



- Example: Build the Routing Table for Node A.
- Find the shortest paths from A to all the other Nodes.
- Note the node to which A should forward to reach the respective nodes.
- Note the distance metric to those nodes.

Shortest Path (Dijstra's) routing algorithm - Example

Destination	Outgoing line	"Distance"
A	-	0
B	B	4
C	F	8
D	F	9
E	F	5
F	F	3
G	F	12
H	F	14

Flooding algorithm

- It is a static flooding algorithm for distributing packets to every part of a connected network in the shortest time.
- Flooding algorithms are used as a part of some routing protocols, generally those used in ad-hoc wireless networks.
- Each node tries to forward every message to every one of its neighbours.
  - This results in every message eventually being delivered to all reachable parts of the network.

Flooding algorithm

- Real-world flooding algorithms have to be more complex than this, since precautions have to be taken to
  - avoid wasted duplicate deliveries
  - avoid infinite loops
  - allow messages to eventually expire from the system.
- Selective Flooding Flood only in the direction of the destination.
- Practically useful in a few settings like Military Applications

Distance Vector routing algorithm

- Distance Vector routing algorithm (also known as Bellman-Ford algorithm) is an adaptive algorithm.
- These routing algorithms operate by having each router maintain a Table (Vector) giving the best known "distance" to each destination and which line to use to get there.
- The "distance" metric is normally the number of hops or the propagation time.
- The table at each router has one entry for each router in the subnet. For each entry, it has two parts:
  - The preferred outgoing line to use for that destination
  - An estimate of the time or 'distance' to that destination

Distance Vector routing algorithm

- Periodically (every 5-10 minutes), each router does the following:
  - Exchange their Tables with all the neighbouring routers every 5-10 minutes.
  - Estimate the “distance” to each neighbour.
  - If the distance metric is hops, the distance to each neighbour is just one hop.
  - If the distance metric is propagation delay, the router can measure it directly by sending ECHO packets to each neighbour (the neighbour just Time-stamps and sends it back)
  - The Routing table is then updated using the Tables received from the neighbours and the distance estimate to the neighbours.

Example 1 - DVR algorithm – Router J updating it's Table

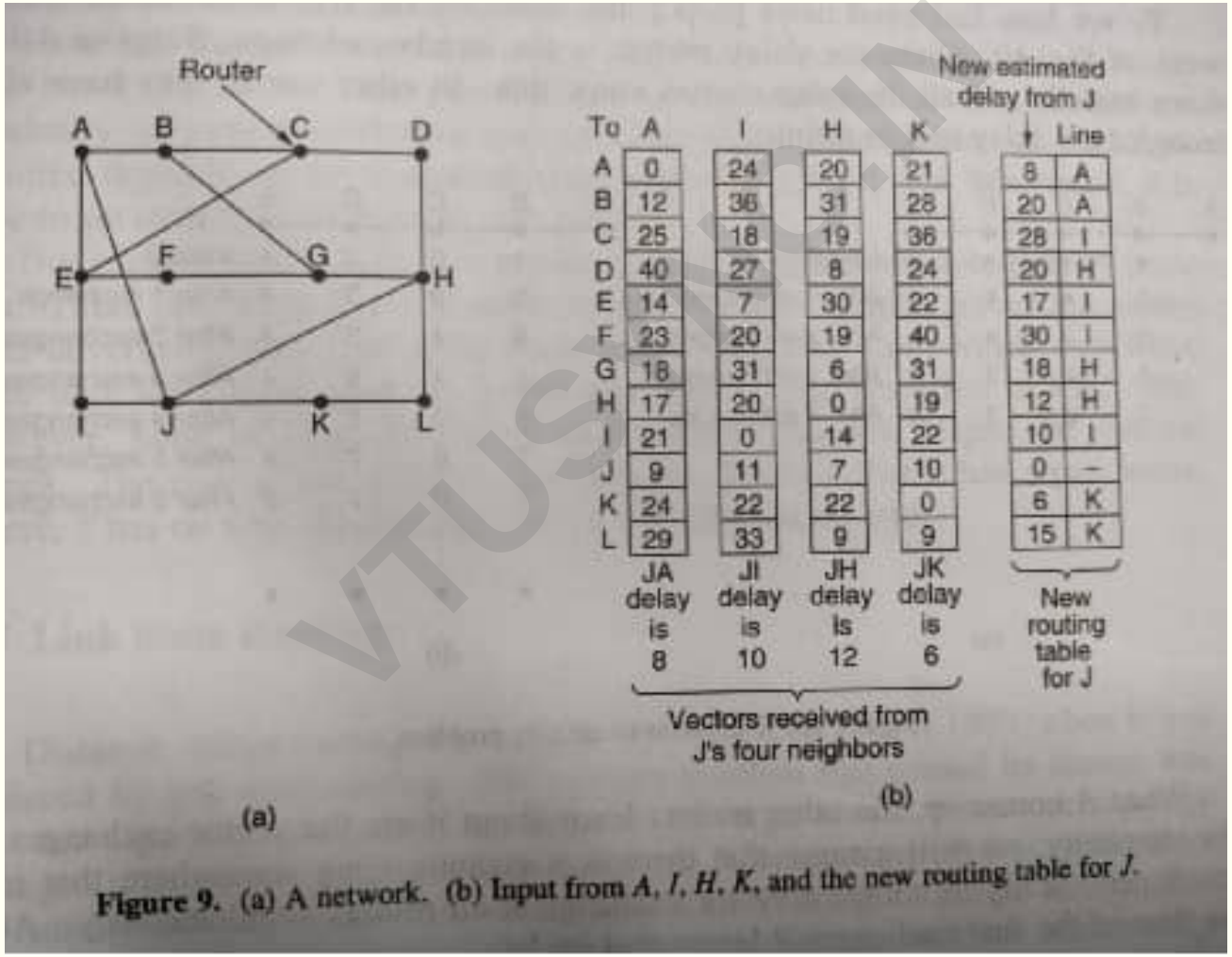
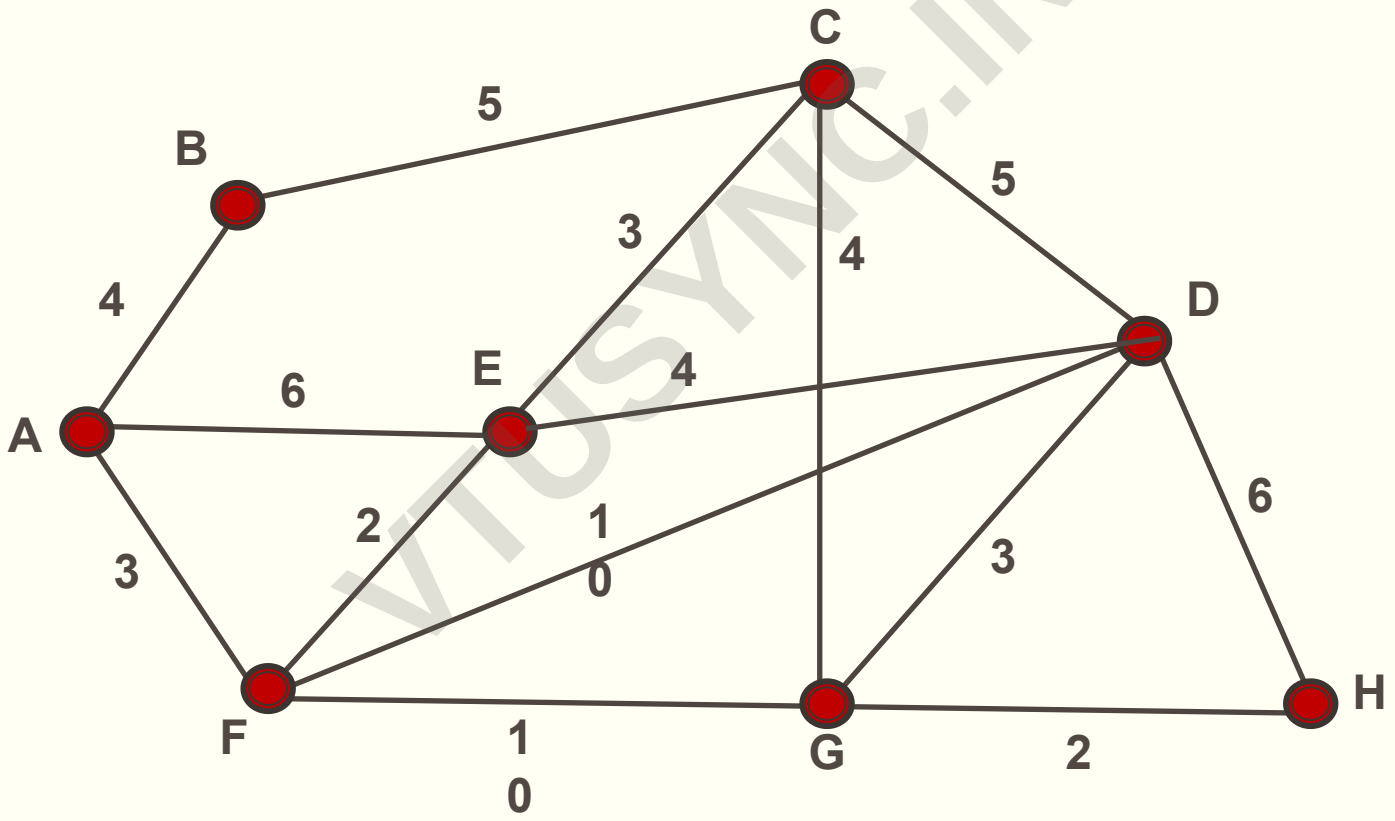


Figure 9. (a) A network. (b) Input from *A*, *I*, *H*, *K*, and the new routing table for *J*.

Example 2 - DVR algorithm – Router E updating it's Table



Existing Distance Vector of Router E

Destination	Outgoing line	Delay (ms)
A	F	5
B	C	8
C	C	3
D	D	4
E	-	0
F	F	2
G	D	7
H	D	9

Distance Vectors received by router E & Measurements

Destination	Delay from A	Delay from C	Delay from D	Delay from F
A	0	9	9	3
B	4	5	14	7
C	8	0	5	5
D	9	5	0	6
E	5	3	4	2
F	3	5	6	0
G	12	4	3	9
H	14	6	5	11

5	3	5	3

Delay from E

Updated Distance Vector of router E

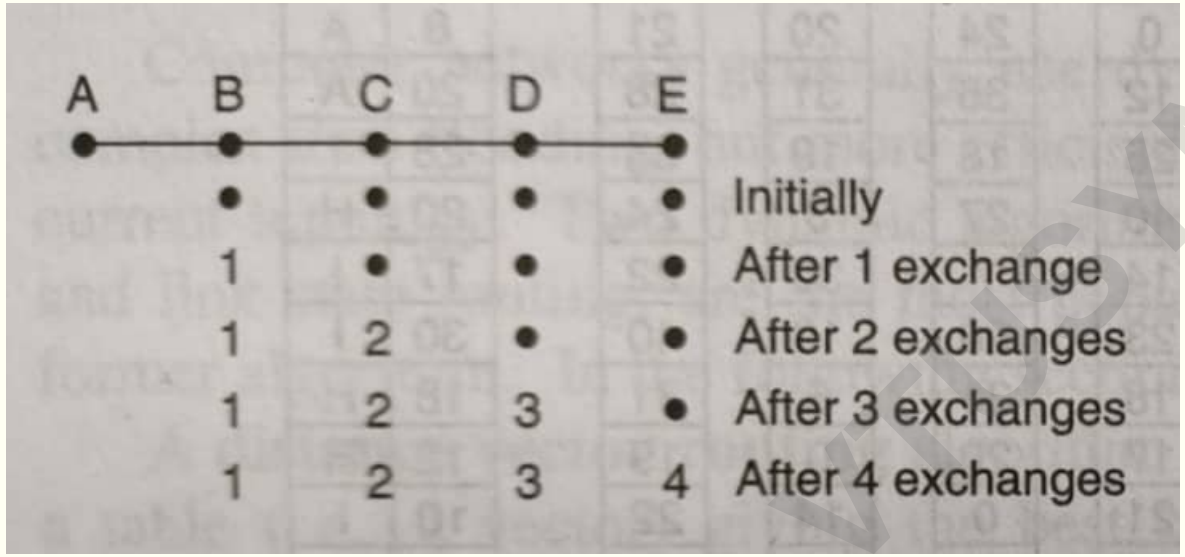
- The entry for Self (E to E) will not be determined by the updating method shown in the previous slide.
- It will always be “0”.

Destination	Outgoing line	Delay (ms)
A	A	5
B	C	8
C	C	3
D	D	5
E	-	0
F	F	3
G	C	7
H	C	9

Count-to-infinity problem

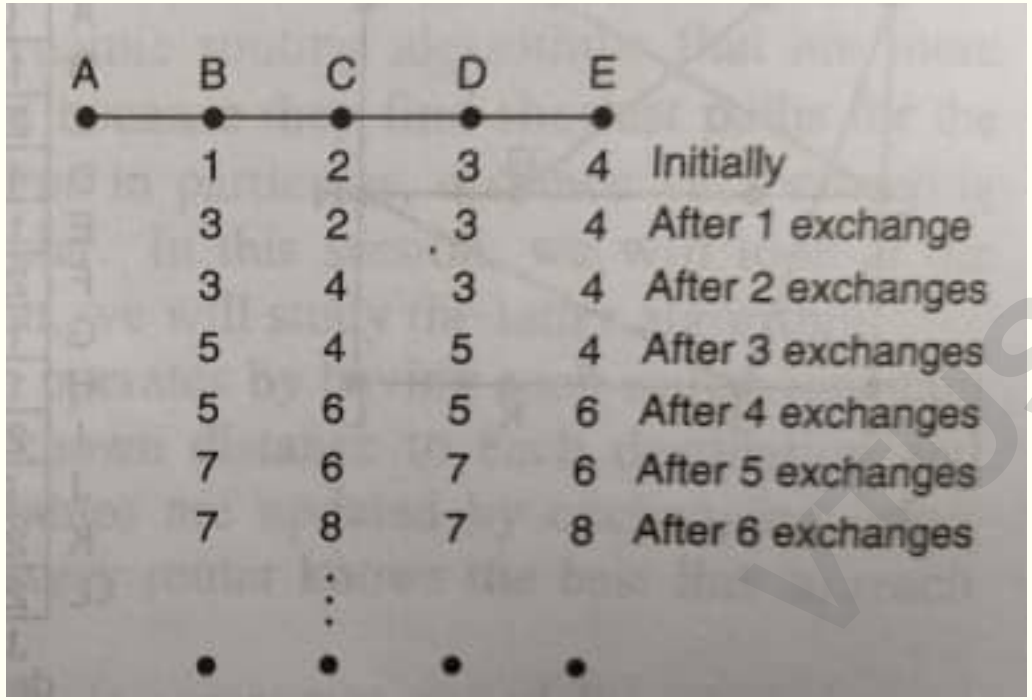
- One of the main drawbacks of the Distance Vector Routing algorithm is that while “good news” travels fast, “bad news” take a long time to reach all the routers in the subnet.
  - The latter part is called “Count-to-infinity” problem.
- “Good news” is, for example, a link that was down earlier, coming back into service.
- “Bad news” is, for example, a link going down.

Count-to-infinity problem – “Good news travels fast”



- “A” was initially down and all the other nodes recorded it as “∞” (a “.” in the figure).
- During the first exchange of Tables, after A comes back, B will know the fact and mention that A is just one hop away.
- On subsequent exchanges, the other too learn about it.
- If there are N nodes in a linear stretch, they will all know about it after (N-1) exchanges.

Count-to-infinity problem – "Bad news takes a long time"



- "A" was initially working and all the other nodes recorded the respective number of hops to A.
- "A" suddenly goes down.
- During the first exchange of Tables, B realises that A is not directly reachable.
- However, C tells B that it can reach A in 2 hops; So, B concludes that it can reach A in 3 hops (one hop to C and 2 hops from C to A).
- On the next exchange, C will update its Table to say that it can reach A in 4 hops (one hop to B and another 3 hops from B to A).

Count-to-infinity problem – “Bad news takes a long time”

- This progresses very slowly as shown in the figure.
- It takes a long time to reach Infinity (generally defined as 16 hops).
- A few approaches have been proposed to overcome this, but in practice, none of them are effective.
  - One such approach was to prevent routers from advertising their best paths back to their neighbours from which they heard about them.

Link State routing algorithm

- Distance vector routing had a major drawback in that it took too long to converge especially to ‘bad’ news (like a link going down).
- Link State routing has the following 5 steps:
  1. Discover the neighbours and learn their addresses
  2. Measure the delay or cost metric to each of the neighbours
  3. Construct a packet holding all that it has just learned
  4. Send this packet to, and receive packets from, all other routers
  5. Compute the shortest path to every other router

Link State routing algorithm

- Building the Link State packets can either be done periodically or when some significant event occurs (such as a line or neighbour going down or coming back again).
- In effect, the complete topology is known to every router.
- Then, Dijkstra's algorithm can be run at each router to find the shortest path to every other router.
- Steps 1 and 2 can be completed by sending an ECHO packet to the neigbours.

Link State routing algorithm

Step 3. Construct a packet holding all that it has just learned

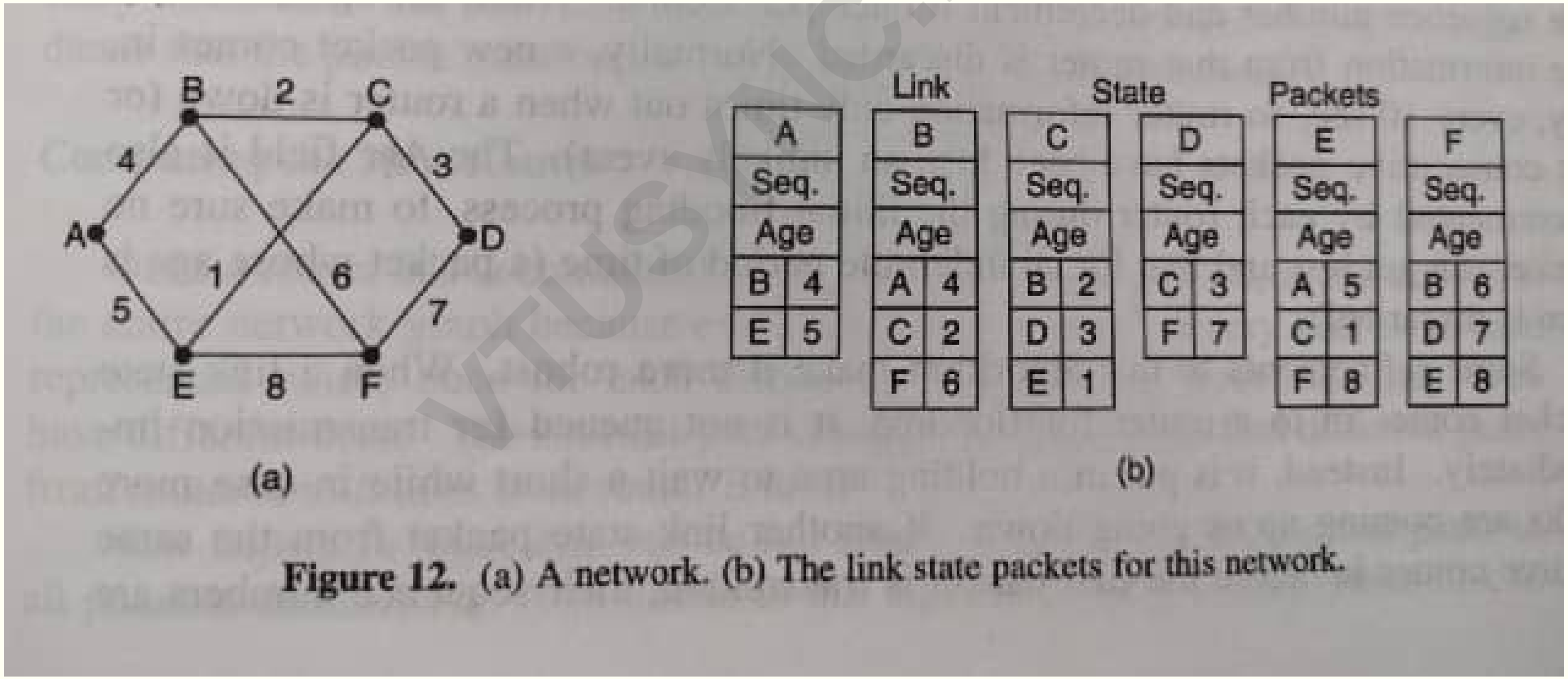


Figure 12. (a) A network. (b) The link state packets for this network.

Link State routing algorithm

Step 4. Distribution of Link State Packets.

- Flooding is used to distribute the packets.
- To keep the flood in check, Sequence Number and Age fields are used.
- For every new Link State Packet sent by a source router the Sequence number is incremented by one.
  - If a packet with a sequence number lower than the highest one seen so far arrives, it is discarded as obsolete, as the router has more recent data.

Link State routing algorithm

- The Age field is a preset number when the packet is formed.
- Each router decrements it by one before flooding it further.
- When the Age becomes zero, it is discarded.
- The idea is to make sure that a packet doesn't live on for a very long time.

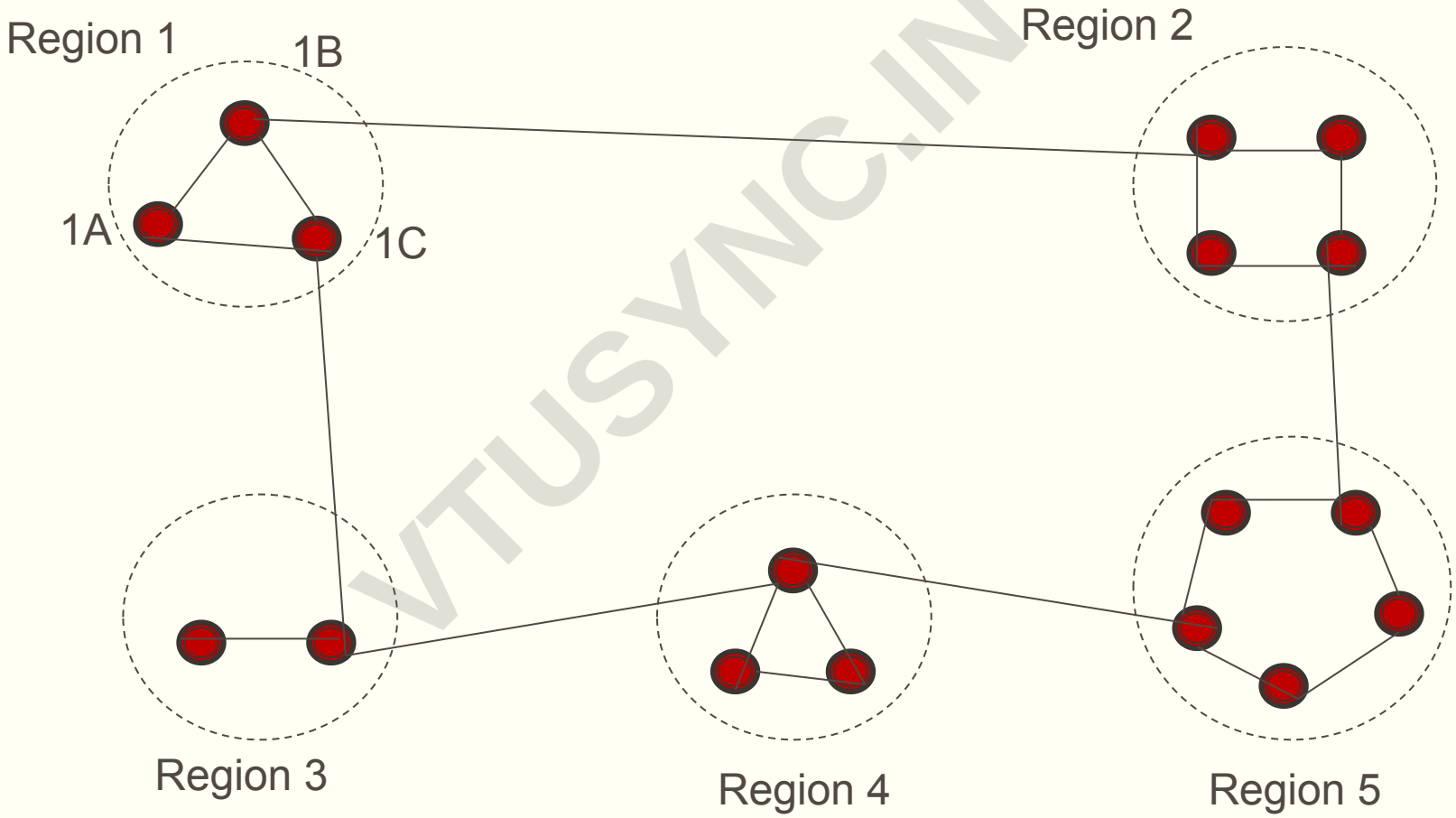
Step 5. Computation of the new shortest paths

- This is done by using Dijkstra's algorithm.
- OSPF (Open Shortest Path First) is a very popular Link State Protocol.

Hierarchical routing

- As networks grow, the router routing table grows proportionally.
- This would consume more memory and more search and updation time.
- In a hierarchical network, routers are organised into **regions**, with each router only knowing all the details of how to route packets to destinations within its own region.
- To route packets to destinations in other regions, they send the packets to a designated router within their own region.
- Regions may be further combined into **clusters**, and so on.

Hierarchical routing



Hierarchical Routing Table for 1A

Destination	Outgoing line	Hops
1A	-	-
1B	1B	1
1C	1C	1
2	1B	2
3	1C	2
4	1C	3
5	1C	4

Hierarchical routing - Examples

- With an eye on minimizing the length of the Routing Table, how do you group 120 routers into:
  1. Regions (2-level hierarchy)?
  2. Regions and Clusters (3-level hierarchy)?
- What will be the length of the Routing Table in both cases?
- Answers:
  - 12 regions of 10 routers each; Table of each router will have 10 entries for each router in its own region, and 11 entries for each of the other 11 regions. **A total of 21 entries**.
  - 4 Clusters each having 5 Regions, with each region having 6 routers each; Table of each router will have 6 entries for each router in its own region, 4 entries for each of the other 4 regions in its own cluster, and 3 entries for each of the other 3 clusters. **A total of 13 entries.**

Routing Information Protocol (RIP)

- RIP is a variation of Distance Vector Routing (DVR) algorithm, where in the "distance" metric is always the number of hops.
- The maximum hop count allowed in the RIP is 15.
- It is mostly used in small to medium-sized networks.
- Routing Tables are broadcast to all the neighbouring routers every 30 seconds.
- That information is used by every router to update it's own routing table.

Border Gateway Protocol (BGP)

- BGP (Border Gateway Protocol) is the protocol that enables the global routing system of the internet.
- It manages how packets get routed from network to network by exchanging routing and reachability information among “edge routers”.
- An edge router is a specialized router located at a network boundary that enables an internal network to connect with another network.
- BGP enables peering to send packets between autonomous systems (ASes), which are networks managed by a single enterprise or service provider. Together, these ASes make up the public internet.

Border Gateway Protocol (BGP)

- BGP is an exterior gateway protocol, which means it is designed to share routing information between different ASes.
  - Alternatively, an interior gateway protocol sends information within a single AS.
- BGP helps provide redundancy by enabling routers to quickly adapt and send packets through another connection if one internet path goes down.
- It is often used in large networks, such as internet service provider networks, wide area networks and infrastructure-as-a-service (IaaS) environments.

Multicast routing

- Multicast is a method of group communication where the sender sends data to multiple receivers or nodes present in the network simultaneously.
- Multicasting uses spanning trees to deliver the packets to members of a particular multicast group.
- Link State routing helps construct these spanning trees since each router knows the entire topology, and which router belongs to which multicast group.

Multicast routing - Example

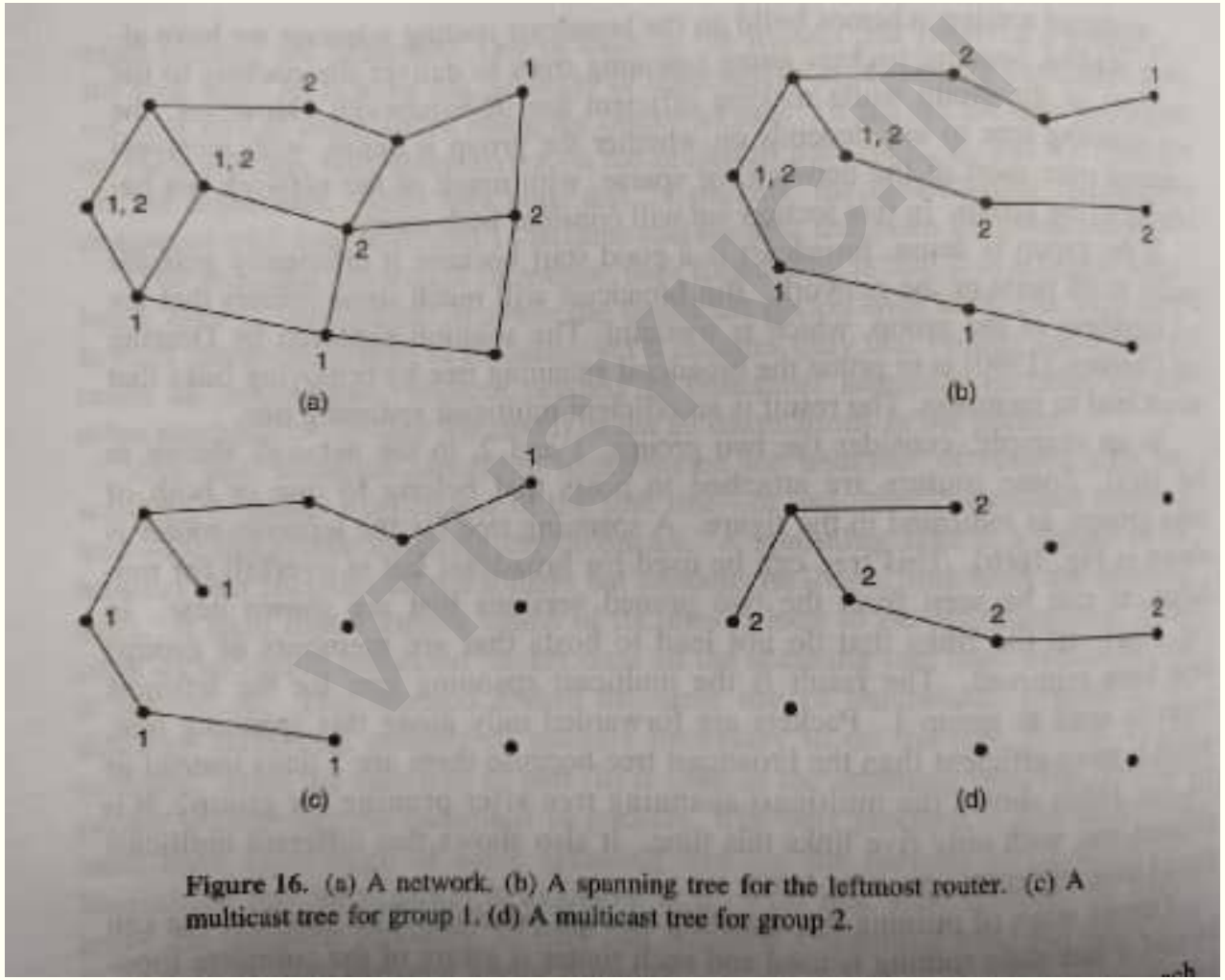


Figure 16. (a) A network. (b) A spanning tree for the leftmost router. (c) A multicast tree for group 1. (d) A multicast tree for group 2.

Multicast OSPF (MOSPF) routing

- It is an Interior Gateway Protocol (IGP) specifically designed to distribute unicast topology information among routers belonging to a single Autonomous System.
- OSPF is based on link-state routing algorithm.
- MOSPF routers are required to implement a "local group database" which maintains a list of directly attached group members.
- One of the routers is designated as a Designated Router (DR) and another one as a Backup Designated Router (BDR).
- The DR is responsible for communicating group membership information to all other routers in the OSPF area by flooding Group-Membership information.