

# Routing Algorithms and Protocols

Lectures-32-34: Introductory Concepts, select Routing Algorithms

**Dr. Rahul Banerjee**

*Associate Professor, Computer Science Group*

**Birla Institute of Technology and Science, Pilani - 333 031, INDIA**

*E-mail: [rahul@bits-pilani.ac.in](mailto:rahul@bits-pilani.ac.in) / [Rahul.Banerjee.CSE@Gmail.com](mailto:Rahul.Banerjee.CSE@Gmail.com)*

*Home Page: <http://www.bits-pilani.ac.in/~rahul/>*

# Interaction Goals

- Routing: The First Principles
- Broad Types of Routing:
  - Static & Dynamic
  - Unicast, Multicast, Anycast & Broadcast
  - Delay-sensitive & Delay-Tolerant
  - Energy-sensitive & Energy-Tolerant
  - Interior (Intra-Domain) & Exterior (Inter-Domain)
  - Frame, Packet & Message
- Introduction to the Unicast Routing
  - Objectives, Issues, Problems & Solutions
- Static Unicast Routing Algorithms and Protocols
- Dynamic Unicast Routing Algorithms and Protocols
- Recommended Readings
- Summary

# Packet Routing Problems

- Loss of packets
- Receipt and circulation of duplicate packets
- Packet Choking / Network Congestion
- Network Cleansing
- Worst-case upper bound problem
- QoS negotiation
- Failure Handling
- Quick Recovery Requirement
- Route Tracing
- Network Management Support

# Static Intra-Domain Routing Algorithms

# Static Packet Routing Schemes

## Shortest Path Routing:

- This is one of the simplest routing schemes and the primary technique involved here is the determination of the shortest available path between a source and a destination.
- The term shortest path may be interpreted in a variety of ways including:
  - path of the least geographical distance
  - path of the least congestion
  - path of the least number of Hops
  - path of the least mean queuing delay
  - path of the least propagation / transmission delay

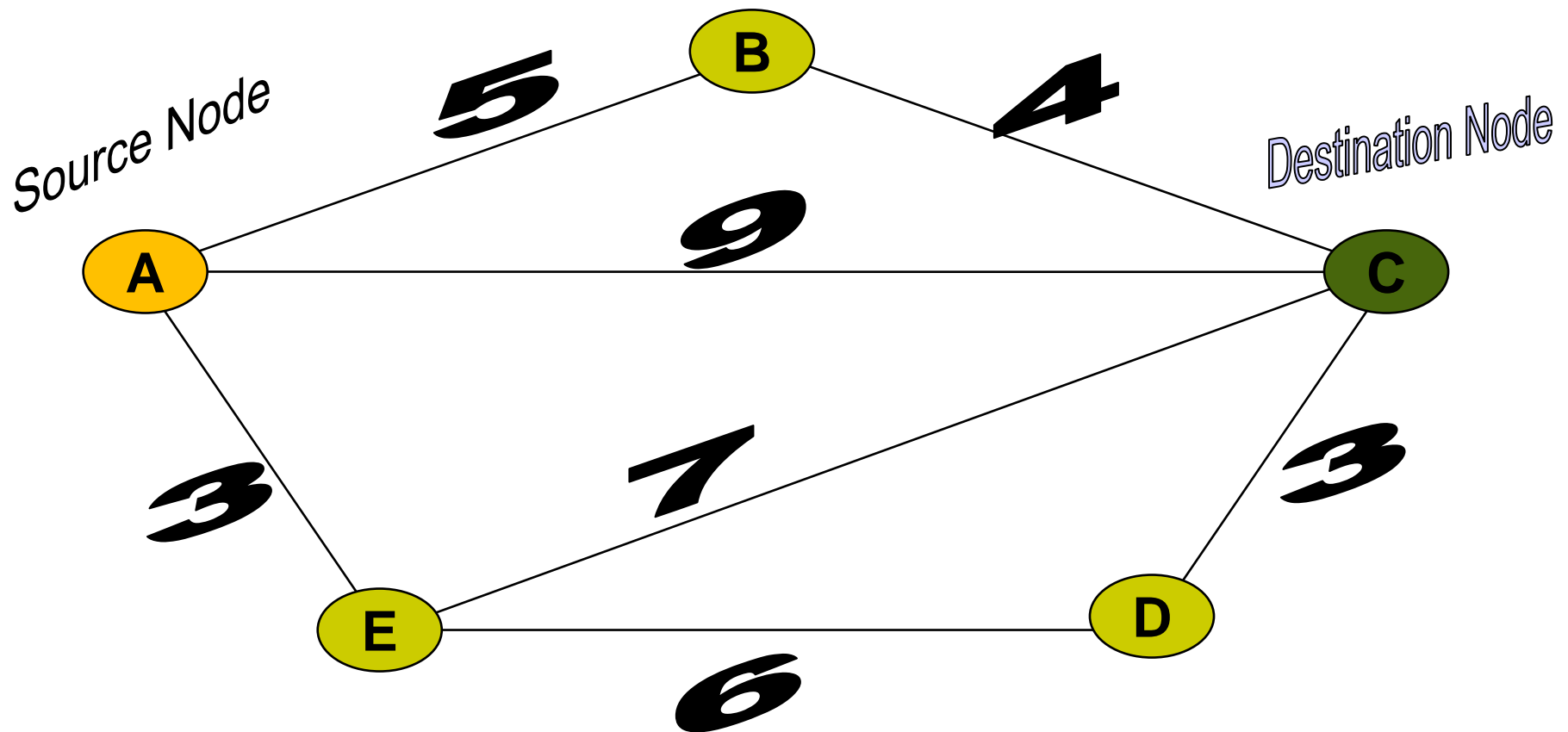
Any weighted average based metric can be yet another choice for employing this scheme.

# Dijkstra's Shortest Path Routing Algorithm

- One of the best known algorithms that may be employed for determination of the shortest path is the one suggested by Dr. E. W. Dijkstra in as early as 1959.
- The gist of this strategy is given below:
- 1. Each node is labelled with the name of the source node and its distance from the current node.
  - Normally, the labelling is done in the reverse order, i.e. the label (9, A) represents distance of the current node from the source node (9) followed by the name of the source node (A).
  - A label may be permanent or tentative.

# Dijkstra's Algorithm ...

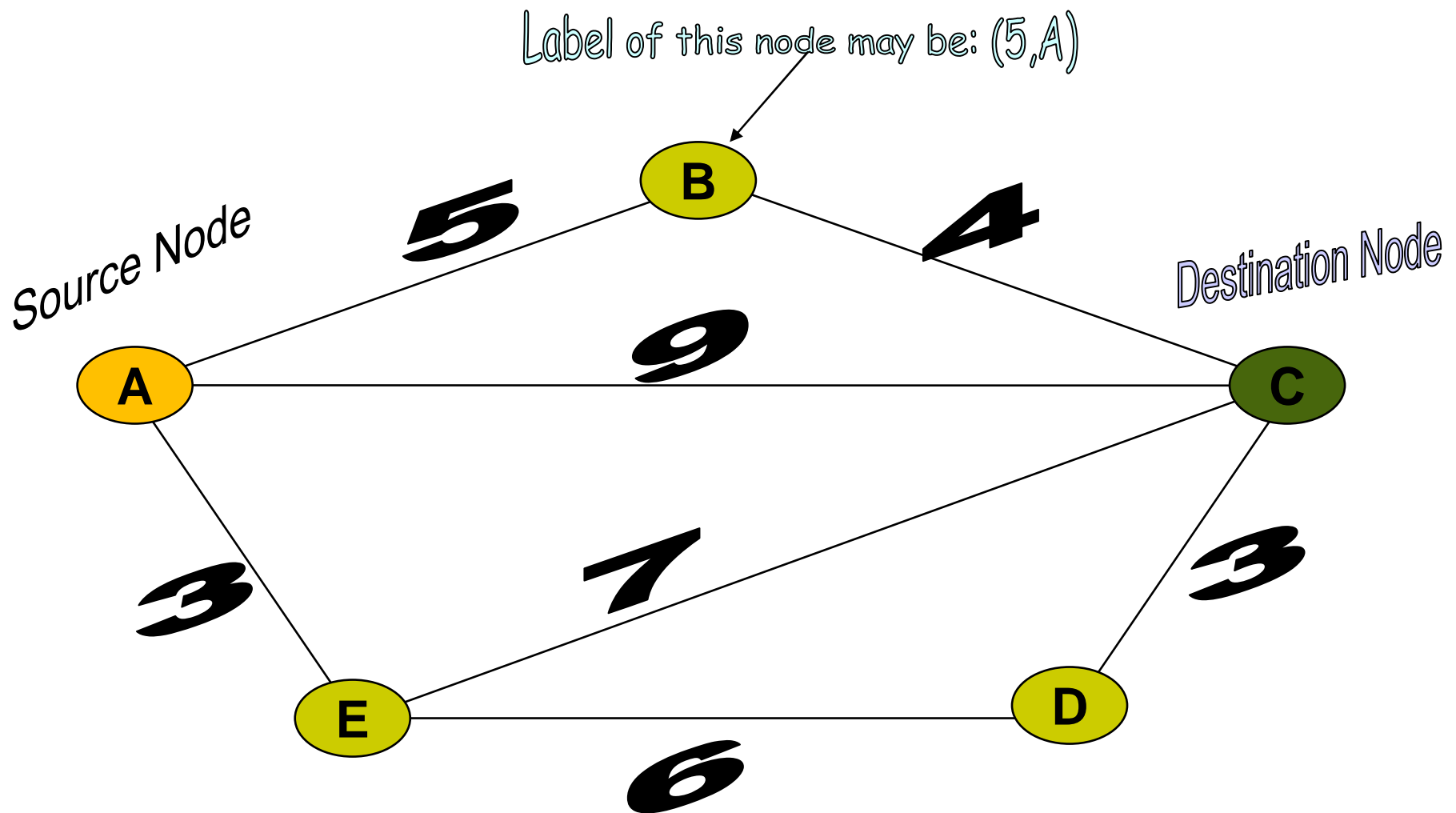
Initially, all nodes, except A, shall be tentatively labeled as (Infinity, A)



# Dijkstra's Shortest Path Routing Algorithm

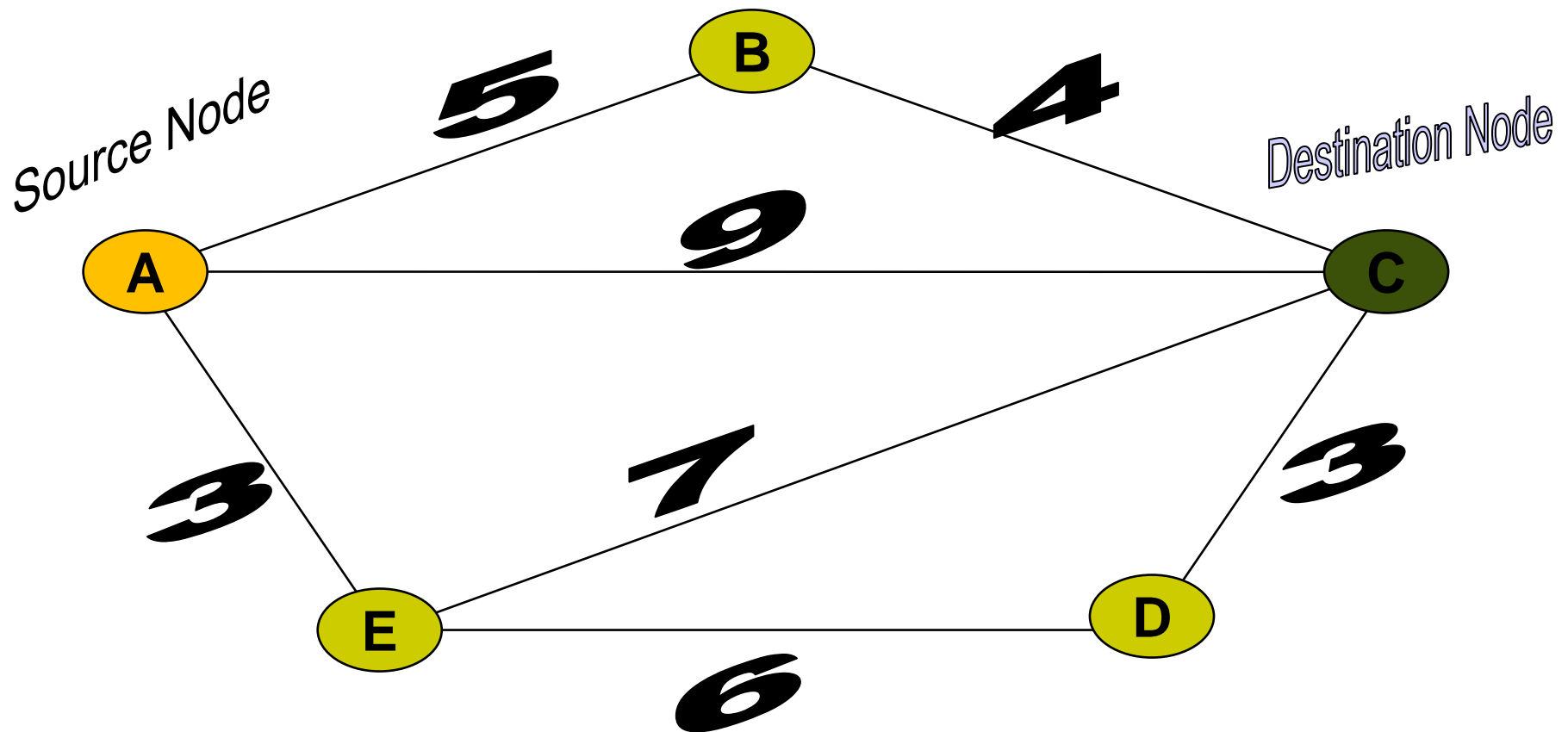
2. At the start of the algorithm all nodes are labelled tentatively.
3. As the algorithm progresses, the labels may change.
4. At any stage, when it becomes clear that the current label represents the smallest distance / shortest path between a node and the source node, former's label is marked as a permanent label.
5. As the algorithm progresses, more and more nodes acquire permanent labels.
6. The algorithm terminates when the destination node gets a permanent label.

# Dijkstra's Algorithm ...



# Dijkstra's Algorithm ...

Finally, each node shall have a permanent label)



# Flooding-based Routing Schemes

- There exist three major variants:
  - Pure / Unconstrained Flooding
  - Hop-Count Based / Constrained Flooding
  - Selective / Direction-Constrained Flooding

Each of these types finds a brief description in the following slides.

# Hop-Count-based Flooding Algorithm

- This algorithm may be expressed as follows:
  1. At any originating node 's',
    - structure a packet such that its header contains a 'hop count' that
      - be initialized to length of the path (if known) or
      - full diameter of the subnet.

## Hop-Count-based Flooding Algorithm ...

2. At every intermediate node 'i' examine the incoming queue of packets, take the packet at the head of the queue and note the packet-id, line on which it arrived on, its hop count and destination address.
3. Decrement the hop count by one (1).
4. If the count becomes zero, discard / drop the packet and flush the corresponding entries in the local table. Otherwise, generate (n-1) replicas of the packet (where 'n' is number of arcs converging at this node) and transmit one replica on all arcs / lines except the one this packet arrived on.
5. Examine the incoming queue and
  - if it is non-empty,
    - repeat steps 2 to 5
  - else wait until a new packet arrives and
    - then repeat steps 2 to 5.

# Selective / Direction-Constrained Flooding Algorithm

1. It is a variant of the basic Flooding Algorithm with the constraint of direction thrown in for the purpose of improved efficiency.
2. In this scheme, packets are selectively flooded by the routers in such a way that they move approximately in the right direction (i.e. leading towards the Destination).

# Flow-based Routing Algorithm

- This is yet another Static Routing Algorithm; but unlike the Shortest Path based Routing Algorithm and the Flooding based Routing Algorithm, which primarily consider the Subnet Topology alone, it considers Subnet Topology as well as Load (Traffic).
- This is particularly suitable for the subnets characterized by nearly stable average data transfer rate / mean data flow rate.
- In other words, this scheme may not prove to be effective if the mean inter-node data flow in a given subnet cannot be reliably predicted / estimated.

# Flow-based Routing Algorithm ...

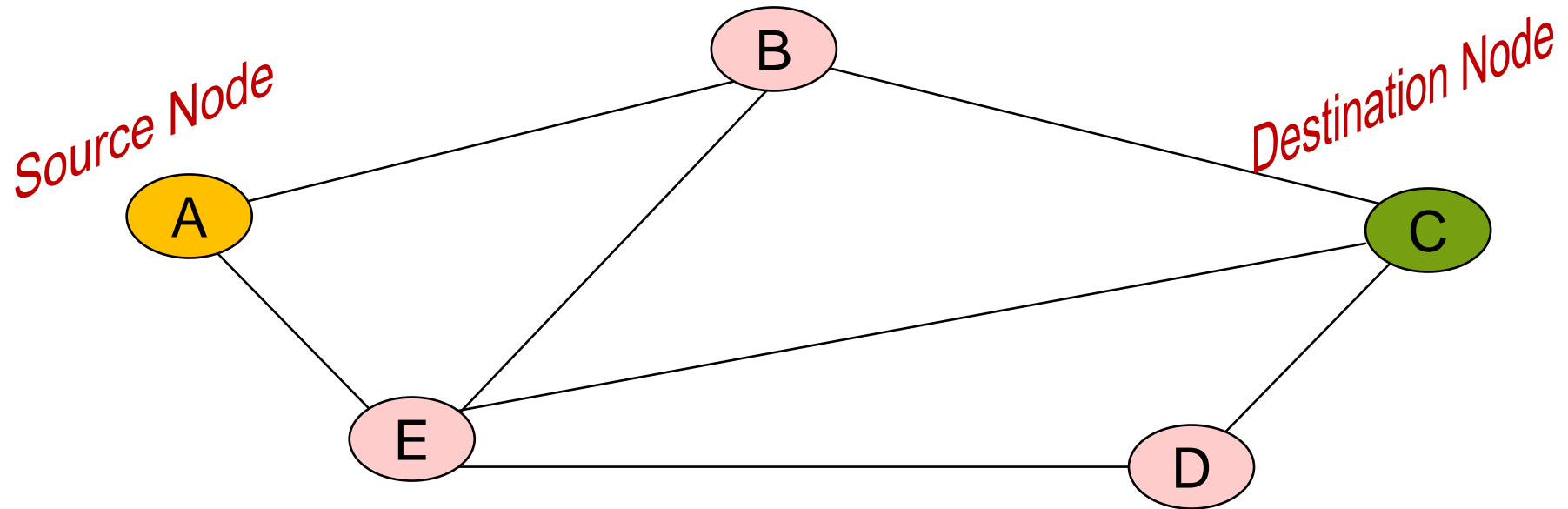
- This algorithm, unlike the other algorithms discussed so far, has several pre-requisites including the following:
  - Subnet topology must be known in advance.
  - Link / Line Capacity Matrix must be known in advance.
  - Traffic Matrix must be available a priori.
  - Mean packet-size must be known.
  - Some preliminary Routing Algorithm must be available.

# Flow-based Routing Algorithm ...

- The scheme makes use of the fact that under the above stated circumstances, for each of the links, if the link-capacity, average rate of data-flow and topology are known and if the traffic-matrix and subnet topology is available in advance, then it is possible to:
  1. Compute the mean delay in packet-delivery per link,
  2. Compute the mean (overall) delay in packet-delivery over the given subnet,
  3. Compute the most appropriate route between any pair of Source and Destination .

# Another example subnet

Link-Id	Link-Traffic In packets per second	Link-Speed	Link-Capacity In packets per second	Mean Delay on the Link in ms	Link-Weight
AB	12 pps	400 kbps	... pps	... ms	...
..	..	..	..	..	..



# Dynamic Intra-Domain Routing Algorithms

# The Distance Vector Routing Algorithm

- This is also known as the Bellman-Ford / Ford-Fulkerson Routing Algorithm.
- It is the original Dynamic Routing Algorithm used in the erstwhile ARPANET.
- For quite some time, it was popular over the Internet where a variant of it called Routing Internet Protocol (RIP) was used.
- One of the reasons of its popularity was that an RIP implementation was distributed free with the BSD version of the UNIX later known as BSD
- Many Routers still use one or other variation of this algorithm.

# The Distance Vector Routing Algorithm ...

- In brief, this scheme may be expressed as:
  - Each Router knows / discovers its distance from its neighbours.
  - Each Router locally maintains a Routing Table indexed by an entry for every other Router in the subnet and identification of a preferred neighbour / link leading to that Router.
  - Metric of estimation may vary. For instance, it may be any one of Physical Distance, Hops, Delay etc.

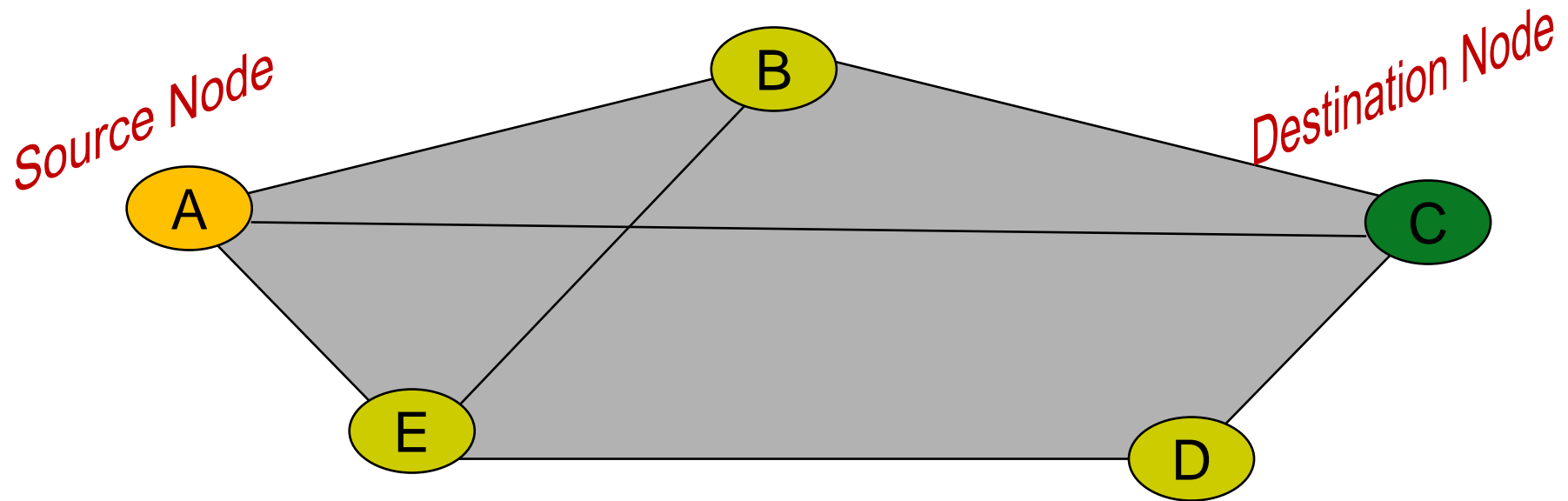
# The Distance Vector Routing Algorithm ...

- Periodically, each Router sends a single-dimensional array of estimated distances called *Distance Vector* to its neighbouring Routers (*directly connected routers are called neighbours*).
- On receipt of such Distance Vectors from its neighbours, every Router re-computes its estimates and updates its local routing table.

*Figure on the following slide presents an example subnet and a sample distance vector generated at one of its nodes.*

# An Sample Routing Subnet

Let a routing subnet look like:



# The DVR Process

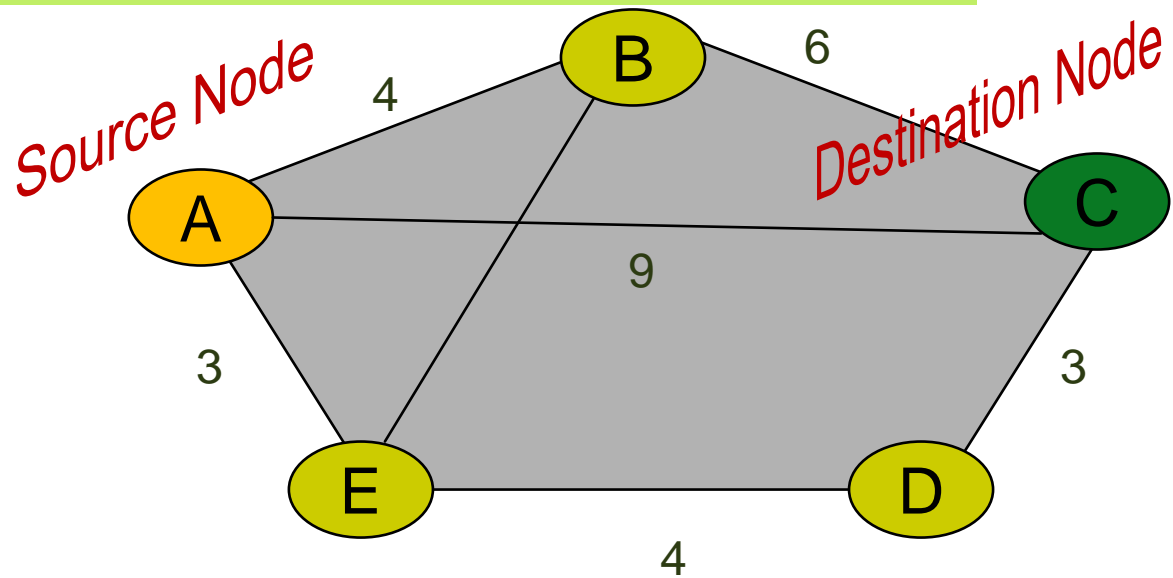
A typical Forwarding Table for a Router 'A' may look like (initially)

Initial Estimated Distance from 'A'	Via Routing Node / Router
0	A
4	B
9	C
-	D
3	E

•This is how a **Distance Vector** generated at A could look like!  
 •Here, the first column indicates Current Estimates and the second column refers to Identification Symbol / Address of the corresponding Router.

4	B	0	A
9	C	4	B
3	E	9	C
		-	D
		3	E

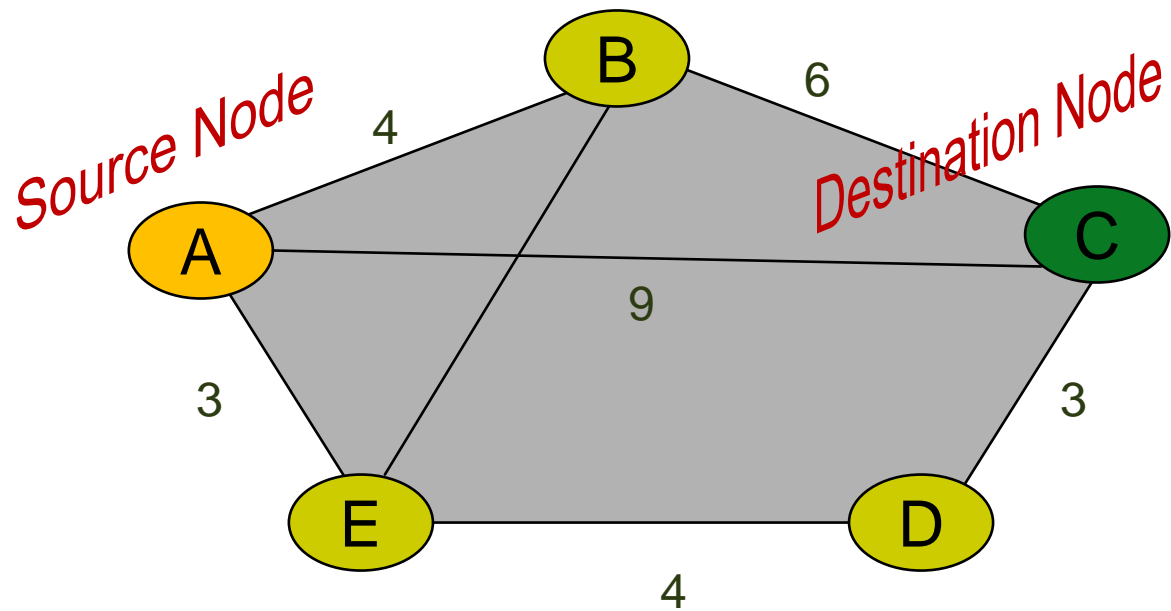
•Trouble in such a DV format is that it makes it impossible to learn gradually about the rest of the topology or later changes in topology.  
 •**Therefore, the DVs are formed as shown on the right side.**  
 •However, there is a catch here! Routers need to know about all routers which may form the subnet. Different methods are employable for this in different situations.  
 •'-' is an expression of 'Infinity'; often taken as 16 in RIP versions, which uses a form of DVR considering 'hop' as its metric..



# The DVR Process ...

After a round of exchange of Distance Vectors with its neighbours, Table at 'A' might look like:

New Estimated Distance from 'A'	Via Routing Node / Router
0	A
4	B
9	C
7	D
3	E



# Issues with the Distance Vector Routing

- The primary drawback of this algorithm is its vulnerability to the 'Count-to-Infinity' problem. There have been proposed many partial solutions but none works under all circumstances.
- Another drawback of this scheme is that it does not take into account Link Bandwidth.
- Yet another problem with this algorithm is that it takes appreciably long time for convergence as the network-size grows.
- A fallout of the Count-to-Infinity issue and slow convergence has been to limit the maximum number of hops to 15 which means more than 16-router subnets, it may not be appropriate routing algorithm.
- However, it is one of the simplest dynamic algorithms in use. 26

# The Link-State Routing Algorithm

In this algorithm, each router:

- Discovers its neighbours and their Network Addresses by sending special packets called 'Hello' packets.
- Estimates delay / cost or any other metric for reaching its neighbours by sending another special packet type called 'Echo' packets.
- Immediately applies its recent knowledge to form Link-state packets which encapsulate this estimate; and, sends copies to all the discovered neighbouring routers.
- Computes the shortest path to every other router using the Shortest Path Algorithm.

# Link-State Routing Algorithm ..

- In this case, fresh link-state packets are built:
  - **periodically** or
  - upon occurrence of an event like node-failure / link-failure / addition of a node or link / revival of a failed node or link.
- The algorithm requires that names / identifiers representing the routers be unique (globally).

# Link-State Routing ...

A typical Link-State Packet for a Router 'A' may look like:

<b>A</b>	
<b>1101..1100</b>	
<b>60</b>	
<b>B</b>	<b>7</b>
<b>E</b>	<b>5</b>

where,

- the first row indicates the Originating Router,
- the second row refers to the Sequence Number (usually a 64-bit or higher number) of the link-state packet,
- third row shows the Age of the packet, the fourth and subsequent rows indicate estimated metrics for each of the neighbouring routers (B and E in this case).

# Link-State Routing Algorithm ..

- Data Structure for the Packet Buffer at the Routers has the format:

Source	Sequence No.	Age	Send Flags	Acknowledgement Flags	Data
--------	--------------	-----	------------	-----------------------	------

- Examples of some of the well known implementations of this scheme include:
  - Open Shortest Path First (OSPF) scheme
  - Intermediate System- Intermediate System (IS-IS) scheme

# The Interior Gateway Routing Protocol (IGRP)

- The Interior Gateway Routing Protocol (IGRP) was originally developed in the mid-1980s by Cisco Systems. This protocol did not support VLSM scheme. Its successor, EIGRP, supports VLSM.
- **Basic objective of the IGRP was to provide a robust protocol for routing within what was called as an 'autonomous system' (AS).**

(An AS is a collection of networks under common administration that share a common routing strategy. Every AS is normally uniquely identified by a 16-bit number.)

# The Interior Gateway Routing Protocol (IGRP) ...

- The most commonly used AS-AS routing protocol prior to the advent of the IGRP was the Routing Information Protocol (RIP).
- As mentioned earlier, very small hop limit (only 16 hops) restricted the size of RIP based internetworks.
- Moreover, as pointed out in the slides related to the Dynamic Routing Algorithms, RIP proved sub-optimal and less flexible.

# The Exterior Gateway Routing Protocol (EGRP):

- **The Exterior Gateway Protocol (EGP) is an inter-domain connectivity / reachability protocol.**
- **The original version of the EGP that enjoyed quite a bit of popularity is gradually giving way to other competing exterior gateway routing protocols (like the BGP and the IDRP).**

(This is because of certain weaknesses that came to light with the exponential growth of the Internet over the years.)

# The Exterior Gateway Routing Protocol (EGRP) ...

- Currently, the most well known Exterior Gateway Routing Protocol is the Border Gateway Protocol Version 4.
- Incidentally, BGP4 also happens to be the first version that is capable of handling the CIDR and Supernetting.

# The Border Gateway Protocol (BGP) ...

- BGP uses the TCP as its transport protocol of choice. (**Advantage of this approach is that BGP can relieve itself of reliability specific concerns.**)
- BGP is a path vector protocol. (Since, the routing information used by the BGP consists of a vector of Autonomous System ID Nos., which actually maps to a traversed path / route, it is called as a path vector protocol.)
- It is primarily used for exchange of information between autonomous systems.

# The Classless Inter-Domain Routing (CIDR) in IPv4 Subnets:

## Primary Objective:

- Finding a temporary solution to the IPv4 Address space depletion

## The basic idea behind the CIDR:

- Allocate the unallocated set of Class-C IPv4 network addresses in variable-sized address blocks.
- These blocks, in effect, refer to contiguous Class-C IPv4 network addresses.

# The Classless Inter-Domain Routing (CIDR) ...

The RFC 1519 allocation rules for the IPv4 world:

1. The whole world was suggested to be divided into four zones each of which could use nearly 32 Million Addresses:

- Asia-Pacific:
- Central-Southern America
- Europe
- North America

# The Classless Inter-Domain Routing (CIDR) ...

## The RFC 1519 allocation rules for the IPv4 world ...

2. A set of nearly 320000000 addresses were suggested to set aside for future use.
3. If a router 'X' get a packet that belongs to the IPv4 addresses of one these four zones, the packet is simply forwarded to the zonal gateway.

# The Classless Inter-Domain Routing (CIDR) ...

## The Supernetting:

- Terms ‘Aggregation’, ‘CIDR Block allocation’, ‘Supernetting’ etc. are often used interchangeably in the IPv4-CIDR literature. (This is however, done in casual discussion alone!)
- In principle, ‘a network whose prefix-boundary has lesser number of bits than the natural mask of the network itself, is called a Supernet’.

Two ways to represent the same CIDR address are :

- 199.28.0.0/16
- 199.28.0.0 255.255.0.0

# Recommended Readings:

- S. Keshav: An Engineering Approach to Computer Networking, AWL, 1997.
- A. S. Tanenbaum: Computer Networks, Fourth Edition, PHI, 2006.
- C. Huitema: IPv6, Second Edition, Prentice-Hall PTR, 1998.
- U. D. Black: Computer Networks, Second Edition, PHI, 1993.
- D. Bertsekas and R. Gallager: Computer Networks, Second Edition, PHI, 1992.
- G. R. McClain (Ed.): Handbook of Networking and Connectivity, AP Professional (Academic Press), 1994.

# Recommended Readings ...

- RFC 1009 (Requirements for Internet Gateways)
- RFC 1254 (Gateway Congestion Control)
- RFC 1360 (Official Protocol Standards of the Internet Architecture Board)
- RFC 1124 (Policy Issues in Interconnecting Networks)
- RFC 1125 (Policy Requirements for Inter-Administrative Domain Routing)
- RFC 781 (IP Timestamp)
- RFC 791 (IP)
- RFC 815 (IP Datagram Reassembly)
- RFC 1042 (IP over IEEE 802.3)
- RFC 1011 (Official IP)

# Recommended Readings ...

- RFC 1883 (IPv6 Specification)
- RFC 1825 (IP Security Architecture)
- RFC 1826 (IP Authentication Header)
- RFC 1827 (IP Encapsulation Security Payload)
- RFC 1828 (IP Authentication using MD5)
- RFC 1175 (FYI : A very useful reference-list on Internetworking related information)
- RFC 1208 (Glossary of Networking Terms)
- Smoot Carl-Mitchell & John S. Quarterman: Practical Internetworking with TCP / IP and UNIX, Addison-Wesley, Reading, 1993. (This book does not really discuss the IPv6. This however, helps the reader to take a look at the pre-IPv6 days and realize the wisdom of evolution of the IP.)

# Recommended Readings

...

- Larry Hughes: Introduction to Data Communication: A Practical Approach, Narosa Publishers, 1997.
- Prakash C. Gupta: Data Communications, PHI, 1996.
- A. Shah: FDDI: A High Speed Network, PTR Prentice Hall, 1994.
- M. R. Tolhurst (Ed.): Open System Interconnection, Macmillan, 1988.
- William Stallings: Data and Computer Communications, Fifth Edition, PHI, 1998.

# Recommended Readings ...

- D. Comer: Internetworking with TCP / IP , Vol..-1, PHI, 1995.
- D. Comer & D. L. Stevens: Internetworking with TCP /IP, Vol.. 2-3, PHI,1994, 1993.
- W. Buchanan: Advanced Data Communication and Networks, Chapman & Hall, London, 1997.
- Uyles D. Black: TCP / IP & Related Protocols, Second Edition, McGraw-Hill, N. Y., 1995.
- RFC 1519 (CIDR)
- RFC 1997 (BGP community attribute)
- Bassam Hallabi: Internet Routing Architectures, Cisco Press, New Riders Publishing, 1997.
- RFC 904 (Exterior Gateway Protocol)