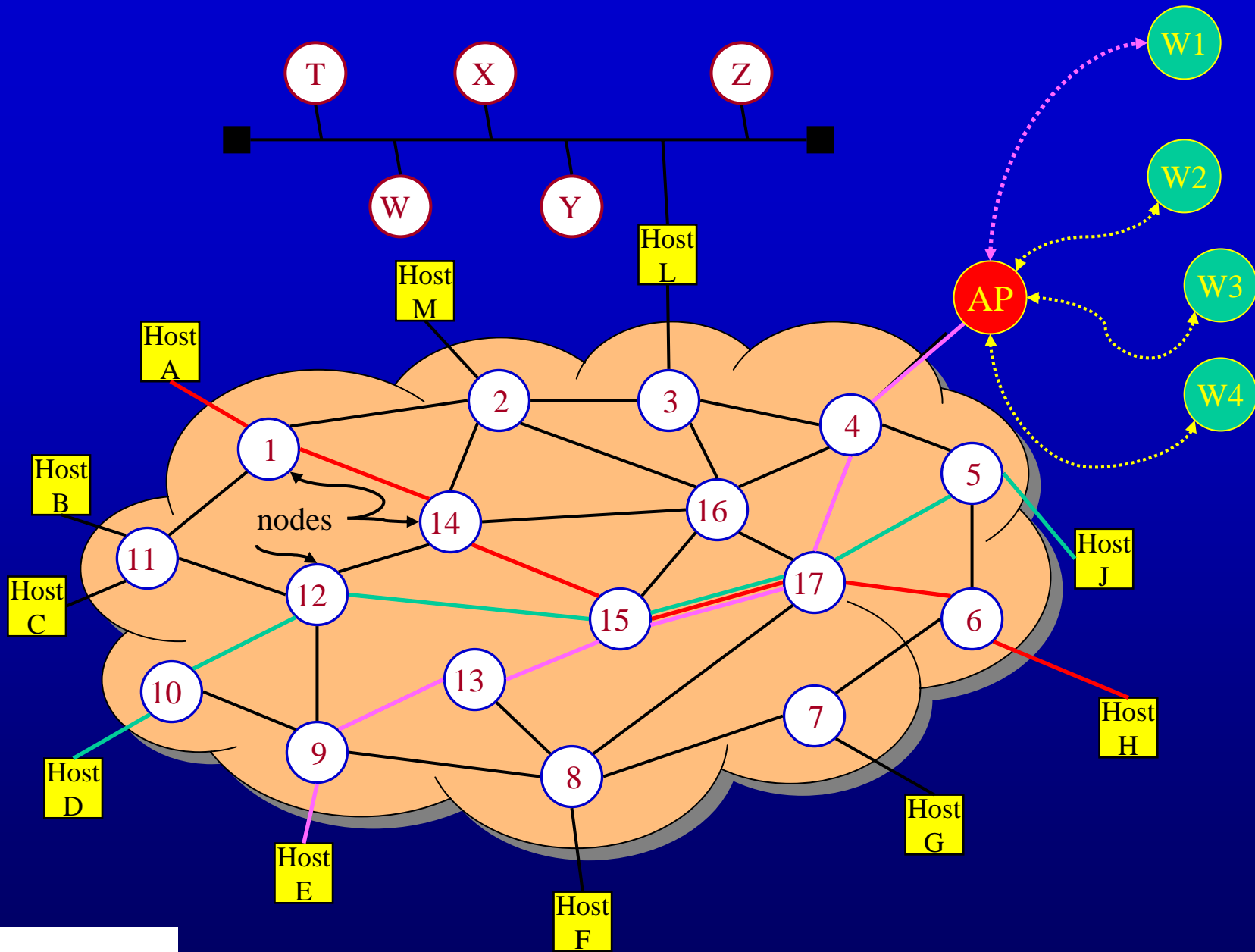
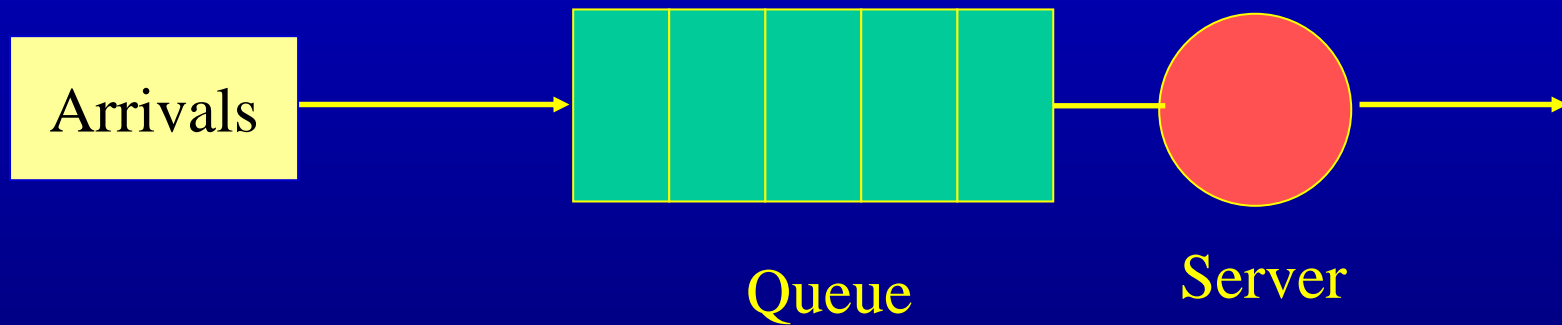


Computer Network Performance Measures

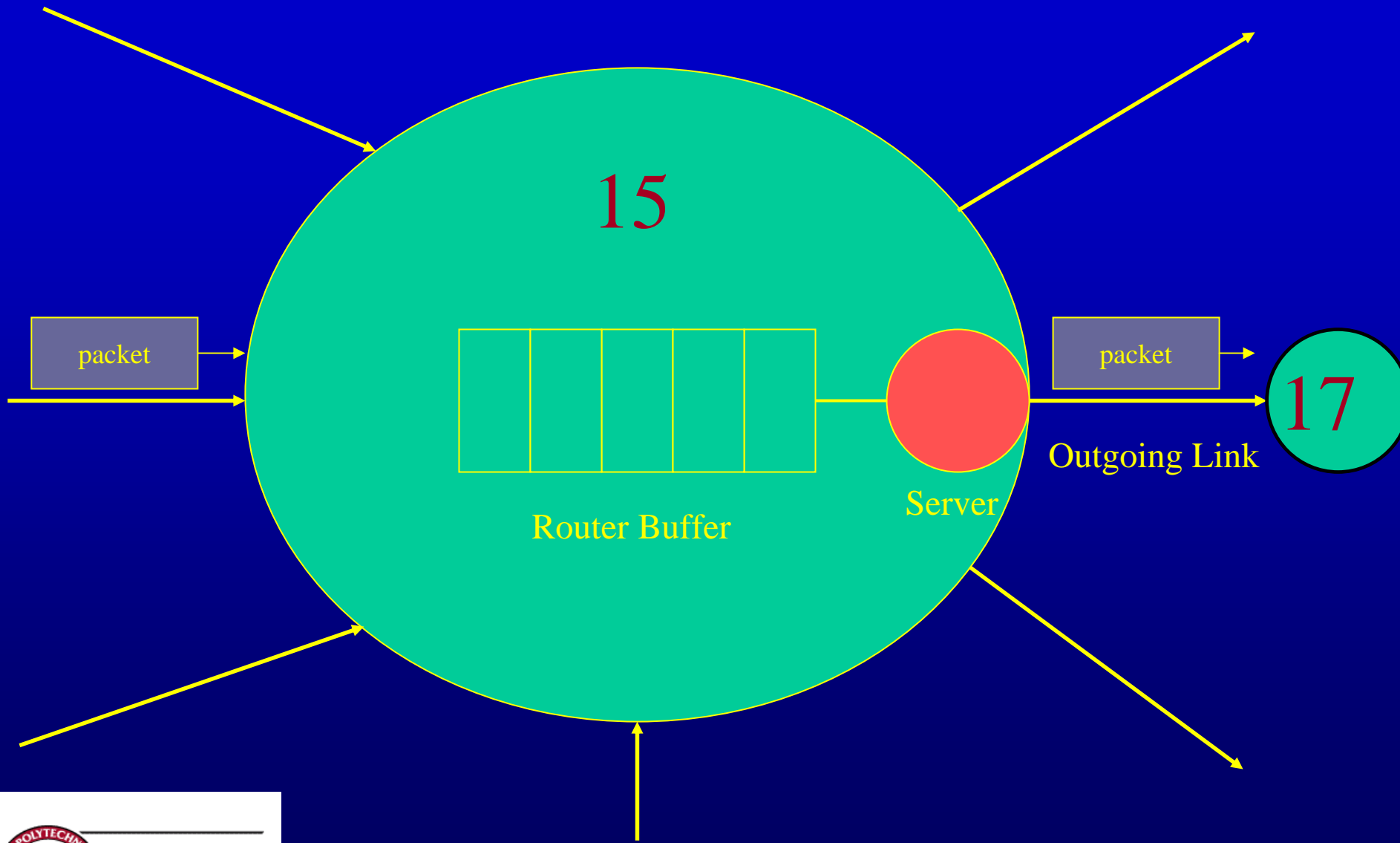




Simple Queuing Model



Router Node



Performance Metrics

(General Definitions)

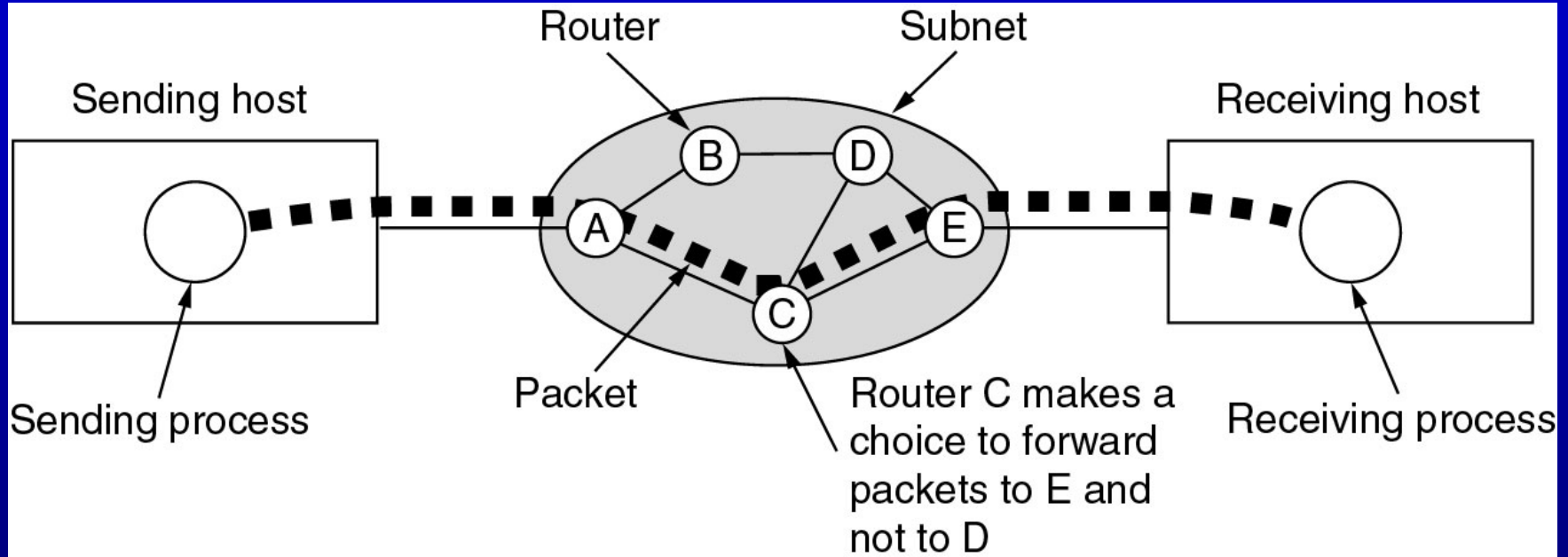
- **Utilization** :: the percentage of time a device is busy servicing a “customer”.
- **Throughput** :: the number of jobs processed by the “system” per unit time.
- **Response time** :: the time required to receive a response to a request (round-trip time).
- **Delay** :: the time to traverse from one end to the other in a system.



Network Performance Measures

- **Channel utilization** :: the average fraction of time a channel is busy [e.g. Util = 0.8]
 - when overhead is taken into account (i.e., excluded from *useful* bits, channel utilization is often referred to as channel efficiency)
- **Throughput** :: bits/sec. successfully transmitted [e.g. Tput = 10 Mbps]

End-to-end packet delay



End-to-end delay includes multiple hop link delays

Tanenbaum slide

End-to-end packet delay

End-to-end packet delay :: the time to deliver a packet from source to destination.

{most often we are interested in the packet delay within the *communications subnet*}

This delay is the sum of the delays on each subnet link traversed by the packet. Each link delay consists of four components

[B&G Bertsekas and Gallager]:



Link Packet Delay

1. The *processing delay* [**PROC**] between the time the packet is correctly received at the head node of the incoming link and the time the packet is assigned to an outgoing link queue for transmission.
2. The *queuing delay* [**QD**] between the time the packet is assigned to a queue for transmission and the time it starts being transmitted. During this time, the packet waits while other packets in the transmission queue are transmitted.



Packet Delay

3. The *transmission delay* **[TRANS]** between the times that the first and last bits of the packet are transmitted.
4. The *propagation delay* **[PROP]** between the time the last bit is transmitted at the head node of the link queue and the time the last bit is received at the next router. This is proportional to the physical distance between transmitter and receiver.



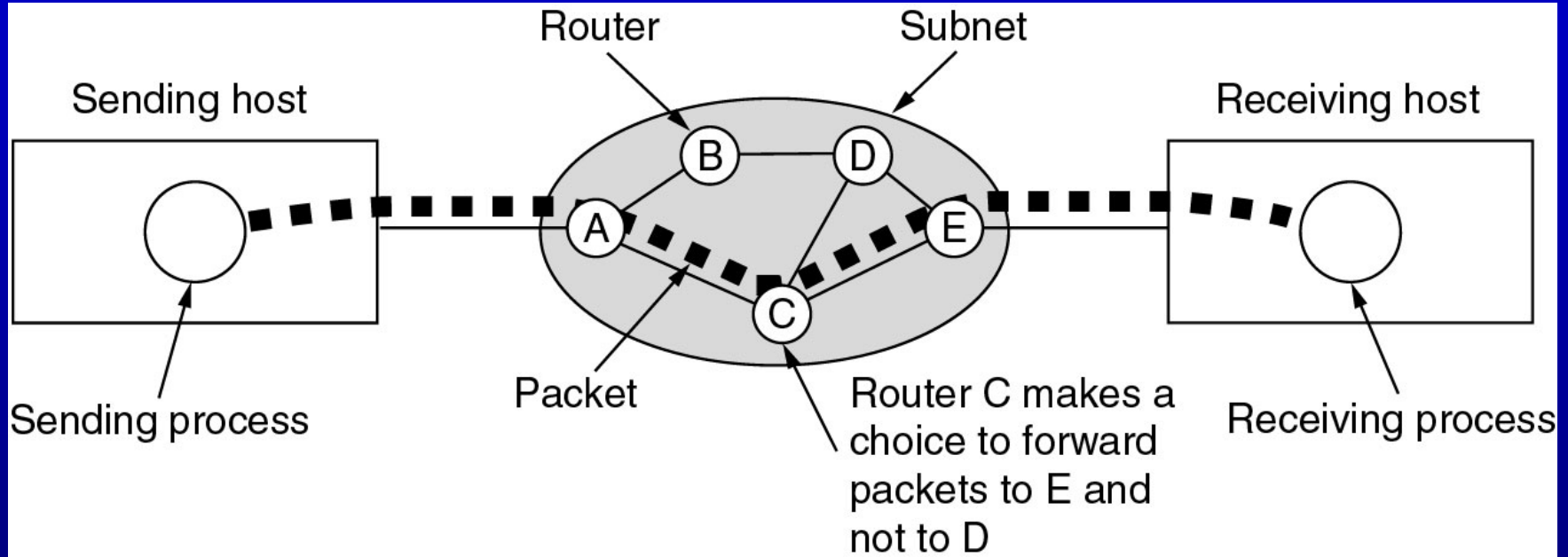
End-to-End Packet Delay

$$\text{Link packet delay} = \text{PROC} + \text{QD} + \text{TRANS} + \text{PROP.}$$

end-to-end packet delay = sum of ALL link packet delays.

Be Careful !! *end-to-end* can be defined either from Host-to-Host or only within the subnetwork.

End-to-end packet delay



End-to-end delay includes multiple hop link delays

Tanenbaum slide

Network Performance Measures

Latency:: usually implies the minimum possible delay.

Latency assumes no queuing delay and no contention encountered along the path.

Goodput:: {*measured at the receiver*} the rate in bits per second of useful traffic received. Goodput excludes duplicate packets and packets dropped along the path.

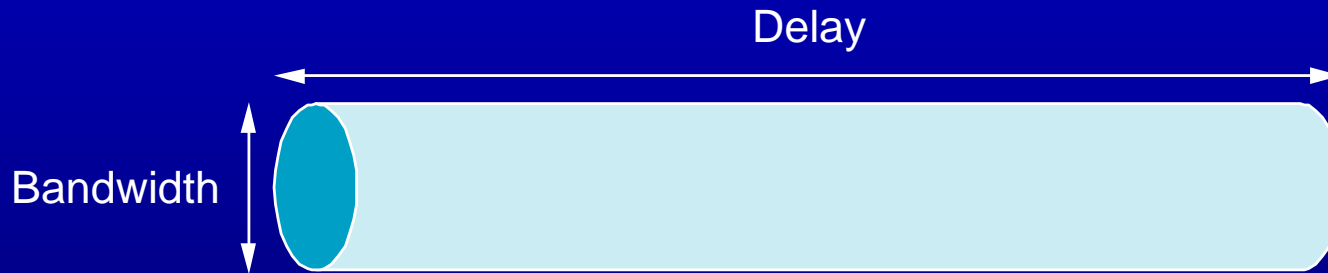
Fairness:: either Jain's fairness or max-min fairness are used to measure fair treatment among competing flows.

Quality of Service:: a QoS measure accounts for importance of specific metric to one type of application [e.g. jitter for streaming media].



Delay x Bandwidth Product

- Amount of data “in flight” or “in the pipe”
- Usually relative to RTT
- Example: $100\text{ms} \times 45\text{Mbps} = 560\text{KB}$



■ *"Bandwidth-delay product"* is the more generally accepted term.

P&D slide

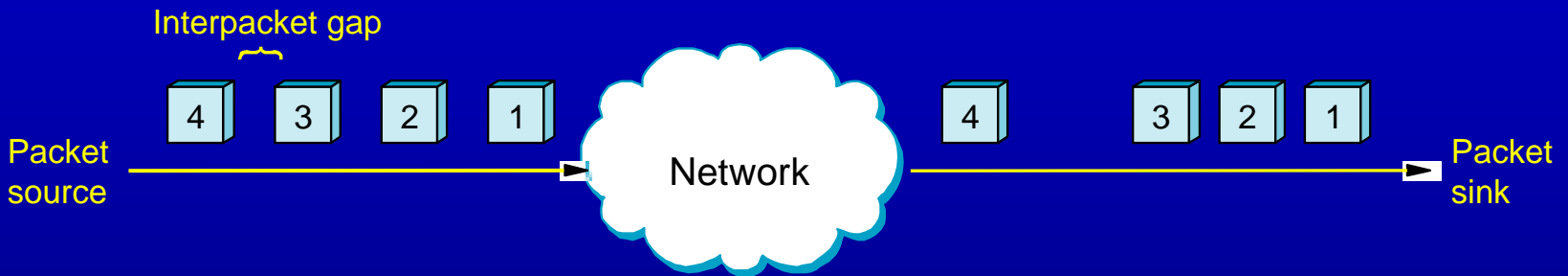


Figure 1.24 Network-induced jitter