Data Communications and Networking Fourth Edition

Forouzan

Chapter 24

Congestion Control and Quality of Service

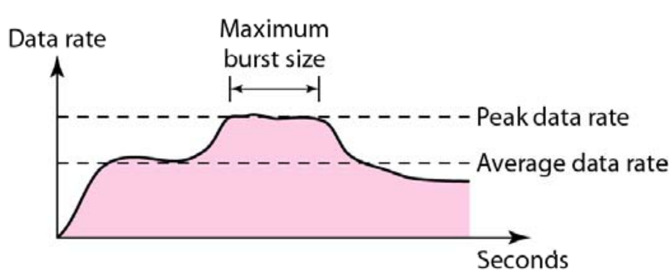
24-1 DATA TRAFFIC

The main focus of congestion control and quality of service is data traffic. In congestion control we try to avoid traffic congestion. In quality of service, we try to create an appropriate environment for the traffic. So, before talking about congestion control and quality of service, we discuss the data traffic itself.

Topics discussed in this section:

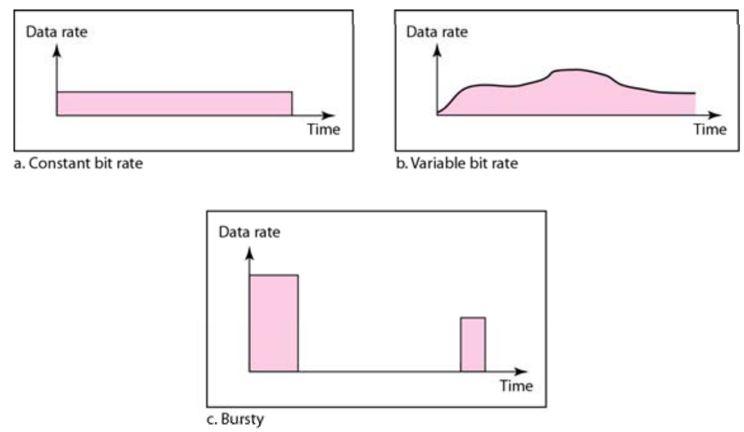
Traffic Descriptor
Traffic Profiles

Figure 24.1 *Traffic descriptors*



Effective bandwidth: the bandwidth that the network needs to allocate for the flow of traffic

Figure 24.2 *Three traffic profiles*



24-2 CONGESTION

Congestion in a network may occur if the load on the network—the number of packets sent to the network—is greater than the capacity of the network—the number of packets a network can handle. Congestion control refers to the mechanisms and techniques to control the congestion and keep the load below the capacity.

Topics discussed in this section:

Network Performance

Figure 24.3 ***Queues in a router***

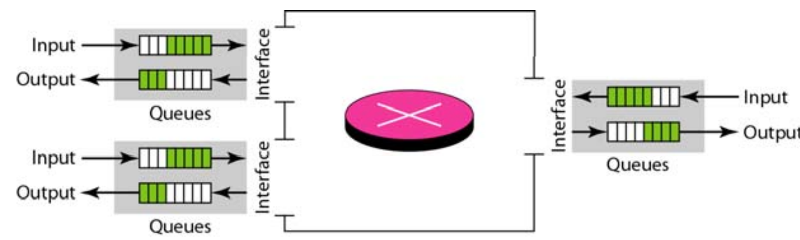
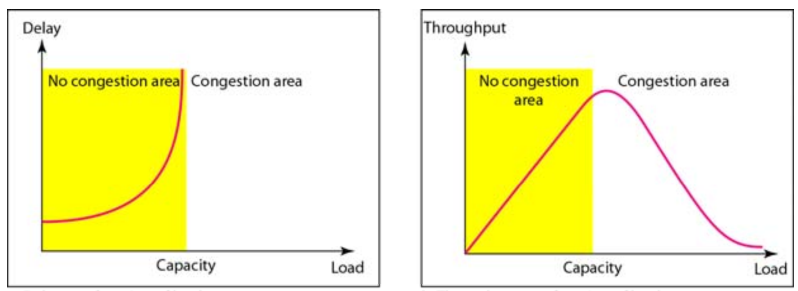


Figure ***Packet delay and throughput as functions of load***



a. Delay as a function of load

b. Throughput as a function of load

Throughput: number of packets passing through the network in a unit of time.

24-3 CONGESTION CONTROL

Congestion control refers techniques and mechanisms that can either prevent congestion, before it happens, or remove congestion, after it has happened. In general, we can divide congestion control mechanisms into two broad categories: open-loop congestion control (prevention) and closed-loop congestion control (removal).

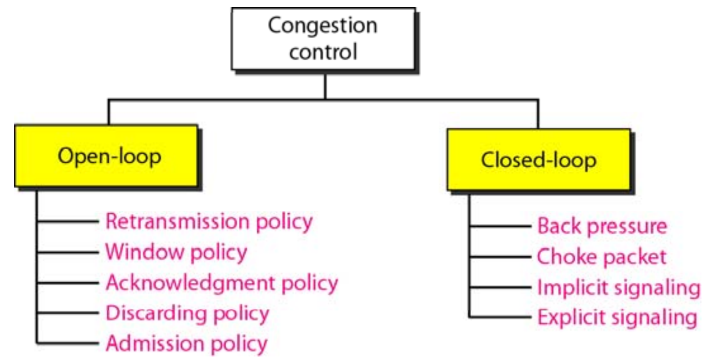
Topics discussed in this section:

Open-Loop Congestion Control
Closed-Loop Congestion Control

Effects of Congestion

- Packets arriving are stored at input buffers (not ATM)
- Routing decision made
- Packet moves to output buffer
- Packets queued for output transmitted as fast as possible
  - Statistical time division multiplexing
- If packets arrive too fast to be routed, or to be output, buffers will fill
- Can discard packets
- Can use flow control
  - Can propagate congestion through network

Figure 24.5 *Congestion control categories*



Congestion Control Objectives

- In general, we want to:
- Minimize discards
- Maintain agreed QoS (if applicable)
- Minimize probability of one end user monopoly over other end users
- Simple to implement
  - Little overhead on network or user
- Create minimal additional traffic
- Distribute resources fairly
- Limit spread of congestion
- Operate effectively regardless of traffic flow
- Minimum impact on other systems
- Minimize variance in QoS

Basic Mechanisms for Congestion Control

<u>Open-Loop Congestion Control</u> (rely on other layers for feedback and control)

Retransmission policy - a good policy can reduce congestion
Window policy - sel-reject better than go-back-N; use a bigger window size
Acknowledgment policy - don't ack each packet individually
Discard policy - a good policy by routers may prevent congestion and at the same time may not harm the integrity of the transmission
Admission policy - QOS mechanism

Basic Mechanisms for Congestion Control

<u>Closed-Loop Congestion Control</u>

- **Backpressure - when a router is congested, it informs the previous upstream router to reduce the rate of outgoing packets**
- **Choke packet of choke point - sent by router to source, similar to ICMP's source quench packet**
- **Implicit signaling - look for delay in some other action**
- **Explicit signaling - router sends an explicit signal**
- **Backward signaling - bit is set in packet moving in the direction opposite to the congestion**
- **Forward signaling - bit is set in packet moving in the direction of congestion. Receiver can use policy such as slowing down acks to alleviate congestion**

Figure 24.6 ***Backpressure method for alleviating congestion***

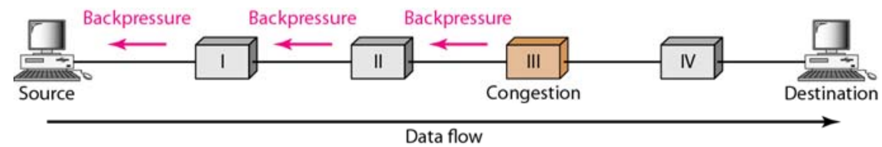
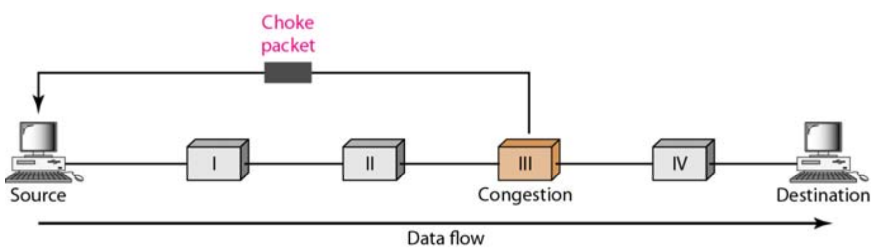
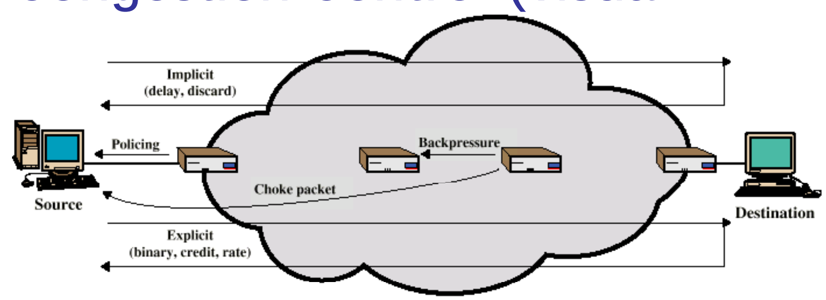


Figure 24.7 ***Choke packet***



Basic Mechanisms for Congestion Control (visual



Backpressure

- If node becomes congested it can slow down or halt flow of packets from other nodes
- May mean that other nodes have to apply control on incoming packet rates
- Propagates back to source
- Can restrict to logical connections generating most traffic
- Used in connection oriented that allow hop by hop congestion control (e.g. X.25)
- Not used in ATM or frame relay
- Only recently developed for IPv6 (PRI field)

Choke Packet

- Control packet
  - Generated at congested node
  - Sent to source node
  - e.g. ICMP source quench
    - From router or destination
    - Source cuts back until no more source quench message
    - Sent for every discarded packet, or anticipated
- Rather crude mechanism

Implicit Congestion Signaling

- Transmission delay may increase with congestion
- Packets may be discarded
- Source can detect these as implicit indications of congestion (source is responsible, not network)
- Useful on connectionless (datagram) networks
  - e.g. IP based (TCP includes congestion and flow control)
- Used in frame relay LAPF

Explicit Congestion Signaling

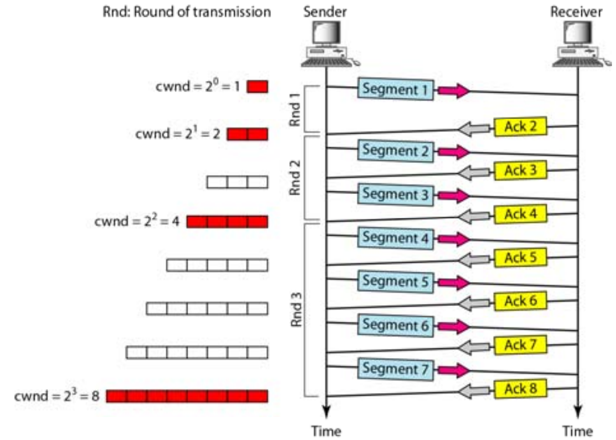
- Network alerts end systems of increasing congestion
- Used on connection-oriented networks
- End systems take steps to reduce offered load
- Backwards
  - Congestion avoidance info sent in opposite direction of packet travel
- Forwards
  - Congestion avoidance info sent in same direction as packet travel - when end system receives info, either sends it back to source or hands it to higher layer to take action

Congestion control in TCP

To better understand the concept of congestion control, we will consider congestion control in TCP.

- ***Slow start***
- ***Congestion avoidance***
- ***Congestion detection***

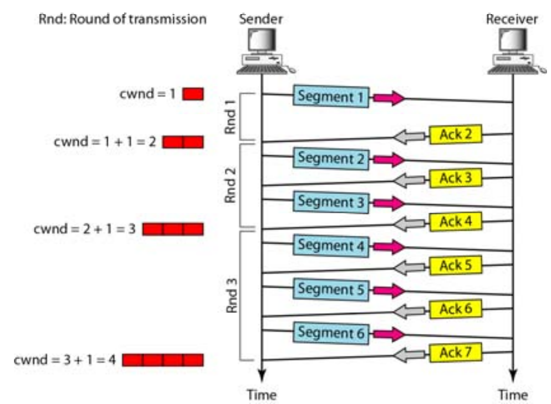
Figure 24.8 ***Slow start, exponential increase***



Note

In the slow-start algorithm, the size of the congestion window increases exponentially until it reaches a threshold.

Figure 24.9 ***Congestion avoidance, additive increase***



Note

In the congestion avoidance algorithm, the size of the congestion window increases additively until congestion is detected.

Note

An implementation reacts to congestion detection in one of the following ways:
- **If detection is by time-out, a new slow start phase starts.**
- **If detection is by three ACKs, a new congestion avoidance phase starts.**

Figure 24.10 *TCP congestion policy summary*

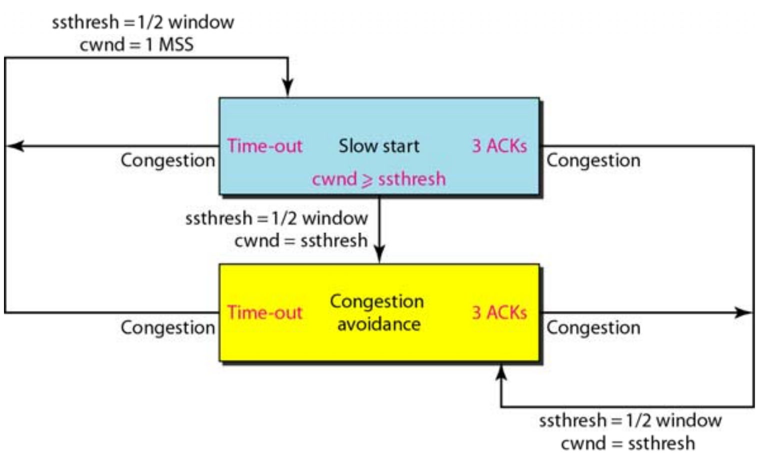
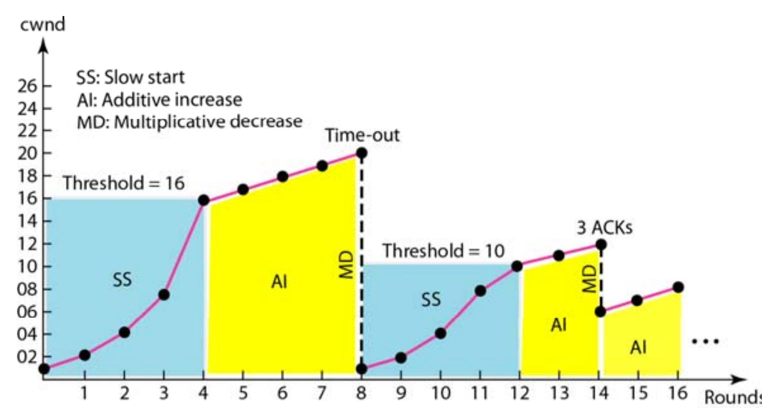


Figure 24.11 *Congestion example – Max-w =32*



24-5 QUALITY OF SERVICE

Quality of service (QoS) is an internetworking issue that has been discussed more than defined. We can informally define quality of service as something a flow seeks to attain.

Topics discussed in this section:

Flow Characteristics
Flow Classes

Figure 24.15 ***Flow characteristics***