Data and Computer Communications

Chapter 13 – Congestion in Data Networks

Eighth Edition

by William Stallings

Lecture slides by Lawrie Brown

Congestion in Data Networks

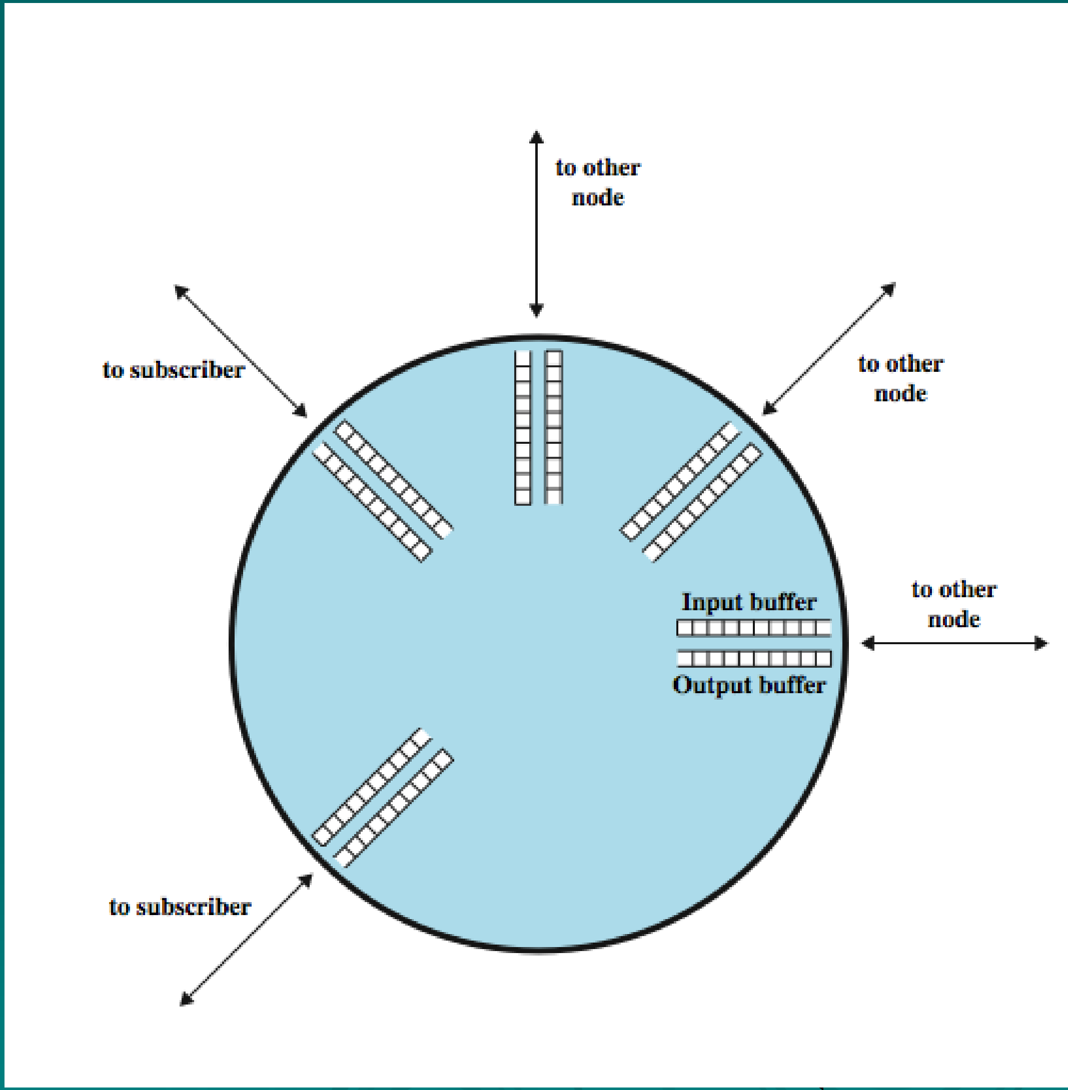
At St. Paul's a great throng crammed the platform. She saw a sea of faces, each stamped with a kind of purposeful, hungry urgency, a determination to get into this train. As before, when she was on the Northern Line, she thought there must be some rule, some operating law, that would stop more than a limited, controlled number getting in. Authority would appear and stop it.

—*King Solomon's Carpet*, Barbara Vine (Ruth Rendell)

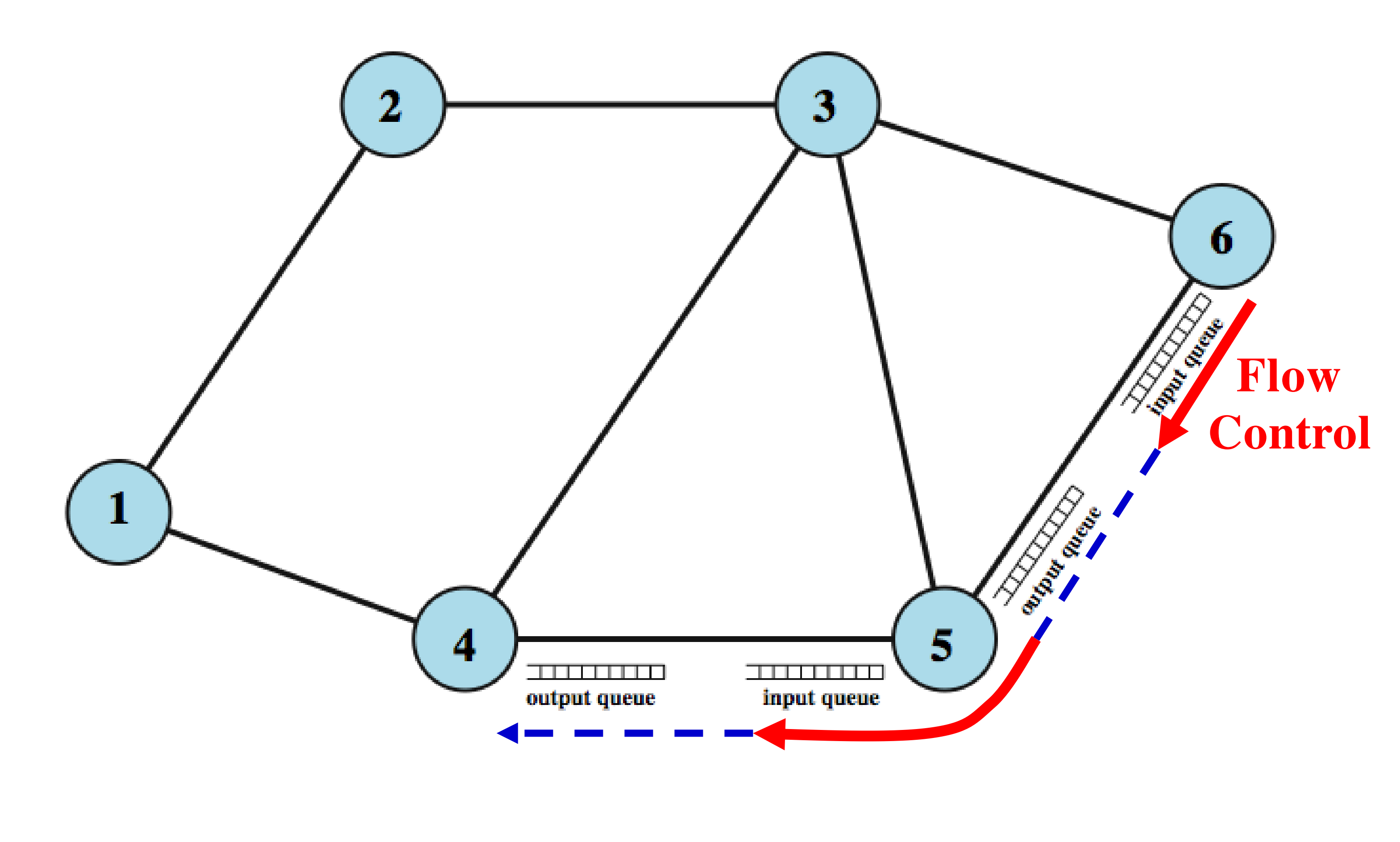
What Is Congestion?

- congestion occurs when the no of packets being transmitted through the network approaches the packet handling capacity of the network
- congestion control aims to keep no of packets below a level at which performance falls off dramatically
- a data network is a network of queues
- generally 80% utilization is critical
- finite queues mean data may be lost

Queues at a Node

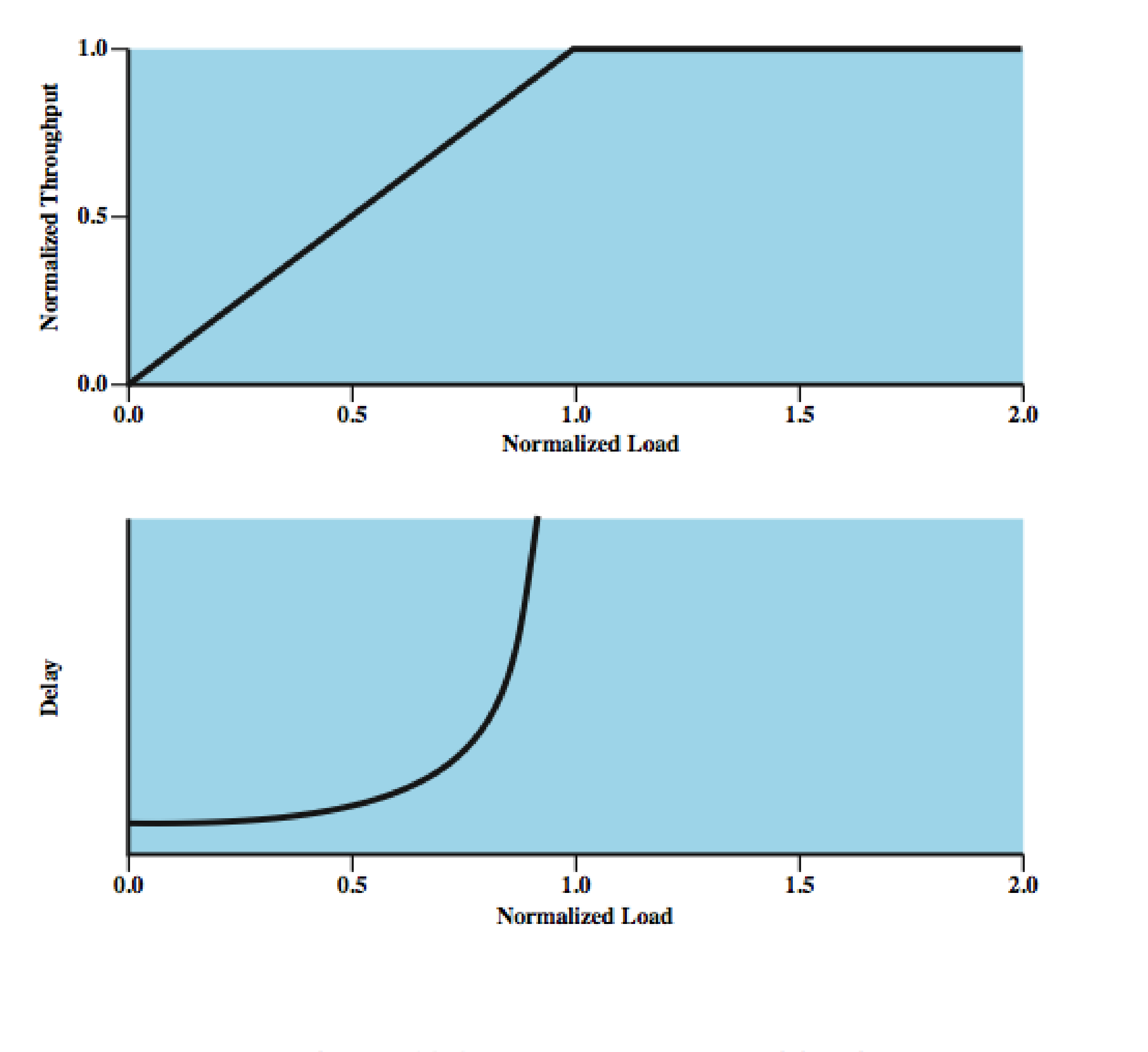


Interaction of Queues

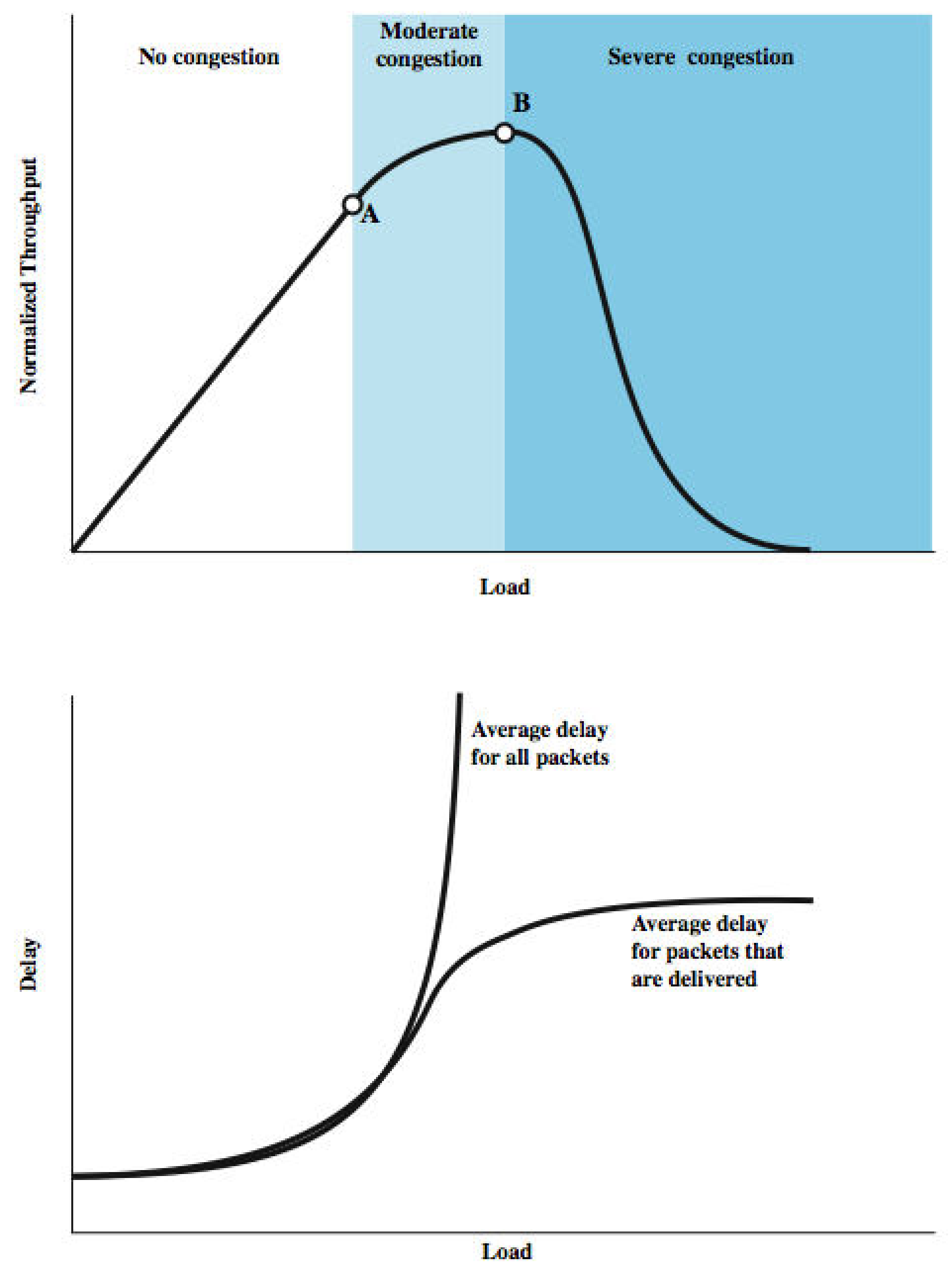


Ideal Network Utilization

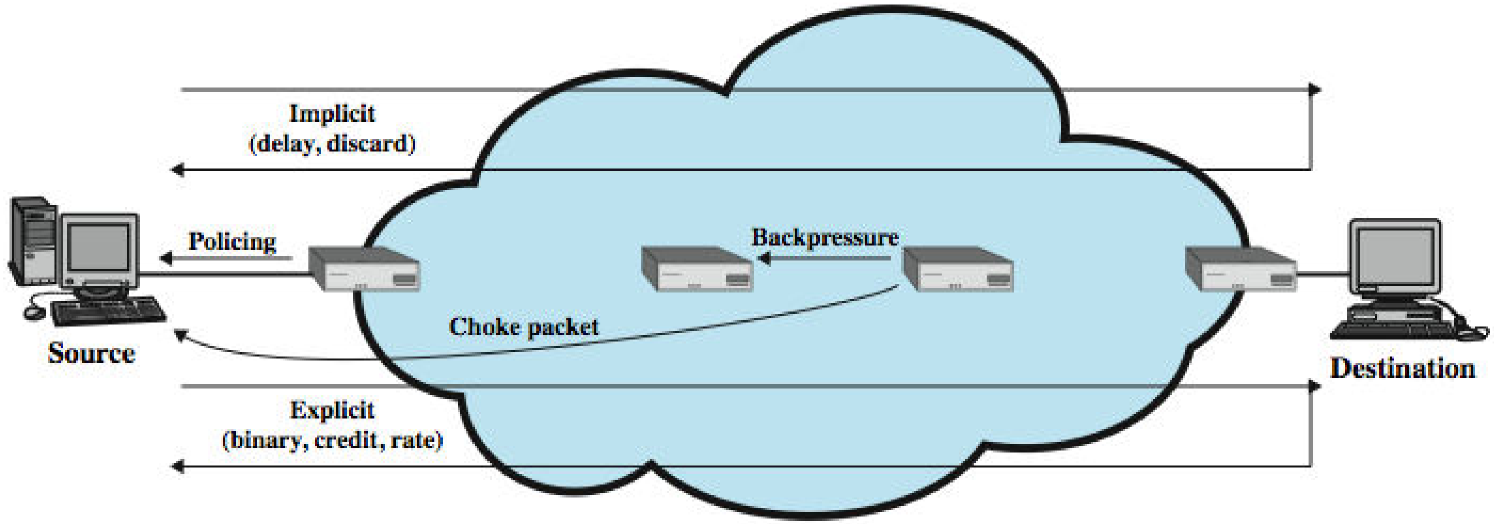
1. Infinite buffers
2. No signaling overhead



Effects of Congestion - No Control



Mechanisms for Congestion Control



Backpressure

- if node becomes congested it can slow down or halt flow of packets from other nodes
  - cf. backpressure in blocked fluid pipe
  - may mean that other nodes have to apply control on incoming packet rates
  - propagates back to source
- can restrict to high traffic logical connections
- used in connection-oriented nets that allow hop by hop congestion control (eg. X.25)
- not used in ATM nor frame relay
- only recently developed for IP

Choke Packet

- a control packet
  - generated at congested node
  - sent to source node
  - eg. ICMP source quench
    - from router or destination
    - source cuts back until no more source quench message
    - sent for every discarded packet, or anticipated
- is a rather crude mechanism

Implicit Congestion Signaling

- transmission delay increases with congestion
- hence a packet may be discarded
- source detects this implicit congestion indication
- useful on connectionless (datagram) networks
  - eg. IP based
    - (TCP includes congestion and flow control - see chapter 17)
- used in frame relay LAPF (control protocol)
  - End-to-end
  - Capable of detecting lost frames and adjusting the flow of data accordingly

Explicit Congestion Signaling

- network alerts end systems of increasing congestion
- end systems take steps to reduce offered load
- Backwards
  - congestion avoidance notification in opposite direction to packet required
- Forwards
  - congestion avoidance notification in same direction as packet required

Explicit Signaling Categories

- Binary
  - a bit set in a packet indicates congestion
- Credit based
  - indicates how many packets source may send
  - common for end to end flow control
- Rate based
  - supply explicit data rate limit
  - nodes along path may request rate reduction
  - eg. ATM

Traffic Management

- fairness
  - provide equal treatment of various flows
- quality of service
  - different treatment for different connections
- reservations
  - traffic contract between user and network
  - carry best-effort or discard excess traffic
  - E.g. ATM, RSVP
  - Traffic policing

Congestion Control in Packet Switched Networks

- send control packet to some or all source nodes
  - requires additional traffic during congestion
- rely on routing information
  - may react too quickly
- end to end probe packets
  - adds to overhead
- add congestion info to packets in transit
  - either backwards or forwards

Frame Relay Congestion Control

- minimize discards
- maintain agreed QoS
- minimize probability of one end user monopoly
- simple to implement
- 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

FR Control Techniques

- difficult for frame-relay
  - Because of limited tools
- joint network & end-system responsibility
- techniques:
  - discard strategy
  - congestion avoidance
  - explicit signaling
  - congestion recovery
  - implicit signaling mechanism

FR Congestion Control

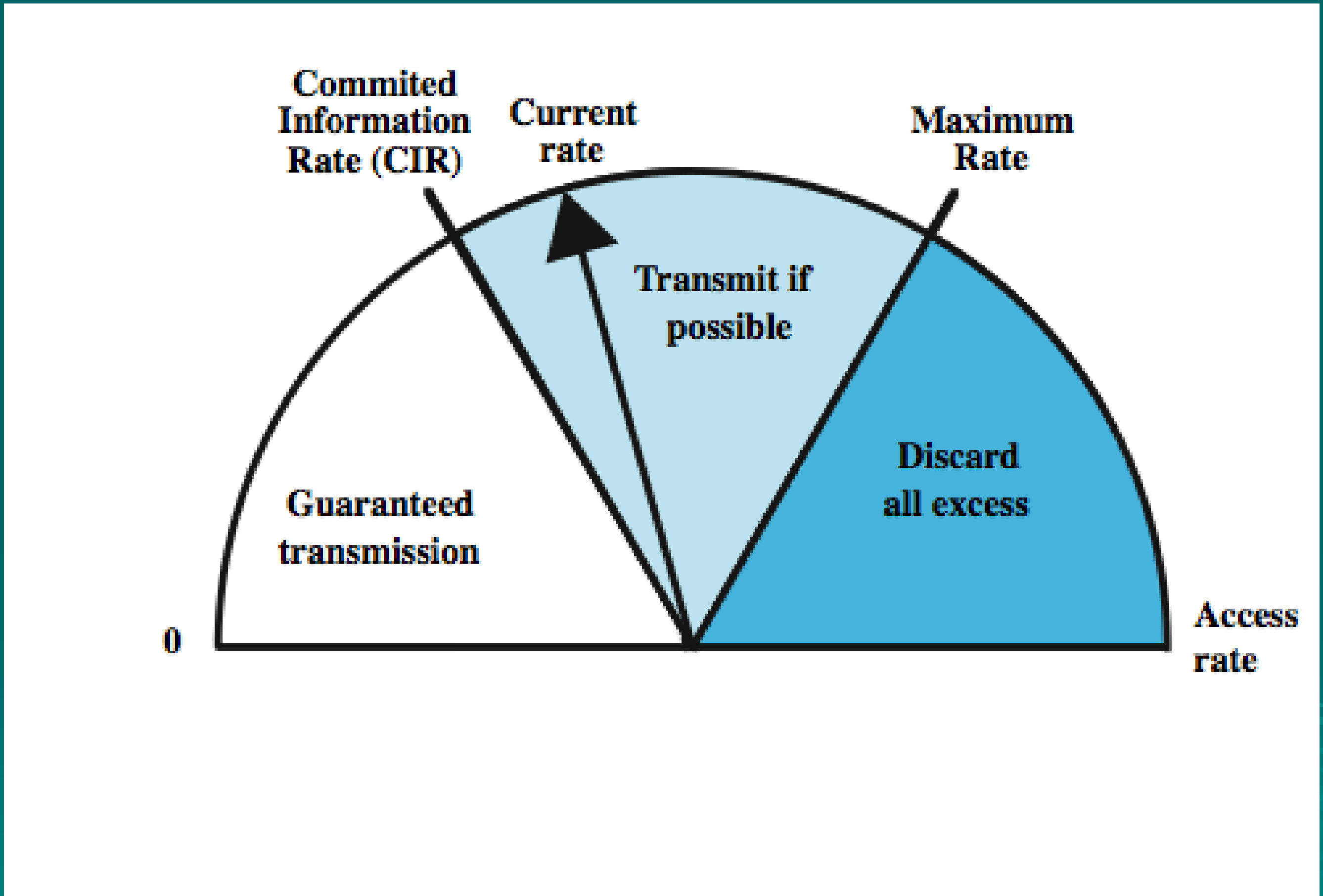
Table 13.1 Frame Relay Congestion Control Techniques

Technique	Type	Function	Key Elements
Discard control	Discard strategy	Provides guidance to network concerning which frames to discard	DE bit
Backward explicit Congestion Notification	Congestion avoidance	Provides guidance to end systems about congestion in network	BECN bit or CLLM message
Forward explicit Congestion Notification	Congestion avoidance	Provides guidance to end systems about congestion in network	FECN bit
Implicit congestion notification	Congestion recovery	End system infers congestion from frame loss	Sequence numbers in higher-layer PDU

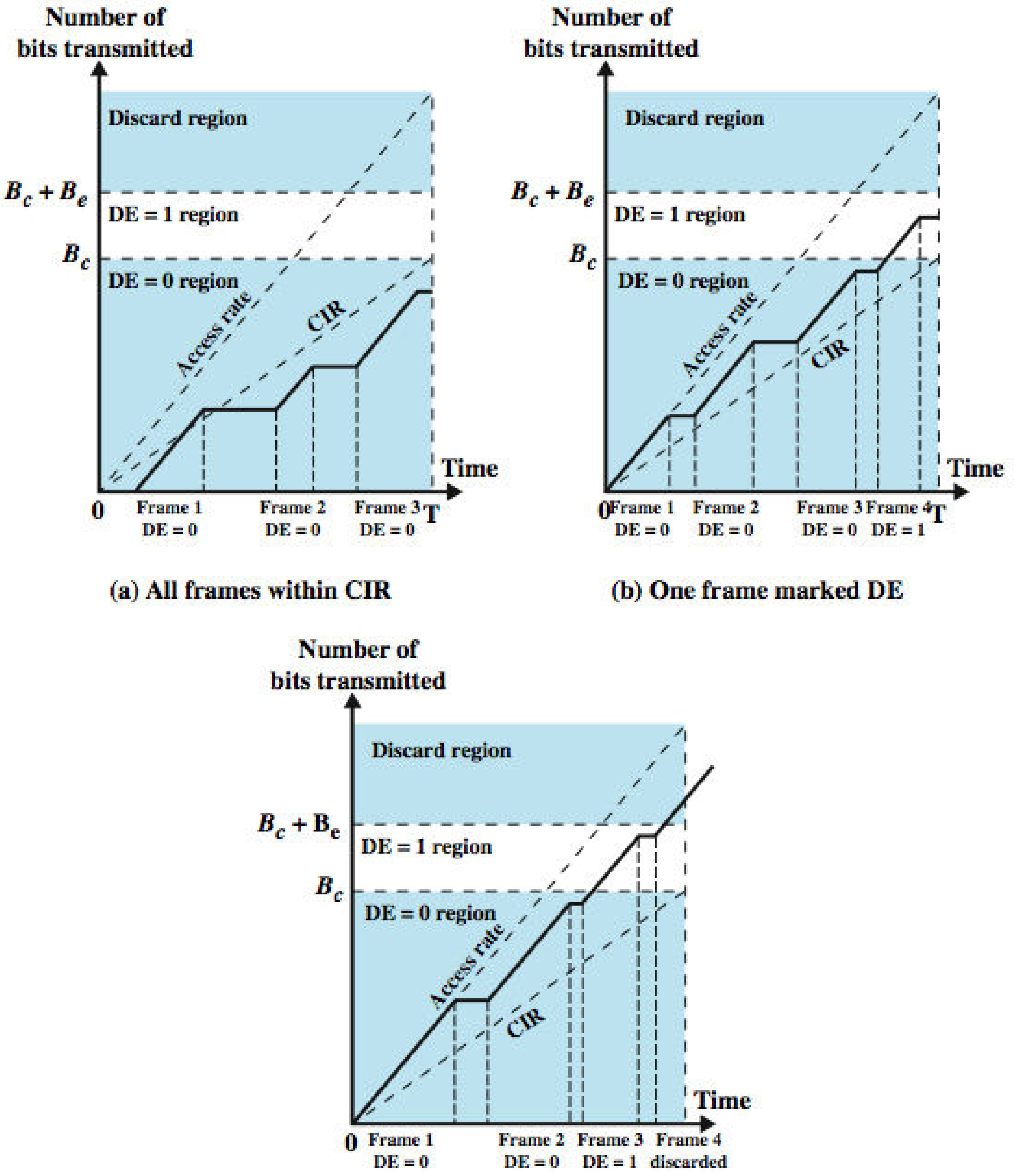
Traffic Rate Management

- must discard frames to cope with congestion
  - arbitrarily, no regard for source
  - no reward for restraint so end systems transmit as fast as possible
  - Committed information rate (CIR)
    - data in excess of this liable to discard
    - not guaranteed in extreme congestion situations
    - aggregate CIR should not exceed physical data rate
- Committed burst size (B_c)
- Excess burst size (B_e)

Operation of CIR



Relationship Among Congestion Parameters



(a) All frames within CIR

(b) One frame marked DE

(c) One frame marked DE; one frame discarded

Congestion Avoidance using Explicit Signaling

- network alerts end systems of growing congestion using
  - backward explicit congestion notification
  - forward explicit congestion notification
- frame handler monitors its queues
- may notify some or all logical connections
- user response: reduce rate

ATM Traffic Management

- high speed, small cell size, limited overhead bits
- still evolving
- reasons existing tools are inadequate for ATM
  - majority of traffic not amenable to flow control
  - feedback slow due to reduced transmission time compared with propagation delay
  - wide range of application demands
  - different traffic patterns
  - different network services
  - high speed switching and transmission increases volatility

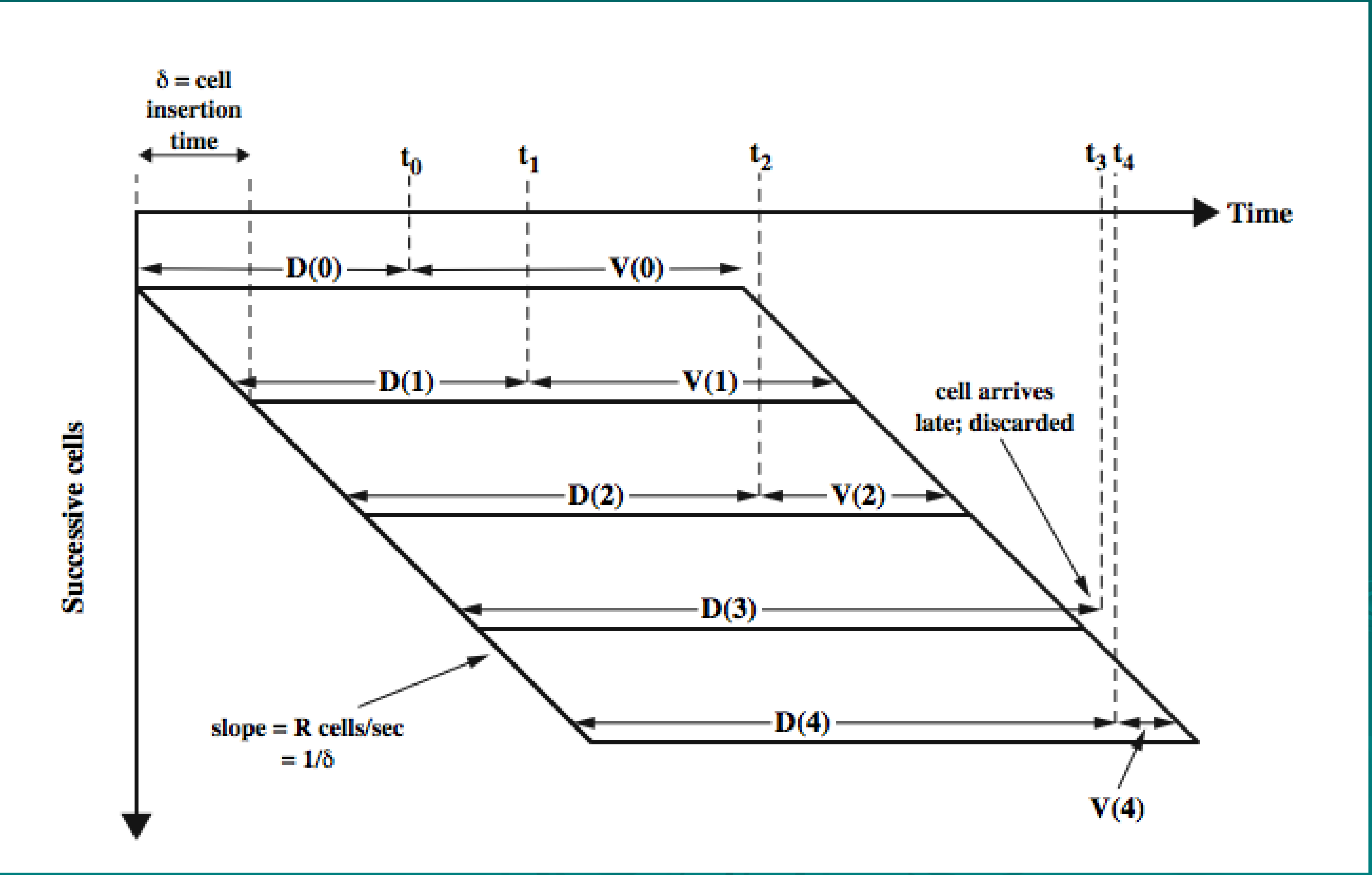
Latency/Speed Effects

- consider ATM at 150Mbps
- takes ~2.8×10^{-6} seconds to insert single cell
- time to traverse network depends on propagation delay and switching delay
- assume propagation at two-thirds speed of light
- if source and destination on opposite sides of USA, propagation time ~ 48×10^{-3} seconds
- given implicit congestion control, by the time dropped cell notification has reached source, 7.2×10^{6} bits have been transmitted
- this is not a good strategy for ATM

Cell Delay Variation

- for ATM voice/video, data is a stream of cells
- delay across network must be short
- rate of delivery must be constant
- there will always be some variation in transit
- delay cell delivery to application so that constant bit rate can be maintained to application

Timing of CBR Cells



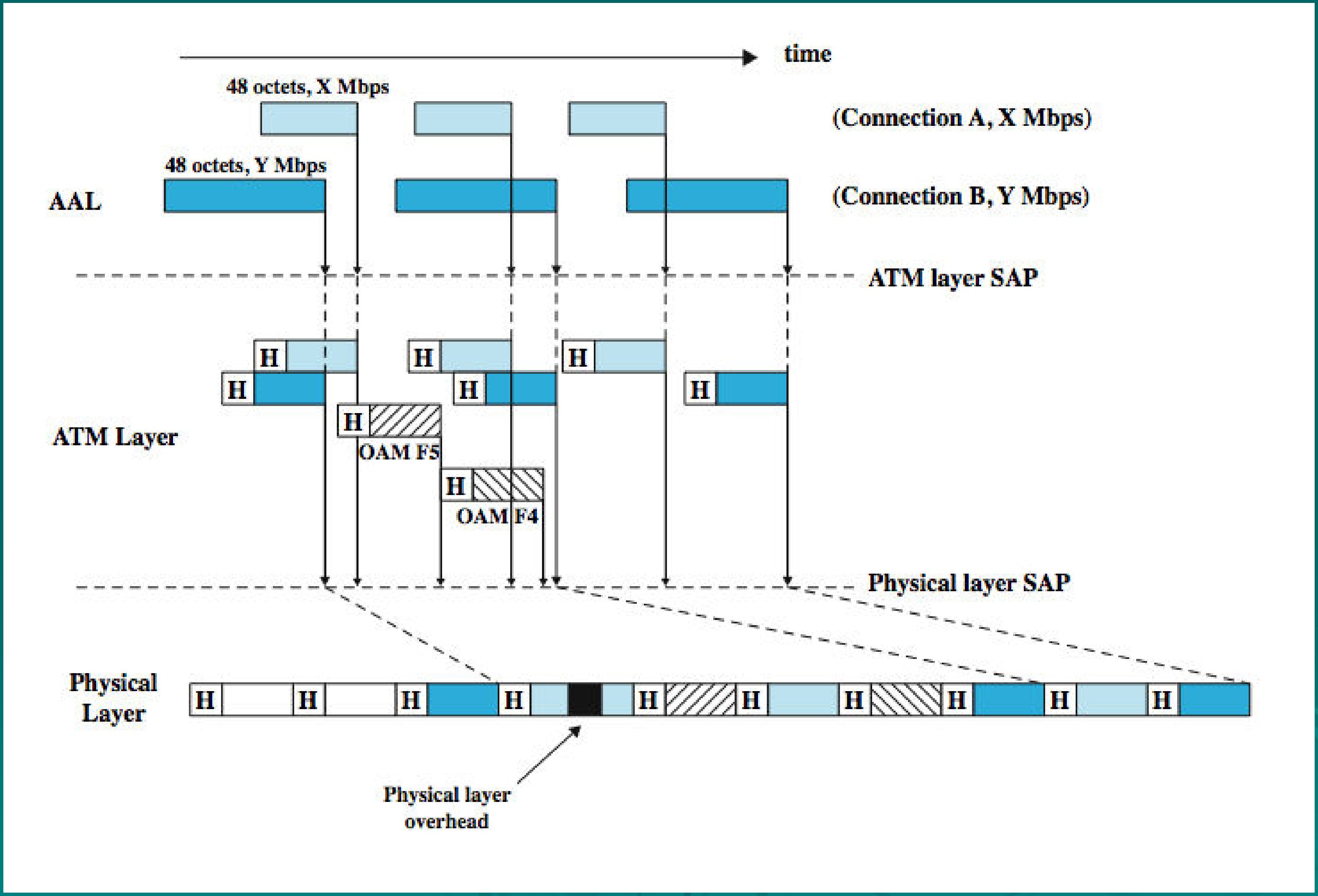
Network Contribution to Cell Delay Variation

- in packet switched networks is due to queuing delays and routing decision time
- in Frame relay networks is similar
- in ATM networks
  - less than frame relay
  - ATM protocol designed to minimize processing overheads at switches
  - ATM switches have very high throughput
  - only noticeable delay is from congestion
  - must not accept load that causes congestion

Cell Delay Variation At The UNI

- application produces data at fixed rate
- 3 layers of ATM processing causes delay
  - interleaving cells from different connections
  - operation and maintenance cell interleaving
  - if using synchronous digital hierarchy frames, these are inserted at physical layer
- cannot predict these delays

Origins of Cell Delay Variation



Traffic and Congestion Control Framework

- ATM layer traffic and congestion control should support QoS classes for all foreseeable network services
- should not rely on AAL protocols that are network specific, nor higher level application specific protocols
- should minimize network and end to end system complexity

Timings Considered

- timing intervals considered:
  - cell insertion time
  - round trip propagation time
  - connection duration
  - long term
- traffic control strategy then must:
  - determine whether a given new connection can be accommodated
  - agree performance parameters with subscriber
- now review various control techniques

Table13.2 Traffic Control and Congestion Control Functions

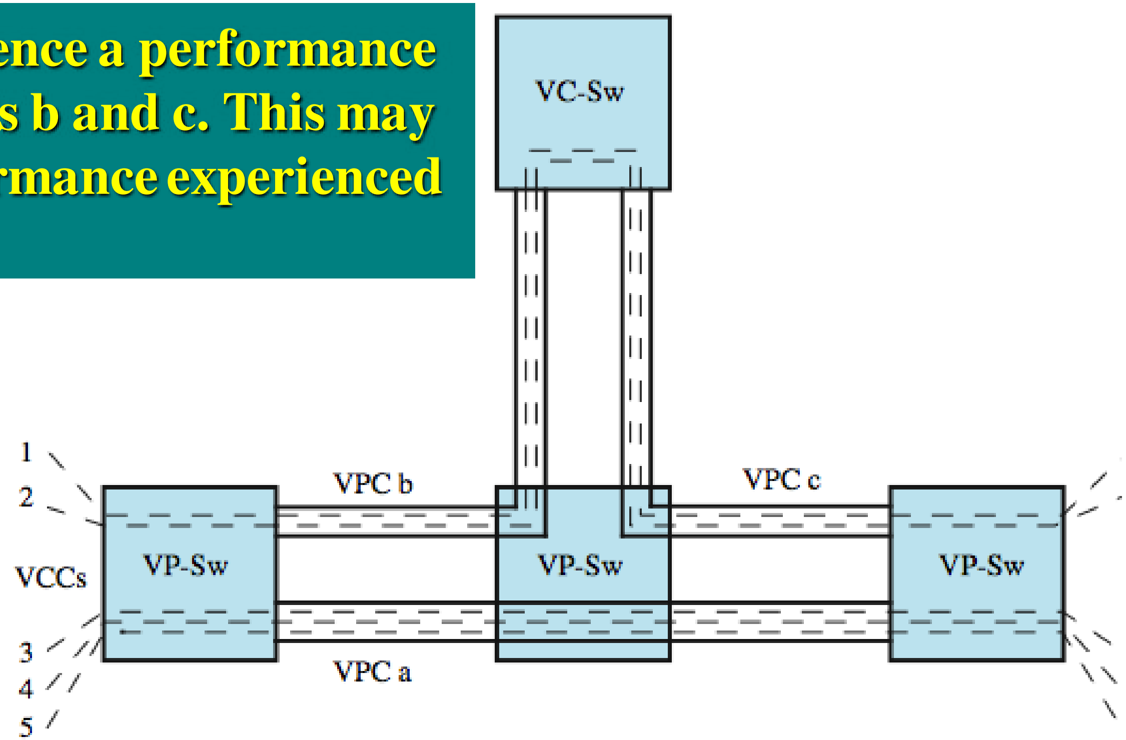
Response Time	Traffic Control Functions	Congestion Control Functions
Long Term	•Resource management using virtual paths	
Connection Duration	•Connection admission control (CAC)	
Round-Trip Propagation Time	•Fast resource management	•Explicit forward congestion indication (EFCI) •ABR flow control
Cell Insertion Time	•Usage parameter control (UPC) •Priority control •Traffic shaping	•Selective cell discard

Resource Management Using Virtual Paths

- separate traffic flow according to service characteristics on a virtual path
  - user to user application
  - user to network application
  - network to network application
- QoS parameters concerned with are:
  - cell loss ratio
  - cell transfer delay
  - cell delay variation

Configuration of VCCs and VPCs

VCCs 1 and 2 experience a performance that depends on VPCs b and c. This may differ from the performance experienced by VCCs 3, 4, and 5



VPC = Virtual path connection
VCC = Virtual channel connection
VP-Sw = Virtual path switching function
VC-Sw = Virtual channel switching function

Allocating VCCs within VPC

- all VCCs within VPC should experience similar network performance
- options for allocation:
  - aggregate peak demand
    - set VPC capacity to total of all peak VCC rates
    - will meet peak demands, but often underutilized
  - statistical multiplexing
    - set VPC capacity to more than average VCC rates
    - will see greater variation but better utilization

Connection Admission Control

- first line of defense
- user specifies traffic characteristics for new connection (VCC or VPC) by selecting a QoS
- network accepts connection only if it can meet the demand
- traffic contract
  - peak cell rate
  - cell delay variation
  - sustainable cell rate
  - burst tolerance

Usage Parameter Control

- UPC function monitors a connection to ensure traffic obeys contract
- purpose is to protect network resources from overload by one connection
- done on VCC and VPC
- peak cell rate and cell delay variation
- sustainable cell rate and burst tolerance
- UPC discards cells outside traffic contract

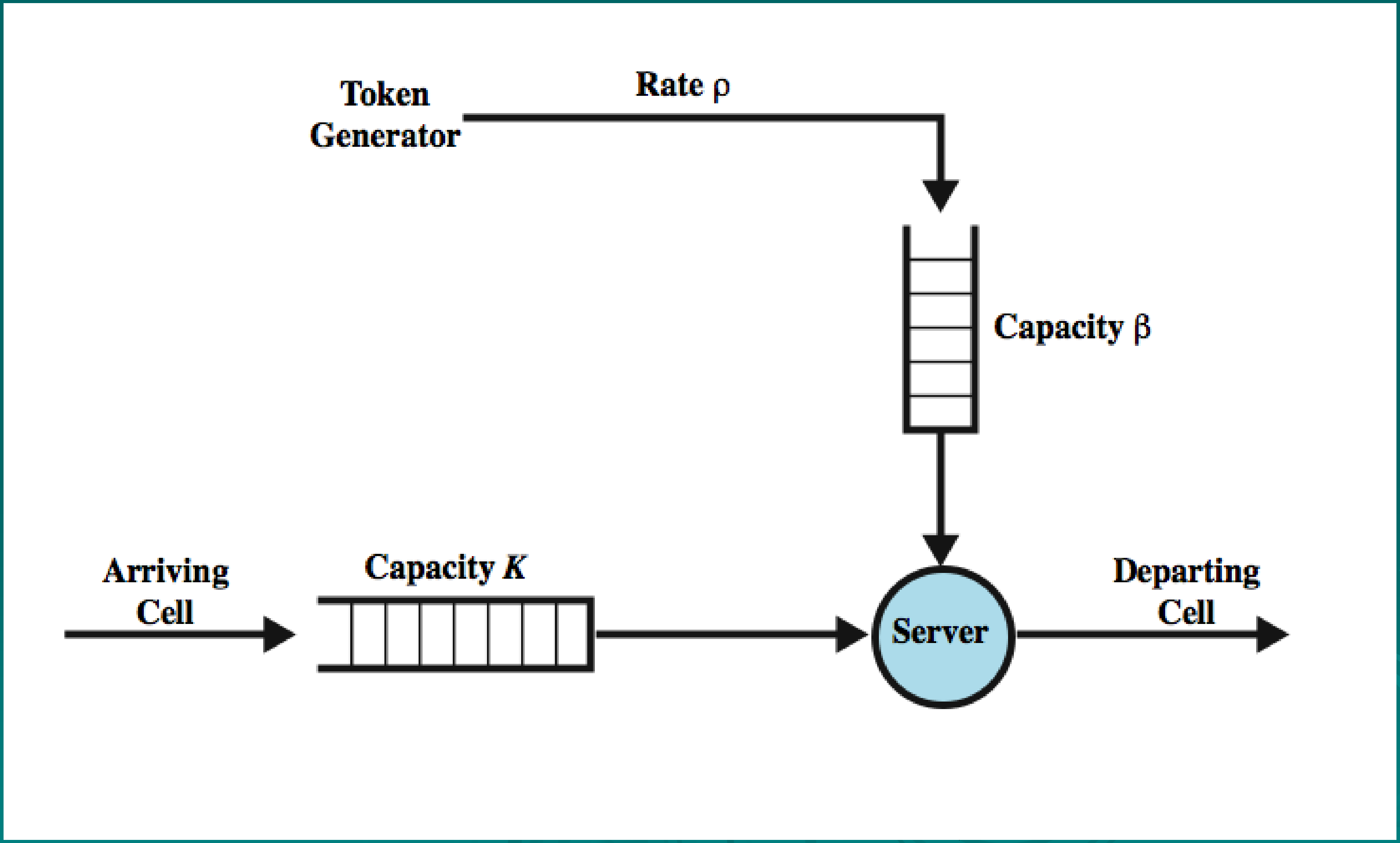
Selective Cell Discard

- when network at point beyond UPC discards (CLP=1) cells
- aim to discard lower-priority cells when congested to protect higher-priority cells
  - note. can't distinguish between cells originally labeled lower priority, verses those tagged by UPC function

Traffic Shaping

- UPC provides a form of traffic policing
- can be desirable to also shape traffic
- smoothing out traffic flow
- reducing cell clumping
- token bucket

Token Bucket for Traffic Shaping



GFR Traffic Management

- guaranteed frame rate (GFR) as simple as UBR from end system viewpoint
- places modest requirements on ATM network
- end system does no policing or shaping of traffic
- may transmit at line rate of ATM adaptor
- no guarantee of frame delivery (under congestion)
  - so higher layer (eg. TCP) must do congestion control
- user can reserve capacity for each VC
  - ensures application can send at min rate with no loss
  - if no congestion, higher rates maybe used

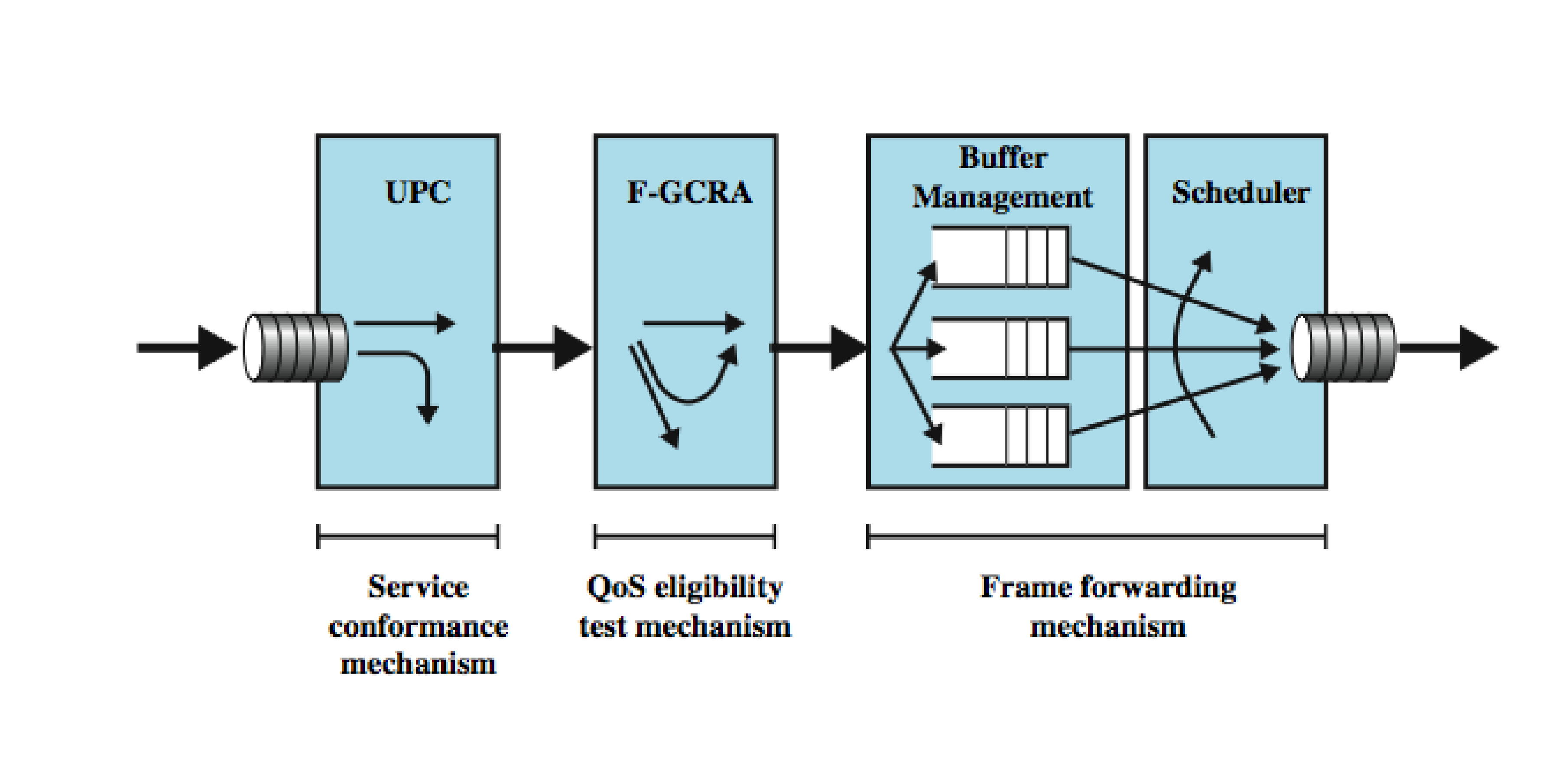
Frame Recognition

- GFR recognizes frames as well as cells
- when congested, network discards whole frame rather than individual cells
- all cells of a frame have same CLP bit setting
- CLP=1 AAL5 frames lower priority (best effort)
- CLP=0 frames minimum guaranteed capacity

GFR Contract Parameters

- Peak cell rate (PCR)
- Minimum cell rate (MCR)
- Maximum burst size (MBS)
- Maximum frame size (MFS)
- Cell delay variation tolerance (CDVT)

Components of GFR System Supporting Rate Guarantees



Tagging and Policing

- discriminates between frames that conform to contract and those that don't
- set CLP=1 on all cells in frame if not
  - gives lower priority
- maybe done by network or source
- network may discard CLP=1 cells
  - policing

Buffer Management

- deals with treatment of buffered cells
- congestion indicated by high buffer occupancy
- will discard tagged cells in preference to untagged cells
  - including ones already in buffer to make room
- may do per VC buffering for fairness
- cell discard based on queue-specific thresholds

Scheduling

- preferential treatment to untagged cells
- separate queues for each VC
- make per-VC scheduling decisions
- enables control of outgoing rate of VCs
- VCs get fair capacity allocation
- still meet contract

GFC Conformance Definition

- UPC function monitors each active VC
- to ensure traffic conforms to contract
- tag or discard nonconforming cells
- frame conforms if all cells conform
- a cell conforms if:
  - rate of cells is within contract
  - all cells in frame have same CLP
  - frame satisfies MFS parameter
    - check if either last cell in frame or cell count < MFS

QoS Eligibility Test

- two stage filtering process
- a frame is tested for conformance to contract
  - if not, may discard or tag
  - set upper bound & penalize cells above upper bound
  - do expect attempt to deliver tagged cells
- determine frames eligible for QoS guarantees
  - under GFR contract for VC
  - set lower bound on traffic
  - frames in traffic flow below threshold are eligible

GFR VC Frame Categories

- nonconforming frame (above upper bound)
  - cells of this frame will be tagged or discarded
- conforming but ineligible frames (in-between)
  - cells will receive a best-effort service
- conforming and eligible frames (under lower bound)
  - cells will receive a guarantee of delivery

- form of cell rate algorithm is used

Summary

- congestion effects
- congestion control
- traffic management
- frame relay congestion control
- ATM congestion control