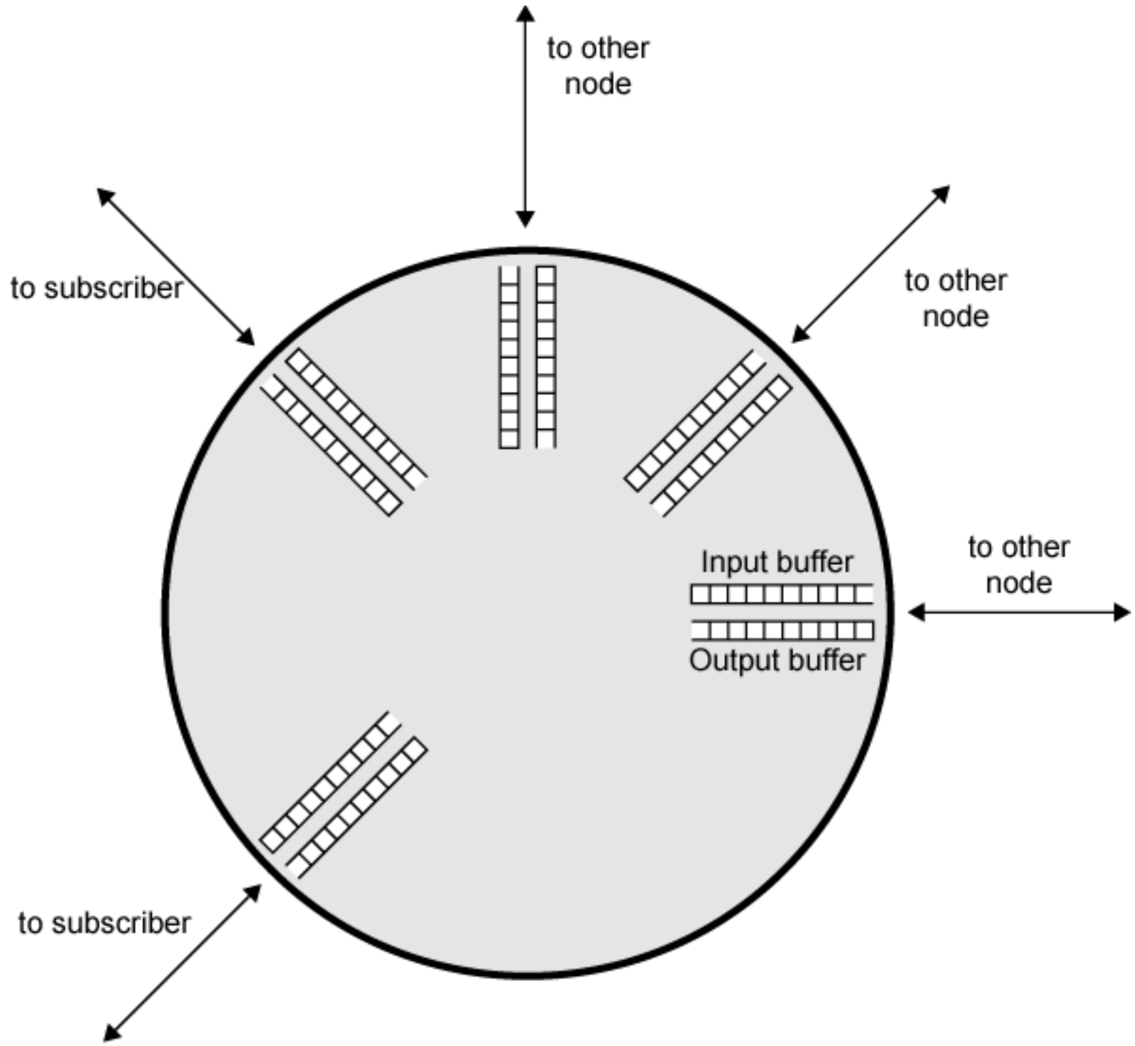
Congestion in Data Networks

Congestion in Data Networks

- What is Congestion?
  - Congestion occurs when the number of packets being transmitted through the network approaches the packet handling capacity of the network
  - Congestion control aims to keep number of packets below level at which performance falls off dramatically
  - Data network is a network of queues
  - Generally 80% utilization is critical
  - Finite queues mean data may be lost

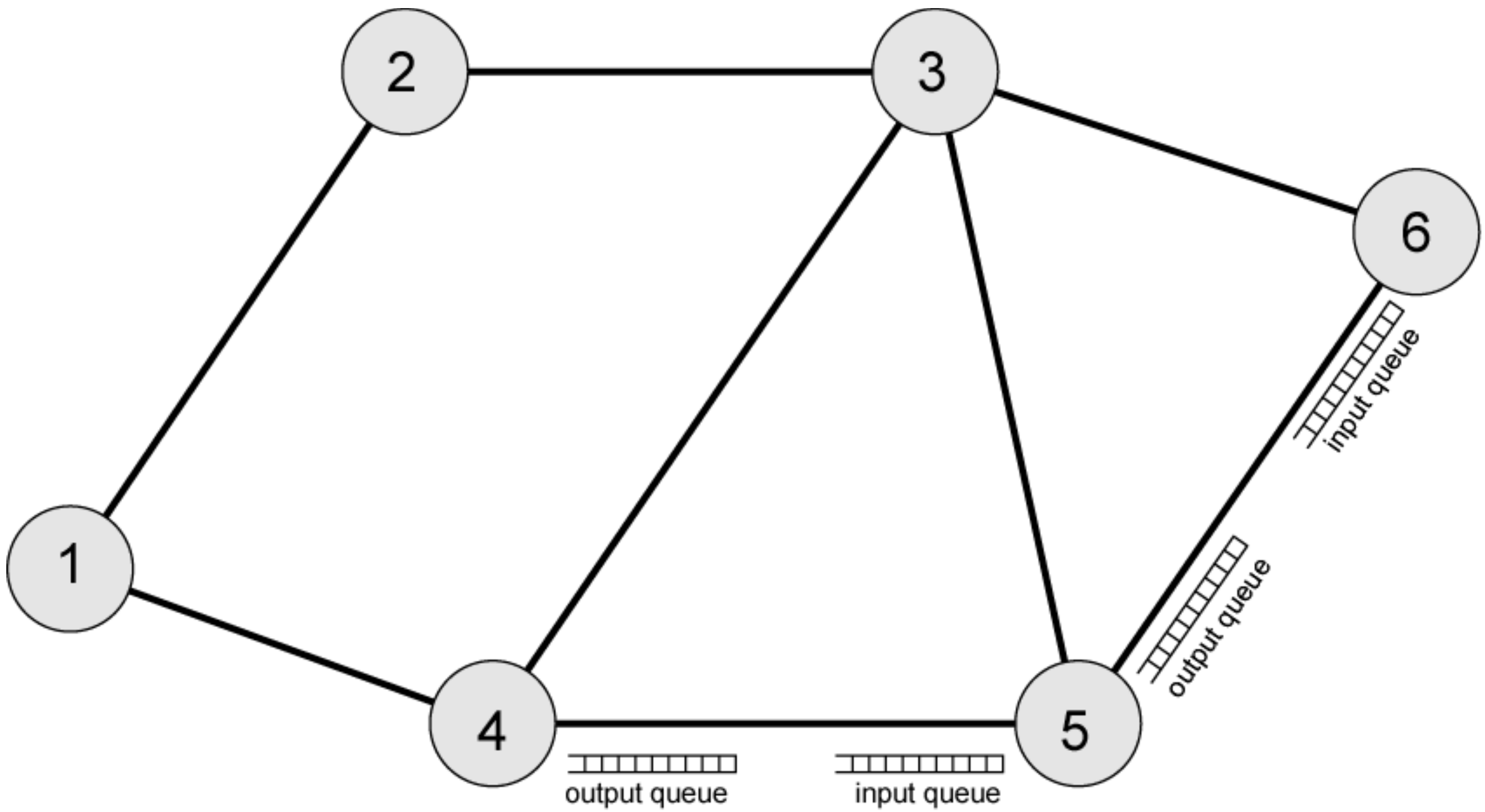
Queues at a Node



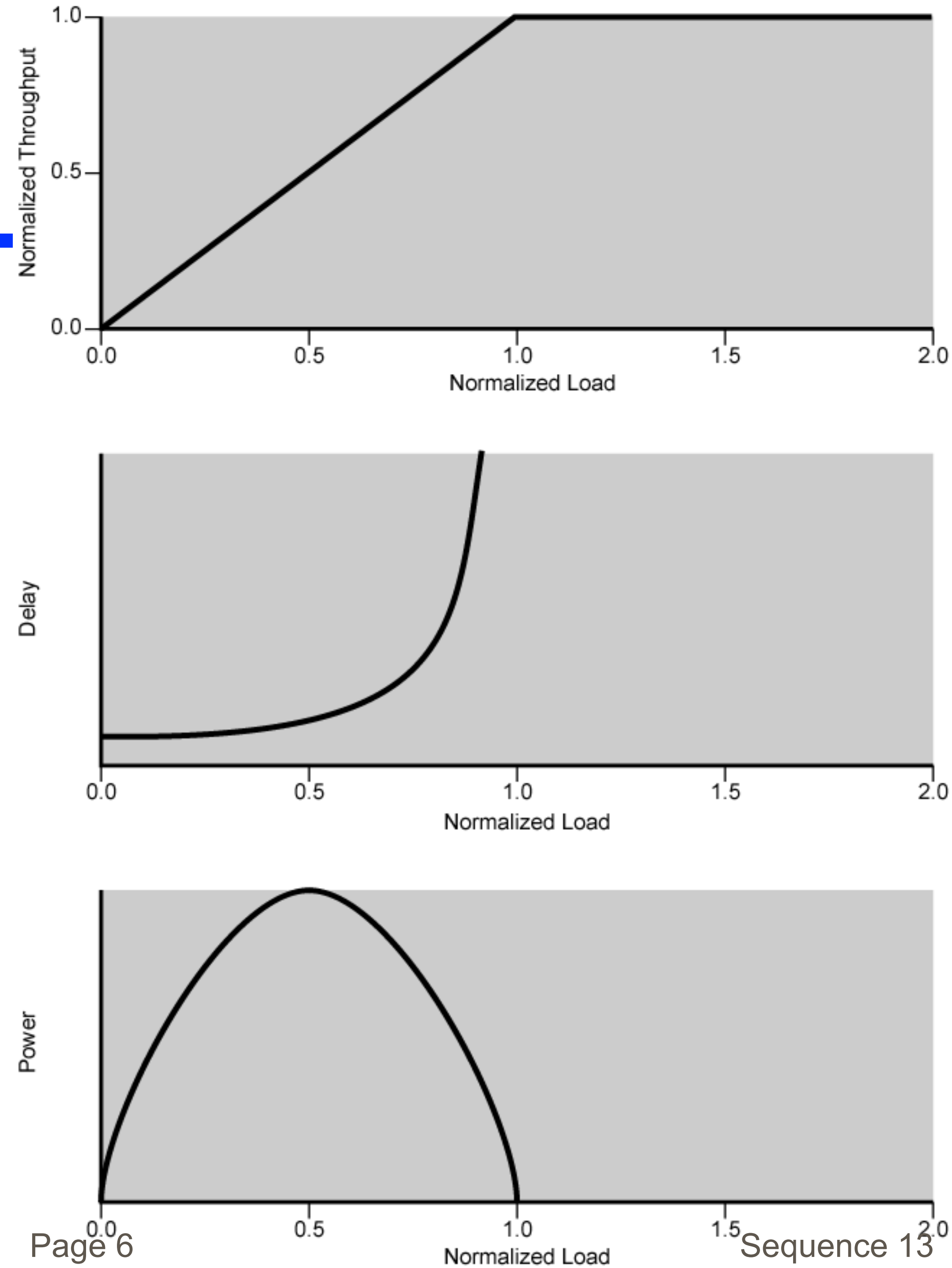
Effects of Congestion

- Packets arriving are stored at input buffers
- Routing decision made
- Packet moves to output buffer
- Packets queued for output transmitted as fast as possible
  - this is, in effect, 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

Interaction of Queues



Ideal Network Utilization

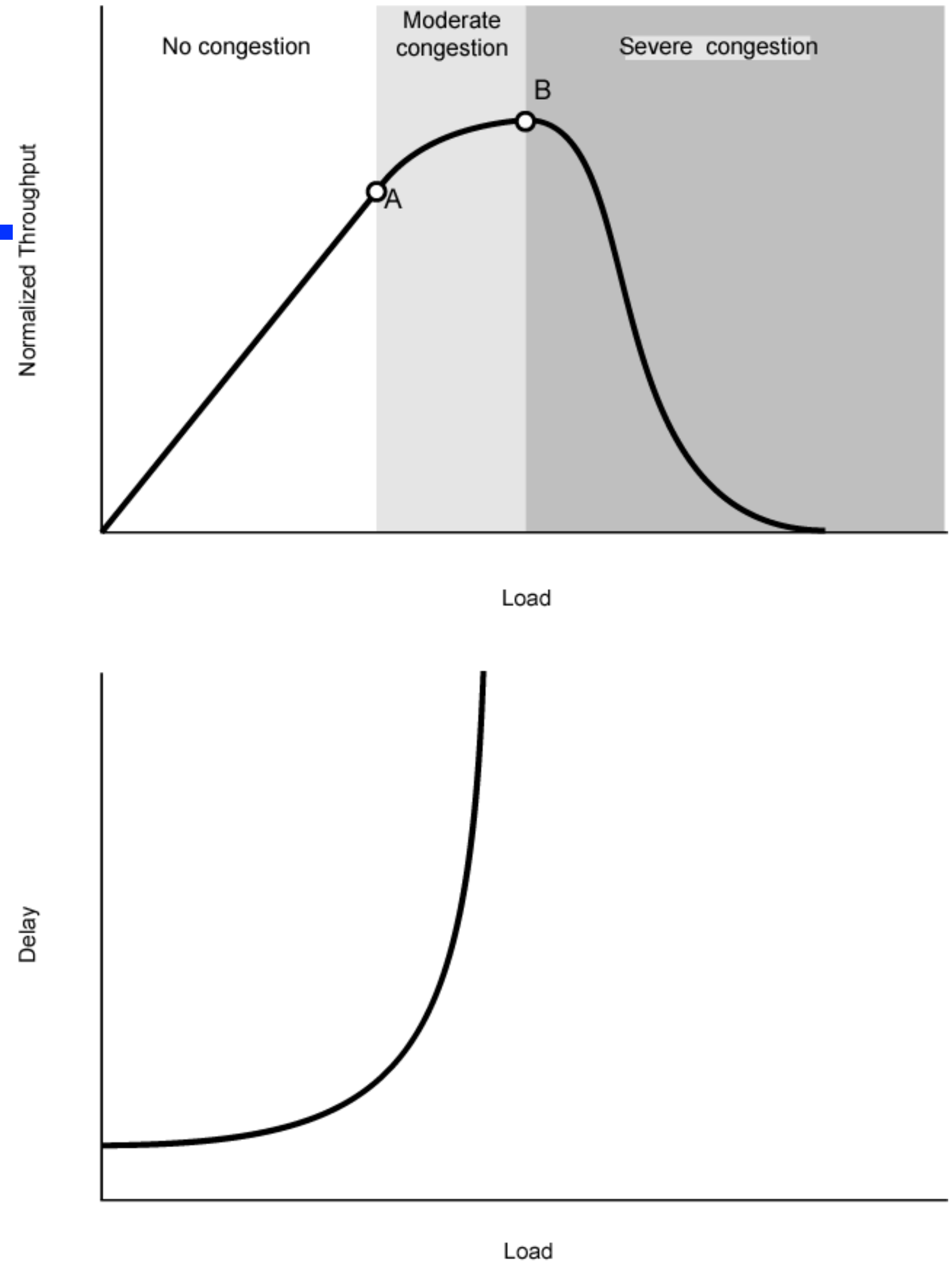


$$power = \frac{throughput}{delay}$$

Practical Performance

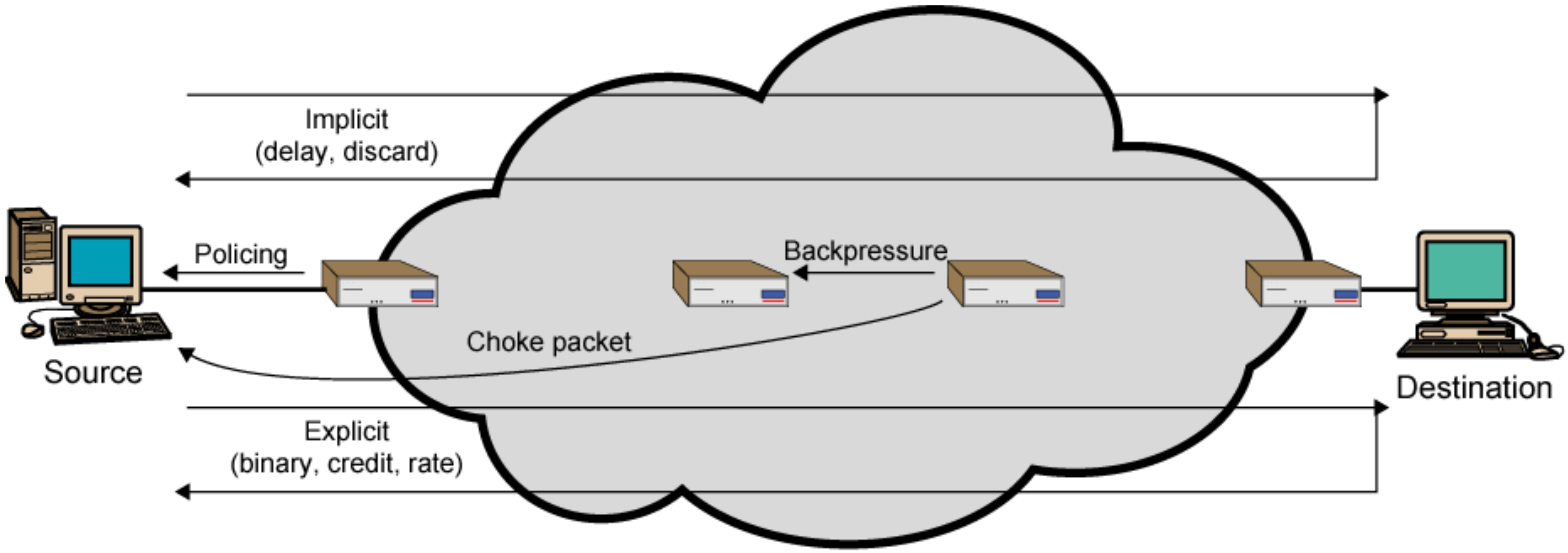
- Ideal
  - assumes infinite buffers and no overhead
- Realistic
  - buffers are finite
  - overheads occur in exchanging congestion control messages

Effects of Congestion - No Control



Note:
at point B, as queues become full and packets are dropped, retransmitted packets worsen the situation…

Mechanisms for Congestion Control



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 nor frame relay
- Only recently developed for IP

Choke Packet

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

Implicit Congestion Signaling

- Transmission delay may increase with congestion
- Packet may be discarded
- Source can detect these as implicit indications of congestion
- Useful on connectionless (datagram) networks
  - e.g. IP based
    - (TCP includes congestion & flow control is coming up later)
- Used in frame relay LAPF
    - recall LAPF:
      link access procedure for frame mode bearer service

Explicit Congestion Signaling

- Network alerts end systems of increasing congestion
- End systems take steps to reduce offered load
- Can work in one of two directions:
  - Backward
    - Notifies the source that congestion avoidance procedures should be initiated
      - Congestion avoidance in opposite direction to packet flow required.
  - Forward
    - Notify user that congestion avoidance procedures should be initiated
      - Congestion avoidance in same direction as packet required

Categories of Explicit Signaling

- 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
  - e.g. ATM

Traffic Management

- Fairness
  - provide equal treatment of various flows
- Quality of service (QoS)
  - May want different treatment for different connections
- Reservations
  - e.g. ATM
  - Traffic contract between user and network

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 as they cross nodes
  - Either backwards or forwards

Frame Relay Congestion Control

- I.370 objectives for frame relay congestion control:
  - Minimize discards
  - Maintain agreed QoS
  - Minimize probability of one end user monopoly
  - 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

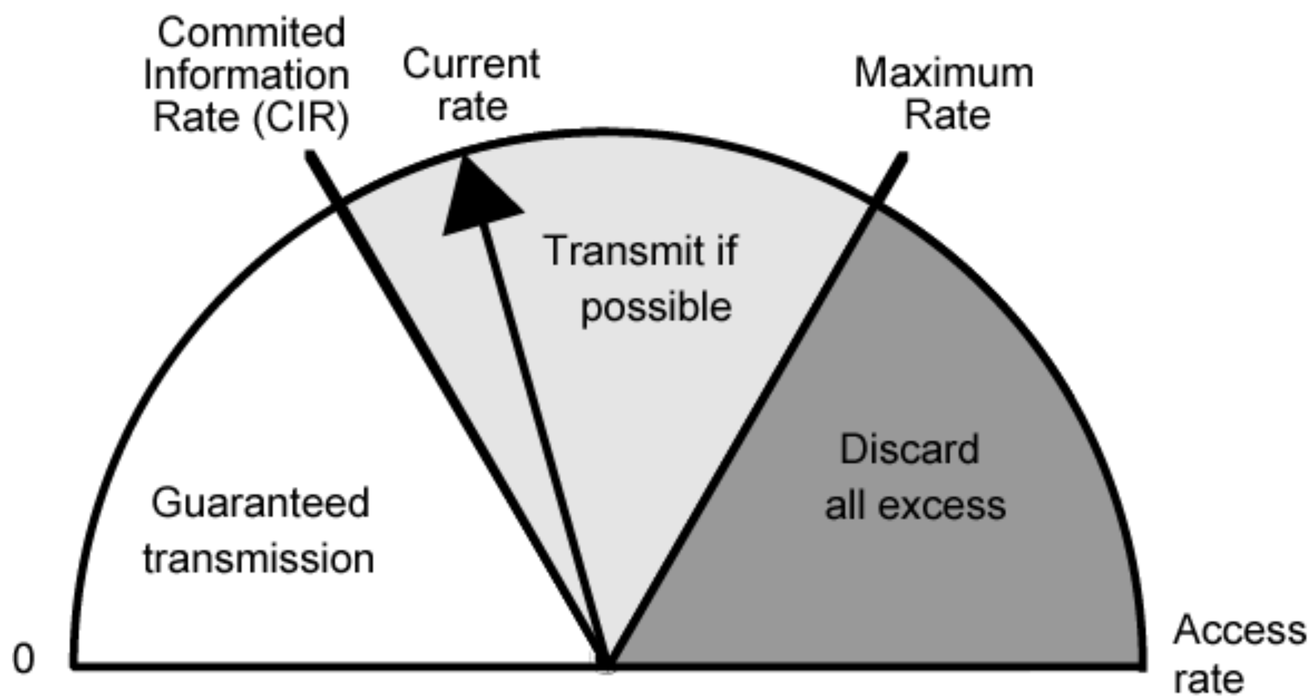
Techniques

- Discard strategy
- Congestion avoidance
- Explicit signaling
- Congestion recovery
- Implicit signaling mechanism

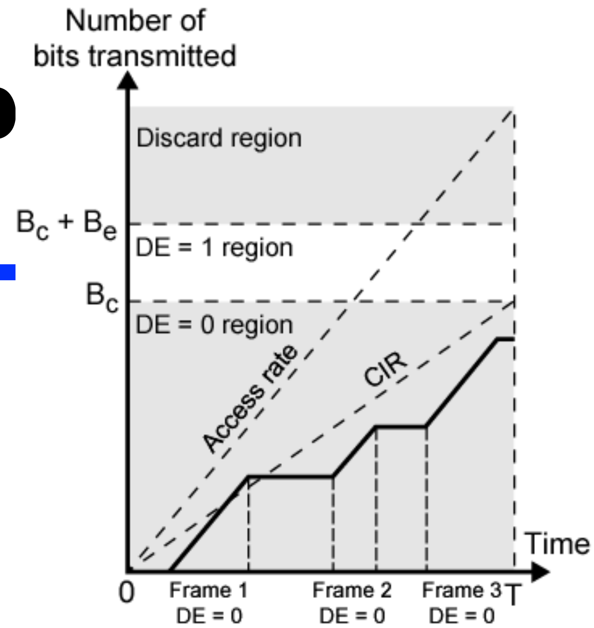
Traffic Rate Management

- 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
    - Aggregate CIR should not exceed physical data rate
- Committed burst size B_c
- Excess burst size B_e

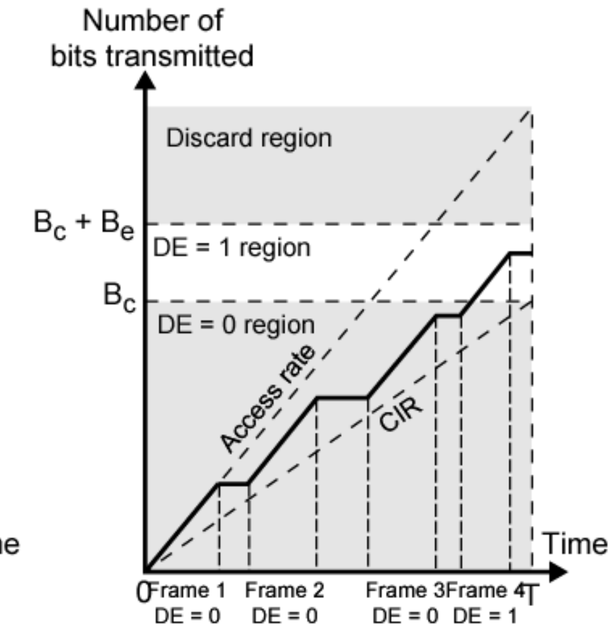
Operation of CIR



Relationship Among Congestion Parameters



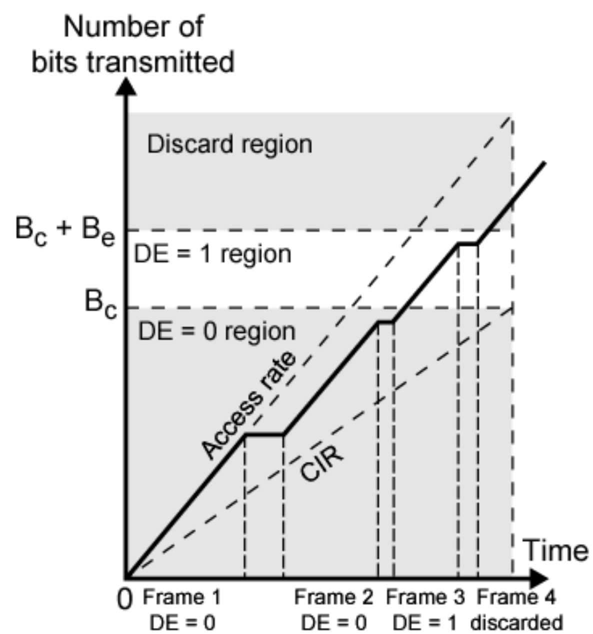
(a) All frames within CIR



(b) One frame marked DE

DE bit (Discard Eligibility)
used in the LAPF frame to discriminate
CIR: committed information rate

B_c = committed burst size
B_e = excess burst size, i.e. the max excess of B_c that the system will attempt to transfer



(c) One frame marked DE; one frame discarded

Explicit Signaling

- Network alerts end systems of growing congestion
- 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
- Requirements
  - 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,
    - e.g. constant vs. variable bit rate
  - Different network services
    - e.g. video, sound, file transfer
  - High speed switching and transmission increases volatility

Latency/Speed Effects

- Consider ATM 150Mbps
  - How long does it take to send single frame?
    - (53 x 8 bits)/(150 x 10^6 bps) ~2.8 x 10^{-6} seconds to insert single cell
- Time to traverse network depends on propagation delay, switching delay
  - Assume propagation at two-thirds speed of light
  - If source and destination on opposite sides of USA, round-trip 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
  - So, this is not a good strategy for ATM

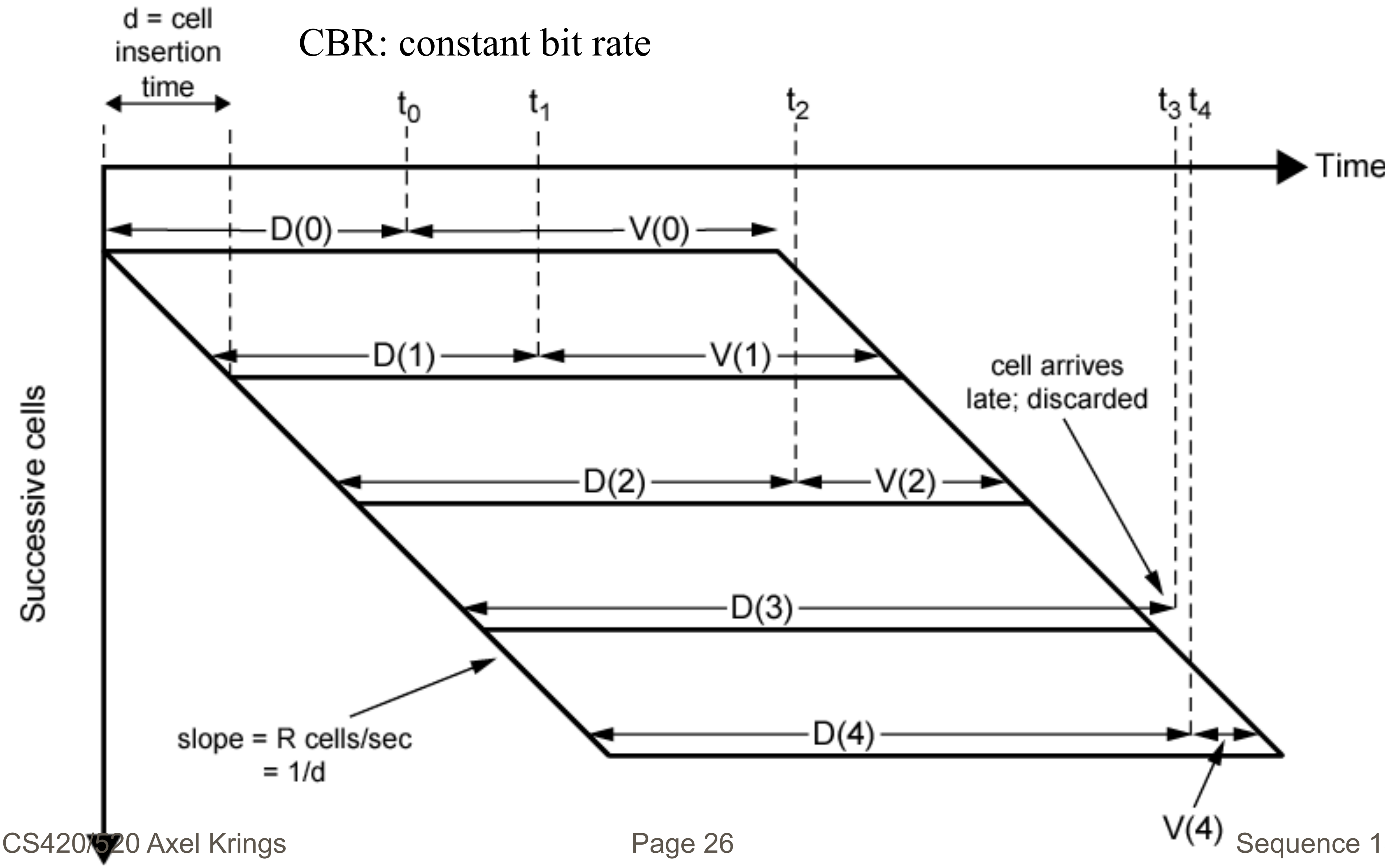
Cell Delay Variation

- 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

- D(i) represents end-to-end delay experienced by the i'th cell
- V(i) additional time that target user delays the i'th cell

Time Re-assembly of CBR Cells

CBR: constant bit rate



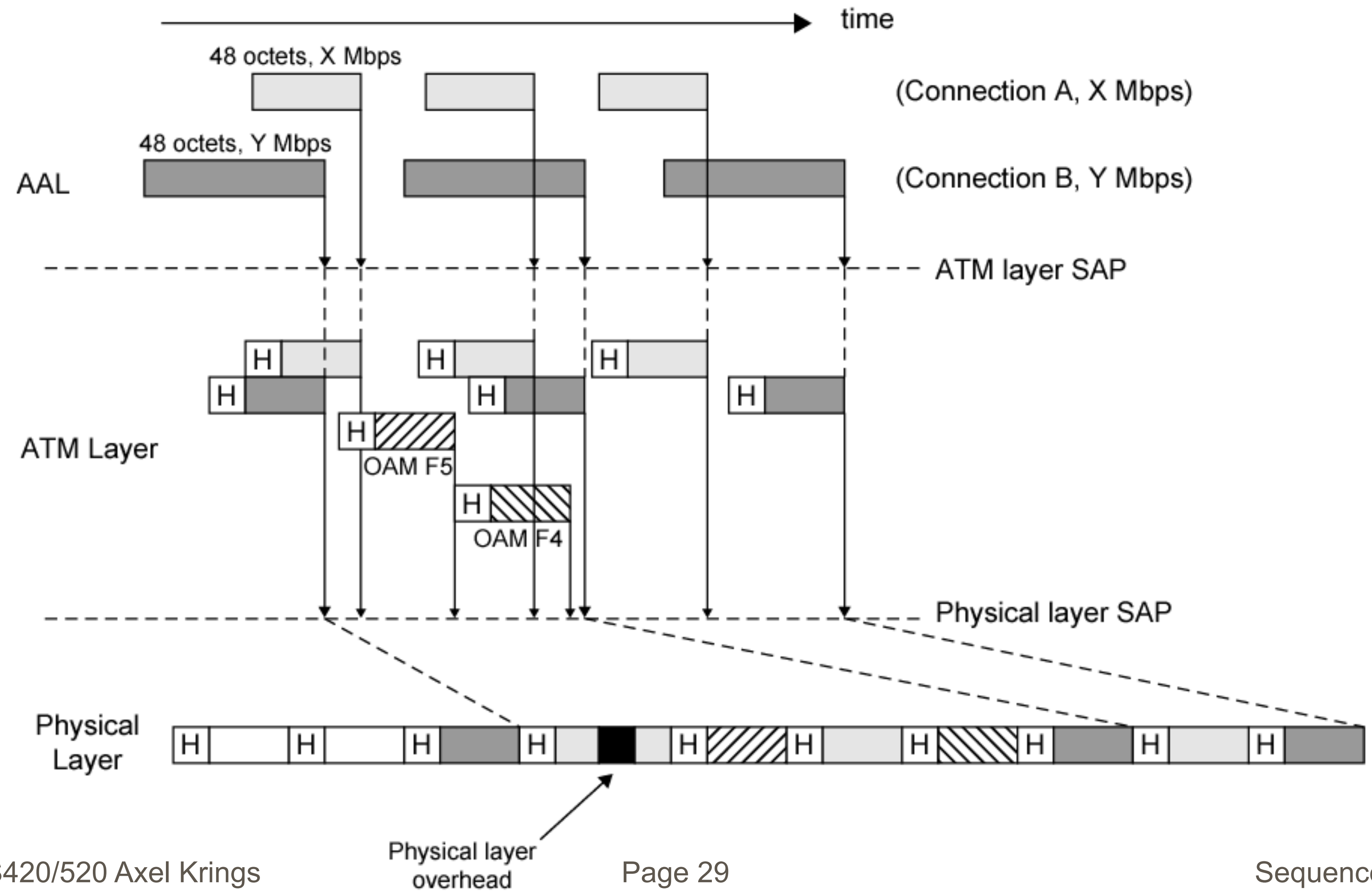
Network Contribution to Cell Delay Variation

- Packet switched networks
  - Queuing delays
  - Routing decision time
- Frame relay
  - As above but to lesser extent
- ATM
  - 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

- Application produces data at fixed rate
- Processing at three layers of ATM causes delay
  - —Interleaving cells from different connections
  - —Operation and maintenance cell interleaving
  - —If using synchronous digital hierarchy frames, these are inserted at physical layer
  - —Can not 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

- Congestion control functions time intervals
  - Cell insertion time
    - react immediately to cells as they are transmitted
  - Round trip propagation time
    - responds within life-time of a cell in the network
  - Connection duration
    - determine whether a new connection at a given QoS can be accommodated
    - agree on performance levels will be agreed to
  - Long term
    - affect more than one ATM connection
    - established for long-term use

Traffic Management and Congestion Control Techniques

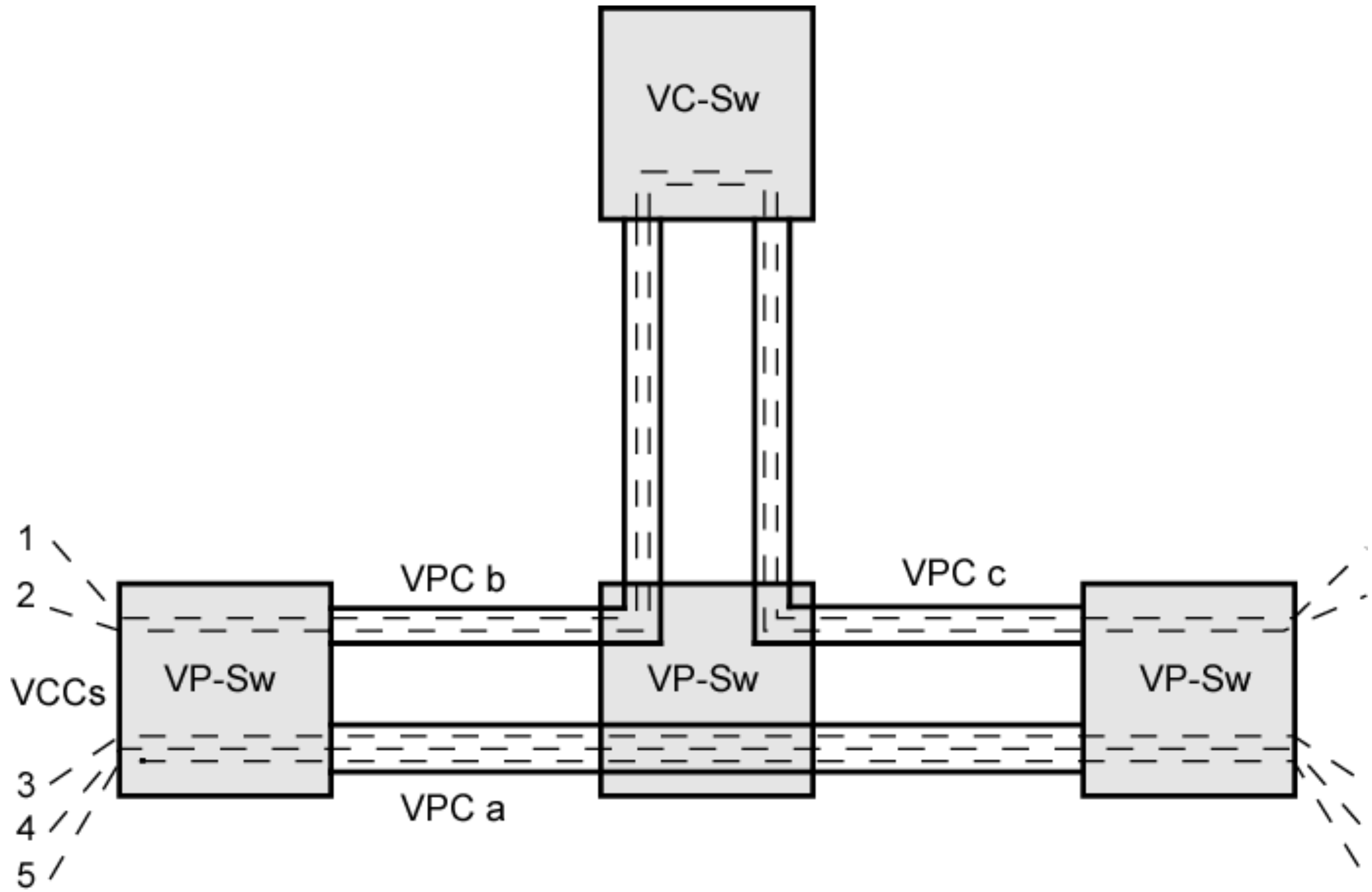
- Resource management using virtual paths
- Connection admission control
- Usage parameter control
- Selective cell discard
- Traffic shaping

Resource Management Using Virtual Paths

- Separate traffic flow according to service characteristics
- User to user application
- User to network application
- Network to network application

- Concern with:
  - Cell loss ratio
  - Cell transfer delay
  - Cell delay variation

Configuration of VCCs and VPCs



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
  - Statistical multiplexing

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

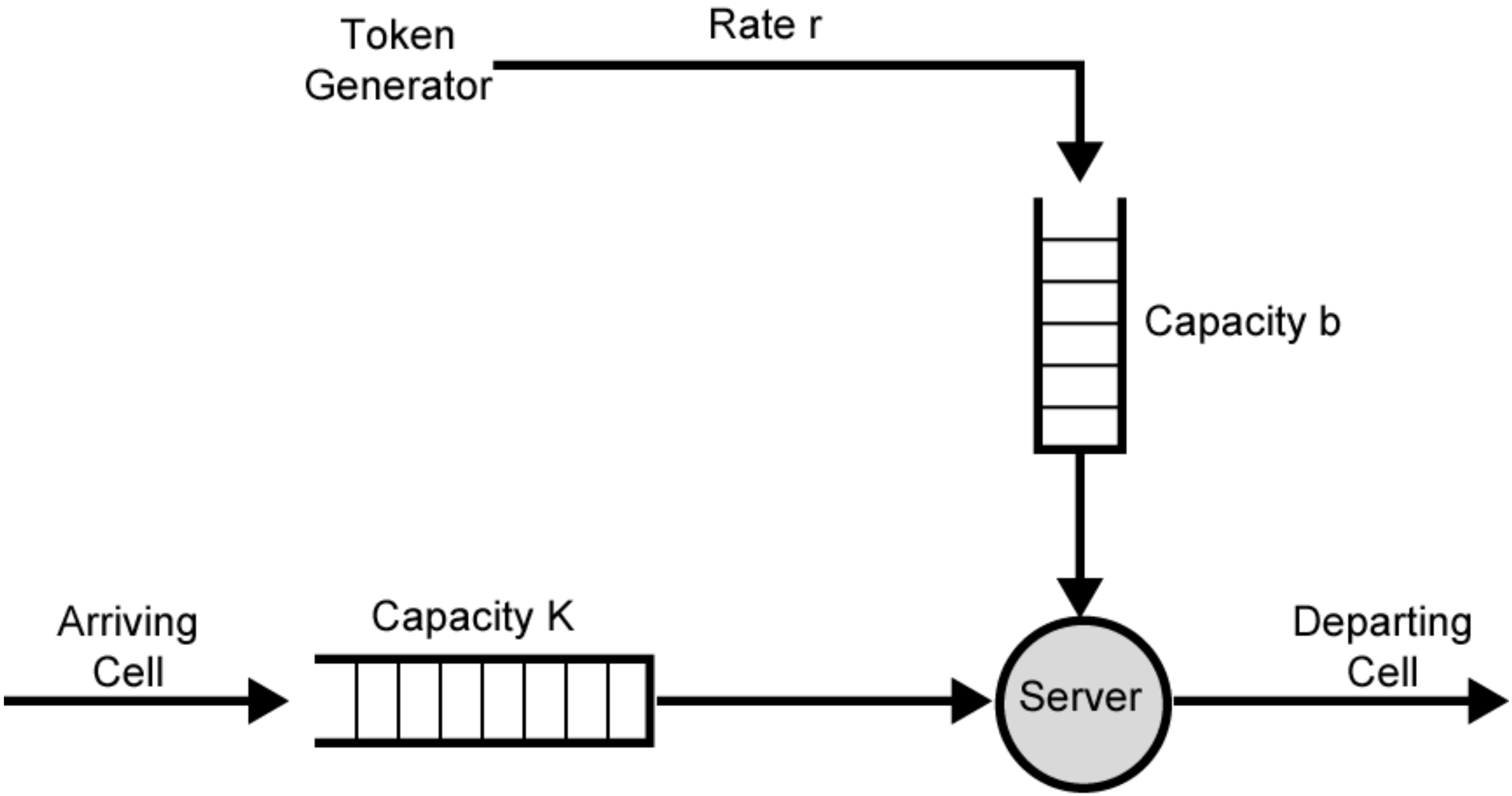
Usage Parameter Control

- Monitor connection to ensure traffic conforms to contract
- Protection of 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
- Discard cells that do not conform to traffic contract
- Called traffic policing

Traffic Shaping

- Smooth out traffic flow and reduce cell clumping
- Token bucket

Token Bucket for Traffic Shaping



GFR Traffic Management

- GFR: *Guaranteed Frame Rate* is as simple as *unspecified bit rate* (UBR) from end system viewpoint
- Places modest requirements on ATM network elements
- End system does no policing or shaping of traffic
- May transmit at line rate of ATM adaptor
- No guarantee of delivery
  - Higher layer (e.g. TCP) must do congestion control
- User can reserve capacity for each VC
  - Assures application may transmit at minimum rate without losses
  - 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 are lower priority
  - Best efforts
- 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)

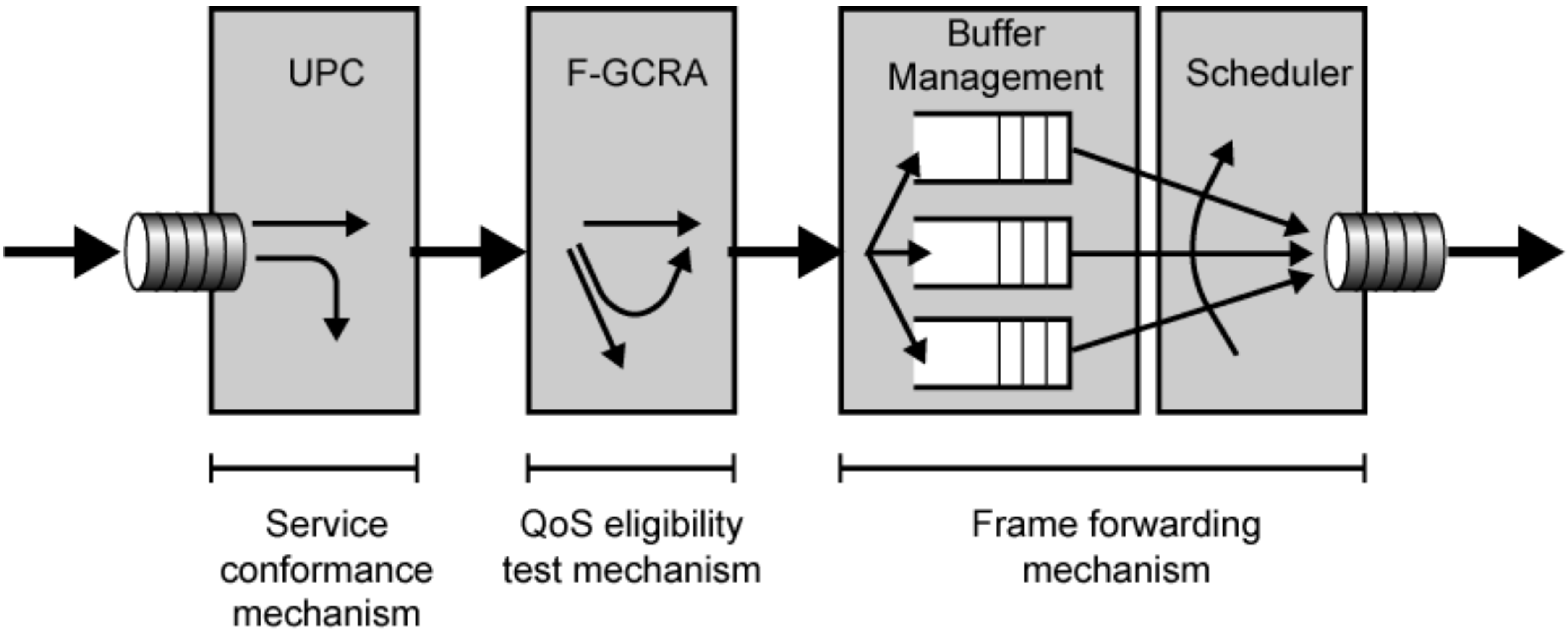
Mechanisms for Supporting Rate Guarantees (1)

- Tagging and policing
  - Discriminate 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
  - Treatment of buffered cells
  - Congestion indicated by high buffer occupancy
  - Discard tagged cells
    - Including ones already in buffer to make room
  - To be fair, per VC buffering
  - Cell discard based on queue-specific thresholds

Mechanisms for Supporting Rate Guarantees (2)

- Scheduling
  - Give 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

Components of GFR System



Conformance Definition

- UPC
  - Monitors each active VC
  - Ensure traffic conforms to contract
  - Tag or discard nonconforming cells
  - Frame conforms if all cells conform
  - Cell conforms if:
    - Rate of cells within contract
    - All cells in frame have same CLP
    - Frame satisfies MFS parameter (check for last cell in frame or cell count < MFS)

QoS Eligibility Test

- Two stage filtering process
  - Frame tested for conformance to contract
    - If not, may discard
    - If not discarded, tag
    - Sets upper bound
    - Penalize cells above upper bound
    - Implementations expected to attempt delivery of tagged cells
  - Determine frames eligible for QoS guarantees
    - Under GFR contract for VC
    - Lower bound on traffic
    - Frames making up traffic flow below threshold are eligible

GFR VC Frame Categories

- Nonconforming frame
  - Cells of this frame will be tagged or discarded
- Conforming but ineligible frames
  - Cells will receive a best-effort service
- Conforming and eligible frames
  - Cells will receive a guarantee of delivery

Summary

- congestion effects
  - ideal and practical performance
- congestion control
  - backpressure, choke packet, implicit/explicit
- traffic management
  - fairness, QoS, reservations
- ATM traffic management
- ATM-GFR traffic management
  - tagging, policing, buffer, scheduling
- QoS eligibility testing