第13章 数据网中的拥塞控制

南京大学计算机系　黄皓教授

2007年11月9日 星期 五

Chapter 13. Congestion in Data Networks

1. Concept of Congestion
2. Congestion Control
3. Congestion Control in Frame Relay Network
4. Congestion Control in ATM Network

1. Concept of Congestion

1. Concept of Congestion

- If， for any time interval，the total sum of demands on a resource is more than its available capacity，the resource is said to be congested for that interval

$$\sum \textbf{Demand > Available Resources}$$

- In computer networks，there are a large number of resources，such as **buffers**，**link bandwidths**，**processor times**，**servers**，and so forth.
- If，for a short interval，the buffer space available at the destination is less than that required for tile arriving traffic，packet loss occurs.
- Similarly，if the total traffic wanting to enter a link is more than its bandwidth，the link is said to be congested.

Congestion

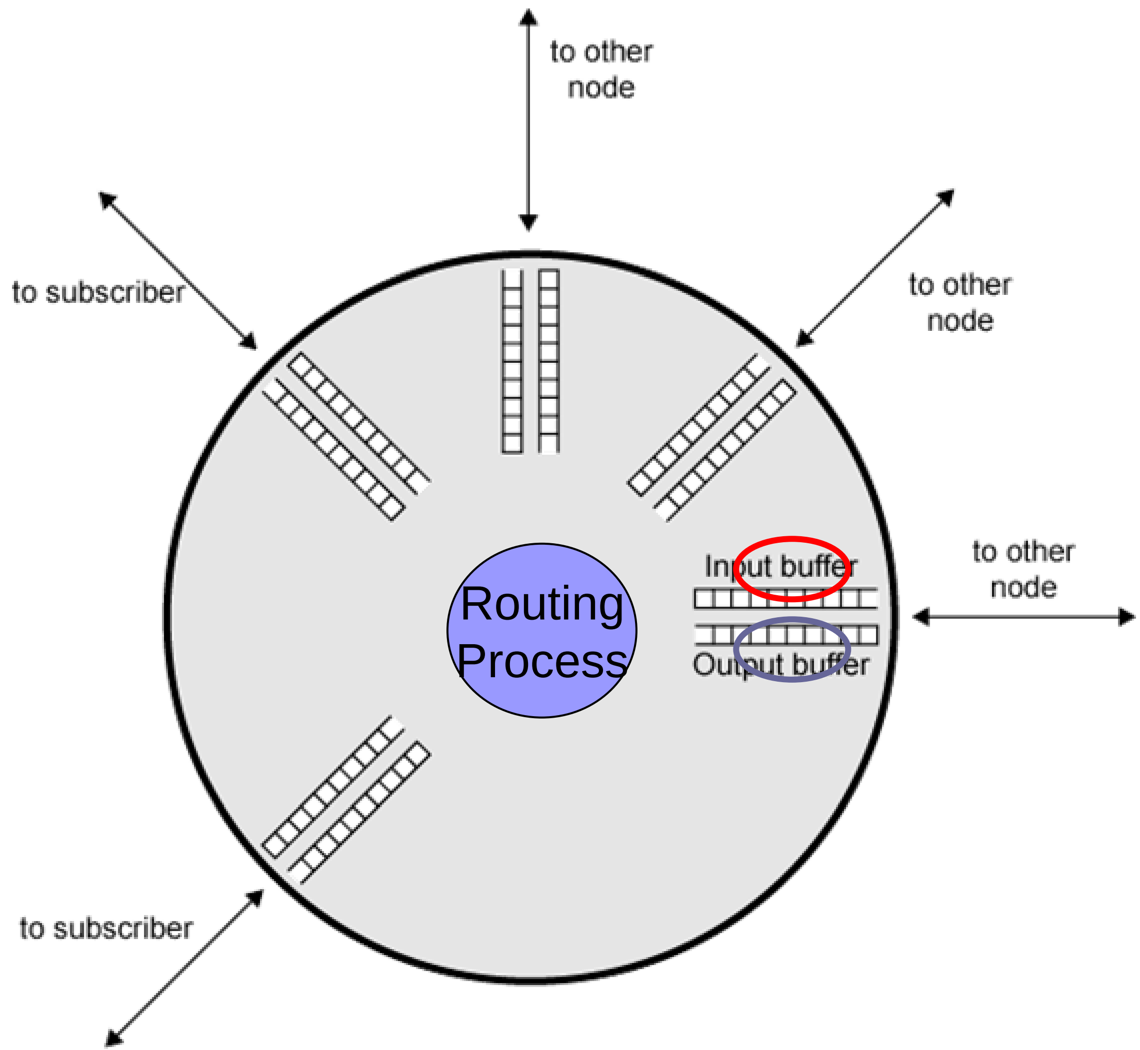
- Congestion
  - **Number of packets** transmitted through the network **approaches the packet handling capacity** of the network

$$L_q = \rho \frac{\lambda}{\mu - \lambda}$$

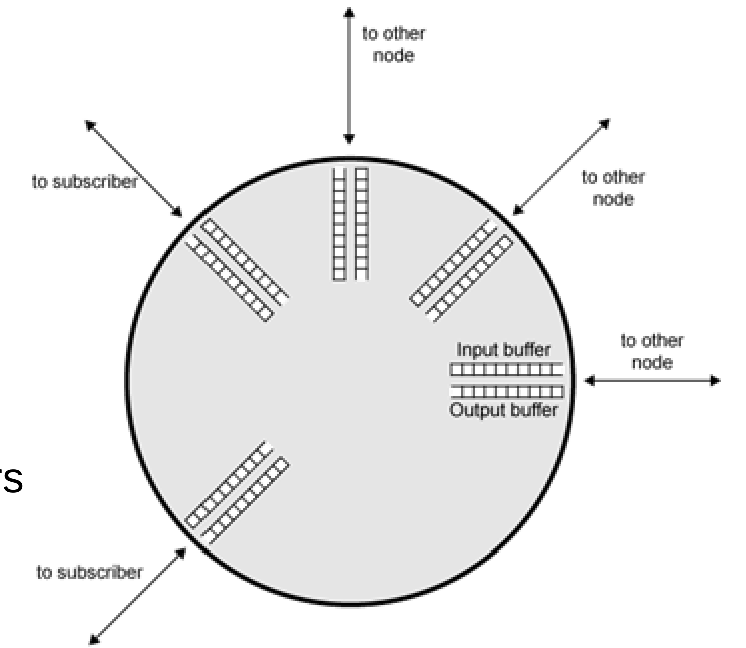
  - Generally **80%** utilization is critical
- Congestion control
  - **Keep number of packets below level** at which performance falls off dramatically
- Data network is a network of queues
  - Finite queues mean data may be lost

Queues at a Node

- Input
- Store
- Input queueing
- Move
- Routing process
- Output queueing
- Transmit

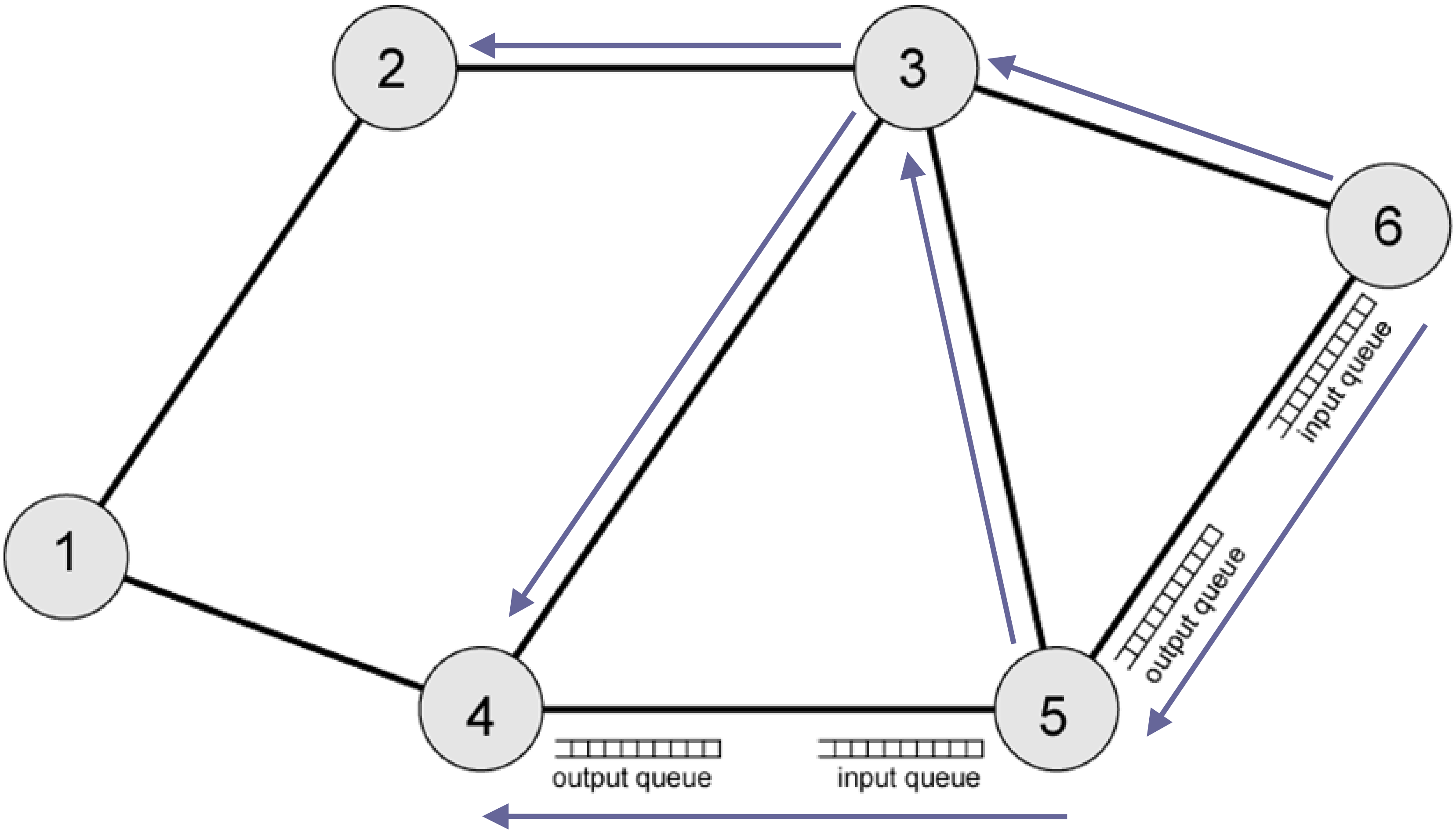


Effects of Congestion



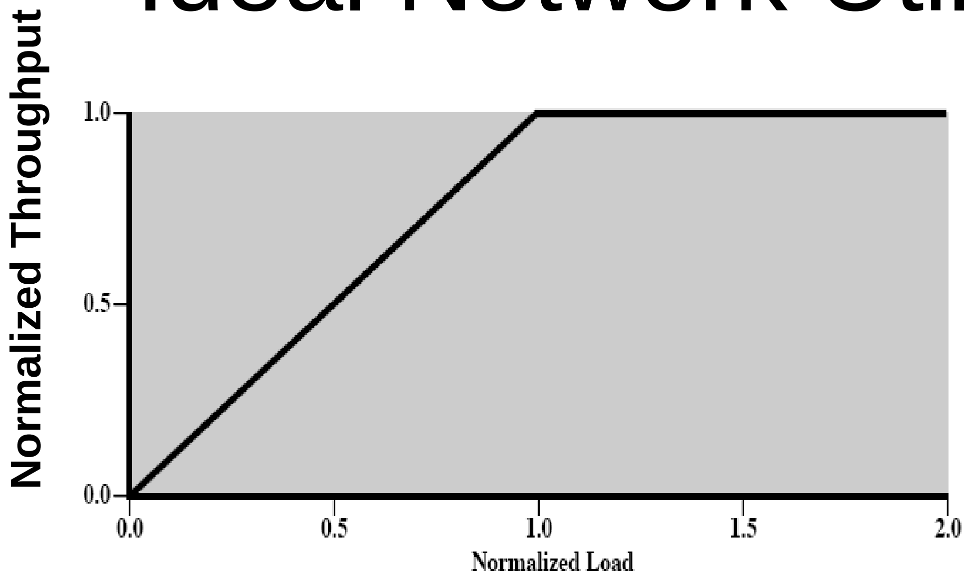
- **Input Buffer**
  - Packets arriving are stored at input buffers
  - Wait for routing decision made
- **Output Buffer**
  - Routed packet moves to output buffer
  - Packets queued for output transmitted as fast as possible
    - Statistical time division multiplexing
- Congestion
  - Packets arrive too fast to be routed, or to be transmitted
  - Can **discard packets**
  - Can **propagate congestion** through network

Interaction of Queues

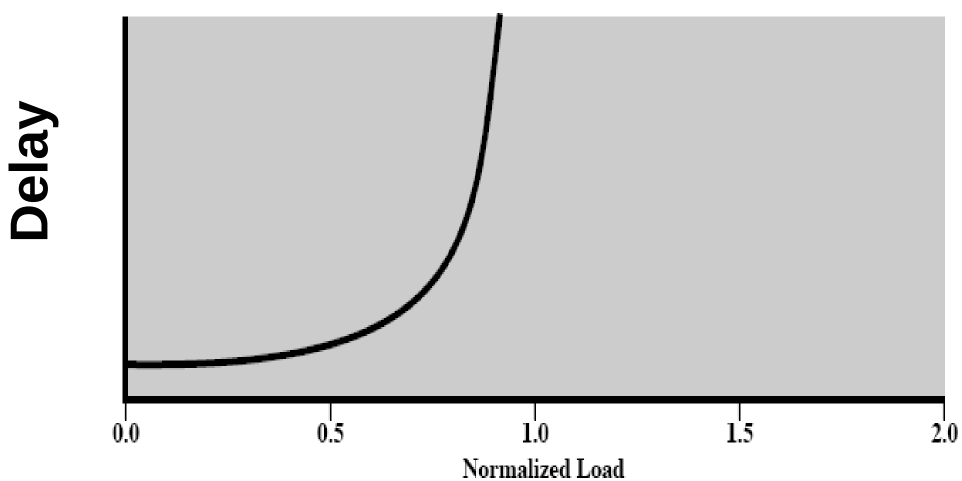


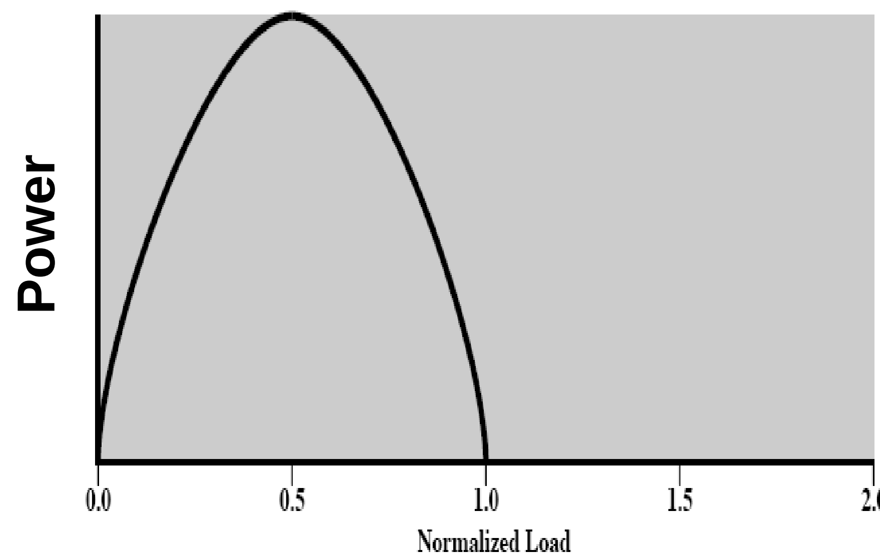
Congestion propagation

Ideal Network Utilization



$$Power = \frac{Throughput}{Delay}$$



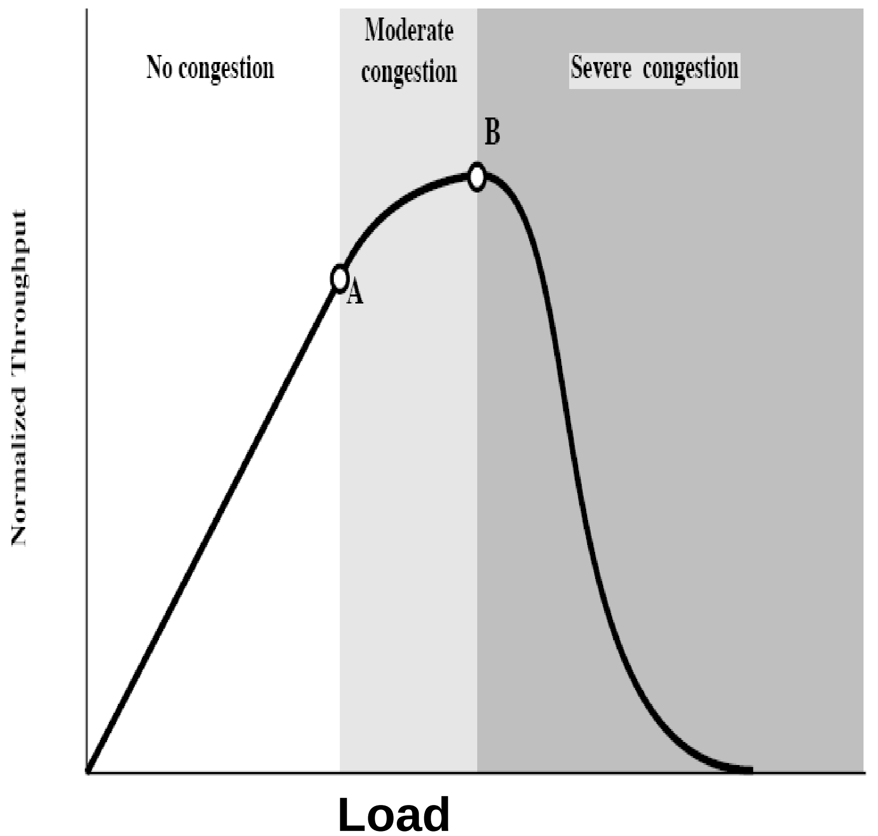


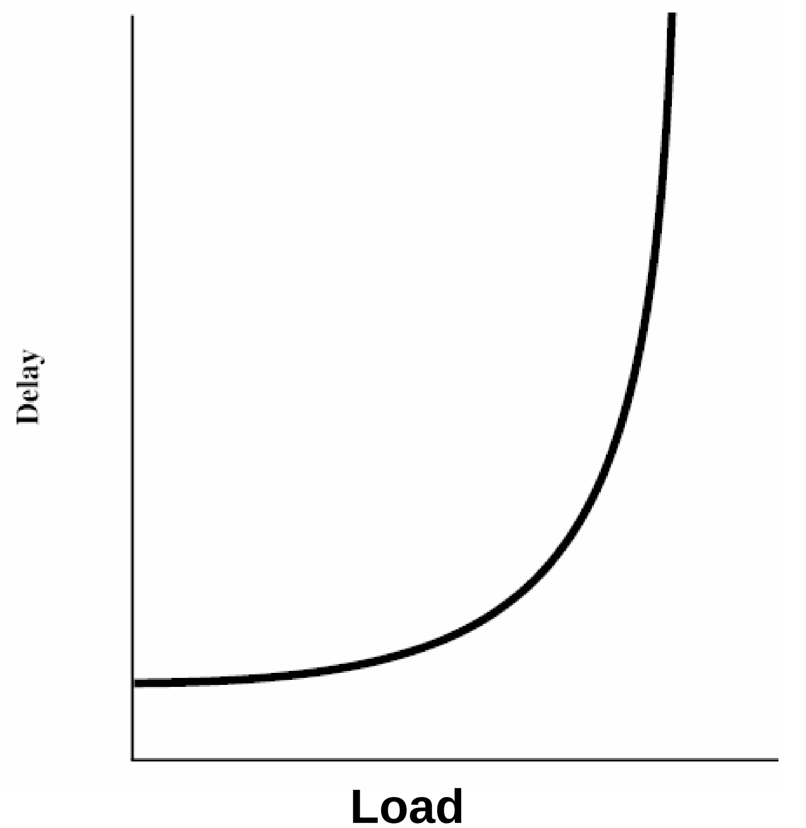
Practical Performance

- Ideal assumes
  - infinite buffers and no overhead
- Practical situation
  - Buffers are finite
  - Overheads occur in exchanging congestion control messages

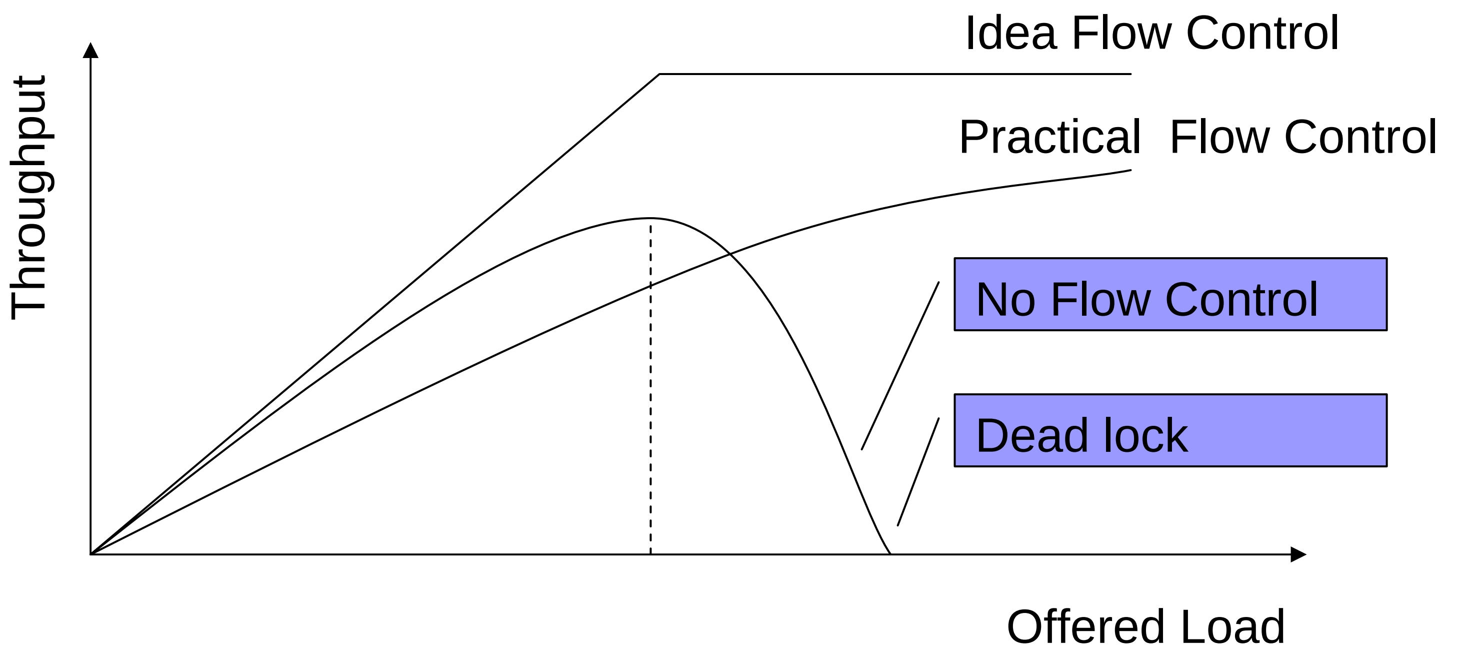
Effects of Congestion

- Start congestion control at point A
- Prevent critical level at point B





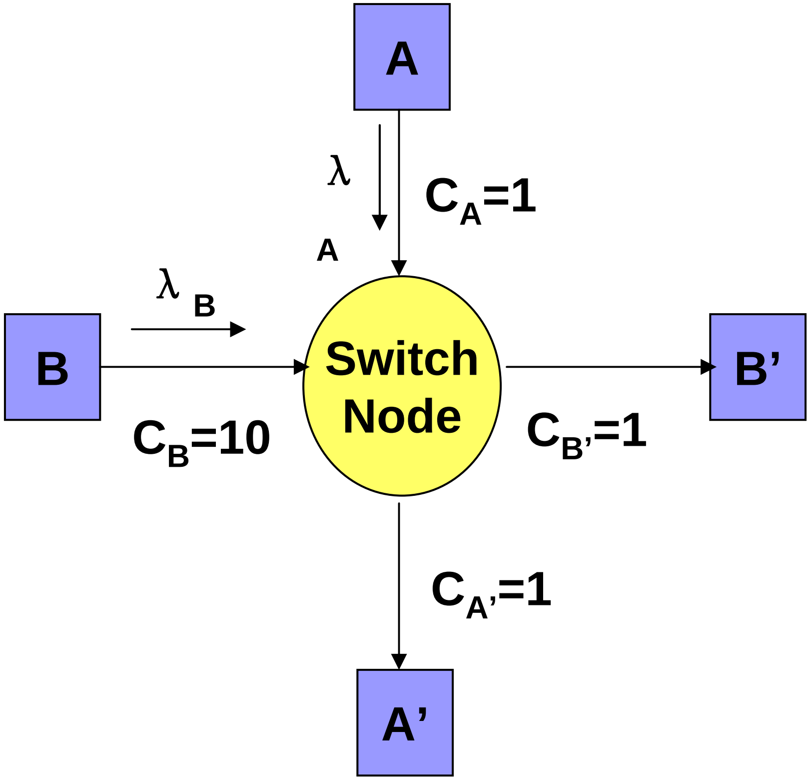
Effects of Congestion



网络的资源被浪费

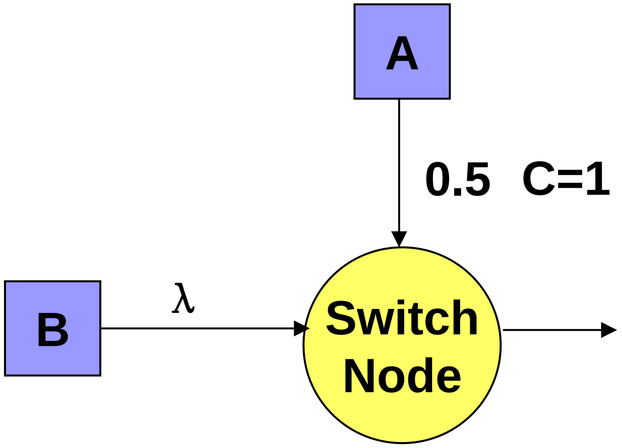
- 如果 $\lambda_A = \lambda_B = 0.8$, 总的吞吐量为1.6。
- 当通信量增大到使主机A和B到交换结点的两条链路都饱和时

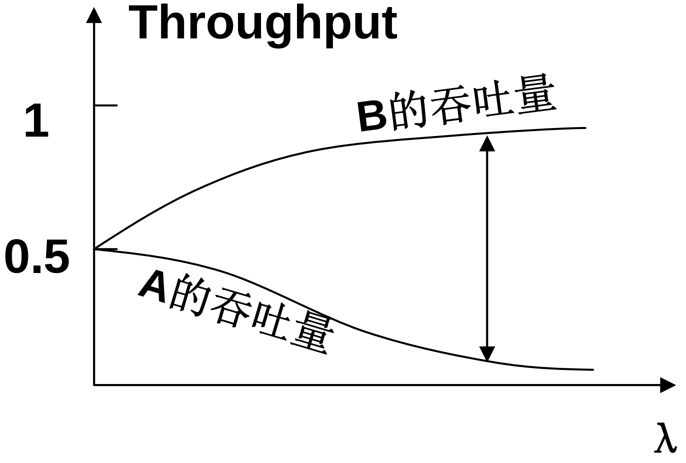
  $\lambda_A = 1$, $\lambda_B = 10$
- 这时在交换结点的等待发送队列长度均会趋于无限大。
- 主机B与主机A相比，它们占有输入缓冲区的机会之比应为10:1。
- 因为最终从交换结点输出到主机B’的数据率不超过1，按上述10:1的关系，可知从交换结点输出到主机A’的数据率不超过0.1。
- 于是整个网络的吞吐量下降到1.1。
- 必须寻求一种有效的方法来控制进入网络的通信量和管理好交换结点的缓冲空间。



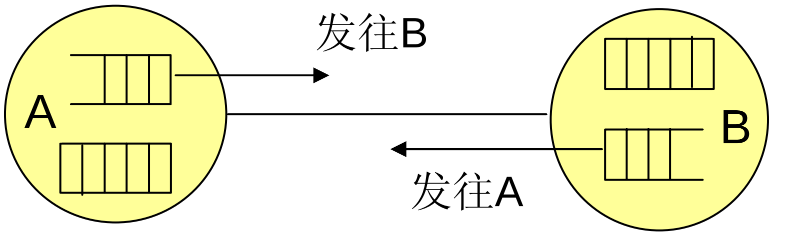
公平性问题

- A数据率为0.5，B数据率为λ。
- 通过交换结点向容量为**C=1**的链路输出数据。
- 当λ **<0.5**时，网络的运行是合理的。
- 特别是当λ **>> 0.5**时，网络以很不公平的方式运行。
- 在缓冲空间中主机**B**的分组和主机**A**的分组的数目之比约为**A:0.5**。
- 因此在总的吞吐量中，主机**B**和主机**A**的吞吐量之比亦约为λ **:0.5**。





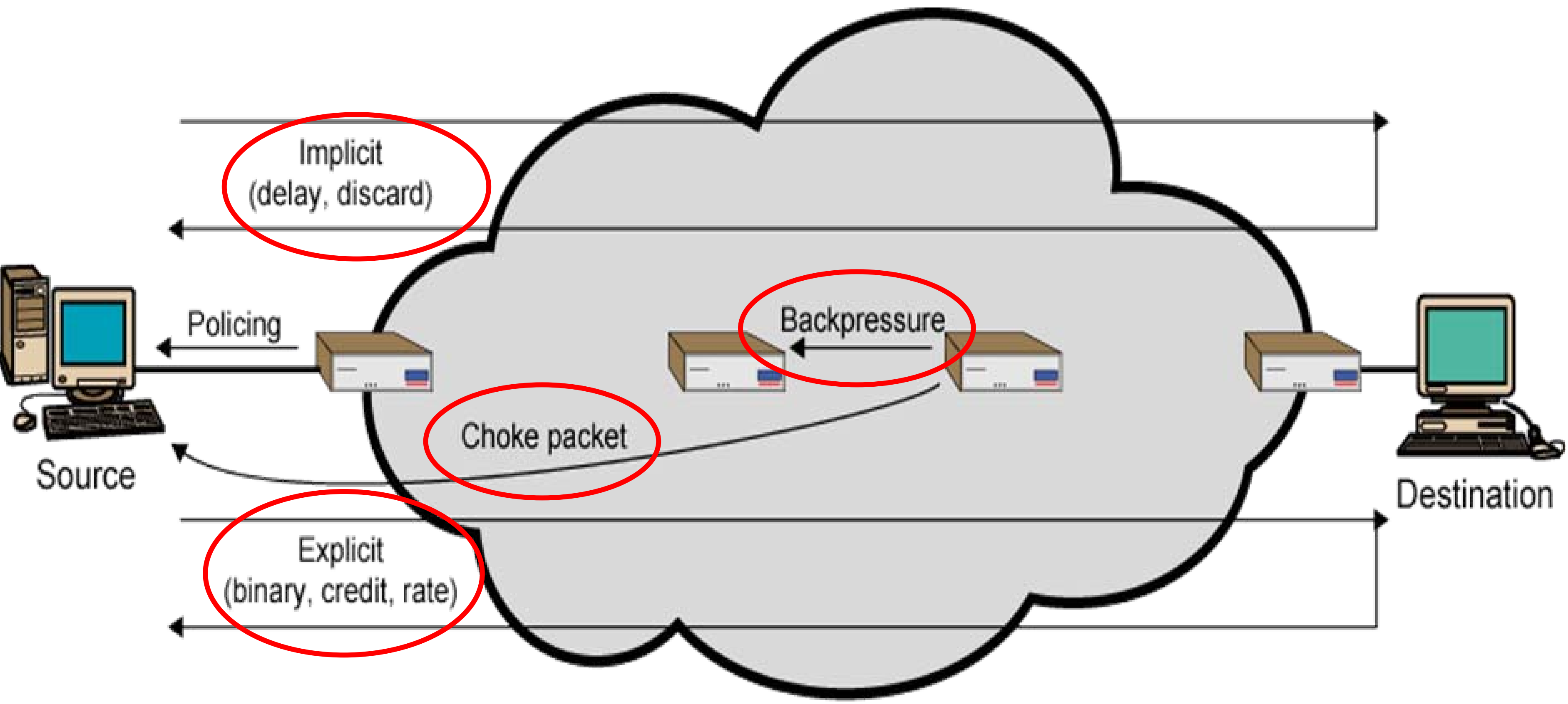
死锁问题



- 两个结点A和B都有大量的分组要发往对方。
- 两个结点中的缓冲区在发送之前就已经全部被占满了。
- 当每个分组到达对方时，由于没有地方存放，只好被丢弃。
- 发送分组的一方因收不到对方发来的确认信息，只能将发送过的分组依然保存在自己结点的缓冲区中。
- 这两个结点就这样一直互相僵持着，谁也无法成功地发送出一个分组。

2. Congestion Control

Mechanisms for Congestion Control



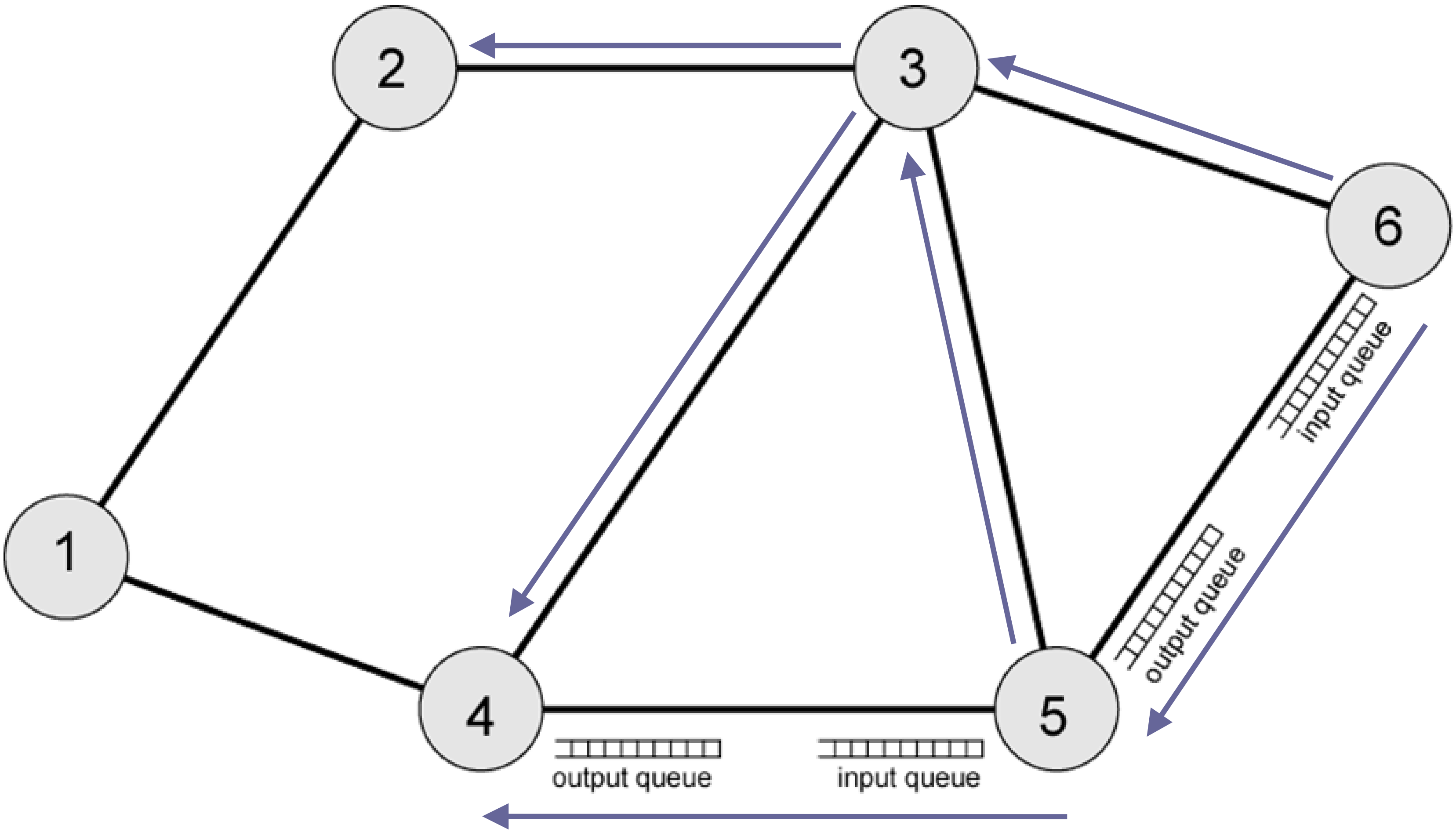
Possible Mechanisms

- Backpressure
- Choke Packet
- Implicit Congestion Signaling
- Explicit Congestion Signaling

Backpressure

- Concept
  - 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
- Usability
  - 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

Backpressure



Congestion propagation

Choke Packet

- Control packet
  - Generated at **congested node** or **destination**
  - 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 based on buffer utilization
- Rather **crude mechanism**

Implicit Congestion Signaling

- Concept
  - **Transmission delay may increase** with congestion
  - Packet may be timeout
  - Source can detect these as **implicit indications** of congestion
- Usability
  - Useful on **connectionless** (datagram) networks
    - e.g. IP based, TCP includes such a congestion and flow control mechanism
  - Used in frame relay LAPF

Explicit Congestion Signaling

- Concept
  - Network **alerts end systems** of increasing congestion
  - End systems take steps to reduce offered load
- Backwards
  - Congestion avoidance **in opposite direction** to packet required
  - Assume congestion will **burst up quickly**
- Forwards
  - Congestion avoidance **in same direction** as packet required
  - Assume congestion will **cumulate slowly**

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

Policies Affecting the Congestion Control Scheme

- Daca Link Layer
  - Data link level retransmission policy
  - Data link level queuing and service policy
  - Data link level packet drop policy
  - Data link level acknowledgment policy
  - Data link level flow control pollicy

Requirement of DataLink Management

- Frame Synchronization.
- Flow Control
- Error Control
- Addressing
- Link Management

Flow Control — Stop and Wait

- Source transmits frame
- Destination receives frame and replies with acknowledgement
- Source waits for ACK before sending next frame
- Destination can stop flow by not send ACK
- Works well for a few large frames

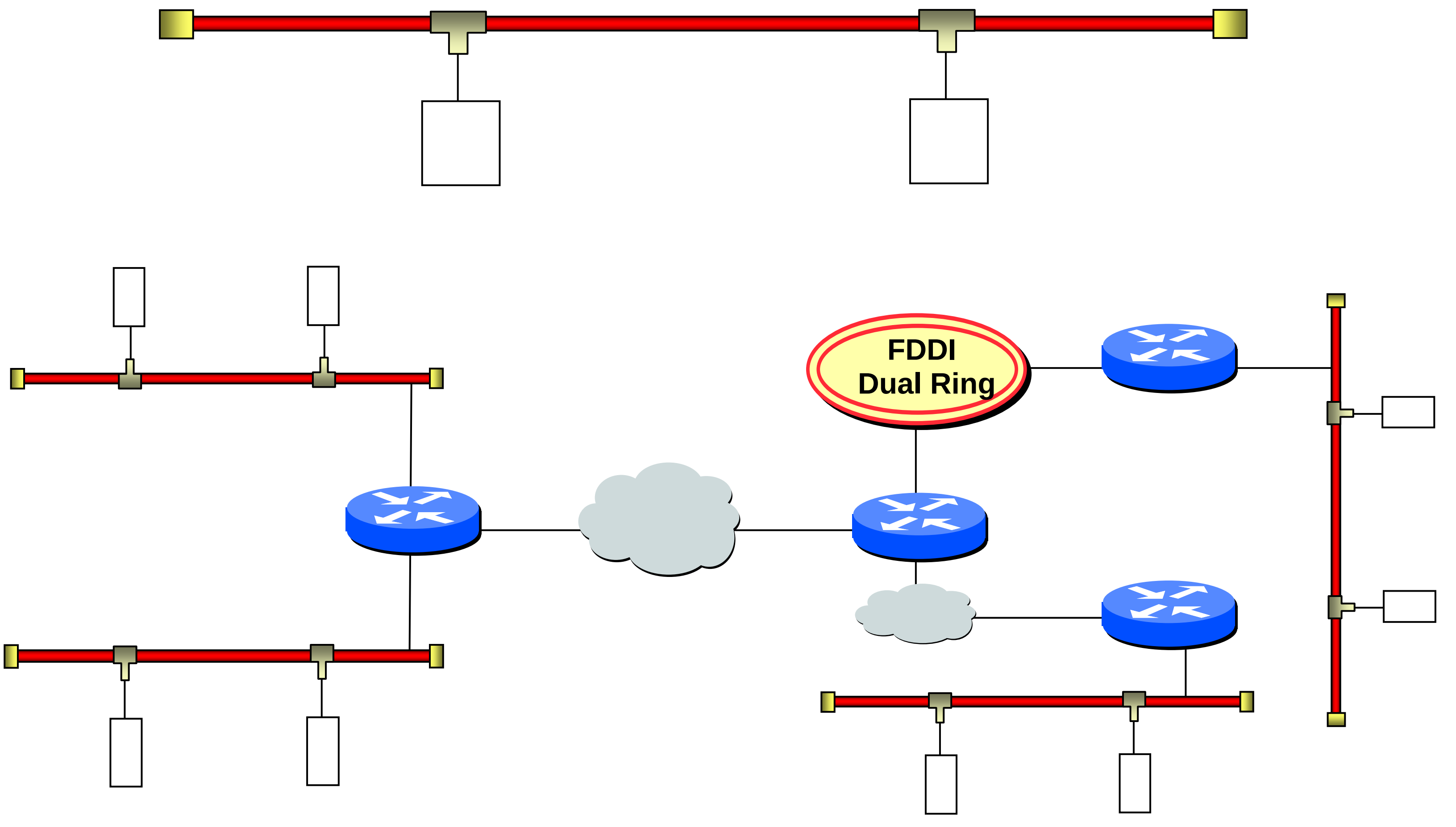
Flow Control — Fragmentation

- Large block of data may be split into small frames
  - Limited buffer size
  - Errors detected sooner (when whole frame received)
  - On error, retransmission of smaller frames is needed
  - Prevents one station occupying medium for long periods
- Stop and wait becomes inadequate

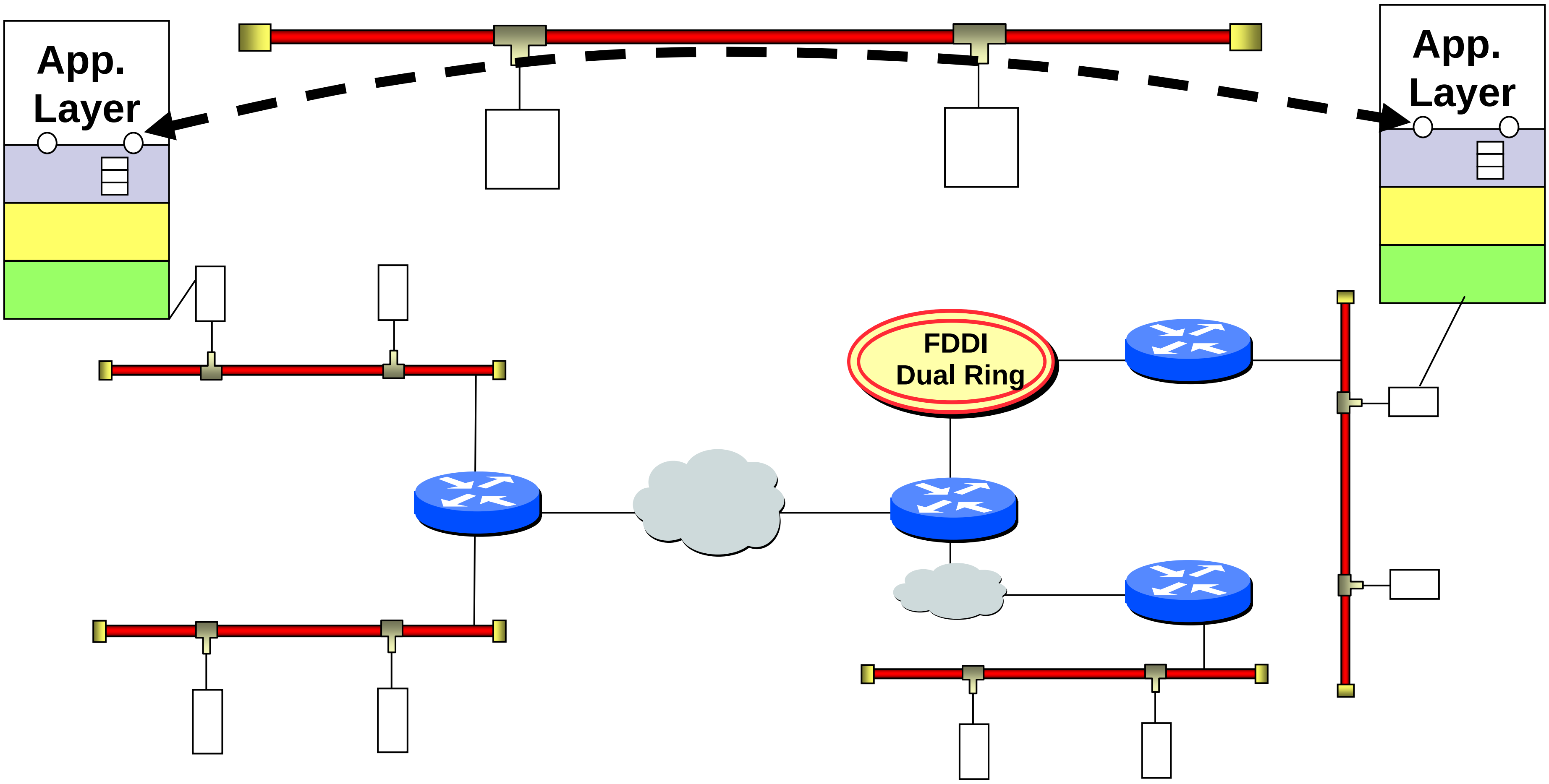
Flow Control — Sliding Window Enhancements

- Receiver can acknowledge frames without permitting further transmission (Receive Not Ready)
- Must send a normal acknowledge to resume
- If duplex, use piggybacking
  - If no data to send, use acknowledgement frame
  - If data but no acknowledgement to send, send last acknowledgement number again, or have ACK valid flag (TCP)

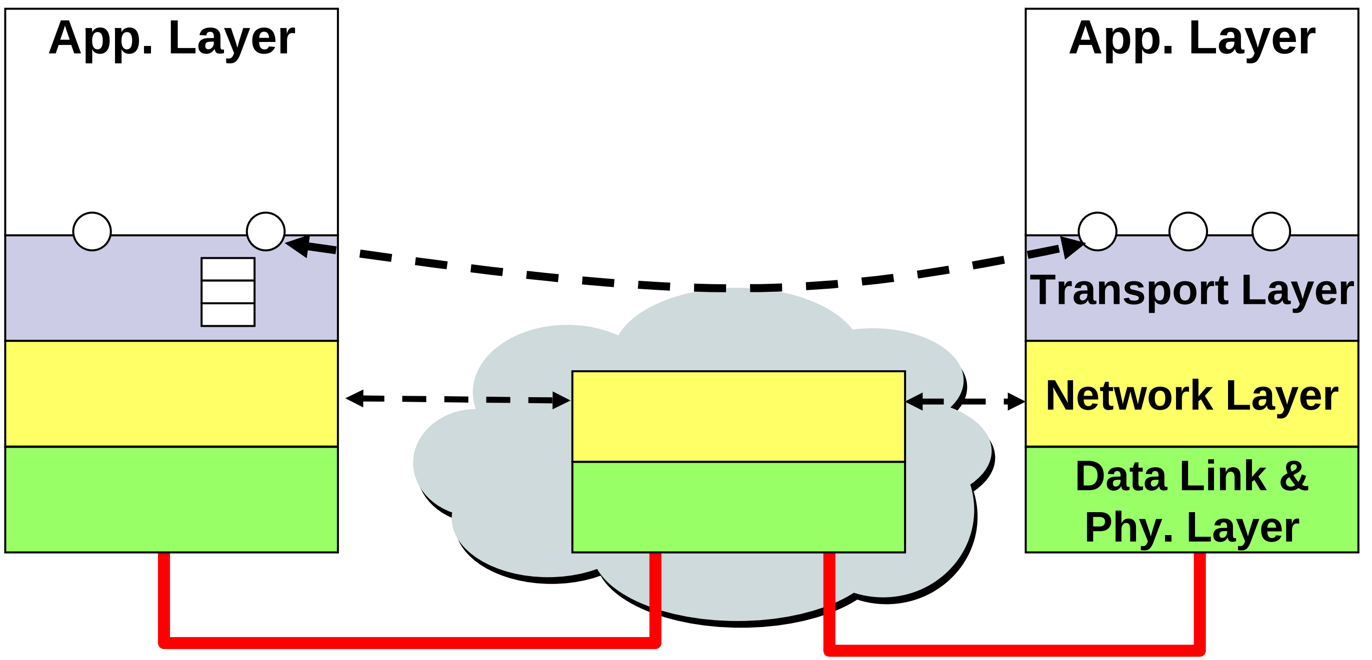
Requirement of Network Layer Flow Control



Requirement of Network Layer Flow Control



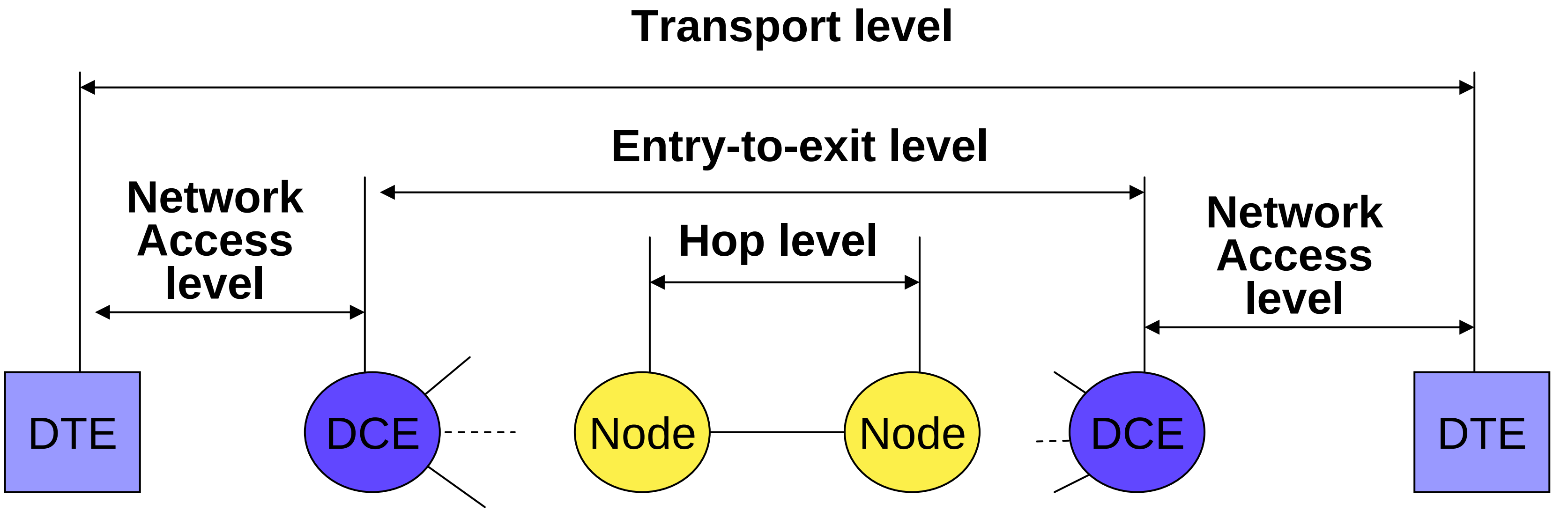
Transport Layer Flow Control



Policies
Affecting the Congestion Control Scheme

- Network Layer
  - Connection mechanism
  - Packet queuing and service policy
  - Packet drop policy
  - Packet routing policy
  - Lifetime control policy
- Transport Layer
  - Round trip delay estimation algorithm
  - Timeout algorithm
  - Retransmission policy
  - Out of order packet caching policy
  - Acknowledgment policy
  - Flow control policy
  - Buffer management policy

按级进行流量控制


Transport level
Entry-to-exit level
Network Access level
Hop level
Network Access level
DTE
DCE
Node
Node
DCE
DTE

段级流量控制

- 段级流量控制是根据网络中局部范围内(一个段的范围内)的通信量来进行控制。
- 这种控制方法是在每个结点监视本地的队列和缓冲区的占用情况。
- 当达到某些事先规定好的门限时(例如，队列的长度)，就将到达此结点且需要转发的分组丢弃。
- 这种检查缓冲区门限以及丢弃(然后发方再重传)分组的功能通常由数据链路层协议来完成。

段级流量控制

- 这种方式称为链路段级流量控制
- 监视每一个队列的缓冲区的占用情况。
- 对每一队列定义一个极限(固定的或动态可调的)。
- 当某一队列中的待转发的分组数超过此极限时，即丢弃该分组。

段级流量控制

- 全部分配
  - 设N为结点中输出队列的总数，n_i为第i个队列中的分组数，而B为缓冲区的大小(单位为分组数)，则约束为：

    对任意的I,　$0 \leqslant n_i \leqslant B/N$
- 最长队列共享
  - 设b_{max}为所允许的最长队列(此处$b_{max} > B/N$)，则我们的约束为

    $0 \leqslant n_i \leqslant b_{max}$，并且$\Sigma n_i \leqslant B$

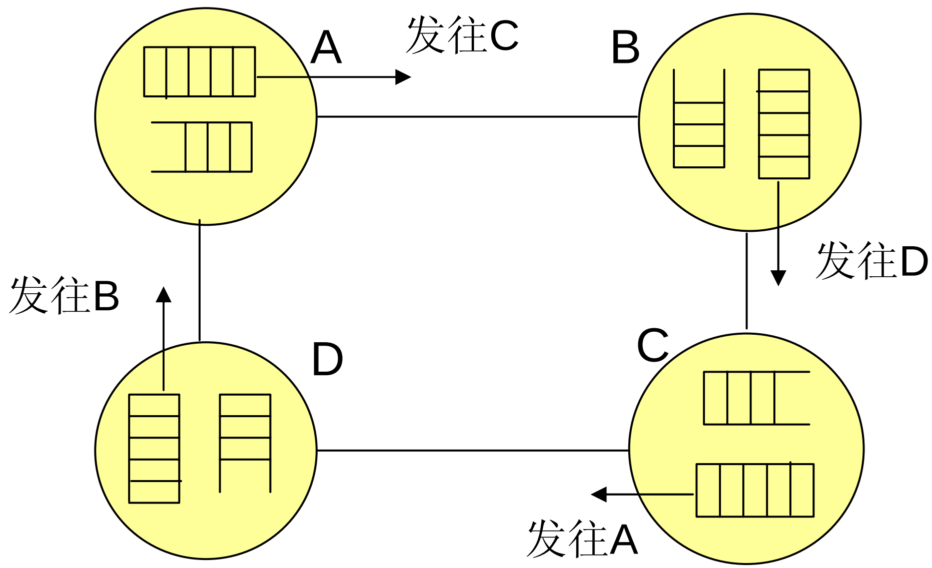
段级流量控制

- 最小分配共享
  - 设$_{bmin}$为保证分配给每一队列的缓冲区大小(当然，$b_{min} \leqslant$ B/N)。约束条件是：

$$\Sigma \, Max(0，n_i - b_{min}) \leqslant B\text{-}N^* \, b_{min}$$

- 最小分配最长队列共享
  - 每一队列保证有一个最小的缓冲区b_{min}供使用，同时其最大长度受b_{max}的限制。

段级流量控制 — 间接死锁问题



- 现在设每个结点的等待发送队列长度均已达到规定的最大值b_{max}。
- 每个结点现在都具有用以存放到达分组的存放空间，但这些到达的分组却无法进入等待发送的队列，因而最终仍要被丢弃。
- 所以在图中的A，B，C，D四个结点中位于发送队列中的分组，没有一个能进入下一个结点中的发送队列。这样的死锁即为间接死锁。

段级流量控制 — 间接死锁问题

- 设子网共有N个结点，其中最长的路由要经过M段链路。
- 给每个结点开辟M+1个缓冲区(每个缓冲区存放一个分组)，并编上号码，从0到M。
- 缓冲区的有向图是这样做出的
  - 从一个结点的缓冲区i连一条弧到与其相邻的每个结点中的缓冲区i+1。
  - 以某个结点A为例，合法的路由就是从结点A的缓冲区i转发到其相邻结点的缓冲区i+1，然后再转发到与结点A相距2段链路的结点，其缓冲区编号为i+2，以后依次类推。

段级流量控制 — 间接死锁问题

- 若主机要想把一个分组发往子网中的某一条通路的源结点，则仅当此源结点中的缓冲区0空闲时方有可能。
- 接着，这个分组只能发往某个相邻结点中的缓冲区1。
- 以后每转发一次，所到达的缓冲区的编号就增大一个号。这样，
- 分组的最后结局一定是以下两种情况之一：
  - 传到所要到达的目的结点，然后交付给主机；
  - 传到编号为肘的缓冲区，然后被丢弃，因而不再消耗网络的资源。

段级流量控制 — 间接死锁问题

- 这种方法能够消除间接死锁。
- 具有最高编号肘的缓冲区不可能总是被占满的(分组到达这里就被丢弃)。
- 当一个分组处于缓冲区i时，它一定能转发到缓冲区i+1，这里i<M。也就是说，缓冲区i+1不可能总是空不出来的。

段级流量控制

- 虚电路段级流量控制的工作原理是
  - 给每条虚电路上的每一个结点规定一个缓冲区个数的上限值 M(固定或可变)。
  - 当掛个缓冲区用完时，这条虚电路就不允许通过更大的流量。
  - M可在虚电路建立时即规定好，也可以根据通信量的情况动态地加以调整。
  - 虚电路段级流量控制的好处是可以更有效地和更迅速地从拥塞状态中恢复过来，因为它可以有选择地抑制分组进入处于拥塞状态的虚电路。
  - 在使用虚电路段级流量控制时，可以将M个缓冲区固定地分配给每一条虚电路，也可以采用完全共享一个公用缓冲池的方法。
  - 由于固定分配缓冲区的办法会使缓冲区的利用率非常低，因此在大多数情况下是使用共享公用缓冲池的方法。
- 美国的TYMNET和法国的TRANSPAC都是采用虚电路段级流量控制的。

入口到出口级流量控制

- 入口到出口级流量控制的主要目的是防止在出口结点(即目的结点)缓冲区的拥塞。
- 这种拥塞的产生是由于目的结点到主机的连接线路过载或是由于主机接收数据的速率太低。
- 当需要在目的结点把到达的分组重新装配成报文时，就有可能出现重装死锁(reassembly deadlock)。

进网级流量控制

- 进网级流量控制是在测量网络内部拥塞的基础上控制输入到网络的通信量。
  - 局部测量：测量入口结点缓冲区的占有程度
  - 全局测量：根据全网所有缓冲区的占用程度
  - 选择测量：根据到某个目的结点的整条通路的拥塞情况
- 将测量结果报告进网结点，以控制从外部进入网内的通信量。

进网级流量控制

- 进网级流量控制是设法抑制从外部进入网内的通信量，从而防止整个网内的缓冲区产生拥塞。

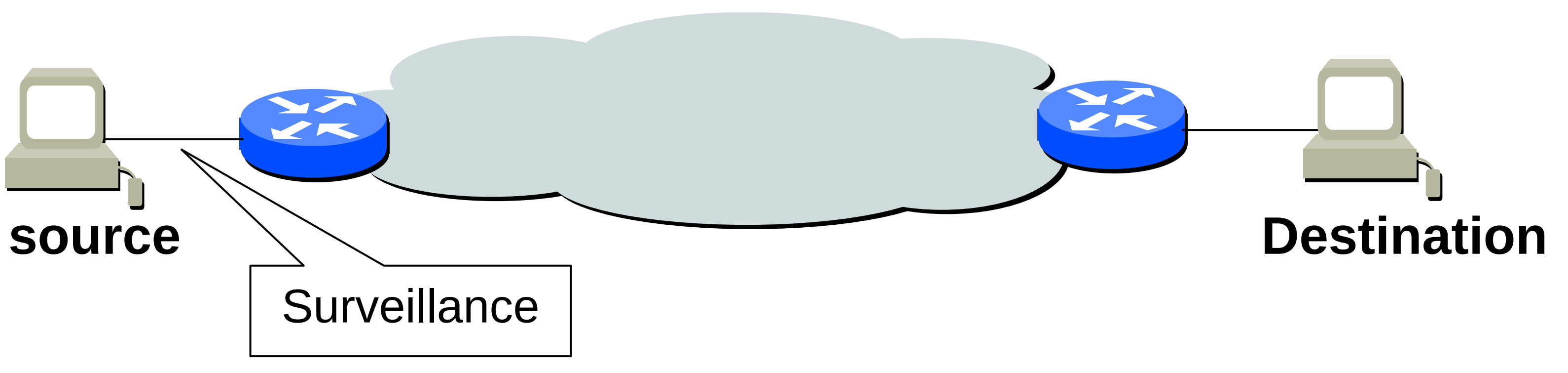
进网级流量控制

- 英国国家物理实验室提出的许可证方案(Isarithmic scheme)
  - 在一开始，网内各结点都持有若干张许可证，许可证的作用是允许分组从入口结点发往目的结点。
  - 主机向网内发送分组时，必须使每一个分组都能得到一张许可证。
  - 每进入网络一个分组，就使网络所存储的许可证数目减1。
  - 当全部许可证发完后，就不再允许新的分组进入网内。
  - 只有当一个分组到达目的站送交主机后，才将原来得到的许可证释放。这个被释放的空闲许可证可供新入网的分组来使用。
- 许可证的分发管理(包括对许可证丢失的处理)是一个相当复杂的问题。

3. Traffic Management

Traffic Management

- Fairness
- Quality of service
  - May want different treatment for different connections
- Reservations
  - e.g. ATM
  - **Traffic contract** between user and network



Packet Switched Networks

- Send **control packet** to some or all source nodes
  - Requires additional traffic during congestion
- Rely on **routing mechanism** and traffic info
  - 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

Congestion Control in Different Networks

- Packet Switched Networks
  - Frame Relay
- ATM (Asynchronous Transfer Mode)

4. Congestion Control in Frame Relay Network

Objectives of Congestion Control in FR

- Minimize discards
- Maintain agreed QOS
- Minimize variance in QOS
- Distribute resources fairly
- Minimize probability of one end user monopoly
- Limit spread of congestion
- Simple to implement
  - Little overhead on network or user
- Create minimal additional traffic
- Operate effectively regardless of traffic flow
- Minimum impact on end systems

FR 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

(1) Traffic Rate Management

- Must discard frames to cope with congestion
- How to discard
  - 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, i.e. not guaranteed
    - Aggregate CIR should not exceed physical data rate

$$\sum_i CIR_{i,j} \le AccessRate_j$$

(1) Traffic Rate Management

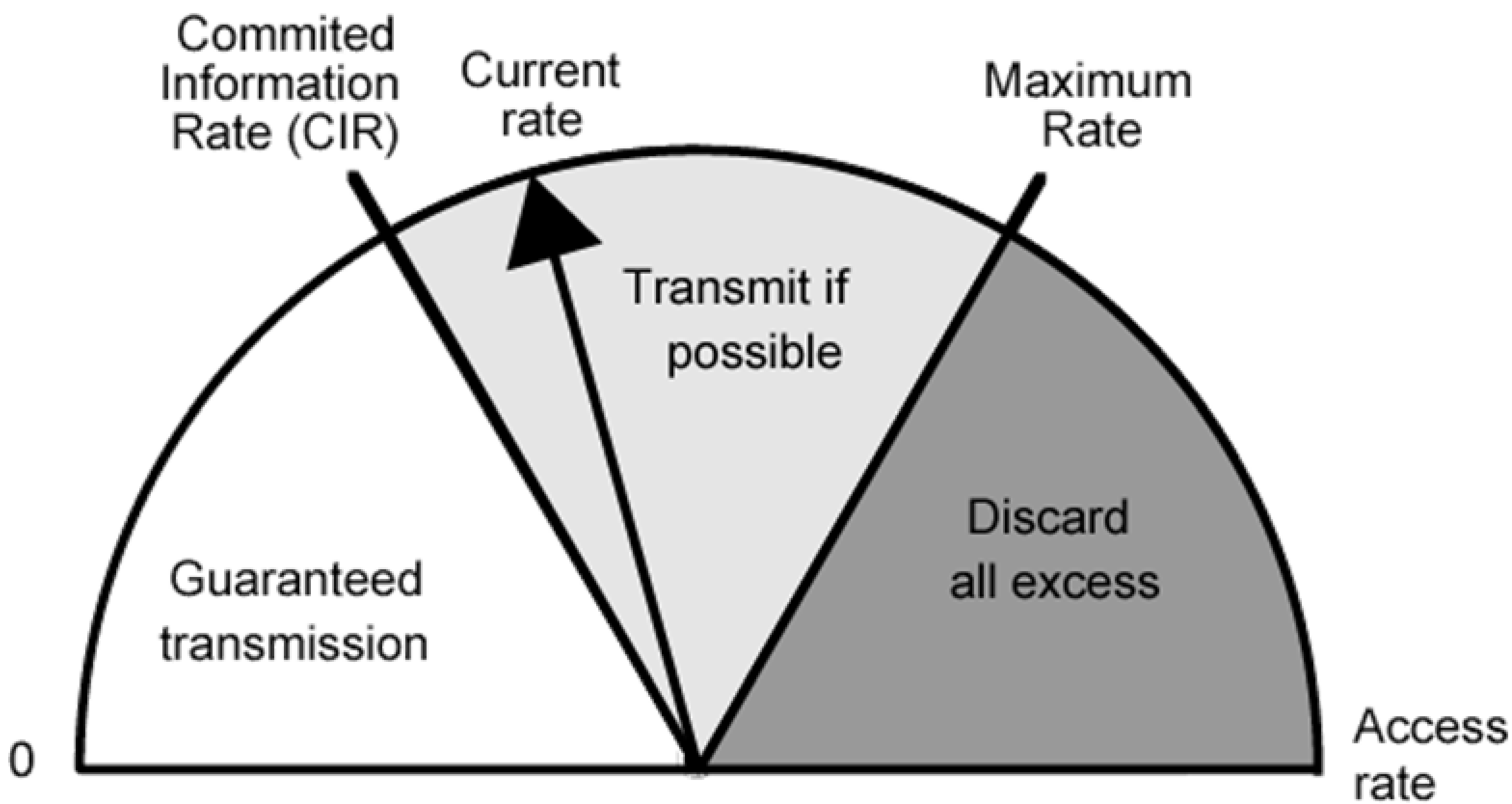
- Terms in CIR
  - Committed burst size (B_c in duration T)
  - Excess burst size (B_e in duration T)
- B_c is **committed to transmit**

$$CIR = B_C / T \quad \text{i.e.} \quad B_C = CIR \times T$$

- Data between B_c+B_e are **permitted but not guaranteed**
- Data above B_c+B_e are **discarded**

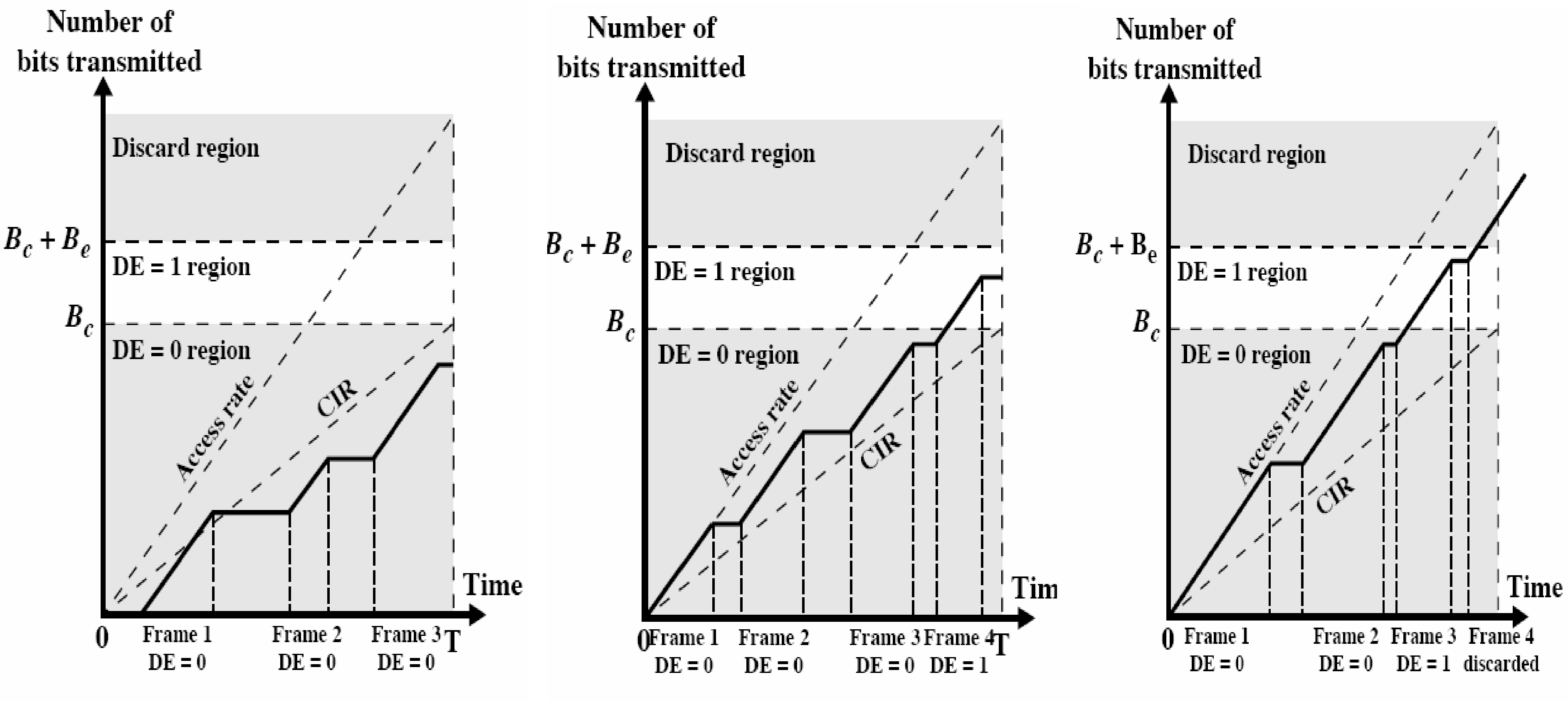
(1) Traffic Rate Management

Operation of CIR



(1) Traffic Rate Management

Relationship Among Congestion Parameters

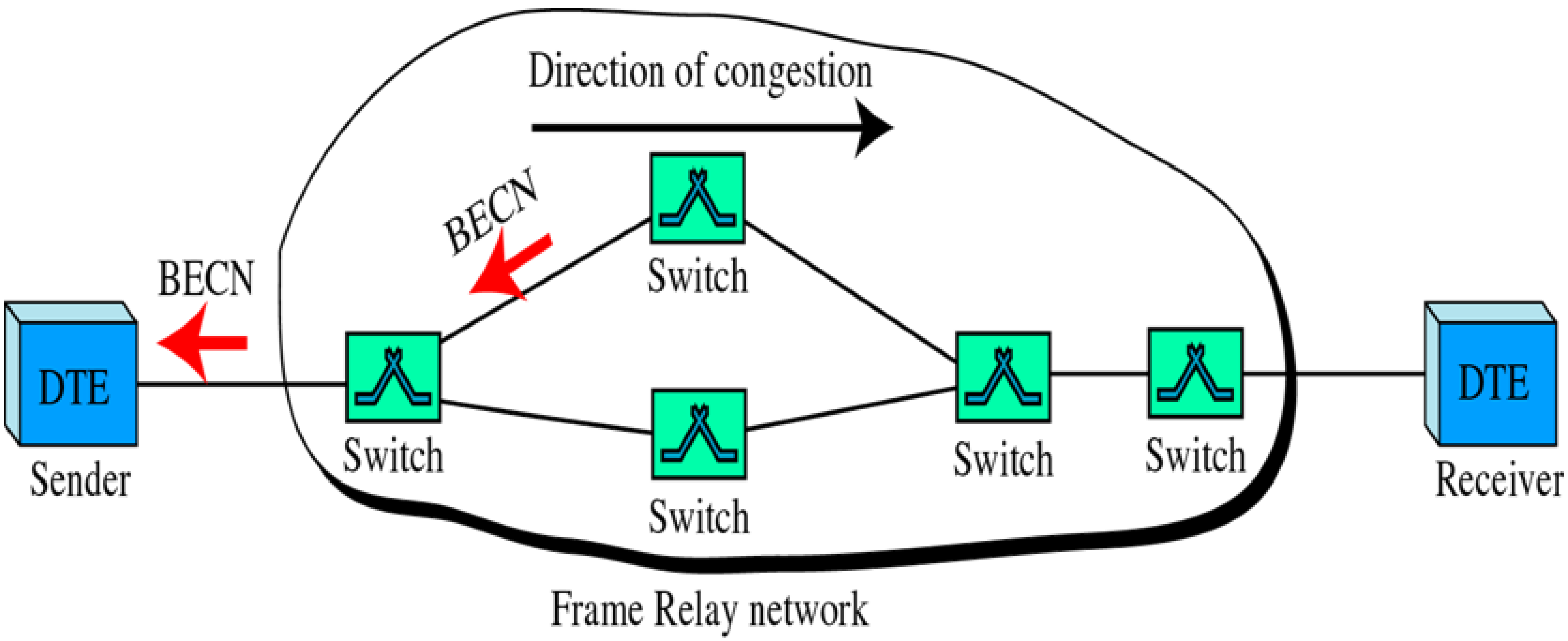


(a) All frames within CIR (b) One frame marked DE (c) One frame marked DE; one frame discarded

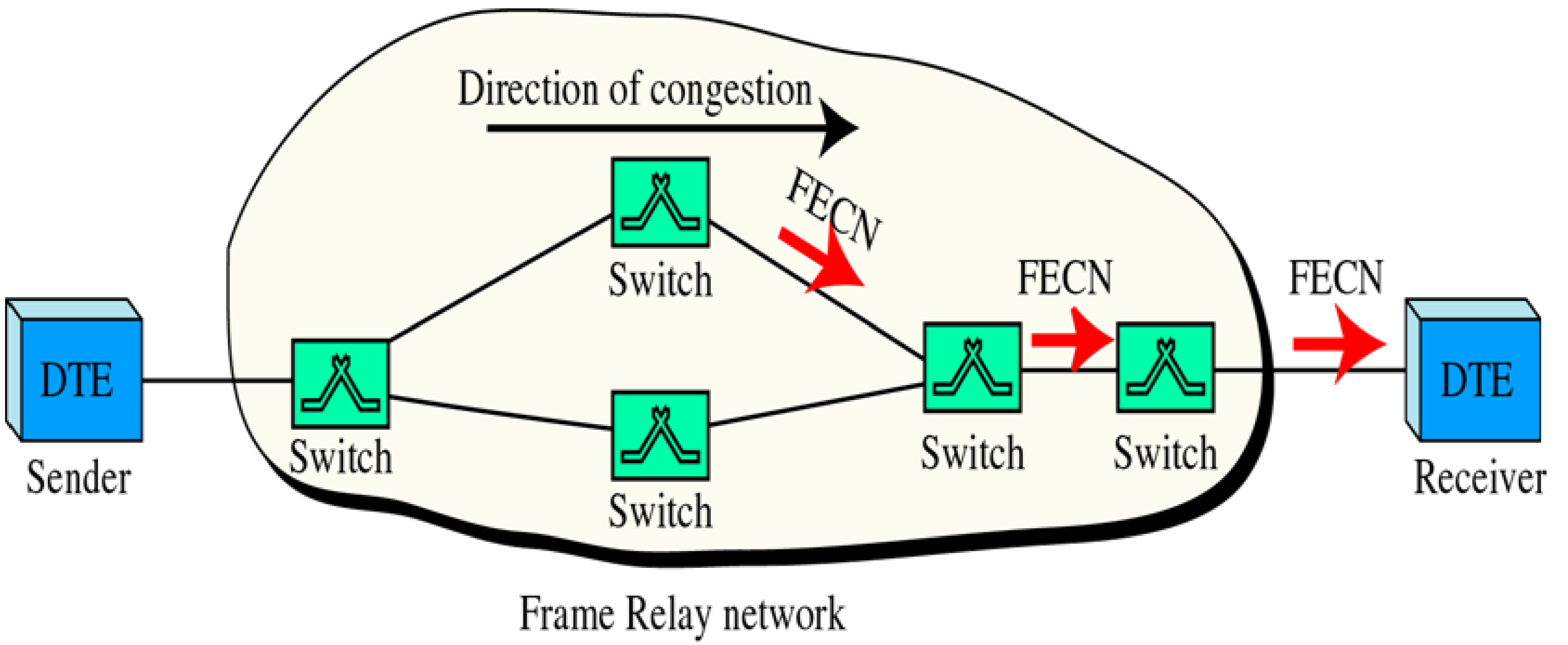
(2) Explicit Signaling

- Network **alerts end systems** of growing congestion
  - Frame handler monitors its queues
  - May notify **some** or **all** logical connections
- Method
  - Backward explicit congestion notification (BECN)
  - Forward explicit congestion notification (FECN)
- Network response
  - Set BECN or FECN
- User response
  - Reduce data rate

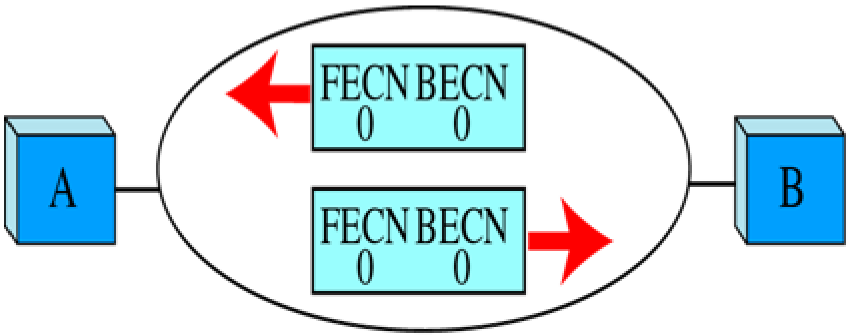
BECN


Direction of congestion
BECN
BECN
DTE
Sender
Switch
Switch
Switch
Switch
Switch
DTE
Receiver
Frame Relay network

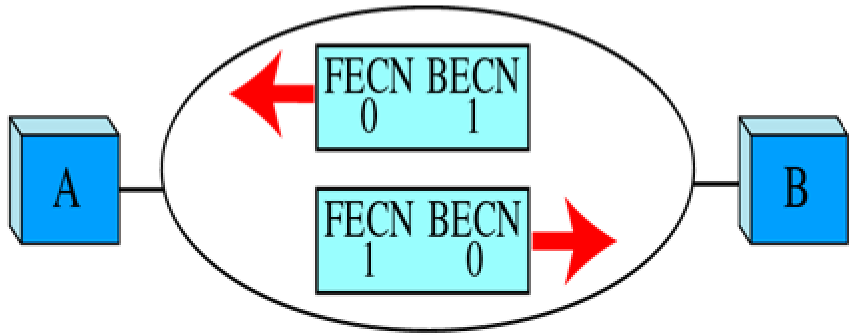
FECN


Direction of congestion
FECN
FECN
FECN
DTE
Sender
Switch
Switch
Switch
Switch
Switch
DTE
Receiver
Frame Relay network

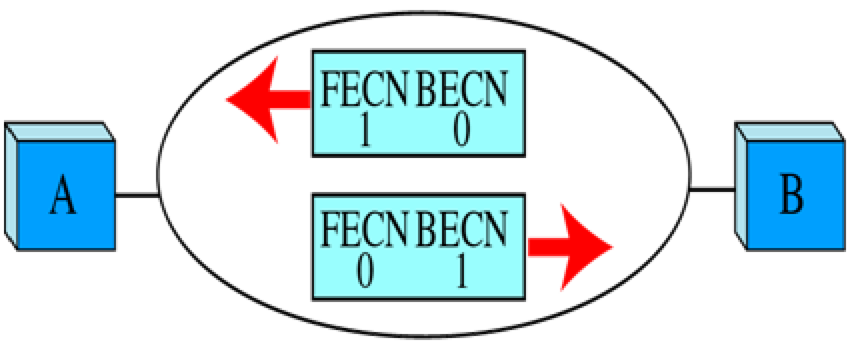
4 Cases of Congestion



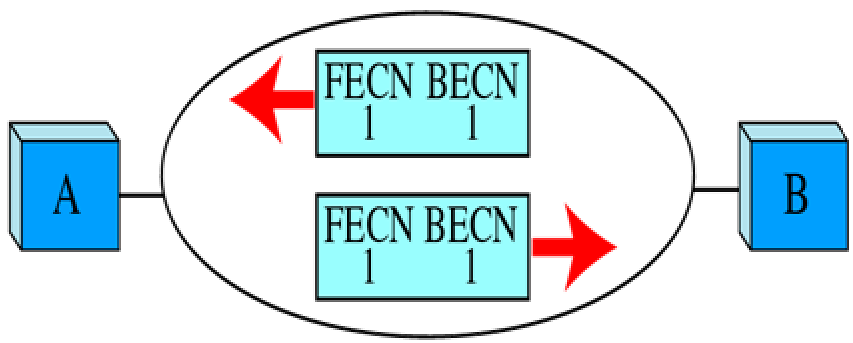
a. No congestion



b. Congestion in the direction A–B



c. Congestion in the direction B–A



d. Congestion in both directions

5. Congestion Control in ATM Network

(1) ATM Traffic Management

- ATM traffic
  - High speed, small cell size, limited overhead (control) bits
- Peculiarities
  - Majority of traffic not amenable to flow control (not draw back)
  - **Feedback slow** due to reduced transmission time compared with propagation delay
  - **Wide range** of application demands, from several kbps to hundreds of Mbps
  - **Different traffic patterns**, from real-time traffic to bursty traffic
  - **Different network QOS**, from lost sensitive to delay sensitive
  - High speed switching and transmission increases volatility

(2) Latency/Speed Effects

- ATM transmission rate is 150Mbps
- Time to **insert a cell**
  - （53•8）/（150•10^6）≈**2.8 • 10^{-6}s**

- Time to **traverse network**
  - Propagation delay + switching delay
  - Assume 0 switching delay
  - Assume propagation at 2/3 speed of light
  - If source and destination on opposite sides of USA, propagation time ≈ **48•10^{-3}** *seconds*

（2） Latency/Speed Effects

- If apply **implicit congestion control** (TCP) and ICMP **source quench**
- By the time dropped cell notification has reached source, **7.2M bits** have been transmitted

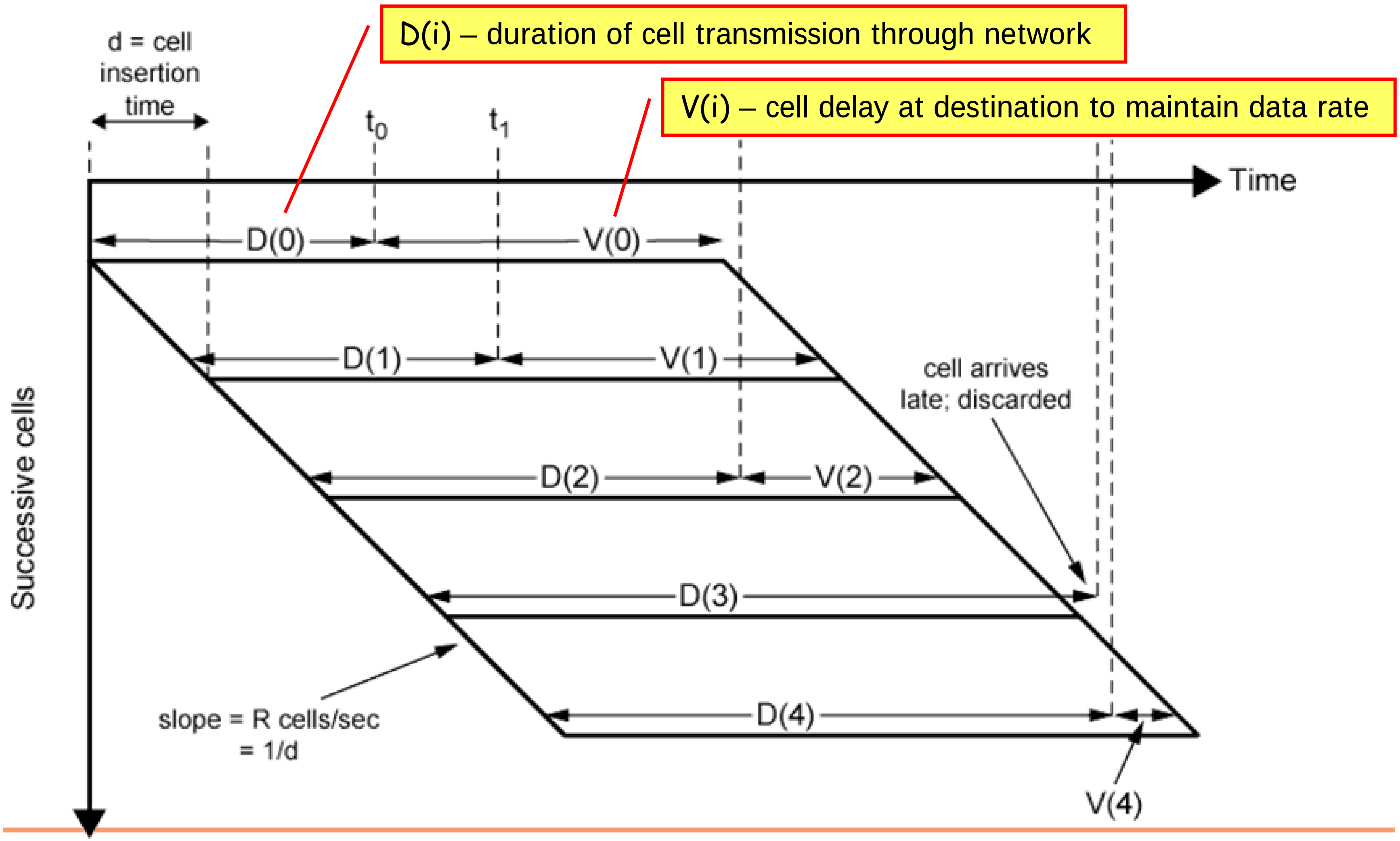
$$N=\frac{48\times10^{-3}}{2.8\times10^{-6}}=1.7\times10^{4}\,cell=1.7\times10^{4}\times53\times8=7.2\times10^{6}\,bits$$

- So, this is not a good strategy for ATM

（3） Cell Delay Variation

- For ATM voice/video, data is **a real-time 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 app so that constant bit rate can be maintained

（2） Cell Delay Variation



（3） Cell Delay Variation

- To achieve:
  - **A constant rate** of reception
  - Same as the cell insertion time pattern
- The next arrived cell is delayed V(i) to satisfy:
  - $t_1+V(1) = t_0+V(0)+\delta$

    t_1 is the arrive time of cell1,

    $t_0+\delta$ is the cell finish time,

    $t_0+V(0)+\delta$ is the deadline for next cell

    so it must satisfies $t_1 \leqslant t_0+V(0)+\delta$ that is

    there exist $V(1) \geqslant 0$ such that $t_1+V(1) = t_0+V(0)+\delta$
  - $V(1) = V(0) - [t_1 - (t_0+\delta)] \geqslant 0$
  - In General

    $V(i) = V(i-1) - [t_i - (t_{i-1}+\delta)]$, and

    $V(i) = V(0) - [t_i - (t_0+i\times\delta)]$
  - Discard the cell if $V(i)<0$

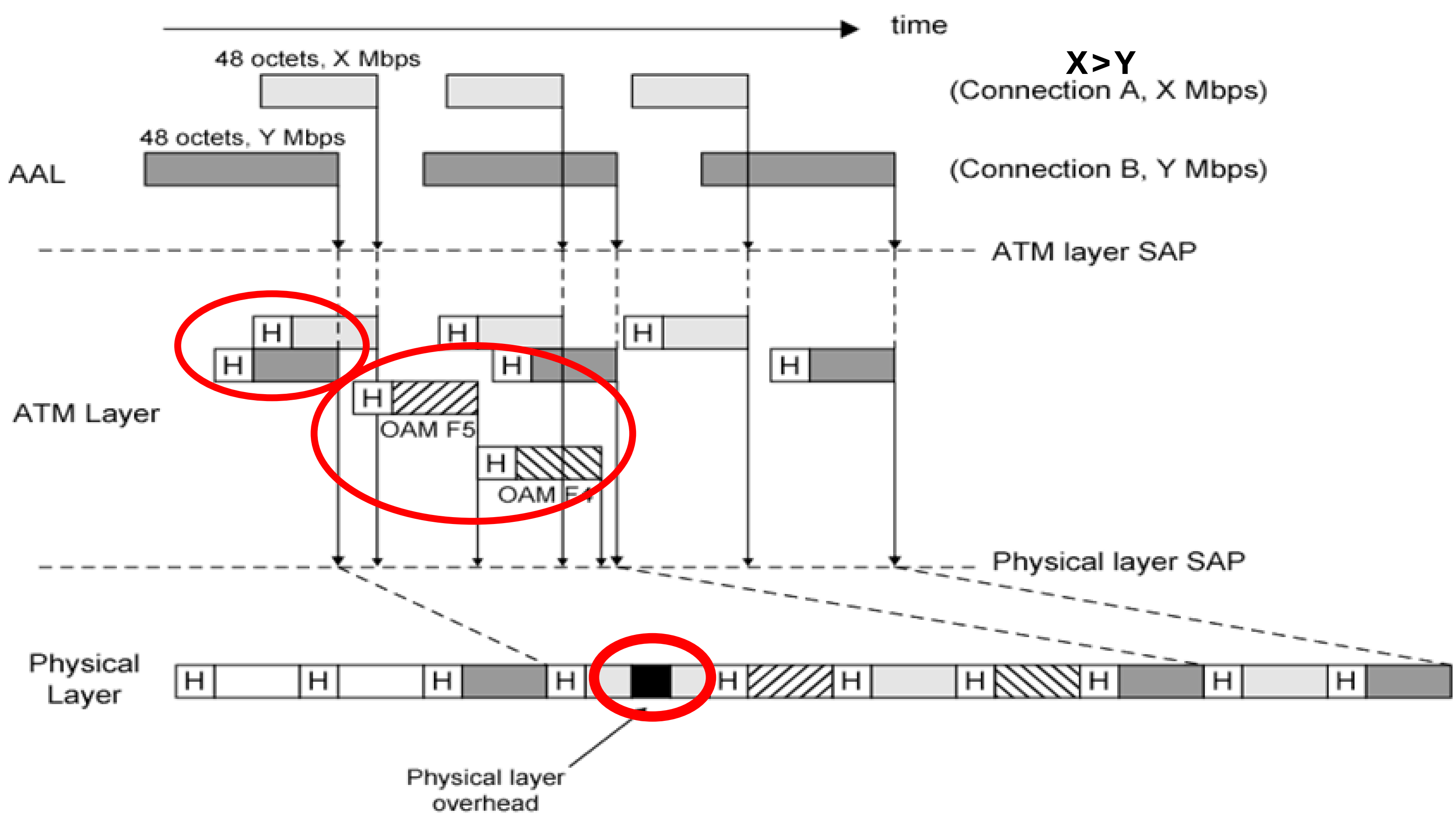
（3） Latency/Speed Effects

UNI Contribution to the Cell Delay

- Application produces data at fixed rate
- Processing at 3 layers of ATM causes delay
  - Interleaving cells **from different connections**
  - **Operation and maintenance** cell interleaving
  - If using synchronous digital hierarchy frames, **extra bits are inserted** at physical layer
- **Can not predict** these delays

（3） Latency/Speed Effects

UNI Contribution to the Cell Delay



（3） Latency/Speed Effects

Network Contribution to the Cell Delay

- 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

（4）Traffic and Congestion Control

- Requirements
  - **Support QOS classes** for all foreseeable network services
  - **Not rely on AAL protocols** that are network specific, **nor higher level** application specific protocols
  - Minimize network and end to end system complexity
- Tactics
  - Determine whether a **given new connection can be accommodated**
  - Agree **performance parameters** with subscriber

（4）Traffic and Congestion Control

- Cell insertion time
  - Usage parameter control
  - Selective cell discard
  - Traffic shaping
  - **Selective cell discard**
- Round trip propagation
  - Fast resource management
  - **Explicit forward congestion indication (EFCI)**
  - **ABR flow control**
- Connection duration
  - Connection admission control (CAC)
- Long term
  - Resource management using virtual paths

The Framework

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

（5）Control Techniques

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

（5）Control Techniques

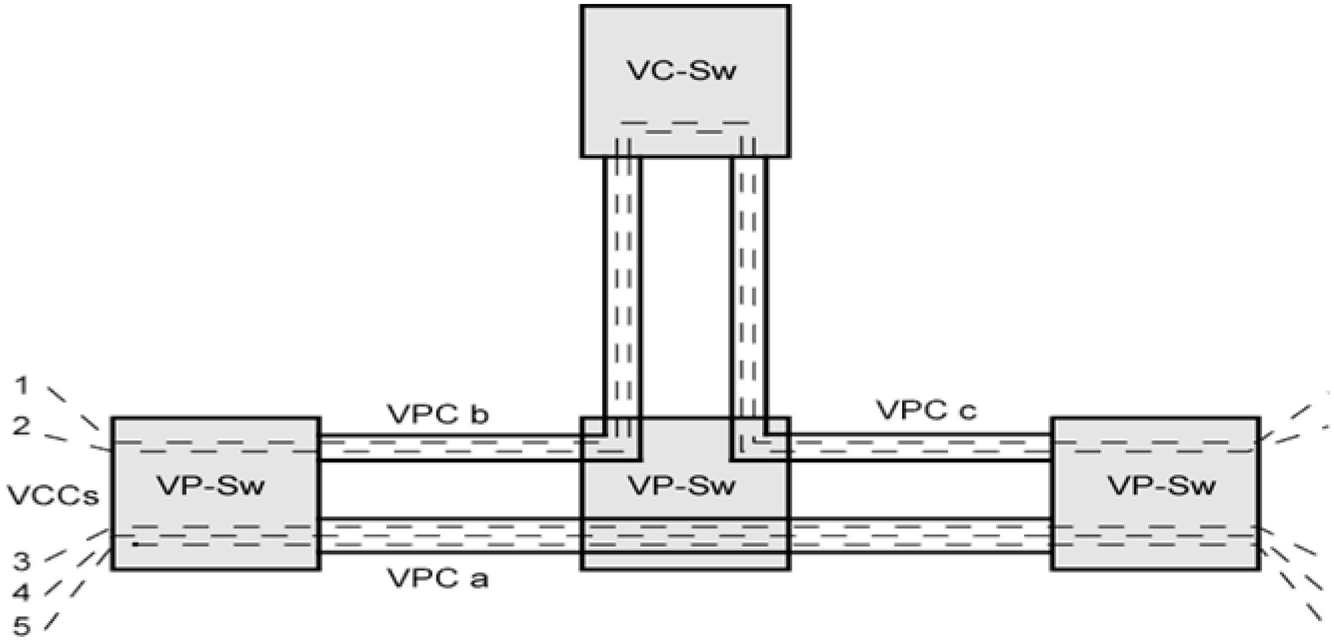
Resource Management Using Virtual Paths

- **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

（5）Control Techniques

Resource Management Using Virtual Paths

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

（5）Control Techniques

Resource Management Using Virtual Paths

Allocating VCCs within VPC

- All VCCs within a VPC should **experience similar network performance**
- Options for allocation
  - Aggregate peak demand
    - Good but **wasteful**
  - Statistical multiplexing
    - May suffer **cell loss** and high cell delay variation

（5）Control Techniques

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 (Tolerance)
  - Sustainable cell rate (VBR only)
  - Burst tolerance (VBR only)

（5）Control Techniques

Traffic Parameters Used in Defining VCC/VPC Quality of Service

Parameter	Description	Traffic Type
Peak Cell Rate (PCR)	An upper bound on the traffic that can be submitted on an ATM connection.	CBR, VBR
Cell Delay Variation (CDV)	An upper bound on the variability in the pattern of cell arrivals observed at a single measurement point with reference to the peak cell rate.	CBR, VBR
Sustainable Cell Rate (SCR)	An upper bound on the average rate of an ATM connection, calculated over the duration of the connection.	VBR
Burst Tolerance	An upper bound on the variability in the pattern of cell arrivals observed at a single measurement point with reference to the sustainable cell rate.	VBR

（5）Control Techniques

Usage Parameter Control

- Objective
  - Monitor connection to **ensure traffic conforms to contract**
  - Protection of network resources from overload by one connection
- Method
  - 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**

（5）Control Techniques

Control with CLP

- **Peak cell rate** and **cell delay variation**
  - Set peak cell rate R
  - Set CDV tolerance limit of τ
- The **inter-arrival time** between cells
  - $d = 1/R$
  - $(d - \tau)$ will be the tolerance limit
- Tag noncompliant cells within tolerance limit with *CLP=1*
  - $d-\tau \sim d$

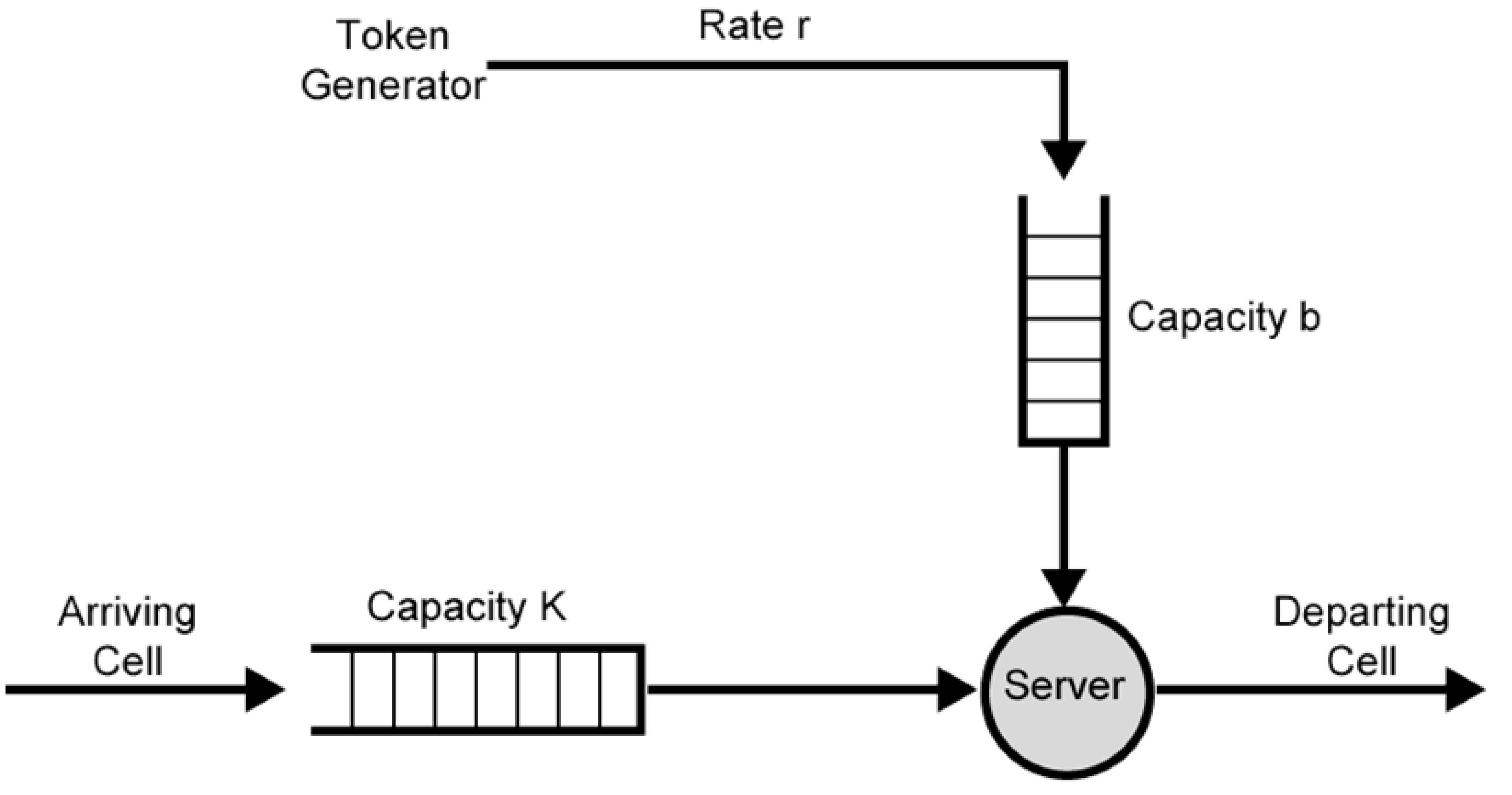
（5）Control Techniques

Traffic Shaping

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

（5）Control Techniques

Traffic Shaping — Token Bucket



6. GFR Traffic Management

GFR Traffic Management

- Control traffic from **end system**
  - **Guaranteed frame rate** is as simple as 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 transmit at minimum data rate without losses
  - If no congestion, higher rates maybe accepted

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, sent with best efforts
  - CLP=0 frames sent with **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)

Burst Size and Time

- PCR set to 20Mbps
- Maximum burst size (MBS) set to 100 cells
- Time for burst
  - 100 cells / PCR

    (100 cells $\times$ 424 bits/cell) 20 $\times$ 10^6 bits/second

    = 2.12 ms
- Burst time is **relatively short**

Mechanisms for Supporting Rate Guarantees

- Tagging and policing
- Buffer management
- Scheduling

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 with policing

How to Tagging

- **For VBR connections**
  - Cells that exceed *PCR* are discarded
  - Cells that exceed *SCR+MBS* are either discarded or tagged with *CLP=1*
  - Cells that exceed *SCR* (*MCR*) may be tagged with *CLP=1*
- For CBR connections
  - Cells that exceed *PCR* are discarded

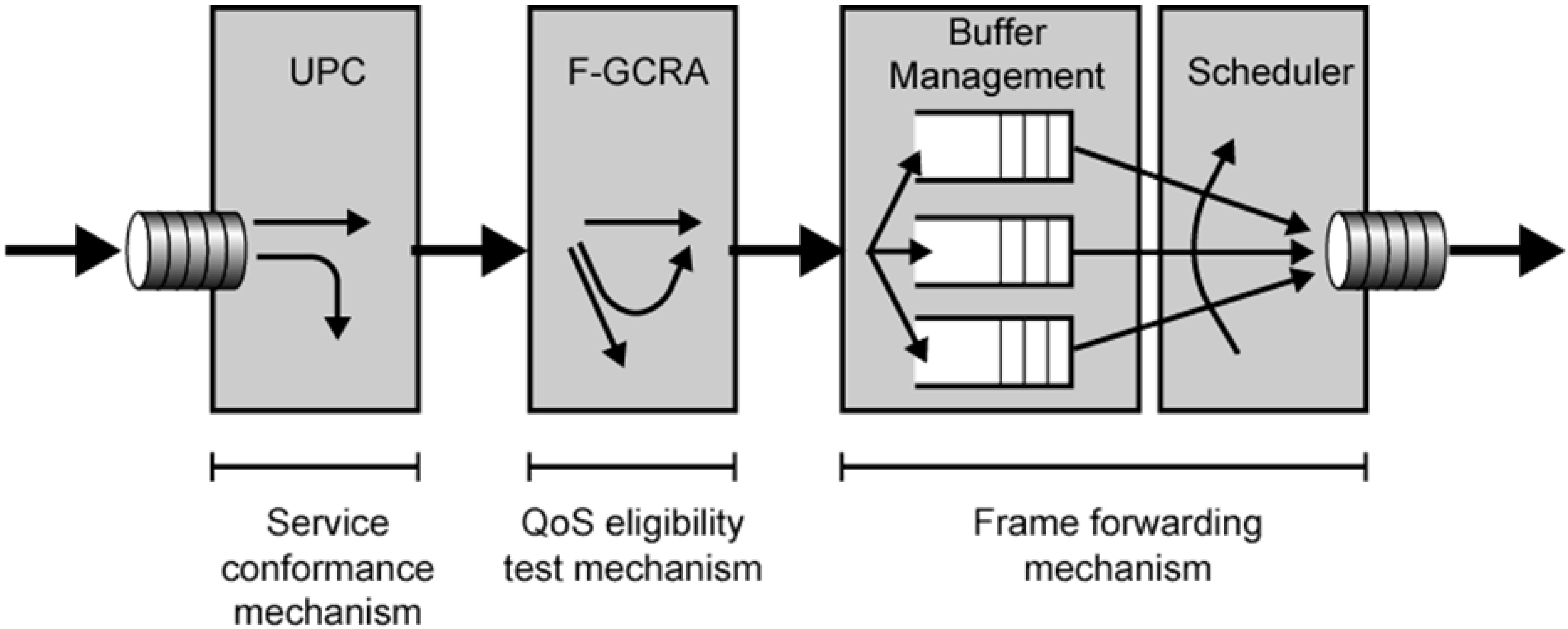
Buffer Management

- 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

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 and meet contract

Components of GFR System



Conformance Definition

- **UPC Function**
  - 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 a frame have same *CLP*
  - Frame satisfies *MFS* parameter (check for last cell in a frame or *Cell count < MFS*)

QOS Eligibility Test

Two stage filtering process

- Frame tested for **conformance to contract**
  - **Set upper bound** (PCR + MBS)
  - Penalize cells above upper bound: may discard or tag
  - Implementations expected to attempt delivery of tagged cells (best-effort)
- Determine frames **eligible for QOS guarantees**
  - **Set lower bound** (MCR), under GFR contract for VC
  - Frames making up traffic flow below threshold are eligible

GFR VC Frame Categories

- **Nonconforming frame**
  - Cells of this frame will be tagged or discarded at source
- **Conforming but ineligible frames**
  - Cells will receive a best-effort service
  - Tagged under congestion control at switch
- **Conforming and eligible frames**
  - Cells will receive a guarantee of delivery

Summary

- **Congestion**
  - Queues, Ideal Network Utilization vs. Practical Performance
- **Congestion Control**
  - Backpressure, Choke Packet, Implicit Congestion Signaling, Explicit Congestion Signaling
  - Traffic Management
- **Packet Switched Networks (FR)**
  - Traffic Rate Management, BECN & FECN
- **ATM**
  - Latency/Speed Effects, Cell Delay Variation
  - Traffic and Congestion Control: Resource management with VP, Usage parameter control, Traffic shaping
  - GFR Traffic Management: Frame Recognition, Contract Parameters, Frame Categories