南京邮电大学
Nanjing University of Posts and Telecommunications

宽带交换技术

第二章 英特网及QoS控制技术

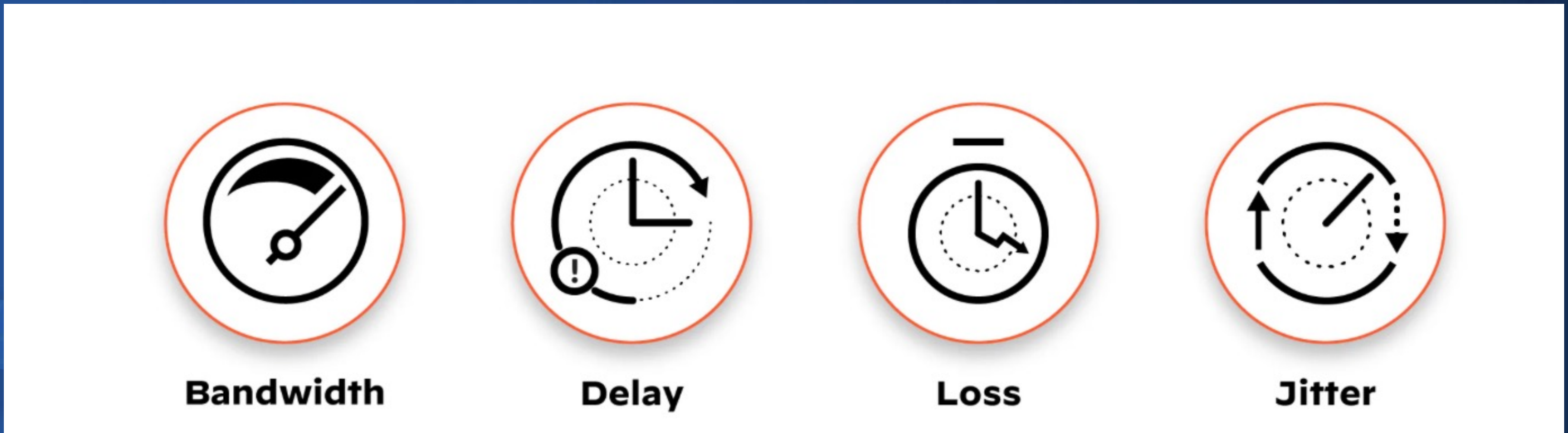
通信与信息工程学院
张喆

英特网及QoS控制技术

- QoS, QoE的基本概念(网络性能参数)
- **Traffic Engineering (TE)的基本概念**
- **流量控制与拥塞控制的基本方法**
- 因特网中的综合服务模型
- 移动性相关的QoS控制

什么是QoS？

- QoS (quality of service)
- QoS is determined by **bandwidth**, **delay**, **jitter**, and **packet loss**.



度量指标

QoS的定义

- ITU-T中的定义：服务质量是一种服务性能的体现，这种服务性能决定了网络多大程度上满足业务用户的需求。
- 狭义QoS：指的是技术，如传输时延，时延抖动，丢失率，带宽需求，吞吐量等。
- 广义的QOS: 资源调配与利用，层与层之间的协商，涉及最终服务的各个层次。

不同业务的QoS需求

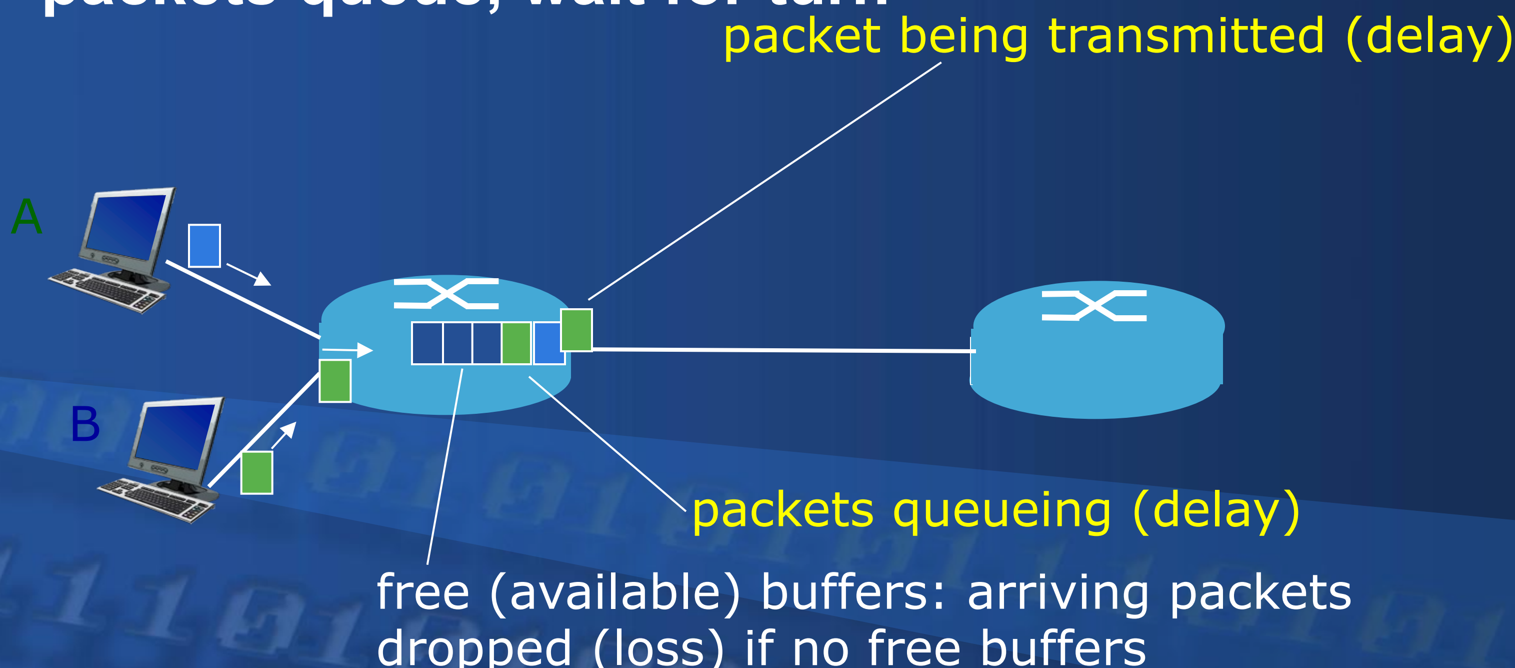
- 存在标准映射将现有服务以信元形式表示：语音，数据，视频
- 业务的质量要求

要求 / 类型	最大时延 S	最大时延波动 m S	速率 Mbps	平均吞吐量	可接受的误比特率	可接受的误分组率
话音	0.25	10	0.064	0.064	$<10^{-2}$	$<10^{-2}$
活动图象	0.25	10	100	100	$<10^{-2}$	$<10^{-3}$
静止图象	1.0	-	2-10	2-10	$<10^{-4}$	$<10^{-9}$
压缩后活动图象	0.25	1.0	2-10	2-10	$<10^{-6}$	$<10^{-9}$
数据文件	1.0	-	2-100	2-100	0	0
实时数据	0.001	-	<10	<10	0	0

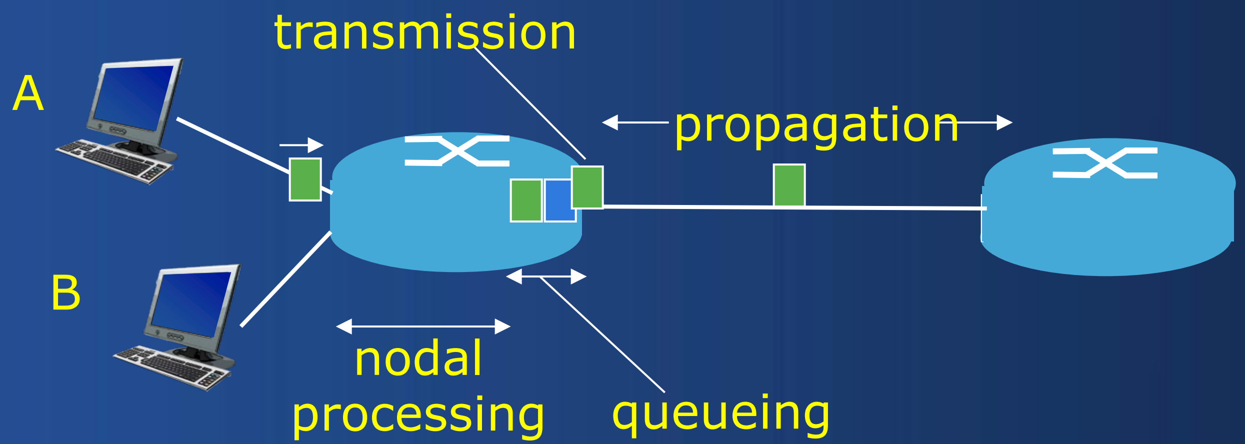
How do loss and delay occur?

packets *queue* in router buffers

- packet arrival rate to link (temporarily) exceeds output link capacity
- packets queue, wait for turn



Four sources of packet delay

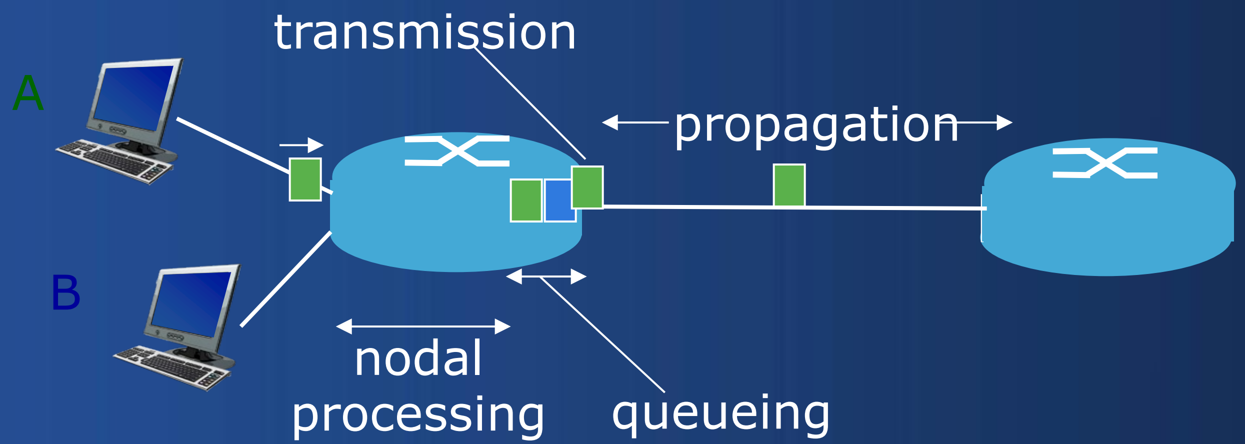


$$d_{nodal} = d_{proc} + d_{queue} + d_{trans} + d_{prop}$$

d_{proc}: nodal processing
- **check bit errors**
- **determine output link**
- **typically < msec**

d_{queue}: queueing delay
- time waiting at output link for transmission
- depends on congestion level of router

Four sources of packet delay



$$d_{nodal} = d_{proc} + d_{queue} + d_{trans} + d_{prop}$$

d_{trans}: transmission delay:

- L: packet length (bits)
- R: link *bandwidth (bps)*
- $d_{trans} = L/R$

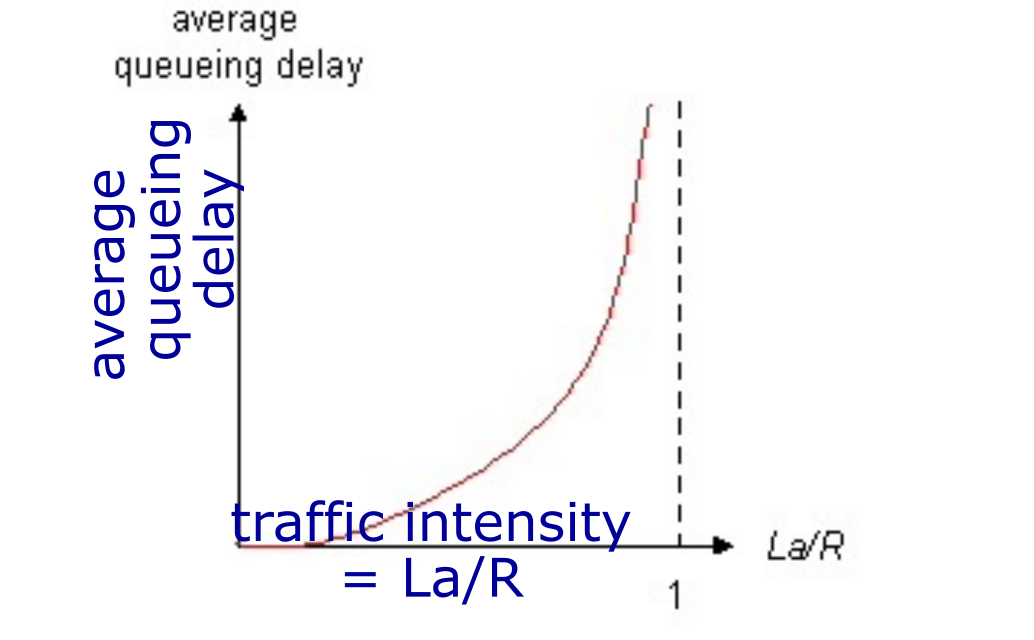
d_{trans} and d_{prop}
very different

d_{prop}: propagation delay:

- d: length of physical link
- s: propagation speed in medium (~2×10^8 m/sec)
- $d_{prop} = d/s$

Queueing delay (revisited)

- *R:* link bandwidth (bps)
- *L:* packet length (bits)
- a: average packet arrival rate



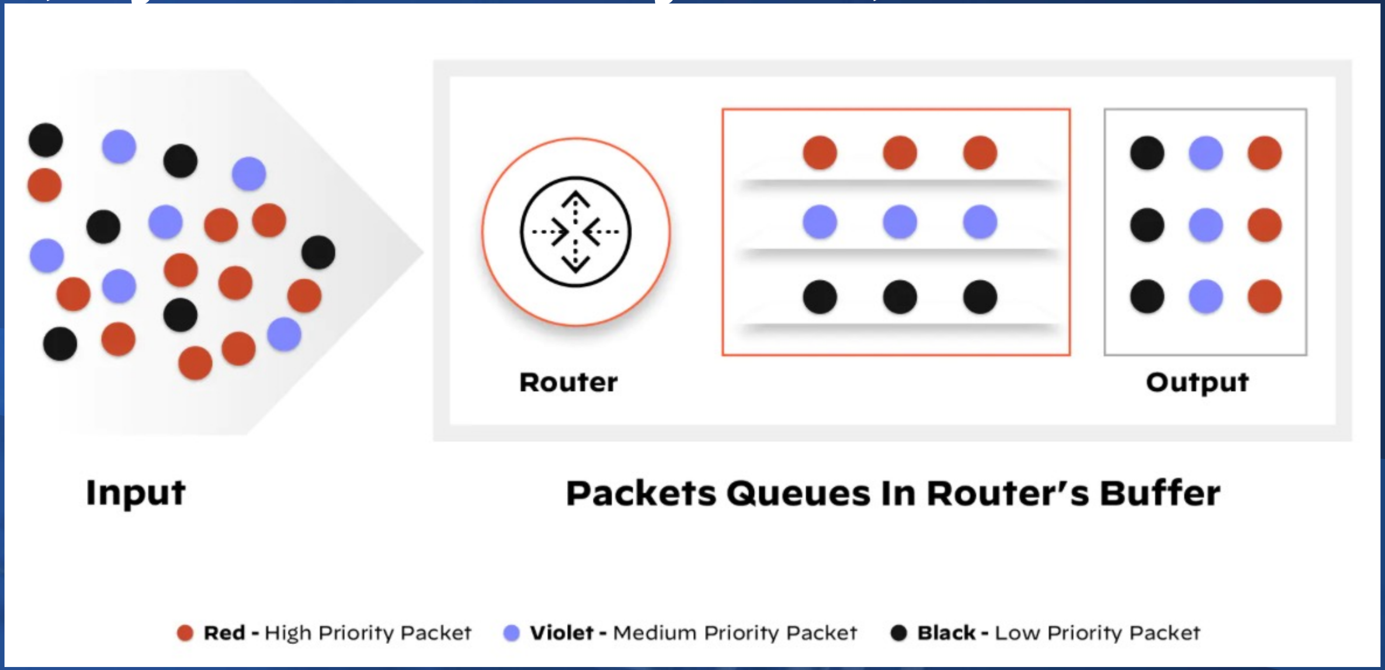
- *La/R* ~ 0: avg. queueing delay small
- *La/R* -> 1: avg. queueing delay large
- *La/R* > 1: more "work" arriving than can be serviced, average delay infinite!

La/R ~ 0

La/R -> 1

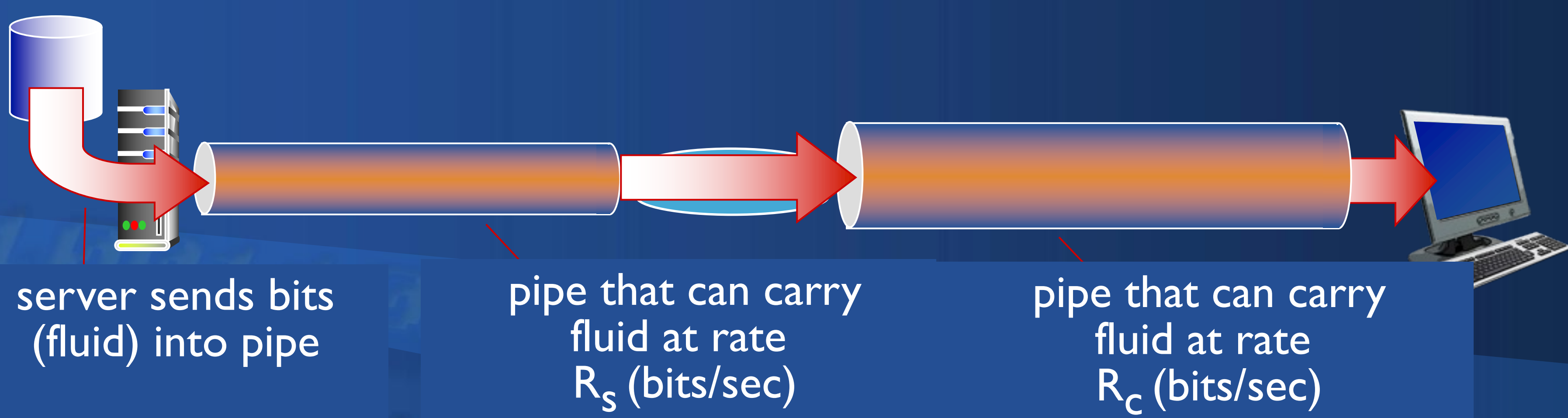
Packet loss

- queue (buffer) preceding link in buffer has finite capacity
- packet arriving to full queue dropped (lost)
- lost packet may be retransmitted by previous node, by source end system, or not at all



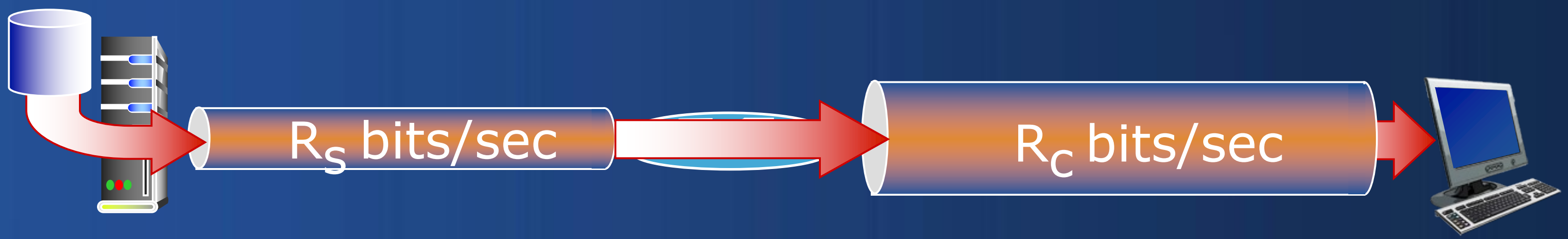
Throughput（吞吐量）

- ***throughput:*** **rate (bits/time unit) at which bits transferred between sender/receiver**
  - ***instantaneous:*** **rate at given point in time**
  - ***average:*** **rate over longer period of time**

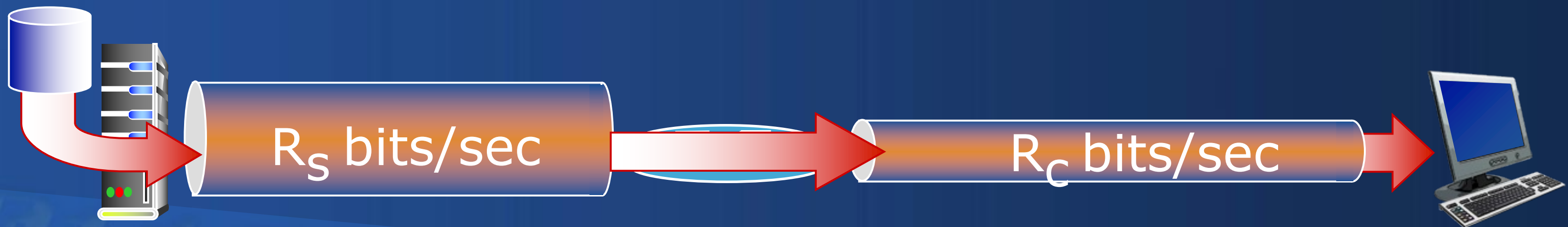


Throughput (more)

- $R_s < R_c$ What is average end-end throughput?



- $R_s > R_c$ What is average end-end throughput?



bottleneck link

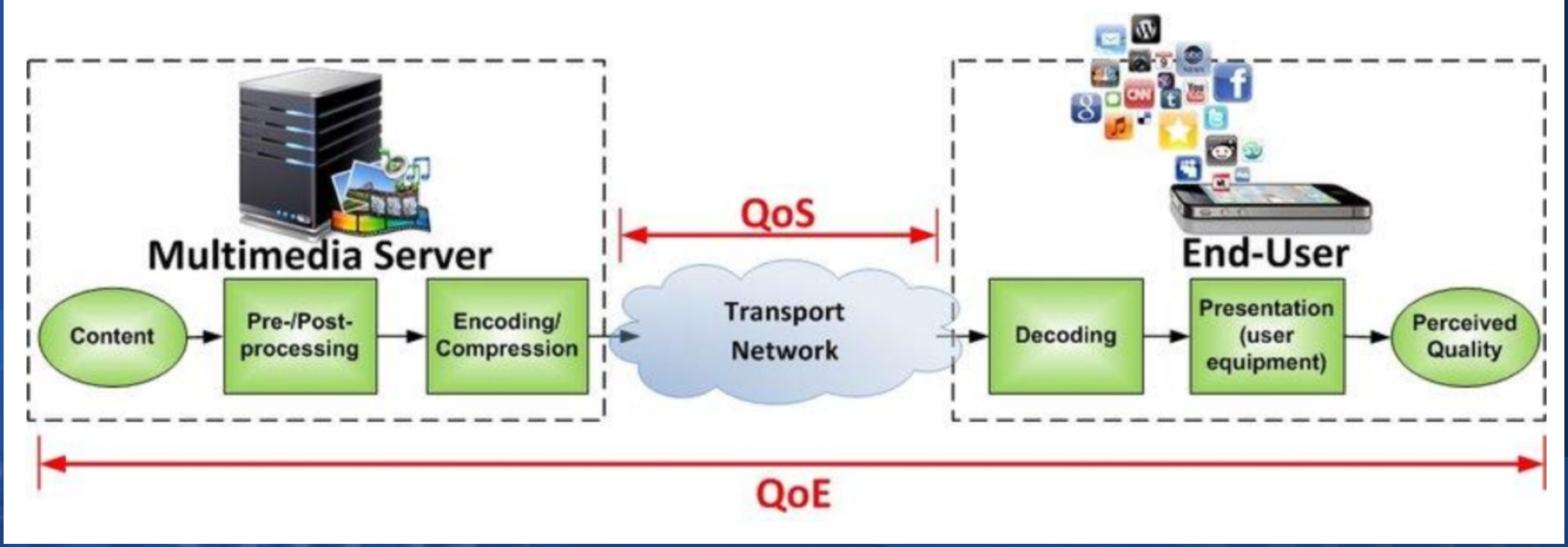
link on end-end path that constrains end-end throughput

QoS技术手段

- 流量整形（Traffic Shaping）: 限制发送速率。
- 包调度（Packet scheduling ）：为数据包分配资源，主要包括带宽、buffer space，CPU cycles。
- 接纳控制（Admission Control）:允许或拒绝流量，确保网络资源不被超额占用。
- Integrated services (IntServ): 需要资源预留协议RSVP来提前申请所需资源。
- Differentiated services (DiffServ): 通过对数据包进行分类和标记来确定其优先级。
-

什么是QoE？

- QoE (quality of experience)：是指用户对设备、网络和系统、应用或业务的质量和性能的主观感受。QoE指的是用户感受到的完成整个过程的难易程度。



英特网及QoS控制技术

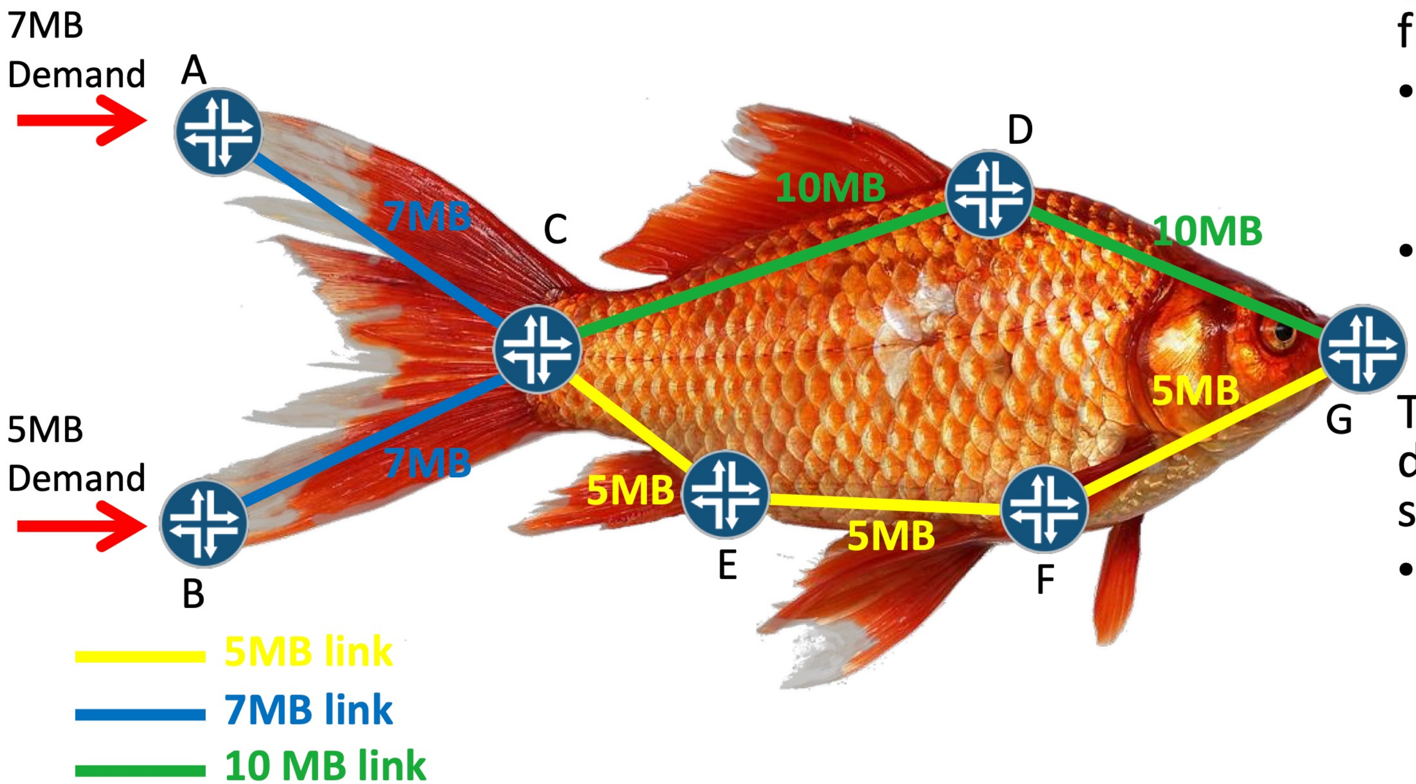
- QoS的基本概念（网络性能参数）
- Traffic Engineering (TE)的基本概念
- 流量控制与拥塞控制的基本方法
- 因特网中的综合服务模型
- 移动性相关的QoS控制

什么是Traffic Engineering（TE）

- **Traffic Engineering (TE): is the process of reconfiguring the network in response to changing traffic loads in order to achieve optimal performance.**

- **Key feature:**
  - No matter how the network resources are limited or if there are links/nodes down.
  - TE is expected to achieve the optimal performance.

It's All a Bit Fishy



Suppose a 5MB demand arrives first

- It gets placed on the shortest path
- Then the 7MB can't be placed

TE allows the 5MB demand to be steered to BCEFG

- Allows the network to support both demands

什么是Traffic Engineering (TE)

- Traffic Engineering (TE) focus on layer 3 and below:
  - Packets/frames: IP, MPLS, Ethernet
  - Transport technologies: TDM, OTN, lambda, port
- Toolkit:
  - QoS
  - Multiprotocol label switching (MPLS)
  - Traffic shaping
  - Load balancing
  - Network monitoring and analysis

QoS VS TE

- QoS: ensure that critical traffic is given priority over less important traffic

- TE: optimize the performance and utilization of network resources to ensure the traffic is efficiently and effectively delivered across the network

英特网及QoS控制技术

- QoS的基本概念（网络性能参数）
- Traffic Engineering (TE)的基本概念
- 流量控制与拥塞控制的基本方法
- 因特网中的综合服务模型
- 移动性相关的QoS控制

流量控制技术

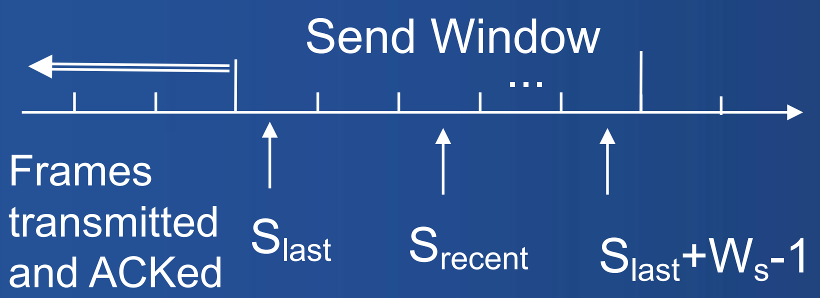
- Flow control：prevent a fast sender from overrunning a slow receiver
- 常见方法：
  - Sliding window protocol（滑动窗口机制，数据链路层和传输层均可采用此方法进行流量控制）
  - Stop and wait protocol (停止等待协议)
  - 二者区别：stop and wait protocol only send one data frame at a time, while sliding window protocol can send multiple data frames at a time.

Sliding Window Protocols

- A One-Bit Sliding Window Protocol
  - Also called stop and wait protocol
- A Protocol Using Go Back N
- A Protocol Using Selective Repeat

Sliding Window Operation

Transmitter

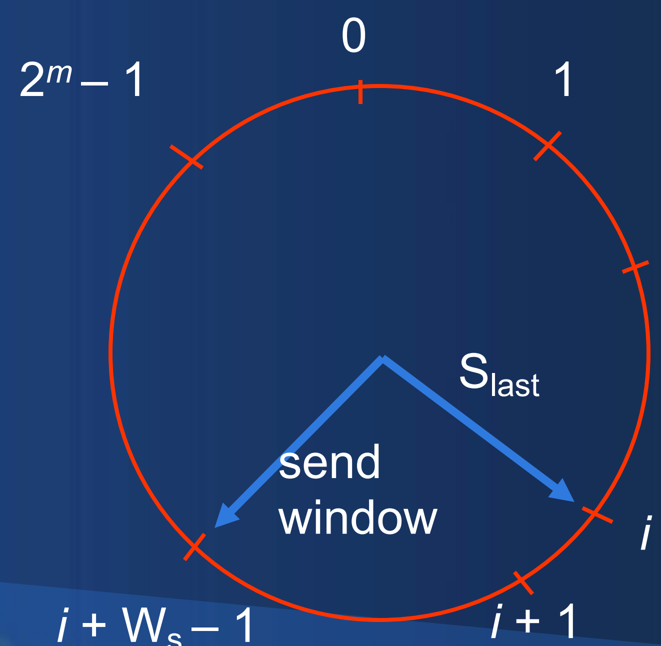


Transmitter waits for error-free ACK frame with sequence number S_{last}

When such ACK frame arrives, S_{last} is incremented by one, and the *send window slides forward* by one

W_s: send window size

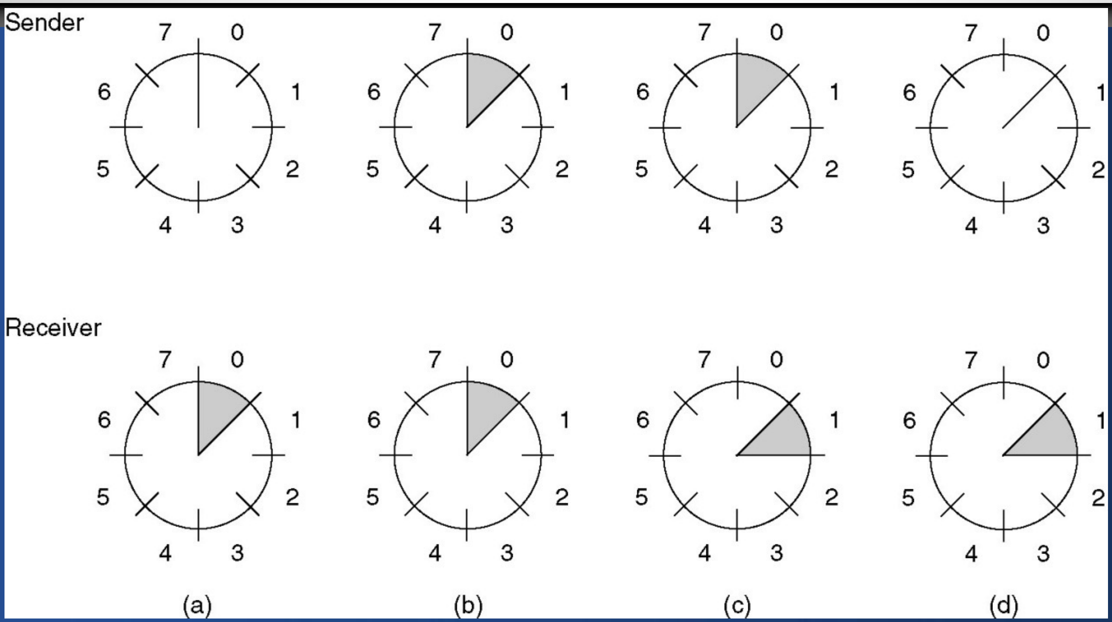
m-bit Sequence Numbering



Sliding Window Protocols (2)

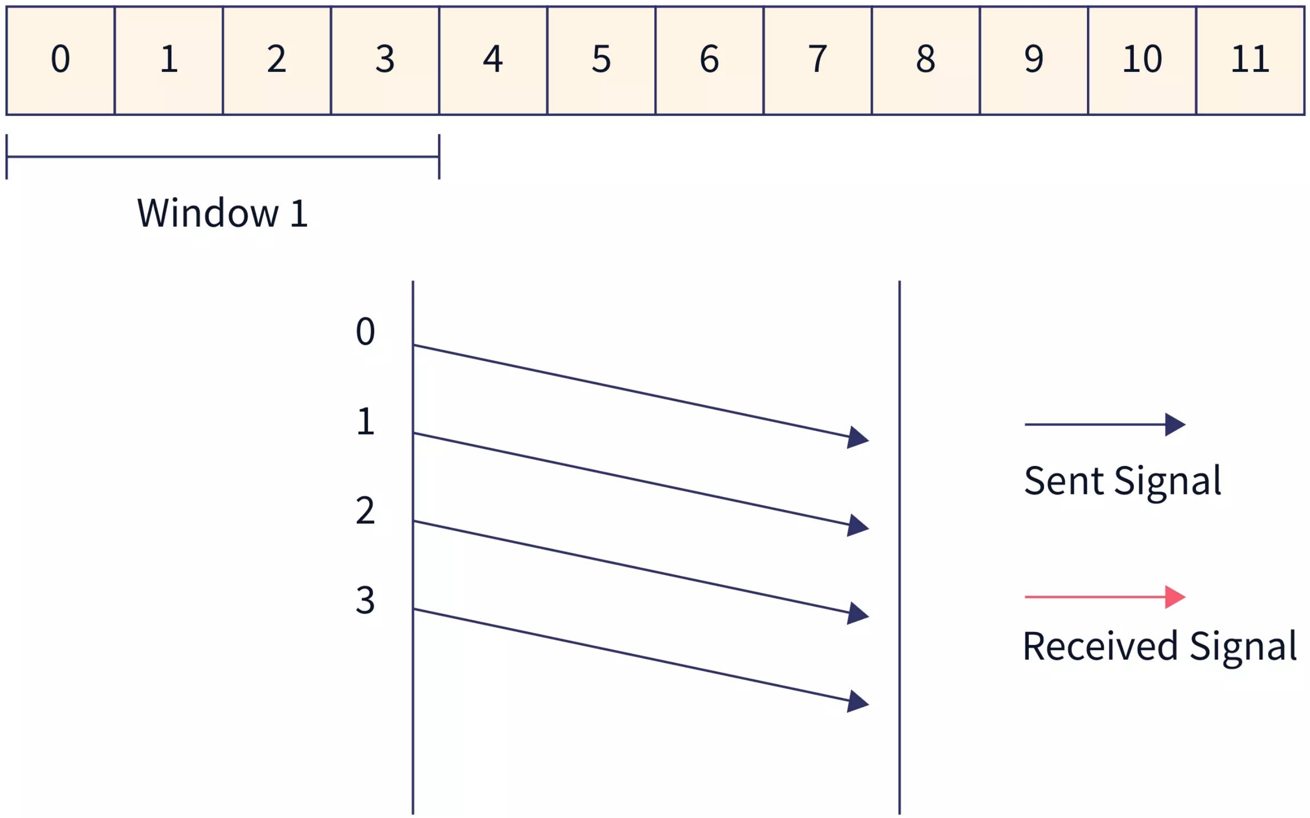
- Principle of sliding window protocols
  - Each transmitter maintains a table of sequence number(in order).The sequence numbers within the sender's window W_S represent frames that have been sent or can be sent but are as yet not acknowledged
  - Each receiver maintains a table of sequence number(in order)， receiving window W_R: corresponding to the set of frames it is permitted to accept.
  - Only when receiving window rotated, send window can slide.
  -

Sliding Window Protocols (3)

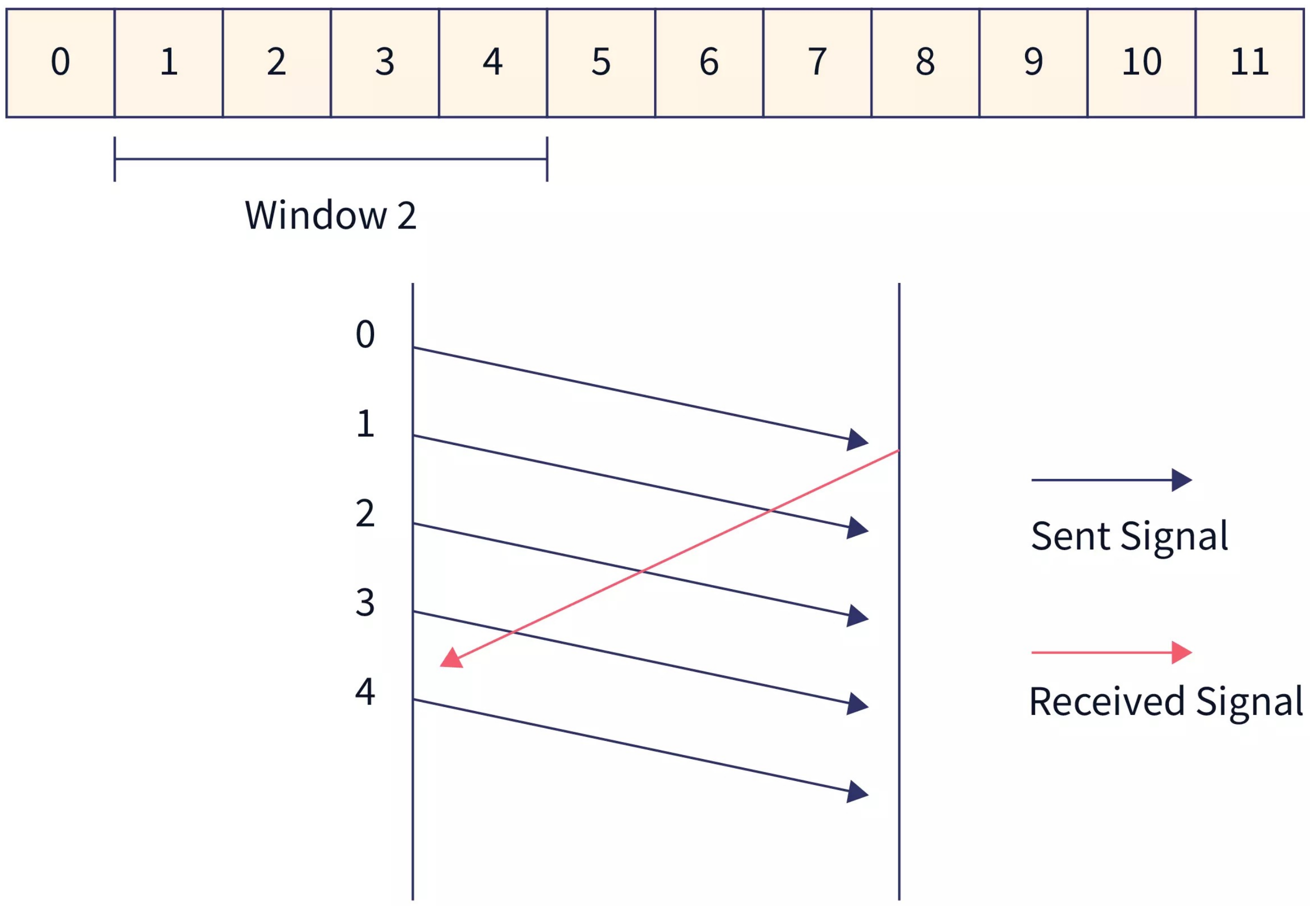


A sliding window of size 1, with a 3-bit sequence number.
(a) Initially.
(b) After the first frame has been sent.
(c) After the first frame has been received.
(d) After the first acknowledgement has been received.

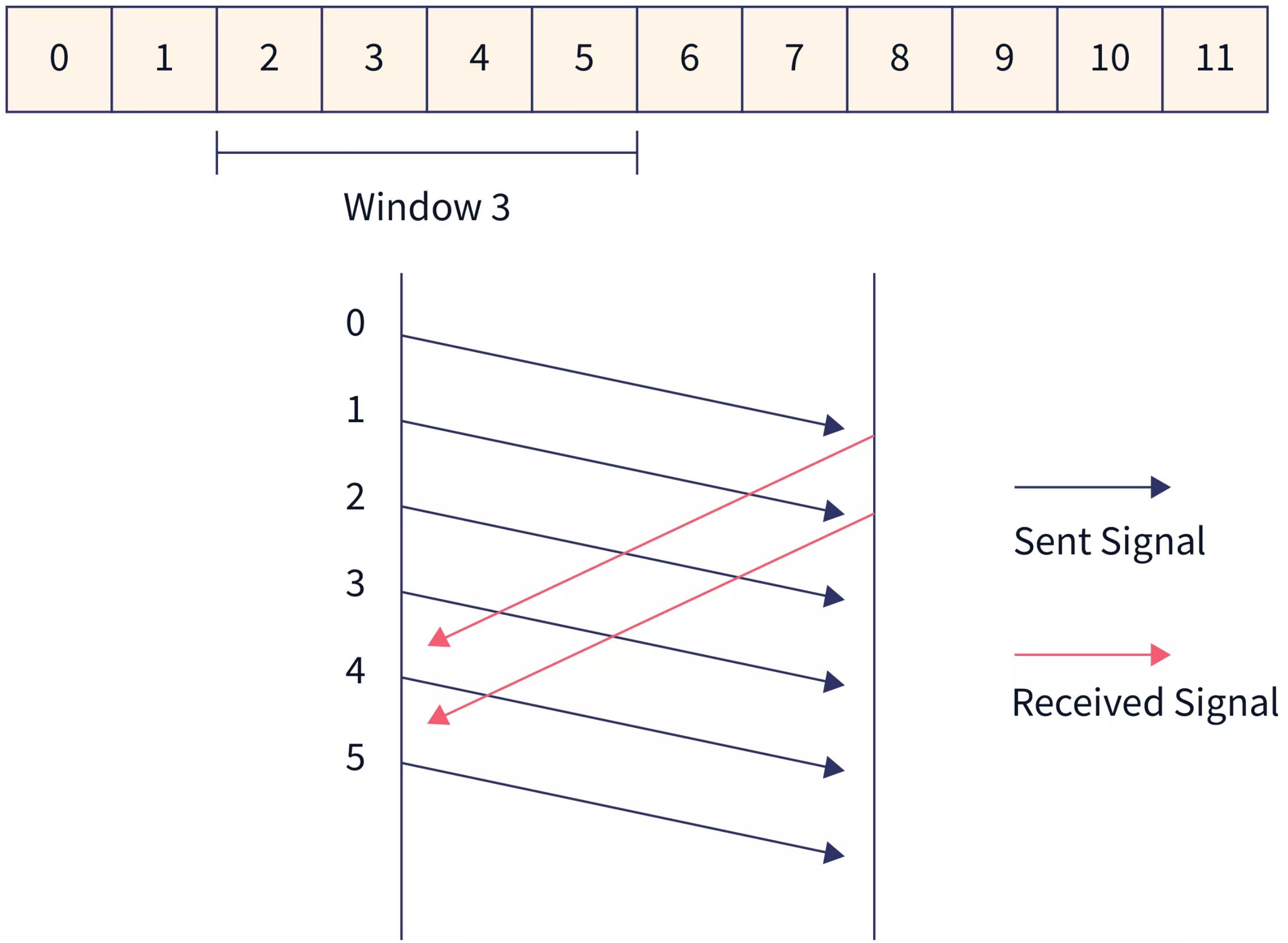
Go-Back-N ARQ (automatic repeat request)



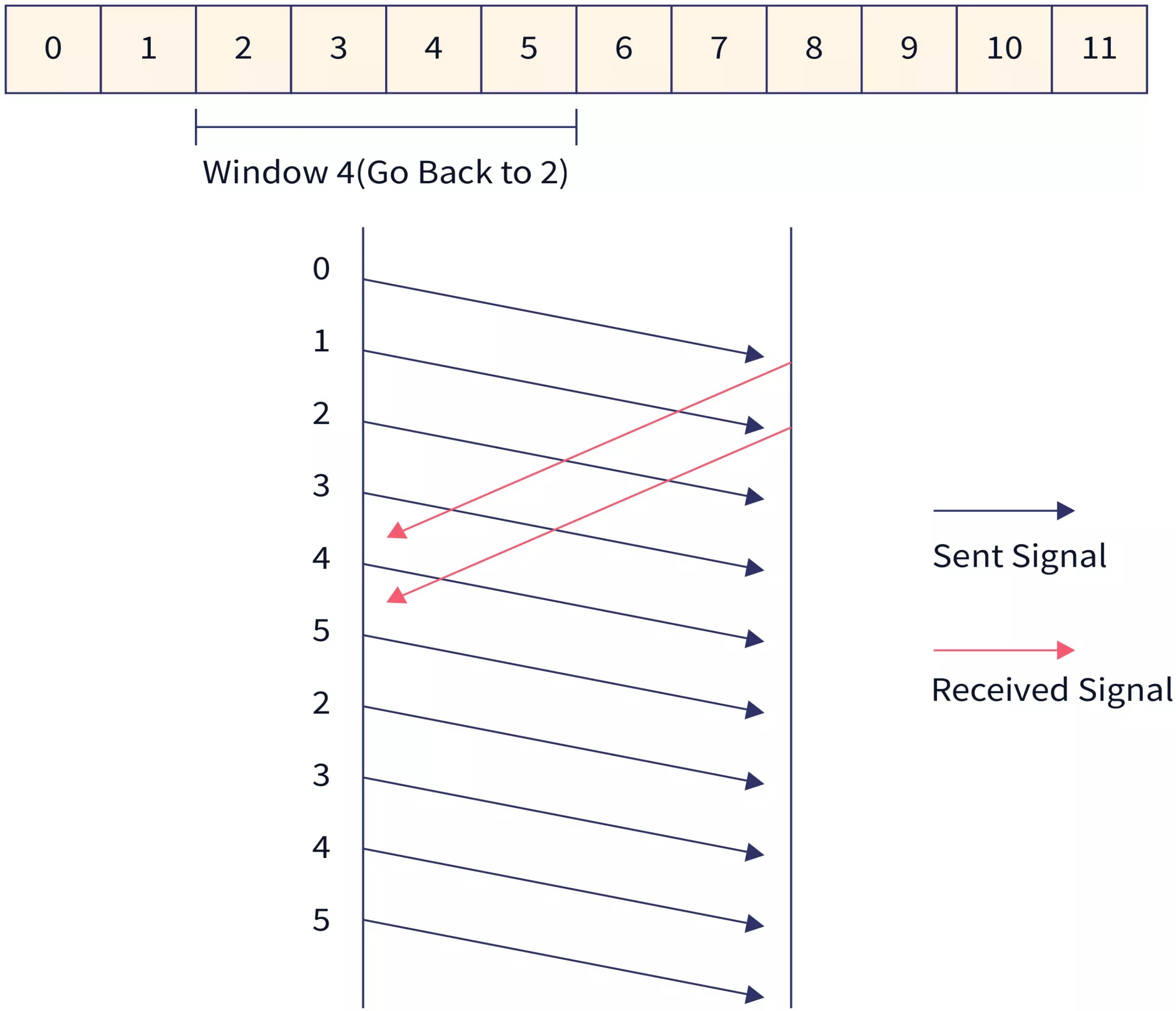
Go-Back-N ARQ (automatic repeat request)



Go-Back-N ARQ (automatic repeat request)



Go-Back-N ARQ (automatic repeat request)



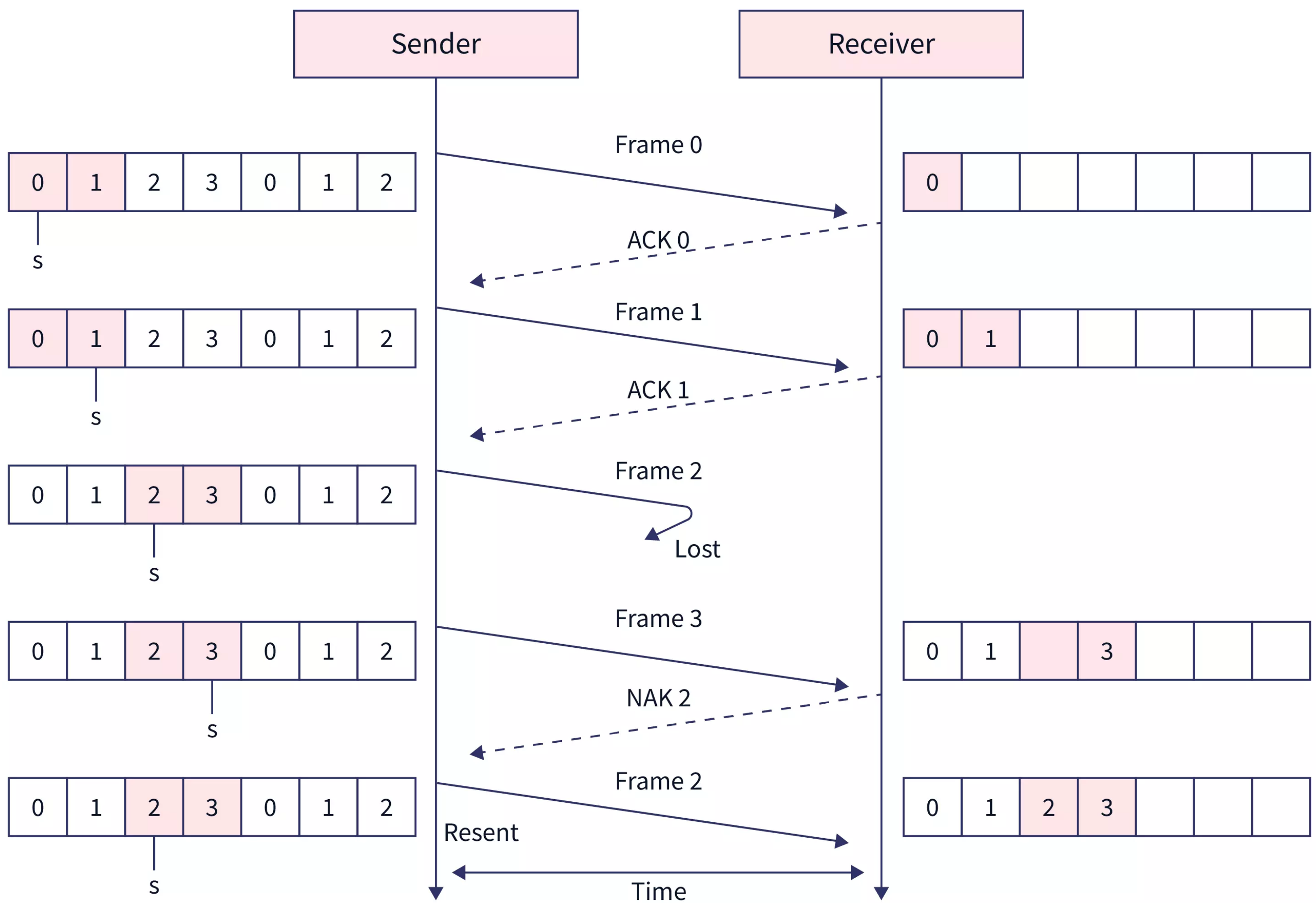
Go-Back-N ARQ (automatic repeat request)

- **Characteristics of Go-Back-N ARQ:**
  - The size of the sender window in Go Back N ARQ is equal to N.
  - The size of the receiver window in Go Back N ARQ is equal to 1.
  - When the acknowledgment for one frame is not received by the sender or the frames received by the receiver are out of order, then the whole window starting from the corrupted frame is retransmitted.

Selective Repeat Protocol (SRP)

- **Why we need SRP?**
  - If the connection is poor, there will be frequent loss of packets and the sender would have to retransmit all the outstanding packets. This wastes the bandwidth of the channel.
  - The size of the receiver window is 1 in go-back-N.
  - This retransmission of packets increases the traffic on the network creating a cumulative increase in congestion.

Selective Repeat ARQ (automatic repeat request)



主观题 10分

设置

拥塞产生的原因?

正常使用主观题需2.0以上版本雨课堂

作答

拥塞产生的原因

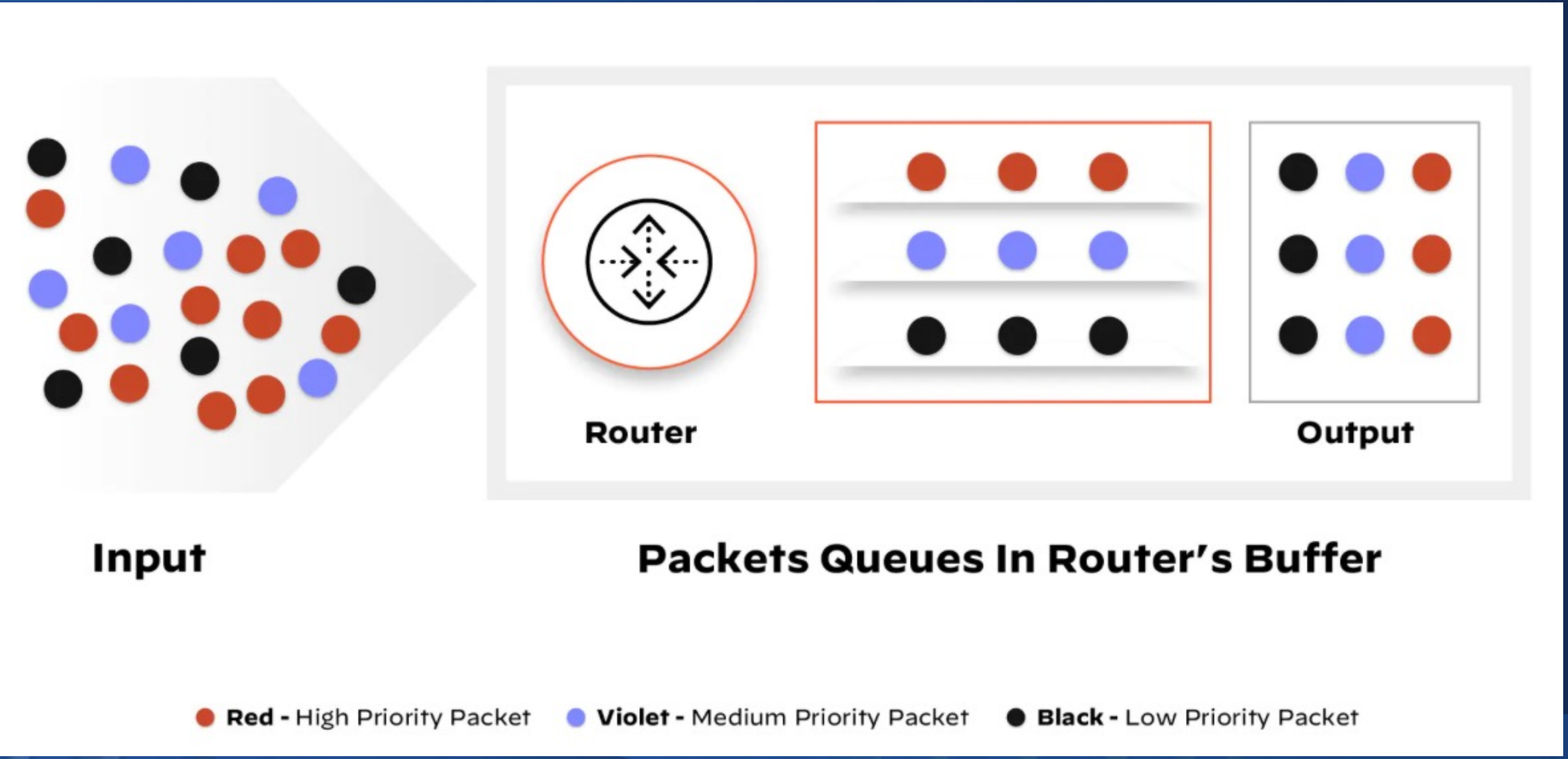
- Root cause for congestion is (dynamic) lack of bandwidth.
- – Demand for bandwidth is greater than capacity.
- – Sudden surge in demand.
- – Unexpected traffic flowing into the links due to routing. .
- Link down (某条链路突然断开)

- 拥塞引起的影响：时延，时延抖动，数据丢失

Congestion scenarios: Speed Mismatch

- LAN – WAN interconnect.
- Interconnection of high bandwidth LAN links to low bandwidth links.
- Problems:
- – Traffic from high bandwidth links gets choked on entering low bandwidth links.
- – Buffer exhaustion on devices.

Buffers can become full



Congestion scenarios: Aggregation

- Traffics from multiple links **aggregate** (聚合） into a single link of small bandwidth.
- Problems
- – Similar to speed mismatch. Here aggregation is the reason for the perceived speed mismatch.
- – Aggregated link is congested.
- – Buffer is full on devices.

拥塞控制的原则

- Monitor system to **detect** when and where congestion occurs
- **Pass** this information to places where action can be taken
- **Adjust** system operation to correct the problem

The approach that the Internet takes is to use a distributed algorithm to converge to an efficient and fair outcome.

How to achieve this?

Approach: sense and react

Example: taking a shower

Use a feedback loop with signals and knobs

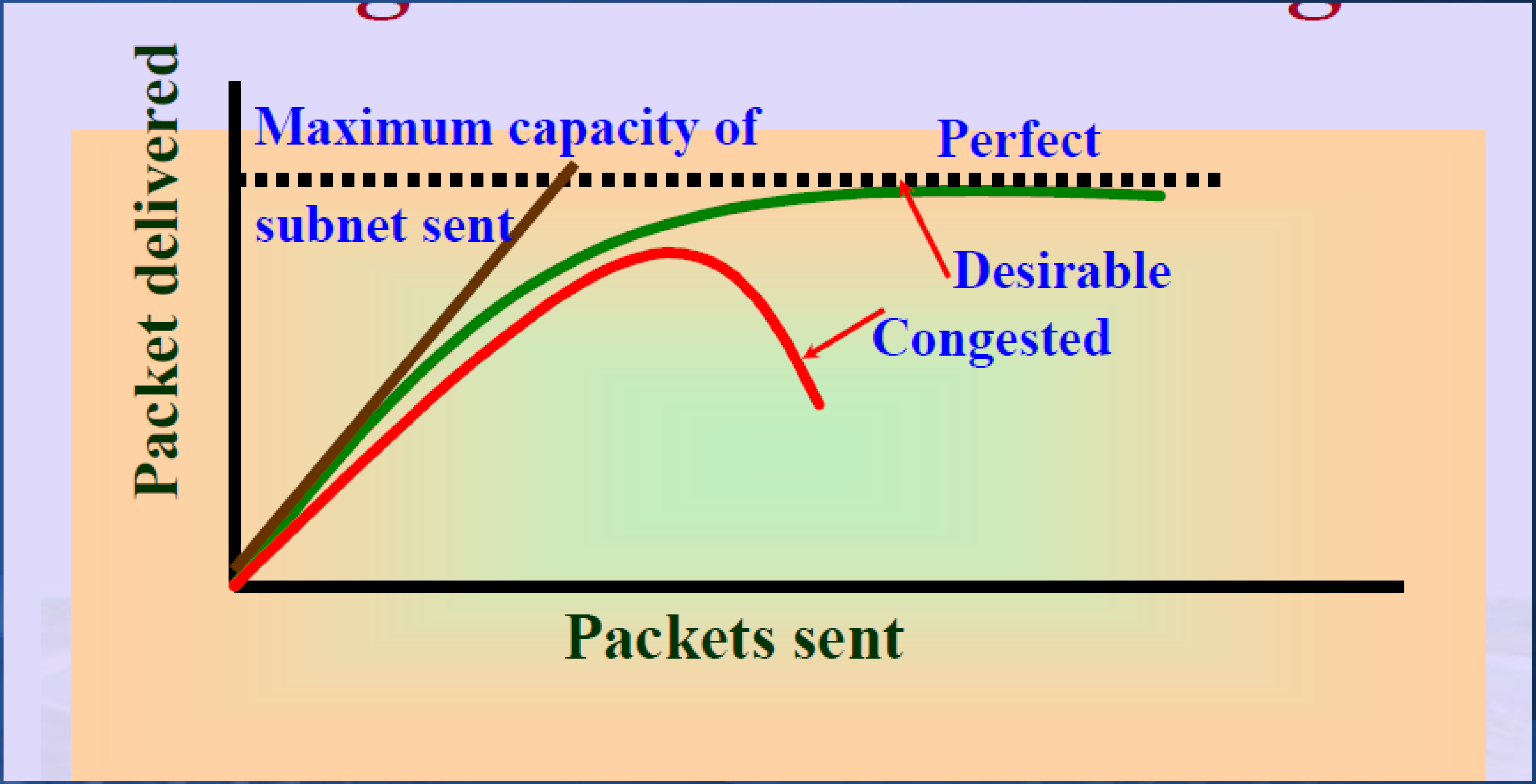
Signals and Knobs in Congestion Control

- **Signals**
  - **Packets being ACK'ed**
  - **Packets being dropped (e.g. RTO)**
  - **Packets being delayed (RTT)**
  - **Rate of incoming ACKs**

Implicit feedback signals measured directly at sender. (There are also explicit signals that the network might provide.)

- **Knobs**
  - **What can you change to "probe" the available bottleneck capacity?**
  - **Suppose receiver buffer is unbounded:**
  - **Increase window/sending rate: e.g., add x or multiply by a factor of x**
  - **Decrease window/sending rate: e.g., subtract x or reduce by a factor of x**

Goodput



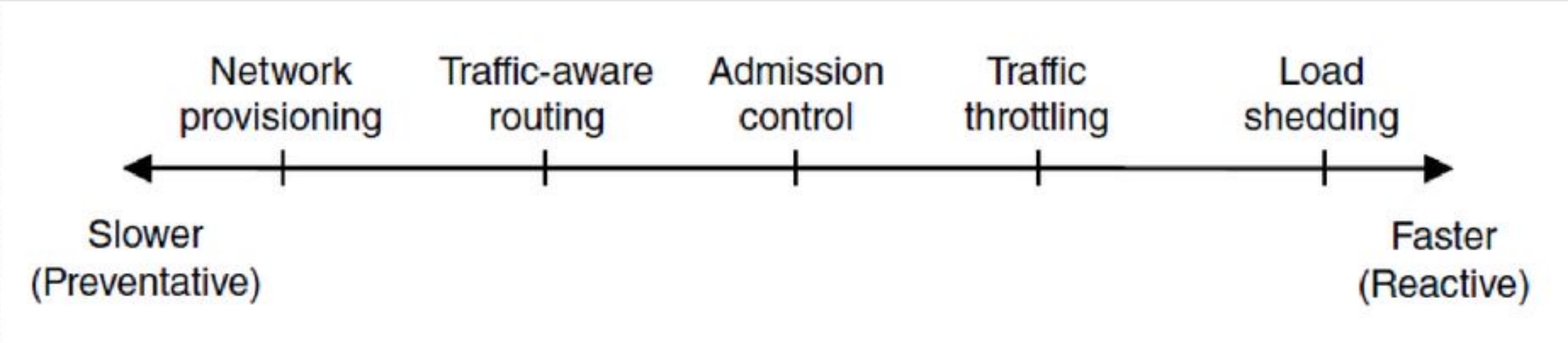
Maximum capacity of subnet sent
Perfect
Desirable
Congested
Packet delivered
Packets sent

拥塞控制技术手段

- Network provisioning （网络配置）
  - Network deployment
- Traffic-aware routing
- Admission control
- Traffic throttling （流量限流）
  - **ECN (explicit congestion notification)**
- Load shedding（减载）

Approaches to Congestion Control

month s-min ms-s ms-s ms-s



- Increase resource
- Reduce load

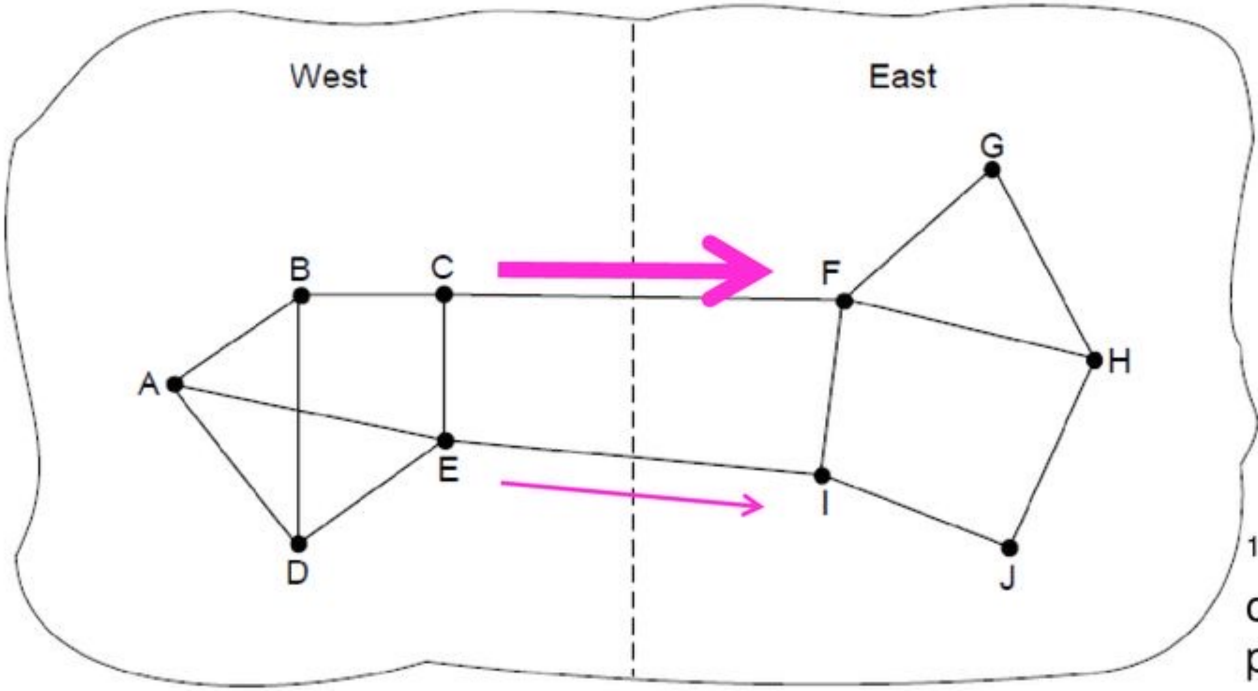
Timescales of approaches to congestion control

Traffic-Aware Routing

Shifting traffic away from congested regions by setting the link weight to be a function of the link bandwidth and propagation delay plus the (variable) measured load or queuing delay. Least weight paths favor paths that are more lightly loaded. Rarely done today, preferring traffic engineering[1] (TE) instead (e.g., QoS).

Choose routes depending on traffic, not just topology

- E.g., use *EI* link for West-to-East traffic if *CF* is loaded
- But take care to avoid oscillations (i.e., convergence issues)



[1]TE done outside of routing protocols

Admission Control

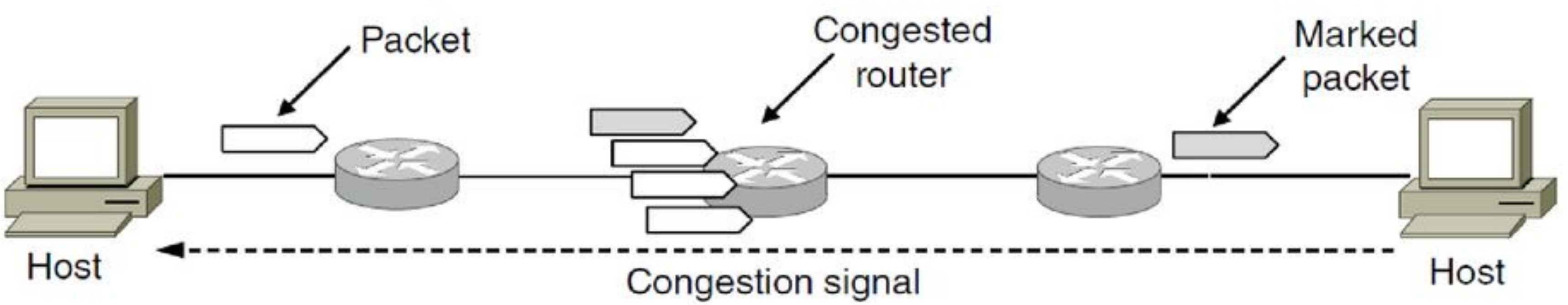
- Admission Control（接纳控制）是用于决定**是否接受**新的连接或请求以**使用网络资源**的一种机制。

Traffic Throttling

Network aims to operate just before the onset of congestion. Requires (1) routers must be able to discern when congestion is (about to) occur (e.g., queueing delay) and (2) routers must be able to deliver timely feedback to senders to throttle back rate

Congested routers signal hosts to slow down traffic

- ECN (Explicit Congestion Notification) marks packets and receiver returns signal to sender
  - Routers set the 2 ECN bits in IP packet header signals that router is experiencing congestion. Destination echoes this back to sender in reply
    - » ECN bits are the 2 least significant (rightmost) bit in DiffServ field in IP header
    - » In TCP, echo reply is indicated using ECE bit of TCP header – sender then knows to throttles back packet rate at the TRANSPORT Layer



Load shedding

Load shedding (减载)

– Discard packets

• question what to discard?

• file transfer – Keep old, discard new

• real-time app – keep new, discard old

• need more intelligence:(?)

– Some packets are more important

» **Video – full frame(don't discard)-difference frame (discard)**

» **Sender prioritizes packets!**

Other solutions

- Traffic prediction?
- Router informs neighbor of possible congestion
- Traffic shaping
- Regulate the packet rate
- VC (virtual circuit) - traffic characteristics
- ...

拥塞控制vs流量控制

- 流量控制：end-to-end approach
  - Ensure that data is not sent faster than the receiver can process
- 拥塞控制：network-level approach
  - Ensure that data is not sent faster than the network can process

小结

- QoS、QoE基本概念
- 流量工程（TE）的基本概念
- QoS的流量控制与拥塞控制方法