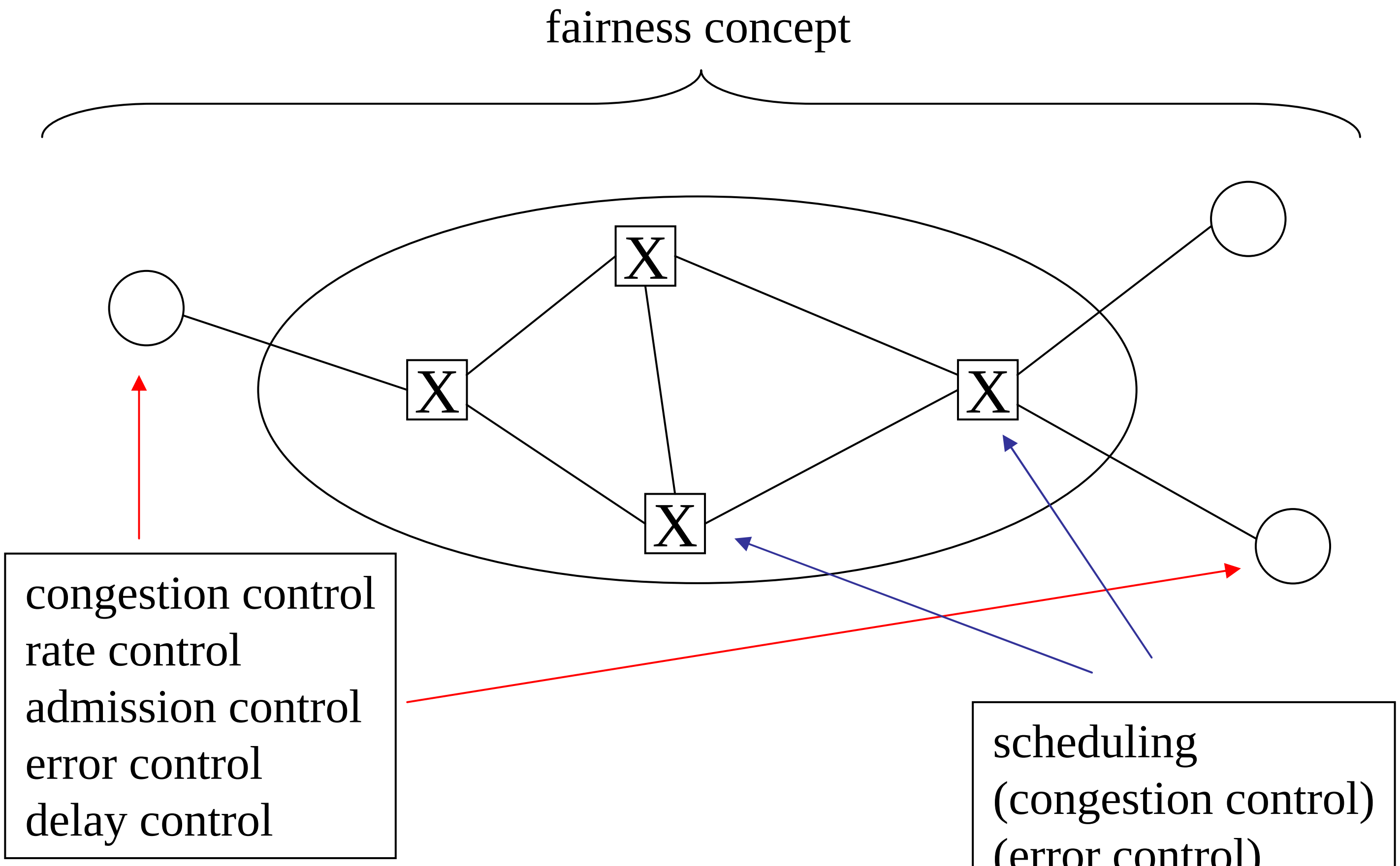
Congestion and flow control

- Lecture material:
  - Bertsekas, Gallager, *Data networks*, 6.1-2
  - J-Y Le Boudec, “Rate adaptation, congestion control and fairness: a tutorial,” Nov. 2005
- Reading for next lecture:
  - J. Padhye, F. Firoiu, D. Towsley, J. Kurose, "Modelling TCP throughput: a simple model and its empirical validation," ACM Sigcomm, 1998
  - T. Lakshman, U. Madhow, "The performance of TCP/IP networks with high bandwidth-delay products and random loss," IEEE TON, June 1997

Control functions in communication networks



Comment: end-to-end argument, read
Saltzer, Reed, Clark, End-to-end arguments in systems design

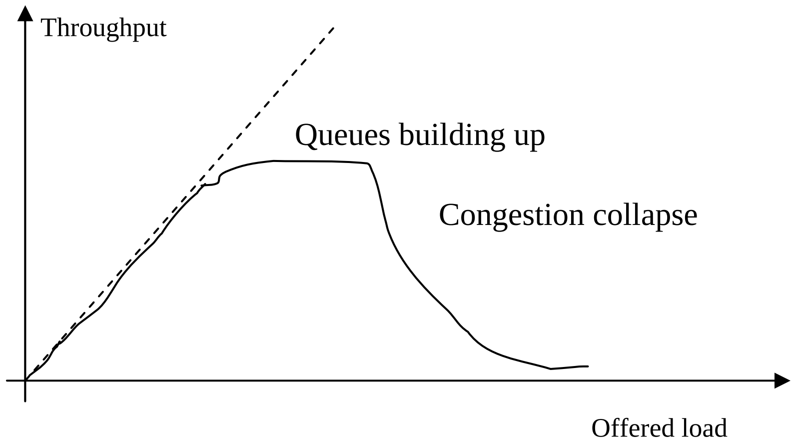
Congestion control

- Congestion control:
  - to regulate the packet population in the network
  - to share resources (link bandwidth, buffer space)
- *Flow control: between two users for speed matching*
  - sometimes means congestion control in the literature
- The main objectives of congestion control
  - Efficiency
    - high utilization (from the network provider's perspective)
    - high per flow throughput, low delay (from users' perspective)
  - Fairness: fair allocation of resources
    - measured with some fairness index, e.g.,

$$f = \frac{\left(\sum_n x_i\right)^2}{n\sum_n x_i^2}, \quad f \le 1$$

$$f = 1 \quad if \quad x_i = x$$

Congestion and flow control



- What happens if the incoming traffic is not restricted?
  - Bottleneck links: the offered traffic is higher than the link transmission capacity: temporarly (bursts arriving) or permanently
- What happens at bottleneck links?
  - queue sizes grow, end-to-end delays increase
  - queue space fills up, packets get dropped
  - packets are retransmitted by the applications, further increasing the load
- Congestion collaps: the network throughtput decreases and delays become excessive

Congestion control techniques

- Should depend on the service requirements of the application
  - voice, video: minimum bandwidth requirement, delay sensitive
  - data: requires strict error control
- Call blocking at the network edge + rate control
  - calls blocked if resources are not available
  - the rate of accepted calls is controlled
- Packet discarding at a network node
  - at buffer overflow or earlier
  - discarding policy - fairness, service differentiation
- Packet blocking at the network edge
  - packet waits in a queue outside the network
- Packet scheduling at the network nodes
  - selective transmission of packets - fairness, service differentiation

Group work: find examples for the techniques above.

Congestion control techniques – examples

- Call blocking at the network edge + rate control
- Packet discarding at a network node
- Packet blocking at the network edge
- Packet scheduling at the network nodes

Congestion control techniques

- Should depend on the service requirements of the application
  - voice, video: minimum bandwidth requirement, delay sensitive
  - data: requires strict error control
- Call blocking at the network edge + rate control
  - calls blocked if resources are not available
  - the rate of accepted calls is controlled
- Packet discarding at a network node
  - buffer overflow or active queue management
  - discarding policy - fairness, service differentiation
- Packet blocking at the network edge
  - packet wait in a queue outside the network
- Packet scheduling at the network nodes
  - selective transmission of packets - fairness, service differentiation

End-to-end window congestion control

- Upper bound on the number of packets transmitted by A and not acknowledged by B (figure!)
- Input rate reduced if acknowledgements arrive slowly - congestion control
- Simple case: constant window size, infinite buffers

  W: window size – fix!

  X: packet transmission time (packet size/link rate [sec])

  rtt: round-trip delay (propagation, transmission and queuing delays)

  r: maximum packet transmission rate [packet/sec]

$$r = \min\left\{\frac{1}{X}, \frac{W}{rtt}\right\}$$

- Rate invers proportional to round-trip delay (determined by the queuing delays)
- Reacts to congestion in W packets transmission time – large window means slow reaction to congestion

Window control - limitations

- Can not guarantee minimum transmission rate – not for streaming

- For elastic traffic (minimum rate guarantee is not required)
— How to select window size?
  — large windows to allow high throughput if the network is low loaded
— Window size vs. end-to-end delay?

Window control - packet delay

May be group work: B-G pp. 503-506

- Packet delay T ?
  - Simple queuing analysis in “steady state”
  - Calculate delay at given network load
  - What is the queuing system?
    - the whole network
  - Input we know:
    - arrival intensity - constant
    - number of packets in the network
  - Use Little-theorem

$$N = \sum_{i=1}^{n} const_i W_i$$

$$T = \frac{N}{\lambda} = \frac{\sum_{i=1}^{n} const_i W_i}{\lambda}$$

n: active flows

W_i : window size for flow *i*

N: number of packets in the network

λ: aggregate throughput of all flows (no loss – infinite buffers!)

- For constant aggregate throughput
- Delay is proportional to the number of active sessions (or the sum of window sizes)

Window control - open questions

- High speed / large delay networks: how to select the window size?
  - Small window: OK packet delay, but low throughput
  - Large window: OK throughput, but high packet delay
- Dynamic window sizes are necessary to follow the network load
  - small window if the network is congested
  - large window if the network is low loaded
  - congestion is controlled by the window size – not by the RTT
  - TCP

- Performance with dynamic window sizes - comes now
- More realistic models for TCP performance - next lecture
- TCP in high bandwidth / large delay networks - next lecture

TCP congestion control

- Window based congestion control
- Dynamic window size:
  - decrease window size if congestion is detected (e.g., by packet loss)
  - increase window size if current rate is sustainable
- How to increase and decrease the window sizes?
  - Additive-increase, multiplicative decrease (AIMD)
  - Efficieant and fair: Chiu and Jain, 1989
- Congestion - decrease phase: $w_{i+1} = aw_i, 0 < a < 1$
- Probing - increase phase in each rtt: $w_{i+1} = w_i + b, 0 < b << w_{\max}$
- Thus, it is linear control: $w_{i+1} = aw_i + b$
- TCP: a=0.5, b=1
- TCP additional phases: slow start, fast recovery – not considered now

Analysis – AIDM model

- How does the throughput depend on the loss rate?

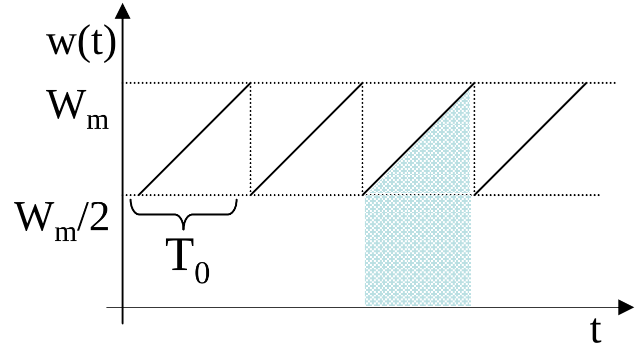
- Assumptions – for a very simplified case
  - stationary conditions (background traffic does not change)
  - loss due to congestion only
  - transients negligible (long flows)

  $\Rightarrow$ we model a static congestion avoidance phase
  - constant round trip time (rtt)
  - rtt >> transmission time
  - constant packet size L
  - low loss probability

  (to simplify calculations)

TCP throughput vs. packet loss

- Statement: throughput in steady state varies as the inverse-square-root of the loss rate



T_0: cycle (s)
W_m: max window (packets)
rtt: round trip time (s)
N: packets transmitted in one cycle
p: packet loss probability
L: packet length (bit)
θ: throughput (bit/s)

$$T_0 = \frac{W_m}{2} rtt, \qquad r = \frac{W(t)}{rtt}$$

$$N = \int_0^{T_0} \frac{W(t)}{rtt} dt = \frac{1}{rtt} \frac{W_m/2 + W_m}{2} T_0 = \frac{3}{8} W_m^2$$

$$p = \frac{1}{N}, \qquad W_m = \sqrt{\frac{3}{8} N} = 2\sqrt{\frac{2}{3}} \frac{1}{\sqrt{p}}$$

$$\theta = \frac{N-1}{T_0} L = \frac{\left(\frac{1}{p} - 1\right) L}{2\sqrt{\frac{2}{3}} \frac{1}{\sqrt{p}} \frac{rtt}{2}} = \sqrt{\frac{3}{2}} \frac{L}{rtt} \left(\frac{1-p}{p}\right) \sqrt{p} \approx \sqrt{\frac{3}{2}} \frac{L}{rtt} \frac{1}{\sqrt{p}}$$

Congestion control for streaming flows

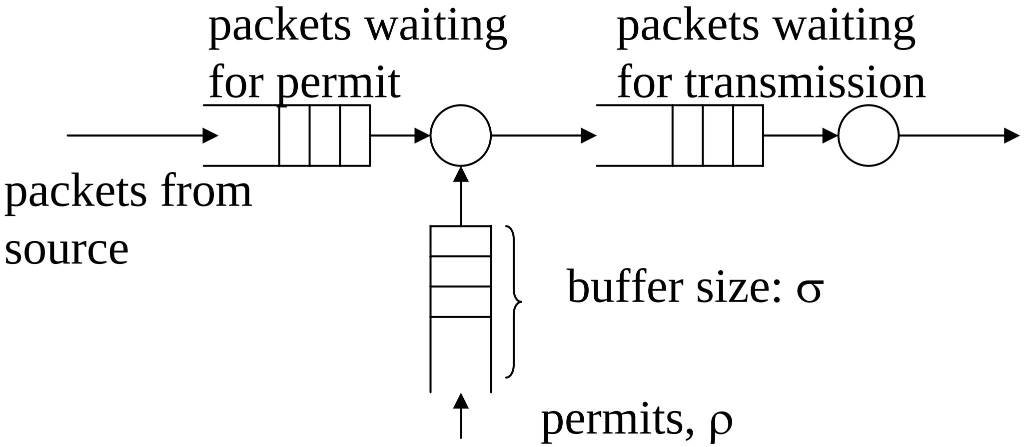
Which one of these techniques should be applied?

- Call blocking at the network edge + rate control
  - calls blocked if resources are not available
  - the rate of accepted calls is controlled
- Packet discarding at a network node
  - buffer overflow or earlier
  - discarding policy - fairness, service differentiation
- Packet blocking at the network edge
  - packet wait in a queue outside the network
- Packet scheduling at the network nodes
  - selective transmission of packets - fairness, service differentiation

Congestion control for streaming flows

Rate-control techniques

- To avoid congestion and packet loss
- Pre-defined flow rate
  - long term average rate
  - maximum rate
  - maximum burstiness – max. number of consecutive packets (or bits)
  - e.g., (σ,ρ) model
- Window based rate control (jumping or sliding window)
- Leaky-bucket scheme



Reading for next lecture

TCP throughput – random loss due to congestion

- J. Padhye, F. Firoiu, D. Towsley, J. Kurose, "Modeling TCP throughput: a simple model and its empirical validation," Sigcomm, 1998
  - read sections 1, 2-2.1, 3, 4, 5
- Home assignment:
  - summarize the main results of the paper in words
  - list the assumptions of the analysis in 2.1
  - how does the model differ from the one considered on the lecture?
- Presentation
  - explain, how TCP Reno works (including slow start, fast retransmit, etc.)
  - list and explain the assumptions used in model 2.1 – why are these assumptions necessary and how are they valid?
  - present the analysis in 2.1 and related performance figures from section 4

Reading for next lecture

TCP throughtput – high bandwidth-delay product networks

- T. Lakshman, U. Madhow, "The performance of TCP/IP networks with high bandwidth-delay products and random loss," TON, June 1997
  - Read sections 1 and 5
- Home assignment
  - summarize the main results of the paper – what is their importance?
  - list the assumptions of the analysis in section 5
- Presentation
  - explain what the expression high bandwidth-delay product means
  - focus on the TCP unfairness towards sessions with high bandwidth-delay product (section 5)
  - give the intuitive explanation of this unfairness
  - give the main ideas of the analytical model
  - discuss the resulting equations of the analysis