CS144
An Introduction to Computer Networks

Congestion

Basic Ideas

Nick McKeown
Professor of Electrical Engineering
and Computer Science, Stanford University

Outline of next few videos

What is congestion control?

Basic approaches to congestion control

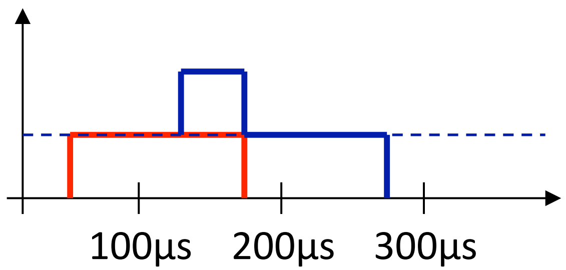
- In the network
- From the end host

TCP Congestion Control

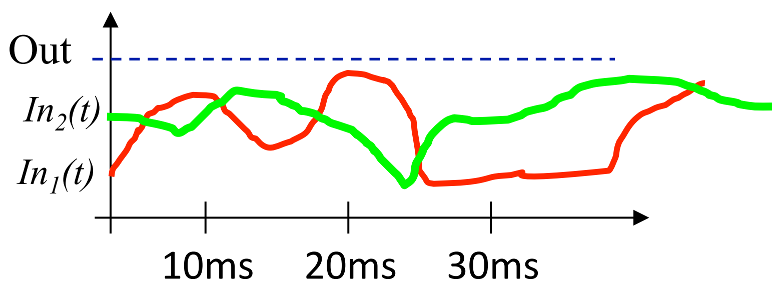
- TCP Tahoe
- TCP Reno
- TCP RTT Estimation
- Performance in practice

Time Scales of Congestion

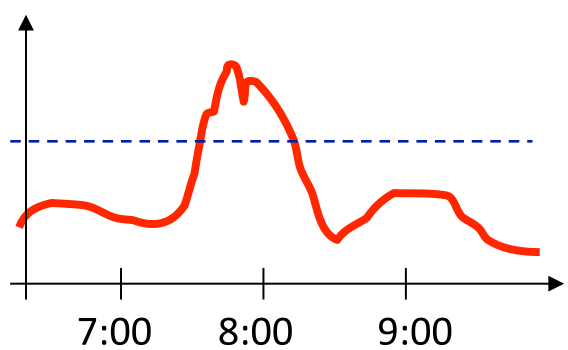
Two packets colliding at a router



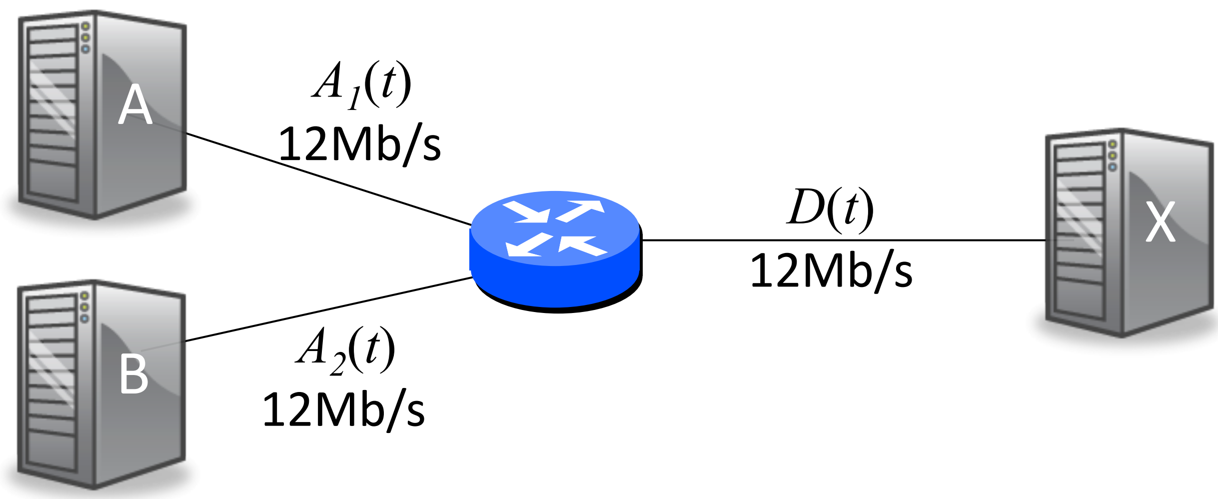
Flows using up all link capacity



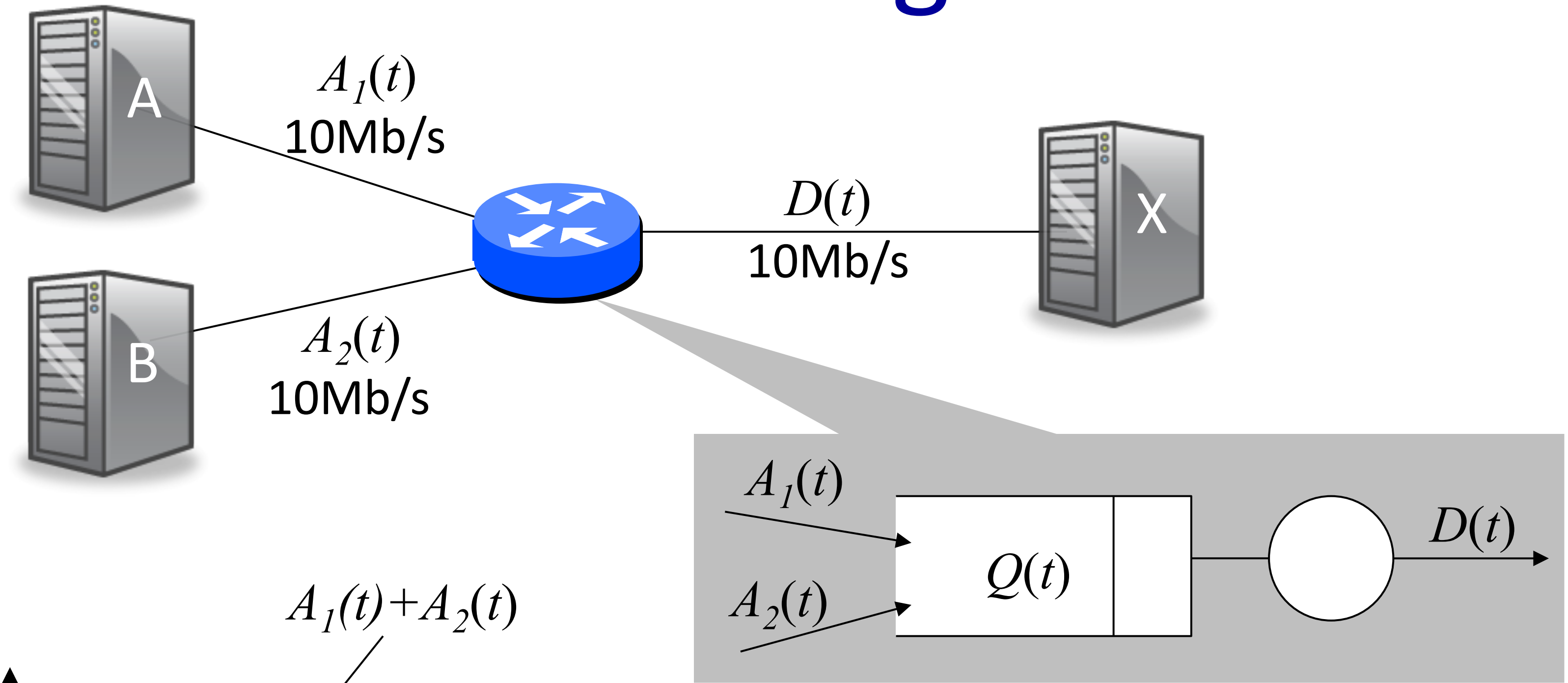
Too many users using a link during a peak hour

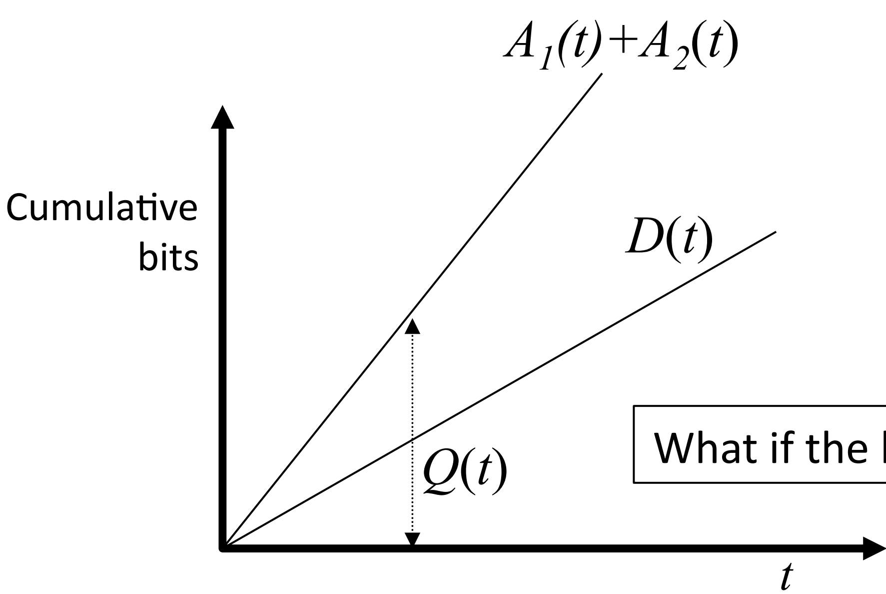


What causes congestion?



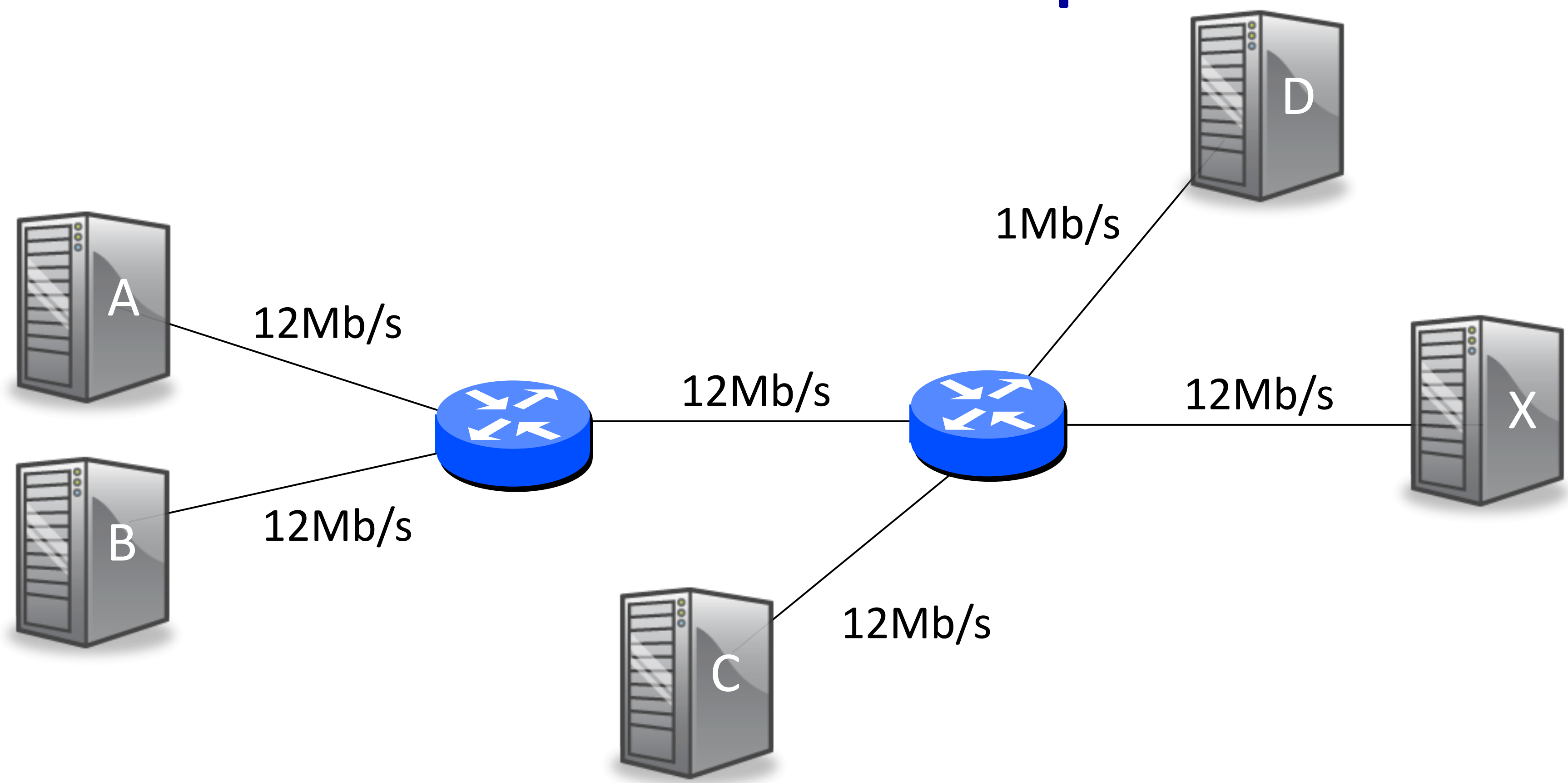
What causes congestion?





What if the buffer is finite?

Another example



Congestion is unavoidable

Arguably it's good!

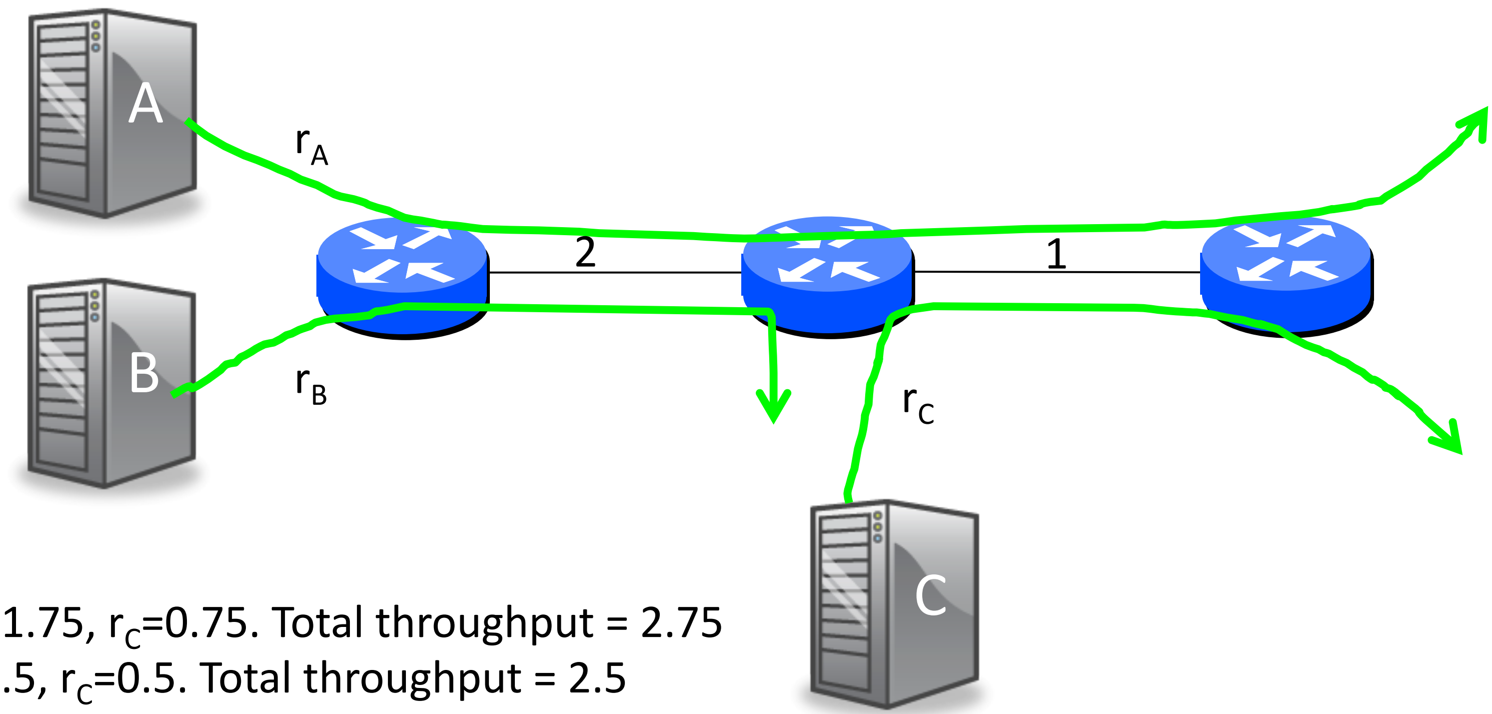
1. We use packet switching because it makes efficient use of the links. Therefore, buffers in the routers are frequently occupied.
2. If buffers are always empty, delay is low, but our usage of the network is low.
3. If buffers are always occupied, delay is high, but we are using the network more efficiently.

Observations

1. Congestion is inevitable, and arguably desirable.
2. Congestion happens at different time scales – from two individual packets colliding, to some flows sending too quickly, to flash crowds appearing in the network.
3. If packets are dropped, then retransmissions can make congestion even worse.
4. When packets are dropped, they waste resources upstream before they were dropped.
5. We need a definition of <u>fairness</u>, to decide how we want flows to share a bottleneck link.

Fairness and throughput

Fairness and throughput



Allocations:

1. r_A=0.25, r_B=1.75, r_C=0.75. Total throughput = 2.75
2. r_A=0.5, r_B=1.5, r_C=0.5. Total throughput = 2.5

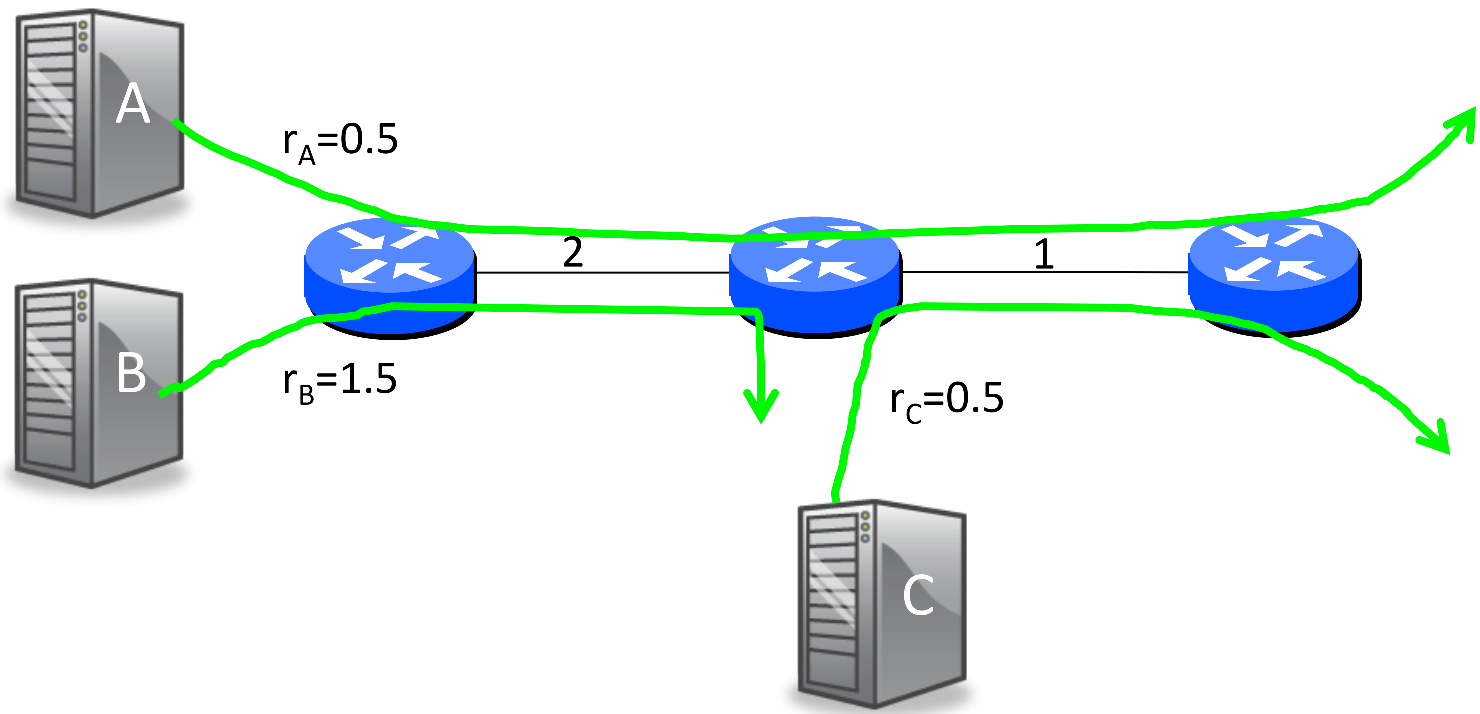
Allocation #2 is "fairer". How can we define fairness?

Max-min Fairness

Definition:

An allocation is max-min fair if you can not increase the rate of one flow without decreasing the rate of another flow with a lower rate.

Max-min fair allocation



Max-min fairness: Single link

Definition is intuitive and simple on a single link.

Goals for congestion control

1. High throughput: Keep links busy and flows fast
2. Max-min fairness
3. Respond quickly to changing network conditions
4. Distributed control

<end>