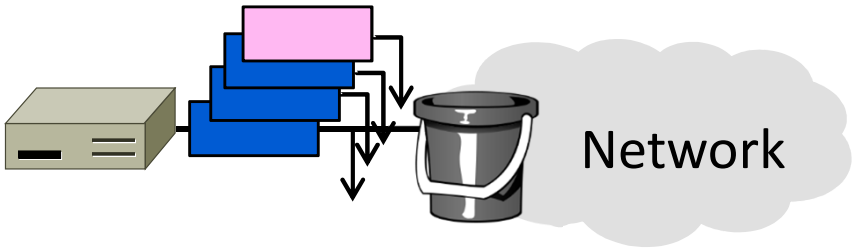
Introduction to Computer Networks

Traffic Shaping (§5.4.2)

David Wetherall (djw@uw.edu)
Professor of Computer Science & Engineering
UNIVERSITY of WASHINGTON

Topic

- Shaping traffic to constrain bursts
  - Token buckets
    - Key building block for QOS

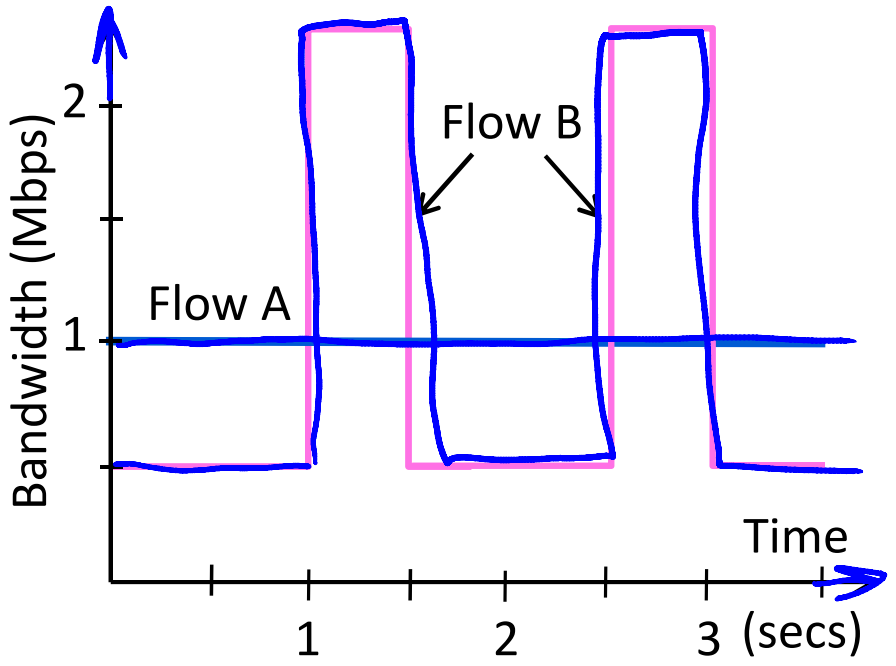


Motivation

- Shaping traffic flows constrains the load they may place on the network
  1. Limiting the total traffic enables bandwidth guarantees
  2. Limiting bursts avoids unnecessary delay and loss
- How should we shape traffic?
  - Real apps generate varying traffic – unrealistic to smooth it out

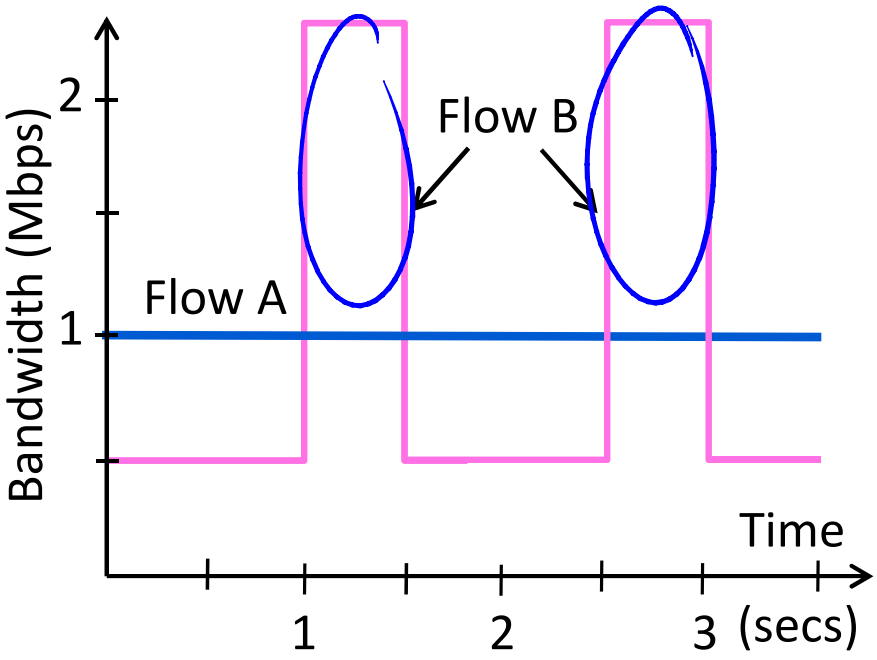
Motivation (2)

- Flow A and flow B have the same average rate
  - 1 Mbps over 3.5 secs
    - But they have very different behaviors!
- Average rate alone is not a good descriptor of behavior …



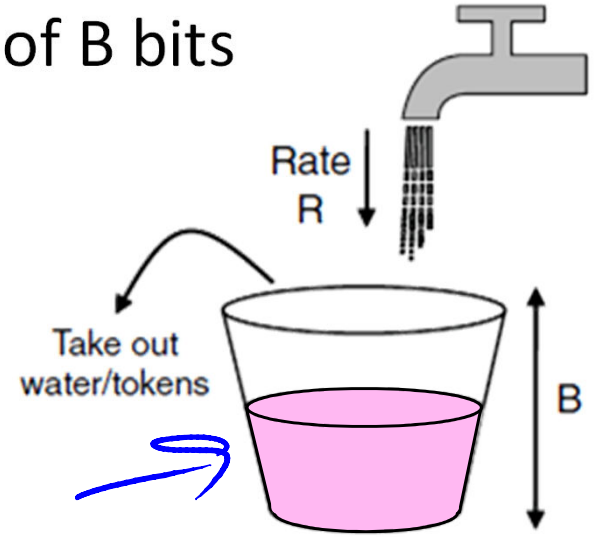
Motivation (3)

- How should we describe traffic flows to the network?
  - Average rate matters; relates to long-term bandwidth
  - Burstiness also matters; relates to short-term bandwidth
- Two characteristics useful
  - More expressive than average
  - Still relatively simple



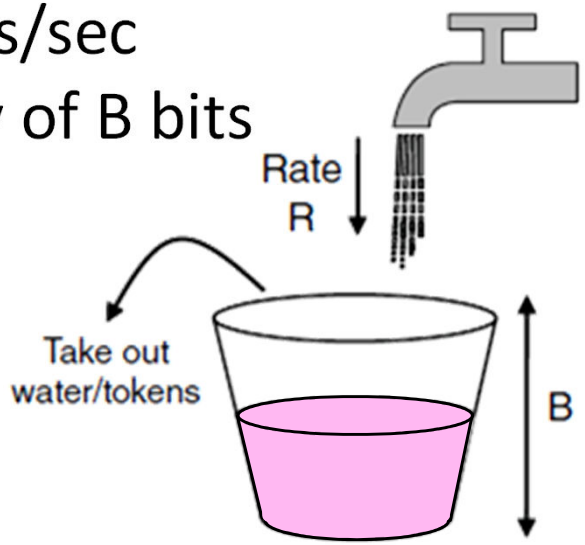
Token Bucket

- (R, B) token bucket constrains:
  - Average rate of R bits/sec
  - Bursts (over R) of B bits



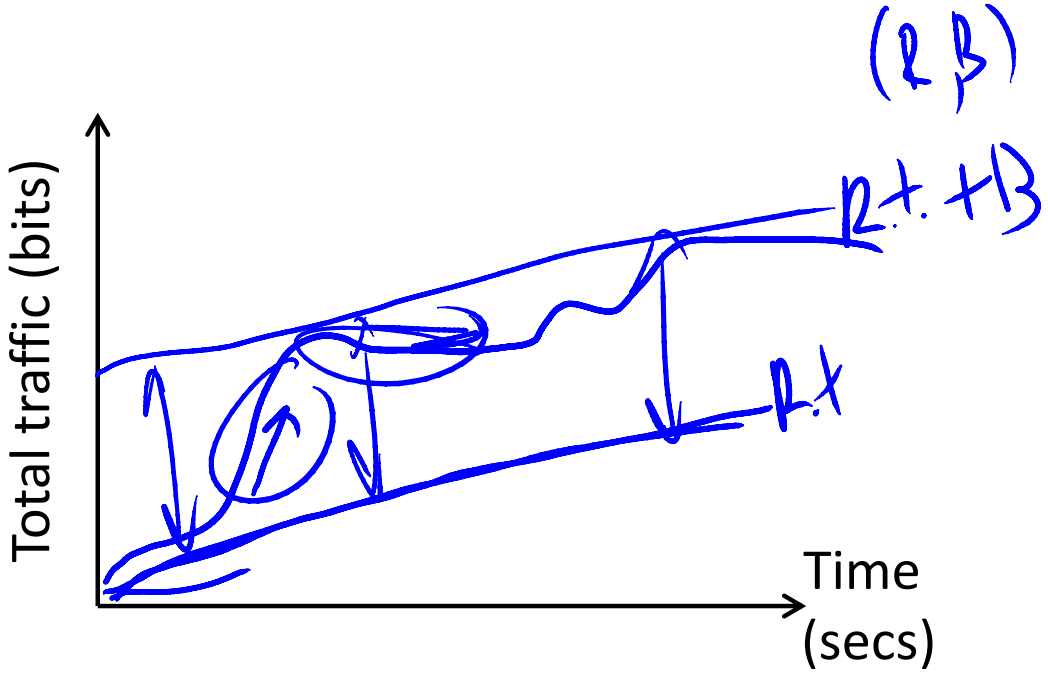
Token Bucket (2)

- Sending removes tokens (or credits) from the bucket; no credit, no send
  - Fill rate of R bits/sec
  - Bucket capacity of B bits



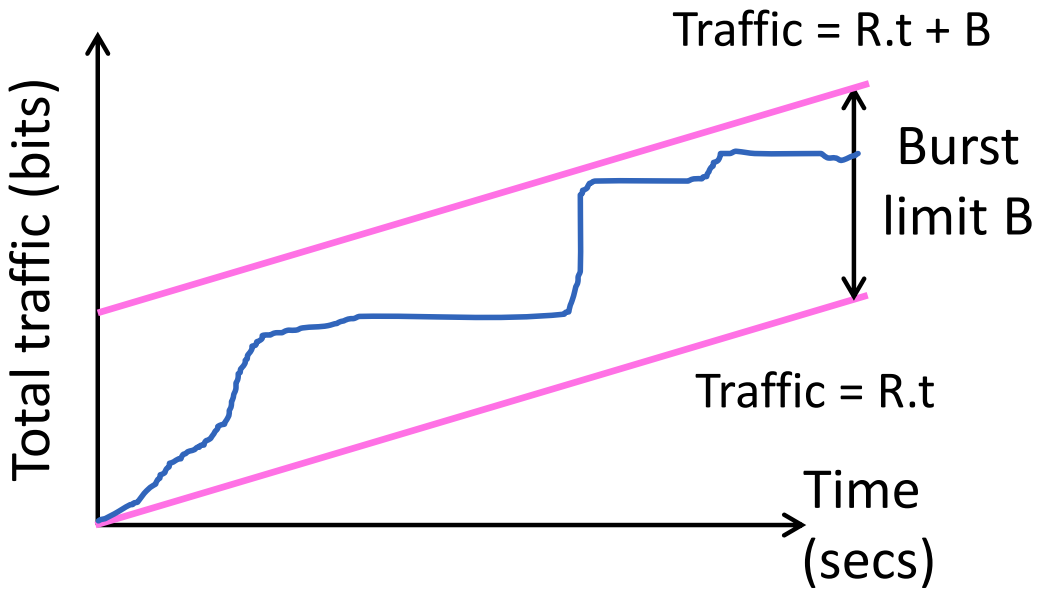
Token Bucket (3)

- Constrains greatest traffic over time



Token Bucket (4)

- Constrains greatest traffic over time



Shaping vs. Policing

- Shaping modifies traffic near the source to fit within an (R, B) profile
  - Run (R, B) token bucket at the source
  - Pass sent packets to the network when there are tokens
  - Delay (queue) packets while more tokens arrive
- Lets user condition their traffic to meet the network contract

Shaping vs. Policing (2)

- Policing verifies that traffic within the network fits an (R, B) profile
  - Run (R, B) token bucket at network edge
  - Let packets into the network when there are tokens
  - Demote or discard packets when there are insufficient tokens
- Lets network check traffic to verify it meets the user's contract

Usage for QOS

- Token buckets help the user and network regulate traffic for QOS
  - Network can limit the traffic for preferential treatment
  - User can flexibly select that traffic
- Special treatment is implemented with other means such as WFQ

END