Advanced Computer Networks

Packet Scheduling

Jianping Pan
Summer 2007

Review: Internet design

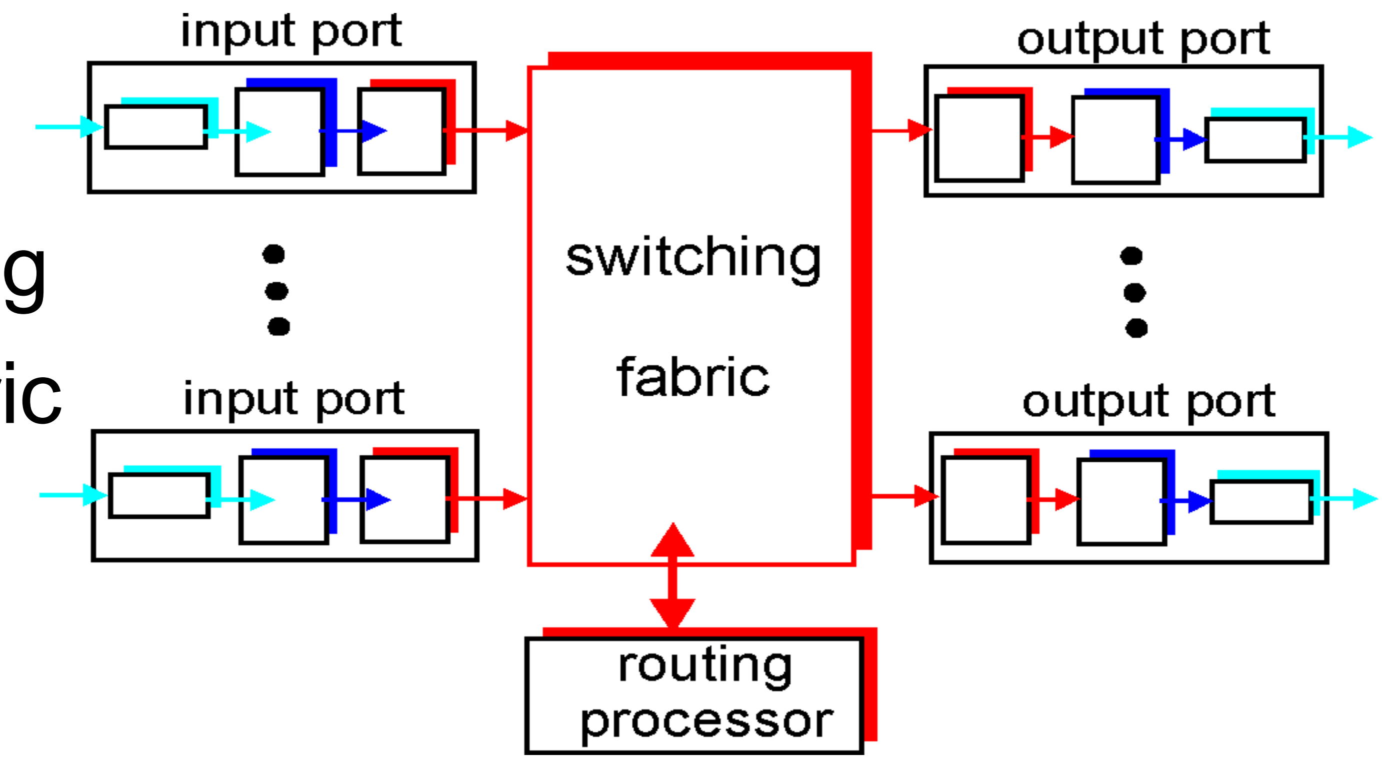
- Design principles
  - store-and-forward packet switching
  - end-to-end argument
- Network design
  - datagram routing
  - stateless network as much as possible

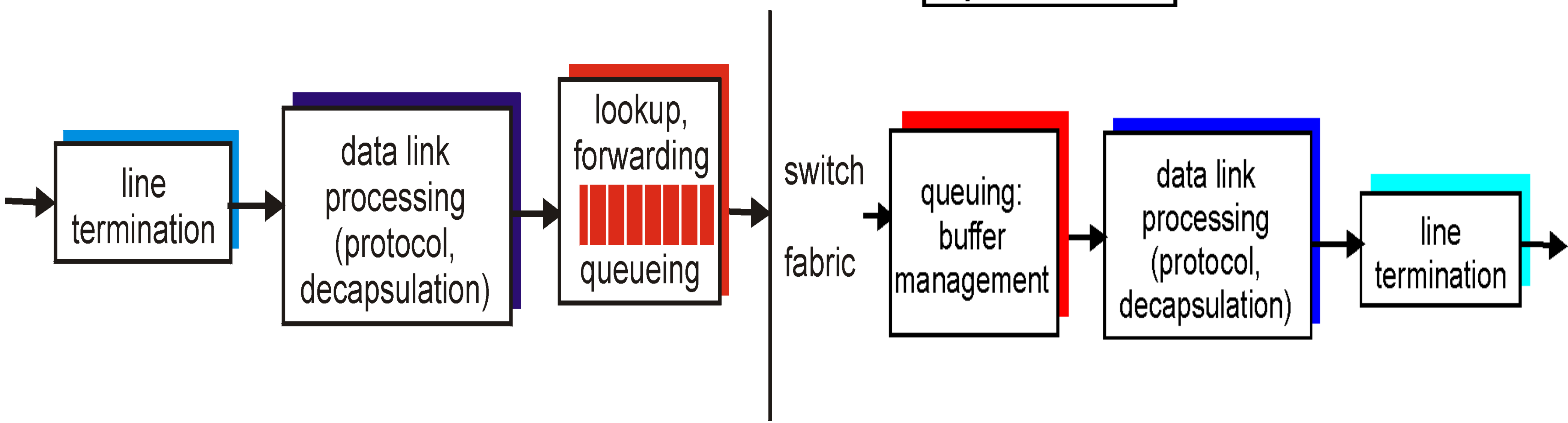
Reality check

- Internet
  - destination-oriented routing
    - inter-domain and intra-domain routing
  - drop-tail queuing
    - when buffer is full, drop incoming packets

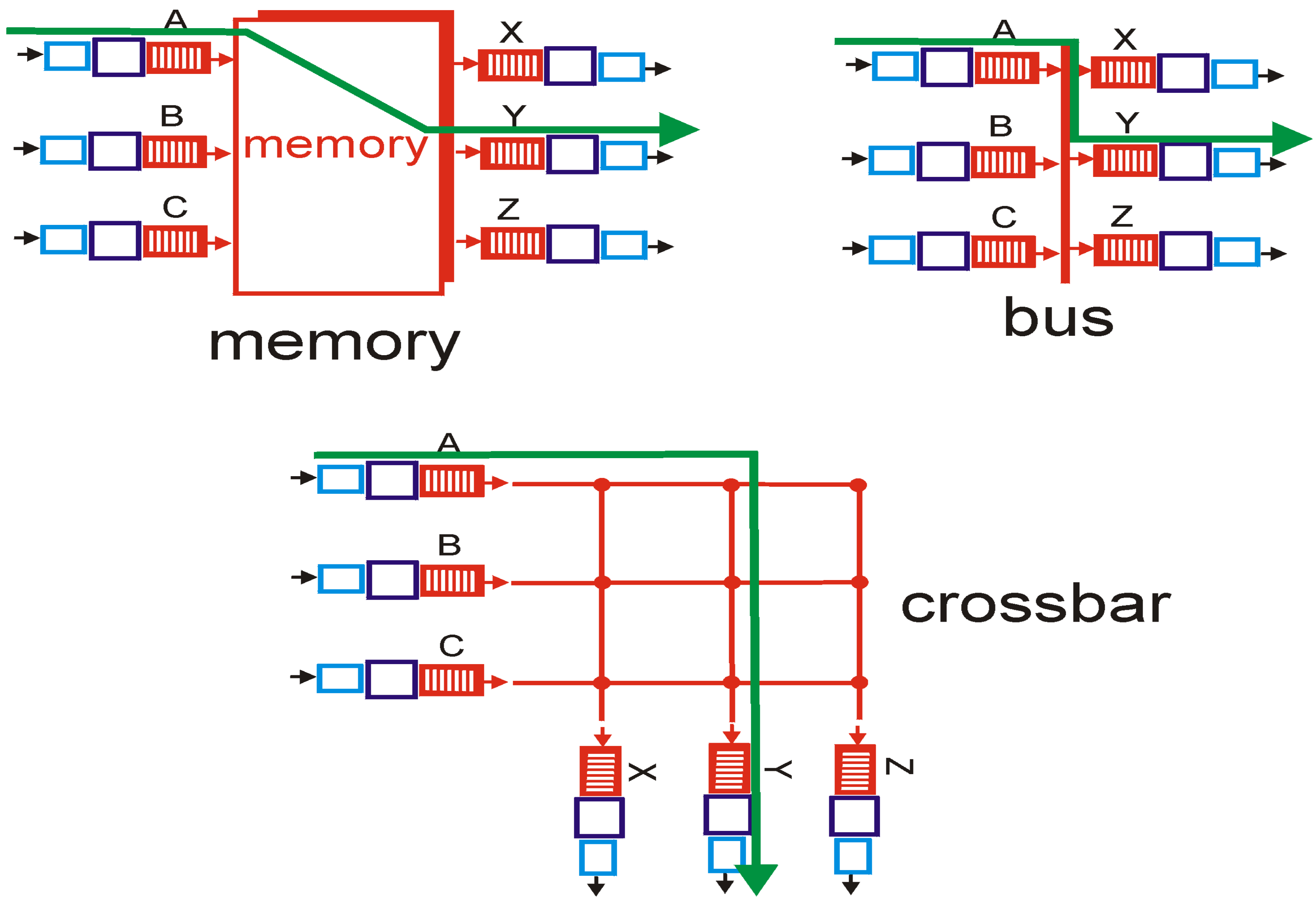
Inside a router

- Input queuing
- Output queuing
- Switching fabric





Switching fabric



Packet buffering

- When buffer is full
  - drop is inevitable
  - possible choices
    - drop-tail
      - the incoming packets
    - drop-head
      - the oldest packets
    - random drop
      - other packets in the queue
- Drop can happen even before the buffer is full
  - random early drop/detection (RED)
  - choke

Packet scheduling

- Scheduling
  - “who's the next?”
- Goals
  - sharing: multiplexing
    - e.g., first-come first-serve and drop-tail
  - isolation: fairness
    - e.g., circuit switching
  - and a balance of the above two!

Fairness measures

- Max-Min fairness
  - no one gets more than required (satisfied)
  - the excess, if any, shared by unsatisfied ones
- Algorithm
  - recursive allocation
- Example
  - A, B, C, D request 1, 2, 3, 4, respectively
  - resource available: 8

Round-robin scheduling

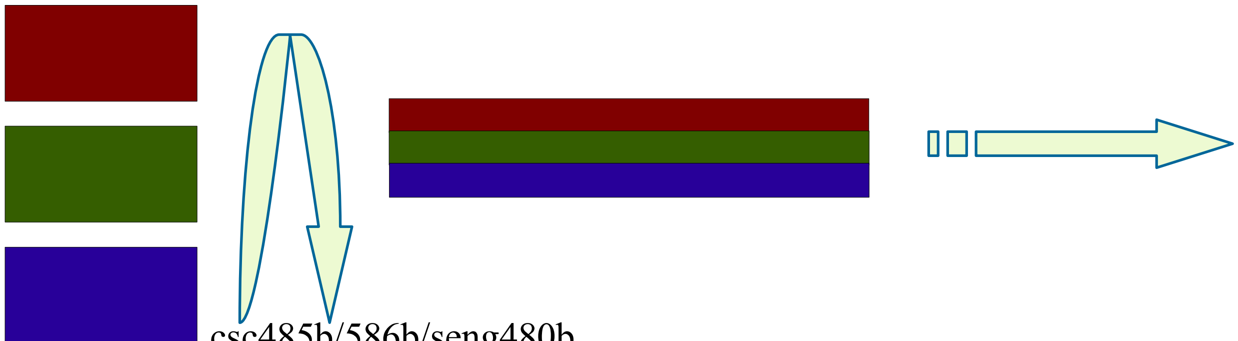
- Round-robin
  - one unit each round
- Weighted round-robin
  - multiple units each round
  - the number of units proportional to weight
- Example
  - weight: 0.2, 0.6, 1.5

Priority scheduling

- Serve the highest priority
  - preemptive
  - non-preemptive
- Static priority can cause starvation
- Dynamic priority
  - e.g., deadline

General processor sharing

- Assumption
  - serving in infinitesimal units
  - fluid-like
- Performance bound
  - for (r,b)-regulated flows, delay is bounded by b/r
- Not implementable
  - serving packet-by-packet



Weighted fair queuing

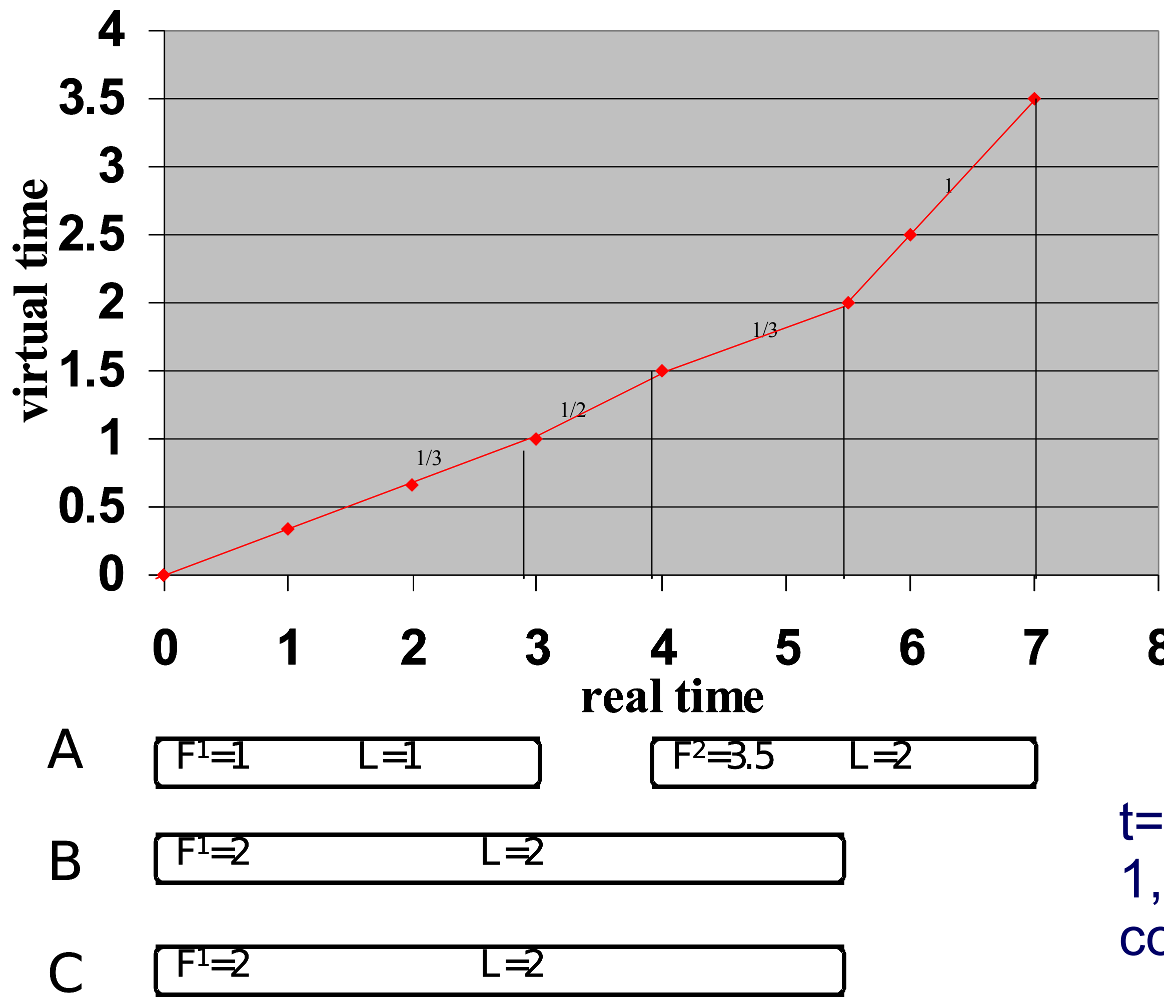
- WFQ: packet-by-packet GPS approximation
  - calculate "finish number" in fluid model
    - in each round, one bit of active flows served
  - schedule according to finish number
- Finish number
  - if backlogged: finish number of the last packet in the same flow + packet size
  - otherwise: current round number + packet size
- Performance bound

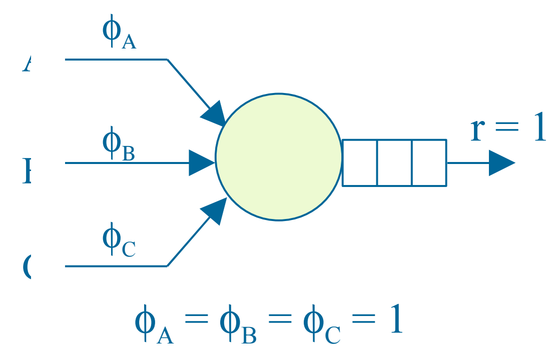
$$\frac{b}{r_i} + \frac{(K-1)L_i}{r_i} + \sum_{m=1}^{K} \frac{L_{\max}}{R_m}$$

WFQ: example

- Example
  - unit service rate
  - equal weight
  - t=0
    - packets of sizes 1, 2, 2 arrive at connection A, B, C
  - t=4
    - a packet of size 2 arrive at connection A

WFQ Example





t=0: Packets of sizes 1,2,2 arrive at connections A, B, C.

t=4: Packet of size 2 arrives at connection A

Student presentation

- Leo Gong: IntServ
  - [CSZ92] D. Clark and S. Shenker and L. Zhang, "Supporting Real-Time Applications in an Integrated Services Packet Network: Architecture and Mechanism". In Proceedings of SIGCOMM '92, Baltimore, Maryland, Aug, 1992, pp 14-26. [IntServ]

Discussion

- IntServ: pros and cons

This lecture

- Integrated services
  - commitments, interfaces, scheduling, admission
- Explore further
  - [DKS89] A. Demers, S. Keshav, and S. Shenker, "Analysis and Simulation of a Fair Queueing Algorithm". In Proceedings of ACM SIGCOMM'89, pp 3-12. [FQ]
  - [ZDESZ93] L. Zhang, S. Deering, D. Estrin, S. Shenker, and D. Zappala, "RSVP: A New Resource Reservation Protocol". IEEE Communications Magazine, 31(9):8-18, September 1993. [RSVP]

Next lecture

- Per-flow states?
  - [SSZ98] I. Stoica , S. Shenker , and H. Zhang , "Core -Stateless Fair Queueing: Achieving Approximately Fair Allocations in High Speed Networks" , Proc. ACM SIGCOMM , Vancouver, Canada, September 1998. [CSFQ]