Scheduling problem:
Differentiation between short and long TCP flows

Eeva Nyberg, Samuli Aalto
Networking Laboratory
Helsinki University of Technology

{eeva.nyberg, samuli.aalto}@hut.fi

Background

- Internet measurements show that
  - a small number of large TCP flows responsible for the largest amount of data transferred
  - most of the TCP flows made of few packets
- Intuition says that
  - favoring short flows reduces the total number of flows, and thus, by Little's law, also the mean "file transfer" time

Mathematical model

- Consider a bottleneck link loaded with elastic flows
  - such as file transfers using TCP
- Assume that
  - the flows arrive according to a Poisson process
  - each flow has a random service requirement (= file size) with a general distribution (typically heavy-tailed)
- So, at the flow level, we have a M/G/1 queueing system

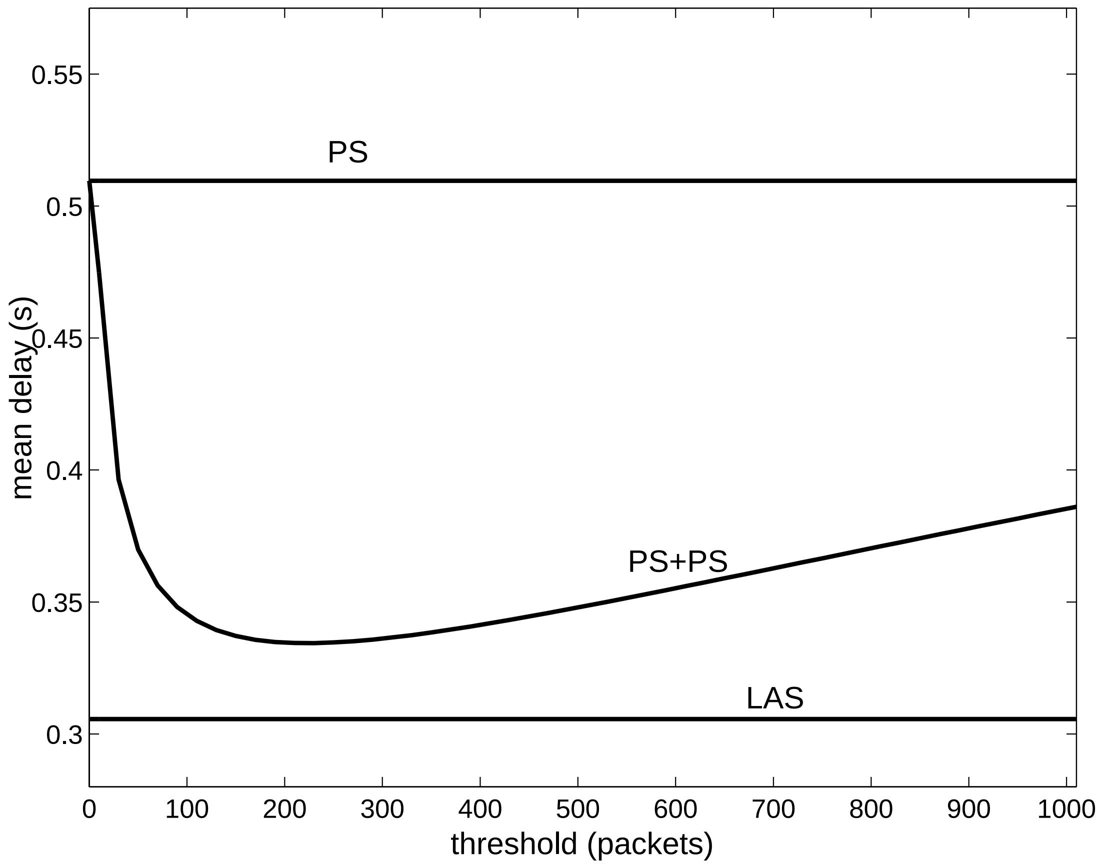
Scheduling disciplines

- PS = Processor Sharing
  - Without any specific scheduling policy, the flows are assumed to divide the bottleneck link capacity evenly
- SRPT = Shortest Remaining Processing Time
  - Choose a packet of the flow with least packets left
  - Hard to implement
- LAS = Least Attained Service
  - Choose a packet of the flow with least packets sent
  - Packet level implementation: RuN
- MLPS = Multilevel PS
  - Choose a packet of the flow with less packets sent than a given threshold
  - Proposed packet level implementation: RuN2C

Known optimality results for M/G/1

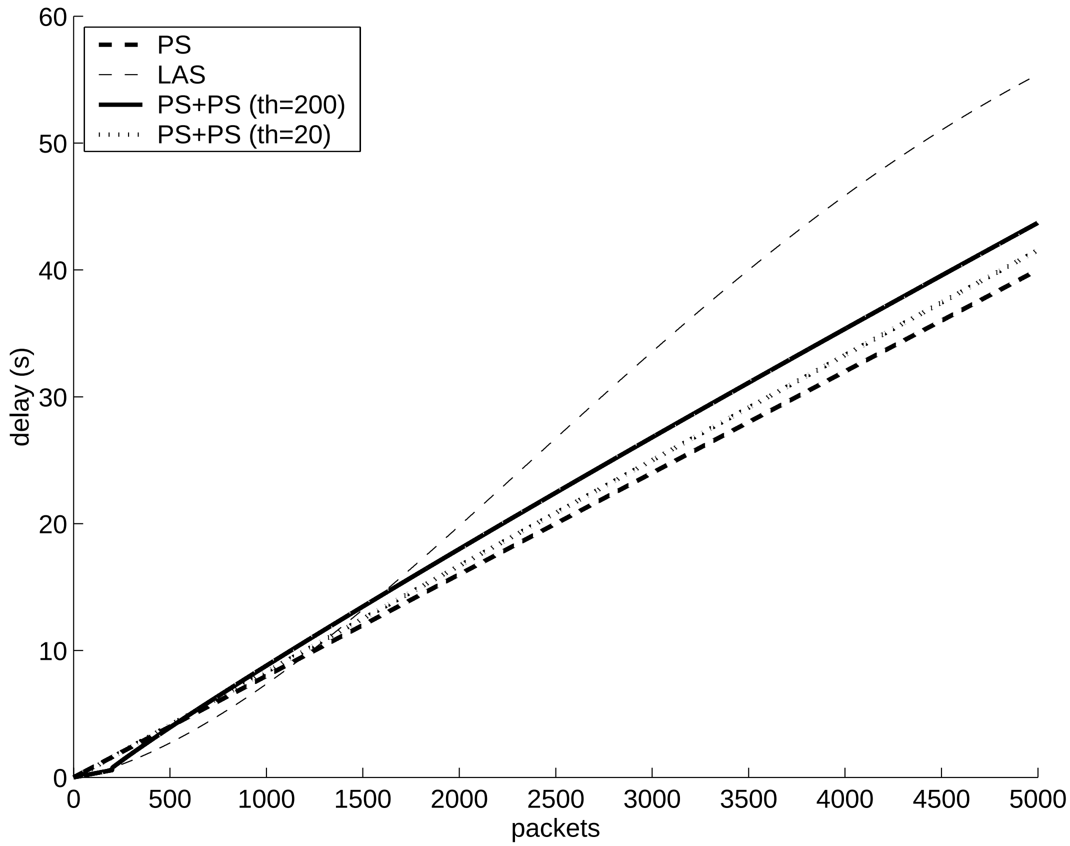
- If the **remaining service times** (= number of packets left) are **known** for each customer (= flow), then
  - Schrage (1968) proves that **SRPT is optimal**, i.e. it minimizes the mean delay (= file transfer time)
- If only the **attained service times** (= number of packets sent) are **known** for each customer (= flow), then
  - Sevcik (1974) conjectures that **SIPT is optimal**
  - Yashkov (1978) proves (?) that, if the **hazard rate** of the service time distribution is **decreasing**, then **LAS is optimal** among work-conserving scheduling disciplines
  - Feng and Misra (2003) proves (?) the same claim as above
  - Wierman et al. (2002) prove that if the **hazard rate** of the service time distribution is **decreasing**, **LAS is better than PS**

Mean delay



- Conclusion: MLPS seems to be better than PS in the mean delay sense (when hazard rate decreasing)

Asymptotic properties of the delay curve



- Conclusion: MLPS seems to be better than LAS in the asymptotic region (when hazard rate decreasing)

Open problem

- Prove that MLPS is better than PS in the mean delay sense (when hazard rate decreasing)
- Steps:
  - Easy to show: for any work-conserving disciplines S_1 and S_2

$$E[T^{S_1}] - E[T^{S_2}] = -\frac{1}{\lambda}\int_0^\infty (E[U_x^{S_1}] - E[U_x^{S_2}])h'(x)dx$$

    - T = delay
    - U_x = remaining truncated service time (min{S,x}) of those customers who have attained service at most x time units
  - Hard to show (but plausible): for any x

$$U_x^{\mathrm{LAS}} \le U_x^{\mathrm{MLPS}} \le U_x^{\mathrm{PS}}$$

THE END