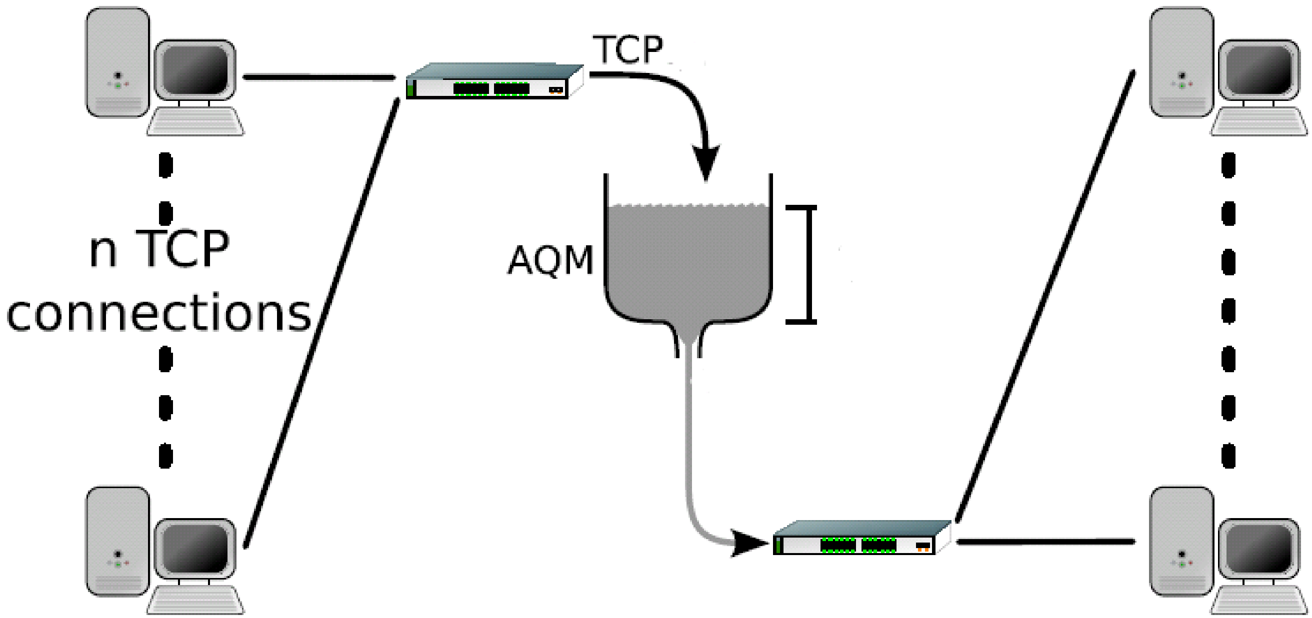
FLUID MODEL OF AN INTERNET ROUTER UNDER THE MIMD CONTROL SCHEME

Urtzi Ayesta (LAAS-CNRS)
Alexei Piunovskiy (University of Liverpool)
Y. Zhang (University of Liverpool)

INFORMS TELECOM
28 march, 2008

– *n* TCP connections sharing a bottleneck

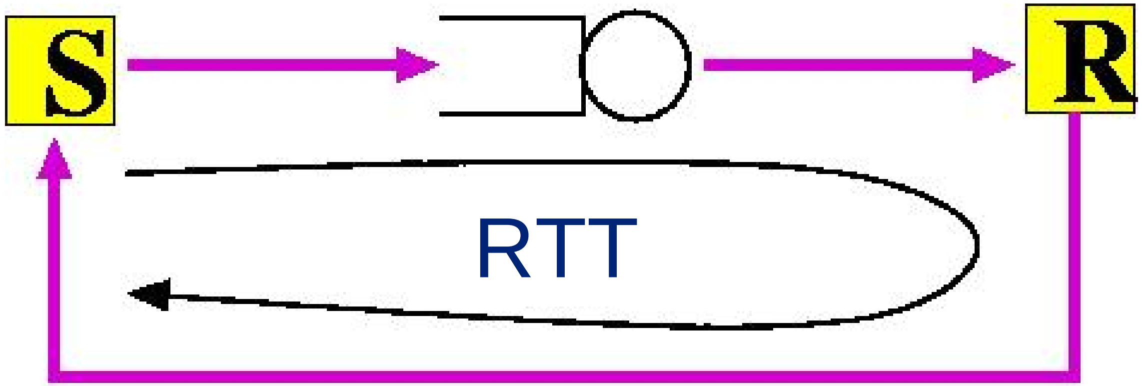


– TCP:

- Distributed Congestion Control algorithm

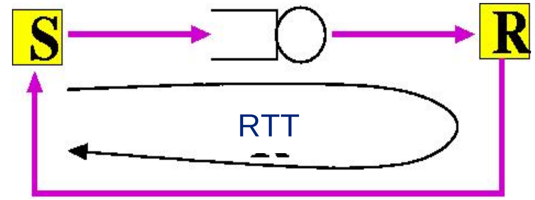
- The TCP sender detects congestion on the network through the feedback received from the TCP receiver

- The packet sending rate (congestion window) is adapted dynamically
  - Increase in absence of packet losses
  - Decrease upon reception of congestion notification

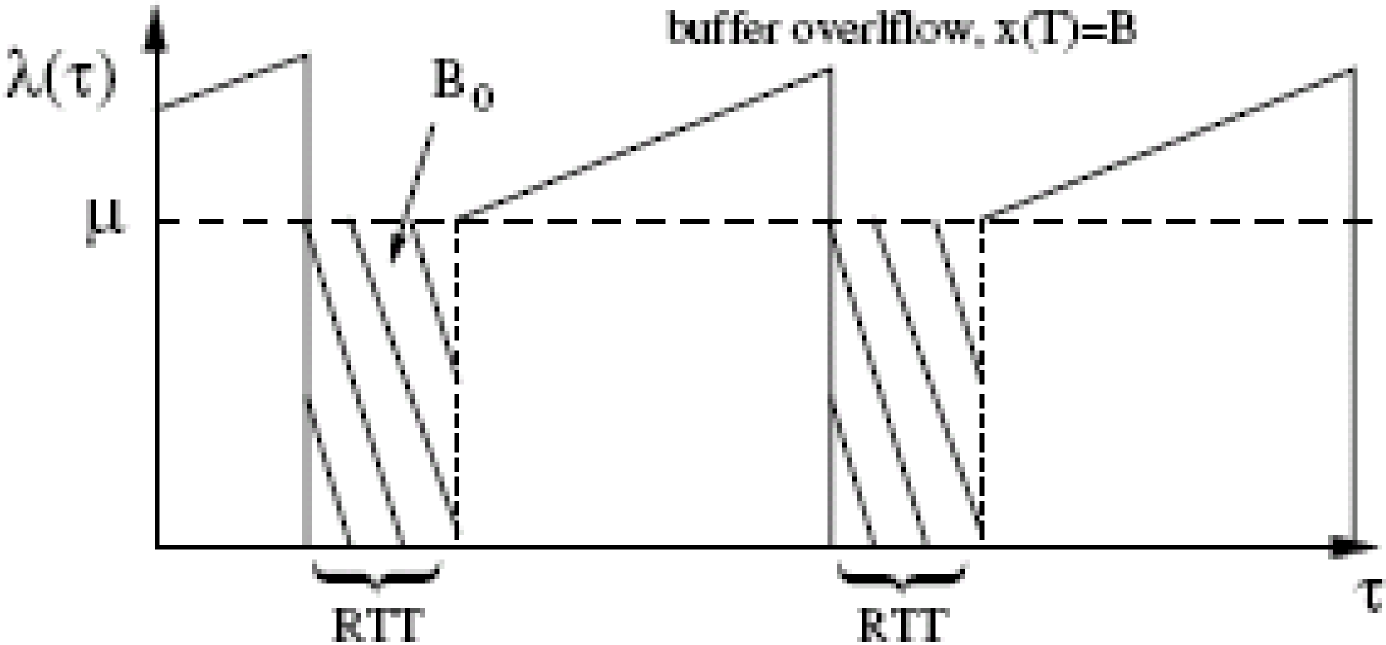


- Current TCP Implementations:
  - AIMD paradigm : The increase is linear in time
  - Problem in fully utilizing the link capacity in the presence of High-Speed links
- The MIMD TCP is one of the possible alternatives
  - MIMD: Exponential Increase in time
  - After a loss, the elapsed time until the congestion window reaches the size before the loss is constant.
  - Scalable TCP is a particular case

Buffer size: BDP rule [VS94]



– One TCP AIMD connection. The minimum buffer size in order to get 100% utilization is B_0=RTT x μ.



– Interes in recent years: [AKMcK04] [BHLO03] [DJD05] [W06] [VS07]

Performance criterion

- Maximize throughput: $\bar{\lambda} = \lim_{t\to\infty} \frac{1}{t} \int_0^t \lambda(s) ds.$

- Minimize queueing delay: $\bar{x} = \lim_{t\to\infty} \frac{1}{t} \int_0^t x(s) ds.$

- Contradictory goals: Example of multicriteria optimization. We construct the Pareto Set of the unconstrained optimization

$$\max \left\{ \lim_{t\to\infty} \frac{1}{t} \int_0^t c_1 x(s) + c_2 \lambda(s) ds. \right\}$$

- We deal with the optimal impulse control problem of a deterministic system with long-run average optimality criterion.
  - Also, our model can be viewed as a hybrid system.

- In principle, the control policy in our model can depend on $x(t)$ and $\lambda(t)$.

- In practice, however, all currently implemented buffer management schemes send congestion signals based only on the state of the buffer. Thus, we also limit ourselves to this case.

- Furthermore, we assume that one can only control B, the size of the buffer.

- Objective: Determine the buffer size that solves the optimization problem

A fluid model

- Let $\lambda_i(t)$ denote the sending rate of connection *i* at time *t*

$$\dot{\lambda}_i(t) = r_i \lambda_i(t)$$

- Assuming the same RTT for all connections

$$\dot{\lambda}(t) = \sum_i \dot{\lambda}_i(t) = r\lambda(t)$$

- Let $x(t)$ denote the queue length at time *t.* Then

$$\dot{x}(t) = \begin{cases} \lambda(t) - \mu, & \text{if } x(t) > 0, \text{or if } x(t) = 0 \text{ and } \lambda(t) > \mu, \\ 0, & \text{if } x(t) = 0 \text{ and } \lambda(t) \leq \mu. \end{cases}$$

A Fluid model (cont.)

- When $x(t) = B$ a packet is dropped, and the sending rate of a connection is reduced instantaneously

$$\lambda_i(t+0) = \beta_0 \lambda_i(t-0) = \beta_0 \lambda_i(t), \quad \beta_0 < 1$$

- Assuming that the sending rates are uniformly distributed in a congestion epoch:

$$\lambda(t+0) \approx \beta \lambda(t-0), \quad \text{where} \quad \beta = (\beta_0 - 1)\frac{\tilde{n}}{n} + 1$$

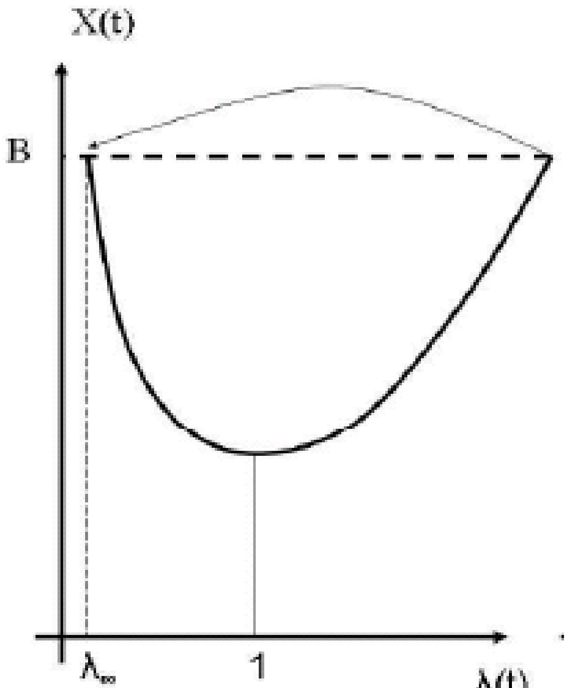
- Hybrid model : Dynamical systems that combine both discrete and continuous behavior.

- Example: Scalable TCP, $r_i = 0.01/RTT$ and $\beta_0 = 0.875$ [Kelly03]

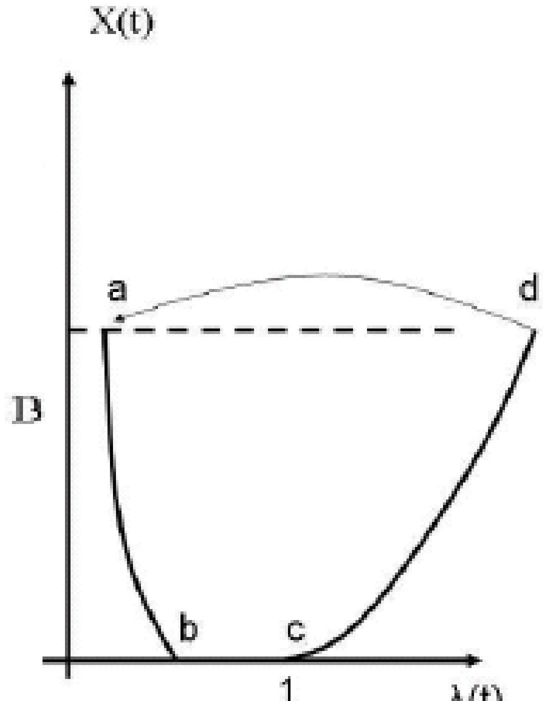
– **Theorem:** Let $B_0 = \frac{\mu}{r}\left(\ln\frac{\beta-1}{\beta\ln\beta} - \frac{\beta\ln\beta}{1-\beta} - 1\right).$

Then unclipped (clipped) cycle exists if and only if $B \geq B_0, (B \leq B_0).$

The duration of the cycle equals $T = -\ln\beta, \quad \forall B \geq 0.$



Unclipped cycle.



Clipped cycle.

– **Lemma:** In the desynchronized case, $\beta = 1 - \frac{1}{2n}, \quad \lim_{n\to\infty} n^2 B_0 = \frac{1}{32}$

- meaning that B_0 decreases to zero as. $\frac{1}{2n^2}$,

– **Theorem:** Let λ_k be the value of $\lambda(t)$ immediately after the *k*-th jump. Then, starting from any initial value λ_0, the limit $\lim_{k\to\infty}\lambda_k = \lambda_\infty$ exists.

- In addition there exist only simple limiting cycles, i.e. instant series of more than one jump are never realized, and all the values of $\lambda(t)$ immediately after a jump coincide with λ_∞ for the trajectory starting from $x_0 = B, \quad \lambda_0 = \lambda_\infty$.The limiting cycle is stable.

– Theorem: If $B \geq B_0$ then

$$\bar{\lambda} = \mu; \quad \bar{x} = \frac{\mu}{r}\left(B + \frac{\ln\beta}{2} + \frac{\beta\ln\beta}{1-\beta} + 1\right),$$

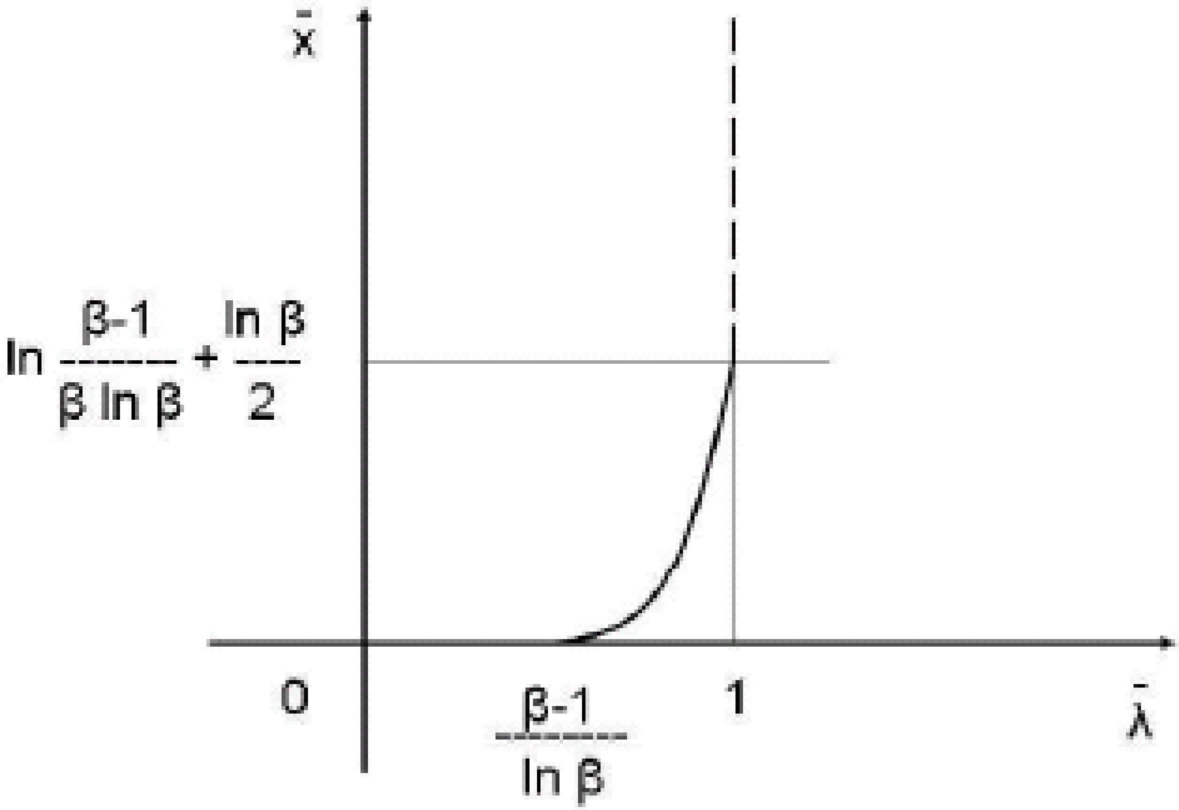
if $B < B_0$ then

$$\bar{\lambda} = \mu\frac{e^{\theta}(\beta-1)}{\ln\beta}; \quad \bar{x} = -\frac{\mu}{r}\frac{\delta e^{\theta}(1-\beta)-\frac{1}{2}(\delta+\theta)^2}{\ln\beta},$$

Pareto Set or Trade-off curve

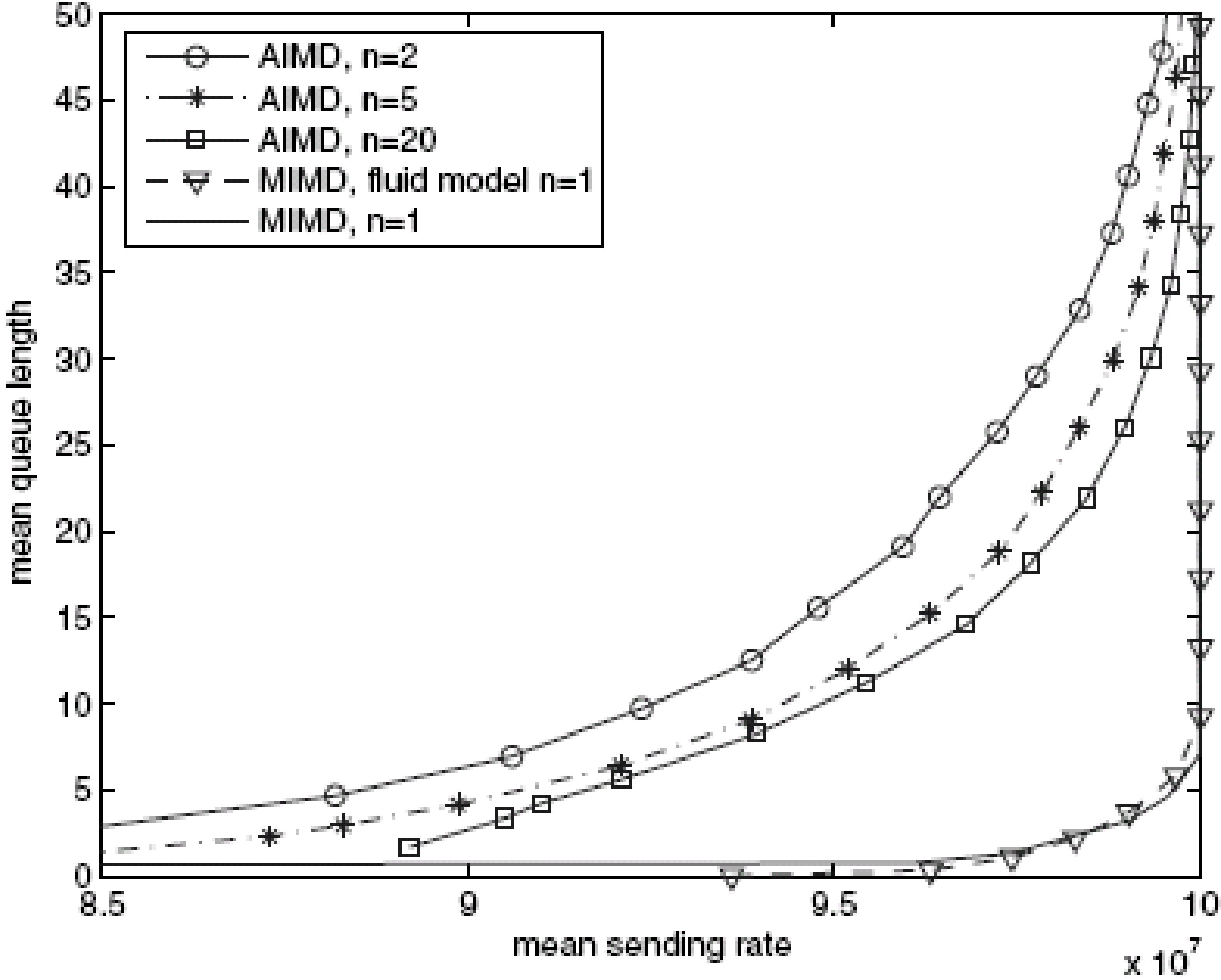
– Corollary: If $B \to 0$, then $\bar{x} \to 0$; $\bar{\lambda} \to \frac{\beta - 1}{\ln \beta}$.

$$\text{If} \quad B \to B_0, \quad \text{then} \quad \bar{x} \to \ln \frac{\beta-1}{\beta \ln \beta} + \frac{\ln \beta}{2}; \qquad \bar{\lambda} \to 1$$



For every (c_1, c_2), we can determine the value of the Buffer that maximizes

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t c_1 x(s) + c_2 \lambda(s) ds.$$



AIMD, n=2
AIMD, n=5
AIMD, n=20
MIMD, fluid model n=1
MIMD, n=1
mean queue length
mean sending rate
x 10^7
0
5
10
15
20
25
30
35
40
45
50
8.5
9
9.5
10

Conclusions

- MIMD : The cycle duration independent of the link capacity (for AIMD it does depend)

- As the number of connections increases, B_0 decreases with $1/n^2$.

- For Scalable TCP and one connection B_0=0,22 x RTT x μ.

- For MIMD, B_0 depends linearly on μ. For AIMD depends on μ^2. MIMD not only scales well with capacity, but it also requires smaller buffer sizes in high-speed networks.

Conclusions (cont.)

- When $B \to 0$, then $\bar{\lambda} \to \mu(\beta - 1)/\ln(\beta)$. Then if $\beta = \beta_0 = 0.875$ and n=1, $\lambda = 0.94\mu$. That is, even with no buffer, the utilization of Scalable TCP would be very high.

- Incorporate to the model : delays, heterogeneous RTT, dynamically arriving and departing connections etc.

- The Pareto Set of the multicriteria optimization problem may be useful in dimensioning routers