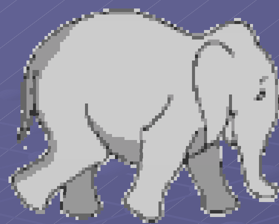
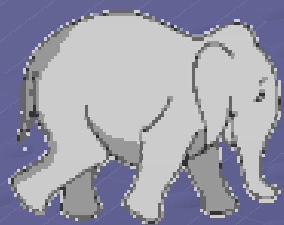
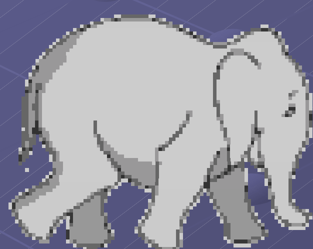
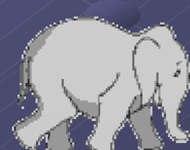





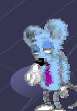
The War Between Mice & Elephants

by,
Matt Hartling
&
Sumit Kumbhar

ACN Presentation
Feb. 5, 2002







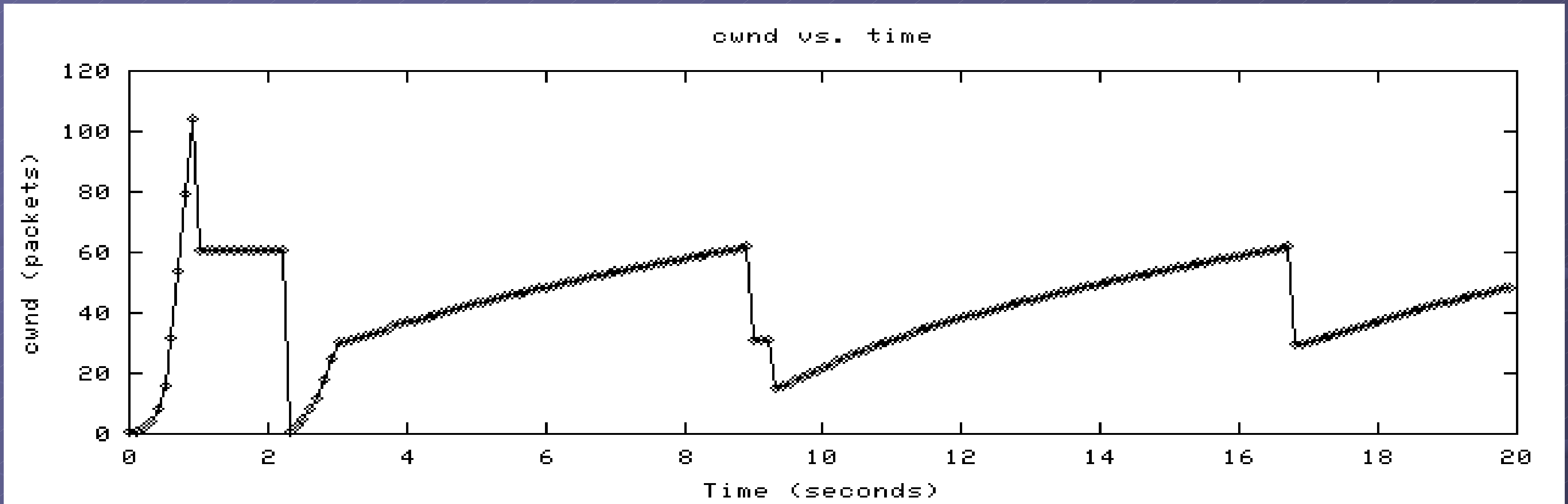
Paper Info…

- ICNP 2001
- Guo, Liang (BU; Graduate Student)
- Matta, Ibrahim (BU; Professor)
  - Matta, Ibrahim; Guo, Liang. *Differentiated Predictive Fair Service for TCP Flows*. In *Proceedings of ICNP'2000: The 8th IEEE International Conference on Network Protocols*, Osaka, Japan, October 2000.
  - Yilmaz, Selma; Matta, Ibrahim. *On Class-based Isolation of UDP, Short-lived and Long-lived TCP Flows*. In *Proceedings of the International Workshop on Modeling, Analysis and Simulation of Computer and Telecommunications Systems - MASCOTS '01*, Cincinnati, Ohio, August 2001.

Outline

- Introduction
- Analyzing Short TCP Flow Performance
- RIO-PS: Architecture and Mechanisms
- Simulations
- Discussion
- Conclusions and Future Work

Short vs. Long Lived Flows



Long-Lived (Elephants)

- Transmit large number of packets
- Operate primarily in TCP congestion avoidance
- TCP mechanisms designed for long-lived flows

Example:

- FTP

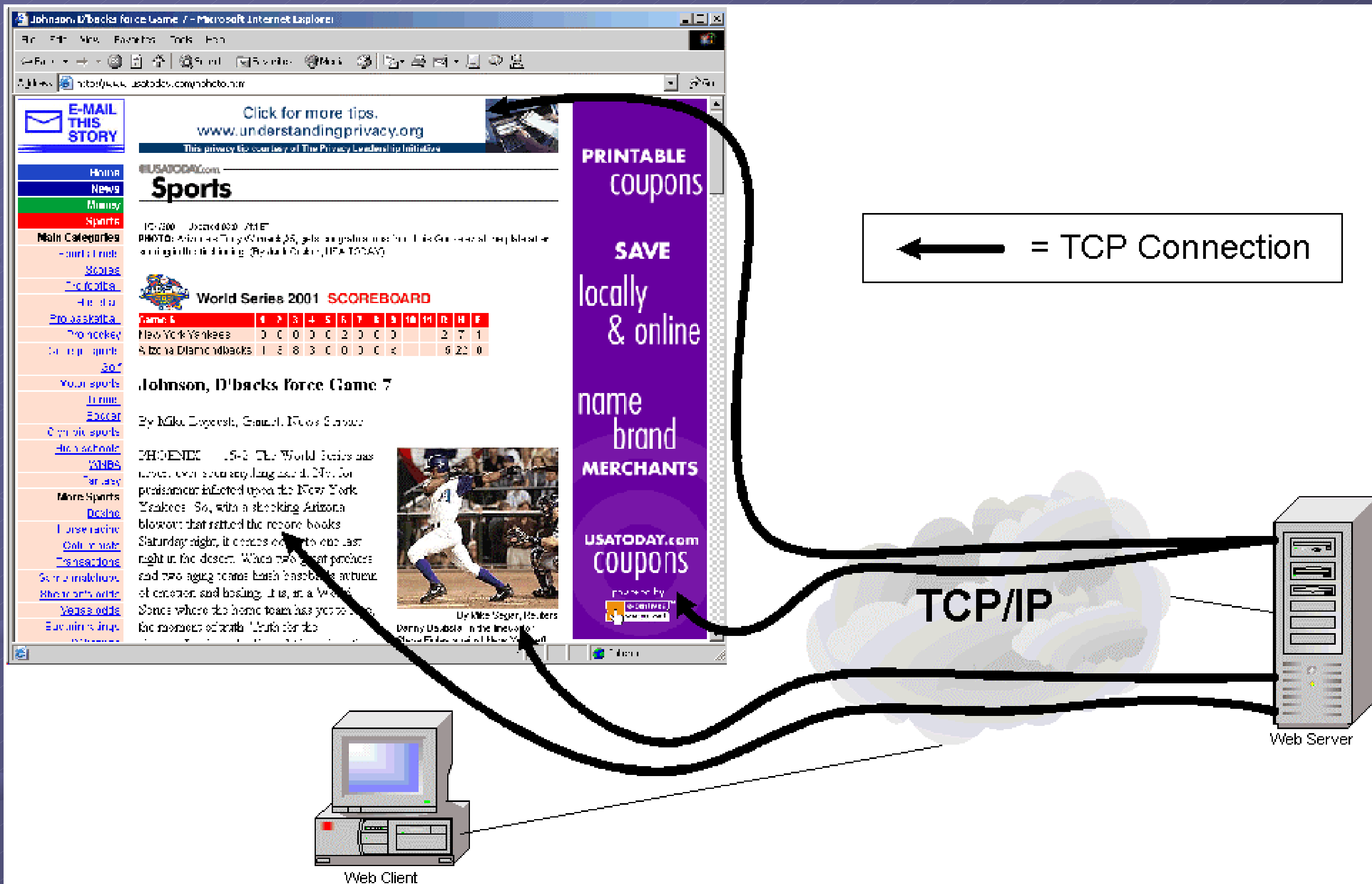
Short-Lived (Mice)

- Transmit small number of packets
- Operate Primarily in slow-start phase

Example:

- HTTP/1.0 transfers

HTTP/1.0

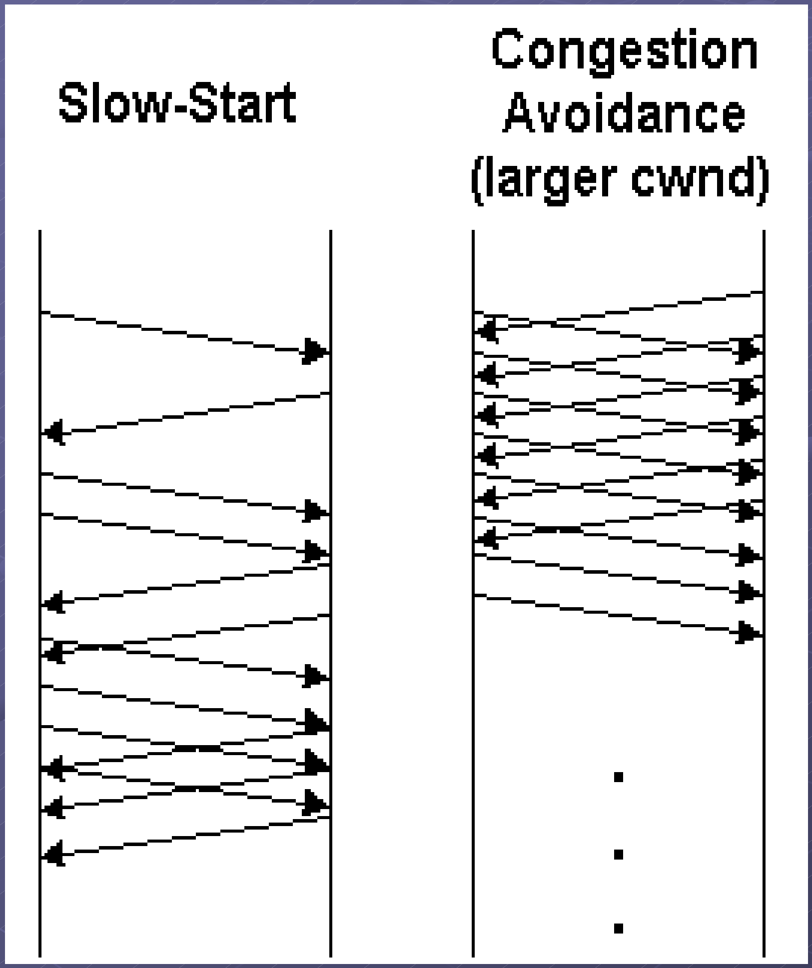


= TCP Connection
TCP/IP
Web Server
Web Client
Sports
World Series 2001 SCOREBOARD
Johnson, D'backs force Game 7
PRINTABLE coupons
SAVE locally & online
name brand MERCHANTS
USATODAY.com coupons

Web Traffic (HTTP/1.0)

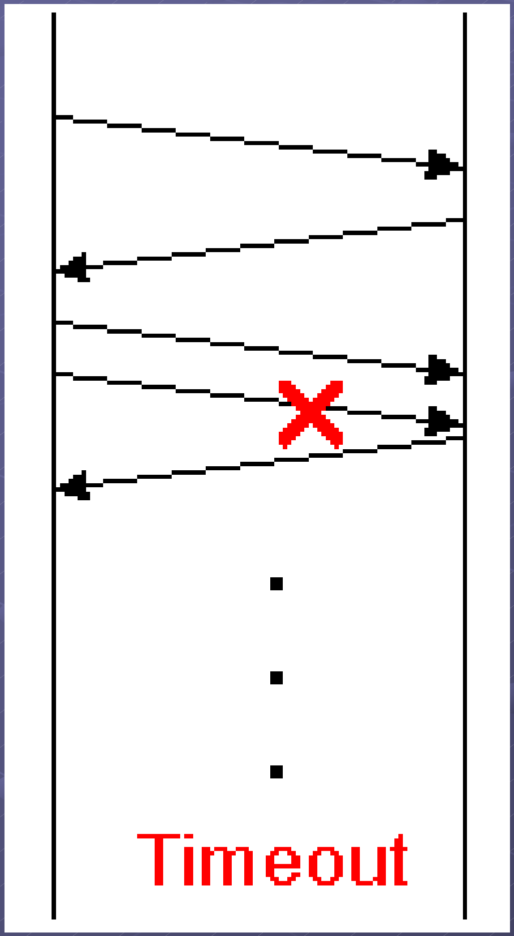
- Bruce Mah's Statistics (HTTP/1.0; 1995)
  - HTTP Reply Lengths:
    - Max = 1 MB
    - Mean = 8-10 KB
    - Median = 2 KB (50% send only 1 or 2 packets!!)
- This paper uses BU Traces from 1995 (old data!)
- HTTP/1.0 reply lengths modeled using Pareto distribution.

Short-Lived Flow Issues (1): Slowstart



- Transmission rate increases slowly
- Added delay
- Bad for flows that don’t get out of the early stages of slowstart.

Short-Lived Flow Issues (2): Small cwnd



- Fast Retransmit needs three duplicate acks
- If one of the first three packets are dropped, a TO occurs
- TO's are BAD!! (Web traffic is somewhat more sensitive to delays.)

Short-Lived Flow Issues (3): Initial Timeout Value

- Initial timeout value (ITO) is set to 3 seconds
- If the SYN, SYN-ACK, or first data packet is dropped, the flow must wait for the 3 second TO before a retransmission

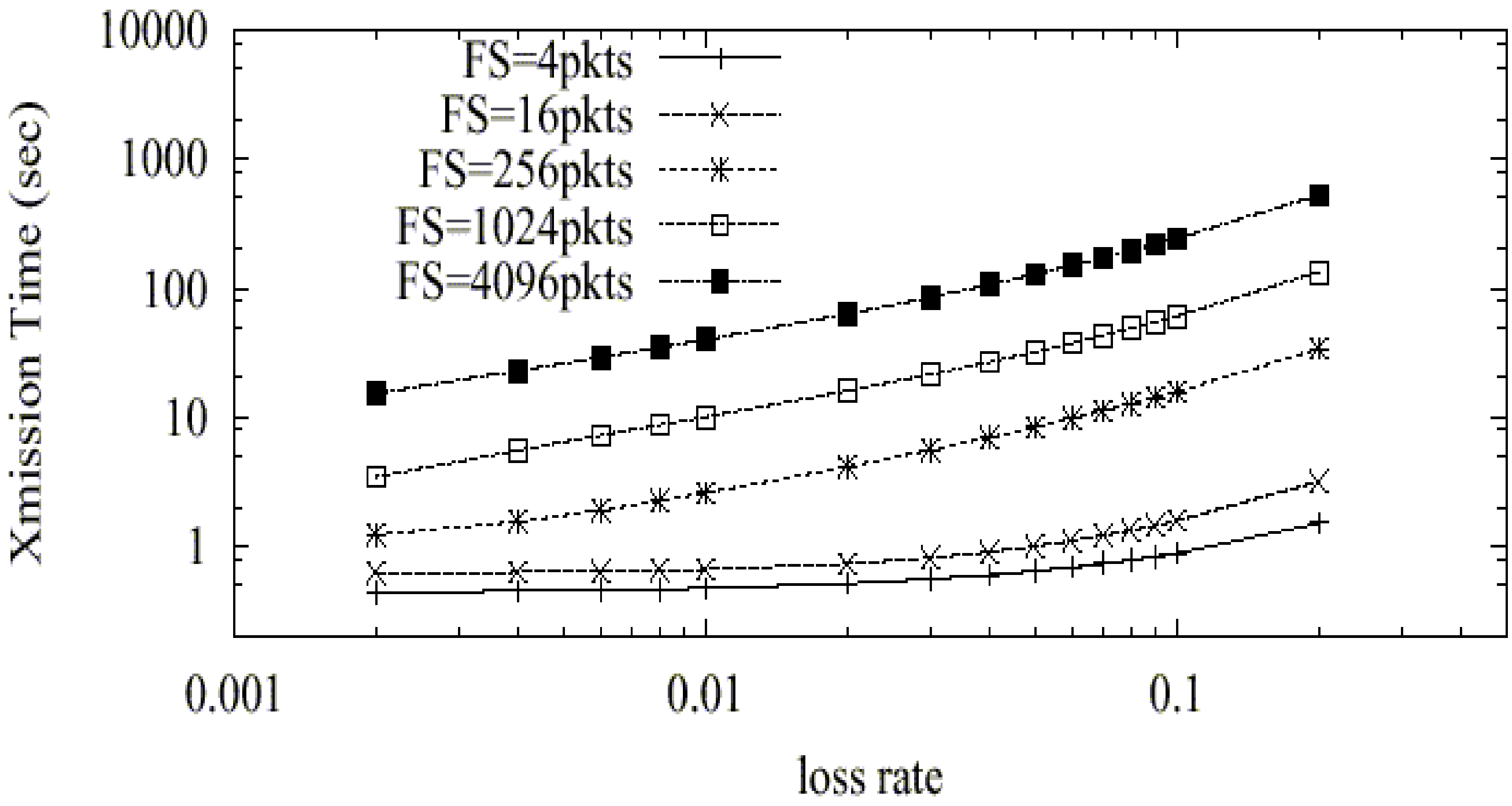
Related Work

- Class Based Isolation of short, long, and UDP [Yilmaz, Selma; Matta, Ibrahim 2001]
- TCP Protocol Solutions:
  - Reduce ITO
  - Larger Initial CWND
- "Size-aware job scheduling" [Crovella et al. 1999 and Bansal et al. 2001]

Outline

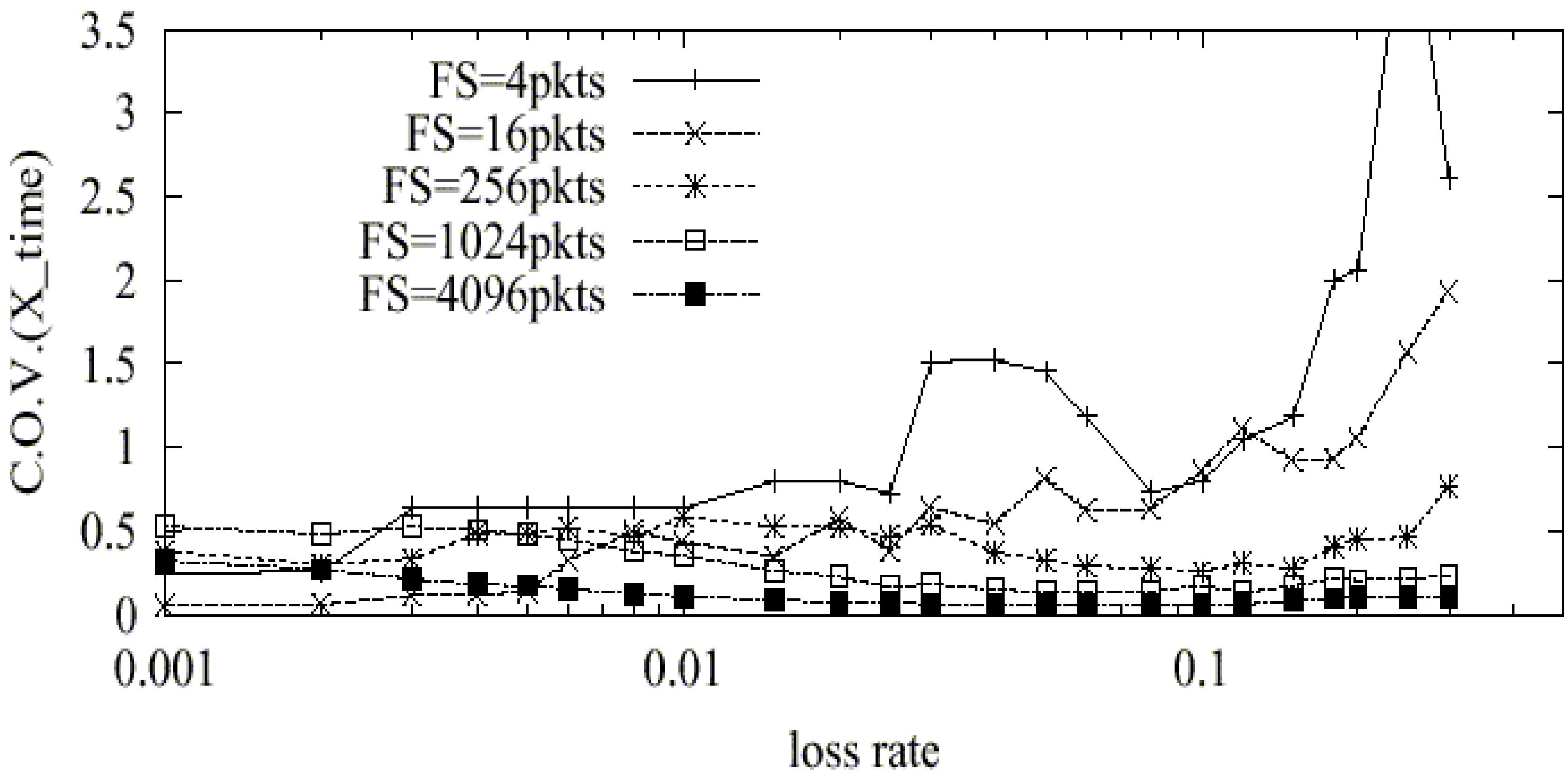
- Introduction
- <u>Analyzing Short TCP Flow Performance</u>
- RIO-PS: Architecture and Mechanisms
- Simulations
- Discussion
- Conclusions and Future Work

Average Transmission Time



(a) Average Transmission Time

Transmission Time Variance



(b) Coefficient of Variation

Comparison of Drop Tail, RED, RIO-PS

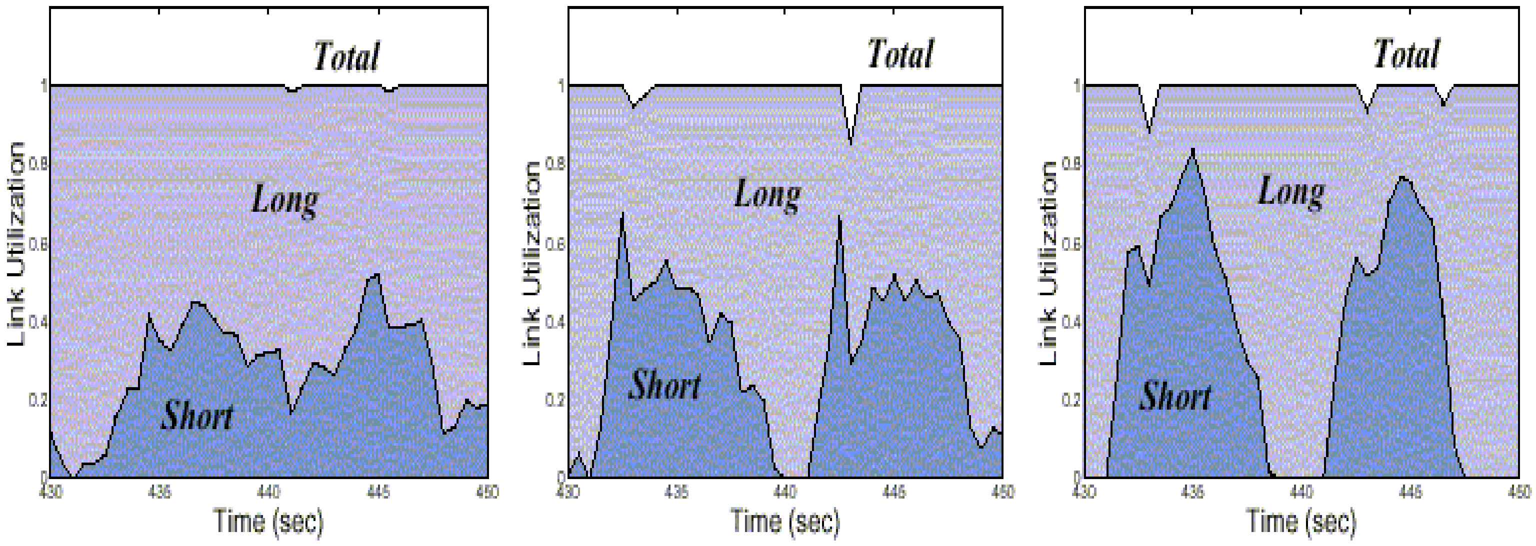


Fig. 2. Impact of Preferential Treatment— Link utilization under Drop Tail (left), RED (middle), and RIO-PS (right)

Goodput

Link B/W	Flows	DropTail	RED	RIO-PS
1.25Mbps	All	153479	154269	154486
	Short	40973	49897	49945
	Long	112506	104372	104541
1.5Mbps	All	185650	184315	183154
	Short	43854	49990	49990
	Long	141796	134325	133164

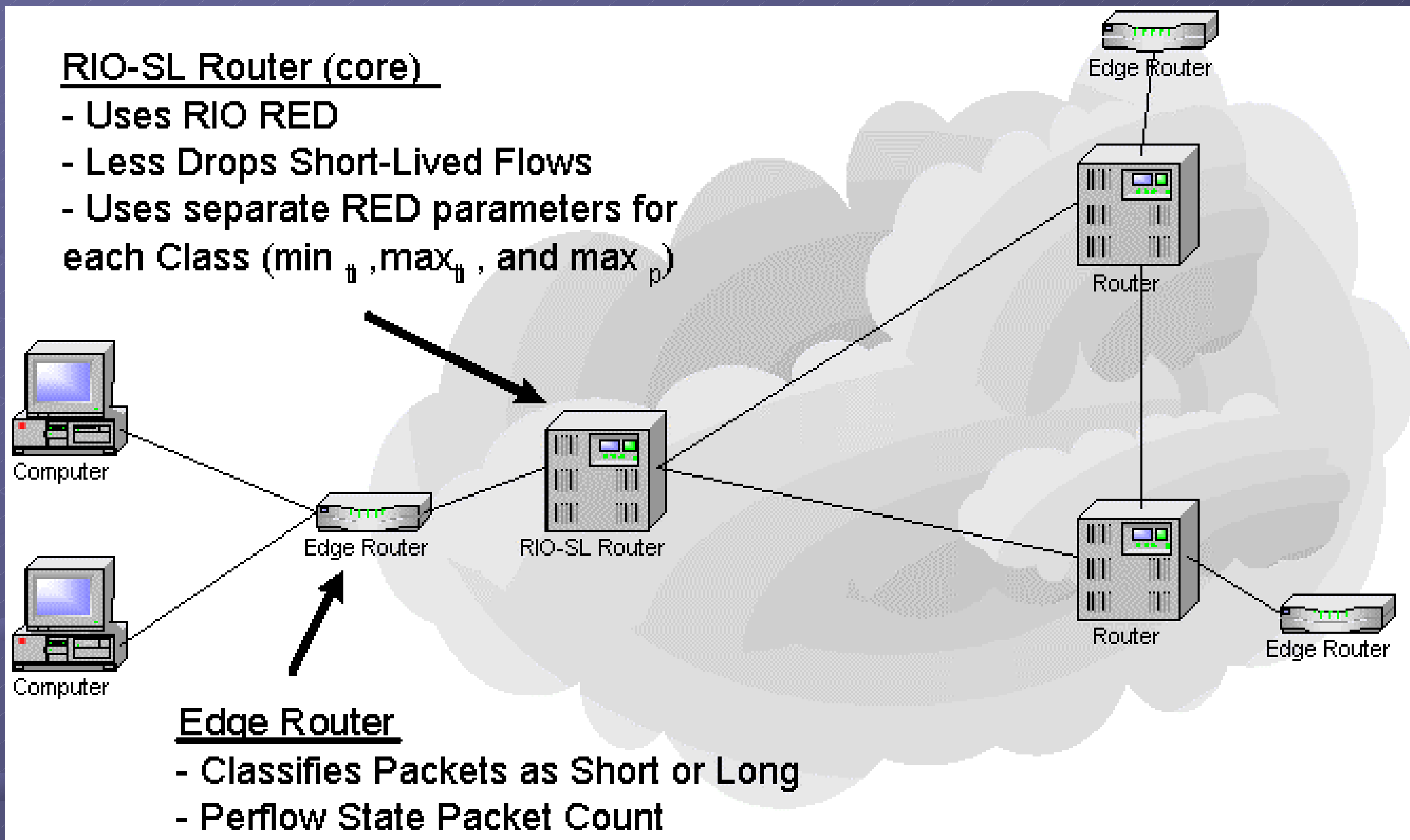
TABLE I

NETWORK GOODPUT UNDER DIFFERENT SCHEMES

Outline

- Introduction
- Analyzing Short TCP Flow Performance
- <u>RIO-PS: Architecture and Mechanisms</u>
- Simulations
- Discussion
- Conclusions and Future Work

RIO-PS


RIO-SL Router (core)
- Uses RIO RED
- Less Drops Short-Lived Flows
- Uses separate RED parameters for
each Class (min th ,max th , and max p)
Edge Router
Router
Computer
Edge Router
RIO-SL Router
Computer
Router
Edge Router
Edge Router
- Classifies Packets as Short or Long
- Perflow State Packet Count

Edge Router Functions

- Maintains per flow packet counts
- Labels the packet as “short” or “long” based on the following parameters:
  - L_t - threshold when exceeded packets are labeled as long (Dynamic or Static)
  - T_u – idle timer; if no packets are received within time T_u seconds, the flow is removed from the system

Edge Router Functions (cont)

- Dynamic L_t parameters:
  - *SLR* – target ratio # *short flows* / # *long flows*.
  - T_c – time between making additive adjustments to L_t to achieve SLR

- All Long flows begin as Short

Core Router Functions

- Uses RIO Mechanism
- RIO (David Clark – MIT and Wenjia Fang – Princeton University)
  - Sender or receiver classifies the packet as “in” or “out” of profile

    (Profile example: transmission rate of 64K)
  - Basic idea – drop out of profile packets more aggressively than in profile.
  - Maintains 2 sets of RED parameters (minth, maxth, maxp, wq, etc.)

Core Router Functions (cont)

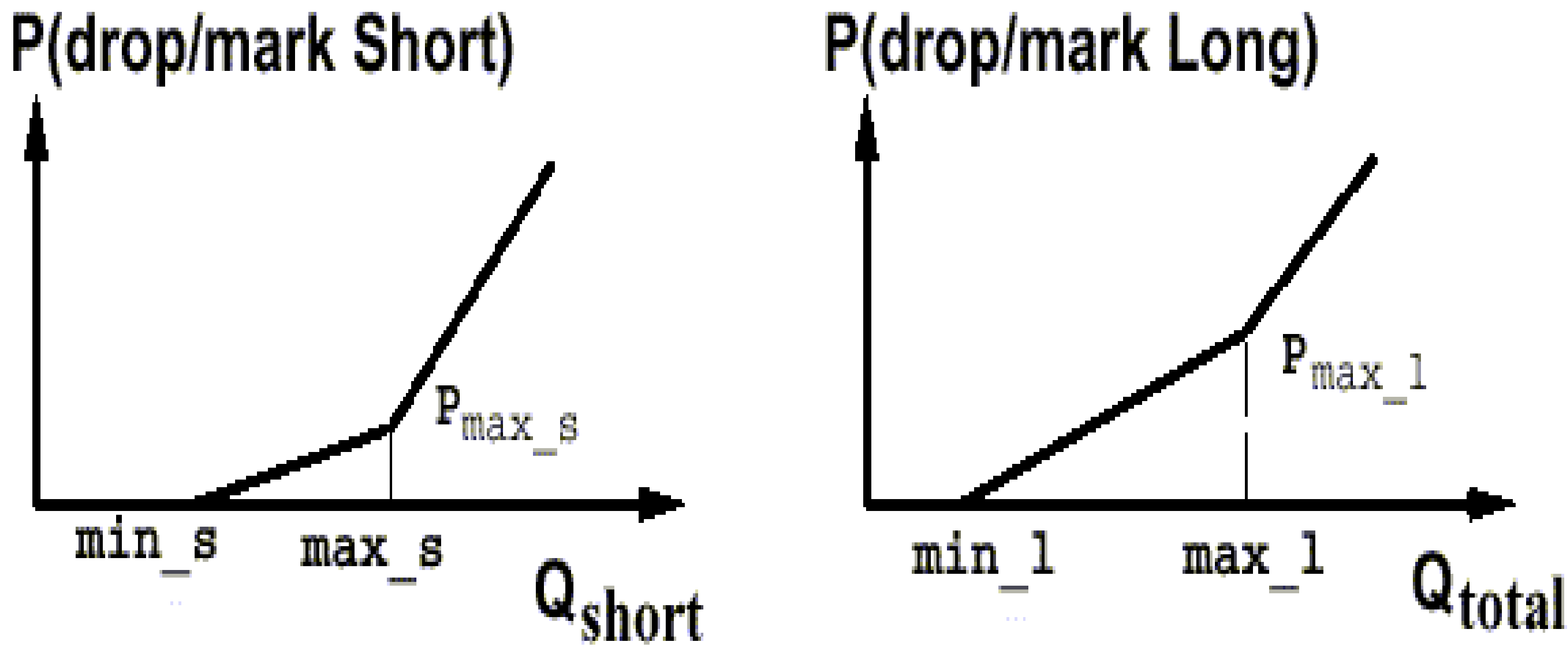
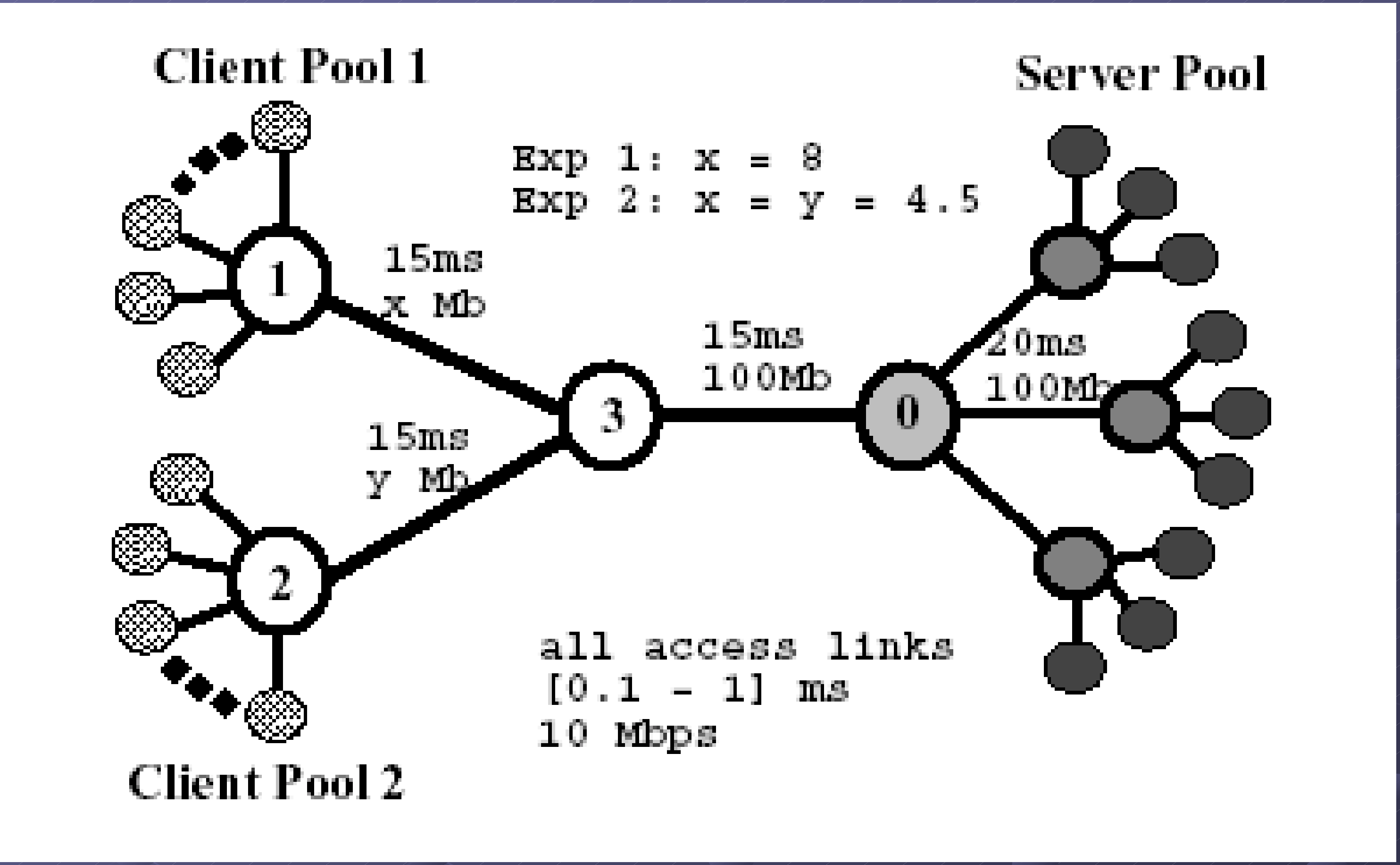


Fig. 4. RIO queue with Preferential treatment to Short flows

Outline

- Introduction
- Analyzing Short TCP Flow Performance
- RIO-PS: Architecture and Mechanisms
- <u>Simulations</u>
- Discussion
- Conclusions and Future Work

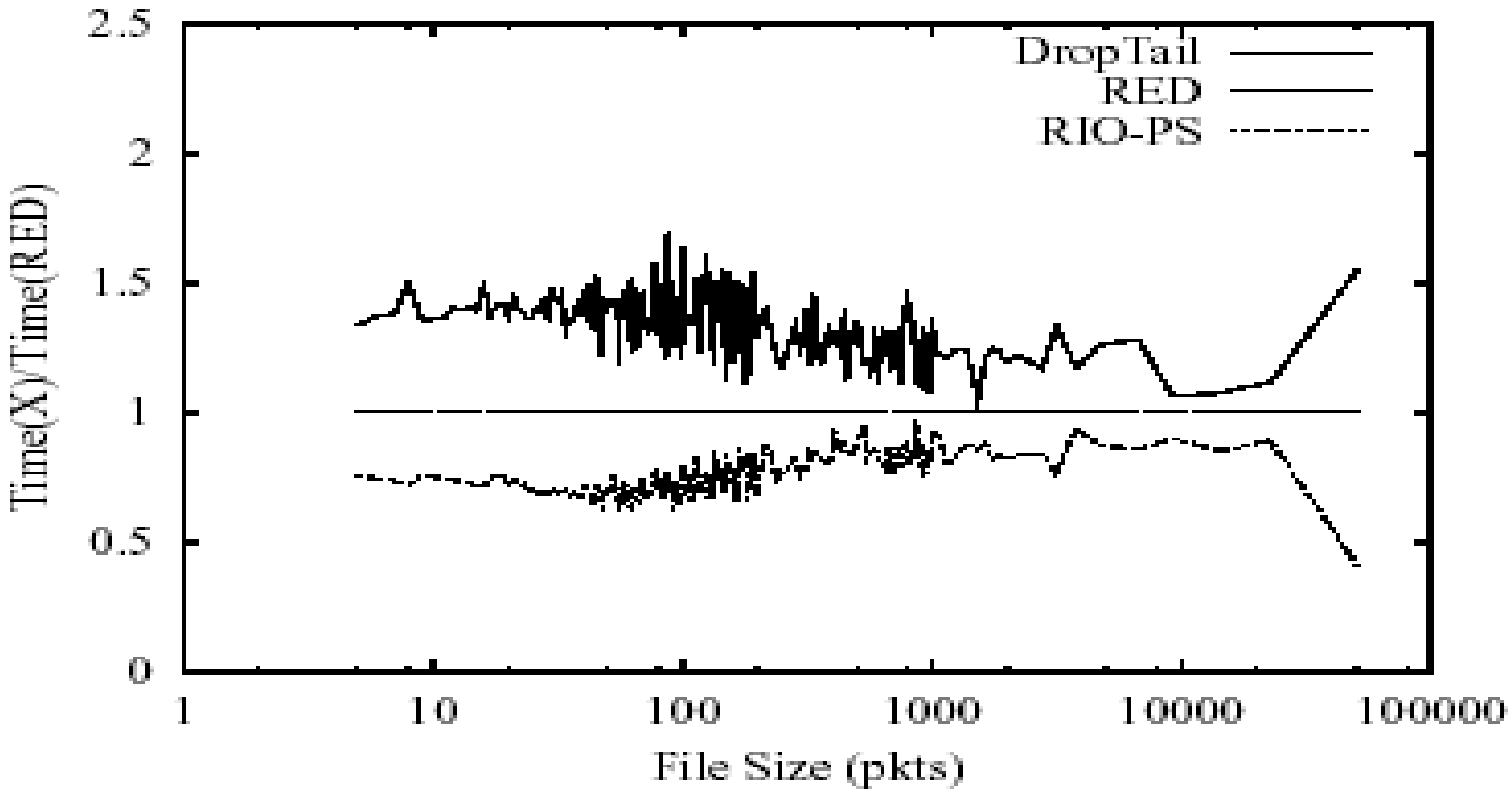


RIO with preference to short flows

Description	**Value**
Packet Size	500 bytes
Maximum Window	128 packets
TCP version	Newreno
TCP timeout Granularity	0.1 seconds
Initial Retransmission Timer	3.0 seconds
B/W delay product (BDP)	≈ 200 pkts (Exp1) ≈ 120 pkts (Exp2)
Bottleneck Buffer Size (B)	DropTail: $1.5\times$ BDP RED/RIO-PS: $2.5\times$BDP
Q. Parameters	$(min_{th}, max_{th}, P_{max}, w_q)$
RED	(0.15B, 0.5B, 1/10, 1/512)
RIO-PS short	(0.15B, 0.35B, 1/20, 1/512)
RIO-PS long	(0.15B, 0.5B, 1/10, 1/512)
RED & RIO-PS	ecn_ on, wait_ on, gentle_ on
Edge Router	$SLR = 3, T_u = 1\ sec, T_c = 10\ sec$
Foreground Traffic	
(Src, Dest)	(Server Pool, Client Pool)
Long Connection Size	1000 packets
Short Connection Size	10 packets

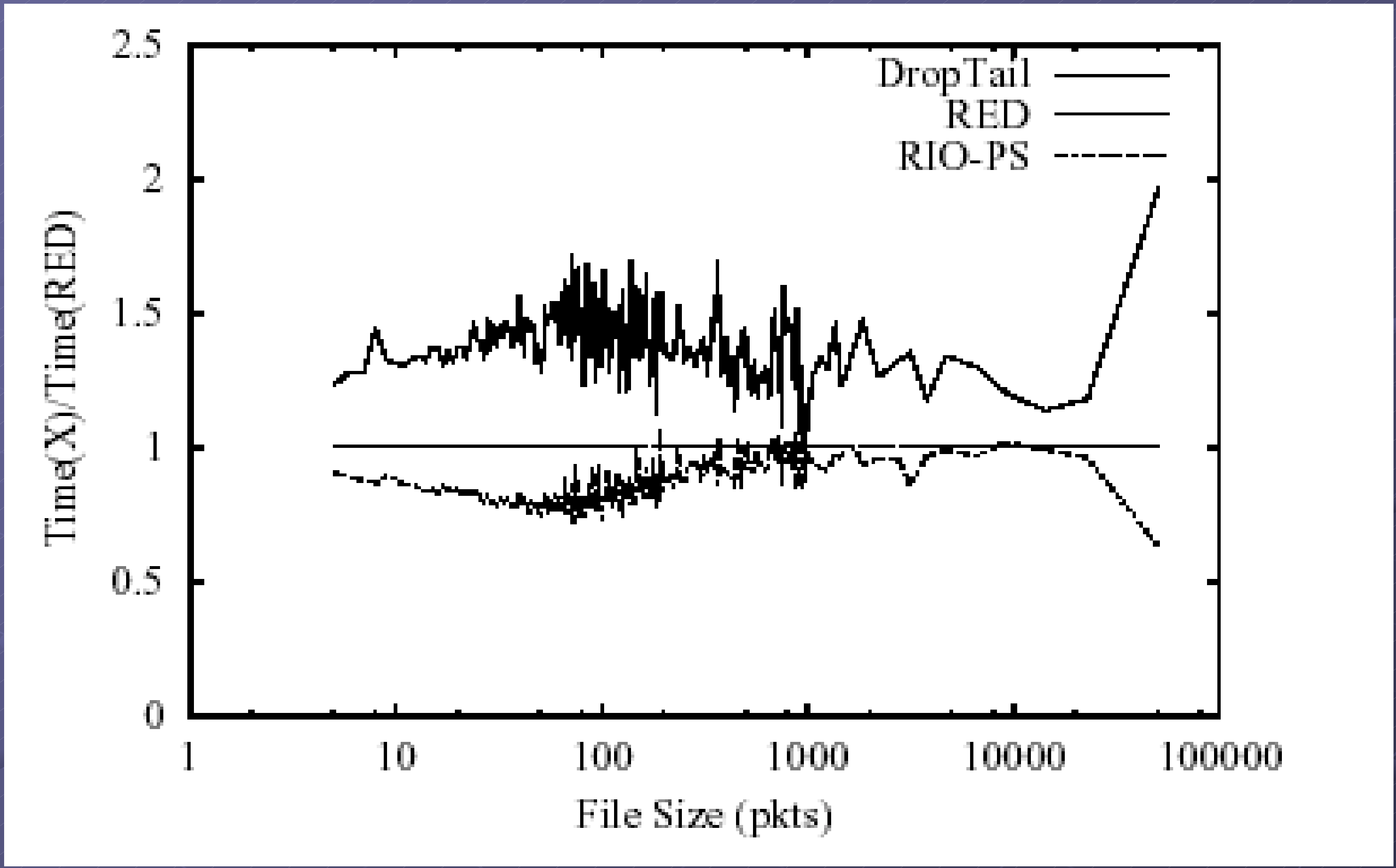
TABLE III

NETWORK CONFIGURATION



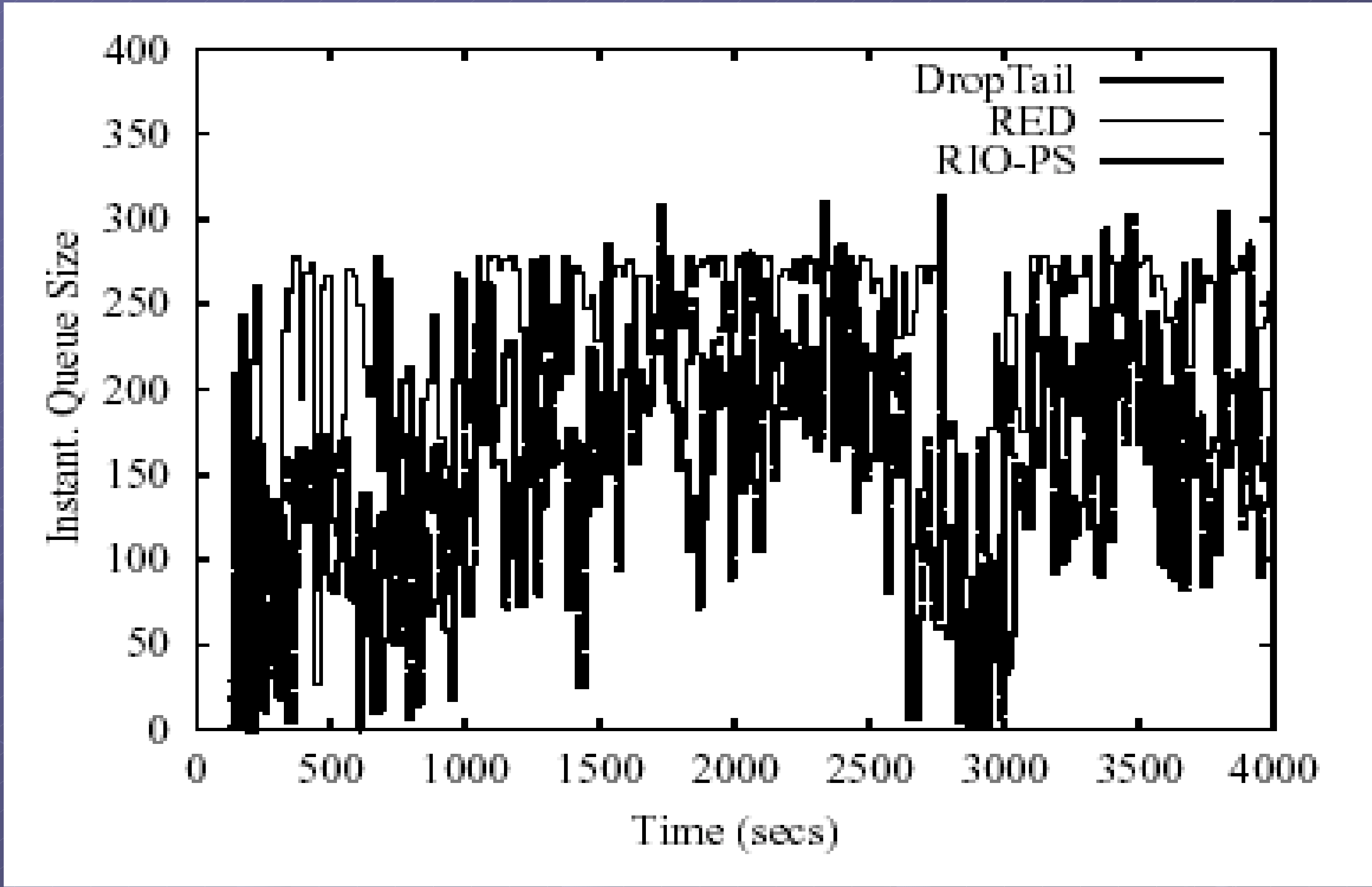
Average response time for different sized objects

Initial retransmission timer = 3 seconds

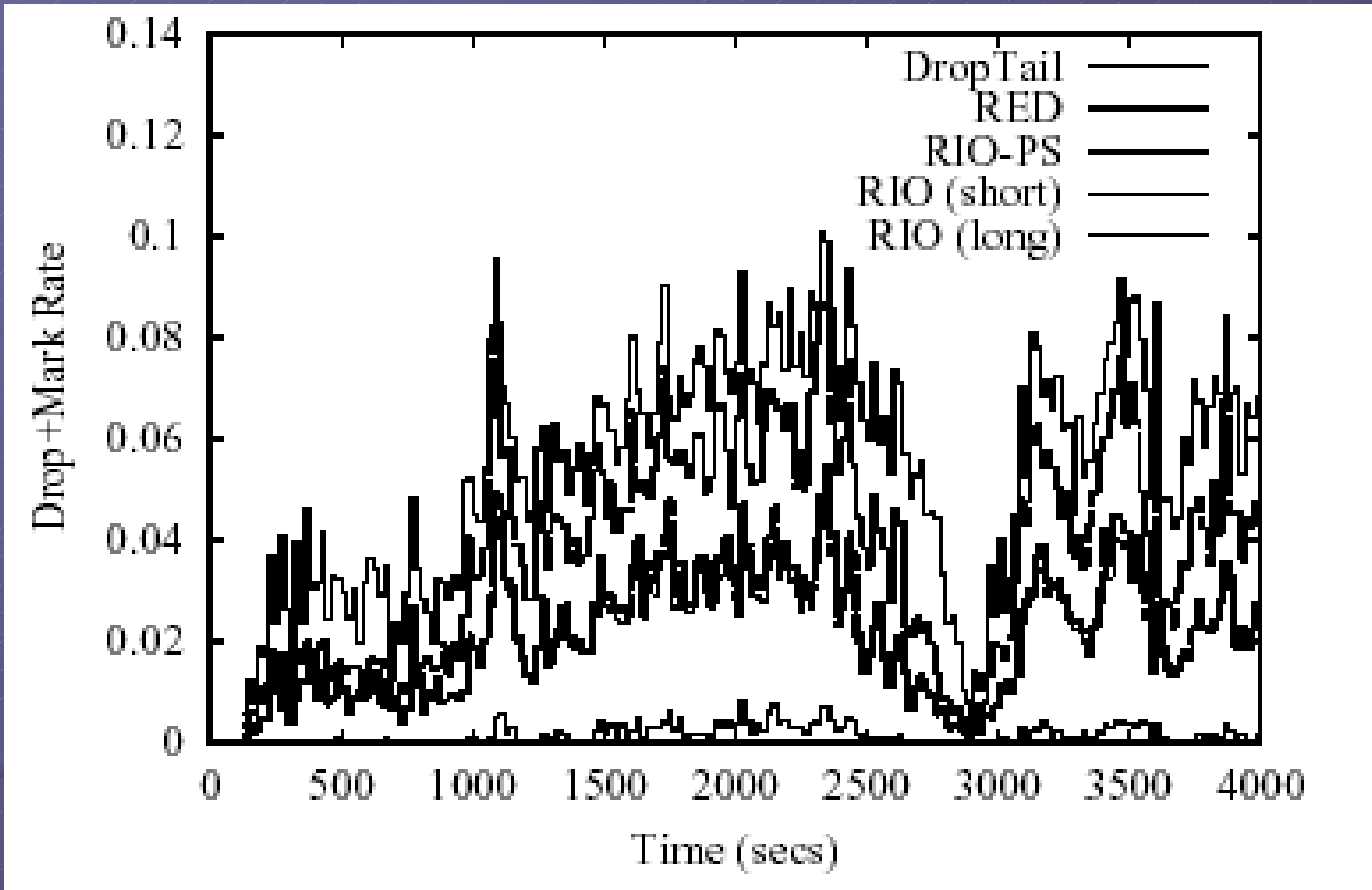


Average response time for different sized objects

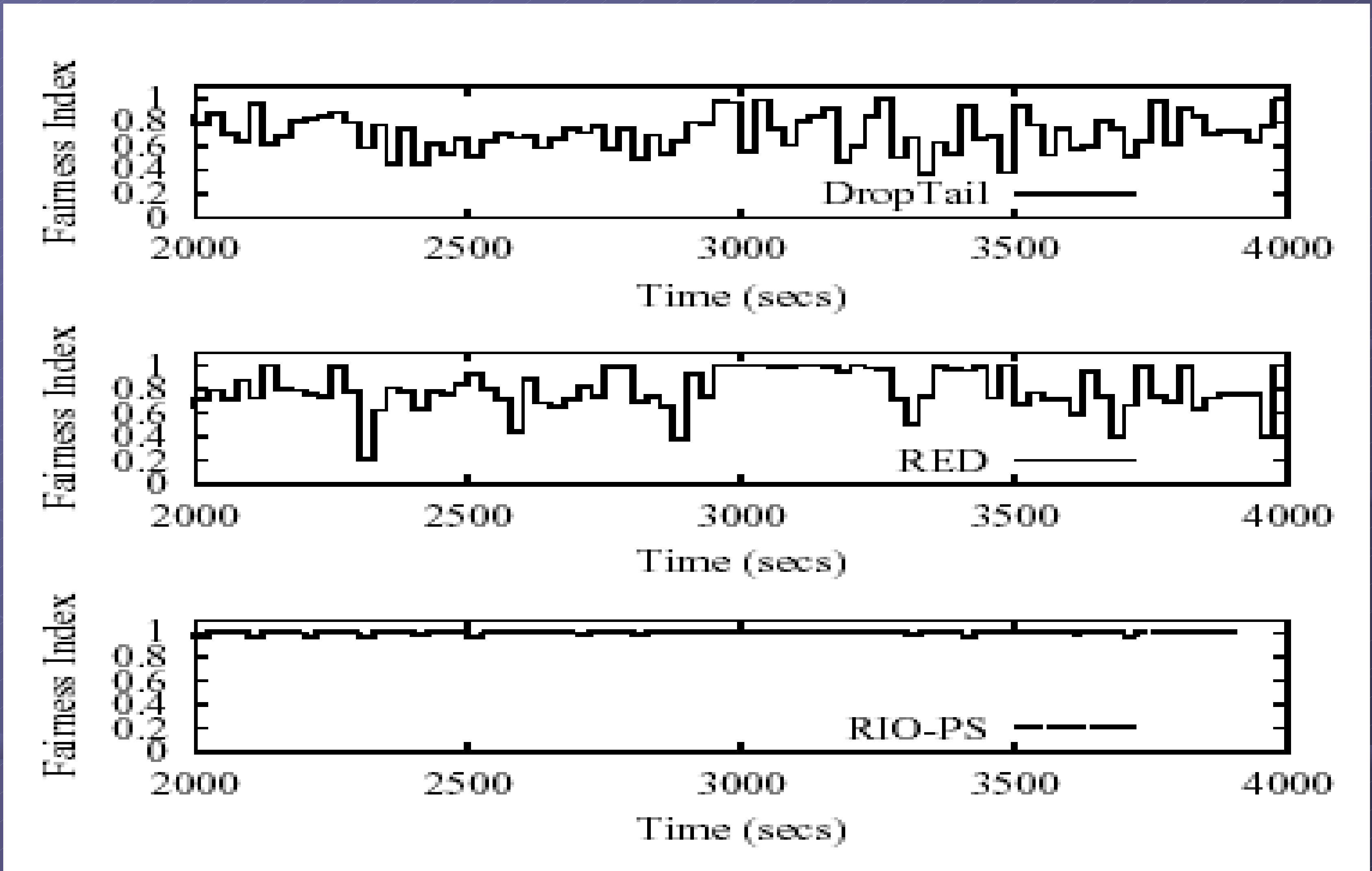
Initial retransmission timer = 1 second



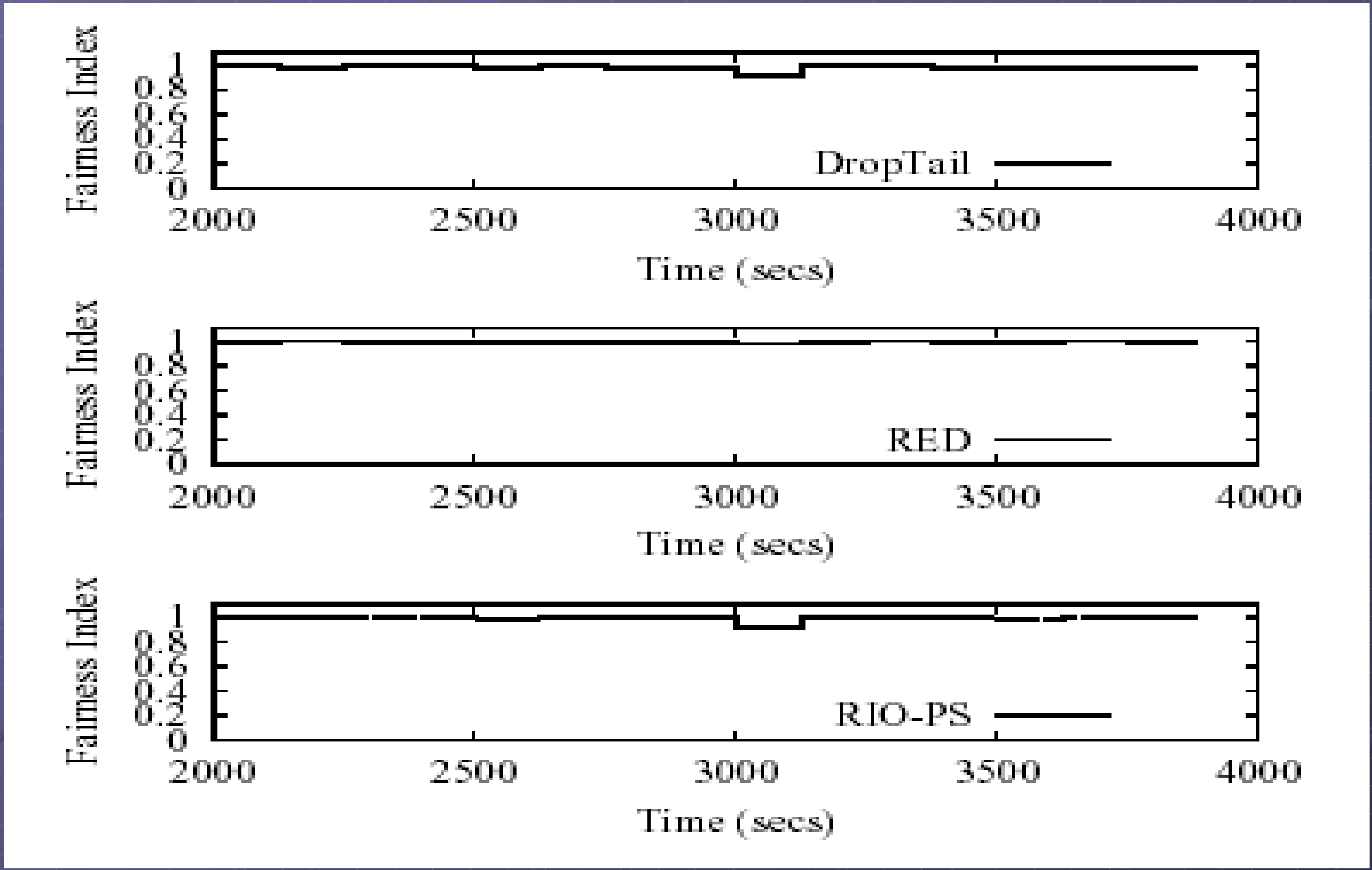
Instantaneous Queue Size



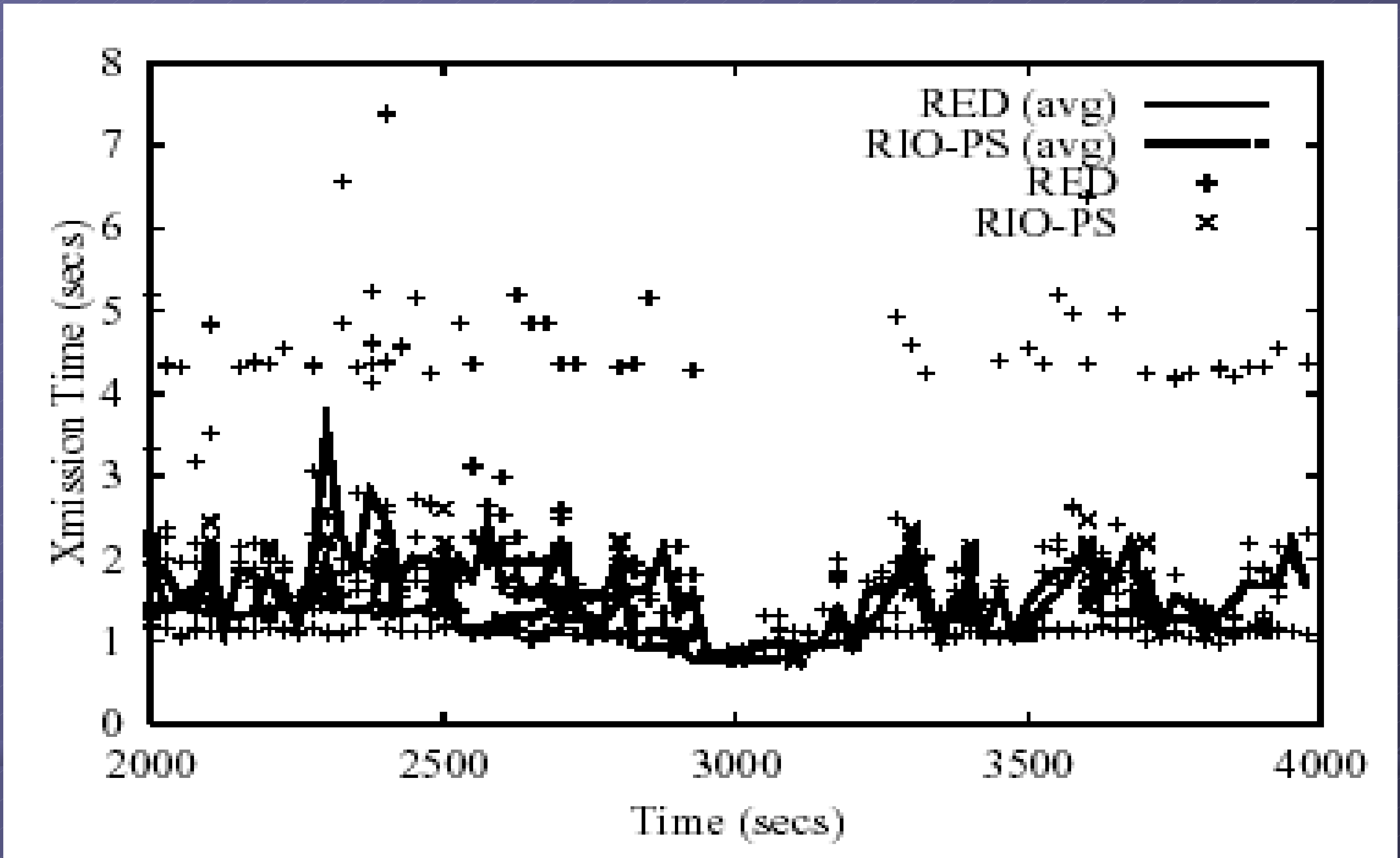
Instantaneous Drop/Mark Rate



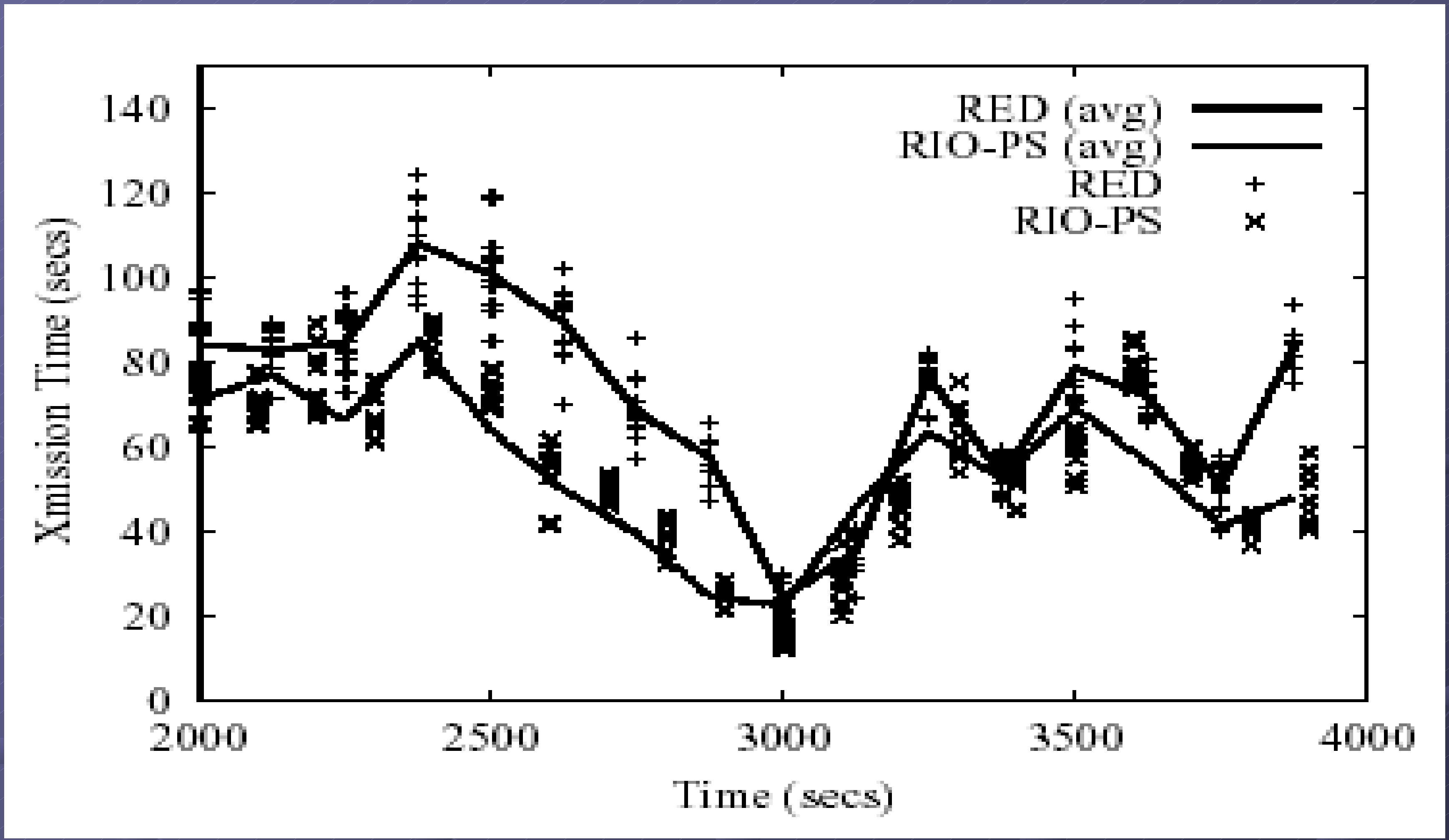
Fairness Index for Short Connections



Fairness Index For Long Connections



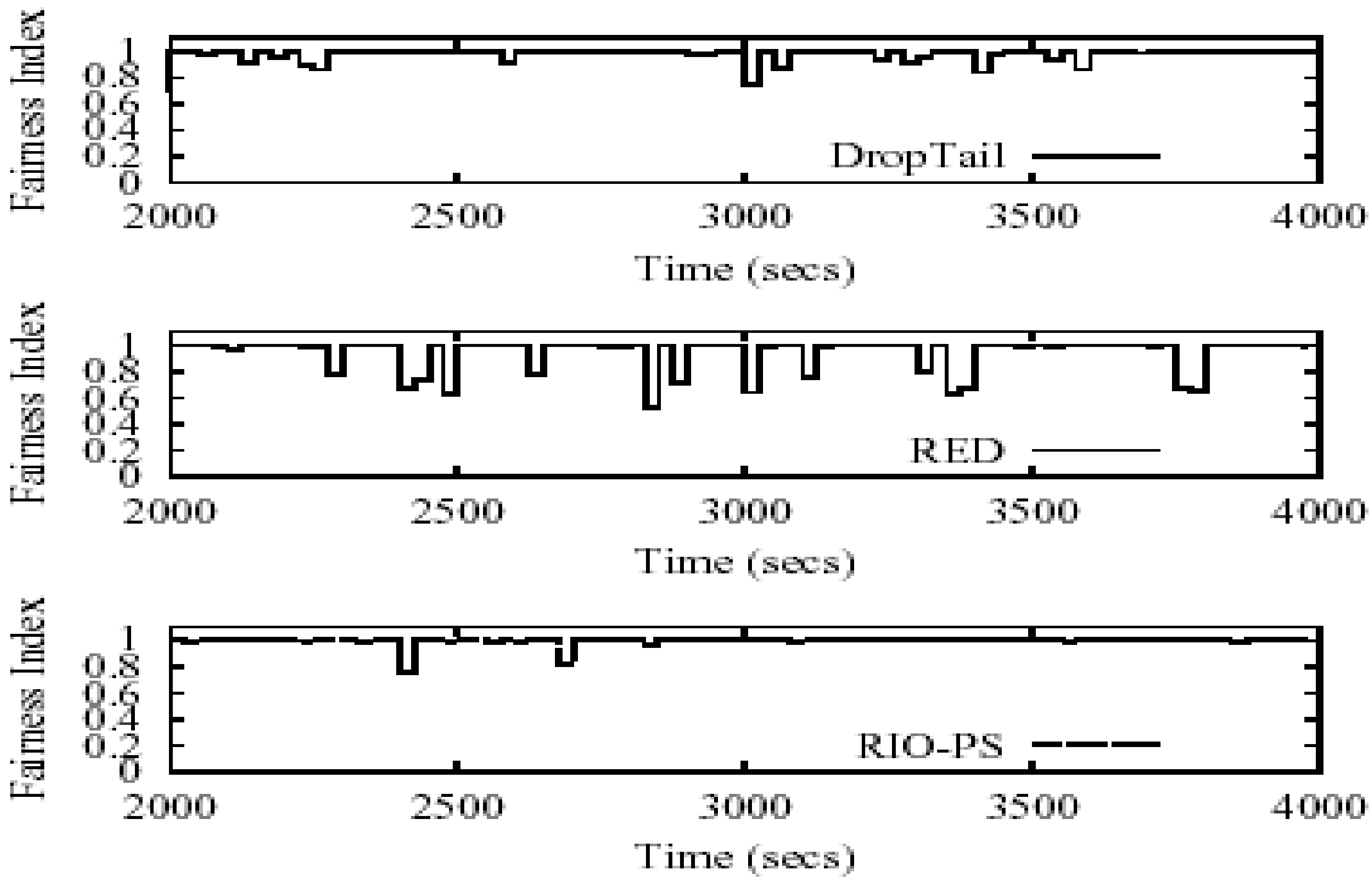
Transmission Time For Short Connections



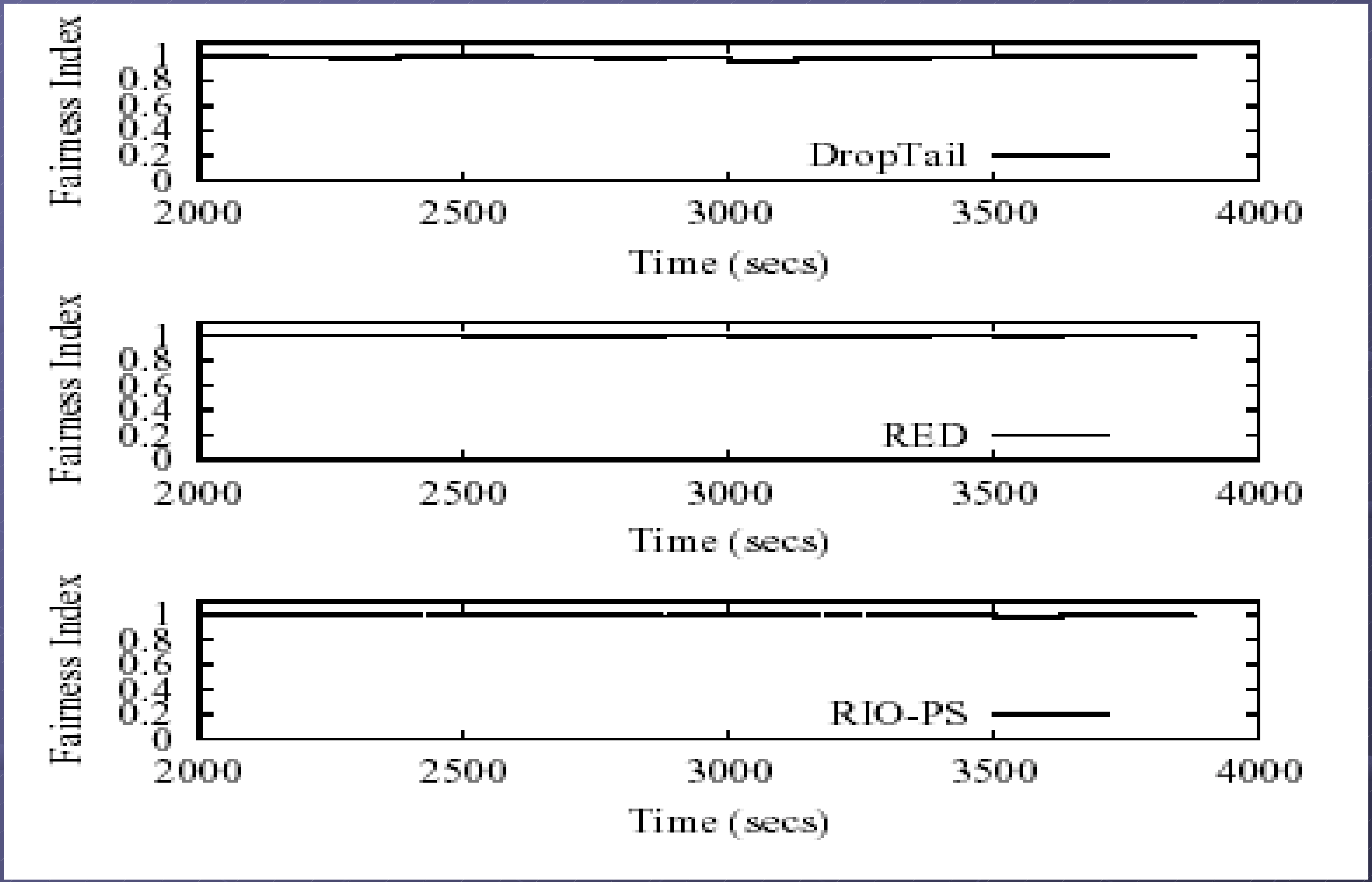
Transmission Time For Long Connections

Scheme	DropTail	RED	RIO-PS
Exp1 (ITO=3sec)	4207841	4264890	4255711
Exp1 (ITO=1sec)	4234309	4254291	4244158
Exp2 (ITO=3sec)	4718311	4730029	4723774

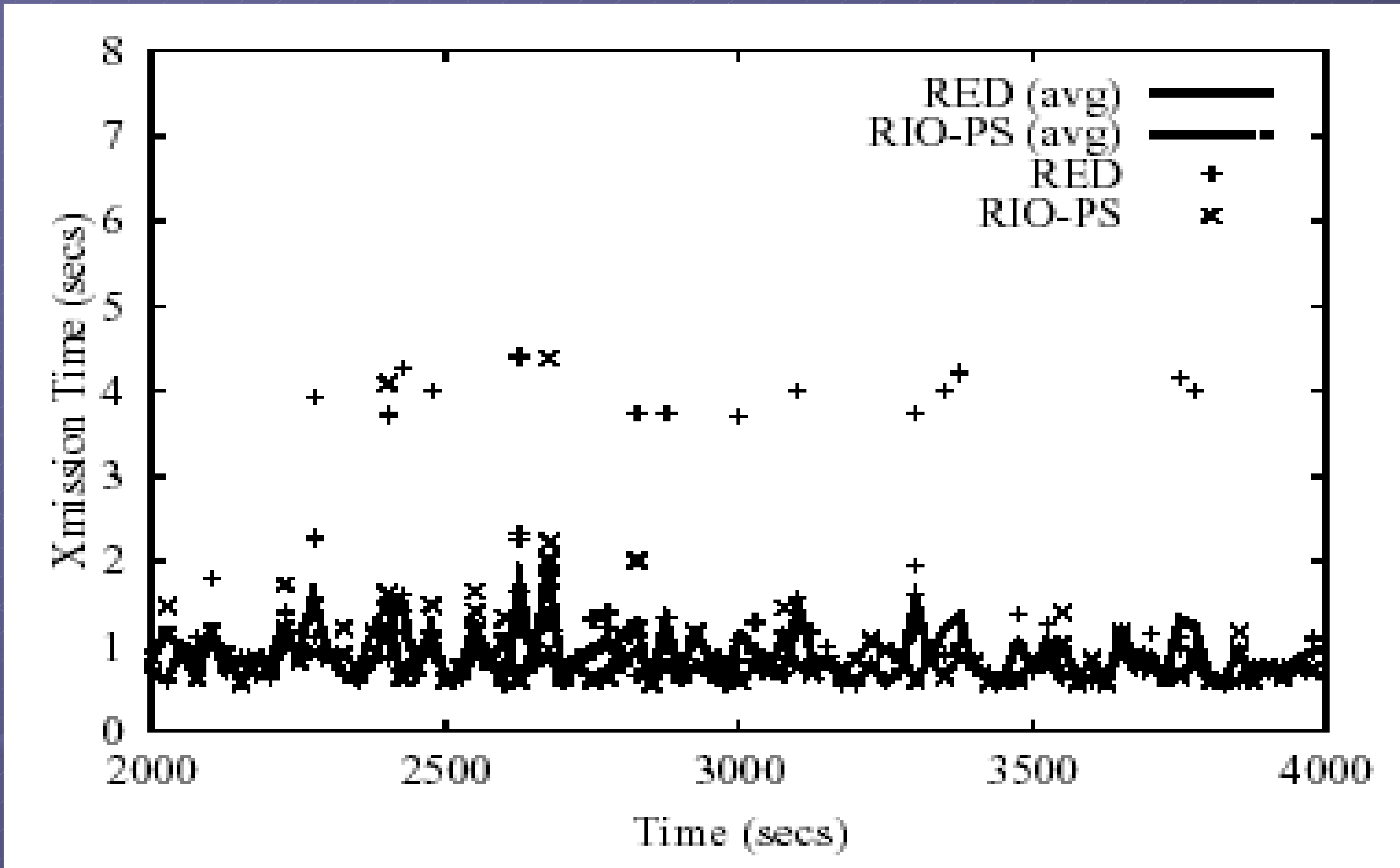
Network Goodput Over The Last 2000 Seconds



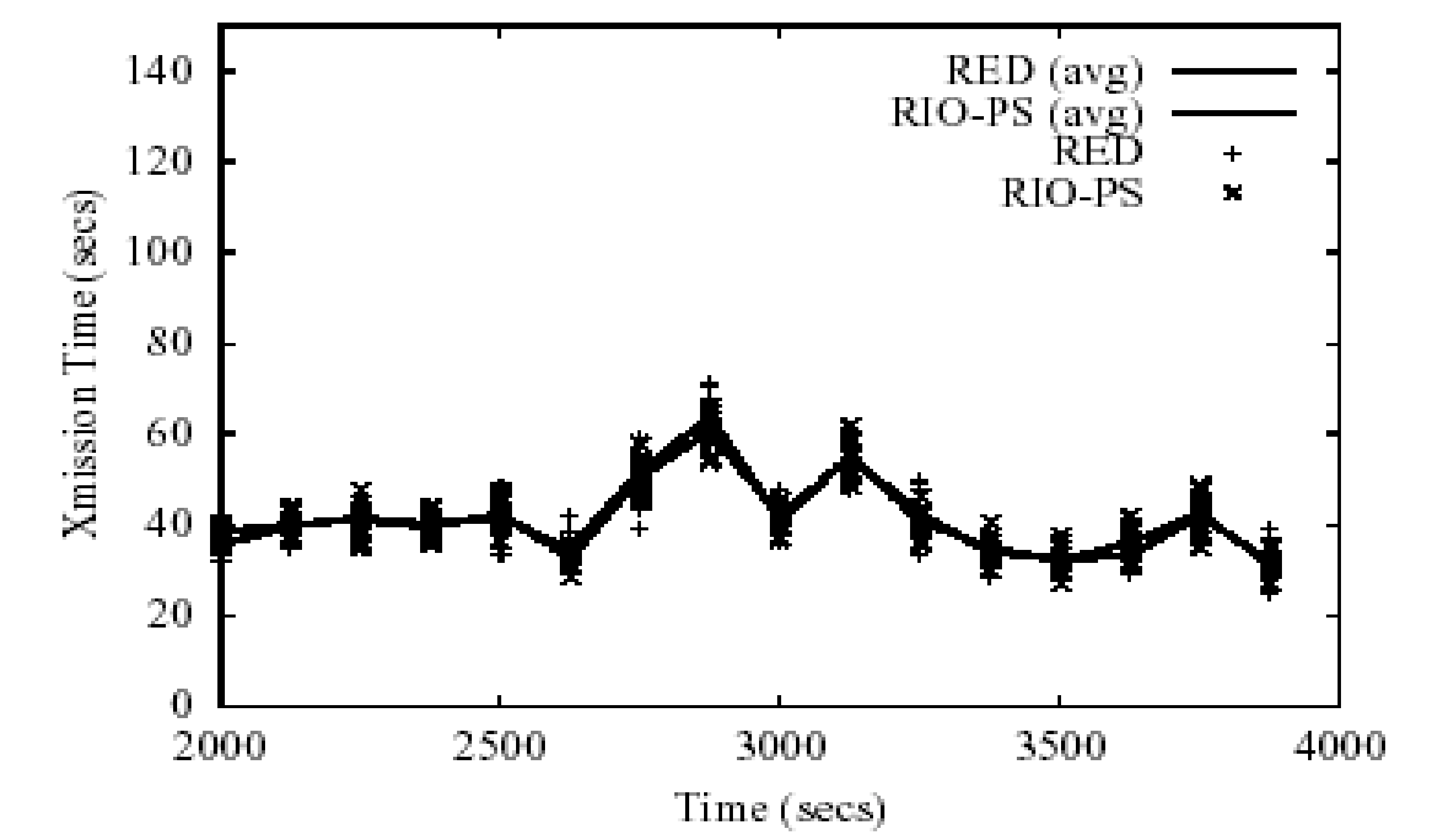
Fairness Index of Short Connections



Fairness Index For Long Connections



Transmission Time For Short Connections



Transmission Time For Long Connections

Outline

- Introduction
- Analyzing Short TCP Flow Performance
- RIO-PS: Architecture and Mechanisms
- Simulations
- Discussion
- Conclusions and Future Work

Discussion

1. Simulation Model

- All TCP connections have similar end to end propagation delays.

- Better performance obtained in the presence of reverse traffic.

2. The Queue Management Policy

- RIO does not provide class based or flow guarantees.

- Other option could be PI controlled RED queue.

Discussion

3. Deployment Issues

   - Success of the scheme depends on how well the edge router can classify the traffic to be long or short.

4. Flow Classification

   - Long connections initially are classified as Short and so are treated with higher preference.

Discussion

5. Controller Design

- "SLR" depends on

  Tc = time after which classification threshold Lt is updated.

  Tu = time after which active flow table is updated.

6. Malicious Users

- Breaking long transmissions into short.
- Problem of overhead.

Outline

- Introduction
- Analyzing Short TCP Flow Performance
- RIO-PS: Architecture and Mechanisms
- Simulations
- Discussion
- Conclusions and Future Work

Conclusions

- Implementation of TCP service that classifies traffic based on their Duration.
- Performance of Short TCP connections is Improved.
- Performance of first few Long connections is also improved.
- Proposed Architecture is good in the terms of deployment.