

# Internet Performance Dynamics

---

Paul Barford

Boston University Computer Science Department

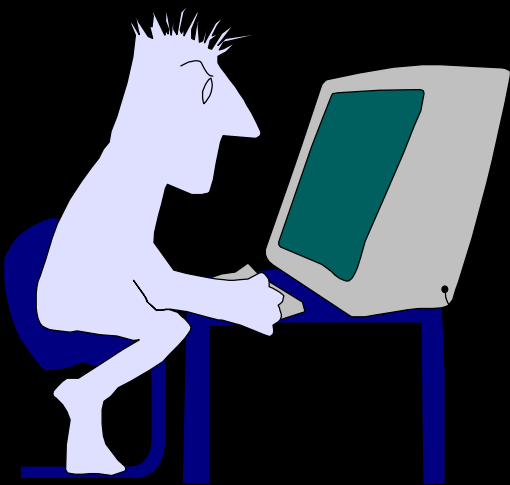
<http://cs-people.bu.edu/barford/>

Fall, 2000



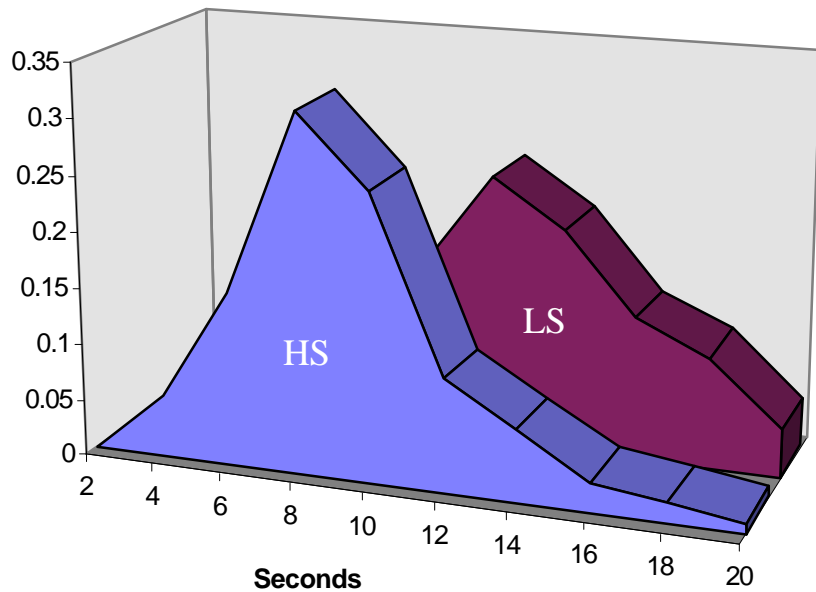
# Motivation

---



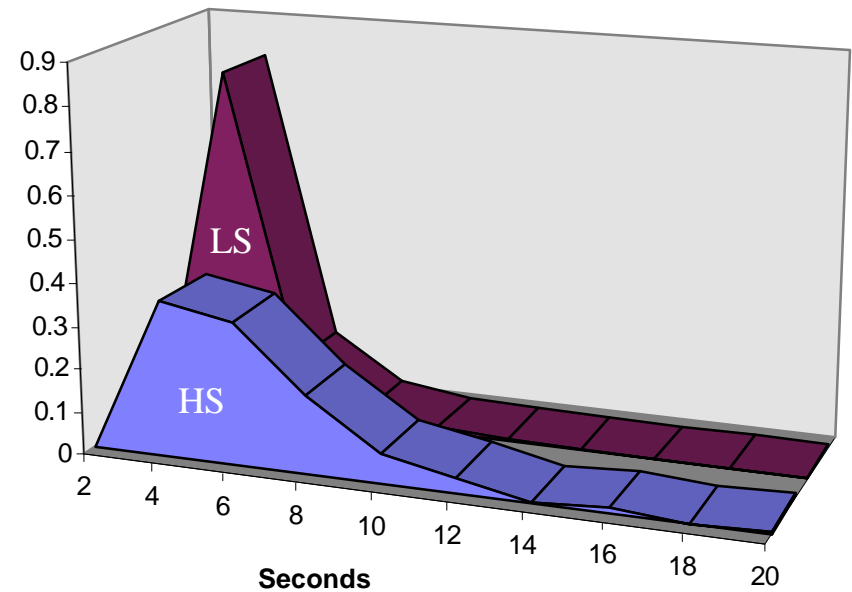
- What are the root causes of long response times in wide area services like the Web?
  - Servers?
  - Networks?
  - Server/network interaction?

# A Challenge



HS mean = 8.3 sec. LS mean = 13.0 sec.

Day 1



HS mean = 5.8 sec. LS mean = 3.4 sec.

Day 2

Histograms of file transfer latency for 500KB files transferred between Denver and Boston

Precise separation of server effects from network effects is difficult

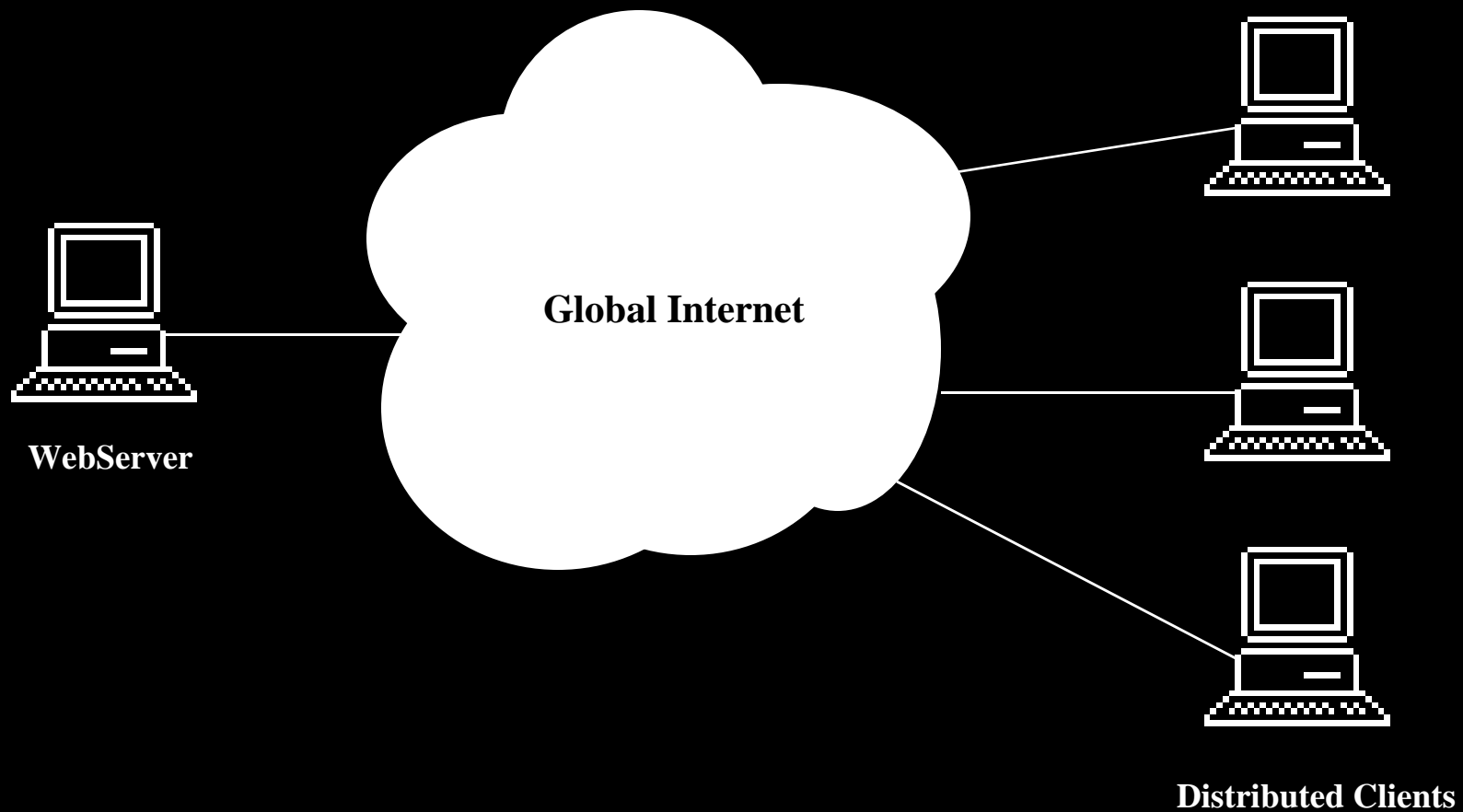
## What is needed?

---

- A laboratory enabling detailed examination of Web transactions (Web “microscope”)
  - Wide Area Web Measurement (WAWM) project testbed
- Technique for analyzing transactions to separate and identify causes of delay
  - Critical path analysis of TCP

# Web Transactions “under a microscope”

---



# Generating Realistic Server Workloads

---

- Approaches:
  - Trace-based:
    - Pros: Exactly mimics known workload
    - Cons: “black box” approach, can’t easily change parameters of interest
  - Analytic: synthetically create a workload
    - Pros: Explicit models can be inspected and parameters can be varied
    - Cons: Difficult to identify, collect, model and generate workload components

# SURGE: Scalable URL Reference Generator

---

- Analytic Web workload generator
  - Based on 12 empirically derived distributions
  - Explicit, parameterized models
  - Captures “heavy-tailed” (highly variable) properties of Web workloads
- SURGE components:
  - Statistical distribution generator
  - Hyper Text Transfer Protocol (HTTP) request generator
- Currently being used at over 130 academic and industrial sites world wide
  - Adopted by W3C for HTTP-NG testbed

# Seven workload characteristics captured in SURGE



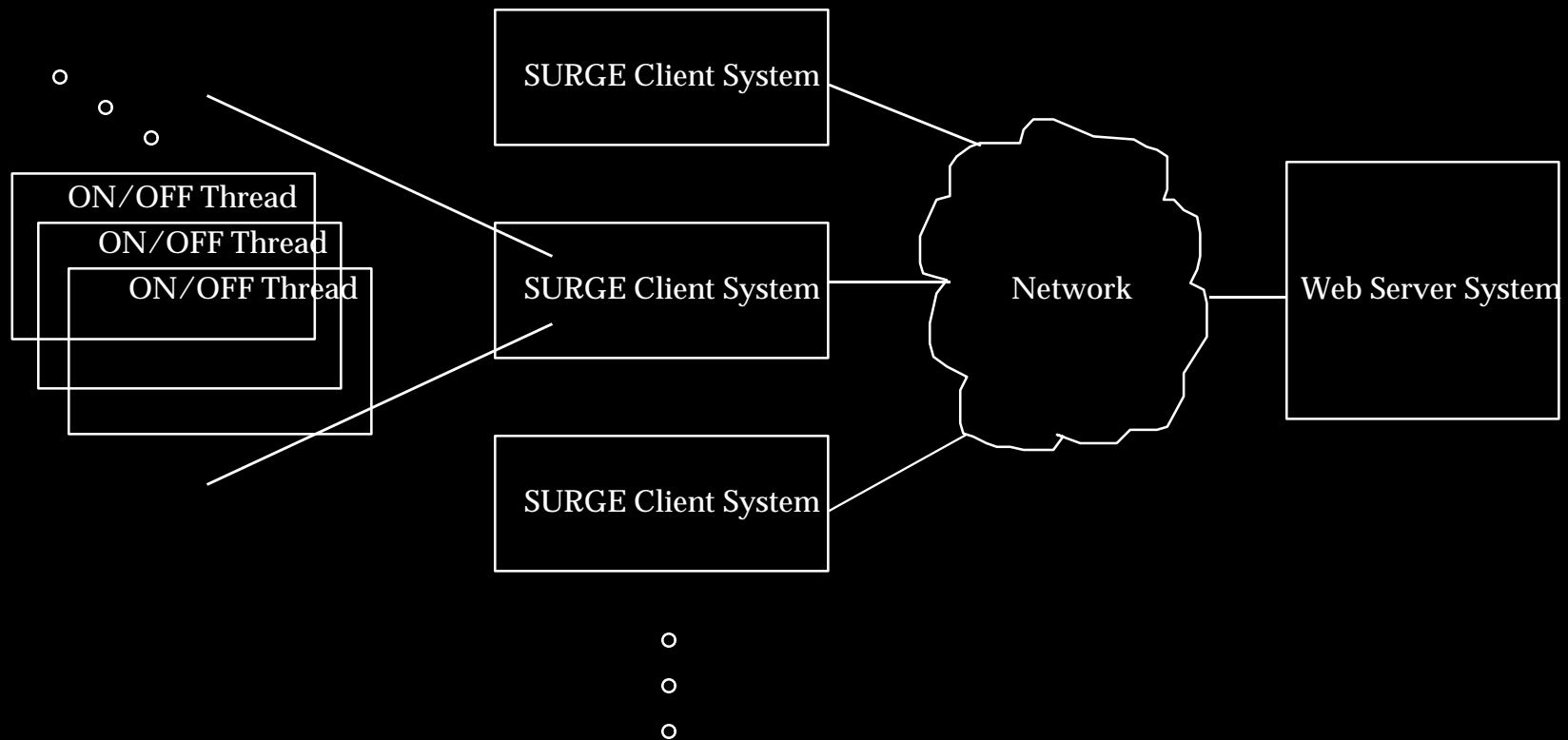
<u>Characteristic</u>	<u>Component</u>	<u>Model</u>	<u>System Impact</u>	
File Size	Base file - body	Lognormal	File System	*
	Base file - tail	Pareto		*
	Embedded file	Lognormal		*
	Single file1	Lognormal		*
	Single file 2	Lognormal		*
Request Size	Body	Lognormal	Network	*
	Tail	Pareto		*
Document Popularity		Zipf	Caches, buffers	
Temporal Locality		Lognormal	Caches, buffers	
OFF Times		Pareto		*
Embedded References		Pareto	ON Times	*
Session Lengths		Inverse Gaussian	Connection times	

\* Model developed during the SURGE project

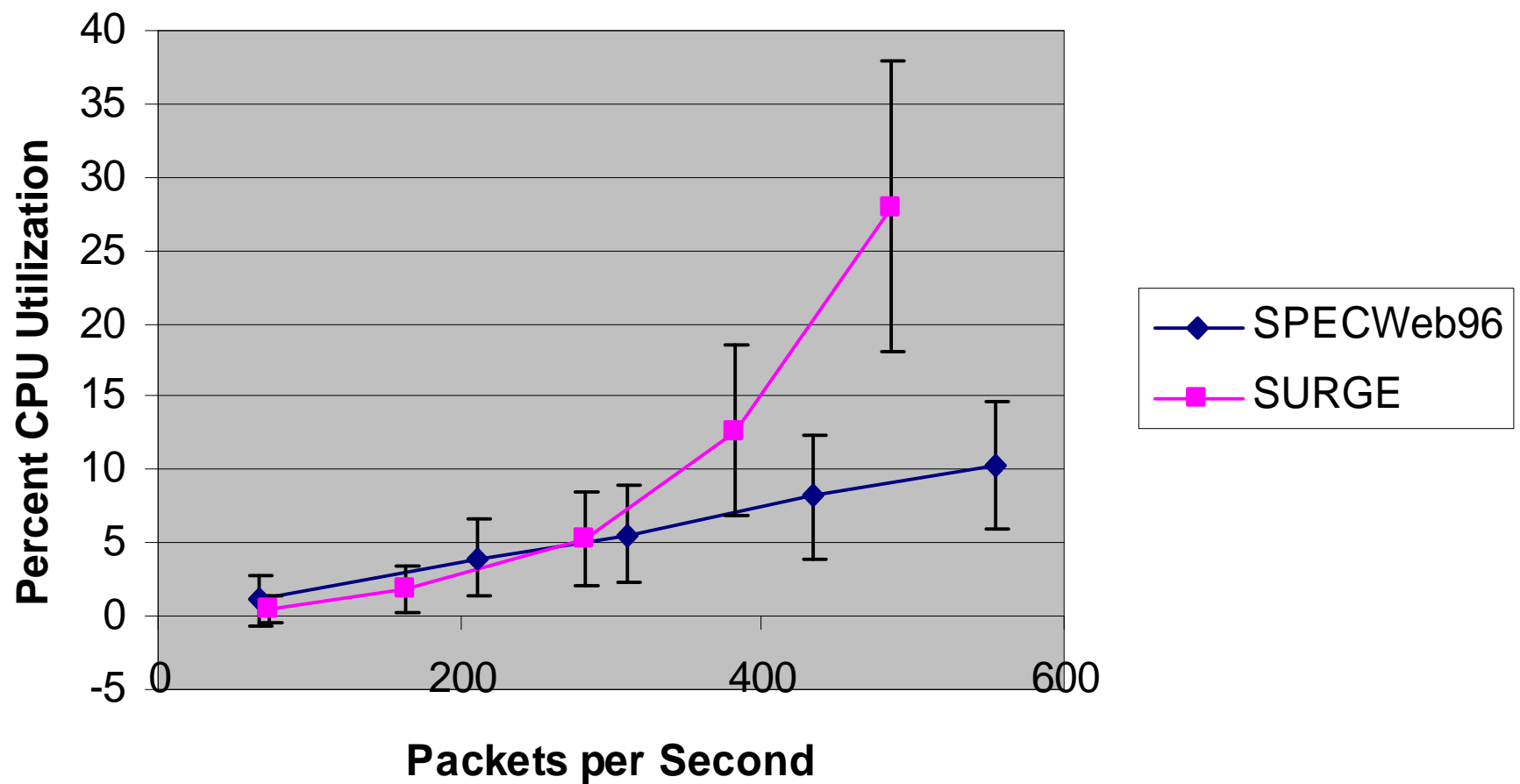
# HTTP request generator

---

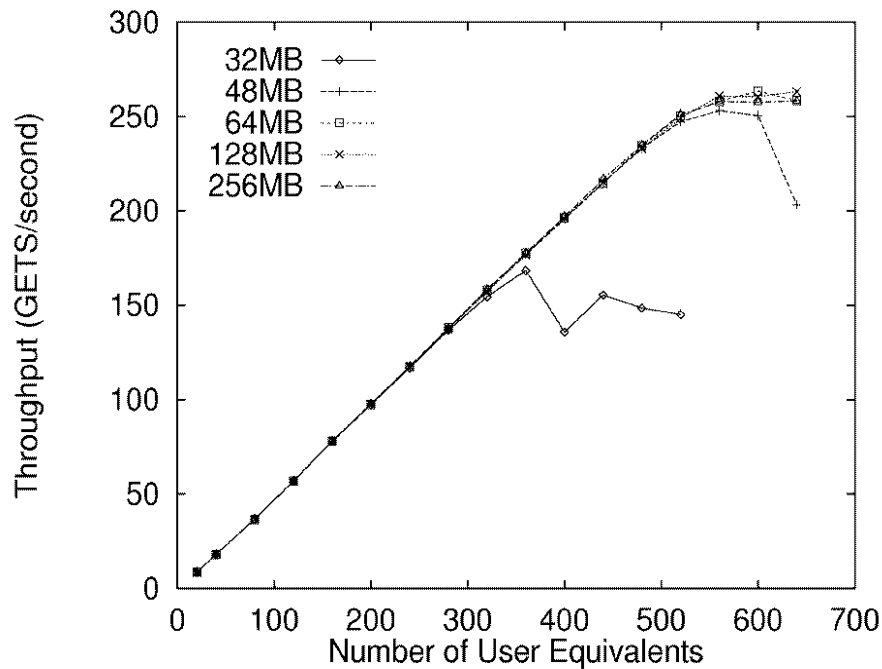
- Supports both HTTP/1.0 and HTTP/1.1
- ON/OFF thread is a “user equivalent”



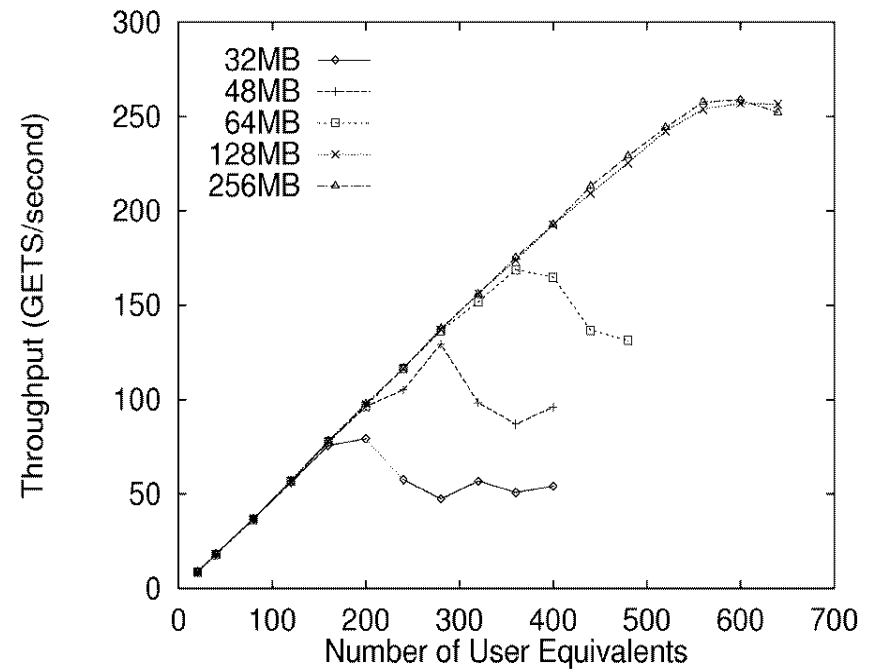
## SURGE and SPECWeb96 exercise servers very differently



# SURGE's flexibility allows easy experimentation



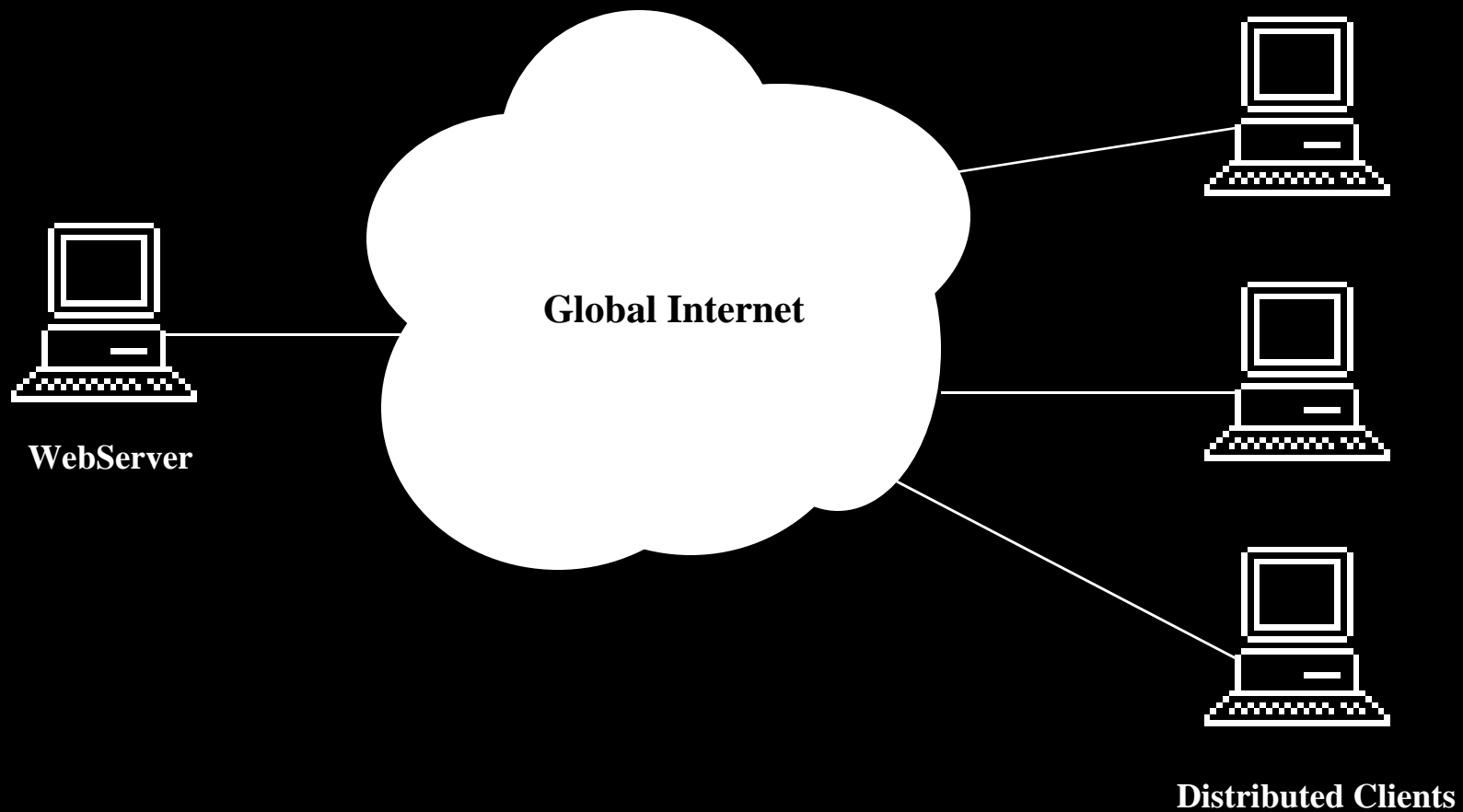
HTTP/1.0



HTTP/1.1

# Web Transactions “under a microscope”

---



## WAWM Infrastructure

---

- 13 clients distributed around the global Internet
  - Execute transactions of interest
- One server cluster at BU
  - Local load generators running SURGE enable server to be placed under any load condition
- Active and passive measurements from both server and clients
  - Packet capture via “tcpdump”
- GPS timers

# WAWM client systems

---

Harvard University, MA

Purdue University, IN

University of Denver, CO

ACIRI, Berkeley, CA

HP, Palo Alto, CA

University of Saskatchewan, Canada

University Federal de Minas Gerais, Brazil

University Simon Bolivar, Venezuela

EpicRealm - Dallas, TX

EpicRealm – Atlanta, GA

EpicRealm - London, England

EpicRealm - Tokyo, Japan

Internet2/Surveyor

Others??

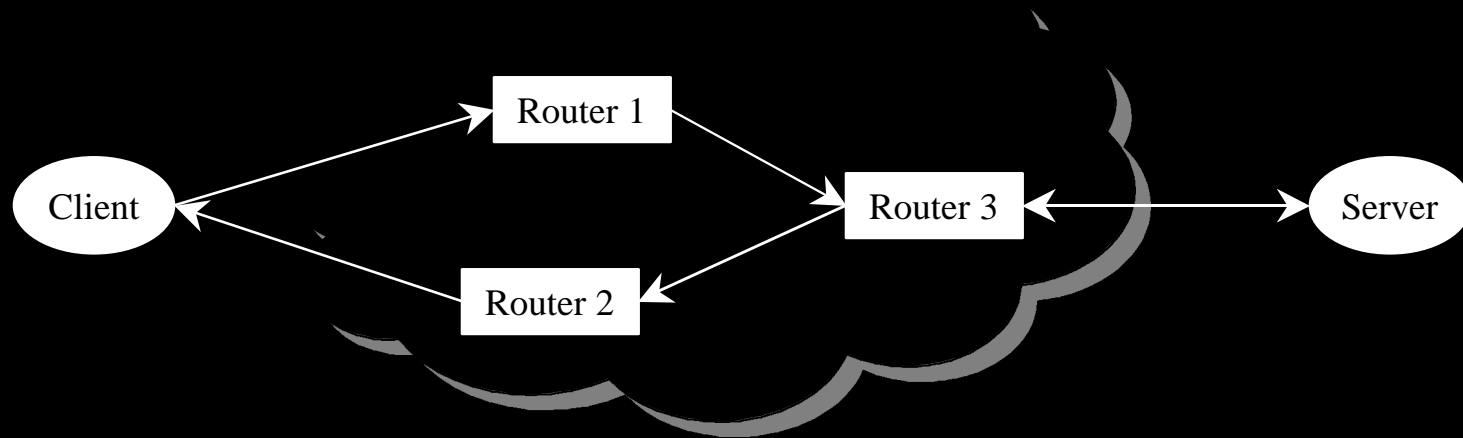
## What is needed?

---

- A laboratory enabling detailed examination of Web transactions (Web “microscope”)
  - Wide Area Web Measurement (WAWM) project testbed
- ■ Technique for analyzing transactions to separate and identify causes of delay
  - Critical path analysis of TCP

# Identifying root causes of response time

---



- Delays can occur at many points along the end-to-end path *simultaneously*
- Pinpointing where delays occur *and which delays matter* is difficult
- Our goal is to identify precisely the determiners of response time in TCP transactions

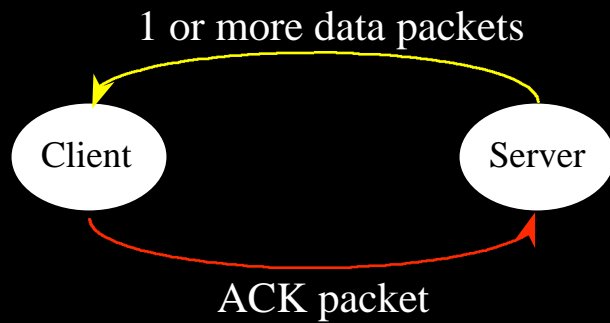
# Critical path analysis (CPA) for TCP transactions

---

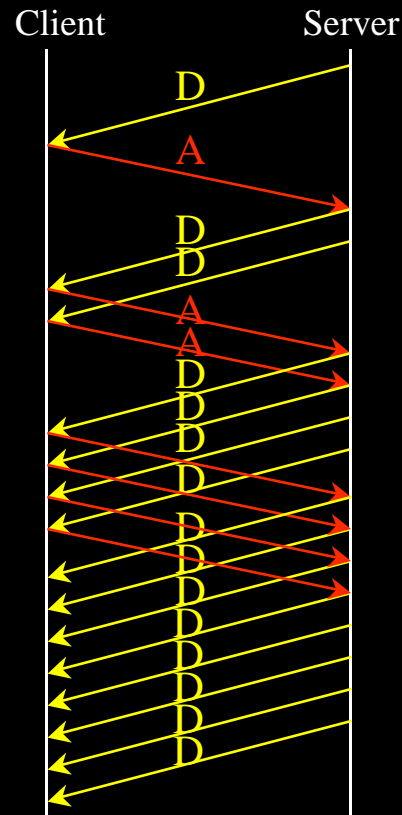
- CPA identifies the precise set of events that determine execution time of a distributed application
  - Web transaction response time
- Decreasing duration of any event on the CP decreases response time
  - not true for events off the CP
- Profiling the CP for TCP enables accurate assignment of delays to:
  - Server delay
  - Client delay
  - Network delay (propagation, network variance and drops)
- Applied to HTTP/1.0
  - Could apply to other applications (eg. FTP)

# Window-based flow control in TCP

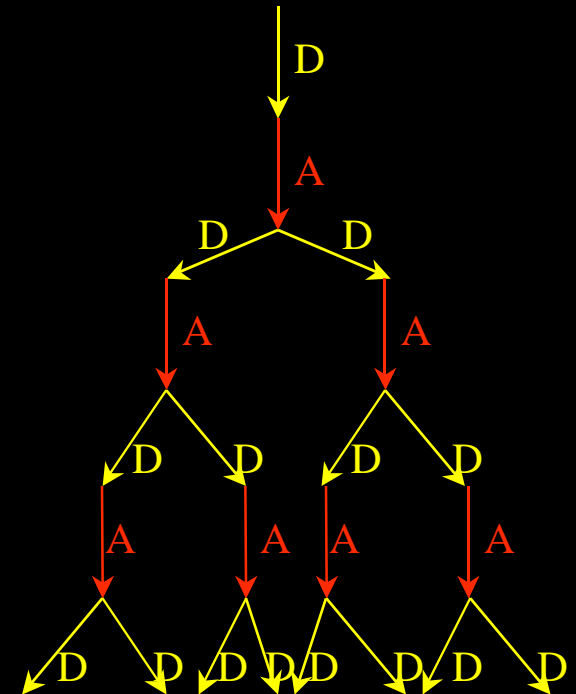
System



Time line



Graph



## TCP flows as a graph

---

- Vertices are packet departures or arrivals
  - Data, ACK, SYN, FIN
- Directed edges reflect Lamport's "happens before" relation
  - On client or server or over the network
- Weights are elapsed time
  - Assumes global clock synchronization
- Profile associates categories with edge types
  - Assignment based on logical flow

Original Data Flow

Rounds

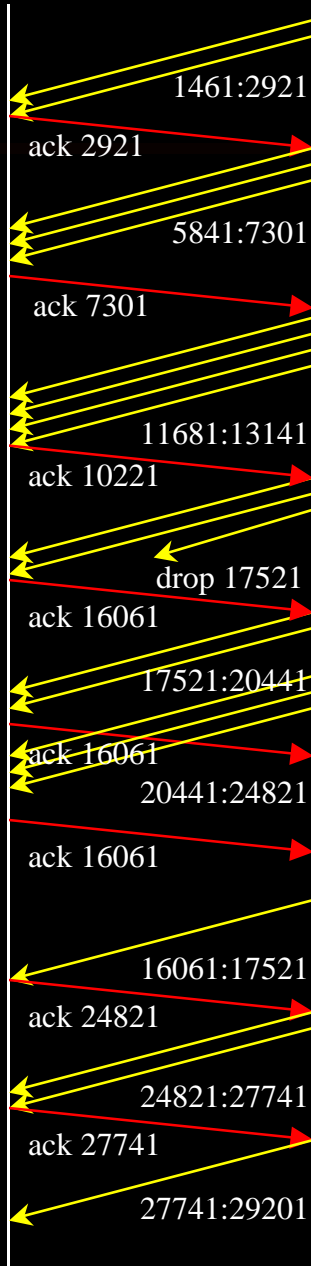
Critical Path

Profile

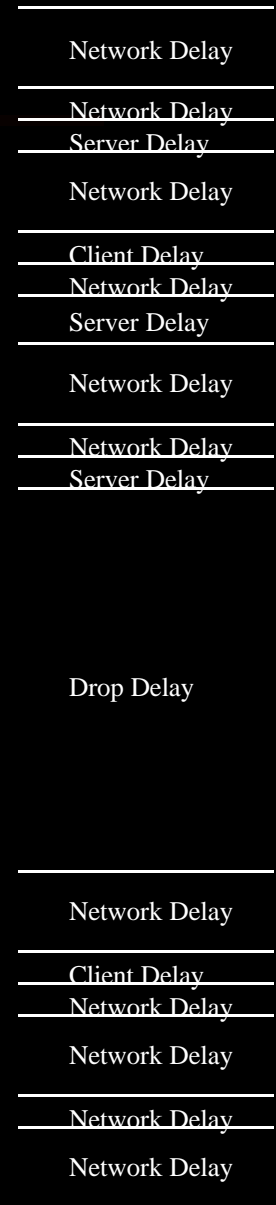
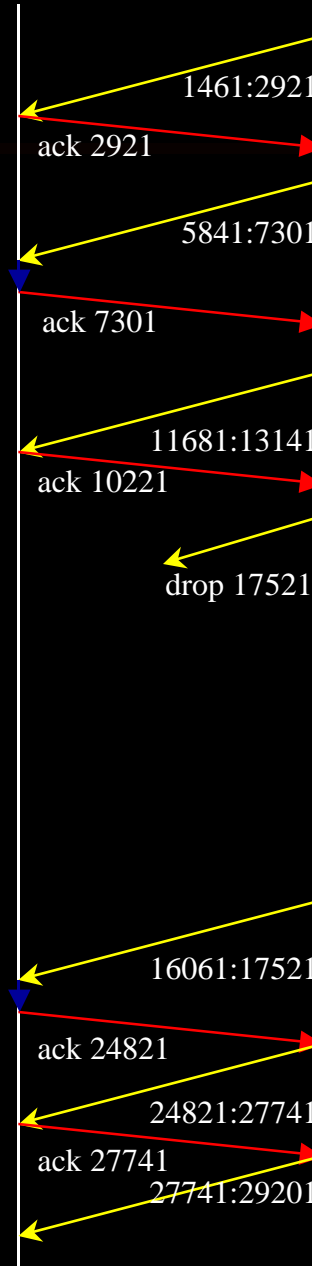
Client Server

Number Bytes Liberated

Client Server



1 1:2920  
2 2921:7300  
3 7301:13140  
4 13141:17520  
5 17520:24820  
6 24821:27740  
7 27741:30660



## tcpeval

---

- Inputs are “tcpdump” packet traces taken at end points of transactions
- Generates a variety of statistics for file transactions
  - File and packet transfer latencies
  - Packet drop characteristics
  - Packet and byte counts per unit time
- Generates both timeline and sequence plots for transactions
- Generates critical path profiles and statistics for transactions
- Freely distributed

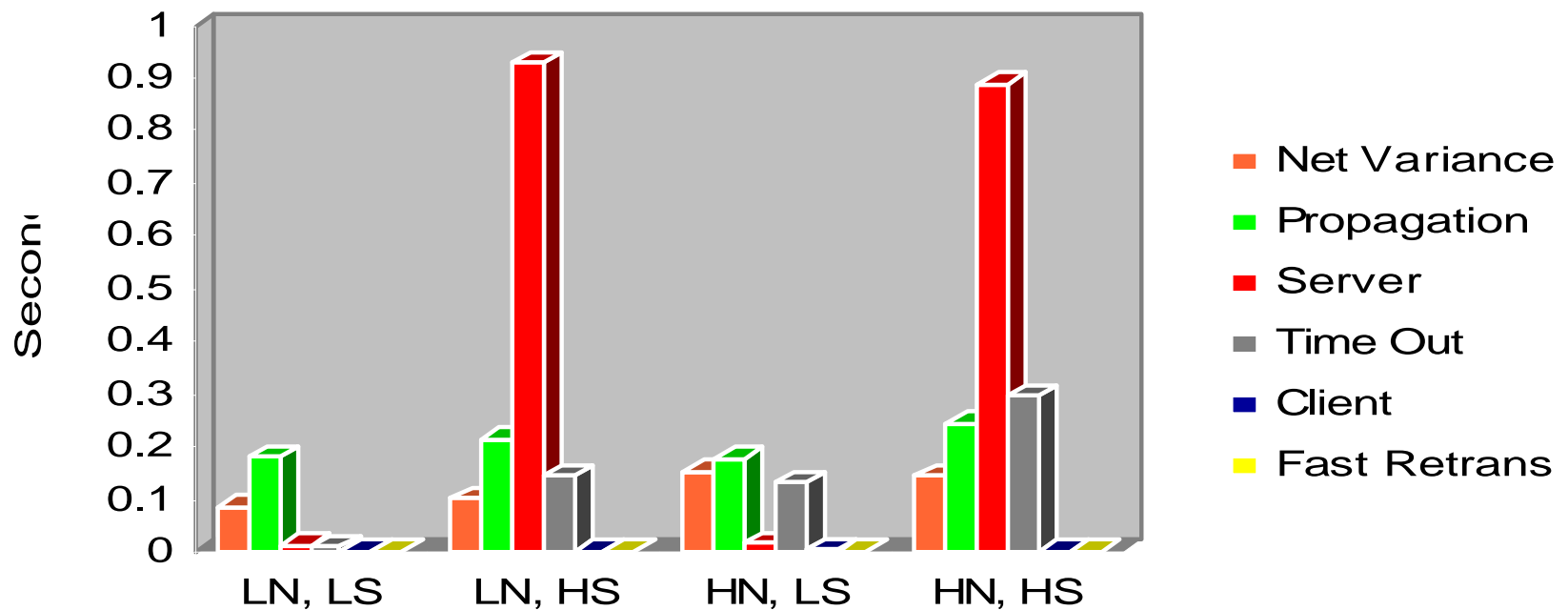
# Implementation Issues

---

- tcpeval must recreate TCP state at end points as packets arrive
  - Capturing packets at end points makes timer simulation unnecessary
  - “Active round” must be maintained
- Packet filter problems must be addressed
  - Dropped packets
  - Added packets
  - Out of order packets
- tcpeval works across platforms for RFC 2001 compliant TCP stacks

# CPA results for 1KB file

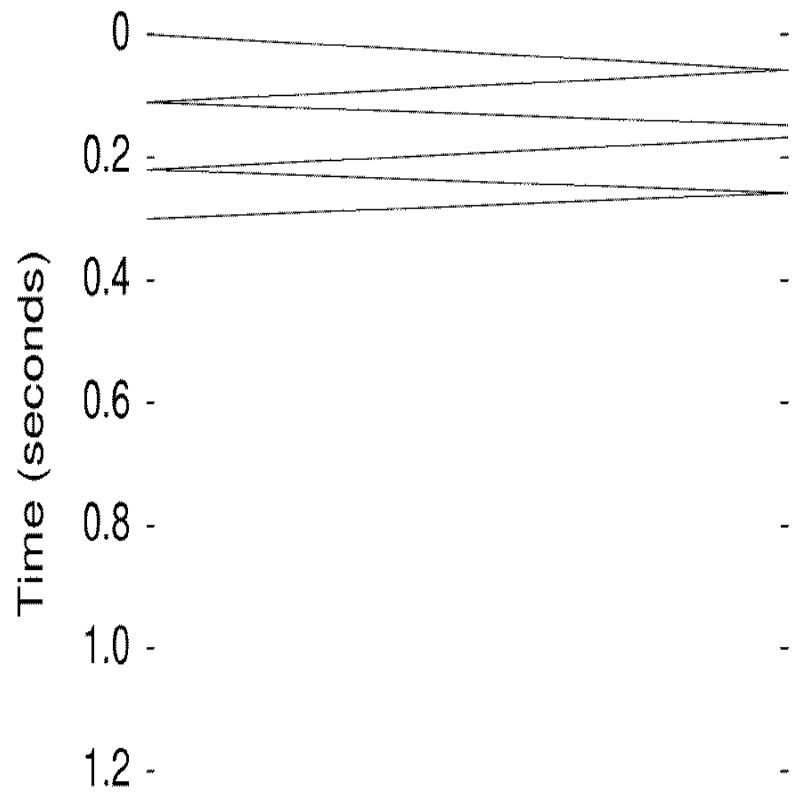
6 packets are typically on the critical path



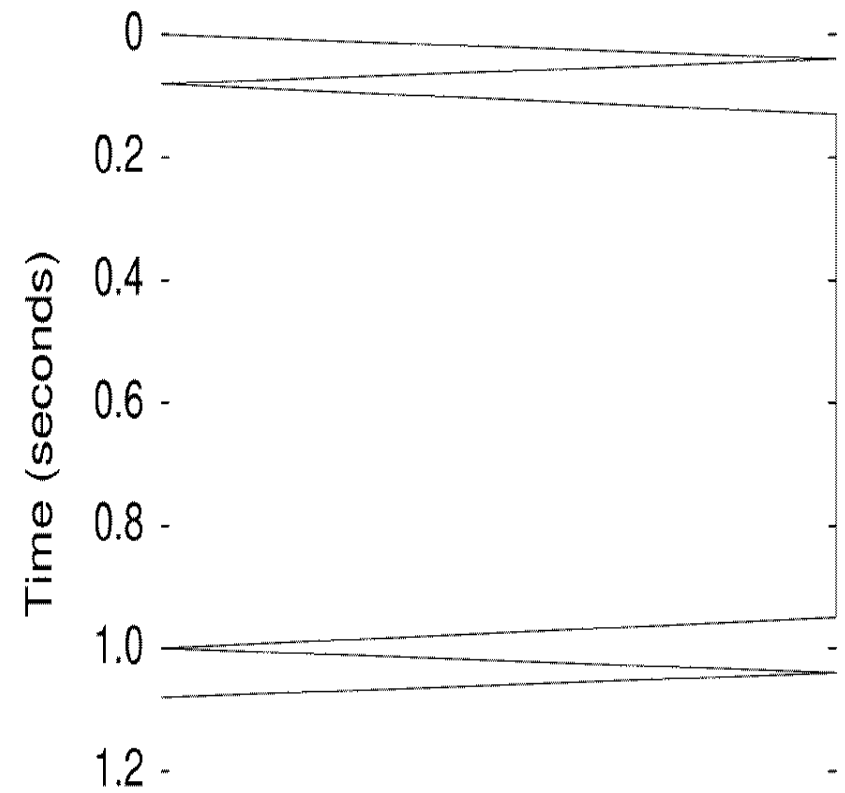
- Latency is dominated by server load for BU to Denver path

# CP time line diagrams for 1KB file

---



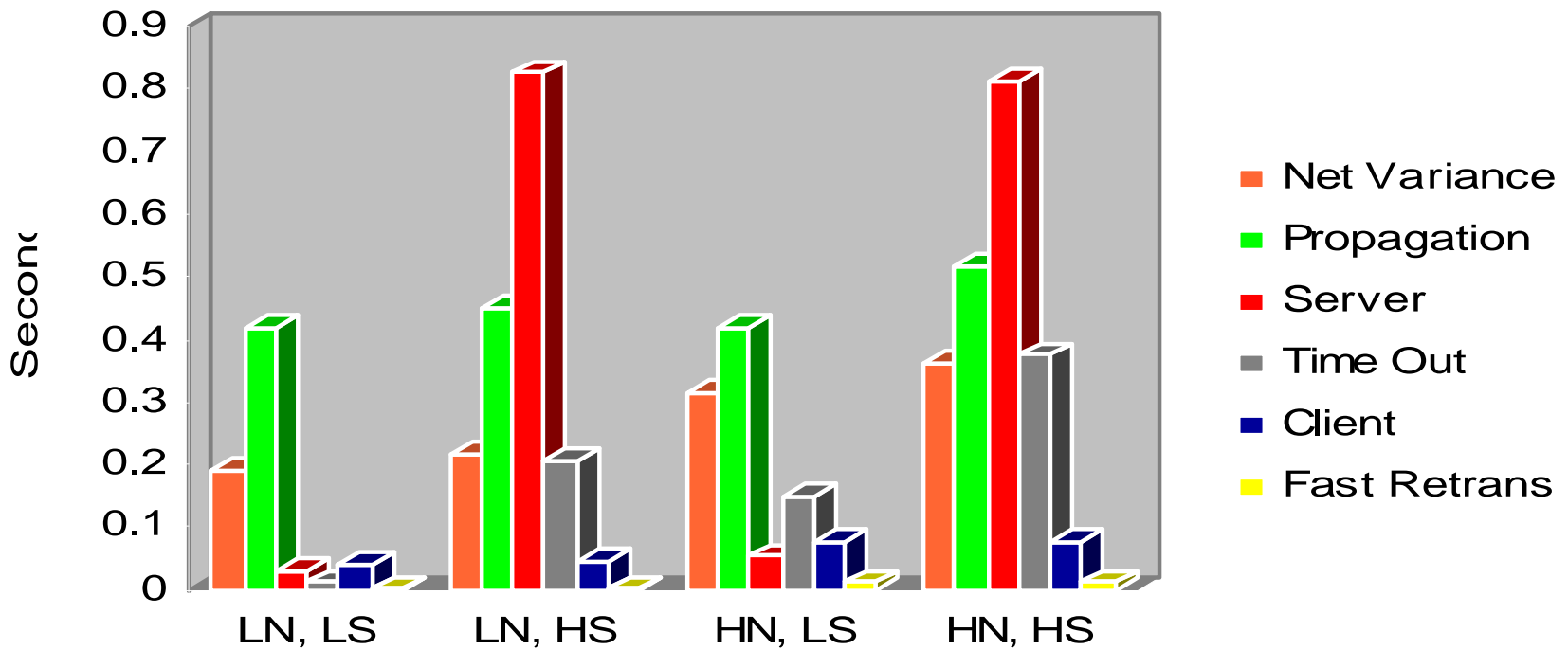
Low Server Load



High Server Load

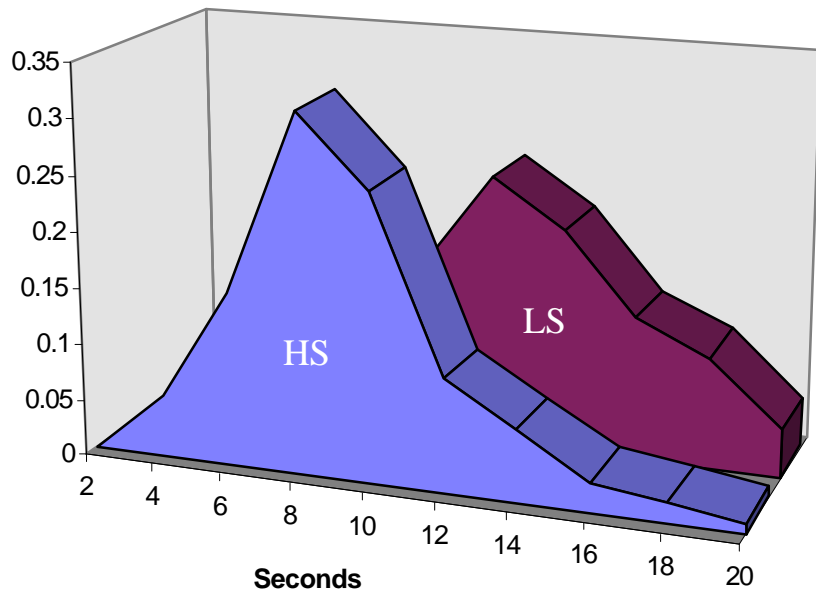
# CPA results for 20KB file

14 packets are typically on the critical path



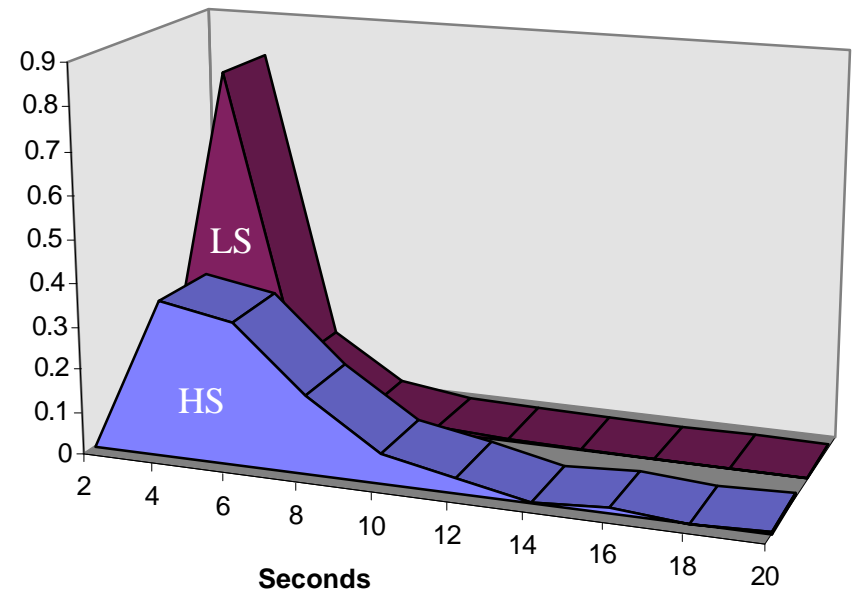
- Both server load and network effects are significant

# The Challenge



HS mean = 8.3 sec. LS mean = 13.0 sec.

Day 1



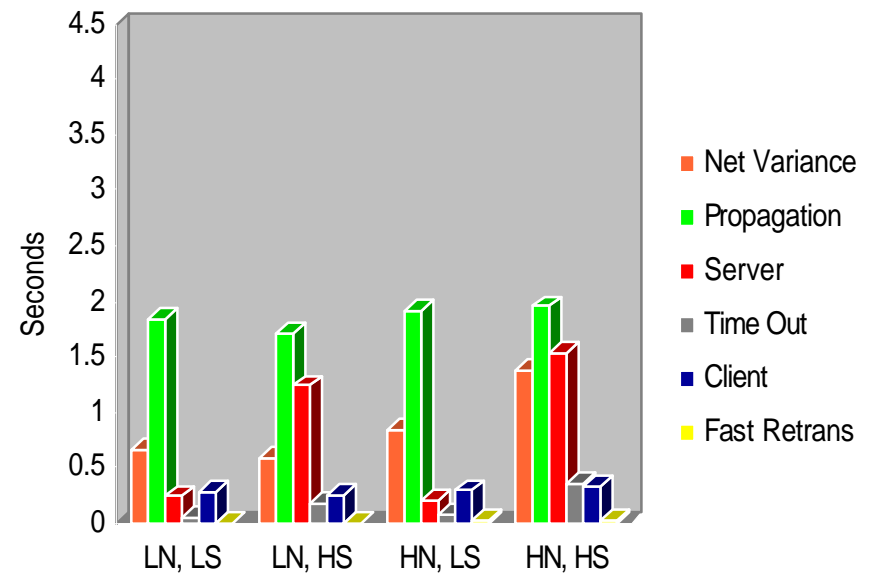
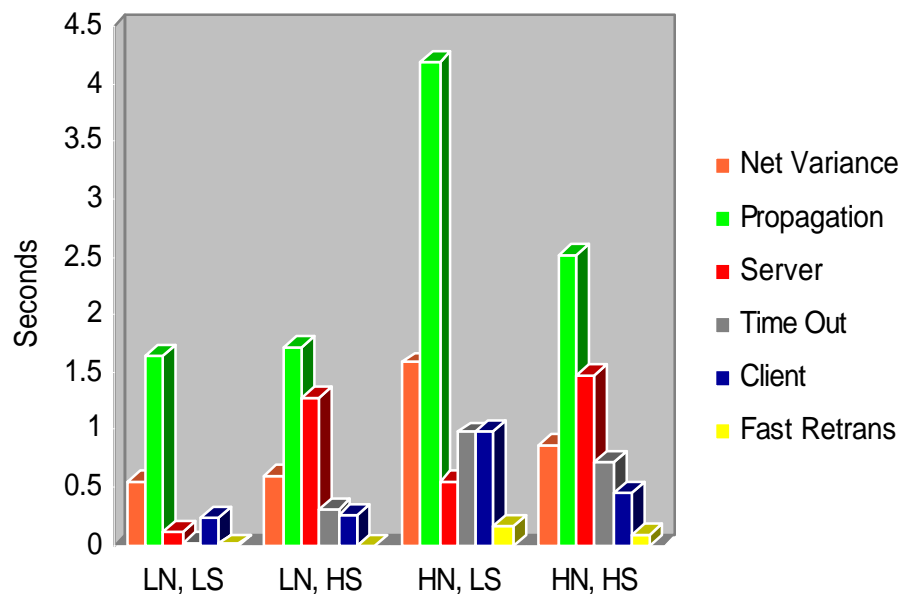
HS mean = 5.8 sec. LS mean = 3.4 sec.

Day 2

Histograms of file transfer latency for 500KB files transferred between Denver and Boston

# CPA results for 500KB file

56 packets are typically on the critical path



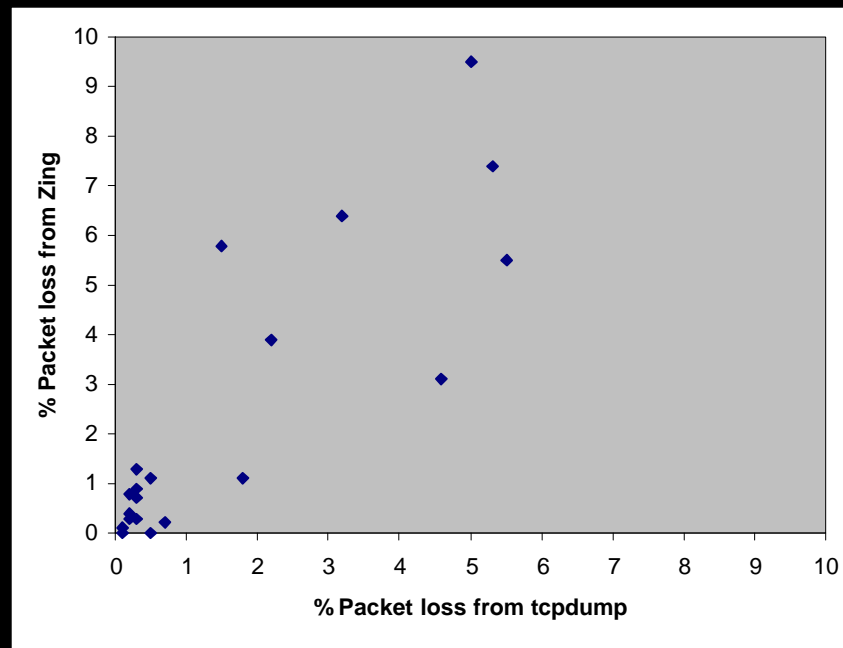
Day 1

Day 2

- Latency is dominated by network effects

# Active versus Passive Measurements

- Understanding active (Zing) versus passive (tcpdump) network measurements
- Figure shows active measures are a poor predictor of TCP performance
  - Goal is to be able to predict TCP performance using active measurements



## Related work

---

- **Web performance characterization**
  - Client studies [Catledge95,Crovella96]
  - Server studies [Mogul95, Arlitt96]
- **Wide area measurements**
  - NPD [Paxson97], Internet QoS [Huitema00], Keynote Systems Inc.
- **TCP analysis**
  - TCP modeling [Mathis97, Padhye98,Cardwell00]
  - Graphical TCP analysis [Jacobson88, Brakmo96]
  - Automated TCP analysis [Paxson97]
- **Critical path analysis**
  - Parallel program execution [Yang88, Miller90]
  - RPC performance evaluation [Schroeder89]

# Conclusions

---

- Using SURGE, WAWM can put realistic Web transactions “under a microscope”
- Complex interactions between clients, the network and servers in the wide area can lead to surprising performance
- Complex packet transactions can be effectively understood using CPA
- CP profiling of BU to Denver transactions allowed precise assignment of delays
  - Latency for small files is dominated by server load
  - Latency for large files is dominated by network effects
- Relationship between active and passive measurement is not well understood
- Future work – lots of things to do!

# Acknowledgements

---

- Mark Crovella
- Vern Paxson, Anja Feldmann, Jim Pitkow, Drue Coles, Bob Carter, Erich Nahum, John Byers, Azer Bestavros, Lars Kellogg-Stedman, David Martin
- Xerox, Inc., EpicRealm Inc., Internet2
- Michael Mitzenmacher, Kihong Park, Carey Williamson, Virgilio Almeida, Martin Arlitt