



Measuring the Behavior of a WWW Server

David Yates

J. Almeida, V. Almeida, D. Yates

UFMG / Boston University

Outline

- Motivation
- Measurement tool - WebMonitor
- Workload - WebStone
- Results
 - Throughput and latency
 - Connection management overhead
- Conclusions

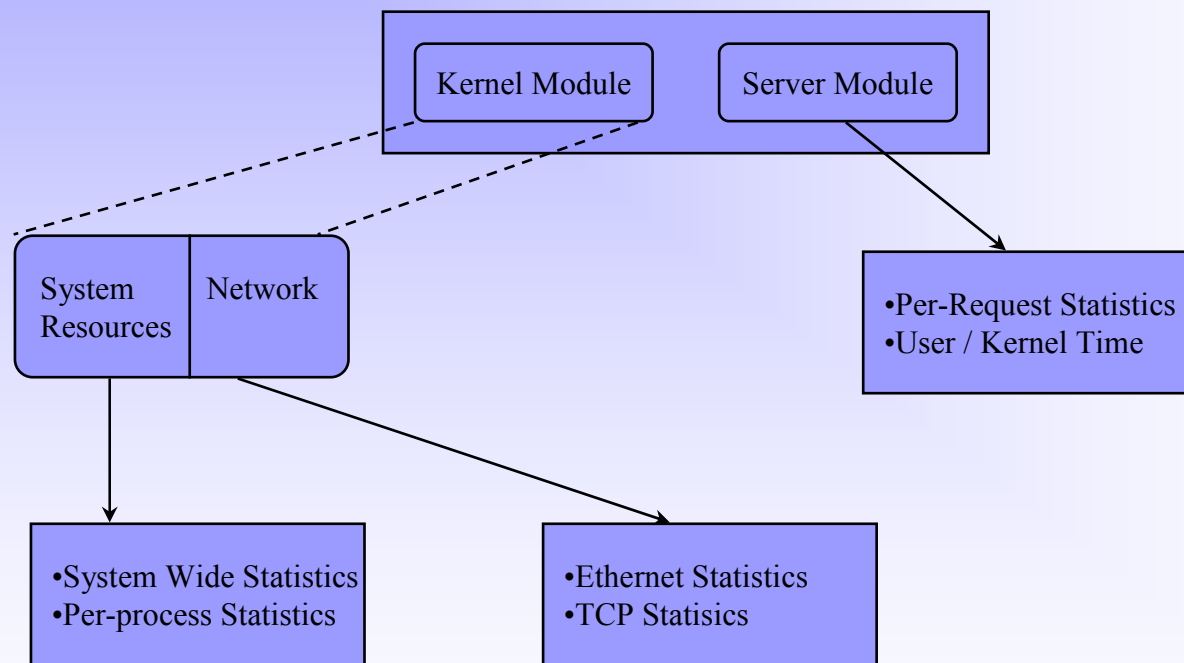


Motivation

- Servers on WWW are a bottleneck
- Don't understand bottlenecks within servers
- Our goals:
 - Measure and understand server behavior
 - Connection management overhead (isolate effect of keeping TCP connections open)



WebMonitor



WebMonitor

☞ Features

- Measures activity in the kernel and in user-space
- Low overhead (< 3%)

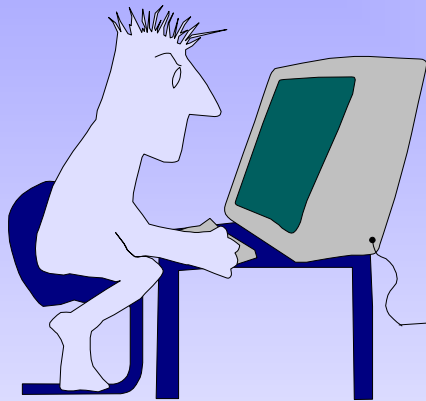
☞ Validation

- Kernel module against /proc
- Server module against WebStone



Experimental Testbed

Web Clients (30)

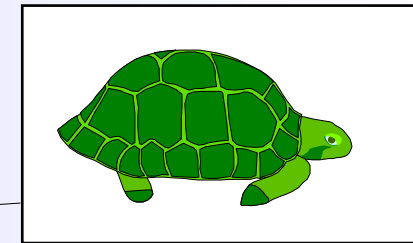


- Sun / Solaris
- WebStone

HTTP Requests



Web Server



- Pentium / Linux (and KM)
- Apache (and SM)

HTTP Responses



Heavy-tailed Workload

Type / Class	Files	Size (KB)	Access Prob.
Image & HTML 1	53	565	0.946
Audio 2	20	3580	0.050
Video 3	4	9216	0.004

Table of “Full Size” workload



Server Module Statistics

	Class 1	Class 2	Class 3
Conn/s	16.40	0.88	0.07
Mbits/s	1.55	1.17	1.19
Reads /conn	0.03	2.67	34.67
Latency /conn	954	4752	61413



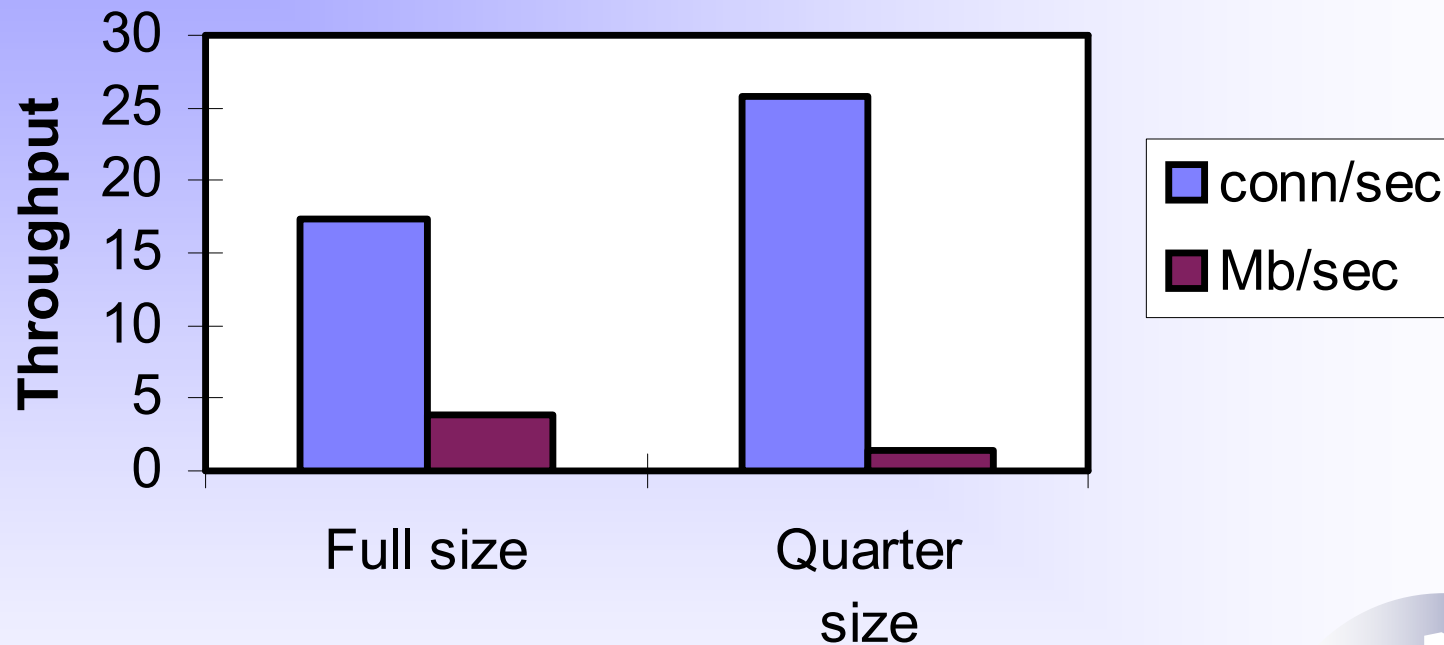
Kernel Module Statistics

CPU User (%)	9.00
CPU System (%)	90.65
Reads / sec	5.17
Writes / sec	4.90
Network + disk interrupts / sec	620 + 290

Over 90 % of the processing is done in the kernel



Throughput vs. File Size



Throughput does NOT scale with file size



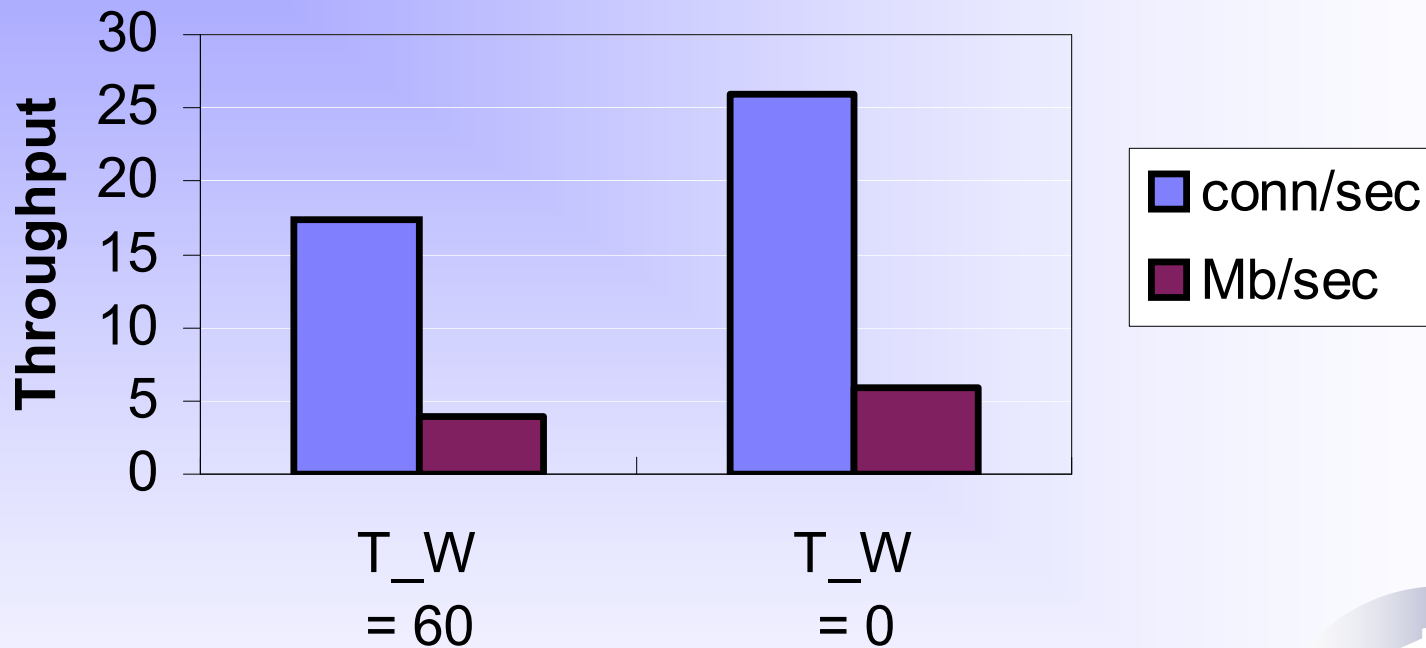
Keeping TCP Connections Open

Per-connection	Class 1	Class 2	Class 3
CPU time	24	111	1383
Latency ($T_W = 0$)	646	3085	37953
CPU time	29	162	2244
Latency ($T_W = 60$)	954	4752	61413

Significant increase in CPU time and elapsed time



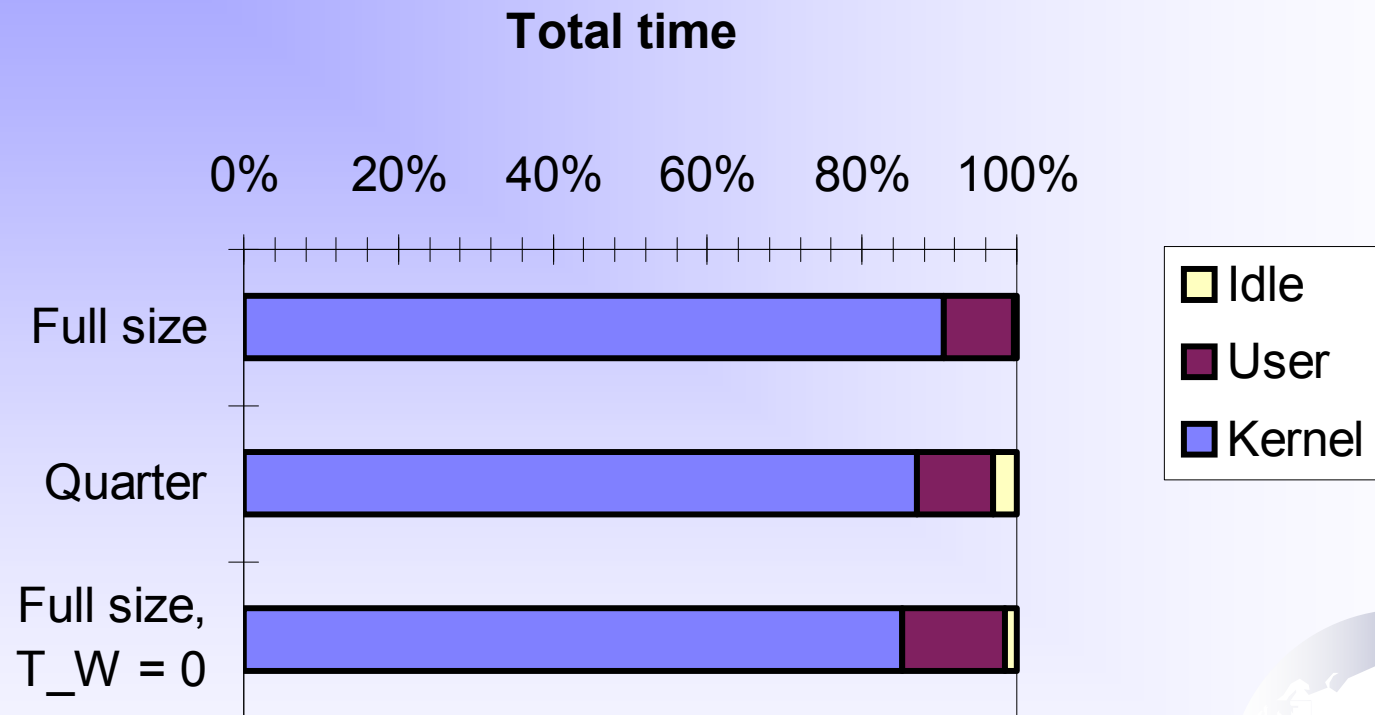
Keeping TCP Connections Open



TIME_WAIT time costs 33 % in performance



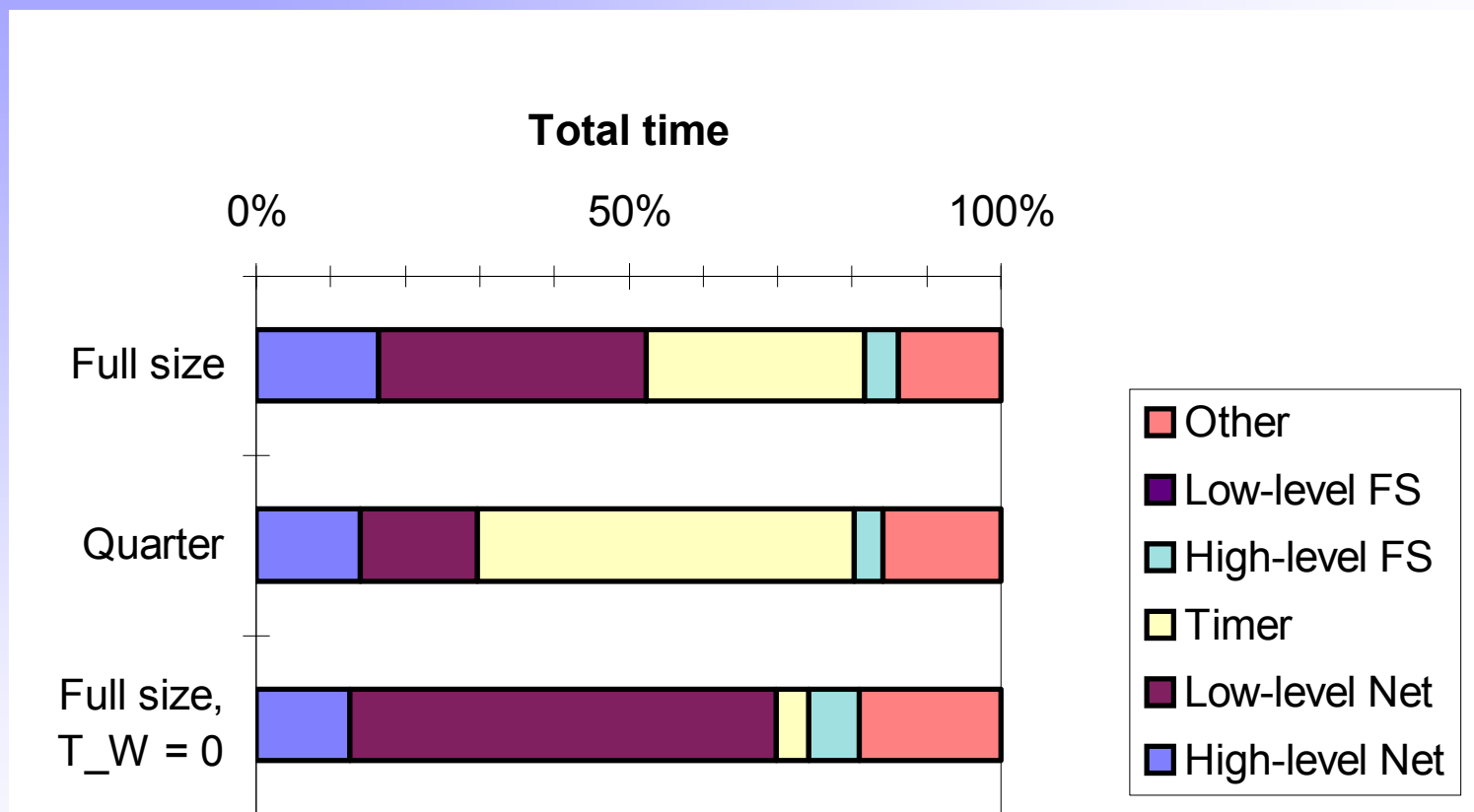
Where's the Time Spent?



In the kernel!



Where in the Kernel?



Conclusions

- WebMonitor and profiling are useful!
- For a saturated Web server ...
 - 90 % of the processing is done in kernel
- Connection management overhead ...
 - Increases CPU time and latency (up to 62 %)
 - Decreases throughput by 33 %
- What's the bottleneck for Web servers ...
 - CPU, memory, or I/O?

