



# BenchLab

An Open Testbed for Realistic  
Benchmarking of Web Applications

<http://lass.cs.umass.edu/projects/benchlab/>

*Emmanuel Cecchet, Veena Udayabhanu,  
Timothy Wood, Prashant Shenoy*

University of Massachusetts Amherst



# WEBAPPS YESTERDAY AND TODAY

The screenshot displays the Best Buy website interface. At the top, there are navigation links for 'Español', 'My Account', 'Order Status', 'Customer Service', and a welcome message. A shopping cart icon shows '0 Items'. The main navigation bar includes 'PRODUCTS', 'SERVICES', 'SHOPS & DEALS', and 'GIFTS'. A search bar is present with the text 'Search by Keyword, SKU # or Item #'. Below the navigation, a banner reads 'MAKE YOUR SUMMER SHINE with great prices on the latest devices. Shop now >'. The central promotional banner features the text 'THE MOST SMARTPHONES UNDER ONE ROOF\*' and 'Free shipping every day on these smartphones.' It includes a 'SHOP NOW >' button and logos for Verizon, at&t, Sprint, and T-Mobile. Below this, a section titled 'THIS WEEK'S OFFERS' lists three deals: a Toshiba Satellite laptop for \$299.99, 5%–15% off on select HDTVs, and 10%–40% off on select GPS receivers. To the right, a video player for 'FOSTER THE PEOPLE' is shown with a 'PLAY' button. At the bottom, there are small thumbnails for 'Backyard Getaway', 'Kung Fu Panda 2: Inside the Game', and 'inFAMOUS 2'.

Español | My Account | Order Status | Customer Service | Welcome. Please create an account or Sign in. CART 0 Items

Best Buy

Search by Keyword, SKU # or Item #

Store Locator Weekly Ad Credit Cards Reward Zone

PRODUCTS SERVICES SHOPS & DEALS GIFTS

MAKE YOUR SUMMER SHINE with great prices on the latest devices. Shop now >

**THE MOST SMARTPHONES UNDER ONE ROOF\***

Free shipping every day on these smartphones.

SHOP NOW >

verizon at&t Sprint T-Mobile

\*Based on a comparison to national retailers as of February 19, 2011.

Mobile phones

THIS WEEK'S OFFERS

See more featured offers >

Only \$299.99  
Toshiba Satellite laptop with Intel® Celeron® processor, 15.6" screen size, 3GB memory.  
shop now >

5%–15% Off  
select HDTVs.  
shop now >

10%–40% Off  
select GPS receivers. Plus free shipping.  
shop now >

**FOSTER THE PEOPLE** PLAY

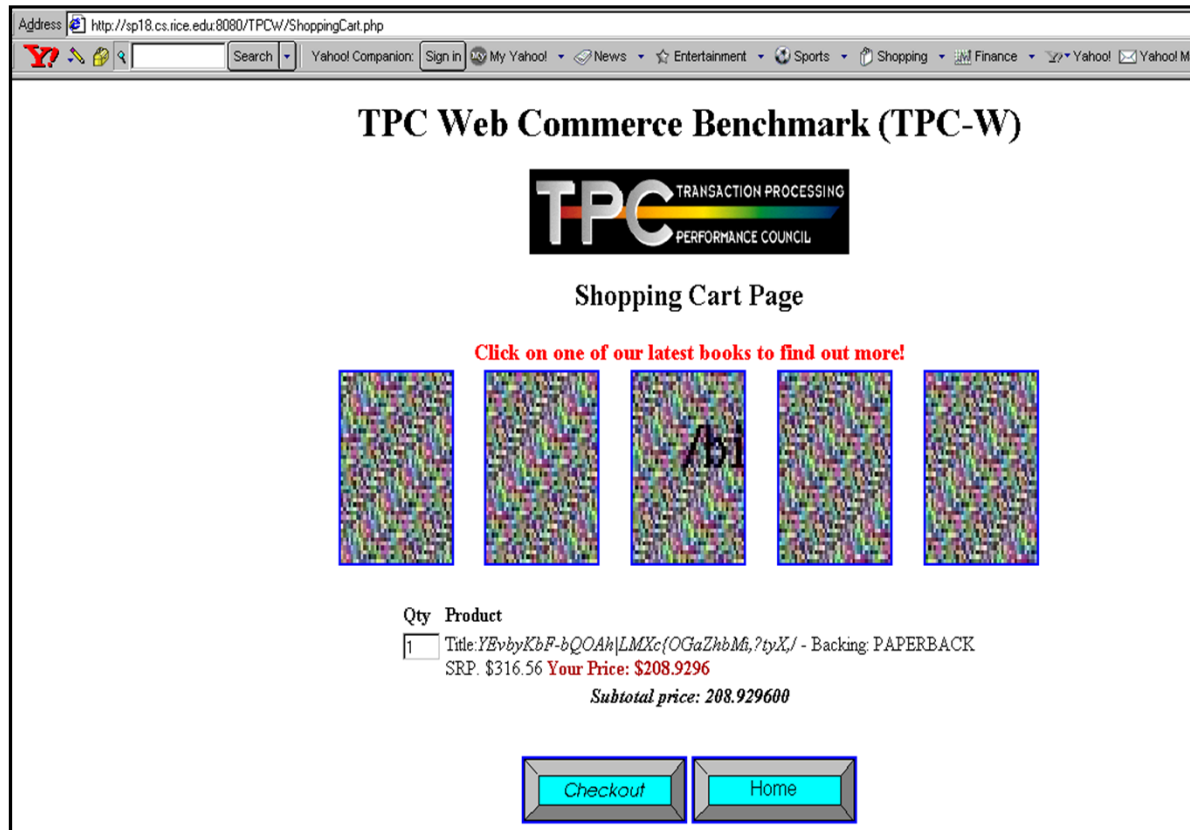
These guys started in the school band, now their first song is a viral hit.

Backyard Getaway Kung Fu Panda 2: Inside the Game Foster the People EXCLUSIVE

See more at Best Buy On >

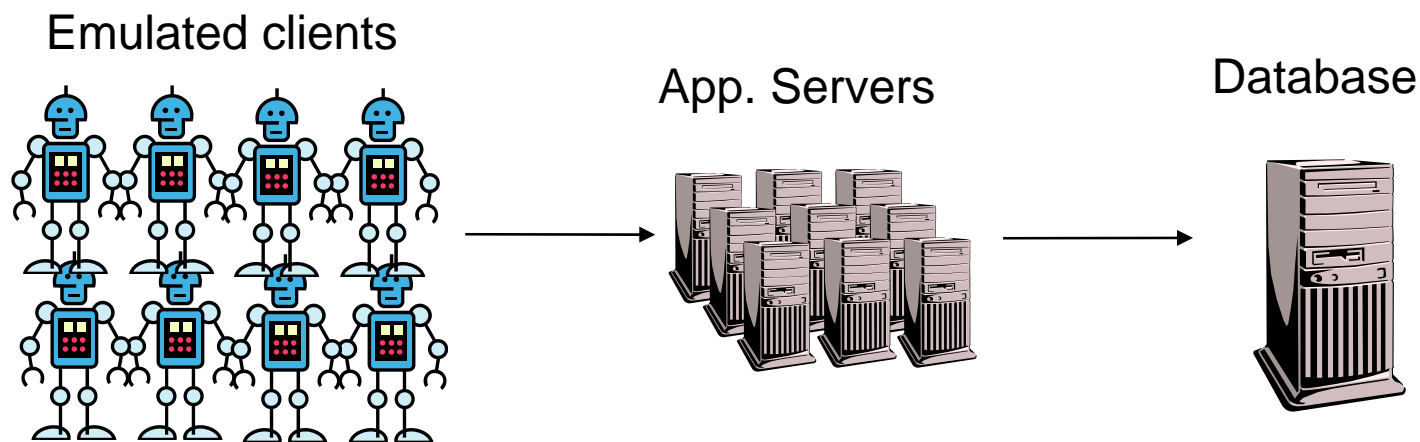
# TPC-W BENCHMARK

- Online bookstore emulating... amazon.com!
- Reference benchmark still in use today in the Systems community



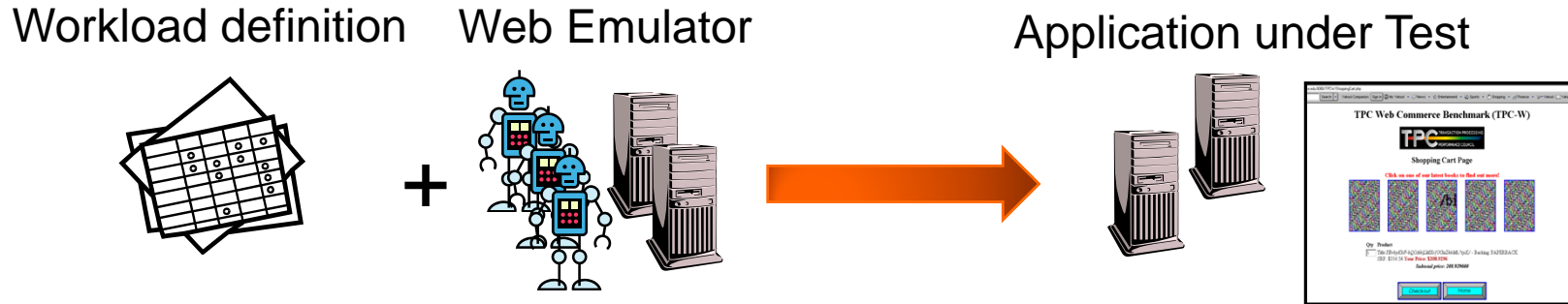
# TYPICAL E-COMMERCE BENCHMARK AVAILABLE TO THE RESEARCH COMMUNITY

- Setup for performance benchmarking
  - Browser emulator
  - Static load distribution
- Missing features
  - Content: Javascript, AJAX, cookies, ...
  - Network: HTTP 1.1 request pipelining, caching, SSL
  - Latencies: WAN, rendering time...



# BENCHMARK DESIGN

## *Traditional approach (TPC-W, RUBiS...)*



## *BenchLab approach*

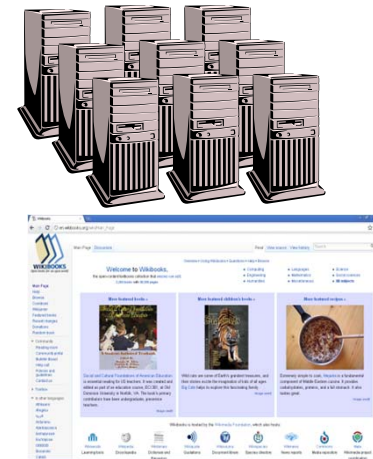
HTTP trace

```
http://...  
http://...  
http://...  
http://...  
http://...  
http://...
```

Real Web Browsers



Application under Test





# OUTLINE



- **What has changed in WebApps**
- Benchmarking real applications with BenchLab
- Experimental results
- Demo

# WEB APPLICATIONS HAVE CHANGED

- ❑ Web 2.0 applications
  - Rich client interactions (AJAX, JS...)
  - Multimedia content
  - Replication, caching...
  - Large databases (few GB to multiple TB)
- ❑ Complex Web interactions
  - HTTP 1.1, CSS, images, flash, HTML 5...
  - WAN latencies, caching, Content Delivery Networks...

# EVOLUTION OF WEB APPLICATIONS

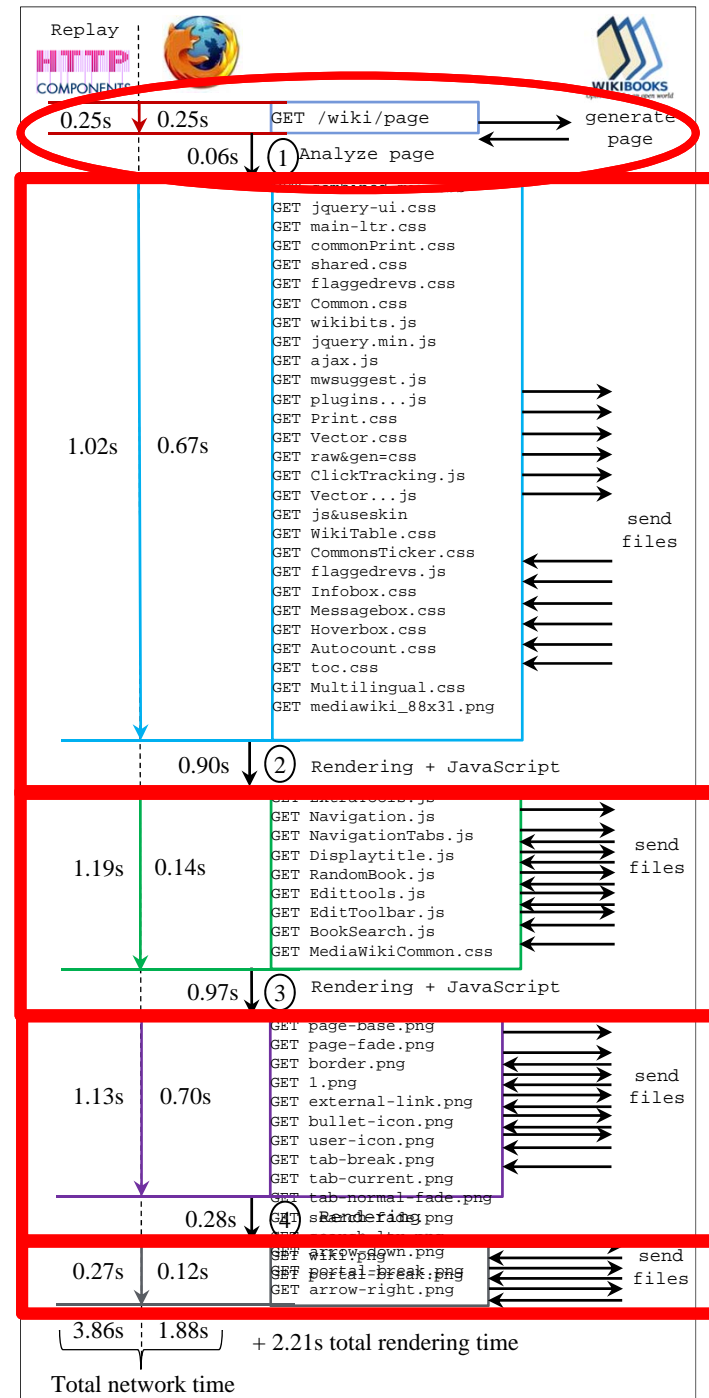
Applications	HTML	CSS	JS	Multimedia	Total
RUBiS	1	0	0	1	2
eBay.com	1	3	3	31	38
TPC-W	1	0	0	5	6
amazon.com	6	13	33	91	141
CloudStone	1	2	4	21	28
facebook.com	6	13	22	135	176
wikibooks.org	1	19	23	35	78
wikipedia.org	1	5	10	20	36

Number of interactions to fetch the home page of various web sites and benchmarks



# HTTP VS BROWSER REPLAY

- GET home page and receive HTML (0.25s)
- Parse HTML (0.06s)
- GET CSS and Javascript on 6 connections in parallel (Firefox)
- Starts rendering and execute Javascript
- GET additional Javascript imports
- Complete Javascript and continue rendering
- Download images
- More rendering
- Download final images
- Browsers are smart
  - Caching, prefetching, parallelism...
  - Javascript can trigger additional requests
- HTTP replay cannot approximate real Web browser access to resources



# TYPING SPEED MATTERS

- Auto-completion in search fields is common
- Each keystroke can generate a query
- Text searches use a lot of resources

```
GET /api.php?action=opensearch&search=W
GET /api.php?action=opensearch&search=Web
GET /api.php?action=opensearch&search=Web+
GET /api.php?action=opensearch&search=Web+2
GET /api.php?action=opensearch&search=Web+2.
GET /api.php?action=opensearch&search=Web+2.0
```

# STATE SIZE MATTERS

- Does the entire DB of Amazon or eBay fit in the memory of a cell phone?
  - TPC-W DB size: 684MB
  - RUBiS DB size: 1022MB
- Impact of CloudStone database size on performance

Dataset size	State size (in GB)	Database rows	Avg cpu load with 25 users
25 users	3.2	173745	8%
100 users	12	655344	10%
200 users	22	1151590	16%
400 users	38	1703262	41%
500 users	44	1891242	45%

**CloudStone Web application server load observed for various dataset sizes using a workload trace of 25 users replayed with Apache HttpClient 3.**

# OUTLINE

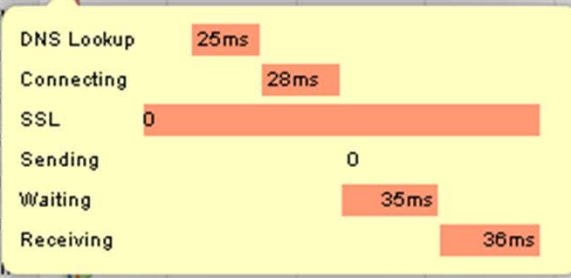


- What has changed in WebApps
- **Benchmarking real applications with BenchLab**
- Experimental results
- Demo

# BENCHLAB

- Capture application workloads
- Replay captured traces in real Web browsers
- Store detailed performance results

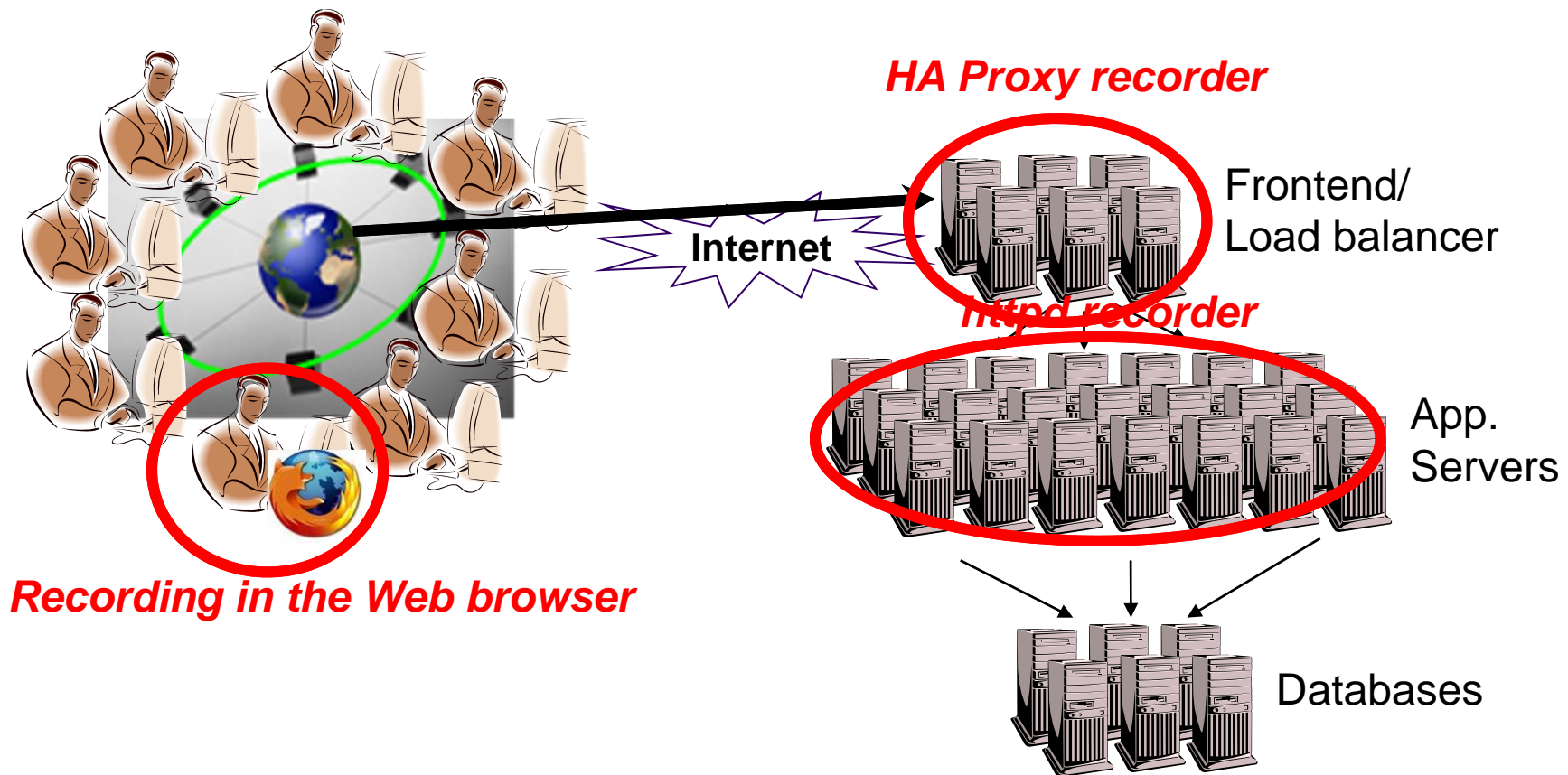
Name	Method	Status	Type	Size	Time	Timeline	7.82s	11.72s	15.63s
site-wide-5674991483.css_V183552863...	GET	200	text/css	30.15KB	141ms	36ms			
websiteGridCSS-websiteGridCSS-358...	GET	200	text/css	21.36KB	148ms				
navPackedSprites-US-22_V183711641_...	GET	200	image/png	7.11KB	200ms				
transparent-pixel_V192234675_gif	GET	(from cache)	image/gif	(from cache)	86ms				
site-wide-13640146130.js_V183531323...	GET	200	application/x-javascript	133.61KB	403ms				
www.amazon.com	GET	Pending	Pending	0B	10ms				
51skuPkUrxL_SL500_Plsitb-sticker-ar...	GET	200	image/jpeg	7.10KB	170ms				
s9-widget-combined-min_V183012506...	GET	200	text/css	8.65KB	127ms				



- Benchmark repository
  - Store virtual machines of applications under test
  - Store test traces, configurations and results
  - Repeat experiments
  - Compare results

# RECORDING HTTP TRACES

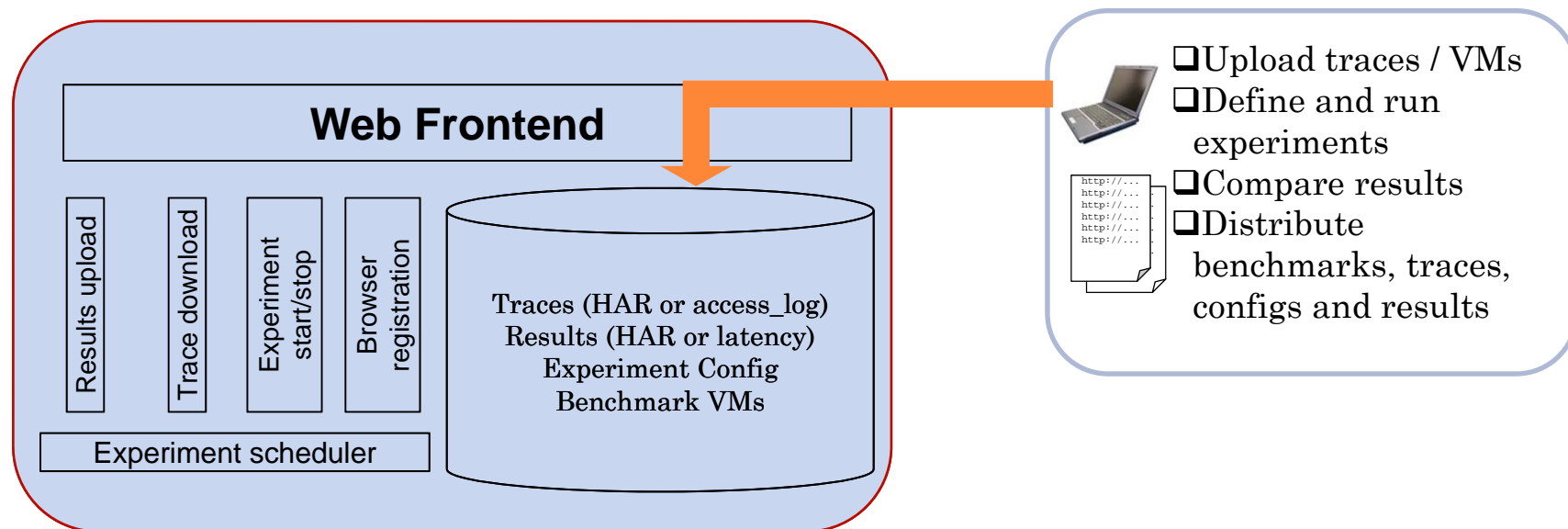
- 3 options to record traces in HTTP Archive (HAR) format
  - directly in Web browser
  - at HA proxy load balancer level
  - using Apache httpd logs





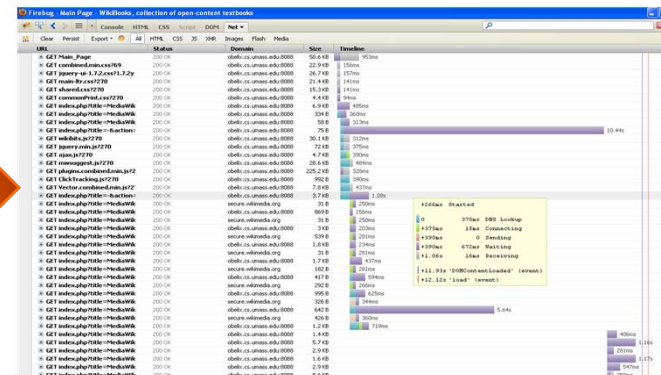
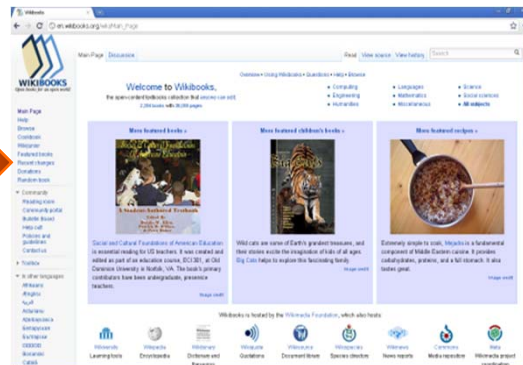
# BENCHLAB WEBAPP

- JEE WebApp with embedded database
- Repository of benchmarks and traces
- Schedule and control experiment execution
- Results repository
- *Can be used to distribute / reproduce experiments and compare results*



# BENCHLAB CLIENT RUNTIME (BCR)

- Replay traces in real Web browsers
- Small Java runtime based on Selenium/WebDriver
- Collect detailed response times in HAR format
- Can record HTML and page snapshots
- Upload results to BenchLab WebApp when done



Web page browsing  
and rendering

HAR results

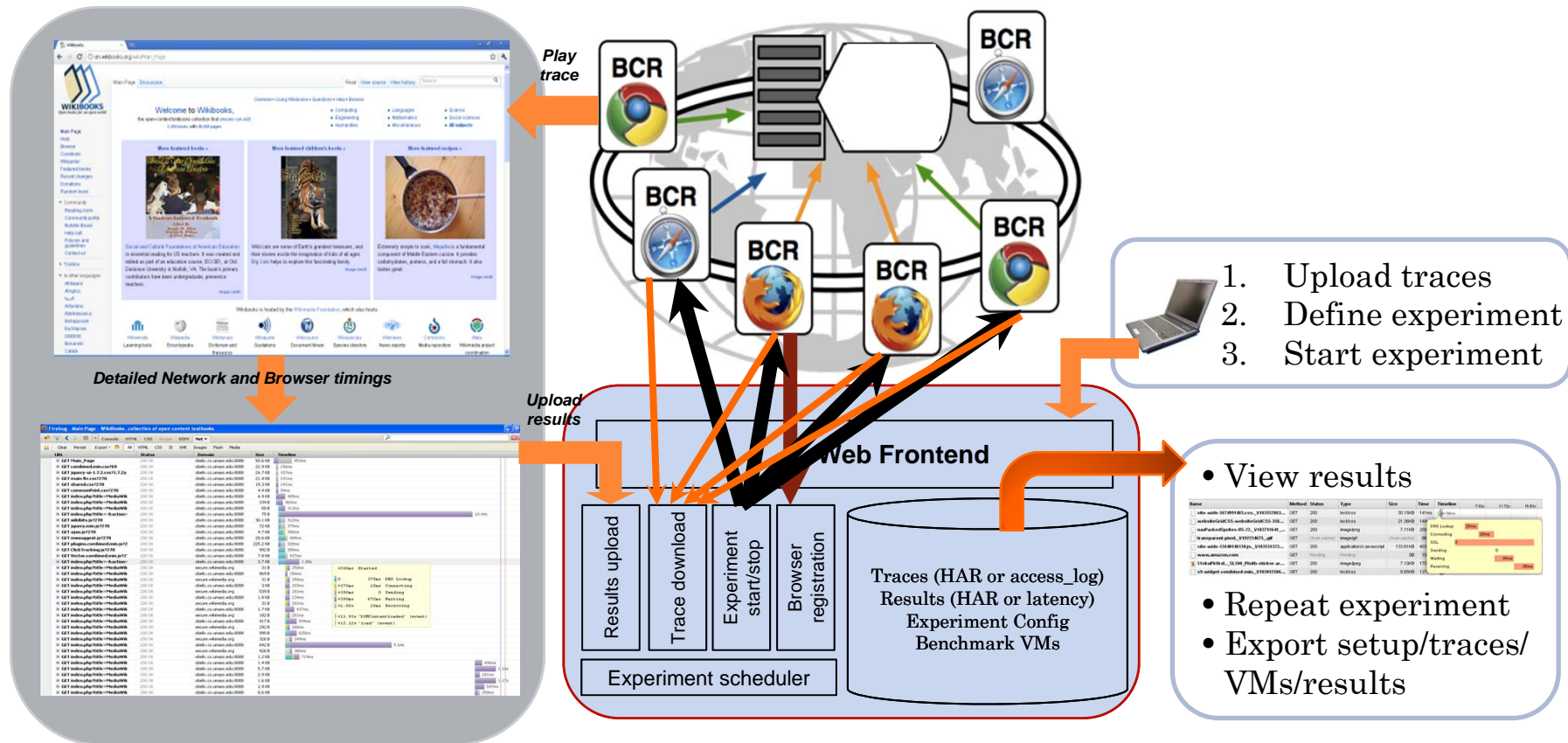
# WIKIMEDIA FOUNDATION WIKIS



- Wikimedia Wiki open source software stack
  - Lots of extensions
  - Very complex to setup/install
- Real database dumps (up to 6TB)
  - 3 months to create a dump
  - 3 years to restore with default tools
- Multimedia content
  - Images, audio, video
  - Generators (dynamic or static) to avoid copyright issues
- Real Web traces from Wikimedia
- Packaged as Virtual Appliances

# RUNNING AN EXPERIMENT WITH BENCHLAB

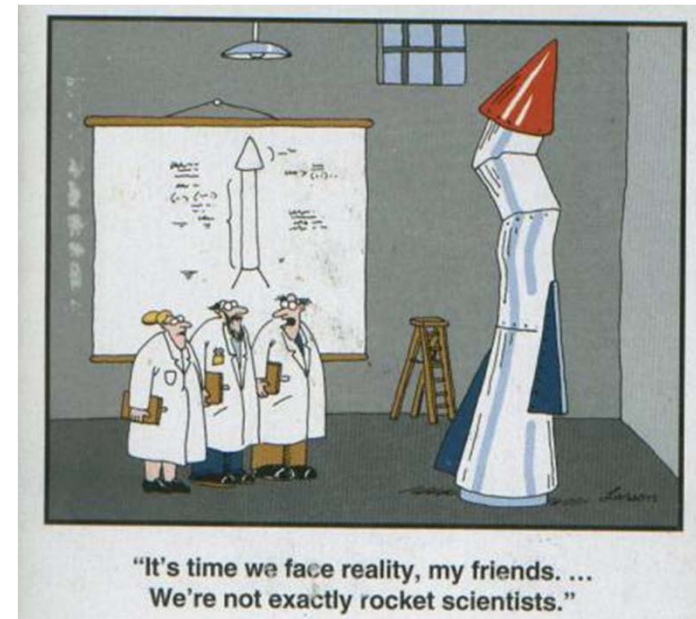
- Use your own deployment framework and monitoring
- Real browser replay
  - Full rendering with AJAX, JS, multimedia...
  - Emulate human typing speed while filling form, click buttons...
- Analyze, compare & share results



# OUTLINE



- What has changed in WebApps
- Benchmarking real applications with BenchLab
- **Experimental results**
- Demo

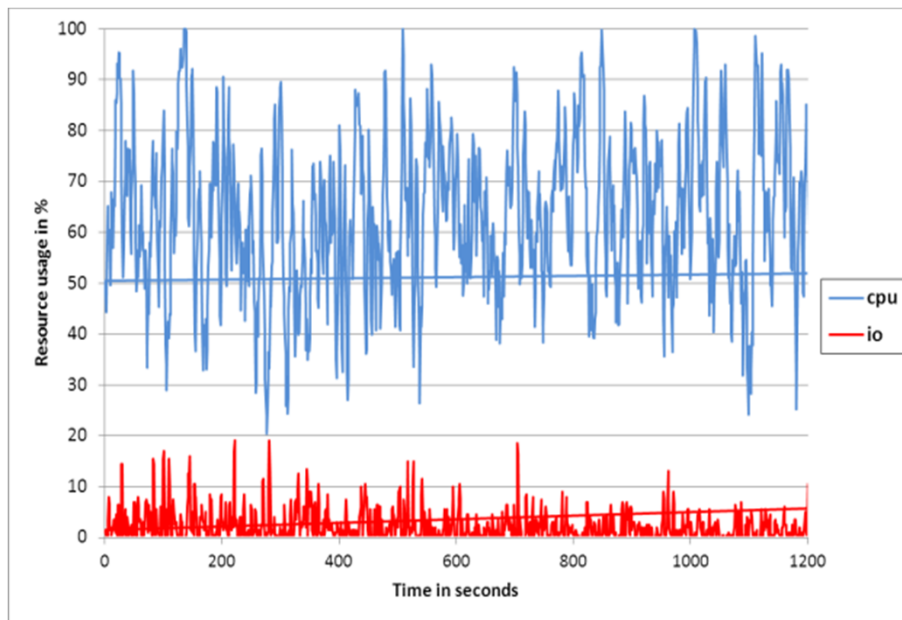




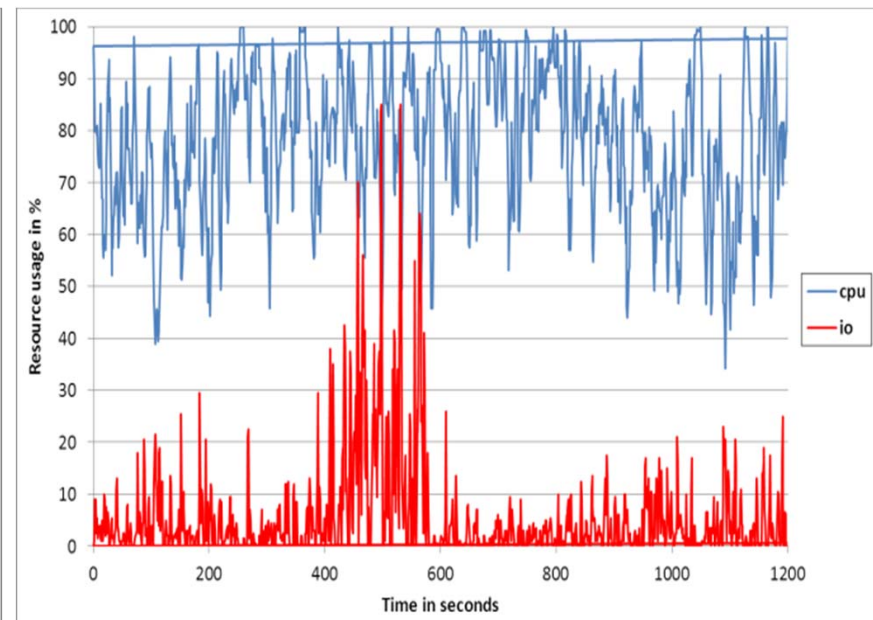
# EMULATED VS REAL BROWSER

- CPU and IO usage varies greatly on server for same workload
- CloudStone with 25 users workload from Amazon EC2 East Coast to Umass Amherst
  - Avg CPU emulated: 63.2%
  - Avg CPU real browser: 77.7%
- Resource access pattern affects file caching on server

Emulated browser



Real browser





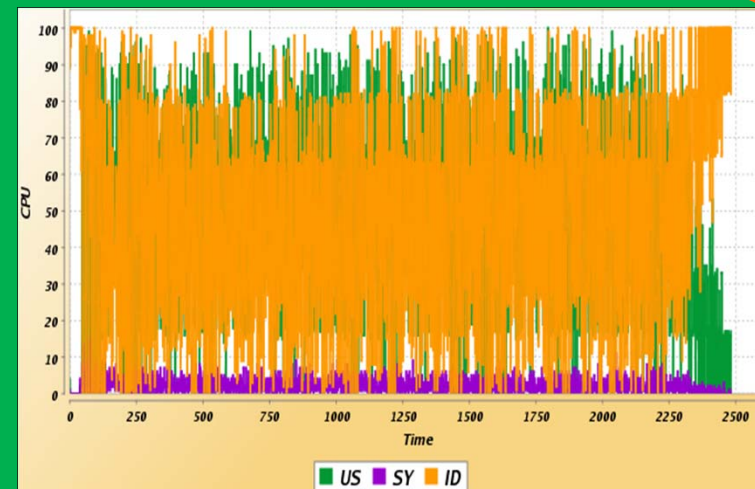
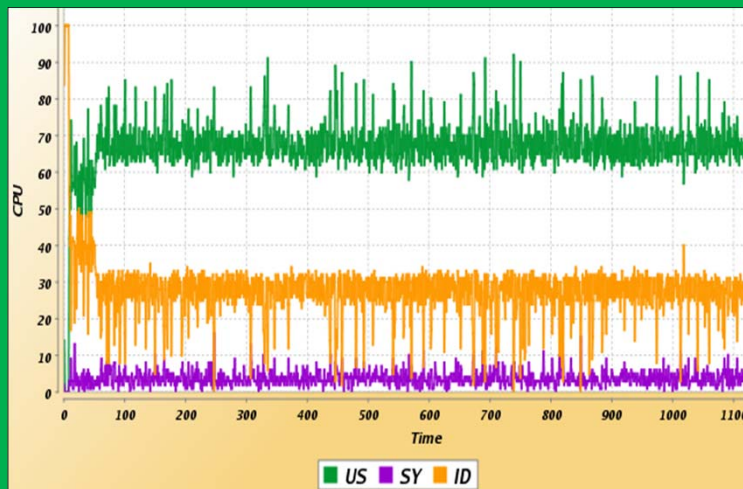
# JAVASCRIPT EFFECTS ON WORKLOAD

- Server side resource usage varies greatly
- Additional queries during form processing

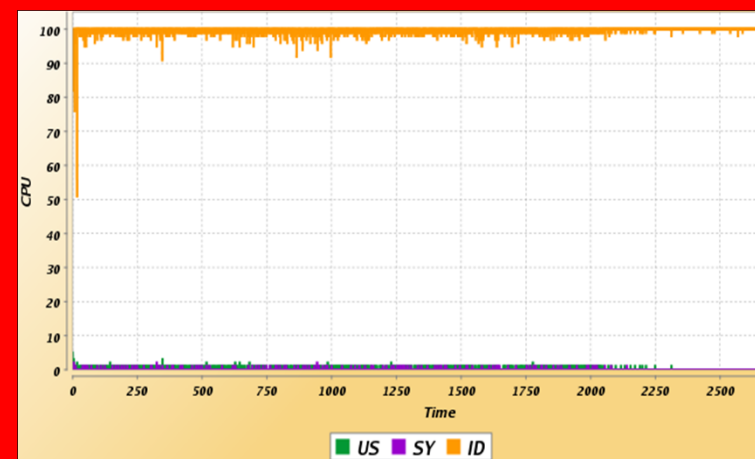
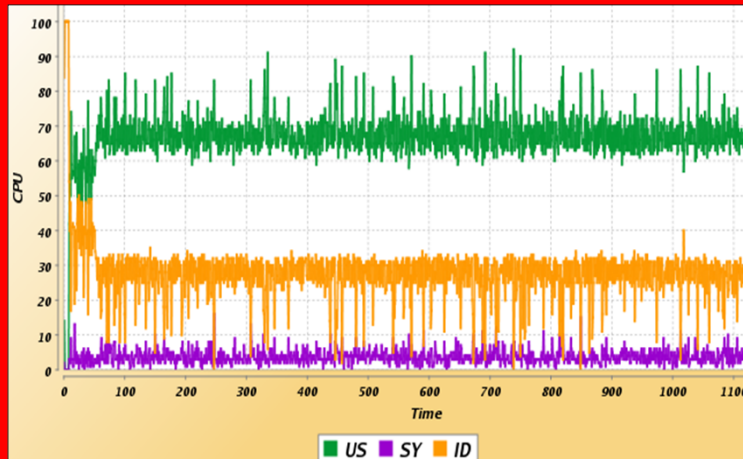
Emulated Browser

Real Browser

Good Input

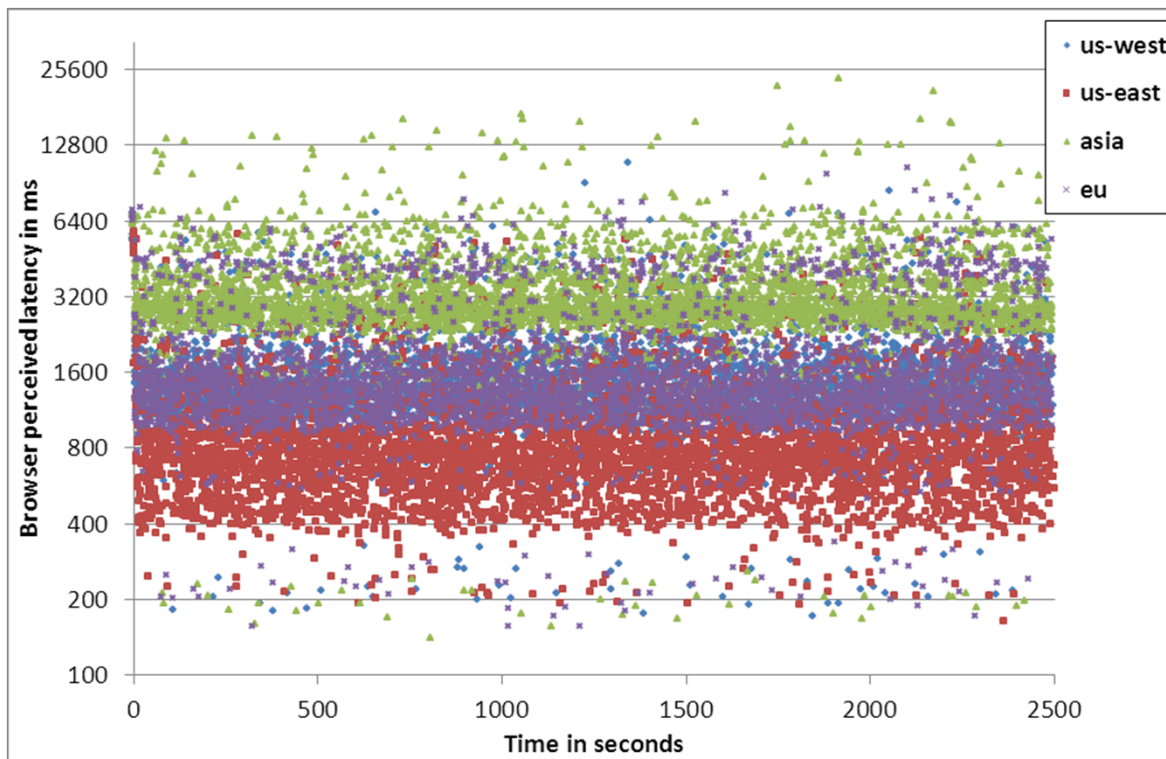


Bad Input



# LAN VS WAN LOAD INJECTION

- Deployed BCR instances in Amazon EC2 data centers
  - As little as \$0.59/hour for 25 instances for Linux
  - Windows from \$0.84 to \$3/hour
- CPU usage varies greatly on server for same workload (LAN 38.3% vs WAN 54.4%)



# CONCLUSION

- Web Applications have changed
- Real browsers needed for modern WebApp benchmarking
- BenchLab provides
  - Infrastructure for Internet scale Benchmarking of real applications
  - Virtual Appliances of real applications
  - Repository of traces, benchmarks and results
- A lot to explore...



# Q&A

COME AND SEE OUR **POSTER** AND **DEMO**

<http://lass.cs.umass.edu/projects/benchlab/>

DILBERT™ by Scott Adams

