Specification and Implementation of Dynamic Web Site Benchmarks

Sameh Elnikety
Department of Computer Science
Rice University
Dynamic Content Is Common

1. THINKPAD A22M P3-1GHZ 30GB 128MB
   - IBM
   - Our Price: $1,740.95
   - Type: System
   - Platform: Pentium III
   - Mfg part#: 2620TTU
   - Quantity Remaining: 96

2. Market Summary
   - Dow: 9863.74, -21.04 (-0.21%)
   - Nasdaq: 1834.21, -12.45 (-0.67%)
   - S&P 500: 1107.50, -4.44 (-0.40%)
   - 10-Yr Bond: 4.911%, +0.012
   - NYSE Volume: 1,090,619,000
   - Nasdaq Volume: 1,621,515,000

3. Amazon.com
   - Welcome to Amazon
   - Your Shopping Cart
   - Valentine's Day
     - Roses for $39.99, romantic books, and more
     - Order flowers, find gifts
Generating Dynamic Content

http

Web Server

Dynamic Content Generator

Database Server
DynaServer Project

- Scale Performance of Dynamic Content Sites
  - Caching
  - Clustering
  - Replication
  - Dynamic content delivery networks
Motivation

• We Need Workloads to Evaluate Different Architectures

• No Workloads for Dynamic Content
Benchmarks

- **Online Bookstore**
  - Models amazon.com

- **Auction Site**
  - Models ebay.com

- **Bulletin Board**
  - Models slashdot.org

Common Applications

Stress Different Parts of the System
Contributions

• **Online Bookstore**
  – Followed TPC-W specification
  – Wrote implementation

• **Auction Site**
  – Developed specification & implementation

• **Bulletin Board**
  – Developed specification & implementation
Outline

• Benchmarks
  – Online Bookstore
  – Auction Site
  – Bulletin Board

• Bottleneck Characterization
  – Experimental Setup
  – Measurement
  – Performance Results
Online Bookstore Benchmark

• Models amazon.com
• Follows TPC-W (TPC Feb 2000)

• Activities
  – Browsing, buying, administrative updates

• Persistent Data
  – Images in Web server file system
  – All others in database
Online Bookstore Benchmark

TPC Web Commerce Benchmark (TPC-W)

Shopping Cart Page

Click on one of our latest books to find out more!

Title: YbwykJF-JQoAh]LMXc(OGaZhhMh,?tX/- Backing: PAPERBACK
SRP: $316.56 Your Price: $208.9296
Subtotal price: 208.929600

Quantity: 1

Image
Promotion
Shopping Cart
Next Interaction
Online Bookstore Benchmark

• 14 Interactions (as specified in TPC-W), e.g.
  – Home
  – Best sellers
  – Secure payment
  – Shopping cart

• Workload Mixes (as specified in TPC-W)
  – Browsing (95% read-only)
  – Shopping (80% read-only)
  – Ordering (50% read-only)
Auction Site Benchmark

- Models Ebay.com
- Activities
  - Browsing, selling, bidding
- User Sessions
  - Visitor, seller, buyer
- Persistent Data
  - Images in Web server file system
  - All others in database
Auction Site Benchmark

• 26 interactions
  – Browsing by category
  – Bidding
  – Buying
  – Selling
  – Leaving comments
  – User page

• Workload Mixes
  – Browsing (100% read-only)
  – Bidding (85% read-only)
Bulletin Board Benchmark

- Models Slashdot.org

- Activities
  - Discussion thread (story & comments)
  - Ratings

- User sessions
  - Visitor, moderator

- Persistent Data
  - Images in Web server file system
  - All others in database
Bulletin Board Benchmark

• 24 interactions
  – Browsing by category
  – Submit new story
  – Search story titles
  – Display story & filter comments

• No Full Text Search

• Workload Mixes
  – Browsing (100% read-only)
  – Submission (85% read-only)
Bottleneck Characterization

- Experimental Setup
- Measurement
- Performance Results
Software

http

Web Server

Dynamic Content Generator

Database Server

http

Apache PHP

sql

MySQL
Hardware

<table>
<thead>
<tr>
<th></th>
<th>CPU</th>
<th>Memory</th>
<th>Disk</th>
<th>Network</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AMD Athlon 1.33GHz</td>
<td>768MB</td>
<td>60GB, 9msec, 5400rpm</td>
<td>100Mbps Ethernet</td>
</tr>
</tbody>
</table>

Apache

http

MySQL

sql
Performance Metric

- Throughput
  - Number of interactions per minute
Resources Monitored

- Web Server
- Database
- CPU
- Memory
- Disk
- Network
Emulated Clients

- Java Emulator
  - Session duration
  - Think time
  - Markov model
Measurement

• Driving the System
  – Several clients
  – Ramp-up time
  – Steady-state measurement
Online Bookstore Benchmark Throughput
Bookstore CPU Utilization (Shopping Mix)
<table>
<thead>
<tr>
<th>Resource</th>
<th>Web Server</th>
<th>Database</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU</td>
<td>30%</td>
<td>71%</td>
</tr>
<tr>
<td>Memory</td>
<td>200 MB</td>
<td>390 MB</td>
</tr>
<tr>
<td>Disk</td>
<td>&lt;20 block/s</td>
<td>&lt;20 block/s</td>
</tr>
<tr>
<td>Network</td>
<td>3.2 + 0.6 Mb/s</td>
<td>0.6 Mb/s</td>
</tr>
</tbody>
</table>

**Conclusion** Database CPU is the bottleneck
Auction Site Benchmark Throughput

![Graph showing the benchmark throughput for browsing and bidding with respect to the number of clients.](chart)
Auction CPU Utilization (Bidding Mix)
## Auction Resources (Bidding Mix)

<table>
<thead>
<tr>
<th>Resource</th>
<th>Web Server</th>
<th>Database</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU</td>
<td>100%</td>
<td>58%</td>
</tr>
<tr>
<td>Memory</td>
<td>70 MB</td>
<td>250 MB</td>
</tr>
<tr>
<td>Disk</td>
<td>&lt;20 block/s</td>
<td>&lt;20 block/s</td>
</tr>
<tr>
<td>Network</td>
<td>55 + 2 Mb/s</td>
<td>2 Mb/s</td>
</tr>
</tbody>
</table>

### Conclusion
Web Server CPU is the bottleneck
Bottlenecks

- **Online Bookstore**
  - Models amazon.com

- **Bidding**
  - Models ebay.com

- **Bulletin Boards**
  - Models slashdot.org

- CPU of Database
- CPU of Web Server
Conclusions

• 3 Benchmarks
  – Online bookstore, auction site, bulletin board
  – Open-source software
  – Available implementations (PHP, Servlets, EJB)

• Capacity Planning

• Research on Dynamic Content
Collaborators

Cristiana Amza, Anupam Chanda, Alan Cox,
Romer Gil, Willy Zwaenepoel
Department of Computer Science - Rice University

Emmanuel Cecchet, Julie Marguerite
INRIA Rhône - Alpes

Karthick Rajamani
IBM Austin Research Lab
Download the Benchmarks

Benchmark Implementations in

- PHP
- Servlets
- EJB