Visual Application Usage Modeling

First Presented for:

Mountain View, CA 4/24/2006

Scott Barber
Chief Technologist
PerfTestPlus, Inc.
Why use \textit{calculus} when all you need are \textit{Crayons}?
Background
The problem...

Application Usage Models tend to be either:

Rigorous

- Time consuming
- Mathematically intensive and/or complex
- High degree of accuracy (when done well)
- Requires empirical data
- Typically overkill

Overly Simplistic

- Quick
- Little to no math needed
- Occasionally accurate (generally by accident)
- Ignores empirical data

There is very little in between to assist a modeler in an industry that desires rigor, but barely has time for simple!
In Practice:

Empirical data is uncommon
Complex math skills are rare
Time is not a luxury we often have
Every sub-team uses different modeling languages
Models are not often shared with or understood by other teams
Few modeling tools and methods are easily available
Most usage models are little more than “semi-educated” guesses
A lot of testing is based on application usage models
Rigorous Techniques to Model Application Usage

Connie U. Smith, PhD. - *Performance Solutions: A Practical Guide to Creating Responsible, Scalable Software*

Alberto Savoia - “Web Load Test Planning: Predicting how your Web site will respond to stress"

Daniel Menasce, PhD. – *Capacity Planning for Web Performance: Metrics, Models and Methods & Scaling for E-Business*

J.D. Meier - *Improving .NET Application Performance and Scalability*

All require empirical data, two require advanced mathematics, none are intuitive to untrained individuals.
“What sense does it make to measure with a laser, mark with chalk and cut with an axe?!?”

- Bob Barber, industrial arts teacher, middle school guidance counselor & my dad.
Summary

Rigorous approaches are frequently impractical

How much value does a model have if only one person understands it?

Why spend the time modeling if the resulting model isn’t useful?
What the experts say
“All models are wrong; some models are useful.”

-George Box, Industrial Statistician.

Dr. Box is a Vilas Professor, the highest honor awarded to faculty by the University of Wisconsin. Dr. Box is concerned with the planning and analysis of industrial experiments. The object is to determine the important factors affecting product quality then adjust them to their best levels. Recent work has concentrated on reduction of variance, as well as adjustment of mean levels. New Bayesian methods of analysis have been devised for highly fractionated designs. Studies of reduction of variance transmissions are in progress.

http://www.engr.wisc.edu/ie/faculty/box_george.html
On Visualization…

“Visual representations of evidence should be governed by principles of reasoning about quantitative evidence. Clear and precise seeing becomes as one with clear and precise thinking.”

-Edward Tufte, data visualization expert.

Edward Tufte has written seven books, including Visual Explanations, Envisioning Information, The Visual Display of Quantitative Information, and Data Analysis for Politics and Policy. He writes, designs, and self-publishes his books on information design, which have received more than 40 awards for content and design. He is Professor Emeritus at Yale University, where he taught courses in statistical evidence, information design, and interface design. His current work includes digital video, sculpture, printmaking, and a new book Beautiful Evidence.

http://www.edwardtufte.com
On Application Usage Models…

“If… your simulated load is realistic and you can draw useful conclusions. **If not, you are wasting your time and money.**”

- Alberto Savoia, CTO Agitar, Inc (Former Engineering Executive of Google)

Before Agitar, Alberto worked at Google as the engineering executive in charge of the highly successful and profitable ads group. In October 1998, he cofounded and became CTO of Velogic Inc., the pioneer and leading innovator in Internet performance and scalability testing. Velogic was acquired in 2000 by Keynote Systems Inc. (NASD:KEYN), where Alberto continued to serve as chief technologist. Prior to Velogic, Alberto had an impressive 13-year career at Sun Microsystems. His last position at Sun was general manager of SunTest, a highly successful business unit he created and established as the technology and market leader in Java testing solutions. Before SunTest, he served Sun as director of the software technology research group for Sun Microsystems Laboratories, where his staff developed mission-critical and groundbreaking new technology in the areas of software development environments, programming languages, and clustering.

http://www.agitar.com
Summary

Models need to be useful
A (good) picture is worth a thousand words
Inaccurate models lead to wasted time and money
Usage models we are (unfortunately) used to seeing
Logs from Technical Administrators

Logs from Technical Administrators
Parsed Logs from Less Technical Administrators

Reports: www.perftestplus.com

Directory and Pages

Top 10 Clicks From

<table>
<thead>
<tr>
<th>Directory</th>
<th>Clicks</th>
<th>Clicks %</th>
<th>G1/Clicks Avg. Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>/aboutpc.htm</td>
<td>233</td>
<td>65.27%</td>
<td>0.00%</td>
</tr>
<tr>
<td>/acceptance.htm</td>
<td>45</td>
<td>12.61%</td>
<td>0.00%</td>
</tr>
<tr>
<td>/advising.htm</td>
<td>29</td>
<td>8.12%</td>
<td>0.00%</td>
</tr>
<tr>
<td>/advisors.htm</td>
<td>18</td>
<td>5.04%</td>
<td>0.00%</td>
</tr>
<tr>
<td>/affiliations.htm</td>
<td>18</td>
<td>5.04%</td>
<td>0.00%</td>
</tr>
<tr>
<td>/alliances.htm</td>
<td>9</td>
<td>2.52%</td>
<td>0.00%</td>
</tr>
<tr>
<td>/analyze.htm</td>
<td>4</td>
<td>1.12%</td>
<td>0.00%</td>
</tr>
<tr>
<td>/approach.htm</td>
<td>3</td>
<td>0.84%</td>
<td>0.00%</td>
</tr>
<tr>
<td>/architecture.htm</td>
<td>3</td>
<td>0.84%</td>
<td>0.00%</td>
</tr>
<tr>
<td>/contact_us.htm</td>
<td>1</td>
<td>0.28%</td>
<td>0.00%</td>
</tr>
</tbody>
</table>

Top 10 Clicks To

<table>
<thead>
<tr>
<th>Directory</th>
<th>Clicks</th>
<th>Clicks %</th>
<th>G1/Clicks Avg. Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>/www.perftestplus.com/index.htm</td>
<td>136</td>
<td>38.10%</td>
<td>0.00%</td>
</tr>
<tr>
<td>/load.htm</td>
<td>63</td>
<td>17.65%</td>
<td>0.00%</td>
</tr>
<tr>
<td>/www.perftestplus.com/index.htm</td>
<td>45</td>
<td>12.61%</td>
<td>0.00%</td>
</tr>
<tr>
<td>/presentations.htm</td>
<td>23</td>
<td>6.44%</td>
<td>0.00%</td>
</tr>
<tr>
<td>/pubs.htm</td>
<td>19</td>
<td>5.32%</td>
<td>0.00%</td>
</tr>
<tr>
<td>/resources.htm</td>
<td>10</td>
<td>2.80%</td>
<td>0.00%</td>
</tr>
<tr>
<td>/training.htm</td>
<td>10</td>
<td>2.80%</td>
<td>0.00%</td>
</tr>
<tr>
<td>/contact_us.htm</td>
<td>8</td>
<td>2.24%</td>
<td>0.00%</td>
</tr>
<tr>
<td>/tech_leadership.htm</td>
<td>7</td>
<td>1.96%</td>
<td>0.00%</td>
</tr>
</tbody>
</table>
UML from Analysts

Hotel Reservations System

Make provisional reservation <<extends>> Accommodation rejected

Make provisional reservation <<extends>> Terms rejected

Make provisional reservation <<extends>> Confirmed immediately

Customer confirms reservation <<uses>> Confirm reservation

Customer confirms reservation <<uses>> Run management reports

Check in guest

Check out guest

Customer requests cancellation <<uses>> Cancel unconfirmed reservation

Customer requests cancellation <<uses>> Cancel reservation

Check in guest

Check out guest

Manager

Receptionist


State Transition Models from Developers

### Diagram:

- **Customer**
  - Start
  - Insert card
  - Enter pin
  - Enter amount
  - Take money from slot
  - Take card

- **ATM Machine**
  - Authorize
    - Branch
      - Valid PIN
      - Invalid PIN
    - Check account balance
      - [balance >= amount]
      - [balance < amount]
    - Debit account
    - Show balance
    - Merge
    - Eject card

- **Bank**
  - Guard expression

**Activities**:
- Start
- Insert card
- Enter pin
- Enter amount
- Take money from slot
- Take card
- Authorize
- Branch
- Valid PIN
- Invalid PIN
- Check account balance
- Debit account
- Show balance
- Merge
- Eject card
“Process Model

All content submitted into the CMS will automatically follow the appropriate workflow in MyTool. A variety of workflows can be generated to mandate the path of information based on the type of data, the role of the creator, and the location within the site. Content providers will be empowered to make changes through the pre-formatted templates that they can access through a standard Web browser. The templates clearly separate the design elements of the site from the text and editable portions. Content providers can edit and submit their changes for approval via an automated workflow cycle, which circulates the proposed changes to authorized reviewers for approval or edits. Once this cycle is complete, the changes can be automatically posted to the site. This reduces the need for the Web team to be involved in every site change…”
Complex Mathematics from Academics

This relationship is called the Service Demand Law, which can also be written as \( D_i = V_i \times S_i \), by definition of the service demand (and since \( D_i = U_i/X_0 = (B_i/T)(C_0/T) = B_i/C_0 = (C_i \times S_i)/C_0 = (C/C_0) \times S_i = V_i \times S_i \)). In many cases, it is not easy to obtain the individual values of the visit counts and service times. However, Eq. (3.2.3) indicates that the service demand can be computed directly from the device utilization and system throughput. The multiclass version of the Service Demand Law is \( D_{i,r} = U_{i,r}/X_{0,r} = V_{i,r} \times S_{i,r} \).

**Example 3.3.**

A Web server is monitored for 10 minutes and its CPU is observed to be busy 90% of the monitoring period. The Web server log reveals that 30,000 requests are processed in that interval. What is the CPU service demand of requests to the Web server?

The observation period \( T \) is 600 (= 10 \times 60) seconds. The Web server throughput, \( X_0 \), is equal to the number of completed requests \( C_0 \) divided by the observation interval; \( X_0 = 30,000/600 = 50 \) requests/sec. The CPU utilization is \( U_{\text{CPU}} = 0.9 \). Thus, the service demand at the CPU is \( D_{\text{CPU}} = U_{\text{CPU}}/X_0 = 0.9/50 = 0.018 \) seconds/request.
Summary

Cross-team communication is, at best, inefficient
Everyone has many modeling techniques to master
Important information is often lost in translation
Teams argue about adherence to modeling standards
None of these models are immediately intuitive

There has *GOT* to be a better approach!!
Visual Application Usage Modeling

Something that may be better
A Visual Model of Application Usage

[Diagram showing a visual model of application usage with nodes labeled 'Login', 'Account Maint.', 'FAQ', 'Register', and 'HomePage'.]
A Visual Model of Application Usage
A Whiteboard Accident

We started using it and named it UCML™…

Value:

- Visually depict system with multiple usage paths and/or users.
- Document system usage and associated parameters.
- Easier to create than many other methods.
- Intuitively understood by all members of the team.
- Is a modeling framework, not a standard.
- Gives team a common language to discuss usage models.
- Can be done with paper, whiteboards, graphics tools or… crayons.
A Visual Model of Application Usage
Sometimes simple gets us further, faster

Being able to use the same medium to talk to everyone from user through developer to academic often yields better models

Simple may not replace other models, but rather feed into them or supplement them

Sometimes, a lack of standards are more unifying than rigid standards

Whiteboards are easier to use for multi-user collaboration than projectors

We could always use another excuse to use…
Summary

Crayons!
Where to go for more Information

George Box  http://www.engr.wisc.edu/ie/faculty/box_george.html
Edward Tufte  http://www.edwardtufte.com
Alberto Savoia  http://www.agitar.com
Connie Smith  http://www.perfeng.com/
Daniel Menasce  http://cs.gmu.edu/faculty/menasce.html
Questions
Contact Info

Scott Barber
Chief Technologist
PerfTestPlus, Inc

E-mail: sbarber@perftestplus.com
Web Site: www.PerfTestPlus.com