Automatic Extraction of Session-Based Workload Specifications for Architecture-Level Performance Models

Christian Vögele, André van Hoorn, Helmut Krcmar

Includes material created by Eike Schulz (Kiel University)
Domain

Background and Research Context

Interactive business-critical software systems

Domain

Background and Research Context

Interactive session-based business-critical software systems

Session:
“A series of consecutive and related requests issued by the same customer”

(Menascé et al. 1999)
Specifying Session-Based Workloads

Background and Research Context

(van Hoorn et al.) – building on (Menascé et al. 1999) and (Krishnamurthy et al. 2006)
Specifying Session-Based Workloads

Background and Research Context

(\textit{van Hoorn et al.}, 2008) – building on (Menascé et al. 1999) and (Krishnamurthy et al. 2006)
Specifying Session-Based Workloads

Background and Research Context

(van Hoorn et al., 2008) - building on (Menascé et al. 1999) and (Krishnamurthy et al. 2006)
Specifying Session-Based Workloads

Background and Research Context

(van Hoorn et al., 2008) – building on (Menascé et al. 1999) and (Krishnamurthy et al. 2006)
Markov4JMeter Reference Implementation

Background and Research Context

(van Hoorn et al., 2008)
Markov4JMeter Reference Implementation

Background and Research Context

Limitations:
1. Modeling language not formalized
2. Only a JMeter supported as reference implementation by Markov4JMeter extension
3. How to get to the workload specifications?

→ WESSBAS:
Workload Extraction and Specification for Session-Based Application Systems

(van Hoorn et al., 2008)
WESSBAS Approach before LT ‘15

1. **WESSBAS-DSL** as tool- and system-agnostic (intermediate) modeling language
2. Extraction of **WESSBAS-DSL instances** from monitoring data (employing clustering)
3. **Transformation** from **WESSBAS-DSL** to JMeter/Markov4JMeter test plan
4. **Evaluation** using SPECjEnterprise

**WESSBAS** is an acronym for Workload Extraction and Specification for Session-Based Application Systems

(van Hoorn et al., 2014)
WESEMBAS Approach before LT ’15

Selected Evaluation Results – SPECjEnterprise Experiments
WEESBAS Approach **before** LT ‘15

Selected Evaluation Results – SPECjEnterprise Experiments

WEESBAS is an acronym for Workload Extraction and Specification for Session-Based Application Systems

Load (Faban)

System Under Test (SUT)

Logging (Kieker)

Session Records

WEESBAS Instance

Extraction (incl. Clustering)

Transformation

JMeter (Markov4J) Testplan

Load Test

SUT'

Load (Faban)

System Under Test (SUT)

Logging (Kieker)

Session Records

WEESBAS Instance

Extraction (incl. Clustering)

Transformation

JMeter (Markov4J) Testplan

Load Test

SUT'
1. Workload characteristics do not differ among each other when using different clustering results.
2. Session lengths and no. of distinct sessions differ from original characteristics.
3. Server-side request counts exactly match the original characteristics.
LT ‘15 Contribution: WESSBAS-DSL to PCM

- Longer term goal: Integration of workload modeling
- LT ‘15: Transformation of WESSBAS-DSL to Palladio Component Model (PCM)
Transformation into Palladio Component Models

Automatically generated (e.g., Brunnert et al. 2013, Brosig et al. 2011)

Generation of large parts of the workload specification into the repository model

Missing possibility to reference usage models from usage models

Becker et al. (2009)

<table>
<thead>
<tr>
<th>WESSBAS-DSL</th>
<th>PCM Model Elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavior Models</td>
<td>Repository Model (Basic Component, RDSEFF)</td>
</tr>
<tr>
<td>Session Layer FSMs</td>
<td>not required</td>
</tr>
<tr>
<td>Protocol Layer FSMs</td>
<td>not required</td>
</tr>
<tr>
<td>Workload Intensity</td>
<td>Usage Model (Closed Workload)</td>
</tr>
<tr>
<td>Behavior Mix</td>
<td>Usage Model (Branch)</td>
</tr>
</tbody>
</table>
Generation into PCM Repository Model
Transformation into Palladio Component Models

WESSBAS-DSL
Generation into PCM Repository Model
Transformation into Palladio Component Models

WESSBAS-DSL

Call to the modelled system operation
Accuracy of PCM Workload Specification?

Evaluation

- Question: How well match the predicted workload characteristics the measured workload characteristics?
- Experimental setting:
  - Transformation and simulation of our VALUETOOLS SPECjEnterprise instances
  - Comparison of measured and predicted request counts

<table>
<thead>
<tr>
<th>Request</th>
<th>Orig. MRC</th>
<th>SRC</th>
<th>PE</th>
<th>2 Behavior Models MRC</th>
<th>SRC</th>
<th>PE</th>
<th>3 Behavior Models SRC</th>
<th>PE</th>
<th>4 Behavior Models SRC</th>
<th>PE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 add to cart</td>
<td>63,761</td>
<td>64,943</td>
<td>1.82%</td>
<td>61,812</td>
<td>3.15%</td>
<td>60,986</td>
<td>4.55%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 cancel order</td>
<td>632</td>
<td>609</td>
<td>3.78%</td>
<td>661</td>
<td>4.39%</td>
<td>625</td>
<td>1.12%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 clear cart</td>
<td>6,047</td>
<td>6,178</td>
<td>2.12%</td>
<td>5,927</td>
<td>2.02%</td>
<td>5,846</td>
<td>3.44%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 defer order</td>
<td>6,782</td>
<td>6,873</td>
<td>1.32%</td>
<td>6,524</td>
<td>3.95%</td>
<td>6,606</td>
<td>2.66%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 home</td>
<td>59,934</td>
<td>61,146</td>
<td>1.98%</td>
<td>58,747</td>
<td>2.02%</td>
<td>58,744</td>
<td>2.03%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 inventory</td>
<td>30,596</td>
<td>30,539</td>
<td>0.19%</td>
<td>29,574</td>
<td>3.46%</td>
<td>29,405</td>
<td>4.05%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 login</td>
<td>61,500</td>
<td>61,156</td>
<td>0.56%</td>
<td>58,747</td>
<td>4.69%</td>
<td>58,745</td>
<td>4.69%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 logout</td>
<td>59,934</td>
<td>61,146</td>
<td>1.98%</td>
<td>58,747</td>
<td>2.02%</td>
<td>58,744</td>
<td>2.03%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 purchase cart</td>
<td>8,360</td>
<td>8,388</td>
<td>0.33%</td>
<td>7,976</td>
<td>4.81%</td>
<td>7,836</td>
<td>6.69%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 remove</td>
<td>3,027</td>
<td>2,986</td>
<td>1.37%</td>
<td>2,876</td>
<td>5.25%</td>
<td>2,949</td>
<td>2.64%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 sell inventory</td>
<td>66,679</td>
<td>66,131</td>
<td>0.83%</td>
<td>63,185</td>
<td>5.53%</td>
<td>63,914</td>
<td>4.33%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 shopping cart</td>
<td>9,074</td>
<td>9,164</td>
<td>0.98%</td>
<td>8,803</td>
<td>3.08%</td>
<td>8,795</td>
<td>3.17%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13 view items</td>
<td>498,601</td>
<td>491,812</td>
<td>1.38%</td>
<td>470,392</td>
<td>6.00%</td>
<td>475,000</td>
<td>4.97%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

MRC: Measured Request Count
SRC: Simulated Request Count
PE: Prediction Accuracy

\[
\begin{array}{c|ccc|ccc|ccc}
\hline
& \text{Orig.} & \text{2 Behavior Models} & \text{3 Behavior Models} & \text{4 Behavior Models} \\
& \text{MRC} & \text{SRC} & \text{PE} & \text{SRC} & \text{PE} & \text{SRC} & \text{PE} \\
\hline
\text{1 add to cart} & 63,761 & 64,943 & 1.82\% & 61,812 & 3.15\% & 60,986 & 4.55\% \\
\text{2 cancel order} & 632 & 609 & 3.78\% & 661 & 4.39\% & 625 & 1.12\% \\
\text{3 clear cart} & 6,047 & 6,178 & 2.12\% & 5,927 & 2.02\% & 5,846 & 3.44\% \\
\text{4 defer order} & 6,782 & 6,873 & 1.32\% & 6,524 & 3.95\% & 6,606 & 2.66\% \\
\text{5 home} & 59,934 & 61,146 & 1.98\% & 58,747 & 2.02\% & 58,744 & 2.03\% \\
\text{6 inventory} & 30,596 & 30,539 & 0.19\% & 29,574 & 3.46\% & 29,405 & 4.05\% \\
\text{7 login} & 61,500 & 61,156 & 0.56\% & 58,747 & 4.69\% & 58,745 & 4.69\% \\
\text{8 logout} & 59,934 & 61,146 & 1.98\% & 58,747 & 2.02\% & 58,744 & 2.03\% \\
\text{9 purchase cart} & 8,360 & 8,388 & 0.33\% & 7,976 & 4.81\% & 7,836 & 6.69\% \\
\text{10 remove} & 3,027 & 2,986 & 1.37\% & 2,876 & 5.25\% & 2,949 & 2.64\% \\
\text{11 sell inventory} & 66,679 & 66,131 & 0.83\% & 63,185 & 5.53\% & 63,914 & 4.33\% \\
\text{12 shopping cart} & 9,074 & 9,164 & 0.98\% & 8,803 & 3.08\% & 8,795 & 3.17\% \\
\text{13 view items} & 498,601 & 491,812 & 1.38\% & 470,392 & 6.00\% & 475,000 & 4.97\% \\
\hline
\text{\textbf{\sum}} & 874,927 & 871,071 & 0.44\% & 833,971 & 4.91\% & 838,195 & 4.38\% \\
\end{array}
\]
Future Work

• Automatic generation of application model → Executable load tests
  – Automatic learning of guards and actions
  – Generation of protocol layer
  – Modeling, extraction and generation of parameters
• Support for workload intensity → LIMBO (Kistowski et al. 2014)
• Additional transformations
  – to alternative workload generators
  – to other architecture-level performance models
  – from PCM to WESSBAS-DSL
• Online clustering to detect evolution of behavior mix
• Industrial case study
  ➢ Supplementary material (software, (meta-)models, data, scripts) publicly available online: http://markov4jmeter.sourceforge.net/lt15
Statements as Input for the LT ’15 Discussion

1. Thought-provoking statement or discussion question about the area
   (e.g., how could this work be validated?)

   • What values/use cases do you see in integrating workload modeling and
     extraction for measurement- and model-based performance evaluation?

2. Thought-provoking statement or discussion question about the area
   (e.g., how can this work be of value to industry).

   • How can we transfer state-of-the-art load modeling/testing approaches
do industrial practice?
   • Do you see a need to improve load testing practice?
References