Modeling Variations in Load Intensity over Time

Third International Workshop on Large-Scale Testing (LT 2014), co-located with ICPE’14
Jóakim v. Kistowski, Nikolas Herbst, Samuel Kounev | March 22nd, 2014
Motivation

- Page Requests for the German Wikipedia
Motivation

Page Requests for the German Wikipedia
Motivation

- Page Requests for the German Wikipedia
Additive decomposition into seasonal part, trend, and remainder. Created using BFAST [1].
Problem: No means to effectively capture and reproduce the varying load intensity of real-world cloud systems

Idea: Close this gap by creating a meta-model for load intensity variations

Benefits: Enable more precise communication and creation of realistic load scenarios for benchmarking

Actions: Creation of the Descartes Load Intensity Model (DLIM) and High-Level DLIM
Problem: No means to effectively capture and reproduce the varying load intensity of real-world cloud systems

Idea: Close this gap by creating a meta-model for load intensity variations

Benefits: Enable more precise communication and creation of realistic load scenarios for benchmarking

Actions: Creation of the Descartes Load Intensity Model (DLIM) and High-Level DLIM
Problem: No means to effectively capture and reproduce the varying load intensity of real-world cloud systems

Idea: Close this gap by creating a meta-model for load intensity variations

Benefits: Enable more precise communication and creation of realistic load scenarios for benchmarking

Actions: Creation of the Descartes Load Intensity Model (DLIM) and High-Level DLIM
Problem: No means to effectively capture and reproduce the varying load intensity of real-world cloud systems

Idea: Close this gap by creating a meta-model for load intensity variations

Benefits: Enable more precise communication and creation of realistic load scenarios for benchmarking

Actions: Creation of the Descartes Load Intensity Model (DLIM) and High-Level DLIM
## Related Work

### User Behavior Models (Markov Chains)
- van Hoorn et al. (2008) [2]: probabilistic and intensity-varying workloads
- Roy et al. (2013) [3]: workload volatility of a VoD system

### Workload Models
- Barford et al. (1998) [4]: file popularity and distribution (Web)
- Casale et al. (2012) [5]: bursts
- Beich et al. (2010) [6]: data popularity and user separation (Cloud)

### Statistical Model
- Feitelson (2002) [7]: workload representatativity through statistical characteristics
Descartes Load Intensity Meta-Model

- Describe arrival rate variations over time
- Provides structure for piece-wise mathematical functions
- Independent of work/request type
DLIM Example Instance

- Created using LIMBO
- Contains *Seasonal* part, *Trends*, and *Burst*

---

LIMBO: [www.descartes-research.net/tools](http://www.descartes-research.net/tools)
High-Level DLIM

- **DLIM Benefits:**
  - Powerful
  - Expressive
  - Deriving arrival rates or request time-stamps is easy

- **DLIM Drawbacks:**
  - Instances can become complex
  - Large trees may be unintuitive

**Solution:** **High-Level DLIM**

- Few parameters for load intensity variation description
- Easier to grasp
- Strictly structured into one **Seasonal**, **Trend**, recurring **Burst**, and **Noise** parts
HLDLIM Seasonal and Trend parts

- HLDLIM **Seasonal** part:

- HLDLIM **Trend** part:
HLDLIM Burst and Noise parts

- **HLDLIM Burst part:**

- **HLDLIM Noise part:**
  - Uniform Distribution
    - Minimum Noise Rate
    - Maximum Noise Rate
HLDLIM Example Instance

- **Seasonal part:**
  - Period: 24
  - Peaks per Seasonal: 1
  - Base Arrival Rate: 4
  - First Peak Arrival Rate: 12

- **Burst part:**
  - First Burst Offset: 46
  - Burst Peak Arrival Rate: 8
  - Burst Width: 4

- **Trend part:**
  - Number of Seasonal Periods within one Trend: 1
  - Trend List: 16, 20, 14
DLIM / HLDLIM Differences

- DLIM:

- HLDLIM:
**Preliminary Evaluation**

- **Trace:** German Wikipedia project-counts
- **Fitted using best effort approach**

<table>
<thead>
<tr>
<th></th>
<th>median error</th>
<th>mean error</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DLIM</strong></td>
<td>8.95%</td>
<td>16.08%</td>
</tr>
<tr>
<td><strong>HLDLIM</strong></td>
<td>9.64%</td>
<td>16.25%</td>
</tr>
</tbody>
</table>

**Motivation**

**Outline**

- Related Work
- Preliminary Evaluation
- Conclusions
- Final Slide

---

Jóakim v. Kistowski, Nikolas Herbst, Samuel Kounev – Load Intensity Model

March 22nd, 2014
Conclusions

- Two Meta-Models for load intensity variation description
  - **DLIM**: Powerful and expressive
  - **HLDLIM**: Abstract and concise

- Available as part of LIMBO

- Enables creation of custom load intensity variations for open workload based benchmarking

- Capable of accurately capturing real-world load intensity variations

---

LIMBO: [www.descartes-research.net/tools](http://www.descartes-research.net/tools)
Thank you for your Interest!

- Which parameters are missing in HLDLIM, which are superfluous?

- How would you go about using DLIM for run-time resource managements? Which additional use cases do you envision for DLIM?

- LIMBO: www.descartes-research.net/tools
References


