WORKLOAD-AWARE SYSTEM MONITORING USING PERFORMANCE PREDICTIONS APPLIED TO A LARGE-SCALE E-MAIL SYSTEM

Christoph Rathfelder (rathfelder@fzi.de)  
FZI Research Center for Information Technology Karlsruhe, Germany

Stefan Becker (stefan.becker@1und1.de)  
1&1 Internet AG, Germany

Klaus Krogmann (krogmann@fzi.de)  
FZI Research Center for Information Technology Karlsruhe, Germany

Ralf Reussner (reussner@kit.edu)  
Karlsruhe Institute of Technology, Germany
1. Quick Facts about the 1&1 Mail System
2. Motivation and Approach
3. Foundation
   1. Model-based Performance Predictions
   2. Palladio Approach
4. Case Study
5. Validation and Applicability of our Approach
6. Lessons Learned
7. Conclusion and Future Work
The 1&1 Mail System

Quick Facts

- Basis for several brands like GMX, WEB.DE, mail.com, india.com
- > 40 Million users
- > 200 Million emails per day
- Build up as service oriented system
  - More than 100 software services
  - Distributed over 2,000 hosts
Ensuring high availability with guaranteed quality of service
- Detailed monitoring of the system

Potential reasons for deviating resource utilisation
- Misbehaving loadbalancer
- Hardware failures
- Service unavailability
- Connection problems

Identified reason was

Motivation

Chris Rathfelder, Stefan Becker: Workload-aware System Monitoring Using Performance Predictions Applied to a Large-scale E-Mail System
© 2012 1&1 Internet AG and FZI Forschungszentrum Informatik
Ensuring high availability with guaranteed quality of service
- Detailed monitoring of the system

Potential reasons for deviating resource utilisation
- Misbehaving loadbalancer
- Hardware failures
- Service unavailability
- Connection problems

Identified reason was

System monitoring should be aware of the current usage profile
Applying performance prediction to derive the expected system behaviour depending on the monitored user behaviour
Applying performance prediction to derive the expected system behaviour depending on the monitored user behaviour.
Foundations
Model-based Performance Predictions

Software System

Software Model

Annotated Software Model

Prediction Results

Performance Model

Measurements/Estimation

Simulation/Analysis

Transformation

Automated and encapsulated in tooling

describes
Automated and encapsulated in tooling

Measurements/Estimation

Software System

Chris Rathfelder, Stefan Becker: Workload-aware System Monitoring Using Performance Predictions Applied to a Large-scale E-Mail System

Foundations Model-based Performance Predictions

Software Model

Prediction Results

CPU Utilisation
Response Time Histogram

Software System describes

Transformation

Simulation/Analysis

Measurement/Estimation

Automated and encapsulated in tooling

CPU Demand: 100 Units

Prob. = 0.5

Call Freq.: 5Hz

CPU: 1GHz

Param1 := 10

Processing Rate CPU: 1GHz

Simulation/Analysis

Measurement/Estimation

Prediction Results

Software Model

Annotated Software Model

Performance Model

[Becker 09]
Foundations
Palladio - Tooling

- Palladio Component Model (PCM)
  - Domain-specific modelling language
  - Aligned with UML2 syntax

- Transformations into predictions models and simulation
  - Performance
  - Reliability

- Palladio Bench
  - Eclipse-based tool
  - Integrated modelling and prediction
  - Graphical editors

- Tool maturity and availability
  - Open source
    - http://www.palladio-simulator.com
  - Development started in 2003
  - > 20 active contributor
Foundations
Palladio Approach

Parameterized Models

- Component Repository
- System Model
- Hardware and Deployment Model
- Usage Model

Response Time

Resource Utilization

Service-Level Prediction

Chris Rathfelder, Stefan Becker: Workload-aware System Monitoring Using Performance Predictions Applied to a Large-scale E-Mail System
© 2012 1&1 Internet AG and FZI Forschungszentrum Informatik
Agenda

1. Quick Facts about the 1&1 Mail System
2. Motivation and Approach
3. Foundation
   1. Model-based Performance Predictions
   2. Palladio Approach
4. Case Study
5. Validation and Applicability of our Approach
6. Lessons Learned
7. Conclusion and Future Work

Chris Rathfelder, Stefan Becker: Workload-aware System Monitoring Using Performance Predictions Applied to a Large-scale E-Mail System
Chris Rathfelder, Stefan Becker: Workload-aware System Monitoring Using Performance Predictions Applied to a Large-scale E-Mail System
The 1&1 Mail System – Top Level View

DNS MX-Lookup – Load-Balancer

MDA MTA MP

MDA MTA MX

CLDR IDENTITY QUOTA EBAY TDIAG CMGR SPAM COMS VIRUS

www FE mobile, apps

TRINITY-MSG-REST

PROXY POP IMAP SGATE

STORE Mails / Folder

BACKUP

SERIE

DBFM

UAS

Chris Rathfelder, Stefan Becker: Workload-aware System Monitoring Using Performance Predictions Applied to a Large-scale E-Mail System

© 2012 1&1 Internet AG and FZI Forschungszentrum Informatik
The 1&1 Mail System – Top Level View

Chris Rathfelder, Stefan Becker: Workload-aware System Monitoring Using Performance Predictions Applied to a Large-scale E-Mail System

© 2012 1&1 Internet AG and FZI Forschungszentrum Informatik
The 1&1 Mail System – Top Level View

**DNS MX-Lookup** → **Load-Balancer**

**MDA** → **MTA**

**MP** → **MX**

**PROXY**
- **POP**
- **IMAP**
- **SGATE**

**STORE** → **BACKUP**

**SERIE**

**DBFM**

**www FE** → **mobile, apps CLIENTS**

**TRINITY-MSG-REST**

**UAS**

Chris Rathfelder, Stefan Becker: Workload-aware System Monitoring Using Performance Predictions Applied to a Large-scale E-Mail System
© 2012 1&1 Internet AG and FZI Forschungszentrum Informatik
The 1&1 Mail System – Top Level View

DNS MX-Lookup

Load-Balancer

MDA MTA
MP

MDA MTA
MX

CLDR IDENT QUOTA EBAY TDIAG CMGR SPAM COMS VIRUS

www FE

mobile, apps

CLIENTS

TRINITY-MSG-REST

PROXY
POP IMAP SGATE

STORE
Mails / Folder

BACKUP

SERIE

DBFM

UAS
Case Study

Analysis
- System architecture
- Hardware environment
- Available monitoring data

Architecture Model
- Structure
- Behaviour
- Deployment

Calibrations
- Measurements [Logs]
- Resource demand estimation
- Usage model
### Challenges

- **Many servers**
  - 113 even in the selected subsystem

- **Live system**
  - High availability must be guaranteed
  - No experiments
  - Expensive generation of load
  - QA System can not be used because of virtual hosts, mocks etc.

- **Distributed knowledge**

- **Outdated documentation**
Case Study

Store System

Analysis → Architecture Model → Calibrations

1. SGATE-Requests
2. POP-Requests
3. IMAP-Requests

PROXY
1. IMAP
2. POP
3. SGATE

STORE-Cluster

STORE

BACKUP

Analysis
Architecture Model
Calibrations

Chris Rathfelder, Stefan Becker: Workload-aware System Monitoring Using Performance Predictions Applied to a Large-scale E-Mail System
© 2012 1&1 Internet AG and FZI Forschungszentrum Informatik
Case Study

Repository

Analysis Architecture Model Calibrations

Chris Rathfelder, Stefan Becker: Workload-aware System Monitoring Using Performance Predictions Applied to a Large-scale E-Mail System

© 2012 1&1 Internet AG and FZI Forschungszentrum Informatik
Case Study

System Model

Analysis Architecture Model Calibrations

Chris Rathfelder, Stefan Becker: Workload-aware System Monitoring Using Performance Predictions Applied to a Large-scale E-Mail System

© 2012 1&1 Internet AG and FZI Forschungszentrum Informatik
## Resource Demand Estimation

### Model function for incoming traffic

<table>
<thead>
<tr>
<th>Bytes received</th>
<th>Function</th>
<th>Relation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AppendMail</td>
<td>Byte</td>
</tr>
<tr>
<td>1</td>
<td>GetMailText</td>
<td>Byte</td>
</tr>
<tr>
<td>1</td>
<td>RETR</td>
<td>Byte</td>
</tr>
<tr>
<td>530966</td>
<td>APPEND</td>
<td>Request</td>
</tr>
<tr>
<td>143349</td>
<td>GetMails (+)</td>
<td>Request</td>
</tr>
<tr>
<td>83067</td>
<td>TOPN</td>
<td>Request</td>
</tr>
<tr>
<td>62479</td>
<td>TOP0</td>
<td>Request</td>
</tr>
<tr>
<td>23911</td>
<td>SortMails (+)</td>
<td>Request</td>
</tr>
<tr>
<td>3089</td>
<td>FETCH</td>
<td>Request</td>
</tr>
<tr>
<td>1298</td>
<td>STORE</td>
<td>Request</td>
</tr>
</tbody>
</table>
APPEND MAIL request on a PROYS server

Case Study

Service Effect Specification

Chris Rathfelder, Stefan Becker: Workload-aware System Monitoring Using Performance Predictions Applied to a Large-scale E-Mail System

© 2012 1&1 Internet AG and FZI Forschungszentrum Informatik
**Prediction Accuracy**

- Comparison between real and predicted utilizations per resource
- 1 reference day
- 48 Measurements
- 30 minutes each
- Over all mean error: 8.6%

<table>
<thead>
<tr>
<th></th>
<th>CPU</th>
<th>NET in</th>
<th>NET out</th>
<th>DISK</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Whole Model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Decile</td>
<td>12.94</td>
<td>5.99</td>
<td>5.92</td>
<td>9.64</td>
</tr>
<tr>
<td>36.7</td>
<td>13.46</td>
<td>12.37</td>
<td>20.1</td>
<td></td>
</tr>
<tr>
<td><strong>PROXY-Server</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Decile</td>
<td>5.9</td>
<td>2.93</td>
<td>5.92</td>
<td>9.64</td>
</tr>
<tr>
<td>10.1</td>
<td>6.4</td>
<td>11.09</td>
<td>20.1</td>
<td></td>
</tr>
<tr>
<td><strong>STORE-Server</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Decile</td>
<td>15.29</td>
<td>7.01</td>
<td>6.01</td>
<td>9.64</td>
</tr>
<tr>
<td>42.1</td>
<td>16.19</td>
<td>13.31</td>
<td>20.1</td>
<td></td>
</tr>
</tbody>
</table>

Model errors in percent
Applicability of our Approach

Simulated System Failure

- Real and predicted resource utilisation over the day
- Planned update between 2:00 p.m. and 6:00 p.m.
- Situation for the model to raise an alarm
Distributed knowledge and outdated documentation increase efforts
- Performance models can serve as documentation

Very large systems can not be measured out by performing experiments
- High availability => too dangerous
- Creating the needed load needs many resources
- Existing logfiles often contain useful data

Redundant parts can be depicted together in a model
- Combining servers with identical hardware software and load

Detection of missconfigurations
- Possible already in early steps of the modelling process
- I.E. we found a server with disabled hyperthreading.
Conclusions

• Enhancing a system monitoring process with performance prediction results
• Applicability of PCM in an industrial environment shown (mean model error of 8.6%)

Future Work

• Improved support for replicated software components
• Automated model extraction and calibration
• Runtime management based on performance predictions

http://www.palladio-simulator.com