Predicting Performance Degradations of Black-Box Microservice Applications

11th Symposium on Software Performance 2020

Martin Straesser, Johannes Grohmann

13 November 2020

Leipzig, Germany

https://se.informatik.uni-wuerzburg.de
Problem: Reactive Monitoring of Microservices

- Microservice applications as main architectural paradigm for cloud applications [1]
- Performance engineers rely on reactive application performance management (APM) tools
- **Proactive** performance management needed to ensure user experience and revenue

---

SuanMing: Enabling Proactiveness for APMs

- **SuanMing** augments an existing reactive monitoring stack with a **proactive component**

- Weaknesses of related work (e.g. [5-10])
  - Lack of explainability
  - Require low-level hardware measurements or application logs

- Goals
  - High **explainability**
  - **No prior knowledge**
  - Modular and extensible

- Prediction process divided into four steps

SuanMing: Predicting Performance Degradations

Initialization ➔ Load Forecasting ➔ Request Propagation ➔ Performance Inference ➔ Root-Cause Detection

Users
SuanMing: Predicting Performance Degradations

Initialization ➔ Load Forecasting ➔ Request Propagation ➔ Performance Inference ➔ Root-Cause Detection

Users ➔ Service A ➔ Service B

Predicting Performance Degradations of Black-Box Microservice Applications

Martin Straesser, Johannes Grohmann
SuanMing: Predicting Performance Degradations

Initialization ➔ Load Forecasting ➔ Request Propagation ➔ Performance Inference ➔ Root-Cause Detection

Users ➔ Service A ➔ Service B ➔ Service C ➔ Service D ➔ Service E ➔ Service F ➔ Service G ➔ Service H ➔ Service I

Predicting Performance Degradations of Black-Box Microservice Applications

Martin Straesser, Johannes Grohmann
SuanMing: Predicting Performance Degradations

Initialization ➔ Load Forecasting ➔ Request Propagation ➔ Performance Inference ➔ Root-Cause Detection

Users

Service A ➔ Service B ➔ Service C ➔ Service D ➔ Service E ➔ Service F ➔ Service G ➔ Service H ➔ Service I

“Sends requests to“
SuanMing: Predicting Performance Degradations

Initialization ➔ Load Forecasting ➔ Request Propagation ➔ Performance Inference ➔ Root-Cause Detection

Service A ➔ Service C ➔ Service E
30 req/s

Service B ➔ Service D ➔ Service E ➔ Service G ➔ Service I
60 req/s

Service H

Predicting Performance Degradations of Black-Box Microservice Applications

Martin Straesser, Johannes Grohmann
SuanMing: Predicting Performance Degradations

Initialization ➔ Load Forecasting ➔ Request Propagation ➔ Performance Inference ➔ Root-Cause Detection

Service A ➔ Service C ➔ Service E ➔ Service H
Service B ➔ Service D ➔ Service F ➔ Service G ➔ Service I

30 req/s
60 req/s
SuanMing: Predicting Performance Degradations

Initialization ➔ Load Forecasting ➔ Request Propagation ➔ Performance Inference ➔ Root-Cause Detection

Service A ➔ 30 req/s ➔ Service C ➔ 30 req/s ➔ Service F ➔ 30 req/s ➔ Service H

Service B ➔ 60 req/s ➔ Service D ➔ 60 req/s ➔ Service G ➔ 60 req/s ➔ Service I

Service E ➔ 80 req/s

Predicting Performance Degradations of Black-Box Microservice Applications

Martin Straesser, Johannes Grohmann
SuanMing: Predicting Performance Degradations

Initialization ➔ Load Forecasting ➔ Request Propagation ➔ Performance Inference ➔ Root-Cause Detection

Service A ➔ 30 req/s ➔ Service C

Service B ➔ 60 req/s ➔ Service D

Service C ➔ 60 req/s ➔ Service E

Service F ➔ 30 req/s ➔ Service H

Service G ➔ 30 req/s ➔ Service I

Service E ➔ 60 req/s ➔ 80 req/s ➔ Service H

Service F ➔ 60 req/s ➔ 30 req/s ➔ Service I

Service G ➔ 60 req/s ➔ 60 req/s ➔ Service I

Predicting Performance Degradations of Black-Box Microservice Applications

Martin Straesser, Johannes Grohmann
SuanMing: Predicting Performance Degradations

Initialization ➔ Load Forecasting ➔ Request Propagation ➔ Performance Inference ➔ Root-Cause Detection

Service A ➔ Service C ➔ Service E
30 req/s ➔ 60 req/s ➔ 80 req/s

Service B ➔ Service D ➔ Service G
60 req/s ➔ 60 req/s ➔ 60 req/s

Service H

Service I
SuanMing: Predicting Performance Degradations

**Initialization** ➔ **Load Forecasting** ➔ **Request Propagation** ➔ **Performance Inference** ➔ **Root-Cause Detection**

Predicting Performance Degradations of Black-Box Microservice Applications

*Martin Straesser, Johannes Grohmann*
SuanMing: Predicting Performance Degradations

Initialization ➔ Load Forecasting ➔ Request Propagation ➔ Performance Inference ➔ Root-Cause Detection

Service A ➔ Service B ➔ Service C ➔ Service D ➔ Service E ➔ Service H ➔ Service F ➔ Service G ➔ Service I

30 req/s ➔ 60 req/s ➔ 30 req/s ➔ 60 req/s ➔ 80 req/s

“performance depends on“

Martin Straesser, Johannes Grohmann
SuanMing: Predicting Performance Degradations

Initialization ➔ Load Forecasting ➔ Request Propagation ➔ Performance Inference ➔ Root-Cause Detection

Predicting Performance Degradations of Black-Box Microservice Applications

Martin Straesser, Johannes Grohmann
SuanMing: Predicting Performance Degradations

Initialization ➔ Load Forecasting ➔ Request Propagation ➔ Performance Inference ➔ Root-Cause Detection

Service A ➔ Service C ➔ Service E
30 req/s ➔ 60 req/s ➔ 80 req/s

Service B ➔ Service D ➔ Service G ➔ Service I
60 req/s ➔ 60 req/s ➔ 60 req/s ➔ 60 req/s

Service H
80 req/s

Predicting Performance Degradations of Black-Box Microservice Applications
Martin Straesser, Johannes Grohmann
SuanMing: Predicting Performance Degradations

Initialization ➞ Load Forecasting ➞ Request Propagation ➞ Performance Inference ➞ Root-Cause Detection

Service A ➞ Service C ➞ Service E, 30 req/s ➞ 60 req/s ➞ 80 req/s

Service B ➞ Service D ➞ Service G, 60 req/s ➞ 60 req/s ➞ 60 req/s

Service A ➞ Service C ➞ Service I, 30 req/s ➞ 60 req/s ➞ 60 req/s

Predicting Performance Degradations of Black-Box Microservice Applications

Martin Straesser, Johannes Grohmann
SuanMing: Predicting Performance Degradations

Initialization ➔ Load Forecasting ➔ Request Propagation ➔ Performance Inference ➔ Root-Cause Detection

Predicting Performance Degradations of Black-Box Microservice Applications

Martin Straesser, Johannes Grohmann
SuanMing: Predicting Performance Degradations

Initialization ➔ Load Forecasting ➔ Request Propagation ➔ Performance Inference ➔ Root-Cause Detection

Predicting Performance Degradations of Black-Box Microservice Applications

Martin Straesser, Johannes Grohmann
SuanMing: Predicting Performance Degradiations

Initialization ➔ Load Forecasting ➔ Request Propagation ➔ Performance Inference ➔ Root-Cause Detection

Service A ➔ Service C ➔ Service E ➔ Service H
30 req/s ➔ 60 req/s ➔ 80 req/s

Service B ➔ Service D ➔ Service F ➔ Service I
60 req/s ➔ 60 req/s ➔ 30 req/s ➔ 60 req/s

Legend:
- Green: Healthy service
- Yellow: Warning
- Red: Critical

Predicting Performance Degradations of Black-Box Microservice Applications

Martin Straesser, Johannes Grohmann
SuanMing: Predicting Performance Degradations

Initialization ➔ Load Forecasting ➔ Request Propagation ➔ Performance Inference ➔ Root-Cause Detection

Predicting Performance Degradations of Black-Box Microservice Applications

Martin Straesser, Johannes Grohmann
Predicting Performance Degradations of Black-Box Microservice Applications

Martin Straesser, Johannes Grohmann

Results

- Tests have been performed both on Teastore [11] and TrainTicket [12] applications
- Deployment in lab and real-world cloud environment
- Example question: Which time is needed to search for a train connection? Does it exceed a fixed threshold?

Conclusion

Problem

State-of-the-art reactive APM tools are unable to predict and mitigate performance degradations.
Conclusion

Problem

State-of-the-art reactive APM tools are unable to predict and mitigate performance degradations

Idea

Four step prediction process with load forecasting, request propagation, performance inference and root-cause detection
## Conclusion

<table>
<thead>
<tr>
<th>Problem</th>
<th>State-of-the-art reactive APM tools are unable to predict and mitigate performance degradations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idea</td>
<td>Four step prediction process with load forecasting, request propagation, performance inference and root-cause detection</td>
</tr>
<tr>
<td>Benefits</td>
<td>Augments existing monitoring stack, high explainability</td>
</tr>
</tbody>
</table>
Conclusion

Problem
State-of-the-art reactive APM tools are unable to predict and mitigate performance degradations

Idea
Four step prediction process with load forecasting, request propagation, performance inference and root-cause detection

Benefits
Augments existing monitoring stack, high explainability

Future Work
Evaluate more model types, further validation of results
Conclusion

Problem
State-of-the-art reactive APM tools are unable to predict and mitigate performance degradations

Idea
Four step prediction process with load forecasting, request propagation, performance inference and root-cause detection

Benefits
Augments existing monitoring stack, high explainability

Future Work
Evaluate more model types, further validation of results

THANK YOU FOR YOUR ATTENTION