University of Stuttgart

unibz

eSulabSolutions

NOVATEC

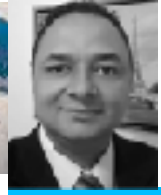
Integrating Automated Scalability Assessment into DevOps

Alberto Avritzer
et al.

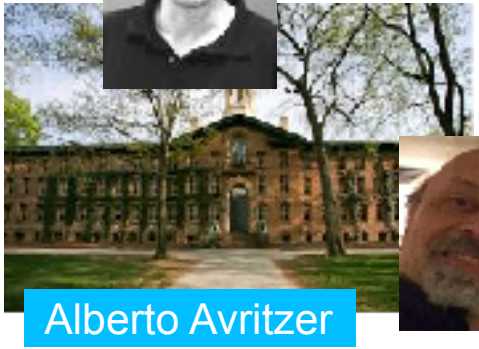
The Eighth International Workshop on Load Testing and Benchmarking of Software Systems (LTB 2020)

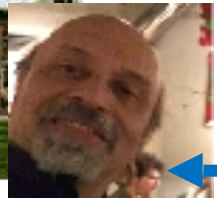
Our Team

Daniel Menasche



Vilc Rufino





Alberto Avritzer

eSulabSolutions,
near Princeton University

UFRJ

André van Hoorn

Universität Stuttgart

Vincenzo Ferme

Henning Schulz

NOVATEC

Barbara Russo

Andrea Janes

About Alberto Avritzer
eSulabSolutions

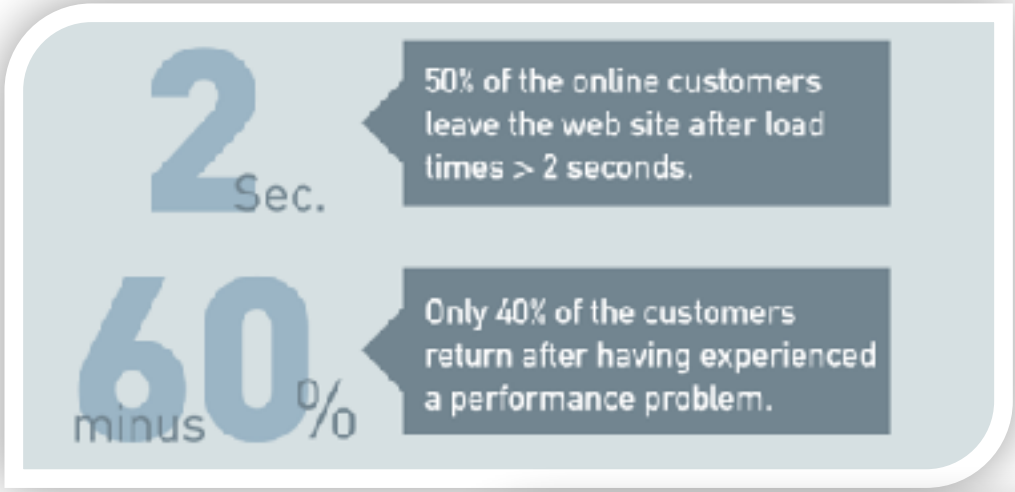
- Senior Researcher at Siemens and AT&T Bell Labs for about 24 years
- Published over 70 papers in journals, refereed conference proceedings, and book chapters in those areas: (http://dblp.uni-trier.de/pers/hd/a/Avritzer:Alberto)
  - **Resilience Assessment and Evaluation of Computing Systems. Springer 2012**
  - **Handbook of Software Aging and Rejuvenation, World Scientific, May 2020, World Scientific (https://worldscientific.com/worldscibooks/10.1142/11673)**
- Siemens/MTA: The NY Subway PA/CIS Scalability Assessment
- Load Testing and Performance Analysis of AT&T Operations support systems:
  - Monitoring for Software aging and Rejuvenation (1993)
  - Performance testing using Markov chain (1995)
- Founder of eSulabSolutions (https://esulabsolutions.godaddysites.com/):
  - Automated scalability assessment in DevOps and micro service architecture

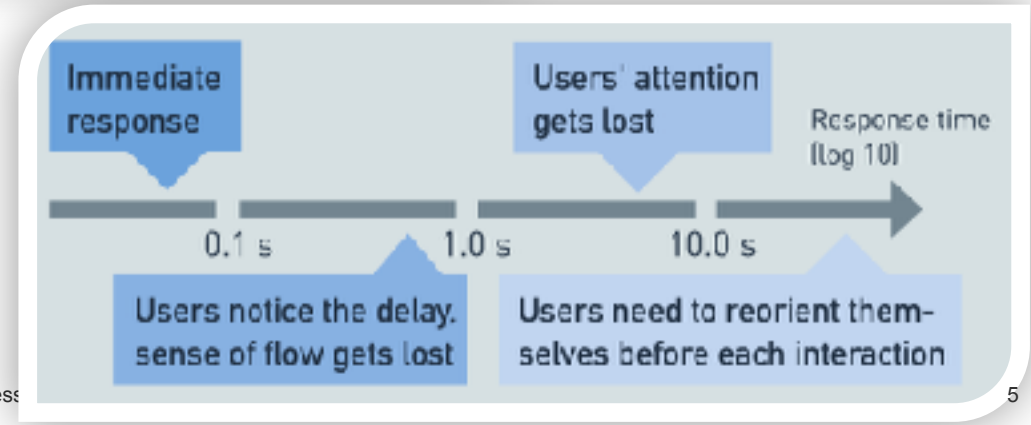
Motivation: Recent Scalability Related Disasters

Americans crash the Canadian immigration

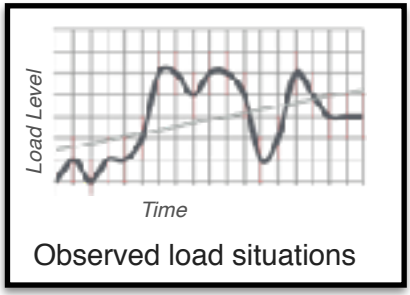
NYS Labor website crashes

Motivation: Influence of Poor Performance on the Success of Businesses



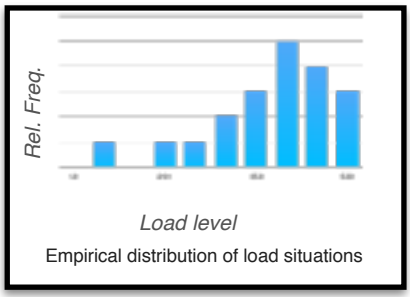


Examples of Operational Profile Representations

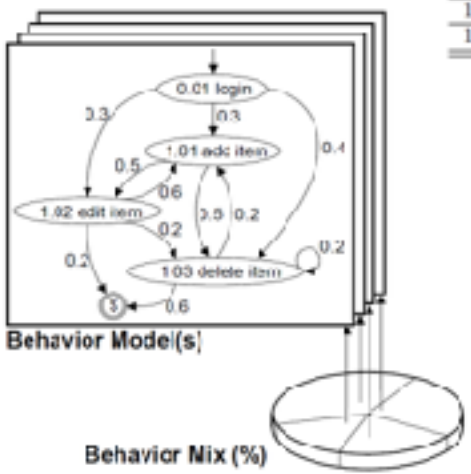


Observed load situations

	Request	*Orig.*	*Rel.*
1	add to cart	63,761	0.07
2	cancel order	632	0.00
3	clear cart	6,047	0.01
4	defer order	6,782	0.01
5	home	59,934	0.07
6	inventory	30,596	0.03
7	login	61,500	0.07
8	logout	59,934	0.07
9	purchase cart	8,360	0.01
10	remove	3,027	0.00
11	sell inventory	66,679	0.08
12	shopping cart	9,074	0.01
13	view items	498,601	0.57
	$\sum$	874,927	1.00



Empirical distribution of load situations



»DevOps is ..

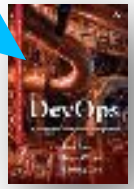
Monitoring
Agile
*-as-Code
Automation
Micro-services
Cultural Change
Chaos Engineering
Docker
CD Pipelines
Fast Feedback
Polyglot
Live Experi-mentation

… a **set of** **Practices**

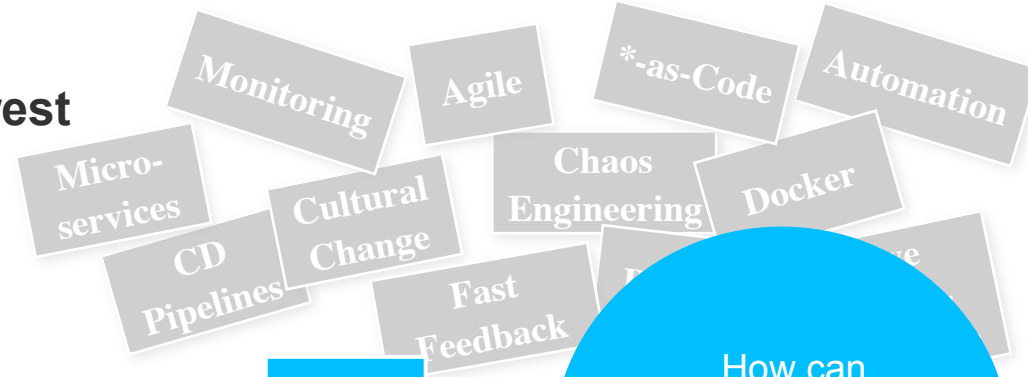
intended to
***reduce the time** between committing a change to a system and the **change being placed into normal production**,*

while ***ensuring*** **HIGH QUALITY** .«

COST-EFFICIENT
ELASTIC
RESILIENT
RELIABLE
FAST
SECURE



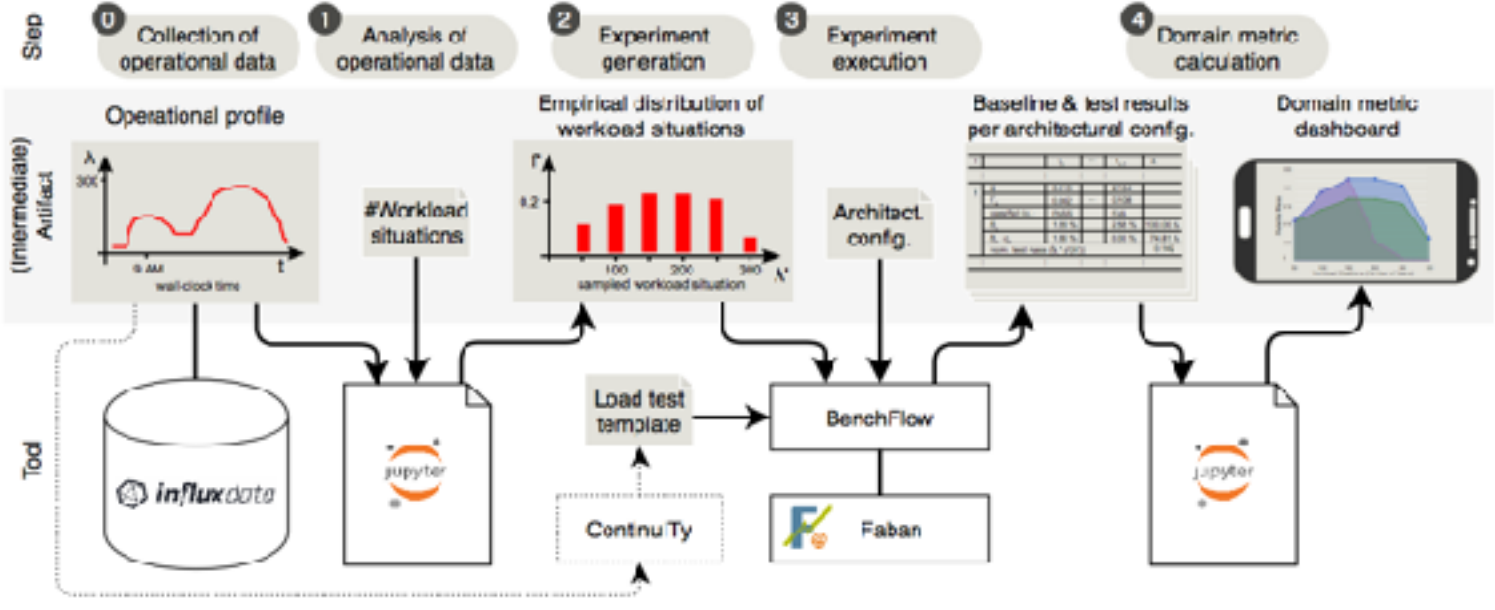
My Questions of Interest



How can DevOps practices support quality assurance?

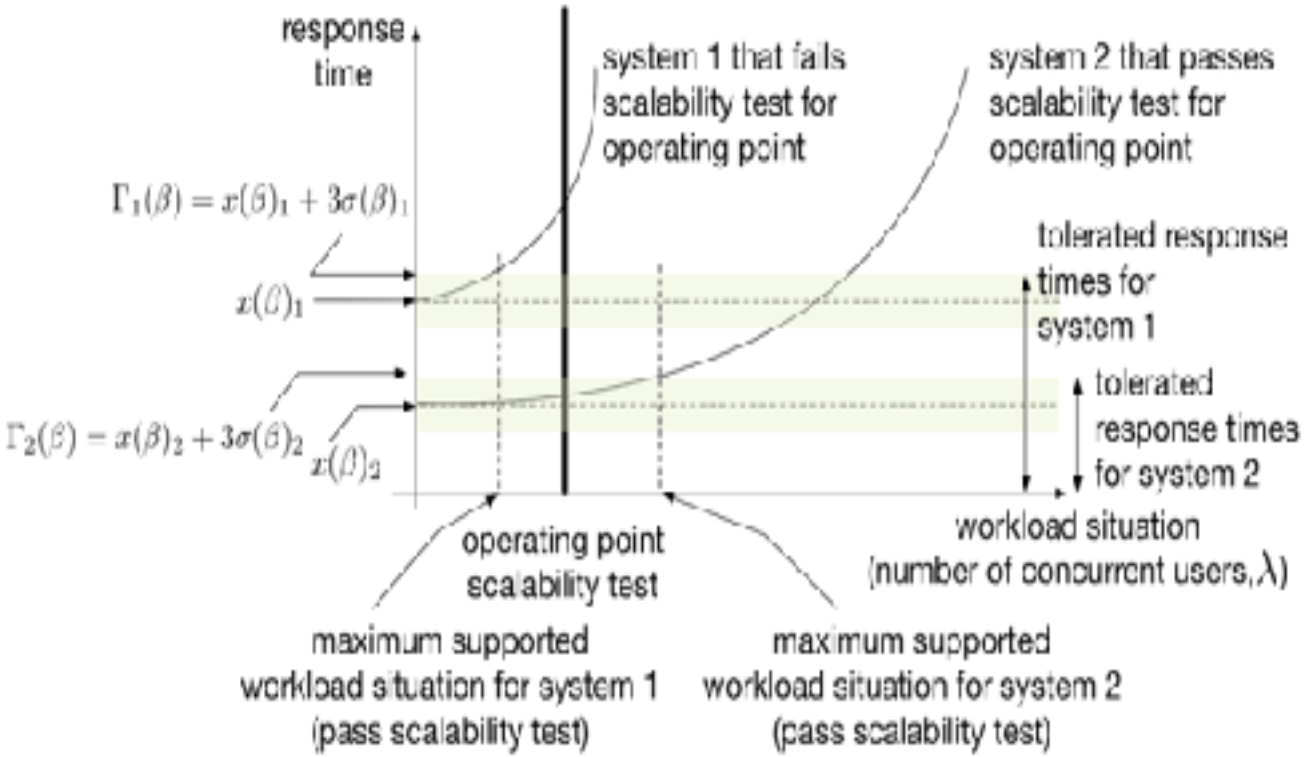
How to integrate „established“ quality assurance with DevOps practices?

- PPTAM: Production and Performance Testing Based Application Monitoring



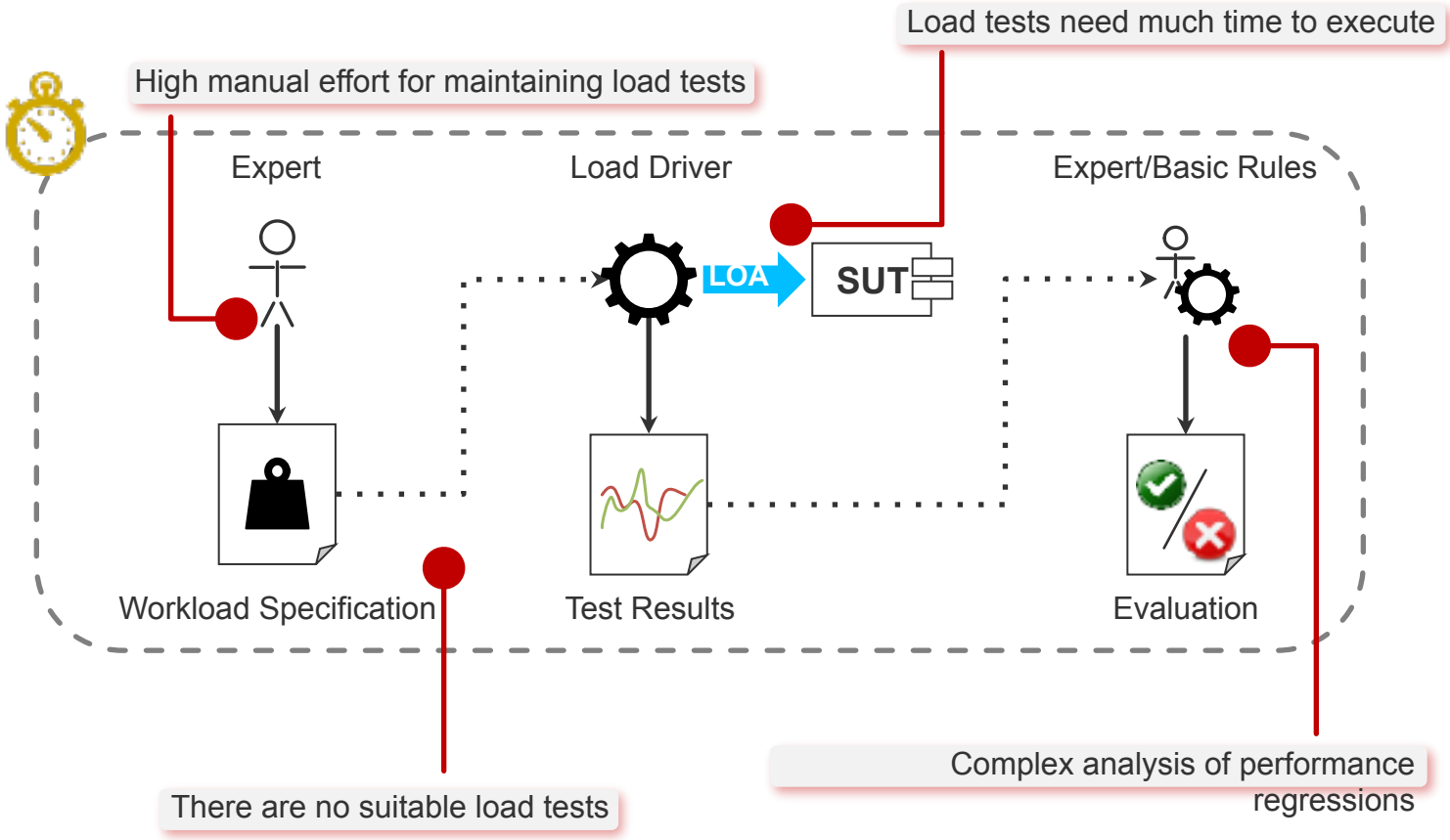
Avritzer et al.: Automated Scalability Assessment in DevOps

Scalability Requirement Automated Measurement



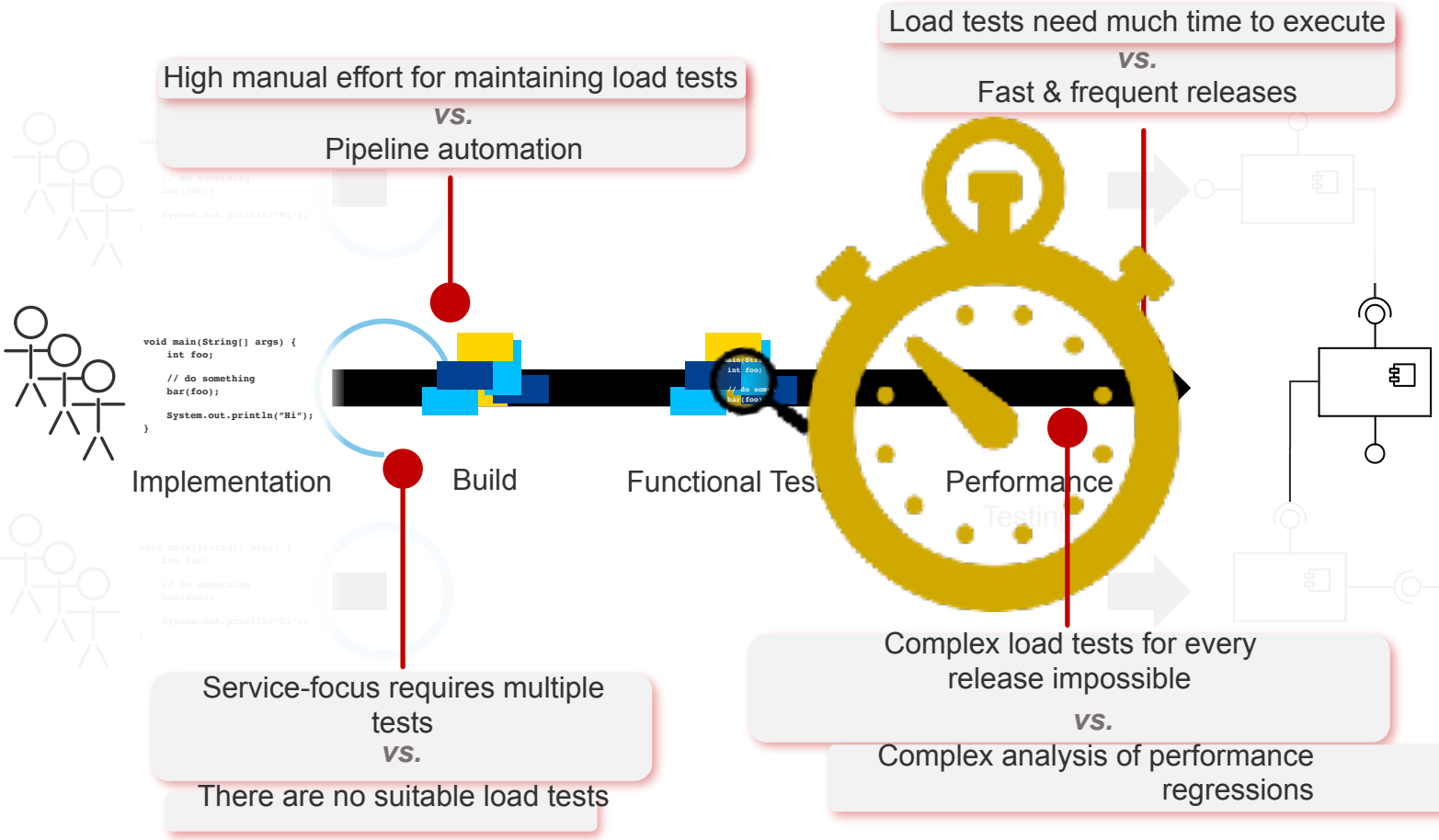
The Classic Load Testing Approach

… and Classic Problems

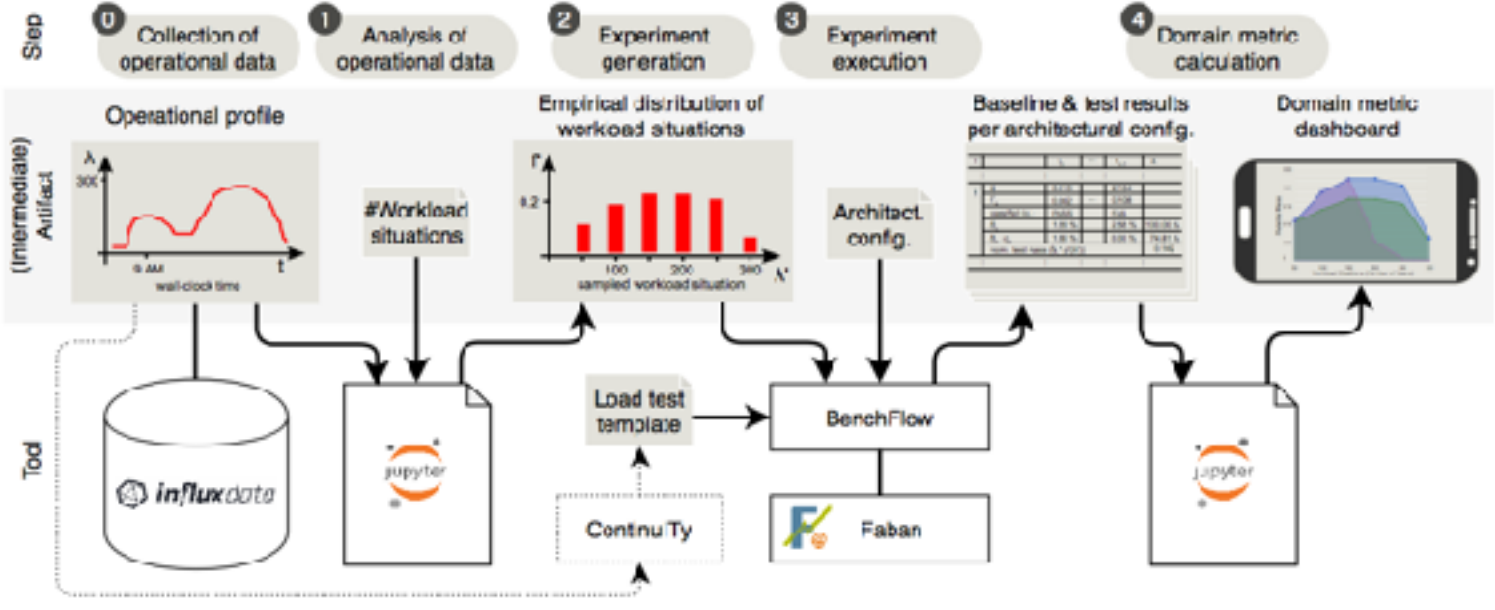


Load Testing in Continuous Delivery Pipelines

… How Problems Get Worse



- PPTAM: Production and Performance Testing Based Application Monitoring



Avritzer et al.: Automated Scalability Assessment in DevOps

Quantitative Assessment of Deployment Alternatives

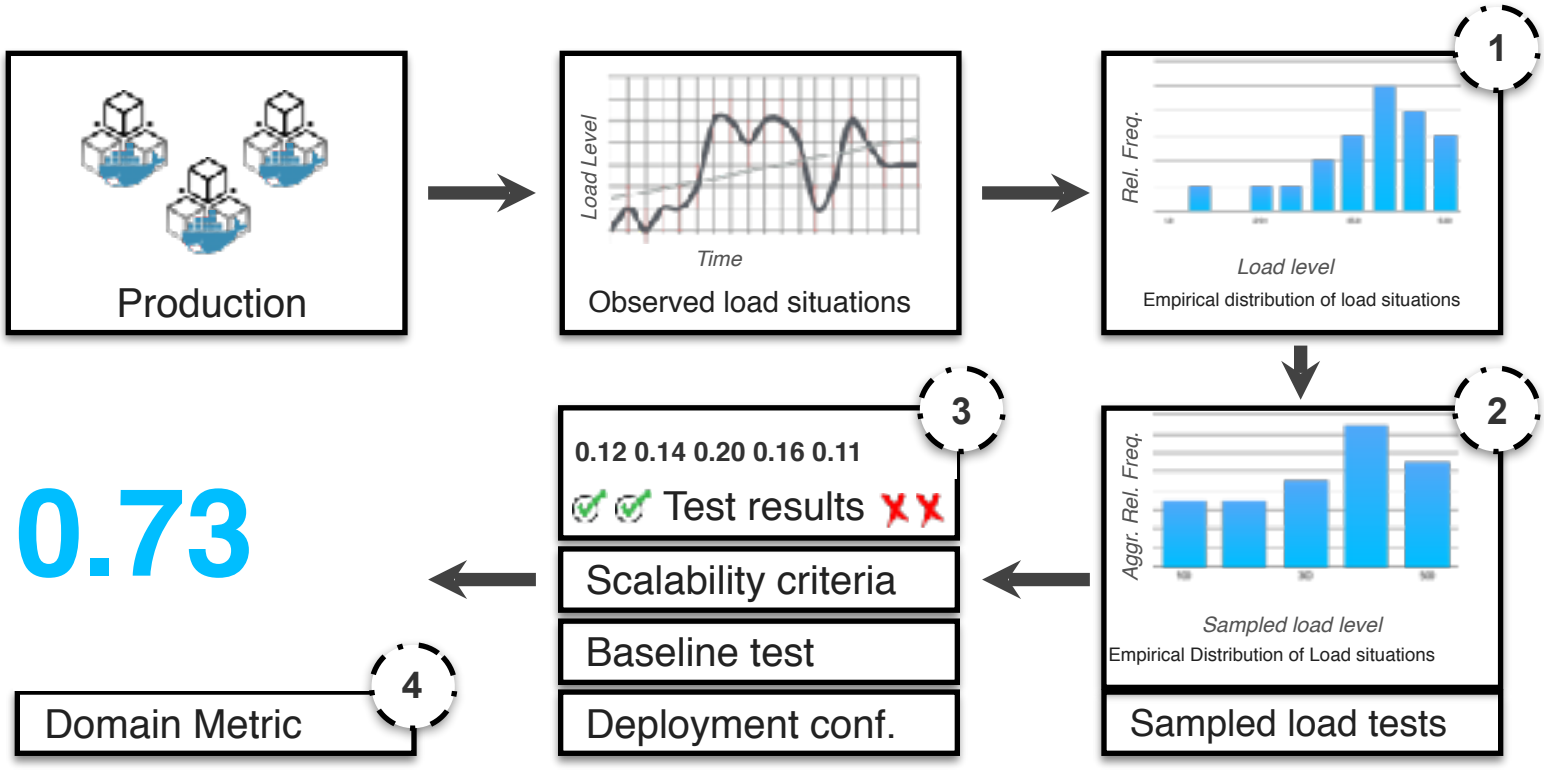
- Challenge: assess performance of architectural deployment alternatives (e.g., number of replicas, CPU/memory allocation, technology stack) under fuzzy requirements
- Our approach
  - Use operational data to generate and weigh load tests
  - Measure baseline requirements
  - Metric allows quantitative comparison of deployment alternatives
  - Builds on previous work from telecommunication systems:

    Avritzer, A., Weyuker, E.J.: The automatic generation of load test suites and the assessment of the resulting software. IEEE Trans. Softw. Eng. 21(9) (Sep 1995)

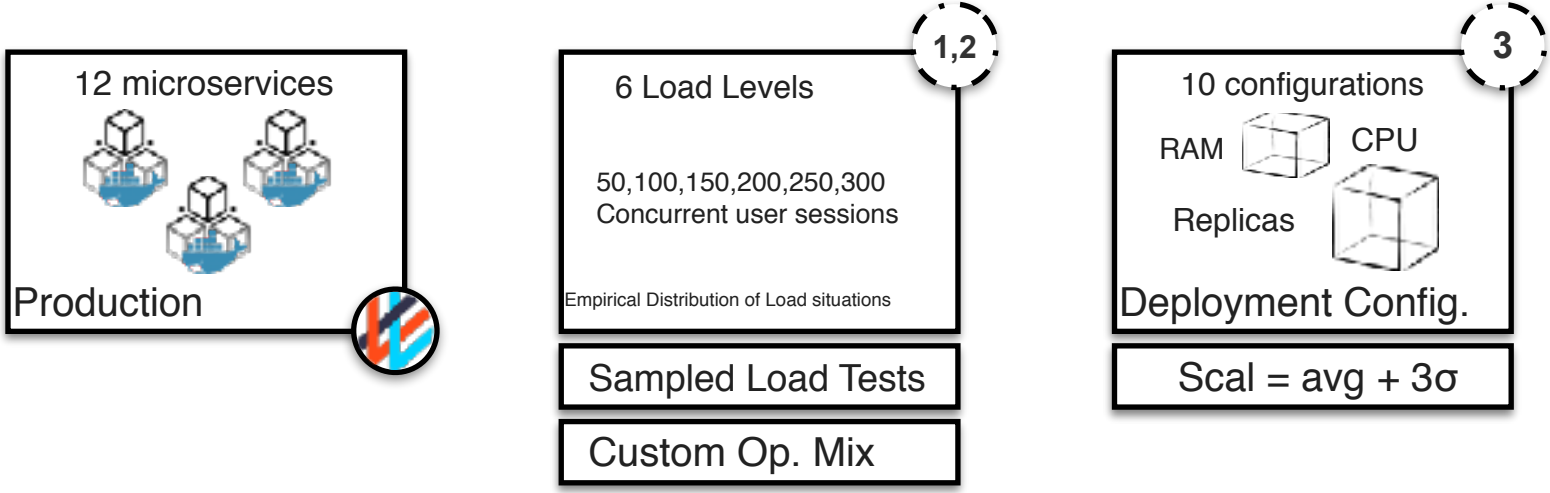
Alberto Avritzer, Vincenzo Ferme, Andrea Janes, Barbara Russo, Henning Schulz, and André van Hoorn:
A Quantitative Approach for the Assessment of Microservice Architecture Deployment Alternatives by Automated Performance Testing.

In Proceedings of the 12th European Conference on Software Architecture (ECSA). LNCS, Springer, 2018 (Accepted)

Overview of Approach



Experiments



Experiment Results: Computation of Domain Metric (1/2)

API	Scalability Criteria
GET /	**PASS**
GET /cart	**PASS**
POST /item	**FAIL**

Custom Op. Mix

Users	Aggr. Rel. Freq.
50	0.10582
100	0.18519
150	0.22222
200	0.22222
250	0.20370
300	0.06085

Aggr. Rel. Freq.

Max: 0.20370

Actual: 0.13580

Contrib. to Domain Metric

Deployment Configuration: 1 GB RAM, 0.25 CPU, 1 Replica

Experiment Results: Computation of Domain Metric (2/2)

Users	Contribution
50	0.10582
100	0.18519
150	0.22222
200	0.07999
250	0.13580
300	0.04729

Contrib. to Domain Metric

Max: 1

Actual:

0.77631

4

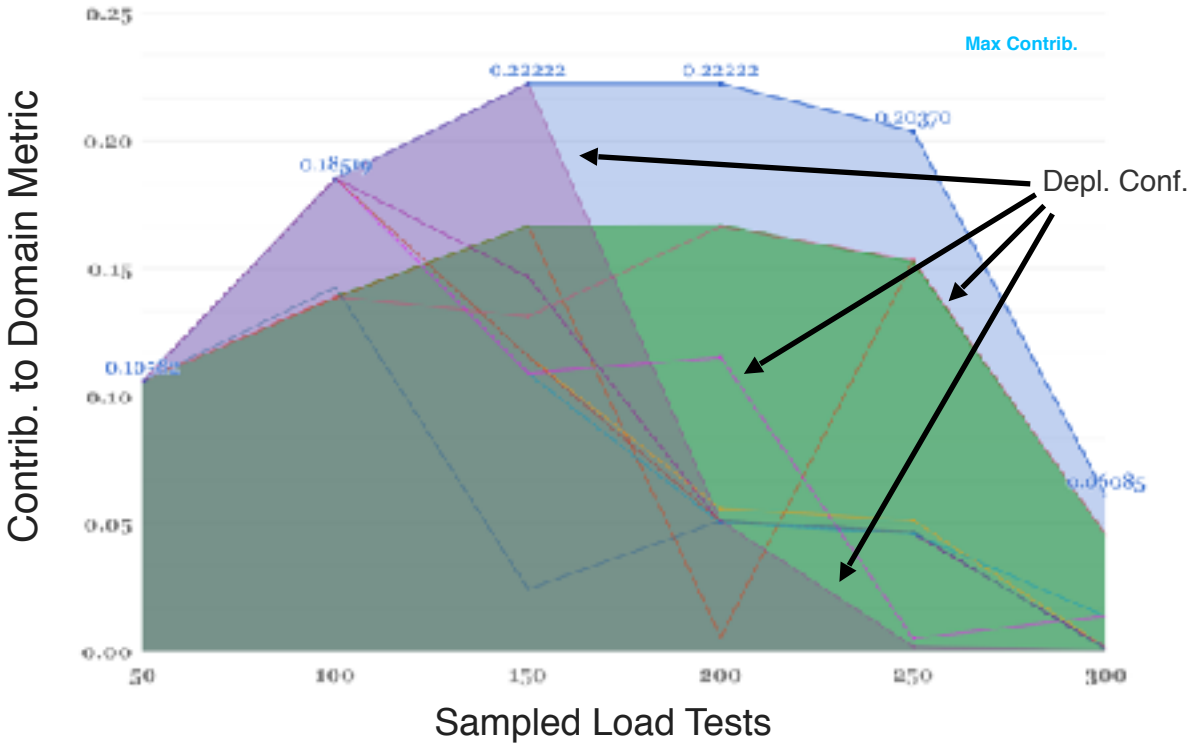
Domain Metric

Deployment Configuration: 1 GB RAM, 0.25 CPU, 1 Replica

Experiment Results: Single-Metric Comparison of Alternatives

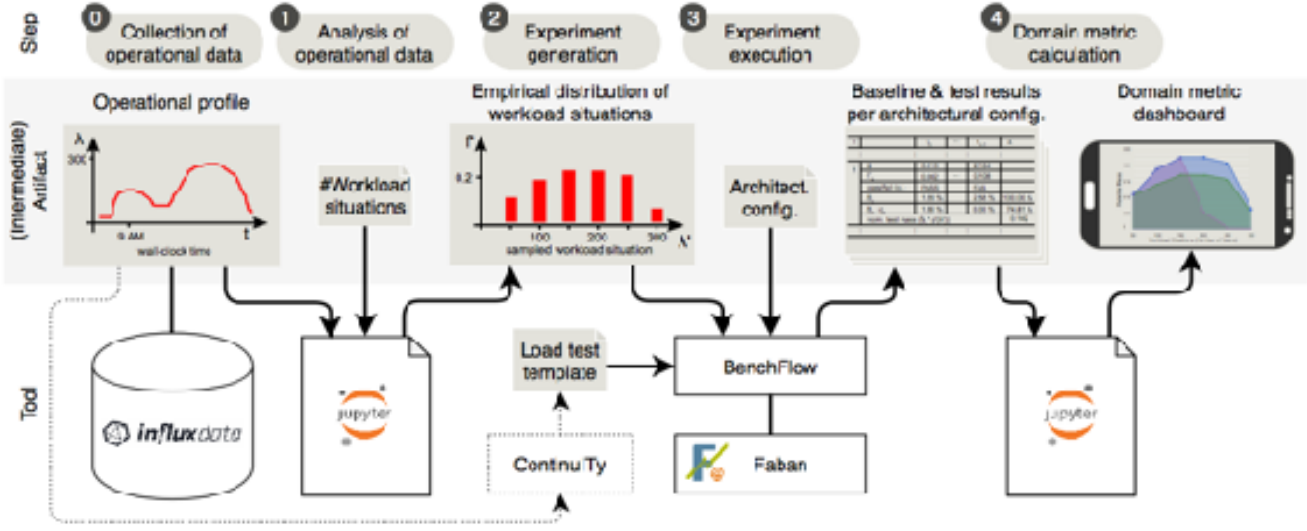
RAM	CPU #	Cart Replicas	Domain Metric (HPI)	Domain Metric (FUB)
0.5 GB	0.25	1	0.61499	0.54134
1 GB	**0.25**	**1**	**0.77631**	0.53884
1 GB	0.5	1	0.53559	0.54106
0.5 GB	0.5	1	0.51536	0.54773
0.5 GB	0.5	2	0.50995	0.54111
1 GB	0.25	2	0.74080	0.54785
1 GB	0.5	2	0.53401	0.54106
0.5 GB	***0.5***	*4*	0.50531	***0.54939***
1 GB	0.25	4	0.37162	0.54272
1 GB	0.5	4	0.56718	0.54271

Experiment Results: Visual Comparison of Alternatives



PPTAM

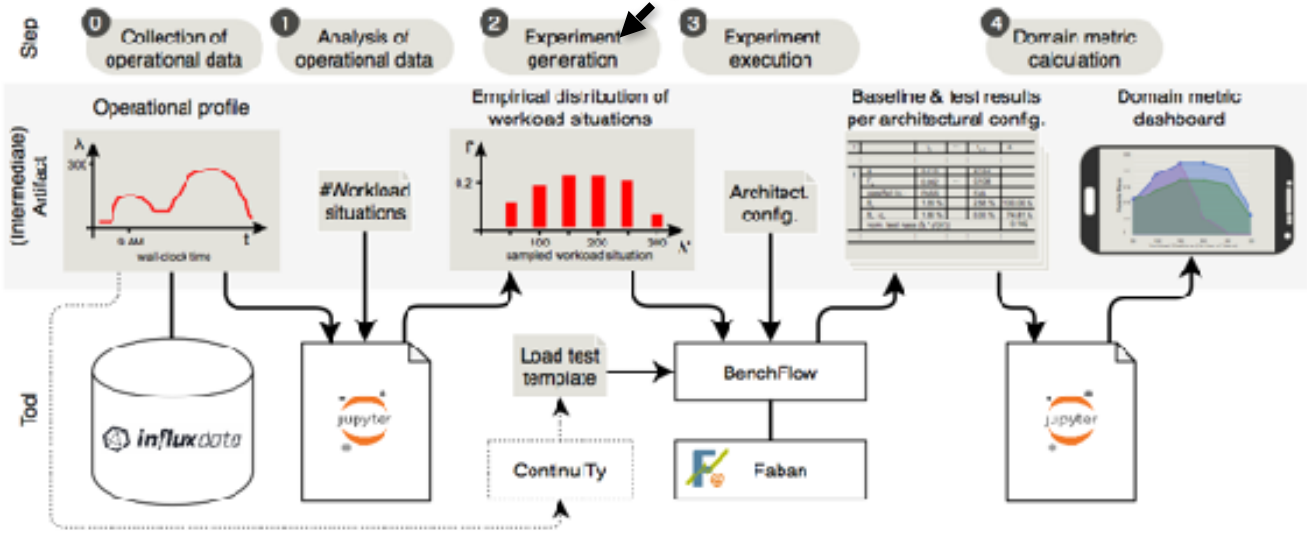
- Production and Performance Testing Based Application Monitoring



Avritzer et al.: Automated Scalability Assessment in DevOps

PPTAM

Mirai

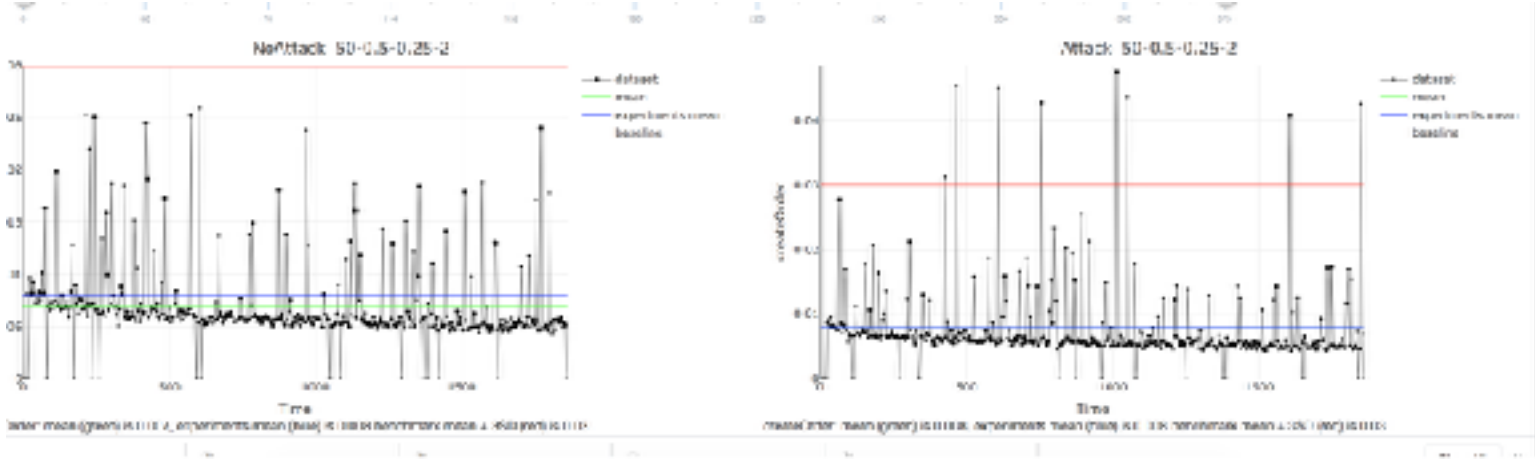


Avritzer and Russo: Operational Profile Data for Continuous Dependability Assessment in DevOps

Mirai BotNet

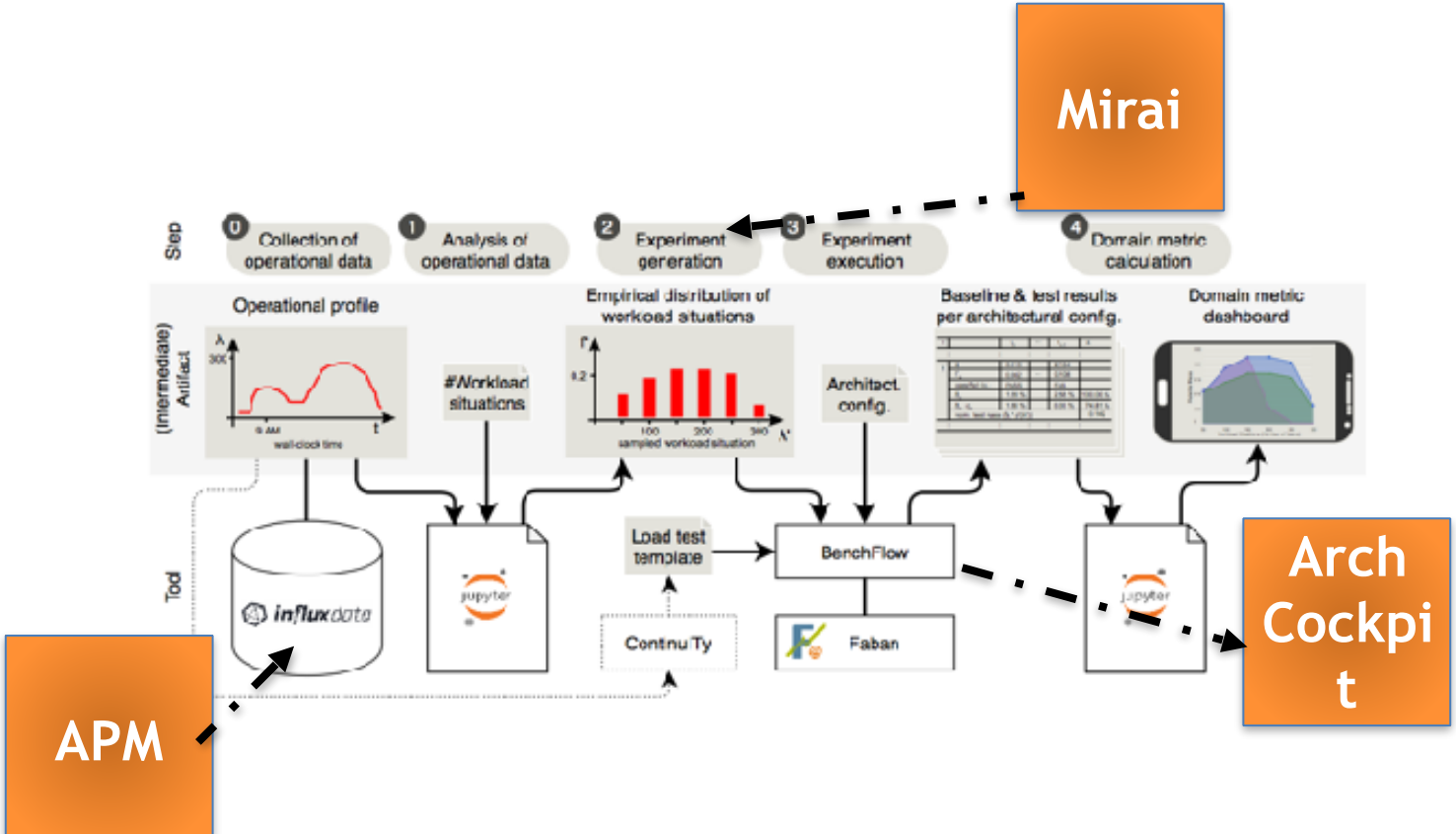
- Mirai is a malware that has been used to turn networked devices running Linux into remotely controlled bots
- We use it to attack the system. It can perform different types of attack
- by now, we have explored http, syn, ack

PPTAM, No attack and attack plots with 50 users



Identify Scalability Impacting Architecture Components

- Apply the approach to a large telecom system



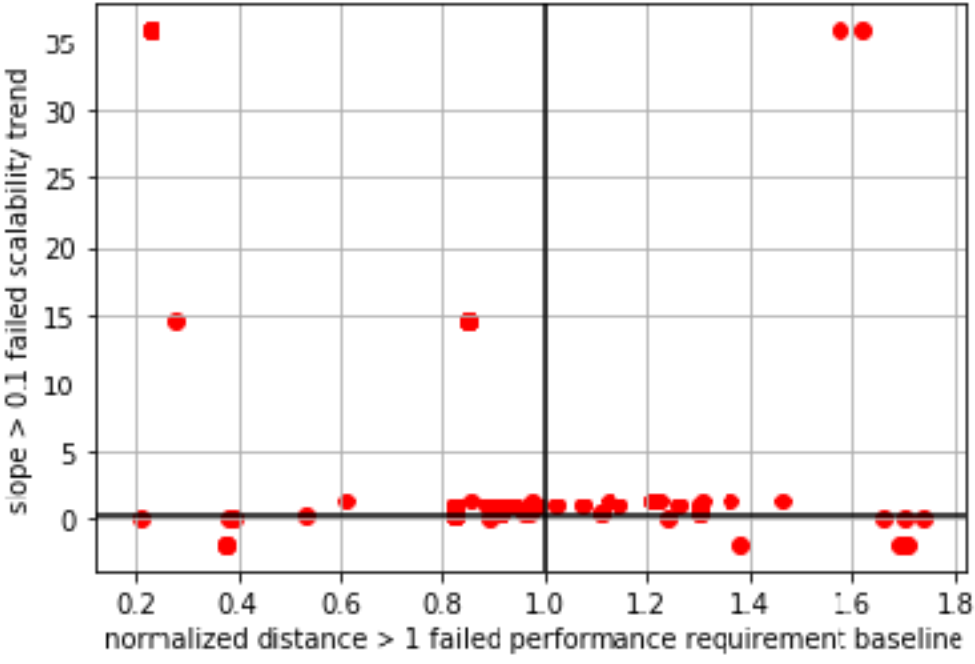
Identify Scalability Impacting Architecture Components from Performance Testing Log

Counter_name	Measurement	Load	Service_name
Response Time_X	Y	L	S
Response Time_X	Y	L	S
Response Time_X	Y	L	S
Response Time_X	Y	L	S

What is the problem with this log?

Compute Scalability Baseline, Normalized Distance from Baseline, and Linear Regression Slope

Counter_Name	Low Load_ RT	Std _RT	Baseline_RT
Response Time X	2.488602	3.513397	12.033352



Compute Scalability Baseline, Normalized Distance from Baseline, and Linear Regression Slope

Counter_Name	Low Load_ RT	Std _RT	Baseline_RT
Response Time X	2.488602	3.513397	12.033352

counter_name	ndistance	slope
Adjustment	0.37025393554043376	-1.6849766996621312
Enquiry	1.2978283298346712	0.42714945996927856
Interrogation	1.1421493733096413	0.8761959026399963
Resources Read	0.9151256042672528	0.45008251512825854
Resources Update	0.8476072668413962	14.582286376009389
Status Updates	1.5746310520634159	35.94647139406611
Control	1.2386583153386215	0.9018691142194458
DB Data Management	1.4807154532199451	1.263218473057755
Internal Communication	0.8228993410883527	0.13000903429414 65
Offline	1.734371251882527	0.011288058555376876
Online	0.9393507653154034	0.9230591413901067
Recompose	1.2213124785523563	1.3542574187250 71

Multivariate analysis used to prioritize re-factoring using Slope and Normalized Distance Profile

Summary

- DevOps provides access to real-time data, enables production and testing integration
- Use cases
  - Performance, Scalability, Security
  - other illities: survivability, resilience
- Provide detailed guidance to architects and developers in real time:
  - Domain-metric based scalability assessment dashboards
  - Prioritize component re-factoring based on load testing results

Publications

- A. Avritzer, D. S. Menasché, V. Rufino, B. Russo, A. Janes, V. Ferme, A. van Hoorn, H. Schulz, *PPTAM: Production and Performance Testing Based Application Monitoring*. ICPE Companion 2019: 39-40

- A. Avritzer, V. Ferme, A. Janes, B. Russo, H. Schulz, A. van Hoorn, *A Quantitative Approach for the Assessment of Microservice Architecture Deployment Alternatives by Automated Performance Testing*. ECSA 2018: 159-174

- A. Avritzer, D. S. Menasché, V. Rufino, B. Russo, A. Janes, V. Ferme, A. van Hoorn, H. Schulz, *Scalability Assessment of Microservice Architectural Configurations: A Domain-based Approach Leveraging Operational Profiles and Load Tests*, Journal of Systems and Software, to appear 2020.