

SAP

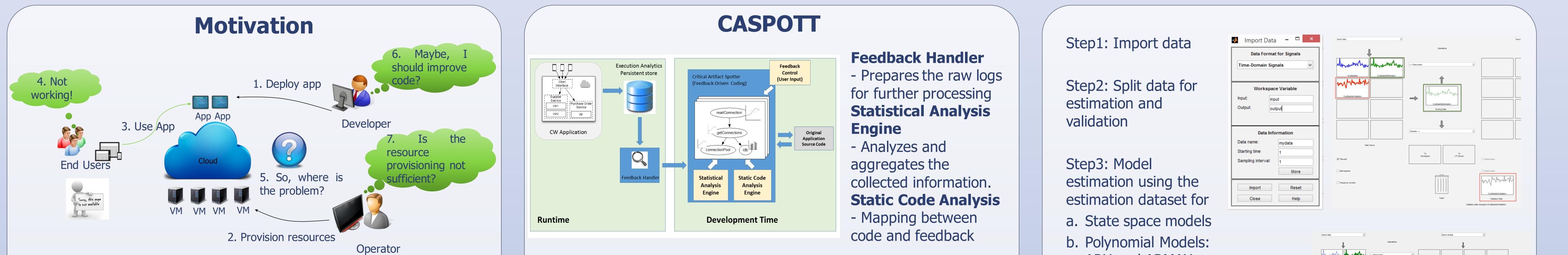
DO-AS-CASPOTT: DevOps Auto-Scaler & Critical Artifact SPOTTer Tool for Cloud Applications



Cloud Wave

Collaboration between Technical University of Darmstadt, SAP Research Germany and University of Zürich

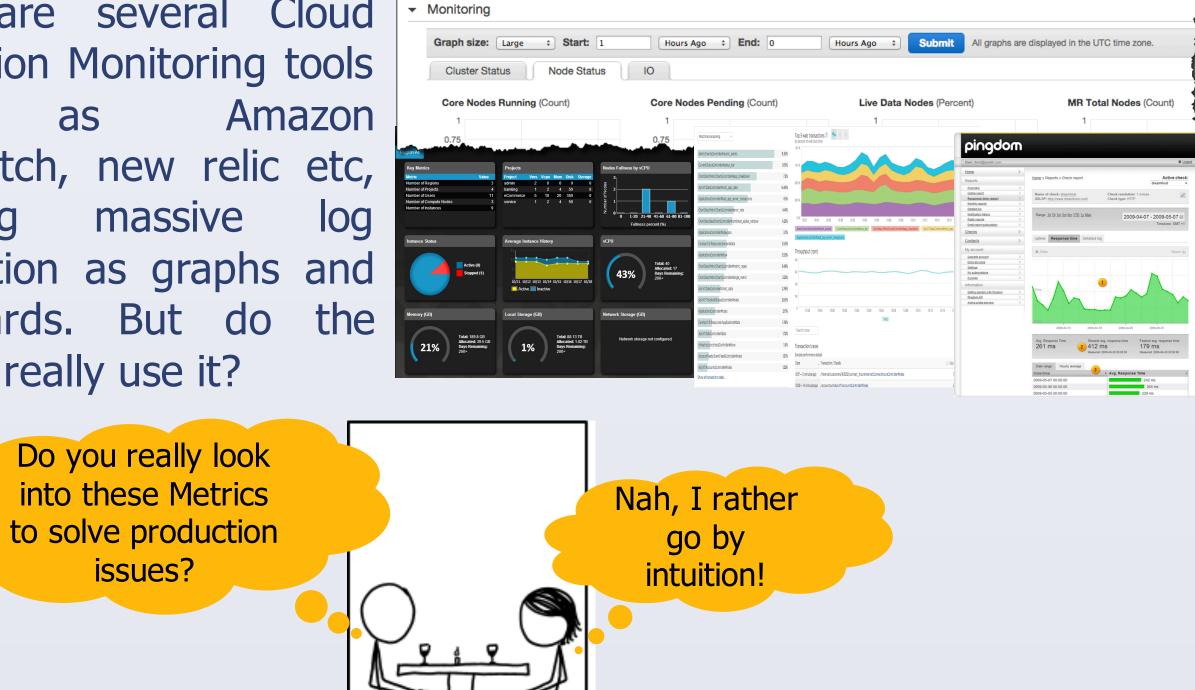
Harini Gunabalan, Technical University of Darmstadt



Auto-Scaler

Problem Statement

Cloud several are There Application Monitoring tools such Amazon as cloudwatch, new relic etc, providing massive information as graphs and dashboards. But do the DevOps really use it?



What is Feedback Driven Development?

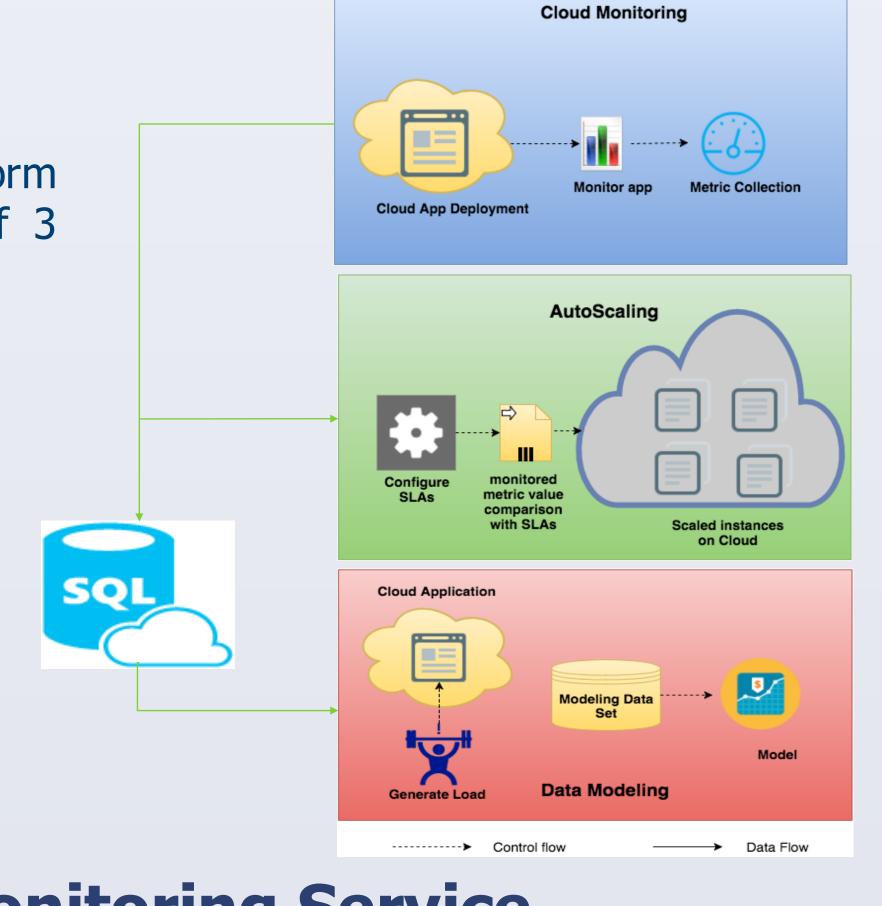
Applications running on Cloud environments

issues?



Auto-scaling at platform level that consists of 3 components:

- a. Monitoring service b. Scaling service
- c. Data modeling

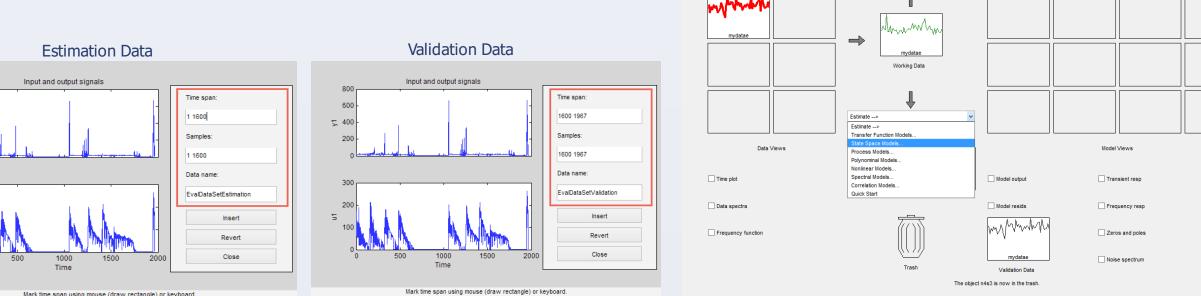


a. Monitoring Service

Data parsing by Monitoring Service

Key metrics monitored Average response time (CF

ARX and ARMAX



State Spa	ace Model estimation	ARX Model Estimation	ARMAX Model Estimation
Polynomial and State Space Mod – 🗖 🛛 🛛	Working Data	🛃 Polynomial and State Space Mod – 🗖 🗙	Polynomial and State Space Mod
tructure: State Space: n rders: 1:10 quation: xnew = Ax + Bu + Ke; y = Cx + Du + e lethod: PEM N4SID omain: Continuous Discrete (1 seconds) eedthrough: false(1,5) orm: Free put delay: zeros(5,1) ame:	Model Order Selection – – × File Options Style Help Select model order in Command Window. 7 Red: Default Choice (3) 6.5 6.5 6.5 6.5 6.5 6.5 6.5 6.5	Structure: ARX: [na nb nk] Orders: [4 [4 4 4 4 4] [1 1 1 1 1]] Equation: Ay = Bu + e Method: Image: ARX Domain: Continuous Imput delay: Zeros(5,1) Name: arx441	Structure: ARMAX: [na nb nc nk] Orders: [2, [2 2 2 2 2], 2, [1 1 1 1 1]] Equation: Ay = Bu + Ce Method: Prediction error method Domain: Continuous
ocus: Prediction Initial state: Auto ✓ ist.model: Estimate K Covariance: Estimate ✓	- ^{Close} 4.5 4	Focus: Prediction Initial state: Auto Dist.model: Estimate Covariance: Estimate	Focus: Prediction Initial state: Auto Dist.model: Estimate Covariance: Estimate
Display progress Stop iterations	0 5 10 15 Model order	Display progress Stop iterations	Display progress Stop iterations
Order Selection Order Editor		Order Selection Order Editor	Iteration Options Order Editor
Estimate Close Help		Estimate Close Help	Estimate Close Help

Evaluation and Results



> 🗉 🔟 🔅 🔻 🜔 🕶 🎭 📲 🖏 🕶 😚 🕶 🎒 🗀 🖋 📲 🚱 🖧 🖢 🖛 🖓 🖛

Auto-scaler evaluated on a Guestbook app. Load generated using Jmeter.

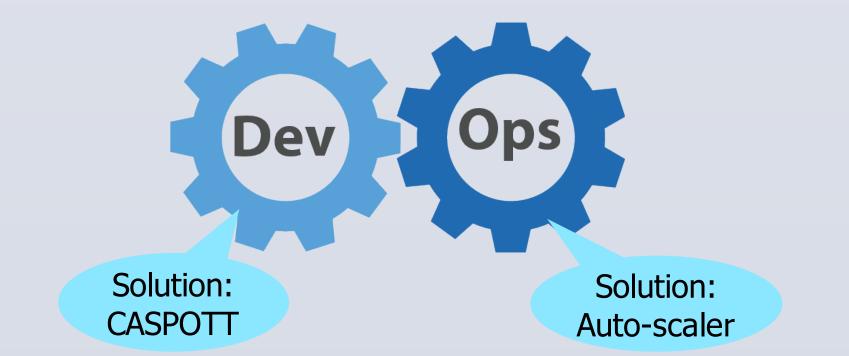
$\Rightarrow \bullet \bullet$	Quick Access 🛛 😰 😰 Java EE			🔒 https://mastert	rthesisdemo-d063995.cfapps.sap.hana.onde	
		Outline 🔀 🗐 Task List			Enter your name:	Eı
	*	s 🗖 🛃 🕅	ξ δ δ δ ¹ ⊂			
		cloudmore.cw.auth			Enter your city:	
		F log : Logger				
		dto: AuthorizeRegDTO			Leave a comm	ent or su

nastert	hesisdemo-d063995.cfapps.sap.h	ana.ondemand.com	숬		
	Enter your name:	Enter vour e-mail address:			Image: Construction of the state of the
				SDesiç	3 applications:
	Enter your city:	Enter your country:		ister]	5 memory: 256M
					6 instances: 1 7 path: masterthesisdemoapp.war
	Leave a comm	ent or suggestion:		r = /1.1c	8

run-time generate Making data. production feedback this run-time available to DevOps in an Problem: How to accessible format to adapt accordingly, is known as Feedback Driven ime information Development (FDD).

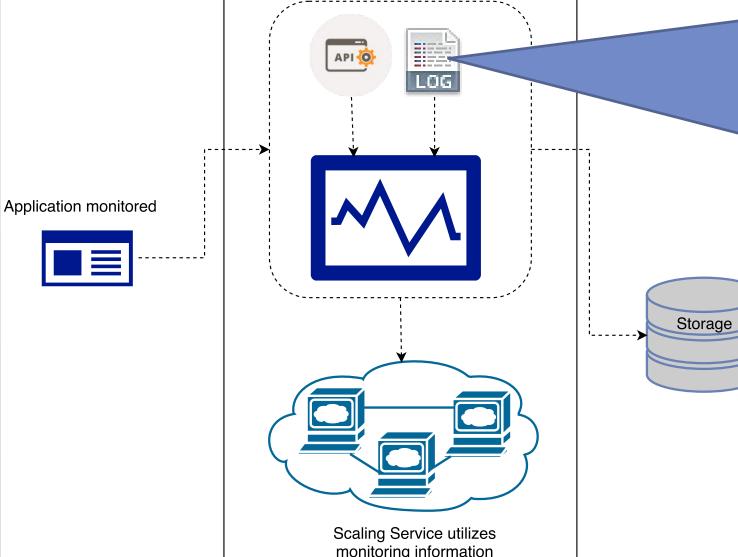
Problem: How to automate Resource identify Source Code provisioning using ssues (root cause the Run-time analysis) using Runinformation

Proposed Solution: DO-AS-CASPOTT



Solution: Combination of plug-ins that support the Developer (CASPOTT) and Operator (Auto-scaler)

Design of DO-AS-CASPOTT

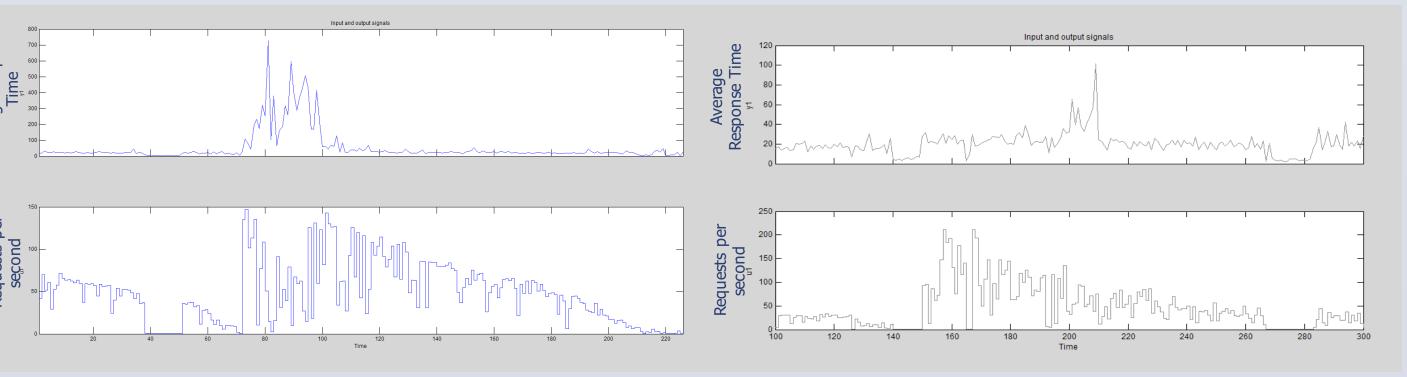


Logs) Number of incoming requests (CF Logs) Number of running app instances (CF API) CPU utilization (CF API) Memory utilization (CF API) • Disk utilization (CF API)

<pre>private AuthorizeReqDTO dto; private AuthorizeReqDTO dto; private Boalean isServiceLimited; private AuthConf conf=ApplicationContex public AuthorizeReq(AuthorizeReqDTO ndt setDto(ndto); if(ndto.getEp()==null getEp().isEm if(ndto.getEp()==null getEp().isEm public List<string>getEp(){ getEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); setEp(); se</string></pre>	 isServiceLimited: isoolean conf: AuthorizeReq(AuthorizeReqDTO) getEp): List<string></string> getEp): List<string< li=""> getGrantType(): String getEpCount(): int hasAudio(): boolean hasVideo(): boolean isAudioOholy(): boolean </string<>			
🗔 Properties 📮 Console 🛱 Timeline 📷 Operator List 🚦	Problems 🗟 CASPOT	T.plv 🛛		
Showing 11 item(s)				
Artifact Name	Load size	Number of Errors	Response Time	
AuthActive.run()	hActive.run() 53 2 0		0	
AuthorizeReq.getConf()	1.001.000	5	185	
AuthorizeReq.isVideoOnly()	22	4	0	
AuthorizeReq.isAudioOnly()	106	1		
AuthorizeReq.getEp() 5.885.880 12 256		256		
AuthorizeReq.getDto()	5.885.880	11	193	
AuthorizeReq.getGrantType()	45	4	1	
AuthActive.stop()	57	1	0	
AuthorizeReg.hasVideo()	642	14	0	

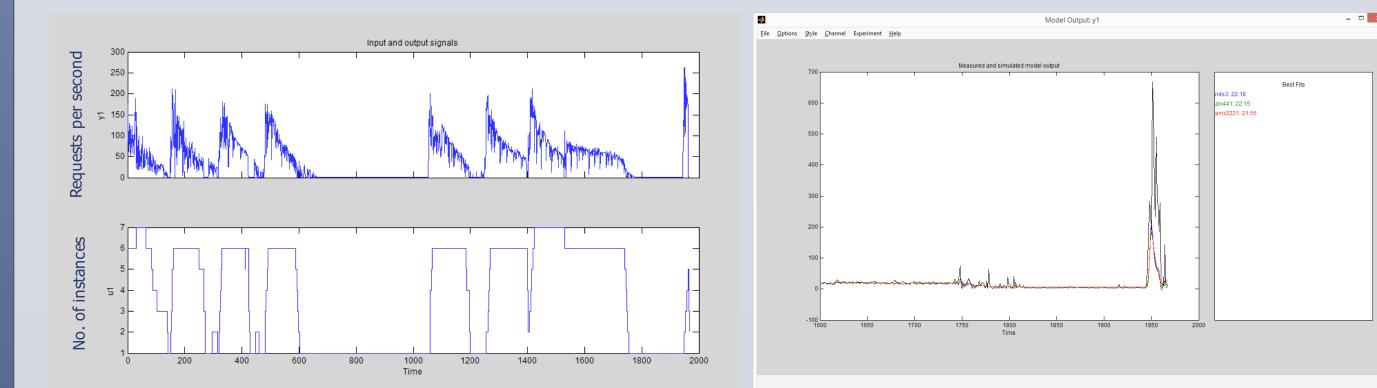
Without auto-scaler, Maximum response time is about 750 ms.

With auto-scaler, Maximum response time is about 100 ms.

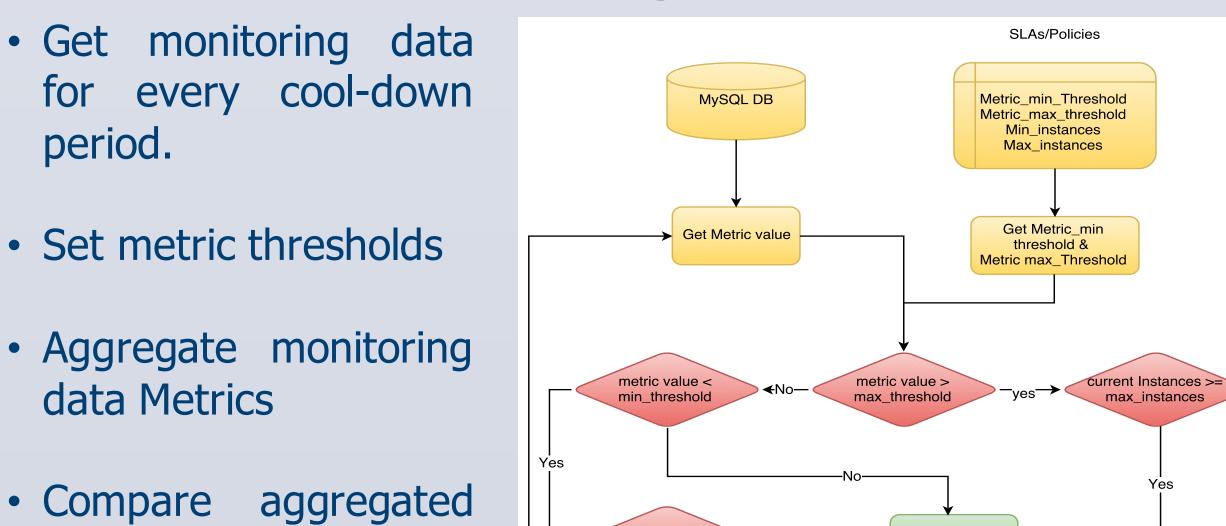


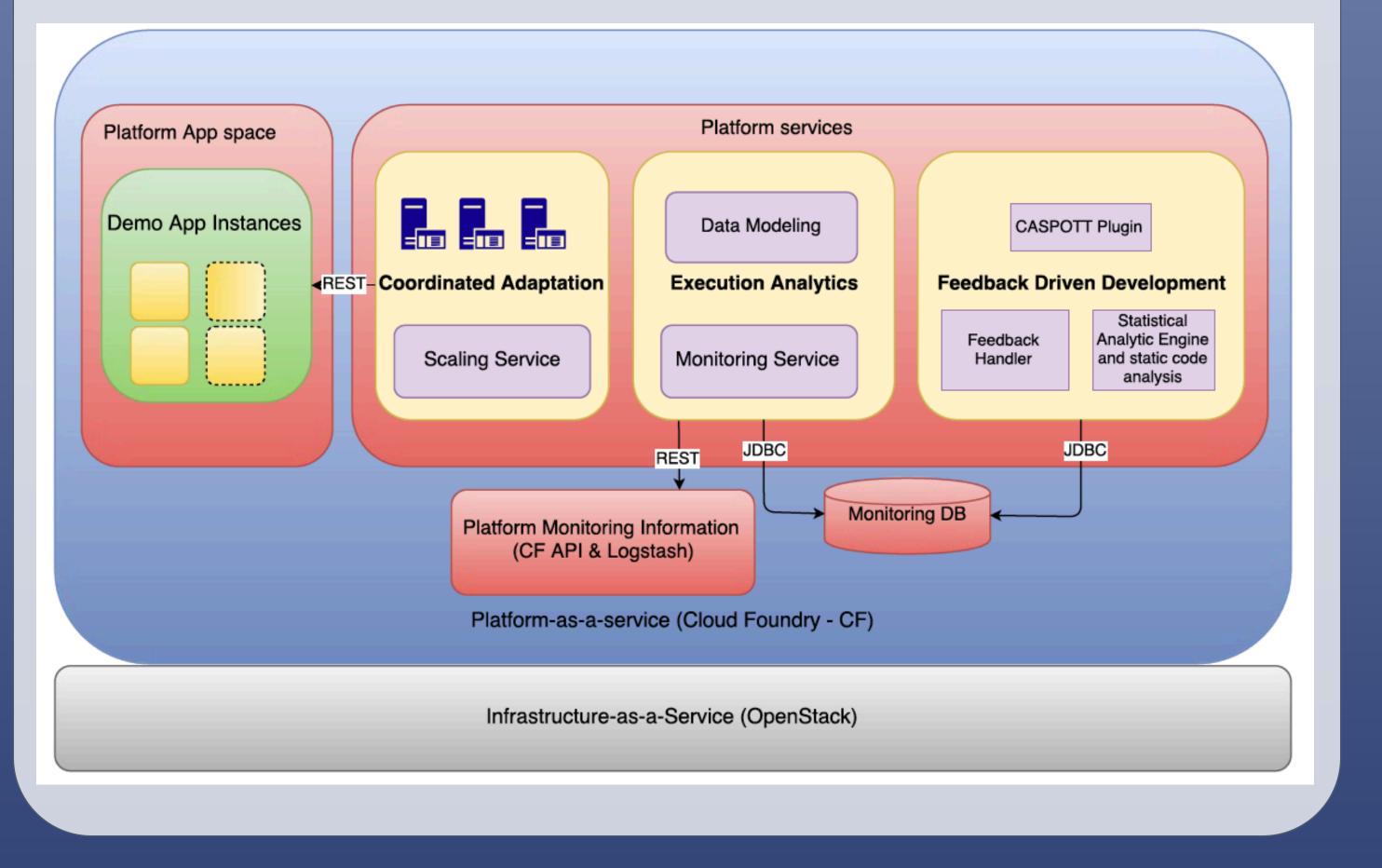
Graph shows load versus number of application instances running.

Model estimation output for State space, ARX and ARMAX.









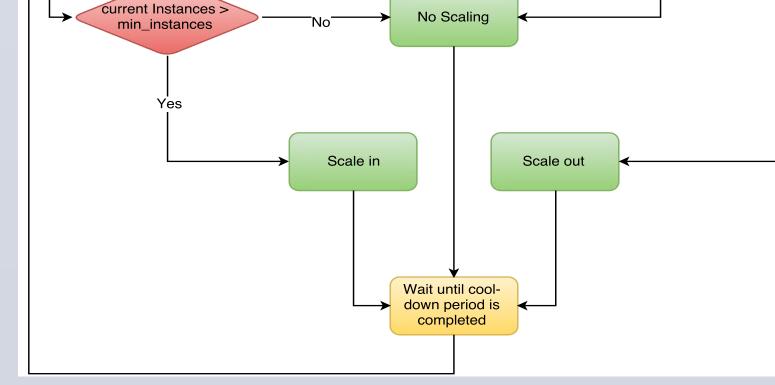
value of the metrics with the thresholds.

• If aggregated value >

threshold, then scale

maximum

out.



• Else if aggregated value < minimum metric threshold, then scale in.

c. Data Modeling

Estimate the following models for the data collected using the MATLAB System Identification Toolbox:

- State-space Model
- Polynomial Models: ARX and ARMAX

metric

References

- 1. Harini Gunabalan. 2016. Feedback Driven Development of Cloud Applications. Master's Thesis. Technical University of Darmstadt, Germany.
- 2. Cito, Jürgen, et al. "Runtime metric meets developer: building better cloud applications using feedback." 2015 ACM International Symposium on New Ideas, New Paradigms, and Reflections on Programming and Software (Onward!). ACM, 2015.

Acknowledgements

This work has been developed partially in the context of the EU-funded CloudWave project (grant 610802), a research project that is enabling a next generation of cloud infrastructure operations and agile development for their hosted applications.



Harini Gunabalan, <u>https://harinigunabalan.github.io/</u> Project webpage: <u>https://harinigunabalan.github.io/views/thesis.html</u> Master Thesis Supervisors: Prof. Dr. –Ing. Mira Mezini, Dr. –Ing. Guido Salvaneschi