



Red Hat Dallas Emerging Tech Summit

December 5, 2019

AI Ops in a cloud-native world

Steven Huels, AI Center of Excellence Director

Introductions



Steven Huels
AI Center of Excellence Director
Based in Raleigh

Steven is the Director of the Red Hat AI Center of Excellence with responsibility for Red Hat's AI strategy and how that impacts internal use cases, products, customers, and partner ecosystems.

Steven has been involved with real-time data management and analytic platforms for over 15 years and has a keen interest in large-scale data problems and using analytics and machine learning to solve those problems.



“By 2022, **50%** of IT assets in enterprise datacenters will have the **ability to run autonomously using embedded AI** functionality that leverages smart IT and facilities systems.”

Common Foundations

AI Ops: Artificial
Intelligence for IT
Operations

!=

ML Ops: Machine
Learning
Operations

The cultural change we
saw brought about by

Dev & Ops = DevOps

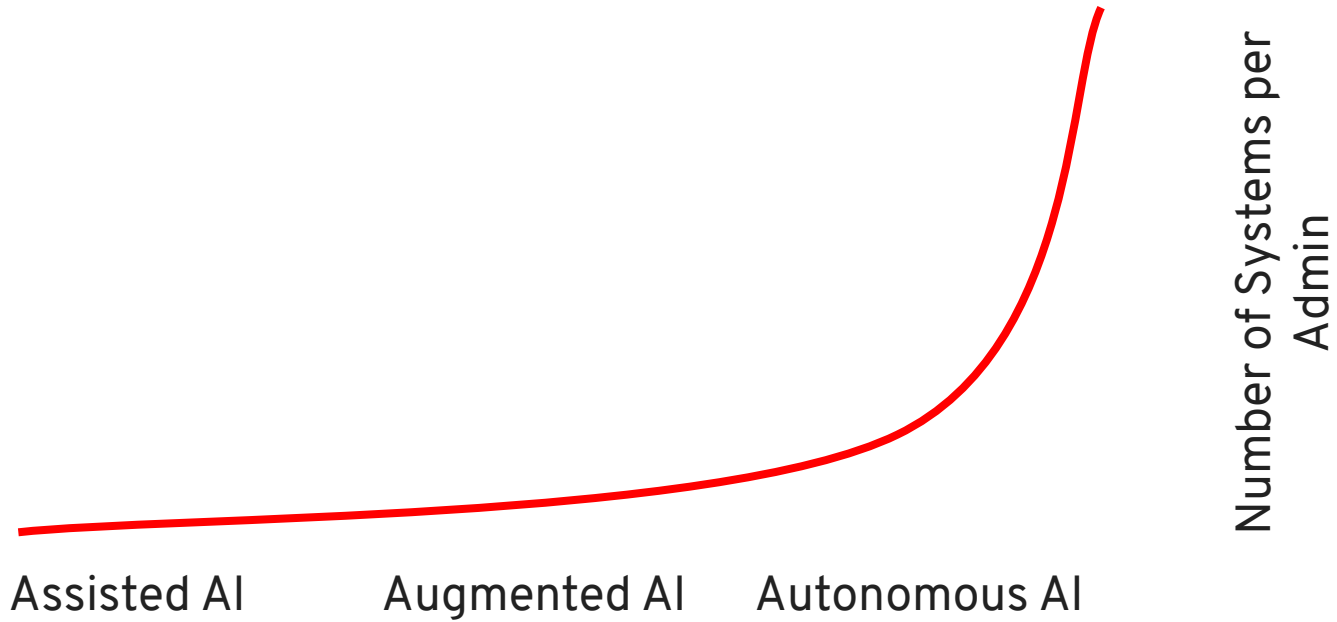
Is being repeated in the
combination of

AI + DevOps = AIOps

AI Ops platforms are software systems that combine big data and **AI** or machine learning functionality to enhance and partially **replace** a broad range of **IT operations** processes and tasks, including availability and performance monitoring, event correlation and analysis, IT service management, and automation.

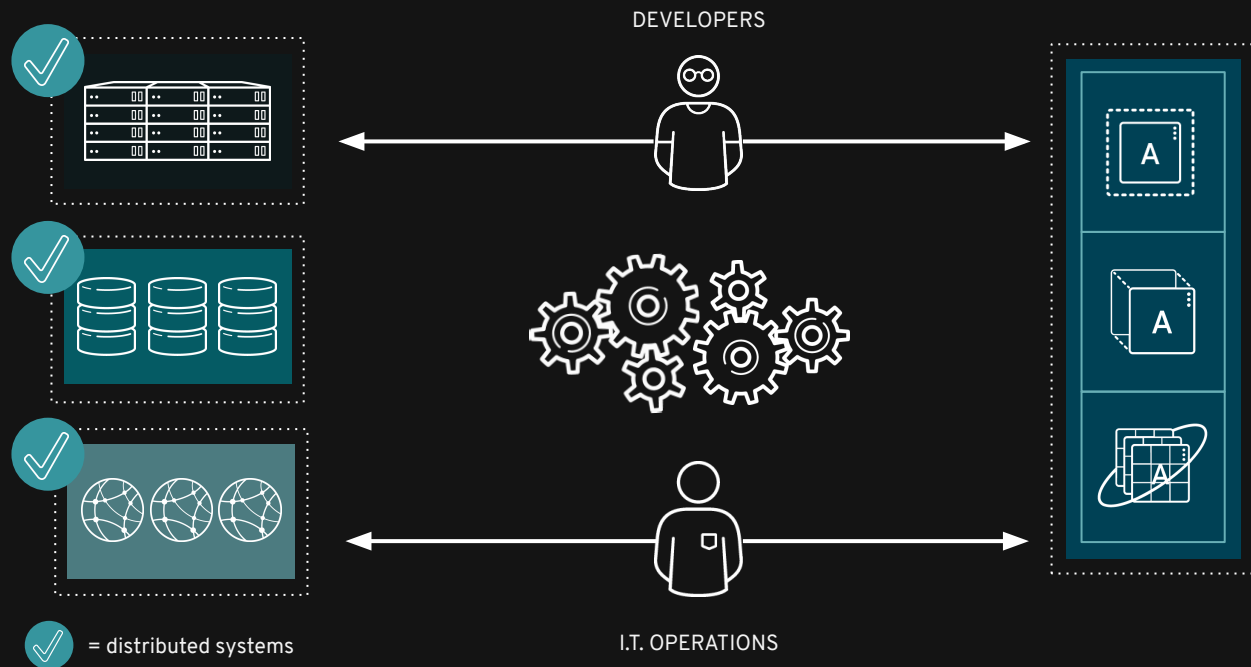
Source: Gartner Market Guide for AIOps Platforms Published: 03 August 2017 ID: G00322184

Increase Efficiency with AI

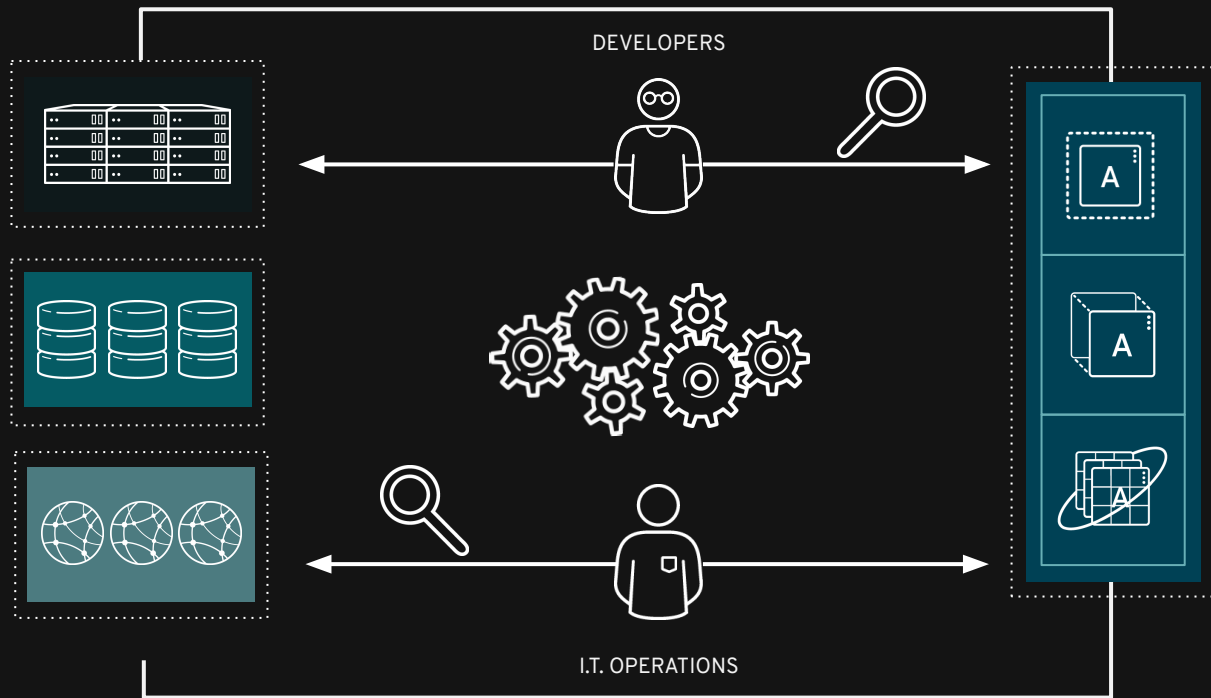


SCALABLE DISTRIBUTED SYSTEMS

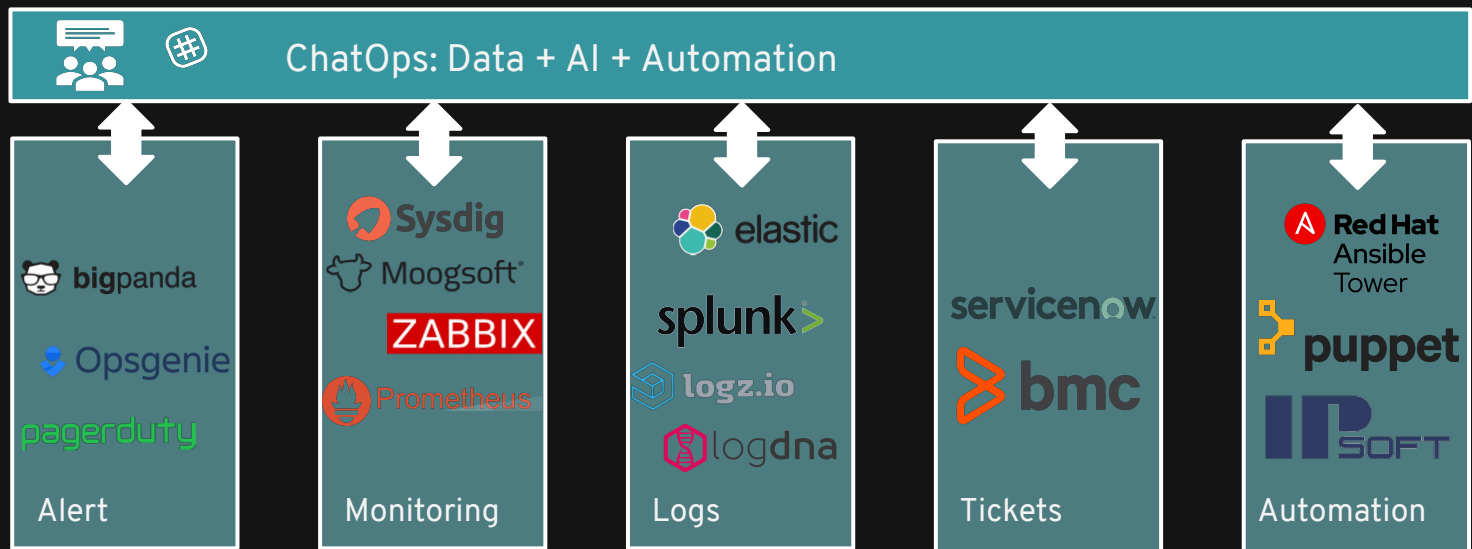
SUPPORT EFFICIENT SEPARATION OF DUTY



TELEMETRY & DIAGNOSTICS REQUIRED DISTRIBUTED SYSTEMS ARE COMPLEX

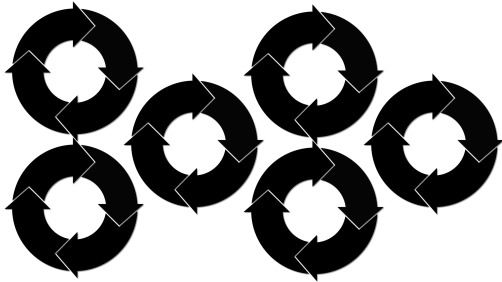


Data, AI and Automation



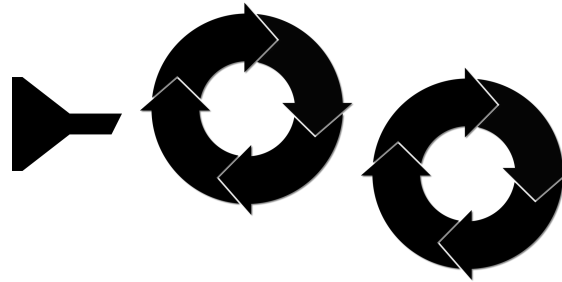
Iterative Approach to Implementing Data Science

EXPERIMENT



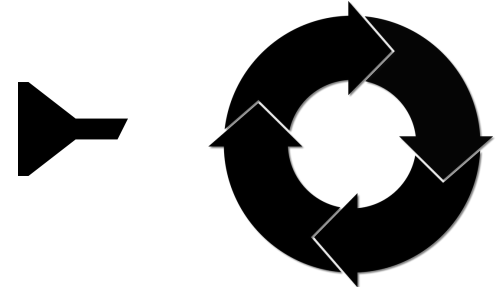
- Rapid iteration & ad-hoc analysis
- Focus on ideation
- Little concern for reproducibility
- Introduction of new technology (stacks)
- Access to pseudonymized or anonymized data

OPTIMIZE & TUNE



- Promising experiments promoted
- Focus on performance, accuracy, repeatability, and auditability
- Concerned with data governance, provenance, and security
- Employs approved technology (stacks)
- Access to production data

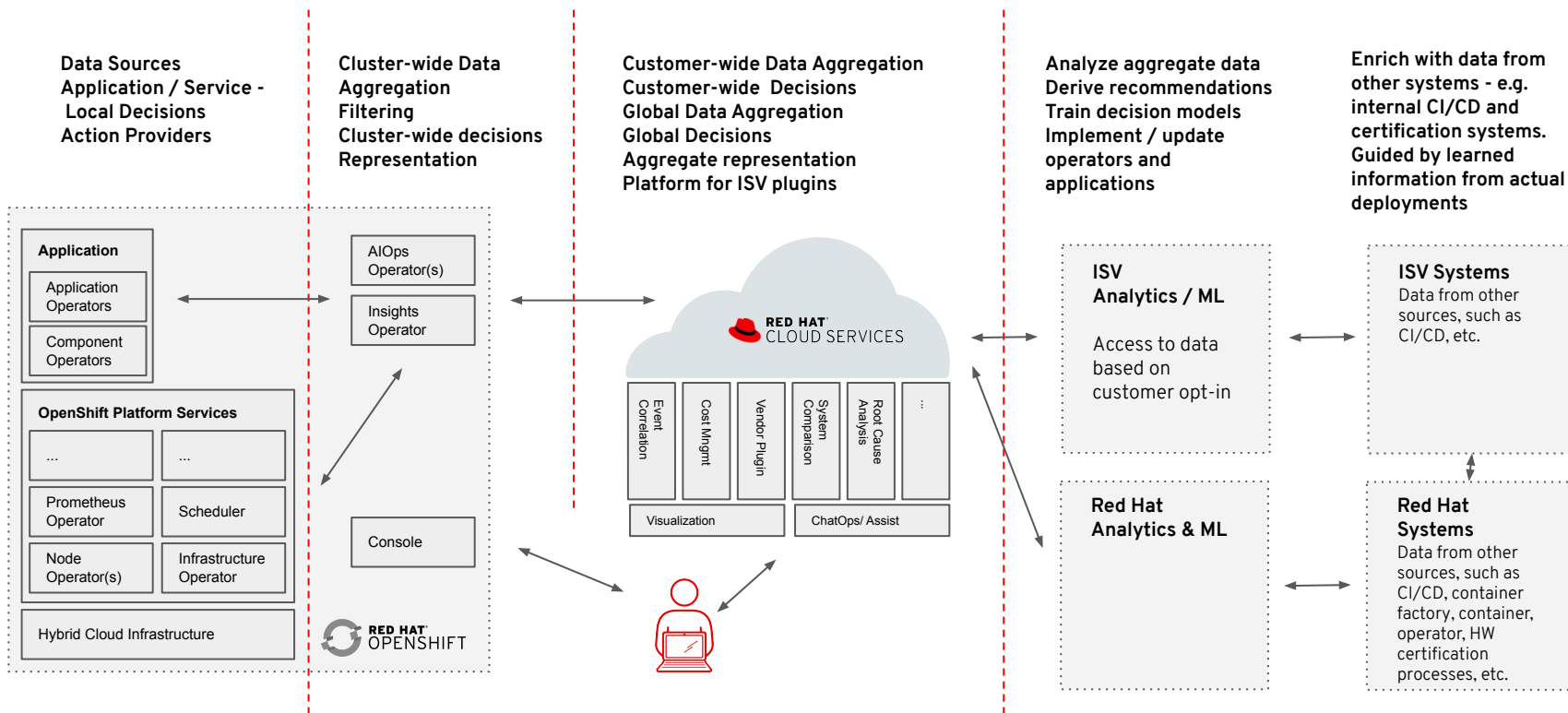
OPERATIONALIZE



- Optimized routines promoted
- Focus on ops (automated deployment/roll-back, workflow, auditing of results)
- Archive old models, decisioning, and automate training and deployment
- Audited access to production data

Red Hat is implementing and enabling AI Ops

AI Ops Architecture





Home

Catalog

Workloads

Networking

Storage

Builds

Monitoring

Administration 

Cluster Settings

Namespaces

Nodes

CLUSTER AUTOSCALER

 [Create Autoscaler](#)

[Update](#) 

Cluster Settings

[Overview](#)

[Global Configuration](#)

[Cluster Operators](#)

CHANNEL	UPDATE STATUS	CURRENT VERSION
fast	4.1.0-0.2	4.0.0-0.2

CLUSTER ID

784ce289-02aa-4d32-8796-cd4a0619499c

CURRENT PAYLOAD

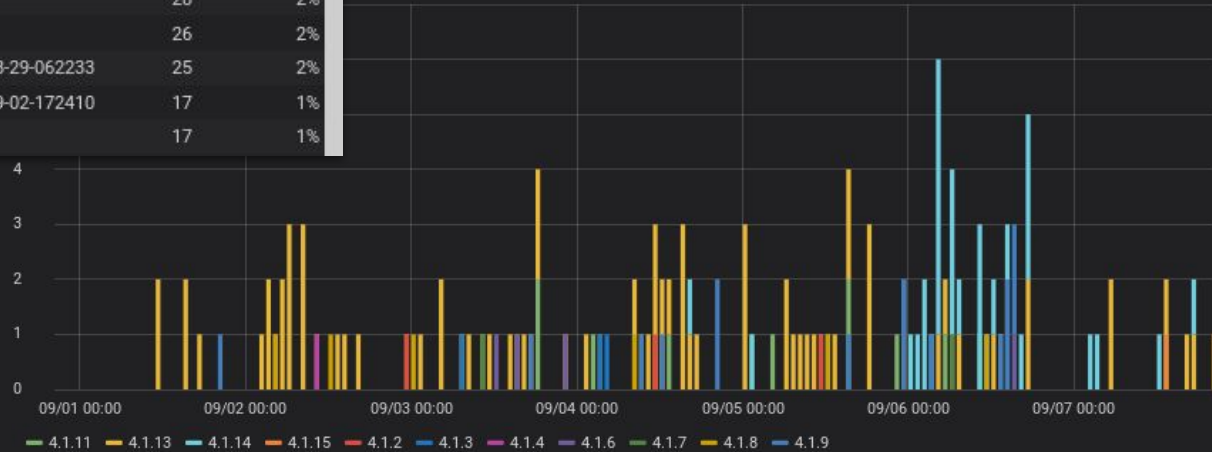
-

CLUSTER AUTOSCALER

 [Create Autoscaler](#)

[Update](#) 

	current ▾	percentage ▾
4.1.13	231	20%
4.1.11	126	11%
4.1.8	116	10%
4.1.9	100	9%
4.1.14	80	7%
4.1.0	64	6%
4.1.4	44	4%
4.2.0-0.nightly-2019-09-04-142146	30	3%
4.1.7	28	2%
4.1.6	28	2%
4.1.2	26	2%
4.2.0-0.nightly-2019-08-29-062233	25	2%
4.2.0-0.nightly-2019-09-02-172410	17	1%
4.1.3	17	1%



Cluster ID 55552ddc-2d33-48ec-a36f-e90a0073f453 | datasource | thanos

Cloud Provider Type	Install Type	Legacy Type	Version	Updating To
None	UPI	UPI	4.1.11	4.1.14

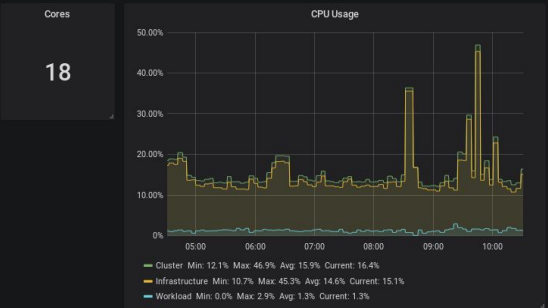
E-mail Domain	Support	Interest Score	Falling For	Updating For
redhat.com	None	2	4.0 min	2.093 hour

RH Org ID	EBS Account	Nodes	Etcd objects	API Error Rate															
11009103	5910538	6	5458	<table border="1"><thead><tr><th>version</th><th>current</th><th>percentage</th></tr></thead><tbody><tr><td>500</td><td>0.0018</td><td>0%</td></tr><tr><td>503</td><td>0.6</td><td>57%</td></tr><tr><td>403</td><td>0.004</td><td>0%</td></tr><tr><td>404</td><td>0.19</td><td>17%</td></tr></tbody></table>	version	current	percentage	500	0.0018	0%	503	0.6	57%	403	0.004	0%	404	0.19	17%
version	current	percentage																	
500	0.0018	0%																	
503	0.6	57%																	
403	0.004	0%																	
404	0.19	17%																	

Age	Since last completed update	Masters	Operators by version								
9.01 week	1.033 week	3	<table border="1"><thead><tr><th>version</th><th>Value</th></tr></thead><tbody><tr><td>4.1.14</td><td>20</td></tr><tr><td>4.1.11</td><td>3</td></tr><tr><td>4.1.13</td><td>1</td></tr></tbody></table>	version	Value	4.1.14	20	4.1.11	3	4.1.13	1
version	Value										
4.1.14	20										
4.1.11	3										
4.1.13	1										

Accessing Insights Data

```
To see the insights data for this cluster, configure supporttel and run  
ssh supportshell1.prod.useraccess-us-west-2.redhat.com \  
-t "cd /ocp/11000103/55552ddc-2d33-48ec-a36f-e90a0073f453/  
bash --login"
```



Degraded Operators

condition	name
Available	storage

Alerts

alertname	alertstate
KubeAPIErrorsHigh	firing
KubeClientErrors	firing
KubeDeploymentReplicasMismatch	firing
Watchdog	firing

Down Nodes

No data to show

Down Services

No data to show

Harnessing the power of AI Ops for your needs...

Detecting Cluster Degradation

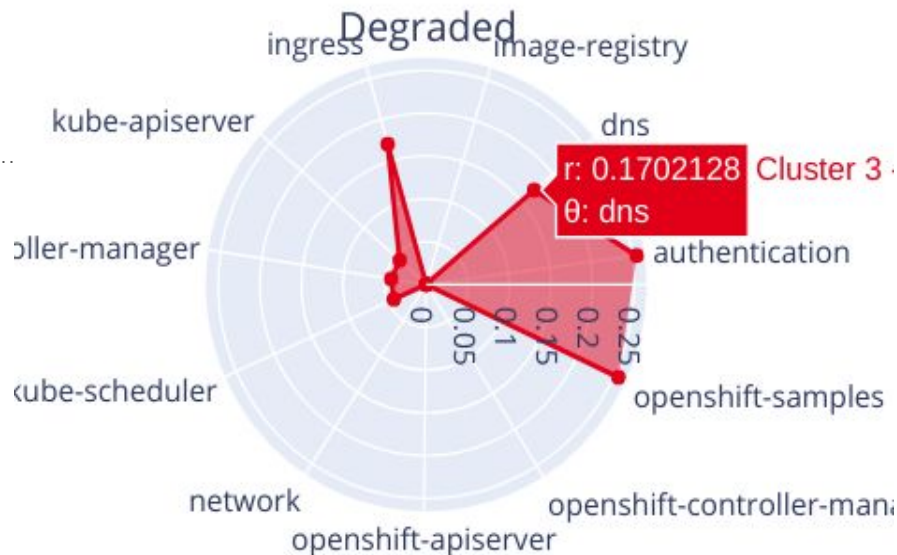


OCP4 Pro-Active Support

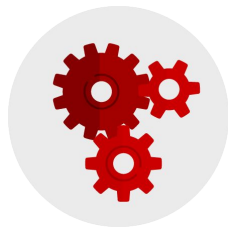
- Clustering of OCP4 deployments
- Guide support to similar issues & resolutions
- Gating for releases

Artifacts

- <https://github.com/AICoE/prometheus-flatliner>



Constructed Using Smaller Building Blocks



Prometheus Metrics Processing

- Stream processing of Prometheus data
- Predict time series
- Alert on anomalies
- Aggregated scoring based on metrics



Artifacts

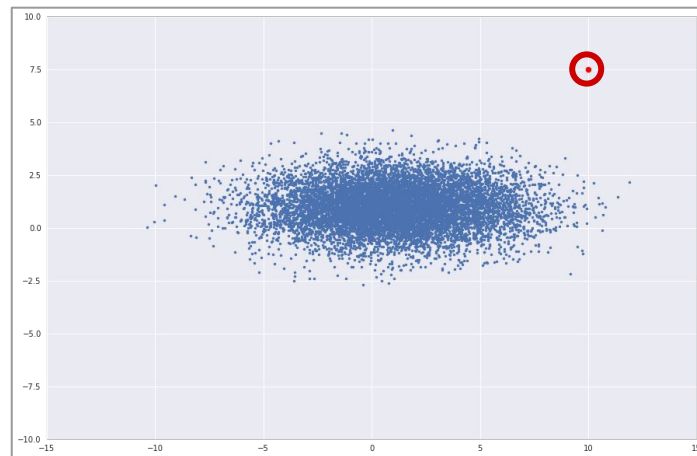
- <https://github.com/AICoE/prometheus-anomaly-detector>
- <https://github.com/AICoE/prometheus-api-client-python/>
- <https://github.com/AICoE/prometheus-flatliner>

Let's move on to
something a bit more
advanced...

Log Anomaly Detection

Motivation

- There is **too much data** to have a human manually search through logs
- Our internal customer wants **tooling** that will assist in **root cause analysis**
- We think there is lots of potential to use **unsupervised machine learning** in **identifying anomalies**
- We would like to have the ability for our **users to give feedback** so we can **improve our models**



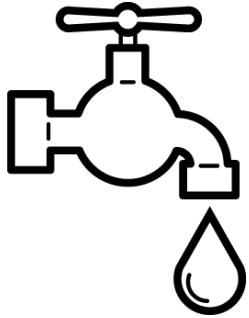
<https://github.com/AICoE/log-anomaly-detector>

Overcoming Common Misconceptions

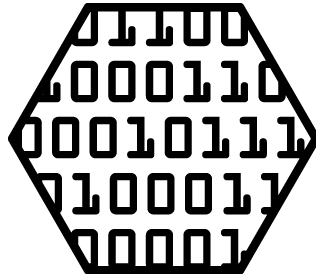
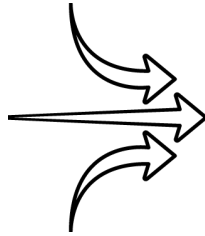
Error messages are a reliable indicator of issues

Timestamp	Severity	Message
2015-10-17 15:37:56,900	INFO	dfs.DataNode\$DataXceiver: Receiving block blk_5102849382819239340 src /10.0.0.1:47915 dest /10.0.0.1:50010
2015-10-17 15:37:57,036	INFO	dfs.FSNamesystem: BLOCK* NameSystem.allocateBlock: /user/root/foo/_temporary/_task_200811101024_0014_m_001575_0/part-01575.blk_5102849382819239340
2015-10-17 15:37:57,634	INFO	dfs.DataNode\$DataXceiver: Receiving block blk_5102849382819239340 src: /10.0.0.1:57800 dest: /10.0.0.1:50010
2015-10-17 15:37:57,720	INFO	dfs.DataNode\$DataXceiver: writeBlock blk_5102849382819239340 received exception java.io.IOException : Could not read from stream

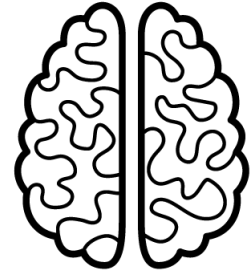
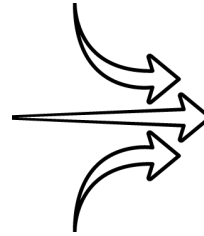
Behind the Scenes



Elastic Search



Word2Vec



**Self Organizing
Map**

Validating the Results and Improving the Model



ElastAlert via redhat.com

to me, mshah, ynanavat, jkaluza ▾

We found the following message in your logs:

message: connecting to amqps://messaging-devops-broker02.web.prod.ext.phx2.redhat.com:5671

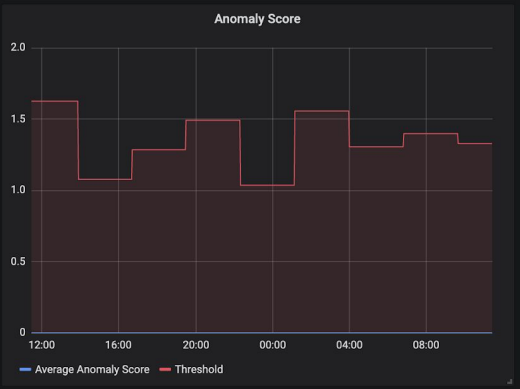
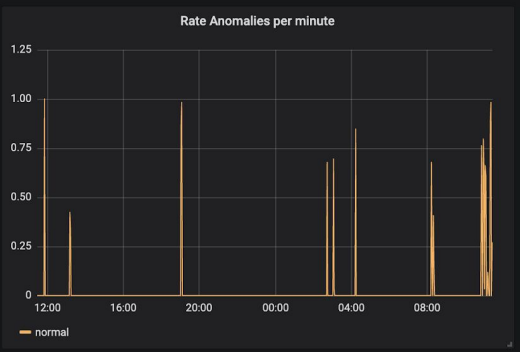
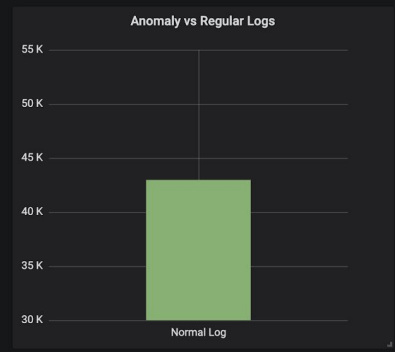
If this is not correct and you would like to report it you can provide feedback with the following link:

http://log-anomaly-detector-factstore-aiops-prod-log-analysis.cloud.paas.psi.redhat.com/?lad_id=AWtKwWIDja48Ks3No14U

- Ask the user to validate whether this was this a false prediction or correct.
- With this feedback can we improve our ML model.
- We prevent our user from receiving the same false-positive notification more than once.

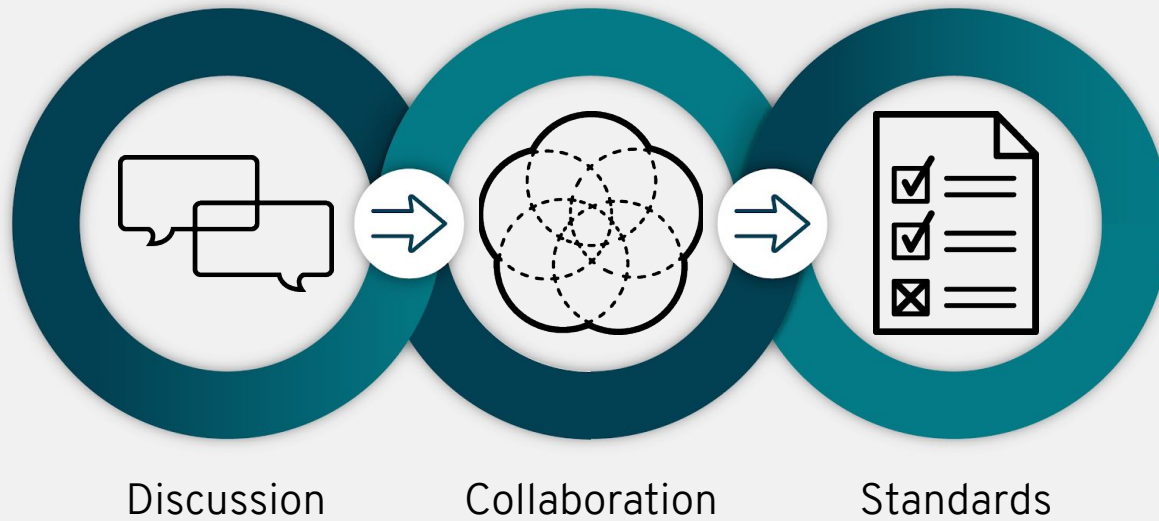
job alops-log-anomaly-detector-greenwave

<p>LAD - Up</p> <p>1</p> <p>Current Log Entries Count</p> <p>60</p>	<p>Percentage of non ano...</p> <p>100%</p> <p>Feedback Count</p> <p>0</p>	<p>LAD FactStore - Up</p> <p>1</p> <p>Total Anomaly</p> <p>0</p>
<p>Total Training Count</p> <p>11</p>	<p>Threshold Value</p> <p>1.3</p>	<p>Total Inference Count</p> <p>1788</p>
<p>False Positive Count</p> <p>483</p>	<p>Feedback Error Count</p> <p>0</p>	



Requests per second

AIOps Community Success



STAY ENGAGED

Developers.redhat.com

Your access point for no-cost developer tools and product subscriptions, how-tos, and demos

Red Hat User Groups

Meetups for networking and tech deep dives
www.meetup.com/Dallas-Red-Hat-Users-Group/

DevNation


Virtual and live events
Catch replays at
<https://developers.redhat.com/devnation/>


Next.redhat.com

Stay in touch with the Office of the CTO

Thank you

Red Hat is the world's leading provider of enterprise open source software solutions. Award-winning support, training, and consulting services make Red Hat a trusted adviser to the Fortune 500.

 [linkedin.com/company/red-hat](https://www.linkedin.com/company/red-hat)

 [youtube.com/user/RedHatVideos](https://www.youtube.com/user/RedHatVideos)

 [facebook.com/redhatinc](https://www.facebook.com/redhatinc)

 twitter.com/RedHat