FN1916 - Splunk Machine Learning and Self Healing at Priceline

Forward-Looking Statements

During the course of this presentation, we may make forward-looking statements regarding future events or plans of the company. We caution you that such statements reflect our current expectations and estimates based on factors currently known to us and that actual events or results may differ materially. The forward-looking statements made in the this presentation are being made as of the time and date of its live presentation. If reviewed after its live presentation, it may not contain current or accurate information. We do not assume any obligation to update any forward-looking statements made herein.

In addition, any information about our roadmap outlines our general product direction and is subject to change at any time without notice. It is for informational purposes only, and shall not be incorporated into any contract or other commitment. Splunk undertakes no obligation either to develop the features or functionalities described or to include any such feature or functionality in a future release.

Pranav Nandedkar
Software Engineer

Mukund Murthy
Software Engineer

splunk> .conf19

Agenda

- About Priceline
- Splunk + MLTK
- Legacy vs Current Architecture
- Numeric Outlier Detection on Errors Percentage
- Custom Apps/Alert Actions.
- Detecting KPI Performance Degradation
- Detecting New Errors – using clustering
- Takeaways
- Useful links & Resources
- Q&A

About Priceline

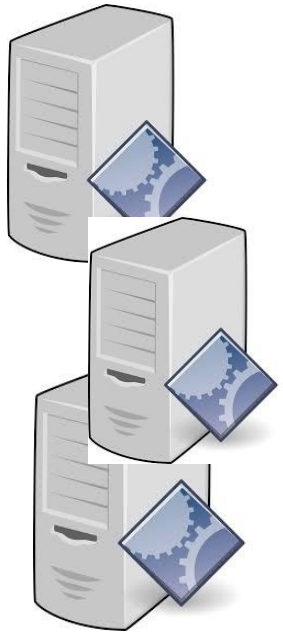
- Priceline.com is part of the Booking Holdings Group.
- The Booking Holdings Group is the world leader in online travel & related services.
- The Group's mission is to help people experience the world.

priceline.com®

Legacy vs Current Architecture

splunk> .conf19

Legacy Architecture



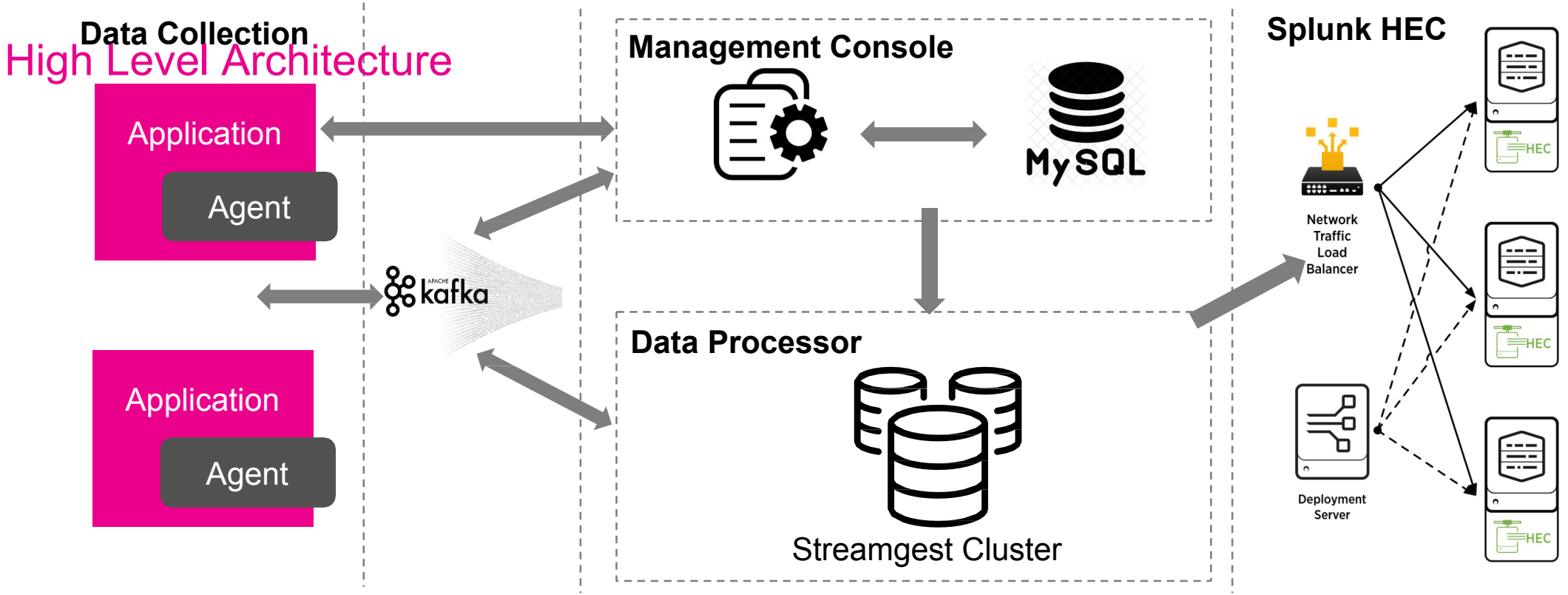
Application Servers

Custom Dashboards

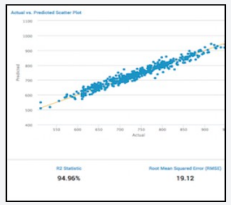
Custom Alerting System

Kibana Dashboards

Priceline Data Collection Platform



Splunk Machine Learning Toolkit (MLTK)



Predict Numeric Fields

Predict the value of a numeric field using a weighted combination of the values of other fields in that event. A common use of these predictions is to identify anomalies: predictions that differ significantly from the actual value may be considered anomalous.

Examples

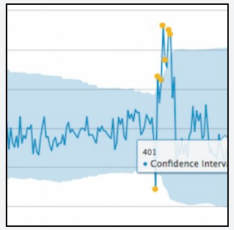
- Predict Server Power Consumption
- Predict VPN Usage
- Predict Median House Value
- Predict Power Plant Energy Output
- Predict Future Logins
- Predict Future VPN Usage (sinusoidal time)
- Predict Future VPN Usage (categorical time)
- Predict Future VPN Usage (categorical time)

Detect Categorical Outliers

Find events that contain unusual combinations of values.

Examples

- Detect Outliers in Disk Failures
- Detect Outliers in Bitcoin Transactions
- Detect Outliers in Supermarket Purchases
- Detect Outliers in Mortgage Contracts
- Detect Outliers in Diabetes Patient Records
- Detect Outliers in Mobile Phone Activity

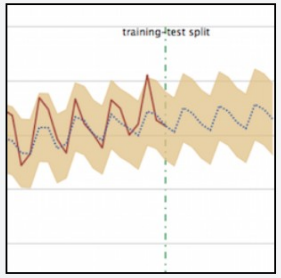


Detect Numeric Outliers

Find values that differ significantly from previous values.

Examples

- Detect Outliers in Server Response Time
- Detect Outliers in Number of Logins (vs. Predicted Value)
- Detect Outliers in Supermarket Purchases
- Detect Outliers in Power Plant Humidity
- Detect Cyclical Outliers in Call Center Data
- Detect Cyclical Outliers in Logins



Forecast Time Series

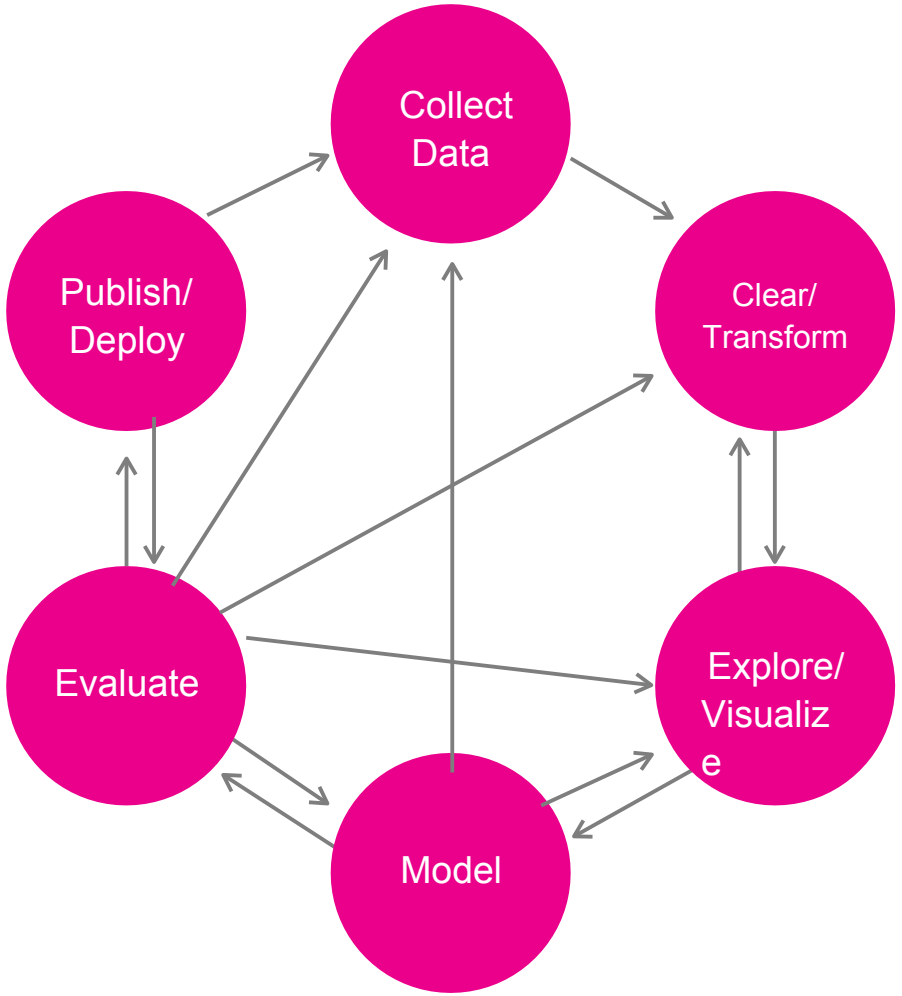
Forecast future values given past values of a metric (numeric time series).

Examples

- Forecast Internet Traffic
- Forecast the Number of Employee Logins
- Forecast Monthly Sales
- Forecast the Number of Bluetooth Devices
- Forecast Exchange Rate TWI using ARIMA

Machine Learning is not Magic

...it's a process

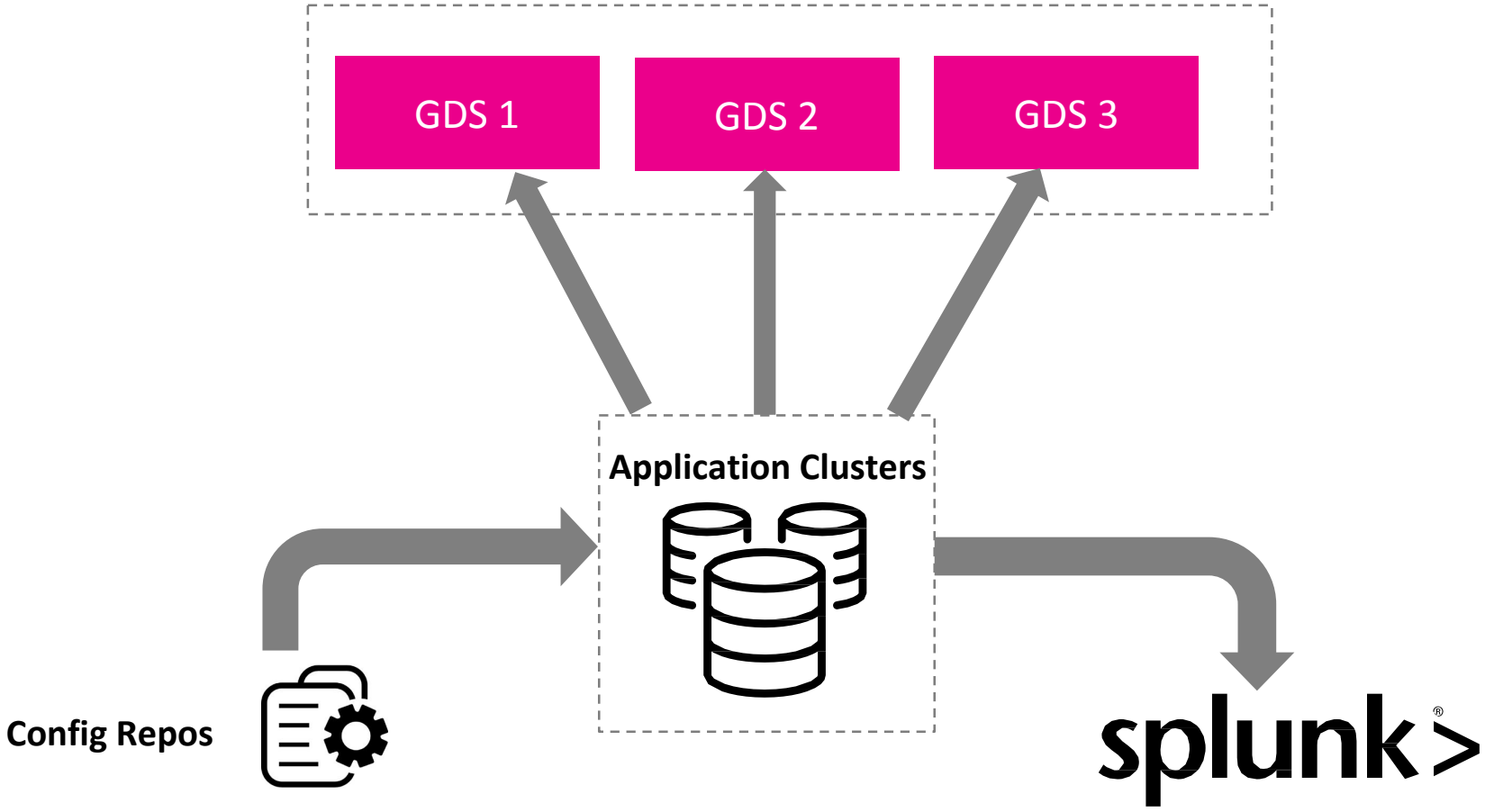


Numeric Outlier Detection on GDS Errors

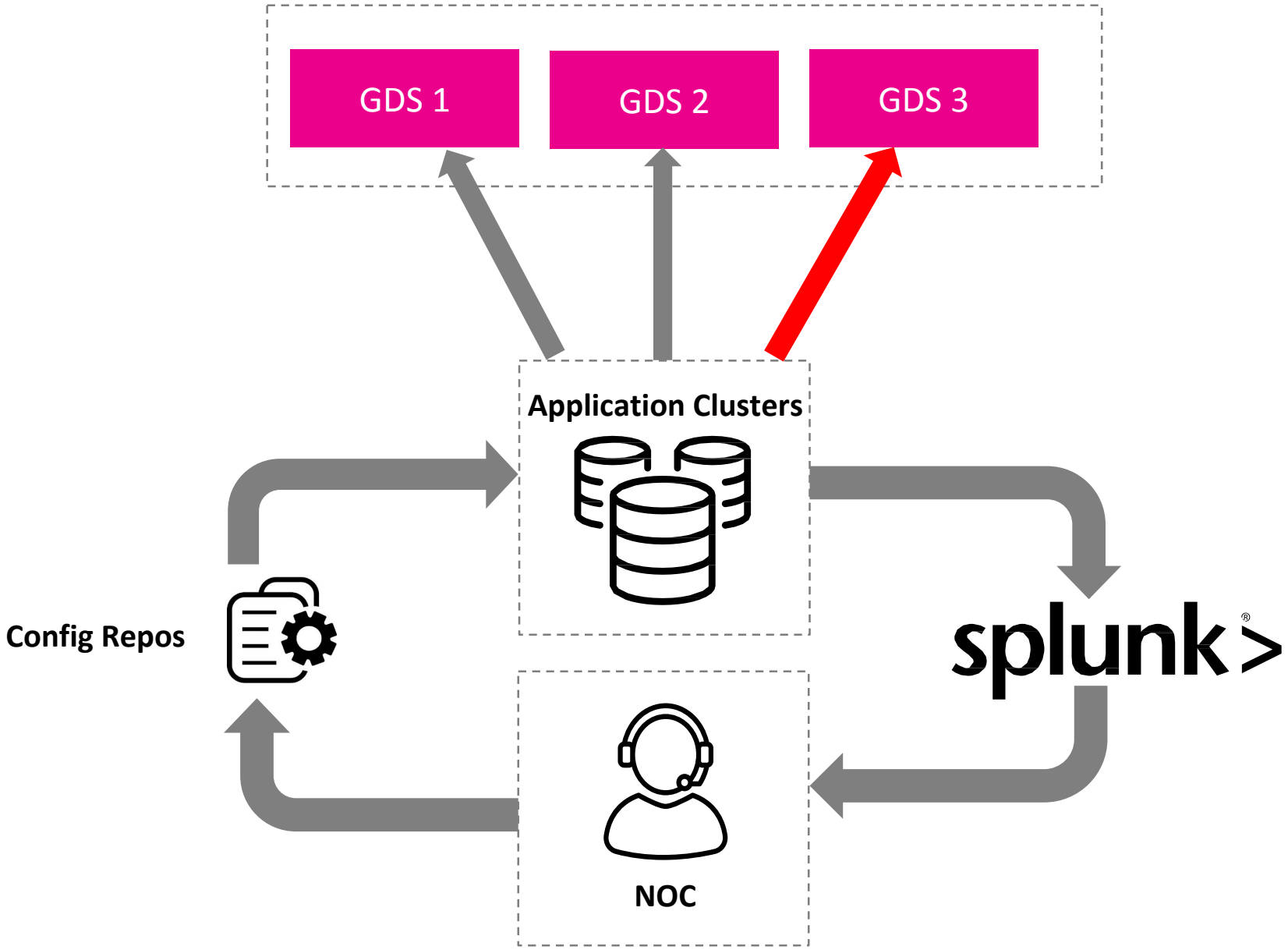
Self Healing with Priceline's Config Management Service

splunk> .conf19

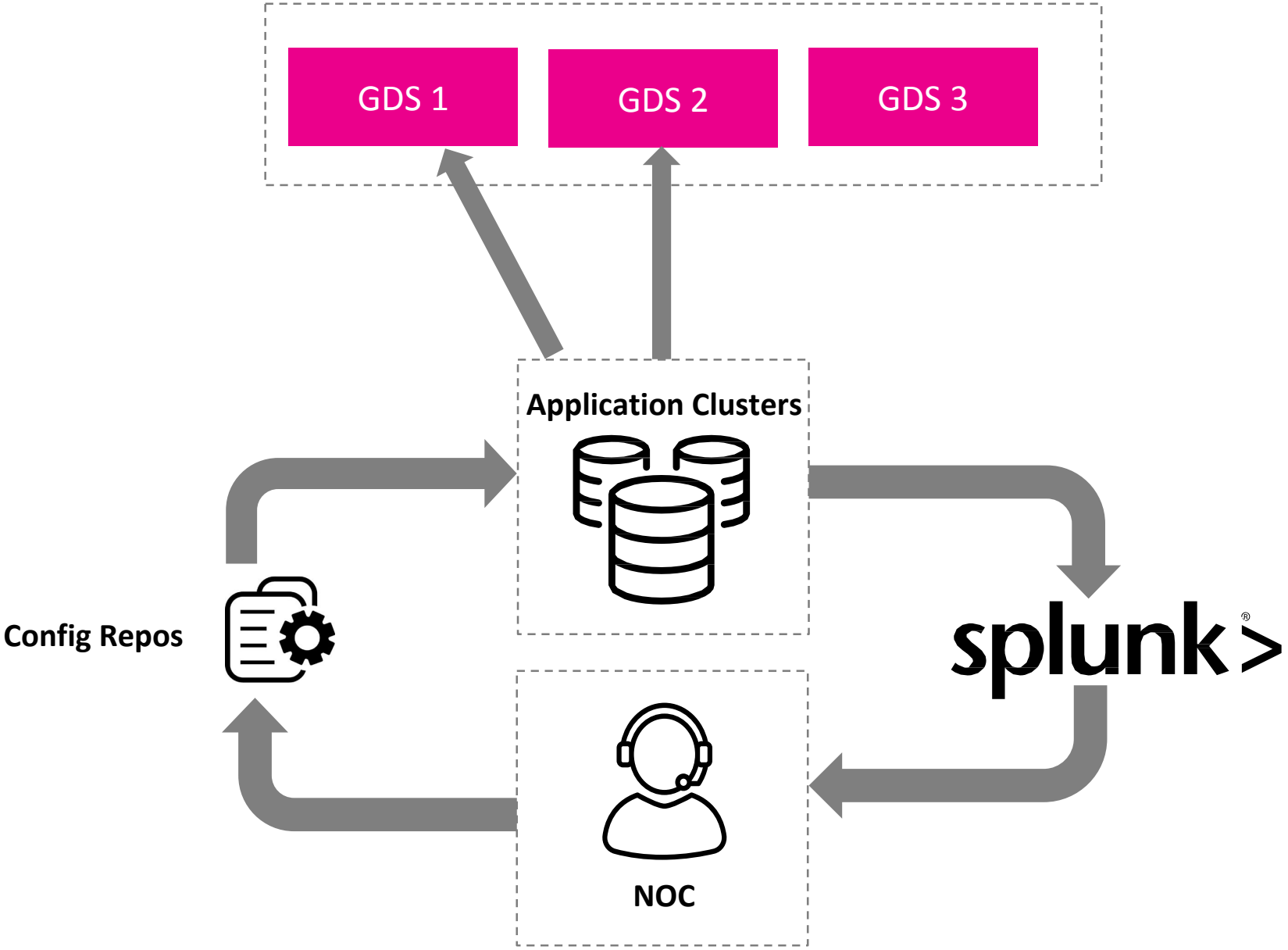
Befor
e



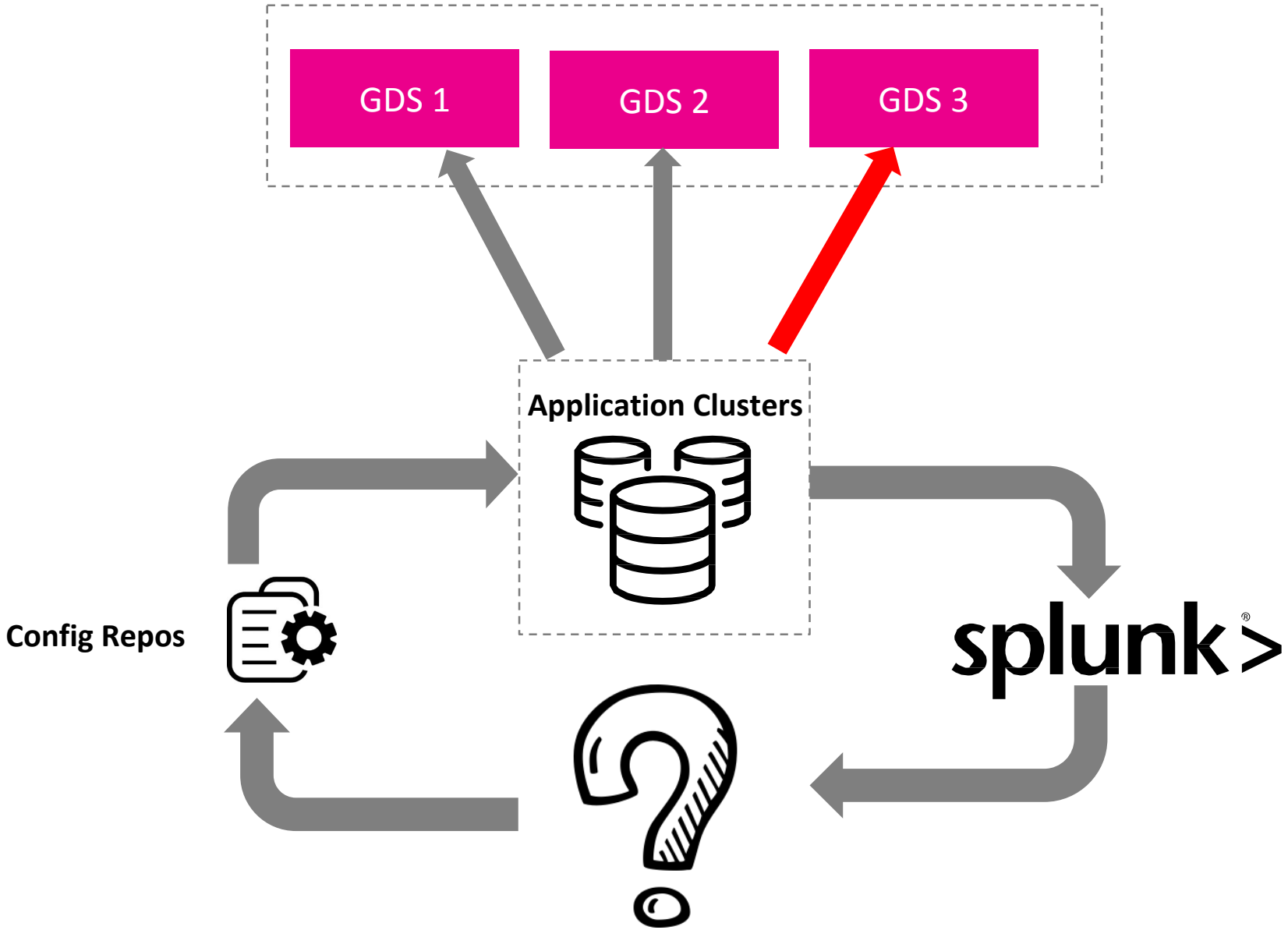
Befor
e


GDS 1
GDS 2
GDS 3
Application Clusters
Config Repos
splunk>
NOC

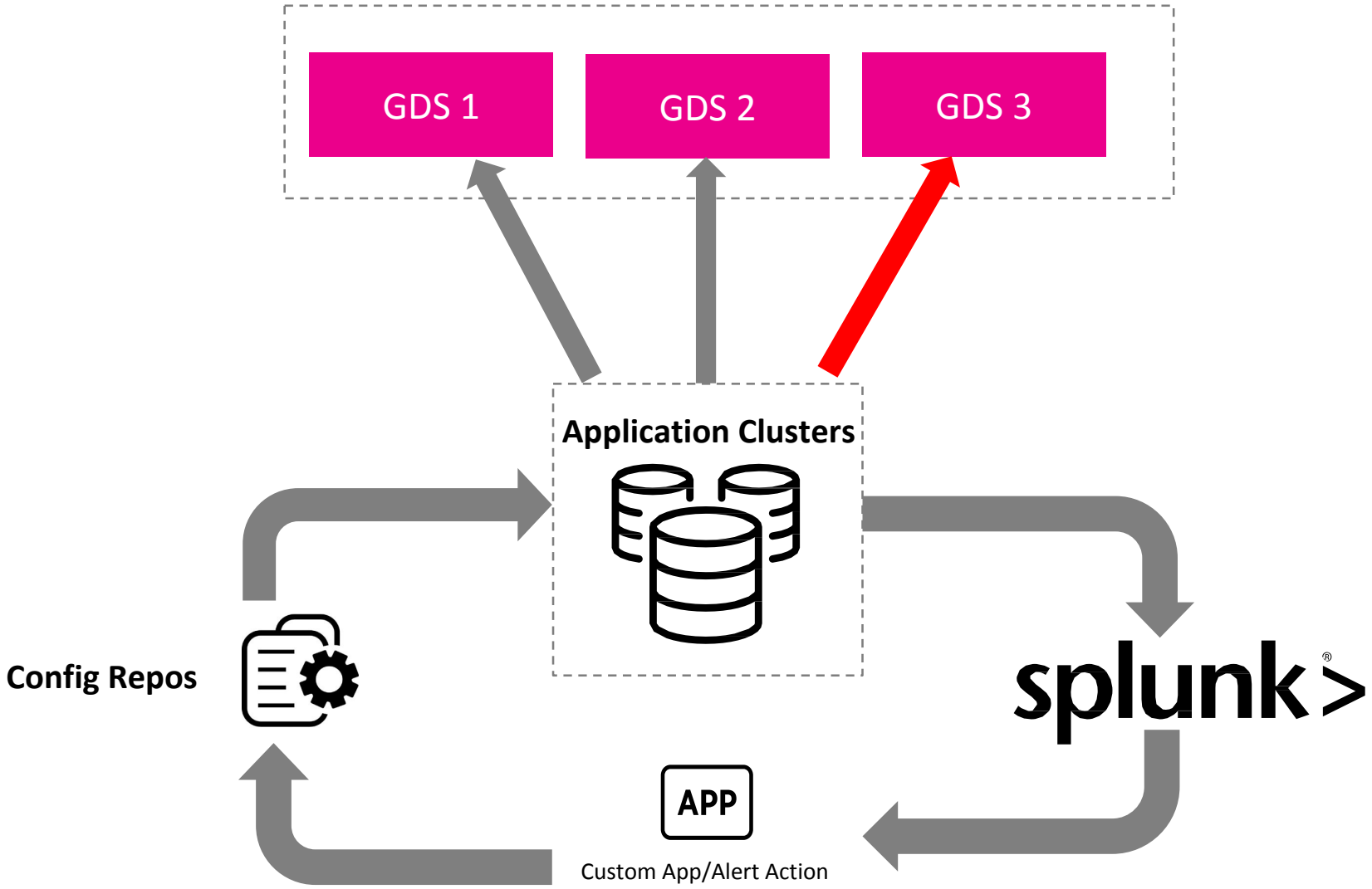
Befor
e


GDS 1
GDS 2
GDS 3
Application Clusters
Config Repos
splunk>
NOC

After



After


GDS 1
GDS 2
GDS 3
Application Clusters
Config Repos
splunk>
APP
Custom App/Alert Action

Custom Apps/Alert Actions

splunk> .conf19

Config Management Service

Acts as a bridge to complete the loop for self-healing

1. Trigger Webhook Plus Alert Action to invoke config service
2. Update Config Repo(s)
3. Trigger Remote Events to Applications
4. Audit config changes to Splunk

Webhook Plus – Custom App/ Alert Action Developed

Original Webhook is capable of making just Rest call.

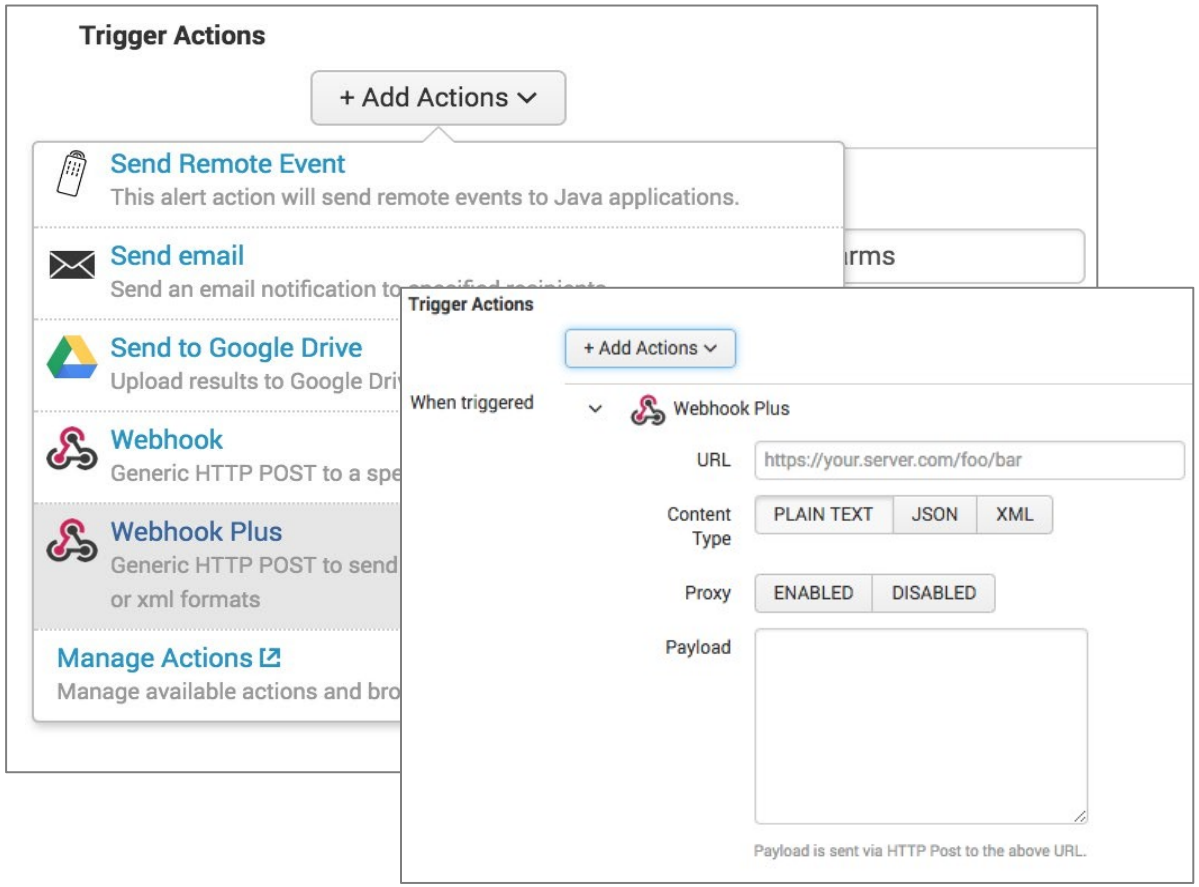
This is an extension to Webhook alert action that can post payloads in different formats to any end-point:

Use Case – 1:

- Turn off GDS by sending respective payload to config management micro-service

Use Case – 2:

- Turn off a consumer group when too many errors are noticed while processing messages into Splunk



Self Healing Loop – Example in Hotels GDS Alert

HOTEL_GDS_ERROR_ALERT

GDS3 has been disabled using config micro service

```
Weighted Threshold= 60
Previous Threshold = 62
ErrorPercent= 65
Alert Type=Outlier
```

Trigger History | Search Results

CONFIG
EnableGDSControl

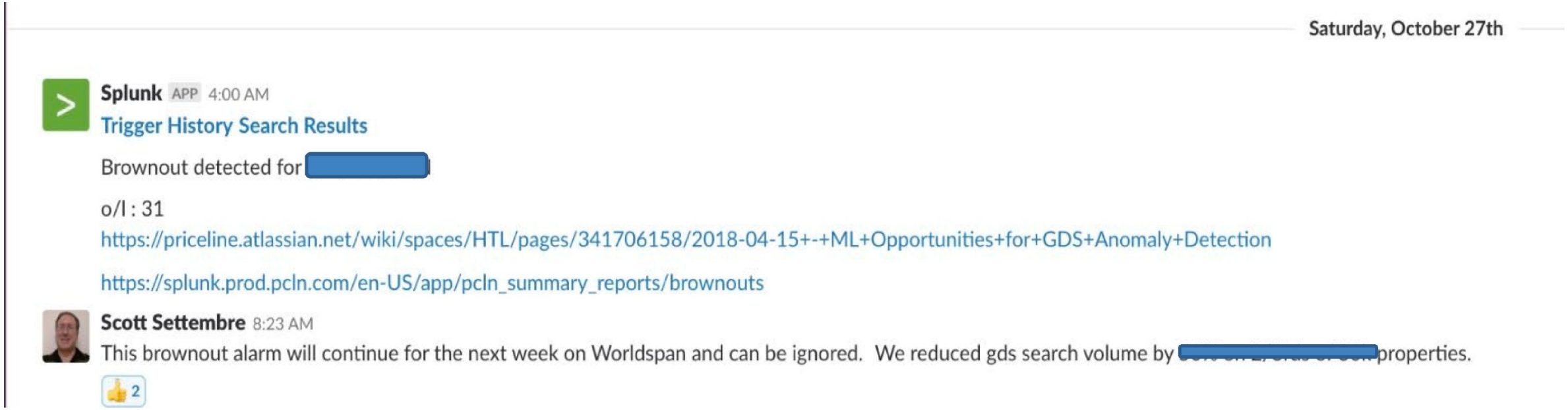
Name	Value
GDS1	true
GDS2	true
GDS3	false

Detecting KPI Performance ~~Degradation~~

Brownout

- Brownout is a Gradual degradation in KPI over time.
- Are difficult to detect unless they are sudden and last for an extended period of time.
- Summarize data.
- Forecast for next day.
- Count number of times it lower bound is breached.

ML for Brownout Use Case...



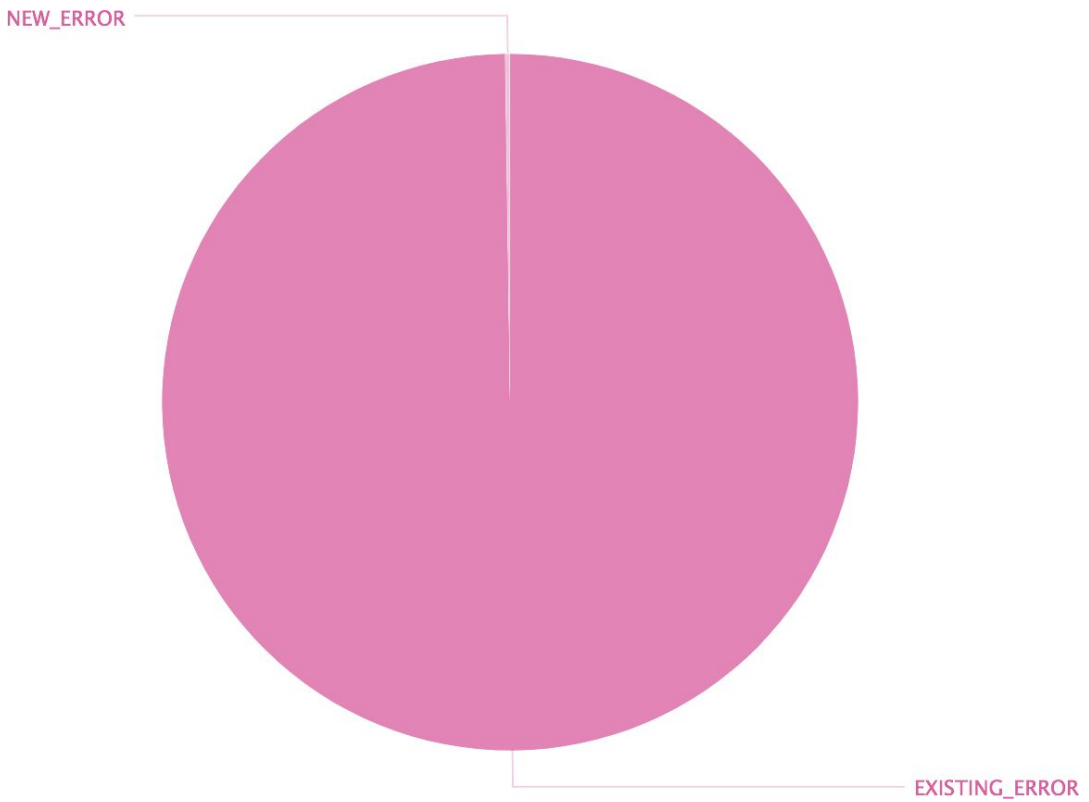
Detecting New Errors – using clustering

splunk> .conf19

Needle in a haystack

Cluster command

- groups events together based on how similar they are to each other.
- Identifying new errors that crop up over time, for instance – any post deploy
- errors or errors that are triggered due to config issues



Error Monitoring – Finding New Errors - Needle In A Haystack

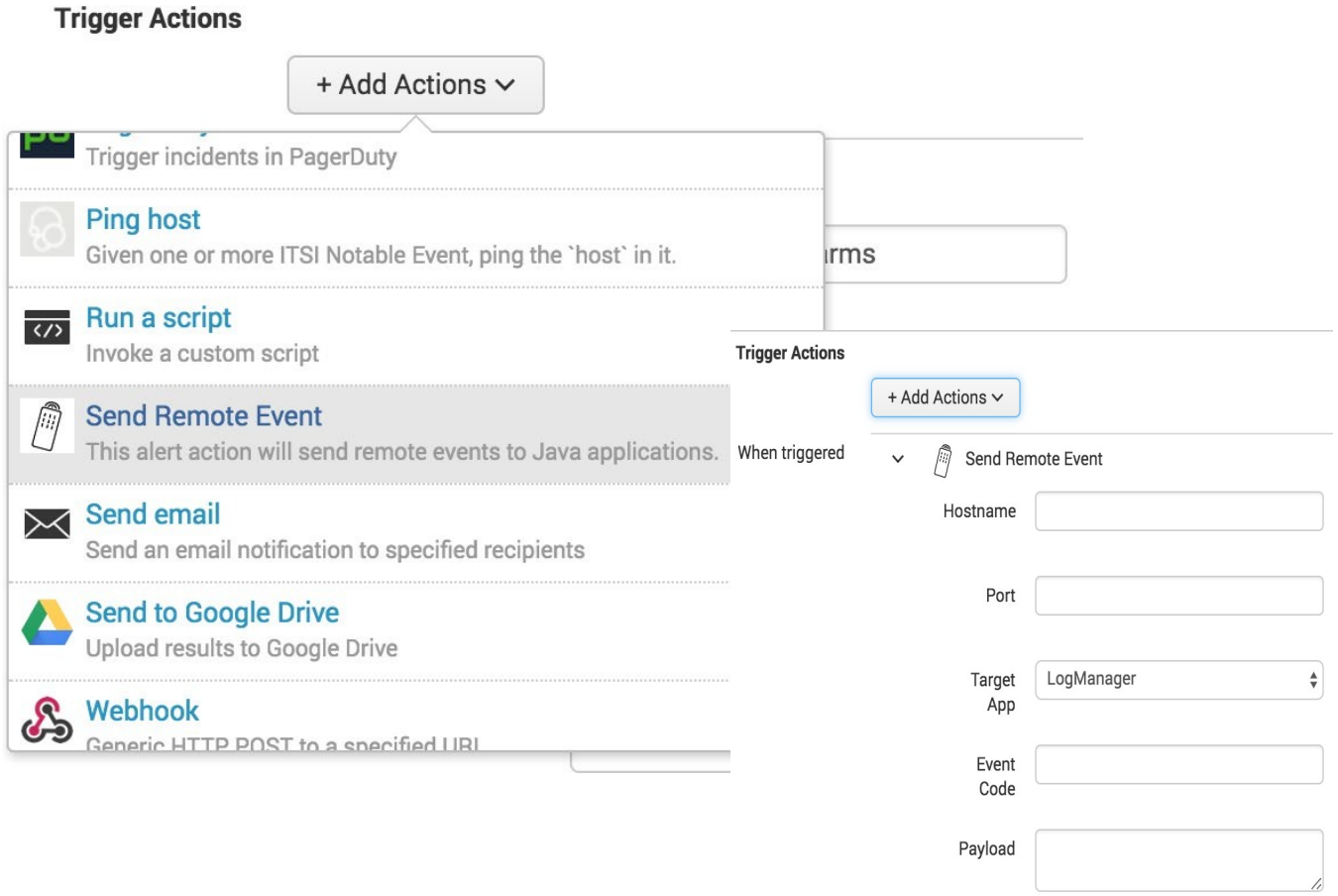
1. Finding anomalies through the concept of clustering
2. Leverage Splunk KVStore to keep track of incoming errors and tag them NEW vs OLD (again through clustering) and only alerts when new errors are triggered.

Send Remote Event – Custom App/ Alert Action Developed

This custom alert action will simulate “Send Remote Events” functionality of LDAP browser, when splunk alerts are triggered.

Use Case – 1:
Change logging levels

Use Case – 2:
Trigger refresh/reload events when certain criteria are met



Key Takeaways

1. Super intuitive to get started with Splunk MLTK in order to build proactive trustworthy alerts without static thresholds.
2. Power of building custom Apps in Splunk to help take corrective actions
3. Self healing use cases are heavily dependent on iteratively tuning and re-training of your alerts
4. All KPI alerts involving multi cyclic /seasonal time Series Metrics can mostly be tackled using Numeric outlier detection

Useful Links

From Zero To 100 In 100 Days – Story on Priceline's Splunk Adoption | .conf2017

- https://conf.splunk.com/files/2017/slides/from-zero-to-100-in-100-days-or-how-quickly-can-you-drive-splunk-_adoption.pdf

HEC Yeah!! How Priceline Uses HEC to Ingest 4TB of Data Every Day? | .conf2018

- https://static.rainfocus.com/splunk/splunkconf18/sess/1523585518806001Vs1b/finalPDF/HEC-Yeah-How-_Priceline-1868_1538788919830001YKjX.pdf

Splunk resources

- https://www.splunk.com/blog/2018/01/19/cyclical-statistical-forecasts-and-anomalies-part-1.html
- https://www.splunk.com/blog/2018/02/05/cyclical-statistical-forecasts-and-anomalies-part-2.html
- https://splunkbase.splunk.com/app/2890

Detect Numerical Outliers | .conf2017

https://conf.splunk.com/files/2017/slides/detect-numeric-outliers-advances.pdf

.conf19

splunk>

Thank You!

Go to the .conf19 mobile app to

RATE THIS SESSION