LOG AGGREGATION

To better manage your Red Hat footprint

Miguel Pérez Colino
Strategic Design Team - ISBU
2017-05-03

@mmmmmmpc

Agenda

Managing your Red Hat footprint with Log Aggregation

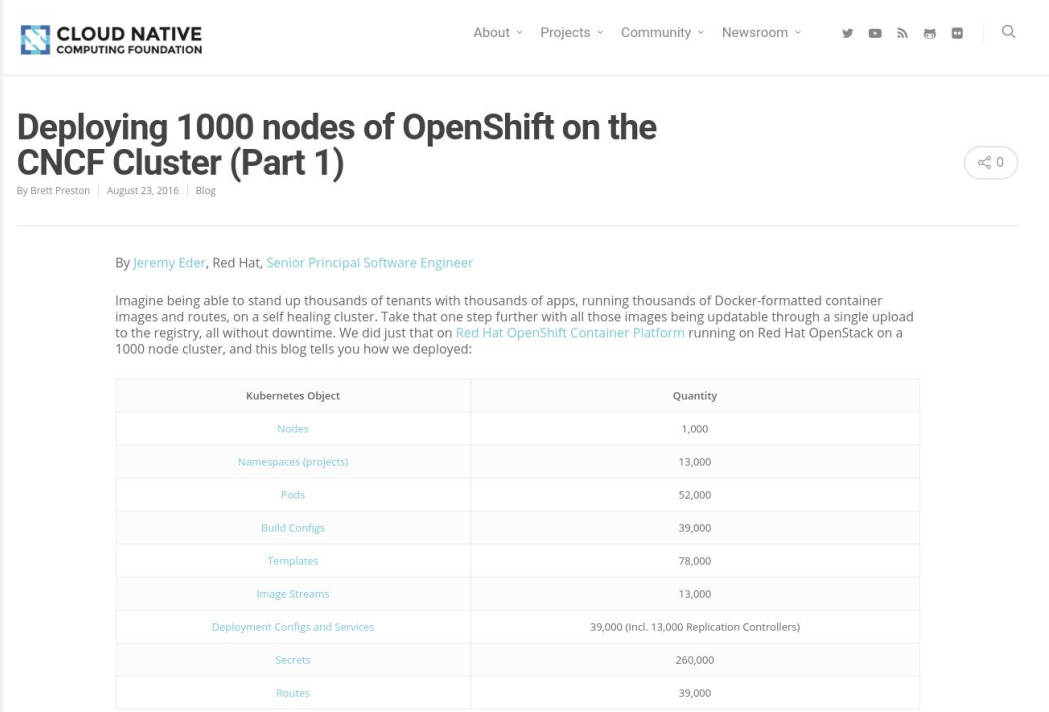
- The Situation
- The Challenge
- The Solution

THE SITUATION

Cloud Deployments

They do really scale …

- Higher scalability
- More workloads per physical machine (multi-tenant)
- Network and Storage also Software Defined
- Containers and Microservices providing more granularity



CLOUD NATIVE COMPUTING FOUNDATION

About Projects Community Newsroom

Deploying 1000 nodes of OpenShift on the CNCF Cluster (Part 1)

By Brett Preston | August 23, 2016 | Blog

By Jeremy Eder, Red Hat, Senior Principal Software Engineer

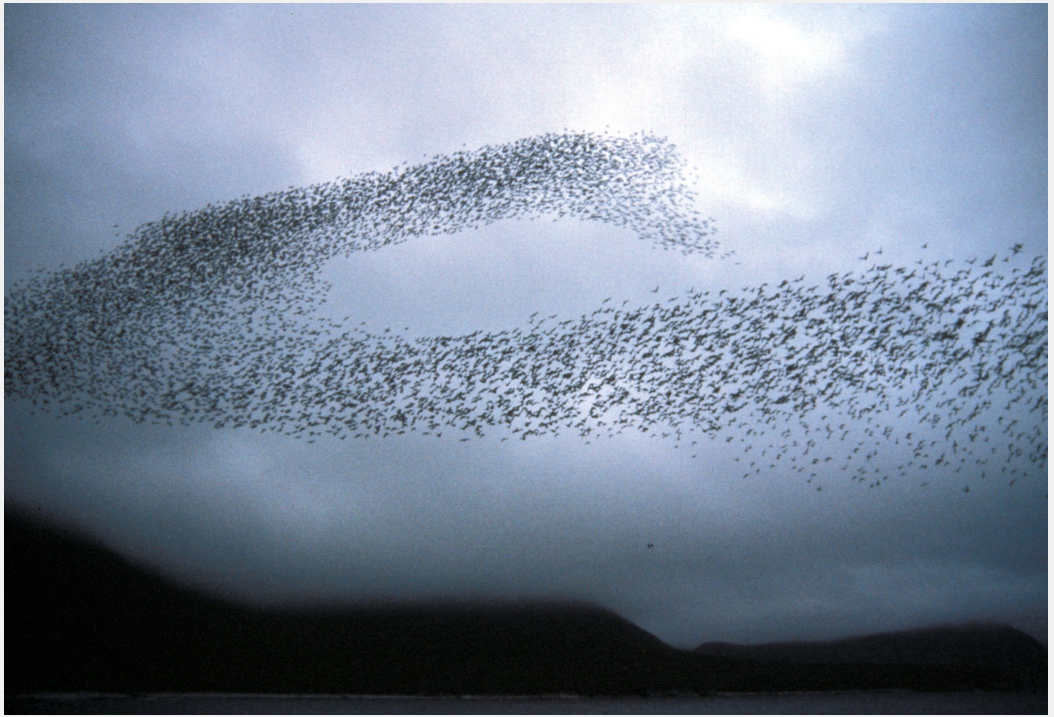
Imagine being able to stand up thousands of tenants with thousands of apps, running thousands of Docker-formatted container images and routes, on a self healing cluster. Take that one step further with all those images being updatable through a single upload to the registry, all without downtime. We did just that on Red Hat OpenShift Container Platform running on Red Hat OpenStack on a 1000 node cluster, and this blog tells you how we deployed:

Kubernetes Object	Quantity
Nodes	1,000
Namespaces (projects)	13,000
Pods	52,000
Build Configs	39,000
Templates	78,000
Image Streams	13,000
Deployment Configs and Services	39,000 (Incl. 13,000 Replication Controllers)
Secrets	260,000
Routes	39,000

https://www.cncf.io/blog/2016/08/23/deploying-1000-nodes-of-openshift-on-the-cncf-cluster-part-1/

Cloud Deployments

Act as one single thing ...



... and need to be managed and operated as one

Source: https://commons.wikimedia.org/wiki/File:Auklet_flock_Shumagins_1986.jpg

THE CHALLENGE

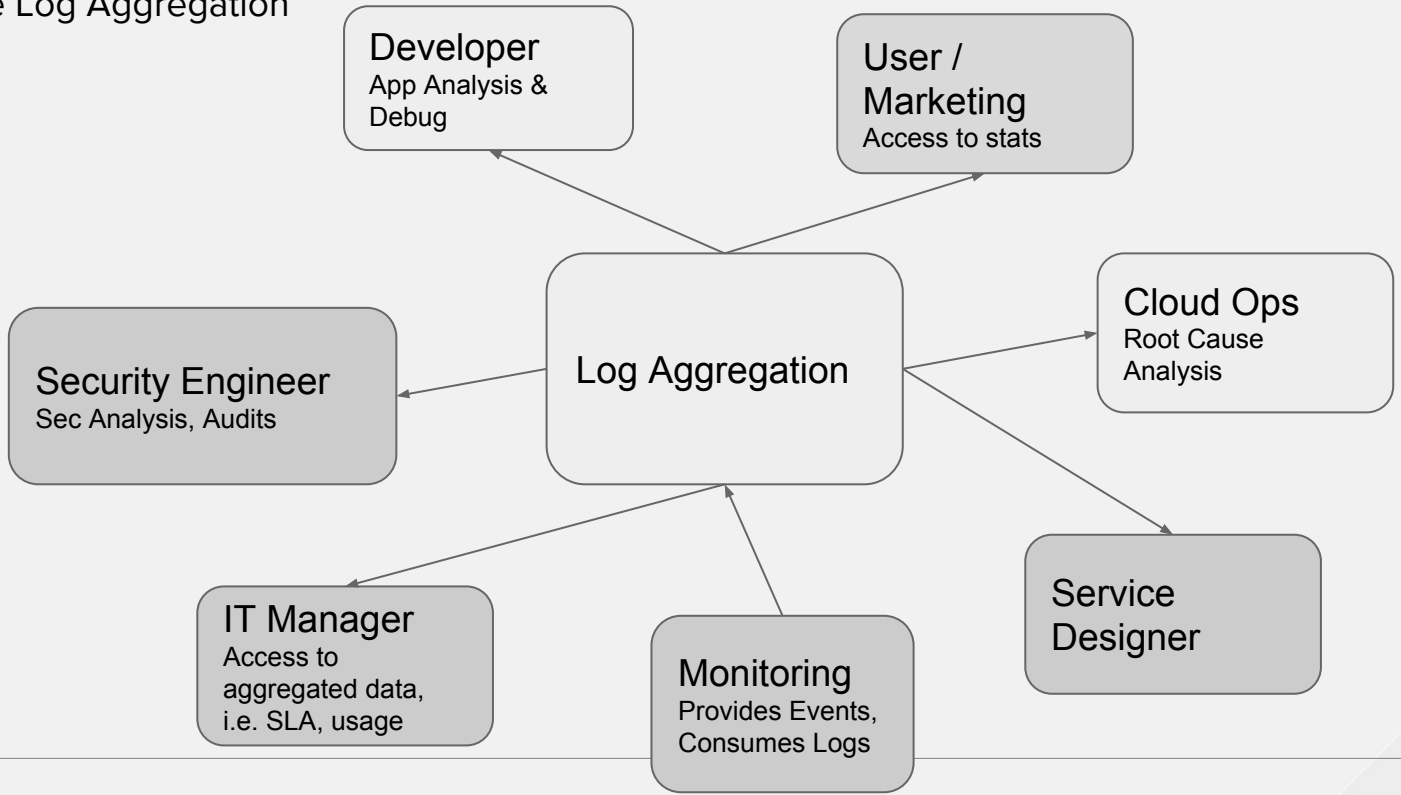
Data (What)

Data + Information flow in Log Aggregation

Derived from: http://www.dataintensive.info/

Personas (Who)

That can use Log Aggregation



Personas (Motivation)

That need Log Aggregation

"Application (multi-tiered) launched from CloudForms returns error"

Cloud Suite User

"I want to proactively know about active or potential degradation of service"

Cloud Ops (Apps)

"User reports that their VM request failed and returned error"

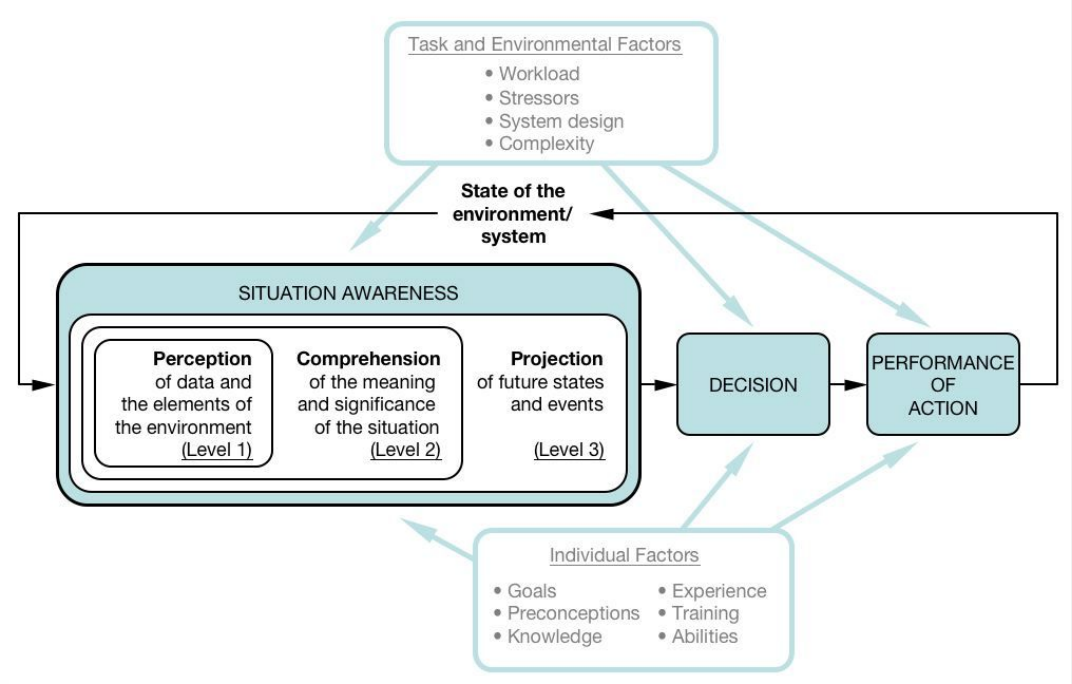
Cloud Ops (OpenStack)

"My recent commit resulted in Jenkins test failure"

Developer (OpenShift)

Situational Awareness (Why)

Or the need of it!

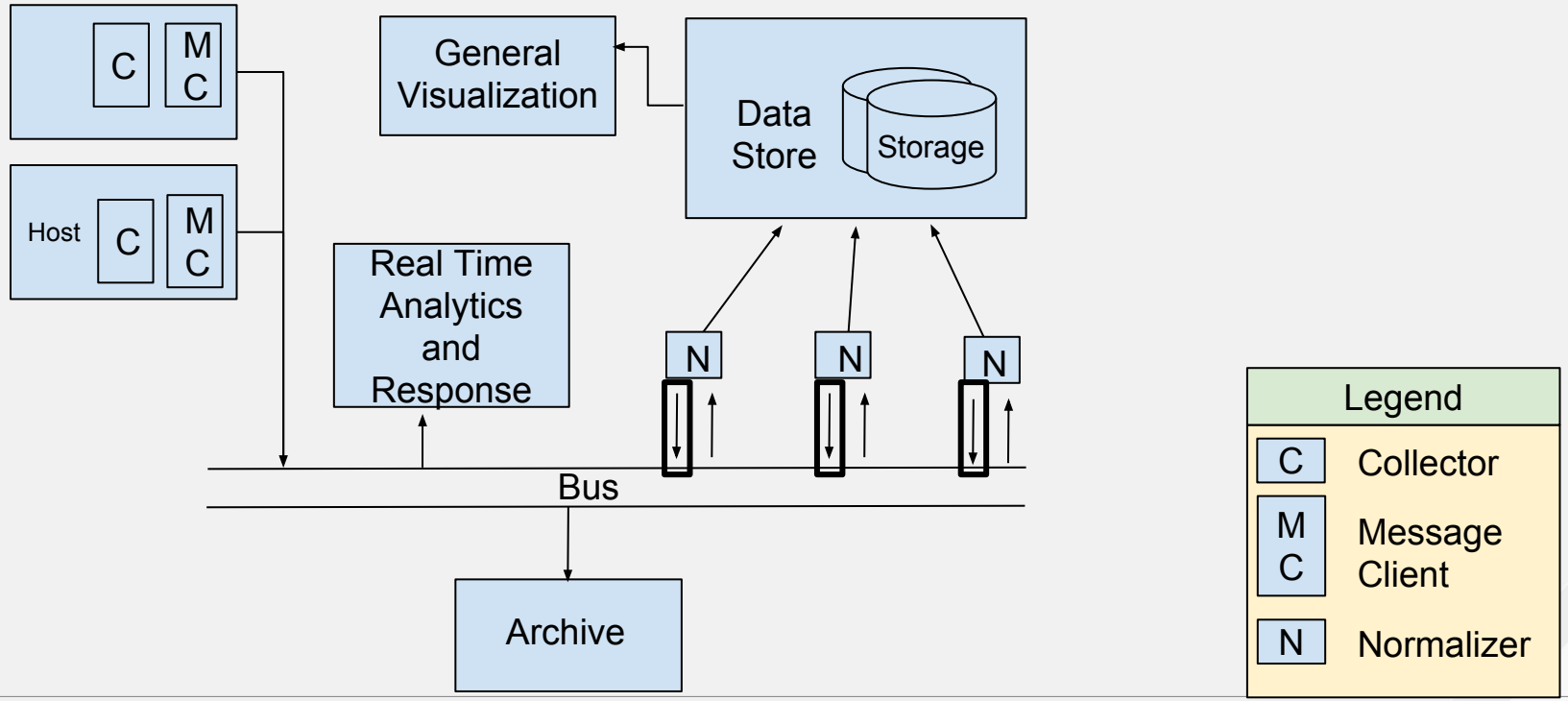


Source: https://en.wikipedia.org/wiki/Situation_awareness

THE SOLUTION

Architecture

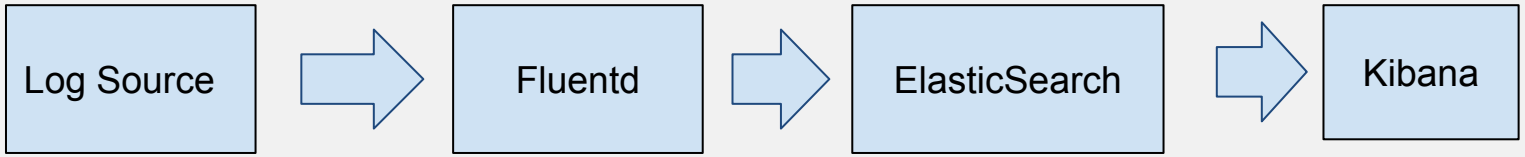
Proposed General Architecture



Slide Credit: Tushar Katarki [@tkatarki]

Implementation

Introduction to EFK



- TCP/UDP
- HTTP
- File: Text
- Stdout: CSV, JSON, MessagePack
- syslog/journal

- Parsing
- Filtering
- Enriching
- Deleting
- Output Buffering

Index and store data and metadata making search fast and reliable

User Interface for:
- Search
- Graph
- Dashboard

Slide Credit: Tushar Katarki [@tkatarki]

Current Status

Being delivered and supported

OpenShift Container Platform 3.5

- Full EFK stack provided as containers

OpenStack Platform 10

- Fluentd as log collector

Red Hat Virtualization

- Coming Soon!

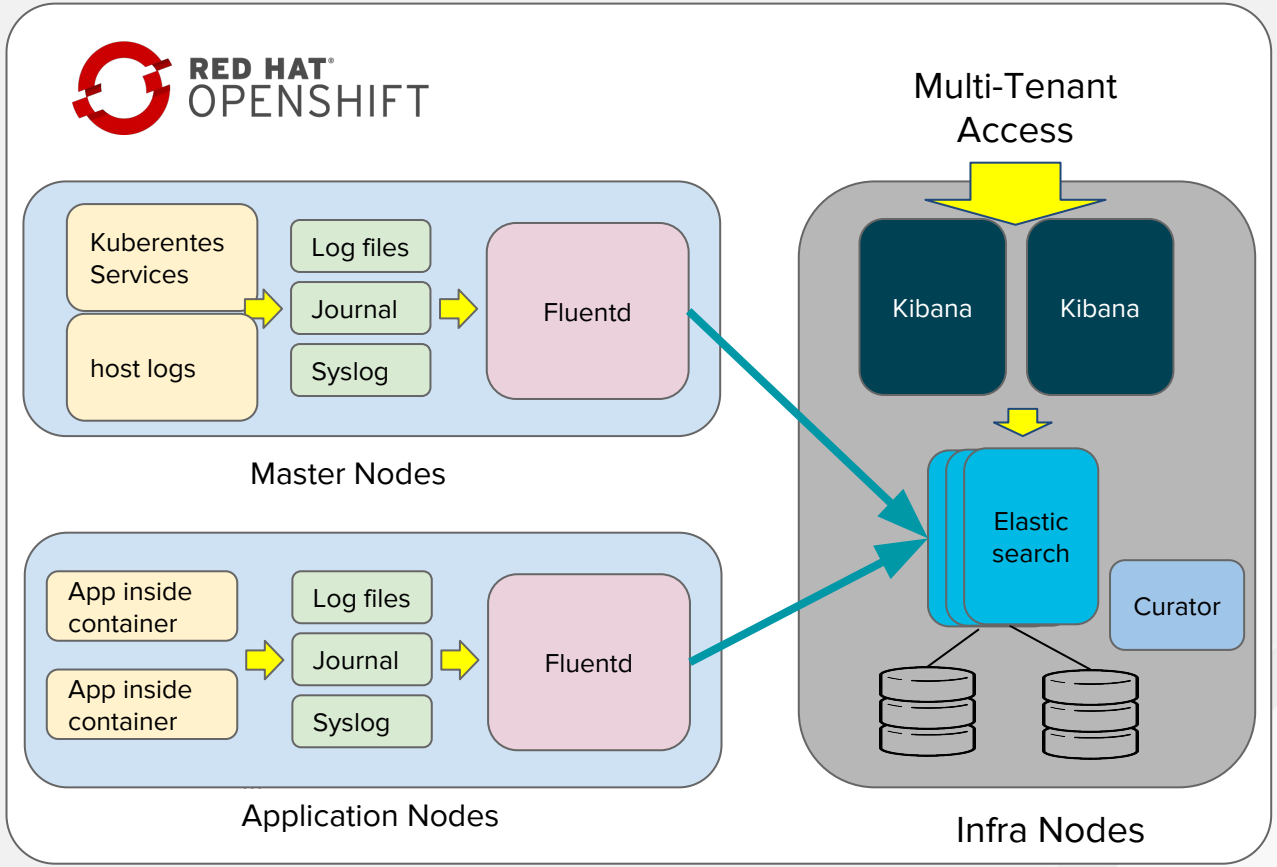


Diagram Credit: Tushar Katarki [@tkatarki]

BEYOND ...

Common Data Model

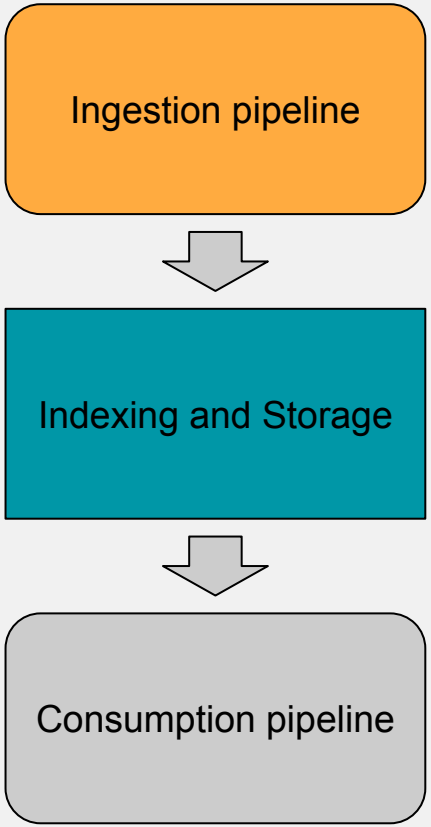
To ensure integration and interoperability

What Is It?

- A Data Model for Logs (and other data) to identify and tag data (i.e. log fields)

Why?

- Alignment/Correlation with different RH products
- Improved maintainability of Data
- Better presentation/data consumption
- Enables 3rd party ecosystem
- Facilitates deep learning analysis of data



Common Data Model

Example …

```
[root@asherkho-ose-sec containers]# tail -1 /var/log/containers/cakephp-example-1-nzx3e_test_cakephp-example-6dcac0cd68b8b56a569505457235c511340e7b9edf7c911ce3ca34af4ea17973.log
{"log":"10.1.0.1 - - [03/Jun/2016:13:53:58 -0400] \"GET / HTTP/1.1\" 200 64124 \"-\" \"Go 1.1 package http\"\n","stream":"stdout","time":"2016-06-03T17:53:59.054842936Z"}
```

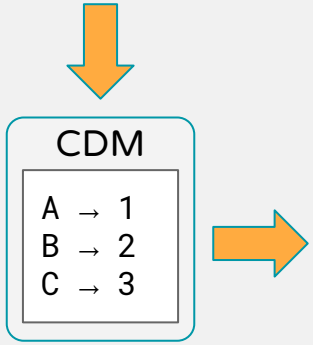
Data extracted:

- Container name
- Pod name
- Namespace name
- Docker container ID

K8S data queried:

- Pod UID
- Pod labels
- Pod host
- Namespace UID.

All merged into output log in JSON Format

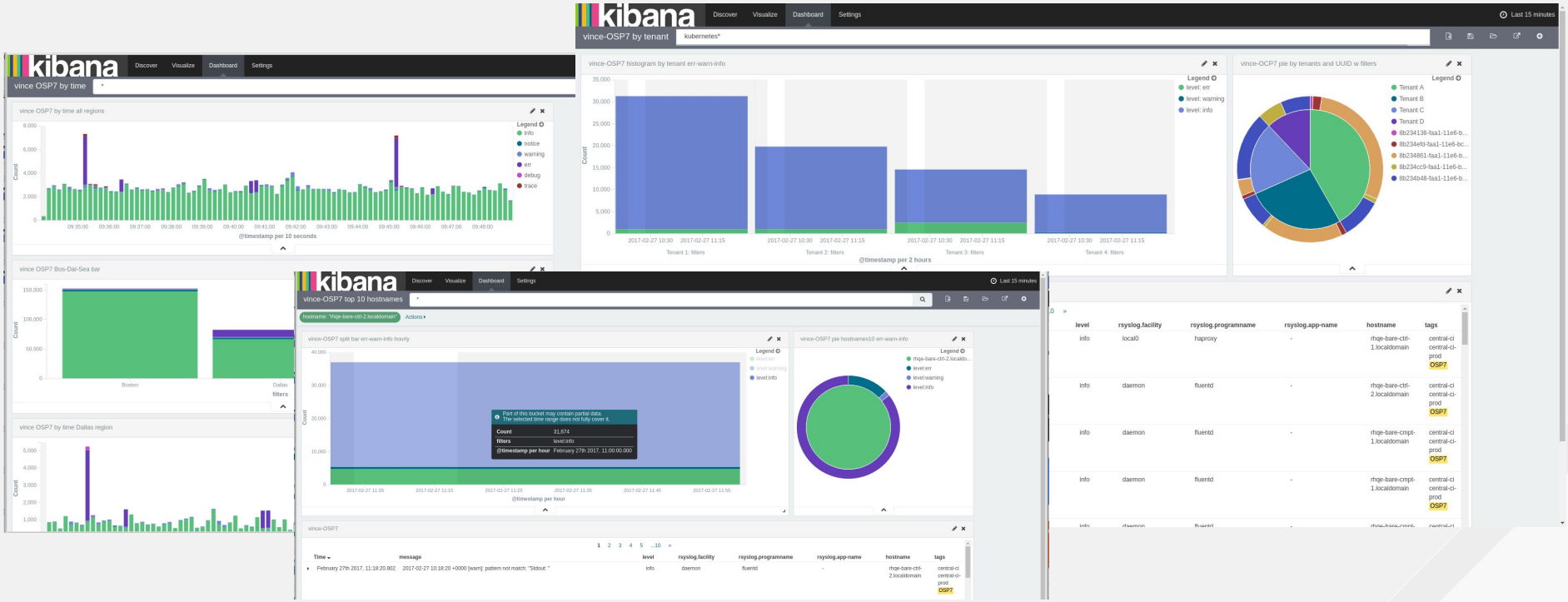


t hostname	asherkho-ose-sec.os1.phx2.redhat.com
t k8s_nodename	asherkho-ose-sec.os1.phx2.redhat.com
t k8s_object_meta.labels	{"deployment"=>"cakephp-example-1", "deploymentconfig"=>"cakephp-example", "name"=>"cakephp-example"}
t k8s_object_meta.name	cakephp-example-1-nzx3e
t k8s_object_meta.namespace	test
t k8s_object_meta.namespace_id	176f3960-2380-11e6-a91f-fa163ebe1970
t k8s_object_meta.uid	3d3269a8-275e-11e6-a91f-fa163ebe1970
t kind	Pod_log
t message	10.1.0.1 - - [03/Jun/2016:13:54:58 -0400] "GET / HTTP/1.1" 200 64124 "-" "Go 1.1 package http"

Images Credit: Anton Sherkhonov [@peatz]

User Experience

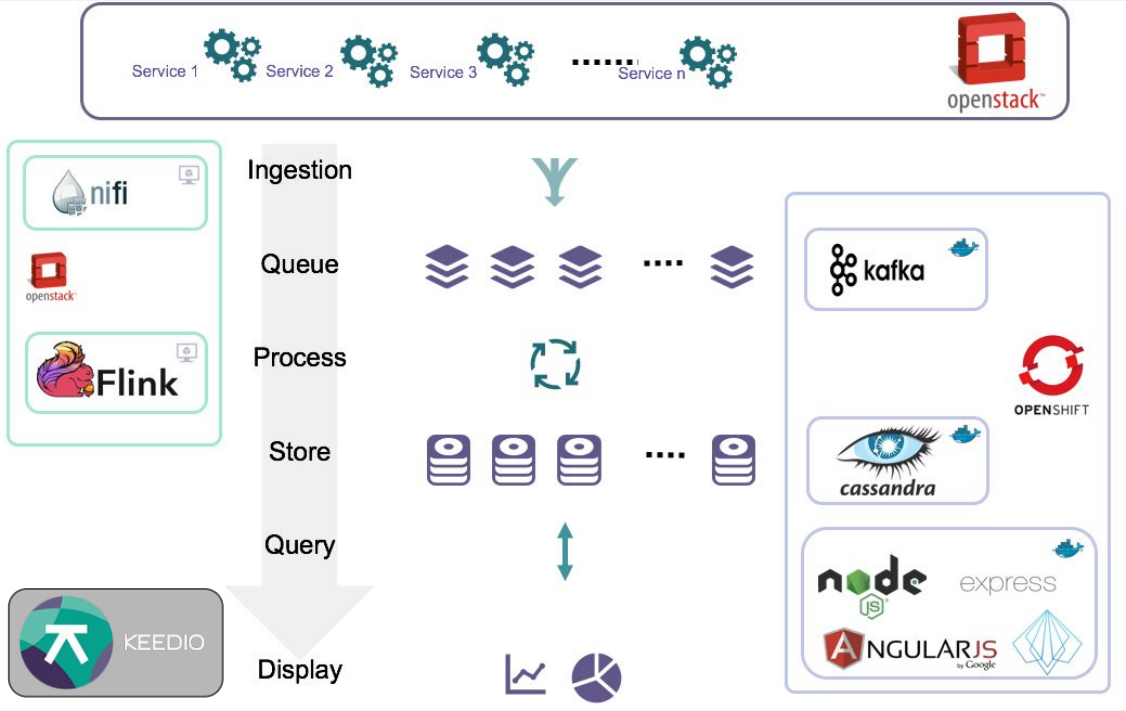
Prototyping and validating dashboards for users



Slide Credits: Peter Portante & Vince Conzola

Exploring different approaches

Prototyping with alternative toolsets with partners



Slide Credits: Luca Rosellini (Keedio)

ACTION!

How are you doing it?

Please, provide your feedback ...

http://bit.ly/log-aggregation

RED HAT SUMMIT

THANK YOU

plus.google.com/+RedHat

facebook.com/redhatinc

linkedin.com/company/red-hat

twitter.com/RedHatNews

youtube.com/user/RedHatVideos

LEARN. NETWORK.
EXPERIENCE
OPEN SOURCE.