Cloud Native Logging

OpenShift Commons Briefing, April 20th 2017

Eduardo Silva @edsiper
eduardo@treasure-data.com

TREASURE DATA

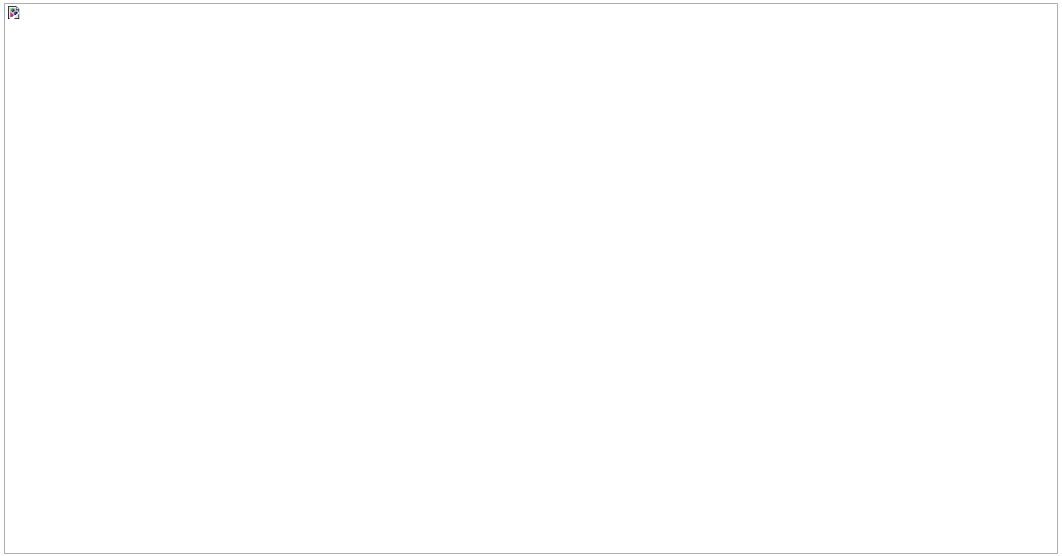
“Logging... Why?”

Logging: Why ?

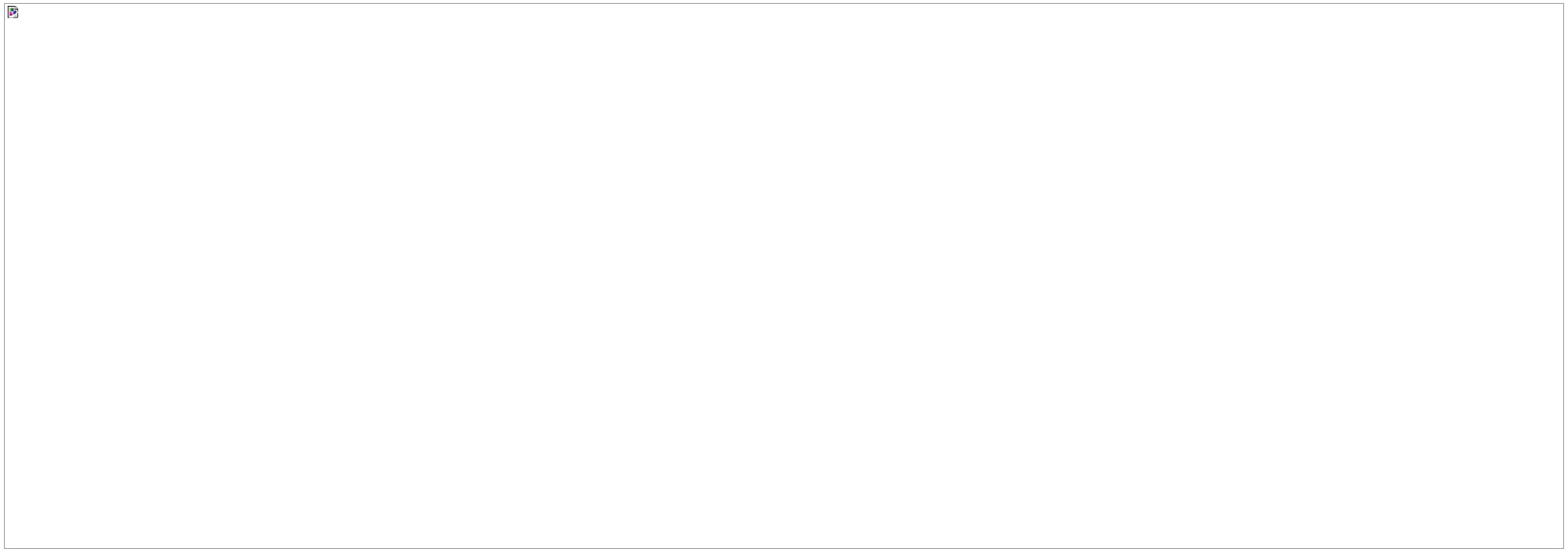
Analyze Applications Behavior

Logging: Why ?

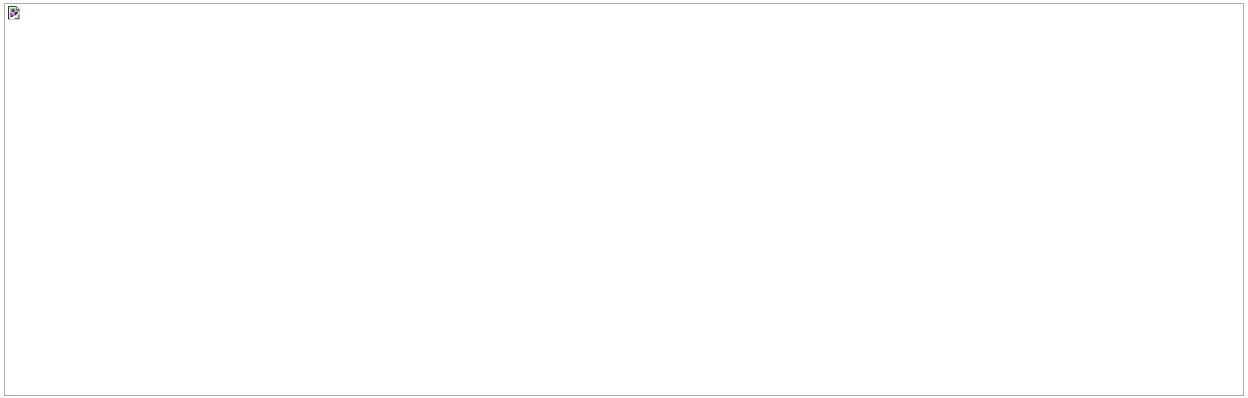
Analyze Applications Behavior



Logging: Why ?



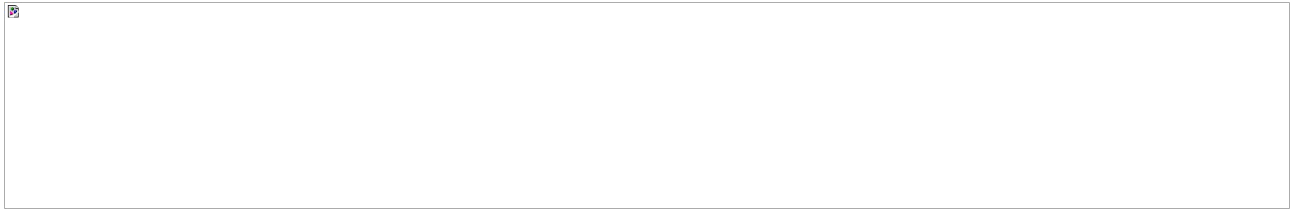
Logging: Why ?



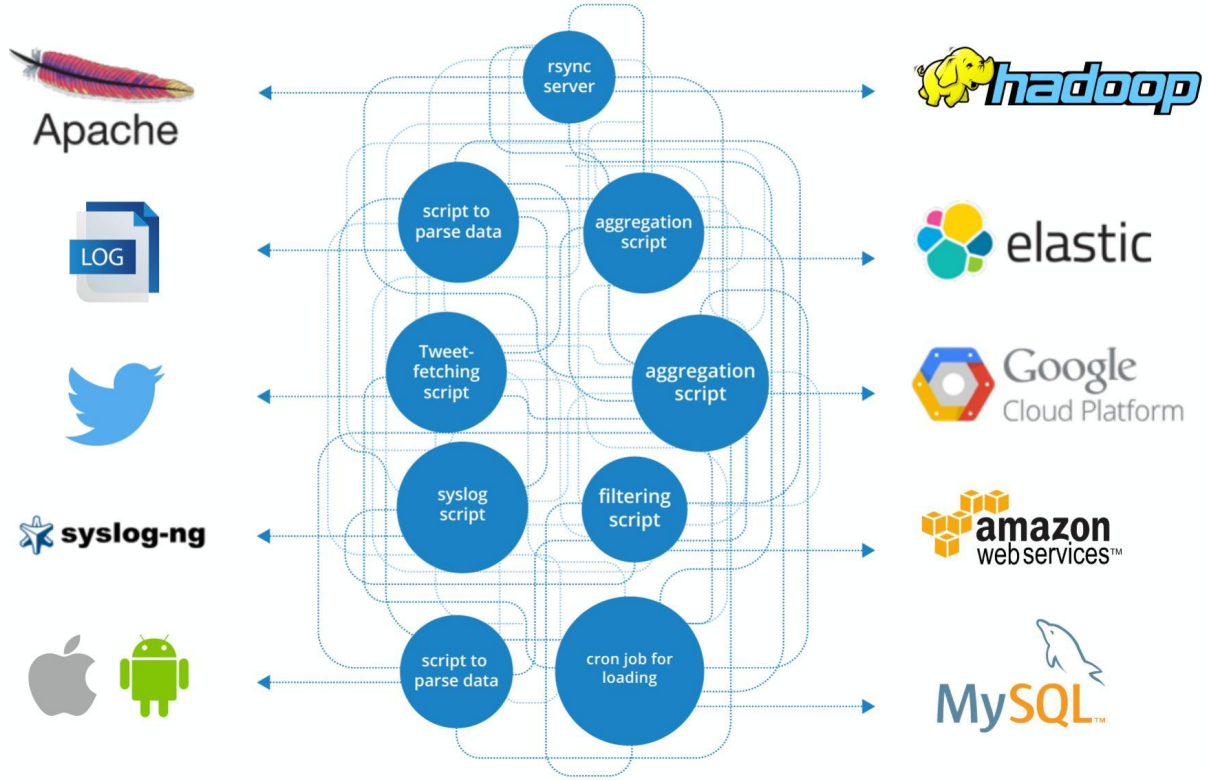
Logging Pipeline

Logging Pipeline

Logging Pipeline



It’s a mess


Apache
LOG
syslog-ng
rsync server
script to parse data
aggregation script
Tweet-fetching script
aggregation script
syslog script
filtering script
script to parse data
cron job for loading
hadoop
elastic
Google Cloud Platform
amazon web services
MySQL

“Logging... Why?” (Micro Services)

OK, Where's My Log in Container Era?

Everything at Google runs in containers

Launching over **2 billion** containers **per week**

Shipping Containers At Clyde, by Steve Gibson

Logging has Never Been Easier

TREASURE DATA


fluentd

About Fluentd

- Invented by

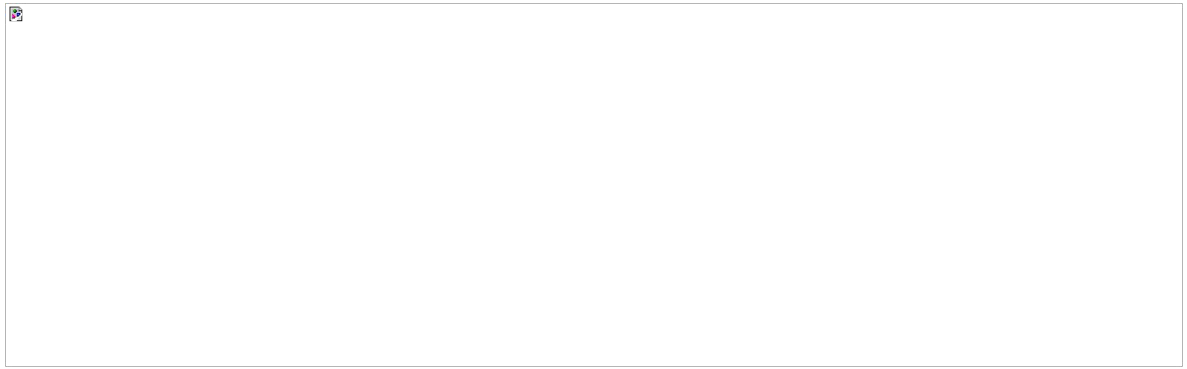
TREASURE DATA

- Now hosted at

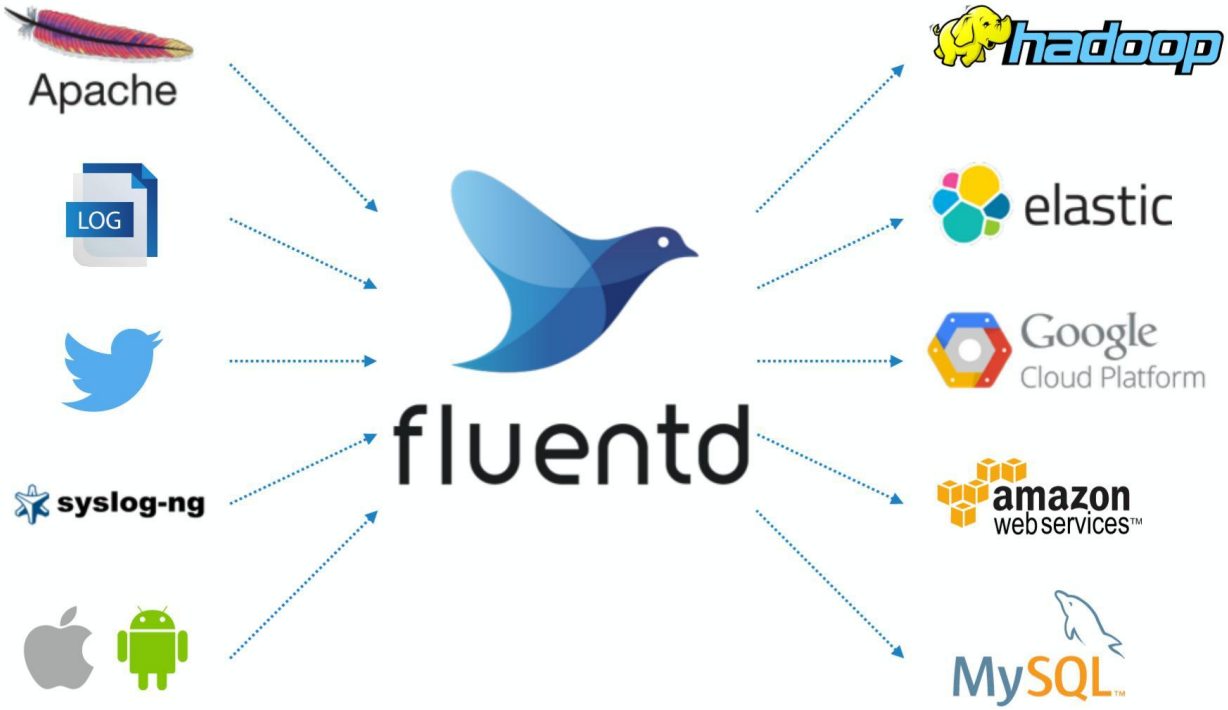
CLOUD NATIVE
COMPUTING FOUNDATION

Logging Pipeline

How Fluentd Simplifies Logging Pipeline



How Fluentd Simplifies Unified Logging Layer



About Fluentd

- More than **600** plugins available
- Pluggable Architecture
- Built-in Reliability
- Full integration with Docker and Kubernetes
- Written in Ruby + C

Fluentd Plugin Ecosystem

fluentd

Schema-less Backends

- Treasure Data
- MongoDB
- InfluxDB
- Hadoop DFS

Syslog Search

- ElasticSearch
- Splunk
- Loggly
- Groonga

Data Archiving

- Azure Blob Storage
- AWS S3
- Google Cloud Storage
- File

Schema-full Backends

- AWS Redshift
- Google Big Query
- MySQL
- PostgresSQL
- SQL Server
- HP Vertica

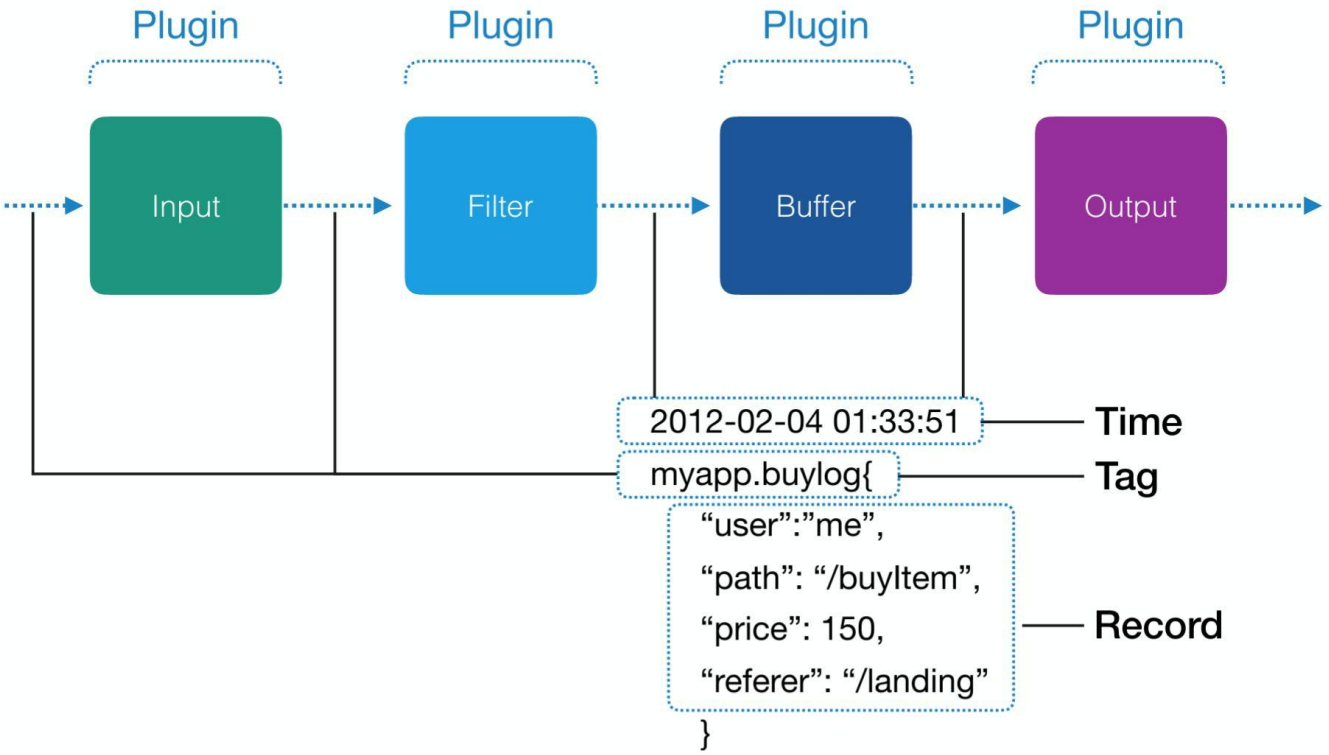
Middleware

- AWS Kinesis
- Kafka
- CouchDB
- Couchbase
- HBase
- RabbitMQ

Monitoring Systems

- Librato
- Ganglia
- Zabbix
- Mackerel
- Datadog
- Sentry

Fluentd Modular Architecture



Fluentd Structured Logging

- Text log

  "This is a test message"

- Structured log

```
{
  "host":"hotdog.infra.treasuredata.com"
  "pod_name":"mypod",
  "container_id":"bfdd5b9...",
  "container_name":"/infallible_mayer",
  "source":"stdout",
  "log": "This is a test message"
}
```

Fluentd at Scale

- Docker
  - Native Docker logging driver to use Fluentd

- Kubernetes
  - Fluentd as main aggregator

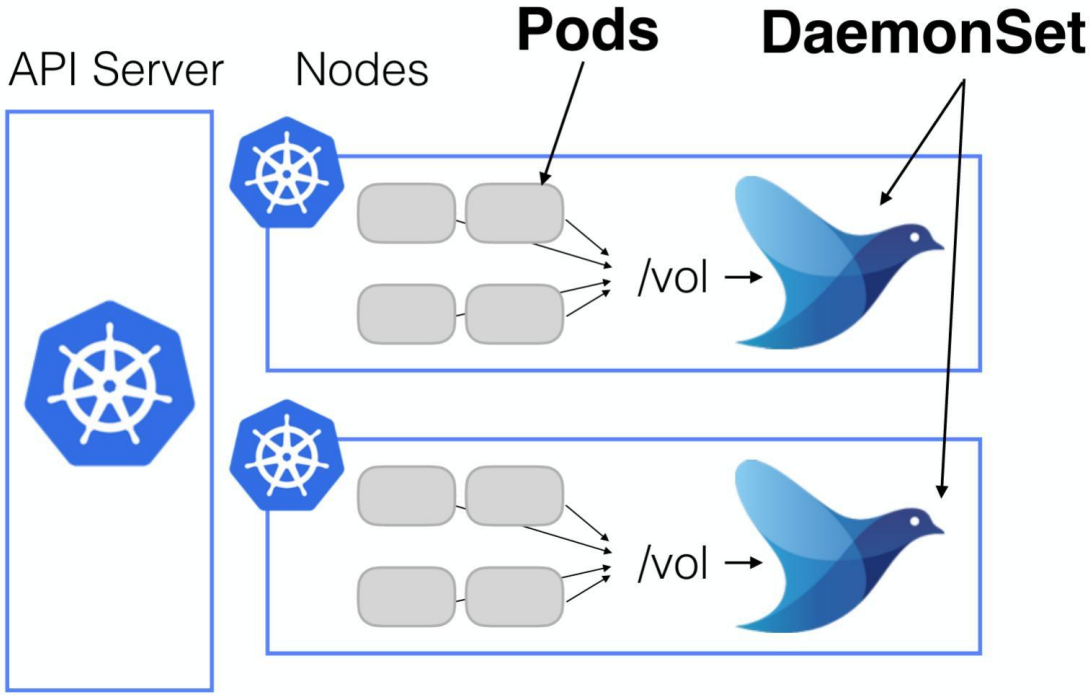
- Google Cloud Platform
  - Fluentd as the logging agent (Stackdriver)

- OpenShift
  - Fluentd as main aggregator

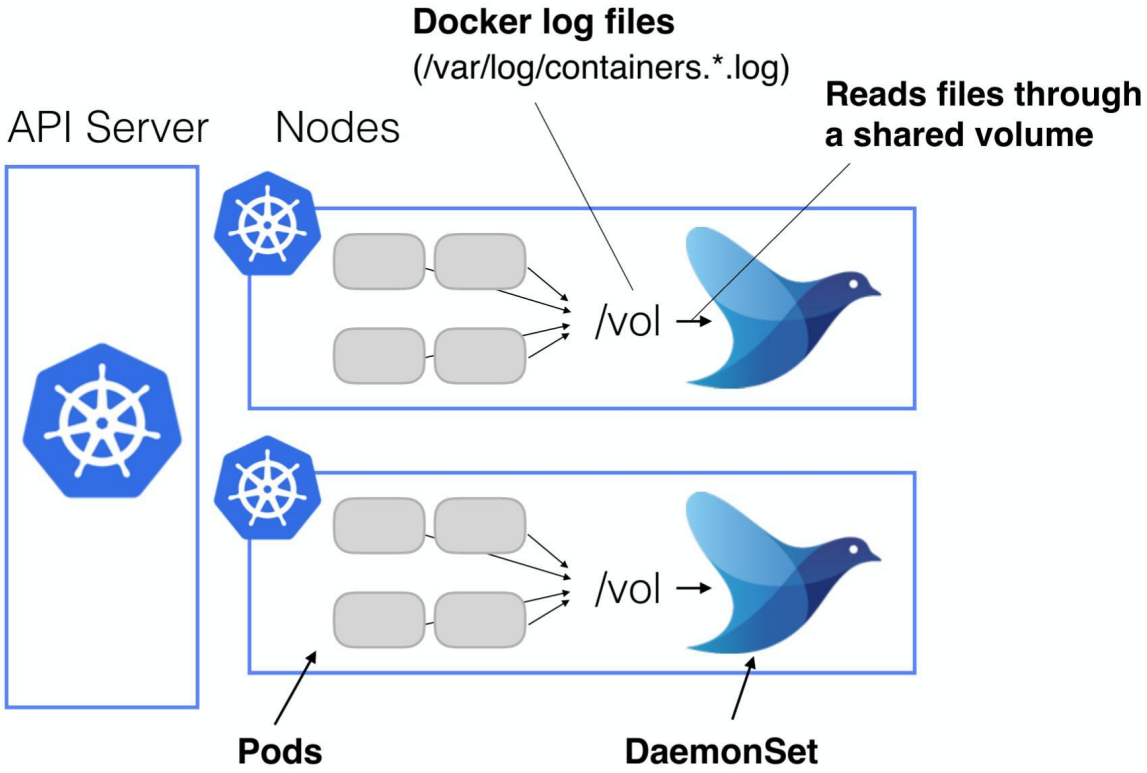
Kubernetes & Fluentd

TREASURE DATA

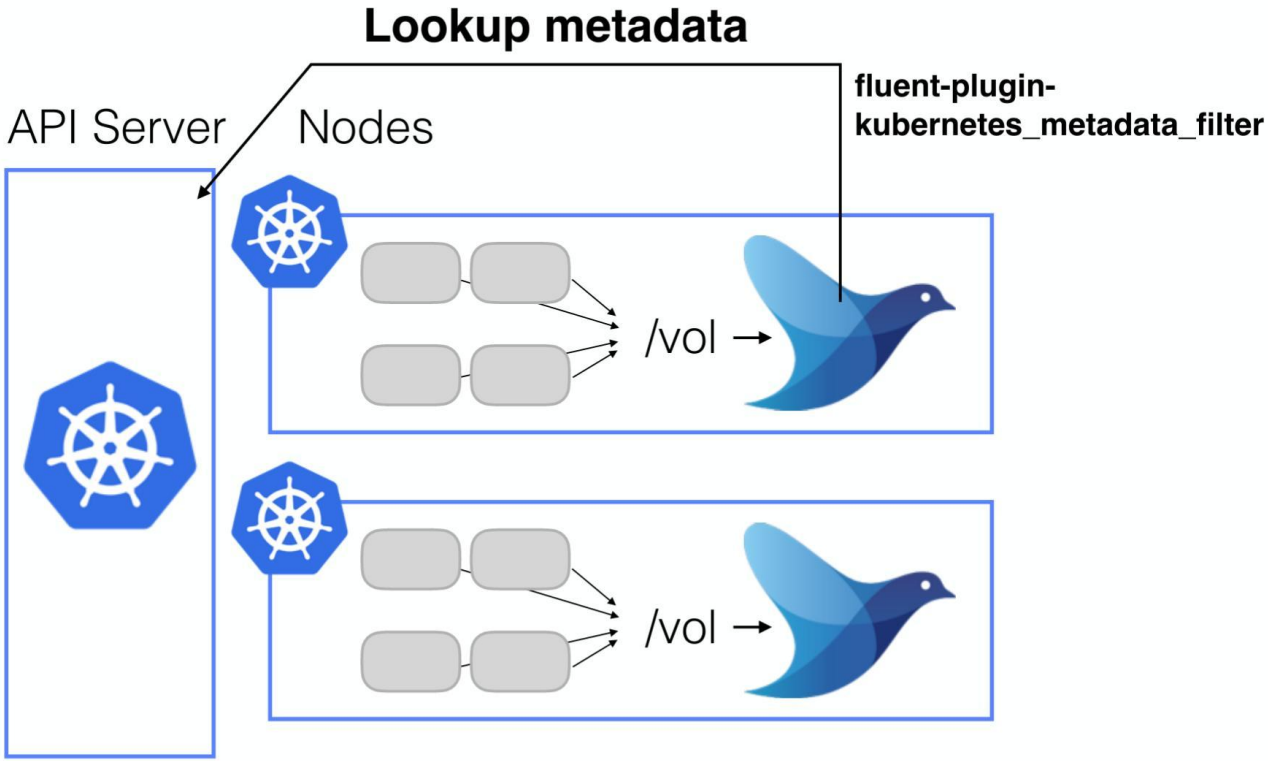
Fluentd Deployment: DaemonSet



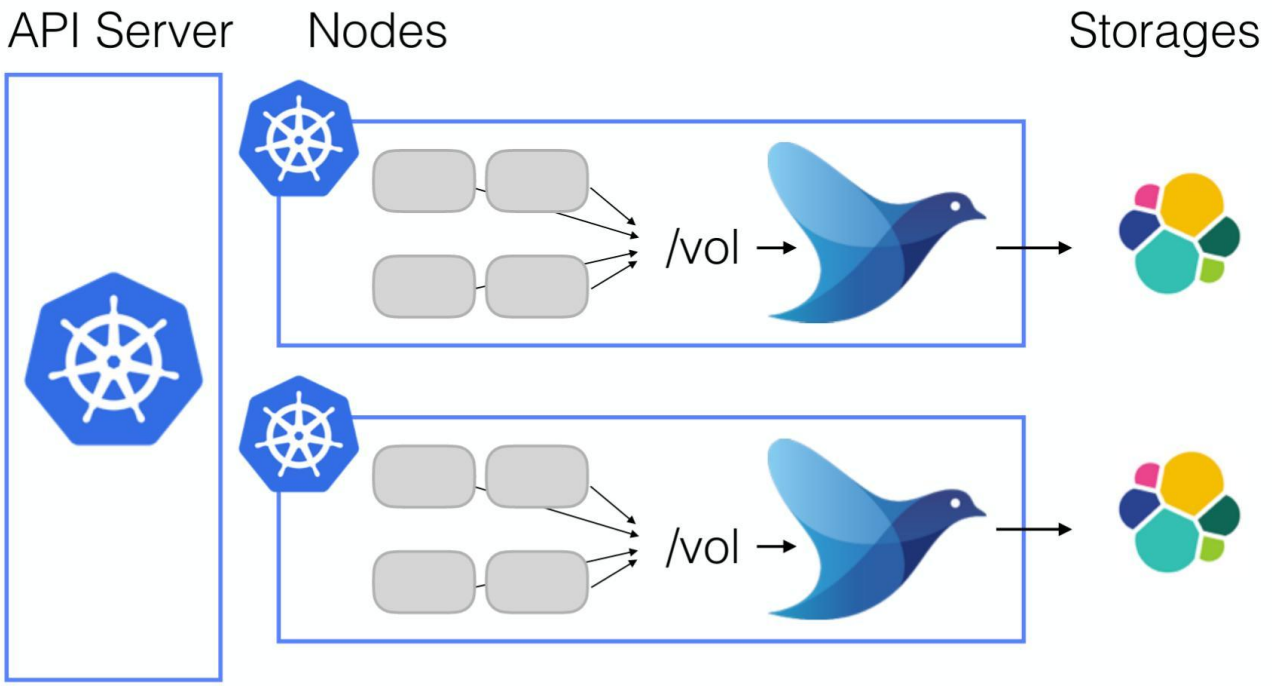
Kubernetes: Fluentd Log Collection



Fluentd: Enrich Logs with Kubernetes Metadata



Storing the Logs



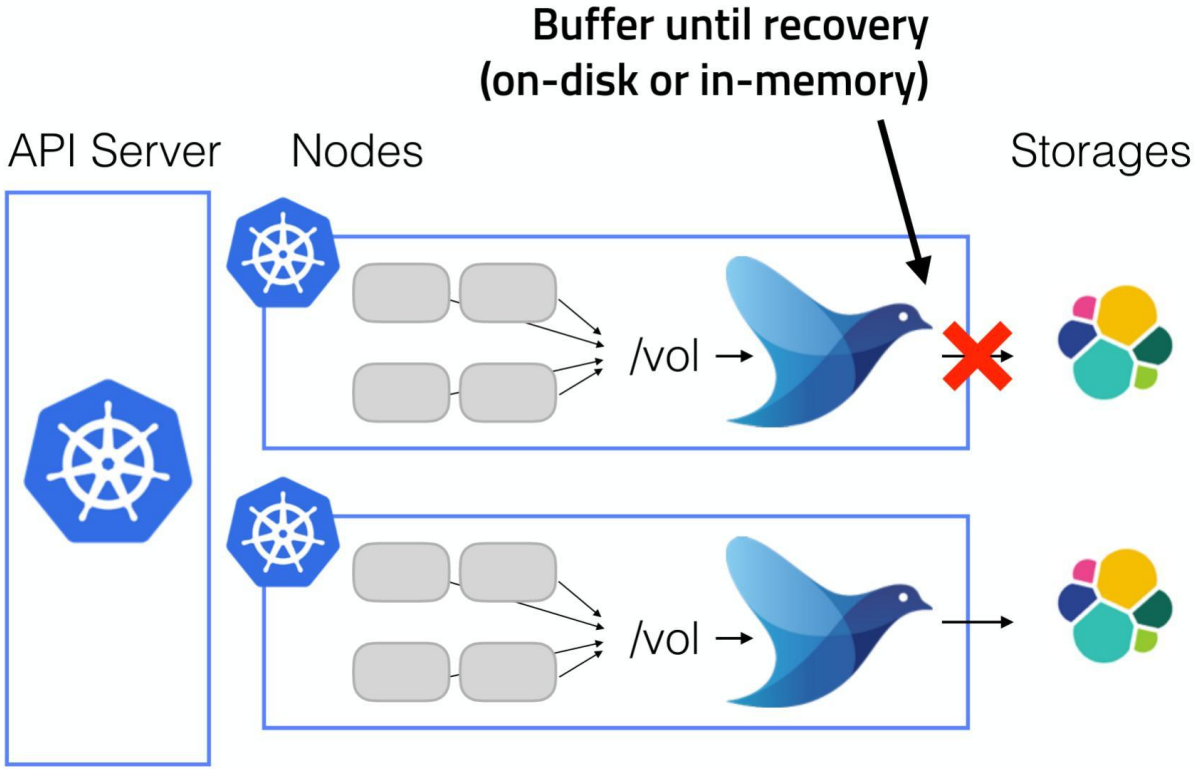
Example: Logs from the Container


```
{
  "log": "Server started\n",
  "stream": "stderr",
  "time": "2015-05-05T19:54:41.240447294Z"
}
```

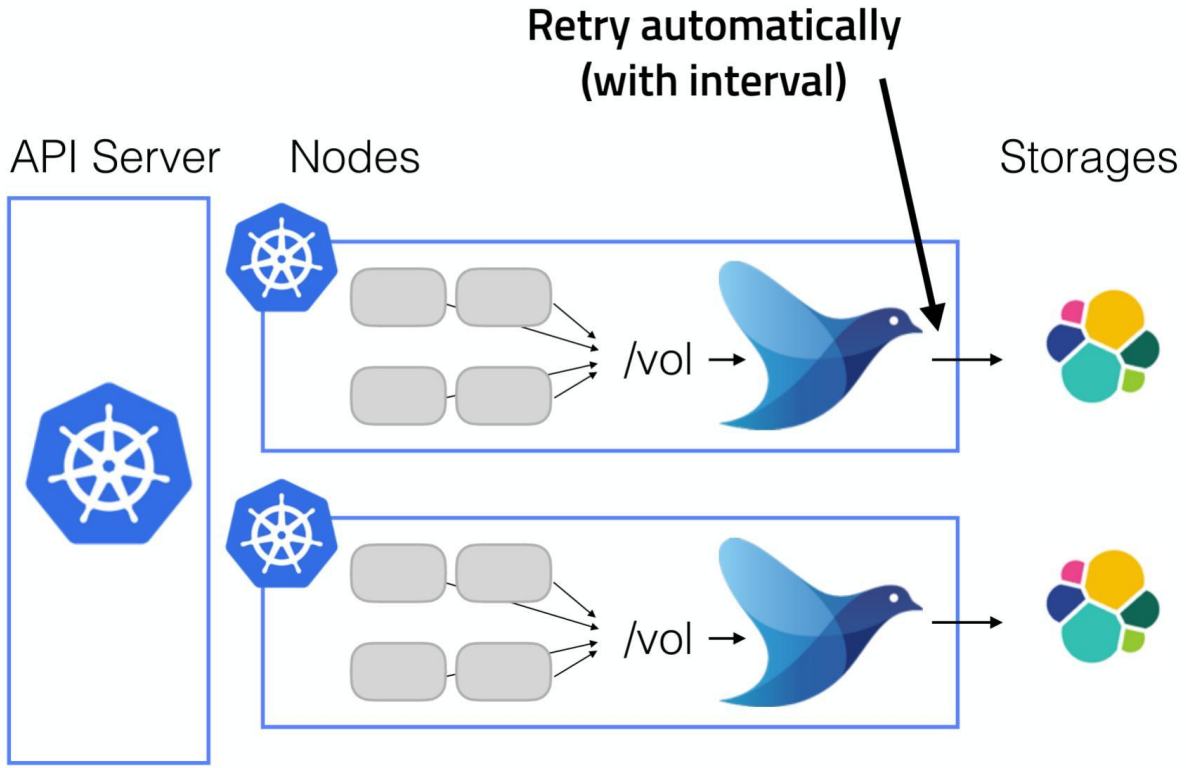
Logs Stored by Fluentd


```
{
 "log": "Server started\n",
 "stream": "stderr",
 "docker": {
  "id": "deadbeef123",
 }
 "kubernetes": {
  "host": "jimmi-redhat.localnet",
  "pod_name":"example-pod",
  "pod_id": "c76927af-f563-11e4-b32d-54ee7527188d",
  "container_name": "example-container",
  "namespace_name": "default",
  "namespace_id": "23437884-8e08-4d95-850b-e94378c9b2fd",
  "labels": {
   "component": "fabric8Console"
  }
 }
}
```

Fluentd: Error Handling



Fluentd: Retrying



Fluentd: Kubernetes DaemonSet

Pre-built configuration:

https://github.com/fluent/fluentd-kubernetes-daemonset

```
17      spec:
18        containers:
19        - name: fluentd
20          image: quay.io/fluent/fluentd-kubernetes-daemonset
21          env:
22            - name:  FLUENT_ELASTICSEARCH_HOST
23              value: "elasticsearch-logging"
24            - name:  FLUENT_ELASTICSEARCH_PORT
25              value: "9200"
```

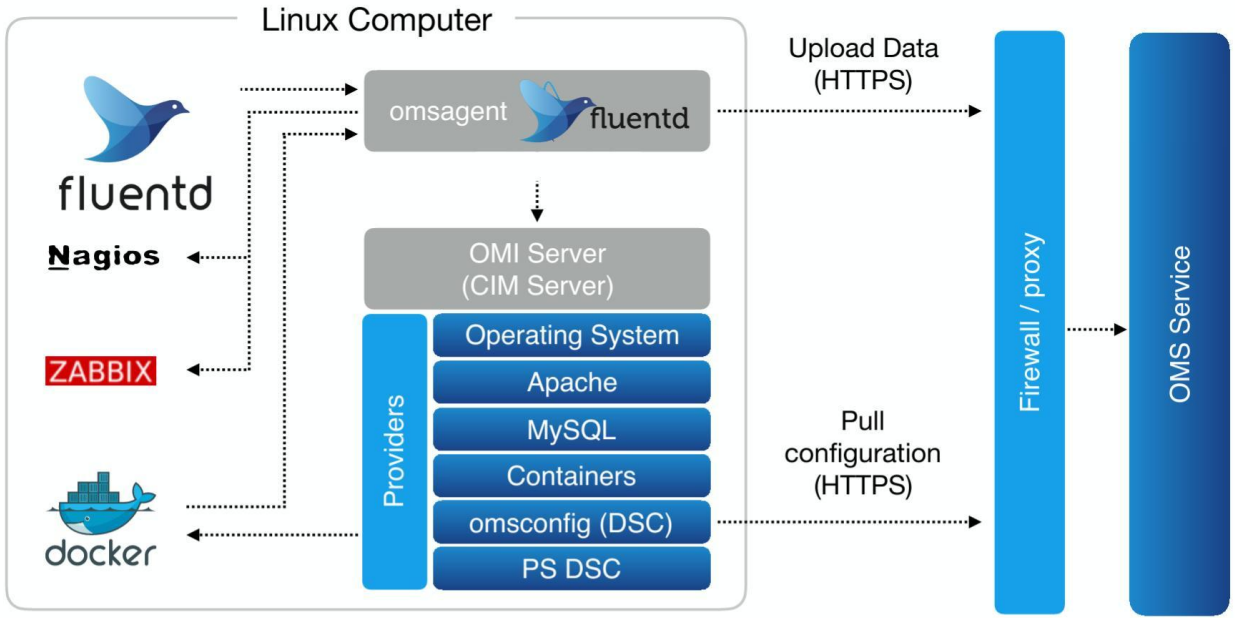
Fluentd Real World Use Cases

TREASURE DATA

Microsoft

Microsoft

Operations Management Suite uses Fluentd: "The core of the agent uses an existing open source data aggregator called Fluentd. Fluentd has hundreds of existing plugins, which will make it really easy for you to add new data sources."



Extending the Fluentd Ecosystem

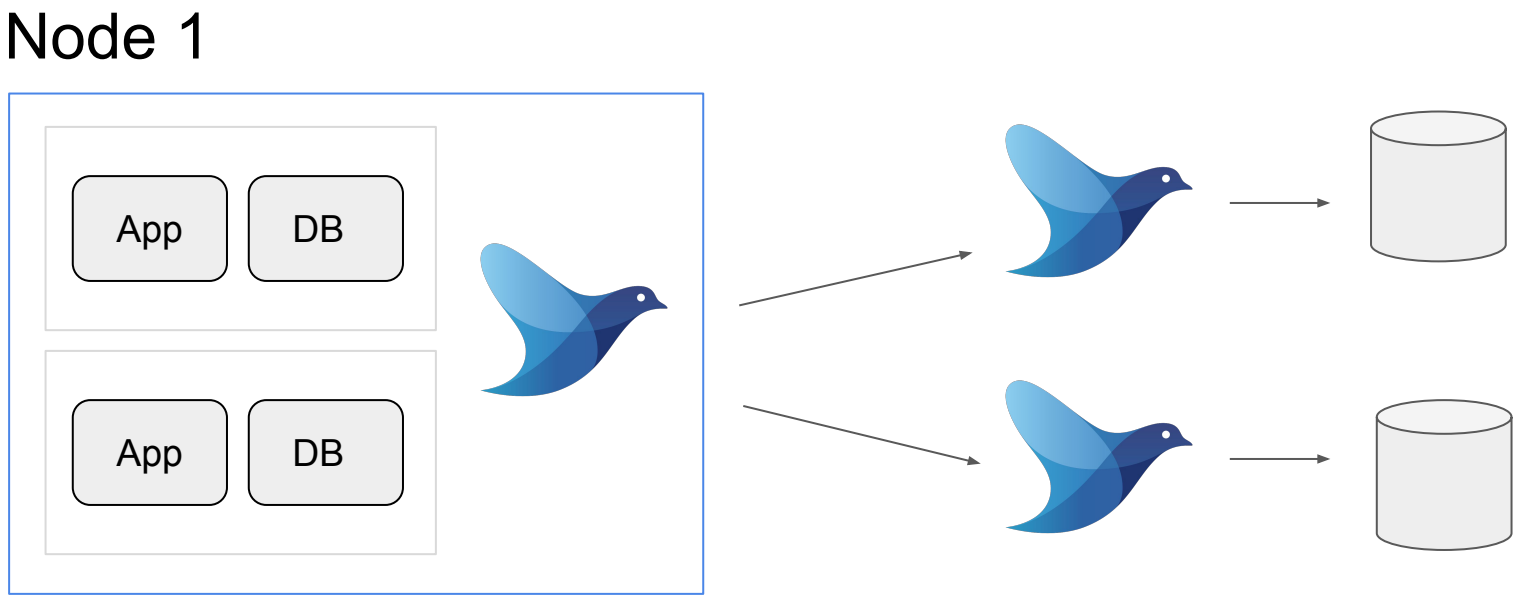
TREASURE DATA

Fluentd Modes

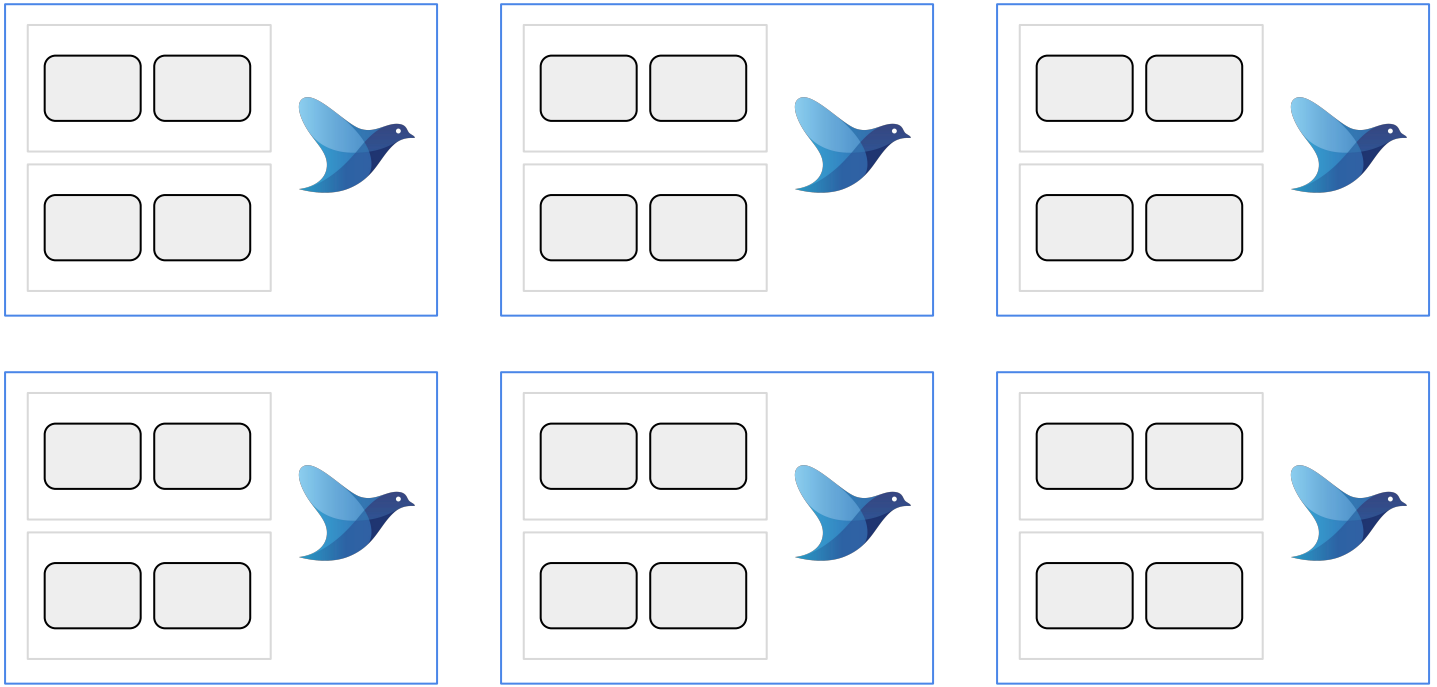
- Log **F**orwarder
- Log **A**ggregator

Log Aggregator =
(Forwarder + Buffering Capabilities)

Edge Nodes / Forward to Aggregators



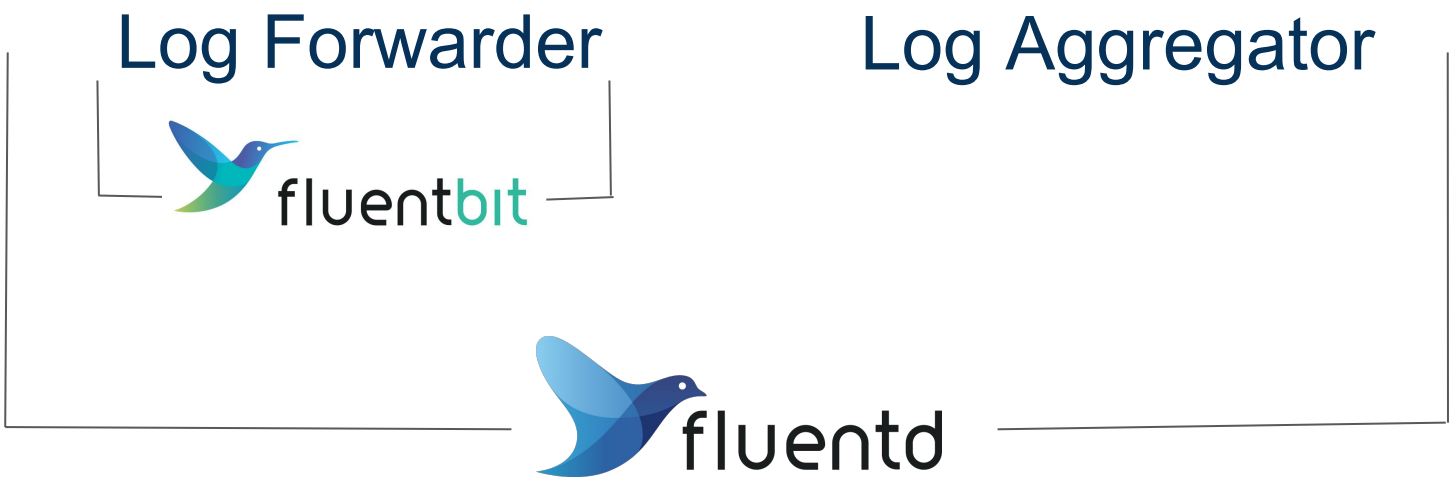
Edge Nodes & Costs



Edge Nodes & Costs

- Fluentd requires ~40MB as minimum
- Deploying a few hundred could be expensive
- Can we make **Forward cheaper** ?

Forwarder & Aggregator



fluentbit

TREASURE DATA

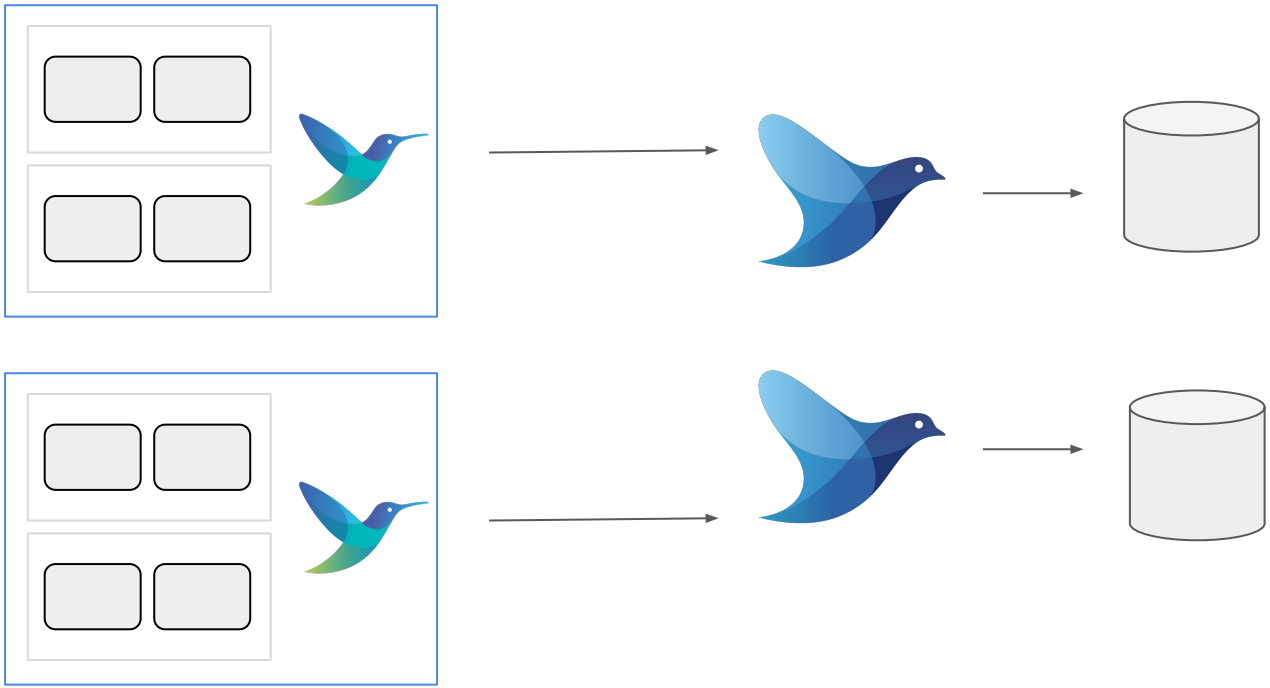
About Fluent Bit

- Written in **C**
- Pluggable Architecture
- Built-in Reliability
- Event Driven - Async I/O

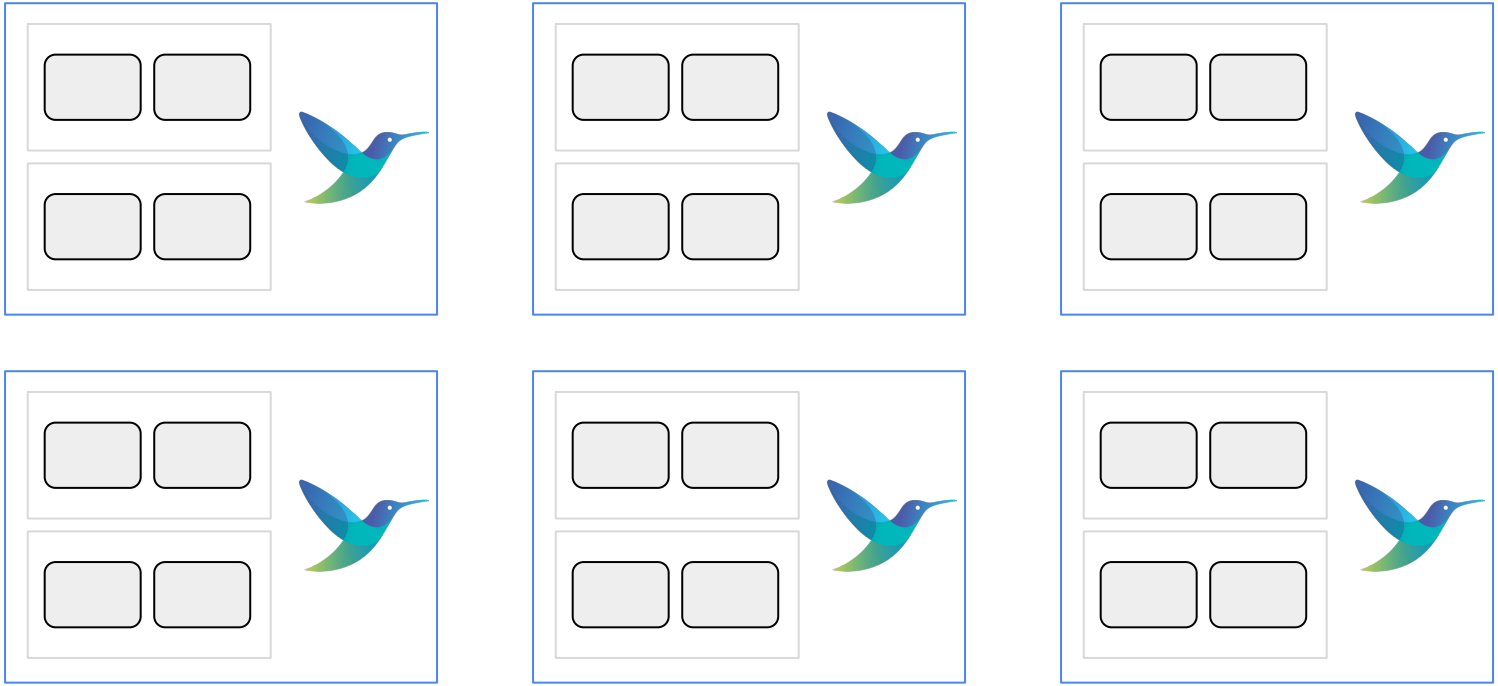
Why Fluent Bit as a Forwarder

- Features
  - Input, Filter and Output Plugins
  - Built-in parsing support
  - Minimum memory required **450KB**

Edge Nodes / Forward to Aggregators

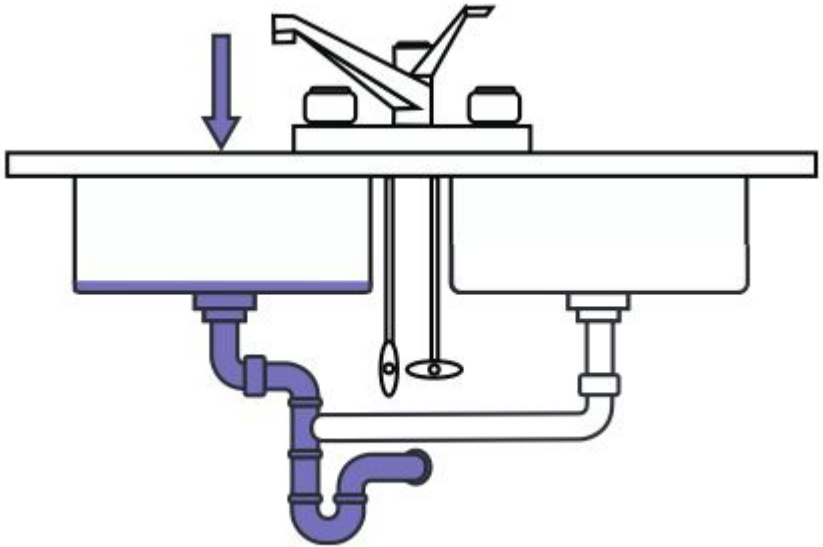


Cheap Forwarding



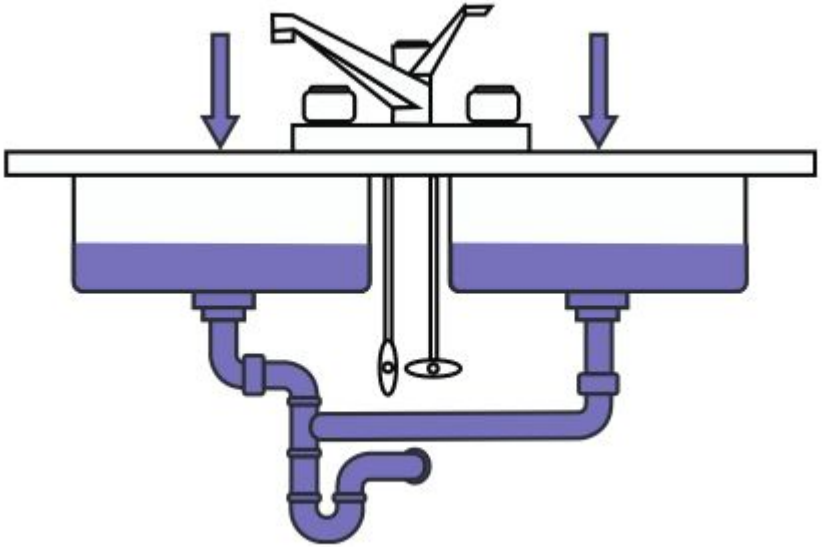
Cloud Native Features

- Backpressure



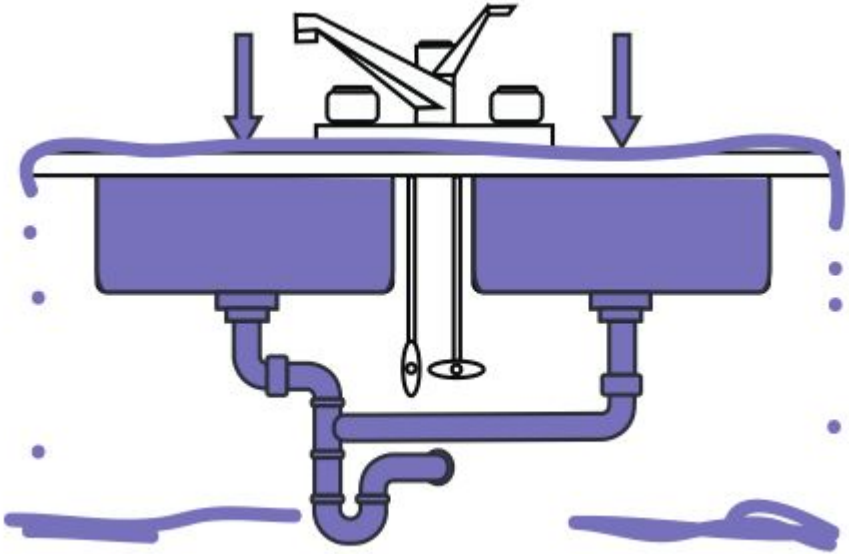
Cloud Native Features

- Backpressure



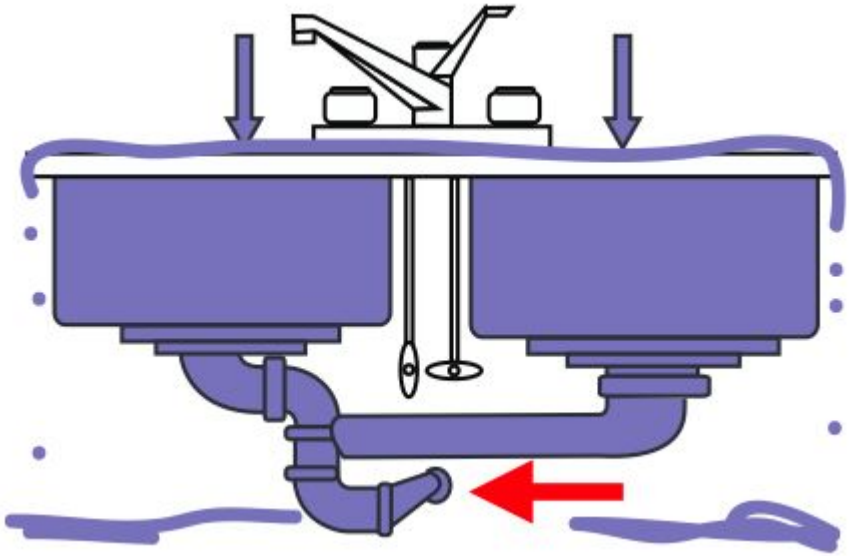
Cloud Native Features

- Backpressure



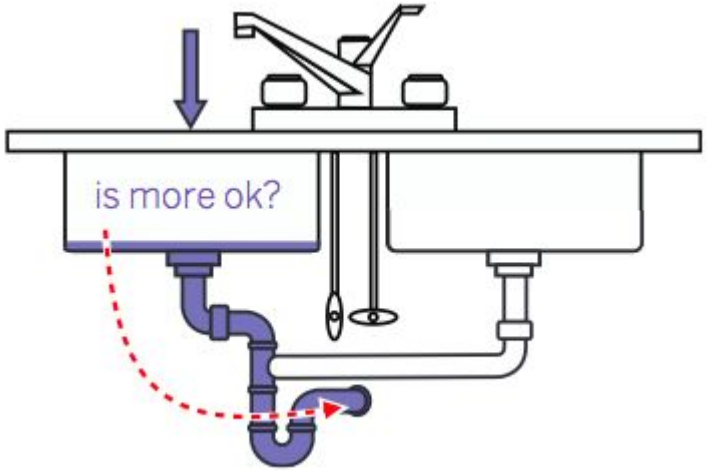
Cloud Native Features

- Backpressure



Cloud Native Features

- Backpressure Solved



Cloud Native Features

- Backpressure Handling
- Security
- Kubernetes Metadata Filter
- Timestamps w/Nanoseconds (upcoming v0.12)

Fluent Bit, what else ?

Networking and Co-routines

Easier implementation of output plugins that interact with networking operations like socket(), connect(), read(), write(), etc.

Fluent Bit provides non-blocking networking API that uses the event-loop with co-routines to implement:

- Network I/O
- TLS/SSL usage
- HTTP Client

Fluent Bit Kubernetes DaemonSet

Github Repository

- https://github.com/fluent/fluent-bit-kubernetes-daemonset

Docker Hub Image

- fluent/fluent-bit-kubernetes-daemonset

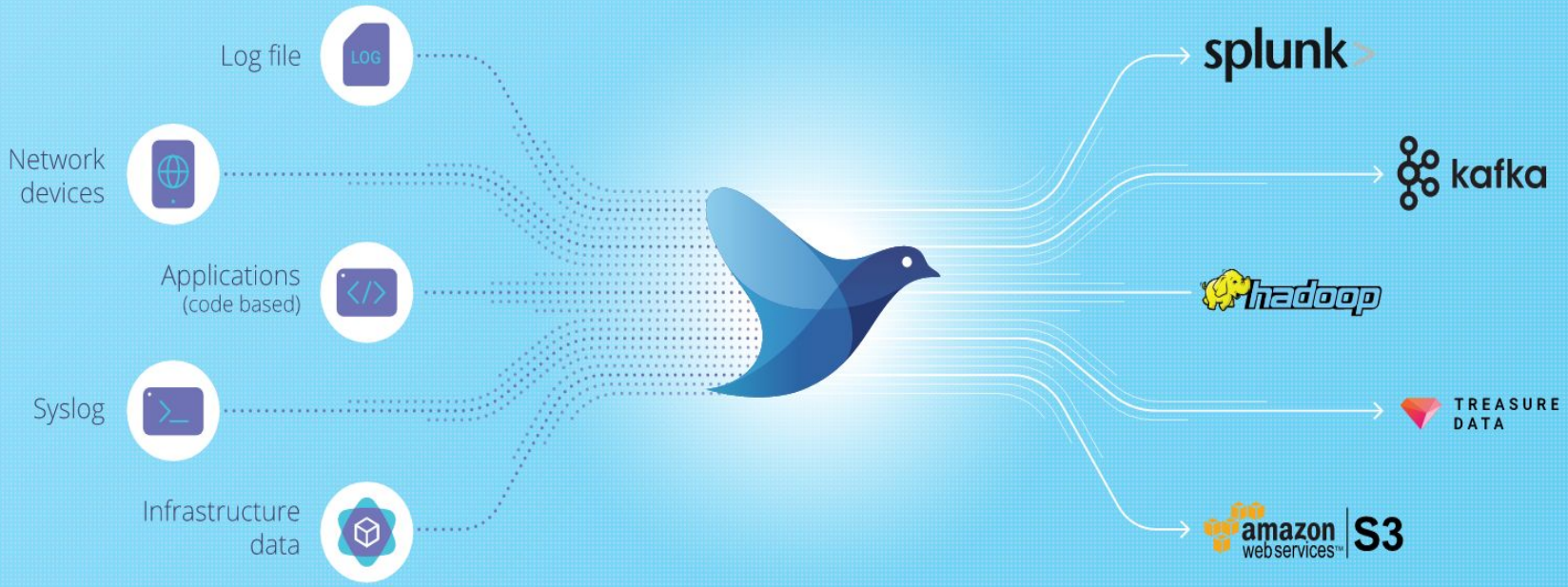
Fluent Bit

Project information

- Web site fluentbit.io
- Github fluent/fluent-bit

Contact / Community

- Slack http://slack.fluentd.org



Fluentd Enterprise

A unified logging layer with security, scale, and reliability built for the enterprise.

fluentd
enterprise

End-to-End Security

Certified Enterprise Plugins

World-Class Support

Questions?

Anurag Gupta - Product Manager Fluentd Enterprise - anurag@treasure-data.com
Eduardo Silvia - Open Source Engineer - eduardo@treasure-data.com

Interested in Fluentd Enterprise?

https://www.treasuredata.com/fluentd-enterprise/

Resources for Fluentd Open Source

http://www.fluentd.org

Community

http://slack.fluentd.org