Centralized
Java Logging Patterns

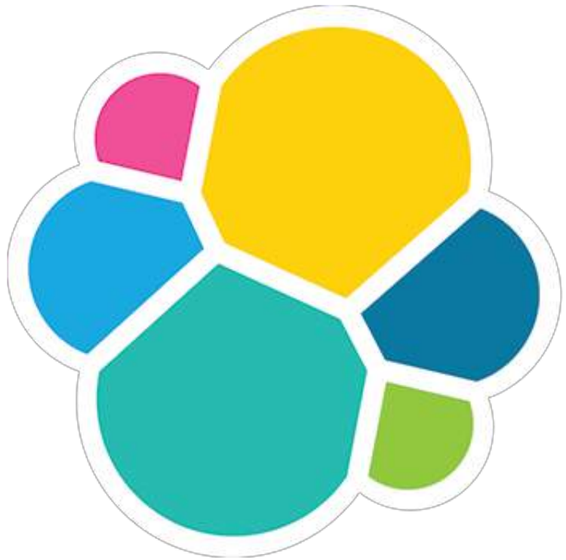
Philipp Krenn @xeraa


ALL THE THINGS!


THIS IS FINE.

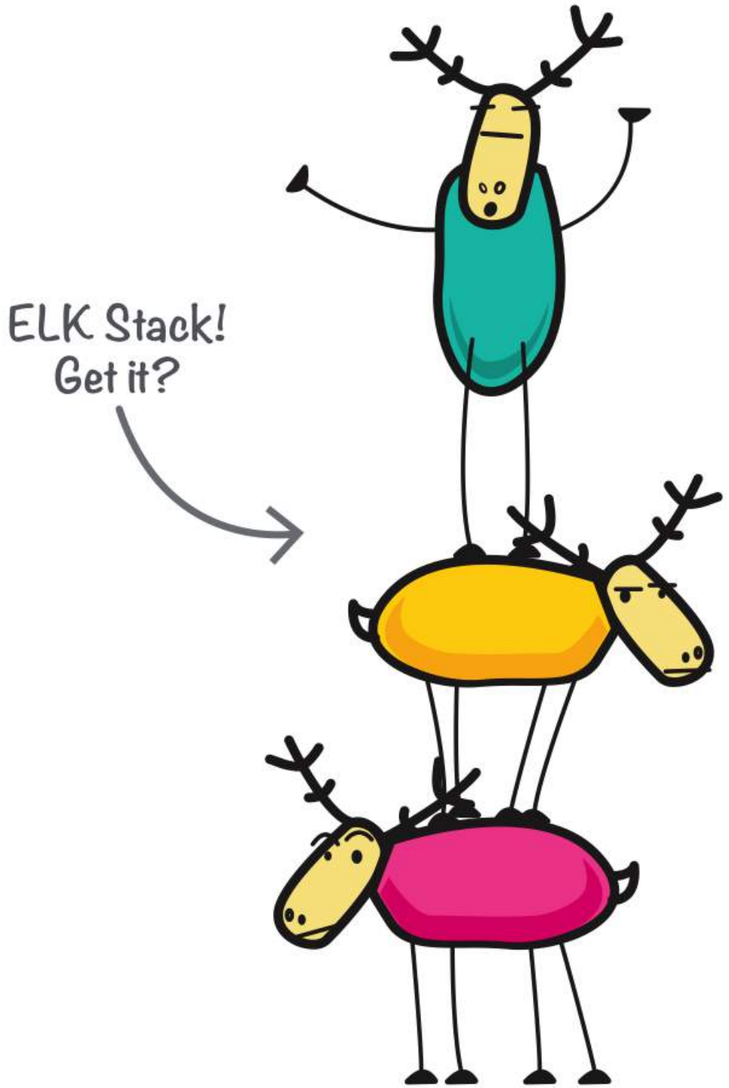

me looking
for the bug
7.2 GB
of log file


TH-
THIS IS NOT FINE!!
OH MY GOD EVERY THINGS ON FIRE
WHAT THE HELL IS MY PROBLEM



elastic

Developer 🥑

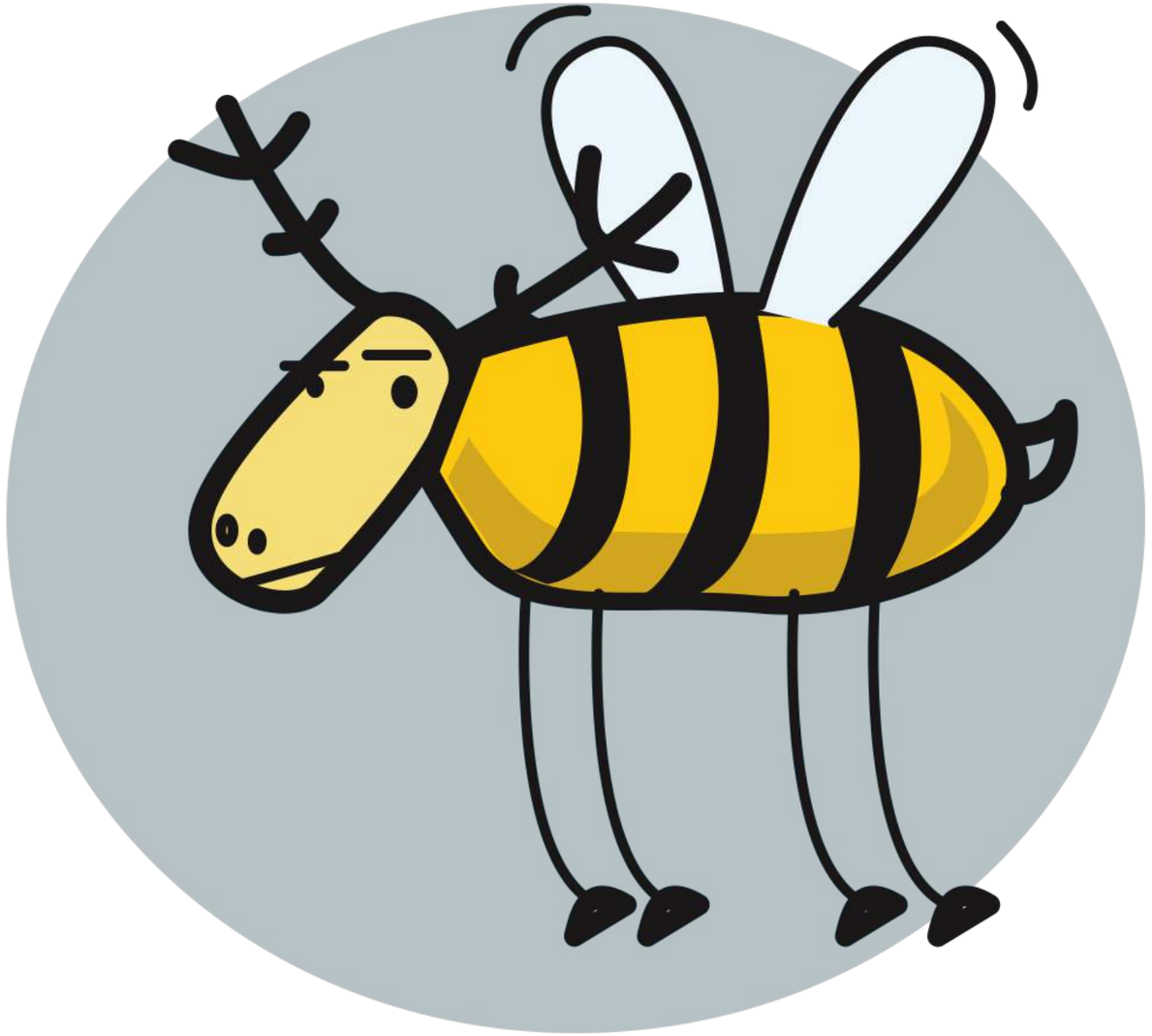


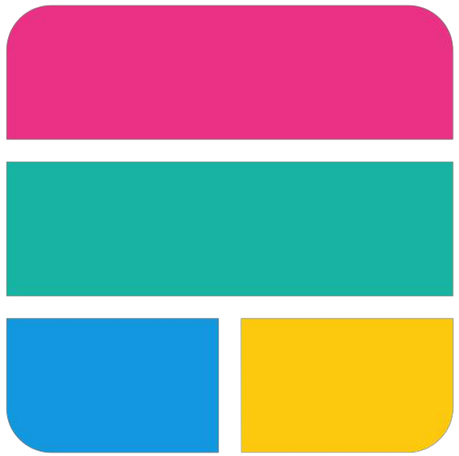
E Elasticsearch

L Logstash

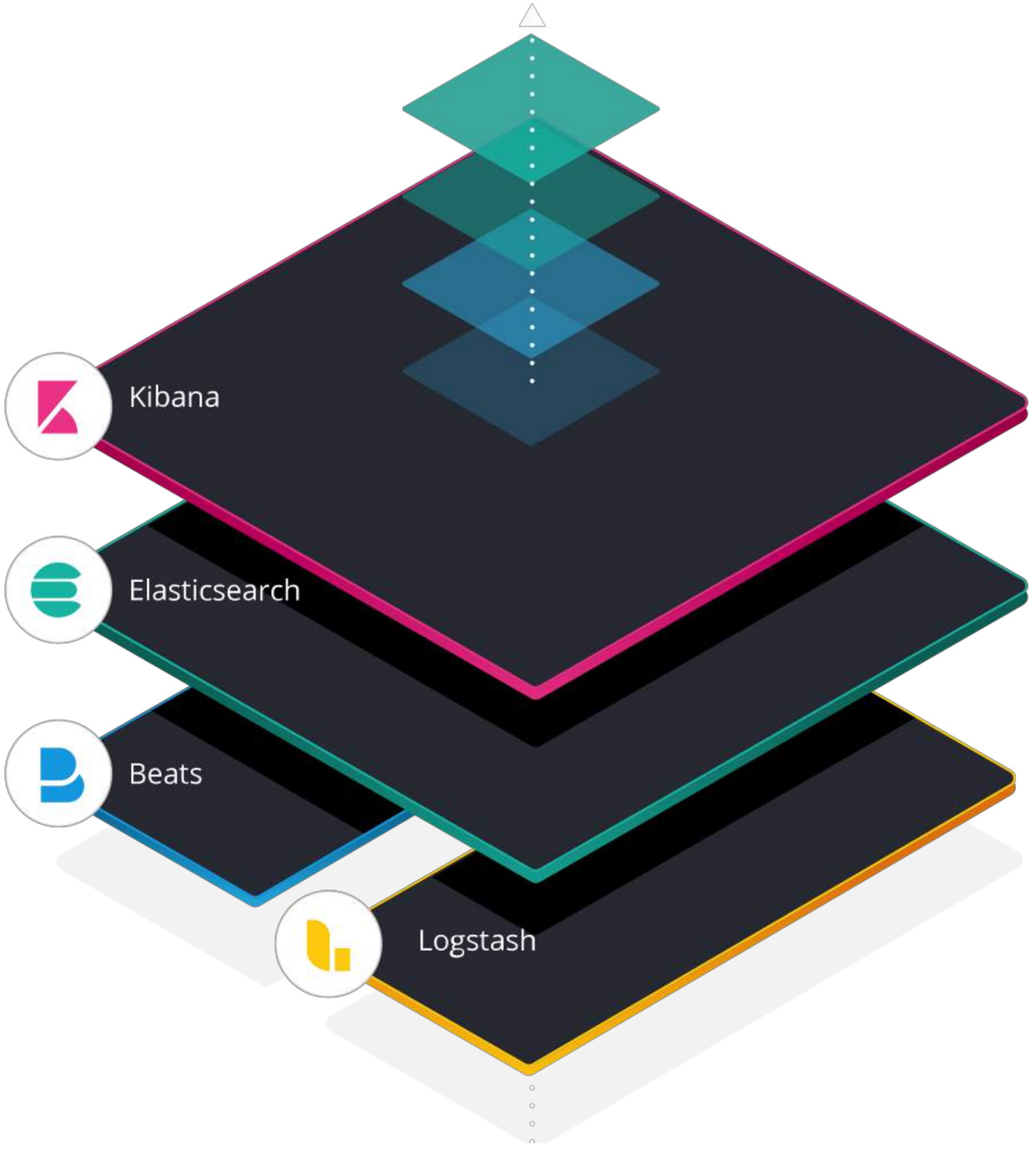
K Kibana

Apparently, I'm an ELKB personality.





elastic stack


Kibana
Elasticsearch
Beats
Logstash

Disclaimer

I build **highly** monitored Hello World apps

Example: Java

SLF4J, Logback, MDC

with https://github.com/elastic/java-ecs-logging
or https://github.com/logstash/logstash-logback-encoder

Alternative https://github.com/vy/log4j2-logstash-layout

And Everywhere Else

.NET: NLog

JavaScript: Winston

Python: structlog

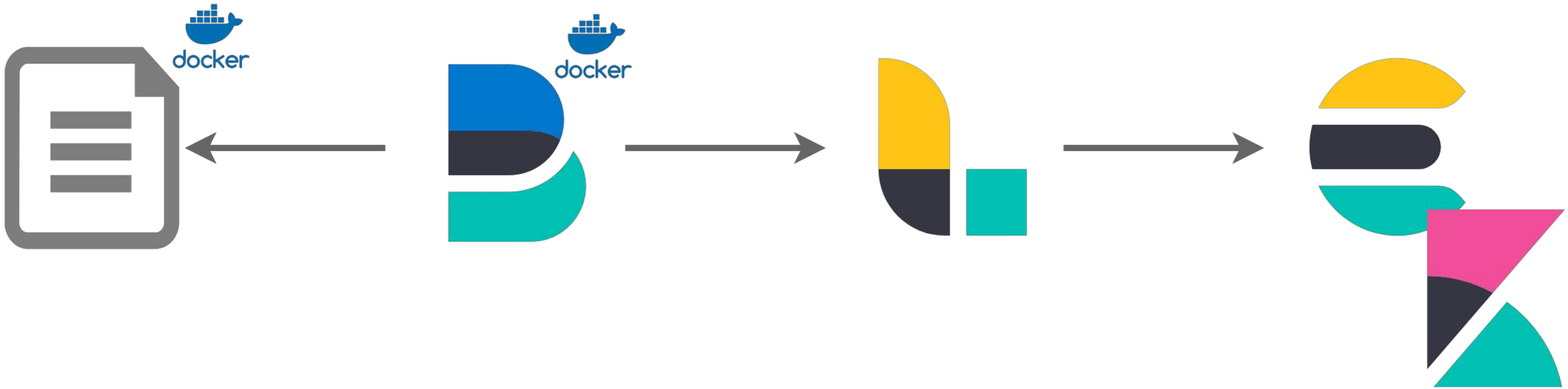
PHP: Monolog

Anti-Pattern: `print`

```
System.out.println("Oops");
```

Anti-Pattern: Coupling

Parse


docker
docker

Bind Mount Logs

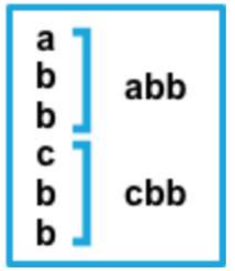
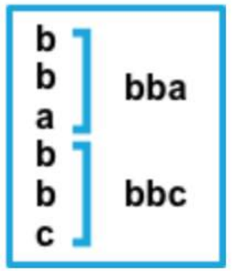
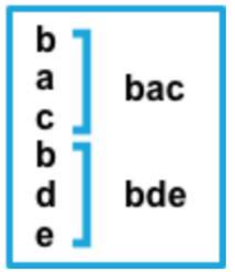
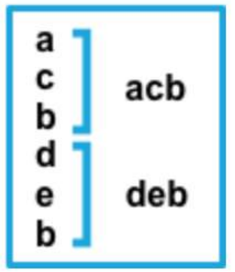
```
java_app:
  volumes:
    - './logs-docker/:/logs/'
  ...


filebeat_for_logstash:
  volumes:
    - './logs-docker/:/mnt/logs/:ro'
  ...
```

Collect Log Lines

```
filebeat.inputs:
- type: log
  paths:
    - /mnt/logs/*.log
```

Metadata

```
processors:
  - add_host_metadata: ~
```

Setting for `negate`	Setting for `match`	Result	Example `pattern: ^b`
`false`	`after`	Consecutive lines that match the pattern are appended to the previous line that doesn't match.	
`false`	`before`	Consecutive lines that match the pattern are prepended to the next line that doesn't match.	
`true`	`after`	Consecutive lines that don't match the pattern are appended to the previous line that does match.	
`true`	`before`	Consecutive lines that don't match the pattern are prepended to the next line that does match.	

Test Multiline Pattern

https://www.elastic.co/guide/en/beats/filebeat/current/_test_your_regexp_pattern_for_multiline.html

The Go Playground Run Format Imports Share

```
1  package main
2
3  import (
4          "fmt"
5          "regexp"
6          "strings"
7  )
8
9  var pattern = `^\[`
10 var negate = true
11
12 var content = `[2019-05-21 09:35:52.004] ERROR net.xeraa.logging.LogMe [main] - user_expe
13 java.lang.RuntimeException: Bad runtime...
14         at net.xeraa.logging.LogMe.main(LogMe.java:30)
15 [2019-05-21 09:35:52.006] TRACE net.xeraa.logging.LogMe [main] - session=59, loop=16 - It
16 [2019-05-21 09:35:52.007] DEBUG net.xeraa.logging.LogMe [main] - session=59, loop=16 - Co
17 `
18
19 func main() {
20         regex, err := regexp.Compile(pattern)
21         if err != nil {
22                 fmt.Println("Failed to compile pattern: ", err)
```

```
matches line
false   [2019-05-21 09:35:52.004] ERROR net.xeraa.logging.LogMe [main] - user_experience=🤬,
true    java.lang.RuntimeException: Bad runtime...
true            at net.xeraa.logging.LogMe.main(LogMe.java:30)
false   [2019-05-21 09:35:52.006] TRACE net.xeraa.logging.LogMe [main] - session=59, loop=16
false   [2019-05-21 09:35:52.007] DEBUG net.xeraa.logging.LogMe [main] - session=59, loop=16
true
```

Grok

https://github.com/logstash-plugins/logstash-patterns-core/blob/master/patterns/grok-patterns

96 lines (85 sloc) | 5.21 KB

```
USERNAME [a-zA-Z0-9._-]+
USER %{USERNAME}
EMAILLOCALPART [a-zA-Z][a-zA-Z0-9_.+-=:]+
EMAILADDRESS %{EMAILLOCALPART}@%{HOSTNAME}
INT (?:[+-]?(?:[0-9]+))
BASE10NUM (?<![0-9.+-])(?>[+-]?(?:(?:[0-9]+(?:\.[0-9]+)?)|(?:\.[0-9]+)))
NUMBER (?:%{BASE10NUM})
BASE16NUM (?<![0-9A-Fa-f])(?:[+-]?(?:0x)?(?:[0-9A-Fa-f]+))
BASE16FLOAT \b(?<![0-9A-Fa-f.])(?:[+-]?(?:0x)?(?:(?:[0-9A-Fa-f]+(?:\.[0-9A-Fa-f]*)?)|(?

POSINT \b(?:[1-9][0-9]*)\b
NONNEGINT \b(?:[0-9]+)\b
WORD \b\w+\b
NOTSPACE \S+
SPACE \s*
DATA .*?
GREEDYDATA .*
QUOTEDSTRING (?>(?<!\\)(?>"(?>\\.|[^\\"]+)+"|""|(?>'(?>\\.|[^\\']+)+')|''|(?>`(?>\\.|[^
UUID [A-Fa-f0-9]{8}-(?:[A-Fa-f0-9]{4}-){3}[A-Fa-f0-9]{12}
# URN, allowing use of RFC 2141 section 2.3 reserved characters
URN urn:[0-9A-Za-z][0-9A-Za-z-]{0,31}:(?:%[0-9a-fA-F]{2}|[0-9A-Za-z()+,.:=@;$_!*'/?#-])

# Networking
MAC (?:%{CISCOMAC}|%{WINDOWSMAC}|%{COMMONMAC})
CISCOMAC (?:(?:[A-Fa-f0-9]{4}\.){2}[A-Fa-f0-9]{4})
WINDOWSMAC (?:(?:[A-Fa-f0-9]{2}-){5}[A-Fa-f0-9]{2})
```

Dev Tools

Grok Debugger

elastic @xeraa

```
[2018-09-28 10:30:38.516] ERROR net.xeraa.logging.LogMe [main] -
                           user_experience=🤬, session=46, loop=15 -
                           Wake me up at night
java.lang.RuntimeException: Bad runtime...
  at net.xeraa.logging.LogMe.main(LogMe.java:30)
```

```
^\[%{TIMESTAMP_ISO8601:@timestamp}\]%{SPACE}%{LOGLEVEL:log.level}
%{SPACE}%{USERNAME:log.package}%{SPACE}\[%{WORD:log.method}\]
%{SPACE}-%{SPACE}%{GREEDYDATA:log.labels}%{SPACE}-%{SPACE}
%{GREEDYDATA:message}(?:\n+(?<stacktrace>(?:.|\r|\n)+))?
```

Elastic Common Schema

https://github.com/elastic/ecs

Event fields

The event fields are used for context information about the data itself.

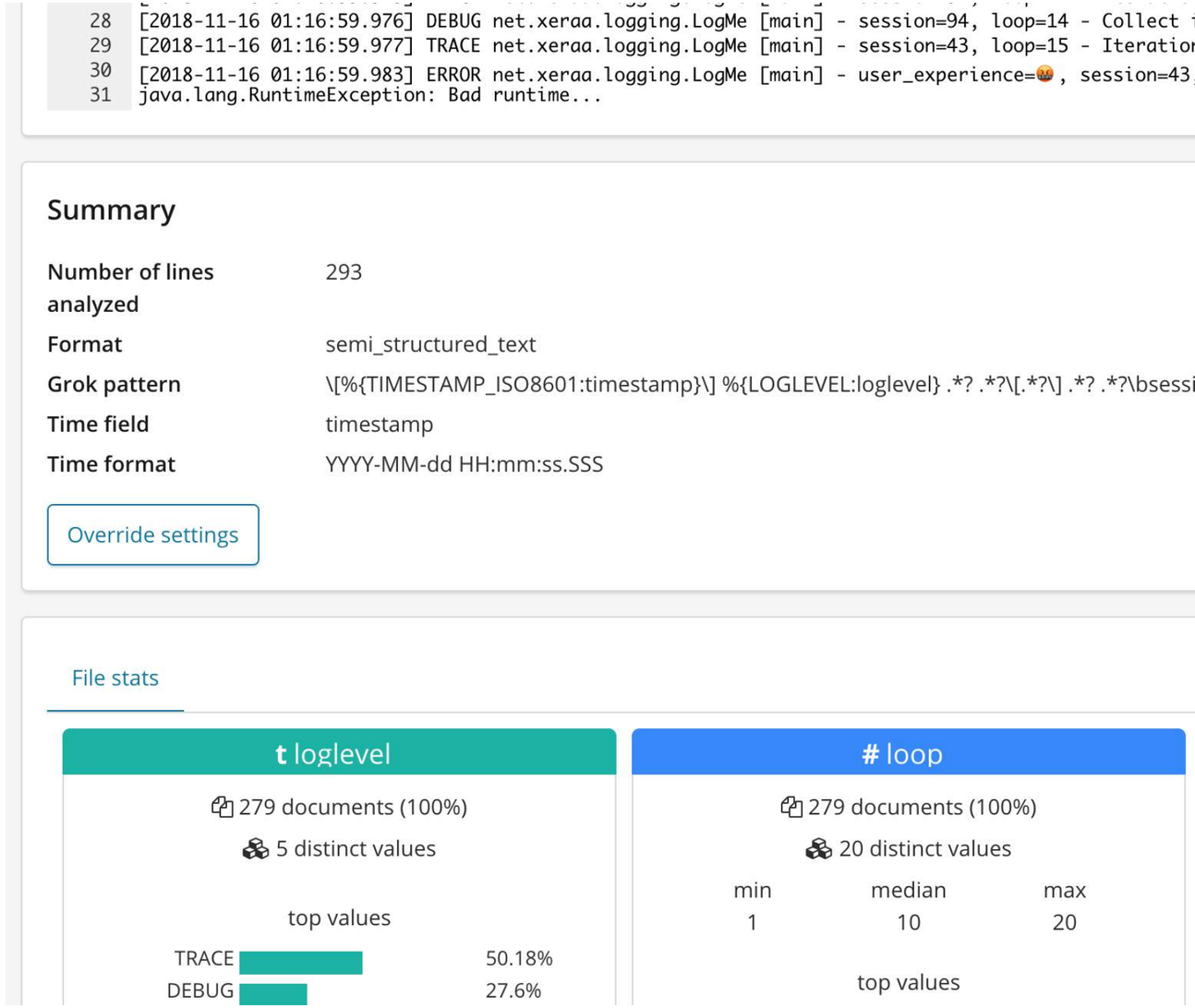
Field	Description	Level	Type	Example
event.id	Unique ID to describe the event.	core	keyword	`8a4f500d`
event.category	Event category. This can be a user defined category.	core	keyword	`metrics`
event.type	A type given to this kind of event which can be used for grouping. This is normally defined by the user.	core	keyword	`nginx-stats-metrics`
event.action	The action captured by the event. The type of action will vary from system to system but is likely to include actions by security services, such as blocking or quarantining; as well as more generic actions such as login	core	keyword	`reject`

elastic

@xeraa

Machine Learning

Data Visualizer



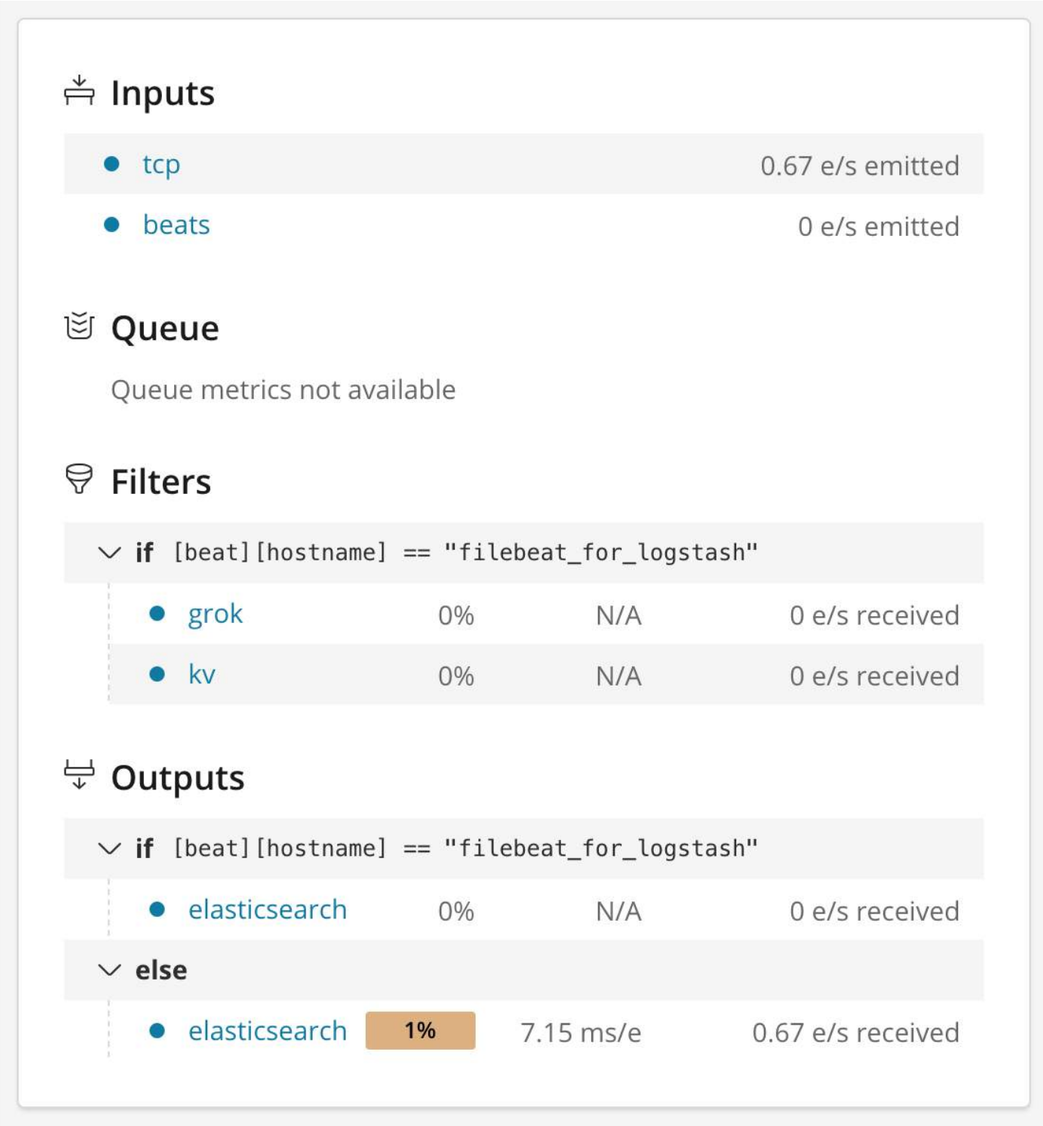
elastic @xeraa

Logstash Key Value Filter for MDC

```
kv {
  source => "labels"
  field_split => ","
  trim_key => " "
}
```

Monitoring: Logstash Pipeline

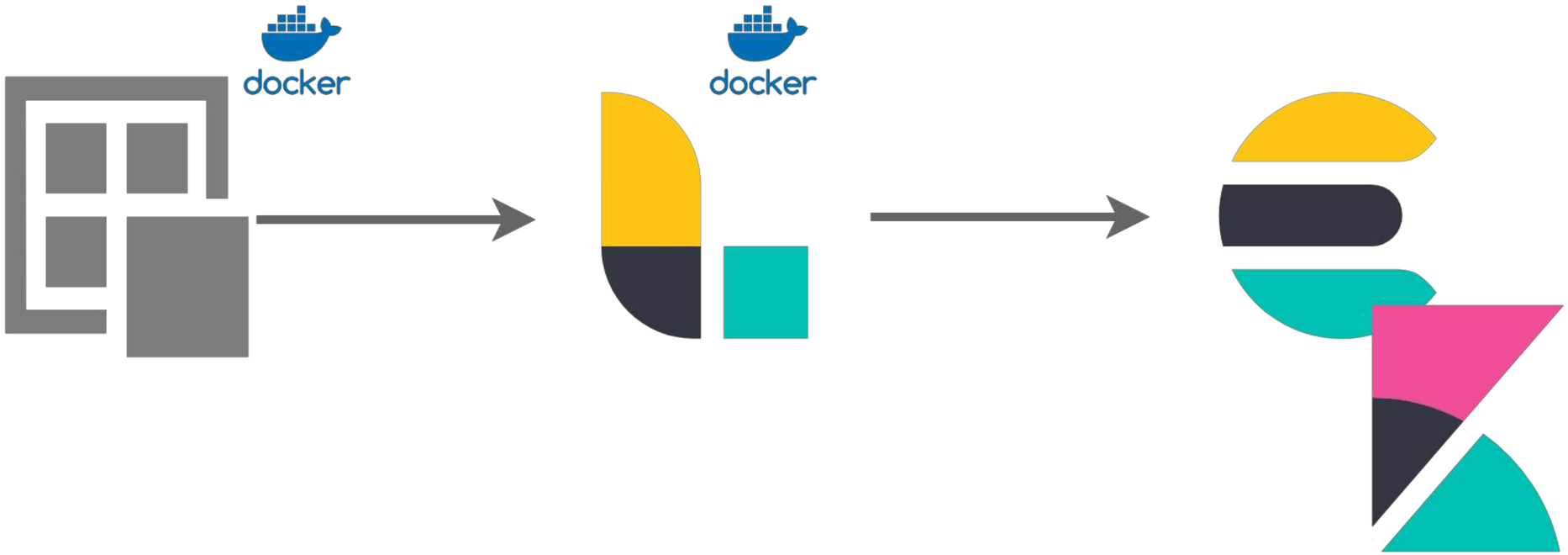
Plus other components



Pro: No change

Con: Regular expression, multiline, format changes

Send 


docker
docker

logback.xml

```
<appender name="logstash" class="net.logstash.logback.appender.LogstashAccessTcpSocketAppender">
    <destination>logstash:4560</destination>
    <encoder class="net.logstash.logback.encoder.LogstashEncoder"/>
</appender>
```

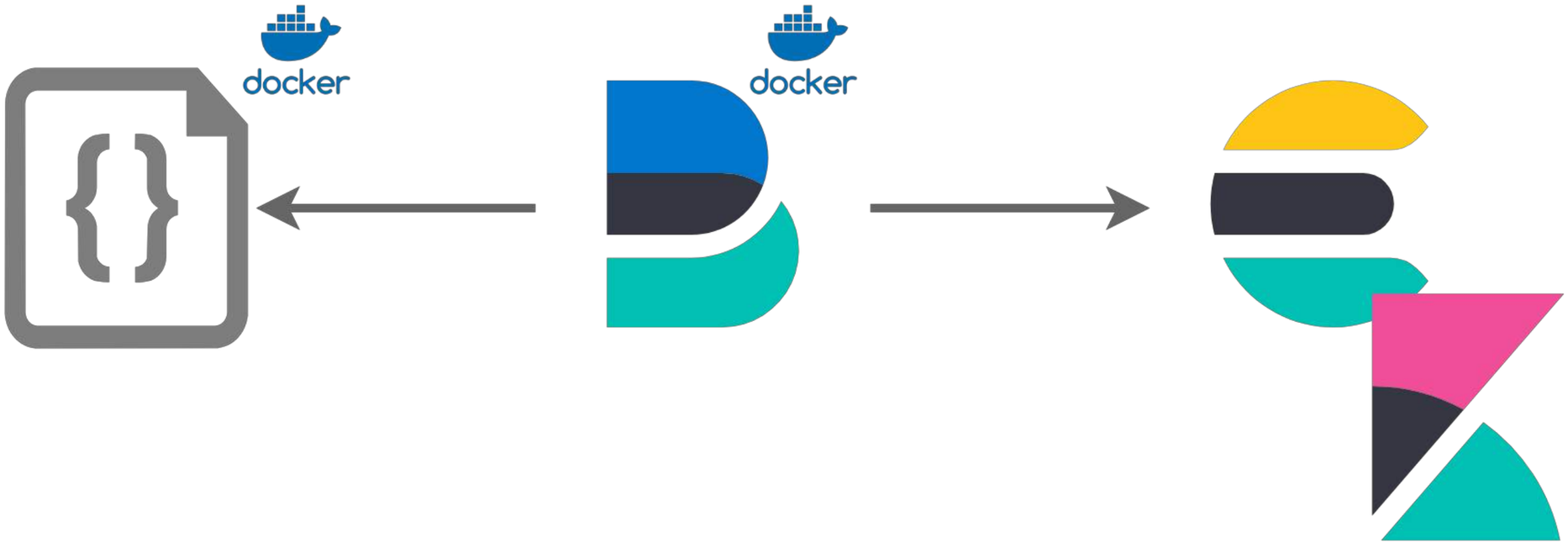
Syslog

```
$WorkDirectory /var/lib/rsyslog
$FileOwner root
$FileGroup adm
$FileCreateMode 0640
$DirCreateMode 0755
$Umask 0022

include(file="/etc/rsyslog.d/*.conf" mode="optional")

module(load="immark")
module(load="imuxsock")
module(load="imklog")

*.*     @@filebeat_syslog:9000
```

Pro: No files

Con: Outages & coupling

Structure


docker
docker

Collect JSON

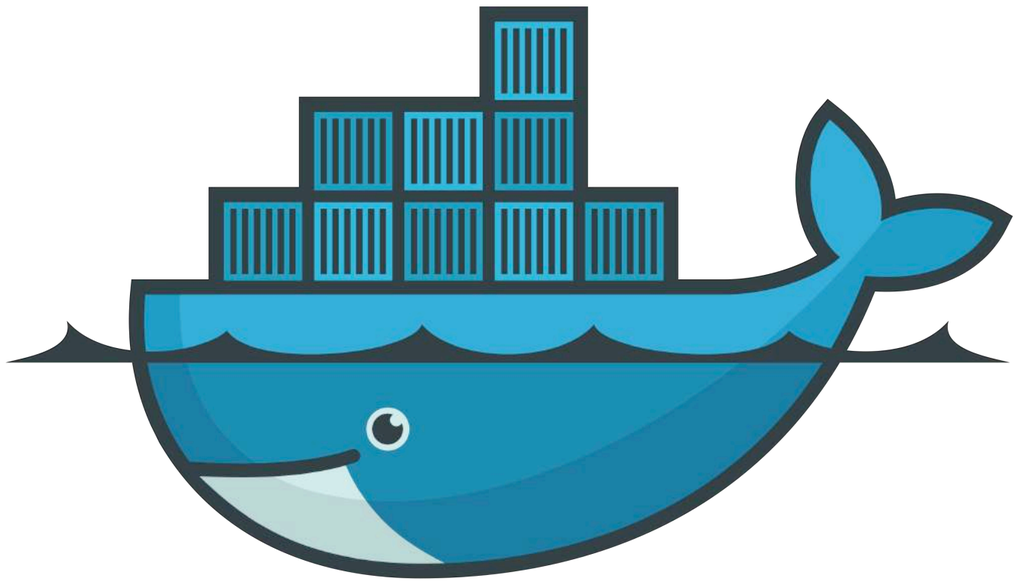
```
filebeat.input:
- type: log
  paths:
    - /mnt/logs/*.json
  json:
    message_key: message
    keys_under_root: true
```

Stack(trace) Hash

Pro: Right format

Con: JSON serialization overhead

Containerize 



docker

elastic

@xeraa

Where to Log?

STDOUT

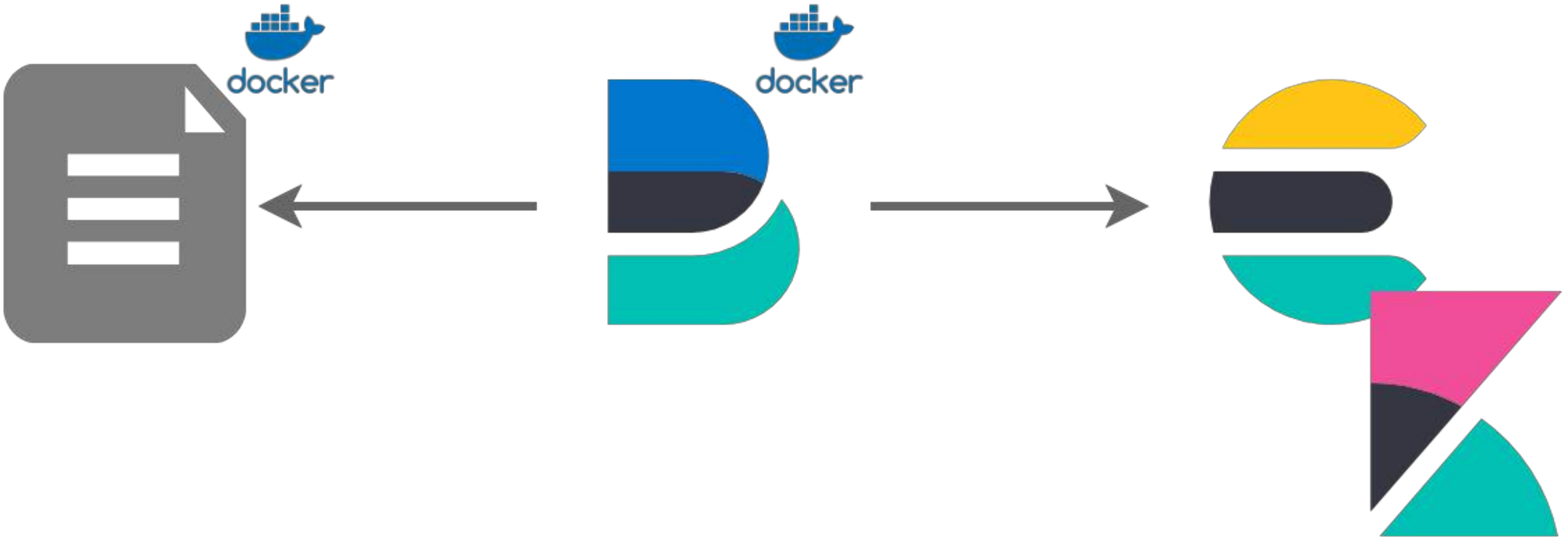
`json-file`

Don't Use Highlighting

```
<appender name="console" class="ch.qos.logback.core.ConsoleAppender">
    <encoder>
        <pattern>[%d{yyyy-MM-dd HH:mm:ss.SSS}] %highlight(%-5level) %logger{36} [%thread] - %mdc - %msg %n</pattern>
        <charset>utf8</charset>
    </encoder>
</appender>
```

Where to Put Filebeat?

Sidecar


docker
docker

https://github.com/elastic/beats/tree/master/deploy/docker

Docker Logs

```
filebeat.autodiscover:
  providers:
    - type: docker
      hints.enabled: true

processors:
  - add_docker_metadata: ~
```

Metadata

No Docker metadata with the other methods

```
"docker": {
  "container": {
    "labels": {
      "app": "fizzbuzz",
      "co_elastic_logs/multiline_match": "after",
      "com_docker_compose_config-hash": "41520c6cf2b6a1f3dae4f16d0a6fd76760cdfc38fbfe43a3a3be2e09bdd1b8b5",
      "environment": "production",
      "co_elastic_logs/multiline_pattern": "^\\[",
      "co_elastic_logs/multiline_negate": "true",
      "com_docker_compose_oneoff": "False",
      "com_docker_compose_project": "java-logging",
      "com_docker_compose_service": "java_app",
      "com_docker_compose_container-number": "1",
      "com_docker_compose_version": "1.23.2"
    }
  }
}
```

Missing the Last Line

Waiting for the newline

Hints

```
labels:
  - "app=fizzbuzz"
  - "co.elastic.logs/multiline.pattern=^\\["
  - "co.elastic.logs/multiline.negate=true"
  - "co.elastic.logs/multiline.match=after"
```

Registry File

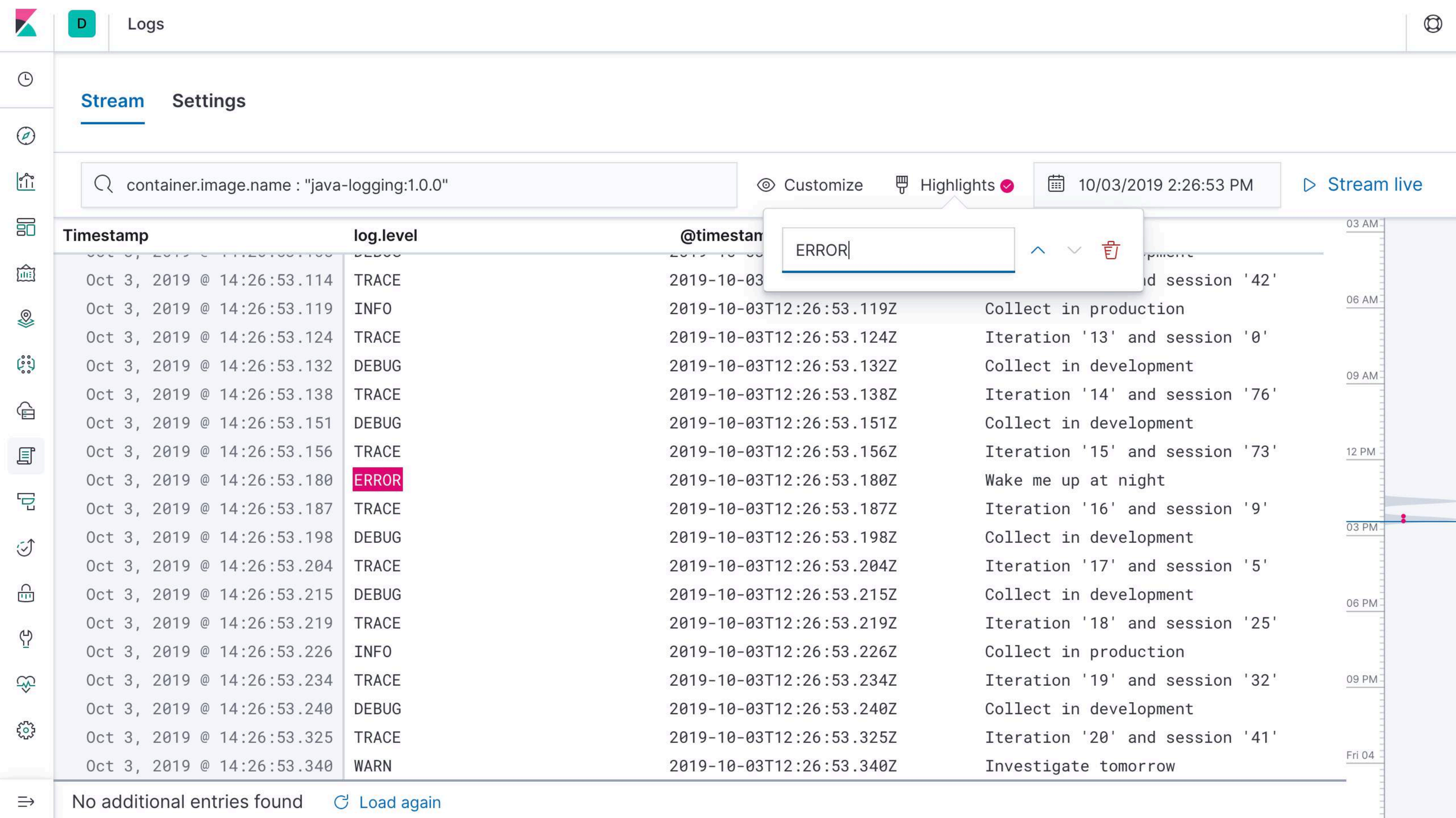
```
filebeat.registry.path: /usr/share/filebeat/data/registry
```

Multi-Index

```
output.elasticsearch:
  hosts: ["http://localhost:9200"]
  indices:
    - index: "docker-java-%{+yyyy.MM}-00"
      when.contains:
        container.name: "java_app"
```

Ingest Pipeline

```
output.elasticsearch:
  hosts: ["http://elasticsearch:9200"]
  index: "docker"
  pipelines:
    - pipeline: "parse_java"
      when.contains:
        container.name: "java_app"
```


D
Logs
Stream
Settings
container.image.name : "java-logging:1.0.0"
Customize
Highlights
10/03/2019 2:26:53 PM
Stream live
ERROR
Timestamp
log.level
@timestam
Oct 3, 2019 @ 14:26:53.114 TRACE 2019-10-03 d session '42'
Oct 3, 2019 @ 14:26:53.119 INFO 2019-10-03T12:26:53.119Z Collect in production
Oct 3, 2019 @ 14:26:53.124 TRACE 2019-10-03T12:26:53.124Z Iteration '13' and session '0'
Oct 3, 2019 @ 14:26:53.132 DEBUG 2019-10-03T12:26:53.132Z Collect in development
Oct 3, 2019 @ 14:26:53.138 TRACE 2019-10-03T12:26:53.138Z Iteration '14' and session '76'
Oct 3, 2019 @ 14:26:53.151 DEBUG 2019-10-03T12:26:53.151Z Collect in development
Oct 3, 2019 @ 14:26:53.156 TRACE 2019-10-03T12:26:53.156Z Iteration '15' and session '73'
Oct 3, 2019 @ 14:26:53.180 ERROR 2019-10-03T12:26:53.180Z Wake me up at night
Oct 3, 2019 @ 14:26:53.187 TRACE 2019-10-03T12:26:53.187Z Iteration '16' and session '9'
Oct 3, 2019 @ 14:26:53.198 DEBUG 2019-10-03T12:26:53.198Z Collect in development
Oct 3, 2019 @ 14:26:53.204 TRACE 2019-10-03T12:26:53.204Z Iteration '17' and session '5'
Oct 3, 2019 @ 14:26:53.215 DEBUG 2019-10-03T12:26:53.215Z Collect in development
Oct 3, 2019 @ 14:26:53.219 TRACE 2019-10-03T12:26:53.219Z Iteration '18' and session '25'
Oct 3, 2019 @ 14:26:53.226 INFO 2019-10-03T12:26:53.226Z Collect in production
Oct 3, 2019 @ 14:26:53.234 TRACE 2019-10-03T12:26:53.234Z Iteration '19' and session '32'
Oct 3, 2019 @ 14:26:53.240 DEBUG 2019-10-03T12:26:53.240Z Collect in development
Oct 3, 2019 @ 14:26:53.325 TRACE 2019-10-03T12:26:53.325Z Iteration '20' and session '41'
Oct 3, 2019 @ 14:26:53.340 WARN 2019-10-03T12:26:53.340Z Investigate tomorrow
03 AM
06 AM
09 AM
12 PM
03 PM
06 PM
09 PM
Fri 04
No additional entries found
Load again

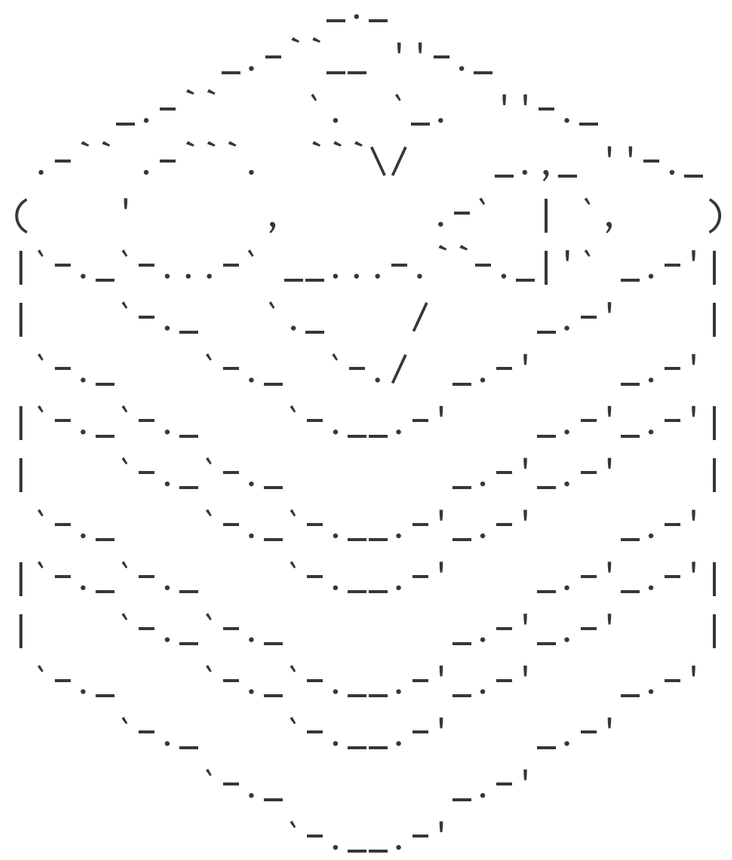
Ingest Pipeline

```
{
  "description" : "Parse Java log lines",
  "processors": [
    {
      "grok": {
        "field": "message",
        "patterns": [ "^\\[%{TIMESTAMP_ISO8601:timestamp}\\]%{SPACE}%{LOGLEVEL:log.level}
            %{SPACE}%{USERNAME:log.package}%{SPACE}\\[%{WORD:log.method}\\]%{SPACE}-
            %{SPACE}%{GREEDYDATA:labels}%{SPACE}-%{SPACE}%{GREEDYDATA:message_parsed}
            (?:\\n+(?<stacktrace>(?:.|\\r|\\n)+))?" ],
        "ignore_failure": true
      }
    }
  ]
}
```

Unknown Fields

?	log.labels	⚠ session=69, loop=20
t	log.level	WARN
?	log.method	⚠ main
#	log.offset	19,744
?	log.package	⚠ net.xeraa.logging.LogMe
t	message	[2019-05-21 05:02:07.458] WARN net.xeraa.logging.LogMe [main] - session=69, loop=20 - Investigate tomorrow
?	message_parsed	⚠ Investigate tomorrow
t	stream	stdout
🕘	suricata.eve.timestamp	May 21, 2019 @ 07:02:07.459
?	timestamp	⚠ 2019-05-21 05:02:07.458

ASCII Art



```
Redis 4.0.9 (00000000/0) 64 bit

Running in stand alone mode
Port: 6379
PID: 55757


      http://redis.io
```

Configuration Templates

```
filebeat.autodiscover:
  providers:
    - type: docker
      templates:
        - condition:
            equals:
              docker.container.image: redis
          config:
            - type: docker
              containers.ids:
                - "${data.docker.container.id}"
              exclude_lines: ["^\\s+[\\-`('.|_]"]
```

Who Logs the Logger

Avoid loops

Process without `-e`

filebeat.yml: `logging.to_files: true`

Pro: Hot 

Con: Complexity

Orchestrate 

kubernetes

elastic

@xeraa

Where to put Filebeat?

DaemonSet

https://github.com/elastic/beats/tree/master/deploy/kubernetes

ConfigMap

```
filebeat.autodiscover:
  providers:
    - type: kubernetes
      host: ${NODE_NAME}
      hints.enabled: true
      hints.default_config:
        type: container
        paths:
          - /var/log/containers/*${data.kubernetes.container.id}.log
```

Metadata

Either in cluster or outside

```
processors:

- add_kubernetes_metadata:
    in_cluster: true

- add_kubernetes_metadata:
    in_cluster: false
    host: <hostname>
    kube_config: ${HOME}/.kube/config
```

```
{
  "host": "172.17.0.21",
  "port": 9090,
  "kubernetes": {
    "container": {
      "id": "382184ecdb385cfd5d1f1a65f78911054c8511ae009635300ac28b4fc357ce51",
      "image": "my-java:1.0.0",
      "name": "my-java"
    },
    "labels": {
      "app": "java",
    },
    "namespace": "default",
    "node": {
      "name": "minikube"
    },
    "pod": {
      "name": "java-2657348378-k1pnh"
    }
  },
}
```

More Metadata

Add: Cloud, local timezone, process

Drop: Events, fields

Rename: Fields

Dissect, DNS reverse lookup

Configuration Templates

```
filebeat.autodiscover:
  providers:
    - type: kubernetes
      templates:
        - condition:
            equals:
              kubernetes.namespace: redis
          config:
            - type: docker
              containers.ids:
                - "${data.kubernetes.container.id}"
              exclude_lines: ["^\\s+[\\-`('.|_]"]
```

Customize Indices

```
output.elasticsearch:
  index: "%{[kubernetes.namespace]:filebeat}-%{[beat.version]}-%{+yyyy.MM.dd}"
```

Pro: Hot 💩💩💩

Con: Complexity++

Conclusion

Examples

https://github.com/xeraa/java-logging

Parse 🔪

Send ✉️

Structure 🕸

Containerize 📦

Orchestrate 🎻

Questions?

Philipp Krenn @xeraa