Module

Logging Fundamentals

elastic

Topics

- What are Logs?
- Getting Started with Filebeat and Logs
- Kibana Visualizations

Logging Fundamentals

Lesson 1

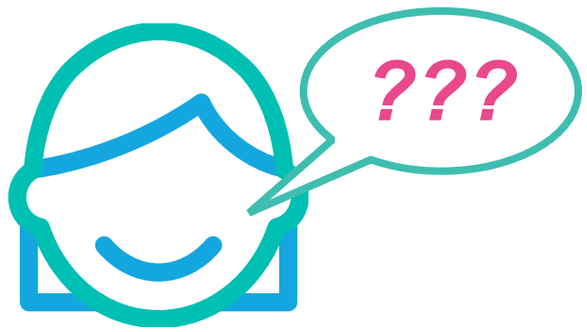
What are Logs?

elastic

Business Questions

How many users visited our new training landing page?

Why is our Javascript application slow?



When should we schedule download service maintenance?

How many webinar signups did we get from Europe?

It's all in the logs!

- Jordan Sissel

What is a log?

- logs are **records of activities**
  - by a system
  - by an application
  - by a device
  - by a human
  - …
- timestamp + data

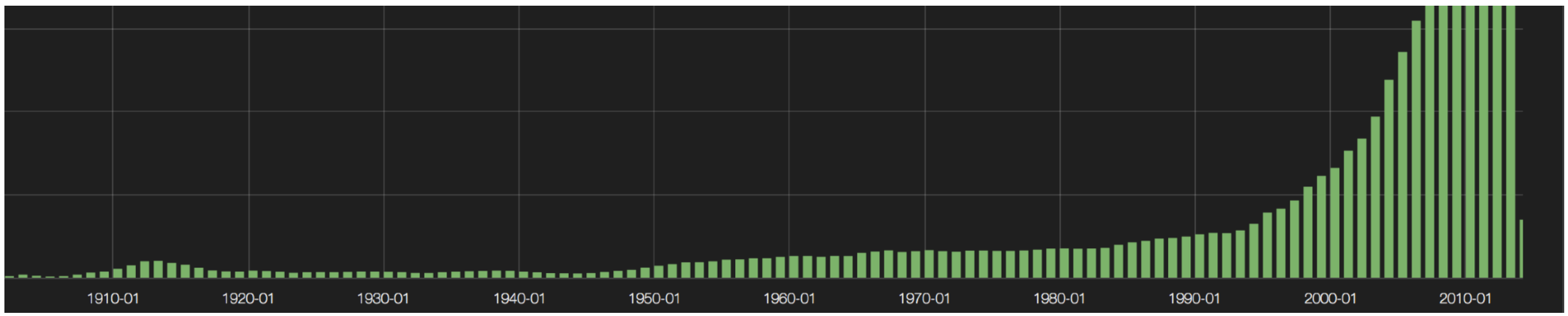
What are Logs?

- Application Logs

```
66.249.73.185 - - [16/Feb/2014:09:47:54 -0500] "GET / HTTP/1.1" 200 37932 "-"
"Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
```

```
[2017-05-18 00:00:05,871][INFO ][cluster.metadata          ] [esndata-2] [.data-
es-1-2017.05.18] creating index, cause [auto(bulk api)], templates [.data-
es-1], shards [1]/[1], mappings [_default_, shards, node, index_stats,
index_recovery, cluster_state, cluster_stats, node_stats, indices_stats]
```

```
120707  0:37:09 [Note] Plugin 'FEDERATED' is disabled.
120707  0:37:09 InnoDB: The InnoDB memory heap is disabled
```



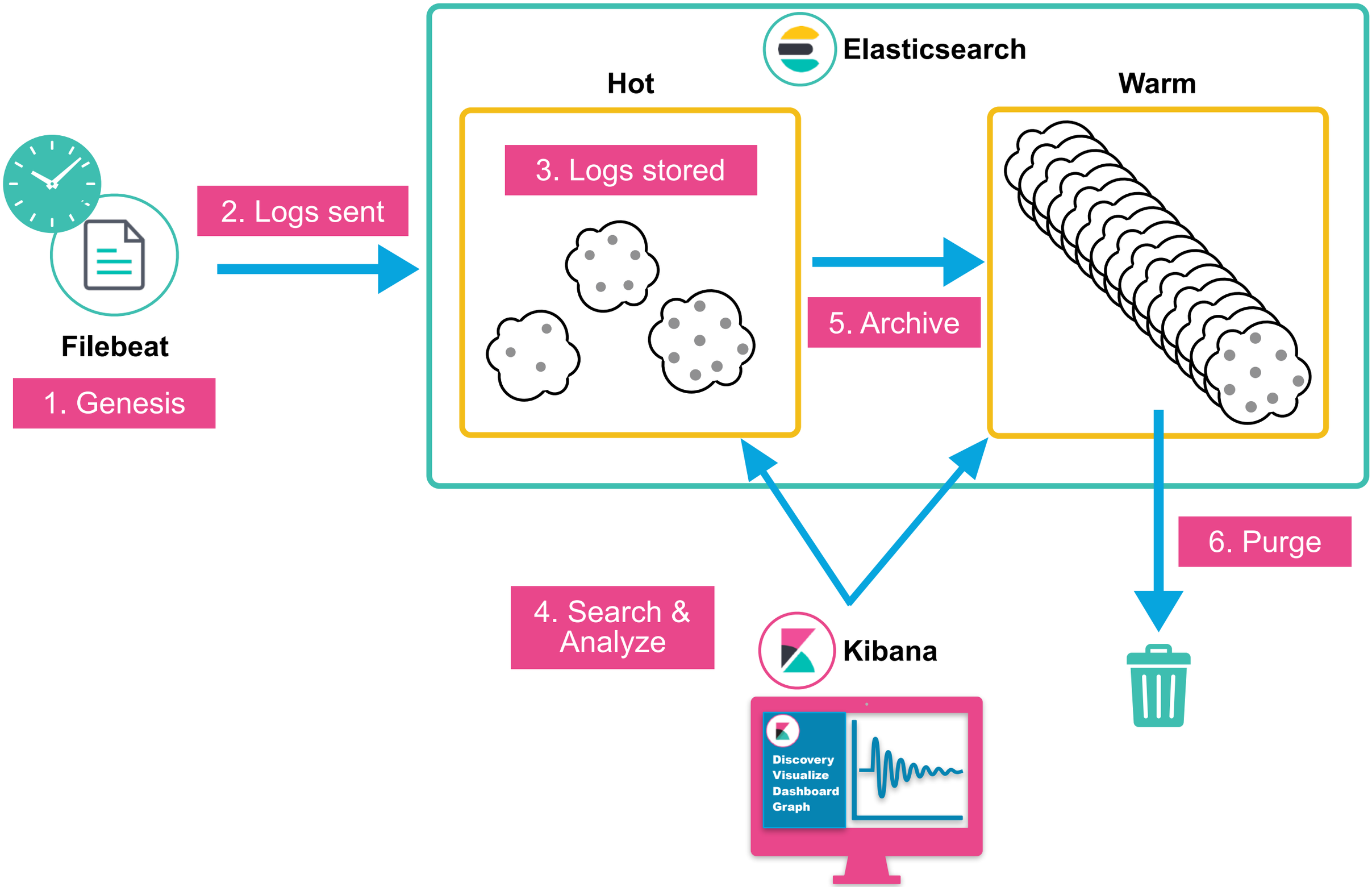
A Note About Timestamp

- Timestamps can be difficult to deal with
  - many different formats
  - time zones can be challenging
- Elasticsearch prefers ***ISO 8601*** format
  - but can ingest multiple time formats, if configured
- The ***ISO 8601*** format was designed to be unambiguous
  - it is a good practice to represent timestamps in ***ISO 8601***
- UI applications (e.g. Kibana) can adjust the displayed time
  - combine the stored time with the user's local time
- Example: **"2018-10-05T14:30:00Z"**

Common Problems

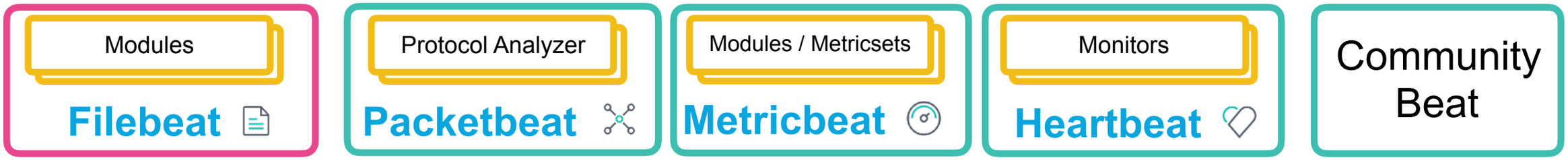
- Consistency
  - every application and device logs in its own format
- Time Formats
  - "Oct 11 20:21:47", "020805 13:51:24"
- Decentralized
  - logs are spread across all of your servers
  - SSH + grep aren't scalable
- Experts Required
  - limited access to log files on servers
  - limited knowledge of the log format

Logs Lifecycle



Beats

- Multiple flavors



- Beats can resolve different issues
  - reading files
  - retrieving metrics
  - retrieving network data
  - testing services availability

Filebeat

- Filebeat Modules
  - simplify the collection, parsing, and visualization of common log formats

Apache

Nginx

MongoDB

MySQL

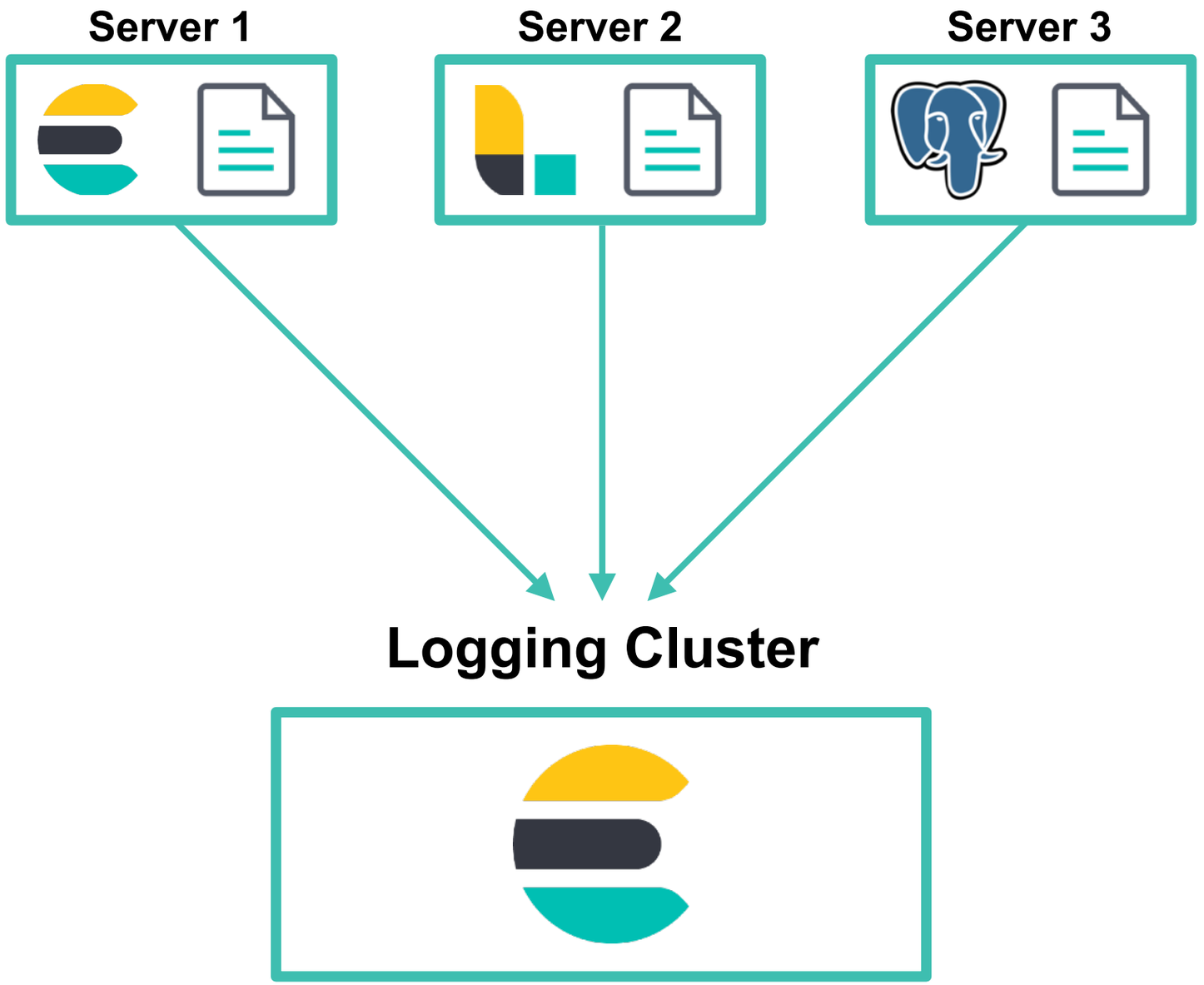
PostgreSQL

Redis

Add your own

Filebeat

- Open source data shippers
- Lightweight agent collecting logs



Lesson 1

Review - What are Logs?

elastic

Summary

- Logs can give us the answers to many questions which we ask of our data
- A log consists of a message with both a timestamp and some piece(s) of data
- Filebeat monitors log directories or specific log files
- Filebeat Modules simplify the collection, parsing, and visualization of common log formats

Quiz

1. What are the two elements of a log message?
2. **True or False**: It's OK to give Scott from Marketing ssh access to the web server so he can see how many people signed up for his latest webinar.
3. **True or False**: Filebeat Modules simplify the collection, parsing, and visualization of common log formats.

Logging Fundamentals

Lesson 1

Lab - What are Logs?

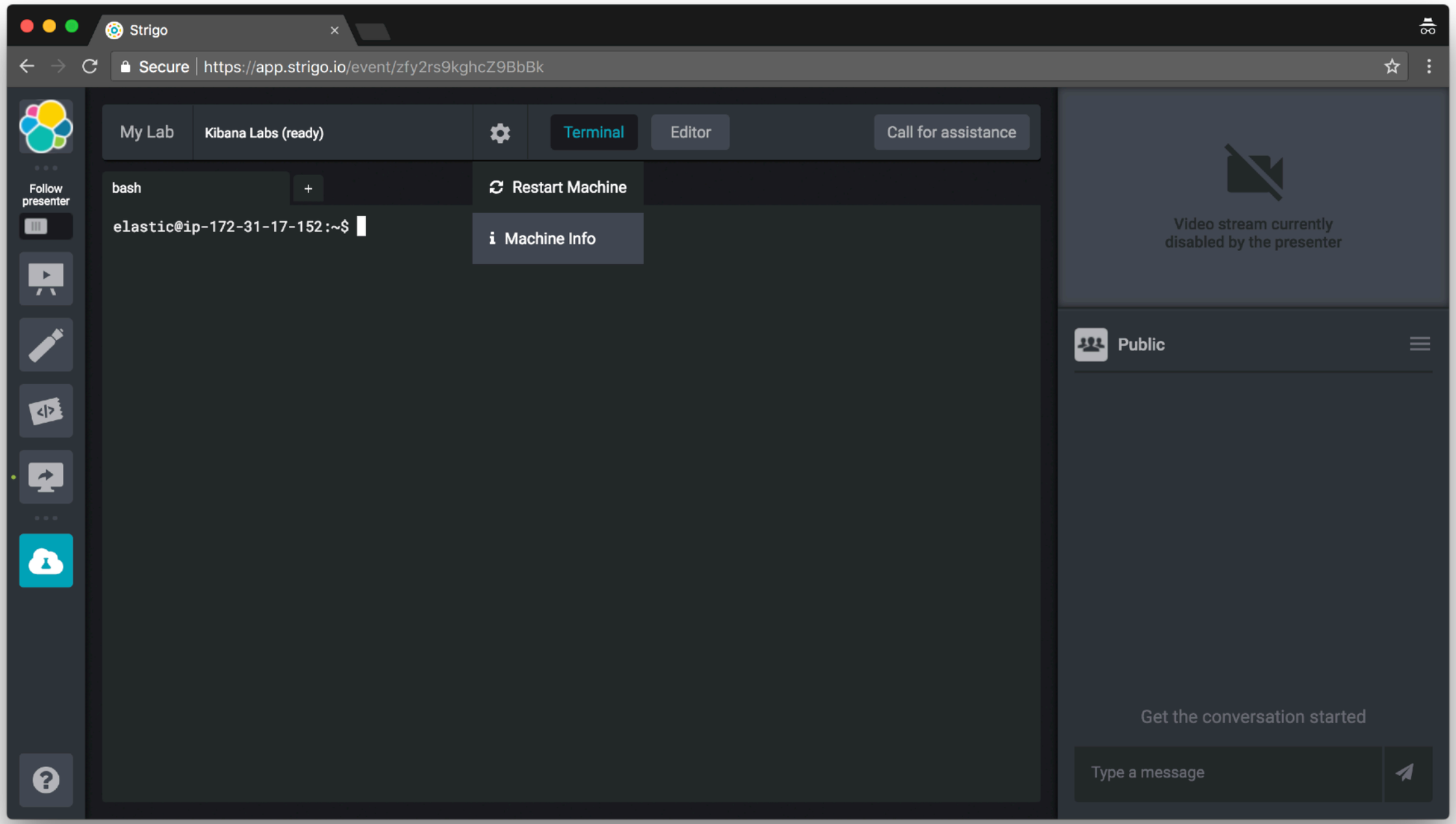
elastic

Lab Environment

- Visit Strigo using the link that was shared with you, and log in if you haven't already done so
- Click on "**My Lab**" on the left

Lab Environment

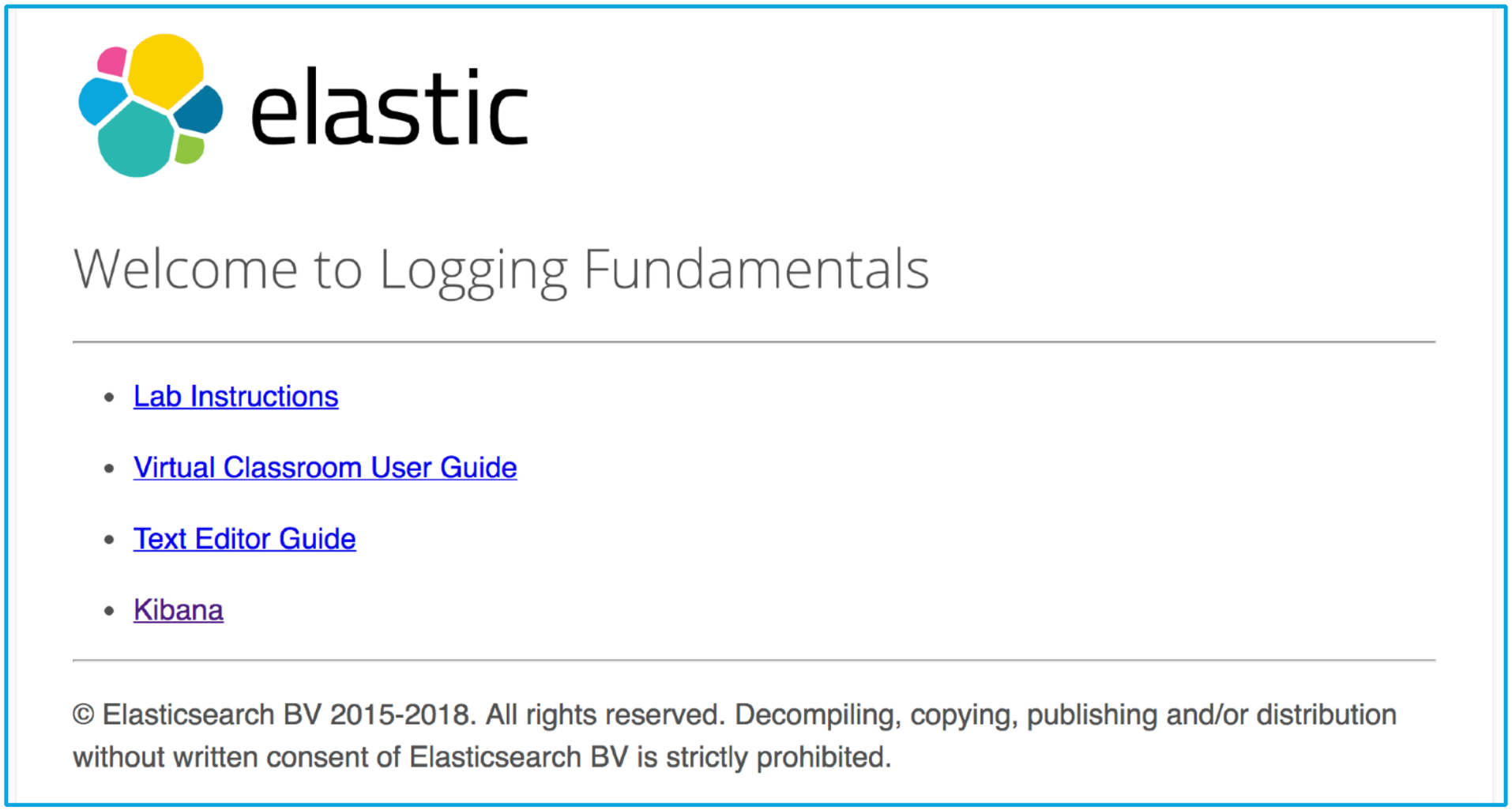
- Click on the gear icon next to "**My Lab**" and select "**Machine Info**"



elastic

Lab Environment

- From here you can access lab instructions and guides
  - You also have them in your .zip file, but it is easier to access and use the lab instructions from here:



Logging Fundamentals

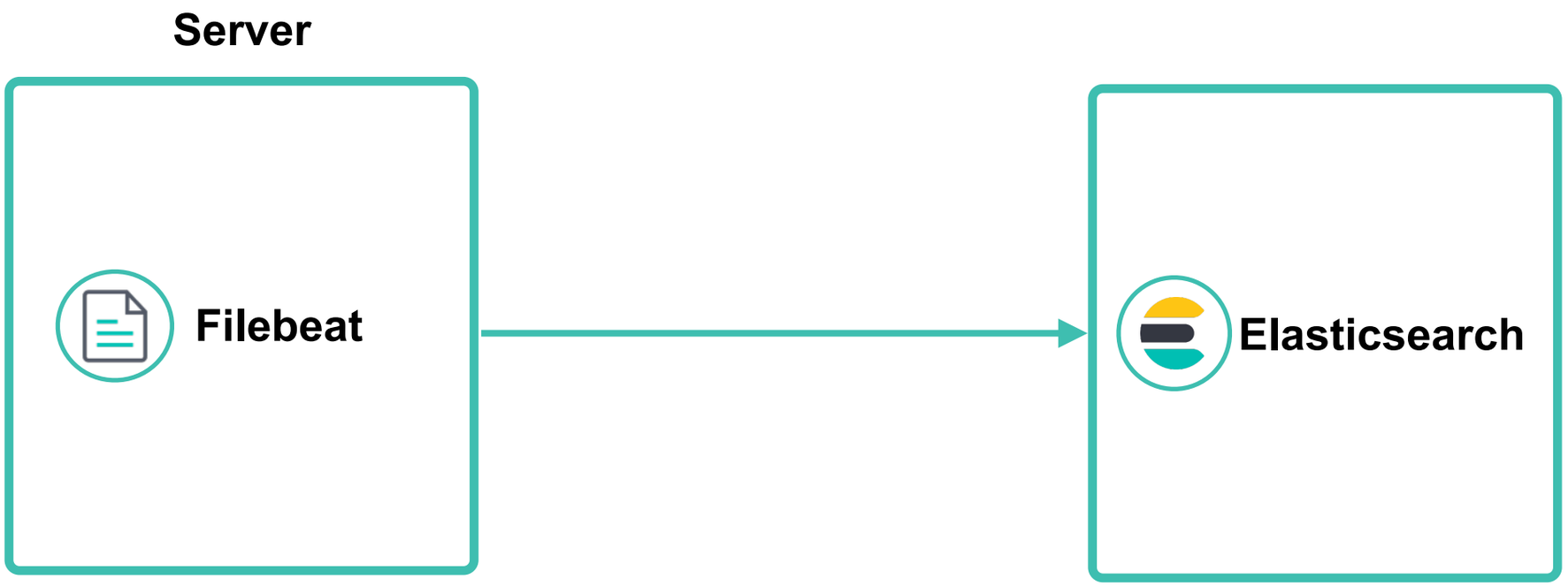
Lesson 2

Getting Started with Filebeat and Logs

elastic

Shipping files from a host to Elasticsearch

- To ship log files from a host you will need to:
  1. download Filebeat on the host where the logs reside
  2. configure Filebeat
  3. start Filebeat
  4. verify that the data has arrived in Elasticsearch



Downloading Filebeat

- Binaries for each environment

Version:	7.3.1	
Release date:	October 01, 2019	
License:	**Elastic License**	
Downloads:	**DEB 32-BIT** sha asc	**DEB 64-BIT** sha asc
	RPM 32-BIT sha asc	**RPM 64-BIT** sha asc
	LINUX 32-BIT sha asc	**LINUX 64-BIT** sha asc
	MAC sha asc	**WINDOWS 32-BIT** sha asc
	WINDOWS 64-BIT sha asc	
Package Managers:	Install with **yum**	
	Install with **apt-get**	
	Install with **homebrew**	
Containers:	Run with **Docker**	Run with **Kubernetes**

- https://www.elastic.co/downloads/beats/filebeat

- APT and YUM repositories are also available

Installing Filebeat

- Mac

```
tar xf filebeat-7.3.1-darwin-x86_64.tar.gz
```

- Linux

```
tar xf filebeat-7.3.1-linux-x86_64.tar.gz
```

- DEB

```
sudo dpkg -i filebeat-7.3.1-amd64.deb
```

- RPM

```
sudo rpm -vi filebeat-7.3.1-x86_64.rpm
```

- Docker

```
docker pull docker.elastic.co/beats/filebeat:7.3.1
```

- Windows

```
PS > cd 'C:\Program Files\filebeat'
PS C:\Program Files\filebeat> .\install-service-filebeat.ps1
```

Configuring Filebeat

- In order to send data, the Filebeat output needs to be configured:
  - multiple outputs available (Elastic cloud, Elasticsearch, Kafka, Logstash, ...)
  - data can be securely sent to Elasticsearch using SSL
- The output can be configured in the ***filebeat.yml*** configuration file:

```
output.elasticsearch:
   hosts: ["localhost:9200"]
```

Starting Filebeat

- Mac

```
sudo chown root filebeat.yml
sudo ./filebeat -e -c filebeat.yml -d "publish"
```

- Linux

```
sudo ./filebeat -e -c filebeat.yml -d "publish"
```

- DEB

```
sudo service filebeat start
```

- RPM

```
sudo service filebeat start
```

- Windows

```
PS C:\Program Files\filebeat> Start-Service filebeat
```

- Docker
  - https://www.elastic.co/guide/en/beats/filebeat/current/running-on-docker.html

Filebeat Indices

- Once the data starts flowing from Filebeat to Elasticsearch, it is possible to look at the data
- By default, all Beats group data by day using the format:
  - {type}beat-{version}-{yyyy-MM-dd}-XXXXXX

Console | Search Profiler | Grok Debugger

```
GET _cat/indices
```

```
green  open .kibana_task_manager          CsizfSNJQ5yrqDcB9oo9Zw 1 0   2  0  45.5kb  45.5kb
yellow open my_index                      U-NmKJE4StKEhr3-l_6uug 1 1   1  0   5.2kb   5.2kb
yellow open filebeat-7.3.1-2019.10.18-000001 WuWTP_tYRVq-ABax7xukog 1 1 381  0 343.9kb 343.9kb
green  open .kibana_1                     dNCPlOgfQK6-f3onehSESA 1 0 705 15 422.9kb 422.9kb
```

Data in Elasticsearch

- To see what the data looks like, it is possible to send a query to Elasticsearch

```
GET filebeat*/_search
```


```
{
  "took" : 0,
  "timed_out" : false,
  "_shards" : {
    "total" : 1,
    "successful" : 1,
    "skipped" : 0,
    "failed" : 0
  },
  "hits" : {
    "total" : {
      "value" : 377,
      "relation" : "eq"
    },
    "max_score" : 1.0,
    "hits" : [
      {
        "_index" :
"filebeat-7.3.1-2019.10.18-000001",
        "_type" : "_doc",
        "_id" :
"mnVV320BHYYYifCddW4a",
        "_score" : 1.0,
        "_source" : {
          "agent" : {
            "hostname" : ...
```

Stopping Filebeat

- Terminal

```
CTRL + C
```

- DEB/RPM

```
sudo service filebeat stop
```

- Windows

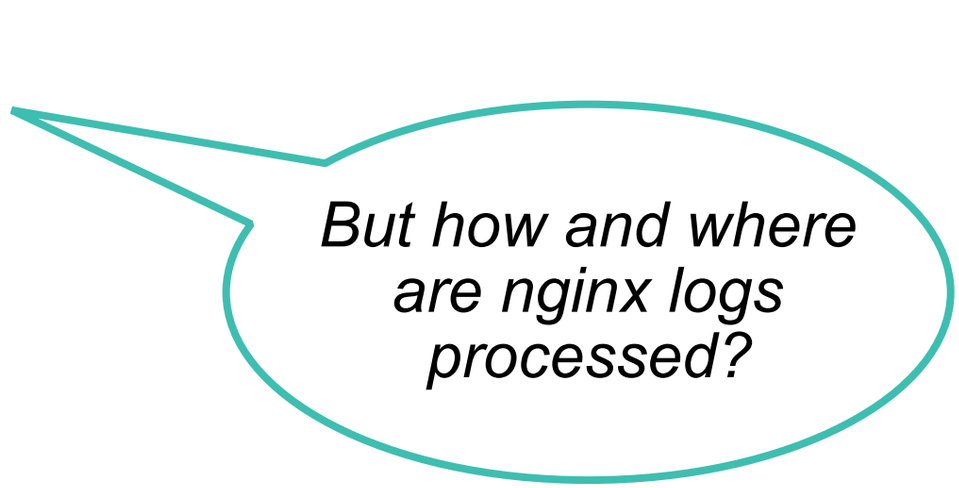
```
PS C:\Program Files\filebeat > Stop-Service filebeat
```

Filebeat Modules

- Simplify the collection, parsing, and visualization of common log formats
- One command:

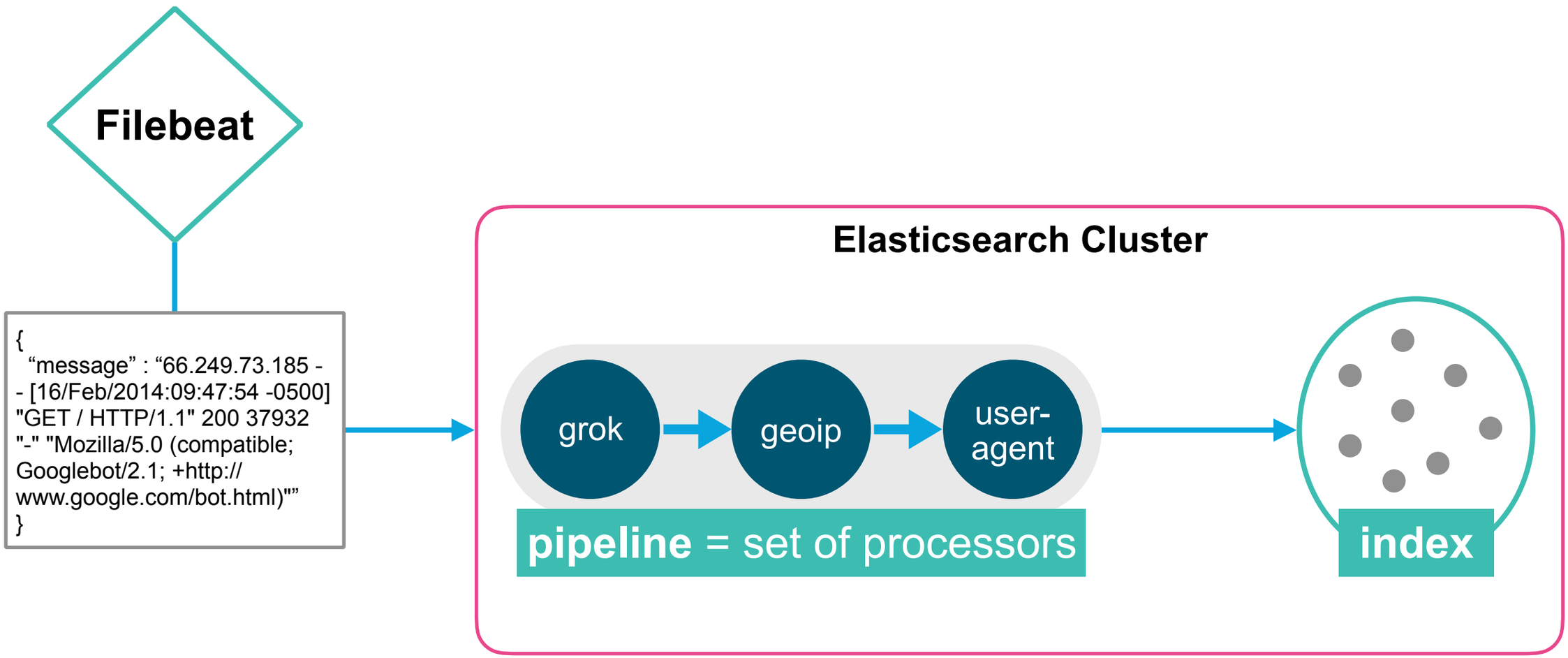
```
./filebeat modules enable nginx
```

  - sets up data collection
  - sets up data processing



Ingest Pipelines

- ***Ingest pipelines*** provide the ability to pre-process a document just before it gets indexed
  - parses, applies transformations, and enriches the data
  - pipelines allow you to configure the processors that will be used



The geoip Processor

- Enriches your documents by looking up geographical location data for an IP address

```
PUT _ingest/pipeline/my_pipeline
{
  "processors" : [
    { "geoip" : { "field" : "ip" } }
  ]
}

PUT my_index/_doc/1?pipeline=my_pipeline
{
  "ip": "8.8.8.8"
}

GET my_index/_doc/1
```

```
{
  "_source": {
    "ip": "8.8.8.8",
    "geoip": {
      "continent_name": "North America",
      "country_iso_code": "US",
      "location": {
        "lat": 37.751,
        "lon": -97.822
      }
} } }
```

The geoip processor is not installed by default. You will see how to install it in the labs.

Logging Fundamentals

Lesson 2

Review - Getting Started with Filebeat and Logs

elastic

Summary

- There are different Filebeat binaries for different distributions
- Filebeat sends the data to a configured output
- Once the data is sent to Elasticsearch, it is possible to query Elasticsearch to explore the data

Quiz

1. **True or False:** Filebeat is sending data to Kibana
2. What is the default output for Filebeat?
3. **True or False:** The default index created by Filebeat is filebeat-{version}-{yyyy-MM-dd}

Logging Fundamentals

Lesson 2

Lab - Getting Started with Filebeat and Logs

elastic

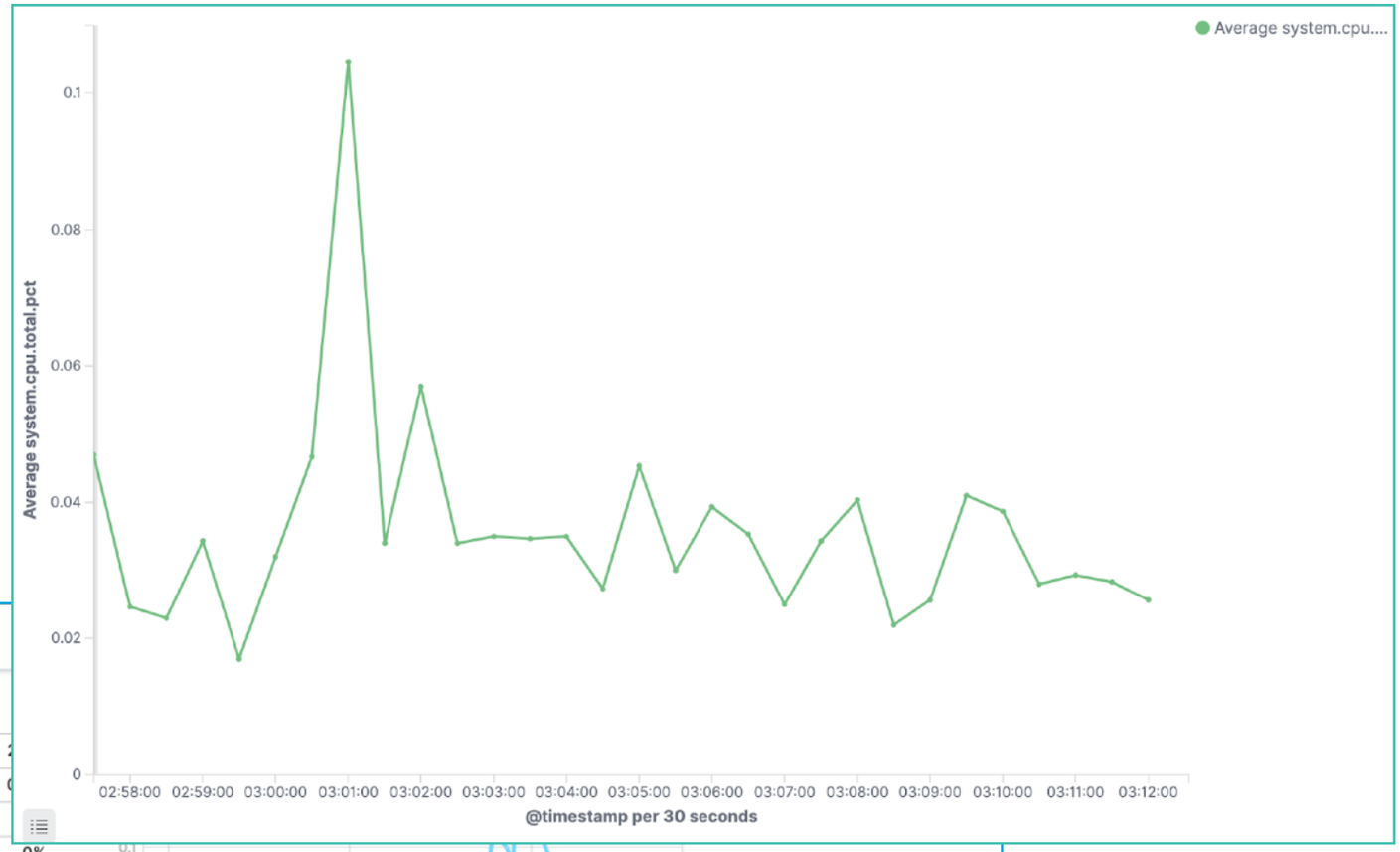
Lesson 3

Kibana Visualizations

elastic

Kibana

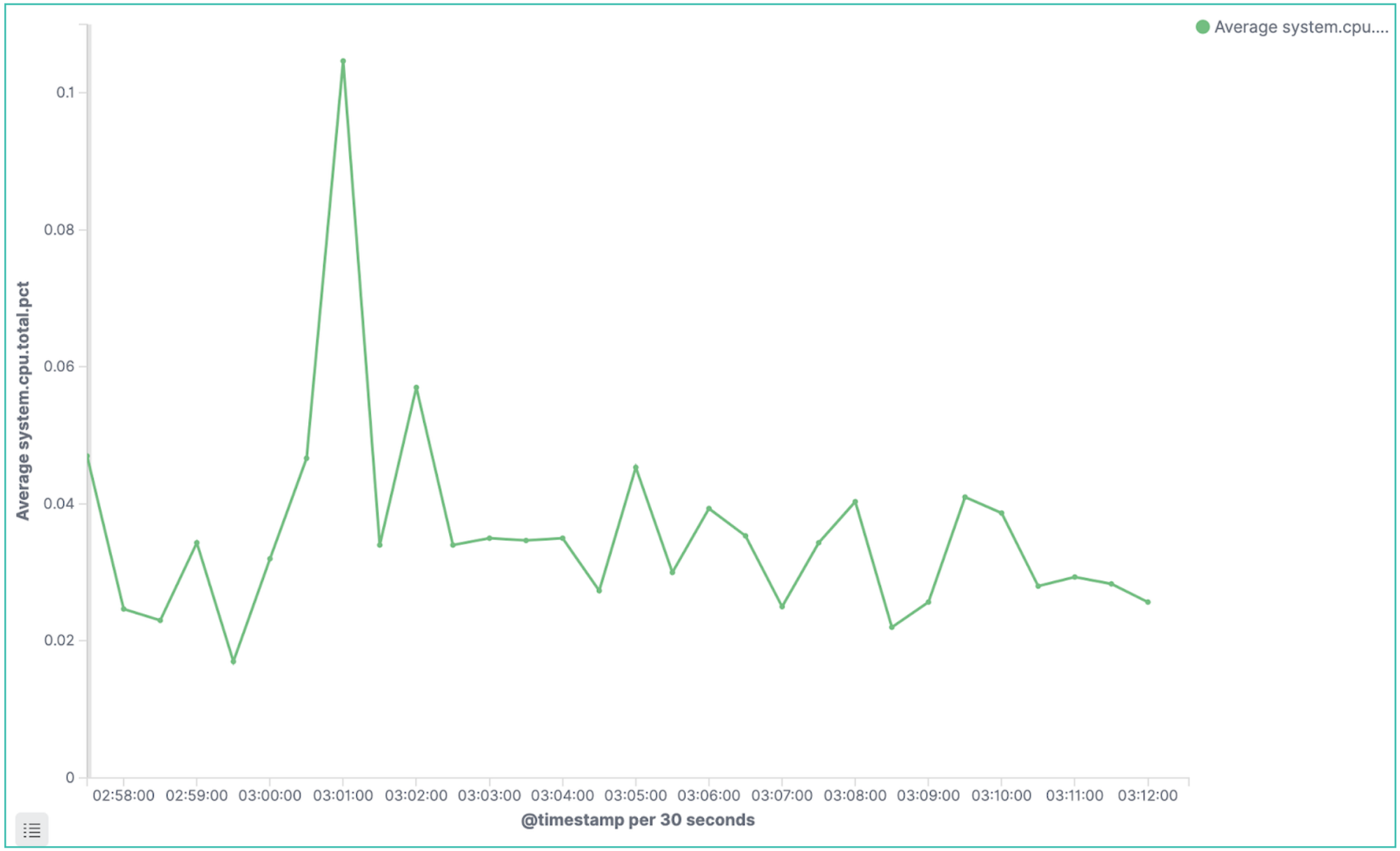
- Data Visualization
- Elastic Stack UI
- Stack Management



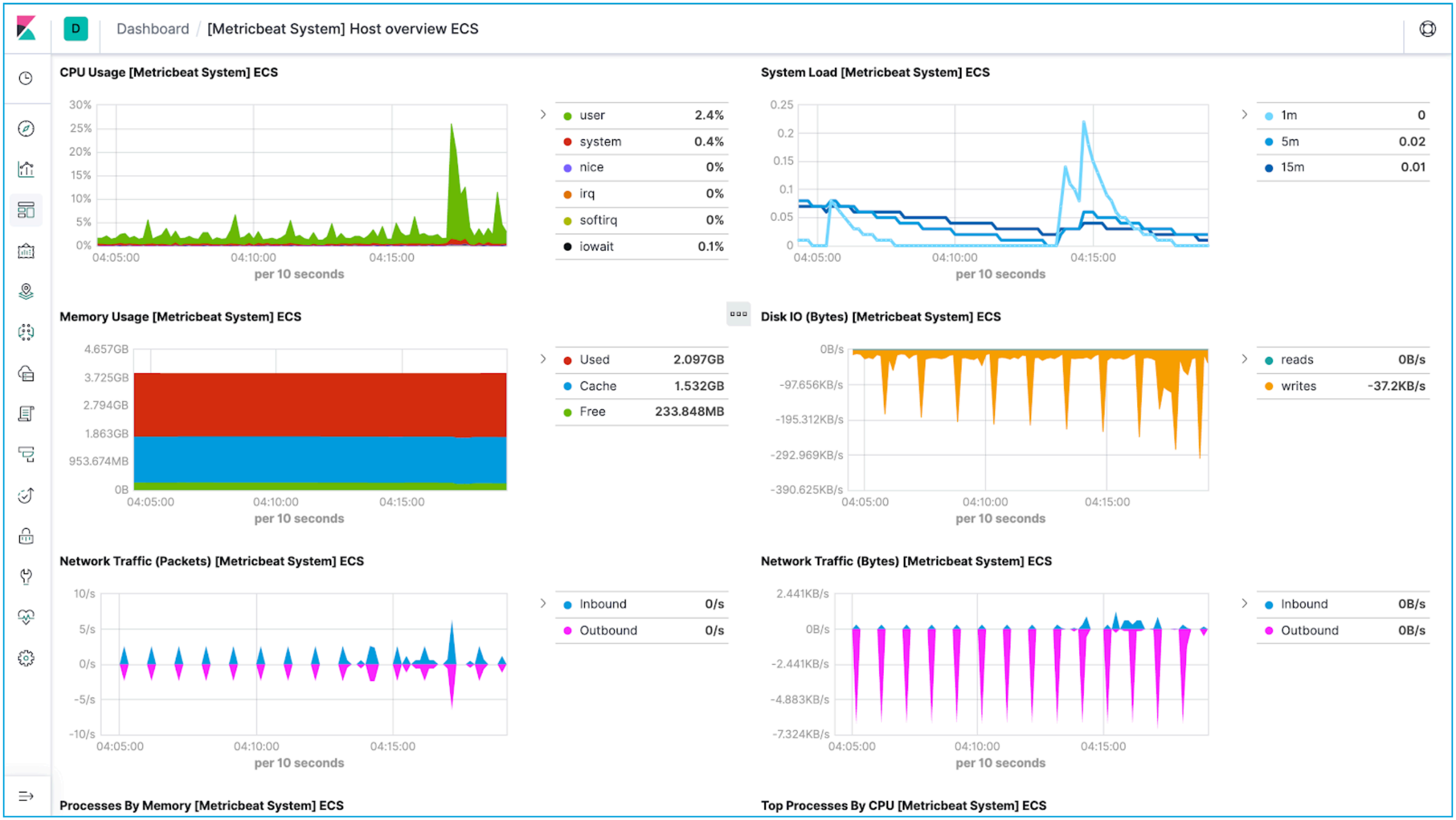
elastic

Discover

Visualize

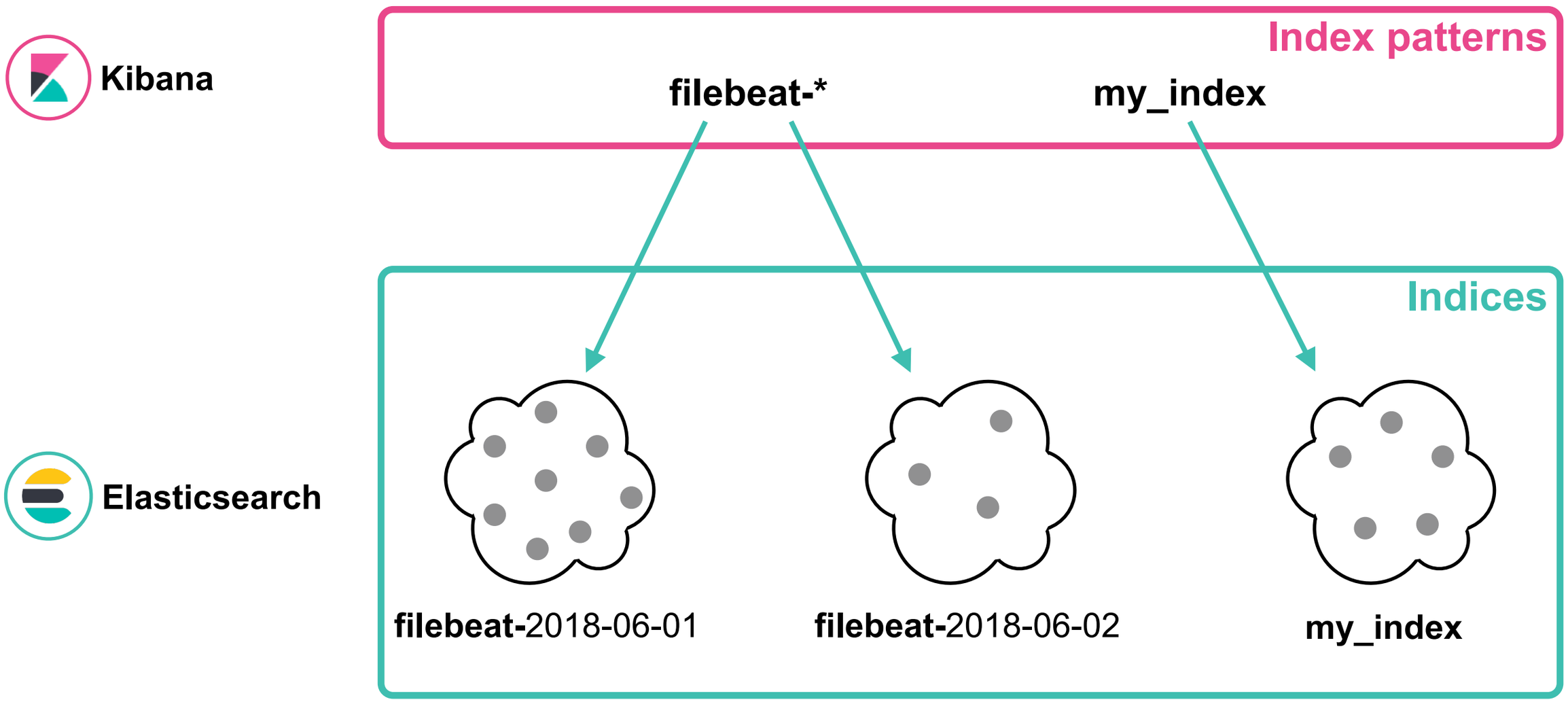


Dashboards

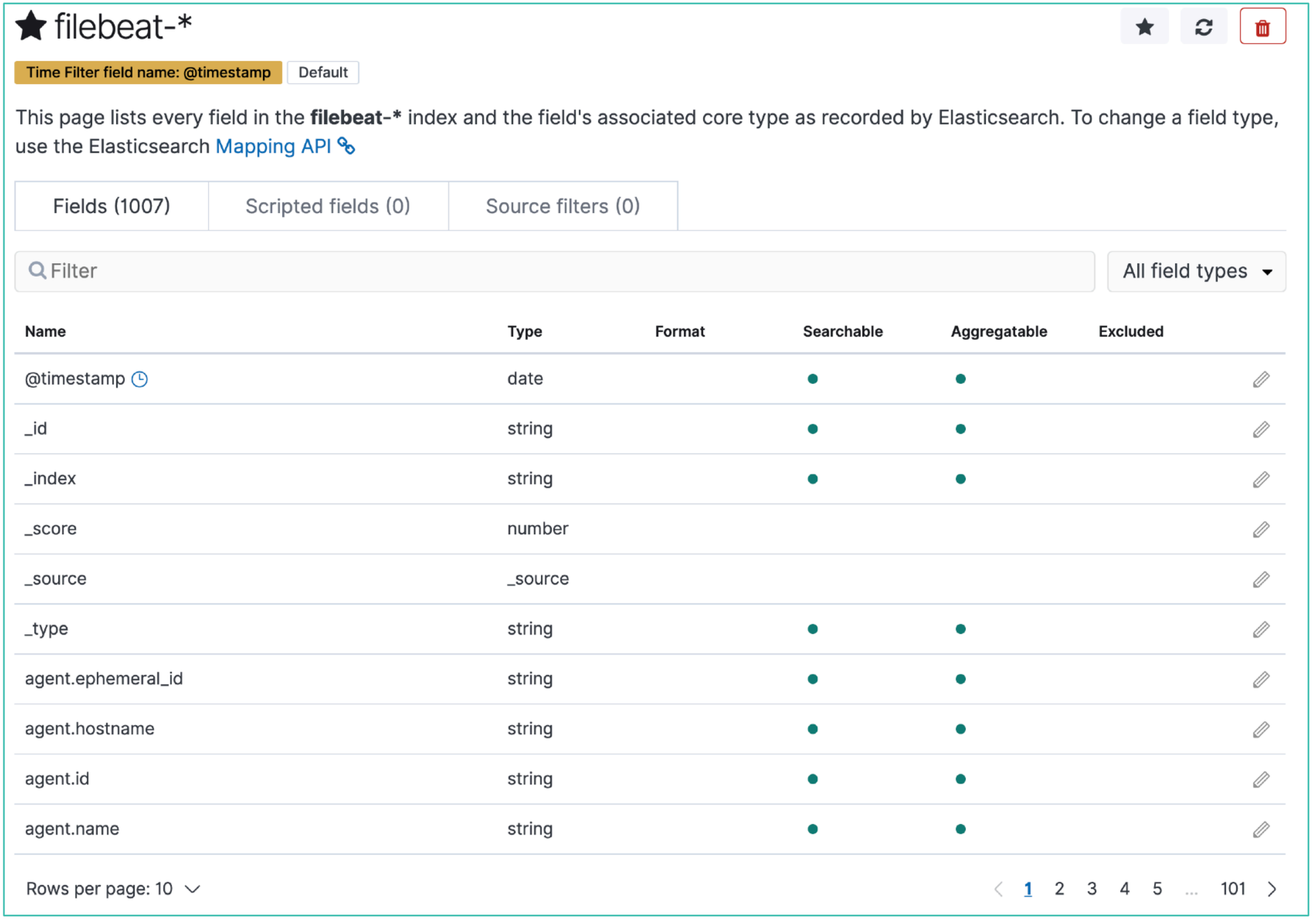


Indices and Index Pattern

- Visualizations are nice, but **which dataset** is being used?
  - when you search or analyze data, you do that on top of a dataset



Index Patterns



Beats and Visualization

- Every Beats comes with a set of pre-build visualizations
- Kibana can help visualizing different kinds of metrics base on the service or the system monitored
- There are two main ways to load Beats' visualizations
  - setup command (easiest way)
  - configuration file (not covered in this training)

Loading the Visualizations

- Make sure Elasticsearch is running
- Make sure Kibana is running
- Make sure Filebeat is properly configured:
  - Elasticsearch output
  - Kibana dashboards
- Run the command:
  - DEB/RPM
  ```
  filebeat setup --dashboards
  ```
  - Mac
  ```
  ./filebeat setup --dashboards
  ```
  - Windows
  ```
  PS > .\filebeat.exe setup --dashboards
  ```
  - Docker
  ```
  docker run docker.elastic.co/beats/filebeat:7.3.1
  setup --dashboards
  ```

Pre-built Dashboards

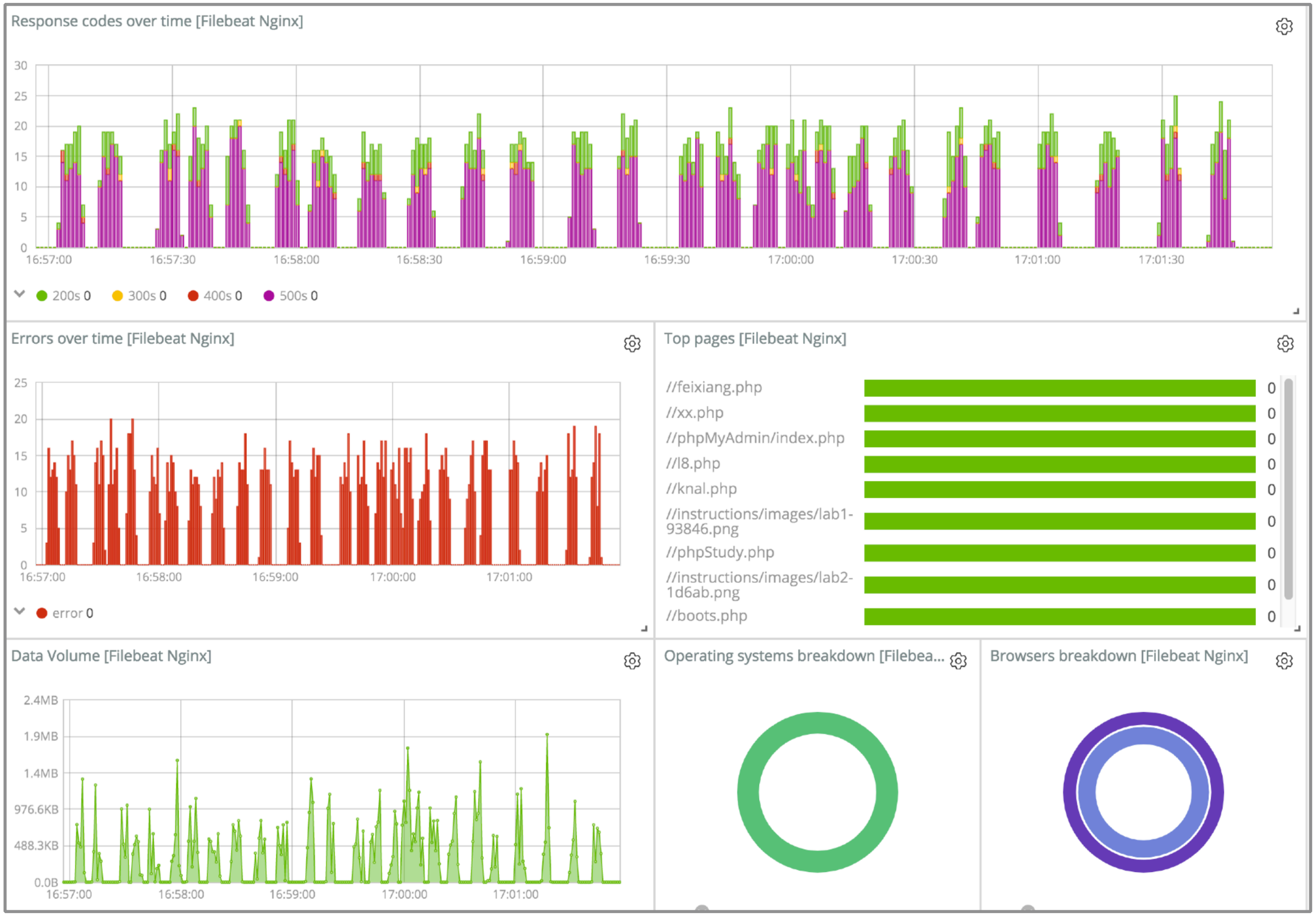
Dashboards

Create new dashboard

Search...

Title	Description	Actions
[Filebeat Apache2] Access and error logs	Filebeat Apache2 module dashboard	
[Filebeat Auditd] Audit Events	Dashboard for the Auditd Filebeat module	
[Filebeat Kafka] Overview	Filebeat Kafka module dashboard	
[Filebeat System] Sudo commands	Sudo commands dashboard from the Filebeat System module	
[Filebeat PostgreSQL] Overview	Overview dashboard for the Filebeat PostgreSQL module	
[Filebeat PostgreSQL] Query Duration Overview	Dashboard for analyzing the query durations of the Filebeat PostgreSQL module	
[Filebeat Icinga] Debug Log	Filebeat Icinga module dashboard for the debug logs	
Overview [Filebeat MongoDB]	Filebeat MongoDB module overview	
[Filebeat Nginx] Overview	Dashboard for the Filebeat Nginx module	
[Filebeat MySQL] Overview	Overview dashboard for the Filebeat MySQL module	
[Filebeat System] New users and groups	New users and groups dashboard for the System module in Filebeat	

NGINX Dashboards



Lesson 3

Review - Kibana Visualizations

elastic

Summary

- Kibana is a tool that allows you to visualize the data stored in your Elasticsearch cluster
- Kibana is the **user interface** of the Elastic Stack
- Filebeat can load pre-built Kibana dashboards using a single command

Quiz

1. **True or False:** Kibana can be used to store data.
2. What is the easiest way to load Kibana dashboards to visualize logging data?
3. **True or False**: By default, Filebeat only loads the system dashboard.

Lesson 3

Lab - Kibana Visualizations

elastic

Lesson 4

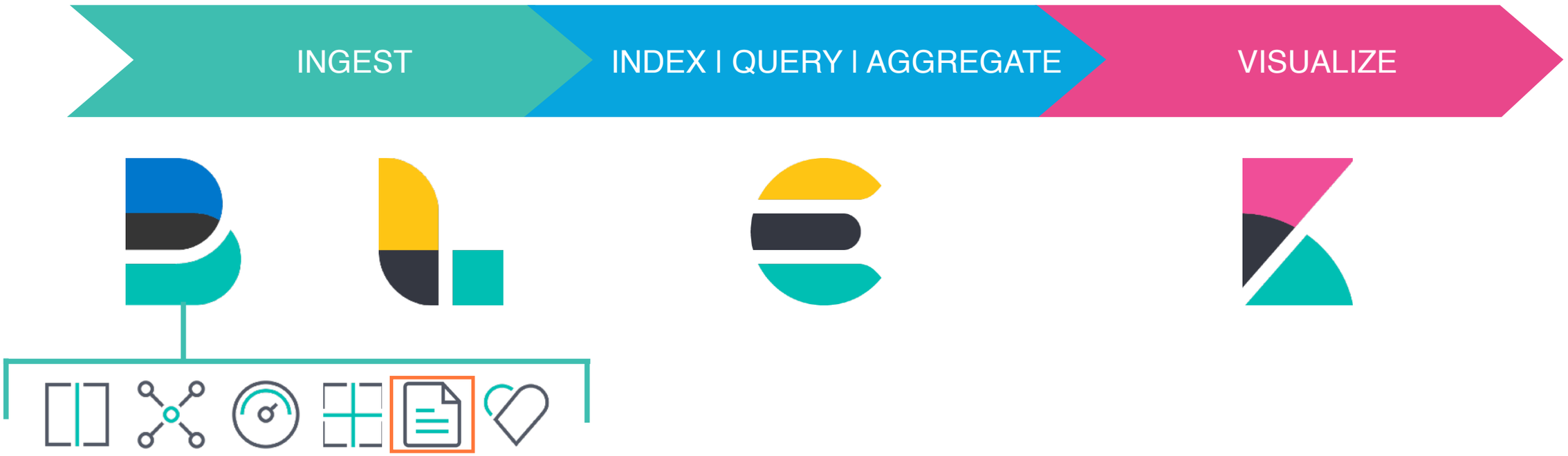
Learn More

elastic

Filebeat Fundamentals

- Monitoring different systems and services is key to any information system
- Filebeat allows you to monitor many different logs
- This is what you built in less than 2 hours:

The Elastic Stack



Next Steps

- This is an introductory course to the logging use case
- Next, use Filebeat to monitor your production system
- Then, enable more modules to understand if the different systems in your infrastructure are having issues
- References
  - https://www.elastic.co/products/beats/filebeat
  - https://www.elastic.co/guide/en/beats/filebeat/current/index.html
  - https://www.elastic.co/solutions/logging
- Finally, to have a deeper understanding of Logging there are other courses you can take...

Other Metrics Courses

- Shipping Log Data
  - Filebeat internal architecture
  - Filebeat modules, tagging data, and conditionals
  - Filebeat resiliency (new files, rollover, recover, etc.)
  - Filebeat multi-line processing
- Log Pipeline Workflows
  - ingest pipelines architecture
  - rename and remove processors
  - foreach and set processors
  - error handling

Other Metrics Courses

- Structuring Log Data
  - structuring unstructured data
  - combining text patterns with grok
  - advanced grok techniques
- Visualizing Log Data with Kibana
  - querying from discover
  - time series visual builder
  - tips and tricks for hunting through your logs

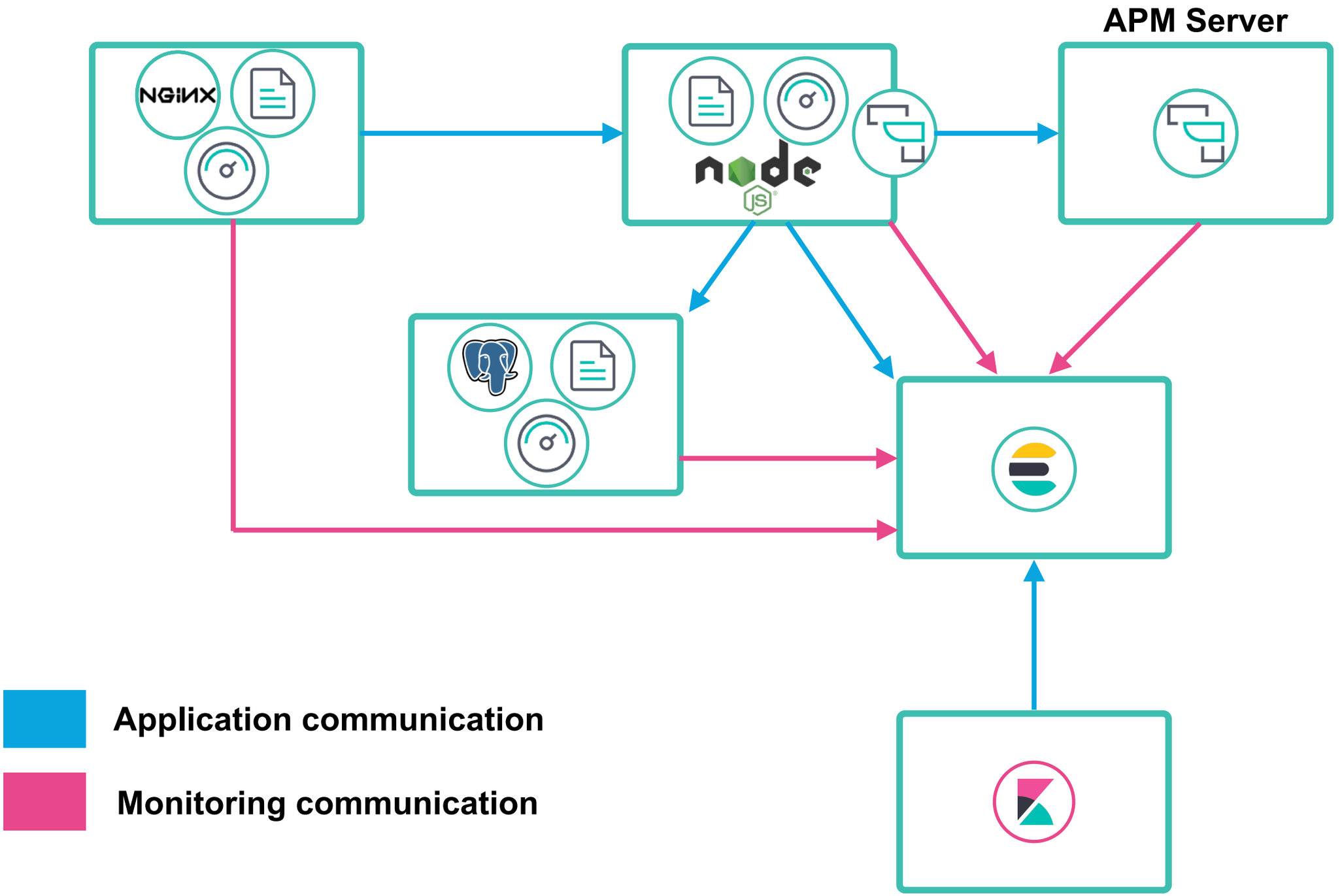
Other Logging Courses

- Monitoring Kubernetes and Docker Container Logs
  - why containers?
  - collect all logs and add metadata
  - filebeat autodiscovery
  - logs from inside a container

Logs are not Enough

- Logs are one aspect of services and systems monitoring
- But there are other aspects:
  - Metrics
  - APM
- To really understand what's happening on your application you should combine them all:
  - Metrics Fundamentals
  - APM Fundamentals

The Bigger Picture



Lesson 4

Review - Learn More

elastic

Summary

- Filebeat is only one tool of the many monitoring tools offered by the Elastic stack
- By combining *Metrics*, *Logging* and *APM* you can have a full understanding of your applications and systems
- There are a many features in Filebeat and ingest pipelines which are not covered in this course, but are paramount for a good logging solution

Quiz

1. **True or False:** Filebeat allows you to drop events based on a condition.
2. Which solutions are complementary to logging when it comes to monitoring an application?
3. What happens to Filebeat if the machine restarts?

Lesson 4

Lab - Learn More

elastic

Quiz Answers

elastic

What are Logs

1. False
2. Metric, Packet, Winlog, File, Audit
3. True

Getting Started with Filebeat and Logs

1. True
2. Elasticsearch
3. True

Kibana Visualizations

1. False
2. Filebeat command line --dashboards
3. False

Learn More

1. True
2. APM and Metrics
3. If the process is setup to start automatically, it will resume from where it stopped