Filebeat

Metricbeat

Elastic Observability Workshop

Lab 2 - Capturing and Visualizing Metrics and Logs

Overview

- Download Metricbeat
- Setup and point to common elasticsearch cluster
- Change back to cloud.id/cloud.auth of your deployment (Lab 1)
- Download Filebeat
- Download nginx logs
- Setup nginx module for Filebeat
- Validate data in Kibana

Windows Install Instructions

Mac Install Instructions

Validate Data in Kibana (Common)

Introduction

In this lab guide we will walk you through how to ingest multiple logs files and metrics into the Elastic stack.

Instructions are provided for both Windows and Mac OS/Linux Operating Systems. Please follow the instructions for your operating system and then follow the steps on the last section **Validate Data in Kibana**.

Local Laptop Installation

Synopsis

Beats agents are data shippers that are designed to be lightweight. Each beat targets a specific type of data set. For the purposes of our lab we will use Metricbeat which will send important metrics like CPU and memory utilization into Elasticsearch. We will also use Filebeat which will not only send log files from services like NGINX and Apache, but also system authorization logs.

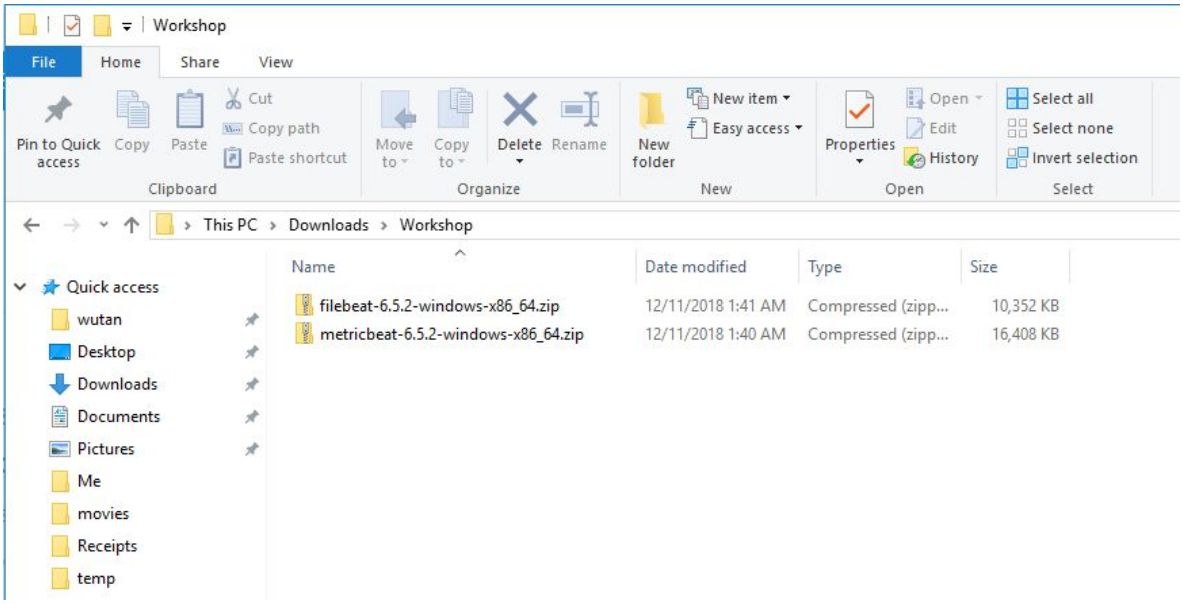
Software Download

Software	**URL**
Metricbeat	https://www.elastic.co/downloads/beats/metricbeat
Filebeat	https://www.elastic.co/downloads/beats/filebeat

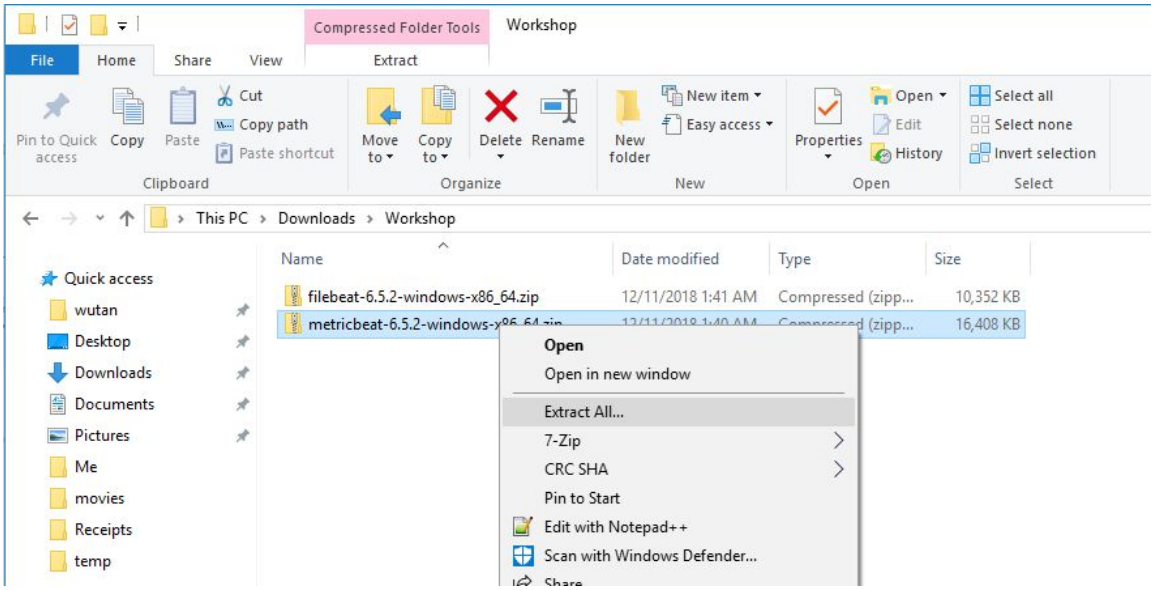
Windows Instructions

Metricbeat

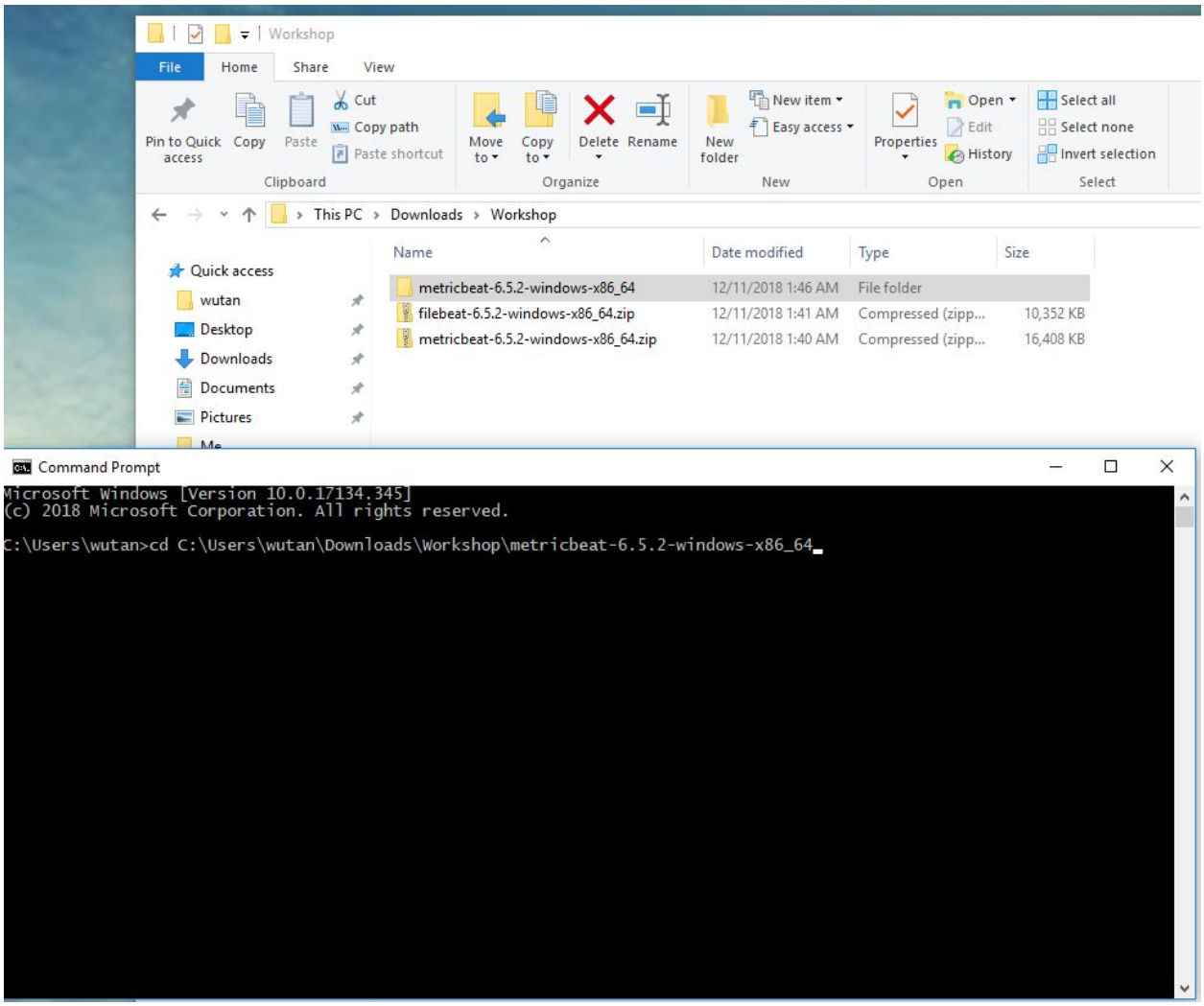
1) Open a Windows Explorer and navigate to the location that you downloaded metricbeat to. (Your version of downloaded beat will be different than the screenshot below based on the latest version available at the time of download)



2) Extract the metricbeat-... zip file that you downloaded



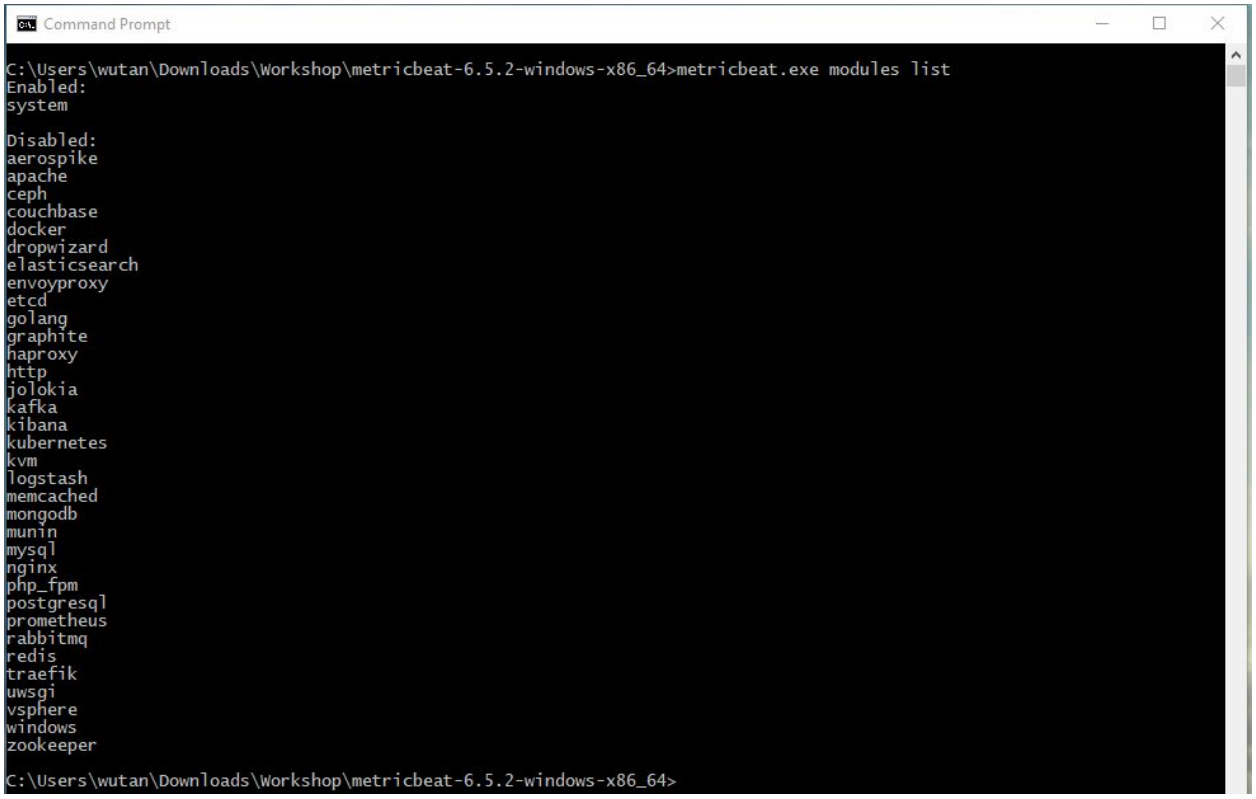
3) Open up a command prompt and type in `cd` (add a space after cd). Now drag-n-drop the extracted folder from step #2 to the command prompt window. Notice how it filled out the full path for you in the command prompt window? You can also type out the full path if you are a glutton for pain. Hit enter.



4) Now list modules that are available

```
.\metricbeat.exe modules list
```

You should see which modules are ***enabled*** and which modules are ***disabled.*** Out of the box the **system** module is the only one that is enabled.
(Elastic actively adds modules to the beats in new versions to make the ingestion easier, the list of modules in the version you downloaded might be different)



5) Before we setup Elasticsearch to accept system metrics we need to tell Metricbeat where Elasticsearch is and provide credentials to login. We do this by editing the configuration file for Metricbeat called metricbeat.yml.

YAML files don't like hard tabs. Do not use them if you are editing a .yml file because they will cause errors. To learn more about .yml files see this link: https://en.wikipedia.org/wiki/YAML

To start with, instead of sending metrics from your laptop to the cluster you have created in Lab 1, we are going to first send it to a shared Elasticsearch cluster created by the instructor. It will be used to demonstrate Kibana Infrastructure UI and also mimics a more realistic scenario - multiple endpoints sending metrics to single Elasticsearch Cluster.

Use your favorite text editor (e.g. Notepad++) to open metricbeat.yml and replace **cloud.id** and **cloud.auth** with the following values:

```
cloud.id:
"observability_workshop:dXMtd2VzdDEuZ2NwLmNsb3VkLmVzLmlvJDAzN2N
hZDU4M2ZiMTRkYTc4MGUzYTE2NjA3OGU0OTc5JDdiZDg5MWFhOGJhMDQ1NzFiMz
g4N2ZkMDAzYjE2MWYz"
cloud.auth: "elastic:ncyRLYhJAl8akCSZLYBnBlED"
```

Example:

```
#============================== Elastic Cloud ==================================

# These settings simplify using metricbeat with the Elastic Cloud (https://cloud.elastic.co/).

# The cloud.id setting overwrites the `output.elasticsearch.hosts` and
# `setup.kibana.host` options.
# You can find the `cloud.id` in the Elastic Cloud web UI.
cloud.id: "CentralizedBeatsMgmt:dXMtZWFzdC0xLmF3cy5mb3VuZC5pbyRkZmRiZTEwOWY2MmI0MTMxODhhZTRmM2U4ODYzNTV1ZiRjYTEzMTg0MGFkMzc0OWZjYThkZWRhZTU5OWE0MjY2OQ=="

# The cloud.auth setting overwrites the `output.elasticsearch.username` and
# `output.elasticsearch.password` settings. The format is `<user>:<pass>`.
cloud.auth: "elastic:uB2AxBXuRlGMO0WwitEZ8VLt"

#================================ Outputs ==================================
```

It's not uncommon that a copy paste from above snippet may result in formatting issues in the file depending on your editor. A good sign is what you copied and pasted ended up in multiple lines. If you notice errors about "Error in line…" after you run the commands below, it is probably because of a formatting error.

6) Run the following command to start sending the metrics information to centralized Elasticsearch cluster (step 5)

```
.\metricbeat.exe -e
```

At this point your system should show up on the Infrastructure UI instructor is sharing on projector.

7) After you saw your system show up on Infrastructure UI of the common cluster, change the ***cloud.id*** and ***cloud.auth*** values in the .yml file back to the ones for the Elasticsearch deployment you created in Lab 1

Now we are ready to setup your Elastic deployment to receive the system metric data, visualize it, and create Machine Learning jobs to detect anomalies. Fortunately it only takes 2 commands! The first command sets up the system Metricbeat module in Elasticsearch and Kibana (index templates, dashboards, ML jobs etc), this only needs to be run once across all your Metricbeat installs, the second command starts Metricbeat.

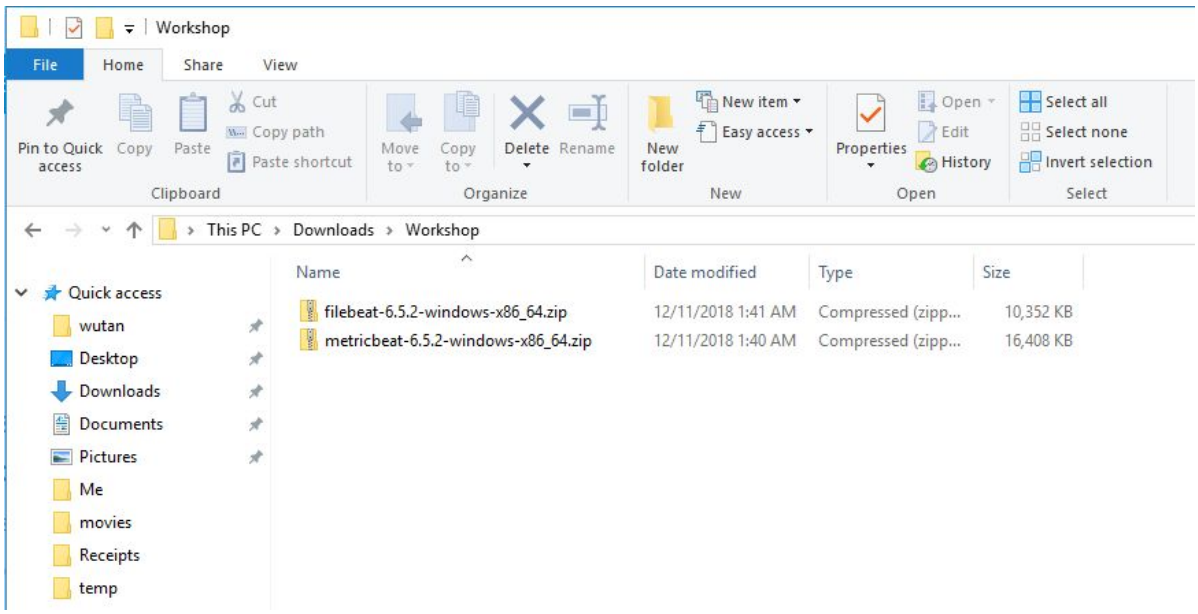
```
.\metricbeat.exe setup
.\metricbeat.exe -e
```

Windows Instructions

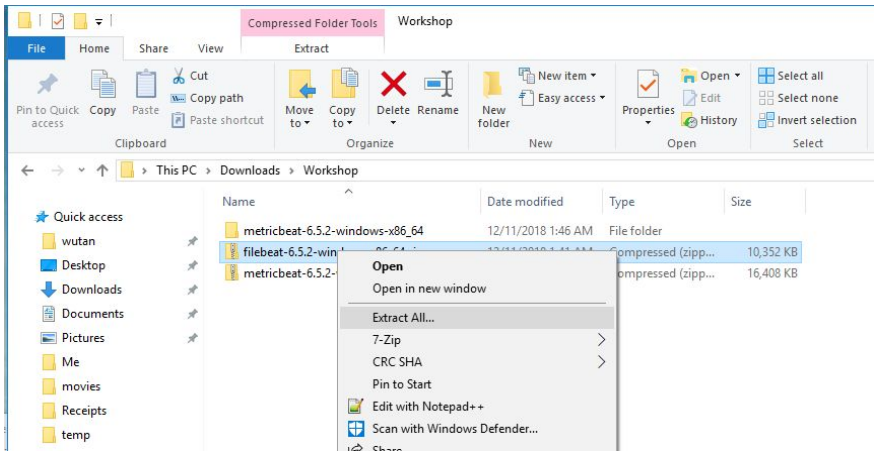
Filebeat

Filebeat

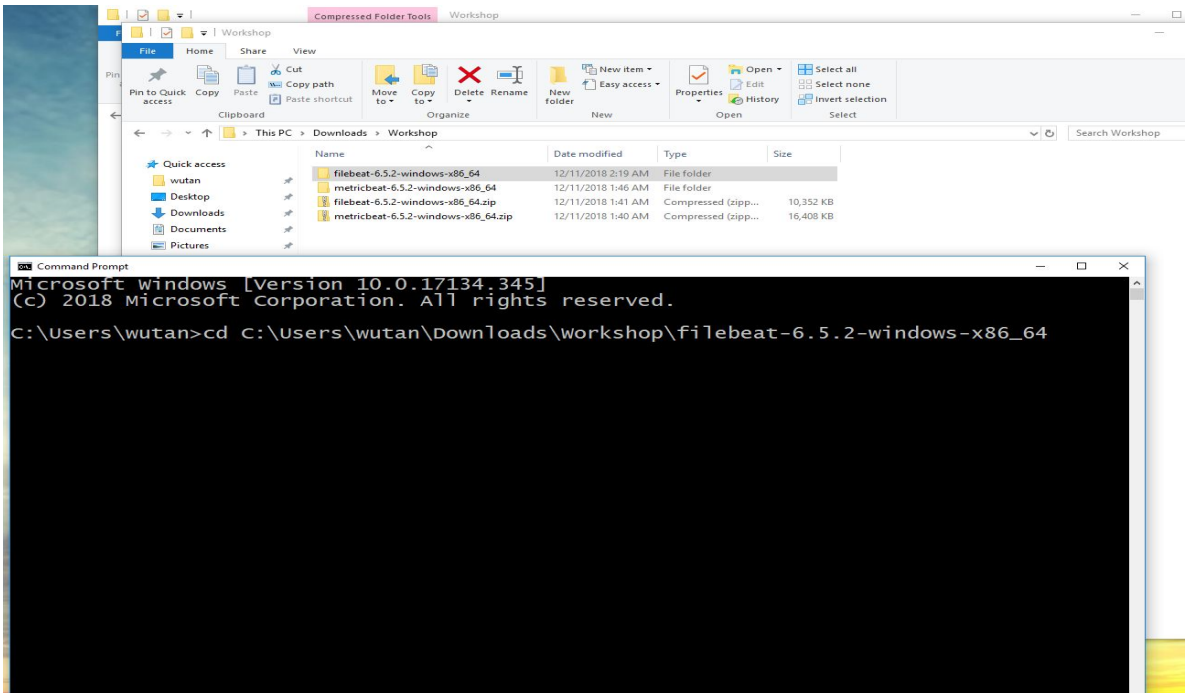
1) Open a Windows Explorer and navigate to the location that you downloaded filebeat to



2) Extract the file that you downloaded. (Your version of downloaded beat will be different than the screenshot below based on the latest version available at the time of download)



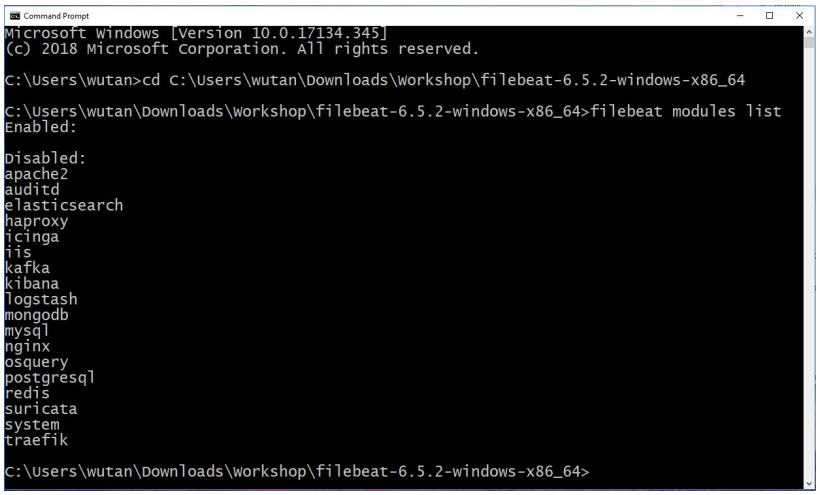
3) Open up a command prompt and type in `cd` (add a space after cd). Now drag-n-drop the extracted folder from step #2 to the command prompt window. Notice how it filled out the full path for you in the command prompt window? You can also type out the full path if you are a glutton for pain. Hit enter.



4) List the modules that are available

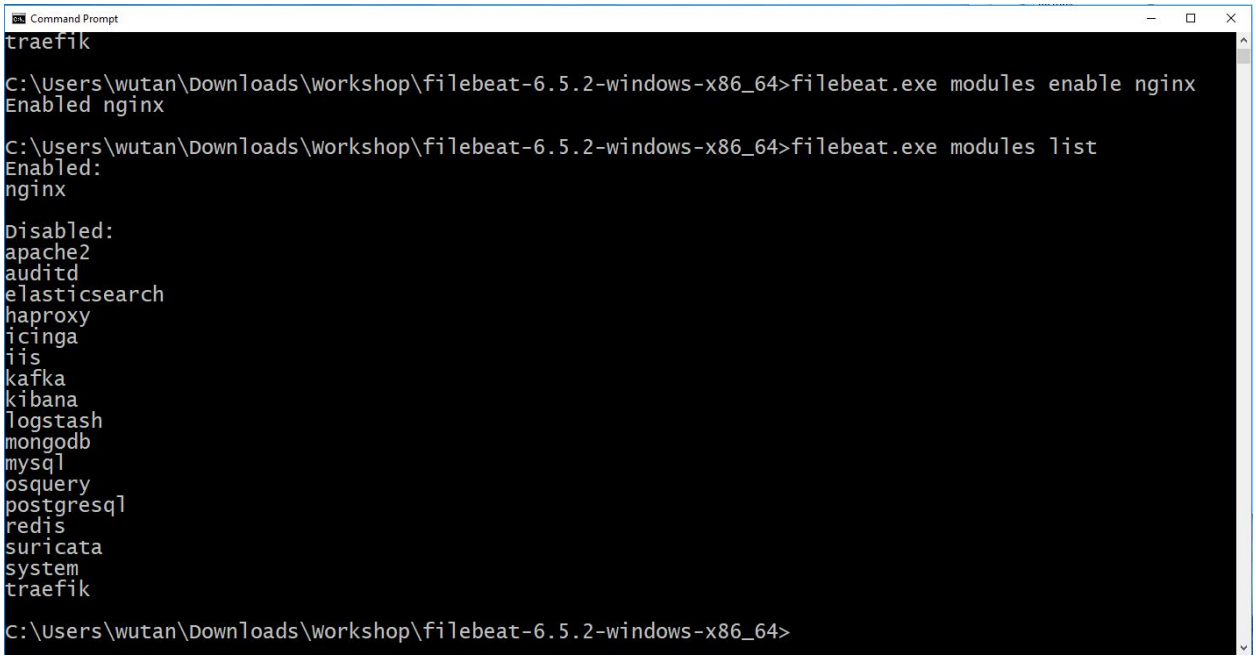
```
.\filebeat.exe modules list
```

You should see which modules are ***enabled*** and which modules are ***disabled.*** Out of the box there are no modules enabled.



5) Now let us enable the NGINX module so we can ingest NGINX logs

```
.\filebeat.exe modules enable nginx
```



6) Before we setup Elasticsearch to accept NGINX logs we need to tell Filebeat where Elasticsearch is and provide credentials to login. We do this by editing the configuration file for Filebeat called filebeat.yml.

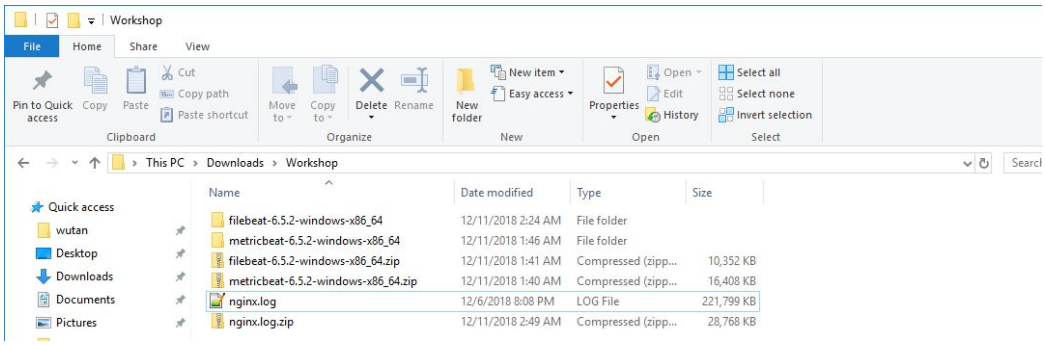
7) Follow the exact same procedure as you did when you setup Metricbeat and add the **cloud.id** and **cloud.auth** for your cluster filebeat.yml using the values from Lab 1.

```
cloud.id: "change-this-to-your-cloud-id"

cloud.auth: "elastic:change-this-to-your-deployment-password"
```

8) Since we don't actually have NGINX installed on our machine we are going to copy some real NGINX log files to the filesystem and tell the NGINX module where they are located.

Download the NGINX logs from the following URL and extract the log file:

[Sample NGINX Logs]



9) In your text editor, open the nginx.yml file under modules.d directory (this exists in the extracted filebeat-.../modules.d/nginx.yml directory. Add the following configuration to the nginx.yml file under the **Access Logs** configuration block (not the **Error Logs**)

```
- module: nginx
  # Access logs
  access:
    enabled: true

    # Set custom paths for the log files. If left empty,
    # Filebeat will choose the paths depending on your OS.
    var.paths: ["C:/Users/wutan/Downloads/Workshop/nginx.log"]

  # Error logs
  error:
    enabled: true

    # Set custom paths for the log files. If left empty,
    # Filebeat will choose the paths depending on your OS.
    #var.paths:

```

- **Change all backslashes in your Windows path to forward slashes**
- **Use the directory you expanded the NGINX log file to**

11) Now we are ready to setup Elasticsearch to receive the NGINX logs, visualize it, and create Machine Learning jobs to detect anomalies. Go one level above modules.d directory (filebeat-…. directory). Fortunately it only takes 2 commands, this first to setup the module like we did for Metricbeat and the second to start Filebeat.

```
cd ..
.\filebeat.exe setup
.\filebeat.exe -e
```

Mac/Linux Instructions

Metricbeat

1) Open a terminal and navigate to the location that you downloaded metricbeat to

```
cd ~/Downloads/
```

2) Expand the file that you downloaded:

```
tar -zxvf metricbeat-<version>-x86_64.tar.gz
```

3) Change directory into the metricbeat

```
cd metricbeat-<version>-x86_64
```

4) List models that are available

```
./metricbeat modules list
```

You should see which modules are ***enabled*** and which modules are ***disabled.*** Out of the box the **system** module is the only one that is enabled.

(Elastic actively adds modules to the beats in new versions to make the ingestion easier, the list of modules in the version you downloaded might be different)

5) Before we setup Elasticsearch to accept system metrics we need to tell Metricbeat where Elasticsearch is and provide credentials to login. We do this by editing the configuration file for Metricbeat called metricbeat.yml.

YAML files don't like hard tabs Do not use them if you are editing a .yml file because they will cause errors. To learn more about .yml files see this link: https://en.wikipedia.org/wiki/YAML

To start with, instead of sending metrics from your laptop to the cluster you have created in Lab 1, we are going to first send it to a shared Elasticsearch cluster created by the instructor. It will be used to demonstrate Kibana Infrastructure UI and also mimics a more realistic scenario - multiple endpoints sending metrics to single Elasticsearch Cluster.

Use your favorite text editor to open metricbeat.yml and replace **cloud.id** and **cloud.auth** with the following values:

```
cloud.id:
"observability_workshop:dXMtd2VzdDEuZ2NwLmNsb3VkLmVzLmlvJDAzN2N
hZDU4M2ZiMTRkYTc4MGUzYTE2NjA3OGU0OTc5JDdiZDg5MWFhOGJhMDQ1NzFiMz
g4N2ZkMDAzYjE2MWYz"
cloud.auth: "elastic:ncyRLYhJAl8akCSZLYBnBlED"
```

Example:

```
#============================ Elastic Cloud ================================

# These settings simplify using metricbeat with the Elastic Cloud (https://cloud.elastic.co/).

# The cloud.id setting overwrites the `output.elasticsearch.hosts` and
# `setup.kibana.host` options.
# You can find the `cloud.id` in the Elastic Cloud web UI.
cloud.id: "CentralizedBeatsMgmt:dXMtZWFzdC0xLmF3cy5mb3VuZC5pbyRkZmRiZTEwOWY2MmI0MTMxODhhZTRmM2U4ODYzNTVlZiRjYTEzMTg0MGFkMzc0OWZjYThkZWRhZTU5OWE0MjY2OQ=="

# The cloud.auth setting overwrites the `output.elasticsearch.username` and
# `output.elasticsearch.password` settings. The format is `<user>:<pass>`.
cloud.auth: "elastic:uB2AxBXuR1GMO0WwitEZ8VLt"

#================================ Outputs =====================================
```

It's very common that a copy paste from above snippet may result in formatting issues in the file depending on your editor. A good sign is what you copied and pasted ended up in multiple lines. If you notice errors about "Error in line…" after you run the commands that follow, it is probably because of a formatting error.

6) Run the following command to start sending the metrics information to centralized Elasticsearch cluster (step 5)

```
./metricbeat -e
```

At this point your system should show up on the Infrastructure UI instructor is sharing on projector.

7) **IMPORTANT:**After you saw your system show up on Infrastructure UI of the common cluster, change the ***cloud.id*** and ***cloud.auth*** values back to the ones for your Elasticsearch deployment you created in Lab 1

Now we are ready to setup your Elastic deployment to receive the system metric data, visualize it, and create Machine Learning jobs to detect anomalies. Fortunately it only takes 2 commands! The first command sets up the system Metricbeat module in Elasticsearch and Kibana (index templates, dashboards, ML jobs etc), this only needs to be run once across all your Metricbeat installs, the second command starts Metricbeat.

```
./metricbeat setup
./metricbeat -e
```

Filebeat

1) Open a terminal and navigate to the location that you downloaded filebeat to

```
cd ~/Downloads/
```

2) Expand the file that you downloaded:

```
tar -zxvf filebeat-<version>-x86_64.tar.gz
```

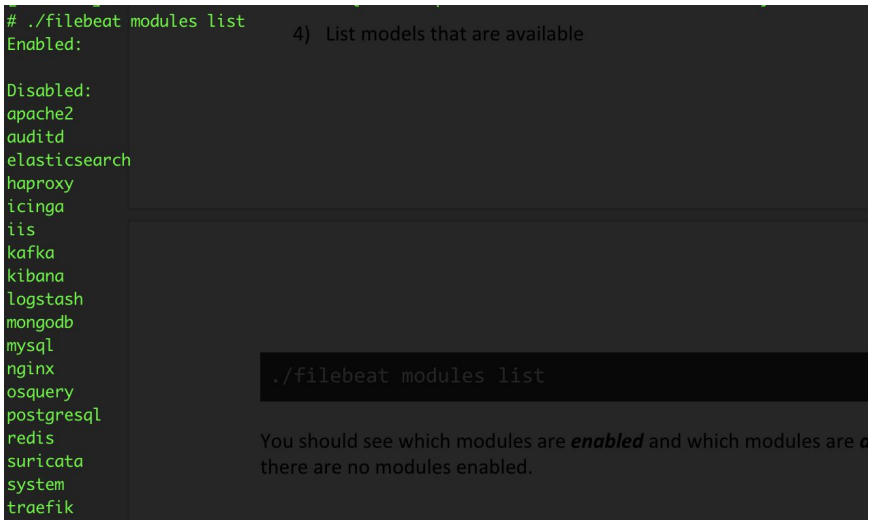
3) Change directory into the filebeat

```
cd filebeat-<version>-x86_64
```

4) List the modules that are available

```
./filebeat modules list
```

You should see which modules are ***enabled*** and which modules are ***disabled.*** Out of the box there are no modules enabled.



5) Now let us enable the NGINX module so we can ingest NGINX logs

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

6) Before we setup Elasticsearch to accept NGINX logs we need to tell Filebeat where Elasticsearch is and provide credentials to login. We do this by editing the configuration file for Filebeat called filebeat.yml.

7) Follow the exact same procedure as you did when you setup Metricbeat and add the **cloud.id** and **cloud.auth** for your cluster filebeat.yml using the values from Lab 1.

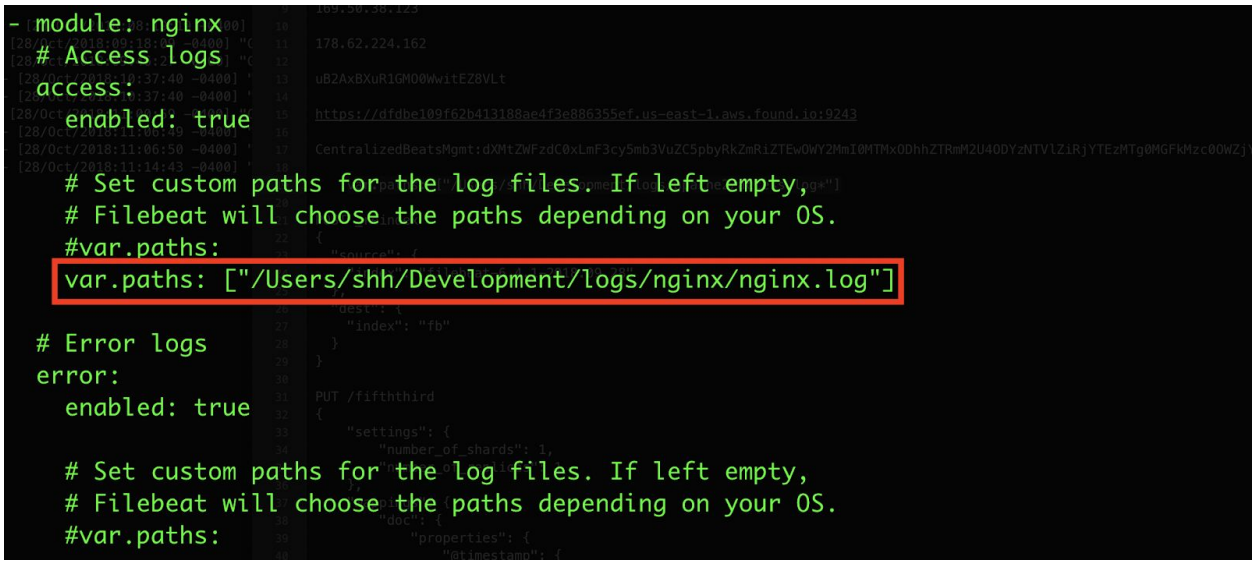
```
cloud.id: "change-this-to-your-cloud-id"

cloud.auth: "elastic:change-this-to-your-deployment-password"
```

8) Since we don't actually have NGINX installed we are going to copy some real NGINX log files to the filesystem and tell the NGINX module where they are located.

Download the NGINX logs from the following URL and extract the log file:

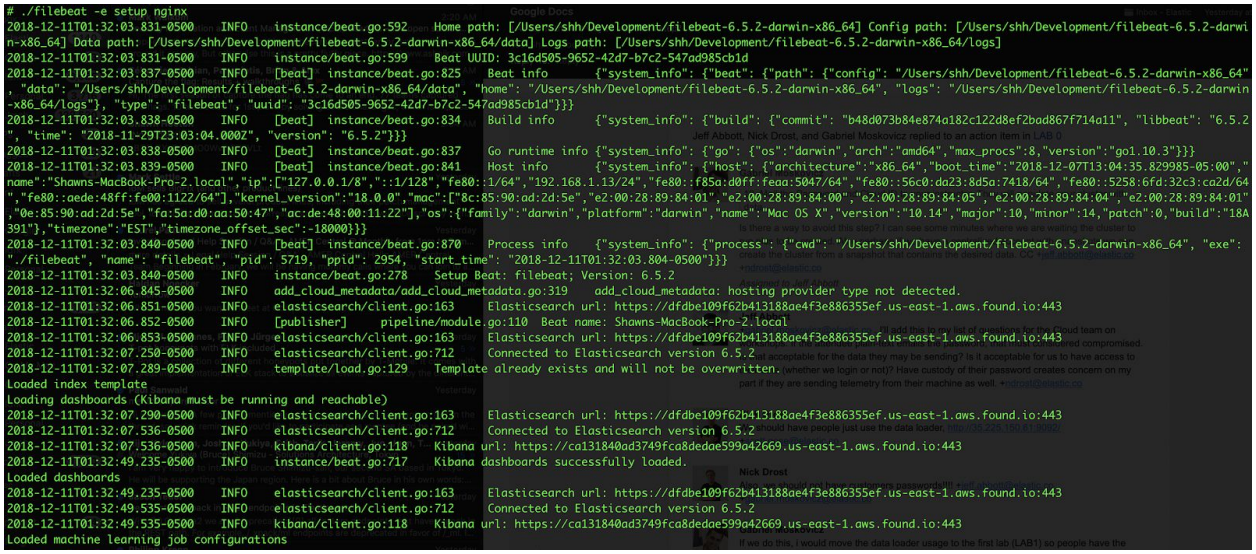
[Sample NGINX Logs]

9) In your text editor, open the nginx.yml file under modules.d directory (this exists in the extracted filebeat-.../modules.d/nginx.yml directory. Add the following configuration to the nginx.yml file under the **Access Logs** configuration block (not the **Error Logs**).



11) Now we are ready to setup Elasticsearch to receive the NGINX logs, visualize it, and create Machine Learning jobs to detect anomalies. Go one level above modules.d directory (filebeat-.... directory). Fortunately it only takes 2 commands, this first to setup the module like we did for Metricbeat and the second to start Filebeat.

```
./filebeat setup
./filebeat -e
```

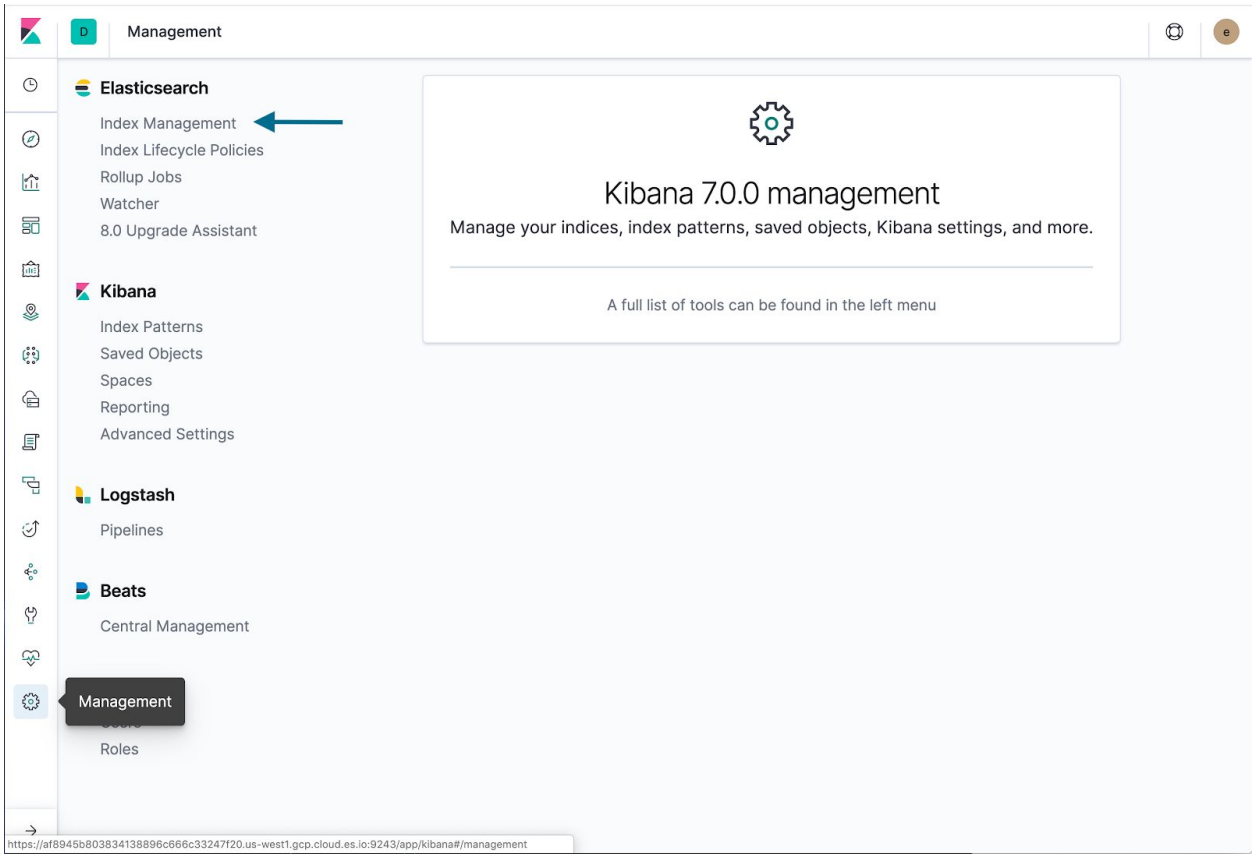

./filebeat -e setup nginx
2018-12-11T01:32:03.831-0500 INFO instance/beat.go:592 Home path: [/Users/shh/Development/filebeat-6.5.2-darwin-x86_64] Config path: [/Users/shh/Development/filebeat-6.5.2-darwin-x86_64] Data path: [/Users/shh/Development/filebeat-6.5.2-darwin-x86_64/data] Logs path: [/Users/shh/Development/filebeat-6.5.2-darwin-x86_64/logs]
2018-12-11T01:32:03.831-0500 INFO instance/beat.go:599 Beat UUID: 3c16d505-9652-42d7-b7c2-547ad985cb1d
2018-12-11T01:32:03.837-0500 INFO [beat] instance/beat.go:825 Beat info {"system_info": {"beat": {"path": {"config": "/Users/shh/Development/filebeat-6.5.2-darwin-x86_64", "data": "/Users/shh/Development/filebeat-6.5.2-darwin-x86_64/data", "home": "/Users/shh/Development/filebeat-6.5.2-darwin-x86_64", "logs": "/Users/shh/Development/filebeat-6.5.2-darwin-x86_64/logs"}, "type": "filebeat", "uuid": "3c16d505-9652-42d7-b7c2-547ad985cb1d"}}}
2018-12-11T01:32:03.838-0500 INFO [beat] instance/beat.go:834 Build info {"system_info": {"build": {"commit": "b48d073b84e874a182c122d8ef2bad867f714a11", "libbeat": "6.5.2", "time": "2018-11-29T23:03:04.000Z", "version": "6.5.2"}}}
2018-12-11T01:32:03.838-0500 INFO [beat] instance/beat.go:837 Go runtime info {"system_info": {"go": {"os":"darwin","arch":"amd64","max_procs":8,"version":"go1.10.3"}}}
2018-12-11T01:32:03.839-0500 INFO [beat] instance/beat.go:841 Host info {"system_info": {"host": {"architecture":"x86_64","boot_time":"2018-12-07T13:04:35.829985-05:00","name":"Shawns-MacBook-Pro-2.local","ip":["127.0.0.1/8","::1/128","fe80::1/64","192.168.1.13/24","fe80::f85a:d0ff:feaa:5047/64","fe80::56c0:da23:8d5a:7418/64","fe80::5258:6fd:32c3:ca2d/64","fe80::aede:48ff:fe00:1122/64"],"kernel_version":"18.0.0","mac":["8c:85:90:ad:2d:5e","e2:00:28:89:84:01","e2:00:28:89:84:00","e2:00:28:89:84:05","e2:00:28:89:84:04","e2:00:28:89:84:01","0e:85:90:ad:2d:5e","fa:5a:d0:aa:50:47","ac:de:48:00:11:22"],"os":{"family":"darwin","platform":"darwin","name":"Mac OS X","version":"10.14","major":10,"minor":14,"patch":0,"build":"18A391"},"timezone":"EST","timezone_offset_sec":-18000}}}
2018-12-11T01:32:03.840-0500 INFO [beat] instance/beat.go:870 Process info {"system_info": {"process": {"cwd": "/Users/shh/Development/filebeat-6.5.2-darwin-x86_64", "exe": "./filebeat", "name": "filebeat", "pid": 5719, "ppid": 2954, "start_time": "2018-12-11T01:32:03.804-0500"}}}
2018-12-11T01:32:03.840-0500 INFO instance/beat.go:278 Setup Beat: filebeat; Version: 6.5.2
2018-12-11T01:32:06.845-0500 INFO add_cloud_metadata/add_cloud_metadata.go:319 add_cloud_metadata: hosting provider type not detected.
2018-12-11T01:32:06.851-0500 INFO elasticsearch/client.go:163 Elasticsearch url: https://dfdbe109f62b413188ae4f3e886355ef.us-east-1.aws.found.io:443
2018-12-11T01:32:06.852-0500 INFO [publisher] pipeline/module.go:110 Beat name: Shawns-MacBook-Pro-2.local
2018-12-11T01:32:06.853-0500 INFO elasticsearch/client.go:163 Elasticsearch url: https://dfdbe109f62b413188ae4f3e886355ef.us-east-1.aws.found.io:443
2018-12-11T01:32:07.250-0500 INFO elasticsearch/client.go:712 Connected to Elasticsearch version 6.5.2
2018-12-11T01:32:07.289-0500 INFO template/load.go:129 Template already exists and will not be overwritten.
Loaded index template
Loading dashboards (Kibana must be running and reachable)
2018-12-11T01:32:07.290-0500 INFO elasticsearch/client.go:163 Elasticsearch url: https://dfdbe109f62b413188ae4f3e886355ef.us-east-1.aws.found.io:443
2018-12-11T01:32:07.536-0500 INFO elasticsearch/client.go:712 Connected to Elasticsearch version 6.5.2
2018-12-11T01:32:07.536-0500 INFO kibana/client.go:118 Kibana url: https://ca131840ad3749fca8dedae599a42669.us-east-1.aws.found.io:443
2018-12-11T01:32:49.235-0500 INFO instance/beat.go:717 Kibana dashboards successfully loaded.
Loaded dashboards
2018-12-11T01:32:49.235-0500 INFO elasticsearch/client.go:163 Elasticsearch url: https://dfdbe109f62b413188ae4f3e886355ef.us-east-1.aws.found.io:443
2018-12-11T01:32:49.535-0500 INFO elasticsearch/client.go:712 Connected to Elasticsearch version 6.5.2
2018-12-11T01:32:49.535-0500 INFO kibana/client.go:118 Kibana url: https://ca131840ad3749fca8dedae599a42669.us-east-1.aws.found.io:443
Loaded machine learning job configurations

Validate Data in Kibana

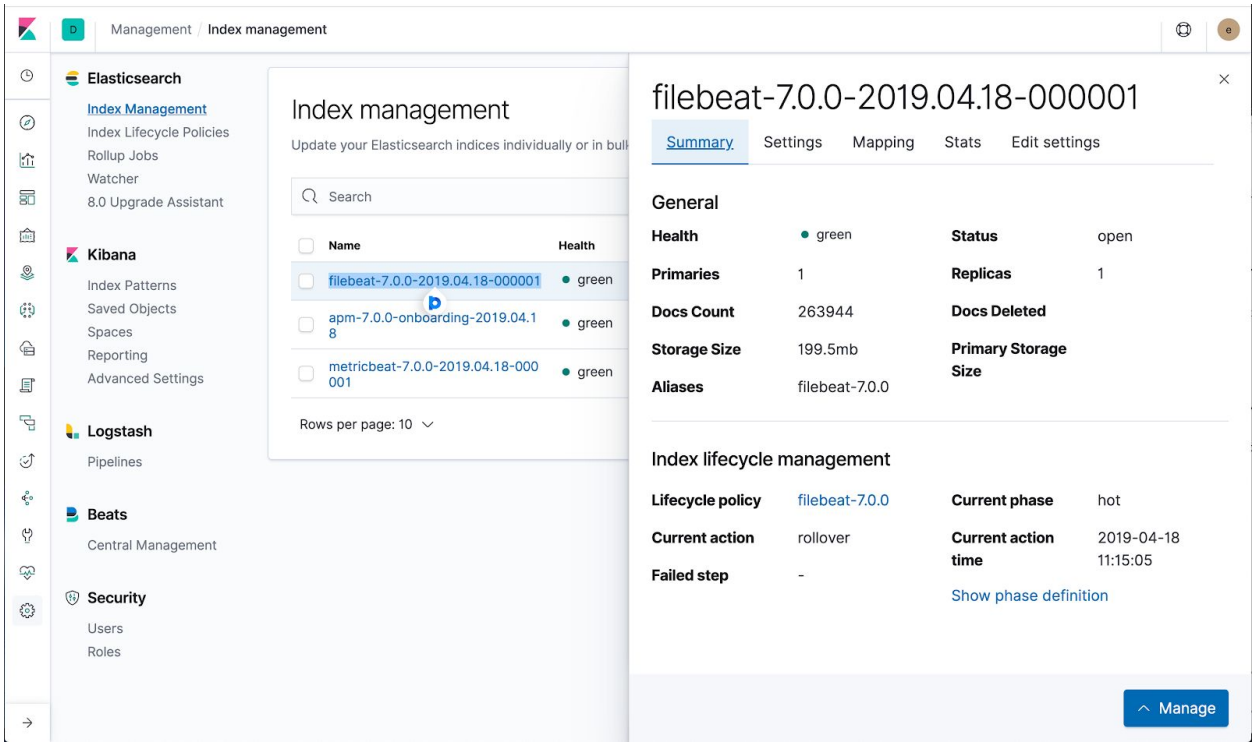
At this point let's look at the data in Kibana by looking at the index.

1) Go to your Kibana (from your deployment in Lab 1)

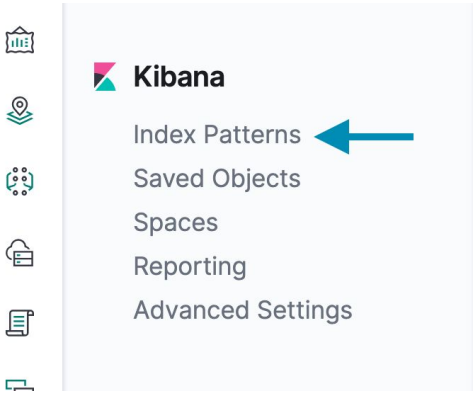
2) Click on the Management App and then on Index Management



3) Look for Indexes named filebeat-<version>-YYYY.MM.DD-000001 where version is the current version of the product and YYYY, MM, and DD represent the year, month, and day respectively. Examine Docs Count, Storage Size, and Primary Storage Size. Realistic data like this provides a wonderful opportunity to look at your data and how much disk space it consumes to help size your environment accurately.

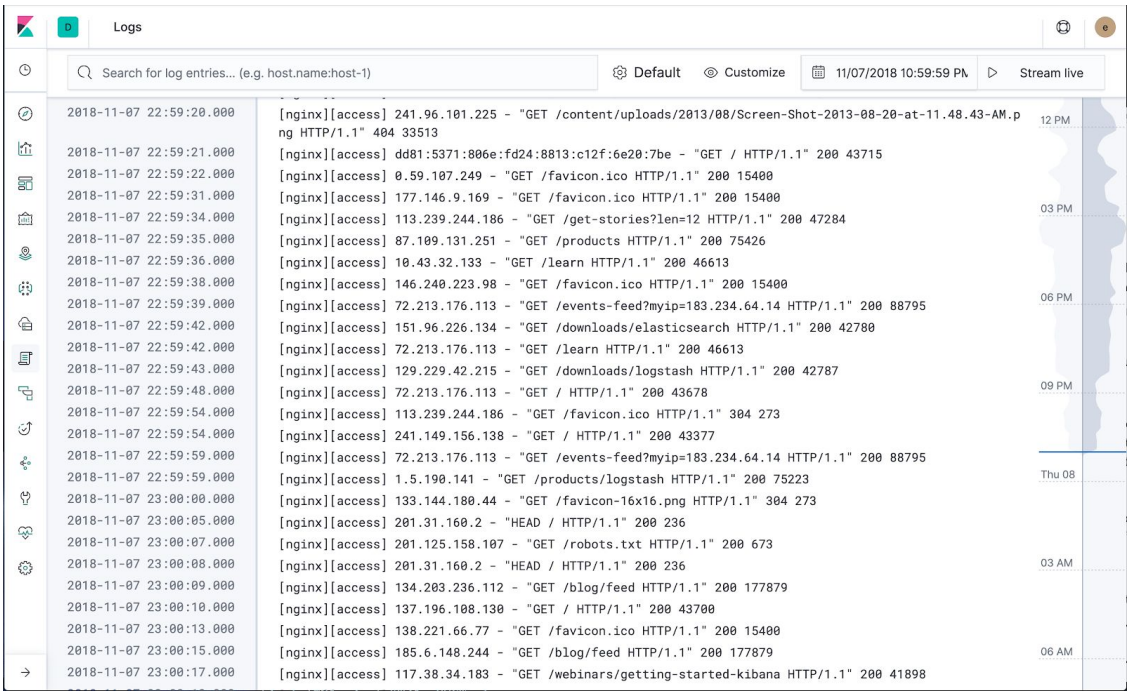


4) Now click on the ”Index Patterns” link on the side navigation of the Management App. Index patterns tell Kibana which Elasticsearch indices you want to explore. An index pattern can match the name of a single index, or include a wildcard (*) to match multiple indices.



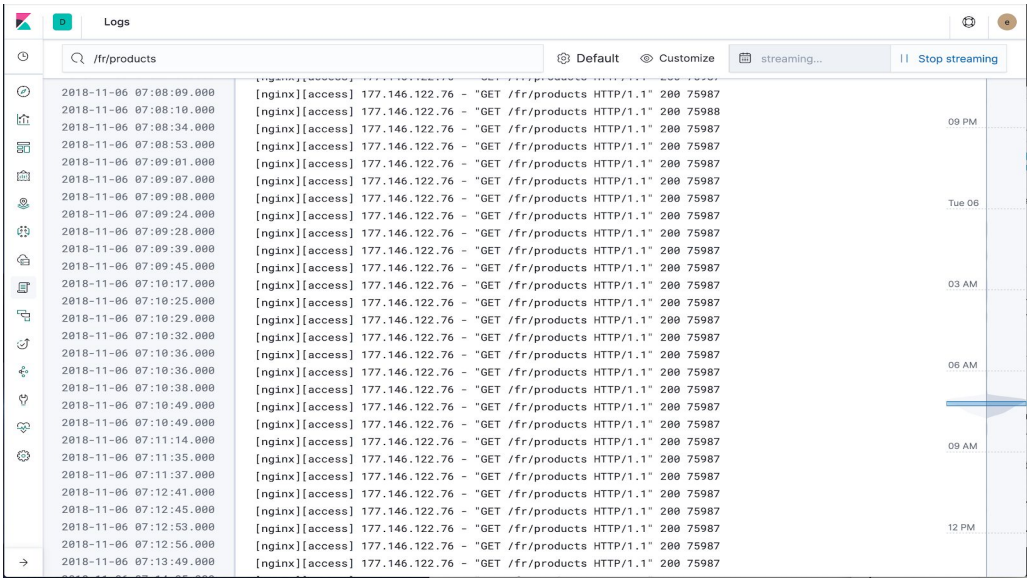
5) Now verify that metricbeat and filebeat index patterns exist. Notice the wildcard pattern. Examine the fields, notice the field data type, whether it is searchable and aggregatable.

6) In Kibana click on “Logs App” item in the side navigation. The data that you see below is coming from filebeat-* indices.



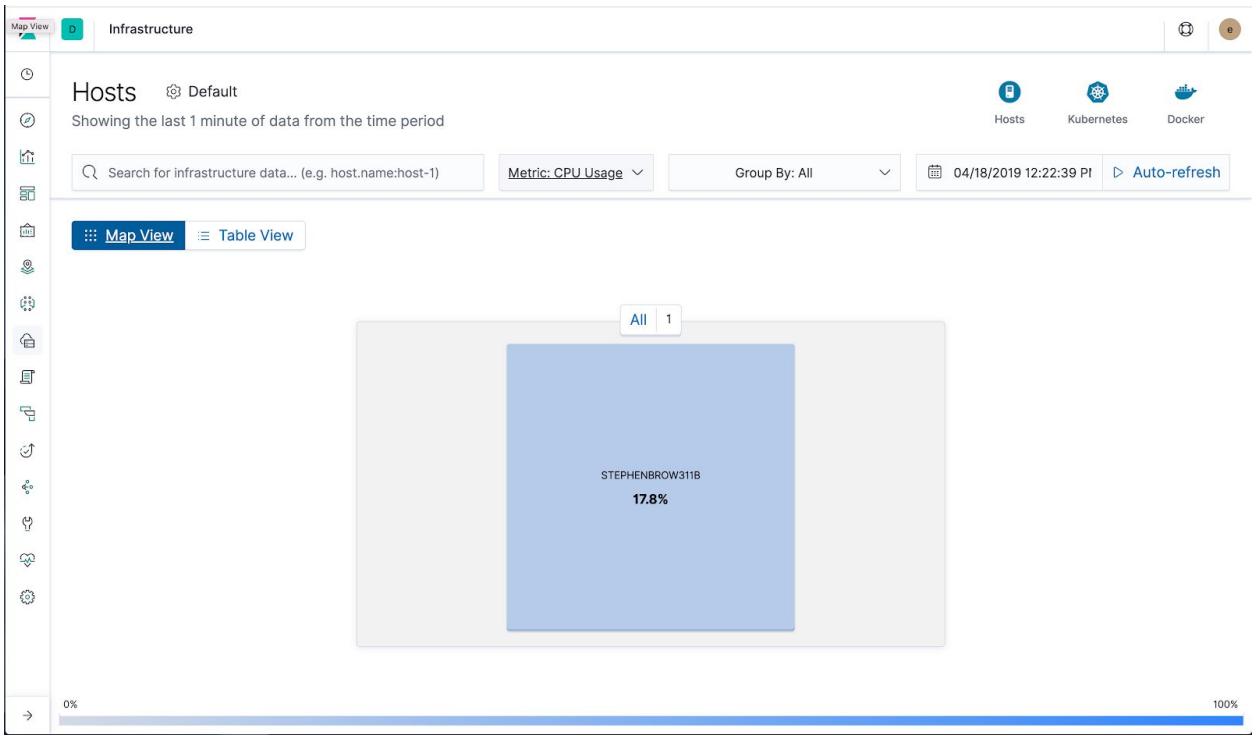
7) In the top right corner click on “Stream live”. Notice how the screen starts to update as more logs flow into Elastic. This feature aims to simplify “tailing the log” experience.

8) In the search bar search for “/fr/products“.
(Based on the WiFi speed the upload of logs to the Elasticsearch might be a little slow, give it a few minutes if you don’t see any results for /fr/products)



This functionality is powered by a search engine (Elasticsearch) and search features are still available to you.

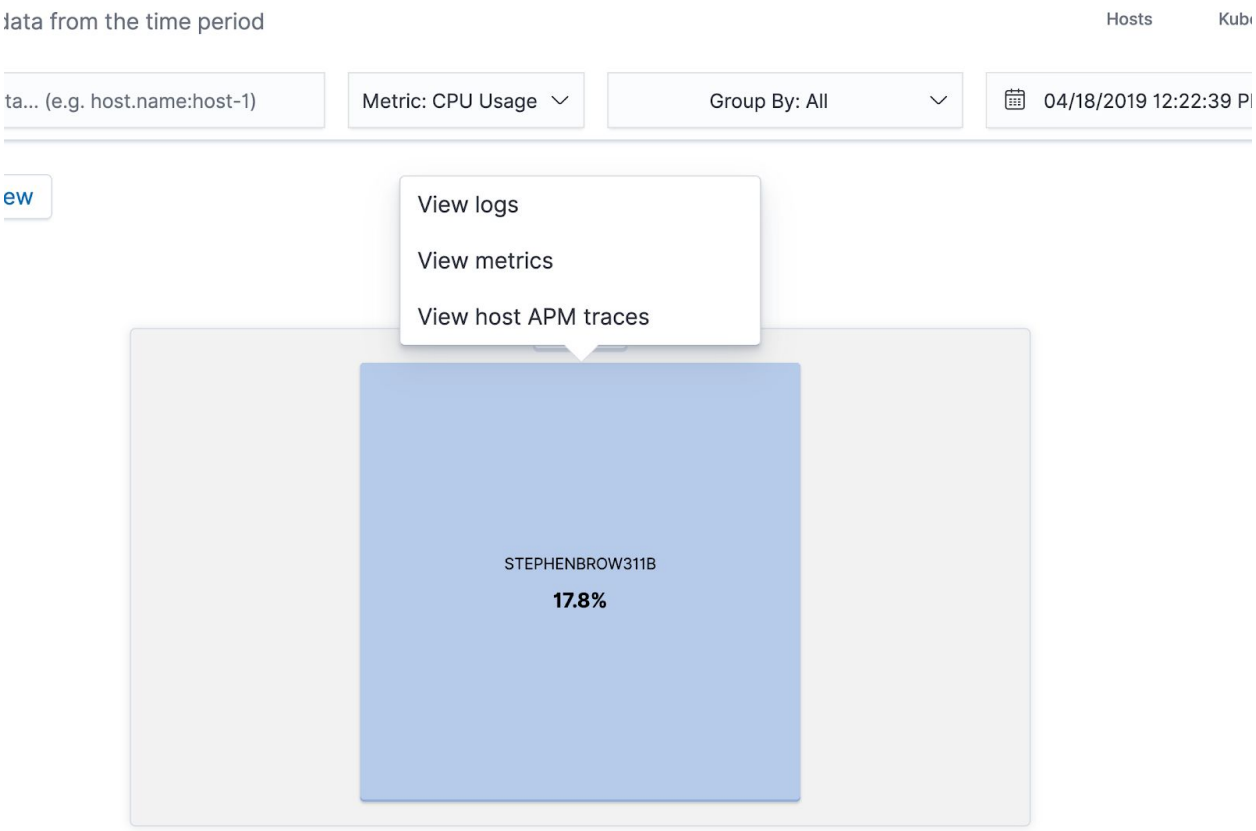
9) Now click on the “Infrastructure App” item in the side navigation.

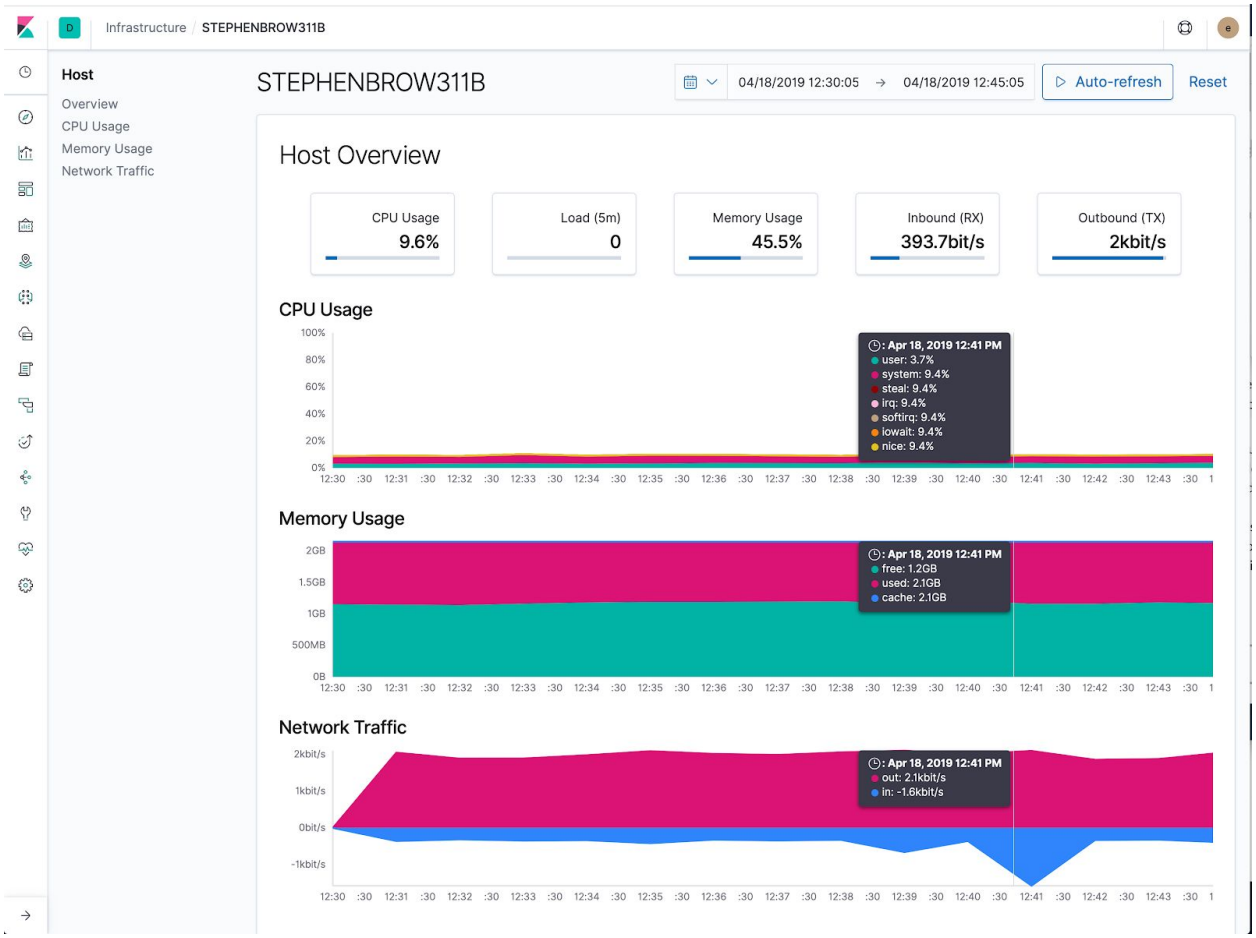


At the moment you see metrics only from one host, but imagine this same view (showcased in the presentation slides and during the instructor’s demo) where you have multiple hosts monitored here on one screen.

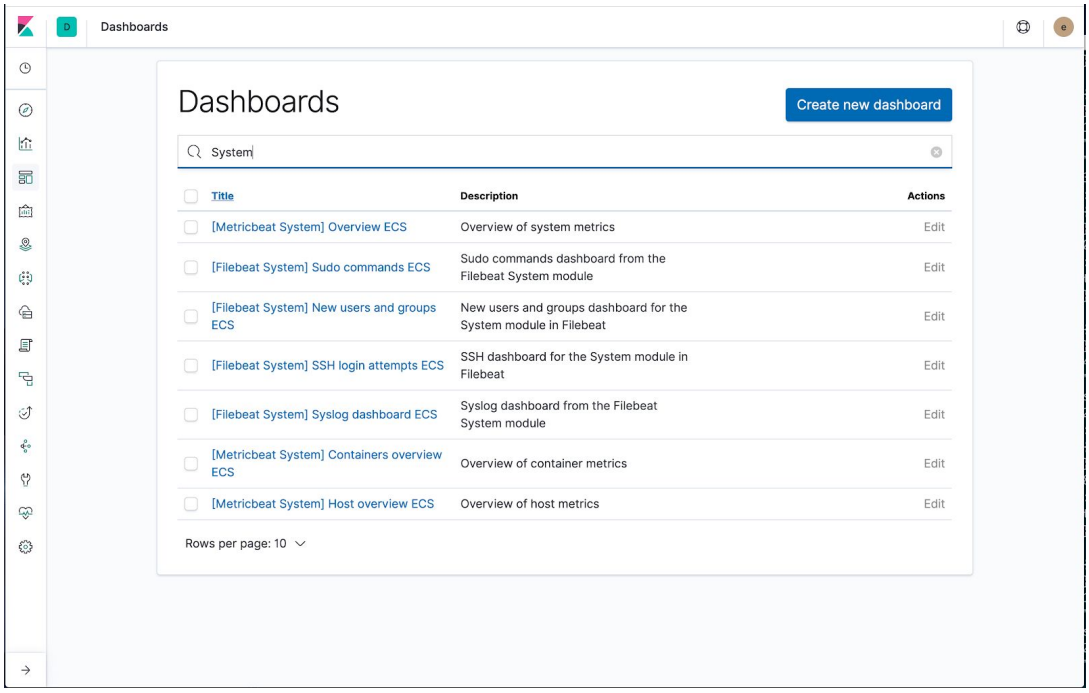
10) Current metric displayed is CPU Usage. Click on the dropdown and select Memory Usage, Load, and other metrics. Note – you might not have data for everything, but this will give you an idea of what kind of metrics could drive the display of the screen.

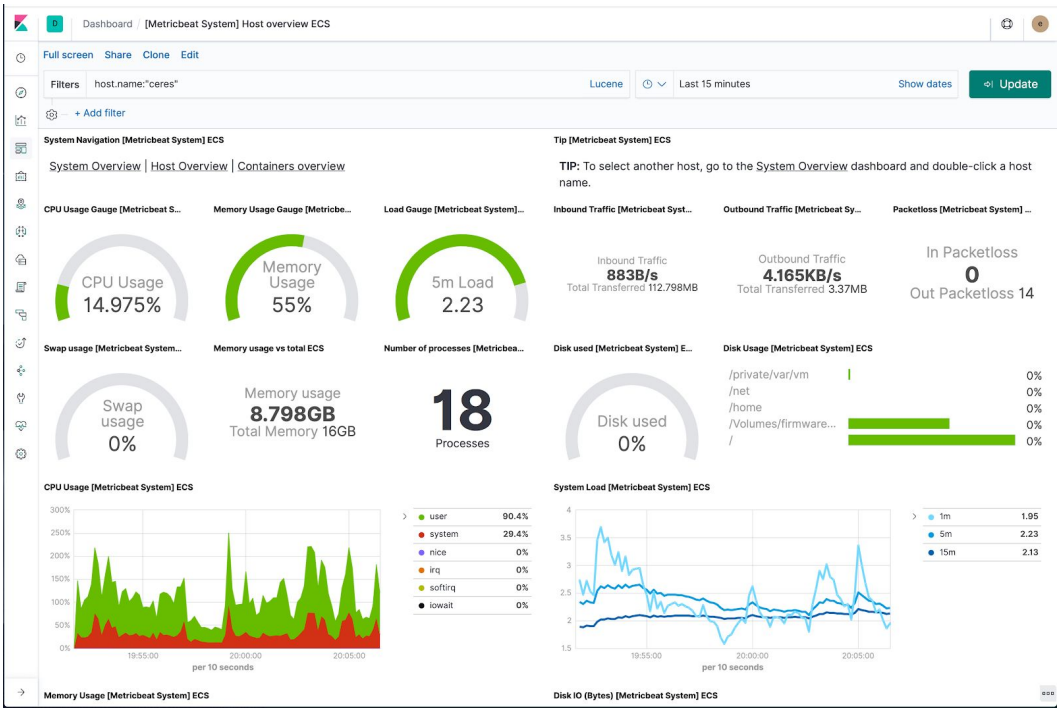
11) Click on the host and then click on “View Metrics”. You will end up on the quick summary metrics screen for the host.


lata from the time period
Hosts
Kube
ta... (e.g. host.name:host-1)
Metric: CPU Usage
Group By: All
04/18/2019 12:22:39 P
ew
View logs
View metrics
View host APM traces
STEPHENBROW311B
17.8%



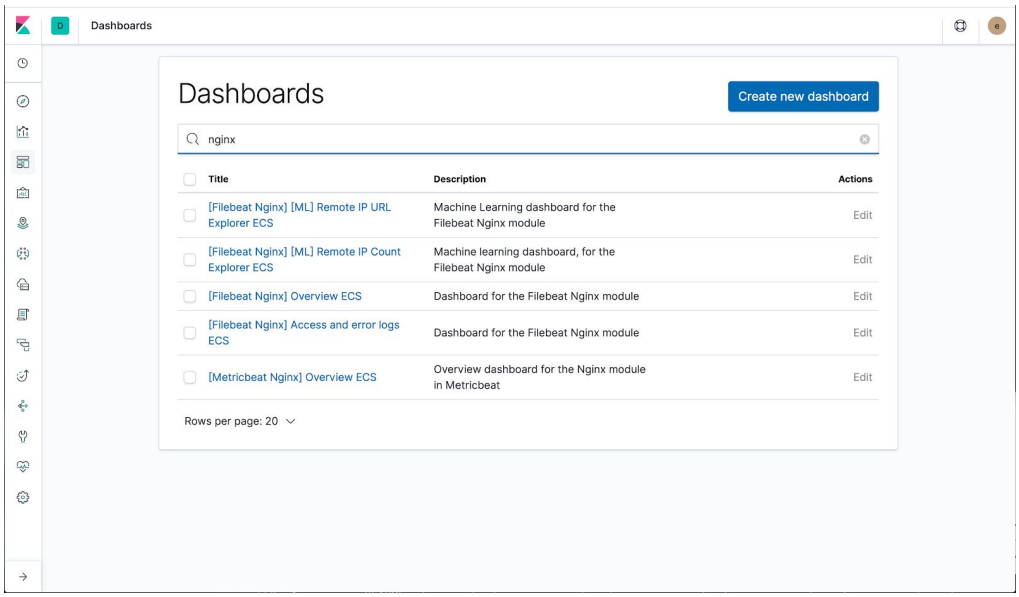
12) Now let's take a look at the Dashboards that come OOB with Metricbeat and Filebeat. Click on Dashboards on the menu. A list with whole bunch of dashboards will display. Type in the search bar "System" and click on **[Metricbeat System] Host overview**. Make sure time picker in the top right corner is set for the "Last 15 min".



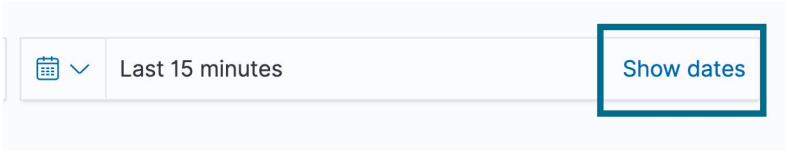


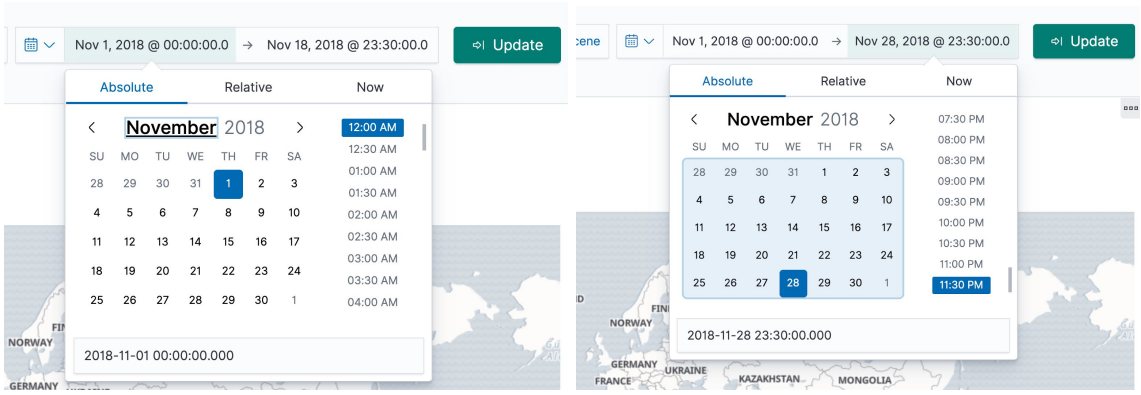
You end up on a dashboard that gives you complete metrics overview of the host where you have metricbeat running. Essentially with just running a few commands you're now able to collect the metrics and have and have a graphic representation of your computer's performance. Imagine running this at scale and having that same real time view of 100s of hosts

13) Click on Dashboard again. This time search for “Nginx”. Click on **[Filebeat Nginx] Overview** dashboard.

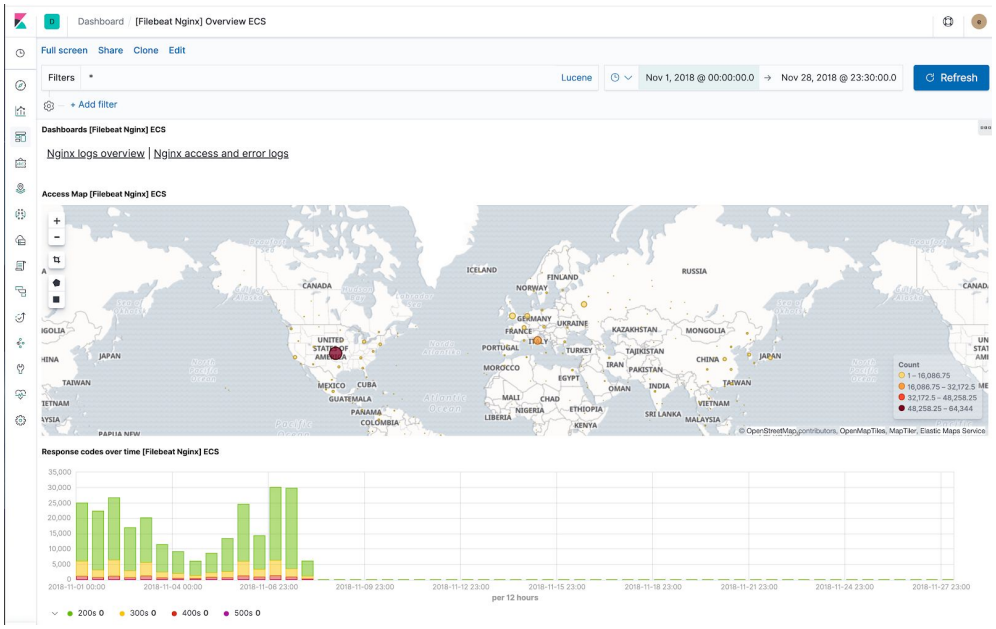


When the dashboard opens up in Timepicker select “Show Dates” and the set option “Absolute” and set the “From” to Nov 1st 2018 and “To” to Nov 28th 2018.



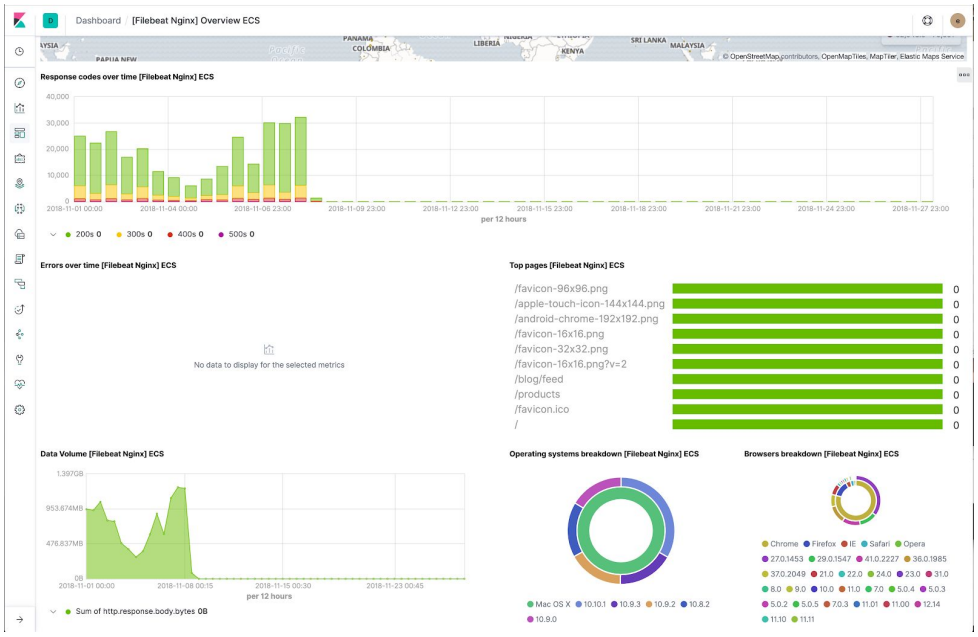


You end up on a dashboard that looks like this:

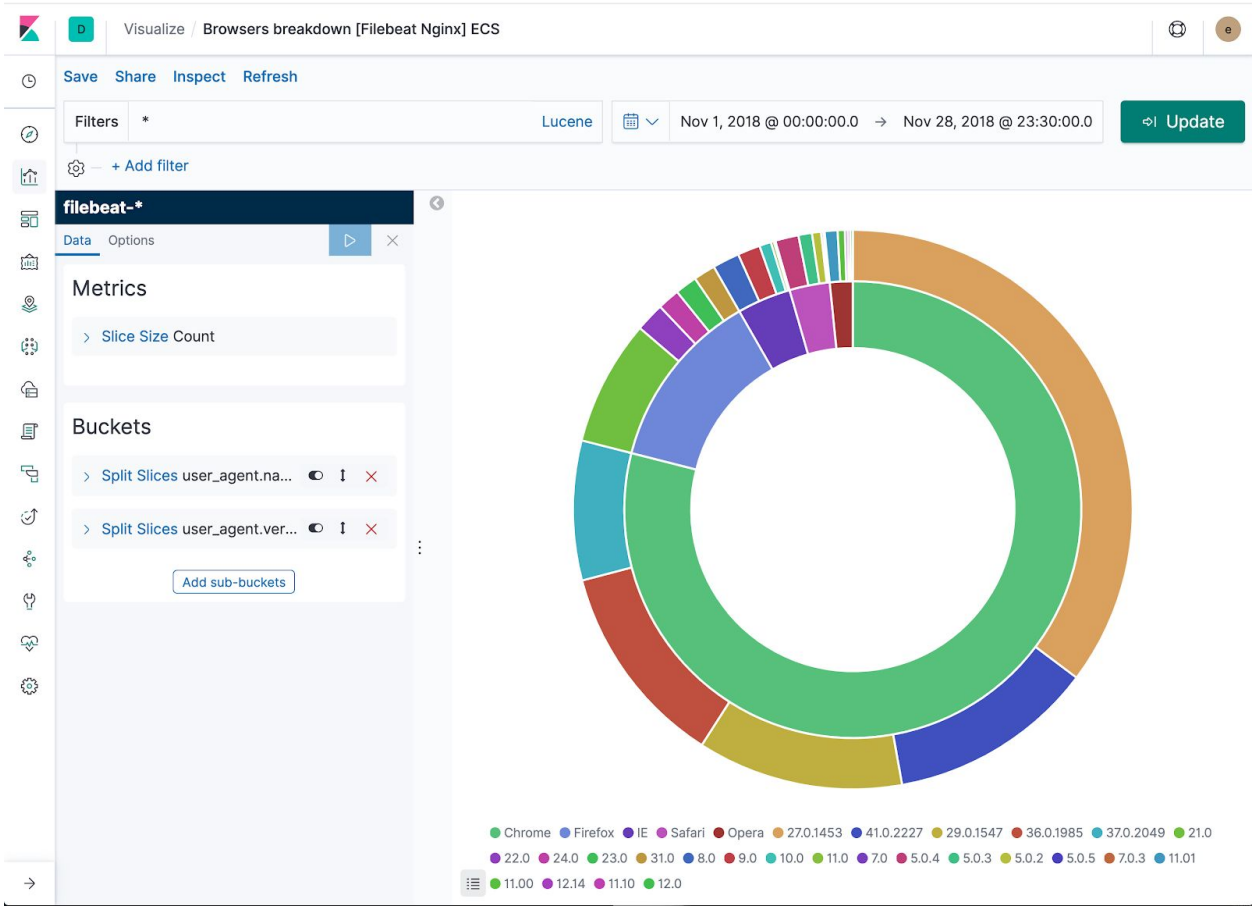


If you do not see the data all the way to November 28th it means it is still loading. Turn on Auto-Refresh (next to date picker) and see how your dashboards keeps updating in real time.

14) These dashboards are also a great example on how to build visualizations in Kibana. Feel free to click on “Edit” (next to Auto-Refresh option) and then edit a particular visualization to see how it was built.



Pie chart visualization:


Visualize / Browsers breakdown [Filebeat Nginx] ECS
Save Share Inspect Refresh
Filters * Lucene
Nov 1, 2018 @ 00:00:00.0 → Nov 28, 2018 @ 23:30:00.0
Update
+ Add filter
filebeat-*
Data Options
Metrics
Slice Size Count
Buckets
Split Slices user_agent.na...
Split Slices user_agent.ver...
Add sub-buckets
Chrome Firefox IE Safari Opera 27.0.1453 41.0.2227 29.0.1547 36.0.1985 37.0.2049 21.0
22.0 24.0 23.0 31.0 8.0 9.0 10.0 11.0 7.0 5.0.4 5.0.3 5.0.2 5.0.5 7.0.3 11.01
11.00 12.14 11.10 12.0