LAB 1

Introduction

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

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.

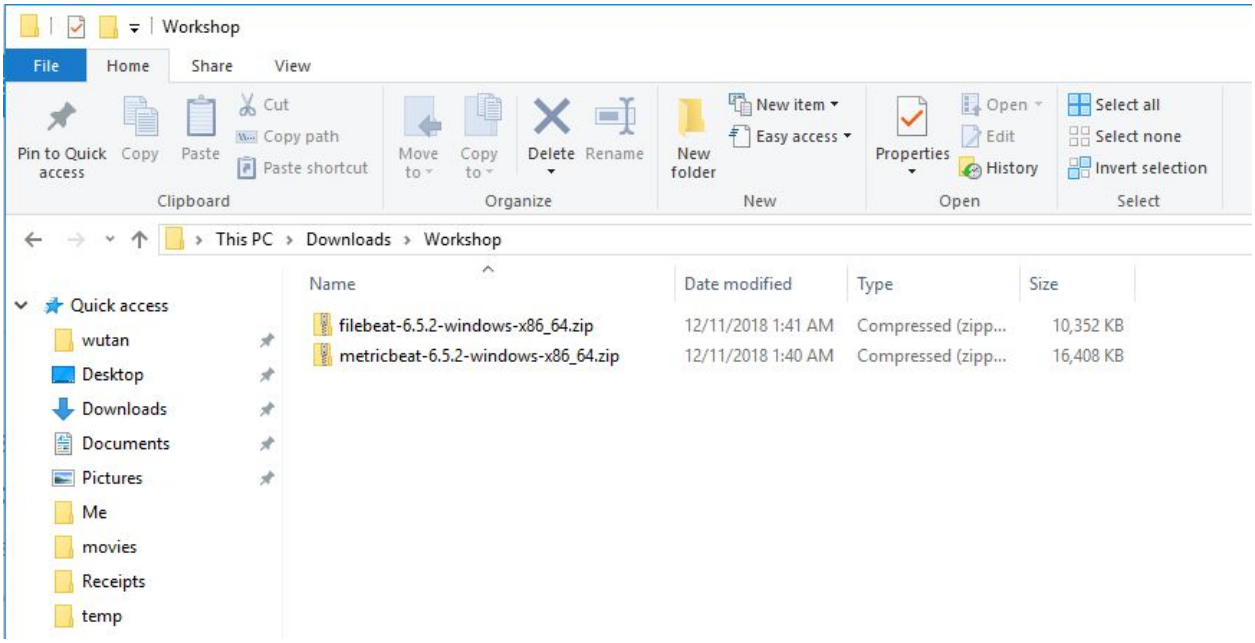
Software Download

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

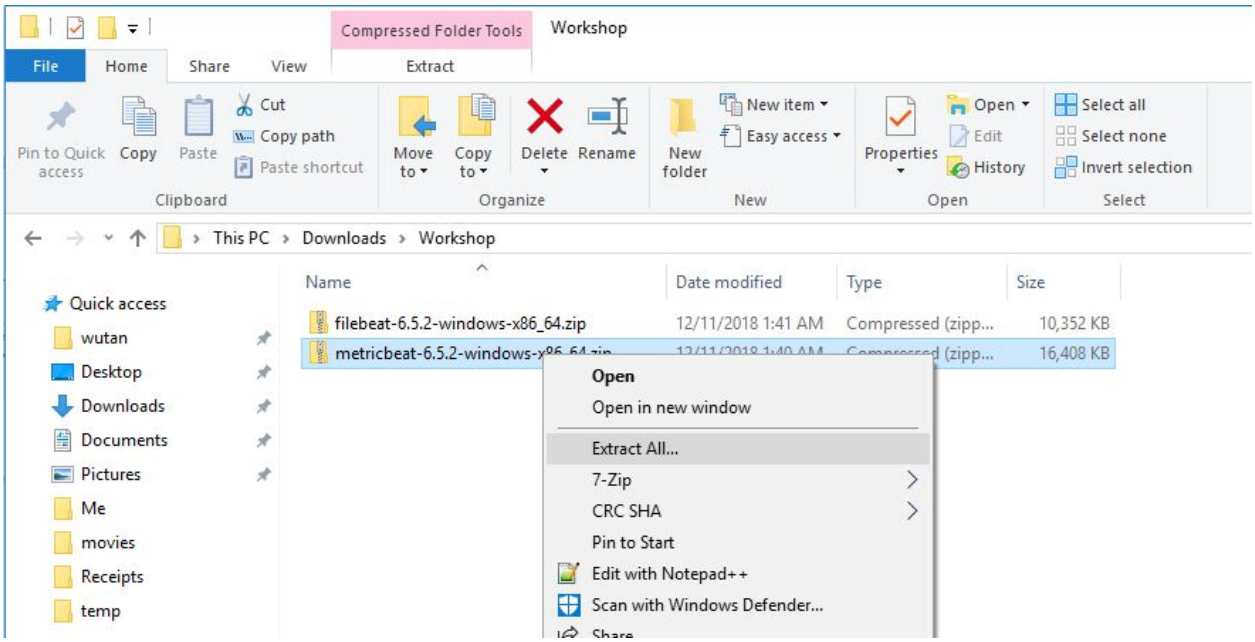
Windows Instructions

Metricbeat

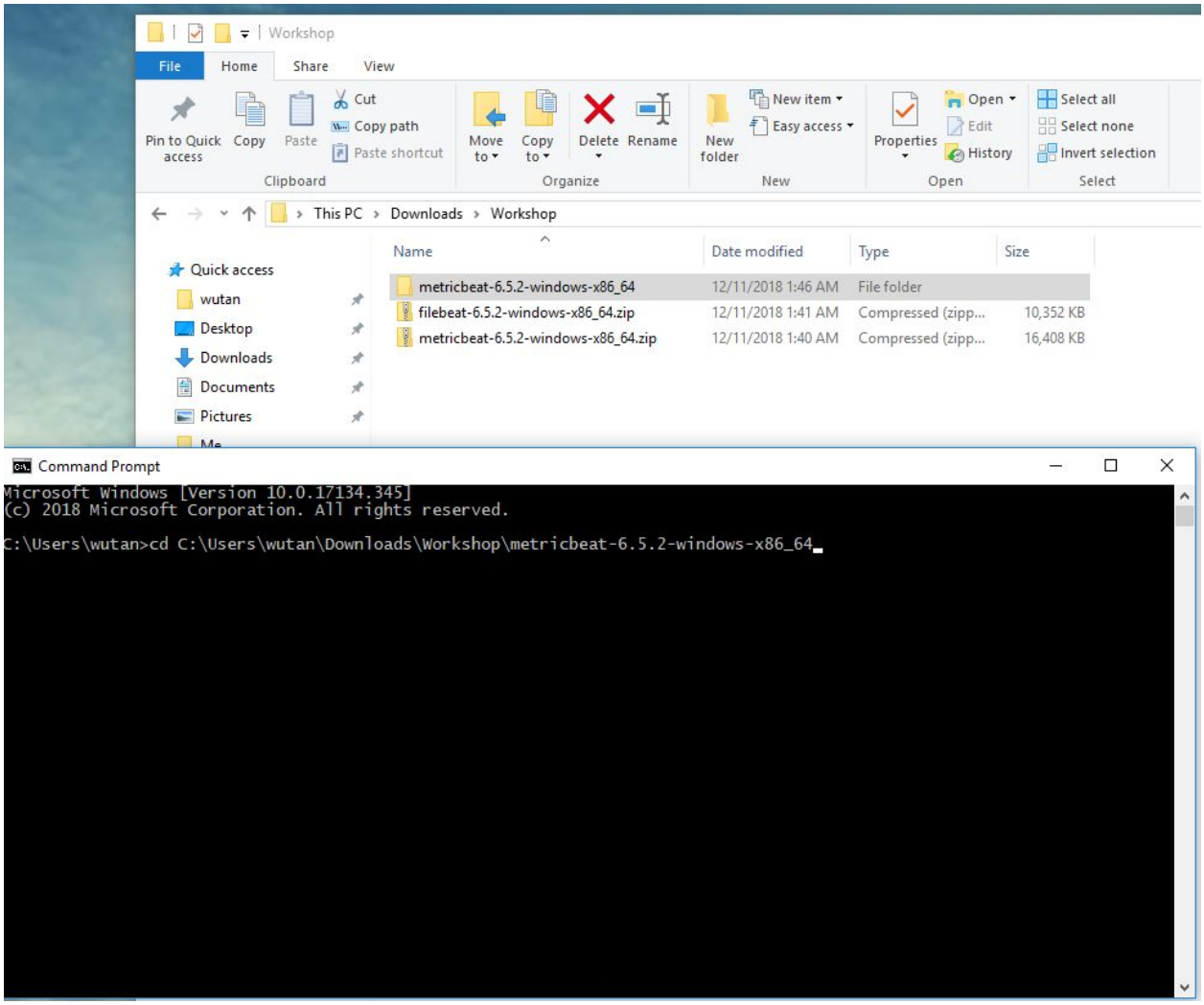
1) Open a Windows Explorer and navigate to the location that you downloaded metricbeat to



2) Expand the file that you downloaded



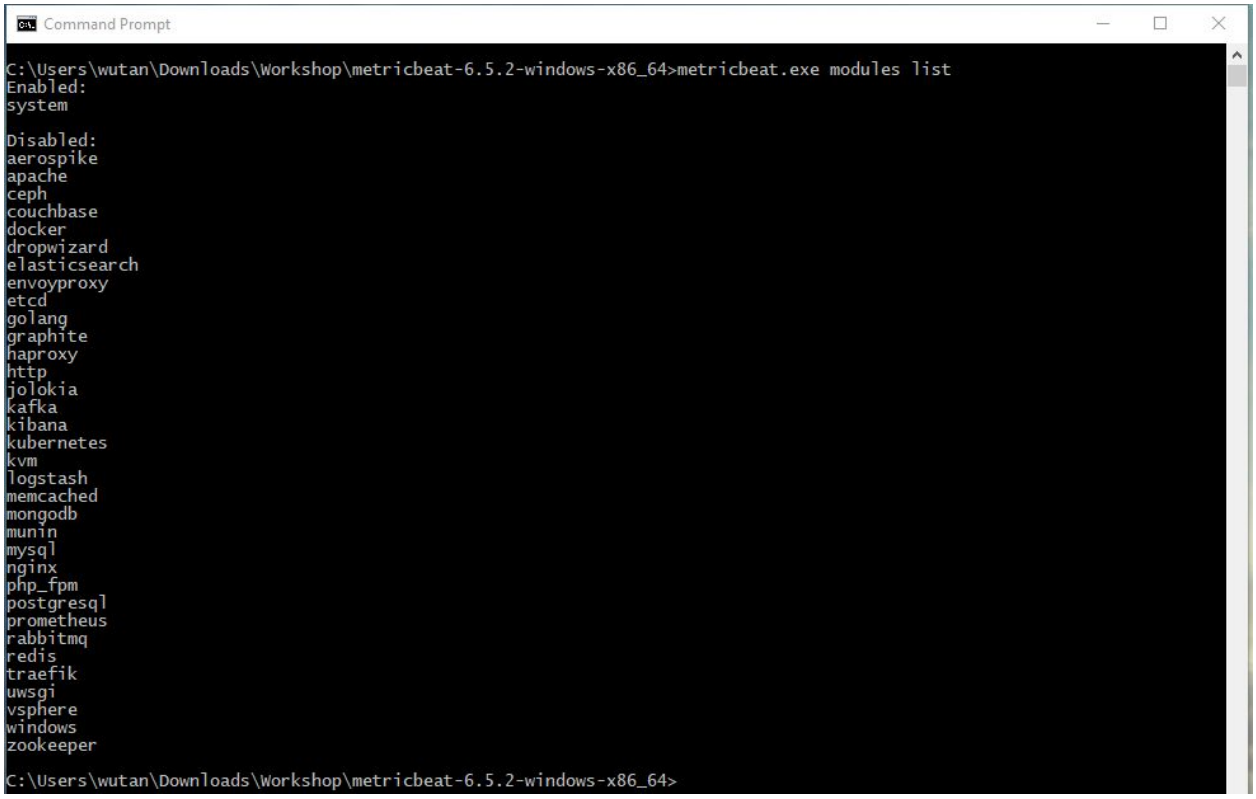
3) Open up a command prompt and type in cd + a space. Now drag-n-drop the extracted folder from step #2 to the command prompt window. Notice how it filled out the full path for you in the command prompt window? You can also type out the full path if you are a glutton for pain. Hit enter.



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.



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`

   Use your favorite text editor to open `metricbeat.yml` and replace **cloud.id** and **cloud.auth** with values obtained from Lab 0.

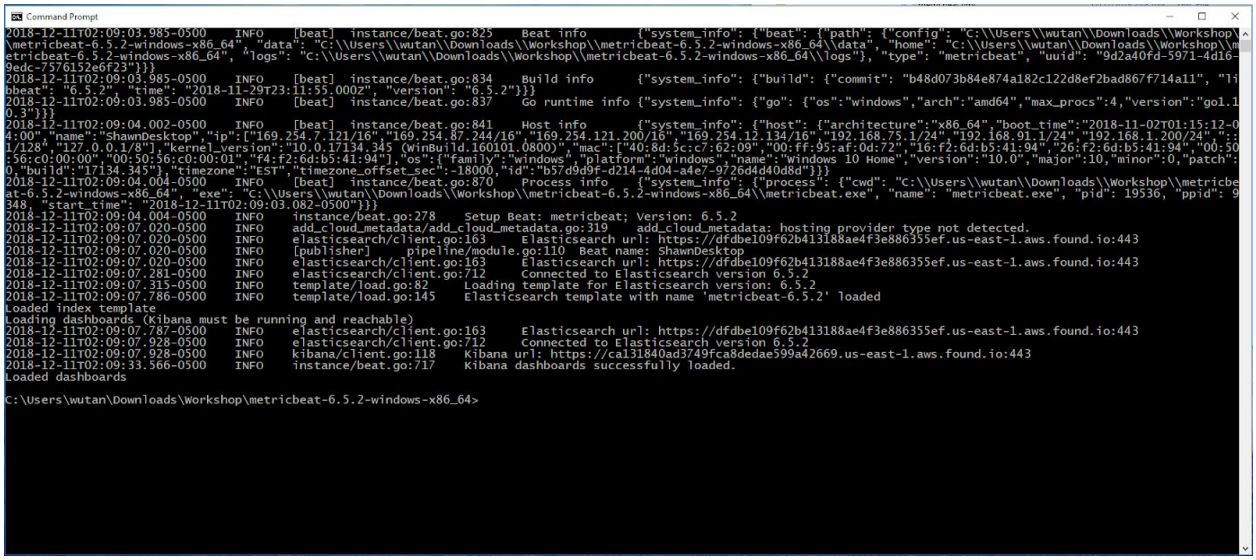
YAML files don't like hard tabs. Do not use them if you are editing a .yml file because they will cause errors. To learn more about .yml files see this link: https://en.wikipedia.org/wiki/YAML

Example:

6) Now we are ready to setup Elasticsearch to receive the system metric data, visualize it, and create Machine Learning jobs to detect anomalies. Fortunately it only takes one command!

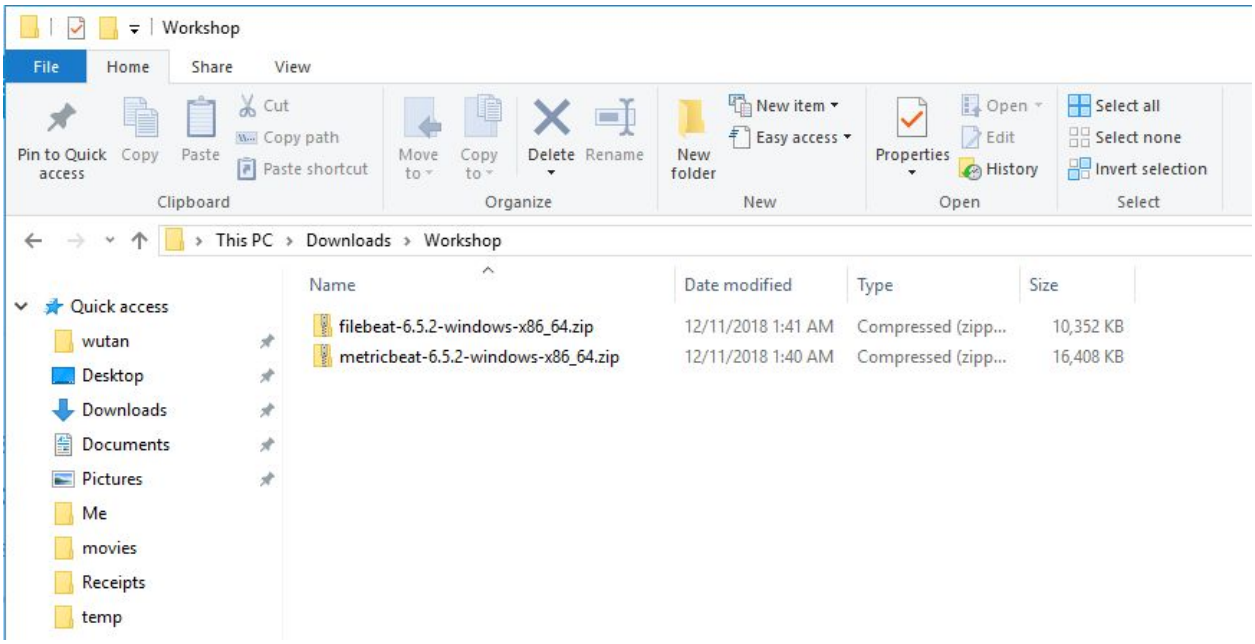
```
metricbeat.exe -e setup system
metricbeat.exe -e
```

Make sure you see the following text at the end of the output:

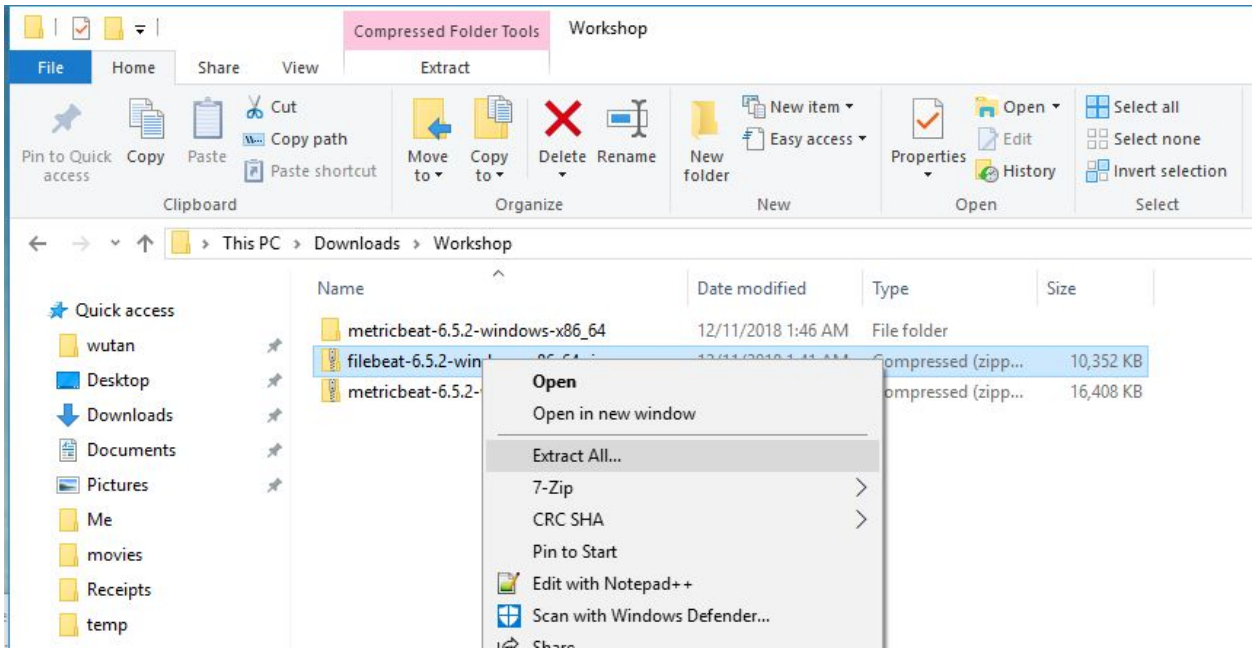


Filebeat

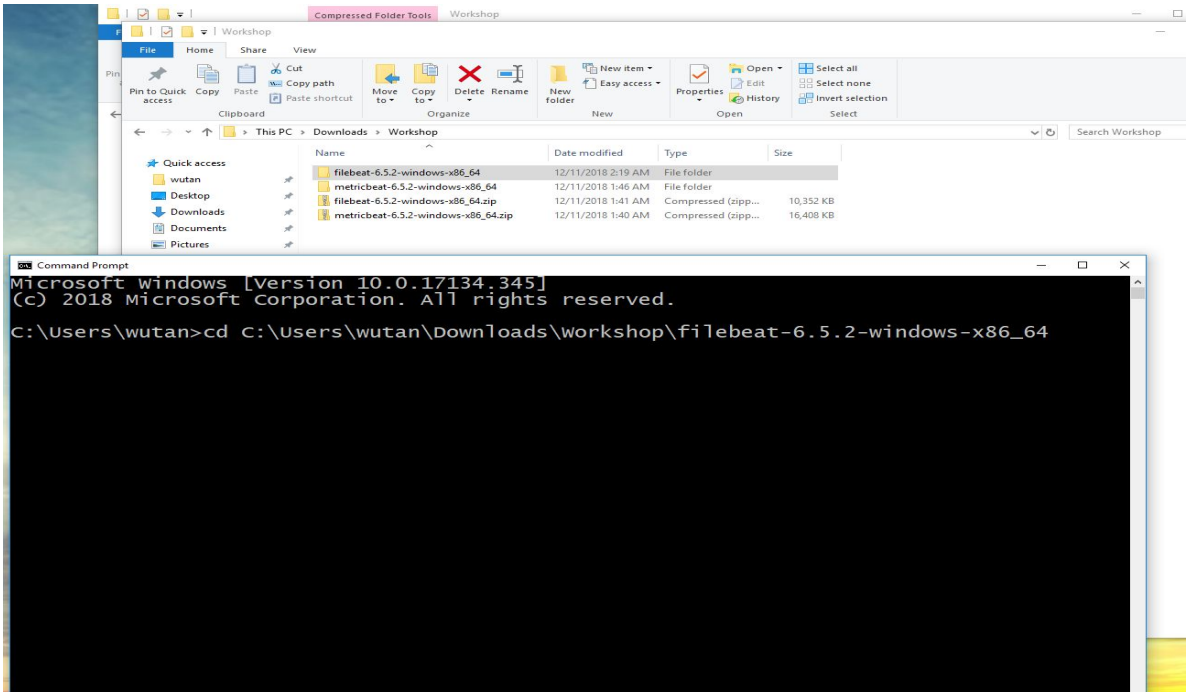
1) Open a Windows Explorer and navigate to the location that you downloaded metricbeat to



2) Expand the file that you downloaded



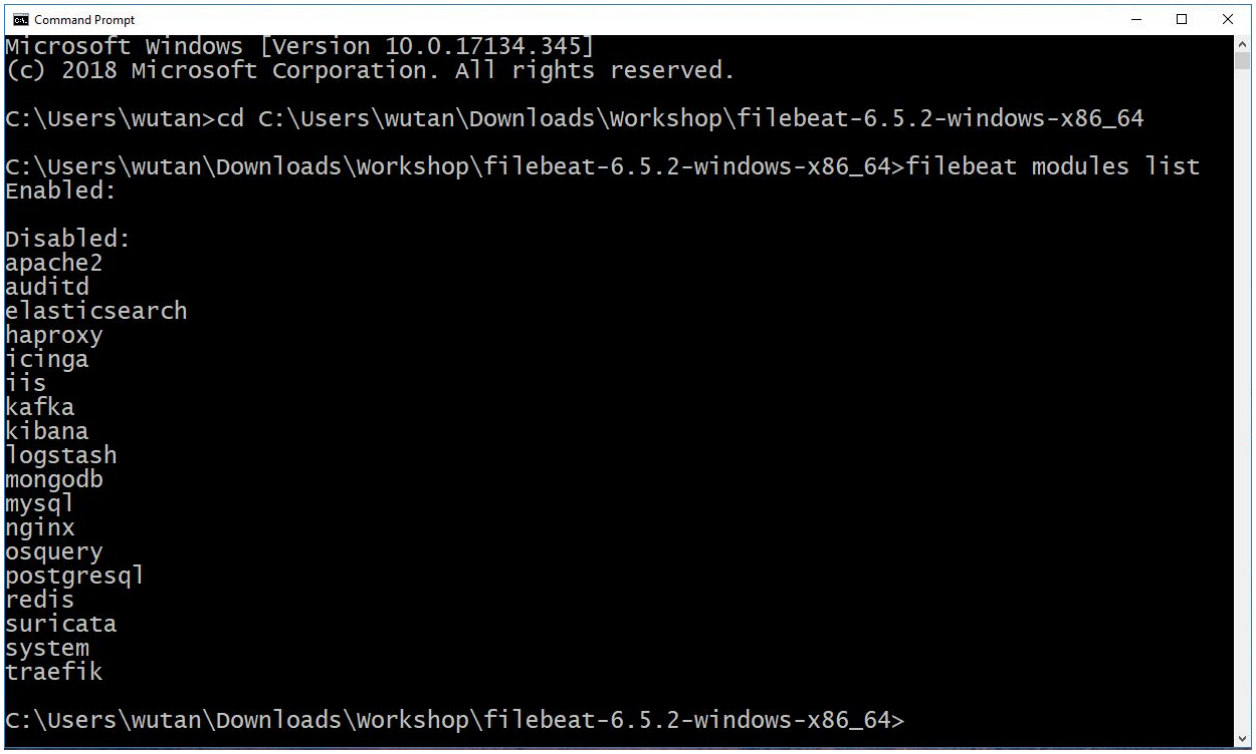
3) Open up a command prompt and type in cd + a space. Now drag-n-drop the extracted folder from step #2 to the command prompt window. Notice how it filled out the full path for you in the command prompt window? You can also type out the full path if you are a glutton for pain. Hit enter.



4) List the modules that are available

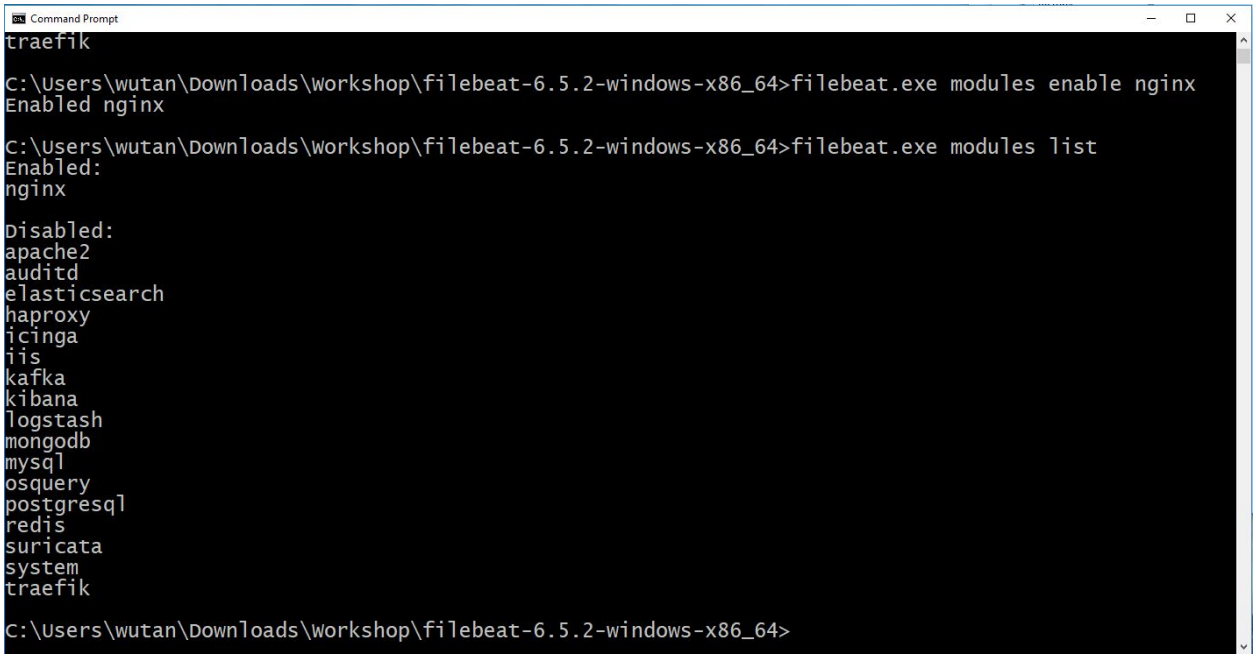
```
filebeat.exe modules list
```

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
```



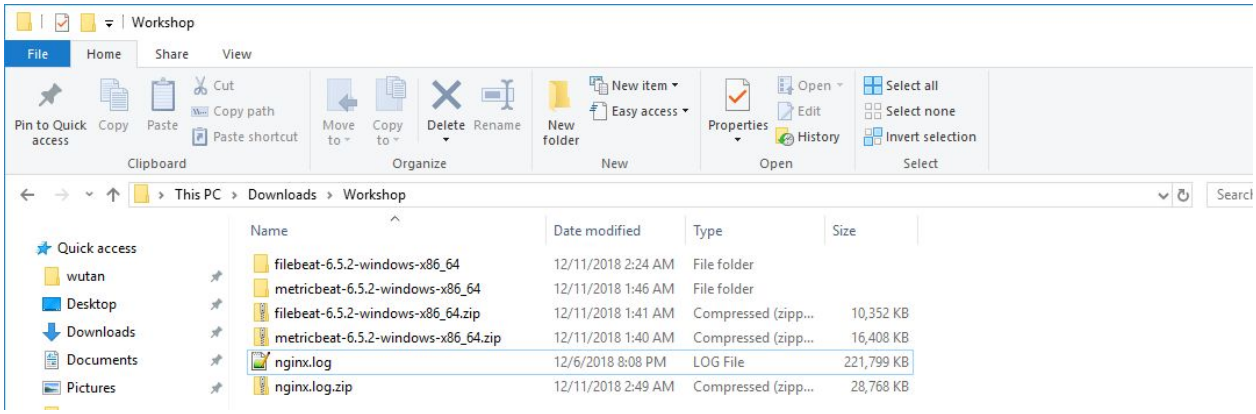
6) Before we setup Elasticsearch to accept NGINX logs we need to tell Filebeat where Elasticsearch is and provide credentials to login. We do this by editing the configuration file for Filebeat called `filebeat.yml`.

7) Follow the exact same procedure that you did in step #5 when you setup Metricbeat and add the **cloud.id** and **cloud.auth** to `filebeat.yml` using the values from Lab 0

8) Normally Filebeat would scan your system in several common locations for NGINX log files. 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:

https://drive.google.com/file/d/1RUDsDI5WOkVAnLw1xRR3H9zX3sIqAHt_/view?usp=sharing



9) Change directory to the following location

```
cd modules.d
```

10) Now let us modify the NGINX module to point to the log file location. We do this by modifying the `nginx.yml` file and creating an entry for the access logs. Add the following configuration to the `nginx.yaml` file.

```
- 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. Fortunately it only takes one command!

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

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`.

   Use your favorite text editor to open `metricbeat.yml` and replace **cloud.id** and **cloud.auth** with values obtained from Lab 0.

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

Example:

```
#cloud.id:
```

to

```
cloud.id:
"testcluster:dXMtZWFzdC0xLmF3cy5mb3VuZC5pbyRkZmRiZTEwOWY2MmI0MTMxODhh
ZTRmM2U4ODYzNTVlZiRjYTEzMTg0MGFkMzc0OWZjYThkZWRhZTU5OWE0MjY2OQ=="
```

and

```
#cloud.auth:
```

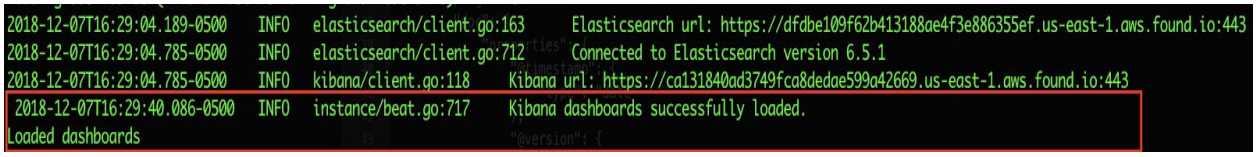
to

```
cloud.auth: "elastic:uB2AxBXuR1GMO0WwitEZ8VLt"
```

6) Now we are ready to setup Elasticsearch to receive the system metric data, visualize it, and create Machine Learning jobs to detect anomalies. Fortunately it only takes one command!

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

Make sure you see the following text at the end of the output:



7) Now run the metricbeat agent. Metrics from your local machine should now be flowing into Elasticsearch!

```
./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
```

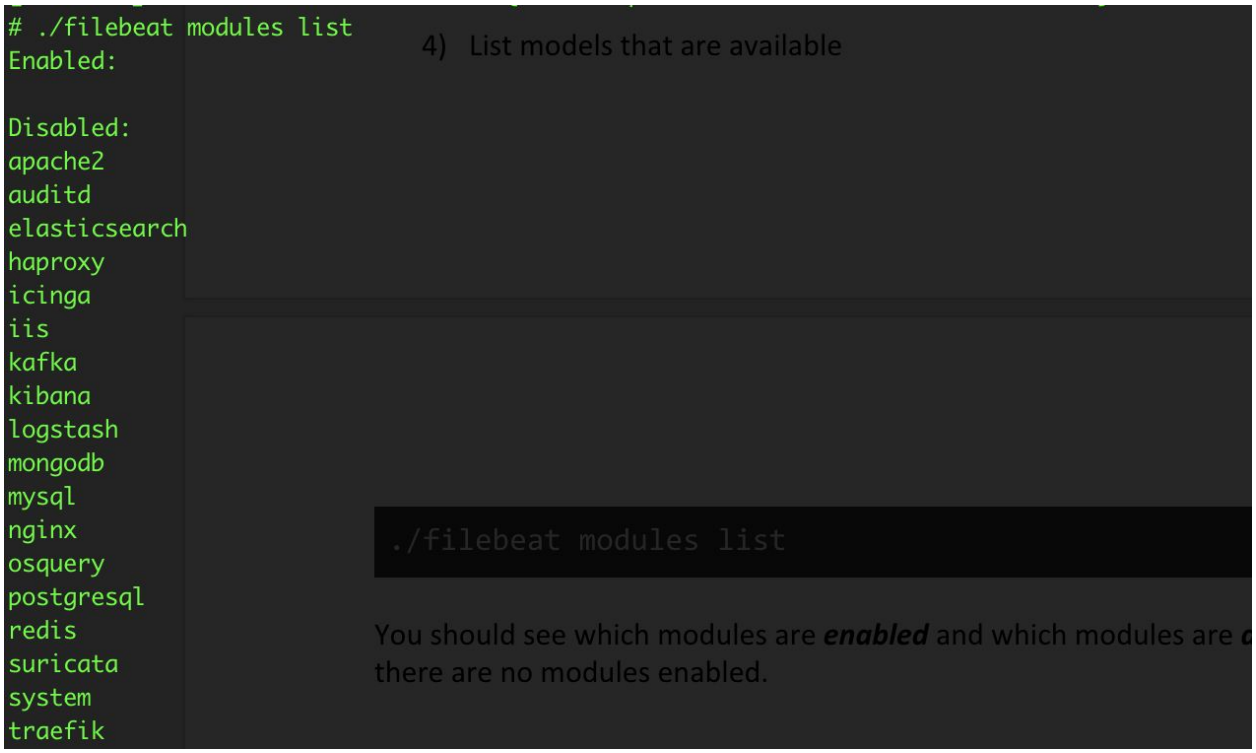
3) Change directory into the filebeat

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

4) List the modules that are available

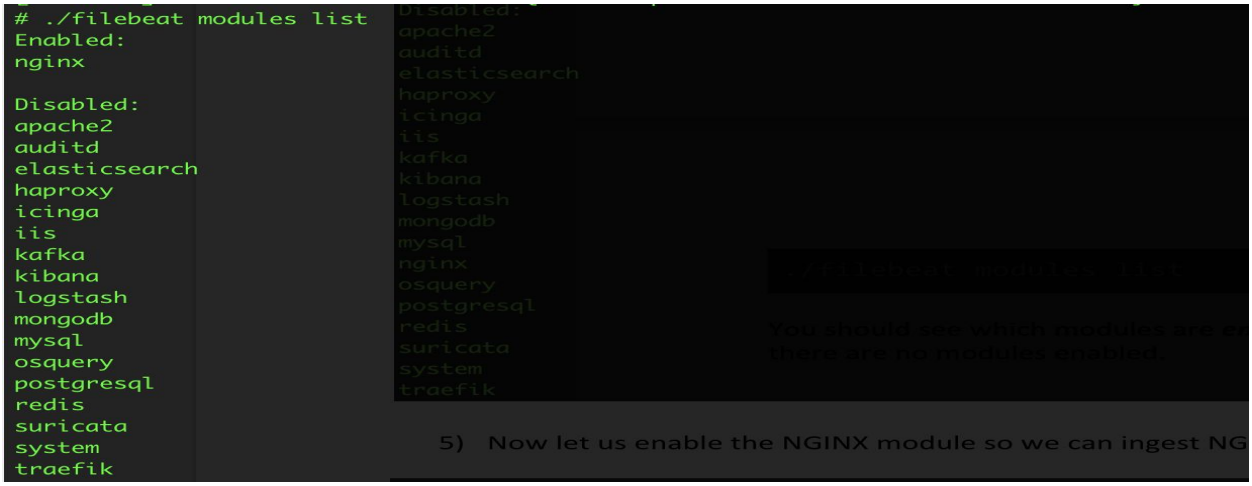
```
./filebeat modules list
```

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 that you did in step #6 when you setup Metricbeat and add the **cloud.id** and **cloud.auth** to `filebeat.yml` using the values from Lab 0

8) Normally Filebeat would scan your system in several common locations for NGINX log files. 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:

https://drive.google.com/file/d/1RUDsDI5WOkVAnLw1xRR3H9zX3sIqAHt_/view?usp=sharing

9) Change directory to the following location

```
cd modules.d
```

10) Now let us modify the NGINX module to point to the log file location. We do this by modifying the `nginx.yml` file and creating an entry for the access logs. Add the following configuration to the `nginx.yaml` file.

Note: Use the directory you expanded the NGINX log file to

```
- 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:
    var.paths: ["/Users/shh/Development/logs/nginx/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:
```

11) Now we are ready to setup Elasticsearch to receive the NGINX logs, visualize it, and create Machine Learning jobs to detect anomalies.

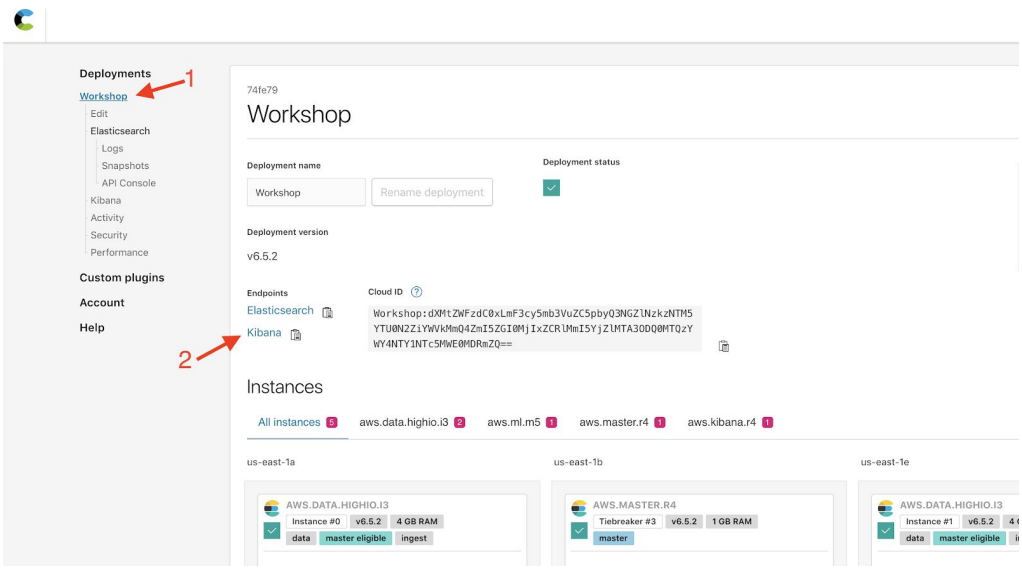
```
cd ..
./filebeat -e setup nginx
./filebeat -e
```

```
# ./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
```

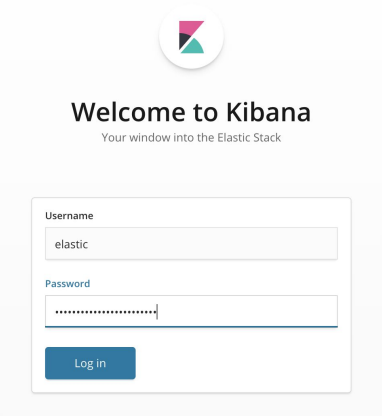
Validate Data in Kibana

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

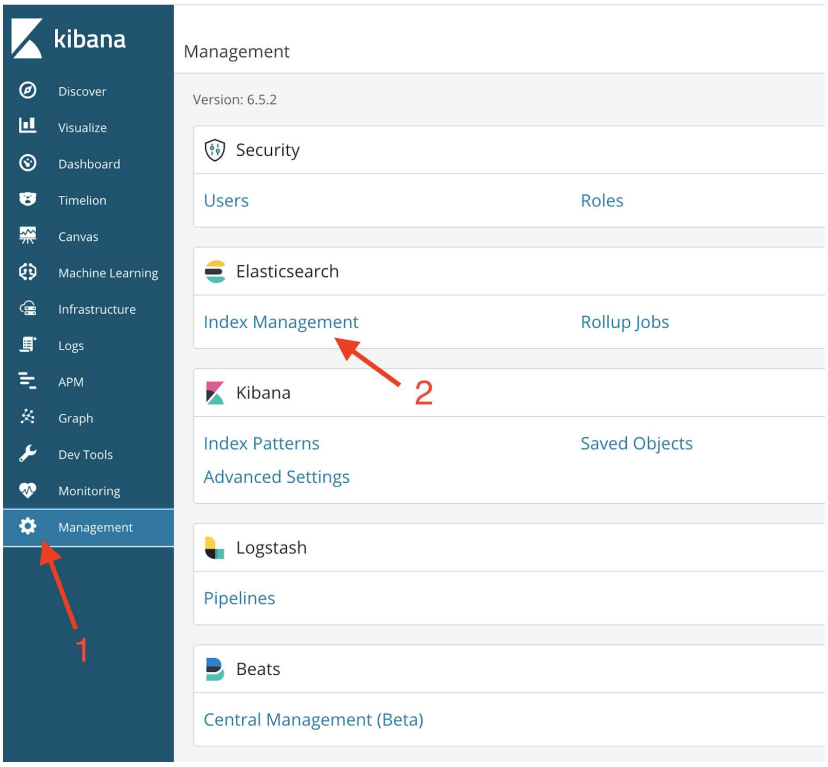
1) Log into your cloud console and click on the Kibana link



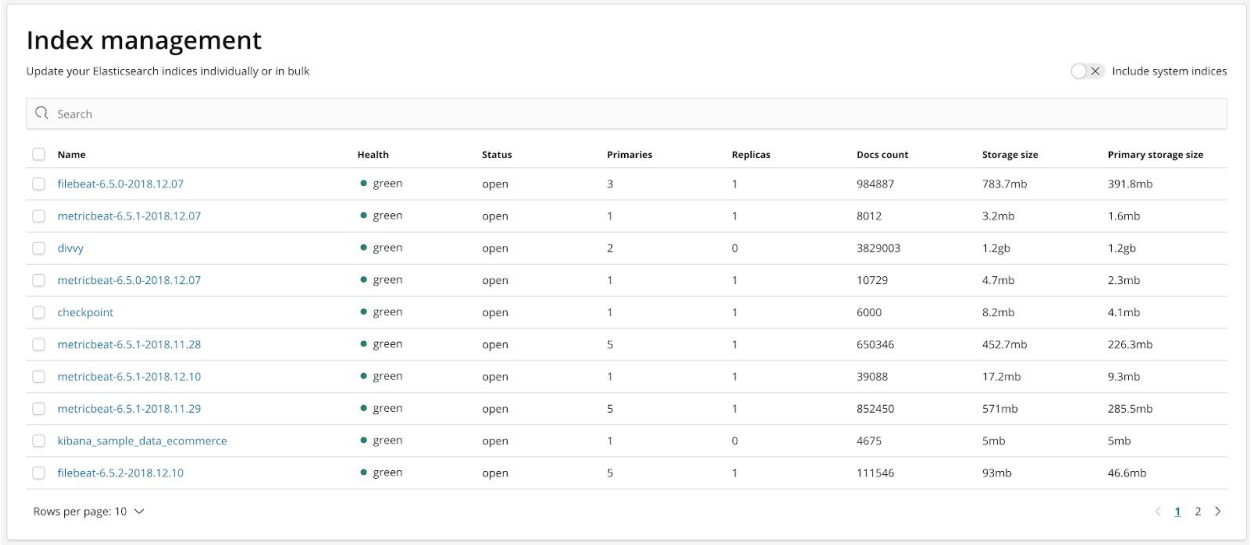
2) Log into Kibana with the credentials you obtained in [Lab 0](#).



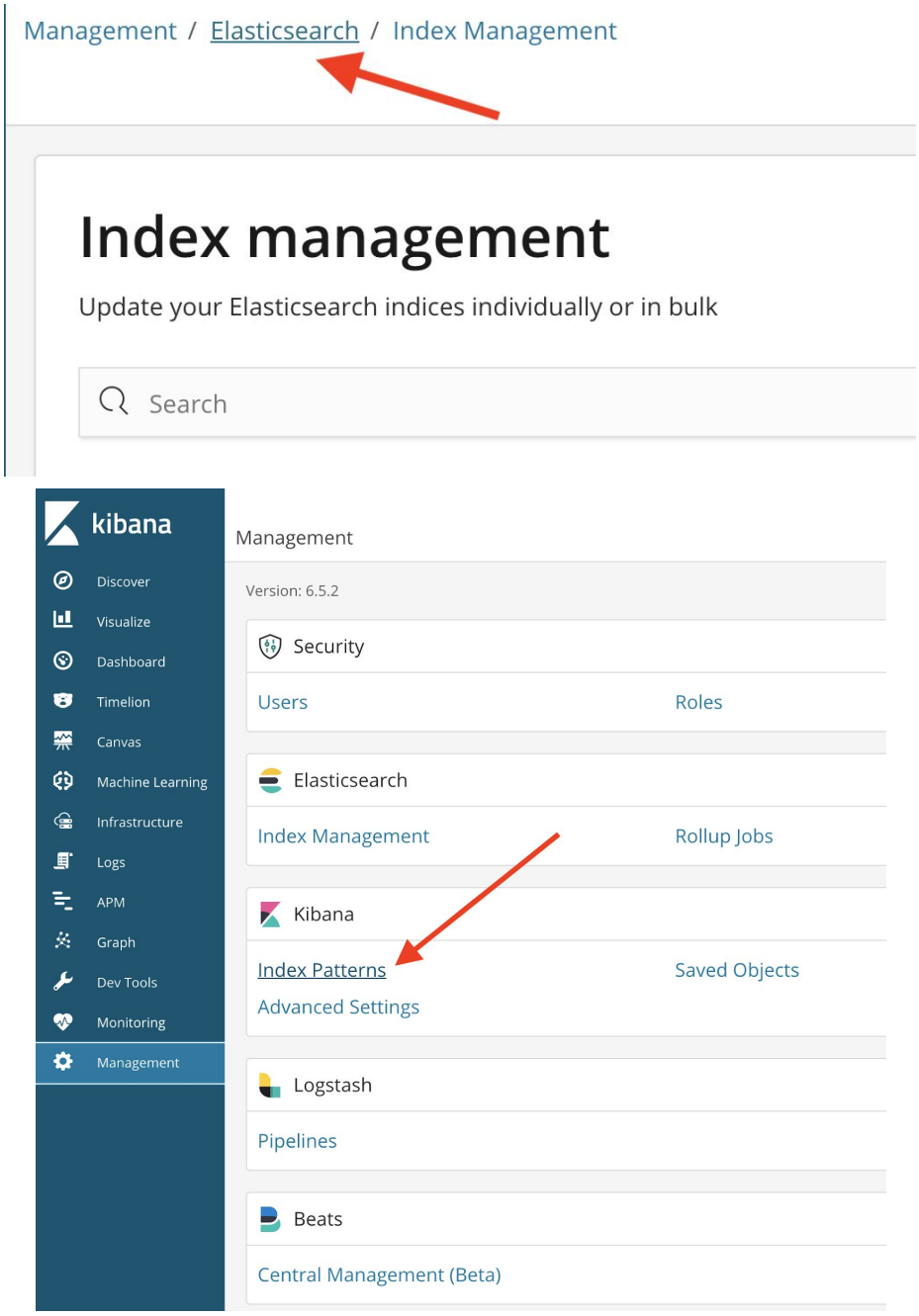
3) Click on the Management Link



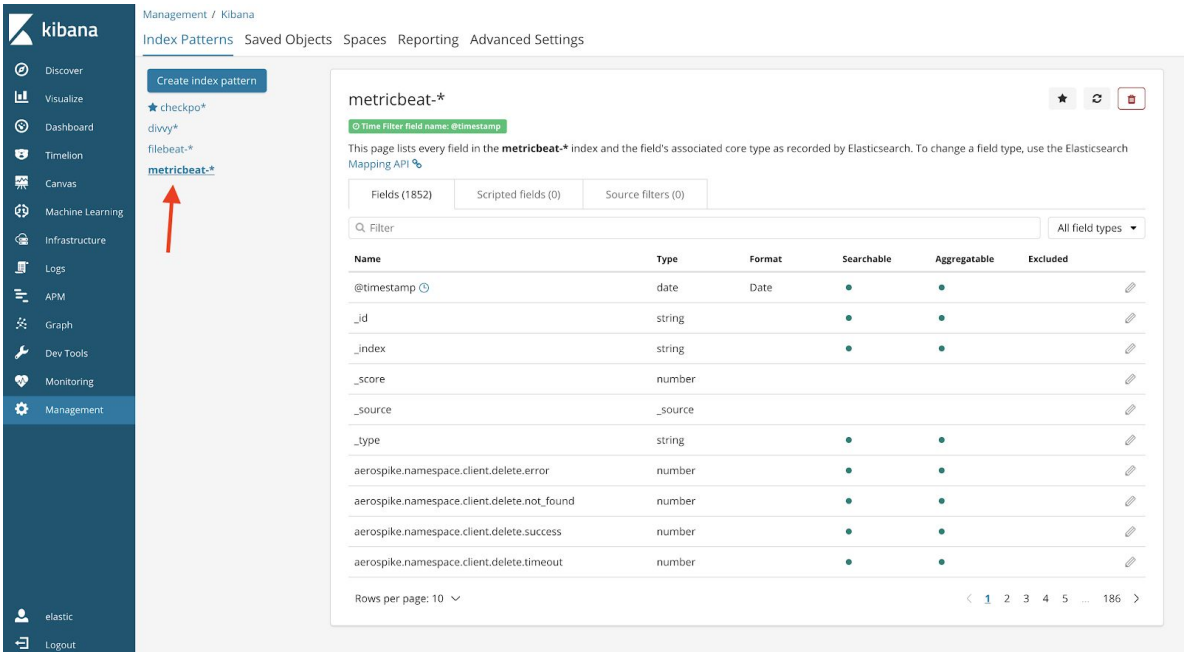
4) Look for Indexes named Metricbeat-<version>-YYYY.MM.DD where version is the current version of the product and YYYY, MM, and DD represent the year, month, and day respectively. Examine Docs Count, Storage Size, and Primary Storage Size. Realistic data like this provides a wonderful opportunity to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



5) Now click on the “elasticsearch” breadcrumb in the upper left hand corner of the screen and click on Index Patterns. 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.



6) 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.



7) Congratulations, you are now ready for Lab 2.