username: yorkfulltime2@exterbox.com user:elastic password: tmh6RZ5tBoQLEqbJyNtMT8PI
cloud_id:york_week_2:dXMtZWFzdC0xLmF3cy5mb3VuZC5pbyRjMWE3ODU2MGNhZTg0ODMwODMzZGR

## Schedule for this week

Day 1: Data Ingestion Day 2: Elasticsearch queries and analytics Day 3: Visualization and Project Day 4: Visualization and Project

## Resources

1. Demos
2. Ingesting CSV into Elastic Search
3. Spark and Elasticsearch

## Lab Exercise 1: Ingesting data into Elastic-search

At the end of this lab you will have:

1. Ingested gdelt events data using the logstash application.
2. Create a logstash configuration file to ingest csv file.
3. Use logstash to transform data before ingestion.
4. Perform basic queries on ingested data.

At the end of this lab exercise you would have created the architecture shown in the figure below:

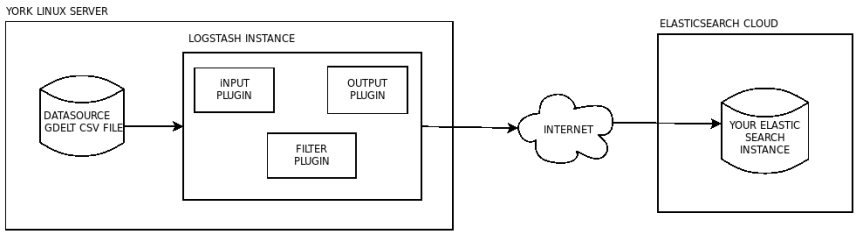


Figure 1: ingestion architecture

There are multiple ways to ingest data into elastic search. 1. Using Logstash 2. Writing code that uses the elasticsearch libraries: - see Python Library - and [Java] (https://github.com/elastic/elasticsearch-hadoop) - there are others

In this lab we will be using logstash to ingest our cleaned gdelt dataset: We will reuse your spark code to cleaned the data set and then ingest that clean dataset into elasticsearch .

**Part 1: Setting up your elastic cloud instance**

1. Signup for a 14 day elastic search trial here: 14 Day Trial

- Enter your email address and press start free trial.
- Then go to inbox and verify your account
- This will take you back to the create cluster page. Go ahead and create a cluster.

2. After creating the cluster be sure to save your user name and password.

You now have the completed setting up your elastic cloud instance.

## Part 2:

Before proceeding read the following introduction to understand what is logstash: Introduction to logstash

In this section we will be using logstash to ingest data into the elastic cloud instance you configure in part 1.

1. Follow the following instructions to login into the york server and open a terminal (provide instructions paul provided)
2. Once you have logged into the server and open a terminal. Enter the following command at the terminal prompt. The following video walks you through the steps:
3. Create a directory to store the files you will be working with. By entering the following command at the terminal prompt:

```
>> mkdir gdelt_ingestion
```

4. Go into the directory by entering the command and the terminal prompt:

```
>> cd gdelt_ingestion
```

5. Enter the follow command at the terminal prompt to download the events-logstash.config file located here: events-logstash-config

```
wget -c  https://raw.githubusercontent.com/LeotisBuchanan/yorkuniv/master/elasticsearch-and-
```

.