

Working with Influx

1) Capturing Statistics in Influx

- a) Using the example `/home/ubuntu/influx/influxTest.py` as a guide, create a Python program to record server information in an influx database named for your userid (e.g., myersjac).
- b) You will need to write a Python script that will take monitor the DB server and store information in Influx. You will be monitoring the amount of disk space used in your `/home/studentX` directory and capturing the user and the top process running on the server in terms of CPU. *** **You will be collecting data every 2 seconds over 1 minute.**

What you need to know/learn:

- Influx commands
- Python
- How to create visualizations in Grafana

```
Blocks: 388976; CPU leader: ubuntu-top
Blocks: 388972; CPU leader: ubuntu-top
Blocks: 388972; CPU leader: root-systemd
Blocks: 388972; CPU leader: neo4j-java
Blocks: 388972; CPU leader: root-systemd
Blocks: 384012; CPU leader: ubuntu-top
Blocks: 384012; CPU leader: root-systemd
Blocks: 384012; CPU leader: ubuntu-top
Blocks: 384012; CPU leader: root-systemd
Blocks: 384012; CPU leader: root-systemd
Blocks: 384012; CPU leader: root-systemd
```

- c) To get the number of blocks used in your directory use this Linux command:

```
du -c /home/<studentX> | grep total | sed 's/\ttotal//g' | tr -d '\n'
```

In order to see changes over time, while your Python script is running, in a separate terminal session you should be creating and deleting dummy files to alter your disk usage.

- d) To get the CPU leader information will take a bit more work. Use this Linux command:

```
top -b -d 5 -n 1 -o %CPU | head -n 8 | sed -n '$p'
```

This will yield a result that looks like this:

```
1 root      20   0 225496  5784  3204 S  0.0  0.6  0:22.71 systemd
```

To extract the required information, you would have to parse the string into an array of words and grab the 2nd word (e.g., root) and the last word (e.g., systemd). Think about how you should store the data. You will need to produce the Grafana charts below.

2) Grafana

- a) Using your browser, connect to the Grafana server.
- b) Make a data source for YOUR Influx database
- c) Using that data source, make the following Grafana dashboards
 - (i) The number of blocks used by your user id
 - (ii) The number of times per minute each user was the top CPU consumer
 - (iii) The number of times per minute each process was the top CPU consumer
- d) For the latter two graphs, you will have to understand the distinct values for user and process, and build multiple queries (one per process) identifying each query (like a series in Excep bargraphs) with an alias.

