Milestone 3

In this Milestone, I researched the solution to configuring the ELK Stack for our use case: to store tens of millions of sflow records a day. I created and modified setup files for each application in the stack. Furthermore, I created a configuration file for Logstash to collect the sflow packets, keeping up with the number of packets we receive per day, and perform an ASN lookup prior to sending the record to Elasticsearch. In Milestone 2, I had to write a script to lookup the ASN of a record and update the document to include the ASN. In this process, the ASN was being adding **after** the record was sent to Elasticsearch. After further testing, Jeronimo and I decided it would be best to have this done **before** the data is sent to elasticsearch to preserve disk space. After having completed this, I created an Elasticsearch index mapping for Logstash to utilize when creating indices so that Elasticsearch can properly map out each field's type, rather than storing everything as text. This will be useful for when we want to create visualizations in Kibana and need to perform operations, such as multiplying or adding two fields, on the data.

This milestone ended with a Bluejeans meeting where I presented the solution, my experience, with the ELK Stack, and what I have to work on going forward. The slides for the presentation can be found at: https://docs.google.com/presentation/d/1TSZnVDdHDiJ9YG5jZq4v4vVM7U4Z4Th4Ip_EN6XvSGc/edit?usp=sharing