BIG DATA CYBERSECURITY

LAB 7 METRON DASHBOARD

CREATING METRON DASHBOARD WITH KIBANA

Lab Description: Cybersecurity dashboards were introduced to the industrial use recently—Honeywell Process Solutions launched the first digital dashboard in 2015 for proactive monitoring, measuring and managing cybersecurity risks for control systems. This tool aimed at providing real-time visibility, understanding and support for informed decision making and action. With cybersecurity dashboards, customers don't have to be cybersecurity experts to be informed on managing risks, vulnerabilities and threats, such as rogue devices, malware, intrusion attempts, insecure configurations and such [1].

Initially, dashboard meant a screen on the front of a horse-drawn vehicle intended to obstruct water, mud or snow [2]. Later, it was used in the automotive world for displaying readouts and controls for multiple systems at a glance of the driver [3].

Few [4] defines a dashboard as a visual display of the most important information needed to achieve one or more objectives; consolidated and arranged on a single screen so the information can be monitored at a glance. Commonly, cybersecurity dashboards provide with remote monitoring of security incident logs and help answering "what," "when," and "where" of cybersecurity responding to the call for better reports and visuals in cyber risk management [5-7].

In this assignment, we will visualize Squid data.

Lab Files that are Needed: metron-cluster.ppk.

LAB EXERCISE/STEP 1

In the AWS cloud, create a new Metron virtual machine (VM) comprising a single-node Hadoop cluster as described in Lab 1. Establish an SSH connection with the VM as specified in Steps 1 through 4 of Lab 2. Note: it is recommended setting up a new VM to avoid misconfiguration because of the settings kept from other assignments.

Switch to superuser and change directory to /root:

```
sudo su
```

```
cd ~
```

Download a Squid proxy access log combined from several sources and stores at SecRepo.com, a repository of sample cybersecurity data.

```
wget https://www.secrepo.com/squid/access.log.gz
```

Unzip the compressed file with gunzip:

```
gunzip access.log.gz
```

View the content of the access.log file using the vi text editor.

Make a screenshot and store it in the submission document.

Notice that the first column in the access.log contains a timestamp in Unix time, also known as POSIX time or Unix Epoch time, which is a number of seconds elapsed after the beginning of day January 1st, 1970 Coordinated Universal Time (UTC) time standard minus leap seconds.

Copy a timestamp value from the first column and convert it to time in your time zone using online EpochConverter at https://www.epochconverter.com/ or another resource.

Make a screenshot and store it in the submission document.

LAB EXERCISE/STEP 2

Set required environment variables. For this, open the /etc/environment file for editing:

```
vi /etc/environment
```

Paste the following values:

METRON_HOME="/usr/metron/0.4.3"

ZOOKEEPER="metronserver.localdomain:2181"

ES_HOST="metronserver.localdomain:9200"

BROKERLIST="metronserver.localdomain:6667"

Press Escape, type :wq and press Enter to save and exit the file.

For this to take effect, close the terminal window, then start a new terminal and reconnect to your single-node cluster.

Verify that the variables have been set correctly:

```
sudo su

cd ~

echo $METRON_HOME

echo $ZOOKEEPER

echo $ES_HOST

echo $BROKERLIST
```

LAB EXERCISE/STEP 3

Pull Zookeeper configuration (single line):

```
$METRON_HOME/bin/zk_load_configs.sh --mode PULL -z $ZOOKEEPER -o $METRON_HOME/config/zookeeper/ -f
```

Create a Kafka topic to read Squid data with this command (single line):

```
/usr/hdp/current/kafka-broker/bin/kafka-topics.sh --zookeeper $ZOOKEEPER --create --topic squid --partitions 1 --replication-factor 1
```

Start the Squid parser:

```
$METRON_HOME/bin/start_parser_topology.sh -z $ZOOKEEPER -s squid
```

Push Zookeeper configuration (single line):

```
$METRON_HOME/bin/zk_load_configs.sh -m PUSH -i $METRON_HOME/config/zookeeper -z $ZOOKEEPER
```

Make a screenshot and store it in the submission document.

LAB EXERCISE/STEP 4

Elasticsearch creates a new index automatically when it receives new data. We will adjust index mapping in order to assign correct data types.

Elasticsearch tokenizes data fields of type string and performs additional processing to enable free-form text search. For the Squid data, we need to use several fields as enumerations. Thus, these fields are marked as not analyzed in the index template. Execute the command below to adjust the Elasticsearch index template:

```
curl -XPOST http://metronserver.localdomain:9200/_template/squid_index
-d '
{
  "template":"squid_index*",
  "mappings":{
    "bro_doc":{
      "properties":{
        "timestamp":{
          "type":"date",
          "format":"epoch_millis"
        },
        "source:type":{
          "type":"string",
          "index":"not_analyzed"
        },
        "action":{
          "type":"string",
          "index":"not_analyzed"
        },
        "bytes":{
          "type":"integer"
        },
        "code":{
          "type":"string",
          "index":"not_analyzed"
        },
        "domain_without_subdomains":{
          "type":"string",
          "index":"not_analyzed"
        },
        "full_hostname":{
          "type":"string",
          "index":"not_analyzed"
        },
        "elapsed":{
          "type":"integer"
        },
        "method":{
```

```
          "type":"string",
          "index":"not_analyzed"
        },
        "ip_dst_addr":{
          "type":"string",
          "index":"not_analyzed"
        }
      }
    }
  }
}'
```

LAB EXERCISE/STEP 5

We are ready to feed the access.log into the Squid Kafka topic, which will pass the data to Elasticsearch for indexing (single-line command).

```
cat access.log | /usr/hdp/current/kafka-broker/bin/kafka-console-producer.sh --broker-list $BROKERLIST --topic squid
```

The file occupies approximately 200MB. It will take some time to feed it completely.

Make a screenshot and store it in the submission document.

LAB EXERCISE/STEP 6

Login to Ambari interface, which is available at http://vmIPaddress:8080, where vmpIPaddress must be substituted with public IPv4 of your EC2 virtual machine using admin as a username and a password. Using the left-hand-side menu, locate and click Elasticsearch.

Use the Quick Links menu to open the Elasticsearch Indexes page. It should show a Squid index and possible and error index, which would index lines not conforming with the index template. The page does not update itself. Press F5 or refresh the webpage using other means. This action will update the document count in the Squid index. The data should look similar to this:

health	status	index	uuid	pri	rep	docs.count	docs.deleted	store.size	pri.store.size
green	open	.kibana	NJNWn8PUQsGI3W08IfX1DA	1	0	2	0	11.5kb	11.5kb
yellow	open	error_index_2019.08.08.11	X09XrVUGTc2nTSmrcmw8SA	5	1	20489	0	69.8mb	69.8mb
yellow	open	squid_index_2019.08.08.11	ZmRn80lDQRSDGAYRI5Y4ZA	5	1	162148	26718	191mb	191mb

Make a screenshot and store it in the submission document.

LAB EXERCISE/STEP 7

It is time to access the Elasticsearch index with Squid data from Kibana, which is a tool for visualizing and custom analytics [8]. The index must be configured first.

Locate and click Kibana in the menu of Ambari services. In Quick Links click Metron UI, then click Discover in the left-hand-side menu. In the Index pattern box type squid*, choose timestamp as a Time Filter field name and click Create.

The new screen will list all fields in the squid* index and their data types as recorded by Elasticsearch.

In text field will prompt for the name of the index. Type `squid*` within the text field. Every hour or day, depending on the specific configuration,a new Squid index will be created. Using this pattern will match against all Squid indices for all time periods.

2. Click outside of that text box and wait for the 'Time-field name' input field to populate. Since there is only one timestamp in the index, thisshould default to a field called `timestamp`. If this does not happen simply choose the field `timestamp`.

3. Then click the 'Create' button.

Make a screenshot and store it in the submission document.

LAB EXERCISE/STEP 8

Before we start working on visualization, a saved search must be created in Kibana. Click Discover in the main menu. Then click the newly created squid* index pattern.

The data chart will be empty, since the default setting is to display the last fifteen minutes of indexed data. Most of the ingested data records in the access.log file we downloaded earlier were created on September 8th, 2006.

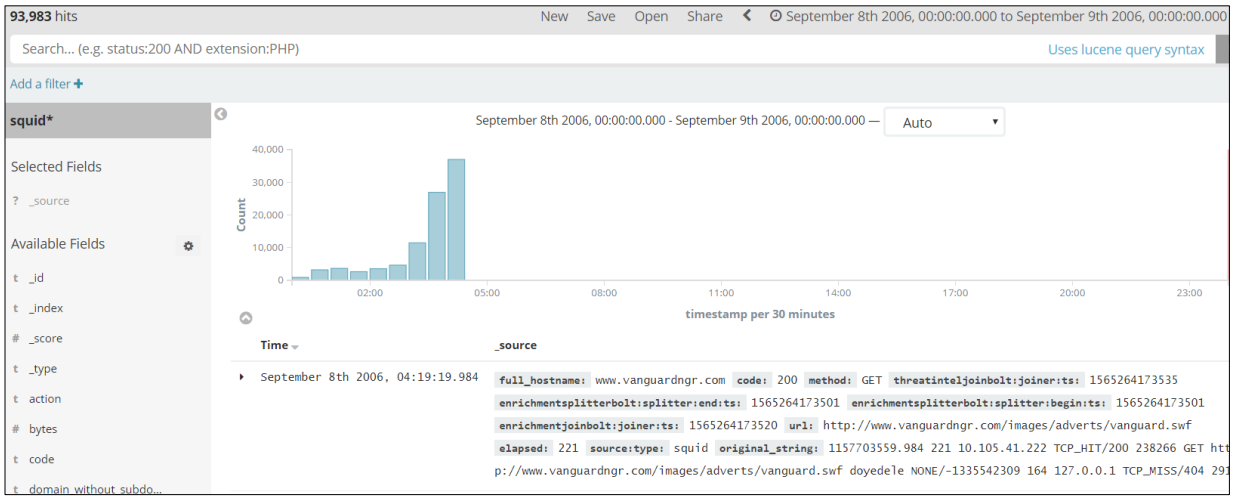
To adjust the time filter, click Last 15 minutes in the top right corner. A time filter panel will show up. Enter the start date as

September 8th, 2006 00:00:00.000

And the end date as

September 9th, 2006 00:00:00.000

And apply the filter. Kibana screen must be updated similar to this:

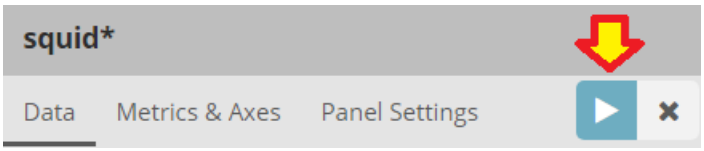


In the Fields panel on the left, add all fields to the saved search by clicking the Add button next to each field. Click 'Save' in the menu on the very top of the screen. Enter a name for the search and save it.

LAB EXERCISE/STEP 9

We will start working on data visualization. Click Visualize in the menu on the left. Then click Create a visualization → Vertical bar chart → From a saved search. Click the saved search name you entered at the previous step.

Under Select bucket types, click the X-Axis. choose Terms for the Aggregation type. For the Field box, choose domain_without_subdomains. Click the green Play button to apply changes.



In the top menu bar, click Save, assign a name to the visualization, and click Save once more. Notice, that moving a mouse pointer above the chart displays additional details.

Make a screenshot and store it in the submission document.

Modify the chart to display more domains. Click the green button with a triangle under the buckets label located on the left of the chart. It will display several options. Increase the value in the Size field to 15. Then click the Play button to apply changes. Click Save in the top menu bar.

Make a screenshot and store it in the submission document.

LAB EXERCISE/STEP 10

Create a dashboard. Click Dashboard in the Kibana menu. Click + Create a dashboard. Click Add in the top menu bar. Click the name of the visualization you have just created. It will appear in the dashboard. It is possible to adjust the size and location of the visualization.

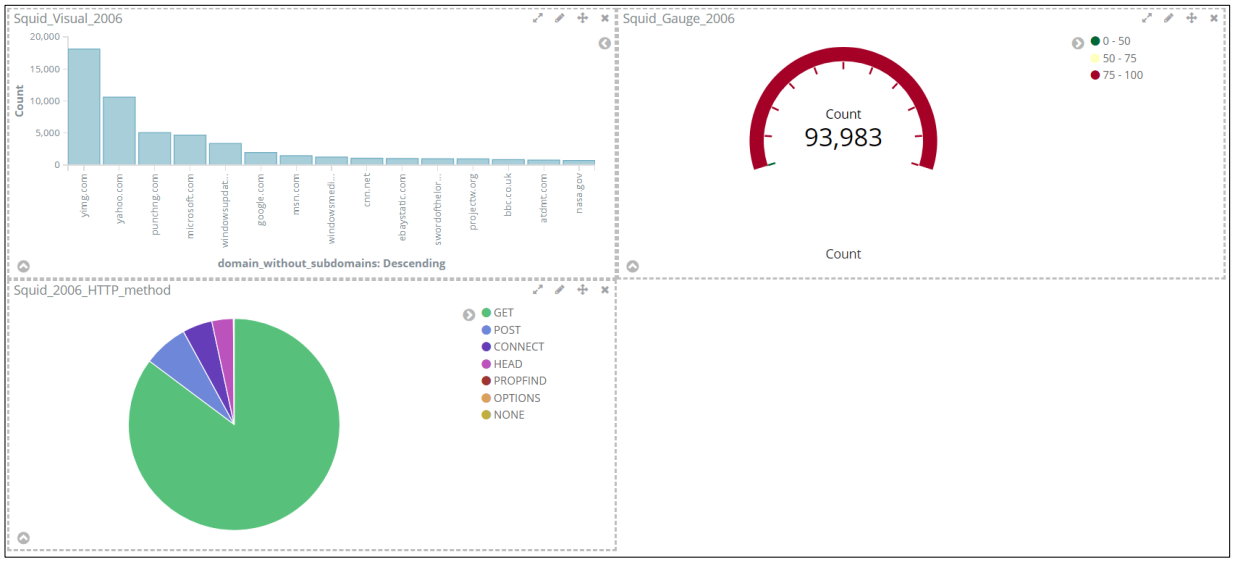
Make a screenshot and store it in the submission document.

Add more visualizations to this dashboard. Click Add in the top menu bar, then click the Add new visualization button located on the right side of the screen. In the Select visualization type screen, click Gauge. Click the name of the saved search. A new visualization showing the record count will appear. Click Save in the top menu bar, enter a name, and click Save and add to the dashboard.

Make a screenshot and store it in the submission document.

In a similar fashion, add the third visualization. Add an Area plot selecting Count as a Slice Size Aggregation, Terms as buckets split slices aggregation, Method as a field, metric:Count as Order By, enter 15 into the Size textbox. Then save it and add it to the dashboard.

At this point, the dashboard should look similar to this:



Make a screenshot and store it in the submission document.

LAB EXERCISE/STEP 11

Be sure to terminate your AWS Metron VM to avoid budget depletion. Resources created in a cloud environment under your account have associated costs. AWS offers more than 60 products at a free tier with associated free tier usage limits [9].

PUZZLER (VERDANA, 12 PT, BOLD, ALL CAPS)

Study threat hunting techniques for Squid proxy log files such as posted at Cyber Huntz https://www.cyberhuntz.com/2016/08/threat-hunting-techniques-av-proxy-dns.html. Come up with an idea for the fourth visualization and add it to the dashboard.

What to submit

Submit a Word (or other text editor) document with embedded screenshots made as requested in the assignment and a brief description for each screenshot.

References

[1] Oil & Gas Product News, "Honeywell Technology First to Proactively Manage Cyber Security Risk for Industrial Sites," [Online]. Available: https://www.oilandgasproductnews.com/article/20779/honeywell-technology-first-to-proactively-manage-cyber-security-risk-for-industrial-sites. [Accessed Aug 8, 2019].

[2] Merriam-Webster, "Dashboard noun," [Online]. Available: https://www.merriam-webster.com/dictionary/dashboard. [Accessed Aug 8, 2019].

[3] J. Engebretson. "What's a Security Dashboard?" *SDM*, vol. 40, no. 3, pp.77-80, 2010. http://search.proquest.com/docview/228467931/. [Accessed Aug 8, 2019].

[4] S. Few. *Information Dashboard Design*. North Sebastopol, CA: O'Reilly Media, 2006.

[5] J. Maier, A. Padmos, M. Bargh, and W. Wörndl, "Influence of mental models on the design of cyber security dashboards," In Proc. Int. Conf. on Information Visualization Theory and Applications, 2017, 128-139.

[6] V. Egeland, "LogWheels: A Security Log Visualizer," M.S. thesis, Norwegian University of Science and Technology, Trondheim, Norway, 2011.

[7] "Accliviti Security Dashboard Prevents Databreaches," *Computer Security Update*, vol. 16, no. 6, 2015.

[8] Elastic, "Kibana. Your window into the Elastic stack," [Online]. Available: https://www.elastic.co/products/kibana. [Accessed Aug 8, 2019].

[9] AWS, "AWS Free Tier," [Online]. Available: https://aws.amazon.com/free/. [Accessed Aug 8, 2019].