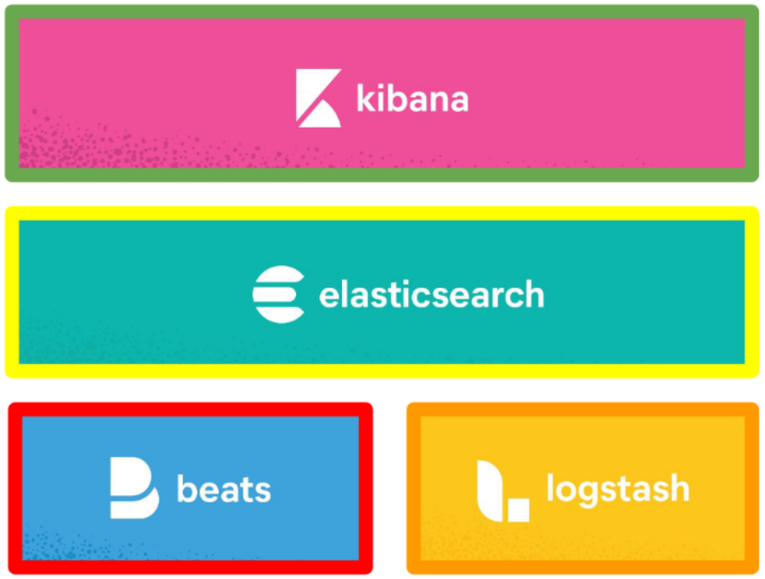
Elasticsearch

By Eric Ng, SE Intern

Introduction to Elasticsearch



01 The Elastic Stack

- Consists of ElasticSearch, Kibana and Logstash

02 ElasticSearch

- The heart of the Elastic Stack
- Used to Store, Search and Analyze data

03 Kibana

- Used to visualize and manage data

04 Logstash

- Used to ingest the data given to it, transforms it into a common format and stores it into the elastic

Understanding GROK

01

Example log:
[Wed Apr 05 04:15:50 2023] [error] [client 162.158.26.186] (70007)The timeout specified has expired: proxy: error reading status line from remote server aiv-app, referer: https://aiv.finexuscards.com/aivwebsdk/v1/selfie-video-review

Grok pattern:
\[%{DATA:timestamp}\] \[%{WORD:response}\] \[client %{IP:clientip}\] %{GREEDYDATA:message}

02 Example log (Stomp.log):

```
[INFO][2023-03-28 18:44:03,012][IDDocument.py:344]:[test][1bd9dac1fdfe4bdaa0494053bdb2212dCfOqUD
SUDSe6hyQ0CEvA]Ghost Match results: True, number of keypoints: 136, meanConfidence: 0.
.3950469195842743
[INFO][2023-03-28 18:44:03,014][StompUtils.py:168]:[test][1bd9dac1fdfe4bdaa0494053bdb2212dCfOqUD
SUDSe6hyQ0CEvA]Sending message to LANDMARK.OUT...
[INFO][2023-03-28 18:44:03,015][StompUtils.py:114]:[test][1bd9dac1fdfe4bdaa0494053bdb2212dCfOqUD
SUDSe6hyQ0CEvA]Message sent to LANDMARK.OUT
[INFO][2023-03-28 18:44:03,015][StompUtils.py:192]:[test][1bd9dac1fdfe4bdaa0494053bdb2212dCfOqUD
SUDSe6hyQ0CEvA]Message sent to LANDMARK.OUT. Processing time: 8.6240003108978 27. Message sent: {
'status': {'errorList': {'msgText': '', 'msgCode': ''}, 'status': 'SUCCESS'},
'myKad_frontLandmarkData': {'extractedData': {'myKadNo': '871231.144242\n\x0c', 'gender': 'F',
'fullName_address': 'A WT\nS\n\x0c', 'placeOfBirth': 'INVALID PLACE OF BIRTH', 'muslimFlg':
 False}, 'landmarkData': {'icPlateMatch': True, 'myKadLogoMatch': True, 'myFlagMatch': False,
'citizenFlg': False, 'ghostMatch': True}}, 'fakeImage': False, 'fakeImageConfidence': '0
.68996805'}
[INFO][2023-03-28 18:44:04,044][DatabaseUtils.py:24]:Successfully connected to database
[INFO][2023-03-28 18:44:04,116][DebuggingUtils.py:183]:[test][1bd9dac1fdfe4bdaa0494053bdb2212dCf
OqUDSUDSe6hyQ0CEvA]Saved debug_image(id:27300)
[INFO][2023-03-28 18:44:17,994][StompUtils.py:150]:[test][1bd9dac1fdfe4bdaa0494053bdb2212dCfOqUD
SUDSe6hyQ0CEvA]Received message
[INFO][2023-03-28 18:44:17,998][Request.py:71]:[test][1bd9dac1fdfe4bdaa0494053bdb2212dCfOqUDSUDS
e6hyQ0CEvA]Processing liveness ...
[INFO][2023-03-28 18:44:18,000][LivenessUtils.py:638]:[test][1bd9dac1fdfe4bdaa0494053bdb2212dCfO
qUDSUDSe6hyQ0CEvA]Message stream type: videoB64
```

Grok pattern:
\[%{LOGLEVEL}\]\[%{TIMESTAMP_ISO8601:timestamp}\].*Message sent to %{WORD:type}\.OUT\. Processing time: %{NUMBER:processing_time}\.

Key Benefits of the ELK Stack for Log Management and Analysis

01 Centralized Log Viewing

- ELK enables centralizing logs from viewing logs from multiple sources, It allows easy searches, filtering and efficient log analysis.

02 Visualization and Dashboards

- Kibana, the visualization tool, enables creation of interactive dashboards. Which allows you to freely choose how you would like to visualize your log data.

03 Scalability

- The ELK stack is highly scalable. It can handle growing log volumes and processing requirements by adding more Elasticsearch nodes to the cluster, ensuring efficient log management

04 Open-source and Community Support

- The ELK stack is open source which means we can leverage the stack's capabilities without incurring any costs.
- The ELK stack is backed by a very supportive community

05 Clients

- Gain insights into application performance, user behavior and system health

06 Developers

- Debug and monitor applications, track errors and improve overall performance

Understanding the Limitations of the ELK Stack

01 Security considerations

- By default, he ELK stack lacks robust security features. Additional measures such as implementing SSL/TLS encryption and configuring secure access are required to enhance the security of the ELK stack in production environments

02 Resource Requirements

- The ELK stack can be resource-intensive, specifically when it is dealing with large datasets.

03 Stability and Uptime Issues

- There has been reports of instability and uptime issues in the past. However, it is usually quickly fixed by the developers