

DATA ANALYSIS USING ELK STACK



F. MOHAMMED
Utech LLC location
– Madison

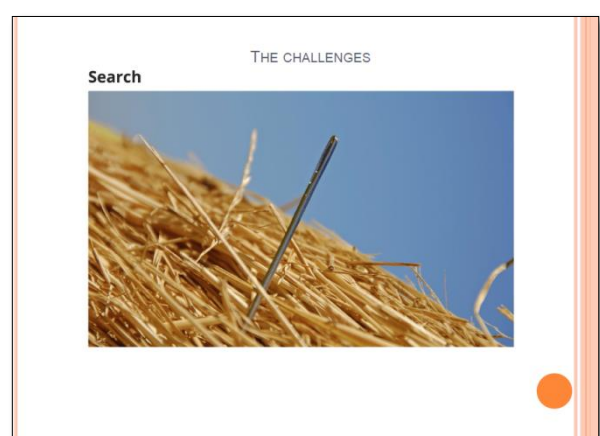
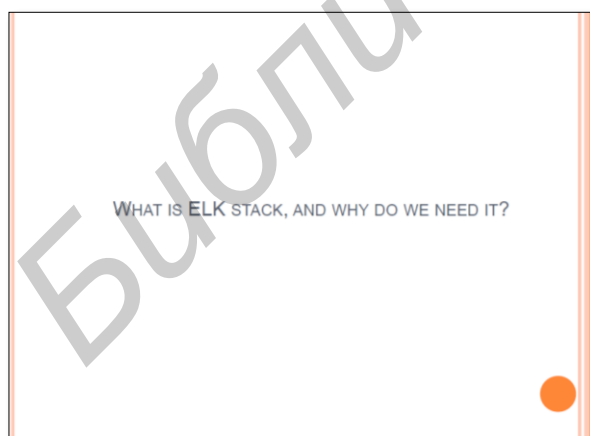
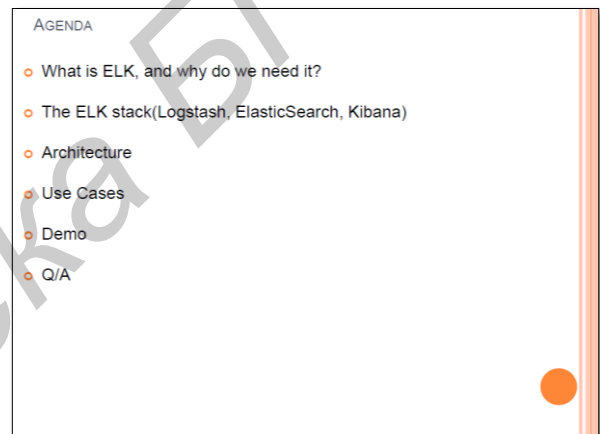
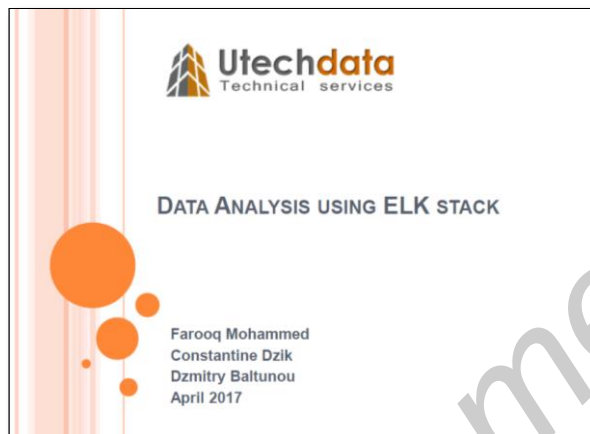


D. BALTUNOU
Utech LLC location
– Madison



C.S. DZIK
Utech LLC location
– Madison

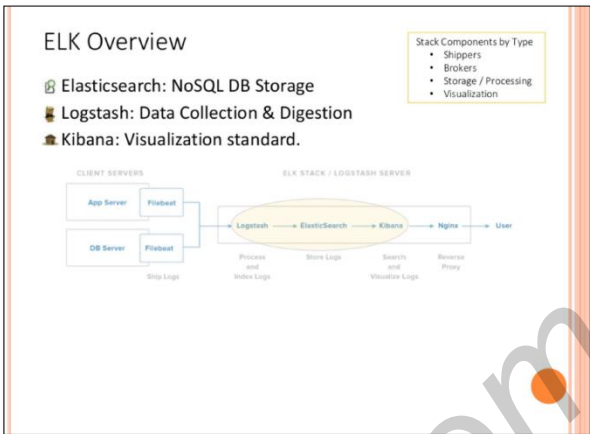
Utech Solution Inc, Madison, USA
E-mail: constantine.dzik@utechdata.com



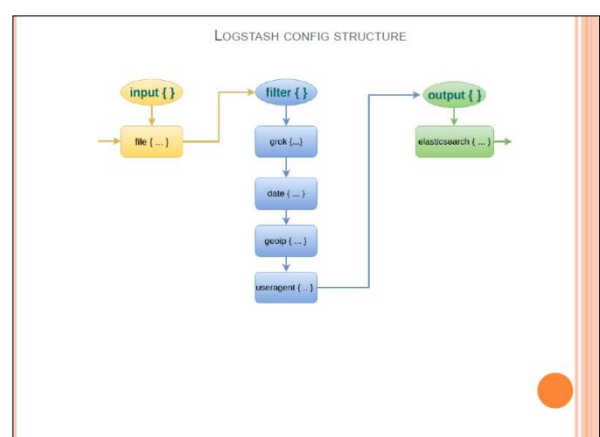
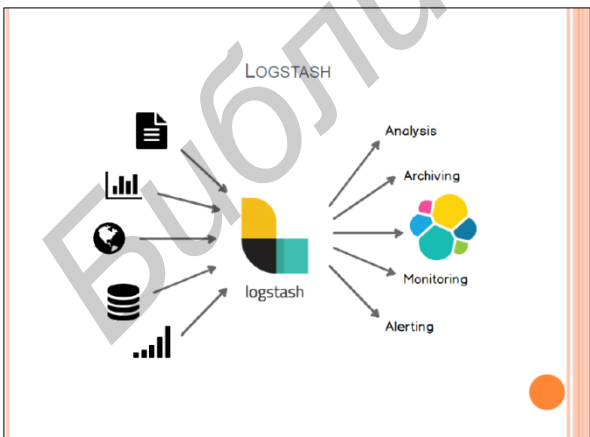


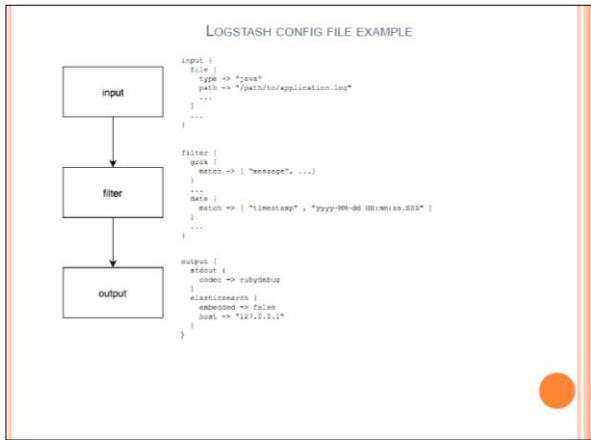
What is the ELK stack

- Elasticsearch
 - Search server
 - Based on Apache Lucene
- Logstash
 - Data pipeline
 - Processes logs and other data
 - Plugins
- Kibana
 - Web frontend for Elasticsearch

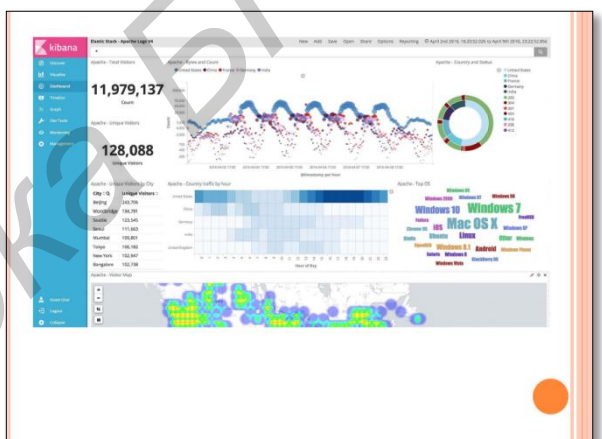


THE ELK STACK (LOGSTASH, ELASTICSEARCH, KIBANA)

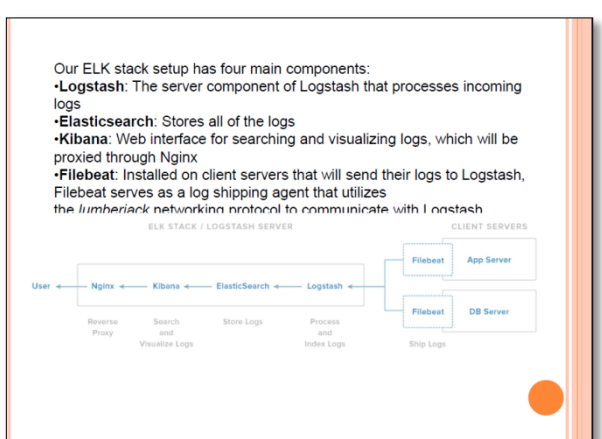




- KIBANA
- Kibana is an open source analytics and visualization platform designed to work with Elasticsearch
 - Use Kibana to search, view, and interact with data stored in Elasticsearch indices.
 - Easily perform advanced data analysis and visualize your data in a variety of charts, tables, and maps.
 - Easy to understand large volumes of data. Its simple, browser-based interface enables you to quickly create and share dynamic dashboards that display changes to Elasticsearch queries in real time.



ARCHITECTURE



USE CASES



USAA reduces security incidents by searching 3-4 billion security events a day, running Python scripts, building custom applications to mine the data, and utilizing Watcher, the Elasticsearch alerting and notification extension.

FICO

FICO

FICO is leveraging Elasticsearch, unstructured and semistructured data to significantly improve the performance of FICO's analytics models. FICO's Analytic Modeler for Text products makes these advanced analytics and visualizations available to end users in an intuitive and interactive way. FICO has integrated advanced descriptive, diagnostic, and predictive analytics with Elasticsearch, and extended Kibana to provide advanced visualizations against same.

NETFLIX

Netflix

Netflix messages millions of customers a day across many channels – email, push notifications, text, voice calls, etc – via its messaging platform: a distributed system made up of a series of separate applications. They use Elasticsearch for higher message deliverability and operational excellence.

GitHub

GitHub

GitHub uses Elasticsearch to continually index the data from an ever-growing store of over 8 million code repositories, comprising over 2 billion documents. Using Elasticsearch, GitHub was able to let users easily search this data. One goal of GitHub's Elasticsearch implementation is to index everything that is publicly available on GitHub.com and make it easy to find. Full-text searching is fully supported, but searching based on a wide variety of criteria is also possible and dead simple. On GitHub you can search for a project that uses Clojure as the primary language, and has had activity over the past month, and all this functionality is powered by Elasticsearch.



Facebook

Facebook has been using Elasticsearch for 3 plus years, having gone from a simple enterprise search to over 40 tools across multiple clusters with 60+ million queries a day and growing.

