

Topics

Introduction to Big Data and Hadoop

Theory

- Introduction to Big Data and prime techniques to handle/process it
- MapReduce algorithm
- Hadoop Framework, it's component and related technologies (e.g. HFDS, HIVE, HBase etc.)
- Run through of one or two real use-cases of Hadoop in the industry
- Step-by-step simple word count problem explained in detail

Lab Sessions

- Implementation of word count problem
- **(and if time permits)** Dictionary translation (or log processing)

NoSQL(Elasticsearch) and Real Time Analytics

Theory

- Big Data Analytics and NoSQL Storage
- Few associated technologies
- Introduction to - Elasticsearch, Logstash and Kibana
- End-2-end data flow with a real use case explained
- Data ingestion and data access mechanism/techniques
- Analytics using Kibana

Lab Sessions

- Few hands-on example of data insertion and data access from Elasticsearch
- Prepare Analytics for Apple Stock Trends
- **(and if time permits)** New York City Traffic Incidents navigation