Science Hack 2018

BASF DATA 2

TIMESERIES HISTORIAN

Balachandra

Team Delta

Nitish

Pallavi
Problem Statement

- More than a billion data points per second
- Need for storing and retrieving data in a reliable and efficient manner
Solution Framework

- Elasticsearch
- Kibana
- HTTP REST API
- HTTP Post
- Sensors
- Sensors
- Sensors
- ......

Diagram:

- Elasticsearch → Kibana
- HTTP POST → HTTP REST API
- Sensors → HTTP REST API
- Sensors → HTTP REST API
- Sensors → HTTP REST API
Tools and their justification

**Elasticsearch:**
- Apache Lucence based search engine.
- Open source & developed using Java

**Kibana**
- Data exploration & visualization tool
- Used for log & time series analytics, application monitoring & operational intelligence.

**Python**
- Used for Sensor emulation.
Framework Capabilities

• Offline capability through database setup on PC
• Created replicas of sensor data
• Generated automatic indexing and query distribution across the cluster
Challenges

• Limited data points
• Minimizing risk of data loss
• Minimizing storage consumption
• Working with multiple nodes
Sensor data emulation
Data Retrieval
Data visualization
Thank you