

CSYE7200 Big-Data Sys Engr Using Scala

Analyzing Ethereum Transactions

Team 3

Charmi Nimishkumar Dalal

Porom Sivam Kalita

Ronit Mankad



1

USERS ACCESSES LIVE
DASHBOARD TO CHECK
THEIR CUSTOMER
TRANSACTIONS IN OUR
SYSTEM



2

USER CAN STUDY
DASHBOARDS AND
FIND LATEST TREND IN
THE ETHEREUM
MARKET



3

USER CAN MODIFY
DASHBOARD DESIGN
AND CHANGE
REPRESENTATION OF
DATA

Use Cases

Methodology

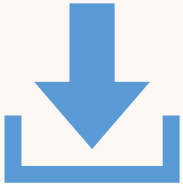
Purpose: To analyze the full set of transactions which have occurred on the Ethereum network from the first transactions 2016 till 2019 using Big data tools such as Kafka and Spark.

- Download chunk of dataset from Bigquery and store locally.
- Use Spark-Streaming API to read the CSV file and write to Kafka topic as stream of records
- Use Spark-Streaming API to consume data stream from the Kafka Topic and load it as a data frame.
- Store the transformed data into MySQL and use Apache Superset to create visualizations of relevant data

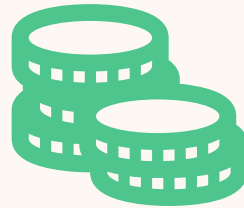
CSV file (Local system) -> Spark (Producer) -> Kafka -> Spark (Consumer) -> MySQL -> Apache Superset



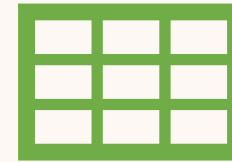
Data Source



Big query crypto currency Ethereum dumps daily. We download it in csv format and save it locally.



It has 9 tables such as transactions, logs, contracts, tokens.



Number of rows are approximately 30,000 per data chunk

Milestones/Sprints

Project

Project Set Up (1 Apr – 7 Apr)



Write

Write Consumer & Producer Logic (8 Apr – 14 Apr)



Perform

Perform data transformation according to need of analysis (15 Apr – 21 Apr)



Create

Create dashboards to visualize data (22 Apr – 27 Apr)

What will
you program
in Scala?

Kafka Topic
configuration

Kafka Producer
for each table

Kafka
Consumer for
each table

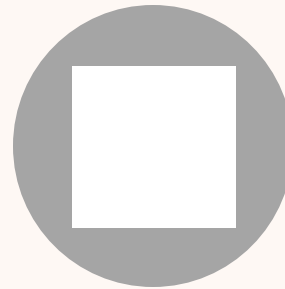
Data
transformation
logic

Store
transformed
data to MySQL

Acceptance criteria



Live dashboard
configured to pull the
data every 10 seconds



Data sent to the kafka
topic loads it into a data
frame every second.



Goals of the project

- **Time Analysis** - Number of transactions occurring every month from Feb 2016 to June 2019.
 - **Top 10 Most Popular Services** - Evaluation of top 10 smart contracts by total Ether received.
 - **How Gas price (Cost in Crypto) is changed over the time**
 - Miscellaneous Analysis
- 



Thank You