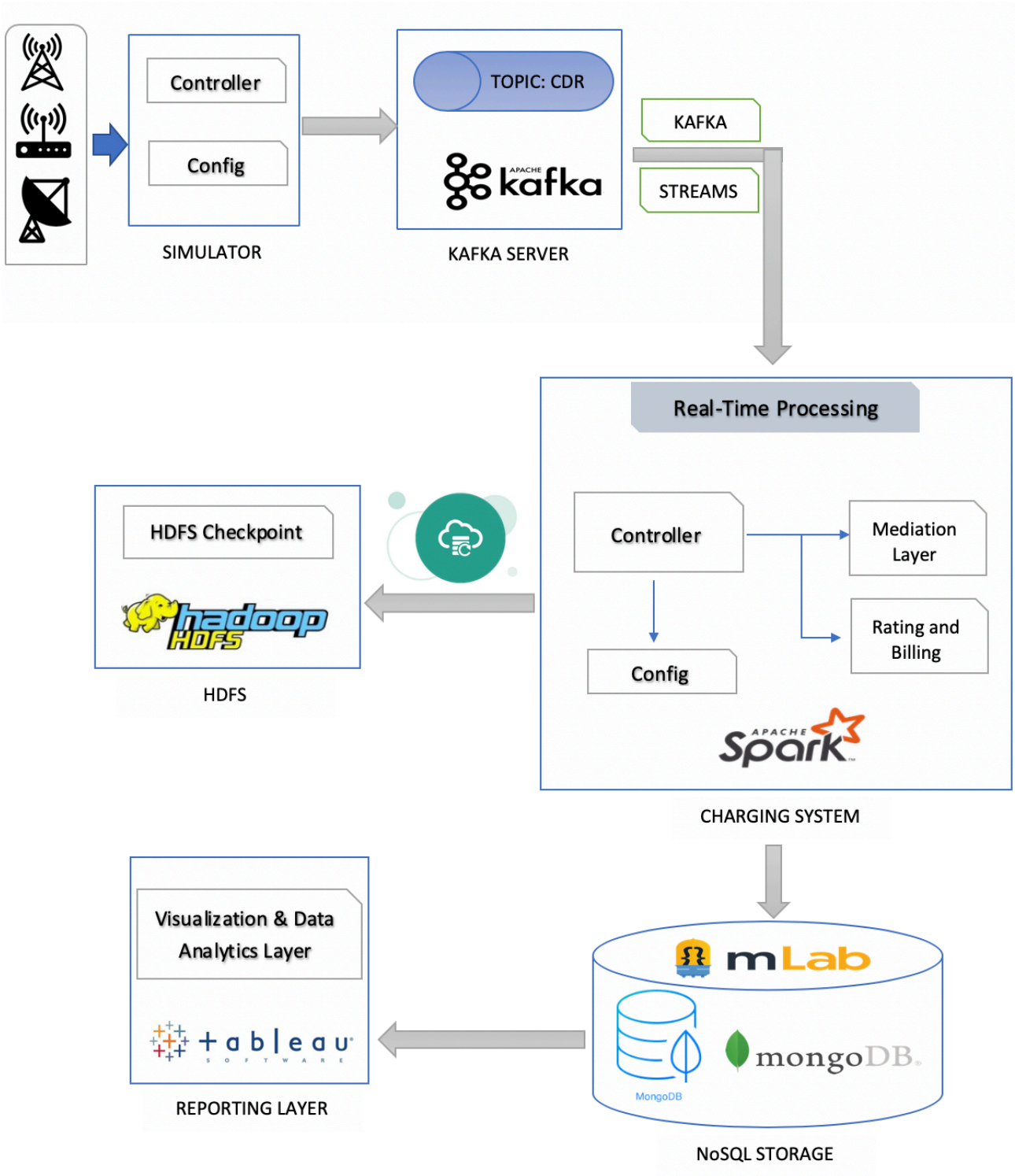
Real-Time Charging System for 5G Network

Mirza Tauqeer Baig - Ekramul Hoque - Mehdi Lebdi - Mehdi Nikkhah



Problem Definition:

Billing engine/Real time charging is one of the components of business support solutions (B.S.S) in telecom industry. Using big data ecosystem, our solution is capable of charging events coming from 4G (fourth generation) of telecom and is adaptive to handle multiple devices creating different event types in 5G (fifth generation). Our model considers only the mandatory fields relevant for charging. Other complex scenarios like call forwarding are exempted.

Methodology:

Architecture proposed for this solution is a microservices architecture where each component can be packaged and deployed separately. We have three main components in this solution:

Simulator:

• It loads a list of customers and cell tower locations from two csv files

• Maps the customers and cell tower locations randomly with given time interval of start date time and end date time

• The message created is then pushed into a Kafka server at a given range of rates. It maintains feeds of messages in topics

• Raw CDR generated by simulator and pushed in Kafka Topic:

2018-01-15T11:08:00.000Z, 2018-01-15T11:36:00.000Z, 43.918900|-79.816100, 44.763900|-79.992800, 4294283714, 8967939604, 0, DATA

Mediation:

• Consumes the Kafka stream and maps the data into a structure

• Filters the call that are dropped to showcase the capabilities of the system

• Data frame is created and passed to rating module

• Using reference from Cisco guide, the CDRs are flattened which are dependent on routers

• Record after CDR is mapped into dataframe by mediation module:

2018-01-15T11:08:00.000Z, 2018-01-15T11:36:00.000Z, 43.918900|-79.816100, 44.763900|-79.992800, 4294283714, 8967939604, 0, DATA

Rating:

• Receives streaming data frames from mediation module

• Loads static data frames from customers and offers collection in mongo

• Joined the two data frames and applied business logic for calls that are rated and billed

• The small static data frames will be broadcasted by framework itself

• Final dataframe after applying business logic in rating module:

C4294283714, 2018-01-15T11:08:00.000Z, 2018-01-15T11:36:00.000Z, 43.918900|-79.816100, 44.763900|-79.992800, 4294283714, 8967939604, 0, DATA, 1680, 0.04, 67.2

Visualization:

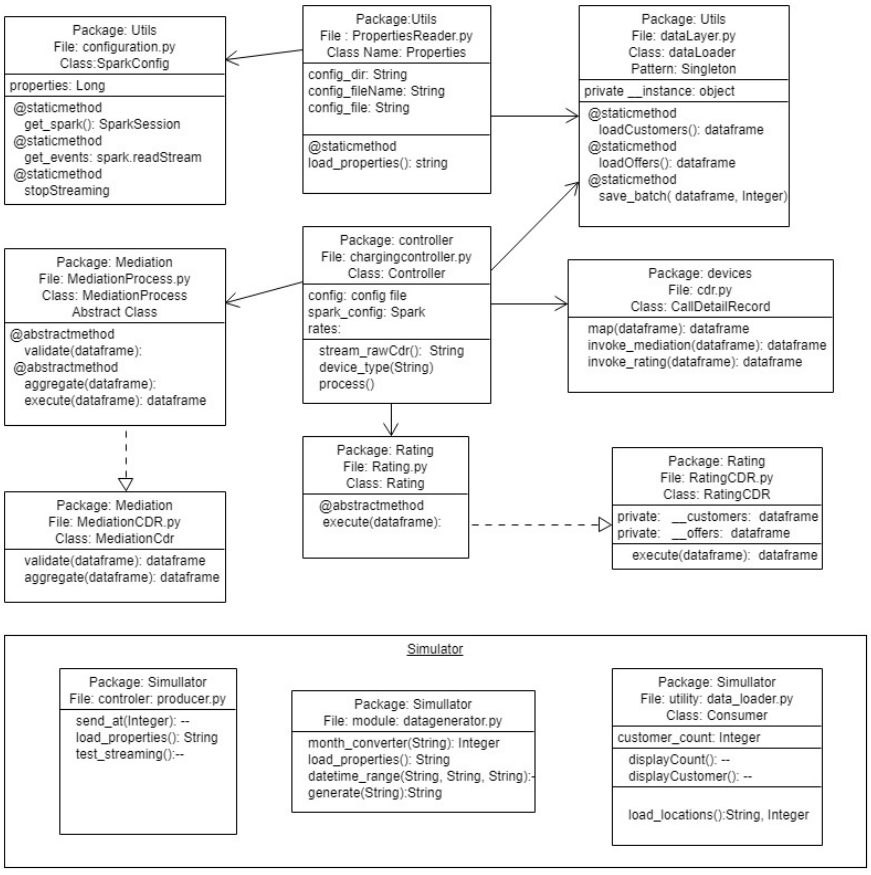
• Tableau loads data in CSV format pulled from mLab

• Graphs are generated for data analysis (more in Results section)

• Data loaded into Tableau for generating insights:


```
{
   "offerId": "Off002",
   "dateTimeConnect": {
      "$date": "2018-01-15T11:08:00.000Z"
   },
   "dateTimeDisconnect": {
      "$date": "2018-01-15T11:36:00.000Z"
   },
   "origNodeId": "43.918900|-79.816100",
   "destNodeId": "44.763900|-79.992800",
   "callingPartyNumber": "4294283714",
   "originalCalledPartyNumber": "8967939604",
   "callStatus": 0,
   "eventType": "DATA",
   "duration": 1680,
   "customerId": "C4294283714",
   "rate": "0.04",
   "bill": 67.2
}
```


Spark-Listener/Charging system listener:



Spark-Listener/Charging system listener:

- Properties are loaded using environment variable $APP_HOME, where config.ini is residing
- Configurations are passed to controller which:
    - Starts streaming process
    - Detects type of device
    - Maps raw events to structured data frames
    - Invokes mediation process
    - Invokes rating process
    - Checks data in hdfs
    - Persist data on configured database which is the MongoDB cloud server (mLab) for this architecture

Design Principles and Patterns:

- Object oriented programming methodology is followed in designing the solution
- Single responsibilities - Each class is responsible only for one responsibility for ex, propertiesReader class is responsible only for loading properties
- Coding to abstractions - We have followed this principle in mediation and rating module, using this principle we can onboard new devices (discussed later in results section)
- Singleton - Data Layer class is only one instance
- Command design pattern - Device instance is a self-aware instance which will have information about which pipeline to follow. Here, the CDR goes through a pipeline of MediationCdr and RatingCdr whereas for other family of devices it would follow the corresponding pipelines

Problems:

Kafka Integration: The Kafka messages produced and consumed were in two separate machines. The network connection required port and firewall configuration to integrate the connection between the two machines. This required knowledge on peer-to-peer ad-hoc wireless network. Properties changes were done in server.properties of Kafka configuration files.

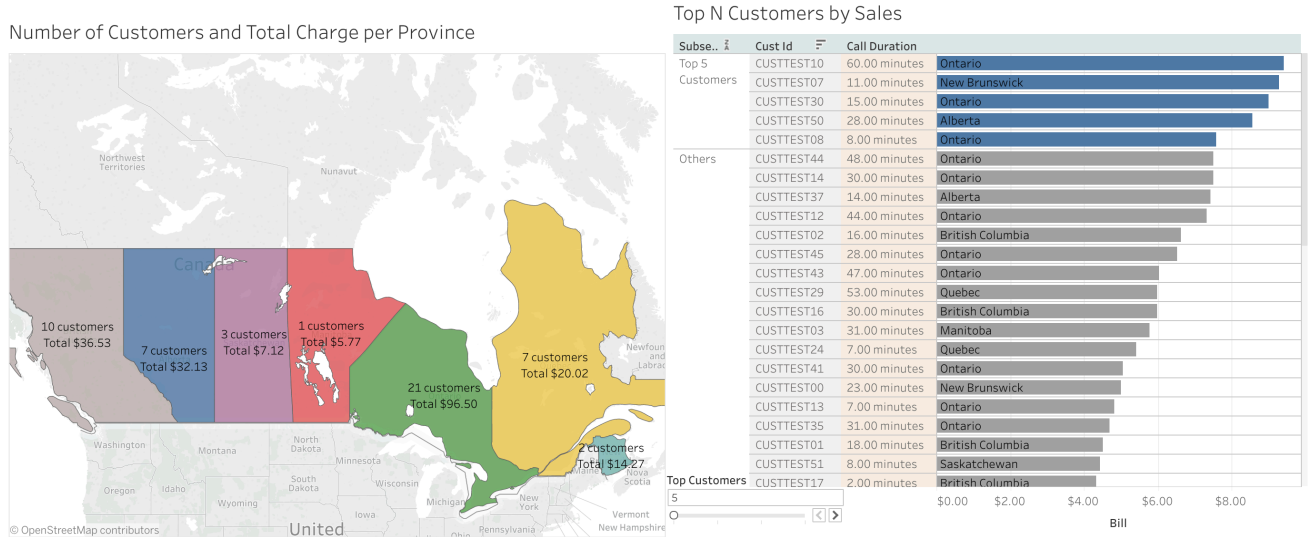
Mongo BI driver for Tableau: Integrating MongoDB into Tableau was an issue because the free version of Tableau doesn't support live databases. In order to solve this, a shell script was created to automatically pull the data from mLab and generate a csv file. The Tableau workbooks are continuously refreshed as more records are being downloaded.

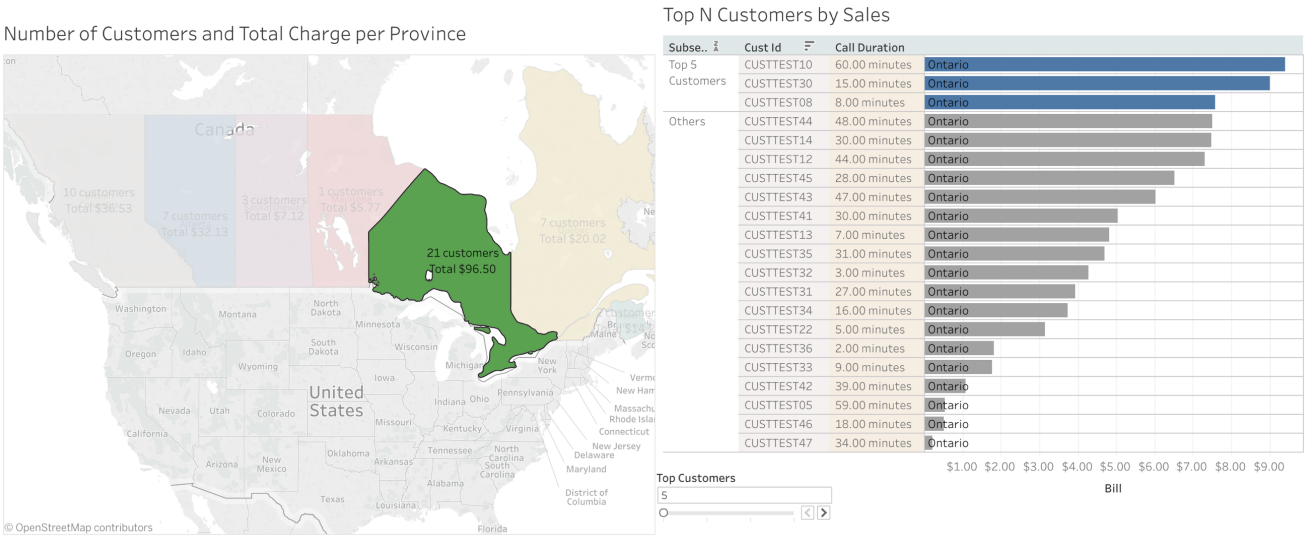
Fault tolerance using hdfs checkpoint: Error mitigation is important in real-time systems. In case of failure, or unintentional shutdown while writing the stream's batch into mLab, a checkpoint was created to provide a fault-tolerant system. The checkpoint feature will save all the progress information to the checkpoint location.

Project Summary:

• Getting the Data: 20/20
• ETL: 20/20
• Problem: 20/20
• Algorithmic Work: 0/20
• Bigness/Parallelization: 20/20
• UI: 0/20
• Visualization: 20/20
• Technologies: 20/20

Results:





The charging system is an end-to-end solution that is implemented to support 4G/5G networks and also adaptable to other domains, for example in the case of **autonomous cars where billing is based on startLocation/endLocation rather than startTime/endTime**. For the implementation, several technologies had to be merged to work in sync, such as Kafka, Spark, HDFS, MongoDB, and Tableau. The outcomes of the project are the insights that Telecom companies use for better decision-making. These insights include resource allocation to concentrated areas as well as improving customer billing offers. **Real-time analytics of events** was done on Tableau, generating graphs that provide insights for Telecom companies. For instance, the graphs above entail that most customers are located in Ontario and in second place in British Columbia as well as insights about top paying customers and location. Link to full Tableau Dashboard - https://public.tableau.com/profile/mehdi.lebdi#!/vizhome/CDR_inTest/Dashboard2