

Saiteja Ramisetty

DATA ENGINEER - Firemark Labs

Singapore

- Email me on Indeed: <http://www.indeed.com/r/Saiteja-Ramisetty/4436312fd65bf34b>

Work Experience

DATA ENGINEER

Firemark Labs

November 2018 to Present

IAG)

- + Involved in analyzing and documenting the requirements for the project.
- + Ingested Real-time data using Kafka and NIFI from Singapore govt provided public APIs like Traffic incident API, Weather Data, Speed band etc.
- + Developed Kafka producer in java and consumed using spark streaming to transform and process the data.
- + Worked on AWS EMR to quick launch and run the spark code on extracted raw dataset to perform computations and automatically terminates once the job is completed successfully.
- + Worked on setting up all components like NIFI, Kafka, Spark, Hive in AWS environment.

BIGDATA SPECIALIST - Verizon wireless

October 2017 to November 2018

Involved in end to end data architecture design.

- + Developed Producer application to load the data into Kafka topic and integrated with Spark Streaming to read the data from Kafka topic.
- + Parsed XML data to JSON data and created Dataframes on top of it.
- + Loaded transformed data into Hive and integrated hive with tableau for visualizing the data.
- + Apache Ignite tool evaluation and bench marking.

BIGDATA DEVELOPER

Accenture

May 2015 to October 2017

Worked on AWS EMR version of Hadoop environment.

- + Developed data pipelines using Data Ingestion framework which involves technologies like Sqoop, Bash scripting, HDFS, Hive, kafka, spark.
- + Import and exporting data in HDFS and Hive using Sqoop.
- + Executed queries using Hive and MapReduce jobs to analyze data
- + Developed the Hive UDF's to preprocess the data for analysis.
- + Worked on Spark SQL, Dataframes for Replacing Hive Queries.
- + Worked on Spark Core - RDDs, Transformations, and Actions.

SYSTEM ENGINEER

IBM Global

February 2013 to May 2015

services

- + Involved in gathering requirements, analyzing and documenting those into specifications
- + Developed data pipeline using Flume, Sqoop, Pig and map reduce to ingest customer behavioral data and purchase histories into HDFS for analysis
- + Used Hive to analyze the partitioned and bucketed data and compute various metrics for reporting on the dashboard.
- + Responsible for developing data pipeline using Sqoop and hive to extract the data from RDBMS databases and store in HDFS

Education

BACHELOR OF ENGINEERING in ECE

Anna University - Chennai, Tamil Nadu
2008 to 2012

12TH

Narayana Junior College
2006 to 2008

Ratnam High School
2005 to 2006

Skills

Hive (6 years), APACHE HADOOP HDFS (4 years), APACHE HADOOP SQOOP (4 years), HADOOP DISTRIBUTED FILE SYSTEM (4 years), HDFS (4 years), Spark (3 years), AWS (3 years)

Additional Information

SKILLS

BigData Eco System Spark Dataframes, Spark SQL, Spark Streaming, Kafka, HDFS, MapReduce, Hive, Sqoop, Flume, Pig, Oozie, Impala Shell, Apache Ignite.

Cloud technologies AWS EMR, AWS ECR, ECS, S3, AWS Glue, AWS Athena, AWS Redshift, IAM, VPC, Code Commit

Programming Languages Java, Scala, Python