Rakesh Adhikari
Email: iamrakesh1502@gmail.com
Mobile: +91-7290002474

Summary

Senior Data Engineer currently associated with Nokia Solutions & Network. I am working in Big data technologies and has leveraged the Hadoop ecosystem to create highly scalable and fault-tolerant systems in domains such as finance, telecom, and healthcare.

I am an experienced Data Engineer with good experience in building and managing big data projects. My areas of expertise are the design and development of Data lakes/data pipelines using Hadoop and Spark. I also have good exposure to various Devops tools like Kubernetes, Docker, Helm.

Previously, I have worked as a Consultant at Deloitte USI and TATA Consultancy Services.

Work Experience

S No	Organization	Roles and Responsibilities	Duration
1)	Nokia Networks & Solutions	Senior Data Engineer, Team Lead	Nov 2018- Till Date
2)	Deloitte Consulting	Consultant, Team Lead, Data Engineer	Jan 2017-Oct 2018
3)	Tata Consultancy Services	IT Analyst, Big Data Developer	Dec 2011-Jan 2017

Technical Skills

- **Technologies**   Apache Spark, Scala, Hive, Pyspark, SQL, Elasticsearch, Kudu, Airflow, Nifi, Impala, Sqoop, MapReduce
- Databases   MySQL, Oracle 10
- Operating Systems   UNIX, Windows
- Tool/Services   Control-M, Oozie, Git lab
- Container   Docker, Kubernetes, Helm
- Certifications: ITIL Foundation, Astronomer Certification for Apache Airflow Fundamentals
- Cloud Services   AWS, Azure

Nokia Solutions & Network, Noida, India
November 2018 to present
Senior Data Engineer

Project 1:

Worked on a solution to predict probable degradation on KPIs at a cell level considering the network KPIs, performance counters and external data like weather information
- Creating a scalable and generic solution in Spark and Scala which can suffice the user requirement for multiple technologies like 2G,3G,4G,5G
- Creating a integrated pipeline in Airflow which will acts as an orchestrator to drive the complete execution of solution covering Data Hub, Data Engineering, Data Science and Data Visualization jobs
- Created pipeline to provide data to ML team for executing Predictive Model, Random Forest

- Storing daily job data to Postgres table for data visualization
- Created helm charts and docker files to build and deploy Spark applications on the Kubernetes environment
- Sprint Planning and doing code reviews of team ensuring continuous and automated deployment using GIT CI/CD pipelines

Technologies Used: Scala, Postgres, AWS S3, Spark, Airflow, Docker, Kubernetes, GIT CI/CD, Unix Shell Scripting.

Project 2:
Worked on a solution to provide an automatic mechanism for analysis of logs which will reduce the time care engineers are consuming to analyse logs and help them on figuring out swiftly the reasons behind a specific incident

- Handled automated Log Analysis of 11 Nokia products with different proprietary log files.
- Created a processing engine in Spark and Scala which helped in processing multiple 20-25GB log bundles in parallel within ~5 mins and store data in Elasticsearch.
- Convert data into structured data on which Rules/Visualizations/Queries can be developed by SMEs
- Worked on microservices to expose data stored in Elasticsearch for end-user consumption.
- Created helm charts and docker files to build and deploy Spark applications on the Kubernetes environment.
- Created pipeline to execute Machine learning Predictive analytics module.
- Sprint Planning and doing code reviews of team ensuring continuous and automated deployment using GIT CI/CD pipelines

Technologies Used: Java, Scala, Confluent Kafka, Elasticsearch, AWS S3, Spark, Spark SQL, Docker, Kubernetes, GIT CI/CD, UNIX Shell Scripting.

Project 3:
Worked on build a generic solution to join PM, FM and CM data into single dataset containing relevant alarm details and configuration parameter values, along with NBH-BBH hour and computed KPI for different customers and technologies(2G,3G,4G).

- Developed NIFI pipeline to generate csv file from .XML and .XLSX file
- Cleansing data and picking up relevant alarm details and concatenate important columns for evaluating NBH (Network Busy Hour)-BBH (Bouncing Busy hour) and KPI computation in Scala Spark
- Merging output from Machine Learning modules and projecting the data in Hive Table using Scala Spark
- Coordinating with Power BI team for any enhancements or development

Technologies Used: Hive, Spark, Spark SQL, Scala, Power BI, AWS s3, Zeppelin

Deloitte Consulting USI - Gurgaon, India
January 2017 to October 2018

Project 2:
Consultant
- Worked on Data Platform Modernization for a Telecom giant in Australia. The project was on the Migration of the client's existing Teradata EDW to the Hadoop platform for a reduced cost, fast processing, and better insights.
  - Involved in the creation of an end-to-end data pipeline using Spark and Hive.

- Created various spark jobs to validate and load the data into hive staging tables.
- Worked on complex hive queries and joins in SPARK SQL to load the Landing layer, Worktables, and Fine grain layer.
- Involved in data analysis and performance tuning methodologies in spark jobs.

- **Technologies Used:** Hadoop, MS Azure, Hive, Impala, Kudu, Spark SQL, Spark, Scala, Azure SQL DB, Cloudera 5.10

Project 2:
Consultant
- Worked on the implementation of a data lake and aggregate layer for an Insurance client to enable "speed to insights" and consumption of structured, semi-structured, and unstructured data.
  - Ingested historical data from the DB2 database to Hive tables using Sqoop.
  - Created various Hive UDFs to Validate, format, and structure data to be loaded in the Initial layer.
  - Developed Oozie workflows to schedule Spark/Hive jobs end to end flow.
  - Worked on complex hive queries and joins in SPARK SQL to load the Landing layer, Worktables, and Fine grain layer.
  - Involved in Apache NiFi development to schedule incremental data load into Initial Layer.
  - Involved in data analysis and performance tuning methodologies like Partitioning and bucketing of hive tables.

- **Technologies Used:** Hadoop, Sqoop, Hive, Spark SQL, Apache NiFi, Oozie, DB2, HDP 2.4

Tata Consultancy Services – Hyderabad, India
July 2015 to January 2017

IT Analyst
Project 1:
The purpose of the project is to make a data lake to handle huge volume of 12 years data efficiently hence their requirement was to migrate data to Hadoop system and then generate objects as required by program to visualize the sales reporting in SAP BO.
- Oracle to Hadoop Migration- Work on Sqoop, Hive and Shell script to import Oracle database to HDFS and made automation tool to run daily import updated data
- Written the Apache PIG scripts to process the HDFS data.
- Involved in creating Hive tables, loading with data, and writing hive queries which will run internally in Map Reduce way.
- Generation of Unix and HQL scripts to load flat file data in Impala and Hive tables

Technologies Used: Hadoop, HDFS, MapReduce, Pig, Hive, Sqoop, DB2, RedHat Linux

Project 2:
To purpose of this project was to deal with the establishment of supply, demand and inventory process, to generate statistical forecast and integrate with Supply and Inventory planning. Build on Big Data platform using technologies like: Hive, Sqoop, Impala, Kudu, and Unix Shell Scripting
Data loaded into Hadoop is transformed and the extracts are generated into the outbound directory. MFT picks up the generated files from the outbound directory and places them HCI inbound. SAP IBP jobs uses those extracts for both Master data and Transactional data.

- Oracle, AS400 to Hadoop Migration- Work on Sqoop, Hive and Shell script to import Oracle, AS400 database to HDFS and made automation tool to run daily import updated data
- Generation of Unix and HQL scripts to load flat file data in Impala and Hive tables
- Involved in writing and implementation of HiveSerdes.

Technologies Used: Hadoop, HDFS, AS400, Hive, Sqoop.

December 2011 to July 2015
System Engineer
As a part of an Insurance client worked in Life Services team in a managed service model providing end to end support, enhancements, maintenance to Client Application
- Involved in requirements gathering and understanding, Analysis, preparing test cases, Design, development, and testing of software modules.
- Reviewing and documenting test results.
- Involved in Monitoring and Error resolution of any failed jobs in both the UAT and Production phase.

Technologies Used: VB 6, Oracle, Java, Webservices Quality Center, MS Access.

Education:
Chhattisgarh Swami Vivekanand Technical University, Bhilai, CG, India
Bachelor of Engineering (Electronics & Communication), 2011