**Naveen R**           Phone: +91 9908938227

Data Engineer E-mail**:** [naveen.it2017@gmail.com](mailto:naveen.it2017@gmail.com)

**PROFESSIONAL SUMMERY:**

* Having overall 5+ years of IT experience in that 2+ years of experience AWS Cloud platform with DevOps tools, 3+ years Big Data Application Development and Support for the Cluster configuration management.
* Understanding the client environment, Administrating and creating solutions using AWS Services, CI/CD, GIT, Dockers.
* Configuring Jenkins job with related plugins for Testing, Artifactory, and Continuous Deployment to accomplish the complete CI / CD.
* Experience with Amazon Web services Creating, configuring and Managing EC2 services across AWS Zone.
* Experience in using version controller tools like GIT.
* Environment provisioning solutions using Github Actions, Docker, Vagrant.
* Hands on expertise to setup, configure continuous build processes using Jenkins, Maven, Git
* Automated Deployment to all environments (Dev, QA Test, Demo, Performance Test and Staging/UAT)
* Hands on installation of application services: Apache HTTPD, Tomcat, MySQL, MongoDB, nginx and write Automation scripts for configure nginx, Apache Tomcat, MySQL, MongoDB
* Versioning and Artifactory Repository Management. Keeping track of all builds and versioning the builds for backup strategy.
* Knowledge in Shell Scripting.
* Strong working experience with ingestion, storage, processing and analysis of big data.
* Experience on creating databases, tables and Views in HIVE
* Worked with Sqoop in importing and exporting data from different databases like MySQ & Oracle
* Good in adopting to new technologies and implementing.

**EXPERIENCE:**

* Working as a Hadoop Devloper &DevOps Engineer for Tech Mahindra Ltd. from Jan 2017 till date.

**EDUCATIONAL QUALIFICATION:**

* I have completed my Master of Computer Applications in 2014 from JNTUH.

**TECHNICAL SKILLS:**

|  |  |
| --- | --- |
| **Operating system** | CentOS, RHEL, Ubuntu and Windows |
| **Version control** | GitLab, GitHub |
| **Build management tool** | Maven, Ant |
| **Configuration management tool** | Ansible, Puppet |
| **Application servers** | Apache Tomcat, Application Server |
| **CI tool** | Github Actions, Jenkins |
| **Scripting language** | Shell scripting |
| **Containerization tool** | Docker |
| **Cloud** | AWS(EC2) |
| **Bigdata** | Hadoop HDFS,MapReduce,Sqoop,Hive & Pig |

**Project Details:**

**Project Name** :CDAP (CASK Data Application Platform)

**Client**: Thomson Reuters

**Role** :Devops Engineer

**Description:** The Cask™ Data Application Platform (CDAP) is an integrated, open source application development platform for the Apache Hadoop® ecosystem that provides developers with data and application abstractions to simplify and accelerate application development, address a broader range of real-time and batch use cases, and deploy applications into production while satisfying enterprise requirements.

**Roles and Responsibilities**

* Interacted with different development teams to provide support in deployment activities

• CreatedBranches and participatedin merging and code changes using Git.

• Creating and maintaining the source code repositories using GIT.

• CI/CD through Jenkins and automate the build Process using Jenkins jobs by writing Pipeline Script.

• Responsible for maintaining/Administration of Jenkins Continuous integration tool.

• While building the code Used Jfrog as corporate Repository.

• Work with development Team and Testing Team to establish a build schedule, execute the builds and troubleshoot build failures.

• Deploy and monitor scalable infrastructure on configuration management using Puppet.

• Perform daily environment health Check of different environments like Dev, QA, Pre-Prod and Prod environments.

• Created Production, development and testing environments by communicating with multiple teams.

• Configured logstash for monitoring log events and viewing the logs using Kibana.

• Used ServiceNow as ticket tracking tool and Jira as Bug/defect tracking tool.

**Project Name** : Analytical Data Mart

Client : Mobistar

**Environment**  : HDFS, Hive, SQOOP, Impala,Java, Ozzie,Unix,Netezza,Hive, Spark, Scala & Spark SQL.

**Role** :Hadoop Developer

**Project Description:**

Mobistar provides telecom services to its customers. The objective of this project is to design and build the data marts for all customers aggregated data like number of customers existed, number of customers joined, Agent commission, Agents Existed, Tariff Amounts, Data packs, time period for payments. Based on these metrics they analyze their business and make the decisions to improve the business. Data mart is developed in Star schema consisting the Facts and Dimensions, using these Facts and Dimensions BI Team generate various types of reports and distributed to different levels like Director Level, Group Director Level and Organization Level.

**Roles and Responsibilities:**

* Moved entire netezza schema tables into HDFS for further processing.
* Developed the sqoop scripts in order to make the interaction between Hdfs and Oracle Database.
* Created the Ozzie workflows to execute shell action.
* Created Hive tables to store the processed results in a tabular format.
* Writing the script files for processing data and loading to HDFS
* Developed the UNIX shell scripts for creating the reports from Hive data.
* Create the Impala Base tables and analytical tables.
* Completely involved in the requirement analysis phase.
* Created External Hive Table on top of parsed data.
* Converting Hive/SQL queries into Spark transformations using Spark RDDs and Scala.

**Project Name :**

**Distributed Network System**

**Operating System**-Linux

**Designation**: Hadoop Developer

**Project Description**: - Purpose of this project is to store terabytes of transactional log information generated by the system. These log files are produced on monthly basis. These log files has to be parsed by set of rules defined in the various formats. Initially these were loaded into the database and retrieving these rules is time consuming activity. With the solution based on the open source BigData (Hadoop), we reduced the time for the whole process. Data will be stored in Hadoop distributed file system and processed using Map/Reduce and Pig scripts. Which intern includes getting the data from the websites, process the files to obtain the analysed information from all the logs, Extract various reports out of this information and Export the information for further processing. This will help to meet the client requirement of revenue increase by delivering comprehensive data, advanced analytics.

**Environment**: Hadoop 2.X, Hdfs,MapReduce, Apache Pig, Hive, SQOOP, Java, UNIX, HBase, MySQL.

**Roles & Responsibilities**:

* Analyzing the requirement to setup a cluster.
* Prepared Design document specifications.
* Moved all log files generated by various network devices into HDFS location.
* Written Map Reduce code that will take input as log files and parse the logs and structure them in tabular format to facilitate effective querying on the log data
* Developed the sqoop scripts in order to make the interaction between Pig and MySQL Database.
* Involved in developing the Pig scripts.
* Involved in creating the External Hive Table on top of parsed data.