Rahul Dabhade

rahul.dabhade@volkswagen.com | +918806431254

Employee ID: 104421

EDUCATION

PUNE UNIVERSITY

B.E IN COMPUTER ENGINEERING

July 2017 | India

MAHARASTRA STATE BOARD OF TECHNICAL EDUCATION

Diploma IN INFORMATION TECHNOLOGY

July 2014 | India

TECHNOLOGY

• Terraform • AWS EC2 • AWS S3 • AWS Autoscaling group • Elastic Load Balancer • Code pipeline • Code deploy • MariaDB • EFS • GitHub Actions • GitHub CICD • Confluence • Spark • Python • PostgreSQL • AWS RDS • HBase • Hive • AWS S3 • EMR Cluster for HBase and Hive.

DATABASES

• MariaDB • MySQL • ORACLE

OPERATING SYSTEM

• Linux • Windows

MANAGEMENT SKILLS

• Agile Methodologies • Problem-Solving and Critical Thinking
• Cross-functional Collaboration • Jira
• Scrum • Collaboration and teamwork

Professional Summary:

- 4+ years of Experience in different domains in IT industry including Big Data Hadoop Development and DevOps.
- Good knowledge of terraform architecture
- Hands on the terraform module and concepts
- Experience on the creating the infrastructure on AWS environment
- Written a terraform code for the AWS to meet the business requirements.
- Hands on the GitHub YML coding for the particular task.
- Good knowledge on the AWS services EC2, S3, RDS, EFS Autoscaling group, secretes manager.
- Have hands on experience Spark Architecture, SparkCore, SparkSql.

EXPERIENCE

VOLKSWAGEN GROUP OF TECHNOLOGY SOLUTION PVT. LTD. |

DATA ENGINEER

December 2023 – Present | Pune, India

Project 1 : Cloud Analytic Platform (B2C)

- The B2C website of Volkswagen Group, accessible at volkswagengroup.fr. This site serves as a public website designed to connect directly with the customers, offering them a comprehensive digital experience of what Volkswagen Group has to offer.

Technology: Terraform, AWS EC2, AWS S3, AWS Autoscaling group, Elastic Load Balancer, Code pipeline, Code deploy, MariaDB. EFS, GitHub Actions, GitHub CICD, Confluence.

Responsibilities:

- Create an infrastructure based on requirement
- Create an two EC2 instance throw the Auto Scaling group
- Attach the VPC, subnets, security group to ec2 instance
- Create IAM Role and Policies for Data Lifecycle Manager
- Create a RDS with MariaDB instance attach VPC, subnets, VPC security group ingress rule.
- Create S3 bucket for App data file used by Code Deploy.
- Attach the Lifecycle Rule for Deployment Bucket and Lifecycle Rule for Pipeline Bucket.
- Create Secret Manager for storing vgf-b2c-app Repository DB credentials.
- Working on the b2p-app repository that are consist of the actual application.
- Uploading the artifact from b2c-app repository to CAP 2.0 deployment bucket so code pipeline and code deploy work.
- Configure the GitHub secrete of CAP 2.0 for the access the credentials from GitHub secretes for three stages Deployment, Pre-production and Production.
- Need to add the GitHub workflow YML file configuration for three AWS environment Deployment, Pre-production and Production.
- In b2c-app Configure AWS Credentials and Assume state role CAP 2.0
- Setup secrets and the Deployment s3 bucket name.
- Site live on pre-production and test the site and work fine.
- Documenting each part of like changes made, issues, Architecture redesign.

Project 2 : Telecom Data Analysis

- **Technologies Skills :** Spark, Python, PostgreSQL, AWS RDS, HBase, Hive, AWS S3, EMR Cluster for HBase and Hive.

- **Responsibilities:**
- Understand the AWS RDS source and work with business cases.
- Written pyspark jobs to extract records from AWS RDS source.
- Involve in creating Hive tables, loading with data and writing Hive queries.
- Created HBase tables to load large sets of structured, semi-structured and unstructured data coming from RDS source
- Extract data from AWS RDS to HBase by using Spark JDBC driver.
- Raise JIRA for Infrastructure, Platform issues.
- Extract data from HBase to Hive by using transformation and stored into Hive.
- Written Hive Queries for Data Analysis to meet the business requirements.
  - Checks duplicate record in Hive and perform Deduplication Job for remove duplicate record.