

Syed Awais Syed Shafiuddin

Email: syed123awais@gmail.com

Mob: +91 8793422042

HADOOP ADMINISTRATOR

5.2 Years of Experienced in Apache Hadoop with Cluster Orchestration Tool like Cloudera Manager, Hortonworks.

RESPONSIBILITIES:

- Hands on Experience in Installation, Configuration, Supporting and Managing Hadoop Cluster using Apache Hadoop and Cloudera.
- Deployed, Configured and Handling Various Components of Hadoop Ecosystem like Hive, Impala, Sqoop, Spark, Hue, Oozie, Zookeeper.
- Experience on Commissioning, Decommissioning, Balancing and Managing Nodes and Tuning server for Optimal Performance of the Cluster.
- Installing Hadoop in the Amazon Cloud and using Amazon Cloud Products like EC2, EMR, and S3 etc.
- Installing Hadoop in the Google Cloud Platform and using Google Cloud Platform Products like Dataproc Google Storage etc.
- As an Admin Involved in Cluster Maintenance, Trouble Shooting, Monitoring and Followed Proper Backup and Recovery Strategies.
- Consolidating Health Monitoring Issues and Configure that Issues.
- Been a Part of POC Team to Help Build New Hadoop Clusters and Hadoop Deployment Decisions.
- Hands on Experience in Analysing Log files for Hadoop and Ecosystem Services and finding Root Cause.
- Configured Various Property files like core-site.xml, hdfs-site.xml, maped-site.xml, and yarn-site.xml.
- As an Admin followed Standard Backup Policies to make sure High Availability (HA) of Cluster.
- Loading data into cluster from RDBMS using Scoop and Kafka, also from local file system to the Hadoop cluster.
- Create And Configure Resource Pools for Job Scheduling in Dynamic Resource Pool configuration.
- Supervising Hadoop jobs using scheduler.
- Manage User Roles by cloudera Manager.
- Setting up New Hadoop Users Roles Policy by Ranger.
- Monitor Hadoop Cluster Connectivity and Performance.
- Experience in being define in Job Flows by Oozie.
- Experience in Handling Namenode Failover and DataNode Failovers.
- Upgraded Cloudera Manager and CDH for Production Clusters.

PROJECT SUMMARY:

1. Hadoop POC

Title: Deployment Cloudera Data Platform CDP 7.1.3

Description: Deploying CDP 7.1.3 on AWS Cloud to Creating POC for Production Upgrade.

Responsibilities:

- Verify and Amend the OS Parameters.
- Set Prerequisites for Hadoop Cluster.
- Installations Cloudera Runtime, Cloudera Manager, Managed Services.
- Deployed and Configure Hadoop Core Ecosystem Services.
- Enable HDFS High Availability.
- Setting up New Hadoop Users Roles Policy by Ranger.
- Performing Benchmark Testing.
- Generating Report.
- Managing Roles to Reduce the Load on Cluster.

2. Hadoop POC

Title: Data Migration to Hadoop.

Description: Deploying Cloudera 5.16.4 and Migrating Data from MySQL to Hadoop for Data Mining.

Responsibilities:

- Deploying CDH 5.16.4 on Google Cloud Platform and Configuring to Import Data from MySQL to Hive.
- Configuring Impala for Data Mining.
- Configuring Hue and Creating Users.
- Tuning Yarn and Hive.
- Importing Data from MySQL via Scoop.
- Monitoring Scoop Jobs.
- Configuring Spark2.
- Installation and Configuration of Streamsets on Cloudera Manager.
- Upgrading Cluster to Java 8.
- Benchmarking and Performance Testing Drills for Customer Satisfaction and Creating Report of Activity.
- Migration of Data from Production to Test Environments.
- Hadoop Cluster Benchmarking Report Generation.

EXPERIENCE SUMMARY:

Name of Company: Defacto Veritas Pvt Ltd

- Designation: Hadoop Administrator
- Duration: 4 year's (FEB 2016 to At Present)

SKILLS:

- Hadoop Ecosystem: HDFS, Yarn, Hive, Scoop, Impala, Hue, Oozie, Spark, and Kafka.
- Hadoop: Cloudera, Hortonworks, EMR, Dataproc and Apache Hadoop.
- Operating System: Redhat, Centos, Ubuntu, Debian and Windows
- RDBMS: MySQL.
- Ticketing Tool: Jira and Servicenow.

EDUCATIONAL QUALIFICATIONS:

- B.E (Mech) : 7.78 CGPA
- HSC : 60.83%
- SSC : 66.15%

PERSONAL INFORMATION

- Marital Status: Unmarried.
 - Sex: Male
 - Language Known: English, Hindi, Marathi, Urdu, Arabic.
 - Nationality: Indian
 - Passport: Yes
 - Pan: Yes
-