

Rishabh Dubey



Lead Big Data Admin and DevOps SRE with 11+ years of experience in Big Data and Cloud

To achieve high career growth through a continuous learning process and keep myself visionary and competitive with the changing scenario of the world.

OVERVIEW:

- 11+ Years of extensive experience in Big Data and Cloud technologies.
- Strong experience and knowledge of Hadoop, Docker, Kubernetes, Portainer, Confluent, Prometheus, Grafana, Zepplin, Anaconda, Jupyter Notebook, Spyder, Python, Ansible, Spark, Kafka, Yarn, Map-Reduce, Hive, Hue, Ambari, Nifi, Cloudera Manager, Ambari, Hbase, Sqoop, Zookeeper, Airflow, Oozie which consists of installing, configuring and performance improvement of DevOps and Hadoop ecosystem components.
- Having experience as Big data Lead regarding end-to-end implementation of Big Data projects which includes requirement gathering, designing, development, testing, troubleshooting, scheduling jobs and production support.
- Strong experience on Hadoop Security like LDAP, SSL, Kerberos and Apache Ranger.
- Having vast experience in Administration of Hadoop Cluster which includes setting up multi node cluster fully automated through Ansible along with how to manage HDFS file system perfectly.
- Huge experience in upgrading different DevOps tools and large Hadoop Clusters.
Setup backup, restore policies and Disk failure process for large Hadoop Clusters.

Work Experience

Lead Big Data Admin and DevOps SRE

Pythian - Hyderabad, India

Client Name: Disney-Hulu, Index Exchange, Rocket Lawyer, Tucows etc

Oct 2019 to Present

Role Description: Pythian has a lot of clients namely Hulu, Index Exchange etc. Primary responsibility is to work with client's BDI Team and look into Big Data Infrastructure issues and work to resolve them. As an SRE main work is of DevOps and Admin role combined and look into OS, Disk, Hardware, Memory related issues present in Big Data Infrastructure On-Premise and Cloud both. Moreover check and monitor the Cloudera and Hortonworks cluster and provide guidance and expert level suggestions to resolve any infrastructure/software related issues.

Roles & Responsibilities:-

- ✓ Guiding team to set up cluster, automate and configure cluster properties through Ansible and Blueprint.
- ✓ Creation of plans to improve existing Hadoop Infra and providing suggestions to improve security.
- ✓ Implemented Kafka and Spark on Zeppelin using Docker and Kubernetes.
- ✓ Guided code/design analysis, strategy development and project planning.
- ✓ Assisted with data capacity planning and node forecasting.
- ✓ Helped team to install and configure Kerberos authentication and Apache Ranger for security purpose.
- ✓ HDFS administration and Troubleshoot for Cloudera and Hortonworks all type of issues including performance tuning.
- ✓ Removal of logs, temp space and trash for unhealthy nodes and as per the cluster issues.
- ✓ Used Sqoop to import data from RDBMS and creation of Linux Scripts to automate jobs.
- ✓ Created Oozie Workflows to automate dataloads.
- ✓ Guided the team regarding any Admin tasks and provided suggestions on both job and cluster level of performance improvements.

Lead Big Data Admin and DevOps

Sears Holdings Corporation - Pune, India

Project Name: Sears Hadoop Infrastructure

Nov 2016 to Oct 2019

Role Description: Sears IT and Management Services India Pvt. Limited provides corporate IT services to Sears Holding Corporation, a leading integrated retailer with more than 2,350 full line and specialty retail stores in the United States and Canada and the home to shop your way, a social shopping platform offering members rewards for shopping at Sears and Kmart as well as with other retail partners. Sears implemented Hadoop to correlate huge quantities of information

about everything from product availability in specific stores to competitor prices on specific products to information about local economic conditions in order to set prices.

Roles & Responsibilities:-

- ✓ Lead Big Data Admin and DevOps team to troubleshoot and resolve regarding any issues related to Big Data clusters including performance tuning.
 - ✓ Minimized involvement of Cloudera and Hortonworks support for critical issues by guiding team to resolve most of the issues by their own.
- Guided team to establish disk replacement task for all Hadoop infra env.

Senior Big Data Admin and DevOps

Saama Technologies Ltd. - Pune, India

Client: American Automobile Association and Motorist Insurance Group

June 2015 to Nov 2016

Project and Role Description: Project was regarding building a data warehouse from RDBMS to HDFS Hadoop to combine customer data for policies, claims, agents, billings for further analysis. This involves combining data from multiple sources, capturing the business requirements and making sure the correct data is in place.

Roles & Responsibilities:-

- ✓ Lead a team of developers and admin to migrate data from RDBMS to Hive HDFS which included cluster planning, requirement gathering, development of Hive queries with UAT, QA testing, dashboarding.
- ✓ HDFS administration and Troubleshoot regarding all type of Hadoop env related issues.

Senior Big Data Admin

Wipro Technologies Ltd. - Bangalore, Karnataka

Projects: 4 different Projects

Project Name: Various Hadoop

Infrastructure Clients: RBS, Visa

June 2010 to June 2015

Role Description: Worked as a Hadoop Admin and Developer wherein installed Hadoop Clusters from scratch and helped Development team regarding any issues.

Roles & Responsibilities

- ✓ Helping business processes by developing, installing and configuring Hadoop ecosystem components that moved data from individual servers to HDFS.

Areas of Expertise:

Big Data Eco System: Hadoop, Cloudera Manager, Ambari, Hue, Yarn, HDFS, Yarn, Spark, Kafka, MapReduce, Hbase, Hue, Hive, Nifi, Flume, Sqoop, Oozie, Spark, ZooKeeper, Solr

Configuration Management Tools:- SaltStack, Ansible, Cobbler

Dev Ops Tools: Docker, Kubernetes, Portainer, Confluent, Prometheus, Grafana, Anaconda, Jupyter Notebook, Spyder, Nagios, GitHub, SVN

Languages:- Java, Python

Reporting/Visualization Tools:- Prometheus-Grafana, Tableau, Google Chart, Hue

Security Tools:- LDAP, SSL, Kerberos, Ranger, Sentry

Workflow Management:- Airflow, Oozie, Control-M

ACADEMIC CREDENTIALS

BOARD / EXAMINATION	SCHOOL / COLLEGE	PERCENTAGE	YEAR OF PASSING
B.Tech in Computer Science Engineering	Jaypee Institute of Information & Technology	70.0%	2010
12 th	Obra Inter College	63.0%	2005
10 th	SVMHS, Obra	68.0%	2003

PERSONAL DETAILS

- **Current Location:** Hyderabad, India
- **Languages Known:** English and Hindi
- **Availability:** Immediately
- **Mobile No:** +91 7066144601
- **EmailId:** rishabhadoop@gmail.com