

Role: Senior Data Devops Engineer / Cloudera Administrator

Overview:

As a senior Cloudera Administrator/Data DevOps Engineer, you will be responsible to set up, shape, administer, and test the applications related to Hadoop platforms. You will be part of a team of DevOps engineers focusing on the day-to-day tasks of managing and maintaining On-Prem and Cloud environments and will be hands-on involved with CI/CD process and monitoring application servers. Candidate must be comfortable working in an agile environment.

Responsibilities:

In this job you will:

- Help install Cloudera-Hadoop from ground up with SDLC cycle methodology including Dev, Test, Cert, Production and Disaster Recovery
- Plan capacity planning, infrastructure planning and version fix to build Hadoop Cluster.
- Implement Hadoop security like Kerberos, Cloudera Key Trustee Server and Key Trustee Management Systems
- Enable Sentry for RBAC (role-based access control) to have a privilege level access to the data in HDFS.
- Perform upgrades to Cloudera Manager, CDP along with support for Linux Server Patching.
- Provide infrastructure and support to software developers to rapidly iterate on their products and services and deliver high-quality results. This includes infrastructure for automated builds and testing, continuous integration, software releases, and system deployment.
- Design and implement Backup and Disaster Recovery strategy based out of Cloudera BDR utility for batch applications and Kafka mirror maker for real-time streaming applications.
- Align with development and architecture teams to propose and deploy new hardware and software environments required for Hadoop and to expand existing environments.
- Support Kafka cluster to enable real-time streaming applications.
- Monitor and coordinate all data system operations, including security procedures, and liaison with infrastructure, security, DevOps, Data Platform and Application team.
- Ensure proper resource utilization between the different development teams and processes.
- Design and implement a toolset that simplifies provisioning and support of a large cluster environment.
- Align with the systems engineering team to propose and deploy new hardware and software environments required for Hadoop and to expand existing environments if needed.
- Apply proper architecture guidelines to ensure highly available services.
- Review performance stats and query execution/explain plans; recommend changes for tuning.
- Create and maintain detailed, up-to-date technical documentation.
- Solve live performance and stability issues and prevent recurrence.

This is the job for you if:

- You have expert level experience in Hadoop-Cloudera administration.
- You have some experience working in a DevOps, CI/CD environment.
- You have knowledge and some experience working with Kubernetes and containerization technologies (Docker).
- You love to attack, analyse, and solve problems with solutions.
- You are excited about implementing new technologies ground up.

- You enjoy working in a fast-paced environment.

Qualifications:

- You have a Bachelor's Degree in Computer Science or equivalent work experience
- You have a minimum 5 years of related experience as a Hadoop Administrator with expert level knowledge of Cloudera Hadoop components including HDFS, Sentry, HBase, Impala, Hue, Spark, Hive, Kafka, YARN, and Zookeeper.
- You have successfully helped in deploying and monitoring Spark jobs.
- You have expertise in administration and implementation of Cloudera Hadoop(CDP), Cloudera Manager, HDFS, Yarn, MapReduce, Hive, Impala, KUDU, Sqoop, Spark, Kafka, HBase, Kerberos, Active Directory, Sentry, TLS/SSL, Linux/RHEL, Unix Windows, SBT, Maven, MS SQL Server, Shell Scripting, Git
- You have prior Hadoop cluster deployment experience in adding and removing nodes, troubleshooting failed jobs, configuring and tuning the clusters and monitoring critical parts of the cluster.
- You are able to set up Linux users and Kerberos principals.
- You have strong Ubuntu Linux administration (security, configuration, tuning, troubleshooting, and monitoring).
- You are experienced in utilizing and implementing ZooKeeper & Kafka Clusters.
- You have strong knowledge of scripting and automation tools and strategies, e.g. Shell, Python, PowerShell, Ansible, Terraform.