

Hadoop Developer Hadoop Developer Hadoop Developer - Vanguard Group Charlotte, NC ? Over 5 years of experience in software design, development, testing and maintenance of web and enterprise applications using Java, python, AWS and Hadoop ? Experienced on Hadoop ecosystem using several components like HDFS, MapReduce, Hive, Kafka, Pig, Hive, Sqoop, Oozie and Spark ? Worked on several AWS components like EMR, S3, EC2 and also familiar with Lambda functions, Kinesis streaming ? Easily adaptable to any new technology and try to explore whenever a situation is needed. ? Can easily coordinate with the small or big team size Authorized to work in the US for any employer Work Experience Hadoop Developer Vanguard Group - Charlotte, NC October 2017 to Present Responsibilities: Built Data Lake in the AWS cloud. Worked on creating an end to end data ingestion pipeline to transfer data on premise to AWS cloud under Agile/Scrum methodology Worked on AWS services like EMR, Cloud Watch, Cloud Formation, EC2, S3, Lambda; Big data technologies like Sqoop, Oozie, Hive, Hue, Tez, Yarn, Zepplin Developed Scala scripts using Dataframes/sql in spark 2.X for data aggregation, queries and writing data back into OLTP system through Sqoop Developed shell scripts to automate the workflows to perform data ingestion, python scripts to generate Json Oozie workflows Experienced in handling large data sets using partitions, Spark in memory capabilities, Effective and efficient joins, Transformations and other during ingestion process Used troposphere in python to auto generate cloud formation templates in JSON format to deploy AWS big data environment Used reporting tool Tableau to connect with Hive for generating daily reports of data Worked with different file formats like TEXTFILE, SEQUENCE FILE, AVRO FILE, ORC and PARQUET for hive querying and processing Worked in migration of several microservices (Webapps and web services) from On-Prem PCF to AWS PCF Worked on setting up Dashboards, monitoring and alerts in Splunk for webapps and web services on AWS Worked on Atlassian CI/CD stack namely Bit Bucket, JIRA, Bamboo, Confluence Hadoop Developer Thales - Melbourne, FL July 2017 to September 2017 Responsibilities: Worked on loading and transforming structured, semi structured and unstructured data Implemented Spark using Scala and Spark SQL for faster testing and processing of data Exploring with Spark, improving performance and optimization of the existing algorithms in Hadoop using Spark Context,

Spark-SQL, Data Frame, and Pair RDD's Created Hive tables, partitions and implemented incremental imports to perform ad-hoc queries on structured data Developed jobs to move inbound files to HDFS file location based on monthly, weekly, daily and hourly partitioning

Analyzed and filtered the logs using Logstash Experienced on tableau desktop/server to visualize the data from csv files and hive server 2.X

Java/ Hadoop Developer Veritis Group Inc - Irving, TX
May 2016 to May 2017 Responsibilities: Ingested files to HDFS using java kafka publisher and consumer API's for streaming application Experienced in using HDFS Java API to put the files to HDFS Implemented cucumber test cases to compare data before and after transformations

Used Pig as to do transformations, event joins, filters and some pre-aggregations before storing the data onto HDFS Created external and internal tables, partitions and data storage format in hive

Worked on Hibernate framework to insert and update the data in MySQL Database Used Log4j for logging purpose and write it to a file for auditing purpose Written shell scripts for file Automation process Used git for code repository and maven to build the project

Java/J2EE Developer Adeptros - Bengaluru, Karnataka January 2014 to June 2014 Responsibilities: Involved in various phases of Software Development Life Cycle (SDLC) such as requirements gathering, modeling, analysis, design, development and testing Used J2EE design patterns for adding new functionality to existing applications Create and execute test cases in JUnit for unit testing of application

Wrote SQL queries to retrieve data from the database using JDBC Utilized frameworks such as Hibernate and Spring for persistence and application layers Utilized object-oriented programming and Java for creating business logic Used Log4J for writing into different logs files Application Log and Error Log Implemented Representational state transfer (REST) Web services for distributed systems to retrieve data from client side Developed UNIX Shell scripts for automating project management tasks To keep track of issues and tasks on individuals, used JIRA ticketing system

Using Eclipse integrated IDE to build the application and git to maintain the version of the files

Involved in Bug fixing of various modules that were raised by the testing teams in the application during integration testing phase

Education M. S in Computer Information Systems Florida Institute of Technology - Melbourne, FL Technology GITAM University - Visakhapatnam, Andhra Pradesh Skills

Hive (1 year), JIRA (2 years), Scala (1 year), sql (2 years), web services (2 years)

Name: James Flores

Email: wangjordan@example.net

Phone: 598-971-2138x329