

Data engineer Data engineer Data engineer - Verizon Atlanta, GA ? 5 years of experience in various IT sectors, which includes hands-on experience in Big Data technologies. ? 4 years of experience as a Hadoop Developer in all phases of Hadoop and HDFS development. ? Hands on experience with HDFS, MapReduce and Hadoop Ecosystem (Pig, NiFi, Hive, Oozie, Hbase, Zookeeper, Flume, and Sqoop) ? Well versed with developing and implementing MapReduce jobs using Hadoop to work with Big Data. ? Have experience with Spark processing Framework such as Spark and Spark Sql. Experience in NoSQL databases like HBase, Cassandra and MongoDB. ? Procedural knowledge in cleansing and analyzing data using HiveQL, Pig Latin, and custom MapReduce programs in Java. ? Experienced in writing custom UDFs and UDAFs for extending Hive and Pig core functionalities. ? Ability to develop Pig UDF'S to pre-process the data for analysis. ? Experience in importing and exporting data using Sqoop from HDFS to Relational Database Systems (RDBMS), Teradata and vice versa. ? Skilled in creating workflows using Oozie for cron jobs. Strong experience in Hadoop Administration and Linux. ? Experienced with Java API and REST to access HBase data. ? Worked extensively with Dimensional modeling, Data migration, Data cleansing, Data profiling, and ETL Processes features for data warehouses. ? Hands on experience in application development using Java, RDBMS, and Linux shell scripting. ? Hands on experience in PERL Scripting and Python. ? Experience working with JAVA, J2EE, JDBC, ODBC, JSP, Java Eclipse, MS SQL Server. ? Extensive experience with SQL, PL/SQL and database concepts. ? Expertise in debugging and optimizing Oracle and java performance tuning with strong knowledge in Oracle 11g and SQL. ? Good experience working with Distributions such as MAPR, Horton works and Cloudera. ? Experience in all stages of SDLC (Agile, Waterfall), writing Technical Design document, Development, Testing and Implementation of Enterprise level Data mart and Data warehouses. ? Having good knowledge on Hadoop Administration like Cluster configuration, Single Node Configuration, Multi Node Configuration, Data Node Commissioning and Decommissioning, Name Node Backup and Recovery, HBase, HDFS and Hive Configuration, Monitoring clusters, Access control List. ? Good Inter personnel skills and ability to work as part of a team. Exceptional ability to learn, master new technologies and to deliver outputs in short deadlines. ? Ability to work

in high-pressure environments delivering to and managing stakeholder expectations. ? Application of structured methods to: Project Scoping and Planning, risks, issues, schedules and deliverables. ? Strong analytical and Problem solving skills. ? Good Inter personnel skills and ability to work as part of a team. Exceptional ability to learn and master new technologies and to deliver outputs in short deadlines

Work Experience

Data engineer Verizon - GA March 2017 to Present

Responsibilities: ? Core person in data ingestion team, involved in designing data flow pipelines and Nifi administration. ? Worked on various data sources and data formats to deliver the data with low latency and accuracy. ? Developed python scripts to monitor and automate Nifi flows using Nifi api. ? Expertise in Nifi to work with various ingestion sources and transforming data on the go. Setup and maintain nifi registry to versioning and CI/CD of flows. ? Worked on creation of hive tables (managed/external) for various use cases. Developed sqoop jobs to import/export data from/to oracle/hdfs. ? Responsible for analyzing and cleansing raw data by performing Hive queries and running Pig scripts on data ? Involved in writing flink jobs to parse near real time data and then push to hive. ? Created Hive tables, loaded data and wrote Hive queries that run within the map. ? Implemented business logic by writing Pig UDF's in Java and used various UDFs from Piggybanks and other sources. ? Used OOZIE Operational Services for batch processing and scheduling workflows dynamically. ? Extensively worked on creating End-End data pipeline orchestration using Oozie. ? Worked on Data Serialization formats for converting Complex objects into sequence bits by using AVRO, PARQUET, JSON, CSV formats. ? Integrated Druid with Hive for High availability and provide data for sla reporting on real time data. ? Developed a framework to extract/load data from/to databases, as a substitute for sqoop. ? Automated splunk indexing to report device and topology metrics.

Environment: Map Reduce, HDFS, Hive, Pig, SQL, Sqoop, Oozie, Shell scripting, Cron Jobs, Apache Nifi, Splunk, Python, Apache Flink, druid, Apache Kafka, J2EE.

Hadoop Developer Bed Bath & Beyond - Union, NJ August 2015 to February 2017

Responsibilities: ? Involved in loading and transforming large sets of structured, semi structured and unstructured data from relational databases into HDFS using Sqoop imports. ? Developed Sqoop scripts to import export data from relational sources and handled incremental loading on the customer, transaction

data by date. ? Developed simple and complex MapReduce programs in Java for Data Analysis on different data formats. ? Developed Spark code using scala and Spark-SQL/Streaming for faster testing and processing of data. ? Import the data from different sources like HDFS/Hbase into Spark RDD. ? Experienced with batch processing of data sources using Apache Spark and Elastic search. ? Experienced in implementing Spark RDD transformations, actions to implement business analysis ? Migrated Hive QL queries on structured into Spark QL to improve performance ? Optimized MapReduce Jobs to use HDFS efficiently by using various compression mechanisms. ? Worked on partitioning HIVE tables and running the scripts in parallel to reduce run-time of the scripts. ? Worked on Data Serialization formats for converting Complex objects into sequence bits by using AVRO, PARQUET, JSON, CSV formats. ? Responsible for analyzing and cleansing raw data by performing Hive/Impala queries and running Pig scripts on data. ? Created Hive tables, loaded data and wrote Hive queries that run within the map. Implemented business logic by writing Hive UDFs in Java. ? Developed Shell scripts and some of Perl scripts based on the user requirement. ? Wrote XML scripts to build OOOIE functionality. ? Used OOOIE Operational Services for batch processing and scheduling workflows dynamically. ? Extensively worked on creating End-End data pipeline orchestration using Oozie. ? Evaluated suitability of Hadoop and its ecosystem to the above project and implementing / validating with various proof of concept (POC) applications to eventually adopt them to benefit from the Big Data Hadoop initiative. Environment: Map Reduce, HDFS, Hive, Pig, SQL, Sqoop, Oozie, Shell scripting, Cron Jobs, Perl scripting, Apache Kafka, J2EE. Java Developer Frost & Sullivan - Chennai, Tamil Nadu May 2012 to June 2013 Responsibilities: ? Used AGILE methodology for developing the application. As part of the lifecycle development prepared class model, sequence model and flow diagrams by analyzing Use cases using Rational Tools. ? Extensive use of SOA Framework for Controller components and view components. ? Involved in writing the exception and validation classes using Struts validation rules. ? Involved in writing the validation rules classes for general server side validations for implementing validation rules as part observer J2EE design pattern. ? Used OR mapping tool Hibernate for the interaction with database. Involved in writing Hibernate queries and Hibernate specific configuration and mapping files. ?

Involved in developing JSP pages and custom tag for presentation layer in Spring framework. ? Developed web services using SOAP and WSDL with Apache Axis 2. ? Developed, implemented, and maintained an asynchronous, AJAX based rich client for improved customer experience using XML data and XSLT templates. ? Developed SQL stored procedures and prepared statements for updating and accessing data from database. ? Development carried out under Eclipse Integrated Development Environment (IDE) ? Used JBoss for deploying various components of application. ? Used JUNIT for testing and check API performance. Involved in fixing bugs and minor enhancements for the front-end modules. Responsible for troubleshooting issues, monitoring and guiding team members to deploy and support the product. Used SVN Version Control for Project Configuration Management. ? Worked with the Android SDK, and implemented Android Bluetooth and Location Connectivity components. ? Worked with business and System Analyst to complete the development in time. ? Implemented the presentation layer with HTML, CSS and JavaScript. ? Developed web components using JSP, Servlets and JDBC. Implemented secured cookies using Servlets. ? Wrote complex SQL queries and stored procedures. Implemented Persistent layer using Hibernate API ? Implemented Transaction and session handling using Hibernate Utils. ? Implemented Search queries using Hibernate Criteria interface. ? Provided support for loans reports for CB&T. Involved in fixing bugs and unit testing with test cases using Junit ? Maintained Jasper server on client server and resolved issues. ? Actively involved in system testing. Fine tuning SQL queries for maximum efficiency to improve the performance ? Designed Tables and indexes by following normalizations. Involved in Unit testing, Integration testing and User Acceptance testing. Utilizes Java and SQL day to day to debug and fix issues with client processes. Environment: Java, Servlets, JSP, Hibernate, Junit Testing, Oracle DB, SQL. Education B.E in Computer Science BITS Pilani - Pilani, Rajasthan M.S in Computer Science Indiana university Purdue University Skills J2EE (5 years), SQL (5 years), APACHE HADOOP HDFS (4 years), APACHE HADOOP OOZIE (4 years), APACHE HADOOP SQOOP (4 years) Additional Information Technical Skills: Technology Hadoop Ecosystem/J2SE/J2EE/Oracle Operating Systems WindowsVista/XP/NT/2000Series, UNIX/LINUX (Ubuntu, CentOS, Redhat) /AIX/Solaris. DBMS/Databases DB2, My SQL, SQL, PL/SQL

Programming Languages C, C++, JSE, XML, Spring, HTML, JavaScript, jQuery, Web services. Big Data Ecosystem HDFS, Nifi, Map Reduce, Oozie, Hive/Impala, Pig, Sqoop, Flume, Zookeeper and Hbase, Spark, Scala Methodologies Agile, Water Fall NOSQL Databases Cassandra, MongoDB, Hbase Version Control Tools SVN, git

Name: James Wright

Email: nhartman@example.com

Phone: +1-236-568-6032