

Hadoop Developer Hadoop Developer Hadoop Developer - Liberty Mutual Dover, NH About 9+years of experience with heavy exposure to Big DataTechnologies, and implementation of various trending technologies in Big Data Eco Systems under various verticals such as Healthcare,Retail and Private Sectors. 4 Years of exposure to Data Science, Machine Learning, Hadoop and Spark Ecosystems implementing innovative solutions for various organizations. Hands on experience in Apache Hadoop ecosystem components like Hive, Pig, Hadoop MapReduce, Sqoop, Flume, Kafka, Storm, Spark, Oozie, Zookeeper, YARN, Impala and HBase along other Hadoop ecosystem components. Authorized to work in the US for any employer Work Experience Hadoop Developer Liberty Mutual - Dover, NH August 2016 to Present Responsibilities: Evaluated suitability of Hadoop and its ecosystem to the above project and implemented various proof of concept (POC) applications to eventually adopt them to benefit from the Big Data Hadoop initiative. Estimated Software & Hardware requirements for the Name Node and Data Node & planning theYARN cluster.

Extracted the needed data from the server into HDFS and Bulk Loaded the cleaned data into HBase and Cassandra. Written the Map Reduce programs, HiveUDFs in Java where the functionality is too complex. Involved in importing and exporting data intoHDFS and Hive using Sqoop. Integrating bulk data into Cassandra file system using MapReduce programs. Develop Hive queries for the analysis, to categorize different items. Designing and creating Hive external tables using shared meta-store instead of derby with partitioning, dynamic partitioning and buckets.

Reviewing peer table creation in Hive, data loading and queries. Used Flume to handle the real time log processing for attribution reports. Developed Pig Latin scripts to extract the data from web server output files to load in HDFS. Used Oozie workflow for Pig and Hive jobs. Involved in working with Spark on top of YARN for interactive and batch analysis. Good understanding on DAG cycle for entire spark application flow on Spark application WebUI. Experienced in writing live Real-time Processing using Spark Streaming with Kafka. Maintained System integrity of all sub-components like Pig, Hive, Spark, Kafka, HBase and Cassandra. Weekly meetings with technical collaborators and active participation in code review sessions with senior and junior developers. Involved in launching, managing and monitoring the Hadoop cluster using

ClouderaManager. Environment: Apache Hadoop, HDFS, HBase, Hive, Pig, MapReduce, Java, Ambari, Cassandra, Sqoop, Flume, Cloudera, Kafka, Scala, Spark, Oozie, MySQL, Linux. Hadoop Developer September 2015 to July 2016 Target, MN Responsibilities: Evaluated business requirements and prepared detailed specifications that follow project guidelines required to develop suitable programs. Worked on analyzing HadoopCluster and different big data analytic tools including Pig, Hive and MongoDB. Involved in importing data from Oracle into HDFS and Hive using Sqoop. Created tables using Impala, and involved in creating Queries which are stored in HBase. Implemented complex scripts to support test driven development and continuous integration. Installed and configured Hadoop, MapReduce, and HDFS and developed multiple MapReduce jobs in Java for data. Involved in data modeling the large data from various sources using NoSQL databases HBase, MongoDB. Created Pig Latin scripts to sort, group, join and filter the enterprise wide to get transformed data sets. Involved in creating Hive tables, loading data and creating Hive queries that will run internally in MapReduce way. Used Hive to analyze the partitioned and bucketed data and compute various metrics for reporting. Experience in Hive-HBase integration by defining external tables in hive and pointing the HBase as data store for better performance and lower I/O. Exported the analyzed data to relational databases using Sqoop for visualization and to generate reports for the BI team. Hands on experience in Tableau which is used for visualization and reports generation. Involved in maintaining and debugging MR programs. Mentoring analyst and test team for writing Hive queries. Installed Oozie workflow engine to run multiple MapReduce jobs. Worked with application teams to install operating systems, Hadoop updates, patches, version upgrades as required. Good experience in monitoring and managing the Hadoop cluster using Cloudera Manager. Environment: Hadoop, HDFS, MapReduce, Hive, Eclipse, Pig, Sqoop, Linux, Flume, Java, Shell Scripting, Tableau, Python, Zookeeper, HBase, Informatica, MongoDB, Cloudera, Oozie. Hadoop Developer Home Depot - Atlanta, GA June 2013 to August 2015 Responsibilities: Responsible for building scalable distributed data solutions using Hadoop. Implemented 20 nodesCDH4 Hadoop cluster on LINUX. Involved in loading data from LINUX file system to HDFS. Estimated Software & Hardware

requirements for the NameNode and DataNode& planning the cluster. Implemented a script to transmit data from Teradata to HBase using Sqoop and created tables out of that data on HBase. Implemented best income logic using Pigscripts and UDFs. Implemented test scripts to support test driven development and continuous integration. Worked on tuning the performance Pig queries. Cluster Coordination service through Zookeeper. Load and transform large sets of structured, semi structured and unstructured data. Responsible for cluster maintenance by Cloudera, adding and removing cluster nodes, cluster monitoring and troubleshooting, manage and review data backups, manage and review Hadoop log files. Installing Oozie workflow engine to run multiple Hive and Pig jobs. Job management using Fair scheduler. Exported the analyzed data to the relational databases using Sqoop for visualization and to generate reports for the BI team. Analyzed large amounts of data sets to determine optimal way to aggregate and report on it. Supported in setting up QA environment and updating configurations for implementing scripts with Pig and Sqoop. Environment: Hadoop, HDFS, Hive, Pig, Sqoop, HBase, Shell Scripting, Ubuntu, Linux, Zookeeper, Teradata, Oozie, Cloudera. Java Developer ChevronCorporation - Irvine, CA November 2010 to May 2012 Responsibilities: Worked with business analyst in understanding business requirements, design and development of the project. Implemented the JSP frame work with MVC architecture. Created new JSP's for the front end using HTML, Java Script, JQuery, and Ajax. Developed the presentation layer using JSP, HTML, CSS and client side validations using JavaScript. Involved in creating Restful web services using JAX\_RS and JERSEY tool. Involved in designing, creating, reviewing Technical Design Documents. Developed DAOs (Data Access Object) using Hibernate as ORM to interact with DBMS - Oracle. Applied J2EE design patterns like Business Delegate, DAO and Singleton. Involved in developing DAO's using JDBC. Worked with QA team in preparation and review of test cases. JUnit was used for unit testing for the integration testing tool. Writing SQL queries to fetch the business data using Oracle as database. Developed UI for Customer Service Modules and Reports using JSF, JSP's and My Faces Components. Log4j used for logging the application log of the running system to trace the errors and certain automated routine functions. Environment: Java, JSP, JavaScript, Servlets, Hibernate,

REST, EJB, JSF, JSP, Ant, Tomcat, Eclipse, SQL, Oracle. Jr. Java Developer July 2008 to October 2010 Project: Employee Billing Management System (EBMS) Responsibilities: Worked and maintained the front end functionality of the Website from scratch Using HTML, CSS and Java Script. Developed server side functionality of the website using JAVA and SPRING MVC. Involved in discussions with clients to gather requirements to fulfill the objectives. Interacted with several teams, to implement the requirements. I was responsible for preparing the technical documentation. Environment: JAVA and SPRING MVC, HTML, CSS and Java Script. Education Hadoop Architecture Secondary NameNode Bachelor's Skills JAVA (6 years), APACHE HADOOP HDFS (4 years), APACHE HADOOP OOZIE (4 years), APACHE HADOOP SQOOP (4 years), APACHE HBASE (4 years), Apache Hadoop, HDFS, HBase, Hive, Pig, MapReduce, Java, Ambari, Cassandra, Sqoop, Flume, Cloudera, Kafka, Scala, Spark, Oozie, MySQL, Linux. Additional Information TECHNICAL SKILLS: Hadoop HDFS, YARN, Map Reduce, Pig, Hive, Sqoop, Flume, Hbase, MongoDB, Oozie, Zookeeper, Kafka, Spark No SQL Database HBase, MongoDB, Cassandra Databases MS SQL Server, Oracle 11g/10g/9i/,DB2, MySQL, MS Access JEE Technologies J2EE, JSP, Servlets, JUnit and JDBC Frameworks Struts, Hibernate, Spring IOC, Spring AOP and Spring JDBC ETL Tools Pentaho, Tableau Operating systems Ubuntu, Windows, iOS, Unix languages C, Java, Python and Shell scripting

Name: Eric Herrera

Email: tmcdonald@example.org

Phone: 4516354940