Sr. Hadoop BigData Engineer Sr. Hadoop BigData Engineer Sr. Hadoop BigData Engineer - Liberty Mutual Insurance Dover, NH Work Experience Sr. Hadoop BigData Engineer Liberty Mutual Insurance - Dover, NH October 2017 to Present Roles & Responsibilities: ? Worked with business partners in discussing the requirements for new projects and enhancements to the existing applications. ? Worked with application teams to install operating system, Hadoop updates, patches, version upgrades as required. ? Developed MapReduce programs to parse the raw data, populate tables and store the refined data in partitioned tables. ? Used Spark API over Cloudera Hadoop YARN to perform analytics on data in Hive. ? Involved in creating Hive Tables, loading with data and writing Hive queries which will invoke and run MapReduce jobs in the backend. ? Developed Pig Latin Scripts to process data in a batch to perform trend analysis. ? Wrote ETL jobs to read from web APIs using REST and HTTP calls and loaded into Cassandra using Java and Talend. ? Analyzed Cassandra database and compared it with other open-source NoSQL databases to find which one of them better suites the current requirement. ? Responsible for Cluster maintenance, Monitoring; commissioning/decommissioning Data Nodes; troubleshooting, manage and reviewing Data Backups and Log Files. ? Ingested data from RDBMS and performed data transformations, and then export the transformed data to Cassandra as per the business requirement. ? Developed automated processes for flattening the upstream data from Cassandra which in JSON format. Used Hive UDFs to flatten the JSON Data. ? Configured Spark Streaming to receive real time data from the Kafka and store the stream data to the database. ? Responsible for developing data pipeline by implementing Kafka producers and consumers. ? Connected to AWS EC2 using SSH and ran Spark-submit jobs. ? Responsible for developing data pipeline with AWS to extract the data from weblogs and store in the database. ? Analyzed user request patterns and implemented various performance optimization measures including implementing partitions and buckets in HiveQL. ? Studied data by performing HiveQL & running Pig Latin scripts to study customer behavior. Environment: Hadoop, Hive, MapReduce, Pig Latin, REST, Java, Cassandra, JSON, Spark, AWS, EC2, HiveQL, Oozie Hadoop Developer NorthShore Medical Group - Lincolnwood, IL November 2015 to August 2017 Roles & Responsibilities: ? Worked on a product

team using Agile/SCRUM methodology to develop, deploy and support solutions that leverage the Client big data platform   ? Documented the systems processes and procedures for future references. ? Supported technical team in management and review of Hadoop log files and data backups. ? Involved in writing MapReduce program and Hive queries to load and process data in Hadoop File System. ? Involved in creating Hive tables, loading with data and writing hive queries which will run internally in MapReduce way. ? Developing and running MapReduce jobs on YARN and Hadoop clusters to produce daily and monthly reports as per user's need. ? Maintain and schedule periodic jobs which range from updates on MapReduce jobs to creating ad-hoc jobs for the business users. ? Scheduling and managing jobs on a Hadoop cluster using Oozie work flow. ? Handled importing data from different data sources into HDFS using Sqoop and performing transformations using Hive, MapReduce and then loading data into HDFS. ? Exporting of result set from HIVE to MySQL using Sqoop export tool for further processing. ? Collecting and aggregating large amounts of log data using Flume and staging data in HDFS for further analysis. ? Developed several REST web services supporting both XML and JSON to perform task such as demand response management. ? Created Maven builds to build and deploy Spring Boot microservices to internal enterprise Docker registry. ? Designed target tables as per the requirement from the reporting team and designed Extraction, Transformation and Loading using Talend. ? Implemented File Transfer Protocol operations using Talend Studio to transfer files in between network folders. ? Optimized MapReduce Jobs to use HDFS efficiently by using various compression mechanisms. ? Responsible for continuous monitoring and managing Elastic MapReduce (EMR) cluster through AWS console. ? Executed speedy reviews and first mover advantages by using workflows like Oozie in order to automate the data. ? Involved in loading data from UNIX file system to HDFS. Environment: Hadoop, MapReduce, HDFS, Hive, Oracle, Java, AWS, Servlets, HTML, XML, SQL, J2EE, JUnit, Tomcat, UNIX, Maven, REST Hadoop Developer Barnes & Noble - Brentwood, TN March 2014 to September 2015 Roles & Responsibilities: ? Involved in developing Hadoop Map Reduce jobs using Java Runtime Environment for the batch processing to search and match the scores. ? Executed speedy reviews and first mover advantages by using workflows like Oozie in

order to automate the data. ? Loading process into the Hadoop distributed File System (HDFS) and Pig language in order to preprocess the data. ? Integrated Oozie with the rest of the Hadoop stack supporting several types of Hadoop jobs out of the box such as Map-Reduce, Pig, Hive, Sqoop, Flume. ? Wrote Hive jobs to parse the logs and structure them in tabular format to facilitate effective querying on the log data. ? Worked with Sqoop to import export data from HDFS to Relational Database system. ? Performed performance tuning and troubleshooting of Map Reduce jobs by analyzing and reviewing Hadoop log files. ? Developed Pig scripts for analyzing large data sets in the HDFS. ? Collecting and aggregating large amounts of log data using Apache Flume and staging data in HDFS for further analysis. ? Designed and built the Reporting Application, which uses the Spark SQL to fetch and generate reports on HBase table data. ? Extracted data from the server into HDFS and Bulk Loaded the cleaned data into HBase. ? Extracted data from the flat files and other RDBMS databases into staging area and populated onto Data warehouse. ? Worked on JVM performance tuning to improve Map-Reduce jobs performance. ? Developed programs to manipulate data & perform CRUD operations on request to database. ? Developed several REST web services supporting both XML and JSON to perform task such as demand response management. ? Used message driven beans for asynchronous processing alerts to the customer. ? Worked on developing Use Cases, Class Diagrams, Sequence diagrams, and Data Models. ? Performed performance tuning and troubleshooting of Map Reduce jobs by analyzing and reviewing Hadoop log files. Environment: Hadoop, Hive, HBase, Linux, MapReduce, HDFS, Hive, Java (JDK), Cloudera, MapReduce, DataStax, IBM DataStage, UNIX Shell Scripting Java Developer Charter Communications - St. Louis, MO February 2012 to December 2013 Roles & Responsibilities: ? Interacted with business analyst to understand the requirements to ensure correct modules been built to meet business requirements. ? Participated in the daily SCRUM meetings to produce quality enhancements within time. ? Developed UML using Case diagrams, Class diagrams, and Sequence diagrams using Rational Software Architect ? Spring MVC model integration for front-end request action controller. ? Developed web screens in JSP, JSTL, CSS and client-side validation using jQuery. ? Developed Web services to allow communication between application through

SOAP over HTTP using Apache CXF ? Configured JMS on Web Sphere Server for asynchronous messaging through implementation of Message Driven Beans (MDB). ? Used Spring ORM module for integration with Hibernate for persistence layer. ? Implemented the application using the concrete principles laid down by several design patterns such as Session Fa ade, Business Delegate, Singleton, Data Access Object, and Service Locator. ? Developed the application in J2EE Application Server environment with IBM WebSphere as deployment server with RAD as development IDE. ? Used JIRA for defect tracking and project management. ? Developed and designed XML Schemas to transport and store data. XML was used to simplify data and allow for Platform Changes, as well as making data more available across the applications distributed platforms. ? Extensively used XSLT to transform XML documents to HTML. ? Wrote custom jQuery plugins and developed JavaScript functions to build a bleeding-edge, AJAX-driven user interface. ? Developed unit and functional test cases using J-Unit. ? Maven and Jenkins used for the automatic build process. ? Used Log4J utility to log error, info and debug messages. ? Used Rational Clear Case for version controlling. ? Worked efficiently in a very tight schedule to meet the deadlines Environment: Java, J2EE, Spring, MVC, Hibernate, HTML, CSS, AJAX, jQuery, JavaScript, JIRA, XML, J-Unit, Maven, Jenkins, Log4J Education Bachelor's Degree in Business Administration in Business Administration Tribhuvan University 2018 MBA in Information Technology Johnson & Wales University - Providence, RI 2012 Skills Cassandra, Ambari, Hdfs, Impala, Oozie, Sqoop, Hbase, Kafka, Db2, Flume, Map reduce, Apache spark, Hbase, Hive, Javascript, Json, Pig, Python, Xml, Zookeeper

Name: Jonathon Bray

Email: wilsonmichele@example.com

Phone: 220-865-3260x763