Sr. Big Data/Hadoop Developer Sr. Big Data/Hadoop Developer Sr. Big Data/Hadoop Developer - NetApp Durham, NC ? Over 9+ years of professional IT experience of Big Data Hadoop Ecosystems experience in ingestion, storage, querying, processing and analysis of big data. ? Experience in using Maven for building and deploying J2EE Application archives (Jar and War) on Web Logic, IBM Web Sphere. ? Experience in building Pig scripts to extract, transform and load data onto HDFS for processing. ? Experience in importing and exporting data using Sqoop from HDFS to Relational Database Systems and vice-versa. ? Experience in Hadoop Shell commands, writing MapReduce Programs, verifying managing and reviewing Hadoop Log files. ? Strong background in Java/J2EE environments. Well experienced in MVC architecture of Spring framework. ? Experience in Big Data analysis using PIG and HIVE and understanding of SQOOP and Puppet. ? Experience with leveraging Hadoop ecosystem components including Pig and Hive for data analysis, Sqoop for data migration, Oozie for scheduling and HBase as a NoSQL data store. ? Good Exposure on Apache Hadoop MapReduce programming, PIG Scripting and Distribute Application and HDFS. ? Expert in developing applications using all J2EE technologies like Servlets, JSP, JDBC, JNDI, JMS. ? Experience in Object Oriented Analysis, Design (OOAD) and development of software using UML Methodology, good knowledge of J2EE design patterns and Core Java design patterns. ? Hands on experience in application development using Java, RDBMS, and Linux shell scripting. ? Strong work ethics with desire to succeed and make significant contributions to the organization. ? Load streaming log data from various web servers into HDFS using Flume. ? Experience in deployment of Hadoop Cluster using Puppet tool. ? Experience in scheduling Cron jobs on EMR, Kafka, and Spark using Clover Server. ? Proficient in using RDMS concepts with Oracle, SQL Server and MySQL. ? Hands on experience with build and deploying tools like Maven and GitHub using Bash scripting. ? Hands on experience with spring tool suit for development of Scala Applications. ? Extensive experience working with structured data using Spark SQL, Data frames, Hive QL, optimizing queries, and incorporate complex UDF's in business logic. ? Experience working with Text, Sequence files, XML, Parquet, JSON, ORC, AVRO file formats and Click Stream log files. ? Experience working with Data Frames, RDD, Spark SQL, Spark Streaming, APIs, System

Architecture, and Infrastructure Planning. ? Experience with Core Java component Collection, Generics, Inheritance, Exception Handling and Multi-threading. ? Developed Java applications using various IDE's like Spring Tool Suite and Eclipse. ? Experience in usage of Hadoop distribution like Cloudera and Hortonworks. ? Strong knowledge in working with UNIX/LINUX environments, writing shell scripts and PL/SQL Stored Procedures. ? Good knowledge in using Hibernate for mapping Java classes with database and using Hibernate Query Language (HQL). ? Operated on Java/J2EE systems with different databases, which include Oracle, MySQL and DB2. ? Knowledge on implementing Big Data in Amazon Elastic MapReduce (Amazon EMR) for processing, managing Hadoop framework dynamically scalable Amazon EC2 instances. ? Build AWS secured solutions by creating VPC with private and public subnets. ? Extensive experience in Application servers likes Web logic, Web Sphere, JBoss, Glassfish and Web Servers like Apache Tomcat. Authorized to work in the US for any employer Work Experience Sr. Big Data/Hadoop Developer NetApp - Durham, NC, US September 2017 to Present Responsibilities: ? Worked as a Sr. Big Data/Hadoop Developer with Hadoop Ecosystems components like HBase, Sqoop, Zookeeper, Hive and Pig with Cloudera Hadoop distribution. ? Worked with data science team to build statistical model with Spark MLLIB and Pyspark. ? Actively involved in designing Hadoop ecosystem pipeline. ? Developed Spark code using Scala and Spark-SQL/Streaming for faster testing and processing of data. ? Implemented Security in Web Applications using Azure and deployed Web Applications to Azure. ? Worked in Agile development environment having KANBAN methodology. ? Used Kibana, which is an open source based browser analytics and search dashboard for Elastic Search. ? Developed Spark code and Spark-SQL/Streaming for faster testing and processing of data. ? Used Java Persistence API (JPA) framework for object relational mapping which is based on POJO Classes. ? Responsible for fetching real time data using Kafka and processing using Spark and Scala. ? Worked on Kafka to import real time weblogs and ingested the data to Spark Streaming. ? Developed business logic using Kafka Direct Stream in Spark Streaming and implemented business transformations. ? Actively involved in daily scrum and other design related meetings. ? Maintained Hadoop, Hadoop ecosystems, and database with

updates/upgrades, performance tuning and monitoring. ? Extensively used JQuery to provide dynamic User Interface and for the client side validations. ? Responsible for defining the data flow within Hadoop eco-system and direct the team in implement them. ? Worked extensively with importing metadata into Hive and migrated existing tables and applications to work on Hive and Spark. ? Build large-scale data processing systems in data warehousing solutions, and work with unstructured data mining on NoSQL. ? Worked with application teams to install operating system, Hadoop updates, patches, version upgrades as required. ? Specified the cluster size, allocating Resource pool, Distribution of Hadoop by writing the specification texts in JSON File format. ? Created Hive tables, and loading and analyzing data using hive queries. ? Wrote Hive Queries for analyzing data in Hive warehouse using Hive Query Language (HQL). ? Developed Hive queries to process the data and generate the data cubes for visualizing. ? Involved in running Hadoop jobs for processing millions of records of text data. ? Created and maintained Technical documentation for launching Hadoop Clusters and for executing Hive queries. ? Used Struts which is an open source MVC framework for creating elegant, modern java web applications. ? Developed customized Hive UDFs and UDAFs in Java, JDBC connectivity with hive development and execution of Pig scripts and Pig UDF's. ? Used Hadoop YARN to perform analytics on data in Hive. ? Developed and maintained batch data flow using HiveQL and Unix scripting ? Used Oozie and Zookeeper operational services for coordinating cluster and scheduling workflows. ? Involved in converting Hive/SQL queries into Spark transformations using Spark RDD, Scala and Python. ? Configured Spark streaming to receive real time data from Kafka and store the stream data to HDFS using Scala. ? Continuous coordination with QA team, production support team and deployment team. ? Developed SQL scripts using Spark for handling different data sets and verifying the performance over MapReduce jobs. ? Used J2EE design patterns like Factory pattern & Singleton Pattern. ? Involved in converting MapReduce programs into Spark transformations using Spark RDD's using Scala and Python. ? Developed and execute data pipeline testing processes and validate business rules and policies ? Built code for real time data ingestion using Java, MapR-Streams (Kafka) and STORM. Environment: HBase 1.4, Sqoop 1.4, Zookeeper 3.4, Oozie 4.3, Hive 2.3, Hadoop 3.0,

Scala 2.12, Spark 2.3, Python 3.7, MS Azure, Agile, NoSQL, JSON, JDBC 4.3, Java, Unix, MapR, Kafka 2.0, Storm 1.0.5 Sr. Hadoop Developer 3M Health - Troy, NY, US May 2016 to August 2017 Responsibilities ? Involved in all phases of Software Development Life Cycle (SDLC) and Worked on all activities related to the development, implementation and support for Hadoop. ? Installed and Configured Apache Hadoop clusters for application development and Hadoop tools like Hive, Pig, HBase, Zookeeper and Sqoop. ? Involved in Agile methodologies, daily scrum meetings, spring planning. ? Worked on MongoDB by using CRUD (Create, Read, Update and Delete), Indexing, Replication and Shading features. ? Developing data pipeline using Flume, Sqoop, Pig and Java MapReduce to ingest customer behavioral data and financial histories into HDFS for analysis. ? Worked on MongoDB, HBase (NoSQL) databases which differ from classic relational databases ? Worked on creating data models for Cassandra from Existing Oracle data model. ? Designed Column families in Cassandra and Ingested data from RDBMS, performed data transformations, and then export the transformed data to Cassandra as per the business requirement. ? Responsible for installation and configuration of Hive, HBase and Sqoop on the Hadoop cluster. ? Configured Spark Streaming to receive real time data from the Apache Kafka and store the stream data to HDFS using Scala. ? Implemented usage of Amazon EMR for processing Big Data across a Hadoop Cluster of virtual servers on Amazon Elastic Compute Cloud (EC2) and Amazon Simple Storage Service (S3). ? Implemented Apache Nifi flow topologies to perform cleansing operations before moving data into HDFS. ? Worked extensively with importing metadata into Hive and migrated existing tables and applications to work on Hive and AWS cloud. ? Implemented a distributed messaging queue to integrate with Cassandra using Apache Kafka and Zookeeper. ? Involved in loading the real-time data to NoSQL database like Cassandra. ? Worked on connecting Cassandra database to the Amazon EMR File System for storing the database in S3. ? Developed UDF's using Scala Scripts, which used in Data frames/SQL and RDD in Spark for Data Aggregation, queries and writing data back into RDBMS through Sqoop. ? Deployed the project on Amazon EMR with S3 connectivity for setting a backup storage. ? Imported data from AWS S3 and into spark RDD and performed transformations and actions on RDD's. ? Responsible for developing data

pipeline with Amazon AWS to extract the data from weblogs and store in Amazon EMR. ? Implemented multiple MapReduce Jobs in java for data cleansing and pre-processing. ? Analyzed large amounts of data sets to determine optimal way to aggregate and report on it. ? Installed and configured local Hadoop Cluster with 3 nodes and set up 4 nodes cluster on EC2 cloud. ? Designed and implemented MapReduce based large-scale parallel relation-learning system. ? Developed the code which will create XML files and Flat files with the data retrieved from Databases and XML files.

Environment: Hadoop 3.0, Hive 2.3, Pig 0.17, HBase 1.4, Zookeeper 3.4, Sqoop 1.4, Agile, MongoDB 4.0, Spark 2.3, Scala 2.12, Apache Kafka 2.0, Nifi 1.7, HDFS, AWS, Cassandra 3.11, NoSQL, java, XML Sr. Java/Hadoop Developer T-Mobile - Bellevue, WA, US January 2015 to April 2016 Responsibilities: ? Designed and Developed application modules using spring and Hibernate frameworks. ? Responsible for building scalable distributed data solutions using Hadoop. ? Experienced in loading and transforming of large sets of structured, semi structured and unstructured data. ? Used MAVEN for developing build scripts and deploying the application onto WebLogic. ? Implemented Spark RDD transformations to map business analysis and apply actions on top of transformations. ? Developed Spark jobs and Hive Jobs to summarize and transform data. ? Involved in converting Hive/SQL queries into Spark transformations using Spark data frames, Scala and Python. ? Implemented MVC architecture using Spring Framework, Coding involves writing Action Classes/Custom Tag Libraries, JSP. ? Expertise in implementing Spark Scala application using higher order functions for both batch and interactive analysis requirement. ? Creating Hive tables with periodic backups, writing complex Hive/Impala queries to run on Impala. ? Implemented partitioning, bucketing and worked on Hive, using file formats and compressions techniques with optimizations. ? Involved in designing and developing modules at both Client and Server Side. ? Worked on JDBC framework encapsulated using DAO pattern to connect to the database. ? Developed the UI Screens using JSP and HTML and did the client side validation with the JavaScript. ? Worked on various SOAP and RESTful services used in various internal applications. ? Developed JSP and Java classes for various transactional/ non-transactional reports of the system using extensive SQL queries. ? Worked on analyzing Hadoop cluster and different

big data analytic tools including MapReduce, Hive and Spark. ? Implemented Storm topologies to pre-process data before moving into HDFS system. ? Implemented POC to migrate MapReduce programs into Spark transformations using Spark and Scala. ? Involved in configuring builds using Jenkins with Git and used Jenkins to deploy the applications onto Dev, QA environments ? Involved in unit testing, system integration testing and enterprise user testing using JUnit. ? Involved in creating Hive tables, loading with data and writing Hive queries which runs internally in MapReduce way. ? Developed Shell, Perl and Python scripts to automate and provide Control flow to Pig scripts. Environment: spring 4.0, Hibernate 5.0.7, Hadoop 2.6.5, Spark 1.1, Hive, Python 3.3, Scala, Sqoop, Flume 1.3.1, Impala, MapReduce, LINUX Sr. Java/J2EE Developer JPMC - New York, NY, US August 2012 to December 2014 Responsibilities: ? As a Java/J2ee developer involved in back-end and front-end developing team. ? Designed and developed various modules of the application with J2EE design architecture, frameworks like Spring MVC architecture and Spring Bean Factory using IOC, AOP concepts. ? Implemented Java/J2EE design patterns such as Factory, DAO, Session Fa ade, and Singleton. ? Used Hibernate in persistence layer and developed POJO's, Data Access Object (DAO) to handle all database operations. ? Used Maven as the build tool, GIT for version control, Jenkins for Continuous Integration and JIRA as a defect tracking tool. ? Developed the user interface components using HTML, CSS, JavaScript, AJAX, JQuery and also created custom tags. ? Implemented the Project structure based on Spring MVC pattern using spring boot. ? Worked on JavaScript to validate input, manipulated HTML elements using JavaScript. ? Developed external JavaScript codes that can be used in several different web pages. ? Developed Web pages using JSP, HTML, CSS, Struts Tag libs and AJAX for the Credit Risk module. ? Used Spring Beans to encapsulate business logic and Implemented Application MVC Architecture using Spring MVC framework. ? Developed XMLs, JavaScript and Java classes for dynamic HTML generation to perform the server side processing on the client requests. ? Used JUnit framework for unit testing of application and Maven to build the application and deployed on Jetty server. ? Developed unit test cases in JUnit and documented all the test scenarios as per the user specifications. ? Implemented Spring framework based on the Model View Controller design

paradigm. ? Involved in development activities using Core Java /J2EE, Servlets, JSP, JSF used for creating web application, XML and springs. ? Used Spring Framework for Dependency injection and integrated with the Hibernate. ? Developed RESTful Web Services client to consume JSON messages using Spring JMS configuration. ? Developed services using Spring IOC and Hibernate persistence layer with Oracle Database. ? Implemented build script using ANT for compiling, building and deploying the application on WebSphere application server.   Environment: J2EE, Java, Spring MVC 3.0, POJO, Jenkins, HTML, JavaScript, AJAX, JQuery, CSS, XML, Maven, JUnit, Hibernate 4.2.8, POJO, ANT, Oracle 10g Java Developer Valuelabs September 2009 to July 2012 Responsibilities: ? Applied Model-View-Controller (MVC) design pattern and Single Tone class design pattern for designing the application. ? Developed JSP pages with Struts and EJB (Enterprise JavaBeans) for implementing different search pages for transaction of each module. ? Used JavaScript and struts validation framework for client side validation. ? Implemented Ant and Maven build tools to build jar and war files and deployed war files to target servers. ? Responsible for writing Struts action classes, Hibernate POJO classes and integrating Struts and Hibernate with spring for processing business needs. ? Developed web application using Struts, JSP, Servlets, and JavaBeans that uses MVC design pattern. ? Used Maven to build the application and deployed on IBM WebSphere Application Server. ? Created bean XML files and row Mappers to map tables and fields in the database. ? Wrote Apache ANT build scripts for building the application and unit test cases using JUnit for performing the unit testing. ? Designed the user interfaces using JSPs, developed custom tags, and used JSTL Tag lib. ? Involved in creating the Hibernate POJO Objects and mapped using Hibernate Annotations. ? Created UML diagrams (use case, class, sequence, and collaboration) based on the business requirements. ? Responsible for designing and developing of Object oriented methodologies using UML. ? Developed Web Services to allow communication between the applications using Rest Web Services. ? Created JUnit test cases for unit testing the code at minute level and used Eclipse IDE. ? Used AJAX frameworks like JQuery, JSON to develop rich GUIs and also involved in performance tuning the website. ? Developed ANT scripts to build and deploy the application in the JBOSS Application Server.   ? Developed

Message-Driven beans in collaboration with Java Messaging Service (JMS). ? Implemented server pages using Apache Tomcat as application server and Log4j for application logging and debugging. ? Developed the presentation layer and content management framework using HTML and JavaScript. ? Extensively used MVC architecture and JBoss for deployment purposes. ? Created JSP, Form Beans for effective way of implementing Model View Controller architecture. ? Developed Servlets to perform business logic and to interact with the database using JDBC.

Environment: JavaScript, Ant, Maven, Struts, JavaBeans, MVC, XML, Apache ANT, Hibernate 4.1.9, POJO, JUnit, Eclipse, JQuery, JSON, AJAX, JBOSS, Apache Tomcat, HTML Education Bachelor's Skills database (9 years), Java (9 years), JavaScript. (6 years), JQuery (6 years), JSON (6 years) Additional Information Hadoop, Big Data, Cloud, Nodes, Azure, Elastic Cloud, Sqoop, Flume, Yarn, Spark, Hive, Hue, Apache Camel, SQL, MySQL, PostgreSQL, MongoDB, HBase, Cassandra and JQuery Technical Skills: ? Hadoop/Big Data Technologies: Hadoop 3.0, HDFS, MapReduce, HBase 1.4, Apache Pig 0.17, Hive 2.3, Sqoop 1.4, Apache Impala 3.0, Oozie 4.3, Yarn, Apache Flume 1.8, Kafka 1.1, Zookeeper 3.4 ? Hadoop Distributions: Cloudera , Hortonworks, MapR ? Cloud: AWS, Azure, Azure SQL Database, Azure SQL Data Warehouse, Azure Analysis Services, HDInsight, Azure Data Lake and Data Factory. ? Programming Language: Java, Scala 2.12, Python 3.6, SQL, PL/SQL, Shell Scripting, Storm 1.0, JSP, Servlets ? Frameworks: Spring 5.0.5, Hibernate 5.2, Struts 1.3, JSF, EJB, JMS ? Web Technologies: HTML5, CSS, JavaScript, JQuery 3.3, Bootstrap 4.1, XML, JSON, AJAX ? Databases: Oracle 12c/11g, SQL ? Database Tools: TOAD, SQL PLUS, SQL ? Operating Systems: Linux, Unix, Windows 10/8/7 ? IDE and Tools: Eclipse 4.7, NetBeans 8.2, IntelliJ, Maven ? NoSQL Databases: HBase 1.4, Cassandra 3.11, MongoDB ? Web/Application Server: Apache Tomcat 9.0.7, JBoss, Web Logic, Web Sphere ? SDLC Methodologies: Agile, Waterfall ? Version Control: GIT, SVN, CVS

Name: Mr. Michael Hill

Email: brownlaura@example.com

Phone: +1-512-397-3469x6966