Big Data Engineer Big Data Engineer Big Data Engineer - AT&T Atlanta, GA ? 9+ years of experience in IT industry with extensive experience in Java, J2ee and Big data technologies. ? 4 +years working of exclusive experience on Big Data technologies and Hadoop stack ? Strong experience working with HDFS, MapReduce, Spark. Hive, Pig, Sqoop, Flume, Kafka, Oozie and HBase. ? Good understanding of distributed systems, HDFS architecture, internal working details of MapReduce and Spark processing frameworks. ? More than two years of hands on experience using Spark framework with Scala. ? Good exposure to performance tuning hive queries, map-reduce jobs, spark jobs. ? Expertise in Inbound and Outbound (importing/exporting) data form/to traditional RDBMS using SQOOP. ? Tuned PIG and HIVE scripts by understanding the joins, group and aggregation between them. ? Extensively worked on HiveQL, join operations, writing custom UDF's and having good experience in optimizing Hive Queries. ? Worked on various Hadoop Distributions (Cloudera, Hortonworks, Amazon AWS) to implement and make use of those. ? Participated in design, development and system migration of high performance metadata driven data pipeline with Kafka and Hive/Presto on Qubole, providing data export capability through API and UI. ? Experience data processing like collecting, aggregating, moving from various sources using Apache Flume and Kafka. ? Hands on experience in NOSQL databases and SQL databases. Authorized to work in the US for any employer Work Experience Big Data Engineer AT&T - Atlanta, GA July 2016 to Present Responsibilities: ? Extensively involved in Design phase and delivered Design documents in Hadoop eco system with HDFS, HIVE, PIG, SQOOP and SPARK with SCALA. ? Collected the logs from the physical machines and the Open Stack controller and integrated into HDFS using Kafka. ? Involved in the high-level design of the Hadoop architecture for the existing data structure and Business process ? Part of Configuring & deployment of Hadoop Cluster in the AWS cloud. ? Worked with clients to better understand their reporting and dash boarding needs and present solutions using structured Agile project methodology approach. ? Worked on analyzing Hadoop cluster and different Big Data Components including Pig, Hive, Spark, HBase, Kafka, Elastic Search, database and SQOOP. ? Involved in loading disparate datasets into Hadoop Data Lake, this would be available to the data science team to predict the future. ? Implemented

Partitioning, Dynamic Partitions and Buckets in HIVE for increasing performance benefit and helping in organizing data in a logical fashion. ? Installed Hadoop, Map Reduce, HDFS, and developed multiple Map-Reduce jobs in PIG and Hive for data cleaning and pre-processing. ? Created tables in HBase to store variable data formats of PII data coming from different portfolios ? Worked on Sequence files, RC files, Map side joins, bucketing, partitioning forHive performance enhancement and storage improvement. ? Experienced in pulling the data from Amazon S3 bucket to Data Lake and builtHive tables on top of it and created data frames in Spark to perform further analysis. ? Used cloud computing on the multi-node cluster and deployed Hadoop application on cloud S3 and used Elastic Map Reduce (EMR) to run aMapReduce. ? Explored MLlib algorithms in Spark to understand the possible Machine Learning functionalities that can be used for use case. ? In preprocessing phase of data extraction, we used Spark to remove all the missing data for transforming of data to create new features. ? Developed data pipeline using Flume, Sqoop, Pig and Java map reduce to ingest customer behavioral data and financial histories into HDFS for analysis. ? Involved in loading data from UNIX file system to HDFS using Flume and HDFSAPI. ? Configured Spark Streaming to receive real time data from the Kafka and storethe stream data to HDFS. ? Developed RDD's/Data Frames in Spark using Scala and Python and applied several transformation logics to load data from Hadoop Data Lake to Cassandra DB. ? Exported the analyzed data to the NoSQL Database using HBase for visualization and to generate reports for the Business Intelligence team using SAS. ? Used various HBase commands and generated different Datasets as per requirements and provided access to the data when required using grant and Revoke ? Created Hive tables as per requirement as internal or external tables, intended for efficiency. ? Developed MapReduce programs for the files generated by hive query processing to generate key, value pairs and upload the data to NoSQL database HBase. ? Implemented installation and configuration of multi-node cluster on the cloud using Amazon Web Services (AWS) on EC2. ? Created and maintained Technical documentation for launching Hadoop Clusters and for executing Hive queries and Pig Scripts ? Worked in tuning Hive & Pig to improve performance and solved performance issues in both scripts ? Worked with Elastic MapReduce (EMR) and setting up

environments on Amazon AWS EC2 instances. ? Developed various data connections from data source to SSIS, Tableau Server for report and dashboard development ? Involved unit testing, interface testing, system testing and user acceptance testing of the workflow tool. ? Used JIRA for bug tracking and GIT for version control. ? Involved in story-driven agile development methodology and actively participated in daily scrum meetings. Environment: Apache Hadoop 3.0, AWS, MLlib, MYSQL, Kafka, HDFS 1.2, Hive 2.3, Pig0.17, MapReduce, Flume 1.8, Cloudera, Oozie, UNIX, Oracle 12c, Tableau 7, GIT, UNIX. Developer Big Data January 2015 to June 2016 Confidential, NJ

Responsibilities: ? Involved in complete Big Data flow of the application data ingestion from upstream to HDFS, processing the data in HDFS and analyzing the data using several tools. ? Imported the data from various formats like JSON, Sequential, Text, CSV, AVRO and Parquet to HDFS cluster with compressed for optimization. ? Experienced on ingesting data from RDBMS sources like - Oracle, SQL Server and Teradata into HDFS using Sqoop. ? Configured Hive and written Hive UDF's and UDAF's Also, created partitions such as Static and Dynamic with bucketing. ? Implemented advanced procedures like text analytics and processing using the in-memory computing capabilities like spark. ? Importing and exporting data into HDFS and hive using Sqoop and Kafka with batch and streaming. ? Experienced with Spark-Streaming APIs to perform transformations and actions on the fly for building the common learner data model which gets the data from Kafka in near real time and Persists into HBase. ? Performance analysis of Spark streaming and batch jobs by using Spark tuning parameters. ? Enhanced and optimized product Spark code to aggregate, group and run data mining tasks using the Spark framework. ? Developed Spark scripts by using Python shell commands as per the requirement. ? Using Hive join queries to join multiple tables of a source system and load them to Elastic search tables. ? Experience in managing and reviewing huge Hadoop log files. ? Expertise in designing and creating various analytical reports and Automated Dashboards to help users to identify critical KPIs and facilitate strategic planning in the organization. ? Involved in Cluster maintenance, Cluster Monitoring and Troubleshooting ? Created Data Pipelines as per the business requirements and scheduled it using Oozie Coordinators. ? Maintaining technical documentation for each and every

step of development environment and launching Hadoop clusters. ? Worked on BI tools as Tableau to create dashboards like weekly, monthly, daily reports using tableau desktop and publish them to HDFS cluster. Environment: Scala, Hadoop, HDFS, Hive, Oozie, SqoopKafka, Elastic Search, Shell Scripting, HBase, Python, GitHub, Tableau, Oracle, MySQL, Teradata and AWS. Hadoop Developer Nationwide Mutual Insurance company - Columbus, OH June 2012 to November 2014 Responsibilities: ? Experience on AWS-EMR, Spark installation, HDFS and MapReduce Architecture. ? Participated in Hadoop Deployment and infrastructure scaling. ? Involved in creating Hive tables, and loading and analyzing data using hive queries. ? Developed Simple to complex Map Reduce Jobs using Hive and Pig. ? Developed workflow in Oozie to automate the tasks of loading the data into HDFS and pre-processing with Pig. ? Parsed high-level design spec to simple ETL coding and mapping standards. ? Maintained warehouse metadata, naming standards and warehouse standards for future application development. ? Worked with Linux systems and RDBMS database on a regular basis to ingest data using Sqoop. ? Implemented Kafka consumers to move data from Kafka partitions into Cassandra for near real-time analysis. ? Involved in Hadoop cluster task like adding and removing nodes. ? Managed and reviewed Hadoop log files and loaded log data into HDFS using Sqoop. ? Created and maintained Technical documentation for launching HADOOP Clusters and for executing Hive queries, Pig Scripts, Sqoop jobs. Environment: Hadoop(Hortonworks stack), HDFS, Oozie, Pig, Hive, MapReduce, Sqoop, Cassandra, Linux. Hadoop Developer US Bank - Denver, CO November 2010 to May 2012 Responsibilities: ? Worked on analyzing, writing HadoopMapReduce jobs using JavaAPI, Pig and Hive. ? Responsible for building scalable distributed data solutions using Hadoop. ? Involved in loading data from edge node to HDFS using shell scripting. ? Created HBase tables to store variable data formats of PII data coming from different portfolios. ? Exported the analyzed data to the relational databases using Sqoop for visualization and to generate reports for the BI team. ? Worked with using different kind of compression techniques to save data and optimize data transfer over network using LZO, Snappy, and Bzip etc. ? Analyze large and critical datasets using Cloudera, HDFS, HBase, MapReduce, Hive, HiveUDF, Pig, Sqoop, Zookeeper, &Spark. ?

Developed custom aggregate functions using Spark-SQL and performed interactive querying. ? Used Scoop to store the data into HBase and Hive. ? Worked on installing cluster, commissioning & decommissioning of DataNode, NameNode high availability, capacity planning, and slots configuration. ? Creating Hive tables, dynamic partitions, buckets for sampling, and working on them using HiveQL. ? Used Pig to parse the data and Store in Avro format. ? Stored the data in tabular formats using Hive tables and Hive Serdes. ? Collecting and aggregating large amounts of log data using Apache Flume and staging data in HDFS for further analysis. ? Worked with NoSQL databases like HBase for creating HBase tables to load large sets of semi structured data coming from various sources. ? Implemented a script to transmit information from Oracle to HBase using Sqoop. ? Implemented MapReduce programs to handle semi/unstructured data like XML, JSON, and sequence files for log files. ? Fine-tuned Pig queries for better performance. ? Involved in writing the shell scripts for exporting log files to Hadoop cluster through automated process. ? Installed Oozie workflow engine to run multiple Hive and pig jobs. ? Analyzed large amounts of data sets to determine optimal way to aggregate and report on it. Environment: Hadoop, Map Reduce, HDFS, Yarn, Sqoop, Oozie, Pig, Hive, HBase, Java, Eclipse, UNIX shell scripting, python, Horton works. Java Developer GE Healthcare - Richmond, TX January 2010 to October 2010 Responsibilities: ? Effectively interacted with team members and business users for requirements gathering. ? Involved in analysis, design, and implementation phases of the software development lifecycle (SDLC). ? Implementation of spring core J2EE patterns like MVC, Dependency Injection (DI), and Inversion of Control (IOC). ? Implemented REST Web Services with Jersey API to deal with customer requests. ? Developed test cases using J Unit and used Log4j as the logging framework. ? Worked with HQL and Criteria API from retrieving the data elements from database. ? Developed user interface using HTML, Spring Tags, JavaScript, JQuery, and CSS. ? Developed the application using Eclipse IDE and worked under Agile Environment. ? Design and implementation of front end web pages using CSS, JSP, HTML, java Script Ajax and, Struts ? Utilized Eclipse IDE as improvement environment to plan, create and convey Spring segments on Web Logic Environment: Java, J2EE, HTML, JavaScript, CSS, J Query, Spring 3.0, JNDI,

Hibernate 3.0, Java Mail, Web Services, REST, Oracle 10g, JUnit, Log4j, Eclipse, Web logic 10.3.

Java Developer BMO Harris Bank - Milwaukee, WI June 2008 to December 2009 Responsibilities: ? Responsible for Daily maintenance and improvement of live website using HTML5, CSS3 JavaScript, JQuery. ? Involved in Software Development Life Cycle phases like requirements gathering, Analysis, Design, Development and Testing. ? Participated in requirement gathering and worked closely with the architect in designing and modeling. ? Worked on development of SQL and stored procedures, trigger and function on MYSQL. ? Developed all the UI pages using HTML5, DHTML, XSL/XSLT, XHTML, DOM, CSS3, JSON, JavaScript, JQuery, Ajax, Adobe Creative suite. ? Updated billing pages using HTML, CSS in Angular.js framework. ? Performed form validations using reactive forms from Angular framework. ? Integrated the Java code (API) in JSP pages and responsible for setting up AngularJS framework for UI development. Developed html views with HTML5, CSS3, bootstrap and AngularJS. ? Implemented code according to coding standards and Created AngularJS Controller, which isolate scopes perform operations. ? Created the Application using AngularJS and NodeJS libraries and used NPM to manage dependencies and gulp to minify, rectify, and babelify the code. ? Involved in developing the web pages using Angular, which are powerful in building the Single page web applications? ? Used the Node.js, backbone.js and Require.js MVC Frameworks in the development of the web applications. ? Used advanced level of HTML5, JavaScript, CSS3 and pure CSS layouts (table less layout). ? Produced content pages with CSS3 layout and style mark-up presentations and also used JavaScript methods. ? Involved in Consuming Restful Services using Angular $http Service. Developed and consumed Restful services using WEB API. Used AJAX coding techniques to update parts of a web page. ? Developed the Client side validation using Java Script and JQuery and server side using Server side validations. ? Built efficient Angular.js backend for client web application. ? Also participated in server side programming with java using JDBC, Servlets and JSP. ? Designed Frontend with in object oriented JavaScript Framework like Backbone.JS, Angular.JS and Ext.JS. ? Used Angular.js to create server side applications. Used its workhorse connectors and libraries relating to HTTP, SSL compression, file system access, etc. ? Developed certain features of the application

functionality i.e. CRUD (Create, read, update, delete) features using Backbone.js, Require.js and Responsive Design. ? Designed and Developed Java Script frame work which is wrapper on top of JQUERY frame work and AJAX based UI frame work for UI Configuration widgets. ? Involved with Mobile development team to make mobile website responsive and fast, added AJAX functionality. ? Involved in developing XML, HTML, and JavaScript for client side presentation and, data validation on the client side with in the forms. ? Involved in redesigning the entire site with CSS styles for consistent look and feel across all browsers and all pages. ? Worked with the team of architects and back-end Developers to gather requirements and enhance the application functionality and add new features. ? Participated in the status meetings and status updating to the management team. Environment: SOAP, Restful web services, JavaScript, Angular JS, SQL, MySQL, JQuery, Apache web server, PHP, SQL Developer tool, JDBC, CSS3, HTML5. Education Bachelor's Skills APACHE HADOOP HDFS (7 years), APACHE HADOOP SQOOP (7 years), Hadoop (7 years), HADOOP (7 years), HADOOP DISTRIBUTED FILE SYSTEM (7 years) Additional Information TECHNICAL SKILLS: Big Data Ecosystem: Hadoop, Spark, Scala, Map Reduce, HDFS, Hive, Pig, Sqoop. Flume, Kafka, HBase Java Technologies: JSP, Servlets, Junit, Spring Hibernate Database Technologies: MySQL, SQL server , Oracle, MS Access Programming Languages: Scala, Python, Java and Linux shell scripting Operating Systems: Windows, LINUX

Name: Darius Fisher

Email: lambertmegan@example.net

Phone: (641)859-6638x25473