

Java-Hadoop Developer Java-Hadoop Developer Java-Hadoop Developer - T-Mobile Bellevue, WA
Certified Apache Spark Developer. Over 7+ years of experience in software architecture, design, development, testing and maintenance of web and enterprise applications with 3+ years of experience in using Hadoop and Hadoop Ecosystems (HDFS, YARN, MAPREDUCE, PIG, HIVE, HBASE, SQOOP, FLUME, ZOOKEEPER, OOZIE) and 4+ years of experience as Java developer. Excellent understanding of Hadoop architecture and various components such as HDFS, YARN, Resource Manager, Node Manager, Name Node, Data Node and Map Reduce programming paradigm. Hands on experience using Cloudera, Horton works and BigInsights Quick start Hadoop Distributions. Hands on experience on Scala and Python. Strong experience in developing Storm topology and ActiveMQ using Java Good understanding how Lambda architecture and Kappa architecture. Responsible for writing MapReduce programs Hadoop, pig, and Hive Experience in importing and exporting data using Sqoop from HDFS to Relational Database Systems and vice-versa. Experience loading data to Hive partitions and creating buckets in Hive Experience developing custom SerDe's in HIVE Experience in extending Hive and Pig core functionality by writing custom UDFs. Knowledge in extending Hive core functionality by writing custom UDFs like UDAFs and UDTFs. Expertise in analysis using PIG, HIVE and MapReduce. Logical Implementation and interaction with HBase Developed MapReduce jobs to automate transfer the data from HBase. Experience in developing ETL process using Map-Reduce Framework in java. Knowledge in efficiently using Spark's memory caching for iterative processing using Scala. Good understanding how Spark instructions are translated into jobs Strong Knowledge developing web service using Playframework and Akka Actors. Good Knowledge on Hadoop Cluster architecture and monitoring the cluster. Knowledge in job/workflow scheduling and monitoring tools like Oozie and Zookeeper. Strong understanding of NoSQL databases like HBase, MongoDB & Cassandra. Good working knowledge with Agile for SDLC. Knowledge in Amazon AWS concepts like EMR, S3 and EC2 web services. Efficient team player, having strong desire and capabilities to efficiently convert Customer's requirements into applications. Authorized to work in the US for any employer

Work Experience Java-Hadoop Developer T-Mobile - Bellevue, WA January 2019 to Present

Project: BEAM **Description:** The transformation program of our billing systems is the main objective of uprising project. It will provide new capabilities for flexibility, efficiency, and new business models. Uprising project covers various domains like Business Intelligence and Engineering etc. Integrated Data warehouse (IDW) is T-Mobile central repository of data in Hadoop and Teradata ecosystems and IDW is part of BI. As part of data ingestion team, we will bring incremental data from various source systems into Hadoop to maintain one central repository.

Responsibilities: Developed Storm topology to ingest data from various source into hadoop data lake Configured ActiveMQ for enterprise and resolved ActiveMQ issues Developed web application using HBase and Hive API to compare schema between HBase and Hive tables. Used JVM monitor to monitor threads and memory usage of HBase and Hive schema check web application. Developed code to generate Hive DDL's from source DDL's Created HBase tables using Sqoop from Relational Database Oracle Developed code to import JSON data into MYSQL database. Developed code to import data SQLServer into HDFS and created Hive views on data in HDFS using Spark Created scripts to append data from temporary HBase table to target HBase table in Spark Developed Hive queries for the analysis. Handling structured and unstructured data and applying ETL processes. Handled importing of data from various data sources, performance optimization of Sqoop. Developed Oozie workflow to compare data in Hive and Source. Assessed existing and EDW technologies and methods to ensure our EDW/BI architecture meet the needs of the business and enterprise and allows for business growth.

Assisted with data capacity planning and node forecasting. Collaborated with the infrastructure, network, database, application and BI teams to ensure data quality and availability. **Environment:**

Hadoop, Map Reduce, HDFS, Hive, HBase, Pig, Java, Hadoop distribution of Hortonworks, SQL,UNIX Shell Scripting, Oracle, SQLServer, MySQL Big Data Developer KCP&L - Kansas City, MO August 2018 to January 2019 **Project:** Meter Data Management **Description:** Collected, managed, and distributed information to the users of databases, applications, and platforms that are supported by the KCP&L team. Key areas were leveraging this capability to drive financial benefit to the company in Information Management, Marketing and Risk Capabilities. Used substantial

technical infrastructure to support the data management model and operational philosophy, which provided the flexibility and reliability, needed for in-class production, analytics, and on-line capability.

Responsibilities: Worked on Big Data Hadoop cluster implementation and data integration in developing large-scale system software. Developed MapReduce programs to parse the raw data, populate staging tables and store the refined data in partitioned tables in the EDW. Capturing data from existing databases that provide SQL interfaces using Sqoop. Worked extensively with Sqoop for importing and exporting the data from HDFS to Relational Database system Oracle and vice-versa. Loading data into HDFS. Develop and maintains complex outbound notification applications that run on custom architectures, using diverse technologies including Java, XML, JMS and Web Services. Created Hive queries that helped market analysts spot emerging trends by comparing fresh data with EDW reference tables and historical metrics. Enabled speedy reviews and first mover advantages by using Oozie to automate data loading into the Hadoop Distributed File System and PIG to pre-process the data. Provided design recommendations and thought leadership to sponsors/stakeholders that improved review processes and resolved technical problems. Shared responsibility for administration of Hadoop, Hive and Pig. Helped business processes by developing and configuring Hadoop ecosystem components that moved data from individual servers to HDFS. Created HBase tables to load large sets of structured, semi-structured and unstructured data coming from UNIX, NoSQL and a variety of portfolios. Developed multiple MapReduce jobs in Java for data cleaning and preprocessing. Developed workflow in Oozie to automate the tasks of loading the data into HDFS and pre-processing with Pig. Production Rollout Support and resolving any issues that are discovered by the client and client services teams.

Environment: Hadoop, Map Reduce, HDFS, Hive, Java (jdk1.6), Hadoop distribution of Hortonworks, Cloudera, PL/SQL, SQL, Toad 9.6, Windows NT, UNIX Shell Scripting. Software Engineer-Big Data Experian - Costa Mesa, CA December 2017 to July 2018

Project: Precise ID for Customer Management

Description: This project involves setting up a data repository and use it for search processing for analytical and research purposes. This application provides the capability for large batch processing using Hadoop map reduce jobs using Java runtime environment as well

as real time search capabilities using Solr cloud environment Responsibilities: Involved in designing and developing Hadoop MapReduce jobs Using Java Runtime Environment for the batch processing to search and match the scores. Used Rational Rose for developing Use case diagrams, Activity flow diagrams, Class diagrams and Object diagrams in the design phase. Used Struts with Tiles in the MVC framework for the application. Experiencing in running MapReduce programs on 20 node cluster (Amazon EC2 spot instances) with Apache Hadoop Developed java web services to upload data from local to Amazon S3, listing S3 objects and file manipulation operations. Successfully ran all Hadoop MapReduce programs on Amazon Elastic MapReduce framework by using Amazon S3 for Input and Output. Experience in Amazon S3 for storing objects and sharing data between Hadoop Involved in developing Hadoop MapReduce jobs for merging and appending the repository data. Used Sqoop to pull the data from Teradata Hands on experience in setting up HBase Column based storage repository for archiving and retro data. Used Java Message Service (JMS 1.1) for reliable and asynchronous communication. Involved in integration of Legacy Scoring and Analytical Models like SMG3 into the new application using Web Services. Responsible for writing Pig UDFs and Hive UDFs. Handled importing of data from various data sources, performance transformation using Hive. Experience in optimization of Map reduce algorithm using combiners and partitions to deliver the best results and worked on Application performance optimization for a HDFS cluster. Experience working with off-shore teams and communicating daily status on issues, road-blocks. Environment: Java, Hadoop, HBase, Zookeeper, Solr cloud, Pig Latin, Oozie scheduler, JavaBeans, Eclipse, Rational Clear case, Servlet 2.3, JUnit, Maven, XML Web services, JDBC, Unix, Windows NT/2000. Hadoop Developer Deutsche Bank - Cary, NC September 2015 to November 2017 Project: Marketing and Securities FDR Description: Designed and developed big data solutions involving Terabytes of data. The big data solution consists of collecting large amounts of log data from distributed sources, transformations and standardizations analysis, statistics, aggregations and reporting etc. Built an on demand elastic Hadoop cluster infrastructure to cater the needs of various Big Data projects, automated various Big Data workflows to process and extracts analytics out of the data using

MapReduce, Pig and Hive. Responsibilities: Installed, configured and maintained Apache Hadoop clusters for application development and Hadoop tools like Hive, Pig, HBase, Zookeeper and Sqoop for POC. Created HDFS (Hadoop Distributed File System), and MapReduce jobs in java. Implemented NameNode backup using NFS for High availability. Used Pig as ETL tool to do transformations, event joins and some pre-aggregations before storing the data onto HDFS. Responsible for developing data pipeline using flume, Sqoop and pig to extract the data from weblogs and store in HDFS. Used Sqoop to import and export data from HDFS to RDBMS and vice-versa. Created Hive tables and involved in data loading and writing Hive UDFs. Exported the analyzed data to the relational database MySQL using Sqoop for visualization and to generate reports. Used Hive to analyze the partitioned and bucketed data and compute various metrics for reporting. Automated workflows using shell scripts to pull data from various databases into Hadoop. Responsible for creating a Solr schema from the Indexer settings. Written Solr queries for various search documents. Environment: Hadoop, MapReduce, Hive, HDFS, PIG, Sqoop, Oozie, Solr, Cloudera, Flume, HBase, Zookeeper, Oracle, NoSQL and Unix/Linux. Java/J2ee Developer Cognolabs Systems Pvt Ltd - Hyderabad, Telangana October 2013 to August 2015 Description: This project was to develop an interface between the Claims Service Records (CSR) and a newly developed website for the Insurance Bureau. Purpose of creating the interface was to minimize the changes to the existing Claim Service Records. On request by a Claims Agent, data was requested for CSR, which was in a flat file format. This was converted to XML by the application and sent to Insurance Bureau website, where it was displayed. Responsibilities: Involved in the software development life cycle coding, testing, and implementation. Worked in the health-care domain. Involved in Using Java Message Service (JMS) for loosely coupled, reliable and asynchronous exchange of patient treatment information among J2EE components and legacy system. Developed MDBs using JMS to exchange messages between different applications using MQ Series. Involved in working with J2EE Design patterns (Singleton, Factory, DAO, and Business Delegate) and Model View Controller Architecture with JSF and Spring DI. Involved in Content Management using XML. Developed a standalone module transforming XML 837 module

to database using SAX parser. Installed, Configured and administered WebSphere ESB v6.x
Worked on Performance tuning of WebSphere ESB in different environments on different platforms.

Configured and Implemented web services specifications in collaboration with offshore team.

Involved in Creating dash board charts (business charts) using fusion charts. Involved in creating reports for the most of the business criteria. Environment: Java, Servlets, JSP, Vignette Tool, XML, Eclipse 3.3.0, Tomcat 5.x Java Developer Skdotcom Technologies - Hyderabad, Telangana
June 2011 to September 2013 Description:SpeedTrial is a web-based enterprise software solution for managing multisite clinical research studies. It helps the research team to manage clinical trials in an easier, faster and more efficient way. SpeedTrial is designed for pharmaceutical, biotech, medical companies, contract research organization (CROs), specialized research organization and academic medical centers (AMCs) to manage thousands of trials in a variety of therapeutic areas. SpeedTrial facilitates creation of study protocol, designing the Case Report Forms (CRFs) by dragging and dropping the sections and questions Electronic Data Capture (EDC), and exhaustive clinical data management and reduces administration and monitoring costs. Responsibilities:
Designed and Developed application using EJB 2.0 and Struts framework. Developed POJO's for Data Model to map the Java Objects with Relational database tables. Designed and developed Service layer using Struts framework. Used MVC based Struts framework to develop the multi-tier web application presentation layer components. Involved in Integration of Struts with Database. Implemented Struts tag libraries like html, logic, tab, bean etc in the JSP pages. Used Struts tiles libraries for layout of web page, and performed struts validations using Struts validation framework.

Implemented Oracle database and JDBC drivers to access the data. Involved in design, analysis and architectural meetings. Created Architecture Diagrams, and Flow Charts using Rational Rose. Followed agile software development practice paired programming, test driven development and scrum status meetings. Developed use case diagrams, class diagrams, database tables, and mapping between relational database tables. Developed Unit test cases using JUnit. Maintained the application configuration information in various properties file. Performed unit testing, system testing and integration testing. Environment: Java, Struts Framework, log4j, Servlets, JSP, JSTL,

I18N, JDBC, HTML, Java Script, CSS, UML, Oracle, Windows NT. Education Bachelor of Technology in Electronics and Communication Engineering in Electronics and Communication Engineering Jawaharlal Nehru Technological University Skills ECLIPSE, J2EE, JAVA, SPRING, JMS, JSP, JVM, HDFS, OOZIE, SQOOP, HBASE, DB2, FLUME, HADOOP, MAP REDUCE, AMAZON WEB SERVICES, Git, GRADLE, Hadoop, HBase

Name: Tracy Thomas

Email: cevans@example.org

Phone: +1-430-653-4868x1488