Senior Bigdata Hadoop Developer Senior Bigdata Hadoop Developer Senior Bigdata Hadoop Developer - Kohls Work Experience Senior Bigdata Hadoop Developer Kohls - Milpitas, CA February 2018 to Present Kohl's is an American department store retailing chain. Kohl's started Marketing Transformation Program (MTP) to create Data Lake for all the data in Google Cloud platform. Ingesting the data from various sources files, real time data, RDBMS and processing the data using Big data technologies and infrastructure. Roles and Responsibilities: Creating and executing Azkaban flows using Java, Hive, Pig, shell scripts to ingest the data into GCP (Google Cloud Platform) from various sources such as server logs, Mount point server files and Real time data from Mobile applications to utilize Google Cloud Platform as Data Lake. Designing Row keys & Schema with NoSQL databases such as HBase to load huge volumes of data for real time transactions. Writing multiple MapReduce Jobs using Java API, Pig and Hive for data extraction, transformation and aggregation from multiple file formats including Parquet, Avro, XML, JSON, CSV, ORCFILE and other compressed file format Codecs like gzip, Snappy, LZO. Building advanced analytical applications by making use of Spark SQL, Big Query to create detail level summary reports and Dashboard using KPI's. Working on Spark-Streaming APIs using Data frames to perform necessary 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 Google Cloud Platform (GCP) and Big Query. Building and deploying the applications using Jenkins and Tonomi with appropriate configurations. Optimizing Spark jobs by tuning executors, memory configurations, D-Streams, Accumulator, Broadcast variables, RDD caching to deliver the best results for the large datasets. Responsible for deploying the scripts into Github version control repository hosting service and deployed the code using Jenkins. Creating and executing apache Airflow jobs in Python language for orchestration purposes Involving in design, code reviews and supporting test teams in testing to fix the identified issues in the developed applications. Involved in business and functional requirements gathering and prepared user stories in JIRA and documentation in Confluence. Environment: Hortonworks Data Platform (HDP), HDFS, MapReduce, YARN, HBase, Java, Hive, Pig, Sqoop, Spark, Parquet, Flume, GCP (Google Cloud Platform), Azkaban, Big query,

Python, Scala, Airflow, Kafka Senior Hadoop Developer AT&T - El Segundo, CA February 2015 to January 2018 DirecTV (Now part of AT&T) provides television and audio services to subscribers through satellite transmissions. Services include the equivalent of many local television stations, broadcast television networks, subscription television services, satellite radio services, and private video services. AMS is a system which enables in DIRECTV set to collect information about events that occur on the Set Top Box (STB) for later reporting. Event information includes live viewing, DVR usage (playback and record), interactive application usage, CA OSD events, advertising data, and diagnostic information.   Roles and Responsibilities:   Designing the applications from the ingestion to reports delivery to third party vendors using big data technologies flume, kafka, sqoop, map-reduce, hive, pig   Process the raw log files from the set top boxes using java map reduce code and shell scripts and stored them as text files in HDFS.   Ingesting the data from legacy and upstream systems to HDFS using apache Sqoop, Flume java map reduce programs, hive queries and pig scripts.   Generating the required reports using Oozie workflow and Hive queries for operations team from the ingested data.   Alert and monitoring mechanism for the oozie jobs for the failure conditions and successful conditions using email notifications   Involved in bluecoat proxy approach while sharing the data from the in-house Hadoop cluster to external vendors   Working with application teams to install operating system, Hadoop updates, patches, version upgrades as required.   Writing Map Reduce code to make un-structured and semi- structured data into structured data and loading into Hive tables.   Worked on debugging, performance tuning of Hive & Pig Jobs   Analyzing system failures, identifying root causes, and recommended course of actions as part of operations support.   Involved in Spark streaming solution for the time sensitive revenue generating reports to match the pace with upstream (STB) systems data   Experience in working on Hbase with Apache phoenix as a data layer to serve the web requests to meet the SLA requirements   Created HBASE tables and load the data using spark-streaming application   Worked on SFDC ODATA connector to get the data from NodeJS services which in turns the fetch the data from HBASE   Utilized AWS S3 services to push/store and pull the data from AWS from external applications   Responsible for functional requirements gathering, code reviews, deployment

scripts and procedures, offshore coordination and on-time deliverables     Environment: Hadoop, HDFS, Pig, Hive, Flume, Kafka, MapReduce, Sqoop, Spark, Oozie, LINUX, NodeJS, SFDC, ODATA and AWS. Hadoop Developer Dun & Bradstreet - Parsippany, NJ August 2014 to January 2015 Dun & Bradstreet (D & B) provides commercial data to businesses on credit history, business-to-business sales and marketing, counterparty risk exposure, supply chain management, lead scoring and social identity matching. It maintains a database contains information on more than 235 million companies across 200 countries worldwide. This database contains over 53 million professional contact names using a variety of sources including public records, trade references, telco providers, telephone interviews, print, digital and trade publications.     Roles and Responsibilities:     Involved in designing and developing the Cascading work flows     Actively participated in validating and cleansing phases of the data flow     Worked with the Data Science team to gather requirements for various data transformations     Involved in writing a new subassembly for a complex query used for lookup the data     Worked on custom functions and subassemblies which helps for code reuse     Involved in running cascading job in local and Hadoop modes     Worked with application teams to install VM, HBASE, Hadoop updates, patches, version upgrades as required   Created and tested JUNIT test cases for different Cascading flows   Worked on Tableau workbooks to perform year over year, quarter over quarter, YTD, QTD and MTD type of analysis.     Utilized Tableau server to publish and share the reports with the business users.     Idea on continuous integration tools like Jenkins to automate the process for code builds and deployments     Environment: Hadoop, HDFS, Cascading, Lingual, HBASE, REDIS, Zookeeper, JUnit, Jenkins, Gradle, Tableau Java/Hadoop Developer Tata Consultancy Services - Hyderabad, Telangana December 2011 to July 2013 India   Client: Verizon, NY     Verizon is one of the world's leading providers of high-growth communications services. Verizon is the nation's largest local exchange carrier and a rapidly emerging long distance provider. Verizon companies are the largest providers of wired line and wireless communications in the United States, serving the equivalent of nearly 135 million access lines and 30 million wireless customers.     Responsibilities:     Created Temporary Tables to store the data from Legacy system.     Used SQL*Loader scripts to load the

data into temporary tables and procedures to validate the data.    Creation of Materialized views and function based indexes    Setting up various notification scripts which will notify if the database is not up and running or if any directory space is filled up.    Used Visual SourceSafe for version control and performing various builds.    Involved in Unit testing, User Acceptance Testing to check whether the data is loading into target    Extracted files from DB2 through Sqoop and placed in HDFS and processed    Extracted from different source systems according to the user requirements.    Involved in POC for analyzing large data sets by running Hive queries and Pig scripts    Worked with the Data Science team to gather requirements for various data mining projects    Worked on Hive tables creation, and loading and analyzing data using hive queries    Analyzed large data sets by running Hive queries and Pig scripts    Worked with the Data Science team to gather requirements for various data mining projects    Created and maintained Technical documentation for launching HADOOP Clusters and for executing Hive queries and Pig Scripts    Providing daily development status, weekly status reports, and weekly development summary and defects report.    Environment: Core Java, JDBC, JavaScript, MySQL, JUnit, Eclipse, QA, Hadoop, Sqoop, Hive, Pig Java/J2EE Developer Zensar Technologies - Hyderabad, Telangana April 2010 to November 2011 India  Client: Assurant Solutions, Miami, FL    Assurant Solutions is one of the leading Auto Insurance companies in the US. Assurant has the various applications within that XYCOR, Non-Bill and Special Products application comes under Insurance and financial application sector. These application mainly deals with Credit insurance. Credit Insurance is an insurance policy and risk management product that covers the payment risk in the event a borrower is unable to repay because of certain events.

Roles and Responsibilities:    Utilized Agile Methodologies to manage full life-cycle development of the project.    Developed front end validations using JavaScript and developed design and layouts of JSPs and custom taglibs for all JSPs.    Used JDBC for database connectivity.    Developed web application using JSP custom tag libraries and Action. Designed Java Servlets and Objects using J2EE standards.    Used JSP for presentation layer, developed high performance object/relational persistence and query service for entire application.    Developed the XML Schema and Web services for the data maintenance and structures.    Used WebLogic Application Server and RAD to

develop and deploy the application.    Worked with various Style Sheets like Cascading Style Sheets (CSS).      Designed database and created tables, written the complex SQL Queries and stored procedures as per the requirements.    Involved in coding for JUnit Test cases, ANT for building the application.      Environment: Java/J2EE, Oracle 10g, SQL, PL/SQL, JSP, EJB , WebLogic 8.0, HTML, AJAX, Java Script, JDBC, XML, JMS, JUnit, log4j, MyEclipse 6.0. Java/J2EE Developer Satyam Computer Services Ltd - Hyderabad, Telangana August 2009 to April 2010 India   Client: Central Bancompany, Jefferson City, MO    Worked on application named Access portal. It is a part of online banking that allows a customer to view quick summary of transactions and account details. It also shows mutual funds associated with account.    Roles and Responsibilities    Responsible for understanding the scope of the project and requirement gathering.    Review and analyze the design and implementation of software components/applications and outline the development process strategies    Coordinate with Project managers, Development and QA teams during the course of the project.      Used Spring JDBC to write some DAO classes to interact with the database to access account information.    Used Tomcat web server for development purpose.    Involved in creation of Test Cases for JUnit Testing.    Used Oracle as Database and used Toad for queries execution and also involved in writing SQL scripts, PL/SQL code for procedures and functions.      Used CVS, Perforce as configuration management tool for code versioning and release.      Developed application using Eclipse and used build and deploy tool as Maven.      Used Log4J to print the logging, debugging, warning, info on the server console.    Environment: Java1.5, J2EE [Servlet, JSP], XML, Spring 3.0, Design Patterns, Log4j, CVS, Maven, Eclipse, Apache Tomcat 6, and Oracle 11g. Java Developer Satyam Computer Services Ltd - Hyderabad, Telangana September 2008 to July 2009 India   Client: OneBeacon Insurance Group, Minnetonka, MN      Disability Income Application is used for the Underwriting and Administration of Disabilities products. This is a Maintenance and Enhancement Project for Individual and Institutional policies of Metlife. Responsibilities:    Coded the business methods according to the IBM Rational Rose UML model. Extensively used Core Java, Servlets, JSP and XML.    Used DB2 Database to store the system data   Used Rational Application Developer (RAD) as Integrated Development Environment (IDE).

Used unit testing for all the components using JUnit    Used Apache log 4j Logging framework for logging of trace and Auditing.    Used Asynchronous JavaScript and XML (AJAX) for better and faster interactive Front-End.    Provide support to resolve performance testing issues, profiling and cache mechanism.    Performs code reviews to ensure consistency to style standards and code quality    Environment: Java 1.6, Servlets, JSP, IBM Rational Application Developer (RAD) 6, Websphere 6.0, iText, AJAX, DB2, log4j. Java/J2EE Developer Satyam Computer Services Ltd - Hyderabad, Telangana October 2007 to August 2008 India   Client: Lowes, Mooresville, North Carolina.    Lowe's Companies Inc is a US-based chain of retail home improvement and appliance stores. Founded in 1946 in North Wilkesboro, North Carolina, the chain now serves over 14+ million customers a week in its 1,616 stores in the United States and Canada.    The import project will provide and support overseas product and relationship growth as well as supported IT systems and processes for both sourcing (item and vendor functions) and purchase order management (creating and managing POs). Mainframes act as a host between PeopleSoft tables and Tradestone. The objective of the project is to design, develop and integrate Lowe's Purchase Order System with TradeStone, a Lowe's web based application for PO maintenance.    Responsibilities:    Designed User Interface using Java Server Pages (JSP) and XML.    Developed the Enterprise Java Beans (Stateless, Stateful Session beans, Entity beans) to handle different transactions such as online funds transfer, bill payments to the service providers.    Implemented Service Oriented Architecture (SOA) using JMS in MDB for sending and receiving messages while creating web services. Worked on Web Services for data transfer from client to server and vice versa using SOAP, WSDL, and UDDI.    Involved in testing the web services using SOAP UI    Extensively worked on JMS using point-point, publisher/subscriber messaging Domains for implementing Exchange of information through Messages.    Environment: Windows, Java 1.4, HTML, JavaScript 1.6, XML, JUnit, JMS, Web Services, SOAP 1.1, UDDI 2, Maven 2.0, Eclipse IDE, CVS, Oracle 10g. Skills Eclipse, J2ee, Java, Intellij idea, Jsp Additional Information    Around 11 years of experience in IT industry which includes around 6 years of experience in Big Data in implementing complete Hadoop solutions, Architecture and Design.    Hands on experience in installing, configuring and using

Apache Hadoop ecosystem components like HDFS, Hadoop MapReduce, Zoo Keeper, Oozie, Hive, Sqoop, Kafka, Spark, Pig and Cascading, Azkaban, Airflow, Big query    Expertise in writing Hadoop Jobs for analyzing data using Hive and Pig    Experience in working with MapReduce programs using Apache Hadoop for working with Big Data    Experience in importing and exporting data using Sqoop from HDFS to Relational Database Systems(RDBMS) and vice-versa    Experience in working with ETL tool Kettle by Pentaho.    Involved in developing Tableau Dashboards with interactive views, trends and drill downs for the users data    In depth understanding/knowledge of Hadoop Architecture and various components such as HDFS, Job Tracker, Task Tracker, Name Node, Data Node and MapReduce concepts    Worked on Cascading API for Hadoop application development and work flows    Good understanding of Data Mining and Machine Learning techniques    Experience in analyzing data using Hive QL, Pig Latin, and custom MapReduce programs in Java.    Knowledge of job workflow scheduling and monitoring tools like Oozie and Zookeeper    Worked on Spark streaming application which consumes the data from Kafka and persist it in to HBASE    Experience in optimization of Map reduce algorithm using combiners and partitioners to deliver the best results    Involved in Spark solution for the time sensitive use cases    Good understanding of NoSQL databases like MongoDB, REDIS    Expertise in core Java, J2EE, Multithreading, JDBC, Shell Scripting and proficient in using Java API's for application development    Proficient in Working with Various IDE tools including Eclipse Galileo, IBM Rational Application Developer (RAD) and IntelliJ IDEA    Worked on different operating systems like UNIX/Linux,Windows XP and Windows 2K    Good knowledge on functional programming language Scala    Integrate Salesforce using ODATA connector for Rest API.    Worked on NODEJS API to pull the data from HBASE    Written Airflow jobs using in-build and custom operators in Python to schedule and run the flows    Very good experience in customer specification study, requirements gathering, system architectural design and turning the requirements into final product.    Strong background in mathematics and have very good analytical and problem solving skills.    Technical Expertise:    Hadoop/Big Data: HDFS, MapReduce, Pig, Hive, Sqoop, Oozie, Zookeeper, Cascading, SPSS, Kafka, Flume, Hbase, Phoenix, Spark, Azkaban, YARN, Airflow  Java & J2EE

technologies: Core Java, JSP, JDBC  Hadoop distributions: Cloudera, Hortonworks, Google Cloud Platform  IDE Tools: Eclipse, IntelliJ IDEA  Programming languages: C, C++, Java, Linux shell scripts, VB.NET, COBOL, Python, Scala  Databases: Oracle 11g/10g/9i, MySQL, DB2, MS-SQL Server, MongoDB, Big Query  Web Technologies: HTML, XML, JavaScript  ETL Tools: Kettle, Reporting Tools: Tableau  Operating Systems: Windows 95/98/2000/XP/Vista/7, LINUX  Monitoring & Reporting: Nagios, Custom shell scripts  Version control: Git, SVN  Testing Tools: JUnit, MRUnit

Name: Julie Deleon

Email: kristen91@example.com

Phone: (964)534-1915