Sr. Hadoop Developer Sr. Hadoop Developer Sr. Hadoop Developer - Directv Miami, FL   Overall 8+ years of professional IT experience in Software Development. This also includes Extensive years of experience in Ingestion, Storage, Querying, Processing and Analysis of Big Data using Hadooptechnologies and solutions.   Excellent understanding / knowledge of Hadoop architecture and various components of Hadoopecosystem such as HDFS, Job Tracker, Task Tracker, Name Node, Data Node, Map Reduce&YARN.   Good understanding/knowledge of Hadoop Architecture and various components such as HDFS, JobTracker, TaskTracker, NameNode, DataNode, Secondry Namenode, and MapReduce concepts.   Experienced managing No-SQL DB on large Hadoop distribution Systems such as: Cloudera, Hortonworks HDP, MapR M series etc.   Experienced developing Hadoop integration for data ingestion, data mapping and data process capabilities.   Experienced in building analytics for structured and unstructured data and managing large data ingestion using technologies like Kafka/Avro/Thift.   Software development in Java Application Development, Client/Server Applications, Internet/Intranet based database applications and developing, testing and implementing application environment using C++, J2EE, JDBC, JSP, Servlets, Web Services, Oracle, PL/SQL and Relational Databases.   Exceptional ability to quickly master new concepts and capable of working in groups as well as independently.   Excellent interpersonal skills and the ability to work as a part of a team.   Experience in debugging, troubleshooting production systems, profiling and identifying performance bottlenecks.   Has good knowledge of virtualization and worked on VMware Virtual Center.   Hands on experience in installing, configuring, and using Hadoop ecosystem components like Hadoop MapReduce, HDFS, HBase, Hive, Sqoop, Pig, Zookeeper, Storm, Spark, Kafka and Flume.   Good Knowledge on Hadoop Cluster architecture and monitoring the cluster.   In-depth understanding of Data Structure and Algorithms.   Experience in managing and troubleshooting Hadoop related issues.   Expertise in setting up standards and processes for Hadoop based application design and implementation.   Experience in importing and exporting data using Sqoop from Relational Database Systems to HDFS and vice-versa.   Experience in managing Hadoop clusters using Cloudera Manager.   Experience in using the Impala usage for the high performance SQL queries.   Very good

experience in complete project life cycle (design, development, testing and implementation) of Client Server and Web applications.   Hands on experience in VPN, Putty, winSCP, VNCviewer, etc. Hands on experience in application development using Java, RDBMS, and Linux shell scripting. Performed data analysis using MySQL, SQL Server Management Studio and Oracle.   Expertise in creating Conceptual Data Models, Process/Data Flow Diagram, Use Case Diagrams and State Diagrams.   Experience with cloud computing platforms like Amazon Web Services(AWS).   Ability to adapt to evolving technology, strong sense of responsibility and accomplishment. Authorized to work in the US for any employer Work Experience Sr. Hadoop Developer Directv - Los Angeles, CA May 2017 to Present Responsibilities: Data Ingestion implemented using SQOOP, SPARK, loading data from various RDBMS, CSV, XML files.   Data cleansing, transformations tasks are handled using SPARK using SCALA and HIVE.   Data Consolidation was implemented using SPARK, HIVE to generate data in the required formats by applying various ETL tasks for data repair, massaging data to identify source for audit purpose, data filtering and store back to HDFS. Responsible for design development of Spark SQL Scripts based on Functional Specifications. Exploring with the Spark improving the Performance and Optimization of the existing algorithms in Hadoop.   ETL development to normalize this data and publish it in IMPALA.   Involved in converting Hive/SQL queries into Spark RDD using Scala.   Responsible for Job management using Fair scheduler and Developed Job Processing scripts using Oozie Workflow.   Responsible for Performance Tuning of Spark Applications for setting right Batch Interval time, correct level of Parallelism and Memory tuning.   Optimizing of existing algorithms in Hadoop using Spark Context, Spark-SQL and Pair RDD's.   Responsible in handling large datasets using Partitions, Spark in Memory capabilities, Broadcasts in Spark, Effective & efficient Joins, Transformations and other during Ingestion process itself.   Importing and exporting data into HDFS and HIVE, PIG using Sqoop.   Involved in creating Hive Tables, loading with data and writing Hive queries which will invoke and run Map Reduce jobs in the backend.   Implemented the workflows using Apache Oozie framework to automate tasks.   Worked with No SQL databases like HBase. Creating HBase tables to load large sets of semi structured data coming from various sources.   Worked with different file

formats such as Text, Sequence files, Avro, ORC and Parquet.   Responsible to manage data coming from different sources.   Responsible on loading and transforming of large sets of structured, semi structured and unstructured data.   Analyzed large amounts of data sets to determine optimal way to aggregate and report on it.   Environment: Scala, Hive, HBase, Flume, Java, Impala, Pig, Spark, Oozie, Oracle, Yarn, Junit, Unix, Cloudera, Flume, Sqoop, HDFS, Java, Python. Hadoop Developer Market St - San Francisco, CA July 2015 to April 2017 Responsibilities:

 Involved in file movements between HDFS and AWSS3 and extensively worked with S3 bucket inAWS.   Developing use cases for processing real time streaming data using tools like Spark Streaming.   Handled large datasets using Partitions, Spark in Memory capabilities, Broadcasts in Spark, Effective & efficient Joins, Transformations.   Imported required tables from RDBMS to HDFS using Sqoop and used Spark and Kafka to get real time streaming of data into HBase. Enhanced and optimized product Spark code to aggregate, group and run data mining tasks using the Spark framework and handled Json Data.   Developed Spark code using Scala and Spark-SQL for faster testing and data processing.   Responsible for batch processing of data sources using Apache Spark.   Developed predictive analytic using Apache Spark Scala APIs.   Developed MapReduce jobs in Java API to parse the raw data and store the refined data.   Developed Kafka producer and consumers, Hbase clients, Spark and Hadoop MapReduce jobs along with components on HDFS, Hive.   Involved in identifying job dependencies to design workflow for Oozie & YARN resource management.   Worked on a product team using Agile Scrum methodology to Design, Develop, Deploy and support solutions that leverage the Client big data platform. Integrated Apache Storm with Kafka to perform web analytics. Uploaded click stream data from Kafka to Hdfs, Hbase and Hive by integrating with Storm.   Design and code from specifications, Analyzes, Evaluates, Tests, Debugs, Documents, and Implements Complex Software Apps. Worked in tuning Hive & Pig to improve performance and solved performance issues in both scripts with understanding of Joins, Group and aggregation and how does it translate to Map Reduce jobs Created Partitions, Buckets based on State to further process using Bucket based Hive joins. Implemented Cloudera Manager on existing cluster.   Extensively worked with Cloudera Distribution

Hadoop, CDH 5.x, CDH4.x Responsible for troubleshooting debugging and fixing the wrong data or data missing problem for Oracle Database (Mysql). Environment: HDFS, MapReduce, JavaAPI, JSP, JavaBean, Pig, Hive, Sqoop, Flume, Oozie, HBase, Kafka,Impala, Spark Streaming, Storm, Yarn, Eclipse, Unix Shell Scripting, Cloudera. Hadoop Developer AT&T - Atlanta, GA April 2013 to June 2015 Responsibilities: Involved in review of functional and non-functional requirements. Facilitated knowledge transfer sessions. Installed and configured Hadoop Mapreduce, HDFS, Developed multiple MapReduce jobs in java for data cleaning and pre-processing. Importing and exporting data into HDFS and Hive using Sqoop. Experienced in defining job flows. Experienced in managing and reviewing Hadoop log files. Extracted files from CouchDB through Sqoop and placed in HDFS and processed. Experienced in running Hadoop streaming jobs to process terabytes of xml format data. Load and transform large sets of structured, semi structured and unstructured data. Responsible to manage data coming from different sources. Exploring with Spark improving the performance and optimization of the existing algorithms in Hadoop using Spark context,Spark-SQL, Data Frame, pair RDD's, Spark YARN. Used Coalesce and repartition on data frames while optimizing the Spark jobs. Developed Spark scripts by using Scala shell commands as per the requirement. Experience in using Spark API over Cloudera Hadoop YARN to perform analytics on data in Hive. Developed Scala scripts, UDFFs using both Data frames/SQL and RDD/MapReduce in Spark 1.6 for Data Aggregation, queries and writing data back into OLTP system through Sqoop and Developed enterprise application using scala as well Involved in creating Hive tables, loading with data and writing hive queries which will run internally in map reduce way. Performed advanced procedures like text analytics and processing, using the in-memory computing capabilities of Spark using Scala. Gained very good business knowledge on health insurance, claim processing, fraud suspect identification, appeals process etc. Developed a custom File System plug in for Hadoop so it can access files on Data Platform. This plugin allows Hadoop MapReduce programs, HBase, Pig and Hive to work unmodified and access files directly. Designed and implemented Mapreduce-based large-scale parallel relation-learning system Extracted feeds form social media sites such as Facebook, Twitter using Python scripts. Setup

and benchmarked Hadoop/HBase clusters for internal use   Environment: Hadoop, MapReduce, HDFS, Hive, Sqoop, HBase, UNIX Shell Scripting,Scala. Java Developer Paycom - Oklahoma City, OK February 2011 to March 2013 Responsibilities: Developed Map Reduce programs in Java for parsing the raw data and populating staging   Worked on both WebLogic Portal 9.2 for Portal development and WebLogic 8.1 for Data Services Programming   Used Eclipse 6.0 as IDE for application development.   Involved in writing test cases by using set of conditions to test the application   Configured Struts framework to implement MVC design patterns   Build sql queries for fetching the required columns and data from database.   Used Subversion as the version control system   Managed the SVN related responsibilities and maintained the versions accordingly. Done SVN check in and check out's.   Used Hibernate for handling database transactions and persisting objects   Used AJAX for interactive user operations and client side validations Developed ANT script for compiling and deployment   Performed unit testing using Junit Extensively used Log4j for logging the log files Environment: Java/J2EE, SQL, PL/SQL, JSP, EJB, Struts, SVN, JDBC, XML, XSLT, UML, JUnit System Engineer Morgan Stanley - New York, NY July 2009 to January 2011 Responsibilities: Involved in Requirement Analysis, Development and Documentation.   Participation in developing form-beans and action mappings required for struts implementation and validation framework using struts.   Development of front-end screens with JSP Using Eclipse.   Involved in Development of Medical Records module.   XML and XSDs are used to define data formats.   Involved in Bug fixing and functionality enhancements.   Designed and developed excellent Logging Mechanism for each order process using Log4J.   Involved in writing Oracle SQL Queries.   Involved in Check-in and Checkout process using CVS.   Developed additional functionality in the software as per business requirements.   Involved in requirement analysis and complete development of client side code.   Followed Sun standard coding and documentation standards.   Participation in project planning with business analysts and team members to analyze the Business requirements and translated business requirements into working software.   Developed software application modules using disciplined software development process. Environment: Java, J2EE, JSP, EJB, ANT, STRUTS1.2, Log4J, Web logic 7.0, JDBC,

MyEclipse, Windows, XP, CVS, Oracle. Education Bachelor of Science in Information Systems Management University of Maryland Baltimore County - Catonsville, MD Skills Hadoop (5 years), XML (7 years), SQL (9 years), ECLIPSE (5 years), JAVA (9 years) Additional Information Technical Skills: Hadoop ECO Systems HDFS, Map Reducing, HDFS, Oozie, Hive, Pig, Sqoop, Flume, Zookeeper and HBase, Cassandra NO SQL HBase, Cassandra, MongoDB Data Bases MS SQL Server 2000/2005/2008/2012, MY SQL, Oracle 9i/10g Languages Languages Java JDK1.4 1.5 1.6 (JDK 5 JDK 6), C/C++, SQL, PL/SQL.   Operating Systems Windows Server 2000/2003/2008, Windows XP/Vista, Mac OS, UNIX ,LINUX Java Technologies Servlets, JavaBeans, JDBC, JNDI Frame Works JUnit and JTest IDE's & Utilities Eclipse, Maven, NetBeans. SQL Server Tools SQL Server Management Studio, Enterprise Manager, QueryAnalyser, Profiler, Export & Import (DTS). WebDev. Technologies ASP.NET, HTML,XML

Name: Heather White

Email: kirbytimothy@example.org

Phone: 001-900-370-0014