

Hadoop Developer Hadoop Developer Hadoop Developer - TIAA-CREF Around 7 years of experience in full Software Development Life Cycle (SDLC), AGILE Methodology and analysis, design, development, testing, implementation and maintenance in Hadoop, Data Warehousing, Linux and Java. 3 years of experience in providing solutions for Big Data using Hadoop 2.x, HDFS, MR2, YARN, Kafka, PIG, Hive, Sqoop, HBase, Cloudera Manager, Zoo keeper, Oozie, Hue, CDH5 & HDP 2.x. Experienced in Big data, Hadoop, NoSQL and various components such as HDFS, Job Tracker, Task Tracker, Name Node, Data Node and Map Reduce2, YARN programming paradigm. Working experience on Cloudera, Horton Works Hadoop distribution. Implementation of Big data batch processes using Hadoop, Map Reduce2, YARN, PIG and Hive. Experienced in using Kafka as a distributed publisher-subscriber messaging system. Experience in importing and exporting data using Sqoop from HDFS/Hive/HBase to Relational Database Systems and vice-versa. Hands on experience in in-memory data processing with Apache Spark. Experience in integration of various data sources like Oracle, DB2, Sybase, SQL server and MS access and non-relational sources like flat files into staging area. Good experience in writing PIG scripts and Hive Queries for processing and analyzing large volumes of data. Experience in optimization of Map Reduce algorithm using Combiners and Partitioners to deliver best results. Experienced in designing, developing and implementing connectivity products that allow efficient exchange of data between the core database engine and the Hadoop ecosystem Extending Hive and Pig core functionality by writing custom UDFs. Hands on experience in using BI tools like Splunk/Hunk, Tableau. Experience in understanding the security requirements for Hadoop and integrate with Kerberos authentication and authorization infrastructure. Experience in managing and reviewing Hadoop log files. Hands on experience in application development using Core JAVA, RDBMS and Linux shell scripting. Having good working experience in Agile/Scrum methodologies, technical discussion with client and Communication using scrum calls daily for project analysis specs and development aspects. Ability to work independently as well as in a team and able to effectively communicate with customers, peers and management at all levels in and outside the organization.

Work Experience Hadoop Developer TIAA-CREF - Charlotte, NC December 2014 to Present The

Project deals with developing a centralized data repository using Hadoop ecosystem for maintaining data flowing in from client locations and provide data querying capabilities to the Pathologists.

Roles & Responsibilities: Worked on analyzing Hadoop cluster using different big data analytic tools including Kafka, Pig, Hive and Map Reduce. Developing parser and loader map reduce application to retrieve data from HDFS and store to HBase and Hive. Importing the data from the MySql and Oracle into the HDFS using Sqoop. Importing the unstructured data into the HDFS using Flume. Written Map Reduce java programs to analyze the log data for large-scale data sets.

Involved in creating Hive tables, loading and analyzing data using hive queries. Involved in using HBase Java API on Java application. Deployed Hadoop Cluster in Pseudo-distributed and Fully Distributed modes. Involved in running Ad-Hoc query through PIG Latin language, Hive or Java Map Reduce. Real time streaming the data using Spark with Kafka. Configured Spark streaming to receive real time data from the Kafka and store the stream data to HDFS using Scale. Wrote the shell scripts to monitor the health check of Hadoop daemon services and respond accordingly to any warning or failure conditions. Importing and exporting data into HDFS using Sqoop and Kafka.

Developed PIG Latin scripts to extract the data from the web server output files to load into HDFS. Automated all the jobs for extracting the data from different Data Sources like MySQL to pushing the result set data to Hadoop Distributed File System. Environment: Hadoop 1.0.0, Oracle 11g/10g, Python, Map Reduce, Hive, HBase, Flume, Sqoop, Pig, Zookeeper, Java, ETL, SQL Server, CentOS, UNIX, Linux, Windows 7/ Vista/ XP. Hadoop Developer Shell Oil Company - Houston, TX January 2013 to November 2014 Shell Oil Company is a multinational oil company, which is amongst the largest oil companies in the world. It is one of America's largest oil and natural gas producers, natural gas marketers, gasoline marketers and petrochemical manufactures. Big data analytics transform big data in shell Exploration and production into sound exploration decisions, high quality wells, reduced costs and lower environment impact. The ability to collect more data and analyze it helped the company to locate oil and gas reserves in areas that were once inaccessible or thought to be depleted.

Roles& Responsibilities: Involved in review of functional and nonfunctional requirements. Develop Map Reduce jobs for the users. Maintain, update and

schedule the periodic jobs which range from updates on periodic Map Reduce jobs to creating ad-hoc jobs for the business users. Importing and exporting data into HDFS and Hive using Sqoop. Experienced in defining job flows. Scheduling all Hadoop/Hive/ Sqoop/ Hbase jobs using Oozie. Experienced in managing and reviewing Hadoop log files. Experienced in running Hadoop streaming jobs to process terabytes of xml format data. Responsible to manage data coming from different sources. Involved in loading data from UNIX file system to HDFS. Involved in creating Hive tables, loading with data and writing hive queries which will run internally in map reduce way. Designed and implemented Map reduce-based large-scale parallel relation-learning system. Setup and benchmarked Hadoop/HBase clusters for internal use. Understanding functional specifications and documenting technical design documents. Performed unit testing for all the components using JUnit. Implemented data access using Hibernate persistence framework.

Environment: Java, Eclipse, Oracle 10g, Sub Version, Hadoop, HBase, Linux, Map Reduce, HDFS, Hive, Java (JDK 1.6), Hadoop Distribution Cloudera, Map Reduce, SQL, Windows, UNIX Shell Scripting Hadoop Developer Celgene - Township of Warren, NJ November 2012 to December 2013 Celgene Pharmaceuticals is a global integrated biopharmaceutical company that delivers truly innovative and life changing drugs for various patients. They are primarily engaged in the discovery, development and commercialization of innovative therapies designed to treat cancer and immune-inflammatory related diseases. Roles & Responsibilities: Worked on analyzing Hadoop cluster using different big data analytic tools including Pig, HBase and Sqoop. Developed Message Handler Adapter, which converts the data objects into XML message and invoke an enterprise service and vice-versa using Java, JMS and MQ Series. Business logic is implemented using Struts action components in the Struts and Hibernate framework. Migrating the needed data from MySQL into HDFS using Sqoop and importing various formats of unstructured data from logs into HDFS using Flume. Used Multithreading for invoking the database and also implemented complex modules which contain business logics using Collection, Reflection, and Generics API. Involved in Pig Latin programming. Importing And Exporting Data from MySQL/Oracle to HiveQL using SQOOP. Experienced in analyzing data with Hive and Pig. Responsible for operational support

of Production system. Loading log data directly into HDFS using Flume. Environment: Apache Hadoop, HDFS, Java Map Reduce, Eclipse, Hive, PIG, Sqoop, Flume, Oozie, Java/J2EE, Oracle 10g, SQL, PL/SQL, JSP, EJB, Struts, Hibernate, Weblogic 8.0, HTML, AJAX, Java Script, JDBC, XML, JMS. Java Developer ICICI Finance HYD September 2010 to September 2012 ICICI Finance uses real time application hosted across the globe for its loans and finance division. This multi-tier architecture application facilitates settlement between various merchant systems and transaction processing units in ICICI. Roles & Responsibilities: Worked on both Web Logic Portal 9.2 for Portal development and Web Logic 8.1 for Data Services Programming. Worked on creating EJBs that implemented business logic. Developed the presentation layer using JSP, HTML, CSS and client validations using JavaScript. Involved in designing and development of the ecommerce site using JSP, Servlet, EJBs, JavaScript and JDBC. Used Eclipse 6.0 as IDE for application development. Validated all forms using Struts validation framework and implemented Tiles framework in the presentation layer. Configured Struts framework to implement MVC design patterns. Designed and developed GUI using JSP, HTML, DHTML and CSS. Worked with JMS for messaging interface. Environment: Java, J2EE, HTML, DHTML, CSS, JavaScript, JSP, Servlets, XML, EJB, Struts, Weblogic 8.1, SQL Server 2008R2, CentOS, UNIX, LINUX, Windows 7/Vista/XP. Java Developer Nalinsoft Pvt Ltd - Hyderabad, Andhra Pradesh May 2009 to August 2010 Roles & Responsibilities: Designed Use Case Diagrams, Class Diagrams and Sequence Diagrams and Object. Diagrams, using IBM Rational Rose model the detail design of the application. Involved in designing user screens using HTML as per user requirements. Used Spring-Hibernate integration in the back end to fetch data from Oracle and MYSQL databases. Used Spring Dependency Injection properties to provide loose-coupling between layers. Implemented the Web Service client for the login authentication, credit reports and applicant information. Used Web services (SOAP) for transmission of large blocks of XML data over HTTP. Used Hibernate object relational data mapping framework to persist and retrieve the data from database. Wrote SQL queries, stored procedures, and triggers to perform back-end database operations by using SQL Server 2005. Implemented the logging mechanism using Log4j

framework. Wrote test cases in JUnit for unit testing of classes. Developed application to be implemented on Windows XP. Created application using Eclipse IDE. Installed Web Logic Server for handling HTTP Request/Response. Used Subversion for version control and created automated build scripts. Environment: JDK 1.5, Rational Rose, Spring, Web Services (JAX-WS) Hibernate, Log4j, Weblogic, SQL Server 2005, Windows XP, HTML, Eclipse, Log4J, XML, JUnit, SVN. Additional Information TECHNICAL SKILLS: Hadoop Ecosystem Hadoop, MapReduce, Sqoop, Hive, Oozie, PIG, HDFS, Zookeeper, Flume, Spark , Kafka NoSQL Databases Hbase, Cassandra, Mongo DB. Java & J2EE Technologies Core Java, Servlets, JSP, JDBC, JNDI, Java Beans. Languages C, C++, JAVA, SQL, PL/SQL, PIG Latin, HiveQL, Unix shell scripting. Frameworks MVC, Spring, Hibernate, Struts 1/2, EJB, JMS, JUnit, MR-Unit Databases Oracle 11g/10g/9i, My SQL, DB2, MS SQL Server. Application Server Apache Tomcat, JBoss, IBM Web sphere, Web Logic. Web Services WSDL, SOAP, Apache CXF, Apache Axis, REST. Methodologies Scrum, Agile, Waterfall.

Name: Nicole Shields

Email: fwilson@example.net

Phone: 228-571-6782x69813