

Hadoop Developer Hadoop Developer Hadoop Developer - Deutsche Bank, Exchange PI, NJ Bloomfield, NJ Highly self-motivated, good technical, communications and interpersonal skills. Able to work reliably under pressure. Committed team player with strong analytical and problem solving skills, ability to quickly adapt to new environments & technologies Over eight years of IT experience in analysis, design and development using Hadoop, Java and J2EE 4+ years of experience with Hadoop, HDFS, MapReduce and Hadoop Ecosystem (Pig, Hive, HBase) Hadoop Eco System: Excellent understanding/knowledge of Hadoop architecture and various components such as HDFS, JobTracker, TaskTracker, Name Node, Data Node and MapReduce programming paradigm Hadoop Administration: Coordinated with administrators in installing, configuring and using Hadoop ecosystem components like Hadoop MapReduce, HDFS, HBase, Oozie, Sqoop, Flume, Pig, Hive and Spark. Hadoop Enterprise Distributions: Experience in installing, maintaining and upgrading Hadoop distributions like Cloudera CDH 5.x/4.x, Hortonworks 2.x, MapR 1.x and DataStax 4.x. Hadoop Tools: Experience in analyzing data using HiveQL, HBase, Pig and custom MapReduce programs in Java. Hands-on experience in creating scripts for performing data-analysis with Pig, Hive and Impala. Experience in migrating applications from MapReduce to Spark using Scala. Cassandra Developer and Modeling: Configuring and setting up Cassandra Cluster. Expertise in data modeling and analysis of Cassandra and Cassandra Query Language. HBase: Ingested Data from RDBMS and Hive to HBase. Based on the requirements implemented custom coded MapReduce for HBase and used the Java client API. Data Ingestion: Using Flume designed the flow and configured the individual components. Efficiently transferred bulk data from and to traditional databases with Sqoop. Data Storage: Experience in maintaining distributed storage like HDFS, Hbase and Cassandra. SQL to NoSQL: Experience in migrating data from traditional database to NoSQL. As well included migration tools such as Sqoop and Flume while ingesting data Management and Monitoring: Maintained and coordinated centralized services using Zookeeper Messaging System: Good knowledge of Kafka for fast messaging transfer across systems Cloud Platforms: Experience in using cloud based Hadoop clusters like OpenStack and Amazon Web Services (AWS). Job Schedulers: Experience in scheduling jobs using tools like

Apache Oozie Scripting: Experienced in Hive, Pig and shell scripting Java and J2EE: Experience in developing applications using Java/J2E, Developed web applications in springframework from the scratch. Worked with Object Relational Mapping (ORM) persistence technologies like Hibernate. Experiences in working with annotation in Spring andHibernateExtensive experience and actively involved in various phases of SDLC. Experience in design patterns like Singleton pattern, Factory pattern, Data Access Objectspattern and Model-View-Controller pattern Web/Application Servers: Experience in the functional usage and deployment of applications on web servers or application servers likeApache Tomcat 8.0/7.0/6.0,WebLogic 12c/11g/10.x,WebSphere 8.x, 7.0, JBoss6.x andIIS 7/6 servers. UI Design: Comprehensive knowledge in HTML5, CSS3, JavaScript and Bootstrap Project Management: Experience in Agile andScrum project management. Authorized to work in the US for any employer Work Experience Hadoop Developer Deutsche Bank,Exchange PI, NJ March 2017 to Present Genre: Risk Management and e-statements Project Description: The Risk Based Capital group (RBC) within regulatory reporting is responsible for the preparation, calculation and reporting of capital components and capital ratios for Basel III (an international standard that banking regulators can use when creating regulations about how much capital banks need to put aside to guard against the types of financial and operational risks banks face). RBC reports the capital components and capital ratios to the Corporate Controller, Treasurer and Chief Financial Officer of bank and the regulatory agencies. In order to calculate the correct amount of capital required to remain Basel III compliant, an existing application was required to get upgradedthat will collect, transform and calculate finance data and assign the appropriate capital calculation and risk weighting. The Deutsche bank has enablede-Statement through email service from Business Credit Card Account. The bank's existing systems collect the account holder data like account details, transaction details and other account related information, which is extracted from various sources and stored into distributed file system. It processes the data to generate the statement file, which is utilized by email sending system to send out emails. Approximately 7 Million e-statements were required to be processed by the existing system. Role and Responsibilities: Handling raw data from various subsystems and load, the data from different subsystems into to

HDFS for further processing    Used Flume to collect, aggregate and store the web log data onto HDFS.    Used Sqoop to import data from Oracle database to HDFS    Developed UDFs in Pig Scripts for data cleaning and preprocessing    Used Hive to do analysis on the data and identify different correlations.    Development and maintenance of the HiveQL, Pig Scripts and MapReduce    Implemented performance-tuning techniques in Hive    Facilitating testing in different dimensions    Coordinated and modified business logic as per the business requirements.    Used Crontab for automation of scripts    Implemented POC for using Apache Impala for real time query processing over Hive    Moved all the required data to NoSQL database Cassandra from HDFS for visualization to generate reports for the BI team.    ENVIRONMENT: Cloudera CDH 4, Hadoop 2.0, Spark, Scala, Pig 0.11.0, Hive 0.10.0, Sqoop 1.4.3, Flume 1.4.0, ZooKeeper 3.4.5, Cassandra, Spark, SparkQL, Cron, Java/J2EE, Oracle 11g, WebLogic Server 11g

**Senior Hadoop Developer**

**Fifth Third Bank - Cincinnati, OH**    October 2015 to February 2017    Genre: Platform re-architecture and Risk Management

**Project Description:** Fifth Third Bank wants to provide risk management services for the entire corporate bank; the current data platform is undergoing a re-architecture in order to meet the changing data storage, processing and maintainability requirements for risk management. The primary objective of this project is to prove that a pivotal based solution will provide a central data platform for management and access to the various risk management datasets while meeting the throughput metrics for rapid data ingestion and processing of this data set with very fast turnaround results. The module of this project was about decision-making, churn analysis and sentiment analysis. To achieve the project objective, Hadoop platform DataStax Enterprise 4.7 and related tools were used.

**Role and Responsibilities:**    Coordinated with Administrators in setting up, configuring, initializing and troubleshooting DS Enterprise 4.7    Involved in loading data from UNIX file system to HDFS    Involved in defining job flows and running data streaming jobs to process terabytes of text data    Developed multiple MapReduce jobs in Java for data cleaning and preprocessing    Wrote MapReduce jobs to discover trends in data usage by users    Involved in managing and reviewing Hadoop log files    Installed and configured Hive and written HiveQL scripts.    Involved in loading and transforming large sets of structured, semi structured and unstructured

data      Involved in creating Hive tables, loading data and writing Hive queries as per business requirements.      Implemented static partitioning,dynamic partitioning and bucketing of data in Hive for improving the performance      Supported Map Reduce programs those are running on the cluster.

Involved in writing both DML and DDL operations in NoSQL database Cassandra      Responsible to manage data coming from different sources      Implemented POC on writing programs in Scala using Spark      Worked on migrating MapReduce programs into Spark using Scala      Assisted the team in their development&deployment activities      Used Web services concepts like SOAP to interact with other project within organization and for sharing information.      Involved in developing database access components using Spring DAO integrated with Hibernate for accessing the data.      Followed Hybrid (Waterfall - Scrum)principles in developing the project      ENVIRONMENT:DataStax Enterprise 4.7, Linux, Hadoop2.4.0, MapReduce, Hive 0.12.0, Pig 0.10.1, Impala, Hbase 0.96.1, Sqoop 1.4.5, Flume 1.4.0, ZooKeeper 3.4.5, Cassandra 2.1.5, Spark 1.2.1, Spark Cassandra Connector 1.2.1, SparkQL,Solr 4.10, SOAP, Spring 4.1, Hibernate 4.3, Oracle 12c Hadoop Developer Comcast - Philadelphia, PA January 2015 to September 2015 Genre: Marketing Research      Project Description:The idea of this project is,distributed batch processing infrastructure system based on Hadoop Distributed File System to access and query large-scale data. It is a high throughput real-time data collection and reporting system used by higher management decision makers in Marketing Research group to analyze future trends for their upcoming product development, marketing campaigns and to enhance business continuity by anticipating changes in customer behavior.      Role and Responsibilities:      DevelopedMapReduce layers using Javato support SequenceandAvroformatted input data.      Used PigLatin scripts to process huge data-sets in parallel for advancedtransformations.      Used HiveQL scripts and Hive User Defined Functionsto extract and analyze data over HBase input format.      Implemented partitioning andbucketing in Hive      Used Sqoop data import and export scripts between various RDBMS structuredsources to HDFSand NoSQL database HBase.      Imported data usingFlume data transfer between various data sources to HDFS.      DevelopedOozieworkflows that chain Hive/MapReduce modules for ingesting periodic/hourly input data      Monitoring Hadoop scripts which take the input from HDFS and load the

data into Hive Monitored Systemhealth, logs, and responds accordingly to any warning or failure conditions using Zookeeper. ENVIRONMENT: MapR 1.2, Hadoop 1.0.3, Hive 0.7.1, Hbase 0.90.4, ZooKeeper 3.3.4, Pig 0.9.0, Sqoop 1.3.0, Oozie 3.0.0, Flume 0.9.4 Java Developer ICICI Bank - Hyderabad, Telangana February 2013 to December 2014 Genre: Fund Transfer Integration Project Description: FTI is fund transfer Integration, which is an application enabling users to transfer fund from ownaccount to third party accounts. The fund transfer can be to any account of same bank account transfer, third party account transfer within same country and Cross border fund transfer. FTI is used almost across 16 APAC countries. Corporate & Consumer segments both uses FTI. FTI process starts with any fund transfer instruction feeds to it from Internal or external systems. Internal Systems like BSI or FC, whereas external instructions coming via SWIFT to CMG to FTI. Role and Responsibilities: Used Spring Framework for dependency injectionusingspring configuration files Developed the presentation layer using JSP, HTML, CSS and client validations using JavaScript Involved in the installation and configuration of Tomcat Server Involved in dynamic form generation, auto completion of forms and user validation functionalities using AJAX Designed, developed and maintained the data layer using Hibernate and performed configuration of Spring application framework. Created stored procedures using PL/SQL for data access layer Worked on tuning of back-end Oracle stored procedures using TOAD Developed test cases for unit testing using JUnit Developed stored procedures to extract data from Oracle database ENVIRONMENT: Java/J2EE, JSP, Servlets, Spring Framework, Hibernate, SQL/PLSQL, Web Services, WSDL, JUnit, Tomcat, Oracle 9i, and Windows. Java Developer Next Step Solutions - Hyderabad, Telangana January 2012 to January 2013 Genre: Payroll System Project Description: The goal of the project was to design an online payroll system, which enabled organizations to maintain the payroll information of their employees. The principle attributes of this payroll system included web based timesheets that recorded working hours,graphical absence calendar system, payroll generationand the ability to re-run the payroll for every single employee. Another prominent feature was the provision to view historical payroll information via the historical pay slip production functionality Role and Responsibilities: \* Involved in designing screen using HTML and CSS \* Used JavaScript to

perform checking and validations at clientside \* Worked with SpringDependencyInjection(DI) \* Involved in implementation of Spring Beans, DAO, Controller and Service layers for fast and efficient processing \* Implemented persistence application layer using Object Relational Model (ORM), Hibernate HQL and annotations \* Involved in writing HQL and SQL Queries for Oracle database \* Used log4j for logging messages \* Developed the classes for Unit Testing by using JUnit

**ENVIRONMENT:** Java/J2EE, Servlets, JSP, Spring MVC, Hibernate, Eclipse, JavaScript, AJAX, XML, Log4j, Oracle 9i, Web Logic. Jr. Java Developer Bank of India - Hyderabad, Telangana May 2009 to December 2011

**Genre:** Bank Portal

**Project Description:** Bank of India (BoI) is an Indian state-owned commercial bank with headquarters in Mumbai, Maharashtra, India. BoI is a founder member of Society for Worldwide Inter Bank Financial Telecommunications, which facilitates provision of cost-effective financial processing and communication services. The goal of this project was to improve the UI design and improve the overall performance of the system.

**Role and Responsibilities:** Worked with the front-end applications using HTML, CSS and Java Script Developed the business components used in the JSP screen. Implemented DAO patterns for building the application Used JUnit for unit testing Involved in the design of JSP screens for the Public Provident Fund and Bond modules Participated in testing phases and provided UAT support. Created jar files and deployed in WebLogicApplicationServer. Involved in writing SQL queries to create tables and stored procedure to fulfill the requirements and accommodate the business rules in Oracle database.

**ENVIRONMENT:** Java/J2EE, JDBC, Servlets, JSP, HTML, CSS, JavaScript, SQL, PL/SQL, Web Logic 6.1, Eclipse Education Bachelors of Technology in Technology JNTU - Hyderabad, Telangana Skills JAVA (8 years), ORACLE (8 years), TESTING (6 years), CSS (5 years), HTML (5 years), Cloudera CDH 4, Hadoop 2.0, Spark, Scala, Pig 0.11.0, Hive 0.10.0, Sqoop 1.4.3, Flume 1.4.0, ZooKeeper 3.4.5, Cassandra, Spark, SparkQL, Cron, Java/J2EE, Oracle 11g, WebLogic Server 11g (8 years)

**Additional Information**

**TECHNICAL SKILLS** Big Data Technologies Apache Hadoop, MapReduce, HDFS, Pig, Hive, HBase, ZooKeeper, Sqoop, Flume, Oozie, HCatalog, Hue, Tez, Avro, YARN, Ambari, Spark, Scala, Solr, Kafka

**Programming Languages** Java, .Net (C#, Vb.Net), SQL, HQL, HiveQL, CQL, Scala,

Unix Shell Scripting   Java Technologies Core Java, Servlets, JDBC, JSP, Spring, Hibernate   Web Technologies HTML 5, CSS 3, Java Script, Bootstrap, Servlets, JSP, ASP, XML, JSON   Web/App Servers Tomcat 8.0/7.0/6.0, WebLogic 12c/11g/10.x, WebSphere 8.x, 7.0, JBoss 6.x and IIS 7/6 servers   Databases RDBMS:Oracle 12c/11g, SQL Server 2012/08, MySQL5.x andNoSQL:HBase, Cassandra   Operating Systems Windows, Unix and Linux   IDE EditPlus, Eclipse, NetBeans, TOAD, SQL Server Management Studio, MySQL Workbench, Visual Studio   Version Control Git, Subversion, CVS   Testing Technologies JUnit   Reporting / ETL Tools Tableau, Informatica, SSIS, Pentaho

Name: Stephanie Brooks

Email: alexandertorres@example.net

Phone: (434)378-4343