**Rekha**

**PROFESSIONAL SUMMARY:**

* Over 6+ years of software experience as a Developer, Designer With cross platform integration experience **with Hadoop, Hive, Pig, Sqoop, HBase,J2EE, JDBC, HTML, and JavaScript.**
* Experience in Design, Development, Testing and implementation of web based applications, Client/server based applications, **Hadoop** and **SOA** based applications.
* Expertise in **HDFS architecture** and **Cluster Deployments**.
* Thorough knowledge on Development life cycle and usage of various Hadoop Eco systems **Hive, Pig, Sqoop, Map Reduce, Flume, Oozie**.
* Extensively worked on **Pig Latin scripting** using Both Local and Map Reducing mode.
* Expertise in **Hive Query Language** as part of data processing in cloud data platform and debugging **Hive** issue.
* Expertise in Optimization of **Hive tables** using optimization techniques like **Partitioning** and **Bucketing.**
* Experience with Managing and Understanding of **Hadoop Log Files.**
* Worked on **NoSQL** database such as **Hbase, Cassandra** for storing high volume data using SQOOP.
* Acquired Skills in writing **Linux Scripts**, **Pig Scripts**, **Hive queries** and **Map Reduce** programs.
* Extensive knowledge on **Hadoop** and **Hive** Security.
* Hands-on experience with **Horton works & Cloudera Distributed Hadoop (CDH)**
* Experience with **Oozie** in scheduling jobs and workflow.
* Good knowledge of Hadoop Architecture and various components such as **HDFS**, **Name Node**, **Data Node** and **MapReduce** concepts.
* Proficient in **Functionality Testing**, **Regression Testing** and **Smoke testing**. Proficient in ETL, Cubes and Reports Testing for BI applications.
* Strong in **Hadoop**, **Bigdata, NoSQL**, **SQL Server**, **Teradata** and **Oracle**. Experience in using SSIS and Informatica tools for validating ETL workflows.
* Experience in defect management tools like **JIRA** and **Quality Centre**/ALM.
* Expertize on **Data Warehousing** and **Pentaho Data Integration** Tool.
* 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.
* Proficiency in writing **SQL Queries**, **Stored Procedures – PL/SQL.**
* Experience in Production Job-scheduling, Implementation and job-monitoring using **Control-m tool.**
* Excellent problem solving skills, high analytical skills, good communication and interpersonal skills.
* Used the **Agile** and **Waterfall methodology** to develop the applications.
* Efficiency to handle emergency and high-pressure circumstances with demanding production support activity.
* Flexible and versatile to adapt to any new environment and work on any project.

**Technical skills:**

|  |  |
| --- | --- |
| **Operating Systems** | **DOS, Windows 98/XP/2000/7, Unix (Sun Solaris),Linux, Unix** |
| **Hadoop/BigData Technologies** | **Pig, Hive, HiveQL, Sqoop, Flume, Oozie, Zoo Keeper, Java Map Reduce, Redshift.** |
| **Scripting** | **Java Script, HTML, CSS, Unix Shell Scripting.** |
| **DBMS** | **ORACLE, MS-Access 2000, SQL-Server, MySQL, HBase** |
| **Web/Application Servers** | **WebSphere, Apache Tomcat 4.1/5.1/6.0/7.0.** |
| **IDE / Tools** | **Eclipse, NetBeans.** |
| **Methodologies** | **Agile, Waterfall** |
| **Distributions** | **Cloudera, Hortonworks** |

**Wells Fargo, Charlotte, NC.**

**Jan 2016 – Present.**

**Hadoop Developer/Tester**

**RESPONSIBILITIES:**

* Experience in Development, Testing and implementation of **Cloudera environment**.
* Responsible for running **Hadoop streaming jobs** to process terabytes of xml format data.
* Handled **importing of data** from various **data sources**, performed transformations using **Hive Map reduce**, loaded data into **HDFS** and extracted data from **MYSQL** into HDFS vice-versa using **Sqoop**.
* Loaded data from **Linux** file system to **HDFS**.
* Defined **job flows,**managing and reviewing **Hadoop log files**.
* Installed and configured Hive and written **Hive UDFs** in **Java**.
* Optimized **Hive joins** for large tables and developed map reduce code for full outer join of two large tables.
* Experienced in using **Hive-vectorization** for faster query process to utilize the CPU usage time.
* **Created Hive** tables, loading with data and writing hive queries which will run internally in map reduce way.
* Written Hive queries for data analysis to meet the business requirements. Validated Hive Managed and External table types.
* Design and develop **Pig Latin** scripts and pig command line transformations for data joins and custom processing of map reduce outputs.
* Was involved in executing SQL queries on Oracle, Hadoop, SQL server tables in order to view successful transactions of data and for validating data.
* Wrote **Pig** and **Hive** scripts to perform data manipulation operations that included data cleansing, joins and aggregations to process and load the data to **Redshift** cluster
* Created **HBase** tables for random read/writes by map reduce programs.
* Installed and configured Apache Hadoop, **Hive** and **Pig** environment on **AWS** platform.
* Wrote Map Reduce programs in Java to achieve the required Output.
* Assisted **SQL** Server Database Developers in code review and optimizing **SQL** queries.
* Used **Oozie** workflow engine to create the workflows and automate the **Map Reduce, Hive, Pig** jobs.
* Tested web services for **ARD** integration using **SOAPUI**.
* Experience in performing manual **User Acceptance Testing** (UAT).
* Supported in setting up **QA environment** and updating configurations for implementing scripts with **Pig** and **Sqoop**. Cluster co-ordination through **Zookeeper**.
* Cluster design, installation and configuration using HDP 2.1 stack on Azure and HDP 2.2 on premise setup.
* Implemented Name Node backup using NFS for High availability.
* Configured Hadoop environment with **Kerberos** authentication, Name nodes, and Data nodes.

**ENVIRONMENT**: Hadoop, MapReduce, HDFS, Pig, Hive, Hbase, Oozie, Cloudera CDH4.5, Cloudera CDH5.1.3, Kerberos security, Cassandra, Kafka, SQL, Linux, JIRA, Amazon Web services (AWS).

**Adobe, San Jose, CA**

**April 2014 – Dec 2015**

**Hadoop Developer/Tester**

* + - **Involved in the architectural decision with of Hadoop for application development and ecosystems like Hive, Pig, Zookeeper, flume, Hbase and Sqoop.**
    - **Performed research &POC on Hive to analyze the partitioned and bucketed data and compute various metrics to determine the performance on Hadoop cluster.**
    - **Used Sqoop to import and export data from RDBMS to HDFS & vice versa.**
    - **Involved in installing and configuring Kerberos for the authentication of users and Hadoop daemons.**
    - **Created Hive tables and involved in data loading and writing Hive UDFs.**
    - **Created Hive External tables on the existing HDFS file systems**
    - **Performed data analysis, queries on hive, pig on AMBARI(Hortonworks)**
    - **Worked on importing and exporting data from SQL Server and Teradata into HDFS and HIVE using SQOOP.**
    - **Performed development, deployment, job scheduling, testing, and validation& troubleshooting of data from development to production environment.**
    - **Involved in writing shell scripts for Access Control List permissions for roles, active directories, UNIX directories& file systems.**
    - **Provided technical and analytical expertise to complex, specialized report requests requiring higher-level data analysis and data management services.**
    - **Experience in performing manual User Acceptance Testing.**
    - **Involved in documentation of high level & detail level design of complete development phase.**
    - **Monitored deployments for successful implementation of data.**
    - **Extensively worked on changed control process of the company required for Hadoop implementation &documented the same to help Hadoop team understand & follow the overall process.**
* Wrote Test Cases from the Test Plan and manually tested the application
  + - **Worked on source target mapping of data to analyze the transformation of data for business analysis**
    - **Established healthy working relationship with team & management to ensure efficient & productive workspace& helped each other at any situation to avoid delay in project & met higher management expectations.**
    - **Learned new technology through company training & self to contribute in continuous innovation in technology.**

**ENVIRONMENT: Hadoop, MapReduce, Hive, HDFS, PIG, Sqoop, Flume, Hbase, ZooKeeper AMBARI(Hortonworks), AWS, Tidal, SQL Server, Teradata, MYSQL, PL/SQL, UNIX, TortoiseGit, MS Visio.**

**UHG, Minnetonka, MN**

**ETL Tester**

**Oct 2012 – Mar 2014.**

**RESPONSIBILITIES**:

* Extensively involved in Business Analysis and Requirements Gathering.
* Performed **ETL Informatica** testing and extensively used SQL functionalities. Used TOAD to connect to Hadoop Database.
* Analyzed business requirements and interfaced with vendors for development and enhancement of a system to support multiple business areas including Asset Management, portfolio management and risk management.
* Developed SQL scripts to validate the data loaded into warehouse and Data Mart tables using Data warehouse **ETL Informatica**.
* Extensively used Informatica power center for extraction, transformation and loading process.
* Performed data analysis and data profiling using **SQL** and **Data warehouse ETL** Informatica Data Explorer on various sources systems including Oracle and **Teradata**
* Performed Regression, **UAT** and **Database testing.** Performed data validation-using SQL.
* Involved in analyzing the requests for new, enhanced, and modified systems, against user and business requirements.
* Involved in the error checking and testing of the Data warehouse ETL procedures and programs Informatica session log.
* Conducted discussions and facilitate requirements gathering, process analysis, and application design sessions with company staff across all levels and functions
* Responsible for validating the files availability, format and data on UNIX server
* Executed **SQL queries** to verify the flow of data between back end and front end.
* Coordinated system walk through, training, **User Acceptance Testing** (UAT) and facilitate issue resolution.
* Experienced in DW projects **ETL testing** against DW testing with Informatica tools.
* Creating and executing SQL queries to perform Data Integrity testing on an Oracle Database to validate and test data using **TOAD/Exceed**.
* Tested reports developed by Business Objects Universes for both AdHoc & Canned Reporting users of Business Objects.
* Executed **SQL queries** to verify the flow of data between back end and front end.
* Performed **Data Driven Testing** of the Application for different users/data sets
* Developed Test cases for the **data warehouse** in Oracle System
* Conducted Functional testing to ensure that the application meets the requirements and to ensure usability readiness
* Experience in creating **UNIX scripts** for file transfer and file manipulation.

Environment: Informatica Power Center 8.1, Hadoop, SQL, PL/SQL, UNIX, Toad, Test Director V7.6, SSIS, Windows NT, Sun Solaris, MS SQL server, Rational Functional Tester, Rational Clear Quest.

**Krvy Group, Hyderabad, India**

**June 2010– Sep 2012.**

**ETL Tester**

**RESPONSIBILITIES**:

* Involved in analyzing system design specifications and developed **Test Plans**, **Test Scenarios** and **Test Cases** to cover overall quality assurance testing.
* Tested all the **ETL** process that was developed to fetch the data from various feeds into the data warehouse.
* Written **SQL scripts** for validating the data loaded by the ETL processes.
* Tested the reports developed using **Business Objects**.
* Perform manual testing and created various test results to track the testing results.
* Export requirement, test cases into the **Quality Center**, run manual and automation test cases from test lab and log defects from defects module.
* Followed up with standard **software testing life cycle** (STLC) to perform different types of testing for the projects.
* Tested request and response **XML's** based web services interfaces using **SOAP UI**.
* Participated in peer review meeting for created Test plan, test cases and analyzing various testing documents and process of testing life cycle.
* Create standard **testing life cycle** flow diagram for projects and presented to test team to follow testing process.
* Participated to create **Test Analysis reports** (TAR) and stored them into the **Quality Center**.
* Updated RTM in daily, send status report and daily tracking report to test manager.
* Performed **UAT testing** in the **UAT environment** prior to notify to the customer to start performing UAT testing.
* Developed **SQL Queries** to retrieve or create test data from different Oracle Test databases.
* Writing complex **SQL queries** for data validation for verifying the **SSIS Packages** and business Rules.
* Tested reports in report studio and query studio by using SQL.
* Interaction client to present future functionality of the application, gather information from them to know about their need and send this information to test manager.
* Automated and scheduled the Data stage jobs using **UNIX** Shell Scripting.
* Worked closely with the application development and design and testing team to resolve technical issues and interacted with them to ensure overall quality of the software.

Environment: ETL, Business Intelligence BI, SOAPUI, SQL, Java, JSP, JDBC, XML, UNIX, Web Logic Application Server, Oracle 9i, Quality Center 8.2, SQL Server, Quick Test Professional 8.2

**Educational Qualification:**

**Bachelor of Engineering** in Electronics and Communications Engineering April 2010

JNTU, Hyderabad, India GPA 3.75/4.0