**Rajesh Kataveni**

**rajesh.k2453@gmail.com**

**704-710-6684**

**Professional Summary:**

* Over 7+ Years of professional IT experience in analysis, architectural design, prototyping, development, Integration, and testing of applications using Java/J2EE Technologies and Good Working experience in Big Data Technologies.
  + Experience in developing Map Reduce Programs using Apache Hadoop for analyzing the big data as per the requirement.
  + Experienced in major Hadoop ecosystem's projects such as Pig, Hive, HBase and monitoring them with Cloudera Manager.
  + Extensive experience in developing Pig Latin Scripts and using Hive Query Language for data analytics.
  + Hands-on experience working on NoSQL databases including HBase, Cassandra and its integration with the Hadoop cluster.
  + Experience in implementing Spark, Scala application using higher order functions for both batch and interactive analysis requirement.
  + Good working experience using Sqoop to import data into HDFS from RDBMS and vice-versa.
  + Good knowledge in using job scheduling and monitoring tools like Oozie and Zookeeper.
  + Experience in Hadoop administration activities such as installation and configuration of clusters using Apache, Cloudera, and AWS.
  + Experienced in designing, built, and deploying a multitude application utilizing almost all the AWS stack (Including EC2, R53, S3, RDS, DynamoDB, SQS, IAM, and EMR), focusing on high-availability, fault tolerance, and auto-scaling.
  + Experienced in writing Business Requirements Document (BRD) and Functional Requirements Document (FRD).
  + Adept at creating and transforming business requirements into functional requirements and designing business models using UML diagrams - Context, Use Case, Sequence, Activity Diagrams in Enterprise Architect, MS Visio and Rational Rose.
  + Extensive experience in loading and analyzing large datasets with the Hadoop framework (MapReduce, HDFS, PIG, HIVE, Flume, Sqoop, SPARK, Impala), NoSQL databases like MongoDB, HBase, Cassandra.
  + Hands on experience in configuring and administering the Hadoop Cluster using major Hadoop Distributions like Apache Hadoop and Cloudera.
  + Developed Web-based applications using Python, Amazon Web Services, jQuery, CSS, and Model View control frameworks like Django, Flask, and JavaScript.
  + Good experience with design, coding, debug operations, reporting and data analysis utilizing python and using python libraries to speed up development.
  + Hands on experience with Bid Data environment on technologies including Hadoop.
  + Experienced in creative and effective front-end development using JSP, JavaScript, HTML 5, DHTML, XHTML Ajax and CSS.
  + Good Working experience in using different Spring modules like Spring Core Container Module, Spring Application Context Module, Spring MVC Framework module, Spring ORM Module in Web applications.
  + Used jQuery to select HTML elements, to manipulate HTML elements and to implement AJAX in Web applications. Used available plug-ins for extension of jQuery functionality.
* Developed multiple POCs using Python Spark API and deployed on the Yarn cluster, compared the performance of Spark, with Hive and SQL.
* Very good knowledge in advanced Big-Data technologies like Spark Ecosystem (Spark SQL, SparkR and Spark Streaming), Kafka and Predictive analytics.
* Involved in converting Hive/SQL queries into Spark transformations using Spark RDDs, Spark SQL and Scala.
  + Worked on reading multiple data formats on HDFS using Python Spark API.
  + Working knowledge of database such as Oracle10g/11g/12c, Microsoft SQL Server, DB2.
  + Experience in writing numerous test cases using JUnit framework with Selenium.
  + Experience in development of logging standards and mechanism based on Log4J.
  + Strong problem-solving skills, good communication, interpersonal skills and a good team player.
  + Have the motivation to take independent responsibility as well as ability to contribute and be a productive team member.

**Technical Skills:**

|  |  |
| --- | --- |
| **Programming Languages** | SQL, C, C++, Java, Core Java, Java 8, J2EE, Python, Scala, Pig Latin, HiveQL and Unix shell scripting |
| **Big Data Technologies** | Hadoop, HDFS, MapReduce, Hive, Pig, HBase, Hue, Sqoop, Storm, Kafka, Oozie, SparkSql, SparkR, Flume, Zookeeper, Cassandra, Spark,Hortonworks. |
| **Databases** | Oracle, MySQL, SQL Server, DB2, Familiar with NoSQL (HBase, Cassandra, MongoDB) |
| **Scripting & Query Languages** | UNIX Shell scripting, SQL and PL/SQL. |
| **Web Technologies** | JSP, Servlets, JavaBeans, JDBC, AWT, Swing, JSF, XML, CSS, HTML, XHTML, JavaScript, AJAX. |
| **Hadoop Paradigms** | MapReduce, YARN, In-memory computing, High Availability, Real-time Streaming. |
| **Operating Systems** | Windows, UNIX, Linux distributions (Centos, Ubuntu), Mac OS |
| **Other Tools** | Eclipse, Tableau 9.0, JUnit, QTP, JIRA, QC (Quality Center) |

**Professional Experience:**

**Client: Baird, Milwaukee, WI Jan’2018 – Till Date Role: Bigdata/Hadoop Developer**

**Responsibilities:**

* Responsible for installation and configuration of Hive, Pig, HBase and Sqoop on the Hadoop cluster and created hive tables to store the processed results in a tabular format.
  + Configured Spark Streaming to receive real-time data from the ApacheKafka and store the stream data to HDFS using Scala.
  + Developed the Sqoop scripts to make the interaction between Hive and vertical Database.
  + Processed data into HDFS by developing solutions and analyzed the data using Map Reduce PIG, and Hive to produce summary results from Hadoop to downstream systems.
  + Build servers using AWS: Importing volumes, launching EC2, creating security groups, auto-scaling, load balancers, Route53, SES and SNS in the defined virtual private connection.
  + Written Map Reduce code to process and parsing the data from various sources and storing parsed data into HBase and Hive using HBase-Hive Integration.
  + Streamed AWS log group into Lambda function to create service now incident.
  + Developed Spark code by using Scala and Spark-SQL for faster processing and testing and performed complex HiveQL queries on Hive tables.
  + Proficient in designing Row keys and Schema Design for NoSQL Database HBase and knowledge of another NoSQL database Cassandra.
  + Used Hive to perform data validation on the data ingested using scoop and flume and the cleansed data set is pushed into HBase.
  + Scheduled several times based Oozie workflow by developing Python scripts.
  + Developed Pig Latin scripts using operators such as LOAD, STORE, DUMP, FILTER, DISTINCT, FOREACH, GENERATE, GROUP, COGROUP, ORDER, LIMIT, UNION, SPLIT to extract data from data files to load into HDFS.
  + Exporting the data using Sqoop to RDBMS servers and processed that data for ETL operations.
  + Worked on S3 buckets on AWS to store Cloud Formation Templates and worked on AWS to create EC2 instances.
  + Designing ETL Data Pipeline flow to ingest the data from RDBMS source to Hadoop using a shell script, Sqoop, package and MySQL.
  + Collaborated with business users to analyze current business process and partnered with them to prepare detailed test scenarios for proof of concept execution.
  + Prepared and presented Business Requirement Document (BRD), System Requirement Specification (SRS) and Functional Requirement Document (FRD).
  + Analyse business requirements and segregated them into Use Cases. Created Use case diagrams, activity diagrams, Sequence Diagrams.
  + Organized JAD sessions to flush out requirements, performed Use Case and work flow analysis, outlined business rules, and developed domain object models.
  + End-to-end architecture and implementation of client-server systems using Scala, Akka, Java, JavaScript and related, Linux
  + Optimized the Hive tables using optimization techniques like partitions and bucketing to provide better.
  + Used Oozie workflow engine to manage interdependent Hadoop jobs and to automate several types of Hadoop jobs such as Java map-reduce Hive, Pig, and Sqoop.
  + Implementing Hadoop with the AWSEC2 system using a few instances in gathering and analyzing data log files.
  + Involved in Spark and Spark Streaming creating RDD's, applying operations -Transformation and Actions.
  + Created partitioned tables and loaded data using both static partition and dynamic partition method.
  + Developed custom Apache Spark programs in Scala to analyze and transform unstructured data.
  + Handled importing of data from various data sources, performed transformations using Hive, MapReduce, loaded data into HDFS and Extracted the data from Oracle into HDFS using Sqoop
  + Using Kafka on publish-subscribe messaging as a distributed commit log, have experienced in its fast, scalable and durability.
  + Test Driven Development (TDD) process and extensive experience with Agile and SCRUM programming methodology.
  + Implemented POC to migrate Map Reduce jobs into Spark RDD transformations using SCALA
  + Involved in Cluster maintenance, Cluster Monitoring, and Troubleshooting, Manage and review data backups and log files.
  + Designed and implemented map reduce jobs to support distributed processing using java, Hive and Apache Pig
  + Analyzing the Hadoop cluster and different BigData analytic tools including Pig, Hive, HBase, and Sqoop.
  + Improved the Performance by tuning of HIVE and map reduce.
  + Research, evaluate and utilize modern technologies/tools/frameworks around Hadoop ecosystem.

Environment: HDFS, Map Reduce, Hive, Sqoop, Pig, Flume, Vertica, Oozie Scheduler, Java, Shell Scripts, Teradata, Oracle, HBase, MongoDB, Cassandra, Cloudera, AWS, JavaScript, JSP, Kafka, Spark, Scala and ETL, Python.

**Client: Lucid Motors, Newark, CA Jan’2017 – Dec’2017 Role: Bigdata/Hadoop Developer**

**Responsibilities:**

* Contributing as a member of a high performing, the agile team focused on next-generation data &analytics
  + Build Big Data Analytics and Visualization platform for handling high-volume batch-oriented and real-time data streams.
  + Utilized Agile Scrum Methodology to help manage and organize a team with regular code review sessions.
  + Built platforms and deployed cloud-based tools and solutions with AWSEMR
  + Analyzed different big data analytics using Hive import data from RDBMS to HDFS.
  + Loaded data from diff servers to AWS S3 bucket and setting appropriate bucket permissions.
  + Reduced the overall EMR production cluster's cost (Amazon Web Services) by obtaining the best configuration for running data.
  + Upgraded the Hadoop Cluster from CDH4 to CDH5 and setup High availability Cluster to Integrate the HIVE with existing applications
  + Implemented complex big data with a focus on collecting, parsing, managing, analyzing, and visualizing large sets of data to turn information into business insights using multiple platforms in the Hadoop ecosystem.
  + Developed workflow in Oozie to automate the tasks of loading the data into HDFS and pre-processing with Pig.
  + Involved in source system analysis, data analysis, and data modeling to ETL (Extract, Transform and Load).
  + Facilitated and managed meeting sessions with committee of SMEs from various business areas including estimating resource and budget requirements.
  + Designed and implemented basic SQL queries for QA testing and report, data validation.
  + Managed plan and design of backup data centre infrastructure build-out and participate in business continuation/disaster recovery planning.
  + Written Spark programs to model data for extraction, transformation and aggregation from multiple file formats including XML, JSON, CSV& other compressed file formats.
  + Developed batch data flow using Spark code in Python, Scala, and Java
  + Build Hive tables using list partitioning and hash partitioning and created Hive Generic UDF's to process business logic with HiveQL.
  + Integrated HBase with MapReduce to move the bulk amount of data into HBase.
  + Developed SQL scripts using Spark for handling different data sets and verifying the performance over Map Reduce jobs.
  + Involved in data warehousing and Business Intelligence systems.
  + Supported MapReduce Programs those are running on the cluster and Wrote MapReduce jobs using JavaAPI.
  + Built code for real-time data ingestion using Java, MapR-Streams (Kafka) and STORM.
  + Designed unit test Data models and applications for data analytics solutions on streaming data
  + Extensively used Apache Sqoop for efficiently transferring bulk data between Apache Hadoop and relational databases (Oracle, MySQL) for predictive analytics.
  + Involved in integration of java search UI, SOLR and HDFS Involved in code deployments using continuous integration tool using Jenkins.
  + Developed Scripts and Batch Job to schedule various Hadoop Program and worked with the raw data, cleanses it and finally polishes it to the format where it can be consumed by Data Scientists to create critical insights.
  + Developed storytelling dashboards in Tableau Desktop and published them on to Tableau Server and used GitHub version controlling tools to maintain project versions.
  + Optimized the mappings using various optimization techniques and also debugged some existing mappings using the Debugger to test and fix the mappings.

**Environment:** Hadoop, Java, MapReduce, HDFS, AWS, Amazon S3, Hive, Linux, XML, Eclipse, Cloudera, CDH4/5 Distribution, Spark, Scala, HBase, MongoDB, Python, GitHub, Jenkins, SQL, QA, Sqoop, Oozie, DB2, SQL Server, Oracle 12c, MySQL.

**Client: Travelers, New York City, NY May’2015 – Dec’2016 Role: Java Developer**

**Responsibilities:**

* Was responsible for designing and developing Manage Purchase Order Web Service and presentation layer with Spring MVC.
* Responsible for Requirement gathering, coding, testing and documentation.
* Architected and deployed Java based web product matching framework that matches products across multiple retailers.
* Responsible for developing and maintaining all the entity and session beans.
* Extensively worked on Drools Rules Engine and Expectation Engine for writing Business rules Validation.
* Designed and developed the messaging framework for communication between workflow manager and other Enterprise Applications over Java class and MQ-Series using JMS.
* Designed and integrated the full-scale Spring/Hibernate persistence solution with the application architectures.
* Worked on generating the web services classes by using Service Oriented Architecture SOA, WSDL, UDDI, and SOAP.
* Developed and implemented the MVC Architectural Pattern using Spring Framework including JSP, Servlets and Action classes.
* Worked on Angular JS 1.2.0 version for client-side JavaScript MVC framework for the development of dynamic web applications.
* Worked on Node.js v0.11.10 version of Node JS framework for server-side JavaScript framework
* Downloaded data from the Amazon Cloud Database for detailed display by the app.
* Worked with the development team to create appropriate cloud solutions for client needs.
* Implemented database layer using EJB 3.0 and Java Persistence API(JPA) in maintenance projects
* Responsible for developing Use case diagrams, Class diagrams, Sequence diagrams and process flow diagrams for the modules using UML and Rational Rose.
* Involved in Preparation of the technical design documents and involved in identifying and implementation of different J2EE design patterns like Singleton and DAO etc.
* Developed the presentation layer using JSP, JSTL, HTML, JSON, XHTML, CSS and client validations using JavaScript.
* Used Backbone.js, Angular.js, React.js for building rich Internet applications &Require.js to optimize in-browser use and to Load the module and to improve the Speed.
* Worked on MongoDB database concepts such as locking, indexes, sharing, replication, schema design.
* Using Apache Camel frame work provided concrete implementations of all the widely used Enterprise Integration Patterns (EIPs) and connectivity to a great variety of transports and APIs.
* Contemporary applications rely on relational databases, NoSQL databases and messaging infrastructure for achieving the internet scale. Cloud Foundry exposes MySQL, PostgreSQL, MongoDB, RabbitMQ and Redis as services that offer the database and messaging capabilities
* Involved in designing and development of web interface using JSP, Servlets, JavaScript and JDBC for administering and managing users and clients.
* Developed the application using My Eclipse 8.5 and used eclipse standard/plug-in features for editing, debugging, compiling, and formatting and build automation.
* Handling user requests such as code enhancements, bug fixes.
* Designed reports per the need of the user.
* Created stored procedures, triggers and functions using SQL Plus to meet user requirements.
* Used Log4j for logging errors, messages and performance.
* Integrated with MDM for managing user data, login and registration information are retrieved from MDM and synced with E commerce database.

**Environment:** Java, J2ee, Java beans, Servlets, JMS, JSP, Drools, Angular 1.2.0, Node.js v0.11.10, Servlets, GWT, Hibernate, EJB, JPA, Spring, Java Script, JDBC, Backbone, UNIX, HTML, XHTML, XSLT, JSON, SOAP, WSDL, Eclipse, Web Services, MySQL, SQL Plus, LDAP, Log4j.

**Client: Sutherland Global Service, Chennai, India Jan’2013 – Dec’2014 Role: Web Developer**

**Responsibilities:**

* Worked on AGILE based development Environment and used SOAP based Web Services from scratch to develop interfaces to integrate between front end system and back end systems.
* Designed and Developed dashboard menu section, Home page, Admin home page, user module (Modify/ search users, create user's screens with assigning various roles) using Spring MVC framework, Hibernate ORM Module, Spring Core Module, XML, JSP, JQuery and XSLT.
* Research and Execution of JavaScript Frameworks, including Angular JS and Node JS.
* Configured and implemented Spring AOP for transaction management, logging and performance tracking and Used Maven to build and deploy the application.
* Involved in creation of Soap Web Services, WSDL and web methods with Annotation in hibernate, implemented Web Services in JAXP based on SOAP protocols using XML and XSLT.
* Hibernate outline work is utilized as a part of perseverance layer for mapping an article situated area model to a social database.
* Used Spring DI, Spring DAP and JDBC Template in the persistence layer for GPS, backend being Oracle10g.
* Responsible for creating tables in development schema and wrote oracle packages, procedures and functions in Oracle 10g.
* Used Web/application servers such as Tomcat and JBoss.
* Created test cases for DAO Layer and service layer using JUnit and bug tracking using JIRA.

**Environment**: Java 1.6, HTML 5, XML, Maven, Hibernate, Spring AOP, JSP, Spring MVC, Junit, JIRA 4.0, JavaScript, JQuery, Angular JS, Node JS, AJAX, JBoss, SOAP, JAX-WS, WSDL, Oracle 10g.  
Additional Information.

**Client: Cyient, Hyderabad, India June’2011 – Dec’2012 Role: Java/J2EE Developer**

**Responsibilities:**

* Involved in requirement analysis and documentation of workflows & functional diagrams.
* Designed J2EE framework, which generates different screens, based on access level using session beans, JSP & XML. Developed HTML forms & client-side validation using JavaScript.
* Wrote classes to parse XML to display data on UI.
* Implemented Session Tracking and User Authentications.
* Involved in designing and implementation of complete admin module using HTML, XML, JSP, Multithreading, servlets and Java Beans classes.
* Developed an environment where a JSP file is requested which in turn uses a Java bean that generates and sends a form associated to XML Schema.
* Experience working with MySQL server, stored procedure, design pattern, etc.
* Created multiple web-based forms using Tiles, Servlets, JSP, JSTL, and JavaScript menus.
* Implemented JSP custom tags, developed Struts Actions, Action Form, and Value Objects for presentation tier.
* Similarly, JavaScript applications were also developed to support validation and interaction for presentation layer.
* Implemented the nested tag concepts of the Struts tag libraries.
* Implemented Web Services to propagate business logics to different clients through SOAP protocol.
* Wrote Test scripts using JUnit.
* Used web services for data exchange using SOAP and WSDL.
* Hibernate was used for object/relational persistence and query service.
* Log4j was used of logging and debugging purposes.
* Interfaced with the client team and development team during acceptance testing and dry runs on the production systems.
* Developed Action Classes and Java Beans using Spring MVC architecture.
* Developed Service Classes for business logic and DAO's for Data Base.

**Environment**: Java/J2EE, Struts, JSP, Web logic, Oracle, SOAP, Multithreading, Web services, Hibernate, MySQL, Design Pattern, Shell script, Log4j, JUnit, Unix.