**SRINIVAS**





**srinivashaddev@gmail.com**

**+1( 347) 669-8792**

**Sr. Hadoop/Java Developer**

**PROFESSIONAL SUMMARY:**

* Over 8 years of IT experience in software development and support with experience in developing strategic methods for deploying **Big Data technologies** to efficiently solve Big Data processing requirement.
* Expertise in Hadoop eco system components **HDFS, Map Reduce, Yarn, HBase, Pig, Sqoop, Spark, Spark SQL, Spark Streaming,**and **Hive** for scalability, distributed computing, and high performance computing.
* Experience in using **Hive Query** Language for data Analytics.
* Experienced in Installing, Maintaining and Configuring Hadoop Cluster.
* Strong knowledge on creating and monitoring Hadoop clusters on **Amazon EC2, VM, Hortonworks Data Platform 2.1 & 2.2, CDH3, CDH4Cloudera Manager on Linux, Ubuntu OS** etc.
* Capable of processing large sets of structured, semi-structured and unstructured data and supporting systems application architecture.
* Having Good knowledge on Single node and Multi node Cluster Configurations.
* Strong knowledge in **NOSQL** column oriented databases like **HBase, Cassandra, MongoDB**, and its integration with Hadoop cluster.
* Expertise on **Scala Programming language** and **Spark Core.**
* Worked with**AWS** based data ingestion and transformations.
* Experienced in job workflow scheduling and monitoring tools like **Oozie and Zookeeper**.
* Good knowledge on **Amazon EMR,Amazon RDS S3 Buckets, Dynamo DB, RedShift**.
* Analyze data, interpret results, and convey findings in a concise and professional manner
* Partner with Data Infrastructure team and business owners to implement new data sources and ensure consistent definitions are used in reporting and analytics
* Promote full cycle approach including request analysis, creating/pulling dataset, report creation and implementation and providing final analysis to the requestor
* Very Good understanding of **SQL, ETL** and **Data Warehousing Technologies**
* Knowledge of **MS SQL Server2012/2008/2005and Oracle 11g/10g/9i** and **E-Business Suite**.
* Expert in TSQL, creating and using Stored Procedures, Views, User Defined Functions, implementing Business Intelligence solutions using **SQL Server 2000/2005/2008**.
* Developed Web-Services module for integration using **SOAP** and **REST**.
* NoSQL database experience on **HBase, Cassandra,DynamoDB**.
* Flexible with Unix/Linux and Windows Environments working with Operating Systems like **Centos 5/6, Ubuntu 13/14, Cosmos**.
* Good experience on Kafkaand Storm
* Knowledge of **java virtual machines (JVM)** and multithreaded processing.
* Strong programming skills in designing and implementation of applications using **Core Java, J2EE, JDBC, JSP, HTML, Spring Framework, Spring batch framework, Spring AOP, Struts, JavaScript, Servlets**.
* Experience in build scripts using **Maven** and do continuous integrations systems like **Jenkins**.
* Java Developer with extensive experience on various **Java Libraries, API’s,**and **frameworks**.
* Hands on development experience with **RDBMS**, including writing complex **Sql queries**, **Stored procedure,**and **triggers**.
* Have sound knowledge on designing data warehousing applications with using Tools like **Teradata, Oracle,**and **SQL Server**.
* Experience on using **Talend ETL tool**.
* Experience in working with job scheduler like Autosys and Maestro.
* Strong in databases like **Sybase, DB2, Oracle, MS SQL,Clickstream**.
* Strong understanding of **Agile Scrum** and **Waterfall SDLC methodologies**.
* Strong Working experience in **snowflake.**
* Hands on experience with automation tools such as **Puppet, Jenkins**.
* Strong communication, collaboration & team building skills with proficiency at grasping new Technical concepts quickly and utilizing them in a productive manner.
* Adept in analyzing information system needs, evaluating end-user requirements, custom designing solutions and troubleshooting information systems.
* Strong analytical and Problem solving skills.

**EDUCATIONAL QUALIFICATION:**

* Bachelor of Engineering from JNTUK, India.

**TECHNICAL SKILLS:**

|  |  |
| --- | --- |
| Hadoop/Big Data Technologies | **HDFS, Map Reduce, Sqoop, Flume, Pig, Hive, Oozie, impala, Spark, Zookeeper and Cloudera Manager,Splunk.** |
| NO SQL Database | **HBase, Cassandra** |
| Monitoring and Reporting | **Tableau, Custom shell scripts,** |
| Hadoop Distribution | **Horton Works, Cloudera, MapR** |
| Build Tools | **Maven, SQL Developer** |
| Programming & Scripting | **JAVA, C, SQL, Shell Scripting, Python, Scala** |
| Java Technologies | **Servlets, JavaBeans, JDBC, Spring, Hibernate, SOAP/Rest services** |
| Databases | **Oracle, MY SQL, MS SQL server, Teradata** |
| Web Dev. Technologies | **HTML, XML, JSON, CSS, JQUERY, JavaScript, angular JS** |
| Version Control | **SVN, CVS, GIT** |
| Operating Systems | **Linux, Unix, Mac OS-X, Cen OS, Windows10, Windows 8, Windows 7, Windows Server 2008/2003** |

**PROFESSIONAL EXPERIENCE:**

**Accolite, Madison, WI July2015-Till Date**

**Senior Hadoop Developer**

**Description:**Accolite is now one of the nation's leading providers of technology solutions and services in healthcare, finance, insurance, telecom, and manufacturing. While dealing with healthcare maintained database related to insurer, policyholder, subscriber, claims, flexible spending accounts and master savings accounts details using hadoop. Enabled Online Renewals and capture request for Endorsements.Provided Online Premium Calculator for NonRegistered/Registered Users Provided Online Customer Support like Chat, Agent Locators, Branch Locators, FAQs, Best Plan Selector, to increase the likelihood of sale. In finance domain , worked for Claiming credit for capturing settlement details. Gets Reports on matched,un- matched and disputed credits.

**Responsibilities:**

* Installed/Configured/Maintained Apache Hadoop clusters forAnalytics, application development and Hadoop tools like **Hive,HSQL Pig, HBase,OLAP, Zookeeper,Avro,parquet**and **Sqoop**on Linux **ARCH**.
* Wrote the shell scripts to monitor the health check of Hadoop daemon services and respond accordingly to any warning or failure conditions.
* Having experience working with **Devops**.
* Installed and configured **Hadoop, MapReduce, HDFS** (Hadoop Distributed File System), developed multiple MapReduce jobs in java for data cleaning.
* Having experience in doing structured modelling on unstructured data models.
* Developed data pipeline using **Flume, Sqoop, Pig** and **Java MapReduce** to ingest customer behavioral data and financial histories into HDFS for analysis.
* Involved in collecting and aggregating large amounts of log data using Apache Flume and staging data in HDFS for further analysis.
* Worked on installing cluster, commissioning & decommissioning of Data Nodes, Name Node recovery, capacity planning, and slots configuration.
* Developed **PIG Latin** scripts to extract the data from the web server output files to load into HDFS.
* Worked on**Hortonworks Data Platform(HDP).**
* Worked with **SPLUNK** to analyze and visualize data.
* Worked on **Mesos cluster** and **Marathon.**
* Used Pig as **ETL** tool to do transformations, event joins and some pre-aggregations before storing the data onto **HDFS**.
* Worked with Orchestration tools like **Airflow**.
* Write test cases, analyze and reporting test results to product teams.
* Good experience on**Clojure, Kafka** and **Storm.**
* Worked with **AWS data pipeline**.
* Worked with **Elastic Search, Postgres, Apache NIFI.**
* Hadoop workflow management using **Oozie, Azkaban,Hamake**.
* Responsible for developing data pipeline using **Azure HDInsight, flume, Sqoop** and **pig** to extract the data from weblogs and store in HDFS.
* Installed **Oozie** workflow engine to run multiple **Hive** and **Pig** Jobs, used Sqoop to import and export data from HDFS to RDBMS and vice-versa for visualization and to generate reports.
* Involved in migration of **ETL** processes from Oracle to Hive to test the easy data manipulation.
* Worked in functional, system, and regression testing activities with agile methodology.1
* Worked on Python plugin on **MySQL** workbench to upload CSV files.
* Used Hive to analyze the partitioned and bucketed data and compute various metrics for reporting.
* Worked with HDFS Storage Formats like **Avro, Orc**.
* Worked with **Accumulo** to Modify server side Key Value pairs.
* Working experience with **shiny** and **R.**
* Working experience with**Vertica, QilkSense, QilkView** and **SAP BOE.**
* Worked with NoSQL databases like **HBase, Cassandra, DynamoDB**
* Worked with**AWS** based data ingestion and transformations.
* Good experience with **Python Pig Sqoop Oozie Hadoop Streaming,Hive and Phoenix.**
* Worked on importing and exporting data from Oracle and DB2 into **HDFS** and **HIVE using Sqoop**.
* Responsible for building scalable distributed data solutions using Hadoop, and responsible for Cluster maintenance, adding and removing cluster nodes, Cluster Monitoring, and Troubleshooting, Manage and review data backups and log files.
* Developed several new MapReduce programs to analyze and transform the data to uncover insights into the customer usage patterns.
* Worked extensively with importing metadata into Hive using Sqoop and migrated existing tables and applications to work on **Hive**.
* Extract, Load, and Transfer data through **Talend**.
* Responsible for running Hadoop streaming jobs to process terabytes of xml's data, utilized cluster co-ordination services through **Zookeeper**.
* Extensive experience in using the MOM with Active MQ**, Apache storm, Apache Spark & Kafka Maven,**and **Zookeeper.**
* Worked on the core and Spark SQL modules of Spark extensively.
* Worked on Descriptive statistics Using **R**.
* Developed Kafka producer and consumers**, HBase clients, Spark,shark,Streams**and **Hadoop MapReduce** jobs along with components on **HDFS, Hive**.
* Strong Working experience in **snowflake,Clickstream.**
* Worked on **Hadoop EMC Greenplum ,Gemstone ,Gemfire**.
* Analyzed the SQL scripts and designed the solution to implement using **PySpark**.
* Experience using **Spark**with **Neo4J**where acquiring the interrelated graphical information of the insurer and to query the data from the stored graphs.
* Experience in writing batch processing huge Scala programs.
* Load and transform large sets of structured, semi structured, and unstructured data using **Hadoop/Big Data** concepts.
* Responsible for creating Hive External tables and loaded the data in to tables and query data using HQL.
* Handled importing data from various data sources, performed transformations using **Hive, MapReduce**, and loaded data into HDFS.

**Environment:Hadoop Cluster, HDFS, Hive, Pig, Sqoop,OLAP,data modelling, Linux, Hadoop Map Reduce, HBase, Shell Scripting, MongoDB, and Cassandra, Apache Spark,Neo4J**.

**Foot Locker, Harrisburg, PA Jan 2014 – June 2015**

**Big Data SoftwareDeveloper**

**Description:**This project mainly deals with pulling the data from various structured and unstructured data sources into Hadoop platform and standardized all data through a series of master data management processes to reach client goal to become efficient and effective within different geographic location in attracting more customers and improve the user experience which helps in developing the market.

**Responsibilities:**

* Performed benchmarking of HDFS and Resource manager using TestDFSIO and Tera Sort.
* Worked on **SQOOP** to import data from various relational data sources.
* Working with Flume in bringing click stream data from front facing application logs
* Worked on strategizing **SQOOP** jobs to parallelize data loads from source systems
* Participated in providing inputs for design of the ingestion patterns.
* Participated in strategizing loads without impacting front facing applications.
* Worked in agile environment using **Jira,Git.**
* Worked on design on **Hive,ANSI** data store to store the data from various data sources.
* Involved in brainstorming sessions for sizing the Hadoop cluster.
* Involved in providing inputs to analyst team for functional testing.
* Worked with source system load testing teams to perform loads while ingestion jobs are in progress.
* Worked with Continuous Integration and related tools (i.e. **Jenkins, Maven**).
* Worked on performing data standardization using **PIG scripts**.
* Worked with query engines**Tez, Apache Phoenix.**
* Worked on installation and configuration Horton works cluster ground up.
* Managed various groups for users with different queue configurations.
* Worked on building analytical data stores for data science team’s model development.
* Worked on design and development of **Oozie** works flows to perform orchestration of **PIG** and **HIVE** jobs**.**
* Worked on performance tuning of **HIVE** queries with partitioning and bucketing process.
* Worked on the core and Spark SQL modules of Spark extensively.
* Developed **Kafka** producer and consumers**, HBase** clients, **Spark,**and **Hadoop MapReduce** jobs along with components on **HDFS, Hive**.
* Worked with big data tools like **Apache Phoenix, Apache Kylin, Atscale, Apache Hue**.
* Worked with securities like **Knox, Ranger, Atlas.**
* Worked with **BI Concepts-Dataguru,Talend.**
* **Worked with**Source Code Management Tools GitHUB, Clearcase SVN, CVS,
* Working experience with Testing tools **JUNIT / SOAPUI.**
* Experienced in analyzing the SQL scripts and designed the solution to implement using **PySpark**.
* Worked with Code Quality Governance related tools (**Sonar,PMD, FindBugs, Emma, Cobertura**, etc)
* Analyzed the **SQL scripts** and designed the solution to implement using **PySpark**.

**Environment:Hadoop, HDFS, Map Reduce, Flume, Pig, Sqoop, Hive, Pig, Sqoop, Oozie, Ganglia, HBase, Shell Scripting, Apache Spark**.

**STIFEL, St. Louis, MO July 2012 – Dec 2013**

**Senior Hadoop Developer**

**Description**:The purpose of the project is to store terabytes of log information generated by the ecommerce website and extract meaningful information out of it. The solution is based on the open source BigData s/w Hadoop. The data will be stored in Hadoop file system and processed using Map/Reduce jobs. Which in turn includes getting the raw html data from the websites, Process the html to obtain product and pricing information, extract various reports out of the product pricing information and Export the information for further processing. This project is mainly for the replat forming of the current existing system which is running on Web Harvest a third-party JAR and in MySQL DB to a new cloud solution technology called Hadoop which can able to process large data sets (i.e. Tera bytes and Peta bytes of data) to meet the client requirements with the increasing competition from his retailers.

**Responsibilities:**

* Worked on Distributed/Cloud Computing (**Map Reduce/Hadoop, Hive, Pig, HBase, Sqoop, Spark AVRO, Zookeeper** etc.), Cloudera distributed Hadoop (CDH4).
* Installed and configured Hadoop **MapReduce, HDFS**, developed multiple Map Reduce jobs in java for data cleaning and processing.
* Involved in installing Hadoop Ecosystem components.
* Importing and exporting data into **HDFS, Pig, Hive and HBase** using **SQOOP**.
* Responsible to manage data coming from different sources.
* Flume and from relational database management systems using **SQOOP**.
* Responsible to manage data coming from different data sources.
* Involved in gathering the requirements, designing, development and testing.
* Worked on loading and transformation of large sets of structured, semi structured data into Hadoop system.
* Developed simple and complex **MapReduce** programs in Java for Data Analysis.
* Load data from various data sources into HDFS using **Flume**.
* Developed the **Pig UDF'S** to pre-process the data for analysis.
* Worked on Hue interface for querying the data.
* Created Hive tables to store the processed results in a tabular format.
* Developed Hive Scripts for implementing dynamic Partitions.
* Developed Pig scripts for data analysis and extended its functionality by developing custom UDF's.
* Extensive knowledge on **PIG** scripts using bags and tuples.
* Experience in managing and reviewing Hadoop log files.
* Developed workflow in **Oozie** to automate the tasks of loading the data into HDFS and pre-processing with Pig.
* Exported analyzed data to relational databases using **SQOOP** for visualization to generate reports for the BI team.

**Environment:Hadoop (CDH4), UNIX, Eclipse, HDFS, Java, MapReduce, Apache Pig, Hive, HBase, Oozie, SQOOP and MySQL.**

**Kraft Foods, Chicago, IL Oct 2011 – June 2012**

**JAVA/J2EE Developer**

**Description:** Kraft Foods, Inc. has approximately 140,000 diverse employees around the world and is mainly involved in creating food products all around United States.

**Responsibilities:**

* Involved in the project from requirements gathering and involved in various stages like Design, testing till production following agile methodology.
* Implemented **Spring MVC framework**, which includes writing Controller classes for handling requests, processing form submissions and performed validations using Commons Validator.
* Implemented the business layer by using Hibernate with Spring DAO and developed mapping files and POJO java classes using **ORM tool**.
* Designed and developed Business Services using Spring Framework (Dependency Injection) and DAO Design Patterns.
* Have Knowledge on Spring Batch, which provides Functions like processing large volumes of records, including job processing statistics, job restart, skip, and resource management.
* Implemented various design patterns in the project such as Business Delegate, Data Transfer Object, Service Locator, Data Access Object, and Singleton.
* Developed web service for web store components using **JAXB** and involved in generating stub and JAXB data model class based on annotation.
* Worked on the platforms **REST APIs, NodeJS.**
* Developed XML configuration and data description using **Hibernate**. Hibernate Transaction Manager is used to maintain the transaction persistence.
* Designed and develop web based application using **HTML5, CSS, JavaScript, AJAX, JSP** framework.
* Used Maven Deployment Descriptor Setting up build environment by writing Maven build XML, taking build, configuring, and deploying of the application in all the servers
* Implementing all the Business logic in the middle-tier **using Java classes, Java beans, used JUnit framework** for Unit testing of application.
* Involved in doing various testing efforts as per the specifications and test cases using Test Driven.
* Applied **MVC pattern of Ajax framework**, which involves creating Controllers for implementing Classes.

**Environment:JDK5.0, J2EE, Servlets, JSP, Spring, HTML, Java Script Prototypes, XML, JSTL, XPath, JQuery, Oracle 10, RAD, TTD, Web Sphere Application, SVN, MAVEN, JDBC, Windows XP, Hibernate**.

**Mind Tree, Hyderabad, India. Aug 2009 – Sept 2011**

**JAVA/J2EE Developer**

**Description:** Mind Tree specializes in application Development and Maintenance, Data Analytics, Digital Services, EAI BPM, Independent Testing, and IT Strategy & Consulting. Mind tree provides various research and Development services such as Bluetooth Solutions, Digital video surveillance, Mind Test.

**Responsibilities:**

* Involved in **Java, J2EE, struts, web services** and Hibernate in a fast-paced development environment.
* Followed agile methodology, interacted directly with the client on the features, implemented optimal solutions, and tailor application to customer needs.
* Involved in design and implementation of web tier using Servlets and JSP.
* Used Apache POI for Excel files reading.
* Developed the user interface using JSP and Java Script to view all online trading transactions.
* Designed and developed **Data Access Objects (DAO)** to access the database.
* Used DAO Factory and value object design patterns to organize and integrate the JAVA Objects
* Coded Java Server Pages for the Dynamic front end content that use Servlets and EJBs.
* Coded **HTML** pages using CSS for static content generation with JavaScript for validations.
* Used **JDBC** API to connect to the database and carry out database operations.
* Used **JSP** and **JSTL** Tag Libraries for developing User Interface components.
* Performing Code Reviews.
* Performed unit testing, system testing and integration testing.
* Involved in building and deployment of application in Linux environment.

**Environment:Java, J2EE, JDBC, Struts, SQL. Hibernate, Eclipse, Apache POI, CSS**.

**Ediko Systems Inc., Hyderabad, India. July 2008 –Aug 2009**

**Software Engineer**

**Description:**Ediko Systems Integrators, an IBM Premier Business Partner, is a specialist company delivering world class business solutions leveraging IBM Technologies. EDIKO ensures the delivery of high quality business integration solutions through the application of sound software architecture principles and using the latest IBM technologies together with agile project management techniques.

**Responsibilities:**

* Used Web Sphere for developing use cases, sequence diagrams and preliminary class diagrams for the system in **UML**.
* Extensively used Web Sphere Studio Application Developer for building, testing, and deploying applications.
* Used **Spring Framework** based on **(MVC) Model View Controller**, designed GUI screens by using HTML, JSP.
* Developed the presentation layer and GUI framework in **HTML, JSP** and Client-Side validations were done.
* Involved in Java code, which generated XML document, which in turn used XSLT to translate the content into **HTML** to present to **GUI**.
* Implemented XQuery and XPath for querying and node selection based on the client input XML files to create Java Objects.
* Used Web Sphere to develop the Entity Beans where transaction persistence is required and JDBC was used to connect to the **MySQL** database.
* Developed the user interface using the JSP pages and **DHTML** to design the dynamic **HTML** pages.
* Developed Session Beans on Web Sphere for the transactions in the application.
* Utilized **WSAD** to create **JSP**, Servlets, and EJB that pulled information from a DB2 database and sent to a front-end GUI for end users.
* In the database end, responsibilities included creation of tables, triggers, stored procedures, sub-queries, joins, integrity constraints and views.
* Worked on MQ Series with J2EE technologies (**EJB, Java Mail, JMS**, etc.) on Web Sphere server.

**Environment:  Java, EJB, IBM Web Sphere Application server, Spring, JSP, Servlets, JUnit, JDBC, XML, XSLT, CSS, DOM, HTML, MySQL, JavaScript, Oracle, UML, Clear Case, ANT**.