**Sri Harsha Gogineni**

**Email:** [**srigogi54@gmail.com**](mailto:srigogi54@gmail.com) **Mobile: (469) 364-4148**

* Around 7 years of IT experience which includes over 3 years’ experience in Big Data technologies/platforms and 5 years of extensive experience in JAVA.
* Expertise in all components of Hadoop Ecosystem - Hue, Hive, Pig, Hbase, Impala, Sqoop, Flume, Zookeeper, Oozie and Apache Spark.
* In depth understanding/knowledge of Hadoop Architecture and various components such as HDFS, Job Tracker, Task Tracker, Name Node, Data Node, and MapReduce concepts and experience in working with MapReduce programs using Apache Hadoop for working with Big Data to analyze large data sets efficiently.
* Hands-onexperience on YARN (MapReduce 2.0) architecture and components such as Resource Manager, Node Manager, Container and Application Master and execution of a MapReduce job.
* Written complex MapReduce code by implementing custom writables and writablecomparables to perform analysis on large datasets.
* Written custom writablecomparables to optimize network traffic.
* Experience in collecting the log data from different sources (webservers and social media) using Flume and storing in HDFS to perform the MapReduce jobs.
* Knowledge in installing, configuring, and using Hadoop ecosystem components like Hadoop Map Reduce, HDFS, HBase, Oozie, Hive, Sqoop, Pig and Flume.
* Strong knowledge of Pig and Hive's analytical functions, extending Hive and Pig core functionality by writing custom UDFs**.**
* Expertise in developing PIG Latin Scripts and Hive Query Language for data Analytics.
* Well-versed in and implemented Partitioning, Dynamic-Partitioning and bucketing concepts in Hive to compute data metrics.
* Integrated BI tool with Impala and analyzed the data.
* Experience with NoSQL databases like Hbase, MongoDB and Cassandra.
* Experience in importing and exporting data using Sqoop from HDFS to Relational Database Systems/ Non-Relational Database Systems and vice-versa.
* Used Oozie job scheduler to schedule MapReduce jobs and automate the job flows and Implemented cluster coordination services using Zookeeper.
* Working Knowledge in architecting Hadoop solutions including hardware recommendations, benchmarking in different levels for optimizing Map Reduce programs which run on large datasets, that are written once, executed over and over, performance tuning, administration and support.
* Reviewed the HDFS usage and system design for future scalability and fault-tolerance.
* Knowledge in creating different visualizations using Bars, Lines and Pies, Maps, Scatter plots, Gantts, Bubbles, Histograms, Heat maps, Highlight tables and application of local and global filters according to the end user requirement.
* Knowledge in designing and creating various analytical reports and Automated Dashboards to help users to identify critical KPIs and facilitate strategic planning in the organization.
* Experience in working with different relational databases like MySQL, MS SQL and Oracle.
* Strong experience in database design, writing complex SQL Queries and Stored Procedures
* Expertise in development of multi-tiered web based enterprise applications using J2EE technologies like Servlets, JSP, EJB(2.0), JDBC, JNDI, JMS, Java Beans, JAXB, JAXP(SAX/DOM), JAX-WS and JAX-RPC.
* Expertise in various faces of Software Development including analysis, design, development and deployment of applications using Servlets, JSP, Java Beans, EJB, JSTL, JMS, Struts, Spring Framework, JSF, JDBC, JNDI and Hibernate.
* Having Experience on Development applications like Eclipse, RAD etc.
* Expertise in Unit Testing, Integration Testing, System Testing and experience in preparing the Test Cases, Test Scenarios and Test plans.
* Proficient in software documentation and technical report writing.
* Versatile team player with good communication, analytical, presentation and inter-personal skills.

**Technical Skills:**

|  |  |
| --- | --- |
| **Hadoop/Big Data** | HDFS, MapReduce, Hue, Hive, Pig, Hbase, Impala, Sqoop, Flume, Zookeeper, Oozie and Apache Spark |
| **Operating Systems** | Windows, Ubuntu, RedHat Linux, Unix |
| **Java & J2EE Technologies** | Core Java, Servlets, JSP, JDBC, JNDI, Java Beans. |
| **Frameworks** | MVC, Struts, Hibernate, Spring |
| **Databases/Database Languages** | Oracle 11g/10g/9i, MySQL, DB2, MS-SQL Server, SQL, PL/SQL, NoSQL (Hbase, MongoDB) |
| **Web Technologies** | JavaScript, AJAX, HTML, XML, SOAP, CSS |
| **Programming Languages** | C, C++, Java, JQuery, Python, Unix shell scripting |
| **IDE’s** | Eclipse, Net beans |
| **Web Servers** | Web Logic, Web Sphere, Apache Tomcat 6 |
| **Build Management Tools** | Maven, Apache Ant |
| **ETL Tools** | Informatica, Pentaho, Talend |
| **Testing Tools** | Junit, Win Runner, Load Runner, QTP |
| **Predictive Modelling Tool/Statistical Programming** | R, SAS Editor, SAS Enterprise guide, SAS Miner, Tableau, IBM Cognos, IBM SPSS Modeler |

**Professional Experience:**

**Client: MetLife, Tampa, FL. July 2013 – Present**

**Role:** Senior Hadoop Developer

**Description:** The aim of this project is to provide big data solutions involving huge amounts of data. This project is designed for serving the customers in an efficient manner which involves in collecting the data of all the historic claims and load that into Big Data platform. Categorizing the diseases encountered frequently based on the claims and on the geographical location. It is Built on a demand elastic Hadoop cluster infrastructure to cater the needs of various Big Data projects, automated various Big Data workflows to process and extract analytics out of the data using MapReduce, Pig and Hive.

**Responsibilities:**

* Installed and configured HadoopMapReduce, HDFS, Developed multiple MapReduce jobs in java for data cleaning and preprocessing.
* Written MapReduce code to process and parsing the data from various sources and storing parsed data into HBase and Hive using HBase-Hive Integration.
* Worked with HBase and Hive scripts to extract, transform and load the data into HBase and Hive.
* Worked on installing cluster, commissioning & decommissioning of Datanode, Namenode recovery, capacity planning, and slots configuration.
* Worked on moving all log files generated from various sources to HDFS for further processing.
* Developed workflows using custom MapReduce, Pig, Hiveand Sqoop.
* Implemented Flume to collect the data from various sources and is loaded in to HDFS.
* Tuned the cluster for optimal performance to process these large data sets.
* Built reusable Hive UDF libraries for business requirements, which enabled users to use these UDF's in Hive Querying.
* Written Hive UDF to sort Structure fields and return complex data type.
* Responsible for loading data from UNIX file system to HDFS.
* Used Sqoop to import from different database sources and file systems to HDFS and vice versa.
* Developed suit of Unit Test Cases for Mapper, Reducer and Driver classes using MR Testing library.
* Developed workflow in Control-M to automate tasks of loading data into HDFS and preprocessing with PIG.
* Used Maven extensively for building jar files of MapReduce programs and deployed to Cluster.
* Implemented Oozie engine to chain multiple MapReduce, Hive jobs.
* Modelled Hive partitions extensively for data separation and faster data processing and followed Pig and Hive best practices for tuning.
* Exported the analyzed data to the relational databases using Sqoop for visualization and to generate reports.
* Integrated BI tool with Impala.

**Environment:** Hadoop, MapReduce, Hive QL, MySQL, HBase, HDFS, HIVE, Impala,PIG, Sqoop, Oozie, Flume, Cloudera,Zookeeper, Hue Editor, Eclipse (Kepler), Oracle 11g, PL/SQL, SQL\*PLUS, Toad 9.6, UNIX, Tableau, Control-M.

**Client: Accenture June 2011 – June 20th 2013**

**Client1: Fidelity Investments, Cary, NC.**

**Role:** Hadoop Developer

**Description:**The purpose of this project is to acquire Core metrics data, export and process in Hadoop. The Core metrics data files to be loaded to Hadoop data warehouse daily to enable business reporting and outbound feeds.We use Hadoop to store vast volumes of unstructured data allows the company to collect web logs, transaction data and social media data.

**Responsibilities:**

* Installed and configured HDFS with MapReduce architecture on 5 data centers with hundreds nodes cluster around the world.
* Interacting with the Business Requirements and the design team and preparing the Low Level Design and high level design documents.
* Provide in-depth technical and business knowledge to ensure efficient design, programming, implementation and on-going support for the application.
* Involved in identifying possible ways to improve the efficiency of the system.
* Developed multiple MapReduce jobs in java for log data cleaning and preprocessing and scheduled the job to collect aggregate the log on an hourly basis.
* Implemented MapReduce programs using Java.
* Logical implementation and interaction with HBase.
* Efficiently put and fetched data to/from HBase by writing Map/Reduce job.
* Developed Map Reduce jobs to automate transfer of data from/to HBase.
* Assisted with the addition of Hadoop processing to the IT infrastructure.
* Used flume to collect all the web log from the online ad-servers and push into HDFS.
* Implemented Map/Reduce job and execute the Map/Reduce job to process the log data from the ad-servers.
* Wrote efficient map reduce code to aggregate the log data from the Ad-server.
* Used Hive to analyze the partitioned and bucketed data and compute various metrics for reporting.
* Worked on HBase, MongoDB, and Cassandra.
* Prepared multi-cluster test harness to exercise the system for better performance.
* Developed high-performance cache, making the site stable and improving its performance.

**Environment:**Hadoop, HDFS, MapReduce, HBase, Hive, Cassandra, Hadoop distribution of Hortonworks, Cloudera, MapR, DataStax, Eclipse (Juno), JMS, MRUnit, Java Batch, SQL\* PLUS and Oracle 10g.

**Client 2: United Health Group, Eau Claire, WI**

**Role:** Senior Java/J2EE Developer

**Description:** The project deals with development of EIVR (Enterprise Interactive Voice Response) application as a call center by providing them with the information like Claims, ID card, Eligibility, and Benefits. . The EIVR is an enhancement to the VRU system and is capable of interacting with the callers both by voice and/or key presses.

**Responsibilities:**

* Involved in VUI design and creation of test plans and development of detailed test scripts based on the functional specs.
* Extracted data to test various IVR call flows such as Claims, E&B and etc.
* Involved in the initial requirements gather session to evaluate the feasibility of the project.
* Gathered and analyzed business requirements and involved in converting the functional specs into Tech Specs.
* Created Call flow diagrams for IVR applications.
* Used STRUTS framework in designing the IVR application.
* Used Java Script and ECMA script for revalidations when caller provides input thru DTMF or thru Voice.
* Performed Unit testing using URLs, JUnit and also by making calls into the system.
* Supported the application from development phase thru QA, UAT, and Production.
* Used Java Servlets and JSPs to generate VXML code dynamically.
* Developed reusable code for ApplicationRoot.jsp, Header, Footer using JSF and JSP.
* Designed the Speech Applications with Call Flow charts using the Visio and was actively involved during the VUI design and grammar development requirements phase.
* Developed API to generate the VXML pages dynamically in Java.
* Analyzed the call logs and provided recommendations where most of the calls are failing or callers getting transferred, recognition failures etc.
* Wrote External grammars and validated using Parse tool.
* Updated Design document with the call flow to reflect new dialog states.
* Assisted QA Team in defining and implementing a defect resolution process including defect priority, and severity.
* Worked in flexible schedules to deliver the project on time.
* Prepared documentation for some of the recurring defects and resolutions and business comments for those defects.

**Environment:** JAVA, J2EE (JSP, Servlets, JDBC), Struts, VXML, XML/XSL, DB2, JUnit, Web Services, IBM RAD, Log4j, SQL.

**HDFC Bank, Hyderabad, India December 2009 – June 2011**

**Role:** Programmer Analyst

**Description:** This Project aims to implement the infrastructure of Java Message Service (JMS). This project developed in J2EE package using JMS API’s provides services for Exchange for message between components in a distributed environment. It supports both Synchronous and Asynchronous messaging and the receiver receives the message according to selection of the message format. The message will be stored in Database and it can be retrieved whenever sender or receiver requires.

**Responsibilities:**

* Developed JMS API using J2EE package.
* Made use of Java script for client side validation.
* Used Struts Framework for implementing the MVC Architecture.
* Wrote various Struts action classes to implement the business logic.
* Involved in the design of the project using UML Use Case Diagrams, Sequence Diagrams, Object diagrams, and Class Diagrams.
* Understand concepts related to and written code for advanced topics such as Java IO, serialization and multithreading.
* Used DISPLAY TAGS in the presentation layer for better look and feel of the web pages.
* Developed Packages to validate data from Flat Files and insert into various tables in Oracle Database.
* Provided UNIX scripting to drive automatic generation of static web pages with dynamic news content.
* Participated in requirements analysis to figure out various inputs correlated with their scenarios in Asset Liability Management (ALM).
* Assisted design and development teams in identifying DB objects and their associated fields in creating forms for ALM modules.
* Also involved in developing PL/SQL Procedures, Functions, Triggers and Packages to provide backend security and data consistency.
* Responsible for performing Code Reviewing and Debugging.

**Environment:** Java, J2EE, UML, Struts, HTML, CSS, Java Script, Oracle 9i, SQL\*Plus, PL/SQL, MS Access, UNIX Shell Scripting.

**Basavatarakam Indo-American Cancer Hospital & Research Institute, Hyderabad, India**

**May 2007 - December 2009**

**Role:** Programmer Analyst

**Description:** The objective of this project is to design a system to keep track of employee data such as personal information, title, working hours, salary, department info, etc. The system allows management to keep track of the employees and optimize the usage of their skills.

**Responsibilities:**

* Gathered requirements from client, analyzed and prepared the requirement specification document.
* Object oriented design using UML. Rational Rose used in implementing UML.
* Configured application connectivity using JDBC.
* Designed and developed all user interfaces using JSP.
* Deployed the application in Apache Tomcat.
* Involved in API development using Core Java concepts.
* Used HTML, CSS, JSP, and JavaScript for Front End User Interface design.
* Worked with the collection libraries.
* Involved in Database designing and developedin SQL Server.
* Used development environment integrated with Eclipse.
* Integrated the Java application to end-users.
* Involved in production support.

**Environment:**Java/J2EE (JSP, Servlet), Eclipse, Struts 2.0, Hibernate 3.3, JPA, XML, WebLogic, DB2, ANT, Clear Case, JUnit, UML, XSLT.