**SURESH KALAKOTA Email:** **sureshk.bigdata51@gmail.com**

Mobile**:** **7676 02 0612**

------------------------------------------------------------------------------------------------------------

**OVERVIEW:**

To build a successful career in software industry and to be a part of highly motivated team wherein I have an opportunity to demonstrate and develop my skills, make a positive contribution and help organization to achieve its goals.

**PROFESSIONAL SUMMARY:**

* Having 3.9+ years of IT-industry, 2 year experience in Big Data and Hadoop Ecosystem.
* Involved in installing Apache Hadoop Ecosystem components and Cloudera.
* Understanding of **Name node, Data node, Job Tracker, Secondary Name node, Task Tracker**.
* Hands on with MapReduce Programming Model.
* Involved in configuring Pig, Hive, Hbase, Zookeeper, Oozie, Phoenix with SquirrelSQL.
* Hands on experience on MapReduce, Hive, Sqoop.
* Sound knowledge on Pig scripts.
* Sound knowledge on No-Sql data bases like MangoDB,Hbase.
* Involved in Integrating Hive with Hbase using Phoenix and SquirrelSQL, Pig with Hbase.
* Strong experience on Cluster setup and commissioning, de-commissioning and tune up of cluster nodes.
* In depth knowledge on Object Oriented Concepts and MVC Architecture.
* Sound knowledge on Scala and Spark.
* Participated in multiple big data POC to evaluate different architectures, tools and vendor products.
* Extensive experience in web applications using HTML, Java Script.
* Having Knowledge on **loading XML file into Hive tables** without using MapReduce.

**PROFESSIONAL EXPERIENCE :**

* Worked as a Hadoop developer with Canarys Automations Pvt Ltd, from Sep 2014 to till now.
* Worked as a Java Developer with Zensar Technologies Ltd, from Oct 2012 to Aug 2014.

**Educational Qualifications:**

* Swarna Bharathi College of engineering, Khammam, in April 2011.

**Project Experience:**

**Hadoop Developer**

**# Current Project : PRTS - RAN**

Environment : Hadoop 2.x, HDFS, Yarn, Hive, Sqoop, HBase, Sqlserver.

Cluster Size : 96 Node Cluster.

Distribution : Cloudera - CDH 4.2

Duration : August 2015 to Till date

Vendor : Alcatel lucent(1X) and Ruckus Wireless

**Description:**

The scope of this project is to maintain and store the operational and parameters data collected from the multiple vendors networks by the mediation team into the OMS data store and make it available for RF engineers to boost the network performance.

**Responsibilities:**

* + Working with Hadoop **Distributed File System.**
  + Involved in **importing data** from MySQL to HDFS using SQOOP.
  + Involved in creating Hive tables, loading with data and writing hive queries.
  + Involved in Preparing **Test cases** Document.
  + Involved in **Integrating Hive and HBase** to store the operational data.
  + Monitoring the Jobs through **Oozie**.

**# POC #1 :**

**Title : Retail Marketing**

Environment : Hadoop 1.2.1, MapReduce, Hive, Sqoop, HBase, MYSQL.

Cluster Size : 10 Node Cluster.

Distribution : Apache Hadoop (1.x)

Duration : Feb 2015 to July 2015

**Description:**

The purpose of the project is to store terabytes of log information generated by the ecommerce website. The solution is based on the open source Big Data s/w Hadoop. we are using Hadoop to Store that log data into Single place like HDFS. Process the data by using Hive to get the Price and Product details once we got the Info we need to share that info for further analysis by using Hive.

**Responsibilities:**

* Working with Hadoop **Distributed File System.**
* Involved in **importing data** from MySQL to HDFS using SQOOP.
* Involved in writing Pig Scripts.
* Involved in writing **Hive Queries**.
* Involved in Integrating Hive and HBase to operational data.
* **Commissioning and Decommissioning** the Hadoop nodes & Data Re-balancing.
* **Balancing Hadoop cluster** using balancer utilities to spread data across the cluster equally.
* Writing Oozie Workflows.
* Involved in creating Hive tables, loading with data and writing hive queries which will run internally in MapReduce way.
* **POC #2**

**Title: Web-log Analysis using Hadoop with Hive Tools : Hadoop1.x, Hive**

**Description:** This POC has been designed to analyze the web-log data. The data gets loaded into hive warehouse. During the loading process, the data gets formatted using Hive RegexSerde. Later on, using Hive the data gets portioned based on country, time and keywords. This structured data gets analyzed by Hive based on the business requirements using Hive queries.

* **POC #3**

**Title: Integrating Hive with Hbase on Hadoop.**

**Tools: Hadoop, Hive, Zookeeper, Hbase, Phoenix, SquirrelSQL**.

**Description:** In this integration initially load the data into Hadoop HDFS and analyze the data by using Hive and Pig. Updating data in Hive can be done by integrating Hive with Hbase. SQL queries have been used to execute on top of Hbase using Phoenix. SquirrelSQL can also be used to achieve better performance**.**

**Project #2:**

**Project Title** : Health Care Unit**.**

#### **Duration** : Oct 2012 to till Aug 2014.

#### **Role** : Developer.

#### **Client** : IHC.

**Description:** IHC is a cloud based system designed to be universally accessed with different health care units like facilities and pharmacies. To maintaining resident medical health records, prescription management, scheduling process, prescription refilling and medical tracking.

**Responsibilities:**

* Attending client meetings to gather Requirements.
* Analyzing the requirements and performing effort estimations.
* Involved in developing JSPs, actions, data access objects.
* Written SQL queries to retrieve data from tables.
* Involved in bug fixing.

**Areas of Experience:**

* + - Big Data Ecosystems : Hadoop-HDFS, MapReduce, Hive, Sqoop, HBase Oozie, Have an idea about Flume, Zookeeper ,Scala,Spark .
    - Programming Languages : Java,Scala
    - Web Technologies : Servlets, HTML
    - Databases : Oracle 10g, MySQL, Sqlserver NoSQL - HBase,.
    - Tools : Eclipse, IntelliJ.
    - Scripting Languages : JSP,PIG
    - Platforms : Windows , Unix .