

SAMPATH KUMAR S

E-mail:sheelam.sampath498@gmail.com

Mobile:+91-9071263188

Experience Summary:

- Currently working as a Software Engineer with 3+ Years of IT Experience in Customizing, Implementing, and Supporting enhance functionality to Cisco Application in E-Commerce as well as Health Care Domain with having good hands on Hadoop Big data.
- Seeking roles in Big Data, Apache Hadoop, HDFS, YARN, Hive, Pig, Sqoop, Spark SQL, Spark Core and Scala.
- Good working knowledge with Apache Sqoop, Hive and Pig.
- Very good understanding of Partitions, Bucketing concepts in Hive and designed both Managed and External tables in Hive to optimize performance.
- Involved in writing the hive, Pig scripts to reduce the job execution time.
- Having good exposure on Hadoop Trouble Shooting, Debugging and Performance Tuning.
- Strong Knowledge of Hadoop and Hive's analytical functions.
- Converting business requirements into functional/technical specifications.
- Documenting the functional and code flows in various aspects related to the product.
- Capable of processing large sets of structured, semi-structured and unstructured data and supporting systems application architecture.
- Experience in importing and exporting the data using Sqoop from HDFS to Relational Database systems and vice-versa.
- Efficient in building Hive, Pig and MapReduce scripts.
- Scheduling all Hadoop/ Pig/ Hive/ Sqoop/Spark Actions using tidal.
- Resolved defects/bugs, environmental issues during QA testing, pre-production, and production.
- Excellent problem solving skills with a strong technical background and result oriented team player with excellent communication skills.

Technical Skills:

Big Data	: Hive, Pig, Sqoop, HDFS, MapReduce and spark with Scala
Database	: Oracle, MySQL and Teradata
Other Tools	: SQL Developer, Winscp, Putty, Tidal Eclipse and SSH
Operating Systems	: Windows and UNIX.
Languages	: SQL, Core Java and Scala

Work Experience:

- Currently working with **Zensar Technology Pvt. Ltd**, as Software Engineer from Apr 2018 to till date
- Worked with **QuintilesIMS (IQVIA) Pvt. Ltd**, Bangalore as Software Engineer from Feb 2016 to Apr 2018.

Academic Profile:

- **B-Tech** from **JNTUH**, Hyderabad.

Project Profiles:

Project 1:

Project Title	: CPS
Client	: Cisco Systems, Inc, Bangalore
Role	: Hadoop Developer
Environment	: Hive, Sqoop, Spark, HDFS, Teradata, Tidal, CCW, Kafka, UNIX and Informatica
Duration	: Apr 2018 - Till Date

Description:

Smart Accounts give you full visibility of the Cisco software assets in your company. Now you can easily manage and control your licenses.

Smart Accounts are a new way to manage your software licenses. They provide your company with a central location where you can manage Cisco licenses across the entire organization. Smart Accounts are similar to online bank accounts. You can view, store, manage, and move your Cisco software assets to where they are needed.

Today, you can experience the benefits of Smart Accounts in many Cisco portals. These include the Cisco Smart Software Manager and Cisco Commerce (formerly CCW). They're available for Enterprise License Agreements (ELAs) and for traditional PAK licenses in the License Registration Portal.

In the future, Smart Accounts will be positioned as the one location where you will be able to manage all your Cisco subscriptions, contracts, hardware, and services.

Responsibilities:

- Responsible for building scalable distributed data solutions using **Hadoop** cluster environment with MapR distribution.

HADOOP-Developer

- Replaced default Derby metadata storage system for Hive with MySQL system.
- Written the Apache Spark scripts to process the HDFS data.
- Created Hive tables to store the processed results in a tabular format.
- Create & update Hive schema to support a highly-available data warehouse for BI teams.
- Developed the Sqoop scripts in order to make the interaction between Hive and Teradata Database.
- TES (Tidal) is use to schedule the job has per the client time.
- Involved in gathering the requirements, designing, development and testing.
- Writing the script files for processing data and loading to HDFS.
- Completely involved in the requirement analysis phase.
- Moved all log/text files generated by various products into HDFS location
- Created External Hive Table on top of parsed data.
- Supporting the production by resolving the issues raised by the client.

Project Profiles:

Project 2:

Project Title	: CSOI&M
Client	: Eli Lilly pharmaceutical, Indianapolis, USA
Role	: Hadoop Developer
Environment	: Hive, Sqoop, HDFS, and MySQL
Duration	: Mar 2017 – Apr 2018

Description:

Eli Lilly and Company(NYSE: LLY) is an American global Pharmaceutical company In addition to internal research and development activities Lilly is also involved in publicly funded research projects with other industrial and academic partners. One example in the area of non-clinical safety assessment is the InnoMedPredTox, a collaboration with pharmaceutical companies, research organizations and improve the safety of drugs.

Eli Lilly is a global contract research organization (CRO) providing discovery, development and post-approval services as well as compound partnering programs. The company's clients and partners include pharmaceutical, biotechnology, medical device, and academic and government organizations. Lilly is mainly known for conducting drug development for its clients, it also invests in developing new products through its compound partnering program. Oracle kind of data box with the same reason Eli Lilly wants to move it Hadoop, where exactly we can handle massive amount of data by means of its cluster nodes and also to satisfy the scaling needs of the Eli Lilly business operation.

Responsibilities:

- Moved all generic data flat files generated from various vender and stores to HDFS for further processing.
- Written the Apache PIG scripts to process the HDFS data.
- Created Hive tables to store the processed results in a tabular format.
- Developed the Sqoop scripts in order to make the interaction between Pig and Oracle Database.
- Oozie is use to schedule the job has per the client time.
- Involved in gathering the requirements, designing, development and testing.
- Writing the script files for processing data and loading to HDFS.
- Completely involved in the requirement analysis phase.
- Moved all log/text files generated by various products into HDFS location
- Created External Hive Table on top of parsed data.
- Supporting the production by resolving the issues raised by the client.

Project 3:

Project Title	: Siro
Client	: SIRO Clinpharm Pvt Ltd, USA
Role	: Hadoop Developer
Duration	: Feb 2016 – Mar2017
Environment	: Hadoop, Apache Pig, Hive, Sqoop, MapReduce and HDFS.

Description:

SIRO Clinpharm's extensive clinical network and experience assist customers in strategy planning and identifying the most efficient and rapid routes of success in conducting clinical trials. Our experience in conducting clinical study feasibility pans across the globe.

Quintiles doe's clinical research projects we helping pharmaceutical research companies bringing the best drugs without any side effects in to the market. Quintiles India has the pride of holding some of the solid relationships with the leading medical centers as well as with some of the experienced principal investigators. All the studies that are conducted by Quintiles are in-line with FDA (Food and Drug Administration) and ICH GCP (International Conference on Harmonization-Good Clinical Practice) standards.

This project aims to move all log data from individual servers to HDFS as the main log storage and management system and then perform analysis on these HDFS data-sets. Flume was used to move the log data periodically into HDFS. Once the data-set is inside HDFS, Pig and Hive were used to perform various analyses.

Responsibilities:

- Attending daily status meeting with business users to discuss on open issues
- Involved in transferring files from Local to HDFS.
- Written the Apache PIG scripts to process the HDFS data.
- Involved in writing queries with Hive QL and Pig.
- Involved in database connection by using SQOOP.
- Process and analyze the data from Hive tables using HiveQL.
- Analyzed transactions using Pig scripts and Hive to generate reports for end users

Date:

Sampath Kumar