Esakki Thangappa Pillai

Email ID: [esakkisankart@gmail.com](mailto:esakkisankart@gmail.com)

Contact No: +91 9500804123

**EXPERIENCE SUMMARY:**

* Over 4.6 years of professional IT experience which includes experience in Core JAVA, Big data ecosystem, Spark related technologies.
* Having 2.6 years of hands on experience in Big Data using Core Java, Hadoop, Sqoop, Hive, Apache Spark and Scala.
* Involved in Creating and Performance Tuning of Spark Jobs.
* Involved in creating tables, partitioning, bucketing of table and creating UDF’s in Hive.
* Good knowledge on HDFS, Hive, Sqoop, Spark, Oozie.
* Excellent communication, interpersonal, analytical skills, and strong ability to perform as part of team.
* Good experience in interacting with client and understand client requirements.
* Exceptional ability to learn new concepts.
* Good knowledge of Agile Methodology

**SOFTWARE SKILLS:**

|  |  |
| --- | --- |
| **Technology Area** | **Tools** |
| Operating System | Linux , Windows |
| Big Data/ Hadoop Ecosystem | Hadoop ,HDFS, Map Reduce, Hive, Pig, Sqoop, Oozie, Flume, Spark |
| Database | Oracle |
| Languages | Java |
| Hadoop Distribution | Horton Works |
| Others | WinSCP, Putty, Jira |

**PROFESSIONAL EXPERIENCE:**

### Organization 1: Computer Science Corporation.

Duration: July 2013 – December 2015

Organization 2: Capgemini India Limited,

Duration: January 2016 – Till Now.

**PROJECTS WORKED:**

**PROJECT 1:**

**Title: Transactional Processing Data Lake.**

**Client: Standard Chartered Bank**

Skills:Spark, Sparksql, Hive, HDFS

Duration: March 2017 to till now

Role:Associate Consultant.

**PROJECT DESCRIPTION:**

SCB sends offline batch and online data to various downstream applications, warehouses, data marts and online messages from various countries/locations. Till date the source systems are sending redundant information to same or various other downstream which increases longer batch run for the source system impacting their EOD process. So the purpose of this project is to source all TP data to support downstream applications. This platform will aide in performing the sourcing, staging, normalizing and reporting functions of the data from the various TP systems.

**PROJECT RESPONSIBILITIES:**

* Participated in design Analysis.
* Gathering requirements from BA Team.
* Integrating with meta App team to make sure that all the tables, Json and avro schema are in place.
* Developed Spark Sql to transform the Business Logics.
* Created Hive External tables as per requirement defined with appropriate partitions, intended for efficiency
* Providing the statistics to the management on daily bases.

**PROJECT 2:**

**Title: Hadoop Sourcing/Extraction project.**

**Client: Standard Chartered Bank**

Skills:Hive, Oozie, HDFS, and Map Reduce

Duration: January 2016 to February 2017

Role:Associate Consultant.

**PROJECT DESCRIPTION:**

Project initiated to warehouse the data from different TP/Interface systems in Hadoop data lake (EDM). Data from TP system will be extracted into a fixed/variable length file and transferred through ISIS/CDM/NDM to Hadoop sever. From there it gets loaded into hive tables through the defined framework. Data in Hadoop Lake is considered as T1 layer where it is a direct dump from Source. From T1 layer data get filtered based on requirements to T2 layer/Hadoop layer by applying the required business logics. . It gets further moved to T3 layer with further business logics from their final data get extracted and transferred to require down streams.

**PROJECT RESPONSIBILITIES:**

* Understanding the Business requirements functionally and translating them

Technically

* Performing Impact Analysis, design, coding and testing for the service requirements.
* Providing assistance for rest of the team on their requirements/issues in the Application
* Contributing ideas to improve the performance of existing application and to reduce the efforts in Maintenance/enhancement projects
* Providing the required assistance for QA.
* Based on the request from downstream systems – Modification of HIVE Tables / Databases.
* Testing the same by creating views, etc.
* Monitoring the batch runs for the applications in PRODUCTION Cluster
* Sourcing the data into the HIVE tables manually for HISTORICAL DATA
* Cleanup of the files in the Temp paths.
* Adding new applications to source and then provide data to downstream systems.
* Raising tickets if there are any issues in Data post the job failure
* Analyzing the areas of improvement in the HIVE Queries / Jobs run.
* Checking the reconciliation of the data/record counts in the source and the HIVE tables manually if there are any alerts triggered.

**PROJECT 3:**

**Title: Insurance Premium and Customers analysis: LF System**

Client: **Lansforsakringar**

Skills:HDFS, Hive, Sqoop, Oozie, Oracle

Duration: January 2015 to December 2015

**PROJECT DESCRIPTION:**

Länsförsäkringar AB is a financial service provider. The company provides banking and insurance products to individual and corporate customers. This project is to develop an analytical result for Market Intelligence which contains in depth information about data. The statistical report utilizes a wide range of primary and secondary sources, which are analysed and presented to business.

**RESPONSIBILITIES:**

* Major responsibilities include Development using Apache Hive, Sqoop, etc.
* Participated in different meetings with business and SME's.
* Involved in loading and transforming large sets of data and analyzed them by running Hive queries
* Worked on different Hadoop file formats such as Row Columnar, Optimized Row Columnar and Sequence File in Hive.

**PROJECT 4:**

**Title: ETL FACTORY.**

**Client: Network Rail.**

Skills:SQL, HP Quality Center, Oracle.

Domain: Core Banking.

Duration: July 2013 to December 2014.

**PROJECT DESCRIPTION:**

Network Rail is the owner and infrastructure manager of most of the rail network in England. The Enhanced Decision Support (EDS) Tools delivered through the ETL (Extract, Transform, Load) Factory will leverage the Asset Data Store to deliver the dashboard and reporting capability as detailed in the appropriate sections of this document. Network Rail requires a trusted source of asset-related information and insight, from which informed business decisions can be made to balance risk, asset performance and funding. Reliable and accurate asset information provides the foundation for improved safety, improved rail system performance and the implementation of a least-whole-life-cost management policy approach to the management of all railway assets and it is through the implementation of the Asset Data Store and Decision Support Tools that this vision will be delivered.

**PROJECT RESPONSIBILITIES:**

* Understanding the requirements and preparing Requirement Clarifications document
* Designing, reviewing and executing test cases based on requirements specification
* Tracking and reporting bugs and
* Preparation of Sql Queries based on Technical Specs.
* Validation of data before and after ETL Loading.
* Reporting and Tracking defects and preparation of Metrics Report
* Testing the Reports by executing SQL queries and comparing the values.

**ACADEMIC CREDENTIALS:**

* B.E from Cape Institute Of Technology in the year 2012 with **8.4 (CGPA)**
* HSC from Govt. Higher Secondary School, Kumarapuram Thoopur in the year 2008 with **87.5%**
* SSLC from Govt. High School, Eraviputhoor in the year 2006 with **92.4%**

**Appreciations and Awards**

* Awarded with many team level Associate Recognition Awards.
* Appreciated by Onsite team on many occasions for good work done.

**PERSONAL VITAE:**

Name : Esakki .T

Date of Birth : March 07, 1991

Gender : Male

Marital Status : Single

Place: Chennai. Signature

Date: