**EZHILARASI CHINNAPAIYAN**

**Email:** [**ezhil.ananth@gmail.com**](mailto:ezhil.ananth@gmail.com)

**PH: +1 469 846 8345**

**Big Data Engineer**

**PROFESSIONAL SUMMARY**

* Overall 9 years of IT experience in Analysis, design, development, implementation, maintenance and support with experience in developing strategic methods for deploying big data technologies to efficiently solve Big Data processing requirement.
* Around 5 years of experience on BIG DATA using HADOOP framework and related technologies such as HDFS, MapReduce, Spark,Scala, HIVE, OOZIE, SQOOP.
* Experience in data analysis using HIVE and custom Map Reduce programs.
* Expert level experience in designing, building and managing applications to process large amounts of data in a Hadoop ecosystem
* Extensive experience with performance tuning applications on Hadoop and configuring Hadoop systems to maximize performance
* Explored the Spark to improve the performance and optimization using various transformations in Hadoop using Spark-Context, Spark-SQL, Data Frame and Pair RDD's.
* Performed analysis on Datamart using spark-SQL joins and window functions.
* Hands of Experience in GCP,BigQuery,GCS bucket, DataProc.
* Hands on experience in resolving production issue in critical situation.
* Responsible for implementing a generic framework to handle different data collection methodologies from the client primary data sources, validate transform using spark and load into S3 or GCP buckets.
* Extensive experience on SQL , Data Analysis ,SPARK, SCALA in GCP environment
* Developed analytical components using SCALA, SPARK, HBASE and SPARK STREAM.
* Performed Hadoop backup Strategy to take the backup of hive, HDFS, HBase, oozie, Airflow etc.
* Experience in importing and exporting data using Sqoop from HDFS to Relational Database Systems (RDBMS) and from RDBMS to HDFS.
* Created the HBASE tables for validation, audit and offset management table.
* Created logical view instead of tables in order to enhance the performance of HIVE queries
* Involved in developing Hive DDLS to create, alter and drop Hive tables
* Pretty Good Knowledge on HIVE Optimization techniques like Vectorization and column-based optimization.
* Written OOZIE , AIRFLOW workflow to invoke the Jobs in predefined Interval.
* On Other Hand working on POC with Kafka and NIFI to pull the real-time events into Hadoop Box.
* Exploring with the Spark for improving the performance and optimization of the existing algorithms in Hadoop using Spark Context, SPARK-SQL, DATA FRAME, PAIR RDD'S and YARN.
* Developed Spark Structured Streaming Job and created STREAMING DATA FRAME.
* Experience in creating BATCH DATA FRAME and STREAMING DATA FRAME.
* Experience with migration to GCP from HDFS.
* Solid understaning of RDBMS database concepts including performance tuning and query optimization.
* Experience in  Java, J2EE, Java Scripting, HTML, JSP
* Good working experience on Hadoop tools related to Data warehousing like Hive and also involved in extracting the data from these tools on to the cluster using Sqoop.

**TECHNICAL SKILLS:**

* **Hadoop/Spark Ecosystem**: Hadoop, Map Reduce, Hive, YARN, Flume, Sqoop, Oozie, Zookeeper, Spark, Airflow, Cassandra, HBase.
* **Hadoop Distribution**: Cloudera distribution and Horton works.
* **Cloud:** GCP – Big Query, GCP buckets, dataProc clusters AWS – S3, EC2,EMR
* **Programming Languages:** Scala, Java, SQL, HQL
* **Databases:** Oracle, MySQL, SQL Server, MS Access, HBase, Cassandra, MongoDB
* **Operating Systems:** Linux, Windows, Ubuntu.
* **IDE** : IntelliJ, Eclipse,Pycharm
* **Programming Language**: C, C++, Java.
* **Scripting Language**: JSP, HTML, XML.
* **Version Control**: BitBucket,git
* **Methodology**: Agile.

**Verizon Technologies, TX April 2022 – Till date**

**Big Data Developer/GCP**

* **Data Engineer:**
* Hands of Experience in GCP,BigQuery,GCS bucket, DataProc.
* Designed, implemented and developed ETL solutions for data ingestion, cleansing, business rules execution as per the business requirements.
* Big data Developer and developed code with various transformations using hive sql and spark scala .
* Expertise On optimizing spark Jobs in GCP.
* Written oozie workflow to invoke the Jobs in predefined Intervals.
* Expertise in Hive/SQL queries into Spark transformations using Spark RDDs and Scala
* Experienced in working with spark eco system using Spark SQL and Scala queries on different formats like Text file, CSV file and Performed necessary Transformations and Aggregation on data model and persists the data in HDFS.
* Familiar with hashing the sensitive data using Jenkins job .
* Knowledge and experience to use code versioning tools such as Git.
* Experience with databases, PostgreSQL preferred and Oracle,
* Team player and individual contributor.
* Ability to accurately identify root cause of technical problems.
* Used Hive to analyze the Partitioned and Bucketed data and compute various metrics for reporting.
* Expertise On optimizing spark Jobs when dealing with Huge joins and data Skew.
* Written oozie workflow to invoke the Jobs in predefined Intervals.
* Performed necessary Transformations and Aggregation on the fly to build the common learner data model and persists the data in HDFS.
* Explored the usage of Spark for improving the performance and optimization of the existing algorithms in Hadoop using Spark Context, Spark SQL and Spark Yarn.
* Used Hive to analyze the Partitioned and Bucketed data and compute various metrics for reporting.
* Involved in converting Hive/SQL queries into Spark transformations using Spark dataframes and Scala
* **Production Support :**
* Responsible for monitoring the production servers, scheduled jobs, incident management and receiving incidents and requests from end-users.
* Analyze the available data and find the root cause of the problem.
* Analyzing the requests and either responding to the end user with a solution or escalating it to the other IT teams.
* Ability to prioritise work to successfully deliver service to agreed levels in a diverse and constantly changing technical environment.

**Environment**:  Spark, Hive, Spark SQL, Oozie, Scala, Maven, Jupiter Notebook, GCP, GCS buckets, Bigquery,Dataproc Clusters, Unix Shell Scripting,PostgresSQL, Intellij,

**Zubaid InfoTech, Chennai, India December 2016 –November 2018**

**Data Engineer**

**Responsibilities:**

* Responsible for building and developing Data Warehouse system in Horton works Hadoop environment.
* Responsible for managing data coming from different sources like RDBMS, Oracle, Mainframe Systems & Web server logs.
* Handled Data Ingestion: Importing and exporting data into HDFS using SQOOP.
* Load and transform large sets of structured and semi structured data like XML, JSON format.
* Experienced in developing Hive Queries which converts as Map Reduce programs on different data formats like Text file, CSV file, Log files etc.
* Responsible for delivery in the areas of big data engineering with Hadoop, Hive
* Involved in creating Hive tables, and loading and analyzing data using hive queries.
* Involved in loading data from LINUX file system to HDFS.
* Assisted in exporting analyzed data to relational databases using Sqoop.
* Implemented performance improvement techniques in HIVE for efficient data access using HQL.
* Created and maintained Technical documentation for launching HADOOP Clusters and for executing Hive queries.

**Environment**:**:** Hadoop, HDFS, Hive,SCALA, Map Reduce, HBase, Sqoop, LINUX, Cassandra, SQL,HQL.

**Siri Data Services Private Ltd March 2015 - November 2016**

**Senior Developer**

**Responsibilities:**

* Responsible for designing database.(database tables design with relationships)
* Involved in modifying the components using Java&J2EE technologies to implement the Business logic as per the needs of the client.
* Responsible for writing all the required SQL Scripts, Joins, Stored Procedures and triggers.
* Involved in developing the code for use bean, DAO classes and POJO classes for data base interactions.
* Responsible for developing Arrays and Collection API concepts and handling Exceptions.
* Responsible for developing data warehousing design based on OLTP.

**Environment**: - Tomcat, MySQL, Maven, Java SE, JSP, Oracle, MySQL

**Siri Data Services Private Ltd January 2012 - February 2015**

**Developer**

**Responsibilities:**

* Involved in client interaction for gathering requirements, building up a solution ti implement the business logic..
* Involvement in designing and creation of DB objects as per the needs of requirements.
* Involved in configuration of JDBC connection to interact with the database and then writing SQL Scripts in order to perform the DB related operations.
* Developed simple web pages using JSP and handled the requests by using Java and Servlets.
* Responsible for deploying the code in a local machine for end to end testing of the changes made, and then deployed the same into production server.
* Development of Masters, Indexing and Quality check screens.
* Created views with business logics to reduce database complexities for easily Adhoc reporting.
* Involved in all stages of software life cycle in the project.

**Environment**:-Java SE, Jdbc, Servlets, Jsp, Html, Java Script, Tomcat, SQL Server

**Asst.Proffessor/Senior Lecturer/Lecturer \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Responsibilities:**

* Head Of the Department for a period of a year and acted as a Department Overall Coordinator  which involves Academic plan & monitoring, Student Performance monitoring, Staff Recruitment, Motivating students towards Technical activities, Industrial visits and training.
* Project coordinator, guided many final year students projects .
* Handled Subjects: Database Systems, Operating Systems, Computer Architecture, Data Structures etc.,
* Convener for the Department Technical Conferences and Faculty Workshops.
* Member of Project Review Committee & Software Maintenance Committee
* Examiner for  University Examinations

**Publications - INTERNATIONAL CONFERENCE**

* Presented a paper titled “Enterprise Integration Using Open Source Middleware” in an International Conference ISSISS’09 held in GCT, Coimbatore, INDIA
* **Speaker** in a Faculty Development Program titled “Enterprise Integration using open source middleware” held at PSG College of Technology, Coimbatore.

**EDUCATION**

* BACHELOR OF ENGINEERING - Computer Science (73%)
* MASTER OF ENGINEERING – Computer Science (9.3/10 CGPA)