Karthikeyan S Email: [kar9475@gmail.com](mailto:kar9475@gmail.com)

Mobile: +91-9500135567

Experience summary

Software professional with 3 years of experience in the IT industry and extensive experience in analysis & design, coding, development, implementation and maintenance of business technology solutions for various business and customer applications

* Having experience in Hortonworks and in analysis, Design, Development, in Software application which includes Big Data using Hadoop HDFS, Hive, Sqoop, Spark SQL ,PIG, HBase, Phoenix, Kafka, NiFi, Flume, YARN and Map Reduce Program.
* Excellent Understanding on Hadoop Architecture viz HDFS, Namenode, Secondarynode, Datanode, Nodemanager and Resource Manager.
* Working knowledge on Hortonworks on a production cluster (Hortonworks distributions HDP 2.6.2).
* Possess knowledge and having diversified experience in Banking domain.
* Ability to meet deadlines and handle multiple tasks, decisive with strong leadership qualities, flexible in work schedules and possess good communication skills.
* Comprehensive technical, oral, written and communicational skills.

Education background

* Bachelor of Engineering in IT from Velammal Engineering College, Chennai, TN, India.

Technical skill set

**Languages** : Scala, SQL, Shell Scripting, COBOL.

**Databases** : Oracle, DB2

**Data analysis tools** : Tableau Desktop 10 and Tableau server

**Hadoop System** : MR 2.0(YARN), spark, hive, pig, Hadoop, HDFS, sqoop, Kafka, Hbase, Phoenix

**Organization: Cognizant**

**Description: Cognizant** is an American [multinational corporation](https://en.wikipedia.org/wiki/Multinational_corporation" \o "Multinational corporation) that provides digital, technology, consulting, and operations services. It is headquartered in [Teaneck](https://en.wikipedia.org/wiki/Teaneck,_New_Jersey" \o "Teaneck, New Jersey), [New Jersey](https://en.wikipedia.org/wiki/New_Jersey" \o "New Jersey), United States. Cognizant is listed in the [NASDAQ-100](https://en.wikipedia.org/wiki/NASDAQ-100" \o "NASDAQ-100) and the [S&P 500](https://en.wikipedia.org/wiki/S%26P_500" \o "S&P 500) indices.

PROJECT PROFILE

1. **Project Title**: **Data Lake Project**

**Role :** Hadoop developer

**Domain :** Banking Domain

**Duration :**  January 2016 to July 2017

**Tools :** Hadoop, HDFS, HBase, Sqoop, Oozie, MapReduce, Hive, Pig, Phoenix

**BI Tools** **:** Tableau Desktop 10 and Tableau server.

**Description:** Client is a leading American banking organization and data involved in Pension Payments of external clients of Key Bank. Moving forward client decides to go with Hadoop instead of Mainframe and other traditional databases. **As part of migration from mainframe and other databases to Hadoop, I was trained in Hadoop in the period of Jan 2016 to Feb 2016.** Source Data is ‘n’ number of databases. The target is to make all the data to be present in converged platform, so we planned to go with Hadoop platform for efficient way storage and processing with the cost effective way.

**Roles and Responsibilities:**

* Handled importing of data from ‘n’ number of data source viz Teradata, oracle, Mysql, DB2 etc. performed transformations using Hive, loaded data into **HDFS** or **Hbase** based on the requirement and Extracted the data into Hive using **Sqoop**.
* Experienced in developing **Hive Queries** on different data formats like Text file, CSV file, Log files and leveraging time based partitioning yields improvement in performance using HiveQL and running **Pig** scripts (Pig Latin) to study customer behavior.
* Designed **Hive** repository with **external tables, internal tables, buckets, partitions** for data load of parsed data for analytical & operational dashboards
* Involved using Apache Sqoop for efficiently transferring bulk data between HDFS and relational databases for product level forecast.
* Experienced on loading and transforming of large sets of structured, semi structured and unstructured data.
* Involved in achieving the performance tuning of the queries by using partitioning, buckets, Index, Vectorization, Hive on tez, ORC format and cost based optimization techniques.
* We were also analyzing the data of customers to find the Frustration scoring, so that we will work on the business pain points to improve the business.
* Also we will receive the variety of files on adhoc basis from client to convert it into structured format, so that it can be uploaded in the respective websites.
* **Phoenix** to create a table on top of **hbase** table to perform low latency queries.
  1. **Project Title**: **Real Time Streaming and Processing on a vehicle which uploads money in ATM:**

**Role :** Hadoop/Spark developer

**Domain :** Banking

**Duration :**  August 2017 to January 2018

**Tools :** Hadoop, HDFS, HBase, Sqoop, Spark , MapReduce, Hive, Pig, Phoenix, Kafka, NiFi, Oozie, Elasticsearch

**BI Tools** **:** Tableau Desktop 10 and Tableau server.

**Description:** Requirement is to track a vehicle which uploads money in ATM on real time. For doing this we will use vehicle sensors(IOT device), NiFi, kafka, Spark streaming, Elastic search and Kibana. The sensors data will be taken into Kafka through Nifi and from Kafka, Spark will take the data to do the processing by joining the kafka data with HDFS data which is brought by Sqoop into Hadoop platform. The processed data will be put into Elastic search which is a full document search data sore(Best for real time processing), from there kibana will do the real time visualization. Based on Kibana visualization we will get the insights.

**Project Flow:**

**Batch:**

* **Sqoop** for injecting the drivers data from database.
* **HDFS** for persisting the data for primary staging.

**Realtime:**

* Sensors(IOT device) data will passed to kafka through NiFi.
* **NIFI** data flow to read the data from the files created by vehicle sensors and push the data to **Kafka.**
* **Spark** Program to read the data from HDFS location where sqoop imported and read from **Kafka**, create **dataframes**, temporary **views** in spark and join the sqoop data with the kafka data and persist into **Elastic search indices**.
* **Kibana** visualizations and dashboards created for reporting.

**Roles and Responsibilities:**

* Started Working on NiFi, kafka and spark to develop the code.

1. **Project Title**: **BPS(Benefit Payment System)**

**Role :** Program Analyst

**Domain :** Banking

**Duration :**  June 2015 to January 2016

**Tools :** Endevor, SCLM

**Programming**  **:**  Mainframe – COBOL, JCL

**Database :** DB2

**BI Tools** **:** Tableau Desktop 10

**Description:** Client is a leading American banking organization and data involved in Pension Payments of external clients of Key Bank.

**Roles and Responsibilities:**

* Analyzing COBOL codes in case of any issues in production to fix the issue.
* Developed COBOL programs for automating set of works which are repeatedly done by us.
* Creating Dashboards for several teams inside my Account using Tableau.

**Current Organization: Capgemini**

**Description: Capgemini** is a French [multinational](https://en.wikipedia.org/wiki/Multinational_corporation" \o "Multinational corporation) [professional services](https://en.wikipedia.org/wiki/Professional_services" \o "Professional services) and [business consulting](https://en.wikipedia.org/wiki/Business_consulting" \o "Business consulting) [corporation](https://en.wikipedia.org/wiki/Corporation" \o "Corporation) headquartered in [Paris, France](https://en.wikipedia.org/wiki/Paris,_France" \o "Paris, France). It provides IT services and is one of the world's largest [IT consulting](https://en.wikipedia.org/wiki/IT_consulting" \o "IT consulting), [outsourcing](https://en.wikipedia.org/wiki/Outsourcing" \o "Outsourcing) and [professional services](https://en.wikipedia.org/wiki/Professional_services" \o "Professional services) companies.

personal information

Name : Karthikeyan S

Place : Chennai