**WORK EXPERIENCE**

**Hadoop/Big-Data Engineer: AT&T Big Data Sunnyvale, CA** Jan 2017–Jan 2018

* **Built Complete pipeline** for **ETL** with Hadoop eco-system Using **Sqoop, Hive, Pig, Flume, Linux, Kafka, HBase, Cassandra**
* **Designed Kafka** to achieve high through put and low latency transfer of Data from various data sources into HDFS and to other data platforms like spark or hive for further analysis and generating meaningful insights
* Generated **Spark event logs** in local mode to optimize and analyze and fix the memory leak using the Spark history server which increased the **Performance 9 hours to 57 minutes** an **increase of 947% in the efficiency**
* Designed, executed & Optimize SQL query creating meaningful Indexes for faster retrieval of data
* Optimized SQL queries to access required data out of columns instead of selecting the whole tables **for where and Limit clause** thereby increasing the overall efficiency by **350%**
* Used **Pig Scripting** for faster execution using filter , group by methods and sorted complex data Structure and transferred the output back to **HDFS** for transferring Data to **HBASE** and other **NoSQL** databases

**Hadoop/Big-Data Engineer : Hewlett-Packard Enterprise, Houston, Texas** Aug 2016–Jan 2017

* **Designed & Developed End-To-End** solution for Data ingestion from a **Map-R** source system into HDInsight **HDFS** system
* Developed Bash shell scripts for automation & data pulling from a Map-R cluster over WebHdfs URL into Hortonworks **data lake**
* **Flume** for processing real time data from **Twitter** Provided multiple demonstrations of functionalities and efficient data

visualizations approaches using Tableau to the senior management

* Compressed Hive data into Hive ORC managed table with snappy compression for executing queries **5 times faster**
* Experience in running spark SQL queries in order to avoid complex algorithms for querying structured data
* Created Hive – Hbase tables for data storage. Hive for Meta-Store and **Hbase** for data storage in Row Key Format
* Developed **Oozie** workflows to automate the **ETL** process. Falcon cluster to schedule Oozie workflows every **12 hours**

**Big Data Developer** **:** **California State University Continental Tires, India** Sep 2015–May 2016

* Implemented and Design pipeline in Hadoop Using Sqoop, Hive, Pig, Flume, Hbase, Cassandra
* MapReduce program in Java and Fetched data from JDBC using Java and transferred it into HDFS.
* Experience working with Time Series Data Using Cassandra database.
* Implemented Flume for the real time data from Twitter, Sensor generated log files to load directly in Hdfs.

**Java and Database Developer : Continental Tires, India** May 2013–Aug 2014

* Developed, Uploaded clean, documented and re-usable code to GIT for **Continuous Integration** and **Deployment** through **Jenkins**
* Mentored other Junior Application developers on table and query/view(SQL) design and **SQL performance tuning**
* Designed Faster and Optimized execution of the SQL query by avoiding correlated Subqueries when required
* JIRA for assigning tasks & issues and confluence for our knowledge base

**SKILLS & EXPERTISE**

Big Data Ecosystems: Hadoop, Sqoop, Hive, Pig, Flume, Oozie, Kafka, Map-Reduce, Spark SQL, Shell

Java Technologies Java, Enterprise Java Bean(EJB) Java database Connectivity (JDBC) technologies

Big Data Platform: Windows Azure, Cloudera CDH5.X, Hortonworks Sandbox

Relational Database: AWS, Postgres, MySQL

NoSQL Database: MongoDB, HBase, Cassandra

Business Intelligence Tools: Tableau, Google Analytics, MS Excel, Power Query, GIS tools, SAS

Web Services: Restful, SOAP, WSDL

Version Control: Github, Gitlab

**EDUCATION**

M.S. In Information Systems, California State University, Los Angeles Graduated: Mar 2016

B.S. in Electronics Communications. Maharishi Markendeshwar University, India Graduated: July 2013