

Big Data Developer Big Data Developer Big Data Developer Charlotte, NC Over 8 years of professional IT experience which includes experience in Big data ecosystem Excellent Experience in Hadoop architecture and various components such as HDFS Job Tracker Task Tracker NameNode Data Node and MapReduce programming paradigm. Have sound exposure to Retail market including Retail Delivery System. Hands on experience in installing configuring and using Hadoop ecosystem components like Hadoop MapReduce HDFS HBase Hive Sqoop Pig Zookeeper and Flume. Good Exposure on Apache Hadoop Map Reduce programming PIG Scripting and Distribute Application and HDFS. Good Knowledge on Hadoop Cluster architecture and monitoring the cluster. In-depth understanding of Data Structure and Algorithms. Experience in managing and reviewing Hadoop log files. Excellent understanding and knowledge of NOSQL databases like MongoDB HBase Cassandra. Implemented in setting up standards and processes for Hadoop based application design and implementation. Experience in importing and exporting data using Sqoop from HDFS to Relational Database Systems and vice-versa. Experience in Object Oriented Analysis Design OOAD and development of software using UML Methodology good knowledge of J2EE design patterns and Core Java design patterns. Experience in managing Hadoop clusters using Cloudera Manager tool. Very good experience in complete project life cycle design development testing and implementation of Client Server and Web applications. Experience in Administering Installation configuration troubleshooting Security Backup Performance Monitoring and Fine-tuning of Linux Redhat. Extensive experience working in Oracle DB2 SQL Server and My SQL database. Hands on experience in VPN Putty WinSCP VNCviewer etc. Scripting to deploy monitors checks and critical system admin functions automation. Hands on experience in application development using Java RDBMS and Linux shell scripting. Experience in Java JSP Servlets EJB WebLogic WebSphere Hibernate Spring JBoss JDBC RMI Java Script Ajax JQuery XML and HTML Ability to adapt to evolving technology strong sense of responsibility and accomplishment. Authorized to work in the US for any employer Work Experience Big Data Developer MorganStanley/Capgemini - Charlotte, NC March 2018 to July 2019 Launched 5 Node Mapr Cluster 5.2 with 80 Cores, 600G Ram on AWS, to implement a data recommendation engine.

Installed Spark, Hive, Oozie Rstudio, Zeepline on to Mapr Hadoop Cluster. Wrote Python and Scala api from positec's rdbms to aws s3 and implemented on Spark Execution framework. Worked Intensively on Spark, Pyspark and SparkR. Created Spark Data Frames while loading the data in to Hive Tables Used Kafka Cluster to load IOT data on the Hadoop Cluster and designed a data lake for IOT on Kafka Cluster on Azure Created a VPG connection From Client Network to Hadoop cluster on top of aws. Used Sqoop & built the data lake From MSSQL TO Hadoop cluster & scheduled jobs using crontab & oozie Built a data pipeline from FTP Server to Hadoop using python api. Created Hive tables in parquet format & connected to Rstudio and zeepline through Sparkly & JDBC Connection & given Infrastructure access to data science team to run their data models on aws cloud Connected Hive tables on Hadoop cluster to visualization tools like PowerBI & Tableau using ODBC and mapr drill connector. Used s3 as data backup from the hadoop cluster. Implemented POC on Launching Hadoop and Spark Cluster on Hdinsights. Developed Spark for the recommendation engine and validated using python. Environment: AWS, Spark, Python, Azure, Hadoop, Kafka, Scala, Hive, Flume, Hbase, Sqoop, PIG, Java JDK 1.6, Eclipse, MySQL, Ubuntu, Zookeeper, Amazon EC2, SOLR Hadoop Developer Positec/Syntelli Solutions Charlotte, North Carolina March 2016 to February 2018 Created Hive, HBase tables using ORC file format and Snappy compression. Wrote different Pig scripts to clean up the ingested data and created partitions for the daily data. Implemented Partitioning and bucketing for Hive tables based on the requirement. Performed Various tasks on data Ingestion from DB2 to the HDFS using Sqoop JDBC Connector Created shell scripts to parameterize the Pig, Hive actions in Oozie workflow. Designed and implemented Incremental Imports into Hive tables. Created different formats of Hive Tables Finally worked on ORC & Avro and picked ORC as fit requirements. Managed and reviewed the Hadoop log files. Provisioning and managing multi-tenant Cassandra cluster on public cloud environment. Created Hive External & Managed Tables & sorted out the best performance and voted to External Tables. Created and maintained Technical documentation for launching Hadoop Clusters and for executing Pig scripts. Worked on aggregating the data using advanced aspects of Map-Reduce. Environment: Hadoop, MapReduce, HDFS, Hive,

Java, Hadoop distribution of Horton Works Cloudera, Pig, Hbase, Linux, XML, MySQL, Java 6 Eclipse, Oracle 10g, PL/SQL SQL PLUS Hadoop Developer KIOO LIMITED - Hyderabad, Telangana March 2012 to July 2014 Worked on a live 16 nodes Hadoop cluster running CDH4.4 Performed Flume & Sqoop imports of data from Data warehouse platform to HDFS. Continuous monitoring and managing the Hadoop cluster through Cloudera Manager. Implemented Data classification algorithms using Map reduce design patterns. Expertise in designing and deployment of Hadoop cluster and different Big Data analytic tools including Pig, Hive, HBase, Oozie, Sqoop, Flume, Spark, Impala with Cloudera distribution. Used Pig to do data transformations, event joins, filter and some pre-aggregations before storing the data into HDFS. Supported MapReduce Programs those are running on the cluster. Extracted the data from Teradata into HDFS using the Sqoop. Environment: Cloudera, Hadoop, MapReduce, HDFS, Hive, Sqoop, Hbase, Linux, XML, MySQL, Java 6 Eclipse Java & J2EE Developer MUFINDI LIMITED - Hyderabad, Telangana April 2009 to March 2012 Developed web components using JSP, Servlets and JDBC. Designed tables and indexes. Designed, Implemented, Tested and Deployed Enterprise Java Beans both Session and Entity using. WebLogic as Application Serve. Environment: Spring3.0.5, Spring batch 2.1.7, Web Services (SOAP), Apache Axis 2, Tomcat v6.0, Eclipse Indigo 3.7, Ant 1.8.2, Log4j2, CVS, ACORD, SoapUI 4.5.3, DB Visualizer Pro 9.1.2, WebSphere MQ 7.0.1.3, ACORD, Visio 2013, BMC Blade Logic Server Admin console, putty, WinSCP Education Master in Electrical Engineering & Computer Science Texas A & M University 2014 Bachelors of Engineering in Engineering Electronics and Instrumentation Engineering BITS Pilani - Pilani, Rajasthan 2008 Skills Big Data, Python, ETL

Name: Kyle Contreras

Email: patricia49@example.net

Phone: 001-785-798-5418x839