Boyina Devendra Kumar

## bdevendrakumar1992[@gmail.com](mailto:abhisekpandey99@gmail.com)

+91 7416224657

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# Career Objective

Experienced data engineer having 2+ experience with a passion for data-driven decision making. Proven ability to develop and manage big data pipelines, build data models, and perform statistical analysis.

# Current Employment

Employer - Amazon

Designation - Quality Specialist

# Roles & Responsibility

* Experienced with the tools in Hadoop Ecosystem including Hive, HDFS, MapReduce, Sqoop, Spark and basics of Kafka.
* Excellent knowledge on Hadoop ecosystems such as HDFS, Resource Manager, Node Manager, Name Node, Data Node and Map Reduce programming paradigm.
* Having working experience in SQL.
* Hands-on experience with importing and exporting data from Relational databases to HDFS, Hive and HBase using Sqoop.
* Analyzed large amounts of data sets writing Hive queries.
* Having good knowledge on Big Data, HDFS, MapReduce, Hadoop ecosystem (Hive, HBase, Sqoop) and Spark
* Hands on Experience on Major Components of Hadoop Ecosystem like HDFS, HIVE, Sqoop

and YARN.

* Worked on Developing ETL processes using Talend to load data from multiple data sources to HDFS using FLUME and SQOOP, and performed structural modifications using Map Reduce, HIVE.
* Good Knowledge of EC2 instance type and designing and implementing security with VPC, IAM and Security groups, Load Balancer, Internal Gateway and Routing Table etc.
* Experience in EMR, Lambda Functions, AWS Glue, AWS Athena, AWS S3.
* Experience of creating architecture on AWS, Deploying Hadoop clusters using AWS EC2.
* Hands on experience in Server less technologies like AWS GLUE to perform ETL operations, Lambda Functions to trigger the pipeline

# Tools used:

* Hadoop Ecosystems: Hadoop-YARN, HDFS, Hive, Sqoop.
* Programming Language: SQL, Python, HiveQL, Spark,
* Database: MSSQL Server, MySQL, Redshift .
* Cloud Services: AWS
* AWS tools: EMR, Lambda, S3, Athena, Glue, RDS
* ETL Tool: Talend
* Project Tool: JIRA

# 

# Amazon (Internal)

# Title – PARS

# Project Description

Due to the increased sales and online purchases, Amazon started with business initiative to solve the real-world project with the developing data warehouse for better product sales, avoid delayed delivery. The project involves developing Talend ETL jobs to extract data from various sources to transform the data in various way and load data into data lake / data warehouse using Big data and AWS cloud technologies. Big Data was used for huge file processing AWS Redshift was used to store data in data warehouse

* Involved in understanding the business requirements and translate them to technical solutions.
* Worked for preparing design documents and interacted with the data modelers to understand the data model and design.
* Used Scoop to efficiently transfer data between Traditional databases to HDFS .
* Involved in creating Hive Tables , loading and analyzing data using Hive queries.
* Writing Spark RDD transformations and actions for the required input data and performed the data transformation using Spark Core.
* Evaluate the workflow and increase the efficiency of data pipelines that process over 5 TB of data daily
* Increased the efficiency of the data fetching by approximately 30% using query optimization and indexing
* Used Scala to store data to HDFS and to implement Spark for faster processing of data (40% faster)

Senior Associate, May,2017 **–** Dec 2018 WIPRO

* Reviewing and analyzing system specifications
* Executing test scripts and reviewing results.
* Reporting and documenting technical issues.

# Education

B.Tech in Electronics and Telecommunication, Silicon Institute of Technology 2014