

First Name: - Goutham

Last Name: - Bokka

Overall IT Experience: 3(Years / 5 Months)

No of Experience in Requested Job Family: 3 (Years/ 5 Months)

Previous Experience in Ford: - CDS ID (Mandatory): -

Global ID from MSXI / GRI: NA  
(If working / worked in Ford Motor Company through MSXi / GRI)

Candidate Status: market

Market / Internal

Notice Period / Time to join:  
LWD: 27-Jan-2023

Hacker Rank Score: / 300  
(Mandatory for SE Positions)

Willing to Relocate: Yes  
(Yes/No)

Availability for Interview  
(Date / Time) :19-Jan-2023

**Employment History:**

- Worked as GCP Data Engineer in Crocus Global Service Pvt.Ltd, Hyderabad – **August 2019 to till now**

**Skill Set:**

- GCS
- BigQuery
- SQL
- Python Scripting
- AirFlow/Composer
- PySpark
- Hadoop components-HDFS
- Kafka
- Pub/Sub
- DataFlow
- Linux Commands
- Agile -IIRA Scrum Sprint GitHub Confluence

Certifications:

**Project (Assignments):**

- Understanding the client requirement by going through DIS sheet given by Business Analyst. Connecting with source team and IT team and testing team as well for smooth ongoing of tasks.
- Loading the data from source to GCS Bucket.
- Loading from the GCS Bucket to BigQuery and then processing the data in BigQuery and creating authorized views on top of that as per the Entitlement Logic.
- Scheduling all the above tasks in a script using tools like Airflow/Composer.
- Working in Agile methodology by using JIRA, Git Hub, Confluence.
- Used GitHub to manage and deploy the code version. Actively participated in all sprint phases/activities.
- Active participation in training new associates.

Additional Information's (If any): Immediately Available



**Career Objective:**

Looking for opportunity as a Cloud Data Engineer in a high-quality engineering environment where my resourceful experience and skills will add value to organizational operations.

**Experience:**

Aug 2019 - Till Now

**Crocus Global Services Pvt Ltd**

**Cloud Data Engineer**

- 3+ years of experience in IT Industry, Development and Implementation as a Data Engineer.
- Experience in Google Cloud Platform (GCP) in the following components like Big Query, GCS, DataProc, Datalow and Cloud Composer.
- Experience in Hadoop and its components like HDFS, Hive and SPARK.
- Excellent reputation for resolving problems and improving customer satisfaction.
- Understanding client requirements by going through DIS
- Created datasets, tables and authorized views in BigQuery.
- Communicated with DBA team for deployment and used GitHub for code version

**Skills:**

- GCS
- BigQuery
- SQL
- Python Scripting
- AirFlow/Composer
- PySpark
- Hadoop components-HDFS
- Kafka
- Pub/Sub
- DataFlow
- Linux Commands
- Agile -JIRA, Scrum, Sprint, GitHub, Confluence
- File Formats - CSV, JSON

**Education:**

Bachelor of Technology from SR ENGINEERING COLLEGE Affiliated to JNTUH

**Project #2:**

Client : Blibli  
Project : Clicks Data Analytics  
Technologies : GCP, Big Query, GCS, Hadoop, Pub/Sub,  
Airflow.Role : GCP Data Engineer  
POD Size : 8  
Project Duration : Sep 2021 to till Now

**Project Description:**

Blibli.com aims to become a leader in Indonesia's ecommerce sector. Founded in 2010 and launched in 2011, Blibli.com plans to be a one-stop shop for consumers seeking products and services in mobile and tablet devices, sports and outdoors activities, home and living, men's and women's fashion the data platform captures clickstream data as well as data from its backend systems. the business uses the clickstream data to capture intelligence about customer behavior on its website including trending products and services.

**Responsibilities:**

- Understanding the client requirement by going through DIS sheet given by Business Analyst. Connecting with source team and IT team and testing team as well for smooth ongoing of tasks.
- Loading the data from source to GCS Bucket.
- Loading from the **GCS Bucket** to **BigQuery** and then processing the data in **BigQuery** and creating authorized views on top of that as per the Entitlement Logic.
- Scheduling all the above tasks in a script using tools like Airflow/Composer.
- Working in Agile methodology by using **JIRA, Git Hub, Confluence.**
- Used GitHub to manage and deploy the code version. Actively participated in all sprint phases/activities.
- Active participation in training new associates.

### **Project #1:**

Client : Target  
Project : Customer Data Migration Service  
Technologies : GCP, Big Query, GCS, Pub/Sub,  
Airflow.Role : GCP Data Engineer  
Pod Size 8  
Project Duration: Sep 2020 to July 2021

### **Project Description:**

Target is known for its addictive shopping experience, and shoppers often joke about going into a store to make one purchase but end up buying far more and also Target is a general merchandise retailer with stores in all 50 U.S. states and the District of Columbia. It is headquartered in Minneapolis, Minnesota, its hometown since the first Target store opened in 1962 under The Dayton Company.

### **Responsibilities:**

- Creating DDL & DML scripts based on schema given in DIS sheet by Business Analyst.
- Worked on creating DAG's and tasks to deploy and run PySpark jobs in Airflow scheduling tool. Awareness in performance tuning, deployment procedures, troubleshooting and solving database issues. Experience on python language scripting.
- Working Experience on Linux and Windows environments. skills and good command over GCP components. Have created Authorized Views based on customers requirement.
- Responsible to manage data coming from different sources. Experience in handling JSON, CSV Data formats.
- Have an idea about Service Account and Users. Actively participating in daily stand up calls.