**Tech Stack used for the Project**

* Apache Airflow – orchestration tool used for building the pipelines for the data flow, this helps us to monitor the jobs. Airflow has a scheduler which helps us to run the jobs on timely manner.
* Apache Hudi – Hudi is framework which helps us to handle the upserts in easiest way, Hudi have hand in hand integration with spark and flink framework, as we are building most of the jobs using glue and spark.
* Python (Programming Language) – this is the primary programming language used for programming.
* AWS S3 – S3 is storage service
* AWS Glue
* AWS EC2