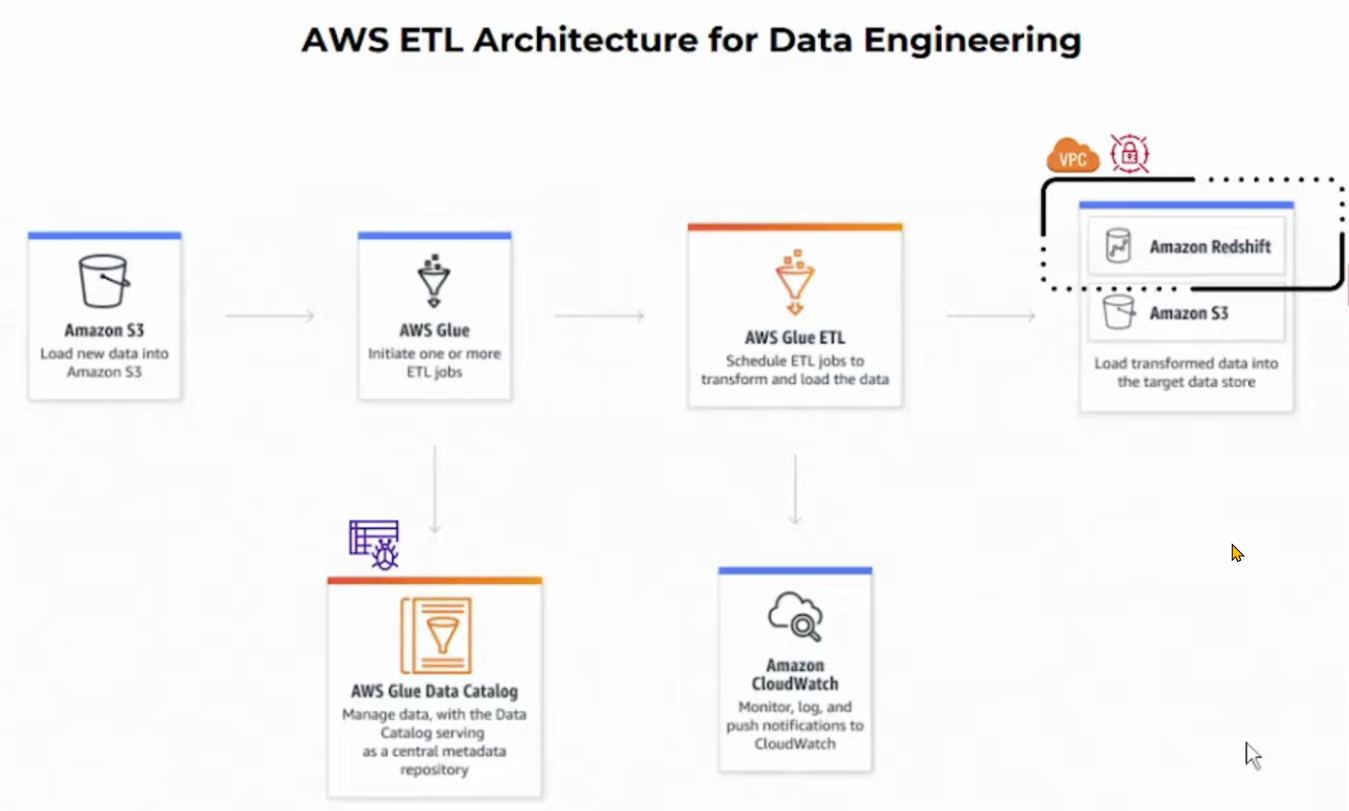
**End to End Data Engineering Poject in AWS using Spark (Pyspark)**



* Create Data Engineering System in AWS using „Infrastructure as Code“
* End to End Data Pipleline
  + S3 Source of Pipeline
  + Build Data Pipeline in Pyspark using Glue Jupyter interactive Notebook
    - Reading data from s3 storage, processing it in Spark and then loading it into Redshift (using dynamic frames and spark data frames

