

- This solution is based on AWS lake formation service.
- AWS lake formation provides features like centralized data catloging, access control, cross account data sharing and security.
- Data in gold layer can be crawled using Glue crawler which creates table metadata in Glue catlog.
- With AWS LF, we can share tables in Glue catlog with multiple cross accounts with simple grant permissions.
- To perform grants we need to register gold layer location in LF service as datalake location.
- Once the tables are shared they should be linked in the consumer account database for querying.

Implementation steps:

- 1. Register gold layer location in s3 as datalake location in LF service.
- 2. Use AWS glue catlog to crawl and update metadata for gold layer tables.
- 3. Enable cross account data sharing by granting the strategy AWS account ID/ IAM role with select/describe access on gold layer database/ tables in LF service.
- 4. The above steps sends an invitation to strategy AWS account to accept data share.
- 5. Accept the invitation in AWS resource manager service.
- 6. The shared database/tables will now be available in strategy team's aws LF service.
- 7. To query the shared data from Athena, we need to link the shared tables in a local schema in Athena.
- 8. Once the link is established, strategy team can query gold layer data in sales account from strategy account Athena service.
- This solution is scalable, meaning in future, if marketing team or any other team request access on gold layer tables, we just need to share gold tables with their account. In a ideal scenario, the data engineering team's account can act as a centralized data catloging account that shares data with all organizational units AWS account.
- There is another solution, without using the lake formation, that is to register the sales AWS account's Glue catalog as a datasource in strategy team Athena service. But it doesn't serve the centralized data governance feature as the lake formation service offers.