

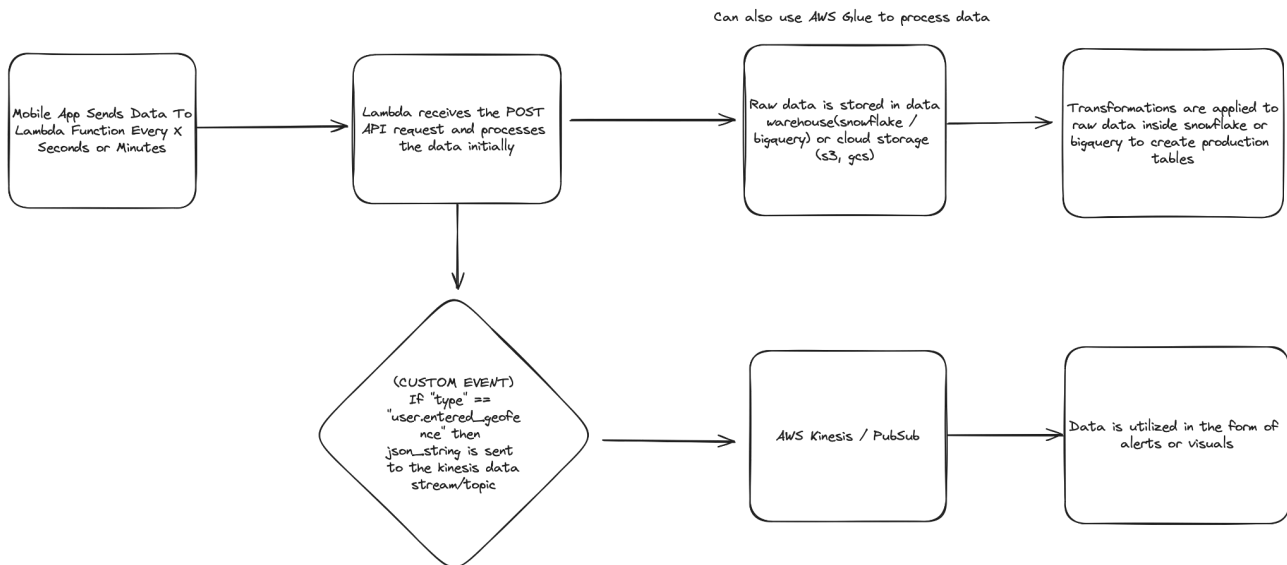
## Milk Moovement Solution Explanation

### Services Used (In Production):

- Milk Moovement Mobile App (Event Emitter/Source)
- AWS Lambda
- Snowflake
- Optional: dbt

### Assumption:

Data that comes from the mobile app is vetted and checked for accuracy inside its own application layer. Our tests are being run in snowflake for accuracy.



### Explanation:

Milk Moovement Mobile App triggers the lambda function every X seconds or minutes in AWS. AWS Lambda / Serverless function receives the POST API call and processes the data initially. If the type of event is "user\_entered\_geofence" which is a custom event, then the data is sent to AWS Kinesis as a JSON string for further use.

Data received through the POST call is then ingested into sqlite/snowflake as a JSON string. This data could also be stored into AWS S3 as a json file that could be picked up by an orchestrator later (if needed). As data is now inside snowflake, dbt can be used to orchestrate transformation jobs to manipulate and clean the data and store it into production tables.

### Further Analysis:

AWS Glue (beam) can be used to perform streaming transformations if needed. If transformations are simpler, then they can also be processed inside the serverless/lambda function itself.

Note: I haven't implemented these as I don't know what to transform for. I stored 2 simple queries in the `/sql` folder to demonstrate some aggregations.

Best option depends on data velocity, and how it's being utilized. If velocity is high, and data is being displayed in a real-time app then processing data either in glue(beam) or inside lambda would make more sense. If velocity is low, and it's being used for analytical purposes then storing it in snowflake and using airflow/dbt to apply transformations would make more sense. The data will eventually end up in the data warehouse in either case.