Assumptions:

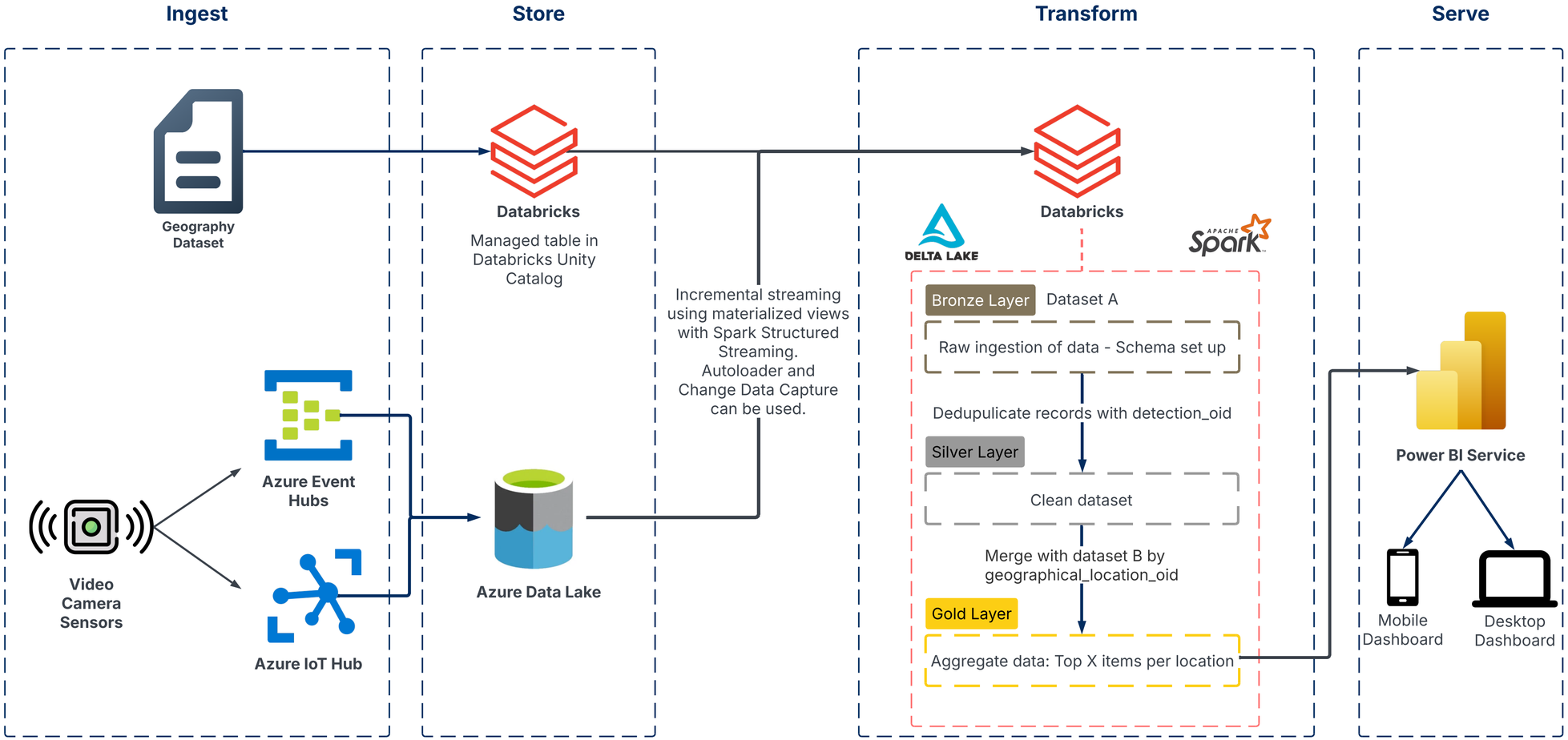
1. Cloud environment: Azure
2. Power BI will be used for the dashboard reports.
3. Periodic report will take place - Allow 1 hour latency periods
4. Dataset B is small and will not have any major changes anytime.
5. detection_oid is a globally unique variable

Questions for Dashboard Users:

1. How often will the dashboard be refreshed?
2. How many users will consume the dashboard at a given time?

Questions for data analysts / business analysts:

1. How large is dataset B?
2. How often will dataset B be updated - can it be kept in memory?
3. How many geographical locations and cameras active?
4. Are the updates of dataset occurring per minute or is there a latent period?
5. Are there patterns in duplication records? Can we run deduplication scripts periodically?



Azure Events Hub and Azure IoT Hub can be used to ingest the data from video camera sensors using Kafka. The data can be batched within these services using message queues. These services will need to be set up on the Azure Resource group.

All incoming data can be stored in ADLS Gen2. ADLS can be managed as an external table in Databricks. The data can be incrementally loaded into Databricks using Autoloader and Change Data Feed within Databricks.

Static dataset B can be stored within Databricks as a Delta Live Table in the Unity Catalog.

All incoming data will be stored as per the medallion architecture where each layer exists as materialized views in Delta Live Tables. Change data feed will ensure only new data will be captured and processed into the bronze layer. This saves compute resources. A Databricks job can be configured to schedule runs based on known data increment periods (i.e. 5 minutes, 10 minutes etc.)

The job pipeline can include ingesting data into the bronze layer, cleaning duplicated records to the silver layer and merging with dataset B and aggregating it in the gold layer. This will be an automated run.

Databricks offers connection to Power BI and this keeps it within the Azure environment as well. From Power BI, the gold layer can be used for the visuals. The dashboard can be served via web or mobile depending on user requirements after publishing it to Power BI Service.