**Flood Monitoring Historical Data Processing**

It works as follows:

1. A script runs every hour and pushes the historical data points received within the last hour to SQS.
2. A second script continuously checks for new historical data points from SQS and processes the same.

N.B.: Here SQS is used instead of directly processing the data points in order to ensure that only one data point is processed at a time. Again, processing one data point at a time is required to ensure data integrity (because it might happen that two points are processed simultaneously and updated the data as per them individually but the changes of both are not reflected finally in the end results).

The scripts will have the following functions:

1. **Script to check for historical data and push to SQS**
   1. Should run every hour at hh:05.
   2. Check for historical data received within the last hour in iot\_device\_raw\_data table
   3. If any historical data points found push the same (complete data packet including debug data) to SQS.
2. **Script to process historical data**
   1. Checks to pull at most 1 data point at a time and checks every 5 mins. for updated data.
   2. Should process only one data point at a time and should fetch the next data point only after the first data point is processed completely.
   3. The data validations and other required conditions applied in case of station raw data processing should also be applied.
   4. The script will:
      1. First update station raw data
      2. Then update corresponding 15 min. avg. data
      3. Then update corresponding 1 hr. avg. data
      4. Then update corresponding 1 day avg. data
   5. For street sensors (E / F) only one raw data point will be inserted (as per the timestamp of the device)
   6. For other sensors, proceed as follows:
      1. Take the next data point of the device
      2. Check if the same is within 15 minutes (because one data point can be maximum repeated upto 15 minutes)
      3. If yes, then pick all data points of the station between raw data and next data point of station time for correction
      4. If no, then pick all the data points between raw data and next 15 min.s time
      5. Replace / update all the raw data points of the station between the picked interval
         1. Check for every minute within the interval
         2. If data point is present for the station, update the values
         3. If data point is not present for the station then insert new ones
      6. For the status change log of the pump stations, delete all the logs between the time interval and insert new ones after calculating against the old data point of the station.