**Azure Event Hub** and **Stream Analytics**

**Project Objective** **-** Implementing real-time fraud detection system for banking transactions based on various business rules. Data Format is as follows :

Transaction ID: 123456789

Timestamp: 2023-06-15 09:30:12

Customer ID: 987654321

Transaction Type: Purchase

Amount: $500.00

Merchant: XYZ Electronics

**Workflow -**

1. **Setting up Azure Event Hub –** add the event hub resource in azure portal and create a new event hub. Setup configurations like no. of partitions and partition key.
2. **Getting real-time data from Logic Apps into Event Hub –** add the logic apps resource in azure portal and apply following :
3. Set up Azure Logic App:
   * Create a new Azure Logic App in the Azure portal.
   * Add a trigger to initiate the generation of transactions. You can use a scheduled trigger or another appropriate trigger based on your requirements.
4. Loop Action:
   * Add a Loop action in your Logic App to generate multiple transactions.
   * Configure the loop to iterate a specific number of times or based on a predefined condition.
5. Compose Action:
   * Within the loop, add a Compose action to create a transaction in the desired format.
   * Use the built-in functions in Logic Apps to generate values for each field, such as Transaction ID, Timestamp, Customer ID, Transaction Type, Amount, and Merchant.
6. Send Data to Event Hubs:
   * Add an action or connector in the Logic App to send the generated transaction data to Azure Event Hubs.
   * Configure the Event Hubs connector with the necessary settings, including the connection string, event hub name, and authentication details.
7. Test and Scale:
   * Test the Logic App to ensure that it generates the desired number of transactions and successfully sends them to Azure Event Hubs.
   * If needed, scale up the Logic App to handle a higher volume of transactions by adjusting the concurrency or using parallel processing techniques.

***NOTE:*** The amount of data generated is dependent on the number of throughput units we define in Event Hub. Say, if we define 1 throughput unit then Logic Apps can generate 1MB or 10000 events per seconds for Event Hub.

1. **Setting up Azure Stream Analytics and SQL DB –** add the stream analytics resource in azure portal and here we can perform SQL-like operations to find the fraudulent transactions. Also define the output bindings for Azure SQL Database in Azure Stream Analytics using following :
2. Open the Azure portal and navigate to your Azure Stream Analytics job.
3. In the job overview page, click on "Outputs" in the left navigation menu.
4. Click on the "Add" button to add a new output.
5. In the "New Output" pane, select "Azure SQL Database" as the sink.
6. Provide the necessary information for the Azure SQL Database output, including the connection details:
   * Provide a name for the output alias, e.g., "NormalSQLTable".
   * Specify the Azure SQL Database server name and database name.
   * Choose the authentication method (SQL authentication or Azure AD authentication).
   * Provide the required credentials for authentication.
   * Select the table or specify a custom query to define the table structure.
7. Click "Create" to add the output binding.
8. Repeat the same steps to add another output binding for the "FraudSQLTable" using the appropriate connection details and table name.

Once you have added the output bindings for Azure SQL Database, you can use the output alias names ("NormalSQLTable" and "FraudSQLTable") in your Stream Analytics query to direct the data into the respective tables.

***NOTE:***Make sure to create the "NormalSQLTable" and "FraudSQLTable" in Azure SQL DB in advance before applying the computations in Stream Analytics that will put the data in sink.

1. **Computations –**

WITH ProcessedTransactions AS (

SELECT

TransactionID,

Timestamp,

CustomerID,

TransactionType,

Amount,

Merchant,

CASE

WHEN ***[z-score\_computation\_here]*** THEN 1

WHEN ***[k-means\_computation\_here]*** THEN 1

WHEN ***[velocity\_computation\_here]***THEN 1

ELSE 0

END AS DetectionResult

FROM

InputEventHub

)

SELECT \*

INTO

NormalSQLTable

FROM

ProcessedTransactions

WHERE

DetectionResult = 0;

SELECT \*

INTO

FraudSQLTable

FROM

ProcessedTransactions

WHERE

DetectionResult = 1;

*NOTE:* There is no internal orchestration in Stream Analytics (like we have in Synapse Analytics), we can choose ‘output to Power BI’ option to send data to Power BI in real-time with Stream Analytics. And in Power BI we can set alerts if fraud transaction is detected.