**Real-Time Data Streaming Analytics for Fraud Detection**

Bootcamp Project - 4

Architecture of **Real-Time Data Streaming Analytics for Fraud Detection:**

A diagram of a data security system

AI-generated content may be incorrect.

The following tells you what tools are required for the project.

### **Tools & Services:**

* **Azure Event Hub** (for ingesting streaming data)
* **Azure Stream Analytics** (for processing and filtering data)
* **Azure Data Lake Storage (ADLS)** (for storing clean data)
* **Azure Cosmos DB** or **Azure SQL Database** (for unique transactions)
* **Azure Synapse Analytics** (for further analysis)

Assuming the data collected from various sources are stored in **raw container** in ADLS Gen 2.

A screenshot of a computer

AI-generated content may be incorrect.

For ingestion of the o/p, we created a new container called **SILVER** in **ADLS** and **Unique Transactions** in **COSMOS DB.**

A screenshot of a computer

AI-generated content may be incorrect.

**Step1: Data ingestion with Azure Event Hub**

* Created an Azure Event Hub Namespace

**Aks-event-004** is the name of the event space.

A screenshot of a computer

AI-generated content may be incorrect.

* Created an Event Hub inside the name space

**Aks-hub-004** is the name of the Event Hub

A white background with black text

AI-generated content may be incorrect.

* Configure Event Hub for Data Ingestion

In the **Event Hub**, a **Shared Access Policy** with **Send** permissions for the data source is created.

A screenshot of a computer

AI-generated content may be incorrect.

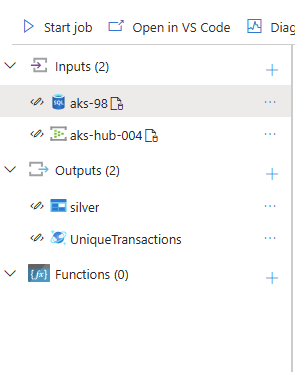
* **Sample iteration-1** file is used as source for Event Hub to send the events to Stream analytics.

A screenshot of a computer

AI-generated content may be incorrect.

**Step 2: Stream Analytics Job for Fraud Detection**

Added **Event hub** (aks-hub-004) **and SQL DB** (aks-98) as Input,



Added **ADLS** and **Cosmos DB** as **Outputs**,

A white screen with text

AI-generated content may be incorrect.

Here we have started the job and sent the event from Event Hub to Stream Analytics and was able to get the o/p as designed.

**o/p:**

**A screen shot of a computer

AI-generated content may be incorrect.**

**A screenshot of a computer

AI-generated content may be incorrect.**

The Code used within the Stream Analytics is as follows:

SELECT

    t.transaction\_id,

    t.user\_id,

    t.timestamp,

    t.amount,

    t.currency,

    t.location,

    t.device,

    t.ip\_address,

    t.merchant\_id,

    t.transaction\_type,

    t.is\_fraud,

    -- Detecting potential fraud reasons

    CASE

        WHEN t.amount > 5000 THEN 'High Amount Fraud'

        WHEN u.location <> t.location THEN 'Unusual Location Fraud'

        ELSE 'Legitimate Transaction'

    END AS fraud\_reason,

    -- Flagging continuous transactions from the same user

    CASE

        WHEN LAG(t.user\_id) OVER (

            PARTITION BY t.user\_id

            LIMIT DURATION(second, 60)  -- Check within 60 seconds window

        ) = t.user\_id

        THEN 'Rapid Transaction Fraud'

        ELSE 'First Transaction'

    END AS rapid\_transaction\_flag

INTO

    [silver]

FROM

    [aks-hub-004] t TIMESTAMP BY t.timestamp

LEFT JOIN

    [aks-98] u

    ON t.user\_id = u.user\_id

WHERE

    t.transaction\_id IS NOT NULL

    AND t.user\_id IS NOT NULL

    AND t.timestamp IS NOT NULL

    AND t.amount IS NOT NULL

    AND t.location IS NOT NULL;

SELECT \*

INTO [UniqueTransactions]  -- Cosmos DB Sink

FROM [aks-hub-004] t TIMESTAMP BY t.timestamp

Now we are going to create synapse for further process of reading the data. A serverless SQL Db is created.

**dbo. Transactions** is created as external table.

A screenshot of a computer

AI-generated content may be incorrect.