A diagram of a computer

Description automatically generated

This diagram represents an incremental data loading process using a Watermark Table to track the last processed value (LPV) for delta processing. Below is a breakdown of each step and component in the ETL workflow:

**1. Watermark Table (Lookup)**

**Purpose:** Maintains metadata for tracking the last processed value (LPV) of each table.

**Columns:**

Id - Unique identifier for each entry in the watermark table.

TableName - The name of the source table being processed.

SchemaName - The database schema name where the table resides.

DeltaCol - The column used to track changes (i.e. ID).

LPV - The last processed value used for incremental data extraction.

**Functionality:**

* A Lookup activity retrieves the last processed value (LPV) from this table.
* This value is used in subsequent steps to determine which new records need to be processed.

**2. Lookup: Fetch Maximum Delta Value from the Source Table**

**Purpose:** Retrieves the latest maximum value of the Delta Column (e.g., MAX(ID) or MAX(LastModifiedDate)) from the source table.

**Functionality:**

* Queries the source system to get the latest delta value (e.g., SELECT MAX(DeltaCol) FROM Table).
* This value represents the most recent change in the data source.

**3. Data Extraction: Copy New/Updated Records**

**Purpose:** Extracts only new or updated records based on the Delta Column and LPV.

**Query Used:**

SELECT \* FROM Table WHERE DeltaCol > LPV;

**Functionality:**

* Uses the retrieved LPV from the Watermark Table to filter only records that have changed since the last run.
* Ensures that only incremental data is copied, reducing processing time and data transfer.

**4. Update Watermark Table**

**Purpose:** Updates the LPV in the Watermark Table after successfully processing new data.

**Process:**

* A Stored Procedure (USP\_WATERMARK) is called to update LPV for the table.
* The LPV is set to the new maximum value from the source lookup step.

**Functionality:**

* Ensures the Watermark Table maintains the latest processed value.
* Prevents reprocessing of the same records in the next ETL run.

**5. Looping Over Multiple Tables (For Each Loop)**

**Purpose:** The entire process is repeated for multiple tables dynamically.

**Functionality:**

* The ETL framework iterates over a list of tables.
* For each table, it performs the lookup, extraction, and update steps.
* Ensures that all necessary tables are processed in an automated manner.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated A screenshot of a computer

Description automatically generated A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generatedforeach activity A screenshot of a computer

Description automatically generated A screenshot of a computer

Description automatically generated A screenshot of a computer

Description automatically generatedlookup2 A screenshot of a computer

Description automatically generated A screenshot of a computer

Description automatically generated A screenshot of a computer

Description automatically generated A screenshot of a computer

Description automatically generated A screenshot of a computer

Description automatically generated A screenshot of a computer

Description automatically generated A screenshot of a computer

Description automatically generated A screenshot of a computer

Description automatically generated A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated A screenshot of a computer

Description automatically generated A screenshot of a computer

Description automatically generated A screenshot of a computer

Description automatically generated A screenshot of a computer

Description automatically generated