**Generic Workflow and DLT**

**Objective**

The generic workflow automates data ingestion by dynamically orchestrating metadata retrieval, ingestion, and processing. It efficiently handles multiple data sources using API-driven triggers and integrates with Delta Live Tables (DLT) for seamless data processing.

**Workflow and DLT Orchestration**

A diagram of a diagram

AI-generated content may be incorrect.

**Orchestration Components**

**1. Trigger Workflow**

There are three ways to trigger the generic workflow as of now:

* Manually, initiated by a user.
* Triggered by another workflow.
* Initiated by an external orchestrator.

**2. Metadata Fetcher**

**Description:** The Metadata fetcher is a component designed to retrieve metadata necessary for the subsequent steps in the workflow.

**Functionality:**

* Fetches all the JSON configuration present at a location based on different data sources.
* Prepares the metadata for use in the Ingest notebook.

**3. Ingest Notebook**

**Description:** The Ingest notebook is responsible for ingesting the data into the system based on the metadata fetched.

**Functionality:**

* Performs Change Data Capture (CDC) operations by fetching source data, transforming it, and comparing it with the target Delta table.
* Validates metadata and ingests data.

**4. Bronze Notebook**

**Description:** The Bronze notebook handles two scenarios based on the presence of a DLT pipeline.

**Functionality:**

* **Start DLT Pipeline**: If the DLT pipeline is already present (same data source running daily), it starts the DLT pipeline.
* **Create DLT Pipeline**: If the DLT pipeline is not present (new data source), it checks if the DLT pipeline exists and creates a new DLT pipeline.