# **Automating Sybase SQL Scripts from Box using Azure Data Factory (ADF)**

ADF is an excellent choice if you prefer a low-code, visual, and fully managed orchestration service. The entire process can be built as a pipeline of activities.

## **High-Level Architecture**

The process involves creating an ADF pipeline that reads files from a Box folder, iterates through each file, and executes its content as a SQL script against a Sybase database. For on-premises databases, a Self-Hosted Integration Runtime (SHIR) is used to bridge the gap between Azure and the private network.

## **Steps to Implement**

### **1. Set up Linked Services**

Linked Services in ADF are like connection strings, defining the connection information needed for Data Factory to connect to external resources.

* **Box Linked Service**:
  + In ADF, create a new **Linked Service** for Box. You will use the credentials from your Box Custom App (Client ID, Client Secret, etc.).
* **Sybase Linked Service**:
  + Install the **Self-Hosted Integration Runtime (SHIR)** on a server that has network access to your Sybase database.
  + Install the appropriate Sybase **ODBC/JDBC driver** on that same SHIR server.
  + In ADF, create a new Linked Service for Sybase, selecting the appropriate driver and pointing it to your configured SHIR.
* **Azure Key Vault Linked Service (Recommended)**:
  + Store your Box and Sybase secrets in an **Azure Key Vault**.
  + Have ADF retrieve them securely via its Managed Identity. This avoids hard-coding credentials.

### **2. Create Datasets**

Datasets represent the data structures within the data stores that you want to use in your activities.

* **Box Dataset**:
  + Create a dataset that points to your Box Linked Service.
  + Specify the folder path containing your SQL scripts.
  + Use a wildcard file path like \*.sql in the file name setting to ensure it only picks up SQL files.

### **3. Build the Pipeline**

The pipeline is the logical grouping of activities that performs the end-to-end task.

* **Get Metadata Activity**:
  1. Use this activity to get a list of all .sql files in your Box folder. Configure it to use your Box dataset and select "Child items" in the field list.
* **ForEach Activity**:
  1. Add a ForEach loop that iterates over the list of files returned by the Get Metadata activity.
  2. Set the "Items" property to: @activity('Get Metadata1').output.childItems.
* **Inside the ForEach Loop**:
  1. **Lookup Activity**:
     + Add a Lookup activity to read the content of the current SQL file from the loop.
     + The file name for the Lookup's dataset should be set dynamically to @item().name.
  2. **Script Activity**:
     + This is the key component for executing the SQL. Add a Script activity connected to your Sybase Linked Service.
     + **Script Type**: Query

**SQL script**: Set this dynamically to insert the output from the Lookup activity. The expression will look like this:  
@activity('Read SQL File Content').output.firstRow.Prop\_0

* + - *(Note: The exact name of the Lookup activity and property may vary slightly based on your configuration.)*

### **4. Schedule the Pipeline**

* Create a **Schedule Trigger** in ADF.
* Configure it to run on your desired schedule (e.g., daily at 2 AM).
* Attach the trigger to your pipeline and publish all changes.