Stephen Xu

July 11, 2023

MIS Automated Download Report

**Task Description**

The task is to automate the process of downloading MIS (Market Information System) files from the ERCOT (Electric Reliability Council of Texas) API. The script should download the files for the current day and extract them into a specified destination folder. The goal is to speed up the process using concurrent operations and thread-lock synchronization.

**Solution Design**

- Location: The solution is implemented as a Python script.

- Steps to Run:

1. Install the required libraries: `requests`, `pandas`.

2. Set up the necessary parameters and variables in the script, such as the destination folder, file paths, API URLs, etc.

3. Execute the script.

- High-level Details:

- The script imports the required libraries and sets up the necessary parameters and variables.

- It defines a function `download\_folder()` that takes a mapping of folder names to report IDs and types, a folder name, and a date as input.

- The `download\_folder()` function queries the ERCOT API for the specified report ID and date, retrieves the ZIP file response, and extracts the relevant files into the destination folder.

- The script reads an Excel sheet to create a mapping of folder names to report IDs and types.

- It creates subfolders in the destination folder for each folder name if they don't already exist.

- Using concurrent.futures and threading, the script submits each folder for processing in parallel, with a maximum number of concurrent operations defined by the `max\_workers` parameter.

- After all the downloads are complete, the script outputs summary statistics, including the execution time.

- It also writes log files for tracking invalid queries and detailed information about the downloaded files.

- The script can be run periodically to download new files for the current day automatically.