Efficiency Redefined: Introducing Our Data Orchestration Solution

Need for such a solution

Apache Airflow is used to schedule and orchestrate data pipelines or workflows.
A DAG (Directed Acyclic Graph) is the core concept of Airflow, collecting tasks together and organizing
them with dependencies and relationships to indicate how they should run. However, manually writing a complex DAG for every job
takes time and effort. Our Data Orchestration framework simplifies Apache Airflow DAG generation, eliminating the need to create
complex DAGs manually. It seamlessly integrates with Azure Data Factory and offers a cost-effective, easy setup.

Designed for developers, it reduces extensive Python coding by allowing workflows to be defined through simple JSON configuration
updates. Leveraging Jinja Templates for automated DAG generation, the framework enhances user experience, especially for
organizations managing workflows across various cloud platforms like Azure. With minimal configuration needs, it's a standout solution
in the market.

Value add for customers

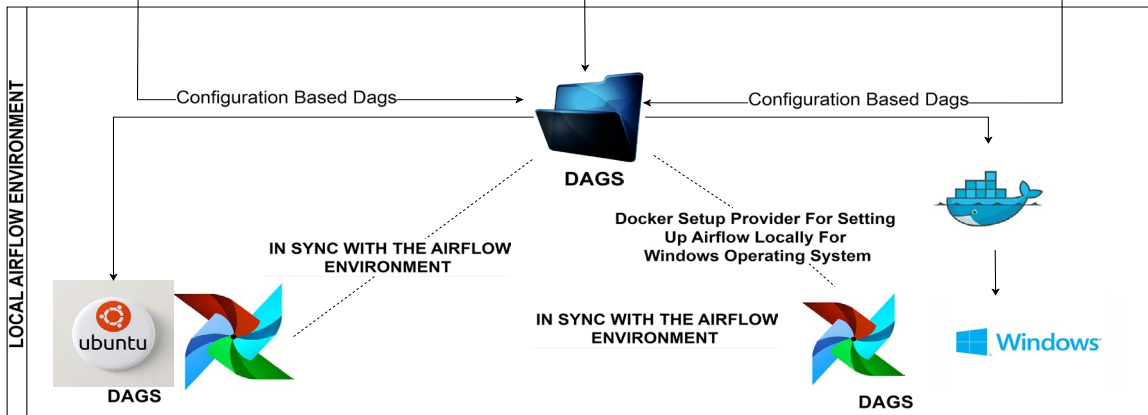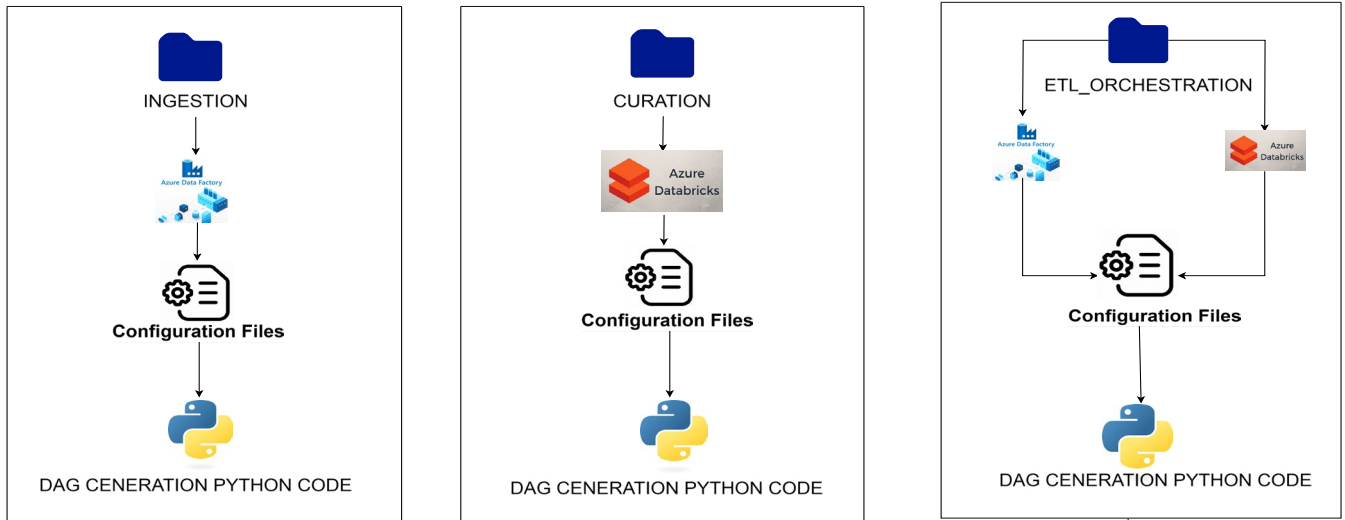| *Automated DAG Creation:*	*User-Friendly Interface:*	*Comprehensive Solution:*
Our framework streamlines DAG creation for Apache Airflow, removing manual complexities and allowing focus on strategic data initiatives.	Using a command-line interface, users can easily select connectors and configure pipelines using intuitive config files, leaving the rest to the framework.	Besides automation, we offer consultation and expertise to identify the best data orchestration tool for various use cases, saving time and effort on research and evaluation.

Overview of the solution

Our Data Orchestration framework modernizes Apache Airflow by automating DAG (Directed Acyclic Graph) generation and offering
cost-effective local setup options. Organizations can efficiently orchestrate data pipelines by seamlessly integrating popular connectors
such as Azure Databricks and Azure Data Factory. This innovative solution requires minimal configuration and provides a user-friendly
experience for developers.

Components of the solution

- **Efficient Data Pipeline Orchestration:** This is focused on the scheduling and management of data pipelines and workflows, distinct from data governance.

- **Streamlined DAG Generation:** Simplifies Apache Airflow Directed Acyclic Graph (DAG) creation, enabling easy setup and cost-effective solutions.

- **Integration Capabilities:** Seamlessly integrates with Azure Databricks and Azure Data Factory for effective data flow management.

- **User-Friendly Interface:** Reduces the need for extensive Python coding, allowing developers to define complex workflows through simple updates to configuration JSON files.

Airflow MetaData Driven Orchestration Framework

INGESTION

Azure Data Factory

Configuration Files

DAG CENERATION PYTHON CODE

CURATION

Azure Databricks

Configuration Files

DAG CENERATION PYTHON CODE

ETL_ORCHESTRATION

Azure Data Factory

Azure Databricks

Configuration Files

DAG CENERATION PYTHON CODE

LOCAL AIRFLOW ENVIRONMENT

Configuration Based Dags → **DAGS** ← Configuration Based Dags

IN SYNC WITH THE AIRFLOW ENVIRONMENT

DAGS (ubuntu)

Docker Setup Provider For Setting Up Airflow Locally For Windows Operating System

IN SYNC WITH THE AIRFLOW ENVIRONMENT

DAGS Windows

Azure Repo

CI

CI PIPELINE PROVIDED FOR GENERATING AND DEPLOYING DAGS TO THE RESPECTIVE AIRFLOW ENVIRONMENT FOR AZURE

DAGS

Apache Airflow

powered by:
Azure Data Factory

Benefits of using this solution

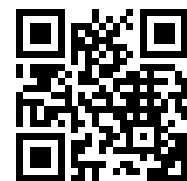| *Accelerated Insights:*	*Streamlined Operations:*	*Seamless Integration:*	*Enhanced Data Integrity:*	*Cost-Efficient Operations:*
Achieve insights 90% faster through automated DAG creation and pipeline execution.	Liberate your team from mundane tasks to concentrate on strategic data endeavours.	Effortlessly connect to diverse connection types for enhanced functionality.	Minimize errors with auto-generated DAGs, improving overall data quality.	Cut costs by leveraging the optional local Airflow setup, avoiding pricier managed Airflow alternatives.

Activities

- Initiate kick-off meetings, gather requirements, and finalize optimal Airflow environment.
- Help set up the finalized Airflow environment configuration.
- Migrate the framework onto the user's system and configure it according to the Airflow environment.
- Assist in a comprehensive review of the architectural blueprint and repository structure.
- Gain insight into the diverse range of connectors and their corresponding configuration file requirements.
- Facilitate the establishment of multiple essential connections from the Airflow environment.
- Conduct thorough testing and deliver the final demo with the necessary configurations.

Deliverables

- Documentation
- Framework Repository
- Airflow Set up Guidance
- Implementation Guidance
- Architectural Blueprint
- Best Practices

Why YASH Technologies

For 24 years, YASH has provided strategic consulting services, maximizing businesses' investments in Microsoft Technology through customized solutions. As a Microsoft Gold Certified Partner with over 700 experts in its Centre of Excellence, YASH ensures that its consultants stay up-to-date with technological advancements through their R&D labs. Specializing in Azure Cloud, Business Applications, Modern Workplace, Data & AI, and Security Solutions, YASH is committed to delivering innovative solutions that drive business success.

YASH Technologies
More than what you think.

Microsoft Solutions Partner

Global Presence: AMERICAS | EUROPE | APAC | MEA

World HQ: 841 Avenue of the Cities East Moline IL-61244 USA

Tel: 309-755-0433 | **Fax:** 309-796-1242 | *www.yash.com*

For more information contact YASH today at
info@yash.com