



Amazon Managed Workflows for Apache Airflow (MWAA): Latest and Greatest!



Aritra Ghosh, Sr. PMT | Jen Dunham, WW GTM

Agenda

01

What is Amazon MWAA?

02

How does AWS contribute to Airflow?

03

MWAA Provisioned: What shipped this year?

04

Serverless: (almost) one year in

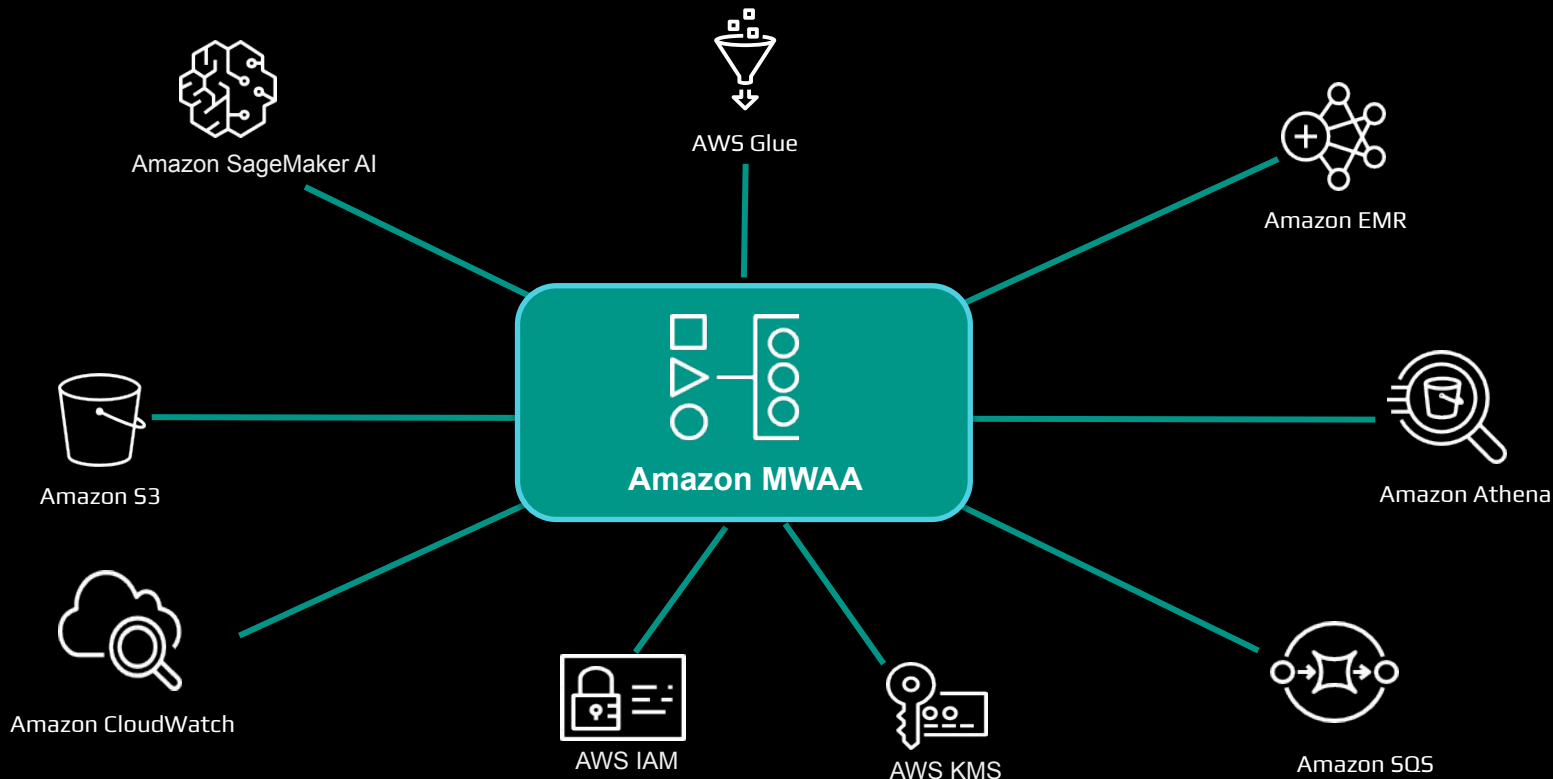
06

Choosing the right deployment for your workload

07

The Future of Workflows!

Amazon MWAA: Seamless Integration with AWS and Beyond



+ 1000s more community operators (Snowflake, dbt, Databricks...)

✓ Fully Managed

✓ Multi AZ

✓ HIPAA/SOC/FedRAMP

✓ Auto-scaling

✓ Hybrid (cloud+on-prem)

✓ 1P+3P Operators

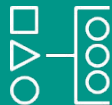
Amazon MWAA: Your Managed Airflow on AWS

Orchestrate complex pipelines seamlessly and reliably



Performance and Scale

Native AWS integrations | Auto scaling | High concurrency DAGs



Amazon MWAA



Lower Total Cost

Prevents overprovisioning | Pay as you go



HA and Enterprise Security

Multi AZ | IAM + KMS | HIPAA, PCI, SOC, FedRAMP



Fully Managed, Open Source

Full Airflow features | Cloud agnostic | On prem

MWAA Used Across a Variety of Use Cases

MWAA orchestrates complex workflows



ETL

Automate ETL and data workflows across AWS and 3P services

- 20,000+ daily jobs at scale
- 70% faster end-to-end runtime
- Cross-system orchestration



DevOps / LLMOps

Automate operational tasks and LLM pipelines

- CI/CD orchestration
- Model deployment automation
- Infrastructure as Code flows

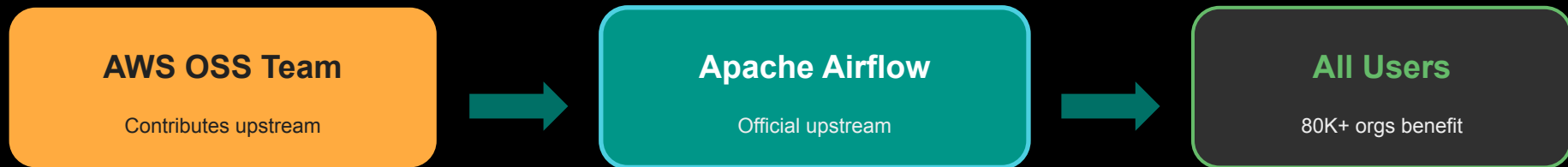


AI/ML and GenAI

Orchestrate end-to-end AI/ML pipelines

- Training + inference pipelines
- Feature engineering
- Model monitoring + retraining

We Invest in Open Source as a Foundational Commitment



Every feature goes into the official OSS release



Airflow Releases



**Bug Fixes +
Improvements**



**PR/AIP Reviews +
Talks**



**MWAA Images on
GitHub**

Active OSS Contributions from Our Team



Multi-Team Support Compatible Auth Managers

- Introduce Teams notion
- UI layer for team resources
- Configuration updates
- KeyCloak AuthManager updates



Deadline Alerts AIP-86 Implementation

- Alerts for data driven workflows
- Alerts for manually triggered runs
- Custom references + callbacks



S3 DAG Bundle Versioning (In progress)

- Version tracked bundles
- Stored natively in S3
- Immutable deploy artifacts

MWAA Provisioned: Enterprise Grade Orchestration

Dedicated, always on Airflow with full control



Always On

Dedicated environments running 24/7



Full Airflow UI

Monitor, debug, iterate visually



Custom Logic

Custom operators, plugins, connections



HA + Scaling

Multi AZ, auto scale workers min to max

What is New in MWAA Provisioned



Airflow 3.3, 3.2, 3.0

Latest major versions available on MWAA



Graceful Environment Updates

Update environments without interrupting tasks



Enhanced Observability with CloudWatch metrics

Updated console with metrics dashboards and recommended alarms



Airflow 2.11 + Python 3.12

Migration path to Airflow 3



Region Expansion

9+ new regions added for global availability

LaunchDarkly scaled to 14,000 tasks per day with Amazon MWAA

Migrating from a vendor-managed Airflow solution to achieve efficiency, isolation, and multi-team orchestration

Challenge

Custom Airflow operators relied on Kubernetes executor behavior from their previous vendor, causing concurrency failures on migration.

Data volume and workload complexity grew exponentially, requiring centralized orchestration across multiple teams.

Onboarding new services and teams took weeks.

Solution

Migrated to Amazon MWAA with Amazon ECS on Fargate for isolated task execution using ECSRunTaskOperator.

Built a reusable Terraform module enabling rapid spin-up of new MWAA environments for additional teams.

Integrated Datadog and CloudWatch for end-to-end monitoring and observability.

Result

Scaled from 100 tasks/day across 1 team to over 14,000 tasks/day across 3 teams and 4 services. Minimal cost increase due to Fargate's pay-per-task model for batch workloads.

Reduced onboarding time for new teams, services, and pipelines from weeks to days.



Looking ahead, we plan to continue iterating on this solution to expand our use of Amazon MWAA to additional AWS services such as AWS Lambda and Amazon Simple Queue Service (Amazon SQS), and further automate our data workflows to support even greater scalability as our company grows.

Engineering Team at LaunchDarkly



Model Distribution

GPT-5.5

Claude Sonnet 4.6

AGENTCONTROL SDK

Python TypeScript

```
# one-time SDK setup
ctx = Context.builder(

def handle_model_call(
    response = tracker
    lambda: openai
```



MWAA Serverless (Almost) Turns 1!



Launched at re:Invent 2025 (November)
Scaling from zero to peak, with zero infrastructure.



MWAA Serverless: Zero Infra, Infinite Scale



Scale to Zero

No cost when idle.
Pay nothing between runs.



Per-Workflow IAM

Granular security.
Each workflow gets its
own IAM role.



SageMaker Unified Studio

Visual interface for
building workflows.



80+ AWS Operators Built In

No pip install needed.
No dependency management.



Pure Pay Per Use

Only task execution time.
No base cost, no minimum.

Serverless: What Shipped in Year 1

Python and Bash Operator Support	Native execution of Python and Bash tasks
Shared VPC Support	Run workflows in shared VPC configurations
EventBridge Notifications	Workflow + task state events to EventBridge
Additional Regions	Expanded global availability
80+ AWS Operators	Native integrations with AWS services

Comparison: OSS vs Provisioned vs Serverless

	OSS	PROVISIONED	SERVERLESS
No Vendor Lock-in	✓	✓	✓
Control Infrastructure	✓	✓	
Vuln Mgmt / Versioning	✓	✓	✓
Bursty Workloads			✓
Airflow UI	✓	✓	*
Community Operators	✓	✓	
Scale to Zero			✓
Dedicated Platform Engineers	✓	✓	
Access to Metadata DB	✓	✓	

Flutter UKI optimizes data pipelines at scale with Amazon MWAA

Transitioning from self-managed Airflow on EC2 to 60,000+ DAG runs daily across managed environments

Challenge

Monolithic, self-managed Airflow on EC2 demanded constant attention, resource scaling, and growing operational overhead.

Complex workloads with 3,500+ dynamically generated DAGs needed isolation and load balancing across environments.

Engineering teams spent time on infrastructure maintenance instead of core business objectives.

Solution

Migrated to multiple Amazon MWAA environments with intelligent DAG distribution and round-robin load balancing across S3 locations.

Delegated task execution to Amazon EKS using Kubernetes Pod Operator with a custom wrapper (KPOw) to simplify DAG authoring and enforce consistency.

Implemented CI/CD with GitLab, Argo Workflows, and dynamic DAG generation outside MWAA with its own pipeline.

Result

Manages ~5,500 DAGs with 30,000+ tasks and over 60,000 DAG runs daily at 450+ concurrent tasks.

Achieved 100% success rate for critical processes in 2025, including 30% load spikes during major sporting events.

Freed engineering teams to focus on DAG development and pipeline efficiency instead of infrastructure management.

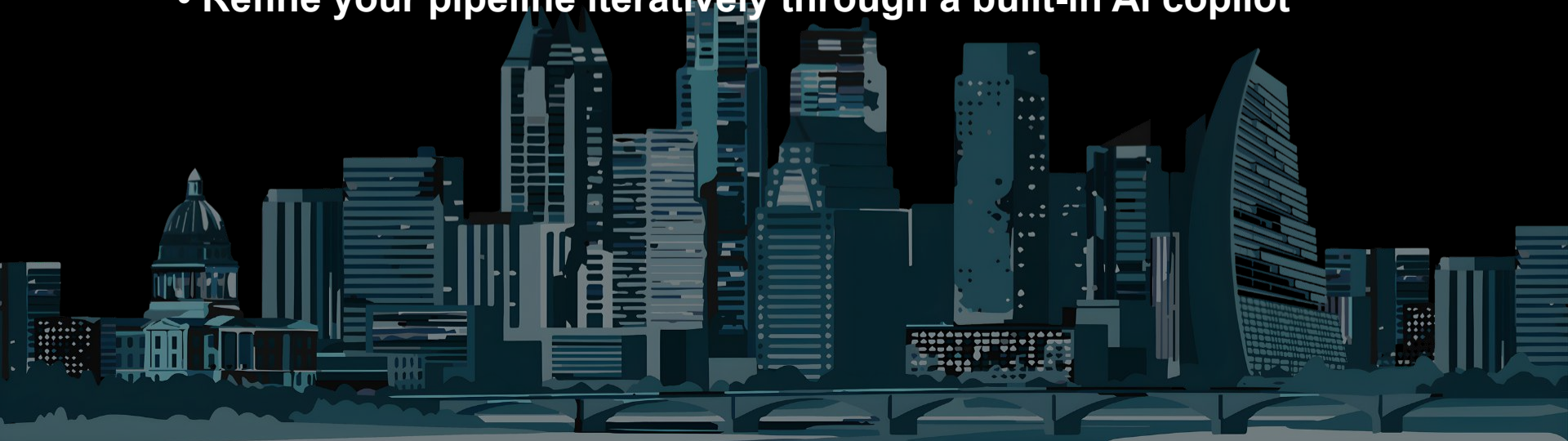




What is Next: Agentic Workflows



- Author workflows in natural language
- Diagnose and auto-fix failures with AI-powered troubleshooting
- Invoke skills natively from your IDE or SageMaker Studio
- Query and transform data without leaving the workflow canvas
- Refine your pipeline iteratively through a built-in AI copilot





Agentic Workflows in Sagemaker Unified Studio

Reach out for access



Kaplan modernized data pipelines with Amazon MWAA

Automating Salesforce-to-Redshift integration with a single DAG as the control center

Challenge

Incremental data loads required manually changing filter dates in Amazon AppFlow flows.

Two-part pipeline (Salesforce → S3, then S3 → Redshift) had no synchronization — no way to know when part one completed before starting part two.

Needed to modernize a legacy integration platform to be agile, adaptive, and straightforward across 250+ objects from three Salesforce instances.

Solution

Built a custom Airflow operator in Amazon MWAA that controls Amazon AppFlow via APIs — creating, updating, starting, and monitoring flows from a single DAG.

Separated ingestion (AppFlow → S3) from warehousing (MWAA COPY → Redshift) for flexibility, archiving, and future transformation needs.

Orchestrated the full end-to-end flow with automated date filter updates, flow status polling, and Redshift loading — zero manual intervention.

Result

Reduced time to create a new pipeline for an existing or new source by 50%.

Cut cost per table by 80–90% by using incremental loads instead of full loads.

Pipelines are now robust, simple to understand, and less error-prone with single-point orchestration from Amazon MWAA.



Check out our other sessions happening today!



Today @ 3:00pm in Texas Ballroom 6
Scaling Airflow for Capacity Forecasting at
Amazon Prime Video

Niko Oliveira



Vincent Beck



Today @ 5:30pm in Texas Ballroom 6
Multi-Team Airflow:
A Customer-Driven Journey



Alex Guglielmino

Today @ 4:30pm in Texas Ballroom 1-2-3
It Works! Now What? Fast Iteration for AI
Capabilities in Airflow

Sneha Rao



Sush Barthakur



Suba Palanisamy



Today @ 5pm in Texas Ballroom 5
Orchestrating GenAI & ML Pipelines
with Apache Airflow

Questions?

Let's Connect on LinkedIn!

Aritra Ghosh

Senior PM, External Services

[linkedin.com/in/aritraghoshpm](https://www.linkedin.com/in/aritraghoshpm)



Jen Dunham

WWSO Data and Analytics GTM

<https://www.linkedin.com/in/jendunham/>





Thank you! 🙏