What’s New With Amazon Managed Workflows for Apache Airflow

John Jackson (he/him)
Principal Product Manager
Agenda

About Amazon MWAA

Architecture Review

Private vs Public Web Servers

Local Runner

Best Practices

Open Source

What’s Next
About Amazon MWAA
What is Amazon MWAA?

Open Source
Same Apache Airflow, no forks

Secure
AWS Identity and Access Management (IAM)

Integrated
AWS Fargate, Amazon CloudWatch, S3

Easy to Deploy
Console, CloudFormation, API, CLI
Key Milestones

November 2020
Amazon Managed Workflows for Apache Airflow is generally available

May 2021
Amazon MWAA adds Airflow 2.0 support

August 2021
Amazon MWAA expands to a total of 15 commercial regions

January 2022
Amazon MWAA adds Airflow 2.2 support with ability to install requirements and plugins on the web server
Amazon MWAA Architecture
Web Server Options
Web Server Options

**Private Network**
For customers that require their Airflow web server to be completely disconnected from the Internet.

**Public Network**
Uses the same IAM authentication as AWS console pages to allow access to the Airflow web server.
Public Web Server Option

- Service VPC (single tenant)
- Load Balancing
- IAM
- Internet
- NAT gateway
- Airflow Web Servers
Private Web Server Option

- Service VPC (single tenant)
- Web Server VPCE
- Load Balancing
- IAM
- Amazon CloudWatch
- Amazon Simple Storage Service (S3)
- Amazon Simple Queue Service
- Amazon Elastic Container Registry
- AWS Key Management Service
- AWS Secrets Manager
MWAA Local Runner
What is MWAA Local Runner?

- Docker Container
- All supported MWAA versions
- Same libraries and options as MWAA
- Requirements management utilities

```
MWAA Local Runner CLI

Syntax: mwaa-local-runner [command]

---commands---
help Print CLI help
build-image Build Image Locally
reset-db Reset local PostgresDB container.
start Start Airflow local environment. (LocalExecutor, Using postgres DB)
test-requirements Install requirements on an ephemeral instance of the container.
package-requirements Download requirements WHL files into plugins folder.
validate-prereqs Validate pre-reqs installed (docker, docker-compose, python3, pip3)
```

https://github.com/aws/aws-mwaa-local-runner
Best Practices
# Things to Watch Out For

<table>
<thead>
<tr>
<th><strong>Logging</strong></th>
<th><strong>Access</strong></th>
<th><strong>Migration</strong></th>
<th><strong>Running Airflow at Scale</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>MWAA updated to Watchtower 2.0.1 to improve performance.</td>
<td>Connectivity to AWS Services (NAT/PrivateLink/Security Groups)</td>
<td>Everything on the Worker, Custom runtimes, Large CPU/Memory</td>
<td>Shameless self-promotion: <a href="https://airflowsummit.org/sessions/2022/managing-apache-airflow-at-scale/">https://airflowsummit.org/sessions/2022/managing-apache-airflow-at-scale/</a></td>
</tr>
<tr>
<td>Can still run into account-wide PutLogEvent throttle</td>
<td>IAM Execution Role access to AWS Services, AOK vs CMK</td>
<td>Ephemeral Workers / Limited Bash Value</td>
<td></td>
</tr>
</tbody>
</table>

© 2023, Amazon Web Services, Inc. or its affiliates.
Open Source
What’s Next?
What the MWAA Team is Working On

**Airflow REST API Access**
AWS has a very high bar for security and scalability of APIs
Working with our MWAA service team and Open Source Team

**New Airflow Versions Faster**
Looking at ways to speed up security reviews and deployments

**Improved Performance and Monitoring**
Incrementally improving performance and stability, and adding tools that will monitor and when possible recover on your behalf
Resources

Stay “In the Know”...

MWAA docs https://docs.aws.amazon.com/mwaa
Amazon MWAA landing page https://aws.amazon.com/mwaa
Amazon MWAA resources page https://aws.amazon.com/managed-workflows-for-apache-airflow/resources/
#airflow-aws Slack Channel: https://apache-airflow.slack.com
Thank you!

John Jackson
jacnjohn@amazon.com
@JohnJacksonPM