Astro\_FAQ GARUNK14 edited this page 20 hours ago · 13 revisions

**Enterprise** 

<> Code

1. Even after removing DAGs, we could able to see it in the UI. Airflow UI for a couple of reasons:

When you remove a DAG from your deployment, it should no longer be active. However, it might still appear in the

1. Airflow scans for DAGs periodically: Airflow scans for new DAGs periodically (every minute by default). If you removed the DAG file very recently, it might still appear in the UI until the next scan. 2. DAG runs history: Even after a DAG is removed, its historical runs will still be visible in the Airflow UI. This is a

feature of Airflow, not a bug. It allows you to see the history of your workflows, even after they've been removed or updated. 3. If you want to completely remove the DAG from the UI, you'll need to delete its records from the Airflow metadata database. Please be careful while doing this, as it involves directly manipulating the Airflow database and can lead

to data loss if not done correctly. (Contact Astronomer support via https://pages.github.ford.com/gcam/astronomer-docs/astronomer/Astronomer-contact/ 4. If the DAG is still appearing in the UI despite waiting for the next scan, it might be due to a caching issue, and you might need to clear the cache. If you're still having issues.

2. Is it possible to change dagbag\_import\_timeout from default value to 160?

Yes, by following this steps you can change the dagbag\_import\_timeout from default value to 160 1. Go to the Astronomer UI and navigate to the deployment you want to add the environment variable to. 2. Click on the Variables tab.

3. Click on the Add Variable button.

4. Enter the key-value pair of the environment variable you want to add. In this case add the environment variable

AIRFLOW CORE DAGBAG IMPORT TIMEOUT and its value.

5. Click on the Add button. 6. After adding the environment variable, you'll need to redeploy your deployment for the changes to take effect. 7. For more information on environment values visit: https://airflow.apache.org/docs/apache-

airflow/stable/configurations-ref.html#dagbag-import-timeout 3. Anyone here have worked on Identifying long running DAGs and send a notification via webex.?

1. Please visit this wiki -> https://pages.github.ford.com/gcam/astronomer-docs/astronomer/airflow-alerting/ for setting up notification.

4. Can anyone assist me in understanding why, when I deploy my DAG changes through this Astro workspace, although

the Tekton pipeline executes as expected and the last updated timestamp refreshes, the corresponding Docker tag does not update as it should? Additionally, why am I unable to locate the deployed DAG in Airflow?

1. Confirm if the Astro build is successful first and check whether the tags value is changed in Astro UI. 2. If build is successful and tag value is not updated in UI, please contact Astro support team using the below link -> https://pages.github.ford.com/gcam/astronomer-docs/astronomer/Astronomer-contact/ Unable to locate the deployed DAGs in Airflow might be because of the following reasons.

1. DAG Parsing Issues: If there are syntax errors in your DAG file, Airflow might not be able to parse it correctly.

Check your Airflow logs for any error messages related to DAG parsing. 2. Incorrect File Location: Make sure your DAG files are in the directory specified by the dags\_folder configuration in your airflow.cfg file.

3. Folder Naming: If you provide the wrong names to the folder, Airflow won't be able to display it.

5. Anyone here have worked on setting up DAG run based on Event based. ? for example a pubsub notification to trigger cloud function --> cloud function will trigger an Airflow DAG or something similar using cloud run.

For event based trigger of Dag runs we can follow below methods:

1. TriggerDagRunOperator: Used when a system-event trigger comes from another DAG within the same Airflow

environment. 2. Sensors: Sensors are useful when you need to coordinate the execution of tasks based on external events or dependencies.

3. Deferrable Operators: An option available to use when sensors, explained above, are ideal but the time of the system event is unknown. Deferrable operators are put in place so you don't have to leave a long-running sensor up all day, or forever, which would increase compute costs.

4. Airflow API: Used when the trigger event is truly random. It's the most reliable and low-cost method of monitoring system events in third-party applications outside of Airflow. 6. I am experiencing a problem with one of my tasks. When I attempt to view the logs, they are taking more than 15

minutes to load. Moreover, once they do load, I only see a small portion of the logs as the rest are truncated.

This issue could be due to a number of factors: 1. If your tasks are writing a large amount of logs, it could be causing delays in fetching and displaying them in the UI. Consider reducing the verbosity of your logging if possible.

amount of log data to display in the UI.

docs/astronomer/astro-localdev/

5. Incorrect function or method calls

failed: failed to execute cmd: exit status 1 "

secret and add in Tekton is as follows:

kubernetes.io/dockerconfigjson

12. Hi team,

can anyone help?

6. Mismatched parentheses, brackets, or braces

2. Syntax errors can occur for a variety of reasons, such as:

2. If you are using remote logging (GCP Bucket, PubSub), there could be network latency or issues with the remote logging service. Ensure your remote logging service is properly configured and functioning correctly. 3. If your Airflow webserver is under heavy load, it could be slowing down the loading of logs. Consider scaling up

your webserver or reducing its load if possible. As for the truncated logs, Airflow by default only fetches a limited

1. If you face any IP exhaustion issues in the dataproc service, please verify whether the VLANs provided in the dataproc repository is correct. 2. For further support please contact the dataproc service team. Link-> https://docs.gcp.ford.com/docs/services/ai---ml---analytics/dataproc/

1. You can follow the below tutorial to set Astronomer Locally -> https://pages.github.ford.com/gcam/astronomer-

9. Hi Team, we are trying to create a DAG to copy a file from one GCS bucket to different folder in same bucket. Getting

error as "Broken Dag: Syntax error" while deploying the script. Could someone please help on this?

4. Improper structure or indentation (especially in languages like Python that rely on indentation)

3. You can find the complete list of airflow component and operators in registry.astronomer.io

1. Please use the following documentation and use the suggested airflow provider operator

7. Could anyone please assist me in resolving an error I'm encountering with a Dataproc task? The error message I'm

exhausted)." How can I address this issue of IP space exhaustion in my GCP project?

8. Is there any document to create Astonomer Local Environment (Astro CLI) with Dbt Core.

2. You can find the provider packages and related operators in registry.astronomer.io

receiving is: "ERROR - Failed to execute job for task dataproc\_preprocessing (409 IP space of 'projects//regions/'

GCSFileTransformOperator to copy a file from one folder to another in the same Google Cloud Storage (GCS) bucket. Source Link -> https://airflow.apache.org/docs/apache-airflow-providersgoogle/stable/\_api/airflow/providers/google/cloud/transfers/gcs\_to\_gcs/index.html

1. Misspelled keywords or identifiers 2. Missing or misplaced punctuation (e.g., commas, semicolons, braces) 3. Incorrect use of operators

10. How to create an image pull secret for Ford Registry and add it to Astro Tekton pipeline? // Hey Guys ... I am having this issue with image registry: failed to fetch anonymous token. "ERROR: failed to solve: registry.ford.com/astro/astro-runtime:6.0.3: failed to authorize: failed to fetch anonymous token:

runtime%3Apull&service=registry.ford.com: 401 Unauthorized Error: command 'docker build -t milo\_09da08/airflow:latest

Check whether the image pull secret is created and the correct secret is used in Tekton. Steps to create an image pull

unexpected status from GET request to https://registry.ford.com/v2/auth?scope=repository%3Aastro%2Fastro-

1. Login to the registry.ford.com, create a robot account in your organisation. 2. If you are not part of any organization, then you need to create an organization.

5. Click on kubernetes secret 6. Click on view secret.yaml 7. Copy the yaml. 8. Below is the example for the pull-secret

9. For creating secrets in the tekton namespace click on workloads 10. Click on secrets 11. Click on create 12. Click on copy from yaml and paste the yaml

i'm getting following error at ftp sensor step, please help me resolve this.

this page. Otherwise, the logs for this task instance may have been removed.]

Could you please help me to resolve the below issue occured during the dag run.

3. Once you are in your organization registry, create robo account.

4. Once robot account is created, click on the account.

13. Click on Create 11. Hi Team,

[\*\*\* Log DUAL\_AUTH\_CDS\_SFTGDIA1\_FORD\_CDS\_AOD\_PACKLIST-ftp-sensor-manual\_\_2024-01-

apiVersion: v1 kind: Secret metadata: name: image-pull data: .dockerconfigjson: XXXXXXX type:

1. Try retriggering the DAGs after ensuring all the environmental variables, connections and dependencies are available to them along with the requisite resources by analysing the metrics tab. 2. If the issue persists, please contact Astronomer support by following the below link -> https://pages.github.ford.com/gcam/astronomer-docs/astronomer/Astronomer-contact/ Note for Astro-support

I ran a dag which should execute series of SQL files but failed with below log, \*\*\* Log IDM\_01\_ACCT\_PROCESSstg-na\_scaid\_guid\_lkup\_tbl\_crt-manual\_\_2024-01-08T05 58.048743+00:00--1-1 not found in Elasticsearch. If your task started recently, please wait a moment and reload this page. Otherwise, the logs for this task instance may have been removed.

1. Try retriggering the DAGs after ensuring all the environmental variables, connections and dependencies are

1. Try retriggering the DAGs after ensuring all the environmental variables, connections and dependencies are

team: Verify all the Airflow Components and StatefulSets are Up and running for Kubernets/Celery Executors.

08T18:35:34.033010+00:00--1-1 not found in Elasticsearch. If your task started recently, please wait a moment and reload

available to them along with the requisite resources by analysing the metrics tab. 2. If the issue persists, please contact Astronomer support by following the below link -> https://pages.github.ford.com/gcam/astronomer-docs/astronomer/Astronomer-contact/ Note for Astro-support team: Verify all the Airflow Components and StatefulSets are Up and running for Kubernets/Celery Executors. 13. Dags are failing or not running without any proper logs available in the ElasticSearch.

https://pages.github.ford.com/gcam/astronomer-docs/astronomer/Astronomer-contact/ Note for Astro-support

team: Verify all the Airflow Components and StatefulSets are Up and running for Kubernets/Celery Executors.

3. Review the workspace owner group, and make sure the role is "workspace admin" 4. Make sure you are part of the FIM group. 5. However if you don't see even a workspace, please raise an it connect ticket w/ the astro team including the above

14. Unable to see workspace / Customer Astro workspace not visible or listed in FCP portal.

1. Log into the environment for which you are facing the issue in Astro platform

2. Open the workspace view and click on the "Teams" tab

16. What are the reasons for dags not able to show in the airflow?

Operator to a DAG, Airflow won't be able to display it.

Here are a few things you can do to troubleshoot this:

facing the issue, contact the astronomer team

2. Increase the resources if it's overused.

2. DAG values: Dag id values should be unique.

follow the same format and namings provided in the document.

reviewed details.

airflow.cfg file.

Service repository is correct.

available to them along with the requisite resources by analysing the metrics tab.

2. If the issue persists, please contact Astronomer support by following the below link ->

15. Hi Team, GWTS dev pipeline failing in dataproc step due to below error, Kindly look into this. IP address exhaustion error. google.api\_core.exceptions.Aborted: 409 IP space of 'projects//regions/' is exhausted. 10: IP space of 'projects//regions/' is exhausted. 1. If you face any IP exhaustion issues in the GCP service, please verify whether the VLANs provided in the GCP

2. For further support please contact the GCP service team. Link-> https://docs.gcp.ford.com/docs/services/

Check your Airflow logs for any error messages related to DAG parsing. 4. Incorrect File Location: Make sure your DAG files are in the directory specified by the dags\_folder configuration in your airflow.cfg file. 5. Scheduler is Not Running: The Airflow scheduler needs to be running for DAGs to show up in the UI.

3. DAG Parsing Issues: If there are syntax errors in your DAG file, Airflow might not be able to parse it correctly.

1. Folder Naming Standard: If you provide the wrong names to the folder, Airflow won't be able to display it. Please

The tag shown in the Airflow UI is a unique identifier generated by a CI tool or a timestamp generated by the Astro CLI on astro deploy. If it is not changing after pushing new changes, it may indicate that the changes were not successfully deployed.

18. Dag is failing at random and not get proper logs for the failed dags. 1. verify the deployment resources utilization under the "metric's" section in the UI to confirm airflow resources components are not over utilized

20. what are the reasons for a deployment to be in unhealthy state? A deployment can be in an unhealthy state due to a variety of reasons:

2. Increase the resources if its overused.

6. If not contact the astronomer team

We didn't find any error in log as well

to start: CreateContainerConfigError

2. Also ensure the secrets are patched with the pipeline.

https://docs.astronomer.io/astro/secrets-backend

27. How can we encrypt password and other data In astronomer variable?

help resolve that?

confirm if the build/pipeline is successful.

pgbouncer etc) are in healthy state. (green)

can cause high latency and result in an unhealthy deployment.

infrastructure, DAGs may remain in a queued state and not run.

deployment UI.

errors, can cause your deployment to be unhealthy. 21. Astronomer deployment in progress since long time. 1. Check the resources usage in "CPU usage" under "Historical" which is present in astronomer deployment "metric".

3. DAGs in queued state: If the number of tasks being scheduled is beyond the capacity of your Airflow

1. Resource constraints: If the resources allocated to your Airflow deployment (CPU, memory, storage) are not

2. Task scheduling latency: If the scheduler does not have enough resources to parse DAGs and schedule tasks, it

sufficient, it can cause your tasks to fail and put your deployment in an unhealthy state.

1. Please wait for 15 minutes for all the pods to be up and deployment to be in healthy state in the Astro UI. 2. In the deployment metric tab section, please check whether the core airflow components (webserver, triggerer, pgbouncer etc) are in healthy state. (green) 3. You can check the live logs of your deployment and its components in the UI for debugging the issue.

4. If you are still facing some issue, please raise a ticket to Astro support team for further investigation and support.

22. One of our deployments is in a constant deployment state even though we havent deployed recently. could someone

1. Try retriggering the DAGs after ensuring all the environmental variables, connections and dependencies are available to them along with the requisite resources by analysing the metrics tab. 2. If the issue persists, please contact Astronomer support by following the below link -> https://pages.github.ford.com/gcam/astronomer-docs/astronomer/Astronomer-contact/

25. Hi Team, I m getting this error "Error: Error waiting to create Service: resource is in failed state "Ready:False",

spec.template.spec.containers[0].env[5].value\_from.secret\_key\_ref.name: Secret

message: Revision 'df-dataingest-cloudrun-astro-trigger-dag-00001-7p9' is not ready and cannot serve traffic.

Deployment GCP Connection Configurations with the appropriate values. 2. Additionally, ensure that the Astronomer Deployment Service account has the necessary permissions for GCP services. 26. Hi team, We are trying to push the code in the dev and the tekton pipeline is failing with the below error: container

1. You can use Airflow Secrets Backend to store sensitive connections and variables in a secure manner using GCP Secret Manager. 2. Refer this doc on how to set up secrets backend.

3. Astronomer uses a Kubernetes Secret to store environment variables added via the Astronomer UI. These variables are encrypted at rest and are automatically mounted to your Airflow pods. 28. Hi Support Team, Latest Changes to Deployment "Dev common" not reflecting. can you please help? // Chnages to deployment not reflected.

resolve it. 29. Hi Team, Our deployment CI pipeline got broke. Any one in this forum can help us get this up. // Pipeline broke or stuck or not working.

1. Check if the pipeline is completed, Astro build is success and Tag is updated.

2. Wait for the deployment to ingest the DAGs / check for DAG import errors

investigation and support. 2. If you are using a custom pipeline, please contact your respective internal teams to resolve the issue. 30. How would i go about getting the logs from an astronomer job that has failed but no longer outputs any logs?

respective team.

Ford Motor Company

available to them along with the requisite resources by analysing the metrics tab. 2. Try refereshing the page 3. If the issue persists, please contact Astronomer support by following the below link -> https://pages.github.ford.com/gcam/astronomer-docs/astronomer/Astronomer-contact/ Note for Astro-support

1. If you are using the standard Astronomer CI pipeline, please raise a ticket with Astro Support team for further investigation and support. 2. If you are using a custom pipeline, please contact your respective internal teams to resolve the issue.

3. Fossa and CheckMarx are custom pipelines, Astro team is not handling these pipelines. Kindly check with the

+ Add a custom footer

▼ Pages 42 Find a page...

★ Edit Pins ▼

មុំ Fork 0

☆ Star 1

Watch 0 ▼

Edit

New page

**▶** Home Airflow Version Update **Astro Deployment Weekend** 

Astro platform post upgrade verification steps Astro Top 30 FAQs

Astro\_FAQ

**Shutdown: Approach and Design** 

Astronomer Platform Architecture **Astronomer Best Practices** 

**Suggested by Astronomer Support Astronomer CI Pipeline Issues and Troubleshooting** 

Astronomer Finops Billing Information

Astronomer Issue Tracker Astronomer Issue trouble shooting

Astronomer Local Dev Setup Astronomer Pipeline List

Show 27 more pages... + Add a custom sidebar

https://github.ford.com/cmsa/As

Astronomer Proxy IDs

Clone this wiki locally

Training

Blog



23. Hello Team, I tried deploying new changes to astro with a new version but looks like the tag is not updated in the deployment. Could you please look into this and help me with the resolution? 1. Confirm if the Astro build is successful first and check whether the tags value is changed in Astro UI.

to fetch, access secrets or tokens. 1. Please verify that the secrets are correctly configured in the GCP Secret Manager and in the Astronomer

projects/712389963333/secrets/astro-token-qa/versions/2 was not found "while configuring the Cloud Run. // Unable

a. https://pages.github.ford.com/gcam/astronomer-docs/astronomer/astro\_external\_secrets\_backend/b.

3. Check if the latest code is reflected in the Code tab of the DAG 4. If you are unable the view the DAG, check if it has a unique name. DAGs should have a unique name across deployment. 5. Try the above steps and if you are still facing issue, please raise a ticket with the Astro support team to help you

1. Try retriggering the DAGs after ensuring all the environmental variables, connections and dependencies are

31. Hi Team, We are facing the Tekton issue. Fossa and CheckMarx failed, for Astro. Even the Pipeline configuration is marked as "enabled": false. Kindly guide us to resolve this issue.

team: Verify all the Airflow Components and StatefulSets are Up and running for Kubernets/Celery Executors.

© 2024 GitHub, Inc.

GitHub Enterprise Server 3.8.8

6. DAG Bag Loading Issues: If there are too many DAGs or if DAG files are too large, it might take some time for Airflow to load all of them. You can check the dagbag\_size and dagbag\_import\_timeout configurations in your 7. Incorrectly Assigned DAG: Every Operator/Task must be assigned to a DAG to run. If you forget to assign an 17. The astronomer deployment Current Tag is not changing after pushing new changes. What might me the cause? 1. Verify that your astro deployment completed successfully without any error logs in the tekton pipeline. If still 3. implement the retry logics in the task to avoid dag failures while execution. 19. Deployment is in unknow state and airflow is showing rocket page. 1. If deployment is in unknown state: Triggering the pipeline can bring the deployment back to online. 2. If deployment is in healthy state and airflow page is showing the rocket/ error page follow the below steps: • verify the resources metrics, if the resources are overused Increase the resources. • if any recent code changes are made, revert back the changes and check the status of the deployment. • if you have set this variable "AIRFLOW\_\_WEBSERVER\_\_WORKERS" remove the variable from astronomer

4. Database issues: Problems with your database, such as memory issues, connectivity problems, or configuration 3. If any recent changes made in code/dockerfile version change/ variables and other configuration items and 4. In the deployment metric tab section, please check whether the core airflow components (webserver, triggerer, 5. You can check the live logs of your deployment and its components in the UI for debugging the issue.

2. If build is successful and tag value is not updated in UI, please contact Astro support team using the below link -> https://pages.github.ford.com/gcam/astronomer-docs/astronomer/Astronomer-contact/ 24. Hi Team, is there any issue with Airflow? Our dag failed in dev with error msg - Exception: Error in BigQuery Procedure: Error encountered during execution. Retrying may solve the problem. Yesterday, we faced similar issue in prod as well.

"step-get-astro-config" in pod "astro-ci-bp-dfd010af5245ca22db52d713df79b10c9df0c058418cd86-pod" is waiting 1. Check whether the secrets are created and configured for workspace, image pull, image push.

1. If you are using the standard Astronomer CI pipeline, please raise a ticket with Astro Support team for further

About