

How Users can figure out if customer has been migrated to Synapse spark cluster in the last runs

Last updated by | Nelson Gomes | Dec 27, 2022 at 1:55 AM PST

Contents

- [Issue](#)
- [Recommendation:](#)
- [Kusto Query:](#)

Issue

How Users can figure out if customer has been migrated to Synapse spark cluster in the last runs

Recommendation:

Customer may face issues running their Dataflows when they are migrated from Databricks spark cluster to synapse spark cluster.

You can get the spark cluster change in specific time window by using kusto query. In the following example, you can see that customer was migrated to Synapse Spark cluster on 2022-12-01.

ClusterType	LastRunDate	LastRunId	LastRunBuild	FirstRunDate	FirstRunId	FirstRunBuild
Databricks	2022-11-30 11:41:00.6920000	f6f9a1b5-54e1-46ef-a4b3-2de2e74e5cea	20220224.12	2022-11-29 05:07:14.4200000	4b92f0eb-b489-4563-ab36-ed4987ef476d	20220224.12
Synapse	2022-12-20 06:31:39.4400000	81fc0ac8-0aea-4cf8-b833-fa959c6cb7f6	20221202.1	2022-12-01 12:35:31.0630000	d4f4127f-adc8-4175-9f6d-fd7fd3a94b13	20221130.6

Kusto Query:

```
let runid='6fb2a74e-8292-4d9f-994d-5e34a06ea561'; // replace with the dataflow activity runId
let LookBackwindowSize=22; // change accordingly to time window you desire to check.
let startDate=datetime_add('Day',-LookBackwindowSize,now());
let Rundetails=cluster('adfneu.kusto.windows.net').database("AzureDataFactory").ActivityRuns| where TIMESTAMP
let MyActivityRuns=Rundetails
|join kind=inner(
cluster('adfneu.kusto.windows.net').database("AzureDataFactory").ActivityRuns
|where TIMESTAMP >=startDate and status == 'Succeeded'|project TIMESTAMP,activityRunId,subscriptionId,dataFacto
on $left.subscriptionId==$right.subscriptionId and $left.dataFactoryName==$right.dataFactoryName and $left.pip
MyActivityRuns|join kind=inner(
cluster('adfneu.kusto.windows.net').database("AzureDataFactory").DataflowClusterLogs|where Caller == "com.micro
|extend ClusterType=case(ClusterType contains "Synapse","Synapse",ClusterType contains "databricks","Databric
|project Timestamp,ActivityRunId,Build,ClusterType,Message
| summarize arg_max(Timestamp,ActivityRunId,Build),arg_min(Timestamp,ActivityRunId,Build) by ClusterType
|project-rename FirstRunDate=Timestamp1,FirstRunId=ActivityRunId1,FirstRunBuild=Build1,LastRunDate=Timestamp,L
```

Additional information

ICM reference:

- **Author:** negome
- **Reviewer:**
- **Keyword:** spark cluster, cluster migration, dataflow on synapse

How good have you found this content?

