

Find Cluster version of an activity run in Monitoring output

Last updated by | Ranjith Katukojwala | Nov 23, 2022 at 1:44 PM PST

Contents

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

Issue

How Users can now figure out which version of spark was used to run the dataflow.

Recommendation:

The version is now present as part of the monitoring output.

The version is currently displayed for all the activities that used Synapse Spark 3.x version. The absence of the version could be deemed to have run on Spark version 2.4.4.

The screenshot displays the Azure Data Factory Monitoring interface. On the left, a sidebar contains navigation links: Dashboards, Runs, Pipeline runs, Trigger runs, Change data capture (previ..., Runtimes & sessions, Integration runtime, Data flow debug, Notifications, and Alerts & metrics. The main area is titled 'All pipeline runs > MapFunctions - Activity runs'. Below this, there's a 'MapFunctions' section with 'List' and 'Gantt' tabs. Action buttons include 'Rerun', 'Rerun from activity', 'Rerun from failed activity', 'Refresh', and 'Update pipeline'. A visual representation of the data flow shows a 'mapFunctions' activity connected to a 'Validate_mapFunctions' stored procedure. An 'Output' window is overlaid, showing a JSON object for 'runStatus' with the following structure:

```
{  "runStatus": {    "sparkVersion": "3.1.3.5.0-73073920",    "computeAcquisitionDuration": 168809,    "version": "20221109.1",    "profile": {      "Dummy": {        "computed": [],        "lineage": []      }    }  }}
```

Below the output window, a table lists the activity runs:

Run start	Duration	Status	Error	Log
Nov 10, 2022, 4:34:01 pm	00:00:05	✓ Succeeded		
Nov 10, 2022, 4:30:32 pm	00:03:28	✓ Succeeded		

However, in the future, PG team could improve on this to provide spark versions irrespective of old or newer spark versions

Kusto Queries :

```
cluster('adfcus.kusto.windows.net').database('AzureDataFactory').DataflowClusterLogs  
| union cluster('adfneu.kusto.windows.net').database('AzureDataFactory').DataflowClusterLogs  
| where ActivityRunId == actId  
| extend SparkConf = iff(Message contains "Initial spark configuration: ", Message, "")  
| extend SparkVersion = extract("/var/lib/(.*)/shs_db", 1, SparkConf)
```

```
cluster('adfneu.kusto.windows.net').database("AzureDataFactory").DataFlowTelemetryEvent  
| union cluster('adfcus.kusto.windows.net').database("AzureDataFactory").DataFlowTelemetryEvent  
| extend ActivityId = extract("runId\\":\"([0-9a-z-]+)\", 1, Event)  
| where ActivityId == actId  
| extend SparkVersion = iff(extract("sparkVersion\\":\"([0-9a-z-\\.]+)\", 1, Event) == "3.1.2", "3.1.2", "2.4.4")
```



How good have you found this content?

