Identify capacity or OOM issue

Last updated by | Jackie Huang | Jan 4, 2022 at 12:24 AM PST

Capacity issue

Symptom

Customer saw errors like the following for his data flow run:

```
"Unexpected failure while waiting for the cluster to be ready.
Cause Unexpected state for cluster (job-57486-run-1):
Could not launch cluster due to cloud provider failures.
azure_error_code: AllocationFailed,
azure_error_message: Allocation failed.
We do not have sufficient capacity for the requested VM size in ..."
```

Root cause and suggestion for customer

1. (most likely) If customer got this issue in low capacity region (North Europe, West Europe...) with instance pool setting (a.k.a. using customized Azure IR in data flow activity with "time to live" property greater than 0), then it happened because the existing pool is still using different sku. Suggestion will be: **ask customer to create a new IR with same setting and refers to it in data flow activity.**
2. If customer got this issue in low capacity region without instance pool setting, please file an ICM ticket to data factory v2 -> data flow channel asap since it's not expected.
3. If customer got this issue in other regions than known low capacity regions, please check the icm query in data factory v2 -> data flow channel to see whether there's a sev2 for that region or not.

Out of memory (OOM) issue

Symptom

Symptom: customer saw errors like and the data flow resources are located in West Europe or North Europe:

```
""errorCode": "4501", "message": "Hit unexpected exception and execution failed.", "failureType": "UserError",
```

Root cause and suggestion for customer

1. Get the support files of data flow from customer.
2. Using the query below (modify the time range/region if necessary) to check whether this activity run id (not pipeline run id) is in the query result or not.
    1. If it does, check whether customer is using low partition in this data flow or not. We can suggest customer to **use larger core count (compute type won't help) in order to avoid OOM issue** if there's no specific setting for low partition. Otherwise please file the ICM ticket for further suggestion from PG team.

2. If it doesn't, file an ICM ticket to PG team since it's not expected behavior.

Query to check OOM:

```
cluster('adfneu.kusto.windows.net').database('AzureDataFactory').ActivityRuns
| where location in ("westeurope", "northeurope")
| where activityType == "ExecuteDataFlow"
| where category == "ActivityRuns"
| where status == "Failed"
| where end >= ago(3d)
| where end <= ago(1h) // in case log delay
| project TIMESTAMP, activityRunId, subscriptionId, dataFactoryName, pipelineName, activityName, location
| join kind=inner (
cluster('azuredmprod.kusto.windows.net').database('AzureDataMovement').CustomLogEvent
| where Level == 3
| where Message contains "Hit unexpected exception and execution failed"
| where TIMESTAMP >= ago(3d)
) on $left.activityRunId == $right.ActivityId
| join kind=leftanti (
ApiOperationEvent
| where operationName contains "PUT SparkJob/OnJobCompleteAsync"
| where env_cloud_location in ("westeurope", "northeurope")
| where env_time >= ago(3d)
| extend activityId = tolower(substring(RequestUri, 60, 96))
| project activityId
) on $left.activityRunId == $right.activityId
```

How good have you found this content?