Pipeline performance lower on Managed VNet IR than Azure IR

Last updated by | Veena Pachauri | Mar 8, 2023 at 10:57 PM PST

Contents

Issue

Simply toggling the Linked Service dropdown in the dataset performs the same pipeline activies but has drastically different run-times. When the dataset is based on the Managed Virtual Network Integrated Runtime it takes over 6 minutes on average to complete while the dataset based on the Default Integrated Runtime takes a total of approximately 20 seconds to complete.

RCA

Checking the slow pipeline is running on Managed VNet IR while the normal one is running on Azure IR. Managed VNet IR taking longer queue time than Azure IR is by design as we are not reserving one compute node per data factory, so there is a warm up time for each copy activity to start as the time mainly happens in VNet join takes extra time compared with Azure IR.

Kusto query to confirm the time:

```
ExecutionApiCall

| where DataFactoryName == 'XXXXX' and PipelineName == 'XXXXX'
| summarize starttime=min(TIMESTAMP), endtime = max(TIMESTAMP) by ActivityId, PipelineJobId
| extend duration = endtime-starttime
| extend duration = endtime-starttime
| project ActivityId, duration, PipelineJobId
| summarize percentiles(duration, 99.99, 99.9, 99, 95)
```

percentile_duration_9...	percentile_duration_99_9	percentile_duration_99	percentile_duration_95
00:04:24.2385996	00:04:24.2385996	00:02:32.4668640	00:02:23.6610908

How good have you found this content?

😊 🙁