

High Memory from Self-IR

Last updated by | Veena Pachauri | Mar 8, 2023 at 11:10 PM PST

Contents

- [Issue](#)
- [Recommendation](#)
- [Owner](#)

Issue

Customer reports that the self hosted IR is consuming resources even when there are no activities running. We saw that, after IR or VM restart it goes back to ~35-40% RAM usage. Then again when ADF pipeline triggers and memory spikes to 80-85% (after some time). The IR worker process takes most of the memory.

But, after pipeline run ends, the memory still stays at that 80-85% usage. All the IR worker process keep memory utilized and no significant memory release happens nor IR worker process closes.

Only remedy is to either restart the VM or restart the IR.

Recommendation

When a IR worker process finished the job it will keep memory is hold and not be released, which is by design as C# Garbage Collection behavior, memory is not released right after job execution and is only released when OS memory is under extreme pressure (like below 10% ~ 15%) or it starts to process another job. Even if there is no running pipeline, we will reserve # active worker processes (diawp.exe) to listen for jobs coming. The # here is equal to the max concurrent jobs count set in Madrid portal.

Customer can follow below TSG to reduce the worker process count to save memory

https://supportability.visualstudio.com/AzureDataFactory/_wiki/wikis/AzureDataFactory/323801/Updating-max-limit-concurrent-jobs-hangs-in-SHIR

Owner

David Zeng Brian Wang