Understand how throughput is calculated for ADF copy activity

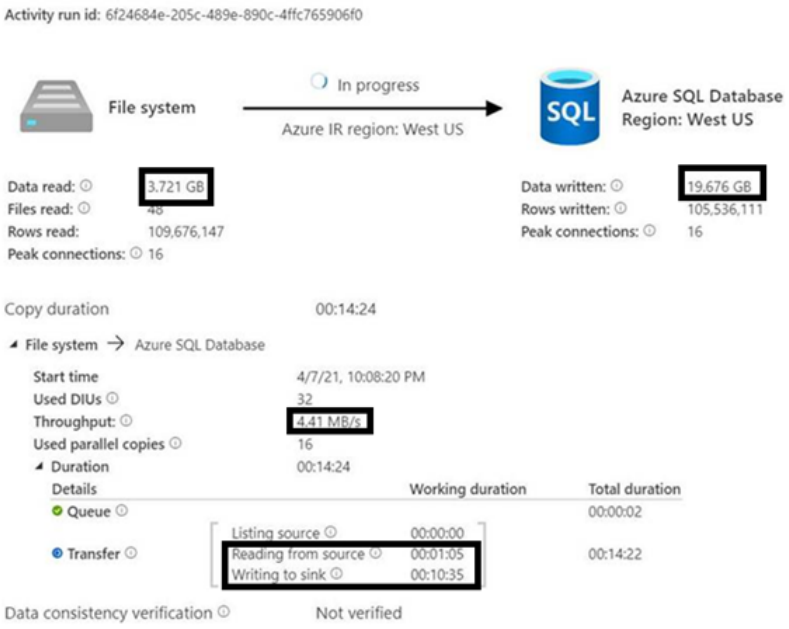
Last updated by | Veena Pachauri | Mar 8, 2023 at 11:57 PM PST

Contents

Scenario

This TSG provides an understanding of how the throughput is calculated for ADF copy activity

Let's take this example of copying data from File system to Azure SQL, to understand how throughput is calculated.



Let us look at the above example and try to understand why customer is getting such a low throughput as 4.41 MBBPS. ADF calculates throughput from source data. In this example data read from the source is 3.721 GB and the total duration for which the job ran is 14:24 min (864 secs). So the total throughput is calculated as –

(Total data read from the source) / (Total time duration)

3810 / 864 = 4.41 MBPS

Though 4.41 MBPS looks very less, but this is not the actual throughput, to understand this further lets check the throughput at source side and at sink side –

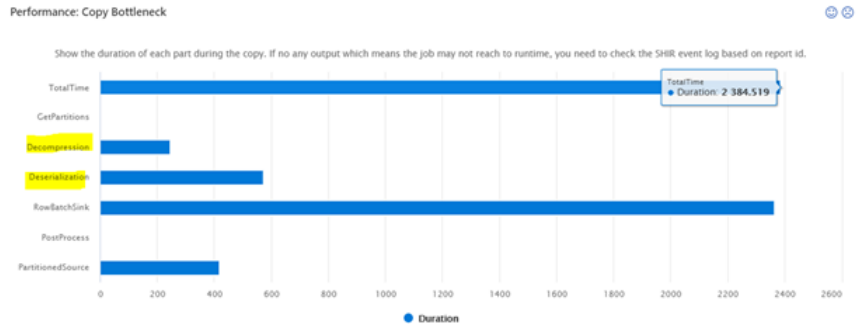
At source-
Data read from source – 3.721 GB / 3810 MB
Time taken in reading source file – 1:05 min / 65 sec
Source side throughput – 58 MBPS

At sink-
Data written at sink – 19.676 GB / 20148 MB
Time taken to write data at sink – 10:35 min / 635 sec
Sink side throughput – 31 MBPS

So in this case even if the ADF shows the throughput at nearly 4 MBPS, but the actual throughput by which data is written at sink is 31 MBPS.

Why we are seeing in difference of Data read and data write is because may the source file is compressed and also when the data is present in form of files then we have to ADF has to deserialize it before inserting into the Azure SQL table.

To validate this point we can check the performance of copy activity from ASC troubleshooter.



Additional Information:

- **Icm References:**
- **Author:** nikhojar
- **Reviewer:** vimals
- **Keywords:**

How good have you found this content?