Oracle Performance

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

Oracle Performance

Thursday, April 26, 2018
11:44

If the throughput is below expectation, the reason could be:

1. ODBC trace log is enabled

The ODBC trace log collects each of the ODBC operation on the machine. Turning on this log will impact all the ODBC related performance on the machine. As our Oracle driver underlying uses ODBC protocol, Oracle performance is impacted.
To turn off the log,

a. Go to “Control Panel” and click “Administrative Tools”
b. Double click “ODBC Data Sources (64-bit)”, and click “Tracing” tab
c. Click “Stop Tracing Now”

2. The data size is not big enough

As we know, there're several one-time actions for each Copy Activity, such as open connection, prepare query, etc., which all take some time. If the data size is not big enough, the overall throughput will be impacted by these one-time actions. **Data size less than 100 MB can be considered as "not big enough".**

3. Source and sink are in different DC

Cross-DC data transfer usually has lower throughput. Please make sure to avoid cross-DC transfer between Oracle -> IR, IR -> sink

4. High workload of Oracle DB

One indication for this case is that the "timeToFirstByte" from "CustomLogEvent" from the Kusto is high (but it's also possible that Oracle DB has a high workload with a low "timeToFirstByte"). The customer's Oracle DBA should have the best knowledge about the workload during the time when the customer's pipeline has poor throughput. Paste some useful links below in case the customer has no DBA.

a. Monitoring General Database State and Workload
b. What is AWR(Automatic workload repository) and How to generate the AWR report
c. 10 Steps to Analyze AWR Report in Oracle

5. Poor network between Oracle -> IR, IR -> sink
6. When copying from Oracle source to Azure Blob Storage/ADLS Gen2, the default value of blockSizeInMB on sink side is 100MB. When the sink peak connection count is observed to be 1 within copy duration, customer can consider reducing the value to achieve higher value of used parallel copies.

Another indication is that, if the poor performance issue cannot be always repro'ed, probably it could be caused by #4 or #5 . Try to let the customer run the pipeline as frequently as possible (e.g. scheduled hourly) for 1 or 2 days, to see whether the issue can be always repro'ed.

If none above seems to be the root cause, use Partition Option to optimize the perf.
Document

How good have you found this content?