Power BI On-premises Data Gateway Jan 2018 Update Implication on TMVGate

Technical Note: TA2019010

Contents

1.0 Overview

In the past, the Power BI On-premises Data Gateway did not have control over the number of concurrent queries that would be executed during the scheduled data refresh. As a result, the data gateway could fire multiple concurrent requests to the data sources defined.

As TMVGate is multithreaded, it would handle these calls as concurrent requests and process them in parallel.

For data sets that are relatively large (e.g. in millions of data cells), this may cause a server resource constraint (specifically for memory usage) if TMVGate is running on the same server as TM1/PA instance.

This was highlighted in the following Technical Documents:

- Power BI Parallel Data Loading Option with TMVGate.pdf
- Optimizing TMVGate Performance with Cube View.pdf

This, however, will change with the Power BI On-premises Data Gateway update in Jan 2018.

This technical document describes a new option in which customers can use to control the number of concurrent queries in the Power BI On-premises Data Gateway, and manage the resources for running TMVGate.

Note that the behaviour is not unique to TMVGate but apply in general to all data sources supported by Power BI.

2.0 Power BI On-premises Data Gateway Mashup Engine Container in Jan 2018 update

The following text was extracted from the Power BI Jan 2018 update blog:

As you probably know, mashup engines are the components that run your queries, extract the data and load it into the data model. They are part of the on-premise gateway structure, and they are the same components that run in the Power BI desktop as well as the Power BI service.

On-premise gateway admins can now control the number of containers running simultaneously on the gateway machine. The more containers there are running, the more queries will run concurrently through the gateway.

Here are the steps to change this configuration:

- On the gateway machine, open the file: [Program Files]\On-premises data gateway\Microsoft.PowerBI.DataMovement.Pipeline.GatewayCore.dll.config
- Find the "MashupDefaultPoolContainerMaxCount" setting.
- Change its value to the desired value.
- Save and restart the gateway.

You will notice that in that file, the default value of this setting is '0'. This means that it will use the default out-of-the-box mashup engine setting of '6' containers. You can override this default by setting your own value.

Keep in mind that the mashup engine containers require machine resources and will affect your machine's performance. You will need to strike a good balance between the efficiency you get by running concurrent queries and optimizing the local machine resources those containers use. A recommended ceiling for the maximum number of containers you can use is roughly twice the number of cores in your local machine processor.

Changing this property value to “1” means that all queries will run sequentially, which could be desired for specific data sources that do not support or do not perform well with concurrent requests.

3.0 Summary

In summary, if you are already running TMVGate on a separate server, you may not have any issues where TMVGate competes with TM1/PA for server resources. Otherwise, we advise you to review the above options to optimize the deployment.

Also note that the Power BI options discussed in this document are applicable to all other data sources and not just TMVGate.

4.0 References

The following are the reference article for the Power BI On-premises Data Gateway Jan 2018 Update.

- https://powerbi.microsoft.com/en-us/blog/on-premises-data-gateway-january-update-is-now-available/