Issues with poor Script Performance

Last updated by | Vitor Tomaz | Nov 22, 2021 at 10:30 AM PST

Contents

Issue

Investigating performance issues for script run taking a very long time

Investigation/Analysis

Check to understand if we are running into Resource Bottlenecks

Data to collect:

- **Q: What is the R script being used with sp_execute_external_script?** Have customer send you the entire command they are running.
- **Q: When did the R query start performing poorly? How long does it take to execute the query**?
- **Q: Did it ever perform well, and if yes, how long did it take to execute?**
- **Q: How long does the script take to run outside of SQL Server?**
- **Q: Does the R Script work outside of SQL Server using something like R Studio?**

Identify the bottle neck for the customer queries using the below script while the script is running

```
SET NOCOUNT ON;

DROP TABLE IF EXISTS UsageStats
GO

DECLARE @startTime datetime2(0) = GETDATE();
DECLARE @secondsToCapture INT = 600;  ---- 10 minutes capture window

CREATE TABLE UsageStats (
        timestamp DATETIME2,
        package_name NVARCHAR(100),
        memory_usage FLOAT,
        cpu_usage FLOAT
)

WHILE (GETDATE() < DATEADD(SECOND, @secondsToCapture, @startTime))
BEGIN
        INSERT INTO UsageStats (timestamp, package_name, memory_usage, cpu_usage)
        SELECT CURRENT_TIMESTAMP as timestamp, dmv.* FROM sys.dm_external_script_resource_usage_stats as dmv
        WAITFOR DELAY '00:00:00.500';
END

SELECT * FROM UsageStats
```

Identify bottleneck for the customer queries

```
MonExtensibilityTrace
| extend phase = case(
    trace_message contains "< SQLSatellite_LaunchSatellite", "LaunchSatellite",
    trace_message contains "<= SatelliteRun()", "SatelliteRun",
    trace_message contains "<= ReadDataset()", "ReadDataSet",
    trace_message contains "<= ExecuteScript()", "ExecuteScript",
    trace_message contains "<= WriteDataSet()", "WriteDataSet",
    "")
| where SubscriptionId == 'Input SubscriptionId Here'
| where isnotempty(phase)
| summarize TIMESTAMP=max(TIMESTAMP) by phase, session_id
| order by TIMESTAMP asc
| evaluate pivot(phase, sum(TIMESTAMP), session_id)
| project session_id,
    Launch=SatelliteRun-LaunchSatellite,
    Input=ReadDataSet-SatelliteRun,
    Execute=ExecuteScript-ReadDataSet,
    Output=WriteDataSet-ExecuteScript
```

If the time taken is outside of customer script execution/data. Pull the events in the phase to drill down further.

TSG-To-be-updated further with few more details

Mitigation

If we are running into resource bottleneck , the customer may have to scale your database to a higher service tier to resolve this.

Note
After Each SLO update you will need the Engineering team to re-enable the dedicated resource limits for R/Python. The Customer will not see a downtime for using ML services after an update, but the 80/20 dedcated resource limits will be set incorrectly and overcomitted , which may lead to OOM errors) In case of a Service

Level Objective (SLO) update, please update the SLO and raise a support ticket to re-enable the dedicated resource limits for R/Python.

The Engineering team will Run RgIsolationSettings CAS command to update the limits to 80 (SQL)- 20 (R/Python external processes). If the customer needs more for Python/R, the Team can set that up as well (Generally not recommended , but it is possible)

If this is not a Resource specific issue , ensure that we collect the above information mentioned in the Investigation and & Analysis. Possibly check if the customer has the same issue runing the External Script with ML Services for On premise SQL Server. Open an ICM for with the Product Engineering Team on this for further investigation

Queue Name Extensibility (ML Services & Language Extensions on SQL Azure)

Public Doc Reference (Preview Limitations)

https://docs.microsoft.com/en-us//azure/azure-sql/managed-instance/machine-learning-services-differences?view=sql-server-ver15#preview-limitations

How good have you found this content?