Sudden increase in data used space for the user database

Last updated by | Vitor Tomaz | Feb 24, 2023 at 3:27 AM PST

Contents

Issue

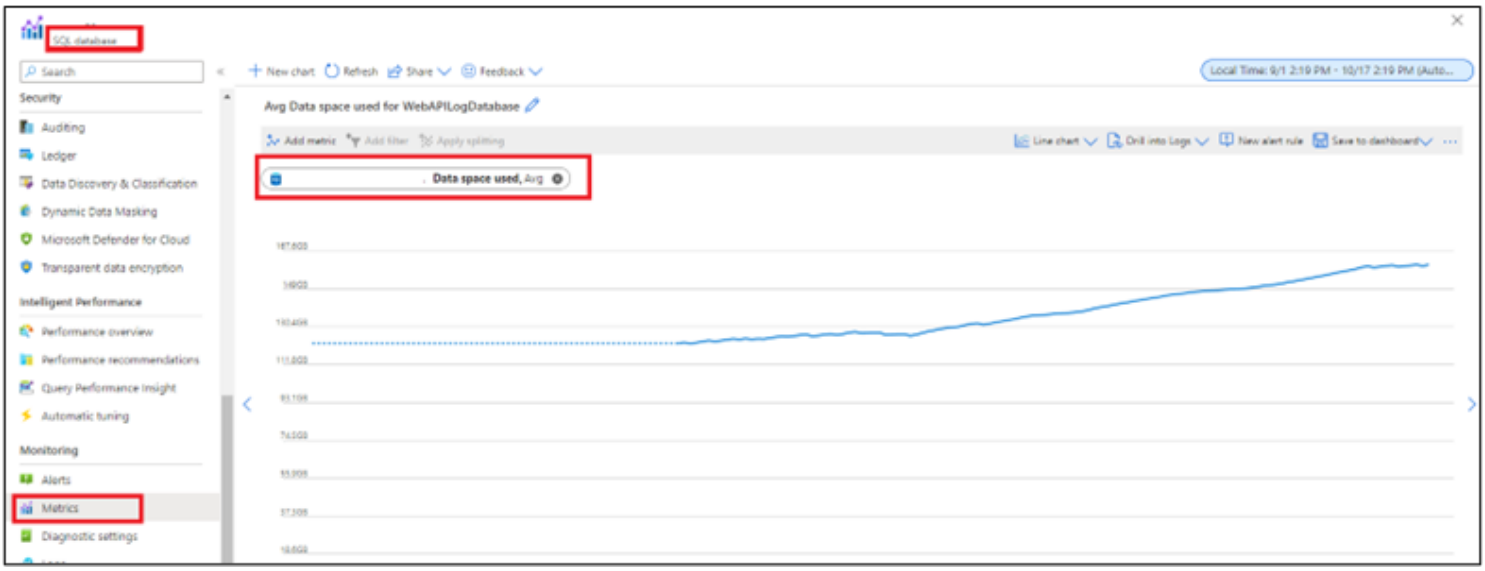
A customer has noticed a sudden increase in **data used space** and have requested help to identify the cause as they were not able to reclaim the space using DBCC SHRINKFILE.

Investigation/Analysis

In these scenarios, the increase could be due to an actual increase in the data size or can be related to other reasons such as fragmented indexes.

Below steps, will help you to troubleshoot and check if the size was increased due to actual data increase to confirm if the cause was related to customer workload.

On customer side, you can check the Metrics (Azure portal -> Azure SQL database -> Metrics blade -> filter by data space used) as below:



You can check MDF and LDF files size from customer side you can use the below T-SQL:

```
select * from sys.database_files
```

For more information: sys.database_files (Transact-SQL) - SQL Server | Microsoft Learn

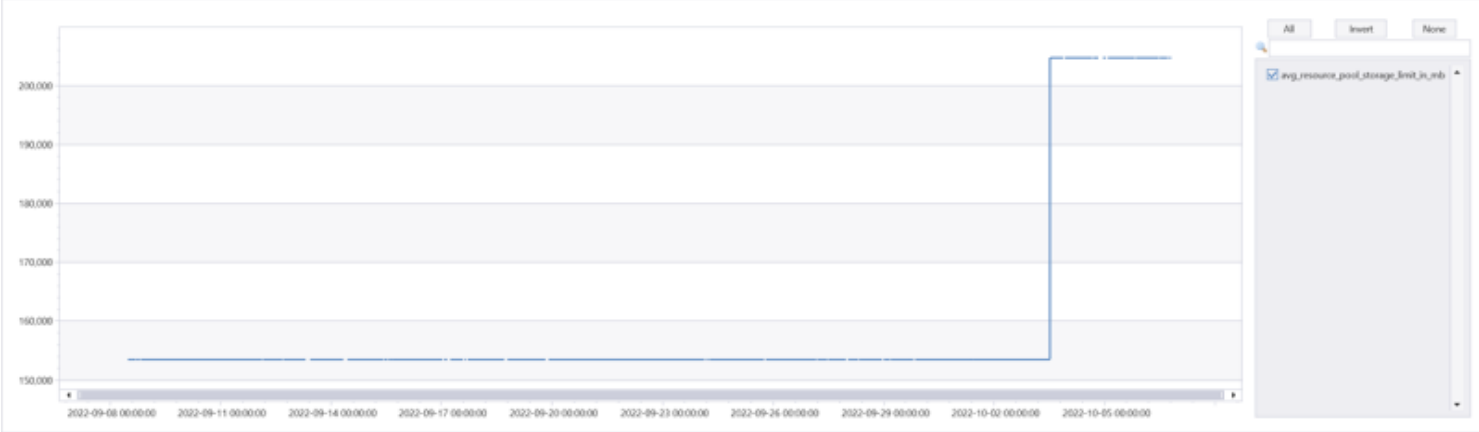
In addition, you can use the below T-SQL to get more information about the database space allocated and compare it with the unused allocated space:

```
SELECT SUM(size/128.0) AS DatabaseDataSpaceAllocatedInMB,
SUM(size/128.0 - CAST(FILEPROPERTY(name, 'SpaceUsed') AS int)/128.0) AS DatabaseDataSpaceAllocatedUnusedInMB
FROM sys.database_files
GROUP BY type_desc
HAVING type_desc = 'ROWS'
```

On CSS side, we can use the below Kusto queries to identify the database size as well as the increase in the row count which will give us an indication about the cause for the increase.

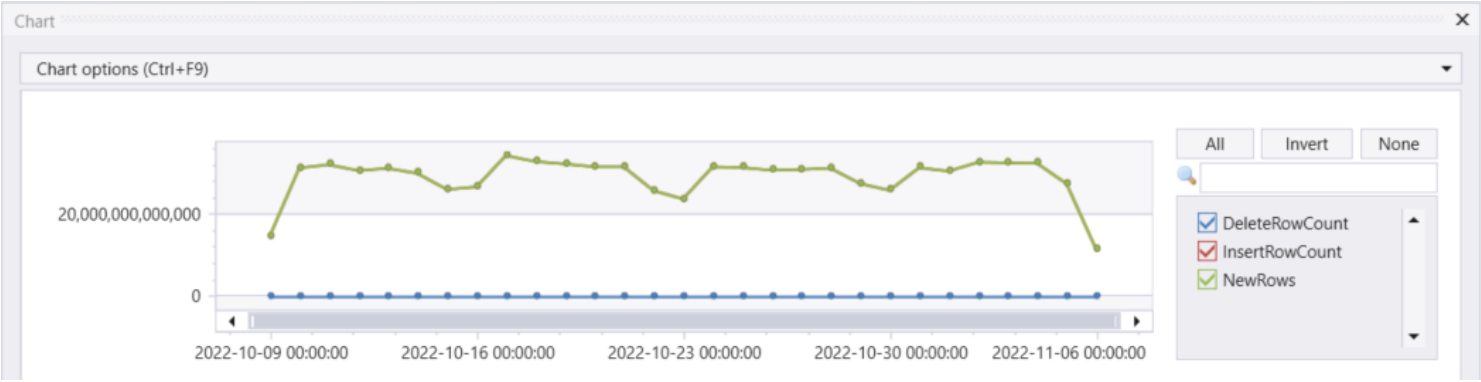
In case the database was in an elastic pool, you can use the below Kusto query to check the pool storage limit in MB:

```
MonAnalyticsRPSnapshot
| where name == "ElasticPoolName"
| summarize avg(resource_pool_storage_limit_in_mb) by bin(TIMESTAMP, 10m)
| render timechart
```



You can use the below query to check the number and increase in new rows, inserted as well as deleted rows (this can give you an indication if the increase in the database size was related to the data increase):

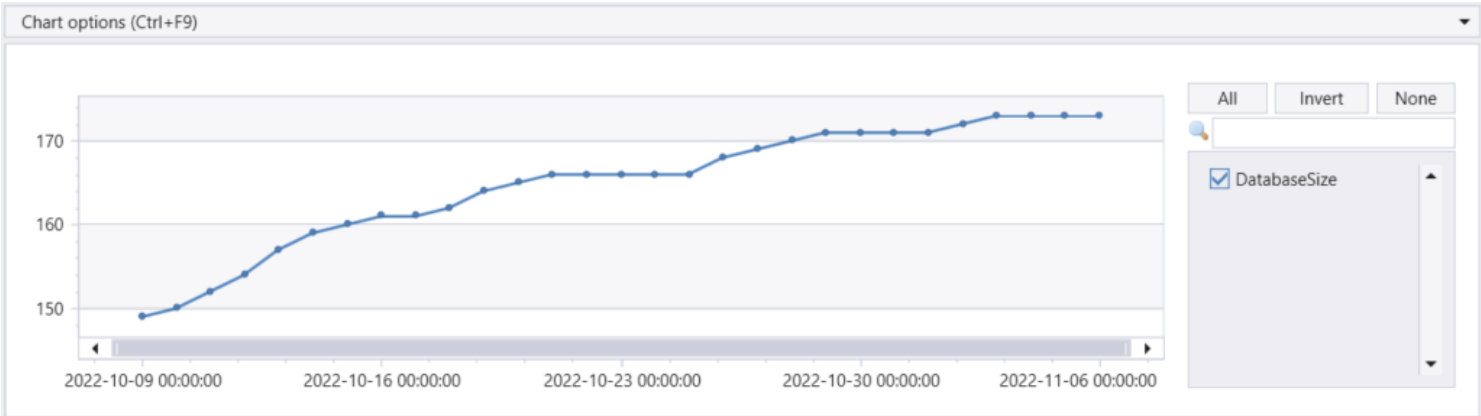
```
let InsertRecordData = MonWiQdsExecStats
| where LogicalServerName == "prodsql50" and statement_type == "x_estypInsert"
| summarize InsertRowCount = sum(rowcount) by bin(originalEventTimestamp,1d);
let DeleteRecordData = MonWiQdsExecStats
| where LogicalServerName == "prodsql50" and statement_type == "x_estypDelete"
| summarize DeleteRowCount = sum(rowcount) by bin(originalEventTimestamp,1d);
InsertRecordData
| join kind=inner DeleteRecordData
on $left.originalEventTimestamp == $right.originalEventTimestamp
|project originalEventTimestamp, InsertRowCount, DeleteRowCount, NewRows = (InsertRowCount-DeleteRowCount)
| render timechart
```



originalEventTimestamp	InsertRowCount	DeleteRowCount	NewRows
2022-10-09 00:00:00.0000000	14665244089925	85930	14665244003995
2022-10-10 00:00:00.0000000	31317090017605	390290	31317089627315
2022-10-11 00:00:00.0000000	32108606004389	571030	32108605433359
2022-10-12 00:00:00.0000000	30658928775339	786105	30658927989234
2022-10-13 00:00:00.0000000	31216481153202	855764	31216480297438
2022-10-14 00:00:00.0000000	29994067475960	805609	29994066670351
2022-10-15 00:00:00.0000000	26062657303884	688076	26062656615808
2022-10-16 00:00:00.0000000	26677494824042	541659	26677494282383
2022-10-17 00:00:00.0000000	34345294935948	785251	34345294150697
2022-10-18 00:00:00.0000000	32869741706676	616709	32869741089967
2022-10-19 00:00:00.0000000	32287511433942	592389	32287510841553
2022-10-20 00:00:00.0000000	31590873821571	629445	31590873192126
2022-10-21 00:00:00.0000000	31557133027327	572867	31557132454460

You can use the below Kusto query to check the overall increase in database size:

```
MonSqlRgHistory
| where  originalEventTimestamp > ago(30d)
| where event =~ 'aggregated_virtual_files_io_history'
    and (database_id == 11 and file_id == 1) and LogicalServerName == "prodsql50"
| extend size_on_disk_gb = size_on_disk_bytes / 1024 / 1024 / 1024
| summarize DatabaseSize = max(size_on_disk_gb) by bin(originalEventTimestamp,1d)
| project originalEventTimestamp,DatabaseSize
| render timechart
```



Mitigation

If using the above queries, we noticed a big increase rate in terms of new rows inserted in this case the sudden increase is related to the customer workload.

In case the above queries don't show an actual increase in the data size due to workload, refer to the TSG and blog article below to troubleshoot other causes:

- Azure SQL DB or SQL MI used data space is larger than expected - Overview (visualstudio.com)
- Azure SQL Database or SQL Managed Instance Database used data space is much larger than expected - Microsoft Community Hub

Internal Reference

ICM: Incident-337654646 Details - IcM (microsofticm.com)

Root Cause Classification

/Root Cause: Azure SQL v3/Performance/Space Management/User DB

How good have you found this content?