Tool for Gathering I/O Resource Manager Metrics: metric_iorm.pl (Doc ID 1337265.1)

To Bottom

In this Document

APPLIES TO:

Oracle Database Products > Exadata > Storage
Information in this document applies to any platform.

MAIN CONTENT

Purpose

Each Exadata storage cell gathers several types of metrics that are important for understanding how the storage is being shared by multiple workloads, their I/O performance, and the effect of I/O Resource Manager. These metrics include:

- Disk and flash usage, by database, pluggable database and consumer group.
- Disk latencies.
- If I/O Resource Manager is enabled, how long I/Os waited to be scheduled, by database, pluggable database and consumer group.

The Enterprise Manager 12c UI displays these metrics in user-friendly graphs. In the absence of the EM 12c UI, CellCLI can be used to view both current and historical values of these metrics. However, these metrics are much easier to interpret if the relevant ones are grouped together.

In this document, we provide an easy-to-use script for gathering I/O metrics and some pointers for how to interpret them. This script can also be used if I/O Resource Manager is not enabled. This script can be used for all versions of the Exadata storage cell.

Scope

This document is intended to be used by Exadata administrators who are using Exadata storage cells for mixed workloads or consolidation.

Tool for Gathering I/O Resource Manager Metrics: metric_iorm.pl

The script, metric_iorm.pl, is attached to this document. Place this script on each Exadata storage cell. We recommend that you put this script in your home directory for several reasons. First, files in your home directory are preserved across upgrades. Second, dcli automatically looks for the script in your home directory. Therefore, if the user name is "celladmin", then the script should be located in /home/celladmin.

You can then invoke it as follows:

1. To query the storage cell's current I/O metrics:

./metric_iorm.pl

2. To view the storage cell's historical I/O metrics, using explicit start and/or end times:

./metric_iorm.pl "where collectionTime > (start_time) and collectionTime < (end_time)"

The start and end times must be provided in a CellCLI-compliant format. For example, to view I/O metrics from 9 AM to 10 AM on 2011-05-25, using the timezone 7 hours behind UTC:

./metric_iorm.pl "where collectionTime > '2011-05-25T09:00:00-07:00' and collectionTime < '2011-05-25T10:00:00-07:00'"

3. To view the storage cell's historical I/O metrics, using relative start and/or end times:

./metric_iorm.pl "where ageInMinutes > (number_minutes) and ageInMinutes < (number_minutes)"

For example, to view I/O metrics between 2 and 3 hours ago:

./metric_iorm.pl "where ageInMinutes > 120 and ageInMinutes < 180"

4. To view I/O metrics from multiple storage cells, use dcli from the compute nodes.

Before running the metric script, you need to copy the script to all the storage cells as shown below. This copies the script to the default directory for user celladmin, /home/celladmin.

dcli -c cel01,cel02 -l celladmin -f metric_iorm.pl

dcli -g cell_names.dat -l celladmin -f metric_iorm.pl
where cell_names.dat contains the names of all the cells (cel01 and cel02), one per line.

You can then run the metric script based on the examples shown below.

dcli -c cel01,cel02 -l celladmin "/home/celladmin/metric_iorm.pl > /var/log/oracle/diag/asm/cell/metric_output"

dcli -g cell_names.dat -l celladmin "/home/celladmin/metric_iorm.pl > /var/log/oracle/diag/asm/cell/metric_output"
where cell_names.dat contains the names of all the cells (cel01 and cel02), one per line.

dcli -g cell_names.dat -l celladmin /home/celladmin/metric_iorm.pl \"where collectionTime \> \'2010-12-15T17:10:51-07:00\' and collectionTime \< \'2010-12-15T17:15:51-07:00\'\" \> /var/log/oracle/diag/asm/cell/metric_output

The Output of the Script

For each minute, metric_iorm.pl will output I/O metrics in the format below. This output can be seen in the file **/var/log/oracle/diag/asm/cell/metric_output** if it is being redirected to a file based on the above commands.

Database: PSFT
Utilization: Small=44% Large=3%
Flash Cache: IOPS=869 Space allocated=3200MB
Disk Throughput: MBPS=44
Small I/O's: IOPS=1246 Avg qtime=0.0ms
Large I/O's: IOPS=38.8 Avg qtime=0.0ms
  Consumer Group: _ORACLE_MEDPRIBG_GROUP_
  Utilization: Small=39% Large=0%
  Flash Cache: IOPS=315
  Disk Throughput: MBPS=9
  Small I/O's: IOPS=1039 Avg qtime=0.0ms
  Large I/O's: IOPS=4.7 Avg qtime=2.1ms
  Consumer Group: OTHER_GROUPS
  Utilization: Small=4% Large=0%
  Flash Cache: IOPS=546
  Disk Throughput: MBPS=0
  Small I/O's: IOPS=161 Avg qtime=0.0ms
  Large I/O's: IOPS=0.0 Avg qtime=0.0ms
  Consumer Group: _ORACLE_BACKGROUND_GROUP_
  Utilization: Small=0% Large=2%
  Flash Cache: IOPS=8.4
  Disk Throughput: MBPS=34
  Small I/O's: IOPS=46.1 Avg qtime=0.0ms
  Large I/O's: IOPS=34.1 Avg qtime=0.0ms

CELL METRICS SUMMARY

Cell Total Utilization: Small=44% Large=3%
Cell Total Flash Cache: IOPS=869 Space allocated=3200MB
Cell Total Disk Throughput: MBPS=44
Cell Total Small I/O's: IOPS=1246 Avg qtime=0.0ms
Cell Total Large I/O's: IOPS=38.8 Avg qtime=0.0ms

Cell Avg small read latency: 8.26 ms
Cell Avg small write latency: 17.23 ms
Cell Avg large read latency: 19.21 ms
Cell Avg large write latency: 10.08 ms

The Output of the Script for Container Databases

Database: ASM
Utilization: Small=0% Large=0%
Flash Cache: IOPS=0.0 Space allocated=0.000MB
Disk Throughput: MBPS=0
Small I/O's: IOPS=5.1 Avg qtime=0.2ms
Large I/O's: IOPS=0.0 Avg qtime=0.0ms
  Consumer Group: _ORACLE_BACKGROUND_GROUP_
  Utilization: Small=0% Large=0%
  Flash Cache: IOPS=0.0
  Disk Throughput: MBPS=0
  Small I/O's: IOPS=5.1 Avg qtime=0.2ms
  Large I/O's: IOPS=0.0 Avg qtime=0.0ms

Database: CDBIORM
Utilization: Small=0% Large=82%
Flash Cache: IOPS=0.0 Space allocated=54491MB
Disk Throughput: MBPS=0
Small I/O's: IOPS=1.5 Avg qtime=0.1ms
Large I/O's: IOPS=1639 Avg qtime=2563ms
  Pluggable Database: CDB$ROOT
  Utilization: Small=0% Large=0%
  Flash Cache: IOPS=0.0 Space allocated=347MB
  Disk Throughput: MBPS=0
  Small I/O's: IOPS=1.5 Avg qtime=0.1ms
  Large I/O's: IOPS=0.0 Avg qtime=0.0ms
    Consumer Group: CDB$ROOT_ORACLE_BACKGROUND_GROUP

Consumer Group: CDB$ROOT._ORACLE_BACKGROUND_GROUP_
Utilization: Small=0% Large=0%
Flash Cache: IOPS=0.0
Disk Throughput: MBPS=0
Small I/O's: IOPS=1.5 Avg qtime=0.1ms
Large I/O's: IOPS=0.0 Avg qtime=0.0ms
Pluggable Database: PDB1
Utilization: Small=0% Large=41%
Flash Cache: IOPS=0.0 Space allocated=52.688MB
Disk Throughput: MBPS=866
Small I/O's: IOPS=0.0 Avg qtime=0.0ms
Large I/O's: IOPS=827 Avg qtime=2781ms
Consumer Group: PDB1.SYS_GROUP
Utilization: Small=0% Large=41%
Flash Cache: IOPS=0.0
Disk Throughput: MBPS=866
Small I/O's: IOPS=0.0 Avg qtime=0.0ms
Large I/O's: IOPS=827 Avg qtime=2781ms
Pluggable Database: PDB3
Utilization: Small=0% Large=0%
Flash Cache: IOPS=0.0 Space allocated=26991MB
Disk Throughput: MBPS=6
Small I/O's: IOPS=0.0 Avg qtime=0.0ms
Large I/O's: IOPS=23.4 Avg qtime=12.2ms
Consumer Group: PDB3.OTHER_GROUPS
Utilization: Small=0% Large=0%
Flash Cache: IOPS=0.0
Disk Throughput: MBPS=6
Small I/O's: IOPS=0.0 Avg qtime=0.0ms
Large I/O's: IOPS=23.4 Avg qtime=12.2ms
Pluggable Database: PDB2
Utilization: Small=0% Large=39%
Flash Cache: IOPS=0.0 Space allocated=26988MB
Disk Throughput: MBPS=825
Small I/O's: IOPS=0.0 Avg qtime=0.0ms
Large I/O's: IOPS=788 Avg qtime=2411ms
Consumer Group: PDB2.OTHER_GROUPS
Utilization: Small=0% Large=39%
Flash Cache: IOPS=0.0
Disk Throughput: MBPS=825
Small I/O's: IOPS=0.0 Avg qtime=0.0ms
Large I/O's: IOPS=788 Avg qtime=2411ms

CELL METRICS SUMMARY

Cell Total Utilization: Small=0% Large=82%
Cell Total Flash Cache: IOPS=0 Space allocated=54379MB
Cell Total Disk Throughput: MBPS=1524.901
Cell Total Small I/O's: IOPS=8.9
Cell Total Large I/O's: IOPS=1544.4

Cell Avg small read latency: 77.26 ms
Cell Avg small write latency: 91.40 ms
Cell Avg large read latency: 142.01 ms
Cell Avg large write latency: 0.80 ms

The "utilization" metric shows you how busy the disks are, on average. This metric is broken down by database, pluggable database and by consumer group. In each case, it is also further broken down by small I/Os (indicative of latency-sensitive workloads) and large I/Os (indicative of throughput-intensive workloads). They correspond to these metrics: DB_IO_UTIL_SM, DB_IO_UTIL_LG, PDB_IO_UTIL_SM, PDB_IO_UTIL_LG, CG_IO_UTIL_SM, and CG_IO_UTIL_LG.

The "small and large IOPS" metrics are an alternative to the "utilization" metrics. They show the rate of I/Os per second to disk, broken down by database, pluggable database, consumer group, and I/O size. They correspond to these metrics: DB_IO_RQ_SM_SEC, DB_IO_RQ_LG_SEC, PDB_IO_RQ_SM_SEC, PDB_IO_RQ_LG_SEC, CG_IO_RQ_SM_SEC, CG_IO_RQ_LG_SEC.

The "throughput" metrics are another alternative to the "utilization" metrics. They show the rate of I/Os to disk in megabytes per second, broken down by database, pluggable database and consumer group. They correspond to these metrics: DB_IO_BY_SEC, PDB_IO_BY_SEC and CG_IO_BY_SEC.

The "Flash Cache IOPS" metric shows you the rate of I/Os to the flash cache. It is also broken down by database, pluggable database and by consumer group. It corresponds to these metrics: DB_FC_IO_RQ_SEC, PDB_FC_IO_RQ_SEC and CG_FC_IO_RQ_SEC.

The "Flash Cache Space Allocated" metric shows you the amount of space allocated to that database or pluggable database in flash cache. It corresponds to these metrics: DB_FC_BY_ALLOCATED and PDB_FC_BY_ALLOCATED.

The "avg qtime" metric shows you the average amount of time in milliseconds that I/Os waited to be scheduled by I/O Resource Manager. They are broken down by database, consumer group, and I/O size. It corresponds to these metrics: DB_IO_WT_SM_RQ, DB_IO_WT_LG_RQ, PDB_IO_WT_SM_RQ, PDB_IO_WT_LG_RQ, CG_IO_WT_SM_RQ, CG_IO_WT_LG_RQ.

The "disk latency" metrics show the average disk latency in milliseconds. They are broken down by I/O size and by reads and writes. They are not broken down by database and consumer group since the disk driver does not recognize I/Os by database and consumer group. Therefore, the disk latencies across databases and consumer groups are the same. They correspond to these metrics: CD_IO_TM_R_SM_RQ, CD_IO_TM_R_LG_RQ, CD_IO_TM_W_SM_RQ, CD_IO_TM_W_LG_RQ.

If your storage cell hosts multiple databases, then a database or pluggable database may not be listed for the following reasons:

- All metrics for the database are zero because the database is idle.
- The database is not explicitly listed in the inter-database IORM plan and the inter-database IORM plan does not have a default "share" directive. Use "list iormplan detail" to view the inter-database IORM plan.

How to Interpret the Output

These metrics can be used to answer the following, common questions:

Which database, pluggable database or consumer group is utilizing the disk the most? Use the disk utilization metrics to answer this question. You can also use the disk IOPS metrics. However, the total number of IOPS that can be sustained by the disk is extremely dependent on the ratio of reads vs writes, the location of the I/Os within the disk, and the ratio of small vs large I/Os. Therefore, we recommend using the disk utilization metrics.

Am I getting good latencies for my OLTP database, pluggable database or consumer group? The I/O latency, as seen by the database, is determined by the flash cache hit rate, the disk latency, and the IORM wait time. OLTP I/Os are small, so you should focus on the disk latency for small reads and writes. You can use IORM to improve the disk latency by giving high resource allocations to the OLTP databases, pluggable databases and consumer groups. If necessary, you can also use the "low latency" objective. You can also decrease the IORM wait time by giving high resource allocations to the critical databases, pluggable databases and consumer groups.

What is the flash cache hit rate for my database, pluggable database or consumer group? For OLTP databases, pluggable databases and consumer groups, you should expect the flash cache to be used for a significant number of I/Os. Since the latency of flash cache I/Os is very low, the I/O response time, as seen by the database, can be optimized by improving the flash cache hit rate for critical workloads.

How much is I/O Resource Manager affecting my database, pluggable database or consumer group? If IORM is enabled, then IORM may delay issuing an I/O when the disks are under heavy load or when a database, pluggable database or consumer group has reached its I/O utilization limit, if any. IORM may also delay issuing large I/Os when it is optimizing for low latencies for OLTP workloads. These delays are reflected in the "average queue time" metrics. You can decrease the delays for critical databases, pluggable database and consumer groups by increasing their resource allocation.

REFERENCES

NOTE:1339769.1:1340181.1

Didn't find what you are looking for? Ask in Community...

Attachments

Script for gathering IORM metrics (40.72 KB)

Related

Products

- Oracle Database Products > Exadata > Storage

Keywords

CELLCLI; CONSOLIDATION; EXADATA; IORM; LATENCY; METRIC; RESOURCE MANAGER

Back to Top