

# Oracle PGA hangolás adattárház környezetben

## PGA memória hangolás AIX operációs rendszeren

Okner Zsolt    Golden Ratio  
Erdősi Zoltán    Erste Bank Hungary

Budapest, 2022 november 30.



# Előzmények

- ODI alapokon futó feldolgozásoknál nagy sok esetben az AIX oprendszeren elfogyott a memória és a feldolgozás megállt.
- Sok a feldolgozási szál és ehhez társult rengeteg parallel szál (parallel\_max\_servers=980).
- Sok esetben PGA Cache Hit % értéke 70 % körül mozgott.
- DB cache-t nem használták az SQL utasítások, hanem direktbe a PGA töltötte az adatokat (sga\_target=400G)
- TEMP-en nagyon sok volt az írás és olvasás.

## IOStat by Function summary

- 'Data' columns suffixed with M,G,T,P are in multiples of 1024 other columns suffixed with K,M,G,T,P are in multiples of 1000
- ordered by (Data Read + Write) desc

Function Name	Reads: Data	Reqs per sec	Data per sec	Writes: Data	Reqs per sec	Data per sec	Waits: Count	Avg Tm(ms)
Direct Reads	68.5T	2153.56	1.1G	368.3G	46.02	6.161M	1951	28.11
Buffer Cache Reads	7.2T	719.88	122.942M	0M	0.00	0M	37.5M	13.40
Direct Writes	79.7G	3.76	1.333M	6.2T	702.21	106.914M	5516	33.20
Others	2.6T	1016.48	44.986M	1.1T	24.65	19.237M	60.3M	2.44
RMAN	925.1G	16.60	15.475M	3.6G	0.14	.06M	46K	17.89
LGWR	37M	0.04	.001M	872.5G	67.75	14.595M	2M	13.33
DBWR	170M	0.18	.003M	785.8G	224.41	13.144M	10.9K	4.23
Streams AQ	3M	0.00	0M	0M	0.00	0M	81	23.72
TOTAL:	79.3T	3910.51	1.3G	9.3T	1065.19	160.11M	99.8M	6.79

Top 10 Foreground Events by Total Wait Time

Event	Waits	Total Wait Time (sec)	Wait Avg(ms)	% DB time
direct path read	23,637,249	2.6M	108.39	22.7
DB CPU		965.1K		8.6
direct path read temp	29,383,557	308.4K	10.50	2.7
db file sequential read	57,799,810	291.8K	5.05	2.6
cursor: pin S wait on X	483,036	194.7K	402.99	1.7
db file scattered read	7,745,856	193.6K	25.00	1.7
read by other session	13,166,467	128.5K	9.76	1.1
direct path write temp	6,810,165	119.1K	17.49	1.1
control file sequential read	26,204,602	80.4K	3.07	.7
log file sync	927,482	44.7K	48.21	.4



# Előzmények 2

- Lassúak voltak a folyamatok, amin gyorsítani kellett.
- Történt memória bővítés az UNIX szerverben 1400 GB-ra lett felemelve a fizikai memória.
- A fizikai memória növelés és a PGA paramétereken nagyobbra állítása nem segített.
- Ezért kerestük meg a Great IT-t.

## PGA Aggr Summary

- PGA cache hit % - percentage of W/A (WorkArea) data processed only in-memory

PGA Cache Hit %	W/A MB Processed	Extra W/A MB Read/Written
76.26	23,206,397	7,225,689

[Back to Advisory Statistics](#)  
[Back to Top](#)

## PGA Aggr Target Stats

- B: Begin Snap E: End Snap (rows identified with B or E contain data which is absolute i.e. not diffed over the interval)
- Auto PGA Target - actual workarea memory target
- W/A PGA Used - amount of memory used for all Workareas (manual + auto)
- %PGA W/A Mem - percentage of PGA memory allocated to workareas
- %Auto W/A Mem - percentage of workarea memory controlled by Auto Mem Mgmt
- %Man W/A Mem - percentage of workarea memory under manual control

	PGA Aggr Target(M)	Auto PGA Target(M)	PGA Mem Alloc(M)	W/A PGA Used(M)	%PGA W/A Mem	%Auto W/A Mem	%Man W/A Mem	Global Mem Bound(K)
B	573,440	514,066	9,585.24	1,930.23	20.14	100.00	0.00	1,048,571
E	573,440	510,473	14,104.37	897.46	6.36	99.89	0.11	1,048,571

## PGA Aggr Target Histogram

- Optimal Executions are purely in-memory operations

Low Optimal	High Optimal	Total Execs	Optimal Execs	1-Pass Execs	M-Pass Execs
2K	4K	14,402,074	14,402,074	0	0
64K	128K	83,506	83,506	0	0
128K	256K	51,247	51,247	0	0
256K	512K	189,481	189,481	0	0
512K	1024K	9,152,281	9,152,281	0	0
1M	2M	140,648	140,648	0	0
2M	4M	77,571	77,543	28	0
4M	8M	51,962	51,859	103	0
8M	16M	48,439	48,366	73	0
16M	32M	29,911	29,898	13	0
32M	64M	25,713	25,602	111	0
64M	128M	31,651	27,423	4,228	0
128M	256M	15,432	12,861	2,571	0
256M	512M	6,524	4,011	2,513	0
512M	1024M	3,039	1,726	1,313	0
1G	2G	1,842	277	1,499	66
2G	4G	295	8	287	0
4G	8G	50	0	49	1
8G	16G	7	2	5	0
16G	32G	6	2	4	0
32G	64G	7	2	5	0
64G	128G	2	2	0	0



# PGA működése

- Hogyan működik az adatbázis PGA szempontjából
  - Kikapcsolat memory management-nél a PGA max mérete fix.
  - Nonshared / nem osztott memória, ellenben a SGA-val.
  - Minden szerverfolyamathoz egy PGA terület tartozik, így a háttér folyamatokhoz is, a parallel szervekhez is.
  - PGA méretezése lehet automatikus vagy manuális.
  - PGA nagy része SQL munkaterületek számára van szentelve.
    - Ezekből az SQL munkamenetekből a következők műveletek igénylik a legtöbb memóriát:
      - HASH-JOIN
      - SORT
      - AGGREGATION / GROUP BY
      - UNION / INTERSECT / MINUS ...
  - Ha nem fér el, akkor jön a TEMP-be való kiírás majd onnan a visszaolvasás
    - TEMP területen végzett műveletek viszont jóval lassabbak.



# PGA működése 2

## PGA allokáció működése globális szinten

PGA init paraméter:

- `pga_aggregate_limit` -- Ez egy határvonal, ami fölé nem mehet az oprendszeren a memória foglaltság. (12.1-től kezdve)
- `pga_aggregate_target` – Ez egy elméleti határ, ezen a határon felüli méretet is lefoglalhat a PGA össz. mérete

19c !!!

19c-ben erre a két paraméterre a következő összefüggés érvényes:

`pga_aggregate_limit = 2 * pga_aggregate_target`

Nem lehet se a limit paraméter kisebbre állítani, mint a target kétszerese a target paraméter állítása nélkül és nem lehet állítani target paramétert se a limit felénél nagyobb értékre állítani a másik állítás nélkül.

Ez a rejtett paraméter szabályozza ezt a összefüggést: `_pga_limit_target_perc`

```
SQL> /
NAME                VALUE DESCRIPTION
-----
_pga_limit_target_perc 200 default percent of pga_aggregate_target for pga_aggregate_limit
SQL>
```

Ezt érdemes átállítani adattárház környezetben 100-ra, ekkor be lehet állítani a target paramétert a limit közeli értékre.

### Limiting Process Size with Database Parameter `PGA_AGGREGATE_LIMIT` (Doc ID 1520324.1):

„The background process CKPT checks every **three** seconds to see whether the amount of memory exceeds the value of the `PGA_AGGREGATE_LIMIT` initialization parameter. If the limit is reached or exceeded, then this process terminates calls for the sessions using the most untunable PGA memory. If these sessions still do not release enough memory, then they are also terminated.”

(12c: Bug 30025002 ORA-4030: (QERHJ hash-join,klcqas:kllsltba) using HASH JOIN)

# PGA működése 3

## PGA allokáció működés session szinten

- `_pga_max_size` -> Maximális PGA mérete egy feldolgozáis szálnak
- `_smm_max_size` / `_smm_max_size_static` -> Maximális memória mérete a munkaterületnek feldolgozási szálnként
- `_smm_px_max_size` maximum -> Parallel lekérdezéseknél maximális memória mérete a munkaterületnek

<http://dba-blog.blogspot.com/2005/08/pgamaxsize-hidden-parameter.html>

*„Another hidden parameters that you must be aware of, are the `_SMM_MAX_SIZE` and `_SMM_PX_MAX_SIZE`.*

*`_SMM_MAX_SIZE` limits the maximum work area size for serial operations. `_SMM_MAX_SIZE` is calculated from the result of  $\min(5\% \text{ of } \text{PGA\_AGGREGATE\_TARGET}, \text{PGA\_MAX\_SIZE}/2)$ .*

*When you are executing a parallel query it is still using only one work area, all slaves are working under that work area.*

*`_SMM_PX_MAX_SIZE` limits the maximum size of used memory from all slave processes. The used size by each slave process inside that work area is limited from the value of `_SMM_MAX_SIZE` parameter. So, the value of `_SMM_MAX_SIZE` is used not only for serial but for parallel operations as well.*

*`_SMM_PX_MAX_SIZE` is used only when you are running query with a degree with a parallelism more than 6. For example, if you have `DOP = 6` then every slave process can takes no more than `_SMM_MAX_SIZE`. When `DOP > 6` then the `_SMM_PX_MAX_SIZE` is taking a place.*

*The values for `_SMM_MAX_SIZE` and `_SMM_PX_MAX_SIZE` are measured in KBytes.*

*After changing of the `_PGA_MAX_SIZE` parameter and restarting of the database, values for both `_SMM_MAX_SIZE` and `_SMM_PX_MAX_SIZE` are recalculated automatically. This means that you don't need to change them manually. But you must know what they are doing.”*

# PGA működése 4

## PGA allokáció

```
SELECT a.ksppinm Param, b.ksppstvl SessionVal FROM x$ksppi a, x$ksppcv b, x$ksppsv c
WHERE
    a.indx = b.indx
AND a.indx = c.indx
AND a.ksppinm in ('_pga_max_size','_smm_max_size','_smm_px_max_size','pga_aggregate_target');
```

PARAM	SESSIONVAL
pga_aggregate_target	21474836480
_pga_max_size	2147483648
_smm_max_size	1048576
_smm_px_max_size	10485760

Látszódik, hogy 20G PGA target esetén is csak 1G a work area size  
Bármekkora méretre állítjuk a PGA target paramétert ez a paraméter nem lesz 1G-nél nagyobb

# Megoldás

## Mi a megoldás ?

- PGA növelés az biztos, de azt hogyan ?
  - Meg kell találni az egyensúlyt a fizikai memória méret és a PGA\_AGGREGATE\_TARGET / PGA\_AGGREGATE\_LIMIT + SGA paraméterek és a rejtett paraméterek között ( \_pga\_max\_size , \_smm\_max\_size )

## Megoldási lépések nagy léptékekben:

- Több körös memória beállítások
- AWR elemzések
- AIX problémák
  - svmon
- Processzek számának változtatása, Parallel szálak változtatása
  - parallel\_max\_servers = 980 -> parallel\_max\_servers = 1500
- Resource Manager állítás
  - 32 maximális parallel feldolgozási szárról -> 48 maximális feldolgozási szál
- ODI oldalon több feldolgozás engedése egyidejűleg:
  - 30 -> 100 szálra



# Több körös memória beállítások

## 1. Kör

Ezen paraméterek lettek beállítva:

- alter system set pga\_aggregate\_limit='320G' scope=spfile; --> biztonsági vonal, hogy ne fogyjon el a hoszt memóriája
- alter system set pga\_aggregate\_target='200G' scope=spfile;
- alter system set "\_pga\_max\_size" = 4294967296 scope=spfile; --> egy session által használható pga memóriaszelet: maximum 4GB
- alter system set "\_smm\_max\_size"=4194304 scope=spfile; --> egy session által használható workarea memóriaszelet : maximum: 4GB
- alter system set "\_smm\_max\_size\_static"=4194304 scope=spfile; --> egy session által használható workarea memória szelet, perzisztálás (az előző beállítás nem marad meg az újraindítások között : maximum: 4GB

Itt előjött egy memória elfogyás probléma, de erre volt PATCH AIX platfromon:

**Patch 18674051: ORA-30009: NOT ENOUGH MEMORY FOR CONNECT BY WHILE THERE IS PLENTY MEMORY**

## 2. Kör

```
pga_aggregate_limit= 512G  
pga_aggregate_target= 512G  
_pga_max_size = 4G  
_smm_max_size / _smm_max_size_static = 4194304
```

N+1. kör-nél parallel\_max\_servers=2200 egy nagy memória elfogyást produkált.

### Legutolsó Kör

```
pga_aggregate_limit= 900G  
pga_aggregate_target= 700G  
_pga_max_size = 8G  
_smm_max_size / _smm_max_size_static = 8388608  
sga_target = 400G  
sga_max_size = 400G  
Parallel_max_server = 1500
```



# Összehasonlítás

2022 októberében ->

## PGA Aggr Summary

- PGA cache hit % - percentage of W/A (WorkArea) data processed only in-m

PGA Cache Hit %	W/A MB Processed	Extra W/A MB Read/Written
75.07	20,495,469	6,807,903

[Back to Advisory Statistics](#)  
[Back to Top](#)

## PGA Aggr Target Stats

- B: Begin Snap E: End Snap (rows identified with B or E contain data which is absolute i.e. not diffed over the interval)
- Auto PGA Target - actual workarea memory target
- W/A PGA Used - amount of memory used for all Workareas (manual + auto)
- %PGA W/A Mem - percentage of PGA memory allocated to workareas
- %Auto W/A Mem - percentage of workarea memory controlled by Auto Mem Mgmt
- %Man W/A Mem - percentage of workarea memory under manual control

	PGA Aggr Target(M)	Auto PGA Target(M)	PGA Mem Alloc(M)	W/A PGA Used(M)	%PGA W/A Mem	%Auto W/A Mem	%Man W/A Mem	Global Mem Bound(K)
B	573,440	510,713	17,316.33	5,199.00	30.02	100.00	0.00	1,048,576
E	573,440	505,825	93,190.29	13,481.38	14.47	99.84	0.16	1,048,576

## PGA Aggr Summary

- PGA cache hit % - percentage of W/A (WorkArea) data processed only in-memory

PGA Cache Hit %	W/A MB Processed	Extra W/A MB Read/Written
87.86	75,270,718	10,398,398

[Back to Advisory Statistics](#)  
[Back to Top](#)

## PGA Aggr Target Stats

- B: Begin Snap E: End Snap (rows identified with B or E contain data which is absolute i.e. not diffed over the interval)
- Auto PGA Target - actual workarea memory target
- W/A PGA Used - amount of memory used for all Workareas (manual + auto)
- %PGA W/A Mem - percentage of PGA memory allocated to workareas
- %Auto W/A Mem - percentage of workarea memory controlled by Auto Mem Mgmt
- %Man W/A Mem - percentage of workarea memory under manual control

	PGA Aggr Target(M)	Auto PGA Target(M)	PGA Mem Alloc(M)	W/A PGA Used(M)	%PGA W/A Mem	%Auto W/A Mem	%Man W/A Mem	Global Mem Bound(K)
B	716,800	621,181	542,228.52	36,051.07	6.65	99.96	0.04	8,388,608
E	716,800	625,499	282,920.22	4,816.12	1.70	99.83	0.17	8,388,608

<- 2020 októberében



# Összehasonlítás 2

2020 októberében ->

## IOStat by Function/Filetype summary

- 'Data' columns suffixed with M,G,T,P are in multiples of 1024 other columns suffixed with K,M,G,T,P are in multiples of 1000
- Ordered by (Data Read + Write) desc for each function

Function/File Name	Reads: Data	Reqs per sec	Data per sec	Writes: Data	Reqs per sec	Data per sec	Waits: Count	Avg Tm(ms)
Direct Reads	46T	1886.92	744.153M	416.7G	37.82	6.58M	2068	5.42
Direct Reads (Data File)	38.4T	636.99	621.633M	43.6G	3.62	.688M	765	2.38
Direct Reads (Temp File)	7.6T	1249.91	122.519M	373.1G	34.20	5.892M	0	
Direct Reads (Control File)	20M	0.02	0M	0M	0.00	0M	1303	7.20
Direct Writes	77.7G	3.41	1.227M	8T	854.36	128.632M	4564	15.59
Direct Writes (Temp File)	28.4G	2.55	.449M	7.6T	833.64	122.951M	0	
Direct Writes (Data File)	49.2G	0.81	.777M	359.8G	20.72	5.681M	973	4.43
Direct Writes (Control File)	56M	0.06	.001M	0M	0.00	0M	3591	18.62
Others	3.4T	969.61	54.753M	870.7G	16.93	13.749M	61.9M	0.98

## IOStat by Function/Filetype summary

- 'Data' columns suffixed with M,G,T,P are in multiples of 1024 other columns suffixed with K,M,G,T,P are in multiples of 1000
- Ordered by (Data Read + Write) desc for each function

Function/File Name	Reads: Data	Reqs per sec	Data per sec	Writes: Data	Reqs per sec	Data per sec	Waits: Count	Avg Time
Direct Reads	260T	15766560.79	3.1G	531.3G	40.93	6.297M	705	1.20ms
Direct Reads (Data File)	244.1T	13120490.45	2.9G	148.6G	20.32	1.761M	36	916.67us
Direct Reads (Temp File)	15.9T	2646070.33	193.053M	382.7G	20.60	4.535M	44	840.91us
Direct Reads (Control File)	9M	0.01	0M	1M	0.00	0M	625	1.24ms
Buffer Cache Reads	23.4T	1693498.13	283.584M	4M	0.00	0M	169.2M	4.34ms
Buffer Cache Reads (Data File)	15.5T	932323.17	188.419M	4M	0.00	0M	147.4M	4.01ms
Buffer Cache Reads (Temp File)	7.8T	761174.96	95.165M	0M	0.00	0M	21.7M	6.59ms
Direct Writes	1.3T	38524.25	15.248M	17.2T	1221.29	208.209M	28	.96ms
Direct Writes (Temp File)	26.9G	17686.91	.319M	16.1T	1199.91	195.383M	0	
Direct Writes (Data File)	1.2T	20836.69	14.928M	1.1T	21.38	12.826M	0	
Direct Writes (Control File)	1M	0.65	0M	0M	0.00	0M	28	.96ms
Others	5.2T	267301.96	62.768M	2.7T	70.51	32.54M	90.9M	1.60ms

<- 2022 októberében



# Összehasonlítás 3

2020 októberében ->

## Foreground Wait Events

- s - second, ms - millisecond, us - microsecond, ns - nanosecond
- Only events with Total Wait Time (s) >= .001 are shown
- ordered by wait time desc, waits desc (idle events last)
- %Timeouts: value of 0 indicates value was < .5%. Value of null is truly 0

Event	Waits	%Time-outs	Total Wait Time (s)	Avg wait	Waits /txn	% DB time
direct path read	147,155,345		3,769,207	25.61ms	62.70	36.94
direct path read temp	77,764,254		517,374	6.65ms	33.13	5.07
db file sequential read	196,189,749		471,210	2.40ms	83.59	4.62
db file scattered read	32,103,060		372,612	11.61ms	13.68	3.65
direct path write temp	27,758,873		270,705	9.75ms	11.83	2.65
read by other session	18,637,794		85,231	4.57ms	7.94	0.84
OJVM: Generic	348,539	98	54.573	156.58ms	0.15	0.53
library cache: mutex X	26,834,679		51,069	1.90ms	11.43	0.50
control file sequential read	30,257,842		39,939	1.32ms	12.89	0.39
SQL*Net message from dblink	8,162,684		36,573	4.48ms	3.48	0.36

## Foreground Wait Events

- s - second, ms - millisecond - 1000th of a second
- Only events with Total Wait Time (s) >= .001 are shown
- ordered by wait time desc, waits desc (idle events last)
- %Timeouts: value of 0 indicates value was < .5%. Value of null is truly 0

Event	Waits	%Time-outs	Total Wait Time (s)	Avg wait (ms)	Waits /txn	% DB time
direct path read	12,160,989		491,758	40.44	14.58	11.99
direct path read temp	44,649,485		160,302	3.59	53.53	3.91
db file sequential read	79,131,258		107,071	1.35	94.87	2.61
cursor: pin S wait on X	264,427		85,902	324.86	0.32	2.09
direct path write temp	10,138,384		67,882	6.70	12.16	1.65
db file scattered read	3,323,087		41,252	12.41	3.98	1.01
control file sequential read	23,970,050		34,006	1.42	28.74	0.83
read by other session	5,455,370		25,138	4.61	6.54	0.61
external table read	3,603		8,625	2393.94	0.00	0.21
SQL*Net more data from dblink	13,909,332		8,538	0.61	16.68	0.21

<- 2022 októberében

# Összehasonlítás 4

2020 októberében ->

## PGA Aggr Target Histogram

- Optimal Executions are purely in-memory operations

Low Optimal	High Optimal	Total Execs	Optimal Execs	1-Pass Execs	M-Pass Execs
2K	4K	165,876,091	165,876,091	0	0
64K	128K	222,842	222,842	0	0
128K	256K	264,138	263,432	706	0
256K	512K	804,259	803,845	414	0
512K	1024K	2,570,343	2,569,758	585	0
1M	2M	775,683	774,786	897	0
2M	4M	321,423	320,621	802	0
4M	8M	235,948	235,112	836	0
8M	16M	168,459	168,166	293	0
16M	32M	111,591	111,409	182	0
32M	64M	74,317	74,317	0	0
64M	128M	59,005	59,005	0	0
128M	256M	28,663	28,585	78	0
256M	512M	23,905	23,896	9	0
512M	1024M	14,289	13,910	379	0
1G	2G	6,214	5,387	827	0
2G	4G	2,244	1,638	606	0
4G	8G	683	268	415	0
8G	16G	201	4	197	0
16G	32G	87	7	80	0
32G	64G	4	0	4	0
64G	128G	27	20	7	0
128G	256G	1	0	1	0
2T	4T	2	2	0	0

## PGA Aggr Target Histogram

- Optimal Executions are purely in-memory operations

Low Optimal	High Optimal	Total Execs	Optimal Execs	1-Pass Execs	M-Pass Execs
2K	4K	8,703,283	8,703,283	0	0
64K	128K	54,368	54,368	0	0
128K	256K	44,473	44,473	0	0
256K	512K	168,407	168,407	0	0
512K	1024K	6,566,310	6,566,310	0	0
1M	2M	131,357	131,357	0	0
2M	4M	68,265	68,265	0	0
4M	8M	45,208	45,077	131	0
8M	16M	42,135	42,051	84	0
16M	32M	30,258	30,258	0	0
32M	64M	26,099	26,077	22	0
64M	128M	30,914	26,197	4,717	0
128M	256M	12,372	9,512	2,860	0
256M	512M	6,569	4,240	2,329	0
512M	1024M	3,121	1,931	1,190	0
1G	2G	1,769	229	1,478	62
2G	4G	292	14	278	0
4G	8G	44	2	41	1
8G	16G	4	0	4	0
16G	32G	6	2	4	0
32G	64G	6	2	4	0
64G	128G	2	2	0	0

<- 2022 októberében



# Köszönöm!

## Kérdések ?

