



Oracle Rdb Row Cache Update

Norman Lastovica / Philippe Vigier

Oracle Rdb Engineering

norman.lastovica@oracle.com

philippe.vigier@oracle.com

www.oracle.com/rdb

ORACLE

Copyright © 2002-2003, Oracle Corporation

April 20, 2004



Agenda

- Row Cache background
- How much is enough?
- Improvements
 - Snapshots in Row Cache
 - Native 64-bit Support
- Performance Characteristics
- Other things to consider

ORACLE

2

April 20, 2004



What is Row Cache?

- Section of globally accessible memory
- Stores data rows with minimum overhead
- Fastest access to frequently accessed rows
- VLM support (64-bit addressing; many GB)
- Allows tables to be pinned in memory
- Works along with global or local buffers
- Introduced in Rdb V7.0

ORACLE

3

20 April 2004



Row Cache Benefits

- Explicit control of what will be cached
- Efficient use of resources for shared data; higher data density in memory
- Lower overhead to access a row in cache than on page in buffer
- Reduced database page locking
- “Lazy” database updates from cache to disk
- VLM support - cache lots of data!

ORACLE

4

20 April 2004



Where Row Cache Wins

- When row is found in cache (a *hit*), there is no page I/O or locking
- If row is not found in cache (a *miss*), after it is read from database, it is inserted into cache
- A row in cache may be read or modified without database page I/O or locking!

ORACLE

5

20 April 2004



Large?

ORACLE

6

April 20, 2004



Rdb Does It

- **1 million rows?**
 - 100 million?
 - 1 billion?
- **10,000 index nodes?**
 - 100,000?
 - 1 million?
- **1 gb?**
 - 10 gb
 - 100 gb?
 - 1 tb?
- **10,000 records per day?**
 - 1 million per hour?
- **100 tps?**
 - 1,000 tps?
 - 5,000 tps?
 - 10,000 tps?

ORACLE

7

April 20, 2004



Handling it

ORACLE

8

April 20, 2004



Native 64-bit Row Cache

- Replace existing VLM/SSB techniques
- Improved performance
- Vastly larger caches viable

ORACLE

9

April 20, 2004



Snapshots in Row Cache

- Store snapshot copy of row in cache
 - Limited snapshot “life time”
- Memory access faster than disk access
 - No page locking or I/O
- Quickly write and read
 - No need to write snapshot page
 - No need to update live page
 - Search in memory

ORACLE

10

April 20, 2004



10+ Million Cached Rows

```

Rate: 3.00 Seconds      Row Cache Overview (Unsorted)      Elapsed: 00:32:04.22
Page: 1 of 1           DISK$RANDOM_SYS:[BIGDB]BIGDB.RDB;1      Mode: Online
-----
Cache.Name..... #Searches Hit% Full% #Inserts #Wrap #Slots Len
C0                91395 99.8 90.9      103    0 1100000 252
C1                91973 99.8 90.9        98    0 1100000 252
C2                91164 99.9 90.9        77    0 1100000 252
C3                91846 99.9 90.9        74    0 1100000 252
C4                91616 99.8 90.9        95    0 1100000 252
C5                91355 99.9 90.9        83    0 1100000 252
C6                91624 99.9 90.9        82    0 1100000 252
C7               180193 55.8 99.9     79585  1 1100000 252
C8                91289 99.9 90.9        77    0 1100000 252
C9                91567 99.8 90.9        98    0 1100000 252
I0               366860 100.0 42.6         0    0  25000 2000
I1               369179 100.0 43.0         0    0  25000 2000
I2               365736 100.0 44.4         0    0  25000 2000
I3               368531 100.0 45.0         0    0  25000 2000
I4               367782 100.0 44.9         0    0  25000 2000
I5               366650 100.0 44.9         0    0  25000 2000
I6               367731 100.0 44.9         0    0  25000 2000
I7               368110 99.9 89.0         0    0  25000 2000
I8               366337 100.0 45.0         0    0  25000 2000
I9               367563 100.0 44.9         0    0  25000 2000

```

ORACLE



100+ Million Cached Rows

```

Rate: 3.00 Seconds      Row Cache Overview (Unsorted)      Elapsed: 1 00:08:28.93
Page: 1 of 1           DISK$RANDOM_TEST:[BIGDB]BIGDB.RDB;1      Mode: Online
-----
Cache.Name..... #Searches Hit% Full% #Inserts #Wrap #Slots Len
C0              18847063 39.8 99.9 18626651  1 11000000 252
C1             10099635  0.9 90.9 10000076  0 11000000 252
C2             10100004  0.9 90.9 10000101  0 11000000 252
C3             10100241  0.9 90.9 10000109  0 11000000 252
C4             10099578  0.9 90.9 10000102  0 11000000 252
C5             10099653  0.9 90.9 10000106  0 11000000 252
C6             10099868  0.9 90.9 10000098  0 11000000 252
C7             10099586  0.9 90.9 10000099  0 11000000 252
C8             10100385  0.9 90.9 10000107  0 11000000 252
C9             10099742  0.9 90.9 10000102  0 11000000 252
I0            128530783 99.8 44.8   215296  0  330000 2000
I1             69350965 99.8 34.1   112657  0  330000 2000
I2             69355896 99.8 34.1   112695  0  330000 2000
I3             69300014 99.8 34.7   114778  0  330000 2000
I4             69596633 99.8 35.8   118162  0  330000 2000
I5             69586400 99.8 35.8   118168  0  330000 2000
I6             69580204 99.8 35.8   118173  0  330000 2000
I7             69587849 99.8 35.8   118186  0  330000 2000
I8             69598640 99.8 35.8   118166  0  330000 2000
I9             69549265 99.8 35.8   118157  0  330000 2000

```

ORACLE



...Under Load

Rate: 3.00 Seconds		Summary IO Statistics			Elapsed: 00:01:52.92	
Page: 1 of 1		DISK\$RANDOM_TEST:[BIGDB]BIGDB.RDB;1			Mode: Online	
statistic.....	rate.per.second.....	total.....	average.....			
name.....	max..... cur..... avg.....	count.....	per.trans....			
transactions	2068 2018 1975.2	223043	1.0			
verb successes	14419 14076 13783.2	1556409	6.9			
verb failures	0 0 0.0	0	0.0			
synch data reads	142 50 49.4	5581	0.0			
synch data writes	4 2 2.1	243	0.0			
asynch data reads	0 0 0.0	0	0.0			
asynch data writes	3 2 1.9	217	0.0			
RUJ file reads	0 0 0.0	0	0.0			
RUJ file writes	0 0 0.0	1	0.0			
AIJ file reads	0 0 0.0	1	0.0			
AIJ file writes	1059 1008 0.0	112858	0.5			
root file reads	0 0 0.0	0	0.0			
root file writes	527 500 17.7	55085	0.2			

TIME IN PROCESSOR MODES	0	400	800	1200	1600
Combined for 16 CPUs	+ - - - - + - - - - + - - - - + - - - - +				
Interrupt State	57		X		
MP Synchronization	30				
Kernel Mode	331		XXXXXXXX		
Executive Mode	665		XXXXXXXXXXXXXXXXXXXX		
User Mode	486		XXXXXXXXXXXXXXXX		
Idle Time	32				

ORACLE



One Final Example

•bil·lion *n.*

1. The cardinal number equal to 10^9
2. Chiefly British. The cardinal number equal to 10^{12}
3. An indefinitely large number

ORACLE



Cache It

- **202+Gb (data & index)**
- **1,000,000,000 data rows**
 - 100 Bytes
- **12,400,000 index nodes**
 - Sorted Ranked 2000 byte nodes
 - 5 levels

ORACLE

15

April 20, 2004



Big System

- **GS1280**
- **32 EV7 processors**
- **256gb**
- **EVA storage**
- **VMS 7.3-2**
- **Rdb 7.1.2.0.1**

ORACLE

16

April 20, 2004



Billion+ Cached Rows

```

Node: MARV64 (1/1/1) Oracle Rdb V7.1-201 Perf. Monitor 8-OCT-2003 12:24:48.51
Rate: 3.00 Seconds Row Cache Overview (Unsorted) Elapsed: 13:27:23.07
Page: 1 of 1 DISK$DATABASE:[BIGDB]BIGDB.RDB;1 Mode: Online
-----
Cache.Name..... #Searches Hit% Full% #Inserts #Wrap #Slots Len
C0 150194926 33.3 99.0 100041567 0 101000000 100
C1 150179641 33.3 99.0 100041746 0 101000000 100
C2 150194907 33.3 99.0 100041744 0 101000000 100
C3 150172693 33.3 99.0 100041135 0 101000000 100
C4 150181844 33.3 99.0 100041620 0 101000000 100
C5 150175533 33.3 99.0 100042025 0 101000000 100
C6 150182697 33.3 99.0 100041568 0 101000000 100
C7 150195409 33.3 99.0 100041419 0 101000000 100
C8 150199047 33.3 99.0 100041586 0 101000000 100
C9 150181963 33.3 99.0 100041668 0 101000000 100
I0 252707121 99.5 74.6 1130743 0 1515000 2000
I1 252725604 99.5 78.9 1195574 0 1515000 2000
I2 252822011 99.5 79.7 1207765 0 1515000 2000
I3 252702306 99.5 79.7 1207814 0 1515000 2000
I4 252754270 99.5 79.7 1207794 0 1515000 2000
I5 252725336 99.5 79.7 1207759 0 1515000 2000
I6 252804413 99.5 81.7 1239244 0 1515000 2000
I7 252867375 99.5 81.8 1239273 0 1515000 2000
I8 252890671 99.5 81.8 1239285 0 1515000 2000
I9 252796357 99.5 81.7 1239239 0 1515000 2000
RDB$SYSTEM_CACHE 501772322 99.9 13.5 675 0 5000 1000

```



A Workload... CPU Utilization

```

+-----+
| CUR |
+-----+
TIME IN PROCESSOR MODES
on node MARV64
7-OCT-2003 19:42:34.89
Combined for 32 CPUs
0 800 1600 2400 3200
+-----+
Interrupt State 69 |
MP Synchronization 89 |X
Kernel Mode 699 |XXXXXXXXXX
Executive Mode 1392 |XXXXXXXXXXXXXXXXXXXXX
User Mode 892 |XXXXXXXXXXXXX
Idle Time 62 |
+-----+
LOCK MANAGEMENT STATISTICS CUR AVE MIN MAX
New ENQ Rate 47053.66 47814.85 46486.00 49408.00
Converted ENQ Rate 64719.33 65602.52 63977.33 67816.00
DEQ Rate 37902.00 38536.00 37434.00 39768.00
Blocking AST Rate 7491.33 7549.33 7344.33 7756.00
ENQs Forced To Wait Rate 18905.33 19152.57 18704.66 19802.00
ENQs Not Queued Rate 9142.33 9292.14 9072.33 9638.00

```



A Workload... 12,757 TPS

Node: MARV64 (1/1/1) Oracle Rdb V7.1-201 Perf. Monitor 7-OCT-2003 19:46:19.27
Rate: 3.00 Seconds Summary IO Statistics Elapsed: 00:05:52.40
Page: 1 of 1 DISK\$DATABASE:[BIGDB]BIGDB.RDB;1 Mode: Online

statistic..... name.....	rate.per.second.....			total..... count.....	average..... per.trans....
	max.....	cur.....	avg.....		
transactions	13228	12672	12757.1	4592520	1.0
verb successes	91904	89550	86881.3	31688388	6.9
verb failures	0	0	0.0	0	0.0
synch data reads	843	341	360.7	241647	0.0
synch data writes	17	14	13.5	7565	0.0
asynch data reads	0	0	0.0	0	0.0
asynch data writes	16	13	11.9	6379	0.0
RUJ file reads	0	0	0.0	0	0.0
RUJ file writes	0	0	0.0	1	0.0
AIJ file reads	21	0	0.5	66	0.0
AIJ file writes	2316	2193	0.0	1250634	0.1
root file reads	0	0	0.0	0	0.0
root file writes	3438	3324	119.8	1875589	0.2

ORACLE



1 Million Transactions Per Minute

Node: MARV64 (1/1/1) Oracle Rdb V7.1-201 Perf. Monitor 9-OCT-2003 00:04:49.67
Rate: 3.00 Seconds Summary IO Statistics Elapsed: 00:03:44.25
Page: 1 of 1 DISK\$DATABASE:[BIGDB]BIGDB.RDB;1 Mode: Online

statistic..... name.....	rate.per.second.....			total..... count.....	average..... per.trans....
	max.....	cur.....	avg.....		
transactions	17014	16836	16356.8	3668051	1.0
verb successes	119091	117838	114477.2	25671742	6.9
verb failures	0	0	0.0	0	0.0
synch data reads	414	16	73.1	16414	0.0
synch data writes	19	7	11.4	2570	0.0
asynch data reads	0	0	0.0	0	0.0
asynch data writes	0	0	0.0	2	0.0
RUJ file reads	0	0	0.0	0	0.0
RUJ file writes	0	0	0.0	0	0.0
AIJ file reads	15	0	0.4	96	0.0
AIJ file writes	2863	2805	0.0	613076	0.1
root file reads	0	0	0.0	0	0.0
root file writes	54	32	28.5	10067	0.0

ORACLE



Performance Test Case

ORACLE

21

April 20, 2004



Snapshots In Row Cache

- Read and modify records from cache
- Maintain record snapshot in cache
- Avoid IO to database and page locks
- Should improve application performance
 - That's what we try to demonstrate here 😊

ORACLE

22

April 20, 2004



Simple Database

```
SQL> create table t1 (  
    f1 integer,  
    f2 integer,  
    f3 char(50));  
  
SQL> create storage map t1_map  
    for t1 store in foo_u1;  
  
SQL> create unique index  
    it1 on t1 (f1)  
    type is sorted  
    store in foo_u1_idx;
```

ORACLE

23

April 20, 2004



Simple Processes

- **Writer**
 - Does 5000 read write transactions
 - Each transaction does 20 updates of the form:

```
update t1 set f2 = :lp where f1 = :k;
```
 - :k is incremented in circular way from 501 to 600
- **Reader**
 - Does 5000 read only transactions
 - Each transaction does 30 selects of the form:

```
select f2 into :r from t1 where f1 = :v;
```
 - :v is incremented in circular way from 501 to 600

ORACLE

24

April 20, 2004



A Lot Of Page Contention

- Reader and Writer processes access :
 - The same 100 records
 - Distributed over the same 4 database pages

ORACLE

25

April 20, 2004



Simple Database

```
SQL> create data file tdb:foo
open is manual
reserve 10 storage areas
reserve 10 journals
async batch writes disabled
async prefetch disabled
detected async prefetch disabled

create storage area foo_ul
filename tdb:foo_ul

create storage area foo_ul_idx
filename tdb:foo_ul_idx;
```

This is to get
the actual IO
to the
database

ORACLE

26

April 20, 2004



Simple Database (cont.)

```
SQL> alter data file db
      reserve 10 cache slots
      row cache is enabled
      number of cluster node 1 -- or galaxy cluster

      add journal fooaij1 filename fooaij1.aij
      add journal fooaij2 filename fooaij2.aij
      add journal fooaij3 filename fooaij3.aij
      add journal fooaij4 filename fooaij4.aij

      journal is enabled
      allocation 50000 blocks,
      fast commit enabled (
      checkpoint interval 10000 blocks));
```

Blue is what is mandatory for row cache

ORACLE

27

April 20, 2004



Sizing The Cache

```
$ RMU/ANAL/LAREA=T1 DB
```

```
Logical area: T1 for storage area : FOO_U1
Larea id: 57, Record type: 30,
Record length: 66, Compressed
```

```
Data records: 20000, bytes used: 259745 (36%)
average length: 13, compression ratio: 0.21
```

```
$ RMU/ANAL/INDEX DB IT1
```

```
Index IT1 for relation T1 duplicates not allowed
Max Level: 3, Nodes: 407,
Used/Avail: 157190/161986 (97%), Keys: 20406,
Records: 20000
```

ORACLE

28

April 20, 2004



Defining The Caches

```
SQL> alter data file db

add cache t1
  cache size 21000 rows row length 66 bytes
  row snapshot enabled ( cache size is 10000 rows )

add cache it1
  cache size 500 rows row length 430 bytes
  row snapshot enabled ( cache size is 5000 rows );
```

ORACLE

29

April 20, 2004



Index Is In Cache

```
Rate: 3.00 Seconds          Row Cache Status          Elapsed: 00:06:41.0
Page: 1 of 1                RAID1:[VIGIER.SNAPINK.DB]FOO.RDB;1          Mode: Online
-----
                          For Cache: IT1
Statistic.Name Stat.Value Percent
Total slots:           500 100.0% Slot Length: 432 Hash slots: 512
Slots full:            407 81.4% Use:           0 0.0%
Slots empty:           93 18.6% Rsv:           0 0.0%
Marked Slots:          56 11.2% Hot:           0 0.0% Cold:           56 100.0%
Clean Slots:           444 88.8% Hot:           0 0.0% Cold:           444 100.0%
Used Space:            175k 81.0% Wstd:          0k 0.3%
Free Space:            40k 18.6%
Hash Que Lengths:      Empty:138 1:341      2:33      3:0      4+:0
Cursor position:       410 of 500 wrapped 0 times
Checkpoints: 2 Last:   2-JUN-2003 09:25:35.06 Location: 1:2
Cache Recovery:        1:2
Snap Slots:            5000 100.0% Ful:          1 0.2% Rcl:          1 0.2%
Snap Cursor:           10 of 5000 (slot 510) wrapped 0 times
```

ORACLE

30

April 20, 2004



Table Is In Cache

```

Rate: 3.00 Seconds                      Row Cache Status                      Elapsed: 00:06:59.53
Page: 1 of 1                            RAID1:[VIGIER.SNAPINK.DB]FOO.RDB;1    Mode: Online
-----
                                For Cache: T1
Statistic.Name Stat.Value Percent
Total slots:                21000 100.0% Slot Length: 68 Hash slots: 32768
Slots full:                  20000 95.2% Use:          0 0.0%
Slots empty:                  1000  4.7% Rsv:          0 0.0%
Marked Slots:                52 0.2% Hot:          0 0.0% Cold:          52 100.0%
Clean Slots:                 20948 99.7% Hot:          0 0.0% Cold:          20948 100.0%
Used Space:                   359k 25.1% Wstd:         1000k 70.0%
Free Space:                   68k  4.7%
Hash Que Lengths: Empty:25323 1:20      2:2295      3:5130      4+:0
Cursor position: 20013 of 21000 wrapped 0 times
Cache latched: No
Cache is full: No                Cache modified: Yes Snapshot is full: No
Checkpoints: 2 Last: 2-JUN-2003 09:25:35.06 Location: 1:2
Cache Recovery: 1:2
Snap Slots: 10000 100.0% Ful: 0 0.0% Rcl: 0 0.0%
Snap Cursor: 0 of 10000 (slot 21000) wrapped 0 times

```



1st Run, Not So Good !

```

Rate: 1.00 Second                      Summary IO Statistics                      Elapsed: 00:49:16.14
Page: 1 of 1                            RAID1:[VIGIER.SNAPINK.DB]FOO.RDB;1    Mode: Online
-----
Statistic..... rate.per.second..... total..... average.....
name..... max..... cur..... avg..... count..... per.trans....
transactions          1565          0          1.2          3624          1.0
verb successes         7863          0          7.7         22931          6.3
verb failures           0            0          0.0           0            0.0

synch data reads       143           1          2.4          7230          1.9
synch data writes       86            2          3.2          9474          2.6
asynch data reads       0            0          0.0           0            0.0
asynch data writes     0            0          0.0           0            0.0
RUJ file reads         0            0          0.0           0            0.0
RUJ file writes        34            1          0.0          3728          1.0
root file reads        0            0          0.0           0            0.0
root file writes        6            0          0.0           260           0.0

```




What Are Those IO ?

```

Rate: 1.00 Second      File IO Overview (Unsorted Total I/O      Elapsed: 00:49:37.19
Page: 1 of 1          RAID1:[VIGIER.SNAPINK.DB]FOO.RDB;1      Mode: Online
-----
File/Storage.Area.Name..... Sync.Reads SyncWrites AsyncReads AsyncWrits PgDis
Database Root                0          1          0          260      0
AIJ (After-Image Journal)    1          241         0          0        0
RUJ (Recovery-Unit Journal)  0          55          0          3694     0
All data/snap files          7275       9522         0          0        0
data FOO                      178        0           0          0        0
data FOO_U1                   4484       5160         0          0        0
data FOO_U1_IDX              0          0           0          0        0
snap FOO                      0          0           0          0        0
snap FOO_U1                   2613       4362         0          0        0
snap FOO_U1_IDX              0          0           0          0        0

```



Look At Cache Statistics

```

Rate: 1.00 Second      Row Cache (T1)      Elapsed: 00:49:58.36
Page: 1 of 1          RAID1:[VIGIER.SNAPINK.DB]FOO.RDB;1      Mode: Online
-----
statistic..... rate.per.second..... total..... average.....
name..... max..... cur..... avg..... count..... per.trans....
latch requests    1          1          0.0          49          0.0
retrieved         0          0          0.0          0           0.0
cache searches   46824       202        36.2        108817       29.2
  found in workset 0          0          0.0          0           0.0
  found in cache   46825       47         20.1        60471       16.2
  found too big    2152       156        16.1        48363       13.0
insert cache      0          0          0.0          0           0.0
row too big       0          0          0.0          0           0.0
cache full        0          0          0.0          0           0.0
collision         0          0          0.0          0           0.0
hash misses      39336       191        30.4        91415       24.6
cache unmark      23         14         0.2          766         0.2

```



Look At Record Statistics

```
Rate: 1.00 Second           Record Statistics           Elapsed: 00:50:08.57
Page: 1 of 1                RAID1:[VIGIER.SNAPINK.DB]FOO.RDB;1           Mode: Online
-----
```

statistic.....	rate.per.second.....			total.....	average.....
name.....	max.....	cur.....	avg.....	count.....	per.trans....
record marked	86	0	1.6	4884	1.3
record fetched	187962	0	144.9	435923	117.3
fragmented	4301	0	31.7	95640	25.7
record stored	11	0	0.0	81	0.0
fragmented	0	0	0.0	0	0.0
pages checked	11	0	0.0	81	0.0
saved IO	10	0	0.0	53	0.0
discarded	0	0	0.0	0	0.0
record erased	0	0	0.0	0	0.0
fragmented	0	0	0.0	0	0.0

ORACLE

35

April 20, 2004



Disable Compression

- Now Reader and Writer processes access :
 - The same 100 records
 - Distributed over the same 9 database pages

```
SQL> alter storage map t1_map
      disable compression;
```

ORACLE

36

April 20, 2004



ZERO I/O

Rate: 1.00 Second File IO Overview (Unsorted Total I/O Elapsed: 00:00:32.39
Page: 1 of 1 RAID1:[VIGIER.SNAPINK.DB]FOO.RDB;3 Mode: Online

File/Storage.Area.Name.....	Sync.Reads	SyncWrites	AsyncReads	AsyncWrites	PgDis
Database Root	0	6	0	5321	0
AIJ (After-Image Journal)	1	0	0	5003	0
RUJ (Recovery-Unit Journal)	0	0	0	0	0
All data/snap files	53	0	0	0	0
data FOO	53	0	0	0	0
data FOO_U1	0	0	0	0	0
data FOO_U1_IDX	0	0	0	0	0
snap FOO	0	0	0	0	0
snap FOO_U1	0	0	0	0	0
snap FOO_U1_IDX	0	0	0	0	0



Few Page Locks

Rate: 1.00 Second Locking (buffer/page locks) Elapsed: 00:00:52.12
Page: 1 of 1 RAID1:[VIGIER.SNAPINK.DB]FOO.RDB;3 Mode: Online

statistic..... name.....	rate.per.second.....			total..... count.....	average..... per.trans....
	max.....	cur.....	avg.....		
locks requested	112	0	2.1	112	0.0
rqsts not queued	16	0	0.3	16	0.0
rqsts stalled	14	0	0.2	14	0.0
rqst timeouts	0	0	0.0	0	0.0
rqst deadlocks	0	0	0.0	0	0.0
locks promoted	46	0	1.1	60	0.0
locks demoted	102	0	2.2	116	0.0
locks released	49	0	2.1	111	0.0
blocking ASTs	14	0	0.2	14	0.0
stall time x100	0	0	0.0	0	0.0



What About The Application ...

- Average transaction duration : 0.01s
- Elapse time for the writer : 20s
- Elapse time for the reader : 16s

ORACLE

39

April 20, 2004



Now: Row Cache Without Snapshots In Cache

```
SQL> alter data file db
      alter cache t1
        cache size 21000 rows
        row length 66 bytes
        row snapshot disabled
      alter cache it1
        cache size 500 rows
        row length 430 bytes
        row snapshot disabled;
```

ORACLE

40

April 20, 2004



Still All Live Rows In Cache

```
Rate: 3.00 Seconds           Row Cache Status           Elapsed: 00:04:20.74
Page: 1 of 1                 RAID1:[VIGIER.SNAPINK.DB]FOO.RDB;1      Mode: Online
-----
                                For Cache: IT1
Statistic.Name Stat.Value Percent
Total slots:           500 100.0% Slot Length: 432 Hash slots: 512
Slots full:            399 79.8% Use:           0 0.0%
Slots empty:           101 20.2% Rsv:           0 0.0%
Marked slots:          131 26.2% Hot:           0 0.0% Cold:          131 100.0%
Clean Slots:           369 73.8% Hot:           0 0.0% Cold:          369 100.0%
Used Space:             171k 79.4% Wstd:          0k 0.3%
Free Space:             43k 20.2%
Hash Que Lengths:      Empty:142 1:341      2:29      3:0      4+:0
Cursor position:       400 of 500 wrapped 0 times
Checkpoints: 2 Last:   2-JUN-2003 11:32:17.04 Location: 1:2
Cache Recovery:        1:2
```

ORACLE

41

April 20, 2004



Still All Live Rows In Cache (cont.)

```
Rate: 3.00 Seconds           Row Cache Status           Elapsed: 00:04:30.31
Page: 1 of 1                 RAID1:[VIGIER.SNAPINK.DB]FOO.RDB;1      Mode: Online
-----
                                For Cache: T1
Statistic.Name Stat.Value Percent
Total slots:           21000 100.0% Slot Length: 68 Hash slots: 32768
Slots full:            20000 95.2% Use:           0 0.0%
Slots empty:           1000 4.7% Rsv:           0 0.0%
Marked Slots:          121 0.5% Hot:           0 0.0% Cold:          121 100.0%
Clean Slots:           20879 99.4% Hot:           0 0.0% Cold:          20879 100.0%
Used Space:             1320k 92.4% Wstd:          40k 2.8%
Free Space:             68k 4.7%
Hash Que Lengths:      Empty:12768 1:20000      2:0      3:0      4+:0
Cursor position:       20013 of 21000 wrapped 0 times
Checkpoints: 2 Last:   2-JUN-2003 11:32:17.04 Location: 1:2
Cache Recovery:        1:2
```

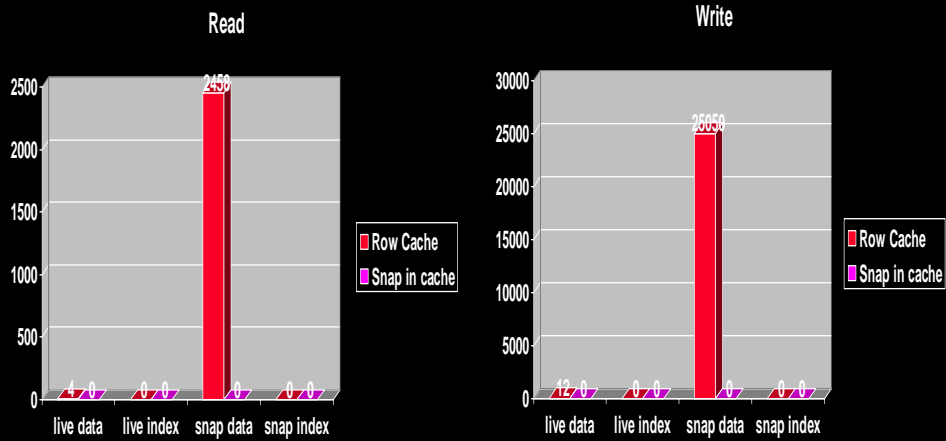
ORACLE

42

April 20, 2004



Results: Database IO



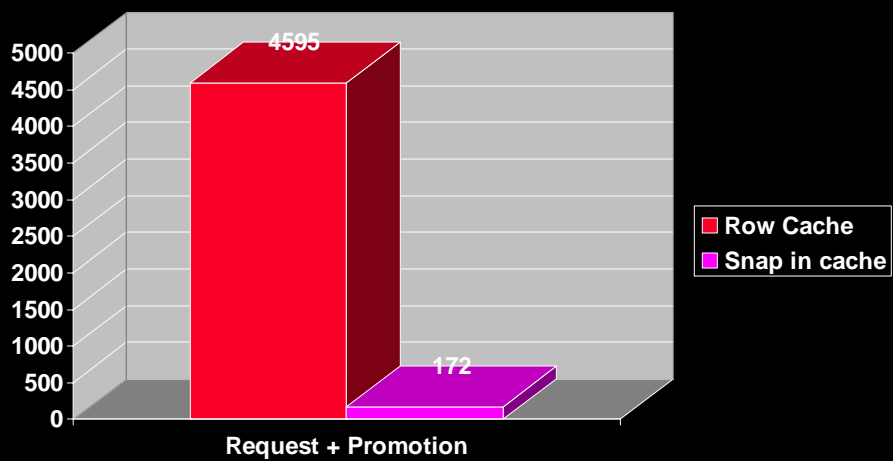
ORACLE

43

April 20, 2004



Results: Page Locking



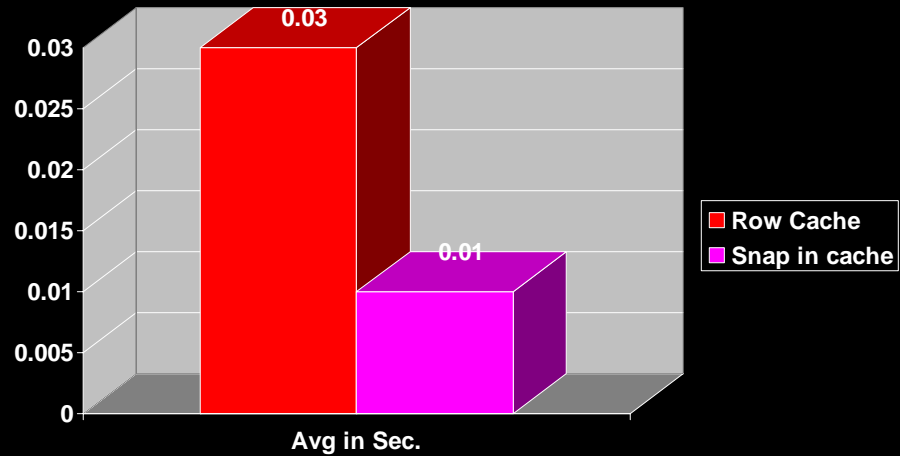
ORACLE

44

April 20, 2004



Results: Average Transaction Duration



ORACLE

45

April 20, 2004



Now: Without Row Cache

- We will have database default
 - 20 local buffers

```
SQL> alter data file db  
      row cache is disabled;
```

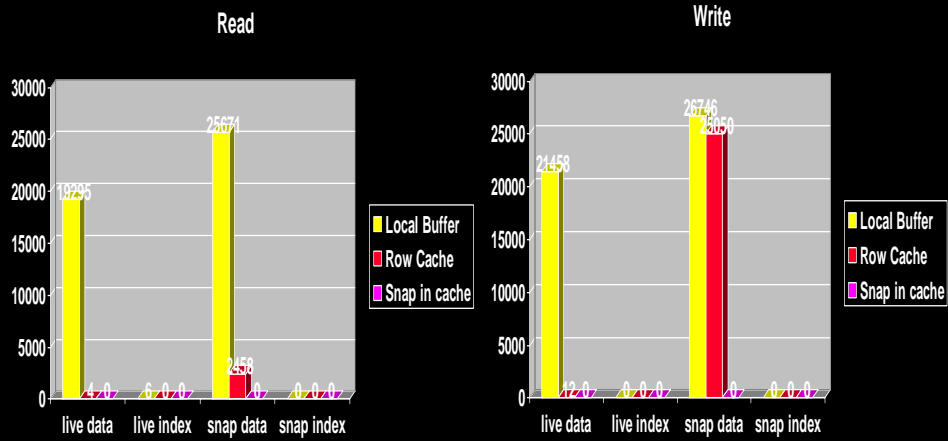
ORACLE

46

April 20, 2004



Results: Database IO



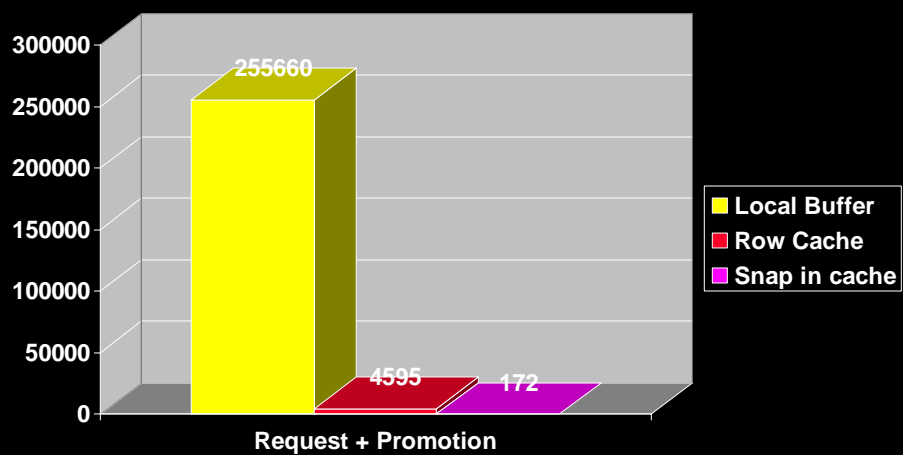
ORACLE

47

April 20, 2004



Results: Page Locking



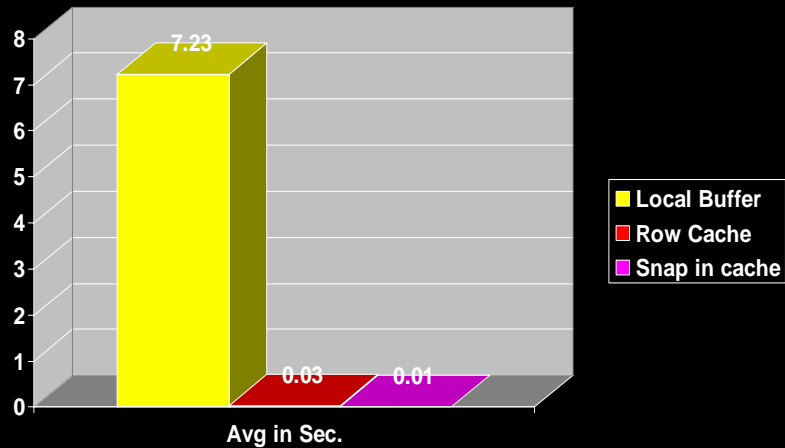
ORACLE

48

April 20, 2004



Results: Average Transaction Duration



ORACLE

49

April 20, 2004



Now: With Global Buffers

- Should save read IO
- We will use default values
 - Number of buffers : 250
 - User limit : 5

```
SQL> alter data file db
      global buffers enabled;
```

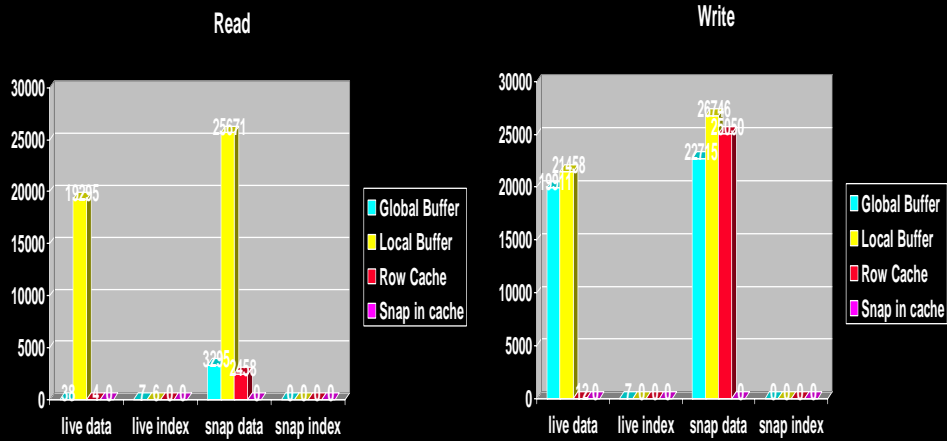
ORACLE

50

April 20, 2004



Results: Database IO



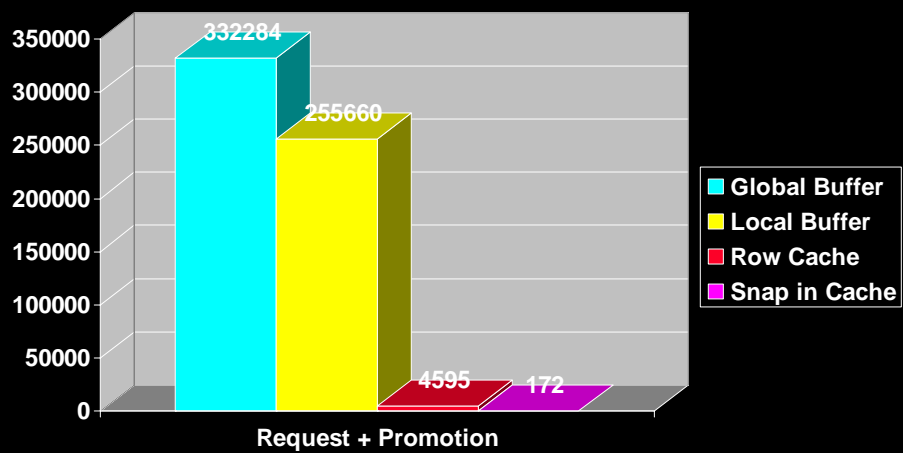
ORACLE

51

April 20, 2004



Results: Page Locking



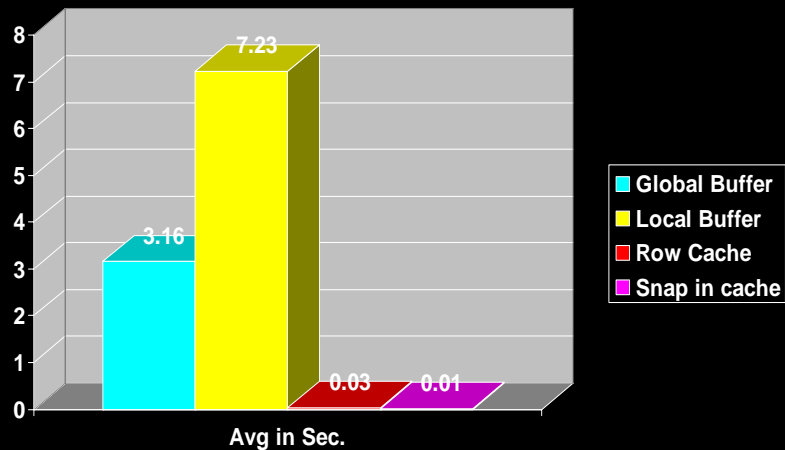
ORACLE

52

April 20, 2004



Results: Average Transaction Duration



ORACLE

53

April 20, 2004



Now: With Global Buffers & Page Transfer Via Memory

- Should save write IO
- Here we use
 - Number of buffers : 1000
 - User limit : 250

```
SQL> alter data file db
      global buffers are enabled (
        number is 1000,
        user limit is 250,
        page transfer via memory);
```

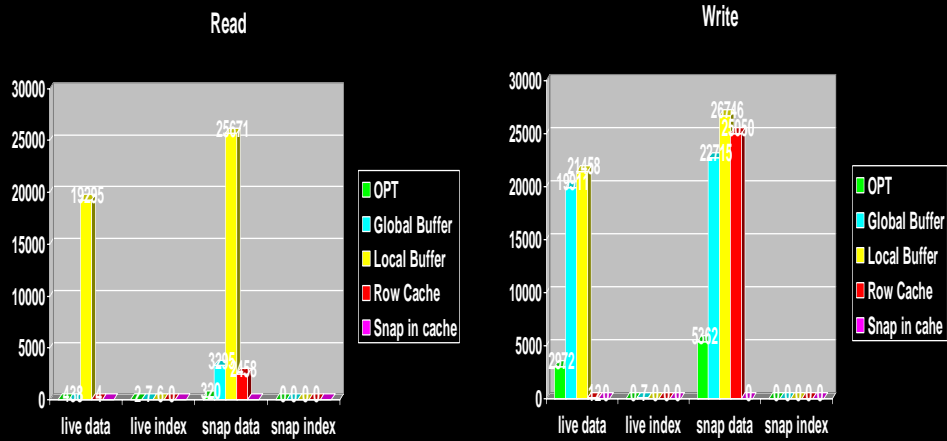
ORACLE

54

April 20, 2004



Results: Database IO



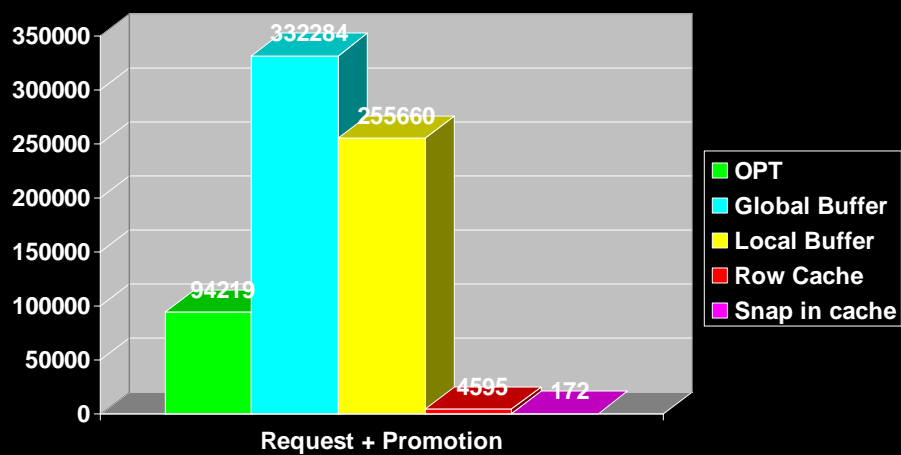
ORACLE

55

April 20, 2004



Results: Page Locking



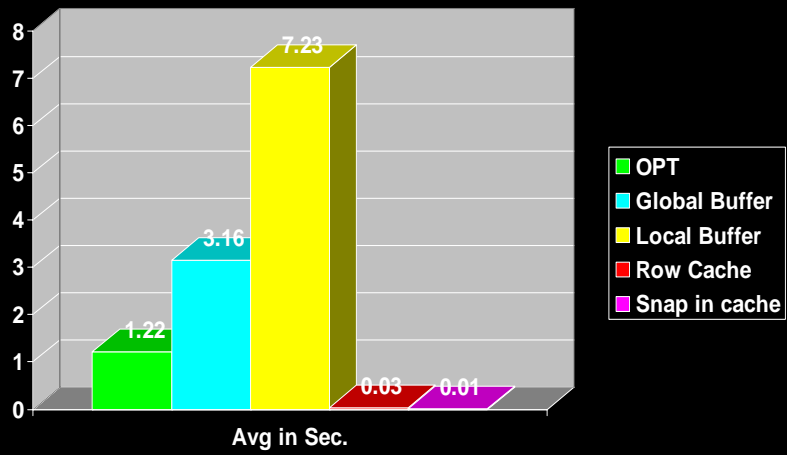
ORACLE

56

April 20, 2004



Results: Average Transaction Duration



ORACLE

57

April 20, 2004



Considerations

ORACLE

58

April 20, 2004



Row Cache Considerations

- **Limits**
 - ~2,100,000,000 pages per snapshot storage area
 - ~2,100,000,000 total slots per cache
- **Reduced I/O and locking stalls can (greatly) reduce CPU idle time**
- **Objects stored in mixed-format areas aren't eligible for snapshots in cache in 7.1.2**
 - Sequential scans are problematic
 - Expect to relax restriction in the future
- **May need to increase GBLPAGES**
- **RMU/SHOW STATISTICS performance**
- **RCS checkpoint to database performance**

ORACLE

59

April 20, 2004



What it all Means

- **Snapshots in cache**
 - Potentially huge reduction in I/O for environments with snapshots enabled
- **64-bit Row Cache**
 - “Boundless” number of records in cache
 - Improved performance
- **Page Transfer Via Memory**
 - Avoid database page I/O with global buffers

ORACLE

60

April 20, 2004



Availability

- **Oracle Rdb Release 7.1.2**
 - Page Transfer Via Memory re-enabled
 - Initial release
 - Snapshots in Row Cache
 - Native 64-bit addressing support

ORACLE

61

April 20, 2004



What's Next?

ORACLE

62

April 20, 2004



Futures

- **Under consideration**
 - More RCS responsibilities
 - Sweeping snapshot rows to database
 - Sweeping “cold” modified rows to database
 - Improved/reduced latch usage
 - Lower contention for heavily accessed cache slots
 - Enhanced hashing algorithms
 - RCS increased buffer counts
 - RMU “/POPULATE_CACHE” command
 - Snapshots in cache for objects in mixed-format areas
 - Improved RMU /SHOW STATISTICS performance
 - Multi-threaded?

ORACLE

63

April 20, 2004



For More Information

- **OpenVMS**
 - Alpha Guide to 64-Bit Addressing & VLM Features
 - System Services Reference Manual
 - Calling Standard
 - Programming Concepts Manual
- **Alpha Architecture Handbook**
- **Oracle Rdb**
 - Release Notes
 - Guide to Performance and Tuning

ORACLE

64

April 20, 2004

COMMENTS ?
Q U E S T I O N S
&
A N S W E R S

ORACLE

ORACLE®

www.oracle.com/rdb

Philippe.vigier@oracle.com