



ORACLE®



Implementing Oracle Rdb Row Caches

Norman Lastovica
Oracle OpenVMS Development Team

13 October 2009

Agenda

- Row Cache Review
- Deciding What To Cache
- Configuring Caches
- Results





Row Cache Review

- Copies of database rows in memory
- Fetch/Modify cached rows with no database IO
- Locking & IO Reduction
 - Page locks
 - Doesn't help sequential scans – can even hurt
- Requires database open on single computer



Deciding What To Cache

- Often follows “80/20” rules of thumb
- Start with “Busy” Sorted indexes & “Hot” tables
- Snapshots in cache
- How much memory is available; how bad is IO and locking?



Snapshots In Cache

- Specify count of snapshot slots
- May be less or more than the cache slot count
- Start by evaluating relationship between storage area size and snapshot storage area size
- Performance penalty if too small



Checkpointing & Sweeping

- RCS checkpoint timer
 - Suggested to backing store with snapshots in cache, to database otherwise
- Per-process - evaluated at end of transaction
 - Fast Commit timer – 5 to 10 minutes?
 - Fast Commit transactions – 1000?
 - Fast Commit AIJ growth – 500,000 blocks?
- Optional cache sweep timer
 - Several minutes?



Checkpointing & Sweeping

- RMU /CHECKPOINT
 - Global checkpoint
- RMU/SERVER RECORD_CACHE CHECKPOINT
 - /WAIT
 - /LOG
- Checkpoint at database close, backup, and verify



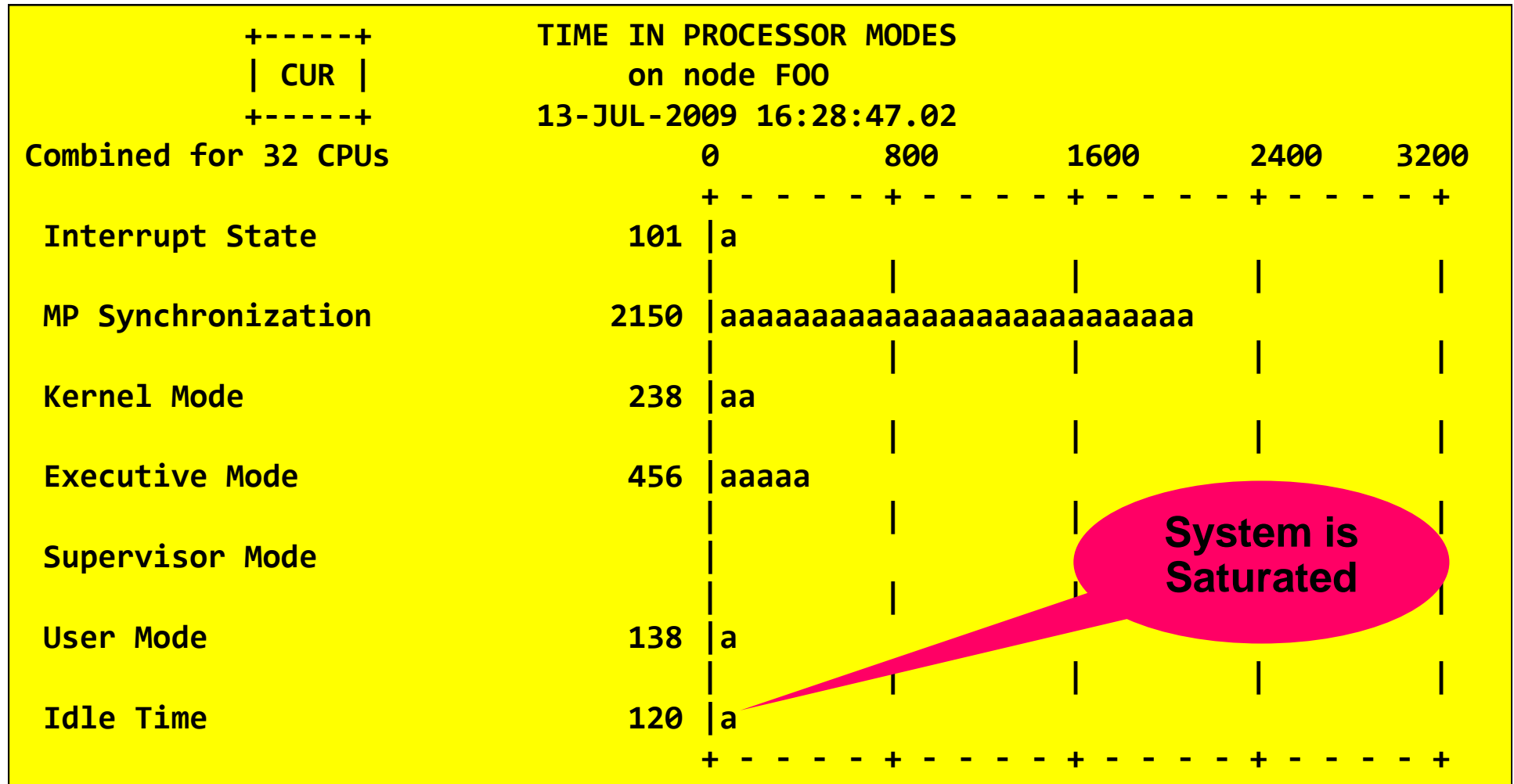
CASE STUDY



Case Study

- Customer production system
- “It is slow”
- “Backups are slow”
- Cluster of 164 8640 with high-end storage

21 of 32 CPUs in MPSYCH



175,000 Locking Operations Per Second

LOCK MANAGEMENT STATISTICS

on node F00

13-JUL-2009 16:41:10.72

	CUR	AVE	MIN	MAX
New ENQ Rate	73395.00	73395.00	73395.00	73395.00
Converted ENQ Rate	23417.00	23417.00	23417.00	23417.00
DEQ Rate	69951.66	69951.66	69951.66	69951.66
Blocking AST Rate	1103.00	1103.00	1103.00	1103.00
ENQs Forced To Wait Rate	846.66	846.66	846.66	846.66
ENQs Not Queued Rate	3723.66	3723.66	3723.66	3723.66
Deadlock Search Rate	0.00	0.00	0.00	0.00
Deadlock Find Rate	0.00	0.00	0.00	0.00
Total Locks	509849.00	509849.00	509849.00	509849.00
Total Resources	271984.00	271984.00	271984.00	271984.00



120 TPS

110 I/O Per Transaction

Node: F00 (1/1/2) Oracle Rdb V7.2-350 Perf. Monitor 15-JUL-2009 17:13:38.85
Rate: 3.00 Seconds Summary IO Statistics Elapsed: 00:41:48.89
Page: 1 of 1 DSA35:[000000.DATABASE]DB.RDB;4 Mode: Online

statistic.....	rate.per.second.....			total.....	average.....
name.....	max.....	cur.....	avg.....	count.....	per.trans....
transactions	785	82	117.7	295280	1.0
verb successes	257550	6984	12916.6	32401684	109.7
verb failures	1	0	0.0	125	0.0
synch data reads	129529	10710	9721.2	24385844	82.5
synch data writes	1229	42	127.2	319143	1.0
asynch data reads	46100	2749	3298.1	8273497	28.0
asynch data writes	1801	104	184.8	463779	1.5
RUI file reads	6	0	0.1	327	0.0
RUI file writes	155	14	16.4	41308	0.1
AUI file reads	215	0	1.2	3191	0.0
AUI file writes	133	11	16.9	42528	0.1
root file reads	147	0	0.3	902	0.0
root file writes	1551	154	222.7	558997	1.8

Top Indexes

Node: F00 (1/1/2) Oracle Rdb V7.2-350 Perf. Monitor 15-JUL-2009 17:12:07.19
Rate: 3.00 Seconds Logical Area Overview (Btree Indexes) Elapsed: 00:40:17.24
Page: 1 of 57 DSA35:[000000.DATABASE]DB;4 Mode: Online

Logical.Area.Name.....	leaf fetches	leaf inserti	leaf removal	discarded
XPK_NR_ACC_ANALYTICAL2	54724472	32	0	0
SI_NR_ACCOUNT_ANALYT_BANK	14281695	32	0	0
XPK_ORGSTRUCTURE	11775100	2	0	0
SI_KR_RADM_ROLE	10598169	0	0	0
SI_NR_ACC_ANALYT_NUM	5102910	32	0	0
XPK_KBK_ADM_RECEIVER	4884833	0	0	0
XPKCD_CLIENT_INFO	3644873	30	0	0
SI_ANK1_1	3626054	53	0	0
I_H_CRED	3344220	0	0	0
SI_NR_DOC_DATE2	3222071	6808	4587	0
SI_NR_DOC_COMPLEX_DATE	3087472	8442	4005	0
SI_NR_KRED_REQUEST_PHYS	2713654	3	0	0
SIR_ACCOUNT_ANALYTICAL_1	2468185	32	0	0



SETUP FOR CACHING



Configure Database

```
ALTER DATABASE FILENAME FOO  
    NUMBER OF CLUSTER NODES IS 1  
    RESERVE <n> CACHE SLOTS  
    ROW CACHE IS ENABLED  
    JOURNALING IS ENABLED  
    (FAST COMMIT IS ENABLED ...)
```




Configuring Checkpointing

```
ALTER DATABASE FILENAME FOO
    JOURNALING IS ENABLED
        (FAST COMMIT IS ENABLED
            (CHECKPOINT EVERY <t> TRANSACTIONS,
              CHECKPOINT TIMED EVERY <s> SECONDS,
              CHECKPOINT INTERVAL IS <b> BLOCKS))

ALTER DATABASE FILENAME MF_PERSONNEL
    ROW CACHE IS ENABLED (SWEEP INTERVAL IS <n> SECONDS);
```



Configuring Caches

- Slot size
- Slot count
- Snapshots in cache

Row / Node Size

```
$ RMU/SHOW AIP/BRIEF FOO SH_EMPLOYEE_ID
```

```
*-----  
* Logical Area Name          LArea PArea   Len Type  
*-----  
SH_EMPLOYEE_ID              58      1    430 UNKNOWN
```

Be careful of non-ranked indexes allowing duplicates



Index Node Count

```
$ RMU/ANALYZE/INDEX FOO
```

```
Index PLUGH for relation XYZZY duplicates allowed
```

```
Max Level: 8, Nodes: 9399698,
```

```
Used/Avail: 2237181111/3741079804 (59%),
```

```
Keys: 117076163,
```

```
Records: 107604684
```



Build a Cache Creation Spreadsheet

- One row per alter cache command
- Easy to cut-n-paste
- A fancy spreadsheet could do slot count analysis from the output of RMU/ANALYZE/INDEX, for example

Build a Cache Creation Spreadsheet

ALTER DATABASE FILE X\$ alter cache	SI_NR_ACCOUNT_ANALYT_BANK	CACHE SIZE IS	75000	rows	ROW LENGTH IS	430	BYTES CHECKPOINT UPDATI
ALTER DATABASE FILE X\$ alter cache	XPB_NR_ACC_ANALYTICAL2	CACHE SIZE IS	75000	rows	ROW LENGTH IS	430	BYTES CHECKPOINT UPDATI
ALTER DATABASE FILE X\$ alter cache	SIR_DOC_CARTOTEC_HIST_1	CACHE SIZE IS	30000	rows	ROW LENGTH IS	430	BYTES CHECKPOINT UPDATI
ALTER DATABASE FILE X\$ alter cache	XPB_ORGSTRUCTURE	CACHE SIZE IS	8193	rows	ROW LENGTH IS	430	BYTES CHECKPOINT UPDATI
ALTER DATABASE FILE X\$ alter cache	XPB_NR_DOC_COMPLEX1	CACHE SIZE IS	10421241	rows	ROW LENGTH IS	430	BYTES CHECKPOINT UPDATI
ALTER DATABASE FILE X\$ alter cache	SI_NR_ACC_ANALYT_NUM	CACHE SIZE IS	75000	rows	ROW LENGTH IS	430	BYTES CHECKPOINT UPDATI
ALTER DATABASE FILE X\$ alter cache	SIR_ACCOUNT_ANALYTICAL_1	CACHE SIZE IS	12000	rows	ROW LENGTH IS	430	BYTES CHECKPOINT UPDATI
ALTER DATABASE FILE X\$ alter cache	SI_NR_DOC_DATE2	CACHE SIZE IS	7670707	rows	ROW LENGTH IS	430	BYTES CHECKPOINT UPDATI
ALTER DATABASE FILE X\$ alter cache	XPB_NR_DOC	CACHE SIZE IS	3097377	rows	ROW LENGTH IS	430	BYTES CHECKPOINT UPDATI
ALTER DATABASE FILE X\$ alter cache	I_R_LS_NUMBER_2	CACHE SIZE IS	600000	rows	ROW LENGTH IS	430	BYTES CHECKPOINT UPDATI
ALTER DATABASE FILE X\$ alter cache	I_R_LS_NUMBER_1	CACHE SIZE IS	200000	rows	ROW LENGTH IS	430	BYTES CHECKPOINT UPDATI
ALTER DATABASE FILE X\$ alter cache	SI_LS_2	CACHE SIZE IS	300000	rows	ROW LENGTH IS	430	BYTES CHECKPOINT UPDATI
ALTER DATABASE FILE X\$ alter cache	SI_KR_RADM_ROLE	CACHE SIZE IS	200	rows	ROW LENGTH IS	430	BYTES CHECKPOINT UPDATI
ALTER DATABASE FILE X\$ alter cache	SI_NR_KRED_REQUEST_PHYS	CACHE SIZE IS	5000	rows	ROW LENGTH IS	430	BYTES CHECKPOINT UPDATI
ALTER DATABASE FILE X\$ alter cache	SI_ANK1_1	CACHE SIZE IS	75000	rows	ROW LENGTH IS	430	BYTES CHECKPOINT UPDATI
ALTER DATABASE FILE X\$ alter cache	SI_ILS_1	CACHE SIZE IS	8324497	rows	ROW LENGTH IS	430	BYTES CHECKPOINT UPDATI
ALTER DATABASE FILE X\$ alter cache	KR_DOC_COMPLEX_CLSB_DC	CACHE SIZE IS	150000	rows	ROW LENGTH IS	430	BYTES CHECKPOINT UPDATI
ALTER DATABASE FILE X\$ alter cache	SI_R_ZCH_TOPSPIS_PRIM	CACHE SIZE IS	5000	rows	ROW LENGTH IS	430	BYTES CHECKPOINT UPDATI
ALTER DATABASE FILE X\$ alter cache	SI_NR_DOC_COMPLEX	CACHE SIZE IS	20934080	rows	ROW LENGTH IS	430	BYTES CHECKPOINT UPDATI
ALTER DATABASE FILE X\$ alter cache	SI_NR_BANDEROL_K8	CACHE SIZE IS	130000	rows	ROW LENGTH IS	430	BYTES CHECKPOINT UPDATI



Create Cache

```
ALTER DATA FILE F00
  ADD cache PLUGH
    cache size is 9999999 rows
    row length is 430 bytes
    shared memory is PROCESS RESIDENT
    checkpoint updated rows to database
    row snapshot is ENABLED
      (cache size is 99999 rows)
```



How Much To Cache

- Use your memory
- Points of diminishing return
- RMU /DUMP /HEADER
 - “Shared memory section requirement”



Ease of Use

- RMU /SET ROW_CACHE
 - /ENABLE
 - /DISABLE
- RMU /CLOSE /WAIT
 - If required, use /ABORT=FORCEX of user processes
- RMU /OPEN /WAIT
- RMU /CHECKPOINT



RESULTS



Results

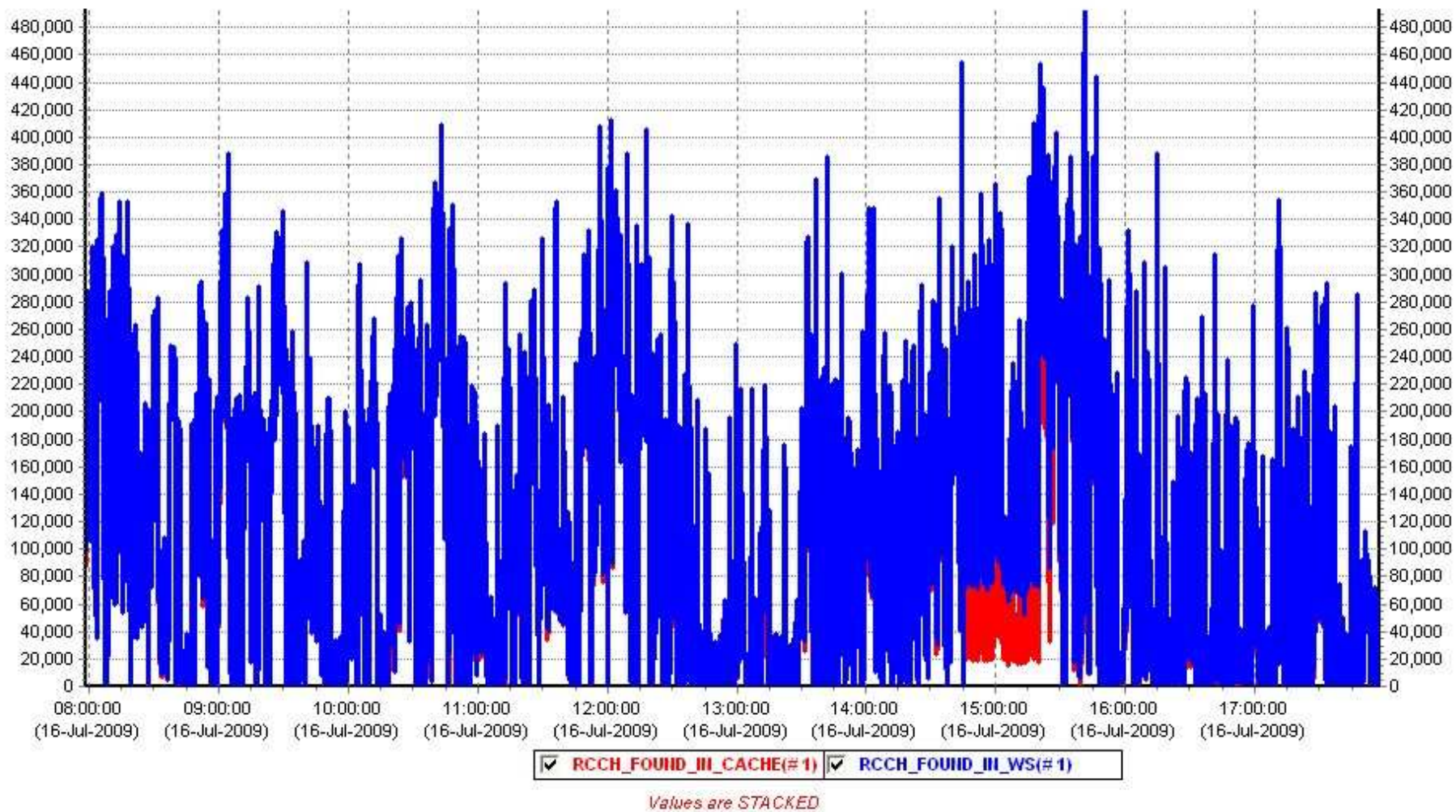
- Look at both “whole day” and “interval” statistics
- There is always more tuning possible

Excellent Hit Rates & A Few Problems

Node: F00 (1/1/1) Oracle Rdb V7.2-350 Perf. Monitor 17-JUL-2009 10:16:31.18
Rate: 3.00 Seconds Row Cache Overview (Unsorted) Elapsed: 02:50:21.66
Page: 1 of 1 DSA35:[VOX_DB]VOX.RDB;1 Mode: Online

Cache.Name.....	#Searches	Hit%	Full%	#Inserts	#Wrap	#Slots	Len
XPK_NR_ACC_AN2	250887097	99.9	72.3	54276	0	75000	432
SI_NR_ACC_ANALYT_NUM	58139058	99.9	61.0	45777	0	75000	432
XPK_KBK_ADM_REC	9841561	99.9	55.5	285	0	513	432
SI_NR_ACC_AN_BAN	768137386	99.9	76.1	57100	0	75000	432
SI_NR_ACC_AN	2075325	99.8	33.9	2782	0	8193	432
SI_NR_ACC_AN_GRO	5536387	99.0	73.6	55227	0	75000	432
SI_NR_RED_REQ	20579202	99.9	41.1	2058	0	5000	432
XPK_ORGST	57380427	99.9	60.0	4920	0	8193	432
SI_KR_RADM	24873333	99.9	37.5	75	0	200	432
XPKCD_INFO	2263936	99.8	87.5	3585	0	4097	432
SI_ANK1_3	2266	74.8	0.0	570	0	600000	432
SI_ANK1_1	12551786	99.5	70.3	52755	0	75000	432
SI_CR_CB	1000646	99.7	71.0	2910	0	4097	432
SI_OSTPH	106262174	0.0	100.0	106234285	47	3	432

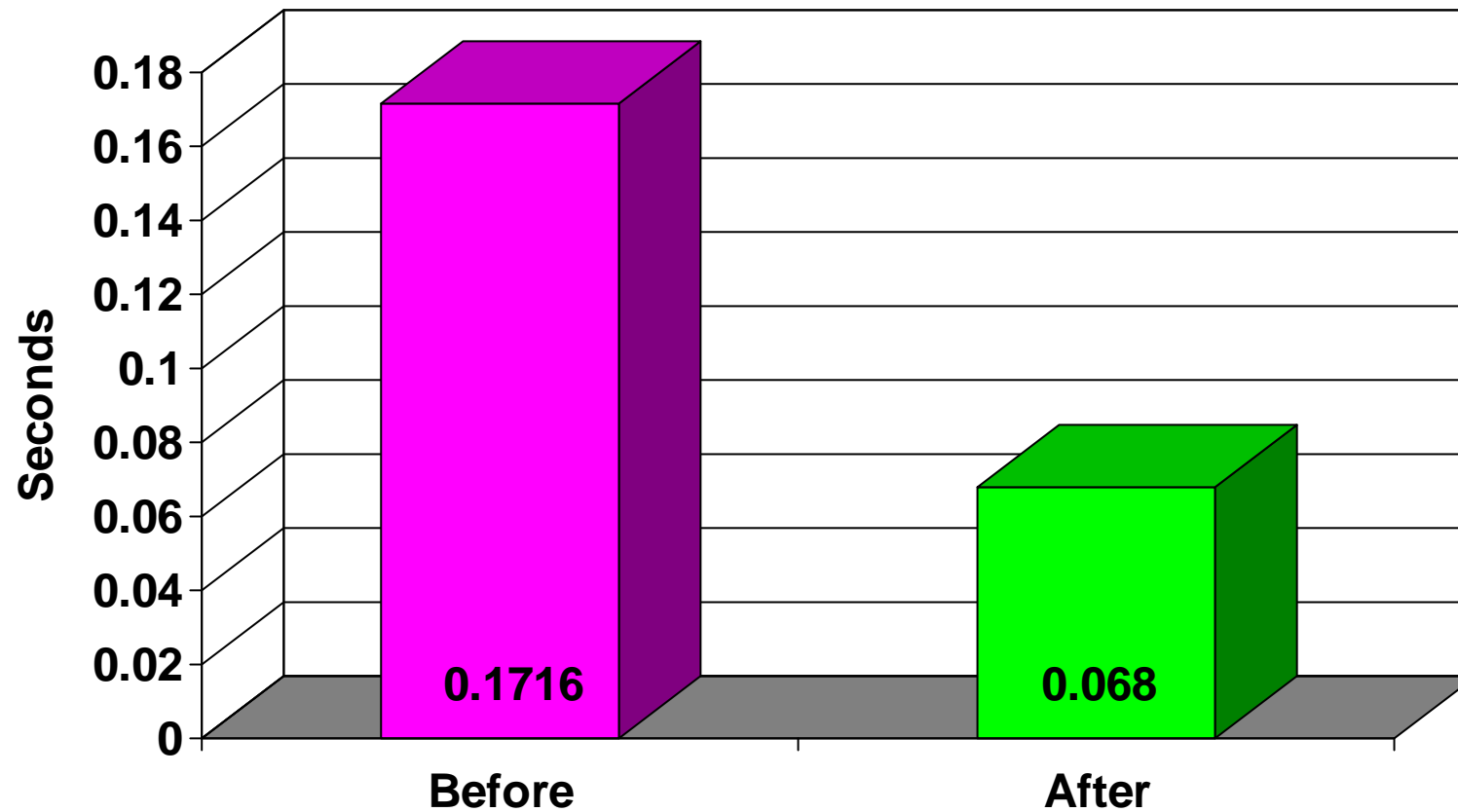
480,000+ Cache Hits Per Second from Working Set



21 Idle CPUs

+-----+ CUR +-----+		TIME IN PROCESSOR MODES on node F00 16-JUL-2009 13:10:10.23				
Combined for 32 CPUs		0	800	1600	2400	3200
		+ - - - - +	+ - - - - +	+ - - - - +	+ - - - - +	+ - - - - +
Interrupt State	17					
MP Synchronization	6					
Kernel Mode	432	aaaaa				
Executive Mode	434	aaaaa				
Supervisor Mode						
User Mode	149	a				
Idle Time	2161	aaaaaaaaaaaaaaaaaaaaaaaaaaaaa				
		+ - - - - +	+ - - - - +	+ - - - - +	+ - - - - +	+ - - - - +

Average Transaction Duration





Test, Analyze, Change, Repeat

- Get the latest VMS patches
- Get the latest Rdb releases
- Use the dedicated CPU lock manager if needed



Evaluate

- See how the caches are doing
- Do not look at every problem as a nail

Node: EPROD1 (1/1/1) Oracle Rdb V7.2-350 Perf. Monitor 7-OCT-2009 11:43:28.60
Rate: 6.00 Seconds Row Cache Status Elapsed: 3 11:31:05.67
Page: 1 of 1 POS_DB_01: [RDB.DATA]ACE_POS_DB.RDB;1 Mode: Online

For Cache: ACE_PAYMENT_DTL_RIDX2

Statistic.Name	Stat.Value	Percent	MRES
Total slots:	100000	100.0%	Slot Length: 3012 Hash slots: 131072
Slots full:	99813	99.8%	Use: 22 99.8%
Slots empty:	187	0.2%	Rsv: 69 36.8%
Marked Slots:	17	0.0%	Hot: 17 100.0% Cold: 0 0.0%
Clean Slots:	99983	99.9%	Hot: 0 0.0% Cold: 99983 100.0%
Used Space:	300636k	99.8%	Wstd: 0k 0.0% Free: 563k 0.2%
Hash Que Lengths:	Empty:31347	1:99637	2:88 3:0 4+:0
Cursor position:	10827	of 100000	wrapped 45 times
Cache latched:	No		
Cache is full:	No	Cache modified:	Yes Snapshot is full: No
Checkpoints:	1031	Last: 7-OCT-2009 11:40:58.97	AIJ Location: 3560:2030190
Cache Recovery:	3560:2153766		
Snap Slots:	30000	100.0%	Ful: 29848 99.4% Rcl: 29328 97.7%
Snap Cursor:	15549	of 30000	(slot 115549) wrapped 8 times

Node: EPROD1 (1/1/1) Oracle Rdb V7.2-350 Perf. Monitor 7-OCT-2009 11:44:24.32
 Rate: 6.00 Seconds Summary Cache Statistics Elapsed: 3 11:32:01.3
 Page: 1 of 1 POS_DB_01:[RDB.DATA]ACE_POS_DB.RDB;1 Mode: Online

statistic..... name.....	rate.per.second.....			total..... count.....	average..... per.trans....
	max.....	cur.....	avg.....		
latch requests	65	28	16.6	5000207	0.0
retrieved	0	0	0.0	5834	0.0
cache searches	13916	9446	6761.0	2033197063	25.1
found in workset	6742	4479	2090.3	628623621	7.7
found in cache	7090	4566	2677.3	805142781	9.9
found too big	1	0	1.9	593214	0.0
insert cache	1094	402	224.3	67469000	0.8
row too big	0	0	0.0	0	0.0
cache full	0	0	0.2	77551	0.0
collision	2	1	2.8	842473	0.0
skipped dirty slot	5	0	0.0	16142	0.0
skipped inuse slot	1	0	1.2	368628	0.0
hash misses	5684	2762	2161.6	650047944	8.0
cache unmark	164	38	47.4	14275770	0.1
snapshot search	74	0	1.9	589897	0.0
snap found cache	74	0	1.9	587164	0.0
snapshot store	220	87	51.6	15532085	0.1

Node: EPROD1 (1/1/1) Oracle Rdb V7.2-350 Perf. Monitor 7-OCT-2009 11:44:59.57
 Rate: 6.00 Seconds Row Cache Overview (%Hit) Elapsed: 3 11:32:36.64
 Page: 1 of 3 POS_DB_01:[RDB.DATA]ACE_POS_DB.RDB;1 Mode: Online

Cache.Name	#Searches	Hit%	Full%	#Inserts	#Wrap	#Slots	Len
SSQ_QUEUE_RIDX1	68251967	100.0	4.5	45	0	1000	3012
ACE_ITEM_DISCOUNT_RIDX1	74936584	99.9	2.9	897	0	30000	3012
SSQ_HISTORY_RIDX_PK	1377019	99.9	3.5	724	0	20000	3012
SSQ_QUEUE_RIDX2	634506	99.9	9.7	98	0	1000	3012
CO_COMPLETED_ABS_TIME_RI	54638428	99.9	3.0	926	0	30000	3012
CO_APPR_TEL_NUM_RIDX	69884636	99.9	97.4	42179	1	30000	3012
ACE_REFUND_ORDER_RIDX1	18426348	99.9	4.8	3421	0	70000	3012
CREDIT_SEQ_NUM_DTL_RIDX_	923900	99.9	0.6	17	0	3000	3012
CREDIT_SEQ_NUM_DTL	134150	99.9	0.3	34	0	10000	40
CREDIT_SEQ_NUM_RIDX_PK	134160	99.9	9.0	9	0	100	3012
CREDIT_SEQ_NUM	402484	99.9	33.0	33	0	100	72
CUST_ORD_VISION_RIDX1	71444210	99.9	56.3	45101	0	80000	3012
SSQ_QUEUE_RIDX_PK	1144796	99.9	14.8	282	0	1000	3012
ACE_ITEM_TRAN_RIDX1	142873292	99.9	63.7	57440	0	90000	3012
ACE_FFL_TRACKING_RIDX1	3645456	99.8	18.7	4625	0	20000	3012
ACE_FFL_TRACKING_RIDX_PK	4787048	99.8	36.8	7530	0	20000	3012
CSFO_RIDX_PK	84582447	99.8	98.1	88673	0	90000	3012
ACE_PAYMENT_HDR_RIDX_PK	16438906	99.8	35.4	31939	0	90000	3012
ACE_ITEM_DISCOUNT_RIDX_P	33773380	99.7	98.4	79177	2	30000	3012
ACE_REMARKS_RIDX	17834721	99.7	42.7	38706	0	90000	3012
ACE_PAY_MAP_RIDX_PK	12391147	99.7	36.9	33227	0	90000	3012
ACE_PAYMENT_DTL_RIDX1	2735547	99.7	7.6	6886	0	90000	3012
ACE_PAYMENT_DTL_RIDX_PK	17781946	99.7	35.3	35763	0	100000	3012
CUSTOMER_ORDER_ENC_RIDX1	5627070	99.6	44.8	22456	0	50000	3012
ACE_PAYMENT_DTL_ENC_RIDX	10903467	99.6	66.5	33298	0	50000	3012
ACE_FFL_TRACKING_RIDX3	895504	99.6	14.3	2874	0	20000	3012
ACE_VISION_PAY_DTL_RIDX_	530015	99.6	10.0	2014	0	20000	3012
ACE_ORDERS_RIDX1	4138717	99.6	17.4	13967	0	80000	3012
CCA_RIDX_PK	37931160	99.6	97.2	125687	1	90000	3012
ACE_PAYMENT_HDR_RIDX2	10527336	99.5	67.3	47125	0	70000	3012
CREDIT_DECISION_RIDX_PK	112072784	99.5	97.8	529893	7	75000	3012
ACE_REFUND_ORDER_RIDX	724660	99.5	3.6	3329	0	90000	3012
ACE_CUSTOMER_ENC_RIDX1	832974	99.4	8.4	4240	0	50000	3012

Row Cache Status

```
Rate: 3.00 Seconds          Row Cache Status          Elapsed: 00:02:02.03
Page: 1 of 1                $1$DGA203:[LASTOVICA.V73]MF_PERSONNEL.RDB;1      Mode: Online
-----
                                For Cache: C1
Statistic.Name Stat.Value Percent

Total slots:          1000  100.0% Slot Length: 16  Hash slots: 1024
Slots full:           0     0.0% Use:           0     0.0%
Slots empty:         1000  100.0% Rsv:           0     0.0%
Marked Slots:         0     0.0% Hot:            0     0.0% Cold:           0     0.0%
Clean Slots:         1000  100.0% Hot:            0     0.0% Cold:        1000  100.0%
Used Space:           0k    0.0% Wstd:           0k    0.0% Free:        16k  100.0%
Hash Que Lengths: Empty:1024 1:0           2:0           3:0           4+:0
Cursor position:      0 of 1000 wrapped 0 times
Cache latched:        No
Cache is full:        No           Cache modified:      No   Snapshot is full: No
Checkpoints: 2 Last:  6-OCT-2009 18:40:05.94 AIJ Location: 1:2
Cache Recovery:       1:2
Snap Slots:           5000  100.0% Ful:           0     0.0% Rcl:           0     0.0%
Snap Cursor: 0 of 5000 (slot 1000) wrapped 0 times
```


Row Cache Summary

```
Rate: 3.00 Seconds                      Row Cache (C1)                      Elapsed: 00:01:04.16
Page: 1 of 1                          $1$DGA203:[LASTOVICA.V73]MF_PERSONNEL.RDB;1      Mode: Online
-----
statistic..... rate.per.second..... total..... average.....
name..... max..... cur..... avg..... count..... per.trans....
latch requests          0          0          0.0          0          0.0
  retried                0          0          0.0          0          0.0
cache searches          0          0          0.0          0          0.0
  found in workset       0          0          0.0          0          0.0
  found in cache         0          0          0.0          0          0.0
  found too big          0          0          0.0          0          0.0
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
skipped dirty slot      0          0          0.0          0          0.0
skipped inuse slot      0          0          0.0          0          0.0
hash misses            0          0          0.0          0          0.0
cache unmark            0          0          0.0          0          0.0
snapshot search         0          0          0.0          0          0.0
  snap found cache       0          0          0.0          0          0.0
snapshot store          0          0          0.0          0          0.0
```




Final Thoughts

- Row Cache & Hot Standby database
- Locking & IO reduction
- CPU consumption increases due to reduced waiting
- Re-open after “node failure”



For More Information

search.oracle.com



or

[**oracle.com/rdb**](https://oracle.com/rdb)



ORACLE IS THE INFORMATION COMPANY