

The middle ground in the Buffer Cache

Lothar Flatz

Senior Principal Consultant

May I introduce myself ...

Who am I?

- 25 Year Oracle Database experience (starting with Version 5)
- 15 Years Oracle Employee
- Oak Table Member
- Ex-Real World Performance Group
- Oracle ACE
- Signatur Project: PVSS (CERN)
- US 8103658 B2 patent with Björn Engsig



bbi Software AG

- Founded 1994
- Legal privately owned incorporated company
- Resides Baar ZG und Oberdorf SO
- Employees 10
- Certification SQS certified ISO 9001
- Software acquisition, License management, Software assembling, Consulting, Enterprise Reporting



How this presentation is organized

- Activity Snapshot of the database segments
- Addressing specific issues
- Segments that cover the middle ground
- Demo
- Buffer cache temperature in 12c
- Conclusions

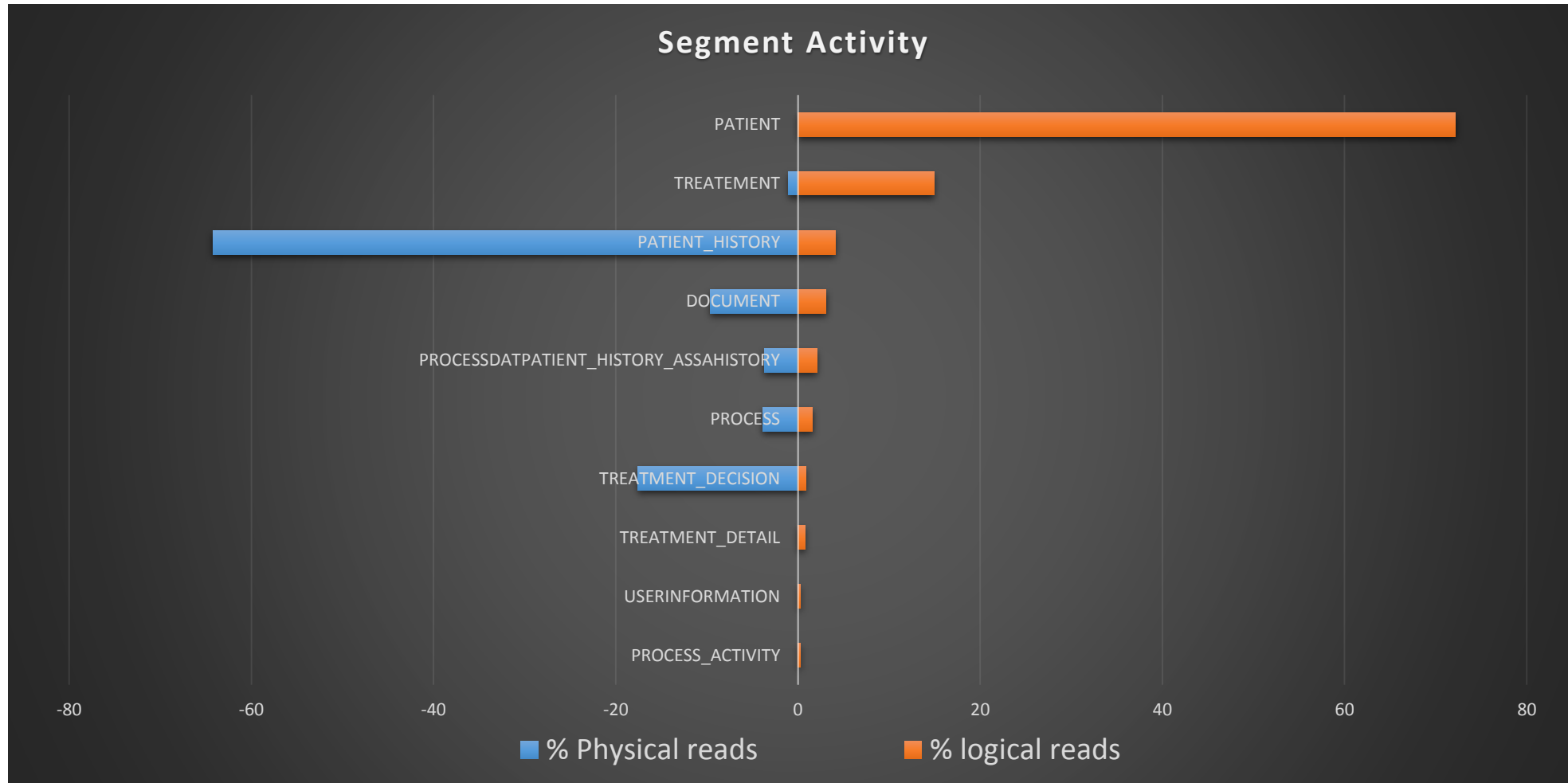
Activity Snapshot of the database segments

```
SELECT OBJECT_NAME, OBJECT_TYPE,  
       "logical reads",  
       "physical reads",  
       ROUND(ratio_to_report("physical reads") over () * 100,2) "% physical reads",  
       ROUND(ratio_to_report("logical reads") over () * 100,2) "% buffergets"  
FROM  
  (SELECT OBJECT_NAME, OBJECT_TYPE,  
         SUM(DECODE(STATISTIC_NAME,'logical reads', value ,NULL)) "logical reads",  
         SUM(DECODE(STATISTIC_NAME,'physical reads', value ,NULL)) "physical reads"  
    FROM v$segment_statistics  
   WHERE owner          =user  
         AND STATISTIC_NAME IN ('logical reads','physical reads')  
   GROUP BY OBJECT_NAME, OBJECT_TYPE  
         ORDER BY "logical reads" DESC  
  )  
WHERE rownum < 11  
/
```

Walk in Clinic - Activity Snapshot of the database segments

Table	logical reads	physical reads	% physical reads	% logical reads
PATIENT	11 027 020 480	6	0	72,16
TREATMENT	2 280 336 592	8 438	1,02	14,92
PATIENT_HISTORY	621 398 800	529 544	64,22	4,07
DOCUMENT	471 252 336	79 163	9,6	3,08
PATIENT_HISTORY_ASS	311 690 656	30 524	3,7	2,04
PROCESS	242 127 360	32 042	3,89	1,58
TREATMENT_DECISION	128 403 280	144 862	17,57	0,84
TEATEMENT_DETAIL	115 552 080	0	0	0,76
USERINFORMATION	42 014 800	0	0	0,27
PROCESS_ACTIVITY	40 782 768	0	0	0,27

Activity Snapshot of the database segments



Middle ground in the Buffer Cache

Before we start tuning



Let us tidy up the buffer cache

Buffer cache care

1. Eliminate damaging SQL
2. Consider clustering
3. Increase if it makes sense

Issue: Patient

- 72% of all Activity on one Table
- => reasonable?
- Shall we look into it?

Middle ground in the Buffer Cache

Issue: Patient – naïve approach

```
SELECT sql_id,  
       executions,  
       buffer_gets  
FROM gv$sql  
WHERE sql_id IN  
      (SELECT sql_id  
       FROM gv$sql_plan  
       WHERE object_name IN ('PATIENT') )  
order by buffer_gets desc
```

Issue Patient

Sql Id	Executions	physical reads
gd5462fjsy1z7	111 337	976 742 521,00
7cfpujxhh9m16	1 257	143 862 516,00
37dpg2tx3f58t	960	110 264 215,00
902jtma0ks90p	284	82 092 156,00
fr3ysnxvkvz8k	54 719	81 029 893,00
3rvjsc4cgbtgx	1 679	52 579 288,00
gw289sq96axxa	993	33 149 899,00
aqx1b1gphkbbk	93	29 641 220,00
1bfxu2a5w32y1	3 038	25 245 020,00

Issue Patient Sql Id gd5462fjsy1z7

Id	Operation	Name	Starts	E-Rows	A-Rows	A-Time	Buffers
0	SELECT STATEMENT		1		1	00:00:00.81	21157
1	NESTED LOOPS		1	25	1	00:00:00.81	21157
2	NESTED LOOPS		1	25	4	00:00:00.81	21153
* 3	TABLE ACCESS FULL	PATIENT	1	25	4	00:00:00.81	21143
* 4	INDEX UNIQUE SCAN	TRTMT_PKC	4	1	4	00:00:00.01	10
* 5	TABLE ACCESS BY INDEX ROWID	TREATMENT	4	1	1	00:00:00.01	4

Predicate Information (identified by operation id):

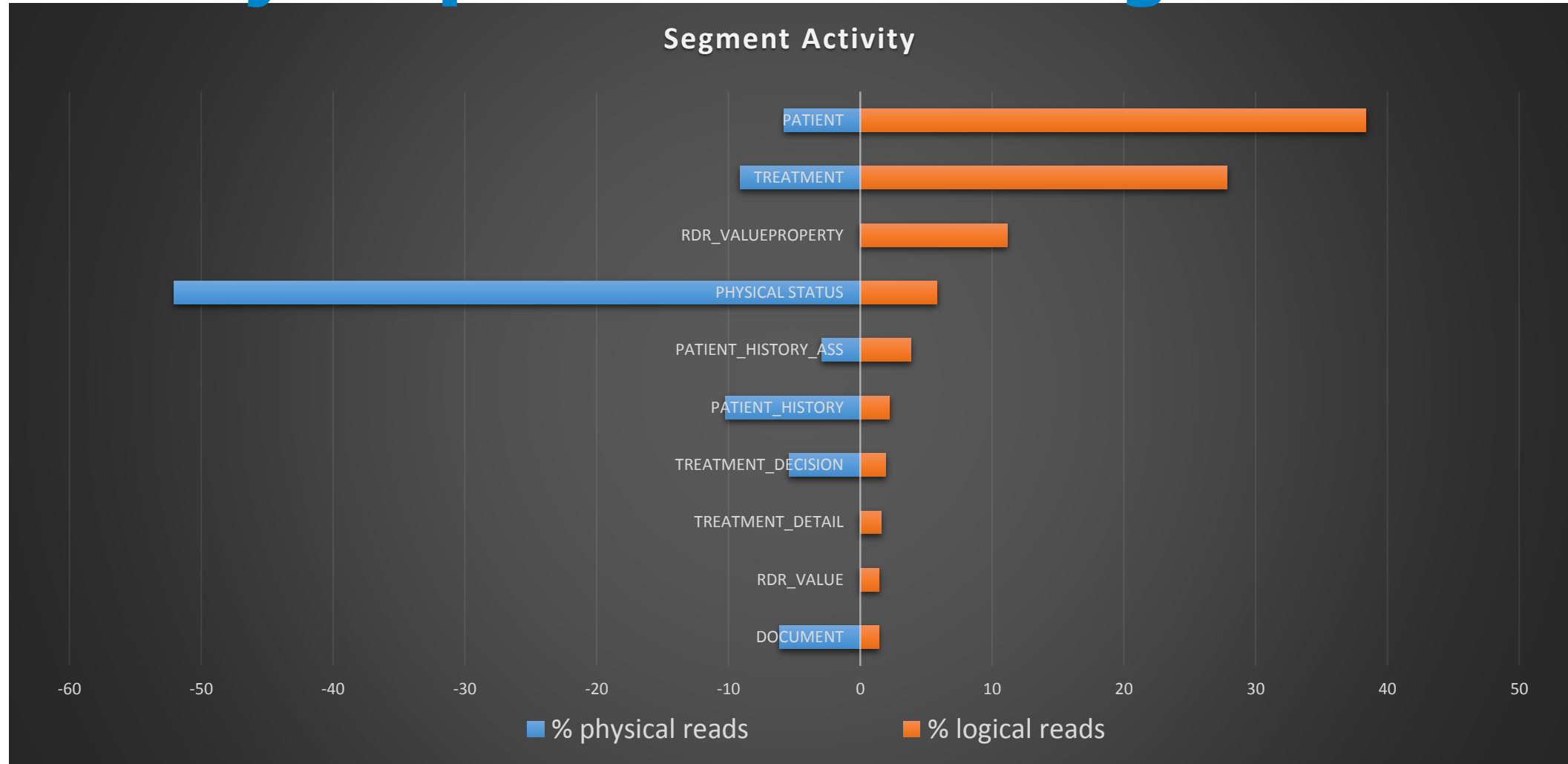
```

3 - filter((LOWER("INSURED1"."FIRSTNAME") LIKE :B1 AND LOWER("INSURED1"."LASTNAME")
          LIKE :B2 AND TRUNC(INTERNAL_FUNCTION("INSURED1"."DATEOFBIRTH"))=TRUNC(TO_DATE(:B3,'MM/DD/YY
          YY HH24:MI:SS'))))
4 - access("CLAIM0"."CLAIM_PK"="INSURED1"."INSURED_PK")
5 - filter(("CLAIM0"."REGION_FK"=:B5 AND "CLAIM0"."SERIALID"=:B4))
    
```

Activity Snapshot after cleaning & Release

Table	logical reads	physical reads	% physical reads	% logical reads
PATIENT	1 668 152 816	2 216 393	6	38
TREATMENT	1 210 539 568	3 501 906	9	28
RDR_VALUEPROPERTY	484 608 144	55	0	11
PHYSICAL STATUS	253 261 616	19 978 864	52	6
PATIENT_HISTORY_ASS	167 372 144	1 124 570	3	4
PATIENT_HISTORY	96 933 184	3 924 844	10	2
TREATMENT_DECISION	84 569 488	2 074 591	5	2
TREATMENT_DETAIL	69 611 904	0	0	2
RDR_VALUE	62 904 864	108	0	1
DOCUMENT	61 246 944	2 339 419	6	1

Activity Snapshot after cleaning & Release



Issue: Patient_history

- A classic
- In a sanatorium you would need to cluster by patient id
- In the ambulance we have on average 1-3 entries per patient, max 16

Issue: Patient_history Sanatorium

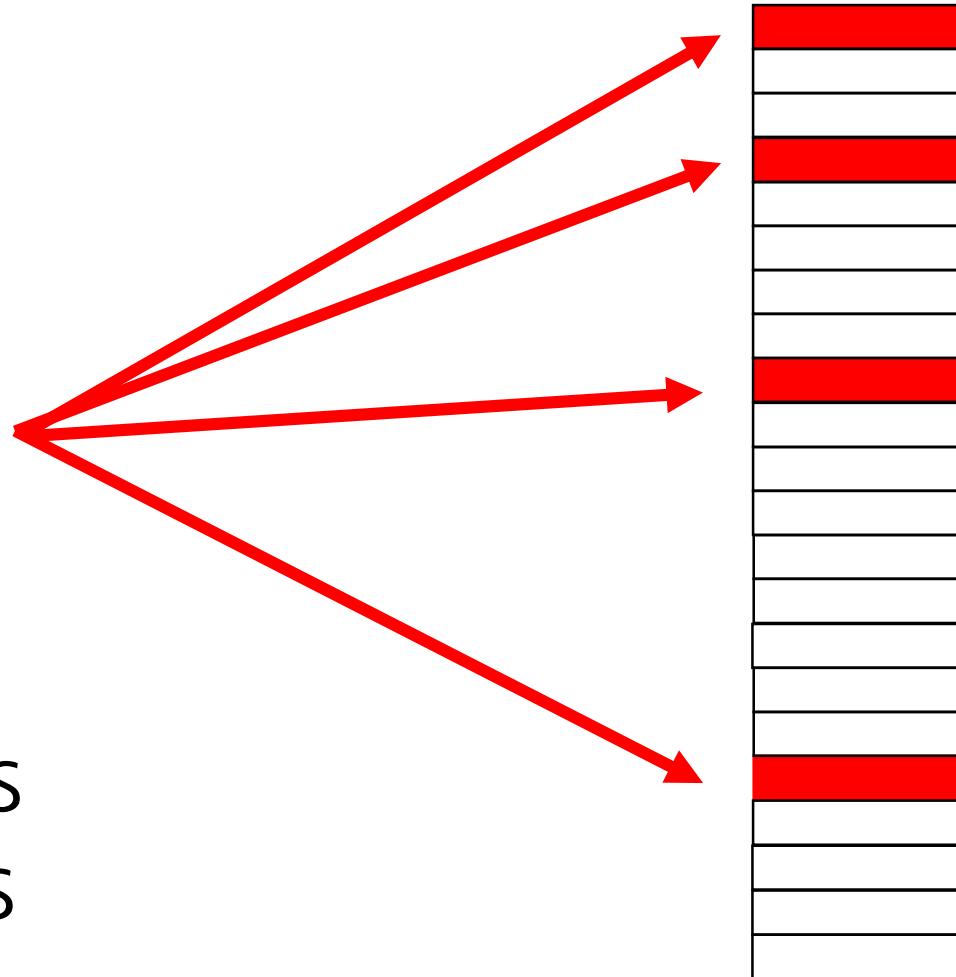
```
SELECT max(cnt), avg(cnt)
FROM (SELECT /*+ Parallel(8) */ pat_id,
      COUNT(*) cnt
      FROM Patient_history
      GROUP BY pat_id
      )
```

MAX (CNT)	AVG (CNT)
10419	77.7088

Issue: Patient_history Sanatorium

Patient id

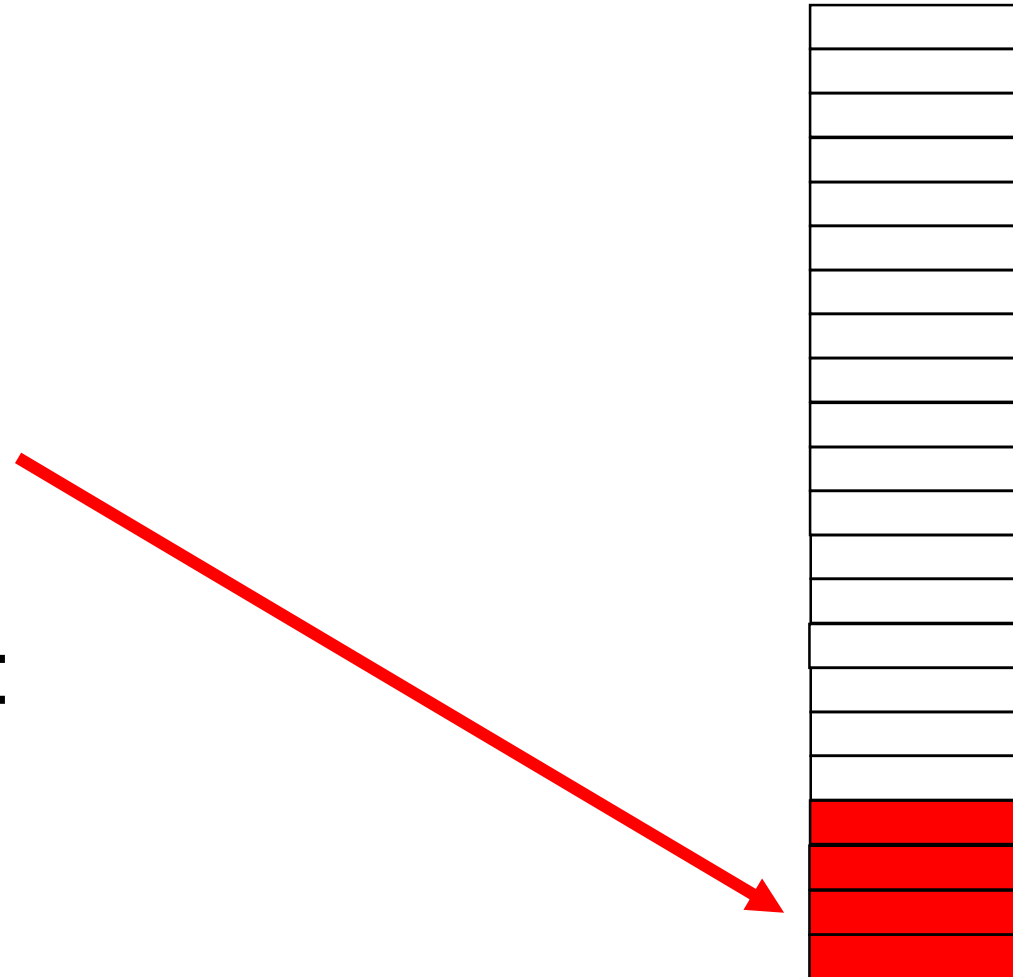
Rows are scattered across many db blocks



Issue: Patient_history Sanatorium

Patient id

Consequence:
Cluster (IOT,
Index Cluster,
Hash Cluster)



Issue: Patient_history Walk in Clinic

```
SELECT AVG(cnt),      AVG(blocks),
       STDDEV(cnt),  STDDEV(blocks)
FROM (SELECT pat_id, count(*) cnt,
          COUNT( distinct
                DBMS_ROWID.ROWID_BLOCK_NUMBER(h.rowid)
          ) blocks
      FROM patient_history H
      GROUP BY pat_id
    )
```

AVG (CNT)	AVG (BLOCKS)	STDDEV (CNT)	STDDEV (BLOCKS)
2.13331786	1.3981599	.458027726	.581603838

Conclusion: Patient_history Walk in Clinic

- Not resolvable by clustering
- Why is it not efficient?

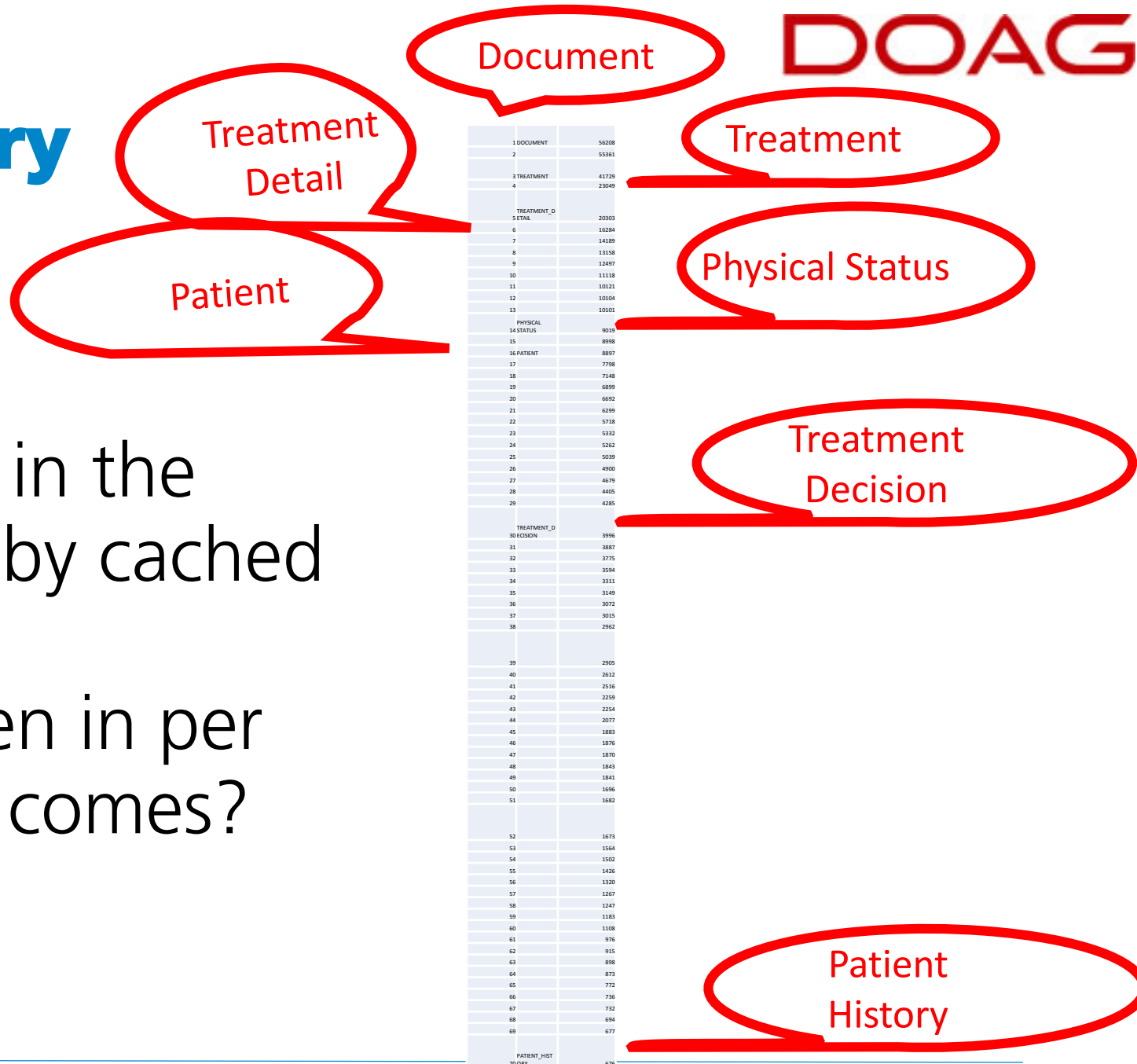


Let us check out the cache

Middle ground in the Buffer Cache

Issue: Patient_history

- A bigger Table
- Only on position 70 in the buffer cache sorted by cached blocks
- But under the top ten in per Table activity – how comes?



Conclusion: : Patient_history Walk in Clinic



- In the top ten segments by activity
- Segment No 70 in cache
- What goes in between and why?

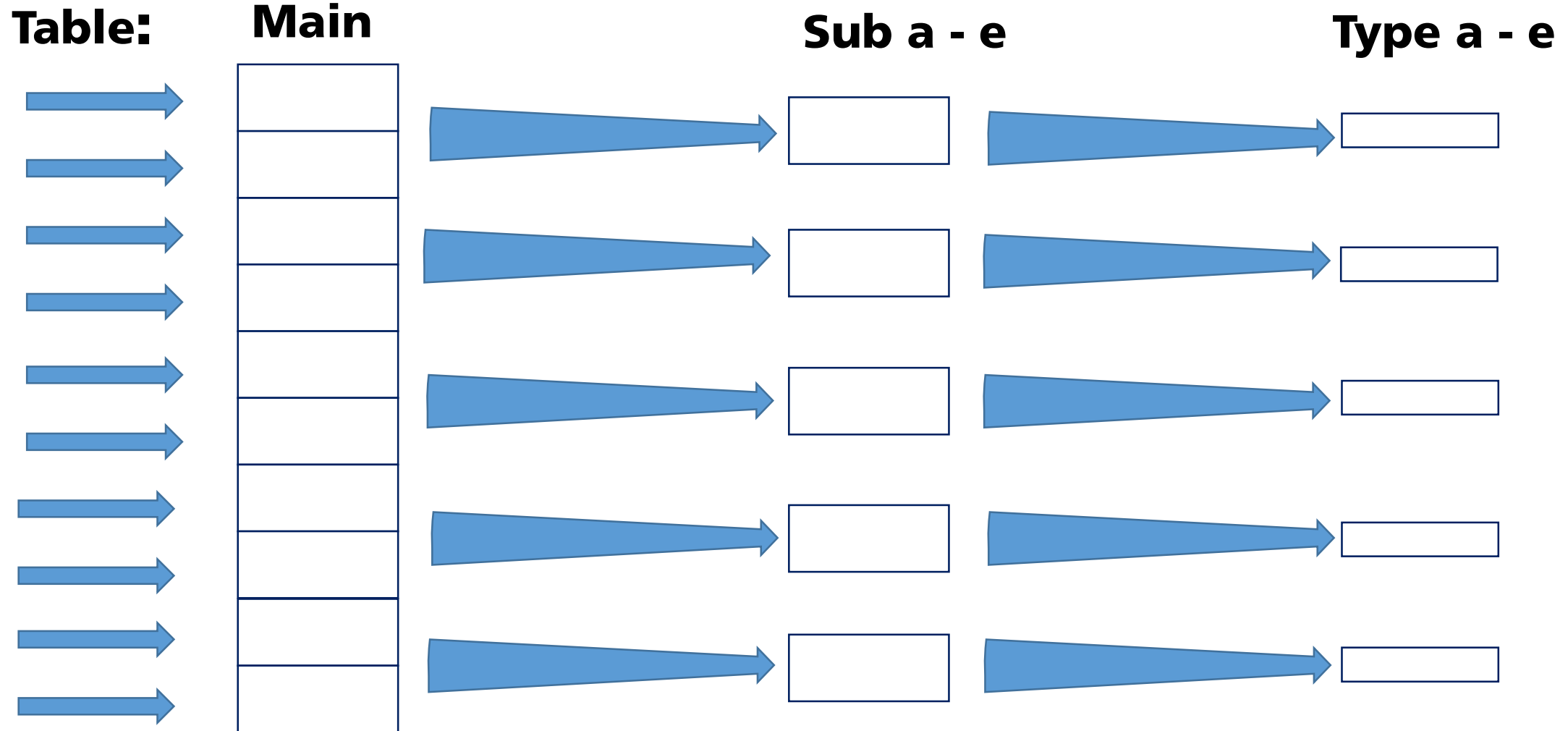
Middle ground in the Buffer Cache

Patient_history Walk in Clinic



Let us set up an experiment !

Experiment set up



Middle ground in the Buffer Cache

Cache content – 2 GB BC

OBJECT_NAME	OBJECT_TYPE	CACHED_BLOCKS	ALL_BLOCKS	% cached	cached MB
MAIN_IDX1	INDEX	28263	28672	98,57	221
MAIN_IDXA	INDEX	5640	5760	97,92	44
TYPE_IDXA	INDEX	5639	5760	97,90	44
TYPE_D	TABLE	7163	7424	96,48	56
TYPE_C	TABLE	7160	7424	96,44	56
TYPE_E	TABLE	7153	7424	96,35	56
TYPE_A	TABLE	7146	7424	96,26	56
TYPE_B	TABLE	7144	7424	96,23	56
TYPE_IDXD	INDEX	5659	5888	96,11	44
MAIN_IDXB	INDEX	5658	5888	96,09	44
MAIN_IDXD	INDEX	5654	5888	96,03	44
MAIN_IDXE	INDEX	5652	5888	95,99	44
MAIN_IDXC	INDEX	5646	5888	95,89	44
TYPE_IDXC	INDEX	5644	5888	95,86	44
TYPE_IDXB	INDEX	7716	8064	95,68	60
TYPE_IDXE	INDEX	7680	8064	95,24	60
SUB_A	TABLE	15175	31744	47,80	119
SUB_D	TABLE	15164	31744	47,77	118
SUB_B	TABLE	15128	31744	47,66	118
SUB_C	TABLE	15036	31744	47,37	117
SUB_E	TABLE	15031	31744	47,35	117
AFTER	TABLE	3	8	37,50	0
MAIN	TABLE	75140	405376	18,54	587
TRACK	TABLE	149	21504	,69	1

Middle ground in the Buffer Cache

Activity – 2 GB BC

OBJECT_NAME	logical reads	physical reads	% physical reads	% buffergets
MAIN_IDX1	3001296	2044	,15	24,97
MAIN	999840	810036	60,67	8,32
TYPE_IDXC	604448	393	,03	5,03
MAIN_IDXD	603104	398	,03	5,02
MAIN_IDXC	602336	389	,03	5,01
MAIN_IDXA	602288	416	,03	5,01
TYPE_IDXE	600784	4159	,31	5
MAIN_IDXB	600496	436	,03	5
TYPE_IDXD	600496	402	,03	5
TYPE_IDXA	598896	414	,03	4,98
MAIN_IDXE	598128	441	,03	4,98
TYPE_IDXB	597312	4047	,30	4,97
SUB_A	201536	99946	7,49	1,68
TYPE_A	201360	1905	,14	1,68
TYPE_E	201248	2048	,15	1,67
SUB_D	200976	99819	7,48	1,67
TYPE_C	200480	1911	,14	1,67
SUB_E	199184	100058	7,49	1,66
TYPE_D	198912	1947	,15	1,66
SUB_C	198800	99940	7,49	1,65
SUB_B	198080	100150	7,50	1,65
TYPE_B	198080	1945	,15	1,65
TRACK	10656	1892	,14	,09
BEFORE	16	0	,00	0
total physical reads		1335136		

Middle ground in the Buffer Cache

Advisor – 2 GB BC

Cache Size (m)	Buffers	Estd Phys Read Factor	Estd Phys Reads
224	27,538	3.40	10,038,113
448	55,076	3.03	8,930,329
672	82,614	2.67	7,887,870
896	110,152	2.34	6,916,797
1,120	137,690	2.04	6,026,850
1,344	165,228	1.77	5,216,021
1,568	192,766	1.53	4,519,210
1,792	220,304	1.33	3,912,795
2,016	247,842	1.15	3,408,100
2,240	275,380	1.02	3,000,601
2,272	279,314	1.00	2,952,114
2,464	302,918	.90	2,667,326
2,688	330,456	.81	2,395,295
2,912	357,994	.73	2,163,887
3,136	385,532	.67	1,965,726
3,360	413,070	.60	1,785,126
3,584	440,608	.55	1,618,599
3,808	468,146	.50	1,466,411
4,032	495,684	.45	1,320,772
4,256	523,222	.40	1,185,718
4,480	550,760	.36	1,063,184

Middle ground in the Buffer Cache

Cache content – 4 GB BC

OBJECT_NAME	OBJECT_TYPE	CACHED_BLOCKS	ALL_BLOCKS	% cached	cached MB
MAIN_IDX1	INDEX	28340	28672	98,84	221
TYPE_IDXE	INDEX	7930	8064	98,34	62
TYPE_IDXB	INDEX	7928	8064	98,31	62
MAIN_IDXA	INDEX	5656	5760	98,19	44
TYPE_IDXA	INDEX	5656	5760	98,19	44
TYPE_B	TABLE	7248	7424	97,63	57
TYPE_C	TABLE	7248	7424	97,63	57
TYPE_E	TABLE	7248	7424	97,63	57
TYPE_D	TABLE	7247	7424	97,62	57
TYPE_A	TABLE	7240	7424	97,52	57
MAIN_IDXC	INDEX	5670	5888	96,30	44
MAIN_IDXB	INDEX	5670	5888	96,30	44
MAIN_IDXD	INDEX	5670	5888	96,30	44
TYPE_IDXC	INDEX	5670	5888	96,30	44
TYPE_IDXD	INDEX	5670	5888	96,30	44
MAIN_IDXE	INDEX	5670	5888	96,30	44
SUB_D	TABLE	28290	31744	89,12	221
SUB_E	TABLE	28269	31744	89,05	221
SUB_B	TABLE	28244	31744	88,97	221
SUB_A	TABLE	28191	31744	88,81	220
SUB_C	TABLE	28181	31744	88,78	220
MAIN	TABLE	226933	405376	55,98	1773
AFTER	TABLE	3	8	37,50	0
TRACK	TABLE	379	21504	1,76	3

Middle ground in the Buffer Cache

Activity – 4 GB BC

OBJECT_NAME	logical reads	physical reads	% physical reads	% buffergets
MAIN_IDX1	3000048	0	,00	24,96
MAIN	999360	421224	84,67	8,32
TYPE_IDXA	605984	2	,00	5,04
TYPE_IDXC	603696	0	,00	5,02
TYPE_IDXD	600624	0	,00	5
MAIN_IDXE	600512	0	,00	5
MAIN_IDXA	599648	2	,00	4,99
MAIN_IDXC	599216	0	,00	4,99
TYPE_IDXE	598544	5	,00	4,98
TYPE_IDXB	598032	12	,00	4,98
MAIN_IDXD	597792	0	,00	4,97
MAIN_IDXB	595696	0	,00	4,96
TYPE_A	205072	0	,00	1,71
SUB_D	203424	14949	3,00	1,69
SUB_C	201632	14895	2,99	1,68
TYPE_B	201584	1	,00	1,68
TYPE_D	201312	3	,00	1,67
SUB_E	199952	15089	3,03	1,66
SUB_B	199584	15067	3,03	1,66
TYPE_C	198624	2	,00	1,65
TYPE_E	198432	3	,00	1,65
SUB_A	198320	14633	2,94	1,65
TRACK	11600	1630	,33	,1
BEFORE	32	0	,00	0
BIN\$gKHcakqERUiTU0d9Lw5Giw==\$0	0	1	,00	0
total physical reads		497518		

Middle ground in the Buffer Cache

Advisor – 4 GB BC

-----	-----	-----	-----	-----
Estd Phys Cache Size (m)	Estd Phys Buffers	Read Factor		Reads
400	49,175	6.80		10,773,279
800	98,350	5.39		8,542,588
1,200	147,525	4.25		6,731,309
1,600	196,700	3.32		5,260,300
2,000	245,875	2.61		4,133,485
2,400	295,050	2.10		3,325,435
2,800	344,225	1.72		2,732,547
3,200	393,400	1.44		2,279,361
3,600	442,575	1.21		1,923,003
4,000	491,750	1.02		1,619,978
4,048	497,651	1.00		1,585,343
4,400	540,925	.84		1,339,365
4,800	590,100	.69		1,087,592
5,200	639,275	.53		845,130
5,600	688,450	.41		649,377
6,000	737,625	.41		649,256
6,400	786,800	.41		649,256
6,800	835,975	.41		649,256
7,200	885,150	.41		649,256
7,600	934,325	.41		649,256
8,000	983,500	.41		649,256

Middle ground in the Buffer Cache

Cache content – 4 GB BC / 2 GB KC

OBJECT_NAME	OBJECT_TYPE	CACHED_BLOCKS	ALL_BLOCKS	% cached	cached MB
MAIN_IDX1	INDEX	28338	28672	98,84	221
MAIN_IDXA	INDEX	5656	5760	98,19	44
TYPE_IDXA	INDEX	5656	5760	98,19	44
TYPE_IDXE	INDEX	7913	8064	98,13	62
TYPE_IDXB	INDEX	7910	8064	98,09	62
TYPE_C	TABLE	7246	7424	97,60	57
TYPE_B	TABLE	7245	7424	97,59	57
TYPE_E	TABLE	7243	7424	97,56	57
TYPE_D	TABLE	7243	7424	97,56	57
TYPE_A	TABLE	7235	7424	97,45	57
MAIN_IDXE	INDEX	5670	5888	96,30	44
MAIN_IDXD	INDEX	5670	5888	96,30	44
MAIN_IDXB	INDEX	5670	5888	96,30	44
MAIN_IDXC	INDEX	5670	5888	96,30	44
TYPE_IDXD	INDEX	5670	5888	96,30	44
TYPE_IDXC	INDEX	5670	5888	96,30	44
SUB_D	TABLE	23685	31744	74,61	185
SUB_A	TABLE	23670	31744	74,57	185
SUB_E	TABLE	23660	31744	74,53	185
SUB_B	TABLE	23618	31744	74,40	185
SUB_C	TABLE	23599	31744	74,34	184
MAIN	TABLE	251848	405376	62,13	1968
AFTER	TABLE	3	8	37,50	0
TRACK	TABLE	252	21504	1,17	2

Middle ground in the Buffer Cache

Activity – 4 GB BC / 2 GB KC

OBJECT_NAME	logical reads	physical reads	% physical reads	% buffergets
MAIN_IDX1	3003824	30	,00	24,99
MAIN	1000944	384103	62,80	8,33
TYPE_IDXC	605520	3	,00	5,04
MAIN_IDXA	603440	4	,00	5,02
TYPE_IDXA	602720	4	,00	5,01
MAIN_IDXC	602416	5	,00	5,01
TYPE_IDXB	601552	238	,04	5,01
MAIN_IDXE	600464	3	,00	5
TYPE_IDXD	598944	3	,00	4,98
MAIN_IDXD	598688	3	,00	4,98
TYPE_IDXE	598640	220	,04	4,98
MAIN_IDXB	597232	3	,00	4,97
SUB_A	202160	44782	7,32	1,68
TYPE_C	200208	44	,01	1,67
TYPE_D	199696	44	,01	1,66
TYPE_E	199520	49	,01	1,66
TYPE_A	199472	42	,01	1,66
SUB_E	199360	45310	7,41	1,66
SUB_C	198640	44751	7,32	1,65
TYPE_B	198368	44	,01	1,65
SUB_D	198192	45119	7,38	1,65
SUB_B	197376	45161	7,38	1,64
TRACK	11312	1631	,27	,09
BIN\$XmZPo7GeT8GoYIeap2y15g==\$0	32	1	,00	0
BEFORE	16	0	,00	0
total physical reads		611597		

Middle ground in the Buffer Cache

Advisor – 4 GB BC / 2 GB KC

Estd Phys Cache Size	Estd Phys (m) Buffers	Read Factor	Reads
192	23,604	3.83	8,930,979
384	47,208	3.43	8,007,610
576	70,812	3.04	7,103,981
768	94,416	2.67	6,221,537
960	118,020	2.33	5,439,607
1,152	141,624	2.01	4,694,970
1,344	165,228	1.73	4,041,922
1,536	188,832	1.48	3,448,489
1,728	212,436	1.26	2,945,147
1,920	236,040	1.07	2,505,773
2,000	245,875	1.00	2,333,352
2,112	259,644	.92	2,142,537
2,304	283,248	.79	1,837,110
2,496	306,852	.68	1,576,852
2,688	330,456	.58	1,352,245
2,880	354,060	.49	1,146,157
3,072	377,664	.41	961,640
3,264	401,268	.34	789,225
3,456	424,872	.28	644,363
3,648	448,476	.21	500,328
3,840	472,080	.15	358,852

Middle ground in the Buffer Cache

Advisor Keep – 4 GB BC / 2 GB KC

Estd Phys Cache Size (m)	Estd Phys Buffers	Read Factor	Reads
192	23,604	3.72	3,074,576
384	47,208	3.30	2,725,167
576	70,812	2.90	2,394,353
768	94,416	2.52	2,081,139
960	118,020	2.17	1,796,068
1,152	141,624	1.84	1,521,898
1,344	165,228	1.57	1,298,957
1,536	188,832	1.06	872,526
1,728	212,436	1.00	826,214
1,920	236,040	1.00	826,214
2,048	251,776	1.00	826,214
2,112	259,644	1.00	826,214
2,304	283,248	1.00	826,214
2,496	306,852	1.00	826,214
2,688	330,456	1.00	826,214
2,880	354,060	1.00	826,214
3,072	377,664	1.00	826,214
3,264	401,268	1.00	826,214
3,456	424,872	1.00	826,214
3,648	448,476	1.00	826,214
3,840	472,080	1.00	826,214

Automatic big table caching

- A cache section reserved for big tables
- caching data for table scans
- primarily designed for warehouse workloads
- also potential for mixed workloads
- uses temperature and object based algorithms

[Oracle Database Online Documentation 12c Release 1 \(12.1\) / Data Warehousing and Business Intelligence](#)

<https://mahmoudhatem.wordpress.com/2014/10/26/oracle-12c-automatic-big-table-caching/>

Middle ground in the Buffer Cache

Configure big table caching

```
alter system set  
DB_BIG_TABLE_CACHE_PERCENT_TARGET=  
80 scope=memory;
```



Only for full table scans !

<https://martincarstenbach.wordpress.com/tag/data-warehousing/>

Conclusion

- Several ways to make the buffer cache efficient
- Sql tuning
- Tuning of the physical structure
- Sometimes you have to extend
- Sub - Caches (secondary, keep)



bbi Software AG
Ruessenstrasse 12
6340 Baar ZG (Switzerland)
Phone: +41 (0)41 511 44 00
Direct: +41 (0)78 626 43 31
Fax: +41 (0)41 511 44 07

<mailto:Lothar.Flatz@bbi.ch>
<http://www.bbi.ch>

