ORA-SOLUTIONS.NET
Mastering Oracle Performance, High Availability, Manageability

Oracle Database
- Advanced Troubleshooting

Martin Decker

ORACLE 10g Certified Master

ORACLE Certified Master
Oracle Database 11g Administrator

ORACLE Certified Master
Oracle Database 12c Administrator

20. - 23. November
in Nürnberg

2018 DOAG Konferenz + Ausstellung

2018.doag.org

Agenda

- Vorstellung
- 5 Troubleshooting Einsätze aus der Praxis vom letzten Jahr
  - **Case 1:** Mutex Contention bei 12.2 RMAN Catalog Database
  - **Case 2:** Locking Problem ohne sichtbaren Blocker
  - **Case 3:** ORA-4031 bei 11gR2 2-Node RAC
  - **Case 4:** Oracle Linux 6 – RAC 12.1.0.2 – Instance Hang
  - **Case 5:** Oracle Linux 7 – RAC 11.2.0.4 - Instance Hang

whoami

- seit 2003 im Oracle Datenbank Umfeld tätig
- seit 2008 unabhängiger Oracle Consultant in D/A/CH
- Spezialisierung auf:
  - Performance Management (Instance / SQL)
  - Hochverfügbarkeit (MAA, RAC, DataGuard)
  - Manageability (OEM)
  - Unix (Linux, Solaris, HP-UX)
- Oracle Certified Master 10g/11g/12c
- Website & Blog: http://www.ora-solutions.net

Case 1:

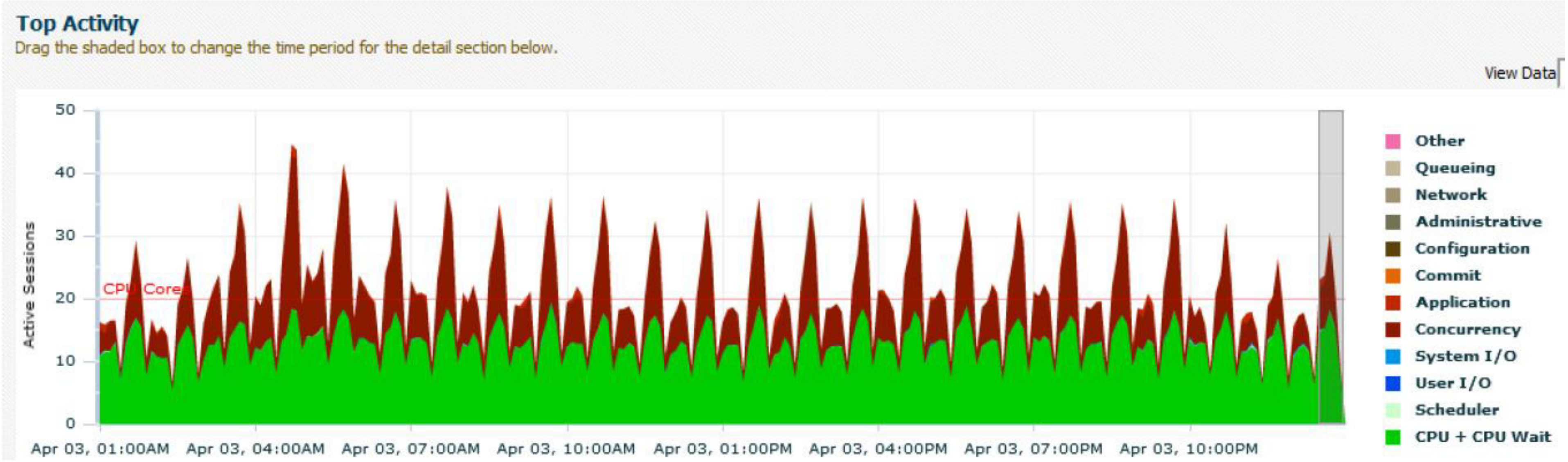
Mutex Contention bei
12.2 RMAN Catalog Database

Case 1: Mutex Contention

- System
  - RMAN Catalog Database – ca. 270 RMAN Catalog Schemas
  - Oracle Database 12.2.0.1.170718
- Symptome
  - PARSE ERRORS im Alert Log
  - Hohe CPU Last
  - Top Activity Chart zeigt Waits im Bereich Concurrency: (cursor: mutex X)

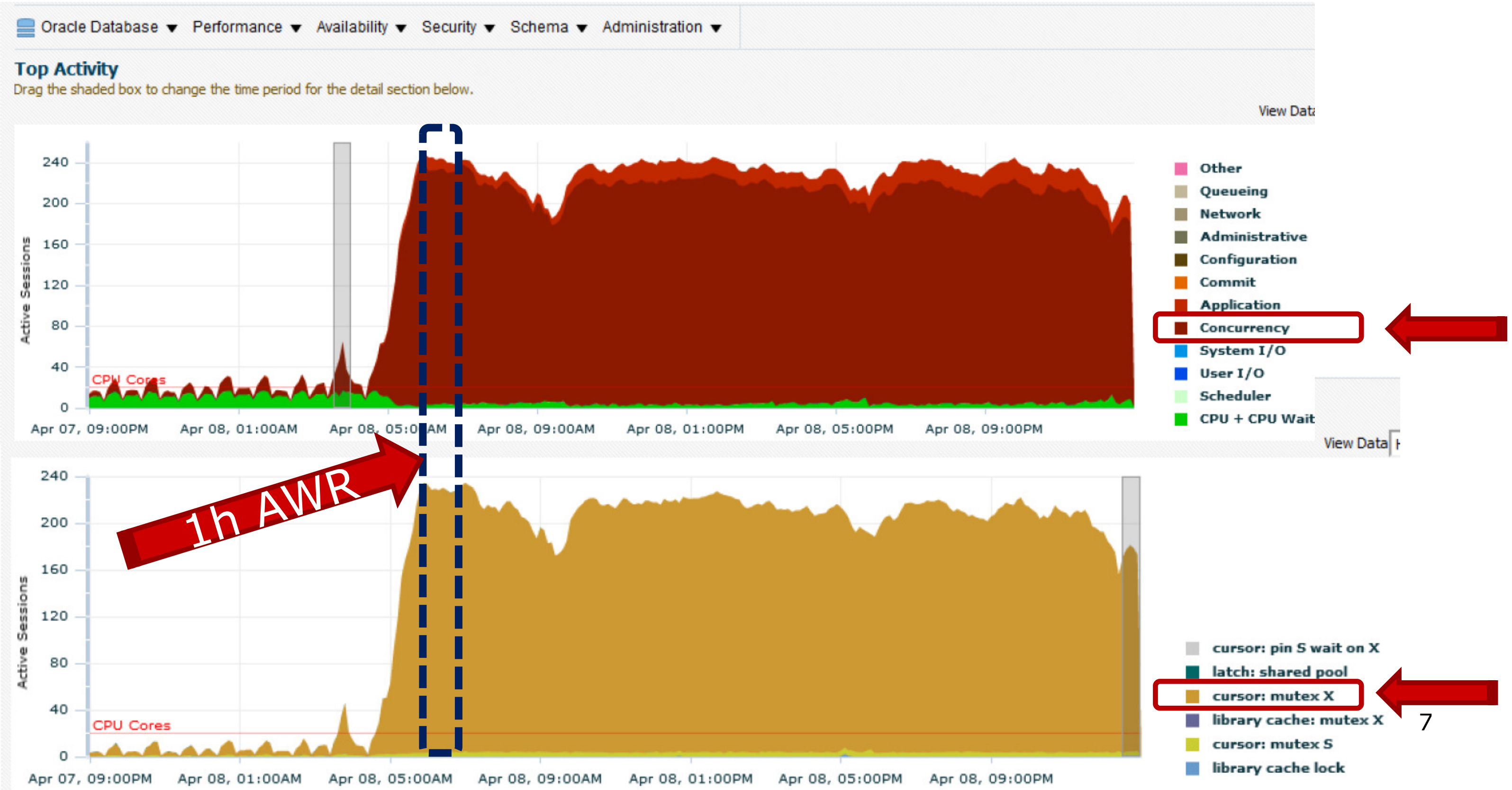
Case 1: Mutex Contention

- Top Activity bei Normalzustand



Case 1: Mutex Contention

- Top Activity bei akutem Problem – cursor: mutex X



Case 1: Mutex Contention

	Snap Id	Snap Time	Sessions	Cursors/Session
Begin Snap:	3257	08-Apr-18 06:00:11	318	10.2
End Snap:	3258	08-Apr-18 07:00:15	309	12.0
Elapsed:		60.08 (mins)		
DB Time:		14,591.53 (mins)		

- AWR Report bei akutem Problem

Load Profile

	Per Second	Per Transaction	Per Exec	Per Call
DB Time(s):	242.9	33.1	0.04	2.18
DB CPU(s):	3.9	0.5	0.00	0.03
Background CPU(s):	0.1	0.0	0.00	0.00
Redo size (bytes):	274,278.4	37,320.9		
Logical read (blocks):	25,734.2	3,501.6		
Block changes:	1,789.6	243.5		
Physical read (blocks):	383.0	52.1		
Physical write (blocks):	113.5	15.4		
Read IO requests:	108.4	14.7		
Write IO requests:	26.6	3.6		
Read IO (MB):	3.0	0.4		
Write IO (MB):	0.9	0.1		
IM scan rows:	0.0	0.0		
Session Logical Read IM:	0.0	0.0		
User calls:	111.3	15.1		
Parses (SQL):	209.1	28.5		
Hard parses (SQL):	24.4	3.3		
SQL Work Area (MB):	23.1	3.1		
Logons:	0.2	0.0		
Executes (SQL):	5,861.1	797.5		
Rollbacks:	0.1	0.0		
Transactions:	7.4			

Case 1: Mutex Contention

Instance Efficiency Percentages (Target 100%)

Buffer Nowait %:	100.00	Redo NoWait %:	100.00
Buffer Hit %:	98.78	In-memory Sort %:	100.00
Library Hit %:	99.77	Soft Parse %:	88.35
Execute to Parse %:	96.43	Latch Hit %:	98.90
Parse CPU to Parse Elapsd %:	1.19	% Non-Parse CPU:	24.81
Flash Cache Hit %:	0.00		

Top 10 Foreground Events by Total Wait Time

Event	Waits	Total Wait Time (sec)	Avg Wait	% DB time	Wait Class
cursor: mutex X	1,915,850	666.8K	348.03ms	76.2	Concurrency
enq: TX - row lock contention	73	36K	493.62 s	4.1	Application
DB CPU		13.9K		1.6	
cursor: mutex S	1,499,854	13.2K	8.78ms	1.5	Concurrency
acknowledge over PGA limit	15,845	172.1	10.86ms	.0	Scheduler
library cache lock	1,828	135.5	74.14ms	.0	Concurrency
cursor: pin S wait on X	2,026	98.7	48.72ms	.0	Concurrency
SQL*Net break/reset to client	14,466	66.2	4.58ms	.0	Application
direct path write temp	11,044	19.8	1.79ms	.0	User I/O
latch: shared pool	33,273	16.9	507.02us	.0	Concurrency

Case 1: Mutex Contention

- AWR Report bei akutem Problem

Statistic Name	Time (s)	% of DB Time	% of Total CPU Time
parse time elapsed	829,520.94	94.75	
sql execute elapsed time	740,268.01	84.55	
inbound PL/SQL rpc elapsed time	141,453.81	16.16	
DB CPU	13,898.22	1.59	98.02

Elapsed Time (s)	Executions	Elapsed Time per Exec (s)	%Total	%CPU	%IO	SQL Id	SQL Module	SQL Text
668,636.00			76.37	0.55		d3bumq6dhnn8v		declare first boolean; conf# n...
650,827.34			74.34	0.56		442f07fswtyr		select count(*) from conf where db_key = ...

d3bumq6dhnn8v declare first boolean; conf# number; begin if (:first > 0) then first := TRUE; else first := FALSE; end if; dbms_rcvman.getConfig(conf#, :name, :value, first); end;

Mutex Sleep Summary

- ordered by number of sleeps desc

Mutex Type	Location	Sleeps	Wait Time (ms)
Cursor Parent	kkscsAddChildNode [KKSPRTLOC28]	64,299,596	287,093,668
Cursor Parent	kkscsPruneChild [KKSPRTLOC27]	12,357,917	1,140,798

kkscs => child cursor

Case 1: Mutex Contention

- Gibt es Statements mit einer hohen Anzahl (>100) von Child Cursors?

```
SELECT sql_id, version_count, sql_text from v$sqlarea
order by 2 desc;
```

SQL_ID	VERSION_COUNT	SQL_TEXT
16dhrs3rhxzmc	3466	UPDATE RSR SET RSR_PH
cfuuuzmmz6zzx	3314	SELECT RSR_KEY FROM R
ddx81jvh8yvm5	3244	UPDATE FB SET OLDEST_
bzq48b5zww4fv	3218	SELECT 0 FROM FB WHER
442f07fsvvtyr	3139	SELECT COUNT(*) FROM
9h3pcpkzpr0vc	3135	SELECT NVL(MIN(SCN),E
0d9k3bnbg8v8h	3038	UPDATE NODE SET NODE.
akbp5gjaf1bxc	3010	SELECT HIGH_BCB_RECID
bj6djbmm28fcp	3005	SELECT HIGH_TS_RECID

Case 1: Mutex Contention

- Was sind "cursor: mutex X" waits?
  - Mutex: **Mut**ual **ex**clusive Synchronisierungsstruktur innerhalb Userland Oracle Code
  - Wird u.a. benötigt, wenn neuer Child-Cursor unter Parent Cursor erstellt (loading) wird, d.h. wenn kein passender Child-Cursor existiert
  - Mehr Details unter
    http://blog.tanelpoder.com/files/Oracle_Latch_And_Mutex_Contention_Troubleshooting.pdf
    http://www.hroug.hr/content/download/4932/75688/file/504_Senegacnik_Mutexes.pdf

C.3.26 cursor: mutex X

The session requests the mutex for a cursor object in exclusive mode, and it must wait because the resource is busy. The mutex is busy because either the mutex is being held in exclusive mode by another session or the mutex is being held shared by one or more sessions. The existing mutex holder(s) must release the mutex before the mutex can be granted exclusively.

Parameter	Description
P1	Hash value of cursor
P2	Mutex value (top 2 bytes contain SID holding mutex in exclusive mode, and bottom two bytes usually hold the value 0)
P3	Mutex where (an internal code locator) OR'd with Mutex Sleeps

Source: Oracle Documentation: Reference

Case 1: Mutex Contention

ORA-SOLUTIONS.NET
Mastering Oracle Performance, High Availability, Manageability

- Was sind Parent/Child Cursor und wann werden mehrere Child Cursor für einen Parent Cursor erstellt?
  - SQL Query wird durch Cursor in Library Cache repräsentiert
  - v$sql zeigt Child Cursor, identifiziert durch "ADDRESS/SQL_ID/CHILD_NUMBER"
  - wenn cursor nicht ge-"shared" werden kann, wird neuer Child Cursor erstellt
  - Child Cursor sind als Elemente einer verketteten Liste
  - beim Parsing muss Liste auf "passenden Child Cursor" untersucht werden
  - mehr als 60 Ursachen für "unshareable Cursors":
    - Rolling Invalidation von DBMS_STATS
    - **identische SQL Statements ausgeführt in unterschiedlichen Schemas** (zB RMAN Catalog Schemas)
    - Adaptive Cursor Sharing (ACS)
    - Reoptimization / Adaptive Statistics
    - unterschiedliche Session Settings zB. Optimizer Settings, NLS, etc.
    - Bugs

Case 1: Mutex Contention

- Liste der ca. 60 möglichen Ursachen "unshared" cursors
=> v$sql_shared_cursor (MOS 120655.1)

Column	Datatype	Description
SQL_ID	VARCHAR2(13)	SQL identifier
ADDRESS	RAW(4 | 8)	Address of the parent cursor
CHILD_ADDRESS	RAW(4 | 8)	Address of the child cursor
CHILD_NUMBER	NUMBER	Child number
UNBOUND_CURSOR	VARCHAR2(1)	(Y | N) The existing child cursor was not fully built (in other words, it was not optimized)
SQL_TYPE_MISMATCH	VARCHAR2(1)	(Y | N) The SQL type does not match the existing child cursor
OPTIMIZER_MISMATCH	VARCHAR2(1)	(Y | N) The optimizer environment does not match the existing child cursor
OUTLINE_MISMATCH	VARCHAR2(1)	(Y | N) The outlines do not match the existing child cursor
STATS_ROW_MISMATCH	VARCHAR2(1)	(Y | N) The existing statistics do not match the existing child cursor
LITERAL_MISMATCH	VARCHAR2(1)	(Y | N) Non-data literal values do not match the existing child cursor
FORCE_HARD_PARSE	VARCHAR2(1)	(Y | N) For internal use
EXPLAIN_PLAN_CURSOR	VARCHAR2(1)	(Y | N) The child cursor is an explain plan cursor and should not be shared
BUFFERED_DML_MISMATCH	VARCHAR2(1)	(Y | N) Buffered DML does not match the existing child cursor
PDML_ENV_MISMATCH	VARCHAR2(1)	(Y | N) PDML environment does not match the existing child cursor
INST_DRTLD_MISMATCH	VARCHAR2(1)	(Y | N) Insert direct load does not match the existing child cursor
SLAVE_QC_MISMATCH	VARCHAR2(1)	(Y | N) The existing child cursor is a slave cursor and the new one was issued by the coordinator (or, the existing child cursor was issued by the coordinator and the new one is a slave cursor)
TYPECHECK_MISMATCH	VARCHAR2(1)	(Y | N) The existing child cursor is not fully optimized

MV_REWRITE_MISMATCH	VARCHAR2(1)	(Y | N) Cursor needs recompilation because an SCN was used during compile time due to being rewritten by materialized view
ROLL_INVALID_MISMATCH	VARCHAR2(1)	(Y | N) Marked for rolling invalidation and invalidation window exceeded
OPTIMIZER_MODE_MISMATCH	VARCHAR2(1)	(Y | N) Parameter OPTIMIZER_MODE mismatch (for example, all_rows versus first_rows_1)
PX_MISMATCH	VARCHAR2(1)	(Y | N) Mismatch in one parameter affecting the parallelization of a SQL statement. For example, one cursor was compiled with parallel DML enabled while the other was not.
MV_STALEOBJ_MISMATCH	VARCHAR2(1)	(Y | N) Cursor cannot be shared because there is a mismatch in the list of materialized views which were stale when the cursor was built
FLASHBACK_TABLE_MISMATCH	VARCHAR2(1)	(Y | N) Cursor cannot be shared because there is a mismatch with triggers being enabled and/or referential integrity constraints being deferred
LITREP_COMP_MISMATCH	VARCHAR2(1)	(Y | N) Mismatch in use of literal replacement
PLSQL_DEBUG	VARCHAR2(1)	(Y | N) Value of the PLSQL_DEBUG parameter for the current session does not match the value used to build the cursor
LOAD_OPTIMIZER_STATS	VARCHAR2(1)	(Y | N) A hard parse is forced to initialize extended cursor sharing
ACL_MISMATCH	VARCHAR2(1)	(Y | N) Cached ACL evaluation result stored in the child cursor is not valid for the current session or user
FLASHBACK_ARCHIVE_MISMATCH	VARCHAR2(1)	(Y | N) Value of the FLASHBACK_DATA_ARCHIVE_INTERNAL_CURSOR parameter for the current session does not match the value used to build the cursor
		…schema used to build the cursor no longer exists
		…ring criterion is deprecated
		…d cursor was previously remote-mapped and is currently not remote-mapped. …cursor must be reparsed.
		… of run-time heap for the new cursor (or reload of aged out cursor) failed
		…ing child cursors have the unsafe literal bind hash values required by the…
		…rsor is marked for purging
		…gth(s) required for the current cursor are longer than the bind length(s) used to… cursor
		…parse is forced so that the optimizer can reoptimize the query with improved …ts (for example, cardinality estimates)
		… id, and reason the cursor is not shared. The content of this column is …ng XML.
		…is null, the FORCE_HARD_PARSE, LOAD_RUNTIME_HEAP_FAILED, and …_FAILED columns may enable you to determine why the child cursor is not…
		…container to which the data pertains. Possible values include: …value is used for rows containing data that pertain to the entire CDB. This value …sed for rows in non-CDBs. …value is used for rows containing data that pertain to only the root …e n is the applicable container ID for the rows containing data

LANGUAGE_MISMATCH	VARCHAR2(1)	(Y | N) The language handle does not match the existing child cursor
TRANSLATION_MISMATCH	VARCHAR2(1)	(Y | N) The base objects of the existing child cursor do not match
BIND_EQUIV_FAILURE	VARCHAR2(1)	(Y | N) The bind value's selectivity does not match that used to optimize the existing child cursor
INSUFF_PRIVS	VARCHAR2(1)	(Y | N) Insufficient privileges on objects referenced by the existing child cursor
INSUFF_PRIVS_REM	VARCHAR2(1)	(Y | N) Insufficient privileges on remote objects referenced by the existing child cursor
REMOTE_TRANS_MISMATCH	VARCHAR2(1)	(Y | N) The remote base objects of the existing child cursor do not match
LOGMINER_SESSION_MISMATCH	VARCHAR2(1)	(Y | N) LogMiner Session parameters mismatch
INCOMP_LTRL_MISMATCH	VARCHAR2(1)	(Y | N) Cursor might have some binds (literals) which may be unsafe/non-data. Value mismatch.
OVERLAP_TIME_MISMATCH	VARCHAR2(1)	(Y | N) Mismatch caused by setting session parameter ERROR_ON_OVERLAP_TIME
EDITION_MISMATCH	VARCHAR2(1)	(Y | N) Cursor edition mismatch
MV_QUERY_GEN_MISMATCH	VARCHAR2(1)	(Y | N) Internal, used to force a hard-parse when analyzing materialized view queries
USER_BIND_PEEK_MISMATCH	VARCHAR2(1)	(Y | N) Cursor is not shared because value of one or more user binds is different and this has a potential to change the execution plan
TYPCHK_DEP_MISMATCH	VARCHAR2(1)	(Y | N) Cursor has typecheck dependencies
NO_TRIGGER_MISMATCH	VARCHAR2(1)	(Y | N) Cursor and child have no trigger mismatch
FLASHBACK_CURSOR	VARCHAR2(1)	(Y | N) Cursor non-shareability due to flashback
ANYDATA_TRANSFORMATION	VARCHAR2(1)	(Y | N) Is criteria for opaque type transformation and does not match
PDDL_ENV_MISMATCH	VARCHAR2(1)	(Y|N) Environment setting mismatch for parallel DDL cursor (that is, one or more of the following parameter values have changed: PARALLEL_EXECUTION_ENABLED, PARALLEL_DDL_MODE, PARALLEL_DDL_FORCED_DEGREE, or PARALLEL_DDL_FORCED_INSTANCES)
TOP_LEVEL_RPI_CURSOR	VARCHAR2(1)	(Y | N) Is top level RPI cursor
DIFFERENT_LONG_LENGTH	VARCHAR2(1)	(Y | N) Value of LONG does not match
LOGICAL_STANDBY_APPLY	VARCHAR2(1)	(Y | N) Logical standby apply context does not match
DIFF_CALL_DURN	VARCHAR2(1)	(Y | N) If Slave SQL cursor/single call
BIND_UACS_DIFF	VARCHAR2(1)	(Y | N) One cursor has bind UACs and one does not
PLSQL_CMP_SWITCHS_DIFF	VARCHAR2(1)	(Y | N) PL/SQL anonymous block compiled with different PL/SQL compiler switches
CURSOR_PARTS_MISMATCH	VARCHAR2(1)	(Y | N) Cursor was compiled with subexecution (cursor parts were executed)
STB_OBJECT_MISMATCH	VARCHAR2(1)	(Y | N) STB is an internal name for a SQL Management Object Mismatch. A SQL Management Object Mismatch means that either a SQL plan baseline, or a SQL profile, or a SQL patch has been created for your SQL statement between the executions. Because a cursor is a read-only entity, a hard parse is forced to be able to create a new cursor that contains information about the new SQL management object related to this SQL statement.
CROSSEDITION_TRIGGER_MISMATCH	VARCHAR2(1)	(Y | N) The set of crossedition triggers to execute might differ
PQ_SLAVE_MISMATCH	VARCHAR2(1)	(Y | N) Top-level slave decides not to share cursor
TOP_LEVEL_DDL_MISMATCH	VARCHAR2(1)	(Y | N) Is top-level DDL cursor
MULTI_PX_MISMATCH	VARCHAR2(1)	(Y | N) Cursor has multiple parallelizers and is slave-compiled
BIND_PEEKED_PQ_MISMATCH	VARCHAR2(1)	(Y | N) Cursor based around bind peeked values

Case 1: Mutex Contention

- Wie viele Child Cursor kann ein Parent-Cursor maximal haben?
  - bis 11.2.0.2: unlimitiert viele
  - ab 11.2.0.3: Parameter "_cursor_obsolete_threshold"

_cursor_obsolete_threshold was first introduced in 11.2.0.3 with the default to 100 and then increased to 1024 in 11.2.0.4

11.2.0.3: _cursor_obsolete_threshold=100
11.2.0.4: _cursor_obsolete_threshold=1024
12.1: _cursor_obsolete_threshold=1024
12.2: _cursor_obsolete_threshold=8192

Case 1: Mutex Contention

- MOS Note 2431353.1 vom 11.08.2018 (ca. 4 Monate nach Problem)

The default value of _cursor_obsolete_threshold is increased heavily (8192 from 1024) from 12.2 onwards in order to support 4096 PDBs which was only 252 PDBs till 12.1. This parameter value is the maximum limit for obsoleting the parent cursors in an multitenant environment and cannot be increased beyond 8192.

Having said, this is **NOT** applicable for non-CDB environment and hence for those databases, this parameter should be set to 12.1 default value manually i.e. 1024. The default value of 1024 holds good for non-CDB environment and the same parameter can be adjusted case-to-case basis should there be a problem.

Case 1: Mutex Contention

- ab welchem Version Count Wert ist die Situation kritisch?
  => it depends :-)

- MOS liefert diese Informationen:

SQL ordered by Version Count

- Only Statements with Version Count greater than 20 are displayed

Any Version count > 20 is large and potentially detrimental to performance.

Version Count	Executions	SQL Id	SQL Module	SQL Text
148,635		6n7hnvnw7fcz8		[redacted]
144,946		06v14z9s8jzc7		[redacted]
1,019	296	[illegible]		[redacted]
979	310	94dhw6q5asj2r		[redacted]
967	281	66w361y5df55r		[redacted]

Version counts in to the thousands, such as those shown here, can be catastrophic.

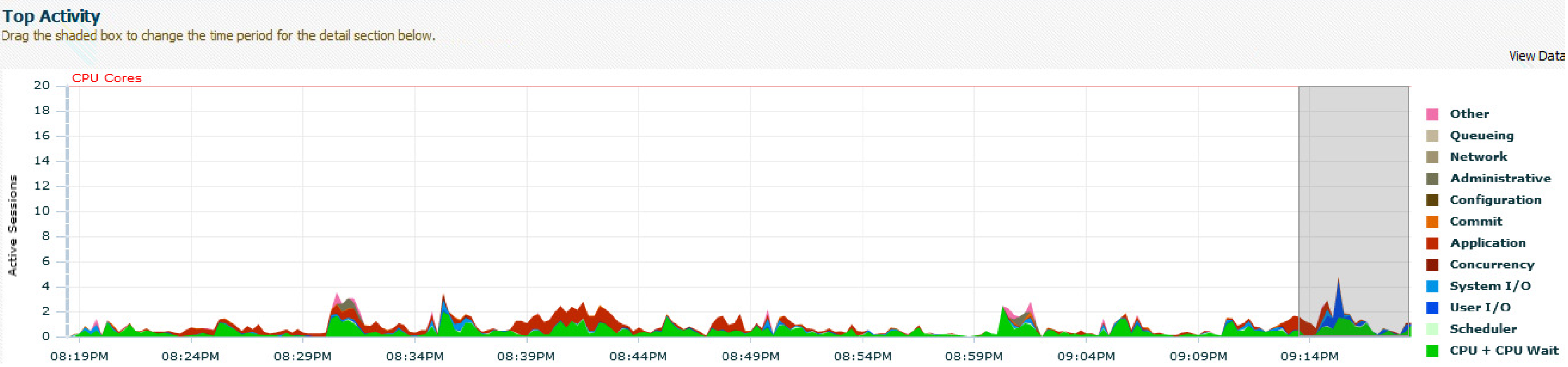
Source: Troubleshooting: High Version Count Issues (Doc ID 296377.1)

Case 1: Mutex Contention

- temporäre Lösung:
  - bei akuter Störung: mehrfach: ALTER SYSTEM FLUSH SHARED POOL;

- permanente Lösung:
  - Reduzierung der max. Child Cursors auf ca. Anzahl von RMAN Catalog Schemas plus etwas Puffer (hier: 512). Falls Anzahl von RMAN Catalog Schemas über 512 steigen sollte, muss Parameter erhöht werden.

```
ALTER SYSTEM SET "_cursor_obsolete_threshold" = 512
comment='Reduzierung Max Childs aufgrund von Mutex X Contention - 11.04.2018';
```



Case 1: Mutex Contention

- Fazit
  - "Corner-Case", weil die Anzahl der User-Schemas mit identischen SQL Statements sehr hoch ist. Allerdings sollte der Default-Wert von Oracle so gewählt werden, dass er für die meisten Systeme passt!
  - **Default-Wert von Parameter _cursor_obsolete_threshold (8192) für fast alle Systeme ungeeignet!**

- Empfehlung
  - regelmäßige Prüfung des Systems auf SQL Statements mit hohen Version Counts
  - bei Fund: Prüfung der Gründe der hohen Version Counts (MOS Note), wenn unerklärbar -> Oracle SR (evtl. Bug)
  - Aktuelle Patches (DBBP, RU)

Case 2:

Locking ohne sichtbaren Blocker

Case 2: Locking ohne sichtbaren Blocker

- System
  - ERP System-Datenbank
  - Oracle Database 11.2.0.4 auf Windows x64
- Symptome
  - viele Benutzer-Sessions warten auf "enq: TX row lock contention"
  - Spalte BLOCKING_SESSION.V$SESSION bleibt leer
  - Kein Eintrag in v$lock für Blocker
  - Hanganalyze zeigt Blocker nicht
  - ORA-02049: timeout: distributed transaction waiting for lock
- Behebung
  - Instance Restart –> nicht besonders beliebt bei Senior Management
  - Oracle Sev 1 Service Request

Case 2: Locking ohne sichtbaren Blocker

ORA-SOLUTIONS.NET
Mastering Oracle Performance, High Availability, Manageability

- Kein Eintrag in v$lock für Blocker

```
SQL> select * from v$lock where sid = 68 and type = 'TX';

ADDR             KADDR            SID        TY ID1        ID2        LMODE      REQUEST    CTIME      BLOCK
---------------- ---------------- ---------- -- ---------- ---------- ---------- ---------- ---------- ----------
0000000778A7FCA8 0000000778A7FD00 68         TX 1245193    735567     0          6          355        0

SQL> select * from v$lock where id1 = 1245193 and id2 = 735567;

ADDR             KADDR            SID        TY ID1        ID2        LMODE      REQUEST    CTIME      BLOCK
---------------- ---------------- ---------- -- ---------- ---------- ---------- ---------- ---------- ----------
0000000778A7FCA8 0000000778A7FD00 68         TX 1245193    735567     0          6          355        0
```

Case 2: Locking ohne sichtbaren Blocker

ORA-SOLUTIONS.NET
Mastering Oracle Performance, High Availability, Manageability

- Hanganalyze / System State Dump

```
SQL> oradebug setmypid
SQL> oradebug unlimit;
SQL> oradebug hanganalyze 4
SQL> oradebug hanganalyze 4
SQL> oradebug hanganalyze 5
```

```
SQL> oradebug setmypid
SQL> oradebug unlimit;
SQL> select * from dual;
SQL> oradebug dump systemstate 258
SQL> oradebug dump systemstate 258
```

Case 2: Locking ohne sichtbaren Blocker

ORA-SOLUTIONS.NET
Mastering Oracle Performance, High Availability, Manageability

```
--------------------------------------------------------------------------------
Chain 1:
--------------------------------------------------------------------------------
    Oracle session identified by:
    {
                instance: 1 (erpprod.erpprod)
                   os id: 4600
              process id: 161, ORACLE.EXE (SHAD)
              session id: 68
        session serial #: 20263
    }
    is waiting for 'enq: TX - row lock contention' with wait info:
    {
                      p1: 'name|mode'=0x54580006
                      p2: 'usn<<16 | slot'=0x130009
                      p3: 'sequence'=0xb394f
           time in wait: 7 min 6 sec
           timeout after: never
                 wait id: 10
                blocking: 0 sessions
             current sql: select * from  ERPAPP.remote_printing_node_tab  WHERE
NAME='XXX-INST05-P' for update
             short stack:
ksedsts()+585<-ksdxfstk()+44<-ksdxcb()+2402<-ssthreadsrgruncallback()+629<-
OracleOradebugThreadStart()+829<-0000000076D559CD<-0000000076E8A561<-0000000076EABD7A<-
000007FEFCE910AC<-skgpwwait()+182<-ksliwat()+2459<-kslwaitctx()+184<-ksqcmi()+7848<-
ksqgtlctx()+6814<-ksqgelctx()+829<-ktuGetTxForXid()+190<-ktcwit1()+526<-kdddgb()+9636<-
kddlkr()+284<-qerfuProcessFetchedRow()+1019<-__PGOSF308_qerfuLockingRowProcedure()+89<-
qertbFetchByRowID()+2708<-qerfuStart()+590<-selexe0()+3495<-opiexe()+22377<-kpoal8()+2624<
-opiodr()
            wait history:
              * time between current wait and wait #1: 0.001442 sec
              1.       event: 'SQL*Net message from client'
                 time waited: 41.751540 sec
                     wait id: 9              p1: 'driver id'=0x42455100
                                             p2: '#bytes'=0x1
              * time between wait #1 and #2: 0.000001 sec
```

- Hanganalyze-Trace Output

fehlende "is blocked by" Section

Case 2: Locking ohne sichtbaren Blocker

- System State Dump
  - Zustand der verschiedenen Ressourcen wird in einem State Object Tree festgehalten und gepflegt
  - Parent / Child Relationship zwischen State Objects
  - Parent (Level 1) State Object ist "Process", Child (Level 2) ist zB "Session"
  - Liste von Process "State Objects" nach "Oracle pid:" sortiert

Case 2: Locking ohne sichtbaren Blocker

- System State Dump – Überblick über Abhängigkeiten

```
grep -n -e "SO: " -e "^\s*(" systemstate.trc | less
```

Level 1

Level 2

Level 3

...

Einrückung bei Childs

```
      SO: 0x000000075C65F078, type: 2, owner: 0x0000000000000000, flag: INIT/-/-/0x00 if: 0x3 c: 0x3
1056: (process) Oracle pid:1, ser:0, calls cur/top: 0x0000000000000000/0x0000000000000000
1062: (post info) last post received: 0 0 0
1089:   (latch info) wait_event=0 bits=0x0
1       SO: 0x00000001498F0E70, type: 5, owner: 0x000000075C65F078, flag: INIT/-/-/0x00 if: 0x3 c: 0x3
109       SO: 0x0000000756B0FC70, type: 87, owner: 0x00000001498F0E70, flag: INIT/-/-/0x00 if: 0x3 c: 0x3
1102:     (circuit holder) disp = 0x0000000756B0FBE8 (0, 1), proc = 0x00000007785CE6F8 (18, 27)
1104: SO: 0x00000007785C9958, type: 2, owner: 0x0000000000000000, flag: INIT/-/-/0x00 if: 0x3 c: 0x3
1110: (process) Oracle pid:2, ser:1, calls cur/top: 0x000000075C94EC70/0x000000075C94EC70
1116: (post info) last post received: 0 0 16
1160:   (latch info) wait_event=0 bits=0x0
1162:   SO: 0x0000000764A07EA8, type: 14, owner: 0x00000007785C9958, flag: INIT/-/-/0x00 if: 0x1 c: 0x1
1173:   (broadcast handle) 0000000764A07EA8 flag: (2) ACTIVE SUBSCRIBER,
1175:   SO: 0x0000000764A07CF8, type: 14, owner: 0x00000007785C9958, flag: INIT/-/-/0x00 if: 0x1 c: 0x1
1186:   (broadcast handle) 0000000764A07CF8 flag: (2) ACTIVE SUBSCRIBER,
1188:   SO: 0x00000007646EBFF0, type: 4, owner: 0x00000007785C9958, flag: INIT/-/-/0x00 if: 0x3 c: 0x3
1391:   (session) sid: 99 ser: 1 trans: 0x0000000000000000, creator: 0x00000007785C9958
1393:   SO: 0x0000000764A07998, type: 14, owner: 0x00000007785C9958, flag: INIT/-/-/0x00 if: 0x1 c: 0x1
1404:   (broadcast handle) 0000000764A07998 flag: (2) ACTIVE SUBSCRIBER,
1407:   SO: 0x0000000760C227B0, type: 24, owner: 0x00000007785C9958, flag: -/-/-/0x00 if: 0x3 c: 0x3
1409:   SO: 0x000000075C94EC70, type: 3, owner: 0x00000007785C9958, flag: INIT/-/-/0x00 if: 0x3 c: 0x3
1413:   (call) sess: cur 7646ebff0, rec 0, usr 7646ebff0; flg:30 fl2:1; depth:0
1415:   SO: 0x0000000760BCDA80, type: 20, owner: 0x00000007785C9958, flag: INIT/-/-/0x00 if: 0x1 c: 0x1
        (osp req holder)
1423:
```

Case 2: Locking ohne sichtbaren Blocker

- System State Dump
  - Step 1:
    Identifizieren einer wartenden (blocked) Session und Ermittlung des "Resource Identifier" für Lock

Case 2: Locking ohne sichtbaren Blocker

```
-------------------------------------
SO: 0x00000007606D9B88, type: 4, owner: 0x000000075C68F8B8, flag: INIT/-/-/0x00 if: 0x3 c: 0x3
 proc=0x000000075C68F8B8, name=session, file=ksu.h LINE:12748, pg=0
(session) sid: 68 ser: 20263 trans: 0x0000000755DB9838, creator: 0x000000075C68F8B8
          flags: (0x41) USR/- flags_idl: (0x1) BSY/-/-/-/-/-
          flags2: (0x40009) -/-/INC
          DID: , short-term DID:
          txn branch: 0x0000000000000000
          edition#: 100              oct: 3, prv: 0, sql: 0x00000007708AA2E8, psql: 0x000000076BDCDED8, user: 0/SYS
ksuxds FALSE at location: 0
service name: SYS$USERS
client details:
  O/S info: user: username, term: hostname, ospid: 5968:4960
  machine: hostname program: sqlplus.exe
  application name: sqlplus.exe, hash value=254292535
Current Wait Stack:
 0: waiting for 'enq: TX - row lock contention'
    name|mode=0x54580006, usn<<16 | slot=0x130009, sequence=0xb394f
    wait_id=10 seq_num=11 snap_id=1
    wait times: snap=8 min 15 sec, exc=8 min 15 sec, total=8 min 15 sec
    wait times: max=infinite, heur=8 min 15 sec
    wait counts: calls=166 os=166
    in_wait=1 iflags=0x15a0
Wait State:
  fixed_waits=0 flags=0x22 boundary=0x0000000000000000/-1
...
  -------------------------------------
  SO: 0x0000000778A7FCA8, type: 8, owner: 0x000000075BC6A370, flag: INIT/-/-/0x00 if: 0x1 c: 0x1
   proc=0x000000075C68F8B8, name=enqueue, file=ksq1.h LINE:380, pg=0
  (enqueue) TX-00130009-000B394F   DID: 0001-00A1-0003F59C
  lv: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00  res_flag: 0x6
  req: X, lock_flag: 0x10, lock: 0x0000000778A7FD00, res: 0x0000000760A2BD80
  own: 0x00000007606D9B88, sess: 0x00000007606D9B88, proc: 0x000000075C68F8B8, prv: 0x0000000760A2BDA0
-------------------------------------
```

SID: 68: Waiting for **"enq: TX – row lock contention"**
enqueue: **TX / slot = 0x130009 / seq=0xb394f**
und Request Mode ist e**X**clusive

Case 2: Locking ohne sichtbaren Blocker

- System State Dump
  - Step 2:
    Suchen nach der angeforderten Enqueue Ressource im restlichen System State Dump

ORA-SOLUTIONS.NET
Mastering Oracle Performance, High Availability, Manageability

Case 2: Locking ohne sichtbaren Blocker

```
BEGIN DUMP RESERVED PROCESSES
END DUMP RESERVED PROCESSES
DETACHED BRANCHES:
----------------------------------------
SO: 0x0000000755E7C0F0, type: 58, owner: 0x0000000000000000, flag: INIT/-/-/0x00 if: 0x3 c: 0x3
 proc=0x0000000000000000, name=branch, file=ktccts.h LINE:420, pg=0
(branch) trn = 0x0000000000000000, flg = 0x40, state = 0x00 bno=0 ser=0 evt=0
 creator = 0x0000000000000000 uid = 0x0 serial# = 0 ttl = 0x0 dtm = 0x0
 2PCrole = (nch = 0x0 flg = 0x0) DTP svc = 0x0
   gti.fmt: 00000000  gti.tln: 0  gti.bln: 0
   gti.tid =
   gti.bid =
  ----------------------------------------
  SO: 0x0000000755E2B378, type: 58, owner: 0x0000000755E7C0F0, flag: INIT/-/-/0x00 if: 0x3 c: 0x3
   proc=0x000000077860EB80, name=branch, file=ktccts.h LINE:420, pg=0
  (branch) trn = 0x0000000755E01068, flg = 0x41090, state = 0x00 bno=1 ser=10 evt=10
   creator = 0x00000007607940E8 uid = 0x42 serial# = 7929 ttl = 0x15180 dtm = 0x5a4db9f5
   2PCrole = (nch = 0x0 flg = 0x0) DTP svc = 0x773ffc978
     gti.fmt: 00020003  gti.tln: 31  gti.bln: 29
     gti.tid = 0x312d61333130306361 3a633831363a3561336333373738 3a32656135373961
     gti.bid = 0x613331303063613a633831363a35613363333737383a32656135376132
    ----------------------------------------
    SO: 0x0000000755E01068, type: 56, owner: 0x0000000755E2B378, flag: INIT/-/-/0x00 if: 0x3 c: 0x3
     proc=0x000000077860EB80, name=transaction, file=ktccts.h LINE:410, pg=0
    (trans) flg = 0x00401e03, flg2 = 0x04080000, flg3 = 0x00000000, prx = 0x0000000000000000, ros = 2147483647, crtses=0x00000007607940E8
```

nach allerletztem Process gibt es
Abschnitt "DETACHED BRANCHES", owner = 0x00

2PCrole -> Hinweis auf Distributed Transactions
(2 PhaseCommit), XA (Global Transactions)

```
      gti.bid = 0x613331303063613a633831363a35613363333737383a3265613
    (enqueue) TX-00130009-000B394F DID: 0001-00E6-000010CD
    lv: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00  res_flag: 0x6
    mode: X, lock_flag: 0x0, lock: 0x0000000755E010E0, res: 0x0000000760A2BD80
    own: 0x0000000755E01068 (TXN), sess: 0x0000000000000000, prv: 0x0000000760A2BD90
```

enqueue: **TX-00130009-000B394F**
mode X gehalten von einem State Object,
das keiner session zugeteilt ist, sondern
einer Section "DETACHED BRANCHES"

Case 2: Locking ohne sichtbaren Blocker

- Step 3: MOS Research
  Solving locking problems in a XA environment (1248848.1)

Case 2: Transaction is not yet prepared but xa_end already called

- Locks can be released by
  - database reboot to cleanup transactions and locks
  - using a JDBC program attaching to that transaction and do a rollback

```
SQL>
SQL> select --+ ORDERED
  2  '----------------------------------------'||'
  3  Curent Time : '|| substr(to_char(sysdate,' HH24.MI.SS'),1,22) ||'
  4  '||'TX start_time: '||t.KTCXBSTM||'
  5  '||'FORMATID: '||g.K2GTIFMT ||'
  6  '||'GTXID: '||g.K2GTITID_EXT ||'
  7  '||'Branch: '||g.K2GTIBID ||'
  8  Local_Tran_Id ='||substr(t.KXIDUSN||'.'||t.kXIDSLT||'.'||t.kXIDSQN,1,15)||'
  9  '||'KTUXESTA='|| x.KTUXESTA ||'
 10  '||'KTUXEDFL='|| x.KTUXECFL ||'
 11  Lock_Info: ID1: ' || ((t.kXIDUSN*64*1024)+ t.kXIDSLT)
 12  ||' ID2: '|| t.kXIDSQN
 13  XA_transaction_INFO
 14  from x$k2gte g, x$ktcxb t, x$ktuxe x
 15  where g.K2GTDXCB =t.ktcxbxba and
 16  x.KTUXEUSN = t.KXIDUSN(+) and
 17  x.KTUXESLT = t.kXIDSLT(+) and
 18  x.KTUXESQN =t.kXIDSQN(+);

XA_TRANSACTION_INFO
--------------------------------------------
--------------------------------------------
-----------------------------------------
Curent Time :  13.46.26
TX start_time: 02/21/18 13:18:46
FORMATID: 131075
GTXID: 312D61333130306361 3A656534653A356138
Branch: 61333130306361 3A656534653A356138633
Local_Tran_Id =13.15.2684227
KTUXESTA=ACTIVE
KTUXEDFL=NONE
Lock_Info: ID1: 851983 ID2: 2684227
```

```
C:\DBA\scripts>set JAVA_HOME=c:\oracle\product\11.2.0.4\dbhome_1\jdk\jre
C:\DBA\scripts>set CLASSPATH=c:\dba\scripts\;c:\oracle\product\11.2.0.4\dbhome_1\jdbc\lib\ojdbc5.jar;%CLASSPATH%
C:\DBA\scripts>c:\oracle\product\11.2.0.4\dbhome_1\jdk\jre\bin\java XA_rb jdbc:oracle:thin:@erphost:1521:ERPPROD -ROLLBACK 13 15 2684227
++ URL: jdbc:oracle:thin:@erphost:1521:ERPPROD
++ Rollback of Local_tran_ID    : 13.15.2684227
++       got      XA resource      handle
++       got        Connection     handle
++ SQL : select g.K2GTIFMT,      g.K2GTITID_EXT, g.K2GTIBID,     rawtohex( g.K2GTITID_EXT), rawtohex(g.K2GTIBID) from sys.vw_x$k2gte      g,
sys.vw_x$ktcxb t,  sys.vw_x$ktuxe x       where    g.K2GTDXCB
SLT(+) and x.KTUXESQN =t.KXIDSQN(+)   and t.KXIDUSN=13 and t.KXIDSLT=15 and t.kXIDSQN= 2684227
++ Getting XID  for Local_Tran_ID:   13.15.2684227
+  Found global XID: -->
...
 -> Would        your really     like to continue (yes or no):
yes
...      doing Rollback
...      Rollback done !
```

Case 2: Locking ohne sichtbaren Blocker

- Fazit
  - Temporärer Workaround: Rollback über Java Procedure von MOS
  - Permanente Behebung durch Applikationshersteller

Case 3:

ORA-4031 bei 11gR2 2-Node RAC

Case 3: ORA-4031 bei 11gR2 2-Node RAC

ORA-SOLUTIONS.NET
Mastering Oracle Performance, High Availability, Manageability

- System
  - ERP System Database
  - Oracle Database – Standard Edition 11.2.0.4
  - 2-Node RAC auf Linux x86-64
  - 2 CPUs á 8 Core => (32 Hyperthreading Cores pro Node)
  - 500 GB RAM pro Node
- Symptome
  - Fehlermeldung:

    ```
    ORA-04031: unable to allocate 3896 bytes of shared memory
    ("shared pool","select a.default_selectivity...","sga heap(2,0)","kglsim object batch")
    ```
  - kein Login möglich
  - keine Möglichkeit "ALTER SYSTEM FLUSH SHARED_POOL;" durchzuführen

Case 3: ORA-4031 bei 11gR2 2-Node RAC

- Step 1: Öffnen von Oracle Support Service Request und Analyse des ORA-4031 Trace Files
  - unqualifiziertes erstes Feedback von Oracle Support verweist auf hohen Wert von open_cursor als Ursache
  - Parameter wurde nicht geändert und ist Anforderung von Application Vendor
  - Fakt: automatisches Trace File des ORA-4031 Fehlers enthält nicht die nötigen Informationen zur sicheren Identifizierung des Problems

Case 3: ORA-4031 bei 11gR2 2-Node RAC

- Step 2: Sammeln der nötigen Diagnose-Daten bei erneutem Auftreten
  - 2 Alternativen:
    - a. Parameter 4031_dump_bitvec so setzen, dass nötiger Heapdump geschrieben wird – ODER -
    - b. Setzen von "smartem" Event: einmaliges Erstellen eines Heapdumps und Errorstacks bei Auftreten (pro Session)

```
alter system set events '4031 trace name HEAPDUMP level 536870914, lifetime 1; name errorstack level 3, lifetime 1';
```

  - Warten bis Fehler erneut auftritt und Heapdump geschrieben wird, dann event deaktivieren

```
alter system set events '4031 trace name HEAPDUMP off; name errorstack off';
```

  - Nach erneutem Auftreten -> Escalation auf Sev1 und Upload von ORA-4031 Trace (inkl. Heapdump und Errorstack)

Case 3: ORA-4031 bei 11gR2 2-Node RAC

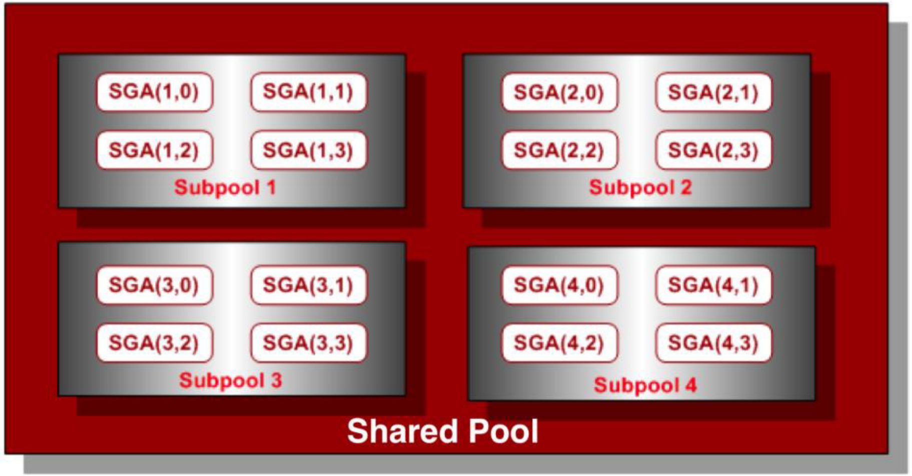
- Memory-Konfiguration:
  - SGA_TARGET: 192 GB
  - DB_CACHE_SIZE: 150 GB
  - **SHARED_POOL_SIZE: 37 GB**
  - JAVA_POOL_SIZE: 3,5 GB
  - STREAMS_POOL_SIZE: 0,5 GB
  - LARGE_POOL_SIZE: 0,5 GB
- Shared Pool zum Zeitpunkt des ORA-4031:

```
Total free memory          4688 MB
Total memory alloc.        32 GB
Grand total                37 GB
```

- Wie kann es zum ORA-4031 kommen, wenn noch 4.6GB frei sind?

Case 3: ORA-4031 bei 11gR2 2-Node RAC

- Shared-Pool-Layout



- Shared Pool wird in bis zu 7 Subpools geteilt: (Parameter _kghdsidx_count)
- <   12.1.0.1: Jeder Subpool wird in 4 Durations geteilt
(0-Instance [perm], 1-Session, 2-Cursor, 3-Execution)
- >= 12.1.0.1: Jeder Subpool wird in 2 Durations geteilt (0-Instance/Session/Cursor, 1-Execution)
- `ORA-04031: unable to allocate` **`3896 bytes`** `of shared memory`
`("`**`shared pool`**`","select a.default_selectivity...","`**`sga heap(2,0)`**`","kglsim object batch")`
- Anforderung eines **Shared-Pool** Chunks von zusammenhängenden
3896 Bytes aus **Subpool 2, Duration 0**: SGA(2,0) fehlgeschlagen

Bild: http://www.aioug.org/sangam/images/stories/conference/Sangam13/Presentations/36321_sharma_pdf.pdf

Case 3: ORA-4031 bei 11gR2 2-Node RAC

ORA-SOLUTIONS.NET
Mastering Oracle Performance, High Availability, Manageability

- Heapdump Trace File enthält relevante Daten:

```
Allocation Request Summary Informaton
=====================================
Allocation request for: kglsim object batch
...
Heap: 0x6007b278, size: 3896
******************************************************
HEAP DUMP heap name="sga heap(3,0)"  desc=0x6007b278
extent sz=0xfe0 alt=248 het=32767 rec=9 flg=-126 opc=2
...
durations enabled for this heap
reserved granules for root 0 (granule size 134217728)
Total heap size    =2684352800. (2559 MB)
Total free space   =     38360
Total reserved free space    =120837856 (115MB)
Permanent space    =2563474760
```

Request von sga heap (3,0)
für "kglsim object batch"
Größe: 3896 bytes

Subpool (3,0) ist 2559 MB groß, hat aber nur noch 38360 Bytes frei!

Case 3: ORA-4031 bei 11gR2 2-Node RAC

- Heapdump Analyzer analysiert 4031 Trace File mit Heapdump
- Heaps nach Größe sortiert und Info für Chunk_type (perm, free, etc.)

```
./heapdump_analyzer  ./PROD1_ora_27396.trc > heap.txt

grep -e "---" -e "Total size" -e "sga heap(3,0)"  heap.txt

-- Heapdump Analyzer v1.00 by Tanel Poder ( http://www.tanelpoder.com )

Total_size #Chunks  Chunk_size,        From_heap,       Chunk_type,  Alloc_reason
---------- ------- ------------ ----------------- ----------------- -----------------

114123720     17    6713160 ,    sga heap(3,0),         R-free,
28442152       1   28442152 ,    sga heap(3,0),          perm,   perm
...
7485000        1    7485000 ,    sga heap(3,0),          perm,   perm
6713080        1    6713080 ,    sga heap(3,0),        R-free,
6712104        1    6712104 ,    sga heap(3,0),        R-perm,   perm
...
3028568        1    3028568 ,    sga heap(3,0),          perm,   perm
...

8928           4    2232 ,    sga heap(3,0),          perm,   perm
7864           1    7864 ,    sga heap(3,0),          perm,   perm
5488           2    2744 ,    sga heap(3,0),          perm,   perm
3792           1    3792 ,    sga heap(3,0),          free,   <<<<<<<<<<<<<<
...
```

Chunk Types:
- **free**
- **recreatable**
- freeable
- permanent

größter **freier** zusammenhängender Chunk = 3792 Bytes (< 3896 requested)

Source: https://github.com/tanelpoder/tpt-oracle/blob/master/tools/unix/heapdump_analyzer

Case 3: ORA-4031 bei 11gR2 2-Node RAC

- Besonderheit:
  - Wenn ORA-4031 Fehler nur für Requests aus Duration "0" sind, z.B. sga heap (1,**0**), (2,**0**), etc. und DB Release < 12.1.0.1, dann sehr wahrscheinlich:

ORA-4031: unable to allocate 4160 bytes of shared memory ("shared pool","unknown object","sga heap(4,0)","modification ") (Doc ID 1675470.1)

SOLUTION

Changes for unpublished Bug 8857940 are included in 12.1.0.1 and permits to group the shared pool durations in 2 groups to allow better shareability of the memory (granule movement) and avoid the ORA-4031. Patch 8857940 is available for some platforms for previous versions.

The issue can be worked around with next actions:

- disable durations by setting "_enable_shared_pool_durations"=false or disabling ASMM i.e. setting SGA_TARGET to 0. The impact with disabling durations is that the shared pool cannot shrink effectively negating that element of auto memory management.

  - AND / OR -

- increase the shared pool size (using SHARED_POOL_SIZE, SGA_TARGET or MEMORY_TARGET depending on memory configuration)

Case 3: ORA-4031 bei 11gR2 2-Node RAC

- Patch 8857940 (Unix only)

  In the shared pool, combines the three unshrinkable durations (instance, session, and cursor) into one pool. In the streams pool, combines the two unshrinkable durations (instance and session) into one pool.

  - oder –

- _enable_shared_pool_durations = false

Case 3: ORA-4031 bei 11gR2 2-Node RAC

- Known Issues mit ORA-4031 (Doc ID 4031.1)

The list below is restricted to show only bugs believed to affect version 12.2.0.1.
Other bugs may affect this version but have not been confirmed as being relevant yet.

ORA-4031 Bugs in 12.2.0.1

There are 8 bugs listed.

NB	Prob	Bug	Fixed	Description
	III	26405036	19.1	Large Allocation Of "ges enqueues" and "ges resource dynamic" In The Shared Pool
	II	26844406	18.1	ORA-4031 Due To Memory Growth Seen For 'execution ID memory' Comment
	-	25209912	18.1	ogg-db: Integrated Parallel Replicat repeatedly abends with batch txns due to ORA-4031-streams pool leak
	II	23011934	18.1	Enabling Flashback or Setting a Restore Point Causes Slow Performance and/or ORA-4031
	III	15931756	11.2.0.3.11, 11.2.0.3.BP21, 11.2.0.4	ORA-4031 / Queries against SYS_FBA_TRACKEDTABLES not shared (do not use binds)
	III	27163928	18.1	ORA-4031 error due to leakage of 'gc index split transaction' permanent memory chunks
	II	25226665	18.1	ORA-4031 LMS waiting for latch free while trying to get lock on message buffer
D	III	5618049	10.2.0.4, 11.1.0.6	"mvobj part des" leaked memory after partition DDL (ORA-4031)

Known Bugs in other Products

13548390	Error 4031 in kgttusec - plsql memory > 1g unreliable

Case 4:

Oracle Linux 6 – RAC 12.1.0.2 – Instance Hang

Case 4: OL6–RAC 12.1.0.2 – Instance Hang

- System
  - ERP System Database
  - Oracle Database – Standard Edition 2 Version 12.1.0.2 (Patchlevel: Jan 2017)
  - 2-Node RAC auf Linux x86-64
  - pro Node: 1 CPU á 18 Core (36 HT Cores pro Node)
  - 378 GB RAM pro Node

Case 4: OL6–RAC 12.1.0.2 – Instance Hang

- Symptome am 18.09.2017
  - während geplanter Hardwarewartung eines Nodes
  - CPU Überlastung von verbleibenden Node
  - Hohe Anzahl von Sessions bei Ausführung von selbem SQL Statement d3z03cqbgcrqb
  - SQL Statement d3z03cqbgcrqb hat 284 Child Cursors
  - ORA-20: maximum number of processes (500) exceeded

Case 4: OL6–RAC 12.1.0.2 – Instance Hang

- Step 1: Killen von einigen Sessions, Login und Diagnose-Daten sammeln
  - v$session
  - v$sql
  - v$sql_shared_cursor
  - v$resource_limit
  - Statspack Report
  - Alert Log

- Step 2: Service wiederherstellen
  - normale Performance nach Flush Shared Pool und ERP Application Restart

- Step 3: Service Request für Root-Cause-Analysis erstellen

Case 4: OL6–RAC 12.1.0.2 – Instance Hang

ORA-SOLUTIONS.NET
Mastering Oracle Performance, High Availability, Manageability

- RAC Reconfiguration / Instance Recovery bei geplantem Instance Shutdown

```
Mon Sep 18 15:55:33 2017
Reconfiguration started (old inc 24, new inc 26)
List of instances (total 1) :
 1
Dead instances (total 1) :
 2
My inst 1
publish big name space -  dead or down/up instance detected, invalidate domain 0
 Global Resource Directory frozen
 * dead instance detected - domain 0 invalid = TRUE
 Communication channels reestablished
 Master broadcasted resource hash value bitmaps
 Non-local Process blocks cleaned out
Mon Sep 18 15:55:33 2017
 LMS 0: 0 GCS shadows cancelled, 0 closed, 0 Xw survived
Mon Sep 18 15:55:33 2017
 LMS 1: 0 GCS shadows cancelled, 0 closed, 0 Xw survived
Mon Sep 18 15:55:33 2017
 LMS 2: 1 GCS shadows cancelled, 0 closed, 0 Xw survived
 Set master node info
 Submitted all remote-enqueue requests
 Dwn-cvts replayed, VALBLKs dubious
 All grantable enqueues granted
Mon Sep 18 15:55:35 2017
 Post SMON to start 1st pass IR
Mon Sep 18 15:55:35 2017
Instance recovery: looking for dead threads
Mon Sep 18 15:55:35 2017
Beginning instance recovery of 1 threads
Mon Sep 18 15:55:35 2017
Started redo scan
Mon Sep 18 15:55:35 2017
minact-scn: master found reconf/inst-rec before recscn scan old-inc#:24 new-inc#:24
Mon Sep 18 15:55:37 2017
Completed redo scan
 read 469279 KB redo, 30960 data blocks need recovery
Mon Sep 18 15:55:41 2017
 Submitted all GCS remote-cache requests
 Fix write in gcs resources
Mon Sep 18 15:55:42 2017
Reconfiguration complete (total time 9.4 secs)
```

Reconfiguration mit GRD freeze (9,4 s)

Instance Recovery (ca. 10 s)

```
Mon Sep 18 15:55:44 2017
* validated domain 0, flags = 0x0
Mon Sep 18 15:55:44 2017
Started redo application at
 Thread 2: logseq 10247, block 2590695
Mon Sep 18 15:55:44 2017
Recovery of Online Redo Log: Thread 2 Group 8 Seq 10247 Reading mem 0
  Mem# 0: +FRA/IFSPROD/ONLINELOG/group_8.1210.931274199
  Mem# 1: +REDO/IFSPROD/ONLINELOG/group_8.269.931274219
Mon Sep 18 15:55:53 2017
Recovery of Online Redo Log: Thread 2 Group 7 Seq 10248 Reading mem 0
  Mem# 0: +FRA/IFSPROD/ONLINELOG/group_7.1286.931273441
  Mem# 1: +REDO/IFSPROD/ONLINELOG/group_7.270.931273465
Mon Sep 18 15:55:54 2017
Completed redo application of 278.20MB
Mon Sep 18 15:55:54 2017
Completed instance recovery at
 Thread 2: logseq 10248, block 182566, scn 27292490664
 30251 data blocks read, 31094 data blocks written, 469279 redo k-bytes read
Mon Sep 18 15:55:54 2017
Thread 2 advanced to log sequence 10249 (thread recovery)
Mon Sep 18 15:55:54 2017
Redo thread 2 internally disabled at seq 10249 (SMON)
```

Case 4: OL6–RAC 12.1.0.2 – Instance Hang

```
Snapshot       Snap Id      Snap Time      Sessions Curs/Sess Comment
~~~~~~~~     ---------- ------------------ -------- --------- ------------------
Begin Snap:      17021 18-Sep-17 14:00:26       174       3.8
  End Snap:      17022 18-Sep-17 15:00:01       172       4.1
   Elapsed:              59.58 (mins) Av Act Sess:       0.8
   DB time:              48.27 (mins)      DB CPU:      46.51 (mins)

Cache Sizes             Begin        End
~~~~~~~~~~~        ---------- ----------
    Buffer Cache:   310,656M              Std Block Size:        8K
     Shared Pool:    27,008M                  Log Buffer:  119,640K

Load Profile                 Per Second   Per Transaction  Per Exec  Per Call
~~~~~~~~~~~~            ---------------   --------------- --------- ---------
      DB time(s):                   0.8               0.0      0.00      0.00
       DB CPU(s):                   0.8               0.0      0.00      0.00
       Redo size:              52,914.8           2,479.1
   Logical reads:              29,611.5           1,387.3
   Block changes:                 299.8              14.1
  Physical reads:                   8.9               0.4
 Physical writes:                   5.4               0.3
      User calls:                 220.0              10.3
          Parses:               1,852.7              86.8
     Hard parses:                   0.4               0.0
W/A MB processed:                   2.5               0.1
          Logons:                  22.6               1.1
        Executes:               3,859.7             180.8
       Rollbacks:                  20.5               1.0
    Transactions:                  21.3

Instance Efficiency Indicators
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
            Buffer Nowait %:   99.99       Redo NoWait %:  100.00
            Buffer  Hit   %:   99.97    Optimal W/A Exec %:  100.00
            Library Hit   %:   99.98        Soft Parse %:   99.98
         Execute to Parse %:   52.00         Latch Hit %:   99.95
Parse CPU to Parse Elapsd %:   75.67     % Non-Parse CPU:   83.07
```

```
Snapshot       Snap Id      Snap Time      Sessions Curs/Sess Comment
~~~~~~~~     ---------- ------------------ -------- --------- ------------------
Begin Snap:      17022 18-Sep-17 15:00:01       172       4.1
  End Snap:      17023 18-Sep-17 16:00:14       446       2.2
   Elapsed:              60.22 (mins) Av Act Sess:      99.6
   DB time:           5,998.39 (mins)      DB CPU:     244.47 (mins)

Cache Sizes             Begin        End
~~~~~~~~~~~        ---------- ----------
    Buffer Cache:   310,656M              Std Block Size:        8K
     Shared Pool:    27,008M                  Log Buffer:  119,640K

Load Profile                 Per Second   Per Transaction  Per Exec  Per Call
~~~~~~~~~~~~            ---------------   --------------- --------- ---------
      DB time(s):                  99.6               5.2      0.03      0.49
       DB CPU(s):                   4.1               0.2      0.00      0.02
       Redo size:              38,454.4           1,988.3
   Logical reads:              33,599.1           1,737.2
   Block changes:                 223.1              11.5
  Physical reads:                  23.9               1.2
 Physical writes:                  14.4               0.7
      User calls:                 201.3              10.4
          Parses:               1,670.5              86.4
     Hard parses:                   0.6               0.0
W/A MB processed:                   3.1               0.2
          Logons:                  20.8               1.1
        Executes:               3,296.0             170.4
       Rollbacks:                  18.6               1.0
    Transactions:                  19.3

Instance Efficiency Indicators
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
            Buffer Nowait %:   99.99       Redo NoWait %:  100.00
            Buffer  Hit   %:   99.93    Optimal W/A Exec %:  100.00
            Library Hit   %:   99.93        Soft Parse %:   99.97
         Execute to Parse %:   49.32         Latch Hit %:   99.94
Parse CPU to Parse Elapsd %:    4.31     % Non-Parse CPU:   12.06

 Shared Pool Statistics        Begin    End
                              ------  ------
             Memory Usage %:   92.66   74.76
    % SQL with executions>1:   79.96   86.86
  % Memory for SQL w/exec>1:   84.63   74.76

Top 5 Timed Events                                                   Avg %Total
~~~~~~~~~~~~~~~~~~                                                  wait   Call
Event                                            Waits    Time (s)   (ms)   Time
----------------------------------------- ------------ ----------- ------ ------
resmgr:cpu quantum                              10,992     341,605  31078   83.9
CPU time                                                    14,075            3.5
```

Case 4: OL6–RAC 12.1.0.2 – Instance Hang

- Findings
  - Finding 1: Oracle SE2 Thread Limit
  - Finding 2: hohe Logon-Frequenz (GRD Freeze bei Instance Shutdown)
  - Finding 3: hohe Anzahl von Child Cursor

Case 4: OL6–RAC 12.1.0.2 – Instance Hang

- Exkurs: Oracle "Standard" Editions
  - Standard Edition One
    - max 2 CPU Sockets auf einem Host
    - kein RAC
    - nur bis inkl. 12.1.0.1
  - Standard Edition
    - Single-Instance: ein Host mit max 4 CPU Sockets und 4 CPUs bestückt
    - RAC: max. 2 Hosts mit je 2 CPU Sockets und max. 2 CPU Sockets bestückt (pro Node)
    - nur bis inkl. 12.1.0.1
  - Standard Edition 2
    - ab 12.1.0.2
    - Single-Instance: ein Host mit max 2 CPU Sockets und max 2 CPUs bestückt
    - RAC: max 2 Hosts mit jeweils max 2 CPU Sockets pro Host und max 1 CPU pro Host bestückt

Case 4: OL6–RAC 12.1.0.2 – Instance Hang

- Finding 1: Standard Edition 2 – Thread Limit (2)
  - Single-Instance:
    - **max. 16 aktive Foreground-Prozesse auf CPU aktiv**
    - Job Processes werden auch mitgerechnet
  - RAC:
    - **pro Node max 8 aktive Foreground-Prozesse auf CPU aktiv**
    - **Clusterwide: 16 aktive Foreground-Prozesse auf CPU aktiv**
    - Job Processes werden auch mit gerechnet
    - **Bei Wartung eines Nodes verbleiben trotzdem bei anderem Node nur 8 mögliche aktive Threads**

Case 4: OL6–RAC 12.1.0.2 – Instance Hang

- Finding 1: Standard Edition 2 – Thread Limit (3)
  - gezählt wird mit:
    - STATUS='ACTIVE' AND
      STATE != 'WAITING' AND
      TYPE = 'USER'
  - CPU Zeit (Time Slices) wird gleichmäßig auf max 8 / 16 "Sessions" verteilt, Sessions warten auf Wait Event "resmgr: cpu quantum" (STATE=WAITING), wenn schon alle 8 / 16 "Slots" belegt sind
  - spannende Frage:
    - Was passiert, wenn Session kritische Ressourcen hält (Latches, Mutexes, Library Cache Locks, Pins, etc.) und dann keine CPU bekommt und Ressourcen deshalb auch nicht freigeben kann ... ;-)

Case 4: OL6–RAC 12.1.0.2 – Instance Hang

- Finding 2: hohe Logon Frequenz
  - 20 Logons pro Sekunde von Apache / PHP Web Application, die ERP Daten ausliest
  - Logons und Logoffs sehr teuer im RAC Umfeld (Freigeben von Shared Pool Chunks -> benötigt Latches)
  - **Lösung: Database Resident Connection Pool (DRCP)**

Case 4: OL6–RAC 12.1.0.2 – Instance Hang

- Finding 2: hohe Logon Frequenz
  - Database Resident Connection Pool (DRCP):
    - eigenen Service erstellen (srvctl add service)
    - Aktivieren DRCP

```
exec dbms_connection_pool.configure_pool(minsize=>25,maxsize=>25);
exec dbms_connection_pool.start_pool();
```

- PHP oci.ini

```
extension=oci8.so
oci8.connection_class = "INTRANET"
```

- Im PHP Source-Code alle Connect Functions ändern:

```
$conn = OCILogon("INTRANET","secret","//rac-scan/PROD",'AL32UTF8');

Nach:
$conn = oci_pconnect("INTRANET","secret","//rac-scan/INTRANET_SERVICE:POOLED", 'AL32UTF8')
```

Case 4: OL6–RAC 12.1.0.2 – Instance Hang

ORA-SOLUTIONS.NET
Mastering Oracle Performance, High Availability, Manageability

- Finding 3: hohe Anzahl von Child Cursor für d3z03cqbgcrqb
  - **289 Sessions** mit
    STATUS='ACTIVE',
    STATE='WAITING',
    EVENT=**'resmgr:cpu quantum**',
    SQL_ID ='d3z03cqbgcrqb'
  - **13 Sessions** warten auf "library cache lock", welcher von SID 37 gehalten wird
  - SID 37 ist eine der 289 Sessions, die wegen Thread Limit warten müssen
  - **11 Sessions** mit
    STATUS='ACTIVE'
    SQL_ID='d3z03cqbgcrqb'

resmgr: cpu quantum
Indikator auf Thread Limit von
Standard Edition 2

SE2 Thread Limit:
alle "Slots" der RAC Instance belegt

Case 4: OL6–RAC 12.1.0.2 – Instance Hang

- Finding 3: hohe Anzahl von Child Cursor für d3z03cqbgcrqb
  - SQL ID: d3z03cqbgcrqb
  - 284 Child Cursor
  - Reasons (v$sql_shared_cursor)
    - BIND_LENGTH_UPGRADEABLE
    - LANGUAGE_MISMATCH
    - ROLL_INVALID_MISMATCH
  - zu hohe Anzahl von Child Cursors für Top SQL Statement
  - hoher CPU Verbrauch bei Soft-Parse
  - Folgeproblem: SE2 Thread Limit
  - **Support hat zusätzlich noch einen Bug identifiziert:**

```
Bug 18529766 - Excessive Child Cursors generated with VPD due to Authorization Check Failed
(Fixed 12.2, Patch für 12.1.0.2)
```

Case 4: OL6–RAC 12.1.0.2 – Instance Hang

- Fazit
  - Überlast-Situation ausgelöst durch Hardware-Wartung von RAC Node
  - Kombination von:
    - High Logon Rate
    - Top SQL mit vielen Child Cursor
    - SE2 Thread Limit

Case 5:

Oracle Linux 7 – RAC 11.2.0.4 - Instance Hang

Case 5: OL7 – RAC 11.2.0.4 - Instance Hang

- System:
  - Rolling Re-Install von 2 Node RAC Cluster für Upgrade von Oracle Linux 5 auf Oracle Linux 7 wegen Storage-Migration und fehlender OL5 Zertifizierung
  - RDBMS/GRID Version 11.2.0.4
- Symptome: **4 Monate nach Re-Install**
  - Extrem schlechte Performance
  - Lange Wartezeiten bei Queries von anderer Instanz
  - Fehlermeldung im RDBMS Alert Log

```
Mon Feb 26 08:53:37 2018
skgxpvfynet: mtype: 61 process 15801 failed because of a resource problem in the OS. The OS has most likely run out of buffers (rval: 4)
Errors in file /u01/app/oracle/diag_p/diag/rdbms/prod/PROD2/trace/PROD2_ora_15801.trc (incident=480004):
ORA-00603: ORACLE server session terminated by fatal error
ORA-27504: IPC error creating OSD context
ORA-27300: OS system dependent operation:sendmsg failed with status: 105
ORA-27301: OS failure message: No buffer space available
ORA-27302: failure occurred at: sskgxpsnd2
Incident details in: /u01/app/oracle/diag_p/diag/rdbms/prod/PROD2/incident/incdir_480004/PROD2_ora_15801_i480004.trc
```

Case 5: OL7 – RAC 11.2.0.4 - Instance Hang

- **Step 1:** Memory-Check

```
cat /proc/meminfo
zzz ***Mon Feb 26 08:53:09 CET 2018
MemTotal:       528028424 kB
MemFree:        14593828 kB
MemAvailable:   78305772 kB
Buffers:        28009752 kB
Cached:         46896496 kB
SwapCached:            0 kB
Active:         22627436 kB
Inactive:       66945168 kB
Active(anon):   14315300 kB
Inactive(anon):  2105748 kB
Active(file):    8312136 kB
Inactive(file): 64839420 kB
Unevictable:      363996 kB
Mlocked:          364020 kB
SwapTotal:      33554428 kB
SwapFree:       33554428 kB
Dirty:               404 kB
Writeback:             0 kB
AnonPages:      15768480 kB
Mapped:           709896 kB
Shmem:            945384 kB
Slab:            2462644 kB
SReclaimable:    2092232 kB
SUnreclaim:       370412 kB
KernelStack:       28336 kB
PageTables:       395568 kB
NFS_Unstable:          0 kB
Bounce:                0 kB
WritebackTmp:          0 kB
CommitLimit:    87853440 kB
Committed_AS:   22087384 kB
VmallocTotal:   34359738367 kB
VmallocUsed:     1106260 kB
VmallocChunk:   34085810172 kB
HardwareCorrupted:     0 kB
AnonHugePages:         0 kB
CmaTotal:          16384 kB
CmaFree:               0 kB
HugePages_Total:   204800
HugePages_Free:    20214
HugePages_Rsvd:     6266
HugePages_Surp:        0
Hugepagesize:       2048 kB
```

MemTotal: 503 GB

MemFree: 13,9 GB

MemAvailable: 74,6 GB

Case 5: OL7 – RAC 11.2.0.4 - Instance Hang

- **Step 2:** SEV1 Service Request und TFA Upload

- **Step 3:** Feedback von Service Request
  - MTU Size of Loopback Interface soll von 65536 auf 16384 reduziert werden
  - Default Wert bei Oracle Linux 7: 65536
  - Default Wert führt zu Problemen mit Memory Fragmentation

```
lo: flags=73<UP,LOOPBACK,RUNNING> mtu 65536
inet 127.0.0.1 netmask 255.0.0.0
inet6 ::1 prefixlen 128 scopeid 0x10<host>
loop txqueuelen 0 (Local Loopback)
RX packets 1965818660 bytes 570345508658 (531.1 GiB)
RX errors 0 dropped 0 overruns 0 frame 0
TX packets 1965818660 bytes 570345508658 (531.1 GiB)
TX errors 0 dropped 0 overruns 0 carrier 0 collisions 0
```

Case 5: OL7 – RAC 11.2.0.4 - Instance Hang

- **Step 4:** Reduzierung lo (Loopback Interface) MTU Size und Node Reboot – Performance wieder normal

- **Step 5:** Reduzierung lo (Loopback Interface) MTU Size von anderem Node in nächstem geplanten Wartungsfenster

Case 5: OL7 – RAC 11.2.0.4 - Instance Hang

- MOS:
  - Oracle Linux: ORA-27301:OS Failure Message: No Buffer Space Available (Doc ID 2041723.1)
  - ORA-27301: OS Failure Message: No Buffer Space Available / ORA-27302: failure occurred at: sskgxpsnd2 Source Script (Doc ID 2322410.1)
- Kritikpunkte:
  - Anforderung ist nicht in Oracle Documentation enthalten
  - weder Cluster Verification Utility (CVU), noch orachk prüfen den Wert bzw. geben Empfehlungen
- Fazit:
  - SEV1 Support hat schnell reagiert und Fehler sofort gefunden
  - Keine Chance für Admin proaktiv das Problem zu vermeiden

Q & A

ORACLE
10g Certified Master

ORACLE
Certified Master
Oracle Database 11g
Administrator

ORACLE
Certified Master
Oracle Database 12c
Administrator

Martin Decker
ora-solutions.net

E-Mail: martin.decker@ora-solutions.net
Internet: http://www.ora-solutions.net
Blog: http://www.ora-solutions.net/web/blog/