

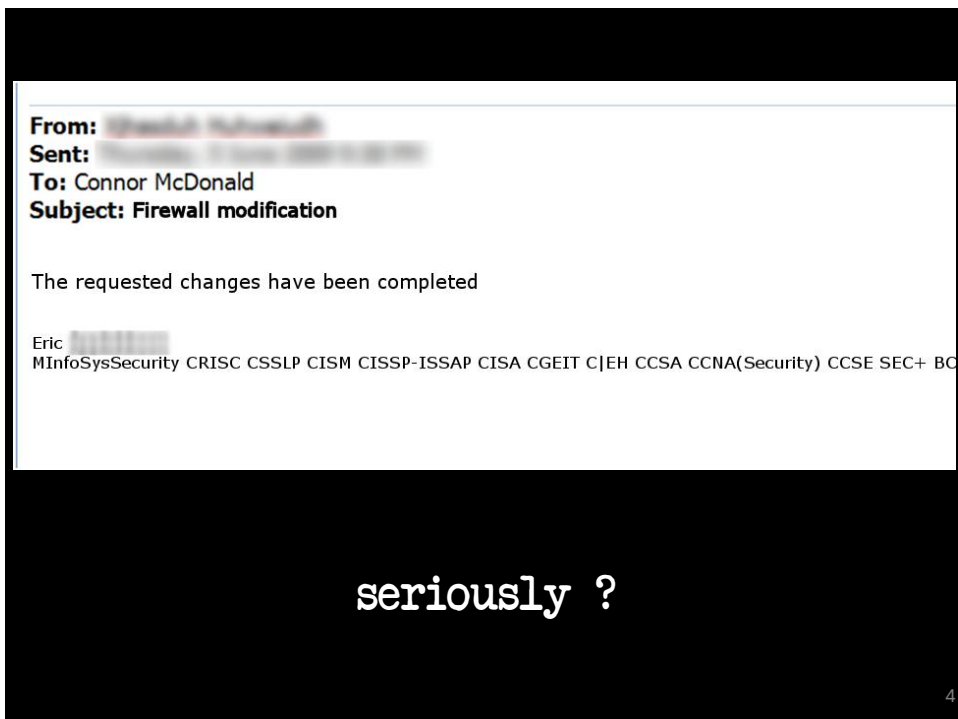
Connor McDonald



1

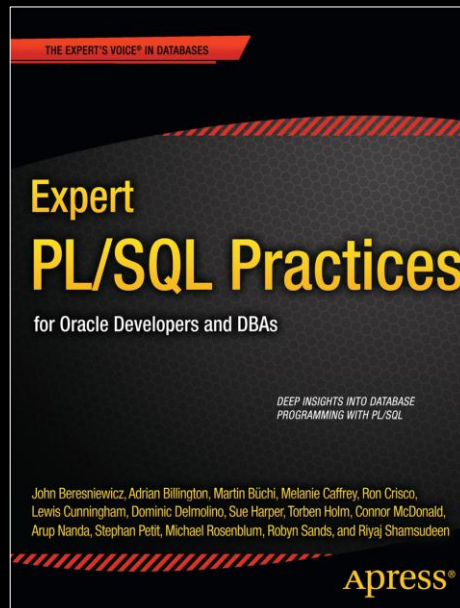


2





gratuitous book plug



today

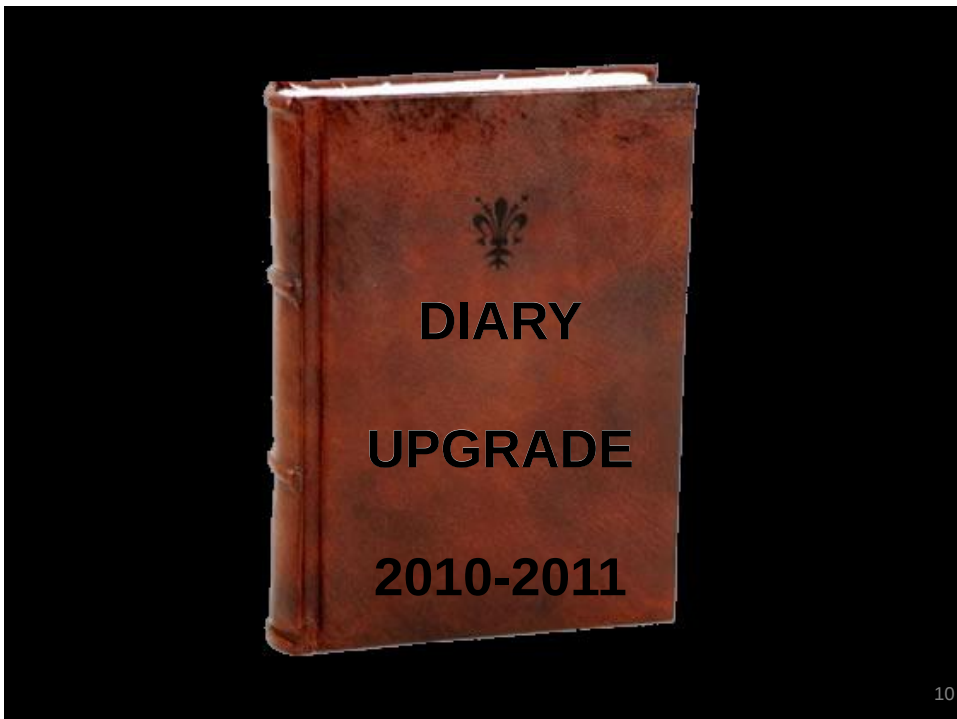
7

upgrade to 11.2

optimizer statistics

SQL techniques

8





11

why ?

12





3,760 km away...

...from our cluster

17

why ?

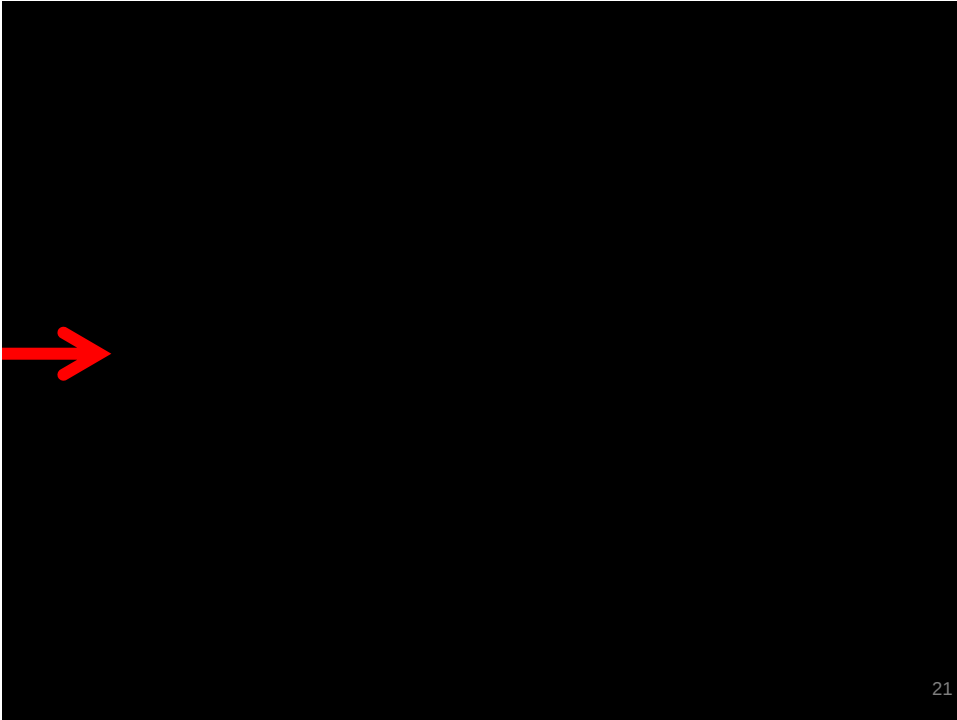
18

pain

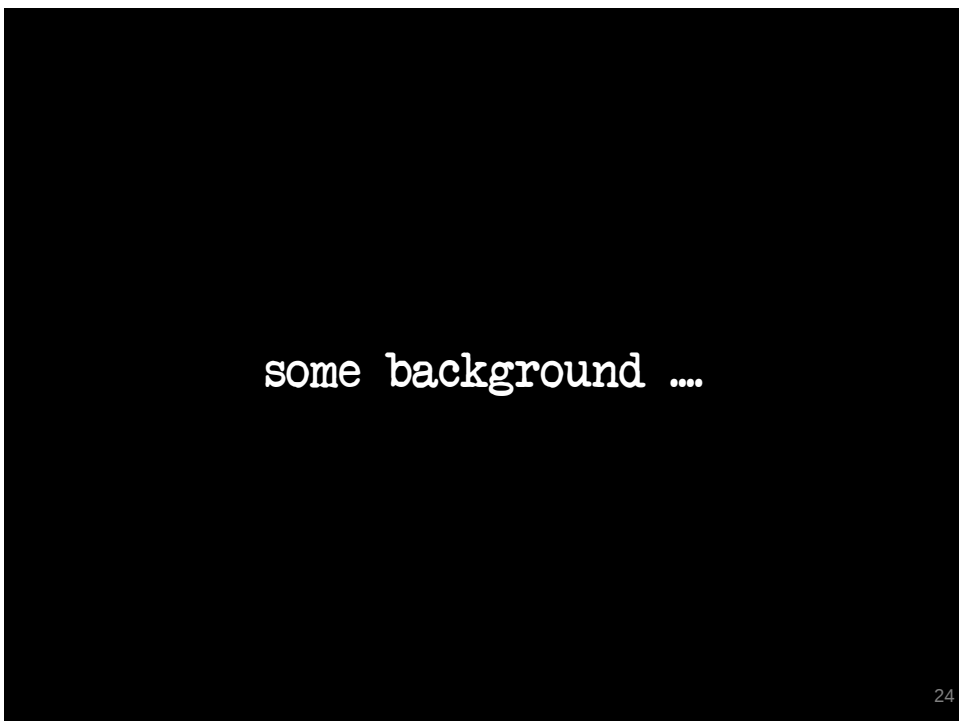
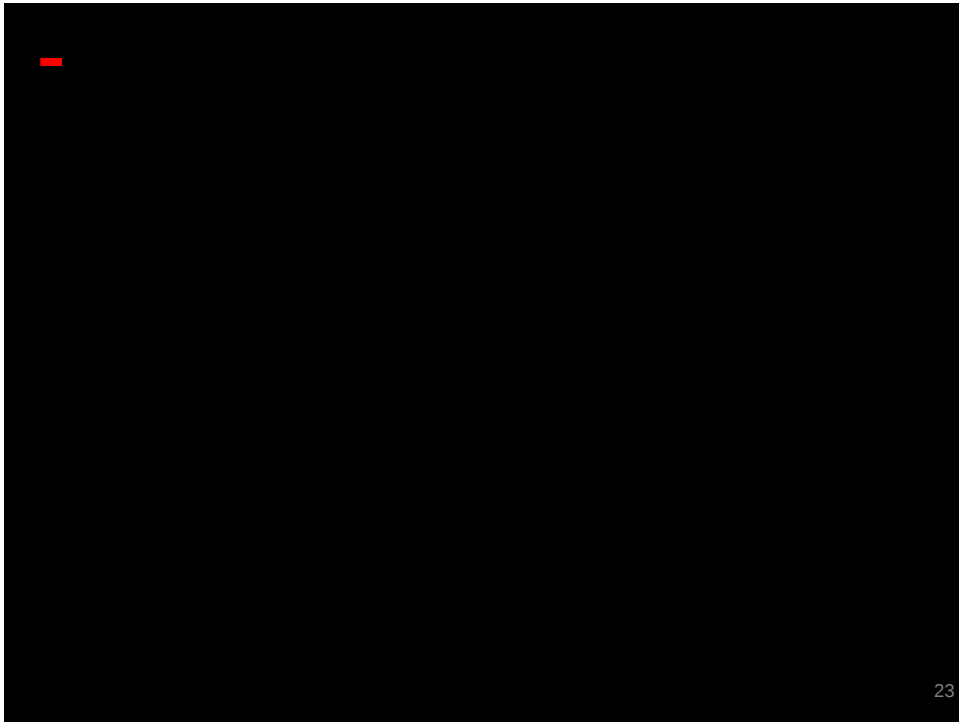
19

RIP
CONNOR

20



*"Experience is what you
get when you were
expecting something else"*
- Anon



a touch of anonymity

25



26

sell a LOT of cases

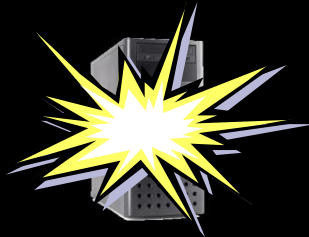
27



28

*"The **case** that stops the nation"*

29



10 minutes = \$1,000,000

35 years

31



32

2007

33

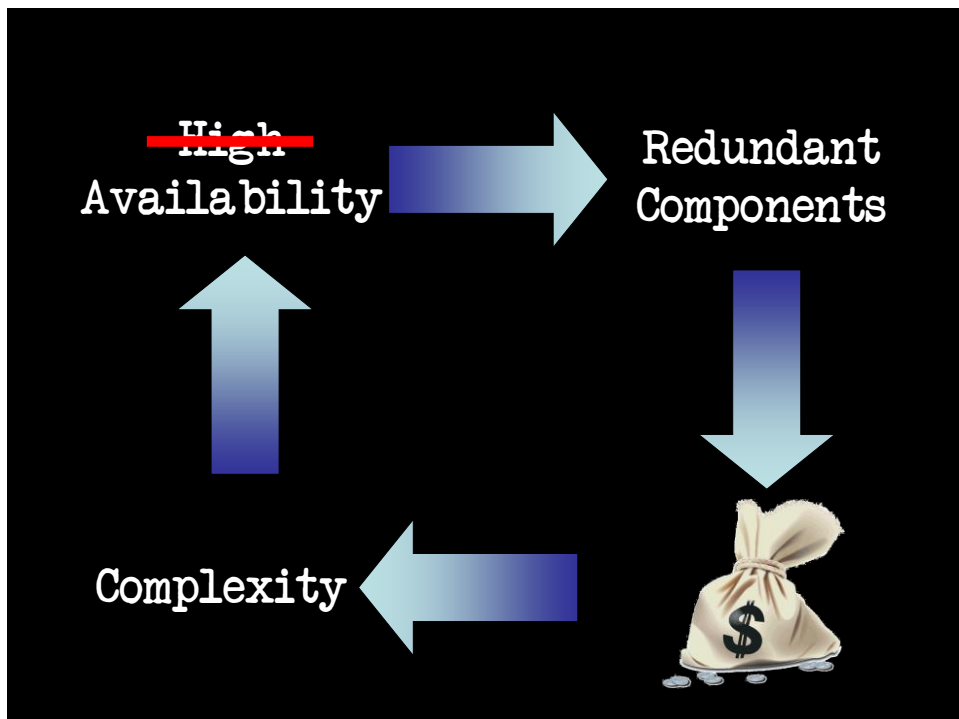
build new system

go-live Nov 2010

34

"challenge loop"

35



The Oracle logo, consisting of the word "ORACLE" in white, uppercase, sans-serif font, followed by a registered trademark symbol (®). The logo is centered within a red rectangular box, which is itself centered on a black background.

HA = RAC

RAC is complex

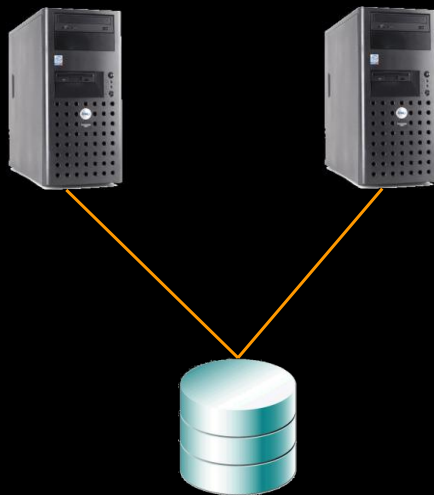
39

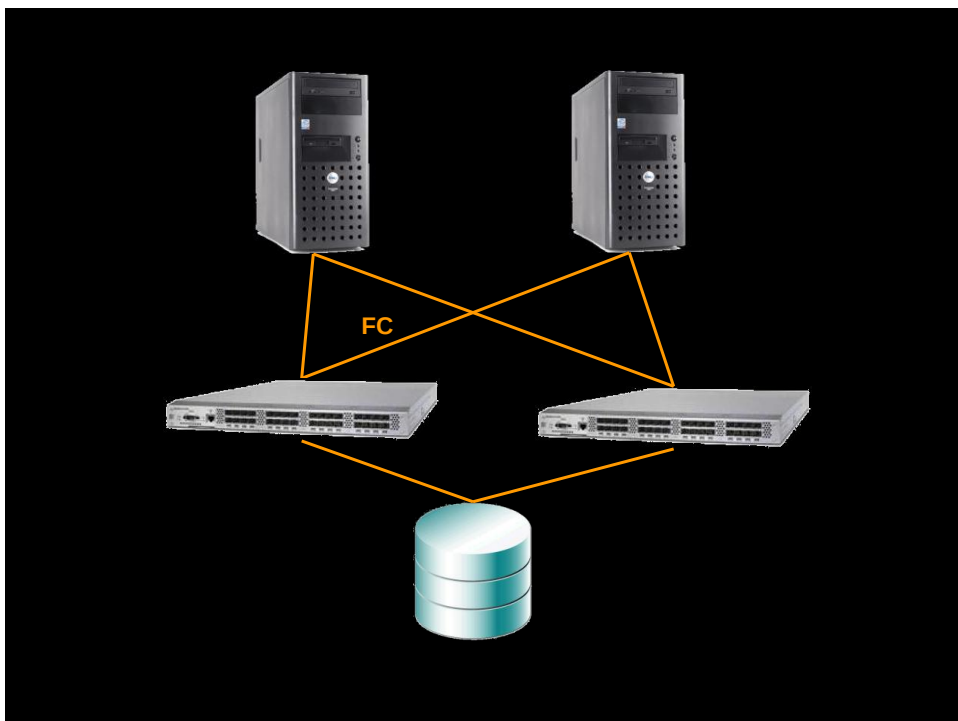
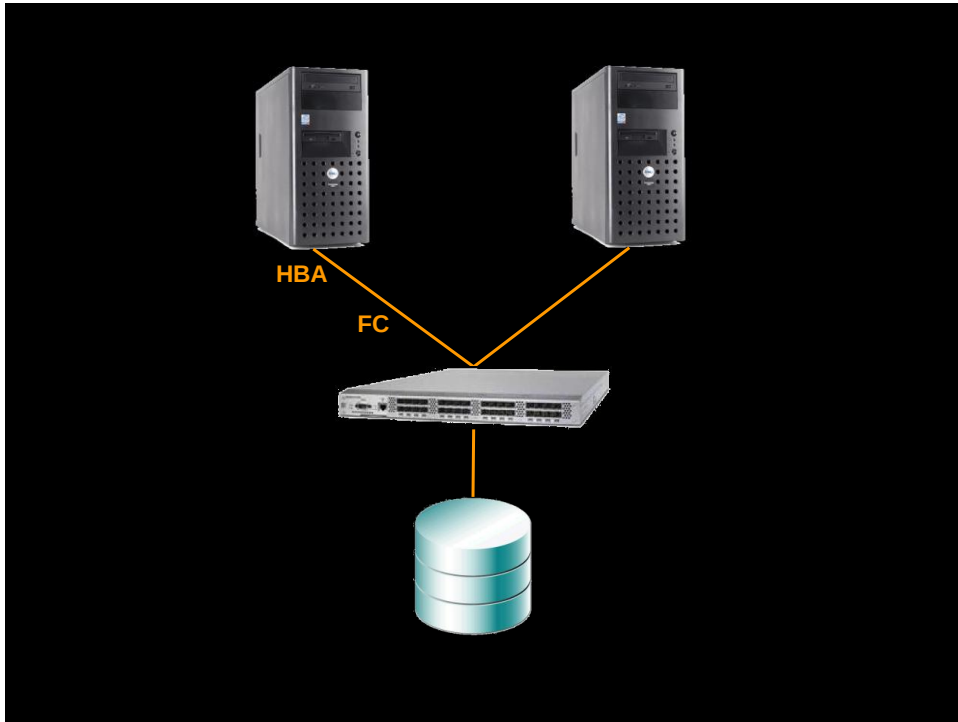
RAC is complex software

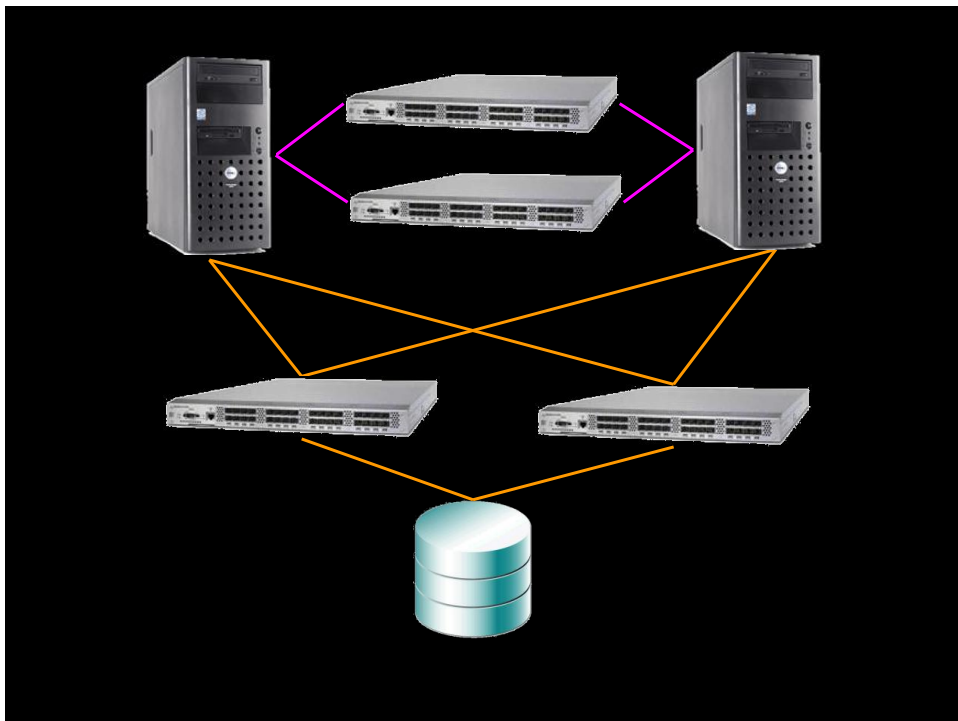
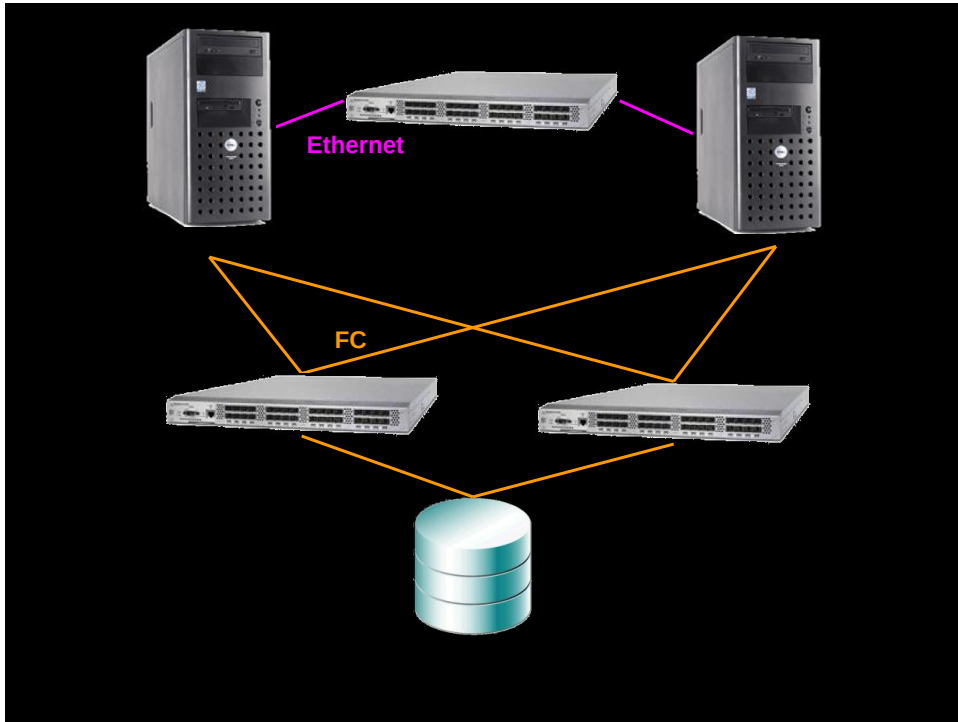
GPNPD GIPCD MDNSD EVMD EMVLOGGER
 OCRDUMP SRVCTL
 RAC GCS CLUVFY
 CRSCTL CSS LMD OCFS2
 LMS OCR PI
 OCRCHECK LCK OCSSD VIP OIFCFG
 RAC is complex software
 VIPCA DIAG FAN CRS
 EVMD LMON ONS
 BAST OCRCONFIG
 OCFS AST ASM GES
 TAF FCF GSD
 LKDEBUG CRS_STAT CSSDAGENT
 DISKMON

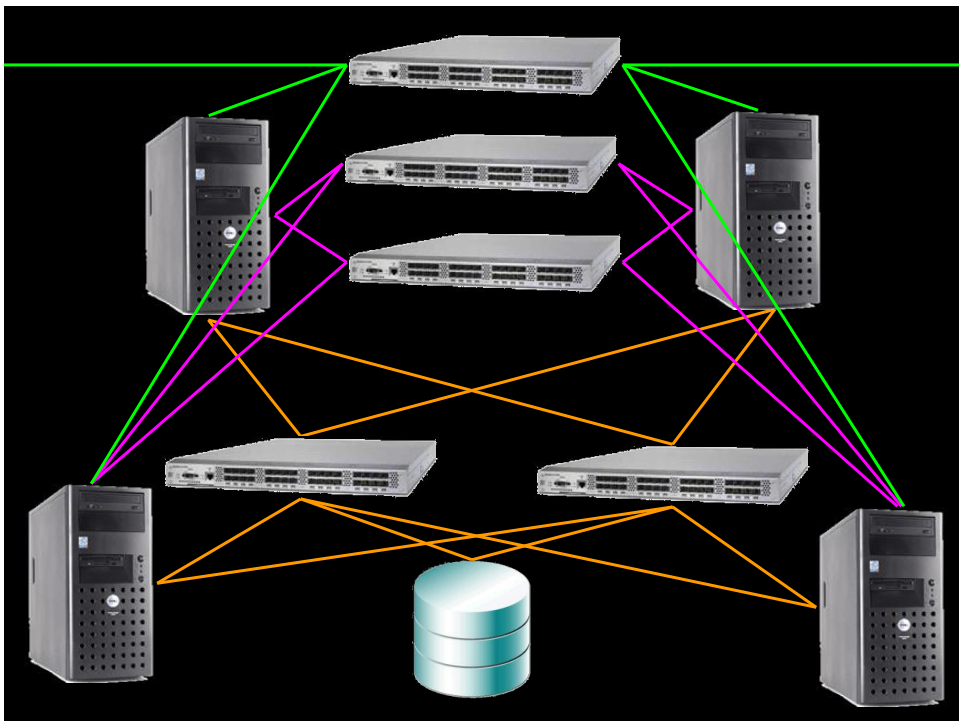
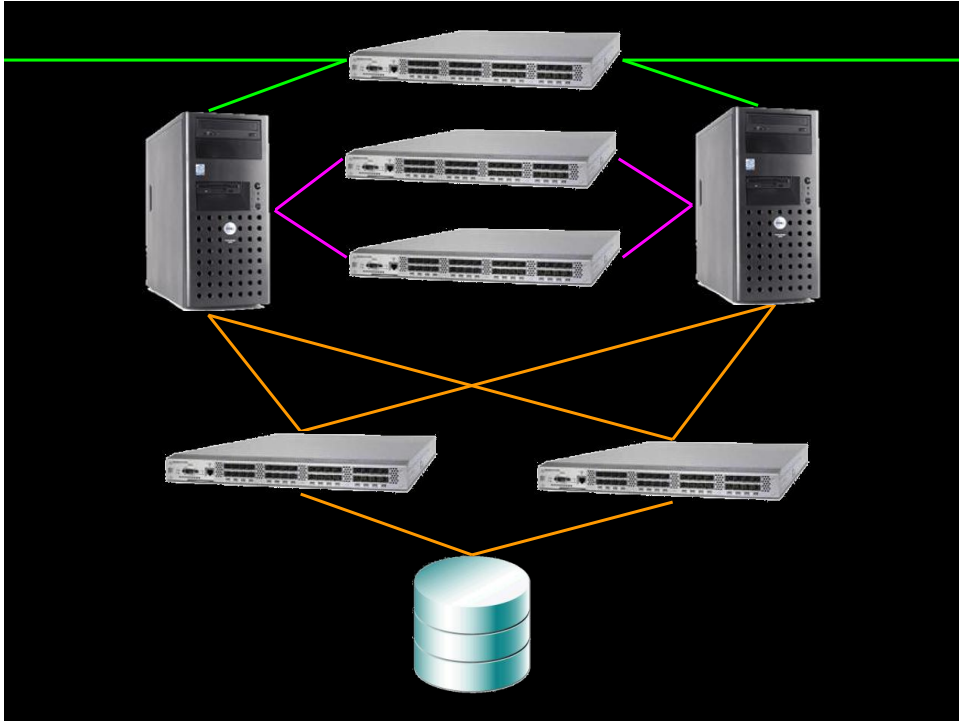
RAC is complex **hardware**

41



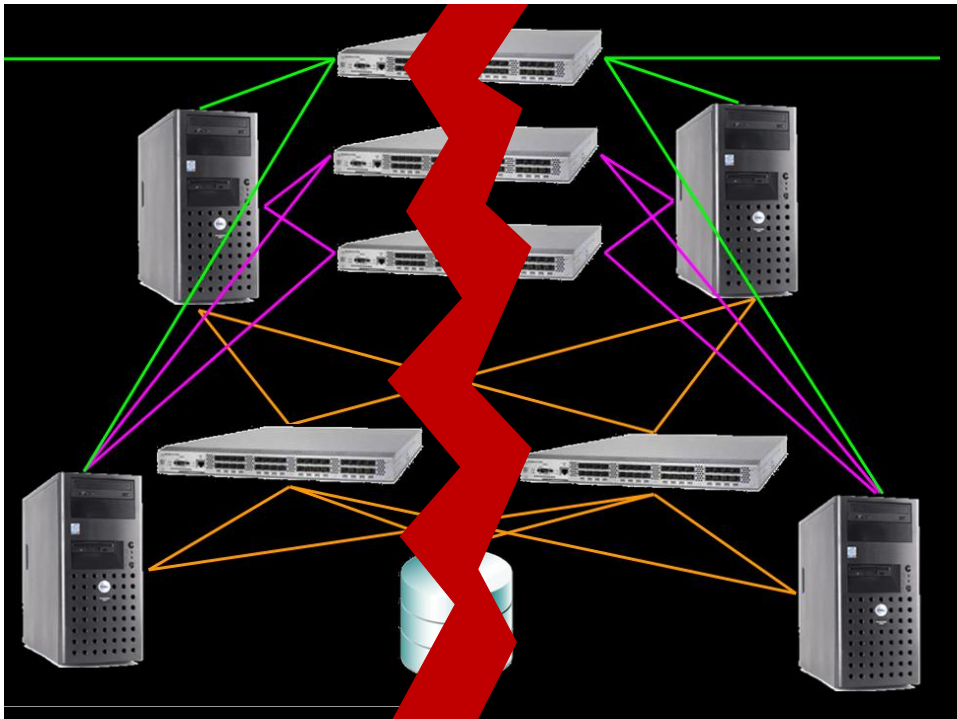






this client takes HA seriously

49

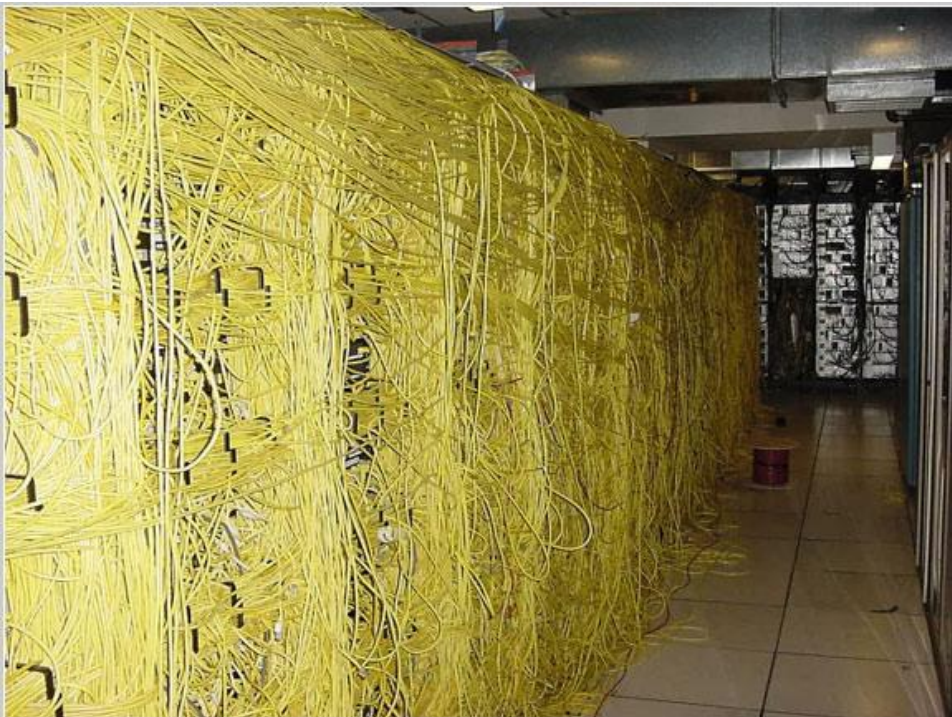
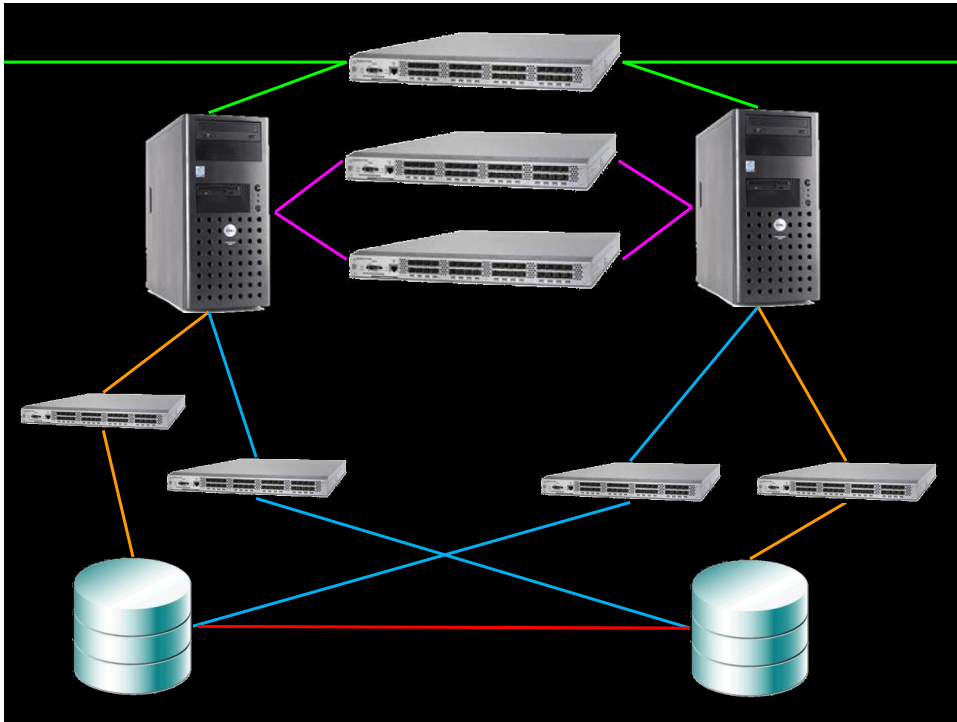


extended distance cluster

51

"stretch" cluster

52



other complexities

55

local file system

OS, crs

56

clustered file system

archives, admin

57

raw / ASM

database

58

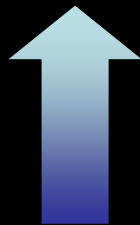
lots of technologies

lots of protocols

lots of file systems

59

~~High~~
Availability



Complexity

as simple (as you can)

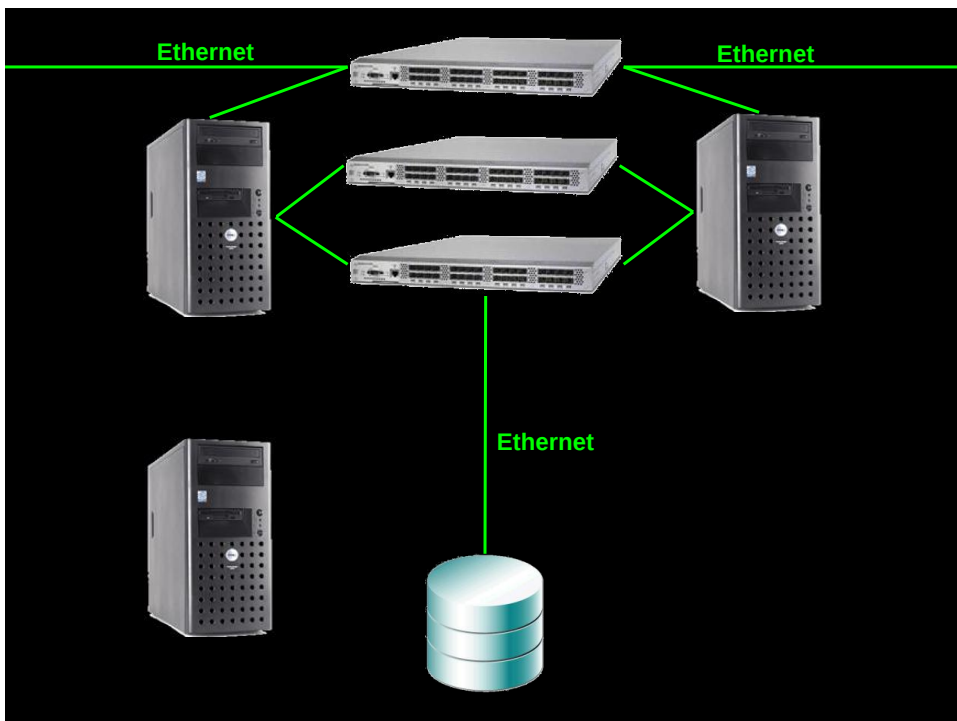
61

one protocol for everything

62

NFS

63



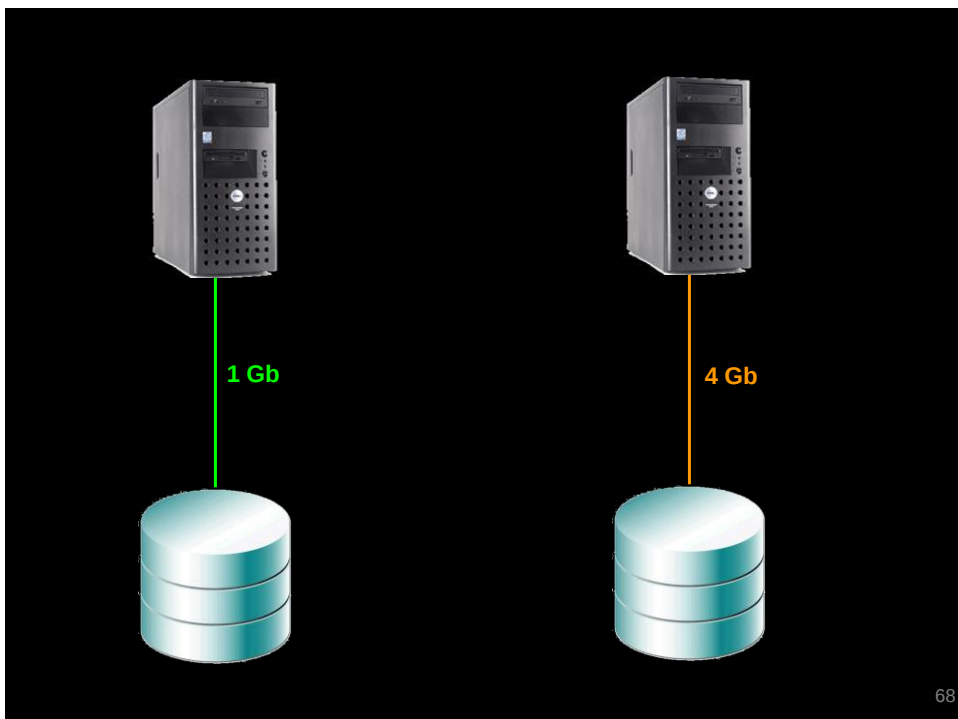


65



"but ethernet is slow"

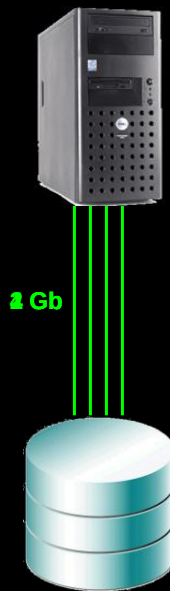
67



68

bonding

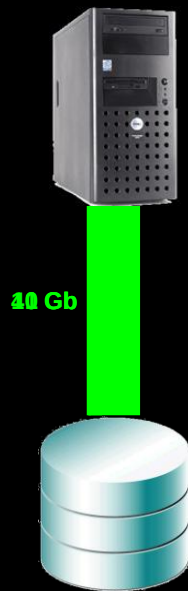
69



70

technology advances

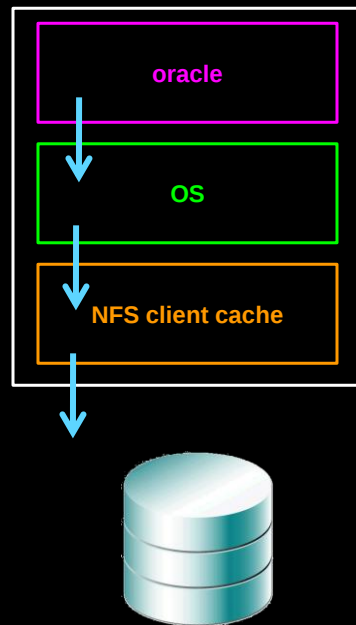
71



72

"but NFS is slow"

73



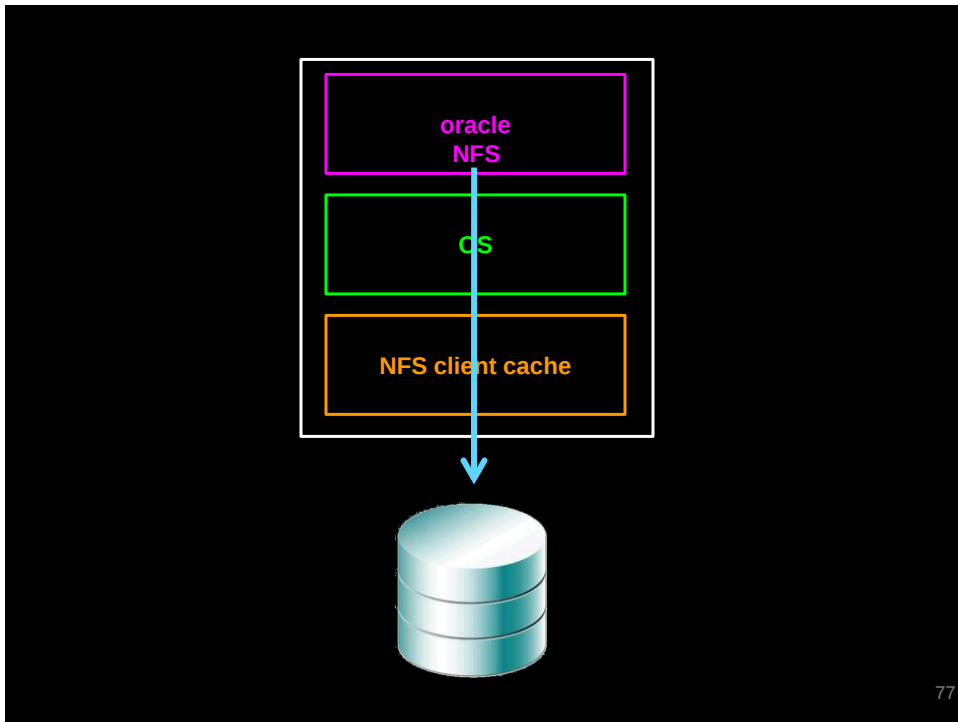
74

solution

75

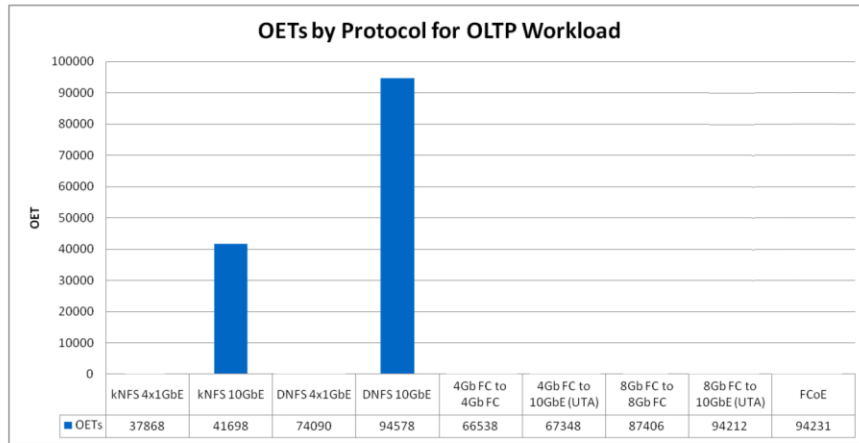
Direct NFS

76



Direct NFS rocks !

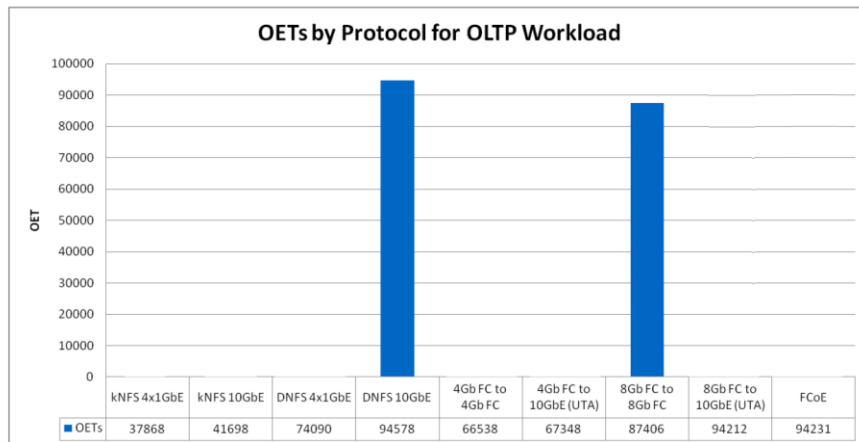
Figure 1) Application throughput comparison using OLTP workload (higher results are better).



Source:
Red Hat Enterprise Linux Protocol Performance Comparison with Oracle Database 11g Release 2 NetApp, TR-3932, June 2011

79

Figure 1) Application throughput comparison using OLTP workload (higher results are better).



Source:
Red Hat Enterprise Linux Protocol Performance Comparison with Oracle Database 11g Release 2 NetApp, TR-3932, June 2011

80

confirmed with local benchmark

81

<http://kevinclosson.wordpress.com>

82

except....

83

AIX 11.1.0.6

our platform / version

84

```

SQL> select sample_time, event, time_waited
2  from    V$ACTIVE_SESSION_HISTORY
3  where   event = 'db file sequential read'
4  and     time_waited > 1000000;

```

SAMPLE_TIME	EVENT	TIME_WAITED
23-MAY-10 08.33.27.255 PM	db file sequential read	60008772
23-MAY-10 07.39.48.301 PM	db file sequential read	60004661
23-MAY-10 07.23.34.517 PM	db file sequential read	60008839
23-MAY-10 06.52.05.100 PM	db file sequential read	60010558
23-MAY-10 06.44.49.290 PM	db file sequential read	60006266
23-MAY-10 06.44.43.280 PM	db file sequential read	60008222
23-MAY-10 06.22.51.793 PM	db file sequential read	60009589

85

500 bets / second

60 second I/O's are bad

86

`v$active_session_history`

(ASH)

87

`amazing`

88

before ASH

10g

89

Statspack

90

3rd party tools

91



92

"server CPU is 45%"

"buffer cache is 99%"

"no problems over 20mins"

93



94

too high level

95

sql trace

96

too low level

intrusive

97

"please repeat"

98

ASH

99

every active session...

100

...every second

101

```
SQL> select * from v$sgastat  
2  where name like 'ASH%';
```

POOL	NAME	BYTES
shared pool	ASH buffers	16777216

102

low overhead

103

MMNL

104

no latching !

105

```
SQL> select sample_time, event, time_waited
2  from   v$active_session_history
3  where  event = 'db file sequential read'
4  and    time_waited > 1000000;
```

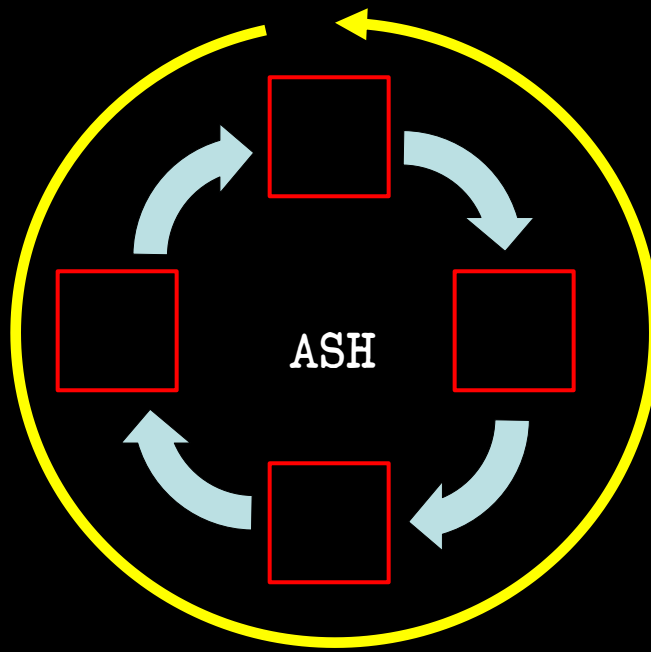
SAMPLE_TIME	EVENT	TIME_WAITED
23-MAR-11 08.33.27.255 PM	db file sequential read	60008772
23-MAR-11 07.39.48.301 PM	db file sequential read	60004661
23-MAR-11 07.23.34.517 PM	db file sequential read	60008839
23-MAR-11 06.52.05.100 PM	db file sequential read	60010558
23-MAR-11 06.44.49.290 PM	db file sequential read	60006266
23-MAR-11 06.44.43.280 PM	db file sequential read	60008222
23-MAR-11 06.22.51.793 PM	db file sequential read	60009589



106



107



108

and keep it !

10 second samples

109

very very cool

110

but...

111



diagnostic pack

worth every cent...

113

alternatives...

114



Snapper

Tanel Poder

<http://tech.e2sn.com/oracle-scripts-and-tools/session-snapper>

115

back to direct NFS...

60 second I/O's are bad

116

bug 6493018

117

thus our upgrade story begins...

118

recommendation

119

→
04/10

upgrade 11.1.0.6 to 11.1.0.7

upgrade #1

120

Readme for 11.1.0.7 AIX (crs)

Known Issues:

Does not work on AIX

121

→
26/05/10

upgrade 11.1.0.6 to 11.1.0.7.1

upgrade #2

122

(CRS) in-place upgrade

backout risk

123

solved in 11.2

124

upgraded without incident

125

~~60 second I/O's~~

direct NFS

3 second I/O's

126

26/05/10 - SR (sev 2)

02/06/10 - response ?

04/06/10 - response ?

back to Support

14/06/10 - escalate

16/06/10 - no response

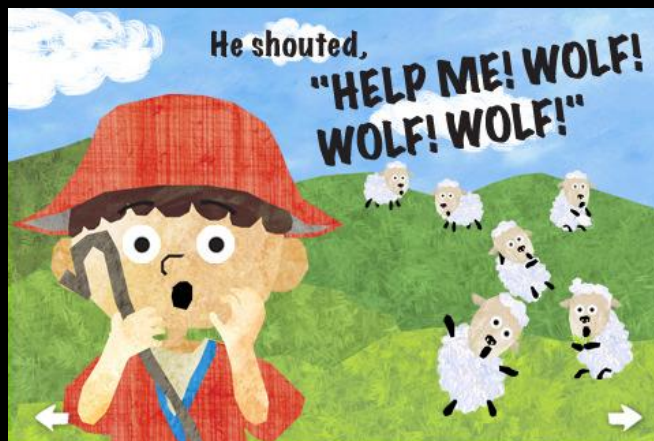
127

Severity 1

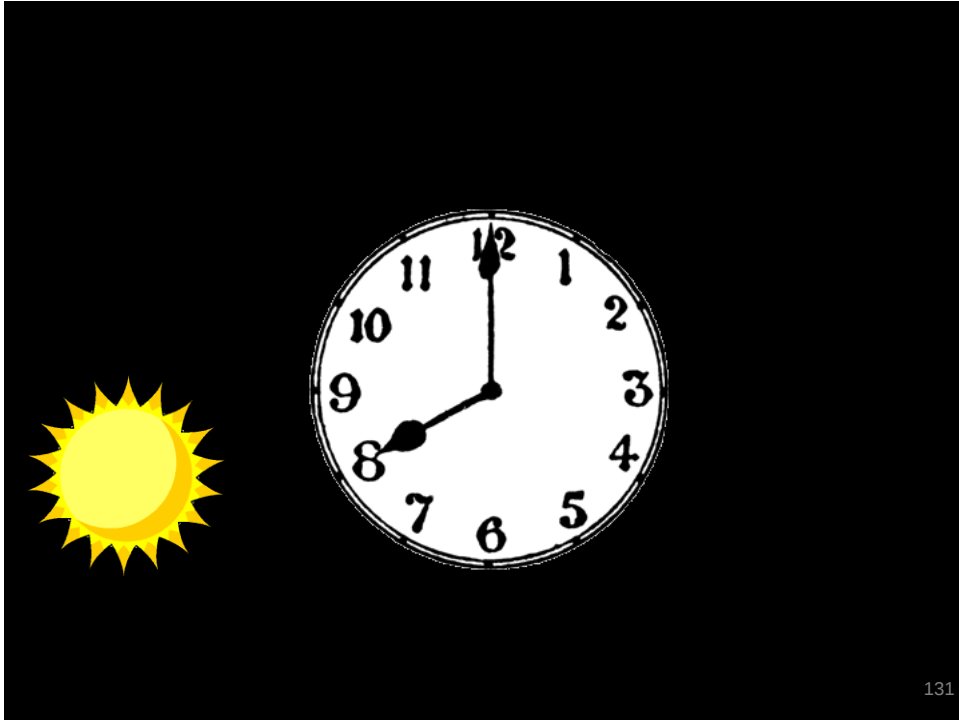
128

don't use Sev 1

129



130



stay with same region

132

June 17th

133

more than 11.1.0.7.1 needed

134

Bug 8418190 DIRECT NFS WARNINGS DURING DATABASE STARTUP
Bug 7682745 RUNNING OLTP WORKLOAD OVER DNFS CORRUPTS DATAFILE
Bug 7286541 11.1.0.7 : ORA-600 : [KSFD_ODMIO1] , [KSFD_ODMRSZ2]
Bug 7827453 DIRECT NFS STALE FILE ERROR RESULTING IN I/O FAILURES
Bug 8579034 DIRECT NFS:READ FAIL STATUS 70 ERROR
Bug 8551597 RMAN CREATES HUGE TRACEFILE TO DNFS FILESYSTEM
Bug 8206872 DNFS ODM ERR:KGNFS OPERATION INPROGRESS ON FAILOVER
Bug 8535912 MERGE LABEL REQUEST ON TOP OF 11.1.0.7
Bug 8565508 MERGE LABEL REQUEST ON TOP OF 11.1.0.7
Bug 8587954 MERGE LABEL REQUEST ON TOP OF 11.1.0.7

135

always clarify...

136

always confirm...

137

patch 8768632

direct NFS

138

~~11.1.0.7.1~~

11.1.0.7.3

139



22/07/10

upgrade 11.1.0.7.1 to 11.1.0.7.3
(database only)

upgrade #3

140

```
try apply patch 8768632
```

141

```
Invoking OPatch 11.1.0.6.9
```

```
Oracle Interim Patch Installer version 11.1.0.6.9  
Copyright (c) 2009, Oracle Corporation. All rights reserved.
```

```
PREREQ session
```

```
Oracle Home : /u01/app/oracle/product/11.1.0/db_1  
Central Inventory : /u01/app/oracle/oraInventory
```

```
Invoking prereq "checkconflictagainsthwithdetail"
```

```
ZOP-40: The patch has conflicts/supersets with other patches.
```

```
Prereq "checkConflictAgainstOHWithDetail" failed.
```

142

patch 8768632

≠

11.1.0.7.3

143

stay on 11.1.0.7.3

144

mutexes

145

quick background

146

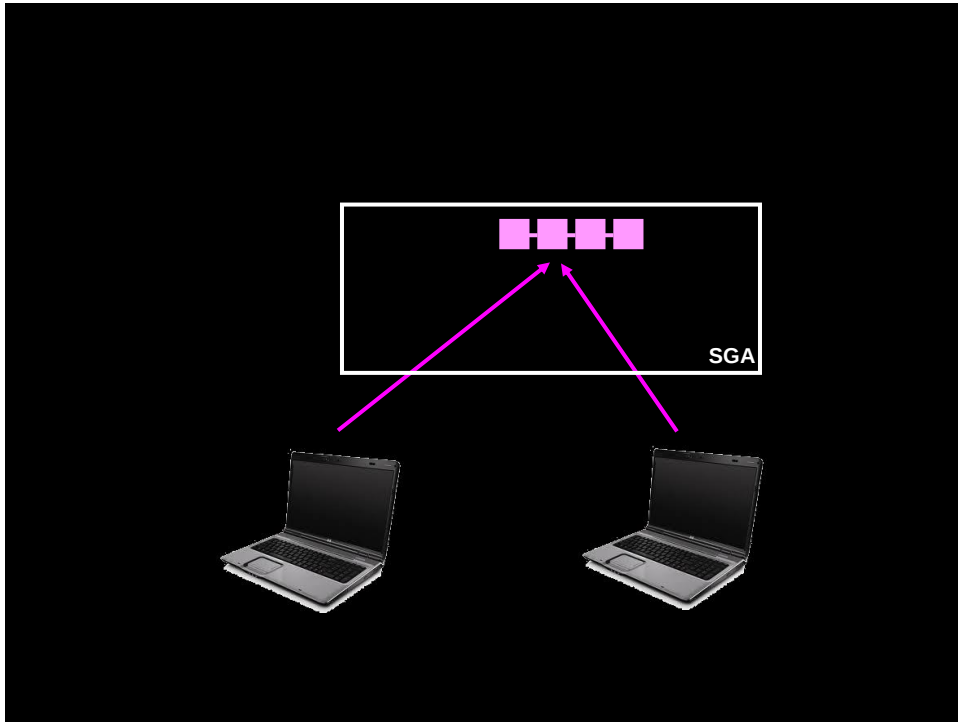
woeful metaphor

147





same with memory

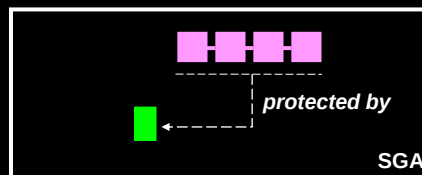


latches

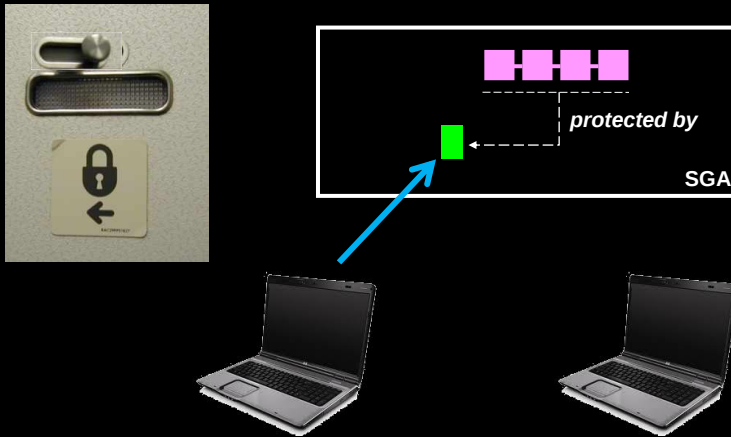
serialised access

aka, someone waits

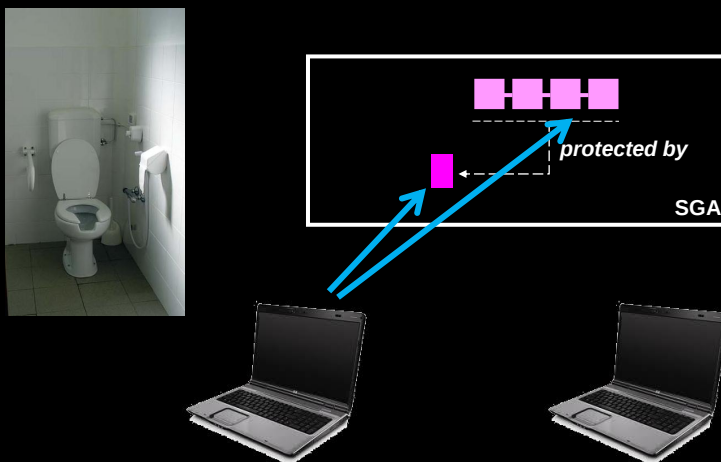
153



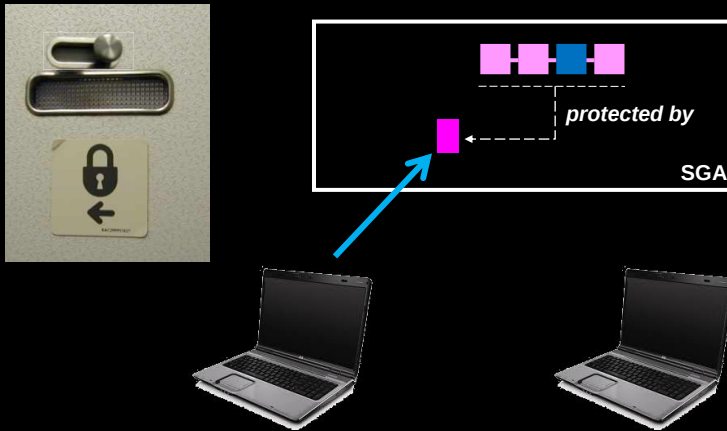
1) get latch



2) access memory



3) release latch



what protects the latch ?

latch acquisition

159

atomic (CAS)

160

library cache

161

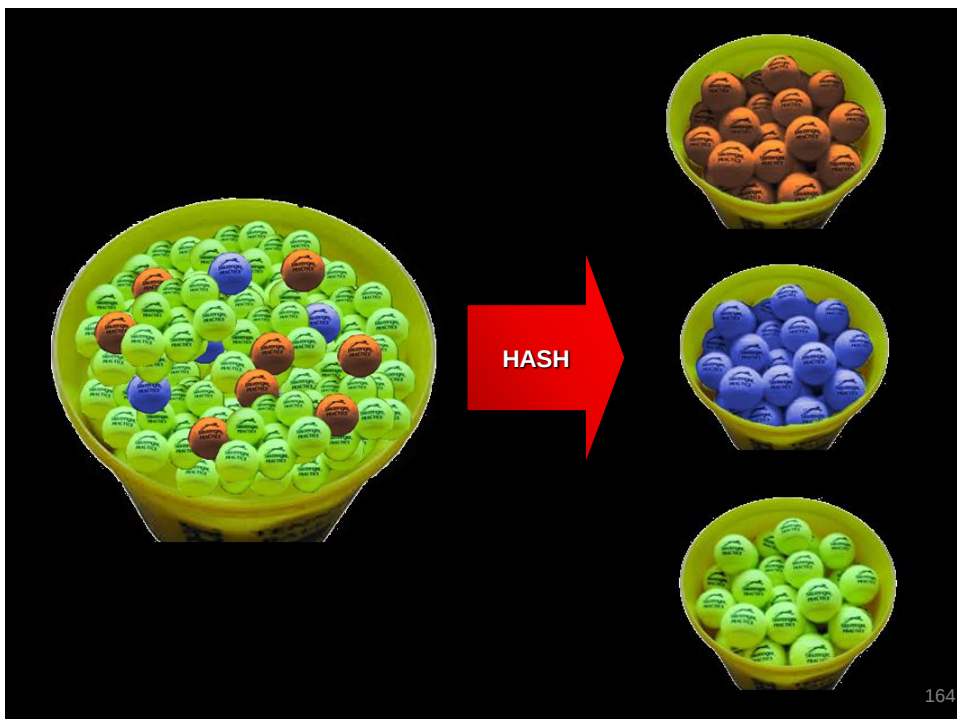
lots of stuff

162

```
SQL> select count(*)  
       2 from v$db_object_cache
```

```
COUNT (*)  
-----  
    39,532
```

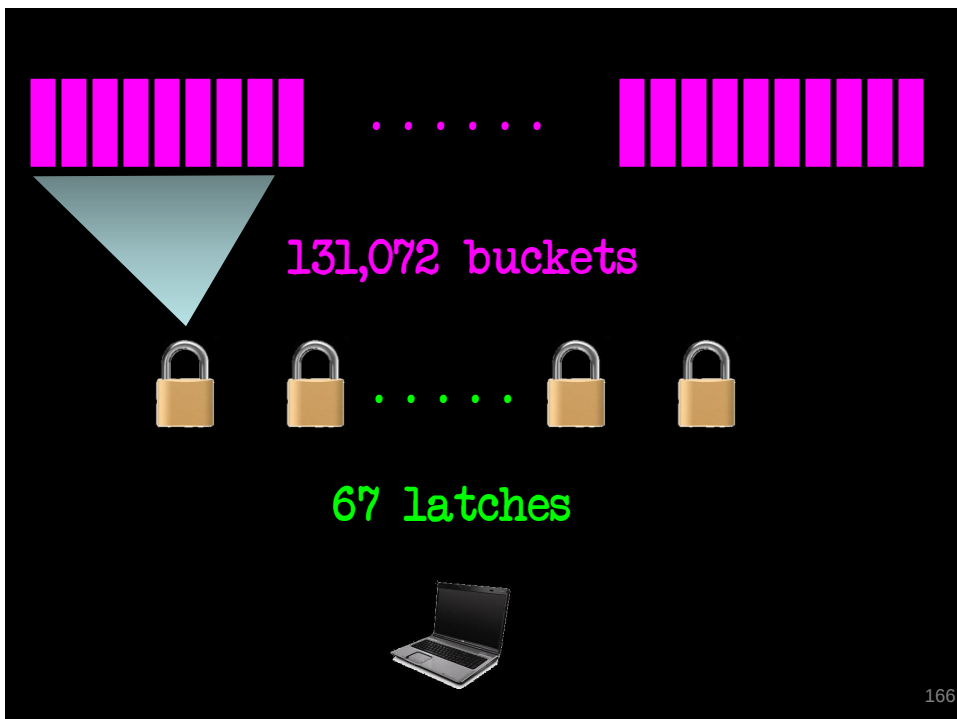
163



164

same with library cache

165



166

(false) latch contention

167

fine for good apps

168

`bad` apps ... latch contention

169

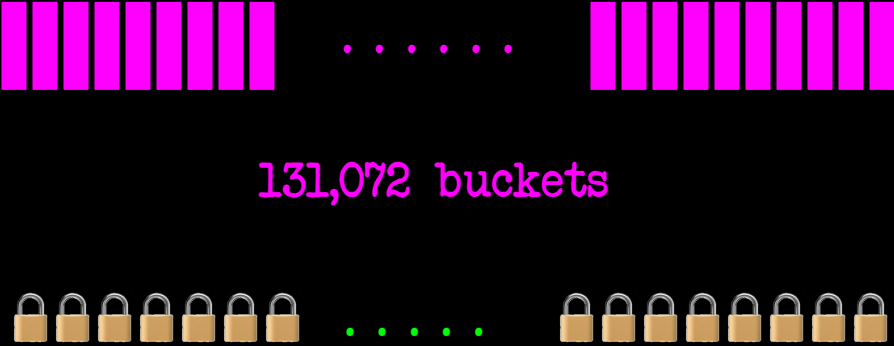


llg ... mutexes

171

lightweight, smaller, faster

172



131,072 buckets

131,072 mutexes



173

lightweight

smaller

faster

new

issues....

174

The screenshot shows the Oracle My Oracle Support Knowledge Base search results for the query "library cache: mutex X". The search results are displayed in a list format, showing various bugs and articles related to the "library cache: mutex X" issue. The left sidebar contains a "Refine Search" section with filters for "All Sources", "Knowledge Base", "All Bugs", "Communities", "Documentation", "Sun System Handbook", "Patches", "Product Category", "All Products", "Enterprise Management (1)", "Internal Applications (1)", "JD Edwards EnterpriseOne (2)", "Middleware (4)", "More Applications (4)", "Oracle Database Products (114)", "PeopleSoft Enterprise (1)", "Siebel (1)", "Sun Microsystems (11)", "Tools/Utilities", "Document Types", and "Updated".

The search results list includes the following entries:

- Oct 20, 2011** **WAITEVENT: "library cache: mutex X"**
In this Document Purpose Scope WAITEVENT: "library cache: mutex X" Definition: Individual Waits: Systemwide
Tags: buging_address [Article ID 727400.1]
- Oct 20, 2011** Bug 9530750 - High waits for "library cache: mutex X" for cursor Build lock.
Bug 9530750 High-wait for "library cache: mutex X" for cursor Build lock This note gives ... Waits for "library cache: mutex X" Stack
[Article ID 9530750.8]
- Sep 17, 2011** Bug 10145558 - Selects on library cache VS/XS views cause "library cache: mutex X" waits
Bug 10145558 Selects on library cache VS/XS views cause "library cache: mutex X" waits This note gives ... Waits for "library cache: mutex X" VSQ
[Article ID 10145558.8]
- Jan 17, 2011** Bug 9239863 - Excessive "library cache:mutex X" contention on hot objects
Bug 9239863 Excessive "library cache:mutex X" contention on hot objects This note gives ... Waits for "library cache: mutex X" _jgl
[Article ID 9239863.8]
- Sep 23, 2011** "Library Cache: Mutex X" On Koka Cursors (LOBs) Non-Shared :
11.1.0.7 Performance 11.1.0.x Child Cursors Library Cache: Mutex X Waitefc child completion NLS_LENGTH_SEMANTICS
[Article ID 758674.1]
- Jul 20, 2010** Bug 7155860 - Spin on lgldde causes "library cache: mutex X" This note gives a brief overview of bug ... Waits for "library cache: mutex X" (none)
Bug 7155860 Spin on lgldde causes "library cache: mutex X" This note gives a brief overview of bug ... Waits for "library cache: mutex X" (none)
[Article ID 7155860.8]
- May 27, 2011** Bug 9312879 - "library cache: mutex x" waits after killing sessions / PMON slow to clean up
Bug 9312879 "library cache: mutex x" waits after killing sessions / PMON slow to clean up ... Waits for "library cache: mutex x" Waits
[Article ID 9312879.8]
- Sep 10, 2010** Bug 9282521 - Excessive "library cache:mutex X" contention on hot objects
Bug 9282521 Excessive "library cache:mutex X" contention on hot objects This note gives ... Waits for "library cache: mutex X" _jgl
[Article ID 9282521.8]
- Jan 27, 2011** Bug 9398685 - High "library cache: mutex X" when using Application Context
Bug 9398685 High "library cache: mutex X" when using Application Context This note ... Waits for "library cache: mutex X" Application
[Article ID 9398685.8]
- Nov 20, 2010** Bug 7307972 - Excessive waits on "library cache: mutex x" This note gives a brief overview bug 7307972 ... Waits for "library cache: mutex X" (none)
Bug 7307972 Excessive waits on "library cache: mutex x" This note gives a brief overview bug 7307972 ... Waits for "library cache: mutex X" (none)
[Article ID 7307972.8]
- Jan 11, 2011** Bug 5928271 - Excessive waits on "library cache: mutex X"
Bug 5928271 Excessive waits on "library cache: mutex X" This note gives a brief overview of bug ... Waits for "library cache: mutex X" (none)
[Article ID 5928271.8]
- Oct 6, 2009** Bug 7226463 - EXECUTE IMMEDIATE no releasing mutex or library cache pin This note gives a brief overview of
Bug 7226463 EXECUTE IMMEDIATE no releasing mutex or library cache pin This note gives a brief overview of
[Article ID 7226463.8]
- May 27, 2011** Bug 7502237 - Unnecessary "library cache: mutex X" waits using stored Java

175

"case that stops the nation"

load test

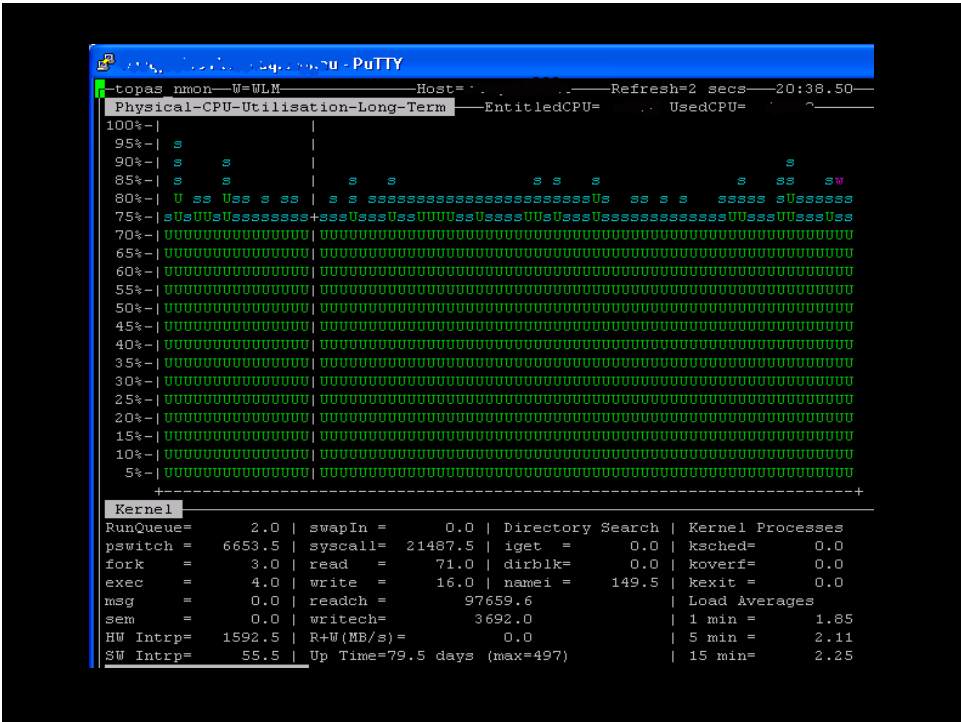
176

AIX ... nmon

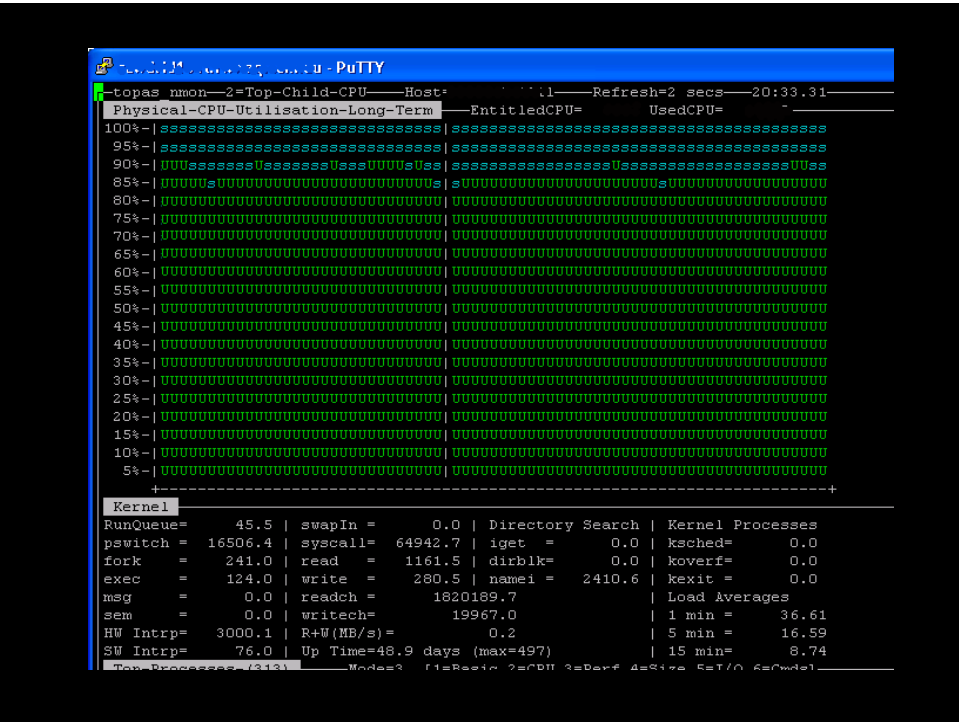
177

11.1.0.7.1

178



11.1.0.7.3



"case that stops the nation"
server

Bug 9711070:

WAITS ON LIBRARY CACHE: MUTEX X WHEN CALLING PL/SQL PACKAGES

183

11.1.0.7.1 to 11.1.0.7.3

184

no **minor** upgrades anymore

185

back to Support

186

"you need to upgrade to 11.1.0.7.2"

187

11.1.0.7.3

188

"you need to upgrade to 11.1.0.7.4"

189

→
04/09/10

upgrade 11.1.0.7.3 to 11.1.0.7.4

upgrade #4

190

October 2010

one month to go-live

191

nothing... changed

192

"you need to upgrade to 11.1.0.7.4"

193

"we are on 11.1.0.7.4 !!!!"

194

escalation

195

don't sit on poor support

we did

196

account managers

local office

197

"there is no bug...

... you have too much PL/SQL"

198

instrumentation

199

instrumentation is gold !

200



"how much
overhead in
instrumentation"

201

"minus 10%"

202

"10% slower ...?"

203

"No ... 10% *faster*..."

204

Pro*C layer

205

```
begin
    dbms_session.set_identifier( <tag> );
    betting.save_bet(<parms>);
    dbms_session.clear_identifier;
end;
```

206

use the **Oracle** facilities

`dbms_session.set_identifier`

207

```
SQL> select TABLE_NAME from DBA_TAB_COLUMNS
       2  where COLUMN_NAME in ('CLIENT_ID','CLIENT_IDENTIFIER')
       3  order by 1;
```

TABLE_NAME

ALL_SCHEDULER_JOBS	GV_\$ACTIVE_SESSION_HISTORY
ALL_SCHEDULER_JOB_LOG	GV_\$CLIENT_STATS
ALL_SCHEDULER_WINDOW_LOG	GV_\$GLOBALCONTEXT
DBA_AUDIT_EXISTS	GV_\$LOGMNR_CONTENTS
DBA_AUDIT_OBJECT	GV_\$SESSION
DBA_AUDIT_SESSION	GV_\$SQL_MONITOR
DBA_AUDIT_STATEMENT	V_\$ACTIVE_SESSION_HISTORY
DBA_AUDIT_TRAIL	V_\$CLIENT_STATS
DBA_COMMON_AUDIT_TRAIL	V_\$DIAG_ALERT_EXT
DBA_FGA_AUDIT_TRAIL	V_\$GLOBALCONTEXT
DBA_HIST_ACTIVE_SESS_HISTORY	V_\$LOGMNR_CONTENTS
DBA_SCHEDULER_JOBS	V_\$SESSION
DBA_SCHEDULER_JOB_LOG	V_\$SQL_MONITOR
DBA_SCHEDULER_WINDOW_LOG	

208

PLSQL layer

209

```
procedure betting.save_bet(...) is
begin

    msg('Starting bet placement');
    msg('  p_txn_timestamp = '||p_txn_timestamp);
    msg('  p_bet_no       = '||p_bet_no);

    select ...

    update ...

    msg('updated '||sql%rowcount||' rows');

end;
```

210

msg = local plsql table

211

```

procedure msg(...) is
begin

  l_request := sys_context('DEBUG_CTX',
                          sys_context('USERENV','SESSIONID'));

  if l_request is not null then

    if instr(l_request,'ENABLE') > 0 then
      g_debug := true;

    elsif instr(l_request,'FLUSH') > 0 then
      dump_to_db_table;

    elsif instr(l_request,'OUTPUT') > 0 then
      g_output := true;

    elsif l_request = 'DISABLE' then
      g_debug := false;

  end if;

```

212

```
...  
  
if g_debug then  
    insert into ...  
  
if g_output then  
    dbms_output.put_line ...  
  
<etc>  
  
<etc>  
  
end;
```

213

global context

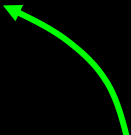
214

cross session

215

```
SQL> create or replace  
2 context GLOBCTX using ctxpkg  
3 ACCESSED GLOBALLY  
4 /
```

Context created.



instance wide

```
SQL> create or replace
  2 package CTXPKG as
  3
  4   procedure set_global_val(p_val number);
  5
  6 end;
  7 /
```

Package created.

```
SQL> create or replace
  2 package body CTXPKG as
  3
  4
  5   procedure set_global_val( p_val number)
  6   is begin
  7       dbms_session.set_context(
  8           NAMESPACE=>'GLOBCTX',
  9           ATTRIBUTE=>'OBJ_ID',
10           VALUE     =>p_val,
11           USERNAME  =>null,
12           CLIENT_ID=>'DEBUG');
13   end;
14 end;
15 /
```

Package body created.

```
SQL> exec ctxpkg.set_global_val(123123);
```

PL/SQL procedure successfully completed.

```
SQL> begin
  2   dbms_output.put_line(
  3       sys_context('GLOBCTX','OBJ_ID'));
  4   end;
  5   /
```

123123

```
SQL> begin
  2   dbms_output.put_line(
  3       sys_context('GLOBCTX','OBJ_ID'));
  4   end;
  5   /
```

123123

very fast (11g)

no dual access

BUG 9398685:
HIGH "LIBRARY CACHE: MUTEX X" WHEN USING APPLICATION CONTEXT

221

compromised...

222

```

begin

    dbms_session.set_identifier( <tag> );

    betting.save_bet(<parms>);

    dbms_session.clear_identifier;

end;

```

223

```

procedure betting.save_bet(...) is
begin

    $IF $$debug_on $THEN
        msg('Starting suitcase purchase);
        msg('  p_txn_tstamp = '||p_txn_tstamp);
        msg('  p_bet_no      = '||p_bet_no);
    $END

```

224

conditional compilation

careful with package state

225

November 15th, 2010

live on 11.10.7.4

226

two weeks **after** ...

227

"The case that stops the nation"

2010

228



229

mutex issue ?

230

351 days....

231

January 2011

232

bug 9239863:

excessive "library cache mutex X" contention

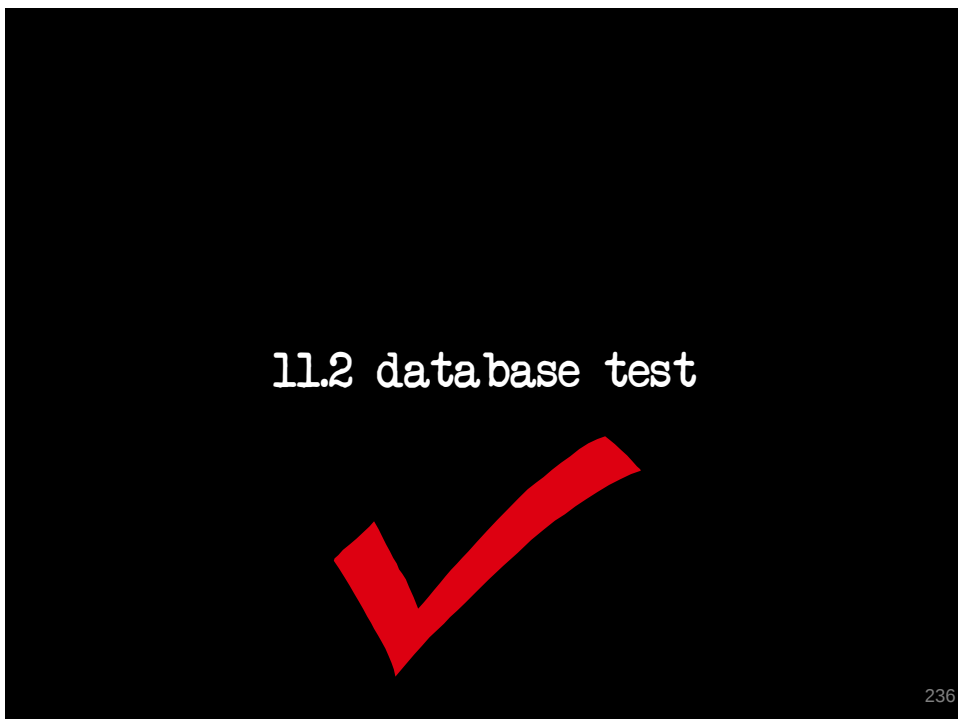
Affects: 11.1.0.7
11.2.0.1

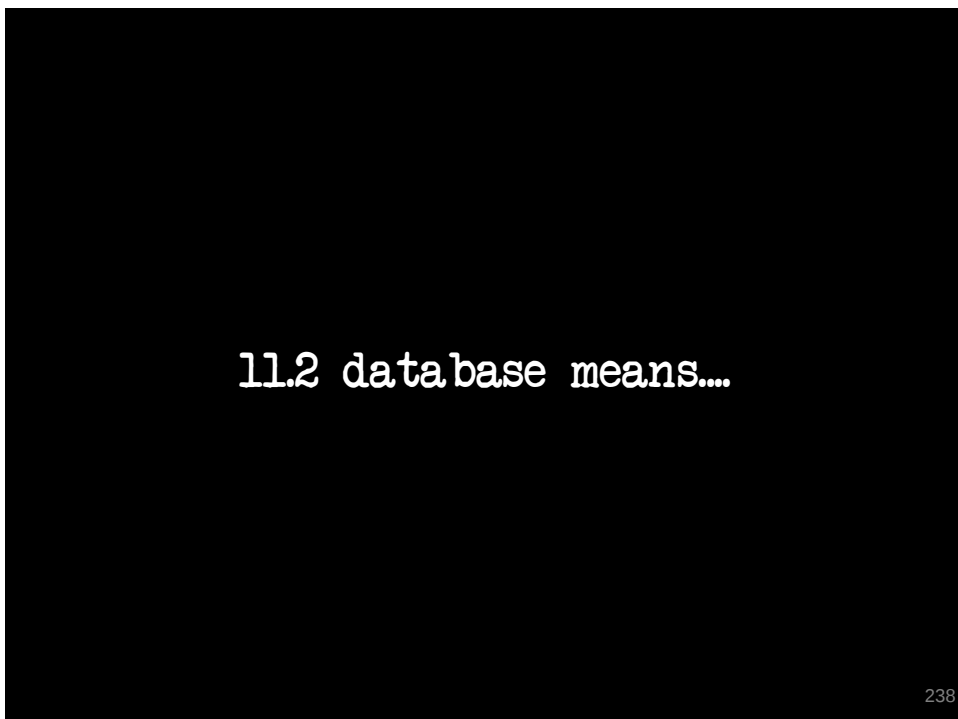
Fixed: 11.2.0.2

233

<http://andreynikolaev.wordpress.com>

234





11.2 clusterware

239

big change

240

10g

cssd

crsd

evmd

241

oclsomon

crsd

oclsomon

ocssd

11g

evmd

evmlogger

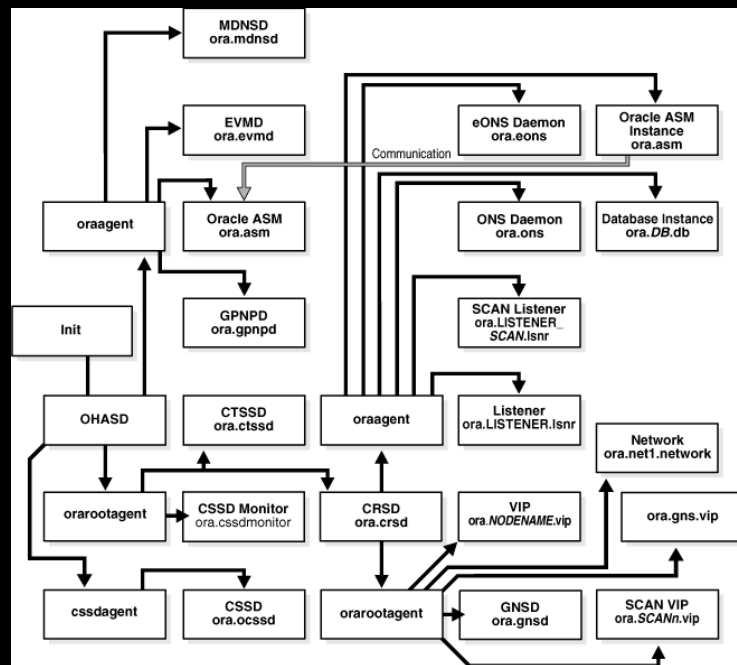
ocssd

oproc

242

11.2

243



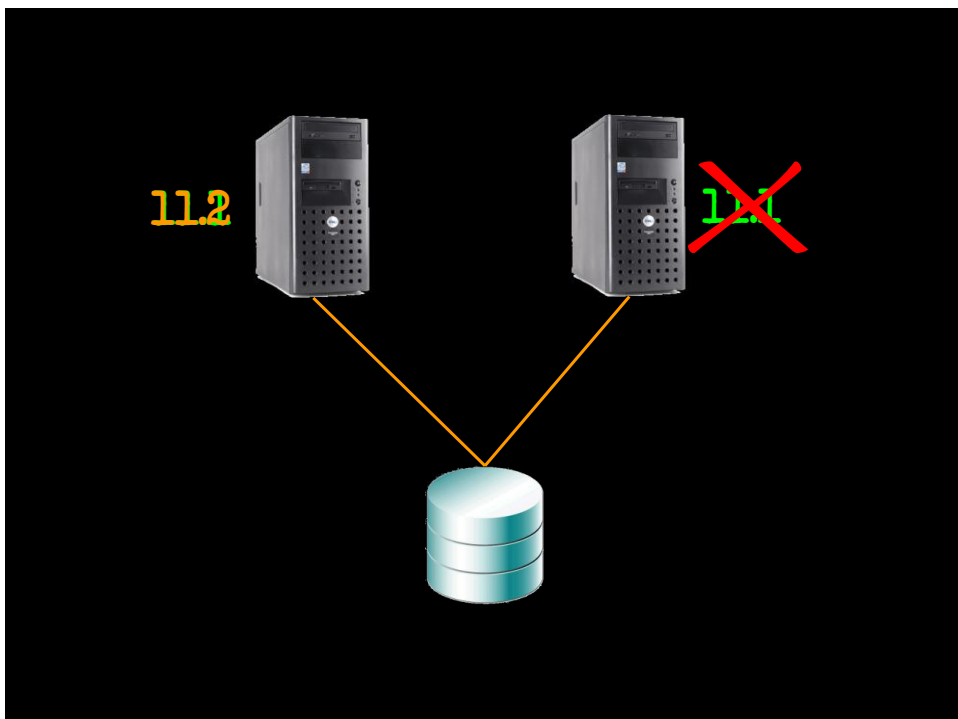
244

04/02/11

upgrade 11.1.0.7.4 to 11.2.0.2
(crs)

*upgrade #5
(dev cluster 1)*

245



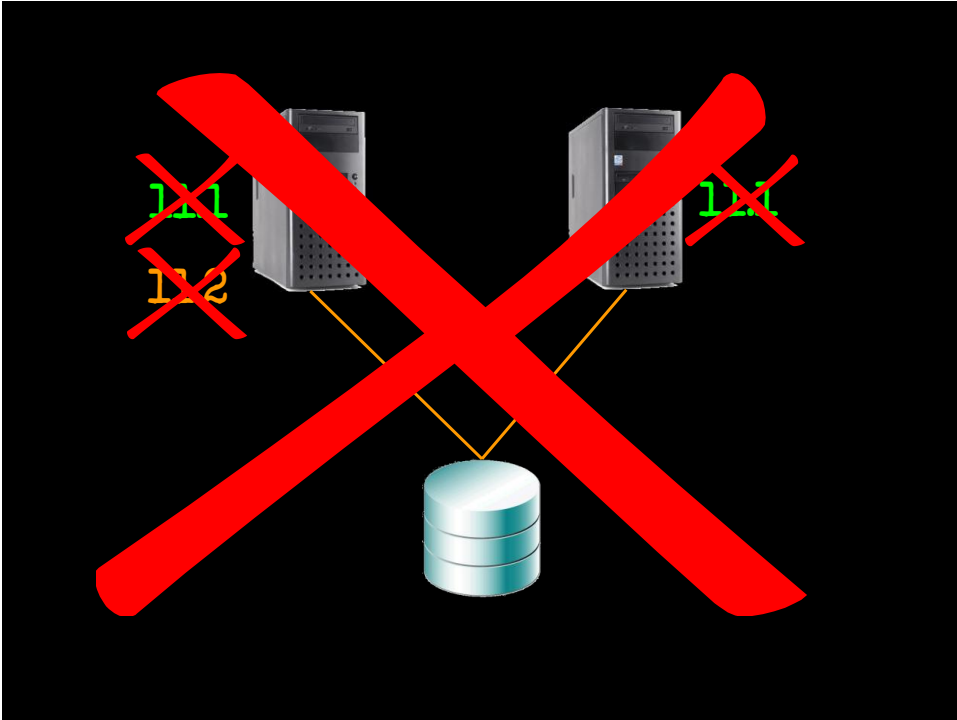
haip

247

11.2 deinstall

... Support recommended

248



International RAC user group

multicast problem

230.0.1.0

251

patch 9974223

224.0.0.251

252

download multicast utility !

or 11.2.0.3

253

→
14/02/11

upgrade 11.1.0.7.4 to 11.2.0.2
(crs, take 2)

*upgrade #6
(dev cluster 1)*

254

success !

255

db upgrade fine !

no mutex issues

256

April 1st

257



01/04/11

upgrade 11.1.0.7.4 to 11.2.0.2
(crs + patch 9974223)

upgrade #7
(dev cluster 2)

258



new patch 10370797



12/04/11

upgrade 11.1.0.7.4 to 11.2.0.2
(crs + patch 9974223 + 10370797)

upgrade #7
(dev cluster 2)

261

Invoking OPatch 11.2.0.5

PREREQ session

ZOP-40: Patch 9974223 conflicts with patch 10370797

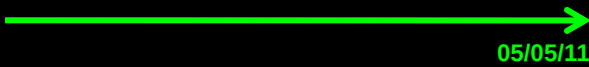
Prereq "checkConflictAgainstOHWithDetail" failed.

262

11.2.0.2.2

fixes merged

263



05/05/11

upgrade 11.1.0.7.4 to 11.2.0.2.2
(crs)

upgrade #8
(dev cluster 2)

264

success !

um.....not quite

265

database upgrade

"Database upgrade cannot proceed.

Cluster is not healthy"

266

back to support...

267

"Database upgrade cannot proceed."

"please proceed...."

7 days later...

268

OTN forum

entire cluster shutdown

269

cluster 1: 11.2.0.2 + patch

cluster 2: 11.2.0.2.2

success !

270



31/05/11

upgrade 11.1.0.7.4 to 11.2.0.2.2
(crs)

upgrade #9
(uat cluster 1)

271

success !

first time !

272

cluster 1: 11.2.0.2 + patch

cluster 2: 11.2.0.2.2

cluster 3: 11.2.0.2.2

success !

273

until....

274

```
SQL> startup
ORA-3113: Instance terminated, Disconnection forced.

PMON (ospid: 770146): terminating the instance due to error 481

Instance terminated by PMON, pid = 770146

kjzdattdlm: Can not attach to DLM (LMON up=[FALSE], DB mounted=[FALSE]).
kjzdattdlm: Can not attach to DLM (LMON up=[TRUE], DB mounted=[FALSE]).
kjzdattdlm: Can not attach to DLM (LMON up=[TRUE], DB mounted=[FALSE]).
```

275

half cluster

276

recall...

277

new patch 10370797

278

merged with 11.20.22 PSU

279

back to Support

280

"actually.....no"

281

"please apply 10370797...

...on top of 11.2.0.2.2"

282



283

```
Invoking OPatch 11.2.0.5
```

```
PREREQ session
```

```
ZOP-40: Patch 10370797 conflicts with 11.2.0.2.2
```

```
Prereq "checkConflictAgainstOHWithDetail" failed.
```

284



285

"not that 10370797...

...the new 10370797 "

286

applied ... no luck

287

Oracle Support is an organisation

288

(Australia) support

very good

289

log calls at particular times

then phone

290

user groups

OTN

use all resources

blogs

forums

291

deconfig CRS

hack crsconfig_params

re-run root.sh

add db, services, etc

292

cluster 1: 11.2.0.2 + patch

cluster 2: 11.2.0.2.2

cluster 3: 11.2.0.2.2 + patch

success !

cluster 4: 11.2.0.2.2 + patch

293

but then....

294

"case that stops the nation"

load test 11.2

295

```

topas mmon--2=Top-Child-CPU--Host: 11--Refresh=2 secs--20:33.31
Physical-CPU-Utilisation-Long-Term--EntitledCPU= UsedCPU=
100%-|#####|#####
95%-|#####|#####
90%-|#####U#####U#####|#####
85%-|#####U#####U#####|#####
80%-|#####U#####U#####|#####
75%-|#####U#####U#####|#####
70%-|#####U#####U#####|#####
65%-|#####U#####U#####|#####
60%-|#####U#####U#####|#####
55%-|#####U#####U#####|#####
50%-|#####U#####U#####|#####
45%-|#####U#####U#####|#####
40%-|#####U#####U#####|#####
35%-|#####U#####U#####|#####
30%-|#####U#####U#####|#####
25%-|#####U#####U#####|#####
20%-|#####U#####U#####|#####
15%-|#####U#####U#####|#####
10%-|#####U#####U#####|#####
5%-|#####U#####U#####|#####
+-----+
Kernel
RunQueue= 45.5 | swapIn = 0.0 | Directory Search | Kernel Processes
pswitch = 16506.4 | syscall= 64942.7 | iget = 0.0 | ksched= 0.0
fork = 241.0 | read = 1161.5 | dirblk= 0.0 | koverf= 0.0
exec = 124.0 | write = 280.5 | namei = 2410.6 | kexit = 0.0
msg = 0.0 | readch = 1820189.7 | Load Averages
sem = 0.0 | writech= 19967.0 | 1 min = 36.61
HW Intrp= 3000.1 | R+W(MB/s)= 0.2 | 5 min = 16.59
SW Intrp= 76.0 | Up Time=48.9 days (max=497) | 15 min= 8.74

```

mutex issue was back !!!

297

11.2.0.2 to 11.2.0.2.2

298

Bug 12431716 - Mutex waits may cause higher CPU usage in 11.2.0.2.2 PSU / GI PSU [ID 12431716.8]

Modified 02-DEC-2011 *Type* PATCH *Status* PUBLISHED

299

another patch...

300

one promising thing...

301

"Currently we are working on a new non-flash based version of MyOracleSupport which will be released early next year"

302

cluster 1: 11.2.0.2.2 + patch

cluster 2: 11.2.0.2.2 + patch

cluster 3: 11.2.0.2.2 + patch

success !

cluster 4: 11.2.0.2.2 + patch + patch



303

cluster 1: 11.2.0.2.2 + patch + patch

cluster 2: 11.2.0.2.2 + patch + patch

cluster 3: 11.2.0.2.2 + patch + patch

success !

cluster 4: 11.2.0.2.2 + patch + patch



304

351 days

305

Production 1
(bet enquiry)

2 clusters left

Production 2
(bet sales)

306

Production 1
(bet enquiry)

September 13th

49 days

307

what was planned.....

308

10pm	full backup
11pm	upgrade CRS node 1
12am	upgrade CRS node 2
1am	database upgrade
4am	ready for 7am business

309

what happened.....

310

10pm

full backup



311

10pm

full backup



11pm

upgrade CRS node 1

- haip
- rebooted
- deconfig
- OK !



312

10pm

full backup



11pm

upgrade CRS node 1

- haip
- rebooted
- deconfig
- OK !



1am

upgrade CRS node 2



313

1am

upgrade CRS node 2

- haip
- rebooted
- deconfig



314

1am upgrade CRS node 2

- haip
- rebooted
- deconfig

2am backout

- stop 11.2 crs
- start 11.1 crs

315

3am full restore



316

3am full restore



4am

backout #2
- start 11.1 crs



317

3am full restore



4am

backout #2
- start 11.1 crs



4:01am

ask for CEO's mobile

318





321

4:05am

shared files
- OCR / voting disk

322

4:05am shared files
- OCR / voting disk

5am backout #3
- restore OCR
- reinit voting disk
- start 11.1 crs



323

4:05am shared files
- OCR / voting disk

5am backout #3
- restore OCR
- reinit voting disk
- start 11.1 crs



6am services restarted

324

7am.... just another day

325

back to support...

326

"known issue with AIX during upgrade..."

set PSALLOC=early"

327

32 days

328

mutex issue = we **must** upgrade

329

~~Production 1~~

2 clusters left

Production 2

330



what was planned....

6pm	full backup
7pm	upgrade CRS node 1
10pm	upgrade CRS node 2
11pm	database upgrade
1am	ready for 7am business

333

what happened.....

334

6pm

full backup



335

6pm

full backup



8pm

upgrade CRS node 1
- PSALLOC=early



336

6pm

full backup



8pm

upgrade CRS node 1
- PSALLOC=early

8:10pm

upgrade CRS node 2
- PSALLOC=early
- *commenced*

337

`$GRID_HOME/cfgtoollogs/crsconfig`

338

back to node 1

341

```

topas mmon-2=Top-Child-CPU-----Host: 11-----Refresh=2 secs-----20:33.31
Physical-CPU-Utilisation-Long-Term-----EntitledCPU=-----UsedCPU=-----
100%-|#####|#####
95%-|#####|#####
90%-|#####|#####
85%-|#####|#####
80%-|#####|#####
75%-|#####|#####
70%-|#####|#####
65%-|#####|#####
60%-|#####|#####
55%-|#####|#####
50%-|#####|#####
45%-|#####|#####
40%-|#####|#####
35%-|#####|#####
30%-|#####|#####
25%-|#####|#####
20%-|#####|#####
15%-|#####|#####
10%-|#####|#####
5%-|#####|#####
+-----+
Kernel
RunQueue= 45.5 | swapIn = 0.0 | Directory Search | Kernel Processes
pswitch= 16506.4 | syscall= 64942.7 | iget = 0.0 | ksched= 0.0
fork = 241.0 | read = 1161.5 | dirblk= 0.0 | koverf= 0.0
exec = 124.0 | write = 280.5 | namei = 2410.6 | kexit = 0.0
msg = 0.0 | readch = 1820189.7 | Load Averages
sem = 0.0 | writch= 19967.0 | 1 min = 36.61
HW Intrp= 3000.1 | R+W(MB/s)= 0.2 | 5 min = 16.59
SW Intrp= 76.0 | Up Time=48.9 days (max=497) | 15 min= 8.74
Top-Processes=(313) -----Mode=3 (1=Basic 2=CPU 3=Perf 4=Size 5=I/O 6=Cmdl)

```

342

Sev 1 to support ...

343

... routed to Europe

344

node 1, 11.2 CRS, 100% cpu

help !

node 2, upgrading, 100% cpu

345

90 minutes silence

346



© DESPAIR.COM

LONELINESS

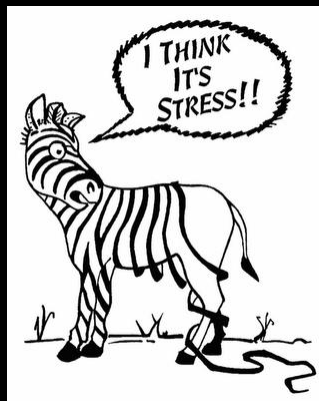
IF YOU FIND YOURSELF STRUGGLING WITH LONELINESS, YOU'RE NOT ALONE.
AND YET YOU ARE ALONE. SO VERY ALONE.

47

12:30am

upgrade CRS node 2

- 20% complete



348

```
truss <upgrade script>
```

349

```
getrlimit64(1, 0xFFFFFFFFFFDFA8) = 0
sigprocmask(2, 0x09001000A01FE9A0, 0xFFFFFFFFFFDE70) = 0
_sigaction(28, 0x0000000000000000, 0xFFFFFFFFFFDFB8) = 0
thread_setmystate(0xFFFFFFFFFFDA60, 0x0000000000000000) = 0
disclaim("", 524288, 1) = 0
disclaim("", 438272, 1) = 0
disclaim("", 118784, 1) = 0
disclaim("", 24576, 1) = 0
disclaim("", 12288, 1) = 0
disclaim("", 20480, 1) = 0
disclaim("", 520192, 1) = 0
disclaim("", 69632, 1) = 0
disclaim("", 1048576, 1) = 0
disclaim("", 438272, 1) = 0
disclaim("", 118784, 1) = 0
...
...
...
```

350

Google

Search About 677 results (0.23 seconds)

Everything [Miscellaneous tunable parameters](#) ★
 publib.boulder.ibm.com/infocenter/...aix.../misc_tun_params.htm [+1]
 Several of the miscellaneous parameters available in AIX are tunable. ... For 32-bit processes, the maximum size of all **memory** segments is 2.75 GB.

Images

Maps

Videos

News [Understanding Paging Space Allocation Policies](#) ★
 www.ualberta.ca/dept/...AIX_43/.../pag_space_order.ht... [+1]
 AIX Version 4.3 System Management Guide: Operating System and Devices ... The

Shopping

351

Paging Space Allocation

PSALLOC=early

352

back to support ...

353

is this the **cause** ?

can we unset ?

how to unset ?

354

"reboot both nodes"

355

in the middle of upgrade ?

356

"we think it will be OK..."

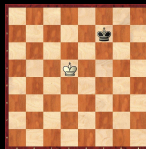
357

*"...if CRS fails to start, you will need to
deinstall and reinstall the 11.1 clusterware"*

358

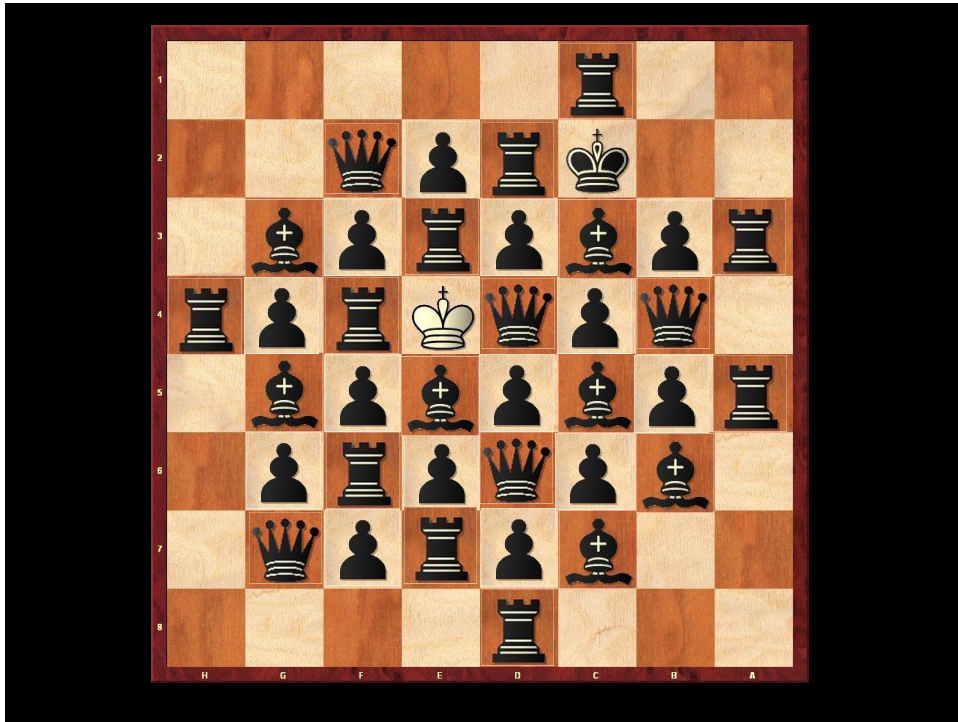
"zugzwang"

359



*"A player who has **no** move that
does not **worsen** his position is
said to be in '**zugzwang**'"*

360



```
server1:/root# reboot
```

"Fork failure -
not enough memory available"

362

```
server1:/root# exec reboot
```

363



364

12:30am upgrade CRS node 2
 - 20% complete

365

12:30am upgrade CRS node 2
 - 20% complete

01:30am

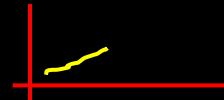


366

12:30am upgrade CRS node 2
 - 20% complete

01:30am power up node 1

01:35am no crs



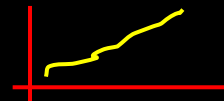
367

12:30am upgrade CRS node 2
 - 20% complete

01:30am power up node 1

01:35am no crs

01:40am no crs



368

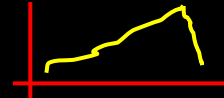
12:30am upgrade CRS node 2
 - 20% complete

01:30am power up node 1

01:35am no crs

01:40am no crs

01:42am CRS started



369

2:00am node 2 upgrade



370

2:00am reboot node 2
 - *mid upgrade*

02:15am rerun rootupgrades.sh

371

2:00am reboot node 2
 - *mid upgrade*

02:15am rerun rootupgrades.sh

03:30am CRS upgraded !

372

should have stopped there

3:30am

373

have staff support

tired decisions

374

3:30am database upgrade



5am success !



6am backup completed



6:30am services restarted

7:00am just another day !

375

30 days later...

376

November 1st, 2011

377

*"The **case** that stops the nation"*

378



fall out

380

Production 1 still on 11.1.0.7

381

11.2.0.2 Grid Infrastructure PSU 4 (Patch 12827731)

11.2.0.2 Grid Infrastructure PSU 3 (Patch 12419353)

11.2.0.2 Grid Infrastructure PSU2 (Patch 12311357)

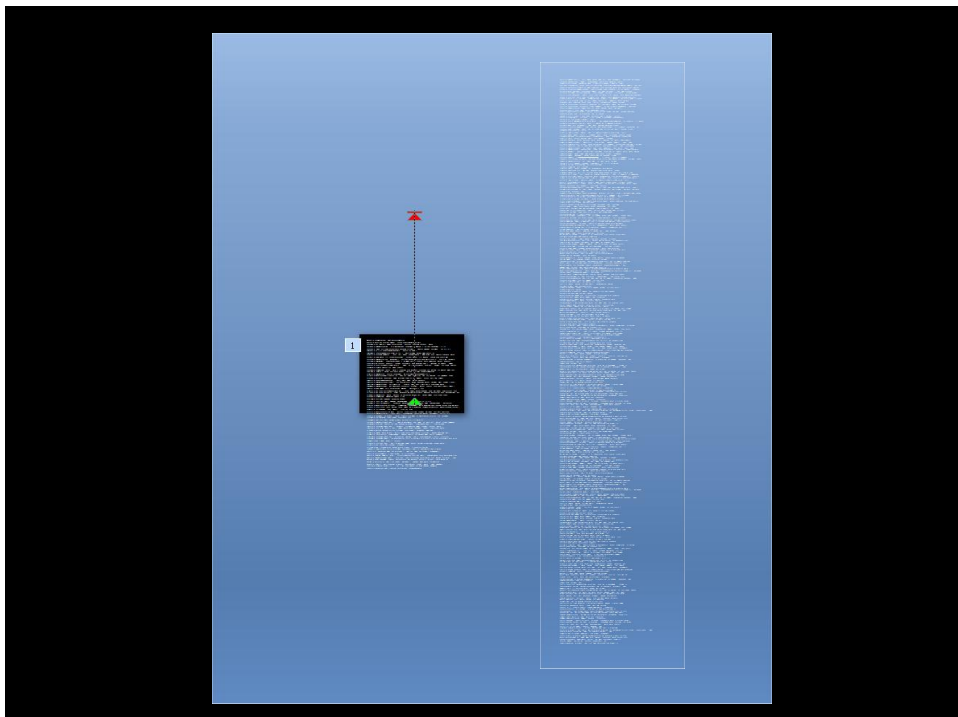
the world keeps moving

11.2.0.2 Grid Infrastructure Bundle 2 (Patch 10425672)

11.2.0.2 Grid Infrastructure Bundle 1 (Patch 10157506)

382

9458055 - CONNECTIONS FAIL WITH ORA-12537
 9795321 - MTU SIZE FOR VIP UNDER 11GR2 GRID INFRASTRUCTURE
 10161846 - BC: CHM/OS: DATA COLLECTION STOPPED ON ONE CLUSTER NODE
 10229297 - LNX64-11202-CRF: OSYSMOND.BIN CONSUMES ABOUT 7.5% CPU IN KERNEL 2.6.32
 10370797 - START OF 'ORA.CLUSTER_INTERCONNECT.HAIP' FAILED DURING UPGRADE TO 11.2.0.2
 10420872 - 11.2.0.2 GRID INFRASTRUCTURE ROOT.SH FAILS
 10428946 - SUSE10-LNX64-11203:OHASD.SLES NOT FOUND WHEN RAN ROOT.SH
 11692271 - 11203-NT:PARALLEL STOP CRS STACK ON ALL NODES AND ASM STOP FAILED ON LAST NODE
 11768055 - ONE NODE CSS JOIN CLUSTER,ALL OTHER NODES CSS ABORT FROM GM PEER LSNR
 11768599 - LNX64-11203:THE ATTRIBUTES OF \$CRS_HOME/CV/LOG/CVUTRACE.LOG.0 FILE ISN'T CORRECT
 11777648 - LNX64-RAC: GMON IN +ASM1, 2, 4 HANG IN 'CSS GROUP REGISTRATION' AFTER REBOOT
 11818524 - AT CLUSTER REBOOT, CEDING RECONFIG HAS WRONG RULES, CAUSING A STARTUP HANG
 11820787 - 11202_GIBTWO_LINUX:CRSCTL START RES WITH ERROR "CRS-5010: UPDATE OF CONF FAILED"
 11834289 - OHASD FAILED TO START TIMELY
 11871469 - ORAAGENT CHECK TASK IS TIMING OUT WHICH IS FORCING THE AGENT TO ABORT AND EXIT
 11925641 - SERVICES DON'T GET REGISTERED IN SUSE IN SIHA CONFIGURATION
 12345151 - LNX64-11203: GIPCD COREDUMP AT GIPCDMONITORTHREAD.C:126
 12359585 - AIX-11203-HA-CSS:ROOT.SH FAIL TO START ORA.CSSD DUE TO FAILURE OF EXADATA FENCE
 12366642 - DB CHECK ACTIONS TAKE A LONG TIME IF THE SHARED POOL SIZE ISN'T TUNED
 12406573 - GRID INFRASTRUCTURE DOES NOT START AFTER REBOOT
 12412303 - LNX64-11203-DIT:PERMS OF \$CH/CV/LOG/* WERE WRONG AFTER INSTALL RAC USING 2 USERS
 12422324 - LNX64-112-CMT: HIT PRVF-7539 : GROUP "DBA" DOES NOT EXIST ON OUDA NODE
 12424479 - LNX64-112-CMT: VIP FAILOVER TOOK 123 SECONDS AFTER PULLED OUT TWO PUBLIC CABLES
 12426375 - ASM HANG 'GCS LOCK OPEN S'<='ENQ: FA - ACCESS FILE'
 12546712 - HPI 11203 CRS:ORAROOTAGENT CORE DUMP WHEN ADDING PRIV NIC BACK VIA OIFCFG SETIF
 12556028 - HPI 11203 CRS:CRSD LEAK MEMROY AT 8M/1.2H ON PE MASTER WITH OCR BUSY WORKLOAD RU
 12597463 - LNX64-11201: NODE SHOULD BE EVICTION WHEN CRS HOME HANG FOR SOME TIME
 12600343 - ORAAGENT DAEMON CRASH ISSUE
 12615394 - GIPC API LEAKING MEMORY IN AIX
 12651595 - STOP CRS FAILS DURING DOWNGRADE (11.2.0.3.0 TO 11.2.0.1.0)
 12680491 - STOP DEPENDENCY BETWEEN NETWORK AND VIP MUST HAVE INTERMEDIATE MODIFIER
 12695202 - AIX-11202GIB3 CRS:CRSD ORAAGENT COREDUMP AT CLSN_AGENT:INSTCONNECTION::BREAKCALL
 12700217 - 11202_GIB3: CRSD AGENT CORE DUMP AT CLSN_AGENT::CONNECTIONPOOL::ABORTCONNECTIONS
 12743755 - SET FENCING TYPE CAUSES CSSD TO SPIN
 12767231 - AIX-11202-GIPC DISABLE SWITCH PORT FOR ONE PRIVATE NIC,HAIP DID NOT FAILOVER



personal fallout

385



386

wrap up
recommendations

387

Oracle software company

388

solutions company

389

hardware and software

390

Oracle Press Release

Oracle Stops All Software Development For Intel Itanium Microprocessor

REDWOOD SHORES, Calif., March 22, 2011

After multiple conversations with Intel senior management Oracle has decided to discontinue all software development on the Intel Itanium microprocessor. Intel management made it clear that their strategic focus is on their x86 microprocessor and that Itanium was nearing the end of its life.

Both Microsoft and RedHat have already stopped developing software for Itanium. HP CEO Leo Apotheker made no mention of Itanium in his long and detailed presentation on the future strategic direction of HP.

Oracle will continue to provide customers with support for existing versions of Oracle software products that already run on Itanium.

391



392

entire stack

393

better product

primary port

better support

394

11.2.0.3

the bare minimum

395

as simple as possible

396

have **lots** of clusters to test

397

consider full MAA

398



399