数据库升级与数据迁移实战分享

DTCC

2015中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA 2015

大数据技术探索和价值发现

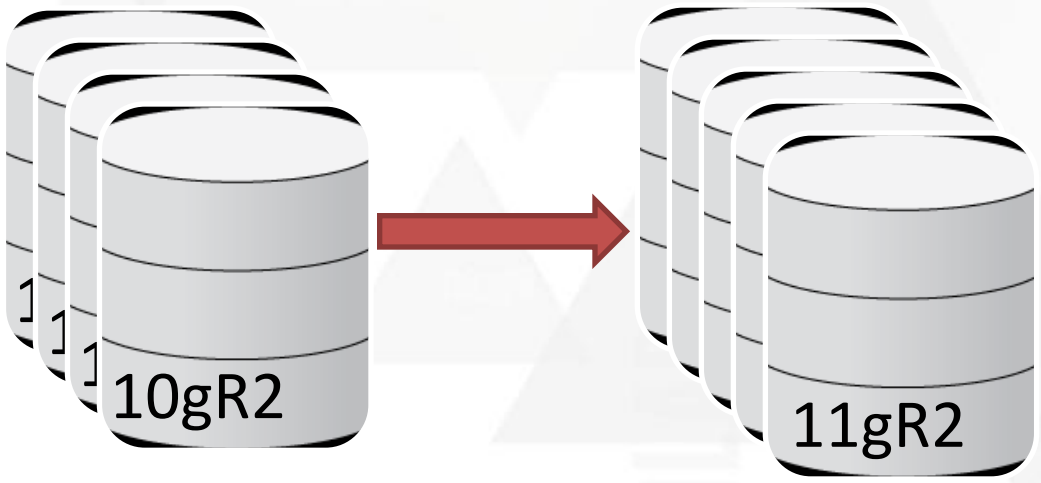
Amdocs 杨建荣

分享思路

- 数据库升级实战
  - 严谨，细心，运气
- 数据迁移实战
  - 挑战传统方法，艰难的技术推广


加油！

升级目标

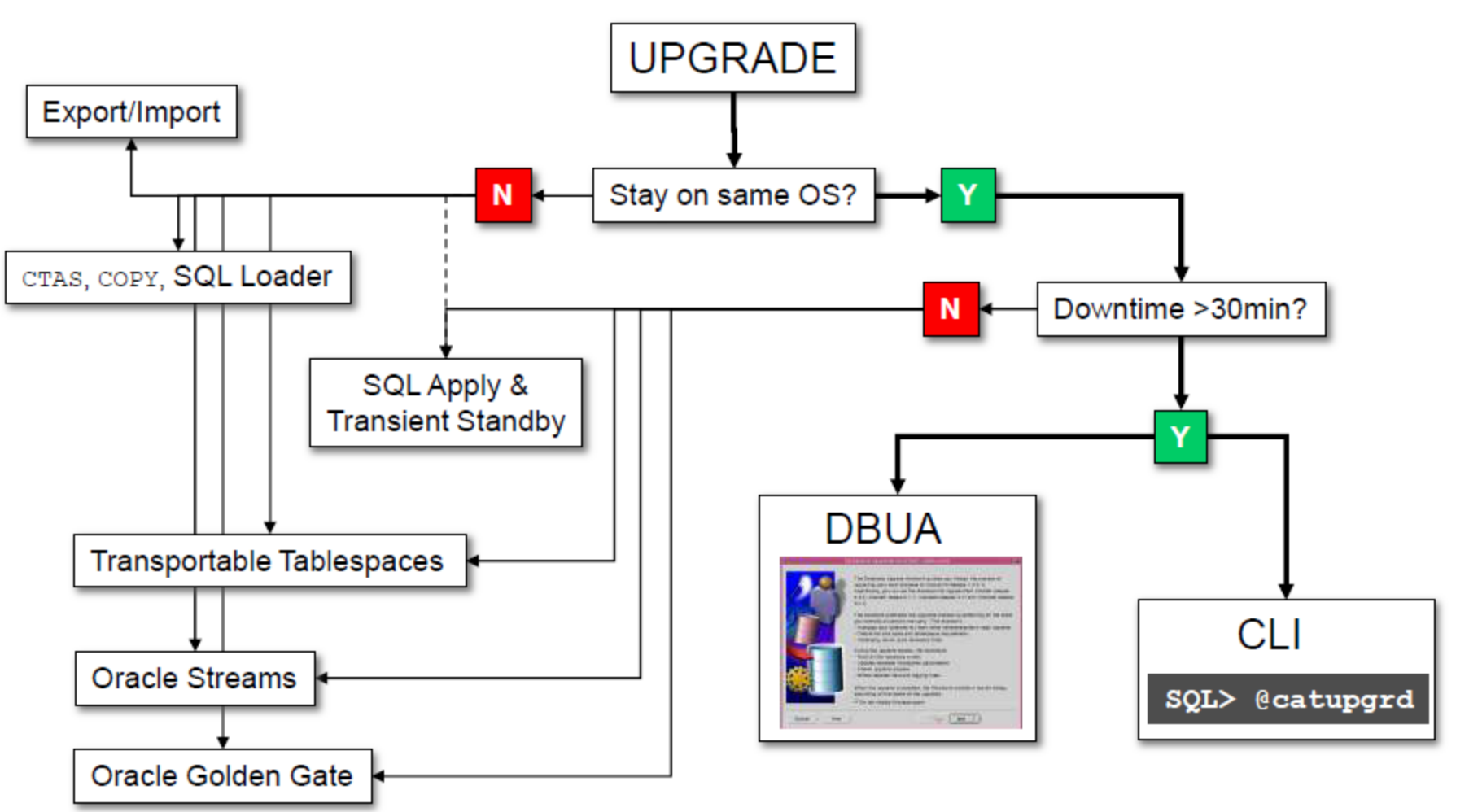


	Full Support	Extended
11.2	Jan 2015	2018
11.1	Aug 2012	2015
10.2.0.5	July 2010	2015

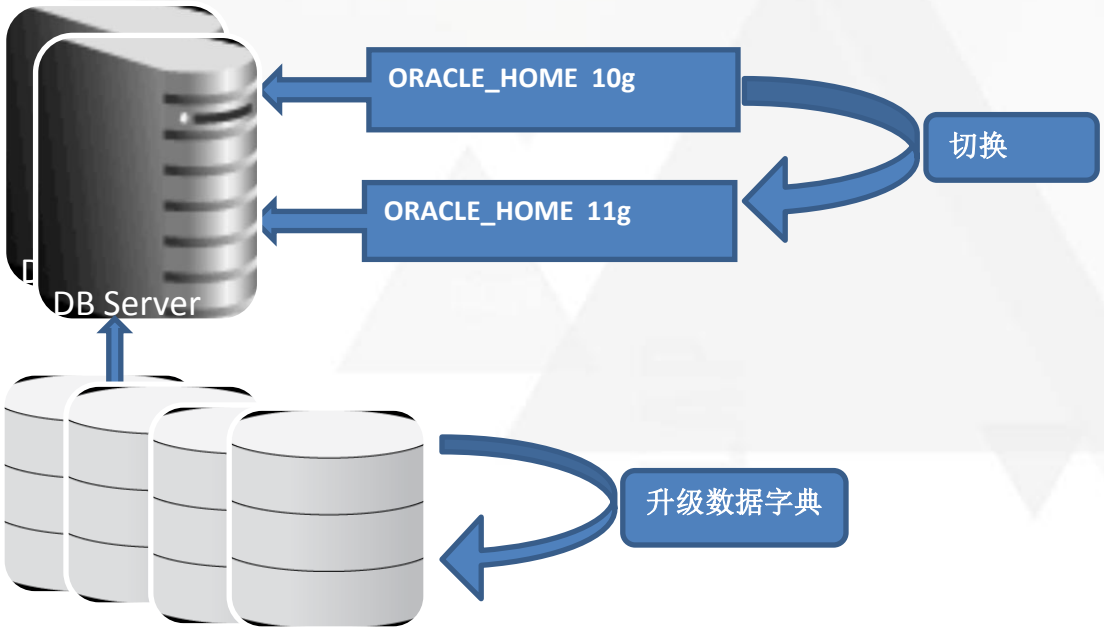
资源情况

- 2台服务器,单实例数据库 4个,10gR2, T级
- Linux 2.6.18-308.el5 #1 SMP
  - x86_64 x86_64 x86_64 GNU/Linux
- OS Sockets / Cores / CPUs：**4 / 40 / 80**
  - CPU model name : **Intel(R) Xeon(R) CPU E7- 4870 @ 2.40GHz**
- OS Memory：**346 GB**
- 至少四轮以上的测试
- 开发测试部门,产品部规划
- 性能DBA,应用DBA,开发DBA
- 客户的DBA
- Oracle的support

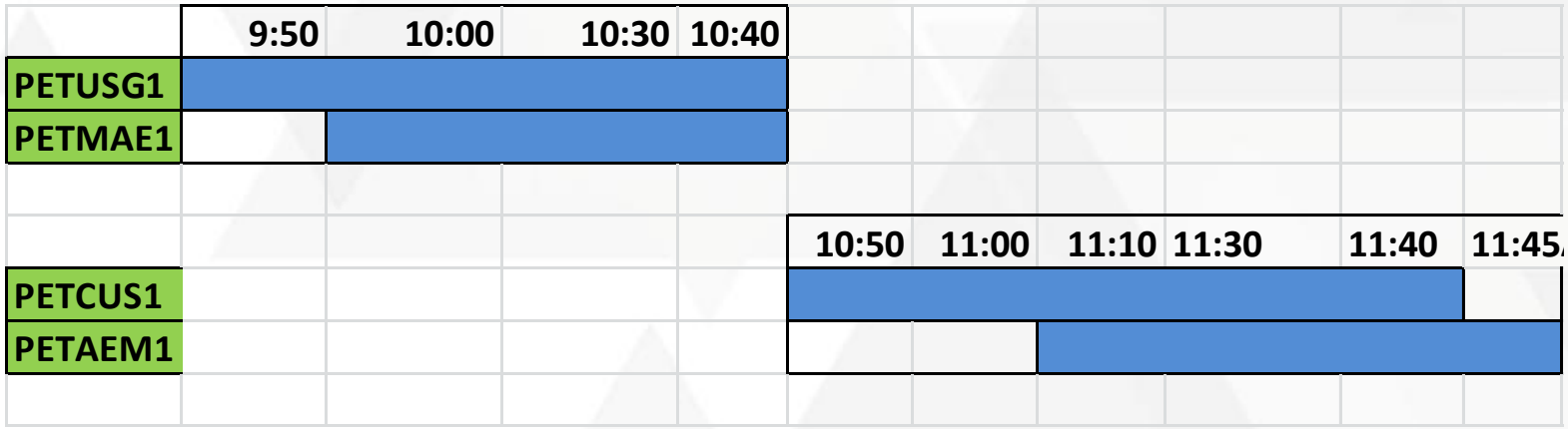
升级可选方案

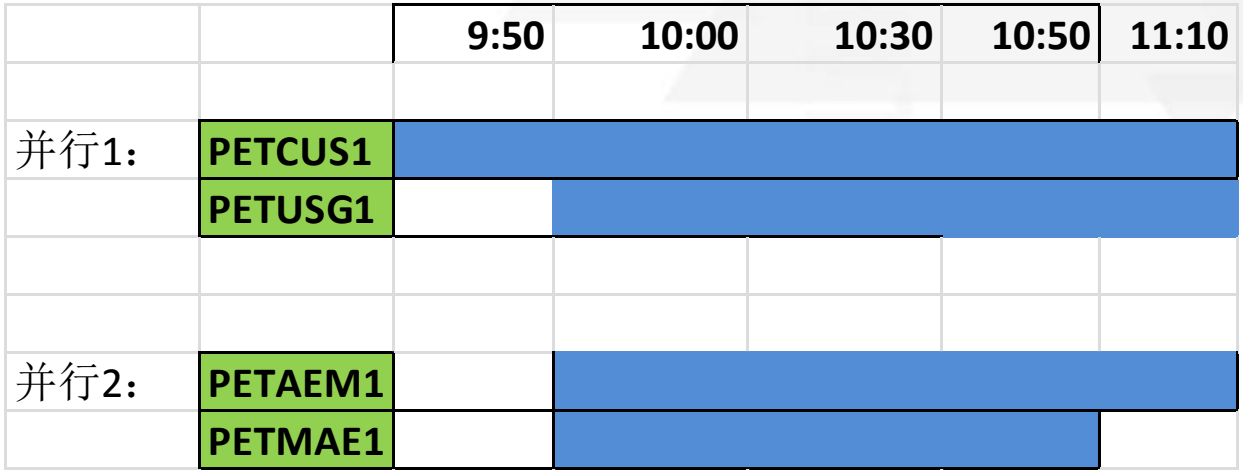

UPGRADE
Export/Import
N
Stay on same OS?
Y
CTAS, COPY, SQL Loader
N
Downtime >30min?
SQL Apply &
Transient Standby
Y
DBUA
CLI
SQL> @catupgrd
Transportable Tablespaces
Oracle Streams
Oracle Golden Gate

手工升级流程图



升级流程规划图


9:50
10:00
10:30
10:40
PETUSG1
PETMAE1
10:50
11:00
11:10
11:30
11:40
11:45
PETCUS1
PETAEM1


9:50
10:00
10:30
10:50
11:10
并行1：
PETCUS1
PETUSG1
并行2：
PETAEM1
PETMAE1

升级中的挑战

- 过程需要完全可控
- 需要同时升级4个数据库实例
- 附加部分业务升级
- 3个小时内全部完成(1个小时的业务升级)
- 未知的突发情况

升级前的准备

安装11gR2 软件，部署PSU

./opatch lsinventory
Interim patches (1) :

Patch 16056267 : applied on Thu Oct 03 16:01:47 ICT 2013
Unique Patch ID: 15994298
Patch description: "Database Patch Set Update : 11.2.0.2.10 (16056267)"
Created on 12 Mar 2013, 10:08:56 hrs PST8PDT
Sub-patch 14727315; "Database Patch Set Update : 11.2.0.2.9 (14727315)"
Sub-patch 14275621; "Database Patch Set Update : 11.2.0.2.8 (14275621)"
Sub-patch 13923804; "Database Patch Set Update : 11.2.0.2.7 (13923804)"
Sub-patch 13696224; "Database Patch Set Update : 11.2.0.2.6 (13696224)"

升级前的参数优化

10g已有的隐含参数

parameter name	parameter value
_optimizer_cost_model	CPU
_optimizer_skip_scan_enabled	FALSE
_optimizer_sortmerge_join_enabled	FALSE
_optimizer_use_feedback	FALSE
_optimizer_dynamic_sampling	0
_optimizer_join_elimination_enabled	FALSE
parallel_adaptive_multi_user	FALSE
parallel_execution_message_size	65536

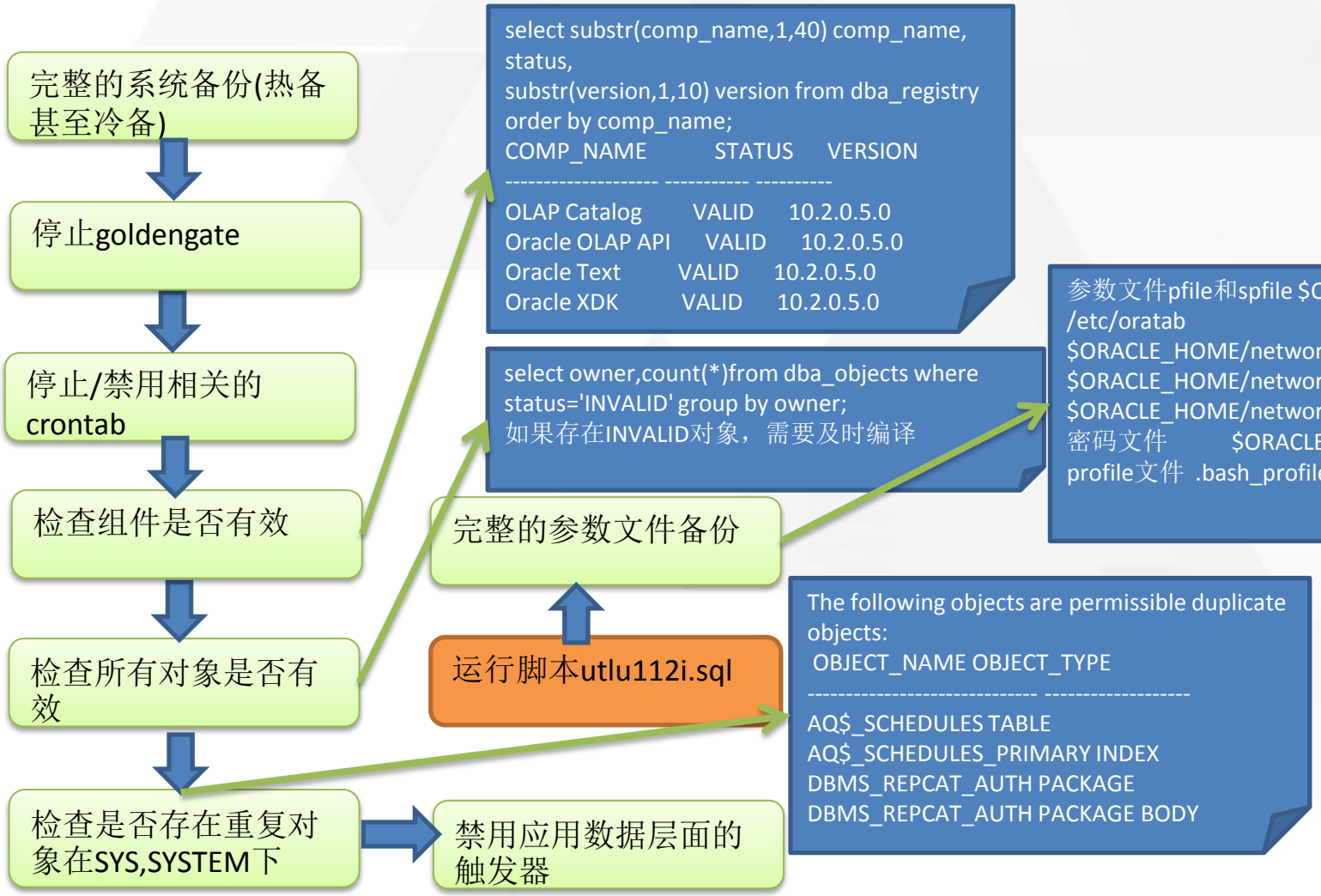
11g中过期的10g参数

background_dump_dest
cursor_space_for_time
fast_start_io_target
log_archive_local_first
log_archive_start
max_enabled_roles
parallel_automatic_tuning
parallel_server
parallel_server_instances
plsql_debug
plsql_v2_compatibility
remote_os_authent

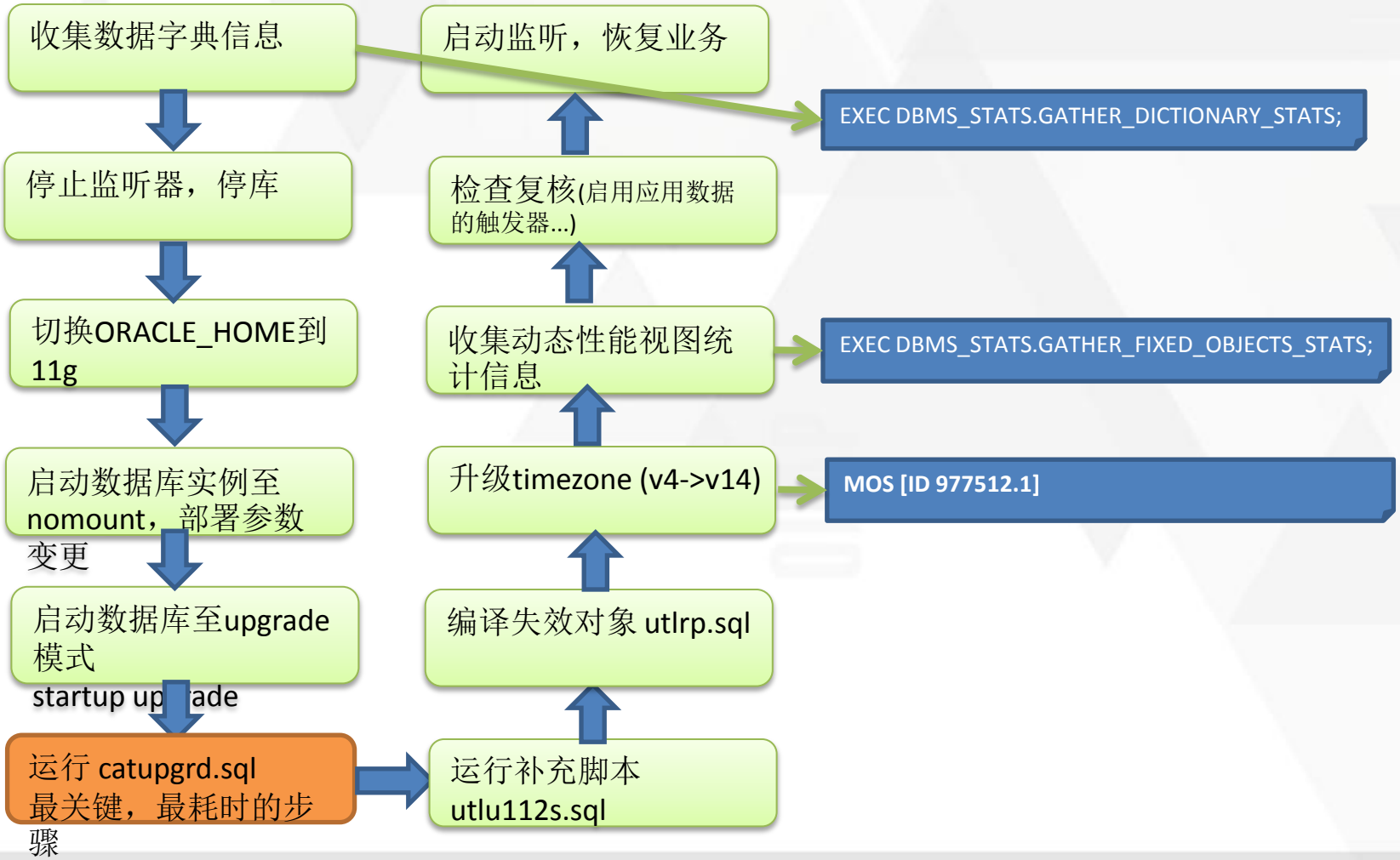
升级前的参数优化

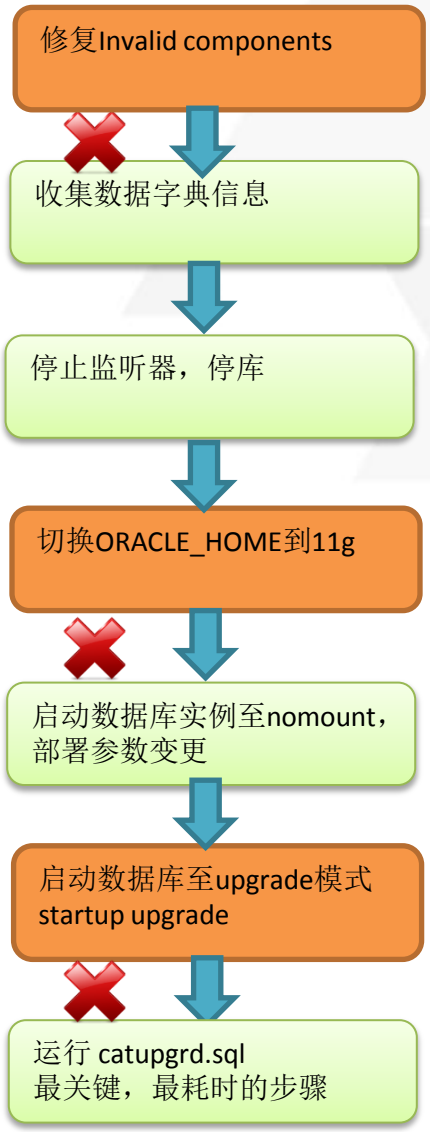
- 11g中新参数的选择
  - sec_case_sensitive_logon
  - sec_max_failed_login_attempts
  - result_cache_max_size
  - deferred_segment_creation
  - diagnostic_dest
  - db_unrecoverable_scn_tracking

升级前的检查


完整的系统备份(热备甚至冷备)
停止goldengate
停止/禁用相关的crontab
检查组件是否有效
检查所有对象是否有效
检查是否存在重复对象在SYS,SYSTEM下
禁用应用数据层面的触发器
select substr(comp_name,1,40) comp_name,
status,
substr(version,1,10) version from dba_registry
order by comp_name;
COMP_NAME STATUS VERSION
OLAP Catalog VALID 10.2.0.5.0
Oracle OLAP API VALID 10.2.0.5.0
Oracle Text VALID 10.2.0.5.0
Oracle XDK VALID 10.2.0.5.0
select owner,count(*)from dba_objects where
status='INVALID' group by owner;
如果存在INVALID对象，需要及时编译
完整的参数文件备份
参数文件pfile和spfile $ORACLE_HOME/dbs
/etc/oratab
$ORACLE_HOME/network/admin/listener.ora
$ORACLE_HOME/network/admin/tnsnames.o
$ORACLE_HOME/network/admin/sqlnet.ora
密码文件 $ORACLE_HOME/dbs
profile文件 .bash_profile或者.profile
运行脚本utlu112i.sql
The following objects are permissible duplicate objects:
OBJECT_NAME OBJECT_TYPE
AQ$_SCHEDULES TABLE
AQ$_SCHEDULES_PRIMARY INDEX
DBMS_REPCAT_AUTH PACKAGE
DBMS_REPCAT_AUTH PACKAGE BODY

升级流程图


收集数据字典信息
停止监听器，停库
切换ORACLE_HOME到11g
启动数据库实例至nomount，部署参数变更
启动数据库至upgrade模式
startup upgrade
运行 catupgrd.sql
最关键，最耗时的步骤
运行补充脚本
utlu112s.sql
编译失效对象 utlrp.sql
升级timezone (v4->v14)
收集动态性能视图统计信息
检查复核(启用应用数据的触发器...)
启动监听，恢复业务
EXEC DBMS_STATS.GATHER_DICTIONARY_STATS;
EXEC DBMS_STATS.GATHER_FIXED_OBJECTS_STATS;
MOS [ID 977512.1]




```
SQL> BEGIN
  2    IF dbms_registry.status('CATJAVA') IS NULL THEN
  3      RAISE_APPLICATION_ERROR(-20000,
  4          'CATJAVA has not been loaded into the database.');
  5    END IF;
```

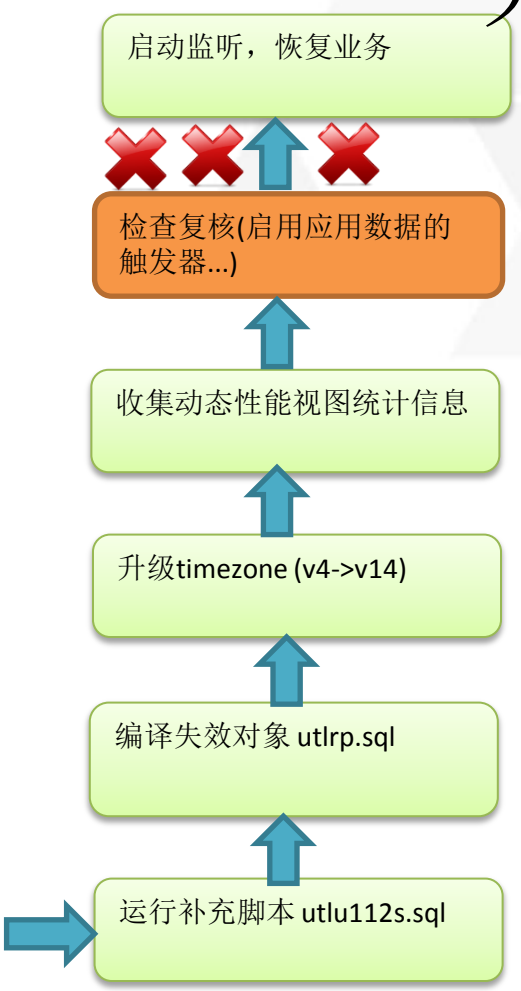
sqlplus / as sysdba
SP2-1503: Unable to initialize Oracle call interface

startup upgrade
ORA-00600: internal error code, arguments: [kccsbck_first], [1], [3125205511], [], [], [], [], [], [], [], [],

解决方法:reload->remove>re-install

解决方法：最后还是先保证**11g**的环境下没有**ora**进程存在，然后切回**10g**的环境，看有没有**ora**的进程，如果有一定要停掉。然后再切回**11g home,**就可以了。

升级中的问题分析



查询user_synonyms持续近10分钟

重建USER_SYNONYMS

```
CREATE OR REPLACE FORCE VIEW "SYS"."USER_SYNONYMS" ("SYNONYM_NAME",
"TABLE_OWNER", "TABLE_NAME", " DB_LINK") AS
    select /*+ RULE */ o.name, s.owner, s.name, s.node
  from sys.syn$ s, sys."_CURRENT_EDITION_OBJ" o
  where o.obj# = s.obj#
   and o.type# = 5
    and o.owner# = userenv('SCHEMAID');
```

```
exp APP_ROLLBK/APP_ROLLBK file=test.dmp
tables=AAAAA  consistent=y
. . exporting table                  AAAAA 76 rows exported
EXP-00008: ORACLE error 1466 encountered
ORA-01466: unable to read data - table definition has changed
Export terminated successfully with warnings.
```

alert日志报错：

```
Archived Log entry 6765 added for thread 1 sequence 6728 ID
0xb8c6d509 dest 1:
Tue Oct 08 04:51:01 2013
ORA-1466 (RO Tx began: 10/07/2013 21:51:00, Last DDL: 10/08/2013
02:07:25, Curr Time: 10/07/2013 21:51:00)
Tue Oct 08 06:04:15 2013
```

关于ORA-01466所做的尝试

expdp XXXX/XXX DIRECTORY=DATA_PUMP_DIR DUMPFILE=test.dmp tables=AAAAA consistent=y **--->OK**

export from **schema1**(without consistent=y)
import into **schema2**
export from **schema2** with consistent=y
还是有问题

检查物理时钟，时间都是同步的。
hwclock;date
Tue 08 Oct 2013 01:18:18 PM ICT -0.267508 seconds
Tue Oct 8 13:18:17 ICT 2013

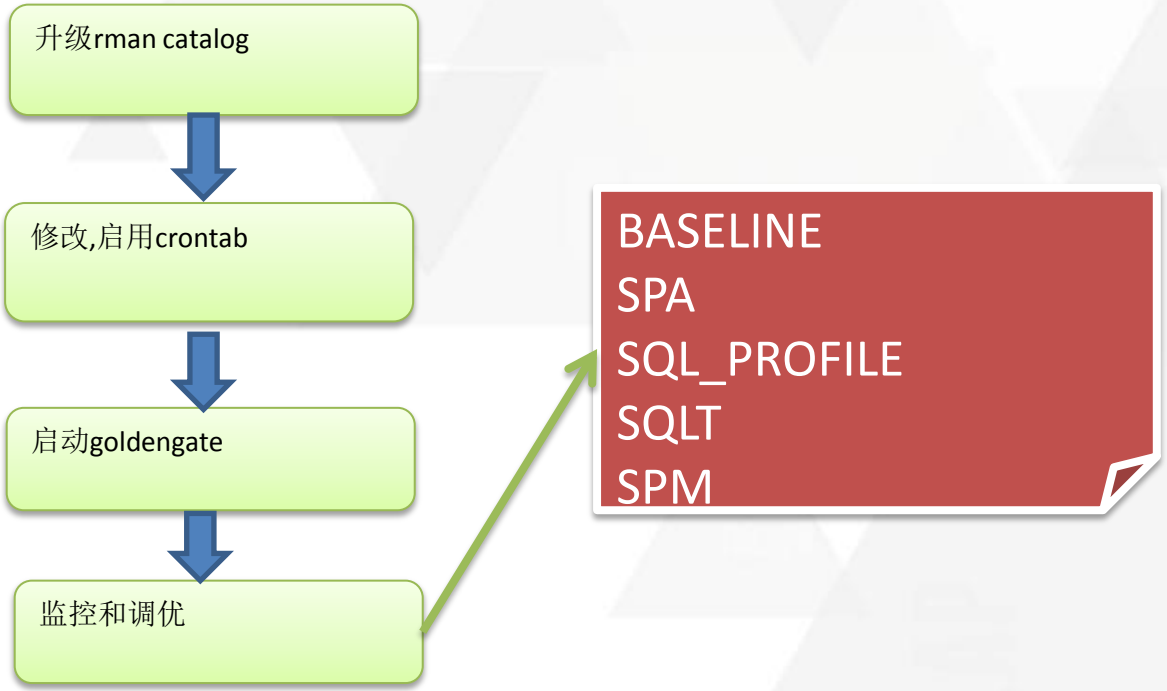
ORA-01466的最终解决方案

解决方案：**PURGE RECYCLEBIN**

sqlplus / as sysdba
SQL> purge recyclebin ;
EXP-00008: ORACLE error 1466 encountered
ORA-01466: unable to read data - table definition has changed
Export terminated successfully with warnings.

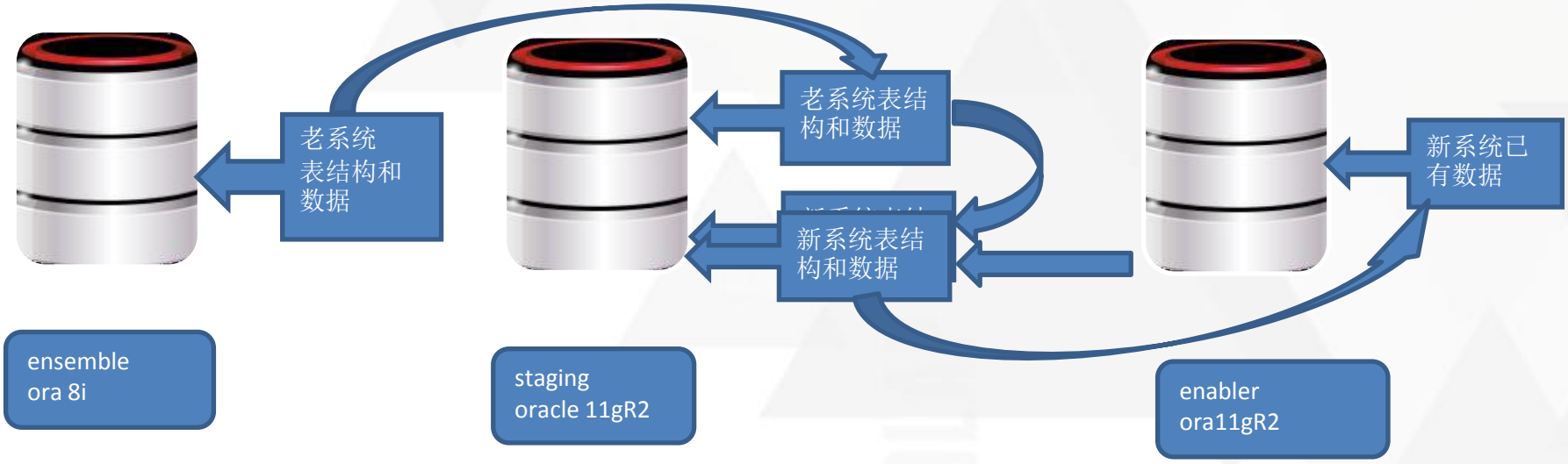
This is a time-based read consistency error for a database object, such as a table or index.
Either of the following may have happened: The query was parsed and executed with a snapshot older than the time the object was changed.
The creation time-stamp of the object is greater than the current system time. This happens, for example, when the system time is set to a time earlier than the creation time of the object.

升级后的补充



Upgrade from/to	% Different execution plans
v9 -> v10	15.31%
v9 -> v11	30.05%
v10 -> v11	21.42%

数据迁移背景


老系统
表结构和
数据
老系统表结
构和数据
新系统表结
构和数据
新系统已
有数据
ensemble
ora 8i
staging
oracle 11gR2
enabler
ora11gR2

已有的数据迁移方案

- perl+csv+sqlldr+datapump
- 脚本由公司资深专家编写
- 在其它项目中已经使用多次
- 数据抽取和加载的过程可控
- 经过多次验证，算是稳定的实现
- 成本低，不依赖第三方的付费服务
- sqlldr加载LOB数据类型的限制

其它升级方案的比较

- **insert /*+append */ into ... from db_link +nologging**

可以实现跨平台, 过度依赖网络和磁盘情况, 迁移的过程不好控制

- **standby（dataguard)**

切换时间极短

整个数据库的迁移. 不能实现数据的重组，即业务数据的增量导入

- **stream/goldengate**

staging库是非归档模式,实现过程相对复杂,过度依赖网络和磁盘情况

- 表空间传输

跨平台,数据文件转换,业务数据的增量导入无法实现

纠结的测试结果

- **需要单独安装perl-DBI模块**
- **datapump的性能考虑**
- **生成大量的csv文件，基本不可读**
- **潜在的乱码问题**
- **sqlldr遇到的数据问题**
  - 主键冲突
    - ORA-00001: unique constraint (PRDAPPO.AR1_MEMO_PK) violated
  - 非空约束问题
    - ORA-01400: cannot insert NULL into ("PRDAPPO"."CL9_CRD_MNTR_TREAT"."ACT_RSN_CODE")
  - 外键数据问题/表插入数据的顺序
    - ORA-02291: integrity constraint (PRDAPPO.CH_OBJECT_ATTRIBUTES_1FK) violated

Datapump带来的问题

- Datapump的导入问题
  - 约束导致的导入回退
    - ORA-31693: Table data object "PRDAPPO"."MO1_MEMO":"PMAX_AMAX_EMAX" failed to load/unload and is being skipped due to error: ORA-00001: unique constraint (PRDAPPO.MO1_MEMO_PK) violated Job "PRDAPPO"."SYS_IMPORT_FULL_01" completed with 1 error(s) at 02:34:33
  - **undo的困扰**
    - ORA-31693: Table data object "MIG_TEST"."MO1_MEMO":"P2_A1000_E3" failed to load/unload and is being skipped due to error: ORA-29913: error in executing ODCIEXTTABLEFETCH callout ORA-30036: unable to extend segment by 8 in undo tablespace 'UNDOTBS1'

Datapump带来的问题

– 加载LOB数据的假并行

Elapsed Time (s)	Executions	Elapsed Time per Exec (s)	%Total	%CPU	%IO	SQL Id	SQL Module
3,553.36	0		1.01	35.03	10.32	6cz7m51m82vqg	Data Pump Worker

6cz7m51m82vqg	INSERT /*+ PARALLEL("MO1_MEMO", 1)+*/ INTO RELATIONAL("APPO"."MO1_MEMO" NOT XMLTYPE) ("APP_ID", "ENTITY_KEY", "PERIOD_KEY", "MEMO_ID", "SYS_CREATION_DATE", "SYS_UPDATE_DATE",xxxxxxxx, "ATTR10VALUE", "CLOB_IND", "MEMO_SYSTEM_TEXT_C") SELECT "APP_ID", "ENTITY_KEY", "PERIOD_KEY", "MEMO_ID", "SYS_CREATION_DATE", "SYS_UPDATE_DATE", xxxxxxxx, "ATTR10VALUE", "CLOB_IND", "MEMO_SYSTEM_TEXT_C" FROM "ET$111D05F70001" KU$ LOG ERRORS INTO "APPO"."ERR$DP111D05F70001" REJECT LIMIT UNLIMITED

使用imp加载CLOB数据的速度测试

SQL> select 620540/15/60 from dual;

620540/15/60

689.488889

让人提心吊胆的问题

- 数据问题总是在最后关头发现，没法提前进行比较
- 大分区表的加载性能问题
- 数据加载的过程可控性不够
- 数据修复总是紧急加急

被逼无奈的大胆尝试

	抽取开始时间	抽取结束时间	时长	产生的dump大小
sqlldr(exclude 2 tables) 68 tables	Tue Jul 15 13:28:15 ICT 2014	Tue Jul 15 13:42:32 ICT 2014	**14 mins**	150G
EXT_DATAPUMP 70 tables	Tue Jul 15 14:03:11 ICT 2014	Tue Jul 15 14:06:20 ICT 2014	**3 mins**	59G

	加载数据开始时间	加载数据结束时间	时长	产生归档文件大小
sqlldr(exclude 2 tables) 68 tables	Tue Jul 15 14:24:24 ICT 2014	Tue Jul 15 15:44:04 ICT 2014	**80 mins**	230G
EXT_DATAPUMP 70 tables	Tue Jul 15 17:06:51 ICT 2014	Tue Jul 15 17:50:04 ICT 2014	**44 mins**	139G

❑ 希望

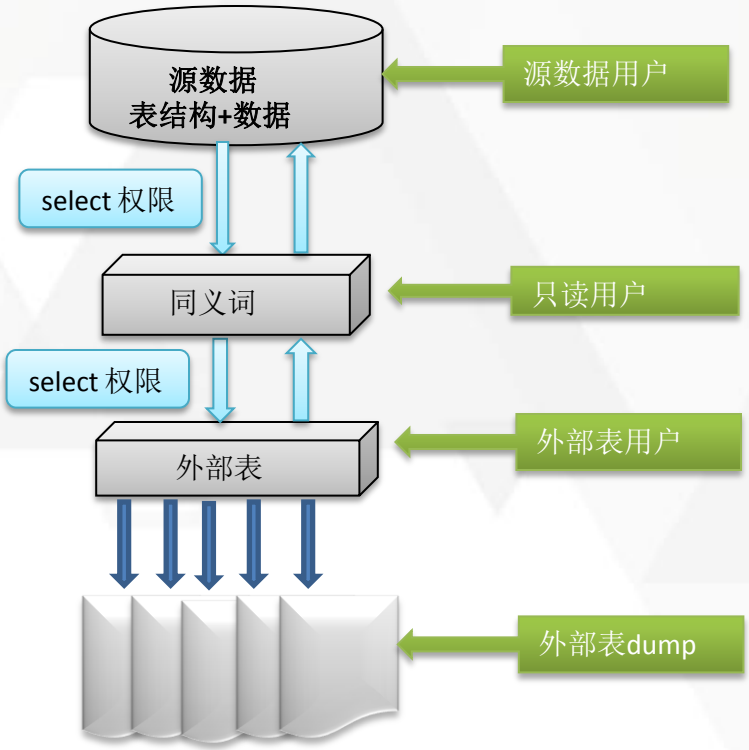
- ❑ 速度提升明显
- ❑ 完整的数据类型支持
- ❑ 解决乱码问题的困扰
- ❑ 稳定性和数据完整性
- ❑ 有质疑变为信任
- ❑ 压力变为动力

❑ 顾虑

- ❑ 在其它项目中目前还没有尝试过
- ❑ 新技术的新鲜感，可能华而不实
- ❑ 技术是否成熟，稳定
- ❑ 数据是否会丢失
- ❑ 测试是否充分
- ❑ 数据安全性，引入不当操作（drop,alter）

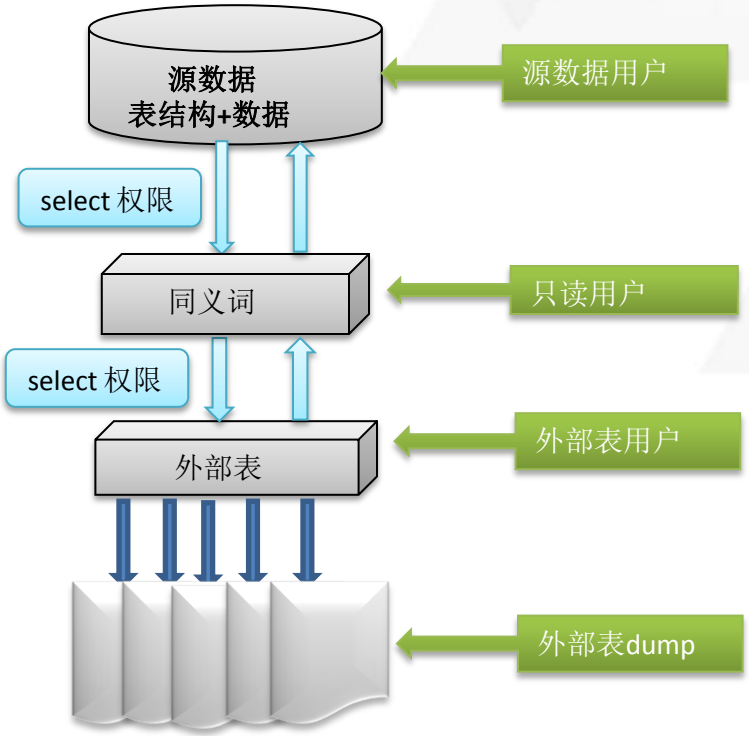
数据抽取示意图

```
CREATE TABLE MO1_MEMO_EXT
   ORGANIZATION EXTERNAL
  (
    TYPE ORACLE_DATAPUMP
    DEFAULT DIRECTORY xxxxx
    LOCATION (
   'xxxx1.dmp','xxxx2.dmp')
   )
   parallel xx   as
   SELECT  /*+ parallel(t xx) */
              *
   FROM 只读用户.MO1_MEMO;
```

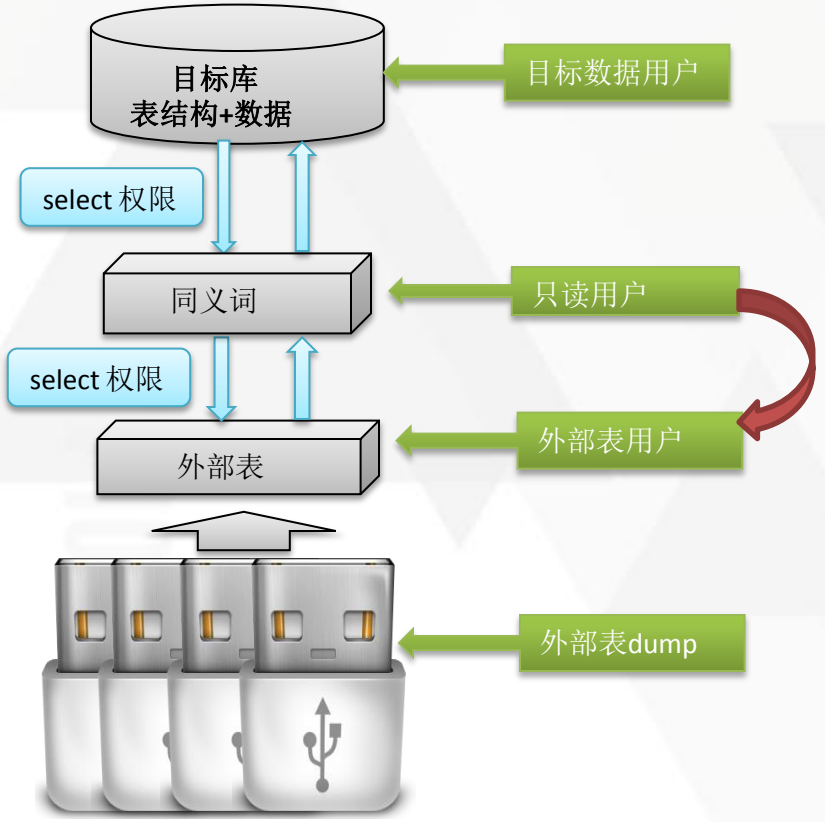


数据加载示意图

数据源

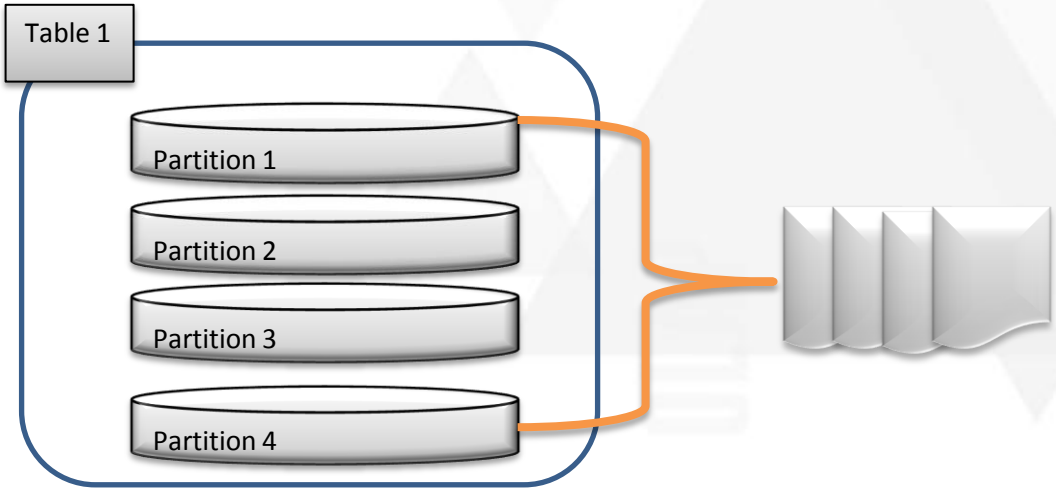


数据加载的流程



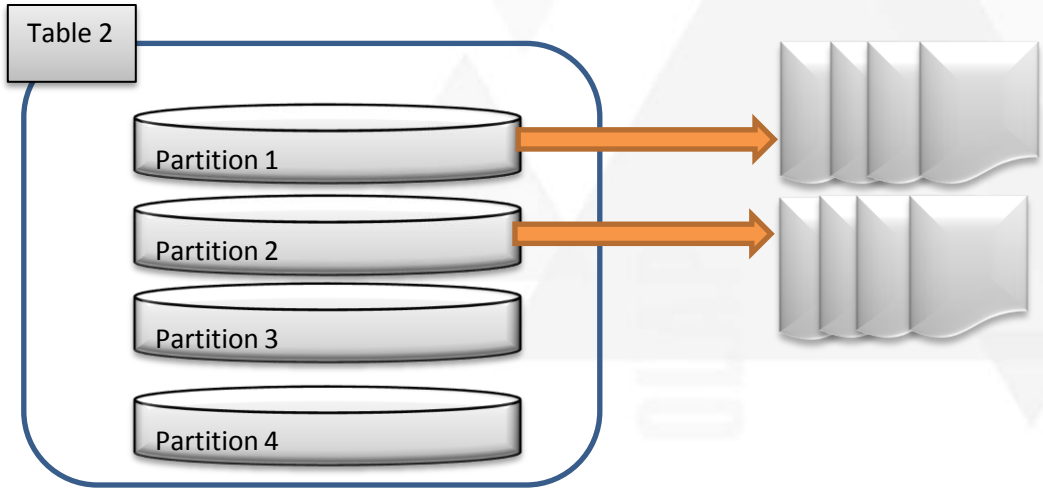
数据抽取的切分思路

full table+parallel



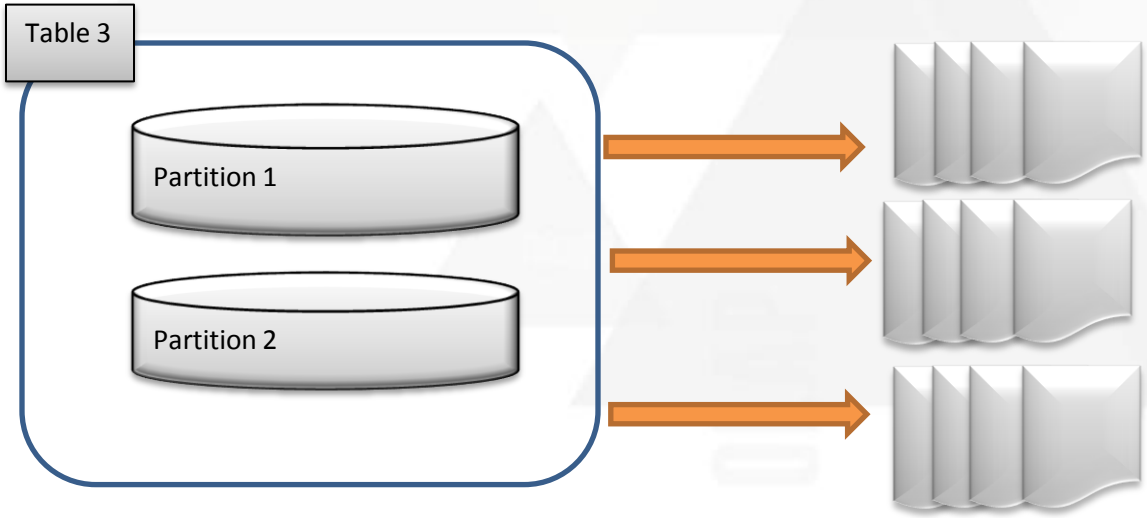
数据抽取的切分思路

partition+parallel



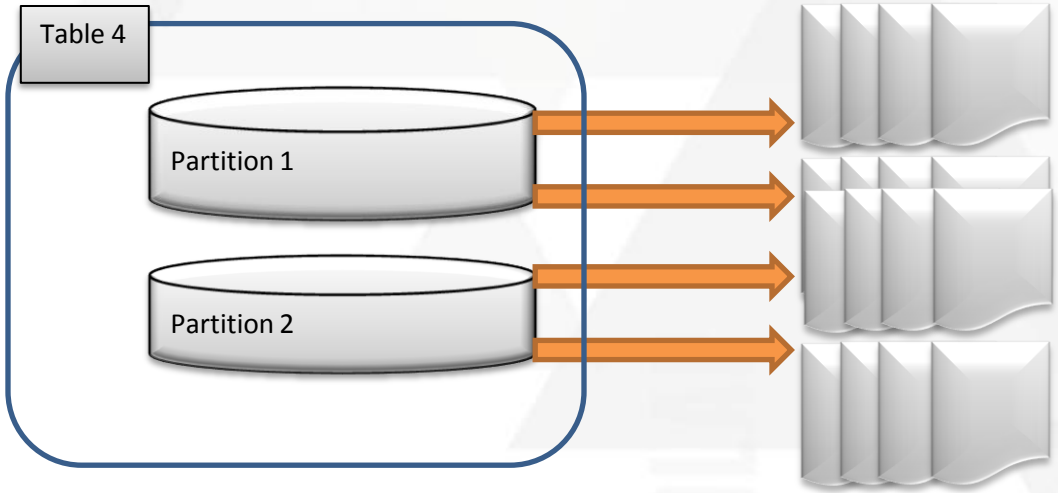
数据抽取的切分思路

full table+rowid+parallel

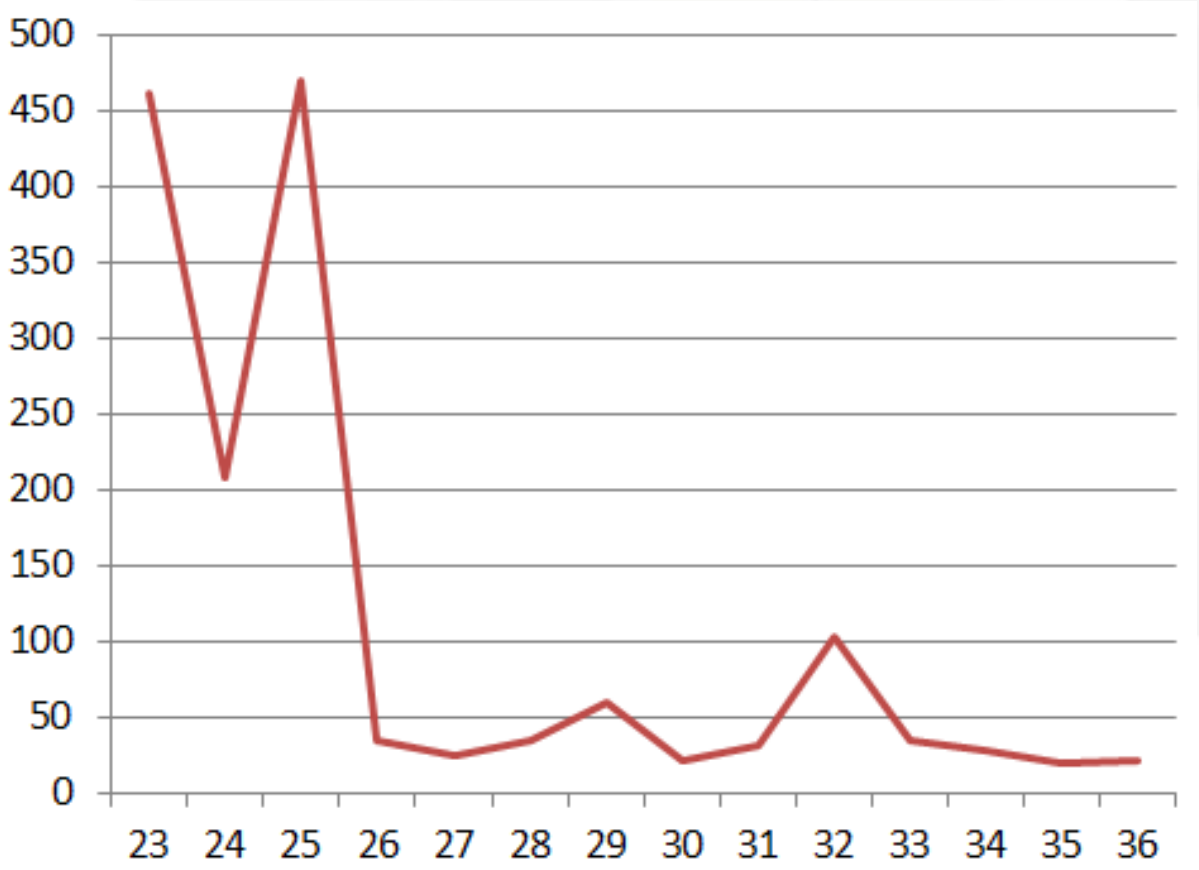


数据抽取的切分思路—rowid切分

partition+rowid+parallel



节外生枝 --严重的I/O问题

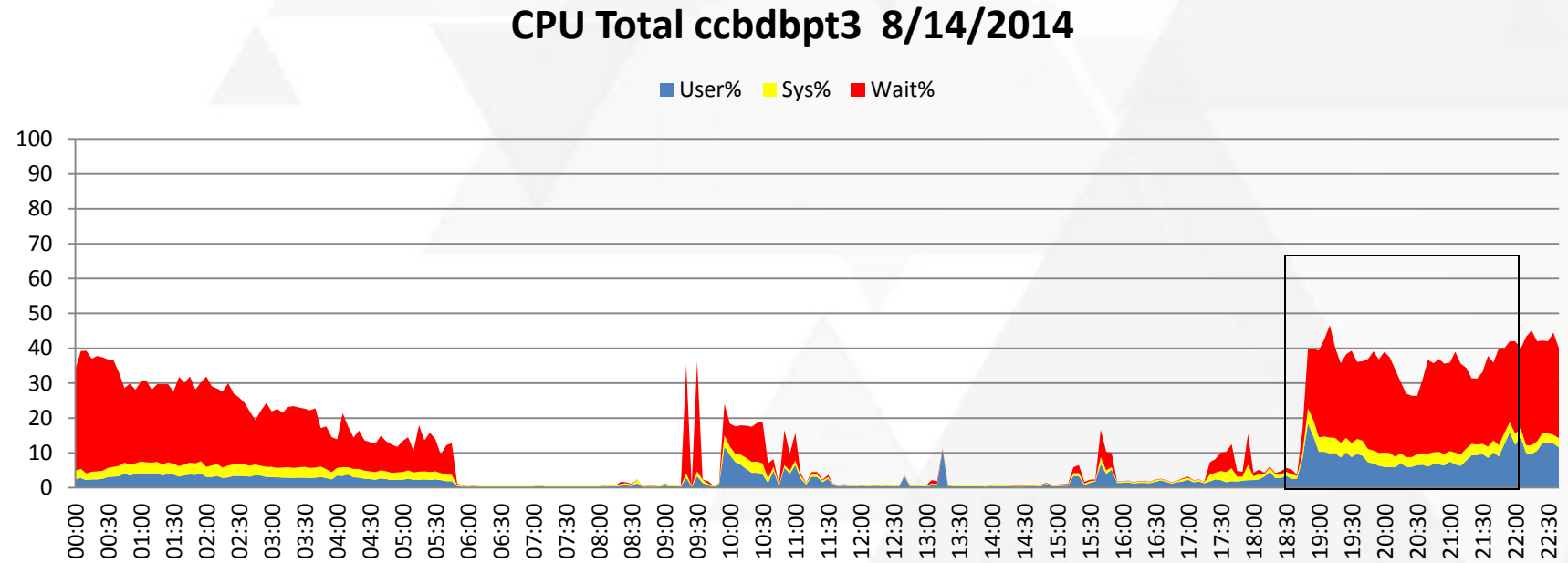


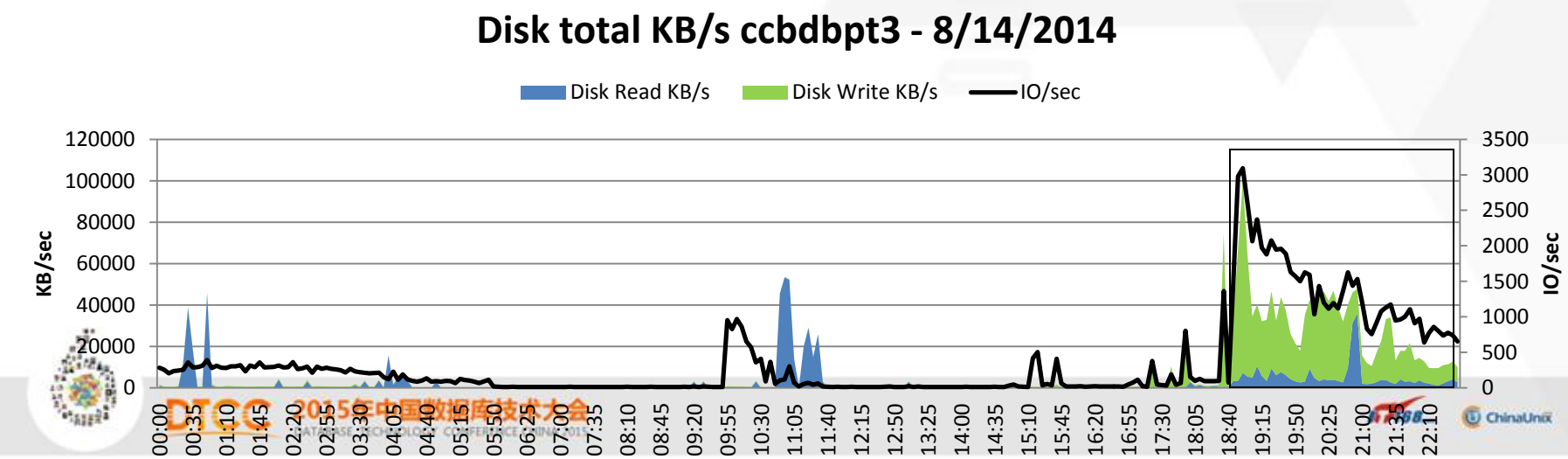
节外生枝

- 极度无奈的rollback
- 客户的发现
- 内部改进建议

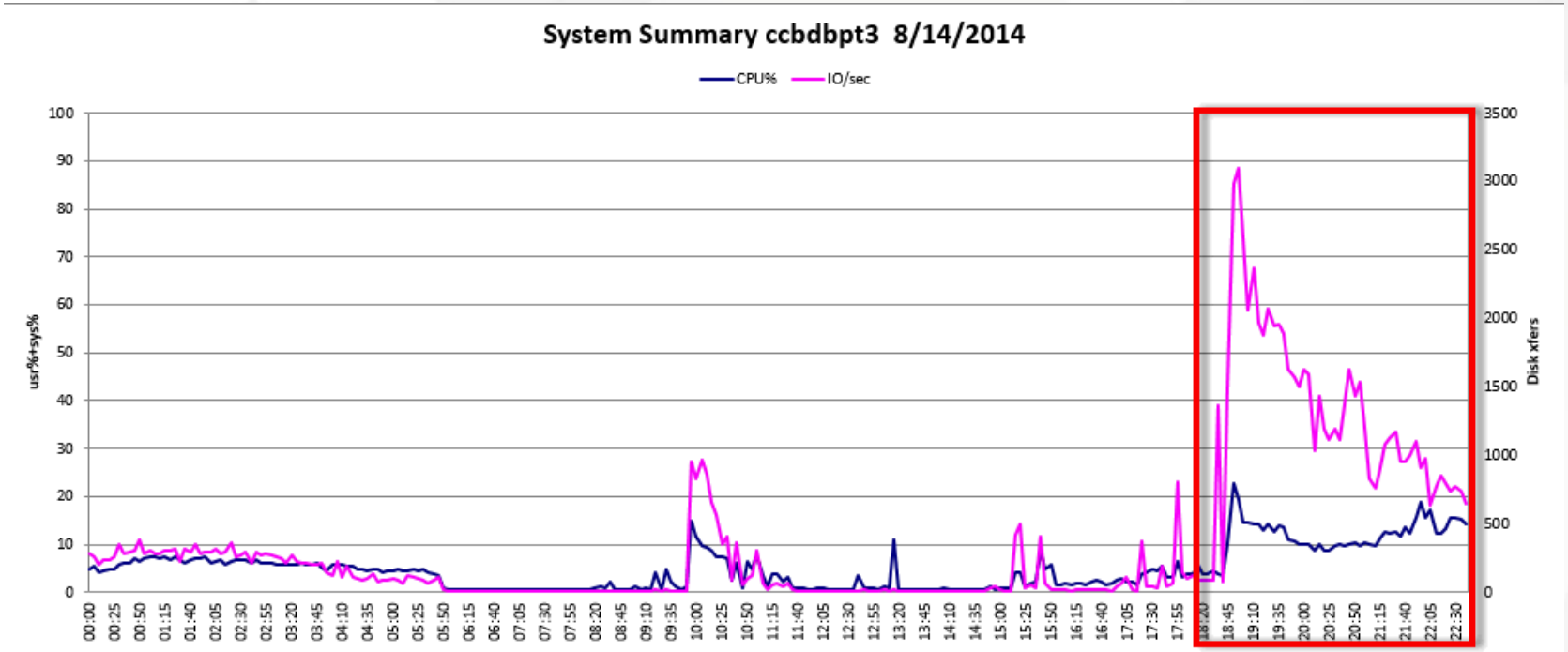
- 持续进行新方案的性能测试
  - 演练中遇到的性能瓶颈
  - 客户和同事的质疑

存储问题之外的性能瓶颈





存储问题之外的性能瓶颈



性能测试中的排除法

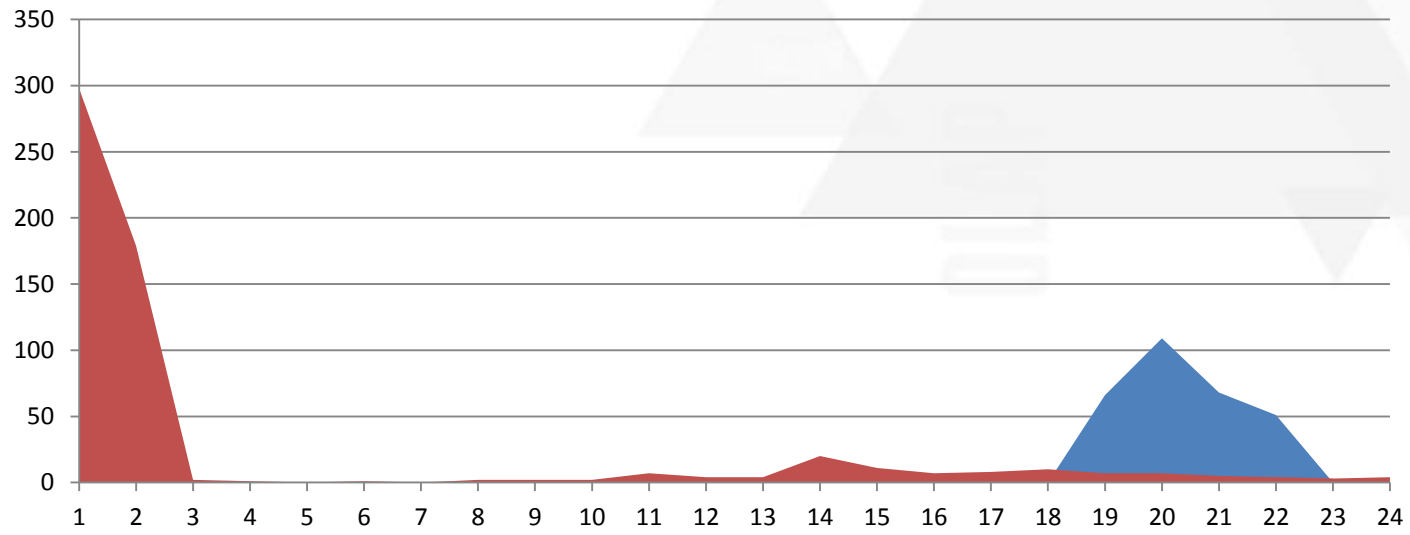
```
  _optimizer_use_feedback  = FALSE
  aq_tm_processes          = 0
Deprecated system parameters with specified values:
  background_dump_dest
  user_dump_dest
End of deprecated system parameter listing
Oracle instance running with ODM: Veritas 6.0.100.000 ODM Library, Version 2.0
Sun Aug 17 23:24:39 2014
PMON started with pid=2, OS id=20704
Sun Aug 17 23:24:39 2014
PSP0 started with pid=3, OS id=20729
```

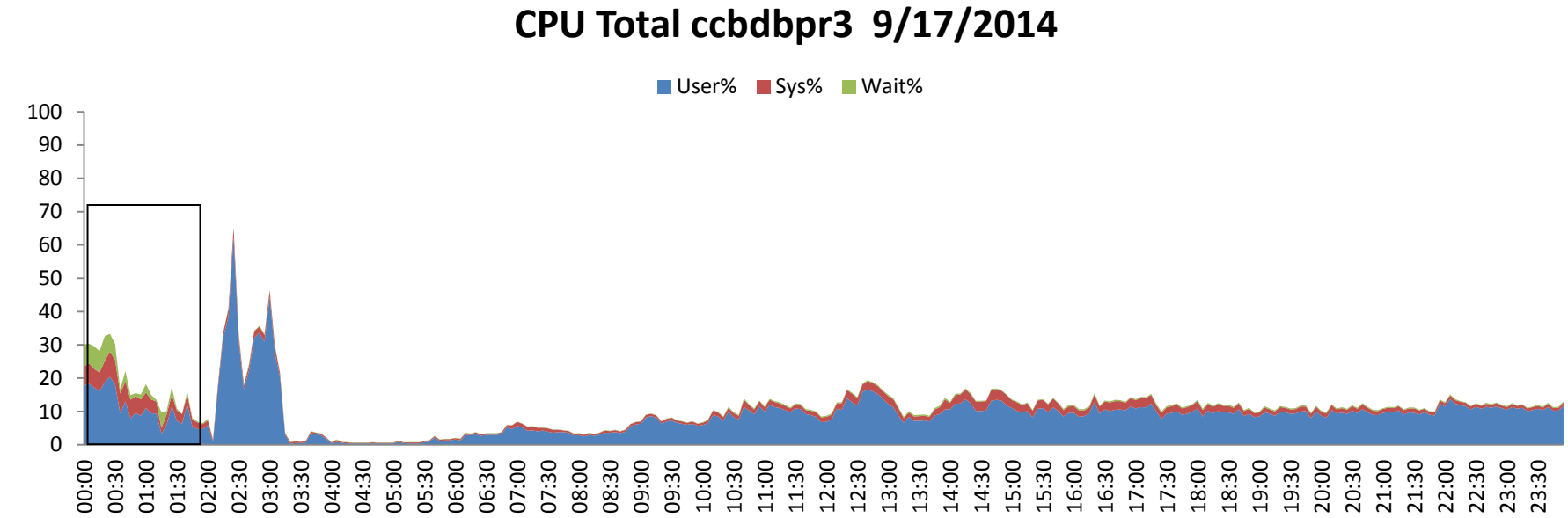
```
> ps -ef|grep odm
root     10615     1  0 Jul23 ?       00:00:17 [vxodm_ioreap]
root     10616     1  0 Jul23 ?       00:00:00 [vxodm_ioclean]
oraccbs1 24858 28913  0 12:58 pts/9    00:00:00 grep odm
```

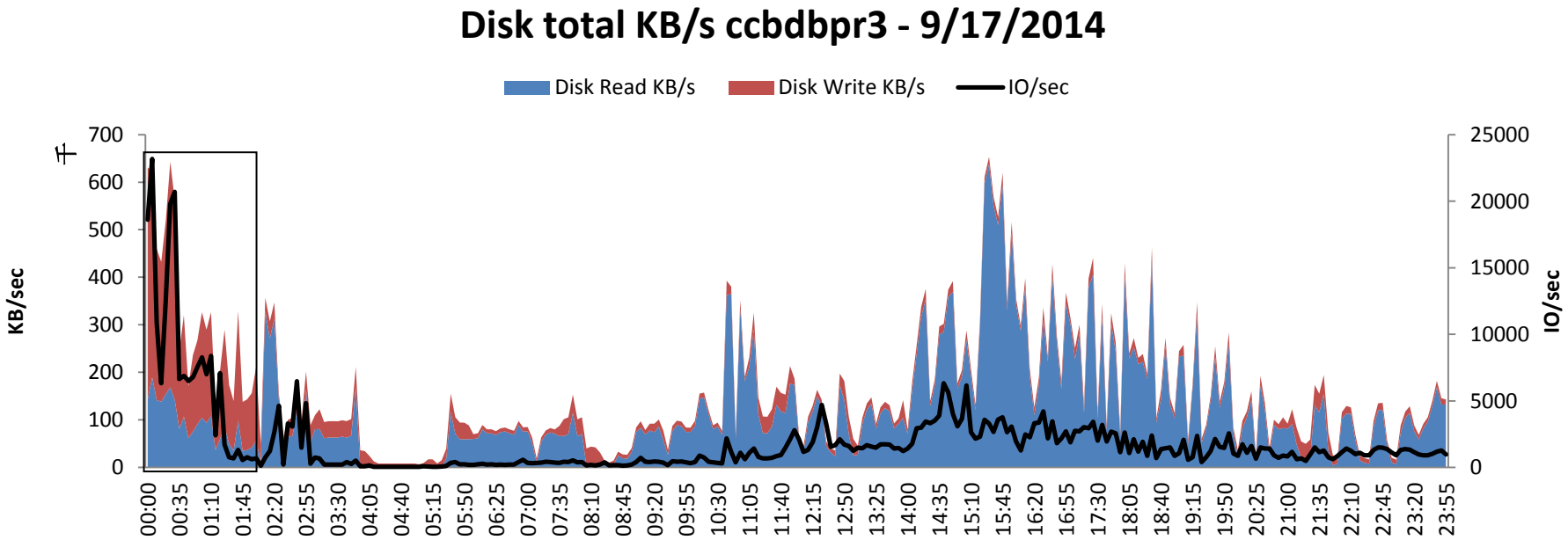
ODM性能的对比图

	时间																							
日期	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
2014.xx.13	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	1	0
2014.xx.14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**66**	**109**	**68**	**51**	xxx	xxx

	时间																							
日期	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
2014.xx.16	1	0	3	3	22	1	0	1	4	3	2	2	4	4	3	4	4	4	3	4	4	2	1	86
2014.xx.17	**298**	**179**	2	1	0	1	0	2	2	2	7	4	4	20	11	7	8	10	7	7	5	4	3	4




CPU Total ccbdbpr3 9/17/2014
User%
Sys%
Wait%
100
90
80
70
60
50
40
30
20
10
0
00:00
00:30
01:00
01:30
02:00
02:30
03:00
03:30
04:00
04:30
05:00
05:30
06:00
06:30
07:00
07:30
08:00
08:30
09:00
09:30
10:00
10:30
11:00
11:30
12:00
12:30
13:00
13:30
14:00
14:30
15:00
15:30
16:00
16:30
17:00
17:30
18:00
18:30
19:00
19:30
20:00
20:30
21:00
21:30
22:00
22:30
23:00
23:30


Disk total KB/s ccbdbpr3 - 9/17/2014
Disk Read KB/s
Disk Write KB/s
IO/sec
KB/sec
IO/sec
700
600
500
400
300
200
100
0
25000
20000
15000
10000
5000
0
00:00
00:35
01:10
01:45
02:20
02:55
03:30
04:05
04:40
05:15
05:50
06:25
07:00
07:35
08:10
08:45
09:20
09:55
10:30
11:05
11:40
12:15
12:50
13:25
14:00
14:35
15:10
15:45
16:20
16:55
17:30
18:05
18:40
19:15
19:50
20:25
21:00
21:35
22:10
22:45
23:20
23:55

数据加载进度的控制

```
 BL1_BILL_STATEMENT    3 of TOTAL    36 completed, |--processing... from      split_par_5_appendata.log
CM1_AGREEMENT_PARAM   52 of TOTAL    74 completed, |--processing... from      split_par_9_appendata.log
          MO1_MEMO    58 of TOTAL   250 completed, |--processing... from      split_par_2_appendata.log
AR1_TRANSACTION_LOG   16 of TOTAL    28 completed, |--processing... from      split_par_8_appendata.log
        AR1_CHARGES   31 of TOTAL   107 completed, |--processing... from      split_par_4_appendata.log
       BL1_DOCUMENT   21 of TOTAL    74 completed, |--processing... from     split_par_10_appendata.log
        BL1_INVOICE   23 of TOTAL    27 completed, |--processing... from      split_par_7_appendata.log
         BL1_CHARGE   30 of TOTAL   126 completed, |--processing... from      split_par_1_appendata.log
```

```
> ksh check_status.sh
837 of 1548 has finished data append process
54.06 % finished
```

数据问题的修复

- DML ERROR LOG

SQL> **EXEC DBMS_ERRLOG.create_error_log**(dml_table_name => 'BL1_DOCUMENT',SKIP_UNSUPPORTED=>true,ERR_LOG_TABLE_NAME=>'BL1_DOCUMENT_DUPLICATION');
PL/SQL procedure successfully completed.

SQL> insert into BL1_DOCUMENT select *from MIG_TMP.BL1_DOCUMENT_EXT_30 **LOG ERRORS INTO** BL1_DOCUMENT_DUPLICATION('duplicate data for BL1_DOCUMENT') **REJECT LIMIT UNLIMITED**;
198842 rows created.

SQL> commit;
Commit complete.

迁移需要注意的细节

- 1）充分的测试,评估时间,总结经验,提升性能
- 2）完整的备份策略
- 3)　网络（网络带宽和网络中断）
- 4）完整的日志
- 5）存储
- 6）归档空间
- 7）表级nologging
- 8）index级nologging
- 9）lob级nologging
- 10）foreign key
- 11）trigger的影响
- 12）goldengate的影响
- 13）主键冲突数据排除
- 14）constraint级的数据不一致


ChinaUnix
ITPUB
THANKS