Oracle 11gR2 RAC预防性优化措施

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

崔华（dbsnake）
allantreycn@gmail.com

关于我

- 中航信架构师团队的负责人
- Oracle ACE总监
- 畅销书《基于Oracle的SQL优化》作者

中国航信 TravelSky

ORACLE ACE Director

打最新PSU和Oracle推荐的Patch

- **Quick Reference to Patch Numbers for Database PSU, SPU(CPU), Bundle Patches and Patchsets (Doc ID 1454618.1)**
- **Oracle Recommended Patches -- Oracle Database (Doc ID 756671.1)**

配置Hugepage

- ALERT: Disable Transparent HugePages on SLES11, RHEL6, OL6 and UEK2 Kernels (Doc ID 1557478.1)
- HugePages and Oracle Database 11g Automatic Memory Management (AMM) on Linux (Doc ID 749851.1)
- Hugepages Not Used when ASM is used (Doc ID 1457842.1)
- ASM & Shared Pool (ORA-4031) (Doc ID 437924.1)
- Shell Script to Calculate Values Recommended Linux HugePages / HugeTLB Configuration (Doc ID 401749.1)
- HugePages on Oracle Linux 64-bit (Doc ID 361468.1)

配置Hugepage（续）

特别注意的是——如果要禁掉ASM实例的AMM，就一定不要同时reset memory_target和memory_max_target，而是应该将memory_target设为0并只reset memory_max_target。

在任意一个RAC节点执行如下操作:

alter system set sga_target=2048M scope=spfile sid='*';

alter system set pga_aggregate_target=1024M scope=spfile sid='*';

alter system set memory_target=0 scope=spfile sid='*';

alter system set memory_max_target=0 scope=spfile sid='*';

alter system reset memory_max_target scope=spfile sid='*';

```
[root@epayrac1 oracle_install]# ./hugepages_settings.sh

This script is provided by Doc ID 401749.1 from My Oracle Support
(http://support.oracle.com) where it is intended to compute values for
the recommended HugePages/HugeTLB configuration for the current shared
memory segments. Before proceeding with the execution please note following:
 * For ASM instance, it needs to configure ASMM instead of AMM.
 * The 'pga_aggregate_target' is outside the SGA and
   you should accommodate this while calculating SGA size.
 * In case you changes the DB SGA size,
   as the new SGA will not fit in the previous HugePages configuration,
   it had better disable the whole HugePages,
   start the DB with new SGA size and run the script again.
And make sure that:
 * Oracle Database instance(s) are up and running
 * Oracle Database 11g Automatic Memory Management (AMM) is not setup
   (See Doc ID 749851.1)
 * The shared memory segments can be listed by command:
       # ipcs -m


Press Enter to proceed...

Recommended setting: vm.nr_hugepages = 17030
```

停掉NTP，配置CTSSD

1、在RAC各个节点先停掉NTPD

2、cluvfy comp clocksync

3、将RAC各个节点的NTP配置文件/etc/ntp.conf改名

4、重启GI

从各个节点的 clusterware 的 alert 日志里可以看到如下内容：

节点 1：

2014-08-22 16:45:50.612:
[cssd(10223)]CRS-1601:**CSSD Reconfiguration complete. Active nodes are epayrac1 .**
2014-08-22 16:45:52.592:
[ctssd(10340)]CRS-2407:**The new Cluster Time Synchronization Service reference node is host epayrac1.**
2014-08-22 16:45:52.593:
[ctssd(10340)]CRS-2401:**The Cluster Time Synchronization Service started on host epayrac1.**

节点 2：

[cssd(10208)]CRS-1601:**CSSD Reconfiguration complete. Active nodes are epayrac1 epayrac2 .**
2014-08-22 16:46:23.794:
[ctssd(10294)]CRS-2407:**The new Cluster Time Synchronization Service reference node is host epayrac1.**
2014-08-22 16:46:23.793:
[ctssd(10294)]CRS-2401:**The Cluster Time Synchronization Service started on host epayrac2.**
2014-08-22 16:46:25.490:
[ohasd(10001)]CRS-2767:Resource state recovery not attempted for 'ora.diskmon' as its target state is OFFLINE
2014-08-22 16:46:28.218:
[ctssd(10294)]CRS-2408:**The clock on host epayrac2 has been updated by the Cluster Time Synchronization Service to be synchronous with the mean cluster time.**

调整ASM实例的LARGE_POOL_SIZE

```
alter system set large_pool_size=128M
scope=spfile sid='*';
```

关闭跨实例并行

```
alter system set parallel_force_local=true
scope=spfile sid='*';
```

设置dump文件大小上限

```
alter system set max_dump_file_size='4096M'
scope=spfile sid='*';
```

关闭UNDO_RETENTION的自动调整

```
alter system set "_undo_autotune"=false
scope=spfile sid='*';

alter system set undo_retention=10800
scope=spfile sid='*';
```

设置并行子进程的数量上限

手工设置并行子进程的数量上限为**CPU个数*2**，避免**Oracle**自动计算的值太大，并行执行时把资源耗光：

```
alter system set parallel_max_servers=128
scope=spfile sid='*';
```

关闭DRM

关闭DRM（因DRM导致的问题非常多）：

alter system set "_gc_policy_time"=0 scope=spfile sid='*';

alter system set "_gc_undo_affinity"=false sid='*';

增加实例延迟降级锁的时长

增加实例延迟降级锁的时长为3毫秒,避免遇到一些导致实例crash的bug：

alter system set "_gc_defer_time"=3 scope=spfile sid='*';

关闭数据文件的自动扩展

否则**undo**和**temp**可能超过你的空间规划，大量吃掉空间

关闭登录、登出的审计日志

11g默认打开了登录、登出审计，如果短连接比较多，则:

1、可能会导致system表空间被撑爆

2、在执行” insert into sys.aud$...” 这个SQL时大量软解析，消耗大量CPU时间，可能导致连接超时

建议关闭记录的每一次登录、登出操作。在sqlplus中执行如下语句：

noaudit create session;

关闭登录、登出的审计日志（续）

Event	Waits	Time(s)	Avg wait (ms)	% DB time	Wait Class
DB CPU		2,444		65.56	
cursor: mutex S	11,981,410	67	0	1.80	Concurrency
kksfbc child completion	400	20	50	0.54	Other
SQL*Net message from dblink	13,753	12	1	0.33	Network
log file sync	5,289	4	1	0.12	Commit

Statistic Name	Time (s)	% of DB Time
connection management call elapsed time	3,659.48	98.16
parse time elapsed	3,658.80	98.14
DB CPU	2,444.31	65.56
sql execute elapsed time	48.36	1.30
PL/SQL execution elapsed time	10.35	0.28
hard parse elapsed time	4.93	0.13
hard parse (sharing criteria) elapsed time	1.55	0.04
inbound PL/SQL rpc elapsed time	1.44	0.04

Version Count	Executions	SQL Id	SQL Module	SQL Text
4,783		4vs91dcv7u1p6		** SQL Text Not Available **

关闭自适应游标共享

```
alter system set
"_optimizer_adaptive_cursor_sharing"=false sid='*'
scope=spfile;

alter system set
"_optimizer_extended_cursor_sharing"=none sid='*'
scope=spfile;

alter system set
"_optimizer_extended_cursor_sharing_rel"=none sid='*'
scope=spfile;
```

关闭Cardinality feedback

11g的Cardinality feedback可能会导致执行计划的不稳定：

alter system set "_optimizer_use_feedback"=false sid ='*' scope=spfile;

设置28401和10949事件

```
alter system set event='28401 trace name context
forever,level 1','10949 trace name context forever,level 1'
sid='*' scope=spfile;
```

使并行会话改为使用large pool

```
alter system set "_px_use_large_pool"=true  sid ='*'
scope=spfile;
```

恢复LGWR的post/wait通知方式

```
alter system set "_use_adaptive_log_file_sync"=false
sid='*' scope=spfile;
```

禁用ora.crf

禁用ora.crf，避免osysmond进程大量吃系统资源（如CPU）

以root用户在每个RAC节点执行以下语句：

crsctl modify res ora.crf -attr "AUTO_START=never" -init

crsctl modify res ora.crf -attr "ENABLED=0" -init

crsctl stop res ora.crf -init

配置OSWatcher

- OSWatcher (Includes: [Video]) (Doc ID 301137.1)
- Document 1531223.1 OSWatcher User Guide
- Document 461053.1 OSWatcher Analyzer User Guide


ChinaUnix
ITPUB
THANKS