

Oracle中checkpoint的讨论 PDF转换可能丢失图片或格式，建议阅读原文

https://www.100test.com/kao_ti2020/253/2021_2022_Oracle_E4_B8_ADc_c102_253993.htm

什么是checkpoint? checkpoint是一个数据库事件，它将已修改的数据从高速缓存刷新到磁盘，并更新控制文件和数据文件。什么时候发生checkpoint? 我们知道了checkpoint会刷新脏数据，但什么时候会发生checkpoint呢？以下几种情况会触发checkpoint。 1.当发生日志组切换的时候 2.当符合LOG_CHECKPOINT_TIMEOUT， LOG_CHECKPOINT_INTERVAL

， fast_start_io_target,fast_start_mttr_target参数设置的时候 3.当运行ALTER SYSTEM SWITCH LOGFILE的时候 4.当运行ALTER SYSTEM CHECKPOINT的时候 5.当运行alter tablespace XXX begin backup， end backup的时候 6.当运行alter tablespace ,datafile offline的时候. 增量检查点（ incremental checkpoint） oracle8以后推出了incremental checkpoint的机制，在以前的版本里每次checkpoint时都会做一个full thread checkpoint,这样的话所有脏数据会被写到磁盘，巨大的i/o对系统性能带来很大影响。为了解决这个问题，oracle引入了checkpoint queue机制，每一个脏块会被移到检查点队列里面去，按照low rdb（第一次对此块修改对应的redo block address）来排列，靠近检查点队列尾端的数据块的low rba值是最小的，而且如果这些脏块被再次修改后它在检查点队列里的顺序也不会改变，这样就保证了越早修改的块越早写入磁盘。每隔3秒钟ckpt会去更新控制文件和数据文件，记录checkpoint执行的情况。 数据字典 完全检查点 0select * from

```
X$KCCRT where indx=0.ADDR INDX INST_ID RTNUM RTSTA  
RTCKP_SCN RTCKP_TIM RTCKP_THR RTCKP_RBA_SEQ  
RTCKP_RBA_BNO RTCKP_RBA_BOF RTCKP_ETB RTOTF  
RTOTB RTNLF RTLHF RTLFT RTCLN RTSEQ RTENB RTETS  
RTDIS RTDIT RTLHP RTSID RTOTS-----  
-----  
-----  
-----  
-----  
-----  
-----  
-----
```

```
-----4084B228 0 1 1 15 720368521 06/25/2004  
18:49:37 1 949 2 16 0600000000000000 2 0 3 1 3 1 949 1 05/16/2004  
13:29:03 0 1389 tbdb2in1 06/12/2004 12:30:50 这里显示了上一次的  
完全检查点是在06/25/2004 18:49:37发生，所以我们推  
断06/25/2004 18:49:37发生了一次日志切换，再去操作系统上  
去看生产的归档，果然18：49有一个归档生产。 -rw-r----- 1  
oracle oinstall 83532800 Jun 25 18:49 1_948.dbf 增量检查点 SQL>  
0select * from X$KCCCP where indx=0.ADDR INDX INST_ID  
CPTNO CPSTA CPFLG CPDRT CPRDBCPLRBA_SEQ  
CPLRBA_BNO CPLRBA_BOF CPODR_SEQ CPODR_BNO  
CPODR_BOF CPODSCPODT CPODT_I CPHBT CPRLS  
CPRLC CPMIDCPSDR_SEQ CPSDR_BNO CPSDR_ADB-----  
-----  
-----  
-----  
-----
```

-----4084B45C 0 1 1 2 0 10762 29753 949 76847 0 949 106814 0
721554970 06/25/2004 21:05:10 529794310 529036227 1 526310932
1413781667 949 1 0 这里显示了low-rba,on-disk rba,checkpoint
time等信息。每隔3秒钟ckpt会去更新控制文件和数据文件，
记录checkpoint执行的情况。这里应该是只更新控制文件，
每3秒不是更新数据文件 说 记录 checkpoint 的执行情况，这
个说法，没错，但不够详细，应该说，由于增量检查点和
checkpoint queue 的原理，ckpt 进程每次只是告诉 dbwr ，
写dirty buffer将要一直写到最新这个位置，仅仅是告诉 dbwr
一个 checkpoint queue 中的结束点，而 ckpt 每3秒中，在控制
文件中报告一下 dbwr 最新写入的位置。这样使得，比如数据
库要做恢复的时候（instance recovery）可以从这个最新位置
开始做恢复，而不是从数据文件中的 checkpoint scn 开始做恢
复，这样将缩短恢复时间，尤其是 instance crash 的情况下启
动更快 另外要注意的是，检查点发生的时候，ckpt 去更新数
据文件头和控制文件，并不是把当前检查点发生时候的 scn
更新进去，而是把上一次dbwr写入已经完成的检查点发生时
候的 scn 更新进去，也就是说，更新控制文件和数据文件头
是 滞后于检查点的发生的，这个从恢复的原理也很容易理解
，因为检查点发生的时候 dirty buffer还没有写入，自然不能立
即更新成当前的 scn 了。 100Test 下载频道开通，各类考试题
目直接下载。详细请访问 www.100test.com