CN SOUG 中国南方ORACLE用户组 | IT大咖说 知识共享平台

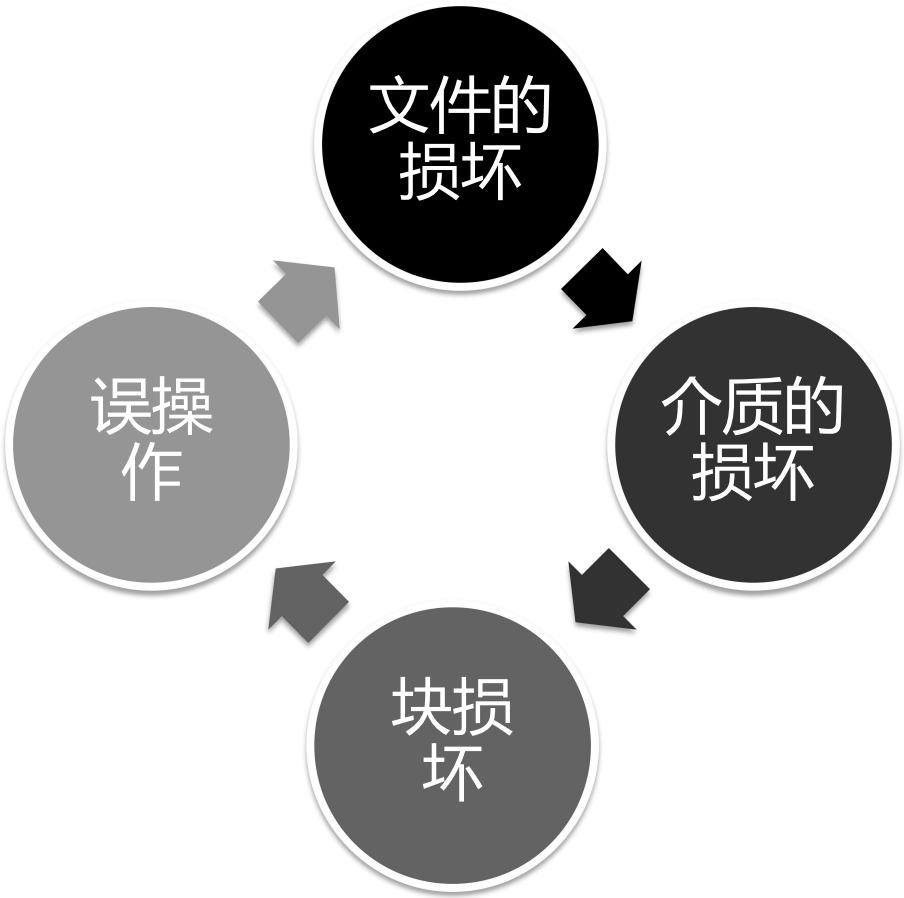
Oracle数据库特殊恢复技术漫谈

李广才 Ludatou

关于我的简介

ORACLE ACE-A. Oracle Mdata(MDU) Author. Founder of China SOUG.

最糟糕的永远在路上



损害的情形有很多,正确的评估方法十分重要;

没有备份,库损坏会让你惊喜不断;

并且恢复的手段千奇百怪;

数据库灾难表象

无法启动

访问报错

数据丢失

空访问

错乱数据

数据库损坏评估

数据库状态评估,启动状态,报错情况

集群日志以及告警日志的提示

系统以及设备日志提示

来自一线使用者的错误提示

数据库损坏评估—文件损坏

- 控制文件 control file ORA-00205等
- 数据文件 datafile ORA-1115,ORA-1110等错误
- 在线日志 online redo log ORA-00316 ORA-01623等
- 回滚表空间 undo tablespace ORA-00600 4193/4194/4037等
- 集群配置文件 ocr,voting,disk header…

数据库损坏评估—介质损坏

- 存储服务崩溃
- 硬盘坏道
- 磁盘卷组操作不当
- 光纤链路的故障或者操作不当
- 服务器内存故障(包括掉电等情况)

数据库损坏评估—最烦人心是坏块

- 坏块在核心字典表或者索引(system)
- 坏块在普通表或者索引
- 坏块在lob段或lob索引中
- 坏块在空块
- 坏块是坏道
- 坏块在redo
- 坏块在undo……等等无所不在

数据库损坏评估—误操作

- Delete 表数据,几行或者整张表
- Truncate 表数据
- Drop 表
- Drop column
- 设备误操作 导致损坏

这么多的可能性导致一个结果

这么多内容40分钟不够的

有限的时间谈坏块

ORA – 01578

ORA – 15081

ORA – 01115

ORA – 00600

.....

处理坏块的--知识储备

数据库的启动过程探索(BOOTSTRAP$以及SQL.BSQ的研究)

主要块结构的理解(堆表,索引,lob块)尝试用DD,UE,BBED操作数据

表空间的存储结构(ASSM的原理)

段的存储原理(表段,索引段,LOB段)

利用BBED以及DUMP命令探索数据块构造

研究坏块发生的场景以及模拟

坏块的屏蔽以及修复技术(进阶)

坏块产生的原因

- 硬件损坏比如存储或者光纤卡,硬盘坏道等
- 操作系统的原因
- Oracle的BUG
- 执行过“UNRECOVERABLE”或“NOLOGGING”操作的数据库进行恢复
- 掉电导致IO闪断
- 人为构造(abort,或者手工截断构造)

Oracle里坏块的几种定义情况

1.坏块只是逻辑校验的问题,本身块结构以及checksum正常的,但是块内容可能混乱或者损坏了

2.本就是空块被污染(比如异常信息覆盖)

3.物理损坏(比如存储数据的块在IO写一半中断或者内容被非块内容覆盖)

4.设备上无法访问的扇区对应的块(俗称坏道)

Oracle逻辑与物理坏块的区别

逻辑与物理坏块的情况判断 [840978.1]

一.物理块损坏：
各种各样的块损坏通常是通过ORA-1578报告出来的，详细的信息被记录在alert日志中。物理损坏的例子包括：

坏头块破坏/不完整 （Fractured Block） 块的checksum不正确 (Bad Checksum)
块的位置错误 (Block Misplaced)
归零的块/ORA-8103

二.逻辑块损坏：

当块包含一个正确的checksum和结构，但是块头一下的部分（块的内容）被损坏，可能引起不同的ORA-600错误。逻辑损坏详细的损坏信息通常不打印在alert日志中，DBV将报告逻辑损坏的块。
逻辑损坏的例子包括：
行被不存在的事务锁定-ORA-600[4512]等 使用的空间大小不等于块的大小 avsp bad等
等
。

Oracle坏块的抢救评估

1.确认坏块的产生是持续增加的还是稳固状态,有条件可以先追寻坏块产生的原因.比如是设备还是bug导致等.

2.确认坏块影响的对象以及文件 [836658.1]

3.根据确认的结果选择合适的恢复方式. [1526911.1]

有非常多的情况会导致坏块,每一种情况处理的方式都不一样,正确的评估非常重要.

Oracle坏块的常见场景

1.坏块在索引上

2.坏块在表上

3.坏块在内存中

4.坏块是数据文件头

5.坏块在redo/undo/temp中

6.坏块涉及到OBJ_ID < 60的对象

7.坏道造成的坏块

这里的每一种情况还要分不通的场景,损害的位置.
比如坏块在表段中,那么是坏在lob段里还是普通的表块,还是空块?

Oracle坏块处理 – 索引上

如果是普通的索引:

◆设置event 跳过索引坏块 38003,10231,10232,10233等,如果IOT另计

◆重建索引

◆不建议标记坏块

如在是小于<obj id 60的核心对象的索引

◆尝试bbed检查块是否完整,如果完整则修改坏块标识为正常

◆修改oracle 二进制文件让sql不扫描索引,不推荐

◆修改obj$以及boogstrap$创建索引行记录,让启动不创建索引,不推荐

◆替换bootstrap$表,在新建bootstrap$表中屏蔽相关索引,不推荐

Oracle坏块处理方法 - 表上

如果是普通的表:

◆标记坏块,让扫描跳过该块
◆通过构造ROWID的方式来抢救出其他好的块存放的表数据
◆通过bbed确认坏块是否存在部分好的的rows,通过构造rowid的方式来逐行尝试恢复
◆如果是坏道,最好先抢救出文件,再通过构建rowid的方式抢救出数据.
◆如果涉及LOB段还需要在标记后挪位log段位置或者新建表.

如在是数据库系统自带核心对象的表

◆常用借尸还魂的方法,创建组件一模一样的库,坏块涉及对象对比评估后(CTAS可以评估或者恢复工具),新建库好块覆盖坏块解决.
◆如果是坏道引起的,需抢救出数据文件,再执行第一种方法,抢救文件可能需要硬件厂商配合.
◆核心表中存储涉及应用对象信息的块损坏,则需要工具配合尽可能的抢救数据,

Oracle坏块处理方法 - 内存中

在这里 如果内存的数据坏块没有特别好的办法,

一般都归于oracle bug,也有内存的原因存在.

处理方式因场景而异.

Oracle坏块处理方法 – 各类文件上

1.控制文件坏块
◆一般损坏的都是第一个控制文件,可以用冗余的文件替代后重启DB
◆如果控制文件都损坏,在其他文件正常的情况下通过重建控制文件

2.Redo文件坏块
◆Inactive的redo文件 可以drop 再起库
◆Current/active的redo文件有坏块 需要设置允许redo文件损坏的参数,然后进行相关不完全恢复操作后进行强制起库操作,这个过程很可能碰到2662,4000等之类的错误,按错误逐一解决即可.

3.Undo文件坏块
◆一般重建undo表空间即可

4.数据文件头
◆BBED重构数据文件头
◆通过恢复工具抢救该文件上的对象

一次存储故障引起的集群宕机恢复案例

平台描述:

1.ORACLE LINUX X86_64
2.ORACLE 11.2.0.4 GRID/DATABASE

设备状况:

1.IBM DS5020 存储 双RAID, 其中RAC使用的RAID 5+1,坏一块盘,同时在换盘前增加另外一个硬盘的坏道告警,亮黄灯.

2.双RAID 5+1,经过确认1个热备盘给2个RAID做热备.

故障现象:

1.集群状态正常,
2.数据库无法启动报错

一次存储故障引起的集群宕机恢复案例

报错描述:

启动DB报错

ORA-01115: IO error reading block from file 13 (block # 133361)
ORA-01110: data file 13: '+DATA/……/ccavaya02.dbf'
ORA-15081: failed to submit an I/O operation to a disk
ORA-15081: failed to submit an I/O operation to a disk

一次存储故障引起的集群宕机恢复案例

损坏评估过程:

1.确认告警日志信息 => 确认13号文件有坏块,预估15000多个
2.确认设备损坏信息 => 坏一个盘,同时一个盘有坏带告警建议更换.坏盘已更换
3.确认集群状态信息 => 集群状态正常,数据库服务无法启动
4.确认数据库历史备份信息 => 数据库非归档状态,无rman备份,2016年3月份dmp若干
5.确认现场人员的近期操作 => 在出现数据库hang时候 现场人员拔过光纤
6.确认硬件乙方的近期操作 => 更换了损坏的盘,坏道盘未敢更换
7.保留一手现场数据 =>尝试备份现有数据文件,发现IO错误,初步判断坏道.CP/DBV/RMAN VALIDATE
8.评估损坏情况 => 经过评估,1号system文件以及13号数据文件怀疑存在坏道,无法全部读取完整既退出
9.初步怀疑坏道 =>坏道的依据 ,进程未能读取文件成功是最直接的表象,普通坏块系统级命令不会出现
10.设法抢救文件 => 通过恢复工具固有的坏道处理方法 恢复出文件.反复再三也没成功,但是恢复出来了前面数据system前面900多M.system的坏块是在3w多号
11.硬件的抢救办法/DBA的建议 =>…
12.文件抢救成功/强制起库 => ..
13.数据的最大程度抢救 / 恢复工具的介入 =>exp以及恢复工具的最大极限恢复

备份.

THANKS FOR WATCHING

感谢大家.