中国邮政储蓄银行
POSTAL SAVINGS BANK OF CHINA

这些年踩过的“坑”

中国邮政储蓄银行
POSTAL SAVINGS BANK OF CHINA

一、邮储银行简介

二、数据库使用情况

三、核心系统整体架构

四、问题及分析解决

一、邮储银行简介

中国邮政储蓄银行
POSTAL SAVINGS BANK OF CHINA

1、邮储银行历程

- 前身可以追溯到1919年，邮政储金汇业总局成立；
- 建国初期，1953年邮政储蓄业务停办；
- 1986年，邮政储蓄正式恢复开办；
- 2007年3月20日，中国邮政储蓄银行正式挂牌成立。

2、现状

- 邮储银行已成为全国网点规模最大、覆盖面最广、服务客户数量最多的商业银行。
- 截至2016年3月31日，邮储银行共有40057个营业网点，覆盖中国所有的城市和98.9%的县域地区。
- 进入2016年，日均交易笔数接近9千万，遇到类似代发养老金的业务高峰日，交易笔数可突破1亿，2016年春节期间，日交易笔数突破1.3亿。

二、数据库使用情况

中国邮政储蓄银行 POSTAL SAVINGS BANK OF CHINA

1、2004之前（Oracle 7 OPS/Oracle 8i/Informix....）

- 此阶段为省中心模式，有数据中心和交换中心之分，各省中心使用的数据库种类不一。

2、2004-2007（Oracle 9i）

- 因为各省的绿卡中心软件版本不一，数据格式转换困难，达不到以后数据大集中的需求，邮政公司决定推进绿卡统版建设。2004年绿卡统版工程正式上线，自此之后，重要生产系统中Oracle数据库基本一统天下。

- 2005-2007年实施邮政金融灾备中心建设，最终确定的使用standby db技术而不是完整的DG体系结构，采用第三方软件传输归档日志。

二、数据库使用情况

中国邮政储蓄银行 POSTAL SAVINGS BANK OF CHINA

3、2007-2011（Oracle 10g RAC + ASM）

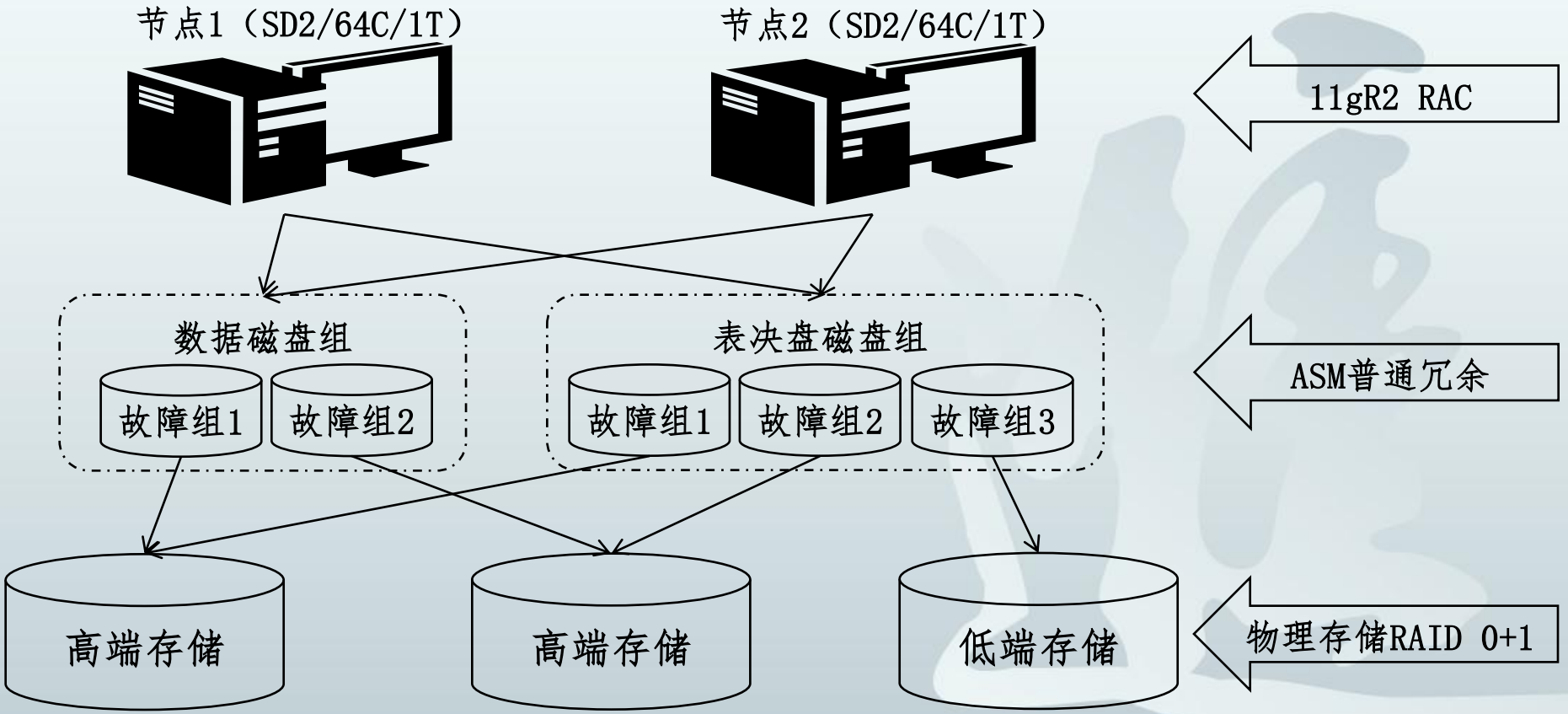
- 自2007年汇兑集中上线，新上线的生产系统基本使用Oracle 10g RAC + ASM的模式。
- 2009.08-2010.02实施物理集中工程，将31省的数据库及应用主机迁至北京
- 2010年验证能否在开放式平台上实施邮政储蓄银行逻辑大集中项目,测试结果TPS值突破1万。

4、2012至今（Oracle 11g RAC + ASM/mysql....）

- 2012年开始，新系统逐步过渡到11g R2 RAC + ASM。
- 面对越来越强的去IOE呼声，也开始使用mysql、postgresql等其它数据库。

三、核心系统整体架构

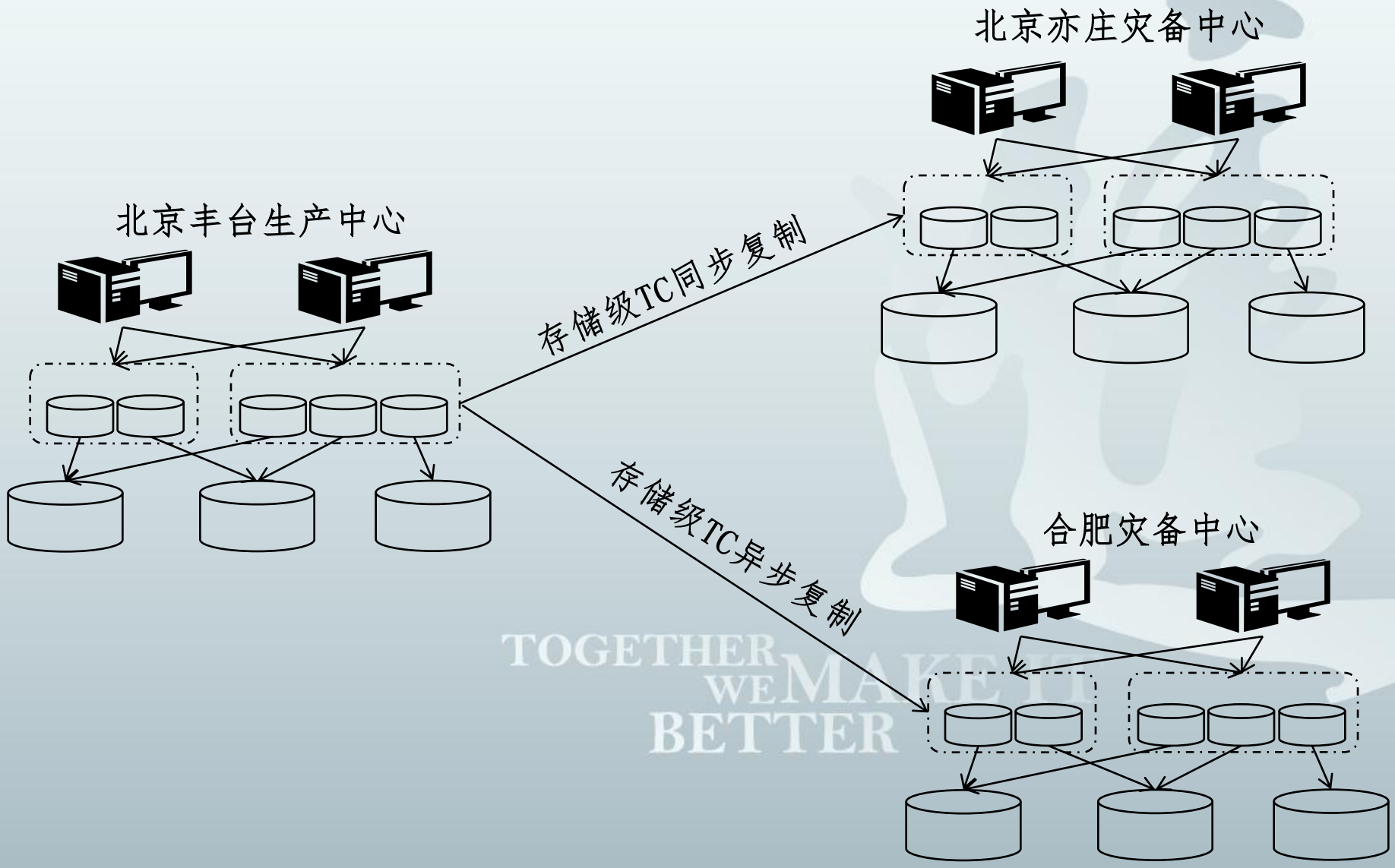
中国邮政储蓄银行 POSTAL SAVINGS BANK OF CHINA



- 上线之初为11套RAC库，目前已扩展至15套
- 特殊设计1：redo映射到存储的Cache LUN中
- 特殊设计2：OS层面上专门指定两颗CPU给lgwr进程专用

三、核心系统整体架构

中国邮政储蓄银行
POSTAL SAVINGS BANK OF CHINA



四、问题及分析解决

中国邮政储蓄银行 POSTAL SAVINGS BANK OF CHINA

1、核心系统数据库特殊设计的由来

- 为何需要如此多的存储
  - 高端存储用于存放数据库，指定一个存储一个故障组是为了防止单个存储失效导致系统不可用；
  - 引入低端存储存放表决盘是为了避免单存储故障导致集群无法启动（启动集群需要>50%的表决盘可用）；
  - 联机日志映射至存储的Cache LUN是为了消除log file sync等待事件中的IO瓶颈；
  - log file sync等待事件的罪魁祸首是等待CPU调度，为提高效率才为lgwr指定专用CPU；（压测环节中由HP实验室的ken给出建议）

四、问题及分析解决

2、好心办坏事的11g新特性

- 如果用户登录错误3次之后，开始锁定这个用户3秒钟，才允许下一次登录。这个锁定时间将从3秒逐渐延长，不断增加，此新特性本意是是为了防止暴力破解密码；
- 然而持续错误登录，早期是错误登录的用户即使使用正确密码也登录缓慢，如果频次很高，可能整个库表现都和挂起类似；
- 解决方案参见Mos Doc ID：1309738.1(Library Cache Locks Due to Invalid Login Attempts)
  - 设置EVENT="28401 TRACE NAME CONTEXT FOREVER, LEVEL 1"
  - 然而在生产环境中发现错误登录频次较高的情况下，应对措施无效
- 如果发现数据库某用户登录缓慢，且user$中相关用户的lcount值持续增加则表明已经中招了；

四、问题及分析解决

中国邮政储蓄银行 POSTAL SAVINGS BANK OF CHINA

2、好心办坏事的11g新特性

- 找出持续错误登录的终端（相关输出在alert日志中）

```
create or replace trigger ncgxq_logerr after servererror on database
begin
  if (is_servererror(1017)) then
    sys.dbms_system.ksdwrt(2,
      'by_ncgxq: DATE        = '||to_char(sysdate,'yyyy-mm-dd hh24:mi:ss')||chr(10)||
      'by_ncgxq: HOST        = '||sys_context('userenv','host')||chr(10)||
      'by_ncgxq: IP          =
'||sys_context('userenv','network_protocol')||'/'||nvl(sys_context('userenv','ip_addr
ess'),'localhost')||chr(10)||
      'by_ncgxq: OS USER     = '||sys_context('userenv','os_user')||chr(10)||
      'by_ncgxq: TERMINAL    = '||sys_context('userenv','terminal')||chr(10)||
      'by_ncgxq: MODULE      = '||sys_context('userenv','module')||chr(10)||
      'by_ncgxq: ACTION      = '||sys_context('userenv','action')||chr(10)||
      'by_ncgxq: CLIENT_INFO = '||sys_context('userenv','client_info'));
  end if;
end;
/
```

四、问题及分析解决

中国邮政储蓄银行 POSTAL SAVINGS BANK OF CHINA

3、索引跳跃扫描惹的祸

- 2012年某系统上线后即报交易缓慢，查看v$session发现大量的索引跳跃扫描事件，查看v$session_longops发现同样的语句最长执行47s
- 相关对象在两个字段建有本地分区复合索引，前导列为开户机构号，是range分区的partition key，另一字段为内部机构号；故障语句以内部机构号为条件查询，
- 解决方案：在内部机构号上建立全局索引；索引建立后本以为问题解决，但发现在新建索引上有大量的并行操作，将索引并行度改为1之后问题解决；

四、问题及分析解决

中国邮政储蓄银行 POSTAL SAVINGS BANK OF CHINA

4、接入系统上午交易堵塞问题

- 监控反馈接入系统数据库1号机堵塞，大致时间段在8:22-8:24左右，使用ashrpt采集08:20-08:26的ASH报告，及7:40-08:40的3个采样片的AWR报告，结合ASH报告及08:20-08:40的AWR报告，发现在故障时间段集中等待US和TA锁

08:22:00 (1.0 min)	96,419	enq: TA - contention	52,458	25.13
		latch: shared pool	25,737	12.33
		enq: SQ - contention	4,626	2.22
08:23:00 (1.0 min)	94,703	enq: US - contention	45,063	21.59
		latch: row cache objects	39,225	18.79

Top 5 Timed Foreground Events

Event	Waits	Time(s)	Avg wait (ms)	% DB time	Wait Class
enq: US - contention	1,457,940	60,862	42	26.33	Other
enq: TA - contention	3,434	57,081	16622	24.69	Other

四、问题及分析解决

中国邮政储蓄银行 POSTAL SAVINGS BANK OF CHINA

4、接入系统上午交易堵塞问题

•TA主要保护回滚段的DDL串行操作，据此认为应该是业务高峰导致大量回滚段从offline切换到online，从而导致交易堵塞。查看3个时间段的AWR报告中undo的统计数据，也可发现从7:40-8:40这3个采样时间段内，最大事务并发量10->123->388

Undo Segment Stats

- Most recent 35 Undostat rows, ordered by Time desc

End Time	Num Undo Blocks	Number of Transactions	Max Qry Len (s)	Max Tx Concy	Tun Ret (mins)
24-Jul 07:53	14,624	227,073	1,048	10	997
24-Jul 07:43	10,765	161,502	741	10	995

End Time	Num Undo Blocks	Number of Transactions	Max Qry Len (s)	Max Tx Concy	Tun Ret (mins)
24-Jul 08:13	36,266	592,405	1,877	16	1,007
24-Jul 08:03	25,216	408,390	1,651	123	1,004

End Time	Num Undo Blocks	Number of Transactions	Max Qry Len (s)	Max Tx Concy	Tun Ret (mins)
24-Jul 08:33	51,110	777,044	1,935	26	1,004
24-Jul 08:23	41,074	651,016	1,331	388	1,008

四、问题及分析解决

中国邮政储蓄银行 POSTAL SAVINGS BANK OF CHINA

4、接入系统上午交易堵塞问题

•当时查出的问题还有两节点事务分布不均，节点1上的联机回滚段3600左右，节点2在380左右，相差巨大，而逻辑集中11个库的不同节点间相差基本在100以下。

•问题定位后，在Oracle工程师的建议下，采取以下措施（需重启数据库）：

- 修改初始化参数_rollback_segment_count=50000（初始回滚段数量）
- events='10511 trace name context forever,level1'（保持回滚段永久联机）

•第二天又因undo表空间不够（初始回滚段过多）、高水位等待事件堵塞，在将undo表空间扩展并通知厂商部署预分配空间脚本后最终解决问题

四、问题及分析解决

中国邮政储蓄银行 POSTAL SAVINGS BANK OF CHINA

5、接入系统轧账堵塞问题

- 接入系统在下午轧账时必堵，生成堵塞时的ASH报告，发现解析占70%的DB TIME，其中硬解析占19%；
- 查看问题时段的AWR报告，在memory resize部分发现有时有缩小db cache，扩展共享池的操作，这进一步加剧了堵塞；
- 数据库配置方面，未使用AMM，但指定了SGA_MAX_SIZE和SGA_TARGET，也指定了DB CACHE的最小值，但共享池大小指定为0；
- 对数据库做如下调整并重启后（预留10G的机动内存以应急），轧账问题解决，但这只是治标，关键还是需要厂商修改应用；
  - SGA_MAX_SIZE=180G
  - SGA_TARGET=170G
  - DB_CACHE_SIZE=110G
  - SHARED_POOL_SIZE=36G

四、问题及分析解决

中国邮政储蓄银行 POSTAL SAVINGS BANK OF CHINA

6、grant造成的“血案”

- 监控系统在某天早9:20左右告警，接入系统交易堵塞，持续时间约7分钟左右；
- 查看问题时段的AWR/ASH报告，发现又是解析占据了大量的DB TIME，但当时不可能做维护操作，厂商也予以否认；
- 当下即怀疑是否做了赋权类操作，初期厂商否认，后威胁使用logminer挖出黑手，厂商维护人员始承认当时为了部署监控系统做了相关操作；

四、问题及分析解决

中国邮政储蓄银行 POSTAL SAVINGS BANK OF CHINA

7、早期灾备为何不直接使用DataGuard

•为了主备库之间解耦，灾备方案测试发现DG存在如下问题

- Oracle DG的WAN中单进程只能利用10Mbps带宽，无法充分利用155Mbps的带宽，因此将日志传输工作交由第三方软件进行；
- 因9i的bug，主备库间log相差个数过多时，易使ARCH进程全部用于远程归档而无法归档本地日志（使用文档中的workaround设置隐含参数无效）；这样在业务繁忙（如业务高峰、批处理特别是结息）或DB维护操作时，极易造成Oracle主库挂起；
- Oracle使用checksum检查归档日志的完整性，但无法保证内容的正确性；因此如果归档损坏，在备库注册归档可能：
  - 无法注册；
  - 注册成功但应用失败；
  - 注册成功且应用成功，但后果无法预料；

中国邮政储蓄银行
POSTAL SAVINGS BANK OF CHINA

谢谢！

TOGETHER WE MAKE IT BETTER