

文章编号: 1005-8451 (2017) 10-0017-04

数据库容灾系统的设计与研究

赵嘉林, 徐 卓, 张 琦

(哈尔滨铁路局 信息技术所, 哈尔滨 150006)

摘 要: 随着铁路信息系统在数据集中和应用集中方向上的不断发展, 系统运行安全成为风险控制关注的焦点, 其中, 数据库容灾系统的规划与实施更是重中之重。文章以运用Oracle Data Guard数据复制技术的哈尔滨铁路局关键业务数据库容灾系统的建设为基础, 阐述了数据库容灾系统的规划与设计过程及相关注意事项。目前, 系统已经建设完成, 可以进行关键业务数据库容灾与监测。

关键词: 数据库; 容灾; Data Guard

中图分类号: U29 : TP39 **文献标识码:** A

Database disaster recovery system

ZHAO Jialin, XU Zhuo, ZHANG Qi

(Institute of Information Technology, Harbin Railway Administration, Harbin 150006, China)

Abstract: With the continuous development of the railway information system in the direction of data centralization and application centralization, the safe operation of the system has become the focus of risk control, the planning and implementation of database disaster recovery system is the priority among priorities. This article was based on the construction of database disaster recovery system which was used Oracle Data Guard of data replication technology, and also the key business of Harbin Railway Administration. The article described the database disaster recovery system planning, design process, related matters. The system was implemented, could be used to recover disaster and monitor database.

Keywords: database; disaster recovery; Data Guard

随着各大中型企业信息数据的急剧增长和数据库应用的高速发展, 企业数据库的安全已经直接影响到企业的管理和运营。自然灾害和突发事件对企业信息系统会造成很大影响, 设备损坏等有形的损失可以弥补, 而数据丢失造成的损失则是无法估量的。数据库容灾系统作为防止数据灾难的最后防线显得更加重要。

以哈尔滨铁路局(简称: 哈局)关键业务数据库容灾系统的设计与实施为基础, 阐述了数据库容灾系统的总体设计思路以及相关解决方案。

1 数据库容灾系统设计

1.1 数据库容灾系统

容灾从层次上分为数据库容灾和应用容灾, 数据库容灾是应用容灾的基础, 是指建立一个备用的数据系统, 该备用系统可以对生产系统的关键数据

进行备份。

1.2 系统IT环境

作为一个适合的数据库容灾系统, 首先要考虑是容灾企业既有的IT环境。哈局关键业务数据库的运行平台统一为IBM POWER7系列小型机, 存储统一采用集中式的IBM的高端DS8700双存储。数据库软件也统一为Oracle 11g R2版本的双节点RAC (Real Application Cluster) 高可用模式, 逻辑存储采用Oracle本身的ASM镜像方式, 配合物理双存储建库; 这种软、硬件平台统一的IT环境现状是设计哈局数据库容灾系统的重要依据。具体系统架构基础, 如图1所示。

1.3 系统硬件架构

从哈局关键业务数据库的IT环境现状可以看到: 系统架构已经具备了防止单点宕机(含单存储)故障的高可用性, 并且搭配虚拟带库的定期自动备份和转储功能, 基本具备了一定的数据库容灾能力, 但依然有着明显的不足之处。所以在建设新的容灾系

收稿日期: 2017-03-24

作者简介: 赵嘉林, 高级工程师; 徐 卓, 高级工程师。

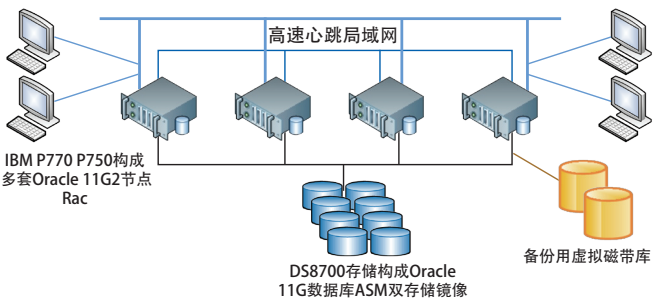


图1 数据库容灾系统架构基础

统时要做到：

(1) 弥补既有系统的不知之处。

(2) 考虑投资回报率和建设成本，新系统规模不可能是原系统的硬件复制，但灾害发生时容灾系统的RTO (Recovery Time Objective)、RPO (Recovery Point Objective) 不但要尽可能小，性能也不能低于既有系统太多。

(3) 从方案可验证性、设备利用率及容灾远期规划方面考虑，新数据库容灾系统必须实时与原数据库保持完全一致，并可验证；同时兼具只读查询等提高设备利用率的其他辅助功能。

(4) 可以方便地进行异地迁移部署，为将来建设完整的异地容灾中心做好准备。

综合上述原因，最终采用IBM的power vm小型机虚拟化平台作为新的数据库容灾系统的硬件平台。power vm小型机虚拟化容灾架构，投资远低于由物理小型机构成的原有关键业务数据库系统，又能够保证与原系统在硬件架构、操作系统及数据库层面的统一；而且每个DLPAR分区可以在线动态调

整CPU和内存资源的大小，甚至整个DLPAR分区也可以跨物理机进行自动或人工迁移，又在安全性和性能上为该容灾系统提供了可靠保证。

1.4 系统软件架构

由于原有数据库系统已具备离线式容灾功能，因此新建数据库容灾系统必须要采用在线式容灾技术；目前，在线式容灾技术均采用数据复制方式实现，根据数据复制方式采取的手段不同，可以分为以下4种类型，如表1所示。

表1 数据复制技术类型比较

类别		基本原理	平台要求	复制性能	资源占用	建设成本	数据验证	主流技术
项目								
存储层数据复制	基于普通存储的数据复制	数据的复制过程通过本地的存储系统和远端的存储系统之间的通信完成。	同构、同型号存储	高	对生产系统存储性能有影响	高，复制功能通常单独按容量收费	无需验证	IBM PPRC, EMC SRDF, HP CA (Continues Access), HDS TrueCopy
	基于虚拟化技术的数据复制	复制技术是伴随着存储局域网的出现引入的，通过构建虚拟存储来实现数据复制。	与平台无关，需要增加专有的复制服务器或带有复制功能的SAN交换机	高	对网络要求高	较高，需要专用设备	无法验证	Brocade Tapestry DMM, UIT SVM, EMC VSM
操作系统层数据复制		通过操作系统或者数据卷管理器来实现对数据的远程复制。	同构主机、异构存储	高	对生产系统主机性能有影响	较高，需要同构主机	无法验证	IBM AIX LVM, HP-UNIX MirrorDisk, Sun Solaris SVM, 专用软件: Symantec SF/VVR
数据库（应用层）数据复制		数据库的异地复制技术，通常采用日志复制功能，依靠本地和远程主机间的日志归档与传递来实现两端的数据一致。	相同操作系统和数据库平台	较高	占用少部分生产系统数据库资源	一般，部分软件免费，如Data Guard	可单独激活容灾库	Oracle DataGuard, Oracle GoldenGate, DNT IDR, DSG RealSync, Quest SharePlex

根据表1比较可以得出，数据库复制技术是当前主流数据复制技术中，建设投资最小，唯一容灾端可单独激活验证且对生产端应用性能影响较小的一种数据复制方式。综合上述考虑，哈局数据库灾备系统采用了Oracle Data Guard 数据复制技术。

1.5 系统总体架构

哈局数据库容灾系统根据机房供电、空调、消防等环境因素，将容灾端部署于相对隔离的机房C区，仅通过企业局域网与原端相连接。同时结合Oracle Data Guard技术的最大性能模式对广域网的支持，一旦条件成熟可以将整个容灾端迁移到异地机房，实现远程容灾的目标。数据库容灾系统总体架构，如图2

所示。

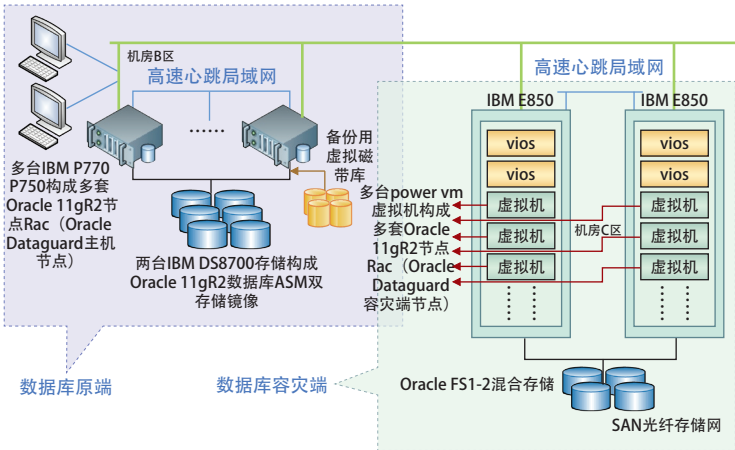


图2 数据库容灾系统总体架构

2 关键技术

2.1 Oracle Data Guard技术

Oracle Data Guard 是 Oracle 数据库从 9i 版本开始自带的数据同步功能，基本原理是将日志文件从原数据库传输到目标数据库，在目标数据库上应用这些日志文件，从而使目标数据库与源数据库保持同步，是一种数据库级别的高可用性方案。Data Guard 可以提供 Oracle 数据库的冗灾、数据保护及故障恢复等，可以实现数据库快速切换与灾难性恢复。在保证生产数据库“事务一致性”的同时，实现使用生产库的物理全备份创建备份库，备份库通过生产库传输过来的归档日志重做，条目式自动维护备用数据库。物理备份库日志应用流程，如图 3 所示。

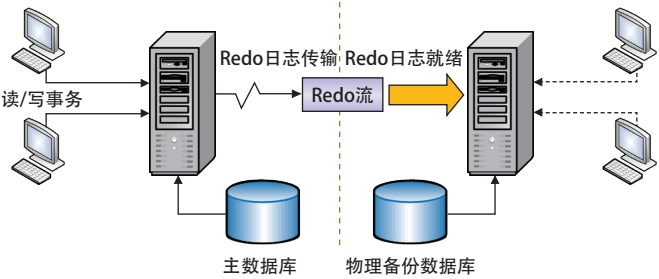


图3 Data Guard备份库日志应用流程

2.2 Oracle Data Guard技术优势

与其他数据复制技术相比，无论是 Oracle Golden Gate, 还是 DSG RealSync 或 Quest SharePlex 等，都需要按 CPU 数量或需要容灾的数据库实例数单独收费且价格不菲。而 Oracle Data Guard 包含在各企业版本的数据库软件中。只是安装与配置方式

不同于独立数据库。对于已经据有企业版的 Oracle 数据库软件的用户来说，Oracle Data Guard 几乎是免费的。

其他复制技术是逻辑层面的数据库复制，容灾数据库采用这些技术不能与源数据库物理级保证一致，其基于日志的复制机制决定了这些技术对于数据表的关联外键、级联修改 trigger 及同义词等是无法实现复制的。Oracle Data Guard 复制方式是基于渐进式的全库恢复的，容灾端和主数据库是实时保持物理级别的全库一致的，创建各种同义词、trigger 及授收权等各种 DML (Data Manipulation Language) 和 DDL (Data Definition Language) 都会自动同步，容灾端和主数据库的切换也非常简单，适合进行定期的容灾切换应急演练。

2.3 Oracle Data Guard与数据库备份的关系

Oracle 11g R2 版本的 Data Guard 可以实现 3 种类型的备用数据库，即物理备用数据库、快照备用数据库和逻辑备用数据库。这 3 种类型的备用数据库虽然应用的场景不同，但都是采用“日志复制”技术对主数据库进行数据复制操作；因此会出现将主数据库中的程序逻辑错误、人为操作错误全部复制过来的问题。为了防止上述情况的发生，既经济又行之有效的方法就是对主数据库的定期备份，包括采用 rman 的全库、归档备份和各种 exp 方式的逻辑备份（那种不惜代价、建立多个不同类型备份数据库定期轮流切换的极端场景不在考虑范围内）。各种数据库备份也是处理一般数据库故障的有力保障，避免频繁地进行 Data Guard 主备数据库的切换。

2.4 Oracle Data Guard物理备用数据库可用性检测

为了保证 Oracle Data Guard 的备用数据库实时可用性，可以通过在主库建立自动更新的心跳表来监视备用库，具体数据库脚本如下：

```
create table dg_heartbeat tablespace users as
select dbid, name, sysdate as last_update from
v$databases;

var job number
exec dbms_job.submit(job => :job, what =>
'update dg_heartbeat set last_update=sysdate;', next_
```



```
date => trunc(sysdate, 'MM') + 1/1440, interval =>
'sysdate + 1/1440');
print job
commit;
JOB
-----
442
```

在备份库执行如下查询, 查看主备数据库的实时差异, 其中, LAST_UPDATE 字段是备份库最后更新的时间, DIFFERENCE 是主库和备份库的时间差。在 Data Guard 的最大性能模式下, 该差值在 1 min 之内表明主备库同步正常。

3 结束语

此设计方案采用基于渐进式、全库恢复的数据同步 Data Guard 解决方案, 具有数据一致性好、配

置简单以及数据“零损失”等特点, 此方案已成功用于哈尔滨铁路局各生产系统数据库容灾应用中, 它实现了运行数据的实时复制, 确保业务数据的安全完整, 也为后期异地容灾做好了坚实基础。

参考文献:

- [1] 冯宝军. 沈阳铁路局客票系统容灾的实现 [J]. 铁路计算机应用, 2003, 12 (11): 40-41.
- [2] 王茂林, 蒲全武, 郭伟. 轨道交通 AFC 系统数据库容灾系统方案设计 [J]. 铁路计算机应用, 2010, 19 (2): 47-49.
- [3] 吴建芳, 王慧丽, 张自立. 铁路车站客票异地灾备系统的设计与实现 [J]. 铁路计算机应用, 2009, 18 (7): 21-25.
- [4] 王亦然. Oracle 数据库灾备技术探讨 [J]. 微电脑世界, 2011, 17 (11): 94-96.
- [5] 王新伟. 基于 Oracle 数据库的逻辑数据同步技术在实践中的应用研究 [J]. 电子技术与软件工程, 2015 (8): 195-197.

责任编辑 徐侃春

· 信息 ·

2017年度《铁路计算机应用》全路通讯员岗位培训圆满结束

2017年9月20日至21日, 由《铁路计算机应用》编辑部主办的全路通讯员岗位培训成功举办。来自铁路总公司、15个铁路局共计23位通讯员齐聚北京, 围绕智能时代的创意与知识进行了认真学习研讨。

中国铁道科学研究院电子计算技术研究所所长史天运出席了开班仪式。史所长指出: 近年来, 《铁路计算机应用》紧跟时代, 强化互联网思维, 推行多媒体融合战略, 但仍然要植根于铁路行业, 要与铁路总公司、各路局及各兄弟单位共同发展, 合力推进铁路信息化建设。

两天的培训时间, 学员先后听取了编辑部王浩主任关于“2016-2017年度《铁路计算机应用》期刊情况”的工作报告, 资深互联网人、知识专家方军的《智能时代的创意与知识》, 编辑部付思关于“2017年度《铁路计算机应用》微信公众号运营情况”的工作汇报等内容, 并前往中国铁道科学研究院院史馆和新华三集团科技展厅进行了参观学习, 最后大家针对如何更好地开展新媒体工作进行了交流与探讨。培训以课堂讲授、专题讨论、参观学习等形式进行; 内容丰富, 形式多样; 既有宏观, 又有微

观; 既有理论, 又有实践; 分析问题, 找准对策。

培训结束后, 学员纷纷表示, 本次学习非常及时和必要, 收获很大, 受益匪浅。培训老师请得好, 课程设计安排得好, 培训研讨成效好。学员之间相互学习, 相互交流, 开阔了视野, 开拓了思路, 为进一步推动《铁路计算机应用》发展、开创编辑部工作新局面起到了很好的促进作用。



《铁路计算机应用》编辑部