



基于 Goldengate 的数据库异地灾备实现

宓正宇

(中国电信股份有限公司上海分公司, 上海 200120)

摘要: 随着电信关键业务系统越来越受到重视, IT 类需求水涨船高。IT 系统应该更可靠、更快速、更易用以及更适应业务的变化发展; 另一方面, 重要业务平台的用户数据和业务数据的集中存放带来的安全风险也不断加大。重点阐述了 Goldengate 同步复制技术的概念和实现原理, 研究了该技术的特性和优势, 通过实际案例的分析, 验证了采用 Goldengate 的数据复制技术的可行性, 最终实现数据实时备份, 确保核心数据库的安全, 同时避免引入过多种类的软件, 降低了运营维护的复杂度和成本投入, 有利于灾备系统的恢复和实时切换。

关键词: Goldengate; 备份; 容灾; 数据库

中图分类号: TP309.3

文献标识码: A

doi:10.11959/j.issn.1000-0801.2018104

Design and implementation on remote disaster recovery system based on Goldengate

MI Zhengyu

Shanghai Branch of China Telecom Co., Ltd., Shanghai 200120, China

Abstract: As telecom business-critical systems become more and more important, the demand for IT classes has also increased. The IT system should be more reliable, faster and easier to use and to adapt better to changing business needs. On the other hand, the security risks are also increasing. The concept of synchronous replication technology and the realization theory of Goldengate was focused on, the characteristics and advantages of the technology were discussed. Through the analysis of actual cases, the feasibility of using Goldengate data replication technology was verified, the data real-time backup was realized, the core database security was ensured while avoiding the introduction of too many categories software, and the complexity of operation and cost inputs was reduced, also it is conducive to the recovery of disaster recovery system and real-time switching.

Key words: Goldengate, backup, disaster recovery, database

1 引言

中国电信股份有限公司上海分公司(以下简称上海电信)业务平台建设从一开始的分散建设

模式到现在的集中化建设模式, 实现了业务平台的集约化运营与管理。业务平台的集中化建设可以减少运营商对硬件的投资, 减少运营的维护成本, 便于集约化运营管理等。但由此带来的安全

隐患也不能忽视,例如火灾、地震等不可抗拒因素,用户数据和业务数据集中存放的单点故障会对公司造成不可预计的损失。因此上海电信统筹安排了临港新城电信机房作为公司IT系统的新建EDC机房,以实现电子化数据异地高速备份,提升IT系统的数据备份能力,并为后续进一步的应用级容灾做好准备。同时异地灾备的数据库还可以分散主节点数据库的系统访问压力,扩展了数据文件的实时读写能力。

本文探讨了基于Oracle的Goldengate技术实现对电信IBP系统用户及业务数据进行异地备份,同时考虑到公众客户售中管控模块的上线会对现有IBP生产数据库系统产生影响,建议该管控模块所需的IBP相关数据从IBP复制库获取;同时以后ODS的数据抽取也将从IBP复制库获取,减少生产系统的压力。

2 Goldengate 灾备系统技术

2.1 灾备技术

灾备技术应用主要源于对数据的保护,其核心的技术思想是在异地创建副本。纵观业界的灾备产品都是基于以下几个数据保护与同步复制技术,包括基于传统备份的灾备技术、基于镜像的灾备技术、基于复制的灾备技术。其中,基于复制的灾备技术又可以分为基于数据库的复制技术、基于存储的复制技术、基于存储交换层的复

制技术、基于主机软件的复制技术。本文主要涉及的是基于数据库的复制技术,因为基于数据库的复制技术总体投资规模较小,而且不要求数据集中存储,对网络带宽要求较低,但是实施难度大,技术成熟度也一般^[1]。

2.2 Goldengate 技术

Goldengate 软件是一个实现异构数据环境间数据复制的综合软件分组,它通过分析源数据库的在线日志或归档日志记录数据的增/删/改变化,然后将这些增/删/改的操作同步到目标数据库,最终实现目标数据库与源数据库的双活和同步。另外该软件提供异构环境下交易数据的捕捉、转换和接收,可以满足异构数据库环境下的变化数据同步,而且即使是异构操作系统或异构数据库都可以实现大量数据亚秒级的同步。Goldengate 能实现一对一、一对多、多对一、双向和点对点等灵活的拓扑结构并应用到计费系统、报表系统、供应链系统等多个场景,满足跟踪、同步、分发和备份容灾的要求^[2]。

(1) 单向数据复制

单向数据复制简单的说就是利用 extract 抓取进程在源数据库捕获在线日志或者归档日志并在 extract 队列中记录源数据库增/删/改等操作,然后通过 deliver 传输进程接收这些变化量,最后通过 replicate 入库进程创建复制或者同步的SQL语句,最终在目标数据库中进行执行操作。单向数据复制过程的示意如图1所示。

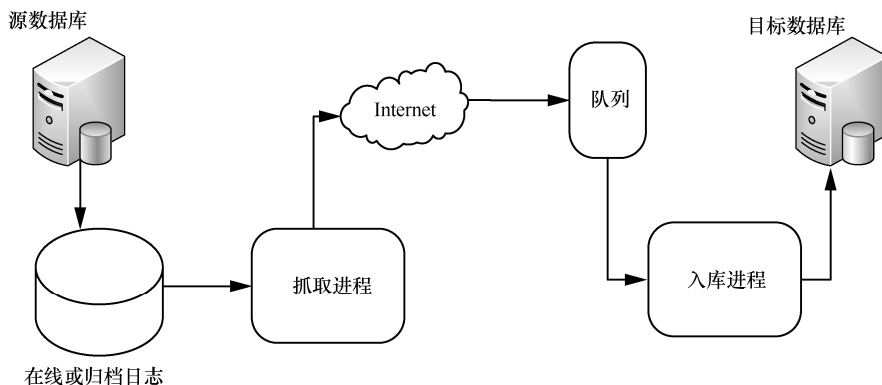


图1 单向数据复制示意



(2) 双向数据复制

双向数据复制的原理与单向数据复制的原理基本一致，只是说两端的源数据/目标数据互相作为复制对象，并且源数据库和目标数据库同时为双活状态。

但是双向数据复制存在一个问题，那就是如果不采取有效的机制，刚被复制进目标数据库的变化数据马上会被目标数据库的抓取进程复制回源数据库，为了避免这种死循环状态，Goldengate采用了一种非常有效的判断机制来识别哪些数据可以用来被抓取。也可以这样理解，这种判断机制其实就是在抓取进程中使用了跟踪表机制，当复制进程和应用程序同时更新同一个对象表的时候，抓取进程就开始启用跟踪表机制，在进行双向数据复制中，抓取进程就通过命令行向源数据库和目标数据库中分别加入跟踪表，一旦抓取进程发现跟踪表有数据更新，就判断为复制进程所产生，并忽略掉该操作；反之，如果更新表没有数据更新，那么抓取进程就判断为由应用程序产生，并将更新操作抓取出来。双向数据复制的过程示意如图2所示。

(3) 其他复制方式

除了单向复制和双向复制方式以外，Goldengate还提供了其他的复制方案，利用灵活

的技术架构可以满足用户不同的复制需求。例如广播复制、集中复制和多层复制等。

2.2.1 关键特性

Goldengate的配置方式十分灵活，包括单向数据复制、双向数据复制和多层次数据复制。可以满足用户本地驻地网或者城域网中的不同场景和需求。

该同步软件有以下3个关键特性。

(1) 实时访问实时信息

Goldengate提供异构系统间数据库事务的实时捕获、转换、路由和交付。该软件在确保事务完整性的同时让更多数据库和平台实现了高性能、低影响的亚秒级数据移动。它采用一个基于组件的体系结构，可帮助企业系统所需的持续可用性和实时集成。

(2) 保持关键系统的持续可用

Goldengate可以帮助系统恢复因意外或者计划造成的终端停机，并提高系统性能和可伸缩性。该软件可以支持下列情形，包括灾难恢复、数据保护、零停机运营、数据分布和查询卸载等。

(3) 稳定的模块化体系结构

Goldengate软件体系包括3个主要组件：抓取、跟踪队列和交付。这种模块化方式让每个组件可以独立运行，从而加快数据复制并确保数据完整性^[3]。

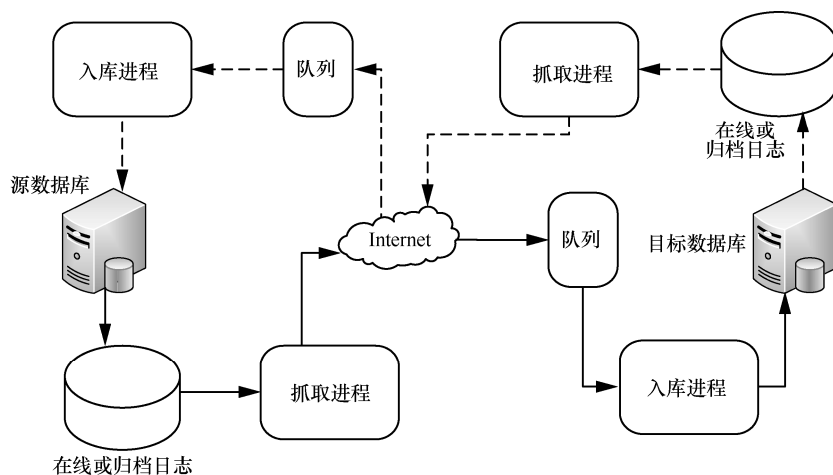


图2 双向数据复制示意

2.2.2 特性和优势

Goldengate 软件具备很多特性和优势,可以在平台系统间快速移动事务数据。包括以下几点。

(1) 实时数据

立即捕获、转换事务数据并将其交付给其他系统,延迟不到 1 s,可以让系统业务平台均可获得及时、准确的信息,改善组织决策。

(2) 异构支持

利用异构数据库和平台提高 IT 部门的灵活性。从现有的 IT 资源中提取数据,在统一所有企业用户业务数据的同时降低了成本。

(3) 可靠性

将所有提交的记录交付给目标系统,即使在网络中断时,也无需中断系统。

(4) 高性能、低影响

每秒移动上千个事务,且不会影响源系统或目标系统。实时访问关键信息,且不影响生产系统。

(5) 事务完整性

在源系统和目标系统间移动数据时,保持数据提交边界的一致性、隔离性、持久性等属性,确保备份系统和报表数据库间数据的一致性和应用的完整性。

2.2.3 Goldengate 与其他备份方式比较

基于数据库备份的软件还有 Dataguard 和 Stream 等,表 1 对 3 种实现方式进行了详细的比较。

表 1 3 种备份实现方式的比较

	Stream	Dataguard	Goldengate
是否支持跨操作系统版本	是	否	是
是否支持异构数据库平台	是(需另外购买组件)	否	是
是否需要数据库归档	是	是	否
是否支持部分数据复制	否	否	是
是否支持目标端的数据转换	是	否	是
是否影响数据库性能	是	否	否
是否支持多种复制场景	是	否	是
备用数据库是否处于活动组状态	是	否	是
是否收费	否	否	是

基于 Goldengate 软件与存储远程镜像的比较见表 2。

表 2 基于 Goldengate 软件与存储远程镜像的比较

	Storage mirror	Goldengate
源端同步采集方式	块复制	归档日志
是否支持跨平台(DB, 系统)	否	是
目标端是否支持读写	否	是
源库一致性	不完全一致	逻辑一致
是否支持断点续传	否	是
使用场景	灾备	灾备+数据同步
传输效率	块复制	高可压缩
带宽要求	高	低
数据层面切换时间	时间较长	无需切换

因此基于上述原因综合评估后,下面主要通过具体案例探讨基于 Goldengate 技术实现 IBP 业务平台的数据库备份及数据库同步以解决业务模块增加访问 IBP 系统数据库压力的方案。

3 案例

3.1 IBP (服务开通) 系统

目前上海本地的 IBP 系统已经综合了 SDH 和基于 SDH 的多业务传送平台 MSTP 业务,因此上海本地的 SDH 业务和 MSTP 业务都是由 IBP 系统统一管控的,IBP 系统承受的压力可想而知。

3.2 IBP 系统现状

业务开通流程的系统支撑现状如图 3 所示。

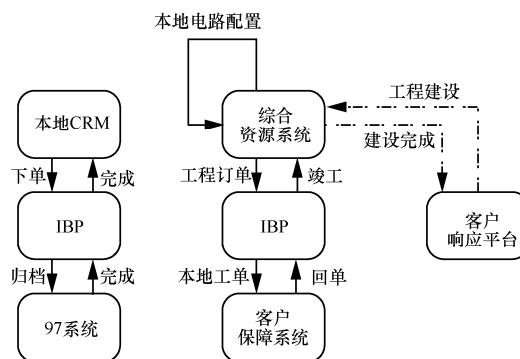


图 3 业务开通流程的系统支撑现状

上海电信 IBP 系统自 2008 年上线后,系统压力不断上升;除了要承受系统本身的新产品和新



业务的压力外，还要满足很多外来系统的业务发展需求，具体如下。

- ODS 系统每天从 IBP 抽取数据。
- 公众客户售中管控模块对服务开通跨系统流程的关键环节进行实时监测要求。
- 公众客户售中管控模块对服务开通跨系统流程的系统衔接环节进行实时监测要求。

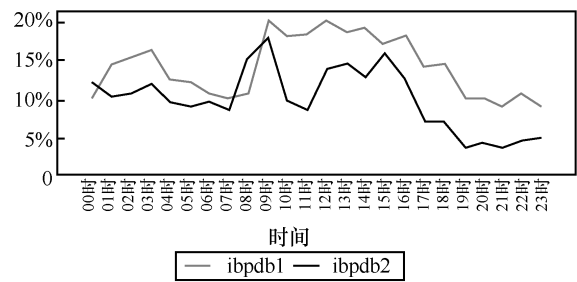
考虑到公众客户售中管控模块的上线会对现有 IBP 生产数据库系统产生影响，因此该管控模块所需的 IBP 相关数据从 IBP 复制库获取；同时以后 ODS 的数据抽取也将从 IBP 复制库获取，减少生产系统的压力。于是就有了本文中基于 Goldengate 来同步生产 IBP 数据库数据到 IBP 复制库的方法。

3.2.1 系统网络拓扑

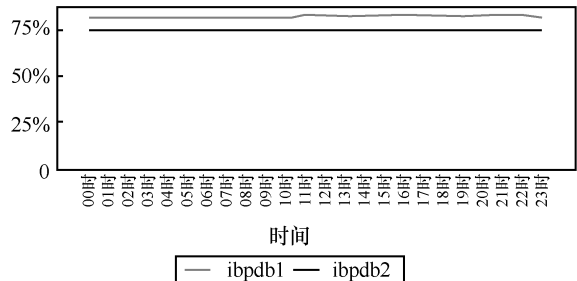
系统网络拓扑如图 4 所示。

3.2.2 数据库服务器性能与日志

数据库服务器如图 5 所示。



(a) CPU利用率



(b) 内存使用率

图 5 数据库服务器性能

固网数据库日志产生量每天约 224 GB。

CDMA 网数据库日志产生量每天约 139 GB。

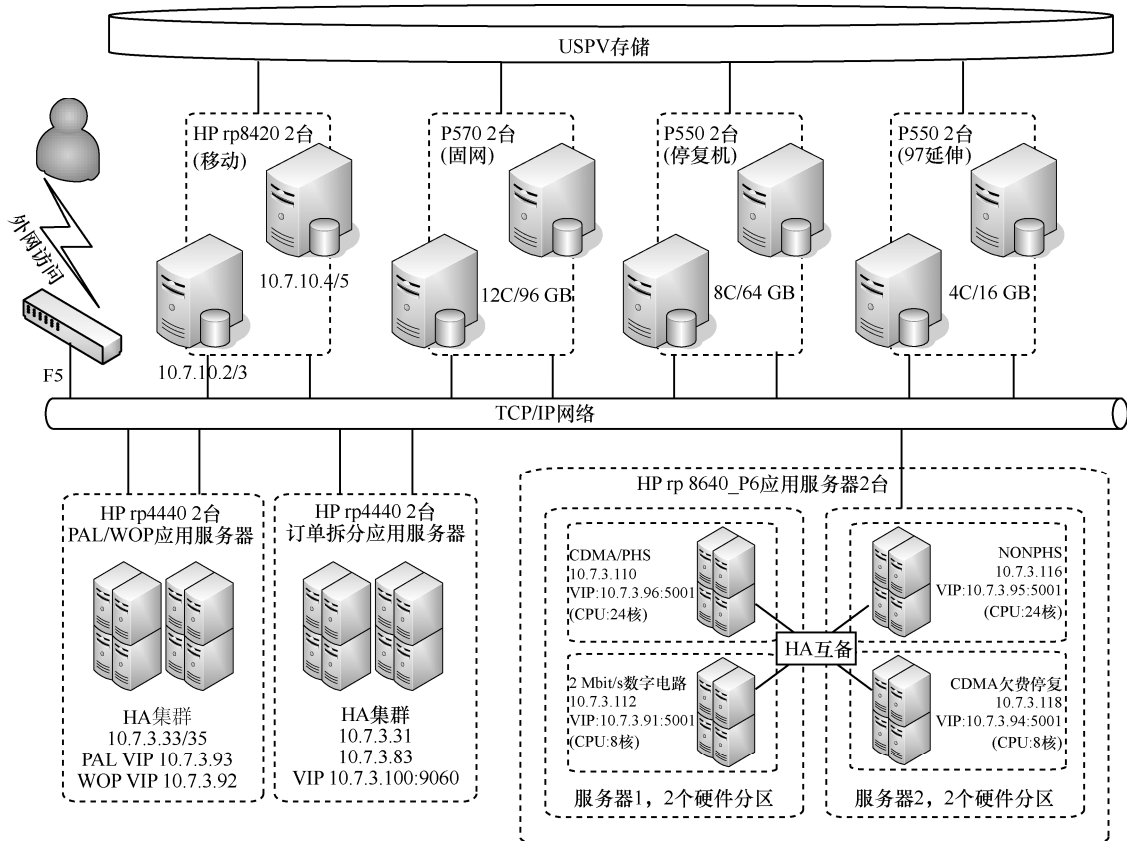


图 4 系统网络拓扑

3.3 基于 Goldengate 复制数据库

Goldengate 数据复制软件当前采用基于 LOG 的方法获取数据变化,可以在相同或不同的 Oracle 数据库版本进行复制,或者可以从 Oracle 数据库复制数据到其他类型的数据库(异构复制)。目前,Goldengate 只配置了 DML 复制,对于所有 DDL 操作均不予以复制^[4]。

3.3.1 Goldengate 支持类型及操作

(1) 数据支持类型

支持 numeric(数字类型),包括 number、binary float、binary double。暂不支持 binary_integer 和 PLS_integer。

支持所有 character(字符类型),包括 char、varchar2、long、nchar、nvarchar2。

支持大对象,包括 clob、nclob、blob。但不支持 bfile。

支持 binary(二进制类型),包括 raw 和 long raw。

支持 date 及 timestamp 类型。支持除了 timezone_region 和 timezone_abbr 类型的所有 timestamp 类型。

暂不支持多字节的 XML 类型。

支持用户自定义类型(UDT),源端的 UDT 与目标端的 UDT 必须相同。

其他支持的数据类型,包括 rowid、varray、interval day、interval year。

Goldengate 暂不支持的其他数据类型: anydata 和 anydataset、anytype、mlslabel、uritype 和 urowid。

(2) 特殊对象支持

Goldengate 支持对于常规表的数据复制,对于一些特殊对象支持如下所示。

有条件支持物化视图(materialized view),在以下有条件限制中支持:

- 源表必须有唯一主键;
- Goldengate 不支持物化视图使用“with

rowid”来创建,但当物化视图 LOG(不是物化视图自己)用“with rowid”来创建时,Goldengate 是支持的;

- 物化视图必须是使用单个表,而不能包含 joins 生成的表;
- Goldengate 不支持 truncates 物化视图,但支持使用“delete from”来替代。
- 索引组织表(index organized table)通过物化视图复制;
- 对于 Goldengate 基于 LOG 的复制方式,IOT 表无法直接从日志中抽取,但是可以为 IOT 表建立物化视图,通过复制该物化视图的变化实现 IoT 数据变化的复制。

(3) 不支持的特殊操作

以下特殊操作由于不写日志,Goldengate 无法予以复制。

- direct-path table load(由于不进行写 LOG 操作)

Oracle 插入数据有两种方式:常规插入和 direct-path 插入。常规插入,重新使用 table 中的自由空间,在已有数据中插入新数据,维护引用完整性约束。direct-path 插入,在表中已有数据之后插入新数据;数据直接插入数据文件,绕过缓存;已有数据中的自由空间没有被重新利用;忽略了引用完整性约束。direct-path 会写较少的 redo log 和 undo log,导致 Goldengate 无法通过日志获取足够信息捕捉数据变化。

常见的 direct path 插入包括:Insert/*+ APPEND */ into table1 as select table2; create table 1 as select table2。

- 并行 DML 模式下缺省为 direct-path
如: ALTER SESSION { ENABLE | FORCE } PARALLEL DML。

- 使用 SQL *loader 并且设置 direct=y

如果遇到上述操作,一般这些表数据不是原生数据,可以在复制中排除掉这些表并通过定时



备份等方式予以复制。

- /*+ APPEND */ hint

如果在应用中使用该 hint，并且表被设置为 nologging，则不会写足够日志，Goldengate 无法复制。

- /*+ BUFFER */ hint

如果在应用中使用该 hint，则不会写足够日志，Goldengate 无法复制。

- 不支持 truncate 操作

因为在数据库维护的时候 truncate 操作是不去写数据库日志的。

3.3.2 复制数据库

(1) 复制内容和数据量

本次复制的内容只需针对 IBP 系统中的 3 个模块，分别见表 3。

表 3 IBP 系统中的 3 个模块

模块名称	数据库用户	备注
P6	PROVISIONING	流程管理系统
WOP	ORDERSPLIT	人工工单系统
SOP	WOPDB	订单拆分系统

通过调研 CDMA 网和固网数据库里的用户，本次复制的数据量见表 4。

以未来一年数据量的等比增长，需要复制的数据量为：

$$\text{CDMA 网: } 1.6+1.6+1.6+1.6=6.4 \text{ TB} \quad (1)$$

$$\text{固网: } 1+0.2+1+0.2=2.4 \text{ TB} \quad (2)$$

(2) 软硬件环境

本次复制工具使用的是 Oracle 公司的 Goldengate，鉴于 Goldengate 是通过将主库所有的

变化信息应用到复制库的方式，所需的主机性能建议和生产库的性能相当，同时由于 I/O 的性能直接影响了复制的效果，建议存储使用高端存储，所配置容量与生成库相当。根据笔者的实际实施经验，目标库的机器性能建议和生产库相当，但可以略微低一点。

在主机配置方面，鉴于生产环境数据库内部的 SCHEMA 的名称又相同，建议建立各自单独的复制库，建议配置见表 5。

在存储配置方面，根据式 (1)、式 (2) 中以一年为基础估算的存储数据量为：

CDMA 网复制库：6.4 TB；

固网复制库：2.4 TB。

(3) Goldengate 部署及测试

本案例 Goldengate 的基本步骤如下所示。

- 整理需要复制的表的相关信息。包括：表的属主、表结构、索引、主外键、表的大小、表的用途、表所在的表空间和表采用的复制方式。
- 安装主机和操作系统，配置网络，安装数据库软件。包括：操作系统的安装、主机命名、IP 地址分配、文件系统的划分。存储的划分和分配，数据库软件的规划和安装及配置。
- 安装 Goldengate 软件及测试。包括 Goldengate 软件在生产库和复制主机之间的安装、部署。复制脚本的编写和测试，复制程序的压力测试。
- 系统割接前的复制内容全备份。
- 系统割接前的生产库和复制库的同步。
- 系统割接时的生产库和复制库的同步。

表 4 CDMA 网和固网数据库里的数据量

数据库	主机	内存	存储大小	数据库用户	存储使用 (1 年时间)
IBP (CDMA 网)	HP 8420	32 GB	6.5 TB	ORDERSPLIT	1.6 TB
				PROVISIONING	1.6 TB
SIBP (固网)	IBP P570	96 GB	6 TB	WOPDB	180 GB
				P6SPTV	1 TB

表5 建议配置

数据库	主机	CPU	内存	同步模块
IBP_COPY1	HP 8640	12 核	32 GB	P6、SOP
SIBP_COPY1	IBP P570	10 核	32 GB	P6、WOP

- 系统割接后的生产库和复制库的同步检测。
- 系统割接后的生产库和复制库的主机性能检测。
- 系统割接后的生产库应用测试。
- 系统割接后的复制库应用测试。

4 使用 Goldengate 的要求及注意事项

4.1 时钟同步

对于 RAC 环境, Goldengate 要求所有节点必须保持时钟同步, 同时必须保持所有 RAC 节点和运行的抽取进程的节点保持时钟同步。因为 Goldengate 会比较本地的系统时间和 commit 的时间戳。所以不能忽略这个设置。否则可能导致数据复制的紊乱, 因此 RAC 环境下务必使用 NTP (网络时间协议) 进行时间同步^[5]。

4.2 共享存储

在使用 Goldengate 数据同步的时候, Goldengate 的相关软件和工作目录建议最好配置在共享存储环境中, 从而保证对所有节点都是可用的, 从任何一个节点都可以启动 Goldengate 的进程, 当其中一个节点出现异常时, 可以在剩余的节点启动而无须修改任何配置参数。否则如果运行在单个节点上的话, 需要将剩余节点中的归档日志通过 NFS 文件共享的方式加载到 Goldengate 运行节点。

4.3 定期清理归档

开启数据库归档后, 需要部署清理归档日志的定时任务, 以防磁盘空间被占满导致数据库无法启动的故障, 包括主节点和容灾节点都要定时清理归档。

5 结束语

业务平台集中建设和集约化运营管理的发展

是大势所趋, 越来越多的业务平台走向集中, 特别是随着中国电信移动业务的蓬勃发展, 各业务或产品的基地化建设和运营, 各种用户数据, 业务数据的集中存放是否安全, 各种重要的业务平台是否有切实可行的容灾措施, 是目前急需解决的问题。本文通过探讨基于 Oracle 的 Goldengate 技术实现业务平台数据库的异地备份, 通过复制 IBP 数据库系统的案例很大程度上说明复制库为缓解生产库的压力起到了很明显的效果, 总之希望 Goldengate 能对现网业务平台的数据库级容灾备份提供有效的帮助。

参考文献:

- [1] 罗圣美, 李明, 叶郁文. 大数据容灾备份技术挑战和增量备份解决方案[J]. 大数据, 2015, 1(3): 106-112.
LUO S M, LI M, YE Y W. Challenge and solution of big data backup and recovery[J]. Big Data Research, 2015, 1(3): 106-112.
- [2] Oracle Goldengate administrator's guide[R]. 2012.
- [3] 曲波, 邓旭东, 姜峰. Oracle Goldengate 数据同步机制研究与应用[J]. 微型电脑应用, 2014, 30(6): 55-58.
QU B, DENG X D, JIANG F, et al. Research and application of Oracle Goldengate data synchronization mechanism[J]. Micro-computer Applications, 2014, 30(6): 55-58.
- [4] 李鹏, 于洪涛, 徐静波. 七号信令监测系统中基于 Oracle 的数据同步方案研究[J]. 电信科学, 2010, 26(2): 60-64.
LI P, YU H T, XU J B. Research on the scheme of Oracle-based data synchronization in the SS7 monitoring system[J]. Telecommunications Science, 2010, 26(2): 60-64.
- [5] 佟敏, 李方村. 关于 BOSS 异地容灾系统建设的讨论[J]. 电信科学, 2004, 20(7): 66-69.
TONG M, LI F C. Discussion on the construction of BOSS disaster recovery system[J]. Telecommunications Science, 2004, 20(7): 66-69.

[作者简介]



宓正宇 (1983-), 男, 中国电信股份有限公司上海分公司信息网络部工程师, 主要从事 IT 项目管理, 云计算项目、全球眼平台项目的规划、立项、实施方面的工作。