

# Oracle10g 分布式数据库技术在云南 地震观测数据库系统中的应用<sup>\*</sup>

武孔春<sup>1</sup>, 牛从达<sup>1</sup>, 倪泰山<sup>2</sup>

(1. 云南省地震局, 昆明 650224 2 云南省红河州地震局, 云南蒙自 661100)

**摘要:** 介绍了 Oracle 10g 分布式数据库技术的基本原理, 以及分布式数据库在地震观测数据处理方面的优势; 应用 Oracle 10g 数据库技术, 设计了云南地震观测分布式数据库系统结构, 阐述了分布式地震观测数据库的实现过程及相关技术, 有效地解决了数据分散和集中管理的矛盾。

**关键词:** 分布式数据库系统; 数据共享; 数据复制; 分布式查询

**中图分类号:** TP392 P315 **文献标识码:** A **文章编号:** 1000-0666(2008)01-0089-06

## 0 引言

在地震信息项目数据库建设中, 目前中国地震局已完成了“十五”前兆数据库、测震数据库等的设计和建库工作, 并将在全国范围内部署和应用。云南省地震局根据前兆规范和地震监测物理数据模型设计规范, 也对云南“十五”前兆数据库、测震数据库建设工作进行了部署。数据库建成后可供全省 16 个地、市地震局的科研人员使用。但随着全省网络系统的建成和 Oracle 10g 数据库系统在全省范围的使用, 各地、市地震局之间的数据交换和处理及其与省中心节点之间的数据交换存在许多问题, 如何处理数据的分散性和管理的集中性, 以及系统的可靠性和稳定性, 是目前需要解决的问题, 分布式数据库系统技术为解决这个问题提供了一个很好的模式。笔者将详细讨论如何利用分布式数据库技术有效实现地震数据的交换和管理。

## 1 Oracle10g 分布式数据库技术

分布式数据库技术是分布式技术与数据库技术的结合, 在数据库研究领域中有多年的历史。从概念上讲, 分布式数据库是物理上分散在计算机网络各节点上, 而逻辑上属于同一个系统的数

据集合(图 1)。它具有数据的分布性和数据库间的协调性两大特点。系统强调整节点的“自治性”而不强调系统的集中控制, 且系统应保持数据的分布透明性, 从而在编写应用程序时可完全不考虑数据的分布情况(Wu, 2000; 周龙骧, 1996)。分布式数据库系统具有如下基本特点: 物理分布性, 逻辑整体性, 场地“自治性”, 场地之间的协作性。为了保持分布式数据库系统的一致性, 还必须建立分布式事务处理。

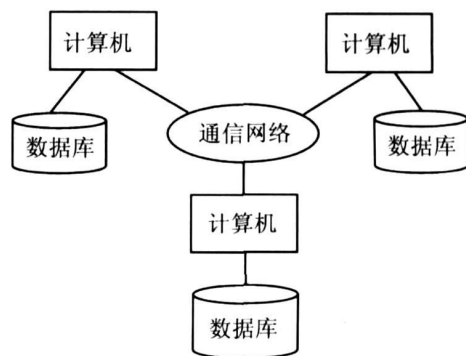


图 1 分布式数据库系统示意图

与集中式数据库系统不同, 数据冗余在分布式系统中被看作是所需要的特性, 其原因在于: 首先, 如果在需要的节点复制数据, 则可以提高局部的应用性; 其次, 当某节点发生故障时, 可以操作其它节点上的复制数据, 从而增加系统的

<sup>\*</sup> 收稿日期: 2007-07-31.

基金项目: 云南省地震信息服务系统建设项目和云南省重点项目(云计地区[2002]54号-02-02)联合资助。

有效性。

在分布式数据库中，数据存储通过以下 3 种途径实现：①复制：系统维护关系的几个完全相同的副本，这些副本存储在不同的节点上；②分片：关系被划分为几个片段，各个片段存储在不同的节点上；③复制 + 分片：关系被划分为几个片段，系统为每个片段维护几个副本。

在地震数据共享系统中，我们使用 Oracle 10g 数据库管理系统实现地震数据的交换和分布式数据库管理。Oracle 高级复制管理器提供了几个功能强大的管理工具，如复制目录、分布式模式管理以及其他一些 Oracle 复制管理工具。Oracle 10g 数据库的复制与 MS SQL Server 数据库的复制有一些差别，它不再涉及出版服务器、订阅服务器、分发服务器、出版物等（长城工作室数据组，2001），但它有自己特有的复制解决方案。与 MS SQL Server 数据库提供的复制工具一样，Oracle 10g 也提供了一个图形界面的复制管理工具，即 Oracle 复制管理器，通过该工具可以配置、定时和通过一个节点来管理整个的复制环境。该管理工具可以单独运行，也可以在 Oracle 企业管理器中通过一个 applet 来激活。通过 Oracle 复制管理器，可以任意部署一个复制组到一个新的节点（路川等，2006 王海亮等，2004）。

云南省 18 个地震分布式数据库节点的复制，可利用 Oracle 高级复制的分布式模式管理功能对数据库进行维护。这种管理功能允许在一个控制点就可以进行复制环境的定义和改变。对复制环境的任何改变，都将自动复制并且应用这些数据定义语言（DDL）命令到其它复制环境节点，同时利用数据库触发器和复制相关过程自动生成复制所需要的一些支持对象，这对云南地震观测数据库共享系统的维护具有极为重要的意义，对“十五”数据库的任何改变都可以通过在省中心完成全省数据库表的更新来实现。

在地震数据共享系统中，通过现已建成的主传输链路和备份传输链路实现远程复制及分片，主传输链路采用 SDH 链路，备份传输链路采用 VPN 连接，以 GRE 实现隧道封装，并使用 SuSE Linux 系统平台，保证了系统的可靠性和稳定性（牛从达等，2005 张云峰等，2000）。由于分片在云南地震数据共享系统中起次要作用，我们主要考虑了数据库复制的问题。Oracle 10g 数据库的

复制包含多主复制和可更新快照复制，根据业务需求，可以单独使用，也可以结合起来使用，同时，根据特定的任务采用各种不同的复制策略。数据库复制是解决各地、市地震局之间数据交换和处理及它们与省中心节点之间数据交换的一种很好的方法。

2 分布式地震观测数据库系统

按照“十五”云南地震信息服务系统行业管理模式，根据分布式数据库系统的技术要求，全省共设置了 18 个地震分布式数据库节点。其中，在云南省地震局建立数据库系统中心节点，设置多个服务器，作为全省地震观测数据中心利用西部防震减灾中心平台，在下关滇西地震预报实验场建立全省地震数据异地备份中心；在云南 16 个地、市地震局建立分布式数据库系统子节点，负责本地、市地震观测数据的管理和运用。云南分布式地震观测数据库系统结构如图 2 所示。

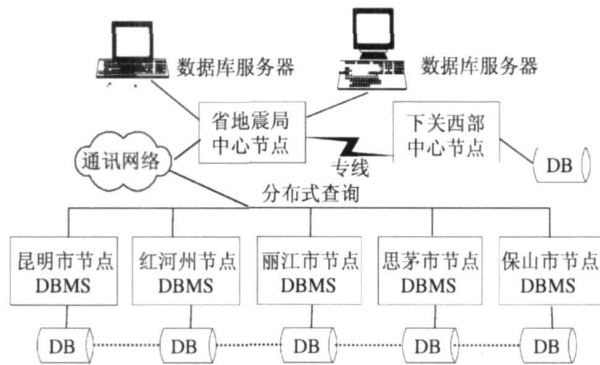


图 2 云南分布式地震观测数据库系统结构图

从全局应用的角度出发，这些数据库自下而上构成分布式数据库系统，实现全局数据的完整性和一致性。各地方节点仍然存放本地的数据，省中心节点的数据库则存放所有的业务数据，并对数据进行完整性和一致性的检查。这种做法虽然有一定的数据冗余，但在不同场地存储同一数据的多个副本，不仅提高系统的可靠性和可用性，也提高了局部应用的效率，减少了通讯代价。该分布式数据库系统可以在对当前机构影响最小的情况下进行扩充，同时也使得各处理机之间的相互干扰降到最低。

复制策略综合考虑各种因素，根据不同的数

据共享策略采用不同的复制方案,如数据联系的紧密程度、网络的可靠性、费用等。在此笔者选择 Oracle 10g 的高级复制功能和异步复制应用模型。Oracle 高级复制支持整个表的复制和基于部分表的复制两种方案。这两种复制方案主要是通过 Oracle 的多主复制和可更新快照复制,以及这两种复制机制的组合来完成。复制策略采用“复制”,但不启用分布式分区视图,使省中心节点汇集各地方节点的数据,而各地方节点保持各自的数据。

### 3 云南分布式数据库的设计

在地震数据共享系统中,根据中国地震局的要求,将使用 Oracle 10g 数据库管理系统来实现“十五”地震观测数据的交换和分布式数据库的管理,其中将通过全省网络系统来实现远程复制,并根据不同的数据共享策略采用不同的复制方案。

在地震数据共享系统中,复制策略选择异步分发数据库技术,这是一种通过延迟远程存取和延迟传播实现数据更新的技术。这种技术具有很高的可用性和很短的响应时间,其传播最终将保证所有复制节点间的数据一致。同时为避免更新冲突,采用动态所有权模型,该模型保证在同一时刻只有一个特定的节点可以对数据进行更新。

实际数据工作流程中,对速报目录的实时性要求最高,通常要求地震发生后 30 分钟内得到地震定位参数,所以速报目录表至少半小时更新一次,月报目录表每月更新一次,地震观测报告表每天更新一次,数字化及模拟观测前兆数据库每天更新一次(牛从达等, 2005)。根据对数据的实际需求,在省中心节点和各地、市节点上采用多主复制、可更新快照复制和混合复制方案。多主复制方案支持全表在各个主节点间的对称复制,允许所有主节点对主表都有更新操作的权利。可更新快照复制既可以是包含一个主表的完全拷贝,也可以是为满足基于值的选择标准的主表中行的子集。快照是由查询语句定义的逻辑数据结构,该查询参考了一个或多个远端主表中的数据,从而决定了快照所包含的数据内容。混合复制将多主复制和可更新快照复制结合在一起,构成一种新的混合配置,这种配置可以完成对全表或者子表的灵活复制,完全满足地震数据共享系统的

要求。

多主复制方式用于省中心节点合并地方节点数据,即多个地方节点数据库向省中心节点数据库进行数据库复制,这样省中心节点就可以保存全省的数据。但地方节点也可能需要其它地方节点的数据,如红河州地方节点数据库需要保山地方节点数据库的数据,这些数据可以通过可更新快照复制由省中心节点实现基于值的选择主表中行的子集来实现。我们设计了 16 个地州的可更新快照复制数据库。

在数据库复制中,首先得保证网络通畅,可通过 Ping 进行测试;其次要保证复制的网络链路中不存在防火墙,即防火墙必须位于网络链路的两侧。目前已建成的全省网络系统完全能满足数据库复制的要求。数据库复制能够成功的另一个条件是数据库系统能够相互注册,且必须以 IP 地址注册。因此对动态 IP 是不能进行数据库复制的。

对于前兆数据库,省中心节点和各地方节点软件系统相同,数据库结构也相同,实际上各节点都采用同一个“十五”数据库系统,在全省部署时已进行分发和安装。利用现已建设完成的数据库系统,修改原来观测数据入库流程,16 个地、市节点的前兆观测数据不再由省中心入库,而是由 16 个地、市节点汇集本地数据并入库,省中心节点合并各地、市节点数据库中的数据。测震数据库存储云南省测震台网产出的速报目录、月报目录和观测报告,这些文件目前保存在 MySQL 数据库中,很容易实现异库数据库转换(DTS),保存为 Oracle 数据库,由省中心节点负责数据入库和维护。因此有不同的复制策略。

在多个地方节点与省中心节点的数据库复制中,由于采用初始同步架构可能出现冲突,因此必须采用手工同步。手工同步首先创建链接服务器到要同步的服务器,再创建触发器实行远程更新,同时还要考虑数据库复制中的帐号问题和更新冲突,这时就需要编写存储过程对系统表进行操作。

省中心节点合并地方节点前兆数据库,但地方节点保存本地的前兆数据,这些数据通过多主复制在省中心节点合并。为了方便复制,需要在地方节点创建一个单独的复制表空间和一个模式(schema),系统中的其他用户通过一个私有同义词来存取这些复制。

但地方节点也需要实现地震数据的交换,这可由省中心节点通过可更新快照复制将数据传递到地方节点。通过 Oracle 复制管理器,可通过省中心节点来配置、定时和管理整个复制环境,以适合不同的数据同步方式。因为节点较多,且复制关系复杂,需要复制的表也很多,所以需要相应增大一些初始化参数的值,这可以通过修改不同的参数试验获得。省中心节点和地方节点需要通过以下步骤来搭建复制环境:

(1) 在省中心节点和各地方节点数据库上增加需要同步的数据库的 `services name`, 可以直接在 `tnsnames.ora` 文件中增加数据库的 `services name` 包括 IP 地址、SD 以及端口号等。所有复制的节点要全部加入到 `tnsnames.ora` 文件中。

(2) 创建一个用于各节点连接数据库的数据库链接 (`dblink`)。出于安全考虑,可以采用一个私有数据链接。

(3) 在各节点创建一个名为 `snapshot- dba` 的表空间来存放快照,并创建一个和该表空间有关的名为 `Snap` 的用户。在省中心节点创建一个名为 `replication- dba` 的表空间来存放多主复制,并创建一个和该表空间有关的名为 `Repl` 的用户。

(4) 运行一个特定的脚本文件来生成创建各节点数据库上用户表的快照脚本或多主复制脚本。

(5) 创建一个定时刷新过程来定时刷新复制。

(6) 省中心节点应保持各节点的统一维护代码表。

## 4 分布式数据库的实现

新设计的“十五”前兆数据库与“九五”前兆数据库相比,有很大的改进,数据库表的结构完全不同,表的数量大幅增加,数字化观测数据和模拟观测数据由不同的标识进行区分。目前“九五”前兆数据库的数字化观测数据表的数据,可以通过程序转入“十五”数字化前兆数据库。目前,“十五”前兆数据库由中国地震局统一下发,已在云南省中心节点和 16 个地、市节点进行安装,并配备相应的数据入库与管理软件。因此在省中心节点建立主站点,地方节点建立实体化视图站点,采用实体化视图复制,即一个主站点提供源复制对象,其它实体化视图站点拷贝主站点数据,这些数据通过可更新快照复制由省中心

节点实现基于值的选择主表中行的子集来实现,即只考虑复制地方节点的数据。

(1) 定义复制组。Oracle 使用复制组来管理复制对象。复制组包含前兆数据库中所有表,可以实现对数据库对象的集中管理。按照数据分类,我们选择 16 个地、市的数据表建立 16 个复制组,16 个地、市地方节点建立实体化视图站点。省中心节点的数据表也建立 1 个复制组。只需用省中心节点的“复制管理员”这一个用户就可以完成大多数的高级复制配置。

(2) 建立复制通道。省中心节点将运行主站点程序,在主站点程序中,主要通过调用 `DBMSREPCAT` 包中的过程来完成对主站点复制环境的配置。16 个地、市将运行快照站点程序,快照站点程序的初始阶段与主站点程序类似,主要是完成对快照站点复制环境的设置。然后创建一个数据库链接来支持快照,从而使快照与主站点的复制对象建立对应关系。为了支持快照的刷新,在主站点上还要创建快照日志。最后调用 `create snapshot` 程序来创建快照,随后将快照加入快照组和刷新组中。复制建立后,当省中心节点的数据发生更新时,复制接收者将该变化应用到复制的数据上。

“九五”前兆模拟观测数据库目前由省中心节点维护和管理,这些数据可通过程序转移到“十五”前兆数据库的模拟观测数据表中。随着“十五”前兆数据库在全省的分发和部署,“十五”前兆数据库的模拟观测数据表将在地方节点进行维护和管理。由于地方节点只存放本地的数据,数据管理和分析功能是由省中心节点的数据库服务器来实现的,各地方节点的数据库需要在省中心节点的数据库上汇总,因此复制的关系流程与省中心节点向各地方节点分发数据的情况相反,主站点程序由 16 个地、市运行,省中心节点则运行快照站点程序,然后通过实体化视图在省中心节点进行数据的合并。

例如在建立红河州本地的“十五”前兆数据库后,模拟前兆观测数据表由红河州本地维护和管理,包括入库和完整性检查,由中国地震局分发的配套软件完成。红河州节点的数据库将运行主站点程序,在主站点程序中,主要通过调用 `DBMSREPCAT` 包中的过程来完成对主站点复制环境的配置。而快照站点程序在省中心节点运行,

通过数据库链接建立的实体化视图由查询语句定义，该查询参考了省中心节点主表中的数据，从而决定了快照所包含的数据内容。省中心节点的数据库可同时订阅 16 个地、市的数据库表，以达

到数据库汇总。  
地方节点也可能需要其它地、市的数据库表，可以采用“十五”前兆数据库的数字化观测数据表的复制方案（表 1）。

表 1 地震观测数据库复制方案

数据表或存储过程	主站点/实体化视图站点	复制过程	复制方式	事务调度间隔
前兆数据库表 (本地观测部分)	本地节点建主站点，省中心节点建实体化视图站点	从本地节点到省中心节点	采用异步复制应用模型，可更新快照复制	1 天
测震数据库表 (由 MySQL 通过 DTS 到 Oracle 10g)	省中心节点建主站点，本地节点建实体化视图站点	从省中心节点到 16 个地、市节点	采用异步复制应用模型，可更新快照复制	30 分钟
前兆数据库表 (本地及临区数据)	省中心节点建主站点，本地节点建实体化视图站点	从省中心节点到 16 个地、市节点	采用异步复制应用模型，可更新快照复制	1 天

## 5 利用分布式技术实现事务处理和查询

分布式数据库系统中数据的分布使查询也具有了分布性，分布式查询可能针对异类的 DB 或 ODBC 数据源。Oracle 10 g 支持分布式查询，即包括来自两个或更多服务器数据的查询，支持服务器间的检索、更新和游标，并保证节点间事务语义，维护服务器间的安全。分布式查询对动态 IP 服务器特别有用。

在地、市节点上，存储在本地的数据可直接实现本地查询，本地查询约占 80%。存储在异地节点的数据，利用 Oracle 10 g 数据库管理系统中的链接服务器功能，创建节点与异地数据库节点的链接服务器，通过建立实体化视图实现对远程数据库的访问，为用户提供对云南地震观测数据库执行分布式查询、更新、命令和事务的功能。

实体化视图是 Oracle 的一种数据库对象，是“概要管理”组件的一部分，它提供了一种对远程数据快速访问和报告数据的方法。物理上它存在于数据库内部，可以包括联接和/或聚合。它能够在执行之前预先计算所占资源大的联接和聚合操作，它的存在缩短了查询执行时间。

因此，分布式查询可以利用 Oracle 数据库复制建立的实体化视图来实现对远程数据库的访问，而且数据库用户完全无需知道存在哪些实体化视图，就像在本地查询一样透明地对远程数据库进行查询。由于省中心节点和各地方节点都建立了

实体化视图，因此在各地方节点之间及地方节点与省中心节点之间，可以方便地对远程 Oracle 数据库进行快速访问和查询。

## 6 结语

将分布式数据库技术引入云南地震信息管理系统后，我们结合先进的网络技术和充分论证的数据库复制方案，有效地解决了数据分散和集中管理的矛盾，减轻了省中心节点数据管理和维护的压力，实现了数据的共享和交换，并建立了数据异地备份机制和双路网络传输链路，保障了数据的安全性和可靠性。实践证明，分布式技术在远程数据管理中具有不可替代的作用。

### 参考文献：

Wu J (美). 2000. 分布式系统设计 [M]. 北京：机械工业出版社.  
周龙骧. 1996. 分布式多媒体数据库系统的分层体系结构 [J]. 计算机学报  
长城工作室数据组. 2001. SQL Server 2000 高级应用 [M]. 北京：人民邮电出版社.  
路川，胡欣杰. 2006. Oracle 10g 宝典 [M]. 北京：电子工业出版社.  
王海亮，张立民，王海凤，等. 2004. 精通 Oracle 10g PL/SQL 编程 [M]. 北京：中国水利水电出版社.  
牛从达，王峰，李倩，等. 2005. 云南分布式地震观测数据库系统设计与实现 [J]. 地震研究，28 (4)：398—402  
张云峰，申文庄. 2000. 对中国数字地震观测系统建设中有关问题的思考 [J]. 地震地磁观测与研究，21 (3)：56—59.

Application of Distributed Database Technology of Oracle10g for  
Database System of Yunnan Earthquake Observation

WU Kong-chun, NIU Cong-da, NI Tai-shan

- (1. Earthquake Administration of Yunnan Province, Kunming 650224, China)
- (2. Earthquake Administration of Honghe State, Mengzi 661100, Yunnan, China)

Abstract

Based on the primarily principle of distributed database technology of Oracle10g and sufficient reasoning on Earthquake information service and database system of Yunnan earthquake observation, and according to technology regulation of Earthquake information service system of China, we set forth process and implementation of building on Database system of Yunnan earthquake observation for Oracle10g.

Key words: distributed database system; communion of earthquake data; replication; distributed query