

基于网格环境下的 野外台站观测数据平台

张耀南^{1,2} 张宝山^{1,2} 陆好^{1,2} 康建芳^{1,2} 赵雪茹^{1,2}

1. 中国科学院寒区旱区环境与工程研究所冻土工程国家重点实验室, 兰州 730000

2. 甘肃省高性能网格计算中心, 兰州 730000

摘 要: 数据主要来源于地理上分布的野外台站、空间观测、互联网服务机构的资源环境数据平台, 是地学研究数据与模型模拟的基础, 其数据采集、管理、融合与共享服务是野外台站观测数据平台的主要任务。基于数据来源分散和服务对象的分布特征, 本文给出了基于网格环境中地理上分布的野外台站资源环境数据平台建设框架和数据汇总、管理、融合与服务的流程, 定义和分析了各模块的关键功能与实现技术, 并在Linux环境下, 模拟野外台站环境, 对数据平台中数据复制、数据同步、统一服务的关键技术和思路进行了功能性实验, 初步建立了网格环境下野外台站的资源环境数据平台。

关键词: 资源环境数据平台; 野外观测站; Linux操作系统; Grid环境

Data Platform of Field Observation Stations Base on Grid Environment

Zhang Yaonan^{1,2}, Zhang Baoshan^{1,2}, Lu Yu^{1,2}, Kang Jianfang^{1,2}, Zhao Xueru^{1,2}

1. State Key Laboratory of Cold Soil Engineering, Cold and Arid Regions Environment and Engineering Research Institute, Chinese Academy of Sciences, Lanzhou 730000, China

2. Gansu High Performance & Grid Computing Center, Lanzhou, Gansu 730000, China

Abstract:

The Data Platform of Resource and Environmental Science is a foundation of data and model simulation for geography research. The data mainly comes from several field observation stations in distributed locations, spatial observation and some data service sites on the Internet. The data collection, management, assimilation and the sharing service are the main tasks of the data platform. According to the distributed characteristics of the data sources, the service objects and new technology of the grid environment, in this paper we discussed the foundation of the data platform of the field observation stations' Resource and Environmental Science, and the process frame of the data together with the management, and analyzed the key functions and implementing techniques for each module. Based on simulated the environment of the field observation stations in Linux operation system, we have conducted some functionality tests of the key techniques and pathways for data replication, data synchronization and a uniformed data service on the data platform, and used acquired techniques to build a primary Data Platform of the field observation stations' Resource and Environmental Science under the grid environment.

Keywords:

Data platform of resource and environmental science; Field observation stations; Linux operation system; Grid environment



1.引言

资源和环境的研究中,已布置了许多野外台站,积累了大量的观测数据,但这些数据分散在不同的野外台站中,数据的管理和使用变得非常复杂。因此,人们希望有一种平台,能整合各个野外台站的数据,可以方便地查询、获取、共享并有效应用这些数据,以减少不必要的数据访问困难,从而提高数据的利用效率^[1]。

基于野外台站观测位置的地理分布性,观测数据的实时性、多样性、长期性和数据量大,以及各野外台站的数据管理系统的差异性等特点,要构建的基于网络环境下的野外台站观测数据平

台(以下简称数据平台)必须具有实用性、稳定性、可靠性、可扩展性和可移植性和易整合性,具备野外台站观测要素的数据规范、数据属性管理的元数据标准、所有地理信息的集成方法,以及对这些信息的获取、处理、存储、分发等所需要的技术和标准。针对目前各野外台站已建立的数据管理系统,基于网络环境,采用网格技术和元数据技术是解决这一空间分布数据集成问题的理想的方法。

网格技术,是将分布在不同地理位置的计算机资源,通过高速互联网组成充分共享的资源集成,从而提供高性能的数据管

理和信息服务的能力。采用网格技术来构建资源环境数据平台,可以有效地将不同野外台站地观测数据加以储存、管理并提供高效服务,降低观测数据的获得成本,提高监测数据的利用率。

元数据一般被认为是对数据的描述,是关于数据集的数据。在资源环境数据平台的解决方案中,元数据已经从一种数据描述与索引的方法,扩展到包括数据发现、数据转换、数据管理和数据使用的整个信息管理过程中,形成不可或缺的工具和方法之一,是资源环境数据平台的核心内容之一。

采用Linux 操作系统、基于

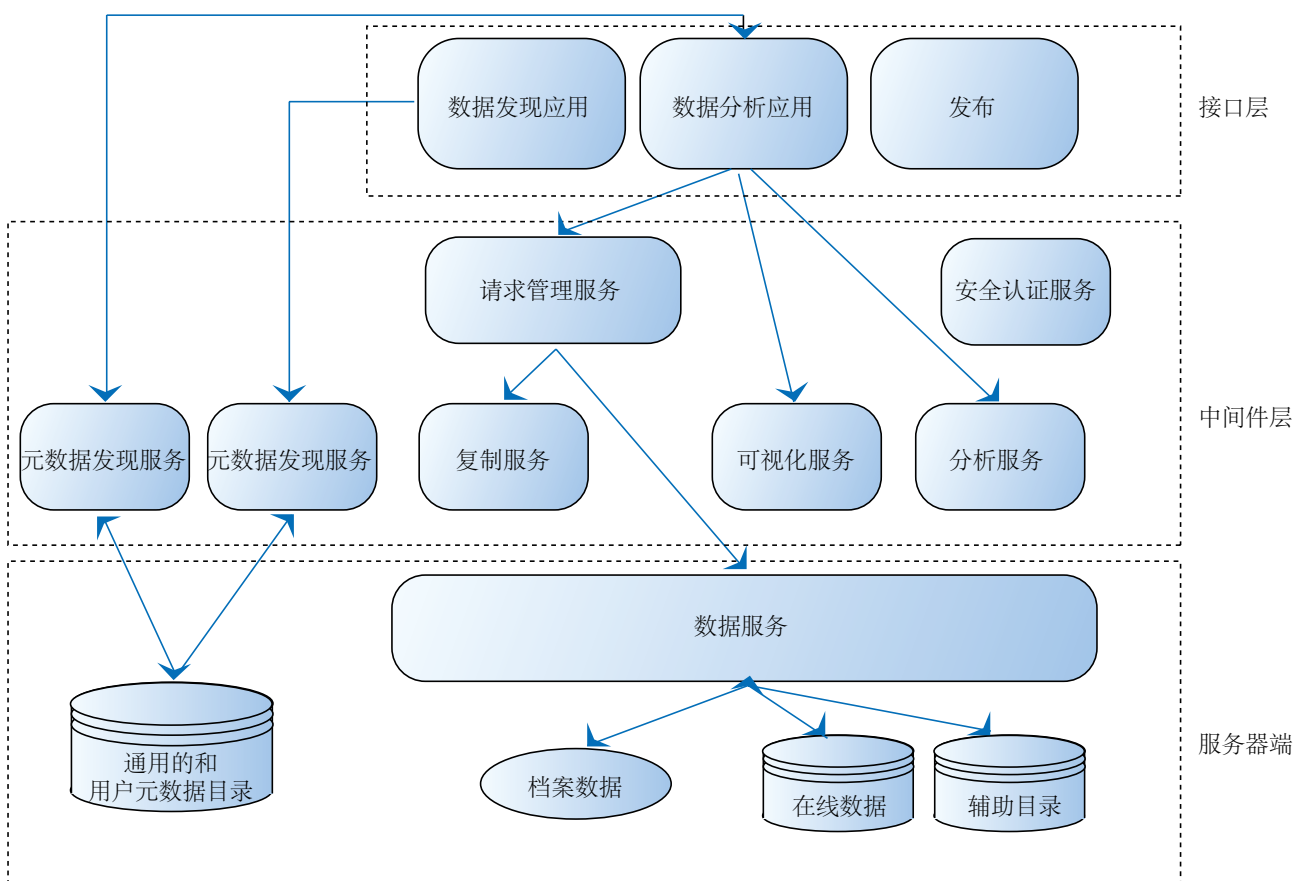


图1 数据平台的体系结构示意图

Linux 环境的网络技术、Oracle 技术和元数据管理技术, 本文讨论了基于网格环境下的地理上分布的野外台站资源环境数据平台建设和数据汇总与管理的流程框架, 定义和分析了各模块的关键功能与实现技术。并在Linux环境下, 模拟野外台站环境, 对数据平台中数据复制、数据同步、统一服务的关键技术和思路进行了功能性验证, 初步获得了在网格环境下构建野外台站资源环境数据平台建设的关键技术。

2. 数据平台结构分析

2.1 数据平台的体系结构

在设计这套数据平台系统时, 参考了Globus的五层沙漏模型和Web Service的模型结构, 按照数据网格的体系结构, 将这套资源环境数据平台分为如图1三部分:

服务器端、中间层、接口层。

图1所示的数据平台的体系结构中, 其中服务器端主要提供数据服务, 包括档案数据、本地数据、在线网上数据、数据的辅助目录、通用元数据有关的服务。中间件层重点提供元数据的发现服务、数据复制服务、元数据复制服务、请求管理服务、安全认证服务、可视化服务、数据分析服务。接口层主要提供数据发布、数据应用分析、数据发现和数据应用服务。

据视图, 并通过一致的方式来访问数据。

如图2发现服务的过程所示, 通过网格存储的API和数据平台的网格接口, 可以通过检索来找到存储于野外台站的文件实例, 并通过网格存储的API来对这些文件实例进行远程的读写以及查看属性的操作, 并通过在文件实例存储服务建立第三方的传输控制来在不同的存储系统之间进行文件实例的传输。

2.3 数据复制服务

数据平台中存储的数据是分布式的, 有时需要的数据分布于不同的野外台站节点上, 如果每次都要依次从每个节点进行数据的集合, 就使得系统的开销加大, 不利于数据的使用。在平台中, 通过利用网格服务中的数据复制服务功能来将符合一定条件

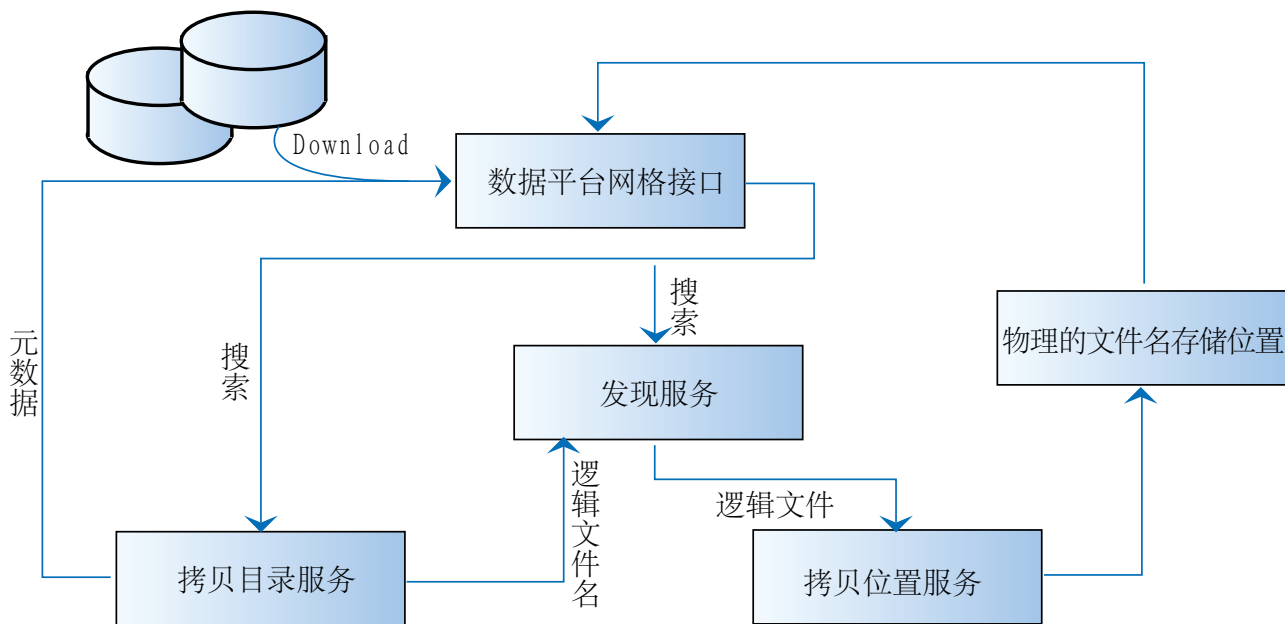


图2 发现服务的过程

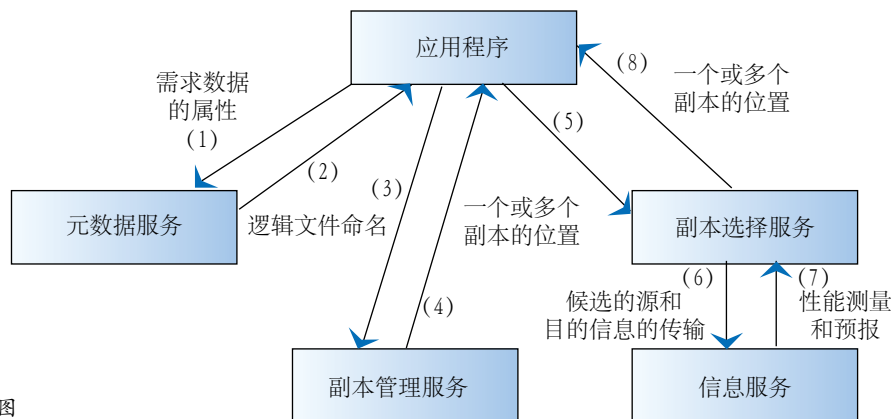


图3 数据复制过程示意图

的数据进行集合，减少了数据存取的时间、提高访问效率。在数据复制的粒度上，将文件作为最小复制粒度。并将相关的元数据和所有相关文件的拷贝均归为元数据配置库中的一个实体，与文件实例一起表现成为一个称为“逻辑文件”的逻辑结构。在数据复制服务中，将相关的逻辑文件聚合在一起，称为拷贝目录。

如图3数据复制过程的示意图可见，数据复制服务从与元数据的集合到服务管理的定位，再到复制的选择，最后到信息服务，中间需要经过图3上的8个步骤，分别是：

(1) 应用程序将所需求数据的属性传送到元数据服务中进行检索；

(2) 元数据服务将符合关

键字的逻辑文件命名传送给应用程序；

(3) 应用程序再将得到的逻辑文件命名传送给副本管理服务；

(4) 副本管理服务将该文件的一个或多个副本的位置再送回应用程序；

(5) 应用程序再将这些副本的位置发送到副本选择服务中；

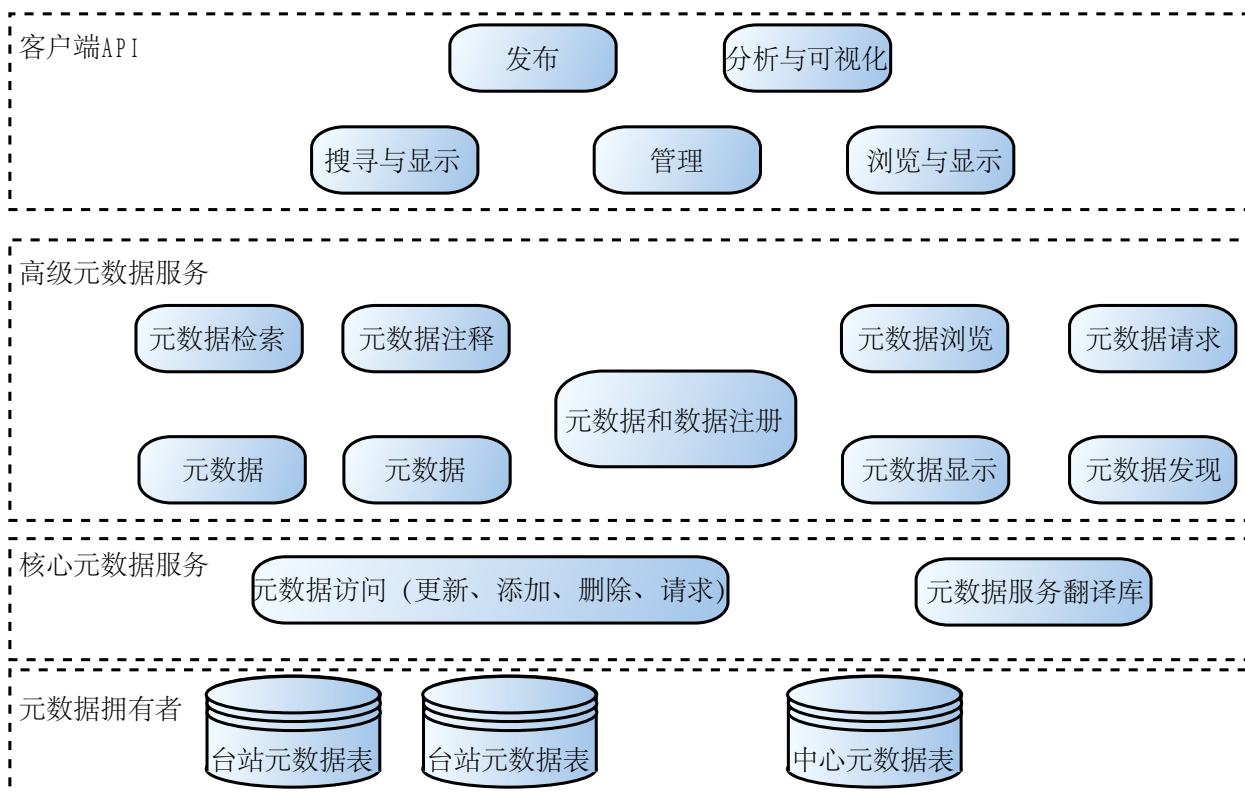


图4 元数据管理与服务模型

(6) 副本选择服务将候选的源以及目的的信息传送给信息服务;

(7) 并由信息服务将获取这些副本所需的性能进行测量和预报并返回给副本选择服务;

(8) 最后副本选择服务将可用的一个或多个副本的位置信息传送给应用服务。

2.4 元数据的管理模型

在资源环境数据平台中,使用了一种索引结构来构造元数据的体系。利用这种索引结构,可以解决网格环境下进行元数据的操作,将可能出现的冲突清空,使得元数据可以保持一致。图4即为建立的元数据管理与服务的模型示意图。

如图4所示,元数据服务具有

一定的层次结构,通过数据组织利用这种层次结构来解决网格中分布式索引高效检索的复杂性。构建一个支持网格中分布式环境层次的命名结构和数据模型,通过这种模型结构和与LDAP相关的目录层次结构,来组织、复制和分布目录信息,不用考虑数据是如何存放的,以及是存放在那里的,形成支持多种方式为用户提供元数据服务和保持元数据一致的机制。

2.5 整体分析

数据平台系统主要分为数据中心部分和数据节点部分,数据中心部分主要作用是协调各个数据节点部分之间的信息和用户权限,并将各个数据节点的数据进行收集,同时对外提供信息服

务;数据节点部分主要负责该节点数据的整理和用户访问该类数据的权限指定工作,同时也对外提供信息服务。它们之间的关系如图5所示。

如图5所示数据中心和数据节点之间几个重要的关系:

(1) 数据中心通过在数据节点上的各种监控服务来完成对数据节点的监控工作,其中包括网络流量的监控、用户访问数据、各项服务的运行情况等。

(2) 数据节点通过信息同步功能将数据传送到数据中心。

(3) 数据节点通过数据中心来完成服务的发现,从而使得客户端可以了解数据信息服务是否存在以及如何与之交互。

(4) 数据中心可以对数据节点的设置信息加以调整。

2.6 模块功能说明

数据平台内部的模块主要是由用户管理系统、信息服务系统、元数据管理系统和数据管理系统所组成,其功能描述如下:

用户管理系统,主要负责对平台系统中的用户信息、用户权限以及用户信息同步进行管理。

信息服务系统,主要负责向外部提供信息服务,这其中包括了元数据信息服务和数据信息服务并涉及到了数据与元数据的发现和分析服务。

元数据管理系统,主要是对于数据平台中用到的元数据

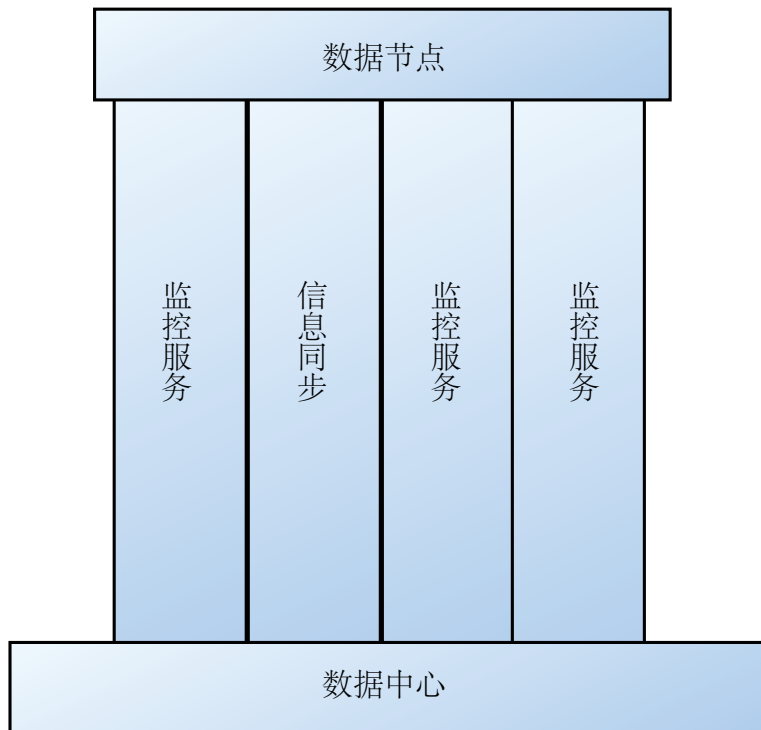


图5 数据中心和数据节点之间的关系

► 的管理、更新、同步、注册和备份工作。

数据管理系统，主要是负责对于数据库中存储的数据进行有效的管理、同步和备份工作。

3. 系统实现

3.1 系统运行环境

操作系统采用稳定、性能突出、安全性好的CentOs^[2]。

Web服务器是包含Servlet API 2.2和JSP 1.1兼容的Servlet/JSP容器，提供了快捷高效的jsp/servlets运行平台的Tomcat。

底层网络选择支持开放源码的网络基础平台Globus Toolkit^[3]。Globus Toolkit是一个基于开放结构、开放服务资源的软件库，并支持网格和网格应用，目的是为构建网格应用提供中间件服务和程序库。

数据库选择加入了对网格的支持的Oracle10g^[4]。

网络存储API使用GridFTP来

扩充API接口，来保证第三方数据传输中的数据安全和并行数据传输。

数据复制工具选择支持Subscription 的GDMP^[5]。

3.2 实现的关键技术

3.2.1 数据同步

目前存储于数据节点的数据有关系型和非关系型两种。对于非关系数据文件，可以采用Linux系统中的rsync^[6] +ssh shell^[7]相结合的方法来实现这个功能。使用rsync通过SSH Shell的安全认证机制，可以将数据节点中的非关系型数据文件全部镜像增量传输到数据中心来。在实际操作中，配置好rsync和ssh服务后，可以通过在Linux系统中的crontab来定时自动完成数据的同步过程而不用人工参与。

对于关系型数据，可以通过使用Oracle 10g中自带的Stream技术来完成数据节点与数据中心之间的关系型数据同步工作。

Oracle Stream利用高级队列技术，通过解析归档日志，将归档日志解析成DDL及DML语句，从而实现数据库之间的同步。这种技术可以将整个数据库、数据库中的对象复制到另一数据库中，通过使用Stream的技术，对归档日志的挖掘，可以在对主系统没有任何压力的情况下，实现对数据库对象级甚至整个数据库的同步。

3.2.2 数据复制服务

数据复制服务的主要目的是通过数据复制来降低系统的访问时间以及节约网络带宽等。复制技术在运行时刻根据系统的应用特征以及用户的访问特征自动将数据项复制到不同的平台节点，其实质就是通过数据在系统的流动，实现系统的整体性能优化。相对于通常的数据拷贝，数据复制主要在技术上要完成数据副本创建、数据传输、副本删除、副本选择、数据一致性管理、安全管理这些工作。

表1 实验环境

	Center	Node-A	Node-B
操作系统	CentOS 5.0	CentOS 5.0	CentOS 5.0
网格支持	Globus Toolkit 4.0.3	Globus Toolkit 4.0.3	Globus Toolkit 4.0.3
数据库	Oracle10g	Oracle10g	Oracle10g
GDMP	3.2.6	3.2.6	3.2.6
IP	210.77.67.xx	210.77.68.xx	210.77.69.xx

在数据复制中, 选择GDMP (Grid Data Mirroring Package) 工具, 通过GDMP的数据复制过程: `gdmp_register_local_file`、`gdmp_publish_catalogue` 和 `gdmp_replicate_get` 来进行数据的复制过程; 然后通过HTTP Redirection来进行数据副本的目录管理; 通过文件替换、增量更新、日志管理来进行数据副本的同步工作; 最后在副本的选择方面, 所需要创建的成本模型必须要考虑网络负载和数据所在服务器的负载方面。

3.2.3 数据入库的过程实现

从野外台站的观测仪器中如何快速有效的将所得数据导入到数据库中, 成为本套系统数据来源中一个重要的问题。在这套方案中, 考虑使用JAVA来构建一套导入工具, 用以实现从原始数据文件导入到Oracle数据库这一过程。这一过程是: 从仪器上得到的符合规范的数据, 将其保存成指定格式的数据文件, 然后利用导入工具连接Oracle数据库, 将其导入指定的表中, 从而完成数据的导入过程。

4. 实 验

4.1 实验环境

为了模拟野外台站的环境, 选用了3台机器做为实验环境来进行准备, 其中一台机器被模拟用于

数据中心, 称之为Center, 另外两台机器被用于模拟野外台站的数据节点, 称之为Node-A和Node-B。它们均放置于Internet环境中, 具有独立的合法IP地址。表1展示了它们的软件环境配置。

4.2 实验测试

4.2.1 数据复制

这里以对地下水水位数据的时间序列拷贝过程为例, 对数据复制做了功能性的验证。Node-A中存放了从2000年1月到2000年5月的甘肃兰州的水位的数据文件, Node-B中存放了从2000年6月到2000年12月的甘肃榆中的水位的数据文件, 其中兰州和榆中可以认为在大尺度上属于同一地点, 要进行的实验过程是将这2个部分的数据进行复制, 形成完整年份的甘肃兰州地区的水位的时间序列数据文件存储于Node-B中安装并配置GDMP 3.2.6版本后, 在Node-A上执行`gdmp_register_local_file -d /var/filedb`进行对目录文件的注册, 然后在Node-A上执行`gdmp_publish_catalogue`来进行目录的发布工作, 最后在需要数据复制的机器Node-B上运行`gdmp_replicate_get`来进行文件从Center复制到Node-B的过程。

4.2.2 数据同步

在这个功能性实验中, 将分别对储存于数据库中的关系型数

据和储存于介质中的非关系型数据的同步工作加以实验验证。

针对关系型数据的同步工作进行验证。在这个方面的实验中, 实验目的是使得Node-B的Oracle数据库中的关系型数据, 可以传输到Center的Oracle中。这里使用的是Oracle 10g的Stream功能。

首先对需要同步的两端数据库做参数调整、设置归档模式、设置同步管理员等操作。然后, 在需要同步的Node端创建Master队列, 并建立捕捉和传播进程。

首先创建Master队列(程序略去), 然后创建捕捉进程(程序略去), 再创建传播进程(程序略去), 最后在需要接收同步数据的Center端建立应用进程(程序略去)。

通过以上配置, 在启动捕捉和应用进程后, 在Node端进行数据库中数据的新增、修改和删除工作, 在指定的时间后, 在Center端检查, 发现已经得到了Node端的数据, 故此功能可以通过这种方法来实现。

针对非关系型数据的同步工作进行验证。对于非关系型数据, 既可以采用数据复制的办法来做, 也可以采用别的方式来做。本文采用的是使用rsync+ssh的方式来完成非关系型数据的同步工作。实验的目的是使Node-A的非关系型数据每天定时自动同步增量备份到Center上。



► 第一步，在Center的磁盘中，设定一个目录做为信息服务所指定的非关系型数据存储目录filedb，其位置为/var/filedb，并根据相应的节点创建子目录为/var/filedb/nodea。同样也在Node-A的磁盘中，也为非关系型数据设定存储目录，其位置为/var/filedb。第二步，在Node-A中配置rsync服务和ssh服务，并可以成功通过Center进行连接。第三步，在Center上建立一个shell脚本文件，并将文件放入crontab中进行定时自动运行。此脚本文件内容如下：

```
#!/bin/sh
BDIR=/var/filedb/nodea
EXCLUDES=/var/excludes
BSERVER=nodea
export RSYNC_PASSWORD=XXXXXX
BACKUPDIR=`date +%A`
OPTS="--force --ignore-errors --delete-excluded --exclude-
from=$EXCLUDES --delete --backup --backup-dir=$BACKUPDIR -a"
export PATH=$PATH:/bin:/usr/bin:/usr/local/bin
[ -d /var/emptydir ] || mkdir /var/emptydir
rsync --delete -a /var/emptydir/ $BSERVER::$USER/$BACKUPDIR/
rmdir /var/emptydir
rsync $OPTS $BDIR $BSERVER::$USER/current
```

4.2.3 统一服务

为了使整个数据平台能统一的对外服务，在各节点和中心之间，通过使用同步技术，使得平台中存储的观测数据和用户的信息、权限，都可以在中心和节点之间互相传输，这为资源的整合和数据的安全性和不间断服务提供了底层支持。

在实验中，通过在数据库中构建视图的方式，来完成相关数据的数据表组合，比如元数据视图通过以下代码构建：

```
CREATE OR REPLACE VIEW DATAPLATFORMDB.V_ME_DSET_
METAINFO_ALL (DSET_META_ID, METAATTRIBUTE_CODE, DSET_CODE, META_
VALUE, TABLENAME)
AS
SELECT "DSET_META_ID", "METAATTRIBUTE_CODE", "DSET_
CODE", "META_VALUE", "TABLENAME"
from DATAPLATFORMDB.V_ME_DSET_METAINFO_ALL1
union
```

```
SELECT "DSET_META_
ID", "METAATTRIBUTE_
CODE", "DSET_CODE", "META_
VALUE", "TABLENAME"
from DP_TGS_DB.V_ME_DSET_
METAINFO_ALL
```

5. 总结

利用本文讨论的网格环境下的资源环境数据平台的总体思路 and 关键技术，正在构建针对地理上分布的野外台站数据节点中数据的采集、管理与共享系统。初步建立的数据平台证明数据平台的整体结构和各主要功能模块的设计是合理的，获得一些主要实现技术是可行的。由于采用了网格技术、Java技术以及Linux系统，使得整个系统结构清晰，代码复用率提高。后期要在初步建立的实验平台上，研究数据复制、数据拷贝、元数据管理以及数据应用的关键技术，完成针对野外台站的数据采集、管理和集成工作。虽然本文针对野外台站的数据集成与服务，但对类似情况下的数据管理问题也具有一定的参考价值。



参考文献:



- [1] 张耀南, 程国栋, 韦五周, 等. 基于曙光3000计算环境的寒区旱区资源环境数据平台建设[J]. 冰川冻土, 2004, 26(2): 224-229.
- [2] CentOS. <http://www.centos.org/>.
- [3] Globus. <http://www.globus.org/toolkit/>.
- [4] Oracle. <http://www.oracle.com/>.
- [5] GDMP. <http://project-gdmp.web.cern.ch/project-gdmp/>.
- [6] rsync. <http://samba.anu.edu.au/rsync/>.
- [7] OpenSSH. <http://www.openssh.com/>.

收稿时间: 2008年12月27日

作者信息



张耀南

中国科学院寒区旱区环境与工程研究所, 研究员, 博士, 主要研究方向为基于数据、模型、计算的e-Science研究。



张宝山

中国科学院寒区旱区环境与工程研究所, 助理工程师, 主要研究方向为数据库与应用研究。