

气象资料省际共享监控系统^①

王建荣, 唐怀瓿, 孔芹芹, 金素文

(安徽省气象信息中心, 合肥 230031)

摘要: 气象资料省际共享为预报预测以及气象服务提供了及时的周边省资料服务. 目前国家气象信息中心按月统计省际共享资料传输质量, 不能及时发现资料收发异常. 参照上行传输监控方案, 设计与实现了省际共享实时监控系系统, 将接收与发送的自动站文件解码得到的站级信息以及雷达基数据站级信息写入省际共享监控库. 通过 Web 展现国家级自动站正点资料、区域自动站正点资料和雷达基数据接收与发送情况. 基于数据库实现各省资料日、月和年统计分析, 为资料应用部门和管理部门提供实时的省际资料传输情况. 系统自动生成省际共享日志并定时上报. 业务应用证明, 该系统运行稳定可靠, 体现了气象资料省际共享实时业务质量监控的优势.

关键词: 省际共享; 三层架构; CTS; 多线程并行; 分区表

引用格式: 王建荣, 唐怀瓿, 孔芹芹, 金素文. 气象资料省际共享监控系统. 计算机系统应用, 2017, 26(12): 71-77. <http://www.c-s-a.org.cn/1003-3254/6085.html>

Monitoring System for Meteorological Data Interprovincial Sharing Service

WANG Jian-Rong, TANG Huai-Ou, KONG Qin-Qin, JIN Su-Wen

(Anhui Meteorological Information Centre, Hefei 230031, China)

Abstract: Meteorological data interprovincial sharing service provides the weather prediction and meteorological service of timely data for neighboring provinces. The National Meteorological Information Center now gets statistics and publishes the transmission quality of interprovincial sharing service by month. It is not easy to find out the abnormality during data transmission timely. We design and implement a real-time monitoring system for interprovincial data sharing service using the monitoring method for data upload transmission. Firstly, the system above decodes AWS-data and then shows the receiving and sending status of the hourly national and regional automatic station data and radar data. This system provides data application department and management department of the daily, monthly and annual statistics results based on SQL Server database and it also generates interprovincial sharing log files. The application result shows that the real-time monitoring system of meteorological data interprovincial sharing service is stable and reliable.

Key words: interprovincial sharing; three-tier architecture; CTS; multi-threaded parallel; partition table

引言

数据共享业务在短时临近预报、突发灾害性天气预报预警和决策气象服务工作中的作用十分显著. 之前, 各省只能通过中国气象局卫星广播系统 (CMACast^[1]) 下行通道获取邻省资料, 从 CMACast 系统下发的全国站点打包文件中逐个查找所需的站点, 由于全国站点

数目大, 下发报文数量多, 导致每时次数据处理耗时长, 不能满足预报预测以及气象服务对资料及时性的需求^[2]. 2012年5月中国气象局组织实施了全国省际气象数据共享工作, 在全国范围开展雷达基数据、国家级和区域气象自动站观测数据的省际数据共享传输业务. 各省陆续开通了新一代国内通信系统共享服务, 实

^① 基金项目: 中国气象局关键技术集成项目 (CMAGJ2015M29); 安徽省气象局科技发展基金项目 (KM201306)

收稿时间: 2017-03-09; 修改时间: 2017-03-27; 采用时间: 2017-03-29

现雷达基数据和自动站资料在省际间快速共享和传输,极大地提高了各省市(区)获取周边省气象资料的时效。

国家气象信息中心负责按照共享气象资料传输时限要求按月统计省际共享气象资料传输质量,通过国家级气象业务内网发布统计结果,各省信息中心每天将前一天的站级记录日志文件通过省级CTS系统上传至国家气象信息中心^[3]。目前,国内对气象资料省际共享实时监控方法的研究尚待完善。

本系统按照省际共享气象资料传输时效要求,对省际共享做实时监控,及时发现收发双向中断问题,保障气象资料省际共享业务质量,并向资料用户和管理部门提供实时的省际传输状况。

1 系统架构

按照中国气象局规定,安徽省与江苏、浙江、上海、江西、湖北、湖南、河南和山东8个省进行气象资料交换共享。中国气象局建设了CMANet(全国气象宽带网络)和新一代国内通信系统以及后续的CTS,省际之间以CMANet为支撑,共享资料,如图1所示。

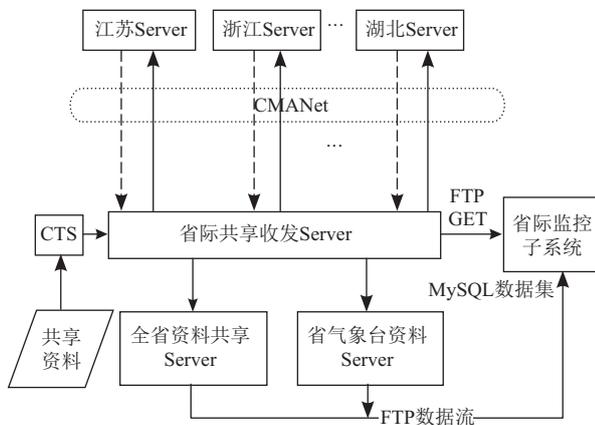


图1 系统架构

CTS是全国综合气象信息共享平台(CIMISS)的数据收集与分发子系统,是新一代国内通信系统的升级^[4,5]。目前,安徽省新一代系统主要运行CMACast下行业务,省际共享收发Server作为新一代系统的单机备份,专门承载省际共享业务。

省际共享的资料包括国家级自动站资料、区域自动站资料和多普勒雷达基数据,本文将国家级自动站、区域自动站统称为自动站资料。

安徽省气象局省际共享业务流程为:资料通过CTS分发至省际共享收发Server,省际共享收发Server将收集的资料放进各省传输队列,分发进程采用并行方式将各队列资料推送至各省服务器并将发送日志入库。省际共享收发Server接收外省的资料后也通过队列方式推送至全省资料共享Server(主机别名:SHARES)省气象台资料Server(主机别名:SQXT)两个本地存储服务器并将接收日志入库。

省际监控子系统通过查询省际共享收发Server的MySQL数据库xxshj_db实时接收表(TR_FILE_RCV)和实时发送表(TR_FILE_SEND),逐行读取记录集,根据文件名字段调用.NET FTP API函数访问资料存储路径中的自动站打包文件,逐个文件解码得到区站号和观测时间等信息后批量插入省际监控子系统数据库。

2 软件系统设计与实现

2.1 功能模块

省际监控子系统包括接收信息入库模块、发送信息入库模块、省际共享日志文件生成模块、收发信息补入库模块、监报告警模块和统计分析模块,如图2。根据中国局要求,每日生成省际共享日志文件并上传,省际共享日志生成模块基于省际监控子系统数据库(数据库ID: sjgx, 基于SQL Server2008 R2),定时生成日志文件。

2.1.1 接收信息入库

省际共享收发Server接收外省的自动站资料后解码入库得到站级信息(包括区站号Iiii和观测时间),而雷达基数据是命名规范的二进制文件,不作解码,解析其文件名即可得到雷达站区站号和观测时间。由资料类型、区站号、观测时间、接收时间等即可批量插入sjgx库自动站接收信息表(dbo.RcvRealInfo)和雷达接收信息表(dbo.RadRcvRealInfo)。

2.1.2 发送信息入库

发送信息包括两个类型目的地:

- (1) 8个外省;
- (2) 两个本地资料服务器SHARES及SQXT。

对于(1),本监控系统需解码省际共享收发Server上本省发送外省成功后转移至备份目录中的自动站资料并入库;对于(2),由于“接收信息入库”模块是解码SHARES和SQXT资料存储服务器中的文件,接收信

息入库的同时, 也根据发送信息表结构生成 DataTable 对象批量写入发送信息表 (dbo.SndRealInfo 和 dbo.RadSndRealInfo)。

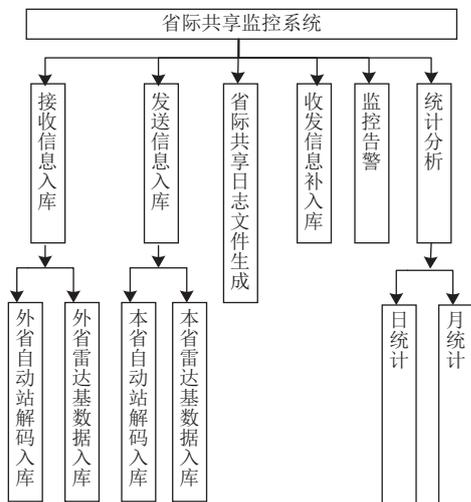


图2 监控系统功能模块

2.1.3 省际共享日志文件生成

按照《省际共享气象资料日志文件传输规范》要求, 每天由 sjgx 数据库的作业调用存储过程 dbo.Create DataShareProvinceDayLogFile 定时生成前一天的站级记录日志文件并压缩成一个文件后通过 CTS 系统上传至国家气象信息中心。

文件名模板为: Z_NOTES_C_CCCC_YYYYMMDD HHmmss_L_SHARE-sType.TXT, 其中 sType 规定见表 1。

表 1 日志文件

资料类型	sType
国家级自动站	NATAWS
区域自动站	REGAWS
雷达基数据	RADA

2.1.4 监控告警

邻省自动站正点资料、雷达基数据收发时效监测。通过检测策略指定第几分钟采集 sjgx 数据库接收信息表和发送信息表, 某省缺报站点总数超过阈值 (或缺报率超过阈值, 因为接收外省的考核站点每年均有不同数量的调整, 因此各省缺报阈值随之调整) 则发出声音提示。

2.1.5 统计分析

实现各省共享资料的接收和发送日统计、月统计和年统计及报表功能, 统计对象为各省自动站逐小时

资料和雷达基数据及时率。

2.1.6 收发信息补入库

系统解码入库程序执行异常中断时, 需要将中断时段内的收发信息补入库, 补入库流程与正常流程一致, 只是从异常起始时间执行至恢复时间。

2.2 软件设计

系统采用三层架构^[6]设计和实现, 分为表示层、业务逻辑层、资源访问层, 如图 3 所示。各层逻辑分离, 降低了耦合度, 增加了系统的可扩展性。

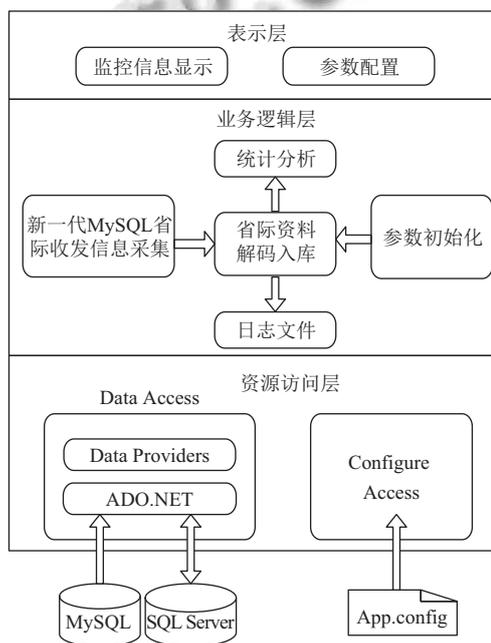


图3 软件架构

2.2.1 表示层

监控信息展示和参数配置 UI。

监控信息显示: 资料接收和发送信息实时显示 (各省每小时应收、实收、接收及时率)、传输统计 (按资料类型、省份、统计时间) 显示、日数据统计分析、月数据统计分析、年数据统计分析, 如图 4 所示。

参数配置 UI: 省际告警配置, 即资料类型、启用接收告警、启用发送告警和告警频次设置接口。

2.2.2 业务逻辑层

业务逻辑层的主要功能: 定时从省际共享收发 Server 数据库 xxshj_db 查询资料接收与发送信息, 解码各省自动站文件得到站级记录集后批量写入 sjgx 库、解析雷达基数据文件名得到雷达站级信息并写入 sjgx 库; 基于 sjgx 库定时生成省际共享气象资料日志文

件; 根据检测策略定时采集省际传输异常信息并告警。

资料走向	资料类型	资料子类型	应收发个数	未接收个数	未发送个数	接收及时率	发送及时率
上海→新一代→安徽省台	自动站资料	区域站四六要素	48	11	11	77.08	77.08
山东→新一代→安徽省台	自动站资料	区域站四六要素	1373	128	128	90.68	90.68
浙江→新一代→安徽省台	自动站资料	区域站四六要素	2157	147	147	93.18	93.18
湖北→新一代→安徽省台	自动站资料	区域站四六要素	865	35	35	95.95	95.95
湖北→新一代→安徽省台	自动站资料	国家级自动站	84	3	3	96.43	96.43
江苏→新一代→安徽省台	自动站资料	区域站四六要素	1227	40	40	96.74	96.74
河南→新一代→安徽省台	自动站资料	区域站四六要素	647	14	14	97.84	97.84
安徽→新一代→湖南	自动站资料	区域站四六要素	524	11	524	97.9	0
安徽→新一代→上海	自动站资料	区域站四六要素	524	11	524	97.9	0
安徽→新一代→浙江	自动站资料	区域站四六要素	524	11	524	97.9	0
安徽→新一代→河南	自动站资料	区域站四六要素	524	11	11	97.9	97.9
安徽→新一代→湖北	自动站资料	区域站四六要素	524	11	11	97.9	97.9
安徽→新一代→江苏	自动站资料	区域站四六要素	524	11	11	97.9	97.9
安徽→新一代→江西	自动站资料	区域站四六要素	524	11	11	97.9	97.9
安徽→新一代→山东	自动站资料	区域站四六要素	524	11	11	97.9	97.9

图4 省际共享实时监测

本文主要介绍外省资料解码入库过程. 应用多线程技术和.NET Framework FTP 类库 API^[7]. 针对 8 个外省, 在 SHARES 服务器上按省名创建 8 个资料存储目录, 每个省名目录分别包含国家自动站目录、区域自动站目录和雷达基数据目录. 除雷达基数据外, 各类自动站报文均需解码.

系统解码入库程序共创建 8 个工作线程并行执行, 并且使用 C# ADO.NET 数据库连接池技术^[8], 各线程执行时访问不同的目录、打开连接池中不同的连接资源, 因此基本不存在线程同步问题.

解码入库进程通过.NET 定时器每 2 分钟执行数据访问步骤:

(1) 查询省际共享服务器的 MySQL 文件级实时接收表 (TR_FILE_RCV) 和文件级实时发送表 (TR_FILE_SEND), 并返回记录集.

(2) 主线程暂停 N 秒钟 (N=1, 2, ...), 等待记录集各数据行 (DataRow) 的文件名字段对应的文件完整写入 SHARES.

(3) 逐行读取记录集, 获取资料类型 (DTYPE)、资料子类型 (CTYPE)、省份代码 (CCCC)、文件名 (FNAME)、接收时间 (ARR_TIME) 和入库时间 (LOG_TIME) 等字段值. C#项目的 App.config 文件配置了 SHARES 等资料存储服务器的 FTP 登录账户、省际资料主目录以及各省 CCCC 对应的资料相对路径. 应用.NET FTP API 访问 SHARES 目录下匹配的文件 FNAME, 从字节流中解码得到站级信息, 包括区站号、观测时间等, 将其存入表变量 (DataTable 对象), 记录遍历结束后, 将表变量作为参数传入存储过程批量插入 sjgx 库的 dbo.RcvReallInfo.

在第 (3) 步, 各省自动站资料均是打包后的 Z 文件, 其中包含多站记录, 因此解码子程序采用正则表达

式 (Regex) 来匹配测站基本信息行, 国家级自动站报文、区域自动站报文的测站基本信息行格式不一样, 分别为:

- i. 国家级自动站 Regex: `\b[A-Z0-9]\d{4}\s\d{14}\s\d{6}\s\d{7}\s\d{5}\b.*\b;`
- ii. 区域自动站 Regex: `\b[A-Z0-9]\d{4}\s\d{6}\s\d{7}\s\d{5}\b.*\b.`

设置不同的高效正则表达式来匹配报头, 保证了解码效率.

外省自动站资料解码入库流程图如图 5 所示.

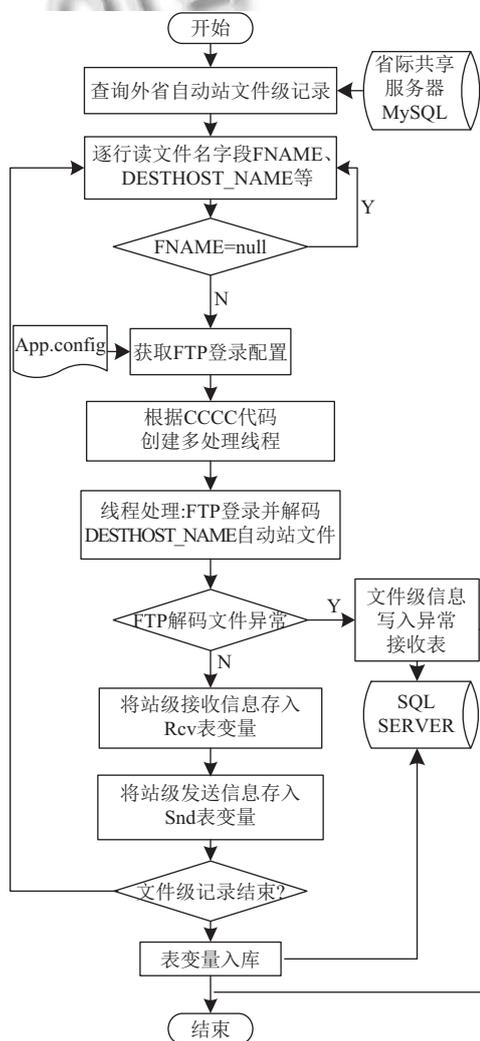


图5 外省自动站解码与入库流程

本省发送至外省的自动站资料解码入库流程基本与图 5 相同, 主要区别是资料所在服务器为省际共享收发 Server, 其 FTP 连接及存储目录结构同样在 App.config 文件中配置.

外省雷达基数据接收无解码步骤,如图6所示。

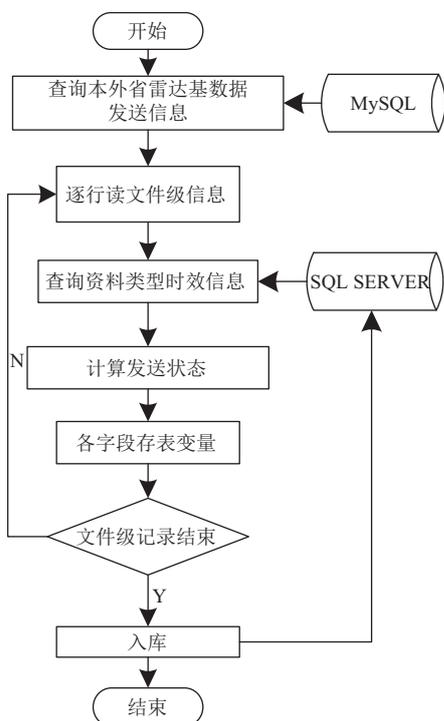


图6 外省雷达基数据入库流程

本省雷达基数据发送入库流程与图6相同。

(1) 入库子程序中表变量的定义。

```
public static
DataTable GetRcvInfoTable()
{
    DataTable dt = new DataTable();
    dt.Columns.AddRange(new DataColumn[] {
        new DataColumn("Dtype", typeof(string)),
        new DataColumn("Ctype", typeof(string)),
        new DataColumn("MTime", typeof(string)),
        new DataColumn("Iiiii", typeof(string)),
        new DataColumn("RcvTime", typeof(string)),
        new DataColumn("RcvFname", typeof(string)),
        new DataColumn("RcvStatus", typeof(string)),
        new DataColumn("DealTime", typeof(string)),
        new DataColumn("DealStatus", typeof(string))
    });
    return dt;
}
```

(2) 数据库端设计。

资料入库子程序使用 SQL Server 2008 支持的表值参数 (Table-valued parameters) 特性,作为入库存储过程的参数,接收传入的表变量 (DataTable) 实现批量插入 sjgx 库收发信息表,从而提高了入库效率,存储过程定义如下:

```
CREATE PROCEDURE BulkInsertRcvInfo
(
    @bulkIns
    dbo.tvbBulkInsertRcvInfoForeignProvince
    READONLY
)
AS
SET NOCOUNT ON
BEGIN
    Insert dbo.RcvRealInfo Select
    Dtype, Ctype, MTime, Iiiii, RcvTime, RcvFname,
    RcvStatus, DealTime, DealStatus
    From @bulkIns;
END
```

2.2.3 资源访问层

包括数据访问子层 (Data access)、配置文件访问子层 (Config access) 和服务访问子层 (Service access)。

基于 ADO.NET 数据库访问技术建立适应 MySQL 和 SQL Server 两种数据库的数据访问子层接口。向业务逻辑层提供 MySQL 数据库的收发数据集 (DataSet); 将业务逻辑层获取的站级收发信息通过调用存储过程的方式持久化到 SQL Server 数据库。

3 数据库模型设计

sjgx 库主要数据表有资料类型字典表 (见表2)、资料-省名配置表、外省台站基本信息表 (见表3)、外省自动站接收节目表、自动站接收信息表 (见表4)、自动站发送信息表 (见表5)、雷达基数据接收信息表、雷达基数据发送信息表、日统计表 (见表6)、月统计表等。

主要存储过程: 解码数据批量入库存储过程、统计分析存储过程、日志文件生成存储过程和 Web 显示存储过程。

为优化 sjgx 库表读写性能,提高数据查询速度,利用分区表技术,对接收信息表和发送信息表按观测时间做分区,具体为每月一个分区^[9,10]。主要数据表均创

建了主键,例如 RcvRealInfo 表选取 (MTime, CType, Dtype, Iiiii) 作为复合主键,也提高了数据检索效率。

3.1 资料类型描述表

表2 资料类型字典表 (DataDescDic)

列名	类型	含义
DTYPE	VARCHAR(10)	数据类型
DtypeDesc	VARCHAR(40)	数据类型描述
JSTime	SMALLINT	考核及时时间

3.2 外省台站基本信息表

表3 外省台站基本信息表 (StationRcvDic)

列名	类型	含义
DTYPE	VARCHAR(10)	资料类型
Iiiii	VARCHAR(9)	区站号
Province	CHAR(4)	CCCC编码

3.3 自动站站级接收信息表

表4 自动站站级接收表 (RcvRealInfo)

列名	类型	含义
DTYPE	VARCHAR(10)	资料类型
Iiiii	VARCHAR(9)	区站号
MTime	SAMLLDATETIME	观测时间
RcvTime	DATETIME	资料接收时间
RcvFName	VARCHAR(255)	接收文件名
RcvStatus	VARCHAR	接收状态(0:及时,1:缺报)
DealTime	DATETIME	入库时间

3.4 自动站站级发送信息表

表5 自动站站级发送表 (SndRealInfo)

列名	类型	含义
DTYPE	VARCHAR(10)	资料类型
Iiiii	VARCHAR(9)	区站号
MTime	SAMLLDATETIME	观测时间
DestName	VARCHAR(15)	发送目的地
SndTime	DATETIME	资料发送时间
SndFName	VARCHAR(255)	发送文件名
SndStatus	VARCHAR	发送状态(0:及时,1:缺报)

雷达基数据接收表和自动站站级接收表结构相同,而雷达基数据发送表和自动站站级发送表结构相同,不再赘述。

3.5 自动站日统计表

雷达基数据日统计表与之相同,不再赘述。

4 系统评估

选取国家级自动站资料作为评估对象,具体为

2013年11月份的省际共享监控数据库自动站站级接收表中8个邻省的国家级自动站资料和CMACast下发的全国自动站打包文件所包含的8省的国家站级自动站资料,在站点和观测时间相同条件下,资料接收的时效性、解码入库效率作评估。

表6 自动站日统计表

列名	类型	含义
DTYPE	VARCHAR(10)	资料类型
Iiiii	VARCHAR(9)	区站号
MTime	SAMLLDATETIME	观测时间
Province	CHAR(4)	省局字母代号(CCCC编码)
TotalNum	INT	每日应收报文数
RcvJSNum	INT	每日接收及时数
RcvJSPer	DECIMAL(5,2)	及时率
SndFName	VARCHAR(255)	发送文件名
SndStatus	VARCHAR	发送状态(0:及时,1:缺报)

4.1 省际共享与CMACast时效对比

省际共享与CMACast下行两种途径的资料接收时效对比,接收时间对比结果:省际共享接收资料的时效要比CMACast下行平均快11.5秒钟,可见省际共享在资料及时性方面有优势。

4.2 系统解码入库效率

解码时间由每次从FTP服务器中采集的自动站文件数和文件大小决定,解码过程最大耗时25.4S,平均耗时14.7S;入库过程平均耗时2.2S。基本能够在2分钟内(系统工作周期)完成外省资料解码入库和本省外省资料解码入库过程。

5 结语

本文针对气象资料省际共享实时监控需求设计和实现了安徽省省际共享监控系统。具体功能如下。

(1) 定时采集MySQL实时收发表得到文件名,再运用.NET FTP API获取相应文件的字节流并解码自动站信息、批量写入sjgx库自动站收发信息表;通过解析雷达基数据文件名得到站级信息并写入雷达基数据收发信息表。

(2) 按照监控告警策略,定时采集sjgx收发信息表中各省缺报信息并及时告警。

(3) 基于sjgx库统计各省日传输、月传输和年传输及时率。

(4) 运用SQL Server存储过程定时生成省际共享上传日志。

(5) 补入库功能, 保障收发信息的完整性。

系统运行稳定, 基本在 2 分钟内完成外省资料解码入库和本省发外省资料解码入库过程。

省际共享资料实时监测和统计信息为资料用户、业务管理部门提供了可靠的依据, 体现了气象资料省际共享实时业务质量监控的优势。

参考文献

- 1 王春芳, 李湘, 陈永涛, 等. 中国气象局卫星广播系统 (CMACast) 设计. 应用气象学报, 2012, 23(1): 113-120.
- 2 李显风, 邹海燕. 基于省际共享的外省自动站报文入库系统研究. 第 31 届中国气象学会年会. 北京, 中国. 2015.
- 3 中国气象局. 预报司关于下发 2014 年全国气象资料传输质量检查工作要求的通知. 气预函[2014]13 号, 2014.
- 4 杨润芝, 马强, 李德泉, 等. 内存转发模型在 CIMISS 数据收发系统中的应用. 应用气象学报, 2012, 23(3): 377-384.
- 5 熊安元, 赵芳, 王颖, 等. 全国综合气象信息共享系统的设计与实现. 应用气象学报, 2015, 26(4): 500-512. [doi: 10.11898/1001-7313.20150412]
- 6 华连生, 王建荣, 金素文, 等. 基于 IMS 的气象信息传输智能语音通知系统设计与实现. 气象科技, 2015, 43(6): 1040-1045.
- 7 Liberty J, Xie D. Programming C# 3.0 中文版. 李愈胜, 刘卫卫, 汪泳, 译. 第 5 版. 北京: 电子工业出版社, 2009.
- 8 帕特里克. ADO.NET 4 从入门到精通. 贾洪峰, 译. 北京: 清华大学出版社, 2012: 113-116.
- 9 Walters RE, Coles M, Ferracchiati F, 等. 深入 SQL Server 2008. 任斌, 刘芳芳, 译. 北京: 人民邮电出版社, 2011.
- 10 吴东丽, 梁海河, 曹婷婷, 等. 中国自动土壤水分观测网运行监控系统建设. 气象科技, 2014, 42(2): 278-282.