

多源气象数据实时推送系统的设计与实现

梁宝龙^{1,2,3}, 崔学林², 谢寒生^{1,2,3}, 贺永兴^{1,2,3}, 李晋峰²

(1. 海南省气象信息中心, 海南 海口 570203

2. 海南省南海气象防灾减灾重点实验室, 海南 海口 570203

3. 海南省气象信息化科技团队, 海南 海口 570203)

摘要:气象服务与人民群众的生活息息相关,特别是在台风、特大暴雨等重大天气过程中,气象预报预警信息的及时发布,对公众的防灾减灾尤为重要。为了整合不同数据源的气象数据,并实时推送给相关政府部门使用,设计并实现了一个多数据源的气象数据实时推送系统。该系统将不同来源的气象数据整合,并根据数据收集和存储方式,将气象资料划分为结构化数据和非结构化数据,并实现了气象数据的多策略推送。通过该系统的应用,可实时地将气象观测数据实况、气象预报服务产品等数据推送到政府部门的业务应用实时展示服务器中,为广大民众了解气象提供了更丰富的渠道,为海南当地政府部门的气象防灾减灾提供了高时效性的基础气象数据支撑。

关键词:气象;结构化数据;推送系统;多策略

中图分类号:TP302

文献标识码:A

文章编号:1673-629X(2018)08-0139-05

doi:10.3969/j.issn.1673-629X.2018.08.029

Design and Implementation of a Real-time Push System for Multi-source Meteorological Data

LIANG Bao-long^{1,2,3}, CUI Xue-lin², XIE Han-sheng^{1,2,3}, HE Yong-xing^{1,2,3}, LI Jin-feng²

(1. Hainan Meteorological Information Center, Haikou 570203, China;

2. Key Laboratory of South China Sea Meteorological Disaster Prevention and Mitigation of Hainan Province, Haikou 570203, China;

3. Hainan Meteorological Information Technology Team, Haikou 570203, China)

Abstract: Meteorological services are closely related to the lives of the people, especially in the typhoon, heavy rain and other significant weather process. The timely release of weather forecasting information is especially important for public disaster prevention and mitigation. In order to integrate multi-source meteorological data and push them to the relevant government departments, we design and implement a real-time push system for multi-source meteorological data. According to the meteorological data collection methods and storage structures, the system divides the meteorological data into structured data and unstructured data, and achieves the multi-policy push of meteorological data. Through the application of this system, many meteorological data such as real-time observation data and weather forecast service products are pushed to the business application server of the government department, which provides an extensive way for the general public to understand the meteorological services. For the local government departments in Hainan meteorological disaster prevention and mitigation, it provides a high efficiency of the basic weather data support.

Key words: meteorology; structured data; push system; multi-policy

0 引言

随着气象服务的广泛传播,气象天气预报、灾害预警等服务也越来越受到广大民众的关注。特别是在台风、特大暴雨等重大天气过程中,气象部门的预报预警

信息及时传播与公众的防灾减灾息息相关。如何让公众快速地了解当前的天气情况和未来天气的预报,除了常规的电视播报、手机短信和市、县、村等室外广播大喇叭等方式外,实时地将气象数据传输到海南省防

收稿日期:2017-09-18

修回日期:2018-01-23

网络出版时间:2018-03-07

基金项目:国家科技支撑计划课题(2013BAK05B03);海南省气象局技术提升项目(HNQXJS201608);海南省气象局重点项目(HNQXZD201406)

作者简介:梁宝龙(1986-),男,工程师,CCF 会员,研究方向为气象信息技术。

网络出版地址: <http://cnki.net/kcms/detail/61.1450.TP.20180307.1432.080.html>

汛防风防旱总指挥部办公室(简称省三防办)等政府门户网站实时展示,也为公众了解气象服务提供了更广泛的途径。通过气象数据的实时推送与更新,民众可以通过权威的政府部门提供的平台(如网站或者手机客户端等)了解实时的天气情况,同时也为管理部门的防灾减灾决策提供气象实况依据。

数据的传输有很多方式,近年来也有很多有指导性的工作^[1-5]。分别对远程文件同步进行了相关的研究工作^[6-8],有的针对远程文件同步^[9-10],有的基于 RSYNC^[11-13]算法的同步系统,均是对文件的同步与同步算法优化的研究,但这些研究仅仅针对的是文件的同步,没有结合文件数据与数据库表同步。

在上述研究的基础上,根据气象行业常用实时观测数据的共享及同步情况,将文件同步和数据库表的同步相结合,设计并实现了基于多源气象数据的快速推送系统。该系统整合不同数据源的气象数据到一个统一平台,其中将多源气象数据分为结构化气象数据和非结构化数据并快速推送,实现对不同数据源的气象数据的实时推送。该系统的主要功能是气象部门为省三防办等政府部门提供实时的气象数据推送,实时地将气象观测数据实况、气象预报服务产品等数据推送到政府部门的业务应用服务器中,为当地的政府部门防灾减灾决策提供气象数据基础支撑。

1 系统设计

1.1 设计原则

(1)时效性。

气象数据的数据传输涉及到业务考核,时效性要求很高,要求各类考核的气象观测数据在考核时间内传输完毕。推送的时效需要在满足业务要求的前提下,尽快完成推送任务。各类数据的推送,必须在采集完毕之后在规定的时间内完成推送工作。

(2)可扩展性。

系统推送的设计方面,尽可能考虑到将来业务的变化情况,保证将来有新的资料推送需求时,在满足现有推送任务的前提下,能灵活配置推送策略满足推送需求变更。

(3)安全性。

系统建立在安全可靠的网络安全措施之上,确保系统软硬件及系统中的数据受到保护,不因偶然的或者恶意的原因而遭受到破坏、更改、泄露,实现系统连续可靠正常的运行,网络服务不中断,推送系统和接收端之间的通信安全可靠。

(4)易用性。

推送的策略可以通过图形界面进行推送策略配置,操作步骤可视化,操作方法简单易学。

1.2 结构设计

各类气象数据收集分布在不同的业务系统中,为了更好地进行控制推送的集中管理,文中对各业务系统中的数据服务器进行访问权限配置,推送服务器的访问权限可以读取所有业务服务器的数据存储目录和文件,便于访问获取数据。同样地,对于实时观测要素数据库的服务器,推送服务器对数据库有读写访问权限。依据气象数据推送业务的需求,将推送系统功能规划为三个子模块,分别为数据收集存储模块、数据推送服务器模块、数据接收服务器模块,如图 1 所示。

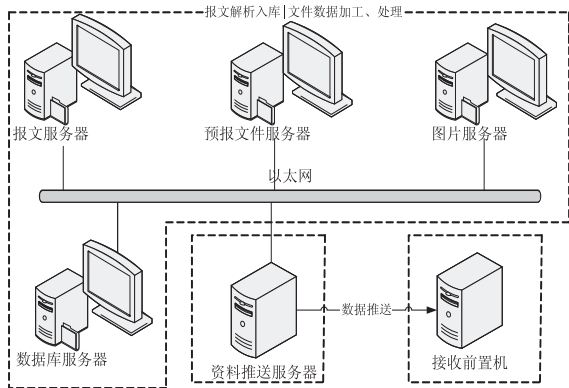


图 1 推送系统结构

1.3 功能设计

推送系统设计主要分为数据收集存储、数据推送服务、数据接收存储三个模块的设计。

(1)数据收集存储模块。

该模块功能是数据源整合的部分,包括常规观测报文收集服务器,预报文件服务器,雷达图片、云图图片服务器等中心站业务服务器。各业务中心站负责一种业务的运行,并保障数据的及时收集和存储。收集存储的各服务器并行执行各自的收集及存储任务,业务之间相互不干涉,极个别业务的暂时中断不影响其他业务数据收集存储的独立运转。该模块的主要功能是对报文要素解析入库,文件数据加工处理,预报产品文件存储,雷达云图产品的加工并处理成 JPG 图片文件格式存储。

(2)数据推送服务器模块。

数据推送服务器模块的设计,通过开发一个多任务平台,实时地将数据收集模块的各类数据同步到接收服务器的指定存储目录中。推送服务器作为数据推送系统的关键模块,对内联系局域网数据收集模块中的各节点服务器,对外负责实时将各类数据推送到省三防办的接收前置机中。分别对推送数据的各类中心站数据服务器进行配置,使得资料推送服务器可以获得需要推送的各类数据的读写访问权限,获取推送的数据的存放目录及文件读取访问权限。

(3)数据接收服务器模块。

数据接收服务器模块,其存储目录和数据库表结构与推送服务器的存储结构一一对应,并且推送服务器对接收服务器的各推送目录均有读写访问权限。接收服务器及时处理推送服务器推送过来的数据,并定时清理过时数据文件。

在系统推送设计中,数据接收服务器模块是气象业务外网物理隔离的节点。基于网络安全的考虑,通过网络防火墙等安全设备的设置,气象局域网的推送服务器是唯一能跟省三防办对端接收前置机通信的服务器,它们之间的通信是点对点的。气象内网的数据收集服务器均与接收前置机服务器物理隔离,相互之间网络访问不可达。同样,接收前置机也不能访问气象内网的其他业务中心站服务器。

2 数据推送系统实现

对于气象数据的推送,文中分别针对非结构化和结构化数据的特点,分为两种方式实现。对于非结构化的气象数据,例如雷达产品、气象云图的以文件形式存在的数据,是以共享目录拷贝的方式实现数据的实时同步,同步的策略是根据观测数据的生成规则,以及数据的观测时效和数据文件生成结束的时间来判定。例如雷达产品,推送策略是在雷达产品的生成之后立即将雷达数据推送至接受服务器的对应目录中,并将已推送数据的文件名入库,避免再次拷贝时的重复推送。

多源气象数据推送系统分别对不同的来源数据做不同的推送响应,有的数据可直接推送,有的数据需要经过加工成产品后再推送,有的数据需要临时处理、转存等步骤之后推送。数据推送系统的推送过程如图 2 所示,分别包括数据收集、收集处理、数据临时存档、推送处理、数据推送、推送控制策略等。

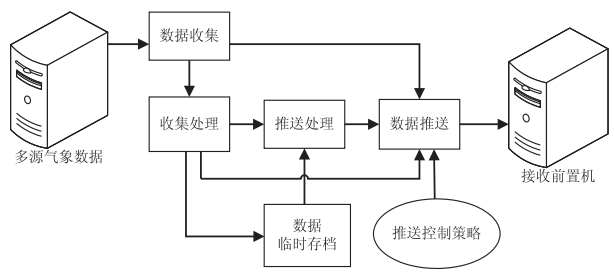


图 2 系统推送过程

数据的收集,是对各类气象数据的收集整合,包括即将推送的结构化和非结构化气象数据。收集处理,是指需要对原始资料加工成产品之后再推送的数据,例如雷达图和卫星云图等数据。数据临时存档,是指推送时需要本地临时存储数据,以便推送中出现异常后,恢复时继续之前的推送任务。推送处理,是指对收集处理之

每一类数据的顺利推送。数据推送,是指直接将多源的各类气象数据实时推送到接收前置机中。推送控制策略,是指对每一类数据的推送方式和推送时效的触发控制,达到对每一类数据推送指令的配置目的,实现对每一类数据推送策略的配置与管理。

2.1 非结构化数据推送

可推送的非结构化常规气象数据有:单站雷达回波图、降雨预报分布图、卫星云图、气温分布图、旱情分布图、预报信息、决策气象服务、气候分析、生活气象等。其中,以 JPG 图片等文件格式推送的气象数据有雷达回波图、降雨预报图、卫星云图、气温分布图;以 WORD 或者 PDF 格式文件形式推送的有决策气象服务稿件、气候分析和生活气象等。

非结构化数据对的推送是以文件形式推送到指定的服务器的指定目录中,具体的推送策略符合《全国气象数据传输质量考核办法》的要求,以气象观测数据的传输考核时效和观测数据生成时间为依据。对于观测时效比较规律的数据,设置好推送主机和接收服务器的同步时间,计算并判断好观测数据生成的时间,一旦当前观测过程结束并生成观测数据,立即触发推送策略响应,将当前生成的最新观测数据推送到接收服务器指定的目录中,实现数据的实时推送。

2.2 结构化数据推送

对于结构化气象数据的推送,文中选择的是数据库表要素的实时同步。笔者认为对数据库表的实时同步,是对结构化气象数据实时展示的时效要求和数据出口一致性的,比较符合气象要素实时展示的常用方法。结构化的数据推送,在省三防办等部门网站的实时显示是通过读取数据库表要素实时更新展示的,文中方法是将气象实时收集数据入库的数据库表要素直接同步到对端的数据库表中。根据每一类数据的观测时效来定义数据库表同步的策略,按类型划分数据时效实现结构化数据的实时同步。这些同步方式可以通过增量的方式同步,也可使用外部的数据库管理工具实现。在数据写入过程中,同时遵循数据库事务的 ACID^[14-15]四个特性,确保数据库写入的表要素一一对应和表结构要素的完整性,确保气象数据要素展示的结果唯一。

3 系统相关技术

3.1 系统的开发环境

开发平台:eclipse;数据库:Oracle 11g;Web 应用服务器:apache-tomcat-7.0.57;数据库中间件:JPA (Java persistence API);实现 Web 层:Web2.0、Struts2、jQuery;推送服务器的操作系统:Windows Server 2012 R2。

3.2 监控和策略设置

多源气象数据的推送系统,针对每一类数据配置相对应的策略,实现每一类数据的推送独立运行且互不干涉。非结构化的数据推送将本地文件数据直接推送到接收服务器的对应目录。结构化数据的推送,选择数据库同步的常用数据库中间件实现。推送系统策略配置如图 3 所示。每一类推送策略都可以通过 Web 配置修改,一般可选的推送策略根据气象数据的采集时次分别定义“全部”、“整点”、“非整点”等配置选项。“全部”选项表示所有数据均推送,“整点”选项表

示只推送每个小时整点的数据,“非整点”选项表示只推送可选非整点的数据。推送策略可根据观测数据的时效自行定义和实现。

为了实时监控推送记录和配置推送策略,将每一数据的推送记录存入数据库,记录推送 IP 地址、推送文件名、推送的时间、推送目的主机地址和推送数据记录等详细信息,推送策略可通过 Web 监控平台配置,从而实现推送数据的实时监控。推送日志查询见图 4,实现推送的全流程监控。



图 3 系统推送策略配置



图 4 系统推送日志

4 系统推送时效分析

较以往三防办等部门通过气象内网获取气象数据的方式,多源气象数据实时推送系统正式上线以后,数据推送到省三防办等部门的时效性也明显提高。笔者对比了几类气象数据,即自动站、雨量站、雷达资料、国家级站、土壤水分,其资料推送时效分析如图 5 所示。

在该实时推送系统中,自动站、雨量站、雷达资料、国家级站、土壤水分等数据推送时效均在资料到达后

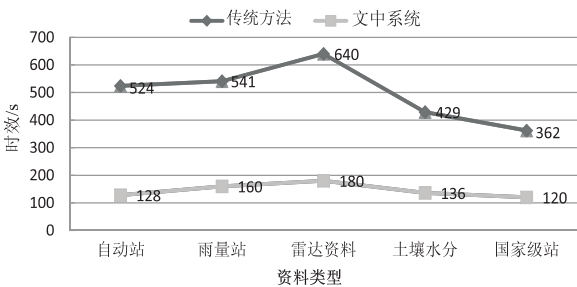


图 5 传输时效对比

三分钟以内推送到对端接收前置服务器中,较短的时间是自动站和国家级站,分别是资料到达后 128 s 内推送到接收前置机的数据库表中。传统的传输方式使得数据获取时效比较慢,原因可能是气象数据源分布在不同的观测系统中,数据分别分散地存储在不同服务器中,获取数据前没有对数据进行有效的整合和预处理困难等等。

推送系统实现后,通过预先整合各类气象数据源到一个统一的平台,对数据进行推送预处理并按需定制策略推送,这样优化了数据推送环节,有效保障了数据的推送时效。同时也可根据业务扩展实现推送资料的种类定制推送,对每一类数据的推送方式和推送时效的触发控制,达到对每一类数据推送指令的配置目的,实现对每一类数据推送策略的配置与管理,提高了传输时效。

5 结束语

从气象数据共享的方式,设计并实现了气象数据的快速推送系统,并在海南省气象局的内网中实现了对省三防办所需的气象数据的实时推送,为海南当地的气象服务提供了便捷高效的气象资料推送服务。该系统上线以来,在强台风、暴雨等重大天气过程中,发挥了积极的作用。下一步,将开发多策略预订等推送服务,不断完善系统功能,提供更高时效性的气象数据推送服务,为海南的经济发展和海南气象服务贡献一份绵薄之力。

参考文献:

[1] 李 强,朱立谷,曾赛峰,等. 分布式目录同步的冲突处理与副本管理研究[J]. 计算机研究与发展,2012,49:257-262.

[2] 李 强,朱立谷,曾赛峰. 弱连接网络环境中的多主机文件同步[J]. 小型微型计算机系统,2012,33(10):2182-2187.

[3] 胡晓勤,卢正添,刘晓洁,等. 远程文件快速同步方法[J].

电子科技大学学报,2008,37(4):594-597.

[4] 徐 旦,生拥宏,鞠大鹏,等. 高效的两轮远程文件快速同步算法[J]. 计算机科学与探索,2011,5(1):38-49.

[5] 李 雪,咸 迪. 跨平台文件共享系统设计与实现[J]. 计算机技术与发展,2012,22(6):191-194.

[6] 敖 莉,舒继武,李明强. 重复数据删除技术[J]. 软件学报,2010,21(5):916-929.

[7] 杨小龙,李 涛,刘晓洁,等. 基于差异的文件同步系统的设计 and 实现[J]. 微计算机信息,2009,25(3-3):67-69.

[8] 曾 珊,周 薇,韩冀中. 面向多应用的文件同步方法[J]. 计算机应用与软件,2016,33(10):297-302.

[9] 何 骞,卓碧华. 一种远程文件同步方法[J]. 计算机应用,2012,32(2):566-568.

[10] 汤晓迪,马晓旭,宁 静,等. 远程文件差异同步系统的设计与实现[J]. 计算机工程与设计,2010,31(20):4389-4392.

[11] 李 贞. 基于 Rsync 算法的远程文件同步系统的设计与实现[D]. 北京:北京邮电大学,2010.

[12] SUEL T, NOEL P, TRENDAFILOV D. Improved file synchronization techniques for maintaining large replicated collections over slow networks[C]//Proceedings of the 20th international conference on data engineering. Boston, MA, USA:IEEE,2004:153-164.

[13] IRMK U, MIHAYLOV S, SUEL T. Improved single-round protocols for remote file synchronization[C]//24th annual joint conference of the IEEE computer and communications societies. Miami, FL, USA:IEEE,2005:1665-1676.

[14] ARMBRUST M, FOX A, PATTERSON D A, et al. SCADS: scale-independent storage for social computing applications[C]//Proceedings of 4th Biennial conference on innovative data systems research. Asilomar, CA, USA:IEEE,2009:86-93.

[15] YANG Fan, SHANMUGASUNDARAM J, YERNENI R. A scalable data platform for a large number of small applications[C]//Proceedings of 4th Biennial conference on innovative data systems research. Asilomar, CA, USA:IEEE,2009:17-27.

(上接第 138 页)

Advanced Materials Research,2014,971-973:1998-2003.

[3] 胡方炜. 基于嵌入式 Linux 系统的 SNMP 代理开发研究[D]. 杭州:浙江工业大学,2004.

[4] 张雪锋. 基于实时内核 $\mu\text{C}/\text{OS-II}$ 的应用研究[D]. 成都:西南交通大学,2004.

[5] LABROSSE J J. 嵌入式实时操作系统 $\mu\text{C}/\text{OS-II}$ [M]. 北京:北京航空航天大学出版社,2003.

[6] XUE Sujing. Discuss the design method of topology discovery system based on SNMP[J]. Advanced Materials Research,2011,267:594-598.

[7] TIAN Jinghua, WANG Jun. Application software based on SNMP in the teaching of networks management[J]. Advanced Materials Research,2011,271-273:1816-1818.

[8] 严蔚敏. 数据结构[M]. 北京:清华大学出版社,1996:305-379.

[9] 马海红. 基于 ARM 和 RTOS 的嵌入式 GUI 应用平台设计[D]. 武汉:武汉理工大学,2005.

[10] ARM 发布 Keil Vision4 集成开发环境为嵌入式软件开发提供简化的开发环境[J]. 电子与电脑,2009(4):70.

[11] 陈承慧. 基于 FS8610 的嵌入式 SNMP 代理的分析与实现[D]. 苏州:苏州大学,2009.

[12] 姜晓琳,王厚军,马 敏. LXI 标准下的网络 VISA 设计与实现[J]. 计算机测量与控制,2010,18(8):1849-1851.

[13] 刘小兵. 基于 LXI 网络化仪器通信模块的设计与实现[D]. 西安:西安邮电大学,2016.

[14] 葛化敏,杨利青,郑 静. 基于 NET-SNMP 的嵌入式网络管理代理的开发[J]. 通信技术,2009,42(11):138-140.