

何文春, 高峰, 许艳, 等. 气候监测业务专题数据服务系统的设计与实现. 应用气象学报, 2012, 23(5): 624-630.

# 气候监测业务专题数据服务系统的设计与实现

何文春\* 高峰 许艳 冯明农 孙超 赵春燕

(国家气象信息中心, 北京 100081)

## 摘 要

为了满足气候监测业务日益突出的个性化资料需求, 设计和开发了气候监测业务专题数据服务系统。在分析用户资料需求的基础上, 重点介绍了系统的设计思路和实现技术。该系统由 1 个专题数据支撑库和 4 个功能模块组成, 包括资料更新发现、专题数据集制作、数据服务和任务调度, 其中, 专题数据集制作方法和 3 种任务调度模式是系统设计的核心。系统实现采用 SSH(Struts Spring Hibernate) 框架, 按数据访问层、业务逻辑层和交互表现层进行分层实现; 同时, 结合 DWR(Direct Web Remoting) 技术实现了同步异步交互方式, 使用灵活的配置管理加强系统的可扩展性。系统投入运行后, 已制作出 13 个专题数据集, 提供了良好的数据服务。

**关键词:** 专题数据服务; 气候监测; 数据集制作; 任务调度; SSH 框架

## 引 言

气象资料服务有在线服务和离线服务两种形式, 在线服务是主要手段, 通过气象资料服务系统实现<sup>[1]</sup>。随着气象事业和气象业务系统的飞速发展, 对在线数据支撑的要求也越来越高, 包括对资料本身的要求和对服务功能的要求。

气象部门气象资料服务系统的建设和业务运行开始较早, 其发展历程大致分为 4 个阶段<sup>[1-5]</sup>: 20 世纪 80 年代初开始的单机文件检索阶段, 90 年代中期开始的接口调用阶段, 90 年代末开始的数据库系统阶段, 21 世纪初开始的分布式数据服务阶段。当前, 业务运行的国家级资料服务系统主要包括国家级气象资料存储检索系统(MDSS)<sup>[4,6]</sup>和气象科学数据共享系统<sup>[4]</sup>, 前者主要服务于国家级气象业务和科研, 后者主要针对行业用户和公众用户。

MDSS 基本满足了国家级气象业务、科研工作对气象资料的使用<sup>[6]</sup>, 但该系统对所有用户提供的服务相同, 包括提供的资料和服务手段, 这与不同业务系统对资料的个性化需求的矛盾逐渐凸显, 从而影响气象业务系统的运行和发展。因此, 气象资料服务系统需要在满足用户普遍需求的同时, 开始发

展专题数据服务。本文在分析气候监测业务资料需求特点的基础上, 重点介绍了气候监测业务专题数据服务系统的设计思路和实现技术。该系统提出和实现了以专题数据支撑库为基础, 以资料更新发现、专题数据集制作、数据服务和任务调度 4 个模块为核心的系统结构, 总结出了具有典型参考和推广价值的专题数据集制作方法、任务调度策略以及 SSH(Struts Spring Hibernate) 框架<sup>[7]</sup>结合 DWR(Direct Web Remoting)<sup>[8]</sup>和配置管理技术的实现手段。通过该系统, 目前已制作出“地面降雪观测资料日值”等 13 个专题数据集, 为“陆面冰雪监测业务系统”等 3 个系统提供了良好的数据服务。

## 1 气候监测业务的资料需求特点

近年来, 极端天气气候事件的发生给社会、经济和人民生活造成了严重的影响和损失<sup>[9]</sup>, 国家级气候监测业务也在迅速发展, 冰雪监测、干旱监测、气候要素监测、台风监测、季风监测、高空气候变化监测等业务系统在不断建设或完善中。这些业务系统须通过国家气象信息中心提供的在线服务系统获取数据, 按期生成气候监测产品, 提供决策服务和气候预测等使用。

2011-09-26 收到, 2012-04-16 收到再改稿。

\* E-mail: hewc@cma.gov.cn



首先,系统研制的数据集制作方法严格遵循相关规范,包括《地面气象观测规范》<sup>[10]</sup>(以下简称《规范》)、《全国地面气候资料(1961—1990)统计方法》<sup>[11]</sup>(以下简称《统计方法》)和《地面气候资料 30 年整编常规项目及其统计方法》<sup>[12]</sup>等。如系统的地面自动站风资料日值统计方法遵循《统计方法》中关于风要素的日值统计方法、《规范》中缺测记录的处理和不完整记录的统计规定。其次,数据集制作方法也符合专题服务的特点,尽可能多地给用户提实用资料。如系统的地面自动站风资料日值统计方法在《规范》和《统计方法》相关规定的基础上,结合要素的质控标识、要素间的关联关系、用户的统计要求,进行了更丰富的补充和扩展。

专题数据集制作方法的具体内容包括 4 个部分:资料分析、筛选整理、统计加工、标准化处理。

• 资料分析 与用户某一特定需求相关的资料可能有多种,对其中的每类原始资料进行分析,分析其站点分布、质量情况、资料时次、更新频率、时间序列等,在此基础上选择符合用户业务需求的一个或几个数据源。如在 MDSS 的实时库中,涉及到土壤湿度观测的有农气旬月报资料、自动站土壤湿度资料等几种资料,经分析并结合气候监测业务的需要,选择从农气旬月报中提取相关要素制作数据集比较合适;它虽然更新频率较慢(每旬更新),7 层的观测层次不及自动站土壤湿度资料 20 层的观测层次多,但资料质量较好,观测台站分布均匀且较多,历史序列长,更符合气候监测业务的要求。

• 筛选整理 按业务需要,从数据源中挑选所需要素,剔除全为缺测的记录和质量错误的要素值,同时补充必要的台站信息,包括经度、纬度、海拔高度、台站名称等。

• 统计加工 对于需要统计加工的数据集,主要是实时资料数据集,制定详细的加工处理方法和步骤,如日值统计、极值统计等。

• 规范化处理 对加工出的数据集产品,进行规范化处理以生成标准的数据产品提供服务。规范化处理过程包括:统一的特征值转换、统一的量纲和单位转换、规范的文件组织、规范的文件命名和规范的文件内容格式。

2.2.4 数据服务

不同的用户业务系统对数据获取方式要求不同,分为主动获取和被动获取两种。因此,本系统提供两种数据服务方式:FTP(文件传输协议)下载和数据推送。其中,FTP 下载的实现完全基于数据文件存储的标准目录结构和文件命名;数据推送由专题数据集制作过程来驱动,当新的数据文件制作好后,根据用户需求配置自动生成 1 个或多个新的推送任务,以完成向 1 个或多个业务系统实时推送的功能。

2.2.5 任务调度

为了满足实时业务要求,保证稳定可靠的数据服务,本系统设计的任务调度由实时自动调度、异常处理调度和人工补做调度 3 种调度模式组成,如图 2 所示。3 种调度模式相互配合,构成一个完备的调度体系。

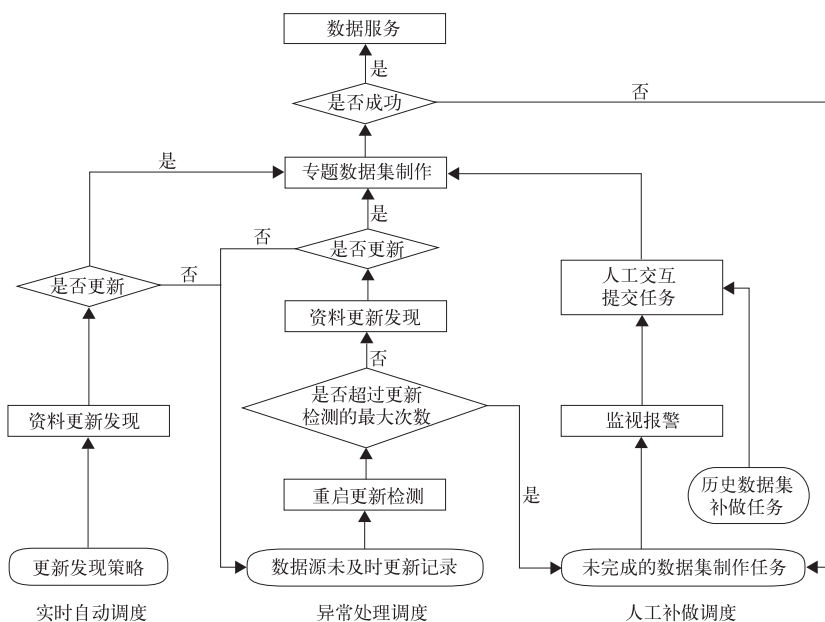


图 2 任务调度流程设计图  
Fig. 2 Job scheduling flow

在正常情况下,实时自动调度完成所有数据集制作和服务。根据资料的更新发现策略,调度资料更新发现模块,若已更新则启动数据集制作和服务,否则,记录资料未更新信息,交付给异常处理调度。

在资料源更新延迟情况下,异常处理调度自动处理。获取数据源未及时更新的信息,在一段时间内多次尝试处理;若数据源始终不更新,则交付给人工补做调度。

对未完成的数据集制作任务和历史数据集补做任务,采用人工补做调度来完成。针对数据源未更新、数据集制作中出现故障等信息,通过监视报警,提醒业务人员进行处理;业务人员指定数据集名称和时间段等条件,通过人机交互方式进行调度,完成补做。

### 3 系统实现

#### 3.1 基于 SSH 的技术架构

该系统以 SSH 框架为基础,结合 DWR 的 AJAX(Asynchronous JavaScript and XML)技术<sup>[13]</sup>和配置管理技术,构成了一个全新的更高效灵活的技术架构。SSH 框架由 3 个开源框架组成:

Struts<sup>[14]</sup>, Spring<sup>[15]</sup> 和 Hibernate<sup>[16]</sup>, 它们有效配合,提供了一种开发 J2EE<sup>[17]</sup> (Java 2 Platform, Enterprise Edition)企业级 Web 应用的轻量级解决方案。Struts 采用 Servlet/JSP<sup>[18]</sup> 技术,实现了基于 MVC(Model View Controller)设计模式<sup>[19]</sup> 的 Web 应用框架,尤其是系统选用的 Struts 2<sup>[14]</sup> 更是对 Struts 1 进行了改进,成为目前 Java 领域最优秀的同步交互 Web 框架。Spring 有效解决了传统企业应用开发模式的复杂度问题,使开发人员使用基本的 JavaBean<sup>[15]</sup> 来完成以前只可能由 EJB<sup>[17]</sup> (Enterprise JavaBean)完成的事情,同时为 J2EE 应用开发提供集成的框架。Hibernate 是一个开源的对象关系映射框架,它对 JDBC(Java Database Connectivity)进行了非常轻量级的对象封装,使得开发人员可以方便地使用对象编程思维来操纵数据库。AJAX 是目前实现异步交互的默认标准,实现该技术的 DWR 是一个轻便和实用的框架;通过将 DWR 集成到 SSH 中,使之弥补 Struts 在异步交互上的不便,与其搭配实现同步、异步交互的完备模式。

该架构在本系统中的应用如图 3 所示,其特点及优势在于其分层结构和配置管理。

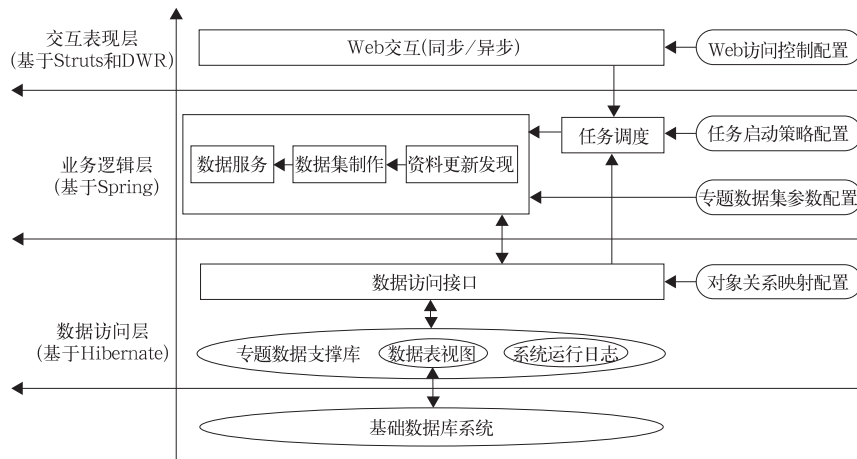


图 3 气候监测业务专题数据服务系统技术架构  
Fig. 3 The technical architecture of Special Data Service System for Climate Monitoring

首先,该架构具有明显的松耦合层次划分,下层对上层提供服务式的接口,上层通过配置和调用接口来访问下层。系统在实现时自下而上分为数据访问层、业务逻辑层和交互表现层等 3 个层次。数据访问层基于 Hibernate,屏蔽数据库结构化的数据结构和 SQL(Structured Query Language)的访问方

式,提供面向对象的访问接口。业务逻辑层基于 Spring,实现系统的业务功能,并提供交互接口。交互表现层基于 Struts,结合 DWR,实现 Web 展示,提供操作人员与系统之间的交互。

其次,该架构通过一系列的配置管理,来支撑系统的灵活度和可扩展性。主要配置包括两类:业务

逻辑配置和框架参数配置。业务逻辑配置是从业务功能实现中抽取出的通用配置信息,配置所用 XML (eXtensible Markup Language) 文件架构 (Schema)<sup>[20]</sup> 由系统自行设计;本系统中该类配置包括专题数据集参数配置、任务启动策略配置等。框架参数配置是支撑开源框架运行的配置,配置所用 XML 文件的架构由框架自身定义;本系统中所使用的开源框架主要有 Hibernate, Spring, Struts 和 DWR, 其配置分别为对象关系映射配置、对象管理参数配置、同步 Web 访问控制配置和异步 Web 访问控制配置。

### 3.2 系统功能的实现

#### 3.2.1 数据访问层的实现

在数据访问层中,通过数据表视图的创建,屏蔽不同的数据库来源,为数据访问接口的实现提供统一的数据源;数据访问接口实现数据表的对象化,封装数据访问方法,为上层提供接口服务(图 3)。

数据表视图存放在专题数据支撑库中,其创建分为两个步骤:①DBLink<sup>[21]</sup>(数据库链接)的创建。基础数据库系统可由多个数据库组成,目前包括 MDSS 中的实时数据库和综合数据库;为了创建数据表视图,需要在专题数据支撑库中创建连接到这些基础库的 DBLink,通过设置基础数据库 IP 地址、实例名、用户名、密码等参数进行创建。②数据表视图的创建。在 DBLink 创建的基础上,创建专题数据制作需要访问的数据表视图,过滤掉数据集制作

不需要的要素。数据表视图包括资料要素表视图和资料更新发现所需的元数据表视图。

在数据访问接口的实现中,通过配置 Hibernate 里的对象关系映射,完成专题数据支撑库中数据表视图和系统运行日志表的持久化映射<sup>[16]</sup>;使用 Hibernate 中 Session 管理数据对象的功能,通过程序开发,完成对数据对象的访问及操作方法的封装,为上层提供访问接口。

#### 3.2.2 业务逻辑层的实现

业务逻辑层基于 Spring 实现,主要是对资料更新发现、专题数据集制作、数据服务和任务调度 4 个核心业务模块的开发。

在实现过程中,充分利用开源软件,提高系统稳定性和开发效率。如数据服务中的 FTP 下载功能,借助 Apache FtpServer<sup>[22]</sup> 软件来实现;任务调度中的实时自动调度功能,使用 Quartz<sup>[23]</sup> 软件来实现。Quartz 是 OpenSymphony 开源组织的一个功能齐全的任务调度系统,将其集成到 Spring,然后辅以任务启动策略配置,配置每个任务的启动时间或重新启动的时间间隔,即可完成实时任务启动控制。

此外,专题数据集参数配置是业务逻辑层实现的关键,它直接指导资料更新发现、专题数据集制作和数据服务的实现及运行,从这几个环节中抽取出的重要参数信息,主要包括专题数据集相关描述信息和源数据集信息(表 1)。

表 1 专题数据集参数配置表

Table 1 The parameters of special dataset processing

参数种类	参数名称	参数取值示例
专题数据集 信息	数据集代码	SURF_CLL_CHN_SNO_DAY
	数据集名称	中国地面气候资料日值积雪数据集
	存储子目录	SURF_CLL_CHN_SNO_DAY/YYYY
	数据文件名称	SURF_CLL_CHN_SNO_DAY-YYYYMMDD.TXT
	数据集每个文件的时间范围及单位	1 d
源数据集 信息	源数据集代码	SURF_CLL_CHN_MUL_DAY
	源数据集每次的更新量及单位	1 个月
	源数据集更新的滞后时间及单位	3 个月
	源数据集更新检查首次启动的时间模式	YYYYMM12HHMISS
	源数据集更新检查的间隔时间及单位	1 d
	源数据集更新检查的最多次数	15

#### 3.2.3 交互表现层的实现

本系统的交互表现层,主要是实现数据管理人员对数据集人工补做任务的提交,使用 Struts 结合 DWR 来实现同步和异步两种交互方式。

系统通过检索专题数据支撑库,获取可制作的专题数据集列表及每个数据集待制作的时间范围,使用 Struts 动态加载到页面上,提供给数据管理员;数据管理员选择需要补做的数据集,指定补做的



时间范围,设置若数据文件已生成是否需要覆盖等参数,通过DWR提交给后台进行异步处理,不需要管理员因后台处理而长期等待。

#### 4 结 论

气候监测业务专题数据服务系统是针对气候监测业务的具体需求,设计和开发的一套服务系统,是在专题服务领域的一次成功尝试。与传统的服务系统相比,该系统的设计和实现具有以下特点:

1) 系统功能结构及调度体系具有典型的专题服务特征,保障用户能稳定、实时、自动地获取到个性化的专题数据。

2) 专题数据集制作方法既具有科学性,又能满足业务实用性。

3) 松耦合的分层实现和灵活的配置管理,使系统具有良好的可扩展性,便于专题服务的持续发展。

今后,本系统也将不断扩展和完善,为更多的业务系统提供支撑,推进专题数据服务工作向用户更满意的方向发展。

#### 参 考 文 献

- [1] 高峰,王国复,喻雯,等. 气象数据文件快速下载服务系统的设计与实现. 应用气象学报, 2010,21(2):243-249.
- [2] 王国复,李集明,邓莉,等. 中国气象科学数据共享服务网总体设计. 应用气象学报, 2004,15(增刊):10-16.
- [3] 李集明. 基于Internet面向社会公益性共享的气象科学数据存储检索策略研究. 气象科技, 2007,35(4):590-592.
- [4] 李集明,沈文海,王国复. 气象信息共享平台及其关键技术研究. 应用气象学报, 2006,17(5):621-628.
- [5] 李集明,王国复. 气象数据库系统总体设计综述. 气象科技, 2007,35(增刊):1-5.
- [6] 沈文海,赵芳,高华云,等. 国家级气象资料存储检索系统的建立. 应用气象学报, 2004,15(6):727-736.
- [7] 李刚. 轻量级J2EE企业应用实战——Struts+Spring+Hibernate整合开发. 北京:电子工业出版社, 2007.
- [8] Zammetti F. Practical DWR 2 Projects. New York: Springer. 2008.
- [9] 刘绿柳,孙林海,廖要明,等. 基于DERF的SD方法预测月降水和极端降水日数. 应用气象学报, 2011,22(1):77-85.
- [10] 中国气象局. 地面气象观测规范. 北京:气象出版社, 2003.
- [11] 国家气象中心. 全国地面气候资料(1961—1990)统计方法. 1990.
- [12] 中国气象局. 地面气候资料30年整编常规项目及其统计方法. 2004.
- [13] Crane D, Pascarello E, James D. Ajax in Action. New York: Manning Publications, 2005.
- [14] Brown D, Davis C M, Stanlick S. Struts 2 in Action. New York: Manning Publications, 2008.
- [15] Walls C, Breidenbach R. Spring in Action. New York: Manning Publications, 2008.
- [16] Bauer C, King G. Hibernate in Action. California: O'reilly & Associates, Inc, 2005.
- [17] Broemmer D. J2EE Best Practices: Java Design Patterns, Automation, and Performance. New Jersey: John Wiley & Sons Inc, 2002.
- [18] Chopra V, Eaves J, Jones R, et al. Beginning JavaServer Pages. New Jersey: John Wiley & Sons Inc, 2005.
- [19] Buschmann F, Mcunier R, Rohnert H, et al. Pattern-oriented Software Architecture: A System of Patterns. New Jersey: John Wiley & Sons Inc, 1996.
- [20] Walmsley P. Definitive Xml Schema. New Jersey: Prentice Hall, 2005.
- [21] Abbey M, Corey M, Abramson I. Oracle9i: A Beginner's Guide. New Jersey: McGraw-Hill, 2002.
- [22] Apache FTPServer. [2011-02-18]. <http://mina.apache.org/FTPserver/>.
- [23] Cavaness C. Quartz Job Scheduling Framework: Building Open Source Enterprise Applications. New Jersey: Prentice Hall, 2005.

## Design and Implementation of Special Data Service System for Climate Monitoring

He Wenchun Gao Feng Xu Yan Feng Mingnong Sun Chao Zhao Chunyan

(National Meteorological Information Center, Beijing 100081)

### Abstract

Currently, the national meteorological data service systems have basically met the demand of operational systems and scientific research. But, as the deep development of the meteorological operations, it's urgent to develop special data services to meet the specific requirements. For instance, the special data service system for climate monitoring is designed and implemented.

The system is made up of one supporting database and four modules: Data-update monitoring, special dataset producing, data serving and task scheduling modules. The supporting database is in charge of fetching meteorological data from basic databases and storing information for the system operation. The data-update monitoring module detects the status of data updating with the strategies designed for each dataset, and triggers the special dataset producing module. The special dataset producing module produces datasets with scientific and practical methods, including analyzing, filtering, arranging, making statistics and standardizing of original data. The data serving module provides two ways for service: Data pushing and downloading via FTP. The task scheduling module dispatches the other modules according to business rules, realizes a complete operational flow perfectly by implementing three scheduling modes: Real-time scheduling, exception handling, and manual calling.

The system is developed based on SSH architecture, DWR technology and configuration management. SSH involves three open-source products: Struts, Spring and Hibernate, which separately implement three layers of the system: Web representation layer, business logic layer and data access layer. The web representation layer provides a web platform for data manager submitting manual calling of producing datasets with the method of synchronous and asynchronous interaction, realized by using Struts and DWR. The business logical layer is based on Spring product, which implements all of the four operational modules, and implements the calling interface for the web representation layer. The data access layer provides data fetching interface by creating views of data in basic databases and encapsulating data access objects with Hibernate. In addition, the skills of configuration management, including the business configuring and the third-party software configuring, enhance the flexibility and scalability of the system.

The system is a successful attempt on special meteorological data service. After being put into operation, it produces thirteen real-time updating special datasets, and serves climate monitoring systems very well. The specific dataset producing method, complete strategy of task scheduling, management skills of configuration, and the technical framework of the system provide a widely reference and promotional value for other special meteorological data service systems' design and implementation.

**Key words:** special data service; climate monitoring; dataset producing; task scheduling; SSH framework