

长江流域气象水文雨量资料一张图分析系统建设与应用^①



杨代才, 张冰松

(湖北省气象信息与技术保障中心, 武汉 430074)
通讯作者: 杨代才, E-mail: 11509460@qq.com

摘要: 本文基于目前对气象和水文共享站点的筛选, 根据云分布式存储技术和气象数据处理技术, 详细阐述在“一张图”目标实现过程中, 系统总体设计的主要思路、功能结构及产品应用, 同时梳理了降水资料的实时汇集流程, 设计数据处理业务系统以及数据产品制作和共享系统, 实现气象与水文部门长江流域雨量观测数据一张图的目标, 实现了与水文部门数据的高效交换与共享应用。

关键词: 气象; 水文; 融合; 一张图; 共享

引用格式: 杨代才, 张冰松. 长江流域气象水文雨量资料一张图分析系统建设与应用. 计算机系统应用, 2019, 28(1): 47-52. <http://www.c-s-a.org.cn/1003-3254/6734.html>

Construction and Application of One Map of Meteorological and Hydrological Rainfall Data on the Yangtze River

YANG Dai-Cai, ZHANG Bing-Song

(Meteorological Information and Technology Support Center, Hubei Province, Wuhan 430074, China)

Abstract: Based on the filter of meteorological and hydrological stations and according to the cloud distributed storage technology and meteorological data processing technology, the main idea, function structure, and product application of the system design are introduced. The real-time collection process of precipitation data is combed, and the data processing and sharing system are established. The system integrates the rainfall data of the Yangtze River from meteorological and hydrological departments, and realizes the efficient data exchange and sharing application with hydrology department.

Key words: meteorology; hydrology; fuse; one map; share

引言

长江流域覆盖范围较大, 包含 12 个省(市), 共计两万多个观测站(水文雨量观测站 3730 个、气象国家站 748 个、气象区域站 19 557 个). 充分利用这些站点, 将水文和气象部门的观测数据进行融合, 通过一个平台进行统一展示, 对长江流域监测预报预警、防汛抗旱具有重要的意义. 本文在对气象和水文站点筛选的基础上, 建立降水资料的实时汇集流程, 设计数据处理业务系统以及数据产品制作和共享系统, 实现降水

数据的融合共享.

1 系统总体设计

1.1 系统设计思想

长江流域气象水文雨量资料一张图分析系统主要是依托气象部门和水文部门现有的观测站, 收集台站参数信息, 根据测站的空间分布、数据完整性以及质量指标筛选出需要共享的观测站, 利用已建成的通信线路实现数据的实时收集汇聚, 收集的数据经过质量

① 基金项目: 公益性行业(气象)科研专项(GYHY201206031)

Foundation item: Special R&D Fund for Public Welfare Industry (Meteorology)(GYHY201206031)

收稿时间: 2018-07-09; 修改时间: 2018-08-09, 2018-08-15; 采用时间: 2018-08-21; csa 在线出版时间: 2018-12-07

控制和加工处理后存储于数据库中,采用 B/S 架构开发数据产品制作和共享系统,实现质控后降水资料的显示和共享.

由于长江流域范围观测站较多,数据量比较大,因此在数据存储管理方面考虑采用 HBase 分布式数据库和 Spark 内存数据库,提高数据读写性能;针对查询可能出现缓慢的情况,在 WebGIS 上进行优化.最终为用户提供一个雨量数据统一、页面响应快捷、数据质量可靠的共享系统,为长江流域监测预警、防汛抗旱提供数据支撑.

1.2 系统总体架构

根据气象、水文部门的实际情况,设计了长江流域气象水文数据融合一张图的解决方案,总体架构如图 1 所示,主要分为水文数据信息收集、气象水文大数据存储管理、数据产品制作和共享三部分.

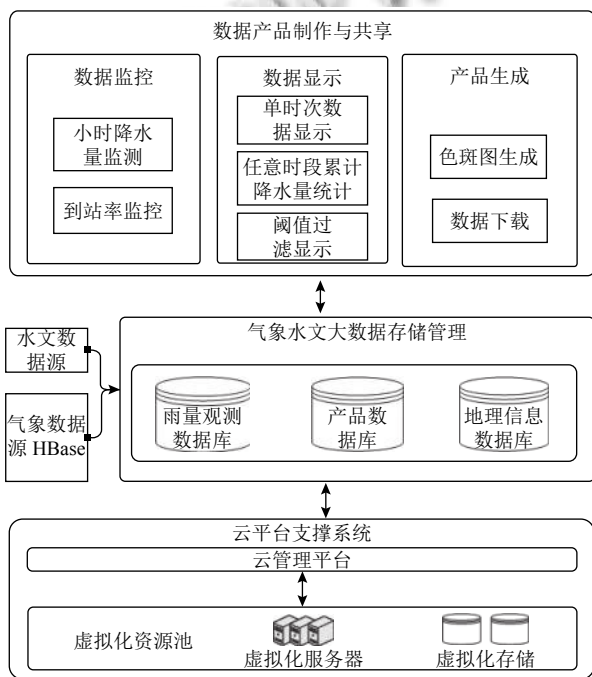


图 1 长江流域气象水文雨量共享分析系统架构图

(1) 水文数据信息收集,收集气象和水文部门的台站参数信息,包括台站的名称、经度、纬度、拔高等内容.通过触发器以及轮询的方式,实现水情信息的实时同步收集.

(2) 气象水文大数据存储管理,通过 Web Service 接口,实时接收水文部门小时降水数据文件,并启动入库模块.结合气象部门台站的降水数据,进行空间一致性检查、持续性检查.根据数据质量控制结果,对观测

数据标注质量控制码,通过 Music 接口(气象数据统一服务接口),提供数据获取、回存、计算、可视化等服务接口,满足用户需要的高效、规范的检索.

(3) 数据产品制作和共享,基于 B/S 架构,开发数据产品加工、数据显示和共享系统,实现长江流域气象与水文部门所有台站的质量控制后降水数据的显示、数据产品的制作与共享.

1.3 关键技术

1.3.1 气象数据分布式存储技术

数据存储管理方面采用 HBase 分布式数据库和 Spark 内存数据库,提高数据读写性能. HBase 分布式数据库与传统数据库不同,它是非关系型的开源数据库,基于 HDFS 文件系统运行^[1].数据采集部分基于 Kafka 开发,消息生产者(RProducer)部署在传感器数据接收端服务器,发送消息到消息队列;消息消费者(RCustomer)部署在服务器端,接收 Kafka 消息并转换成可用格式存入 HBase. Spark 是基于内存的迭代计算框架,其核心是弹性分布式数据集,复杂的算法和运算都是在内存中进行多次快速迭代^[2,3]. Spark Streaming 针对实时数据进行流式处理,具有高通量和容错性^[4],对 Kafka 消息进行操作处理,并将结果保存到 HBase 数据库.

1.3.2 数据共享交换管理

“一张图”共享数据库服务器(气象)是水文、气象数据汇聚的枢纽.而传统数据库同步方式极为繁琐,需要对数据库做诸多配置,并对源数据库与目的数据库有限制和要求,且传输实时性、稳定性均有较大瓶颈.为实现数据的实时同步,采用了 DIP 工具,依托物理机或虚拟机的方式部署,与源和目标端数据相连,原理等同于 Oracle ogg,但无需在前置机上部署任何 Agent.

此次数据交换共享的实施,主要分为两个步骤:一是从实时水雨情数据库服务器到“一张图”共享数据库服务器(气象侧)的同步,源数据库与目标数据库分别为 Oracle.二是“一张图”共享数据库服务器(气象侧)到“一张图”共享数据库服务器(水利侧)的同步,源数据库与目标数据库均为 Oracle.简而言之,本次的数据交换共享实现就是从 Oracle 到 Oracle 的同步过程.原理如图 2 所示.

1.3.3 应用平台显示优化技术

在 WebGIS 上进行优化,一是增加一个缓存机制,定时获取数据,每次收到页面请求时直接从缓存中获取区域站的数据,这样可以将查询和结果返回时间大

大缩短,二是优化前端站点绘制速度,用绘制 SVG 矢量图形替换原来的图片叠加方式,三是采用分级策略显示站点,逐级增大显示站点数^[5].

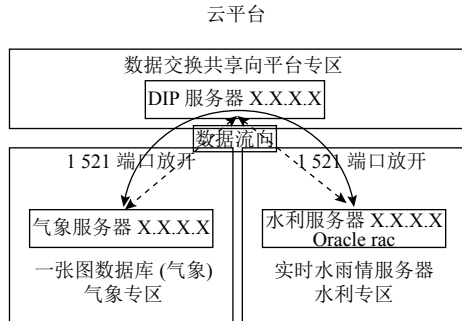


图2 “一张图”共享数据库服务器数据同步原理图(气象端)

2 主要功能设计

2.1 水文数据信息收集功能

结合目前实时雨水情数据库的使用情况,为了最大限度的减少对现有系统的影响,同时能够保证数据交换

的时效性,系统通过触发器,在水文部门的数据库平台内将发生变动(新增、修改或删除)的数据写入到待交换区,通过轮询方式监控交换区是否有新的数据,如果有新的记录将数据发送给气象部门的数据接收节点,从而实现水情信息的实时交换收集,数据交换处理过程如图3.

(1) 水利部门将实时雨水情数据库中所管辖的站号信息类、雨量、水文的数据进行插入(修改、删除)操作时,启动数据库表的触发器;

(2) 将变化的数据保存到待交换区中;

(3) 系统定时对待交换信息记录表进行轮询检查,发现有待发送数据后,根据系统设置的转发关系,生成相应的数据文件,并启动发送模块;

(4) 发送模块负责把数据文件,分别发送给各接收节点;

(5) 接收节点通过 Web Service 接口,接收上传的文件,并启动入库模块;

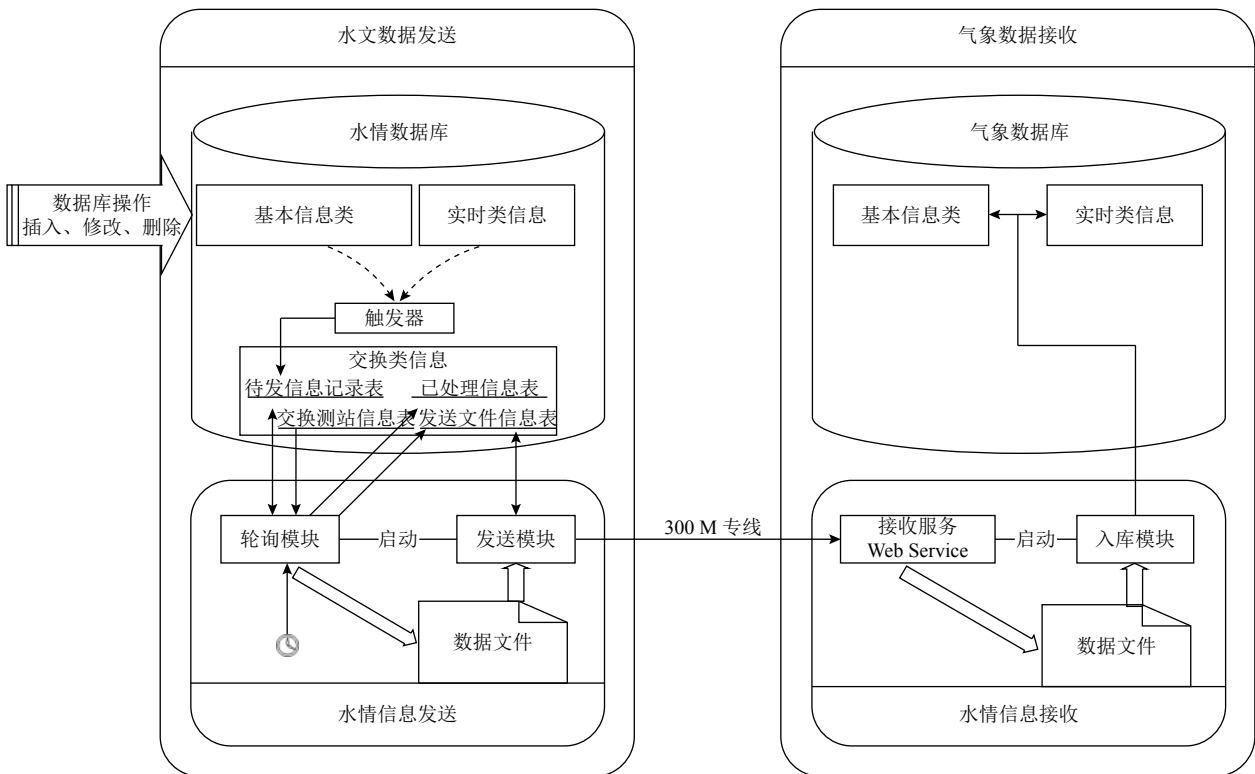


图3 水文数据信息收集处理过程图

(6) 入库模块负责解读数据文件,并把数据记录插入到已接收信息记录表中,同时插入(修改、删除)相应的基类、实时类信息.

2.2 气象水文大数据存储管理

长江流域气象、水文观测站多、频次密,对高并发需求数据,采用内存数据库技术(Spark);对非高并发

数据,采用非关系型数据库(HBase),存储管理结构如图4所示。通过五台服务器及万兆交换机构建集群数据库环境,利用云分布式存储技术,采用数据主键、分块和数据压缩相结合的方式提高大数据的存储能力,实现降水数据要素级的数据库存储管理。

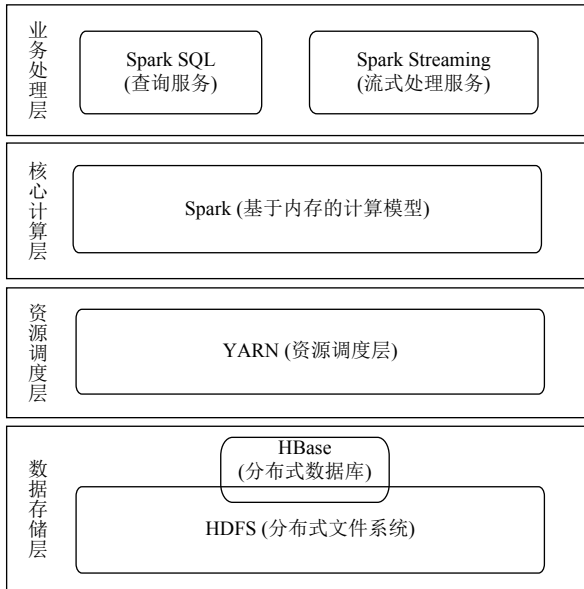


图4 气象水文大数据存储管理结构图

据统计,使用HBase存储45天数据,数据量达到280GB,其中区域站173415988条记录,国家站12303617条数据,湖北省6000多个站点雨量查询显示在3s以内,形成矢量图片在20s以内;长江流域20000多个站点雨量查询显示在10s左右,形成矢量图片在25秒左右;长江、汉江流域的水位、流量、水势等数据能够每小时正点实时显示。相比传统关系型数据库存储显示速度有明显提升。

2.3 数据融合质量控制

根据降水观测数据的特点,参考天气因素等条件设计以下质量控制方法,包括界限值检查、范围值检查、空间一致性检查、持续性检查、累计降水量秩相关系数检查、综合检查等方法^[6-8]。

2.3.1 界限值检查方法

气象部门《地面气象观测规范》中对每一个要素的界限值范围进行了规定,可以根据该范围判断降水资料质量情况。

2.3.2 范围值检查方法

范围值检查是通过插值,计算某一地点的多年日极值,根据日数据变化规律,求得降水资料的逐月阈值。

2.3.3 空间一致性检查

空间一致性检查方法是检查观测数据的变化是否符合其空间规律,利用与被检站下垫面及周围环境相识的一个或多个邻近站观测数据计算被检站降水量,对被检站观测值和计算值进行比较,比较结果超出给定的阈值,即认为被检站降水观测数据为可疑资料。

2.3.4 持续性检查

如果某个要素的观测值在一段时间内没有变化,则有可能观测设备或者数据传输出现问题,在降水资料的持续性检查中,若降水量长期持续不变(有降水现象时,即降水量大于0),被检站在被检时段内的小时降水观测数据均为可疑资料。

2.3.5 累计降水量秩相关系数检查

根据降水现象的特点,在时间、空间分布上对被检站与邻近站滑动过去6小时累计降水量进行比较,比较结果超出给定的阈值,即认为被检站降水观测数据为可疑资料。

2.3.6 综合一致性检查等方法

综合一致性检查是根据以上所有检查方法得出的结论来确定该类数据最终的质控码。

2.4 数据产品制作和共享

数据产品制作和共享功能^[9,10]包括数据监控、数据显示和产品生成三个功能(见图5)。数据监控包括到站率监控和小时降水量监测;数据显示功能实现了降水量的查询、显示和统计,具体包括单时次数据显示、任意时段累计降水量统计、阈值过滤显示、数据表格显示、单站曲线显示、色斑图显示和疑误数据显示等;产品生成是指对查询、统计结果进行保存和下载,包括图片保存和Excel格式下载。

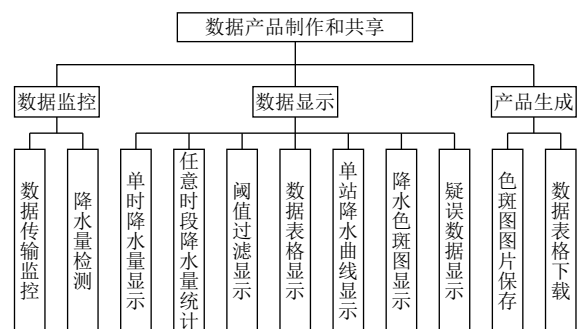


图5 数据产品制作和共享系统功能结构图

2.4.1 数据监控

(1) 数据传输监控

数据传输监控模块逐小时从数据库中统计已到资

料的总站数,并统计到站率.在GIS页面上显示应到站、实到站和到站率,并将已到站按绿色、未到站按红色进行标示.

(2) 降水量监测及对比

降水量监测功能是系统每小时自动统计所有站点的小时降水量,自动统计出大于设定值的站数,并显示最大降水量和所在站点.设定值可以根据用户进行修改.统计的时间将按照到站率和时间策略进行综合设计.在显示页面上,以气象、水文两个复选框给出,用户根据需要可以形成三个组合,即气象与水文单独、气象与水文合并的显示方式,做到雨量的对比融合显示.

2.4.2 数据产品制作

(1) 色斑图生成

提供不同时间段累计降水量的色斑图制作功能,生成图形产品.

(2) 面雨量统计

按流域,统计不同时间段累计面雨量.

2.4.3 数据显示

(1) 单时次降水量显示

根据用户选择,单时次降水量显示功能在WebGIS页面上所有站点标示附近显示站点的过去1小时(3小时、6小时、12小时、24小时)降水量,用户可通过放大地图直观的查看区域内的降水量情况.

(2) 任意时段降水量统计

任意时段降水量统计功能是查询开始时间和结束时间内所有站点的累计雨量,并在WebGIS页面上显示站点的累计雨量值,用户可通过放大地图直观的查看区域内的降水量情况.

(3) 按阈值过滤显示

对查询结果提供阈值过滤显示功能,在阈值文本框中输入数值,滤掉降水小于阈值的站点,只显示值大于等于阈值的站点,并在文本框后显示站点数目.

(4) 数据表格显示

所有查询结果除以WebGIS方式叠加在地图上展示,也提供数据表格显示,并提供数据下载.

(5) 单站降水曲线显示

提供单站降水曲线显示功能,单击站点可查看过去24小时降水曲线图或柱状图,并可保存图片到本地.

(6) 色斑图实时显示

对降水量查询结果进行插值算法,并根据指定图例生成色斑图.

(7) 面雨量显示

显示各流域不同时间段累计面雨量.

3 系统应用

3.1 系统应用

长江流域气象水文雨量资料一张图分析系统(见图6),部署在湖北省气象局,水文和气象部门的用户可以通过数据库客户端获取数据,也可以通过浏览器查询、分析雨情,雨量资料的对比,做到了气象、水利部门数据的一致性,提高了数据发布的权威性.系统主要包括数据监控、数据显示和产品生产三大功能.其中,数据监控包括到站率监控和小时雨量监测;数据显示功能包括雨量时查询、显示和统计,具体包括单时次数数据显示、任意时段累计雨量统计、阈值过滤显示、数据表格显示、单站曲线显示、色板图显示和疑误数据显示等;产品生成可对查询、统计结果进行保存和下载.



图6 长江流域气象水文雨量资料一张图分析系统

为提高产品应用的便捷性,增强平台对业务的支撑度,还开发了按流域、行政区域(地市级、县级)查询显示的功能,同时还根据业务需要形成气象与水文雨量显示的三种组合,即气象与水文单独、气象与水文合并的显示方式,实现雨量的对比融合显示.

3.2 效益分析

2017年7月25日,湖北省气象局正式下发了《关于做好湖北气象业务内网水文部门雨量资料产品应用的函》,水文部门的雨量、水文数据正式在全省气象部门得到应用.8月3日和8月10日,“长江流域气象水文雨量资料一张图分析系统”在省水利厅和长江委正式使用,普遍反映良好.在我省今年盛夏8月7-8日、11-13日强降雨过程以及多次局地强对流天气影响期间,为武汉中心气象台更加有效捕捉“坨子

雨”强度、提高预警精度、精准防范发挥了重要作用。长江委、省防办、省水利厅在今年秋汛精准调度中也很好的发挥了作用。下一步将继续优化站点,完善雨量数据质量控制,开发满足水利部门雨量产品和各种极值统计产品,更好的为流域防汛抗旱提供服务。

4 结语

长江流域气象水文雨量共享分析系统构建了集气象、水文雨量融合为一张图的可视气象信息服务平台,主要承担为长江流域气象、长江水利委员会水文部门提供实时融合雨情查询、监视、分析统计。系统基于分布式数据库和 WebGIS 技术进行优化处理,响应速度快捷、功能完善、界面丰富友好,为流域气象服务工作提供了很好的平台,得到了用户的一致肯定。

参考文献

- 褚福银, 张林, 何坤鹏. 基于 Hadoop 平台海量数据的快速查询与实现. 电脑知识与技术, 2016, 12(21): 3-5.
- 雷明涛, 李绪福. 内存数据库与 Oracle 数据库的数据同步设计与实现. 信息化研究, 2015, 41(2): 61-65.
- 王宏伟, 方群, 陈伟. 基于内存数据库的 OPC 监测系统的设计与实现. 计算机技术与发展, 2013, 23(7): 242-244, 248.
- 陈海燕. 一种分布式大数据管理系统的设计与实现. 计算机科学, 2014, 41(11A): 393-395.
- 姜彤, 罗静. 基于 WebGIS 的遵义市水雨情实时监测系统. 人民黄河, 2012, 34(12): 75-77. [doi: 10.3969/j.issn.1000-1379.2012.12.028]
- 任芝花, 赵平, 张强, 等. 适用于全国自动站小时降水资料的质量控制方法. 气象, 2010, 36(7): 123-132.
- 吴书成, 魏爽, 王丽吉, 等. 自动气象站降水实时快速质控算法设计. 气象科技, 2016, 44(6): 882-888. [doi: 10.3969/j.issn.1671-6345.2016.06.005]
- 刘雨佳, 陈洪滨, 金德镇, 等. 加密自动气象站雨量计资料的质量控制及其相关关系的研究. 大气科学, 2014, 38(1): 159-170.
- 华韵子, 邵玲玲, 林红. 长江三角洲气象数据实时共享技术研究. 气象科技, 2014, 42(2): 261-265. [doi: 10.3969/j.issn.1671-6345.2014.02.014]
- 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. [doi: 10.3969/j.issn.1001-7313.2006.05.012]