

DOI: 10.3969/j.issn.1004-4701.2019.01-14

江西水利数据资源整合共享实践

朱松挺, 胡应龙

(江西省防汛信息中心, 江西 南昌 330009)

摘要: 水利数据涉及对象繁多, 持续面向业务建设导致用户和数据自成体系, 出现一数多源、跨业务共享困难、新建业务应用调数据效率低、应用访问入口不一等问题。针对这些问题, 本文提出了分级分部门分权限的水利数据更新维护模式, 构建了基于微服务的一整套共享服务体系, 提供了水利数据从生产到消费的全流程实践参考。

关键词: 水利数据; 整合共享; 微服务

中图分类号: TP311.13 **文献标识码:** B **文章编号:** 1004-4701(2019)01-0072-04

1 背景

全国水利行业及其它行业都在大力开展数据资源整合共享。数据资源整合前, 江西水利信息化建设经过多年的快速发展, 防汛、水资源、水保、农水等业务系统间数据及用户自成体系, 跨业务调用数据困难、同类数据多头采集, 水政执法、河长制管理等新业务系统对其它业务数据的共享需求尤为迫切, 数据、用户、平台的割裂为深入应用融合制造了障碍。为打通系统、数据、应用间的共享壁垒, 解决数据跨部门、跨业务共享的难题^[1], 急需开展水利数据资源整合, 建立一个遵循国家标准和规范的整合共享体系^[2], 实现全省水利信息统一的数据采集、存储、管理、交换、共享等服务^[3], 规范数据资源从生产到消费的全流程, 提高信息资源的利用效率和共享水平。

2 现状及问题分析

2.1 水利信息资源现状

江西水利信息资源包括水利普查数据、防洪工程

数据、业务系统数据、系统用户数据等, 共有 40 类约 18 万个对象数据, 各类业务监测数据记录过亿。地理空间数据存有各类基础矢量数据 70GB、DEM 数据 80GB、各类影像数据约 2.5TB, 各类专题涉及 45 类对象超过 150 个服务。水利各业务应用持续建设, 活跃业务应用及基础应用系统达 60 余个, 涉及到的各级用户过万。数据种类繁多, 各类数据存在大量冗余, 各系统间用户信息没有打通, 给用户认证带来了麻烦。

2.2 数据共享现状

在数据资源整合并建立统一的共享平台之前, 信息共享存在多种方式: ①一次性复制带走; ②通过系统或网站提供信息查询; ③通过接口服务提供给第三方应用调用; ④通过交换软件进行数据传输; ⑤直接连数据库拉取; ⑥根据需求订阅推送等等^[4]。这些共享方式在水利各应用间或独立存在或几种并存, 数据交叉传输导致数据一致性及传输可靠性都不可控。需要建立统一的数据共享系统、接口服务、数据交换以解决这些问题。

2.3 数据环境现状

整合共享还需要计算、网络、存储、安全、标准等环境基础支撑。已建成的全覆盖省、市、县、乡四级水利网

收稿日期: 2018-11-26

作者简介: 朱松挺(1982-), 男, 硕士, 工程师。

络,保证监测信息及时入网,为各业务信息的生产、传输、共享提供了通道。但数据计算及存储资源没有统一管理,不利于硬件资源的共享利用,基于此提出了水利云整合计算与存储资源,“云平台+微服务”的运营模式为信息共享提供硬件环境支撑^[5]。

3 整合共享关键技术

3.1 数据整合技术

数据整合在充分调研分析各信息系统数据库基础上,采取数据抽取方式将各系统基础数据抽取到数据中心数据库。数据抽取过程中优先选用生产库数据,同时抽取水利普查数据和防洪工程数据作为补充。

数据抽取工具采用 kettle-spoon,Kettle 是一款开源的 ETL 工具,可以在 Window、Linux、Unix 多环境中运行,数据抽取高效稳定。数据抽取过程分四个阶段:①对象模型的录入与建设;②数据源的调研与抽取;③原始数据的校验与转化;④生成对象数据库^[6]。抽取清洗入库流程如图 1 所示。

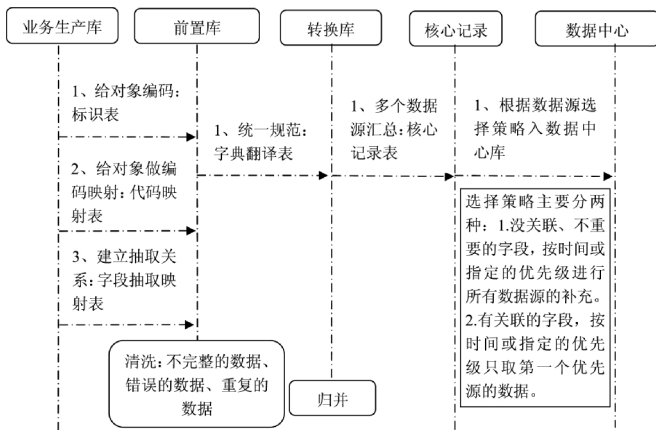


图 1 数据抽取清洗入库流程图

3.2 微服务技术

微服务对原有庞大的单体系统进行模块化拆分,进行轻量级独立开发部署。技术的使用是为了解决系统扩展性差、复杂度高、部署慢、资源使用弹性等问题。其技术架构如图 2 所示。

系统整体按照 SpringCloud 技术栈作为微服务形式进行开发,其中使用到的组件有 Eureka 注册中心、

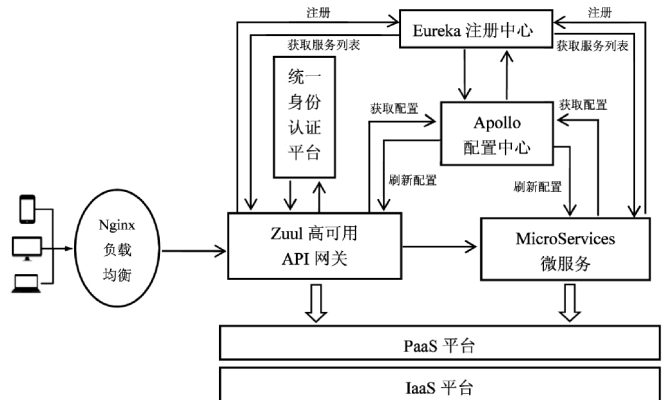


图 2 微服务技术架构图

Apollo 配置中心、Zuul 网关等。注册中心采用 Spring Cloud Netflix Eureka。当集群模式中有分片发生故障的时候,Eureka 会自动转入自我保护模式。集群中的不同服务注册中心通过异步模式互相复制各自的状态,保证正确获取服务。Apollo 能够集中化管理应用不同环境、不同集群的配置,配置修改后能够实时推送到应用端,并且具备规范的权限、流程治理等特性。API 网关采用 Spring Cloud Netflix Zuul,微服务场景下,每一个微服务对外暴露了一组细粒度的服务。Zuul 组件可以做反向代理的功能,通过路由寻址将请求转发到后端的粗粒度服务上,并做一些通用的逻辑处理。在 API 网关中接入省水利厅内部统一身份认证平台,确保人员部门等数据的一致性。

4 整合共享方法

4.1 整合共享体系架构

水利数据资源整合共享按照统一数据库、统一身份、统一门户的思路,目标是实现“一数之源”、“融合共享”。在统一网络、云平台、统一标准、统一安全基础上进行数据整合共享,采用物理集中与逻辑集中相结合的思路建立统一的库模型,整编入库数据,建立统一身份认证、支撑服务体系、统一门户平台。其整体架构体系如图 3 所示。

4.2 库模型整合方法

库模型的构建基于面向对象思想,建立水利对象基础信息库模型。在雨水情库、水资源库等国家专业业

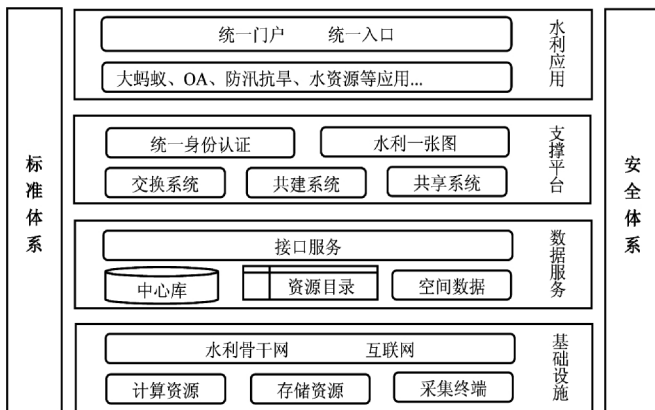


图3 体系架构图

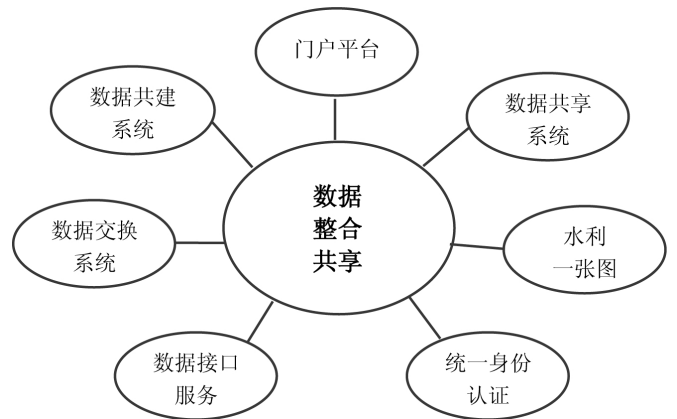


图4 数据整合共享服务体系结构图

务库模型上统一业务库,与对象基础库建立关联关系。库模型根据业务接入和使用的需要不断调整优化,持续迭代的。对象基础库与新建业务库物理集中在中心库管理,对早于中心库建设的成熟业务库采用逻辑集中的方式纳入中心库管理。

4.3 信息整合共享

作为各系统的身份认证支撑的用户信息,以即时通信系统中的人员信息为主,整合形成人员(用户)信息库。开发水利人员信息管理系统进行信息维护,维护职责分级分部门划分,各部门管理员保障本部门人员信息的正确性与及时性。人员信息通过接口方式共享,目前统一身份认证系统、大蚂蚁系统、通讯录系统均通过该系统获取统一的人员信息。

水利工程及业务信息整编工作分为元数据与资源目录、水利对象基础数据、业务数据3类。经过梳理整编,形成资源目录与元数据并入库,为数据共享提供支撑。在水利普查与防洪工程数据的基础上完成水利对象基础数据的整编,建立统一的对象数据库,为业务应用提供数据共享。

5 整合共享实践与应用

通过对整合共享方法研究及技术应用,以微服务方式建立数据接口服务、数据共建系统、数据共享系统等应用。其整个应用体系包括数据接口服务、统一身份认证、数据共建系统、数据共享系统、水利一张图、门户平台。体系结构如图4所示。

(1)数据接口服务提供接口的注册、管理、监控,接口通过统一的域名对外发布。应用与库之间通过接口进行分离,接口对应用保持不变,对库根据模型调整对应调整,灵活配置。

(2)在统一的人员信息库基础上,搭建了全省统一身份认证平台,实现用户集中管理、访问控制、系统单点登陆,保证了资源的融合共享。

(3)数据共建系统提供各类基础数据及业务数据的填报维护功能,根据业务需要灵活配置填报模板和任务。对于数据库中需更新数据、错误数据或缺失数据,建立数据更新维护机制,分级分部门分权限具体落实各类数据更新维护部门,明确责任,通过数据共建系统统一更新维护。各业务系统能通过接口直接更新数据。

(4)数据共享系统包括资源目录、数据资源查阅、数据资源调用、数据订阅推送等;解决有哪些资源、怎么获取的问题。共享内容有基础工程数据、实时监测数据、API接口、文档,逐步加入视频、计算、存储、网络等资源的共享。

(5)水利数据资源具有空间强关联性,能够直观通过地图平台进行分析决策与展示,为各系统的建设提供良好的支撑。水利一张图提供全省基础地图、水利专题数据、影像数据的服务,为防汛指挥等水利业务系统提供服务。

(6)数据交换系统配合共建共享系统解决数据的交换共享问题,特别是跨层级、跨部门的实时业务数据的共享,例如雨水情数据的汇集、气象数据的交换等。

(7) 门户平台作为信息共享窗口,已打通多系统的用户身份,以即时通讯应用作为统一入口,身份自动识别,集中展示提供水利信息服务。为解决移动端的使用需求,参照PC门户内容开发了移动门户。门户建立了统一的接入标准,支持接入各水利业务系统,已接入OA、短信、邮件、网盘、水利一张图、防汛指挥系统等,用户可根据自身需求定义所需访问的业务模块。

6 整合共享成效

数据整合共完成40类数据整编,入库数据对象名录超过18万,各类业务数据记录以亿计。这些数据已通过统一共享平台提供给防汛抗旱、河湖管理、工程标准化等业务系统使用。

数据整合共享体系的建设以业务需求为导向,以实用为目标,以对象整合为纲目,及时迭代改进,灵活应对各业务应用系统支撑需求。通过划清责任,制定责任清单,分级分部门分权限实施数据更新维护,按各业务需求自行定制共享,库标准建设统一由中心把关,有

效解决了重复建设、数出多源、共享困难等问题。

数据资源整合共享建设为新一代信息技术在水利的深度应用,实现与水利应用的深度融合,提高省水利治理与管理的工作水平和协同能力,可为各类政务数据的共享提供切实可行的解决方案。

参考文献:

- [1] 江西省人民政府公报. 关于加快推进全省政务数据共享的工作方案[Z]. 2018-09-29.
- [2] 水利部信息化工作领导小组办公室. 水利信息化标准指南[Z]. 2003-06.
- [3] 宋蕊,曾剑秋. 物联网应用的发展状况及建设“感知中国”的策略探讨[A]. 两化融合与物联网发展学术研讨会论文集[C]. 2010年.
- [4] 胡应龙,朱松挺,陈杰. 水利数据共享推送平台设计与实现[J]. 江西水利科技, 2015(6): 463~465.
- [5] 郭栋,王伟,曾国荪. 一种基于微服务架构的新型云件PaaS平台[J]. 信息安全, 2015(11): 15~19.
- [6] 裴莹. ETL中数据清洗方法研究与实现[D]. 中国科学院大学, 2010.

编辑: 张绍付

Practice of integration and sharing of water resources data in Jiangxi province

ZHU Songting, HU Yinglong

(Jiangxi Flood Control Information Center, Nanchang 330009, China)

Abstract: Water resources data involves a wide range of objects. Continuous business-oriented software leads to having one's own system of user and data. There are many problems, such as data from multisources, difficulties in sharing, especially for new applications, and different entrances for applications. In order to solve these problems, this paper proposes a data maintenance model with areas, departments and authorities, and constructs a whole set of sharing system based on Microservices, which provides a reference for the whole process of data from production to consumption.

Keywords: Water resources data; Integration and sharing; Microservices

翻译: 朱松挺