

西南诸河流域水利信息化综合管理平台设计研究

段仙琼¹, 金晨曦², 张阳帆¹, 甘坤¹, 龚德林¹, 白思哲¹

¹长江水利委员会水文局西南诸河水文水资源勘测局, 云南 昆明

²云南省水文水资源勘测局, 云南 昆明

收稿日期: 2022年4月24日; 录用日期: 2022年6月13日; 发布日期: 2022年6月29日

摘要

随着国家对加快水利信息化升级转型的不断推进, 以及天空地一体化水利感知网的建立, 水利智能感知水平提升和水利大数据的建立成为必然趋势。水利大数据时代的到来使其数据量容量发生指数增长, 对数据获取方式、数据类型、数据管理方式、空间分析能力等提出了新的要求。研究基于水利大数据技术阐述了水利大数据的基本概念, 结合西南诸河流域的水利数据资源情况, 从水利数据资源池的建立、水利大数据运维管理体系设计、基于Hadoop分布式数据存储模式的平台架构设计等方面进行论述, 为平台应用落地提供技术保障。

关键词

大数据, 水利信息化, Hadoop

Research on the Design of Water Information Integrated Management Platform of Southwest Rivers

Xianqiong Duan¹, Chenxi Jin², Yangfan Zhang¹, Kun Gan¹, Delin Gong¹, Sizhe Bai¹

¹Bureau of Hydrology and Water Resources Survey of Southwest Rivers, CWRC, Kunming Yunnan

²Bureau of Hydrology and Water Resources of Yunnan Providence, Kunming Yunnan

Received: Apr. 24th, 2022; accepted: Jun. 13th, 2022; published: Jun. 29th, 2022

Abstract

With the continuous promotion of to accelerate the upgrading and transformation of water informatization and the establishment of the sky earth integrated water conservancy perception network, the im-

作者简介: 段仙琼(1990-), 女, 工程师, 硕士, 研究方向: 水环境及水资源, Email: 921546732@qq.com

文章引用: 段仙琼, 金晨曦, 张阳帆, 甘坤, 龚德林, 白思哲. 西南诸河流域水利信息化综合管理平台设计研究[J]. 水资源研究, 2022, 11(3): 320-327. DOI: 10.12677/jwrr.2022.113035

provement of water conservancy intelligent perception level and the establishment of water conservancy big data have become an inevitable trend. The arrival of the era of big data in water conservancy makes its data volume and capacity grow exponentially, which puts forward new requirements for data acquisition methods, data types, data management methods, spatial analysis capabilities, etc. Based on the water conservancy big data technology, this paper expounds the basic concept of water conservancy big data. Combined with the water conservancy data resources of the southwest rivers of China, it discusses the establishment of water conservancy data resource pool, the design of water conservancy big data operation and maintenance management system, and the design of platform architecture based on Hadoop distributed data storage mode, so as to provide technical support for the implementation of platform applications.

Keywords

Big Data, Water Informatization, Hadoop

Copyright © 2022 by author(s) and Wuhan University.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

西南诸河位于我国的西南边陲，地貌类型气候条件复杂，水能资源丰富，开发条件优渥，而且涉及多条国际河流，包括雅鲁藏布江、怒江、伊洛瓦底江、澜沧江等国际河流。其流经地域广阔，使用传统的结构化数据及 GIS 分析愈来愈难以满足对流域立体化感知、监视、管理及指导开发利用的需求，对水利信息数据的感知、汇集、存储、管理及深度挖掘分析提出了更高的要求。

随着水文现代化建设逐步向监测站网立体化、监测手段自动化、信息采集自动化、数据处理智能化、服务产品多样化的“五化”要求水平提升。充分利用区块链、大数据、云计算、人工智能等先进技术，建设监测评价和趋势分析、预报预警智能化综合信息系统，不断推进水文现代化是水利行业发展的大势所趋。

水文数据体量庞大，结构复杂，形式多样化，在短时间内以常规手段难以获取、处理与分析，而大数据正是针对这一类数据衍生的一种新型数据处理技术。

2. 水利大数据应用现状

2.1. 水利大数据应用技术概况

随着社会经济的快速发展和人民生活水平的逐步提高，水资源、水生态、水环境、水旱灾害的问题也逐步凸显。根据习近平总书记提出的“节水优先、空间均衡、系统治理、两手发力”的十六字治水思路，以及网络强国、数字中国、智慧社会的智能化发展趋势。根据李国英部长提出的推动水利高质量发展的要求，要推进智慧水利建设，以数字化、网络化、智能化为主线，构建数字孪生流域，开展智慧化模拟，支撑精准化决策，全面推进算据、算法、算力建设，加快构建具有预报、预警、预演、预案功能的智慧水利体系。

2013 年 Nature 提出利用大数据技术，实现水资源可持续发展目标[1]。2019 年智慧水利(Smart Water Conservancy)的概念也在我国水利行业逐步兴起。智慧水利是指应用云计算、物联网、大数据、移动互联网和人工智能等新一代信息技术，对水利对象，如：河流、湖泊、水库、水电站、水闸、监测站和取水口等，进行透彻感知、网络互联、广泛共享、智能分析和泛在服务，为水旱灾害防范与抵御、水资源开发与配置、水环境监管与保护、河湖生态监督与管理等水利业务提供现代化支撑的新型技术。智慧水利的发展使水利行业发展与大数据深度融合成为必然[2]。2022 年 1 月，水利部印发关于推进智慧水利建设的指导意见和实施方案。意见指出

要建设水利智能业务应用体系，水利网络安全体系，智慧水利保障体系。

过去大数据平台一般使用基于集中式存储的虚拟化平台提供各系统的存储、计算、运行环境，存在重量大、启动慢、性能损耗较大等问题。西南地区地域广阔，近年来基于 GIS、RS 以及卫星影像等技术的非结构化数据积累日渐完善，过去基于集中式存储的扩展性、可靠性、性价比不再能满足水利大数据日益增长的存储、计算需求。

2.2. 水利大数据的定义

在维克托·迈尔-舍恩伯格编写的《大数据时代》[3]中，大数据指不用随机分析法这样的捷径，而采用所有数据进行分析处理。大数据是一种规模大到在获取、存储、管理、分析方面大大超出了传统数据库软件工具能力范围的数据集合。大数据具有海量的数据规模、快速的数据流转、多样的数据类型和价值密度低四大特征。常常包括结构化、半结构化和非结构化数据，非结构化数据构成了数据的主要部分。大数据分析常和云计算联系在一起，因为实时的大型数据集分析需要像 MapReduce 一样的框架来向数十、数百或甚至数千的电脑分配工作。

随着建设智慧水利方针的逐步落实，我国水利正在向着构建天空地一体化水利感知网方向努力。通过采用卫星、雷达、高清视频监控、无人机、遥控船、物联网通信等新型监测技术，水利数据已经达到了大数据量级[4][5][6]。水利大数据具备一般大数据的基本特征，同时还具备水利领域特征，如广泛交互性、时空融合性、高效能性、深共情性[7]。大数据创造价值的过程被不同的学者提炼成不同的价值链体系，有三环价值链体系[8]，四环价值链体系[9]，五环价值链体系[10]。结合大数据五环价值链体系，蒋云钟等人将大数据价值链体系提炼成：数据采集、数据存储、数据计算、数据分析、数据可视等环节。

2.3. 水利大数据的存储模式

水利数据结构多样化，包括了结构型，半结构性，非结构性数据。水利大数据的存储模式主要包括：分布式文件系统、分布式数据库和关系型数据库，其中关系型数据库又包含了传统关系型数据库和分布式并行数据库。水利大数据系统中结构型数据主要存储于关系型数据库中，如实时雨水情数据、基础水文测验数据、水质分析数据等都属于结构型数据[11]。关系型数据库是水利大数据存储结构的重要组成部分，且广泛服务于智慧水文的各个领域，如国家防汛抗旱指挥系统[12]，全国水利一张图系统等[13]。半结构型和结构型数据主要采用分布式文件系统进行存储，其中 HDFS 分布式文件系统随着 Hadoop 一起出现并逐步发展成熟，因其可移植性、高容错性、可大规模水平扩展性[14]，HDFS 作为开源分布式文件系统被各大数据领域广泛采用[15]。

在设计大数据环境的存储模式时，吴飞将数据的存储分成了三层模式：操作数据存储层、数据仓库层、分析结果存储层[16]。操作性存储层主要使用关系型数据库为数据仓库层提供数据源，保证仓库层的稳定，同时筛选不符合要求的数据。数据仓库层使用 Hive 进行存储，将汇总的数据按照不同数据类型进行分别存储。分析结果存储整合了关系型和非关系型数据库存储模式，为上层应用访问提供数据。

2.4. 水利大数据的计算、分析模式

根据水利大数据特征的不同，计算方式包括：流式计算、批量式计算、内存计算和图计算[17]。水利领域的计算目前以 MapReduce 批量计算为主，该计算方式被广泛应用于卫星遥感数据处理、旱情分析、降雨等级预测等领域，流式计算包括 Spark Streaming 流计算[18]，Storm 降雨径流预测计算等[19]。

水利大数据分析是指采用相应的分析方法，如数据挖掘、统计分析和新兴分析等方法，利用大数据模型或专业模型提取大数据中隐藏的信息，作为水利事件提供决策、预测的参考依据。其中数据挖掘技术在水利分析中应用比较广泛，例如暴雨时空分布规律，需水量预测，流域水质空间评估等[20]。数据挖掘还可实现对文本文件的挖掘，使半结构化和结构化水利数据中隐含的新的水利价值信息得以全面揭示。

3. 水利信息化综合管理平台开发

利用大数据技术相，开发一套可以实现西南诸河水利水资源数据共享、计算机自动作业、智能处理、服务产品丰富多样、服务手段便捷先进的水利信息综合管理平台是本次研究的重点。

3.1. 西南诸河水利大数据体系

西南诸河流域包含了雅鲁藏布江、伊洛瓦底江、澜沧江、怒江、藏南藏西诸河等水系，西南诸河流域水利大数据包含的数据类型多样，数据来源广泛，数据结构类型复杂。建立涵盖洪水、干旱、水利工程安全运行、水利工程建设、水资源开发利用、城乡供水、节水、江河湖泊、水土流失和水利监督等业务需求的水利大数据平台是水利信息化升级转型，实现智慧水利建设的总体目标。西南诸河流域水利大数据体系如图 1 所示：包含了水文水资源数据、水利工程数据、空间地理数据、气象数据、生态环境数据及其他非结构化，半结构化数据。在大量水利数据整合的基础上，根据国家、行业标准建立水利数据资源池，建立水利数据资源目录，形成数据资源汇聚体系，建立多层次的数据管理体系。



Figure 1. Southwest rivers big data system of water conservancy
图 1. 西南诸河水利大数据体系

3.2. 西南诸河流域水利大数据运维管理体系

西南诸河流域水利大数据体系汇集了体量庞大，结构多样的各类水利数据，为了实现数据的有效管理利用和确保数据安全得以保障，需要加强对水利大数据的综合运维。传统的数据运维方式无法满足水利大数据的运维需求，需开发建立更加自动化、智能化、流程化的数据管理模式。

西南诸河流域水利大数据运维管理体系流程如图 2 所示。该体系根据分布式文件 HDFS 为依据，运用 Hive、Presto、Kafka 大数据插件优化水利数据管理体系的各项功能。水利大数据运维管理体系包含了数据集成层、数据存储层、数据计算层、数据应用层。

1) 数据集成层：水利大数据结构化多样，数据集成层能将数据运行监控平台、数据运维管理系统、ZABBIX 系统中的各类数据进行清洗、检验、管理及规范化。Hive、Presto、Kafka、Flume 数据插件实现各类实时结构化数据、非实时结构化数据、非结构化数据的集成。数据集成管理方便后期数据的应用及计算，可对数据质量进行判断，发现错误及异常值，并对数据进行优化调整。

2) 数据存储层：使用 HDFS 数据存储模式，该模式具有存储模式扩展性强、运行环境安全可靠、数据吞吐

量大的优点。HDFS 用于非结构化数据存储，能够对大型文件进行访问，并进行错误节点处理。数据仓储使用 Hive 做结构化数据存储媒介，可建立定向分析的数据集成化运行环境，便于进行决策性工作。使用 Kafka 插件可实现实时数据的提取与存储工作。

3) 数据计算层：数据计算层提供数据处理运行引擎及数据计算功能，对离线数据进行批量计算，对在线实时数据进行在线计算。

4) 数据应用层：该层包含了数据使用申请管理，数据访问管理，资源管理。资源管理需要对源数据及分析计算结果数据进行设置，数据访问管理可对账号进行授权管理。

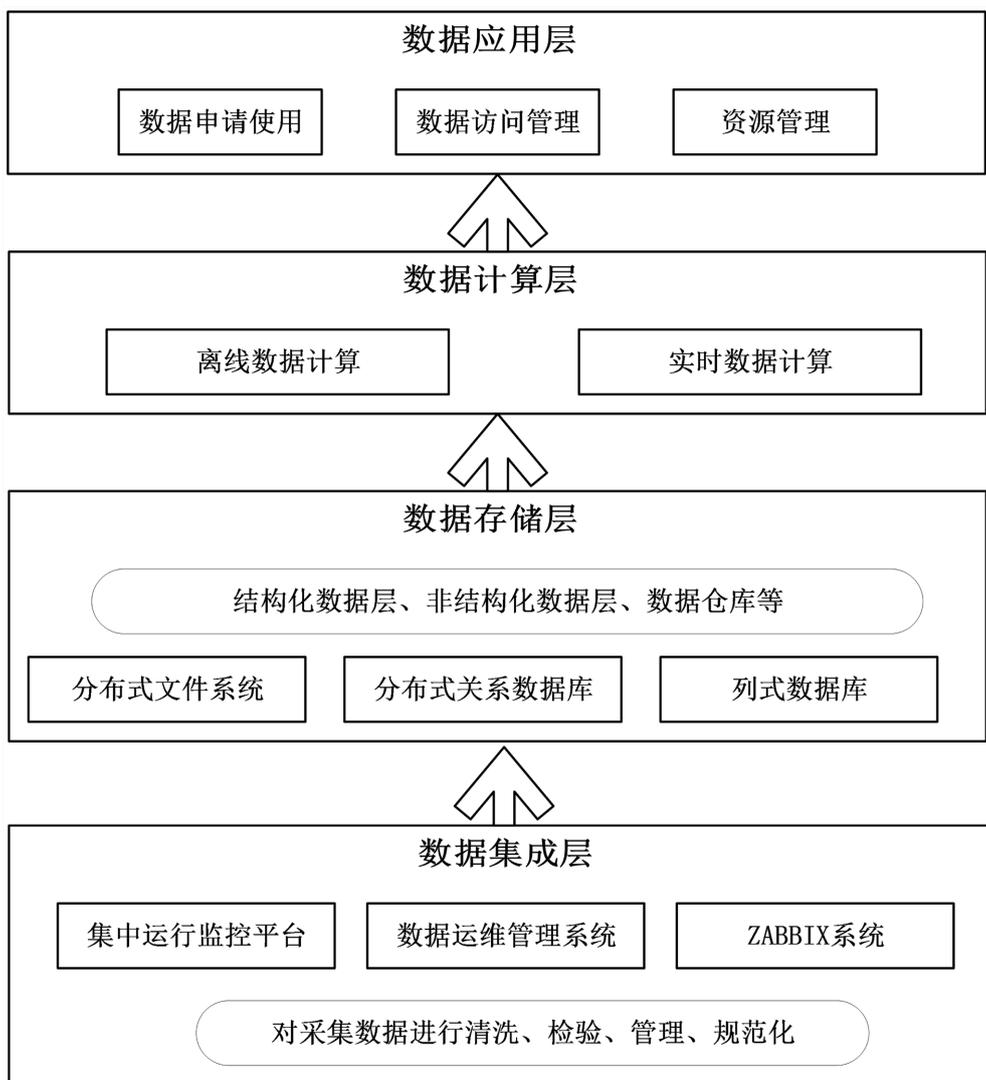


Figure 2. Southwest rivers water conservancy big data operation and maintenance management system
图 2. 西南诸河流域水利大数据运维管理体系

3.3. 西南诸河流域水利信息化综合管理平台体系架构

西南诸河流域水利信息化大数据综合管理平台以服务器作为硬件支撑部署数据接受、存贮、分析处理、应用等数据处理软件。平台体系架构如图 3 所示：

1) 数据资源池建立：数据接收汇聚包含了全流域水文、气象、环境等全方位监测数据，融合了水利、社会

和其他相关数据，建立一套结构型、半结构型、非结构型数据综合存储模式，方便数据实时更新，数据安全有保障的数据资源池。

2) 数据分析计算与存储：从源数据池中对数据进行抽取、清洗、规范化处理、除重等程序，采用 Hadoop 分布式数据存储模式，利用 Hadoop 提供的分布式文件系统(HDFS)和非关系型数据库(HBASE)对源数据进行分布式调用，同时对数据进行筛选，使用 Hive 将汇总的数据按照不同数据类型进行分别存储，依托水利大数据数据源和数据分析工具，对数据进行挖掘分析。强化数据安全，提供数据安全隔离访问，暴露数据层的 API 接口，将应用层的直接读库需求，转为提供查询、分析结果为上层应用访问提供数据支持。

3) 平台支撑层：支撑层是本系统的基石，在该层中包含了搭建系统的所有基础要素。构建 Ceph 分布式存储系统提供整个系统的存储环境，提高存储层的横向扩展能力和容错能力。构建基于 Kubernetes、Docker 和 OpenStack 的分类混合管理平台，实现对虚拟机、容器的分类管理以及容器的统一编排，提供不同级别的隔离环境，提供对于上层应用层 APPS 的运行环境支持，提供敏捷的应用开发、测试、部署手段。构建基于软件定义存储、软件定义网络、网络安全探针以及应用防火墙等系统，提供对安全事件的实时侦测、管理、隔离的信息安全动态感知主动防御层，实现网络、系统以及数据安全的全面提升。在其上实现包括用户访问服务、权限管理服务、GIS 服务、资源管理服务、关联分析、实现模型分析库集成如 SWAT 模型、TOPMODEL 模型、SAC 模型、新安江模型、SCLS 模型、CRFPDP 模型、水箱(Tank)模型等服务。在该系统中，把这些模块有机组织为



Figure 3. Southwest rivers architecture of water information integrated management platform
图 3. 西南诸河流域水利信息化综合管理平台体系架构

一个体系，形成一套应用支撑平台，来支撑终端用户和业务系统搭建，为今后的水资源管理、河长制以及数字孪生流域提供弹性运行环境支撑。

4) 数据应用层：传统的数据应用层是静态单一化的，无法对水利数据进行动态的分析，科学预警。基于大数据的应用层是面向系统业务的子模块，这些子模块包含了本系统规划的所有业务功能。包括水利信息多维关系展示和水利业务应用。依靠大数据使水利决策从经验到科学过度，提高水利预测能力。结合大数据元数据信息，引入低代码平台，实现水文数据分析人员对于大数据综合运用的快速迭代开发、二次开发，提供海量数据动态分析的手段。为流域综合调度、水旱灾害预警、雨水情与工情分析预测、水生态保护预警、水资源监控等工作提供决策依据。通过对水资源量和质的动态分析，可分析研判水资源取用水变化趋势，水资源污染情况等，可指导水资源管理决策和助力河长制服务。

水利大数据技术实现了西南诸河水利数据多领域、全对象、多类型数据的汇集，形成了海量大数据的存储、管理与分析计算服务。水利大数据技术汇集调动了各级水利部门及不同行业之间水利数据的共享交互，全面提升了数据的质量和覆盖范围。

该平台有效的利用了大数据技术，实现多时空尺度、多数据类型、多方式的数据深度融合，在此基础上经过数据挖掘分析，形成有价值的多维水利信息，为水利决策提供依据。从而实现水利业务“网络化管理”，形成一整套规范统一的管理标准和流程，且使得管理步骤形成一个闭环，从而提升管理的能力和水平，将过去传统、被动、定性和分散的水利管理模式，转变为现代、主动、定量和系统的水利管理，使水利决策从经验到科学转变。

4. 存在的不足

1) 水利大数据运维管理系统有待提升：随着水利大数据体系的不断扩大，对水利数据的运维管理工作越来越受重视，面对数据体量庞大、结构多样的大数据体系，运维管理的自动化、智能化是目前大数据体系需要重点解决的问题。充分利用人工智能、机器学习、大数据技术进行智能化数据处理是未来发展的方向。

2) 数据安全保障有待提升：1、水利大数据的前提是感知，为了大数据分析的需要，有待引入物联网关、物联网服务器等技术进一步完善数据感知网络，进一步扩展水利数据收集的深度和广度。2、建立完善的大数据体系运维管理制度，建立基于 AI 的运维智能分析平台，通过大数据分析、机器学习等技术，洞察运行状态，准确进行故障风险预警，深入挖掘故障根源，建立故障自动恢复体系，全面掌握运维薄弱环节，提升自动化运维能力。

3) 模拟预测能力有待加强：水利大数据技术与专用模型整合，增加了不同时空尺度下复杂环境要素的模拟精度和预测速度，水利预测实现从被动到主动，风险识别和预报预警能力提升。但仍缺乏基于大数据技术的高精度存储、分析和集成模型，缺少更专业、更开放的大数据分析。还缺少适用于西南地区水质、蒸发、干旱时空分析及预测等的模型集成。要进一步完善水利模型库，建设机器认知库，增强大数据智能算法算力建设。

参考文献

- [1] 维克托·迈尔·舍恩伯格, 周涛. 大数据时代[J]. 房地产导刊, 2013(4): 113.
VICTOR, M. S., ZHOU Tao. Big data era. Real Estate Guide, 2013(4): 113. (in Chinese)
- [2] 蒋云钟, 冶运涛, 赵红莉, 梁犁丽, 曹引, 顾晶晶. 水利大数据研究现状与展望[J]. 水力发电学报, 2020, 39(10): 1-32.
JIANG Yunzhong, YE Yuntao, ZHAO Hongli, LIANG Lili, CAO Yin and GU Jingjing. Research status and prospect of water conservancy big data. Journal of Hydropower, 2020, 39(10): 1-32. (in Chinese)
- [3] GIJZEN, H. Development: Big data for a sustainable future. Nature, 2013, 502(7469): 38. <https://doi.org/10.1038/502038d>
- [4] 陈能成, 肖长江, 杨超, 等. 地理空间传感网融合服务技术与应用[J]. 地球信息科学学报, 2020, 22(1): 11-20.
CHEN Nengcheng, XIAO Changjiang, YANG Chao, et al. Geospatial sensor network fusion service technology and application. Journal of Earth Information Science, 2020, 22(1): 11-20. (in Chinese)

- [5] 冶运涛, 蒋云钟, 梁犁丽, 等. 智慧水利大数据理论与方法[M]. 北京: 科学出版社, 2020.
YE Yuntao, JIANG Yunzhong, LIANG Lili, et al. Smart water conservancy big data theory and method. Beijing: Science Press, 2020. (in Chinese)
- [6] 岩腊, 龙笛, 白亮亮, 等. 基于多源信息的水资源立体监测研究综述[J]. 遥感学报, 2020, 24(7): 787-803.
YAN La, Long Di, BAI Liangliang, et al. Review on three-dimensional monitoring of water resources based on multi-source information. Journal of Remote Sensing, 2020, 24(7): 787-803. (in Chinese)
- [7] 蒋云钟, 冶运涛, 赵红莉. 智慧水利大数据内涵特征、基础架构和标准体系研究[J]. 水利信息化, 2019(4): 6-19.
JIANG Yunzhong, YE Yuntao and ZHAO Hongli. Research on connotation characteristics, infrastructure and standard system of smart water conservancy big data. Water Conservancy Informatization, 2019(4): 6-19. (in Chinese)
- [8] MILLER, H. G., MORK, P. From data to decisions: A value chain for big data. IT Professional, 2013, 15(1): 57-59.
<https://doi.org/10.1109/MITP.2013.11>
- [9] KRIKSCIUNIENE, D., SAKALAIUSKAS, V. and KRIKSCIUNAS, B. Process optimization and monitoring along big data value chain. In ABRAMOWICZ, W. Lecture notes in business information processing. Cham: Springer, 2015: 75-86.
https://doi.org/10.1007/978-3-319-26762-3_8
- [10] CURRY, E. The big data value chain: definitions, concepts, and theoretical approaches. In CAVANILLAS, J., CURRY, E. and WAHLSTER, W. New horizons for a data driven economy. Cham: Springer, 2016: 29-37.
https://doi.org/10.1007/978-3-319-21569-3_3
- [11] 陈华, 徐坚, 肖志远, 等. 水文大数据共享平台研究与设计[J]. 水资源研究, 2018, 7(1): 10-18.
CHEN Hua, XU Jian, XIAO Zhiyuan, et al. Research and design of hydrological big data sharing platform. Journal of Water Resources Research, 2018, 7(1): 10-18. (in Chinese)
- [12] 孙洪林. 国家防汛抗旱指挥系统中数据汇集平台的设计和实现[J]. 中国防汛抗旱, 2020, 30(6): 20-26.
SUN Honglin. Design and implementation of data collection platform in national flood control and drought relief command system. China Flood Control and Drought Relief, 2020, 30(6): 20-26. (in Chinese)
- [13] 刘业森, 郭良, 张晓蕾, 等. 全国山洪灾害调查评价成果数据管理平台设计[J]. 南水北调与水利科技, 2017, 15(6): 196-202.
LIU Yesen, GUO Liang, ZHANG Xiaolei, et al. Establishment of data management platform for national mountain flood disaster investigation and evaluation results. South to North Water Transfer and Water Conservancy Technology, 2017, 15(6): 196-202. (in Chinese)
- [14] 周江, 王伟平, 孟丹, 等. 面向大数据分析的分布式文件系统关键技术[J]. 计算机研究与发展, 2014, 51(2): 382-394.
ZHOU Jiang, WANG Weiping, MENG Dan, et al. Key technologies of distributed file system for big data analysis. Computer Research and Development, 2014, 51(2): 382-394. (in Chinese)
- [15] 金国栋, 卞昊穹, 陈跃国, 等. HDFS 存储和优化技术研究综述[J]. 软件学报, 2020, 31(1): 137-161.
JIN Guodong, BIAN Haoqiong, CHEN Yueguo, et al. Review of HDFS storage and optimization technology. Journal of Software, 2020, 31(1): 137-161. (in Chinese)
- [16] 程学旗, 靳小龙, 王元卓, 等. 大数据系统和分析技术综述[J]. 软件学报, 2014, 25(9): 1889-1908.
CHENG Xueqi, JIN Xiaolong, WANG Yuanzhuo, et al. Overview of big data system and analysis technology. Journal of Software, 2014, 25(9): 1889-1908. (in Chinese)
- [17] 吴飞. 大数据环境下的数据存储及应用研究[J]. 科技视界, 2021(32): 10-12.
<https://doi.org/10.19694/j.cnki.issn2095-2457.2021.32.04>
WU Fei. Research on data storage and application in big data environment. Science and Technology Horizon, 2021(32): 10-12.
<https://doi.org/10.19694/j.cnki.issn2095-2457.2021.32.04> (in Chinese)
- [18] 上官博屹, 乐鹏. 流计算与传感网集成的城市内涝模型实时感知研究[J]. 地理信息世界, 2018, 25(3): 1-6.
SHANGGUAN Boyi, LE Peng. Research on real-time sensing of urban waterlogging model integrated with flow computing and sensor network. Geographic Information World, 2018, 25(3): 1-6. (in Chinese)
- [19] 王慧敏, 李臣明, 彭建华, 等. 基于流式大数据处理方式的降雨径流预测计算方法[P]. 中国, CN106446359B. 2019-05-03.
WANG Huimin, LI Chenming, PENG Jianhua, et al. Calculation method of rainfall runoff prediction based on streaming big data processing. CN106446359b. 2019-05-03. (in Chinese)
- [20] 王一旭, 王飞儿, 俞洁. 基于自组织映射与随机森林耦合模型的流域水质空间差异性评估[J]. 环境科学学报, 2020, 40(6): 2278-2285.
WANG Yixu, WANG Feier and YU Jie. Spatial difference assessment of watershed water quality based on self-organizing mapping and random forest coupling model. Journal of Environmental Science, 2020, 40(6): 2278-2285. (in Chinese)