

DOI: 10.19364/j.1674-9405.2019.04.004

智慧永定河云数据中心建设思路探讨

张涛, 黄锐

(水利部海河水利委员会, 天津 300181)

摘要: 随着《永定河综合治理与生态修复总体方案》的出台, 智慧永定河成为智慧水利北方流域试点, 需要建设数据中心, 为流域机构、地方水利部门、相关企业和社会公众提供广泛的, 专业的, 海量的数据共享和支撑服务。云数据中心作为新型的分布式数据中心, 可以在适应管理机制的基础上, 利用大数据、云计算、人工智能等技术, 构建基础设施云、数据资源池和业务支撑能力, 打破资源壁垒, 提高流域综合治理的信息化水平和协同能力, 是较适合智慧永定河的数据中心模式, 可以作为支撑永定河流域综合治理的重要平台和手段, 成为智慧永定河的“大脑”。

关键词: 永定河; 智慧; 云数据中心; 水利信息化; 建设目标; 总体框架

中图分类号: TP39

文献标识码: A

文章编号: 1674-9405(2019)04-0025-05

0 引言

永定河流域发源于内蒙古高原的南缘和山西高原的北部, 东邻潮白、北运河系, 西临黄河流域, 南为大清河系, 北为内陆河。流域地跨内蒙古、山西、河北、北京、天津等5个省(自治区、直辖市), 面积约为4.70万km²。2017年2月, 国家发展和改革委员会、水利部、原国家林业局印发《永定河综合治理与生态修复总体方案》, 提出利用5—10a时间将永定河恢复成为“流动的河、绿色的河、清洁的河、安全的河”, 并规划了大数据中心建设。同年, 智慧永定河被作为智慧水利试点之一, 是北方智慧流域典型设计试点。永定河综合治理和生态修复, 是中国首例跨行政区域实行全流域协同治理的项目, 将有效突破以往以行政区域划界限的治理模式, 是京津冀晋地区协同发展、创新和完善永定河流域协同治理机制的一项重大举措。

智慧流域是流域管理信息化的更高层次, 智慧永定河的“大脑”与“中枢”——大数据中心需要结合流域管理实际, 根据永定河流域内各类单位对流域治理的不同侧重点与领域、数据资源的提供与需求、资源集约化需求, 采用云计算和大数据等

技术建设覆盖流域管理机构、地方水利主管部门、相关企业的新型云数据中心, 实现跨地域、部门、层级、行业的数据共享和交换, 并承载各类智慧应用, 全面支撑永定河流域综合治理与生态修复工程, 推进新一代信息技术在流域管理服务核心领域的应用和融合^[1]。为此, 对永定河云数据中心建设思路进行探讨, 以提高流域综合治理的信息化水平和协同能力。

1 云数据中心介绍

根据IDC(互联网数据中心)的报告显示, 全球大数据市场规模年增长率达到40.0%。其中, 大数据技术及服务市场复合年增长率(CAGR)将达到31.7%, 其增速约为信息通信技术(ICT)市场整体增速的7倍多, 因此需要进行大数据中心建设, 以满足日益增大的数据查询、统计、分析、挖掘和备份等数据存储, 管理和处理需求。

随着采集汇聚的业务、政务、互联网等数据的不断拓展, 大数据中心未来可能面临“数据丰富、信息贫乏”的尴尬局面, 原因主要有以下3点: 1) 缺乏有效的数据分析技术; 2) 由于缺乏业务协同,

收稿日期: 2019-06-29

作者简介: 张涛(1987-), 男, 天津人, 工程师, 主要从事水利信息化相关工作。E-mail: zhangtao_jlu@163.com

数据交换频率低,时效性差;3)数据质量不高,例如出现滥用缩写词、惯用语,数据输入错误,重复记录,丢失值,拼写变化,不同的计量单位和过时的编码等各种问题。随着永定河流域涉水单位今后业务的开展,数据将不断发生变化,因此要保证数据的质量,必须建立数据动态更新机制、手段,配备专职人员保证数据的及时更新,将云技术和大数据中心技术相结合的云数据中心应运而生。

云数据中心,是采用虚拟化、自动化、并行计算、安全策略及能源管理等新技术,解决目前数据中心存在的成本增加过快和能源消耗过度等问题,并通过标准化、模块化、动态弹性部署和自助服务的架构方式,实现对业务服务的敏捷响应和按需获取的新型数据中心。云数据中心本质上由云计算平台和服务构成,云计算服务包括通过各种通信手段提供给用户的应用、软件、工具及计算资源服务等,云计算平台包括用来支撑这些服务的安全可靠和高效运营的软硬件平台。通过云计算平台将1个或多个数据中心的软硬件整合起来,形成一种分层的虚拟计算资源池,并提供可动态调配和平滑扩展的计算、存储和网络通信能力,用以支撑云计算服务的实现。与传统数据中心相比,云数据中心提供的不仅是计算、存储、管理和监控服务,更重要的是提供数据、基础组件、算法、模型、通用支撑和部分核心业务。云数据中心目前除广泛应用于大型云服务提供商外,在政府、行业应用中也逐步开展,如雄安新区规划以云数据中心模式建设城市大脑。

智慧永定河是典型的行业信息化顶层设计,具有非常强的专业性,且海河水利委员会(以下简称海委)与两省两市及永定河投资公司需要协同共享的核心是数据、模型、业务等能力,这些能力分散在各个单位,很难实现简单集中。同时,各自建设管理的信息系统相互独立存在和运行,形成“信息孤岛”,业务数据繁杂,涉密情况复杂,数据资源共享交换、采集汇聚存在多部门协同需求;传统大数据中心,通过数据交换和报送、页面采集等手段将数据采集汇总,并进行清洗、比对等数据提质操作,最终将所需的数据存储于大数据平台中的方式已不能满足性能和功能需求。因此,云数据中心利用云计算技术,可以调用各个节点的资源、数据和能力,是最适合智慧永定河的技术路线。

2 云数据中心建设目标

充分利用云计算、大数据等技术,构建永定河基础设施云、数据资源池和业务支撑能力,实现计算、存储、网络等资源的弹性分配,统一运维,实现永定河流域水利行业、其他行业和互联网数据的归集,治理,共享和开放^[2],实现跨部门信息共享、业务协同和决策支撑,实现各类应用的资源化、服务化,从根本上解决信息孤岛问题,促进信息资源的融合和数据的一体化,为永定河流域涉水管理政府机构、相关企业、社会公众提供所需要的,基础、数据和应用等服务能力,赋能永定河流域智慧化发展这一必然趋势^[3]。

3 云数据中心总体框架

在深入分析智慧永定河“透彻感知、全面互联、深度融合、广泛共享、智能应用、泛在服务”^[4]总体需求的基础上,按照水利信息资源整合“五统一”的总体要求^[5],提出云数据中心总体框架,总体框架如图1所示。永定河云数据中心由3个层次构成,包括基础设施云、数据资源池、业务支撑能力3个部分,其中基础设施云整合各类基础计算、存储、网络、机房环境等资源,通过云管理平台提供云IaaS(基础设施即服务)能力,为用户提供弹性、可定义、按需分配、特性丰富的云化资源;数据资源池主要包括存储空间、基础、监测、业务、多媒体、元数据及互联网数据等各类数据资源,并提供数据资源交换共享服务,提供云DaaS(数据即服务)能力;业务支撑能力由信息资源交换服务、应用支撑和智慧使能等平台组成,提供云PaaS(平台即服务)和部分云SaaS(软件即服务)能力。

4 云数据中心主要建设内容

4.1 基础设施云

基础设施云主要包括计算、存储、安全、机房、会商环境及云管理运维平台。安全部分应符合国家关于网络安全的相关要求,机房及会商环境因各节点现实情况及差异性相对独立,暂不介绍。

1) 计算资源池。主要通过底层数据的处理,将物理服务器进行虚拟化构建虚拟化资源池,解除应用和操作系统与硬件的紧耦合关系,使物理设备的差异性、兼容性与上层应用透明,不同虚拟机之间相互隔离,互不影响。

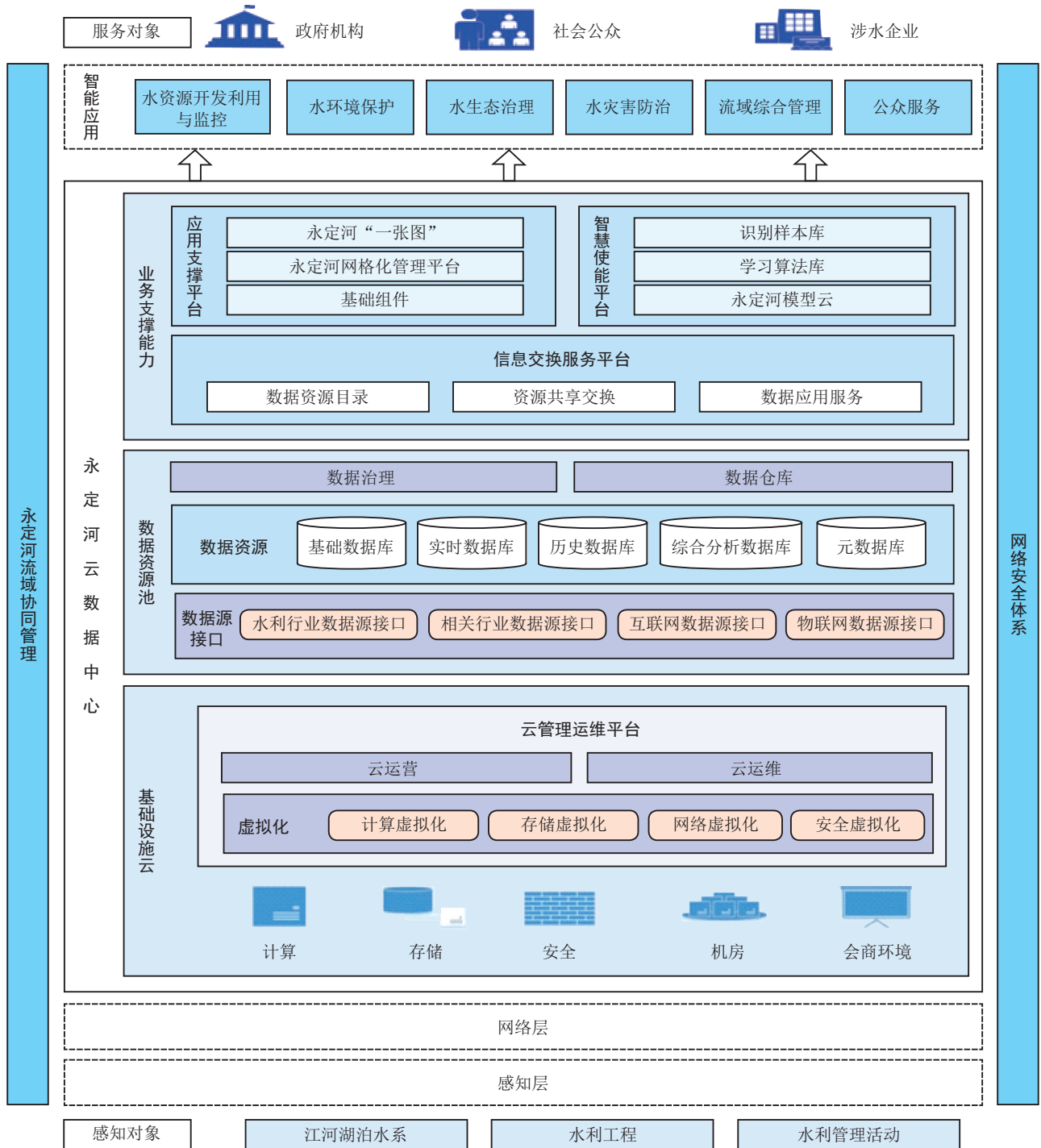


图 1 永定河云数据中心总体架构

2) 存储资源池。采用集中式存储和分布式存储相结合的方式，分为块和文件等存储资源池。其中块存储资源池采用集中存储方式，主要存储业务数据；文件存储资源池采用分布式存储方式，主要存储视频和图片等数据。

3) 备份资源池。将采用云备份解决方案，用户可以通过管理平台申请对本节点的部分或全部虚拟

机或者虚拟机的磁盘做备份，可以自行设置备份策略，包括备份周期、执行时间点、备份时保留的副本数目、全备或增备策略。

4) 云管理运维平台。是整个大数据中心后台的管理、调度、运维中心，实现对主、运营、备份 3 个中心及 4 个分节点的管理运维，包括资源编排、统一运营、综合分析等。

4.2 数据资源池

整合永定河流域水利行业数据,融合相关行业和社会数据,在永定河基础设施云各级节点构建数据资源池,通过多元化采集、主体化汇聚构建全域化原始数据,基于“一数一源、一源多用”原则,汇聚全域数据,开展数据治理,形成标准一致的数据资源池。

根据永定河流域各级水利业务应用和综合决策需要,按照数据资源目录,分类分级汇聚至省级分节点和海委核心节点形成数据库,包括以下数据库:

1) 基础类数据库。包括基础地理信息、社会经济信息与水利基础等数据库。

2) 实时类数据库。包括雨水情、地下水、旱情、灾情、气象、水质、工程、水利建设信息等实时数据库,以及实时遥感动态与突发事件数据库。

3) 历史数据库。历史数据库主要与实时数据库相对应,实时数据库中的数据是实时采集监测得到的,而历史数据库中的内容是根据实时数据库中的内容整理得到的。根据国家和行业的整编规范,例如《水文资料的整理规范》等,定期(如1a)对实时数据库中的内容进行整理形成历史数据库。

4) 综合分析数据库。包括经过各业务应用和决策支持等系统综合分析与处理后的统计分析报表、报告、预测信息、各种方案。

5) 元数据库。通过数据仓库技术实现对数据库中海量数据的载入、存储、提取、使用等的管理,而数据仓库中的一个重要建设内容就是元数据库的建设。

4.3 业务支撑能力

业务支撑能力完整地作为上层各类智能业务应用提供数据、共性和专业支撑、模型算法等 Paas 服务能力,具体包括信息资源交换服务、应用支撑和智慧使能3个平台。3个平台具体分析如下:

1) 信息交换服务平台。主要通过建立数据资源目录服务功能,以信息服务、共享接口和数据传输等多种形式实现部门业务驱动,基础先行,按需整合的共享交换策略,并支撑基于统一信息资源的各类综合业务、创新示范等应用的发布和管理,提供信息资源目录的统一展示。

2) 应用支撑平台。按照大中台、小前台、微服务的架构,将共性应用资产下沉,重点聚焦于水利业务共性剥离和微服务构建应用支撑平台,通过为上层水利业务应用统一提供公共基础服务支撑,避

免不同水利业务应用之间的重复建设,支持前台快速开展业务创新。主要任务包括基础组件整合、永定河网格化管理平台构建和“永定河一张图”建设3个部分。

3) 智慧使能平台,对涉及永定河的各类水利机理模型、机器学习算法、数据模型及相应的样本进行统一管理,验证模型精度并不断迭代,为各级单位提供模型算法服务。

5 云数据中心部署模式

根据目前永定河流域管理的现状,以1个主中心、1个备份中心、1个运营中心、4个分支节点“三中心、四节点”的模式进行部署。

1) 主中心。在海委机关建设永定河基础设施云的主中心,主中心是基础设施云的核心,负责流域内重要系统及数据的部署及存储,负责各单位之间的数据交换服务,承担基础设施云的统一管理及运维。

2) 备份中心。在天津选址建设永定河基础设施云备份中心,定期将基础设施云上的重要数据进行备份,为保证链路质量,备份中心与主中心距离不宜太远,备份中心的选址将采取与主中心同城异址的方式进行选择。

3) 运营中心。在永定河投资公司建设永定河基础设施云运营中心,通过日常运营管理收集永定河流域重要基础信息,供海委和其它分支节点分享,同时实现永定河投资公司与主中心和其它分支节点的整合共享。

4) 分支节点。在流域内两省两市分别建设永定河基础设施云分支节点,负责永定河流域在本省市范围内的数据收集汇总管理工作,负责与主中心之间的数据交换。

6 结语

数据中心是智慧化应用的物质基础,云数据中心则是提供计算、存储、数据、算法、模型、业务等能力的数据中心升级版。云数据中心模式以往多采用于IT企业、云服务商等,针对智慧永定河既有流域管理,也有区域管理,并有强烈的数据、业务互通需求,云数据模式提供了较好的支撑手段,具有一定的普遍性。本研究偏重于建设思路与模式选

择，在后期进行具体设计时，还要根据各地政务云模式、数据需求、地方标准等多种情况进行详细研究。总之，通过建设资源动态分配的永定河基础设施云，汇集共享海量数据的数据资源池，支撑各类智慧应用的业务支撑能力，云数据中心可以成为智慧永定河的“大脑”，为实现全景监控仿真、精准预测预警、联合调度会商、多维决策指挥、综合分析评估等智慧应用，构建涵盖水资源、水安全、水环境、水生态、水管理等领域的全业务体系^[6]，推动永定河综合管理模式的创新，实现“流动的河、绿色的河、清洁的河、安全的河”，提升永定河绿色生态河流廊道协同管理能力提供强有力支撑。

参考文献：

- [1] 蔡阳. 水利信息化“十三五”发展应着力解决的几个问题[J]. 水利信息化, 2016 (1): 1-6.
- [2] 蔡阳. 智慧水利建设现状分析与发展思考[J]. 水利信息化, 2018 (4): 1-5.
- [3] 刘辉. 水利工程智慧之路的探讨——从通用 IT 到知识自动化再到数据智能[J]. 水利规划与设计, 2017 (12): 64-68.
- [4] 水利部参事咨询委员会. 智慧水利现状分析及建设初步设想[J]. 中国水利, 2018 (5): 1-4.
- [5] 水利部信息化工作领导小组办公室. 水利信息化资源整合共享顶层设计[R]. 北京: 水利部信息化工作领导小组办公室, 2015: 3-6.
- [6] 海河水利委员会, 智慧永定河技术方案[R]. 天津: 海河水利委员会, 2019: 13-14.

Discussion on construction of Cloud Data Center from Smart Yongding River

ZHANG Tao, HUANG Rui

(Haihe River Water Conservancy Commission, the Ministry of Water Resources, Tianjin 300181, China)

Abstract: With the publication of the General Plan for Comprehensive Control and Ecological Restoration of Yongding River, the Smart Yongding River becomes the pilot project of Smart Water Conservancy in northern basin. It needs to build a data center to provide extensive, professional and massive data services and support services for river basin organizations, local water conservancy departments, relevant enterprises and the public. As the new form of distributed data center, Cloud Data Center based on adapting to management mechanism, by using big data, cloud computing and artificial intelligence technology, can build Infrastructure Cloud, data resource pool and business support ability, break the resource barrier and improve the informatization level. So using the Cloud Data Center mode is fit to the Smart Yongding River, Yongding River Cloud Data Center is the important platform and means for supporting comprehensive management of Yongding River Basin, can be the “brain” of Smart Yongding River.

Key words: Yongding River; Smart; Cloud Data Center; water conservancy information; construction objectives; overall framework

(上接第 24 页)

resources monitoring platform, etc. However, there are still many issues. The general protection standard is not high enough. Many departments' secure systems are constructed independently. Network threat perception is not sufficiently capable to identify risks. And the system lacks methods for cloud computing and other new technologies. Based on the facts of water industry, this paper introduces the overall scheme of network security, which includes defense in depth approach, monitoring, warning, emergency response ability construction, secure management and secure operation. The scheme points out the direction of constructing an all-round coordinated active intelligent cyber security system in Smart Water Conservancy Era, which is significant to the water industry.

Key words: Smart Water Conservancy; monitoring and warning; cyber security; secure technology; secure management; secure operation