

云技术在水利地理信息服务平台建设中的应用

刘仲刚, 陈 辉, 黄章羽, 陆 明

(江苏省水利信息化工作领导小组办公室, 江苏 南京 210029)

摘 要: 为解决水利信息化建设中重复建设、资金浪费等问题, 在江苏省水利地理信息服务平台中采用云计算技术。介绍云计算与云 GIS 的概念与基本特征, 以江苏省水利地理信息服务平台建设为例, 阐述云 GIS 平台的总体设计与架构, 分析虚拟化和移动数据采集与同步等关键技术的应用。江苏省水利云 GIS 平台的建成, 可为江苏各级水利部门提供全省基础地理信息数据、高分辨率遥感影像及航片、水利地理信息数据等数据服务; 高效能服务器、海量数据存储等基础设施服务; 基于水利专网提供各种水利业务应用系统开发的平台服务。

关键词: 云计算; 地理信息服务; 水利信息化; 江苏省

中图分类号: TP393; P208.3

文献标识码: A

文章编号: 1674-9405(2014)02-0015-05

0 引言

水利信息化是水利现代化的基础和标志之一, 水利地理信息服务平台是水利信息化的基础平台, 也是国家空间基础设施框架的重要组成部分。水利信息化主要包括: 防汛防旱指挥、水情测报、水资源管理、防汛会商、办公自动化、信息服务等系统建设, 实现信息技术为水利建设和管理服务^[1]。江苏省地处平原河网地区, 河湖众多, 水网密布, 但水资源时空分布不均衡, 江苏水利以南水北调、江水东引、引江济太 3 大工程为重点, 建设了一大批水利工程, 实现了长江、淮河、沂沭泗 3 大流域水源互调互济。江苏水利历来重视水利信息化建设, 很早就开始投入防汛防旱指挥、水资源管理、重点工程自动监控等业务系统的建设, 水利专网已覆盖省、市、县三级水利部门, 水利信息化在水利现代化建设中发挥了重要作用。水利信息兼具空间性和时间性特征, 是一种典型的地理空间数据。2009 年 8 月, 江苏省水利厅提出江苏省水利地理信息系统是数字水利、现代水利的基础设施, 是现代水利业务应用和社会公众服务的基础平台, 是水利信息化进程的一项标志性工程^[2]。目前, 江苏省水利地理信息系统一期工程已顺利完成, 制定了江苏省水利地理信息分类与编码、采集、信息共享等标准, 初步建

成了基于全省 1:10 000 基础地理信息数据和最新影像资料的省级水利地理信息数据库, 开发了水利地理空间信息数据发布及管理系统, 开展了地理信息数据维护、服务管理和信息发布等工作, 为建设全省信息共享、标准统一、服务方式规范的水利地理信息系统提供了技术保障和基础框架。

2010—2012 年, 国务院开展了第一次全国水利普查工作。对江河湖泊、水利工程、经济社会用水、河湖开发治理保护、水土保持、水利行业能力建设等对象进行了全面的调查^[3]。江苏省已经完成了第一次水利普查任务, 初步集成了全省水利地理信息, 涵盖河湖基本情况、水利工程、地下水取水井、灌区、经济社会用水、河湖开发治理保护、水土保持、行业能力 8 大专题。如何扩大水利普查数据应用范围, 整合各个业务系统, 更有效地发挥省级水利信息资源, 在省、市、县三级水利部门实现统一的门户入口、身份认证、应用平台、数据标准、数字地图, 以解决市县水利信息化重复建设、资金浪费等问题, 即成为江苏省水利地理信息服务平台二期工程的关键。云计算作为新兴信息技术, 提供了一种解决此类问题的新技术思想。下面结合江苏省水利地理信息服务平台建设, 初步分析云计算在水利地理信息服务平台中的应用。

收稿日期: 2013-10-11

作者简介: 刘仲刚(1974—), 男, 河北玉田人, 工程师, 主要从事水利信息化科研与管理工作。

1 云计算与云 GIS

1.1 云计算

网络时代的信息和数据呈现爆炸式增长,海量数据、大规模并发请求客观上需要不断加大软硬件投入,更提高了后期的运维成本。John McCarthy 在 1960 年指出,计算迟早会变成一种公用基础设施,即计算能力可以作为一种商品进行流通,就像煤气、水电一样,取用方便,费用低廉^[4]。2006 年,谷歌公司首次提出云计算的概念。根据美国国家标准与技术研究院(NIST)的定义,云计算是一种利用互联网实现随时、随地、按需、便捷地访问共享资源池(如计算设施、存储设备、应用程序等)的计算模式。计算机资源服务化是云计算重要的表现形式,它为用户屏蔽了数据中心管理、大规模数据处理、应用程序部署等问题。通过云计算,用户可以根据业务负载快速申请或释放资源,并以按需支付的方式对所使用的资源付费,在提高服务质量的同时降低运维成本^[5]。

云计算是网格计算、分布式计算、并行计算、效用计算、网络存储、虚拟化、负载均衡等技术发展融合的产物,被认为是信息产业的又一次革命,一经提出便得到各界的广泛关注。谷歌、微软、亚马逊等信息产业领导者和斯坦福大学等科研机构,合作开展了大量的研究工作,英美等国已将其列为国家战略^[6]。2012 年,国务院批复了《中国云科技发展“十二五”专项规划》,明确把云计算列为重点发展的国家级战略性新兴产业。云计算具有以下 4 个显著特点:1)云计算能够提供最可靠、安全的数据存储中心;2)云计算对客户端的硬件设备配置要求较低,软件使用服务方便;3)云计算可以轻松实现不同设备间的数据与应用共享服务;4)云计算为今后使用网络数据提供了几乎无限多的可能,扩展潜力巨大^[7]。云计算的概念提出以来,即得到水利行业的重视。目前,在水利模型计算、流域洪水预报、云服务平台搭建等方面有较多的应用与探讨。

从已发表的文献资料看,云计算在水利信息化方面已经有较多应用,中国水利学会 2013 年论文集中关于水利信息化方面的文章大约有 1/5 与云技术有关,但提供水利地理信息云服务还未见报道^[8]。

1.2 云 GIS

云 GIS 是利用云基础设施获得大规模计算能力,解决 GIS 中海量空间数据的分布式存储、处理

任务划分、查询检索、互操作和虚拟化等关键性科学问题,提高 GIS 数据处理与管理能力,为计算和数据密集型的各类 GIS 服务提供高性能处理的技术^[9]。云 GIS 除了利用基础设施即服务(IaaS)、平台即服务(PaaS)、软件即服务(SaaS)等云技术服务外,由于地理空间数据的独特性,还以数据即服务(DaaS)的形式为用户提供空间信息应用^[10-11]。

云 GIS 的服务模式有根本性的转变,具体体现如下:

1) GIS 的传统建设模式是每建立 1 项业务系统,就需要 1 套 GIS 软件加空间地理数据库。云 GIS 建立在统一平台之上,只要同一套 GIS 软件提供同一个数据库中的地理空间信息给各个业务系统,可解决数据多源异构、业务无法协同的问题。

2) 云 GIS 将地理信息行业由提供产品转为提供服务,个性化的专题制图将更为容易方便,将使 GIS 用户和应用范围得到极大的扩展。

3) 集中管理地理空间数据资源。云 GIS 平台建设方的技术和资金实力都非常雄厚,可为整个系统或行业提供基础设施、平台、数据和应用服务,从而达到资源的有效控制和合理使用,特别适合于我国目前的行业管理体制。

云 GIS 有 3 种建设模式:公有云、私有云和混合云 GIS。公有云 GIS 是开放的、所有用户共同使用的云环境,业务功能以服务的方式,通过互联网提供给用户,如 ESRI 的 ArcGIS Online、超图地理信息云服务平台等;私有云 GIS 由某个组织独立构建和使用,通过局域网为企业内部用户提供服务,私有云的所有者不与其他企业或组织共享任何资源;混合云 GIS 整合了公有云与私有云 GIS 所提供服务的云环境,用户根据自身因素和业务需求选择合适的整合方式,制订使用混合云的规则和策略^[12-13]。根据水利部门的业务特点和数据要求,江苏省水利地理信息服务平台先以私有云 GIS 模式建设,随着应用需求的增加,向混合云发展,充分利用公共资源。

2 云 GIS 平台架构与关键技术

2.1 平台总体架构

江苏省水利地理信息服务平台旨在为全省各级水利部门提供信息的采集、存储、处理、提取、交换和应用等方面的服务,这就要求平台运行在江苏省水利信息专网上。江苏省水利信息专网是省、

市、县三级多层次广域网，省厅广域网的节点连接 13 个市水利局、13 个市水文局、9 个厅直属管理处。江苏省云 GIS 平台基于虚拟化云计算平台搭建而成，从传统的 GIS 建设模式转变为“云 GIS”平台模式，实现服务器、数据存储、网络等物理资源，以及操作系统、数据库、GIS 软件平台等软件资源的共享。在江苏省水利地理信息服务平台基础上，市、县水利部门可以根据自身业务需求和信息

化能力，自主选择分布式或集中式方式建设市、县级水利地理信息服务平台。江苏省云 GIS 平台总体架构如图 1 所示。

2.1.1 平台软件环境

平台软件系统由操作系统、数据库管理系统、虚拟化管理软件、GIS 软件等组成。操作系统选用 Windows 2008 Server，数据库管理系统选用 Oracle 10g，虚拟化管理软件选用 VMware vCenter

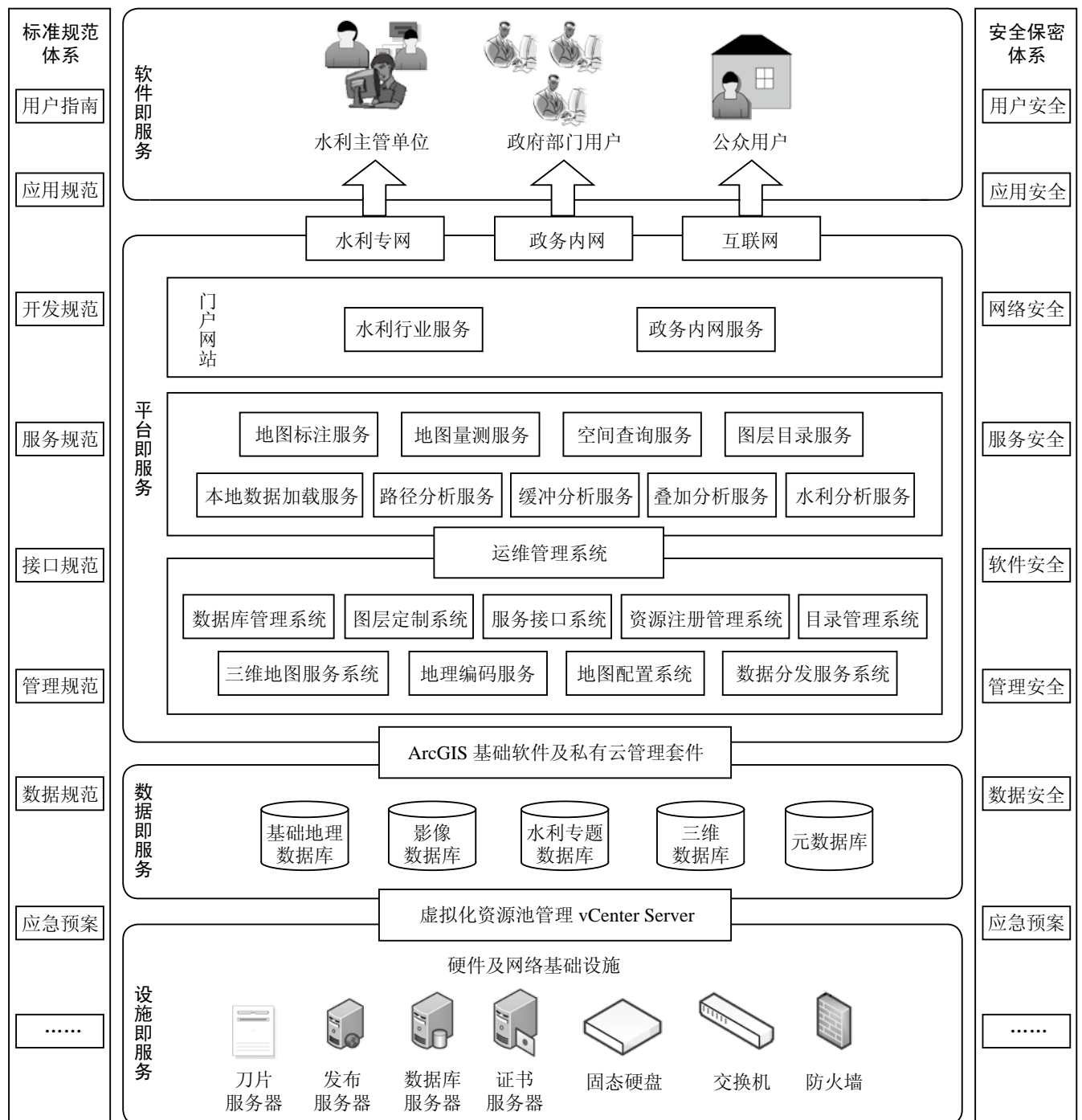


图 1 江苏省云 GIS 平台总体架构

Server, GIS 软件选用 ArcGIS 10 系列产品。对于云 GIS 平台来说, 关键软件是虚拟化和云 GIS 这 2 个管理软件。

1) 虚拟化管理软件。VMware vCenter Server 提供了用于管理虚拟基础架构的集中式可扩展平台, 管理虚拟化涉及的方方面面, 监控整个虚拟化架构的运行状况。虚拟化管理软件可以集中管理物理主机及虚拟机, 提供统一的 B/S 架构方式管理平台进行集中化管理和性能监测, 优化资源配置, 提高响应速度, 利用虚拟机模板快速部署, 以及设置安全访问权限管理机制等。

2) ArcGIS 私有云管理套件。ArcGIS 私有云管理套件由云端自服务门户和运维管理 2 个子系统组成。云端自服务门户子系统, 主要面向业务部门用户, 以自服务的方式具有申请/取消、性能监控、资源调整、资源度量等功能; 云端运维管理子系统, 主要面向数据中心的运维管理人员, 实现对数据中心云环境中 ArcGIS 资源的监控管理等。ArcGIS 云管理套件在基础设施层, 采用 VMware 的基础设施云管理套件 vCloud Director, 将数据中心的基础架构资源整合成虚拟数据中心资源池, 包括计算、存储、网络等资源池, 根据用户业务的实际需求消费这些资源。

2.1.2 平台硬件环境

平台的硬件系统由计算机网络设备、数据库服务器、数据存储系统、应用服务器、工作站、信息接收设备及相关安全防护设备等组成。

1) 服务器。服务器系统包括应用和数据库服务器。平台建成后需要部署十多个应用软件系统向全省水利部门提供地理信息服务, 选用高端刀片服务器 5 把, 做虚拟化资源池。应用代理服务器、虚拟化管理平台各用 1 台虚拟机, 其它虚拟机安装 GIS 服务平台软件。水利地理信息平台数据格式复杂多样, 为有效存储数据, 满足业务处理的高效性需求, 配置 2 台企业级数据库服务器做集群系统, 以满足高可用性和负载均衡服务要求。

2) 数据存储。目前, 江苏省水利网络数据中心存储的数据大约为 2 TB, 按照线性增长模型估算, 每年数据生成量为 1 TB, 若平台应用得到迅速推广, 数据将以非线性方式增长, 预计 5 年内可达到 20~30 TB。另外, 平台存储的数据多为矢量地图、栅格及空间三维数据, 发布的服务多为图片及 XML 格式文件, 水利专网并发访问最高按照 150 个计

算, 每个访问的数据传输至少在 300 kB, 而每个访问可能会产生 5~20 个网络层会话, 这对存储平台的输入输出要求较高。考虑到平台数据的重要性, 数据备份至关重要, 还需提供 20 TB 以上的备份空间。江苏省水利云 GIS 平台采用 50 TB 固态硬盘阵列用于存储和备份。

2.1.3 标准体系与安全保障

标准体系是水利信息化建设中涉及的所有标准的集合。江苏省水利地理信息服务平台标准体系以现有国家和省级法律法规为依托, 遵照水利、测绘、计算机等行业标准规范, 制定了适应水利信息化应用需求的标准规范计 31 项, 其中《江苏省水利地理信息服务平台电子地图数据规范》等数据资源类 15 项, 《江苏省水利地理信息服务平台数据维护与更新规范》等建设管理类 9 项, 《江苏省水利地理信息服务平台浏览器端应用开发接口规范》等应用服务类 7 项。

系统安全是保证水利地理信息服务平台正常运行的关键, 江苏省水利地理信息服务平台建立了网络、数据、软件系统的安全与系统安全管理等一整套安全保障体系。网络安全主要用于保障主机与网络设备的正常运行; 数据安全主要对平台数据源进行控制, 保证部署在平台上的数据服务安全; 软件系统安全用于解决系统软件、应用软件、访问服务等的安全; 系统安全管理通过制定和执行人员, 文档, 系统环境, 软硬件系统的选购、使用与维护, 应用系统运营与开发, 应急预案等强制性规章制度来减少外界环境因素的干扰和破坏, 消除安全隐患, 提高系统的安全性、可靠性。

云计算是一种新的技术方法, 目前还没有专门针对云计算的安全保障标准, 还是信息系统级的软硬件安全要求, 一般采用《计算机信息系统安全保护等级划分准则》、《信息系统安全等级保护基本要求》等标准, 江苏省水利地理信息服务平台按安全保护等级二级进行保护。

2.2 关键技术

2.2.1 虚拟化技术

虚拟化技术是云计算最重要的特征, 应用系统在虚拟的而非真实的基础上运行。在传统的信息化建设中, 一般来说, 每个项目建设都要配置 1 套应用和数据库服务器, 但这些高配置、大容量的服务器的计算能力并没有得到充分利用, 而且机房占用的空间趋于饱和, 配电、制冷系统需要消耗大量能

源。从理论上讲, 虚拟化技术可根据用户需求, 在服务器集群上配置无限多的服务器, 通过软件管理虚拟化资源池, 简化软、硬件的重新配置过程, 节省占地, 降低能耗。简单的讲, 虚拟化技术是在 1 台计算机上, 通过虚拟化软件模拟出多个虚拟机, 每个虚拟机有独立的操作系统和硬件设施, 可以独自承担不同的应用负载, 这些计算资源和硬件设施都是弹性的、可配置的。

根据江苏省水利地理信息服务平台的需求, 配置 5 把高端刀片服务器, 每台物理服务器上安装虚拟化软件, 根据应用需求生成多台虚拟服务器和虚拟机的应用网络、ESX 服务器 (VMware Elastic Sky X Server) 管理软件及高可用性 (HA) 的心跳网络。每个虚拟服务器, 从功能、性能和操作方式上, 等同于传统的单台物理服务器, 在配置虚拟化服务器时, 根据应用软件运算能力的不同要求, 分配服务器数据处理资源。

2.2.2 移动数据采集与同步技术

水利行业的野外采集工作任务较多。常规的野外采集需要 GPS、便携计算机等设备支持, 实际工作中存在诸多不便。移动互联网的发展丰富了云 GIS 的外沿, 江苏省水利地理信息服务平台为移动数据采集提供了 iOS, Android 和 Windows Phone 系统的服务支持, 在移动平台设计移动采集系统, 实现地图放大、缩小、漫游、全景、浏览等功能。数据采集人员基于现有水利电子地图数据进行数据采集, 并实时发布更新的数据服务, 给其他用户共享。

由于江苏省水利地理信息平台运行在水利专网上, 而外业数据采集时, 用户采集到的数据保存在用户的移动设备上, 因此需通过无线网络上传到与外网相连的服务器上。为保证水利专网运行安全和数据的互联互通, 在外网部署 1 台移动采集数据服务器, 实时接收移动采集系统上传的数据, 在水利专网也部署 1 台移动采集数据服务器, 专网服务器到外网的服务器获取数据, 实现同步接收移动采集系统上传的数据, 外网与专网之间通过防火墙进行逻辑隔离。

3 结语

云计算技术改变了传统的 GIS 建设模式, 将独占资源变为资源共享, 可以最大化地利用计算资源。江苏省水利地理信息服务平台为江苏省各级水利部门提供高效优质的计算资源、海量的存储空

间、安全的数据中心, 可以避免水利 GIS 建设的重复投资。同时, 县市水利部门可以将自己的业务应用系统和数据中心部署到省水利 GIS 平台上, 可大大降低 GIS 建设成本, 使得县市一级的水利信息化水平实现跨越式的发展。

江苏省云 GIS 平台仍然处于初建阶段, 尚有许多理论、技术、制度上的问题有待在实际工作中进一步研究和完善, 尤其在大规模并行计算和海量数据分布存储等方面有待深入研究。在建设水利地理信息服务平台的同时, 需要建立健全与测绘、环保、气象、林业、国土等部门的地理信息资源共建共享交换机制, 为科学决策、应急管理、社会服务提供全面的地理信息服务。

参考文献:

- [1] 苑希民. 水利信息化技术应用现状及前景展望[J]. 水利信息化, 2010 (2): 5-8.
- [2] 柏屏, 高祥涛. 江苏省水利地理信息系统的设计与实现[J]. 水利信息化, 2011 (4): 61-65.
- [3] 雷莹, 刘丽芬, 许礼林, 等. 全国第一次水利普查的基础地理信息保障服务[J]. 地理信息世界, 2013, 20 (2): 94-98.
- [4] 张建勋, 古志民, 郑超. 云计算研究进展综述[J]. 计算机应用研究, 2010, 27 (2): 429-433.
- [5] MELL P, GRANCE T. Definition of Cloud Computing [R]. Colorado Boulder: National Institute of Standards and Technology, 2011.
- [6] 罗军舟, 金嘉晖, 宋爱波, 等. 云计算: 体系架构与关键技术[J]. 通信学报, 2011, 32 (7): 3-21.
- [7] 李水英. 基于云计算的地理空间信息公共服务平台建设构想[J]. 数字技术与应用, 2011 (8): 233-236.
- [8] 中国水利学会. 中国水利学会 2013 学术年会论文集[C]. 北京: 中国水利水电出版社, 2013: 129-148.
- [9] 林德根, 梁勤欧. 云 GIS 的内涵与研究进展[J]. 地理科学进展, 2012, 31 (11): 1519-1528.
- [10] 陈维崧, 陈庆秋. 基于云计算的 GIS 研究[J]. 测绘与空间地理信息, 2011, 34 (1): 157-161.
- [11] 冯军. 基于云计算的政务 GIS 平台研究[J]. 测绘, 2013, 36 (2): 80-84.
- [12] 张桂芬. 基于云计算的城市地理信息公共服务平台设计与实现[J]. 城市勘测, 2012 (4): 12-15.
- [13] 贾萍, 刘聚海, 王远. 基于云计算及物联网的 GIS 综述[J]. 国土资源信息化, 2012 (6): 11-14.

(下转第 35 页)

- [3] Salton G, Wong A, Yang Chungshu. A vector space model for automatic indexing[J]. Communications of the ACM, 1995, 18 (11): 613-620.
- [4] LI Sujian, ZHANG Jian, HUANG Xiong, et al. Semantic computation in Chinese question-answering system [J]. J. Comput. Sci. & Technol, 2002, 17 (6): 933-939.
- [5] Chatterjee N. A statistical approach for similarity measurement between sentences for EBMT[C]//Proceedings of Symposium on Translation Support System. Washington: IEEE Computer Society, 1999: 15-17.
- [6] 李彬, 刘挺, 秦兵, 等. 基于语义依存的汉语句子相似度计算[J]. 计算机应用研究, 2003, 20 (12): 15-17.
- [7] 范立新. 改进的中文近似字符串匹配算法[J]. 计算机工程与应用, 2006, 42 (34): 172-174.
- [8] 黄林晟, 邓志鸿, 唐世渭, 等. 基于编辑距离的中文组织机构名称简称-全称匹配算法[J]. 山东大学学报: 理学版, 2012, 47 (5): 43-48.
- [9] 刁兴春, 谭明超, 曹建军. 一种融合多种编辑距离的字符串相似度计算方法[J]. 计算机应用研究, 2010, 27 (12): 4523-4527.
- [10] F.Fábris, I.Drago, F.M.Varejão. A multi-measure nearest neighbor algorithm for time series classification[C]//Proceedings of the 11th Ibero-American Conference on AI (IBERAMIA '08). Berlin:Springer-Verlag, 2008: 153-162.
- [11] Evgeniou T, Pontil M, Elisseeff A. Leave one out error, stability, and generalization of voting combinations of classifiers[J]. Machine Learning, 2004, 55(1): 71-97.

Data Fusion of Water Resources Census Based on Similarity Searching

WANG Jimin¹, XU bo², ZHU Yuelong¹, WANG Weijun³, LI Shijin¹, WAN Dingsheng¹

(1. College of Computer and Information, HoHai University, Nanjing 210098, China;

2. Development Research Center, the Ministry of Water Resources, Beijing 100038, China;

3. Nanjing Social Insurance Administration Department, Nanjing 210017, China)

Abstract: Aiming at the existing problems of data fusion in the first National Water Resources Census, adaptive edit distance similarity measure is proposed in this paper. The traditional edit distance is improved by adjusting the weight of editing operations and heuristic learning weight to improve the accuracy of similarity search. The data fusion approach and process based on edit distance is proposed. The approach is validated by the data processing of water resources census.

Key words: edit distance; water census; similarity searching; data fusion

(上接第19页)

Application of Cloud Technology in Water Resources Geographic Information Services Platform Construction

LIU Zhonggang, CHEN Hui, HUANG Zhangyu, LU Ming

(Water Resources Informatization Leading Group of Jiangsu Province, Nanjing 210029, China)

Abstract: Cloud computing is adopted in the Water Resources Geographic Information Service Platform (WRGIS) of Jiangsu province to avoid or reduce redundant construction and waste of capital. In this paper, concepts and fundamental features of cloud computing and cloud GIS are introduced. Taking WRGIS of Jiangsu province as an example, the general scheme and frame of cloud GIS platform are proposed, and virtualization, mobile data acquisition and synchronization are analyzed. The implementation and accomplishment of WRGIS of Jiangsu province will benefit and provide services for departments of water resources of Jiangsu province at all levels, such as providing data service - fundamental geographic information data of whole province, high resolution remote sensing images and aerial photograph, geography information data for water resources departments of Jiangsu province, as well as other data services; providing infrastructure services - high performance server and mass data storage; providing platform services for developing all kinds of professional application system based on intranet of departments of water resources.

Key words: cloud computing; geographic information services; water resources informatization; Jiangsu Province