

云计算环境下的 GIS 研究

彭义春^{1 2 3} 王云鹏² 牛熠¹

(1. 东莞理工学院城市学院 计算机与信息科学系, 广东东莞 523106;

2. 中国科学院 广州地球化学研究所, 广州 510640;

3. 中国科学院 研究生院, 北京 100049)

摘要: 云计算是近几年来在计算机领域十分热门的一个词汇。在云计算时代, GIS 也面临着新的技术革新。对云计算、云 GIS 的概念及其优势进行了探讨, 分析了 GIS 与云计算结合的必要性和可行性, 重点论述了云 GIS 的构建模式与关键技术。

关键词: 云计算; 云 GIS; GIS

中图分类号: TP393.03

文献标识码: A

文章编号: 1009-0312 (2013) 01-0017-07

随着 GIS 与主流 IT 技术的日益加速融合, GIS 的大规模、大众化应用趋势已十分明显, 涉及到多个部门和行业的 GIS 的应用的需求也越来越大, 用户对最新数据的需求也越来越快, 这样一来, 一方面, 原有的 GIS 系统需要不断更新升级, 将需要花费大量的资金在软硬件的升级上, 另一方面, 由于不同行业, 不同部门之间缺乏沟通与合作, 不同 GIS 系统之间壁垒比较分明, 数据更新与共享以及服务共享都很困难, 造成了 GIS 软硬件建设上的重复和浪费。因此, 如何解决大众化应用对超大规模并发访问给 GIS 平台架构带来的严峻挑战? 如何解决重复建设投资的问题? 如何解决我们长期面临的信息孤岛的问题? 云计算为上述问题的解决找到了新的方法, 云计算为 GIS 提供一种稳定、高效、低成本而又环保的支撑架构, 使 GIS 彻底突破既有的“专业圈子”, 将空间信息的服务和增值带给大众, GIS 的各项功能能够以弹性的、按需获取的方式提供最广泛的基于 WEB 的服务, GIS 用户可以将 GIS 应用部署在云计算供应商所提供的云计算平台中, 以实现能动态的调整软件和硬件的需求。因此, GIS 与云计算的结合必将成为行业应用和产业发展最重要的趋势之一。

1 云计算概述

近年来, 随着互联网应用的普及与深化, 网络信息与服务, 无处不在, 无时不用, 广大用户对网络信息与服务的需求也在不断提升, 同时“去软件化”趋势也在不断增强。作为一种无需自购软硬件和托管、无需关心服务提供者、只需关注所需资源和服务的全新技术, 云计算一诞生就受到人们的热情追捧。相对于复杂的海量数据处理、分布异构、硬软件更新频繁、数据安全等问题, 自从 Google 在 2006 年提出云计算的概念后, 云计算犹如一夜春风, 迅猛吹遍全球各个角落, 云计算派生出的云存储、云安全、云引擎、云推理、云服务、云娱乐不绝于耳, 各国政府也在纷纷加大对云计算的投入力度, 在国际上 Google、亚马逊、IBM、微软、SAP 和雅虎等大公司是先行者, 他们已经利用云计算技术建立了自己的云计算平台。在国内, 云计算发展势头也很迅猛, 公有云和私有云建设典型案例日益增多, 涌现出了一批如北京“祥云计算”、上海“云海计划”、苏州“风云在线”、广州“天云计划”、中国移动“大云计划”、联通“沃云计划”、电信“星云计划”等云计算项目。

什么是云计算? 目前并没有统一的标准定义, 不同的企业和专家有自己的定义, 这些定义是结合企业的产品以及商业利益来提出的。中国电子学会云计算专家委员会给出的定义: 云计算就是一种基于互

收稿日期: 2012-09-04

基金项目: 广东省科技计划项目 (2009B010800042)。

作者简介: 彭义春 (1974—), 男, 讲师, 博士生, 主要从事 GIS 与 RS 领域的研究。

联网的、大众参与的计算模式，其计算资源（包括计算能力、存储能力、交互能力等）是动态、可伸缩、被虚拟化的，以服务的方式提供，实现分享和交互，形成群体智能。广大用户基于互联网上某些结点强大的信息资源，包括存储资源、计算资源、软件资源、数据资源、管理资源，以服务的方式为个体所用。

加利福尼亚大学圣塔芭芭拉分校的博士生 Lamia Youseff 和纽约 IBM T. J. Watson 研究中心的研究员 Maria Butrico、Dilma Da Silva 在 2008 年发表了一份题为《Toward a Unified Ontology of Cloud Computing》的研究报告，该报告建立了如图 1 所示的 5 层模型，



图 1 云计算的五层模型

当前，云计算主要有三种服务模式：即基础设施即是服务（Infrastructure as a Service: IaaS）、平台即是服务（Platform as a Service: PaaS）和软件即是服务（Software as a service: SaaS）。

IaaS: 通过网络提供了数据中心、基础设施硬件和软件资源。IaaS 可以提供服务器、操作系统、磁盘存储、数据库和/或信息资源。IaaS 通常会按照“弹性云”的模式引入其他的使用和计价模式，也就是用户可只在需要时才接入这些基础设施资源，并只为自己使用的部分。SaaS 和 PaaS 都将建立在 IaaS 上。IaaS 将成为企业云计算整体战略和方案的重要基础。目前，IaaS 的代表产品有亚马逊公司的 Elastic Compute Cloud，不过 IBM、Vmware 和惠普以及其他一些传统 IT 厂商也提供这类的服务。

PaaS: 是把计算环境、开发环境等平台作为一种服务提供的商业模式。云计算服务提供商可以将操作系统、应用开发环境等平台级产品通过 Web 以服务的方式提供给用户，软件开发者可以在这个基础架构之上建设新的应用，或者扩展已有的应用，或者创建个性化的应用，也允许独立软件厂商或者其他的第三方机构针对垂直细分行业创造新的解决方案，却不必购买开发、质量控制或生产服务器。Salesforce.com 的 Force.com、Google 的 App Engine 和微软的 Azure（微软云计算平台）都采用了 PaaS 的模式。

SaaS: 是一种完全创新的软件应用模式。大家可以将它理解为一种软件分布模式，在这种模式下，应用软件安装在厂商或者服务供应商那里，用户可以通过某个网络来使用这些软件，通常使用的网络是互联网。这种模式通常也被称为“按需应变（on demand）”软件，这是最成熟的云计算模式，因为这种模式具有高度的灵活性、已经证明可靠的支持服务、强大的可扩展性，因此能够降低客户的维护成本和投入，而且由于这种模式的多宗旨式的基础架构，运营成本也得以降低。Salesforce.com、NetSuite、Google 的 Gmail、Zimbra、Zoho、IBM Lotus Live 和 SPSCoMerce.net 都是这方面非常好的例子。

数据库技术领域正在随着云计算的深化而面临变革，从传统的关系数据库、内存数据网格再到如今大火的 NoSQL；技术的变革孕生第四种云计算模式：数据即是服务（Data as a Service, DaaS）。

DaaS: 是 SaaS 部署的一个有力补充，是网络上提供虚拟存储的一种服务方式，通过基于 Web 的连接，它可以为业务用户和商业智能用户简化信息检索的过程，客户可以根据实际存储容量来支付费用。将数据库迁移到云中的另一个好处就是数据整合，通常大型企业中的数据库需要在不同的部门当中进行共用，云服务可以将它整合成一个单一的托管 DBMS，因此就减少了企业内部数据库不断扩张的问题。主要产品如：Amazon 公司提供了一个这样的服务，称为 SimpleDB。Google 的 AppEngine 提供访问 BigTable 的 DataStore API 接口以及中国电信上海公司与 EMC 合作的“e 云”等。

将 SaaS、PaaS、IaaS 结合在一起，再加上 DaaS，整个就可以构成一个云计算的概念。IaaS 是云计

算的基础, PaaS 和 DaaS 基于 IaaS 构建, 而 SaaS 可以是基于 PaaS 或者直接部署于 IaaS 之上, 其次 PaaS 可以构建于 IaaS 之上, 也可以直接构建在物理资源之上。

2 云 GIS 概述

将所有数据存储在网上, 由云计算平台提供强大的计算资源, 由全球最顶尖的专家提供数据维护和保密服务, 困扰 GIS 开发者的数据存储、管理、计算、传输问题在云计算面前已不足为道, 云计算与 GIS 的结合也成为 GIS 领域里令人关注的技术方向之一。

2.1 云 GIS 的定义

所谓云 GIS, 就是将云计算的各种特征用于支撑地理空间信息的各要素, 包括建模、存储、处理等等, 从而改变用户传统的 GIS 应用方法和建设模式, 以一种更加友好的方式, 高效率、低成本的使用地理信息资源。云 GIS 是: 一个集中的空间信息存储环境、一个以服务为基础的空间信息应用平台、一个以租赁为主要形式的商业运营模式。云 GIS 是 WebGIS、网格 GIS、分布式 GIS 的一种集合和扩展, 它支持 WebGIS、网格 GIS、分布式 GIS 等技术标准, 是在这些技术基础上融合商业云计算平台发展起来的技术。通过云 GIS, 用户无需了解、也不用担心系统应用运行的具体位置, 用户随时随地只需要一台笔记本或者一部手机, 能在 CDMA、GPRS 等无线互联网上, 连接 PDA、手机等智能移动信息终端等, 通过 Web 服务的方式提供空间数据存取与交换服务、空间信息查询服务、空间信息分析服务以及空间信息应用接口服务, 能够实现分布式跨平台的空间数据集成, 为用户提供分布式协同信息处理和按需服务。

2.2 云 GIS 的优势

云计算与 GIS 的结合, 可充分发挥云 GIS 的各种优势, 迅速扩大服务器能力, 提供安全的数据中心, 降低企业投资成本, 从而给用户带来绿色、高效的科技体验。

1) 降低了对 GIS 用户的要求

一方面, 基于云计算的 GIS 用户不需像 WebGIS 和网格 GIS 一样在自己的计算机上安装软件, 也不需要购买数据, 甚至不需要有硬件基础。用户只需要有一个网络浏览器就能以他所需的方式(文本、图像等)获取现有 GIS 软件的所有功能; 另一方面, 云计算的一个组成部分是效用计算, 用户可以按需支付自己需要的服务, 可以用最低的代价实现真正意义上的移动 GIS。

2) 降低投资和运营成本

我们知道, GIS 系统建设与维护的最大的成本也在于数据。采用云计算模式, 可以集中统一的维护数据, 通过共享的方式为所有的客户端提供数据服务和软件服务, 降低软硬件的投资和运行过程中的维护与升级成本。

采用云计算技术, 可以集中统一的管理 GIS 云数据, 可以通过共享的方式为所有的客户端提供数据服务, 使用者无需关注数据如何采集, 更新或维护, 也无需购买数据, 需要时间采用公共 GIS 云服务, 只需按流量的方式付费, 大大节省成本。

通过现有的云计算平台, GIS 企业可以租用他们的硬件服务, 利用其基础设施, 将数据或服务部署在他们提供的云平台上, 面对用户需求的不断变化时, 只需动态、弹性地增加或移除硬件设备就可以应对, 而无需增加重新部署或编码的工作量, 提高应用程序和基础设施的灵活性。

对于企业级网络 GIS 用户来说, 他们采用云计算技术, 可以大大简化 GIS 服务器的部署流程, 减少复杂的服务器管理, 他们可以通过增加或减少 GIS Server 进程的数量来快速满足不同的负载需求, 不需要投资新硬件, 节约资本。

3) 降低了 GIS 系统的开发时间和工作量

使用基于云计算的 GIS, 用户只需对云计算平台提出资源申请就可以获得超级计算机般的数据处理能力, 能够快速完成空间数据的分析处理, 而无需开发人员进行算法的优化和构建复杂的并行计算、调度模型。另一方面, 对于一个传统的 GIS 系统中的软硬件的安装与维护过程都将运行在云端, 即无需再重复考虑做这些工作, 系统建设周期将大大缩短。

4) 提高了资源利用效率

云计算平台的一个主要特点是超大规模,如现阶段的 Google 云计算平台已拥有超过 100 万的服务器(在可预见的将来其规模还会不断扩大),这些服务器处于 Google 的完全支配之下,此时如果用户提出计算申请,云计算平台就能从整体上进行全局的统筹分配(而不需要利用他人的空闲计算能力),合理利用资源,有效杜绝资源浪费。由于用户对 GIS 计算能力的要求极其不平均,如:简单的导航、最优路线计算和深入的数据挖掘相比,它们之间计算量的差别千倍不止,此时,基于云计算的 GIS 用户就可以根据自身需要向云计算平台申请合理的资源,按需使用。

5) 降低了网络的负担

网络 GIS 利用网络节点上的空闲计算机来提供所需的计算能力,在计算过程中势必涉及空间数据的传出和传回,增大了网络的负担。基于云计算的 GIS 只需用户向云计算平台提出申请,数据存储和处理都在云内部完成。在网络传输的只是最后的处理结果——一个简单的数据集,因此大大减少了网络传输的数据量。

3 GIS 与云计算结合的必要性和可行性

3.1 GIS 与云计算结合的必要性

目前,在 GIS 领域存在以下两方面的问题,一是数据方面,数据来源广泛,导致坐标和格式不能互换、数据不兼容、语义不统一、分享和共享难、互操作难等问题,如何对这些信息进行存储、管理和分析等操作也较难实现;二是应用方面,由于空间数据产生单位较少而使用者众多,任何一个单一的系统都很难拥有全部的资源和处理能力,用户无法从单一源得到所有所需的数据,同时也没有必要得到所有的数据,另外,GIS 往往需要大量的数据存储和高效的计算资源,但具有基础数据量庞大而更新频率低、并发访问数据量大等应用特点。结合 GIS 上述问题和特点,根据前文对云计算的分析,GIS 有必要也适合应用云计算技术,云计算将使 GIS 在海量数据存储、大规模计算、深度数据挖掘方面获得更加强大的优势。

3.2 GIS 与云计算结合的可行性

由于 GIS 应用的特点,非常适合采用云计算模式:①空间数据的产生单位相对较少,而数据使用者众多且多样化;②基础数据多,数据量庞大,更新频率低,适合采用云存储服务方式共享;③并发用户数很大,但每次使用量较小,适合云计算的大规模分布式计算;④需要海量数据存储,进行数据处理和数据挖掘,适合云计算的并行化分布式处理。

当前,国内外已经进行了基于云计算的 GIS 的初步尝试,国外的有:谷歌公司的 Google Earth、Google Moon 和 GoogleMars,ESRI 公司的 ArcGIS Online、ArcGIS10.1;国内的有:超图的 SuperMap GIS 6R,中地数码的 MapGIS K9 SP3,武大吉奥的 GeoCloud 等。

4 云 GIS 的构建模式

根据云计算的四种模式,云 GIS 也即从 IaaS、PaaS、SaaS、DaaS 四方面来构建,即:地理信息内容即服务(也即 DaaS)、地理信息软件即服务(也即 SaaS)、地理信息平台即服务(也即 PaaS)、地理信息基础设施即服务(也即 IaaS)。

4.1 地理信息内容即服务

地理信息内容即服务,就是把地理信息的内容作为一种服务向外提供。地理信息内容即服务是云 GIS 应用中的最低层次。地理信息内容即服务现在一般是由在线地图网站提供,这些网站提供地图信息和简单的查询服务。如:百度地图,Google Map、Bing 地图、雅虎地图等。这些地图一般提供 API,供开发者使用它们的云服务。这些 API 一般是一套 JavaScript 或 Flash 语言编写的应用程序接口,它能够帮助用户在网站中构建功能丰富、交互性强的地图应用程序。图 2 为超图公司的地理信息资源服务平台框架。

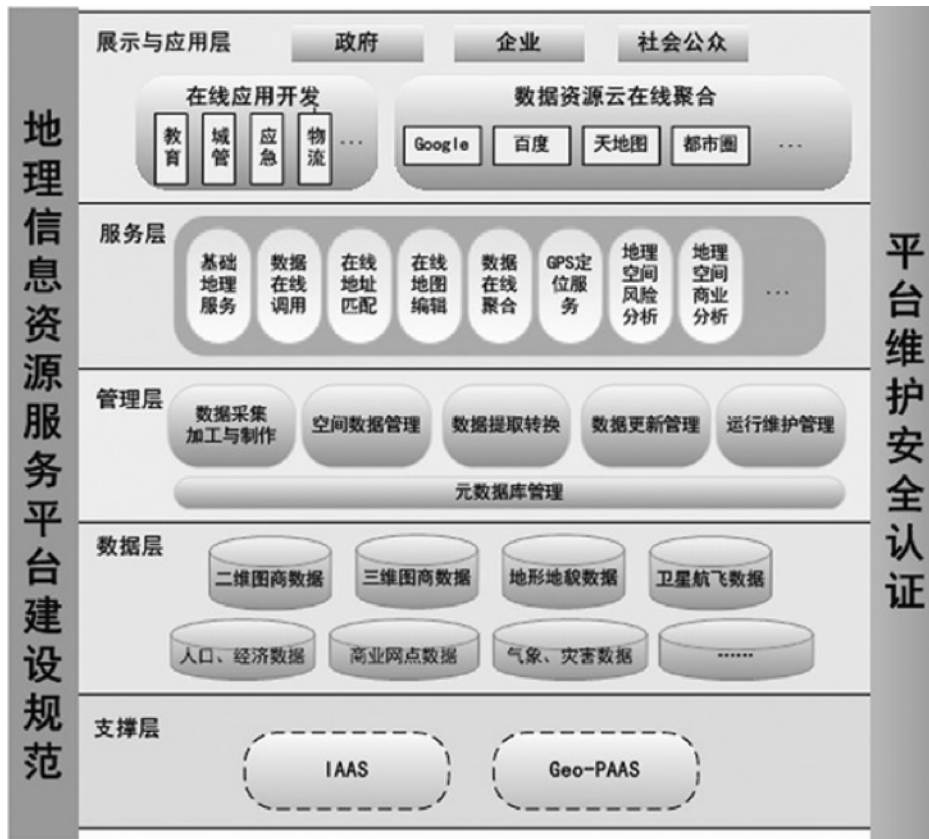


图 2 Geo-DaaS 地理信息资源服务平台框架

4.2 地理信息软件即服务

地理信息软件即服务是指利用互联网提供在线地理信息处理的服务。这种服务以往是以单机版地理信息软件完成的。主要服务内容应该包含地图发布服务、数据格式转化服务、空间分析服务等。云地理信息在线服务, 最上层采用 SOA 架构模式, 将地理信息各种服务封装成标准的 Web Services, 并纳入到 SOA 体系中进行管理和使用, 其内容包括服务接口、服务注册、服务查找、服务访问等。管理中间件是云地理信息管理部分, 负责对用户对地理信息进行计费, 负责负载均衡, 地图切片服务等。最后任务有 GIS Server 提供底层数据服务。GIS 系统开发人员可以利用云计算的技术和模式来构建自己的解决方案, 如利用 GIS 平台商的 PaaS 服务来研发自己的 GIS 解决方案, 也可利用云计算的模式为自己的客户提供服务。如超图地理信息云服务之云 GIS 主机服务: 提供云在线虚拟服务器主机和 SuperMap iServer 平台租用服务。GIS 平台以租代买, 优化技术实现手段, 降低系统初始化建设投入和运维成本。支持用户开发在线 SaaS 应用; 支持用户将 Geo-SaaS 应用服务部署在云上托管运营; 支持用户数据与云地图数据的叠加等。

4.3 地理信息平台即服务

地理信息平台即服务即把地理信息整个开发环境作为服务向外提供。地理信息平台即服务是提供 GIS 的一个开发平台服务。目前国外, 提供平台即服务较为著名的是 Google App Engine。地理信息系统开发者可以在 Google App Engine 上开发地理信息软件, 并运行在 Google 的基础设施上。对于国内 GIS 平台提供商而言, 基本上只能利用云计算的模式为自己的客户和合作伙伴提供服务。例如, GIS 平台商可将自己的 GIS 平台打造成 PaaS 服务。这样, 合作伙伴在开发自己的各种 GIS 解决方案时, 就可以不需要购买 GIS 平台的许可证费用, 而是直接采用云计算的租用模式, 甚至可以不用本地部署与安装 GIS 平台, 直接在网上就能进行 GIS 解决方案的开发。这将为 GIS 解决方案合作伙伴带来极大的便利, 节省大量的成本。

4.4 地理信息基础设施即服务

地理信息服务的构建可以运行在其他商业公司构建的云基础设施中。目前, 提供硬件基础设施服务

的有亚马逊、IBM，以及一些电信运营商，如中国电信、中国移动，这些企业正在或已经搭建了基础设施服务环境，并以此为基础提供相应的计算资源或弹性租赁服务。这是“云”模式的基础，地理信息基础设施即服务是地理信息软件即服务，地理信息内容即服务（即 PaaS、SaaS 和 CaaS）的基础，因此，基础设施即服务对 GIS 而言，这种基础环境和服务模式是不可或缺的。通过云计算技术，云 GIS 用户可以租用商业云计算平台的软件和硬件资源，把地理信息服务部署在云中提供。目前，提供这种云设施的公司主要是 Amazon。它提供了弹性云计算（Elastic Compute Cloud, EC2）和简单存储服务（Simple Storage Service, S3）两种租用方式。这种存储服务按照每个月类似租金的形式进行付费。

5 云 GIS 的关键技术

云 GIS 必须贯穿数据、软件、开发等几个层面用户才能真正随时获取所需的 GIS 资源，因此云 GIS 在云计算平台与数据互操作，GIS 空间数据存储、管理、处理和分析，终端接入等方面具有自身独特的技术。

5.1 云计算平台与数据互操作技术

云 GIS 要实现从单机到局域网、互联网应用的跨越，能够以一套统一的体系架构满足私有云和公共云环境，并实现私有云、公有云的一体化连通和交互，同一技术能够支持服务器、桌面、Web 和移动端的全面应用。它可以支持单点发布、自动同步、频度统计和自动优化，支持云内部的数据互操作、私有公有云的互操作和云中心之间的互操作。为了实现上述操作，必须制定云计算互相操作和集成标准、云计算的服务接口标准和应用程序开发标准、云计算不同层面之间的接口标准、云计算服务目录管理、不同云之间无缝迁移的可移植性标准、云计算商业指标标准、云计算架构治理标准、云计算安全和隐私标准等。云计算的技术架构采用多层结构，从上到下：业务逻辑层、应用层、分布式文件和操作系统层、虚拟化层、硬件层和数据中心基础设施，如图 3 所示。

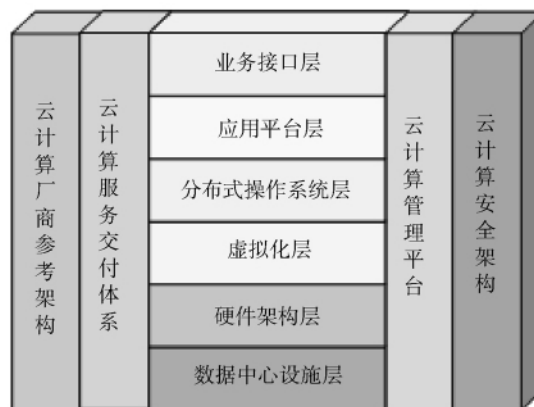


图3 云计算技术架构

5.2 GIS 空间数据存储、管理、处理和分析技术

存储技术：空间数据存储从文件系统发展到分布式文件系统以及完全基于互联网技术的云存储系统；空间数据库从企业级数据库向分布式空间数据库发展，并在将来支持 BigTable、HBase、NoSQL 等数据库技术存储和管理空间数据；支持通过标准的空间数据库接口和 REST 接口进行统一访问。当前主要有谷歌的非开源的 GFS（Google File System）和 Hadoop 开发团队开发的 GFS 的开源实现 HDFS（Hadoop Distributed FileSystem）。大部分 IT 厂商，包括 yahoo、Intel 的“云”计划采用的都是 HDFS 的数据存储技术。未来的发展将集中在超大规模的数据存储、数据加密和安全性保证、以及继续提高 I/O 速率等方面。

数据管理：采用虚拟化技术实现空间数据库的统一管理，支持在系统之间、部门之间、层级之间快速迁移和自动同步数据，具有离线应用和在线更新技术，实现分布式、多级别、支持多终端的空间数据保障流程。数据管理往往采用数据库领域中列存储的数据管理模式，将表按列划分后存储。当前主要有谷歌的 BigTable 数据管理技术和 Hadoop 的类似 BigTable 的开源数据管理模块。

数据处理: 具有面向任务的异步空间数据处理架构, 支持大型集群的并发处理和流程控制, 支持长时间运行、长事务处理, 支持移动终端操作处理大型空间数据库, 具有处理流程的可视化设计和运行状态的实时监控功能, 可以跨平台、跨地域整合空间数据的处理流程, 并能够即时将处理结果进行发布。

空间分析: 具有统一的空间分析框架和丰富的空间分析模型, 建立标准化的分析模型库, 支持空间分析流程的快速构建和自动化运行、分析结果的即时发布。

5.3 终端接入技术

基于云计算的 GIS 平台的最终目的是让用户只要通过一个浏览器就可以获得 GIS 的所有功能。为满足海量数据在不同网络中通畅的传输, 要求系统具有统一的内核和接口, 能实现多类型桌面、Web 和移动客户端的服务访问, 最终达到数据同步存取、处理结果一致和用户体验优雅的效果。

6 结语

随着云计算技术的发展和 GIS 应用的不断深入, 云计算与 GIS 将会进一步的进行融合, 从而使得 GIS 以云计算的形式向用户提供服务, 并逐渐发展成熟。不过真正意义上的基于云计算的 GIS 还有待深入研究, 适合云计算平台 GIS 还有一段漫长的路要走, GIS 冲上云端还要面临很大的挑战。但云时代的到来已是一股不可阻挡的发展潮流, 基于云计算的 GIS 必将是未来 GIS 发展的主要方向。

参考文献

- [1] Youseff L, Butrico M, Da Silva D. Toward a Unified Ontology of Cloud Computing[C]. Grid Computing Environments Workshop, 2008: 1-10
- [2] 李永楠, 李德忠. 浅谈云计算优势、应用及发展[J]. China New Technologies and Products, 2011, 03.
- [3] 戴立乾, 陈娜. 浅谈云计算时代下 GIS 的发展[J]. 安徽农业科学, 2009, 37(31): 15556-15557.
- [4] 杨柳. 基于云计算的 GIS 应用模式研究[D]. 郑州: 河南大学, 2011.
- [5] 王尔琪. 云计算与 GIS 技术革新[J]. 新经济导刊, 2011(10): 83-87.
- [6] 卢敏. 软件世界: 云中漫步从 GIS 开始[EB/OL]. [2011-01-27]. [2012-06-27] <http://www.esrichina-bj.cn/2011/0127/730.html>.
- [7] Esri 中国(北京)有限公司, 译. ArcGIS 与云计算技术[G/OL]. [2012-05-25] <http://www.docin.com/p-97128111.html>.
- [8] 王尔琪. 云 GIS 技术路线与发展探析[J/OL]. 超因通讯, 2011, 28. [2012-06-29] <http://www.supermap.com.cn/sup/xwtxpage.asp?orderID=416>.

GIS Under Cloud Computing Environment

PENG Yi-chun^{1 2 3} WANG Yun-peng² NIU Yi¹

(1. City College of Dongguan University of Technology, Dongguan 523106, China;

2. Guangzhou Institute of Geochemistry, Chinese Academy of Sciences, Guangzhou 510640, China;

3. Graduate University of the Chinese Academy of Sciences, Beijing 100049, China)

Abstract Cloud computing in recent years is a very popular word. In the era of cloud computing, GIS also faces a new technological innovation. This paper discusses cloud computing, the concept and advantage of cloud GIS, analyzing the necessity and feasibility of the Combination of GIS and cloud computing, and relating the constructing mode and key technology of cloud GIS.

Key words Cloud Computing; Cloud GIS; GIS