

云计算在蔬菜安全预警和追溯平台中的应用

刁海亭¹,郑成良¹,郑晓梅²

1. 山东农业大学信息科学与工程学院, 山东 泰安 271018
2. 山东省五莲县国土资源局, 山东 五莲 262300

摘要: 蔬菜作为中国出口创汇的重要作物和广大人民群众日常生活的必需品,其质量安全非常重要。传统方式建立的蔬菜质量安全监管系统比较庞大,网络终端用户访问的速度较慢。本文讨论了云计算的概念、服务模型及国内外应用,介绍了反向代理服务器 Nginx 的背景及应用。利用 Nginx,结合 C#开发语言,借助 3 台电脑,探讨了蔬菜安全预警和追溯平台中图片数据的分布式动态存储和管理。云计算技术的应用减少了溯源数据库本地服务器的存储数据,提升了用户访问网站的速度,降低了平台维护成本。

关键词: 云计算; 蔬菜; 安全预警

中图分类号: TS207.7; TP315

文献标识码: A

文章编号: 1000-2324(2016)05-0768-05

The Application of the Cloud Computing in Vegetable Safety Early Warning and Traceability Platform

DIAO Hai-ting¹, ZHENG Cheng-liang¹, ZHENG Xiao-mei²

1. College of Information Science and Engineering/Shandong Agricultural University, Tai'an 271018, China
2. Wulian Land and Resources Bureau of Shandong, Wulian 262300, China

Abstract: Vegetables, as China's important export crops and the people's necessities in daily life, their quality and safety are very important. The supervision system for traditional vegetable quality and safety is relatively large, so the access speed of the network terminal users is slow. The concept of cloud computing, service model and application in domestic and abroad were discussed in this paper and also introduced the background and the application of the reverse proxy server Nginx. Using Nginx and C# development language with 3 computers, the distributed dynamic storage and management of the picture data in vegetable safety early warning and traceability platform were explored in this paper. The use of cloud computing technology reduced data quantity of the local database, enhanced access speed of the user and reduced maintenance cost of the platform.

Keywords: Cloud computing; vegetable; safety early warning

蔬菜从农田到餐桌要经过生产、加工、储存、运输、销售等环节,整个产业链工艺、流程错综复杂,存在数据多源、异构。基于这些海量数据建立相关系统往往比较庞大,不仅占用大量存储空间,还降低了系统的效率。因此,需要采取措施应对日益庞大的蔬菜管理系统。目前涌现的云计算(Cloud Computing)技术正是解决这一问题的有力武器。云计算被视为信息技术的第三次浪潮,已从新兴技术发展成为当今的热点技术,它的分布式存储管理理念在国际上和国内都引起了研究的热潮。它将各种资源集中、结合起来提供各种服务,利用虚拟化技术向用户提供高效、可靠和稳定的计算和存储服务,解决海量数据的存储与处理问题。

1 云计算

1.1 云计算概念

云计算真正概念的提出是 2006 年 8 月 9 日的谷歌(Google)搜索引擎大会(SES San Jose 2006)上,当时的谷歌首席执行官埃里克·施密特(Eric Schmidt)第一次提出了“云计算”的概念。此后这一思想引起了全世界的重视,关于云计算的概念也众说纷纭,至少有几十种。倪光南院士指出“人们感受到的丰富多样的云服务属于商业模式的范畴,而可以按需提供强大计算资源的云计算平台则属于技术的范畴”^[1]。

简而言之,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(如计算资源、存储设备、应用程序等)的计算模式(来自美国国家标准与技术研究院的定义)。这个概念明确

收稿日期: 2015-03-14

修回日期: 2015-04-15

基金项目: 国家科技支撑计划课题资助(2012BAK17B05)

作者简介: 刁海亭(1978-),女,讲师,博士,主要从事地理信息系统及国土资源信息化的研究. E-mail:htdiao@sdau.edu.cn

了云计算的实质就是在互联网上将 IT 资源当作服务来提供。在云计算的环境之下，一切都是服务。

1.2 云计算的服务模型

云计算有三种主要的服务模型，图 1 是三种服务模型的层次关系^[2]。

(1) SaaS: Soft-ware as a Service，软件即服务。将应用程序作为服务提供给用户，面对的是普通用户；

(2) PaaS: Platform as a Service，平台即服务。将服务器平台或者开发环境作为服务提供给用户，面对的是开发人员；

(3) IaaS: Infrastructure as a Service，基础设施即服务。将服务器、存储和网络硬件以及相关软件等资源作为服务提供给用户，面对的是系统管理员^[3]。

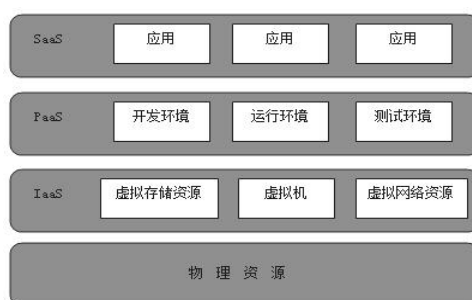


图 1 云服务模型层次关系

Fig.1 The hierarchy relationship of cloud service models

1.3 云计算的国内外研究

自从云计算的概念被提出以来，国外许多国家如美国、日本、德国、韩国等都进行了大量的研究；一些公司如：Google、IBM、Amazon、微软、雅虎等都是云计算的先行者，在云计算上做了大量的实践工作。Google 的技术核心就是云计算，其早期发布的一系列关于分布式文件系统、并行计算、数据和分布式资源管理方面的文章，为全球云计算的发展奠定了技术基础，可以说 Google 是云计算的先驱。IBM 于 2007 年 8 月推出“蓝云（Blue Cloud）”计划，为企业客户搭建分布式、可通过互联网访问的云计算体系。Amazon 于 2008 年 8 月推出弹性计算云（Elastic Compute Cloud, EC2）和简单存储服务（Simple Storage Service, S3），在云计算的发展中也起到了重要的作用。微软于 2008 年 10 月推出基于云计算的操作系统—Windows Azure，为开发者提供一个平台，帮助他们在 PC、Web、云服务器、数据中心上运行应用程序^[6]。我国在云计算上起步较晚，在北京、上海、深圳、杭州和无锡等 5 个城市先行开展云计算创新发展试点示范工作，随后向全国各城市推开。

随着云技术的广泛应用，云技术和 GIS 的结合也开始日益得到重视，如美国环境系统研究所公司（Environmental Systems Research Institute, ESR）ArcGIS 10.1 推出的私有云 GIS 服务在美国密苏里州西普兰市的招商引资、印尼国家测绘局的 NSDI 项目中都有典型应用；中地数码公司推出的 MapGIS K9 SP3 采用了面向服务的悬浮式体系架构，在数字郴州共享服务平台、数字宜宾共享服务平台等项目得到了应用；中国超图公司研发的 SuperMap GIS 6R（2012）支持虚拟化、64 位 CPU、支持二三维一体化、跨平台等，在福建省地理信息云服务平台、全球建筑机械管理 GIS 云平台等得到了很好的研究。

2 Nginx

2.1 Nginx 介绍

随着网络技术的提高和应用的普及，各种网站也越来越复杂。而一个网站同时为成百上千甚至是几百万用户同时服务的情况也并不少见，这就导致客户端浏览网站的速度变慢。因此，为了能够持续性处理随着日益增加的用户量而导致的不断增加的负载以及获得更高的并发性，一个网站必须基于一系列非常高性能的模块来构建。Web 服务器需要能够通过非线性扩展来满足每秒钟不断增长

的并发连接和请求数。Nginx 正是能满足这种需求的 Web 服务器。

Nginx 是 2004 年俄罗斯人 Igor Sysoev (塞索耶夫) 为俄罗斯访问量第二的 Rambler.ru 站点开发的一款高性能的 HTTP 和反向代理服务器, 也是一个 IMAP/POP3/SMTP 代理服务器。Nginx 本身就可以托管网站, 进行 HTTP 服务处理, 也可以作为反向代理服务器使用^[7,8]。

随着 Nginx 的功能越来越强大, 其在国内外的普及率也越来越高。在俄罗斯, 大约超过 20% 的虚拟主机平台采用 Nginx 作为反向代理服务器。在国内, 已经有淘宝、新浪、网易、豆瓣、迅雷在线、六间房等多家网站使用 Nginx 作为 Web 服务器或反向代理服务器。本文利用 Nginx 探讨平台中图片数据分布式动态存储, 提升用户访问网站的速度。

2.2 Nginx+IIS 服务器搭建服务器集群

利用 Nginx+IIS 服务器搭建服务器集群, 架构图如图 2 所示。

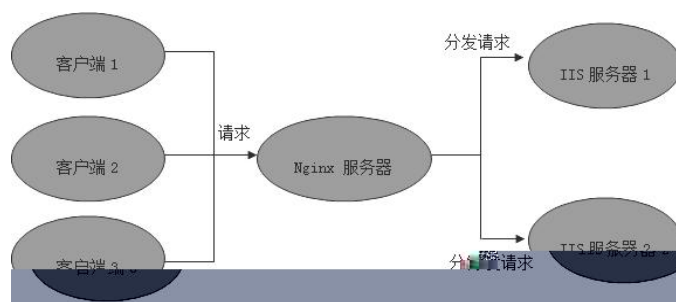


图 2 Nginx+IIS 服务器搭建服务器集群
Fig.2 The servers cluster from Nginx+IIS server

3 云技术在平台中的应用

本文前期开发的蔬菜安全预警与追溯平台基于网络设计, 数据量较大, 终端用户访问的速度受到影响。平台中保存有大量蔬菜产地图片, 本文实验借助 3 台电脑, 探讨蔬菜安全预警和追溯平台中图片数据的分布式动态存储和管理。

3.1 环境设计

电脑的环境设计如下:

(1) Nginx 服务器:

硬件: 华硕笔记本、内存 4G

软件: Nginx1.4.7

局域网 IP:192.168.191.1:8080

(2) Image 服务器:

硬件: 联想台式、内存 4G

软件: IIS+imgServerX2

局域网: 192.168.191.2:80; 192.168.191.2:90

说明: 可弹性添加 Image 服务器节点个数

(3) 主站服务器:

硬件: 联想笔记本 (ThinkPad Edge)、内存 4G

软件: IIS+主站点

局域网 IP:192.168.191.3:80

说明: 可弹性添加主站节点个数

3.2 分布式存储图片服务器状态设计

Web 服务器需要及时掌握所有图片服务器的状态和信息, 才能动态决定把图片保存到哪一台图片服务器。因此, 需要把所有的图片服务器的状态信息全部记录到数据库服务器中, 记录图片服务

器信息和状态的表格如 1 所示。

表 1 图片服务器状态信息表
Table 1 Image server state information table

字段 Word information	数据类型 Data type	长度 Length	是否允许 Null 值 Null values yes or no	字段描述 Description
ServerId	int	4	否	自增字段主键
ServerName	Nvarchar	32	否	图片服务器名称
ServerUrl	Nvarchar	100	否	图片服务器地址
PicRootPath	Nvarchar	100	否	图片存储物理路径
MaxPicAmount	Int	4	否	存储最多图片数
CurPicAmount	Int	4	否	当前已保存图片数
FlgUsable	Bit	1	否	标识图片服务器是否可用

状态信息表中的 ServerId 字段为主键自增列，唯一代表一条图片服务器纪录。ServerName 字段记录服务器的名称，方便管理员识别该记录代表哪台服务器。ServerUrl 字段标识了图片服务器上图片主目录的 URL 根路径。PicRootPath 字段标识了保存图片的物理主目录。MaxPicAmount 字段表示图片服务器能保存的最大图片数，该数可以根据图片服务器的硬件配置和性能以及用户实际需要而进行动态调整。CurPicAmount 字段表示当前已保存的图片数，当 CurPicAmount ≥ MaxPicAmount 时系统将不再把图片上传到该服务器。FlgUsable 字段表示图片服务器是否可用。

表 2 为平台中涉及到的图片相关信息。

表 2 图片信息表
Table 2 Image information

字段 Word information	数据类型 Data type	长度 Length	是否允许 Null 值 Null values yes or no	字段描述 Description
Id	int	4	否	自增字段主键
ImgName	nvarchar	100	否	图片的路径
ImgServerID	int	4	否	图片服务器 Id
X	float		是	图片所在地 X 坐标
Y	float		是	图片所在地 Y 坐标

3.3 平台分布式架构实现

由于图片保存在物理硬盘上，访问图片需要频繁进行 I/O 操作，因此随着并发用户的数量越来越多，I/O 操作就会影响整个系统的性能，最终影响用户访问速度。本文探讨采用图片数据分布式动态存储及负载均衡的方案，以解决用户访问平台中蔬菜产地图片信息时的速度问题。该方案只需增加很少的硬件成本，即可提升网站的访问速度，并且可以根据需要动态调整图片服务器的数量及图片的存储目录，确保系统具有可扩展性和伸缩性。

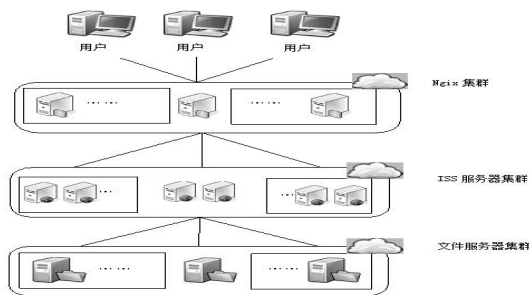


图 4 分布式平台存储设计图
Fig.4 The design for the storage of distributed platform

结合 Nginx 与文件分布式存储方案，本平台的分布式存储设计如图 4 所示，其中方框中的服务器代表可扩展的节点。平台实施过程中，通过 Nginx，将用户的请求分发给 IIS 服务器。本文在设计时，根据实际情况选择了一台 Nginx 服务器，两台 IIS 服务器，在实际应用中可以建立 Nginx 服务器集群。

3.3.1 图片文件上传 当生产用户上传产地图片时，Web 服务器需要及时掌握所有图片服务器的状态

和信息才能动态决定把图片保存到哪一台图片服务器。因此，需要把所有的图片服务器的状态信息全部记录到数据库服务器中，记录图片服务器信息和状态如表 1 所示。但由于 B/S 架构本身技术限制，图片无法通过 Web 服务器直接上传到不同的图片服务器，因此需要在所有图片服务器上部署一个 Web Service 以便 Web 服务器可通过调用不同图片服务器上的 Web Service 执行保存。

从状态表筛选出可用的图片服务器集合记作 A，获取集合的总记录数 N。然后用随机函数产生一个随机数 R1，并用 R1 与 N 进行取余运算，记作 $I=R1\%N$ 。则 A[I]即为要保存图片的图片服务器。

3.3.2 图片浏览 客户端用户通过浏览器向 Web 服务器发出浏览某页面的请求，Web 服务器从数据库服务器中获取该页面的所有图片 URL 信息，并根据 URL 信息去搜索图片服务器的状态信息表，判断该 URL 所指向的图片服务器的状态字段 FlgUsable，若 FlgUsable=false 表示该图片服务器当前因某种原因处于不可用状态，则把该图片的 URL 替换成 Web 服务器上保存的一个默认图片的 URL，否则把该 URL 直接返回给客户端。客户端再根据图片的 URL 路径自动从不同的图片服务器上下下载并显示相应的图片。由于图片 URL 路径直接指向具体的图片服务器，因此需要在每个图片服务器的保存图片的主目录上建立一个 Web 站点。由于客户端浏览器所需要的图片是从多个图片服务器上直接下载，因此浏览器可以并发地从多台服务器上同时下载图片，这样就缩短了图片下载时间，同时也减轻了 Web 服务器的 I/O 请求及性能压力，提高了网站的访问速度。实现效果如图 5 所示，向 192.168.191.1:8080 地址发送请求，Nginx 接到请求后分发给 192.168.191.3（主站点），然后主站点向两台图片服务器发送请求，由 192.168.191.2:80 和 192.168.191.2:90 做出响应。



图 5 分布式平台实现效果图
Fig.5 The implementation effect of the distributed platform

4 结论

本文结合云计算技术利用 Nginx 探讨了蔬菜安全预警和追溯平台中图片数据的分布式动态存储，提升用户网站的访问速度。对于 Web 服务器而言，用户对图片信息的访问是很消耗服务器资源的。页面中图片越多 Web 服务器受到的压力也就越大。平台设立单独的图片服务器来专门存放图片，可以根据需要动态调整图片服务器的数量及图片数据存储目录，提高用户的访问速度，确保所建平台具有可扩展性和伸缩性。但蔬菜管理系统中的数据多源异构，本文仅仅探讨了图片数据的分布式动态存储，下一步将对其它数据格式的云存储展开研究。

参考文献

- [1] 姚宏宇,田溯宁.云计算:大数据时代的系统工程[M].北京:电子工业出版社,2013
- [2] 林 利,石文昌.构建云计算平台的开源软件综述[J].计算机科学,2012,39(11):1-7,28
- [3] 孙香花.云计算研究现状与发展趋势[J].计算机测量与控制,2011,19(5):998-1001
- [4] 修长虹,梁建坤,董鸿晖.云计算技术综述[J].网络安全技术与应用,2012(3):9-11
- [5] 姚有真.云计算及其发展趋势[J].通信管理与技术,2011(4):13-15
- [6] 刘 越.云计算综述与移动云计算的应用研究[J].信息通信技术,2010(2):14-20
- [7] 田纯青.利用 Nginx 实现基于 URI 的 Web 负载分配[J].现代计算机:专业版,2009(7):187-191
- [8] 聂 鹏.基于 Nginx 的云计算访问控制网关的设计与实现[D].北京:北京交通大学,2011