

云计算中的 大数据技术与应用

YUNJISUAN ZHONGDE
DASHUJU JISHU YU YINGYONG

梁 凡 / 著

 吉林大学出版社

图样设计(GB)标准

：第一，普通应用已不能满足大规模的中算长

吉林出版集团，2018.3

ISBN 978-7-5633-3183-1

云计算中的 大数据 技术与应用

YUNJISUAN ZHONGDE
DASHUJU JISHU YU YINGYONG

梁 凡 / 著

全书共分5章，包括云计算、大数据、大数据存储、大数据处理、大数据应用、云存储、云计算技术中的网络安全问题、云计算人才培训等。本书可作为高等院校计算机专业及相关专业的教材，也可供从事云计算、大数据工作的工程技术人员参考。

在本书的编写过程中，得到了作者所在单位的领导和同事的大力支持，以及部分同事的帮助，在此表示衷心的感谢。

由于作者水平有限，书中难免存在不足之处，恳请广大读者批评指正。

 吉林大学出版社

吉林出版集团

图书在版编目(CIP)数据

云计算中的大数据技术与应用/梁凡著. —长春:
吉林大学出版社, 2018. 3
ISBN 978-7-5692-2166-4

I. ①云… II. ①梁… III. ①云计算—研究②数据处理—研究 IV. ①TP393.027②TP274

中国版本图书馆 CIP 数据核字(2018)第 089029 号

书 名 云计算中的大数据技术与应用
YUNJISUAN ZHONG DE DASHUJU JISHU YU YINGYONG

作 者 梁凡 著
策划编辑 李伟华
责任编辑 李伟华
责任校对 安 萌
装帧设计 墨创文化
出版发行 吉林大学出版社
社 址 长春市人民大街 4059 号
邮政编码 130021
发行电话 0431-89580028/29/21
网 址 <http://www.jlup.com.cn>
电子邮箱 jdcbs@jlu.edu.cn
印 刷 北京富泰印刷有限责任公司
开 本 787mm×1092mm 1/16
印 张 9.5
字 数 210 千字
版 次 2018 年 6 月第 1 版
印 次 2018 年 6 月第 1 次
书 号 ISBN 978-7-5692-2166-4
定 价 38.00 元

版权所有 翻印必究

前言

QIANYAN

可以说,我们身处于一个大数据时代。利用大数据,我们可以建立城市规划、预测犯罪发生、预测禽流感暴发……“大数据”一词,既代表着信息技术的发展与创造,更代表了崭新的生产、生活方式。

作为大数据的基础技术支撑——云计算,它是以网络为基础,以资源服务为目标的一种计算模式,其“侧重计算效能提升,通过虚拟化、分布式、并行计算等多种技术手段,解决海量数据环境下的计算复杂性和时效性问题”。而大数据本质上是“一种信息价值获取方式,侧重数据的应用分析,采用数据存储、数据处理、分析应用及数据展现等多个交叉学科技术,解决各种海量、异构、多模态数据的价值获取问题”。简单来说,云计算解决了“如何算”,大数据则解决了“如何用”。大数据在互联网、电信、企业、物联网等行业还有很大的发展空间,大数据问题将挑战企业的存储架构及数据中心基础设施等,也会引发云计算、数据仓储等的应用的连锁反应。

全书共分9章,包括云计算、大数据、大数据存储、大数据处理、数据查询分析计算系统、云存储、云计算技术中的网络安全问题、云计算入侵检测,以及大数据技术的应用。

在本书的编写过程中,得到了作者所在单位南宁职业技术学院各级领导的大力支持与帮助,以及部分同事提出许多宝贵的意见,还提供了有效案例。同时,也参考了同行们无私分享在互联网上的大量资料,在此一并表示衷心的感谢!

由于作者水平有限,书中难免存在不足之处,恳请读者批评指正。

编者:梁凡

2018年2月

目 录

第1章 云计算	1
1.1 云计算简介	1
1.2 云计算的技术分类及优点	6
1.3 国内外云计算技术的发展状况及供应商	9
第2章 大数据	16
2.1 大数据简介	16
2.2 主要的大数据处理系统	24
2.3 大数据处理的基本流程	34
第3章 大数据存储	36
3.1 非结构化数据	37
3.2 (半)结构化数据	46
第4章 大数据处理	50
4.1 离线数据处理	51
4.2 实时数据处理	66
第5章 数据查询分析计算系统	79
5.1 HBase 的搭建与使用	80
5.2 Hive 的搭建与使用	85
第6章 云存储	90
6.1 云存储技术	91
6.2 云存储的应用及面临的问题	94
第7章 云计算技术中的网络安全问题	98
7.1 云计算引发的新的安全问题	99

7.2	云计算技术层面怎么应对网络安全	100
第8章	云计算入侵检测	104
8.1	云计算入侵检测系统简介	104
8.2	云计算网络入侵检测的分类	105
8.3	云计算网络安全入侵检测	105
第9章	大数据技术的应用	113
9.1	大数据技术在电子商务领域的应用	113
9.2	大数据及云计算技术在智慧校园中的应用研究	116
9.3	大数据在网络购物中的应用	119
9.4	网络虚拟化	128
参考文献		146



第 1 章 云计算

国家重点研发计划“云计算和大数据”重点专项 2017 年度立项项目“大数据驱动的人类智能感知与情感交互关键技术”项目启动会在合肥召开。中科院科技促进发展局相关领导、中国科学技术大学朱长飞副校长、总体专家组责任专家、项目咨询专家、课题负责人、专项办有关人员出席会议。项目负责人吴枫教授主持了会议。

会上，专项办有关同志对云计算和大数据重点专项管理情况、专项组织实施总体考虑和专项总体管理原则进行了介绍，并对后续工作提出了相应的要求。朱长飞副校长作为项目牵头单位给 7 位专家颁发了专家聘书。紧接着，项目负责人汇报了项目的研究背景与意义、研究内容与关键科学问题、项目技术路线和实施计划、项目管理机制、项目成果呈现形式及测试方法等。最后，专家们进行了研讨，建议将项目主线进一步聚焦，将项目研究子任务形成一个闭环，提出了更为具体的应用需求，应用范围更为明确，并考虑项目成果的展示方式。

通过本次会议，对项目课题及参与人员的任务进行了进一步的明确分工，为下一阶段项目实施方案指明了论证方向，确保项目能够顺利实施以及目标能够按期完成。

1.1 云计算简介

本节将介绍现代企业为什么需要云计算、云计算的含义及特点、云计算的发展历史。

1.1.1 为什么我们需要云计算

以前由于条件的限制，个人使用计算机软件与企业建立和开发系统，都需要一定的预算。例如，个人首先需要要在自己的电脑上安装各种软件，这些软件有些免费，而有些软件需要额外付费。即使是不经常使用的付费软件，也需要购买后才能使用。而对于企业来说，如果需要建立一套软件系统，除了需要购买硬件等基础设施外还需要购买软件的许可证，同时，需要由专门的人员维护。随着企业规模的扩张和需求的增加，各种软、硬件设施需要通过不断升级来完成工作、获取盈利、提高效率。但事实上，企业真正所需要的并不是计算机的硬件和软件本身，如何通过租用和共享来减少



支出,对企业来说,真的是再好不过。

部分服务提供商抓住这个机会,纷纷开始思考:为给个人和企业用户提供更多的便捷,是否可以提供某种服务,例如,将软件以租赁的方式提供给用户?这样,用户只需要交纳少量租金,就可使用这些软件服务,不仅能够节省许多购买软、硬件的资金,还能够及时更新服务资源。在计算机应用中推广这种服务模式的想法最终导致了云计算的产生。

云计算改变了人们的生活和工作方式,为人们的生活提供了无限的可能。用户的计算机只需要通过浏览器给“云”发送请求然后接收数据,就能便捷地使用云服务。这样一来,计算机不再需要过大的内存,甚至也不需要购买硬盘和安装各种应用软件,但仍然能获得海量的计算资源、存储空间和各种应用软件等。



小资料

赵阿姨退休后,开始了自己的旅游之路,游山玩水间少不了拍照留念,然而手机里存储越来越多的照片成为了赵阿姨的烦恼。刚开始她把照片都上传到 QQ 空间里,这既是一种分享,又能节省手机的存储空间,但是她又有些担心害怕 QQ 号被盗,上传的照片存在丢失的风险。于是,赵阿姨的孩子又帮她注册了百度云,将所有的照片都存入了百度云中,这样就相当于上了“双保险”。赵阿姨不禁感慨:“以前,照相要到照相馆,照完的照片还要洗出来,放进相册里,底片也要保存好。现在有了手机、电脑、‘云’,照片保存方便多了!”

1.1.2 云计算的含义及特点

由于人们对云计算的认识还不够全面,云计算也在不断地发展和变化中,因此目前云计算并没有非常严格和准确的定义。

1. 云计算的含义

在计算机还没有普及的 20 世纪 60 年代,就有科学家曾经提出“计算机可能变成一种公共资源。”2006 年,谷歌首席执行官艾里克·施密特在搜索引擎大会上第一次提出了云计算的概念。

最近几年,云计算这一概念经常成为各大报道的头条,虽然大部分人对云计算的真正含义还不是很了解,但是不得不承认,云计算技术在社会生活的诸多领域中已经开始运用。云计算是一种具有开创性的新计算机技术,它是传统计算机和网络技术发展到现在一定阶段融合的产物。通过互联网提供计算能力即,就是云计算的原始含义。

2012 年,国务院政府工作报告将云计算作为国家战略性新兴产业给出了定义:“云计算是基于互联网服务的增加、使用和交付模式,通常涉及通过互联网来提供动态、易扩展且经常是虚拟化的资源。”



云计算主要包括两个方面：一方面是厂商通过分布式计算和虚拟化技术搭建数据中心或超级计算机，主要有免费和按需租用的方式向技术开发者或者企业客户提供数据存储、分析以及科学计算等服务。另一方面是指厂商建立网络服务器集群，把在线软件使用、硬件租借、数据存储、计算分析等不同类型的服务提供给不同类型的客户。云计算能提供更多的厂商和服务类型。云计算的应用和影响力日益扩大，并成为新兴战略性产业之一。云计算体系结构如图 1-1 所示。

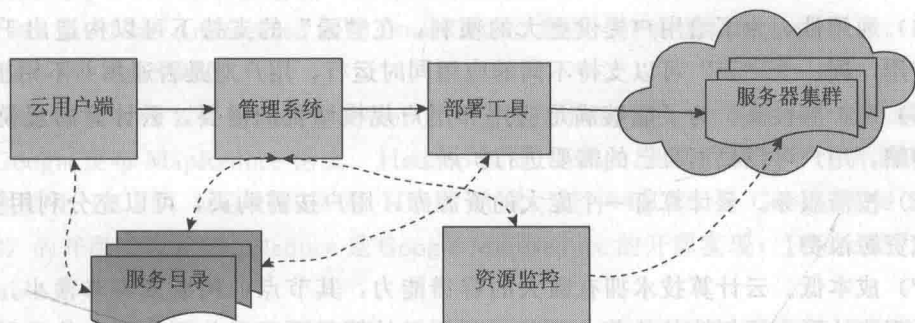


图 1-1 云计算体系结构

2. 云计算的组成

云计算的“云”是指存在于互联网上的服务器集群上的资源，包括硬件资源和软件资源。硬件资源主要指服务器、存储器、CPU 等，软件资源主要指应用软件、集成开发环境等。云计算是一种“共享”，本地计算机 A 只要通过互联网发送一个需求信息，远端就会有成千上万的计算机（B、C、D……N）为其提供资源，并将搜索结果返回到本地计算机 A，本地计算机 A 几乎不用做什么，这是因为云计算提供商所提供的计算机群能够完成所有的处理。

在云计算环境下，用户形成了“购买服务”的使用观念，他们面对的不再是复杂的硬件和软件，而是最终的服务。用户不需要购买硬件设施实物，节省了购买费用，同时可以节省等待时间（漫长的供货周期和冗长的项目实施时间），只需要把钱汇给云计算服务提供商，就能立刻享受服务。云计算的最终目标是将计算、服务和应用作为一种公共设施提供给公众。

3. 云计算的特点

目前，大众普遍接受的云计算具有以下特点。

（1）规模化。云计算“资源库”拥有的规模相当大，一般由较多台机器组成“云”的集群，企业的云系统一般拥有几十万台到一百多万台服务器，企业的私有云一般也拥有成百上千台服务器不等。

（2）虚拟化。在互联网的基础上建立了云计算，而互联网本身就是一个虚拟的世界，因此，云计算技术也是虚拟的。事实上，可以把云计算类比成一个存在于网络虚拟世界里的“资源库”，所有用户请求的来源都出自该“资源库”，并非一个个固定的实体。



(3) 可靠性高。“将资料存储在硬盘里或计算机中，硬盘或计算机一旦出现故障，或者云系统一旦崩溃，自己的资料会不会无法找回？”这是很多用户的担忧。实际上，“云”使用了数据多副本容错、计算节点同构可互换等措施来保障服务的高可靠性，使用云计算比本地计算机的可靠性要高。因为数据被复制到了多个服务器节点上拥有多个副本（备份），即使遇到意外删除或硬件崩溃存储在云里的数据也不会受到影响。

(4) 通用性。为了给用户提供更便利，在“云”的支持下可以构造出千变万化的应用，同一个“云”可以支持不同的应用同时运行，用户对是否通用并不担心。

(5) 可扩展性高。为了满足应用和用户规模增长的需要，云计算的规模可以动态伸缩，用户可以根据需要进行扩展。

(6) 按需服务。云计算有一个庞大的资源库，用户按需购买，可以充分利用资源，不造成资源浪费。

(7) 成本低。云计算技术拥有强大的容错能力，其节点的构成成本非常小。用户和企业都能认可它所创造的价值。例如，利用云计算只要花费几百美元和几天时间就能完成以前需要数万美元和历经数月才能完成的任务。

云计算技术作为一项涵盖范围广且对产业影响深远的技术，未来将逐步渗透到信息产业和其他产业的方方面面，并将深刻改变产业的结构模式、技术模式和产品销售模式，进而深刻影响人们的生活。云计算的重要性也会日益凸显。同时移动互联网的出现促使云计算应用走向人们的指间，推动了云计算技术的应用发展，今后云计算将是一项随时、随地、随身为我们提供服务的技术。

1.1.3 云计算的发展历史

非常重视客户体验的在线零售商 Amazon 推动了云计算快速发展。在 Amazon 发展到一定规模后，发现自己的数据中心在大部分时间都只有不到 10% 的利用率，这表明有 90% 的资源都被闲置了。为了让自己的数据中心得到更加充分的利用，Amazon 开始考虑将计算资源从单一、特定的业务中解放出来，在空闲时提供给其他有需要的用户使用，于是就有了 AWS（Amazon Web Service，亚马逊网络服务）。初期的 AWS，只是一个线上资源库，因此，人们对其关注并不是很多。2006 年，Amazon 发布了 EC2（Elastic Compute Cloud），它第一次面向公众提供基础构架，云计算进入了更广阔的服务对象领域。

1956 年 6 月，Christopher Strachey 发表虚拟化论文，虚拟化是今天云计算基础架构的基石。

1961 年，John McCarthy 提出计算力和通过公用事业销售计算机应用的思想。

1962 年，J. C. R. Licklider 提出“星际计算机网络”设想。

1965 年，美国电话公司 Western Union 一位高管提出建立信息公用事业的设想。

1984 年，太阳电脑（Sun Microsystems）公司的联合创始人 John Gage 说出了“网



络就是计算机”，用于描述分布式计算技术带来的新世界，今天的云计算正在将这一理念变成现实。

1996年，网格计算 Globus 开源网格平台起步。

1997年，南加州大学教授 Ramnath K. Chellappa 提出云计算的第一个学术定义，认为计算的边界可以不是技术局限，而是经济合理性。

1998年，VMware（威睿公司）成立并首次引入 X86 的虚拟技术。

1999年，Marc Andreessen 创建 LoudCloud，是第一个商业化的 IaaS 平台。

2000年，SaaS 兴起。

2004年，Web 2.0 会议举行，Web 2.0 成为技术流行词，互联网发展进入新阶段。

Google 发布 MapReduce 论文。Hadoop 就是 Google 集群系统的一个开源项目总称，主要由 HDFS、MapReduce 和 Hbase 组成，其中 HDFS 是 Google File System (GFS) 的开源实现；MapReduce 是 Google MapReduce 的开源实现；Hbase 是 Google BigTable 的开源实现。

同年，Doug Cutting 和 Mike Cafarella 实现了 Hadoop 分布式文件系统（HDFS）和 Map-Reduce，Hadoop 并成为了非常优秀的分布式系统基础架构。

2005年，Amazon 宣布 Amazon Web Services 云计算平台。

2006年，Amazon 相继推出在线存储服务 S3 和弹性计算云 EC2 等云服务。

2006年，太阳电脑公司推出基于云计算理论的“BlaceBox”计划。

2007年，Google 与 IBM 在美国大学校园推广云计算的计划。

2007年3月，戴尔成立数据中心解决方案部门，先后为全球 5 大云计算平台中的三个（包括 Windows Azure、Facebook 和 Ask.com）提供云基础架构。

2007年7月，亚马逊公司推出了简单队列服务（Simple Queue Service, SQS），这项服务使托管主机可以存储计算机之间发送的消息。

2007年11月，IBM 首次发布云计算商业解决方案，推出“蓝云”（Blue Cloud）计划。

2008年1月，Salesforce.com 推出了按需应变平台 DevForce，Force.com 平台是世界上第一个平台即服务的应用。

2008年2月，EMC 中国研发集团云架构和服务部正式成立，该部门结合云基础架构部、Mozy 和 Pi 两家公司共同形成 EMC 云战略体系。

2008年2月，IBM 宣布在中国无锡太湖新城科教产业园为中国的软件公司建立第一个云计算中心。

2008年4月，Google AppEngine 发布。

2008年，Gartner 发布报告，认为云计算代表了计算的方向。

2008年5月，Sun 在 2008 JavaOne 开发者大会上宣布推出“Hydrazine”计划。

2008年6月，EMC 公司中国研发中心启动“道里”可信基础架构联合研究项目。

2008年6月，IBM 宣布成立 IBM 大中华区云计算中心。

2008年7月，HP、Intel 和 Yahoo 联合创建云计算试验台 OpenCirrus。

2008年8月3日，美国专利商标局（以下简称“SPTO”）网站信息显示，戴尔正



在申请“云计算”(Cloud Computing)商标,此举旨在加强对这一未来可能重塑技术架构的术语的控制权。戴尔在申请文件中称,云计算是“在数据中心和巨型规模的计算环境中,为他人提供计算机硬件定制制造。”

2008年9月,Google公司推出Google Chrome浏览器,将浏览器彻底融入云计算时代。

2008年9月,甲骨文和亚马逊AWS合作,用户可在云中部署甲骨文软件、在云中备份甲骨文数据库。

2008年9月,思杰公布云计算战略,并发布新的思杰云中心(Citrix Cloud Center, C3)产品系列。

2008年10月,微软发布其公共云计算平台——Windows Azure Platform,由此拉开了微软的云计算大幕。

2008年12月,Gartner披露十大数据中心突破性技术,虚拟化和云计算上榜。

2008年,亚马逊、Google和Flexiscale的云服务相继发生宕机故障,引发业界对云计算安全的讨论。

2009年,思科先后发布统一计算系统(UCS)、云计算服务平台,并与EMC、Vmware建立虚拟计算环境联盟。

2009年1月,阿里软件在江苏南京建立首个“电子商务云计算中心”。

2009年4月,VMware推出业界首款云操作系统VMwarev Sphere4。

2009年7月,中国首个企业云计算平台诞生(中化企业云计算平台)。

2009年9月,VMware启动vCloud计划构建全新云服务。

2009年11月,中国移动云计算平台“大云”计划启动。

2010年1月,HP和微软联合提供完整的云计算解决方案。

2010年1月,IBM与松下达成迄今为止全球最大的云计算交易。

2010年1月,Microsoft正式发布Microsoft Azure云平台服务。

2010年,微软宣布其90%员工将从事云计算及相关工作。

2010年4月,戴尔推出源于DCS部门设计的PowerEdgeC系列云计算服务器及相关服务。

2011年2月,思科系统正式加入OpenStack,重点研制OpenStack的网络服务。

1.2 云计算的技术分类及优点

1.2.1 云计算的技术分类

1. 按服务对象分类

按服务对象可分为公有云、私有云和混合云。这种分类主要出现在商业领域中。



公有云是服务对象面向公众的云计算服务。企业/机构利用外部云为企业/机构的用户服务,即企业/机构将云服务外包给公共云的提供商,由此来减少构建云计算设施的成本。例如,Amazon、Google、Apps、Windows Azure。

私有云通常由企业/机构自己拥有,私有云特定的云服务功能不会直接对外开放。例如 Ebay。

混合云包含公有云和私有云的混合应用。可以在通过外包减少成本的同时通过私有云保证对敏感数据等部分的控制。混合云在实践中应用较少。

2. 按技术路线分类

按技术路线分类,可分为资源整合型云计算和资源切分型云计算。

资源整合型云计算的云计算系统和技术实现方面大多体现为集群架构,通过整合大量节点的计算资源和存储资源后输出。这类系统通常能构建跨节点弹性化的资源池,分布式计算和存储技术为其核心技术。

资源切分型云计算是目前应用较为广泛的技术。虚拟化系统是最为典型的类型,这类云计算系统运用系统虚拟化对单个服务器资源实现弹性化切分,从而有效地利用服务器资源,虚拟化技术为其核心资源,此技术的优点在于用户的系统可以不进行任何改变接入采用虚拟化技术的云系统,尤其在桌面云计算技术上应用得较为成功,其缺点是跨节点的资源整合成本较高。

3. 按服务模式分类

按服务模式分类可分为基础设施即服务 (Infrastructure as a Service, IaaS)、平台即服务 (Platform as a Service, PaaS) 和软件即服务 (Software as a Service, SaaS)。

(1) 基础设施即服务

基础设施即服务指用户通过互联网可以从计算机基础设施中获得相应的服务,服务商把多台服务器组成庞大的基础设施来为客户提供服务,这需要网格计算、集群和虚拟化等技术实现。

(2) 平台即服务

平台即服务指提供一种软件研发平台的服务,将可以访问的完整或部分应用程序的开发平台提供给用户。

(3) 软件即服务

通过互联网把软件作为一种服务提供给用户,用户不需要单独购买想要的软件,而是向服务商租用基于 Web 的软件,进行软件的使用。软件作为一种服务来提供完整可直接使用的应用程序,在平台层以 SOA 方法为主,使用不同的体系应用构架,具体需要用不同的技术支持来得以实现,表示在软件应用层使用 SaaS 模式,如图 1-2 所示。



图 1-2 云计算技术分类

1.2.2 云计算技术优点

1. 信息的扩展性

使用云计算技术的用户群体与日俱增，云计算的服务量也在随之逐渐增大，因此在云计算的过程中，用户就可以更方便地在数据库中获得所需要的资源。如果在某一环节出现了安全问题，云计算系统轻松地就能将这节点进行隔离，然后迅速排除安全隐患，等到问题得以解决后，才会将这一节点的信息重新投入使用。

2. 资源的共享性

达到资源共享是云计算运行的目的，同时也是对用户的主要贡献之一。其可以不受地域的限制，即便用户处于世界的另一端，只要被网络覆盖，用户对云数据的需求都能够得到满足。拥有庞大的计算机服务器系统的云计算系统的服务商，它们能够通过网络，建立起一个足够大的平台，然后在这个平台中，用户的计算机或者手机能够获取所需的服务，这样极大地增加了知识和信息的共享性，同时服务商的运营成本也得以降低，真正优化配置了资源。

3. 管理的灵活性

由于不同用户的需求千差万别，云计算技术可以满足用户各种各样不同需求。云计算系统在为用户提供服务之前，对用户的需求应当有所了解，在了解完相应情况以后，根据用户的具体要求来制订服务计划，为用户提供所需的数据、配置相应的能力和应用。在增加或撤销的一些应用上，也可以让用户参与到云计算模式的管理和维护中来，从而了解用户的需求，进而满足用户需求。



4. 较高的性价比

云计算虽然是一项处于高科技领域的新技术,但是其成本投入并不高,而对用户来说,云计算系统也是十分经济的,因为既不需要为云计算系统配置高端的设备,也不需要对其进行实时的更新,客户只需要对云计算系统的服务商提出自己定制的需求,系统机型的维护由云计算服务商的专业人员进行,这样可以减少用户对云计算的投入,用最少的投入就可以体验最优质的服务。

5. 可靠的服务系统

云计算系统的服务商拥有非常庞大的计算机服务器系统,同时,数据和信息的存储、传递都是在一个虚拟网络平台上进行的。计算机服务器可以做到分工明确,各司其职。目前,云计算系统趋于可靠和稳定,只要某一个服务器出现安全问题,备用服务器会迅速投入到工作中来,来完成问题服务器所进行的任务。云计算技术相比于传统的互联网应用模式,它不仅能够从各个方面确保服务的灵活性、高效性和精确性,还能够为用户带来更完美的网络体验以及为企业创造更多的效益。

1.3 国内外云计算技术的发展状况及供应商

1.3.1 国内外云计算技术的发展状况

云计算可以说是信息时代中生产工具的一次巨大的飞跃,20世纪80年代是个人电脑和局域网的时代,20世纪90年代至今是互联网的时代,而未来则很有可能是云计算的时代。云计算时代的来临,预示着以设备为中心计算模式的终结,以应用互联网为中心的新的计算模式将会取而代之。现有的工作方式和商业信息服务模式等方面也会因此而改变。这对企业来说既是巨大的商机又是挑战,各大企业纷纷高调宣布“入云”“建云”等行动。2006年,亚马逊提出简单存储服务 and 弹性计算云,这标志着云计算服务开始走向成熟。2008年,IBM推出了“蓝云计划”,并提出了“共有云”“私有云”等概念,2009年又发布了基于云端的协作平台,2011年,基于Power7指挥系统,Waston Box、Cloud Box、Mason Box等三个宝盒构建云基础架构。

目前,我国政府将云计算服务正式纳入了采购项目之中,这足以表明我国政府对云计算技术的重视。在我国,政府对IT行业投入巨大,而在这部分投入的资金中,云计算系统领域占据了相当大的比例,我国政府重视云计算技术的原因在于希望我国信息产业能够迅速地发展,从而可以更好地为公共事业服务。从全球范围来看,各国政府对云计算技术发展都持支持的态度,比如欧盟、美国、日本等国家都对云计算技术投入了大量的资金,并制定了相关政策,用来增强公共云服务。相比于发达国家政府



针对云计算所采取的政策来说,我国在云计算组织架构、行业标准以及制度流程等方面还有待完善,建设力度仍需加大。经过了这些年的发展,我国的云计算市场开始逐步成型,加入云计算技术市场的软件公司、网络企业以及市场调研公司越来越多,并且已经取得了一定的成效。云计算市场处于形成阶段,我国政府对这一行业的干预并不多,而是希望通过适当的市场调控,使其健康发展。而随着云计算市场规模逐步扩大,云计算技术对于社会、经济的影响也日益凸显,越来越多的企业想要进入云计算这一领域,面对这样的市场状况,我国政府对云计算领域的监管力度必须加大,具体包括:对互联网行业进行分门别类;制定相应的管理制度;健全云计算组织结构;制定云计算行业准则;完善其相关流程,等等。

另外,对以发展云计算为理由进行固定资产投资以及圈地等不良做法引起重视。随着云计算技术快速地发展,对计算机硬件设施的要求将会越来越高,同时,为了发展云计算系统,企业必将会增加对相关设备的定制数量。相关数据显示,定制云计算服务器已然占到了全球服务器总量的10%以上。IT制造企业的利润并不会随设施定制数量的增加而增大,相反,可能会吸引其他行业的投资者进入到这一行业。IT制造业面对这样的情况,必须要加强其产品设计能力、压缩生产成本、打造自身品牌。只有这样,才能够对行业发展趋势有所把握,实现长足发展。其次,云计算的发展将会对IT行业商业模式产生重大的影响。传统IT行业的商业模式主要是以实体软件、硬件产品为主,而随着云计算技术的快速发展,这种商业模式开始转变为向用户提供软件、硬件服务,就是利用云计算,通过互联网将用户所需的应用系统传递给用户。Google公司宣称,在未来能够利用云计算技术将所有软件都转移到网络中,然后利用服务来代替软件,未来的云计算技术能够以其强大的信息处理功能,基本上能够帮助用户解决所有问题。这样一来,云计算将会对IT行业的商业模式产生重大的影响,IT行业的经营理念与发展模式也将进行转变。

如今,我国也紧跟国际步伐,现已启动“商用云计算中心”“中国云谷”“祥云工程”等项目。云计算是时代发展的趋势,真正的云计算是可以去除机器的硬盘存储,完全利用网络硬盘来储存资料。这对于网络带宽的要求很高,国内家庭网络宽带实际上平均只有1M到2M,虽然有的宽带宣称能够达到10M,但基于共享网络技术的宽带传输并不是很稳定,所以目前来说很难实现真正的云计算。但是像日本、韩国、新加坡等国家的网络带宽已经达到独享10M以上的带宽网络。实际上,云计算在我们的生活中已经很常见了,现在所用到的电子商务平台,比如京东、淘宝、唯品会等网站。哪个城市、哪台服务器存储了网店中商品的信息,卖主其实不必关心,只需在家卖货就行了。再比如,视频云计算是专门为广播电视行业处理大批量的音频、视频而精心打造的,基于平台即服务和基础设施即服务的架构方式的云平台。在未来,只需要一台手机或者一个移动终端设备,就可以通过无线网络连接到互联网,通过网络获得所需的一切服务,甚至能够获得只有超级计算机才能完成的服务。对于云计算的这些服务,最终的使用者和受益者是我们的用户。现今用户的使用习惯是目前云计算推广所面临的最大问题。用户已经对计算机的传统使用模式形成了自己的使用习惯,即软件、



硬件、网络供应商可以被看成是相互独立的部门，这些相互独立的部门按照国际规定的统一标准能够统一协作。而云计算的出现和发展，将给这些相互独立的部门协作带来更大的挑战，为了让云计算能真正地普及到用户的工作、生活和学习中，首先需要制定大量的国际标准，用以协调软件、硬件和网络之间的互联关系，这些标准相较于现在的协作标准，其制定难度可能要远远加大。

1.3.2 主流云计算供应商

1. Amazon

现如今，Amazon 免费提供 12 个月全球云服务。存储服务器费用、带宽费用、CPU 资源费用以及月租费为 Amazon 云服务的收费项目。月租费类似于电话月租费、存储服务器、带宽根据容量收费，CPU 则根据时长（小时）运算量收费。

Amazon 作为互联网上最大的线上零售商，不仅是第一个互联网云计算提供商也是目前最大的公有云服务提供商，主要为独立开发人员/开发商提供云计算服务平台。

Amazon 公司提供弹性计算云（Elastic Compute Cloud, EC2）服务、简单存储服务、弹性块存储服务、关系型数据库服务和 NoSQL 数据库服务，同时还提供与网络、数据分析、机器学习、物联网、移动服务开发、云管理、云安全等有关的云服务。

Amazon 把自己的云计算平台叫作弹性计算云，用户可以在客户端与 Amazon 的 EC2 内部实例进行交互，用户可以基于 Linux 应用程序的虚拟的集群环境运行。用户能够根据自己的使用状况，选择所使用计算平台实例的付费方式，节省了自行搭建云计算平台所需的设备和维护费用。Amazon 的弹性计算云不仅满足了软件开发人员对集群系统的需求，同时也减少了设备的维护费用。

Amazon 公司目前对云计算的研究仍在不断深化，弹性计算云平台的功能不断扩大，争取为用户提供更多的便利。

2. IBM

IBM Cloud 云服务器，将工作负载移动到高性能的全球云基础架构。IBM 的“蓝云”计算平台，为企业提供可通过 Internet 访问的分布式云计算体系。

“蓝云”计算平台结合了 IBM 的先进技术和原有的软、硬件系统，支持开放标准与开放源代码软件，它的组成部分包括数据中心、应用服务器、部署管理软件、数据库、监控软件和一些开源信息处理和虚拟化软件。集群文件系统和基于块设备方式的存储区域网络组成了“蓝云”的存储体系结构，这两个部分相互协作，为用户提供高质量的可扩展云计算服务。

3. Google

Google 是云计算研究的先行者，它推出的 GAE（Google App Engine）平台允许用户在上面编写程序，并可以在其基础架构上运行，应用运行的一切平台资源都由 GAE 提供，用户无需担心运行时所需的资源问题。GAE 平台是一种典型的云计算