



云计算与物联网关键技术 研究及应用

尚凤军 © 著



中国工信出版集团



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
<http://www.phei.com.cn>

重庆邮电大学出版基金资助

云计算与物联网关键技术 研究及应用

尚风军 著



电子工业出版社
Publishing House of Electronics Industry
北京·BEIJING

内 容 简 介

云计算与物联网是国内外备受关注、多学科高度交叉的新兴前沿研究领域，已经成为当今社会的两大支柱产业。在大数据时代，计算和存储能力成为衡量一个国家综合国力的重要标准；云计算是大数据处理的平台，对大数据的研究和利用提供支撑。计算存储能力的提升加速了社会信息化技术发展，推动了通信系统与互联网技术的高度融合，产生了物联网（Internet of Things, IoT）。本书重点介绍云计算和物联网关键技术及其应用，包括 6 部分。第一部分主要介绍云计算产生、发展和应用，以及经典的云计算平台，如 Google、Amazon、Microsoft 和 OpenStack 等；第二部介绍云平台优化技术、云存储关键技术、数据中心网络优化技术等；第三部分主要介绍新闻云、导航云等应用案例；第四部分介绍物联网体系结构及其关键技术；第五部分介绍物联网之触觉——传感网通信协议、节点定位关键技术及数据传输机制等；第六部分介绍物联网应用技术，包括儿童位置服务、智能实验室、基于 SDN 的视频应用和物联网数据融合技术等。

本书适合从事云计算和物联网研究与应用的科学工作者和工程技术人员参考使用。

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有，侵权必究。

图书在版编目（CIP）数据

云计算与物联网关键技术研究及应用 / 尚凤军著. —北京：电子工业出版社，2017.11

ISBN 978-7-121-33005-6

I. ①云… II. ①尚… III. ①云计算—应用—研究 ②智能技术—应用 IV. ①TP393.027 ②TP18

中国版本图书馆 CIP 数据核字（2017）第 271990 号

策划编辑：窦 昊

责任编辑：窦 昊

印 刷：北京京华虎彩印刷有限公司

装 订：北京京华虎彩印刷有限公司

出版发行：电子工业出版社

北京市海淀区万寿路 173 信箱 邮编：100036

开 本：787×980 1/16 印张：17 字数：435 千字

版 次：2017 年 11 月第 1 版

印 次：2017 年 11 月第 1 次印刷

定 价：69.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：(010)88254888，88258888。

质量投诉请发邮件至 zltz@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式：(010)88254466，douhao@phei.com.cn。

前 言

云计算与物联网是当今国内外备受关注的、由多学科高度交叉形成的新兴前沿研究领域。信息技术产业迅猛发展，极大丰富了人们生活的方方面面。与此同时，企业、组织和个人所产生的数据量也与日俱增，我们生活在一个数据增长比以往任何时候都要快的时代。IDC 预测，2020 年将有 40 亿人接入网络，全球数据量将达到 44ZB，是 2013 年的 10 倍之多。特别是在互联网领域，随着 Web 2.0 技术的发展，以超大规模、高并发为特点的社交网站逐渐兴起，随之而来的海量用户数据需要处理。Web 2.0 的“瘦客户端+服务”运行模式将这些庞大的数据处理任务交给服务器端，对服务器端的数据计算和存储能力要求越来越高。2012 年，Google 在世界范围内拥有数百万台服务器总量的数据中心，平均每天要处理 33 亿次的搜索请求，每月要处理的用户生成数据更是超过 400PB；同年，Facebook 公布其数据中心平均每天接收 3 亿幅用户上传的图片，数据库中新增数据也超过 500TB。这些海量数据所蕴含的 4V 特征，即体量大（Volume）、模式多（Variety）、速度快（Velocity）、价值密度低（Value），正是大数据管理和信息提取的困难度和复杂性所在。随着社会全面信息化，我们进入了“大数据时代”。传统数据存储能力和数据处理技术逐渐显得乏力，云计算技术应运而生。

面对这些海量的数据，它们的存储、挖掘、实时处理等，都需要云计算作为技术支撑，所以云计算是让大数据发挥价值的关键。云计算采取集群计算，将若干独立计算实体通过网络整合成一个具有强大计算能力的资源池，并借助基础设施即服务（IaaS）、平台即服务（PaaS）、软件即服务（SaaS）和管理服务提供商（MSP）等先进的商业模式，把计算资源池中强大的计算能力按需分配到用户手中。工业界的云计算实例，包括 Google 公司的云计算平台以及云计算的网络应用程序（如 Google App Engine），IBM 公司的“蓝云”平台，Amazon 公司的可伸缩弹性计算云等。其中，Google 云计算平台成为了互联网云计算的领头羊。

在硬件系统方面，Google 已经拥有超过 200 万台服务器。而随着大数据时代的深化以及云计算技术的发展，许多云服务提供商的云计算数据中心集群规模也持续扩展，服务器节点数目不断增多。这就带来了大规模集群和底层网络支撑之间矛盾的日

益凸显，这一问题逐渐得到了学者和工业界的广泛关注。传统网络架构已不能满足越来越灵活的网络需求，软件定义网络（Software-Defined Networking, SDN）在这种背景下被人们提出。它可为网络提供实时、快速和灵活的配置能力，使得云计算中心集群可以根据需求来调用网络资源。近年来，Google 公司已经在云计算数据中使用了 SDN 技术。

在软件系统方面，Google 以学术论文的形式陆续公开了其 GFS、MapReduce、BigTable 等分布式海量数据处理框架，同时证明了该框架的高可扩展、高性能等优点。基于这些技术，Doug Cutting 将其运用到了全网搜索引擎项目（Nutch）中。2006 年年初，开发人员将这个开源实现移出了 Nutch，成为 Lucene 的一个子项目，称为 Hadoop。同年 2 月，Apache Hadoop 项目正式启动，以支持 MapReduce 和 HDFS 的独立发展。其中，HDFS（Hadoop Distributed File System）是 Hadoop 的分布式文件系统，也是 GFS 的开源实现，它主要提供在集群中的流式高吞吐量数据访问。MapReduce 则是 Hadoop 的分布式数据处理框架，也是 Google MapReduce 计算模型的开源实现，它支持大规模数据集的并行运算。

在大数据时代，计算能力已经成为衡量一个国家综合国力的重要标准，开源的 Apache Hadoop 提供了稳定、廉价、高效的大数据处理工具，是大数据计算的事实标准，成为云计算研究的重要样本之一，被广泛应用在海量数据的存储、计算、挖掘、分析、查询以及机器学习等领域。随着云计算技术的发展，不断有类似的大数据处理工具诞生，例如 Spark、Storm 等，这说明在市场需求方面有广阔的空间。

除此之外，从 Apache 基金会推出 Hadoop 平台到目前为止不过短短数年时间，虽然得到了学术界和工业界的重视，但在很多方面仍然需要进一步完善和提高。其中，调度策略作为一项关键技术还有很长的路要走。Hadoop 中的调度策略组件负责系统中所有作业与其子任务的整个调度过程，包括如何选择作业和它的子任务，以及如何选择适合的计算节点来执行它们。调度的结果会直接影响 Hadoop 系统的整体性能和集群资源利用率。所以衡量调度策略优劣的主要指标就是系统对作业任务的响应时间（即周转时间），和集群资源（例如计算资源和带宽资源）的利用率。而目前 Hadoop 系统中仍在广泛使用的调度策略均存在许多不足之处。

社会信息化技术快速发展，通信系统与互联网技术的融合，使得科学研究工作不断更新，其中，物联网相关的各种技术研究状况较为热烈。物联网（Internet of Things, IoT）将物与物之间、人与物之间通过互联网联系起来，采取智能感知识别技术、计算机通信技术等，以互联网为核心基础，在其层面上深层次研究拓展而形成网络，用户通过物与物的通信过程实现信息交换，如远程管理、智能监控等信息化操作。从 1999 年中科院对 IoT 的研究工作开始，我国各项工程都针对 IoT 的相关技术进行了深层次的探索，如 2006 年

在重庆成立的移动物联网运营中心，2009年建立的国家传感信息中心，2010年在上海开幕的IoT大会，将物联网视为国家重点发展新型产业之一。由于传感技术的逐渐成熟与政府的大力支持，智能交通、安防、家居等市场已经融入IoT的相关技术，多个城市已经投入规模化的应用，目前我国IoT的产业链已经基本形成。与此同时，云计算、大数据、移动互联网的时代已经来临，目前互联网的数据传递方式从最传统的形式变为海量数据互联，随着国内云计算服务的建设，IoT的广泛应用推动着云计算的发展。

本书作者经过近5年对云计算与物联网的学习和研究，积累了一定的成果，为本书的出版奠定了扎实的基础。我们首先从云计算的工作原理出发，通过吸收国内外大学和研究所以来的研究成果，对云计算技术进行了研究并开发了基于云平台的应用案例，提出自己的看法和思路；接着从物联网方面进行研究并开发了基于物联网的应用案例，取得一定的研究成果。随着不断地研究开发积累，通过对云计算与物联网的理论研究和实践经验总结，构成本书的全部内容。根据当前国家有关部门对云计算与物联网技术和应用的密切关注程度，以及国家未来15年中长期科技发展战略部署和国家重点基础研究发展计划纲要的主题，本书划分为18章、6大部分。第一部分主要介绍云计算产生、发展和应用，以及经典的云计算平台，包括Google、Amazon、Microsoft和OpenStack等。第二部介绍云平台优化技术、云存储关键技术、数据中心网络优化技术等内容。第三部分主要介绍新闻云、导航云等应用案例。第四部分介绍物联网体系结构及其关键技术。第五部分介绍物联网之触觉——传感网通信协议、节点定位关键技术及数据传输机制等。第六部分介绍物联网应用技术，包括儿童位置服务、智能实验室、基于SDN的视频应用和物联网数据融合技术等。

本书由尚凤军著，程克非博士、黄颖博士、雷建军博士、张旭博士等多次参与有关本书的技术讨论，提供了云计算与物联网的相关资料。另外，感谢重庆邮电大学计算机科学与技术学院对本书出版工作的支持。同时，参与资料整理的人员有付强、闫辰云、龚文娟、邓小林、王楹颖、吴江、陈炫伶、李燕、胡尚平、李雅琳、梁聪、刘海昇、何德祥、王文凯、晨星、李路中。另外，在本书编写过程中，引用和参考了部分其他作者和网上的相关文献材料，在此一并表示感谢。同时，被引用的内容版权归站点或其他作者所有。

本书的出版得到了国家自然科学基金项目（项目编号：61672004）、重庆邮电大学出版基金、重庆邮电大学教材建设项目（项目编号：JC2012-05）、重庆市基础与前沿研究计划项目（项目编号：cstc2016jcyjA0590）、重庆市研究生教育教学改革研究项目（项目编号：yjg143056）、重庆市高等教育教学改革研究项目（项目编号：173068）、重庆邮电大学教育教学改革研究项目（XJG1513）等的资助，特此表示感谢！

由于水平有限，加之编写时间仓促，书中疏漏之处在所难免，欢迎广大读者和同行批评指正。无线传感器网络正处在飞速发展的阶段，我们愿在吸取大家意见和建议的基础

上,不断修改和完善书中有关内容,为推动无线传感器网络应用领域的发展与进步尽微薄之力。

为充分展现本书编写特色,帮助读者深刻理解本书的编写意图与内涵,进一步提高对本书教学的使用效率,我们给出本书使用指导联络方式,欢迎读者将图书使用过程中的问题与各种探讨、建议反馈给我们,本书编者将竭诚为您服务。我们的联系方式是 E-mail: shangfj@cqupt.edu.cn。

编 者

2017年9月

目 录

云 基 础 篇

第 1 章 云计算概述	2
1.1 云计算产生	2
1.2 云计算体系结构	3
1.3 云计算发展	5
1.4 云计算应用	5
参考文献	7
第 2 章 云计算基础	8
2.1 Google 云计算	8
2.2 Amazon 云计算	11
2.3 Microsoft 云计算	13
2.4 VMware 云计算	18
2.5 OpenStack 云计算	21
2.6 开源云计算系统	23
2.7 云计算关键技术	25
2.8 国内云平台	26
参考文献	30

云 研 究 篇

第 3 章 Hadoop 云计算中任务调度策略研究	34
3.1 Hadoop 系统基本框架	34
3.1.1 Hadoop 作业调度模型	35
3.2 MapReduce 简介	35
3.2.1 MapReduce 作业执行过程	36

3.2.2	MapReduce 的任务调度策略	37
3.2.3	云计算应用案例	38
3.3	Hadoop 作业调度优化	42
3.3.1	MNBC 调度策略	42
3.3.2	代价计算模型及调度算法	43
3.3.3	带宽节省效率分析	47
	参考文献	50
第 4 章	OpenStack 云计算中资源调度机制研究	51
4.1	OpenStack 虚拟机初始资源调度算法	51
4.1.1	虚拟机初始资源调度算法概述	51
4.1.2	国内外研究现状	53
4.2	OpenStack 的初始放置机制设计	54
4.2.1	虚拟机放置模型	55
4.2.2	虚拟机初始放置机制	56
4.3	OpenStack 虚拟机初始放置算法仿真实现	60
4.3.1	MHD-VMP 算法仿真流程	60
4.3.2	MHD-VMP 算法代码仿真实现	61
4.4	仿真实验	62
4.4.1	实验条件	63
4.4.2	实验结果与分析	63
	参考文献	65
第 5 章	数据中心网络传输机制研究	68
5.1	数据中心网络概述	68
5.1.1	多约束 QoS 路由算法	68
5.1.2	数据中心流量负载均衡	69
5.1.3	国内外研究现状	70
5.2	大象流碰撞回避流量调度算法	70
5.2.1	数据中心网络中大象流碰撞问题	71
5.2.2	基于 FatTree 的路由模型	71
5.2.3	单流调度优化算法	73
5.2.4	基于非等概率的大象流碰撞回避流量调度算法	74
5.3	仿真实验及分析	75

5.3.1 对比模型	75
5.3.2 实验方法	76
5.3.3 实验结果及分析	78
参考文献	81
第 6 章 云计算平台中混合存储机制研究	82
6.1 混合主存系统	82
6.1.1 混合主存系统概述	82
6.1.2 国内外研究现状	83
6.2 基于热度感知的页面划分管理算法	84
6.2.1 算法结构	84
6.2.2 页面迁移策略	88
6.2.3 页面管理策略	91
6.3 实验和结果分析	93
6.3.1 模拟器研究	93
6.3.2 实验环境	94
6.3.3 实验结果对比	95
参考文献	98

云 应 用 篇

第 7 章 新闻云	102
7.1 背景	102
7.2 新闻云搭建方案	102
7.3 实现流程及系统框架	104
7.3.1 新闻云服务	104
7.3.2 数据分析处理及存储	106
7.4 测试	106
7.4.1 新闻云手机客户端测试	107
7.4.2 客户机测试	108
参考文献	113
第 8 章 基于 Hadoop 平台的导航系统设计	115
8.1 背景	115
8.2 使用环境及软件介绍	115

8.3	原理	116
8.4	实现流程及系统框架	117
8.5	测试	118
8.5.1	测试环境	119
8.5.2	功能测试方案	119
8.5.3	性能测试方案	120
8.5.4	测试数据及结果	121
	参考文献	123
第 9 章	基于 Hadoop 和 Storm 的交通预测	124
9.1	背景	124
9.2	应用技术	124
9.3	系统架构图	125
9.4	实现原理	125
9.5	测试	126
	参考文献	128

物联网基础篇

第 10 章	物联网概述	130
10.1	物联网的产生	130
10.2	物联网的体系结构	132
10.3	物联网关键技术	133
10.4	物联网应用	134
	参考文献	135
第 11 章	物联网基础	137
11.1	EPC 与物联网	137
11.2	RFID 技术	139
11.2.1	RFID 的工作原理	140
11.2.2	RFID 体系结构	140
11.2.3	RFID 系统分类	141
11.2.4	RFID 应用	142
11.3	智能技术	143
11.3.1	智能	143

11.3.2 智能技术的应用	143
11.4 纳米技术	145
11.4.1 纳米材料	145
11.4.2 纳米材料的特性	146
参考文献	146

物联网研究篇

第 12 章 传感器网络通信协议研究	150
12.1 传感器网络简介	150
12.2 无线传感器网络体系结构	150
12.3 传感器网络通信协议	152
12.3.1 CHTD 簇头的产生	152
12.3.2 CHTD 簇头数目的确定	153
12.3.3 CHTD 最优簇半径	155
12.3.4 CHTD 描述	156
12.4 CHTD 算法的改进	157
12.5 仿真分析	161
12.5.1 生命周期	162
12.5.2 接收数据包量	164
12.5.3 能量消耗	165
12.5.4 负载均衡	166
参考文献	169
第 13 章 传感器网络定位研究	171
13.1 无线传感器网络定位技术	171
13.2 存在的问题	171
13.3 基于测距的定位方法	171
13.4 非测距定位算法	173
13.5 传感器网络应用	174
13.6 定位算法实现	175
13.7 定位算法仿真及分析	177
参考文献	180
第 14 章 基于随机游走的安全传输机	181
14.1 无线传感器安全	182

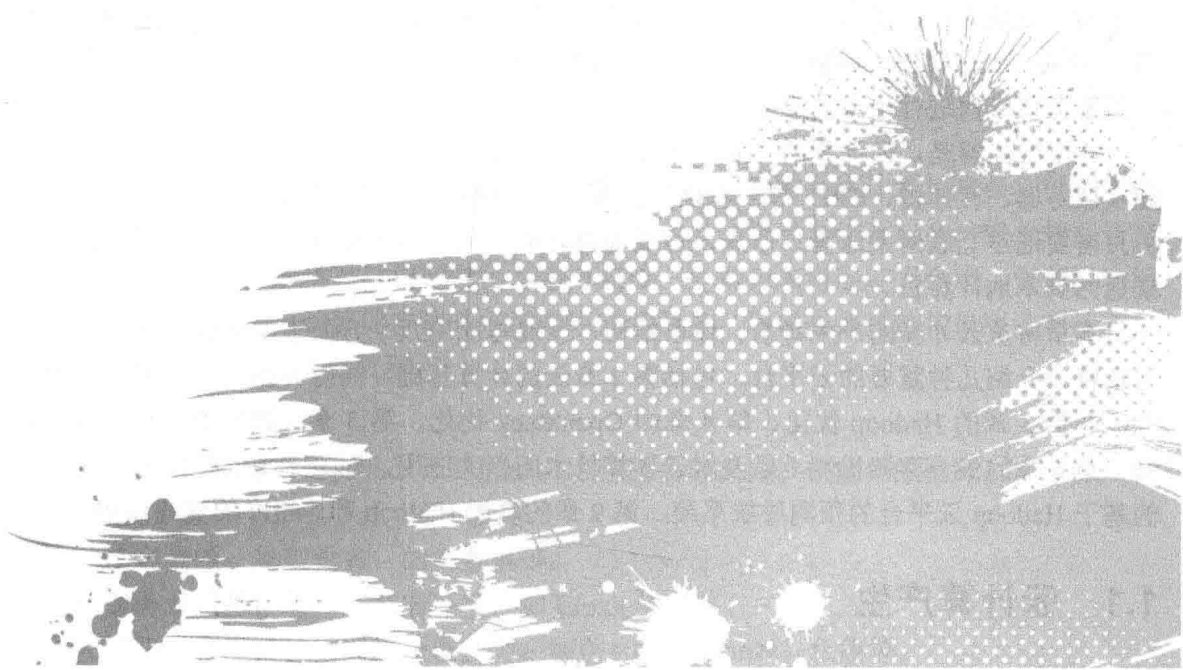
14.2	基于数据分片的安全路由机制	182
14.2.1	安全路由机制介绍	183
14.2.2	安全通信模型的实现与改进	183
14.3	仿真与分析	195
	参考文献	198

物联网应用篇

第 15 章	儿童定位技术	202
15.1	背景	202
15.2	使用环境及软件	202
15.3	具体设计方案	202
15.3.1	整体架构	203
15.3.2	定位方法	205
15.4	测试	206
15.4.1	测试设备	206
15.4.2	测试内容	207
15.4.3	测试数据	209
15.4.4	测试结果分析	216
	参考文献	218
第 16 章	智能实验室系统	220
16.1	智能实验室介绍	220
16.2	系统方案	220
16.3	实施原理及过程	222
16.4	硬件框图与测试结果	224
	参考文献	227
第 17 章	软件定义网络应用	228
17.1	SDN 技术及其在视频业务带宽保障中的应用	228
17.1.1	研究问题及本设计的特色	229
17.1.2	设计架构与程序流程	230
17.1.3	实验设计	232
17.1.4	实验结果	233
17.2	SDN 中的链路故障快速重构应用	235
17.2.1	背景及应用场景	235

17.2.2 具体设计	236
17.2.3 实验测试	242
参考文献	246
第 18 章 物联网数据融合技术	248
18.1 多聚焦图像融合概述	248
18.2 基于 DCT 域多聚集图像融合算法及评判标准	249
18.2.1 几种重要的基于 DCT 域多聚集图像融合算法	249
18.2.2 常用的图像融合评价标准	253
18.3 一种基于 DCT 块与系数相结合的融合算法	254
18.3.1 无线视觉传感器中的图像融合	254
18.3.2 算法描述	254
18.3.3 实验结果	257
参考文献	258

云基础篇



第 1 章

云计算概述

云计算 (Cloud Computing) 是由分布式计算 (Distributed Computing)、并行计算 (Parallel Computing)、网格计算 (Grid Computing) 发展而来的一种新兴商业计算模型。目前对于云计算的认识在不断发展变化中, 云计算仍没有普遍一致的定义。美国国家标准与技术研究院 (NIST) 将其定义为: 云计算是一种按使用量付费的模式, 这种模式提供可用的、便捷的、按需的网络访问, 进入可配置的计算资源共享池 (资源包括网络、服务器、存储、应用软件、服务等), 这些资源能够被快速提供, 只需投入很少的管理工作, 或服务供应商进行很少的交互。

通俗的理解是, 云计算的“云”是存在于互联网服务器集群上的资源, 它包括硬件资源 (服务器、存储器、CPU 等) 和软件资源 (如应用软件、集成开发环境等), 本地计算机只需要通过互联网发送需求信息, 远端就会有成千上万的计算机提供所需的资源并将结果返回到本地计算机。

本章将从云计算的产生背景、体系结构、发展与应用几个方面来阐述云计算, 接下来几章将会介绍几种重要的云平台, 包括第 2 章云的计算基础, 同时介绍与云计算相关的研究, 即第 3 章的 Hadoop 优化、第 4 章的 OpenStack 优化、第 5 章的数据中心网络传输机制、第 6 章的混合存储机制, 以及对云计算技术的相关应用, 即第 7 章的新闻云、第 8 章的基于 Hadoop 云平台的预测导航系统、第 9 章的基于 Hadoop 和 Storm 的交通预测。

1.1 云计算产生

云计算由一系列可以动态升级和被虚拟化的资源组成, 这些资源被所有用户共享并且可以方便地通过网络访问, 用户无需掌握云计算的技术, 只需要按照个人或者团体的需要租赁云计算的资源。云计算是继个人计算机变革、互联网变革之后的第三次 IT 浪潮, 将带来生活、生产方式和商业模式的根本性改变, 云计算成为当前社会关注的热点。

云计算还处于萌芽阶段, 有庞杂的各类厂商在开发不同的云计算服务。目前, 云计算

的主要服务形式有 SaaS (Software as a Service), PaaS (Platform as a Service), IaaS (Infrastructure as a Service)。

(1) 基础设施即服务 (IaaS)。将厂商提供的服务器组成“云端”基础设施, 并作为计量服务提供给客户。它将内存、I/O 设备、存储和计算能力整合成一个虚拟的资源池, 为整个业界提供所需要的存储资源和虚拟化服务器等服务。

(2) 平台即服务 (PaaS)。将开发环境作为一种服务来提供。这是一种分布式平台服务, 厂商提供开发环境、服务器平台、硬件资源等服务给客户, 用户在其平台基础上开发自己的应用程序并通过其服务器和互联网传递给其他客户。PaaS 能够给企业或个人提供研发的中间件平台, 提供应用程序开发、数据库、应用服务器、试验、托管及应用服务。

(3) 软件即服务 (SaaS)。SaaS 服务提供商将应用软件统一部署在自己的服务器上, 用户根据需求通过互联网向厂商订购应用软件服务, 服务提供商通过浏览器向客户提供软件应用服务, 并根据客户所定软件的数量、时间的长短等因素收费。这种服务模式的优势是, 由服务提供商维护和管理软件、提供软件运行的硬件设施, 用户只需拥有能够接入互联网的终端, 即可随时随地使用软件。

1.2 云计算体系结构

云计算平台是一个强大的“云”网络, 连接了大量并发的网络计算和服务, 可利用虚拟化技术扩展每台服务器的能力, 将各自的资源通过云计算平台结合起来, 提供超级计算和存储能力。云计算的体系结构通常由 5 部分组成, 分别为应用层、平台层、资源层、用户访问层和管理层。云计算的本质是通过网络提供服务, 所以其体系结构以服务为核心。云计算的体系结构如图 1.1 所示。

1、资源层

资源层是指基础架构里的云计算服务, 这些服务可以提供虚拟化的资源, 从而隐藏物理资源的复杂性。包括以下几个组成部分:

- 物理资源: 物理设备, 如服务器等。
- 服务器服务: 操作系统的环境, 如 Linux 集群等。
- 网络服务: 提供的网络处理能力, 如防火墙、VLAN、负载等。
- 存储服务: 为用户提供存储能力。

2. 平台层

平台层为用户提供对资源层服务的封装, 使用户可以构建自己的应用, 包括:

- 数据库服务: 提供可扩展的数据库处理能力。
- 中间件服务: 为用户提供可扩展的消息中间件或事务处理中间件等服务。