



21 世纪高等院校
云计算和大数据人才培养规划教材

INTRODUCTION TO CLOUD COMPUTING

Concepts, Frameworks and
Applications

云计算导论

概念 架构与应用

· 武志学 编著 ·



中国工信出版集团



人民邮电出版社
POSTS & TELECOM PRESS



21 世纪高等院校
云计算和大数据人才培养规划教材

INTRODUCTION TO CLOUD COMPUTING

Concepts, Frameworks and
Applications

云计算导论

概念 架构与应用

· 武志学 编著 ·

人民邮电出版社
北京

图书在版编目 (C I P) 数据

云计算导论：概念 架构与应用 / 武志学编著. —
北京：人民邮电出版社，2016.9
21世纪高等院校云计算和大数据人才培养规划教材
ISBN 978-7-115-42405-1

I. ①云… II. ①武… III. ①计算机网络—高等学校—教材 IV. ①TP393

中国版本图书馆CIP数据核字(2016)第102663号

内 容 提 要

本书全面介绍云计算的概念、框架与应用。全书共 8 章，主要内容包括云计算的基本概念、云计算平台体验、IaaS 服务模式、PaaS 服务模式、SaaS 服务模式、桌面云、云存储、典型的云计算平台。本书内容实用，实验丰富，将实验内容融合在课程内容中，使理论紧密联系实际。

本书主要是面向大学本专科教学的云计算技术概论性入门教材，通过学习本书，可以了解今后需要学习哪些课程和技术来系统掌握云计算工作原理和开发基于云计算的应用。本书不仅适用于高校本专科教学使用，也可以作为培训教材用于相关技术培训。

◆ 编 著 武志学

责任编辑 马小霞

责任印制 焦志炜

◆ 人民邮电出版社出版发行，北京市丰台区成寿寺路 11 号

邮编 100164 电子邮件 315@ptpress.com.cn

网址 <http://www.ptpress.com.cn>

固安县铭成印刷有限公司印刷

◆ 开本：787×1092 1/16

印张：15

2016 年 9 月第 1 版

字数：385 千字

2016 年 9 月河北第 1 次印刷

定价：39.80 元

读者服务热线：(010)81055256 印装质量热线：(010)81055316

反盗版热线：(010)81055315

前 言

在过去的半个多世纪中，信息技术的不断发展极大地改变了政府和企业的运行模式，也给人们的生活方式带来了巨大的变化。继个人计算机变革、互联网变革之后，云计算被看作第三次 IT 浪潮，是中国政府“互联网+”战略的重要组成部分。云计算已经给人们的生活、生产方式和商业模式带来了根本性改变，是近年来全社会关注的热点。

很少有一种技术能够像云计算这样使全球所有的 IT 巨头都同时关注和推动其发展。在 2005 年 Amazon 公司推出以 S3 和 EC2 为核心的 AWS 云服务之后，云计算服务受到整个 IT 产业的重视，人们认识到一种新的 IT 服务模式已经开始形成。Google、IBM、Microsoft 等互联网公司和 IT 行业巨头在发现云计算服务市场的巨大潜力之后，分别从不同的方面进入云计算市场，提供不同层面的云计算服务，促进云计算服务进入了快速发展的阶段。经过十多年的发展，云计算已经逐渐从快速发展阶段向成熟阶段迈进。云计算服务已经得到了政府、企业及个人的认可和使用。产业界对待云计算不再是抱着疑虑和试探的态度，而是越来越务实地接纳它、拥抱它，不断去挖掘云计算中蕴藏的巨大价值。

云计算和众多新行业一样，面临着一个重要问题，即人才问题。云计算行业的就业机会增长迅速，在 2015 年，全球最知名职业人士社交网站 LinkedIn 公布的最受雇主喜欢同时炙手可热的技能中，“云计算”排名第一，“数据分析”位列第二。对于云计算和大数据人才的稀缺也同样体现在国内。据 IDC 报告，中国和印度在 2015 年产生了 670 万个云计算就业机会。同时，在公有云和私有云 IT 服务领域将创造 1380 万个就业机会，并预测，超过一半的人才需求来自 500 人以下的中小企业。随着云计算进入成熟阶段，云计算服务的普及率会越来越高，对满足产业发展的人才需求将呈现空前增长态势，尤其是对优质产业人才的需求将不断扩大。

造成云计算产业人才需求数量迅猛增长的主要原因有几个：一是云计算产业市场规模快速增长，使得云计算产业人才的需求数量不断增加；二是相关云计算企业加大了对核心技术的投入，提高了对客户的服务，因此无论是技术层面，还是运营商层面、集成与服务提供层面，都对高精尖人才有着巨大的需求；三是随着云计算新市场、新业务、新应用的不断出现，国内外各大知名软件企业加速占据国内云计算产业高地，在全国加速建立分公司和研发中心。

面对云计算人才紧缺的现状，编者基于自己在国际著名虚拟化和云计算公司思杰（Citrix）十几年的工作经验，与电子科技大学成都学院在 2012 年 5 月联合创办了云计算科学与技术系，重点培养具有完整知识体系的本科云计算人才。经过近四年的实践，已经为社会输出了近千名的云计算本科人才，形成了一套完整的云计算本科阶段的培养方案。

作为云计算技术的入门教材，本书将对云计算技术的起源、系统架构、核心技术、使用模式、部署模式、发展现状做深入浅出的全面介绍，使读者清晰了解云计算的整体概念和应用前景，以及在后续课程中所需要学习的技术。实践是掌握知识的最佳途径。本书不但把云计算知识的介绍与国际知名云计算平台的具体实践的描述相结合，并且为学生提供了实际使用各类云平台的实践环境，从而使学生可以通过实践加深对云计算知识的了解和认知。

本书共分为 8 章，下面对各章内容分别进行简单介绍。

第 1 章对云计算进行了概述。首先通过对云计算的思想起源和发展历程的介绍使读者了解云计算在 IT 革命中的历史地位；在给出云计算的定义之后，对云计算的三种服务类型、四类部署模式和五大核心特性进行了阐述，以加深对云计算概念的理解；随后讨论了云计算的使用场景和云计算给 IT 行业及社会带来的变革；最后描述了云计算的产业链结构以及各著名互联网和 IT 行业巨头在云计算产业链中的地位。

为了能更好地理解云计算的概念，第 2 章描述了如何实际使用三大类云计算服务，包括 OpenStack 基础设施即服务（Infrastructure as a Service, IaaS），Cloud Foundry 平台即服务（Platform as a Service, PaaS）和 Microsoft 的软件即服务（Software as a Service, SaaS）。通过实际使用这些云计算平台为深入学习各类云计算服务的原理和实现打下良好的基础。

第 3 章～第 7 章分别对三大类云计算核心服务和两类云计算衍生服务从实现功能、体系架构与核心技术方面进行了分析和介绍，并通过描述典型云计算平台对具体实现方法进行了讨论。

第 3 章对 IaaS 进行了介绍。本章描述了 IaaS 服务应提供的基本功能，包括用户管理、资源抽象、资源监控、资源调度、资源部署、数据管理、安全管理和任务管理；描述了 IaaS 平台的整体架构；深入介绍了实现 IaaS 平台的核心技术：服务器虚拟化；通过描述开源云计算平台 OpenStack 的功能和架构讨论了实现 IaaS 平台的方法。

第 4 章对 PaaS 进行了介绍。本章讨论了 PaaS 平台的两大主流类型（事务处理类和数据分析类）及其各自的驱动力；介绍了 PaaS 平台的三种功能角色，即共享的中间件平台、集成的软件和服务平台及虚拟的应用平台；介绍了 PaaS 平台核心系统的主要功能；分别以 Cloud Foundry 和 Hadoop 两个开源平台讨论了两大主流类型 PaaS 平台的功能、系统架构和实现技术。

第 5 章对 SaaS 进行了介绍。本章讨论了 SaaS 服务的发展历程和基本特征；介绍了 SaaS 平台的四种架构类型：定制开发、可配置架构、多租户架构和可伸缩性的多租户架构；描述了实现 SaaS 平台的关键技术和参考模型；介绍了对 SaaS 应用的分类以及典型的 SaaS 应用的实例。

第 6 章对桌面云（Desktop as a Service）进行了介绍。本章介绍了桌面云的业务价值和发展历史；描述了桌面云的整体架构；分别对桌面云的三大核心技术进行了深入介绍，包括虚拟桌面架构（VDI）技术、虚拟桌面交付协议和应用发布技术。本章特别对不同桌面云产品所提供的虚拟桌面交付协议进行了深入的分析和对比。

第 7 章对云存储进行了介绍。本章描述了云存储系统的结构模型和两大类云存储架构类型：紧耦合对称架构和松弛耦合非对称架构；详细讨论了云存储的三大类型及其适合的应用，包括块存储、文件存储和对象存储；对实现云存储的关键技术做了深入介绍，包括存储虚拟化技术、分布式存储技术、数据容错技术、数据备份技术和数据缩减技术；介绍了三个典型的云存储服务系统：EMC ATMOS、Amazon 云存储服务和 Google 的云存储服务。

第 8 章介绍了云计算的业界动态，对当前国际上典型的云计算平台做了详细介绍，主要涉及国际上领先的云计算厂商，包括 Amazon、Google、Salesforce 和 Microsoft 等，以及当前比较热门的开源云计算平台。本章比较详细地介绍了每个云计算厂商的云计算

产品线，并分析其产品的功能和特点，使读者能够对著名云计算厂商及其产品有一个总体的认识。本章还在对开源云计算平台进行描述的基础上按类型进行了分析和比较，以使读者了解每个开源平台的服务对象和主要特点。

云计算引起的变化，不只局限于 IT 领域，它和人们的整个生活方式都有关系。不论是对 IT 企业（硬件商、软件商还是平台商），对企业（大型企业、中型企业还是小型企业），还是对个人和政府，云计算都带来了革命性的改变。本书不仅从 IT 角度解释了什么是云计算，还从非 IT 角度来描述云计算给社会带来的变化，以及如何使用云计算为人们的生活和工作服务。本书系统地说明云计算的概念和发展历程、现有云计算企业的战略、云计算的核心技术、云计算的未来发展以及如何利用云计算的优势来改变你的生活和企业发展。

本书的编者具有多年国内外高校教学经验，拥有深厚的虚拟化、云计算和大数据技术理论基础功底，曾担任国际著名云计算公司主任研究员。丰富的虚拟化云计算技术研发经历，使编者对虚拟化云计算技术的发展历程和体系架构有着深入和系统的理解。参与的多项新技术普遍应用于各行业。例如，广泛应用于欧洲各大学的 ANSA 系统、应用于欧洲各大银行的 CAGE 互联网安保系统、Xen 虚拟化技术、XenApp 应用虚拟化系统、XenDesktop 桌面云系统、XenServer 服务器虚拟化系统、IBM 抗灾云系统、目前国际上最热的开源云平台 OpenStack 和 PaaS 云平台 Cloud Foundry。作为世界云计算技术前瞻性研究和应用推广的参与者，开源云平台初期的开发者及核心代码贡献者，国内首个云计算专业系的创办者，编者基于多年的工作、研究及教学经验编写了本书，不仅仅是将相关技术内容简单地告诉读者，而且将复杂问题简单化，以深入浅出的方式描述了云计算的方方面面，目的是希望读者通过对本书的学习，了解什么是云计算，云计算的优势有哪些，云计算的服务类型有哪些，云计算能够为个人、企业、政府、社会带来什么变化，云计算的核心技术有哪些，云计算的将来发展方向是什么，从而使读者可以有意识地利用云计算为个人或企业服务。

在编写过程中，编者得到了电子科技大学成都学院领导和同事的不少帮助，获得了学院教材建设基金给予的经费资助，在此对学院的支持表示感谢。同时，感谢协助完成本书编写工作的电子科技大学成都学院的赵阳老师，感谢成都创智云公司宋怡、田盛、郭岩婷、杨莹、吕珊珊等同事在代码验证、体验案例测试、插图绘制、内容校订等方面给予的大力协助；感谢合作企业广州五舟公司的一贯支持和帮助。

编者

2016年3月于成都

目 录 CONTENTS

第 1 章 云计算的基本概念 1

1.1 云计算概述	1	1.2.3 SaaS: 软件即服务	11
1.1.1 云计算的思想起源	1	1.3 云计算部署模式	13
1.1.2 云计算的定义	2	1.4 云计算的使用场景	16
1.1.3 云计算的发展历程	4	1.5 云计算带来的变革	17
1.2 云计算的服务类型	7	1.6 云计算产业链结构	18
1.2.1 IaaS: 基础设施即服务	8	习题	20
1.2.2 PaaS: 平台即服务	10		

第 2 章 云计算平台体验 21

2.1 IaaS 体验	21	2.2.1 体验对象	27
2.1.1 体验对象	21	2.2.2 安装客户端	27
2.1.2 安装部署	21	2.2.3 部署应用	30
2.1.3 添加镜像	22	2.2.4 使用应用	31
2.1.4 登录管理界面	23	2.3 SaaS 体验	32
2.1.5 创建云主机	24	2.3.1 注册账号	32
2.1.6 操作云主机	25	2.3.2 登录 OFFICE ONLINE	34
2.1.7 使用云主机	25	2.3.3 使用 OFFICE ONLINE	34
2.1.8 挂载磁盘	26	2.4 总结	36
2.2 PaaS 体验	27	习题	36

第 3 章 IaaS 服务模式 37

3.1 概述	37	3.4.2 实现方式	44
3.2 基本功能	38	3.4.3 关键特性	44
3.2.1 用户管理	39	3.4.4 核心技术	45
3.2.2 资源抽象	39	3.4.5 虚拟化与云计算	52
3.2.3 资源监控	39	3.5 OpenStack	53
3.2.4 资源调度	39	3.5.1 简介	53
3.2.5 资源部署	40	3.5.2 OpenStack Compute: Nova	55
3.2.6 数据管理	40	3.5.3 OpenStack Block Storage: Cinder	60
3.2.7 安全管理	41	3.5.4 OpenStack Network: Neutron	61
3.2.8 任务管理	41	3.5.5 OpenStack Image Service: Glance	63
3.3 整体架构	41	3.5.6 OpenStack Object Storage: Swift	64
3.4 服务器虚拟化技术	43	3.5.7 小结	67
3.4.1 IaaS 的基本资源	43	习题	67

第4章 PaaS 服务模式 68

4.1 概述	68	4.3.5 部署模式	81
4.1.1 驱动力	68	4.4 Hadoop	84
4.1.2 主流类型	69	4.4.1 概述	84
4.1.3 功能角色	71	4.4.2 Hadoop 简史	84
4.2 核心系统	73	4.4.3 Hadoop 组成部分	85
4.2.1 简化的应用开发和部署模型	73	4.4.4 HDFS	85
4.2.2 自动资源获取和应用激活	74	4.4.5 MapReduce	87
4.2.3 自动的应用运行管理	75	4.4.6 MapReduce 计算举例	88
4.2.4 平台级优化	76	4.4.7 HDFS 与 MapReduce 组合	90
4.3 Cloud Foundry	77	4.4.8 MapReduce 的优势与劣势	90
4.3.1 简介	77	4.4.9 小结	91
4.3.2 特点	77	4.5 总结	91
4.3.3 逻辑结构	78	习题	91
4.3.4 整体架构	80		

第5章 SaaS 服务模式 92

5.1 概述	92	5.2.3 支撑平台的参考实现	105
5.1.1 特征	92	5.3 SaaS 应用	107
5.1.2 发展历程	93	5.3.1 SaaS 应用的分类	107
5.1.3 实现层次	94	5.3.2 云应用的典型示例	108
5.2 支撑平台	95	5.4 SaaS 发展趋势	111
5.2.1 支撑平台的类型	95	5.5 总结	111
5.2.2 支撑平台的关键技术	96	习题	112

第6章 桌面云 113

6.1 概述	113	6.6.4 PCoIP 协议	132
6.2 业务价值和缺点	114	6.6.5 SPICE 协议	133
6.3 发展历史	116	6.6.6 对比分析	136
6.4 桌面云架构	119	6.6.7 小结	137
6.5 虚拟桌面架构 (VDI) 技术	123	6.7 应用发布	139
6.6 虚拟桌面交付协议	124	6.7.1 应用流	140
6.6.1 概述	125	6.7.2 应用虚拟化	140
6.6.2 RDP 协议	125	6.8 总结	141
6.6.3 ICA/HDX 协议	127	习题	142

第7章 云存储 143

7.1 概述	143	7.5.2 分布式存储技术	151
7.2 结构模型	144	7.5.3 数据容错	153
7.3 云存储架构	146	7.5.4 数据备份	154
7.4 云存储类型及其适合的应用	146	7.5.5 数据缩减技术	155
7.4.1 块存储	147	7.6 典型的云存储服务	159
7.4.2 文件存储	147	7.6.1 EMC ATMOS	159
7.4.3 对象存储	148	7.6.2 Amazon 云存储服务	160
7.4.4 小结	149	7.6.3 Google 的云存储服务	162
7.5 关键技术	150	7.7 总结	163
7.5.1 存储虚拟化	150	习题	163

第8章 典型的云计算平台 164

8.1 Amazon 云计算平台	164	8.3.6 应用打包服务	185
8.1.1 AWS 产品	164	8.4 Microsoft Azure	186
8.1.2 常用 AWS 之间的关系	171	8.4.1 Microsoft Azure 简介	186
8.1.3 Amazon EC2	171	8.4.2 Windows Azure	187
8.1.4 Amazon EBS	173	8.4.3 SQL Azure	190
8.1.5 Amazon Simple Storage Service (S3)	175	8.4.4 Windows Azure AppFabric	191
8.1.6 Amazon SimpleDB	176	8.4.5 Windows Azure Marketplace	192
8.2 Google 云计算平台	177	8.4.6 Microsoft Azure 服务	192
8.2.1 GAE 平台简介	177	8.5 开源 IaaS 平台	195
8.2.2 分布式存储服务	178	8.5.1 OpenStack	195
8.2.3 应用程序运行环境	179	8.5.2 CloudStack	202
8.2.4 应用开发套件	180	8.5.3 Eucalyptus	208
8.2.5 Google 应用	181	8.5.4 三大开源 IaaS 平台的比较	212
8.3 Salesforce 云计算平台	181	8.6 开源 PaaS 平台	217
8.3.1 Salesforce 的整体架构	182	8.6.1 Cloud Foundry	217
8.3.2 Force.com	182	8.6.2 OpenShift 3	221
8.3.3 基础服务	183	8.6.3 OpenShift 与 Cloud Foundry	226
8.3.4 数据库服务	184	8.7 其他云计算公司	227
8.3.5 应用开发服务	184	8.8 总结	227
		习题	228

云计算 (Cloud Computing) 是一个内涵丰富而定义模糊的名词。当前, 云计算已经席卷了 IT 行业的各个领域, 人们似乎很难清晰地把握住云计算的本质。本章通过讲解云计算的概念、起源及发展历程, 使大家逐步认识云计算的特征、服务模式和类型, 从而对云计算有一个初步的了解。

1.1 云计算概述

云计算的出现并非偶然, 早在 20 世纪 60 年代, 麦卡锡就提出了把计算能力作为像水和电一样的公用事业提供给用户的理念, 这成为云计算思想的起源。在 20 世纪 80 年代网格计算、90 年代公用计算, 21 世纪初虚拟化技术、面向服务的体系结构 (Service-Oriented Architecture, SOA)、软件即服务 (Software as a Service, SaaS) 应用的支撑下, 云计算作为一种新兴的资源使用和交付模式逐渐为学界和产业界所认知。继个人计算机变革、互联网变革之后, 云计算被看作第三次 IT 浪潮。

云计算是指 IT 资源的交付和使用模式, 通过网络以按需、易扩展的方式获得所需的资源 (硬件、平台、软件)。典型的云计算提供商往往提供通用的网络业务应用, 可以通过浏览器等软件或者其他 Web 服务来访问, 而软件和数据都存储在远程数据中心的服务器上。用户通过计算机、手机等方式接入数据中心, 按自己的需求进行运算。提供资源的网络被称为“云”。“云”中的资源在使用者看来是可以无限扩展的, 并且可以随时获取、按需使用、随时扩展、按使用付费。

1.1.1 云计算的思想起源

在传统计算模式下, 企业建立一套 IT 系统不仅仅需要购买硬件等基础设施, 还需要购买软件的许可证, 需要专门的人员维护。当企业的规模扩大时还要继续升级各种软硬件设施以满足需要。对于企业来说, 计算机等硬件和软件本身并非他们真正需要的, 它们仅仅是完成工作、提供效率的工具而已。对个人来说, 如果想正常使用计算机就需要安装多种软件, 而许多软件是收费的, 对不经常使用该软件的用户来说购买是非常不划算的。可不可以有这样的服务, 能够提供我们需要的所有软件供我们租用? 这样我们只需要在使用时支付少量“租金”即可“租用”到这些软件服务, 为我们节省许多购买软硬件的资金。

我们每天都要用电, 但不是每家自备发电机, 而是由电厂集中提供; 我们每天都要用自来水, 但不是每家都需要有水井, 它由自来水厂集中提供。这种模式极大地节约了资源, 方便了我们的生活。面对计算机给我们带来的困扰, 我们可不可以像使用水和电一样使用计算

机资源? 这些想法最终导致了云计算的产生。

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, 使人们能够像使用水、电、煤气和电话那样使用计算机资源。

云计算模式即为电厂集中供电模式。在云计算模式下, 用户的计算机会变得十分简单, 或许不大的内存、不需要硬盘和各种应用软件, 就可以满足我们的需求, 因为用户的计算机除了通过浏览器给“云”发送指令和接收数据外基本上什么都不用做, 便可以使用云服务提供商的计算资源、存储空间和各种应用软件。这就像连接“显示器”和“主机”的电线无限长, 从而可以把显示器放在使用者的面前, 而主机放在远到甚至计算机使用者本人也不知道的地方。云计算把连接“显示器”和“主机”的电线变成了网络, 把“主机”变成云服务提供商的服务器集群。

在云计算环境下, 用户的使用观念也会发生彻底的变化, 从“购买产品”到“购买服务”转变, 因为他们直接面对的将不再是复杂的硬件和软件, 而是最终的服务。用户不需要拥有看得见、摸得着的硬件设施, 也不需要为机房支付设备供电、空调制冷、专人维护等费用, 并且不需要等待漫长的供货周期、项目实施等冗长的时间, 只需要向云计算服务提供商支付一定的费用, 我就可以马上得到需要的服务。

1.1.2 云计算的定义

云计算的概念自提出之日起就一直处于不断的发展变化中, 很多机构和学者对云计算进行了解读, 但没有形成公认的定义。本节列出几个典型的定义, 使读者从多个角度了解云计算的含义。

(1) 维基百科给出的云计算的定义: 云计算是一种基于互联网的计算方式, 通过这种方式, 共享的软硬件资源和信息可以按需求提供给计算机和其他设备。云计算描述了一种基于互联网的新的 IT 服务增加、使用和交付模式, 通常涉及通过互联网来提供动态易扩展而且经常是虚拟化的资源。

(2) 百度百科给出的云计算的定义: 云计算是分布式计算技术的一种, 其最基本的概念, 是通过网络将庞大的计算处理程序自动分拆成无数个较小的子程序, 再交由多部服务器所组成的庞大系统经搜寻、计算分析之后将处理结果回传给用户。通过这项技术, 网络服务提供者可以在数秒之内, 处理数以千万计甚至亿计的信息, 达到和“超级计算机”同样强大效能的网络服务。

(3) IBM 认为云计算是一种革新的信息技术与商业服务的消费与交付模式。在这种模式中, 用户可以采用按需的自助模式, 通过访问无处不在的网络获得任何地方资源池中被快速分配的资源, 并按实际使用情况进行付费。

(4) Salesforce.com 认为云计算就是一种更友好的业务运行模式。在这种模式中, 用户的应用程序运行在共享的数据中心, 用户只需要通过登录和个性化定制就可以使用这些数据中的应用程序。

(5) 美国国家标准与技术研究院 (National Institute of Standards and Technology, NIST) 对云计算的定义: 云计算是一种无处不在、便捷且按需对一个共享的可配置计算资源 (包括网络、服务器、存储、应用和服务) 进行网络访问的模式, 它能够通过最少量的管理以及与服务提供商的互动实现计算资源的迅速供给和释放。该定义是目前较为公认的云计算的定义。

图 1.1 形象地描述了云计算的定义。云端的计算资源池包含了服务器、计算机桌面、软件

平台、软件应用和存储/数据等计算资源。用户可以使用台式计算机、笔记本电脑、智能手机、平板电脑等终端设备联网获取云端的计算资源。

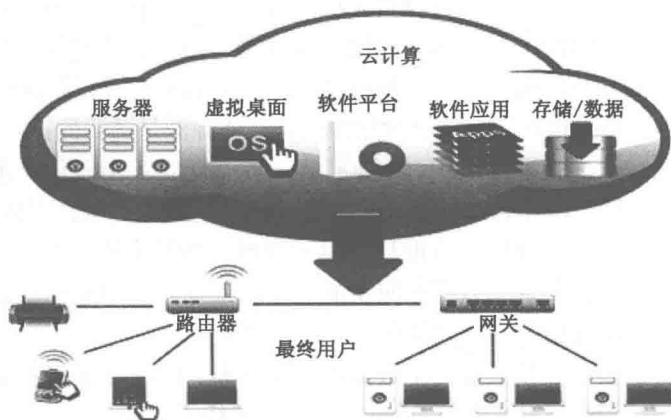


图 1.1 云计算的定义

按照 NIST 对云计算的定义，自助式服务、随时随地使用、可度量的服务、快速资源扩缩和资源池化是云计算的基本特征，如图 1.2 所示。

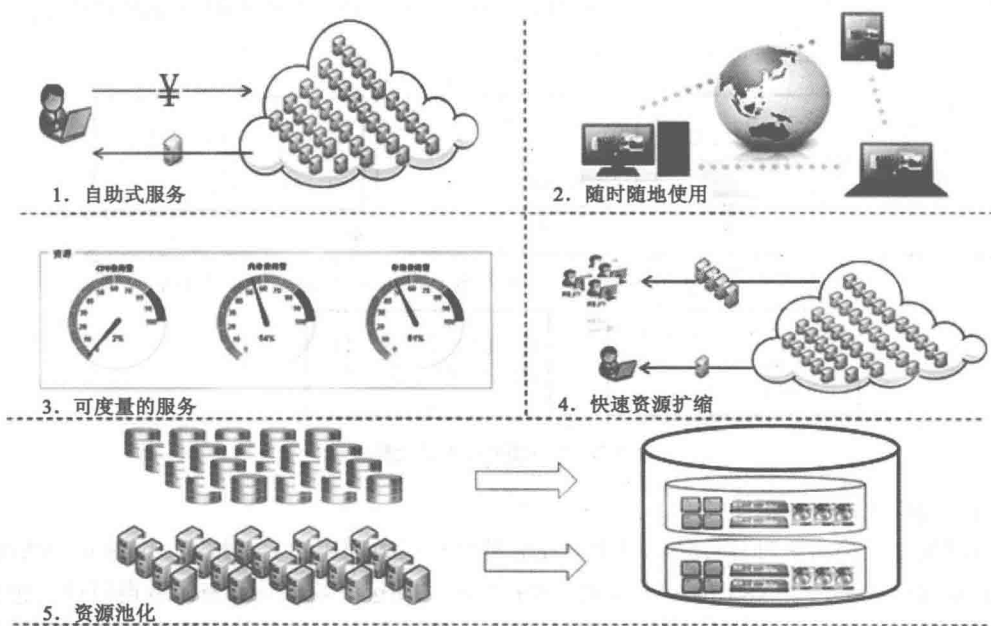


图 1.2 云计算的基本特征

(1) 自助式服务：使用云计算的用户大多是通过自助方式获取资源的。例如，当使用 Amazon 的 EC2 云服务时，用户可以自助选择服务器的操作系统类型、服务器的配置 (micro、small、large) 以及磁盘大小，Amazon EC2 平台便会根据用户的设置分配一台云主机供用户使用。

(2) 随时随地使用：可以通过各种移动设备或瘦客户端（如手机、平板电脑等）连接云端，使用云计算平台提供的服务，打破地理位置和硬件部署环境的限制。

(3) 可度量的服务：云计算平台会对存储、CPU、带宽等资源保持实时跟踪，根据这些

可量化指标对后台资源进行调整和优化。

(4) 快速资源扩缩：用户可以根据自己的需求申请或释放虚拟资源，实现资源的快速扩缩。

(5) 资源池化：云服务提供商的计算资源集中成一个巨大的资源池，这些资源以多租户的方式供用户共享使用。资源的种类包含存储、处理器、内存、带宽等。对云服务的提供者而言，各种底层资源的异构性被屏蔽，边界被打破，所有的资源可以被统一管理和调度，成为云计算资源池，为用户提供按需服务；对用户而言，资源池是透明的，可以按需使用付费。

从技术角度来看，云端软硬件维护、数据管理、安全防护均由云计算服务提供商负责，用户端只需将设备接入即可使用云端的服务，降低了用户的技术门槛，提高了数据的安全性；使用云计算模式，用户将所要使用的数据和应用上传到云端即可随时随地通过任意终端设备进行数据访问和应用体验，实现了数据和应用的共享；云计算资源池具有很强的弹性，用户可以按需使用资源，并仅对使用的资源付费。

从商业角度来看，云计算模式提高了数据中心服务器、网络设备等资源的利用率，降低了用户的 IT 设备建设和维护费用，使用户可以专注于其核心业务。

1.1.3 云计算的发展历程

1. 云计算产生的五大契机

云计算是分布式计算、并行计算、效用计算、虚拟化、网络存储、负载均衡、热备份冗余等传统计算机和网络技术发展融合的产物，更是 SaaS、SOA 等技术混合演进的结果，图 1.3 所示的五大契机更是直接促进了云计算的诞生。

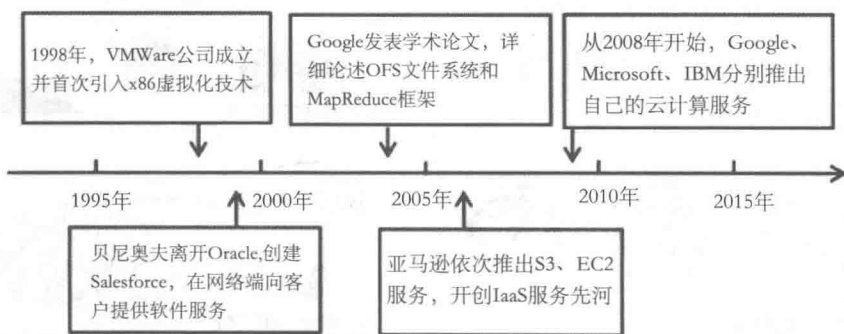


图 1.3 云计算产生的五大契机

(1) 虚拟化技术的诞生

虚拟化技术是将各种计算及存储资源充分整合和高效利用的关键技术。1998 年，VMWare 公司成立并首次引入了 x86 虚拟化技术。虚拟化技术使远程获取计算资源变得可行，也使得资源可以进行动态切分和组合，是将各种计算及存储资源充分整合和高效利用的关键技术。

(2) SaaS 的诞生

1999 年贝尼奥夫离开 Oracle 公司与几个 Oracle 公司原高管一起创办了 Salesforce.com，主要向企业客户销售基于云的 SaaS，也就是从服务器端向客户传递服务，号称“软件”终结者。用户通过瘦客户端自主选择合适的功能模块，即可使用客户关系管理（Customer Relationship Management, CRM）系统。这种模式的优势显而易见：由于软件的维护、升级都在服务器端，客户不必购买昂贵的软硬件而且避免了 CRM 等应用程序的配置和版本更新，软件开发商也不必储备大量的运维服务人员，软件新功能升级可以在服务器端快速进行，自

然而地完成了从 License 到服务的商业模式转型。

Salesforce.com 对云计算的重要意义不仅在于它通过销售 SaaS 服务，成长为一家年销售额超过 20 亿美元的公司，更重要的是 Salesforce.com 第一次将 SaaS 服务大规模地销售给了企业用户，包括像通用电气、荷兰航空、NBC 和 Comcast 这样的国际知名企业。Salesforce.com 的成功第一次证明了基于云计算的服务不仅仅是大型业务系统的廉价替代品，它还可以是真正提高企业运营效率、促进业务发展的解决方案，同时可以在可靠性方面维持一个极高的标准。

（3）Google 发表三大论文

2003 年，Google 在 SOSP 大会上发表了有关分布式存储系统（Google File System, GFS）的论文，在 2004 年的 OSDI 大会上发表了有关 MapReduce 分布式处理技术的论文，在 2006 年的 OSDI 大会上发表了关于 BigTable 分布式数据库的论文。这三篇重量级论文的发表，不仅使大家了解了 Google 搜索引擎背后强大的技术支撑，而且克隆这三项技术的开源产品也大量涌现。比较著名的有克隆 MapReduce 的 Hadoop 系统，克隆 GFS 的 HDFS（Hadoop Distributed File System），以及克隆 BigTable 的 HBase、Hypertable 和 Cassandra 等。

这三篇论文和相关的开源技术极大地普及了云计算中非常核心的分布式技术。

（4）AWS 的推出

在线零售商 Amazon 的 B2C 业务类似于中国的淘宝和京东的业务，平时的流量就很大，特别是到每年特定的时刻（比如圣诞节），流量会急速攀升。为了应对这种情况，Amazon 需要购置远超过其平常使用量的硬件资源以应对这些增加的流量。显然这会造成很大的资源浪费，当发展到一定规模的时候，它发现自己的数据中心在大部分时间只有不到 10% 的利用率，剩下 90% 的资源都是闲置着，这些资源仅有的作用就是被用来缓冲圣诞节购物季这种高峰时段的流量。于是 Amazon 开始寻找一种更有效的方式来利用自己的数据中心，其目的就是将计算资源从单一的、特定的业务中解放出来，在空闲时提供给其他需要的用户。Amazon 首先在企业内部进行实施，得到了很好的回报。随后，Amazon 便将这个服务开放给外部的用户。通过租借硬件资源给公众，以减少浪费，这也就是 Amazon 推出 AWS（Amazon Web Services）的最主要的原因。

在 2006 年初，Amazon 推出了 AWS 的第一款产品 S3（Simple Storage Service，简单存储服务）云存储服务；2006 年 8 月推出了另一款产品 EC2（Elastic Compute Cloud，弹性计算）云基础设施服务。AWS 产品的出现使人们惊奇地发现计算资源原来可以像亚马逊的其他商品一样被销售。之所以说 Amazon 的 EC2 是云计算的一个里程碑，是因为它是业界第一个将基础架构大规模开放给公众用户的云计算服务 IaaS（Infrastructure as a Service，基础设施即服务）。

继 Amazon AWS 之后，各种类似的云计算产品开始大量涌现，关于云计算的报道和解读遍布市场。多个企业纷纷投入云计算市场。在基础设施方面，除 Amazon 之外，IBM、微软等颇具实力的 IT 企业也开始提供云计算服务。

（5）Google 对外提供 GAE 服务

2009 年，Google 开始对外提供 GAE（Google App Engine）服务，这是一个 PaaS（Platform as a Service，平台即服务）服务。Google 搭建了一个完整的 Web 应用开发环境，给用户提供了主机、数据库、互联网接入带宽等资源，用户不必自己购买设备，而是可以在浏览器里面开发和调试自己的代码，然后直接部署到 Google 的云平台上，并对外发布服务。这样的好处是用户不必再担心主机、托管商、互联网接入带宽等一系列运营问题。因此 GAE 可以看作是

托管网络应用程序的平台。

从架构上看，GAE 提供了一套 API，帮助你获取网络数据、发送邮件、存储数据、操作图片、缓存数据，相信以后还会有更多的 API 推出。开发人员在 GAE 的框架内开发，不用再考虑 CPU、内存、分布等复杂和难以控制的问题，初级的程序员按照 GAE 的规范也可以写出高性能的应用。当然，实现高性能也是有代价的。比如不能使用 socket，文件操作、数据查询必须有索引，不同时支持两个不等式做条件的查询等。对于开发而言，多了些约束，少了些选择。可以让开发更加简单，更关注业务。

作为 PaaS 平台的旗帜，Google App Engine 补齐了云计算的产品版图，从此用户可以在基于云的环境中找到绝大部分计算资源。

2. 云计算发展历程的点滴事件

云计算被视为科技界的一次革命，带来工作方式和商业模式的根本性改变，与并行计算、分布式计算和网格计算关系紧密，更是虚拟化、效用计算、SaaS、SOA 等技术混合演进的结果。这几十年来，云计算是怎样一步步演变的呢？让我们回顾一下云计算发展历程中的点滴事件，如表 1.1 所示。

表 1.1 云计算发展历程的点滴事件

时 间	事 件
1959 年	Christopher Strachey 发表虚拟化论文
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	Sun 公司的联合创始人 John Gage 说出了“网络就是计算机”的名言
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
1998 年	VMware 公司成立并首次引入 x86 的虚拟技术
1999 年	Marc Andreessen 创建 LoudCloud，第一个商业化的 IaaS 平台 Salesforce.com 公司成立，宣布“软件终结”革命开始
2004 年	Google 发布 MapReduce 论文 Hadoop 分布式文件系统（HDFS）和 Map-Reduce 被实现
2005 年	Amazon 推出 Amazon Web Services 云计算平台
2006 年	Amazon 推出在线存储服务 S3 和弹性计算云 EC2 等云服务 Sun 推出基于云计算理论的“BlackBox”计划
2007 年	IBM 首次发布云计算商业解决方案，推出“蓝云”计划
2008 年	Salesforce.com 推出 DevForce 平台，Force.com 成为世界上首个 PaaS 应用平台 Google App Engine 发布
2009 年	VMware 推出业界首款云操作系统 Vmware vSphere 4 Google 宣布将推出 Chrome OS 操作系统

时 间	事 件
2010 年	微软正式发布 Windows Azure 公共云计算平台
	开源 IaaS 云计算平台项目 OpenStack 发布
2011 年	Citrix 收购 CloudStack 的前身 Cloud.com 并将其开源
2012 年	VMware 推出业界第一个开源 PaaS 云平台 Cloud Foundry
	Eucalyptus 重新全面开源，并迁移到 Github
2013 年	第一个版本的 Docker 正式发布
	思科等公司发起成立了 Open Daylight，与 Linux 基金会合作，开发 SDN 控制器
2014 年	Spark 成为了 Apache 软件基金会的顶级项目，其开源生态系统得到了大幅增长
	OpenStack 开始支持 Docker 容器
2015 年	基于 Docker 和 Kubernetes 容器技术的 OpenShift 3 发布
	Docker 宣布其正式在 Linux 基金会指导下建立产业联盟

1.2 云计算的服务类型

云计算是一种新的技术，也是一种新的服务模式。云计算服务提供方式包含基础设施即服务（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）3 种类型。IaaS 提供的是用户直接使用计算资源、存储资源和网络资源的能力，PaaS 提供的是用户开发、测试和运行软件应用的能力，SaaS 是将软件以服务的形式通过网络提供给用户使用。

这三类云计算服务的层次关系如图 1.4 所示。IaaS 处于整个架构的底层；PaaS 处于中间层，可以利用 IaaS 层提供的各类计算资源、存储资源和网络资源来建立平台，为用户提供开发、测试和运行环境；SaaS 处于最上层，既可以利用 PaaS 层提供的平台进行开发，也可以直接利用 IaaS 层提供的各种资源进行开发。

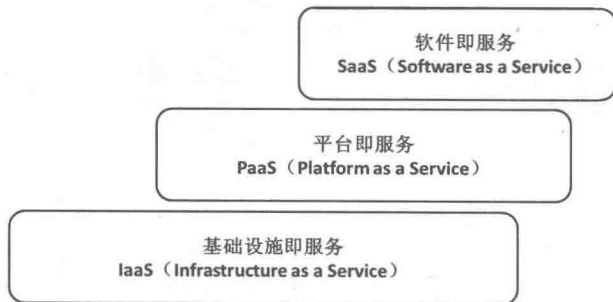


图 1.4 云计算服务层次

云计算服务提供商可以专注于自己所在的层次，无需拥有三个层次的服务能力，上层服务提供商可以利用下层的云计算服务来实现自己计划提供的云计算服务。例如，提供 PaaS 服务的云计算提供商可以基于 Amazon EC2 的 IaaS 平台来建立自己的 PaaS 服务平台，向用户提

供 PaaS 服务;提供 SaaS 服务的云计算提供商可以使用 Google App Engine 平台或者微软的 Azure 平台开发、测试和运行自己的软件,向用户提供 SaaS 服务。

1.2.1 IaaS: 基础设施即服务

IaaS 是把计算、存储、网络以及搭建应用环境所需的一些工具当成服务提供给用户,使得用户能够按需获取 IT 基础设施。IaaS 主要由计算机硬件、网络、存储设备、平台虚拟化环境、效用计费方法、服务级别协议等组成。

通过使用 IaaS,用户无需购买计算机和相关系统软件,也不需要购买存储设备,更省去了维护和升级计算机的烦恼。用户只需要购买 IaaS 的服务就可以获得计算和存储资源,并在这些资源上构建自己的平台和应用。IaaS 服务模式如图 1.5 所示。

IaaS 具有以下特点:

- 把 IT 资源以服务的方式提交给用户;
- 基础设施可以动态扩展,即可以根据应用的需求动态增加或者减少资源;
- 计费服务灵活多变,按实际使用的资源进行计费;
- 多租户,相同的基础设施资源可以同时提供给多个用户共享使用;
- 企业级的基础设施,不仅仅可以为个人用户提供 IT 资源,而且可以满足中小企业的 IT 资源需求,使得它们可以从聚集的计算资源池中获利。

从业务上来看,IaaS 需要把计算、存储、网络等 IT 基础设施通过虚拟化整合和复用后,通过互联网提交给用户。提供的 IT 基础设施要能够根据应用的运行情况进行动态扩展或收缩,并按照实际的使用量进行计费。作为 IaaS 服务提供商,需要重点解决资源提供和运营管理两个问题。



图 1.5 IaaS 服务模式图

Amazon EC2 是第一个商用的 IaaS 平台,它的底层使用了 Citrix 的 Xen 虚拟化技术,以 Xen 虚拟机的形式向用户动态提供计算资源。此外,Amazon 公司还提供了简单存储服务 (Simple Storage Service, S3),以及弹性块存储服务 (Elastic Block Store, EBS) 等多种 IT 基础设施服务。Amazon EC2 向虚拟机提供动态 IP 地址,并且具有相应的安全机制来监控虚拟机节点间的网络,限制不相干节点间的通信,从而保证了用户通信的私密性和数据的安全性。从收费模式来看,Amazon EC2 按照用户使用资源的数量和时间进行计费,从真正意义上实现了云计算的“按使用付费”的模式。