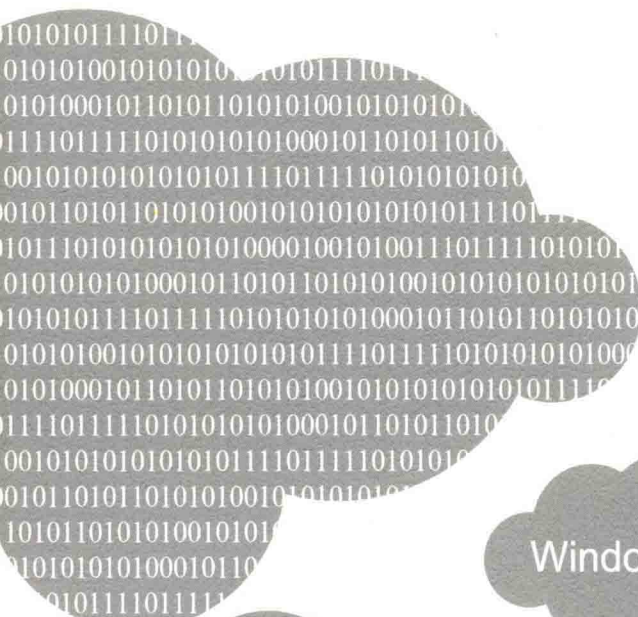




Hadoop

VMware

Amazon

Windows Azure



典型
云计算
赵新芬 编著
平台与应用教程

典型云计算平台 与应用教程

赵新芬 编著

电子工业出版社

Publishing House of Electronics Industry

北京 • BEIJING



内 容 简 介

本书从应用的角度出发,首先概括了云计算的相关概念、专业工具及程序设计等知识,从而揭开云计算的神秘面纱,带领读者全面了解云计算的发展状况;接着分别介绍了云计算在各企业中应用相关平台,主要包括微软云计算、Google 云计算、Amazon 云计算、VMware 云计算等内容,从而让读者切身体会到云计算在各企业中的应用;最后介绍了云计算的仿真器及其实际应用,让读者全盘掌握云计算。

本书可以作为高等学校教材,同时可作为培训教材,也可供广大科研人员、学者、程序员自学或参考使用。

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

典型云计算平台与应用教程/赵新芬编著. —北京:电子工业出版社, 2013.4

ISBN 978-7-121-19709-3

I. ①典... II. ①赵... III. ①计算机网络—教材 IV. ①TP393

中国版本图书馆 CIP 数据核字(2013)第 038351 号

策划编辑:陈韦凯 特约编辑:刘海霞

责任编辑:陈韦凯

印 刷:三河市鑫金马印装有限公司

装 订:三河市鑫金马印装有限公司

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:787×1 092 1/16 印张:28.5 字数:730 千字

印 次:2013 年 4 月第 1 次印刷

印 数:4 000 册 定价:65.00 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。

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。

服务热线:(010) 88258888。

前言

信息技术从诞生以来一直在经历快速的发展和变革。新技术、新产品、新思维的层出不穷催生了一代又一代崭新的商业模式和理念，它们在使得市场竞争越发激烈的同时，也带来了令人振奋的无限机遇。

早在大型机盛行的 20 世纪五六十年代，计算资源就是采用“租借”的方式对外提供服务的。IBM 公司当时的首席执行官 Thomas Watson 曾预言道：“全世界只需要五台计算机”，过去三十年的 PC 大繁荣似乎正在推翻这个论断，人们也常常引用这个例子，来说明信息产业的不可预测性。然而，信息技术变革并不总是直线前进，而是螺旋式上升的。半导体、互联网和虚拟化技术的飞速发展使得业界不得不重新思考这一构想，这些支撑技术的成熟让我们有可能把全世界的数据中心进行适度的集中，从而实现规模化经济效应。人们只需远程租用这些共享资源，从而消除企业为了使用信息技术而进行的数据中心构建、硬件采购、软件安装、系统维护等环节。

“云计算（Cloud Computing）”被称为继个人计算机、互联网之后的第三次信息化革命，通过与相关技术创新要素、商业模式创建要素形成有机互动，“云计算”将成为推动电信业乃至广义 ICT 产业下一轮突破发展的重要驱动力。

“云计算”时代已经来临，未来不管是软件服务、信息分析、影音发送还是联机游戏等，都会由云计算来运行，因为云计算系统具有庞大的计算能力与存储空间，用户只要通过网络与云计算系统连接，就可取得云计算服务资源，而且消费者只要按自己的使用量付费，就像付水电费一样方便。如果是个人使用者，则可享受云计算软件服务在生活与工作所带来的便利性；如果为一般中小企业，则可以通过云计算平台开发设计应用程序，来经营云计算软件服务事业；如果是有能力提供云计算基础设施的大型企业，致力于打造数以万计的云计算硬件与架构，则需要自行构建云计算系统，打造自己的私有云与企业云。

“云计算”的影响将是深远的，其将彻底改变 IT 产业的架构和运行方式。可预见，高性能计算机、高端服务器、高端存储器和高端处理器的市场将被数量众多、低成本、低能耗和高性价比的“云计算”硬件市场所挤占；绝大多数软件将以服务方式呈现，甚至连大多数游戏都将在“云”里运行；呼叫中心、网络会议中心、智能监控中心、数据交换中心、视频监控中心和销售管理中心等，将越来越向某些“云计算”设施集中而获取更多的性价比。放眼远眺，“云计算”将与网络计算融为一体，实现“云计算”平台间的互操作和资源共享，实现紧耦合高性能科学计算与松耦合高吞吐量商业计算的融合，使互联网上的主要计算设施融为一个有机整体，称为“云格”。

云计算平台可以划分为 3 类：以数据存储为主的存储型云平台、以数据处理为主的计算型云平台，以及计算和数据存储处理兼顾的综合云计算平台。

云计算平台具有以下特点。

（1）服务无处不在

用户只需要一台具备基本计算能力的计算设备，以及一个有效的互联网连接，就可以随时随地使用该服务。从这个意义来讲，任何联网的应用，都具备成为云计算平台的

潜力。

(2) 具备进入成本

用户具备使用该服务的需求，但是并不具备独立提供该服务的经济或者技术条件。例如，某些企业需要定期地进行大规模的运算，但是并不值得专门为此购置一台具备大规模运算能力的计算设备。超算中心通过发展客户群让多个用户来分担超级计算机的成本，使得其用户能够在不拥有计算设备的情况下以较小的成本完成计算任务。

(3) 用户决定应用

云计算平台提供计算能力（包括处理器、内存、存储、网络接口），但是并不关心用户的应用类型。用户利用云计算平台所提供的计算能力，并且充分考虑云计算平台所设定的（技术和经济）限制，开发出丰富多彩的应用。

本书共 8 章，分别如下。

第 1 章：云计算简介，主要包括云计算基础、云计算关键技术研究、云标准及云安全等内容。

第 2 章：介绍云计算的专业工具及程序设计，主要包括云计算应用 Java 软件、云计算应用软件 Platform Symphony、云计算应用软件 Python 及云计算应用软件 MATLAB 等内容。

第 3 章：介绍微软云计算，主要包括 Windows Azure 计算平台、微软云计算数据库 SQL Azure 及 Windows Azure AppFabric 等内容。

第 4 章：介绍 Google 云计算，主要包括 Hadoop 概述、Hadoop 常用命令、分布式处理数据 MapReduce 及分布式服务 Zookeeper 等内容。

第 5 章：介绍 Amazon 云计算，主要包括弹性计算云 EC2、存储服务 S3、简单队列服务 SQS 及简单数据库服务 Simple DB 等内容。

第 6 章：介绍 VMware 云计算，主要包括 VMware 云主品简介、ESX/ESXi 概述、VMware vSphere 分布式服务等内容。

第 7 章：介绍云计算仿真器 CloudSim，主要包括 CloudSim 体系结构、CloudSim 的功能及扩展、CloudSim 的使用方法等内容。

第 8 章：介绍云计算的综合应用，主要包括提升办公效率及云计算在移动通信信令监控与查询的综合实例等内容。

本书主要由赵新芬编写，此外，参加编写的还有李晓东、丁伟雄、雷晓平、李娅、杨文茵、何正风、赵书兰、赵书梅、栾颖、刘志为、周灵、周品、张德丰、余智豪。

由于作者的水平有限，加之时间紧凑，书中难免会存在不足之处，敬请广大读者批评指正。

编著者

目 录

第 1 章 云计算简介	1
1.1 云计算基础	1
1.1.1 云产生的背景	1
1.1.2 云计算定义	4
1.1.3 云计算体系架构	4
1.1.4 云计算的分类	7
1.1.5 云计算发展趋势	12
1.1.6 云计算研究方向	13
1.1.7 云计算的实现机制	15
1.2 云计算关键技术研究	16
1.2.1 虚拟化技术	17
1.2.2 数据存储技术	18
1.2.3 资源管理技术	20
1.2.4 能耗管理技术	20
1.2.5 云监测技术	22
1.3 云计算应用实例	24
1.3.1 Google 的云计算平台应用实例	24
1.3.2 IBM “蓝云” 计算平台应用实例	25
1.3.3 Amazon 的弹性计算云应用实例	26
1.4 云标准	27
1.4.1 云标准背景	27
1.4.2 云计算潜在标准化需求分析	27
1.4.3 云计算标准制定	29
1.4.4 云标准现状	29
1.4.5 云标准性能评估	31
1.4.6 云计算标准发展趋势	33
1.5 云计算产业的优势	33
1.5.1 优化产业布局	33
1.5.2 推进专业分工	34
1.5.3 提升资源利用率	35
1.5.4 减少初期投资	36
1.5.5 降低运营成本	37
1.5.6 产生新创价值	37
1.6 云安全	38
1.6.1 云安全与网络安全的差别	39

1.6.2	云安全常见问题	39
1.6.3	云安全研究方向	41
1.6.4	云安全的难点问题	41
1.6.5	云安全新增及增强功能	42
1.6.6	云安全应用	43
1.6.7	安全示例	49
1.6.8	云安全的问题	53
1.7	云生命周期	54
1.8	云计算优缺点	55
第 2 章	云计算的专业工具及程序设计	57
2.1	云计算应用 Java 软件	58
2.1.1	创建 Java 云计算项目	58
2.1.2	创建 HTTP Servlet 程序类别	59
2.1.3	Web.xml 文档设置	60
2.1.4	设置 appengine-web.xml 文档	60
2.1.5	添加 Google 用户服务	61
2.1.6	设计用户界面	62
2.1.7	访客留言板表单的创建	63
2.1.8	利用 JDO 访问 datastore	64
2.1.9	创建 JDO 类别	65
2.1.10	创建 PersistenceManager 类别	67
2.1.11	JOD 的创建及保存	68
2.1.12	JDO Query Language 查询存储物件	69
2.1.13	界面外观修饰	71
2.2	云计算应用软件 Platform Symphony	72
2.2.1	Symphony DE 简介	72
2.2.2	Platform Symphony 架构	74
2.2.3	使用 Symphony DE 编写程序	74
2.3	云计算应用软件 Python	84
2.3.1	创建 Python 项目	84
2.3.2	app.yaml 设置文档	84
2.3.3	创建 webapp 程序架构	85
2.3.4	利用 datastore 保存资料	87
2.3.5	大胜 Django 留言范本	89
2.3.6	界面外观修饰	91
2.3.7	创建首页 index.html	91
2.3.8	app.yaml 文档更改	92
2.3.9	测试 Python 云计算程序	92

2.3.10	发布 Python 云计算程序	93
2.4	云计算应用软件 MATLAB	93
2.4.1	MATLAB 的简单使用	93
2.4.2	利用 PCT 提高计算速度	97
2.4.3	更多的工具运行在云计算环境中	106
2.5	云计算应用软件 C#与 VB.NET	113
2.5.1	C#与 VB.NET 开发环境创建	113
2.5.2	C#与 VB.NET 云计算程序设计	114
2.5.3	使用 C#处理云计算服务回应的 HML 资料	118
2.5.4	使用 VB.NET 处理云计算服务回应的 XML 资料	118
2.6	动手做自己的云计算	119
2.6.1	系统总体分析	119
2.6.2	管理节点程序设计与分析	120
2.6.3	子节点程序分析	130
2.6.4	客户端 API 设计	140
第 3 章	微软云计算	141
3.1	Windows Azure 计算平台	141
3.1.1	Windows Azure 概述	142
3.1.2	在 Azure 开发平台并创建运行项目	143
3.1.3	Windows Azure 存储服务	149
3.1.4	Windows Azure 存储服务特点	176
3.2	微软云计算数据库 SQL Azure	176
3.2.1	SQL Azure 架构	177
3.2.2	使用 SQL Azure 的理由	177
3.2.3	使用 SQL Azure Database 的好处	179
3.2.4	SQL Azure 与 SQL Server 的比较	179
3.2.5	SQL Azure 关键技术	181
3.2.6	在应用程序中使用 SQL Azure	190
3.3	Windows Azure AppFabric	192
3.3.1	服务总线	193
3.3.2	访问控制 (Access Control Service, ACS)	207
3.3.3	分布式缓存	214
第 4 章	Google 云计算	216
4.1	Hadoop 概述	216
4.1.1	Hadoop 的功能与作用	216
4.1.2	Hadoop 的优点	217
4.1.3	Hadoop 的主要子项目	217
4.1.4	Hadoop 体系结构	217

4.2	Hadoop 部署与开发	220
4.2.1	在 Linux 下安装 Hadoop	220
4.2.2	在 Windows 下安装 Hadoop	223
4.3	Hadoop 常用命令	227
4.3.1	FS shell	227
4.3.2	管理与更新	231
4.3.3	用户命令	232
4.4	Hadoop 输入与输出	238
4.4.1	MapReduce 输入与输出	238
4.4.2	HDFS 的输入与输出	238
4.5	分布式处理数据 MapReduce	241
4.5.1	编程模型	241
4.5.2	MapReduce 实现机制	242
4.5.3	MapReduce 实例分析	244
4.6	分布式表 HBase	250
4.6.1	数据模型	250
4.6.2	概念视图	251
4.6.3	物理视图	252
4.6.4	系统架构	253
4.6.5	HBase 的安装与部署	254
4.6.6	HBase 的实例分析	257
4.7	分布式服务 Zookeeper	260
4.7.1	Zookeeper 概述	260
4.7.2	Zookeeper 工作原理	260
4.7.3	Zookeeper 安装与部署	261
4.7.4	Zookeeper API	265
4.7.5	Zookeeper 实例分析	268
4.8	数据仓库 Hive	276
4.8.1	Hive 结构	276
4.8.2	Hive 数据存储	277
4.8.3	Hive 安装	278
4.8.4	Hive 的扩展特性	280
4.8.5	Hive 实例分析	285
4.9	数据序列化系统 Avro	288
4.9.1	数据序列化	289
4.9.2	数据排序顺序	292
4.9.3	Avro 的安装	292
4.9.4	RPC 通信实现	293

4.9.5	Avro 的 Java 实现	296
第 5 章	Amazon 云计算	302
5.1	Amazon 相关概念	302
5.2	弹性计算云 EC2	304
5.2.1	EC2 主要特性	305
5.2.2	EC2 基本架构	305
5.2.3	Amazon EC2 的优势与特点	306
5.2.4	EC2 实战	309
5.3	存储服务 S3	312
5.3.1	S3 基本概念	313
5.3.2	S3 一致性	313
5.3.3	S3 安全措施	314
5.3.4	Amazon Web Services 和 S3 入门	316
5.3.5	使用 GAE 与 AWS S3 部署	322
5.4	简单队列服务 SQS	328
5.4.1	SQS 的模型及特性	328
5.4.2	SQS 框架的概念	329
5.4.3	Amazon Web Services 和 SQS 入门	330
5.4.4	使用 Amazon SQS 进行基于云计算的消息传送	334
5.4.5	通过 Zend Framework 使用 Amazon SQS	337
5.5	简单数据库服务 Simple DB	343
5.5.1	SDB 的基本概念	344
5.5.2	SDB 的使用	346
第 6 章	VMware 云计算	350
6.1	VMware 云产品概述	350
6.1.1	VMware 云操作系统	350
6.1.2	VMware 功能	351
6.1.3	VMware 特征	352
6.1.4	VMware 带来的好处	353
6.1.5	VMware 云计算工作模式	353
6.2	VMware 云产品简介	354
6.2.1	VMware 云三层框架	355
6.2.2	VMware vSphere 架构	356
6.2.3	云操作系统 vSphere	359
6.2.4	底层架构服务 vCloud Service Director	360
6.2.5	虚拟桌面产品 VMware View	361
6.3	ESX/ESXi 概述	362
6.3.1	ESX 与 EXSi 的比较	363

6.3.2	VMware ESXi 虚拟化管理程序体系结构的优点	364
6.4	VMware vSphere 分布式服务	364
6.5	VMware 管理端 vCenter Server	366
6.6	VMware 的安装	368
第 7 章	云计算仿真器 CloudSim	375
7.1	CloudSim 体系结构	375
7.1.1	CloudSim 核心模拟引擎	376
7.1.2	CloudSim 层	378
7.2	CloudSim 的功能及扩展	380
7.3	CloudSim 的使用方法	381
7.3.1	CloudSim 环境配置	381
7.3.2	使用 CloudSim 仿真的步骤	381
7.4	CloudSim 实例分析	382
第 8 章	云计算的综合应用	407
8.1	提升办公效率	407
8.1.1	Excel 概述	407
8.1.2	用 Excel 对蒙特卡罗模拟	414
8.1.3	云计算与 Excel 的集成	418
8.2	云计算在移动通信信令监控与查询的应用实例	428
8.2.1	分析与设计	428
8.2.2	实现代码	431
参考文献		444

第1章 云计算简介

云计算（Cloud Computing）是当前发展十分迅猛的新兴产业，被认为微型计算机、互联网后的第三次 IT 革命，是互联网发展的大趋势。它不仅是互联网技术发展、优化和组合的结果，也为整个社会信息化带来了全新的服务模式，为形成信息化经济时代基于访问权的商业模式提供了可能，将对人类社会生活带来重大变革。现在，越来越多的国家把云计算发展提升到了国家战略层面，我国也将云计算作为新一代信息技术列为“十二五”规划重点扶植的战略新兴产业，各级政府和社会各界对此均高度重视并开始积极探索。

鉴于云计算在信息化进程中的战略地位，云计算安全的重要性也不言而喻，它不仅是广大用户选择云计算应用时的首要考虑因素，是云计算实现健康可持续发展的基础，也是网络安全领域新的探索。充分考虑并加强云计算安全不仅能增强用户对云计算服务的信心，促进云计算市场的成熟和发展，也能极大推动网络安全领域技术和应用的发展。

1.1 云计算基础

什么是云计算？云计算是一种基于因特网的超级计算模式，在远程的数据中心，几万甚至几千万台电脑和服务器连接成一片。因此，云计算甚至可以让你体验每秒超过 10 万亿次的运算能力，如此强大的运算能力几乎无所不能。用户通过电脑、笔记本、手机等方式接入数据中心，按各自的需求进行存储和运算。

1.1.1 云产生的背景

有人说云计算是技术革命的产物，也有人说云计算只不过是已有技术的最新包装，是设备厂商或软件厂商新瓶装旧酒的一种商业策略。我们认为，云计算是社会、经济的发展和需求的推动、技术进步，以及商业模式转换共同作业的结果。

1. 经济方面

1) 全球经济一体化

后危机时代加速了全球经济一体化的发展。实践证明，国家和地区的区位优势 and 比较优势：自发地寻租，基于成本考虑，价值链的协作者自发整合；基于效率考虑，协同效应需要弹性的业务流程支持。对成本和效率的需求将促进云计算的加速发展。

2) 日益复杂的世界和不可确定性的“黑天鹅”现象

在复杂的世界面前，不确定因素在更快、更广地涌现，计划跟不上变化，任何一台精于

预测的机器也无法准确预测到“黑天鹅”现象的发生（不可预知的未来，一旦发生，影响力极大，事前无法预测，事后有诸多理由解释）。实时的信息获取和全面的信息分析有助于管理复杂性，而按需即用的计算资源、随需应变的业务流程将“黑天鹅”的两面影响降到最小。实时的、覆盖全网的、随需应变的云计算的作业显而易见。

3) 需求是云计算的发展动力

IT 设施要成为社会基础设施，现在面临高成本的瓶颈，这些成本至少包括人力成本、资金成本、时间成本、使用成本、环境成本。云计算带来的益处是显而易见的：用户不需要专门的 IT 团队，也不需要购买、维护、安放有形的 IT 产品，可以低成本、高效率、随时按需使用 IT 服务；云计算服务提供商可极大地提高资源（硬件、软件、空间、人力能源等）的利用率和与业务响应速度，有效聚合产业链。

2. 社会层面

1) 数字一代的崛起

未来的世界在网上，世界的未来在云中。根据埃森哲的调查，中国网民数在 2009 年达到 3.84 亿，超过美国和日本的总和，预计这一数字到 2015 年将增加到 6.5 亿以上。到 2015 年，预计互联网的渗透率将从目前的 29%增加到接近 50%，在中国广大的农村人口中渗透率接近 40%以上。

2) 消费行为的改变

社交网络将现实生活中的人际关系以实名制的方式复制到虚拟世界中，未来的网络的发展将是实名制、基于信任和社交化。在线上线下两个世界，半人马型消费者（美国沃顿商学院营销系主任约瑞姆·杰瑞·温德等，《聚合营销——与半人马并驾齐驱》）互相影响，进而影响着为之服务的商业社会和政府行为（如 Dell 基于 Twitter 的营销，广东警方使用微博与民众交流，香港官员使用 Facebook 与民众直接对话）。如图 1-1 所示，历经 10 年的互联网技术和市场的发展，云计算使得“数字一代”崛起成为可能。

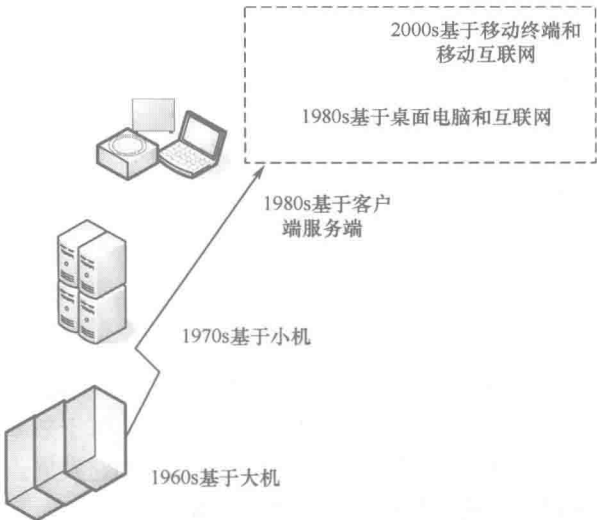


图 1-1 云计算对“数字一代”提供服务效果图



3 亿中国宽带用户中, 92% (年龄大于 13 岁) 的用户参与到社会化媒体中, 而美国仅仅为 76%; 中国拥有超大规模的社会化媒体的内容贡献者, 他们使用博客、微博、社区、视频和图片分享等形式; 43% 的中国宽带用户 (约 1.05 亿) 会使用论坛和 BBS; 在中国, 25~29 岁的年轻上班族是社交媒体的最活跃用户, 和其他的年龄段的互联网用户相比, 他们更依赖在线交流的方式; 37% 的博主 (约 2900 万) 每天都会更新博客; 以一个星期为例, 4100 万的中国人是重度的社会化媒体使用者 (有 6 个以上的线上活动) 会和 54 个人建立联系。云计算是对数字一代消费者提供服务的回答。

3. 政治层面

1) 社会转型

出口型向内需型社会转型, 如何满足人民大众日益增长并不断个性化的需求是一项严峻的挑战。

2) 产业升级

制造型向服务型、创新型的转变。

3) 政策支持

“十二五”规划对物联网、三网融合、移动互联网, 以及云计算战略的大力支持。

4. 技术方面

1) 技术成熟

技术是云计算发展的基础。首先是云计算自身核心技术的发展, 如硬件技术、虚拟化技术 (计算虚拟化、网络虚拟化、存储虚拟化、桌面虚拟化、应用虚拟化)、海量存储技术、分布式并行计算、多租户架构、自动管理与部署; 其次是云计算赖以存在的移动互联网技术的发展, 如高速、大容量的网络, 无处不在的接入, 灵活多样的终端, 集约化的数据中心, Web 技术。

可以将云计算理解为 8 个字: “按需即用, 按需应变”, 使之实现的各项技术已基本成熟 (分布式计算、网络计算、移动计算等)。

2) 企业 IT 的成熟和计算能力过剩

社会需求的膨胀、商业规模的扩大导致企业 IT 按峰值设计, 但需求的波动性却事实上使大量计算资源被闲置。企业内部的资源平衡带来私有云需求, 外部的资源协作促进公有云的发展。

商业模式是云计算的内在要求, 是用户需求的外在体现, 并且云计算技术为这种特定商业模式提供了现实可能性。从商业模式的角度看, 云计算的主要特征是以网络为中心、以服务为产品形态、按需使用与付费, 这些特征分别对应于传统的用户自建基础设施、购买有形产品或介质 (含 licence)、一次性买断模式是一个颠覆性的革命。

从纯粹的技术角度看, 云计算是很多技术的自然发展、精心优化与组合的结果, 是这些技术的集大成者; 另一方面, 如果同时考虑到商业模式, 那么可以断言, 云计算将给整个社会信息化带来革命性的改变。所以, 绝不能离开技术谈云计算, 否则, 有“忽悠”之嫌; 也不能离开商业模式谈云计算, 否则, 云计算就是无源之水、无根之木。



1.1.2 云计算定义

在 IBM 的技术白皮书“Cloud Computing”中对云计算定义：“云计算一词用来同时描述一个系统平台或者一种类型的应用程序。一个云计算的平台按需进行动态的部署（provision）、配置（configuration）、重新配置（reconfigure）及取消服务（deprovision）等。在云计算平台中的服务器可以是物理的服务器或者虚拟的服务器。高级的计算云通常包含一些其他的计算资源，如存储区域网络（SANs）、网络设备、防火墙及其他安全设备等。云计算在描述应用方面，它描述了一种可以通过互联网 Internet 进行访问的可扩展的应用程序。“云应用”使用大规模的数据中心，以及功能强劲的服务器来运行网络应用程序与网络服务。任何一个用户可以通过合适的互联网接入设备，以及一个标准的浏览器就能够访问一个云计算应用程序。”

上述定义给出了云计算两个方面的含义：一方面描述了基础设施，用来构造应用程序，其地位相当于 PC（个人计算机）上的操作系统；另一方面描述了建立在这种基础设施之上的云计算应用。在与网格计算的比较上，网格程序是将一个大任务分解成很多小任务并行运行在不同的集群及服务器上，注重科学计算应用程序的运行。而云计算是一个具有更广泛含义的计算平台，能够支持非网格的应用，例如，支持网络服务程序中的前台网络服务器、应用服务器、数据库服务器三层应用程序架构模式，以及支持当前 Web 2.0 模式的网络应用程序。云计算是能够提供动态资源池、虚拟化和高可用性的下一代计算平台。现有的云计算实现使用的技术体现了以下 3 个方面的特征。

（1）硬件基础设施架构在大规模的廉价服务器集群之上。与传统的性能强劲但价格昂贵的大型机不同，云计算的基础架构大量使用了廉价的服务器集群，特别是 x86 架构的服务器。节点之间的巨联网络一般也使用普遍的千兆以太网。

（2）应用程序与底层服务协作开发，最大限度地利用资源。传统的应用程序建立在完善的基础结构（如操作系统）之上，利用底层提供的服务来构造应用。而云计算为了更好地利用资源，采用了底层结构与上层应用共同设计的方法来完善应用程序的构建。

（3）通过多个廉价服务器之间的冗余，使用软件获得高可用性。由于使用了廉价的服务器集群，节点的失效将不可避免。并且会有节点同时失效的问题。因此，在软件设计上需要考虑节点之间的容错问题，使用冗余的节点获得高可用性。

通过上面的技术手段，云计算达到了两个分布式计算的重要目标：可扩展性和高可用性。可扩展性表达了云计算能够无缝地扩展到大规模的集群之上，甚至包含数千个节点同时处理。高可用性代表了云计算能够容忍节点的错误，甚至有很大一部分节点发生失效也不会影响程序的正确运行。

1.1.3 云计算体系架构

前面提到，云计算不仅是一种技术，更是一种服务模式。从技术的角度看，它包括了互联网技术（包括虚拟化、分布式处理、在线软件等）的发展演进。然而当我们跳出技术的视野，从服务模式的角度来梳理云计算，会发现云计算所提供的 3 类服务模式几乎将整个 IT

生态系统全面覆盖，因此，从服务模式的角度理解云计算将更为清晰。

目前，普遍认为云计算的服务模式可以分为基础设施服务（IaaS，Infrastructure as a Service）、平台即服务（PaaS，Platform as a Service）和软件即服务（SaaS，Software as a Service）3类。其中，IaaS 通常面向企业用户，提供包括服务器、存储、网络和管理工具在内的 IT 基础设施，可以帮助企业削减 IT 资源的建设成本和运维成本。PaaS 通常面向互联网应用开发者，提供简化的分布式软件开发、测试和部署环境，它屏蔽了分布式软件开发底层复杂的操作，使得开发人员可以快速开发出基于云计算平台的高性能、高可扩展的互联网应用。SaaS 通常面向个人用户，提供各种各样的在线软件服务。这 3 类服务模式被认为云计算体系架构的 3 个层次，但它们在技术实现上并没有必然的联系，SaaS 可以在 IaaS 的基础上实现，也可以在 PaaS 的基础上实现，也可以独立实现；类似地，PaaS 可以在 IaaS 的基础上实现，也可以独立实现。云计算体系架构如图 1-2 所示。



图 1-2 云计算技术体系架构

1. PaaS

PaaS是指云计算服务商提供应用服务引擎，如互联网应用程序接口（API）或运行平台，用户基于服务引擎构建该类服务。PaaS是基于SaaS发展起来的，其将软件研发的平台作为一种服务，以SaaS的模式提交给用户，可以加快SaaS的发展，尤其是加快SaaS应用的开发速度。从用户角度来说，这意味着他们无须自行建立开发平台，也不会不同平台兼容性方面遇到困扰；从供应商的角度来说，可以进行产品多元化和产品定制化。

Salesforce.com公司的云计算结构称为Force.com。该平台作为一个服务运行在Internet 上，是完全即时请求的，收费是以每登录为基础的。让更多的独立软件提供商成为其平台的客户，从而开发出基于他们平台的多种SaaS 应用，使其成为多元化软件服务供货商（Multi Application Vendor），扩展了其业务范围。

2. IaaS

IaaS 是把厂商的由多台服务器组成的“云端”基础设施，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。这是一种托管型硬件方式，用户付费使用厂商的硬件设施。例



如, Amazon Web 服务(AWS)、IBM 的 BlueCloud 等均是將基础设施作为服务出租。IaaS 的优点是用户只需低价的硬件, 按需租用相应计算能力和存储能力, 大大降低了用户在硬件上的开销。目前, 以 Google 云应用最具代表性, 例如, GoogleDocs、GoogleApps、Googlesites、云计算应用平台 GoogleApp Engine。GoogleDocs 是最早推出的云计算应用, 是软件(即服务思想)的典型应用。它是类似于微软的 Office 的在线办公软件, 它可以处理和搜索文档、表格、幻灯片, 并可以通过网络和其他人分享并设置共享权限。Google 文件是基于网络的文字处理和电子表格程序, 可提高协作效率, 多名用户可同时在线更改文件, 并可以实时看到其他成员所作的编辑。用户只需一台接入互联网的计算机和可以使用 Google 文件的标准浏览器即可在线创建和管理、实时协作、权限管理、共享、搜索能力、修订历史记录功能, 以及随时随地访问的特性, 大大提高了文件操作的共享和协同能力。GoogleApps 是 Google 企业应用套件, 使用户能够处理日渐庞大的信息量, 随时随地保持联系, 并可与其他同事、客户和合作伙伴进行沟通、共享和协作。它集成了 Cmail、GoogleTalk、Google 日历、GoogleDocs 及最新推出的云应用 GoogleSites、API 扩展及一些管理功能, 包含了通信、协作与发布、管理服务三方面的应用, 并且拥有着云计算的特性, 能够更好地实现随时随地协同共享。另外, 它还具有低成本的优势和托管的便捷, 用户无须自己维护和管理搭建的协同共享平台。Googlesites 是 Google 最新发布的云计算应用, 作为 GoogleApps 的一个组件出现。它是一个侧重于团队协作的网站编辑工具, 可利用它创建一个各种类型的团队网站, 通过 Googlesites 可将所有类型的文件包括文档、视频、相片、日历及附件等与好友、团队或整个网络分享。Google AppEngine 是 Google 在 2008 年 4 月发布的一个平台, 使用户可以在 Google 的基础架构上开发和部署运行自己的应用程序。目前, Google AppEngine 支持 Python 语言和 Java 语言, 每个 Google AppEngine 应用程序可以使用达到 500MB 的持久存储空间及可支持每月 500 万综合浏览量的带宽和 CPU。并且, Google AppEngine 应用程序易于构建和维护, 并可根据用户的访问量和数据存储需要的增长轻松扩展。同时, 用户的应用可以和 Google 的应用程序集成, Google AppEngine 还推出了软件开发套件(SDK), 包括可以在用户本地计算机上模拟所有 Google AppEngine 服务的网络服务器应用程序。

IaaS 具有以下几个主要特点:

- IT 资源当做服务传送给客户。
- 基础设施可动态扩展, 即可以根据应用的需求动态增加/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。
- 企业级的基础设施, 使得中小企业可以从聚集的计算资源池中获利。

3. SaaS

SaaS 是一种基于互联网来提供软件服务的应用模式, 其通过浏览器把服务器端的程序软件传给千万用户, 供用户在线使用。SaaS 提供商为用户搭建信息化所需要的所有网络基础设施及软硬件动作平台, 并负责所有前期的实施、后期的维护等一系列服务; 而用户则根据自己的实际需要, 向 SaaS 提供商租赁软件服务, 无须购买软硬件、建设机房、招聘 IT 人员, 即可通过互联网使用信息系统。SaaS 具有以下几方面特点。