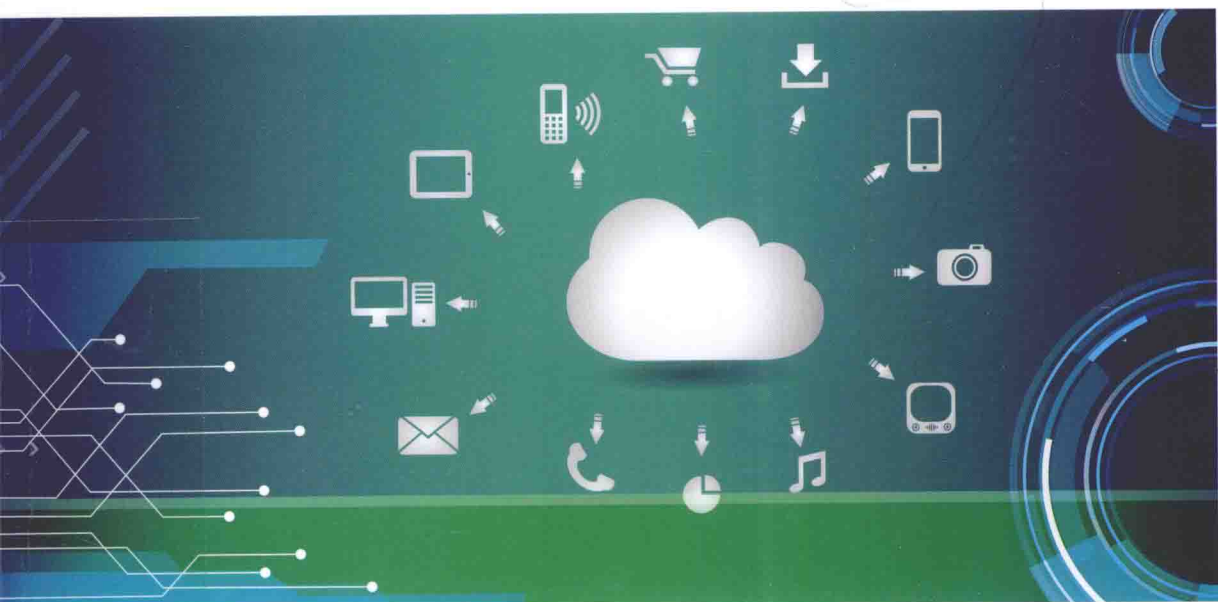


全国高校采用的标准教材 再次大幅升级 透视云计算最新技术

云计算 (第三版)

CLOUD COMPUTING, Third Edition

刘 鹏 主编



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>

云 计 算

(第三版)

刘 鹏 主编

電子工業出版社

Publishing House of Electronics Industry

北京 · BEIJING

内 容 简 介

本书是国内销量最大、被众多高校采用的教材《云计算》的最新升级版，是中国云计算专家咨询委员会秘书长刘鹏教授团队的心血之作。在应对大数据挑战的过程中，云计算技术日趋成熟，拥有大量的成功商业应用。本书追踪最新技术，相比第二版更新了60%以上的内容，包括大数据与云计算、Google 云计算、Amazon 云计算、微软云计算、Hadoop 2.0 及其生态圈、虚拟化技术、OpenStack 开源云计算、云计算数据中心、云计算核心算法和中国云计算技术等。刘鹏教授创办的中国云计算（chinacloud.cn）、中国大数据（thebigdata.cn）网站和刘鹏微信公众号（lpoutlook）为本书学习提供技术支持。

“让学习变得轻松”是本书的初衷。通过本书可掌握云计算的概念和原理，学习主要的云计算平台和技术，还可了解云计算核心算法和发展趋势。本书适合作为相关专业本科和研究生教材，也可作为云计算研发人员和爱好者的学习和参考资料。

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 / 刘鹏主编. —3 版. —北京：电子工业出版社，2015.8
ISBN 978-7-121-26386-6

I. ①云… II. ①刘… III. ①计算机网络 IV. ①TP393

中国版本图书馆 CIP 数据核字（2015）第 137325 号

责任编辑：董亚峰 特约编辑：王 纲

印 刷：北京中新伟业印刷有限公司

装 订：三河市皇庄路通装订厂

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：26.75 字数：688 千字

版 次：2010 年 3 月第 1 版

2015 年 8 月第 3 版

印 次：2015 年 8 月第 1 次印刷

定 价：59.00 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。

编 写 组

主 编：刘 鹏

副主编：陈卫卫

编 委：叶晓江 慈 祥 任桐炜 李志刚 鲍爱华
唐艳琴 付印金 吴海佳 李 涛 余 俊
王 真 张晓燕 沈大为 杨震宇 张海天
宋春博 王 磊

第三版前言

《云计算》第一版于2010年3月出版，第二版于2011年5月出版。时隔四年，在读者的翘首以待中，最受欢迎的云计算教材，终于出第三版啦！

有趣的是，2010年初，我曾经发表了《云计算开启潘多拉星球时代》一文，对2015年的云计算发展进行了预测，我们来看看准不准。

“云计算的影响将是深远的，它将彻底改变IT产业的架构和运行方式。作者在此（2010年1月）做出大胆预测，请广大读者在5年后（2015年1月）回过头来检验这些预测的正确性：在短期之内，高性能计算机、高端服务器、高端存储器、高端处理器的市场的增长率将进入拐点，这些高端硬件市场将被数量众多、低成本、低能耗、高性价比的云计算硬件市场所挤占。紧接着，成本远高于云计算的传统数据中心（IDC），将因其过高的硬件、网络、管理和能耗成本，以及过低的资源利用率，而迅速被云计算数据中心取代，已建的数以万计的数据中心将被迫转换成云计算运行模式。很快地，绝大多数软件将以服务方式呈现，用户通过浏览器访问，数据都存储在‘云’中。甚至连大多数游戏都将在‘云’里运行，用户终端只负责玩家输入和影音输出。在不远的将来，会出现‘泛云计算化’的现象，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、销售管理中心等，将越来越向某些超大型专业运营商集中而获得高的性价比。”

可以说，云计算发展到今天，与预测的结果非常吻合！现在几乎是一切皆云了！

我们唯一能做的就是跟上变化。学习最新的云计算知识，应对大数据挑战，武装自己，迎接未来！

一些同志参加了《云计算》第二版的编写工作，第三版部分地继承了他们的成果，在此记载他们的贡献。他们是：朱军、田浪军、程浩、张洁、张贞、李浩、邓鹏、刘楠、张建平、邓谦、魏家宾、王昊、李松、马少兵、冯颖聪、陈秋晓、傅雷扬等。在此一并致谢！

作者时间充分，但水平有限，欢迎大家不吝赐教。可以通过“刘鹏看未来”（lpoutlook）微信公众号或刘鹏的邮箱 gloud@126.com 与刘鹏取得联系。

刘鹏 教授
2015年7月1日

第二版前言

《云计算》第一版于2010年3月出版。承蒙大家的喜爱，一年中印刷了4次，在当当网云计算书籍中销量保持领先。由于云计算技术发展迅猛，我们的云计算研发团队封闭数月，紧密跟踪，及时推出了第二版。新版《云计算》增加了40%内容，并对原有内容进行全面改写或扩充，以确保能更准确地反映云计算技术的最新面貌。

为了使第二版能够更好地满足大家的需要，本书在改版时先进行了读者调查。调查结果显示大家已经普遍跨越了概念理解阶段，而对云计算的动手实践环节和核心技术原理有着迫切的需求。因此，本书强化了Hadoop、Eucalyptus、CloudSim等动手性强的内容，充实了Google、Amazon、微软云计算原理，增补了VMware虚拟化技术，还同步更新了对云计算理论研究热点的综述。

一些同志参加了《云计算》第一版的编写工作，第二版内容部分地继承了他们的成果。由于编写组署名空间的限制，只好在此记载他们的贡献。他们是：文艾、罗太鹏、龚传、薛志强、朱扬平、王晓璇、王晓盈、鲍爱华、伊英杰、吕良干、周游等。

虽然云计算起步于企业界，但在发展过程中有许多挑战性的技术问题需要解决，希望学术界与企业界密切协作，共同迎接挑战。本着这个思想，我们团队与华为、中兴通讯、360安全卫士、华胜天成、天威视讯、世纪鼎利等知名企业建立了紧密的联合研究关系，研究内容紧跟市场需求和技术发展，研究成果能够迅速转化成生产力。在这本书里，我们将和大家分享其中一些研究成果。

解放军理工大学 刘鹏

2011年5月18日

第一版前言

随着网络带宽的不断增长，通过网络访问非本地的计算服务（包括数据处理、存储和信息服务等）的条件越来越成熟，于是就有了今天我们称作“云计算”的技术。之所以称作“云”，是因为计算设施不在本地而在网络中，用户不需要关心它们所处的具体位置，于是我们就像以前画网络图那样，用“一朵云”来代替了。其实，云计算模式的形成由来已久（Google 公司从诞生之初就采用了这种模式），但只有当宽带网普及到一定程度，且网格计算、虚拟化、SOA 和容错技术等成熟到一定程度并融为一体，又有业界主要大公司的全力推动和吸引人的成功应用案例时，它才如同一颗新星闪亮登场。

既然云计算的服务设施不受用户端的局限，就意味着它们的规模和能力不可限量。Google、亚马逊、微软和 IBM 等的云计算平台已经达到几十万乃至上百万台计算机的规模。由于规模经济性和众多新技术的运用，加之拥有很高的资源利用率，云计算的性能价格比较之传统模式可以达到惊人的 30 倍以上——这使得云计算成为一种划时代的技术。

云计算与当今同样备受关注的 3G 和物联网是什么关系呢？是互为支撑、交相辉映的关系。3G 为云计算带来数以亿计的宽带移动用户。移动终端的计算能力和存储空间有限，却有很强的联网能力，如果有云计算平台的支撑，移动用户将获得前所未有的服务体验；物联网使用数量惊人的传感器、RFID 和视频监控单元等，采集到极其海量的数据，通过 3G 和宽带互联网进行传输，如果汇聚到云计算设施进行存储和处理，则可以更加迅速、准确、智能、低成本地对物理世界进行管理和控制，大幅提高社会生产力水平和生活质量。

云计算的影响将是深远的，它将彻底改变 IT 产业的架构和运行方式。可以预见，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；传统互联网数据中心（IDC）将迅速被成本低一个数量级的云计算数据中心所取代；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高得多的性价比。放眼远眺，云计算将与网格计算融为一体，实现云计算平台之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——作者称之为云格（Gcloud，即 Grid+Cloud）。

因为云计算如此重要，与云计算相关的书籍应运而生。但由于云计算技术起源于企业界而非学术界，各种技术文献很难寻获，目前还未见到对云计算技术进行全面、深入剖析的教科书式出版物。本书编写团队核心成员自 2000 年起就从事网格计算研发，并一

直紧跟国际形势从事云计算领域研发，运营了中国网格 (<http://www.chinagrid.net>) 和中国云计算 (<http://www.chinacloud.cn>) 网站，并承担了知名企业的云计算技术培训工作。我们能够感受到广大读者渴望弄清云计算技术本质和细节的迫切心情，集中力量编写了这本书，希望有所裨益。

本书适合不同层次的读者阅读。根据作者的经验，读一本书，面面俱到的方法不可取——耗时过长、印象不深。建议读者带着自己的疑问，寻找感兴趣的阅读点，直奔主题而去：希望了解云计算的概念、本质和发展趋势的读者，可以重点阅读第 1、11 章；希望学习云计算技术原理的读者，可以将重点放在第 2、3、4、5 章；希望动手从事云计算开发工作的读者，可重点阅读第 6、7、8 章；希望从事云计算理论研究的学术界同仁，可重点阅读第 9、10 章。

此书非常适合作为高校教材使用。建议高校为高年级本科生和研究生开设《云计算》课程。目前解放军理工大学、南京大学等多所高校已经为本科生、研究生开设了《云计算》课程。本课程教学时数建议为 60 学时，其中实验教学占 10~20 学时为宜。建议各位老师在中国云计算网站上共享自己的教案和课件，争取依靠大家的共同努力把它做成精品课程。

感谢中国云计算专家委员会主任委员李德毅院士和林润华秘书长对我们云计算研究工作的指导和鼓励。感谢在我攻读硕、博士学位期间，我的导师谢希仁教授和李三立院士分别在计算机网络和网格计算方向对我的悉心指导。

由于云计算技术较为前沿，加之作者水平有限、时间较紧，书中难免存在谬误，恳请读者批评指正。意见和建议请发到 glood@126.com。欢迎在本书配套网站中国云计算 (<http://www.chinacloud.cn>) 上获取更多资料，并交流与云计算相关的任何问题。我们将密切跟踪云计算技术的发展，吸收您的意见，适时编撰本书的升级版本。

解放军理工大学 刘鹏

2010 年 3 月 1 日

目 录

第 1 章 大数据与云计算	1
1.1 大数据时代	1
1.2 云计算——大数据的计算	3
1.3 云计算发展现状	6
1.4 云计算实现机制	7
1.5 云计算压倒性的成本优势	9
习题	12
参考文献	12
第 2 章 Google 云计算原理与应用	13
2.1 Google 文件系统 GFS	13
2.1.1 系统架构	14
2.1.2 容错机制	16
2.1.3 系统管理技术	17
2.2 分布式数据处理 MapReduce	17
2.2.1 产生背景	17
2.2.2 编程模型	18
2.2.3 实现机制	19
2.2.4 案例分析	20
2.3 分布式锁服务 Chubby	22
2.3.1 Paxos 算法	23
2.3.2 Chubby 系统设计	24
2.3.3 Chubby 中的 Paxos	26
2.3.4 Chubby 文件系统	28
2.3.5 通信协议	29
2.3.6 正确性与性能	31
2.4 分布式结构化数据表 Bigtable	32
2.4.1 设计动机与目标	32
2.4.2 数据模型	33
2.4.3 系统架构	34
2.4.4 主服务器	35
2.4.5 子表服务器	36
2.4.6 性能优化	40

2.5	分布式存储系统 Megastore	41
2.5.1	设计目标及方案选择	41
2.5.2	Megastore 数据模型	42
2.5.3	Megastore 中的事务及并发控制	44
2.5.4	Megastore 基本架构	45
2.5.5	核心技术——复制	47
2.5.6	产品性能及控制措施	51
2.6	大规模分布式系统的监控基础架构 Dapper	52
2.6.1	基本设计目标	52
2.6.2	Dapper 监控系统简介	53
2.6.3	关键性技术	56
2.6.4	常用 Dapper 工具	57
2.6.5	Dapper 使用经验	59
2.7	海量数据的交互式分析工具 Dremel	61
2.7.1	产生背景	61
2.7.2	数据模型	62
2.7.3	嵌套式的列存储	63
2.7.4	查询语言与执行	68
2.7.5	性能分析	70
2.7.6	小结	71
2.8	内存大数据分析系统 PowerDrill	71
2.8.1	产生背景与设计目标	71
2.8.2	基本数据结构	72
2.8.3	性能优化	73
2.8.4	性能分析与对比	75
2.9	Google 应用程序引擎	76
2.9.1	Google App Engine 简介	77
2.9.2	应用程序环境	78
2.9.3	Google App Engine 服务	80
	习题	88
	参考文献	88
第 3 章 Amazon 云计算 AWS		90
3.1	基础存储架构 Dynamo	90
3.1.1	Dynamo 概况	90
3.1.2	Dynamo 架构的主要技术	91
3.2	弹性计算云 EC2	99
3.2.1	EC2 的基本架构	99

3.2.2	EC2 的关键技术	100
3.2.3	EC2 的安全及容错机制	102
3.3	简单存储服务 S3	103
3.3.1	S3 的基本概念和操作	103
3.3.2	S3 的数据一致性模型	105
3.3.3	S3 的安全措施	105
3.4	非关系型数据库服务 SimpleDB 和 DynamoDB	108
3.4.1	非关系型数据库与传统关系数据库的比较	108
3.4.2	SimpleDB	108
3.4.3	DynamoDB	110
3.4.4	SimpleDB 和 DynamoDB 的比较	111
3.5	关系数据库服务 RDS	111
3.5.1	RDS 的基本原理	111
3.5.2	RDS 的使用	112
3.6	简单队列服务 SQS	113
3.6.1	SQS 的基本模型	113
3.6.2	SQS 的消息	114
3.7	内容推送服务 CloudFront	115
3.7.1	CDN	116
3.7.2	CloudFront	117
3.8	其他 Amazon 云计算服务	119
3.8.1	快速应用部署 Elastic Beanstalk 和服务模板 CloudFormation	119
3.8.2	DNS 服务 Router 53	120
3.8.3	虚拟私有云 VPC	120
3.8.4	简单通知服务和简单邮件服务	120
3.8.5	弹性 MapReduce 服务	121
3.8.6	电子商务服务 DevPay、FPS 和 Simple Pay	122
3.8.7	Amazon 执行网络服务	127
3.8.8	土耳其机器人	128
3.8.9	数据仓库服务 Redshift	129
3.8.10	应用流服务 AppStream 和数据流分析服务 Kinesis	129
3.9	AWS 应用实例	130
3.9.1	照片和视频共享网站 SmugMug	130
3.9.2	视频制作网站 Animoto	131
3.9.3	网站排名 Alexa	132
3.10	小结	132
	习题	132
	参考文献	133

第4章 微软云计算 Windows Azure	136
4.1 微软云计算平台	136
4.2 微软云操作系统 Windows Azure	137
4.2.1 Windows Azure 概述	137
4.2.2 Windows Azure 计算服务	138
4.2.3 Windows Azure 存储服务	140
4.2.4 Windows Azure Connect	147
4.2.5 Windows Azure CDN	148
4.2.6 Fabric 控制器	149
4.3 微软云关系数据库 SQL Azure	150
4.3.1 SQL Azure 概述	150
4.3.2 SQL Azure 关键技术	151
4.3.3 SQL Azure 和 SQL Server 对比	153
4.4 Windows Azure AppFabric	154
4.4.1 AppFabric 概述	154
4.4.2 AppFabric 关键技术	155
4.5 Windows Azure Marketplace	159
4.6 Windows Azure 服务平台	160
4.6.1 网站	160
4.6.2 虚拟机	161
4.6.3 云服务	162
4.6.4 移动服务	163
4.6.5 大数据处理	164
4.6.6 媒体支持	165
习题	166
参考文献	166
第5章 Hadoop 2.0: 主流开源云架构	168
5.1 引例	168
5.1.1 问题概述	168
5.1.2 常规解决方案	168
5.1.3 分布式下的解决方案	169
5.1.4 小结	173
5.2 Hadoop 2.0 简述	173
5.2.1 Hadoop 2.0 由来	173
5.2.2 Hadoop 2.0 相关项目	174
5.2.3 Hadoop 应用	175
5.3 Hadoop 2.0 部署	175

5.3.1	部署综述	175
5.3.2	传统解压包部署	179
5.4	Hadoop 2.0 体系架构	182
5.4.1	Hadoop 2.0 公共组件 Common	182
5.4.2	分布式文件系统 HDFS	184
5.4.3	分布式操作系统 Yarn	190
5.4.4	Hadoop 2.0 安全机制简介	198
5.5	Hadoop 2.0 访问接口	199
5.5.1	访问接口综述	199
5.5.2	浏览器接口	199
5.5.3	命令行接口	200
5.6	Hadoop 2.0 编程接口	202
5.6.1	HDFS 编程	202
5.6.2	Yarn 编程	205
	习题	214
	参考文献	214
第 6 章	Hadoop 2.0 大家族	215
6.1	Hadoop 2.0 大家族概述	215
6.1.1	分布式组件	215
6.1.2	部署概述	217
6.2	ZooKeeper	223
6.2.1	ZooKeeper 简介	223
6.2.2	ZooKeeper 入门	224
6.3	Hbase	226
6.3.1	Hbase 简介	226
6.3.2	Hbase 入门	229
6.4	Pig	232
6.4.1	Pig 简介	232
6.4.2	Pig 入门	233
6.5	Hive	234
6.5.1	Hive 简介	234
6.5.2	Hive 入门	235
6.6	Oozie	238
6.6.1	Oozie 简介	238
6.6.2	Oozie 入门	238
6.7	Flume	241
6.7.1	Flume 简介	241

6.7.2	Flume 入门	243
6.8	Mahout	246
6.8.1	Mahout 简介	246
6.8.2	Mahout 入门	246
6.9	小结	247
	习题	248
	参考文献	248
第 7 章	虚拟化技术	250
7.1	虚拟化技术简介	250
7.2	服务器虚拟化	251
7.2.1	服务器虚拟化的层次	251
7.2.2	服务器虚拟化的底层实现	252
7.2.3	虚拟机迁移	253
7.2.4	隔离技术	256
7.2.5	案例分析	257
7.3	存储虚拟化	258
7.3.1	存储虚拟化的一般模型	258
7.3.2	存储虚拟化的实现方式	259
7.3.3	案例分析	260
7.4	网络虚拟化	262
7.4.1	核心层网络虚拟化	262
7.4.2	接入层网络虚拟化	262
7.4.3	虚拟机网络虚拟化	262
7.4.4	案例分析: VMware 的网络虚拟化技术	263
7.5	桌面虚拟化	265
7.5.1	桌面虚拟化简介	266
7.5.2	技术现状	266
7.5.3	案例分析	267
	习题	267
	参考文献	268
第 8 章	OpenStack 开源虚拟化平台	269
8.1	OpenStack 背景介绍	269
8.1.1	OpenStack 是什么	269
8.1.2	OpenStack 的主要服务	270
8.2	计算服务 Nova	272
8.2.1	Nova 组件介绍	272

8.2.2	Libvirt 简介	273
8.2.3	Nova 中的 RabbitMQ 解析	275
8.3	对象存储服务 Swift	282
8.3.1	Swift 特性	282
8.3.2	应用场景	283
8.3.3	Swift 主要组件	284
8.3.4	Swift 基本原理	286
8.3.5	实例分析	290
8.4	镜像服务 Glance	291
8.4.1	Glance 的作用	291
8.4.2	Glance 的组成部分	291
	习题	292
	参考文献	292
第 9 章	云计算数据中心	293
9.1	云数据中心的特征	293
9.2	云数据中心网络部署	294
9.2.1	改进型树结构	294
9.2.2	递归层次结构	296
9.2.3	光交换网络	299
9.2.4	无线数据中心网络	301
9.2.5	软件定义网络	301
9.3	绿色节能技术	305
9.3.1	配电系统节能技术	305
9.3.2	空调系统节能技术	307
9.3.3	集装箱数据中心节能技术	310
9.3.4	数据中心节能策略和算法研究	312
9.3.5	新能源的应用	315
9.3.6	典型的绿色节能数据中心	318
9.4	自动化管理	320
9.4.1	自动化管理的特征	320
9.4.2	自动化管理实现阶段	321
9.4.3	Facebook 自动化管理	322
9.5	容灾备份	324
9.5.1	容灾系统的等级标准	325
9.5.2	容灾备份的关键技术	325
9.5.3	云存储在容灾备份中的应用	327
	习题	328

参考文献	328
第 10 章 云计算核心算法	332
10.1 Paxos 算法	332
10.1.1 Paxos 算法背景知识	332
10.1.2 Paxos 算法详解	333
10.1.3 Paxos 算法举例	336
10.2 DHT 算法	338
10.2.1 DHT 原理介绍	339
10.2.2 Chord 中 DHT 的具体实现	340
10.2.3 Pastry 中 DHT 的具体实现	342
10.2.4 CAN 中 DHT 的具体实现	343
10.2.5 Tapestry 中 DHT 的具体实现	345
10.3 Gossip 协议	345
10.3.1 Gossip 协议的特点	346
10.3.2 Gossip 协议的通信方式及收敛性	346
10.3.3 Gossip 节点管理算法	348
10.3.4 Cassandra 中 Gossip 协议的具体实现方式	349
10.3.5 CoolStreaming 系统中 Gossip 协议的具体实现方式	352
习题	354
参考文献	354
第 11 章 中国云计算技术	355
11.1 国内云计算发展概况	355
11.2 国产云存储技术	356
11.2.1 淘宝分布式文件系统 TFS	356
11.2.2 云创存储 cStor 分布式文件系统	359
11.2.3 A8000 超低功耗云存储一体机	364
11.3 国产大数据库技术	367
11.3.1 阿里巴巴 OceanBase	367
11.3.2 云创存储数据立方 (DataCube)	371
11.4 云视频监控技术	376
11.4.1 cVideo 云视频监控系统	376
11.4.2 cVideo 智能分析系统	378
11.4.3 cVideo 云转码系统	381
11.5 阿里巴巴阿里云服务	383
11.5.1 阿里云计算体系架构	383
11.5.2 弹性计算服务 (ECS)	384

11.5.3	开放存储服务 (OSS)	385
11.5.4	开放结构化数据服务 (OTS)	385
11.5.5	开放数据处理服务 (ODPS)	385
11.5.6	关系型数据库 (RDS)	386
11.6	云创存储万物云服务	386
11.6.1	平台简介	386
11.6.2	系统架构	387
11.6.3	功能服务	388
11.6.4	应用举例	391
	习题	392
	参考文献	392
第 12 章	总结与展望	394
12.1	主流商业云计算解决方案比较	394
12.1.1	应用场景	394
12.1.2	使用流程	396
12.1.3	体系结构	397
12.1.4	实现技术	397
12.1.5	核心业务	398
12.2	主流开源云计算系统比较	400
12.2.1	开发目的	401
12.2.2	体系结构	402
12.2.3	实现技术	403
12.2.4	核心服务	403
12.3	云计算的历史坐标与发展方向	404
12.3.1	互联网发展的阶段划分	404
12.3.2	云格 (Gloud) ——云计算的未来	407
	习题	409
	参考文献	409