中国计算机学会学术著作丛书

对等网络：
结构、应用与设计

Peer-to-Peer Network

Structure, Application and Design

陈贵海　李振华　著

倪明选　刘云浩　校

清华大学出版社

北　京

内 容 简 介

对等网络是分布式系统和计算机网络相结合的产物，它在应用领域和学术界获得了广泛的重视和成功，被称为“改变 Internet 的新一代网络技术。”本书由浅入深，全面系统地介绍了对等网络的各个方面内容，重点在于对三代 P2P 网络、P2P 网络结构和设计机制的讲解。首先介绍对等网络的概念、历史、特点，重点介绍三代 P2P 网络，同时讲述各种 P2P 应用系统和软件；然后深入讨论 P2P 网络设计所要考虑的核心机制、优化网络性能的增强机制和 P2P 模拟器的设计；最后分析了 P2P 的现状和发展趋势。

本书是有关对等网络的学术专著，包含对等网络产生以来在学术界、商业领域有价值、有影响力的成果，内容编写力求前沿、具体、实用。本书既可作为大学计算机、电子信息等专业教材，也可作为研究所、商业公司中对等网络研究者的重要参考书。

图书在版编目(CIP)数据

责任编辑：
责任校对：
责任印制：

出版发行：清华大学出版社　　**地　　址：**北京清华大学学研大厦 A 座
http://www.tup.com.cn　　**邮　　编：**100084
c-service@tup.tsinghua.edu.cn
社 总 机：010-62770175　　**邮购热线：**010-62786544
投稿咨询：010-62772015　　**客户服务：**010-62776969
印 刷 者：
装 订 者：
经　　销：全国新华书店
开　　本：　　**印张：**　　**字数：**　千字
版　　次：2007 年　月第 2 版　　**印次：**2007 年　月第1次印刷
印　　数：1～　000
定　　价：　.00 元

本书如存在文字不清、漏印、缺页、倒页、脱页等印装质量问题，请与清华大学出版社出版部联系调换。
联系电话：010-62770177 转 3103　产品编号：

清华大学出版社计算机学术著作出版基金

评审委员会

中国计算机学会学术著作丛书

名誉主任委员：张效祥
主　任　委　员：唐泽圣
副主任委员：陆汝钤
委　　员：（以姓氏笔画为序）
王　珊　李晓明　吕　建
林惠民　罗军舟　郑纬民
施伯乐　焦金生　谭铁牛

序

Preface

第一台电子计算机诞生于20世纪40年代，到目前为止，计算机的发展已远远超出了其创始者的想象。计算机的处理能力越来越强，应用面越来越广，应用领域也从单纯的科学计算渗透到社会生活的方方面面：从工业、国防、医疗、教育、娱乐直至人们的日常生活，计算机的影响可谓无处不在。

计算机之所以能取得上述地位并成为全球最具活力的产业，原因在于其高速的计算能力、庞大的存储能力以及友好灵活的用户界面。而这些新技术及其应用有赖于研究人员多年不懈的努力。学术研究是应用研究的基础，也是技术发展的动力。

自1992年起，清华大学出版社与广西科学技术出版社为促进我国计算机科学技术与产业的发展，推动计算机科技著作的出版，设立了“计算机学术著作出版基金”，并将资助出版的著作列为中国计算机学会的学术著作丛书。时至今日，本套丛书已出版学术专著近50种，产生了很好的社会影响，有的专著具有很高的学术水平，有的则奠定了一类学术研究的基础。中国计算机学会一直将学术著作的出版作为学会的一项主要工作。本届理事会将秉承这一传统，继续大力支持本套丛书的

出版，鼓励科技工作者写出更多的优秀学术著作，多出好书，多出精品，为提高我国的知识创新和技术创新能力，促进计算机科学技术的发展和进步做出更大的贡献。

中国计算机学会

2002 年 6 月 14 日

前 言

Foreword

对等网络(Peer-to-Peer Network,简称 P2P 网络)是分布式系统和计算机网络相结合的产物,它在网络协议的应用层,打破过去的“客户/服务器”模式,让所有网络成员享有“自由、平等、互联”的功能,不再有客户、服务器之分,任何两个网络结点之间都能共享文件、传递消息。对等网络起源于 1999 年风行一时的音乐文件共享软件 Napster,随后则是一系列人们耳熟能详的网络软件:Gnutella*,KaZaA,BitTorrent,eDonkey/eMule,Skype,等等。虽然从产生到现在只有短短几年的历史,但是对等网络在应用领域和学术界获得了广泛的重视和成功,并占据了当前 Internet 超过一半的带宽资源,被称为“改变 Internet 的新一代网络技术”。

在应用领域,1999 年出现的第一个 P2P 网络 Napster,在美国 6 个月内即拥有 5000 万注册用户,成为网络时代的一个奇迹。此后许多计算机公司开始投资 P2P 网络的开发,如著名的无结构 P2P 网络 Gnutella、KaZaA 和 eDonkey/eMule。2002 年,家喻户晓的 BitTorrent 网络出现,其简称“BT”已成为“自由下载、文件

* 注意 Gnutella 的发音为[nju:ˈtelə],同 new-tella,字母 G 不发音。

共享”的代名词。BT 在中国的应用尤为广泛，2006 年发布的中国互联网统计报告显示：中国 1.11 亿网民中有 27.8%的人使用过 BT 软件（超过 3000 万人），其流行程度由此可见一斑。除了文件共享，P2P 软件在其他多个应用领域也带来了冲击性的变革：网络语音电话软件 Skype 直接威胁到电话公司的利益，网络电视软件如 PPLive、TvAnts 等使得人们能通过 Internet 流畅地观看电视直播，诸如此类的应用还有很多。

在学术界，虽然 P2P 的思想起源很早并一直有人提及，但直到 2001 年“P2P 网络”这一概念才获得人们的共识并成为研究热点。计算机领域许多重要会议（如 SIGCOMM，SPAA，PODC，ICNP，INFOCOM，ICDCS 等）和刊物（如 IEEE/ACM Transactions on Networking，IEEE Transactions on Parallel and Distributed Systems，IEEE Journal on Selected Areas in Communications 等）从 2001 年开始发布和刊登 P2P 领域的论文，同年出现了 P2P 领域最著名的结构化网络模型：Chord，CAN，Tapestry 和 Pastry。P2P 领域自身的两个专业会议 IEEE P2P 和 IPTPS 分别于 2001 年、2002 年举行。人们提出了许多新颖、独特的 P2P 网络模型，开发了许多用以实现、完善 P2P 网络的核心机制及增强机制，其中不少理论成果已经转化为 Internet 上的实用软件。

到目前为止，P2P 还是一项比较新的技术，其中还存在很多问题，同时也存在着各种挑战和机遇。虽然有太多的问题和障碍，但 P2P 从无到有，从一个民间小软件发展到改变 Internet 面貌、改变人们交流方式的一项新技术，这一过程是值得人们回味的，其中蕴藏的力量和意义，值得我们付出勇气、热情、耐心去努力研究并发掘。

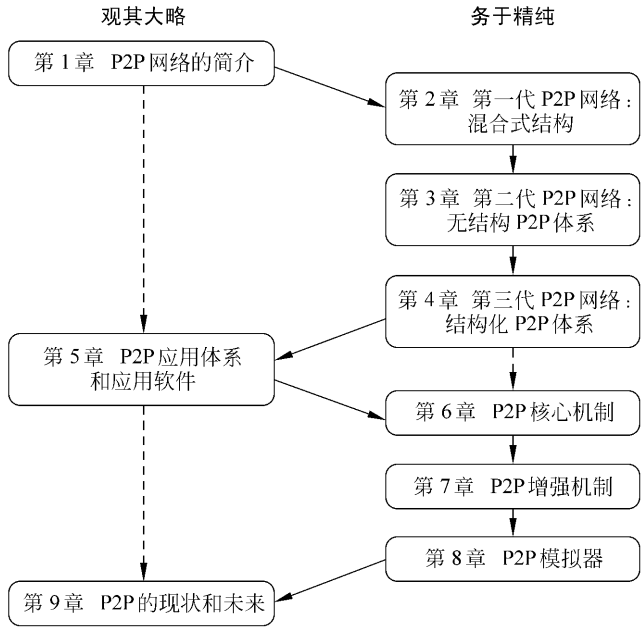
阅读本书的方法

追比圣贤，原是读书人的本意。《三国演义》用这样一段话来描述诸葛亮读书：“孔明与博陵崔州平、颍川石广元、汝南孟公威与徐元直四人为密友。此四人务于精纯，惟孔明独观其大略。尝抱膝长吟，而指四人曰：‘公等仕进可至刺史、郡守。’众问孔明之志若何，孔明但笑而不答。每常自比管仲、乐毅，其才不可量也。”

我们以为这段话，正说出了读书的要义。做学问不必执着于一事一物，对于大多数事情，只要能“观其大略”，明白是怎么回事就足够了；生命有限，青春短暂，我们要做的，是关注那些真正对自己有价值的东西，将精力集中于真正值得去研究的问题，努力“务于精纯”。

这也正是我们希望的读者阅读本书的方法。我们真心希望您付出时间之

后能有所得，所以建议读者按照如下线索来阅读本书：



给读者的鼓励

第一次读这本书时你或许不能完全理解，但不用担心，实际上作者在撰写本书时很多相关论文第一次也没看懂，比如 Freenet 的论文就至少看了 5 遍！读者完全可以跳过暂时看不懂的内容，因为我们在写作时已努力使各章各节尽量独立，当你往后看得更多时，前面感到困惑的内容很可能就豁然开朗了。

“事理因人言而悟者，有悟还有迷，总不如自悟之了了。”在阅读本书的时候，千万不要忘记您自己的批判和思考。无论是批评之言，还是建议之语，希不吝赐教。我们的信箱是：gchen@nju.edu.cn 或 lizhenhua@dislab.nju.edu.cn。

或许世上只有两种人能成为作家：痛苦的或快乐的，因为只有不一样的心情才能承受写作的漫漫孤寂。在作者心中，只能靠想象去构思他的读者空间，所以我们希望能看到、听到您任何形式的反馈，感受到您阅读时的欢欣、困惑与体会。只有这样，我们所构思的空间才能变得真实、饱满，让我们觉得写作时所有的单调、枯燥与艰辛，都是值得付出的。

愿本书带给您的，不仅仅是技术和知识！

参考文献来源

本书所引用的参考文献绝大多数来自于网络通信领域的重要国际会议和刊物，文献发表时间集中于2001—2007年之间，以保证内容的权威性和时效性；除此之外，我们也参考了大量Web资源，尤其是P2P相关网站上的资料。我们所关注的会议、刊物列举如下：

缩写名称	英文全称	中文名称
SIGCOMM	ACM Annual Conference of the Special Interest Group on Data Communication	ACM数据通信特别兴趣组年会
SPAA	ACM Symposium on Parallel Algorithms and Architectures	ACM并行算法和体系讨论会
PODC	ACM Symposium on Principles of Distributed Computing	ACM分布式计算原理讨论会
ICNP	International Conference on Network Protocols	国际网络协议会议
INFOCOM	IEEE Conference on Computer Communications	IEEE计算机通信会议
NOSSDAV	ACM International Workshop on Network and Operating System Support for Digital Audio and Video	ACM数字音频和视频的网络和操作系统支持讨论会
SOSP	ACM Symposium on Operating Systems Principles	ACM操作系统原理讨论会
OSDI	USENIX Symposium on Operating Systems Design and Implementation	USENIX操作系统设计和实现讨论会
USITS	USENIX Symposium on Internet Technologies and Systems	USENIX因特网技术和系统讨论会
ASPLOS	International Conference on Architectural Support for Programming Languages and Operating Systems	国际程序设计语言和操作系统的体系支持会议
ICDCS	International Conference on Distributed Computing Systems	国际分布式计算系统会议
ICPP	International Conference on Parallel Processing	国际并行处理会议
IPDPS	International Parallel and Distributed Processing Symposium	国际并行和分布式处理讨论会
IPTPS	International Workshop on Peer-to-Peer Systems	国际P2P系统讨论会
IEEE P2P	IEEE International Conference on Peer-to-Peer Computing	IEEE国际P2P计算会议
刊物名一般不缩写	IEEE/ACM Transactions on Networking	IEEE/ACM网络学报
	IEEE Transactions on Parallel and Distributed Systems	IEEE并行和分布式系统学报
	IEEE Journal on Selected Areas in Communications	IEEE通信精选领域杂志

致谢

本书的写作、出版得到两大项目的支持：一是 2005 年国家自然科学基金项目“实用化对等网络技术的研究”；二是 2005 年江苏省自然科学基金前期预研项目“新型 P2P 计算技术的基础研究”。香港科技大学倪明选教授和刘云浩教授对本书的写作给予了有益的指导，华中科技大学金海教授和国防科技大学肖侬教授通读了本书的初稿，并提出了许多有价值的建议。邱彤庆、谢军峰、袁瑞峰、叶懋、刘之育、陈欢、于南南等同学对部分章节提出了修改建议。在此谨表示衷心的感谢！

陈贵海　李振华

2007 年 3 月

于南京大学鼓楼校区

目 录

Contents