



国际信息工程先进技术译丛

P2P系统及其应用

Peer-to-Peer Systems
and Applications

(德) Ralf Steinmetz 等著
Klaus Wehrle

王玲芳 陈焱 译



机械工业出版社
CHINA MACHINE PRESS

TP393.4/207

2008

国际信息工程先进技术译丛

P2P 系统及其应用

(德) **Ralf Steinmetz** 等著
Klaus Wehrle

王玲芳 陈 焱 译

机械工业出版社

本书首先介绍了有关 P2P 的概念、领域、历史和未来,然后逐步展开,详细阐述了 P2P 系统结构、基于 P2P 的应用、P2P 的自组织特征等人们普遍关注的不同热点技术。本书内容涵盖广泛,包括 P2P 概念、领域、历史和未来、无结构的 P2P 系统、有结构的 P2P 系统、基于 P2P 的应用、自组织特征、查找和搜索、P2P 流量和性能评估、移动和泛在环境中的 P2P、商业应用和市场以及其他问题等。本书以实际系统为根据,采用丰富的图表和案例,详细解释了 P2P 的各个方面,为读者描绘了一个关于 P2P 的清晰图景。

本书适合于网络系统和应用工程技术人员、大专院校学生以及对 P2P 技术感兴趣的商业管理人员阅读或参考,也可为相关领域的研究人员了解国外该领域的最新进展提供参考。本书还可作为高等院校、培训基地等开展 P2P 系统和应用方面的教材使用。

Translation from the English language edition;

Peer-to-Peer Systems and Applications, edited by Ralf Steinmetz and Klaus Wehrle, Copyright © 2005 Springer-Verlag Berlin Heidelberg. Springer is a part of Springer Science + Business Media.

All Rights Reserved.

本书中文简体字版由德国 Springer 出版社授权机械工业出版社独家出版。版权所有,侵权必究。

本书版权登记号:图字:01-2007-2074

图书在版编目(CIP)数据

P2P 系统及其应用/(德)斯太门兹(Steinmetz,R.)
等著. 王玲芳,陈焱译. —北京:机械工业出版社,
2008.5

(国际信息工程先进技术译丛)

ISBN 978-7-111-23647-4

I. P… II. ①斯…②王…③陈… III. 因特网—基
本知识 IV. TP393.4

中国版本图书馆 CIP 数据核字(2008)第 032007 号

机械工业出版社(北京市百万庄大街 22 号 邮政编码 100037)

策划编辑:张俊红 责任编辑:顾 谦 责任校对:陈延翔

封面设计:马精明 责任印制:洪汉军

北京铭成印刷有限公司印刷

2008 年 5 月第 1 版第 1 次印刷

169mm×239mm·17.125 印张·646 千字

0001—4000 册

标准书号:ISBN 978-7-111-23647-4

定价:78.00 元

凡购本书,如有缺页、倒页、脱页,由本社发行部调换

销售服务热线电话:(010)68326294

购书热线电话:(010)88379639 88379641 88379643

编辑热线电话:(010)88379768

封面无防伪标均为盗版

译者序

在互联网领域,对等端到对等端(Peer to Peer,以下简称 P2P)模式是作为客户服务器模式的对立面出现的,其成功的系统典范是 Napster 和 Gnutella。P2P 广泛地用来描述各种网络系统,通常运行于表示层、会话层和应用层,虽然自组织网络和其他系统从物理层起开始就使用相同的概念。特别地,P2P 已经扩展到了内容分发、协作、缓冲、商务过程自动化、供应链管理、网格计算、分布式计算、商务往来、数据部署、用户间通信、用户社团、自组织网络以及 Internet 自身。也许 P2P 应该看作是一种架构性的方法,而不是一种特定技术或商务方法。

目前 P2P 网络系统及应用迅速地在 Internet 中蔓延,产生了各方面积极的和消极的影响,对电信运营商网络规划、运营以及对终端用户的行为、体验等的影响都是不可忽视的。无论是互联网、电信网还是广播电视网的行业人员如何看待这个潘多拉魔盒,但都普遍认为这是未来业务发展的一种崭新模式。

在此之前,国际上 P2P 领域的知识是分散的、支离破碎的。2004 年底,在我们开始申请 CNGI(中国下一代互联网)视频多媒体点播系统(CNGI-04-15-2A)项目时,就苦于找不到关于 P2P 比较全面的专著。国内更是缺乏这方面系统介绍的书籍,工程技术/研究人员迫切需要有一本结合基本理论知识、实践、商务等方面的 P2P 书籍。2006 年当我们看到 Ralf Steinmetz、Klaus Wehrle 等编写的《Peer-to-Peer Systems and Applications》一书时,有感于本书的经典,遂决定将之翻译成中文,奉献给国内在 P2P 领域耕耘的各界人士。

本书首先介绍了有关 P2P 概念、领域、历史和未来,然后逐步展开,详细阐述了 P2P 系统结构、基于 P2P 的应用、P2P 的自组织特征等人们普遍关注的不同热点技术。本书内容涵盖广泛,包括 P2P 概念、领域、历史和未来、无结构的 P2P 系统、有结构的 P2P 系统、基于 P2P 的应用、自组织特征、查找和搜索、P2P 流量和性能评估、移动和泛在环境中的 P2P、商业应用和市场以及其他问题。本书内容全面而翔实,是不可多得的必备参考书。

本书译者长期从事网络技术方面的科研工作,具有丰富的理论基础和实践经验。本书由王玲芳负责全书统稿和校对工作,并负责第 1~10、21~33 章的翻译,河南省信阳职业技术学院的陈焱负责第 11~20 章的翻译工作。本书在翻译过程中得到了国家网络新媒体技术工程研究中心王劲林、倪宏两位主任的大力支持,同时中心成立了由博士生、硕士生组成的 P2P 技术研究讨论小组。本书翻译中的问题,曾经经过热烈讨论,在此对中心的工作人员邓浩江、赵志强、张

武、孙鹏、张辉、陈君，以及讨论组的李挺屹、齐卫宁、武广柱、尤佳莉、单明辉、李虹、游庆珍、王弟英、吴璟、李山坡、潘东升、李冬梅、史岩、卫铮、刘磊、稽智辉、杨群、吴秋义、段世惠、宋磊、匡振国等表示感谢，他们都对本书的具体翻译工作做出了贡献。另外，机械工业出版社的张俊红在本书的出版方面给予了大力支持，特此感谢。尤其不能忘记的是，家人在翻译过程的无私奉献，是他们的帮助促成了本书的翻译。

不过，需要指出的是，本书的内容是译者在尽量忠实于原著的基础上翻译的，书中所述并不代表译者及其所在单位的观点和见解，这点请读者注意。另外，由于P2P技术的日新月异，翻译时间比较仓促，疏漏错误之处在所难免，敬请读者原谅和指正。

译者

2008年3月

前 言

Ion Stoica(美国加州大学伯克利分校)

从 Napster 和 Gnutella 开始, P2P 系统就已成为 Internet 吸引数百万用户的不可分割部分。根据几个大型 ISP(Internet 服务提供商)的最近测量显示, P2P 流量已经超过了 Web 流量, 而 Web 流量曾经是 Internet 上的主要流量。当最流行的 P2P 应用继续是文件共享和内容分发时, 新的应用, 如 Internet 电话开始出现。

并不令人惊奇的是, P2P 系统的流行已经掀起了学术研究的热潮。在一个非常短的时间内, P2P 已经演化成为了一个令人激动的研究领域, 它将来自系统、网络 and 理论方面的研究人员联合了起来。在过去 5 年中, P2P 研究工作成果已经出现在几乎所有顶级系统和联网会议的论文集中。

但是, 虽然 P2P 系统的巨大流行性和 P2P 的研究已经产生了知识的爆炸, 但这个知识整体还是没有结构的。关于 P2P 系统的综述和提供关于 P2P 技术的全面论述的书籍还是很少的, 且内容零散。P2P 仍然是一个快速演进领域的事实使这种资料的相对缺乏变得更加急迫。

本书通过包括有代表性文章的集合填补了这个空白, 给出了最新的和全面的 P2P 领域的现状。涵盖如此广泛和相对新的领域, 任何一本这种书籍所要面对的主要挑战之一就是如何架构材料。本书通过将材料分成粗略的三部分解决了这个难题。

本书的第一部分涵盖了 P2P 设计的基础知识、有结构的和无结构的系统基本知识, 并提供了很多应用, 包括 E-mail、组播、网格计算和 Web 服务。之后, 本书又避免了仅描述传统系统的思路, 讨论了 P2P 系统的一般特征, 即 P2P 系统的组织本质, 且讨论了评估这些系统的最重要专题。另外, 本书通过讨论在两个计算机科学领域中 P2P 技术的影响, 描绘出 P2P 的广泛应用性, 这两个领域是搜索和信息检索及移动计算。如果不讨论商业模型、记账和安全, P2P 书籍将是不完全的, 本书在最后一部分涉及到了这些专题。

对于本书, Steinmetz 和 Wehrle 为提供 P2P 领域的海量知识作出了成功的尝试, 这些知识以一种相关的和有结构的形式呈现, 是过去几年收集的。本书包括了本领域中关于最新发展的文章, 这使本书无论对于想了解本领域最新观点的读者, 而且对于想进入本领域的研究人员都是有用的。传统 P2P 设计和应用的结合、其自组织属性的讨论和它们对计算机科学其他领域的影响, 这三方面使本书成为 P2P 领域的有价值的新资料。

Ion Stoica 于伯克利

原书作者名单^①

Ion Stoica
645 Soda Hall
Computer Science Division
University of California, Berkeley
Berkeley, CA 94720-1776
USA

Ralf Steinmetz
TU Darmstadt
KOM – Multimedia Communications
Merckstraße 25
64283 Darmstadt
Germany

Rüdiger Schollmeier
TU München
Institute of Communication Networks
Arcisstraße 21
80290 München
Germany

Kai Fischbach
Universität zu Köln
Seminar für Wirtschaftsinformatik,
insb. Informationsmanagement
Pohligstr. 1
50969 Köln
Germany

Vasilios Darlagiannis
TU Darmstadt
KOM – Multimedia Communications
Merckstraße 25
64283 Darmstadt
Germany

Klaus Wehrle
Universität Tübingen
Protocol-Engineering &
Distributed Systems Group
Morgenstelle 10c
72076 Tübingen
Germany

Jörg Eberspächer
TU München
Institute of Communication Networks
Arcisstraße 21
80290 München
Germany

Detlef Schoder
Universität zu Köln
Seminar für Wirtschaftsinformatik,
insb. Informationsmanagement
Pohligstr. 1
50969 Köln
Germany

Christian Schmitt
Universität zu Köln
Seminar für Wirtschaftsinformatik,
insb. Informationsmanagement
Pohligstr. 1
50969 Köln
Germany

Katharina Anna Lehmann
Universität Tübingen
Arbeitsbereich für Paralleles Rechnen
WSI – Am Sand 13
72076 Tübingen
Germany

① 此名单为本书英文原版书作者名单，排名次序与原书次序一致。——译者注

Michael Kaufmann
 Universität Tübingen
 Arbeitsbereich für Paralleles Rechnen
 WSI – Am Sand 13
 72076 Tübingen
 Germany

Stefan Götz
 Universität Tübingen
 Protocol-Engineering &
 Distributed Systems Group
 Morgenstelle 10c
 72076 Tübingen
 Germany

Karl Aberer
 School of Computer and
 Communication Sciences
 Ecole Polytechnique Fédérale
 de Lausanne (EPFL)
 1015 Lausanne
 Switzerland

Manfred Hauswirth
 School of Computer and
 Communication Sciences
 Ecole Polytechnique Fédérale
 de Lausanne (EPFL)
 1015 Lausanne
 Switzerland

Kostas Katrinis
 ETH Zürich, TIK
 Gloriastrasse 35
 8092 Zürich
 Switzerland

Andreas Hæberlen
 Rice University & MPI-SWS
 Distributed Systems Group
 3007 Duncan Hall, 6100 Main St.
 Houston TX 77005
 USA

Peter Druschel
 Rice University & MPI-SWS
 Distributed Systems Group
 3007 Duncan Hall, 6100 Main St.
 Houston TX 77005
 USA

Simon Rieche
 Universität Tübingen
 Protocol-Engineering &
 Distributed Systems Group
 Morgenstelle 10c
 72076 Tübingen
 Germany

Heiko Niedermayer
 Universität Tübingen
 Computer Networks & Internet
 Morgenstelle 10c
 72076 Tübingen
 Germany

Anwitaman Datta
 School of Computer and
 Communication Sciences
 Ecole Polytechnique Fédérale
 de Lausanne (EPFL)
 1015 Lausanne
 Switzerland

Martin May
 ETH Zürich, TIK
 Gloriastrasse 35
 8092 Zürich
 Switzerland

Alan Mislove
 Rice University & MPI-SWS
 Distributed Systems Group
 3007 Duncan Hall, 6100 Main St.
 Houston TX 77005
 USA

Ansley Post
 Rice University & MPI-SWS
 Distributed Systems Group
 3007 Duncan Hall, 6100 Main St.
 Houston TX 77005
 USA

Andreas Mauthe
 Lancaster University
 Computing Department
 Lancaster, LA1 4YR
 UK

Oliver Heckmann
 TU Darmstadt
 KOM – Multimedia Communications
 Merckstraße 25
 64283 Darmstadt
 Germany

Paul Müller
 TU Kaiserslautern
 AG ICSY
 Gottlieb-Daimler-Straße
 67663 Kaiserslautern
 Germany

Christian Koppen
 Universität Passau
 Computer Networks & Computer
 Communications Group
 Innstraße 33
 94032 Passau
 Germany

Jan Mischke
 McKinsey Company & Inc.
 Switzerland

Wolfgang Nejd
 Universität Hannover, KBS
 Appelstraße 4
 30167 Hannover
 Germany

Wolf-Tilo Balke
 L3S Research Center
 Expo Plaza 1
 30539 Hannover
 Germany

Kurt Tutschku
 Universität Würzburg
 Institut für Informatik, Lehrstuhl III
 Am Hubland
 97074 Würzburg
 Germany

Wolfgang Kellerer
 DoCoMo Communications
 Laboratories Europe GmbH
 Landsberger Straße 312
 80687 München
 Germany

Markus Hillenbrand
 TU Kaiserslautern
 AG ICSY
 Gottlieb-Daimler-Straße
 67663 Kaiserslautern
 Germany

Hermann de Meer
 Universität Passau
 Computer Networks & Computer
 Communications Group
 Innstraße 33
 94032 Passau
 Germany

Burkhard Stiller
 Universität Zürich, IFI
 Communication Systems Group
 Winterthurerstraße 190
 8057 Zürich
 Switzerland

Danny Raz
 Technion IIT
 Department of Computer Science
 Haifa 32000
 Israel

Wolf Siberski
 Universität Hannover, KBS
 Appelstraße 4
 30167 Hannover
 Germany

Gerhard Hasslinger
 T-Systems Technologiezentrum
 Deutsche-Telekom-Allee 7
 64307 Darmstadt
 Germany

Phuoc Tran-Gia
 Universität Würzburg
 Institut für Informatik, Lehrstuhl III
 Am Hubland
 97074 Würzburg
 Germany

Andreas Heinemann
 TU Darmstadt
 FG Telekooperation
 Hochschulstraße 10
 64289 Darmstadt
 Germany

Max Mühlhäuser
 TU Darmstadt
 FG Telekooperation
 Hochschulstraße 10
 64289 Darmstadt
 Germany

Christoph Lindemann
 Universität Dortmund
 Rechnersysteme und
 Leistungsbewertung
 August-Schmidt-Straße 12
 44227 Dortmund
 Germany

Thomas Hummel
 Accenture European
 Technology Park
 449, Route des Crêtes
 06902 Sophia Antipolis
 France

Jan Gerke
 ETH Zürich, TIK
 Gloriastrasse 35
 8092 Zürich
 Switzerland

Michael Conrad
 Universität Karlsruhe
 Institute of Telematics
 Zirkel 2
 76128 Karlsruhe
 Germany

Hannes Hartenstein
 Universität Karlsruhe
 Institute of Telematics
 Zirkel 2
 76128 Karlsruhe
 Germany

Martina Zitterbart
 Universität Karlsruhe
 Institute of Telematics
 Zirkel 2
 76128 Karlsruhe
 Germany

Oliver P. Waldhorst
 Universität Dortmund
 Rechnersysteme und
 Leistungsbewertung
 August-Schmidt-Straße 12
 44227 Dortmund
 Germany

Jussi Kangasharju
 TU Darmstadt
 FG Telekooperation
 Hochschulstraße 10
 64289 Darmstadt
 Germany

Steffen Muhle
 Universität zu Köln
 Seminar für Wirtschaftsinformatik,
 insb. Informationsmanagement
 Pohlstr. 1
 50969 Köln
 Germany

David Hausheer
 ETH Zürich, TIK
 Gloriastrasse 35
 8092 Zürich
 Switzerland

Jochen Dinger
 Universität Karlsruhe
 Institute of Telematics
 Zirkel 2
 76128 Karlsruhe
 Germany

Marcus Schöller
 Universität Karlsruhe
 Institute of Telematics
 Zirkel 2
 76128 Karlsruhe
 Germany

Daniel Rolli
 Universität Karlsruhe
 Lehrstuhl für
 Informationsbetriebswirtschaftslehre
 Englerstr. 14
 76128 Karlsruhe
 Germany

Ralf Ackermann
 TU Darmstadt
 KOM – Multimedia Communications
 Merckstraße 25
 64283 Darmstadt
 Germany

Nicolas C. Liebau
 TU Darmstadt
 KOM – Multimedia Communications
 Merckstraße 25
 64283 Darmstadt
 Germany

Luka Divic-Krnic
 TU Darmstadt
 KOM – Multimedia Communications
 Merckstraße 25
 64283 Darmstadt
 Germany

Timothy Roscoe
 Intel Research Berkeley
 2150 Shattuck Avenue
 Berkeley, CA 94704
 USA

关于本书

开始于Napster和Gnutella,P2P系统已成为Internet网络不可分割的部分,它吸引了数百万的用户。依据最近的评估,P2P流量现在超过了Web流量—Web流量曾经是Internet上的主要流量。虽然最受欢迎的P2P应用仍然是文件共享和内容分发,但像Internet电话这样的新应用正在出现。

在仅仅数年中,P2P系统迅速流行起来,对P2P的大量而深入的研究引发了P2P知识的海量增长。但是,本书是第一本教材性的综述,它提供了对该领域最新的和最具有深度的介绍。



本书特色

这本最新内容的综述系统地将各个领域中的必要条件集中起来，以一种具有原则性的和一致的方式讲解技术和方法，并就P2P体制的众多应用领域给出了深入完全的总体概貌。领军性的研究人员向本书贡献了他们的专业知识，每种知识都处于他／她自己特定的关注领域之中。

读者可从本书的所有方面的紧密组合的32章这样的大范围内容中选取章节，并因此可以以其独有的强调侧重而设计其课程的提纲。活跃于P2P领域的研发(R&D)人员将从本书中受益颇丰，并将其作为必备参考资料和深受鼓舞的一个有价值的源泉。

目 录

译者序

前言

原书作者名单

第 1 章 引言	1
1.1 写本书的意图	1
1.2 结构和内容	2
1.3 教学材料和书籍网站	4
1.4 致谢	4
第 1 部分 P2P：概念、领域、历史和未来	7
第 2 章 对等端到对等端(P2P)	9
2.1 定义	10
2.1.1 Internet 通信中体制的转换	11
2.2 在 P2P 系统和应用中的研究挑战	11
2.2.1 非结构化的 P2P 系统	14
2.2.2 结构化的 P2P 系统	14
2.3 小结	15
第 3 章 P2P 的过去和未来	16
3.1 现状：P2P 流量充满了网络	16
3.2 所有这些是如何开始的：从 ARPANET 到 P2P	17
3.3 Napster 的故事	18
3.4 Gnutella 和其近似物：完全非中心化架构	19
3.5 P2P 背后的驱动力	21
第 4 章 应用领域	23
4.1 信息	23
4.2 文件	24
4.3 带宽	26
4.4 存储空间	27
4.5 处理器周期	28

第 2 部分 非结构化 P2P 系统	31
第 5 章 第一代和第二代 P2P 系统	33
5.1 早期 P2P 系统的一般特征	33
5.2 中心化的 P2P 网络	35
5.2.1 基本特征	35
5.2.2 信令特征	36
5.2.3 讨论	38
5.3 纯粹 P2P 网络	39
5.3.1 基本特征	39
5.3.2 信令特征	41
5.3.3 讨论	42
5.4 混合 P2P 网络	45
5.4.1 基本特征	45
5.4.2 信令特征	47
5.4.3 讨论	48
第 6 章 随机图、小世界和规模无关网络	51
6.1 引言	51
6.2 定义	52
6.3 谜——真实网络的分析	54
6.4 随机图、小世界和规模无关网络的族类和模型	55
6.4.1 随机图	55
6.4.2 小世界——谜的第一个解	57
6.4.3 规模无关网络：富人是如何变得更富的	59
6.5 应用到 P2P 系统	61
6.5.1 在小世界中航行	61
6.5.2 P2P 系统中的小世界重叠网	63
6.5.3 P2P 系统中的规模无关重叠网	66
6.6 小结	67
第 3 部分 结构化 P2P 系统	69
第 7 章 分布式散列表	71
7.1 数据的分布式管理和检索	71
7.1.1 数据检索策略的比较	72
7.1.2 中心式服务器	72
7.1.3 洪泛搜索	73
7.1.4 分布式索引——分布式散列表	74

7.1.5 查找概念的比较	76
7.2 分布式散列表的基础知识	77
7.2.1 数据的分布式管理	77
7.2.2 在分布式散列表中寻址	77
7.2.3 路由	79
7.2.4 数据存储	79
7.3 DHT 机制	80
7.3.1 综述	80
7.3.2 节点到达	80
7.3.3 节点失效	81
7.3.4 节点离开	82
7.4 DHT 接口	82
7.4.1 路由接口	82
7.4.2 存储接口	82
7.4.3 客户接口	83
7.5 小结	83
第8章 DHT 算法选讲	85
8.1 Chord	85
8.1.1 标识符空间	85
8.1.2 路由	86
8.1.3 自组织	87
8.2 Pastry	89
8.2.1 标识符空间	89
8.2.2 路由信息	89
8.2.3 路由过程	91
8.2.4 自组织	91
8.2.5 路由性能	94
8.3 内容寻址网络	95
8.3.1 标识符空间	95
8.3.2 路由信息	96
8.3.3 路由过程	96
8.3.4 自组织	96
8.3.5 路由性能	98
8.4 Symphony	99
8.5 Viceroy	100
8.6 Kademlia	102
8.7 小结	103

第 9 章 分布式散列表中的可靠性和负载均衡	105
9.1 分布式散列表中数据的存储负载均衡	105
9.1.1 定义	106
9.1.2 统计分析	107
9.1.3 DHT 中负载均衡的算法	109
9.1.4 负载均衡方法的比较	114
9.2 分布式散列表中的数据可靠性	115
9.2.1 冗余	115
9.2.2 复制	116
9.3 小结	118
第 10 章 P-Grid: 结构化 P2P 系统中自组织过程的动态性	120
10.1 自组织的概念	120
10.2 非结构化 P2P 系统中自组织的一个示例	121
10.3 结构化 P2P 系统中的自组织	122
10.3.1 P-Grid 重叠网的结构	123
10.3.2 P-Grid 重叠网的动态性	125
10.3.3 启动一个 P-Grid 重叠网	125
10.3.4 路由表维护	128
10.3.5 维护机制的分析	130
10.4 小结	132
第 4 部分 基于 P2P 的应用	135
第 11 章 应用层组播	137
11.1 在应用层上组播的原因	137
11.2 设计的各方面及其分类	138
11.3 非结构化重叠网	139
11.3.1 中心化系统	139
11.3.2 完全的分布式系统	140
11.4 结构化的重叠网	142
11.4.1 基于洪泛的复制	143
11.4.2 基于树的复制	144
11.4.3 性能/开销评估	146
11.5 热点话题	147
11.6 小结	148
第 12 章 ePOST	149
12.1 有限范围的重叠网	150
12.1.1 设计	151