

Comprehensive Analysis of
P2P Technology

P2P技术 全面解析

■ 张春红 裘晓峰 弭伟 纪阳 等 编著

- ❖ 中国工程院院士刘韵洁作序
- ❖ 国家级、部级、企业合作项目成果总结
- ❖ 内容全面，覆盖整个P2P技术研究领域
- ❖ 阐述透彻，让读者真正理解P2P技术



人民邮电出版社
POSTS & TELECOM PRESS

Comprehensive Analysis of
P2P Technology

P2P技术 全面解析

▪ 张春红 裘晓峰 弭伟 纪阳 等 编著

人民邮电出版社
北京

图书在版编目(CIP)数据

P2P技术全面解析 / 张春红等编著. — 北京: 人民邮电出版社, 2010.5
ISBN 978-7-115-22189-6

I. ①P… II. ①张… III. ①因特网—基本知识 IV
①TP393.4

中国版本图书馆CIP数据核字(2010)第004556号

内 容 提 要

本书对 P2P 技术进行了全面而又深入的讲解, 内容涵盖了 P2P 的基础知识: P2P 概述、无结构的 P2P 系统、有结构的 P2P 系统、负载均衡以及可靠性等; 目前 P2P 技术的应用: P2P 网络中的 NAT 穿越技术、内容分发技术、P2P VoIP 技术等; P2P 与电信网的结合和应用: P2PSIP、P2P 与 IMS 的结合; P2P 领域中一些高级问题: P2P 安全、P2P QoS、移动 P2P 等。

本书适合从事 P2P 技术研发、应用、产业管理以及信息通信产业界关心 P2P 技术的相关从业人员阅读, 也可供高等院校通信、计算机等相关专业师生参考。本书也适合作为相关培训班的教材。

P2P 技术全面解析

-
- ◆ 编 著 张春红 裘晓峰 弭伟 纪阳 等
责任编辑 姚予疆
执行编辑 刘洋
 - ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街14号
邮编 100061 电子函件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
北京艺辉印刷有限公司印刷
 - ◆ 开本: 787×1092 1/16
印张: 27.5
字数: 672千字
印数: 1-3500册
- 2010年5月第1版
2010年5月北京第1次印刷

ISBN 978-7-115-22189-6

定价: 69.00元

读者服务热线: (010)67129264 印装质量热线: (010)67129223
反盗版热线: (010)67171154

序

在互联网领域，P2P（Peer to Peer，对等端到对等端）模式是作为 C/S（客户端/服务器）模式的对立面出现的。P2P 技术以其特有的自组织性、分布性，在互联网上迅速发展，已成为互联网网络不可分割的部分。而 P2P 技术的应用更是层出不穷，已扩展到商业、政府、通信等各大领域，包括内容分发、数据资源缓冲、商务自动化、应用业务管理、网格计算、分布式计算、数据部署、用户间通信、自组织网络等。

P2P 技术以其独具特色的互联网应用，迅速在网络中蔓延，产生了各方面积极和消极的影响，对电信运营商的网络规划、运营以及对终端用户的行为、体验等方面的影响都是不可忽视的。目前产业界对 P2P 的发展前景给予了很大关注，如电信业界对 P2PSIP 技术给予了充分重视，有人认为该技术有可能成为未来网络体系的备选方案。除了利用 P2P 技术补充完善已有的电信核心网络外，业界也在积极探索基于 P2P 技术本身的新的电信核心网络平台的构架。

本书基于作者多年来在 P2P 领域的研究成果和积累，力图从各个方面对 P2P 技术进行较全面的讲解。本书首先介绍了有关 P2P 的概念、技术原理、基本应用、系统结构等基础知识；然后逐步展开，详细地阐述了基于 P2P 的各种应用系统以及相关的热点技术；特别地，本书从 P2P 与电信网的结合方面进行了较为系统的探讨，并对 P2P 技术未来在电信网中的应用与走向做出了前瞻性的预测；最后，本书从安全性、QoS、移动 P2P、开发工具等方面对 P2P 领域的一些关键问题进行了阐述。

本书源于作者实验室长期从事 P2P 领域的科研工作，凝聚了实验室所有研究人员的科研成果。我希望本书能够为推动我国的 P2P 技术研究工作向前发展贡献一份力量。

中国工程院 院士



2010 年 3 月

前 言

P2P 技术，主要指网络的参与者共享他们所拥有的一部分硬件资源，这些资源提供的服务和内容能被 P2P 网络中的节点访问，访问不需要经过 P2P 网络外的其他中间实体。P2P 技术引导网络计算模式从集中式向分布式偏移，也就是说网络应用的核心从中央服务器向网络边缘的终端设备扩散。这使人们在 Internet 上的共享行为被提到了一个更高的层次，使人们以更主动的方式参与到网络活动中去。所以，P2P 给互联网的分布、共享精神带来了无限的遐想，很多新的应用被开发出来。

P2P 作为一种与服务器/客户端对立的网络架构，其应用近些年在互联网中发展十分迅速，新的 P2P 应用更是层出不穷。在 P2P 文件共享领域，如 eMule、BT 已分别拥有了数量庞大的用户群。尤其是 BT 的成功，带来了 P2P 应用的一个高潮。在 P2P 协同计算方面，Groove 已有了相对成熟的产品。虽然这方面应用软件总体数量较少，但随着协同计算概念的兴起，其需求将呈现急剧增长的趋势，拥有广阔的前景。在 P2P 的流媒体方面，国内外已有很多成功的产品，像 PPlive、PPS 这些较早的应用都已拥有了众多的用户，新的 P2P 流媒体应用也是在不断涌现。在 P2P 即时通信与 VoIP 方面，基于 P2P 技术的 VoIP 产品 Skype 的巨大成功更是把 P2P 应用再次推向了一个高潮。

P2P 技术在被广泛应用于互联网的同时，对电信运营商也产生了重大影响，P2P 应用在吸引大量客户的同时，会使网络流量急剧上升，导致电信运营商运营成本上升，扩容压力增大。同时，传统话务量开始逐渐被 P2P 语音通信技术应用大量分流。如何应对 P2P 带来的冲击和挑战已经成为摆在电信运营商面前的重要课题。随着 P2P 技术向电信运营商传统业务领域的进军，电信网络运营商对待这个新生互联网宠儿态度的变迁，可以看到一种新兴技术的出现带来的与既有格局的激烈博弈过程，体现出一种共同发展、互利双赢的趋势。因此，可以认为未来的网络将成为一个更加开放的平台，能够灵活和广泛地采用 P2P 技术，越来越丰富的业务将改变未来的生活。未来，P2P 将会走得更远。

作者所在的研究组从 2005 年开始关注基于 P2P 覆盖网络的多媒体业务研究，重点跟踪并开展了 P2PSIP (Peer to Peer Session Initiation Protocol) 方向的技术研究。研究组在 P2PSIP 的网络体系架构、异构 P2PSIP 网络互通机制、P2PSIP 网络性能分析与优化、P2PSIP 网络负载均衡、热点资源等问题上取得了大量研究成果。作者承担了国家“863”项目“对等网络会话初始化协议系统 P2P SIP 互通性研究”，并向 IETF P2PSIP 工作组提交了三篇相关草案。

近年来，研究组还开展了与 P2P 相关新业务的研究，承担了国家“863”项目“基于 P2P

P2P 技术全面解析

的新型家庭多媒体系统—新兴业务的研究”，提出了一种基于对等网络技术的新型家庭多媒体网络系统，包括系统终端侧和网络侧的功能架构，以及相关业务应用流程和系统部署方案。此方案较好地解决了传统客户端/服务器架构的性能瓶颈问题，从整体上提高了多媒体业务网络的性能，改善了用户的体验，有着很好的应用前景。

另外，研究组在 P2P 网络架构和算法、P2P 流量监测和安全方面也有深入的研究，尤其是对 P2P 技术的主流应用以及未来电信网中 P2P 的应用等进行了前瞻性研究。研究组在新型业务应用、多网络融合业务提供、可重配置无线网络、对等网络等技术领域推出了具有国际先进水平的创新成果，引领我国在相关领域的技术发展，形成一批自主知识产权成果。研究组通过对新型业务应用的研究带动产业发展并催生新型服务业，通过对新型网络体系结构的研究来探索和发现新型业务支撑能力和技术，推动源头创新，并积极推进引导相关领域国内和国际的标准化工作。

本书基于作者多年来在 P2P 领域的研究积累和成果，力图从各个方面对 P2P 给予一个全面的介绍。书中既有对 P2P 技术的介绍，也有关于经典 P2P 应用系统的详细探讨，更有作者自身对于 P2P 未来在电信网中应用与走向的大胆预测。

作为一本全面介绍 P2P 各研究方向的书，本书涵盖了 P2P 的概念、技术原理、应用等各个方面。

第一部分（第 1~2 章）对于 P2P 的概念、应用、发展历史及前景给出了初步的概述，这部分有利于读者对于 P2P 有初步的认识。

第二部分（第 3~6 章）从 P2P 网络的模型出发，介绍了非结构化和结构化 P2P 的基本原理，并对其中的关键问题，如资源搜索、负载均衡、可靠性等问题给予重点阐述。

第三部分（第 7~9 章）对于目前 P2P 技术的应用及遇到的问题，从技术的角度进行讲述，包括 P2P 网络中的 NAT 穿越技术、P2P 网络中的内容分发技术、P2P VoIP 技术。

第四部分（第 10~12 章）主要从 P2P 与电信网的结合方面，探讨了目前业界的研究现状，既包含了电信运营商对待 P2P 进行监测、利用的技术分析，又包含了目前主要的 P2P 与电信网结合的技术方向，如 P2PSIP、P2P 与 IMS 网络的结合等。

第五部分（第 13~16 章）从 P2P 安全、P2P 网络中的 QoS、移动 P2P、P2P 仿真与开发几个方面对 P2P 领域中一些高级问题分别给予阐述。

本书由北京邮电大学移动生活与新媒体实验室张春红、裘晓峰、纪阳老师，博士生弭伟、李漓春、马涛、王刚，以及硕士生王岩、岳晓雯、杨鹏、张永辉、王辉、熊淼、孙崇伟、刘玉婷、蒋旭昂、成城、陈全、高雅同学共同完成。由于 P2P 技术发展的日新月异，加上作者水平有限，因此书中难免出现错误和疏漏，希望广大读者批评指正。读者可通过本书编辑的电子邮箱（liuyang@ptpress.com.cn）与我们联系。

作者
2010 年 2 月于北京

目 录

第一部分 P2P 基本概念

| | |
|-----------------------------------|----|
| 第 1 章 P2P 技术基本概念 | 3 |
| 1.1 P2P 网络的定义 | 3 |
| 1.2 P2P 的特点 | 4 |
| 1.3 P2P 的功能和对应技术 | 5 |
| 1.4 P2P 的主要应用 | 6 |
| 1.5 P2P 技术发展历史及前景 | 7 |
| 1.5.1 P2P 的发展历史 | 7 |
| 1.5.2 发展前景 | 8 |
| 1.6 相关标准 | 9 |
| 1.6.1 IETF 组织制定的标准 | 9 |
| 1.6.2 中国通信标准化协会 (CCSA) 制定的标准 | 10 |
| 1.6.3 美国分布式计算产业协会 (DCIA) 的 P4P 标准 | 10 |
| 1.6.4 中国 P2P 标准化工作组的 DDP 标准 | 10 |
| 第 2 章 主流 P2P 应用 | 11 |
| 2.1 文件共享 | 11 |
| 2.1.1 Napster | 11 |
| 2.1.2 Gnutella | 12 |
| 2.1.3 BitTorrent | 14 |
| 2.1.4 eMule | 16 |
| 2.1.5 Maze | 17 |
| 2.2 即时通信 | 18 |

| | |
|--------------------|----|
| 2.2.1 Skype | 18 |
| 2.2.2 GTalk | 19 |
| 2.2.3 QQ | 20 |
| 2.3 流媒体 | 20 |
| 2.3.1 PPLive | 20 |
| 2.3.2 PPStream | 21 |
| 2.3.3 AnySee | 22 |
| 2.4 共享存储 | 22 |
| 2.4.1 CFS | 22 |
| 2.4.2 Total Recall | 23 |
| 2.4.3 OceanStore | 23 |
| 2.4.4 Granary | 24 |
| 2.5 对等计算与协同处理 | 24 |
| 2.5.1 SETI@home | 24 |
| 2.5.2 Groove | 25 |
| 参考文献 | 25 |

第二部分 P2P 原理分析

| | |
|------------------------|----|
| 第 3 章 P2P 网络模型概述 | 29 |
| 3.1 常见网络模型 | 29 |
| 3.1.1 随机网络 | 29 |
| 3.1.2 规则网络 | 30 |
| 3.1.3 小世界网络 | 31 |
| 3.2 集中目录式 P2P 网络模型 | 33 |
| 3.2.1 原理简介 | 33 |
| 3.2.2 典型应用: BitTorrent | 34 |
| 3.3 纯 P2P 网络模型 | 35 |

| | | | |
|--|----|---|-----|
| 3.3.1 纯 P2P 非结构化 网络模型 | 36 | 参考文献 | 99 |
| 3.3.2 纯 P2P 结构化网络 模型 | 38 | 第 6 章 DHT 网络中的负载均衡和 可靠性 | 101 |
| 3.4 分层式 P2P 网络模型 | 39 | 6.1 DHT 网络中的负载均衡概述 .. | 101 |
| 3.4.1 原理简介 | 39 | 6.2 DHT 负载均衡算法 | 103 |
| 3.4.2 查询机制 | 40 | 6.2.1 改变节点 ID 的负载 均衡算法 | 103 |
| 3.4.3 簇管理 | 40 | 6.2.2 移动资源的负载均衡 算法 | 105 |
| 3.4.4 性能分析 | 40 | 6.2.3 虚拟节点的负载均衡 算法 | 105 |
| 参考文献 | 41 | 6.2.4 热点问题的解决 ^[12-14] | 111 |
| 第 4 章 P2P 网络中的资源定位 方法 | 43 | 6.3 DHT 网络中的数据可靠性 | 112 |
| 4.1 资源搜索算法 | 43 | 6.3.1 DHT 网络中的数据 冗余机制 | 113 |
| 4.1.1 盲目搜索算法 | 43 | 6.3.2 DHT 网络中的冗余 数据分发机制 | 114 |
| 4.1.2 启发式智能搜索算法 | 53 | 6.3.3 DHT 网络中的错误 检查及一致性维护 机制 | 118 |
| 4.2 资源查询算法 | 55 | 6.3.4 DHT 网络中的数据 恢复机制 | 120 |
| 4.2.1 关键字查询 | 56 | 6.3.5 DHT 网络中数据的 安全问题 | 121 |
| 4.2.2 向量空间模型的检索 方案 | 58 | 参考文献 | 121 |
| 4.2.3 隐含语义的检索方案 | 60 | 第三部分 P2P 技术应用 | |
| 参考文献 | 61 | 第 7 章 P2P 网络中的 NAT 穿越 技术 | 127 |
| 第 5 章 结构化 P2P 系统的算法 | 64 | 7.1 P2P 网络中的 NAT 穿越问题 .. | 127 |
| 5.1 分布式散列表 (DHT) | 64 | 7.1.1 NAT 的原理与类型 | 127 |
| 5.1.1 DHT 背景 | 64 | 7.1.2 常见 NAT 穿越解决 方案 | 129 |
| 5.1.2 DHT 原理 | 64 | 7.2 P2P 网络中典型 NAT 穿越 技术详解 | 131 |
| 5.1.3 DHT 功能及实现 | 65 | 7.2.1 P2P 网络中 UDP 穿越 NAT 的原理与实现 | 131 |
| 5.2 DHT 的分类 | 66 | 7.2.2 P2P 网络中 TCP 穿越 | |
| 5.2.1 Multi-hop DHT | 66 | | |
| 5.2.2 $O(1)$ DHT | 66 | | |
| 5.2.3 One-hop DHT | 66 | | |
| 5.3 DHT 算法选讲 | 66 | | |
| 5.3.1 Chord | 66 | | |
| 5.3.2 Pastry | 73 | | |
| 5.3.3 CAN ^[6] | 79 | | |
| 5.3.4 Kademlia | 83 | | |
| 5.3.5 Kelips ^[9] | 88 | | |
| 5.3.6 Beehive | 92 | | |
| 5.3.7 One-hop ^[12] | 95 | | |
| 5.3.8 DIHT ^[13] | 98 | | |

| | | | |
|---|------------|--|------------|
| NAT 的原理与实现..... | 138 | 9.1.1 VoIP 基本概念和原理..... | 183 |
| 7.2.3 P2P 网络中 NAT 的 特性..... | 141 | 9.1.2 VoIP 中的关键技术..... | 184 |
| 7.3 基于 P2P 结构的 NAT 穿越 技术..... | 143 | 9.2 P2P VoIP 技术..... | 189 |
| 7.3.1 Skype 的 NAT 穿越 方案..... | 143 | 9.2.1 P2P VoIP 概述..... | 189 |
| 7.3.2 P2PSIP 中 ICE 的解决 方案..... | 144 | 9.2.2 P2P VoIP 中的关键 技术..... | 189 |
| 参考文献..... | 148 | 9.3 Skype 技术..... | 191 |
| 第 8 章 P2P 网络中的内容分发 技术..... | 149 | 9.3.1 Skype 概述..... | 191 |
| 8.1 内容分发技术分类..... | 149 | 9.3.2 Skype 的工作原理..... | 191 |
| 8.1.1 中心化服务器内容 分发..... | 149 | 9.3.3 Skype 的中继系统 研究..... | 197 |
| 8.1.2 CDN 网络内容分发..... | 150 | 9.4 基于 SIP 的 P2P VoIP..... | 198 |
| 8.1.3 P2P 网络内容分发..... | 152 | 9.4.1 基于 SIP 的 VoIP..... | 198 |
| 8.1.4 混合内容分发..... | 154 | 9.4.2 基于 SIP 的 P2P VoIP 基本体系结构..... | 199 |
| 8.2 P2P 中的内容分发技术..... | 155 | 9.4.3 基于 SIP 的 P2P VoIP 高级问题..... | 203 |
| 8.2.1 内容定位技术..... | 155 | 参考文献..... | 204 |
| 8.2.2 内容存储技术..... | 157 | | |
| 8.2.3 内容分片技术..... | 158 | 第四部分 P2P 与电信网的结合 | |
| 8.2.4 数据调度技术..... | 160 | 第 10 章 P2P 与电信运营商..... | 209 |
| 8.2.5 P2P 内容分发系统 举例..... | 163 | 10.1 P2P 对电信运营商的影响..... | 209 |
| 8.3 编码技术..... | 164 | 10.2 电信运营商对 P2P 流量的 监管..... | 211 |
| 8.3.1 网络编码概述..... | 164 | 10.2.1 P2P 监管问题概述..... | 212 |
| 8.3.2 几种主要的网络编码 方式..... | 167 | 10.2.2 P2P 流量识别技术..... | 214 |
| 8.3.3 网络编码的应用..... | 168 | 10.2.3 P2P 流量控制技术..... | 219 |
| 8.3.4 网络编码在 P2P 内容 分发中的作用..... | 174 | 10.2.4 P2P 流量监控系统 解决方案..... | 219 |
| 8.4 应用层多播技术..... | 175 | 10.3 电信运营商与 P2P 应用的 协作..... | 222 |
| 8.4.1 应用层多播简介..... | 175 | 10.3.1 基于网络拓扑信息的 协作方案..... | 222 |
| 8.4.2 P2P 应用层多播结构..... | 176 | 10.3.2 基于 P2P Cache 的 协作方案..... | 230 |
| 8.4.3 P2P 应用层多播系统..... | 177 | 10.4 电信运营商对 P2P 技术的 利用..... | 234 |
| 参考文献..... | 180 | 10.4.1 可运营、可管理的 | |
| 第 9 章 P2P VoIP 的技术分析..... | 183 | | |
| 9.1 VoIP 概述..... | 183 | | |

| | | | |
|--|-----|--|-----|
| P2P 网络..... | 234 | 12.1.1 IMS 基本概念..... | 285 |
| 10.4.2 发展 P2P 重叠网络 进行互联网的分布式 控制管理..... | 236 | 12.1.2 IMS 框架和功能实体 介绍..... | 285 |
| 10.4.3 分布式业务网络 (DSN)..... | 241 | 12.1.3 IMS 与 NGN、 软交换的关系..... | 287 |
| 10.4.4 P2P 核心网平台..... | 243 | 12.1.4 为什么选择 IMS..... | 288 |
| 参考文献..... | 244 | 12.1.5 IMS 的现状..... | 289 |
| 第 11 章 P2PSIP | 246 | 12.2 P2P 与 IMS 网络的融合..... | 294 |
| 11.1 SIP..... | 246 | 12.2.1 IMS 如何和 P2P 融合..... | 294 |
| 11.1.1 SIP 简介..... | 246 | 12.2.2 P2P 引入 IMS ^[8] | 296 |
| 11.1.2 传统 SIP 会话流程..... | 246 | 参考文献..... | 305 |
| 11.1.3 传统 SIP 中的注册 机制..... | 247 | | |
| 11.2 P2PSIP 网络的结构及其 实现..... | 250 | 第五部分 P2P 高级问题 | |
| 11.2.1 P2PSIP 简介..... | 250 | 第 13 章 P2P 网络中的安全问题 | 309 |
| 11.2.2 P2PSIP 常用术语..... | 251 | 13.1 P2P 应用带来的安全问题..... | 309 |
| 11.2.3 P2PSIP 网络架构..... | 251 | 13.1.1 知识产权问题..... | 309 |
| 11.2.4 用户注册过程..... | 255 | 13.1.2 P2P 应用对现有 网络的威胁..... | 311 |
| 11.2.5 用户呼叫过程..... | 256 | 13.2 P2P 的安全需求..... | 312 |
| 11.3 P2PSIP 的协议..... | 257 | 13.2.1 P2P 应用的安全需求..... | 312 |
| 11.3.1 RELOAD 协议综述..... | 257 | 13.2.2 P2P 网络安全威胁与 安全需求..... | 313 |
| 11.3.2 RELOAD 的术语 定义..... | 259 | 13.2.3 P2P 网络攻击之间的 关系..... | 316 |
| 11.3.3 RELOAD 的覆盖 网络管理需求及 协议的设计..... | 260 | 13.3 P2P 中的安全技术..... | 318 |
| 11.3.4 RELOAD 协议支持的 应用概述及具体设计..... | 273 | 13.3.1 认证..... | 318 |
| 11.3.5 RELOAD 协议的安全 考虑..... | 278 | 13.3.2 密钥管理..... | 323 |
| 11.4 P2PSIP 的互联互通问题..... | 280 | 13.3.3 P2P 与路由安全..... | 331 |
| 11.4.1 分级 P2PSIP..... | 280 | 参考文献..... | 334 |
| 11.4.2 P2PSIP 与传统 SIP 互通解决方案..... | 283 | 第 14 章 P2P 网络中的 QoS | 336 |
| 参考文献..... | 283 | 14.1 QoS 综述..... | 336 |
| 第 12 章 P2P 与 IMS 的结合 | 285 | 14.1.1 QoS 概念..... | 336 |
| 12.1 IMS 概述..... | 285 | 14.1.2 QoS 中的性能指标..... | 337 |
| | | 14.1.3 电信网的 QoS..... | 340 |
| | | 14.2 P2P QoS 综述..... | 344 |
| | | 14.2.1 文件下载..... | 345 |
| | | 14.2.2 流媒体..... | 346 |

| | | | |
|---|-----|-----------------------------------|-----|
| 14.2.3 VoIP..... | 347 | 15.5.2 Chord 在无线环境中的 性能分析..... | 391 |
| 14.3 P2P QoS 机制与算法..... | 349 | 15.5.3 流媒体在 MP2P 中的 性能分析..... | 397 |
| 14.3.1 测量机制与算法..... | 349 | 15.6 本章总结..... | 401 |
| 14.3.2 节点选择机制与算法..... | 354 | 参考文献..... | 402 |
| 14.3.3 激励机制与算法..... | 357 | 第 16 章 仿真和开发 | 404 |
| 14.3.4 覆盖层路由..... | 358 | 16.1 PlanetLab..... | 404 |
| 参考文献..... | 365 | 16.1.1 PlanetLab 简介..... | 404 |
| 第 15 章 移动 P2P (MP2P) | 367 | 16.1.2 PlanetLab 的典型 应用..... | 405 |
| 15.1 MP2P 的重要性..... | 367 | 16.1.3 组织原则..... | 405 |
| 15.2 无线环境的特点及 Ad Hoc 介绍..... | 368 | 16.1.4 责任人和信任关系..... | 407 |
| 15.2.1 无线环境与有线 环境的区别..... | 368 | 16.2 OverSim..... | 408 |
| 15.2.2 Ad Hoc 网络..... | 368 | 16.2.1 OverSim 简介..... | 408 |
| 15.3 MP2P 概述..... | 371 | 16.2.2 OverSim 的平台结构..... | 409 |
| 15.3.1 MP2P 的特点..... | 371 | 16.2.3 OverSim 的开发 环境..... | 411 |
| 15.3.2 MP2P 的应用前景..... | 373 | 16.3 JXTA..... | 422 |
| 15.4 MP2P 网络体系结构..... | 374 | 16.3.1 JXTA 简介..... | 422 |
| 15.4.1 集中式结构..... | 374 | 16.3.2 JXTA 的基本概念..... | 422 |
| 15.4.2 半分布式结构..... | 375 | 16.3.3 JXTA 的协议及其 应用体系结构..... | 425 |
| 15.4.3 中间件体系结构..... | 376 | 16.3.4 JXTA 开发环境..... | 426 |
| 15.5 MP2P 性能分析..... | 377 | 参考文献..... | 428 |
| 15.5.1 Gnutella 及 XL-Gnutella 在无线环境中的性能 分析..... | 377 | | |

第一部分

P2P 基本概念

第 1 章 P2P 技术基本概念

第 2 章 主流 P2P 应用



第 1 章

P2P 技术基本概念

要了解 and 认识 P2P 技术，就必须从 P2P 的基本概念入手。本章首先给出一个相对详细的 P2P 网络的定义，随后介绍 P2P 网络的一些特点。P2P 的功能和具体技术以及目前的应用有很多，本章主要对 P2P 的功能作一个分类并简要介绍 P2P 的几种应用，具体内容及相应技术将在后面的章节中详细介绍。本章的最后两节将分别介绍 P2P 的发展历史及前景以及目前有关 P2P 的一些相关标准。

1.1 P2P 网络的定义

本书给出的 P2P 网络定义为：P2P (Peer-to-Peer) 网络是一种在 IP 网络之上的应用层的分布式网络，网络的参与者即对等节点 (peer) 共享他们所拥有的一部分硬件资源 (如处理能力、存储能力、网络连接能力等)。P2P 网络中的这些共享资源提供的服务和内容能被 P2P 网络中的节点访问，访问不需要经过 P2P 网络外的其他中间实体。在 P2P 网络中的对等节点既是资源 (服务和内容) 提供者，又是资源 (服务和内容) 获取者。

由于 P2P 文件共享以及 P2P 流媒体在线播放应用是人们最为熟悉的 P2P 应用，因此很多人将 P2P 狭义地理解为 P2P 文件共享和内容分发。实际上，P2P 网络的应用非常广泛，除了即时文件共享和内容分发外，P2P 还可用于实时通信、网络游戏、协同工作等应用。P2P 网络也不一定是由用户 PC 组成的，也有可能是由服务器组成的，还有可能是由特殊终端组成的，比如机顶盒、PDA、传感器等。

也有人将 P2P 定义为：利用网络边缘的普通用户 PC 来提供服务。本文认为这种定义并不准确，不能涵盖服务器之间组成的 P2P 网络，也包括了 master/slave 架构的利用普通用户 PC 的分布式计算，如 SETI@home 等。这种分布式计算方式虽然利用普通用户的资源提供服务，但是服务的对象是 master 服务器，普通用户间不通信，不相互提供服务。这种架构当中没有对等的关系。

P2P 与分布式计算是一种从属关系，分布式计算是一个更大范畴的概念。分布式计算最初是指在一台计算机多个处理器之间的协同计算。随着计算机网络的发展，分布式计算一般是指多台计算机协同完成任务。P2P 属于分布式计算，但某种分布式计算不一定是 P2P。分布式计算中的各台计算机之间的角色可能是不对等的，比如，master/slave 架构的网格和 SETI@home。

P2P 技术全面解析

Ad Hoc 网络是一种无线环境中的分布式网络，它与 P2P 有很多相似的地方，但也有一些差别。Ad Hoc 是一个网络层（IP 层）上的概念，Ad Hoc 网络中所有节点都执行路由算法和协议，承担 IP 层路由。P2P 是一个应用层的概念，P2P 网络是建立在 IP 层网络之上的。Ad Hoc 网络中的节点在应用层上的关系一般也是对等的，一般也组成了 P2P 网络。

1.2 P2P 的特点

P2P 网络作为一种分布式网络，打破了传统的 Client/Server (C/S) 模式，网络中每个节点地位都是对等的，具备了分布式网络的很多特点，图 1-1 给出了一种简单的 C/S 模型与 P2P 网络模型的对比图。

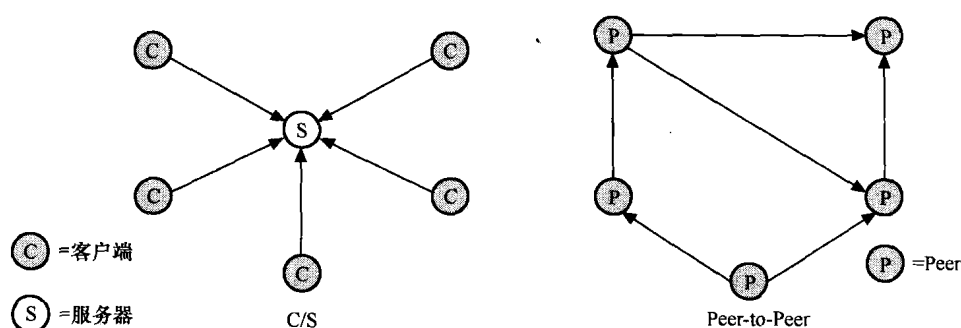


图 1-1 C/S 模型与 P2P 网络模型的对比

传统的 C/S 模式中，客户端之间进行交互都需要依赖中心化的服务器，服务器就容易成为网络的瓶颈，当网络规模越大时，服务器的负担就会越重，一旦服务器崩溃，整个网络就会瘫痪。另外由于服务器的能力有限，往往不能及时地对每一个客户端的请求作出响应，因此会造成资源利用率低下。而 P2P 网络中没有中心化的服务器，不存在系统瓶颈，每个节点既充当客户端又充当服务器，因而具有很高的资源利用率。

总的来说，P2P 网络的特点主要体现在下面几个方面。

(1) 可扩展性：不管是用户节点组成的 P2P 网络还是服务器组成的 P2P 网络，都是完全分布式的，不存在单点性能上的瓶颈。对于用户节点组成的 P2P 网络，随着用户的加入，不仅服务的需求增加了，系统整体的资源和服务能力也在同步地扩充，始终能较容易地满足用户的需要，理论上其可扩展性几乎可以认为是无限的。对于服务器组成的 P2P 网络，只需向 P2P 网络中增加服务器即可平滑扩容。由于 P2P 自组织、自配置、自动负载均衡的特性，系统扩容变得非常容易。

(2) 健壮性：P2P 架构天生具有耐攻击、高容错的优点。由于服务是分散在各个节点之间进行的，部分节点或网络遭到破坏对其他部分的影响很小。P2P 网络一般在部分节点失效时能够自动调整整体拓扑，保持其他节点的连通性。P2P 网络通常都是以自组织的方式建立起来的，并允许节点自由地加入和离开。P2P 网络还能够根据网络带宽、节点数、负载等变化不断地作自适应式的调整。

(3) 高性能/价格比：一方面，对于用户节点组成的 P2P 网络来说，采用 P2P 架构可以有

效地利用互联网中散布的大量普通用户节点的空闲资源，不需要部署服务器或需要的服务器很少。P2P 网络可以将计算任务或数据分布到所有用户节点上，利用其中闲置的带宽、计算能力或存储空间，达到高性能计算、海量数据传输、海量数据存储的目的。另一方面，对于服务器组成的 P2P 网络来说，采用 P2P 这种分布式架构后，可以使用一群高性价比的普通服务器来取代价格极其昂贵的超级服务器。

(4) 私密性：在 P2P 网络中，由于信息的传输分散在各节点之间进行而无需经过某个集中环节，用户的隐私信息被窃听和泄漏的可能性大大缩小。用户组成的纯 P2P 网络中也不会出现服务提供商滥用个人信息、出售个人信息的情况。此外，采用中继转发的技术方法，可以将通信的参与者隐藏在众多的网络实体之中。传统的一些匿名通信通常使用中继服务器来实现匿名通信的目的。而在 P2P 中，所有参与者都可以提供中继转发的功能，因而大大提高了匿名通信的灵活性和可靠性，能够为用户提供更好的隐私保护。

(5) 流量均衡：P2P 网络环境下，硬件资源和数据内容分布在多个节点，而 P2P 节点可以分布在网络各个角落，可以很好地实现整个网络的流量均衡。

(6) 自组织、低部署维护成本：P2P 网络环境下，一般没有中心管理者，普通用户也缺乏高级的计算机知识。因此，P2P 网络采用了自动计算技术，以实现 P2P 网络的自组织、自配置、自愈等特性，从而大大降低了对人为干预的需要。对于采用 P2P 技术的服务运营者来说，这大大降低了系统的部署维护成本，降低了出现人为配置错误的可能性。

1.3 P2P 的功能和对应技术

P2P 网络主要的功能可分为 3 种：基于 P2P 的数据发布和传输、基于 P2P 的数据存储和检索、基于 P2P 的分布式数据处理。

基于 P2P 的数据发布和传输是指 P2P 网络中点到点的具体数据传输过程，传输的方式有一对一、一对多、多对多三种模式。一对一模式即一个源端到一个目的端的传输，较为简单，主要见于即时通信和 VoIP 等应用的用户间直接通信，目前的即时通信与 VoIP 在连接建立后的数据传输过程中基本都是基于 P2P 的。一对多模式即一个源端到多个目的端的传输，应用层多播、广播都是这种模式，主要用于群组通信和消息广播，还常用于 P2P 网络的维护。多对多模式即多个源端到多个目的端的传输，主要见于内容分发应用，比如 BT、电骡等文件分发和 PPLive、PPS 等流媒体内容分发。

P2P 数据存储和检索是指在 P2P 网络中的分布式存储和检索方法，包括基于结构化 P2P 网络的方法和基于非结构化 P2P 网络的方法两类。结构化 P2P 网络是一种基于分布式散列表 (DHT, Distributed Hash Table) 的 P2P 网络。基于结构化 P2P 网络的数据存储和检索建立在确定性拓扑结构的基础上，并基于 DHT 算法实现存储和数据检索。非结构化 P2P 数据存储和检索基于非结构化 P2P 网络，资源存储通常与 P2P 系统的拓扑结构无关，一般只放置在本地，资源检索则可分为两类：盲目搜索和启发式搜索。有关非结构化 P2P 网络和结构化 P2P 网络中的存储和检索方法会在第 4、5 章中详细介绍。

大多分布式数据处理采用的是 master/slave 架构。利用 P2P 对等网络，也能够进行分布式数据处理技术。基于 P2P 的分布式数据处理与分布式计算技术联系更为紧密，涉及大量分布式计算与并行计算的知识，因此本书不涉及基于 P2P 的分布式数据处理。

1.4 P2P 的主要应用

目前，P2P 技术的应用已扩展到军事、商业、政府、通信等各大领域，根据具体应用不同，可以把 P2P 分为以下几种类型。

(1) 用于内容共享和数据分发，例如文件共享的 Napster、Gnutella、电骡、BT 和流媒体分发的 PPLive、PPS。P2P 技术使互联网上任意两台计算机间共享数据成为了可能，利用 P2P 技术可以使计算机间不通过中心服务器直接进行内容交互和数据分发，比起传统 C/S 模式中需要把数据上传到服务器再从服务器下载的方式，利用 P2P 的方式既方便了用户又省去了部署大量中心服务设备的开销。Napster 就是利用人们对 MP3 资源的需求提供给人们一种方便下载 MP3 资源的方式从而引发了网络的 P2P 技术革命。电骡、BT 更是被广泛地应用来下载各种文件。基于 P2P 方式的流媒体技术很好地解决了基于 C/S 模式的传统流媒体中带宽不足的问题，使得媒体播放更加流畅，PPLive、PPS 正是由于其流畅的播放而获得了广泛的应用。

(2) 挖掘 P2P 对等计算能力和存储共享能力，例如 SETI@home、Avaki、Popular Power 等。利用 P2P 网络的分布式结构可以构造分布式的存储系统实现存储共享，它可以提供高效率的、顽健的和负载均衡的文件存取功能。尤其是当存储规模较大时，集中方式往往不再适合，这种情况下就需要构建一个分布式的存储系统来管理系统中的数据。P2P 网络的节点除了可以共享存储能力之外，还可以共享 CPU 处理能力，通过把网络中众多计算机的计算能力联合起来就可以执行超级计算机的任务。目前已有很多公司投入对等计算的开发，Inter 也利用对等计算技术来设计其 CPU，可以为其节省极大的费用，SETI@home 则是迄今为止最成功的分布式计算试验项目。

(3) 基于 P2P 方式的协同处理与服务共享平台，例如 JXTA、Magi、Groove 等。协同处理和服务共享是指多用户之间利用网络中的协同计算平台互相协同来共同完成计算任务、共享信息资源等。通过采用 P2P 技术使得互联网上任意两台 PC 都可建立实时的联系，采用多种方式建立在线、非在线的协同应用环境。协同应用一般包括：实时通信、聊天室、文件共享、语音通信等基本功能以及共享白板、协同写作、视频会议，甚至包括工程人员的协作开发软件。比起传统协同工作的 Web 实现方式，采用 P2P 技术使协同工作不再需要中心服务器，参与协同工作的计算机可以建立点对点的连接。Groove 就是一个典型的基于 P2P 的协同软件平台，目前已经被微软公司收购。

(4) 采用 P2P 方式的即时通信交流，包括 ICQ、QQ、Skype、P2PSIP 等。即时通信与 VoIP 应该是日常使用最多的应用了，QQ、Skype 这些都是目前最流行的即时通信与 VoIP 软件。这些应用软件都采用了 P2P 技术作为其实现的一种技术方式，只是应用的程度不同。以 QQ 为代表的传统即时通信软件只是在连接建立后数据传输时采用 P2P 技术，而以 Skype 为代表的 P2P VoIP 软件则是从连接建立到数据传输过程中都是采用 P2P 实现的。Skype 良好的通话质量令其拥有了众多的注册用户，成为目前最成功的 P2P VoIP 应用，这也证明了采用 P2P 方式进行即时通信与 VoIP 具有良好的应用前景。

基于 P2P 的各种应用所利用的 P2P 功能也不太一样。P2P 的文件和其他内容共享主要利用的是 P2P 数据传输这一功能，有的也会采用 P2P 数据存储和查询这一功能来进行内容的定