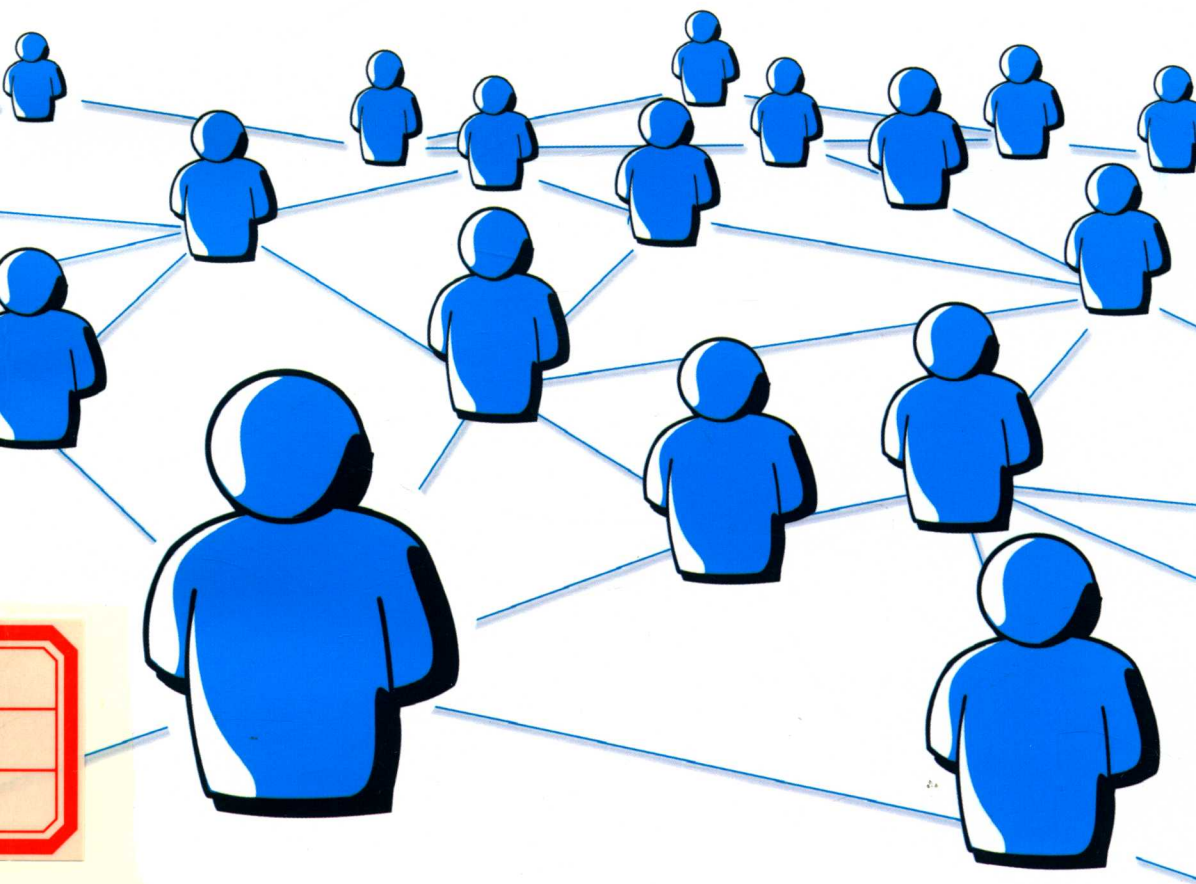
P2P Network
Measurement and Analysis

P2P网络测量与分析

张宏莉 叶麟 史建焘◎著

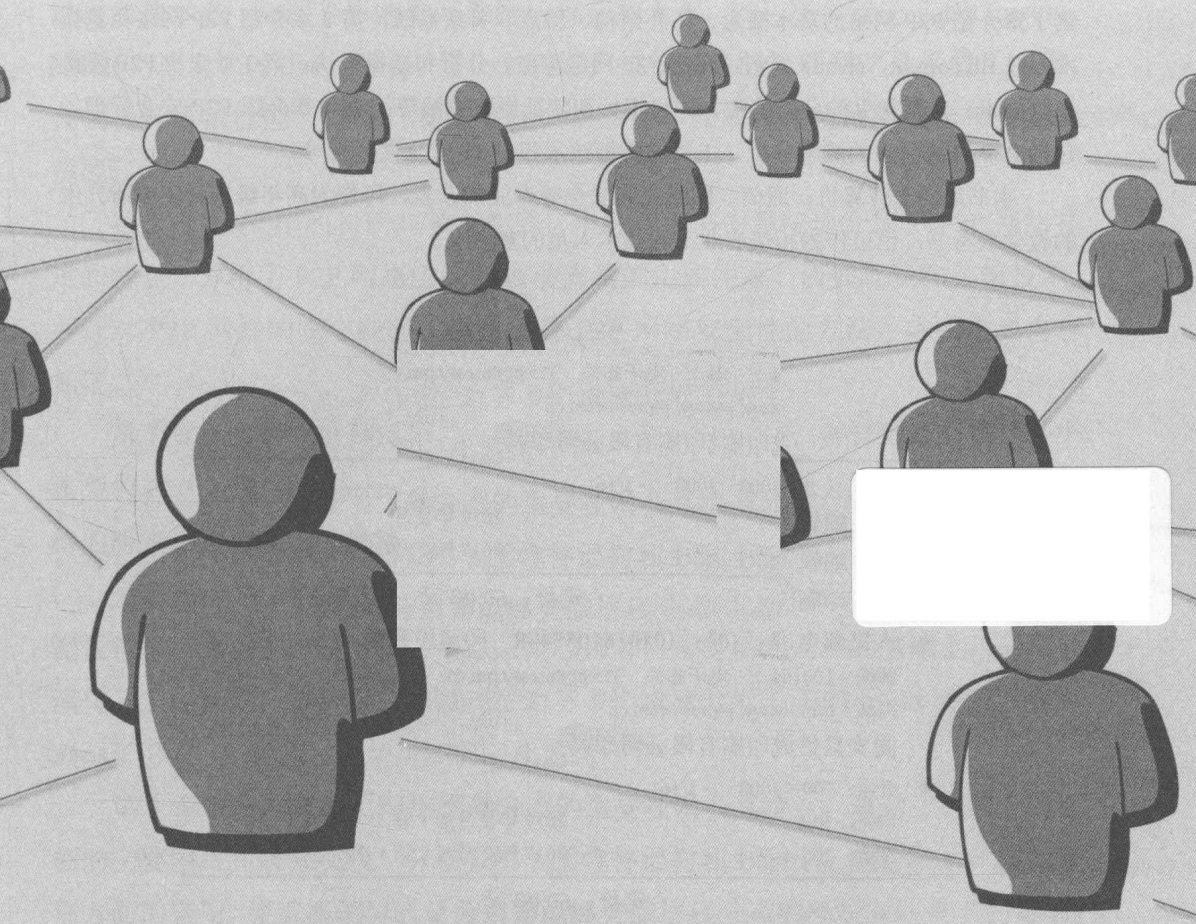


中国工信出版集团
人民邮电出版社
POSTS & TELECOM PRESS

P2P Network
Measurement and Analysis

P2P网络测量与分析

张宏莉 叶麟 史建焘◎著



人民邮电出版社
北京

图书在版编目（CIP）数据

P2P网络测量与分析 / 张宏莉，叶麟，史建焘著. -- 北京 : 人民邮电出版社，2017.2
ISBN 978-7-115-43165-3

Ⅰ. ①P… Ⅱ. ①张… ②叶… ③史… Ⅲ. ①计算机网络－研究 Ⅳ. ①TP393

中国版本图书馆CIP数据核字(2017)第019746号

内 容 提 要

本书全面、深入地介绍了 P2P 网络测量、流量识别、安全管理等技术。全书分为 4 章：第 1 章介绍 P2P 网络的基本概念、拓扑结构、特点和典型系统；第 2 章介绍 P2P 网络测量技术，以 BitTorrent、eMule 为例，介绍 P2P 网络测量、分析和建模方法；第 3 章介绍 P2P 流量识别技术，包括常见的分类算法、评价指标和特征提取方法等；第 4 章介绍 P2P 安全管理，包括 P2P 系统面临的威胁分析、主要的防御技术和 DHT 安全性分析等。

本书适用于计算机、通信、网络空间安全等专业的读者，可作为高年级本科生或研究生的教学参考书，也可作为相关专业工程技术人员的参考书。

◆ 著　　张宏莉　叶　麟　史建焘
责任编辑　邢建春
执行编辑　肇　丽
责任印制　彭志环

◆ 人民邮电出版社出版发行　　北京市丰台区成寿寺路 11 号
邮编　100164　　电子邮件　315@ptpress.com.cn
网址　http://www.ptpress.com.cn
固安县铭成印刷有限公司印刷

◆ 开本：700×1000　1/16
印张：10.5　　2017 年 2 月第 1 版
字数：206 千字　　2017 年 2 月河北第 1 次印刷

定价：69.00 元

读者服务热线：(010)81055488　印装质量热线：(010)81055316
反盗版热线：(010)81055315

前　言

对等网络（P2P，Peer-to-Peer）技术被誉为 21 世纪初期互联网上发展最为迅猛的技术之一，深刻影响着互联网的现在和未来。P2P 技术催生了 BitTorrent、eMule、迅雷、PPlive、Skype 等一系列用户规模过亿的互联网“杀手级”应用，大幅提升通信性能。同时，其匿名性、动态性和开放性也给网络安全带来新的挑战。

结合过去 10 年的科研与教学实践，作者在本书中系统地介绍了 P2P 网络技术和特点，分析了 P2P 网络主动和被动测量的主要困难、研究现状和关键技术，并针对 P2P 网络的安全缺陷，提出提升 P2P 网络安全性的方法。本书主要内容如下。

第 1 章介绍了 P2P 网络的起源、发展、结构以及特点，并详细地介绍了 3 个典型的 P2P 系统：BitTorrent、eMule 和 Kademlia，在此基础上总结了当前 P2P 网络的研究热点和面临的挑战。

为了揭示 P2P 网络的内在机制和运行规律，第 2 章介绍了高效的 BitTorrent 测量技术，涵盖了种子采集、节点采集和内容测量等功能。在此基础上，从拓扑演化、宏观行为、节点行为和微观行为等多个角度分析了 BitTorrent 系统，深入刻画了 BitTorrent 的节点到达和离开模型。

第 3 章系统地介绍了网络流量分类的主要技术及其特点，总结了流量分类面临的主要困难和研究现状。分别介绍了柔性神经树和 Bagging 算法在 P2P 流量识别方面的应用，使机器学习流量分类器能够克服协议分布的不平衡性带来的影响。同时，针对面向轻量级深分组检测的负载特征提取问题，分别应用层次聚类方法

迭代提取字节特征和关联挖掘方法从不随机位中提取位特征，实现负载特征有效提取。

第 4 章以典型的 P2P 文件共享系统的行为周期模型为脉络，分别总结了资源发布/搜索、节点选择、分块选择以及数据传输 4 个不同阶段下的主要安全威胁，并从系统设计的角度，有针对性地提出了假块污染攻击防御方法和 DHT 安全加固方法。

感谢哈尔滨工业大学计算机网络与信息安全研究中心的教师和研究生对本书的支持。本书虽经多次修改和校对，但由于时间仓促，疏漏及缺点在所难免，热忱欢迎广大读者批评指正。

作　者

2016 年 12 月

目　录

第1章 绪 论

1.1 P2P 网络概述

1.1.1 P2P 网络的起源与发展

P2P 概念与 Internet 的历史一样悠长，Internet 设计的初衷就是对等共享，但受限于计算能力，其对等性一般体现在 Internet 核心部分。近年来，主机计算、存储能力和网络带宽的不断提高使 P2P 重新成为人们关注的焦点，P2P 概念有了更为广阔的延伸，使位于网络边缘的主机都能参与网络构建与共享，极大地推动了 Internet 的发展。

但 P2P 概念并没有形成一个统一而严格的定义，研究者仅对其进行如下的定性描述与概括。

P2P 是一类利用位于网络边缘的存储、计算、内容、人工资源的网络应用[1]。P2P 分布式访问意味着 P2P 应用运行在一个无稳定连接、无预知 IP 的高度动态环境中，必须具有相当的自治能力。

一个没有等级和中心控制的分布式系统，是位于 IP 层之上并提供一系列如广域路由、数据查询、近邻选择、冗余存储、信任机制、匿名、容错等特性的自组织网络[2]。简而言之，P2P 是一个处于应用层的对等共享的分布式系统，其抽象的体系结构如图 1-1 所示。位于最底层的网络层描述节点间的网络性质；节点层负责节点的管理、发现和路由；特性层实现安全、资源整合、可靠、容错等特性的 P2P 网络管理；服务层针对不同的应用提供并行的任务、内容和文件管理支持；最终应用层在底层的构建基础之上实现具体的应用。

所有参与系统的节点处于完全对等的地位，没有客户端和服务器之分，既向别人提供服务，也享受来自别人的服务[3]。

真正促进 P2P 概念快速发展的推动力来自第一代 P2P 网络的典型代表 Napster[4]，

其采用中央索引服务器记录所有用户共享的资源，并为用户提供检索服务。中央服务器的存在使其检索极为高效，但也使 Napster 在后续发展中遭受版权的困扰，并最终被迫关闭。为了应对版权诉讼威胁，第二代 P2P 网络应运而生，如早期的 Gnutella[5]，其设计者试图采用无中心的方式组织节点避免单点失效，但其简单的洪泛（Flooding）查询方式极大地降低了 P2P 网络的可用性。在前两代 P2P 的基础上，第 3 代 P2P 网络主要从两个方面进行改进。一方面，节点组织采用动态分层方式，部分节点被动态选取扮演局部中心节点，如 KaZaA[6]；另一方面，节点组织采用分布式散列表（Distributed Hash Table）技术，节点以结构化方式组织提高查询效率，如 Kadmilia[7]。而新一代 P2P 网络不仅关注自身的优化，也开始考虑 Internet 底层基础设施因素[8]，尝试与 ISP 相互协作进一步优化 P2P 网络，降低 P2P 网络对底层设施的冲击。

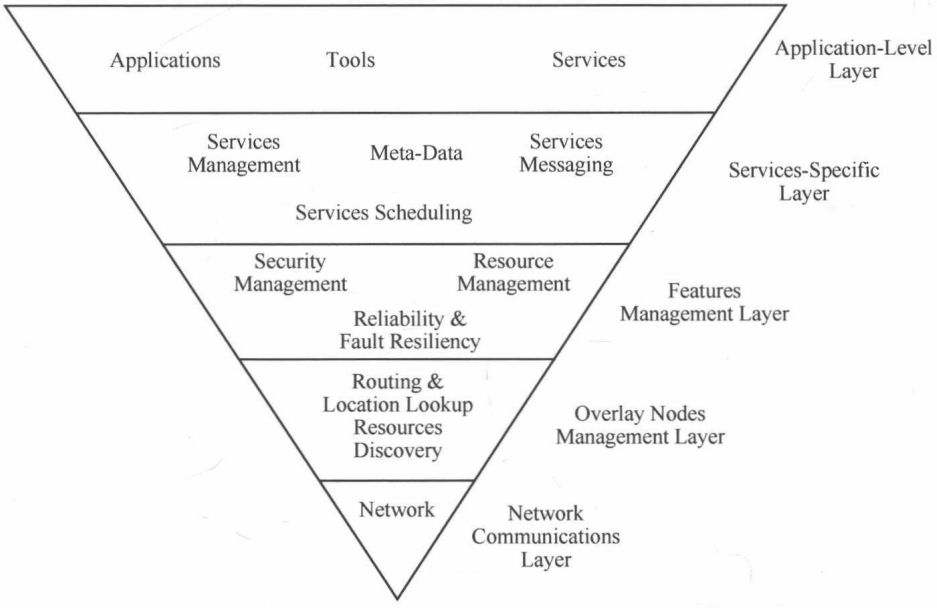


图 1-1　抽象 P2P 覆盖网络体系结构

从 P2P 的发展历程可以看出，P2P 网络经历了盲目发展—激烈对抗—和谐共存的 3 个发展阶段，无论 P2P 网络的整体行为还是客户端的个体行为，在这个过程中均产生了不同的变化。如今，一个客户端的网络边界已趋于模糊，单个客户端可以跨越多个不同类型的 P2P 网络，P2P 网络的行为变得更加复杂、有趣。

1.1.2　P2P 系统的拓扑结构

拓扑结构是 P2P 设计过程的核心问题，直接决定了整个网络的基本形态，并

且影响着系统中大量对等节点的命名、组织管理、节点加入/退出方式、容错机制和资源定位机制的设计。由于 P2P 网络的拓扑维护和路由是在叠加于 IP 层之上的应用层进行的，因此也被称为覆盖网络（Overlay Network）。P2P 网络的发展史实际上就是 P2P 覆盖网络拓扑结构的发展史，其发展过程是伴随着节点间的链路连通方式以及共享资源的定位技术的不断变革而进行的。到目前为止，P2P 网络的拓扑结构如图 1-2~图 1-5 所示，大致经历了以下几个过程：中心式 P2P 网络、无结构 P2P 网络、结构化 P2P 网络以及混合式 P2P 网络。

（1）中心式 P2P 网络。最初的 P2P 程序 Napser 诞生于 1998 年美国的波士顿大学，由当时还是一年级学生的肖恩·范宁编写，其拓扑结构如图 1-2 所示。服务器对用户共享的资源建立索引并提供查询服务，下载用户从服务器获得文件所有者的信息后直接与相应节点进行数据传输。由于资源的查询和传输过程是分离的，与传统的 C/S 模式相比，大大降低了服务器的负载。随后，发展起来的很多 P2P 文件共享系统都沿用了 Napser 的拓扑结构，包括当前著名的 BitTorrent[9]、eDonkey2000[10]和 FS2You[11]等。虽然中心化拓扑很好地保证了网络的性能，但是，由于中心组件的存在，制约了整个系统的扩展性和顽健性，同时也容易受到法律诉讼的威胁[12]。

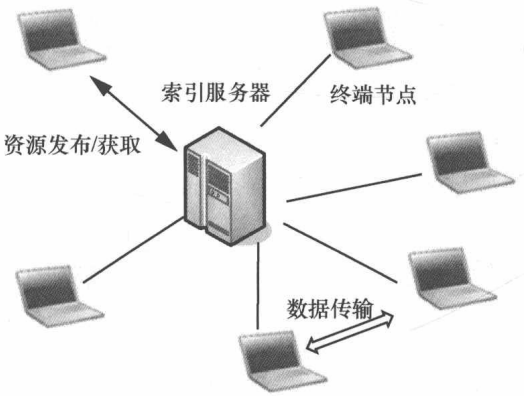


图 1-2　中心式 P2P 拓扑结构

（2）无结构 P2P 网络。为了消除中心组件所带来的系统瓶颈，随后出现的 Gnutella[5]系统去掉了中心索引服务器，将资源文件的索引信息分散到所有网络节点上。其网络拓扑如图 1-3 所示，在这一拓扑结构中，网络中所有节点在功能和访问权限上是完全对等的，当一个节点需要查询所需要的资源时，会向所有的邻居节点发送查询请求，收到请求的节点会以泛洪[13]、BFS[14]或随机漫步[15]的方式将消息向下转发。为了限制消息的转发次数，在每个查询消息中携带了 *TTL* 参数，当 *TTL* 减为 0 时，查询停止，用户会基于每次的查询结果选择合适的文件进行下载。由于网络节点在组织上的自由性和随意性，这种拓扑结构被称为无结构 P2P

网络。虽然消除了网络中的中心组件，提高了系统的顽健性，但由于网络中节点仅知道自己邻居节点的信息，只能采取盲目扩散的搜索方式，速度慢且对网络带宽消耗很大，造成了网络扩展性差，查询结果不完全，无法提供服务性能保障。这些都是造成无结构 P2P 系统不适合商业化的原因，目前，典型的无结构 P2P 系统的实现有 Gnutella 和 Freenet[16]。

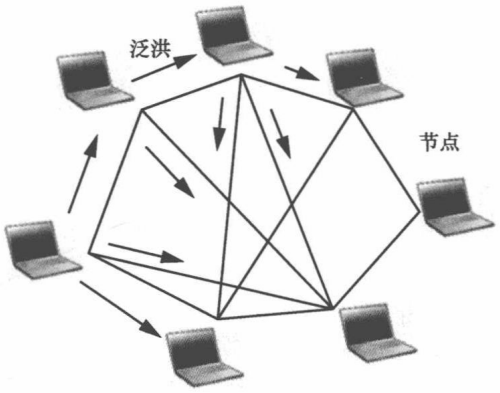


图 1-3　无结构 P2P 拓扑结构

（3）结构化 P2P 网络。为了消除无结构 P2P 网络搜索盲目性的缺点，提高资源定位的准确性，分布式散列表（DHT，Distributed Hash Table）技术被应用到了 P2P 网络的拓扑设计中。这种网络具有相对稳定和规则的拓扑结构，如图 1-4 所示。DHT 是由网络中所有节点共同维护的一个巨大的散列表，每个节点被分配了一个逻辑地址来标识自己，同时全局散列表也按照逻辑地址区域分配给不同节点。每个节点只对散列值落入自己负责区域的资源提供存储和查询服务，搜索过程也按照相应的路由算法在整个网络中进行。在 DHT 网络中，节点的散列值就像

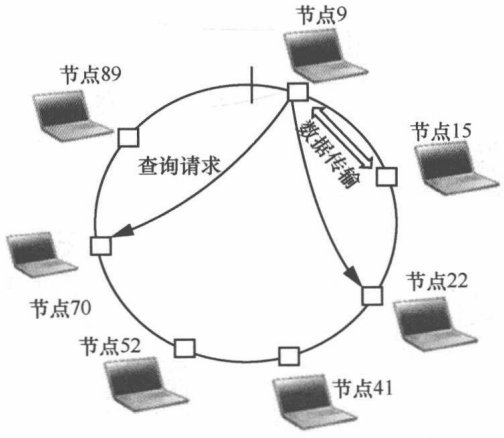


图 1-4　结构化 P2P 拓扑结构

街道和门牌号一样，将不相关的节点组织在一起，使在网络中的任何操作都有址可查。因此，这种基于 DHT 的 P2P 网络也被称作结构化 P2P。结构化 P2P 网络能够保证查询结果在一定跳数内收敛，可以自适应地维护节点进出和均衡节点负载，具有良好的可扩展性和顽健性。但是，由于资源和节点间的精确映射关系使 DHT 结构只适用于精确资源对象的定位和查找，而不适用于一般性的语义查询，这也阻碍了结构化 P2P 的大规模商业使用。目前，研究领域提出的有代表性的 DHT 结构有 CAN（Content-Addressable Network）[17]、Chord[18]、Pastry[19]、Tapestry[20]、P-Grid[21]和 Kademlia[7]等。

（4）混合式 P2P 网络。这是结合中心式拓扑与分布式拓扑特点的一种折中的选择，通过在完全分布式的 P2P 网络中引入超级节点 （Supernode）来实现的。网络拓扑如图 1-5 所示，节点分为叶子节点和超级节点，其中，叶子节点同超级节点之间通过星形结构相连，而超级节点之间则采用无结构的随机组织方式相连接。超级节点负责存储所辖叶子节点的索引信息，处理和转发叶子节点的查询请求。在这种结构中，节点之间不再是完全对等的实体，超级节点负责了大部分的工作，大大降低了叶子节点的负载，使网络流量负荷大幅减小，查询速度加快。但是，混合式 P2P 网络的健壮性一定程度上依赖于超级节点，当超级节点受到攻击时，网络可能部分瘫痪，但与中心式拓扑相比，这种威胁的强度已经被大大降低了。当前，大多数 P2P 软件都开始采用这种混合式拓扑结构，包括 KaZaA、Skype 和 eMule 等。

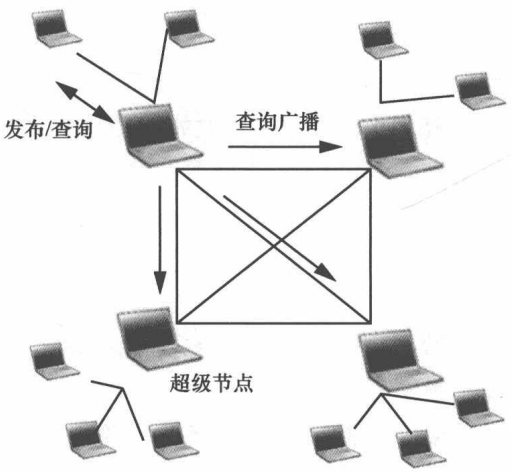


图 1-5　混合式 P2P 拓扑结构

表 1-1 是对中心式 P2P、无结构 P2P 以及结构化 P2P 优缺点的对比。如何利用不同拓扑的优点，回避其缺点是 P2P 网络拓扑结构发展的主线。而 P2P 网络也正是这样经历了一个从绝对集中到绝对分散再到混合结构的演变过程。

表 1-1　不同拓扑结构的优缺点对比

类型	优点	缺点
中心式 P2P	可以进行模糊查询 终端节点消耗低 网络节点易于管理	服务可持续性差 服务器成本高 服务器有可信性问题
无结构 P2P	全分布式网络 不存在单点失效问题 支持复杂查询 受节点频繁进出影响小	泛洪式搜索限制了可扩展性 通信代价非常高 网络可控性差
结构化 P2P	全分布式结构 可扩展性好 网络同构性高 网络节点可寻址	不支持模糊查询 网络可控性差

1.1.3　P2P 系统的特点

P2P 网络从字面上可以理解为参与到网络中的所有主体都是平等的，具体到某一通信过程，就是指交互双方为了达到某种目的而进行的双向的、直接的信息和服务交换，所有个体既可以是服务的提供者，也可以是服务的索取者[3]。从网络模型上来看，节点地位对等这一 P2P 网络模型的基础思想也是与 Internet 设计的初衷一致的，在最基本的 TCP/IP 协议中，并没有明确的客户端与服务器的概念，可以说 P2P 并不是一个新的概念，而是互联网整体架构的基础。但是，由于早期的互联网受限于计算机性能和网络资源等因素，大多数互联网节点不具备服务提供能力，于是逐步形成了以服务器节点为中心，连接众多客户端的 C/S 架构。近年来，随着主机计算存储能力和网络性能的不断提升，以及网络用户对更直接、更广泛和更自由的信息交流产生需求，P2P 架构重新成为了人们关注的焦点。P2P 的概念也得到了广阔的延伸，互联网的计算和存储模式正在由集中式向分布式偏移，由中心服务器向网络边缘的终端设备扩散。无论是服务器、个人电脑还是移动终端设备都能直接加入到网络的构建当中，极大地推动了 Internet 的发展。

与传统的 C/S 架构相比，P2P 架构具备了如下特点。

（1）对等性。在 P2P 网络中的所有节点都身兼服务器和客户端两个职责，对等连接和直接交换资源，不需要集中式服务器的参与。这有效地解决了传统网络的单点瓶颈问题，同时也提高了网络的可扩展性和顽健性。

（2）可扩展性。由于网络中每个节点都是服务的提供者，所以随着系统中节点数目的增加，系统的服务能力也不断增强，始终能够满足不断增加的用户需求。从理论上讲，P2P 系统就像一个无限大的信息仓库，其扩展性几乎可以认为是无限的。

（3）顽健性。由于 P2P 网络不存在对集中式服务器的依赖，所以不存在由单

一节点性能引起的系统瓶颈。P2P 网络中的数据和资源分散在所有网络节点上，具有天生的高容错的优点。由于 P2P 网络具有自组织的能力，网络始终能够保持高连通性，即便少数节点失效，对整个系统的影响也是微乎其微的。在节点不断动态加入和离开网络的同时，系统还能够保证极高的性能。

（4）高效性。随着硬件水平的不断提高，个人计算机的计算存储能力和网络带宽也在不断提高。P2P 网络充分利用了个人计算机闲置资源，以更低的成本消耗换取了更高的计算和存储能力。

表 1-2 是 P2P 结构和 C/S 结构多角度的对比，可以看出 P2P 结构无论是在成本、服务能力还是在网络扩展性方面都具有很大优势。

表 1-2　P2P 与 C/S 网络架构对比

性能	C/S 结构	P2P 结构
网络基础设施成本	高	低
系统容错能力	低	高
终端设备是否参与网络服务	否	是
系统负载分布	集中	分散
系统可扩展性	差	非常好（无限扩展）
网络可管理性	高	低
服务质量可控性	高	低
安全性	高	低

1.2　常见 P2P 系统简介

1.2.1　BitTorrent

BitTorrent（简称 BT，俗称 BT 下载）是一种内容分发协议，是 2002 年由布拉姆·科恩自主开发并发布在 CodeCon 上的，原始的 BitTorrent 以 Python 语言编成，3.2 版以前的源代码基于 MIT 许可证发布；4.x 版与 5.x 版的源代码基于 BitTorrent 开放源代码许可证（一个 Jabber 开放源代码许可证的改作版）发布；5.3 版被重新以 GPL 发布；自 6.0 版起，BitTorrent 成为μTorrent 的改名版，因此不再开放源代码，并仅支持 Windows 与 Mac OS X 10.5.x 版。如今，经过了十几年的发展，BitTorrent 成为了互联网领域不可或缺的重要工具，亦成为数百万人日常生活的一部分，它伴随了整整一代人的成长。每天都有大量的数据通过 BitTorrent 传输，而且这个数字仍在增加。

BitTorrent 采用高效的软件分发系统和点对点技术共享大体积文件（如一部电影或电视节目），并使每个用户像网络重新分配节点那样提供上传服务。一般的下载服务器为每一个发出下载请求的用户提供下载服务，而 BitTorrent 的工作方式与之不同。分配器或文件的持有者将文件发送给其中一名用户，再由这名用户转发给其他用户，用户之间相互转发自己所拥有的文件部分，直到每个用户的下载都全部完成。这种方法可以使下载服务器同时处理多个大体积文件的下载请求，而无须占用大量带宽。

以下是 BitTorrent 协议中重要的名词定义和算法介绍。

种子文件（Torrent 文件）。BitTorrent 是通过一个扩展名为.torrent 的种子文件进行下载部署的，它由文件最初发布者创建，发布到互联网上，供感兴趣的用户下载。种子文件记录了负责管理该文件所在分发网络的 Tracker 服务器的地址、文件名、文件长度以及每个文件分块的 SHA-1 校验值。

种子节点（Seed 节点）。Seed 节点是指在一个 P2P 共享下载网络中，拥有完整文件拷贝的节点。这类节点只提供上传服务，而没有下载请求。

下载节点（Leecher 节点）。共享网络中相对于 Seed 节点的是 Leecher 节点，它只拥有部分的文件拷贝，在提供这部分拷贝的同时，还会向其他节点请求自己缺少的那部分文件。

跟踪服务器（Tracker 服务器）。Tracker 是一个中心服务器，负责跟踪系统中所有的参与节点，收集和统计节点状态，帮助参与节点互相发现，维护共享网络中文件的下载。一个 Tracker 服务器可以同时维护和管理多个文件共享网络。

共享网络（Swarm 网络）。一个 Swarm 共享网络是拥有和传输同一个文件资源的所有节点所构成的一个覆盖网络，包括共享该文件的 Seed 节点、Leecher 节点和 Tracker 服务器。

分片机制。BitTorrent 像其他文件共享软件一样对文件进行了分片（Piece），Piece 是最小的文件共享单位，每个 Leecher 在下载完一个完整的分片后才会进行完整性校验，完整性校验成功后通知其他节点自己拥有这部分数据。为了加快文件传输的并行性，每个分片还会分成更小的分块（Block），Block 是最小的文件传输单位，数据请求者每次向数据提供者请求一个 Block 的数据。

片选择机制。为了保证共享网络的健壮性，延长一个共享网络的生命周期，BitTorrent 通过局部最少块优先（Rarest-First）策略在节点间交换数据。下载节点根据自己周围的邻居节点拥有的数据块信息，选择拥有节点最少的分块优先下载，从而维护局部的数据块相对平衡。

节点选择机制。BT 系统采用了基于“Tit-for-Tat”的激励机制来抵御“Free-riding”行为，其中 Choking/Unchoking 算法最为关键。每个 BT 节点通过 Interest/Uninterest 消息来维护与多个节点的并发连接，但是只能为少数节点提供上传。服务提供节点

在收到上传请求后会通过 Choking/Unchoking 机制决定是否对文件请求节点提供上传服务，可以拒绝服务（Choking）或者允许服务（Unchoking）。该机制决定了两个相连的节点是否共享彼此的资源。为了防止部分节点只下载不上传的自私行为，Choking/Unchoking 算法优先选择曾经为自己提供过上传数据并拥有高下载速率的节点，前者可以鼓励节点上传以获取下载，后者有助于最大化系统资源利用率。此外，Choking/Unchoking 算法每隔 30 s 将不考虑过去的贡献随机选择一个节点进行上传，一方面有利于发现可能存在更高下载速率的节点，另一方面可以避免新节点因从未进行过上传而无法获得有效的下载连接。

图 1-6 给出了 BT 系统下一个节点的生命周期。首先，一个新节点通过论坛或网站获得 Torrent 文件。随后该节点通过 Torrent 中的 URL 链接与对应的 Tracker 服务器通信，获得由 Tracker 提供的一个包含 30~80 个节点的随机邻居列表。成功加入 Swarm 网后，该节点成为一个新的 Leecher 节点，并与其邻居节点建立 TCP 连接，请求数据分片。当拥有了有效数据分片后，该 Leecher 节点会通过 Choking/Unchoking 机制同其他节点交易数据。当拥有了所有数据分片后，Leecher 节点成为 Seed 节点，为其他节点提供数据。在整个生命周期过程中节点会定期与 Tracker 服务器通信获得活跃的邻居节点，并更新自身的下载状态。

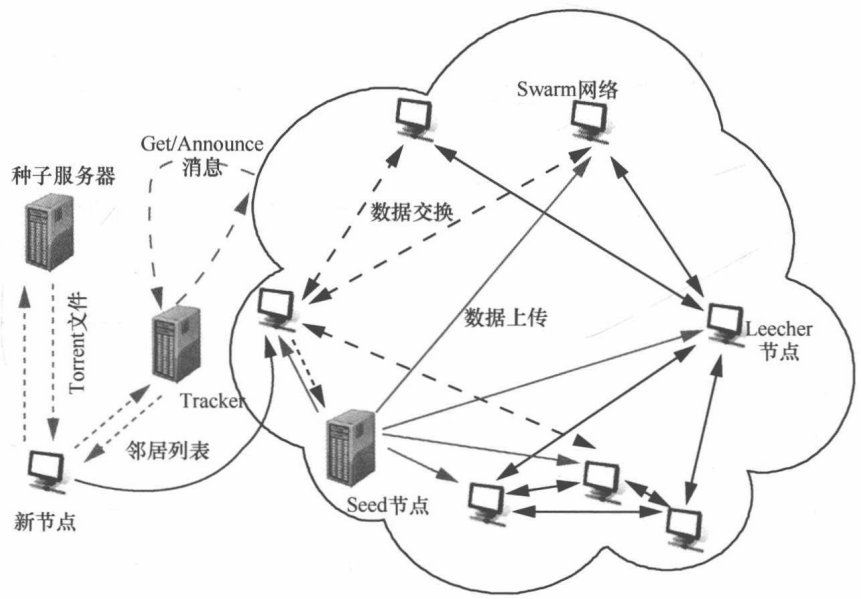


图 1-6　BT 网络中节点的生命周期

随着 BT 的发展，相继产生了基于结构化 P2P 的 Mainline DHT 协议，以及基于节点间邻居节点交换的 PEX 协议。Mainline DHT 为 BT 共享网络提供了结构化拓扑，是减轻 Tracker 负担的一种扩展协议，其协议大致基于 Kademlia 协议开发。