

Peer To Peer 计算技术的应用领域与典型问题讨论^{*}

王庆波¹ 代亚非² 田 敬²

(哈尔滨工业大学计算机科学与技术学院 哈尔滨150001)¹

(北京大学计算机科学与技术系 北京100871)²

The Applied Fields and Typical Problems in Peer To Peer Computing

WANG Qing-Bo¹ DAI Ya-Fei² TIAN Jing²

(School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001)¹

(Department of Computer Science and Technology, Peking University, Beijing 100871)²

Abstract With the rapid development of computer networks there are more and more resources involved in it, such as CPU cycles, PC's disk space and network bandwidth. Many people believe that peer to peer computing can harness these disruptive powers. In this paper we describe the applied fields, research issues and typical problems in peer to peer computing.

Keywords Peer To Peer, Distributed computing, Pervasive computing

1 引言

网络技术的飞速发展及迅速普及使其成为数据通信的重要手段,网络的发展大大超出了网络的提出者以及早期的建立者的构想。网络规模越来越大,联入网络中的设备以及计算单元的数量和种类也越来越多,然而这些设备以及计算单元并没有得到充分的利用,如果能够将这些设备及计算单元的处理能力、磁盘存储能力以及网络带宽资源等进行充分利用将会有效缓解目前互联网所面临的一些问题。

P2P(Peer To Peer)计算技术的出现目的就是希望能够充分利用互联网中所蕴含的潜在计算资源。P2P 中文称为对等网络,是指分布式系统中的各个节点是逻辑对等的,与目前互联网上比较流行的 C/S 计算模型不同的是:P2P 计算模型中不再区别服务器以及客户端,系统中的各个节点之间可以直接进行数据通信而不需要通过中间的服务器。

随着像 Napster^[1]等 P2P 信息共享应用程序的流行,P2P 计算技术受到人们的广泛关注。由于该计算技术所具有的一些技术特点,很多著名的公司、研究部门都认为 P2P 计算模型蕴含着巨大的商业和技术潜在价值,并从不同的角度应用和研究这种技术。文中接下来分别对 P2P 计算技术的应用领域以及研究 P2P 计算技术所面临的一些典型问题进行了讨论。

2 P2P 计算技术的应用领域

目前人们从很多不同的角度来应用 P2P 计算技术,主要应用的角度包括:信息资源共享、普及计算、协同工作、实时通信技术、信息检索技术、广域网络存储系统等。下面将从上述这几个角度对 P2P 计算技术的应用领域分别进行介绍。

2.1 信息资源共享

信息资源共享一直是网络技术发展的重要推动力,也是 P2P 技术中最典型的应用。目前人们主要采用 Web 技术来实现信息资源共享,在基于 Web 的方式进行信息资源共享时,

Web Server 需要能够对大量用户的访问提供有效的服务,Web Server 经常成为这类系统的性能瓶颈所在。Napster 是提供给用户在互联网上共享 Mp3 音乐文件的 P2P 应用,与传统的音乐共享技术不同的是 Napster 把音乐文件存储在客户节点上而不是存储在服务器节点上,中心服务器上存储的仅仅是文件的索引信息,用户之间可以直接共享、传输音乐文件而不需要通过中心索引服务器。采用这种方式来共享信息资源可以更加充分地利用网络中的带宽资源,从而提高了系统数据通信的效率。目前有很多研究项目都是针对 P2P 的文件共享的,包括 Freenet^[2]、Gnutella^[3]、Free Haven^[4]、Ohaha^[5]等,这些研究项目均从不同的角度尝试解决目前网络中的信息资源共享所存在的一些问题。

2.2 普及计算

普及计算技术研究的是如何充分利用网络中各种各样的计算单元来共同完成大规模的计算任务。由于单一计算单元的计算能力总是有限的,因此人们一般采用并行技术、分布式技术将多个计算单元节点联合起来共同完成大规模的计算任务,同时目前网络中的计算机的计算能力一直利用得不是很充分,人们期望能够充分利用网络中的闲散计算能力来完成大规模的计算任务,这样将会使得网络中所蕴含的海量计算能力得到更加充分的利用。P2P 计算技术则为普及计算技术的发展提供了新的机遇。

SETI@home^[6]是 Berkeley 大学启动的普及计算的研究项目,目前大约吸引了一百万台计算机参与研究。该项目是利用该大学的空间科学实验室开发的屏幕保护程序来使用计算机的空闲机时,该屏幕保护程序在运行时分析在外星系文明研究项目中所获得的无线电信号,程序运行节点从中心服务器节点下载数据后进行计算然后再将计算结果上传到该实验室的中心服务器上,因为不是完全的 P2P 计算模式,所以节点之间不能直接利用彼此计算的数据。

普及计算可以帮助企业完成大规模的数据处理,参与计算的计算机之间可以直接共享计算中的中间结果。通过整合

^{*} 本文得到国家973高新技术发展计划(G1999032708)资助。相关研究工作均在北京大学计算机系网络与分布式实验室完成。

这些以前尚未使用的闲散计算能力和资源可以将企业的计算能力相比以前得到很大的提升,同时因为利用了多个节点上的计算能力使得计算任务可以高效廉价地完成。网格计算 GRID^[7]是研究普及计算的典型代表,IBM 公司也启动了自组织计算计划^[8]来研究普及计算。

2.3 协同工作

协同工作是指多个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,计算机支持的系统工作 CSCW 是协同工作的典型研究方向之一。通过采用 P2P 计算技术个人和组织可以随时采用各种方式建立在线、非在线的协同应用环境。例如目前许多企业都使用协同工具来完成公司内部员工之间的通信,但是传统的协同应用平台一般都是采用电子邮件方式来完成工作人员之间的协同工作,而并不能很好地完成企业与合作伙伴、客户、供应商之间的沟通与交流。下一代的 P2P 的解决方案试图要很好地解决这些问题,通过使用各种形式的个人对个人,不存在中心服务器的协同工具来提供给企业一个包含项目管理等功能的协同应用平台。协同工作使得在不同地点的参与者可以在一起工作,因为采用文件直接共享的方式可以保证系统中的每个人所获得的信息总是最新的,同时节省了采用单点服务器时对该服务器存储以及性能的要求。

一般的协同应用包括:实时通信、聊天室、好友列表、文件共享、语音通讯等基本功能,除了这些基本的功能,用户之间还可以共享白板,协同写作进行视频会议等。由于协同应用的用户数量一般都比较大,数据量也比较多,采用传统的单一中心节点服务方式很难满足这种应用,如果采用 P2P 的计算方式可以不再需要目前协同工作中的中心服务器,参与协同工作的两台计算机可以直接建立联系进行协同工作。Groove^[9]是基于 Internet 的 P2P 协同应用软件的典型代表,其用户可以直接进行实时的协同工作。

2.4 实时通信技术

实时通信技术是网络中重要的通信技术,成功的实时通信技术吸引了数以万计的在线用户。目前的实时通信技术一般也采用一个中心服务器控制着用户的认证等基本信息,节点之间直接进行数据通信。ICQ、OICQ、AIM 等是典型的实时通信系统,这些系统也包含好友列表等基本功能。Jabber^[10]是一个开放源码的实时通信平台,Jabber 提出了一个采用 XML 表示的在不兼容的各种实时通信平台之间进行消息交换的协议。

2.5 信息检索技术

搜索引擎是目前人们在网络中检索信息资源的主要工具,目前的搜索引擎如:天网^[11]、Google^[12]等都是集中式地搜索引擎,同时这类搜索引擎所搜索的主要媒体信息是超文本信息,并没有对多媒体信息进行很好地支持。由于目前网络的信息数量成指数增长,搜索引擎的增长速度已经不能有效地跟踪数据总量的增长,这种信息搜索方式是一种被动的搜索方式,用户不能主动选择将自己的信息发布到搜索引擎上,也不能保持搜索引擎所采集数据的实时性,有时用户并不想将自己的信息进行网页快照,这些在信息的版权等问题上都面临巨大的挑战。在 JXTA Search^[13]中认为采用 P2P 的搜索技术可以有效地跟踪数据的更新、提高访问的有效性以及检索的效率。

2.6 广域网络存储系统

存储技术一直是人们所关注的一项技术,SAN、NAS 是

目前在广泛应用于局域网络的存储技术。分布式文件系统也是广泛使用的分布式文件存储技术,典型的分布式文件包括 NFS、AFS、Coda 等。由于网络规模的扩大,人们对网络的使用也变得十分灵活,人们开始将传统的分布式操作系统、局域存储技术向基于 Internet 的文件存储系统发展。一些研究项目开始使用 P2P 技术来组织和存储文件,典型的系统包括:Oceanstore^[14]、Farsite^[21]等,这些项目的目标都是提供面向全球规模的文件存储服务。

2.7 P2P 计算技术的其它应用领域

除了前面介绍的 P2P 计算技术的几个典型的应用领域,许多公司也从其它不同角度来应用该技术。Net 技术是微软公司提出并正在开发的一个基于 Internet 的操作系统,该技术是围绕着以 SOAP XML 通信协议的 Web 服务为主。JXTA^[15]是 SUN 公司提出的一个 P2P 的网络底层支撑平台,该平台的目标是允许用户在其上开发各种 P2P 应用。

3 P2P 计算技术所面临的典型问题

P2P 计算技术可以归结为一种特殊的分布式计算技术,从而 P2P 计算技术所面临的很多问题都可以利用目前的分布式计算技术的研究成果来解决,如并发控制、事务处理等基本分布式系统所面临的问题。但是很多问题在 P2P 系统中具有其自身的特点,目前大家一般从下面的几个角度来研究和解决 P2P 中所面临的问题。

3.1 P2P 网络拓扑结构的研究

拓扑结构是指分布式系统中各个计算单元之间的物理或逻辑的互联关系,节点之间的拓扑结构一直是确定系统类型的重要依据,目前互联网络中广泛使用集中式、层次式等拓扑结构,Internet 本身是世界上最大的非集中式的互联网络,但是 20 世纪 90 年代所建立的一些网络应用系统却是完全的集中式系统,很多 Web 应用都是运行在集中式的服务器系统上。集中式拓扑结构系统目前面临着过量存储负载、Dos 攻击等一些难以解决的问题。层次式拓扑结构是一种应用比较广泛的分布式拓扑结构,DNS 系统是其最典型的应用。P2P 系统一般要构造一个非集中式的拓扑结构,在构造过程中需要解决系统中所包含的大量节点如何命名、组织以及确定节点的加入、离开方式、出错恢复等问题。CAN^[16]、Pastry^[17]、Chord^[18]、Tapestry^[19]、Grid^[20]等都提出了自己的 P2P 网络拓扑结构模型。

3.2 数据索引、查找、定位、路由机制以及访问路径

在典型的 P2P 网络中数据资源分布在各个独立的节点上,如何高效地索引、查找、定位以及访问这些数据信息资源是另一个需要关注的重要问题,在分布式系统中这些问题同样也是正在研究的热点问题。URL 是目前在 Web 上使用最普遍的信息定位策略,DNS 则提供了一套层次式的查找机制,一般来说在 P2P 共享应用中所采用的检索方式是采用关键字来查询自己所需的信息资源,同时人们也期望能够将数据源的索引信息存放在系统中的每一个节点上而不是像 Napster 那样存储在中心服务器上。路由机制是指节点之间通信的消息传递路径,合适的路由机制可以充分地利用网络带宽资源并使系统具有很好的容错性、可扩展性,目前很多系统中的路由机制都是和这些系统的逻辑拓扑结构紧密相关的。在数据的访问过程中则期望能够采用流水、并行或者选择传输路径的方式来加快数据的访问速度。

(下转第 45 页)

- 17 Floyd S, Handley M, Padhye J. A Comparison of Equation-Based and AIMD Congestion Control. URL: <http://www.aciri.org/tfrc/>.
- 18 Bansal D, Balakrishna H. Binomial Congestion Control Algorithms. In: Proc. IEEE INFOCOM, April 2001
- 19 Golestani J, Sabnani K. Fundamental Observations on Multicast Congestion Control in the Internet. In: Proc. IEEE INFOCOM, 1999
- 20 Wu D, Hou Y T, Zhang Y. Transporting Real-Time Video over the Internet: Challenges and Approaches. In: Proc. of the IEEE. VOL. 88, NO. 12, Dec. 2000
- 21 Alanese A, Blomer J, Edmonds J, Luby M. Priority Encoding Transmission: [Tech. Rep. RT-94-039]. Department of Computer Science, University of California, Berkeley, 1994

- 22 Tan W, Zakhov A. Multicast Transmission of Scalable Video Using Receiver-driven Hierarchical FEC. In: Proc. Packet Video, New York, April 1999
- 23 Davis G, Danskin J. Joint Source and Channel Coding for Internet Image Transmission. In: Proc. SPIE Conf. Wavelet Applications of Digital Image Processing XIX, Denver, CO, Aug. 1996
- 24 Li X, Paul S, Pancha P, Armbrast M. Layered Video Multicast with Retransmissions (LVMR): Evaluation of Error Recovery Schemes. In: Proc. IEEE INFOCOM, 1998
- 25 Xu X R, Myers A C, Zhang H, Yavatkar R. Resilient Multicast Support for Continuous-media Applications. In: Proc. IEEE Int. Workshop Network and Operating System Support for Digital Audio and Video (NOSSDAV'97), May 1997, 183~194

(上接第9页)

3.3 元数据组织与表示

P2P 网络面向的是异构网络与操作系统,这样就需要在这些系统之间交换数据资源,但是因为这些系统的数据表示并不都是完全相同的,这样就需要一个能够在多个系统之间确定一个通用的元数据表示方案。关于元数据的组织包括数据资源的表示、消息通信协议等,很多系统都支持 SOAP 或者 XML-RPC 等协议。

3.4 匿名性的支持

对匿名性进行支持是很多 P2P 系统所要实现的一个重要功能,所谓的匿名性是指对系统中信息资源的操作者的操作行为的隐私进行保护,具体包括对一个信息的作者、发布者、阅读者、信息提供服务器、存储位置、检索的用户行为的隐私性进行支持。在 Freenet、Free Haven 等系统都对这项功能进行了支持。

3.5 P2P 网络的支撑技术

Internet 技术的发展使得连入互联网络中的设备不再局限于计算机,在 P2P 的计算环境中要求任何设备都可以在任何地点很容易地加入到这个环境中,所谓的计算设备既包括有线设备也包括无线设备,这样就需要很多很多网络传输的支撑技术来支持各种不同设备连入整个 P2P 网络。Bluetooth、Jini 都是目前所研究和采用的 P2P 网络支撑技术。

3.6 P2P 网络的安全问题

安全问题是一直伴随着互联网发展的重要问题,安全问题包括很多相关的问题,比如应该防止他人控制整个系统,增加恶意信息等,同时系统应能够保证系统中信息资源的正确性。在 P2P 系统中系统安全同样面临着巨大的挑战。P2P 系统需要在没有中心节点的情况下,提供身份的认证、授权以及数据信息的安全存储、数字签名、加密、安全传输等工具,同时 P2P 系统要有能力抵抗过量存储负载、Dos 攻击等攻击行为。

总结 P2P 技术的兴起使得人们从信息资源共享、普及计算、协同工作、实时通信技术、信息检索技术、广域网络存储系统等领域来应用该技术,同时为了能够充分地发挥其技术优势需要解决与该技术相关的一些典型的问题,文中对这些应用领域以及这些典型的问题分别进行了讨论。

参考文献

- 1 Napster. <http://www.Napster.com/>
- 2 Clarke I, et al. Freenet: A Distributed Anonymous Information Storage and Retrieval System. ICSI Workshop on Design Issues in Anonymity and Unobservability, July 2000
- 3 Gnutella. <http://gnutella.wego.com/>
- 4 Free Haven. <http://www.freehaven.com/>
- 5 Ohaha. <http://www.ohaha.com/>
- 6 SETI@Home. <http://setiathome.ssl.berkeley.edu/>
- 7 Foster I, Kesselman C. The Grid: Blueprint for a New Computing Infrastructure. Morgan Kaufmann Publishers, Inc. 1999
- 8 <http://www.research.ibm.com/autonomic/>
- 9 Groove. <http://www.groove.net/>
- 10 Jabber. <http://jabber.org/>
- 11 WebGather. <http://e.pku.edu.cn/>
- 12 google. <http://www.google.com/>
- 13 JXTA Search: Distributed Search for Distributed Networks. <http://www.jxta.org/>
- 14 Kubiawicz J, et al. OceanStore: An architecture for global-scale persistent storage. In: Proc. of ACM ASPLOS. ACM, Nov. 2000
- 15 JXTA. <http://www.jxta.org/>
- 16 Ratnasamy S, Francis P, Handley M, Karp R, Schenker S. A scalable content-addressable network. Submission to ACM SIGCOMM, 2001
- 17 Druschel P, Rowstron A. Pastry: Scalable, distributed object location and routing for large-scale peer-to-peer systems. Submission to ACM SIGCOMM, 2001
- 18 Stoica I, et al. Chord: A scalable peer-to-peer lookup service for internet applications. Submission to ACM SIGCOMM, 2001
- 19 Zhao B Y, Kubiawicz J, Joseph D. Tapestry: An infrastructure for fault-resilient wide-area location and routing. Submitted for publication, 2001
- 20 LI J, et al. A scalable location service for geographic ad hoc routing. In: Proc. of the 6th ACM Intl. Conf. on Mobile Computing and Networking (Boston, Massachusetts, August 2000), pp. 120~130
- 21 Bolosky W, Douceur J, Ely D, Theimer M. Feasibility of a serverless distributed file system deployed on an existing set of desktop pcs. In: Proc. of Sigmetrics, June 2000