

P2P 入网机制研究

唐九阳 李 榴 汤大权 肖卫东

(国防科技大学信息系统与管理学院 长沙 410073)

摘要 P2P 节点入网机制是构建 P2P 网络的基础。不同拓扑结构的 P2P 网络,节点加入机制也各不相同。提出了 4 种 P2P 入网模型,给出了相应的性能指标,对 P2P 入网机制进行分析、对比,有助于 P2P 网络自组织构造的进一步研究。

关键词 对等网络,网络拓扑,入网机制

Research on P2P Bootstrapping Mechanism

TANG Jiu-yang LI Liu TANG Da-quan XIAO Wei-dong

(School of Information System and Management, National University of Defense Technology, Changsha 410073, China)

Abstract In recent years, peer-to-peer (P2P) systems have evolved to accommodate growing numbers of participating peers. The availability of a P2P system is heavily affected by the P2P bootstrapping mechanism. There are some differences of bootstrapping mechanisms among P2P networks which have different topology. A survey of the recent work about the P2P bootstrapping mechanisms is made. And some performance indexes were proposed and the similarities and differences of them were compared. It is helpful to the further study of P2P network self-organization.

Keywords P2P network, Network topology, Bootstrapping

P2P 节点加入机制是构建 P2P 网络的基础,是决定 P2P 网络性能的重要指标之一。为了使网络具有可扩展性和低负载特性,新节点加入网络不应该引起大量消息传递而加重网络负载,因此新节点加入只能对网络中已存在的少量节点产生影响^[1]。目前的 P2P 入网 (bootstrapping) 机制中,节点初始加入 P2P 网络时需要了解网络中至少一个节点的地址信息,并以其为入口节点加入网络。入口节点需要向新节点提供网络中其他节点的信息,并引导新节点的加入。实际上,由于 P2P 网络拓扑的不同,导致相应的入网机制存在很大差异。本文针对现存的各种 P2P 网络拓扑,提出了 4 种 P2P 入网模型,通过给出的性能指标对相应的节点加入机制进行了分析、对比。

1 P2P 网络的分类

从节点信息资源存储的角度,P2P 网络可以划分为结构化 P2P 网络和非结构化 P2P 网络两大类^[2]。结构化 P2P 网络将每个资源精确放置在确定的节点上,提供了资源标识 ID 到资源所在位置的映射关系,从而确保在有限步数内定位到资源。结构化 P2P 网络以基于分布式哈希表 (Distributed Hash Table, DHT) 的 Tapestry^[3], Pastry^[4], CAN^[5] 和 Chord^[6] 系统为代表。非结构化 P2P 网络中资源放在每个节点上,当用户需要获取资源时,预先并不知道资源的存储位置,因而资源搜索带有一定的盲目性。

根据资源搜索时对维护索引的节点的依赖程度,非结构化 P2P 网络又可以分为集中式 (centralized)、全分布式 (decentralized) 和混合式 (hybrid) 3 种^[7]。集中式 P2P 网络 (如 Napster^[8]) 采用目录服务器作为网络中心,节点及其拥有的资源均在目录服务器中记录,资源的搜索只需在目录服务器中查找便可完成。虽然目录服务器只记录资源索引而不保存资源文件,但仍然存在服务器瓶颈与单点失效的问题。完全分布式 P2P 网络以 Gnutella^[9] 和 Freenet^[10] 为代表,节点间的连接是任意的,数据的放置对其他节点来说不可知。资源搜索过程中所采用的泛洪发现与随机转发机制使得消息总量呈指数增长,存在报文泛滥的问题,对带宽消耗严重。混合式 P2P 网络从所有节点中挑选出负载能力强的节点作为超级节点 (super-peer),组成上层网络,而能力较弱的节点处于网络下层并聚合在不同的超级节点下。下层网络采用集中式网络结构,上层网络的超级节点之间则可采用任意网络结构。典型的混合式 P2P 网络有 KaZaA^[11], Gnutella2^[12] 和 Fast-track^[13]。

P2P 网络的分类图如图 1 所示。

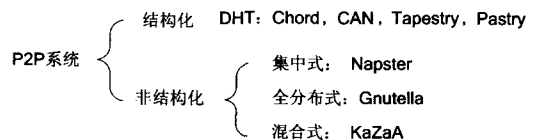


图 1 P2P 网络分类图

到稿日期:2008-07-11 本文受国家自然科学基金项目(60172012),湖南省自然科学基金项目(03JJY3110)资助。

唐九阳(1978-),男,博士,讲师,主要研究方向为对等计算、信息集成、知识管理等,E-mail:jiuyang_tang@sina.com;李榴(1986-),男,硕士生,主要研究方向为对等计算、无线传感器网络等;汤大权(1971-),男,博士,副教授,主要研究方向为信息资源管理、信息检索;肖卫东(1968-),男,博士,教授,主要研究方向为信息管理、信息系统集成、对等计算。

2 P2P 入网模型

2.1 主动模型

在完全分布式 P2P 网络中,节点之间的连接是任意的,不存在全局信息中心,当新节点准备加入网络时,它将自己的请求包通过泛洪、组播等方式,在网络中进行发送,寻找存活节点,以寻求加入网络。早期的 Gnutella 网络使用泛洪发现方式,而基于 JXTA 的 P2P 应用则使用组播方式与其他节点取得联系。主动模型如图 2 所示。

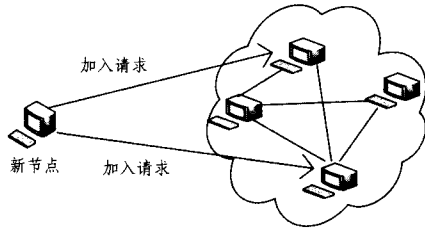


图 2 主动模型

节点加入步骤描述如下:

- 1)新节点 A 产生加入请求包 p ;
- 2)A 通过泛洪或组播方式发送 p ;
- 3)节点 B 收到请求,若 B 不是 P2P 网络中的节点,不予应答,转发 p ;
- 4)若 B 是 P2P 网络中的节点,将 A 存入〈连接节点表〉,向 A 发送应答,丢弃 p ;
- 5)A 收到 B 的应答,将 B 存入〈连接节点表〉,加入过程结束。
- 6)若 A 在限定时间内没有收到任何应答,提高 p 的 TTL,继续发送 p ,直至获得应答。

这种模型的优点是动态性很强,自动化程度很高。每个节点都可自己决定如何处理请求并做出相应的反应。其不足在于当网络中的节点很多时,由于节点请求泛洪的原因,对带宽的要求是指数级增加的,网络带宽会迅速达到饱和,大大降低了效率。

2.2 目录模型

在集中式 P2P 网络中,新节点直接向网络的中心服务器发出请求,登记自己的 IP、信誉度等信息,寻求加入 P2P 网络。目录服务器作为网络中心,通过保存所有节点的资源索引信息来实现资源的搜索。图 3 是一个目录模型,其典型应用有 Napster, Freenet, BitTorrent^[14] 等。

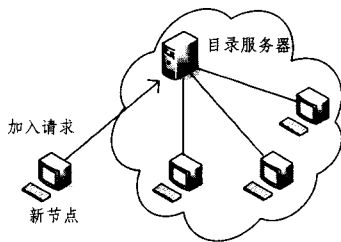


图 3 目录模型

节点加入步骤描述如下:

- 1)新节点 A 产生加入请求包 p ;
- 2)A 向目录服务器 S 发送 p ;
- 3)S 收到请求 p ,向 A 发送应答,丢弃 p ;

- 4)A 收到 S 的应答,将自身资源的索引列表发送给 S;
- 5)S 收到 A 的资源索引列表,存入〈资源索引表〉,返回应答;

- 6)A 收到应答,加入过程结束。

在这种模式中,可以便捷发现其他节点并保持数据的更新,但也容易导致网络崩溃。例如,目录服务器关闭后,注册到服务器上的节点将会失去和其他所有节点的连接。而且,当节点数目急剧增加时,目录服务器的带宽与硬件能力将受到挑战,成为网络的瓶颈。

2.3 WellKnow 模型

Chord, CAN, Kad 等基于 DHT 的结构化 P2P 网络,使用 WellKnow 机制完成节点的入网。即 P2P 网络中存在一个众所周知的节点 (WellKnow 节点),新加入的节点首先向该节点发出请求,该节点根据新节点的哈希值,将其散列到网络空间中,新节点以 WellKnow 节点为中介,加入 P2P 网络。WellKnow 模型如图 4 所示。

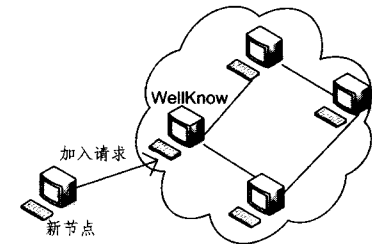


图 4 WellKnow 模型

节点加入步骤描述如下:

- 1)新节点 A 产生加入请求包 p ,其中包含 A 的 IP 地址的 HashID;
- 2)A 向 WellKnow 节点 W 发送 p ;
- 3)W 收到请求 p ,按照相应规则使用 A 的 HashID 将 A 添加到 P2P 网络中,向 A 发送应答,其中包括 A 在网络中的新位置的路由信息;
- 4)A 收到应答,将 W 提供的 P2P 网络路由信息进行存储,并使用该信息通知网络拓扑中的邻居节点,加入过程结束。

在这种模式中,WellKnow 节点作为 P2P 网络的核心,承担着所有新节点的加入服务,一旦 WellKnow 节点瘫痪,新节点将无法与其获得连接,继而无法加入 P2P 网络,可通过提供多个 WellKnow 节点解决单点失效问题。

2.4 WebCache 模型

Gnutella2, KaZaA 等非结构化 P2P 网络,均使用 WebCache 机制。即在节点之外,另外设立一台称为 WebCache 的服务器,用于登记网络中的活跃节点的信息。WebCache 收到新节点发送的请求后,将活跃节点的列表返回给新节点。新节点获得该列表后,将对列表中的所有节点进行连接,以加入 P2P 网络。WebCache 模型如图 5 所示。

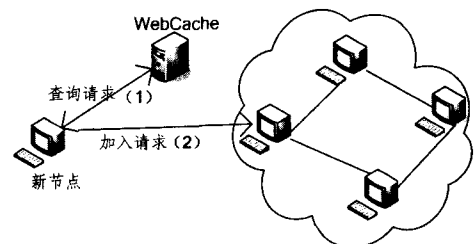


图 5 WebCache 模型

节点加入步骤描述如下:

1) 新节点 A 产生查询请求 q ;

2) A 向 WebCache 服务器 W 发送 q ;

3) W 收到请求 q , 从已登记的活跃节点列表中, 挑选出能力最强的几个活跃节点, 将它们的 IP 地址应答给 A;

4) A 收到活跃节点列表后, 产生加入请求包 p , 向已知的活跃节点发送 p ;

5) 活跃节点 B 收到 p , 将 A 存入〈连接节点表〉, 向 A 发送应答;

6) A 收到 B 的应答, 将 B 存入〈连接节点表〉, 加入过程结束。

在该模型中, WebCache 作为 P2P 网络中的独立单元, 承担新节点加入引导的服务, 并通过对活跃节点列表的动态管理, 调整节点的连接度, 有效地实现网络负载均衡。

3 比较分析

下面从负载均衡 (Load balance)、网络开销 (Network overhead)、公平性 (Fairness)、容错性 (Fault tolerance) 和扩展性 (Scalability) 5 个方面对上述 4 种 P2P 入网模型进行对比分析。

3.1 负载均衡

新节点加入过程中, 以全网节点承担工作量的均衡性来进行负载均衡的判别。主动模型中, 离新节点距离最近的节点承担的工作量最大。随着数据包的不转发, 节点负载以新节点为圆心向外辐射越来越小。目录模型与 WellKnow 模型中, 目录服务器与 WellKnow 节点分别承担了新节点加入的所有工作, 网络中的其他节点并不参与, 唯一的区别是当有多个目录服务器存在时, 目录服务器之间可实现负载均衡, 而 WellKnow 模型无法控制 WellKnow 节点负载。WebCache 模型中由 WebCache 进行节点入口选择的控制, 从而实现活跃节点之间的负载均衡。

3.2 网络开销

设请求数据包为 p , 答复数据包为 r , 下面针对 4 种入网模型, 对新节点加入过程中数据包发送的数量进行网络开销的判别。

主动模型采用泛洪机制或组播机制实现节点入口的发现, 网络开销用公式表示为 $Cost_{Positive} = p \cdot n^{TTL} - \sum_{i=0}^N p \cdot n^{TTL-x_i}$, 其中 n 为泛洪基数, TTL 为请求数据包的存活时间, N 是所有入口节点的个数, x_i 为发现入口节点 $node_i$ 时已经走过的跳数。由此可见, 请求数据包的数量在算法较差 (即 n 较大) 与网络结构复杂 (即 x_i 较大) 的情况下会呈几何级数增长, 网络开销最大。

目录模型的网络开销用公式表示为 $Cost_{Directory} = (p+r) \cdot n_{DS}$, 其中 n_{DS} 表示目录服务器的数目, Napster 和 Freenet 系统中, 一般 $n_{DS} = 1$, 因此目录模型在 4 种入网模型中网络开销最小。

WellKnow 模型中, 新节点需要与 WellKnow 节点及邻居节点进行数据交互, 用公式表示为 $Cost_{WellKnow} = p_w + r_w + (p+r) \cdot n_{NeighborPeer}$, 其中 $n_{NeighborPeer}$ 为邻居节点的数目, 以 Chord 网络为例, $0 \leq n_{NeighborPeer} \leq 2$, 比目录模型的网络开销略大。

WebCache 模型中由于新节点需要与 WebCache 及活跃节点交互, 网络开销用公式表示为 $Cost_{WebCache} = [p_w + r_w +$

$(p+r) \cdot n_{ActivePeer}] \cdot n_{WebCache}$, 其中 $n_{WebCache}$ 为收到新节点请求的 WebCache 的数目, $n_{ActivePeer}$ 为某一 WebCache 向新节点返回的活跃节点的数目。对于大规模应用的比较成熟的基于 WebCache 模型的对等网络 (如 Gnutella, KaZaA), $n_{WebCache} \geq 10$, $n_{ActivePeer} \geq 5$, 网络开销较大。

根据以上分析可以得出 4 种入网模型的网络开销具有以下关系:

$$Cost_{Positive} \geq Cost_{WebCache} \geq Cost_{WellKnow} \geq Cost_{Directory}。$$

3.3 公平性

新节点加入行为中, 以全网节点参与的机率来进行公平性的判别。在主动模型中, 由于新节点出现位置的随机性, 网络中的所有节点都有可能参与新节点的加入行为, 即每一个节点成为加入入口的机率相同, 向新节点做出应答最快的节点最终成为加入入口。在目录模型中, 除目录服务器外的其他网络节点并不参与新节点加入行为, 目录服务器成为所有新节点的加入入口。WellKnow 模型为避免出现目录模型中低公平性的问题, 允许用户节点保存多个节点信息, 在入网时从保存的多个节点信息中选择应答最快的节点作为 WellKnow 节点。WebCache 模型中, 首先使用在网时间、健壮程度等指标从全网节点里筛选出活跃节点, 登记在 WebCache 中, 接着 WebCache 会对活跃节点进行排序, 选择能力最强的节点作为新节点的加入入口, 即能力越强的节点, 成为加入入口的机率越高。

3.4 容错性

新节点加入行为中, 以发生错误对入网的影响来进行容错性的判别。在主动模型中, 某一节点的失效导致该节点无法对请求进行应答或转发, 但由于泛洪与组播对节点的覆盖能力广, 因此单个节点的失效并不影响新节点的加入。在目录模型中, 目录服务器的失效会导致所有新节点无法成功加入网络, 此时考虑设置多台目录服务器以提高容错性。在 WellKnow 模型中, 用户可从自身保存的节点列表中选择 WellKnow 节点, 只有列表中所有节点全部失效的情况下才会出现新节点加入失败。WebCache 模型使用 WebCache 能够筛选出失效节点, 且用户可多次请求 WebCache 以防止提供的活跃节点失效。

3.5 扩展性

从网络节点的增加对新节点入网影响的角度来进行扩展性的判别。主动模型中, 节点数目的增加能够提高加入入口发现的机率, 这是以延长转发时间与搜索次数为代价实现的。在目录模型中, 目录服务器的增加只能有限地提高新节点加入行为的效率。WellKnow 模型中, 由于每一个节点均具有加入入口功能, 因此 WellKnow 节点的选择与网络节点数量无关, 网络节点数量的增多只能保证用户自身保存的节点列表的数量不会低于限定值。在 WebCache 模型中, 由于加入入口是由活跃节点能力所决定, 当网络节点数量增加时, 节点能力的比较结果越优, 很大程度上提高新节点加入效率。

从以上 5 个方面的对比分析中可以得出 4 种入网模型的性能比较, 如表 1 所列。

表 1 入网模型性能比较

	负载均衡	网络开销	公平性	容错性	扩展性
主动模型	Normal	High	High	Normal	High

(下转第 44 页)

参考文献

- [1] Akyildiz I F, Su W, Sankarasubramanian Y, et al. Wireless sensor networks: a survey[J]. *Computer Networks*, 2002, 38(4): 393-422
- [2] Cardei M, Wu J. Energy-Efficient Coverage Problems in Wireless Ad Hoc Sensor Networks[J]. *Journal of Computer Communications. Special Issue on Sensor Networks*, 2005
- [3] Stojmenovic I. Position based routing in ad hoc networks[J]. *IEEE Communications Magazine*, 2002, 40(7): 128-134
- [4] 毛鸞池, 刘明, 陈力军, 等. DELIC: 一种高效节能的与节点位置无关的传感器网络覆盖协议[J]. *计算机研究与发展*, 2006, 43(2): 187-195
- [5] Slijepcevic S, Potkonjak M. Power Efficient Organization of Wireless Sensor Networks[C]// *Proceedings of IEEE Conference on Communications*, 2001, 2: 472-476
- [6] Cardei M, MarCallum D, Cheng X, et al. Wireless Sensor Networks with Energy Efficient Organization[J]. *Journal of Interconnection Networks*, 2002, 3(3/4): 213-229
- [7] Tian D, Georganas N D. A Coverage-preserving node scheduling scheme for large wireless sensor networks[C]// *Proceedings of First ACM International Workshop on Wireless Sensor Networks and Applications*, 2002: 32-41
- [8] Zhang H, Hou J-C. Maintaining scheme coverage and connectivity in large sensor networks[J]. *International Journal of Ad Hoc & Sensor Wireless Networks*, March 2005
- [9] Cardei M, Wu J, Lu M, et al. Maximum Network Lifetime in Wireless Sensor Networks with Adjustable Sensing Ranges[C]// *Proceedings of IEEE International Conference on Wireless and Mobile Computing, Networking and Communications*, August 2005: 22-24
- [10] Kumar S, Lai T H, Balogh J. On k-coverage in a mostly sleeping sensor network[C]// *Proceedings of the 10th Annual International Conference on Mobile Computing and Networking*, 2004: 144-158
- [11] Ye F, Zhong G, Lu S, et al. PEAS: A robust energy conserving protocol for long-lived sensor networks[C]// *Proceedings of the 23rd International Conference on Distributed Computing Systems*, 2003
- [12] Wu K, Gao Y, Li F, et al. Lightweight deployment-aware scheduling for wireless sensor networks[J]. *ACM/Kluwer Mobile Networks and Applications (MONET) Special Issue on "Energy Constraints and Lifetime Performance in Wireless Sensor Networks"*, 2004
- [13] Wu T-T, Su K-F. Determining active sensor nodes for complete coverage without location information[J]. *International Journal of Ad Hoc and Ubiquitous Computing*, 2005, 1(1/2): 38-46
- [14] 毛鸞池, 冯国富, 陈力军, 等. 一种与位置无关的无线传感器网络连通性覆盖协议[J]. *软件学报*, 2007, 18(7): 1672-1684
- [15] Heinzelman W. Application - Specific Protocol Architecture for Wireless Sensor Networks[D]. June 2000

(上接第 29 页)

目录模型	Bad	Low	Low	Low	Low
WellKnow 模型	Bad	Low	Normal	Normal	Normal
WebCache 模型	Good	Normal	High	High	High

结束语 通过对不同 P2P 网络拓扑的研究, 提出了 4 种 P2P 入网模型, 并从负载均衡、管理开销、公平性、容错性、扩展性 5 个方面对入网机制进行了分析、对比。考虑到网络拓扑自身的一些特点, 上述 P2P 入网机制的侧重点不同, 希望通过本文的分析描述, 有助于 P2P 自组织构造的进一步研究。

参考文献

- [1] Li Gong. Peer-to-Peer networks in action [J]. *IEEE Internet Computing*, 2002, 6(1): 37-39
- [2] Lua E K, Crowcroft J, Pias M, et al. A survey and comparison of Peer-to-Peer overlay network schemes [J]. *Journal of IEEE Communications Survey and Tutorial*, 2005, 7(2)
- [3] Zhao B Y, Kubiatowicz J D, Joseph A D. Tapestry: a fault-tolerant wide-area application infrastructure [J]. *ACM SIGCOMM Computer Communication Review*, 2002, 32(1): 81
- [4] Rowstron A, Druschel P. Pastry: Scalable, decentralized object location, and routing for large-scale Peer-to-Peer systems[C]// *Proceedings of the IFIP/ACM International Conference on Distributed Systems Platforms (Middleware'01)*. Heidelberg, Germany: Springer-Verlag, 2001: 329-350
- [5] Ratnasamy S, Francis P, Handley M, et al. A scalable content-addressable network[C]// *Proceedings of the ACM SIGCOMM 2001*. Govindan R, ed. ACM Press, 2001: 161-172
- [6] Stoica I, Morris R, Karger D, et al. Chord: A scalable Peer-to-Peer lookup service for internet applications[C]// *Proceedings of the ACM SIGCOMM 2001*. Govindan R, ed. ACM Press, 2001: 149-160
- [7] Stephanos A T, Diomidis S. A survey of Peer-to-Peer content distribution technologies [J]. *ACM Computing Surveys*, 2004, 36(4): 335-371
- [8] Saroiu S, Gummadi P K, Gribble S. Measuring and analyzing the characteristics of Napster and Gnutella hosts [J]. *Multimedia Systems*, 2003, 9(9): 170-184
- [9] Jovanovic M A. Modeling large-scale peer-to-peer networks and a case study of Gnutella [D]. Cincinnati: University of Cincinnati, 2001
- [10] Clarke I, Sandberg O, Wiley B, et al. Freenet: A distributed anonymous information storage and retrieval system[C]// *Proceedings of the ICSI Workshop on Design Issues in Anonymity and Unobservability*. Berkeley: Int'l Computer Science Institute, 2000: 43-58
- [11] Liang J, Kumar R, Ross K W. Understanding KaZaA[EB/OL]. 2004. <http://cis.poly.edu/~ross/papers/UnderstandingKaZaA.pdf>
- [12] Daniel S, Reza R. Characterizing the two-tier Gnutella topology [C]// *ACM SIGMETRICS'05*. Bamf, Canada, 2005
- [13] Liang J, Kumar R, Ross K W. The FastTrack overlay: a measurement study [J]. *Computer Networks: The International Journal of Computer and Telecommunications Networking*, 2006, 50(6): 842-858
- [14] Pouwelse J P, Garbacki P, Epema D H, et al. The BitTorrent P2P file-sharing system: Measurements and analysis[C]// *4th Int'l Workshop of IPTPS 2005*. LNCS 3640, Ithaca: Springer-Verlag, 2005: 205-216