文章编号:1001 - 9081(2008)08 - 2105 - 04

基于超立方体拓扑结构的发布订阅系统

李宏杰[1],薛小平[2],王小平[1]
(1. 同济大学 计算机科学与技术系,上海 200092;　2. 同济大学 信息与通信工程系,上海 200092)
(jonathan_lee84@ live. com)

摘　要:从网络拓扑结构着手,提出基于超立方体拓扑结构的发布订阅系统,并在此基础上设计对应的路由算法。基于超立方体拓扑结构网络不仅仅具有一般结构化 P2P 网络的特性,即能够支持大规模、动态的分布式应用,而且由于自身结构的特点使其更适合于发布订阅系统中对大量事件传播的要求。仿真结果表明,超立方体结构的引入能显著降低系统负载,节省带宽。

关键词:发布订阅系统;路由算法;网络拓扑;事件代理网络
中图分类号: TP393.02　　**文献标志码**:A

Publish-subscribe system built with hypercube-based topology

LI Hong-jie[1], XUE Xiao-ping[1], WANG Xiao-ping[2]
(1. *Department of Computer Science and Technology, Tongji University, Shanghai* 200092, *China*;
2. *Department of Information and Communication Engineering, Tongji University, Shanghai* 200092, *China*)

Abstract: Starting from the topology of network, this paper proposed a kind of publish-subscribe system based on hypercube with its related routing algorithm. The hypercube-based topology shares the merits of general structured peer-to-peer network, which can be applied dynamically in large scale. Moreover, due to its own characteristics, the needs of propagating a large number of events can be met. The result of simulation indicates that the system load can be well balanced and bandwidth can be saved, when hypercube-based network is introduced.

Key words: publish-subscribe system; routing algorithm; network topology; event brokering network

0　引言

发布订阅系统(Publish/Subscribe System,P/S 系统)是一种面向分布式应用的通信系统。在 P/S 系统中,信息的发布者以事件的形式将信息发送给事件代理,订阅者向事件代理订阅感兴趣的内容,事件在事件代理网络中传播,最终到达感兴趣的订阅者。P/S 系统使得信息的生产者和消费者在时间、空间和控制流 3 个方面解耦合[1],因而能很好地满足大规模、高度动态的分布式消息系统的需要。

P/S 系统中,分布了大量的事件代理,并在 Overlay 层构成具有一定特性的网络,称为拓扑结构。由于通信对象的动态性和不确定性,需要采用遍历所有节点的方法来寻找通信对象。在这过程中,导致严重的通信开销,阻碍了 P/S 系统的最终实现。而拓扑结构作为事件传播的基础,对 P/S 系统的事件匹配和路由性能起了决定性作用。研究表明,合理的拓扑结构可以提高事件传播的效率。

本文侧重研究 P/S 系统的拓扑结构,引入了超立体结构,并基于这种结构,设计了 P/S 系统的事件路由机制以及应对节点动态变化时的策略。研究结果表明:超立方体结构的引入能显著降低系统负载,节省带宽。

1　P/S 系统及其拓扑结构

P/S 系统一般包括拓扑结构、事件模型、订阅模型、匹配算法和路由算法。拓扑结构定义了事件代理之间的组织结构,它在很大程度上影响着 P/S 系统的性能;事件模型定义了事件的数据结构;订阅模型定义了系统能够支持的订阅条件,指明了订阅者如何表达对事件的兴趣;匹配算法负责寻找有共同兴趣的双方;路由算法则负责选择适当的路径,将一个事件从发布者传送到订阅者,并且它与拓扑结构息息相关,需要针对具体的拓扑结构进行优化。发布者和订阅者之间所交互的信息被称为事件。发布者将事件发送给 P/S 事件代理;而订阅者则向 P/S 事件代理发出订阅请求,表示对系统中的某个事件感兴趣;P/S 事件代理则保证将发布的事件及时、可靠地传送给所有对之感兴趣的订阅者。图 1 所示为一个基本的 P/S 系统结构。

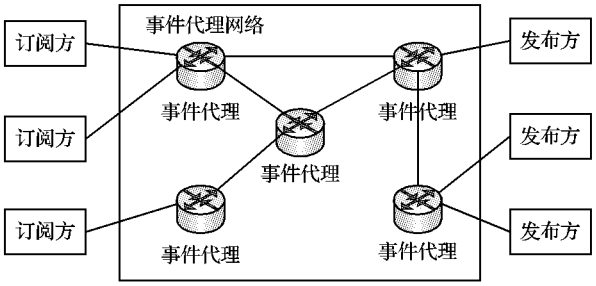

图 1　P/S 系统结构

按照拓扑结构的稳定性把 P/S 系统在结构上分为静态模式和 P2P 模式[2]。

收稿日期:2008 - 02 - 28;**修回日期**:2008 - 04 - 10。　**基金项目**:国家自然科学基金资助项目(60475019)。

作者简介:李宏杰(1984 -),男,浙江嘉兴人,硕士研究生,主要研究方向:计算机网络、可信计算;　薛小平(1964 -),男,江苏常州人,副教授,主要研究方向:计算机网络、移动通信、智能网络;　王小平(1965 -),男,江苏杨中人,教授,博士,主要研究方向:分布式人工智能、复杂适应系统。

静态模式常见的是分层结构。在分层结构中,事件代理之间构成了层次式结构关系,事件代理之间、客户端和事件代理之间的通信采用同一种协议。事件代理会从所有的客户端接收发布的事件和订阅,并转发给关心该事件的子树。该结构中,层级越高的事件代理负载越重,容易产生性能瓶颈。很多P/S的原型系统诸如SIENA[5]、Gryphon[6]和JEDI[7]等采用了分层拓扑结构。

此外,研究人员也在P2P网络之上构建P/S系统,这类结构又可以分为纯P2P网络和混合的P2P网络两种。采用纯P2P网络的P/S系统是指P2P网络中的每个节点同时作为事件代理和客户端,如Scribe[8]系统。而混合的P2P网络的P/S系统是指P2P网络中的每个节点只作为事件代理,每个代理连接若干个客户端,客户端不在P2P网络中,如Hermes[9]。

目前,对P2P网络的研究进入第三代结构化P2P网络,它一般都是基于分布式哈希表(Distributed Hash Table,DHT)技术,运用DHT技术建立具有一定结构的逻辑拓扑,使节点与资源之间形成一定关系,每个节点按一定规则保存系统中部分其他节点的信息,为搜索提供一定的信息,其代表系统有Chord、CAN、Tapestry、Pastry等。由于查询是通过对资源名称所经过哈希函数得到的关键值进行的,因此基于DHT的算法只支持精确查询,无法进行模糊匹配,这就导致了数据查询能力非常薄弱[4],无法满足P/S系统订阅灵活性的要求。

2 基于超立方体的P/S系统

2.1 基于超立方体结构的P2P网络

文献[2]提出了建立于超立方体结构之上的,具有高效广播和搜索的P2P网络——HyperCuP(HyperCube P2P)。这个图形拓扑结构具有很好的特性,例如,各个节点具有相同的职责,网络直径小于$O(n)$,负载比较均衡,整体冗余性较好等。

完整的超立方结构图有$N = b^{L_{\max}+1}$个节点,并且每个节点有$(b-1)\times(L_{\max}+1)$个邻居节点,其中$L_{\max}+1$是由立方体生成的维的度数。超立方体结构是平衡的,有利于负载的平衡。因为,在立方体中每个节点都能成为一个信息发布源,不会造成热点的存在。立方体基数b决定网络的直径和节点度数。这里假定基数都为2,文献[3]介绍了当出现基数$b>2$的情况。假定,网络的节点没有达到$b^{L_{\max}+1}$个,也能很好地工作[3]。因此结构具有很强扩展性,允许任意数目的节点。图2是一个维度为3的完整超立方体结构。

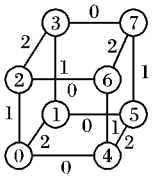

图2 维度为3的完整超立方体结构

在结构中,如果X节点是Y节点的邻接点,则定义Y节点是X节点编号为i的邻居节点,即$Y = iN(X)$。在图2中,节点5是节点4的2号邻居节点。结构中的边是无向的,即节点4也是节点5的2号邻居节点。一个节点的可达节点记为$Y = N(X) = \{x_0, x_1\cdots\}(X)$,其中$N$是达到目标节点的边集,表示从$X$到$Y$所经历各条边的序列。例如:1 = {0,1}(7)和7 = {0,1}(1)。网络中每条边的编号从0开始,节点的邻居等级最高标记为$L_{\max}$。

2.2 基于超立方体结构P2P网络广播特性

基于超立方体结构P2P网络保证了每个节点只会接受和转发一次用来广播的消息,所被广播的消息总和为$N-1$个,并且达到最后一个节点的广播消息经过$\log_b N$跳。数据广播算法如下:

某个节点发起一次广播,它先将消息广播到它所有的邻接点。接收到消息的邻接点把这一消息向节点编号比收到的更大的邻接点进行组播。如图3所示,0号节点发起广播。广播消息会送到所有邻接点4、2、1。节点4从节点编号为"0"的节点收到消息。继而,它向编号比"0"大的节点发送消息,即节点5、6。同时,节点2从节点编号为"1"的节点收到消息。继而,它向编号比"1"大的节点发送消息,即节点3。因为节点1从节点编号为"2"的节点收到消息,而且没有比编号"2"更高的节点,所以节点1不会转发消息。通过这一过程的迭代,这条消息会传播到立方体中的各个节点,完成广播任务。从节点0开始广播到结束的过程可以通过建立一棵生成树的形式表现出来。在此基础上,可以得出任意两个节点之间的平均跳数,以三维超立方体结构为例,平均跳数约为1.714。

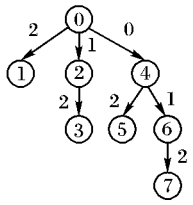

图3 节点0为广播源所建立的组播树

N维超立方体路经长度可以通过式(1)进行计算:

$$L = \frac{1}{N-1}\cdot\sum_{i=1}^{\log_b N}\frac{(b-1)^{\log_b N-i+1}}{(\log_b N-i)!}\cdot\prod_{j=0}^{\log_b N-i}(i+j) \qquad (1)$$

其结果近似于$0.5\cdot\log_b N$。该算法保证了每个节点只接收到一次广播消息,而且路经长度为$O(\log_b N)$,所以具有很好的广播效率。

2.3 基于超立方体结构的P/S系统的事件路由算法

由于P/S系统本身的特性,必须遍历整个代理网络,以确定合适的通信对象。新结构的引入就是针对由于网络不确定性需要大规模数据广播的特点。图4表示了一个基于三维的完整超立方体结构的发布订阅系统模型。图中节点簇和节点簇之间按照超立方体结构组成,簇内由客户端、事件代理和超节点组成。簇内的具体结构不在本文考虑范围,为了方便描述路由算法,可以把簇理解为单个代理节点以及与其相连的订阅方和发布方。具体算法如下。

1)代理节点收到客户端的订阅请求后,在订阅表中查询本地的是否有满足订阅请求的发布方。如果没有匹配,代理节点发起一次订阅请求的广播。它将请求沿着各个邻边发送给相邻代理节点;如果存在匹配则将信息返回给订阅方,并询问是否满足,如不满足,也将发起一次订阅请求的广播。

2)代理节点收到其他代理节点的请求后将请求插入订阅表中。首先检查是否有与之匹配的订阅,如果有则将相关信息沿着请求信息逆向路径返回给订阅方;接着,代理节点将订阅请求向节点编号比收到的更大的邻接代理节点进行组播。

3)重复 2)的操作,直到广播完成。

4)订阅方收到了相关回复后根据需要,连接发布方,进行后续操作。

由于超立方体结构保证了:每个节点只会接受和转发一次用来广播的消息,这样就避免了广播消息在网络中无限度的洪泛,各节点的负载均匀;此外每个节点都能成为广播源,使得连接代理的各个客户端都具有相同的功能。订阅方在订阅信息的同时,也能够进行取消订阅操作。其算法类似订阅广播,作用是在各个代理网络中取消相应的订阅信息。

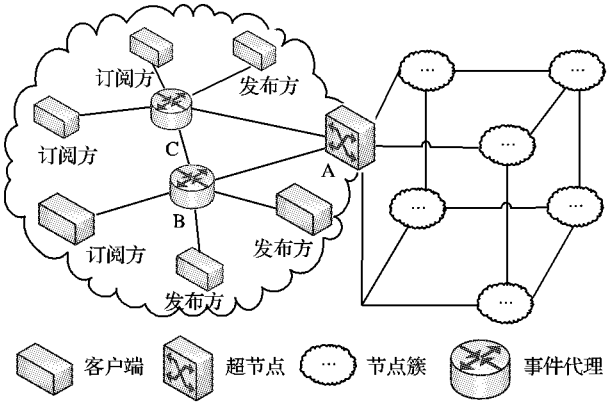

图 4　基于三维完整超立方体结构的发布订阅系统模型

2.4　算法改进分析

事件代理定期检查订阅表,如果事件代理 B 发现其中订阅方 A 经常对某个事件发生兴趣,并且此事件发布者与代理节点 E 互联。在这种情况下,代理节点 B 会通知 A 放弃与它的连接,而与代理节点 E 相连,结果如图 5 所示。与代理节点 E 连接之后,该订阅方所发出的订阅请求将直接得到本地响应,而不会占用外部网络的资源和各个超节点的内部订阅表空间。

由于节点频繁地加入和离开,将使得本身结构的平衡性被打破,虽然在文献[3]中已经说明在不平衡的状态下也能工作,但是这样就无法发挥超立方体结构本身具有的优势。这里提出两种方法:1)当一个代理节点申请加入时,结构中的节点会在一定范围进行协商,找出网络中最需要填补的空位置,然后委托这个位置附近的节点来安置新加入的节点。2)在网络处于平衡状态下,新加入的网络节点选择一个已存节点,与其结合成一个节点簇,共同接入网络。从整体角度来看,新加入的节点不会破坏原来结构的平衡性,而且这种方法的网络代价相对于 1)而言较小。但是如果一个簇过于庞大,也会引起一定的热点效应,此时可以将这个簇进行重组,将一部分节点分配到规模比较小的簇之内。也可以重新开辟一个维度空间,进行重新整合。

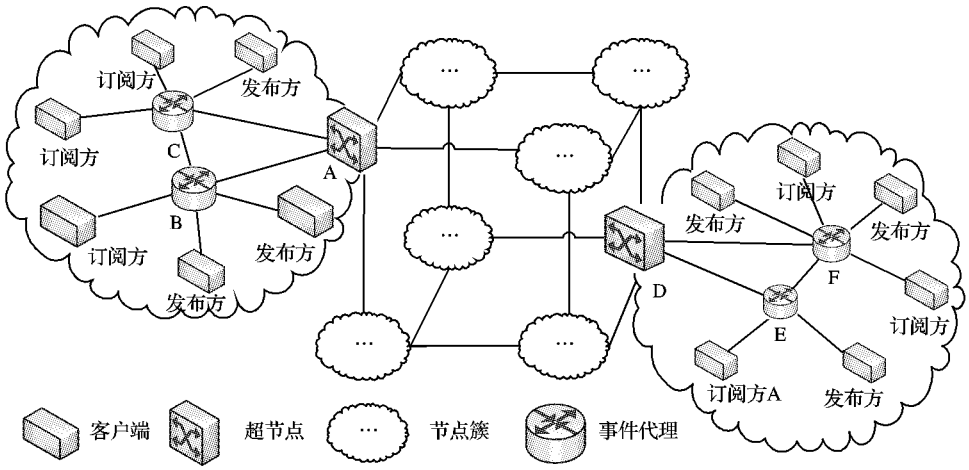

图 5　代理节点 B 放弃与订阅方的连接,并让其与代理节点 E 相连

2.5　系统仿真分析

仿真建立在一台 Pentium4 2.4 GHz、内存 1 GB 主机,运行在 Linux 操作系统的 J-SIM 仿真平台。

对于一个完整的三维超立方体,实验首先模拟各个代理节点随机地发起订阅消息广播,即每个节点发起次数是一定范围内的随机值,这个范围从 10^2 逐级上升到 10^5。每个代理节点收到并转发的广播消息数量如图 6 所示。

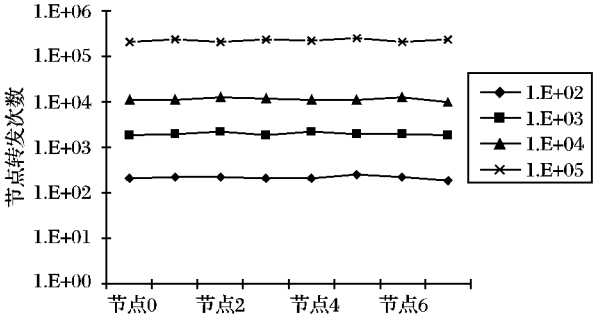

图 6　各个代理节点收到并转发的广播消息数量

虽然每个代理节点发起的消息广播次数是随机的,各个节点之间相差十分剧烈。但从图 6 可以看出,即使在广播消息发起次数不均衡的条件下,每个代理节点所承受的压力非常均衡。从纵向比较,可以得出节点转发消息的数量与节点发起次数呈现线形关系的规律。

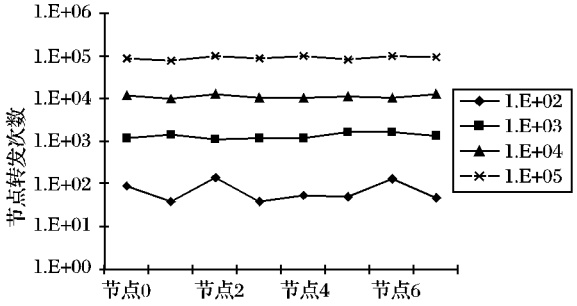

图 7　各个代理节点收到并转发的回复消息数量

假定每个节点对满足其订阅表内的订阅进行响应,发出用于回复给订阅方的消息。这类消息被发起的次数也是一定的范围内随机值,并且范围从 10^2 逐级上升 10^5。在这种情况下,每个节点会根据订阅信息中的路径进行逆向传播,每个节

点所收到并转发的消息数量如图7所示。

同样,虽然每个代理节点发起的回复次数是随机的,节点之间相差十分剧烈。即使在不均衡条件下,每个代理节点所承受的压力十分均衡。从纵向比较,可以发现节点之间压力的均衡度有随着规模上升而上升的趋势,并且节点转发消息的数量与节点发起次数仍然呈现线形关系。另外从实验结果来看,每个事件平均在网络中的转发跳数平均为1.709,满足式(1)。综上所述,超立方体结构的引入,在广播和逆向传播时能够很好地均衡节点之间承载的压力。

3 结语

P/S系统出于自身特点,它的路由算法设计比较复杂。其原因在于事件发布或者事件订阅没有指明具体的接收者,有着不确定性。所以,需要对已有的以广播协议为基础的算法进行优化,以避免不必要的消息转发。这里引入超立方体结构,提供一种高效的广播机制和搜索机制,从底层结构上增强了现有的P/S系统。本文涉及到超立方体结构的P/S系统,还有必要进一步研究其改进的策略。例如:由于在P/S系统中订阅事件、取消订阅事件、结构维护消息等都会频繁地在网络中传播,这将会增加网络链接的开销。其次,对于节点的语义聚集性也有必要进一步研究的方向,引入语义网的概念提高搜索效率。

参考文献:

[1] EUGSTER P T, FELBER P A, GUERRAOUI R, *et al*. The many faces of publish/subscribe [J]. ACM Computing Surveys, 2003, 35(2):114-131.

[2] 马建刚,黄涛,汪锦岭,等.面向大规模分布式计算发布订阅系统核心技术[J].软件学报,2006,17(1): 134-147.

[3] SCHLOSSER M, SINTEK M, DECKER S, *et al*. Shaping up peer-to-peer networks[EB/OL].[2007-08-23]. http://infolab.stanford.edu/~schloss/docs/HyperCuP-DISC2002.pdf.

[4] RATNASAMY S, SHENKER S, STOICA I. Routing algorithms for DHTs: Some open questions[EB/OL].[2007-08-26]. http://www.cs.rice.edu/Conferences/IPTPS02/174.pdf.

[5] CARZANIGA A A, ROSENBLUM D, WOLF A. Design and evaluation of a wide-area event notification service[J]. ACM Transactions on Computer Systems, 2001, 19(3):332-383.

[6] STROM R, BANAVAR G, CHANDRA T, *et al*. Gryphon: An information flow based approach to message brokering[C]// International Symposium on Software Reliability Engineering'98. Washigton D C, USA: IEEE Computer Society, 1998:89-94.

[7] CUGOLA G, NITTO E D, FUGGETTA A. The JEDI event-based infrastructure and its application to the development of the OPSS WFMS[J]. IEEE Transaction on Software Engineering, 2001, 27(9):827-850.

[8] ROWSTRON A, KERMARREC A M, CASTRO M, *et al*. SCRIBE: The design of a large-scale event notification infrastructure[C]// Proceeding of the 3rd International Workshop on Networked Group Communication. London: Springer-Verlag, 2001:30-43.

[9] PIETZUCH P R. Hermes: A scalable event-based middleware [D]. UK: University of Cambridge, 2004.

(上接第2096页)

同的谱范围,采用HMM识别多类特有的L值谱映射优势进行进一步识别,得到了更为理想的识别结果,平均识别率达到了92.8%。

表2 ISVM、ISVM+HMM和标准SVM比较

启动子类别	样本数	平均准确率(MART)/%		
		ISVM	ISVM+HMM	标准SVM
棘皮类动物	44	90.0	93.1	88.0
软体动物	3	100.0	100.0	100.0
非洲爪蛙	28	89.3	92.8	86.0
玉蜀黍花柱	8	87.5	100.0	75.0
节肢动物	300	91.3	91.6	91.0

注:正类准确率为 $A_T = TP/m^+$;负类准确率为 $A_N = TN/n^-$;总体准确率为 $A_{all} = (TP+TN)/k$;$m^+ = TP+FN$,$n^- = TN+FP$,$k = m^+ + n^-$。

5 结语

进一步的实验还发现,超平面始终偏向对角权重系数大的类,而且权重系数可以是任意大于等于零的实数,于是我们将Hessian矩阵对角参数范围进一步拓宽到[0,∞)。这一结论与传统结论不同[1-2],我们认为要根据风险控制进行自动迭代调整而不是将这些参数值固定,这将有利于提高非均衡样本的识别率。对于多类识别,ISVM算法经过若干步后如果仍未达到风险控制,结合隐马尔可夫离散谱映射进行第二次识别,使得识别效果比单独使用非均衡算法更显著。

参考文献:

[1] SONNENBURG S, ZIEN A, RATSCH G. ARTS: Accurate recognition of transcription starts in human[J]. Bioinformatics, 2006, 22(14):472-480.

[2] POLAT K, GUNES S, ARSLAN A. A cascade learning system for classification of diabetes disease: Generalized discriminant analysis and least square Support Vector Machine[J]. Expert Systems with Applications: An International Journal, 2008, 34(1): 482-487.

[3] YU T, DEBENHAM J, JAN T, *et al*. Combine vector quantization and support vector machine for imbalanced datasets[C]// Artificial Intelligence in Theory and Practice. Boston: Springer, 2006, 217:81-88.

[4] ERTEKIN S, HUANG J, BOTTOU L, *et al*. Learning on the border: active learning in imbalanced data classification[C]// Proceedings of The ACM 16th Conference on Information and Knowledge Management. New York: ACM press, 2007:127-136.

[5] BAUM L E, SELL G R. Growth functions for transformations on manifolds[J]. Pacific Journal of Mathematics, 1968, 27(2):211-227.

[6] BAKER J K. The dragon system - an overview [J]. IEEE Acoustic Speech Signal Processing, 1975, 23(1):24-29.

[7] SHA F, SAUL L K. Large margin Gaussian mixture modeling for phonetic classification and recognition[C]// Proceedings of ICASSP 2006. Toulouse: IEEE Signal Processing Society, 2006:265-268.

[8] SHA F, SAUL L K. Comparison of large margin training to other discriminative methods for phonetic recognition by hidden Markov models[C]// Proceedings of ICASSP, 2007. Hawaii: IEEE Signal Processing Society, 2007, 4:313-316.