

一种移动 P2P 网络拓扑生成方法^①

陈世平, 刘 富, 杨晓辉, 方 芳

(上海理工大学 光电信息与计算机工程学院, 上海 200093)

摘 要: P2P 技术是当前研究的热点方向, 众多科研机构和企业加入其中, 取得了令人瞩目的研究成果。由于受到移动网络自身特点制约, 一些在其他网络中应用良好的技术并不一定适合移动 P2P 网络。提出一种基于移动 P2P 网络拓扑生成方法, 该方法通过改进 Chord 算法的方法生成覆盖网网络拓扑, 并提出了一种处理节点移动的方法。实验表明, 该方法在降低覆盖网平均时延的同时, 显著提高了网络的搜索能力。

关键词: 移动网络; P2P; 拓扑; 覆盖网

A Topology Generation Method of Mobile P2P Network

CHEN Shi-Ping, LIU Fu, YANG Xiao-Hui, FANG Fang

(School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: P2P technology is the focus direction of current research, with many research institutions and enterprises joining in, who have made remarkable achievements. Due to Constrained by its own characteristics of mobile network, the good technology of some other network is not suitable for mobile P2P network. This paper presents a method of generate P2P topology based on mobile network. We use this method to generate overlay network topology by improved Chord algorithm, and propose a method of process movement of nodes. Experiments show that the method can obviously reduce the average delay of overlay network, significantly improve the network's search capabilities.

Keywords: mobile network; P2P; topology; overlay network

P2P 计算模式^[6]具有完全的分布式、自组织等特征, 在许多领域已得到广泛的应用。随着网络基础设施逐步建立和移动通信技术的成熟, 以及移动计算设备快速发展, 计算和待机能力的瓶颈被逐步突破, 数据格式也日趋统一。因此, 提出了通过便携机、个人数字助理、移动电话等移动设备, 以方便完成移动用户共享、交互、协作的移动性应用需求。P2P 计算模式在学术研究和工业开发中已经开始扩展到移动设备和移动网络领域^[2]。但是, 移动网络是资源受限的系统, 在移动环境下构建 P2P 网络存在着其所固有的一些缺陷和特点, 如移动计算设备的能量和资源的相对短缺、通信能力相对较弱、网络动态多变和网络连接不可预见性的中断等问题, 这些都给在移动环境下构建 P2P 网络带来了一定的困难。

针对上述问题, 构建合理的网络拓扑对于移动环境下 P2P 网络显得尤为重要。本文提出一种基于移动 P2P 网络拓扑生成方法, 该方法通过改进 Chord 算法的方法生成覆盖网网络拓扑, 能合理地构建移动环境下 P2P 网络拓扑, 有效地降低网络的平均时延, 显著提高网络性能。

1 概况及国内外研究状况

移动通信的快速发展使得移动网络带宽逐步增大、移动终端的处理能力越来越强。GPRS(General Packet Radio Service), UMF5(Universal Mobile Telecommunications System), B3G(Beyond 3G)等蜂窝移动网络和 Wi-Fi(Wireless Fidelity), WiMax(Worldwide Interoperability for Microwave Access),

① 基金项目:国家自然科学基金(60573142);上海市重点学科建设项目(S30504)
收稿时间:2010-08-31;收到修改稿时间:2010-10-19

UWB(Ultra WideBand) 等带宽无线接入网络为用户提供了在任何时间、任何地点访问移动网络服务的可能性。移动应用环境的逐渐成熟使研究者关注由于分布式更强、参与性更广、更具有对等自治特征的移动对等网络环境。

移动对等网络具有以下特征^[2]:

(1) 网络拓扑结构不断变化: 与传统 P2P 网络拓扑结构相对稳定不同, 节点的移动性使移动 P2P 网络拓扑结构不断发生变化, 造成覆盖层(Overlay)与底层物理网络连接状态不匹配, 引起资源发现和数据传输的低效。

(2) 节点自身资源受限: 对移动设备的便携要求, 使当前移动终端的计算处理能力、存储能力、能量供应等受到限制, 这使其在贡献资源的同时必须考虑自身的能耗等因素; 同时, 服务连接的数量也受到限制, 部分节点受所支持协议的限制在无代理服务器的情况下无法直接接入 P2P 网络。

(3) 网络层编址和标识机制不统一: 传统 P2P 系统的网络层采用固定不变的唯一 IP(Internet protocol) 地址作为节点标识, 底层网络的一致使得传统 P2P 网络具有很好的扩展性, 而移动网络环境的异构性使得其网络层所采用的编址及通信方式有很大不同, 要在其上建立统一面向应用的移动 P2P 网络, 就必须屏蔽网络层中不同的网络设备标识的差别。

目前, 国内外对移动 P2P 网络的研究主要集中在 MANET 方面, 对基于蜂窝移动网络和宽带无线接入网络移动 P2P 网络研究较少, 还存在很多未能很好地解决的问题。目前, 国内外的研究方向主要集中在网络体系结构、资源发现策略、网络结构一致性、数据分发及安全 and 隐私问题^[3]。

2 相关工作

2.1 临近节点信息的获取

目前, 搜集底层物理网络的临近信息, 一般有三种获取信息的方式: 地标聚类、洪泛或启发式搜索及采用 IP 地址的分配策略^[3]。地标聚类的基本思想是: 如果两个节点相邻近, 那么它们与那些地标节点的距离应该相似。这种搜集方式存在的问题在于获取的信息精确度不够高。同时, 地标节点也存在着单点失效的问题。泛洪或启发式搜索的方法与 P2P 系统中搜索

的过程相似, 但目的不是为了查找数据, 而是测量节点之间的时延。相对而言, 这种方法的精度更高。可是, 如果不合理地加以限制, 这种探测性的搜索会产生很大网络开销。采用 IP 地址的分配策略的思想是: 如果两个节点的 IP 地址越相似, 那么它们之间的物理距离就越临近。实际上, 网络中并没有对此做严格的限制, 基于这个原则可能导致信息准确度不高。

在移动网络中, 路由到节点的周边邻居仅在底层网络中产生一跳, 这会使它非常高效, 保持到这些节点的连接仅会产生少量的本地网络流量。同时由于移动网络中节点的移动性和网络层编址和标识机制不统一, 使得地标聚类 and 采用 IP 地址的分配策略方法在移动网络中获取临近节点信息并不适用。基于上述特性, 本文中利用洪泛式的方法去获取移动网络中节点的临近信息, 结合跳数和延时时间来综合衡量节点的物理距离。具体做法分为两步:

1) 临近节点信息的获取。实现思想是探测节点通过洪泛式向覆盖网中发出探测信息, 设置 TTL (即生存时间), 其他节点接收到探测信息向探测节点发出响应信息 (包含探测节点到该节点的跳数), 同时将 TTL 减 1, 从而可以获取到探测节点与该节点之间的响应时延 Tdelay 和所经历的网络层路由跳数。

在局部的角度看, 每个节点都是出自它的邻居节点组成的环境中, 每个节点的邻居节点都是在有限跳条件下到达, 我们通过限制节点探测的范围和控制探测信息的大小来减少其带来的网络开销。这里使用参数 TTL 表示节点探测的跳数, 它用来确定节点需要探测多少跳; 参数 Tdelay 表示探测节点与被探测节点之间的时延。开始时, TTL 设置初始值为 init_value。每个临近节点收到一条探测信息, 就将 TTL 减 1, 它向其他节点洪泛这个探测消息, 详细的伪代码如下所示。

算法:

TTL=init_value, Tdelay=0

Detection Node send a probe, Tinit_value

While(TTL > 0) Do

 If Node receive the probe then

 TTL=TTL-1

 Node Send a response message

 Node send this probe to other

nodes

End if

Detection Node receive a response message,

Trreceive

Tdelay =Trreceive-Tinit_value

End While

其中 $init_value$ 的合理设置对减少网络开销有很大影响。恰当的参数值应当根据具体系统做具体设置。

2) 临近节点的确定。我们首先通过跳数来确定探测节点的临近节点,跳数越少,认为该节点与探测节点的物理距离越近。由于路由表存储的信息是有限的,若出现大量临近节点的跳数信息相同时,我们利用延时时间来确定临近节点,延时时间越小,我们认为该节点与探测节点的物理距离越近。

通过上面两步,获取到覆盖网络中所有节点的临近信息后,下一步我们就利用获取到临近节点信息构建节点路由表,形成覆盖网的网络拓扑。

3 基本方法

采用 DHT^[1] (即分布式散列表) 形式进行信息管理, DHT 的每个节点负责数据项的一个特定范围,而且每个节点存储整个分布式系统的部分视图,高效地分布路由信息。本文采用改进 Chord 算法方法去生成和维护移动 P2P 网络。

3.1 标识符空间

Chord^[5]采用相容散列的一种变体计算方法为节点分配关键字。相容散列函数为每个节点和关键字分配 m bit 的标识符,把节点和键值映射到一个大小为 $2m$ 环形空间上。此标识符可以用相容散列 Hash 函数产生,作用于网络中每个节点的 IP,从而得到每个节点的标识;同样,使用 SHA-1 作用于网络资源的“关键字”,从而得到每个网络资源的标识 ID。

每个数据项和节点与一个标识符关联。一个数据项的标识符指的是一个关键字,一个节点的标识符指的是一个 ID。在一个 Chord 环中具有顺时针增长 ID 的一个节点负责其逆时针方向之前的所有关键字。在形式上, (key, value)对(k, v)由 ID 大于或等于 k 的节点持有即在一个 Chord 环中具有顺时针增长 ID 的一个节点负责其逆时针方向之前的所有关键字。

图 1 画出了一个初始化的标识符环,其中 $m=6$,

即有 $26=64$ 个标识符、10 个节点和 5 个数据域。关键字 K10 的后继按顺时针方向其下一个节点是节点 N14,这样 K10 就被定位到 N14 上。

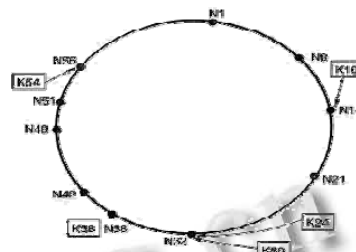


图 1 Chord 环

3.2 路由

3.2.1 扩展路由表

在一个 Chord 标识符环上,每个节点维护一个路由表,即指针表,指向标识符环上的其他节点,本文扩展 Chord 路由表包括后继指针表、邻居集合以及记录下节点到后继指针表与邻居集合中节点的跳数信息。邻居集合的构建原则为:在构建节点的路由表邻居集合时,利用获取到的节点临近信息,按照将跳数少、延迟小的节点设置较高的优先级;而跳数多、延迟大的节点设置较低的优先级别原则对覆盖网中节点的路由表邻居集合进行构建。这样在进行路由选择时,我们总是选择跳数少、延时小且表现稳定的节点作为优先路由表项,这样可以显著提高网络的信息搜索能力,提高网络性能。记录下节点到后继指针表与邻居集合中节点的跳数信息可以为解决移动网络中节点移动问题提供依据。

3.2.2 路由表构建及路由过程

给定一个 Chord 标识符环,关键字和节点被唯一地关联。因此,每个(key,value)对在单个、良好定义的节点上被定位和管理。DHT 有一个标识符环所有节点之上的所有(key,value)对集合形成。

给定一个环,具有 m bit 标识符,一个后继指针表或临近节点表最多有 m 个表项。在节点 n 上,在后继指针表中,行 i 表项标识 n 后至少 $2^i - 1$ 远的第一个节点;在邻居集合中,存储着和节点 n 距离最临近的 m 个节点。在构建路由表同时,节点发出探测“信息”以获取节点到后继指针表与邻居集合中节点所经历的跳数,并记录到指针表中。

路由过程分为两步：第一步，一个节点检查关键字 k 是否在其邻居集合的范围内。如果判断为真，则隐含着 k 位于邻居集合的临近节点之一上。因此，节点将请求转发到数值上与 k 最近的邻居集合节点上。如果判定是节点本身，则路由过程完成。第二步，如果 k 没有落在叶子集合节点的范围内，就需要使用路由表经过一个更长的距离转发请求。节点将查找关键字 k 的请求转发到 k 的最近先驱，该先驱是根据节点的指针表而确定的标识符环上的 k 的先驱。当请求到达一个节点 n ，满足 k 位于 n 及标识符环上 n 的后继之间时，节点 n 回报其后继作为请求的应答，完成路由过程。

3.3 拓扑维护

采用 Chord 算法对节点的加入和离开系统及节点故障的处理方法进行拓扑维护，同时由于本文提出对路由表进行扩展（即加入邻居集合），为了保证邻居信息的准确，也需要对其进行维护。

3.3.1 节点加入

为了加入一个 Chord 标识符环，新节点首先确定某个标识符 n 。对于新节点 n ，在 Chord 系统中必须知道另一个已经加入的节点 o 。通过向 o 提交查找 n 自身 ID 的请求， n 检索其后继。它通知其后继 s 关于它的存在，导致 s 的先驱指针更新为 n 。之后节点 n 同样重复地向 $n+2^1, n+2^2$ 和 $n+2^3$ 等的后继请求 o 而构建其指针，并将节点 o 的邻居集合作为新节点 n 的邻居集合。

3.3.2 节点故障

为了检测节点故障，它和其他节点的所有通信需要检查超时。在一个查找过程中，当一个节点检测到一个指针故障时，节点从其指针表中选中下一个最好的先驱节点，并确保故障节点从指针表中被去除。

3.3.3 节点离开

一个将离开的节点应该将其关键字传输给其后继，并通知其后继和先驱。这将能确保数据不会丢失并会使路由信息保持完整。

3.3.4 邻居集合维护

在移动网络中，存在着节点的频繁移动，这就导致网络路由中经常出现节点故障的情况。但是，由于移动网络资源的有限性，无法做到出现节点故障就去触发重新构造其邻居集合，这样会给网络带来很大的

开销，会显著降低网络的性能，本文采用设置阈值的方式去触发更新邻居集合信息，当故障节点达到阈值时，就触发节点发送“信号”去更新其的邻居信息，流程如图 2 所示：

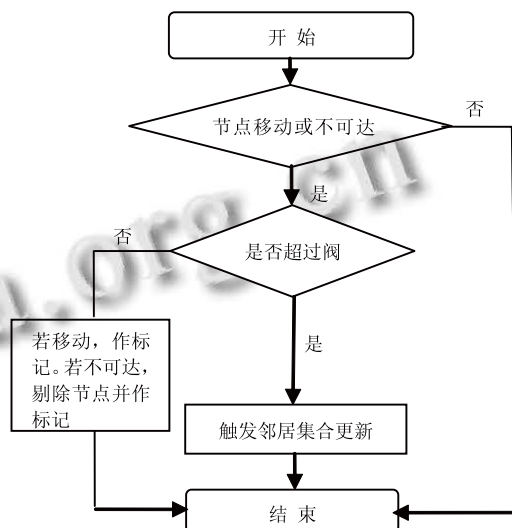


图 2 节点邻居集合维护流程

4 节点移动处理

在移动环境中，考虑到节点之间连接的暂时性，在不同时刻节点将与不同传输范围的节点相交而建立新的物理连接，同时，原有连接可能因为节点距离变远而不可用^[6]。当节点 A 的邻居节点 B 因为某种原因不可达时，节点 A 首先进行失败重传，直至确认节点 B 不再可用。此后，节点 A 将调整路由表，把节点 B 移出邻居集合。此时，若节点 B 又在网络的另一个位置出现，即与网络中其他节点建立物理连接，若节点 A 仍然视节点 B 为邻居节点，则所建立起来的 Overlay 结构将不能反映物理连接的情况；若节点 A 将节点 B 视为新加入的节点，而据此动态调整维护的路由表，Overlay 层就会因为节点无序的移动而变得不稳定，当移动节点频繁加入/退出时，会导致 Overlay 层发生振荡，称为移动性扰动。

例如，在 $T1$ 时刻，节点 A 和节点 B 处在同一区域网络中，节点 A 和 B 之间可以进行通信。而在 $T2$ 时刻，由于节点 B 发生了移动，节点 A 和 B 之间距离边远而变得不可达，导致无法正常通信，如图 3、4 所示。

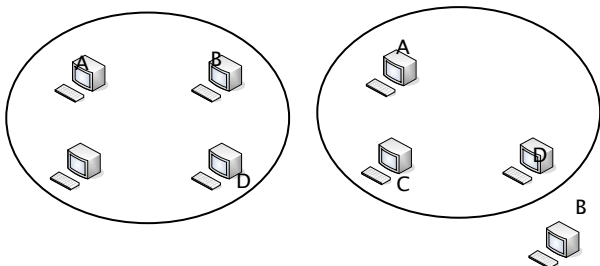


图 3 T1 时刻,节点 A 和 B 位置 图 4 T2 时刻,节点 A 和 B 位置

为了有效地判断节点是否发生移动, 本文通过扩展路由表结构, 在节点的路由表中记录路由表中节点的跳数信息, 通过跳数信息来判断临近节点是否发生了移动, 以此作为解决节点的“移动性扰动”提供了重要的依据。

本文将移动 P2P 网络中节点移动分成两种情况:

1) 节点无法到达, 视为节点故障处理。在进行路由时, 若发生路由表中节点无法到达的情况, 采用处理节点故障的方式处理, 将该节点从路由表中剔除, 并做出记录。

2) 节点可到达, 但路由信息发生变化。在进行路由时, 若发生路由表中目标节点的路由跳数与路由表中记录的跳数不同的情况, 将路由表中目标节点跳数更新为当前跳数, 并做出节点移动记录。

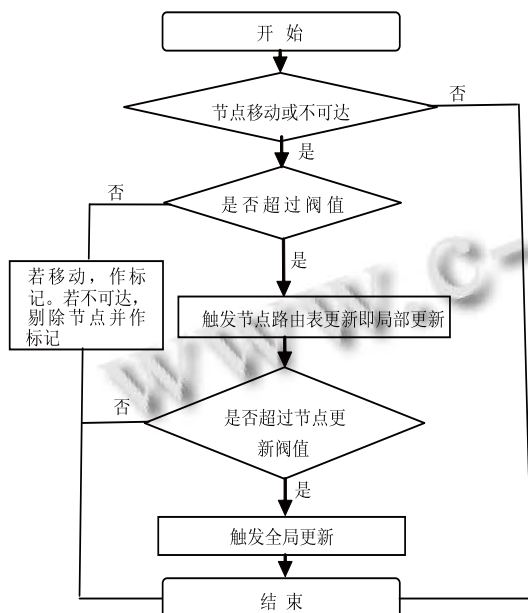


图 5 节点移动处理流程

对于节点的移动, 本文提出事件触发方式处理节

点移动, 分为两种方式处理:

1) 当某个节点路由表中发生移动的节点达到一定数目时, 将采用局部更新的方式重新构建节点的路由表信息, 更新的主要是路由表中邻居集合信息。

2) 在整个移动 P2P 网络中, 若进行路由表更新的节点数目达到一定数目时, 将采用全局更新的方式重新构建节点路由表信息, 流程上图 5 所示。

5 仿真实验及结果分析

本文采用仿真软件来模拟该移动 p2p 网络中拓生成方法, 采用图的方式来构建网络结构, 网络中的端系统用节点来表示, 节点与节点之间的时延采用随机算法模拟, 以节点之间时延的大小作为节点邻居集合生成的依据, 利用邻接表来存储节点的邻居集合, 以节点之间的步长来表示节点之间的跳数。设节点数、时延、伸长系数、时间分别用 N, L, S, T 表示。

本文将仿真实验分两组进行, 第一组对比实验为传统 Chord 算法构建网络拓扑与本文提出的改进的 Chord 算法构建网络拓扑网络性能的比较。第二组对比实验为采用事件触发方法处理节点移动和采用传统优化机制处理节点移动对网络性能影响的比较。上述对比实验每组分别进行 100 次, 取平均值作为实验结果。

性能评价指标主要为获取网络数据的平均时延与伸长系数。平均时延越少, 表明网络性能越好。伸长系数^[4]为平均逻辑链路延迟与平均物理链路延迟的比值, 其中逻辑链路指覆盖网络中相邻节点之间的虚链路。算法中伸长系数反映了系统的逻辑拓扑和物理拓扑之间的匹配程度, 伸长系数越小, 二者匹配程度越高。

第一组实验按网络规模大小分别进行比较。实验结果表明, 改进后的 Chord 算法能显著减少平均时延, 提高网络性能, 如图 6 所示。第二组实验随时间变化按照伸长系数大小比较。实验结果表明, 事件触发方式能有效地解决由于节点移动等原因带来了拓扑不匹配情况, 如图 7 所示。

本实验表明, 本文提出的移动 P2P 网络拓扑生成方法构建出的覆盖网具有良好的网络搜索性能, 有着很高的网络匹配度。

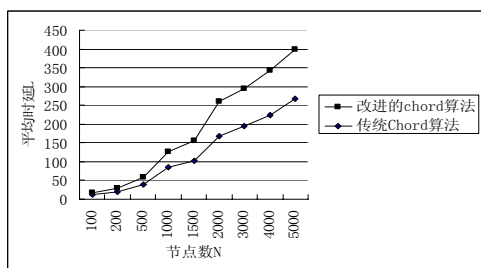


图6 第一组实验结果对比图

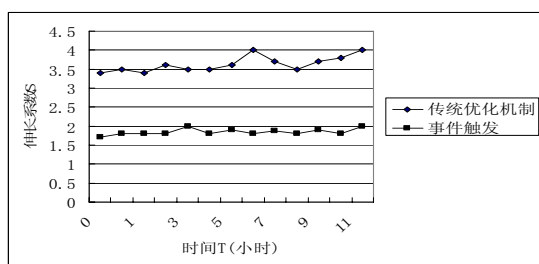


图7 第二组实验结果对比图

6 总结

本文提出了一种构建移动环境下 P2P 网络拓扑结构的方法, 这种方法与协议无关, 适用于各种结构化 P2P 网络, 能有效的构建与物理层网络相匹配的覆盖层网络拓扑。此外, 本文还提出了处理移动 P2P 网络中节点移动的方法, 显著地提高了物理网络拓扑与覆

盖网拓扑匹配程度。实验表明, 通过本方法构建的移动 P2P 网络是一个高效的覆盖网络, 网络搜索性能优越, 拓扑匹配度高。

本文提出的方法未考虑移动设备的异构性, 下一步我们将就降低网络开销及负载均衡等问题展开研究。此外, 移动 P2P 网络还存在着诸多的研究方向如资源发现策略、数据分发及安全及隐私问题, 笔者也将在以后的工作对上述问题展开研究。

参考文献

- 1 斯太门兹等.王玲芳,陈焱译.P2P 系统及其应用.北京:机械工业出版社,2008.
- 2 欧中洪,宋美娜,战晓苏,宋俊德.移动对等网络关键技术.软件学报,2008,2(19):1-4,8-9.
- 3 程久军,李玉宏,程时端,马建.移动 P2P 系统体系结构与关键技术的研究.北京邮电大学学报,2006,4(29):1-2.
- 4 邱彤庆,陈贵海.一种令 P2P 覆盖网络拓扑相关的通用方法.软件学报,2007,2(18):1-3.
- 5 Krcro S, Cleary D, Parker D. P2P Mobile Sensor Networks. Proc. of the 38th Annual Hawaii Int'l Conf. on System Sciences, 2005: 1-9.
- 6 Szer H, Tekkalmaz M, Krpeolu. A Peer-to-Peer File Sharing System for Wireless Ad-hoc Networks. Proc. of the 3rd Annual Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net 2004), 2004: 1-9.
- 5 Sharifian S, Motamedi SA, Akbari MK. A content-based load balancing algorithm with admission control for cluster web servers. Future Generation Computer Systems, 2008,24(8): 775-787.
- 6 谢红薇,谢显宇.基于内容的网络集群负载均衡算法模型.计算机应用与软件,2010,27(1):131-133.
- 7 Youn C. Performance Improvement of Cluster System by Server Status Information. Proc. of the Fourth Annual ACIS International Conference on Computer and Information Science. 2005: 282-287.
- 8 郭成城,晏蒲柳.一种异构 Web 服务器集群动态负载均衡算法.计算机学报,2005,28(2):179-184.
- 9 郑祺.一种基于 WEB 集群的负载均衡算法研究.计算机系统应用,2009,18(7):76-79.
- 10 马双良,张英敏,宋丽君.基于 LVS 和计算任务的实时集群负载均衡方法.计算机工程与设计,2007,28(20):4934-4937.

(上接第 50 页)

- Distribution with Load Balance of Web Servers. Proc. of IEEE International Conference on Cluster Computing, 2001:274-281.
- 2 Vivek S, Mohit A, Gaurav Banga, et al. Locality-aware request distribution in cluster based network servers. ACM SIGPLAN Notices, 1998,33(11):205-216.
- 3 Casalicchio E, Colajanni M. A Client Aware Dispatching Algorithm for Web Clusters Providing Multiple Services. Proc. of the 10th International World Wide Web Conference. Hong Kong, 2001:535-544.
- 4 Cherkasova L, Karlsson M. Scalable Web Server Cluster Designwith Workload-aware Request Distribution Strategy WARD. Proc. of the IEEE International Workshop on Advanced Issues in E-Commerce and Web-Based Information Systems. San Jose, CA, 2001: 212-221.