

基于名的平坦路由

Name-based Flat Routing

中图分类号: TP393 文献标识码: A 文章编号: 1009-6868 (2009) 06-0018-05

摘要:路由是信息网络的基础,分布于网络的各个层面。现有互联网采用了基于IP地址的网络层路由技术和基于域名解析的域名服务器(DNS)应用层路由技术。由于IP地址和域名的结构化特点和身份位置紧耦合特点,使得互联网面临路由可扩展性、安全性、资源服务迁徙移动支持性差等严峻问题。基于名的平坦路由采用了平面化的身份命名方式,实现了身份位置分离,可有效提高网络路由的可扩展性、可配置性、安全性、服务支持的鲁棒性、资源移动的支持性等。

关键词:网络路由;平坦路由;路由体系

Abstract: Routing plays a basic role in information networks, and it works on each layer of networks. Current Internet adopts IP-based network-layer routing and name resolution-based Domain Name System (DNS) application-layer routing. With the structured IP addresses and domain names, as well as with the closely coupled identification and location of both IP and domain names, the Internet are now facing several intractable problems such as bad routing scalability, weak security, weak ability of supporting resource migration and moving. Name-based flat routing uses flat names for the identifications, which achieves the splitting of identification and location. The name-based flat routing can effectively improve network routing scalability, configurability, security, robustness of service providing and ability of supporting resource migration and moving.

Key words: network routing; flat routing; routing system

路由技术是分组网络的核心支撑技术,目的是为分组包确定一条从通信源端到目的端的传输路径,用以解决多跳连接逻辑设备或物理设备间的可达性问题。路由技术分布于网络的各个层面。应用层的域名服务器(DNS)技术、P2P技术、内容传送网络(CDN)路由技术,叠加于IP层以上的Overlay路由技术,网络层的IP路由技术以及链路层的以太网桥技术^[1]等都属于网络路由的范畴。网络路由性能的优劣直接决定着网络通信能力和业务质量的高低。

基金项目:国家重点基础研究发展规划(“973”计划)课题(2007CB307101、2007CB307106)

随着互联网规模的不断扩大,网络层的路由可扩展问题日益严重,已成为制约网络进一步发展的重大隐患。网络的移动、多家乡、流量工程等导致路由表前缀的聚合能力越来越差,虽然硬件的处理速度在不断提高,但相对网络的发展和路由表的增大却明显显得力不从心。2009年7月的统计显示,核心网BGP路由表的前缀数目已接近30万条,而且仍在高速增长^[2]。除此之外,现有路由机制对移动的支持能力也很差。2009年的调查显示全世界范围内已有40亿移动用户,其中高于20%的用户开通了数据业务。未来的手机、个人数字助理(PDA)等手持终端的功能会愈来愈

裘正定/QIU Zhengding
熊轲/XIONG Ke

(1. 北京交通大学计算机与信息技术学院信息科学研究所,北京 100044)
(1. Institute of Information Science, School of Computer Science & Information Technology, Beijing Jiaotong University, Beijing 100044, China)

强,开通互联网数据业务量也会随之飞速增加。这将给网络带来前所未有的挑战。安全性问题是互联网的又一严峻问题,而且日益紧迫。虽然网络已引入入侵检测、防火墙等技术,但互联网的安全性问题仍未得到有效解决。这些问题都与路由体系紧密相关,仅在基于IP的现有路由体系上进行小的修补,难以很好解决互联网的这些问题。

为方便用户的使用,现有网络在应用层引入了域名机制来标志网络资源。采用DNS将资源与地址进行对应,达到网络资源共享的目的。由于域名相对IP地址更加方便用户的记忆,因而,DNS得到了迅速发展,成为了互联网上最大的解析系统。然而,随着网络业务量的不断增长,基于树状结构的DNS逐步暴露出一些致命缺点,如不支持数据移动和复制^[3],域名解析系统鲁棒性不强等问题,并引发了人们对互联网域名解析与资源查找系统的重新思考^[4-7]。

与路由密切相关的就是网络节点或资源的地址定义,不同的地址定义需要不同的路由运行机制与之相配合以实现网络路由功能。研究分析发现,无论是应用层的域名解析系统还是网络层的IP路由技术都具有一个共同特点,那就是它们都是基于层次化、结构化的网络实体地址。IP路

由基于结构化的IP地址,DNS基于层次化的域名定义。这些结构化的地址定义方式给网络早期的发展带来了许多实现好处,而今却成为网络所遇问题难以解决的根本性原因。与结构化地址相对的便是平坦(面)化的地址定义方式。平面化地址打破了传统的结构化的地址定义,随之也产生了与现有路由机制大相径庭的路由方式。近年来,很多研究者的研究都体现了平坦路由的思想。平坦路由打破了现有基于结构化地址的路由方式与网络体系机构,为有效解决现有路由难以解决的问题提供了新的研究思路和方法,已成为新一代路由研究的热点方向之一。

1 现有路由存在的问题

1.1 网络层路由存在的问题

(1)路由可扩展问题

因为IP地址具有身份与位置的双重特性,因而现有互联网的可扩展性存在一个基本假设即IP地址按网络拓扑分配或按地理位置分配。身份特性要求IP地址保持稳定性和唯一性,而位置特性要求IP与网络拓扑地址相对应。因而,网络要支持移动性、多家乡等就会出现IP的两个属性冲突的问题,这造成路由表难以很好聚合。造成IP难以聚合另外一个原因是,现有IP的分配很不合理。美国作为国际互联网的领头军,获得了70%以上的IP资源。众所周知,美国麻省理工学院一所大学拥有的IP地址比中国全国的拥有量还多。其他国家为使用互联网不得不引入其他技术,如网络地址转换(NAT)技术,这造成网络的复杂性越来越高,亦与身份属性要求的唯一性相矛盾。

(2)地址分配不灵活

正是因为IP的双重属性和IP资源拥有的“贫富”差异巨大,造成IP地址的分配受限很多,且很多情况下的分配都不合理。这反过来又制约了网络的发展和路由的扩展性。

(3)IP地址的配置方便性差

对于网络和计算机知识较缺乏的用户,要实现上网,在获得IP地址后,要进行IP的配置无疑是一件很不方便的事情。尽管因引入了动态主机配置协议(DHCP),简化了用户的操作,却给网络带来了很大的负担,而且DHCP打破了一个身份对应唯一的一个终端的假定。

(4)移动性支持差

现有网络中,因为IP具有地址信息,终端在移动后不得不改变地址以实现通信,由于IP的身份与位置紧耦合,将导致移动中终端的身份不得发生变化。而身份的变化会导致通信的中断,为实现IP网络的移动性,不得不引入大量的操作来管理和实现转交地址与原IP地址的转换。

(5)安全性差

由于IP地址同时包含身份与位置信息,网络的恶意窃听者都可以根据一个IP地址获得其身份信息和拓扑位置信息。

1.2 应用层路由存在的问题

应用层的DNS系统由于其设计初期考虑的比较简单,随着用户数量的飞速攀升和数据移动性的增强,在当前的应用中亦逐步暴露出很多致命问题。

(1)DNS管理权限的不平衡性

DNS根服务器扮演了十分重要的角色,所有本地DNS服务器不能解析的域名解析请求都需要被送到DNS根服务器进行处理。由于这些DNS根服务器的管理权与控制权在美国政府的手中,美国可以对DNS根服务器中的ROOTZONE文件的记录进行修改,从而使得一些国家从互联网世界中消失成为可能。

(2)配置复杂,鲁棒性、安全性差

DNS需要手动配置,随着系统的复杂程度逐步提高,配置出错的概率以不断增高,给网络的可靠性和安全性带来了隐患和大量的损失。

(3)易引起网络域名所有权纠纷

因为网络域名带有语义,不可避免涉及到所有权归属问题,从而带来可能的法律纠纷,如www.yaoming.com是否可以被姚明以外的别人注册,一旦被他人注册,所发布的信息很可能对姚明的某些利益造成损害。

(4)不支持数据的移动与迁移

因为DNS是面向主机的,当主机上面的数据被移动或复制后,DNS无法再次为该数据提供服务,用户用以前的名字便无法找到相应的资源。

(5)负载不均衡

因为域名的不同字段的不同符号具有不同含义。在DNS服务器中,处于业务需要,负责.com域名解析的服务器相对于负责.org等域名解析的服务器负载要重很多,这导致.org的DNS服务器资源利用不足,而.com的服务器负荷过重。

为解决现有IP层和应用层路由的问题,很多研究者将目光发在了平面路由方式上。VRR[18]研究了基于分布式哈希表(DHT)环的路由协议,直接工作于链路层之上的协议。CANON[19]给出了一般性构造DHT环的方法,研究了多个DHT环的合并算法。ROFL[10]对CANON进行了扩展,使之能够支持更多的互联网策略。SEATTLE针对小规模以太网给出了平坦路由的实现方法。在文献[7]中提出名为DONA的映射系统,用扁平名字方式提供新服务的接入接口。在文献[5]中提出了DHT形式的命名与解析,实现了身份与地址的分离,身份信息只用于名字解析,地址信息用于路由。在文献[6]中提出用改进的DHT系统OpenDHT来实现命名与解析。在文献[7]中提出了双层结构来提高DHT系统的查询效率,这种结构给我们名字解析映射系统的设计提供了新的思路。

2 基于名的平坦路由

2.1 基于名的单DHT环平坦路由

采用单个DHT环管理整个网络的路由运作。下面给出基于名的单DHT

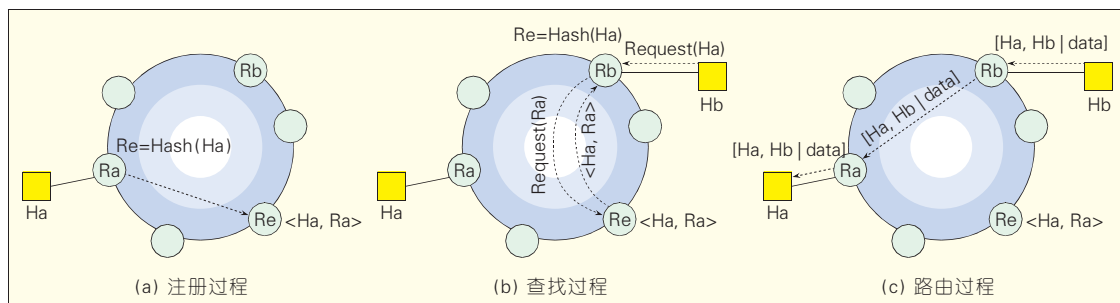


图1
基于名的单DHT环
平坦路由

环平坦路由的工作原理与过程。

在基于名的平坦路由体系中,网络每个终端节点拥有一个平面化的无含义的标志来代表其身份。用网络路由器节点的子集构成一个逻辑相互可达的DHT环,环上的节点分布式管理并负责一部分终端的位置注册与查找。其工作过程主要可分为3个阶段,即位置注册、位置查找和数据路由3部分。位置注册过程的主要任务是,终端在接入网络时向对应的网络位置管理节点注册自己的接入位置。位置查找过程主要是指,通信终端通过通信对端的名字向网络查询其对应的接入位置。而数据路由过程主要负责的是数据的路由优化及正确传输。

以图1为例,大圆环代表DHT环,小圆圈节点代表位于DHT环的路由节点,正方形节点代表终端节点。

(1)位置注册过程

如图1(a)所示,当终端Ha接入网络时,其接入路由器Ra根据Ha的名字Ha采用Hash算法计算出Ha对应的位置注册节点为Re,然后向Re发送<Ha, Ra>消息,在Re上注册Ha的接入位置Ra。

(2)位置查找过程

如图1(b)所示,终端Hb的接入路由器为Rb。当Hb想向Ha发送数据时,Rb需要知道Ha的接入位置,才能进行路由。首先,Hb向Rb发送Request(Ha)查询Ha的接入位置,若Rb的Cache里没有Ha的位置表项,便用Ha哈希出其位置注册节点为Re,然后向Re发送Request(Ha)请求;Re收到查询消息后,查得位置信息对<Ha, Ra>,

并将此信息回复给Rb;至此,Rb获知Ha的接入位置为Ra。

(3)数据路由过程

如图1(c)所示,Hb以[Ha, Hb]为通信的目的、源标志向网络发送数据包;Rb收到Hb发送给Ha的数据包,知道Ha的接入位置为Ra,便将数据包转发给Ra;Ra收到数据包后,将其发送给Ha。

上述3个阶段可实现基于名的单DHT环的平坦路由过程与通信。

2.2 基于名的多DHT环平坦路由实现过程

显然,单DHT环的平坦路由只适合于小规模网络。考虑到大规模网络的路由扩展性问题和DHT环的管理难度,需对主机名进行扩展。这里提出了一种扩展方法,采用了层次DHT的方式,将整个互联网分成若干个逻辑域,每个逻辑域由一个DHT环负责

管理和路由。每个域中设置一个或多个域间位置管理节点(可由高性能路由器担当),各个域的域间位置管理节点逻辑上亦相互连通,并组成一个二级DHT路由环,如图2所示。

记扩展后的主机名为E_Ha,其高x位表示所在域的域号Ai,低n-x位为域内身份标志Ha,因而,E_Ha=Ai+Ha。其中Ai,Ha均为平面化的一串字符。下面以图2为例详细叙述双层DHT环的域间路由实现过程。图2中标号为“1”的虚箭头所示为主机E_Ha接入注册过程,标号为“2”的虚箭头表示主机E_Hb向E_Ha发送数据前查询E_Ha位置的过程,标号“3”为数据路由过程。

(1)位置注册过程

当E_Ha在Ra上接入时,Ra首先根据E_Ha的Ha字段Hash出其域内位置管理节点Re,然后将<E_Ha, Ra>注册到Re上,与此同时,Ra也向域I的域间

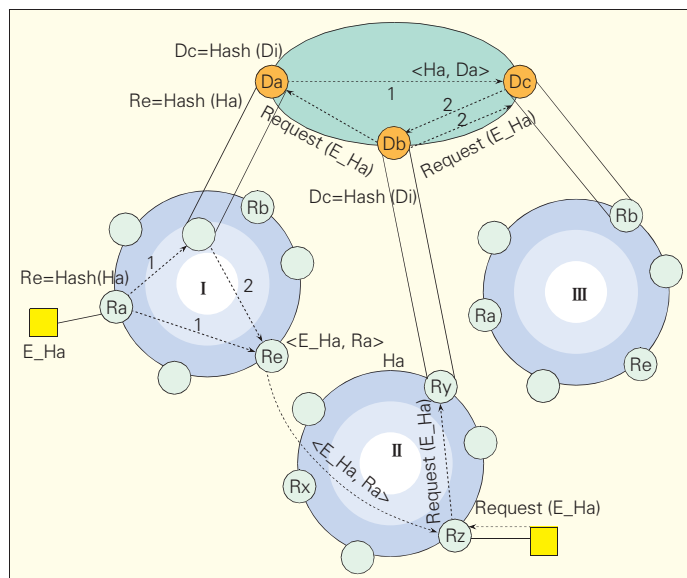


图2
基于名的二层
DHT平坦路由注
册、位置查找和路
由过程

位置管理节点Da发送E_Ha的接入消息。Da收到E_Ha的接入消息后,用E_Ha的Ai字段Hash出其域间的位置注册节点Dc,然后将E_Ha的接入域位置信息<E_Ha, Da>注册到Dc上。至此,E_Ha完成域内、域间的位置注册过程。

(2)位置查找过程

当域II的E_Hb想向E_Ha发送数据时,E_Hb首先向自己的接入路由器Rz发送位置查询消息,若Rz的Cache中没有E_Ha的位置缓存,而且根据E_Ha的Ai字段判断发现E_Ha的位置在II域内,便向域II的域间位置管理节点Db发送Request(E_Ha)消息查询E_Ha的接入位置;若Db的Cache中亦没有E_Ha的位置缓存表项,Db就根据E_Ha的Ai字段Hash出其域间位置管理节点Dc,于是向Dc查询得E_Ha的域间接入节点位置为Da;接着,Db向Da发出E_Ha的接入位置查询信息,若Da的Cache中亦没有E_Ha的位置缓存表项,Da便根据E_Ha的Ai字段Hash出其域内接入位置Re,然后将查询消息发送给Re;Re收到查询请求消息后,直接将查询结果<E_Ha, Ra>返回给Rz。于是Rz获知E_Ha的接入位置为Ra。

(3)数据路由过程

此后,E_Hb开始以[E_Ha, E_Hb]通信的目的、源标志向E_Ha发送数据;Rz接收到数据包后直接向Ra转发;Ra再将数据包送至E_Ha。至此,完成域间路由和数据传输过程。

需要说明的是上述两小节仅给出了基于名的平坦路由在小网络和大网络的最基本方案和工作过程。在具体实施时,还需对各个细节进行优化以提高系统性能,如,可在相关节点加入高速缓存机制,提高查询速度,减小查询开销等。

3 基于名的平坦路由优点分析

基于名的平坦路由,由于采用了无含义平面化的标志来代表网络实

体的身份,实现了网络实体的身份ID与位置ID的分离。能够有效解决现有网络路由体系难以处理的问题,具体表现如下。

3.1 工作于网络层的优势

将基于名的平坦路由应用于网络层,其方案与工作过程前面已有了较为全面的描述。本节重点分析在网络层实施平坦路由带来的好处。

(1)实现了用户(主机)名空间与网络名空间的分离,提高了网络优化的灵活性和可行性。在平坦路由体系中,用户名不再具有地址信息,与网络拓扑没有联系,因而用户与网络被“隔离”开来。对于网络而言,路由器节点的数量要远远小于用户主机的数量,因而可以考虑采用不等长的地址长度定义路由器的地址和主机名。用较少位数的二进制串定义路由器节点的地址,因路由器移动的可能性很小,故仍可采用如IPv4实现。对于用户名空间,由于数量庞大需采用较多二进制位表示,如128位或160位。这样的好处是,数据一旦进入网络可以按短地址进行传输,提高传输效率,而不像现有的IPv4或IPv6网络,主机地址与路由节点地址长度相同。此外,用户空间名与网络名空间分离后,运营商在对路由设备进行调整或对核心网重新部署、调整和优化时,能够使用户感觉不到网络的变化,这大大增加了网络维护的灵活性和可操作性。而现有网络因主机IP与网络位置密切相关,一旦接入路由器的地址发生变化,用户主机也需重新配置,这是严重制约现有网络进行优化的重要因素。

(2)实现了更简单、灵活的主机名分配方式。现有IP地址是身份与位置的紧耦合体,既代表了网络终端节点的身份也包含着网络终端的位置信息。因此IP地址的在分配时除了考虑全局唯一性外还必须与网络拓扑密切联系。然而,由于IP资源占有的不均衡性和网络发展的无尺度特性,使

得IP地址的分配与理想差距越来越远。正是因为IP地址分配的不合理性,使得IP网络发展到今天出现了很多难以解决的问题。在平坦路由体系中,平面化的名称是一串无含义的二进制码串,只需保障全局唯一即可,因此分配与管理起来要方便和灵活得多。

(3)缩小了核心网络路由节点的路由表规模,提高了路由可扩展性。因为用户名空间被分离开来,大大减小了核心网路由器所需存储的路由条目。通过统一的Hash函数确定终端位置,网络中终端位置信息被均衡分布在所有路由节点上存储,相比全平面化基于主机路由方式构建的路由系统,减轻了每个节点的路由表存储开销,提高了系统资源使用的均衡性。在平坦路由体系中,因身份名与地址无关,终端的移动并不会引起路由表的条目和聚合难度的增长,因而可保证路由的可扩展性。

(4)更安全、可靠。对于主机名的分配可采用文献[10]所述的自认证公私钥体制方式,能有效杜绝和降低现有的IP地址欺骗等攻击行为,增强网络的安全性。在基于名的平坦路由体系中,通信的双方仅通过“名字”相互通信,并不知道对方的具体位置,保护了用户位置信息的隐秘性,提高了通信的安全性。此外,由于用户空间与核心网空间的解耦合,使得恶意用户难以利用地址对核心网路由设备实施攻击,这也大大增强了核心网的安全性。

(5)对移动、多家乡的有效支持。根据第2节的工作原理描述可清晰看出,平坦路由体系通过在DHT环上的相应节点存储终端接入位置信息,终端在网络内发生微移动时,仅通过统一的Hash函数便可确定位置,进行通信。终端的地址和“名”不需要发生任何变化,有效减少了网络维护终端移动的开销,提高了对移动的支持性。对于多家乡而言,一个终端若两个位置进行了接入,只需在位置注

册节点进行注册,不需要再分配底层通信的临时地址,也不用增加路由表项,因而基于名的平坦路由可有效支持多家乡。

(6)用户配置简单。用户上网不用再进行复杂配置,特别对于移动用户,在不断更换接入位置的情况下,不用像在IP网中一样,换一个接入网关就需要重新配置一次IP信息。

3.2 工作于应用层的优势

将基于名的平坦路由体系用于应用层的DNS系统或P2P网络,其过程与第3节所述过程相似。不同之处在于,此时代表“名”的二进制码串标志的是一个资源实体而不是主机。注册过程和查找过程与图1相似,但代表的是资源位置的注册过程和资源的位置查找过程。基于名的平坦路由应用于服务资源查找系统的优势在于。

(1)数据安全性增强。因平面名字不带有语义信息,所以不存在某个国家对整个解析映射系统的控制,因为无法从没有语义信息名字中区分哪些名字属于那些国家或单位,因此打破了现有DNS系统权限不均衡性,有效消除了现有DNS系统的隐患。

(2)配置简单,系统的鲁棒性、安全性得到增强。平面名字不需要手动配置,生成过程由系列算法实现,从而避免了手动配置的任务,消除了人为错误配置所产生的安全隐患。

(3)消除了域名注册中的商业纠纷隐患。因平面名字不带有语义信息,所以在全球唯一的前提下随便使用不会带来商业纠纷。

(4)支持数据移动和迁移。因平面名字是面向数据的,与资源的位置无关。当主机上面的数据移动或复制以后,数据的名字永远不变是唯一的,所以可以用以前的名字来找到相应的资源。

(5)实现了资源查找系统的负载均衡。平面解析映射系统一般采用Hash算法,而Hash算法是一种散列算

法,他可以将不均匀的分布变为均匀分布,所以,实现解析映射系统的负载均衡。

4 结束语

随着互联网的不断发展和用户数量的不断增长,现有的基于结构化地址的路由方式已暴露出许多难以解决的问题。作为网络的核心支撑技术——路由技术,直接决定着网络的继续发展与成败。为此,本文给出了基于名的平坦路由的实施方案,首次提出了多域间的平坦路由实现过程。由于平面化名称使得网络资源和主机等实体与其位置信息分离开来,可很好地提高网络的路由可扩展性、移动支持性、路由安全性、可配置性、迁移与复制资源查找的支持性以及网络负载均衡等性能。基于名的平坦路由打破了传统的路由体系,是有效解决网络现有路由问题的一种新的研究思路与方案技术。然而,平坦路由的研究仍处于起步阶段,很多实现方案细节还很不完善,如平面化名称如何定义更有效、大规模网络的平坦路由问题,如何在平坦路由中实施路由策略,流量工程等问题都还有待于进一步研究。

5 参考文献

- [1] KIM C, CAESAR M, REXFORD J. Floodless in SEATTLE: A scalable Ethernet architecture for large enterprises[C]// Proceedings of Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications (SIGCOMM' 08), Aug 17-22, 2008, Seattle, WA, USA. New York, NY, USA: ACM, 2008: 3-14.
- [2] AS65000 BGP routing table analysis report [EB/OL]. [2009-6-26]. <http://bgp.potaroo.net/as2.0/bgp-active.html>.
- [3] WALFISH M, BALAKRISHNAN H, SHENKER S. Untangling the Web from DNS[C]// Proceedings of the 1st Conference on Symposium on Networked Systems Design and Implementation (NSDI' 04), Mar 29-31, 2004, San Francisco, CA, USA. Berkeley, CA, USA: USENIX Association, 2004: 225-238.
- [4] RAY S, GUERIN A, SOFIA R. A distributed hash table based address resolution scheme for large scale Ethernet networks[C]// Proceedings of IEEE International Conference on Communications (ICC' 07), Jun 24-27, 2007, Scotland, UK. Piscataway, NJ, USA: IEEE, 2007: 6446-6453.

- [5] KOPONEN T, CHAWLA M, BYUNG G C, et al. A data-oriented (and beyond) network architecture[C]// Proceedings of Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications (SIGCOMM' 07), Aug 27-31, 2007, Kyoto, Japan. New York, NY, USA: ACM, 2007: 181-192.
- [6] RHEA S, GODFREY B, KARP B, et al. OpenDHT: A public DHT service and its uses [C]// Proceedings of Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications (SIGCOMM' 05), Aug 22-26, 2005, Philadelphia, PA, USA. New York, NY, USA: ACM, 2005: 73-84.
- [7] MIZRAK A, CHENG Y, KUMAR V, et al. Structured superpeers: Leveraging heterogeneity to provide constant time lookup [C]// Proceedings of the 3rd Workshop on Internet Applications (WIAPP' 03), Jun 23-24, 2003, San Jose, CA, USA. Los Alamitos, CA, USA: IEEE Computer Society, 2003: 104-111.
- [8] CAESAR M, CASTRO M., NIGHTINGALE E B, et al. Virtual ring routing: network routing inspired by DHTs[C]// Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications (SIGCOMM' 06), Sep 11-15, 2006, Pisa, Italy. New York, NY, USA: ACM, 2006: 351-362.
- [9] GANESAN P, GUMMADI K, GARCIA-MOLINA H. Canon in G major: Designing DHTs with hierarchical structure[C]// Proceedings of the 24th International Conference on Distributed Computing Systems (ICDCS' 04), Mar 23-26, 2004, Tokyo, Japan. Piscataway, NJ, USA: IEEE, 2004: 263-272.
- [10] CAESAR M, CONDIE T, KANNAN J. ROFL: Routing on flat labels[C]// Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications (SIGCOMM' 06), Sep 11-15, 2006, Pisa, Italy. New York, NY, USA: ACM, 2006: 363-374.

收稿日期: 2009-08-30

作者简介



裘正定, 北京交通大学计算机与信息技术学院信息科学研究所教授, 中国通信学会会员、中国电子学会会员、铁道学会高级会员, 主要研究领域为多媒体通信、下一代互联网络、信号与信息处理等。已发表SCI/EI检索论文100余篇。



熊轲, 北京交通大学计算机与信息技术学院信息科学研究所读博士生, 主要研究方向为多媒体通信、下一代互联网络、信号与信息处理等。已发表论文10余篇。