ISSN 1000-9825, CODEN RUXUEW
Journal of Software, Vol.20, No.11, November 2009, pp.2988−3000
doi: 10.3724/SP.J.1001.2009.03603
© by *Institute of Software, the Chinese Academy of Sciences*. All rights reserved.
E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel/Fax: +86-10-62562563

互联网覆盖路由技术研究*

王旸旸[1,2+], 毕 军[2], 吴建平[1,2]

[1](清华大学 计算机科学与技术系,北京 100084)
[2](清华大学 信息网络工程研究中心,北京 100084)

Research on Internet Overlay Routing

WANG Yang-Yang[1,2+], BI Jun[2], WU Jian-Ping[1,2]

[1](Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China)
[2](Network Research Center, Tsinghua University, Beijing 100084, China)
+ Corresponding author: E-mail: wangyy-06@mails.tsinghua.edu.cn

Wang YY, Bi J, Wu JP. Research on Internet overlay routing. *Journal of Software*, 2009,20(11):2988−3000.
http://www.jos.org.cn/1000-9825/3603.htm

Abstract: This paper surveys related research work of Internet overlay routing, and focuses on the structure and methods of Internet overlay routing on network layer and transport layer according to Internet layering scheme. Apart from that, other related aspects of overlay routing are discussed, including the impacts on overlay routing performance, and interactions between multiple overlay networks routing, etc. The advantages and disadvantages in the key technologies of related research work are analyzed. At last, the possible future research directions are discussed as references for the research of Internet overlay routing.
Key words: routing; overlay routing; overlay network topology; multi-path routing; Internet routing

摘 要: 对互联网覆盖路由相关研究工作进行了调研;按照 Internet 分层概念重点考查了网络层和传输层的覆盖路由的结构和方法;并且考查了覆盖路由其他相关问题的研究,包括覆盖路由性能的影响因素、多个覆盖网络路由之间的相互作用等.分析了已有研究工作的关键技术的优缺点.最后讨论了未来可能的研究方向,为互联网覆盖路由的研究提供参考.
关键词: 路由;覆盖路由;覆盖网络拓扑;多路径路由;互联网路由
中图法分类号: TP393 **文献标识码**: A

当前在 Internet 上,覆盖网络(overlay network)得到了广泛应用,包括文件共享和流媒体服务的 P2P 覆盖网络、内容分发网络(content deliver network,简称 CDN)、应用层组播(application layer multicast,简称 ALM)、虚拟实验床 EmuLab 和 PlanetLab 等.

覆盖网络是以底层物理网络为基础,在其上构建的虚拟网络系统.在覆盖网络中,节点之间的虚拟链路是逻

* Supported by the National Basic Research Program of China under Grant No.2009CB320501 (国家重点基础研究发展计划(973)); the Specialized Research Fund for the Doctoral Program of Higher Education of China under Grant No.200800030034 (高等学校博士学科点专项科研基金)

Received 2008-02-27; Accepted 2009-02-24

© 中国科学院软件研究所 http://www.c-s-a.org.cn

辑上的,通常对应于底层网络的物理路径.覆盖网络可以根据应用环境和需求定义自己的拓扑结构和路由模式,结构比较灵活,可以用来构建特定于应用(application-specific)的服务,大大扩展了 Internet 的服务.

覆盖路由(overlay routing)是通过覆盖网络进行的路由模式,是覆盖网络的关键部分.图 1 演示了覆盖路由的基本概念.如图 1 所示,*A*,*B*,*C* 是覆盖网络节点,当 Internet 路径 *AB* 发生故障或者拥塞时,*A* 通过节点 *C*,经过路径 *AC* 和 *CB*,与 *B* 进行通信,从而绕过原来的故障路径.

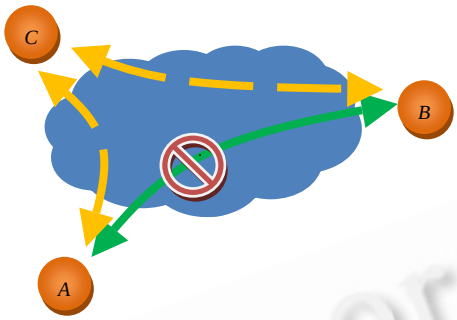

Fig.1 Illustration of overlay routing

图 1 覆盖路由示例

与底层 Internet 路由相比,覆盖路由具有更大的灵活自主性.覆盖网络中的转发节点可被看成“智能网关”,它不仅可以根据“源地址、目的地址”,也可以根据报文的内容,并且结合其他需求决策下一跳节点.然而这种灵活性伴随较大的开销,使用范围和规模会有一定的局限性.

近年来,覆盖网络的相关研究备受关注.我们不仅可以使用覆盖网络提供各种新型应用,而且可以利用其中的覆盖路由改善覆盖网络或者 Internet 端到端服务性能.对于后者,可以把覆盖路由看成一种服务.在这方面有许多研究做了尝试.在早期,Detour 和 RON 项目提出使用覆盖路由改善覆盖网络的性能,随着研究深入,开始使用覆盖路由提高 Internet 端到端的服务性能,比如传输速率、可靠性、QoS 等.然而这种覆盖路由若要在 Internet 得到广泛部署和应用,还有许多关键问题有待解决.本文将对相关的覆盖路由研究进展进行总结,并探讨覆盖路由存在的关键问题和研究前景.本文第 1 节对当前主要的覆盖路由结构和方法进行分类介绍,并分析一些方法的不足之处.第 2 节介绍覆盖路由相关问题的研究.第 3 节分析当前的关键问题,并展望未来可能的研究方向.第 4 节对全文进行总结.

1 覆盖路由主要结构和方法

1.1 概 述

覆盖路由具有很强的灵活性,可以针对应用环境和需求采用特定的路由方法,因此当前有许多不同的覆盖路由结构和方法.Internet 的 TCP/IP 分层模型在物理层上有网络层、传输层和应用层,每一层关注特定的需求目标.本文按照这种层次和相应的需求目标,对当前主要的覆盖路由方法进行了分类.重点关注覆盖网络节点之间路径的可用性和性能(丢包率、延迟)的覆盖路由划分在网络层面;主要目标是提高覆盖网络节点之间传输层会话的吞吐率和可靠性的覆盖路由划分在传输层面.还有一些覆盖路由主要是应用层资源的查找定位和分发,比如 P2P 覆盖网络和应用层组播网络中的资源查找分发模式,它们可以看作应用层面上的覆盖路由.这种分类方法的优点是可以比较全面地覆盖已有的和未来可能的覆盖路由,但是并不十分严格.有些覆盖路由可能需要考虑多个层面上的问题,这时可以根据其所针对的重点问题所在层面进行归类.表 1 列举了主要覆盖路由方法的分类情况.

© 中国科学院软件研究所 http://www.c-s-a.org.cn

Table 1 Taxonomy of overlay routing methods

表 1 覆盖路由方法的分类

Category	Methods	Purpose
Network layer	Resilient overlay routing	Fast route failure detect and recovery
	Routing underlay for overlay network	Aggregate Internet route and topology information to serve upper overlay routing
	QoS overlay routing	provide QoS for overlay network or Internet
Transport layer	Overlay routing for TCP	Improve the throughput of TCP transport
	Overlay routing for UDP	Improve the reliability of UDP transport
Application layer	P2P overlay application layer multicast	Application resource querying, locating and dissemination

本文重点考查改善覆盖网络或者 Internet 端到端性能的覆盖路由相关研究,对应于分类中的网络层和传输层.而应用层面的应用层组播和 P2P 覆盖中的覆盖路由关注的共享资源的查找分发在本文中不作具体讨论.文献[1,2]分别对 P2P 覆盖网络和应用层组播的资源定位分发模式进行了详细的介绍和讨论.

下面将按照表 1 中的分类对主要覆盖路由结构和方法进行具体的介绍和分析.在每一层,将针对更具体的目标和问题进一步分类.

1.2 网络层面的覆盖路由

该层面的覆盖路由关注覆盖节点之间路径的可用性和性能.也有些研究用它来提高互联网端到端的服务性能,为互联网网络层提供 QoS 支持等.下面介绍相关的研究.

1.2.1 弹性覆盖路由

- 弹性覆盖网络

这里,弹性覆盖网络是指以 RON(resilient overlay network)为代表的覆盖网络,我们把其中的路由称为弹性覆盖路由.其目的是实现路径故障快速检测和恢复,这也是弹性的最初含义.另外,也可以通过覆盖节点绕过拥塞路径,提高互联网端到端路由路径性能.重要的相关研究有 Detour 和 RON.

Washington大学早在 1998 年前后就提出了Detour[3]项目.Savage等人[4]在Internet上测量表明,有 30%~80%的情况,通过中间节点的间接路径的RTT(round trip time)小于直接路径RTT,这是Detour的动机来源之一.Detour在网络上部署覆盖节点,称为智能网关,它们实时探测彼此之间在底层网络的延迟、丢包率等性能.智能网关之间通过隧道转发数据,绕过故障路径.智能网关也可以根据数据包的传输层信息,做“第 4 层”的路由,对突发流量做流量整形等.Detour规模小,功能复杂,开销较大.

Andersen等人在文献[5]中提出了RON网络结构,主要是为了网络路径的快速故障检测和恢复.RON周期性地测量覆盖节点之间虚拟链路的性能,及时发现故障.这些信息存储在数据库中,由各个节点查询使用.节点之间通过类似于OSPF(open shortest path first)的链路状态路由协议把拓扑和路径信息分发到每个邻居节点.RON节点之间通过IP隧道来转发流量,绕过故障链路.实验中发现,采用仅经过 1 个中间节点的单跳间接路由,在往返延迟、丢包率以及TCP吞吐率方面可以获得大部分的收益,这一点在后来的相关研究中[6]也得到验证和采用.然而,路径性能周期探测、路由和性能信息分发等开销不利于RON大规模部署.

有些研究在 RON 基础上,把覆盖网络和 multi-homing 结合起来,称为多穴覆盖网络 MON(multi-homed overlay network).Multi-Homing 加强了覆盖节点之间的路径差异,有利于提高弹性覆盖路由的收益.Andersen 等人在文献[7]中把 Web 代理覆盖网络和 multi-homing 结合起来,在客户和 Web 代理之间提供多条访问路径,从而提高 Web 服务的访问质量.Zhu 等人在文献[8]中提出了更具一般意义的 MON 结构.MON 可以由专门的服务商 OSP(overlay service provider)来运营,以提供比本地 ISP 更优的网络访问服务.然而 MON 节点的部署是 NP 难问题,文献[8]中提出了 4 个启发式设计原则.Akellazai 等人在文献[9]中对弹性覆盖路由和 multi-homing 可以带来的性能收益进行了比较.通过互联网范围的测量和分析比较,他们认为 multi-homing 连接 3~4 个精心选择的 ISP 即可获得覆盖路由所能带来的收益中的大部分,覆盖路由在 multi-homing 的基础上只能带来 5%~15%的延迟提高和 1%~10%的吞吐量提高.这个结果在一定程度上肯定了 multi-homing 带来的多路径效用.

由此我们可以看出弹性覆盖网络路由的共性问题:

© 中国科学院软件研究所 http://www.c-s-a.org.cn

(1) 路径的实时探测.需要实时探测节点间虚拟链路的性能.探测周期间隔过长则不能及时获得路径信息;间隔过短则对底层网络造成较大负担[6].考查了覆盖路由提高端到端延迟和丢包的收益以及为此付出的开销.发现连接度和探测频度降低一半,开销可以缩减 1/4,但延迟和丢包的收益会减小 30%~40%.如何在较小的开销下准确测量或者推测出虚拟链路的性能,值得深入探索.

(2) 路由模式.这里的路由模式比较灵活.比如,RON 采用了只经过 1 个中间节点的单跳间接路由.与 Internet 上采用最短路径算法的 OSPF,BGP 等路由协议不同.这里除了考虑跳数最短之外,还要考虑路径性能等信息.

(3) 路由信息更新.拓扑信息和虚拟链路状态信息可以分布存储或者集中存储.前者需要较大的更新分发开销,后者的开销来自每个节点的更新和查询请求.

(4) 网络规模.随着覆盖网络规模的增大,带来的收益将会增加,但是运行开销也会增大.在一定范围的底层网络上,部署多大规模的这种网络,其可能带来的增益和开销比值达到最佳,需要深入探索.

- 路径选择问题

弹性覆盖路由是从源端和目的端节点之间存在的冗余路径中选择一条“更优”的路径,利用路径差异绕过故障或拥塞位置,或者获得性能上的提高.如何选择“更优”路径是一个关键问题.最基本的方法是采集覆盖网络中所有链路的延迟情况,使用延迟作为权重,计算源和目的节点之间的最短路径.这种方法可以选出接近最优的路径,但是显然这样开销很大,不能对网络变化及时反应.下面是一些相关的路径选择方法.

Random-*K*:Gummadi 等人在文献[10]中提出了一种单跳源路由技术.在 RON 中就已经发现,经过 1 跳中间节点可以获得大部分收益.关于如何选取中间一跳节点,他们提出了 Random-*K* 方法.主要过程是在候选节点集合中随机选择 *K* 个节点,分别通过它们向目的端转发数据包,选择最先返回响应的那个节点作为中间一跳.在候选节点集合位置一定时,Random-*K* 在延迟和路径故障避免上可以获得最优情况的大部分收益.作者认为增加一些启发式信息,Random-*K* 还可以进一步改进.

Clustering-Based 启发式方法:覆盖网络路径在底层 IP 路径上的重叠会限制故障绕避和性能提升的能力,文献[11]提出了部署和选择覆盖节点的 Clustering-Based 启发式方法.尽量增大覆盖网络路径在底层 IP 路径上的差异,从而提高故障绕避和性能提升的能力.在这个方法中,设 $I(s,t)$表示通过覆盖节点 I 的从源节点 s 到目的节点 t 的 IP 路径和直接从 s 到 t 的 IP 路径之间的路由器重叠数目,用来衡量通过覆盖节点 I 的间接路由与直接路由之间的路径差异.变量 I 和 J 的相关系数定义为节点 I 和节点 J 引起的路由差异模式的近似性.该值超过一定阈值的覆盖节点聚为一类.在同一个 ISP 中选择部署覆盖节点的位置时,只要在每个聚类中随机选择一个点.调整阈值,可以改变聚类的数量,也改变了可以选择的覆盖节点的数量.作者把这个方法成为 clustering-based 启发式方法.这种方法需要静态的全局拓扑信息,适用于比较稳定的拓扑环境,比如在大范围网络内规划部署覆盖节点.

Earliest-Diversity 启发式规则:文献[12]提出了 Earliest-Diversity 启发式方法.与上述方法类似,该方法基于两个直观假设:1) 路径重叠越小,路径的相关性越小;2) 用 AS 级路径的重叠程度来近似 IP 级路径的重叠程度.测量经验表明,上述假设在实际大多数情况下成立.该启发式规则很简单:选择最早路径分离节点作为下一跳,这也是基于直觉上的假设:与原路径越早分离的路径,从分离点开始的那部分路径与原路径发生重合也越迟.文中对该规则进行加强提出了 EDD(earliest-diversity with delay)方法,选择延迟较大,或者说离原路径较远的最早分离节点.其暗含的直观假设是:如果备选路径远离原路径,那么它与原路径很快重合回来的可能性减小.实验结果表明,该方法优于 Random-*K* 方法,接近最优路径选择.

选择“更优”路径时获得的底层路径信息越多,越容易逼近更优的目标,但是开销也会越大.因此,这里有一个开销和收益的折衷,针对具体的环境和需求可以采用不同的方法.另外,上述方法主要考虑了路径差异,以便有效地绕过故障.如果把性能也考虑进来,在多个需求目标上高效地选择最优则是比较困难的.

1.2.2 支持 QoS 的覆盖路由

目前Internet上QoS方法主要考虑转发网关的队列调度,不会改变数据包转发路径.Stoica曾提出SCORE网络结构[13],用类似主动网络的方法实现QoS支持,但是SCORE也没有改变转发路径.而覆盖路由可以通过改变数

© 中国科学院软件研究所 http://www.c-s-a.org.cn

据包转发路径,为覆盖网络应用本身或者Internet提供一定的QoS支持.但是,覆盖网络的规模对QoS作用有一定局限性.

Subramanian 等人在文献[14]中提出了基于覆盖网络来提高 Internet 的 QoS 结构 OverQoS.OverQoS 只是在覆盖节点之间的虚拟链路上提供了 QoS 保证,称为 CLVL(controlled-loss virtual link),节点间的路由还要结合其他机制,比如 RON 等.CLVL 使用 FEC(forward error correction)和 ARQ 包重传机制限制丢包率在指定的界限之内,并且覆盖节点使用 CLVL 控制带宽和丢包在单个流之间的分配.这样为覆盖节点之间的汇聚流量提供了统计上的 QoS 保证.

Li 等人在文献[15]中提出了 QoS 感知路由的覆盖网络 QRON(QoS-aware routing in overlay networks).它的节点称为 OB(overlay bokers),每个 OB 上的协议栈在传输层和应用层之间增加了 overlay service 层,为上层覆盖应用提供 QoS 支持.它提供的功能主要有:

- 拓扑发现:OB 之间连接拓扑的探测.
- 性能测量:使用 Ping 或者 Sting 探测 OB 之间虚拟链路的性能.
- 资源分配:使用 Opus 的基于效用的资源分配方法根据 QoS 需求为覆盖应用分配资源.Opus 的模型目标是最大化全局效用.QRON 根据用户的服务级别协定(service level agreement,简称 SLA)和 Opus 模型作出分配决定.
- QoS 服务供应:由 OB 节点为上层特定应用的覆盖应用提供 QoS 路径.

与 OverQoS 架构不同,QRON 有自己的路由模式.它提出了 MSDP(modified shortest distance path)和 PBSP(proportional bandwidth shortest path)路由算法.

设 OB 节点 i 和 j 之间的链路为 Lij,Bij 表示 Lij 的可用带宽,Cj 表示链路 Lij 另一端 OB 节点 j 的计算能力,RB 表示请求带宽,RC 表示请求计算开销,则 MSDP 使用 $\max\left(\frac{RB}{B_{ij}},\frac{RC}{C_j}\right)$ 计算链路 Lij 的权重,然后使用 Dijkstra 最短路径算法计算最优路径;PBSP 与 MSDP 类似,只是链路权重定义为 $\left(\frac{B_{ij}}{B_{ij}-RB}\times\frac{C_j}{C_{ij}-RC}\right)$. 它们在满足 QoS 需求的同时都注意到均衡覆盖链路的流量和 OB 节点计算开销,各有优势.

QRON与OverQoS相比,更充分地利用了覆盖路由灵活的路径选择带来的性能收益.国内有研究在QRON的基础上作了改进[16],在权重中考虑进了OB节点之间底层物理链路的数量,取得了一定的效果.

1.2.3 路由铺垫(routing underlay)

覆盖网络在提供特定于应用的服务性能时,需要探测底层网络状态,如果每个网络都要进行探测,重复性的探测必然增加底层网络的负载.Nakao等人[17]提出在覆盖网络之下建立共享的路由铺垫,它位于Internet网络层和覆盖网络之间.它的功能是实时探测底层网络,为上层覆盖网络提供拓扑、路径性能等信息,上层覆盖网络根据这些信息作出特定的路由决策.

路由铺垫并不执行覆盖路由,但它为覆盖路由提供辅助服务,涉及到底层路由信息的采集和控制,所以这里把它归为网络层的覆盖路由分类.

路由铺垫提供的服务主要有:发现到一个节点最近的邻居;发现两个节点之间的多个分离路径;建立路由 mesh 网络等.下面是该框架提供的主要服务原语:

- *GetGraph*():构建粗粒度(AS 级)的 Internet 拓扑图.
- *GetPath*(*src*,*dst*):返回从 *src* 到 *dst* 的数据包穿越的 AS 路径.
- *GetDistance*(*target*,*metric*):从本地节点到远程目标节点的距离,指标可以是 AS 的条数、路由器的跳数或者 RTT.
- *DisjointPaths*(u,v,N,k):对于给定的覆盖节点 u 和 k 以及候选中间节点的集合 N,返回从 u 到 v 的,经过 N 中一个节点的,其穿越的 AS 路径与从 u 到 v 穿越的默认 AS 路径不同的路径.
- *BuildMesh*(N):对于给定的覆盖节点集合 N,返回 N 中节点构建的 mesh 网络中本地节点的邻居集合.

© 中国科学院软件研究所 http://www.c-s-a.org.cn

路由铺垫带来的好处是满足每个特定于应用的覆盖网络对底层网络信息采集的需求,避免了大量重复的探测.但是,底层网络的拓扑、路由以及路径性能信息的采集、查询都会带来开销,这就限制了路由铺垫的效率和扩展性.

1.3 传输层面的覆盖路由

Internet 的传输层只能遵循网络层默认的唯一路径进行传输.而在覆盖网络中,覆盖节点可以作为传输层上的转发网关,利用覆盖节点间的冗余路径,选择单一最优路径或者同时在多个路径上进行传输,以提高端到端传输的可靠性和吞吐量等性能.下面根据面向连接和无连接两种传输模式,对传输层面的覆盖路由进行具体介绍.

1.3.1 面向连接的覆盖传输

TCP 是面向连接的传输协议,它使用丢包重传机制保证传输的可靠性.传输过程中通过拥塞控制、速率控制机制调节发送速率,以合理而高效地使用网络资源.如何提高 TCP 的传输效率,并且保证带宽使用的公平性和拥塞控制的稳定性,仍然在不断深入地研究.覆盖路由为改善 TCP 性能提供了一条途径.这里分成 TCP 的多路径传输和 TCP 的链式传输两类来介绍.

- TCP 的多路径传输

这类方法把通信节点之间的 TCP 连接分成若干部分,在覆盖网络提供的多条可用路径上同时传输.如图 2 所示,从 *S* 到 *D* 的 TCP 连接被分成 3 部分在不同路径上并行传输.

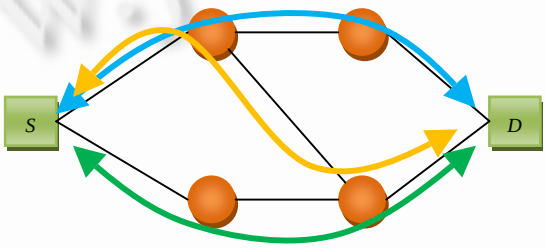

Fig.2 Multi-Path transportation of TCP connection

图 2 TCP 连接的多路径传输

Kokku等人在文献[18]中提出了采用TCP多路径传输的覆盖网络HARP.HARP的提出是为了解决网络中实时性较强的前景传输和数据量大但实时性要求不高的背景传输之间相互干扰的矛盾.它在网络边缘靠近用户的地方部署HARP入口节点,在网络中间部署中继节点,对TCP连接采用多路径传输.前景流量优先传输,背景传输通过冗余路径分散到网络中空闲的区域,从而利用路径差异、负载差异和空闲容量调节前景和背景传输的矛盾.它的TCP多路径传输控制是以KV多路径路由和拥塞控制机制[19]以及TCP Nice[20]为基础,使用流控和路由联合优化,从而提高网络空闲资源的使用效用和公平性.

TCP 的多路径传输的路径选择需要考虑的问题与网络层不同.网络层主要考虑选择符合路由策略的最短路径或者性能较优的路径;而这里主要考虑多路径的拥塞和流速控制机制、控制的稳定性、带宽使用的公平性等.这方面有许多相关研究.

SCTP[21]为TCP多路径传输提供了简单的支持.它把一条flow分成多个stream进行传输.在每个stream内给数据包编号,当在一个stream出现包乱序而等待下一序号的数据包时,其他stream仍然可以继续传输.Iyengar在文献[22]中提出了multi-homing环境下多路径传输机制CMT.它以SCTP为基础,并针对以下问题作了改进:1) 发送者不必要的快速重传;2) 发送者过于保守的拥塞窗口增长;3) 因为被接受者延迟的ack而增加的ack流量.pTCP[23]和RMTP[24]研究了如何跨越多条路径传输数据来获得多路径的汇聚带宽.其中,RMTP为了能够在动态环境中选择较优的多个信道,周期地探测信道,估算带宽和拥塞情况.而pTCP只是考虑在固定的路径上如何

© 中国科学院软件研究所 http://www.c-s-a.org.cn

分配传输流量来获得最优的汇聚带宽.mTCP[25]的研究则更进一步,在获得多条路径汇聚带宽的同时,考虑了路径故障恢复和对共享拥塞链路可用带宽的公平使用.然而,mTCP并没有提供全网范围的资源公平分配,拥有较多访问路径的用户可以占用较多的网络资源.解决这个问题要进行全局优化.

另外,文献[26]讨论了多路径情况下拥塞控制算法的稳定性.Ron Banner[27]使用优化理论考查拥塞控制,最小化多路径上的拥塞.一些研究[28,29]给出了多路径情况下的包乱序的分析模型和评价方法,如何减少多路径情况下包乱序还有待探索.

- TCP 的链式传输

这是把TCP连接分成若干部分,在多条路径上并发传输.文献[30,31]提出通过覆盖网络的中间节点,把端到端大延迟的 TCP 连接分成若干段小延迟的子连接.如图 3 所示,从 *S* 到 *D* 的 TCP 连接由端节点(*S* 和 *D*)和覆盖节点(*a*,*b*,*c*,*d*)之间的 TCP 子连接前后相接组成,称为 TCP 链式传输.传输链每段的延迟小于端到端连接的延迟,相应每段的吞吐量要大于端到端连接的吞吐量,从而提高端到端吞吐量.链式传输的瓶颈是延迟最大的子连接.

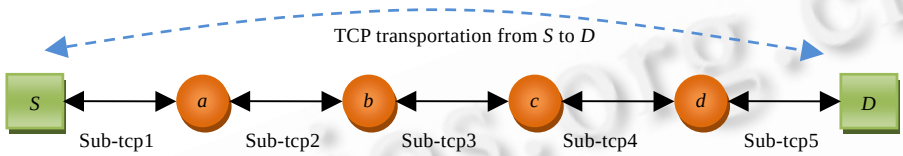

Fig.3 Chain-Style transportation of TCP connection

图 3 TCP 连接的链式传输

这种链式传输可以减小抖动和丢包损耗,绕避故障路径,提高端到端传输性能.然而,它也有许多新问题,比如中间节点故障会带来很大的丢包和恢复开销,速率控制和数据重传,资源占用公平性等,都增加了这种方法广泛部署的复杂性.文献[32,33]使用 TCP 的覆盖组播(overlay multicast)中也采用了类似的 TCP 链式传输.它们关心的是组播网络转发节点的可靠性和整体转发效率,对中间节点缓冲机制和下游转发方式作了改进.

1.3.2 无连接的覆盖传输

对于 VoIP、网络流媒体、在线商务等实时性要求高的应用,一般使用 UDP 无连接传输协议,无法保证可靠性.文献[34]提出了端到端可靠的隧道服务 SmartTunnel,使用覆盖路由来降低端到端路径上的突发丢包事件引起的损失,提高 UDP 这种无连接传输的可靠性.它的方法包括两部分:1) 通过覆盖节点,把端到端隧道流量分配到底层网络多条路径上同时传输;2) 计算原始数据的前向纠错码(FEC),把它与原始数据一起传输.文中使用扩展 Gilbert 模型作为丢包模型,计算在多路径间分配冗余包和数据包的路由优化方案,使每个 FEC 流量组丢包数量的期望值最小.然而无论如何,无连接传输都无法保证绝对可靠.

1.4 应用层面的覆盖路由

覆盖网络中根据资源内容进行的资源定位、存储和分发模式可被看成应用层面上的覆盖路由,比如P2P覆盖网络和应用层组播网络中的覆盖路由.P2P 覆盖网络一般用于大规模分布式应用.有CAN,Chord,Pastry等基于分布式哈希表(distributed Hash table,简称DHT)的结构化模式和Freenet,Gnutella,KaZaA等非结构化模式[1].应用层组播在不同的组管理模式、节点约束和需求等条件下,有最短路径、最小生成树、聚类分层以及P2P结构等多种模式[2,35].国内在应用层组播协议及其QoS路由方面有一些研究[36,37].应用层的覆盖路由不仅考虑路径的性能,更主要考虑如何根据应用层载荷内容进行高效地资源定位和分发.本文考查改善覆盖网络或者Internet端到端性能的覆盖路由相关研究,主要在网络层和传输层,应用层的覆盖路由在本文中不作具体讨论.

2 覆盖路由相关问题研究

在这一节,我们讨论覆盖路由的其他相关问题研究,包括:

1) 覆盖路由的影响因素.有些研究分析了覆盖网络的拓扑结构、节点部署位置对覆盖路由性能的影响,并

© 中国科学院软件研究所 http://www.c-s-a.org.cn

提出部署和构建拓扑的方法.

2) 覆盖路由的博弈.覆盖网络的流量在底层网络上不可避免地互相影响.有些研究用博弈论方法分析了互相作用的结果,并提出方法来削弱其中的不利因素.

下面将逐类介绍这些问题的研究工作.

2.1 覆盖节点部署

Roy等人在文献[38]中考虑了给定M个可能的位置和K个覆盖节点,如何部署节点优化特定应用的覆盖路由的性能.他们主要结合SOSR弹性路由[10]和SLOT[39]中TCP吞吐量的两个具体环境进行了分析,证明该部署问题是一个NP难问题,并提出可增量部署的启发式方法,主要分为以下 3 类:

不知情的完全随机选择:此类方法完全不知道拓扑信息、客户端和服务器之间的流量等网络环境,节点的部署位置完全随机选取.该方法可作为所有部署方法的下界.

部分知情的启发式方法:此类方法只依据网络的静态属性信息,比如网络拓扑,不知道网络中的动态变化.他们提出了称为 NDB(node degree based)的启发式方法.对于 SLOT,该方法选择高度数节点作为部署位置,因为这样的节点连通性强,更有利于作为 TCP 链的中继节点.而对于 SOSR 则选择度数低的节点,因为度数高的节点会增加路径重叠的可能性.

完全知情的启发式方法:完全知情是指可以获得有关流量、延迟、拓扑等全面的网络状态信息.在此类中,他们提出了 TAG(traffic aware greedy)方法.它在每一步迭代中,从余下的候选节点中选择可以使目标函数取得最优值的节点作为部署位置.该算法遍历了所有可能的结果,并取其中的最优结果,是所有渐增部署算法的理想上界,复杂度很高.TAG-p 方法是在 TAG 基础上进行的扩展.它在每一步迭代中,从余下的候选节点中随机选择一部分,比如 p 个,计算并选择其中使得目标函数取最优值的节点.显然,该方法是对性能和开销的折衷.

2.2 覆盖拓扑结构和构建

覆盖节点在逻辑上的邻接关系构成了覆盖网络的拓扑结构.它会对覆盖路由产生约束.覆盖网络的拓扑结构对覆盖路由有怎样的影响、如何构建有效的拓扑结构?下面介绍此类问题的相关研究.

文献[40]中提出 Cost-based 模型来刻画覆盖拓扑的结构特性,可以用来评价和比较覆盖拓扑.模型主要考查一个节点加入覆盖网络必须花费的开销.开销函数的参数包括节点的连通性(节点度和该节点到其他节点的距离)以及经过这个节点的流量.文中用这个模型分析了 Full-Mesh 和星型拓扑结构达到全局最优状态和纳什均衡状态需要满足的特征,并评价了 Chord 环、CAN、De Bruijn 图等 P2P 网络的拓扑结构.作者发现这些拓扑结构会潜在地引发大范围节点负载的不均衡.

Li 等人在文献[41]中研究了覆盖网络中拓扑结构对覆盖路由性能收益和开销的影响.他们基于故障恢复率、覆盖路径跳数惩罚、覆盖路由开销这些指标,考查了 Full-Mesh、最小生成树等拓扑结构对弹性覆盖路由性能的影响,得到如下结论:

- 覆盖拓扑对覆盖路由性能确实有重要影响.
- 如果考虑底层网络信息构造覆盖拓扑,则可以获得更好的性能
- 邻接连接(adjacent-connection)拓扑和 Topology-aware-K 最小生成树拓扑能够提供更高的路由性能.

上述分析是在模拟环境下进行的,还需要在实际的 Internet 中进行评价.Li 等人认为,拓扑构建问题的建模以及随底层网络环境变化动态构建拓扑的方法这些方面值得继续探索.

Chun 等人在文献[42]中分析了如果覆盖节点自私地选择链路构建拓扑,其博弈结果产生的覆盖路由网络的拓扑特性.作者发现改变开销函数,控制最大节点度约束,自私节点的博弈能够产生不同特性的拓扑结构.降低最大节点度约束,提高节点度分布均衡性,能够提高拓扑的弹性(resilience),但是会增大路由路径的长度(付出延展代价).这种在路由效率和弹性之间的权衡也说明了拓扑结构对路由性能的影响.

文献[43,44]提出了广域网范围的覆盖网络的拓扑构建方法.它们利用了底层网络信息,并且考虑了构建开销和性能之间的均衡,尽量用粗粒度的信息构建较优的拓扑.

© 中国科学院软件研究所 http://www.c-s-a.org.cn

Xu等人在文献[43]中分析认为,基于DHT的覆盖网络,如CAN,Pastry,Chord是在应用层面上构建的逻辑拓扑结构,不能反应和利用底层网络存在的转发容量、可用性、负载方面的区域差异.他们以Brocade[45]为基础,提出构建一种次覆盖网络层,称为Expressway.与Brocade不同的是,Expressway使用了底层网络AS级拓扑信息.他们提出两种构建方法:1) 直接使用AS级拓扑信息;2) 使用基于路标节点的聚合(landmark clustering).Expressway的拓扑结构可以根据节点的加入和退出而动态改变,更好地适应底层网络的变化.

在文献[44]中,Ratnasamy 等人提出了 Binning-based 方法.该方法把覆盖节点分成若干个 bins 区域,在同一个 bins 的节点,按照网络延迟关系,彼此互相靠近.从每个 bins 中选择一个节点作为路标来代表这个 bins 的位置.他们把这个策略应用到覆盖网络的构建和服务器节点选择上.实验结果表明,应用可以根据 binning 框架提供粗粒度的拓扑信息而获得较大的性能提升.

2.3 覆盖路由中博弈

覆盖网络的流量和底层网络的流量不可避免地发生相互作用,因此,覆盖路由和底层网络的流量工程及其他覆盖网络的路由会互相影响.有许多研究使用博弈论的方法对此进行了建模和分析.然而这些分析大多基于简单的模型讨论,实际中的相互作用将更为复杂.

2.3.1 覆盖路由之间的博弈

Jiang等人[46]使用非合作博弈的模型,研究了当多个覆盖网络共存时,覆盖网络内的路由策略在底层共享网络上互相作用的情况,以及这种作用对网络稳定性、性能、以及资源分配的公平性可能带来的影响.他们以基于流量的加权平均延迟为指标,分析了底层网络全局优化路由策略和覆盖节点自私路由策略.结果表明,全局优化路由可以使网络整体延迟最低.而自私路由单个节点只考虑自己获得最优性能,这些自私节点间的博弈在达到纳什均衡点时,不仅使底层网络整体延迟最大,而且每个覆盖网络的整体延迟也最大,性能最差.Jiang等人把覆盖网络整体看作一个自私的对象,提出了以覆盖网络整体性能作为优化目标的覆盖优化路由策略,它实际上是全局优化路由和自私路由策略之间的折衷.他们建立了相应的约束凸优化问题模型来求解该策略的纳什均衡点.结果表明,使用该路由策略达到纳什均衡时,底层网络整体延迟优于自私路由,接近全局优化路由,覆盖网络的整体延迟也没有大幅度的降低,但是它会引起覆盖网络之间资源分配不均,而且在纳什均衡点不一定是Pareto最优.Jiang等人引入流量计价因素来消减资源分配不均衡问题.

Jiang 等人只是分析了共享底层网络链路的多个覆盖网络路由之间的博弈,但是多个覆盖网络还可能通过底层网络的流量工程发生间接的相互作用.这个过程更加复杂,有待继续探索.

2.3.2 覆盖路由和底层网络的博弈

Qiu 等人在文献[47]中研究了在 Internet-like 环境下,使用覆盖路由或者源路由的自私路由对网络性能和流量工程的影响.结果发现,自私路由可以与其他覆盖网络流量和常规 IP 流量较好地共存,但是会导致某些链路的过载;并且自私路由与流量工程之间相互影响,引起网络性能的波动和下降.Liu 等人在文献[48]中使用非合作非零和二人博弈方法,以及在实际单一 AS 内的实验考查了覆盖路由和 MPLS 流量工程的互相作用,也发现覆盖路由的自私行为导致底层网络整体性能的降低.

Seetharaman 等人在文献[49]中指出,底层网络和覆盖网络独立的动态路由行为的互相作用将会造成覆盖网络出现次优路由和路由震荡,并提出 Layer-Aware 方案来减轻这些问题.顾名思义,该方法是两个层面根据对方的路由行为作出自己的路由决策.Seetharaman 等人提出了 3 种 layer-aware 策略:1) 概率抑制的覆盖重路由:以概率 p 抑制覆盖独自进行的重路由,以减少可能的路由震荡.2) 推迟覆盖重路由:通过推迟覆盖重路由一段固定时间使得本地底层网络有机会避免或者恢复故障.3) 跟随抑制覆盖重路由:这需要覆盖层面了解本地底层网络的路由 Keep-Alive 时钟周期.如果覆盖网络当前检测到故障,并且它又知道底层网络的 Keep-Alive 时钟紧跟其后,如果这个时间间隔小于某个阈值,则覆盖层面的重路由被抑制.上述多种方法对各种重路由性能指标之间的权衡控制提供了很好的灵活性.

Seetharaman 等人在文献[50]中进一步研究覆盖网络和底层网络流量工程之间的互相作用并提出策略,使得博弈过程向提高某一个层面的网络性能并稳定的方向进行.它把一个层面看成主导层,把其他层面看成跟随

© 中国科学院软件研究所 http://www.c-s-a.org.cn

层.主导层面可以预测跟随层面的反应,采取对策阻止性能下降,策略分为两种:1) 不引发对方路由变化的友好型策略;2) 迫使对方作出改变的敌对型策略.在不同网络特征下的模拟发现,主导层采用抢先策略可以获得接近最优的性能;并且只有当其他层面完全受到侵犯时,某一特定层面才能达到最优性能.

3 研究展望

覆盖路由及其相关研究虽然取得了很多进展,但如果广泛使用,在许多方面还需要继续深入地探索:1) 覆盖路由所能带来的收益和开销还不清楚.对于ISP,部署覆盖路由带来的性价比与其他方法,比如扩展网络硬件能力相比,是否具有竞争力.2) 覆盖路由需要覆盖网络对底层网络进行测量,在覆盖规模较大时,开销也很大,如何降低测量开销,同时满足精度需求,仍需在互联网测量方法上继续探索.3) 覆盖路由的博弈会造成整体上资源使用的不平衡和不充分.在寻求优化自私目标的行为策略的同时,如何阻止网络整体的不稳定和性能下降,是值得深入探索的.4) 覆盖路由的安全问题.虽然有些研究使用覆盖网络保护目标系统避免遭受DDoS攻击[51,52],但是覆盖网络的路由系统本身需要考虑DDoS攻击、假冒攻击等安全问题.

覆盖路由对 Internet 路由结构的改善也有一定的启发:1) 互联网路由拓扑信息和决策的集中控制.文献[17]中的路由铺垫层把底层网络拓扑信息采集和覆盖路由决策分开,进一步认为可以把路由拓扑和决策集中到专门的层面,而 BGP 只提供可达信息.这种控制层面的集中也许有利于增强路由的稳定性,但是能否及时响应路径故障和路由变化,是否有利于路由收敛,还需要深入探索.2) 域间路由和覆盖路由的融合,提供多层面的路由服务.文献[53]把覆盖路由与 IP 层路由结合起来,用这种多层面路由的交融提高覆盖路由的扩展能力和 IP 路由的可用性.文献[54,55]分别研究了如何利用域间的冗余路径提供灵活的路由策略,以及提高域间路由的健壮性.

4 总 结

本文从 TCP/IP 分层模型的角度考查了覆盖路由结构和方法以及相关问题的研究进展.在网络层介绍了弹性覆盖路由和支持 QoS 的覆盖路由方法;在传输层介绍了 TCP 多路径和链式传输方法,以及提高 UDP 传输可靠性的方法.

随后,本文总结了覆盖路由相关问题的研究进展.有些研究工作分析了覆盖网络拓扑结构、节点部署等因素对覆盖路由的影响,并提出了一些改善方法.另外,有些研究使用博弈论方法分析了覆盖路由在底层网络上互相作用,并作出改进,以提高路由行为的稳定性,有利于资源的公平使用.

随着 Internet 的不断发展,边缘接入方式以及 ISP 连接关系越来越复杂.覆盖路由方法对于改善覆盖网络或 Internet 端到端的性能,或者改进 Internet 路由体系结构,是一条值得探索的途径.

致谢 感谢清华大学信息网络工程研究中心网络体系结构与 IPv6 研究室各同学的帮助和支持,也感谢网络运行与管理技术研究室杨家海教授和同学的支持.

References:

[1] Lua EK, Crowcroft J, Pias M, Sharma R, Lim S. A survey and comparison of peer-to-peer overlay network schemes. IEEE Communications Survey and Tutorial, 2005,7(2):72−93.

[2] Hosseini M, Ahmed DT, Shirmohammadi S, Georganas ND. A survey of application-layer multicast protocols. IEEE Communications Surveys and Tutorials, 2007,9(3):58−74.

[3] Savage S, Anderson T, Aggarwal A, Becker D, Cardwell N, Collins A, Hoffman E, Snell J, Vahdat A, Voelker G, Zahorjan J. Detour: A case for informed internet routing and transport. IEEE Micro, 1999,19(1):50−59.

[4] Savage S, Collins A, Hoffman E, Snell J, Anderson TE. The end-to-end effects of Internet path selection. ACM SIGCOMM Computer Communication Review, 1999,29(4):289−299.

[5] Andersen DG, Balakrishnan H, Kaashoek MF, Morris R. Resilient overlay networks. In: Marzullo K, Satyanarayanan M, eds. Proc. of the 18th ACM Symp. on Operating Systems Principles (SOSP 2001). New York: ACM, 2001. 131−145.

© 中国科学院软件研究所 http://www.c-s-a.org.cn

[6] Rewaskar S, Kaur J. Testing the scalability of overlay routing infrastructures. In: Chadi B, Ian P, eds. Proc. of the 5th Int'l Workshop on Passive and Active Network Measurement (PAM 2004). Berlin: Springer-Verlag, 2004. 33−42.

[7] Andersen DG, Balakrishnan H, Kaashoek MF, Rao RN. Improving Web availability for clients with MONET. In: Vahdat A, Wetherall D, eds. Proc. of the 2nd Conf. on Symp. on Networked Systems Design & Implementation (NSDI 2005). Berkeley: USENIX Association, 2005. 115−128.

[8] Zhu Y, Dovrolis C, Ammar M. Combining multihoming with overlay routing (or, how to be a better ISP without owning a network). In: Proc. of the IEEE INFOCOM 2007. Piscataway: IEEE, 2007. 839−847.

[9] Akella A, Pang J, Maggs B, Seshan S, Shaikh A. A comparison of overlay routing and multihoming route control. In: Yavatkar R, Zegura EW, Rexford J, eds. Proc. of the ACM SIGCOMM 2004. New York: ACM, 2004. 93−106.

[10] Gummadi KP, Madhyastha HV, Gribble SD, Levy HM, Wetherall D. Improving the reliability of Internet paths with one-hop source routing. In: Brewer E, Chen P, eds. Proc. of the 6th Symp. on Operating System Design and Implementation (USENIX OSDI 2004). Berkeley: USENIX Association, 2004. 183−197.

[11] Han J, Watson D, Jahanian F. Topology aware overlay networks. In: Makki K, Knightly E, eds. Proc. of the IEEE INFOCOM 2005. Piscataway: IEEE, 2005. 2554−2565.

[12] Fei T, Tao S, Gao LX, Guérin R. How to select a good alternate path in large peer-to-peer systems? In: Domingo-Pascual J, ed. Proc. of the IEEE INFOCOM 2006. Piscataway: IEEE, 2006. 1−13.

[13] Stoica I. Stateless core: A scalable approach for quality of service in the Internet [Ph.D. Thesis]. Pittsburgh: Carnegie Mellon University, 2000. http://www.cs.berkeley.edu/~istoica/thesis.pdf

[14] Subramanian L, Stoica I, Balakrishnan H, Katz RH. OverQoS: An overlay based architecture for enhancing Internet QoS. In: Proc. of the 1st Symp. on Networked System Design and Implementation (NSDI 2004). Berkeley: USENIX Association, 2004. 71−84.

[15] Li Z, Mohapatra P. QRON: QoS-Aware routing in overlay networks. IEEE Journal on Selected Areas in Communications, 2004, 22(1):29−40.

[16] Ying BD, Chen HF, Zhu SB, Qiu PL. RBLCP: Improved QoS-aware routing algorithm for overlay networks. Journal of Zhejiang University (Engineering Science), 2007,41(4):646−649 (in Chinese with English abstract).

[17] Nakao A, Peterson L, Bavier A. A routing underlay for overlay networks. In: Proc. of the ACM SIGCOMM 2003. New York: ACM, 2003. 11−18.

[18] Kokku R, Bohra A, Ganguly S, Venkataramani A. A multipath background network architecture. In: Proc. of the IEEE INFOCOM 2007. Piscataway: IEEE, 2007. 1352−1360.

[19] Kelly F, Voice T. Stability of end-to-end algorithms for joint routing and rate control. ACM SIGCOMM Computer Communication Review, 2005,35(2):5−12.

[20] Venkatramani A, Kokku R, Dahlin M. TCP-Nice: A mechanism for background transfers. In: Proc. of the 5th Symp. on Operating System Design and Implementation (OSDI 2002). Berkeley: USENIX Association, 2002. 329−343.

[21] Zhang L, Paxson V. Stream control transmission protocol. RFC2960, 2000. http://www.ietf.org/rfc/rfc2960.txt

[22] Iyengar JR, Amer PD, Stewart R. Concurrent multipath transfer using SCTP multihoming over independent end-to-end paths. IEEE/ACM Trans. on Networking, 2006,14(5):951−964.

[23] Hsieh HY, Sivakumar R. pTCP. An end-to-end transport layer protocol for striped connections. In: Proc. of the 10th IEEE Int'l Conf. on Network Protocols (ICNP 2002). Piscataway: IEEE, 2002. 24−33.

[24] Magalhaes L, Kravets R. Transport level mechanisms for bandwidth aggregation on mobile hosts. In: Proc. of the 9th Int'l Conf. on Network Protocols (ICNP 2001). Piscataway: IEEE, 2001. 165−171.

[25] Zhang M, Krishnamurthy A, Peterson L, Wang R. A transport layer approach for improving end-to-end performance and robustness using redundant paths. In: Proc. of the Annual Conf. on USENIX Annual Technical Conf. Berkeley: USENIX Association, 2004. 99−112.

[26] Han H, Shakkottai S, Hollot CV, Srikant R, Towsley D. Multi-Path TCP: A joint congestion control and routing scheme to exploit path diversity in the internet. IEEE/ACM Trans. on Networking, 2006,16(6):1260−1271.

[27] Banner R, Orda A. Multipath routing algorithms for congestion minimization. IEEE/ACM Trans. on Networking, 2007,15(2): 413−424.

© 中国科学院软件研究所 http://www.c-s-a.org.cn

[28] Piratla N, Jayasumana A. Reordering of packets due to multipath forwarding: An analysis. In: Proc. of the IEEE Int'l Conf. on Communications (ICC 2006). Piscataway: IEEE, 2006. 28−36.

[29] Han Y, Makowski AM. Resequencing delays under multipath routing: Asymptotics in a simple queueing model. In: Proc. of the IEEE INFOCOM 2006. Piscataway: IEEE, 2006. 1−12.

[30] Amir Y, Danilov C. Reliable communication in overlay networks. In: Proc. of the 2003 Int'l Dependable Systems and Networks (DSN 2003). Piscataway: IEEE, 2003. 511−520.

[31] Pucha H, Hu YC. Overlay TCP: Ending end-to-end transport for high throughput. In: Proc. of the ACM SIGCOMM 2005 (Poster). New York: ACM, 2005. 22−26.

[32] Kwon GI, Byers JW. ROMA: Reliable overlay multicast with loosely coupled TCP connections. In: Proc. of the IEEE INFOCOM 2004. Piscataway: IEEE, 2004. 385−395.

[33] Baccelli F, Chaintreau A, Liu Z, Riabov A. The one-to-many TCP overlay: A scalable and reliable multicast architecture. In: Makki K, Knightly E, eds. Proc. of the IEEE INFOCOM 2005. Piscataway: IEEE, 2005. 1629−1640.

[34] Li Y, Zhang Y, Qiu L, Lam S. SmartTunnel: Achieving reliability in the Internet. In: Proc. of the IEEE INFOCOM 2007. Piscataway: IEEE, 2007. 830−838.

[35] Wu JG. A survey of routing algorithms and protocols in overlay multicast. Computer Science, 2007,34(6):7−12 (in Chinese with English abstract).

[36] Wu JG, Yang YY, Chen YX, Ye XG. A QoS-aware overlay multicast routing protocol. Chinese Journal of Computers, 2005,29(11): 1937−1947 (in Chinese with English abstract).

[37] Zhang L, Yu ZW, Zhang Y, Li N. QoS-Based distributed dynamic routing of multicast overlay network. Computer Engineering, 2006,32(13):103−113 (in Chinese with English abstract).

[38] Roy S, Pucha H, Zhang Z, Hu YC, Qiu LL. Overlay node placement: Analysis, algorithms and impact on applications. In: Proc. of the 27th IEEE Int'l Conf. on Distributed Computing Systems (ICDCS 2007). Piscataway: IEEE, 2007. 25−29.

[39] Pucha H, Hu YC. Slot: Shortened loop Internet transport using overlay networks. Technical Report, TR-ECE-05-12, West Lafayette: Purdue University, 2005. http://docs.lib.purdue.edu/ecetr/66/

[40] Christin N, Chuang J. A cost-based analysis of overlay routing geometries. In: Makki K, Knightly E, eds. Proc. of the IEEE INFOCOM 2005. Piscataway: IEEE, 2005. 2566−2577.

[41] Li Z, Mohapatra P. The impact of topology on overlay routing service. In: Proc. of the IEEE INFOCOM 2004. Piscataway: IEEE, 2004. 408−418.

[42] Chun BG, Fonseca R, Stoica I, Kubiatowicz J. Characterizing selfishly constructed overlay routing networks. In: Proc. of the IEEE INFOCOM 2004. Piscataway: IEEE, 2004. 1329−1339.

[43] Xu ZC, Mahalingam M, Karlsson M. Turning heterogeneity into an advantage in overlay routing. In: Proc. of the IEEE INFOCOM 2003. Piscataway: IEEE, 2003. 1499−1509.

[44] Ratnasamy S, Handley M, Karp RM, Shenker S. Topologically-aware overlay construction and server selection. In: Proc. of the IEEE INFOCOM 2002. Piscataway: IEEE, 2002. 1190−1199.

[45] Zhao BY, Duan Y, Huang L, Joseph A, Kubiatowicz J. Brocade: Landmark routing on overlay networks. In: Druschel P, Kaashoek MF, Rowstron AIT, eds. Proc. of the 1st Int'l Workshop on Peer-to-Peer Systems (IPTPS 2002). LNCS 2429, Berlin: Springer-Verlag, 2002. 34−44.

[46] Jiang WJ, Chiu DM, Liu JCS. On the interaction of multiple overlay routing. Performance Evaluation, 2005,62(1-4):229−246.

[47] Qiu LL, Yang YR, Zhang Y, Shenker S. On selfish routing in Internet-like environments. In: Feldmann A, Zitterbart M, Crowcroft J, Wetherall D, eds. Proc. of the ACM SIGCOMM 2003. New York: ACM, 2003. 151−162.

[48] Liu Y, Zhang HG, Gong WB, Towsley DF. On the interaction between overlay routing and underlay routing. In: Makki K, Knightly E, eds. Proc. of the IEEE INFOCOM 2005. Piscataway: IEEE, 2005. 2543−2553.

[49] Seetharaman S, Ammar MH. On the interaction between dynamic routing in native and overlay layers. In: Proc. of the IEEE INFOCOM 2006. Piscataway: IEEE, 2006. 1−12.

[50] Seetharaman S, Hilt V, Hofmann M, Ammar MH. Preemptive strategies to improve routing performance of native and overlay layers. In: Proc. of the IEEE INFOCOM 2007. Piscataway: IEEE, 2007. 463−471.

© 中国科学院软件研究所 http://www.c-s-a.org.cn

[51] Shi E, Stoica I, Andersen D, Perrig A. OverDoSe: A generic DDoS protection service using an overlay network. Technical Report, CMU-CS-06-114, Pittsburgh: Carnegie Mellon University, 2006. http://www.cs.cmu.edu/~dga/papers/CMU-CS-06-114.pdf

[52] Keromytis AD, Misra V, Rubenstein D. SOS: Secure overlay services. In: Proc. of the ACM SIGCOMM 2002. New York: ACM, 2002. 61−72.

[53] Zhu YP, Rexford J, Bavier A, Feamster N. UFO: A resilient layered routing architecture. ACM SIGCOMM Computer Communication Review, 2008,38(5):59−62.

[54] Xu W, Rexford J. MIRO: Multi-Path interdomain Routing. In: Rizzo L, Anderson TE, McKeown N, eds. Proc. of the ACM SIGCOMM 2006. New York: ACM, 2006. 171−182.

[55] Kushman N, Kandula S, Katabi D, Maggs B. R-BGP: Staying connected in a connected world. In: Proc. of the 4th Symp. on Networked Systems Design and Implementation (NSDI 2007). Berkeley: USENIX Association, 2007. 341−354.

附中文参考文献:

[16] 应必娣,陈惠芳,朱孙斌,仇佩亮.RBLCP:一种覆盖网络的 QoS 路由改进算法.浙江大学学报(工学版),2007,41(4):646−649.

[35] 吴家皋.覆盖多播路由的算法及协议研究综述.计算机科学,2007,34(6):7−12.

[36] 吴家皋,杨音颖,陈益新,叶晓国.一种新的 QoS 覆盖多播路由协议的研究.计算机学报,2005,29(11):1937−1947.

[37] 张丽,余镇危,张扬,李宁.支持 QoS 的覆盖组播分布式动态路由研究.计算机工程,2006,32(13):103−113.

王旸旸(1979－),男,安徽淮北人,博士生,主要研究领域为计算机网络体系结构,网络测量.

吴建平(1953－),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为计算机网络体系结构,协议工程学,下一代互联网.

毕军(1972－),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为计算机网络体系结构,下一代互联网.

© 中国科学院软件研究所 http://www.c-s-a.org.cn