ISSN 1000-9825, CODEN RUXUEW
Journal of Software, Vol.21, No.6, June 2010, pp.1326−1340
doi: 10.3724/SP.J.1001.2010.03797
© 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

标识路由关键技术*

侯 婕+, 刘亚萍, 龚正虎

(国防科学技术大学 计算机学院,湖南 长沙 410073)

Key Techniques of Identifier-Based Routing

HOU Jie+, LIU Ya-Ping, GONG Zheng-Hu

(School of Computer, National University of Defense Technology, Changsha 410073, China)

+ Corresponding author: E-mail: houhou1983@163.com

Hou J, Liu YP, Gong ZH. Key techniques of identifier-based routing. *Journal of Software*, 2010,21(6): 1326−1340. http://www.jos.org.cn/1000-9825/3797.htm

Abstract: With the rapid development of Internet technology, the existing routing system has been confronting many serious challenges of scalability, mobility, multi-homing and traffic engineering. Based on the idea of Identifier/Locator Split, the concept of identifier-based routing is proposed, and its research scope is accurately defined in this paper. On the basis of the design goals of identifier-based routing, some related researches are introduced and compared. Finally, some key issues and the development trends are discussed.
Key words: identifier-based routing; identifier/locator split; scalability; multi-homing; traffic engineering

摘 要: 现有路由系统面临来自可扩展性、移动性、多宿主以及流量工程等方面的挑战,已经不能满足客观需求.基于位置与标识分离的思想,提出了标识路由的概念并对其研究范围进行了精确的界定.根据标识路由的设计目标,还详细介绍和比较了国内外相关的研究工作.最后提出标识路由研究的若干关键问题,并对未来的研究方向做出展望.
关键词: 标识路由;位置与标识分离;可扩展性;多宿主;流量工程
中图法分类号: TP393 文献标识码: A

随着互联网的飞速发展,路由系统在可扩展性、移动性、多宿主和流量工程、提供商锁定(lock-in)效应等方面存在许多亟待解决的问题.首先,可扩展性差是路由系统面临的最主要问题,导致该问题的直接原因是核心路由表规模的急剧增长.Internet 的 DFZ(default-free zone)目前地址前缀已经达到 30 万条[1],而且由于 PI(provider independent)地址的大量使用、多宿主及流量工程的需求以及聚合故障等原因迫使其每年还以超线速(super-linear)的速度持续增长着[2],预计在 2020 年达到 200 万条左右[3].查找和匹配这些数量巨大的地址前缀既增加了路由器的负担,又对路由器的性能提出了更高的要求.此外,Internet 频繁出现每秒接近 1 000 条地址前

* Supported by the National Natural Science Foundation of China under Grant No.90604006 (国家自然科学基金); the National Key Technology R&D Program of China under Grant Nos.2008BAH37B02, 2008BAH37B03 (国家科技支撑计划); the National High-Tech Research and Development Plan of China under Grant No.2008AA01A325 (国家高技术研究发展计划(863)); the National Basic Research Program of China under Grant No.2009CB320503 (国家重点基础研究发展计划(973)).

Received 2009-04-29; Revised 2009-07-21; Accepted 2009-12-02

© 中国科学院软件研究所 http://www.c-s-a.org.cn

缀的更新事件[4].大量路由通告和撤销消息的传播也加重了路由器的负担,同时耗费了大量的可用资源.其次,在互联网设计之初,并未考虑节点移动性的情况,使用与拓扑位置密切相关的 IP 地址同时作为节点的标识和地址信息.当节点移动时,IP 地址发生改变,通信双方无法在基于原有地址建立的连接上传送数据,从而导致通信中断.正是由于节点缺少与拓扑位置无关的固定标识,路由系统才表现出对移动性支持的先天不足.再次,为了提高接入网络的可靠性和资源的利用率,同时实现灵活的策略控制,越来越多的接入网络选择使用多宿主和流量工程技术.如今的多宿主和流量工程应用大多采用地址细分或者 PI 地址的方式,无形之中加速了路由表规模的膨胀.另外,由于多宿主技术的应用,已经向核心路由表引入了 20%~30%左右的额外前缀[5].流量工程一般用来负载平衡[6],负载均衡也已经向全球路由表中引入了 20%~25%的额外前缀[5].同时,由于接入网络在互联网服务提供商之间切换网络流量时,路由表更新报文将被发送到相应的核心网络路由器,从而增加了路由收敛时间,影响了整个网络的稳定性.最后,所谓的提供商锁定效应,也称为换制成本效应,是指企业更换ISP需要付出的代价.基于 IP 地址的标识结构缺乏稳定性,当改变 ISP,网络拓扑发生变化时,需要对企业接入网络进行重新编址.而为了避免此类开销,越来越多的企业偏好使用 PI 地址,在减少换制成本的同时却为核心路由表引入了额外的地址前缀.

针对上述问题,人们已经开始了对互联网 TCP/IP 体系结构的重新思考,认为必须对路由结构进行重新设计,才能很好地适应未来互联网的发展[7,8].互联网现在使用IP地址既表示节点的位置信息又表示节点的身份信息,混淆了位置(locator)和标识(identifier)的功能界限,也就是所谓的 IP 地址语义过载.这种命名机制将原本技术和经济目标相矛盾的用户网络和 ISP 网络紧耦合,阻碍了整个网络的发展.国内外研究者一致认为这就是当前路由系统诸多问题的症结所在.

IRTF(Internet Routing Task Force)的路由研究组 RRG(Routing Research Group)提出基于位置与标识分离(locator/identifier split)的思想来解决路由系统的问题[7,9].国际 IAB 组织提出引入 PI 地址空间和 PA(provider allocated)地址空间来分别表示节点标识和位置信息,解决 IP 地址语义过载的问题.

基于位置与标识分离的思想,研究者提出了多种方案来解决当前互联网面临的诸如路由可扩展性、主机移动、多宿主及安全等问题.其中,FARA[10]提出一种通用抽象模型,将端系统标识与网络地址分离,并定义了一组抽象对象以及这些对象之间的关系.HIP[11],NIIA[12],I3[13],ROFL[14],HLP[8]着重解决路由可扩展性问题.IPNL[15]主要解决端到端透明通信问题,实现公用网络和私有网络之间的互访;TRIAD[16]基于内容路由解决如何为用户定位一个距离最近的所需内容的复本节点;NUTSS[17]提出一种针对中间盒的管理机制,解决端到端跨中间盒通信问题;UIP[18]解决了边缘网络的管理问题;SHIM6[19]主要是为了支持多宿主;DONA[20]旨在设计新的路由系统取代现行互联网 DNS 系统;LISP[21]和 Six/One[22]则关注如何减小 BGP 路由表规模.国内的各项研究如清华大学开展的关于标识分离映射机制评估的研究[23];北京交通大学提出的 DHT-MAP[24]和 4I[25];华为公司提出的 HRA[6],RANGI[26];国防科技大学开展的关于位置与标识分离的命名和寻址体系结构研究[27]以及我们提出的可扩展标识映射系统 SILMS[28]和一种标识路由架构[29];解放军信息工程大学提出的 Hidra[30]等.

针对互联网路由系统的研究已经持续了很多年,而且一直是计算机网络领域的研究热点.国内外研究者提出了如上所述的很多机制,但是各种机制所解决问题的侧重点不同,很多机制倾向于对现有路由系统的简单修补,未能解决路由系统的本源问题,无法适应未来互联网的发展[31].

事实上,就整个互联网路由系统而言,许多问题的研究是相关联的.如何减少核心路由表规模,如何高效地支持节点移动,如何便于多宿主和流量工程技术的发展,如何避免重新编址等是亟待解决的问题.位置与标识分离后,如何利用节点标识的唯一性、持久性等特点,设计满足未来互联网发展需求的新型路由系统正在成为学术界研究的热点.

本文基于现有路由体系结构存在的问题和未来应用的需求,在第 1 节提出了标识路由的概念并界定了其研究范围.第 2 节指出了标识路由的设计目标.第 3 节对实体命名规则、ID/Locator 解析机制以及解析完成以后的路由机制进行了详细地分析总结,并对不同的模型或者机制做出了横向评价比较,同时对几种典型的体系结构从安全性、可扩展性、多宿主以及流量工程等方面进行纵向评价比较.第 4 节阐述了标识路由研究中面临的

© 中国科学院软件研究所 http://www.c-s-a.org.cn

关键问题,提出了解决问题的出发点;最后总结全文,并展望进一步的研究方向.

1 标识路由的基本描述

狭义上讲,标识路由是将数据对象的标识直接映射为路由,而不是将标识先映射为包含拓扑信息的地址,再去获取路由信息.也就是将位置与标识的解析路径和路由信息的取回路径合为一体,只研究如何使用标识进行路由.典型的机制 ROFL[14]就是使用标识完全取代地址在全网范围内路由.

然而大多数的标识路由机制,使用具有扁平特性的标识,都是在主机或者路由器完成标识与地址的映射.由于整个地址的生成和使用过程对用户和应用透明,用户和应用感觉整个过程是始终基于标识完成的路由.这便是广义上的标识路由,研究的范畴包括 ID/Locator 解析机制以及完成解析后的路由机制.图 1 为本文研究的标识路由示意图.

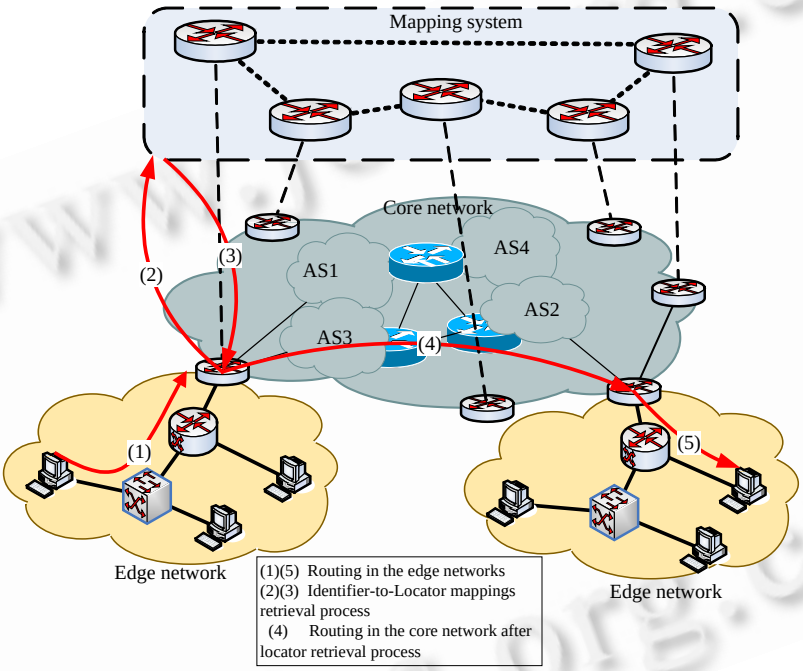

Fig.1 The sketch map of identifier-based routing

图 1 标识路由示意图

本文的研究内容属于广义的标识路由的范畴,主要包括对网络实体命名机制、ID/Locator 解析机制以及路由机制的研究.首先,研究满足新型应用需求、适合下一代互联网的名字空间,必须区分 Identifier,Locator, Address 以及 Label 这 4 个概念.标识即身份标识,是指表示网络实体身份(identity)的抽象符号[32].本文认为标识根据网络实体类型的不同可分为端标识和网络标识两种.其中,端标识通常使用在应用程序中,一般使用用户友好的字符串,可以是全局唯一的,也可以只具备局部含义,一般用于端节点.例如,端标识可以使用(用户名,域名,服务名)的三元组,如(rockie@sina.com,ftpd).网络标识通常使用与网络实体位置无关的身份标识,具有扁平特性和全局唯一性,与拓扑信息无关,当接入点改变时,Identifier 不会变化,一般用于网络节点.例如,在一个企业网中的一个不可路由且与提供商无关的 IP 前缀,就可以作为这个企业网的网络标识.Locator,一般采用 PA 地址,用来表示节点的拓扑位置信息,具有层次特性,在全网范围内有效.当接入点改变时,Locator 会随之发生改变.现有的地址 Address 具有严格的层次特性,如果属于 PA 地址空间,可被用作 Locator,而如果属于 PI 地址空间,可被用作 Identifier.在 MPLS 中,标签(label)是指用于表示转发等价类 FEC(forwarding equivalence class)的固定长度(一般为 20 比特)的标识符[33,34].它用于唯一地表示一个分组所属的 FEC,代表着不同的转发路径 LSP(label switched

© 中国科学院软件研究所 http://www.c-s-a.org.cn

path).MPLS label 是随机生成的、不包含拓扑信息、只具有局部意义的信息内容[33,34].而 IPv6 流标签(flow label)[35]不同于 MPLS 标签,由一个五元组(源地址,目的地址,源端口,目的端口,传输层协议)或者一个三元组(流标签,源地址,目的地址)构成,由发起流的节点来随机生成和分配.同时,IPv6 流标签在全网范围内有效,在数据包到达目的地之前不能改变.表 1 给出上述几个概念的主要区别.其次,对于 ID/Locator 映射关系问题,目前存在两种解决方法[36]:一是分离(separation)的方法;二是消除(elimination)的方法.分离的方法,即将边缘网络与核心网络分离,需要引入一个映射系统,负责边缘网络所使用的地址与核心网络地址之间的映射.典型的机制有 APT[37],LISP[22],Ivip[38],TRRP[39],Six/One[21],Six/One Router[40].消除的方法,即消除所有 PI 地址,将 PA 前缀细分,方便聚合.消除的方法不必改变现有路由结构,但需要改变边缘网络和主机协议栈,使其能够利用多个地址进行报文传输,同时能够检测和处理链路失效的情况.典型的机制是 SHIM6[19].此外,还有一种极端的解决方法,取消现有 IP 地址的概念,完全采用扁平标识进行路由.这种方法应该归为消除的方法.最后,对于 ID/Locator 路由机制的研究,目前可以分为基于拓扑信息的路由、基于标识的路由以及基于地理位置信息的路由.

Table 1 Comparison of some concepts

表 **1** 几个概念的比较

Concept		Scope	Length	Structure	Aggregation	Topology aware
Identifier	Endpoint id	Global/Local	Variable-Size	Hierarchical	Moderate	No
	Network id	Global	Fixed-Size	Flat	Poor	No
Locator		Global	Fixed-size	Hierarchical	Good	Yes
Address		Global/Local	Fixed-size	Hierarchical	Medium	Yes
Label		Global/Local	Fixed-size	Hierarchical/Flat	Poor	No

2 标识路由的设计目标

标识路由设计的总体目标是适应并有利于未来互联网的发展.当然,不同的设计角度具有不同的设计标准.从网络(或者 ISP)的角度考虑,设计标准包括可扩展性、稳定性、鲁棒性、可部署性以及安全性;从应用(或者用户)的角度,设计标准包括多宿主和流量工程、移动性、异构融合性以及安全性.其中,安全性是网络和用户都关心的问题.网络侧重于路由的安全性,而用户更关注网络能否提供安全的服务,以及能否有效地抵制 DDoS 攻击等.实质上,网络角度的设计标准与用户角度的设计标准存在很大的矛盾性.因此,在设计新的路由机制时必须根据需求统筹考虑.

3 标识路由的研究现状

3.1 互联网实体命名规则的研究

随着互联网技术的发展,现有名字空间在语义、功能等多方面的设计上已经不能满足新型应用的需求[41].为了解决现有互联网名字空间在结构和解析机制方面的问题,设计更适合下一代互联网的名字空间,国内外研究者针对名字的结构提出了许多设计方案.

对于名字的结构,HIP[11],SFS[42]最先提出采用扁平的、具有自验证功能的名字对实体命名.随后,MIT 在 2004 年提出的基于无管理协议的互联网体系结构[18],加州大学伯克利分校的研究者在 2004 年提出的互联网层次命名体系结构[43]以及 2007 年提出的面向数据的体系结构[20]中,都采用这种命名规则.采用这种命名机制的名字空间,虽然增强了网络的安全性,但是由于扁平名字空间只保证名字的唯一性,没有内部结构,无法进行聚合,导致可扩展性不好.当前互联网使用的 IP 地址采用层次式的名字空间,体现了一定的拓扑结构.层次名字空间包含了一定的结构特性,有利于聚合,可扩展性较好,但是不能很好地支持移动性.组合名字空间介于二者之间,结合了扁平和层次两种名字空间的优点.例如,1999 年 MIT 计算机科学实验室提出 INS[44](intentional naming system),旨在解决移动环境下的资源发现和服务定位问题.INS 使用组合命名机制,基于属性和数据的组合方法对服务和资源命名,使得资源描述名具有很强的表达能力.

Bengt 等人指出下一代互联网使用扁平名字空间的好处要大于使用层次空间[45].因此,在设计网络实体名

© 中国科学院软件研究所 http://www.c-s-a.org.cn

字空间时,要综合考虑可扩展性(基于层次式的名字有利于聚合)和名字的长期一致性(扁平名字不受拓扑结构的影响)等因素.

3.2 ID/Locator解析机制的研究

对于 ID/Locator 解析机制,现有的研究主要通过边缘网络和 transit 网络之间的间接机制(indirection)[46]完成.具体可以分为数据平面的研究和控制平面的研究.其中,数据平面的研究主要针对报文的处理方式和处理级别.报文的处理方式主要分为地址重写(addressing rewriting)和“映射&封装”(map-and-encapsulate)[36].报文的处理级别分为基于主机的处理、基于网络的处理以及基于主机和网络相结合的处理.不管数据平面采用哪种报文处理方式,在哪个级别完成报文处理,都需要依赖其他附加的机制才能完成 ID/Locator 的转换,这就是控制平面研究的内容.这里主要介绍针对 ID/Locator 映射信息的获取和存储方式,主要分为 push 模型、pull 模型以及 push 和 pull 混合模型.

3.2.1 数据平面的研究

(1) 报文的处理方式

- 映射&封装

映射&封装思想萌芽于 Hinden 提出的 ENCAPS 协议[47].该机制将 ISP 网络和用户网络分离,分别使用不同的地址空间,对报文的处理主要分为映射和封装两个阶段:1) 映射阶段.当通信发起方要与本区域网络之外的节点通信,目的节点标识为 EID(endpoint identifier),报文首先到达本区域网络的边缘路由器.边缘路由器通过查询映射系统将目的 EID 映射为 RLOC(routing locator).2) 封装阶段.当边缘路由器获得了映射信息后,在原始报文头外封装一个包含目的节点 RLOC 信息的报文头,用于在核心网络中路由.这样,经过映射&封装后的报文,源和目的地址是 RLOCs,而原始报文头中的源和目的地址是 EIDs.当该报文到达目的节点所在的边缘网络后,边缘路由器去掉发送方边界路由器封装的报文头,根据原始报文头中目的节点标识 EID,将其发送到目的节点.典型的机制是 eFIT[48],LISP[22]等.

映射&封装机制不需要改变主机和核心路由基础设施,可支持 IPv4/IPv6,同时保持源地址不被改变,有利于过滤机制的实施.但是报文封装和隧道技术引入了新的开销.此外,如果经常通过隧道传送过大的数据包,会存在严重的报文丢失等可靠性问题.

- 地址重写

地址重写思想最先由 Clark 提出,后来被 O’Dell 润色,主要利用 IPv6 地址 128 位的优势,用前 64 位作为路由地址 RG(routing goop),后 64 位作为节点标识 EID.当报文到达本地出口路由器时,报文头中的源 RG 地址被填写;当报文到达目的节点所在网络的入口路由器时,目的 RG 地址被重写,这样保证用户无法感知网络拓扑或者前缀信息.典型的机制是 Six/One[21],Six/One Router[40]等.

Six/One 提出基于 PA 地址空间的解决方案,通过为边缘网络和 Transit 网络维护一个 IPv6 地址空间来避免间接机制带来的问题.IPv6 地址被分为 64 位的子网前缀和 64 位的接口标识,其中前 64 位的子网前缀可分为用于标识提供商的路由前缀和子网 ID.Six/One 借鉴 Shim6[19]机制,多宿主的边缘网络使用每个提供商分配的地址空间,主机可以不必中断会话进行地址切换.主机地址的差异仅表现在地址的高位部分,边缘网络或者提供商可以在报文传输的过程中重写报文头的高位部分.用户发送报文时选择的目的地址,可以作为边缘网络的参考.边缘网络也可以通过重写地址的高位部分为报文选择特定的提供商,在这种情况下,主机很快得到报文头被边缘网络重写的信息,在后续报文直接按照边缘网络选择的地址进行地址重写.此外,Six/One 在报文中附加一部分信息,使得接收方能够根据此信息实现反向的地址重写.

地址重写的机制实现相对简单,但是只适用于 IPv6 网络和应用.此外,存在重写开销,如果地址经常变化,网络中的中间盒需要进行升级,否则无法进行过滤.

(2) 报文的处理级别

- 基于主机的处理机制

该机制采用主机可见的新名字空间,在主机协议栈增加了新的标识层,由主机完成对 ID/Locator 的解析转

© 中国科学院软件研究所 http://www.c-s-a.org.cn

换.典型机制有 HIP[11],SHIM6[19]等.

Shim6(Level 3 Shim for IPv6)提出了一种基于主机的多宿主机制.该机制修改了主机的协议栈,将 IP 层分为 IP 端点子层、Shim6 子层以及 IP 路由子层.Shim6 没有引入新的地址空间,节点标识 ULIDs(upper layer identifiers)和位置都使用 IPv6 地址,Shim6 子层提供了节点标识到位置的绑定.多宿主的主机可以从提供商分配的地址集合中选择地址,同时实现对于上层透明地切换.Shim6 使用 PA 地址,减小了核心路由表规模和路由表的更新速度.此外,主机可以动态选择地址,并实现了对多播的支持.但是该机制过于复杂,部署开销大,而且只支持 IPv6.

在主机协议栈完成映射和报文处理,优点在于核心路由的扩展性较好,支持多宿主和灵活的流量工程,但是主要存在以下 4 个方面的问题:1) 需要更新主机协议栈和上层应用,存在部署难问题;2) 由于主机节点可以获取大量标识与位置的对应关系,网络存在严重的安全性问题;3) 缺乏对站点多宿主与流量工程的支持;4) 由于标识映射数据规模可以达到 $O(10^{14})$[24],DNS 中的标识映射查询效率低,大规模的标识映射表难以维护.

- 基于网络的处理机制

基于网络的处理机制在网络边缘设备上完成映射和报文处理,通过查询映射服务系统完成对 ID/Locator 的解析转换,整个过程对主机透明.典型机制有 LISP[22]等.

思科公司提出的位置与标识分离协议 LISP(locator identity separation protocol)使用 Routing Locators(RLOCs)和 Endpoint Identifiers(EIDs)分别表示设备对于网络的隶属关系以及设备自身的身份信息.LISP 不需要改变主机和核心路由设施,采用映射&封装和隧道的方式由边缘网络路由器 ITR/ETR 来完成报文头中目的 EID 与目的 RLOC 的转换.LISP 封装的报文头中预留了 4 个字节的 Locator Reach Bits[49],对应于 EID-to-RLOC 映射中可达 RLOC 的列表.此外,在 EID-to-RLOC 映射数据属性中增加版本号,有利于快速检测过期无效数据[50].由于这个协议可以通过对边缘路由器进行软件升级的方式部署,并且能够与 IPv4 协议和 IPv6 协议一起使用,无须改变现有名字空间,具有很好的向后兼容性和可部署性.

该机制不需要改变现有主机的协议栈,不采用新的名字空间,提高了地址空间的利用率,避免了大规模主机更新的困境,在安全性、移动性以及可部署性上都具有一定的优越性.但基于网络的机制在标识映射、站点多宿主机制等方面仍然面临巨大的挑战和缺乏深入的研究,仍然没有很好地解决由于标识数据规模巨大所带来的标识映射问题.

- 基于主机和网络相结合的处理机制

基于主机和网络相结合的处理机制,主机和网络共同完成对 ID/Locator 的解析.典型机制有 Six/One[21].该机制中,对于用户发送报文时在报文头填写的目的地址,边缘网络可以参考该值,也可以根据提供商的选择重写地址.当后一种情况发生时,主机根据边缘网络选择的地址对后续报文的地址进行重写.该机制具备基于主机和基于网络的处理机制的优点,在安全性、移动性以及站点多宿主与流量工程方面都有一定的优势.但是由于需要改变主机和网络设备,部署开销较大.此外,对于大规模标识映射表的问题仍没有很好地解决.

3.2.2 控制平面的研究

在构建映射服务时,需要考虑 3 个主要因素:映射数据库的更新速度、映射数据的状态规模、映射查询延时.在设计映射系统时,必须减小映射数据的状态规模(state)与映射数据库的更新速度(rate),同时优化查询延时(latency).

(1) push 模型

push 模型是指边缘路由器获得标识映射表的过程不依赖于数据的驱动.互联网中的边缘路由器通过周期性的交互标识映射表拥有全网的标识映射表,标识与位置的转换过程直接在边缘路由器完成,不存在映射查询延迟问题.但是,由于标识映射数据规模的巨大化(远大于核心路由表),通常,push 模型存在严重的状态规模问题,标识映射表难以收敛.典型机制有 LISP-NERD[51]等.

LISP-NERD(not-so-novel EID to RLOC database)提供了一种用于存放映射信息的网络数据库.与 APT 不同的是,每个 ITR 都需要存储整个映射数据库,报文的映射转发在本地进行,避免了查询机制所带来的延迟和报文

© 中国科学院软件研究所 http://www.c-s-a.org.cn

丢失.NERD 使用 HTTP 协议维护和发布数据库中的数据.但是 ITR 的周期性查询带来了额外的开销;当映射数据的状态规模很大时,每个 ITR 都维护一个完整的映射数据库也是不现实的.

(2) pull 模型

pull 模型,即边缘路由器获得标识映射表的过程受数据驱动,边缘路由器收到报文后通过映射服务查询标识映射表获得标识映射,再进行标识与位置的转换.采用 pull 模型,标识映射表分布放置,可以减少单个网络节点存储的标识映射表的规模,但是增加了映射查询/响应的延时,容易导致边缘路由器由于缓冲区溢出引起的报文丢弃.因此,pull 模型下的标识映射的性能取决于边缘路由器的 Cache 的大小、标识映射服务的查询时间、边缘路由器缓冲区的大小.增大边缘路由器中 cache 的大小,可以增大边缘路由器 cache 中存储的标识映射表的规模,从而提高本地查询命中率.增大边缘路由器缓冲区大小可以有效地降低报文丢弃概率.典型机制有 DHT-MAP[24],LISP-EMACS[52],LISP-DHT[53]等.

LISP-DHT(distributed hash tables)是一种基于 DHT 的分布式映射系统.该机制充分利用 DHT 具有自组织、自维护、可扩展性好等优点,在维护映射局部性特征的同时建立了一个高效、安全的映射查询系统.映射数据库分布式地存储在 LISP-DHT 节点上.LISP-DHT 采用了 Chord[54]算法并且借鉴 Stealth DHT[55]的思想,使用 EID 作为 ChordID,将节点构成逻辑环.该机制基于 DHT 的方式完成映射服务,继承了 DHT 的所有优势,但是映射查询过程,对应实际网络中的跳数无法确定.

DHT-MAP 是 Luo 等人提出的一种基于 DHT 的位置与标识的映射机制,旨在解决现有的映射服务机制不支持扁平标识的问题.DHT-MAP 采用 CAN[56]算法,引入解析器(resolver)、EID-to-Resolver、转发器(forwarder)的概念.其中,解析器负责解析收到的请求,为每个 EID 在 CAN 上注册.每个 AS 都有一个或者多个解析器,存储 AS 内节点 locator 与 EID 的映射.一个 EID-to-Resolver 的映射数据指出了存放 EID-to-locator 映射数据的解析器.转发器是 CAN 中的节点,在它的主映射表和后备映射表中存储了 EID-to-Resolver 的映射关系.此外,DHT-MAP 为了增强网络的鲁棒性并且更好地支持移动性,对 CAN 进行了修改:1) 为每个 CAN 节点设置一个后备节点,实现(key,value)键值对的备份存储.2) 修改节点加入和离开规则:当节点加入时,总是与最大坐标区域平分区域;当节点想离开 CAN,发送一个 HANDOVER 消息到后备节点完成区域的移交.3) 当目的坐标点失效时,将消息转发给后备节点,而不是失效节点.

(3) push 和 pull 混合模型

无论是 push 模型还是 pull 模型均不能很好地满足大规模互联网下通信的要求.push 模型的主要缺点是规模太大,标识映射表难以收敛,扩展性差;pull 模型的主要缺点是查询延迟长,易导致报文丢弃.因此,push 和 pull 混合的模型可能更有效.目前,大多数的研究也是基于 push 和 pull 相结合的模型,在映射查询延迟和映射状态规模之间寻求平衡.典型机制有 APT[37],Ivip[38],LISP-ALT[57],LISP-CONS[58]等.

Jen 等人提出一个新的路由体系结构 APT(a practical tunneling architecture).APT 将网络分为 Transit 空间和 Delivery 空间,类似于核心网络空间和边缘网络空间的划分.Delivery 空间的报文通过 UDP 隧道到达 Transit 空间.APT 通过在 Transit 网络中部署默认的映射器(default mapper)存储整个映射数据库,使用隧道路由器 TR(tunnel router)代替边缘路由器.TR 除了完成对报文进行封装和解封,还缓存从默认映射器获得的映射信息,用于以后的位置和标识转换.此外,APT 对数据库格式没有特殊要求,使用 BGP 协议维护和发布数据库中的数据.

LISP-ALT(alternative logical topology)利用层叠网来完成映射服务.该机制在 ETR(出口隧道路由器)中分布式存储映射数据库,层叠网只负责转发查询和响应消息.在层叠网中,ALT 路由器之间通过 GRE[59](generic routing encapsulation)隧道转发报文.ALT 路由器利用 BGP 协议承载 EID 前缀路由信息,运行单独的 BGP 实例完成 EID 前缀路由信息的通告、聚合以及学习过程.ITR 将本地 cache 未命中的 EID 请求发送到 ALT 层叠网中,属于 pull 模型的行为.此外,ITR 也可以采用数据探测的方法获取映射信息,也属于 pull 模型的行为.收到查询请求报文的 ETR 发送映射响应报文,其中包含 EID 前缀对应的 RLOCs,属于 push 模型的行为.ITR 收到响应消息后,将 EID 前缀和 RLOC 的映射关系缓存在本地 cache 中.该机制扩展性好,部署开销不大.但由于承载于 BGP

© 中国科学院软件研究所 http://www.c-s-a.org.cn

协议,需要运行单独的 BGP 实例,面临映射信息的收敛问题.

3.3 ID/Locator路由机制

完成 ID/Locator 解析后,将标识映射为与拓扑相关的位置信息,接下来就是路由的问题.在位置与标识分离的网络中,现有的路由机制主要分为 3 种:1) 使用传统的基于拓扑信息的路由机制,是利用网络链路状态信息实现数据传输的路由机制.2) 基于标识的路由机制.根据节点在网络中的地位是否平等,也分为层次路由和扁平路由.3) 基于地理位置信息的路由机制,网络中每个节点都有一个可精确定位和测量自身位置及状态的设备,利用该设备所获取的信息进行路由[60].本文重点介绍前两种路由机制.

3.3.1 基于拓扑信息的路由机制

在位置与标识分离的网络中,一般将路由问题分为入口边界域、核心网络域及出口边界域 3 部分.核心网络域可以重用传统的路由机制,也可以采用新型路由机制,增强路由的可扩展性.边界域可以采用单独的路由机制,重点考虑进出核心网络域的路由问题.

互联网是由多个地位不等的网络连接而成的,这种连接关系影响着网络的鲁棒性、端到端的通信性能以及小网络之间的竞争关系.基于网络层次结构的路由,有利于地址聚合和分配.因此,无论从网络规模还是网络管理的角度分析,层次路由一直都是增强路由可扩展性的重要手段.传统的基于路径向量的域间路由协议 BGP 和基于链路状态算法的开放最短路径优先路由协议 OSPF 都属于层次路由机制.此外,典型的机制还有 HLP[8],Hidra[30],ENCAPS[47]等.

Subramanian 等人提出了一种链路状态和路径向量相结合的协议 HLP(hybrid link-state and path-vector protocol)用来代替 BGP,有更好的扩展性、孤立性(isolation)和收敛性.该协议在 AS 内使用链路状态路由,在层次之间使用路径向量路由.链路状态路由改进了收敛性,减少了层次内的路由振荡;路径向量路由增强全局可扩展性,通过层次结构隐藏层次内部的路由更新信息.实验表明,使用 BGP 的路由信息,HLP 能够显著减小路由更新的扰动(churn)速率.此外,有效地将路由事件影响限定在一个区域内,这个区域的大小比 BGP 的事件限定区域缩小了 100 倍.对于大部分路由,HLP 保证最坏情况下的在线性时间内收敛.

Wang 等人提出了一个分级域间路由架构 Hidra(hierarchical inter-domain routing architecture).Hidra 的核心思想是隔离网络边界与核心:相对稳定的核心网络位于高阶路由层,运行高阶域间路由协议,以维持核心网络的可达性;在变化相对剧烈的边界,引入一个低阶映射层和相应的映射服务,以维持边界网络与核心网络之间的可达性.Hidra 引入一个标识传送自治系统位置的域间路由标识 RID(routing identity),由传送自治系统及其提供商自治系统唯一确定.该机制增强了核心网络路由的稳定性,显著降低了全球路由表的规模,增强了域间路由系统的可扩展性.

3.3.2 基于标识的路由机制

(1) 扁平路由机制

目前还有一部分研究者摒弃对标识与位置映射机制的研究,着眼于研究扁平标识路由.扁平标识路由机制是指在网络中直接使用标识路由,不再使用地址信息进行路由,例如 I3[13],ROFL[14],UIP[18]等.

加州大学伯克利分校的 Caesar 等人提出扁平标识路由 ROFL(routing on flat labels).扁平标识路由完全没有使用位置信息.报文头中不包括位置信息,而是直接基于标识进行路由.该方法除继承位置与标识分离的优点外,还有一些独有的特点:无须建立单独的名字解析系统;报文分发不依赖于数据路径之外的其他信息;标识分配简单,只需保持唯一性,无须如 IP 地址一样既要保证唯一性又要保证与网络拓扑的一致性,但是存在可扩展性问题.

扁平标识路由机制的共同特点是不考虑标识和地址的映射问题,直接使用标识进行路由.该机制最主要的问题是可扩展性问题,由于互联网标识规模的庞大,最长路由跳数不确定,无法预知原标识发出的报文是否能在有限时间内一定到达目的标识.

(2) 层次路由机制

基于标识的路由机制中,除了上述的扁平路由机制以外,还有一类基于标识结构的层次特性的层次路由机

© 中国科学院软件研究所 http://www.c-s-a.org.cn

制,典型机制有 NIIA[12],HRA[6]等.

Bengt 等人提出基于位置与标识分离的思想,提出一种基于节点 ID 的互联网体系结构 NIIA(node identity internetworking architecture),旨在解决异构网络的融合问题,为设计新的路由机制提供基础框架.在 NIIA 中,节点标识 NID(node identity)采用非对称密钥对中的公钥值,NIFT(node identity forwarding tag)是 NID 的长度固定的哈希值,与 HIP 中 HIT 类似.采用两层路由结构,在 LD(locator domain)内部采用基于 IP 地址的路由,而在 LD 之间采用基于 NID 注册信息的默认路由.节点用 NID/NIFT 向从 LD 到核心 LD 路径上所有的 NID 路由器注册,NID 路由器形成基于注册信息的路由表.

华为公司的 Xu 等人提出一种新的路由和命名体系结构 HRA(hierarchical routing architecture),旨在解决路由可扩展性问题.该体系结构基于位置与标识分离的思想,对主机标识空间进行了重新设计,引入层次式的标识,采用层次路由机制.HRA 中的每个主机都有一个 HI(host identifier)和一个 HIT(host identity tag).其中,HIT 是 HI 的 hash 值与 AD ID(administrative domain ID)的结合.这种层次结构的标识空间,能够体现主机的归属区域,同时保证安全性和唯一性,改善了映射系统的查询效率.此外,提出基于 LD 的路由和基于地址前缀的路由相结合的层次路由机制.在 LD 之间进行报文转发时使用基于 LD 的路由机制,在 LD 内部使用基于地址前缀的路由机制.HRA 为下一代多层互联网设计了一个通用的架构,但是为了支持 IPv6 需要改变主机和边缘路由器.

此外,紧凑路由(compact routing)[61,62],是指能够满足地址规模和报文头大小以对数级增长,并保证路由表规模达到近似线性增长,同时路由伸展度(stretch)不随网络规模增长能达到一个常数值的路由机制[5],可分为标识具有拓扑信息和没有拓扑信息(扁平)两种方式.紧凑路由揭示了影响路由可扩展性的主要因素,并提出一个满足这些限制因素的路由算法,旨在解决路由表规模与路由伸展度(stretch)之间的权衡问题,研究如何减小路由表规模,同时尽量减少对路由伸展度的膨胀.路由机制的伸展度是指消息路由的路径长度(代价)与该消息使用最短路径路由的路径长度(代价)之比的最大值.减小路由表规模可以使用紧凑路由表结构的方法,最简单的方法是使用节点度数作为路由表入口的匹配值.其次,很多研究针对节点的分组方法,主要思想是根据关联关系将节点分组,每个组对应路由表一个入口.此外,传统的最短路径路由算法需要路由器获得关于网络拓扑的全部信息以及对于每个可能的目的节点的最佳下一跳信息,这就意味着巨大的信息量,可扩展性差,不能保证全网路由表规模的增长速度小于网络规模的线性增长速度[63].因此,为了减小路由表规模,可以去掉一些拓扑信息以及对于某些目的节点的最佳下一跳信息,即不必提供网络拓扑的完整信息.很多紧凑路由的研究,消除最短路径路由的约束,实现路由表规模近似线性的增长.与如今互联网的路由机制相比,紧凑路由机制提供了很好的扩展性.现有的域间路由实现路由表规模指数级增长,而采用紧凑路由机制可以达到对数级增长.研究表明,只有对数级的增长才能满足未来路由体系结构无限扩展的需求[61].但是由于紧凑路由没有考虑策略,并不能直接应用于域间路由,即使支持策略,是否能够提供很好的可扩展性还有待研究.

3.4 评价比较

表 2 主要对数据平面的相关机制进行分析比较,比较的范围包括报文的处理方式、报文的处理级别、节点的标识类型、映射方式以及第 4 节提出的关键技术指标.其中,MS(mapping service)表示映射服务,RVS (rendezvous server)表示聚合点服务器.表 3 主要对控制平面的相关机制进行分析比较.

© 中国科学院软件研究所 http://www.c-s-a.org.cn

Table 2 Comparison of the schemes on the data plane

表 2 数据平面机制的特点比较

Solution	Packet processing approach	Packet processing level	Namespaces	Resolution of names	Scalability	Multi-Homing and TE	Routing aggregation	Routing security	Mobility
eFIT	Map-and-Encapsulate	Network	EID	MS	Medium	Yes	Medium	Poor	Yes
HIP	Addressing rewriting	Host	HIT	DNS+RVS	Medium	No	Poor	Good	Yes
SHIM6	Addressing rewriting	Host	ULID	DNS	Poor	Yes	Medium	Poor	Yes
GSE	Addressing rewriting	Network	ESD	DNS	Poor	Yes	Medium	Poor	No
Six/One	Addressing rewriting	Host+Network	Lower 64 bits of IPv6 address	DNS	Medium	Yes	Medium	Poor	Yes
LISP	Map-and-Encapsulate	Network	EID	MS	Good	Yes	Good	Medium	Yes

Table 3 Comparison of the schemes on the control plane

表 3 控制平面机制的特点比较

Solution	Distribution model	Transmit information	Update sensitive	Mapping approach	Scalability
APT	Push+Pull	The whole mapping database	No	Access default mapper	Poor
LISP-NERD	Push	The whole mapping database	No	Access network database	Poor
LISP-ALT	Push+Pull	EID prefix	Yes	Overlay	Good
LISP-EMACS	Pull	–	No	Multicast	Medium
LISP-CONS	Push+Pull	EID prefix	Yes	Hierarchical mapping	Medium
LISP-DHT	Pull	–	No	Chord	Good
DHT-MAP	Pull	–	No	CAN	Good

4 关键问题

4.1 标识映射服务的可扩展性问题

位置/标识分离体系的核心是标识映射的可扩展性.如何克服影响标识映射可扩展性的 3 个主要因素:映射数据的状态规模、更新频率以及映射查询延迟,是研究的重点和难点.

4.1.1 标识映射数据的规模问题

标识映射数据的大规模是目前标识映射面临的主要困难之一.采用集中式或泛洪式全分布的映射服务显然不能满足要求.因此,如何降低标识映射数据的规模,减少标识映射表的大小是位置与标识分离网络中必须解决的问题.研究思路包括:1) 充分利用标识具有逻辑扁平、物理可聚合的特点,最大程度地将可以聚合的标识聚合,减少总体标识映射表的规模;2) 结合标识的层次特性,标识映射采用层次化结构和 DHT 相结合的方式.

4.1.2 标识映射数据的更新频率问题

导致标识映射数据更新的主要原因是网络波动.网络波动的速度直接影响到标识映射数据的更新机制,仅根据拓扑变化触发标识映射数据的更新也存在收敛和振荡问题.研究思路是采用网络状态主动测量和信息隐藏机制的方式,控制映射数据的更新频率,以最小的代价维护标识映射数据的全局一致性.具体实施可以是:当网络波动速度较快时,可以采用特殊的更新消息通告方式,以减少更新消息的通信量,加快映射状态的收敛;当网络波动速度较慢或者处于零变化时,可以充分借鉴信息隐藏机制,利用多宿主的优势对报文进行重定向,对于可以不通告的更新信息,采用不通告的信息隐藏方式,减少更新的消息开销.

4.1.3 标识映射服务的查询延迟问题

标识映射的查询延迟问题是指在大规模标识空间中,如何减少标识映射的查询时间,提高映射服务性能.现有标识映射服务模型,无论是基于 DNS 还是基于 overlay 网络,都存在查询/响应延迟较长的问题.如何解决标识

© 中国科学院软件研究所 http://www.c-s-a.org.cn

映射的查询延迟问题是建立标识路由模型的重点与难点.研究应根据网络应用呈现的小世界、无标度等特点,依据复杂网络理论,在标识映射延迟和标识映射状态规模之间寻求平衡.研究思路包括:1) 采用基于 overlay 网络的标识映射服务模型,研究构建 overlay 时的服务节点选择问题;2) 采用类似 DHT 的分布式存储机制,研究逻辑结构节点的逻辑位置与物理位置的一致性问题,减小查询请求的路径长度;3) 利用 cache 机制,研究基于机器学习的标识映射的主动推送机制.

4.2 节点快速移动时的通信保障问题

节点在不同区域网络间快速移动时,如何保证通信不间断是研究的主要问题.传统的针对固定 IP 的各种服务连接协议无法保证移动节点的通信质量.对于节点的移动管理问题,目前较为成熟的解决方案有基于应用层的 SIP[64]协议和基于网络层的 MIP[65,66]协议.但因为 SIP 是应用层协议,无法实现对终端移动的及时检测,所以当终端连续移动时,SIP 很难保证可靠的服务,可能导致原有会话的丢失.而 MIP 存在三角路由以及双向隧道路由问题.在位置与标识分离的网络中,当节点移动时,位置与标识的映射关系需要及时更新.当节点快速移动时,标识映射数据库应当及时准确地更新,尽量减少报文延迟和报文丢失的发生,避免映射服务数据不一致,保证主机在不同区域网络间移动时的通信质量.研究思路可以是将 MIP 的思想应用到位置与标识分离的网络中,通过部署服务保障节点,协助服务节点完成标识映射数据的快速及时更新.

4.3 标识路由的聚合问题

位置/标识分离体系的优点是位置采用层次式空间,以方便网络地址前缀的聚合和控制核心路由表的规模.然而,在当前的 Internet 域间路由协议 BGP 中,没有考虑网络的类型(例如 Stub 网络、Transit 网络等),不加区分地将所有的子网前缀载入路由表,每一次 Stub 网络的变化都会激发核心路由表的更新,可能引起整个路由系统的不稳定.另一方面,目前的 BGP 协议由于是一种距离-矢量协议,为避免可能的路由冲突和震荡,难以实现路由的自动聚合,路由聚合需要管理员依据网络拓扑信息进行配置.路由自动聚合问题是实现核心路由表规模可扩展目标的关键问题,即如何在位置/标识分离机制下,从区分网络类型的角度,研究适合路由自动聚合的域间路由网络拓扑表示与模型以及具有路由自动聚合算法.

4.4 标识路由的安全性问题

安全性问题一直都是标识路由研究的难点问题.研究应着眼标识路由系统在遭受攻击、故障或意外事故时,能否及时完成其关键任务,即标识路由是否具有抗毁性.标识路由的抗毁性,主要包含两个方面:1) 标识映射服务的抗毁性.当映射服务系统遭受攻击或者出错的情况下,怎样保证映射信息的真实性和可用性.2) 核心网络路由的抗毁性.当出现路径失效等故障时,如何保证路由的快速重构,迅速从失效状态恢复.

此外,标识路由同样在路由发布、传播、接收等环节缺乏有效的信任机制.在实际情况中,为了窃取网络流量,降低竞争对手的网络服务质量,恶意自治系统可能会通过发布错误路由和劫持网络前缀来欺骗其他自治系统按照错误的路径进行网络流量转发.因此,从路由安全的角度来看,自治系统之间应该是一种有限信任,应该研究基于 IPSec 的消息传输机制,采用 TCP MD5 等签名机制对映射更新会话以及路由会话进行保护,使用公钥证书和数字签名保证会话内容的可信性.

5 总结和展望

在当前的互联网体系结构中,IP 地址存在语义过载的问题,即 IP 地址在语义上具有双重含义,既代表网络节点的拓扑位置,又是节点的标识.它影响了路由的可扩展性,不利于节点或网络的移动,带来了一系列安全问题,限制了多宿主和流量工程的发展.因此,为了解决 IP 地址语义过载问题,国际组织 IAB 提出了位置与标识分离的思想,使用 Identifier 和 Locator 两套名字空间来分别表示节点的标识和位置,即所谓的 Locator/Identifier Split.但是由于 Identifier 不支持全局路由,为了将报文从一个自治域转发到远程自治域,需要将 Identifier 映射为其对应的 Locator,即标识映射技术.对于完成标识映射后的路由问题,可以分为出口/入口域路由和核心域路由,分别采用各自的路由机制实施.

© 中国科学院软件研究所 http://www.c-s-a.org.cn

位置与标识分离后,可以很大程度地减小核心路由表规模,增强了核心路由的可扩展性,更好地支持多宿主和流量工程的发展,增强了网络的安全性和移动性.但是由于引入了位置与标识映射系统,带来了一些新的问题亟待解决:1) 位置与标识映射系统的可扩展性问题,如何减小映射数据的状态规模和映射数据库的更新速度,同时优化查询延时.2) 节点快速移动带来的标识映射数据的快速及时更新问题.3) 位置与标识映射数据的可信性问题.此外,路由优化问题也值得进一步研究,充分利用核心路由地址空间的聚合性,研究路由聚合问题;利用基于节点标识的身份信息,研究路由的信任和抗毁性问题,增强标识路由的安全性.

下一步的研究应把握如下几个方面:1) 标识结构的设计.高效的标识应采用具有自认证功能的组合结构标识,具备扁平特性,但可以选择性地加入部分体现层次特性的元素.2) 路由结构的设计.可以将 Locator/Identifier Split 与层次路由相结合,同时将映射服务系统层次化,将更新消息局部化.3) 映射服务系统的设计.映射服务系统采用基于 DHT 的逻辑结构组织,充分利用 DHT 特点,优化映射服务性能.在设计映射服务系统的过程中,不应该关注能否设计实现一个支持映射信息快速更新的映射系统,也不是在意在映射关系发生改变后,更新消息是否应该发送给每个拥有该映射关系的实体,而应该关注在一个可以接受的代价的前提下,能否达到预期的目标.

致谢 在此,我们向对本文的工作给予支持和建议的同行,尤其是国防科技大学计算机学院网络与信息安全研究所彭伟、胡宁、涂睿、王海龙表示感谢.

References:

[1] Growth of the bgp table-1994 to present. 2009. http://bgp.potaroo.net/

[2] Feldmann A, Cittadini L, Muhlbauer W, Bush R, Maennel O. HAIR: Hierarchical architecture for Internet routing. In: Proc. of the ACM CoNext 2009 Workshop on Rearchitecting the Internet (ReArch 2009) Conf. Rome: ACM Press, 2009. 43−48. http://www.net.t-labs.tu-berlin.de/papers/FBCMM-HAIR-08.pdf

[3] Fuller V. Scaling issues with routing + multi-homing. In: Proc. of the Plenary session at Asia Pacific Regional Internet Conf. on Operational Technologies (APRICOT 2007). 2007. http://www.vaf.net/~vaf/apricot-plenary.pdf

[4] Huston G. BGP routing table analysis reports. http://bgp.potaroo.net/

[5] Guo YF, Zhu XY, Wang N. Considerations on novel network routing mechanisms. ZTE Communications, 2008,14(1):21−26 (in Chinese with English abstract).

[6] Xu XH, Guo DY. Hierarchical routing architecture (HRA). In: Proc. of the Next Generation Internet Design and Engineering (NGI) Conf. Krakow. IEEE Computer Society Press, 2008. 92−99. http://www.ops.ietf.org/lists/rrg/2008/pdfQGp9mwlnHk.pdf

[7] Meyer D, Zhang L, Fall K. Report from IAB workshop on routing and addressing. RFC 4984, 2007.

[8] Subramanian L, Caesar M, Ee CT, Handley M, Mao M, Shenker S, Stoica I. HLP: A next-generation inter-domain routing protocol. ACM SIGCOMM Computer Communication Review, 2005,35(4):13−24.

[9] Quoitin B, Iannone L, Launois CD, Bonaventure O. Evaluating the benefits of the locator/identifier separation. In: Proc. of the 2nd ACM/IEEE Int'l Workshop on Mobility in the Evolving Internet Architecture (MobiArch) Conf. Kyoto: ACM Press, 2007. 1−6. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.125.8411&rep=rep1&type=pdf

[10] Clark D, Braden R, Falk A, Pingali V. FARA: Reorganizing the addressing architecture. ACM SIGCOMM Computer Communication Review, 2003,33(4):313−321.

[11] Moskowitz R, Nikander P. Host identity protocol (HIP) Architecture. RFC 4423, 2006.

[12] Ahlgren B, Arkko J, Eggert L, Rajahalme J. A node identity internetworking architecture. In: Proc. of the IEEE INFOCOM Global Internet Symp. Barcelona: IEEE Computer Society Press, 2006. 1−6. http://net.cs.uni-tuebingen.de/gi2006/papers%20in%20program/Eggert_Lars_A%20Node%20Identity%20Internetworking%20Architecture_final-node-id-arch.pdf

[13] Stoica I, Adkins D, Zhuang S, Shenker S, Surana S. Internet indirection infrastructure. IEEE/ACM Trans. on Networking (TON), 2004,12(2):205−218.

[14] Caesar M, Condie T, Kannan J, Lakshminarayanan K, Stoica I, Shenker S. ROFL: Routing on flat labels. In: Proc. of the ACM SIGCOMM Conf. Pisa: ACM Press, 2006. 363−374. http://www.cs.uiuc.edu/homes/caesar/papers/rofl.pdf

© 中国科学院软件研究所 http://www.c-s-a.org.cn

[15] Francis P, Gummadi R. IPNL: A NAT-extended Internet architecture. In: Proc. of the ACM SIGCOMM Conf. New York: ACM Press, 2001. 69−80. http://nms.csail.mit.edu/~ramki/ipnl.pdf

[16] Gritter M, Cheriton DR. An architecture for content routing support in the Internet. In: Proc. of the 3rd Conf. on USENIX Symposium on Internet Technologies and Systems (USITS 2001) Conf. San Francisco: ACM Press, 2001. 4−15. http://www-dsg.stanford.edu/triad/usits.ps.gz

[17] Guha S, Francis P. An end-middle-end approach to connection establishment. In: Proc. of the ACM SIGCOMM Conf. Kyoto: ACM Press, 2007. 193−204. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.119.2058&rep=rep1&type=pdf

[18] Ford B. Unmanaged internet protocol: Taming the edge network management crisis. In: Proc. of the ACM Hotnets Workshop 2003. Cambridge: ACM Press, 2003. 93−98. http://pdos.csail.mit.edu/papers/uip:hotnets03.pdf

[19] Nordmark E, Bagnulo M. Shim6: Level 3 multihoming shim protocol for IPv6. RFC 5533, 2009.

[20] Koponen T, Chawla M, Chun BG, Ermolinskiy A, Kim KH, Shenker S, Stoica I. A data-oriented (and beyond) network architecture. In: Proc. of the ACM SIGCOMM Conf. Kyoto: ACM Press, 2007. 181−192. http://berkeley.intel-research.net/bgchun/dona-sigcomm07. pdf

[21] Vogt C. Six/One: A solution for routing and addressing in IPv6. Internet draft: draft-vogt-rrg-six-one-02, 2007.

[22] Meyer D. The locator identity separation protocol (LISP). The Internet Protocol Journal, 2008,11(1):23−36.

[23] Zhang H, Chen MK, Zhu YC. Evaluating the performance on ID/Loc mapping. In: Proc. of the IEEE GLOBECOM 2008 Conf. New Orleans: IEEE Computer Society Press, 2008. 1−5. http://zhuyc.info/globecom08idloc.pdf

[24] Luo HB, Qin YJ, Zhang HK. A DHT-based identifier-to-locator mapping approach for a scalable Internet. IEEE Trans. on Parallel and Distributed System, 2009,20(12):1790−1802. [doi: 10.1109/TPDS.2009.30]

[25] Luo HB, Wang HC, Zhang HK, Qin YJ, Qiao CM. 4I: A secure and scalable routing architecture for future Internet. Technical Report, Beijing Jiaotong University, 2009.

[26] Xu XH. Routing architecture for the next generation Internet (RANGI). Internet Draft, draft-xu-rangi-01.txt, 2009.

[27] Tu R, Su JS, Peng W. A survey of research on naming and addressing architecture based on locator/identifier split. Journal of Computer Research and Development, 2009,46(11):1777−1786 (in Chinese with English abstract).

[28] Hou J, Liu YP, Gong ZH. SILMS: A scalable and secure identifier-to-locator mapping service system design for future Internet. In: Proc. of the 2nd Int'l Workshop on Computer Science and Engineering Conf. Qingdao: IEEE Computer Society Press, 2009. 54−58. http://www.computer.org/portal/web/csdl/doi/10.1109/WCSE.2009.765

[29] Hou J, Liu YP, Gong ZH. A framework for identifier-based routing for future Internet. In: Proc. of the Routing-09 Symp. on 2009 8th IEEE Int'l Conf. on Dependable, Autonomic and Secure Computing Conf. Chengdu: IEEE Computer Society Press, 2009. http://doi.ieeecomputersociety.org/10.1109/DASC.2009.12

[30] Wang N, Ma HL, Chen DN, Wang BQ. Hidra: A hierarchical inter-domain routing architecture. Chinese Journal of Computers, 2009,32(3):377−390 (in Chinese with English abstract).

[31] Feldmann A. Internet clean-slate design: What and why? ACM SIGCOMM Computer Communication Review, 2007,37(3):59−64.

[32] Cao R, Wu JP, Xu MW. Research on Internet naming. Journal of Software, 2009,20(2):363−374 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/3389.htm [doi: 10.3724/SP.J.1001.2009.03389]

[33] Rosen E, Viswanathan A, Callon R. Multiprotocol label switching architecture. RFC 3031, 2001.

[34] MPLS. http://baike.c114.net/view.asp?MPLS

[35] Deering S, Hinden R. Internet protocol, version 6 (IPv6) Specification. RFC 2460, 1998.

[36] Jen D, Meisel M, Yan H, Massey D, Wang L, Zhang BC, Zhang LX. Towards a new internet routing architecture: arguments for separating edges from transit core. In: Proc. of the 7th ACM Workshop on Hot Topics in Networks (HotNets-VII). Calgary: ACM Press, 2008. 1−6. http://irl.cs.ucla.edu/~mrm/papers/hotnets08.pdf

[37] Jen D, Meisel M, Massey D, Wang L, Zhang BC, Zhang LX. APT: A practical tunneling architecture for routing scalability. Technical Report, 080004, University of California Los Angeles (UCLA), 2008.

[38] Whittle R. IVIP - A new routing and addressing architecture for the Internet. 2008. http://www.firstpr.com.au/ip/ivip/

[39] Herrin W. Tunneling route reduction protocol (TRRP). http://bill.herrin.us/network/trrp.html

© 中国科学院软件研究所 http://www.c-s-a.org.cn

[40] Vogt C. Six/One Router: A scalable and backwards-compatible solution for provider-independent addressing. In: Proc. of the ACM SIGCOMM MobiArch Workshop. Seattle: ACM Press, 2008. 13−18. http://www.sigcomm.org/sigcomm2008/workshops/mobiarch/papers/p13.pdf

[41] Li D, Wu JP, Cui Y, Xu K. Research on the structures and resolutions of Internet namespaces. Journal of Software, 2005,16(8):1445−1455 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/16/1445.htm [doi: 10/1360/jos161445]

[42] Mazières D, Kaminsky M, Kaashoek MF, Witchel E. Separating key management from file system security. Operating Systems Review, 1999,34(5):124−139. [doi: 10.1145/319344.319160]

[43] Balakrishnan H, Lakshminarayanan K, Ratnasamy S, Shenker S, Stoica I, Walsh M. A layered naming architecture for the Internet. In: Proc. of the ACM SIGCOMM Workshops. ACM Press, 2004. 343−352. http://nms.csail.mit.edu/papers/layerednames-sigcomm04.pdf

[44] Winoto WA, Schwartz E, Balakrishnan H, Lilley J. The design and implementation of an intentional naming system. Operating Systems Review, 1999,34(5):186−201. [doi: 10.1145/319344.319164]

[45] Ahlgren B, D'Ambrosio M, Dannewitz C, Marchisio M, Marsh I, Ohlman B, Pentikousis K, Rembarz R, Strandberg O, Vercellone V. Design considerations for a network of information. In: Proc. of the Re-Architecting the Internet Conf (ReArch 2008). Madrid: ACM Press, 2008. http://conferences.sigcomm.org/co-next/2008/CoNext08_proceedings/ReArch08Papers/1569148930.pdf

[46] Lin C, Lei L. Research on next generation Internet architecture. Chinese Journal of Computers, 2007,30(5):377−390 (in Chinese with English abstract).

[47] Hinden R. New Scheme for Internet routing and addressing (ENCAPS) for IPNG. RFC 1955, 1996.

[48] Massey D, Wang L, Zhang B, Zhang L. A proposal for scalable Internet routing and addressing. Internet Draft, draft-wang-ietf-efit-00.txt, 2007.

[49] Farinacci D, Fuller V, Meyer D, Lewis D. Locator/ID separation protocol (LISP). Internet Draft, Draft-ietf-lisp-05.txt, 2009.

[50] Iannone L, Saucez D, Bonaventure O. LISP mapping versioning. Internet draft: draft-iannone-lisp-mapping-versioning-00.txt, 2009.

[51] Lear E. NERD: A not-so-novel EID to RLOC database. Internet Draft, draft-lear-lisp-nerd-04.txt, 2007.

[52] Brim S, Meyer D, Curran J. EID mappings multicast across cooperating systems for LISP. Internet Draft, draft-curran-lisp-emacs-00.txt, 2007.

[53] Mathy L, iannone L. LISP-DHT: Towards a DHT to map identifiers onto locators. In: Proc. of the Re-Architecting the Internet Conf (ReArch 2008). Madrid: ACM Press, 2008. 7−12. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.139.735&rep=rep1&type=pdf

[54] Stoica I, Morris R, Liben-Nowell D. Chord: A scalable peer-to-peer lookup service for Internet applications. In: Proc. of the ACM SIGCOMM Conf. San Diego: ACM Press, 2001. 149−160. http://pdos.csail.mit.edu/papers/chord:sigcomm01/chord_sigcomm.pdf

[55] Brampton A, MacQuire A, Rai IA, Race NJP, Mathy L. Stealth distributed hash table: A robust and flexible super-peered DHT. In: Proc. of the 2nd Conf. on Future Networking Technologies (CoNEXT 2006). Lisbon: ACM, 2006.

[56] Ratnasamy S, Francis P, Handley M, Karp R, Shenker S. A scalable content-addressable network. In: Proc. of the ACM SIGCOMM Conf. San Diego: ACM Press, 2001. 161−172. http://www.sigcomm.org/sigcomm2001/p13-ratnasamy.pdf

[57] Fuller V, Farinacci D, Meyer D, Lewis D. LISP alternative topology (LISP-ALT). Internet Draft, draft-ietf-lisp-alt-01.txt, 2009.

[58] Brim S, Chiappa N, Farinacci D, Fuller V, Lewis D, Meyer D. LISP-CONS: A content distribution overlay network service for LISP. Internet Draft: draft-meyer-lisp-cons-04.txt, 2008.

[59] Farinacci D, Li T, Hanks S, Meyer D, Traina P. Generic routing encapsulation (GRE). RFC 2784, 2000.

[60] Oliveira R, Lad M, Zhang BC, Zhang LX. Geographically informed inter-domain routing. In: Proc. of the 15th IEEE Int'l Conf. on Network Protocols (ICNP 2007). Beijing: IEEE Computer Society Press, 2007. 103−112. http://www.cs.ucla.edu/~lixia/papers/07SIGPoster_giro.pdf

[61] Krioukov D, Claffy K, Fall K, Brady A. On compact routing for the Internet. ACM SIGCOMM Computer Communication Review (CCR), 2007,37(3):41−52.

© 中国科学院软件研究所 http://www.c-s-a.org.cn

[62] Krioukov D, Fall K, Yang XW. Compact routing on Internet-like graphs. In: Proc. of the IEEE INFOCOM 2004 Conf. Hong Kong: IEEE Computer Society Press, 2004. 219−220. http://www.ieee-infocom.org/2004/Papers/05_4.PDF

[63] Afek Y, Gafni E, Ricklin M. Upper and lower bounds for routing schemes in dynamic networks. In: Proc. of the 30th Symp. on Foundations of Computer Science Conf. IEEE Computer Society Press, 1989. 370−375. http://doi.ieeecomputersociety.org/10.1109/SFCS.1989.63505

[64] Rosenberg J, Schulzrinne H, Camarillo G, Johnston A, Peterson J, Sparks R, Handley M, Schooler E. Session initiation protocol. RFC 3261, 2002.

[65] Perkins C. IP Mobility support for IPv4. RFC 3344, 2002.

[66] Johnson D, Perkins C, Arkko J. Mobility support in IPv6. RFC 3775, 2004.

附中文参考文献:

[5] 郭云飞,朱宣勇,王娜.对新型网络路由机制的思考.中兴通讯技术,2008,14(1):21−26.

[27] 涂睿,苏金树,彭伟.位置与标识分离的命名和寻址体系结构研究综述.计算机研究与发展,2009,46(11):1777−1786.

[30] 王娜,马海龙,程东年,汪斌强.Hidra:一个分级域间路由架构.计算机学报,2009,32(3):377−390.

[32] 曹锐,吴建平,徐明伟.互联网命名问题研究.软件学报,2009,20(2):363−374. http://www.jos.org.cn/1000-9825/3389.htm [doi: 10.3724/SP.J.1001.2009.03389]

[41] 李丹,吴建平,崔勇,徐恪.互联网名字空间结构及其解析服务研究.软件学报,2005,16(8):1445−1455. http://www.jos.org.cn/1000-9825/16/1445.htm [doi: 10/1360/jos161445]

[46] 林闯,雷蕾.下一代互联网体系结构研究.计算机学报,2007,30(5):693−711.

侯婕(1983－),女,河北吴桥人,博士生,主要研究领域为计算机网络体系结构,网络与信息安全.

龚正虎(1945－),男,教授,博士生导师,主要研究领域为路由与交换技术,网络管理,网络测量.

刘亚萍(1973－),女,博士,副研究员,主要研究领域为高性能路由器技术,域间路由技术.

© 中国科学院软件研究所 http://www.c-s-a.org.cn