SAMS:一种新型身份/位置分离方案

杨 芫[1),5)] 徐明伟[1),5)] 杨家海[2),5)] 董庆洲[3)]
陈文龙[4)] 王 会[2),5)] 张志超[1),5)]

[1)](清华大学计算机科学与技术系 北京 100084)
[2)](清华大学网络科学与网络空间研究院 北京 100084)
[3)](北京邮电大学网络与交换国家重点实验室 北京 100876)
[4)](首都师范大学信息工程学院 北京 100037)
[5)](清华信息科学与技术国家实验室(筹) 北京 100084)

摘 要 当前互联网中的 IP 地址同时标识主机身份和主机位置,这种语义重载主要导致了两方面的问题.一方面,它使得核心网络路由表项数量急剧增长,引起路由可扩展性问题.另一方面,主机难以在不改变身份标识的情况下实现多宿主和移动中的高速切换.解决这两个问题的根本办法是主机的身份和位置分离,即分别使用相互独立的身份标识和位置标识.目前身份/位置分离方案得到了研究人员的广泛关注,然而现有的方案只是针对某一个问题,不能同时解决这两个问题.此外身份位置分离之后的真实身份问题也很重要.文中提出了一种新型身份/位置分离方案 SAMS(Scalable Authentic Mobile identifier-locator Separation scheme),它将用户身份标识、主机身份标识、边缘网络位置标识和核心网络位置标识分离,并设计合理的体系结构将这 4 种标识结合在一起,对路由可扩展性,主机多宿主和移动能力都有很大的提高,并支持真实身份.文中实现了 SAMS 的原型系统,并在 CERNET2 主干网上进行了规模部署和实验,验证了方案的有效性和系统的兼容性.

关键词 身份/位置分离;路由可扩展;身份认证;主机移动;多宿主
中图法分类号 TP393 **DOI 号** 10.3724/SP.J.1016.2014.00000

SAMS: A Novel ID-Locator Separation Scheme

YANG Yuan[1),5)] XU Ming-Wei[1),5)] YANG Jia-Hai[2),5)] DONG Qing-Zhou[3)]
CHEN Wen-Long[4)] WANG Hui[2),5)] ZHANG Zhi-Chao[1),5)]

[1)](*Department of Computer Science and Technology, Tsinghua University, Beijing* 100084)
[2)](*Institute for Network Sciences and Cyberspace, Tsinghua University, Beijing* 100084)
[3)](*State Key Laboratory of Networking and Switching, Beijing University of Posts and Telecommunications, Beijing* 100876)
[4)](*Information Engineering College, Capital Normal University, Beijing* 100037)
[5)](*Tsinghua National Laboratory for Information Science and Technology (TNList), Beijing* 100084)

Abstract In the current Internet, IP address has the semantics of both endpoint identifier (ID) and locator. The overloading of IP Address semantics introduces two problems. On one hand, routing entries in the DFZ (default-free-zone) are growing rapidly, which is called the routing scalability problem. On the other hand, mobile or multi-homed endpoints cannot switch smoothly from one interface to another. The prime solution to these two problems is to separate identifier and locator. ID-locator separation is gaining increasing attention recently, but each existing approach can only solve one of the two problems. Furthermore, authentication of identifiers is

收稿日期:2012-01-20;最终修改稿收到日期:2013-10-21.本课题得到国家自然科学基金(61073166,61170211)、国家"八六三"高技术研究发展计划项目基金(2009AA01z251,2011AA01A101)、国家"九七三"重点基础研究发展规划项目基金(2012CB315803)、教育部高等学校博士学科点专项科研基金(20110002110056)资助.**杨 芫**,男,1984 年生,博士研究生,主要研究方向为计算机网络体系结构、互联网路由.E-mail: yyang@csnet1.cs.tsinghua.edu.cn.**徐明伟**,男,1971 年生,博士,教授,主要研究领域为计算机网络体系结构、高速路由器体系结构、互联网路由.**杨家海**,男,1966 年生,博士,研究员,主要研究领域为互联网管理、网络测量、流量工程与大规模网络运行.**董庆洲**,男,1985 年生,硕士研究生,主要研究方向为网络与交换、网络体系结构.**陈文龙**,男,1976 年生,博士,讲师,主要研究方向为网络体系结构、网络协议.**王 会**,女,1977 年生,博士,助理研究员,主要研究方向为网络测量、网络管理、网间路由协议、ISP 经济行为分析.**张志超**,男,1986 年生,硕士研究生,主要研究方向为互联网传输协议.

significant to the Internet. In this paper, we propose a novel ID-locator separation scheme called SAMS (Scalable Authentic Mobile identifier-locator Separation scheme). SAMS separates user ID, host ID, edge locator and routing locator, and combines them with a reasonable architecture. SAMS makes large improvement on routing scalability as well as multi-homing and mobility, and it also supports authentic identifiers. We implemented a prototype of SAMS, and did large-scale deployment on CERNET2 backbone. The experiment results verified the effectiveness and compatibility of SAMS.

Keywords ID-locator separation; routing scalability; authentication; host mobility; multi-homing

1 引 言

在当前的互联网中,IP 地址同时起着标识主机身份和标识主机位置的作用[1].这种 IP 地址语义的重载主要带来了两方面的影响.一方面,它严重影响了互联网路由的可扩展性[2],尤其是当大量用户自治系统(customer's AS)以多宿主(multi-homing)的方式连接到多个提供者自治系统(provider's AS)并向外公告前缀时,相同的地址前缀将出现在多条域间路由中,造成核心网络路由表急剧膨胀[3-4].另一方面,它限制了终端主机的移动能力,通常情况下主机位置的改变伴随着 IP 地址的改变,这必将导致传输连接的中断,虽然提出了移动 IP 等技术[5-6],但其带来的隧道开销问题和三角路由问题仍未得到很好的解决.

解决这两个问题最根本的思路是将 IP 地址的这两种语义分离[1],即设计和使用相互独立的身份标识和位置标识.这一解决思路受到了广泛关注,已经有很多方案被提出[7-8],不妨将它们统称为身份/位置分离方案.然而已有的方案只能解决一部分问题,例如 LISP 方案[9]是重点针对路由可扩展性问题设计的,主机移动能力较差;Shim6[10]方案仅仅解决主机多宿主切换的问题;HIP 方案[11]能提供对主机多宿主和移动中切换的支持,但难以在短期内提升核心网路由可扩展性.此外还有一些其它方案,但都存在较大的不足.为了更加适应新一代互联网的发展和需求,亟需提出一种更好的方案.

本文提出了一种新型身份/位置分离方案 SAMS(Scalable Authentic Mobile identifier-locator Separation scheme).SAMS 将用户身份标识、主机身份标识、边缘网络位置标识和核心网络位置标识相分离,设计合理的网络体系结构来管理不同标识间的映射并进行正确的路由转发,最终将多种标识有机地结合到一起,提供对多方面的同时支持:提高核心网络路由可扩展能力,多宿主和移动中高速透明切换,真实身份认证和会话保证.该方案还具有很好的兼容性和可逐步部署能力,这也是现有方案难以做到的.本文在真实网络设备上实现了 SAMS 的原型系统,并在 CERNET2[12]主干网上进行了规模部署和实验,结果表明该原型系统在上述几个方面都具有较好的性能.

2 问题分析

互联网采用的 TCP/IP 体系结构存在跨层操作的现象,这是由互联网最初的设计目标和一定的历史原因造成的.跨层操作虽然在一定程度上简化了协议栈的实现,但随着互联网的发展却带来了一些问题,其中 IP 地址语义的重载对当今互联网带来了很大的影响.所谓 IP 地址语义重载[1],指的是 IP 地址既用来标识主机或网络所处的位置,又用来标识该主机或网络的身份,即 IP 地址同时拥有位置标识和身份标识的语义.IP 地址的语义重载造成的主要问题包括路由可扩展性、多宿主与移动性两方面.

路由可扩展性问题指的是由于网络中路由表项数量随着互联网规模的扩大而急剧增长,从而带来巨大的存储开销、路由信息通信开销和路由计算开销,大量的路由信息甚至严重影响了互联网路由的收敛性.造成这一问题的主要原因是用户 AS 为了提高网络连通度或实现流量工程,向多个提供者 AS 通告相同的前缀,而其中一些前缀只是某些提供者 AS 提供的前缀的子前缀,这就导致了位置标识空间的聚合性下降.在这种情况下,远程 AS 将计算出多条到达该用户 AS 的路由,这种路由随着网络连通度的提高以乘法的速度增长.为了提高位置标识空间的聚合性,应该设法将边缘网络位置标识与核心网络位置标识空间分离,使边缘网络保持较完整的前缀.这种前缀可以作为该边缘网络的身份标识,再通过设计合理的映射机制将它们与核心网络

的位置标识对应起来，将本来较复杂的边缘网络的路由问题转化成相对简单的核心网络的路由问题.

多宿主和移动性已成为当今互联网端系统发展的新方向.一台多宿主的主机可以同时接入到不同提供者的网络，作为客户端的多宿主主机可以随时根据应用的需求和网络的具体情况选择能提供较好服务质量的网络，而作为服务器的多宿主主机可以直接为它所接入的网络的用户提供访问服务.而端系统移动性则要求网络能够提供随时随地的接入，在主机移动的过程中，正在进行的通信不会中断，服务质量不会剧烈变化.然而端系统多宿主与移动性的发展受到了 IP 语义重载的制约，因为 IP 地址既作为网络层路由寻址的依据，又与端口号组成二元组作为传输层的访问点来标识主机上运行的进程.一方面，如果要保证多宿主主机切换接入网络或移动主机跨域移动时通信不被中断，IP 地址就不能改变，但这样就无法通过路由及时找到该主机正确的位置；另一方面，如果要保证路由的正确性，传输层就会由于 IP 地址的改变而无法将分组提交到正确的进程，最终导致为了继续通信只能重新建立连接.虽然移动 IP 已经为解决这一问题提供了思路，但最根本的方法还是将 IP 地址的语义分离，即单独设计身份标识来取代 IP 地址标识身份的职能.该身份标识被传输层所使用，不包含位置信息，且应该具有全球唯一的特性.

互联网的安全在很大程度上需要依赖认证技术，其中主机和用户的身份认证是保证可信通信的重要手段.IP 地址同时标识着主机的位置和身份，因此难以作为身份认证的依据，给攻击者利用 IP 地址的这一弱点进行攻击提供了机会，例如伪造源 IP 地址进行大规模拒绝服务(DoS)攻击等.虽然可以利用 MAC 地址等信息来进行主机身份认证，但 MAC 地址并不包含用户身份的语义，且也有可能被伪造.设计安全性更强的身份标识，并将其与位置标识即 IP 地址分离有利于提高互联网安全性、可控性和可信性，便于网络安全管理.

3　相关研究

截止到目前，学术界已经提出了很多身份/位置分离方案.其中 LISP 已经在思科路由器上实现并成为 IETF 标准，HIP 和 Shim6 也已经成为 IETF 标准.

LISP 用于实现边缘网络地址空间与核心网络地址空间的分离，从而解决核心网络路由可扩展问题.具体来说，LISP 把 IPv4 地址空间划分为端节点身份标识(Endpoint Identifier，EID)和路由标识(Routing Locator，RLOC).主机发送的分组到达边缘网络入口路由器(Ingress Tunnel Router，ITR)后，ITR 利用一整套映射系统查询得到目的主机所在边缘网络的出口路由器(Egress Tunnel Router，ETR)的 RLOC，接着 ITR 用含有该 RLOC 的头部封装原来的分组，并使用隧道机制将其发往 ETR，ETR 对分组解封装后发往边缘网络中的目的主机.LISP 利用映射系统来保存 EID 与 RLOC 之间的映射，路由器只需要保存 RLOC 的路由，从而减少了路由表项数量.LISP 不对终端主机作任何修改，尽管它显示地提供了对多宿主和流量工程的支持，然而它本身对终端主机提供的安全性以及移动能力的支持较为有限.此外，从本质上说，LISP 是把原有路由表相关的存储和通信开销转移到映射系统中，因此其映射系统的存储和通信开销仍然是一个问题，尤其是由 ITR 发起查询这一点[13]会带来一些问题，对于触发查询的分组，如果 ITR 对其进行缓存，会给高速存储资源有限的路由器带来较大的负担且不可扩展，而如果 ITR 丢弃该分组，则会导致终端计时器超时重传，对发起通信时的时延增加较大.

HIP 用于支持主机身份标识与位置标识的分离.HIP 采用全球唯一的 IPv6 地址作为主机的身份标识，并在分组的网络层头部与传输层头部之间加入一个新的头部存放身份标识.为了处理身份标识，HIP 在传统互联网体系结构的传输层和网络层之间加入了主机标识层(Host Identify Layer，HIL)，HIL 以上使用主机标识(Host Identifier，HI)作为访问点，HIL 以下则采用传统的 IP 地址进行位置寻址.身份标识和位置标识的映射由专门的设备管理，HIL 的功能是查询并缓存这一映射以便正确处理分组头部并完成路由或递交.HIP 支持主机移动和多宿主[14-15]，但它的不足之处在于：不能解决由用户 AS 多宿主导致的路由可扩展性问题；IPv6 形式的身份标识不具有用户身份语义且不便于记忆，能提供的安全保障比较有限[16].此外，HIP 协议栈下能运行的应用程序需要重新开发，它们必须使用一类新的套接字(socket)接口，这使得 HIP 不具有兼容性，很难在现实中部署.

Shim6 协议用于支持主机使用多个位置标识访问互联网，即主机的多宿主.与 HIP 类似，它也将身份标识(Shim6 中称为 Upper Layer Identify，ULID)

设计成 IPv6 地址的形式，并且也增加了新的分组头部和相应的新协议栈层次 Shim 层. 不同之处在于，Shim 层并非独立的一层，而是嵌入到 IPv6 模块中. 此外，Shim6 也没有专门的设备来保存 ULID 与位置标识（Shim6 中称为定位符）的映射，而是双方 Shim 层进行通信协商来获取映射信息. 最重要的协商消息包括初始时的四次握手和定位符变换时的通知消息. 与 HIP 一样，Shim6 也能支持主机移动与多宿主[17-18]，但它同样也没有解决路由可扩展性问题，此外 Shim6 也采用了不具有用户身份语义的 IPv6 地址，且映射关系仅由通信双方自行维护，其安全性不如 HIP.

在身份/位置分离方案的其它研究成果中，一部分研究以现有的标准或方案为基础，因此基本继承了这些方案的优缺点，例如文献[19]和文献[20]使用分布式散列表（DHT）来构建 LISP 的映射系统. 还有的方案设计了新的体系结构，例如文献[21]提出的 MILSA 方案设计了一种类 URI 的层次化主机身份标识，能提供多种场景下的移动能力和安全信任机制，但是没有从根本上解决路由可扩展性问题.

表 1 对现有身份/位置分离方案提供的功能和存在的问题进行了概括总结. 总的来说，目前已经提出的方案都不能同时满足路由的可扩展性、安全性以及多宿主和移动性的需求. 本文提出一种新型身份/位置分离方案 SAMS，并分析它在解决这 3 个问题方面的特性.

表 1 现有方案特性比较

方案	路由可扩展性提高	终端多宿主/移动性提高	安全性提高	可部署性
LISP	较多	较少	无	较好
HIP	短期效果不明显，需要长期大量部署	较多	较多	较差
Shim6	短期效果不明显，需要长期大量部署	较多	无	较好
MILSA	中等	较多	较多	很差

4 SAMS 体系结构

4.1 设计思想

为了同时支持路由可扩展以及主机移动与多宿主，不能简单地照搬现有的标识设计与分离方法. 首先，路由可扩展性的好坏依赖于路由标识能否很好的聚合，尤其是在核心网络中应当尽可能减少路由数量，这可以通过将边缘网络的地址空间与核心网络的地址空间分离来实现，也就是让边缘网络中的前缀在核心网络中仅仅作为该边缘网络的身份标识而存在. 其次，主机移动与多宿主需要传输层能适应网络层位置标识的改变，因此可以在网络层以上传输层以下添加一个新的层次即“身份标识层”，并改变传统传输层以〈IP 地址，端口号〉作为访问点的方式，用〈主机身份标识，端口号〉取而代之，而网络层仍然使用传统的 IP 地址. 这样 IP 地址仅作为主机的位置标识，不起标识主机身份的作用，可以在移动和多宿主情况下进行改变. 此外，为了实现真实身份并适应现实中用户与主机多对多的关系，可以在主机身份标识之外再设计用户身份标识，用户身份标识可以采取便于记忆的形式以便于查询，并应当与用户真实身份绑定从而为身份认证提供基础.

总的来说，SAMS 的核心思想是将用户身份标识、主机身份标识、主机位置标识（边缘网络位置标识）和核心网络位置标识相互分离. 下面将对 SAMS 的体系结构进行详细说明.

4.2 体系结构

用户身份标识（User ID）用于唯一标识一个用户（个人或组织），由用户向权威机构申请获得，并与用户的真实身份信息绑定；主机身份标识（Host ID）用于唯一标识一台主机，它与端口号一起组成主机的传输层的访问点；边缘网络位置标识（Edge Locator，ELOC）指出了主机在本地网络中的位置，是边缘网络路由器决定分组向何处转发的依据；核心网络位置标识（Routing Locator，RLOC）指出了一个边缘网络在核心网络中的位置，是核心网络路由器决定分组向何处转发的依据. 这 4 种标识在协议栈中所处的位置如图 1 所示.

与传统的互联网协议栈相比，SAMS 的协议栈在网络层之上、传输层之下增加了一个新的层次“身份标识层”用于处理身份标识. User ID 可以采用类 Email 地址的格式，由一个标识用户名的字符串、字符“@”和标识管理本区域服务器的主机名（例如 tsinghua.edu.cn）组成. 这样的设计既方便记忆，也有利于借助现有 DNS 系统实现聚合. Host ID 需要携带于分组中进行传输，因此可以通过对 User ID 进行编码得到定长的 Host ID，例如使用 32 比特来标识域（例如可以将一个 AS 作为一个域），24 比特作为用户编号，8 比特作为主机编号. 用户在申请获得 User ID 的同时还将获得一组属于该用户的公私钥对，以便通过数字签名和加密机制来提供对安全性的支持. ELOC 和 RLOC 可以使用现有地址族，

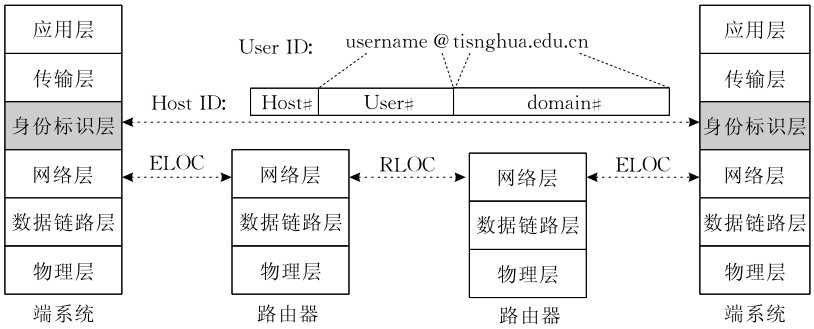


图 1　4 种标识在协议栈中的位置

如 IPv4 或 IPv6. 新型身份/位置分离方案采用如图 2 所示的体系结构来将这 4 种标识统一起来.

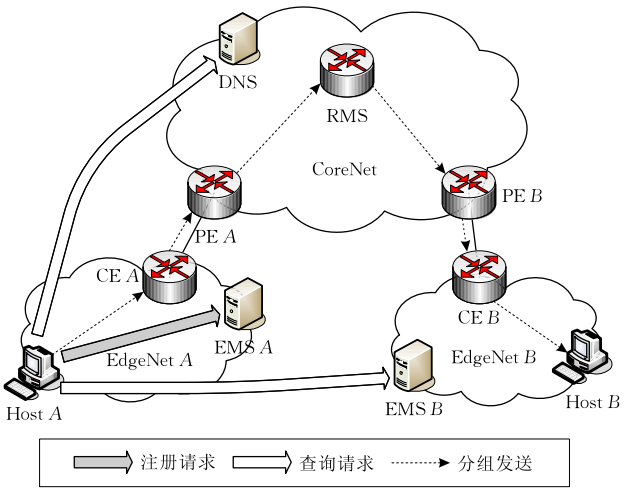


图 2　新型身份/位置分离方案总体框架

该体系结构包括的主要成员和要素有边缘网络映射服务器 EMS(Edge Mapping Server)，扩展的 DNS 系统，核心网络边缘路由器 PE(Provider Edge)以及核心网络映射服务器 RMS(Routing Mapping Server).

EMS 保存了用户真实身份信息和 User ID，它的作用主要有三方面. 首先，EMS 接受主机的注册，认证用户的身份并为其分配 Host ID. 其次，EMS 为已经注册的主机保存 User ID、Host ID 和 ELOC 的实时映射关系，并通过定期与主机交换消息来随时更新 ELOC. 第三，EMS 接受其它主机的查询请求，并用当时的映射关系予以应答. 不同的 EMS 负责不同管理范围内的用户，例如可以在每个 AS 内部部署一台 EMS.

扩展的 DNS 系统保存了 3 类新的记录，第 1 类是域名与 User ID 之间的映射关系，第 2 类是域名或 User ID 与负责该域的 EMS 的 ELOC 之间的映射关系，第 3 类是域名或 User ID 与 Host ID 之间的映射关系. 其中第三类记录主要是针对一些长期在固定位置提供服务，并且 Host ID 很少变化的服务器.

PE 是连接核心网络与边缘网络的路由器，对边缘网络分组进行封装或解封装，并在核心网络中转发. PE 只保存少量常用 ELOC 前缀(popular prefix)到 RLOC 的映射关系，对于其它分组则封装发送到 RMS. PE 还负责通过扩展的 BGP 协议将它所连接的边缘网络的前缀通告给 RMS.

RMS 是核心网络中部署的功能较强的路由器，不仅能从相同 AS 内的 PE 那里学到 ELOC 前缀与 RLOC 之间的映射关系，而且能通过与相邻 AS 建立 BGP 会话学习到全局映射信息. 因此，RMS 拥有全局路由表，能将接收到的封装分组解封装后再次封装，并发送到正确的出口 PE. RMS 还负责将常用 ELOC 前缀通告给相同 AS 内的 PE.

4.3　工作流程

以图 2 中 Host A 发起与 Host B 的通信为例，说明 SAMS 的基本工作流程.

(1) Host A 向其所在管理域的 EMS(即 EMS A)发起注册，经过身份认证后获得 Host ID；

(2) Host A 向扩展的 DNS 系统发起查询，查询关键字为 Host B 的域名，DNS 应答返回 Host B 的 User ID 和 EMS B 的 ELOC；

(3) Host A 用 Host B 的 User ID 作为关键字向 EMS B 发起查询，EMS B 应答返回 Host B 的 Host ID 和 ELOC，(2)和(3)两步中查询的结果可以暂时保存在 Host A 的缓存(Cache)中，在缓存超时之前如果又有查询需求则可以直接从缓存中获取；

(4) Host A 使用(3)中得到的 Host ID 和 ELOC 组装分组，分组被发往该边缘网络的边界路由器(Customer Edge，CE)，并被转发到 PE A；

(5) PE A 用 RMS 的 RLOC 对分组进行封装，封装后的分组在核心网络中被发送到 RMS；

(6) RMS 将分组解封装之后,在映射表中查找目的 ELOC 对应的出口 PE 的 RLOC,用此 RLOC 将分组再次封装,封装后的分组在核心网络中被发送到 PE B;

(7) PE B 将分组解封装,转发到 Host B 所在的边缘网络,分组在该边缘网络中被发送到 Host B;

(8) 最终,Host B 检查出分组的目的 Host ID 与自己的 Host ID 一致,于是接收该分组.

5 SAMS 特性分析

在上一节提出的 SAMS 方案中,路由可扩展性、移动与多宿主以及真实身份认证都能得到很好的支持,本节将分别对这些特性进行分析.

5.1 路由可扩展性

SAMS 的第一个重要特性是能提高核心网络路由的可扩展性.路由可扩展性问题主要是由用户 AS 向多个提供者 AS 通告前缀造成的.然而在 SAMS 中,虽然 PE 会通告它所连接的边缘网络中包含的前缀,但 PE 本身和核心网络中的普通路由器都不会保存这些前缀.换句话说,只有在拥有全局路由表的 RMS 上会产生一些开销,这些开销主要有两部分.

第一是存储路由表和计算路由的开销,这部分开销与现有的核心网络路由器的开销相同.但是,由于一个 AS 内部只有少量 RMS(部署多台 RMS 而不是一台可以预防单点失效问题,并提供可扩展性),因此可以部署高性能的路由器作为 RMS 来存储和计算路由,也就是以较低的成本来解决问题.

第二是路由信息更新的开销,这部分开销与现有的情况相比得到了有所减小.在传统的 BGP 协议中,位于同一个 AS 内部的 BGP 路由器之间以全连通(full-mesh)的形式建立 IBGP 会话,所有的路由更新消息在每一对 BGP 路由器之间传送,造成极大的带宽消耗,虽然目前很多 AS 采用了路由反射器(reflector)技术[22]来取代全连通结构以求减小开销,但 PE 与 reflector 之间的路由更新消息的数据量仍然不可忽视.在 SAMS 中,AS 内部的路由更新消息仅仅涉及与该 AS 所连接的边缘网络的前缀,这只占据了全部路由更新消息的很小一部分;而且两个相邻 AS 之间仅在 RMS 之间建立 EBGP 会话,域间路由更新消息的数量也得到了有效控制.

5.2 移动与多宿主

SAMS 的第二个重要特性是支持大量主机跨 AS 移动或多宿主.在当前的互联网中,IP 地址在标识用户位置的同时还作为传输层的访问点的一部分起着标识主机身份的作用.当主机跨越 AS 移动的时候 IP 地址发生变化,会导致两方面的问题,一是远程用户不能及时得到移动主机最新的位置信息,难以与其建立连接,二是已经建立的传输层连接会被中断,需要重新建立连接才能继续传输,这对于一些实时性要求较高的应用如实时视频会造成较大的影响.

将 User ID、Host ID 与 ELOC 分离可以使边缘网络 IP 地址较好地聚合起来作为位置标识使用,而当主机发生移动/多宿主时也能通过 User ID 和 Host ID 标识出主机的身份.一方面,当主机 ELOC 发生变化时,该主机会立即通知管理它的 EMS 和在之前一段时间内通信过的主机,因此在 ELOC 变化之后其它主机也能通过本地缓存或 EMS 查询到该主机当前的 ELOC,从而发起通信.这种方式能适应终端主机游牧方式的移动,即主机从一个地方移动到另一个地方后停留一段时间,通常移动速度较慢,且移动之后需要重新建立新的连接.另一方面,当主机 ELOC 变化时,该主机会立即通知正与它进行通信的主机,收到通知的对端主机只需要改变目的 ELOC,而不用改变目的 Host ID,所以传输连接不会中断,仅仅会产生一定的带宽和延迟变化.这种方式能适应终端主机漫游方式的移动,即主机可以任意移动且在移动过程中已建立的传输连接不会被中断.

主机多宿主的情况与跨域移动的情况十分类似,不同之处在于多宿主主机可以同时使用多个网络接口卡接入到网络.这样,多宿主主机在从一个 ELOC 切换到另一个 ELOC 之前可以先由两个接口分别同时接入到提供这两个 ELOC 的网络,切换时分别通知 EMS 和对端主机同时更改本地路由.在这个过程中,通信双方的分组仍然能够进行正确收发,所以数据传输受到的影响很小,尤其是当切换前后带宽和端到端延迟比较接近的时候.借鉴多宿主切换的这一优点,可以在主机移动时也使用两个接口,这样一来,即使在获取新的 ELOC 和检测 ELOC 变化的时候也能进行正常通信.

SAMS 不仅支持单个主机的快速移动和多宿主,也支持包含多个主机的网络(例如一辆行使中的列车上的网络)的快速移动和多宿主.这种网络中需要使用特殊的网关,称为移动网络边缘路由器(Mobile Custom Edge, MCE). MCE 一方面充当

NAT 网关，将分组中 ELOC 映射为网关的外部地址和传输层端口号组成的二元组，由于新型身份/位置分离方案中的分组在网络层头部与传输层头部之间增加了一类新的头部，因此 MCE 需要在传统 NAT 网关的基础上进行扩展以便能识别此分组结构；另一方面，MCE 还要负责实现网络整体移动时的相关功能. MCE 可以使用两个外部接口来帮助切换，这与主机多宿主的情况类似. 如图 3 所示，当边缘网络需要整体移动时，MCE 的一个接口负责在原来的路径上继续收发分组，另一个接口则执行建立新的链接的过程. 当新的链接建立完成后，由 MCE 而不是主机负责发送通知给当前有通信的所有对端主机，由于 MCE 作为 NAT 网关保存了动态端口映射表，将其做一定扩展就能实现发送消息给通信对端的功能. 当对端主机收到消息时将发送分组到新的目的地址，此时 MCE 将更新端口映射表，这样就将传输流量转移到了新的路径上.

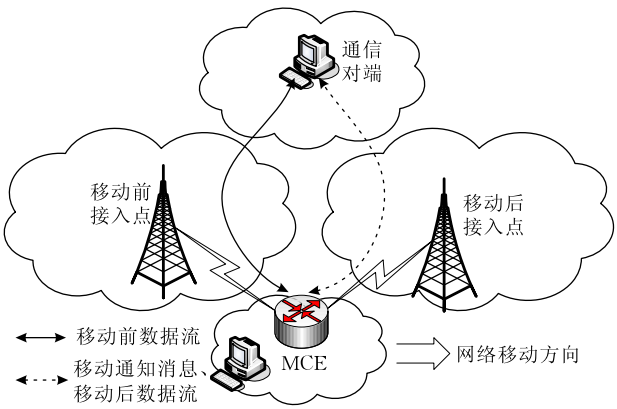


图 3　边缘网络整体移动

5.3　真实身份认证

在现实中，用户与主机之间存在多对多的映射关系. 一方面，同一用户可以使用不同的主机连接到互联网，另一方面，有些主机也会被多个用户交替使用甚至同时使用. 现有的身份/位置分离方案仅定义了一种身份标识即主机身份标识，难以提供对用户身份真实性认证的支持. SAMS 将用户身份标识与主机身份标识相分离的思想能够提供更好的身份认证和安全机制，这与已有的研究有本质上的不同，为验证互联网用户身份的真实性与合法性提供了保障，也为网络攻击源追踪和定位提供了基础.

用户在向权威机构申请用户身份标识的时候，提供给用户的包括一个类 Email 地址格式的 User ID 和一个用于身份验证的 Key 文件，其中包含用户的公私钥对以及管理该用户的 EMS 的公钥. 在 SAMS 中身份认证与标识映射机制是集成在同一设备即 EMS 上来实现的，这就大大减小了系统实现和部署的复杂程度.

SAMS 提供两方面的身份认证机制. 一方面是 EMS 对请求注册的用户的认证. 用户在注册的时候需要使用 Key 文件中包含的用户私钥对注册消息签名，以证明注册用户身份的真实性，同时还可以使用 EMS 的公钥加密以防止消息被篡改. 已注册的用户需要定时向 EMS 发送保活消息并在 ELOC 发生变化时立即通知 EMS，这些消息中包含 EMS 与用户主机协商好的序列号，每交换一次消息序列号增大 1，如果消息丢失则重新协商序列号. 在此过程中 EMS 可以随时要求用户主机进行基于签名的身份验证. 另一方面是通信一端的主机对另一端主机身份的验证，其基本方法仍然是基于数字签名机制. 主机可以在任何它认为需要的时候向对端索取签名，例如对方主机的 ELOC 发生变化之后.

6　讨　论

这一节将对 SAMS 在性能上的一些折衷以及其它一些相关问题进行讨论.

(1) 端到端时延

SAMS 方案可能增大端到端时延的因素包括封装、解封装的时间开销，路径拉伸，以及发起通信前的映射查询时延. 首先，现有核心路由器的封装和解封装技术已极为成熟，能做到线速转发，因此带来的额外时延时间开销很小. 其次，路径拉伸仅发生在目的地址为少数非常用前缀的情形中，且可以通过 RMS 的合理部署来减小. 本文 7.1 节中通过实验评价了不同 RMS 部署对路径拉伸的影响. 最后，通过与 LISP 和 Shim6 的比较来讨论 SAMS 的映射查询时延.

在 LISP 方案中，主机先查询 DNS 从名字解析对端的 EID，然后发起通信. 当第一个分组到达 ITR 后可能会出现 Cache Miss，此时需要借助于映射系统解析对端 EID 到 RLOC 的映射，会引入额外的一个 RTT 的时延，当 ITR 已经缓存了映射时就不会有额外时延. 然而当 Cache Miss 时，ITR 对于触发查询的分组有两种可能的处理方式，一是丢弃该分组，二是缓存该分组直到查询到相应的映射信息或计时器超时. 在前者的情况下，发起通信的终端需要等待计时器超时并重传(例如 TCP 协议)，造成的时延将远不止一个 RTT，而后者的情况会占用大量的

路由器资源，不具有可扩展性. 在 Shim6 方案中，通信双方在发起通信前需要执行四次握手协议，额外时延为 2 个 RTT. 而 SAMS 在最坏的情况下，需要对 DNS 和 EMS 各查询一次，引入的额外时延与 Shim6 相同，为 2 个 RTT. 然而此额外时延产生于发起通信前的查询请求，可以类比为现有网络中多次查询 DNS 的情况，而实际的通信过程中不会有额外的时延. 因此与 LISP 相比，SAMS 将处理查询的开销从网络转移到终端，将时延的代价从通信中转移到通信前，获得了一定的改进效果.

(2) 映射系统可扩展性

一方面讨论 EMS 的可扩展性. EMS 只需要维护它所管辖区域内的用户，即只保存局部的有限的信息，与拥有全局信息的 DNS 相比数据量要少得多. 跨越 EMS 管辖区域的映射查询是通过扩展的 DNS 系统来实现的，而 DNS 系统具有较好的可扩展性. 对于用户数量较多的 AS，可以根据系统的处理能力增设多台 EMS 分别管理不同的用户(例如按地址段划分)和处理查询请求. 而对于少量访问较为频繁的地址(如知名站点)，可以通过扩展的 DNS 系统来保存这些少量的映射信息. 此外，终端会对映射信息做缓存，因此同一台主机不会在短时间内频繁地向同一台 EMS 查询相同的映射信息.

另一方面讨论扩展的 DNS 系统的可扩展性. 扩展的 DNS 系统主要保存两类映射. 第一类是域名到 EMS 的 ELOC 的映射，由于并不是每个用户的 User ID 都需要有一个域名与之对应(一般用户仅拥有 User ID，其映射记录保存在 EMS 中)，因此这类映射的管理与现有 DNS 系统享有相同的可扩展能力. 第二类是个别用户(如知名站点)的域名与其 User ID 及 ELOC 的映射，这类映射主要是为了加快查询速度，以及减小 EMS 处理查询的负担，由于这类映射记录的数量最多与现有 DNS 记录相同，因此不会对 DNS 系统带来可扩展性问题.

(3) 核心网流量策略

SAMS 方案中，部分流量会经过 RMS 来转发，而不是通常的 PE 到 PE 的路由方式，这会使得原本的网络流量发生一定变化. 然而 PE 也缓存了一部分映射，虽然这部分映射数量很少，但是它们对应的是网络中的常用前缀，即大部分流量还是由 PE 来路由，只有少部分流量需要经过 RMS. 此外，运营商还可以部署多个 RMS，并通过控制路由的发布和学习来实现自己的流量策略.

(4) RMS 会话可扩展性与安全性

RMS 建立 BGP 会话的主要作用是学习和发布全局映射信息，可以分为域间会话和域内会话两种情况来讨论其可扩展性. 在域间会话方面，RMS 与现有 BGP 边界路由器之间建立会话的方式一致，部署多少台 RMS 以及每台 RMS 与哪些邻居 AS 的 RMS 建立会话都是通过管理员配置决定的，因此在会话方面的可扩展性与安全性与现有的 BGP 协议一致. 而在域内会话方面，因为 RMS 只负责维护全局映射信息，会话建立的方式不影响域内路由，所以不需要在每一对 RMS 之间都建立会话，只需要使用较少量的会话以保证全局映射信息在域内的传播即可，这样一来就不会引入 iBGP 协议的域内会话数量平方增长的问题(参见路由反射器[22]).

(5) 封装/解封装开销

与 LISP 相比，SAMS 中可能会多引入一次封装/解封装的开销，这次封装/解封装发生在流量经过 RMS 的情况下. 然而这种情况仅出现在少数非常用前缀的流量中，而大多数流量因为 PE 对常用前缀的映射做了缓存，所以不需要经过 RMS，因而与 LISP 一样，只需要一次封装/解封装. 此外正如前面所说，现有路由系统的封装/解封装技术已极为成熟，在我们实现的原型系统中已经能够达到线速处理. 因此，封装的开销主要在于带宽的额外消耗，这部分开销与 LISP 引入的一样，属于能够承受的范围内.

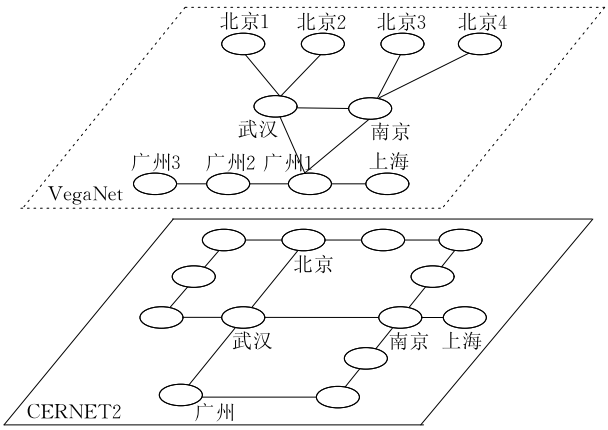
(6) 链路失效处理

SAMS 方案并不改变和影响现有网络的故障恢复能力，很多已有方案都可以与 SAMS 共存. 即使当 PE 和 CE 之间的链路发生故障时，也有一些解决方法，例如可以采用 AS 级别的多宿主连接，当链路出现故障时通过现有的路由系统达到收敛，或者通过构建冗余备份链路来达到处理故障的目的.

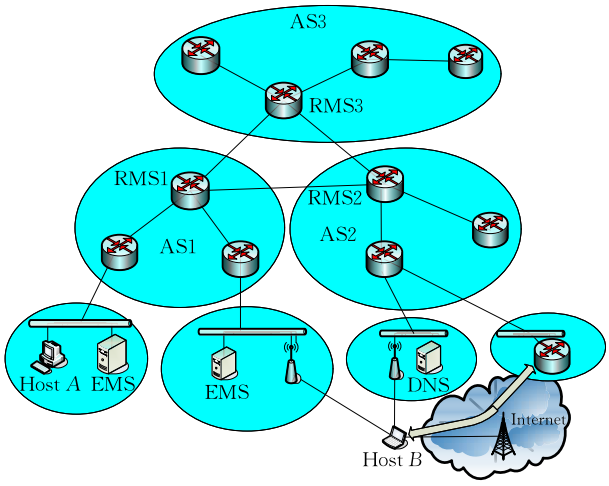
7 性能分析与评价

为了更好地评价 SAMS 方案，我们实现了 SAMS 原型系统. 开发的设备包括扩展协议栈的主机、扩展的 DNS 系统和 EMS，以及支持新型身份/位置分离方案的 PE 路由器和 RMS. 原型系统以高性能虚拟网络 VegaNet[23] 为平台，在 CERNET2 主干网的 10 个路由器节点上进行了部署，部署的节点和拓扑结构如图 4 所示. 实验网络由 3 个核心 AS 和 4 个边缘网络组成. Host A 与 Host B 分别属于不同的边缘网络，且 Host B 能通过不同的接入方式

分别接入 3 个边缘网络，包括无线局域网(WiFi)接入和 3G(TD-SCDMA)接入，其中 3G 接入将 Host B 连接到互联网，通过建立一条 VPN 隧道将流量引回到一个边缘网络. Host B 采用多种接入方式接入到多个边缘网络的目的是实验验证当主机移动或多宿主时的通信情况. 下面将分别对 SAMS 的多个方面的性能进行分析与评价.



(a) 虚拟拓扑的构建



(b) 在虚拟拓扑中搭建实验环境

图 4　原型系统 CERNET2 主干网部署

7.1　路由可扩展能力

图 5 给出了采用 SAMS 前后 CERNET2 核心网络路由表项数量的情况. 总体上，采用 SAMS 之后 25 个核心路由器节点中的平均路由表数量得到了大幅减少. 具体来说，IGP 路由基本保持不变，从用户 AS 学到的路由减少，这两部分在原来总的路由表中占的比例很小；从 Peer AS 学到的路由大幅减少，从占原来路由表的绝大多数减少到与 IGP 路由大致相当，具有很好的可扩展性. 此外，采用一个 RMS 或采用多个 RMS 对路由表大小的影响不大. 然而，采用多个 RMS 一方面可以预防单点失效问题，另一方面还可以减小因为部署 SAMS 而产生的分组转发路径拉伸，从而减小 SAMS 对端到端延迟的影响. 图 6 对端到端路径拉伸的情况进行了比较，被比较的方案包括仅部署一台 RMS 以及同时部署三台 RMS. 从图中可以看出，选择拓扑结构中更合适的节点作为 RMS 可以减少路径拉伸，例如将北京节点作为 RMS 时有 70% 的端到端路径长度不变，而如果将上海节点作为 RMS 则只有 30%. 同时使用三台 RMS 可以获得最小的路径拉伸，88% 的路径长度不变，只有不到 5% 的路径拉伸超过 2，而这部分路径拉伸可以通过在 PE 路由器上配置常用前缀来消除.

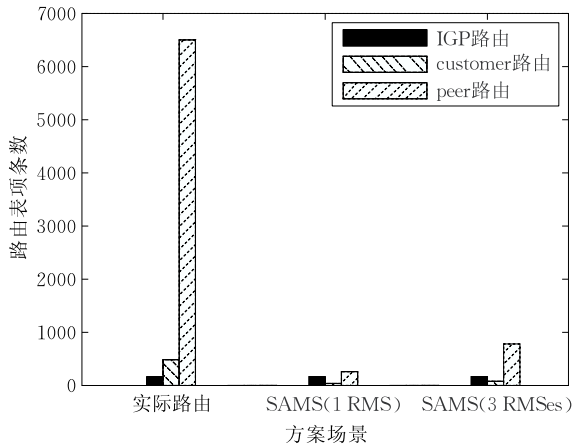


图 5　核心路由器前缀数量实验结果

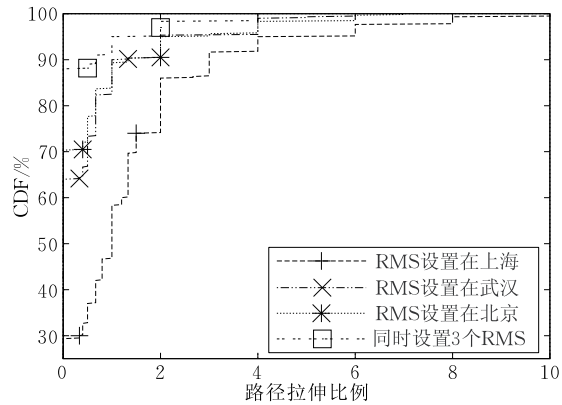


图 6　路径拉伸实验结果

7.2　标识快速切换

为了评价 SAMS 的标识快速切换能力，我们测量了 Host B 在两个接入网络之间每次切换所需的时间，图 7 和图 8 给出了实验结果，其中图 7 按照测试的时间顺序给出了每一次移动切换的时间，而图 8 给出了切换时间的累积分布. 可以看到切换的时间大体较为稳定，有大约 50% 的切换的切换时间在 1.5s～2s 之间，平均切换时间为 2.28s. 由于一般的

流媒体应用都会有 2～3 s 的缓存,因此这一切换时间能为这类应用提供移动中的平滑播放.这一实验结果说明 SAMS 能提供很好的多宿主和移动主机跨域高速切换的性能.

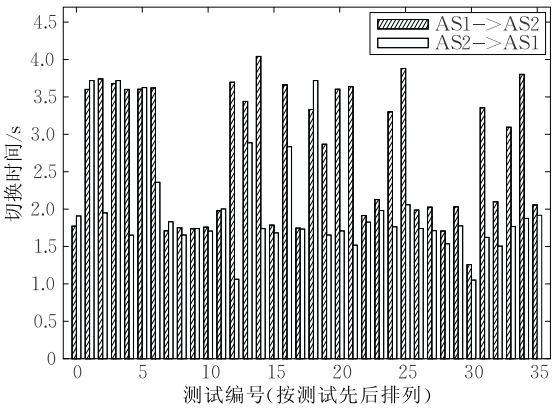


图 7 连续多次测试移动切换时间结果

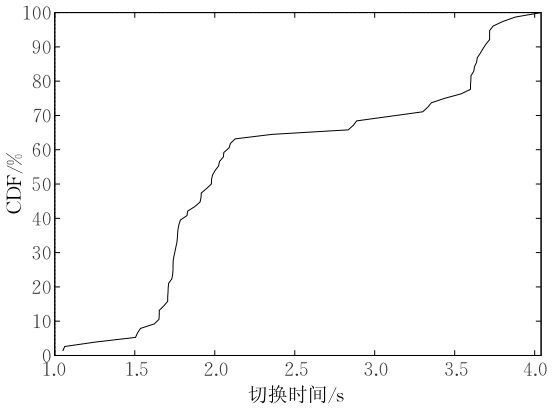


图 8 移动切换时间累积分布

为了分析影响标识切换时间的因素,为进一步优化奠定基础,我们建立了如图 9 所示的理论模型,它使用广义随机 Petri 网[24](Generalized Stochatisc Petri Net,GSPN)描述了移动主机在 ELOC 变化时通信双方的状态.表 2 给出了模型中各位置和变迁的含义.

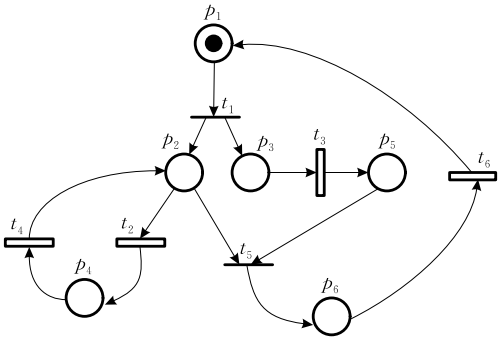


图 9 主机 ELOC 切换机制的 GSPN

表 2 主机移动 GSPN 中位置和变迁含义

符号	含义	符号	含义
p_1	主机 A 发生移动前的瞬间	t_1	主机 A 发生移动
p_2	对端主机 B 尝试发送分组	t_2	对端主机 B 向旧的 ELOC 发送分组
p_3	主机 A 获取 ELOC,检测 ELOC 变化,通知端主机 ELOC 变化	t_3	主机 A 完成获取 ELOC,检测 ELOC 变化,通知对端主机 ELOC 变化
p_4	对端主机 B 等待重传计时器超时	t_4	重传计时器超时
p_5	对端主机 B 收到了 ELOC 变化的消息,映射缓存更新完成	t_5	对端主机 B 选择向新的 ELOC 发送分组
p_6	对端主机 B 完成发送分组到新的 ELOC	t_6	对端主机 B 向新的 ELOC 发送的分组被传输

其中 t_2、t_3、t_4、t_6 是时间变迁,速率分别是 λ_2、λ_3、λ_4、λ_6.通过求解此模型可以得到一个标记从离开位置 p_1 到再次返回 p_1 的平均时间,即主机进行 ELOC 切换时不能接收到对端发来的分组的平均时间间隔:

$$T=\frac{1}{\lambda_3}+\frac{2}{\lambda_4}+\frac{1}{\lambda_6}.$$

这一时间间隔由三部分组成,一是获取新 ELOC 的时间加上检测本机 ELOC 变化的时间再加上发送消息通知对端的时间,二是重传计时器超时时间的 2 倍,三是对端到本主机新路径的传输延迟.针对这一结论,可以从两方面入手对主机移动时产生的延迟变化进行优化:(1)改进传输协议在 ELOC 切换时的重传时间间隔;(2)缩短检测本机 ELOC 变化的时间,可以依靠在内核设计特殊的机制来实现当网卡一旦变成可用状态时就立即发送消息.此外,采用双网卡进行位置标识切换也有利于减小切换时间.

7.3 身份认证开销

从功能看,SAMS 将身份标识与位置标识进行区分,并通过签名机制提供对用户身份真实性的认证,从而保证会话的可信与可靠.这种身份认证机制的安全性能是由签名算法来保证的,在本文中不做进一步的讨论,这里主要对身份认证的开销进行实验评价,以说明 SAMS 提供的真实身份认证特性不会对基本的转发功能造成影响.图 10 给出了在开启签名和不开启签名这两种情况下,用户主机与 EMS 之间的保活消息的往返时间实验结果,这段时间既包括了用户主机和 EMS 构造消息、计算和验证签名的时间,也包括了消息在网络中传输的时间.从图中可以看出开启签名与不开启签名相比额外开销很

少，两种情况下的平均处理延迟分别为 69.28 ms 与 64.44 ms. 这就说明了 SAMS 可以在不影响正常通信的情况下的提供安全性保证.

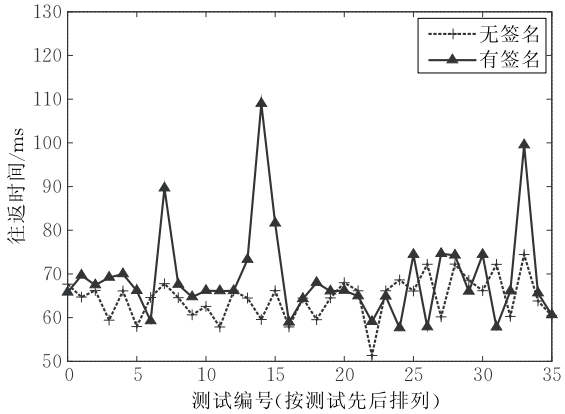


图 10　保活消息往返时间测试结果

7.4　可部署性

SAMS 既需要改动网络，也需要改动主机的协议栈，但它仍然提供了较好的对可部署性的支持. 总的来说，SAMS 既能提高核心网的路由可扩展性，也能为终端用户提供很多好处. 对于网络运营商来说，部署 SAMS 可以帮助他们有效地解决核心网路由可扩展性问题，且不影响流量工程、故障恢复等其它功能. 对于用户来说，SAMS 提供了对移动性和 site 级别的多宿主的支持，并且由于 SAMS 支持真实身份认证，当发生攻击时，部署了 SAMS 的非攻击者用户能够最早被排除嫌疑，这种先部署先获益的特性可以促进 SAMS 的部署. 此外，SAMS 对主机协议栈的修改很轻量，且与现有协议栈兼容，这使得在终端部署 SAMS 可以通过操作系统升级或打补丁的方式实现，无须用户自己动手.

SAMS 原型系统进行了长期实验，在峰值流量达到 600Mbps 的情况下，主干网设备仍然能很好地完成 ELOC 到 RLOC 之间的映射和分组转发功能. 此外，SAMS 原型系统在主机协议栈和应用程序的各种组合情况下都能很好地工作，这些都说明 SAMS 具有很好的可部署性. 下面给出 SAMS 端到端通信的几种典型的场景.

当通信双方的协议栈均为新协议栈且已分别向各自的 EMS 注册的时候，可以使用新开发的应用程序(用 User ID 或 Host ID 进行通信)，也可以使用传统的应用程序(用 IP 进行通信). 当使用传统应用程序通信时，需要事先在本地缓存中保存对端主机的 Host ID 和 ELOC 的映射信息，这可以通过先运行其它新开发的应用程序来实现，例如以对端的 User ID 为对象运行扩展的 ping 程序. 无论使用哪种应用程序，双方都是通过发送和接收包含 ID 头部的分组相互通信的.

当通信双方中只有一端的协议栈为新协议栈，而另一端为传统协议栈时，因为传统协议栈无法对含有 ID 头部的分组进行处理，所以只能依靠发送传统的 IP 分组进行通信. 此时，使用新协议栈的主机查询不到对端主机的 Host ID，映射缓存中也不会保存此信息，所以协议栈虽然包含身份标识层，但不会对分组添加 ID 头部，分组的发送和接收按照传统的流程进行. 总的来说，新协议栈的主机能够一边发送包含 ID 头部的分组与另一新协议栈主机通信，同时一边访问互联网上的传统站点，充分体现了 SAMS 的逐步部署能力.

8　总　结

互联网路由可扩展性问题亟待解决，与此同时，安全性与多宿主以及移动性在未来互联网中的重要程度也越来越高. 学术界普遍认为将 IP 地址的身份标识语义和位置标识语义分离开来是解决这些问题的根本方法，但现有的身份/位置分离方案不能同时很好地解决这些问题. 本文提出了一种新型身份/位置分离方案 SAMS. SAMS 将用户身份标识，主机身份标识，边缘网络位置标识和核心网络位置标识分离，构成方案体系结构的要素包括边缘网络映射服务器 EMS，扩展的 DNS 系统，核心网络边缘路由器 PE 以及核心网络映射服务器 RMS，很好地分离了 IP 地址的语义. SAMS 大幅提高了核心网路由可扩展性，为用户真实身份认证提出了一种较好的解决方案，并能很好地支持多宿主和大规模移动. 我们实现了 SAMS 的原型系统，在 CERNET2 主干网节点上部署了大规模的实验网络并进行了长时间的实验，实验结果表明原型系统具有预期的功能和良好的性能，并具备很强的兼容性和可逐步部署能力.

参　考　文　献

[1] Meyer D, Zhang L, Fall K. Report from IAB workshop on routing and addressing. IETF RFC 4984, 2007 [EB/OL]. http://tools.ietf.org/html/rfc4984

[2] Tang Ming-Dong, Zhang Guo-Qing, Yang Jing, Zhang Guo-Qiang. Scalable routing for the Internet. Journal of Software, 2010, 21(10): 2524-2541(in Chinese)

(唐明董，张国清，杨景，张国强. 互联网可扩展路由. 软件学报，2010，21(10)：2524-2541)

[3] Zhang Wei, Bi Jun, Wu Jian-Ping. Scalability of Internet inter-domain routing. Journal of Software, 2011, 22(1): 84-100(in Chinese)
(张威，毕军，吴建平. 互联网域间路由可扩展性. 软件学报，2011，22(1)：84-100)

[4] Huston G. Analyzing the Internet's BGP routing table. The Internet Protocol Journal, 2001, 4(1): 2-15

[5] Perkins C. IP mobility support for IPv4. IETF RFC 3344, 2002[EB/OL]. http://www.ietf.org/rfc/rfc3344.txt

[6] Johnson D, Perkins C, Arkko J. Mobility support in IPv6. IETF RFC 3775, 2004[EB/OL]. http://www.ietf.org/rfc/rfc3775.txt

[7] Hou Jie, Liu Ya-Ping, Gong Zheng-Hu. Key techniques of identifier-based routing. Journal of Software, 2010, 21(6): 1326-1340(in Chinese)
(侯婕，刘亚萍，龚正虎. 标识路由关键技术. 软件学报，2010，21(6)：1326-1340)

[8] Quoitin B, Iannone L, Launois C, Bonaventure O. Evaluating the benefits of the locator/identifier separation//Proceedings of the 2nd ACM/IEEE International Workshop on Mobility in the Evolving Internet Architecture. Kyoto, Japan, 2007: Article No. 5

[9] Farinacci D, Fuller V, Meyer D, Lewis D. The locator/ID separation protocol (LISP). IETF RFC 6830, 2013[EB/OL]. http://tools.ietf.org/html/rfc6830

[10] Nordmark E, Bagnulo M. Shim6: Level 3 multihoming Shim protocol for IPv6. IETF RFC 5533, 2009[EB/OL]. http://tools.ietf.org/html/rfc5533

[11] Moskowitz R, Nikander P. Host identity protocol architecture. IETF RFC 4423, 2006[EB/OL]. http://www.ietf.org/rfc/rfc4423.txt

[12] 第二代中国教育和科研计算机网 CERNET2[EB/OL]. http://www.cernet2.edu.cn

[13] Iannone L, Bonaventure O. On the cost of caching locator/ID mappings//Proceedings of the 3rd International Conference on Emerging Networking EXperiments and Technologies (CoNEXT). New York, USA, 2007: Article No. 7

[14] Nikander P, Henderson T, Vogt C, Arkko J. End-host mobility and multihoming with the host identity protocol. IETF RFC 5206, 2008[EB/OL]. http://tools.ietf.org/html/rfc5206

[15] Nikander P, Ylitalo J, Wall J. Integrating security, mobility, and multi-homing in a HIP way//Proceedings of the 10th Annual Network and Distributed Systems Security Symposium (NDSS'03). San Diego, USA, 2003: 87-98

[16] Zan Feng-Biao, Xu Ming-Wei, Wu Jian-Ping. Survey of host identity protocol (HIP). Journal of Chinese Computer Systems, 2007, 28(2): 224-228(in Chinese)
(昝风彪，徐明伟，吴建平. 主机标识协议(HIP)研究综述. 小型微型计算机系统，2007，28(2)：224-228)

[17] Dhraief A, Montavont N. Toward mobility and multihoming unification the Shim6 protocol: A case study//Proceedings of the Wireless Communications and Networking Conference (WCNC). Las Vegas, USA, 2008: 2840-2845

[18] Barre S, Bonaventure O. Improved path exploration in Shim6-based multihoming//Proceedings of the SIGCOMM Workshop on IPv6 and the Future of the Internet. Kyoto, Japan, 2007: 29-35

[19] Mathy L, Iannone L. LISP-DHT: Towards a DHT to map identifiers onto locators//Proceedings of the 4th International Conference on Emerging Networking EXperiments and Technologies (CoNEXT). Madrid, Spain, 2008: Article No. 61

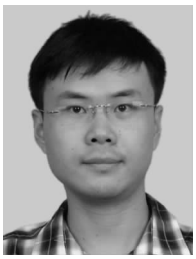
[20] Luo H, Qin Y, Zhang H. A DHT-based identifier-to-locator mapping approach for a scalable Internet. IEEE Transactions on Parallel and Distributed Systems, 2009, 20(12): 1790-1802

[21] Pan J, Jain R, Paul S, So-in C. MILSA: A new evolutionary architecture for scalability, mobility, and multihoming in the future internet. IEEE Journal on Selected Areas in Communications—Special Issue Title on Scaling the Internet Routing System: An Interim Report, 2010, 28(8): 1344-1362

[22] Bates T, Chen E, Chandra R. BGP route reflection-An alternative to full mesh internal BGP (IBGP). IETF RFC 4456, 2006[EB/OL]. http://tools.ietf.org/html/rfc4456

[23] Chen Wen-Long, Xu Ming-Wei, Yang Yang, Li Qi, Ma Dong-Chao. Virtual network with high performance: VegaNet. Chinese Journal of Computers, 2010, 33(1): 63-73(in Chinese)
(陈文龙，徐明伟，杨扬，李琦，马东超. 高性能虚拟网络 VegaNet. 计算机学报，2010，33(1)：63-73)

[24] Holiday M A, Venon M K. A generalized timed Petri net model for performance analysis. IEEE Transactions on Software Engineering, 1987, 11(10): 1216-1225



YANG Yuan, born in 1984, Ph. D. candidate. His research interests include computer network architecture and Internet routing.

XU Ming-Wei, born in 1971, Ph. D., professor. His research interests include computer network architecture, high-speed router architecture and Internet routing.

YANG Jia-Hai, born in 1966, Ph. D., professor. His research interests include Internet management, network measurement, traffic engineering and network operation.

DONG Qing-Zhou, born in 1985, M. S. candidate. His

research interests include networking and switching, network architecture.

CHEN Wen-Long, born in 1976, Ph. D., lecturer. His research interests include network architecture and network protocols.

WANG Hui, born in 1977, Ph. D., assistant professor. Her research interests include network measurement, network management, inter-domain routing protocols and economic analysis of ISPs' behavior.

ZHANG Zhi-Chao, born in 1986, M. S. candidate. His research interests include Internet transport protocols.

Background

This work is supported by the National Natural Science Foundation of China under Grant Nos. 61073166 and 61170211, the 863 Program of China under Grant Nos. 2009AA01z251 and 2011AA01A101, the 973 Program of China under Grant No. 2012CB315803, Research Fund for the Doctoral Program of Higher Education under Grant No. 20110002110056.

The addressing architecture of the current Internet faces many problems. The cause of several problems is that IP address has the semantics of both endpoint identifier (ID) and locator. Such double semantics, on one hand, make the address space flatter and flatter. So the routing entries cannot be aggregated effectively and are growing rapidly. This causes the rouging scalability problem of the DFZ (default-free-zone). On the other hand, mobile or multi-homed endpoints cannot switch smoothly from one interface to another. Such problems are becoming serious as the Internet is becoming larger and more complex, and the need for mobility and multi-homing is also in a fast growth.

There are studies on the design of new architectures that separate ID and locator. Many schemes are proposed, such as LISP, HIP, and Shim6, and some of them have become IETF standards. Nevertheless, existing approaches can solve only one of the routing scalability problem and mobility/multi-homing problem, but not both; and most approaches can solve the problems under full deployment and in a long term running. Furthermore, the authentication of users in an ID-locator separation scheme is also very important.

To overcome such challenges, we, in this paper, propose a novel Scalable Authentic Mobile identifier-locator Separation scheme (SAMS). SAMS separates user ID, host ID, edge locator and routing locator, and combines them with a reasonable architecture. SAMS makes large improvement on routing scalability as well as multi-homing and mobility, and it also supports authentic identifiers. We implemented a prototype of SAMS, and did large-scale deployment on CERNET2 backbone. The experiment results verified the effectiveness and compatibility of SAMS.

The work in this paper is a major part of the supporting 863 Program, and solves the key problem of routing scalability of the supporting 973 Program.