文章编号:1001 - 9081(2006)04 - 0922 - 04

基于探测对 Symmetric NAT 与端口受限 NAT 的穿透方案

王 勇,崔修涛,吕 钊,李子成
(华东师范大学计算机科学系,上海 200062)
(ywang@ica.stc.sh.cn)

摘 要:在采用 P2P 架构的网络音视频应用中,不可避免会涉及到 NAT 穿透的问题。IETF 为此开发了 STUN 标准,但是 STUN 协议中提出的方案无法穿透 Symmetric NAT。通过分析对称 NAT 原理以及一些实际网络中的 NAT 布署的情况,提出了基于探测的穿透方案。经过多种网络环境的测试,达到了预计的目标。

关键词:Cone NAT;端口受限;对称 NAT;穿透;STUN
中图分类号: TP393.09 **文献标识码**:A

Detection-based traversal solution for connection between symmetric NAT and port-restricted NAT

WANG Yong, CUI Xiu-tao, Lü Tao, LI Zi-cheng
(*Department of Computer Science, East China Normal University, Shanghai* 20062)

Abstract: To resolve the problem of NAT traversal in the applications of audio and video based on P2P technology, STUN as a IETF standard was proposed. But it can't traverse the Symmetric NAT. A new solution to resolve the problem was presented by detection with the analysis of Symmetric NAT and the its deploy. Some tests were taken in real situations and the results reached the goal.

Key words: Cone NAT; port-restricted; symmetric NAT; traversal; STUN

0 引言

在网络音视频应用中,因为其数据流量大而通常会采用客户机直接连接的 P2P 技术。但是在目前 Internet 体系架构下由于 IPv4 地址短缺,使得许多客户机都是通过 NAT 技术来共用一个 Internet 网络(通常称为公网)IP 地址[1]。位于不同内部网络的客户机的 IP 都是私网地址而无法直接相连接。网络音视频应用通常是采用无连接的 UDP 协议来传送数据,根据 NAT 工作的特点,UDP 报文对于传统的 NAT 具有可穿透性。

在如何进行不同内部网用户之间的数据交换,众多的研究已经提出了一些切实可行的技术方案。主要可以分为两大类:1)通过不同结构和协议的某种结构的代理转发的方案;2)先直接进行 NAT 穿透的方案,无法穿透的时候再采用代理转发[6]。采用高效的结构和多个代理的转发方式可以减轻单个服务器的压力而提供良好的性能,但是对于网络带宽资源的消耗和传递延迟却没有任何改变。

能够穿透 NAT 进行直接的点到点数据传输是最理想的方案。为此,IETF 制定了 STUN 协议来解决 UDP 报文对于传统 NAT 的穿透问题[2]。但是 STUN 协议并不能穿透当需要建立连接的双方 NAT 设备都为 Symmetric NAT 或一方为 Symmetric NAT,另外一方为端口限制型 NAT(IP 限制型 NAT 对于 P2P 穿透与端口限制型 NAT 类似,后面都不再特别说明)的情况。随着网络的安全性的不断提高,防火墙的大量部署使得非限制型 NAT 也在 P2P 连接中表现出限制型 NAT 的特征,新部署的 NAT 设备越来越多的是 Symmetric NAT。因此,如何能够穿透这种类型的网络连接,提高点对点应用(特别是数据量非常大的网络音视频应用)的网络利用效率,是一项具有现实意义的研究。本文主要分析了 Symmetric NAT 的特点以及各种 NAT 在网络部署的情况,提出了基于探测的穿透方案。

1 NAT 原理与分类

当位于 NAT 设备内部网络中的客户机向外界发送 UDP 报文的时候,NAT 设备会根据输入 UDP 报文中的四元组[源 IP,源端口,目标 IP,目标端口]的全部或者部分信息分配一个 NAT 设备的 UDP 端口。同时建立该端口和输入地址对的映射表,然后将 UDP 报文中的源 IP 和源端口修改为自己的公网 IP 和该分配的输出端口。当目标主机返回 UDP 报文的时候,NAT 设备根据从外部收到的报文的四元组[源 IP,源端口,目标 IP,目标端口],通过查找之前定义的映射表,查找到该报文是发送给哪个内部客户机的。然后相应的将该报文的目标地址替换为内部客户机的 IP 地址和端口后,转发给内部客户机,过程见图 1。

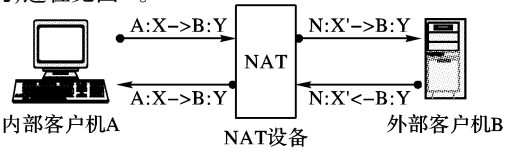

图 1 NAT 原理

收稿日期:2005 - 11 - 01;**修订日期**:2005 - 12 - 22 **基金项目**:上海市科技攻关计划资助项目(035115038)

作者简介:王勇(1974 -),男,重庆人,工程师,博士研究生,主要研究方向:流媒体编码与传输、数字电视; 崔修涛(1976 -),男,四川绵阳人,博士研究生,主要研究方向:数据会议系统、地理信息系统; 吕钊(1970 -),女,四川江油人,副教授,博士,主要研究方向:多媒体技术、CSCW、信息安全; 李子成(1982 -),男,安徽人,硕士研究生,主要研究方向:多媒体技术.

根据 NAT 设备对于收到的内部和外部 UDP 报文时候,对于四元组[源 IP,源端口,目标 IP,目标端口]的检查机制和转换机制的不同, STUN 协议将 NAT 分为四种类型[2,4]:

(1)锥型 NAT (Full Cone)

a)内部客户机 A 用同一个端口 X 发出的任何 UDP 报文,NAT 设备都为其分配同一个输出端口 X′,并建立对应关系表条目[A:X <->X′];

b)随后,NAT 设备将任何公网上主机 B 向 NAT 的端口 X′发送来的 UDP 报文都转发给 A:X。

(2) IP 受限型 NAT (Restricted Cone)

a)内部客户机 A 用同一个端口 X 发出的任何 UDP 报文,NAT 设备都为其分配同一个输出端口 X′。并建立对应关系表条目[A:X <->X′]。和(1)相同;

b)随后,对于任何公网上主机 B 向 NAT 的端口 X′发送来的 UDP 报文,检查是否该内部客户机 A 向 B 发送过数据:如果发送过,就将该报文转发给 A:X;否则,就拒绝接收。

(3) 端口受限型 NAT(Port Restricted Cone)

a)内部客户机 A 用同一个端口 X 发出的任何 UDP 报文,NAT 设备都为其分配同一个输出端口 X′。并建立对应关系表条目[A:X <->X′]。和(1)相同;

b)随后,对于任何公网上主机 B 用端口 Y 向 NAT 的该端口发送来的 UDP 报文,检查该内部客户机 A 是否曾用端口 X 向 B:Y 发送过数据,如果发送过,就将该报文转发给 A:X;否则,就拒绝接收。

(4) 对称型(Symmetric)

a)当 A 用端口 X 向 B 的 Y 端口发送数据的时候,NAT 为其分配端口 X′。当[A,X,B,Y]中任何一项变化的时候,NAT 都为其分配新的端口号。这个限制比前面严格;

b)同(3)的 b)限制。

2　基于 STUN 协议的 NAT 穿透方案

图 2 解释:A 和 B 分别为 2 个拥有内部网客户机,NAT A 和 NAT B 为 2 个不同局域网的 NAT 设备。S 为一个公网的服务器,在 5000 号端口接收 A 和 B 的 UDP 报文,对穿透过程进行协调。

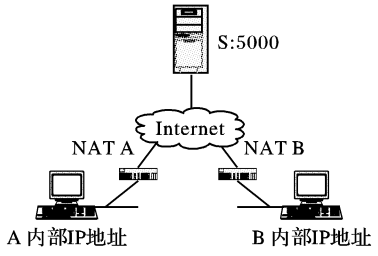

图 2　NAT 部署

STUN 协议通过架设在公网上的 STUN 服务器帮助内网用户得到 NAT 设备为它与公网连接分配的输出映射端口号,从而通过该端口号实现 NAT 穿透和点到点直接连接。下面对根据 STUN 协议进行的穿透过程各种情况进行说明(如图 2):

1) NAT A 和 NAT B 至少有一个是 Cone NAT 的情况,根据 STUN 协议可以进行穿透。

假设 NAT A 为 Cone NAT,NAT B 为任意 NAT,穿透过程如下:

(1)A 向 S 发送一个 UDP 报文[A:X ->S:5000];

(2)NAT A 为其分配一个端口号 X′,然后修改 UDP 报文源 IP 和源端口为[NAT A:X′->S:5000];

(3)B 也同样发送一个 UDP 报文给服务器 S,建立 S 与 B 之间的连接;

(4)S 利用和 B 的连接,发送命令给 B:请发送数据到[NAT A:X′];

(5)B 按照 S 的要求,发送 UDP 报文到[NAT A:X′]之后,NAT A 就根据映射表将该报文转发给 A:X;

(6)这样,A 可以根据该报文中的 NAT B 的地址返回数据,UDP 报文在 NAT B 处被同样的转发给了 B;

(7)A 和 B 直接连接完成。

2) NAT A 和 NAT B 均为端口受限 NAT,根据 STUN 协议可以进行穿透,过程如下:

(1)A 向 S 发送一个 UDP 报文[A:X ->S:5000];

(2)NAT A 为其分配一个端口号 X′,然后修改报文源 IP 和源端口为[NAT A:X′->S:5000];

(3)B 也同样发送一个 UDP 报文给服务器 S,建立 S 与 B 之间的连接;

简化为[(B:Y ->NAT B:Y′) ->S:5000]

(4)S 收到 UDP 报文之后获得信息:NAT A 的端口 X′和 NAT B 的端口 Y′;

S 利用和 B 的连接发送命令给 B:请用 Y 端口发送数据到[NAT A:X′];

(5)B 按照 S 的要求,发送 UDP 报文到[NAT A:X′];

因为 NAT A 是限制型,对于该 UDP 报文会拒绝接收。因为只有 NAT A 没有发送过数据给 NAT B 的这个端口,因此将该数据当作非法的 UDP 报文而拒绝。但是对于 NAT B 来说,虽然该 UDP 报文被拒绝,但在 NAT B 上留下了一个入口:如果 NAT A 此时发送[NAT A :X′->NAT B:Y′],NAT B 会接收并且转发给内部的 B:Y;

(6)服务器 S 继续发出命令给 A:请用 X 端口发送数据到[NAT B:Y′];

(7)根据(5)的分析:[(A:X ->NAT A:X′) ->NAT B:Y′]被 NAT B 接收,并转发给 B:Y;以后 NAT A 就不再拒绝接收[(B:Y ->NAT B:Y′) ->NAT A :X′]的 UDP 报文了;

(8)A 和 B 直接连接完成。

但是,在这个过程中有一个限制:对于某些防火墙或者 NAT,如果它检测到某个端口收到过外部来的非法数据,当作网络攻击而暂时关闭该端口,直接连接就无法建立。

3) 当其中一个 NAT 为端口受限 Cone NAT,而另外一个为对称 NAT 的情况,根据 STUN 协议无法进行穿透。

如果 NAT A 为对称型 NAT,在 2)的步骤(7)时,不存在[(A:X ->NAT A:X′) ->NAT B:Y′],而是变成了[(A:X ->NAT A:X′+Δx) ->NAT B:Y′]。对称型 NAT 会将目标端口不同的连接当作是新的连接而分配一个新的端口号,而端口限制型的 NAT B 不会接收[NAT A:X′+Δx ->NAT B:Y′]的 UDP 报文,导致穿透失败。

3　对称 NAT 的穿透方案

对存在 Symmetric NAT 的网络环境分析,可以概括具有如下的几种情况:

1)Symmetric NAT 对于从内部网络依次接收到的新连接(即<源 IP,源端口,目标 IP,目标端口>四元组在 NAT 映射

表中找不到完全匹配的项),分配的输出端口是按照空闲端口依次连续分配。如果 NAT 为某个内部网络的新连接分配的端口号是 X,则其将为第 N 个新连接分配的端口号为:$X+n$。

2)针对 NAT 从内网依次接收到的新连接,分配的输出端口是在一定端口范围内随机分配。如果 NAT 为某个内部网络的新连接分配的端口号是 X,则其将为后面到来的第 N 个新连接分配的端口号为:$X+\Delta xn$(每个 Δxn 都是由于 NAT 自行产生的随机数)。

3)对于在内部客户到公网之间存在多个 NAT 串联的情况:如果其中存在一个或者多个 Symmetric NAT,对于连接到公网的那个 NAT 的输出端口号的分配变化情况分析如下:

a)只要存在 Symmetric NAT,对于内部产生的每个连接:[源 IP,源端口,目标 IP,目标端口]任何一项变化,通过 Symmetric NAT 的时候,必然被分配一个新的端口号。因此不管 Symmetric NAT 串联客户到公网之间的哪个位置,在最后 NAT 的输出端口号必然选择一个新的端口号。

b)新端口号的分配规律通常是根据最后一个 NAT 自己对新连接分配输出端口的策略。

此种情况的输出端口号分配情况属于上面两种情况之一。因此,可以不特别考虑 NAT 串联的情况。

3.1 对于第一种情况下的 NAT 穿透方案分析

可以合理的作如下假设:

(1)假设 NAT A 为 Symmetric NAT,NAT B 为端口受限型 NAT;且 NAT A 对每个新连接分配的输出端口每次增加 1。

(2)当 NAT A 内部有多个客户同时在利用 NAT A 访问外部网络的时候,设 A 发送数据给 S 和给 B 的连接之间的产生了 Δz 个新连接。同时考虑到可能有少量端口(设为 Δw 个)已经在使用中,分配的端口号将顺延。

则可以预测当 A 发送数据给 B 的时候,NAT A 为其分配的输出端口号相对于 X' 的偏移量为:

$\Delta x=\Delta z+\Delta w+1$,且 $\Delta z\geqslant 0,\Delta w\geqslant 0$

因为在穿透过程中,两次数据的发送间隔不会很长,Δx 是一个较小范围内的正数。

通过上面的设定分析,可以修改穿透过程为:

a)A 向 S 发送一个 UDP 报文[A:X -> S:5000],建立 S 与 A 之间的连接;

[(A:X -> NAT A: X') -> S:5000]

b)B 也同样发送一个 UDP 报文给服务器 S,建立 S 与 B 之间的连接;

[(B:Y -> NAT B:Y') -]S:5000]

c)S 收到 UDP 报文之后知道:NAT A 的端口 X'和 NAT B 的端口 Y';

d)S 首先发送命令给 A: 请用 X 端口发送数据到[NAT B:Y'];

e)当 A 收到命令并执行,NAT A 为该连接分配新的输出端口 X'+Δx;

[(A:X -> NAT A:X'+Δx) -> NAT B:Y']

f)该 UDP 报文被 NAT B 收到后作为非法 UDP 报文而拒绝接收,但是此时 NAT A 上留下了一个入口:将会接收[NAT B:Y' ->(NAT A:X'+Δx -> a:X)]的 UDP 报文

g)S 发送命令给 B:用端口 Y 连续发送多个 UDP 报文到[NAT A:X+1] - - -[NAT A:X+MAXTEST];

MAXTEST(穿透测试次数最大值)由 S 根据经验和性能的需求来设定,当 MAXTEST > Δx,穿透就可以成功;

h)B 收到命令并开始执行探测过程(Δt 为每次分配的端口号的偏移值):

当 $\Delta t<\Delta x$ 的时候,[(B:Y -> NAT B: Y') -> NAT A: X'+Δt],这些 UDP 报文都会被 NAT A 拒绝接收;

而当 $\Delta t=\Delta x$ 的时候,[(B:Y -> NAT B: Y') -> NAT A:X'+Δx],NAT A 接收该数据并转发给 A,同时在 NAT B 上也留下了一个入口,NAT B 将接收 A 对该 UDP 报文的回复 UDP 报文[(A:X -> NAT A:X'+Δx) -> NAT B:Y']。因为 B 不知道 Δx 的数值,只有当 B 收到 A 的回复数据之后[NAT A:X'+Δx] -> B:Y],才知道直接连接已经建立成功。这个时候就可以停止探测了。

i)A 和 B 直接连接成功。

测试结果:Δx 的数值不会很大,B 只需要发送较少的探测连接就可以成功地和 A 建立直接的连接。在不同的时间,对实际使用的该类对称 NAT 的网络进行测试,结果如图 3、图 4 所示。图 3 中,Δx 的数值比较小,集中在 20 次以内;图 4 中,当设定最大穿透次数为 50 (横坐标的偏移量缩小了 10 倍),就能够获得 90% 以上的穿透成功率;当设置最大穿透次数为 120,就能获得接近 100% 的穿透成功率。

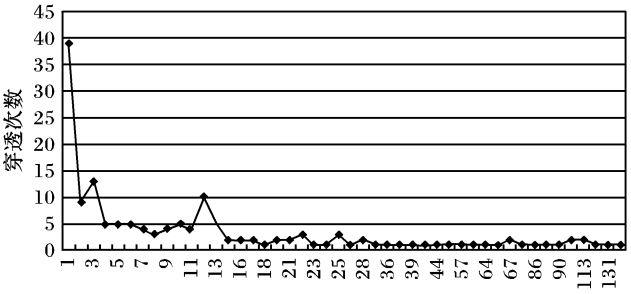

图 3 端口偏移/穿透次数

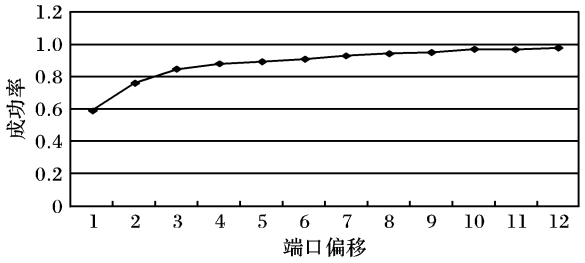

图 4 穿透成功率/端口偏移(缩小 10 倍)

3.2 对于第二种情况下的 NAT 穿透方案分析

现在很多安全性较高的 NAT 设备对于输出端口的分配是随机的。并且在很多防火墙和 NAT 功能合并的设备,其输出端口的分配也是随机的。

可以合理的假设:

a)NAT A 对新连接的输出端口分配是每次随机变化,两次之间偏移量为 Δx (Δx 可能为负);

b)当 NAT A 内部有多个客户同时利用 NAT A 访问外部网络。设 A 发送数据给 S 和 B 的连接之间产生了 z 个新连接;

c)虽然每次 Δx 值的选择是由 NAT 设备通过某种方式随机选择,对具体的设备和网络环境而极大的不同。但通常每次分配的端口号之间都会具有一定的函数关系或者统计上的相关性。

当 A 连接 S 的时候,如果 NAT A 为其分配输出端口号 X',则 A 发往 NAT B 的时候,NAT A 为其分配输出端口号为:

$X' + \Delta z, \Delta z = f(X', z)$

通过在多种网络条件下的测试和分析,也证实了多数情况下存在具有某种特征的分布特性的 Δz 值。因此,可以通过其分布特性来预测 Δz 所属范围,并对该范围端口进行试探的来实施穿透。

测试的结果:在不同时间对部署该类 NAT 的网络进行了测试,结果都如图 5、图 6 所示。图 5 中,Δz 的数值均匀分布在中心点的两边;图 6 中,当设定最大穿透测试端口偏移量为 1200(横坐标的偏移量缩小了 100 倍),能获得 80% 以上的穿透成功率;当设置最大穿透测试端口偏移量接近 2000,能获得接近 100% 的穿透成功率。

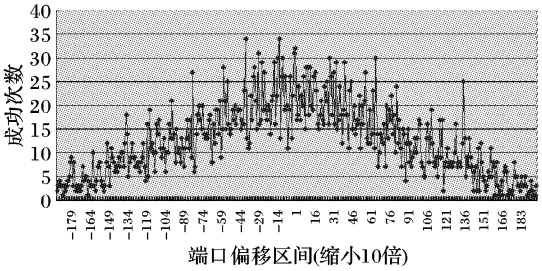

图 5　穿透成功次数/端口偏移区间(缩小 10 倍)

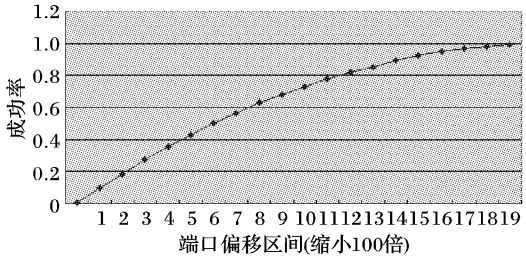

图 6　成功率/端口偏移区间(缩小 100 倍)

3.3　穿透方案的实施和改进

在实际实施中,还采用一些对于特定网络环境的经验办法来提高性能:通常在发送穿透消息前,先随机发送几个假的穿透包(即不是向目标 NAT B 的指定端口发送的 UDP 报文),然后再发送穿透消息,这样常常能减少试探次数并提高穿透的成功率。对于特定网络环境中的 Δz,可以通过一些测量方法来进行估计。这些测量方法在实际应用中,可以通过客户端与服务器之间多建立几个连接来进行。同时在应用规模不大的情况下,服务器端也可以保存一些关于某些网络的 NAT 分配的特征,用来调整 Δz 的估计范围。在具体应用中,采用在探测直接连接的过程中同时进行服务器中转数据的方法,可以减少探测对应用的影响。

4　结语

在实际的网络中,常常都部署有防火墙。对于网络中即使不是端口限制 NAT,但是因为防火墙的原因,也使得内部的普通 Cone NAT 的行为变成了端口限制型 NAT。而且现在很多 NAT 与防火墙合并的软硬件设备都采用了对称 NAT。因此 P2P 应用中,常常无法用已有的稳定方法(例如 STUN)进行直接的 P2P 连接。根据对称 NAT 的特征和 NAT 具体实现时候采用的不同技术和策略,本文试着采用一些经验方法和试探方法,提出了穿透方案,并在实际网络中的测试,得到了预期的效果。此方案虽然不具有稳定的穿透性能,但能够在较多的场合尽可能的建立 P2P 连接。此方案已经在某项目中得到应用,提高了该项目平台中音视频流穿透 NAT 的能力,取得了良好的效果。

参考文献:

[1] SRISURENSH P, NETWORKS J, EGEVANG K. Traditional IP Network Address Translator(Traditional NAT), RFC 3022[S]. IETF, 2001.

[2] ROSENBERG J, WEINBERGER J. STUN Simple Traversal of User Datagram Protocol (UDP) Through Network Address Translators (NATs), RFC3489[S]. IETF, 2003.

[3] STUKAS M, SICKER DC. An Evaluation of VoIP Traversal of Firewalls and NATs within an Enterprise Environment[J], Information Systems Frontiers, 2004, 6(3): 219 - 228.

[4] 何宝宏. 浅析 NAT 的类型[J]. 电信网技术, 2004, (8).

[5] 吴煜卓, 童恒庆, 刘喜雨. 基于 UDP 协议穿透 NAT 代理的研究与设计[J]. 微机发展, 2005, 15(1): 122 - 124.

[6] 柯金水, 王芙蓉, 戴彬. 基于软交换的 NAT/防火墙穿透技术研究[J]. 天津通信技术, 2004, (2): 34 - 39.

(上接第 921 页)

[4] FU XD, SHI WS, AKKERMAN A. CANS: Composable, Adaptive Network Services Infrastructure[A]. Proceedings of 3rd USENIX Symposium Internet Technologies and Systems[C]. 2001.

[5] ZHAO BY, KUBIATOWICZ JD, JOSEPH AD. Tapestry: An Infrastructure for Fault-tolerant Wide-area Location and Routing[R]. UC Berkeley Technical Report UCB//CSD- 01-1141, 2000.

[6] GANESAN P, MANKU GS. Optimal Routing in Chord[D]. Stanford University, SODA 2004.

[7] DABEK F, KAASHOEK MF, KARGER D, *et al*. Wide-area cooperative storage with CFS[A]. ACM SOSP'01[C]. Banff, Canada, 2001.

[8] DABEK F. A Cooperative File System[D]. Master's Thesis, MIT, 2001.

[9] STANDARD SH. National Institute of Standards and Technology, FIPS PUB 180-1[S]. 1995.

[10] CANNY J. Secure Hash Algorithms[EB/OL]. www. cs. berkeley. edu/ ~ jfc/cs174/lecs/lec22/lec22. pdf, 2005.

[11] JOVANOVIC MA, ANNEXSTEIN FS, BERMAN KA. Scalability Issues in Large Peer-to-Peer Networks - A Case Study of Gnutella Network[EB/OL]. http://www. ececs. uc. edu / ~ mjovanov/Research /paper. html, 2001.

[12] Kazaa[EB/OL]. http://www. kazaa. com/, 2002.

[13] YANGB, GARCIA - MOLINA H. Efficient Search in Peer - to - Peer Networks[R]. Technical Report of Stanford Database Group, In ICDCS, 2002.

[14] CRESPO A, GARCIA-MOLINA H. Routing indices for peer-to-peer systems[R]. Proceedings of the 22nd International Conference on Distributed Computing Systems(ICDCS 2002)[C]. 2002.

[15] KUBIATOWICZ J, BINDEL D, CHEN Y. OceanStore: An Extremely Wide-Area Storage System[R]. U. C. Berkeley Technical Report UCB//CSD- 00-1102, 1999.

[16] ZHAO BY, DUAN Y, HUANG L. Brocade: landmark routing on overlay networks[A]. First International Workshop on Peer-to-Peer Systems(IPTPS)[C]. Cambridge, MA, 2002.

[17] LIBEN-NOWELL D, BALAKRISHNAN H, KARGER D. Observations on the Dynamic Evolution of Peer-to-Peer Networks[A]. Proceedings of the First International Workshop on Peer-to-Peer Systems(IPTPS '02)[C]. Cambridge, MA, 2002.

[18] CASTRO M, DRUSCHEL P, HU YC. Exploiting network proximity in distributed hash tables[A]. FuDiCo 2002: International Workshop on Future Directions in Distributed Computing. University of Bologna Residential Center Bertinoro(Forli)[C]. Italy, 2002.