

# 一种新的 IPv6 网络带宽测量方法

李智涛<sup>①</sup> 徐雅静<sup>①</sup> 刘利宏<sup>②</sup> 徐惠民<sup>①</sup>

<sup>①</sup>(北京邮电大学电信工程学院 北京 100876)

<sup>②</sup>(总装备部工程设计研究总院总体室 北京 100028)

**摘要:** 该文提出一种新的应用于 IPv6 网络的带宽测量方法 PTTS(Packet Train Time Stamp)。源端主动向网络发送报文序列(M<sup>h</sup>-L-M<sup>t</sup> packets train), 序列中负载报文反应网络流量特征, 测试报文带有 IPv6 时间戳扩展报文头, 逐跳记录路由器的当前时间, 获得报文序列通过链路所花费时间, 得到可用带宽。同时利用 IPv6 基本报文头中流标签字段定义测试流, 确保报文中测试报文和负载报文路径一致; 利用流量类型字段, 增设用于网络测量的测试级消除其与背景业务的相互影响。仿真证明, 报文设计合理, 测量方法可行。

**关键词:** IPv6 网络; 可用带宽; 报文序列; IPv6; 时间戳

中图分类号: TP393.06

文献标识码: A

文章编号: 1009-5896(2008)09-2283-04

## An Approach to Available Bandwidth Measurement in IPv6 Networks

Li Zhi-tao<sup>①</sup> Xu Ya-jing<sup>①</sup> Liu Li-hong<sup>②</sup> Xu Hui-min<sup>①</sup>

<sup>①</sup>(Telecommunication Engineering School, Beijing University of Posts and Telecommunication, Beijing 100876, China)

<sup>②</sup>(Beijing Special Engineering, Design Institute, Beijing 100028, China)

**Abstract:** An approach to available bandwidth measurement in IPv6 networks called Packet Train Time Stamp (PTTS) is proposed. A new packet train (M<sup>h</sup>-L-M<sup>t</sup>) is designed in which the measurement packets have the IPv6 time stamp extension header and the load packets reflect the characteristic of the networks. The source sends probe packet train (M<sup>h</sup>-L-M<sup>t</sup>), the current router time is recorded in each measurement packets. Processing these time stamp by the bandwidth statistic algorithm, the value of available bandwidth is acquired. Using the flow label and traffic class fields in the IPv6 header, the test class and stream is designed in order to reduce the influence of cross traffic. This approach is verified through simulation.

**Key words:** IPv6 network; Available bandwidth; Packet train; IPv6; Time stamp

### 1 引言

网络中的传输路由由从数据源到目的地一系列存储转发链路组成。带宽是指网络上数据报文的最大传输速率。瓶颈带宽指源节点到目的节点之间处理能力最低的链路所能达到的最大数据传输速率。然而, 业务流很少能独占一条路径, 经常是和其他业务流共享网络, 此时业务流所能得到的带宽就是通常所说的可用带宽。

近年来人们设计了大量的带宽测试算法和测试系统<sup>[1-7]</sup>。算法从广义上主要有两类, 一是数据包对(Packet pair)算法, 衍生算法有 bprobe 改进算法, TOPP 算法, (改进型)Potential Bandwidth Filtering(PBF)算法, Packet Tailgating(PT)算法, PBM 算法等; 二是 Pathchar 算法, 衍生的有非对称链路算法。

本文提出的测量方法是应用于 IPv6 体系类属于数据包对算法的路由器协作主动测量方法——报文序列时间戳方法(PTTS)。本方法中源端背靠背地向网络发送由测试报文

和负载报文组成的探测报文序列(M<sup>h</sup>-L-M<sup>t</sup>)。利用时间戳机制记录测试报文(M<sup>h</sup> 和 M<sup>t</sup>)通过路由器的时间, 间接获得每个报文序列通过链路的时间, 最终得到链路带宽。在测量过程中利用 IPv6 报头中的流标签和业务流字段, 设计专门的测试流和测试级, 保证探测报文序列路径一致并且不被其它竞争流量分离。本文安排如下, 第 2 节介绍 PTTS 方法的设计思想, 第 3 节提出带宽统计算法, 第 4 节是仿真验证, 最后一节给出结论。

### 2 IPv6 中 PTTS 带宽测量方法设计

#### 2.1 基本概念与数学模型

传送过程中分组  $k$  经过链路  $i$  经历排队时延( $t_{queuei}^k$ ), 传播时延( $t_{propi}^k$ ), 传输时延( $t_{trani}^k$ )和处理时延( $t_{proc}^k$ ) 即

$$T_i^k = t_{queuei}^k + t_{propi}^k + t_{trani}^k + t_{proc}^k \quad (1)$$

$$t^k = \sum_{i=1}^n (t_{queuei}^k + t_{propi}^k + t_{trani}^k) \quad (2)$$

其中  $T_i^k$  表示分组  $k$  经过链路  $i$  的时间,  $t^k$  是分组  $k$  经过整个路径的时间。当测试分组序列的发送间隔满足一定要求, 可以求出链路可用带宽  $B$ 。

$$B = \frac{LEN}{\Delta T} \quad (3)$$

2007-03-05 收到, 2007-07-31 改回

国家 863 计划项目(2006AA01Z235)和国家自然科学基金 (90604019) 资助课题

其中 LEN 表示探测分组序列长度,  $\Delta T$  表示序列在链路的分离时间。

包对法测量就是通过测量报文的分离时间获得可用带宽, 在测量过程中做如下假设:

(1) 路由器对于到达的数据包采用先进先出(FIFO)的方式调度、转发;

(2) 探测序列中的所有分组路径一致。

本文提出的方法类属于包对法测量, 设计了全新的测试报文序列, 增设了测试级和测试流解决包对法固有的测量假设问题。同时采用离散化滤波方法对测量的样本空间进行处理, 获得链路的可用带宽。

## 2.2 测试报文序列的设计

为了更好地反映网络真实情况, PTTS 方法不发送等长探测报文, 而是发送探测报文序列。其结构如图 1 所示。序列由测试报文( $M^H, M^T$ )和负载报文(Load)两部分组成。其中负载报文用于反映网络的流量特性<sup>[8]</sup>。文献[8]介绍 IP 网络中, 40byte 的数据包约占 50%, 552 或者 576byte 的数据包占 20%, 1500byte 的数据包占 10%。PTTS 方法中负载报文符合上述分布情况。测试报文采用带有时间戳扩展头的 TCP/UDP IPv6 报文, 每跳路由器将当前时间记录在扩展报文中, 目的端负责提取报文的时间戳序列。通过时间戳序列获得探测报文经过每跳链路的分离时间, 对其做统计处理得到路径可用带宽。那么如何获得报文每一跳的时间戳呢? RFC781 中定义了 IPv4 的时间戳选项, 目前对于 IPv6 无相关定义, 本文设计一种新的 IPv6 时间戳扩展报文来记录报文通过路由器的时间。

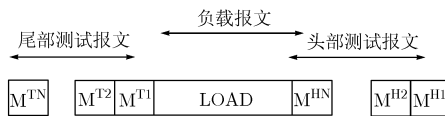


图 1 测试报文序列的结构

IPv6 扩展头通过 IPv6 头中“下一报头”字段标识。根据国际组织 IANA 所给出的“Protocol Number”列表, 目前未指派的编号是 138-252。对于时间戳扩展头, 使用 138 来标识。根据 RFC2460<sup>[9]</sup>规定的格式, IPv6 时间戳选项扩展头格式如图 2。其中: 下一报头(next header): 占用 1byte。所有 IPv6 扩展头均包含此字段, 用于标识下一头类型。

扩展头长度(Hdr Ext Len): 占用 1byte, 用以标识时间戳选项扩展头的长度, 该长度以 8byte 为单位, 不包含扩展头第一个 8byte, 如果扩展头只有 8byte 长, 该字段为 0。此

Next header(8)	Edr ext len(8)	Type(8)	Data pointer(8)	PADS(32)
Data(64)				

图 2 IPv6 时间戳选项报文格式

字段限制了扩展头最多为 2048byte。

类型(type): 占用 1byte, 标识时间戳选项头的类型。根据 RFC2460 建议, Type 字段的高两位值为 00, 表示“若 IPv6 节点不支持该选项类型, 则跳过本扩展头处理下一个扩展头”, 本文以 4 表示“逐跳路由器时间戳”, 故 Type 字段为 00000100。

数据指针(data pointer): 占用 1byte, 标识当前已经记录的时间戳数量, 同时表明下一个时间戳存放在扩展头中的位置。Data pointer 初始值是 0, 当带有时间戳扩展头的 IPv6 报文经过路由器时, 路由器在“data”字段写入当前时间, 同时将 Data pointer 值加 1, 除非“data point 值已经是 255”超过可记载最大时间戳数量。路由器记录时间戳的位置计算方法如下: 该位置的真实地址是从时间戳扩展报文的首部算起第 $(8 \times \text{data pointer} + 1)$ 个字节。

填充(PADS): 占用 4byte, 仅用于填充, 保证时间戳报文头在去除“data”字段后长度是 8byte 的整数倍。

数据(data): 扩展头中数据, 记录报文通过路由器的具体时间(时间戳)。每个时间戳占用 8byte。时间戳是从 UTC 午夜开始到当前时间所经过的微秒数。受“Hdr Ext Len”限制, 最多记 255 个时间戳。

## 2.3 IPv6 中测试流和测试级的设计

多数包对法测量方法中, 都不能很好地解决 2.1 节中假设条件第(2)点。原因在于: (1)测试报文和背景流量具有相同优先级, 当测试报文序列被首尾相连地发送时, 并不能像事先设想的那样被分开, 即第一个和最后一个报文到达目的地的时间间隔并不能很好地反应带宽的情况; (2)所有测试报文转发过程中选择的路径并非一致。

本文利用 IPv6 报头中的流标签和流量类型字段来解决此问题。IPv6 包头中定义了 20bit 的流标签字段和 8bit 的流量类型字段<sup>[9]</sup>。流标签字段用于标识 IPv6 路由器需要特殊处理的包序列。PTTS 方法将探测报文序列中所有报文的流标签字段定义为统一值, 这样保证序列中所有报文在转发过程中选择固定且唯一的路径。为了解决报文相互影响问题, 利用 IPv6 报头中流量类型字段, 增加特殊的测试流级别(简称测试级)。测试级被定义为 IPv6 报文中的最低优先级, 源节点和转发路由器只有在没有其他数据进行处理时才会处理测试报文。这样测试报文不会对网上已有的业务产生影响。因此, 测试时可以很小的间隔发送大量测试报文序列得到测试结果, 提高测试效率。

## 3 带宽统计算法

探测报文序列中第  $k$  个头部测试报文  $P_h^k$  获得的时间戳序列为  $T^{Hk} = \{t_1^{Hk}, t_2^{Hk}, \dots, t_M^{Hk}\}$ , 第  $k$  个尾部测试报文  $P_t^k$  的时间戳序列为  $T^{Tk} = \{t_1^{Tk}, t_2^{Tk}, \dots, t_M^{Tk}\}$ , 其中  $M$  为被测路径的链路数。

$$G_M^k = T^{Tk} - T^{Hk} \quad (4)$$

$G_M^k$  是长度  $L_{total} = Len_{load} + N \times Len_{meas}$  报文经过测试路径中每跳所花费的时间序列(由时间戳增加的报文长度忽略不计),  $N$  表示头部测试报文数目, 亦即尾部测试报文数目。

$$b_M^k = \frac{L_{total}}{G_M^k} = \frac{len_{load} + N \times Len_{meas}}{G_M^k} \quad (5)$$

$b_M^k$  为第  $k$  个头部与尾部测试报文探测得到的通路中每跳链路的可用带宽序列。

在  $W$  时间内, 以频率  $f$  对链路进行探测, 得到每跳链路带宽的样本数为  $n = WfN$ , 对  $n$  个样本进行滤波处理, 最终得到链路的可用带宽。本方法中采用离散化低通滤波器, 滤波公式为

$$B_k = \alpha_k B_{k-1} + (1 + \alpha_k) \frac{b_k + b_{k-1}}{2} \quad (6)$$

其中  $\alpha_k = (2\tau - \Delta k) / (2\tau + \Delta k)$ ;  $\tau$  为滤波器的取样周期;  $\Delta k$  为时间差;  $B_{k-1}$  为  $k-1$  时刻滤波处理后得到的带宽统计值,  $b_k, b_{k-1}$  分别为  $k, k-1$  时刻带宽样本值。离散化低通滤波器能有效地将突发的干扰流量滤除, 利用相邻测量结果的相关性, 收敛速度快, 计算量较小。能够很好地满足 PTTS 方法对测量实时性、算法运算量和收敛速度等各方面要求。

### 4 仿真及分析

仿真配置环境如图 3(a), 3(b)所示, 图 3(a)中背景流量为单跳背景流。X, Y, Z 链路容量分别是 80Mbps, 50Mbps, 80Mbps; C1, C2, C3 代表背景流量, 分别流经链路 X, Y 和 Z。图 3(b)中背景流量为路径背景流(非单跳)。本文首先考察在链路利用率为 20%, 80%, 50%时包对法的测量结果。

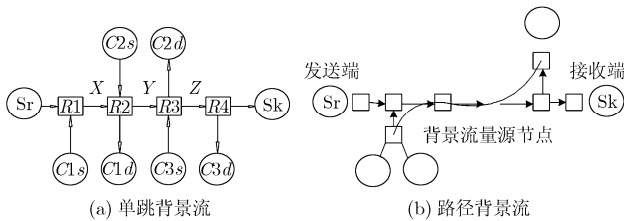


图3 仿真配置图

对于本文提出的方法, 从源端向目的端发送带有 IPv6 时间戳选项的 ICMPv6 Echo Request 报文, 目的端接收此报文后, 生成 ICMP Echo(请求回显)报文, 并将原报文中的 IPv6 扩展报文头中数据复制到回显报文中, 发回源端。发送间隔为 10ms, 背景流量为 pareto 分布( $\alpha = 1.9$ ), 考察在不同可用带宽下的测量结果。

图 4 为包对列法( $N=3$ )在图 3(a)配置情况下 3 种链路利用率下 1500 次试验中结果分布情况。图中横坐标表示有效带宽, 纵坐标表示 1500 次试验中每种带宽值的统计数量。可以看出包对列法测量结果逐渐收敛, 集中在文献[1]中提到的 ADR(Asymptotic Dispersion Rate)附近。但是, 无论链路的负载情况如何, 包对列法无法测出可用带宽。

图 5 给出了实际可用带宽值在 0M,5M,10M,15M,20M, 25M,30M,35M,40M,45M,50M 的测量结果。可以看出 PTTS 方法在两种网络模型均能够测量出链路可用带宽的真实值。

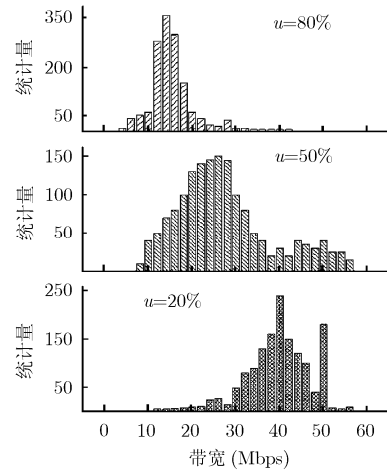


图4 包对列( $N=3$ )仿真结果

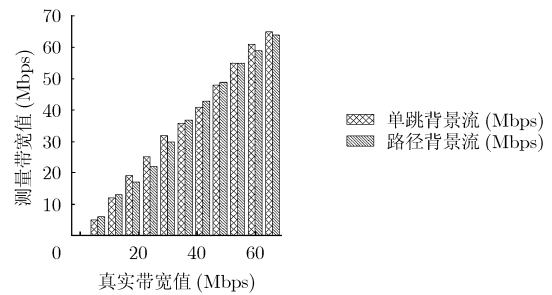


图5 PTTS 方法的测量结果

### 5 结束语

本文提出一种适用于 IPv6 网络的可用带宽测量方法 PTTS。该方法需要路由器支持 IPv6 时间戳扩展报文。此方法克服了已有带宽测量方法所固有的测量时间长, 对网络中已有业务流影响大的缺点。因此可以适用于无线网络等低带宽的网络环境中。仿真验证探测报文序列设计合理, 测量方法可行。

在未来的工作中将考虑能否将这种思想推广到支持区分服务的网络环境中, 以解决更多测量问题。

### 参考文献

- [1] Lai K and Baker M. Nettimer: A tool for measuring bottleneck link bandwidth. In proceedings of USITS'01, San Francisco, 2001: 122-133.
- [2] Lai K and Baker M. Measuring link bandwidths using a deterministic model of packet delay. In proceedings of ACM SIGCOMM, Stockholm, 2000: 283-294.
- [3] Mah B A. Pchar: A tool for measuring internet Path characteristics. <http://www.employees.org/bmah/software/pchar/>, June 2000.

- [4] Carter R L and Crovella M E. Measuring bottleneck link speed in packet switched networks. *Performance Evaluation*, 1996, 27 & 28: 297-318.
- [5] Saroiu S, Gummadi P K, and Gribble S D. SProbe : A fast technique for measuring bottleneck. In proceedings of IEEE INFOCOM, San Francisco, 2002: <http://sprobe.cs.washington.edu/sprobe.ps>.
- [6] Dovrolis C, Ramanathan P, and Moore D. What Do Packet Dispersion Techniques Measure? In proceedings of IEEE INFOCOM, Anchorage, 2001: 905-914.
- [7] Paxson V. End-to-end Internet packet dynamics. *IEEE/ACM Transaction on Networking*, 1999, 7(3): 277-292.
- [8] Thompson K, Miller J G, and Wilder R. Wide area Internet traffic patterns and characteristics. *IEEE Network*, 1997, 11(6): 10223.
- [9] RFC 2460, Internet Protocol, Version 6 (IPv6) Specification, IETF.
- [10] Melander B, Bjorkman M, and Gunningberg P. A new end-to-end probing and analysis method for estimating bandwidth bottlenecks. In the proceedings of IEEE GLOBECOM, San Francisco, 2000: 415-420.
- 李智涛: 女, 1978 年生, 博士生, 研究方向为 IP 网络测量和行为测量及 QoS 管理.
- 徐雅静: 女, 1977 年生, 讲师, 研究方向为无线宽带接入技术.
- 刘利宏: 男, 1978 年生, 工程师, 研究方向为数值计算与仿真.
- 徐惠民: 男, 1941 年生, 教授, 博士生导师, 研究方向为宽带网络通信.