

一种适用于 IPv6 的高效瓶颈带宽测量方法

李 雯¹, 潘 乔¹, 朱畅华¹, 裴昌幸²

(1. 西安电子科技大学 ISN 国家重点实验室, 西安 710071; 2. 西安电子科技大学计算机网络与信息安全教育部重点实验室, 西安 710071)

摘要: 提出了一种下一代网络的瓶颈带宽测量方法——IPv6-pckt-pr。该方法在 OPNET 仿真环境下, 以自相似业务流为背景流量, 给探测报文赋予相同的流标签, 同时将其业务类别设置为最高级, 通过发送不等长的探测包对测量端到端瓶颈带宽。分析与实验表明, 该方法的测量结果准确度较好, 分布比较集中, 测量时间较短, 并能有效减少测量带宽。

关键词: IPv6; 瓶颈带宽; 带宽测量; OPNET

Efficient Bottleneck Bandwidth Measurement Method in IPv6 Networks

LI Wen¹, PAN Qiao¹, ZHU Chang-hua¹, PEI Chang-xing²

(1. State Key Lab of Integrated Services Networks, Xidian Univ., Xi'an 710071;

2. Key Lab. of Computer Networks and Information Security, Ministry of Edu., Xidian Univ., Xi'an 710071)

【Abstract】 An end-to-end measurement method, called IPv6-pckt-pr, is proposed to estimate the bottleneck bandwidth in the next generation Internet. This method gives same flow labels to the probing packet pairs, sets the traffic class of the probing packets to the highest level, estimates the end-to-end bottleneck bandwidth by sending packet pairs of different size. The explicitly generated self-similarity cross traffic in OPNET simulation allows the research to be more efficient. Theoretical analysis and simulations show this method is accurate, with less probing time and low bandwidth expense.

【Key words】 IPv6; bottleneck bandwidth; bandwidth measurement; OPNET

IPv6 具有长度为 128 位的地址空间, 可以彻底解决地址不足的问题。同时 IPv6 在简化报头, 对移动 IP 技术的支持, 提高安全性、服务质量等方面有较大的改进。IPv6 网络带宽测量是更加有效地利用网络带宽, 进行网络控制, 实施 QoS, 进一步提高网络性能的关键。目前, 使用 IPv6 协议网络的瓶颈带宽测量方法尚处于初始研究阶段, 针对 IPv6 特性测量的研究成果相对较少。因此, 本文在研究下一代网络 IPv6 协议的基础上, 提出了一种基于包对法的高效瓶颈带宽测量方法——IPv6-pckt-pr。

1 瓶颈带宽测量技术的现状

端到端瓶颈带宽是指一条端到端链路 $P_H = \{A_0, A_1, \dots, A_H\}$ 上, 所有 H 条单跳链路中最小的链路带宽:

$$C_{bw} = \min_{i=1,2,\dots,H} C_i \quad (1)$$

其中, C_i 为连接节点 A_{i-1} 和节点 A_i 的链路 L_i 的瓶颈带宽(或容量)。可以看出, 端到端瓶颈带宽与网络中背景流量的实时变化无关。

目前已有许多单跳瓶颈带宽测量工具用来测量 IPv4 网络中的瓶颈带宽, 如 Pathchar、Clink 和 Pchar 等。对于端到端的瓶颈带宽的测量, 也有很多测量工具, 如 Bprobe、Nettimer、sprobe 和 Pathrate 等^[1]。而端到端瓶颈带宽测量技术大致分为 3 种: 单个分组技术, 分组对技术以及多分组技术。

单个分组技术通过发送不同长度的分组, 测量端到端路径上各节点与源节点之间的往返时延 RTT, 计算出分组在每一条相邻节点组成的链路上的时延, 也可以叫做可变包长探

测技术。传输时延与分组长度成正比, 与链路的带宽成反比, 找出最小的链路带宽值, 即为所求。则瓶颈带宽

$C_{BW} = \min_{i=0,\dots,H} \frac{L}{RTT/2}$, 其中, L 是探测报文长度。这种技术存

在如下问题: 采用 RTT 时, 若前向链路和反向链路不对称, 就会增大测量误差; 逐链路进行时延测试, 要发送大量分组, 给网络过多测量负载。

分组对技术假设端到端链路上没有背景流量, $\Delta t_m = L/C_0$ 为背靠背报文对进入网络时的时间间隔(其中, L 为探测报文长度; C_0 为源端分组对的发送速率); 此报文对经过所有路径(没有背景流量)之后, 它们之间的时间间隔就变为

$$\Delta t_R = \max_{i=0,\dots,H} \left(\frac{L}{C_i} \right) = \frac{L}{\min_{i=0,\dots,H} (C_i)} = \frac{L}{C} \quad (2)$$

由 $C = \frac{L}{\Delta t_R}$ 得到的带宽值, 就是端到端路径上最小的链路

带宽 $C = \min_{i=0,\dots,H} C_i$, 即端到端链路上的瓶颈带宽。然而实际并不满足分组对技术的前提条件, 因为实际网络中总是存在背

基金项目: 国家自然科学基金资助项目(60572147, 60132030); 西安市科技攻关资助项目(GG04018)

作者简介: 李 雯(1982 -), 女, 硕士, 主研方向: 下一代网络测量技术; 潘 乔, 博士研究生; 朱畅华, 副教授、博士; 裴昌幸, 教授、博士生导师

收稿日期: 2007-01-04 **E-mail:** vivianlee0324@163.com

景流量, 往往会在两个探测包之间随机地加入背景流量报文, 而使测量结果不可避免地总是掺杂噪声。背景噪声会导致测量结果偏离正常值。因此分组对技术中的关键问题是寻找合适的滤波方法去除噪声。而由于背景流量的引入是随机的, 要消去背景流量对于瓶颈带宽的影响, 目前还没有特别行之有效的方法。

多分组技术采用 3 个或 3 个以上的分组技术, 对带宽进行测量。发送分组长度可以相同, 也可以不同, 第 k 个分组长度为 L^k , 则瓶颈带宽 $C_{BW} = \frac{\sum_{i=0}^{N-1} L^i}{t^N - t^1}$ 。背景流量的存在会直接导致由这个公式计算得到的瓶颈带宽误差加大。

2 IPv6-pckt-pr 原理与实现

2.1 IPv6-pckt-pr 测量原理

目前 IPv4 网络下的各种带宽测量方法都假定在测量过程中, 测量路径固定且惟一, 即不改变路由。而实际的 IPv4 网络为让每个报文单独寻径以达到其目的地, 即每个报文都分别进行处理, 其结果是两个从相同源发往相同目的地的报文可以采用不同的路由。所以在当前实际的 IPv4 网络中带宽测量方法的前提假设很难真正成立。在以下一代网络 IPv6 为基础的研究中, 可以发现 IPv6 中具有相同流标签的 IPv6 报文的路由是相同的, 利用 IPv6 的这个特性, 使得各种带宽测量方法所共有的前提假设在 IPv6 网络中完全成立。本文利用 IPv6 报头的流标签避免了带宽测量方法所固有的路径不固定并且不唯一的问题。同时, 具有相同路由的报文其底层协议对于报文对时间间隔的影响是常量, 可以通过发送两对探测报文的方法, 消除底层协议对测量结果的影响。此外, 利用 IPv6 的业务流类别将探测报文设置为最高级别, 探测路径中的所有路由器会对探测报文进行优先处理, 探测数据处理完毕后再进行其他报文的处理。这样就在很大程度上减少了探测报文对之间的背景流量的干扰。

一条端到端链路 $P_H = \{A_0, A_1, \dots, A_H\}$ 由 H 条链路组成, C_i 为连接节点 A_{i-1} 和节点 A_i 的链路 l_i 的带宽。报文对经过这条链路后, 二者之间的时间差是 $\Delta t_R = \frac{L}{C}$, 考虑到链路层、物理层协议的影响以及背景流量的插入, 可得

$$\Delta t_R = \frac{L + d + \psi}{C} \quad (3)$$

其中, C 是端到端链路的瓶颈带宽; L 是探测报文长度; d 是测量包经过数据链路层和物理层后加上的额外字节; ψ 是测量报文对间插入的背景流量。

目前的瓶颈带宽测量方法中, 由于实际网络中总是存在背景流量, 往往在两个探测报文之间会随机地加入背景流量报文, 因此测量结果不可避免地掺杂噪声。如果测量时只有第 1 个数据分组被背景流量分组所延迟, 二者的间隔将减小, 导致时间压缩(time compression); 如果只有第 2 个数据分组被背景流量所延迟, 二者的间隔将加大, 导致时间扩展(time extended)。

IPv6 中定义的流概念可以解决实际网络背景流量的干扰问题。IPv6 的包头(图 1)中包括了 8 位的业务流类别(traffic class)字段, 它也可被路由器用来识别和区分不同类型或具有不同优先级的 IPv6 报文。它将业务量分为受拥塞控制的业务量和不受拥塞控制的业务量。本文利用业务流类别, 将探测报文设置为最高级。当报文进入转发路由器的时候, 即使有

其他等级的报文插入排队序列, 路由器也会对探测报文进行优先处理, 探测数据处理完毕再进行其他报文的处理, 以达到瓶颈带宽测量中假设的链路中没有任何背景流量的效果。这样在测量过程中就不需要再考虑背景流量对测量结果的影响。即 ψ 可以从式(3)中消去, 此时:

$$\Delta t_R = \frac{L + d}{C} \quad (4)$$

版本号 (4位)	业务流类别(8位)	流标签(20位)	
数据长度(16位)		下一包头(8位)	跳数限制(8位)
源地址(128位)			
目的地址(128位)			
扩展包头(不定)			

图 1 IPv6 的包头格式

此外, IPv6 中定义的流是指从某个源点向信宿发送的分组群中, 源点要求中间路由器作特殊处理的那些报文。IPv6 网络按照流来对数据进行分别处理, 属于同一个流的报文的路由一定相同。那么将探测报文放在一个具有相同源地址和目的地址的数据流中, 这些报文就具有相同的流标签(flow label), 也就具有相同的路由。对于具有相同路径的探测报文对, 底层协议的影响 d 就是一个常量。因此, 可以利用相减的关系, 去掉 d 对测量结果的影响: 假设第 1 个报文对中每个分组长度为 L^1 , 第 2 个报文对中每个分组长度为 L^2 , 为了使第 2 个报文对中的第 1 个分组不因第 1 个报文对产生排队, 令 $L^2 > L^1$ 。两对报文对到达目的端的时间间隔分别为 $\Delta t_R^1 = L^1 + d/C$ 和 $\Delta t_R^2 = L^2 + d/C$, 两式相减有: $\Delta t_R^2 - \Delta t_R^1 = (L^2 - L^1 + d - d)/C = (L^2 - L^1)/C$, 从而推导出瓶颈带宽测量公式:

$$C = \frac{L^2 - L^1}{\Delta t_R^2 - \Delta t_R^1} \quad (5)$$

式(5)排除了底层额外字节和背景流量的影响。

2.2 仿真测试环境

本文采用 OPNET 对 IPv6-Pckt-Pair 和现有的瓶颈带宽测量方法分别进行了仿真测试, 采取了最具有代表性的两种网络拓扑模型^[2], 如图 2、图 3 所示, 使用自相似业务流作为背景流量。图 2 中的背景流量总是在进入端到端路径的一跳后离开(one-hop persistent cross traffic), 图 3 中的背景流量和测试报文的传输路径相同(path persistent cross traffic)。通过对这两种拓扑模型的仿真来逼近实际的网络环境。仿真中, 测试路径各链路带宽为 $P_H = \{100, 80, 35, 20, 55, 90\}$ (单位: Mb/s), 瓶颈带宽为 20Mb/s, 使用 5 个路由器来建立连接。

网络中的数据流统计特性不再是泊松过程可以描绘的^[3], 而是具有长时间相关性(long range dependence)和慢衰减方差特性——自相似性的。准确的业务建模可以获得对实际网络设计有指导意义的结果。为了使得仿真业务源能够反映实际业务的统计特征, 本文选用了自相似业务流^[4]来仿真背景流量。仿真中, 每条背景流量由 4 个 Pareto 源产生, 图 2 的背景流量由 4×4 个 Pareto 源产生。每个 Pareto 源都采用 Sup-FRP(superposition of the fractal renewal point processes)^[5], 取 Hurst 参数 $H = 0.7$, 则 $\alpha = 2 \times H - 1 = 0.4$, 背景流量包长取 1500B。Pareto 源的平均到达率(average arrival rate)由 $AAR = \frac{C_i \times u}{4 \times L_{CR}}$ 分别计算得到。其中, C_i 为第 i 条链路

的带宽； L_{CR} 为背景流量包长； u 为每段链路的链路利用率。为计算方便，假设测量路径中的各条链路利用率 u 相同。

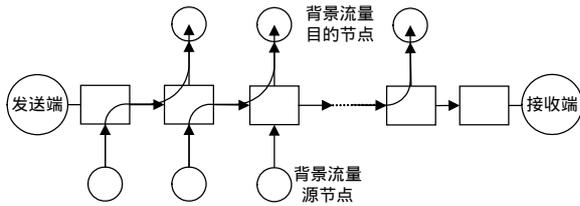


图2 One-hop persistent 背景流量

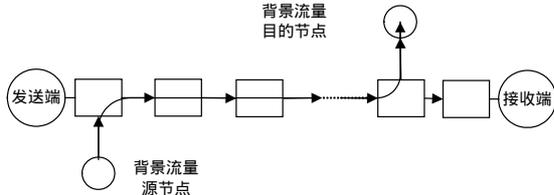


图3 Path persistent 背景流量

3 仿真结果分析

将测量链路上的所有节点都设置为 IPv6 节点，给发送端的探测报文赋予相同的流标签和最高级的业务流类别。每次测量进行 60 次发包，每次发送两对背靠背报文对（即两对报文对之间的时间间隔为 0s），第一对报文的测量分组长度 L^1 取 600B，第二对报文分组 L^2 取 1 500B。在 60 个结果中剔除方差较大的测量值，对方差在可接受范围内的数据进行统计平均，作为此次测量的结果。

仿真中背景流量采取图 2 中的 One-hop persistent cross traffic。在链路利用率 u 分别为 20%，50% 和 80% 的情况下，将 IPv6-pckt-pr 和 Pathrate 在 IPv6 下的实现进行了对比。从表 1 可以看出，IPv6-pckt-pr 和经典端到端瓶颈带宽测量工具 Pathrate（同样是包对法）相比，具有更高的准确度。Pathrate 在网络处于轻负载时（ $u = 20\%$ ），可以测出瓶颈带宽，但是当网络处于重负载时（ $u \geq 50\%$ ），此方法就会产生大量错误结论。与此同时，IPv6-pact-pr 可以得到更加准确的测量结果。

表 1 本文方法和 Pathrate 对比

瓶颈带宽 (Mb·s ⁻¹)	链路利用 率 u (%)	瓶颈带宽测量值/(Mb·s ⁻¹)	
		IPv6-pckt-pr	Pathrate
20	20	19.4	19.1
20	50	18.9	18.6
20	80	18.2	16.3

在相同的网络链路带宽情况下对两种方法 10 次测量得到的瓶颈带宽测量值范围进行对比。分别将经过筛选后留下的测量值的上下限在图 4、图 5 中进行标注。对比图 4、图 5 可知，本方法的结果更加接近真实值，并且 Pathrate 方法测量结果的分布范围较大，而本方法的测量结果集中在较小的区间里面。

现有的端到端瓶颈带宽测量方法中，测试报文对网络上已有的业务流量的影响较大，再者这些方法的测量时间长，并不能及时地反映当前网络的带宽情况。IPv6-pckt-pr 所产生的探测业务量较小，而且测试时间很短。虽然将探测报文设置为最高级，会对网络本身的业务量有影响，但是由于持续时间很短，因此这种影响也可以忽略。

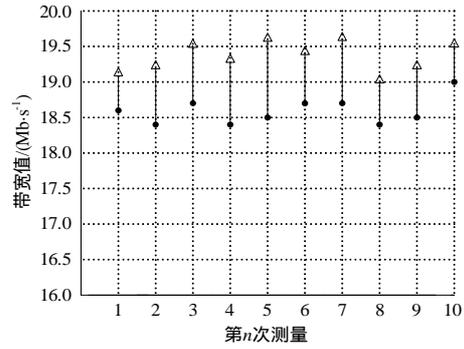


图4 Pathrate 测量结果范围分布

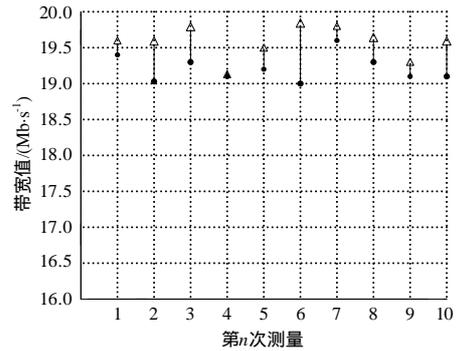


图5 IPv6-pckt-pr 测量结果范围分布

4 结论

本文提出了下一代网络的端到端瓶颈带宽的测量方法——IPv6-pckt-pr。此方法利用 IPv6 包头中的流标签，满足带宽测量中关于测量路由一致的假设；将探测报文对业务类别设置为最高级，以达到探测报文对之间不插入背景流量的效果。并在分析背景流量统计特性的基础上，用 Pareto 源仿真背景流量更好地逼近网络实际状况。理论分析与实验表明，此方法测量结果更加准确，测量结果的分布比较集中，测量时间较短，对网络已有的业务量影响较小。在以后的工作中，将这种思想进一步推广，以解决下一代网络中 Qos 测量及管理等问题。

参考文献

- 1 Prasad R, Dovrolis C, Murray M, et al. Bandwidth Estimation: Metrics, Measurement Techniques, and Tools[J]. IEEE Network, 2003, 17(6): 27-35.
- 2 Jain M, Dovrolis C. End-to-end Available Bandwidth: Measurement Methodology, Dynamics, and Relation with TCP throughput[J]. IEEE/ACM Trans. on Networking, 2003, 11(4): 537-549.
- 3 Paxson V, Floyd S. Wide-area Traffic: The Failure of Poisson Modeling[J]. IEEE/ACM Transactions on Networking, 1995, 3(3): 226-244.
- 4 Addie R G, Zukerman M, Neame T. Fractal Traffic: Measurements, Modeling and Performance Evaluation[C]//Proc. of IEEE INFOCOM'95. 1995.
- 5 Ryu B, Lowen S. Fractal Traffic Models for Internet Simulation[C]//Proc. of IEEE Int'l Symposium on Computer Communication. 2000.