SD-WAN传输限速策略分析

Analysis of Speed Limiting Strategy for SD-WAN Transmission

赵纯熙[1]，童　博[1]，刘锦波[2]（1. 中讯邮电咨询设计院有限公司，北京 100048；2. 中国联合网络通信集团有限公司，北京 100033）

Zhao Chunxi[1]，Tong Bo[1]，Liu Jinbo[2]（1. China Information Technology Designing & Consulting Institute Co.，Ltd.，Beijing 100048，China；2. China United Network Communications Group Co.，Ltd.，Beijing 100033，China）

摘　要：

随着工业互联网发展，将SDN技术应用在广域网中的SD-WAN技术进入高速发展阶段。对某种SD-WAN跨域组网模型进行分段测试，分别对比SD-WAN终端设备在流量监管及流量整形2种限速方式下的广域网传输速率，结合TCP拥塞机制进行分析，并尝试加入TCP BBR技术，在确保链路安全可靠的情况下，提出一种提升SD-WAN跨域组网传输性能的方法。

关键词：

SD-WAN；流量整形；流量监管；TCP CUBIC；TCP BBR

doi：10.12045/j.issn.1007-3043.2021.10.016

文章编号：1007-3043（2021）10-0078-05

中图分类号：TN919

文献标识码：A

开放科学（资源服务）标识码（OSID）：

Abstract：

With the development of industrial Internet，SD-WAN technology which applies SDN technology in wide area network has entered the stage of rapid development. A segmented test is carried out for a certain SD-WAN cross-domain networking model，and the WAN transmission rates of SD-WAN terminal devices under two speed limiting modes of policing and shaping are respectively compared. Combined with the analysis of TCP congestion mechanism，and adding the TCP BBR technology，and under the condition of ensuring the safety and reliability of links，a method to improve the performance of SD-WAN cross-domain networking is proposed.

Keywords：

SD-WAN；Traffic shaping；Traffic policing；TCP CUBIC；TCP BBR

引用格式：赵纯熙，童博，刘锦波. SD-WAN传输限速策略分析[J]. 邮电设计技术，2021（10）：78-82.

1 概述

在工业互联网场景下，网络基础设施的建设需要通盘考虑异构组网和敏捷运维的部署需求，灵活、低成本的SD-WAN广域网组网方式受到越来越多的关注。根据IDC近期发布的调研数据显示，2018年度全球SD-WAN市场收入达到413.7亿美元，同比增长了64.9%。IDC进一步预测2019—2024年，SD-WAN的市场规模将突破500亿美元。SD-WAN广域网技术进一步完成了网络控制与数据转发的双平面分离，实现了网络的虚拟化和网络软件自动化，有效地解决了企业网络资源灵活配置的复杂性问题。SD-WAN也成为可以有效地帮助互联网企业快速构建高性价比、简易运维、即需即用的智能企业网络服务专线。

本文以一种SD-WAN跨域组网模型为实例，通过试验对比流量监管与流量整形、TCP缺省拥塞控制与TCP BBR技术在SD-WAN跨域组网中对数据传输速率的影响，提出一种限速策略与TCP传输技术相结合的优化应用方法，有效地提升了SD-WAN跨域组网的带宽利用效率，具有较大的推广价值。

收稿日期：2021-09-12

2 SD-WAN跨域组网模型

SD-WAN跨域组网模型中客户侧路由及数据由终端盒子接入传输至SD-WAN接入侧网关、骨干网侧

网关后通过运营商骨干网到达对端骨干网PE设备。

a）终端盒子：客户侧SD-WAN终端盒子为客户提供SD-WAN组网的最后一公里接入，SD-WAN客户终端设备支持通过MPLS、互联网、4G以及5G多种网络接入方式。终端盒子通过互联网与SD-WAN接入侧网关建立IPSEC隧道完成网络流量的加密及转发，同时SD-WAN终端设备支持服务质量（QoS）以及多种路由协议配置。

b）SD-WAN接入段网关：向下与客户SD-WAN终端建立overlay的3层IPSec隧道形成数据平面，通过BGP完成路由转发。向上与运营商骨干网网关通过BGP建立邻居关系，通过VRF划分的方式实现基于VPN的路由隔离。将用户本地网络连接到骨干网边缘网关，向骨干网边缘网关发布本地路由并学习远端站点路由。

3 测试方法及基线测试

采用打流工具Iperf进行测试，保持相同的Iperf测试参数，在不同的配置策略下（包括无限速策略），分别测试PC至接入段网关后VM和PC至骨干段网关后VM之间的性能带宽。Iperf测试参数如下。

a）客户端不限带宽测试如下。

（a）客户端单线程上行：iperf3 -c 172.160.30.10 -M 1400。

（b）客户端单线程下行：iperf3 -c 172.160.30.10 -M 1400 -R。

b）客户端指定20 Mbit/s带宽测速如下。

（a）客户端单线程上行：iperf3-c 172.160.30.10 -M 1400 -b 20M。

（b）客户端单线程下行：iperf3 -c 172.160.30.10 -M 1400 -b 20M -R。

（c）客户端多线程下行：iperf3 -c 172.160.30.10 -M 1400 -b 20M -t 20 -R。

其中，通过-M 1400将TCP MSS设置为1400，避免报文在广域网口被分片，从而使得性能降低。

如图1所示，SD-WAN终端通过5G CPE接入Internet，与北京POP节点接入段网关1建立IPSec隧道。在北京POP节点部署2个虚拟机，分别位于接入段网关和骨干侧网关之后，虚拟机安装IPerf、FTP及文件共享工具。测试得到PC到接入段网关后VM虚机、骨干侧网关后VM虚机的时延分别为30.20 ms、27.86 ms，测试的IP地址分配如下：

测试PC（Ubuntu 18.04，iperf 3.1.3）：192.168.30.236。

前置VM（Redhat 7.5，iperf 3.1.7）：172.160.30.10。

汇聚VM（Redhat 7.5，iperf 3.1.7）：172.160.31.10。

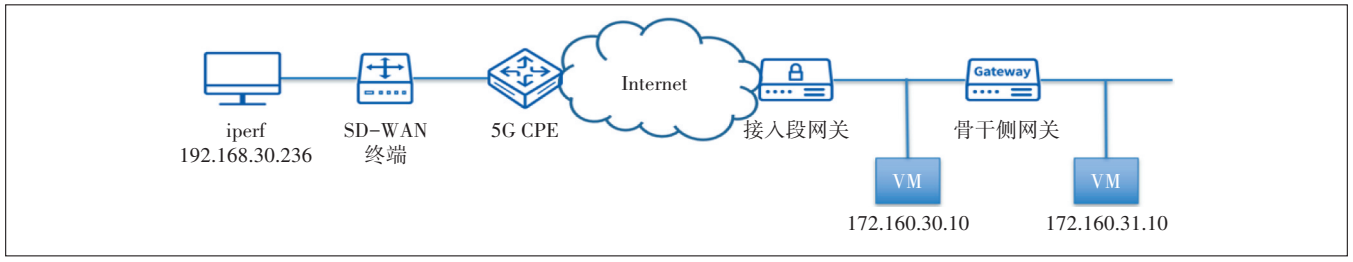

图1　SD-WAN分段测试拓扑

基于上述组网环境，首先进行了SD-WAN终端设备无限速策略（骨干网PE设备保留流量监管限速，20 Mbit/s）下的基础线路测试。采用缺省TCP拥塞控制算法CUBIC，分别从SD-WAN终端（192.168.30.236）到接入段网关后的虚机（172.160.30.10）以及骨干侧网关的虚机（172.160.31.10）进行iperf数据打流测试，得到结果如表1所示。

对比PC到接入网关后虚机VM（结果A）和PC到骨干网关后虚机VM（结果B），网络性能吞吐基本保持一致，即网络流量经过骨干侧网关，其性能吞吐并没有明显降低。理论上，因为转发路径多一跳，时延增大，结果B会小于等于结果A。但实际上，在测试中出现部分情况是结果A大于结果B，或者结果B大于结果A，这多半是属于Internet的带宽和时延得不到保障导致不同时间点测试的偏差。整体测试并没有发现结果B显著低于结果A的情况。因此基本可以得出，在SD-WAN终端设备未作带宽限制的情况下，iperf测

表1　盒子无限速下的传输性能测试

传输速率/（Mbit/s）	平均上传速率	最小上传速率	最大上传速率	平均下载速率	最小下载速率	最大下载速率
接入侧网关	29.40	8.60	49.20	41.00	19.10	59.40
骨干侧网关	31.00	9.89	49.20	42.40	34.80	55.70

试速率基本可达到最大带宽值，网络流量经过接入侧及骨干侧网关，其性能吞吐并没有明显降低。

4 流量监管与流量整形

目前广泛使用的流量控制策略分别是流量监管（traffic policing）与流量整形（traffic shaping），两者都是通过监督流量规格，以达到限制流量及资源使用的目的。

在骨干侧网关启用流量监管的情况下，首先在SD-WAN终端VPN隧道入方向启用流量监管策略。流量监管通过审查数据报文，针对数据报文流量过大的网络连接，采用自动丢弃数据报文或重新设置数据报文优先级的流量监管方式完成数据流量的限速。经过测试发现，在流量监管限制带宽为20 Mbit/s的情况下，采用缺省TCP拥塞控制算法CUBIC，测试得到从SD-WAN终端到接入段网关后虚机和骨干段网关后虚机的网络传输速率过低，均无法达到期望带宽速率。尝试在SD-WAN终端设备VPN隧道出口方向启用流量整形再次测速。流量整形与流量监管的本质区别在于，流量整形面对超过限速的流量报文及突发流量，会将这部分报文先行储存在缓冲区，后续通过令牌桶的控制再匀速的发送出去，在完成流量限速控制的同时，可以更加有效地利用带宽通道。经过测试发现，采用流量整形限制网络传输带宽为20 Mbit/s的情况下，实际网络速率可以达到预期带宽效果，采用缺省TCP拥塞控制算法CUBIC，从SD-WAN终端到接入段网关后虚机的平均上传、下载速率分别为15.4 Mbit/s、20.0 Mbit/s；从SD-WAN终端到骨干侧网关后的虚拟机的平均上传、下载速率分别为15.6 Mbit/s、20.0 Mbit/s，测试结果如表2所示。

表2　2种限速策略对传输质量的影响

传输质量		平均上传/（Mbit/s）	平均下载/（Mbit/s）	上传丢包/个	下载丢包/个
接入侧网关	流量监管	4.29	2.89	33	126
	流量整形	15.40	20.00	3	4
骨干侧网关	流量监管	4.13	3.07	45	65
	流量整形	15.60	20.00	3	0

跑满RTT时延50 ms、20 Mbit/s带宽的链路，需要20 000 000×0.05/8=125 000 B的窗口。如果传输过程丢包率很低，那么TCP的传输窗口可以达到125 000 B，性能就可以提高；如果丢包率较高，TCP传输窗口会回退，传输速率就达不到链路带宽。在SD-WAN跨域组网模型中，普遍使用的SD-WAN终端设备的WAN接口是千兆接口，而实际带宽限速为20 Mbit/s，终端设备无从得知实际链路带宽，会以千兆的速率向外发送报文至接入段网关，触发Policing策略，导致突发流量被丢弃。这些被丢弃的流量会导致TCP CUIC的拥塞控制机制被触发，发送窗口会被按照固定的乘法减小因子β进行降低，而后进入基于一个立方函数的快速恢复阶段。如图2所示，iperf打流结果显示有大量丢包，导致TCP传输速率无法上升，使用Wireshark抓包分别对流量监管与流量整形进行性能分析，可以看出TCP流量出现锯齿现象。

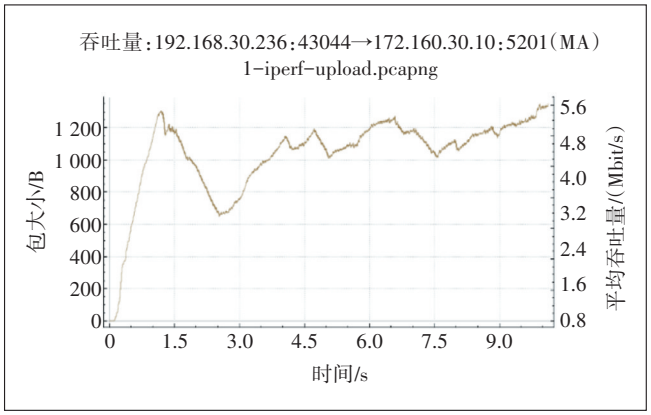

图2　流量监管下的锯齿状流量

在SD-WAN终端设备WAN口使用Shaping限速后，同样的测试，TCP重传的现象减少90%，性能提升2.5倍以上。使用Wireshark抓包进行性能分析，可以看出TCP流量趋于平滑接近实际带宽（见图3）。

通过以上对比不难发现，在SD-WAN的应用场景中，Policing会产生流量锯齿，TCP性能受影响；Shaping则具有典型的去峰填谷作用，使得流量均匀，TCP性能

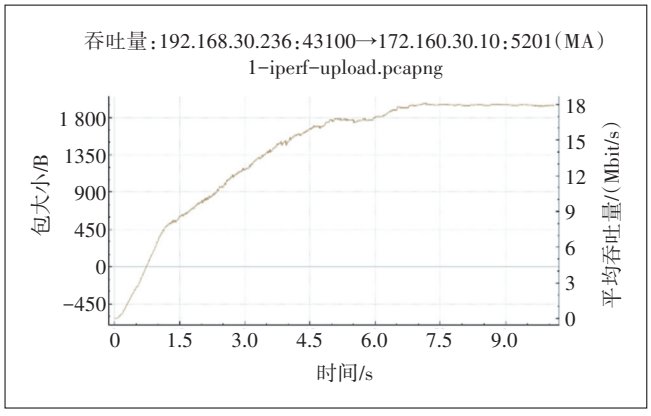

图3　流量整形下“去峰填谷”

接近实际带宽。受限制于试验使用的SD-WAN终端设备，Shaping无法单独作用于VPN隧道，只能应用在物理WAN口上，这就会限制客户互联网访问的上行带宽，但是不影响下行带宽。考虑到互联网访问主要以下行为主，该解决方案可满足大部分应用场景。故建议在SD-WAN终端使用流量整形技术进行限速，该技术可获得接近实际带宽的吞吐性能。

5 缺省TCP拥塞控制与BBR

通过上述流量监管与流量整形的测试对比，可以发现流量监管中出现锯齿状流量及较低网络传输速率的根本原因在于TCP拥塞控制机制。CUBIC通过“加性增，乘性减”的机制，收到的每一个确认消息缓慢地增加发送窗口，而每当确认一个丢包出现的时候就迅速地减小发送窗口。CUBIC发送窗口的增长函数仅取决于连续2次拥塞时间发生的时间间隔，窗口增长不受RTT的影响。但是，CUBIC依然需要基于是否出现丢包来判断网络是否处于拥塞状态，而不去区分造成丢包的原因。可是在本次实验的SD-WAN跨域组网模型中，流量监管的限速策略本身会造成大量突发流量被丢弃，这使得TCP CUBIC“错误地”认为网络出现拥塞影响TCP传输速率。同时，广域网传输中本身具有一定程度上的错误丢包，同样的这些错误丢包也被CUBIC视为发生了网络拥塞。

TCP BBR机制最重要的变化，是将TCP的拥塞控制与可控传输进行了解耦，TCP专注于传输的可靠性，而BBR用来判断可以发送多少、以什么速度发送TCP数据。TCP BBR不再以丢包作为网络拥塞的判断依据，为了尽可能地获得更大的发送窗口，BBR将不断测量最大带宽值与最小延迟值，并将这两者的乘积作为发送窗口的大小。需要注意的是，传输链路的带宽最大值与延迟最小值无法同时测出。测量最大带宽需要用数据包充满整个TCP传输通道，意味着传输缓冲区会有一定量的数据包存在，使链路延迟升高；测量最小延迟则需要排空整个缓冲区的数据包。所以TCP BBR采取交替测量带宽最大值与延迟最小值的方法，以此探测出链路的最大容量，获得最大的发送窗口。这就意味着无论TCP连接是处于默认开始状态、乱序状态还是重传状态，都不会影响BBR的最大发送窗口，这使得在BBR的作用下，输出到网络的流量总是靠近管道容量，这样既能有效保证网络带宽的利用率，同时有效避免了传统TCP的数据缓冲区拥塞膨胀的问题。

在SD-WAN终端盒子启用流量监管策略，采用Linux缺省的TCP拥塞控制CUBIC，可以看到丢包对性能影响较大，采用TCP BBR，虽然重传次数上升30倍以上，但是性能提升了1倍以上。在SD-WAN终端盒子启用流量整形策略，丢包率较高时，TCP BBR比TCP CUBIC表现得更激进，可获得更高吞吐率，因此，在同样的链路质量下，采用TCP BBR更容易测出链路最大容量（见表3）。

表3　TCP BBR对传输性能的影响

传输性能		传输速率/（Mbit/s）	丢包/个
流量监管	CUBIC拥塞控制	4.18	33
	TCP BBR	8.49	1 106
流量整形	CUBIC拥塞控制	15.00	4
	TCP BBR	16.80	15

在SD-WAN终端盒子侧抛弃流量监管而使用出方向的流量整形，可以从根本上避免突发流量导致TCP传输性能受限的问题。同时，TCP BBR算法的引入最大程度优化了数据发送端的发送窗口。在本次实验中通过客户端SD-WAN终端设备部署BBR只能对上传速率起到优化作用，如表4所示，客户端部署的BBR并不能优化下载速率，如果需要优化下载速率，需在服务器端启用TCP BBR。

通过对比上述多组实验数据可以得到以下结论：SD-WAN终端盒子启用流量整形限速策略并配合TCP BBR优化算法可以最大程度上提升网络传输质量，使其更加接近于理论最大带宽。

6 端到端验证及结论

将限速策略改为盒子出方向Shaping，开启BBR TCP算法，PE侧仍然采用流量监管策略，按照图4所示的拓扑进行端到端的文件传输测试。

如图4所示，Ubuntu Server接入SD-WAN终端，经5G CPE连接至Internet，同北京POP节点建立IPSec VPN隧道，接入运营商骨干网；Windows 10 PC位于上海，接入SD-WAN终端，经防火墙接入Internet，同上海POP节点建立IPSec VPN隧道，接入运营商网。两者之间的往返平均时延约为60 ms，分别对SD-WAN终端设备启用流量监管、流量整形及TCP BBR优化，分别进行了FTP文件传输端到端下载、上传测试，测试结果如表5所示。

表4　3种终端限速条件下的数据对比

传输速率/(Mbit/s)			平均上传速率	最小上传速率	最大上传速率	平均下载速率	最小下载速率	最大下载速率
TCP采用缺省拥塞控制	接入侧网关	无限速策略	29.40	8.60	49.20	41.00	19.10	59.40
		流量监管	4.18	2.90	5.43	2.89	2.41	3.03
		流量整形	15.00	8.20	18.60	10连接并发，速率为60.80		
	骨干侧网关PE	无限速策略	31.00	9.89	49.20	42.40	34.80	55.70
		流量监管	4.04	3.14	5.26	3.07	2.76	3.30
		流量整形	15.30	8.90	18.10	10连接并发，速率为67.20		
盒子端PC开启TCP BBR	接入侧网关	无限速策略	19.60	16.10	22.60	33.50	27.50	41.20
		流量监管	8.49	7.77	11.30	3.09	2.67	3.43
		流量整形	16.80	16.10	19.90	10连接并发，速率为71.90		
	骨干侧网关PE	无限速策略	19.10	16.10	22.10	31.70	20.00	41.30
		流量监管	8.54	7.74	11.40	3.05	2.76	3.47
		流量整形	16.60	15.50	18.20	10连接并发，速率为69.30		

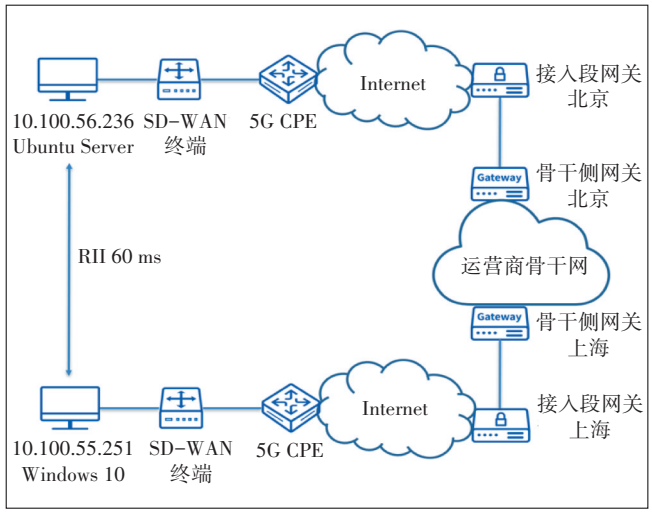

图4　端到端验证拓扑

测试结果验证了修改限速策略并开启BBR优化后SD-WAN广域网组网的TCP吞吐性能有明显改善。在实际的应用场景中，流量监管一般用于普通音频、视频通话等更关注网络时延抖动，对流量丢弃率不敏感的应用。流量整形则用于对传输速度要求较高，但对时延和抖动不敏感的网络应用，如文件传输。在SD-WAN广域网数据传输中，最终将二者配合使用，在最后一公里接入部分的SD-WAN终端盒子进行流量整形并开启BBR优化，在运营商骨干网边缘PE设备上使用流量监管，这样既能最大效率地进行数据传输，又保障了对时延抖动较敏感的网络应用质量。

表5　FTP端到端验证

传输速率/(Mbit/s)	盒子采用流量监管 TCP采用缺省拥塞控制	盒子采用流量整形 开启TCP BBR优化
平均上传速率	0.355	1.800
最小上传速率	0.117	1.600
最大上传速率	0.710	2.100
平均下载速率	0.355	2.000
最小下载速率	0.113	1.600
最大下载速率	0.694	2.300

参考文献：

[1] 侯辛. 分组传送网的QoS实现方式分析[J]. 信息技术，2011(11)：150-153.

[2] 毛奇凰，王岩. TCP拥塞控制算法研究[J]. 洛阳工业高等专科学校学报，2006(3)：4-7.

[3] 吕为. 下一代互联网的QoS分析[J]. 电力系统通信，2005(9)：26-33.

[4] 阮剑飞，林锦贤. 基于三网合一的IP QoS的实现[J]. 福建电脑，2003(4)：4-5.

[5] 李吉良. IP网络服务质量保证技术[J]. 无线电工程，2008(6)：1-4.

[6] 刘芳. 网络流量监测与控制[M]. 北京：北京邮电大学出版社，2009.

[7] 胡云. 对网络流量管理与拥塞管理的研究[J]. 电脑开发与应用，2009(5)：64-66.

[8] 李博杰. BBR算法与之前的TCP拥塞控制相比有什么优势?[EB/OL]. [2021-07-05]. https://www.zhihu.com/question/5355943 3/answer/136002384.

[9] 郑学炜. TCP停止等待、超时重传、滑动窗口、拥塞控制、快重传和快恢复[EB/OL]. [2021-07-22]. https://blog.csdn.net/u014590757/article/details/80214540.

[10] 邓丽山. 基于IP通信网络服务质量技术的阐述[J]. 广东科技，2009(20)：75-77.

[11] 王黎明，赵冰，郭珩，等. 网络拥塞控制专利技术综述[J]. 中国发明与专利，2018(z2)：85-90.

[12] 曹世宏. 流量监管和流量整形[EB/OL]. [2021-07-19]. http s://blog.csdn.net/qq_38265137/article/details/80466790.

作者简介：

赵纯熙，助理工程师，硕士，主要从事网络创新产品研发设计工作；童博，工程师，硕士，主要从事网络创新产品研发设计工作；刘锦波，助理工程师，学士，主要从事边缘计算与CDN相关研究工作。