

像客户一样体验网络 - 杜绝启动差异

通常，2/3 层的启动测试（如 RFC 2544）会在安装以太网服务时进行。提供商在通过 RFC 2544 测试（抑或全新的 Y.1564 测试）“认证”其网络后，还是会从使用视频会议、YouTube、Facebook 或云应用的业务终端客户处收到对糟糕应用性能的投诉。

安装测试中的差异，即传输控制协议 (TCP) 层测试（这是最优化终端用户应用层性能的关键所在）的遗漏，就是这类投诉的根本原因。图 1 简化的协议堆栈以及启动测试和最终用户体验之间的差异。



图 1. 简化的协议堆栈以及启动测试和最终用户体验之间的差异

该测试差异使网络提供商不会像客户一样体验网络性能，因此需要一种可在最终用户激活前验证 TCP 层性能的方案。TCP 层的测试可以避免额外的上门服务、技术支持呼叫以及客户的不满情绪，从而对提供商的运营开支 (OpEx) 产生持久的积极影响。

本白皮书:

- 简要介绍 TCP 协议
- 概述某些常见的客户驻地设备 (CPE) 和可能严重影响 TCP 及应用性能的网络问题
- 介绍全新的 IETF RFC 6349 TCP 测试方法
- 介绍若网络提供商进行基于 RFC 6349 的 TCP 层安装测试可以享有的优势

严重影响 TCP 的网络和 CPE 问题

TCP 在公开系统互联 (OSI) 第 4 层运行, 并驻留在 IP 层 3 顶端。TCP 最重要的特点之一便是可靠; 并且如果数据包丢失, TCP 可确保将数据包重新传输至接收方。

此外, 对于广域网 (WAN) 连接, TCP 必须正确配置, 调节发送方在从接收方接收确认 (ACK) 前可以传输的字节数。该“飞行中的”字节数通常被定义为 TCP 窗口; 即使事实上存在多个运行中的 TCP 窗口机制。

图 2 说明 45 Mbps WAN 连接速度、25 毫秒往返时延 (RTD) 或延迟的“飞行中的”TCP 数据字节数概念。

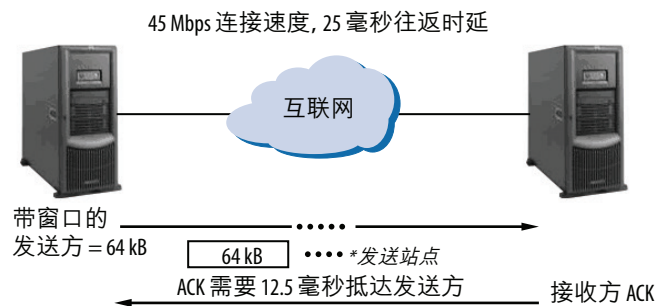


图 2. 45 Mbps WAN 连接速度、25 毫秒往返时延的“飞行中的”TCP 数据字节数图示。

在图 2 中, TCP 窗口未经正确调节, 发送方在需要 ACK 前仅传输了 64 kB。

如 RFC 6349 所述, 带宽时延乘积 (BDP) 即是最适宜的 TCP 窗口, 其计算公式如下:

$$BDP = \frac{\text{连接瓶颈带宽} \times \text{往返时间}}{8}$$

在此例中, BDP 可达到 140 kB, 为发送方的 64 kB 窗口速度的两倍多, 而发送方仅达到了约 20 Mbps 的流量。

TCP 的另一个关键属性是其具有突发性, 而非恒定比特率。因此, 千兆以太网 (GigE) 局域网 (LAN) (WAN 为 100 Mbps) 将引发多个实例, 此时, WAN 网络无法正确处理 GigE“突发”, 导致丢包与 TCP 重发。较高的网络延时将极大地影响 TCP 流量, 因为 TCP 状态机必须调升至最优化的传输速率 (且无丢包现象)。

从 LAN 将 TCP“降档”至 WAN 的主要方法为缓冲和流量整形。图 3 说明了网络缓冲和网络设备的输出队列。输出队列基于多种服务质量 (QoS) 机制 (如差分服务代码点 (DSCP)、虚拟局域网 (VLAN) 标签等) 对流量进行优先等级处理, 并根据流量等级分配队列深度。使用缺省队列深度可能丢失突发性 TCP 流量。丢包会导致 TCP 重发, 从而严重制约最终用户体验。

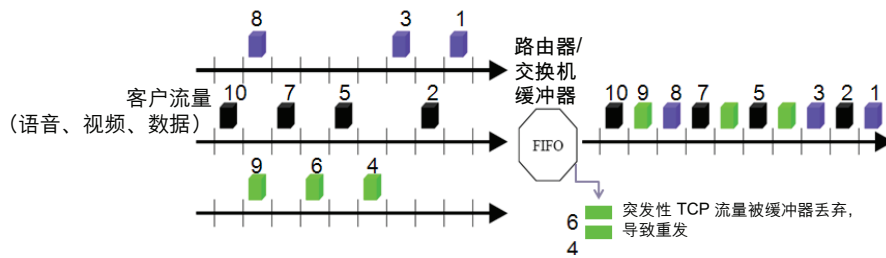


图 3. 网络缓冲和网络设备的输出队列

从 LAN 将 TCP “降档”至 WAN 的第二种方法为流量整形或“智能”网络缓冲，在这种情况下，网络设备将根据承诺信息速率 (CIR) 调整流量。流量整形应在 CPE 边缘设备上进行，但网络提供商也可调整流量，以极大地改善 TCP 性能和最终用户体验。

TCP 流量从较高速度接口降至较低速度接口时，如不进行调整，“网络监管器”可能会制约 TCP 性能。与整形相反，流量监管将“砍下”CIR 上的额外流量，这会导致 TCP 重发并严重制约最终用户性能。图 4 对比了流量整形器与监管器的功能。

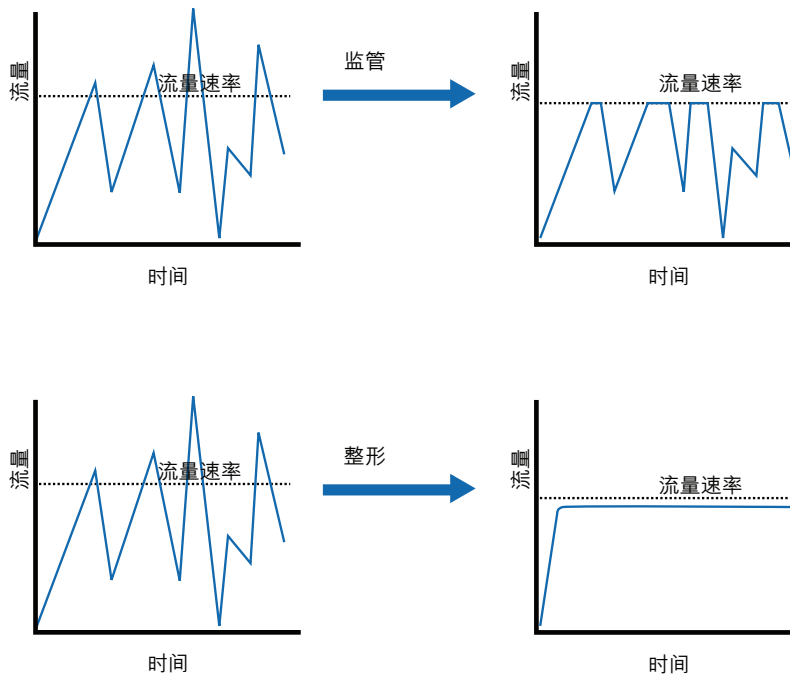


图 4. 流量整形与监管的对比

在客户激活前，使用 RFC 6349 测试方法调整缓冲器队列和流量整形可对 WAN 的 TCP 流量进行优化。

RFC 6349 TCP 测试方法

RFC 6349 是一种创新型 TCP 流量测试方法，由 Viavi 联合 Bell Canada 与 Deutsche Telecom 的代表合作开发。RFC 6349 最近由互联网工程任务组 (IETF) 发布，可为 TCP 流量分析提供可重复的测试方法，凭借系统化流程、指标和准则来优化网络和服务水平。

RFC 6349 建议在 TCP 测试前始终进行 2/3 层启动测试。在 2/3 层验证网络后，RFC 6349 将具体进行下面的三个测试步骤。

- **路径 MTU 检测（基于 RFC 4821）** 使用活动 TCP 段大小测试验证网络的最大传输单位 (MTU)，以确保 TCP 有效载荷保持一体性。
- **基准往返时延和带宽预测最优化 TCP 窗口大小**，以自动计算 TCP BDP。
- **单个及多个 TCP 连接流量测试验证 TCP 窗口大小预测**，以支持自动化的“全管道”TCP 测试。

在任何 TCP/IP 网络通信中，TCP 重发都属于正常现象。如仅凭数字本身，要确定影响性能的重发数量并不容易。RFC 6349 定义了一种全新标准，可深入洞悉因有效负载重发而使用的网络传输的相对比例。

该标准为 TCP 效率标准，或未重发的字节数比例，其计算方式如下：

$$\frac{\text{已传输字节数} - \text{重发字节数}}{\text{已传输字节数}} \times 100$$

已传输字节数为所传输的 TCP 有效负载字节数总数，包括原始和重发字节数。该标准可在各种 QoS 机制（流量管理、拥塞避免）和各种 TCP 实施（Windows XP 和 Linux 等）之间进行比较。

例如，如果发送了 100,000 字节，其中有 2,000 字节重发，TCP 效率的计算方法为：

$$\frac{102,000 - 2,000}{102,000} \times 100 = 98.03\%$$

注意：2/3 层的丢包比例与重发字节数比例并无直接关联，因为丢包的分配可在很大程度上影响 TCP 重发的方式。TCP 效率标准可让网络提供商针对各种服务类别 (CoS) 等级建立一个 TCP 丢失阈值。

RFC 6349 还定义了缓冲器时延比例，可表示在 TCP 流量测试期间从基准往返时间 (RTT) 增加的量，基准 RTT 为在无拥塞的情况下网络路径所固有的 RTT。

缓冲器时延比例定义如下：

$$\frac{\text{传输中的平均 RTT} - \text{基准 RTT}}{\text{基准 RTT}} \times 100$$

例如，如果网络路径的基准 RTT 为 25 毫秒，而在 TCP 传输期间平均 RTT 增加到 32 毫秒，那么缓冲器时延比例的计算公式即为：

$$\frac{32 - 25}{25} \times 100 = 28\%$$

换言之，TCP 传输拥有 28% 的额外 RTD（拥塞），这将导致 TCP 总流量的成比例降低，而最终用户会面临更长时间的时延。

RFC 6349 是对 RFC 2544 测试的完善，可弥补最终用户体验和提供商的网络测试方法之间的差距。下表说明了 RFC 2544 和 Y.1564sam 的网络测试适用性以及进行基于 RFC 6349 的 TCP 测试的额外优势。

启动相关问题	RFC 2544	Y.1564sam	RFC 6349
单一服务、2/3 层 SLA 问题，如丢失和抖动XX不适用	X	X	不适用
多种服务、2/3 层 SLA 问题，如服务优先化、丢失和抖动X不适用		X	不适用
演示最终用户 TCP 窗口流量大小的影响（CPE 问题）			X
设备缓冲器不足，无法处理突发性用户应用			X
监管对 TCP 性能的影响			X

一个姊妹应用注释《Viavi TrueSpeed 的 RFC 6349 测试》提供这一重要新方法以及此新 IETF RFC 自动化 Viavi 测试实施所隐含的详细信息。

造成业务客户不满意的两种情形

通常，两种原因导致业务客户对其网络的糟糕应用性能进行投诉：

- 业务客户错误配置客户设备或运行有缺点的“速度测试”程序。如高带宽 WAN 网络使用缺省 TCP 窗口大小运行，则会严重降低可达到的 TCP 性能。用户在性能糟糕的 PC 上运行使用开源工具（如 Iperf）的 TCP 测试，或该测试无法达到 SLA 流量，这时即为有缺点的速度测试。
- 网络提供商可能会遇到需要调节的网络问题，如端口缓冲器大小不足以应对 TCP 突发。传统意义上的 2/3 层测试无法像突发性 TCP 流量一样对网络造成压力。突发性业务客户流量可导致重发和糟糕性能，因此需要监管。增加端口缓冲器大小（一种流量整形方式）可极大地减少丢包现象并改善 TCP 性能。

两个现实案例的总结将在随后的子章节中说明各自的结果。每个案例均使用传统的 2/3 层安装程序，并且在客户激活前未进行 TCP 层测试。两个案例都突显了传统的故障排除成本和客户不满情绪，而这些提供商均可在客户激活前通过实施 TCP 测试予以避免。

CPE 问题：未最优化的 TCP 配置

对于此情形，拥有两个地点的业务客户从业务提供商处购买了 100 Mbps 的透明 LAN 服务，其网络配置见图 5。

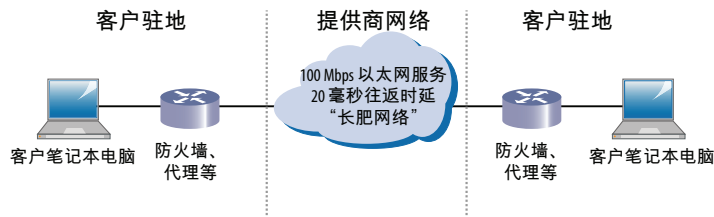


图 5. 业务客户的 100 Mbps 透明 LAN 配置

下表显示了 RFC 2544 测试的结果和网络提供商通过 FTP 下载和上传运行的客户速度测试结果：

RFC 2544 结果	FTP 结果
100 Mbps 流量	25 Mbps 流量

自然，业务客户对 FTP 性能不满并质疑所提供的服务。提供商的技术人员重新运行 RFC 2544 测试时，收到大量客诉并且需要上门服务。

这种情形的问题在于 100 Mbps、20 ms 的延迟网络，这种网络通常被称为“长肥网络”（LFN），需要在终端主机电脑中具有更大的 TCP 窗口设置。适用于这种情形的理想 TCP 窗口设置（与在 RFC 6349 中指定的相同）可以使用下列公式算出：

$$\frac{100 \text{ Mbps} \times 20 \text{ ms}}{8} = 250,000 \text{ 字节}$$

尽管传统的网络提供商不提供 CPE 设备，但也应该在分界点之间进行 RFC 6349 TCP 测试。RFC 6349 提供针对各种网络条件的指南和相关的最佳 TCP 窗口尺寸，这可以帮助网络提供商优化其终端用户的网络体验，实现更高的满意度和忠诚度。

对于此白皮书，网络提供商提供了四种不同的 TCP 窗口尺寸，下表反映了收到的实际效果。

TCP 窗口尺寸	流量
32 kB	12.5 Mbps
64 kB	25 Mbps
128 kB	50 Mbps
256 kB	100 Mbps

测试清楚地显示了网络运行正常，终端客户需要在 CPE 主机中启用 TCP 窗口缩放选项。这证明终端客户没有在其中一个服务器上使用窗口缩放，最大窗口只有 64 kB（只达到了 100 Mbps CIR 中的 25 Mbps）。

最后，我们估计提供商的额外 OpEx 成本将达到 7,000 美元。此提供商目前在其标准的服务激活流程中使用 RFC 6349 测试方法。

网络提供商问题：网络缓冲不足

对于此情形，拥有两个位置的业务客户从业务提供商处购买了 300 Mbps 透明 LAN 服务，网络配置显示在图 6 中。请注意，两个地点之间的延迟约为 50 ms。



图 6. 业务客户的 300 Mbps 透明 LAN 配置

下表显示了 RFC 2544 测试以及客户的速度测试结果，结果由网络提供商使用 iperf 程序测试得到。

RFC 2544 结果	iperf 结果
300 Mbps 流量	95 Mbps 流量

同样，提供商的技术人员重新运行 RFC 2544 测试时，收到大量客诉并且需要上门服务。终端客户要求提供商进行某种形式的 TCP 层速度测试，因为他们对提供商展示的结果缺乏信心。

在此测试情形中，BDP 很大，按照下列方式算出：

$$\frac{300 \text{ Mbps} \times 50 \text{ ms}}{8} = 1,875,000 \text{ 字节}$$

确实，这个 BDP 很大，但是具备 TCP 常识的终端客户在运行并行的 TCP 会话（此案例中为 30），尝试正确地“填补空缺”。

RFC 6349 包括并行 TCP 连接测试，这是适用于超大 BDP 的重要测试方法。提供商使用 Viavi TrueSpeed 解决方案复制此测试情形，并使用 Viavi 硬件 TCP 解决方案排除运行 iperf 程序的终端用户工作工作站可能影响测试结果的不确定性。

对于此情形，提供商实际上犯了相同的错误，只能达到终端客户要求索赔的 95 Mbps 流量。

提供商识别出问题在于边界路由设备中的默认缓冲不足。通过增加设备中的缓冲大小，提供商可以重新测试并证明正确的 300 Mbps TCP 性能。

最后，我们估计提供商的额外 OpEx 成本为 15,000 美元。此提供商目前在其标准的服务激活流程中使用 RFC 6349 测试方法。

结论

此白皮书着重介绍了当前以太网服务激活方法 RFC 2544 和 Y.1564 以及终端客户体验。客户应用程序位于 TCP 层之上，其性能在从 LAN 转移到 WAN 时将受到严重影响。

通过基于 RFC 6349 的测试提高 RFC 2544 和/或 Y.1564 服务激活，使供应商能够对客户的网络体验感同身受，有效地节约 OpEx 成本，极大增加首次客户满意度，带来更高的客户忠诚度和更强的盈利能力。

Viavi 采用 RFC 6349 TCP 测试，并且已经创造性地实现了自动化。Viavi TrueSpeed 测试功能是业内首个兼容 RFC 6349 的测试解决方案。TrueSpeed 使网络提供商在上门服务期间，由同等水平的技术人员在不到五分钟的时间内即可运行传统的 RFC 2544 安装和 TCP 测试。

姊妹应用注释和通过 *Viavi TrueSpeed* 进行的 RFC 6349 测试，详细介绍了新的 RFC 并着重说明了 TrueSpeed 测试功能。



北京
上海
深圳

电话: +8610 6476 1300
传真: +8610 6476 1302
电话: +8621 6859 5270
传真: +8621 6859 5265
电话: +86755 8691 0100
传真: +86755 8691 0001

© 2015 Viavi Solutions Inc.
本文档中的产品规格及描述可能会有所更改，
恕不另行通知。
rfc6349.wp.tfs.tm.zh-cn
30173043 000 0911