

通过 VIAVI Solutions 进行的 RFC 6349 测试 — 像客户一样体验您的网络

RFC 6349 是 VIAVI 与来自 Bell Canada 和 Deutsche Telecom 的代表合作开发的一种新的传输控制协议 (TCP) 吞吐量测试方法。RFC 6349 由互联网工程任务组(IETF) 组织发布，提供了一种可重复的 TCP 吞吐量分析测试方法，以及用于优化网络和服务器性能的系统化流程、指标和准则。

本应用指南对“TCP 吞吐量测试框架”（即 RFC 6349）进行了简要说明，并且重点描述了 TrueSpeed™ 这一自动化并且完全符合标准的 VIAVI RFC 6349 的实现。

TrueSpeed 现在可在 MTS-5800 手持式网络测试仪、MTS 5800-100G、MAP-2100 和 VIAVI Fusion（一种虚拟网络测试和保障系统）上提供。

本应用指南还讨论了 TrueSpeed RFC 6349 与 ITU Y.1564 以太网服务激活标准的集成。这一功能强大的测试组合提供了一种全面的方法来确保在多服务（例如三网融合）环境中实现优化的最终客户体验。

RFC 6349 TCP 测试方法

RFC 6349 指定了一种用于在托管 IP 网络中测量端对端 TCP 吞吐量的实用方法，其目标是为了更好地展现用户体验情况。在 RFC 6349 框架中，还指定了 TCP 和 IP 参数来优化 TCP 吞吐量。

RFC 6349 建议在 TCP 测试之前始终执行第 2/3 层开通测试。验证位于第 2/3 层的网络之后，RFC 6349 指定执行以下三个测试步骤。

- 路径 MTU 检测（依照 RFC 4821），用于验证网络最大传输单元 (MTU)，并进行活动 TCP 段大小测试来确保 TCP 净荷保持完整
- 基准往返延迟和带宽测试，预测最佳 TCP 窗口大小，用于自动计算 TCP BDP
- 单一和多个 TCP 连接吞吐量测试，用于验证实现自动“满管”TCP 测试的 TCP 窗口大小预测

以下各个小节提供了每个 RFC 6349 测试步骤的详细信息。

路径 MTU 发现 (依照 RFC 4821)

TCP 实现应使用依赖于互联网控制消息协议 (ICMP) “需要分片” 消息的路径 MTU 发现技术 (PMTUD) 来了解路径 MTU。如果设备要发送的数据包在 IP 头中设置了不分片 (DF) 位, 并且该数据包比下一个中继段的 MTU 大, 则会丢弃该数据包, 且该设备会将一条 ICMP 需要分片消息发回到最初发出该数据包的主机。ICMP 需要分片消息包括下一个中继段的 MTU, PMTUD 使用该 MTU 来对自身进行调整。遗憾的是, 由于许多网络管理员完全禁用了 ICMP, 因此这一技术可能多少有些不可靠。

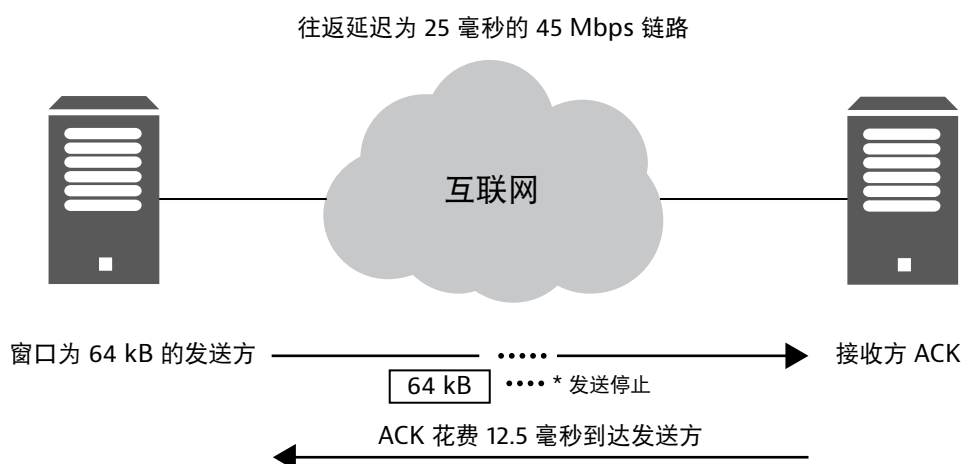
因此, RFC 6349 建议执行依照 RFC 4821 的分组层路径 MTU 发现 (PLPMTUD) 来验证网络路径 MTU, 因为它可在有或者没有 ICMP 的情况下使用。PLPMTUD 指定使用实时 TCP 流量轮询网络来确定 MTU。它采用了与设置 IP 数据包 DF 位相同的技术, 但由于使用实时 TCP 会话, 因此不依赖于 ICMP。该算法使用 TCP 重发条件来搜索 MTU, 用于避免在所有后续步骤中分片。

基准往返延迟和带宽

务必要确定基准环回时延 (RTT) 或非堵塞固有延迟以及端到端网络的瓶颈带宽 (BB), 然后才能开始 TCP 测试。这些基准测量用于计算 BDP 和估算 TCP 接收窗口 (RWND) 大小, 并发送将在后续测试步骤中使用的套接字缓冲区。

在广域网 (WAN) 链路上, 必须对 TCP 进行适当的配置, 来调整发送方在接收来自接收方的确认 (ACK) 之前可传输的字节数。这一 “in-flight” 字节数通常称为 TCP 窗口; 尽管事实上存在若干个工作中的 TCP 窗口机制。

图 1 描绘了往返延迟 (RTD) 为 25 毫秒的 45 Mbps WAN 链路上的 TCP 传输中数据字节数这一概念。



1107.0422

图 1. RTD 为 25 毫秒的 45 Mbps WAN 链路上 TCP 传输中数据字节数的图示

在图 1 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，BDP 为最佳 TCP 窗口，计算方式为：

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

在此示例中，BDP 将为 140 kB，是发送方 64 kB 窗口大小的两倍多，并且发送方将只获得大约 20 Mbps 的吞吐量。

RFC 6349 定义了以下机制用于测量 RTT：

- 第 2/3 层的活动流量生成，以及从一端到另一端的环回
- 数据包捕获
- 网络设备中的扩展管理信息库 (MIB) (RFC 4898)
- ICMP ping

BDP 同时依赖于 RTT 和 BB，因此它还需要测量 BB。为运营网络采用的第 2/3 层测试（例如 RFC 2544）被指定为用于测量 BB 的一种方法。知道了 RTT 和 BB 之后，RFC 6349 将能够计算后续 TCP 吞吐量测试的预期 TCP 性能。

单个和多个 TCP 连接的吞吐量测试

决定采用单一还是多个 TCP 连接测试取决于与最终用户环境中配置的 TCP RWND 相关的 BDP 大小。例如，长肥网络 (LFN) 的 BDP 为 2 MB，则使用多个连接测试此网络路径可能更加现实。假定是 64 kB 的典型主机 TCP RWND 大小（例如，Windows XP），则使用 32 个 TCP 连接将模拟出小型办公室场景。

尽管 RFC 6349 不要求测试多个连接，但作为最现实的方法，强烈建议进行此测试来准确地验证 TCP 吞吐量。RFC 6349 还定义了要在 TCP 吞吐量测试过程中要测量的特定指标，接下来将加以讨论。

RFC 6349 指标

下面介绍了 RFC 6349 TCP 指标，以及使用这些指标来诊断 TCP 性能不理想原因的示例。

TCP 传输时间

第一个 RFC 6349 TCP 指标是 TCP 传输时间，该指标仅测量跨同步 TCP 连接传输数据块所花费的时间。理想的 TCP 传输时间来源于网络路径 BB 以及与网络路径关联的各种第 1/2/3 层开销，例如，通过 500 Mbps 以太网服务在五个同步 TCP 连接上进行的 100MB 批量传输，每个连接上传 100 MB。每个连接在测试期间可能会获得不同的吞吐量，因此确定总体吞吐率并不总是很容易，特别是在连接数增加的情况下尤为如此。

理想的 TCP 传输时间大约为 8 秒，但在此示例中，实际 TCP 传输时间为 12 秒。TCP 传输指数将为 $12 \div 8 = 1.5$ ，表明跨所有连接进行的传输花费的时间长达理想时间的 1.5 倍。

TCP 效率

TCP 重发在任何 TCP/IP 网络通信中都是正常现象。仅仅使用数字本身来确定会对性能产生影响的重发次数很困难。RFC 6349 定义了一个新指标来深入了解由于净荷重发而使用的网络传输的相对百分比。

此指标是“TCP 效率”指标或未重发字节百分比，定义方式如下：

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

传输的字节数是指传输的 TCP 净荷字节总数，包括原始字节数和重发的字节数。此指标对各种服务质量 (QoS) 机制（例如流量管理、拥塞避免）以及各种 TCP 实现（仅举几例，例如 Reno 和 Vegas）进行比较。

举例来说，如果发送了 100000 字节，并且 2000 字节必须重发，则 TCP 效率的计算方式为：

$$\frac{102000 - 2000}{102000} \times 100 = 98.03\%$$

请注意，第 2/3 层的数据包丢失百分比并不直接与重发字节百分比相关联，原因是数据包丢失的分布可能会对 TCP 重发的方式产生广泛影响。

缓冲区延迟百分比

RFC 6349 还定义了缓冲区延迟百分比，表示 TCP 吞吐量测试期间 RTT 相对于基准 RTT 的增加，基准 RTT 是指在无拥塞的情况下网络路径固有的 RTT。

缓冲区延迟百分比定义为：

$$\frac{\text{传输期间的平均 RTT} - \text{基准 RTT}}{\text{基准 RTT}} \times 100$$

举例来说，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，那么缓冲区延迟百分比的计算公式即为：

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

换言之，TCP 传输的 RTD 增加了 28%（拥塞），可能已经导致总体 TCP 吞吐量成比例减少，从而为最终用户带来更长的延迟。

RFC 6349 TCP 调优准则

对于 TCP 性能未达预期的情况，RFC 6349 针对可能的原因提供了相应准则。

- 中间网络设备可能会主动重新生成 TCP 连接，并可能会改变 TCP RWND 尺寸、MTU 和其他参数
- 通过流量管制而非流量整形来限制速率，结果导致因尾部丢弃而造成过多的 TCP 重传

- 最大 TCP 缓冲区空间

所有操作系统都有一个对 TCP 连接使用的系统内存量进行限制的全局机制。在某些系统上，每个连接都受到内存限制的约束，该限制应用于为输入数据、输出数据和控制使用的总内存。在其他系统上，每个连接的输入和输出缓冲区空间存在单独的限制。为客户端/服务器 IP 主机配置的最大 TCP 缓冲区空间限制对于高性能网络而言可能太小。

- 套接字缓冲区大小

大多数操作系统都支持可在最大内存限制范围内调整的单独每连接发送和接收缓冲区限制。这些套接字缓冲区必须足够大，才能容纳完整的 TCP BDP 字节以及开销。可以使用若干方法来调整套接字缓冲区大小，但 TCP 自动调优会根据需要自动调整这些值，以实现 TCP 性能和内存使用率的最佳平衡。

有关网络/主机问题和建议解决方案的完整列表，请参阅 RFC 6349。

RFC 6349 的 VIAVI 实现

VIAVI 已经将 RFC 6349 测试方法集成到其多个便携式网络测试设备、一个机架式测试设备和一个基于软件的网络测试系统中。TrueSpeed 使用测试配置文件，因此技术人员只需加载测试配置，按“执行”，即可完成测试并发布包含结果的测试报告。

图 2 阐释了一个使用 VIAVI TrueSpeed 测试功能的场景。

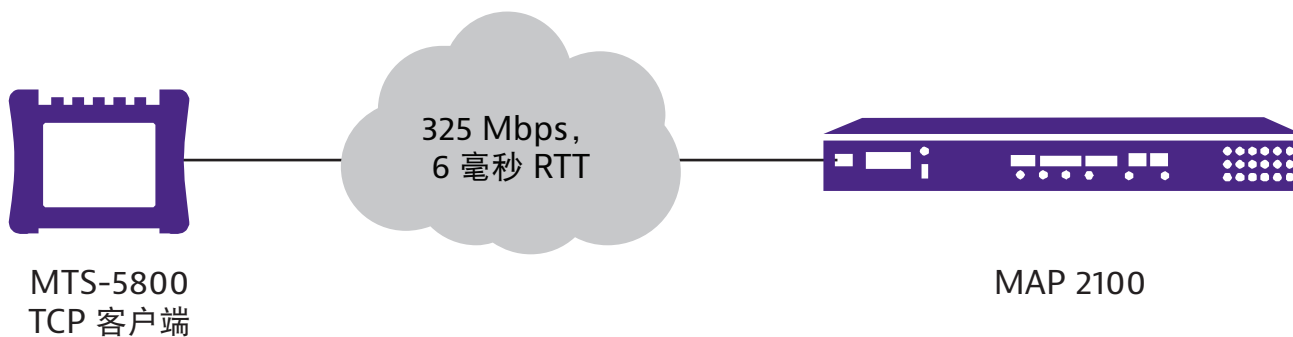


图 2. TrueSpeed 吞吐量测试的测试方案

1106.0422

这是一个客户承诺信息速率 (CIR) 为 325 Mbps、RTT 为 ~6 毫秒并且 BDP 为 ~250 kB 的 LFN。在本例中，MTS-5800 充当 TCP 客户端，该客户端执行针对 TCP 服务器（即 MTS-6000A）的上传吞吐量测试。

测试随后使用建议的默认设置自动运行，并在平均 3 分钟内完成。每个测试步骤都提供图形结果。

测试按 RFC 6349 中指定的顺序运行，第一项测试是路径 MTU 测试。图 11 显示了使用路径 MTU 为 1500 字节的示例网络时此测试的测试结果。

TrueSpeed 测试有两个工作流程：

- 安装测试模式：用户只需输入寻址和 CIR 值。MTS 依照 RFC 6349 自动填充所有 TCP 参数
 - 故障排查测试模式：更高级的用户可以控制 TCP 测试的许多方面来执行重点分析，该分析还包括高级流量整形测试
- 以下主题概述了这两种不同的测试模式。

安装测试模式

在此模式中，将派遣技术人员开通/安装新的最终客户服务，并将首先运行 RFC 2544 或 Y.1564 第 2/3 层测试。然后，使用所有相同的 MTS 寻址信息（例如，IP 地址、VLAN、QoS）来执行自动 TrueSpeed 安装测试。

为远程 MTS 配置了 IP 地址后，所有测试将从本地 MTS 中执行（一人 RFC 6349 测试）。下面概述了测试序列。

技术人员输入 CIR 和测试时间。

- MTS 自动填充 TCP 窗口大小和连接计数的所有字段
- MTS 从本地装置中依次运行上传和下载（速度测试）
- 报告简单的通过/未通过结果，并向本地 MTS 报告

下面连同 MTS 参考屏幕截图一起呈现了更详细的分步指南。

1. 技术人员为本地和远程 MTS 配置 IP 地址（并配置 VLAN，如果使用），然后还可以发出 ping 命令来验证第 3 层连接。

本地 MTS 连接到远程 MTS，并使用 TCP 端口 3000 进行所有测试配置和结果检索。

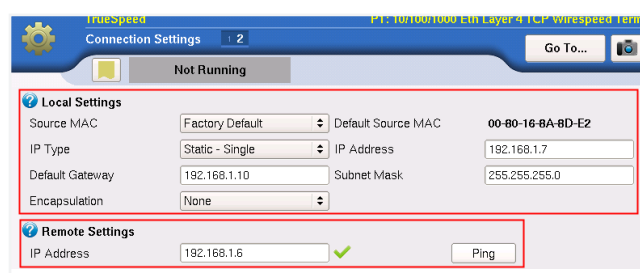


图 3. IP 地址配置

2. 技术人员配置一个屏幕来测试第 4 层的 SLA，如下所示：



图 4. SLA 测试配置

1. 所有 TCP 测试的总测试时间（最小值为 30 秒）。
2. 本地和远程 QoS/VLAN 设置（未显示 VLAN）。
3. 要测试的服务的第 1/2 层 CIR。

没有要配置的复杂 TCP 窗口大小或连接数。MTS 使用 RFC 6349 为用户自动计算这些值。

3. 技术人员单击**运行测试**。

本地 MTS 在上行和下行方向自动执行 RFC 6349 测试（像速度测试一样按顺序执行）。

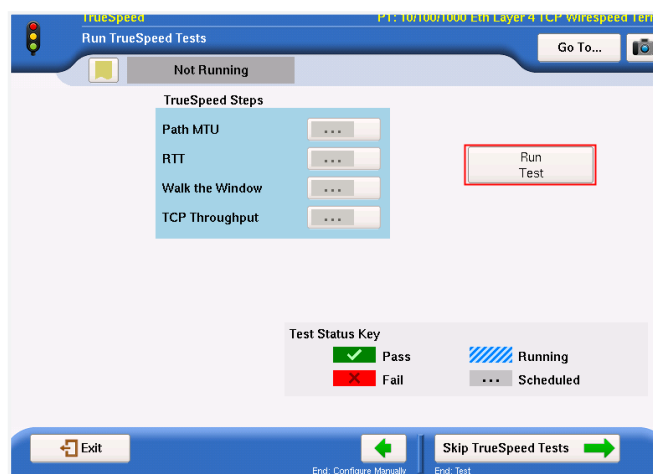


图 5. 运行 RFC 6349 测试

7 通过 TrueSpeed 进行的 RFC 6349 测试 — 像最终用户一样体验您的网络

将依照 RFC 6349 运行以下测试（下面做简要说明）；以下“故障排查测试模式”主题中提供了更详细的说明。

- 路径 MTU 检测（依照 RFC 4821）— 使用活动 TCP 段大小测试验证网络 MTU 以确保 TCP 净荷未分片
 - RTT 测试 — 测量服务的 RTT，并预测最佳 TCP 窗口大小以自动计算 TCP BDP
 - 遍历窗口 — 执行四种不同的 TCP 窗口大小测试，使第 4 层 CIR 吞吐量在 25% 到 100% 的范围内进行匀变
 - TCP 吞吐量 — 在 CIR 处执行更详细的吞吐量测试，并提供通过/未通过结论、RFC 6349 指标和详细图形
- 通过单击结果旁边的框可以显示和访问遍历窗口测试的结果。

请注意，测试有“上行”和“下行”按钮。在本例中，“上行”有 40 Mbps 监控程序，所有窗口设置都有显著的性能问题。CIR 窗口设置始终是测试的第四个窗口，在这种情况下应已生成 40 Mbps 的结果。

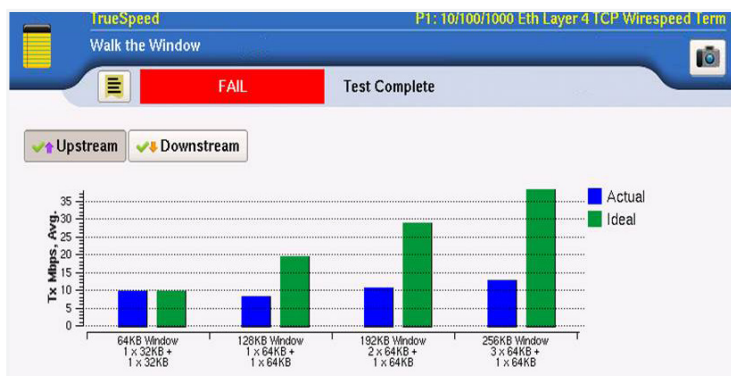


图 6. 遍历窗口测试屏幕—上行

在图 7 中，“下行”方向中没有监控程序，并且吞吐量符合每种情况中的理想值，包括第四个窗口大小（相当于 CIR 窗口大小）。

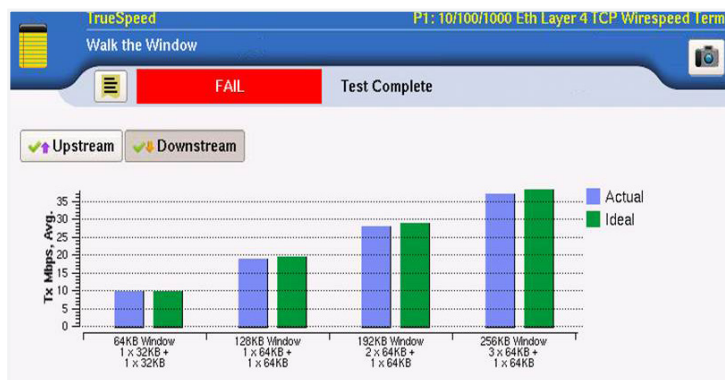


图 7. 遍历窗口测试屏幕—下行

如前所述，对 CIR 窗口大小（“遍历窗口”系列的第四个）执行了 TCP 吞吐量测试，并提供了更详细、更长的测试。

测试完成后，向用户呈现了简单的通过/未通过结论（图 8）以及详细的吞吐量测试结果屏幕（图 9）；在此示例中，由于 40 Mbps 监控程序的原因，测试在“上行”方向中失败。在这种情况下，实际客户吞吐量将仅为 12.3 Mbps。此外，“TCP 效率”和“缓冲区延迟”指标可帮助诊断 TCP 性能不良的原因。在此示例中，监控程序丢弃了数据包。

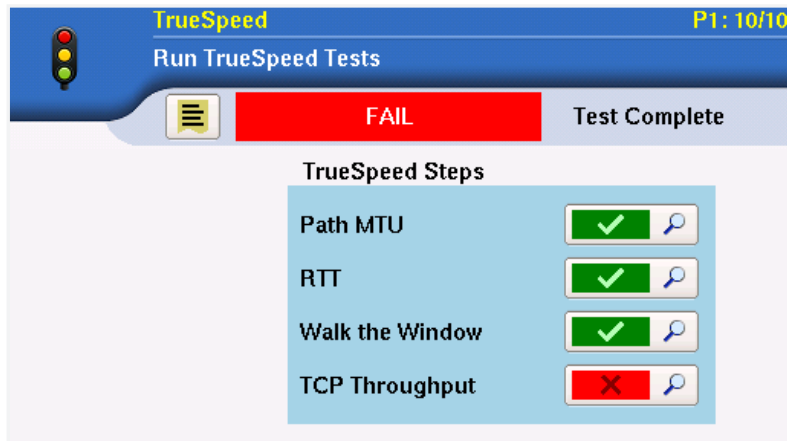


图 8. 通过/未通过测试结果

测试完成后，将生成一个图形测试报告，并且还可以保存测试配置。

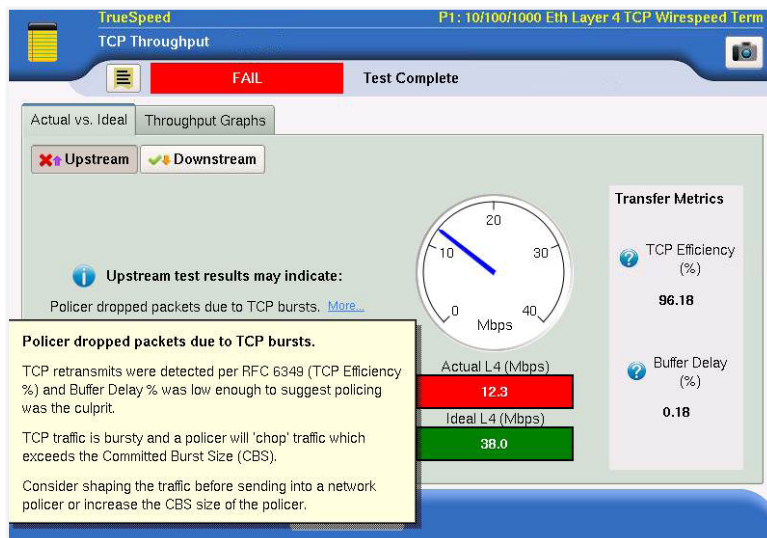


图 9. 详细的 TCP 吞吐量测试结果

故障排查测试模式

在此模式中，用户还可以加载测试配置或手动配置测试。高级现场技术人员可对此模式进行深层配置，并且，通过对 TCP 理论和 RFC 6349 结果进行更详细的解释，我们探讨了一个更详细的测试场景。

用户可执行所有 RFC 6349 测试步骤或这些测试的一小组，如图 10 所示。在本例中，CIR 为 325 Mbps，RTT 为 6.5 毫秒。

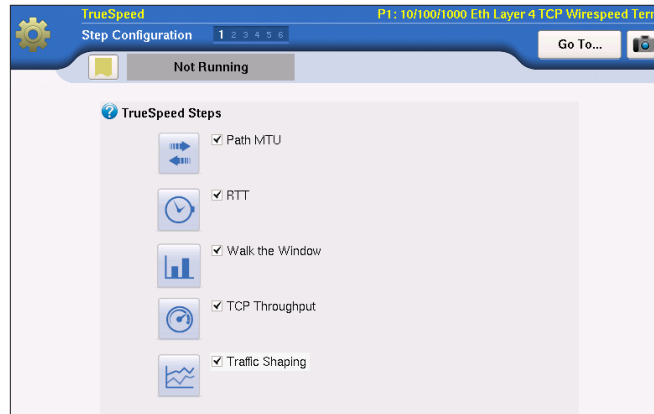


图 10. TrueSpeed 测试配置设置

测试随后使用建议的默认设置自动运行，并在平均 3 分钟内完成。每个测试步骤都提供图形结果。

测试按 RFC 6349 中指定的顺序运行，第一项测试是路径 MTU 测试。图 11 显示了使用路径 MTU 为 1500 字节的示例网络时此测试的测试结果。

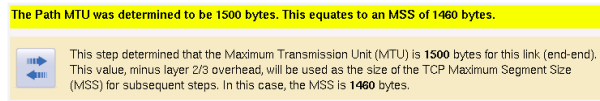


图 11. 路径 MTU 测试结果

完成路径 MTU 测试后，TrueSpeed 将继续执行 RTT 测试。由于 BDP 指定理想 TCP 窗口，因此该测试必不可少。BDP 用于在后续的测试步骤中预测理想 TCP 吞吐量。

图 12 显示了当此示例中的 RTT 为 6.5 毫秒时的 RTT 测试结果。

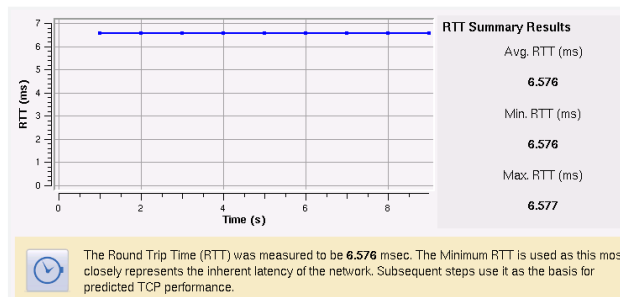


图 12. RTT 测试结果

遍历窗口测试提供测试的窗口大小结果和预期结果的信息性特征分析。遍历窗口测试使用路径 MTU 和 RTT 测试中的参数来执行窗口大小吞吐量测试。图 13 显示遍历窗口测试的结果。

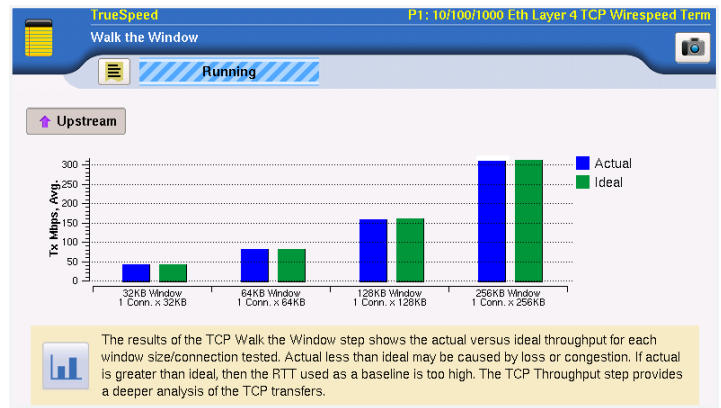


图 13. 遍历窗口测试结果

在图 13 的示例中，实际 TCP 吞吐量将只会使 325 Mbps 的 CIR 达到饱和，TCP 窗口大小配置为 256 kB。很多情况下，终端主机计算机使用小很多的窗口（例如 64 kB），产生的结果比预期吞吐量要低很多。在这里，64 kB 的窗口仅得到约 80 Mbps。

接下来，TCP 吞吐量测试可对有问题的窗口大小进行详细分析，并提供 RFC 6349 指标结果来帮助进行诊断。在图 10 中，TCP 窗口增加到 384 kB（使用大小为 128 kB 的三个连接），从而大幅超额预订了 325 Mbps 的 CIR。最终用户经常会产生“窗口越大就越好”这种极端想法。但是，如图 14 中的这种 WAN 环境所示，网络监控在出现 325 Mbps 的 CIR 时激活，并使 TCP 性能大幅降低。

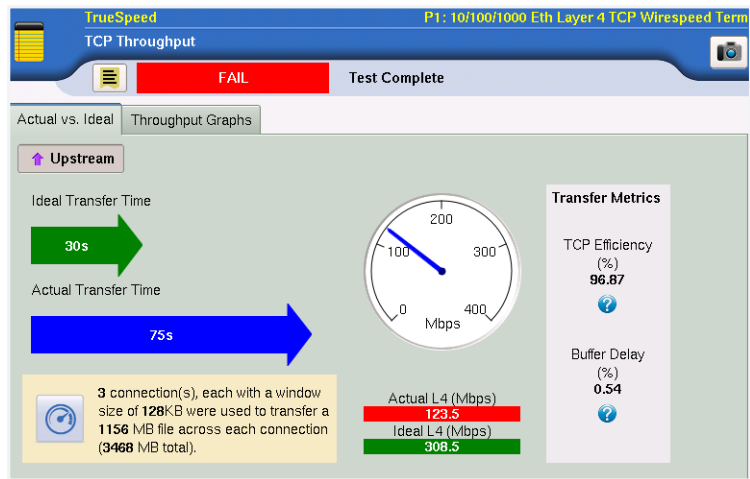


图 14. TCP 吞吐量测试结果（基本视图）

在这里，96.87%的“TCP效率”指标和只有0.54%的“缓冲区延迟百分比”表明是损耗（而不是缓冲延迟）导致了性能差距。图15显示了吞吐量图形的更详细检查。

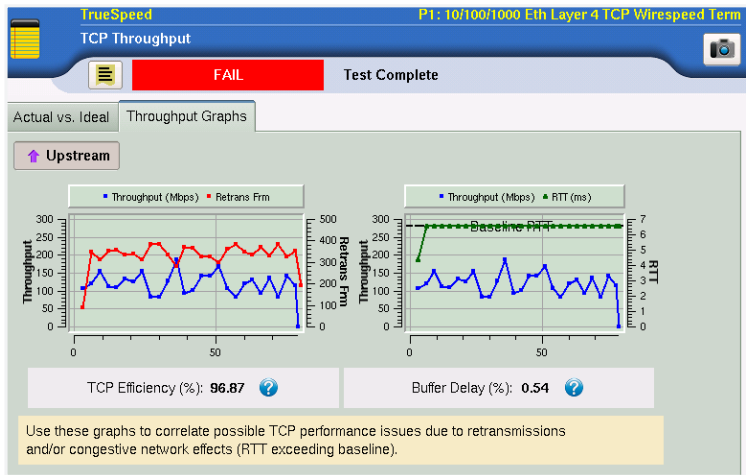


图 15. TCP 吞吐量测试图形

VIAMI 扩展了 RFC 6349 测试，并提供了流量整形测试。流量整形是指智能网络缓冲，在其中网络设备将依据 CIR 对流量进行整形。流量整形应在用户端设备 (CPE) 边缘设备处执行，但网络提供商也可以对流量进行整形来大幅改善 TCP 性能和最终客户体验。

如果在 TCP 流量从较高速度接口降档到较低速度时不对 TCP 流量进行整形，网络监控程序可能会对 TCP 性能产生不利影响。与整形相反，监管会削减超出 CIR 的过多流量，从而导致 TCP 重发，并使最终用户性能严重降低。图 16 对比了流量整形程序与监控程序的功能。

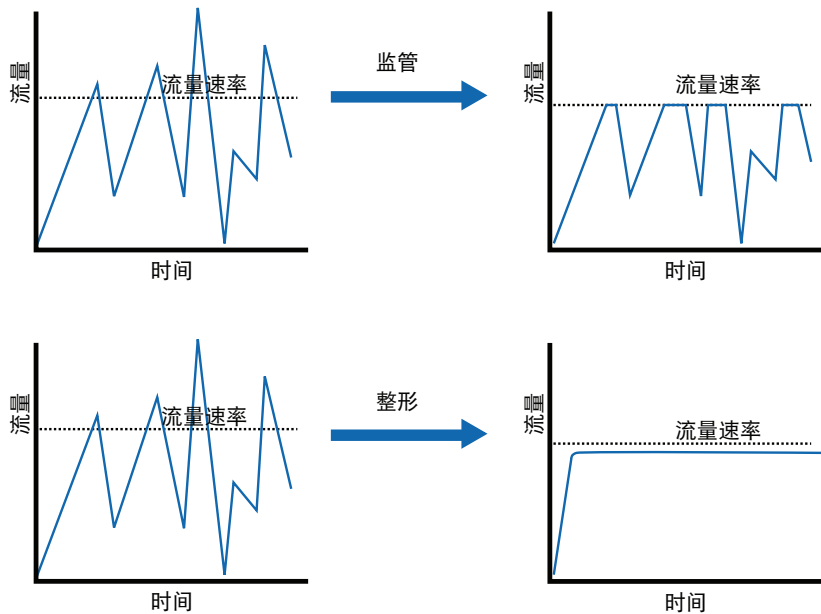


图 16 对比了流量整形程序与监控程序的功能

TrueSpeed 提供一个流量整形测试结果，该结果清晰显示进行了整形与监管处理的流量。图 17 显示进行了监管处理的流量，带宽分布在四个 TCP 连接之间非常参差不齐。

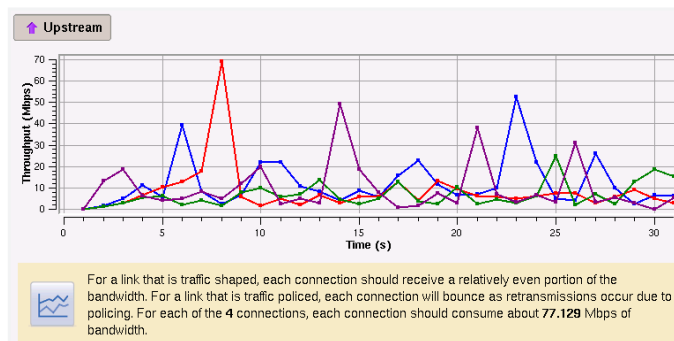


图 17. TrueSpeed 流量整形结果（其中的流量已经过监管处理）

图 18 显示了四个 TCP 连接之间的带宽均匀分布的流量整形。

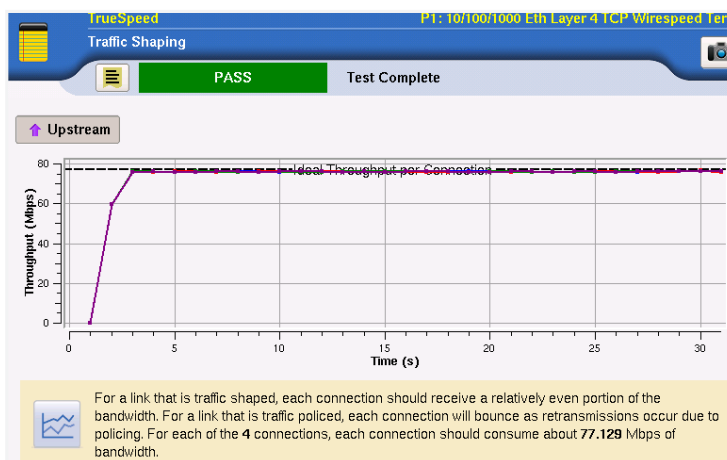


图 18. TrueSpeed 流量整形结果（其中的流量已经过整形处理）

将 TrueSpeed RFC 6349 与 Y.1564 集成

ITU Y.1564 是以太网服务激活的一项 ITU 标准。特点包括：

- 多项服务现场开通和安装测试可满足客户 SLA
- 在远端使用环回进行的自动端到端、多以太网/IP 服务测试
- 非常适合于 LTE/4G IP 服务和三网融合测试

Y.1564 检测的问题包括：

- 网络错误配置 — VLAN ID 和优先级、IP TOS、最大吞吐量
- 服务质量不良 — 太多延迟、抖动或损耗
- 在负载情况下，同一网络上的服务无法正常配合工作

由于定义 Y.1564 的目的只是为了验证第 2 层（以太网）和第 3 层（IP）性能，因此未测试 TCP 层的测试空白。最终结果是，Y.1564 可能会提供“通过”结果，最终客户的性能由于 TCP 相关性能问题（在前面各节中定义）而仍然可能不良。

在服务激活过程中将 TrueSpeed RFC 6349 测试与 Y.1564 集成可解决这一测试缺点。图 19 显示了 TrueSpeed 如何能与 Y.1564 服务性能测试集成。

在图 19 中，将语音和视频服务作为恒定比特率、基于 UDP 的数据流进行测试。但是，数据服务是使用 TrueSpeed 符合 RFC 6349 的流量（基于 TCP 并具有突发特性）测试的。TCP 应用的突发特性可能会对网络 QoS 产生压力，并导致性能问题。如果运行单纯的 Y.1564 测试，将检测不到这些问题。



图 19. 包含集成 TrueSpeed 服务的 Y.1564 性能测试阶段

这种集成方法的 VIAVI 实现称为 SAMComplete，它是业内唯一将 RFC 6349 与 Y.1564 集成的服务激活方法。SAMComplete 提供 TrueSpeed 服务的自动配置。用户只需指定 CIR，SAMComplete 将针对网络情况自动配置适当数量的 TCP 会话。这一集成测试结束时，将向用户提供 TrueSpeed 服务的简单通过/未通过状态，就像传统的 Y.1564 服务一样，如图 20 中所示。

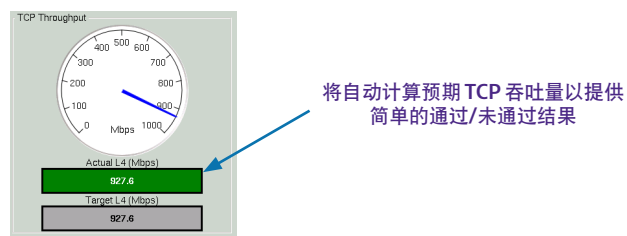


图 20. TrueSpeed RFC 6349 测试的简单通过/未通过结论

结论

此应用指南概述了 RFC 6349 中指定的 TCP 测试方法，这些方法可借助 TCP 吞吐量测试的分步最佳实践方法消除 TCP 测试方法中的显著可变性。RFC 6349 内指定的 TCP 指标可对网络问题（损耗和延迟）以及这些问题对总体 TCP 性能的影响程度进行客观测量。

在实际 TCP 吞吐量未达到理想值的情况下，RFC 6349 可提供用于对网络和/或终端主机进行调优的实用准则。

VIAMI TrueSpeed 测试是一种完全自动、符合 RFC 6349 标准的实现。由于它操作简单，只需按下按钮便可执行，并且具备自动报告功能（经验更丰富的网络工程师可用来验证和实现 SLA），因此，即使经验不足的技术人员也只需五分钟便可执行。



北京
上海
上海

深圳
网站:

电话: +8610 6539 1166
电话: +8621 6859 5260
电话: +8621 2028 3588
(仅限 TeraVM 及 TM-500 产品查询)
电话: +86 755 8869 6800
www.viavisolutions.cn

© 2022 VIAMI Solutions Inc.
本文档中的产品规格和描述如有更改，恕不另行通知。
rfc6349-an-tfs-zh-cn
30173049 904 0422