

使用 JDSU 的 TrueSpeed™ 进行 RFC 6349 测试——像您的客户一样体验您的网络

RFC 6349 是 JDSU 与加拿大贝尔和德国电信的代表共同编写的新传输控制协议 (TCP) 吞吐量测试方法。最近由互联网工程任务组 (IETF) 组织发布的 RFC 6349 提供了一种可重复的 TCP 吞吐量分析测试方法，具有优化网络和服务器性能的系统过程、指标和指南。

本应用说明总结了 RFC 6349 “TCP 吞吐量测试框架”，并强调了自动和完全兼容的 JDSU RFC 6349 实现——TrueSpeed，现在可在 JDSUT-BERD®/MTS-6000A 多服务应用模块 (MSAM) 和 T-BERD/MTS-5800 手持网络测试仪上使用。

本应用说明还讨论了 TrueSpeed RFC 6349 与 ITUY.1564 以太网服务激活标准的集成。这种强大的测试组合提供了一种全面的手段，确保在多服务（如三重播放）环境中优化的终端客户体验。

RFC 6349 TCP 测试方法

RFC 6349 指定了一种实用的方法，用于测量受管 IP 网络中的端到端 TCP 吞吐量，其目标是提供更好的用户体验指示。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 测试之前始终进行第 2/3 层的开通测试。在验证第 2/3 层的网络后，RFC 6349 指定进行以下三个测试步骤。

- 路径 MTU 检测 (每个 RFC 4821)，以验证具有主动 TCP 分段大小测试的网络最大传输单元 (MTU)，确保 TCP 有效负载保持不分片
- 为自动计算 TCP BDP 而预测最佳 TCP 窗口大小的基线往返延迟和带宽
- 验证启用自动“全管道”TCP 测试的 TCP 窗口大小预测的单个和多个 TCP 连接吞吐量测试

下一个小节说明了每个 RFC 6349 测试步骤的详细信息。

路径 MTU 发现(每个 RFC 4821)

TCP 实现应该使用路径 MTU 发现技术(PMTUD)，这项技术依赖于互联网控制消息协议(ICMP)“需要分片”消息来学习路径 MTU。当设备有一个要发送的数据包，该数据包在 IP 报头集中有一个不分片(DF)位，并且该数据包大于下一跳的 MTU，该数据包将被丢弃，该设备发送需要将消息分片的 ICMP 返回到发起该数据包的主机。需要分片消息的 ICMP 包括下一跳 MTU，PMTUD 用它来调整自己。不幸的是，由于许多网络管理器完全禁用 ICMP，这种技术可能有点不可靠。

因此，RFC 6349 建议对每个 RFC 4821 进行分组层路径 MTU 发现(PLPMTUD)来验证网络路径 MTU，因为它可以与 ICMP 一起使用或不与 ICMP 一起使用。PLPMTUD 指定实时 TCP 流量用于轮询 MTU 的网络。实现了设置 IP 数据包的 DF 位的相同技术，但它不依赖 ICMP，因为它使用的是实时 TCP 会话。该算法使用 TCP 重传条件搜索 MTU，用于避免后续所有步骤中的分片。

基线往返延迟和带宽

在 TCP 测试开始之前，应确定端到端网络的基线往返时间(RTT)，或非拥塞固有延迟和瓶颈带宽(BB)，这很重要。这些基线测量用于计算 BDP，并对 TCP 接收窗口(RWND)的大小进行预测，发送将在随后的测试步骤中使用的套接字缓冲。

在广域网(WAN)链路上，必须正确配置 TCP，以调整发送方在从接收方接收确认(ACK)之前可以发送的字节数。这个字节数“飞行中”通常被称为 TCP 窗口；尽管在现实中，有几个 TCP 窗口机制正在工作。

图 1 描述了具有 25 ms 往返延迟(RTD)或延迟的 45 Mbps 广域网链路上 TCP 飞行中数据字节的概念。

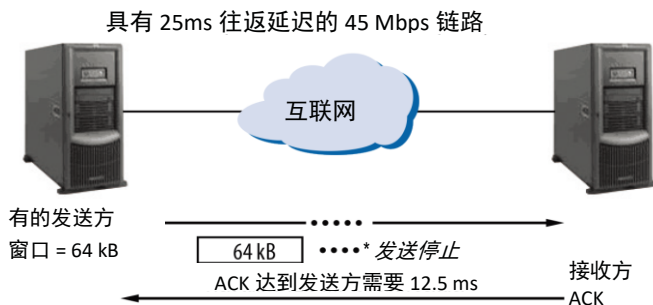


图 1 在具有 25ms RTD 的 45Mbps WAN 链路上的 TCP 飞行中数据字节的说明

在图 1 中，TCP 窗口被不正确地调优，在请求 ACK 之前，只从发送方传输 64 kB。

正如 RFC 6349 所描述的那样，BDP 是最佳的 TCP 窗口，如下计算：

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

在这个例子中，BDP 应该是 140 kB，这是发送方 64 kB 窗口大小的两倍多，发送方只会实现大约 20 Mbps 的吞吐量。

RFC 6349 定义了这些测量 RTT 的机制：

- 在第 2/3 层的主动流量生成和从一端到另一端的回环
- 数据包捕获
- 来自网络设备的扩展管理信息库(MIB)(RFC 4898)
- ICMP ping

BDP 同时取决于 RTT 和 BB，因此也需要测量 BB。第 2/3 层测试，如用于操作网络的 RFC 2544，被指定为测量 BB 的一种手段。一旦 RTT 和 BB 都已知，RFC 6349 就可以计算后续 TCP 吞吐量测试的预期 TCP 性能。

单个和多个 TCP 连接吞吐量测试

决定是否进行单个或多个 TCP 连接测试取决于相对于在最终用户环境中配置的 TCP RWND 的 BDP 大小。例如，如果用于长肥网络(LFN)的 BDP 是 2MB，那么测试有多个连接的网络路径可能更现实。假设使用 32TCP 连接的典型主机 TCP RWND 大小为 64kB(例如 Windows XP)将模拟一个小办公室场景。

虽然 RFC 6349 不要求测试多个连接，但强烈建议将其作为准确验证 TCP 吞吐量的最现实的手段。RFC 6349 还定义了了在 TCP 吞吐量测试期间要测量的特定指标接下来将讨论。

RFC 6349 指标

下面介绍 RFC 6349TCP 指标指标以及使用它们为次优 TCP 性能诊断原因的示例。

TCP 传输时间

第一个 RFC 6349TCP 指标是 TCP 传输时间，它简单地测量在同时 TCP 连接之间传输数据块所需的时间。

理想的 TCP 传输时间是根据网络路径 BB 和与网络路径相关的第 1/2/3 层开销推算出来，例如，在 500Mbps 以太网服务上的五个同时 TCP 连接上批量传输 100MB，每个连接上传 100MB。在测试期间，每个连接都可能实现不同的吞吐量，因此，确定总体吞吐率并不总是容易的，特别是连接数量增加的时候。

理想的 TCP 传输时间约为 8 秒，然而，在本例中，实际的 TCP 传输时间为 12 秒。TCP 传输指数将是 $12 \div 8 = 1.5$ ，表明横跨所有连接的传输比理想的时间长 1.5 倍。

TCP 效率

在任何 TCP/IP 网络通信中，TCP 重传都是正常现象。当简单地使用数字本身时，很难确定将影响性能的重传次数。RFC 6349 定义了一个新的指标，以了解由于重传有效负载而使用的网络传输的相对百分比。

此指标是 TCP 效率指标，或未重传的字节百分比，定义为：

$$\frac{\text{传输字节} - \text{重传字节}}{\text{传输字节}} \times 100$$

传输字节是传输的 TCP 有效负载字节的总数，包括原始字节和重传字节。这个指标提供了各种服务质量(QoS)机制之间的比较，例如流量管理、拥塞避免和各种 TCP 实现，例如 Reno 和 Vegas。

例如，如果发送了 100,000 个字节，并且必须重传 2,000 个字节，则 TCP 效率如下计算：

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

请注意，第 2/3 层的丢包百分比与字节重传百分比没有直接关联，因为丢包的分布会广泛影响 TCP 重传的方式。

缓冲延迟百分比

RFC 6349 还定义了缓冲延迟百分比，它表示从基线 RTT 开始的 TCP 吞吐量测试期间 RTT 的增加，而基线 RTT 是没有拥塞的网络路径固有的 RTT。

缓冲延迟百分比定义为：

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

例如，使用以下公式计算 25ms(在平均 RTT TCP 传输期间增加到 32ms)的基线 RTT 路径的网络的缓冲延迟百分比。

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

换句话说，TCP 传输经历了 28% 的额外 RTD (拥塞)，这可能导致 TCP 总吞吐量的比例下降，导致最终用户的更长的延迟。

RFC 6349 TCP 调优指南

对于 TCP 性能不符合预期的情况，RFC 6349 为可能的原因提供了指南。

- 中间网络设备可以主动地重新生成 TCP 连接，并可以更改 TCP RWND 大小、MTU 和其他事项
- 通过监管而不是整形进行速度限制造成由于尾部丢弃导致的过度 TCP 重传
- 最大 TCP 缓冲空间
所有操作系统都有一个全局机制，限制 TCP 连接使用的系统内存量。在某些系统上，每个连接都受内存限制，该内存限制应用于用于输入数据、输出数据和控件的总内存。在其他系统上，每个连接的输入和输出缓冲区空间都有单独的限制。客户端/服务器 IP 主机可以配置最大的 TCP 缓冲空间限制，这对于高性能网络来说太小了。
- 套接字缓冲大小
大多数操作系统支持单独的每个连接发送和接收缓冲区限制，这些限制可以在最大内存限制范围内进行调整。这些套接字缓冲必须足够大，以容纳完整的 TCP 字节加上开销。可以使用几种方法来调整套接字缓冲区大小，但是 TCP 自动调优会根据需要自动调整，以优化 TCP 性能和内存使用的平衡。

有关网络/主机问题和推荐解决方案的完整列表，请参照 RFC 6349。

JDSU 实现 RFC 6349

JDSU 将 RFC 6349 测试方法集成到其 T-BERD/MTS-5800 和 T-BERD/MTS-6000A MSAM 以太网分析仪 TrueSpeed 中，它是业界第一个完全自动化的 RFC 6349 实现。TrueSpeed 使用测试配置文件，这样技术人员可以简单地加载测试配置、按开始并发布带有结果的测试报告。

图 2 展示了使用 JDSU TrueSpeed 测试能力的情景。

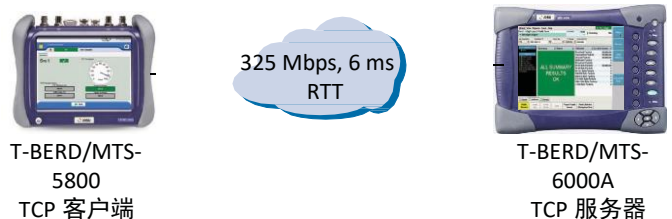


图 2. TrueSpeed 吞吐量测试的测试场景

这是一个 LFN，客户提交的信息速率(CIR)为 325Mbps，RTT 为 6ms 以下，BDP 为 250KB 以下。在这个例子中，T-BERD/MTS-5800 充当一个 TCP 客户端，它将吞吐量测试上传到 TCP 服务器，而服务器是 T-BERD/MTS-6000A。

然后，测试自动运行，并使用推荐的默认设置在平均 3 分钟内完成。每个测试步骤提供图形结果。

测试按照 RFC 6349 中指定的顺序运行，第一个是路径 MTU 测试。图 11 显示了使用我们的路径 MTU 为 1500 字节的示例网络进行此测试的测试结果。

TrueSpeed 有两个工作流程：

- 安装测试模式：用户只需输入寻址和 CIR 值即可。T-BERD/MTS 自动填入各 RFC 6349 的所有 TCP 参数
- 故障排除测试模式：更高级的用户可以控制 TCP 测试的许多方面，以执行集中分析，其中也包括高级流量整形测试

以下主题对两种不同的测试模式进行了总结。

安装测试模式

在这种模式下，技术人员被派去提供/安装新的终端客户服务，并将首先运行 RFC2544 或 Y.1564 第 2/3 层测试。然后，使用所有相同的 T-BERD/MTS 寻址信息(例如 IP 地址、VLAN、QoS)进行自动 TrueSpeed 安装测试。

使用已配置 IP 地址的远程 T-BERD/MTS，所有测试都是从本地 T-BERD/MTS(一人进行 RFC 6349 测试)进行的。以下是测试顺序的概述。

技术员输入 CIR 和测试时间。

- T-BERD/MTS 自动为 TCP 窗口大小和连接数填入所有字段
- T-BERD/MTS 运行上传，然后从本地单元下载(速度测试)
- 报告一个简单的通过/失败结果，并向本地 T-BERD/MTS 报告。

下面是一个更详细的分步指南以及 T-BERD/MTS 参考截图。

1. 技术人员为本地和远程 T-BERD/MTS 配置 IP 地址(如果使用，配置 VLAN)，然后还可以发出 ping 来验证第 3 层的连通性。
本地 T-BERD/MTS 连接到远程 T-BERD/MTS，并使用 TCP 端口 3000 进行所有测试配置和结果检索。

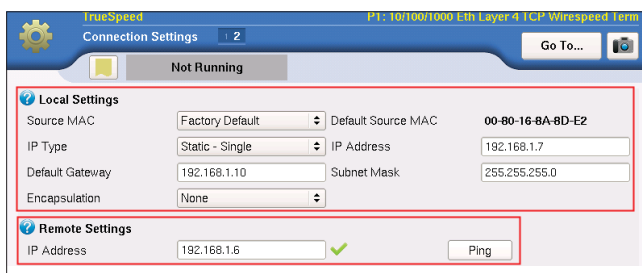


图 3 IP 地址配置

2. 如下所示，技术人员配置一个画面来测试第 4 层的 SLA：

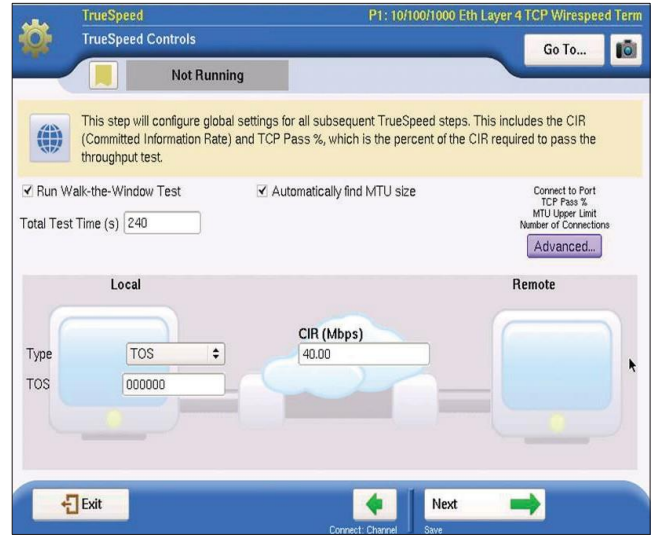


图 4 SLA 测试配置

1. 所有 TCP 测试的总测试时间（最短为 30 秒）。
2. 本地和远程 QoS/VLAN 设置(VLAN 不显示)。
3. 用于要测试的服务的 1/4 层 CIR。

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

3. 技术人员点击“运行测试”。

本地 T-BERD/MTS 自动在上行流和下行流方向进行 RFC 6349 测试（按顺序进行，如速度测试）。

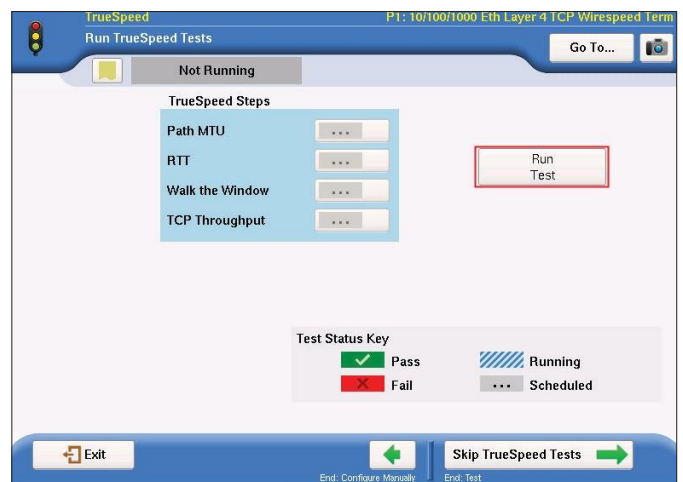


图 5 运行 RFC 6349 测试

下面的测试是按照每个 RFC 6349 运行的，下面有一个简短的描述；在下面的故障排除测试模式主题中提供了更详细的描述。

- 路径 MTU 检测(每个 RFC 4821)——验证网络 MTU 与主动 TCP 段大小测试，以确保 TCP 有效净荷不会被分片
- 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
- Walk-the-Window ——进行四种不同的 TCP 窗口大小测试，并将第 4 层 CIR 的吞吐量从 25%提高到 100%
- TCP 吞吐量——在 CIR 进行更详细的吞吐量测试并提供通过/失败判定、RFC6349 量度和详细图表

显示 Walk-the-Window 测试的结果，并通过单击结果旁边的框来访问。

请注意，所有测试都有一个上行流和下行流按钮。在这个例子中，上行流有一个 40M bps 的策略器，并且在所有窗口设置中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，应该产生 40 Mbps 的结果。

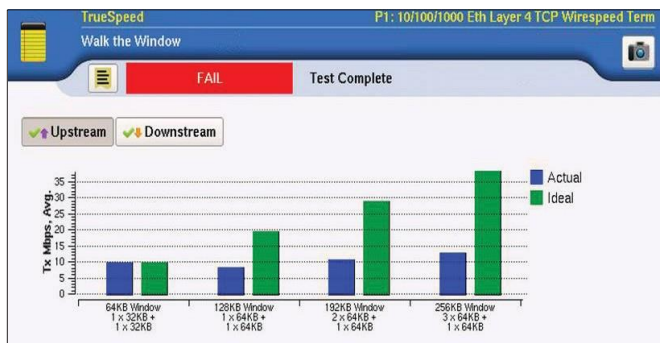


图 6. Walk-the-Window 测试画面 —— 上行流

在图 7 中，下行流方向没有策略器，吞吐量在每种情况下都达到理想状态，包括第四个窗口大小(等于 CIR 窗口大小)。

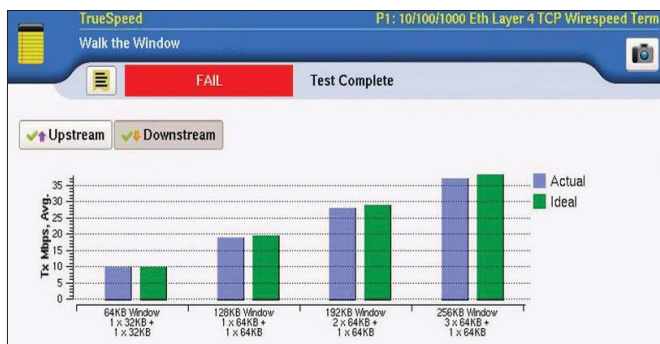


图 7 Walk-the-Window 测试画面— 下行流

如前所述，TCP 吞吐量测试是在 CIR 窗口大小(Walk-the-Window 系列的第 4 个)进行的，并提供了一个更详细、更长的测试。

在测试完成后，用户将得到一个简单的通过/失败判定 (图 8)以及一个详细的吞吐量测试结果画面(图 9)；在本例中，由于 40 Mbps 的策略器，测试在上行流方向失败。在这种情况下，实际客户吞吐量仅为 12.3 Mbps。此外，TCP 效率和缓冲延迟指标有助于诊断 TCP 性能差的原因。在这个例子中，策略器正在丢弃数据包。

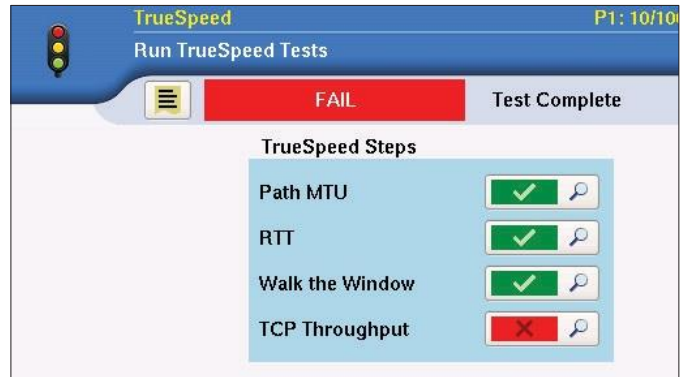


图 8 通过/失败测试结果

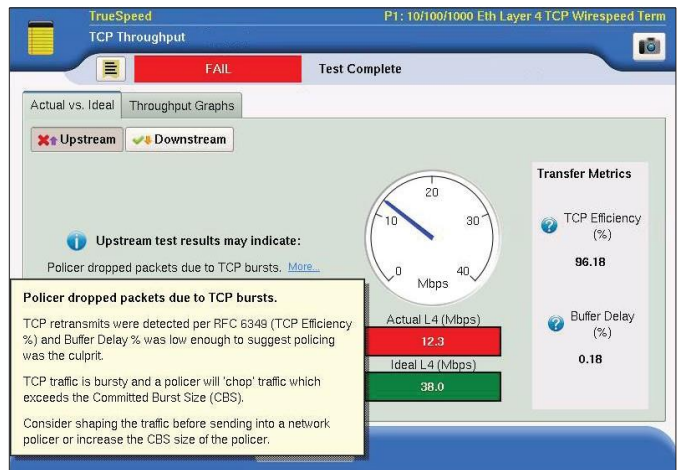


图 9 详细的 TCP-吞吐量测试结果

测试完成后，生成图形测试报告，也可以保存测试配置。

故障排除测试模式

在这种模式下，用户也可以加载测试配置或手动配置测试。对于高级现场技术人员来说，这种模式是高度可配置的，并通过对 TCP 理论和 RFC 6349 结果的更详细的解释来探索更详细的测试场景。

如图 10 所示，用户可以执行所有 RFC 6349 测试步骤或这些测试的子集。在本例中，CIR 为 325Mbps，RTT 为 6.5ms。

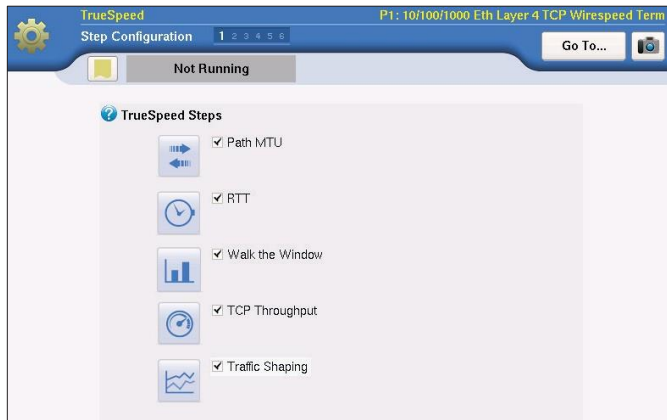


图 10. TrueSpeed 测试配置设置

然后，测试自动运行，并使用推荐的默认设置在平均 3 分钟内完成。每个测试步骤都提供图形结果。

测试按照 RFC 6349 中指定的顺序运行，第一个是路径 MTU 测试。图 11 显示了使用我们的路径 MTU 为 1500 字节的示例网络进行此测试的测试结果。

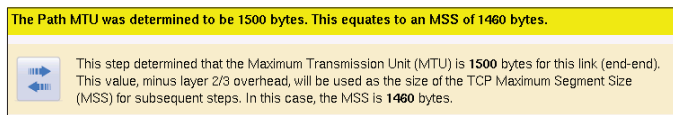


图 11 路径 MTU 测试结果

完成路径 MTU 测试后，TrueSpeed 继续进行必不可少的 RTT 测试，因为 BDP 指定了理想的 TCP 窗口。在随后的测试步骤中使用 BDP 来预测理想的 TCP 吞吐量。

图 12 显示了 RTT 为 6.5ms 的这个示例的 RTT 测试结果。

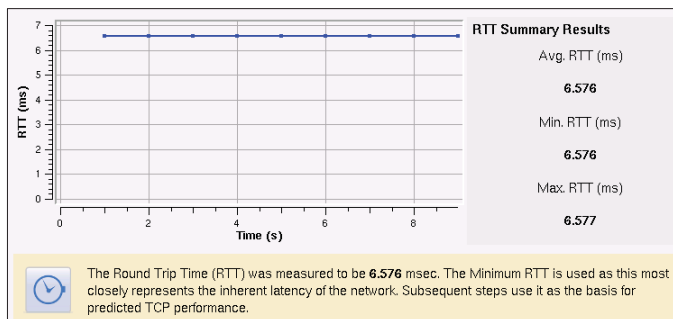


图 12RTT 测试结果

Walk the Window 测试提供了测试窗口大小结果和预期结果的信息特性。Walk the Window 测试使用来自路径 MTU 和 RTT 测试的参数来进行窗口大小吞吐量测试。图 13 显示了 Walk the Window 测试的结果。

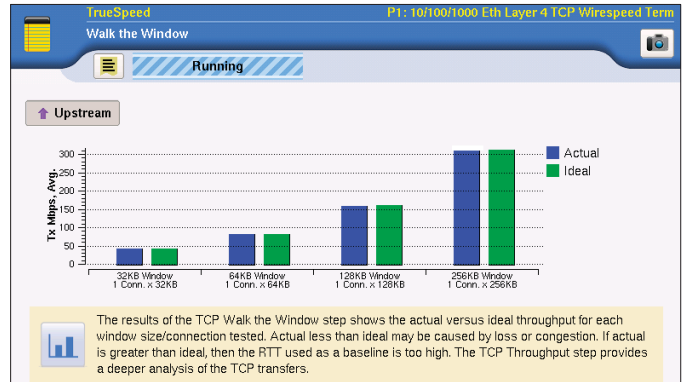


图 13Walk the Window 测试结果

在图 13 中的示例中，实际 TCP 吞吐量只会使配置为 256kB 的 TCP 窗口大小的 CIR 达到 325Mbps。很多时候，端主机计算机使用的窗口要小得多，比如 64kB，导致吞吐量远低于预期。在这里，一个 64kB 窗口只实现了不超过 80Mbps。

接下来，可以对详细分析有问题的窗口大小进行 TCP 吞吐量测试，并提供 RFC 6349 指标结果以帮助诊断。在图 14 中，TCP 窗口增加到 384 kB(使用大小为 128kB 的三个连接)，这大大超过了 325 Mbps 的 CIR。最终用户往往会有这种极端的想法：“窗口越大越好”。然而，如图 14 中所示，WAN 环境显示了在 325 Mbps CIR 上激活的网络监管，并显著降低了 TCP 性能。

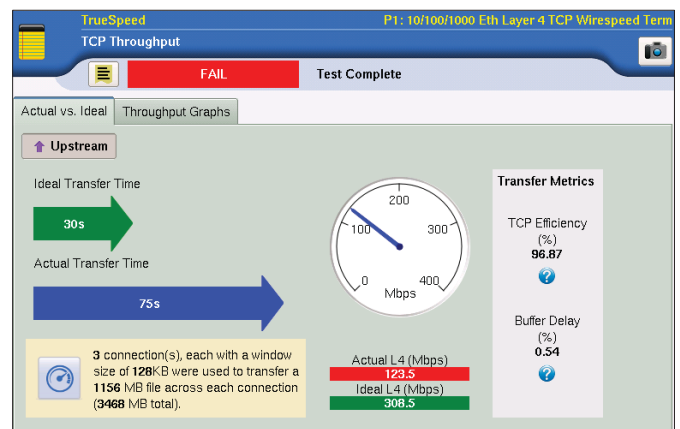


图 14. TCP 吞吐量测试结果(基本视图)

在这里，TCP 效率指标为 96.87%，缓冲延迟百分比仅为 0.54%，这表明损失导致了性能差距，而不是缓冲延迟。图 15 显示了对吞吐量图表更详细的研究。

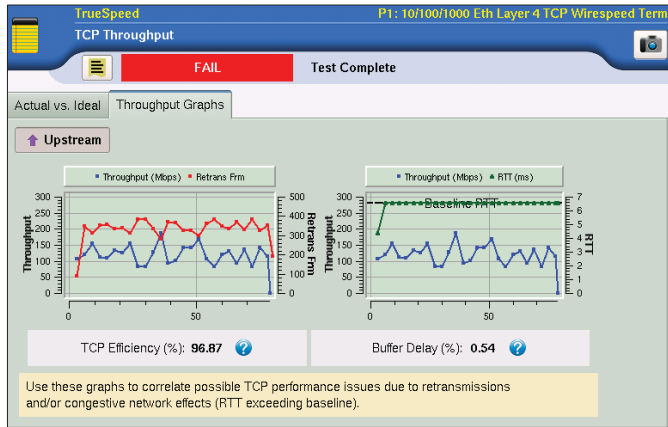


图 15. TCP 吞吐量测试图

JDSU 扩展了 RFC 6349 测试，并提供了流量整形测试。流量整形是智能网络缓冲，在其中，网络设备根据 CIR 对流量进行整形。流量整形应该在客户端设备(CPE)边缘设备上进行，但网络提供商也可以对流量进行整形，从而大大提高 TCP 性能和最终客户体验。

当 TCP 流量从较高的速度接口向下移动到较低的速度时，不对 TCP 流量进行整形，这样网络策略器不会影响 TCP 性能。与整形相反，在 CIR 之上进行监管砍去了多余的流量，导致 TCP 重传并严重降低最终用户的性能。图 16 对流量整形器和策略器的功能进行了对比。

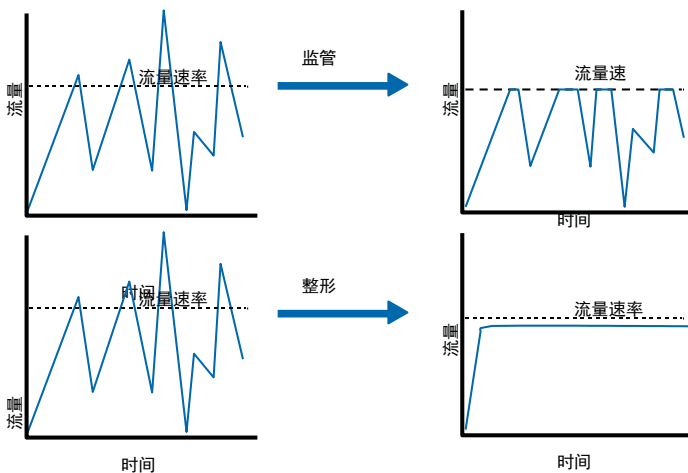


图 16 网络流量的监管与整形处理

TrueSpeed 提供了一个流量整形测试结果，它清楚地显示了正在被整形的流量与被监管的流量。图 17 显示了正在被监管的流量，并且四个 TCP 连接之间的带宽分布非常参差不齐。

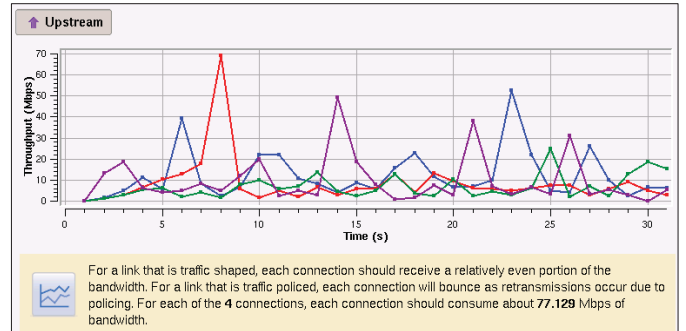


图 17. TrueSpeed 流量整形结果(其中的流量被监管)

图 18 显示了四个 TCP 连接之间带宽分布非常均匀的流量整形。

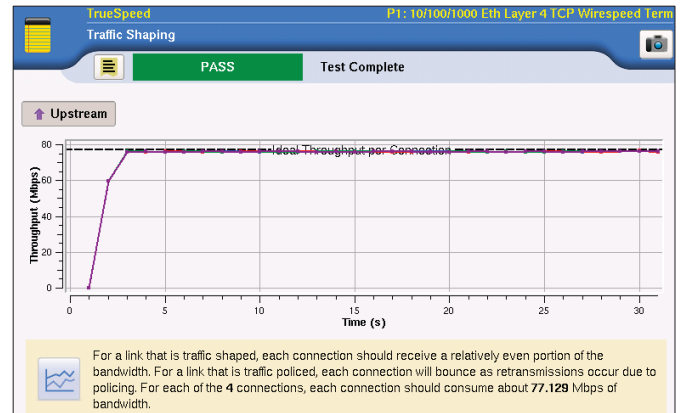


图 18. TrueSpeed 流量整形结果(其中的流量被整形)

整合 TrueSpeed RFC 6349 与 Y.1564

ITU.Y.1564 是激活以太网服务的 ITU 标准。重点如下：

- 多服务现场开通及安装测试，以满足客户 SLA
- 在远端使用回环的自动端到端、多以太网/IP 服务测试，
- 非常适合 LTE/4G IP 服务和三播放测试。

Y.1564 检测到的问题有：

- 网络配置错误 —— VLAN ID 和优先级、IPTOS、最大吞吐量
- 服务质量差 —— 延迟太多、抖动或丢失
- 在负载条件下，多项服务不能很好地一起在同一个网络上工作。

因为定义 Y.1564 只是为了验证第 2 层（以太网）和第 3 层(IP) 的性能，所以 TCP 层的测试间隙未经测试。净结果是 Y.1564 可以提供“传递”结果，但是由于前几节中定义的与 TCP 相关的性能问题，最终客户的性能仍然很差。

解决这个测试缺陷的方法是在服务激活时将 TrueSpeed RFC 6349 测试与 Y.1564 集成到一起。图 19 说明如何将 TrueSpeed 与 Y.1564 服务性能测试集成到一起。

在图 19 中，将语音和视频服务作为基于 UDP 的恒定比特率流测试。然而，数据服务使用基于 TCP 的突发 TrueSpeed RFC 6349 兼容流量测试。TCP 应用的突发特性会给网络 QoS 带来压力，导致运行纯 Y.1564 测试时未被发现的性能问题。



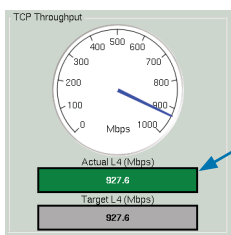
图 19. 含有集成的 TrueSpeed 服务的 Y.1564 性能测试阶段 JDSU 实现的这种集成方法称为 SAMComplete，它是业界唯一将 RFC 6349 与 Y.1564 集成在一起的服务激活方法。SamComplete 提供了 TrueSpeed 服务的自动配置。用户只需要指定 CIR 和 SAMComplete，将自动为网络条件配置适当数量的 TCP 会话。如图 20 所示，在此集成测试结束时，将为用户提供 TrueSpeed 服务的简单通过/失败状态，就像传统的 Y.1564 服务一样。

结论

本应用说明总结了 RFC 6349 中指定的 TCP 测试方法，该方法可以通过逐步的、最佳的 TCP 吞吐量测试方法来消除 TCP 测试方法中的显著变异性。在 RFC 6349 中指定的 TCP 指标揭示了网络问题（丢失和延迟）的客观指标，以及它们是如何影响整个 TCP 性能的。

在实际 TCP 吞吐量不等于理想的情况下，RFC 6349 提供了调优网络和/或终端主机的实用指南。

JDSU True Speed 测试是一个完全自动化的 RFC 6349 兼容的实现，即使是新手技术人员也可以在短短的 5 分钟内执行，因为它具有简单的按钮执行方式和自动报告能力，更有经验的网络工程师可以使用它来验证和实现 SLA。



预期 TCP 吞吐量被自动计算，以提供简单的通过/失败结果

图 20 TrueSpeed RFC 6349 测试的简单通过/失败判定



北美	免费: 1 855 ASK-JDSU	(1 855 275-5378)
拉丁美洲	电话: +1 954 688 5660	传真: +1 954 345 4668
亚太	电话: +852 2892 0990	传真: +852 2892 0770
欧洲、中东、非洲	电话: +49 7121 86 2222	传真: +49 7121 86 1222