

お客様と同じ視点でネットワークを見てください - 発見されたギャップの解消

従来、RFC 2544などのレイヤ2/3のターンアップ試験は、Ethernetサービスをインストールする時に実施されていました。プロバイダはRFC 2544試験（またはさらに最新のY.1564）で自身のネットワークを認証しているにもかかわらずYouTube、Facebook、その他クラウド・ベースのアプリケーションなど、ビデオ会議を使った実際の業務でエンドユーザからアプリケーションの性能が悪いと不満を聞かされていました。

インストール時の試験におけるギャップ、すなわちエンドユーザのアプリケーション層で最高性能を出すために重要な要素である、トランスミッション・コントロール・プロトコル (TCP) レイヤの試験を省いていることがこのギャップの原因です。図1. 簡略化したプロトコルスタックと、ターンアップ試験手法とエンドユーザの期待との間のギャップ。

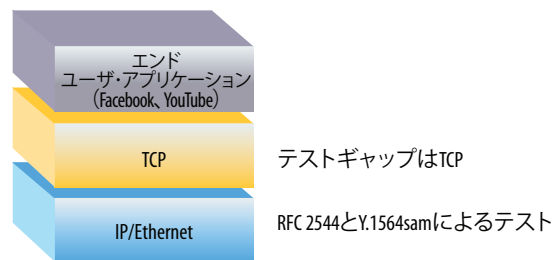


図1. 簡略化したプロトコルスタックと、ターンアップ試験手法とエンドユーザの期待との間のギャップ

このテスト・ギャップは、ネットワーク・プロバイダが自分たちのお客様のようにはネットワーク性能を知ることができないため、エンドユーザが使い始める前にTCPレイヤの性能を検証できるソリューションが必要です。TCPレイヤでの試験は、顧客先で再調整調整作業、技術サポート、顧客離れを無くし、プロバイダの運用費用 (OpEx) に非常によい影響を与えます。

このホワイトペーパーでは次のような点を解説します。

- TCPプロトコルの簡単な紹介
- TCPおよびアプリケーションの性能に悪影響を与える顧客設備 (CPE) やネットワーク上の典型的課題
- 新しいNETF RFC 6349 TCP試験手法の紹介
- RFC 6349ベースのTCPレイヤ設置試験を実施するネットワーク・プロバイダの利点

TCPに悪影響を与えるネットワークおよびCPEでの課題

TCPは、開放型システム間相互接続 (OSI) のレイヤ4で動作し、IP レイヤ3の上に位置します。TCPの重要な側面のひとつは、これが信頼できるということで、パケットが失われた場合にはTCPが受信側に再送することを保証します。

さらに、ワイドエリア・ネットワーク (WAN) リンク上では、TCPは送信側が、受信側からの確認 (ACK) を受信する前に送信できるバイト数の調整を適切に設定しなければなりません。この「途中を飛んでいる」バイト数はTCPウィンドウと呼ばれることがよくありますが、実際には複数のTCPウィンドウ機構が動作しています。

図2は、25ミリ秒の往復遅延 (RTD、レイテンシとも呼ばれます) を持つ45MbpsのWANリンク上を流れるTCPデータバイトの概念図です。

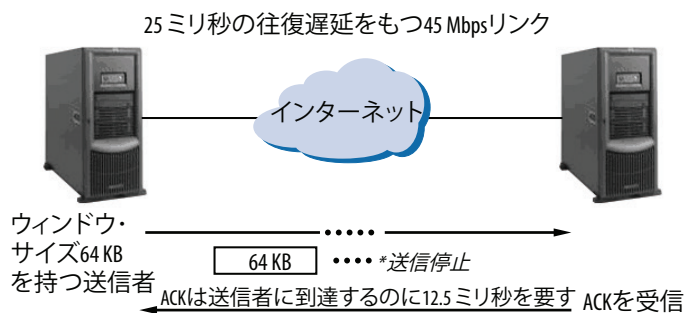


図2. 25ミリ秒のRTDをもつ45Mbps WANリンク上を流れるTCPデータバイトの概念図

図2.では、TCPウィンドウの動きがよくなく、ACKを要求する前に送信側から64kBしか送信されません。

RFC 6349に記載されているように、帯域幅遅延積 (BDP) は最適化されたTCPウィンドウで、以下のように算出されます。

$$\text{BDP} = \frac{\text{リンク・ボトルネック帯域幅} \times \text{往復時間}}{8}$$

この例では、BDPは 140 kBとなります。これは、送信側の 64 kBウィンドウの2倍以上のサイズで、送信者は約 20 Mbps スループットしか達成できません。

TCPのもうひとつの主要な属性として、ビット・レートが一定でなくバースト性があるということが言えます。そのため、100MbpsのWANにつながれたギガビットEthernet (GigE) のローカルエリア・ネットワーク (LAN) では、GigEのバーストをWANネットワークが適切に扱えないと起こります。これは、TCPのステートマシンが最適な転送レートまで上げなければならないからです (パケット損失がない範囲で)。

TCPがLANからWANに減速するための主要な手段はバッファリングとトラフィック・シェーピングです。図3.にネットワーク・バッファリングとネットワーク装置の出力キューを示してあります。出力キューは、DiffServ コードポイント (DSCP)、仮想LAN (VLAN) タグなどさまざまなサービス品質 (QoS) 機構に基づいてトラフィックを優先順位付けし、トラフィックごとにキューの深さを割り当てます。デフォルトのキューの深さをを用いるとバースト性の強いTCPトラフィックでロスが発生します。パケットロスが起きるとTCPの再送が行われるためエンドユーザに大きな悪影響が出ます。

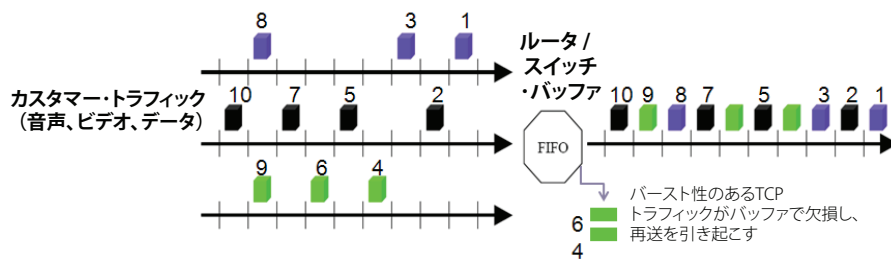


図3. ネットワークバッファリングとネットワーク装置の出力キュー

2つ目の点は、LANからWANへの減速のためのトラフィック・シェーピングや「インテリジェント」なネットワーク・バッファリングで、ここではネットワーク装置が認定情報速度 (CIR) に従ってトラフィックの流量を調節します。トラフィック・シェーピングはCPEのエッジ装置で行われるべきものですが、ネットワーク・プロバイダもトラフィックを調整でき、TCPの性能とエンドユーザの満足度を大幅に向上させることができます。

TCPトラフィック・シェーピングを行わないと、高速インターフェイスから低速側に減速する際にネットワーク・ポリサーがTCPの性能を大きく損なってしまいう可能性があります。シェーピングとは逆に、ポリシングはCIRを超えるトラフィックをカットしてしまい、TCPの再送が起こり、エンドユーザ側での性能に悪影響がでます。図4. トラフィック・シェーピングとポリサーの対比。

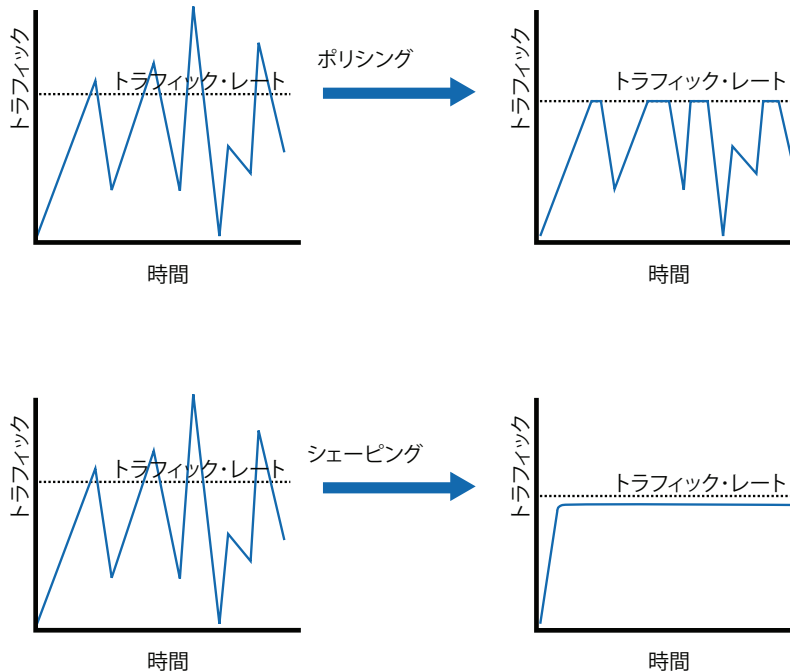


図4. トラフィック・シェーピングとポリサーの対比

お客様側でバッファリングへのキューイングとトラフィック・シェーピングを実施する前にRFC 6349テスト手法を用いることで、WAN越えのTCPスループットを最適化できます。

RFC 6349 TCPのテスト方法

RFC 6349は、Viavi がベル・カナダとドイツテレコムからの代表者たちと協同で創りだした、新しいTCPスループット試験手法です。インターネット技術タスクフォース (IETF) によって発行されたばかりのRFC 6349は、ネットワークとサービスの性能を最適化する、システムティックなプロセス、メトリック、ガイドラインを備えたTCPスループット解析のための繰り返しして利用できる試験手法です。

RFC 6349は、TCPテストを実施する前に、必ずレイヤ2/3 ターンアップ・テストを実施するように推奨しています。レイヤ2/3 でネットワークを検証後、RFC 6349では、以下の3段階のテストの実施を指定しています。

- **パスMTUの検出 (RFC 4821 に基づく)** は、TCPペイロードがフラグメントされないようにするために、アクティブなTCPセグメント・サイズを確認し、ネットワークの最大伝送ユニット (MTU) を検証します。
- **ベースライン往復遅延と帯域幅** は、TCP BDPを自動的に計算するための光TCPウィンドウ・サイズを予測するためのものです。
- **単一と複数のTCP接続スループット・テスト** は、自動「フルパイプ」TCPテストを可能にするTCPウィンドウ・サイズ予測を検証します。

TCPの再送はどのTCP/IPネットワーク通信においても一般的な現象です。再送回数の決定は、性能に影響を与えるため、その数字だけを使う場合は簡単ではありません。RFC 6349では、ペイロードの再送によって使われるネットワークの伝送の相対比率を考慮した新しいメトリックを規定しています。

このメトリックはTCP効率メトリックで、伝送されていないバイトの割合を示すもので、次のように定義されます。

$$\frac{\text{伝送済バイト} - \text{再送バイト}}{\text{伝送済バイト}} \times 100$$

伝送済バイトは、伝送されたTCPペイロードバイトの総数で、オリジナルおよび再送されたバイトを含みます。このメトリックは、トラフィック管理、輻輳回避、Windows XPやLinuxなどさまざまなTCPの実装でのQoS機構の間の比較が可能です。

たとえば、100,000バイトが送られ、2,000バイトが再送しなければならなかったとすると、TCP効率は次のように計算されます。

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

パケット・ロスが起きる分布次第でTCPの再送の状況は大きく変わるため、レイヤ2/3でのパケット・ロスの割合とバイト数での再送割合には相関がないことに注意してください。TCP効率メトリックによって、ネットワーク・プロバイダはさまざまなクラスオブサービス (CoS) レベルに対して単一のTCPロス境界値を設定できるようになります。

RFC 6349は、バッファ遅延率も定義します。これはTCPスループット試験中に往復時間(RTT) が、輻輳のない状態でネットワーク・パスが固有にもっているベースラインRTTからどれくらい増えるかを示します。

バッファ遅延率は、以下のよう定義されます。

$$\frac{\text{転送中の平均RTT} - \text{ベースラインRTT}}{\text{ベースラインRTT}} \times 100$$

たとえば、ベースラインRTT/パスが25ミリ秒で、平均RTT TCP伝送の間に32ミリ秒まで増加した時のネットワークのバッファ遅延率は次のような計算になります。

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

言い換えれば、TCP伝送が28パーセント増えたRTD (輻輳) の影響を受け、それに比例して全体的なTCPスループットが落ち、エンドユーザ・レベルでの遅延が長くなったわけです。

RFC 6349はRFC 2544を補完し、エンドユーザ・レベルでの状況とプロバイダがネットワークを試験する方法との間のギャップを埋めるものです。次の表は、RFC 6349ベースのTCP試験を実施するさらなるメリットであるRFC 2544とY.1564samのネットワーク試験の適用可能性を示したものです。

ターンアップに関連した問題	RFC 2544	Y.1564sam	RFC 6349
シングル・サービス、レイヤ2/3 ロスやジッタなどのSLA問題	X	X	N/A
マルチサービス、レイヤ2/3 サービス優先順位付、ロス、ジッタなどのSLA問題		X	N/A
エンド・ユーザの影響を示す スループットに関するTCPウィンドウ・サイズ (CPEの問題)			X
バースト性の強いユーザ・アプリケーションを扱うには不適な装置バッファ			X
ポーリングがTCPの性能に影響			X

役に立つ応用編 *Viavi TrueSpeed™* によるRFC 6349 試験では、この重要な新手法にまつわる詳細が提供され、IETFのこの新しいRFCの試験の実施がViaviによって自動化されています。

ビジネス・カスタマーが不満を感じた2つのケース

ネットワーク経由でのアプリケーションの性能が悪いというビジネス・カスタマーの不満には、典型的に2つの理由があります。

- ビジネス・カスタマーが顧客設備を正しく設定していないかまたは、誤ったスピード・テスト手順を実行しているからです。デフォルトのTCPウィンドウ・サイズのまま高帯域WANネットワークを使用すると、TCPの性能が大きく落ちます。誤ったスピード・テストの例は、ユーザがiperfのようなオープン・ソースのツールを用いて性能の低いPC上でTCP試験を実行していく場合で、SLAで定めたスループットを達成できなくなります。
- ネットワーク・プロバイダは、TCPバーストに対してポート・バッファサイズが不適切であるような、チューニングを必要とするネットワーク上の問題を経験することがあります。従来のレイヤ2/3試験では、バースト性のあるTCPトラフィックのような形でネットワークに負荷をかけることができません。バースト性の強いビジネス・カスタマーのトラフィックは、ポーリッシュ制御され、再送が発生し、性能が低下します。ポート・バッファサイズを増やすことにより(トラフィック・シェーピングの一形態)、パケットロスが激減し、TCPの性能を向上させることができます。

ふたつの実例をまとめると、次の項にあるような2つの結論を示すことができます。それぞれのケース・スタディでは、従来のレイヤ2/3が導入され、カスタマーが使用する前にTCプレイヤ試験が実施されていません。これらのふたつのケース・スタディで言えることは、TCP試験をカスタマーが使用する前に実施することでプロバイダは余計なトラブルシューティング・コストや顧客離れといった不利益を避けることができるということです。

CPE上の問題:最適化されていないTCP設定

このシナリオでは、ビジネス・カスタマーは2点間を結ぶ100Mbps透過型伝送LANサービスを、図5. に示したようなネットワーク構成でビジネス・プロバイダから購入しています。

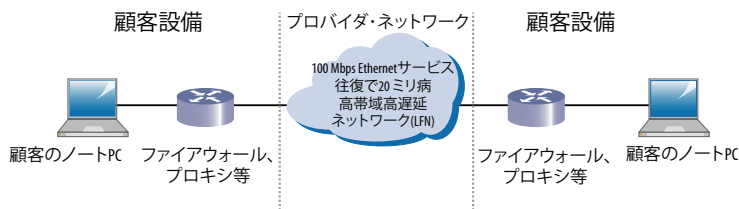


図5. ビジネスカスタマーの100Mbps透過型LAN設定

次の表、RFC 2544の結果と、ネットワーク・プロバイダが実施したFTPのダウンロードとアップロードによる顧客のスピード・テストを示します。

RFC 2544 の結果	FTPの結果
100 Mbps スループット	25 Mbps スループット

当然のことながらビジネス・カスタマーはFTPの性能には満足できず、提供されているサービスに疑問をもちます。プロバイダの技術者が何度も現地に行き、RFC 2544 試験を再実行し、互いに相手が悪いという結果になってしまいました。

このシナリオでの問題点は、一般に高帯域高遅延 (LFN, Long Fat Network) と呼ばれる20ミリ秒のレイテンシをもった100Mbpsネットワークで、エンド側ホスト・コンピュータのTCPウィンドウ設定はもっと大きなものが必要だったのです。このシナリオにおけるTCPウィンドウ設定の理想的な値は (RFC 6349で示されています)、次の式で計算できます。

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

従来のネットワーク・プロバイダにCPE装置に対する責任はないとはいえ、彼らは責任分界点でRFC 6349 TCP試験を実施すべきでした。RFC 6349では、さまざまなネットワーク条件でのガイドラインを最適なTCPウィンドウ・サイズと共に提供しており、ネットワーク・プロバイダがエンドユーザが実感するネットワークを最適化し、満足度とロイヤリティを高めるのに役立ちます。

このホワイトペーパーのために、ネットワーク・プロバイダは4つの異なったTCPウィンドウ・サイズでの試験を実施し、その結果を次の表で示します。

TCPウィンドウ・サイズ	スループット
32 kB	12.5 Mbps
64 kB	25 Mbps
128 kB	50 Mbps
256 kB	100 Mbps

この試験で明らかなのは、このネットワークが適切に動作していて、エンドユーザが自身のCPEホスト上のTCPウィンドウスケール・オプションを有効にする必要があったということです。エンドユーザは1台のサーバでウィンドウ・スケールを使用しておらず、最大ウィンドウ・サイズはわずか64kBでした (100Mbps CIRで25Mbpsしか出ません)。

最終的に我々が予測したプロバイダにとっての追加の運用費用は7,000ドル台でした。このプロバイダは現在、自社の標準のサービス実施手順に RFC 6349 試験手法を取り入れています。

ネットワーク・プロバイダの問題:不適切なネットワーク・バッファ

このシナリオでは、ビジネス・カスタマーは2地点間を結ぶ300Mbps透過型伝送LANサービスを、図6に示したようなネットワーク構成でビジネス・プロバイダから購入しています。2地点間のレイテンシは約50ミリ秒であることに注意してください。

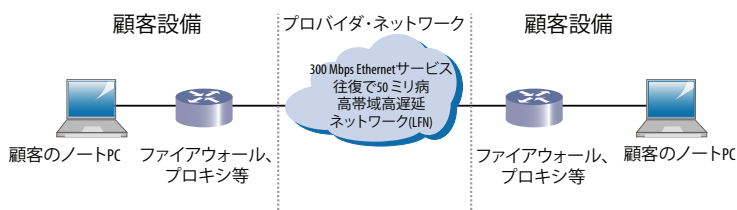


図6. ビジネス・カスタマーの300Mbps透過型LAN構成

次の表は、RFC 2544の結果と、ネットワーク・プロバイダが実施したiperfプログラムによる顧客のスピード・テストを示します。

RFC 2544 の結果	iperf 結果
300 Mbps スループット	95 Mbps スループット

この場合も、プロバイダの技術者がRFC 2544試験を繰り返し実行したターンアップ試験で互いの非を主張することになりました。エンド・ユーザは、プロバイダが示した結果が信頼できず、プロバイダにTCPLレイヤでの何らかのスピード・テストを実施するよう要求しました。

このシナリオでは、BDPがつぎのような計算で非常に大きくなっていました。

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

これは実際非常に大きなBDPですが、TCPの経験が長いエンド・ユーザはパイプをいっぱいにするためにTCPセッションを並行して走らせていました(この場合30)。

RFC 6349には並行TCP接続試験があり、これはBDPが非常に大きい場合の試験手法として重要なものです。プロバイダはViaviのTrueSpeedソリューションを使用し、ViaviのハードウェアTCPソリューションによってこの試験シナリオを再現することで、エンド・ユーザのワークステーションで動いているiperfプログラムの処理性能が試験結果に影響しているかどうかについて不確実性を排除できました。

このシナリオでは、プロバイダは実際にこの問題を再現し、エンド・ユーザが言ったように95Mbpsのスループットしか達成できませんでした。

プロバイダは問題がエッジ・ルーティング装置のデフォルト・バッファリングが不適切であることに気づきました。この装置のバッファ・サイズを増やすことで、プロバイダは再試験の結果正しい300MbpsのTCP性能を検証することができました。

最終的に我々が予測したプロバイダにとっての追加の運用費用は15,000ドル台でした。このプロバイダは現在、自社の標準のサービス実施手順に RFC 6349 試験手法を取り入れています。

結論

このホワイトペーパーで強調したことは、現在のEthernetサービスのアクティベーション手法であるRFC 2544ならびにY.1564と、エンド・ユーザ側で実際に経験することにギャップがあるということでした。お客様のアプリケーションはTCPレイヤの上で動いており、その性能はLANからWAN環境へと動いていく場合に大きな影響を受けます。

RFC 6349ベースの試験を取り入れたRFC 2544やY.1564によるサービス・アクティベーションが増加することで、プロバイダとおお客様の経験は同じになり、運用費用が大きく削減され、お客様の第一印象に大きな違いが生まれ、未永くおつきあいいただけることで企業の収益性も向上します。

Viaviは RFC 6349によるTCP試験を採用し、さらにそれを画期的な形で自動化しました。Viavi TrueSpeed試験機能は、業界初のRFC 6349準拠の試験ソリューションです。TrueSpeedにより、ネットワーク・プロバイダは、従来のRFC 2544による導入とTCP試験を一回の訪問で実施でき、しかもそれに必要なスキルレベルは変わらず、時間も5分以下で可能です。

役に立つ応用ノート *Viavi TrueSpeed™* によるRFC 6349 試験は、この新しいRFCを詳細に説明し、TrueSpeed試験機能を解説しています。



〒163-1107
東京都新宿区西新宿6-22-1
新宿スクエアタワー7F

電話: 03-5339-6886
ファックス: 03-5339-6889
Email: support.japan@viavisolutions.com

© 2015 Viavi Solutions Inc.
この文書に記載されている製品仕様および内容は予告
なく変更されることがあります
rfc6349-wp-tps-tm-ja
30173044 901 1214