

アプリケーション使用だけなぜ 10Mbps リンク 1Gbps ですか。

目次

[概要](#)

[背景説明](#)

[問題の外観](#)

[帯域幅遅延プロダクト](#)

[確認](#)

[解決策](#)

[2つの場所間の Round Trip Time \(RTT \) を言う方法か。](#)

概要

この資料は高速と関連する問題を高いレイテンシー ネットワーク記述したものです。それは BDP からある特定の状態の実質帯域幅の利用を計算するために数式を得ます。

背景説明

どんどん増えるように企業は持ち、または地理的に分散させたデータセンターの構築の過程においてあり、高速リンクによってデータセンターを相互接続します。よりよくの必要は帯域幅を増加しています利用します。

帯域幅遅延プロダクト (BDP) は数年幾年もの間インターネットで送達されました。ただし問題が見えるもののに、実世界例がありません。TCP ウィンドウ サイズの BDP 数式フォーカス。それ doesn か。t は私達に距離に基づいて可能性のある 帯域幅の利用を計算する方法を与えます。この資料は簡潔に BDP を説明したり、また、問題、また、解決を示したものです。この技術情報はまたある特定の状態の帯域幅の利用を計算するために数式を得ます。

問題の外観

会社は 2 データセンターがあります。1 データセンターからの別のデータセンターへの会社 バックアップ ビジネス 重大なデータ。バックアップ admin はそれらがネットワーク遅さによる Backup ウィンドウ内のバックアップを終えることができないことを報告しました。ネットワーク admin としてネットワーク遅さ問題を調査するために、割り当てられます。これらのファクタを知っています:

- これら二つのデータセンターは別 1000KM です。
- これらのデータセンターは 1Gbps リンクによって相互接続されます。

調査に、注意しました:

- 利用可能な帯域幅が十分あります。
- ネットワークハードウェアまたはソフトウェア上の問題がありません。
- バックアップアプリケーションは 10Mbps 帯域幅のまわりだけで利用します 990Mbps 帯域幅の他にによってが自由である。
- バックアップアプリケーションはデータを転送するのに TCP を使用します。

帯域幅遅延プロダクト

バックアップアプリケーション使用だけ 10Mbps の質問に答えるために、それは帯域幅遅延プロダクト (BDP) をもたらします。

BDP はそれを単に示します:

$BDP \text{ (ビット)} = \text{total_available_bandwidth (ビット/秒)} \times \text{round_trip_time (秒)}$

または $RWIN/BDP$ がバイトに通常ある、およびレイテンシー ミリ秒にので測定されます:

$BDP \text{ (バイト)} = \text{total_available_bandwidth (Kバイト/秒)} \times \text{round_trip_time (ms)}$

サーバが受け取り パケットの確認応答を停止し、待っている前にどの位データが転送することができるか判別するこれは TCP ウィンドウがバッファであることを意味します。スループットは BDP によって本質的に結合されます。BDP (か RWIN) レイテンシーおよび利用可能な帯域幅の製品より下部のである場合、クライアントが十分に速く確認応答を送信できないので行を一杯にできません。伝達はそう TCP ウィンドウ (RWIN) 十分に大きい必要があります $\text{maximum_anticipaded_delay} \times \text{maximum_available_bandwidth} \times$ に合うには (RWIN/レイテンシー) 値を超過できません。

上の数式を使って。得られた帯域幅 計算式は次のとおりです:

帯域幅の利用 (キロビット/秒) = $BDP \text{ (bytes)} / RTT \text{ (ms)} * 8$

注: この数式は最大理論的な帯域幅の利用を計算します。それ doesn't が。t は OS を奪取しますか。多くのファクタによって含まれる例えば利用可能なメモリ、NIC ドライバ、ローカル NIC 速度、キャッシュまたは時々ディスク 速度があるので考慮事項への s パケット伝送時間。その結果、TCP ウィンドウ サイズが大きいとき、計算された帯域幅は実際の帯域幅より大きいです。TCP ウィンドウ サイズが非常に大きいとき、偏差は同様に大きい場合もあります。

得られた数式を使うと、答えるバックアップアプリケーションが計算かの下ですることによってしか 10Mbps をなぜの使用できないか質問にすることができます。

- 一般に、1000KM のための RTT は ~15 です。つまり $RTT=15ms$
- デフォルトで、Windows 2003 オペレーティング システム Windows サイズは 17,520 バイトです。つまり $BDP=17,520$ バイト
- 数式にこれらの数を入れて下さい:

帯域幅の利用 (キロビット/秒) = $17520/15*8$ 。

結果は 9344Kbps または 9.344Mbps です。 TCP および IP ヘッダーと 9.344Mbps。 最終結果は ~10Mbps です。

確認

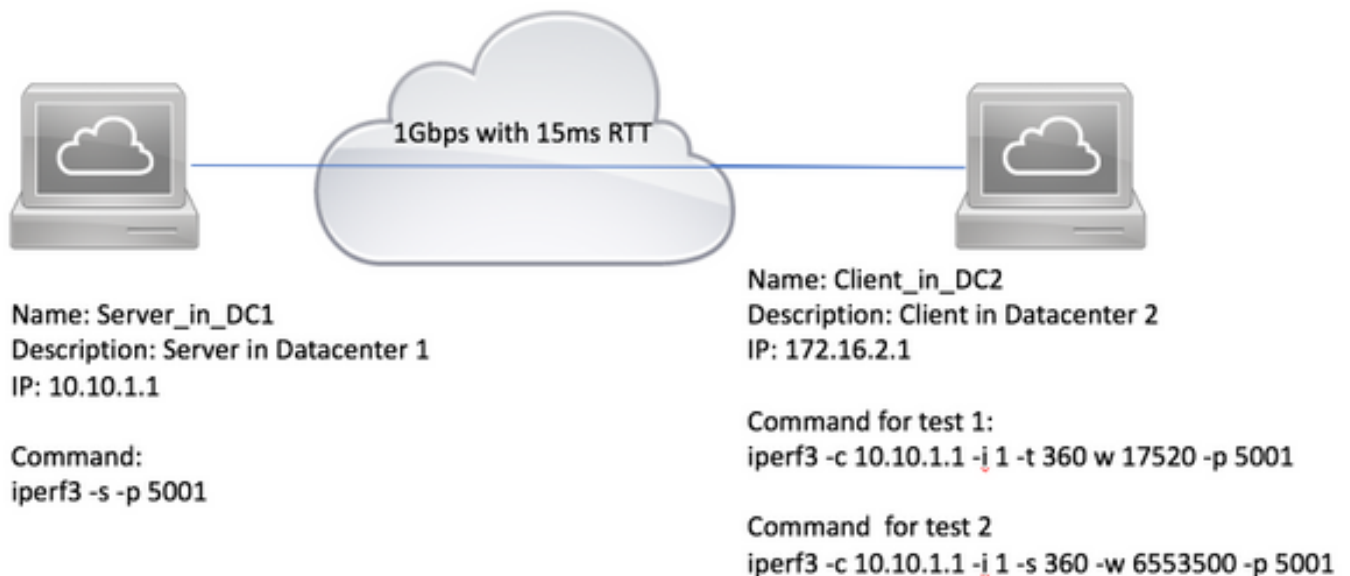
ネットワーク admin として、論理上質問に答えました。 この場合現実の 世界の理論を確認する必要があります。

理論を確認するのにネットワークパフォーマンス テスト ツールを使用できます。 問題および解決を示すために **iperf** を実行することにしました。

これはラボの セットアップです:

1. IP アドレス 10.10.1.1 が付いている datacenter 1 のサーバ。
2. IP アドレス 172.16.2.1 が付いている datacenter 2 のクライアント。

トポロジーはイメージに示すようにあります:



確認するために次の手順に従って下さい:

1. **iperf3** を **-s** -それにサーバをし、TCPポート 5001 で受信する 10.10.1.1 の **p 5001** 実行して下さい。
2. デフォルト TCP ウィンドウとテストするために 17,520 バイトを大きさで分類して下さい。 それにクライアントをするために **iperf3** を **-c 10.10.1.1 -i 1 -t 360 -w 17520 -p 5001** 172.16.2.1 の **p 5001** 実行して下さい。 このコマンドは iperf をポート 5001 のサーバに接続するように告げます 360 秒の間動作し、帯域幅の利用に TCP ウィンドウ サイズの各 1秒を 17,520 バイト報告します。
3. 例えば 6,553,500 バイトをカスタマイズされた TCP ウィンドウ サイズとテストするために、**iperf3-c 10.10.1.1** を **-i 1 -t 360 -w 6553500 -p 5001** 実行して下さい

これはデフォルト TCP ウィンドウ サイズのラボ テスト結果 17,520 バイトです。 帯域幅の利用

が ~10Mbps であることを見ることができます。

```
C:\Tools>iperf3.exe -c 10.10.1.1 -t 360 -p 5001 -i 1 -w 17520Connecting to host 10.10.1.1,
port 5001[ 4] local 172.16.2.1 port 49650 connected to 10.10.1.1 port 5001[ ID] Interval
Transfer      Bandwidth[ ID] Interval      Transfer      Bandwidth[ 4] 0.00-1.00 sec
1.30 MBytes  10.9 Mbits/sec[ 4] 1.00-2.02 sec  919 KBytes  7.41 Mbits/sec[ 4] 2.02-3.02
sec 1.28 MBytes  10.7 Mbits/sec[ 4] 3.02-4.02 sec  1.14 MBytes  9.59 Mbits/sec[ 4]
4.02-5.01 sec 1.24 MBytes  10.4 Mbits/sec[ 4] 5.01-6.01 sec  1.33 MBytes  11.3
Mbits/sec[ 4] 6.01-7.01 sec 1.15 MBytes  9.65 Mbits/sec[ 4] 7.01-8.01 sec  1.12
MBytes  9.36 Mbits/sec[ 4] 8.01-9.01 sec  1.22 MBytes  10.3 Mbits/sec[ 4] 9.01-10.01
sec 1.13 MBytes  9.49 Mbits/sec[ 4] 10.01-11.01 sec  1.30 MBytes  10.8 Mbits/sec[ 4]
11.01-12.01 sec 1.17 MBytes  9.84 Mbits/sec[ 4] 12.01-13.01 sec  1.13 MBytes  9.48
Mbits/sec[ 4] 13.01-14.01 sec 1.28 MBytes  10.7 Mbits/sec[ 4] 14.01-15.01 sec  1.40
MBytes  11.8 Mbits/sec[ 4] 15.01-16.01 sec  1.24 MBytes  10.4 Mbits/sec[ 4] 16.01-17.01
sec 1.30 MBytes  10.9 Mbits/sec[ 4] 17.01-18.01 sec  1.17 MBytes  9.78 Mbits/sec
```

これは TCP ウィンドウ サイズのラボ テスト結果 6,553,500 バイトです。帯域幅の利用が ~200Mbps であることを見るすることができます。

```
C:\Tools>iperf3.exe -c 10.10.1.1 -t 360 -p 5001 -i 1 -w 6553500Connecting to host 10.10.1.1,
port 5001[ 4] local 172.16.2.1 port 61492 connected to 10.10.1.1 port 5001[ ID] Interval
Transfer      Bandwidth[ 4] 0.00-1.00 sec 29.1 MBytes  244 Mbits/sec[ 4] 1.00-2.00
sec 25.4 MBytes  213 Mbits/sec[ 4] 2.00-3.00 sec 26.9 MBytes  226 Mbits/sec[ 4]
3.00-4.00 sec 18.2 MBytes  152 Mbits/sec[ 4] 4.00-5.00 sec 25.8 MBytes  217
Mbits/sec[ 4] 5.00-6.00 sec 28.8 MBytes  241 Mbits/sec[ 4] 6.00-7.00 sec 26.1
MBytes  219 Mbits/sec[ 4] 7.00-8.00 sec 21.1 MBytes  177 Mbits/sec[ 4] 8.00-9.00
sec 22.5 MBytes  189 Mbits/sec[ 4] 9.00-9.42 sec  9.54 MBytes  190 Mbits/sec
```

解決策

ソフトウェア 開発観点から、複数の同時 TCP 脱退を実行する複数のスレッディングは帯域幅の利用を改善できます。ただしソースコードを修正するために、それはネットワークか System Admin のために実用的ではないです。ことができる何をする OS を最適化することです。

RFC1323 は高いパフォーマンス TCP のための複数の TCP 拡張を定義します。これらはウィンドウスケール オプションおよび選択的な ACK が含まれています。それらは主要なオペレーティングシステムによって実装されています。ただし、デフォルトで、OS はそれらをディセーブルにしますそれらをサポートするために TCP/IPスタックが書かれている。

- これらの OS ディセーブル RFC1323 デフォルトで: 先のカーネルとの Windows 2000、Windows 2003、Windows XP および Linux より 2.6.8。

Microsoft Windows システムの問題に直面する場合、TCP を最適化するためにこのリンクに従って下さい。 <https://support.microsoft.com/en-au/kb/224829>。

他の OS に関しては、ベンダーを参照して下さいか。それらを設定する方法の s ドキュメント

。

- これらの OS イネーブル RFC1323 デフォルトで: Windows 2008 およびそれ以降、Windows Vista およびそれ以降、カーネル 2.6.8 およびそれ以降の Linux。これらの機能を改善するためにパッチを加える必要がある場合もあります。ある状況では、それらをディセーブルにすることを望みます。ベンダーを参照して下さいか。それらをディセーブルにする方法の s ドキュメント。

- いくつかのアプリケーションは Microsoft Windows 2000、Windows 2003 または組み込まれたオペレーティングシステムの上に構築されます。例えば NAS、ヘルスケア ハードウェア。RFC1323 が有効になるかどうか確かめるためにベンダーのドキュメントをチェックして下さい。

2つの場所間の Round Trip Time (RTT) を言う方法か。

一般に、RTT は距離と関連付けられます。下記の表は距離および関連した RTT をリストしたものです。また正常なネットワークの状態の RTT の概念を得るのに ping テストを使用できます。

距離 (KM)	RTT (ms)
1,000	15
4,000	50
8,000	120

注: ガイドだけは、実質 RTT 時間ある場合もあります変わります上であります。また、レイテンシーは使用されるテクノロジーによって影響を与られます。たとえば、3G レイテンシーは関係ない頻繁に 100ms である場合もあります距離。それはサテライトに同様にあてはめます。