

アプリケーションが10Mbpsしか使用しない理由 リンクが1Gbpsであっても？

内容

[概要](#)

[背景説明](#)

[問題の概要](#)

[帯域幅遅延製品](#)

[確認](#)

[解決方法](#)

[2つのロケーション間のラウンドトリップ時間\(RTT\)を通知する方法](#)

概要

このドキュメントでは、高速、高遅延ネットワークに関連する問題について説明します。特定の条件での実際の帯域幅使用量を計算する式をBDPから導出します。

背景説明

地理的に分散したデータセンターを構築し、高速リンクを介してデータセンターを相互接続する企業が増加している、または現在進行中の企業です。帯域幅をより有効に活用するニーズが高まっています。

Bandwidth-Delay Product(BDP)は数年間インターネット上で公開されています。ただし、問題の実際の例はありません。BDP式は、TCPウィンドウサイズに焦点を当てています。距離に基づいて帯域幅の使用量を計算する方法はありません。このドキュメントでは、BDPについて簡単に説明し、問題と解決策を示します。また、この記事では、特定の条件での帯域幅使用量を計算する式を導出しています。

問題の概要

貴社には2つのデータセンターがあります。企業は、あるデータセンターから別のデータセンターにビジネスクリティカルなデータをバックアップします。バックアップ管理者から、ネットワークの速度低下により、バックアップ時間内にバックアップを完了できないことが報告されました。ネットワーク管理者として、ネットワークの速度低下の問題を調査するように割り当てられます。次の要因が分かります。

- これら2つのデータセンターは1000KM離れています。
- これらのデータセンターは1 Gbpsリンクで相互接続されています。

調査の結果、次のことに気付きました。

- 使用可能な帯域幅が十分にある。
- ネットワークハードウェアやソフトウェアの問題はありません。
- バックアップアプリケーションは、残りの990Mbps帯域幅が空いていても、約10Mbpsの帯域幅しか使用しません。
- バックアップアプリケーションはTCPを使用してデータを転送します。

帯域幅遅延製品

バックアップアプリケーションの質問に答えるには、10Mbpsのみを使用するため、Bandwidth-Delay Product(BDP)が導入されています。

BDPは単に次のように述べています。

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

または、RWIN/BDPは通常バイト単位で、遅延はミリ秒単位で測定されます。

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

つまり、TCPウィンドウは、サーバが停止して受信パケットの確認応答を待つ前に転送できるデータ量を決定するバッファです。スループットは本質的にBDPによって制限されます。BDP (またはRWIN) が遅延と使用可能な帯域幅の積よりも低い場合、クライアントは確認応答を十分な速さで返信できないため、回線を埋めることはできません。伝送は (RWIN/遅延) の値を超えることはできません。したがって、TCPウィンドウ(RWIN)は、 $\text{maximum_available_bandwidth} \times \text{maximum_predicated_delay}$ に収まるほど大きくする必要があります。

上の式で。算出された帯域幅計算式は次のとおりです。

帯域幅使用量(Kbps)=BDP (バイト) /RTT (ミリ秒) * 8

注：この式では、理論上の最大帯域幅使用量が計算されます。使用可能なメモリ、NICドライバ、ローカルNIC速度、キャッシュ、時にはディスク速度など、多くの要因が関係するため、OSの packets 送信時間を考慮しません。その結果、TCPウィンドウサイズが大きい場合、計算された帯域幅は実際の帯域幅よりも大きくなります。TCPウィンドウのサイズが非常に大きい場合、偏差も大きくなる可能性があります。

派生式を使用すると、バックアップアプリケーションが10Mbpsしか使用できない理由について質問に回答できます。次の計算を行います。

- 一般に、1000KMのRTTは約15です。したがって、RTT=15ミリ秒になります
- デフォルトでは、Windows 2003オペレーティングシステムのWindowsサイズは17,520バイトです。したがって、BDP=17,520バイト
- 数式に次の数値を入力します。

帯域幅使用量(Kbps)= 17520/15*8。

結果は9344Kbpsまたは9.344Mbpsです。9.344MbpsにTCPおよびIPヘッダーを加える。最終結果は最大10Mbpsです。

確認

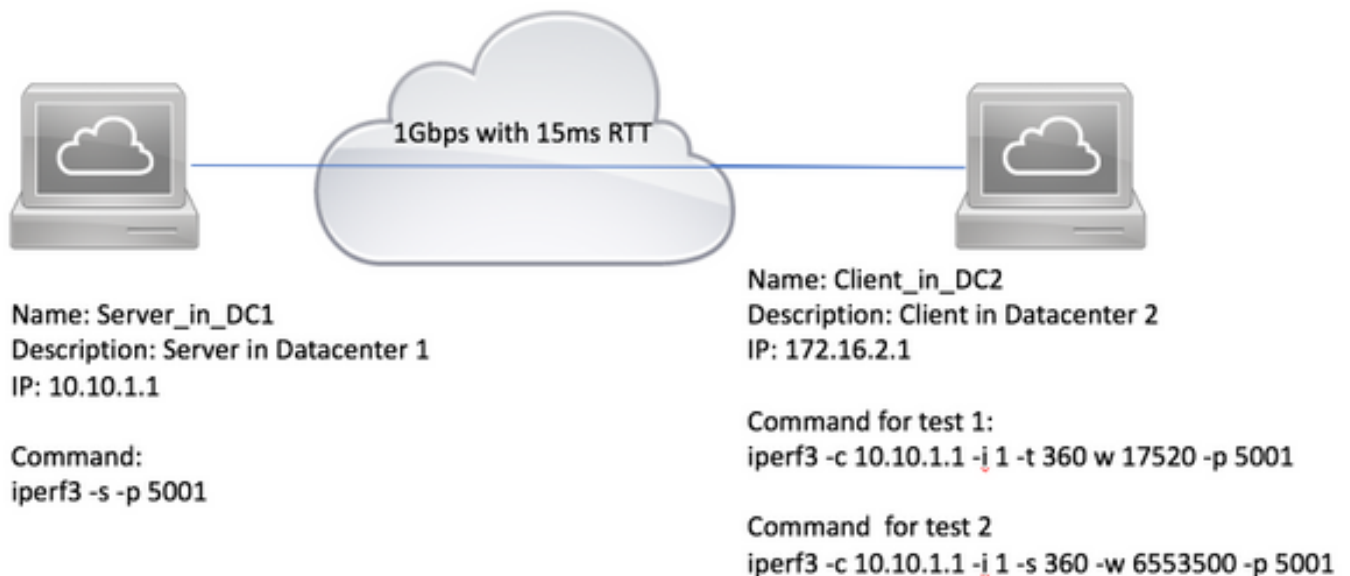
ネットワーク管理者として、この質問に理論的に回答しました。次に、実社会の理論を確認する必要があります。

ネットワークパフォーマンスのテストツールを使用して、理論を確認できます。iperfを実行して問題と解決を実証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のデータセンター1のサーバ。
2. IPアドレス172.16.2.1のデータセンター2のクライアント。

トポロジは次の図のようになります。



次の手順に従って確認してください。

1. 10.10.1.1でiperf3 -s -p 5001を実行し、サーバとして動作させ、TCPポート5001でリッスンします。
2. デフォルトのTCPウィンドウサイズ17,520バイトでテストします。iperf3 -c 10.10.1.1 -i 1 -t 360 -w 17520 -p 5001を172.16.2.1で実行し、クライアントにします。このコマンドは、iperfにポート5001上のサーバに接続し、360秒間実行し、TCPウィンドウサイズ17,520バイトで1秒ごとに帯域幅使用量を報告するように指示します。
3. カスタマイズされたTCPウィンドウサイズ (6,553,500バイトなど) でテストするには、
iperf3 -c 10.10.1.1 -i 1 -t 360 -w 6553500 -p 5001を実行します

これは、デフォルトのTCPウィンドウサイズが17,520バイトのラボテスト結果です。帯域幅の使用量は最大10Mbpsであることがわかります。

```
C:\Tools>iperf3.exe -c 10.10.1.1 -t 360 -p 5001 -i 1 -w 17520
```

```
Connecting 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
[4]	0.00-1.00	sec	1.30 MBytes	10.9 Mb/s
[4]	1.00-2.02	sec	919 KBytes	7.41 Mb/s
[4]	2.02-3.02	sec	1.28 MBytes	10.7 Mb/s
[4]	3.02-4.02	sec	1.14 MBytes	9.59 Mb/s
[4]	4.02-5.01	sec	1.24 MBytes	10.4 Mb/s
[4]	5.01-6.01	sec	1.33 MBytes	11.3 Mb/s
[4]	6.01-7.01	sec	1.15 MBytes	9.65 Mb/s
[4]	7.01-8.01	sec	1.12 MBytes	9.36 Mb/s
[4]	8.01-9.01	sec	1.22 MBytes	10.3 Mb/s
[4]	9.01-10.01	sec	1.13 MBytes	9.49 Mb/s
[4]	10.01-11.01	sec	1.30 MBytes	10.8 Mb/s
[4]	11.01-12.01	sec	1.17 MBytes	9.84 Mb/s
[4]	12.01-13.01	sec	1.13 MBytes	9.48 Mb/s
[4]	13.01-14.01	sec	1.28 MBytes	10.7 Mb/s
[4]	14.01-15.01	sec	1.40 MBytes	11.8 Mb/s
[4]	15.01-16.01	sec	1.24 MBytes	10.4 Mb/s
[4]	16.01-17.01	sec	1.30 MBytes	10.9 Mb/s
[4]	17.01-18.01	sec	1.17 MBytes	9.78 Mb/s

これは、TCPウィンドウサイズ6,553,500バイトのラボテスト結果です。帯域幅の使用量は最大200 Mbpsであることがわかります。

```
C:\Tools>iperf3.exe -c 10.10.1.1 -t 360 -p 5001 -i 1 -w 6553500
```

```
Connecting 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 Mb/s

[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

解決方法

ソフトウェア開発の観点から、複数の同時TCPセッションを実行するマルチスレッディングを使用すると、帯域幅の使用量を改善できます。ただし、ネットワーク管理者やシステム管理者がソースコードを変更することは現実的ではありません。OSの微調整が可能です。

RFC1323では、高性能TCP用の複数のTCP拡張が定義されています。これには、ウィンドウスケールオプションと選択的ACKが含まれます。これらはメインのオペレーティングシステムによって実装されます。ただし、デフォルトでは、TCP/IPスタックがサポートするように書き込まれていても、一部のOSは無効にします。

- 次のOSは、デフォルトでRFC1323を無効にします。2.6.8より前のカーネルを搭載したWindows 2000、Windows 2003、Windows XP、およびLinux。

Microsoft Windowsシステムで問題が発生する場合は、このリンクに従ってTCPを微調整してください。 <https://support.microsoft.com/en-au/kb/224829>。

その他のOSについては、ベンダーのドキュメントを参照して設定方法を確認してください。

- 次のOSは、デフォルトでRFC1323を有効にします。Windows 2008以降、Windows Vista以降、カーネル2.6.8以降のLinux。これらの機能を改善するには、パッチを適用する必要があります。状況によっては、これらは無効にすることが望ましい。無効にする方法については、ベンダーのマニュアルを参照してください。
- 一部のアプライアンスは、Microsoft Windows 2000、Windows 2003または組み込みオペレーティングシステムの上に構築されています。例：NAS、医療ハードウェアRFC1323が有効になっているかどうかをベンダーのドキュメントで確認してください。

2つのロケーション間のラウンドトリップ時間(RTT)を通知する方法

一般に、RTTは距離に関連付けられています。次の表に、距離と関連するRTTを示します。また、ping testを使用して、通常のネットワーク状態でRTTに関する情報を取得することもできます

。

距離(KM) RTT(ms)

1,000 15

4,000 50

8,000 120

注：上記はガイドであり、実際のRTT時間は異なることがあります。また、遅延は使用されるテクノロジーの影響を受けます。たとえば、3G遅延は、距離に関係なく100ミリ秒に頻繁に発生します。それはサテライトにも当たる。