

インターネットと計測技術 (第2回)

長 健二郎

株式会社インターネットイニシアティブ 技術研究所

1 はじめに

今回は ping コマンドを使った計測について話をします。ping コマンドは、他のノードへの到達性を調べるツールです。おそらく ping はネットワークコマンドの中で一番よく知られているコマンドでしょう。ping コマンドは、ほぼ全ての IP 機器に実装されていて、ほとんどのエンジニアは、ネットワークに機器を繋ぐと最初に ping コマンドで接続ができていないか確認するのではないのでしょうか。

ping は、もともと管理用のツールで、計測ツールではありません。したがって、計測用に十分な精度を持ってはいませんが、使い方によっては計測への応用も可能です。

2 ping コマンド

ping というコマンド名は、潜水艦が反射によって障害物を検知するために出すピーンというソナー音から来ているそうです。ping コマンドも相手ノードにメッセージを送り、相手からエコーされる応答メッセージを受信する事で到達性を確認します。メッセージには、ICMP エコーリクエスト/リプライを使います。ICMP エコーリクエストメッセージを受信したノードは、送信ノードに ICMP エコーリプライメッセージを返さず決まりになっています。送信ノードは、リプライが戻ってくるまでの時間から、おおよそのラウンドトリップタイムを知る事ができます。

ping で他のノードとの到達性が確認できれば、ネットワークの基本機能は動作していると言えます。しかし、ping が問題なく通るからといって、ネットワークが正常に動作している保証にはなりません。例えば、最大パケットサイズ (MTU) の不整合が発生している場合でも、小さいパケットは問題なく通るので、ping

は通るがウェブサイトは見られないというような現象が起こります。

逆に、世の中には、ICMP をフィルタするように設定されたファイアウォールも存在し、ping は通らないが、TCP なら通信が出来るというような場合もあります。本来、ICMP はネットワーク運用や Path MTU 発見のために不可欠なので、ICMP をフィルタするとさまざまな弊害がでます。このような設定が広がった要因として、増加する外部からのスキャンに過剰反応した一部の運用者の存在や、一時期 ICMP エコーリクエストを偽装する smurf 攻撃が頻発した事などが上げられます。

3 ラウンドトリップタイム測定

ping コマンドは、ラウンドトリップタイムを手軽に計測する手段として広く使われています。しかし、ping コマンドで得られるラウンドトリップタイムは必ずしも正確とはいえません。高機能のルータでは、TCP 等の通信はハードウェア処理で転送する機能を持ちますが、ICMP パケットは CPU 処理となるので処理時間が余計にかかります。また、ICMP にはレート制限がかかっている場合もあります。したがって、ICMP パケットのラウンドトリップタイムは、通常パケットのラウンドトリップタイムとは異なる可能性があります。さらに、ping コマンドは通常アプリケーションとして実装されているので、そこで測った時間はプロセススケジューリング等の影響を受け、精度に限界があります。これについては、負荷の軽いホストで測定を実行することで、ある程度影響を軽減できます。

このように測定値が誤差を含んでいる事に加え、測定値の分布についても注意を払う必要があります。一般にネットワークの測定データの分布には大きな偏りがあります。測定値が正規分布しないので、平均や標

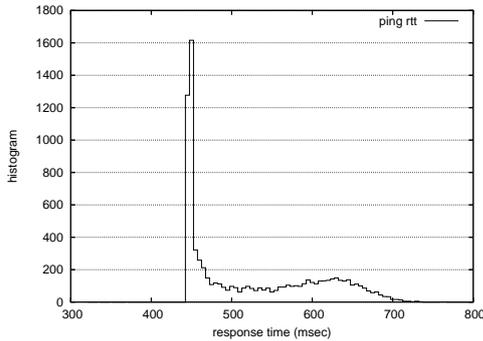


図 1: ping による RTT 測定値の分布

準偏差だけでは十分な情報は得られません。一般に、ラウンドトリップタイムの測定値は、通信路の伝搬遅延に中間ルータでの最小処理時間を加算したあたりに多くの測定値が集まり、それより大きい所にキュー待ちやその他の処理の遅延が加わった測定値が観測されます。しかし、実際に測定をしてみると、あり得ないような小さい値や、とんでもなく大きい値が観測される事があります。

ここでは、実際に ping コマンドを使って測定したラウンドトリップタイムを例に話をします。以下のデータは、国内のあるホストから東南アジアにあるウェブサーバに対して 10 秒間隔で 24 時間測定を行なった結果です。パーセンタイルは、データを小さい順にソートして、その位置を小さい方からのパーセントで示すものです。50th-percentile は、中間値 (median) にあたります。

エコーリクエスト送信:	8640
エコーリプライ受信:	8241
平均 RTT:	526 ms
最小 RTT:	322 ms
最大 RTT:	1852 ms
10th-percentile:	450 ms
50th-percentile:	488 ms
90th-percentile:	648 ms
ロス率:	4.6 %

図 1 は、ラウンドトリップタイムの測定値をヒストグラム表示したものです。多くの測定値 (全体の約 35%) は、445-455ms の間にあります。一方で、これより大きな値も多く観測されているため、単純に平均値をとると 526ms となってしまいます。このように偏った分布を持つ場合、特に分布から大きく外れた極端な値が含まれるような場合には、平均値は必ずしも適切な指標とはなりません。この場合、目的にもよりますが、

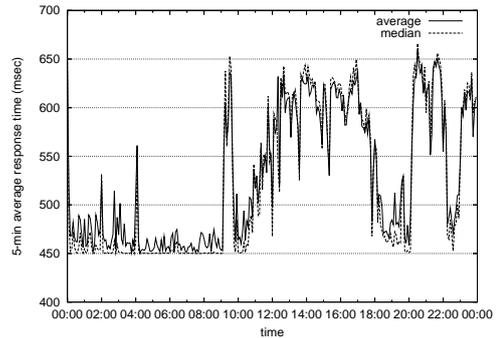


図 2: RTT の平均値、中間値の時系列変化

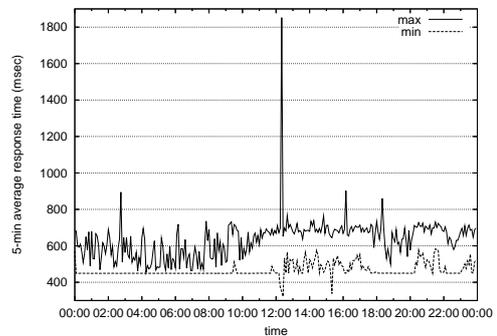


図 3: RTT の最小値、最大値の時系列変化

中間値を使う方がより適切です。また、ごく一部の極端な観測値の影響を除くためには、最大値と最小値の代わりに 10%と 90%というようなパーセンタイル値を使う事も有効です。一方で、平均値には加算則が成り立つので、5 分毎の平均値を平均して 1 日の平均値を求める事ができるという利点があります。中間値やその他のパーセンタイル値ではそのような事はできません。

図 2 は同じデータから時間別の平均値と中間値を 5 分間隔でプロットしたものです。図 3 は同様に最大値、最小値を示しています。まず時間帯によってラウンドトリップタイムが増大している、時間によって通信路のどこかが輻輳していると思われます。また、12:20 頃に最大値、最小値が大きく変動している事が分かります。特に、最小値が落ち込んでいる事から、経路変動が起こって一時的により短い経路が選択された後、再び元の経路に戻ったと考えられます。測定した経路は通常 US まわりで東南アジアに戻って来ているので、

このような現象が起こったようです。

このように、ラウンドトリップタイムを測る場合、単に平均値だけでなく、最大値や最小値を含めて、時系列で比較して変化を見る事によってより多くの情報が得られます。

4 パケットロス率の測定

ping コマンドでパケットロス率を測定する事もできますが、これは誤解を招きやすい測定値です。もちろん、大雑把にパケットロスが起こっている事を確認するためには、ping コマンドで十分ですが、ロス率を正確に測る事は意外に難しい作業です。

前述のように、高性能ルータでは、ICMP パケットは別処理される場合があり、ping での測定では、実際より高い確率でパケットロスが起きる可能性があります。逆に、ping は小さいパケットを一定間隔で送るだけなので、大きいパケットのロスや、バースト的なトラフィックによって発生するロスを過小に評価する可能性もあります。

そもそも通信路のパケットロス率を正確に測定するのは容易ではありません。通常ロス率は、往路と復路では異なり、正確に測定するには、通信路の両端で全てのパケットを観測して比較する必要があります。

つまり、ping を使ってパケットロス率何%という具体的な値にはあまり精度がないという事です。しかし、ping によりパケットロス率を定常的にモニタして、時系列変化を観測する事には意味があります。

5 おわりに

ping は測定ツールではありません。それゆえ、ping 測定では十分な精度は得られません。その一方で、その制約を理解して利用すれば、いろいろ応用の可能性があります。ここでは、ping を例として取り上げましたが、他のツールに関しても同様の事が言えます。