

WWW 環境における遅延要因の分析*

4Aa-6

久田 裕介†

NTT ソフトウェア研究所‡

1995年1月16日

1 はじめに

HTML(HyperText Markup Language)の機能拡張や ImageMap 技術によって、画像を多用したページのデザインが可能になった。この結果 WWW(World-Wide Web)はさまざまな分野で広く使われるようになった。その一方で、転送するデータ容量が増えるために輻輳が発生する可能性も高くなっている。従って、WWW 環境での輻輳回避技術を確立することが必要である。

このために、本稿では WWW 環境での輻輳の要因を分析する。特に、輻輳が発生すると Ack 送信数が増加する現象に注目する。Ack 送信数が増えると輻輳をさらに深刻にするためである。送信数を増加させる要因はセグメントの再転送と TCP の遅延 Ack 機能である。

2 遅延 Ack (Delayed Ack)

シー・ウィンドウ回避のために、受信側 TCP は受信バッファにある程度の空き容量ができるまで送信側 TCP へのウィンドウ告知を遅らせる。このために用いられるアルゴリズムの1つが遅延 Ack である。遅延 Ack では最大 500 ミリ秒(多くの実装では 200 ミリ秒)まで Ack の送信を遅らせる。さらに、タイマがエクスパイアするまでにフルサイズのセグメントを順序番号通りに 2 個受信した場合には、バッファあふれを防ぐために Ack を送信する [1]。このアルゴリズムによって Ack 送信数を 1/2 までに削減できるため、大容量のデータを転送する際には、輻輳を回避する効果がある。

しかしながら、実装が適切でなければ Ack 削減効果が十分に得ることはできない。例えば、タイマ設定値がセグメントの平均的な時間間隔や、フルサイズのセグメント 2 個分のバッファ空き容量を確保する時間に比べて短い場合には、Ack 送信数を削減することができない。また、タイマ設定値が長い場合には送信側 TCP がセグメントを誤って再転送するケースが増える [2][3]。つまり、遅延 Ack の実装要件は WWW が爆発的に流行する以前の仕様のため現在の Internet 環境への適合性を検証する必要がある。

3 観測

WWW の典型的な使用条件(遅延 Ack のタイマ設定値 200 ミリ秒、受信ウィンドウサイズ 4 キロバイト、最大セグメント長 536 バイト)で、日米間のリソース転送処理に

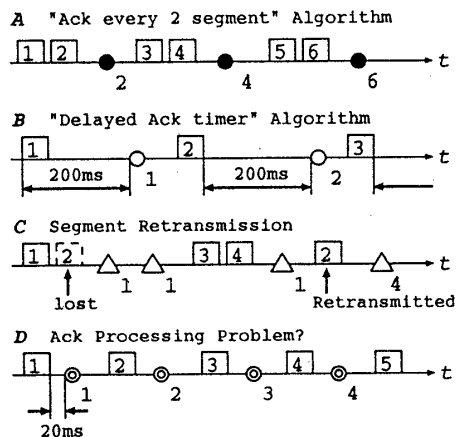


図 1: Ack 送信のタイミング

における遅延 Ack の動作を tcpdump3.1.2 で観測する。クライアントホストの OS は SunOS4.1.3 である。

3.1 Ack 送信のタイミング

受信側 TCP が Ack を送信したタイミングには 4 種類のパターンがある(図 1)。パターン A では 2 個のセグメントを順序通りに受信してから Ack (●) が送信されたために Ack 送信数が削減されている。輻輳の程度と関係なく 2 個のセグメント間の時間間隔が 30 ミリ秒未満の事例が 90% 以上を占め、100 ミリ秒を超えた事例は高々 4% である。つまり、パターン A では タイマ処理を必要とした事例は少なく、ほとんどの事例で RFC1122 で追加されたアルゴリズムによって送信されている。

一方、パターン B では タイマがエクスパイアしたために Ack (○) が送信されている。パターン C ではセグメント損失のために遅延 Ack が動作していない(△)。パターン D では、セグメントを 1 個受信してからタイマがエクスパイアする前に Ack (◎) が送信されている。つまり、パターン B ~ D では RFC1122 で追加されたアルゴリズムを用いて Ack 送信数を削減することはできていない。

3.2 Ack 削減効果

231 キロバイトと 85 キロバイトのリソースはそれぞれ、453 個と 167 個のセグメントで送信される。Ack 送信数を最も削減できるケース、つまり、最後の奇数番目のセグ

* Analysis of the World-Wide Web Problems

- Delayed Ack Performance Problems -

† Yusuke Hisada

‡ NTT Software Laboratories

表 1: Ack 送信数 (下段は削減効果 %)

容量 (バイト)	min	med.	mode	mean	max	SD
231,540	238	257	259	259.1	304	9.6
85,227	89	95	95	95.4	119	4.0
	94.0	86.7	86.7	86.2	57.8	

標本数は 246 (231540 バイト), 251 (85227 バイト).

らつきがみられるため、最高の削減効果は常には得られていない。

3.3 効果減少のメカニズム

セグメントの総数はデータ容量から一意に決まるため、パターン B か D が 1 回発生するとパターン A は 1 回減少する。これに伴ってパターン B か D がさらに 1 回増加する。パターン B と D は Ack 送信数を削減できないため、Ack 送信数は 1 回増加する。

一方、パターン C が発生すると A は 1 回以上減少するが、必ずしも Ack 送信数を増加させるとは限らない。送信側 TCP はパターン C が 1 回発生したことを知るまでの間に、次のいくつかのセグメントを送信する。一方、受信側 TCP は目的のセグメントを受信するまで Ack を繰り返し転送し、受信に成功した場合にはそれまでに受信した最後のセグメントへの Ack を送信する。つまり、損失が生じてから回復するまでの間の Ack 送信数が受信したセグメント数の 1/2 以下ならば、パターン A で処理された場合よりも Ack 送信数は増加しない。逆に、1/2 よりも多い場合には Ack 削減効果を減少させる。

3.4 効果減少の要因

パターン B はセグメントの時間間隔が長い事例が増えるほど増加し、パターン D は受信側 TCP の Ack 送信プロセスの問題が頻繁に発生するほど増加している。また、セグメントの再転送数が増加するほどパターン C の発生回数も増加し、Ack 送信数を削減できた事例が極めて少ないため、全体的には Ack 送信数を増加させている。つまり、輻輳もしくは Ack 送信プロセスの問題のために、Ack 送信数が増加する (Ack 削減効果は減少する)。

3.4.1 セグメントの時間間隔 (パターン B の要因)

セグメント間の時間間隔は輻輳が発生していない場合には高々 300 ミリ秒であるが、輻輳が発生すると 500 ミリ秒を超えるケースもみられる (図 2)。パターン B は時間間隔が 200 ミリ秒を超えると発生するため、輻輳が重大になるほど Ack 削減効果への影響は大きい。

パターン B の発生を防ぐためには、タイマ設定値を 200 ミリ秒よりも長くすればよい。例えば、500 ミリ秒と仮定すると、重大な輻輳が発生してもタイマ処理によって Ack 送信数の増加を抑えることが可能になる。しかしながら、誤った再転送が発生する恐れもあるため、実際に検証する必要がある。

3.4.2 セグメントの再転送 (パターン C の要因)

輻輳がさらに深刻な状況になると、パターン C が頻繁に発生する (図 2)。しかも、セグメント 1 個あたりの Ack 送信数が 0.5 回以下の事例はごくわずかである。231 キロバイトの場合で、0.5 回以下の事例は 3% で、0.5 回より大きく 1 回未満の事例が 55%、1 回の事例が 17%、1 回より大きく 2 回以下の事例が残りの 25% である。85 キロバイトの場合にも同様の傾向を示す。パターン C で転送されたセグメント数が全体の 10% を超える事例は 231 キロバイトの場合には 20%、85 キロバイトの場合には 30% あり

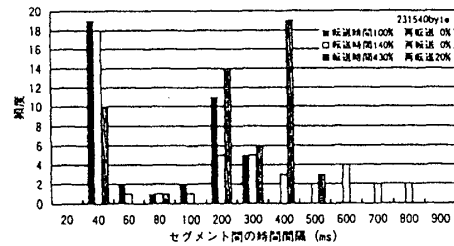


図 2: セグメント間の時間間隔 (パターン B と D の合計)

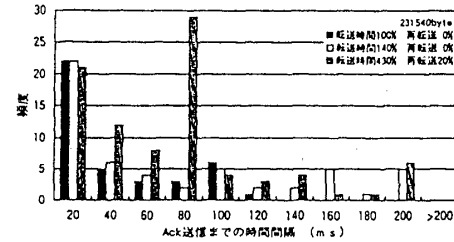


図 3: Ack 送信までの時間間隔 (パターン B と D の合計)

パターン C による影響が小さくない事例が多い。これらの事例では、輻輳をさらに悪化させる恐れがある。

3.4.3 Ack 送信プロセスの問題 (パターン D の要因)

タイマ設定値の 1/2 以下の時間間隔で Ack を送信している点で、パターン D は B と異なっている。また、輻輳の程度に関係なくパターン B よりも出現回数が多く、輻輳が発生していない場合には 55% の事例で、一方輻輳が発生している場合でも 25% の事例でセグメントを 1 個受信してから 20 ミリ秒以内に Ack を送信している (図 3)。さらに、それぞれ 98%、72% の事例で 100 ミリ秒以内に Ack を送信している。このためパターン D の Ack 削減効果への影響は大きく、特に、輻輳が発生していない場合には Ack 送信数を増加させる主な要因である。しかしながら輻輳が発生した場合にも無視できない要因である (図 2)。

パターン D の発生原因はタイマ処理の不安定さ、もしくはセグメント 1 個分しか確保できていない状態で新たにセグメントを受信したことのおいずれかと考えているが、現在までには特定できていない。

4 まとめ

日米間のリソース転送処理では、遅延 Ack タイマを 200 ミリ秒に設定することにより高い Ack 削減効果を得ることができる。しかしながら、輻輳が引き起こす大きなラウンドトリップ遅延やセグメント損失が遅延 Ack の動作を妨げるため、また、受信側 TCP における Ack 送信プロセスの問題のために、転送するデータ容量が大きければ数十回のオーダーで Ack 送信数を削減できない場合も少なくない。つまり、輻輳に深刻な影響を及ぼす恐れがある。

参考文献

- [1] Clark, D.D., "Window and Acknowledgement Strategy in TCP", RFC813, 1982
- [2] Braden, R., "Requirements for Internet Hosts - Communications Layer", RFC1122, 1989
- [3] Jacobson, V., "Congestion avoidance and control", in Proc. ACM SIGCOMM'88, pp.353-359, 1988
- [4] Wright, G.R., Stevens, W.R., "TCP/IP Illustrated, Vol.2 The Implementation", Addison-Wesley, 1995