Evaluation of TCP Acknowledgment Mechanism Influence on Router Performance

L. Pavilanskas, A. Statkus
Telecommunications Engineering Department, Vilnius Gediminas Technical University,
Naugarduko str. 41, LT-03227 Vilnius, Lithuania, e-mails: lukas.pavilanskas@el.vgtu.lt, arunas.statkus@teo.lt

Introduction

Nowadays with growth of data traffic [1], it is essential to control channel congestion efficiently in Internet for successful utilizing network resources (routers, switches, links, and etc). The main protocol for current objective is Transport Control Protocol (TCP) today. It is reliable connection-oriented protocol, which implements the flow control on heterogeneous network by means of a sliding window, congestion avoidance, and acknowledgment mechanisms, simultaneously. TCP runs over more network technologies than any other protocol and provides the highest degree of interoperability.

The main problem with TCP is that typical protocol implementations (RFC 3782) are successful at low speed (up to 100 Mbps) but it is unfit for high speed networks due to slow grow-up of congestion window, and poor evaluation of upper bound of the channel [2]. This is because the congestion window growth function – \( W(t) \) has square features on initial phase (RFC 2581) and depends on channel delay: if latency is high, the congestion window growth rate will be low. To eliminate the following drawback, many TCP congestion control algorithms were developed. Most common, oriented for large volumes of data transfers, are the following: HSTCP (RFC 3649), Fast (RFC 3782), STCP (RFC 3286), Cubic [3].

All of these versions also rely on end-to-end ACK and congestion window [4], but general difference is that \( W(t) \) is cubic-root on initial phase and depends on latency marginally, comparing with typical versions. The congestion window growth is defined in real-time and depends on last congestion event – received ACK [3]. In other words, the sender is allowed to increase the TCP data rate for each incoming ACK. As a result, unhampered transmission of ACK for successful flow control is very important, and any ACK rate suspension can influence the TCP functionality and performance [5–7].

So, in high speed network we have another undesirable feature – growth of technological expenditures [8]. Moreover, TCP is mainly used in Internet for bulk traffic (HTTP, FTP, SMTP, IMAP, and etc) [9]. Bulk traffic sources are sending large packets in one direction only [10]. Therefore, uplink traffic is composed of short packets as showed in [5]. In such conditions, encapsulated Ethernet frame with TCP message without data (only ACK flag and number) is 64 Bytes, not including the preamble.

The poor performance of short frame traffic is well known and analyzed in detail [11], and [12]. This degradation occurs for two reasons: technological expenditures [8] and routers CPU overload [13]. The last concerns to the fact that router CPU load depend on the packet rate and not to the packet length [14]. It’s clear, that on high speed networks this is observable better [13]. Therefore, with growth of bandwidth the TCP ACK message rate is increasing too. For a large number of short packets in traffic more efficient routers will be needed. Hereby, in following situation the routers with slow CPUs will become bottlenecks and will lead to the degradation of network performance too. The critical threshold of router functionality is ~ 45-50% of CPU load [15].

One of the ways to increase performance in modern high speed networks is by lowering the number of TCP ACK messages. This can be done through ACK filtering mechanism [16], which is well known and used in asymmetric and wireless networks environments [5], [6], [7], but is considered to be inconsistent with TCP traffic.

The aim of current paper is to find, how ACK filtering can influence the TCP functionality, and how significant number of TCP ACK messages affects the performance of the network channel – chain of the routers.

TCP acknowledgment mechanism

The acknowledgement mechanism is at the heart of TCP and relies on demand for the receiver to communicate with the sender by sending back an ACK as it receives data. TCP uses a cumulative ACK technique (RFC 813). The ACK number field in a TCP message header received by the sender indicates that all bytes of data with sequence
The congestion window is being lost in routers queues. For this purpose it uses a data can be outstanding in the network without packets to increase the received ACK. After initializing the TCP sender is allowed on the feedback sender gets from the network through the congestion window, which growth function depends on the same direction. A typical TCP uses delay the last window reduction, multiplication decrease factor applied for data flows in Internet [9]. Delayed ACK is usable for an to degradation of the congestion window to its minimal value. Meanwhile, on phase when congestion event becomes less important. In that case, TCP operation and acknowledgment mechanism diagram

Normally TCP does not send an ACK the instant it receives data. It delays ACK (RFC 2581), hoping to have data going in the same direction. A typical TCP uses delay of 200 ms and sends ACK for every other data message. However, there is not much TCP traffic with bidirectional data flows in Internet [9]. Delayed ACK is usable for an interactive applications (Telnet, SSH, and etc.) only [10].

The TCP sender’s uses ACK to estimate how much data can be outstanding in the network without packets being lost in routers queues. For this purpose it uses a congestion window, which growth function $W(t)$ depends on the feedback sender gets from the network through the received ACK. After initializing the TCP sender is allowed to increase the $W(t)$ for each incoming ACK.

In this paper for research we use TCP Cubic were the congestion window is determined by function [3]

$$W(t)_c = C \left( t - \frac{W(t)_\text{max} - \beta}{C} \right)^3 + W(t)_{\text{max}},$$

where $C$ is a scaling factor, $t$ [s] – the elapsed time from the last window reduction, $W(t)_{\text{max}}$ [bytes] – the maximal window size before the last reduction; $\beta$ – a constant multiplication decrease factor applied for $W(t)_c$ reduction to minimal at the time of loss event ($\beta W(t)_{\text{max}}$) [3]

$$W(t)_{\text{min}} = W(t)_{\text{max}} - \beta \cdot W(t)_{\text{max}}.$$ (2)

Consequently, the ACK reception for TCP sender is very important process. Especially at initial phase, where growth of $W(t)_c$ is very intense [3] and ACK loss can lead to degradation of the congestion window to its minimal value. Meanwhile, on phase when $W(t)_c$ is close or equal to maximal value - $W(t)_{\text{max}}$, growth becomes minimal (2) and congestion event becomes less important. In that case, ACK is mainly used for detection of data message loss and this is a second very important function of the TCP acknowledgment mechanism (RFC 813).

The TCP data message retransmission can occur because of two main reasons [17]: after expiration of retransmission timer (RFC 813) and after receiving three or more duplicated ACK (Fig. 1: $ack_3$, $ack_4$, $ack_5$). In first case the retransmission timer expires when no new data is acknowledged for a set of threshold time (RFC 2581).

The retransmission timeout ($t_{\text{RTO}}$) is taken as a loss indication, and it triggers retransmission of the unacknowledged segments. The threshold time during which the confirming ACK message must be received is

$$t_{\text{RTO}} = \frac{t_{\text{RTT}} + 4 \cdot \sigma_{\text{RTT}}}{3},$$ (3)

where $t_{\text{RTT}}$ is the average of time of successful TCP message transmission (Fig. 1: $t_{\text{RTT}}$ measurements); $\sigma_{\text{RTT}}$ – the root-mean-square of deviation of $t_{\text{RTT}}$.

If during $t_{\text{RTO}}$ the ACK is not received, the data segments loss will be detected and the TCP sender will set $W(t)_c$ to 1 segment (RFC 2988); since $t_{\text{RTO}}$ indicate that channel utilization has changed dramatically [17].

Second reason occurs after receiving three or more duplicated ACK [17]. As shown in Fig. 1 the message $data_4$ was lost. The TCP receiver accepts messages $data_3$, $data_5$, $data_6$, and so on (Fig.1.), but not $data_4$. The $ack_4$, $ack_5$ with same ACK number 512 were sent for received messages. After duplicated ACK is detected, the transmitter waits for $t_{\text{RTO}}$ to expire. TCP does not yet know whether a duplicate ACK is caused by a loss or just reordering of segments. TCP sender waits the timeout (3) and assumes that if there is just a reordering it will not get ACK with number 512 anymore. However, after a while, a third duplicated ACK is received in a row ($ack_5$). It is a strong indication that segments have been lost. TCP performs retransmission of segments 513-1023 with $data_6$, without waiting for a $t_{\text{RTO}}$ to expire. After this the sender maintains the number of outstanding segments by sending a new segment for each incoming ACK (1). It should be noted that only retransmission of identified as lost (timed-out) TCP segments, are implemented in the "conservative" TCP versions (RFC 3517). Meanwhile, in "aggressive" TCP implementation after the loss, all unacknowledged messages are retransmitted [17].

The TCP contains only a general assertion that data should be acknowledged promptly, but gives no more specific indication as to how quickly and as how frequently an ACK must be sent. In RFC’s clearly is indicated, that current mechanism must maintain two very important functions: to prevent data retransmission, and as soon as possible to make ACK to permitting further data to be sent. In addition to this argument, the fact that ACK message are very important in the initial phase (1) of data transmission and the fact that rate of ACK rely on data rate (segments loss due to receiving 3 or more duplicated ACK) and cannot be less than $1/t_{\text{RTO}}$ (3), must be evaluated.

**ACK filtering technique**

The concept of ACK filtering was discussed well in [16]. The idea is fairly simple. When router needs to send the current ACK, it scans queue for any early TCP ACK. If
ACK already exist in the queue, simply drop them before a queuing the new ACK. Since acknowledgments are cumulative, the newest one obsoletes all older ACK. So, there's no need for more than one per connection to ever be in the queue. Dropping old ACK when a new one is queued means there would never be more than one ACK on the queue at any time, so this is pretty much the same as replacing the earlier ACK with the newer one.

In the same reference [16] the main drawbacks of ACK filtering were reviewed. It is noted that ACK clocking scheme, which for both drop and congestion control is used, with the ACK filtering can be destroyed. The current propositions are compelling. However, the discussed apprehensions can be challenged with arguments of [5], where the TCP performance in the asymmetric links was analyzed. It is shown that asymmetry affects the TCP performance, because it relies on feedback of cumulative ACK from the receiver. In addition, typical TCP is ACK clocked, so the arrivals of ACK on the reverse channel have significant effects on the forward channel throughput. In the networks with bandwidth asymmetry the ACK filtering can work well. This improves the forward TCP throughput and the fairness of competing connections greatly. In paper ACK filtering was modeled on Opnet.

Work [6] presents a quite similar study. Analysis of unfairness problem between TCP upstream data and ACK downstream on the unevenly shared wireless channel is provided. It is shown that ACK filtering increases 802.11 channel utilization without any dependence on \( f_{RTT} \) (3).

The extensive simulations with ACK filtering in [7] were proposed. It is demonstrated that lowering ACK number can improved TCP performance significantly: achieving up to 25% gain in chain networks and 35% in a complex grid network, compare with typical TCP. In work the ACK filtering motivation follows from fact that short ACK messages consume channel capacity comparable to data packets when the transmission is high rate.

The methodology of experiments

In order to find out, what real influence the ACK filtering makes to the TCP functionality and how it affects the performance of the network channel devices (routers) an experiments were performed. For this reason Ethernet network with IP routing, and TCP session between two independent nodes (PC0 and PC1/PC2, Linux OS, 2.6.32 kernel, and TCP Cubic version enabled [3]) were created. The structure of network is presented in Fig. 2.

In current network the following equipment has been used: PC0 as FTP server – the transmitter of data TCP messages, and PC1/PC2 as FTP client – receiver of TCP data messages (TCP ACK sender), the transparent Ethernet bridge – BR, target router – R, and additional router – R0 (Cisco 881) as well. For experiments two routers of different generations Cisco 881 and Cisco 1841 were used. The main difference between them is CPU power. All devices were connected with 100 Mbps Ethernet links (100BaseTx). The ACK filtering was implemented on BR device. It has been created on Linux based (2.6.32 kernel) PC with Ethernet bridging \(<brctl>\) application. This tool was taken because Linux bridging is faster, work as simple switch, and don't make significant impact to flow parameters comparing with routing.

The ACK filtering has been pursued only in one direction from PC1/PC2 to PC0, whereas in opposite direction the traffic passed through BR without any alterations. Filtering was based on exact frame rate control with Committed Information Rate (CIR) and Committed Burst Size (CBS=5 kB). The last one was used to avoid degradation of congestion window on initial growth phase (1). Filtering was made using \(<tc>\) application, which drop/policed ACK messages if it exceeds specified rate. The scripting code of ACK filtering is shown in Fig. 3.

![Fig. 2. Structure of network used in experiments](image_url)

**Fig. 2. Structure of network used in experiments**

**Fig. 3. Scripting source of \(<tc>\) policing**

The second router (R0) has been used on purpose to keep more realistic IP based network with all routing and switching functionalities. The target router parameters during the experiments with SNMP protocol were collected throw independent router interface. For data transmissions the FTP application has been used. In all experimental iterations file of 400 MB size was transferred. The data speed was controlled on PC0 (FTP server) with \(<tc>\) script, which shaped to desirable speed without packet loss (delay of traffic only). For this Token Bucket Filtering – TBF (RFC 5624) was used.

Performance evaluation experiments

The target of first experiment scenario was to find how the ACK filtering can influence the TCP performance and data transfer integrity. For this, a file of 400 MB size from PC0 to PC1 was repeatedly transferred (Fig. 2.). On the each iteration the ACK filtering rate (0%, 20%, 40%, 60%, and 80%; values of the maximal ACK rate without filtering) was changed.

As shown in Fig. 4 TCP message rate for entire period was stable in all iterations. The growth of message rate (1) is high and equal at all ACK drop values (events...
up to ~3s). This occurred because of two reasons: $W(t)_C$ of used TCP version slightly depends $t_{RTO}$ (3) and ACK filtering is activated only after CBS is exceeded – when the $W(t)_C=W(t)_{max}$ (1). At the end of transfer we have “decline” – the finish of data transfer.

Fig. 4. TCP messages rate during file transmission for various ACK drop values

The result shows that ACK filtering does not affect data transfers of single TCP session. We observed that transfer remains stable for up to 80% of ACK drops. However, if losses are above 85% the ACK rate becomes less than $1/t_{RTO}$ (3), and data rate degrades fatally.

Fig. 5. Cumulative growth of TCP segments count during file transmission for various ACK drop values

The comparison of cumulative TCP segments (count of segments on TCP layer) and FTP bytes (count of data on FTP layer) rates during period $(t)$ of file transfer is presented in Fig. 5. It is shown that on FTP and TCP layers (lines are coincident) the same amount of bytes was received at any given time period. Consequently, it is possible to do the suggestion, that count of duplicated ACK and TCP data retransmissions (3) are not increased respectively (Fig. 1: $t_{RTO}^\prime=t_{data}+t_{ACK}$ is less then $t_{RTO}$).

The second experiment goal was to find what kind of influence cumulative ACK mechanism has on network equipment performance. For this purpose a file transfer of 400 MB, was performed and CPU load of router (Cisco 881, Cisco 1841) was measured (Fig. 2). The experiments were performed at various data rates (1, 4, 8, 16, 32, 65, and 90 Mbps) in scenarios with and without filtering and in scenario when 80% of ACK is filtered.

As shown in Fig. 6 the CPU load is decreasing when the ACK filtering is used. The same situation is observable for both routers. It is clear that results do not depend on router type and amount of data in TCP message. It depends on amount of processed packets by the router CPU. This is confirming the results presented in [14]. Moreover, in Fig. 6 is observable that CPU load utilization is linear, and depends on frame rate: if the TCP data rate increases, the ACK rate is increasing too. The functions of the CPU load curves are quite similar for both routers. The difference is only in designed CPU’s power.

Fig. 6. Target routers CPU load for various FTP traffic rate (TCP goodput) when 80% ACK is dropped

The relation between target router performance and CPU load is shown in Fig. 7. Performance increase should be understood as relative CPU load reduction caused by employing of ACK filtering. With ACK drop of 80% the performance can be increased by 30%, comparing with case when CPU load is 25%, and ACK filtering is not used. Meanwhile, if we have the CPU load of 60% (more than routers overload threshold [14]) with 80% of ACK drop we can increase the performance by 32%. It means that in current situation CPU load will be approx. 40%.

Fig. 7. Performance increase on target router CPU load for various ACK drop values
The third experiment issue was to find how the ACK filtering can influence two concurrent and independent TCP sessions. During investigation the PC0 was used as FTP server and PC1/PC2 - as FTP clients. With each TCP session the 400 MB files were transferred. The ACK filtering was performed on both sessions simultaneously.

The dependence of two TCP sessions message rate during file transmissions is presented in Fig. 8. Graph clearly shows that both independent sessions remain concurrent and divide channel almost equally. In fact, this cannot be so, because TCP objective is not equal channel sharing. At 3s, as in situation with one session (Fig. 4), the TCP1 begins to increase the message rate according to \( W(t) = \frac{1}{C RTO} \) (1). While at 6s TCP2 session is starting too, and about 9s the message rate of both sessions becomes approx.: \( R_{TCP1} \approx 44 \text{Mbps} \); \( R_{TCP2} \approx 49 \text{Mbps} \); \( R_{TCP} \approx 93 \text{Mbps} \). While the TCP1 session is degraded due to congestion (2).

After 78s the TCP1 rate is decreasing since finishing of file transfer, while the TCP2 conversely starts to grow-up rate (1). At 81s TCP2 is finishing transfer too. Current fine competition between two independent session’s show, that both of them from TCP point of view are working well, and the ACK filtering does not make significant influence on TCP functionality in current conditions.

The dependence of cumulative data during two file transmissions is shown in Fig. 9. It can be seen that both sessions collect messages well in TCP layer and in application layer (Fig.1.). It means that TCP goodput (transmitted data to upper layer) of both sessions are close to maximal, while the count of duplicated ACK is minimal.

TCP interoperability defines whether a protocol is fair to other TCP sessions. Therefore, it’s important to find how ACK filtering increases unfairness of TCP. For this purpose fourth experiment was performed. The scenario was the same as in previous, only the filtering for one session and the PC0 without shaping was used (Fig. 2.).

The dependence of cumulative data during two file transmissions is shown in Fig. 9. It can be seen that both sessions collect messages well in TCP layer and in application layer (Fig.1.). It means that TCP goodput (transmitted data to upper layer) of both sessions are close to maximal, while the count of duplicated ACK is minimal.

TCP interoperability defines whether a protocol is fair to other TCP sessions. Therefore, it’s important to find how ACK filtering increases unfairness of TCP. For this purpose fourth experiment was performed. The scenario was the same as in previous, only the filtering for one session and the PC0 without shaping was used (Fig. 2.).

Conclusions

In current work the investigation of TCP functions and ACK filtering was presented. It is shown that TCP congestion window and acknowledgment mechanism are dependent processes with common variables. Therefore, the growth of network bandwidth is coherent with ACK rate: when the network capacity is increasing, the ACK rate on channel is increasing too. This leads to network equipment CPU performance degradation. To resolve this weakness the usage of ACK filtering was proposed.

The experiments were performed to find how significantly number of ACK affects the router CPU and how impact of ACK filtering can influence the TCP functionality and performance. We conclude that:

1. ACK filtering does not significantly affect the TCP data transfer in normal network conditions. The TCP data transfer remains stable for up to 80% of ACK drops. However, if losses are above 85% the ACK rate becomes less than \( \frac{1}{RTO} \), and session is terminating immediately.

2. The performance of router CPU depends on ACK count and can be increased with ACK filtering. The CPU
load utilization is linear, and depends on data rate: if the TCP data rate increases, the ACK rate is increasing too.

3. On CPU load of 25% with 80% of ACK drop, it is possible to increase the router performance by 30%. Meanwhile, on CPU load of 60%, on the same filtering conditions, the performance can be increased by 32%.

4. The ACK filtering can be used not only with single TCP session but also with concurrent sessions. Results show that two sessions work well, without any evident signs of the instabilities; although a slight unfairness among TCP sessions were observed.

In conclusion, the presented results show that ACK filtering does not affect the functionality of TCP in normal network conditions, but allows more efficient use of network equipment. Therefore, to implement the suggested network conditions, but allows more efficient use of network equipment.

References


Received 2010 02 11