Transmission and Retransmission Scheduling for Terrestrial and Space Internetworks

Ioannis Psaras
ipsaras@ee.duth.gr
Advisor: Prof. Vassilis Tsaoussidis
Dept. of Electrical and Computer Engineering
Democritus University of Thrace
12 Vas. Sofias Str., 67100
Xanthi, Greece

May 2008

Ph.D. Dissertation Preface

Transmission and Retransmission Scheduling for transport protocols has been traditionally designed on the basis of *demand and supply*. *Demand* refers to the amount of user data that need to be transferred over the Internet, while *supply* refers to the capacity of the Internet infrastructure. In the early days of the Internet, demand did not exceed supply and therefore, congestion was not a matter of concern. The *Transmission Control Protocol* (TCP) was designed accordingly. That is, *feedback-based* strategies were not included in TCP's basic design principles.

The Internet evolution resulted in increased amounts of data that had to be transferred over the network. When demand exceeded supply, buffer overflows led to congestion events. Given that TCP was based on a feedbackless model, fixed (i.e., non-decreasing) transmission rates caused congestive collapses.

In the mid 80's, when congestive collapses obscured the operational fidelity of the Internet, the networking research community optimized TCP in order to avoid such annoying and destructive events. Feedback-based *Transmission and Retransmission Scheduling* rules, were defined in order to: i) avoid collapses, ii) guarantee stability and iii) achieve fair resource sharing among users. This way, TCP evolved as a *closed-loop, binary feedback, congestion control system.*

The success of the Traditional Internet relies mainly on its ability to provide system utilization and fairness. In turn, system utilization and fairness rely on the transmission scheduling strategy of the transport layer protocol. When transmission scheduling fails, due to symptomatic events for example, the retransmission strategy has to take over in order to provide reliability. Similarly, the behavior of the retransmission rule determines the transmission scheduling itself. For example, aggressive and recurring retransmissions reduce transmission opportunities. Therefore, the transmission and retransmission scheduling strategies have to be studied jointly. Furthermore, they have to be studied on the basis of flow contention. That is, low contention allows for easy and efficient flow multiplexing and in turn leads to fair resource allocation. In contrast, when the level of contention increases, current transmission and retransmission scheduling rules tend to synchronize flows. We show that flow synchronization leads to sharp load fluctuations, unfair resource allocation and system instability.

The extension of Internet into space calls for further investigation on the applicability of Internet rules to space environments. Although the Traditional Internet evolves towards a highly-sophisticated system following the demands of users and applications, Space Internet still faces several technical challenges, from an operational perspective. Both systems, for example, require reliability, which further means that a retransmission strategy is necessary. Retransmission on the terrestrial Internet, however, has the possibility to rely on feedback-based models. In contrast, huge propagation delays experienced in the space environment call for feedbackless retransmission scheduling.

Our initial goal is to identify mismatches between the service requirements and the system constraints, focusing on strategies for data transmission and retransmission. For example, real-time applications (such as on-line TV) require smooth load fluctuations and low delay paths. In contrast, mobile users demand for fast service regardless of the impact to the system load. Finally, mission critical, non-congestive, telecommand or sensor applications require both fast and reliable data delivery. That said, service requirements have increased in extent and complexity, but the rules that govern data transmission and retransmission have not. Our goal, in the present thesis, is to adjust the rules according to the service requirements in order to provide high-quality services to Internet users.

In the following, we briefly discuss the main characteristics and requirements for *Modern* and *Space Internetworking.*

1. **Modern Internet.**

   The Internet has expanded to reach almost every house, enterprise and university campus. Modern Internet is characterized by application and infrastructure diversity.

   The huge amount of applications that have been developed, caused evolution of the Traditional Internet into a multipurpose Internet. That is, Modern Internet serves as the ultimate tool for business, entertainment, education and lately for data storage purposes as well.

   On the other hand, advances in communication systems have brought innovative changes to transmission media technologies. For example, high-speed transmission links have already been deployed in backbone routes; Internet data can be transferred over wireless transmission media; links that exploit the geographical location of orbiting satellites serve as intermediate routes for data transmission between transatlantic Internet endpoits.

Application and infrastructure heterogeneity has changed dramatically the Traditional Internet's traffic characteristics. For instance, various applications download data and leave the system, causing drastic load fluctuations. As a consequence, demand may only temporarily exceed supply. On the other hand, generous supply may be provided permanently (e.g., high-speed paths). Finally, demand may exceed supply for long time periods, during rush hours for example. In all cases, transport protocols inevitably adjust their transmission and retransmission rates according to feedback provided either by the receiver side or by the network itself. Feedback, however, arrives at the sender side, typically, in binary form. That is, the sender is not informed about the intensity of contention and therefore, responses are identical in all cases. Moreover, link errors, which result in packet losses, are falsely interpretted as congestion events and trigger similar responses as well. Summarizing, the growth of the Internet and its subsequent effects on the characteristics of Internet traffic have not been incorporated into the *Transmission and Retransmission Scheduling* rules of transport layer protocols.

2. **Space Internet.**

   During the last decades there has been an increasing interest for space exploration by most space agencies worldwide. A large number of missions has been launched in order to explore climate and surface conditions of outer space planets. The success of a mission, however, depends largely on the efficiency of its communication infrastructure. That is, once scientific data is gathered from satellites, rovers, landers, sensors etc., the data has to travel through deep-space links in order to reach scientists and research labs on Earth.

   Traditionally, the communication infrastructure of space missions, was based on the telecommunication systems technology, namely *circuit switching* networks. Although circuit switched networking comprises an elegant approach for some applications, such as telephony for example, it appears to be inefficient at least, for data transfers in the outer space environment. Circuit switching networks, for example, guarantee reliability by multiple recurring transmissions of the whole file, in case a part of it gets lost or corrupted. Instead, according to *packet switched* networking the file is fragmented into *packets*. A packet, however, is more flexible than a file, size-wise. Thus, *packet switched* networking, the main technology that led to the success of the Internet, allows for retransmission of corrupted or lost packets. Hence, packet switched networking provides reliable, on-time and efficient, energy consumption-wise, delivery of data to end-hosts. In that context, the Internet has gained one extra dimension. Apart from its terrestrial nature, the Internet has expanded to reach outer space planets.

The communication framework is called InterPlaNetary (IPN) or Space Internet. The new era of space communications and networking is called *Delay/Disruption Tolerant Networking*, to reflect its main characteristics, namely huge delays and high packet corruption rates.

Traditional Internet can be considered as a closed-loop system. However, the rules change when we extend communication into space: extraordinary long propagation delays as well as high packet error rates render the use of feedback-based strategies prohibitive. Therefore, conversational mechanisms, algorithms and protocols, such as TCP for example, are not expected to perform efficiently (if at all). *Transmission and Retransmission Scheduling* need, therefore, be defined on the basis of a feedbackless system.

We depart from the following question:

Which are the important components that are missing from current transmission and retransmission scheduling strategies in order to achieve utilization and fairness in Modern and Space Internetworks?

In a *feedback-based system*, such as the Modern Internet:

- **The transmission strategy has to be *contention-adaptive*.** Adjustment of the transmission rate itself has been studied extensively in the related literature, but the acceleration (either positive or negative) in order to reach the desired transmission rate has not. Typical systems adjust their rate, inevitably, only when contention leads to congestion; then, timeouts and multiplicative decreases force flows to reduce their windows. However, the transmission policy itself (i.e., increase/decrease rules) is not adjusted according to network contention. That is, although systems are adaptive to network dynamics, this adaptivity is limited: window size can be regulated but the window increase rate cannot. *This is similar to a car regulating its velocity scale but with fixed acceleration.*

- **The retransmission strategy has to be *contention-aware*.** Statistical Multiplexing, the main technology of packet networks, allows for bandwidth sharing on demand and cancels the possibility for flow starvation. When the resource demand exceeds the supply, that is, when flow contention grows, multiplexing exhibits different characteristics. For example, since packets are dropped due to high demand and limited supply, current demand, and consequently the strategy for multiplexing is determined mainly by the timeout algorithm. That is, *the timeout becomes the scheduler for the link.*

  Although the design of the timeout algorithm has been studied extensively in the past, its association with link scheduling has not. In addition, its scheduling properties have not really been evaluated adequately. Instead, much attention has been paid on its ability to reflect network delay accurately, allowing for speedy retransmission when conditions permit and avoiding double submission due to early expiration. However, network delay as it is captured by measuring the RTT alone, cannot really capture network contention. This way, increased amounts of network contention lead to flow synchronization. That is, congested queues stabilize RTT Variations and minimize RTT Deviations. In turn, timeout values get fixed and deterministic values. Therefore, some flows always find a slot for transmission in the router's buffer space, while some others always fail, leading to flow starvation and unfair resource allocation.

In a *feedbackless system*, such as the Space Internet:

- **The transmission strategy has to be *rate-based*.** Traditional feedback-based increase/decrease rules do not perform efficiently in challenged environments, such as the Space Internet. In particular, extraordinary long propagation delays decelerate rate adjustments to prohibitive levels. Therefore, the transmission strategy has to be decoupled from all kinds of potential feedback. In that sense, we arrive to the notion of an Open-loop system. Moreover, provided that access to specific nodes or links in the outer space is not public, but rather private, fixed and predetermined, rate-based transmission seems to be the most appropriate scheme for the Space Internet.

- **The retransmission scheduling strategy has to be *proactive*.** We observed that feedback-based retransmission strategies are inefficient at least, for huge propagation delay and highly-error-prone links. In case the sender awaits for signals in order to retransmit lost or corrupted packets, valuable time is spent in transmission-idle periods without any actual gain in system performance. Instead, appropriately scheduled proactive retransmission attempts achieve faster file transfer times, without extra retransmission overhead.

In the present thesis, we provide a novel solution framework to address the aforementioned observations. The main contribution of the thesis is summarized as follows:

*Data transmission and retransmission scheduling rules are designed according to service requirements for Terrestrial and Space Internetworking. That is, data transmission and retransmission scheduling becomes **contention-aware** for Terrestrial Internetworks and **proactive** for Space Internetworks. This way, system stability is preserved, while system efficiency and fairness increase.*

The contribution of the present thesis has to be evaluated considering the following facts/arguments:

- The Internet is growing further. Application diversity and infrastructure heterogeneity is expected to expand. The number of networked devices will increase exponentially to include, for instance, moving vehicles (e.g., cars, airplanes), household devices etc. Hence, service requirements will increase further. In that case, *contention-aware rules will preserve system stability.*

- Critical applications, such as online medical diagnosis or spacecraft command and control, are incrementally deployed. In such cases, loss of a single packet may have irreversible and destructive consequences. However, chances for packet loss increase with the extend of resource overutilization. Thus, *keeping contention at regular levels leaves available space for extra, demanding or not, possibly critical applications.*

- Congestion (or high levels of network contention) will always exist. Although currently high-speed links provide plenty of bandwidth, in the future, applications will become more demanding (e.g., ultra high definition music and movies transferred over the Internet). In this context, *contention-adaptive congestion control is required in order to provide efficiency and fairness for Modern Internetworking.*

- Scientific data from outer space planets become pretty valuable if we consider the investments required for a space mission. Moreover, interoperability actions, such as link rental for example, increase expenses further. Therefore, *reliable and fast data transmission in the Space Internet becomes a matter of mission viability.*

In the present thesis, we initially evaluate the performance of the central congestion control algorithm of the Internet, namely the *Additive Increase/Multiplicative Decrease* (AIMD). We argue that a fixed value for the Additive Increase rule of AIMD comprises a natural limit for the responsive behavior of TCP over the modern Internet infrastructure.

We propose AIRA, an *Additive Increase Rate Accelerator* [3], [5]. AIRA extends AIMD functionality towards adaptive increase rates, depending on the level of network contention and bandwidth availability. In this context, acceleration grows when resource availability is detected by goodput/throughput measurements and slows down when increased throughput does not translate into increased goodput as well. Thus, the gap between throughput and goodput determines the behavior of the rate accelerator. We study the properties of the extended model and propose, based on analysis and simulation, appropriate rate decrease and increase rules. Furthermore, we study conditional rules to guarantee operational success even in the presence of symptomatic, extraordinary events. We show that analytical rules can be derived for accelerating, either positively or negatively, the increase rate of AIMD in accordance with network dynamics. Indeed, we find that the blind, fixed Additive Increase rule can become an obstacle for the performance of TCP, especially when contention increases.

Our extensive performance evaluation reveals that when the packet loss rate increases, the retransmission scheduling strategy becomes more crucial than the transmission scheduling policy for the performance of the transport layer protocol. Based on that observation, we evaluate the retransmission scheduling properties of the TCP Retransmission Timeout algorithm (TCP-RTO). Initially, we investigate the efficiency of the TCP-RTO algorithm over short delay, regional paths. We observe that the conservative, 1-second *Minimum TCP-RTO* value, specified in RFC 2988, is indeed an obstacle to fast retransmission. We evaluate the correctness of the suggested policy and investigate the impact of the Minimum TCP-RTO under modern networking conditions. We show, through simulations, that an *Adaptive MINRTO* improves TCP performance significantly, especially in case of next generations high-speed, wireless-access networks [6], [4]. The performance of the proposed mechanism depends on several network conditions. For example, the gain of the Adaptive MINRTO increases with the speed of the transmission link (i.e., high-speeds).

We go beyond typical assumptions to also investigate the performance of the retransmission timeout algorithm of TCP, when the level of contention increases. We notice that increased contention may stabilize RTT variation, minimize the deviation and, in turn, shorten the timeout. We show that this behavior is undesirable indeed, since it leads to unfair resource utilization. We propose the *Contention-Aware Retransmission Timeout* (CA-RTO) [9], an algorithm that incorporates a contention parameter and a randomization technique into the Retransmission Timeout. We report significant improvement in fairness, great reduction of retransmitted packets and slight improvements in application goodput.

During that study, we observed that the design principles of the TCP Retransmission Timeout algorithm inherently lead to flow synchronization, unnecessary retransmission effort and unfair resource allocation. We present a new, contention-aware *Window-Based Retransmission Timeout* (WB-RTO) [7], [8] algorithm for TCP, which exhibits two major properties: (i) it cancels retransmission synchronization, which dominates when resource demand exceeds resource supply and (ii) it reschedules flows on the basis of their contribution to congestion. WB-RTO achieves better fairness and slightly better goodput with significant less retransmission effort.

In order to evaluate viability of the geographical extension of the Internet into Space, we investigate requirements for the transition to Space Internetworking. In particular, deep-space communication links are characterized by extraordinary long propagation delays, high BERs, intermittent connectivity (i.e., blackouts) and bandwidth asymmetries. Common approaches to deal with the above unique characteristics are: rate-based, open-loop protocols to deal with huge propagation delays; regular retransmissions to deal with high BERs; transmission suspension to deal with blackouts; SNACKs to deal with bandwidth asymmetries. We adopt some of the above approaches, namely, the open-loop, rate-based transmission and the SNACKs and focus rather on the optimization of the rest, namely, the retransmission strategy of the transport protocol to deal with high BERs and blackouts.

We introduce the *Double Automatic Transmission* (DAT) technique. DAT sends each packet twice, importing some delay (Rd) between the original transmission and the retransmission. Therefore, in the presence of communication gaps (i.e., errors or blackouts), corrupted packets will eventually be replaced by the same correct packets that arrive with delay Rd. Rd, however, is much smaller than the traditional TCP-RTO value.

Our solution framework is incorporated into the *Deep-Space Transport Protocol* (DS-TP) [2], [1], a new reliable protocol for deep-space communication links. DS-TP's main advantage is its ability to complete file transfers faster than conventional TCP, SCPS-TP and Saratoga. Therefore, missions with small connectivity time are greatly favored. Our theoretical performance evaluation results reveal that DS-TP presents high potential for deployability. In particular, we show that for PER=50%, DS-TP completes a file transfer in half as much time as conventional protocols need.

The main conclusions of the present thesis are summarized below.

1. *Contention-Adaptive Transmission Scheduling* rules preserve system stability and reduce the retransmission effort, without reducing the goodput performance of TCP, when contention is high. Instead, in case of low contention, contention decrease events or random errors, Contention-Adaptive Transmission Scheduling can exploit available resources faster and therefore achieve high utilization.

2. *Adaptive values for the Minimum TCP-RTO* allow for faster retransmission attempts, when conditions permit (i.e., low end-to-end delays). In particular, the lower the path-delay, the greater the gain. Therefore, regional flows are greatly favoured.

3. *Contention-Aware Retransmission Scheduling* rules cancel flow synchronization and consequently flow starvation as well. As a result system fairness increases. Moreover, retransmission overhead is significantly reduced, increasing this way system stability.

4. In contrast to Terrestrial Internetworks, Space Internetworks require *Proactive Transmission and Retransmission Scheduling* rules. Proactive rules achieve i) fast link utilization and ii) fast retransmission of lost or corrupted packets. Therefore, file transfers over deep-space communication links are completed faster. According to our solution framework, the file transfer completion time reduces as the link propagation delay increases.

References

[1] Giorgos Papastergiou, Ioannis Psaras, and Vassilis Tsaoussidis. Deep-Space Transport Protocol: A Novel Transport Scheme for Space DTNs. *Submitted to Elsevier Computer Communications, Special Issue on Delay and Disruption Tolerant Networking.*

[2] Ioannis Psaras, Giorgos Papastergiou, Vassilis Tsaoussidis, and Nestor Peccia. DS-TP: Deep-Space Transport Protocol. In *Proc. of the IEEE Aerospace Conference, 2008, Big Sky, Montana, USA*.

[3] Ioannis Psaras and Vassilis Tsaoussidis. AIRA: Additive Increase Rate Accelerator. In *Proceedings of IFIP NETWORKING 2008, Singapore*.

[4] Ioannis Psaras and Vassilis Tsaoussidis. On the Properties of an Adaptive TCP Minimum RTO. *Submitted to Elsevier Computer Computer Communications*.

[5] Ioannis Psaras and Vassilis Tsaoussidis. On the Properties of an Additive Increase Rate Accelerator. *Submitted to Elsevier Computer Networks*.

[6] Ioannis Psaras and Vassilis Tsaoussidis. The TCP *Minimum RTO* Revisited. In *Proceedings of IFIP NETWORKING 2007, Atlanta, GA, USA*.

[7] Ioannis Psaras and Vassilis Tsaoussidis. WB-RTO: A Window-Based Retransmission Timeout for TCP. In *In Proc. of the 49th IEEE Global Telecommunications Conference GLOBECOM 2006, San Francisco, USA*.

[8] Ioannis Psaras and Vassilis Tsaoussidis. Why TCP timers (still) don't work well. *Computer Networks*, 51(8):2033–2048, 2007.

[9] Ioannis Psaras, Vassilis Tsaoussidis, and Lefteris Mamatas. CA-RTO: A Contention-Adaptive Retransmission Timeout. In *Proceedings of ICCCN*, October 2005.