Abstract—Transmission Control Protocol (TCP), the dominating transport protocol in the Internet, was originally mainly designed for wired networks where packet loss is mostly due to congestion. Deploying standard TCP in hybrid networks leads to a performance degradation due to the wireless link characteristics. To deal with this issue, intermediary nodes on the network path called Performance Enhancing Proxies (PEPs) were designed. In this paper, we first survey PEPs for TCP to assess their relevance in the current research. We then focus on current PEP approaches for QUIC, a new transport protocol incompatible with TCP PEPs. We find practical approaches that benefit the performance of QUIC and discuss their limitations mostly regarding scalability and security.

Index Terms—performance enhancing proxies, satellite networks, tcp, quic, masque

1. Introduction

TCP is a transport layer protocol that provides a reliable connection-oriented ordered data transfer while offering mechanisms for flow and congestion control [1]. It was originally designed mainly for wired networks where packet losses are mostly due to network congestion. Standard TCP interprets packet losses always as congestive and reduces the data transfer rate unnecessarily in case of high Bit Error Rate (BER) which leads to a significant performance degradation.

This unwanted behavior is observed in networks that contain both wired and wireless connections, called hybrid networks. The wireless section is in particular characterized by a high BER. TCP has to adapt to the different needs of these networks to provide the required performance. To this end, several enhancement approaches have been proposed: Wireless TCP variants that introduce modifications to the standard protocol were designed and tailored for the specific needs of the different wireless networks [2]. Other approaches introduced intermediary middleboxes on the network path called Performance Enhancing Proxies (PEPs) [3]. The common approach is to split the TCP connection transparently into two independent connections. Congestion and error control can therefore be tailored for the specific characteristics of each connection.

The development of QUIC [4] as a secure transport protocol based on end-to-end encryption of not only the payload but also most of the control information renders transparent PEPs useless, since they rely on the inspection of unencrypted TCP headers for their functionality. However, the benefits of connection splitting for performance enhancement motivate the exploration of PEPs for QUIC [5].

In this paper, we first provide a general overview of the challenges posed by hybrid networks, an explanation of PEPs’ functionality and QUIC. We then investigate TCP PEPs. Finally, we present, analyze and discuss the limitations of different approaches to QUIC PEPs gleaned from the current research.

2. Background

We first provide a general overview of the challenges posed by hybrid networks, an explanation of PEPs’ functionality and QUIC.

2.1. Problems of Wireless Networks

There are various wireless network architectures. We mention for instance satellite networks and cellular networks. Despite the challenges posed by their specific needs, all wireless networks face the problem of high BER since they use air as transmission medium and are therefore more prone to random factors, such as bad weather conditions and the mobility of end users [2]. In cellular networks, a base station interconnects the wired fast network and the wireless mobile network to provide internet access for mobile users. Such infrastructure is characterized by frequent handoffs and low wireless bandwidth. Satellite networks are of paramount importance in today’s internet communication. They provide connectivity for ships [6], planes, users in less populated areas and in times of disaster. Traditional geostationary (GEO) satellites are widely used despite the rising of new technologies such as Low Earth Orbit (LEO) satellites. Staying at an altitude of 35°786km, GEO satellite communication results in a high round trip time (RTT) of about 600 ms leading, coupled with local packet losses, to drastic goodput degradation. PEPs, which we introduce in the following are, among other approaches [7], deployed to tackle this problem.

2.2. Performance Enhancing Proxies

PEPs are intermediary nodes on the communication path designed to improve the performance of TCP [3]. They can operate at link layer, transport layer and application layer. At transport layer, PEPs split the TCP connection in two separate connections: The first one is terminated at the PEP and the new one is established
from the PEP to the server or another PEP. They can be deployed as integrated PEPs: A single PEP splitting the connection or distributed PEPs: Two PEPs isolate the link between them. Connection splitting leads to a shorter RTT for both connections which increases the TCP sender transmission rate as the time to get acknowledgements gets shorter. It also enables the PEP to apply congestion and error control tailored for the specific needs of the network segment providing significant performance gains in wireless networks [3]. Despite their wide deployment [8] and advantages, connection-splitting PEPs violate the end-to-end semantics of TCP and contribute to the ossification of the transport layer [9] [10]: Their functionality is based on assumptions about TCP headers and the protocol operations so a new update to the protocol becomes difficult since it requires a modification of their design when they are already deployed in the Internet.

2.3. QUIC

QUIC is a transport protocol originally developed by Google [4] and standardized by the IETF [11]. It is based on the end-to-end encryption of not only the payload but also most of the control information [12]. Some of its main advantages are facilitating the process of rolling out updates due to its user space implementation, stream multiplexing to support multiple streams within a single connection avoiding the head-of-line blocking of TCP, reducing handshake overhead by making use of 0-RTT and circumventing the ossification of the protocol, being based on UDP traffic [4].

3. TCP PEPs

In this section, we will investigate PEPs for TCP. We first succinctly present a number of approaches we do not consider to be recent. We then focus on more recent ideas.

3.1. Overview of PEP Implementations

The benefits of PEPs performing connection splitting have been investigated in multiple environments. XCP-PEP [13] is a transparent PEP that leverages the eXplicit Control Protocol (XCP), developed as a new congestion control mechanism [14] to enhance performance over a satellite link. Experiments showed a fast end-to-end link utilization and a quick fairness convergence when deploying XCP-PEP in high latency networks such as satellite networks. HTTPPEP [15] is another PEP example that provides a faster web browsing experience when using a satellite based network by applying protocol and transport layer optimizations coupled with data compression. In mobile systems such as Long Term Evolution (LTE) networks, performance gains in web browsing and video streaming using PEPs have been observed [16]. We focus on some of the ideas in the following.

3.2. TCP Performance Enhancement over Satellite Networks

Satellite networks pose several challenges to TCP given their characteristics. PEPs were employed among other approaches to improve TCP performance over satellite networks. PEPsal was developed by Caini et al. [17] as the first open source integrated splitting approach available for Linux OS under the GNU GPL license. It is a TCP PEP approach that improves the performance of the satellite connection by employing TCP Hybla, a TCP enhancement scheme specifically designed for the needs of satellite networks [18]. Evaluating PEPsal in the presence of congestion and link losses showed performance gains. Some of the results can be attributed to using TCP Hybla on the satellite segment [17]. Although PEPsal is not a recent TCP PEP, we highlight its relevance and wide usage in research.

3.3. Translation between Network Architectures

The deployment of new network architectures to solve the issues faced by the current TCP/IP architecture due to the exponential growth of data traffic is met by the inflexibility of the internet infrastructure [19]. To circumvent this issue, the idea of translating between the existing architecture and a new one was thoroughly investigated by Ciko et al. [20], introducing a Performance Enhancing Proxy for Deploying Network Architectures (PEP-DNA).

PEP-DNA is the first TCP PEP developed with the goal of enabling sending data from TCP/IP applications along a network path with a different underlying architecture and in the other direction translating traffic to be compatible with the TCP/IP architecture. It is fully implemented in kernel space to be deployed on Linux. It has been tested for the translation in a TCP-TCP connection, between TCP and Recursive InterNetwork Architecture (RINA) [21] as well as between TCP and an Information Centric Networking (ICN) architecture [22].

The performance evaluation of PEP-DNA showed its efficiency and scalability achieving good throughput with low CPU and memory utilization. These results make the deployment of new network architectures a realistic objective. Future work should investigate the possibility of PEP-DNA supporting other protocols besides TCP, rendering the proxying explicit to allow a consent based translation and using dynamic proxying mechanisms [20].

3.4. Multi-Domain Congestion Control

Applying different congestion control mechanisms optimized for the different domains in hybrid networks leads to a performance improvement. The concept of Multi-Domain Congestion Control (MDCC) [23] could be achieved through the deployment of transparent connection splitting PEPs that apply an appropriate Congestion Control (CC) for each connection. However, traditional PEPs introduce a processing overhead and contribute to the ossification of the transport layer. Approaches applying end-to-end encryption render transparent PEPs impractical. In order to deploy PEPs without violating the end-to-end semantics, Middlebox Cooperation Protocols (MCPs) [24] propose the idea of middleboxes sharing explicit useful information to the endpoints that can be safely ignored. Based on this concept, Mihály et al. [25] developed a lightweight PEP (LwPEP) that supports MDCC, overcoming the additional communication and processing...
overhead and circumventing the problem of transport ossification without modification at the client side. LwPEP enabled the design of a MDCC showing performance gains in LTE [25] and mmWave 5G Networks [26]. LwPEP supports both TCP and QUIC traffic [26].

4. QUIC PEPs

![Figure 1: MASQUE proxy setup with QUIC tunneling (27)](image)

We first start by addressing the fundamental question of whether exploring the possibilities for QUIC PEPs is justifiable in terms of potential performance gains. Kosek et al. [5] investigated this question for satellite networks using both GEO and LEO satellites, comparing QUIC vs. TCP and HTTP/3 vs HTTP/1.1 with and without PEPs.

Evaluating the performance regarding goodput over time showed a better performance when using PEPs for both GEO and LEO scenarios with different link loss rates, especially with higher RTTs and lower loss rates. Analysing the web performance also demonstrated a real performance gain using QUIC PEPs in GEO orbits. These promising results motivate the further exploration of QUIC PEPs. However, the used implementation in [5] is a proof-of-concept: The developed proxy for connection splitting operates on clear data, thereby violating the principle of end-to-end encryption. We therefore present the following practical possibilities for QUIC PEPs with the common goal of preserving end-to-end encryption. For each discussed concept, we highlight the important points, give an overview of the implementation and discuss the limitations.

4.1. Explicit User-consent based Proxying

**Motivation:** As transparent PEPs are incompatible with QUIC, explicit proxying should be investigated. In this context, we introduce Multiplexed Application Substrate over QUIC Encryption (MASQUE) [28]. It is a QUIC proxying protocol that defines a new HTTP CONNECT extension to enable a tunneled QUIC connection between a client and a proxy, as illustrated by Figure 1: An end-to-end connection over the proxy encapsulated in an outer QUIC tunnel connection. By using MASQUE, the client can explicitly request forwarding of UDP and IP traffic towards a specific server. The proxy receives packets from the client wrapped in an outer QUIC tunnel connection, unwraps them and sends them to the target server. The proxy operates in the other direction as well by encapsulating received packets from the server and relaying them to the client. MASQUE replaces transparent PEPs that operate transparently with explicit user-consent based proxying [28] and provides web proxying without interception and end-to-end security problems which some HTTP based explicit proxies deployed in access networks might suffer from [29].

**Implementation:** MASQUE is still under development. The IETF MASQUE working group developed two specifications for defining a new method to proxy UDP in HTTP. The first is using CONNECT UP HTTP method [30] to create a tunnel to a proxy server over the HTTP request stream to enable sending tunneled UDP payloads using HTTP datagrams to the proxy. The second specification [31] describes two ways of data transmission: using QUIC datagrams [32] for unreliable data transfer called _datagram mode_ or using datagrams encoded as CAPSULE frames [30], a new HTTP type frame, for reliable transmission called _stream mode_.

**Evaluation:** The impact of using QUIC based MASQUE proxying on QUIC performance has been studied by Kühlewind et al. in [28] based on a modified version of aioquic, a QUIC and HTTP/3 python implementation supporting HTTP datagrams and CAPSULE frames for HTTP/3 and modifiable packet sizes and congestion control for QUIC. Experiments show a reduction in overhead and transmission time for increasing packet sizes which increases performance. Investigating the impact of RTTs and nested congestion control shows that with increased RTT, lower transmission times are generally observed in datagram mode compared to stream mode with the Reno CC algorithm showing better performance than Cubic or no CC. A significant improvement in transmission time is witnessed in stream mode for high loss rates with low delays on the link between the client and the proxy advocating the use of reliable streams in MASQUE proxying for lossy local links [28].

**Discussion:** The evaluation of MASQUE proxying suggests the potential loss recovery benefits of using stream mode for reliable data transmissions over wireless networks characterized by high non-congestive loss rates. Their work introduces the possibility of having simple link layer loss recovery mechanisms while offering explicit reliable data transfer with loss recovery when an application needs it. However, these promising results are limited by the used python based setup which is not optimized for performance. Further work using different stacks and emulation setups is required for better performance evaluation. Furthermore, security problems have to be considered when employing MASQUE proxying since it does not inspect the sent packets, therefore opening the possibility of malicious packets and replay attacks.

4.2. Combining Performance and Security

**Motivation:** GEO satellites suffer from long delays which negatively impact the _performance_ and security of the satellite link. To deal with this issue, PEPs are employed by ISPs. The problem with PEPs is that they operate on clear data, thus raising security concerns for the satellite link. On the other hand, employing
approaches that support encryption, for instance VPN, is incompatible with PEP as the latter have to inspect clear TCP headers. The trade-off between performance and security is the crux of the problem in satellite communication motivating the development of QPEP by Pavur et al. [33]. QPEP is a new hybrid protocol between traditional PEP and VPN aiming to enhance the performance of secure satellite traffic making use of QUIC.

Figure 2: Simplified QPEP distributed architecture [33])

**Implementation**: Figure 2 illustrates the basic QPEP architecture. It is implemented as a transparent distributed PEP. The client establishes a persistent QUIC tunnel with the server for multiplexed and encrypted QUIC connection and maps incoming TCP connections to unique QUIC streams. It selectively terminates TCP connections discarding spurious ACKs and sends only relevant data converted to QUIC packets via the tunnel. We note that currently there is only an available TCP/IP stack implementation for tunneling [33].

**Evaluation**: QPEP was tested in simulations in [33] showing better performance than traditional PEPs. These promising results motivated the evaluation of its performance in a real world environment over commercially available satellite links by Huwyler et al. [34]. The metrics for measurements are goodput by measuring the download speed at the client side with varying payload sizes and Page Load Time (PLT) as an important metric for good user web experience.

An evaluation of different scenarios was conducted: Plain connection with neither performance enhancement nor security support, PEPsal as the only publicly available PEP, OpenVPN as an encrypted protocol and QPEP. Regarding goodput, QPEP outperformed all scenarios in a setup where the GEO provider proprietary PEP was not activated. Activating it did not enhance QPEP throughput in contrast to the other scenarios. Overall, QPEP attained higher goodput than OpenVPN especially in case of low payload sizes. The differences in PLTs between the different scenarios were less pronounced in the real testbed compared to the simulated one. QPEP outperformed OpenVPN in the setup with activated provider PEP improving PLT. Without it QPEP demonstrated the best performance among all protocols. Overall, QPEP proved its performance benefits over traditional PEPs and OpenVPN.

**Discussion**: The results obtained are limited by the "black box" nature of the provider networks causing certain ambiguities in the explanation of QPEP performance in specific scenarios. Further parameter tuning for the satellite link could help gain more insight into the performance benefits of QPEP [35]. Further work regarding the collaboration with different GEO providers and testing QPEP for LEO networks should be pursued.

4.3. Secure Middlebox Insertion in QUIC Connections

**Motivation**: QUIC with its end-to-end encryption opposes the idea of splitting the end-to-end connection to inspect or modify the exchanged information for performance enhancement purposes. A possible QUIC enhancement idea consists in the selective and controllable exposure of information to intermediary nodes in the network to allow the conscious insertion of middleboxes by the endpoints without violating the security of the communication. This idea is introduced by Kosek et al. as Secure Middlebox Assisted QUIC (SMAQ) [36].

**Implementation and Evaluation of SMAQ**: SMAQ inserts middleboxes that preserve the end-to-end security aspects while simultaneously being capable of adjusting certain functional aspects of QUIC to improve the performance. As illustrated in Figure 3 a state handover mechanism enables the endpoint to share its protocol state with an inserted middlebox by sharing the necessary keying material. This enables splitting of the end-to-end connection in two independent connections using an enhanced QUIC connection migration mechanism coupled with an added encryption layer during the QUIC handshake. SMAQ also supports the use of multiple middleboxes through transitive state handover. We direct the reader to [36] for a detailed design description.

In the same paper, a study case was conducted with SMAQ to use distributed PEPs in a QUIC satellite connection with both GEO and LEO satellite orbits. To evaluate the performance, a PEP optimized SMAQ using Hybla-Westwood [18] [37] as CC algorithm was compared to an end-to-end QUIC connection using NewReno w.r.t Bulk Download measured by the bytes received by the client over several time intervals and web performance measured by PLT. The results showed an overhead of a bit more than one RTT for SMAQ-PEP connection setup, an increase of bytes received with higher loss compared to normal QUIC and an overall better performance with higher RTT, loss rate, and byte transfer sizes.

**Discussion**: Using SMAQ raises important security concerns. Its design is based on the assumption that middleboxes are trusted to manipulate and modify the connection violating the principles of privacy, integrity and authenticity. The current design only supports state handover started by the client and only during QUIC handshake which limits the potential performance benefits. Future work addressing these limitations should be pursued.
5. Conclusion and Future Work

Performance enhancing proxies are an important tool to enhance TCP performance in hybrid networks. In this survey we first investigate current PEP approaches for TCP. Although the number of open source PEP implementations is limited, we were able to find novel approaches in the current research. We investigate TCP performance enhancement over satellite networks presenting an older approach widely used in research. We discuss the idea of translating between different network architectures to enable the deployment of new architectures. We also introduce novel approaches for Multi-Domain Congestion Control. We then focus on exploring PEP possibilities for QUIC. We discuss interesting concepts such as explicit user-consent based proxying, combining the conflicting goals of performance and security in satellite networks, inserting middleboxes in a QUIC connection without violating the end-to-end security. Overall, the discussed ideas show promising performance gains for QUIC. However, they are limited by the used implementations that raise problems of scalability and security. Future work should further investigate these ideas while dealing with the mentioned limitations.

References


[19] D. John, “Patterns in network architecture: a return to fundamen-


