Cooperative Performance Enhancement Using QUIC Tunneling in 5G Cellular Networks

Zsolt Krämer
Budapest University of Technology and Economics
Budapest, Hungary
kramer@tmit.bme.hu

Mirja Kühlewind, Marcus Ihlar, Attila Mihály
<fistname>.<lastname>@ericsson.com
Ericsson

ABSTRACT

Multiplexed Application Substrate over QUIC Encryption (MASQUE) is a new protocol mechanism that is currenty under standardization in the IETF. MASQUE defines an extension to the HTTP CONNECT method in order to support QUIC-based tunneling and forwarding of UDP and IP traffic. In this paper we discuss use cases for a MASQUE-based proxy setup that addresses challenges in performance optimization in cellular networks. The presented use cases realize different services based on the supported level of cooperation between the three involved parties, i.e., the client, the proxy, and the target server.

CCS CONCEPTS

• **Networks** → **Transport protocols**; **Middle boxes / network appliances**.

ACM Reference Format:
Zsolt Krämer and Mirja Kühlewind, Marcus Ihlar, Attila Mihály. 2021. Cooperative Performance Enhancement Using QUIC Tunneling in 5G Cellular Networks. In *Applied Networking Research Workshop (ANRW '21), July24–30, 2021, Virtual Event, USA.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3472305.3472320

1 INTRODUCTION

When the end-to-end network path consists of a wired- and a wireless domain with different characteristics, it is challenging for the congestion control in the end-to-end transport protocol to optimize for any specific characteristics as it is unknown where the bottleneck is. This motivated the deployment of Performance Enhancing Proxies (PEPs) in cellular networks. These proxies split the end-to-end TCP connection and provide two separate, shorter control loops for the two domains. 5G networks are going to provide very high peak data rates (around 10Gbps) and significantly decreased delay, however, due to the propagation properties of the new mmWave radio in 5G, the available bandwidth can also show high volatility. These factors intensify the need for a shorter control loop for congestion control and local optimization in 5G networks [5]. QUIC [1], a new and inherently encrypted transport protocol, which is being increasingly deployed instead of TLS over TCP, makes traditional connection splitting performance optimizations impossible without breaking end-to-end encryption, and as such requires new approaches to realize similar services as used by currently deployed in-network traffic management solutions.

In this paper, we discuss a cooperative approach to enable network-assisted performance enhancements for encrypted transport protocols. Rather than intercepting any connection at the proxy, with our cooperative approach using MASQUE as the signaling protocol towards the proxy, there are now two layers of connections: a tunnel connection between the proxy and the client and the end-to-end connection between the client and the target server with an own, unmodified end-to-end security context that guarantees confidentiality, source authentication, and integrity between the endpoints. As such, using QUIC-based tunneling also establishes a secure communication channel between the mobile terminal and the proxy as shown in Figure 1. It provides an opportunity to offer additional services like faster loss recovery by the proxy or exposure of up-to-date network information that can be used to assist congestion control. Separation of communication channel data and tunnel data is achieved by establishing separate QUIC streams or datagram flows, within the same connection. This new design approach, where explicit consent in requesting a service is required, enables proxy services without breaking the end-to-end principle of the transport and application layers.

2 USE CASES

2.1 Local loss recovery

If the QUIC tunnel between the client and MASQUE proxy covers a local high loss link, QUIC streams can be used which provide reliable and in-order delivery. Even though loss of IP

ANRW '21, July 24–30, 2021, Virtual Event, USA

ACM ISBN 978-1-4503-8618-0/21/07.
https://doi.org/10.1145/3472305.3472320

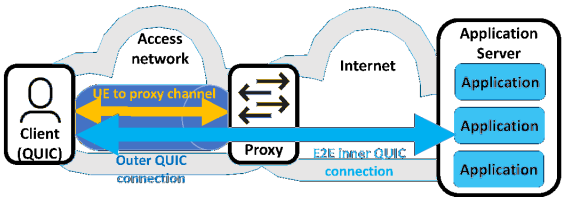

Figure 1: Proxy setup using QUIC tunneling.

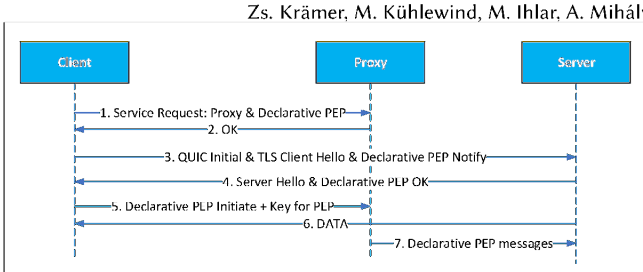

Figure 2: Declarative PEP negotiation/key exchange

packets is a rare event in LTE and 5G access networks when operated in acknowledged mode, the loss can still occur due to buffer overflow, reordering or late loss due to excessive retransmission attempts. Furthermore, efficient transport layer loss recovery has potential to improve performance over LTE and 5G radio using unacknowledged mode instead and leaving the decision about the need of local recovery to the endpoints.

2.2 Promise Signaling

In a situation where the currently available link capacity is limited between the proxy and the client, a promise signal can be used to indicate the reception of an end-to-end QUIC packet by the proxy to the client. If the promise signal was requested by the client and the proxy receives data from the server at a higher rate than it is transmitting towards the client, it appends information of which packets it has received to "promise" delivery in near future. The client then uses this information to identify which packet of the inner end-to-end connection are already successfully received and currently buffered at the proxy and provides this information to the server, by either progressively acknowledging the data or indicating to the server that delays are not to be interpreted as congestion, if this is supported by the server. In both cases the server will not interpret delays as a sign of congestion but keeps sending at the current rate. That allows the proxy to buffer sufficient packets in order to quickly send them out when more capacity becomes available again, similar as split-TCP PEPs operate today.

2.3 Declarative Messages to the Server

Our collaborative proxy approach can further be utilized to enable a trusted and secure communication with the server. One option is to similarly establish a QUIC tunnel to the server, if explicitly supported by the server, or alternatively a more light-weight approach is possible by sending encrypted declarative messages to the server which provide additional information about the network status at the proxy but are safe-to-ignore. Fig. 2 depicts a possible realization how the declarative message service and the respective keys can be negotiated.

Similar as for the promise signaling, this services can be realized by sending sufficient data of the encrypted QUIC header to cover the packet sequence number [3] or alternatively the client directly shares the Header Protection Key of the end-to-end QUIC connection between the client and server. This enables the proxy to decrypt the packet number of packets sent towards the client in order to only generate ACKs or even NACKs, which reduce the overhead significantly. Further, this would simplify key management as the same key could be used for the declarative signaling to the server. This service could then also be realized on an opportunistic basis without additional client-server communication. One specific realization of such a declarative proxy is the Lightweight PEP (LwPEP) concept introduced in [4]. By utilizing both the ACK/NACK information generated by the proxy as well as the regular end-to-end ACKs, the server can apply a Multi-Domain congestion control (MD CC) scheme as also described in [2] that supports fair link sharing in the Internet domain but enables the use of a more aggressive congestion control in the cellular domain.

3 CONCLUSIONS

The paper shows the flexibility of the cooperative approach achieving performance enhancement by presenting three mechanisms with different modes of cooperation between the client, the proxy, and the target server, i.e., local loss recovery, promise signaling, and sending of declarative messages to the server. Moreover, if an explicit tunnel connection is also supported to the server, another approach for performance optimisation can be realised by implementing domain specific congestion control on each side of the proxy while disabling the end-to-end congestion control. Our initial testing have confirmed that the presented approaches can achieve comparable or better performance compared to transparent TCP splitting PEPs as used today. In addition, they are realized based on explicit consensus of both endpoints enabling even more opportunities for collaboration and optimisation in the future with better security.

REFERENCES

[1] Janardhan Iyengar and Martin Thomson. 2020. QUIC: A UDP-Based Multiplexed and Secure Transport. In *IETF Internet-Draft*.

[2] Zsolt Krämer, Sándor Molnár, Attila Mihály, and Szilveszter Nádas. 2019. Towards multi-domain congestion control in next-generation networks. In *2019 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 1–7.

[3] Zsolt Krämer, Sándor Molnár, Marcus Pieska, and Attila Mihály. 2020. A Lightweight Performance Enhancing Proxy for Evolved Protocols and Networks. In *2020 IEEE 24th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD), in press*. IEEE.

[4] Attila Mihály, Szilveszter Nádas, Sándor Molnár, Zsolt Krämer, Robert Skog, and Marcus Ihlar. 2017. Supporting multi-domain congestion control by a lightweight PEP. In *Internet of Things, Embedded Systems and Communications (IINTEC), 2017 International Conference on*. IEEE, 105–110.

[5] Menglei Zhang, Michele Polese, Marco Mezzavilla, Jing Zhu, Sundeep Rangan, Shivendra Panwar, and Michele Zorzi. 2019. Will TCP work in mmWave 5G cellular networks? *IEEE Communications Magazine* 57, 1 (2019), 65–71.