Network Working Group
Internet-Draft
Intended status: Informational
Expires: September 3, 2018

X. de Foy
M. Perras
InterDigital Communications
U. Chunduri
Huawei USA
K. Nguyen
M. Kibria
K. Ishizu
F. Kojima
NICT

March 2, 2018

Considerations for MPTCP operation in 5G
draft-defoy-mptcp-considerations-for-5g-00

Abstract

This document describes scenarios where the behavior of the 5G mobility management framework is different from earlier systems, and may benefit from some form of adaptation of MPTCP implementations and/or the 5G system.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on September 3, 2018.

Table of Contents

1. Introduction

MPTCP [RFC6824] is being deployed and widely adopted in today's smart devices, which typically have multiple network interfaces such as Cellular and Wifi. It provides reliability, bandwidth aggregation capability, and handover efficiency.

This document describes scenarios where the behavior of the 5G mobility management framework is different from earlier systems, and may benefit from some form of adaptation of MPTCP implementations and/or the 5G system.

2. Impact of 5G Session and Service Continuity on MPTCP

One of the goals of 5G [3GPP.23.501] is to enable low latency in some use cases. Mobility in the Evolved Packet System (EPS) was based on a central mobility solution which could hinder that goal, and therefore 5G uses a distributed mobility solution based on multiple anchors providing different IP addresses as the device moves from one area to another.

The base scenario in this section is: a 5G device connected to a single-homed server is in an area with no usable Wifi coverage. An application using MPTCP sends traffic over a single subflow, over the cellular air interface. Then, as the device moves, the 5G device reacts to mobility events. Additionally, we also discuss briefly

cases where switching from Wifi to cellular backup, and cases where both MPTCP peers are 5G mobile devices.

In 5G, every unit of network service (PDU session) can have an IP (IPv4 or IPv6), Ethernet or unstructured type. While session continuity is supported for all types, we will focus on IP-type PDU sessions primarily. Different PDU sessions will typically correspond to distinct network interfaces on the device (though this is not explicit in the standard, and some implementations may possibly behave differently).

In the EPS, session continuity was enabled by having the P-GW and IP address of the mobile device's PDU session maintained over time, even when the device moved around. In 5G, different types of session continuity can be provided, and are indicated by a "Session and Service Continuity" (SSC) mode value of 1, 2 or 3 (defined in [_3GPP.23.501] section 5.6.9). Every PDU session is associated with a single SSC mode, which cannot be changed on this PDU session. The following sub-sections will study how 5G handles each SSC mode, and potential effects on MPTCP.

In 5G multiple applications running on a device may end up using different PDU sessions (e.g. different network interfaces), for example because they require different network slices or SSC modes. It is therefore important for an MPTCP implementation to be aware of which network interfaces are available to which local applications. This mapping information will be known to the 5G stack, and can be made available to an MPTCP implementation running on the device.

2.1. SSC mode 1

In SSC mode 1 the same network anchor is kept regardless of device location. An application running on the device will therefore be able to keep using the same IP address on the same interface.

Additionally, in SSC mode 1, the network may decide to add and remove, dynamically, additional network anchors (and therefore IP addresses) to the PDU session, while always keeping the initial one.

The MPTCP stack will therefore be able to create new subflows and benefit from a potentially shorter path, when the device is far from its initial network anchor, with the caveats that those additional subflows will be available on a temporary basis only. MPTCP must not close the initial subflow in this SSC mode, since this is the only one guaranteed to be maintained over time.

2.2. SSC mode 2

SSC mode 2 has a break-before-make behavior. When the device leaves the service area of its first network anchor, the network stops using it and starts using a new second network anchor closer to the device. (Such service areas may have a highly variable size depending on network deployments.) On the device, this can result in the currently used network interface being brought down, and after a short time a new network interface being brought up. The time between these 2 events is not standardized and implementation dependent.

Break-before-make within cellular technology

When MPTCP is active on cellular only, this break-before-make behavior is similar to the existing break-before-make capability usually used in cellular/Wifi offload (section 3.1.3 of [RFC6897] and section 2.2 of [RFC8041]). A similar MPTCP behavior is therefore needed: wait for a given time for a new IP address to be configured. As per [RFC6897], to benefit from this MPTCP resilience feature, the application should not request using a specific network interface.

Cellular and Wifi

Additionally, when Wifi is active and cellular is used as backup, MPTCP implementations should also support this break-before-make behavior to maintain a usable backup IP address on cellular. In rare cases where a switch-to-cellular backup would be needed during a break-before-make transition on cellular, MPTCP's existing break-before-make capability can ensure MPTCP waits for a new cellular-facing IP address to be available.

2.3. SSC mode 3

SSC mode 3 has a make-before-break behavior. When the device leaves the service area of its first network anchor, the network selects a second network anchor closer to the device, and either creates a new PDU session (i.e. new IP address on new network interface) or share the existing PDU session (i.e. new IP address on same network interface). The first network anchor keeps being used for a given time period, which is communicated to the device by the network using the "valid lifetime" field of a prefix information option in a router advertisement ([RFC4861], [RFC4862]). 5G does not mandate a specific range for this valid lifetime. The first/older IP address should not be used to create any new traffic ([RFC4862] section 5.5.4). In some implementations, the network (SMF) may decide to release the first network anchor as soon as it stops carrying traffic.

There is no limit set by the 5G standard for the number of concurrently used network anchors. We expect that in usual cases the first network anchor will be released before a third network anchor starts being used. Nevertheless, to our knowledge nothing prevents a 5G system deployment to allow a third network anchor to be selected while the first one is still in use.

On the 5G device, when using SSC mode 3, mobility will therefore result in a new IP address being configured, either on the same network interface initially used, or on a different interface. In general an application will see a single cellular-facing IP address, and during transient phase it will see 2 IP addresses (with a possibility for more than 2 concurrent IP addresses on some 5G system implementations). In cases where the server is single-homed and the Wifi interface is down, and assuming a full-mesh path manager policy, there will be in general one subflow, and from time to time, temporarily 2 subflows (or more on some 5G systems). In cases where two mobile 5G devices are communicating with each other over MPTCP and with the same assumptions on Wifi and path manager policy, there will be in general one subflow, and from time to time, temporarily 2 or even more rarely 4 subflows (again, possibly more on some 5G systems).

MPTCP must create new subflows when a new IP address on a same or a new cellular-facing network interface becomes available to the application. MPTCP may keep using the first subflow during a transient phase. Here are some considerations related to this transient phase:

- When compared with simply waiting for the first IP address to be brought down, ramping down usage of the first subflow will not incur inefficiencies from resending lost segments. This may especially help low-latency applications by avoiding throughput drop.

- Assuming a lowest-rtt-first scheduling policy is used, after the initial TCP slow start, the shortest path subflow should typically carry the most traffic. Ramping down should ideally start after the initial slow start is over.

- To make sure the ramping down completes before the interface is brought down by the network, the MPTCP stack should be aware of how long will the first network anchor be kept in use, e.g. through configuration or communication with the local 5G stack.

- Ramping down and closing flows on the first network anchor as soon as possible will help recycling network resources more rapidly.

This is especially true in cases where more than 2 network anchors may be used concurrently.

- There may be some level of contention between subflows during the transient phase, since they share the same air interface, and especially if they share the same PDU session and QoS marking. The shortest path subflow may therefore not reach its full capacity during the transient phase.

- Additionally, the shortest subflow must not be closed during the transient phase (even if it is less efficient for some reason), to avoid losing all connectivity at the end of the transient phase. To avoid this issue, the MPTCP stack could for example follow a policy not to close any subflow created using the latest IP address, during the transient period (in SSC mode 3).

In cases where cellular is used for backup, there is a possibility that the switch to using backup occurs during a transient phase. To support this case, MPTCP should keep creating and releasing subflows as described above, even when cellular subflows are used as backup, to ensure that the backup is always usable. When a backup event occurs during a transient phase, MPTCP should use the subflows associated with the most recent cellular-facing IP address, i.e. corresponding to the latest/closest network anchor.

3. MPTCP with 5G Dual Connectivity

One of the key features of 5G [_3GPP.23.501] is dual connectivity (DC). With DC, a 5G device can be served by two different base stations. DC may play an essential role in leveraging the benefit of 5G new radio, especially in the evolving architecture with the coexistence of 4G and 5G radios.

On a 5G device with DC, an application is able to send data to the destination (e.g., a single-home server) through multiple radio links, over one or more PDU sessions. Some PDU sessions may be over a single radio link, while others may have flows over each radio link. Therefore, in a first case, DC can be made visible to applications through different IP addresses, while in a second case, DC can be used by different flows terminated at the same IP address on the device.

In any of those cases, the issues of out of order delivery and diverse latency values need to be supported in DC. However, such reliable communication scenarios have not been addressed in the current DC architecture. Based on the design history of DC in earlier systems, the 5G system will need to incorporate features to support robustness/reliability (e.g. backup and duplication), that

will likely result in added complexity. On the other hand, in order to benefit the most from DC, the 5G device is expected to send/ receive an optimal amount of traffic over each radio link (e.g., for the sake of minimizing overall latency). Hence, the device needs to select dynamically the most suitable path for a given radio condition. Additionally, algorithms for shifting, based on congestion, ongoing traffic between paths are also necessary.

MPTCP, which includes path manager, scheduler, and congestion control functions, shows a lot of potential to address the aforementioned issues. MPTCP could therefore be integrated with DC and the 5G protocol stack, as an alternative to developing 5G-specific solutions. As part of this integration, the MPTCP stack should be aware of the presence of multiple radio links, whether they are exposed using multiple IP addresses or hidden under a single IP address. MPTCP's scheduler should optimally partition traffic or duplicate a data flow over different links, depending on the application's need, network policy and conditions.

4. Summary of Requirements and Conclusion

With regards to 5G session continuity mechanism, MPTCP stack behavior (including path manager and scheduling) should be updated to achieve optimal performance. As a summary:

- MPTCP should obtain information from the local 5G stack (SSC mode, mapping between interfaces and applications, valid lifetime on first network anchor in SSC mode3)

- In SSC mode 3 during the transient period following a mobility event, MPTCP should gracefully stop using old cellular-facing interface(s), and must not release subflow(s) using the latest cellular-facing IP address.

- In SSC mode 1 MPTCP must not close the initial subflow.

- When cellular is used as backup, MPTCP should actively maintain the backup path in SSC mode 2 and 3.

With regards to dual connectivity, MPTCP can be closely integrated with the 5G stack to avoid duplicating its feature in 5G. As a summary:

- MPTCP should be aware of the presence of multiple DC radio links, which may be exposed as a single or distinct network interfaces/IP addresses.

- MPTCP should optimally partition traffic or duplicate a data flow over DC links, depending on the application's need, network policy and conditions.

5. IANA Considerations

This document requests no IANA actions.

6. Security Considerations

No new security considerations are identified at this time.

7. Acknowledgements

The following people contributed to the present document:

- Debashish Purkayastha
- Akbar Rahman
- Ulises Olvera-Hernandez

8. Informative References

[_3GPP.23.501] 3GPP, "System Architecture for the 5G System", 3GPP TS 23.501 1.4.0, 9 2017, <http://www.3gpp.org/ftp/Specs/html-info/23501.htm>.

[RFC4861] Narten, T., Nordmark, E., Simpson, W., and H. Soliman, "Neighbor Discovery for IP version 6 (IPv6)", RFC 4861, DOI 10.17487/RFC4861, September 2007, <https://www.rfc-editor.org/info/rfc4861>.

[RFC4862] Thomson, S., Narten, T., and T. Jinmei, "IPv6 Stateless Address Autoconfiguration", RFC 4862, DOI 10.17487/RFC4862, September 2007, <https://www.rfc-editor.org/info/rfc4862>.

[RFC6824] Ford, A., Raiciu, C., Handley, M., and O. Bonaventure, "TCP Extensions for Multipath Operation with Multiple Addresses", RFC 6824, DOI 10.17487/RFC6824, January 2013, <https://www.rfc-editor.org/info/rfc6824>.

[RFC6897] Scharf, M. and A. Ford, "Multipath TCP (MPTCP) Application Interface Considerations", RFC 6897, DOI 10.17487/RFC6897, March 2013, <https://www.rfc-editor.org/info/rfc6897>.

[RFC8041] Bonaventure, O., Paasch, C., and G. Detal, "Use Cases and Operational Experience with Multipath TCP", RFC 8041, DOI 10.17487/RFC8041, January 2017, <https://www.rfc-editor.org/info/rfc8041>.

Authors' Addresses

Xavier de Foy
InterDigital Communications, LLC
1000 Sherbrooke West
Montreal
Canada

Email: Xavier.Defoy@InterDigital.com

Michelle Perras
InterDigital Communications, LLC
Montreal
Canada

Email: Michelle.Perras@InterDigital.com

Uma Chunduri
Huawei USA
2330 Central Expressway
Santa Clara, CA 95050
USA

Email: uma.chunduri@huawei.com

Kien Nguyen
National Institute of Information and Communications Technology
YRP Center No.1 Building 7F, 3-4 Hikarinooka, Yokosuka
Kanagawa 239-0847
Japan

Email: kienng@nict.go.jp

Mirza Golam Kibria
National Institute of Information and Communications Technology
YRP Center No.1 Building 7F, 3-4 Hikarinooka, Yokosuka
Kanagawa 239-0847
Japan

Email: mirza.kibria@nict.go.jp

Kentaro Ishizu
National Institute of Information and Communications Technology
YRP Center No.1 Building 7F, 3-4 Hikarinooka, Yokosuka
Kanagawa 239-0847
Japan

Email: ishidu@nict.go.jp

Fumihide Kojima
National Institute of Information and Communications Technology
YRP Center No.1 Building 7F, 3-4 Hikarinooka, Yokosuka
Kanagawa 239-0847
Japan

Email: f-kojima@nict.go.jp