INTAREA J. Zhu
Internet Draft Intel
Intended status: Standards Track S. Seo
Expires: April 1,2020 Korea Telecom
S. Kanugovi
Nokia
S. Peng
Huawei
October 1, 2019

User-Plane Protocols for Multiple Access Management Service
draft-zhu-intarea-mams-user-protocol-08

Status of this Memo

This Internet-Draft will expire on April 1,2020.

Copyright Notice

Abstract

Today, a device can be simultaneously connected to multiple communication networks based on different technology implementations and network architectures like WiFi, LTE, and DSL. In such multi-connectivity scenario, it is desirable to combine multiple access networks or select the best one to improve quality of experience for a user and improve overall network utilization and efficiency. This document presents the u-plane protocols for a multi access management services (MAMS) framework that can be used to flexibly select the combination of uplink and downlink access and core network paths having the optimal performance, and user plane treatment for improving network utilization and efficiency and enhanced quality of experience for user applications.

Table of Contents

1. Introduction

Multi Access Management Service (MAMS) [MAMS] is a programmable framework to select and configure network paths, as well as adapt to dynamic network conditions, when multiple network connections can serve a client device. It is based on principles of user plane interworking that enables the solution to be deployed as an overlay without impacting the underlying networks.

This document presents the u-plane protocols for enabling the MAMS framework. It co-exists and complements the existing protocols by providing a way to negotiate and configure the protocols based on client and network capabilities. Further it allows exchange of network state information and leveraging network intelligence to optimize the performance of such protocols. An important goal for MAMS is to ensure that there is minimal or no dependency on the actual access technology of the participating links. This allows the scheme to be scalable for addition of newer access technologies and for independent evolution of the existing access technologies.

2. Terminologies

Anchor Connection: refers to the network path from the N-MADP to the Application Server that corresponds to a specific IP anchor that has assigned an IP address to the client.

Delivery Connection: refers to the network path from the N-MADP to the C-MADP.

"Network Connection Manager" (NCM), "Client Connection Manager" (CCM), "Network Multi Access Data Proxy" (N-MADP), and "Client Multi Access Data Proxy" (C-MADP) in this document are to be interpreted as described in [MAMS].

3. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

The terminologies "Network Connection Manager" (NCM), "Client Connection Manager" (CCM), "Network Multi Access Data Proxy" (N-MADP), and "Client Multi Access Data Proxy" (C-MADP) in this document are to be interpreted as described in [MAMS].

4 MAMS User-Plane Protocols

Figure 1 shows the MAMS u-plane protocol stack as specified in [MAMS].

```
             +-----------------------------------------------------+
             |      User Payload (e.g. IP PDU)                     |
             |-----------------------------------------------------|
          +--|-----------------------------------------------------|--+
          |  |-----------------------------------------------------|  |
          |  | Multi-Access (MX) Convergence Sublayer              |  |
          |  |-----------------------------------------------------|  |
          |  |-----------------------------------------------------|  |
          |  | MX Adaptation   | MX Adaptation  | MX Adaptation    |  |
          |  | Sublayer        | Sublayer       | Sublayer         |  |
          |  | (optional)      | (optional)     | (optional)       |  |
          |  |-----------------------------------------------------|  |
          |  | Access #1 IP    | Access #2 IP   | Access #3 IP     |  |
          |  +-----------------------------------------------------+  |
          +-----------------------------------------------------------+
```

Figure 1: MAMS U-plane Protocol Stack

It consists of the following two Sublayers:

- o Multi-Access (MX) Convergence Sublayer: This layer performs multi-access specific tasks, e.g., access (path) selection, multi-link (path) aggregation, splitting/reordering, lossless switching, fragmentation, concatenation, keep-alive, and probing etc.
- o Multi-Access (MX) Adaptation Sublayer: This layer performs functions to handle tunneling, network layer security, and NAT.

The MX convergence sublayer operates on top of the MX adaptation sublayer in the protocol stack. From the Transmitter perspective, a User Payload (e.g. IP PDU) is processed by the convergence sublayer first, and then by the adaptation sublayer before being transported over a delivery access connection; from the Receiver perspective, an IP packet received over a delivery connection is processed by the MX adaptation sublayer first, and then by the MX convergence sublayer.

4.1 MX Adaptation Sublayer

The MX adaptation sublayer supports the following mechanisms and protocols while transmitting user plane packets on the network path:

- o UDP Tunneling: The user plane packets of the anchor connection can be encapsulated in a UDP tunnel of a delivery connection between the N-MADP and C-MADP.

o IPsec Tunneling: The user plane packets of the anchor connection are sent through an IPsec tunnel of a delivery connection.
o Client Net Address Translation (NAT): The Client IP address of user plane packet of the anchor connection is changed, and sent over a delivery connection.
o Pass Through: The user plane packets are passing through without any change over the anchor connection.

The MX adaptation sublayer also supports the following mechanisms and protocols to ensure security of user plane packets over the network path.

o IPsec Tunneling: An IPsec [RFC7296] tunnel is established between the N-MADP and C-MADP on the network path that is considered untrusted.
o DTLS: If UDP tunneling is used on the network path that is considered "untrusted", DTLS (Datagram Transport Layer Security) [RFC6347] can be used.

The Client NAT method is the most efficient due to no tunneling overhead. It SHOULD be used if a delivery connection is "trusted" and without NAT function on the path.

The UDP or IPsec Tunnelling method SHOULD be used if a delivery connection has a NAT function placed on the path.

4.2 GMA-based MX Convergence Sublayer

Figure 2 shows the MAMS u-plane protocol stack based on trailer-based encapsulation [GMA]. Multiple access networks are combined into a single IP connection. If NCM determines that N-MADP is to be instantiated with GMA as the MX Convergence Protocol, it exchanges the support of GMA convergence capability in the discovery and capability exchange procedures [MAMS].

```
+-----------------------------------------------------+
|                       IP PDU                        |
|-----------------------------------------------------|
|              GMA  Convergence Sublayer              |
|-----------------------------------------------------|
| MX Adaptation    | MX Adaptation | MX Adaptation    |
| Sublayer         | Sublayer      | Sublayer         |
| (optional)       | (optional)    | (optional)       |
|-----------------------------------------------------|
| Access #1 IP     | Access #2 IP  | Access #3 IP     |
+-----------------------------------------------------+
```
Figure 2: MAMS U-plane Protocol Stack with GMA as MX Convergence Layer

Figure 3 shows the trailer-based Multi-Access (MX) PDU (Protocol Data Unit) format [GMA]. If the MX adaptation method is UDP tunneling and "MX header optimization" in the "MX_UP_Setup_Configuration_Request" message [MAMS] is true, the "IP length" and "IP checksum" header fields of the MX PDU SHOULD remain unchanged. Otherwise, they should be updated after adding or removing the GMA trailer in the convergence sublayer.

```
          +-------------------------------------------------------+
          | IP hdr |        IP payload              | GMA Trailer |
          +-------------------------------------------------------+
```

Figure 3: GMA PDU Format

4.3 MPTCP-based MX Convergence Sublayer

Figure 4 shows the MAMS u-plane protocol stack based on MPTCP. Here, MPTCP is reused as the "MX Convergence Sublayer" protocol. Multiple access networks are combined into a single MPTCP connection. Hence, no new u-plane protocol or PDU format is needed in this case.

```
          |-----------------------------------------------------|
          |                         MPTCP                       |
          |-----------------------------------------------------|
          |  TCP             |   TCP          |      TCP        |
          |-----------------------------------------------------|
          | MX Adaptation    | MX Adaptation | MX Adaptation    |
          | Sublayer         | Sublayer      | Sublayer         |
          | (optional)       | (optional)    | (optional)       |
          |-----------------------------------------------------|
          | Access #1 IP     | Access #2 IP  | Access #3 IP     |
          +-----------------------------------------------------+
```

Figure 4: MAMS U-plane Protocol Stack with MPTCP as MX Convergence Layer

If NCM determines that N-MADP is to be instantiated with MPTCP as the MX Convergence Protocol, it exchanges the support of MPTCP capability in the discovery and capability exchange procedures [MAMS]. MPTCP proxy protocols [MPProxy][MPPlain] SHOULD be used to manage traffic steering and aggregation over multiple delivery connections.

4.4 GRE as MX Convergence Sublayer

Figure 5 shows the MAMS u-plane protocol stack based on GRE (Generic Routing Encapsulation) [GRE2784]. Here, GRE is reused as the "MX Convergence sub-layer" protocol. Multiple access networks are combined

into a single GRE connection. Hence, no new u-plane protocol or PDU format is needed in this case.

```
+------------------------------------------------------+
|      User Payload (e.g. IP PDU)                      |
|------------------------------------------------------|
|              GRE as MX Convergence Sublayer          |
|------------------------------------------------------|
|        GRE Delivery Protocol (e.g. IP)               |
|------------------------------------------------------|
| MX Adaptation    | MX Adaptation | MX Adaptation     |
| Sublayer         | Sublayer      | Sublayer          |
| (optional)       | (optional)    | (optional)        |
|------------------------------------------------------|
| Access #1 IP     | Access #2 IP  | Access #3 IP      |
+------------------------------------------------------+
```

Figure 5: MAMS U-plane Protocol Stack with GRE as MX Convergence Layer

If NCM determines that N-MADP is to be instantiated with GRE as the MX Convergence Protocol, it exchanges the support of GRE capability in the discovery and capability exchange procedures [MAMS].

4.4.1 Transmitter Procedures

Transmitter is the N-MADP or C-MADP instance, instantiated with GRE as the convergence protocol that transmits the GRE packets. The Transmitter receives the User Payload (e.g. IP PDU), encapsulates it with a GRE header and Delivery Protocol (e.g. IP) header to generate the GRE Convergence PDU.

When IP is used as the GRE delivery protocol, the IP header information (e.g. IP address) can be created using the IP header of the user payload or a virtual IP address. The "Protocol Type" field of the delivery header is set to 47 (or 0X2F, i.e. GRE)[IANA].

The GRE header fields are set as specified below,

- If the transmitter is a C-MADP instance, then sets the LSB 16 bits to the value of Connection ID for the Anchor Connection associated with the user payload or sets to 0xFFFF if no Anchor Connection ID needs to be specified.
- All other fields in the GRE header including the remaining bits in the key fields are set as per [GRE_2784][GRE_2890].

4.4.2 Receiver Procedures

Receiver is the N-MADP or C-MADP instance, instantiated with GRE as the convergence protocol that receives the GRE packets. The receiver processes the received packets per the GRE procedures [GRE_2784, GRE_2890] and retrieves the GRE header.

- If the Receiver is an N-MADP instance,
    o Unless the LSB 16 Bits of the Key field are 0xFFFF, they are interpreted as the Connection ID of Anchor Connection for the user payload. This is used to identify the network path over which the User Payload (GRE Payload) is to be transmitted.
- All other fields in the GRE header, including the remaining bits in the Key fields, are processed as per [GRE_2784][GRE_2890].

The GRE Convergence PDU is passed onto the MX Adaptation Layer (if present) before delivery over one of the network paths.

4.5 Co-existence of MX Adaptation and MX Convergence Sublayers

MAMS u-plane protocols support multiple combinations and instances of user plane protocols to be used in the MX Adaptation and the Convergence sublayers.

For example, one instance of the MX Convergence Layer can be MPTCP Proxy [MPProxy][MPPlain] and another instance can be Trailer-based. The MX Adaptation for each can be either UDP tunnel or IPsec. IPsec may be set up for network paths considered as untrusted by the operator, to protect the TCP subflow between client and MPTCP proxy traversing that network path.

Each of the instances of MAMS user plane, i.e. combination of MX Convergence and MX Adaptation layer protocols, can coexist simultaneously and independently handle different traffic types.

5. MX Convergence Control Message

A UDP connection may be configured between C-MADP and N-MADP to exchange control messages for keep-alive or path quality estimation. The N-MADP end-point IP address and UDP port number of the UDP connection is used to identify MX control PDU. Figure 6 shows the MX control PDU format with the following fields:

o Type (1 Byte): the type of the MX control message
o CID (1 Byte): an unsigned integer to identify the anchor and delivery connection of the MX control message
    + Anchor Connection ID (MSB 4 Bits): an unsigned integer to identify the anchor connection

   + Delivery Connection ID (LSB 4 Bits): an unsigned integer to identify the delivery connection
- o MX Control Message (variable): the payload of the MX control message

Figure 7 shows the MX convergence control protocol stack, and MX control PDU goes through the MX adaptation sublayer the same way as MX data PDU.

```
                        <----MX Control PDU Payload --------------->
+-------------------------------------------------------------------+
| IP header | UDP Header| Type | CID |       MX Control Message     |
+-------------------------------------------------------------------+
                      Figure 6: MX Control PDU Format

          |------------------------------------------------------|
          |          MX Convergence Control Messages             |
          |------------------------------------------------------|
          |                  UDP/IP                              |
          |------------------------------------------------------|
          | MX Adaptation   | MX Adaptation | MX Adaptation      |
          | Sublayer        | Sublayer      | Sublayer           |
          | (optional)      | (optional)    | (optional)         |
          |------------------------------------------------------|
          | Access #1 IP    | Access #2 IP  | Access #3 IP       |
          +------------------------------------------------------+
              Figure 7: MX Convergence Control Protocol Stack
```

5.1 Keep-Alive Message

The "Type" field is set to "0" for Keep-Alive messages. C-MADP may send out Keep-Alive message periodically over one or multiple delivery connections, especially if UDP tunneling is used as the adaptation method for the delivery connection with a NAT function on the path.

A Keep-Alive message is 6 Bytes long, and consists of the following fields:

- o Keep-Alive Sequence Number (2 Bytes): the sequence number of the keep-alive message
- o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.

5.2 Probe Message

The "Type" field is set to "1" for Probe messages.

N-MADP may send out the Probe message for path quality estimation. In response, C-MADP may send back the ACK message.

A Probe message consists of the following fields:

- o Probing Sequence Number (2 Bytes): the sequence number of the Probe REQ message
- o Probing Flag (1 Byte):
  - + Bit #0: a ACK flag to indicate if the ACK message is expected (1) or not (0);
  - + Bit #1: a Probe Type flag to indicate if the Probe message is sent during the initialization phase (0) when the network path is not included for transmission of user data or the active phase (1) when the network path is included for transmission of user data;
  - + Bit #2: a bit flag to indicate the presence of the Reverse Connection ID (R-CID) field.
  - + Bit #3~7: reserved
- o Reverse Connection ID (1 Byte): the connection ID of the delivery connection for sending out the ACK message on the reverse path
- o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.
- o Padding (variable)

The "R-CID" field is only present if both Bit #0 and Bit #2 of the "Probing Flag" field are set to "1". Moreover, Bit #2 of the "Probing Flag" field SHOULD be set to "0" if the Bit #0 is "0", indicating the ACK message is not expected.

If the "R-CID" field is not present but the Bit #0 of the "Probing Flag" field is set to "1", the ACK message SHOULD be sent over the same delivery connection as the Probe message.

The "Padding" field is used to control the length of Probe message.

5.3 Packet Loss Report (PLR) Message

The "Type" field is set to "2" for PLR messages.

C-MADP may send out the PLR messages to report lost MX SDU for example during handover. In response, C-MADP may retransmit the lost MX SDU accordingly.

A PLR message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection which the ACK message is for;

- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the anchor connection which the ACK message is for;
- o ACK number (4 Bytes): the next (in-order) sequence number (SN) that the sender of the PLR message is expecting
- o Number of Loss Bursts (1 Byte)
  For each loss burst, include the following
  - + Sequence Number of the first lost MX SDU in a burst (4 Bytes)
  - + Number of consecutive lost MX SDUs in the burst (1 Byte)

```
          C-MADP                                              N-MADP
              |                                                  |
              |<------------------- MX SDU (data packets)--------|
              |                                                  |
             +-----------------------------------+               |
             |Packet Loss detected               |               |
             +-----------------------------------+               |
              |                                                  |
              |----- PLR Message ------------------------------->|
              |<--------------retransmit(lost)MX SDUs -----------|
```

Figure 8: MAMS Retransmission Procedure

Figure 8 shows the MAMS retransmission procedure in an example where the lost packet is found and retransmitted.

5.4 First Sequence Number (FSN) Message

The "Type" field is set to "3" for FSN messages.

N-MADP may send out the FSN messages to indicate the oldest MX SDU in its buffer if a lost MX SDU is not found in the buffer after receiving the PLR message from C-MADP. In response, C-MADP SHALL only report packet loss with SN not smaller than FSN.

A FSN message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection which the FSN message is for;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the anchor connection which the FSN message is for;
- o First Sequence Number (4 Bytes): the sequence number (SN) of the oldest MX SDU in the (retransmission) buffer of the sender of the FSN message.

Figure 9 shows the MAMS retransmission procedure in an example where the lost packet is not found.

```
          C-MADP                                               N-MADP
              |                                                  |
              |<------------------ MX SDU (data packets)--------|
              |                                                  |
             +----------------------------------+                |
             |Packet Loss detected              |                |
             +----------------------------------+                |
              |                                                  |
              |----- PLR Message ------------------------------>|
              |                              +---------------------+
              |                              |Lost packet not found|
              |                              +---------------------+
              |<-------------FSN message -----------------------|
```

Figure 9: MAMS Retransmission Procedure with FSN

5.5 Coded MX SDU (CMS) Message

The "Type" field is set to "4" for CMS messages.

N-MADP (or C-MADP) may send out the CMS message to support downlink (or uplink) packet loss recovery through coding, e.g. [CRLNC], [CTCP], [RLNC]. A coded MX SDU is generated by applying a network coding algorithm to multiple consecutive (uncoded) MX SDUs, and it is used for fast recovery without retransmission if any of the MX SDUs is lost.

A Coded MX SDU message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection of the coded MX SDU;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the coded MX;
- o Sequence Number (4 Bytes): the sequence number of the first (uncoded) MX SDU used to generate the coded MX SDU.
- o Fragmentation Control (FC) (1 Byte): to provide necessary information for re-assembly, only needed if the coded MX SDU is too long to transport in a single MX control PDU.
- o N (1 Byte): the number of consecutive MX SDUs used to generate the coded MX SDU
- o K (1 Byte): the length (in terms of bits) of the coding coefficient field
- o Coding Coefficient (N x K / 8 Bytes)
  - + a(i): the coding coefficient of the i-th (uncoded) MX SDU

     + padding
- o Coded MX SDU (variable): the coded MX SDU

If K = 0, the simple XOR method is used to generate the Coded MX SDU from N consecutive uncoded MX SDUs, and the a(i) fields are not included in the message.

If the coded MX SDU is too long, it can be fragmented, and transported by multiple MX control PDUs. The N, K, and a(i) fields are only included in the MX PDU carrying the first fragment of the coded MX SDU.

```
          C-MADP                                              N-MADP
              |                                                  |
              |<------------------ MX SDU #1 -------------------|
              |      lost<-------- MX SDU #2 -------------------|
              |<---- CMS Message (MX SDU #1 XOR MX SDU #2)------|
            +-----------------------+                            |
            | MX SDU #2 recovered  |                             |
            +-----------------------+                            |
              |                                                  |
```

Figure 10: MAMS Packet Recovery Procedure with XOR Coding

5.6 Traffic Splitting Update (TSU) Message

The "Type" field is set to "5" for TSU messages.

N-MADP (or C-MADP) may send out a TSU message if downlink (or uplink) traffic splitting configuration has changed.

A TSU message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class;
- o Sequence Number (2 Bytes): an unsigned integer to identify the TSU message.
- o Flags (1 Byte)
  - + Bit #0: a Reverse Path bit flag to indicate if the traffic splitting configuration is for the reverse path (1) or not (0);
  - + Bit #1: a Bit-Reversal bit flag to indicate if bit-reversal is used in traffic splitting
  - + Others: reserved.
- o Traffic Splitting Configuration Parameters (5 + (N -1) Bytes):

+ StartSN (4 Bytes): the sequence number of the first MX SDU using the traffic splitting configuration provided by the TSU message
+ L (1 Byte): the traffic splitting burst size
+ K(i): the traffic splitting threshold of the i-th delivery connection, where connections are ordered according to their Connection ID.

Let’s use f(x) to denote the traffic splitting function, which maps a MX SDU Sequence Number "x" to the i-th delivery connection.

    f(x)=i, if K[i-1]< or = mod(x - StartSN, L) < K[i]

Wherein, 1 < or = i < N, K[0]=0, and K[N]=L.

N is the total number of connections for delivering a data flow, identified by (anchor) Connection ID and Traffic Class ID.

When the bit-reversal bit is set to 1, the burst size L MUST be a power of 2, and the traffic splitting function is

    f(x)=i, if K[i-1]< or = F(mod(x - StartSN, L)) < K[i]

Wherein F(.) is the bit reversal function [BITR] of the input variable.

5.7 Acknowledgement Message

The "Type" field is set to "6" for ACK messages.

C-MADP (or N-MADP) SHOULD send out the ACK message in response to the successful reception of a PLR, FSN, or TSU message.

C-MADP SHOULD send out the ACK message in response to a Probe message with the ACK flag set to "1".

The ACK message consists of the following fields:

- o Acknowledgment Number (2 Bytes): the sequence number of the received message.
- o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.

6 Security Considerations

User data in MAMS framework rely on the security of the underlying network transport paths. When this cannot be assumed, NCM configures use of appropriate protocols for security, e.g. IPsec [RFC4301] [RFC3948], DTLS [RFC6347].

7 IANA Considerations

This draft makes no requests of IANA.

8 Contributing Authors

The editors gratefully acknowledge the following additional contributors in alphabetical order: Salil Agarwal/Nokia, Hema Pentakota/Nokia.

9 References

9.1 Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC4301] Kent, S. and K. Seo, "Security Architecture for the Internet Protocol", RFC 4301, DOI10.17487/RFC4301, December 2005, <http://www.rfc-editor.org/info/rfc4301>.

9.2 Informative References

[RFC6347] Rescorla, E. and N. Modadugu, "Datagram Transport Layer Security Version 1.2", RFC 6347, January 2012, <http://www.rfc-editor.org/info/rfc6347>.

[RFC7296] Kaufman, C., Hoffman, P., Nir, Y., Eronen, P., and T. Kivinen, "Internet Key Exchange Protocol Version 2 (IKEv2)", STD 79, RFC 7296, DOI 10.17487/RFC7296, October 2014, <http://www.rfc-editor.org/info/rfc7296>.

[RFC3948] Huttunen, A., Swander, B., Volpe, V., DiBurro, L., and M. Stenberg, "UDP Encapsulation of IPsec ESP Packets", RFC 3948, DOI 10.17487/RFC3948, January 2005, <http://www.rfc-editor.org/info/rfc3948>.

[MPProxy] X. Wei, C. Xiong, and E. Lopez, "MPTCP proxy mechanisms", https://tools.ietf.org/html/draft-wei-mptcp-proxy-mechanism-02

[MPPlain] M. Boucadair et al, "An MPTCP Option for Network-Assisted MPTCP", https://www.ietf.org/id/draft-boucadair-mptcp-plain-mode-09.txt

[MAMS] S. Kanugovi, S. Vasudevan, F. Baboescu, and J. Zhu, "Multiple Access Management Protocol", https://tools.ietf.org/html/draft-kanugovi-intarea-mams-protocol-03

[GMA] J. Zhu, "Trailer-based Encapsulation Protocols for Generic Multi-Access Convergence", https://tools.ietf.org/html/draft-zhu-intarea-gma-01

[GRE2784] D. Farinacci, et al., "Generic Routing Encapsulation (GRE)", RFC 2784 March 2000, <http://www.rfc-editor.org/info/rfc2784>.

[GRE2890] G. Dommety, "Key and Sequence Number Extensions to GRE", RFC 2890 September 2000, <http://www.rfc-editor.org/info/rfc2890>.

[IANA] https://www.iana.org/assignments/protocol-numbers/protocol-numbers.xhtml

[LWIPEP] 3GPP TS 36.361, "Evolved Universal Terrestrial Radio Access (E-UTRA); LTE-WLAN Radio Level Integration Using Ipsec Tunnel (LWIP) encapsulation; Protocol specification"

[RFC791] Internet Protocol, September 1981

[CRLNC] S Wunderlich, F Gabriel, S Pandi, et al. Caterpillar RLNC (CRLNC): A Practical Finite Sliding Window RLNC Approach, IEEE Access, 2017

[CTCP] M. Kim, et al. Network Coded TCP (CTCP), eprint arXiv:1212.2291, 2012

[RLNC] J. Heide, et al. Random Linear Network Coding (RLNC)-Based Symbol Representation, https://www.ietf.org/id/draft-heide-nwcrg-rlnc-00.txt

[BITR] Alan H. Karp, "Bit reversal on uniprocessors", SIAM Review, 38 (1): 1-26, 1996.

Authors' Addresses

Jing Zhu

Intel

Email: jing.z.zhu@intel.com

SungHoon Seo

Korea Telecom

Email: sh.seo@kt.com

Satish Kanugovi

Nokia

Email: satish.k@nokia.com

Shuping Peng

Huawei

Email: pengshuping@huawei.com

INTAREA J. Zhu
Internet Draft Intel
Intended status: Standards Track S. Seo
Expires: April 1,2020 Korea Telecom
S. Kanugovi
Nokia
S. Peng
Huawei
October 1, 2019

User-Plane Protocols for Multiple Access Management Service

draft-zhu-intarea-mams-user-protocol-07

Status of this Memo

Copyright Notice

Abstract

Today, a device can be simultaneously connected to multiple communication networks based on different technology implementations and network architectures like WiFi, LTE, and DSL. In such multi-connectivity scenario, it is desirable to combine multiple access networks or select the best one to improve quality of experience for a user and improve overall network utilization and efficiency. This document presents the u-plane protocols for a multi access management services (MAMS) framework that can be used to flexibly select the combination of uplink and downlink access and core network paths having the optimal performance, and user plane treatment for improving network utilization and efficiency and enhanced quality of experience for user applications.

Table of Contents

1. Introduction

Multi Access Management Service (MAMS) [MAMS] is a programmable framework to select and configure network paths, as well as adapt to dynamic network conditions, when multiple network connections can serve a client device. It is based on principles of user plane interworking that enables the solution to be deployed as an overlay without impacting the underlying networks.

This document presents the u-plane protocols for enabling the MAMS framework. It co-exists and complements the existing protocols by providing a way to negotiate and configure the protocols based on client and network capabilities. Further it allows exchange of network state information and leveraging network intelligence to optimize the performance of such protocols. An important goal for MAMS is to ensure that there is minimal or no dependency on the actual access technology of the participating links. This allows the scheme to be scalable for addition of newer access technologies and for independent evolution of the existing access technologies.

2. Terminologies

Anchor Connection: refers to the network path from the N-MADP to the Application Server that corresponds to a specific IP anchor that has assigned an IP address to the client.

Delivery Connection: refers to the network path from the N-MADP to the C-MADP.

"Network Connection Manager" (NCM), "Client Connection Manager" (CCM), "Network Multi Access Data Proxy" (N-MADP), and "Client Multi Access Data Proxy" (C-MADP) in this document are to be interpreted as described in [MAMS].

3. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

The terminologies "Network Connection Manager" (NCM), "Client Connection Manager" (CCM), "Network Multi Access Data Proxy" (N-MADP), and "Client Multi Access Data Proxy" (C-MADP) in this document are to be interpreted as described in [MAMS].

4 MAMS User-Plane Protocols

Figure 1 shows the MAMS u-plane protocol stack as specified in [MAMS].

```
          +-----------------------------------------------------+
          |      User Payload (e.g. IP PDU)                     |
          |-----------------------------------------------------|
       +--|-----------------------------------------------------|--+
       |  |-----------------------------------------------------|  |
       |  | Multi-Access (MX) Convergence Sublayer              |  |
       |  |-----------------------------------------------------|  |
       |  |-----------------------------------------------------|  |
       |  | MX Adaptation   | MX Adaptation | MX Adaptation     |  |
       |  | Sublayer        | Sublayer      | Sublayer          |  |
       |  | (optional)      | (optional)    | (optional)        |  |
       |  |-----------------------------------------------------|  |
       |  | Access #1 IP    | Access #2 IP  | Access #3 IP      |  |
       |  +-----------------------------------------------------+  |
       +-----------------------------------------------------------+
```

Figure 1: MAMS U-plane Protocol Stack

It consists of the following two Sublayers:

- o Multi-Access (MX) Convergence Sublayer: This layer performs multi-access specific tasks, e.g., access (path) selection, multi-link (path) aggregation, splitting/reordering, lossless switching, fragmentation, concatenation, keep-alive, and probing etc.
- o Multi-Access (MX) Adaptation Sublayer: This layer performs functions to handle tunneling, network layer security, and NAT.

The MX convergence sublayer operates on top of the MX adaptation sublayer in the protocol stack. From the Transmitter perspective, a User Payload (e.g. IP PDU) is processed by the convergence sublayer first, and then by the adaptation sublayer before being transported over a delivery access connection; from the Receiver perspective, an IP packet received over a delivery connection is processed by the MX adaptation sublayer first, and then by the MX convergence sublayer.

4.1 MX Adaptation Sublayer

The MX adaptation sublayer supports the following mechanisms and protocols while transmitting user plane packets on the network path:

- o UDP Tunneling: The user plane packets of the anchor connection can be encapsulated in a UDP tunnel of a delivery connection between the N-MADP and C-MADP.

o IPsec Tunneling: The user plane packets of the anchor connection are sent through an IPsec tunnel of a delivery connection.
o Client Net Address Translation (NAT): The Client IP address of user plane packet of the anchor connection is changed, and sent over a delivery connection.
o Pass Through: The user plane packets are passing through without any change over the anchor connection.

The MX adaptation sublayer also supports the following mechanisms and protocols to ensure security of user plane packets over the network path.

o IPsec Tunneling: An IPsec [RFC7296] tunnel is established between the N-MADP and C-MADP on the network path that is considered untrusted.
o DTLS: If UDP tunneling is used on the network path that is considered "untrusted", DTLS (Datagram Transport Layer Security) [RFC6347] can be used.

The Client NAT method is the most efficient due to no tunneling overhead. It SHOULD be used if a delivery connection is "trusted" and without NAT function on the path.

The UDP or IPsec Tunnelling method SHOULD be used if a delivery connection has a NAT function placed on the path.

4.2 GMA-based MX Convergence Sublayer

Figure 2 shows the MAMS u-plane protocol stack based on trailer-based encapsulation [GMA]. Multiple access networks are combined into a single IP connection. If NCM determines that N-MADP is to be instantiated with GMA as the MX Convergence Protocol, it exchanges the support of GMA convergence capability in the discovery and capability exchange procedures [MAMS].

```
          +-----------------------------------------------------+
          |                        IP PDU                       |
          |-----------------------------------------------------|
          |               GMA  Convergence Sublayer             |
          |-----------------------------------------------------|
          | MX Adaptation   | MX Adaptation | MX Adaptation     |
          | Sublayer        | Sublayer      | Sublayer          |
          | (optional)      | (optional)    | (optional)        |
          |-----------------------------------------------------|
          | Access #1 IP    | Access #2 IP  | Access #3 IP      |
          +-----------------------------------------------------+
```
Figure 2: MAMS U-plane Protocol Stack with GMA as MX Convergence Layer

Figure 3 shows the trailer-based Multi-Access (MX) PDU (Protocol Data Unit) format [GMA]. If the MX adaptation method is UDP tunneling and "MX header optimization" in the "MX_UP_Setup_Configuration_Request" message [MAMS] is true, the "IP length" and "IP checksum" header fields of the MX PDU SHOULD remain unchanged. Otherwise, they should be updated after adding or removing the GMA trailer in the convergence sublayer.

```
          +--------------------------------------------------------+
          | IP hdr |         IP payload              | GMA Trailer |
          +--------------------------------------------------------+
                         Figure 3: GMA PDU Format
```

4.3 MPTCP-based MX Convergence Sublayer

Figure 4 shows the MAMS u-plane protocol stack based on MPTCP. Here, MPTCP is reused as the "MX Convergence Sublayer" protocol. Multiple access networks are combined into a single MPTCP connection. Hence, no new u-plane protocol or PDU format is needed in this case.

```
          |-----------------------------------------------------|
          |                        MPTCP                        |
          |-----------------------------------------------------|
          |  TCP             |   TCP           |      TCP       |
          |-----------------------------------------------------|
          | MX Adaptation    | MX Adaptation | MX Adaptation    |
          | Sublayer         | Sublayer      | Sublayer         |
          | (optional)       | (optional)    | (optional)       |
          |-----------------------------------------------------|
          | Access #1 IP     | Access #2 IP  | Access #3 IP     |
          +-----------------------------------------------------+
```

Figure 4: MAMS U-plane Protocol Stack with MPTCP as MX Convergence Layer

If NCM determines that N-MADP is to be instantiated with MPTCP as the MX Convergence Protocol, it exchanges the support of MPTCP capability in the discovery and capability exchange procedures [MAMS]. MPTCP proxy protocols [MPProxy][MPPlain] SHOULD be used to manage traffic steering and aggregation over multiple delivery connections.

4.4 GRE as MX Convergence Sublayer

Figure 5 shows the MAMS u-plane protocol stack based on GRE (Generic Routing Encapsulation) [GRE2784]. Here, GRE is reused as the "MX Convergence sub-layer" protocol. Multiple access networks are combined

into a single GRE connection. Hence, no new u-plane protocol or PDU format is needed in this case.

```
          +------------------------------------------------------+
          |      User Payload (e.g. IP PDU)                      |
          |------------------------------------------------------|
          |              GRE as MX Convergence Sublayer          |
          |------------------------------------------------------|
          |        GRE Delivery Protocol (e.g. IP)               |
          |------------------------------------------------------|
          | MX Adaptation    | MX Adaptation | MX Adaptation     |
          | Sublayer         | Sublayer      | Sublayer          |
          | (optional)       | (optional)    | (optional)        |
          |------------------------------------------------------|
          | Access #1 IP     | Access #2 IP  | Access #3 IP      |
          +------------------------------------------------------+
```

Figure 5: MAMS U-plane Protocol Stack with GRE as MX Convergence Layer

If NCM determines that N-MADP is to be instantiated with GRE as the MX Convergence Protocol, it exchanges the support of GRE capability in the discovery and capability exchange procedures [MAMS].

4.4.1 Transmitter Procedures

Transmitter is the N-MADP or C-MADP instance, instantiated with GRE as the convergence protocol that transmits the GRE packets. The Transmitter receives the User Payload (e.g. IP PDU), encapsulates it with a GRE header and Delivery Protocol (e.g. IP) header to generate the GRE Convergence PDU.

When IP is used as the GRE delivery protocol, the IP header information (e.g. IP address) can be created using the IP header of the user payload or a virtual IP address. The "Protocol Type" field of the delivery header is set to 47 (or 0X2F, i.e. GRE)[IANA].

The GRE header fields are set as specified below,

- If the transmitter is a C-MADP instance, then sets the LSB 16 bits to the value of Connection ID for the Anchor Connection associated with the user payload or sets to 0xFFFF if no Anchor Connection ID needs to be specified.
- All other fields in the GRE header including the remaining bits in the key fields are set as per [GRE_2784][GRE_2890].

4.4.2 Receiver Procedures

Receiver is the N-MADP or C-MADP instance, instantiated with GRE as the convergence protocol that receives the GRE packets. The receiver processes the received packets per the GRE procedures [GRE_2784, GRE_2890] and retrieves the GRE header.

- If the Receiver is an N-MADP instance,
  - o Unless the LSB 16 Bits of the Key field are 0xFFFF, they are interpreted as the Connection ID of Anchor Connection for the user payload. This is used to identify the network path over which the User Payload (GRE Payload) is to be transmitted.
- All other fields in the GRE header, including the remaining bits in the Key fields, are processed as per [GRE_2784][GRE_2890].

The GRE Convergence PDU is passed onto the MX Adaptation Layer (if present) before delivery over one of the network paths.

4.5 Co-existence of MX Adaptation and MX Convergence Sublayers

MAMS u-plane protocols support multiple combinations and instances of user plane protocols to be used in the MX Adaptation and the Convergence sublayers.

For example, one instance of the MX Convergence Layer can be MPTCP Proxy [MPProxy][MPPlain] and another instance can be Trailer-based. The MX Adaptation for each can be either UDP tunnel or IPsec. IPsec may be set up for network paths considered as untrusted by the operator, to protect the TCP subflow between client and MPTCP proxy traversing that network path.

Each of the instances of MAMS user plane, i.e. combination of MX Convergence and MX Adaptation layer protocols, can coexist simultaneously and independently handle different traffic types.

5. MX Convergence Control Message

A UDP connection may be configured between C-MADP and N-MADP to exchange control messages for keep-alive or path quality estimation. The N-MADP end-point IP address and UDP port number of the UDP connection is used to identify MX control PDU. Figure 6 shows the MX control PDU format with the following fields:

- o Type (1 Byte): the type of the MX control message
- o CID (1 Byte): an unsigned integer to identify the anchor and delivery connection of the MX control message
  - + Anchor Connection ID (MSB 4 Bits): an unsigned integer to identify the anchor connection

       + Delivery Connection ID (LSB 4 Bits): an unsigned integer to
         identify the delivery connection
  o MX Control Message (variable): the payload of the MX control
    message

Figure 7 shows the MX convergence control protocol stack, and MX control PDU goes through the MX adaptation sublayer the same way as MX data PDU.

```
                        <----MX Control PDU Payload --------------->
+--------------------------------------------------------------------+
| IP header | UDP Header| Type | CID |       MX Control Message     |
+--------------------------------------------------------------------+
                      Figure 6: MX Control PDU Format
```

```
          |-----------------------------------------------------|
          |           MX Convergence Control Messages           |
          |-----------------------------------------------------|
          |                   UDP/IP                            |
          |-----------------------------------------------------|
          | MX Adaptation   | MX Adaptation | MX Adaptation     |
          | Sublayer        | Sublayer      | Sublayer          |
          | (optional)      | (optional)    | (optional)        |
          |-----------------------------------------------------|
          | Access #1 IP    | Access #2 IP  | Access #3 IP      |
          +-----------------------------------------------------+
              Figure 7: MX Convergence Control Protocol Stack
```

5.1 Keep-Alive Message

The "Type" field is set to "0" for Keep-Alive messages. C-MADP may send out Keep-Alive message periodically over one or multiple delivery connections, especially if UDP tunneling is used as the adaptation method for the delivery connection with a NAT function on the path.

A Keep-Alive message is 6 Bytes long, and consists of the following fields:

  o Keep-Alive Sequence Number (2 Bytes): the sequence number of the
    keep-alive message
  o Timestamp (4 Bytes): the current value of the timestamp clock of
    the sender in the unit of 100 microseconds.

5.2 Probe Message

The "Type" field is set to "1" for Probe messages.

N-MADP may send out the Probe message for path quality estimation. In response, C-MADP may send back the ACK message.

A Probe message consists of the following fields:

- o Probing Sequence Number (2 Bytes): the sequence number of the Probe REQ message
- o Probing Flag (1 Byte):
    - + Bit #0: a ACK flag to indicate if the ACK message is expected (1) or not (0);
    - + Bit #1: a Probe Type flag to indicate if the Probe message is sent during the initialization phase (0) when the network path is not included for transmission of user data or the active phase (1) when the network path is included for transmission of user data;
    - + Bit #2: a bit flag to indicate the presence of the Reverse Connection ID (R-CID) field.
    - + Bit #3~7: reserved
- o Reverse Connection ID (1 Byte): the connection ID of the delivery connection for sending out the ACK message on the reverse path
- o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.
- o Padding (variable)

The "R-CID" field is only present if both Bit #0 and Bit #2 of the "Probing Flag" field are set to "1". Moreover, Bit #2 of the "Probing Flag" field SHOULD be set to "0" if the Bit #0 is "0", indicating the ACK message is not expected.

If the "R-CID" field is not present but the Bit #0 of the "Probing Flag" field is set to "1", the ACK message SHOULD be sent over the same delivery connection as the Probe message.

The "Padding" field is used to control the length of Probe message.

5.3 Packet Loss Report (PLR) Message

The "Type" field is set to "2" for PLR messages.

C-MADP may send out the PLR messages to report lost MX SDU for example during handover. In response, C-MADP may retransmit the lost MX SDU accordingly.

A PLR message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection which the ACK message is for;

- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the anchor connection which the ACK message is for;
- o ACK number (4 Bytes): the next (in-order) sequence number (SN) that the sender of the PLR message is expecting
- o Number of Loss Bursts (1 Byte)
  For each loss burst, include the following
  - + Sequence Number of the first lost MX SDU in a burst (4 Bytes)
  - + Number of consecutive lost MX SDUs in the burst (1 Byte)

```
          C-MADP                                            N-MADP
              |                                                 |
              |<------------------ MX SDU (data packets)--------|
              |                                                 |
             +----------------------------------+               |
             |Packet Loss detected              |               |
             +----------------------------------+               |
              |                                                 |
              |----- PLR Message ------------------------------>|
              |<-------------retransmit(lost)MX SDUs -----------|
```

Figure 8: MAMS Retransmission Procedure

Figure 8 shows the MAMS retransmission procedure in an example where the lost packet is found and retransmitted.

5.4 First Sequence Number (FSN) Message

The "Type" field is set to "3" for FSN messages.

N-MADP may send out the FSN messages to indicate the oldest MX SDU in its buffer if a lost MX SDU is not found in the buffer after receiving the PLR message from C-MADP. In response, C-MADP SHALL only report packet loss with SN not smaller than FSN.

A FSN message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection which the FSN message is for;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the anchor connection which the FSN message is for;
- o First Sequence Number (4 Bytes): the sequence number (SN) of the oldest MX SDU in the (retransmission) buffer of the sender of the FSN message.

Figure 9 shows the MAMS retransmission procedure in an example where the lost packet is not found.

```
          C-MADP                                               N-MADP
              |                                                  |
              |<------------------ MX SDU (data packets)--------|
              |                                                  |
             +----------------------------------+                |
             |Packet Loss detected              |                |
             +----------------------------------+                |
              |                                                  |
              |----- PLR Message ------------------------------>|
              |                              +---------------------+
              |                              |Lost packet not found|
              |                              +---------------------+
              |<-------------FSN message -----------------------|
```

Figure 9: MAMS Retransmission Procedure with FSN

5.5 Coded MX SDU (CMS) Message

The "Type" field is set to "4" for CMS messages.

N-MADP (or C-MADP) may send out the CMS message to support downlink (or uplink) packet loss recovery through coding, e.g. [CRLNC], [CTCP], [RLNC]. A coded MX SDU is generated by applying a network coding algorithm to multiple consecutive (uncoded) MX SDUs, and it is used for fast recovery without retransmission if any of the MX SDUs is lost.

A Coded MX SDU message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection of the coded MX SDU;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the coded MX;
- o Sequence Number (4 Bytes): the sequence number of the first (uncoded) MX SDU used to generate the coded MX SDU.
- o Fragmentation Control (FC) (1 Byte): to provide necessary information for re-assembly, only needed if the coded MX SDU is too long to transport in a single MX control PDU.
- o N (1 Byte): the number of consecutive MX SDUs used to generate the coded MX SDU
- o K (1 Byte): the length (in terms of bits) of the coding coefficient field
- o Coding Coefficient (N x K / 8 Bytes)
    - + a(i): the coding coefficient of the i-th (uncoded) MX SDU

+ padding
o Coded MX SDU (variable): the coded MX SDU

If K = 0, the simple XOR method is used to generate the Coded MX SDU from N consecutive uncoded MX SDUs, and the a(i) fields are not included in the message.

If the coded MX SDU is too long, it can be fragmented, and transported by multiple MX control PDUs. The N, K, and a(i) fields are only included in the MX PDU carrying the first fragment of the coded MX SDU.

```
          C-MADP                                             N-MADP
             |                                                  |
             |<------------------ MX SDU #1 -------------------|
             |      lost<-------- MX SDU #2 -------------------|
             |<---- CMS Message (MX SDU #1 XOR MX SDU #2)------|
           +-----------------------+                            |
           | MX SDU #2 recovered  |                            |
           +-----------------------+                            |
             |                                                  |
```

Figure 10: MAMS Packet Recovery Procedure with XOR Coding

5.6 Traffic Splitting Update (TSU) Message

The "Type" field is set to "5" for TSU messages.

N-MADP (or C-MADP) may send out a TSU message if downlink (or uplink) traffic splitting configuration has changed.

A TSU message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class;
- o Sequence Number (2 Bytes): an unsigned integer to identify the TSU message.
- o Flags (1 Byte)
  - + Bit #0: a Reverse Path bit flag to indicate if the traffic splitting configuration is for the reverse path (1) or not (0);
  - + Bit #1: a Bit-Reversal bit flag to indicate if bit-reversal is used in traffic splitting
  - + Others: reserved.
- o Traffic Splitting Configuration Parameters (5 + (N -1) Bytes):

+ StartSN (4 Bytes): the sequence number of the first MX SDU using the traffic splitting configuration provided by the TSU message
+ L (1 Byte): the traffic splitting burst size
+ K(i): the traffic splitting threshold of the i-th delivery connection, where connections are ordered according to their Connection ID.

Let's use f(x) to denote the traffic splitting function, which maps a MX SDU Sequence Number "x" to the i-th delivery connection.

f(x)=i, if K[i-1]< or = mod(x - StartSN, L) < K[i]

Wherein, 1 < or = i < N, K[0]=0, and K[N]=L.

N is the total number of connections for delivering a data flow, identified by (anchor) Connection ID and Traffic Class ID.

When the bit-reversal bit is set to 1, the burst size L MUST be a power of 2, and the traffic splitting function is

f(x)=i, if K[i-1]< or = F(mod(x - StartSN, L)) < K[i]

Wherein F(.) is the bit reversal function [BITR] of the input variable.

5.7 Acknowledgement Message

The "Type" field is set to "6" for ACK messages.

C-MADP (or N-MADP) SHOULD send out the ACK message in response to the successful reception of a PLR, FSN, or TSU message.

C-MADP SHOULD send out the ACK message in response to a Probe message with the ACK flag set to "1".

The ACK message consists of the following fields:

- o Acknowledgment Number (2 Bytes): the sequence number of the received message.
- o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.

6 Security Considerations

User data in MAMS framework rely on the security of the underlying network transport paths. When this cannot be assumed, NCM configures use of appropriate protocols for security, e.g. IPsec [RFC4301] [RFC3948], DTLS [RFC6347].

7 IANA Considerations

This draft makes no requests of IANA.

8 Contributing Authors

The editors gratefully acknowledge the following additional contributors in alphabetical order: Salil Agarwal/Nokia, Hema Pentakota/Nokia.

9 References

9.1 Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC4301] Kent, S. and K. Seo, "Security Architecture for the Internet Protocol", RFC 4301, DOI10.17487/RFC4301, December 2005, <http://www.rfc-editor.org/info/rfc4301>.

9.2 Informative References

[RFC6347] Rescorla, E. and N. Modadugu, "Datagram Transport Layer Security Version 1.2", RFC 6347, January 2012, <http://www.rfc-editor.org/info/rfc6347>.

[RFC7296] Kaufman, C., Hoffman, P., Nir, Y., Eronen, P., and T. Kivinen, "Internet Key Exchange Protocol Version 2 (IKEv2)", STD 79, RFC 7296, DOI 10.17487/RFC7296, October 2014, <http://www.rfc-editor.org/info/rfc7296>.

[RFC3948] Huttunen, A., Swander, B., Volpe, V., DiBurro, L., and M. Stenberg, "UDP Encapsulation of IPsec ESP Packets", RFC 3948, DOI 10.17487/RFC3948, January 2005, <http://www.rfc-editor.org/info/rfc3948>.

[MPProxy] X. Wei, C. Xiong, and E. Lopez, "MPTCP proxy mechanisms", https://tools.ietf.org/html/draft-wei-mptcp-proxy-mechanism-02

[MPPlain] M. Boucadair et al, "An MPTCP Option for Network-Assisted MPTCP", https://www.ietf.org/id/draft-boucadair-mptcp-plain-mode-09.txt

[MAMS] S. Kanugovi, S. Vasudevan, F. Baboescu, and J. Zhu, "Multiple Access Management Protocol", https://tools.ietf.org/html/draft-kanugovi-intarea-mams-protocol-03

[GMA] J. Zhu, "Trailer-based Encapsulation Protocols for Generic Multi-Access Convergence", https://tools.ietf.org/html/draft-zhu-intarea-gma-01

[GRE2784] D. Farinacci, et al., "Generic Routing Encapsulation (GRE)", RFC 2784 March 2000, <http://www.rfc-editor.org/info/rfc2784>.

[GRE2890] G. Dommety, "Key and Sequence Number Extensions to GRE", RFC 2890 September 2000, <http://www.rfc-editor.org/info/rfc2890>.

[IANA] https://www.iana.org/assignments/protocol-numbers/protocol-numbers.xhtml

[LWIPEP] 3GPP TS 36.361, "Evolved Universal Terrestrial Radio Access (E-UTRA); LTE-WLAN Radio Level Integration Using Ipsec Tunnel (LWIP) encapsulation; Protocol specification"

[RFC791] Internet Protocol, September 1981

[CRLNC] S Wunderlich, F Gabriel, S Pandi, et al. Caterpillar RLNC (CRLNC): A Practical Finite Sliding Window RLNC Approach, IEEE Access, 2017

[CTCP] M. Kim, et al. Network Coded TCP (CTCP), eprint arXiv:1212.2291, 2012

[RLNC] J. Heide, et al. Random Linear Network Coding (RLNC)-Based Symbol Representation, https://www.ietf.org/id/draft-heide-nwcrg-rlnc-00.txt

[BITR] Alan H. Karp, "Bit reversal on uniprocessors", SIAM Review, 38 (1): 1-26, 1996.

Authors' Addresses

Jing Zhu

Intel

Email: jing.z.zhu@intel.com

SungHoon Seo

Korea Telecom

Email: sh.seo@kt.com

Satish Kanugovi

Nokia

Email: satish.k@nokia.com

Shuping Peng

Huawei

Email: pengshuping@huawei.com

INTAREA J. Zhu
Internet Draft Intel
Intended status: Standards Track S. Seo
Expires: April 1,2020 Korea Telecom
S. Kanugovi
Nokia
S. Peng
Huawei
October 1, 2019

User-Plane Protocols for Multiple Access Management Service draft-zhu-intarea-mams-user-protocol-07

Status of this Memo

Copyright Notice

Abstract

Today, a device can be simultaneously connected to multiple communication networks based on different technology implementations and network architectures like WiFi, LTE, and DSL. In such multi-connectivity scenario, it is desirable to combine multiple access networks or select the best one to improve quality of experience for a user and improve overall network utilization and efficiency. This document presents the u-plane protocols for a multi access management services (MAMS) framework that can be used to flexibly select the combination of uplink and downlink access and core network paths having the optimal performance, and user plane treatment for improving network utilization and efficiency and enhanced quality of experience for user applications.

Table of Contents

1. Introduction

Multi Access Management Service (MAMS) [MAMS] is a programmable framework to select and configure network paths, as well as adapt to dynamic network conditions, when multiple network connections can serve a client device. It is based on principles of user plane interworking that enables the solution to be deployed as an overlay without impacting the underlying networks.

This document presents the u-plane protocols for enabling the MAMS framework. It co-exists and complements the existing protocols by providing a way to negotiate and configure the protocols based on client and network capabilities. Further it allows exchange of network state information and leveraging network intelligence to optimize the performance of such protocols. An important goal for MAMS is to ensure that there is minimal or no dependency on the actual access technology of the participating links. This allows the scheme to be scalable for addition of newer access technologies and for independent evolution of the existing access technologies.

2. Terminologies

Anchor Connection: refers to the network path from the N-MADP to the Application Server that corresponds to a specific IP anchor that has assigned an IP address to the client.

Delivery Connection: refers to the network path from the N-MADP to the C-MADP.

"Network Connection Manager" (NCM), "Client Connection Manager" (CCM), "Network Multi Access Data Proxy" (N-MADP), and "Client Multi Access Data Proxy" (C-MADP) in this document are to be interpreted as described in [MAMS].

3. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

The terminologies "Network Connection Manager" (NCM), "Client Connection Manager" (CCM), "Network Multi Access Data Proxy" (N-MADP), and "Client Multi Access Data Proxy" (C-MADP) in this document are to be interpreted as described in [MAMS].

4 MAMS User-Plane Protocols

Figure 1 shows the MAMS u-plane protocol stack as specified in [MAMS].

```
             +-----------------------------------------------
     ------+
             |      User Payload (e.g. IP PDU)
     |
             |-----------------------------------------------
     ------|
          +--|-----------------------------------------------
     ------|--+
          |  |-----------------------------------------------
     ------|  |
          |  | Multi-Access (MX) Convergence Sublayer
     |  |
          |  |-----------------------------------------------
          ------|  |
          |  |-----------------------------------------------
     ------|  |
          |  | MX Adaptation   | MX Adaptation | MX Adaptation
     |  |
          |  | Sublayer        | Sublayer      | Sublayer
     |  |
          |  | (optional)      | (optional)    | (optional)
     |  |
          |  |-----------------------------------------------
     ------|  |
          |  | Access #1 IP    | Access #2 IP  | Access #3 IP
     |  |
          |  +-----------------------------------------------
     ------+  |
          +--------------------------------------------------
---------+
```

Figure 1: MAMS U-plane Protocol Stack

It consists of the following two Sublayers:

- o Multi-Access (MX) Convergence Sublayer: This layer performs multi-access specific tasks, e.g., access (path) selection, multi-link (path) aggregation, splitting/reordering, lossless switching, fragmentation, concatenation, keep-alive, and probing etc.
- o Multi-Access (MX) Adaptation Sublayer: This layer performs functions to handle tunneling, network layer security, and NAT.

The MX convergence sublayer operates on top of the MX adaptation sublayer in the protocol stack. From the Transmitter perspective, a User Payload (e.g. IP PDU) is processed by the convergence sublayer first, and then by the adaptation sublayer before being transported over a delivery access connection; from the Receiver perspective, an IP packet received over a delivery connection is processed by the MX adaptation sublayer first, and then by the MX convergence sublayer.

4.1 MX Adaptation Sublayer

The MX adaptation sublayer supports the following mechanisms and protocols while transmitting user plane packets on the network path:

- o UDP Tunneling: The user plane packets of the anchor connection can be encapsulated in a UDP tunnel of a delivery connection between the N-MADP and C-MADP.
- o IPsec Tunneling: The user plane packets of the anchor connection are sent through an IPsec tunnel of a delivery connection.
- o Client Net Address Translation (NAT): The Client IP address of user plane packet of the anchor connection is changed, and sent over a delivery connection.
- o Pass Through: The user plane packets are passing through without any change over the anchor connection.

The MX adaptation sublayer also supports the following mechanisms and protocols to ensure security of user plane packets over the network path.

- o IPsec Tunneling: An IPsec [RFC7296] tunnel is established between the N-MADP and C-MADP on the network path that is considered untrusted.

o DTLS: If UDP tunneling is used on the network path that is considered "untrusted", DTLS (Datagram Transport Layer Security) [RFC6347] can be used.

The Client NAT method is the most efficient due to no tunneling overhead. It SHOULD be used if a delivery connection is "trusted" and without NAT function on the path.

The UDP or IPsec Tunnelling method SHOULD be used if a delivery connection has a NAT function placed on the path.

4.2 GMA-based MX Convergence Sublayer

Figure 2 shows the MAMS u-plane protocol stack based on trailer-based encapsulation [GMA]. Multiple access networks are combined into a single IP connection. If NCM determines that N-MADP is to be instantiated with GMA as the MX Convergence Protocol, it exchanges the support of GMA convergence capability in the discovery and capability exchange procedures [MAMS].

```
          +---------------------------------------------------
     ---+
          |                        IP PDU
     |
          |---------------------------------------------------
     ---|
          |               GMA  Convergence Sublayer
     |
          |---------------------------------------------------
     ---|
          | MX Adaptation  | MX Adaptation | MX Adaptation
     |
          | Sublayer       | Sublayer      | Sublayer
     |
          | (optional)     | (optional)    | (optional)
     |
          |---------------------------------------------------
     ---|
          | Access #1 IP   | Access #2 IP  | Access #3 IP
     |
          +---------------------------------------------------
     ---+
```

Figure 2: MAMS U-plane Protocol Stack with GMA as MX Convergence Layer

Figure 3 shows the trailer-based Multi-Access (MX) PDU (Protocol Data Unit) format [GMA]. If the MX adaptation method is UDP tunneling and "MX header optimization" in the "MX_UP_Setup_Configuration_Request" message [MAMS] is true, the "IP length" and "IP checksum" header fields of the MX PDU SHOULD remain unchanged. Otherwise, they should be updated after adding or removing the GMA trailer in the convergence sublayer.

```
          +---------------------------------------------------
     ----+
          | IP hdr |        IP payload             | GMA
          Trailer |                   +---------------------
          ---------------------------------+
```

Figure 3: GMA PDU Format

4.3 MPTCP-based MX Convergence Sublayer

Figure 4 shows the MAMS u-plane protocol stack based on MPTCP. Here, MPTCP is reused as the "MX Convergence Sublayer" protocol. Multiple access networks are combined into a single MPTCP connection. Hence, no new u-plane protocol or PDU format is needed in this case.

```
     |-----------------------------------------------------|
     |                        MPTCP                        |
     |-----------------------------------------------------|
     |  TCP            |   TCP           |       TCP       |
     |-----------------------------------------------------|
     | MX Adaptation   | MX Adaptation  | MX Adaptation    |
     | Sublayer        | Sublayer       | Sublayer         |
     | (optional)      | (optional)     | (optional)       |
     |-----------------------------------------------------|
     | Access #1 IP    | Access #2 IP   | Access #3 IP     |
     +-----------------------------------------------------+
```

Figure 4: MAMS U-plane Protocol Stack with MPTCP as MX Convergence Layer

If NCM determines that N-MADP is to be instantiated with MPTCP as the MX Convergence Protocol, it exchanges the support of MPTCP capability in the discovery and capability exchange procedures [MAMS]. MPTCP proxy protocols [MPProxy][MPPlain] SHOULD be used to manage traffic steering and aggregation over multiple delivery connections.

4.4 GRE as MX Convergence Sublayer

Figure 5 shows the MAMS u-plane protocol stack based on GRE (Generic Routing Encapsulation) [GRE2784]. Here, GRE is reused as the "MX Convergence sub-layer" protocol. Multiple access networks are combined into a single GRE connection. Hence, no new u-plane protocol or PDU format is needed in this case.

```
     +------------------------------------------------------+
     |      User Payload (e.g. IP PDU)                      |
     |------------------------------------------------------|
     |              GRE as MX Convergence Sublayer          |
     |------------------------------------------------------|
          |         GRE Delivery Protocol (e.g. IP)
   |
     |------------------------------------------------------|
     | MX Adaptation    | MX Adaptation | MX Adaptation       |
     | Sublayer         | Sublayer      | Sublayer            |
     | (optional)       | (optional)    | (optional)          |
     |------------------------------------------------------|
          | Access #1 IP    | Access #2 IP   | Access #3 IP
   |
     +------------------------------------------------------+
```

Figure 5: MAMS U-plane Protocol Stack with GRE as MX Convergence Layer

If NCM determines that N-MADP is to be instantiated with GRE as the MX Convergence Protocol, it exchanges the support of GRE capability in the discovery and capability exchange procedures [MAMS].

4.4.1 Transmitter Procedures

Transmitter is the N-MADP or C-MADP instance, instantiated with GRE as the convergence protocol that transmits the GRE packets. The Transmitter receives the User Payload (e.g. IP PDU), encapsulates it with a GRE header and Delivery Protocol (e.g. IP) header to generate the GRE Convergence PDU.

When IP is used as the GRE delivery protocol, the IP header information (e.g. IP address) can be created using the IP header of the user payload or a virtual IP address. The "Protocol Type" field of the delivery header is set to 47 (or 0X2F, i.e. GRE)[IANA].

The GRE header fields are set as specified below,

- If the transmitter is a C-MADP instance, then sets the LSB 16 bits to the value of Connection ID for the Anchor Connection associated with the user payload or sets to 0xFFFF if no Anchor Connection ID needs to be specified.
- All other fields in the GRE header including the remaining bits in the key fields are set as per [GRE_2784][GRE_2890].

4.4.2 Receiver Procedures

Receiver is the N-MADP or C-MADP instance, instantiated with GRE as the convergence protocol that receives the GRE packets. The receiver processes the received packets per the GRE procedures [GRE_2784, GRE_2890] and retrieves the GRE header.

- If the Receiver is an N-MADP instance,
  - o Unless the LSB 16 Bits of the Key field are 0xFFFF, they are interpreted as the Connection ID of Anchor Connection for the user payload. This is used to identify the network path over which the User Payload (GRE Payload) is to be transmitted.
- All other fields in the GRE header, including the remaining bits in the Key fields, are processed as per [GRE_2784][GRE_2890].

The GRE Convergence PDU is passed onto the MX Adaptation Layer (if present) before delivery over one of the network paths.

4.5 Co-existence of MX Adaptation and MX Convergence Sublayers

MAMS u-plane protocols support multiple combinations and instances of user plane protocols to be used in the MX Adaptation and the Convergence sublayers.

For example, one instance of the MX Convergence Layer can be MPTCP Proxy [MPProxy][MPPlain] and another instance can be Trailer-based. The MX Adaptation for each can be either UDP tunnel or IPsec. IPsec may be set up for network paths considered as untrusted by the operator, to protect the TCP subflow between client and MPTCP proxy traversing that network path.

Each of the instances of MAMS user plane, i.e. combination of MX Convergence and MX Adaptation layer protocols, can coexist simultaneously and independently handle different traffic types.

5. MX Convergence Control Message

A UDP connection may be configured between C-MADP and N-MADP to exchange control messages for keep-alive or path quality estimation. The N-MADP end-point IP address and UDP port number of the UDP connection is used to identify MX control PDU. Figure 6 shows the MX control PDU format with the following fields:

- o Type (1 Byte): the type of the MX control message
- o CID (1 Byte): an unsigned integer to identify the anchor and delivery connection of the MX control message
    - + Anchor Connection ID (MSB 4 Bits): an unsigned integer to identify the anchor connection
    - + Delivery Connection ID (LSB 4 Bits): an unsigned integer to identify the delivery connection
- o MX Control Message (variable): the payload of the MX control message

Figure 7 shows the MX convergence control protocol stack, and MX control PDU goes through the MX adaptation sublayer the same way as MX data PDU.

```
                        <----MX Control PDU Payload ---------
------>
+-------------------------------------------------------------
------+
| IP header | UDP Header| Type | CID |       MX Control
Message    |                  +-----------------------------
---------------------------------------+
```

Figure 6: MX Control PDU Format

```
     |------------------------------------------------------|
     |          MX Convergence Control Messages             |
     |------------------------------------------------------|
          |                  UDP/IP
   |
     |------------------------------------------------------|
     | MX Adaptation    | MX Adaptation   | MX Adaptation     |
     | Sublayer         | Sublayer        | Sublayer          |
     | (optional)       | (optional)      | (optional)        |
     |------------------------------------------------------|
```

```
          | Access #1 IP   | Access #2 IP  | Access #3 IP
   |
     +-------------------------------------------------------+
```

Figure 7: MX Convergence Control Protocol Stack

5.1 Keep-Alive Message

The "Type" field is set to "0" for Keep-Alive messages. C-MADP may send out Keep-Alive message periodically over one or multiple delivery connections, especially if UDP tunneling is used as the adaptation method for the delivery connection with a NAT function on the path.

A Keep-Alive message is 6 Bytes long, and consists of the following fields:

- o Keep-Alive Sequence Number (2 Bytes): the sequence number of the keep-alive message
- o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.

5.2 Probe Message

The "Type" field is set to "1" for Probe messages.

N-MADP may send out the Probe message for path quality estimation. In response, C-MADP may send back the ACK message.

A Probe message consists of the following fields:

- o Probing Sequence Number (2 Bytes): the sequence number of the Probe REQ message
- o Probing Flag (1 Byte):
  - + Bit #0: a ACK flag to indicate if the ACK message is expected (1) or not (0);
  - + Bit #1: a Probe Type flag to indicate if the Probe message is sent during the initialization phase (0) when the network path is not included for transmission of user data or the active phase (1) when the network path is included for transmission of user data;
  - + Bit #2: a bit flag to indicate the presence of the Reverse Connection ID (R-CID) field.
  - + Bit #3~7: reserved
- o Reverse Connection ID (1 Byte): the connection ID of the delivery connection for sending out the ACK message on the reverse path

- o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.
- o Padding (variable)

The "R-CID" field is only present if both Bit #0 and Bit #2 of the "Probing Flag" field are set to "1". Moreover, Bit #2 of the "Probing Flag" field SHOULD be set to "0" if the Bit #0 is "0", indicating the ACK message is not expected.

If the "R-CID" field is not present but the Bit #0 of the "Probing Flag" field is set to "1", the ACK message SHOULD be sent over the same delivery connection as the Probe message.

The "Padding" field is used to control the length of Probe message.

5.3 Packet Loss Report (PLR) Message

The "Type" field is set to "2" for PLR messages.

C-MADP may send out the PLR messages to report lost MX SDU for example during handover. In response, C-MADP may retransmit the lost MX SDU accordingly.

A PLR message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection which the ACK message is for;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the anchor connection which the ACK message is for;
- o ACK number (4 Bytes): the next (in-order) sequence number (SN) that the sender of the PLR message is expecting
- o Number of Loss Bursts (1 Byte)
  For each loss burst, include the following
    - + Sequence Number of the first lost MX SDU in a burst (4 Bytes)
    - + Number of consecutive lost MX SDUs in the burst (1 Byte)

```
          C-MADP
          N-MADP
              |
|
```

```
              |<------------------ MX SDU (data packets)-----
---|
              |
|
             +-----------------------------------+
|
             |Packet Loss detected              |
|
             +-----------------------------------+
|
              |
|
              |----- PLR Message ----------------------------
-->|
              |<-------------retransmit(lost)MX SDUs --------
---|
```

Figure 8: MAMS Retransmission Procedure

Figure 8 shows the MAMS retransmission procedure in an example where the lost packet is found and retransmitted.

5.4 First Sequence Number (FSN) Message

The "Type" field is set to "3" for FSN messages.

N-MADP may send out the FSN messages to indicate the oldest MX SDU in its buffer if a lost MX SDU is not found in the buffer after receiving the PLR message from C-MADP. In response, C-MADP SHALL only report packet loss with SN not smaller than FSN.

A FSN message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection which the FSN message is for;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the anchor connection which the FSN message is for;
- o First Sequence Number (4 Bytes): the sequence number (SN) of the oldest MX SDU in the (retransmission) buffer of the sender of the FSN message.

Figure 9 shows the MAMS retransmission procedure in an example where the lost packet is not found.

```
          C-MADP
          N-MADP
              |
|
              |<------------------ MX SDU (data packets)-----
---|
              |
|
             +-----------------------------------+
|
             |Packet Loss detected              |
|
             +-----------------------------------+
|
              |
|
              |----- PLR Message ----------------------------
-->|
              |                              +---------------
          ------+
              |                              |Lost packet not
          found|
              |                              +---------------
------+
              |<-------------FSN message --------------------
---|
```

Figure 9: MAMS Retransmission Procedure with FSN

5.5 Coded MX SDU (CMS) Message

The "Type" field is set to "4" for CMS messages.

N-MADP (or C-MADP) may send out the CMS message to support downlink (or uplink) packet loss recovery through coding, e.g. [CRLNC], [CTCP], [RLNC]. A coded MX SDU is generated by applying a network coding algorithm to multiple consecutive (uncoded) MX SDUs, and it is used for fast recovery without retransmission if any of the MX SDUs is lost.

A Coded MX SDU message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection of the coded MX SDU;

o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the coded MX;
o Sequence Number (4 Bytes): the sequence number of the first (uncoded) MX SDU used to generate the coded MX SDU.
o Fragmentation Control (FC) (1 Byte): to provide necessary information for re-assembly, only needed if the coded MX SDU is too long to transport in a single MX control PDU.
o N (1 Byte): the number of consecutive MX SDUs used to generate the coded MX SDU
o K (1 Byte): the length (in terms of bits) of the coding coefficient field
o Coding Coefficient (N x K / 8 Bytes)
    + a(i): the coding coefficient of the i-th (uncoded) MX SDU
    + padding
o Coded MX SDU (variable): the coded MX SDU

If K = 0, the simple XOR method is used to generate the Coded MX SDU from N consecutive uncoded MX SDUs, and the a(i) fields are not included in the message.

If the coded MX SDU is too long, it can be fragmented, and transported by multiple MX control PDUs. The N, K, and a(i) fields are only included in the MX PDU carrying the first fragment of the coded MX SDU.

```
          C-MADP
          N-MADP
              |
|
              |<------------------ MX SDU #1 ----------------
---|
              |      lost<-------- MX SDU #2 ----------------
---|
              |<---- CMS Message (MX SDU #1 XOR MX SDU #2)---
          ---|
             +-----------------------+
|
             | MX SDU #2 recovered  |
|
             +-----------------------+
|
```

|
|

Figure 10: MAMS Packet Recovery Procedure with XOR Coding

5.6 Traffic Splitting Update (TSU) Message

The "Type" field is set to "5" for TSU messages.

N-MADP (or C-MADP) may send out a TSU message if downlink (or uplink) traffic splitting configuration has changed.

A TSU message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class;
- o Sequence Number (2 Bytes): an unsigned integer to identify the TSU message.
- o Flags (1 Byte)
  - + Bit #0: a Reverse Path bit flag to indicate if the traffic splitting configuration is for the reverse path (1) or not (0);
  - + Bit #1: a Bit-Reversal bit flag to indicate if bit-reversal is used in traffic splitting
  - + Others: reserved.
- o Traffic Splitting Configuration Parameters (5 + (N -1) Bytes):
  - + StartSN (4 Bytes): the sequence number of the first MX SDU using the traffic splitting configuration provided by the TSU message
  - + L (1 Byte): the traffic splitting burst size
  - + K(i): the traffic splitting threshold of the i-th delivery connection, where connections are ordered according to their Connection ID.

Let’s use f(x) to denote the traffic splitting function, which maps a MX SDU Sequence Number "x" to the i-th delivery connection.

f(x)=i, if K[i-1]< or = mod(x - StartSN, L) < K[i]

Wherein, 1 < or = i < N, K[0]=0, and K[N]=L.

N is the total number of connections for delivering a data flow, identified by (anchor) Connection ID and Traffic Class ID.

When the bit-reversal bit is set to 1, the burst size L MUST be a power of 2, and the traffic splitting function is

f(x)=i, if K[i-1]< or = F(mod(x - StartSN, L)) < K[i]

Wherein F(.) is the bit reversal function [BITR] of the input variable.

5.7 Acknowledgement Message

The "Type" field is set to "6" for ACK messages.

C-MADP (or N-MADP) SHOULD send out the ACK message in response to the successful reception of a PLR, FSN, or TSU message.

C-MADP SHOULD send out the ACK message in response to a Probe message with the ACK flag set to "1".

The ACK message consists of the following fields:

- o Acknowledgment Number (2 Bytes): the sequence number of the received message.
- o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.

6 Security Considerations

User data in MAMS framework rely on the security of the underlying network transport paths. When this cannot be assumed, NCM configures use of appropriate protocols for security, e.g. IPsec [RFC4301] [RFC3948], DTLS [RFC6347].

7 IANA Considerations

This draft makes no requests of IANA.

8 Contributing Authors

The editors gratefully acknowledge the following additional contributors in alphabetical order: Salil Agarwal/Nokia, Hema Pentakota/Nokia.

9 References

9.1 Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC4301] Kent, S. and K. Seo, "Security Architecture for the Internet Protocol", RFC 4301, DOI10.17487/RFC4301, December 2005, <http://www.rfc-editor.org/info/rfc4301>.

9.2 Informative References

[RFC6347] Rescorla, E. and N. Modadugu, "Datagram Transport Layer Security Version 1.2", RFC 6347, January 2012, <http://www.rfc-editor.org/info/rfc6347>.

[RFC7296] Kaufman, C., Hoffman, P., Nir, Y., Eronen, P., and T. Kivinen, "Internet Key Exchange Protocol Version 2 (IKEv2)", STD 79, RFC 7296, DOI 10.17487/RFC7296, October 2014, <http://www.rfc-editor.org/info/rfc7296>.

[RFC3948] Huttunen, A., Swander, B., Volpe, V., DiBurro, L., and M. Stenberg, "UDP Encapsulation of IPsec ESP Packets", RFC 3948, DOI 10.17487/RFC3948, January 2005, <http://www.rfc-editor.org/info/rfc3948>.

[MPProxy] X. Wei, C. Xiong, and E. Lopez, "MPTCP proxy mechanisms", https://tools.ietf.org/html/draft-wei-mptcp-proxy-mechanism-02

[MPPlain] M. Boucadair et al, "An MPTCP Option for Network-Assisted MPTCP", https://www.ietf.org/id/draft-boucadair-mptcp-plain-mode-09.txt

[MAMS] S. Kanugovi, S. Vasudevan, F. Baboescu, and J. Zhu, "Multiple Access Management Protocol", https://tools.ietf.org/html/draft-kanugovi-intarea-mams-protocol-03

[GMA] J. Zhu, "Trailer-based Encapsulation Protocols for Generic Multi-Access Convergence", https://tools.ietf.org/html/draft-zhu-intarea-gma-01

[GRE2784] D. Farinacci, et al., "Generic Routing Encapsulation (GRE)", RFC 2784 March 2000, <http://www.rfc-editor.org/info/rfc2784>.

[GRE2890] G. Dommety, "Key and Sequence Number Extensions to GRE", RFC 2890 September 2000, <http://www.rfc-editor.org/info/rfc2890>.

[IANA] https://www.iana.org/assignments/protocol-numbers/protocol-numbers.xhtml

[LWIPEP] 3GPP TS 36.361, "Evolved Universal Terrestrial Radio Access (E-UTRA); LTE-WLAN Radio Level Integration Using Ipsec Tunnel (LWIP) encapsulation; Protocol specification"

[RFC791] Internet Protocol, September 1981

[CRLNC] S Wunderlich, F Gabriel, S Pandi, et al. Caterpillar RLNC (CRLNC): A Practical Finite Sliding Window RLNC Approach, IEEE Access, 2017

[CTCP] M. Kim, et al. Network Coded TCP (CTCP), eprint arXiv:1212.2291, 2012

[RLNC] J. Heide, et al. Random Linear Network Coding (RLNC)-Based Symbol Representation, https://www.ietf.org/id/draft-heide-nwcrg-rlnc-00.txt

[BITR] Alan H. Karp, "Bit reversal on uniprocessors", SIAM Review, 38 (1): 1-26, 1996.

Authors' Addresses

Jing Zhu

Intel

Email: jing.z.zhu@intel.com

SungHoon Seo

Korea Telecom

Email: sh.seo@kt.com

Satish Kanugovi

Nokia

Email: satish.k@nokia.com

Shuping Peng

Huawei

Email: pengshuping@huawei.com

INTAREA   J. Zhu
Internet Draft   Intel
Intended status: Standards Track   S. Seo
Expires: April 1,2020   Korea Telecom
                                                     S. Kanugovi
                                                           Nokia
                                                         S. Peng
                                                          Huawei
                                                 October 1, 2019

User-Plane Protocols for Multiple Access Management Service   draft-zhu-intarea-mams-user-protocol-07

Status of this Memo

Copyright Notice

Abstract

Today, a device can be simultaneously connected to multiple communication networks based on different technology implementations and network architectures like WiFi, LTE, and DSL. In such multi-connectivity scenario, it is desirable to combine multiple access networks or select the best one to improve quality of experience for a user and improve overall network utilization and efficiency. This document presents the u-plane protocols for a multi access management services (MAMS) framework that can be used to flexibly select the combination of uplink and downlink access and core network paths having the optimal performance, and user plane treatment for improving network utilization and efficiency and enhanced quality of experience for user applications.

Table of Contents

1. Introduction

Multi Access Management Service (MAMS) [MAMS] is a programmable framework to select and configure network paths, as well as adapt to dynamic network conditions, when multiple network connections can serve a client device. It is based on principles of user plane interworking that enables the solution to be deployed as an overlay without impacting the underlying networks.

This document presents the u-plane protocols for enabling the MAMS framework. It co-exists and complements the existing protocols by providing a way to negotiate and configure the protocols based on client and network capabilities. Further it allows exchange of network state information and leveraging network intelligence to optimize the performance of such protocols. An important goal for MAMS is to ensure that there is minimal or no dependency on the actual access technology of the participating links. This allows the scheme to be scalable for addition of newer access technologies and for independent evolution of the existing access technologies.

2. Terminologies

Anchor Connection: refers to the network path from the N-MADP to the Application Server that corresponds to a specific IP anchor that has assigned an IP address to the client.

Delivery Connection: refers to the network path from the N-MADP to the C-MADP.

"Network Connection Manager" (NCM), "Client Connection Manager" (CCM), "Network Multi Access Data Proxy" (N-MADP), and "Client Multi Access Data Proxy" (C-MADP) in this document are to be interpreted as described in [MAMS].

3. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

The terminologies "Network Connection Manager" (NCM), "Client Connection Manager" (CCM), "Network Multi Access Data Proxy" (N-MADP), and "Client Multi Access Data Proxy" (C-MADP) in this document are to be interpreted as described in [MAMS].

4 MAMS User-Plane Protocols

Figure 1 shows the MAMS u-plane protocol stack as specified in [MAMS].

```
             +-----------------------------------------------
     ------+
             |      User Payload (e.g. IP PDU)
     |
             |-----------------------------------------------
     ------|
          +--|-----------------------------------------------
     ------|--+
          |  |-----------------------------------------------
     ------|  |
          |  | Multi-Access (MX) Convergence Sublayer
     |  |
          |  |-----------------------------------------------
          ------|  |
          |  |-----------------------------------------------
     ------|  |
          |  | MX Adaptation  | MX Adaptation | MX Adaptation
     |  |
          |  | Sublayer       | Sublayer      | Sublayer
     |  |
          |  | (optional)     | (optional)    | (optional)
     |  |
          |  |-----------------------------------------------
     ------|  |
          |  | Access #1 IP   | Access #2 IP  | Access #3 IP
     |  |
          |  +-----------------------------------------------
     ------+  |
          +--------------------------------------------------
---------+
```

Figure 1: MAMS U-plane Protocol Stack

It consists of the following two Sublayers:

- o Multi-Access (MX) Convergence Sublayer: This layer performs multi-access specific tasks, e.g., access (path) selection, multi-link (path) aggregation, splitting/reordering, lossless switching, fragmentation, concatenation, keep-alive, and probing etc.
- o Multi-Access (MX) Adaptation Sublayer: This layer performs functions to handle tunneling, network layer security, and NAT.

The MX convergence sublayer operates on top of the MX adaptation sublayer in the protocol stack. From the Transmitter perspective, a User Payload (e.g. IP PDU) is processed by the convergence sublayer first, and then by the adaptation sublayer before being transported over a delivery access connection; from the Receiver perspective, an IP packet received over a delivery connection is processed by the MX adaptation sublayer first, and then by the MX convergence sublayer.

4.1 MX Adaptation Sublayer

The MX adaptation sublayer supports the following mechanisms and protocols while transmitting user plane packets on the network path:

- o UDP Tunneling: The user plane packets of the anchor connection can be encapsulated in a UDP tunnel of a delivery connection between the N-MADP and C-MADP.
- o IPsec Tunneling: The user plane packets of the anchor connection are sent through an IPsec tunnel of a delivery connection.
- o Client Net Address Translation (NAT): The Client IP address of user plane packet of the anchor connection is changed, and sent over a delivery connection.
- o Pass Through: The user plane packets are passing through without any change over the anchor connection.

The MX adaptation sublayer also supports the following mechanisms and protocols to ensure security of user plane packets over the network path.

- o IPsec Tunneling: An IPsec [RFC7296] tunnel is established between the N-MADP and C-MADP on the network path that is considered untrusted.

o DTLS: If UDP tunneling is used on the network path that is considered "untrusted", DTLS (Datagram Transport Layer Security) [RFC6347] can be used.

The Client NAT method is the most efficient due to no tunneling overhead. It SHOULD be used if a delivery connection is "trusted" and without NAT function on the path.

The UDP or IPsec Tunnelling method SHOULD be used if a delivery connection has a NAT function placed on the path.

4.2 GMA-based MX Convergence Sublayer

Figure 2 shows the MAMS u-plane protocol stack based on trailer-based encapsulation [GMA]. Multiple access networks are combined into a single IP connection. If NCM determines that N-MADP is to be instantiated with GMA as the MX Convergence Protocol, it exchanges the support of GMA convergence capability in the discovery and capability exchange procedures [MAMS].

```
          +---------------------------------------------------
     ---+
          |                        IP PDU
     |
          |---------------------------------------------------
     ---|
          |               GMA  Convergence Sublayer
     |
          |---------------------------------------------------
     ---|
          | MX Adaptation  | MX Adaptation | MX Adaptation
     |
          | Sublayer       | Sublayer      | Sublayer
     |
          | (optional)     | (optional)    | (optional)
     |
          |---------------------------------------------------
     ---|
          | Access #1 IP   | Access #2 IP  | Access #3 IP
     |
          +---------------------------------------------------
     ---+
```

Figure 2: MAMS U-plane Protocol Stack with GMA as MX Convergence Layer

Figure 3 shows the trailer-based Multi-Access (MX) PDU (Protocol Data Unit) format [GMA]. If the MX adaptation method is UDP tunneling and "MX header optimization" in the "MX_UP_Setup_Configuration_Request" message [MAMS] is true, the "IP length" and "IP checksum" header fields of the MX PDU SHOULD remain unchanged. Otherwise, they should be updated after adding or removing the GMA trailer in the convergence sublayer.

```
          +---------------------------------------------------
     ----+
          | IP hdr |        IP payload            | GMA
          Trailer |                  +---------------------
          ----------------------------------+
```

Figure 3: GMA PDU Format

4.3 MPTCP-based MX Convergence Sublayer

Figure 4 shows the MAMS u-plane protocol stack based on MPTCP. Here, MPTCP is reused as the "MX Convergence Sublayer" protocol. Multiple access networks are combined into a single MPTCP connection. Hence, no new u-plane protocol or PDU format is needed in this case.

```
     |-----------------------------------------------------|
     |                       MPTCP                         |
     |-----------------------------------------------------|
     |  TCP             |  TCP           |      TCP        |
     |-----------------------------------------------------|
     | MX Adaptation    | MX Adaptation  | MX Adaptation   |
     | Sublayer         | Sublayer       | Sublayer        |
     | (optional)       | (optional)     | (optional)      |
     |-----------------------------------------------------|
     | Access #1 IP     | Access #2 IP   | Access #3 IP    |
     +-----------------------------------------------------+
```

Figure 4: MAMS U-plane Protocol Stack with MPTCP as MX Convergence Layer

If NCM determines that N-MADP is to be instantiated with MPTCP as the MX Convergence Protocol, it exchanges the support of MPTCP capability in the discovery and capability exchange procedures [MAMS]. MPTCP proxy protocols [MPProxy][MPPlain] SHOULD be used to manage traffic steering and aggregation over multiple delivery connections.

4.4 GRE as MX Convergence Sublayer

Figure 5 shows the MAMS u-plane protocol stack based on GRE (Generic Routing Encapsulation) [GRE2784]. Here, GRE is reused as the "MX Convergence sub-layer" protocol. Multiple access networks are combined into a single GRE connection. Hence, no new u-plane protocol or PDU format is needed in this case.

```
     +-----------------------------------------------------+
     |      User Payload (e.g. IP PDU)                     |
     |-----------------------------------------------------|
     |              GRE as MX Convergence Sublayer         |
     |-----------------------------------------------------|
          |         GRE Delivery Protocol (e.g. IP)
   |
     |-----------------------------------------------------|
     | MX Adaptation    | MX Adaptation | MX Adaptation      |
     | Sublayer         | Sublayer      | Sublayer           |
     | (optional)       | (optional)    | (optional)         |
     |-----------------------------------------------------|
          | Access #1 IP    | Access #2 IP   | Access #3 IP
   |
     +-----------------------------------------------------+
```

Figure 5: MAMS U-plane Protocol Stack with GRE as MX Convergence Layer

If NCM determines that N-MADP is to be instantiated with GRE as the MX Convergence Protocol, it exchanges the support of GRE capability in the discovery and capability exchange procedures [MAMS].

4.4.1   Transmitter Procedures

Transmitter is the N-MADP or C-MADP instance, instantiated with GRE as the convergence protocol that transmits the GRE packets. The Transmitter receives the User Payload (e.g. IP PDU), encapsulates it with a GRE header and Delivery Protocol (e.g. IP) header to generate the GRE Convergence PDU.

When IP is used as the GRE delivery protocol, the IP header information (e.g. IP address) can be created using the IP header of the user payload or a virtual IP address. The "Protocol Type" field of the delivery header is set to 47 (or 0X2F, i.e. GRE)[IANA].

The GRE header fields are set as specified below,

- If the transmitter is a C-MADP instance, then sets the LSB 16 bits to the value of Connection ID for the Anchor Connection associated with the user payload or sets to 0xFFFF if no Anchor Connection ID needs to be specified.
- All other fields in the GRE header including the remaining bits in the key fields are set as per [GRE_2784][GRE_2890].

4.4.2 Receiver Procedures

Receiver is the N-MADP or C-MADP instance, instantiated with GRE as the convergence protocol that receives the GRE packets. The receiver processes the received packets per the GRE procedures [GRE_2784, GRE_2890] and retrieves the GRE header.

- If the Receiver is an N-MADP instance,
    - o Unless the LSB 16 Bits of the Key field are 0xFFFF, they are interpreted as the Connection ID of Anchor Connection for the user payload. This is used to identify the network path over which the User Payload (GRE Payload) is to be transmitted.
- All other fields in the GRE header, including the remaining bits in the Key fields, are processed as per [GRE_2784][GRE_2890].

The GRE Convergence PDU is passed onto the MX Adaptation Layer (if present) before delivery over one of the network paths.

4.5 Co-existence of MX Adaptation and MX Convergence Sublayers

MAMS u-plane protocols support multiple combinations and instances of user plane protocols to be used in the MX Adaptation and the Convergence sublayers.

For example, one instance of the MX Convergence Layer can be MPTCP Proxy [MPProxy][MPPlain] and another instance can be Trailer-based. The MX Adaptation for each can be either UDP tunnel or IPsec. IPsec may be set up for network paths considered as untrusted by the operator, to protect the TCP subflow between client and MPTCP proxy traversing that network path.

Each of the instances of MAMS user plane, i.e. combination of MX Convergence and MX Adaptation layer protocols, can coexist simultaneously and independently handle different traffic types.

5. MX Convergence Control Message

A UDP connection may be configured between C-MADP and N-MADP to exchange control messages for keep-alive or path quality estimation. The N-MADP end-point IP address and UDP port number of the UDP connection is used to identify MX control PDU. Figure 6 shows the MX control PDU format with the following fields:

- o Type (1 Byte): the type of the MX control message
- o CID (1 Byte): an unsigned integer to identify the anchor and delivery connection of the MX control message
    - + Anchor Connection ID (MSB 4 Bits): an unsigned integer to identify the anchor connection
    - + Delivery Connection ID (LSB 4 Bits): an unsigned integer to identify the delivery connection
- o MX Control Message (variable): the payload of the MX control message

Figure 7 shows the MX convergence control protocol stack, and MX control PDU goes through the MX adaptation sublayer the same way as MX data PDU.

```
                        <----MX Control PDU Payload ---------
------>
+------------------------------------------------------------
------+
| IP header | UDP Header| Type | CID |       MX Control
Message    |                      +----------------------------
---------------------------------------+
```

Figure 6: MX Control PDU Format

```
     |------------------------------------------------------|
     |          MX Convergence Control Messages             |
     |------------------------------------------------------|
          |                  UDP/IP
   |
     |------------------------------------------------------|
     | MX Adaptation    | MX Adaptation   | MX Adaptation       |
     | Sublayer         | Sublayer        | Sublayer            |
     | (optional)       | (optional)      | (optional)          |
     |------------------------------------------------------|
```

```
          | Access #1 IP   | Access #2 IP  | Access #3 IP
   |
     +------------------------------------------------------+
```

Figure 7: MX Convergence Control Protocol Stack

5.1 Keep-Alive Message

The "Type" field is set to "0" for Keep-Alive messages. C-MADP may send out Keep-Alive message periodically over one or multiple delivery connections, especially if UDP tunneling is used as the adaptation method for the delivery connection with a NAT function on the path.

A Keep-Alive message is 6 Bytes long, and consists of the following fields:

- o Keep-Alive Sequence Number (2 Bytes): the sequence number of the keep-alive message
- o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.

5.2 Probe Message

The "Type" field is set to "1" for Probe messages.

N-MADP may send out the Probe message for path quality estimation. In response, C-MADP may send back the ACK message.

A Probe message consists of the following fields:

- o Probing Sequence Number (2 Bytes): the sequence number of the Probe REQ message
- o Probing Flag (1 Byte):
  - + Bit #0: a ACK flag to indicate if the ACK message is expected (1) or not (0);
  - + Bit #1: a Probe Type flag to indicate if the Probe message is sent during the initialization phase (0) when the network path is not included for transmission of user data or the active phase (1) when the network path is included for transmission of user data;
  - + Bit #2: a bit flag to indicate the presence of the Reverse Connection ID (R-CID) field.
  - + Bit #3~7: reserved
- o Reverse Connection ID (1 Byte): the connection ID of the delivery connection for sending out the ACK message on the reverse path

- o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.
- o Padding (variable)

The "R-CID" field is only present if both Bit #0 and Bit #2 of the "Probing Flag" field are set to "1". Moreover, Bit #2 of the "Probing Flag" field SHOULD be set to "0" if the Bit #0 is "0", indicating the ACK message is not expected.

If the "R-CID" field is not present but the Bit #0 of the "Probing Flag" field is set to "1", the ACK message SHOULD be sent over the same delivery connection as the Probe message.

The "Padding" field is used to control the length of Probe message.

5.3 Packet Loss Report (PLR) Message

The "Type" field is set to "2" for PLR messages.

C-MADP may send out the PLR messages to report lost MX SDU for example during handover. In response, C-MADP may retransmit the lost MX SDU accordingly.

A PLR message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection which the ACK message is for;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the anchor connection which the ACK message is for;
- o ACK number (4 Bytes): the next (in-order) sequence number (SN) that the sender of the PLR message is expecting
- o Number of Loss Bursts (1 Byte)
  For each loss burst, include the following
  - + Sequence Number of the first lost MX SDU in a burst (4 Bytes)
  - + Number of consecutive lost MX SDUs in the burst (1 Byte)

```
          C-MADP
          N-MADP
              |
|
```

```
              |<------------------ MX SDU (data packets)-----
---|
              |
|
             +-----------------------------------+
|
             |Packet Loss detected             |
|
             +-----------------------------------+
|
              |
|
              |----- PLR Message ----------------------------
-->|
              |<-------------retransmit(lost)MX SDUs --------
---|
```

Figure 8: MAMS Retransmission Procedure

Figure 8 shows the MAMS retransmission procedure in an example where the lost packet is found and retransmitted.

5.4 First Sequence Number (FSN) Message

The "Type" field is set to "3" for FSN messages.

N-MADP may send out the FSN messages to indicate the oldest MX SDU in its buffer if a lost MX SDU is not found in the buffer after receiving the PLR message from C-MADP. In response, C-MADP SHALL only report packet loss with SN not smaller than FSN.

A FSN message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection which the FSN message is for;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the anchor connection which the FSN message is for;
- o First Sequence Number (4 Bytes): the sequence number (SN) of the oldest MX SDU in the (retransmission) buffer of the sender of the FSN message.

Figure 9 shows the MAMS retransmission procedure in an example where the lost packet is not found.

```
          C-MADP
          N-MADP
              |
|
              |<------------------- MX SDU (data packets)-----
---|
              |
|
             +-----------------------------------+
|
             |Packet Loss detected              |
|
             +-----------------------------------+
|
              |
|
              |----- PLR Message ----------------------------
-->|
              |                              +----------------
          ------+
              |                              |Lost packet not
          found|
              |                              +----------------
------+
              |<-------------FSN message ---------------------
---|
```

Figure 9: MAMS Retransmission Procedure with FSN

5.5 Coded MX SDU (CMS) Message

The "Type" field is set to "4" for CMS messages.

N-MADP (or C-MADP) may send out the CMS message to support downlink (or uplink) packet loss recovery through coding, e.g. [CRLNC], [CTCP], [RLNC]. A coded MX SDU is generated by applying a network coding algorithm to multiple consecutive (uncoded) MX SDUs, and it is used for fast recovery without retransmission if any of the MX SDUs is lost.

A Coded MX SDU message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection of the coded MX SDU;

- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the coded MX;
- o Sequence Number (4 Bytes): the sequence number of the first (uncoded) MX SDU used to generate the coded MX SDU.
- o Fragmentation Control (FC) (1 Byte): to provide necessary information for re-assembly, only needed if the coded MX SDU is too long to transport in a single MX control PDU.
- o N (1 Byte): the number of consecutive MX SDUs used to generate the coded MX SDU
- o K (1 Byte): the length (in terms of bits) of the coding coefficient field
- o Coding Coefficient (N x K / 8 Bytes)
  - + a(i): the coding coefficient of the i-th (uncoded) MX SDU
  - + padding
- o Coded MX SDU (variable): the coded MX SDU

If K = 0, the simple XOR method is used to generate the Coded MX SDU from N consecutive uncoded MX SDUs, and the a(i) fields are not included in the message.

If the coded MX SDU is too long, it can be fragmented, and transported by multiple MX control PDUs. The N, K, and a(i) fields are only included in the MX PDU carrying the first fragment of the coded MX SDU.

```
          C-MADP
          N-MADP
              |
|
              |<------------------ MX SDU #1 ----------------
---|
              |      lost<-------- MX SDU #2 ----------------
---|
              |<---- CMS Message (MX SDU #1 XOR MX SDU #2)---
          ---|
             +-----------------------+
|
             | MX SDU #2 recovered  |
|
             +-----------------------+
|
```

|
|

Figure 10: MAMS Packet Recovery Procedure with XOR Coding

5.6 Traffic Splitting Update (TSU) Message

The "Type" field is set to "5" for TSU messages.

N-MADP (or C-MADP) may send out a TSU message if downlink (or uplink) traffic splitting configuration has changed.

A TSU message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class;
- o Sequence Number (2 Bytes): an unsigned integer to identify the TSU message.
- o Flags (1 Byte)
  - + Bit #0: a Reverse Path bit flag to indicate if the traffic splitting configuration is for the reverse path (1) or not (0);
  - + Bit #1: a Bit-Reversal bit flag to indicate if bit-reversal is used in traffic splitting
  - + Others: reserved.
- o Traffic Splitting Configuration Parameters (5 + (N -1) Bytes):
  - + StartSN (4 Bytes): the sequence number of the first MX SDU using the traffic splitting configuration provided by the TSU message
  - + L (1 Byte): the traffic splitting burst size
  - + K(i): the traffic splitting threshold of the i-th delivery connection, where connections are ordered according to their Connection ID.

Let’s use f(x) to denote the traffic splitting function, which maps a MX SDU Sequence Number "x" to the i-th delivery connection.

    f(x)=i, if K[i-1]< or = mod(x - StartSN, L) < K[i]

Wherein, 1 < or = i < N, K[0]=0, and K[N]=L.

N is the total number of connections for delivering a data flow, identified by (anchor) Connection ID and Traffic Class ID.

When the bit-reversal bit is set to 1, the burst size L MUST be a power of 2, and the traffic splitting function is

f(x)=i, if K[i-1]< or = F(mod(x - StartSN, L)) < K[i]

Wherein F(.) is the bit reversal function [BITR] of the input variable.

5.7 Acknowledgement Message

The "Type" field is set to "6" for ACK messages.

C-MADP (or N-MADP) SHOULD send out the ACK message in response to the successful reception of a PLR, FSN, or TSU message.

C-MADP SHOULD send out the ACK message in response to a Probe message with the ACK flag set to "1".

The ACK message consists of the following fields:

- o Acknowledgment Number (2 Bytes): the sequence number of the received message.
- o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.

6 Security Considerations

User data in MAMS framework rely on the security of the underlying network transport paths. When this cannot be assumed, NCM configures use of appropriate protocols for security, e.g. IPsec [RFC4301] [RFC3948], DTLS [RFC6347].

7 IANA Considerations

This draft makes no requests of IANA.

8 Contributing Authors

The editors gratefully acknowledge the following additional contributors in alphabetical order: Salil Agarwal/Nokia, Hema Pentakota/Nokia.

9 References

9.1 Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC4301] Kent, S. and K. Seo, "Security Architecture for the Internet Protocol", RFC 4301, DOI10.17487/RFC4301, December 2005, <http://www.rfc-editor.org/info/rfc4301>.

9.2 Informative References

[RFC6347] Rescorla, E. and N. Modadugu, "Datagram Transport Layer Security Version 1.2", RFC 6347, January 2012, <http://www.rfc-editor.org/info/rfc6347>.

[RFC7296] Kaufman, C., Hoffman, P., Nir, Y., Eronen, P., and T. Kivinen, "Internet Key Exchange Protocol Version 2 (IKEv2)", STD 79, RFC 7296, DOI 10.17487/RFC7296, October 2014, <http://www.rfc-editor.org/info/rfc7296>.

[RFC3948] Huttunen, A., Swander, B., Volpe, V., DiBurro, L., and M. Stenberg, "UDP Encapsulation of IPsec ESP Packets", RFC 3948, DOI 10.17487/RFC3948, January 2005, <http://www.rfc-editor.org/info/rfc3948>.

[MPProxy] X. Wei, C. Xiong, and E. Lopez, "MPTCP proxy mechanisms", https://tools.ietf.org/html/draft-wei-mptcp-proxy-mechanism-02

[MPPlain] M. Boucadair et al, "An MPTCP Option for Network-Assisted MPTCP", https://www.ietf.org/id/draft-boucadair-mptcp-plain-mode-09.txt

[MAMS] S. Kanugovi, S. Vasudevan, F. Baboescu, and J. Zhu, "Multiple Access Management Protocol", https://tools.ietf.org/html/draft-kanugovi-intarea-mams-protocol-03

[GMA] J. Zhu, "Trailer-based Encapsulation Protocols for Generic Multi-Access Convergence", https://tools.ietf.org/html/draft-zhu-intarea-gma-01

[GRE2784] D. Farinacci, et al., "Generic Routing Encapsulation (GRE)", RFC 2784 March 2000, <http://www.rfc-editor.org/info/rfc2784>.

[GRE2890] G. Dommety, "Key and Sequence Number Extensions to GRE", RFC 2890 September 2000, <http://www.rfc-editor.org/info/rfc2890>.

[IANA] https://www.iana.org/assignments/protocol-numbers/protocol-numbers.xhtml

[LWIPEP] 3GPP TS 36.361, "Evolved Universal Terrestrial Radio Access (E-UTRA); LTE-WLAN Radio Level Integration Using Ipsec Tunnel (LWIP) encapsulation; Protocol specification"

[RFC791] Internet Protocol, September 1981

[CRLNC] S Wunderlich, F Gabriel, S Pandi, et al. Caterpillar RLNC (CRLNC): A Practical Finite Sliding Window RLNC Approach, IEEE Access, 2017

[CTCP] M. Kim, et al. Network Coded TCP (CTCP), eprint arXiv:1212.2291, 2012

[RLNC] J. Heide, et al. Random Linear Network Coding (RLNC)-Based Symbol Representation, https://www.ietf.org/id/draft-heide-nwcrg-rlnc-00.txt

[BITR] Alan H. Karp, "Bit reversal on uniprocessors", SIAM Review, 38 (1): 1-26, 1996.

Authors' Addresses

Jing Zhu

Intel

Email: jing.z.zhu@intel.com

SungHoon Seo

Korea Telecom

Email: sh.seo@kt.com

Satish Kanugovi

Nokia

Email: satish.k@nokia.com

Shuping Peng

Huawei

Email: pengshuping@huawei.com

INTAREA J. Zhu
Internet Draft Intel
Intended status: Standards Track S. Seo
Expires: April 1,2020 Korea Telecom
S. Kanugovi
Nokia
S. Peng
Huawei
October 1, 2019

User-Plane Protocols for Multiple Access Management Service
draft-zhu-intarea-mams-user-protocol-07

Status of this Memo

Copyright Notice

Abstract

Today, a device can be simultaneously connected to multiple communication networks based on different technology implementations and network architectures like WiFi, LTE, and DSL. In such multi-connectivity scenario, it is desirable to combine multiple access networks or select the best one to improve quality of experience for a user and improve overall network utilization and efficiency. This document presents the u-plane protocols for a multi access management services (MAMS) framework that can be used to flexibly select the combination of uplink and downlink access and core network paths having the optimal performance, and user plane treatment for improving network utilization and efficiency and enhanced quality of experience for user applications.

Table of Contents

1. Introduction

Multi Access Management Service (MAMS) [MAMS] is a programmable framework to select and configure network paths, as well as adapt to dynamic network conditions, when multiple network connections can serve a client device. It is based on principles of user plane interworking that enables the solution to be deployed as an overlay without impacting the underlying networks.

This document presents the u-plane protocols for enabling the MAMS framework. It co-exists and complements the existing protocols by providing a way to negotiate and configure the protocols based on client and network capabilities. Further it allows exchange of network state information and leveraging network intelligence to optimize the performance of such protocols. An important goal for MAMS is to ensure that there is minimal or no dependency on the actual access technology of the participating links. This allows the scheme to be scalable for addition of newer access technologies and for independent evolution of the existing access technologies.

2. Terminologies

Anchor Connection: refers to the network path from the N-MADP to the Application Server that corresponds to a specific IP anchor that has assigned an IP address to the client.

Delivery Connection: refers to the network path from the N-MADP to the C-MADP.

"Network Connection Manager" (NCM), "Client Connection Manager" (CCM), "Network Multi Access Data Proxy" (N-MADP), and "Client Multi Access Data Proxy" (C-MADP) in this document are to be interpreted as described in [MAMS].

3. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

The terminologies "Network Connection Manager" (NCM), "Client Connection Manager" (CCM), "Network Multi Access Data Proxy" (N-MADP), and "Client Multi Access Data Proxy" (C-MADP) in this document are to be interpreted as described in [MAMS].

4 MAMS User-Plane Protocols

Figure 1 shows the MAMS u-plane protocol stack as specified in [MAMS].

```
             +-----------------------------------------------------+
             |      User Payload (e.g. IP PDU)                      |
             |-----------------------------------------------------|
          +--|-----------------------------------------------------|--+
          |  |-----------------------------------------------------|  |
          |  | Multi-Access (MX) Convergence Sublayer               |  |
          |  |-----------------------------------------------------|  |
          |  |-----------------------------------------------------|  |
          |  | MX Adaptation    | MX Adaptation | MX Adaptation     |  |
          |  | Sublayer         | Sublayer      | Sublayer          |  |
          |  | (optional)       | (optional)    | (optional)        |  |
          |  |-----------------------------------------------------|  |
          |  | Access #1 IP     | Access #2 IP  | Access #3 IP      |  |
          |  +-----------------------------------------------------+  |
          +-----------------------------------------------------------+
```

Figure 1: MAMS U-plane Protocol Stack

It consists of the following two Sublayers:

- o Multi-Access (MX) Convergence Sublayer: This layer performs multi-access specific tasks, e.g., access (path) selection, multi-link (path) aggregation, splitting/reordering, lossless switching, fragmentation, concatenation, keep-alive, and probing etc.
- o Multi-Access (MX) Adaptation Sublayer: This layer performs functions to handle tunneling, network layer security, and NAT.

The MX convergence sublayer operates on top of the MX adaptation sublayer in the protocol stack. From the Transmitter perspective, a User Payload (e.g. IP PDU) is processed by the convergence sublayer first, and then by the adaptation sublayer before being transported over a delivery access connection; from the Receiver perspective, an IP packet received over a delivery connection is processed by the MX adaptation sublayer first, and then by the MX convergence sublayer.

4.1 MX Adaptation Sublayer

The MX adaptation sublayer supports the following mechanisms and protocols while transmitting user plane packets on the network path:

- o UDP Tunneling: The user plane packets of the anchor connection can be encapsulated in a UDP tunnel of a delivery connection between the N-MADP and C-MADP.

o IPsec Tunneling: The user plane packets of the anchor connection are sent through an IPsec tunnel of a delivery connection.
o Client Net Address Translation (NAT): The Client IP address of user plane packet of the anchor connection is changed, and sent over a delivery connection.
o Pass Through: The user plane packets are passing through without any change over the anchor connection.

The MX adaptation sublayer also supports the following mechanisms and protocols to ensure security of user plane packets over the network path.

o IPsec Tunneling: An IPsec [RFC7296] tunnel is established between the N-MADP and C-MADP on the network path that is considered untrusted.
o DTLS: If UDP tunneling is used on the network path that is considered "untrusted", DTLS (Datagram Transport Layer Security) [RFC6347] can be used.

The Client NAT method is the most efficient due to no tunneling overhead. It SHOULD be used if a delivery connection is "trusted" and without NAT function on the path.

The UDP or IPsec Tunnelling method SHOULD be used if a delivery connection has a NAT function placed on the path.

4.2 GMA-based MX Convergence Sublayer

Figure 2 shows the MAMS u-plane protocol stack based on trailer-based encapsulation [GMA]. Multiple access networks are combined into a single IP connection. If NCM determines that N-MADP is to be instantiated with GMA as the MX Convergence Protocol, it exchanges the support of GMA convergence capability in the discovery and capability exchange procedures [MAMS].

```
          +-----------------------------------------------------+
          |                       IP PDU                        |
          |-----------------------------------------------------|
          |               GMA  Convergence Sublayer             |
          |-----------------------------------------------------|
          | MX Adaptation   | MX Adaptation | MX Adaptation     |
          | Sublayer        | Sublayer      | Sublayer          |
          | (optional)      | (optional)    | (optional)        |
          |-----------------------------------------------------|
          | Access #1 IP    | Access #2 IP  | Access #3 IP      |
          +-----------------------------------------------------+
```
Figure 2: MAMS U-plane Protocol Stack with GMA as MX Convergence Layer

Figure 3 shows the trailer-based Multi-Access (MX) PDU (Protocol Data Unit) format [GMA]. If the MX adaptation method is UDP tunneling and "MX header optimization" in the "MX_UP_Setup_Configuration_Request" message [MAMS] is true, the "IP length" and "IP checksum" header fields of the MX PDU SHOULD remain unchanged. Otherwise, they should be updated after adding or removing the GMA trailer in the convergence sublayer.

```
+-----------------------------------------------------+
| IP hdr |         IP payload              | GMA Trailer |
+-----------------------------------------------------+
```

Figure 3: GMA PDU Format

4.3 MPTCP-based MX Convergence Sublayer

Figure 4 shows the MAMS u-plane protocol stack based on MPTCP. Here, MPTCP is reused as the "MX Convergence Sublayer" protocol. Multiple access networks are combined into a single MPTCP connection. Hence, no new u-plane protocol or PDU format is needed in this case.

```
|-----------------------------------------------------|
|                        MPTCP                        |
|-----------------------------------------------------|
|  TCP             |   TCP          |      TCP        |
|-----------------------------------------------------|
| MX Adaptation    | MX Adaptation | MX Adaptation    |
| Sublayer         | Sublayer      | Sublayer         |
| (optional)       | (optional)    | (optional)       |
|-----------------------------------------------------|
| Access #1 IP     | Access #2 IP  | Access #3 IP     |
+-----------------------------------------------------+
```

Figure 4: MAMS U-plane Protocol Stack with MPTCP as MX Convergence Layer

If NCM determines that N-MADP is to be instantiated with MPTCP as the MX Convergence Protocol, it exchanges the support of MPTCP capability in the discovery and capability exchange procedures [MAMS]. MPTCP proxy protocols [MPProxy][MPPlain] SHOULD be used to manage traffic steering and aggregation over multiple delivery connections.

4.4 GRE as MX Convergence Sublayer

Figure 5 shows the MAMS u-plane protocol stack based on GRE (Generic Routing Encapsulation) [GRE2784]. Here, GRE is reused as the "MX Convergence sub-layer" protocol. Multiple access networks are combined

into a single GRE connection. Hence, no new u-plane protocol or PDU format is needed in this case.

```
+-----------------------------------------------------+
|      User Payload (e.g. IP PDU)                     |
|-----------------------------------------------------|
|              GRE as MX Convergence Sublayer         |
|-----------------------------------------------------|
|        GRE Delivery Protocol (e.g. IP)              |
|-----------------------------------------------------|
| MX Adaptation    | MX Adaptation | MX Adaptation    |
| Sublayer         | Sublayer      | Sublayer         |
| (optional)       | (optional)    | (optional)       |
|-----------------------------------------------------|
| Access #1 IP     | Access #2 IP  | Access #3 IP     |
+-----------------------------------------------------+
```

Figure 5: MAMS U-plane Protocol Stack with GRE as MX Convergence Layer

If NCM determines that N-MADP is to be instantiated with GRE as the MX Convergence Protocol, it exchanges the support of GRE capability in the discovery and capability exchange procedures [MAMS].

4.4.1 Transmitter Procedures

Transmitter is the N-MADP or C-MADP instance, instantiated with GRE as the convergence protocol that transmits the GRE packets. The Transmitter receives the User Payload (e.g. IP PDU), encapsulates it with a GRE header and Delivery Protocol (e.g. IP) header to generate the GRE Convergence PDU.

When IP is used as the GRE delivery protocol, the IP header information (e.g. IP address) can be created using the IP header of the user payload or a virtual IP address. The "Protocol Type" field of the delivery header is set to 47 (or 0X2F, i.e. GRE)[IANA].

The GRE header fields are set as specified below,

- If the transmitter is a C-MADP instance, then sets the LSB 16 bits to the value of Connection ID for the Anchor Connection associated with the user payload or sets to 0xFFFF if no Anchor Connection ID needs to be specified.
- All other fields in the GRE header including the remaining bits in the key fields are set as per [GRE_2784][GRE_2890].

4.4.2 Receiver Procedures

Receiver is the N-MADP or C-MADP instance, instantiated with GRE as the convergence protocol that receives the GRE packets. The receiver processes the received packets per the GRE procedures [GRE_2784, GRE_2890] and retrieves the GRE header.

- If the Receiver is an N-MADP instance,
  - o Unless the LSB 16 Bits of the Key field are 0xFFFF, they are interpreted as the Connection ID of Anchor Connection for the user payload. This is used to identify the network path over which the User Payload (GRE Payload) is to be transmitted.
- All other fields in the GRE header, including the remaining bits in the Key fields, are processed as per [GRE_2784][GRE_2890].

The GRE Convergence PDU is passed onto the MX Adaptation Layer (if present) before delivery over one of the network paths.

4.5 Co-existence of MX Adaptation and MX Convergence Sublayers

MAMS u-plane protocols support multiple combinations and instances of user plane protocols to be used in the MX Adaptation and the Convergence sublayers.

For example, one instance of the MX Convergence Layer can be MPTCP Proxy [MPProxy][MPPlain] and another instance can be Trailer-based. The MX Adaptation for each can be either UDP tunnel or IPsec. IPsec may be set up for network paths considered as untrusted by the operator, to protect the TCP subflow between client and MPTCP proxy traversing that network path.

Each of the instances of MAMS user plane, i.e. combination of MX Convergence and MX Adaptation layer protocols, can coexist simultaneously and independently handle different traffic types.

5. MX Convergence Control Message

A UDP connection may be configured between C-MADP and N-MADP to exchange control messages for keep-alive or path quality estimation. The N-MADP end-point IP address and UDP port number of the UDP connection is used to identify MX control PDU. Figure 6 shows the MX control PDU format with the following fields:

- o Type (1 Byte): the type of the MX control message
- o CID (1 Byte): an unsigned integer to identify the anchor and delivery connection of the MX control message
  - + Anchor Connection ID (MSB 4 Bits): an unsigned integer to identify the anchor connection

       + Delivery Connection ID (LSB 4 Bits): an unsigned integer to
         identify the delivery connection
  o MX Control Message (variable): the payload of the MX control
    message

Figure 7 shows the MX convergence control protocol stack, and MX control PDU goes through the MX adaptation sublayer the same way as MX data PDU.

```
                        <----MX Control PDU Payload --------------->
+-------------------------------------------------------------------+
| IP header | UDP Header| Type | CID |       MX Control Message     |
+-------------------------------------------------------------------+
                      Figure 6: MX Control PDU Format

          |-------------------------------------------------------|
          |          MX Convergence Control Messages              |
          |-------------------------------------------------------|
          |                  UDP/IP                               |
          |-------------------------------------------------------|
          | MX Adaptation   | MX Adaptation | MX Adaptation       |
          | Sublayer        | Sublayer      | Sublayer            |
          | (optional)      | (optional)    | (optional)          |
          |-------------------------------------------------------|
          | Access #1 IP    | Access #2 IP  | Access #3 IP        |
          +-------------------------------------------------------+
              Figure 7: MX Convergence Control Protocol Stack
```

5.1 Keep-Alive Message

The "Type" field is set to "0" for Keep-Alive messages. C-MADP may send out Keep-Alive message periodically over one or multiple delivery connections, especially if UDP tunneling is used as the adaptation method for the delivery connection with a NAT function on the path.

A Keep-Alive message is 6 Bytes long, and consists of the following fields:

- o Keep-Alive Sequence Number (2 Bytes): the sequence number of the keep-alive message
- o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.

5.2 Probe Message

The "Type" field is set to "1" for Probe messages.

N-MADP may send out the Probe message for path quality estimation. In response, C-MADP may send back the ACK message.

A Probe message consists of the following fields:

- o Probing Sequence Number (2 Bytes): the sequence number of the Probe REQ message
- o Probing Flag (1 Byte):
  - + Bit #0: a ACK flag to indicate if the ACK message is expected (1) or not (0);
  - + Bit #1: a Probe Type flag to indicate if the Probe message is sent during the initialization phase (0) when the network path is not included for transmission of user data or the active phase (1) when the network path is included for transmission of user data;
  - + Bit #2: a bit flag to indicate the presence of the Reverse Connection ID (R-CID) field.
  - + Bit #3~7: reserved
- o Reverse Connection ID (1 Byte): the connection ID of the delivery connection for sending out the ACK message on the reverse path
- o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.
- o Padding (variable)

The "R-CID" field is only present if both Bit #0 and Bit #2 of the "Probing Flag" field are set to "1". Moreover, Bit #2 of the "Probing Flag" field SHOULD be set to "0" if the Bit #0 is "0", indicating the ACK message is not expected.

If the "R-CID" field is not present but the Bit #0 of the "Probing Flag" field is set to "1", the ACK message SHOULD be sent over the same delivery connection as the Probe message.

The "Padding" field is used to control the length of Probe message.

5.3 Packet Loss Report (PLR) Message

The "Type" field is set to "2" for PLR messages.

C-MADP may send out the PLR messages to report lost MX SDU for example during handover. In response, C-MADP may retransmit the lost MX SDU accordingly.

A PLR message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection which the ACK message is for;

- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the anchor connection which the ACK message is for;
- o ACK number (4 Bytes): the next (in-order) sequence number (SN) that the sender of the PLR message is expecting
- o Number of Loss Bursts (1 Byte)
  For each loss burst, include the following
  - + Sequence Number of the first lost MX SDU in a burst (4 Bytes)
  - + Number of consecutive lost MX SDUs in the burst (1 Byte)

```
          C-MADP                                           N-MADP
              |                                               |
              |<------------------ MX SDU (data packets)--------|
              |                                               |
             +----------------------------------+             |
             |Packet Loss detected              |             |
             +----------------------------------+             |
              |                                               |
              |----- PLR Message ------------------------------>|
              |<-------------retransmit(lost)MX SDUs -----------|
```

Figure 8: MAMS Retransmission Procedure

Figure 8 shows the MAMS retransmission procedure in an example where the lost packet is found and retransmitted.

5.4 First Sequence Number (FSN) Message

The "Type" field is set to "3" for FSN messages.

N-MADP may send out the FSN messages to indicate the oldest MX SDU in its buffer if a lost MX SDU is not found in the buffer after receiving the PLR message from C-MADP. In response, C-MADP SHALL only report packet loss with SN not smaller than FSN.

A FSN message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection which the FSN message is for;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the anchor connection which the FSN message is for;
- o First Sequence Number (4 Bytes): the sequence number (SN) of the oldest MX SDU in the (retransmission) buffer of the sender of the FSN message.

Figure 9 shows the MAMS retransmission procedure in an example where the lost packet is not found.

```
          C-MADP                                              N-MADP
              |                                                  |
              |<------------------- MX SDU (data packets)--------|
              |                                                  |
            +-----------------------------------+                |
            |Packet Loss detected               |                |
            +-----------------------------------+                |
              |                                                  |
              |----- PLR Message ------------------------------>|
              |                              +---------------------+
              |                              |Lost packet not found|
              |                              +---------------------+
              |<-------------FSN message ------------------------|
```

Figure 9: MAMS Retransmission Procedure with FSN

5.5 Coded MX SDU (CMS) Message

The "Type" field is set to "4" for CMS messages.

N-MADP (or C-MADP) may send out the CMS message to support downlink (or uplink) packet loss recovery through coding, e.g. [CRLNC], [CTCP], [RLNC]. A coded MX SDU is generated by applying a network coding algorithm to multiple consecutive (uncoded) MX SDUs, and it is used for fast recovery without retransmission if any of the MX SDUs is lost.

A Coded MX SDU message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection of the coded MX SDU;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the coded MX;
- o Sequence Number (4 Bytes): the sequence number of the first (uncoded) MX SDU used to generate the coded MX SDU.
- o Fragmentation Control (FC) (1 Byte): to provide necessary information for re-assembly, only needed if the coded MX SDU is too long to transport in a single MX control PDU.
- o N (1 Byte): the number of consecutive MX SDUs used to generate the coded MX SDU
- o K (1 Byte): the length (in terms of bits) of the coding coefficient field
- o Coding Coefficient (N x K / 8 Bytes)
  - + a(i): the coding coefficient of the i-th (uncoded) MX SDU

     + padding
  o Coded MX SDU (variable): the coded MX SDU

If K = 0, the simple XOR method is used to generate the Coded MX SDU from N consecutive uncoded MX SDUs, and the a(i) fields are not included in the message.

If the coded MX SDU is too long, it can be fragmented, and transported by multiple MX control PDUs. The N, K, and a(i) fields are only included in the MX PDU carrying the first fragment of the coded MX SDU.

```
          C-MADP                                            N-MADP
              |                                                 |
              |<------------------ MX SDU #1 -------------------|
              |      lost<-------- MX SDU #2 -------------------|
              |<---- CMS Message (MX SDU #1 XOR MX SDU #2)------|
             +----------------------+                           |
             | MX SDU #2 recovered  |                           |
             +----------------------+                           |
              |                                                 |
```

Figure 10: MAMS Packet Recovery Procedure with XOR Coding

5.6 Traffic Splitting Update (TSU) Message

The "Type" field is set to "5" for TSU messages.

N-MADP (or C-MADP) may send out a TSU message if downlink (or uplink) traffic splitting configuration has changed.

A TSU message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class;
- o Sequence Number (2 Bytes): an unsigned integer to identify the TSU message.
- o Flags (1 Byte)
  - + Bit #0: a Reverse Path bit flag to indicate if the traffic splitting configuration is for the reverse path (1) or not (0);
  - + Bit #1: a Bit-Reversal bit flag to indicate if bit-reversal is used in traffic splitting
  - + Others: reserved.
- o Traffic Splitting Configuration Parameters (5 + (N -1) Bytes):

+ StartSN (4 Bytes): the sequence number of the first MX SDU using the traffic splitting configuration provided by the TSU message
+ L (1 Byte): the traffic splitting burst size
+ K(i): the traffic splitting threshold of the i-th delivery connection, where connections are ordered according to their Connection ID.

Let’s use f(x) to denote the traffic splitting function, which maps a MX SDU Sequence Number "x" to the i-th delivery connection.

f(x)=i, if K[i-1]< or = mod(x - StartSN, L) < K[i]

Wherein, 1 < or = i < N, K[0]=0, and K[N]=L.

N is the total number of connections for delivering a data flow, identified by (anchor) Connection ID and Traffic Class ID.

When the bit-reversal bit is set to 1, the burst size L MUST be a power of 2, and the traffic splitting function is

f(x)=i, if K[i-1]< or = F(mod(x - StartSN, L)) < K[i]

Wherein F(.) is the bit reversal function [BITR] of the input variable.

5.7 Acknowledgement Message

The "Type" field is set to "6" for ACK messages.

C-MADP (or N-MADP) SHOULD send out the ACK message in response to the successful reception of a PLR, FSN, or TSU message.

C-MADP SHOULD send out the ACK message in response to a Probe message with the ACK flag set to "1".

The ACK message consists of the following fields:

o Acknowledgment Number (2 Bytes): the sequence number of the received message.
o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.

6 Security Considerations

User data in MAMS framework rely on the security of the underlying network transport paths. When this cannot be assumed, NCM configures use of appropriate protocols for security, e.g. IPsec [RFC4301] [RFC3948], DTLS [RFC6347].

7 IANA Considerations

This draft makes no requests of IANA.

8 Contributing Authors

The editors gratefully acknowledge the following additional contributors in alphabetical order: Salil Agarwal/Nokia, Hema Pentakota/Nokia.

9 References

9.1 Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC4301] Kent, S. and K. Seo, "Security Architecture for the Internet Protocol", RFC 4301, DOI10.17487/RFC4301, December 2005, <http://www.rfc-editor.org/info/rfc4301>.

9.2 Informative References

[RFC6347] Rescorla, E. and N. Modadugu, "Datagram Transport Layer Security Version 1.2", RFC 6347, January 2012, <http://www.rfc-editor.org/info/rfc6347>.

[RFC7296] Kaufman, C., Hoffman, P., Nir, Y., Eronen, P., and T. Kivinen, "Internet Key Exchange Protocol Version 2 (IKEv2)", STD 79, RFC 7296, DOI 10.17487/RFC7296, October 2014, <http://www.rfc-editor.org/info/rfc7296>.

[RFC3948] Huttunen, A., Swander, B., Volpe, V., DiBurro, L., and M. Stenberg, "UDP Encapsulation of IPsec ESP Packets", RFC 3948, DOI 10.17487/RFC3948, January 2005, <http://www.rfc-editor.org/info/rfc3948>.

[MPProxy] X. Wei, C. Xiong, and E. Lopez, "MPTCP proxy mechanisms", https://tools.ietf.org/html/draft-wei-mptcp-proxy-mechanism-02

[MPPlain] M. Boucadair et al, "An MPTCP Option for Network-Assisted MPTCP", https://www.ietf.org/id/draft-boucadair-mptcp-plain-mode-09.txt

[MAMS] S. Kanugovi, S. Vasudevan, F. Baboescu, and J. Zhu, "Multiple Access Management Protocol", https://tools.ietf.org/html/draft-kanugovi-intarea-mams-protocol-03

[GMA] J. Zhu, "Trailer-based Encapsulation Protocols for Generic Multi-Access Convergence", https://tools.ietf.org/html/draft-zhu-intarea-gma-01

[GRE2784] D. Farinacci, et al., "Generic Routing Encapsulation (GRE)", RFC 2784 March 2000, <http://www.rfc-editor.org/info/rfc2784>.

[GRE2890] G. Dommety, "Key and Sequence Number Extensions to GRE", RFC 2890 September 2000, <http://www.rfc-editor.org/info/rfc2890>.

[IANA] https://www.iana.org/assignments/protocol-numbers/protocol-numbers.xhtml

[LWIPEP] 3GPP TS 36.361, "Evolved Universal Terrestrial Radio Access (E-UTRA); LTE-WLAN Radio Level Integration Using Ipsec Tunnel (LWIP) encapsulation; Protocol specification"

[RFC791] Internet Protocol, September 1981

[CRLNC] S Wunderlich, F Gabriel, S Pandi, et al. Caterpillar RLNC (CRLNC): A Practical Finite Sliding Window RLNC Approach, IEEE Access, 2017

[CTCP] M. Kim, et al. Network Coded TCP (CTCP), eprint arXiv:1212.2291, 2012

[RLNC] J. Heide, et al. Random Linear Network Coding (RLNC)-Based Symbol Representation, https://www.ietf.org/id/draft-heide-nwcrg-rlnc-00.txt

[BITR] Alan H. Karp, "Bit reversal on uniprocessors", SIAM Review, 38 (1): 1-26, 1996.

Authors' Addresses

Jing Zhu

Intel

Email: jing.z.zhu@intel.com

SungHoon Seo

Korea Telecom

Email: sh.seo@kt.com

Satish Kanugovi

Nokia

Email: satish.k@nokia.com

Shuping Peng

Huawei

Email: pengshuping@huawei.com

INTAREA J. Zhu
Internet Draft Intel
Intended status: Standards Track S. Seo
Expires: April 1,2020 Korea Telecom
S. Kanugovi
Nokia
S. Peng
Huawei
October 1, 2019

User-Plane Protocols for Multiple Access Management Service
draft-zhu-intarea-mams-user-protocol-07

Status of this Memo

Copyright Notice

Abstract

Today, a device can be simultaneously connected to multiple communication networks based on different technology implementations and network architectures like WiFi, LTE, and DSL. In such multi-connectivity scenario, it is desirable to combine multiple access networks or select the best one to improve quality of experience for a user and improve overall network utilization and efficiency. This document presents the u-plane protocols for a multi access management services (MAMS) framework that can be used to flexibly select the combination of uplink and downlink access and core network paths having the optimal performance, and user plane treatment for improving network utilization and efficiency and enhanced quality of experience for user applications.

Table of Contents

1. Introduction

Multi Access Management Service (MAMS) [MAMS] is a programmable framework to select and configure network paths, as well as adapt to dynamic network conditions, when multiple network connections can serve a client device. It is based on principles of user plane interworking that enables the solution to be deployed as an overlay without impacting the underlying networks.

This document presents the u-plane protocols for enabling the MAMS framework. It co-exists and complements the existing protocols by providing a way to negotiate and configure the protocols based on client and network capabilities. Further it allows exchange of network state information and leveraging network intelligence to optimize the performance of such protocols. An important goal for MAMS is to ensure that there is minimal or no dependency on the actual access technology of the participating links. This allows the scheme to be scalable for addition of newer access technologies and for independent evolution of the existing access technologies.

2. Terminologies

Anchor Connection: refers to the network path from the N-MADP to the Application Server that corresponds to a specific IP anchor that has assigned an IP address to the client.

Delivery Connection: refers to the network path from the N-MADP to the C-MADP.

"Network Connection Manager" (NCM), "Client Connection Manager" (CCM), "Network Multi Access Data Proxy" (N-MADP), and "Client Multi Access Data Proxy" (C-MADP) in this document are to be interpreted as described in [MAMS].

3. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

The terminologies "Network Connection Manager" (NCM), "Client Connection Manager" (CCM), "Network Multi Access Data Proxy" (N-MADP), and "Client Multi Access Data Proxy" (C-MADP) in this document are to be interpreted as described in [MAMS].

4 MAMS User-Plane Protocols

Figure 1 shows the MAMS u-plane protocol stack as specified in [MAMS].

```
             +-----------------------------------------------------+
             |      User Payload (e.g. IP PDU)                     |
             |-----------------------------------------------------|
          +--|-----------------------------------------------------|--+
          |  |-----------------------------------------------------|  |
          |  | Multi-Access (MX) Convergence Sublayer              |  |
          |  |-----------------------------------------------------|  |
          |  |-----------------------------------------------------|  |
          |  | MX Adaptation  | MX Adaptation | MX Adaptation      |  |
          |  | Sublayer       | Sublayer      | Sublayer           |  |
          |  | (optional)     | (optional)    | (optional)         |  |
          |  |-----------------------------------------------------|  |
          |  | Access #1 IP   | Access #2 IP  | Access #3 IP       |  |
          |  +-----------------------------------------------------+  |
          +-----------------------------------------------------------+
```

Figure 1: MAMS U-plane Protocol Stack

It consists of the following two Sublayers:

- o Multi-Access (MX) Convergence Sublayer: This layer performs multi-access specific tasks, e.g., access (path) selection, multi-link (path) aggregation, splitting/reordering, lossless switching, fragmentation, concatenation, keep-alive, and probing etc.
- o Multi-Access (MX) Adaptation Sublayer: This layer performs functions to handle tunneling, network layer security, and NAT.

The MX convergence sublayer operates on top of the MX adaptation sublayer in the protocol stack. From the Transmitter perspective, a User Payload (e.g. IP PDU) is processed by the convergence sublayer first, and then by the adaptation sublayer before being transported over a delivery access connection; from the Receiver perspective, an IP packet received over a delivery connection is processed by the MX adaptation sublayer first, and then by the MX convergence sublayer.

4.1 MX Adaptation Sublayer

The MX adaptation sublayer supports the following mechanisms and protocols while transmitting user plane packets on the network path:

- o UDP Tunneling: The user plane packets of the anchor connection can be encapsulated in a UDP tunnel of a delivery connection between the N-MADP and C-MADP.

o IPsec Tunneling: The user plane packets of the anchor connection are sent through an IPsec tunnel of a delivery connection.
o Client Net Address Translation (NAT): The Client IP address of user plane packet of the anchor connection is changed, and sent over a delivery connection.
o Pass Through: The user plane packets are passing through without any change over the anchor connection.

The MX adaptation sublayer also supports the following mechanisms and protocols to ensure security of user plane packets over the network path.

o IPsec Tunneling: An IPsec [RFC7296] tunnel is established between the N-MADP and C-MADP on the network path that is considered untrusted.
o DTLS: If UDP tunneling is used on the network path that is considered "untrusted", DTLS (Datagram Transport Layer Security) [RFC6347] can be used.

The Client NAT method is the most efficient due to no tunneling overhead. It SHOULD be used if a delivery connection is "trusted" and without NAT function on the path.

The UDP or IPsec Tunnelling method SHOULD be used if a delivery connection has a NAT function placed on the path.

4.2 GMA-based MX Convergence Sublayer

Figure 2 shows the MAMS u-plane protocol stack based on trailer-based encapsulation [GMA]. Multiple access networks are combined into a single IP connection. If NCM determines that N-MADP is to be instantiated with GMA as the MX Convergence Protocol, it exchanges the support of GMA convergence capability in the discovery and capability exchange procedures [MAMS].

```
          +-----------------------------------------------------+
          |                        IP PDU                       |
          |-----------------------------------------------------|
          |               GMA  Convergence Sublayer             |
          |-----------------------------------------------------|
          | MX Adaptation   | MX Adaptation | MX Adaptation     |
          | Sublayer        | Sublayer      | Sublayer          |
          | (optional)      | (optional)    | (optional)        |
          |-----------------------------------------------------|
          | Access #1 IP    | Access #2 IP  | Access #3 IP      |
          +-----------------------------------------------------+
```
Figure 2: MAMS U-plane Protocol Stack with GMA as MX Convergence Layer

Figure 3 shows the trailer-based Multi-Access (MX) PDU (Protocol Data Unit) format [GMA]. If the MX adaptation method is UDP tunneling and "MX header optimization" in the "MX_UP_Setup_Configuration_Request" message [MAMS] is true, the "IP length" and "IP checksum" header fields of the MX PDU SHOULD remain unchanged. Otherwise, they should be updated after adding or removing the GMA trailer in the convergence sublayer.

```
          +-------------------------------------------------------+
          | IP hdr |         IP payload              | GMA Trailer |
          +-------------------------------------------------------+
```

Figure 3: GMA PDU Format

4.3 MPTCP-based MX Convergence Sublayer

Figure 4 shows the MAMS u-plane protocol stack based on MPTCP. Here, MPTCP is reused as the "MX Convergence Sublayer" protocol. Multiple access networks are combined into a single MPTCP connection. Hence, no new u-plane protocol or PDU format is needed in this case.

```
          |-----------------------------------------------------|
          |                        MPTCP                        |
          |-----------------------------------------------------|
          |  TCP            |   TCP          |       TCP        |
          |-----------------------------------------------------|
          | MX Adaptation   | MX Adaptation | MX Adaptation     |
          | Sublayer        | Sublayer      | Sublayer          |
          | (optional)      | (optional)    | (optional)        |
          |-----------------------------------------------------|
          | Access #1 IP    | Access #2 IP  | Access #3 IP      |
          +-----------------------------------------------------+
```

Figure 4: MAMS U-plane Protocol Stack with MPTCP as MX Convergence Layer

If NCM determines that N-MADP is to be instantiated with MPTCP as the MX Convergence Protocol, it exchanges the support of MPTCP capability in the discovery and capability exchange procedures [MAMS]. MPTCP proxy protocols [MPProxy][MPPlain] SHOULD be used to manage traffic steering and aggregation over multiple delivery connections.

4.4 GRE as MX Convergence Sublayer

Figure 5 shows the MAMS u-plane protocol stack based on GRE (Generic Routing Encapsulation) [GRE2784]. Here, GRE is reused as the "MX Convergence sub-layer" protocol. Multiple access networks are combined

into a single GRE connection. Hence, no new u-plane protocol or PDU format is needed in this case.

```
          +------------------------------------------------------+
          |      User Payload (e.g. IP PDU)                      |
          |------------------------------------------------------|
          |              GRE as MX Convergence Sublayer          |
          |------------------------------------------------------|
          |        GRE Delivery Protocol (e.g. IP)               |
          |------------------------------------------------------|
          | MX Adaptation    | MX Adaptation | MX Adaptation     |
          | Sublayer         | Sublayer      | Sublayer          |
          | (optional)       | (optional)    | (optional)        |
          |------------------------------------------------------|
          | Access #1 IP     | Access #2 IP  | Access #3 IP      |
          +------------------------------------------------------+
```
Figure 5: MAMS U-plane Protocol Stack with GRE as MX Convergence Layer

If NCM determines that N-MADP is to be instantiated with GRE as the MX Convergence Protocol, it exchanges the support of GRE capability in the discovery and capability exchange procedures [MAMS].

4.4.1 Transmitter Procedures

Transmitter is the N-MADP or C-MADP instance, instantiated with GRE as the convergence protocol that transmits the GRE packets. The Transmitter receives the User Payload (e.g. IP PDU), encapsulates it with a GRE header and Delivery Protocol (e.g. IP) header to generate the GRE Convergence PDU.

When IP is used as the GRE delivery protocol, the IP header information (e.g. IP address) can be created using the IP header of the user payload or a virtual IP address. The "Protocol Type" field of the delivery header is set to 47 (or 0X2F, i.e. GRE)[IANA].

The GRE header fields are set as specified below,

- If the transmitter is a C-MADP instance, then sets the LSB 16 bits to the value of Connection ID for the Anchor Connection associated with the user payload or sets to 0xFFFF if no Anchor Connection ID needs to be specified.
- All other fields in the GRE header including the remaining bits in the key fields are set as per [GRE_2784][GRE_2890].

4.4.2 Receiver Procedures

Receiver is the N-MADP or C-MADP instance, instantiated with GRE as the convergence protocol that receives the GRE packets. The receiver processes the received packets per the GRE procedures [GRE_2784, GRE_2890] and retrieves the GRE header.

- If the Receiver is an N-MADP instance,
  - o Unless the LSB 16 Bits of the Key field are 0xFFFF, they are interpreted as the Connection ID of Anchor Connection for the user payload. This is used to identify the network path over which the User Payload (GRE Payload) is to be transmitted.
- All other fields in the GRE header, including the remaining bits in the Key fields, are processed as per [GRE_2784][GRE_2890].

The GRE Convergence PDU is passed onto the MX Adaptation Layer (if present) before delivery over one of the network paths.

4.5 Co-existence of MX Adaptation and MX Convergence Sublayers

MAMS u-plane protocols support multiple combinations and instances of user plane protocols to be used in the MX Adaptation and the Convergence sublayers.

For example, one instance of the MX Convergence Layer can be MPTCP Proxy [MPProxy][MPPlain] and another instance can be Trailer-based. The MX Adaptation for each can be either UDP tunnel or IPsec. IPsec may be set up for network paths considered as untrusted by the operator, to protect the TCP subflow between client and MPTCP proxy traversing that network path.

Each of the instances of MAMS user plane, i.e. combination of MX Convergence and MX Adaptation layer protocols, can coexist simultaneously and independently handle different traffic types.

5. MX Convergence Control Message

A UDP connection may be configured between C-MADP and N-MADP to exchange control messages for keep-alive or path quality estimation. The N-MADP end-point IP address and UDP port number of the UDP connection is used to identify MX control PDU. Figure 6 shows the MX control PDU format with the following fields:

- o Type (1 Byte): the type of the MX control message
- o CID (1 Byte): an unsigned integer to identify the anchor and delivery connection of the MX control message
  - + Anchor Connection ID (MSB 4 Bits): an unsigned integer to identify the anchor connection

 + Delivery Connection ID (LSB 4 Bits): an unsigned integer to
 identify the delivery connection
- MX Control Message (variable): the payload of the MX control message

Figure 7 shows the MX convergence control protocol stack, and MX control PDU goes through the MX adaptation sublayer the same way as MX data PDU.

```
                       <----MX Control PDU Payload --------------->
+-------------------------------------------------------------------+
| IP header | UDP Header| Type | CID |       MX Control Message     |
+-------------------------------------------------------------------+
                     Figure 6: MX Control PDU Format

          |------------------------------------------------------|
          |           MX Convergence Control Messages            |
          |------------------------------------------------------|
          |                   UDP/IP                             |
          |------------------------------------------------------|
          | MX Adaptation   | MX Adaptation | MX Adaptation      |
          | Sublayer        | Sublayer      | Sublayer           |
          | (optional)      | (optional)    | (optional)         |
          |------------------------------------------------------|
          | Access #1 IP    | Access #2 IP  | Access #3 IP       |
          +------------------------------------------------------+
              Figure 7: MX Convergence Control Protocol Stack
```

5.1 Keep-Alive Message

The "Type" field is set to "0" for Keep-Alive messages. C-MADP may send out Keep-Alive message periodically over one or multiple delivery connections, especially if UDP tunneling is used as the adaptation method for the delivery connection with a NAT function on the path.

A Keep-Alive message is 6 Bytes long, and consists of the following fields:

- Keep-Alive Sequence Number (2 Bytes): the sequence number of the keep-alive message
- Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.

5.2 Probe Message

The "Type" field is set to "1" for Probe messages.

N-MADP may send out the Probe message for path quality estimation. In response, C-MADP may send back the ACK message.

A Probe message consists of the following fields:

- o Probing Sequence Number (2 Bytes): the sequence number of the Probe REQ message
- o Probing Flag (1 Byte):
  - + Bit #0: a ACK flag to indicate if the ACK message is expected (1) or not (0);
  - + Bit #1: a Probe Type flag to indicate if the Probe message is sent during the initialization phase (0) when the network path is not included for transmission of user data or the active phase (1) when the network path is included for transmission of user data;
  - + Bit #2: a bit flag to indicate the presence of the Reverse Connection ID (R-CID) field.
  - + Bit #3~7: reserved
- o Reverse Connection ID (1 Byte): the connection ID of the delivery connection for sending out the ACK message on the reverse path
- o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.
- o Padding (variable)

The "R-CID" field is only present if both Bit #0 and Bit #2 of the "Probing Flag" field are set to "1". Moreover, Bit #2 of the "Probing Flag" field SHOULD be set to "0" if the Bit #0 is "0", indicating the ACK message is not expected.

If the "R-CID" field is not present but the Bit #0 of the "Probing Flag" field is set to "1", the ACK message SHOULD be sent over the same delivery connection as the Probe message.

The "Padding" field is used to control the length of Probe message.

5.3 Packet Loss Report (PLR) Message

The "Type" field is set to "2" for PLR messages.

C-MADP may send out the PLR messages to report lost MX SDU for example during handover. In response, C-MADP may retransmit the lost MX SDU accordingly.

A PLR message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection which the ACK message is for;

- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the anchor connection which the ACK message is for;
- o ACK number (4 Bytes): the next (in-order) sequence number (SN) that the sender of the PLR message is expecting
- o Number of Loss Bursts (1 Byte)
  For each loss burst, include the following
  - + Sequence Number of the first lost MX SDU in a burst (4 Bytes)
  - + Number of consecutive lost MX SDUs in the burst (1 Byte)

```
      C-MADP                                               N-MADP
         |                                                   |
         |<------------------ MX SDU (data packets)--------|
         |                                                   |
        +----------------------------------+                 |
        |Packet Loss detected              |                 |
        +----------------------------------+                 |
         |                                                   |
         |----- PLR Message ----------------------------->|
         |<-------------retransmit(lost)MX SDUs -----------|
```

Figure 8: MAMS Retransmission Procedure

Figure 8 shows the MAMS retransmission procedure in an example where the lost packet is found and retransmitted.

5.4 First Sequence Number (FSN) Message

The "Type" field is set to "3" for FSN messages.

N-MADP may send out the FSN messages to indicate the oldest MX SDU in its buffer if a lost MX SDU is not found in the buffer after receiving the PLR message from C-MADP. In response, C-MADP SHALL only report packet loss with SN not smaller than FSN.

A FSN message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection which the FSN message is for;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the anchor connection which the FSN message is for;
- o First Sequence Number (4 Bytes): the sequence number (SN) of the oldest MX SDU in the (retransmission) buffer of the sender of the FSN message.

Figure 9 shows the MAMS retransmission procedure in an example where the lost packet is not found.

```
          C-MADP                                              N-MADP
              |                                                  |
              |<------------------- MX SDU (data packets)--------|
              |                                                  |
            +----------------------------------+                 |
            |Packet Loss detected              |                 |
            +----------------------------------+                 |
              |                                                  |
              |----- PLR Message ------------------------------->|
              |                               +---------------------+
              |                               |Lost packet not found|
              |                               +---------------------+
              |<-------------FSN message ------------------------|
```

Figure 9: MAMS Retransmission Procedure with FSN

5.5 Coded MX SDU (CMS) Message

The "Type" field is set to "4" for CMS messages.

N-MADP (or C-MADP) may send out the CMS message to support downlink (or uplink) packet loss recovery through coding, e.g. [CRLNC], [CTCP], [RLNC]. A coded MX SDU is generated by applying a network coding algorithm to multiple consecutive (uncoded) MX SDUs, and it is used for fast recovery without retransmission if any of the MX SDUs is lost.

A Coded MX SDU message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection of the coded MX SDU;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the coded MX;
- o Sequence Number (4 Bytes): the sequence number of the first (uncoded) MX SDU used to generate the coded MX SDU.
- o Fragmentation Control (FC) (1 Byte): to provide necessary information for re-assembly, only needed if the coded MX SDU is too long to transport in a single MX control PDU.
- o N (1 Byte): the number of consecutive MX SDUs used to generate the coded MX SDU
- o K (1 Byte): the length (in terms of bits) of the coding coefficient field
- o Coding Coefficient (N x K / 8 Bytes)
  - + a(i): the coding coefficient of the i-th (uncoded) MX SDU

+ padding
- o Coded MX SDU (variable): the coded MX SDU

If K = 0, the simple XOR method is used to generate the Coded MX SDU from N consecutive uncoded MX SDUs, and the a(i) fields are not included in the message.

If the coded MX SDU is too long, it can be fragmented, and transported by multiple MX control PDUs. The N, K, and a(i) fields are only included in the MX PDU carrying the first fragment of the coded MX SDU.

```
          C-MADP                                            N-MADP
              |                                                 |
              |<------------------ MX SDU #1 -------------------|
              |      lost<-------- MX SDU #2 -------------------|
              |<---- CMS Message (MX SDU #1 XOR MX SDU #2)------|
            +-----------------------+                           |
            | MX SDU #2 recovered  |                            |
            +-----------------------+                           |
              |                                                 |
```

Figure 10: MAMS Packet Recovery Procedure with XOR Coding

5.6 Traffic Splitting Update (TSU) Message

The "Type" field is set to "5" for TSU messages.

N-MADP (or C-MADP) may send out a TSU message if downlink (or uplink) traffic splitting configuration has changed.

A TSU message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class;
- o Sequence Number (2 Bytes): an unsigned integer to identify the TSU message.
- o Flags (1 Byte)
  - + Bit #0: a Reverse Path bit flag to indicate if the traffic splitting configuration is for the reverse path (1) or not (0);
  - + Bit #1: a Bit-Reversal bit flag to indicate if bit-reversal is used in traffic splitting
  - + Others: reserved.
- o Traffic Splitting Configuration Parameters (5 + (N -1) Bytes):

+ StartSN (4 Bytes): the sequence number of the first MX SDU using the traffic splitting configuration provided by the TSU message
+ L (1 Byte): the traffic splitting burst size
+ K(i): the traffic splitting threshold of the i-th delivery connection, where connections are ordered according to their Connection ID.

Let's use f(x) to denote the traffic splitting function, which maps a MX SDU Sequence Number "x" to the i-th delivery connection.

f(x)=i, if K[i-1]< or = mod(x - StartSN, L) < K[i]

Wherein, 1 < or = i < N, K[0]=0, and K[N]=L.

N is the total number of connections for delivering a data flow, identified by (anchor) Connection ID and Traffic Class ID.

When the bit-reversal bit is set to 1, the burst size L MUST be a power of 2, and the traffic splitting function is

f(x)=i, if K[i-1]< or = F(mod(x - StartSN, L)) < K[i]

Wherein F(.) is the bit reversal function [BITR] of the input variable.

5.7 Acknowledgement Message

The "Type" field is set to "6" for ACK messages.

C-MADP (or N-MADP) SHOULD send out the ACK message in response to the successful reception of a PLR, FSN, or TSU message.

C-MADP SHOULD send out the ACK message in response to a Probe message with the ACK flag set to "1".

The ACK message consists of the following fields:

o Acknowledgment Number (2 Bytes): the sequence number of the received message.
o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.

6 Security Considerations

User data in MAMS framework rely on the security of the underlying network transport paths. When this cannot be assumed, NCM configures use of appropriate protocols for security, e.g. IPsec [RFC4301] [RFC3948], DTLS [RFC6347].

7 IANA Considerations

This draft makes no requests of IANA.

8 Contributing Authors

The editors gratefully acknowledge the following additional contributors in alphabetical order: Salil Agarwal/Nokia, Hema Pentakota/Nokia.

9 References

9.1 Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC4301] Kent, S. and K. Seo, "Security Architecture for the Internet Protocol", RFC 4301, DOI10.17487/RFC4301, December 2005, <http://www.rfc-editor.org/info/rfc4301>.

9.2 Informative References

[RFC6347] Rescorla, E. and N. Modadugu, "Datagram Transport Layer Security Version 1.2", RFC 6347, January 2012, <http://www.rfc-editor.org/info/rfc6347>.

[RFC7296] Kaufman, C., Hoffman, P., Nir, Y., Eronen, P., and T. Kivinen, "Internet Key Exchange Protocol Version 2 (IKEv2)", STD 79, RFC 7296, DOI 10.17487/RFC7296, October 2014, <http://www.rfc-editor.org/info/rfc7296>.

[RFC3948] Huttunen, A., Swander, B., Volpe, V., DiBurro, L., and M. Stenberg, "UDP Encapsulation of IPsec ESP Packets", RFC 3948, DOI 10.17487/RFC3948, January 2005, <http://www.rfc-editor.org/info/rfc3948>.

[MPProxy] X. Wei, C. Xiong, and E. Lopez, "MPTCP proxy mechanisms", https://tools.ietf.org/html/draft-wei-mptcp-proxy-mechanism-02

[MPPlain] M. Boucadair et al, "An MPTCP Option for Network-Assisted MPTCP", https://www.ietf.org/id/draft-boucadair-mptcp-plain-mode-09.txt

[MAMS] S. Kanugovi, S. Vasudevan, F. Baboescu, and J. Zhu, "Multiple Access Management Protocol", https://tools.ietf.org/html/draft-kanugovi-intarea-mams-protocol-03

[GMA] J. Zhu, "Trailer-based Encapsulation Protocols for Generic Multi-Access Convergence", https://tools.ietf.org/html/draft-zhu-intarea-gma-01

[GRE2784] D. Farinacci, et al., "Generic Routing Encapsulation (GRE)", RFC 2784 March 2000, <http://www.rfc-editor.org/info/rfc2784>.

[GRE2890] G. Dommety, "Key and Sequence Number Extensions to GRE", RFC 2890 September 2000, <http://www.rfc-editor.org/info/rfc2890>.

[IANA] https://www.iana.org/assignments/protocol-numbers/protocol-numbers.xhtml

[LWIPEP] 3GPP TS 36.361, "Evolved Universal Terrestrial Radio Access (E-UTRA); LTE-WLAN Radio Level Integration Using Ipsec Tunnel (LWIP) encapsulation; Protocol specification"

[RFC791] Internet Protocol, September 1981

[CRLNC] S Wunderlich, F Gabriel, S Pandi, et al. Caterpillar RLNC (CRLNC): A Practical Finite Sliding Window RLNC Approach, IEEE Access, 2017

[CTCP] M. Kim, et al. Network Coded TCP (CTCP), eprint arXiv:1212.2291, 2012

[RLNC] J. Heide, et al. Random Linear Network Coding (RLNC)-Based Symbol Representation, https://www.ietf.org/id/draft-heide-nwcrg-rlnc-00.txt

[BITR] Alan H. Karp, "Bit reversal on uniprocessors", SIAM Review, 38 (1): 1-26, 1996.

Authors' Addresses

Jing Zhu

Intel

Email: jing.z.zhu@intel.com

SungHoon Seo

Korea Telecom

Email: sh.seo@kt.com

Satish Kanugovi

Nokia

Email: satish.k@nokia.com

Shuping Peng

Huawei

Email: pengshuping@huawei.com

INTAREA J. Zhu
Internet Draft Intel
Intended status: Standards Track S. Seo
Expires: April 1,2020 Korea Telecom
S. Kanugovi
Nokia
S. Peng
Huawei
October 1, 2019

User-Plane Protocols for Multiple Access Management Service
draft-zhu-intarea-mams-user-protocol-07

Status of this Memo

Copyright Notice

Abstract

Today, a device can be simultaneously connected to multiple communication networks based on different technology implementations and network architectures like WiFi, LTE, and DSL. In such multi-connectivity scenario, it is desirable to combine multiple access networks or select the best one to improve quality of experience for a user and improve overall network utilization and efficiency. This document presents the u-plane protocols for a multi access management services (MAMS) framework that can be used to flexibly select the combination of uplink and downlink access and core network paths having the optimal performance, and user plane treatment for improving network utilization and efficiency and enhanced quality of experience for user applications.

Table of Contents

1. Introduction

Multi Access Management Service (MAMS) [MAMS] is a programmable framework to select and configure network paths, as well as adapt to dynamic network conditions, when multiple network connections can serve a client device. It is based on principles of user plane interworking that enables the solution to be deployed as an overlay without impacting the underlying networks.

This document presents the u-plane protocols for enabling the MAMS framework. It co-exists and complements the existing protocols by providing a way to negotiate and configure the protocols based on client and network capabilities. Further it allows exchange of network state information and leveraging network intelligence to optimize the performance of such protocols. An important goal for MAMS is to ensure that there is minimal or no dependency on the actual access technology of the participating links. This allows the scheme to be scalable for addition of newer access technologies and for independent evolution of the existing access technologies.

2. Terminologies

Anchor Connection: refers to the network path from the N-MADP to the Application Server that corresponds to a specific IP anchor that has assigned an IP address to the client.

Delivery Connection: refers to the network path from the N-MADP to the C-MADP.

"Network Connection Manager" (NCM), "Client Connection Manager" (CCM), "Network Multi Access Data Proxy" (N-MADP), and "Client Multi Access Data Proxy" (C-MADP) in this document are to be interpreted as described in [MAMS].

3. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

The terminologies "Network Connection Manager" (NCM), "Client Connection Manager" (CCM), "Network Multi Access Data Proxy" (N-MADP), and "Client Multi Access Data Proxy" (C-MADP) in this document are to be interpreted as described in [MAMS].

4 MAMS User-Plane Protocols

Figure 1 shows the MAMS u-plane protocol stack as specified in [MAMS].

```
             +-----------------------------------------------------+
             |      User Payload (e.g. IP PDU)                     |
             |-----------------------------------------------------|
          +--|-----------------------------------------------------|--+
          |  |-----------------------------------------------------|  |
          |  | Multi-Access (MX) Convergence Sublayer              |  |
          |  |-----------------------------------------------------|  |
          |  |-----------------------------------------------------|  |
          |  | MX Adaptation   | MX Adaptation | MX Adaptation     |  |
          |  | Sublayer        | Sublayer      | Sublayer          |  |
          |  | (optional)      | (optional)    | (optional)        |  |
          |  |-----------------------------------------------------|  |
          |  | Access #1 IP    | Access #2 IP  | Access #3 IP      |  |
          |  +-----------------------------------------------------+  |
          +-----------------------------------------------------------+
```

Figure 1: MAMS U-plane Protocol Stack

It consists of the following two Sublayers:

- o Multi-Access (MX) Convergence Sublayer: This layer performs multi-access specific tasks, e.g., access (path) selection, multi-link (path) aggregation, splitting/reordering, lossless switching, fragmentation, concatenation, keep-alive, and probing etc.
- o Multi-Access (MX) Adaptation Sublayer: This layer performs functions to handle tunneling, network layer security, and NAT.

The MX convergence sublayer operates on top of the MX adaptation sublayer in the protocol stack. From the Transmitter perspective, a User Payload (e.g. IP PDU) is processed by the convergence sublayer first, and then by the adaptation sublayer before being transported over a delivery access connection; from the Receiver perspective, an IP packet received over a delivery connection is processed by the MX adaptation sublayer first, and then by the MX convergence sublayer.

4.1 MX Adaptation Sublayer

The MX adaptation sublayer supports the following mechanisms and protocols while transmitting user plane packets on the network path:

- o UDP Tunneling: The user plane packets of the anchor connection can be encapsulated in a UDP tunnel of a delivery connection between the N-MADP and C-MADP.

o IPsec Tunneling: The user plane packets of the anchor connection are sent through an IPsec tunnel of a delivery connection.
o Client Net Address Translation (NAT): The Client IP address of user plane packet of the anchor connection is changed, and sent over a delivery connection.
o Pass Through: The user plane packets are passing through without any change over the anchor connection.

The MX adaptation sublayer also supports the following mechanisms and protocols to ensure security of user plane packets over the network path.

o IPsec Tunneling: An IPsec [RFC7296] tunnel is established between the N-MADP and C-MADP on the network path that is considered untrusted.
o DTLS: If UDP tunneling is used on the network path that is considered "untrusted", DTLS (Datagram Transport Layer Security) [RFC6347] can be used.

The Client NAT method is the most efficient due to no tunneling overhead. It SHOULD be used if a delivery connection is "trusted" and without NAT function on the path.

The UDP or IPsec Tunnelling method SHOULD be used if a delivery connection has a NAT function placed on the path.

4.2 GMA-based MX Convergence Sublayer

Figure 2 shows the MAMS u-plane protocol stack based on trailer-based encapsulation [GMA]. Multiple access networks are combined into a single IP connection. If NCM determines that N-MADP is to be instantiated with GMA as the MX Convergence Protocol, it exchanges the support of GMA convergence capability in the discovery and capability exchange procedures [MAMS].

```
          +-----------------------------------------------------+
          |                       IP PDU                        |
          |-----------------------------------------------------|
          |               GMA  Convergence Sublayer             |
          |-----------------------------------------------------|
          | MX Adaptation   | MX Adaptation | MX Adaptation     |
          | Sublayer        | Sublayer      | Sublayer          |
          | (optional)      | (optional)    | (optional)        |
          |-----------------------------------------------------|
          | Access #1 IP    | Access #2 IP  | Access #3 IP      |
          +-----------------------------------------------------+
```
 Figure 2: MAMS U-plane Protocol Stack with GMA as MX Convergence Layer

Figure 3 shows the trailer-based Multi-Access (MX) PDU (Protocol Data Unit) format [GMA]. If the MX adaptation method is UDP tunneling and "MX header optimization" in the "MX_UP_Setup_Configuration_Request" message [MAMS] is true, the "IP length" and "IP checksum" header fields of the MX PDU SHOULD remain unchanged. Otherwise, they should be updated after adding or removing the GMA trailer in the convergence sublayer.

```
          +-------------------------------------------------------+
          | IP hdr |         IP payload              | GMA Trailer |
          +-------------------------------------------------------+
```

Figure 3: GMA PDU Format

4.3 MPTCP-based MX Convergence Sublayer

Figure 4 shows the MAMS u-plane protocol stack based on MPTCP. Here, MPTCP is reused as the "MX Convergence Sublayer" protocol. Multiple access networks are combined into a single MPTCP connection. Hence, no new u-plane protocol or PDU format is needed in this case.

```
          |-----------------------------------------------------|
          |                       MPTCP                         |
          |-----------------------------------------------------|
          |  TCP             |   TCP           |      TCP       |
          |-----------------------------------------------------|
          | MX Adaptation    | MX Adaptation | MX Adaptation    |
          | Sublayer         | Sublayer      | Sublayer         |
          | (optional)       | (optional)    | (optional)       |
          |-----------------------------------------------------|
          | Access #1 IP     | Access #2 IP  | Access #3 IP     |
          +-----------------------------------------------------+
```

Figure 4: MAMS U-plane Protocol Stack with MPTCP as MX Convergence Layer

If NCM determines that N-MADP is to be instantiated with MPTCP as the MX Convergence Protocol, it exchanges the support of MPTCP capability in the discovery and capability exchange procedures [MAMS]. MPTCP proxy protocols [MPProxy][MPPlain] SHOULD be used to manage traffic steering and aggregation over multiple delivery connections.

4.4 GRE as MX Convergence Sublayer

Figure 5 shows the MAMS u-plane protocol stack based on GRE (Generic Routing Encapsulation) [GRE2784]. Here, GRE is reused as the "MX Convergence sub-layer" protocol. Multiple access networks are combined

into a single GRE connection. Hence, no new u-plane protocol or PDU format is needed in this case.

```
+------------------------------------------------------+
|      User Payload (e.g. IP PDU)                      |
|------------------------------------------------------|
|              GRE as MX Convergence Sublayer          |
|------------------------------------------------------|
|        GRE Delivery Protocol (e.g. IP)               |
|------------------------------------------------------|
| MX Adaptation    | MX Adaptation | MX Adaptation     |
| Sublayer         | Sublayer      | Sublayer          |
| (optional)       | (optional)    | (optional)        |
|------------------------------------------------------|
| Access #1 IP     | Access #2 IP  | Access #3 IP      |
+------------------------------------------------------+
```

Figure 5: MAMS U-plane Protocol Stack with GRE as MX Convergence Layer

If NCM determines that N-MADP is to be instantiated with GRE as the MX Convergence Protocol, it exchanges the support of GRE capability in the discovery and capability exchange procedures [MAMS].

4.4.1 Transmitter Procedures

Transmitter is the N-MADP or C-MADP instance, instantiated with GRE as the convergence protocol that transmits the GRE packets. The Transmitter receives the User Payload (e.g. IP PDU), encapsulates it with a GRE header and Delivery Protocol (e.g. IP) header to generate the GRE Convergence PDU.

When IP is used as the GRE delivery protocol, the IP header information (e.g. IP address) can be created using the IP header of the user payload or a virtual IP address. The "Protocol Type" field of the delivery header is set to 47 (or 0X2F, i.e. GRE)[IANA].

The GRE header fields are set as specified below,

- If the transmitter is a C-MADP instance, then sets the LSB 16 bits to the value of Connection ID for the Anchor Connection associated with the user payload or sets to 0xFFFF if no Anchor Connection ID needs to be specified.
- All other fields in the GRE header including the remaining bits in the key fields are set as per [GRE_2784][GRE_2890].

4.4.2 Receiver Procedures

Receiver is the N-MADP or C-MADP instance, instantiated with GRE as the convergence protocol that receives the GRE packets. The receiver processes the received packets per the GRE procedures [GRE_2784, GRE_2890] and retrieves the GRE header.

- If the Receiver is an N-MADP instance,
  - o Unless the LSB 16 Bits of the Key field are 0xFFFF, they are interpreted as the Connection ID of Anchor Connection for the user payload. This is used to identify the network path over which the User Payload (GRE Payload) is to be transmitted.
- All other fields in the GRE header, including the remaining bits in the Key fields, are processed as per [GRE_2784][GRE_2890].

The GRE Convergence PDU is passed onto the MX Adaptation Layer (if present) before delivery over one of the network paths.

4.5 Co-existence of MX Adaptation and MX Convergence Sublayers

MAMS u-plane protocols support multiple combinations and instances of user plane protocols to be used in the MX Adaptation and the Convergence sublayers.

For example, one instance of the MX Convergence Layer can be MPTCP Proxy [MPProxy][MPPlain] and another instance can be Trailer-based. The MX Adaptation for each can be either UDP tunnel or IPsec. IPsec may be set up for network paths considered as untrusted by the operator, to protect the TCP subflow between client and MPTCP proxy traversing that network path.

Each of the instances of MAMS user plane, i.e. combination of MX Convergence and MX Adaptation layer protocols, can coexist simultaneously and independently handle different traffic types.

5. MX Convergence Control Message

A UDP connection may be configured between C-MADP and N-MADP to exchange control messages for keep-alive or path quality estimation. The N-MADP end-point IP address and UDP port number of the UDP connection is used to identify MX control PDU. Figure 6 shows the MX control PDU format with the following fields:

- o Type (1 Byte): the type of the MX control message
- o CID (1 Byte): an unsigned integer to identify the anchor and delivery connection of the MX control message
  - + Anchor Connection ID (MSB 4 Bits): an unsigned integer to identify the anchor connection

   + Delivery Connection ID (LSB 4 Bits): an unsigned integer to
     identify the delivery connection
- o MX Control Message (variable): the payload of the MX control
    message

Figure 7 shows the MX convergence control protocol stack, and MX control PDU goes through the MX adaptation sublayer the same way as MX data PDU.

```
                        <----MX Control PDU Payload --------------->
+------------------------------------------------------------------+
| IP header | UDP Header| Type | CID |       MX Control Message    |
+------------------------------------------------------------------+
                      Figure 6: MX Control PDU Format

          |-----------------------------------------------------|
          |          MX Convergence Control Messages            |
          |-----------------------------------------------------|
          |                  UDP/IP                             |
          |-----------------------------------------------------|
          | MX Adaptation    | MX Adaptation | MX Adaptation    |
          | Sublayer         | Sublayer      | Sublayer         |
          | (optional)       | (optional)    | (optional)       |
          |-----------------------------------------------------|
          | Access #1 IP     | Access #2 IP  | Access #3 IP     |
          +-----------------------------------------------------+
              Figure 7: MX Convergence Control Protocol Stack
```

5.1 Keep-Alive Message

The "Type" field is set to "0" for Keep-Alive messages. C-MADP may send out Keep-Alive message periodically over one or multiple delivery connections, especially if UDP tunneling is used as the adaptation method for the delivery connection with a NAT function on the path.

A Keep-Alive message is 6 Bytes long, and consists of the following fields:

- o Keep-Alive Sequence Number (2 Bytes): the sequence number of the keep-alive message
- o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.

5.2 Probe Message

The "Type" field is set to "1" for Probe messages.

N-MADP may send out the Probe message for path quality estimation. In response, C-MADP may send back the ACK message.

A Probe message consists of the following fields:

- o Probing Sequence Number (2 Bytes): the sequence number of the Probe REQ message
- o Probing Flag (1 Byte):
  - + Bit #0: a ACK flag to indicate if the ACK message is expected (1) or not (0);
  - + Bit #1: a Probe Type flag to indicate if the Probe message is sent during the initialization phase (0) when the network path is not included for transmission of user data or the active phase (1) when the network path is included for transmission of user data;
  - + Bit #2: a bit flag to indicate the presence of the Reverse Connection ID (R-CID) field.
  - + Bit #3~7: reserved
- o Reverse Connection ID (1 Byte): the connection ID of the delivery connection for sending out the ACK message on the reverse path
- o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.
- o Padding (variable)

The "R-CID" field is only present if both Bit #0 and Bit #2 of the "Probing Flag" field are set to "1". Moreover, Bit #2 of the "Probing Flag" field SHOULD be set to "0" if the Bit #0 is "0", indicating the ACK message is not expected.

If the "R-CID" field is not present but the Bit #0 of the "Probing Flag" field is set to "1", the ACK message SHOULD be sent over the same delivery connection as the Probe message.

The "Padding" field is used to control the length of Probe message.

5.3 Packet Loss Report (PLR) Message

The "Type" field is set to "2" for PLR messages.

C-MADP may send out the PLR messages to report lost MX SDU for example during handover. In response, C-MADP may retransmit the lost MX SDU accordingly.

A PLR message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection which the ACK message is for;

  o Traffic Class ID (1 Byte): an unsigned integer to identify the
     traffic class of the anchor connection which the ACK message is
     for;
  o ACK number (4 Bytes): the next (in-order) sequence number (SN)
     that the sender of the PLR message is expecting
  o Number of Loss Bursts (1 Byte)
     For each loss burst, include the following
       + Sequence Number of the first lost MX SDU in a burst (4 Bytes)
       + Number of consecutive lost MX SDUs in the burst (1 Byte)

```
          C-MADP                                               N-MADP
              |                                                   |
              |<------------------ MX SDU (data packets)--------|
              |                                                   |
             +----------------------------------+                 |
             |Packet Loss detected              |                 |
             +----------------------------------+                 |
              |                                                   |
              |----- PLR Message ----------------------------->|
              |<-------------retransmit(lost)MX SDUs -----------|
```

                Figure 8: MAMS Retransmission Procedure

Figure 8 shows the MAMS retransmission procedure in an example where the lost packet is found and retransmitted.

5.4 First Sequence Number (FSN) Message

The "Type" field is set to "3" for FSN messages.

N-MADP may send out the FSN messages to indicate the oldest MX SDU in its buffer if a lost MX SDU is not found in the buffer after receiving the PLR message from C-MADP. In response, C-MADP SHALL only report packet loss with SN not smaller than FSN.

A FSN message consists of the following fields:

  o Connection ID (1 Byte): an unsigned integer to identify the anchor
     connection which the FSN message is for;
  o Traffic Class ID (1 Byte): an unsigned integer to identify the
     traffic class of the anchor connection which the FSN message is
     for;
  o First Sequence Number (4 Bytes): the sequence number (SN) of the
     oldest MX SDU in the (retransmission) buffer of the sender of the
     FSN message.

Figure 9 shows the MAMS retransmission procedure in an example where the lost packet is not found.

```
          C-MADP                                              N-MADP
              |                                                  |
              |<------------------- MX SDU (data packets)--------|
              |                                                  |
             +----------------------------------+                |
             |Packet Loss detected              |                |
             +----------------------------------+                |
              |                                                  |
              |----- PLR Message ------------------------------->|
              |                                 +---------------------+
              |                                 |Lost packet not found|
              |                                 +---------------------+
              |<-------------FSN message -----------------------|
```

Figure 9: MAMS Retransmission Procedure with FSN

5.5 Coded MX SDU (CMS) Message

The "Type" field is set to "4" for CMS messages.

N-MADP (or C-MADP) may send out the CMS message to support downlink (or uplink) packet loss recovery through coding, e.g. [CRLNC], [CTCP], [RLNC]. A coded MX SDU is generated by applying a network coding algorithm to multiple consecutive (uncoded) MX SDUs, and it is used for fast recovery without retransmission if any of the MX SDUs is lost.

A Coded MX SDU message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection of the coded MX SDU;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the coded MX;
- o Sequence Number (4 Bytes): the sequence number of the first (uncoded) MX SDU used to generate the coded MX SDU.
- o Fragmentation Control (FC) (1 Byte): to provide necessary information for re-assembly, only needed if the coded MX SDU is too long to transport in a single MX control PDU.
- o N (1 Byte): the number of consecutive MX SDUs used to generate the coded MX SDU
- o K (1 Byte): the length (in terms of bits) of the coding coefficient field
- o Coding Coefficient (N x K / 8 Bytes)
  - + a(i): the coding coefficient of the i-th (uncoded) MX SDU

       + padding
  o Coded MX SDU (variable): the coded MX SDU

If K = 0, the simple XOR method is used to generate the Coded MX SDU from N consecutive uncoded MX SDUs, and the a(i) fields are not included in the message.

If the coded MX SDU is too long, it can be fragmented, and transported by multiple MX control PDUs. The N, K, and a(i) fields are only included in the MX PDU carrying the first fragment of the coded MX SDU.

```
          C-MADP                                               N-MADP
              |                                                  |
              |<------------------ MX SDU #1 -------------------|
              |      lost<-------- MX SDU #2 -------------------|
              |<---- CMS Message (MX SDU #1 XOR MX SDU #2)------|
             +----------------------+                            |
             | MX SDU #2 recovered  |                            |
             +----------------------+                            |
              |                                                  |
```

Figure 10: MAMS Packet Recovery Procedure with XOR Coding

5.6 Traffic Splitting Update (TSU) Message

The "Type" field is set to "5" for TSU messages.

N-MADP (or C-MADP) may send out a TSU message if downlink (or uplink) traffic splitting configuration has changed.

A TSU message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class;
- o Sequence Number (2 Bytes): an unsigned integer to identify the TSU message.
- o Flags (1 Byte)
  - + Bit #0: a Reverse Path bit flag to indicate if the traffic splitting configuration is for the reverse path (1) or not (0);
  - + Bit #1: a Bit-Reversal bit flag to indicate if bit-reversal is used in traffic splitting
  - + Others: reserved.
- o Traffic Splitting Configuration Parameters (5 + (N -1) Bytes):

+ StartSN (4 Bytes): the sequence number of the first MX SDU using the traffic splitting configuration provided by the TSU message
+ L (1 Byte): the traffic splitting burst size
+ K(i): the traffic splitting threshold of the i-th delivery connection, where connections are ordered according to their Connection ID.

Let’s use f(x) to denote the traffic splitting function, which maps a MX SDU Sequence Number "x" to the i-th delivery connection.

```
f(x)=i,  if K[i-1]< or = mod(x - StartSN, L) < K[i]
```

Wherein, 1 < or = i < N, K[0]=0, and K[N]=L.

N is the total number of connections for delivering a data flow, identified by (anchor) Connection ID and Traffic Class ID.

When the bit-reversal bit is set to 1, the burst size L MUST be a power of 2, and the traffic splitting function is

```
f(x)=i,  if K[i-1]< or = F(mod(x - StartSN, L)) < K[i]
```

Wherein F(.) is the bit reversal function [BITR] of the input variable.

5.7 Acknowledgement Message

The "Type" field is set to "6" for ACK messages.

C-MADP (or N-MADP) SHOULD send out the ACK message in response to the successful reception of a PLR, FSN, or TSU message.

C-MADP SHOULD send out the ACK message in response to a Probe message with the ACK flag set to "1".

The ACK message consists of the following fields:

o Acknowledgment Number (2 Bytes): the sequence number of the received message.
o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.

6 Security Considerations

User data in MAMS framework rely on the security of the underlying network transport paths. When this cannot be assumed, NCM configures use of appropriate protocols for security, e.g. IPsec [RFC4301] [RFC3948], DTLS [RFC6347].

7 IANA Considerations

This draft makes no requests of IANA.

8 Contributing Authors

The editors gratefully acknowledge the following additional contributors in alphabetical order: Salil Agarwal/Nokia, Hema Pentakota/Nokia.

9 References

9.1 Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC4301] Kent, S. and K. Seo, "Security Architecture for the Internet Protocol", RFC 4301, DOI10.17487/RFC4301, December 2005, <http://www.rfc-editor.org/info/rfc4301>.

9.2 Informative References

[RFC6347] Rescorla, E. and N. Modadugu, "Datagram Transport Layer Security Version 1.2", RFC 6347, January 2012, <http://www.rfc-editor.org/info/rfc6347>.

[RFC7296] Kaufman, C., Hoffman, P., Nir, Y., Eronen, P., and T. Kivinen, "Internet Key Exchange Protocol Version 2 (IKEv2)", STD 79, RFC 7296, DOI 10.17487/RFC7296, October 2014, <http://www.rfc-editor.org/info/rfc7296>.

[RFC3948] Huttunen, A., Swander, B., Volpe, V., DiBurro, L., and M. Stenberg, "UDP Encapsulation of IPsec ESP Packets", RFC 3948, DOI 10.17487/RFC3948, January 2005, <http://www.rfc-editor.org/info/rfc3948>.

[MPProxy] X. Wei, C. Xiong, and E. Lopez, "MPTCP proxy mechanisms", https://tools.ietf.org/html/draft-wei-mptcp-proxy-mechanism-02

[MPPlain] M. Boucadair et al, "An MPTCP Option for Network-Assisted MPTCP", https://www.ietf.org/id/draft-boucadair-mptcp-plain-mode-09.txt

[MAMS] S. Kanugovi, S. Vasudevan, F. Baboescu, and J. Zhu, "Multiple Access Management Protocol", https://tools.ietf.org/html/draft-kanugovi-intarea-mams-protocol-03

[GMA] J. Zhu, "Trailer-based Encapsulation Protocols for Generic Multi-Access Convergence", https://tools.ietf.org/html/draft-zhu-intarea-gma-01

[GRE2784] D. Farinacci, et al., "Generic Routing Encapsulation (GRE)", RFC 2784 March 2000, <http://www.rfc-editor.org/info/rfc2784>.

[GRE2890] G. Dommety, "Key and Sequence Number Extensions to GRE", RFC 2890 September 2000, <http://www.rfc-editor.org/info/rfc2890>.

[IANA] https://www.iana.org/assignments/protocol-numbers/protocol-numbers.xhtml

[LWIPEP] 3GPP TS 36.361, "Evolved Universal Terrestrial Radio Access (E-UTRA); LTE-WLAN Radio Level Integration Using Ipsec Tunnel (LWIP) encapsulation; Protocol specification"

[RFC791] Internet Protocol, September 1981

[CRLNC] S Wunderlich, F Gabriel, S Pandi, et al. Caterpillar RLNC (CRLNC): A Practical Finite Sliding Window RLNC Approach, IEEE Access, 2017

[CTCP] M. Kim, et al. Network Coded TCP (CTCP), eprint arXiv:1212.2291, 2012

[RLNC] J. Heide, et al. Random Linear Network Coding (RLNC)-Based Symbol Representation, https://www.ietf.org/id/draft-heide-nwcrg-rlnc-00.txt

[BITR] Alan H. Karp, "Bit reversal on uniprocessors", SIAM Review, 38 (1): 1-26, 1996.

Authors' Addresses

Jing Zhu

Intel

Email: jing.z.zhu@intel.com

SungHoon Seo

Korea Telecom

Email: sh.seo@kt.com

Satish Kanugovi

Nokia

Email: satish.k@nokia.com

Shuping Peng

Huawei

Email: pengshuping@huawei.com

INTAREA J. Zhu
Internet Draft Intel
Intended status: Standards Track S. Seo
Expires: April 1,2020 Korea Telecom
S. Kanugovi
Nokia
S. Peng
Huawei
October 1, 2019

User-Plane Protocols for Multiple Access Management Service
draft-zhu-intarea-mams-user-protocol-07

Status of this Memo

Copyright Notice

Abstract

Today, a device can be simultaneously connected to multiple communication networks based on different technology implementations and network architectures like WiFi, LTE, and DSL. In such multi-connectivity scenario, it is desirable to combine multiple access networks or select the best one to improve quality of experience for a user and improve overall network utilization and efficiency. This document presents the u-plane protocols for a multi access management services (MAMS) framework that can be used to flexibly select the combination of uplink and downlink access and core network paths having the optimal performance, and user plane treatment for improving network utilization and efficiency and enhanced quality of experience for user applications.

Table of Contents

1. Introduction

Multi Access Management Service (MAMS) [MAMS] is a programmable framework to select and configure network paths, as well as adapt to dynamic network conditions, when multiple network connections can serve a client device. It is based on principles of user plane interworking that enables the solution to be deployed as an overlay without impacting the underlying networks.

This document presents the u-plane protocols for enabling the MAMS framework. It co-exists and complements the existing protocols by providing a way to negotiate and configure the protocols based on client and network capabilities. Further it allows exchange of network state information and leveraging network intelligence to optimize the performance of such protocols. An important goal for MAMS is to ensure that there is minimal or no dependency on the actual access technology of the participating links. This allows the scheme to be scalable for addition of newer access technologies and for independent evolution of the existing access technologies.

2. Terminologies

Anchor Connection: refers to the network path from the N-MADP to the Application Server that corresponds to a specific IP anchor that has assigned an IP address to the client.

Delivery Connection: refers to the network path from the N-MADP to the C-MADP.

"Network Connection Manager" (NCM), "Client Connection Manager" (CCM), "Network Multi Access Data Proxy" (N-MADP), and "Client Multi Access Data Proxy" (C-MADP) in this document are to be interpreted as described in [MAMS].

3. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

The terminologies "Network Connection Manager" (NCM), "Client Connection Manager" (CCM), "Network Multi Access Data Proxy" (N-MADP), and "Client Multi Access Data Proxy" (C-MADP) in this document are to be interpreted as described in [MAMS].

4 MAMS User-Plane Protocols

Figure 1 shows the MAMS u-plane protocol stack as specified in [MAMS].

```
             +-----------------------------------------------------+
             |      User Payload (e.g. IP PDU)                     |
             |-----------------------------------------------------|
          +--|-----------------------------------------------------|--+
          |  |-----------------------------------------------------|  |
          |  | Multi-Access (MX) Convergence Sublayer              |  |
          |  |-----------------------------------------------------|  |
          |  |-----------------------------------------------------|  |
          |  | MX Adaptation   | MX Adaptation | MX Adaptation     |  |
          |  | Sublayer        | Sublayer      | Sublayer          |  |
          |  | (optional)      | (optional)    | (optional)        |  |
          |  |-----------------------------------------------------|  |
          |  | Access #1 IP    | Access #2 IP  | Access #3 IP      |  |
          |  +-----------------------------------------------------+  |
          +-----------------------------------------------------------+
```

Figure 1: MAMS U-plane Protocol Stack

It consists of the following two Sublayers:

- o Multi-Access (MX) Convergence Sublayer: This layer performs multi-access specific tasks, e.g., access (path) selection, multi-link (path) aggregation, splitting/reordering, lossless switching, fragmentation, concatenation, keep-alive, and probing etc.
- o Multi-Access (MX) Adaptation Sublayer: This layer performs functions to handle tunneling, network layer security, and NAT.

The MX convergence sublayer operates on top of the MX adaptation sublayer in the protocol stack. From the Transmitter perspective, a User Payload (e.g. IP PDU) is processed by the convergence sublayer first, and then by the adaptation sublayer before being transported over a delivery access connection; from the Receiver perspective, an IP packet received over a delivery connection is processed by the MX adaptation sublayer first, and then by the MX convergence sublayer.

4.1 MX Adaptation Sublayer

The MX adaptation sublayer supports the following mechanisms and protocols while transmitting user plane packets on the network path:

- o UDP Tunneling: The user plane packets of the anchor connection can be encapsulated in a UDP tunnel of a delivery connection between the N-MADP and C-MADP.

- o IPsec Tunneling: The user plane packets of the anchor connection are sent through an IPsec tunnel of a delivery connection.
- o Client Net Address Translation (NAT): The Client IP address of user plane packet of the anchor connection is changed, and sent over a delivery connection.
- o Pass Through: The user plane packets are passing through without any change over the anchor connection.

The MX adaptation sublayer also supports the following mechanisms and protocols to ensure security of user plane packets over the network path.

- o IPsec Tunneling: An IPsec [RFC7296] tunnel is established between the N-MADP and C-MADP on the network path that is considered untrusted.
- o DTLS: If UDP tunneling is used on the network path that is considered "untrusted", DTLS (Datagram Transport Layer Security) [RFC6347] can be used.

The Client NAT method is the most efficient due to no tunneling overhead. It SHOULD be used if a delivery connection is "trusted" and without NAT function on the path.

The UDP or IPsec Tunnelling method SHOULD be used if a delivery connection has a NAT function placed on the path.

4.2 GMA-based MX Convergence Sublayer

Figure 2 shows the MAMS u-plane protocol stack based on trailer-based encapsulation [GMA]. Multiple access networks are combined into a single IP connection. If NCM determines that N-MADP is to be instantiated with GMA as the MX Convergence Protocol, it exchanges the support of GMA convergence capability in the discovery and capability exchange procedures [MAMS].

```
          +-----------------------------------------------------+
          |                        IP PDU                       |
          |-----------------------------------------------------|
          |               GMA  Convergence Sublayer             |
          |-----------------------------------------------------|
          | MX Adaptation   | MX Adaptation | MX Adaptation     |
          | Sublayer        | Sublayer      | Sublayer          |
          | (optional)      | (optional)    | (optional)        |
          |-----------------------------------------------------|
          | Access #1 IP    | Access #2 IP  | Access #3 IP      |
          +-----------------------------------------------------+
```

Figure 2: MAMS U-plane Protocol Stack with GMA as MX Convergence Layer

Figure 3 shows the trailer-based Multi-Access (MX) PDU (Protocol Data Unit) format [GMA]. If the MX adaptation method is UDP tunneling and "MX header optimization" in the "MX_UP_Setup_Configuration_Request" message [MAMS] is true, the "IP length" and "IP checksum" header fields of the MX PDU SHOULD remain unchanged. Otherwise, they should be updated after adding or removing the GMA trailer in the convergence sublayer.

```
+-------------------------------------------------------+
| IP hdr |         IP payload               | GMA Trailer |
+-------------------------------------------------------+
```

Figure 3: GMA PDU Format

4.3 MPTCP-based MX Convergence Sublayer

Figure 4 shows the MAMS u-plane protocol stack based on MPTCP. Here, MPTCP is reused as the "MX Convergence Sublayer" protocol. Multiple access networks are combined into a single MPTCP connection. Hence, no new u-plane protocol or PDU format is needed in this case.

```
|-----------------------------------------------------|
|                          MPTCP                      |
|-----------------------------------------------------|
|  TCP             |   TCP          |      TCP        |
|-----------------------------------------------------|
| MX Adaptation    | MX Adaptation | MX Adaptation    |
| Sublayer         | Sublayer      | Sublayer         |
| (optional)       | (optional)    | (optional)       |
|-----------------------------------------------------|
| Access #1 IP     | Access #2 IP  | Access #3 IP     |
+-----------------------------------------------------+
```

Figure 4: MAMS U-plane Protocol Stack with MPTCP as MX Convergence Layer

If NCM determines that N-MADP is to be instantiated with MPTCP as the MX Convergence Protocol, it exchanges the support of MPTCP capability in the discovery and capability exchange procedures [MAMS]. MPTCP proxy protocols [MPProxy][MPPlain] SHOULD be used to manage traffic steering and aggregation over multiple delivery connections.

4.4 GRE as MX Convergence Sublayer

Figure 5 shows the MAMS u-plane protocol stack based on GRE (Generic Routing Encapsulation) [GRE2784]. Here, GRE is reused as the "MX Convergence sub-layer" protocol. Multiple access networks are combined

into a single GRE connection. Hence, no new u-plane protocol or PDU format is needed in this case.

```
+-------------------------------------------------------+
|      User Payload (e.g. IP PDU)                       |
|-------------------------------------------------------|
|              GRE as MX Convergence Sublayer           |
|-------------------------------------------------------|
|        GRE Delivery Protocol (e.g. IP)                |
|-------------------------------------------------------|
| MX Adaptation    | MX Adaptation | MX Adaptation      |
| Sublayer         | Sublayer      | Sublayer           |
| (optional)       | (optional)    | (optional)         |
|-------------------------------------------------------|
| Access #1 IP     | Access #2 IP  | Access #3 IP       |
+-------------------------------------------------------+
```
Figure 5: MAMS U-plane Protocol Stack with GRE as MX Convergence Layer

If NCM determines that N-MADP is to be instantiated with GRE as the MX Convergence Protocol, it exchanges the support of GRE capability in the discovery and capability exchange procedures [MAMS].

4.4.1 Transmitter Procedures

Transmitter is the N-MADP or C-MADP instance, instantiated with GRE as the convergence protocol that transmits the GRE packets. The Transmitter receives the User Payload (e.g. IP PDU), encapsulates it with a GRE header and Delivery Protocol (e.g. IP) header to generate the GRE Convergence PDU.

When IP is used as the GRE delivery protocol, the IP header information (e.g. IP address) can be created using the IP header of the user payload or a virtual IP address. The "Protocol Type" field of the delivery header is set to 47 (or 0X2F, i.e. GRE)[IANA].

The GRE header fields are set as specified below,

- If the transmitter is a C-MADP instance, then sets the LSB 16 bits to the value of Connection ID for the Anchor Connection associated with the user payload or sets to 0xFFFF if no Anchor Connection ID needs to be specified.
- All other fields in the GRE header including the remaining bits in the key fields are set as per [GRE_2784][GRE_2890].

4.4.2 Receiver Procedures

Receiver is the N-MADP or C-MADP instance, instantiated with GRE as the convergence protocol that receives the GRE packets. The receiver processes the received packets per the GRE procedures [GRE_2784, GRE_2890] and retrieves the GRE header.

- If the Receiver is an N-MADP instance,
    - o Unless the LSB 16 Bits of the Key field are 0xFFFF, they are interpreted as the Connection ID of Anchor Connection for the user payload. This is used to identify the network path over which the User Payload (GRE Payload) is to be transmitted.
- All other fields in the GRE header, including the remaining bits in the Key fields, are processed as per [GRE_2784][GRE_2890].

The GRE Convergence PDU is passed onto the MX Adaptation Layer (if present) before delivery over one of the network paths.

4.5 Co-existence of MX Adaptation and MX Convergence Sublayers

MAMS u-plane protocols support multiple combinations and instances of user plane protocols to be used in the MX Adaptation and the Convergence sublayers.

For example, one instance of the MX Convergence Layer can be MPTCP Proxy [MPProxy][MPPlain] and another instance can be Trailer-based. The MX Adaptation for each can be either UDP tunnel or IPsec. IPsec may be set up for network paths considered as untrusted by the operator, to protect the TCP subflow between client and MPTCP proxy traversing that network path.

Each of the instances of MAMS user plane, i.e. combination of MX Convergence and MX Adaptation layer protocols, can coexist simultaneously and independently handle different traffic types.

5. MX Convergence Control Message

A UDP connection may be configured between C-MADP and N-MADP to exchange control messages for keep-alive or path quality estimation. The N-MADP end-point IP address and UDP port number of the UDP connection is used to identify MX control PDU. Figure 6 shows the MX control PDU format with the following fields:

- o Type (1 Byte): the type of the MX control message
- o CID (1 Byte): an unsigned integer to identify the anchor and delivery connection of the MX control message
    - + Anchor Connection ID (MSB 4 Bits): an unsigned integer to identify the anchor connection

   + Delivery Connection ID (LSB 4 Bits): an unsigned integer to
     identify the delivery connection
- o MX Control Message (variable): the payload of the MX control
  message

Figure 7 shows the MX convergence control protocol stack, and MX control PDU goes through the MX adaptation sublayer the same way as MX data PDU.

```
                       <----MX Control PDU Payload --------------->
+-------------------------------------------------------------------+
| IP header | UDP Header| Type | CID |       MX Control Message     |
+-------------------------------------------------------------------+
                     Figure 6: MX Control PDU Format

          |-----------------------------------------------------|
          |          MX Convergence Control Messages            |
          |-----------------------------------------------------|
          |                  UDP/IP                             |
          |-----------------------------------------------------|
          | MX Adaptation   | MX Adaptation | MX Adaptation     |
          | Sublayer        | Sublayer      | Sublayer          |
          | (optional)      | (optional)    | (optional)        |
          |-----------------------------------------------------|
          | Access #1 IP    | Access #2 IP  | Access #3 IP      |
          +-----------------------------------------------------+
              Figure 7: MX Convergence Control Protocol Stack
```

5.1 Keep-Alive Message

The "Type" field is set to "0" for Keep-Alive messages. C-MADP may send out Keep-Alive message periodically over one or multiple delivery connections, especially if UDP tunneling is used as the adaptation method for the delivery connection with a NAT function on the path.

A Keep-Alive message is 6 Bytes long, and consists of the following fields:

- o Keep-Alive Sequence Number (2 Bytes): the sequence number of the keep-alive message
- o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.

5.2 Probe Message

The "Type" field is set to "1" for Probe messages.

N-MADP may send out the Probe message for path quality estimation. In response, C-MADP may send back the ACK message.

A Probe message consists of the following fields:

- o Probing Sequence Number (2 Bytes): the sequence number of the Probe REQ message
- o Probing Flag (1 Byte):
  - + Bit #0: a ACK flag to indicate if the ACK message is expected (1) or not (0);
  - + Bit #1: a Probe Type flag to indicate if the Probe message is sent during the initialization phase (0) when the network path is not included for transmission of user data or the active phase (1) when the network path is included for transmission of user data;
  - + Bit #2: a bit flag to indicate the presence of the Reverse Connection ID (R-CID) field.
  - + Bit #3~7: reserved
- o Reverse Connection ID (1 Byte): the connection ID of the delivery connection for sending out the ACK message on the reverse path
- o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.
- o Padding (variable)

The "R-CID" field is only present if both Bit #0 and Bit #2 of the "Probing Flag" field are set to "1". Moreover, Bit #2 of the "Probing Flag" field SHOULD be set to "0" if the Bit #0 is "0", indicating the ACK message is not expected.

If the "R-CID" field is not present but the Bit #0 of the "Probing Flag" field is set to "1", the ACK message SHOULD be sent over the same delivery connection as the Probe message.

The "Padding" field is used to control the length of Probe message.

5.3 Packet Loss Report (PLR) Message

The "Type" field is set to "2" for PLR messages.

C-MADP may send out the PLR messages to report lost MX SDU for example during handover. In response, C-MADP may retransmit the lost MX SDU accordingly.

A PLR message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection which the ACK message is for;

- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the anchor connection which the ACK message is for;
- o ACK number (4 Bytes): the next (in-order) sequence number (SN) that the sender of the PLR message is expecting
- o Number of Loss Bursts (1 Byte)
  For each loss burst, include the following
  - + Sequence Number of the first lost MX SDU in a burst (4 Bytes)
  - + Number of consecutive lost MX SDUs in the burst (1 Byte)

```
          C-MADP                                            N-MADP
              |                                                 |
              |<------------------ MX SDU (data packets)--------|
              |                                                 |
             +----------------------------------+               |
             |Packet Loss detected              |               |
             +----------------------------------+               |
              |                                                 |
              |----- PLR Message ------------------------------>|
              |<-------------retransmit(lost)MX SDUs -----------|
```

Figure 8: MAMS Retransmission Procedure

Figure 8 shows the MAMS retransmission procedure in an example where the lost packet is found and retransmitted.

5.4 First Sequence Number (FSN) Message

The "Type" field is set to "3" for FSN messages.

N-MADP may send out the FSN messages to indicate the oldest MX SDU in its buffer if a lost MX SDU is not found in the buffer after receiving the PLR message from C-MADP. In response, C-MADP SHALL only report packet loss with SN not smaller than FSN.

A FSN message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection which the FSN message is for;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the anchor connection which the FSN message is for;
- o First Sequence Number (4 Bytes): the sequence number (SN) of the oldest MX SDU in the (retransmission) buffer of the sender of the FSN message.

Figure 9 shows the MAMS retransmission procedure in an example where the lost packet is not found.

```
         C-MADP                                           N-MADP
             |                                                |
             |<------------------ MX SDU (data packets)--------|
             |                                                |
            +----------------------------------+              |
            |Packet Loss detected              |              |
            +----------------------------------+              |
             |                                                |
             |----- PLR Message ----------------------------->|
             |                               +----------------------+
             |                               |Lost packet not found|
             |                               +----------------------+
             |<-------------FSN message -----------------------|
```

Figure 9: MAMS Retransmission Procedure with FSN

5.5 Coded MX SDU (CMS) Message

The "Type" field is set to "4" for CMS messages.

N-MADP (or C-MADP) may send out the CMS message to support downlink (or uplink) packet loss recovery through coding, e.g. [CRLNC], [CTCP], [RLNC]. A coded MX SDU is generated by applying a network coding algorithm to multiple consecutive (uncoded) MX SDUs, and it is used for fast recovery without retransmission if any of the MX SDUs is lost.

A Coded MX SDU message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection of the coded MX SDU;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the coded MX;
- o Sequence Number (4 Bytes): the sequence number of the first (uncoded) MX SDU used to generate the coded MX SDU.
- o Fragmentation Control (FC) (1 Byte): to provide necessary information for re-assembly, only needed if the coded MX SDU is too long to transport in a single MX control PDU.
- o N (1 Byte): the number of consecutive MX SDUs used to generate the coded MX SDU
- o K (1 Byte): the length (in terms of bits) of the coding coefficient field
- o Coding Coefficient (N x K / 8 Bytes)
  - + a(i): the coding coefficient of the i-th (uncoded) MX SDU

      + padding
- o Coded MX SDU (variable): the coded MX SDU

If K = 0, the simple XOR method is used to generate the Coded MX SDU from N consecutive uncoded MX SDUs, and the a(i) fields are not included in the message.

If the coded MX SDU is too long, it can be fragmented, and transported by multiple MX control PDUs. The N, K, and a(i) fields are only included in the MX PDU carrying the first fragment of the coded MX SDU.

```
          C-MADP                                              N-MADP
              |                                                   |
              |<------------------ MX SDU #1 -------------------|
              |      lost<-------- MX SDU #2 -------------------|
              |<---- CMS Message (MX SDU #1 XOR MX SDU #2)------|
             +-----------------------+                           |
             | MX SDU #2 recovered  |                           |
             +-----------------------+                           |
              |                                                   |
```

Figure 10: MAMS Packet Recovery Procedure with XOR Coding

5.6 Traffic Splitting Update (TSU) Message

The "Type" field is set to "5" for TSU messages.

N-MADP (or C-MADP) may send out a TSU message if downlink (or uplink) traffic splitting configuration has changed.

A TSU message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class;
- o Sequence Number (2 Bytes): an unsigned integer to identify the TSU message.
- o Flags (1 Byte)
  - + Bit #0: a Reverse Path bit flag to indicate if the traffic splitting configuration is for the reverse path (1) or not (0);
  - + Bit #1: a Bit-Reversal bit flag to indicate if bit-reversal is used in traffic splitting
  - + Others: reserved.
- o Traffic Splitting Configuration Parameters (5 + (N -1) Bytes):

+ StartSN (4 Bytes): the sequence number of the first MX SDU using the traffic splitting configuration provided by the TSU message
+ L (1 Byte): the traffic splitting burst size
+ K(i): the traffic splitting threshold of the i-th delivery connection, where connections are ordered according to their Connection ID.

Let's use f(x) to denote the traffic splitting function, which maps a MX SDU Sequence Number "x" to the i-th delivery connection.

f(x)=i, if K[i-1]< or = mod(x - StartSN, L) < K[i]

Wherein, 1 < or = i < N, K[0]=0, and K[N]=L.

N is the total number of connections for delivering a data flow, identified by (anchor) Connection ID and Traffic Class ID.

When the bit-reversal bit is set to 1, the burst size L MUST be a power of 2, and the traffic splitting function is

f(x)=i, if K[i-1]< or = F(mod(x - StartSN, L)) < K[i]

Wherein F(.) is the bit reversal function [BITR] of the input variable.

5.7 Acknowledgement Message

The "Type" field is set to "6" for ACK messages.

C-MADP (or N-MADP) SHOULD send out the ACK message in response to the successful reception of a PLR, FSN, or TSU message.

C-MADP SHOULD send out the ACK message in response to a Probe message with the ACK flag set to "1".

The ACK message consists of the following fields:

- o Acknowledgment Number (2 Bytes): the sequence number of the received message.
- o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.

6 Security Considerations

User data in MAMS framework rely on the security of the underlying network transport paths. When this cannot be assumed, NCM configures use of appropriate protocols for security, e.g. IPsec [RFC4301] [RFC3948], DTLS [RFC6347].

7 IANA Considerations

This draft makes no requests of IANA.

8 Contributing Authors

The editors gratefully acknowledge the following additional contributors in alphabetical order: Salil Agarwal/Nokia, Hema Pentakota/Nokia.

9 References

9.1 Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC4301] Kent, S. and K. Seo, "Security Architecture for the Internet Protocol", RFC 4301, DOI10.17487/RFC4301, December 2005, <http://www.rfc-editor.org/info/rfc4301>.

9.2 Informative References

[RFC6347] Rescorla, E. and N. Modadugu, "Datagram Transport Layer Security Version 1.2", RFC 6347, January 2012, <http://www.rfc-editor.org/info/rfc6347>.

[RFC7296] Kaufman, C., Hoffman, P., Nir, Y., Eronen, P., and T. Kivinen, "Internet Key Exchange Protocol Version 2 (IKEv2)", STD 79, RFC 7296, DOI 10.17487/RFC7296, October 2014, <http://www.rfc-editor.org/info/rfc7296>.

[RFC3948] Huttunen, A., Swander, B., Volpe, V., DiBurro, L., and M. Stenberg, "UDP Encapsulation of IPsec ESP Packets", RFC 3948, DOI 10.17487/RFC3948, January 2005, <http://www.rfc-editor.org/info/rfc3948>.

[MPProxy] X. Wei, C. Xiong, and E. Lopez, "MPTCP proxy mechanisms", https://tools.ietf.org/html/draft-wei-mptcp-proxy-mechanism-02

[MPPlain] M. Boucadair et al, "An MPTCP Option for Network-Assisted MPTCP", https://www.ietf.org/id/draft-boucadair-mptcp-plain-mode-09.txt

[MAMS] S. Kanugovi, S. Vasudevan, F. Baboescu, and J. Zhu, "Multiple Access Management Protocol", https://tools.ietf.org/html/draft-kanugovi-intarea-mams-protocol-03

[GMA] J. Zhu, "Trailer-based Encapsulation Protocols for Generic Multi-Access Convergence", https://tools.ietf.org/html/draft-zhu-intarea-gma-01

[GRE2784] D. Farinacci, et al., "Generic Routing Encapsulation (GRE)", RFC 2784 March 2000, <http://www.rfc-editor.org/info/rfc2784>.

[GRE2890] G. Dommety, "Key and Sequence Number Extensions to GRE", RFC 2890 September 2000, <http://www.rfc-editor.org/info/rfc2890>.

[IANA] https://www.iana.org/assignments/protocol-numbers/protocol-numbers.xhtml

[LWIPEP] 3GPP TS 36.361, "Evolved Universal Terrestrial Radio Access (E-UTRA); LTE-WLAN Radio Level Integration Using Ipsec Tunnel (LWIP) encapsulation; Protocol specification"

[RFC791] Internet Protocol, September 1981

[CRLNC] S Wunderlich, F Gabriel, S Pandi, et al. Caterpillar RLNC (CRLNC): A Practical Finite Sliding Window RLNC Approach, IEEE Access, 2017

[CTCP] M. Kim, et al. Network Coded TCP (CTCP), eprint arXiv:1212.2291, 2012

[RLNC] J. Heide, et al. Random Linear Network Coding (RLNC)-Based Symbol Representation, https://www.ietf.org/id/draft-heide-nwcrg-rlnc-00.txt

[BITR] Alan H. Karp, "Bit reversal on uniprocessors", SIAM Review, 38 (1): 1-26, 1996.

Authors' Addresses

Jing Zhu

Intel

Email: jing.z.zhu@intel.com

SungHoon Seo

Korea Telecom

Email: sh.seo@kt.com

Satish Kanugovi

Nokia

Email: satish.k@nokia.com

Shuping Peng

Huawei

Email: pengshuping@huawei.com

INTAREA J. Zhu
Internet Draft Intel
Intended status: Standards Track S. Seo
Expires: April 1,2020 Korea Telecom
S. Kanugovi
Nokia
S. Peng
Huawei
October 1, 2019

User-Plane Protocols for Multiple Access Management Service
draft-zhu-intarea-mams-user-protocol-07

Status of this Memo

Copyright Notice

Abstract

Today, a device can be simultaneously connected to multiple communication networks based on different technology implementations and network architectures like WiFi, LTE, and DSL. In such multi-connectivity scenario, it is desirable to combine multiple access networks or select the best one to improve quality of experience for a user and improve overall network utilization and efficiency. This document presents the u-plane protocols for a multi access management services (MAMS) framework that can be used to flexibly select the combination of uplink and downlink access and core network paths having the optimal performance, and user plane treatment for improving network utilization and efficiency and enhanced quality of experience for user applications.

Table of Contents

1. Introduction

Multi Access Management Service (MAMS) [MAMS] is a programmable framework to select and configure network paths, as well as adapt to dynamic network conditions, when multiple network connections can serve a client device. It is based on principles of user plane interworking that enables the solution to be deployed as an overlay without impacting the underlying networks.

This document presents the u-plane protocols for enabling the MAMS framework. It co-exists and complements the existing protocols by providing a way to negotiate and configure the protocols based on client and network capabilities. Further it allows exchange of network state information and leveraging network intelligence to optimize the performance of such protocols. An important goal for MAMS is to ensure that there is minimal or no dependency on the actual access technology of the participating links. This allows the scheme to be scalable for addition of newer access technologies and for independent evolution of the existing access technologies.

2. Terminologies

Anchor Connection: refers to the network path from the N-MADP to the Application Server that corresponds to a specific IP anchor that has assigned an IP address to the client.

Delivery Connection: refers to the network path from the N-MADP to the C-MADP.

"Network Connection Manager" (NCM), "Client Connection Manager" (CCM), "Network Multi Access Data Proxy" (N-MADP), and "Client Multi Access Data Proxy" (C-MADP) in this document are to be interpreted as described in [MAMS].

3. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

The terminologies "Network Connection Manager" (NCM), "Client Connection Manager" (CCM), "Network Multi Access Data Proxy" (N-MADP), and "Client Multi Access Data Proxy" (C-MADP) in this document are to be interpreted as described in [MAMS].

4 MAMS User-Plane Protocols

Figure 1 shows the MAMS u-plane protocol stack as specified in [MAMS].

```
             +-----------------------------------------------------+
             |      User Payload (e.g. IP PDU)                     |
             |-----------------------------------------------------|
          +--|-----------------------------------------------------|--+
          |  |-----------------------------------------------------|  |
          |  | Multi-Access (MX) Convergence Sublayer              |  |
          |  |-----------------------------------------------------|  |
          |  |-----------------------------------------------------|  |
          |  | MX Adaptation  | MX Adaptation | MX Adaptation      |  |
          |  | Sublayer       | Sublayer      | Sublayer           |  |
          |  | (optional)     | (optional)    | (optional)         |  |
          |  |-----------------------------------------------------|  |
          |  | Access #1 IP   | Access #2 IP  | Access #3 IP       |  |
          |  +-----------------------------------------------------+  |
          +-----------------------------------------------------------+
```

Figure 1: MAMS U-plane Protocol Stack

It consists of the following two Sublayers:

o Multi-Access (MX) Convergence Sublayer: This layer performs multi-access specific tasks, e.g., access (path) selection, multi-link (path) aggregation, splitting/reordering, lossless switching, fragmentation, concatenation, keep-alive, and probing etc.

o Multi-Access (MX) Adaptation Sublayer: This layer performs functions to handle tunneling, network layer security, and NAT.

The MX convergence sublayer operates on top of the MX adaptation sublayer in the protocol stack. From the Transmitter perspective, a User Payload (e.g. IP PDU) is processed by the convergence sublayer first, and then by the adaptation sublayer before being transported over a delivery access connection; from the Receiver perspective, an IP packet received over a delivery connection is processed by the MX adaptation sublayer first, and then by the MX convergence sublayer.

4.1 MX Adaptation Sublayer

The MX adaptation sublayer supports the following mechanisms and protocols while transmitting user plane packets on the network path:

o UDP Tunneling: The user plane packets of the anchor connection can be encapsulated in a UDP tunnel of a delivery connection between the N-MADP and C-MADP.

o IPsec Tunneling: The user plane packets of the anchor connection are sent through an IPsec tunnel of a delivery connection.
o Client Net Address Translation (NAT): The Client IP address of user plane packet of the anchor connection is changed, and sent over a delivery connection.
o Pass Through: The user plane packets are passing through without any change over the anchor connection.

The MX adaptation sublayer also supports the following mechanisms and protocols to ensure security of user plane packets over the network path.

o IPsec Tunneling: An IPsec [RFC7296] tunnel is established between the N-MADP and C-MADP on the network path that is considered untrusted.
o DTLS: If UDP tunneling is used on the network path that is considered "untrusted", DTLS (Datagram Transport Layer Security) [RFC6347] can be used.

The Client NAT method is the most efficient due to no tunneling overhead. It SHOULD be used if a delivery connection is "trusted" and without NAT function on the path.

The UDP or IPsec Tunnelling method SHOULD be used if a delivery connection has a NAT function placed on the path.

4.2 GMA-based MX Convergence Sublayer

Figure 2 shows the MAMS u-plane protocol stack based on trailer-based encapsulation [GMA]. Multiple access networks are combined into a single IP connection. If NCM determines that N-MADP is to be instantiated with GMA as the MX Convergence Protocol, it exchanges the support of GMA convergence capability in the discovery and capability exchange procedures [MAMS].

```
+-----------------------------------------------------+
|                        IP PDU                       |
|-----------------------------------------------------|
|               GMA  Convergence Sublayer             |
|-----------------------------------------------------|
| MX Adaptation   | MX Adaptation | MX Adaptation     |
| Sublayer        | Sublayer      | Sublayer          |
| (optional)      | (optional)    | (optional)        |
|-----------------------------------------------------|
| Access #1 IP    | Access #2 IP  | Access #3 IP      |
+-----------------------------------------------------+
```

Figure 2: MAMS U-plane Protocol Stack with GMA as MX Convergence Layer

Figure 3 shows the trailer-based Multi-Access (MX) PDU (Protocol Data Unit) format [GMA]. If the MX adaptation method is UDP tunneling and "MX header optimization" in the "MX_UP_Setup_Configuration_Request" message [MAMS] is true, the "IP length" and "IP checksum" header fields of the MX PDU SHOULD remain unchanged. Otherwise, they should be updated after adding or removing the GMA trailer in the convergence sublayer.

```
          +--------------------------------------------------------+
          | IP hdr |         IP payload              | GMA Trailer |
          +--------------------------------------------------------+
```

Figure 3: GMA PDU Format

4.3 MPTCP-based MX Convergence Sublayer

Figure 4 shows the MAMS u-plane protocol stack based on MPTCP. Here, MPTCP is reused as the "MX Convergence Sublayer" protocol. Multiple access networks are combined into a single MPTCP connection. Hence, no new u-plane protocol or PDU format is needed in this case.

```
          |------------------------------------------------------|
          |                        MPTCP                         |
          |------------------------------------------------------|
          |  TCP            |  TCP          |      TCP           |
          |------------------------------------------------------|
          | MX Adaptation   | MX Adaptation | MX Adaptation      |
          | Sublayer        | Sublayer      | Sublayer           |
          | (optional)      | (optional)    | (optional)         |
          |------------------------------------------------------|
          | Access #1 IP    | Access #2 IP  | Access #3 IP       |
          +------------------------------------------------------+
```

Figure 4: MAMS U-plane Protocol Stack with MPTCP as MX Convergence Layer

If NCM determines that N-MADP is to be instantiated with MPTCP as the MX Convergence Protocol, it exchanges the support of MPTCP capability in the discovery and capability exchange procedures [MAMS]. MPTCP proxy protocols [MPProxy][MPPlain] SHOULD be used to manage traffic steering and aggregation over multiple delivery connections.

4.4 GRE as MX Convergence Sublayer

Figure 5 shows the MAMS u-plane protocol stack based on GRE (Generic Routing Encapsulation) [GRE2784]. Here, GRE is reused as the "MX Convergence sub-layer" protocol. Multiple access networks are combined

into a single GRE connection. Hence, no new u-plane protocol or PDU format is needed in this case.

```
          +------------------------------------------------------+
          |      User Payload (e.g. IP PDU)                      |
          |------------------------------------------------------|
          |              GRE as MX Convergence Sublayer          |
          |------------------------------------------------------|
          |        GRE Delivery Protocol (e.g. IP)               |
          |------------------------------------------------------|
          | MX Adaptation    | MX Adaptation | MX Adaptation     |
          | Sublayer         | Sublayer      | Sublayer          |
          | (optional)       | (optional)    | (optional)        |
          |------------------------------------------------------|
          | Access #1 IP     | Access #2 IP  | Access #3 IP      |
          +------------------------------------------------------+
```

Figure 5: MAMS U-plane Protocol Stack with GRE as MX Convergence Layer

If NCM determines that N-MADP is to be instantiated with GRE as the MX Convergence Protocol, it exchanges the support of GRE capability in the discovery and capability exchange procedures [MAMS].

4.4.1 Transmitter Procedures

Transmitter is the N-MADP or C-MADP instance, instantiated with GRE as the convergence protocol that transmits the GRE packets. The Transmitter receives the User Payload (e.g. IP PDU), encapsulates it with a GRE header and Delivery Protocol (e.g. IP) header to generate the GRE Convergence PDU.

When IP is used as the GRE delivery protocol, the IP header information (e.g. IP address) can be created using the IP header of the user payload or a virtual IP address. The "Protocol Type" field of the delivery header is set to 47 (or 0X2F, i.e. GRE)[IANA].

The GRE header fields are set as specified below,

- If the transmitter is a C-MADP instance, then sets the LSB 16 bits to the value of Connection ID for the Anchor Connection associated with the user payload or sets to 0xFFFF if no Anchor Connection ID needs to be specified.
- All other fields in the GRE header including the remaining bits in the key fields are set as per [GRE_2784][GRE_2890].

4.4.2 Receiver Procedures

Receiver is the N-MADP or C-MADP instance, instantiated with GRE as the convergence protocol that receives the GRE packets. The receiver processes the received packets per the GRE procedures [GRE_2784, GRE_2890] and retrieves the GRE header.

- If the Receiver is an N-MADP instance,
    - o Unless the LSB 16 Bits of the Key field are 0xFFFF, they are interpreted as the Connection ID of Anchor Connection for the user payload. This is used to identify the network path over which the User Payload (GRE Payload) is to be transmitted.
- All other fields in the GRE header, including the remaining bits in the Key fields, are processed as per [GRE_2784][GRE_2890].

The GRE Convergence PDU is passed onto the MX Adaptation Layer (if present) before delivery over one of the network paths.

4.5 Co-existence of MX Adaptation and MX Convergence Sublayers

MAMS u-plane protocols support multiple combinations and instances of user plane protocols to be used in the MX Adaptation and the Convergence sublayers.

For example, one instance of the MX Convergence Layer can be MPTCP Proxy [MPProxy][MPPlain] and another instance can be Trailer-based. The MX Adaptation for each can be either UDP tunnel or IPsec. IPsec may be set up for network paths considered as untrusted by the operator, to protect the TCP subflow between client and MPTCP proxy traversing that network path.

Each of the instances of MAMS user plane, i.e. combination of MX Convergence and MX Adaptation layer protocols, can coexist simultaneously and independently handle different traffic types.

5. MX Convergence Control Message

A UDP connection may be configured between C-MADP and N-MADP to exchange control messages for keep-alive or path quality estimation. The N-MADP end-point IP address and UDP port number of the UDP connection is used to identify MX control PDU. Figure 6 shows the MX control PDU format with the following fields:

- o Type (1 Byte): the type of the MX control message
- o CID (1 Byte): an unsigned integer to identify the anchor and delivery connection of the MX control message
    - + Anchor Connection ID (MSB 4 Bits): an unsigned integer to identify the anchor connection

+ Delivery Connection ID (LSB 4 Bits): an unsigned integer to identify the delivery connection
- o MX Control Message (variable): the payload of the MX control message

Figure 7 shows the MX convergence control protocol stack, and MX control PDU goes through the MX adaptation sublayer the same way as MX data PDU.

```
                       <----MX Control PDU Payload --------------->
+-------------------------------------------------------------------+
| IP header | UDP Header| Type | CID |       MX Control Message    |
+-------------------------------------------------------------------+
```

Figure 6: MX Control PDU Format

```
          |------------------------------------------------------|
          |          MX Convergence Control Messages             |
          |------------------------------------------------------|
          |                  UDP/IP                              |
          |------------------------------------------------------|
          | MX Adaptation   | MX Adaptation | MX Adaptation      |
          | Sublayer        | Sublayer      | Sublayer           |
          | (optional)      | (optional)    | (optional)         |
          |------------------------------------------------------|
          | Access #1 IP    | Access #2 IP  | Access #3 IP       |
          +------------------------------------------------------+
```

Figure 7: MX Convergence Control Protocol Stack

5.1 Keep-Alive Message

The "Type" field is set to "0" for Keep-Alive messages. C-MADP may send out Keep-Alive message periodically over one or multiple delivery connections, especially if UDP tunneling is used as the adaptation method for the delivery connection with a NAT function on the path.

A Keep-Alive message is 6 Bytes long, and consists of the following fields:

- o Keep-Alive Sequence Number (2 Bytes): the sequence number of the keep-alive message
- o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.

5.2 Probe Message

The "Type" field is set to "1" for Probe messages.

N-MADP may send out the Probe message for path quality estimation. In response, C-MADP may send back the ACK message.

A Probe message consists of the following fields:

- o Probing Sequence Number (2 Bytes): the sequence number of the Probe REQ message
- o Probing Flag (1 Byte):
  - + Bit #0: a ACK flag to indicate if the ACK message is expected (1) or not (0);
  - + Bit #1: a Probe Type flag to indicate if the Probe message is sent during the initialization phase (0) when the network path is not included for transmission of user data or the active phase (1) when the network path is included for transmission of user data;
  - + Bit #2: a bit flag to indicate the presence of the Reverse Connection ID (R-CID) field.
  - + Bit #3~7: reserved
- o Reverse Connection ID (1 Byte): the connection ID of the delivery connection for sending out the ACK message on the reverse path
- o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.
- o Padding (variable)

The "R-CID" field is only present if both Bit #0 and Bit #2 of the "Probing Flag" field are set to "1". Moreover, Bit #2 of the "Probing Flag" field SHOULD be set to "0" if the Bit #0 is "0", indicating the ACK message is not expected.

If the "R-CID" field is not present but the Bit #0 of the "Probing Flag" field is set to "1", the ACK message SHOULD be sent over the same delivery connection as the Probe message.

The "Padding" field is used to control the length of Probe message.

5.3 Packet Loss Report (PLR) Message

The "Type" field is set to "2" for PLR messages.

C-MADP may send out the PLR messages to report lost MX SDU for example during handover. In response, C-MADP may retransmit the lost MX SDU accordingly.

A PLR message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection which the ACK message is for;

- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the anchor connection which the ACK message is for;
- o ACK number (4 Bytes): the next (in-order) sequence number (SN) that the sender of the PLR message is expecting
- o Number of Loss Bursts (1 Byte)
  For each loss burst, include the following
  - + Sequence Number of the first lost MX SDU in a burst (4 Bytes)
  - + Number of consecutive lost MX SDUs in the burst (1 Byte)

```
          C-MADP                                              N-MADP
              |                                                  |
              |<------------------ MX SDU (data packets)--------|
              |                                                  |
             +----------------------------------+                |
             |Packet Loss detected              |                |
             +----------------------------------+                |
              |                                                  |
              |----- PLR Message ----------------------------->|
              |<-------------retransmit(lost)MX SDUs -----------|
```

Figure 8: MAMS Retransmission Procedure

Figure 8 shows the MAMS retransmission procedure in an example where the lost packet is found and retransmitted.

5.4 First Sequence Number (FSN) Message

The "Type" field is set to "3" for FSN messages.

N-MADP may send out the FSN messages to indicate the oldest MX SDU in its buffer if a lost MX SDU is not found in the buffer after receiving the PLR message from C-MADP. In response, C-MADP SHALL only report packet loss with SN not smaller than FSN.

A FSN message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection which the FSN message is for;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the anchor connection which the FSN message is for;
- o First Sequence Number (4 Bytes): the sequence number (SN) of the oldest MX SDU in the (retransmission) buffer of the sender of the FSN message.

Figure 9 shows the MAMS retransmission procedure in an example where the lost packet is not found.

```
          C-MADP                                              N-MADP
              |                                                  |
              |<------------------ MX SDU (data packets)--------|
              |                                                  |
            +----------------------------------+                 |
            |Packet Loss detected              |                 |
            +----------------------------------+                 |
              |                                                  |
              |----- PLR Message ------------------------------>|
              |                              +---------------------+
              |                              |Lost packet not found|
              |                              +---------------------+
              |<-------------FSN message -----------------------|
```

Figure 9: MAMS Retransmission Procedure with FSN

5.5 Coded MX SDU (CMS) Message

The "Type" field is set to "4" for CMS messages.

N-MADP (or C-MADP) may send out the CMS message to support downlink (or uplink) packet loss recovery through coding, e.g. [CRLNC], [CTCP], [RLNC]. A coded MX SDU is generated by applying a network coding algorithm to multiple consecutive (uncoded) MX SDUs, and it is used for fast recovery without retransmission if any of the MX SDUs is lost.

A Coded MX SDU message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection of the coded MX SDU;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the coded MX;
- o Sequence Number (4 Bytes): the sequence number of the first (uncoded) MX SDU used to generate the coded MX SDU.
- o Fragmentation Control (FC) (1 Byte): to provide necessary information for re-assembly, only needed if the coded MX SDU is too long to transport in a single MX control PDU.
- o N (1 Byte): the number of consecutive MX SDUs used to generate the coded MX SDU
- o K (1 Byte): the length (in terms of bits) of the coding coefficient field
- o Coding Coefficient (N x K / 8 Bytes)
  - + a(i): the coding coefficient of the i-th (uncoded) MX SDU

+ padding
- o Coded MX SDU (variable): the coded MX SDU

If K = 0, the simple XOR method is used to generate the Coded MX SDU from N consecutive uncoded MX SDUs, and the a(i) fields are not included in the message.

If the coded MX SDU is too long, it can be fragmented, and transported by multiple MX control PDUs. The N, K, and a(i) fields are only included in the MX PDU carrying the first fragment of the coded MX SDU.

```
          C-MADP                                             N-MADP
              |                                                 |
              |<------------------ MX SDU #1 -------------------|
              |      lost<-------- MX SDU #2 -------------------|
              |<---- CMS Message (MX SDU #1 XOR MX SDU #2)------|
             +-----------------------+                          |
             | MX SDU #2 recovered  |                           |
             +-----------------------+                          |
              |                                                 |
```

Figure 10: MAMS Packet Recovery Procedure with XOR Coding

5.6 Traffic Splitting Update (TSU) Message

The "Type" field is set to "5" for TSU messages.

N-MADP (or C-MADP) may send out a TSU message if downlink (or uplink) traffic splitting configuration has changed.

A TSU message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class;
- o Sequence Number (2 Bytes): an unsigned integer to identify the TSU message.
- o Flags (1 Byte)
  - + Bit #0: a Reverse Path bit flag to indicate if the traffic splitting configuration is for the reverse path (1) or not (0);
  - + Bit #1: a Bit-Reversal bit flag to indicate if bit-reversal is used in traffic splitting
  - + Others: reserved.
- o Traffic Splitting Configuration Parameters (5 + (N -1) Bytes):

+ StartSN (4 Bytes): the sequence number of the first MX SDU using the traffic splitting configuration provided by the TSU message
+ L (1 Byte): the traffic splitting burst size
+ K(i): the traffic splitting threshold of the i-th delivery connection, where connections are ordered according to their Connection ID.

Let's use f(x) to denote the traffic splitting function, which maps a MX SDU Sequence Number "x" to the i-th delivery connection.

f(x)=i, if K[i-1]< or = mod(x - StartSN, L) < K[i]

Wherein, 1 < or = i < N, K[0]=0, and K[N]=L.

N is the total number of connections for delivering a data flow, identified by (anchor) Connection ID and Traffic Class ID.

When the bit-reversal bit is set to 1, the burst size L MUST be a power of 2, and the traffic splitting function is

f(x)=i, if K[i-1]< or = F(mod(x - StartSN, L)) < K[i]

Wherein F(.) is the bit reversal function [BITR] of the input variable.

5.7 Acknowledgement Message

The "Type" field is set to "6" for ACK messages.

C-MADP (or N-MADP) SHOULD send out the ACK message in response to the successful reception of a PLR, FSN, or TSU message.

C-MADP SHOULD send out the ACK message in response to a Probe message with the ACK flag set to "1".

The ACK message consists of the following fields:

- o Acknowledgment Number (2 Bytes): the sequence number of the received message.
- o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.

6 Security Considerations

User data in MAMS framework rely on the security of the underlying network transport paths. When this cannot be assumed, NCM configures use of appropriate protocols for security, e.g. IPsec [RFC4301] [RFC3948], DTLS [RFC6347].

7 IANA Considerations

This draft makes no requests of IANA.

8 Contributing Authors

The editors gratefully acknowledge the following additional contributors in alphabetical order: Salil Agarwal/Nokia, Hema Pentakota/Nokia.

9 References

9.1 Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC4301] Kent, S. and K. Seo, "Security Architecture for the Internet Protocol", RFC 4301, DOI10.17487/RFC4301, December 2005, <http://www.rfc-editor.org/info/rfc4301>.

9.2 Informative References

[RFC6347] Rescorla, E. and N. Modadugu, "Datagram Transport Layer Security Version 1.2", RFC 6347, January 2012, <http://www.rfc-editor.org/info/rfc6347>.

[RFC7296] Kaufman, C., Hoffman, P., Nir, Y., Eronen, P., and T. Kivinen, "Internet Key Exchange Protocol Version 2 (IKEv2)", STD 79, RFC 7296, DOI 10.17487/RFC7296, October 2014, <http://www.rfc-editor.org/info/rfc7296>.

[RFC3948] Huttunen, A., Swander, B., Volpe, V., DiBurro, L., and M. Stenberg, "UDP Encapsulation of IPsec ESP Packets", RFC 3948, DOI 10.17487/RFC3948, January 2005, <http://www.rfc-editor.org/info/rfc3948>.

[MPProxy] X. Wei, C. Xiong, and E. Lopez, "MPTCP proxy mechanisms", https://tools.ietf.org/html/draft-wei-mptcp-proxy-mechanism-02

[MPPlain] M. Boucadair et al, "An MPTCP Option for Network-Assisted MPTCP", https://www.ietf.org/id/draft-boucadair-mptcp-plain-mode-09.txt

[MAMS] S. Kanugovi, S. Vasudevan, F. Baboescu, and J. Zhu, "Multiple Access Management Protocol", https://tools.ietf.org/html/draft-kanugovi-intarea-mams-protocol-03

[GMA] J. Zhu, "Trailer-based Encapsulation Protocols for Generic Multi-Access Convergence", https://tools.ietf.org/html/draft-zhu-intarea-gma-01

[GRE2784] D. Farinacci, et al., "Generic Routing Encapsulation (GRE)", RFC 2784 March 2000, <http://www.rfc-editor.org/info/rfc2784>.

[GRE2890] G. Dommety, "Key and Sequence Number Extensions to GRE", RFC 2890 September 2000, <http://www.rfc-editor.org/info/rfc2890>.

[IANA] https://www.iana.org/assignments/protocol-numbers/protocol-numbers.xhtml

[LWIPEP] 3GPP TS 36.361, "Evolved Universal Terrestrial Radio Access (E-UTRA); LTE-WLAN Radio Level Integration Using Ipsec Tunnel (LWIP) encapsulation; Protocol specification"

[RFC791] Internet Protocol, September 1981

[CRLNC] S Wunderlich, F Gabriel, S Pandi, et al. Caterpillar RLNC (CRLNC): A Practical Finite Sliding Window RLNC Approach, IEEE Access, 2017

[CTCP] M. Kim, et al. Network Coded TCP (CTCP), eprint arXiv:1212.2291, 2012

[RLNC] J. Heide, et al. Random Linear Network Coding (RLNC)-Based Symbol Representation, https://www.ietf.org/id/draft-heide-nwcrg-rlnc-00.txt

[BITR] Alan H. Karp, "Bit reversal on uniprocessors", SIAM Review, 38 (1): 1-26, 1996.

Authors’ Addresses

Jing Zhu

Intel

Email: jing.z.zhu@intel.com

SungHoon Seo

Korea Telecom

Email: sh.seo@kt.com

Satish Kanugovi

Nokia

Email: satish.k@nokia.com

Shuping Peng

Huawei

Email: pengshuping@huawei.com

INTAREA
Internet Draft
Intended status: Standards Track
Expires: April 1,2020

J. Zhu
Intel
S. Seo
Korea Telecom
S. Kanugovi
Nokia
S. Peng
Huawei

October 1, 2019

User-Plane Protocols for Multiple Access Management Service
draft-zhu-intarea-mams-user-protocol-07

Status of this Memo

Copyright Notice

Abstract

Today, a device can be simultaneously connected to multiple communication networks based on different technology implementations and network architectures like WiFi, LTE, and DSL. In such multi-connectivity scenario, it is desirable to combine multiple access networks or select the best one to improve quality of experience for a user and improve overall network utilization and efficiency. This document presents the u-plane protocols for a multi access management services (MAMS) framework that can be used to flexibly select the combination of uplink and downlink access and core network paths having the optimal performance, and user plane treatment for improving network utilization and efficiency and enhanced quality of experience for user applications.

Table of Contents

1. Introduction

Multi Access Management Service (MAMS) [MAMS] is a programmable framework to select and configure network paths, as well as adapt to dynamic network conditions, when multiple network connections can serve a client device. It is based on principles of user plane interworking that enables the solution to be deployed as an overlay without impacting the underlying networks.

This document presents the u-plane protocols for enabling the MAMS framework. It co-exists and complements the existing protocols by providing a way to negotiate and configure the protocols based on client and network capabilities. Further it allows exchange of network state information and leveraging network intelligence to optimize the performance of such protocols. An important goal for MAMS is to ensure that there is minimal or no dependency on the actual access technology of the participating links. This allows the scheme to be scalable for addition of newer access technologies and for independent evolution of the existing access technologies.

2. Terminologies

Anchor Connection: refers to the network path from the N-MADP to the Application Server that corresponds to a specific IP anchor that has assigned an IP address to the client.

Delivery Connection: refers to the network path from the N-MADP to the C-MADP.

"Network Connection Manager" (NCM), "Client Connection Manager" (CCM), "Network Multi Access Data Proxy" (N-MADP), and "Client Multi Access Data Proxy" (C-MADP) in this document are to be interpreted as described in [MAMS].

3. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

The terminologies "Network Connection Manager" (NCM), "Client Connection Manager" (CCM), "Network Multi Access Data Proxy" (N-MADP), and "Client Multi Access Data Proxy" (C-MADP) in this document are to be interpreted as described in [MAMS].

4 MAMS User-Plane Protocols

Figure 1 shows the MAMS u-plane protocol stack as specified in [MAMS].

```
             +-----------------------------------------------------+
             |      User Payload (e.g. IP PDU)                     |
             |-----------------------------------------------------|
          +--|-----------------------------------------------------|--+
          |  |-----------------------------------------------------|  |
          |  | Multi-Access (MX) Convergence Sublayer              |  |
          |  |-----------------------------------------------------|  |
          |  |-----------------------------------------------------|  |
          |  | MX Adaptation    | MX Adaptation  | MX Adaptation   |  |
          |  | Sublayer         | Sublayer       | Sublayer        |  |
          |  | (optional)       | (optional)     | (optional)      |  |
          |  |-----------------------------------------------------|  |
          |  | Access #1 IP     | Access #2 IP   | Access #3 IP    |  |
          |  +-----------------------------------------------------+  |
          +-----------------------------------------------------------+
```

Figure 1: MAMS U-plane Protocol Stack

It consists of the following two Sublayers:

- o Multi-Access (MX) Convergence Sublayer: This layer performs multi-access specific tasks, e.g., access (path) selection, multi-link (path) aggregation, splitting/reordering, lossless switching, fragmentation, concatenation, keep-alive, and probing etc.
- o Multi-Access (MX) Adaptation Sublayer: This layer performs functions to handle tunneling, network layer security, and NAT.

The MX convergence sublayer operates on top of the MX adaptation sublayer in the protocol stack. From the Transmitter perspective, a User Payload (e.g. IP PDU) is processed by the convergence sublayer first, and then by the adaptation sublayer before being transported over a delivery access connection; from the Receiver perspective, an IP packet received over a delivery connection is processed by the MX adaptation sublayer first, and then by the MX convergence sublayer.

4.1 MX Adaptation Sublayer

The MX adaptation sublayer supports the following mechanisms and protocols while transmitting user plane packets on the network path:

- o UDP Tunneling: The user plane packets of the anchor connection can be encapsulated in a UDP tunnel of a delivery connection between the N-MADP and C-MADP.

- o IPsec Tunneling: The user plane packets of the anchor connection are sent through an IPsec tunnel of a delivery connection.
- o Client Net Address Translation (NAT): The Client IP address of user plane packet of the anchor connection is changed, and sent over a delivery connection.
- o Pass Through: The user plane packets are passing through without any change over the anchor connection.

The MX adaptation sublayer also supports the following mechanisms and protocols to ensure security of user plane packets over the network path.

- o IPsec Tunneling: An IPsec [RFC7296] tunnel is established between the N-MADP and C-MADP on the network path that is considered untrusted.
- o DTLS: If UDP tunneling is used on the network path that is considered "untrusted", DTLS (Datagram Transport Layer Security) [RFC6347] can be used.

The Client NAT method is the most efficient due to no tunneling overhead. It SHOULD be used if a delivery connection is "trusted" and without NAT function on the path.

The UDP or IPsec Tunnelling method SHOULD be used if a delivery connection has a NAT function placed on the path.

4.2 GMA-based MX Convergence Sublayer

Figure 2 shows the MAMS u-plane protocol stack based on trailer-based encapsulation [GMA]. Multiple access networks are combined into a single IP connection. If NCM determines that N-MADP is to be instantiated with GMA as the MX Convergence Protocol, it exchanges the support of GMA convergence capability in the discovery and capability exchange procedures [MAMS].

```
+-----------------------------------------------------+
|                        IP PDU                       |
|-----------------------------------------------------|
|               GMA  Convergence Sublayer             |
|-----------------------------------------------------|
| MX Adaptation    | MX Adaptation | MX Adaptation    |
| Sublayer         | Sublayer      | Sublayer         |
| (optional)       | (optional)    | (optional)       |
|-----------------------------------------------------|
| Access #1 IP     | Access #2 IP  | Access #3 IP     |
+-----------------------------------------------------+
```

Figure 2: MAMS U-plane Protocol Stack with GMA as MX Convergence Layer

Figure 3 shows the trailer-based Multi-Access (MX) PDU (Protocol Data Unit) format [GMA]. If the MX adaptation method is UDP tunneling and "MX header optimization" in the "MX_UP_Setup_Configuration_Request" message [MAMS] is true, the "IP length" and "IP checksum" header fields of the MX PDU SHOULD remain unchanged. Otherwise, they should be updated after adding or removing the GMA trailer in the convergence sublayer.

```
          +--------------------------------------------------------+
          | IP hdr |        IP payload               | GMA Trailer |
          +--------------------------------------------------------+
```

Figure 3: GMA PDU Format

4.3 MPTCP-based MX Convergence Sublayer

Figure 4 shows the MAMS u-plane protocol stack based on MPTCP. Here, MPTCP is reused as the "MX Convergence Sublayer" protocol. Multiple access networks are combined into a single MPTCP connection. Hence, no new u-plane protocol or PDU format is needed in this case.

```
          |------------------------------------------------------|
          |                        MPTCP                         |
          |------------------------------------------------------|
          |  TCP            |   TCP          |      TCP          |
          |------------------------------------------------------|
          | MX Adaptation   | MX Adaptation | MX Adaptation      |
          | Sublayer        | Sublayer      | Sublayer           |
          | (optional)      | (optional)    | (optional)         |
          |------------------------------------------------------|
          | Access #1 IP    | Access #2 IP  | Access #3 IP       |
          +------------------------------------------------------+
```

Figure 4: MAMS U-plane Protocol Stack with MPTCP as MX Convergence Layer

If NCM determines that N-MADP is to be instantiated with MPTCP as the MX Convergence Protocol, it exchanges the support of MPTCP capability in the discovery and capability exchange procedures [MAMS]. MPTCP proxy protocols [MPProxy][MPPlain] SHOULD be used to manage traffic steering and aggregation over multiple delivery connections.

4.4 GRE as MX Convergence Sublayer

Figure 5 shows the MAMS u-plane protocol stack based on GRE (Generic Routing Encapsulation) [GRE2784]. Here, GRE is reused as the "MX Convergence sub-layer" protocol. Multiple access networks are combined

into a single GRE connection. Hence, no new u-plane protocol or PDU format is needed in this case.

```
          +-------------------------------------------------------+
          |      User Payload (e.g. IP PDU)                       |
          |-------------------------------------------------------|
          |              GRE as MX Convergence Sublayer           |
          |-------------------------------------------------------|
          |        GRE Delivery Protocol (e.g. IP)                |
          |-------------------------------------------------------|
          | MX Adaptation    | MX Adaptation | MX Adaptation      |
          | Sublayer         | Sublayer      | Sublayer           |
          | (optional)       | (optional)    | (optional)         |
          |-------------------------------------------------------|
          | Access #1 IP     | Access #2 IP  | Access #3 IP       |
          +-------------------------------------------------------+
```

Figure 5: MAMS U-plane Protocol Stack with GRE as MX Convergence Layer

If NCM determines that N-MADP is to be instantiated with GRE as the MX Convergence Protocol, it exchanges the support of GRE capability in the discovery and capability exchange procedures [MAMS].

4.4.1 Transmitter Procedures

Transmitter is the N-MADP or C-MADP instance, instantiated with GRE as the convergence protocol that transmits the GRE packets. The Transmitter receives the User Payload (e.g. IP PDU), encapsulates it with a GRE header and Delivery Protocol (e.g. IP) header to generate the GRE Convergence PDU.

When IP is used as the GRE delivery protocol, the IP header information (e.g. IP address) can be created using the IP header of the user payload or a virtual IP address. The "Protocol Type" field of the delivery header is set to 47 (or 0X2F, i.e. GRE)[IANA].

The GRE header fields are set as specified below,

- If the transmitter is a C-MADP instance, then sets the LSB 16 bits to the value of Connection ID for the Anchor Connection associated with the user payload or sets to 0xFFFF if no Anchor Connection ID needs to be specified.
- All other fields in the GRE header including the remaining bits in the key fields are set as per [GRE_2784][GRE_2890].

4.4.2 Receiver Procedures

Receiver is the N-MADP or C-MADP instance, instantiated with GRE as the convergence protocol that receives the GRE packets. The receiver processes the received packets per the GRE procedures [GRE_2784, GRE_2890] and retrieves the GRE header.

- If the Receiver is an N-MADP instance,
  - o Unless the LSB 16 Bits of the Key field are 0xFFFF, they are interpreted as the Connection ID of Anchor Connection for the user payload. This is used to identify the network path over which the User Payload (GRE Payload) is to be transmitted.
- All other fields in the GRE header, including the remaining bits in the Key fields, are processed as per [GRE_2784][GRE_2890].

The GRE Convergence PDU is passed onto the MX Adaptation Layer (if present) before delivery over one of the network paths.

4.5 Co-existence of MX Adaptation and MX Convergence Sublayers

MAMS u-plane protocols support multiple combinations and instances of user plane protocols to be used in the MX Adaptation and the Convergence sublayers.

For example, one instance of the MX Convergence Layer can be MPTCP Proxy [MPProxy][MPPlain] and another instance can be Trailer-based. The MX Adaptation for each can be either UDP tunnel or IPsec. IPsec may be set up for network paths considered as untrusted by the operator, to protect the TCP subflow between client and MPTCP proxy traversing that network path.

Each of the instances of MAMS user plane, i.e. combination of MX Convergence and MX Adaptation layer protocols, can coexist simultaneously and independently handle different traffic types.

5. MX Convergence Control Message

A UDP connection may be configured between C-MADP and N-MADP to exchange control messages for keep-alive or path quality estimation. The N-MADP end-point IP address and UDP port number of the UDP connection is used to identify MX control PDU. Figure 6 shows the MX control PDU format with the following fields:

- o Type (1 Byte): the type of the MX control message
- o CID (1 Byte): an unsigned integer to identify the anchor and delivery connection of the MX control message
  - + Anchor Connection ID (MSB 4 Bits): an unsigned integer to identify the anchor connection

+ Delivery Connection ID (LSB 4 Bits): an unsigned integer to identify the delivery connection

o MX Control Message (variable): the payload of the MX control message

Figure 7 shows the MX convergence control protocol stack, and MX control PDU goes through the MX adaptation sublayer the same way as MX data PDU.

```
                        <----MX Control PDU Payload --------------->
+--------------------------------------------------------------------+
| IP header | UDP Header| Type | CID |       MX Control Message     |
+--------------------------------------------------------------------+
```

Figure 6: MX Control PDU Format

```
          |------------------------------------------------------|
          |           MX Convergence Control Messages            |
          |------------------------------------------------------|
          |                   UDP/IP                             |
          |------------------------------------------------------|
          | MX Adaptation   | MX Adaptation | MX Adaptation      |
          | Sublayer        | Sublayer      | Sublayer           |
          | (optional)      | (optional)    | (optional)         |
          |------------------------------------------------------|
          | Access #1 IP    | Access #2 IP  | Access #3 IP       |
          +------------------------------------------------------+
```

Figure 7: MX Convergence Control Protocol Stack

5.1 Keep-Alive Message

The "Type" field is set to "0" for Keep-Alive messages. C-MADP may send out Keep-Alive message periodically over one or multiple delivery connections, especially if UDP tunneling is used as the adaptation method for the delivery connection with a NAT function on the path.

A Keep-Alive message is 6 Bytes long, and consists of the following fields:

o Keep-Alive Sequence Number (2 Bytes): the sequence number of the keep-alive message

o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.

5.2 Probe Message

The "Type" field is set to "1" for Probe messages.

N-MADP may send out the Probe message for path quality estimation. In response, C-MADP may send back the ACK message.

A Probe message consists of the following fields:

- o Probing Sequence Number (2 Bytes): the sequence number of the Probe REQ message
- o Probing Flag (1 Byte):
  - + Bit #0: a ACK flag to indicate if the ACK message is expected (1) or not (0);
  - + Bit #1: a Probe Type flag to indicate if the Probe message is sent during the initialization phase (0) when the network path is not included for transmission of user data or the active phase (1) when the network path is included for transmission of user data;
  - + Bit #2: a bit flag to indicate the presence of the Reverse Connection ID (R-CID) field.
  - + Bit #3~7: reserved
- o Reverse Connection ID (1 Byte): the connection ID of the delivery connection for sending out the ACK message on the reverse path
- o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.
- o Padding (variable)

The "R-CID" field is only present if both Bit #0 and Bit #2 of the "Probing Flag" field are set to "1". Moreover, Bit #2 of the "Probing Flag" field SHOULD be set to "0" if the Bit #0 is "0", indicating the ACK message is not expected.

If the "R-CID" field is not present but the Bit #0 of the "Probing Flag" field is set to "1", the ACK message SHOULD be sent over the same delivery connection as the Probe message.

The "Padding" field is used to control the length of Probe message.

5.3 Packet Loss Report (PLR) Message

The "Type" field is set to "2" for PLR messages.

C-MADP may send out the PLR messages to report lost MX SDU for example during handover. In response, C-MADP may retransmit the lost MX SDU accordingly.

A PLR message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection which the ACK message is for;

o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the anchor connection which the ACK message is for;
o ACK number (4 Bytes): the next (in-order) sequence number (SN) that the sender of the PLR message is expecting
o Number of Loss Bursts (1 Byte)
  For each loss burst, include the following
    + Sequence Number of the first lost MX SDU in a burst (4 Bytes)
    + Number of consecutive lost MX SDUs in the burst (1 Byte)

```
          C-MADP                                              N-MADP
              |                                                  |
              |<------------------- MX SDU (data packets)--------|
              |                                                  |
             +-----------------------------------+               |
             |Packet Loss detected               |               |
             +-----------------------------------+               |
              |                                                  |
              |----- PLR Message ------------------------------->|
              |<-------------retransmit(lost)MX SDUs ------------|
```

Figure 8: MAMS Retransmission Procedure

Figure 8 shows the MAMS retransmission procedure in an example where the lost packet is found and retransmitted.

5.4 First Sequence Number (FSN) Message

The "Type" field is set to "3" for FSN messages.

N-MADP may send out the FSN messages to indicate the oldest MX SDU in its buffer if a lost MX SDU is not found in the buffer after receiving the PLR message from C-MADP. In response, C-MADP SHALL only report packet loss with SN not smaller than FSN.

A FSN message consists of the following fields:

o Connection ID (1 Byte): an unsigned integer to identify the anchor connection which the FSN message is for;
o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the anchor connection which the FSN message is for;
o First Sequence Number (4 Bytes): the sequence number (SN) of the oldest MX SDU in the (retransmission) buffer of the sender of the FSN message.

Figure 9 shows the MAMS retransmission procedure in an example where the lost packet is not found.

```
          C-MADP                                              N-MADP
              |                                                  |
              |<------------------ MX SDU (data packets)--------|
              |                                                  |
            +----------------------------------+                 |
            |Packet Loss detected              |                 |
            +----------------------------------+                 |
              |                                                  |
              |----- PLR Message ------------------------------>|
              |                              +---------------------+
              |                              |Lost packet not found|
              |                              +---------------------+
              |<-------------FSN message -----------------------|
```

Figure 9: MAMS Retransmission Procedure with FSN

5.5 Coded MX SDU (CMS) Message

The "Type" field is set to "4" for CMS messages.

N-MADP (or C-MADP) may send out the CMS message to support downlink (or uplink) packet loss recovery through coding, e.g. [CRLNC], [CTCP], [RLNC]. A coded MX SDU is generated by applying a network coding algorithm to multiple consecutive (uncoded) MX SDUs, and it is used for fast recovery without retransmission if any of the MX SDUs is lost.

A Coded MX SDU message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection of the coded MX SDU;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the coded MX;
- o Sequence Number (4 Bytes): the sequence number of the first (uncoded) MX SDU used to generate the coded MX SDU.
- o Fragmentation Control (FC) (1 Byte): to provide necessary information for re-assembly, only needed if the coded MX SDU is too long to transport in a single MX control PDU.
- o N (1 Byte): the number of consecutive MX SDUs used to generate the coded MX SDU
- o K (1 Byte): the length (in terms of bits) of the coding coefficient field
- o Coding Coefficient (N x K / 8 Bytes)
  - + a(i): the coding coefficient of the i-th (uncoded) MX SDU

      + padding
  o Coded MX SDU (variable): the coded MX SDU

If K = 0, the simple XOR method is used to generate the Coded MX SDU from N consecutive uncoded MX SDUs, and the a(i) fields are not included in the message.

If the coded MX SDU is too long, it can be fragmented, and transported by multiple MX control PDUs. The N, K, and a(i) fields are only included in the MX PDU carrying the first fragment of the coded MX SDU.

```
          C-MADP                                             N-MADP
              |                                                  |
              |<------------------ MX SDU #1 -------------------|
              |      lost<-------- MX SDU #2 -------------------|
              |<---- CMS Message (MX SDU #1 XOR MX SDU #2)------|
             +----------------------+                            |
             | MX SDU #2 recovered  |                            |
             +----------------------+                            |
              |                                                  |
```

Figure 10: MAMS Packet Recovery Procedure with XOR Coding

5.6 Traffic Splitting Update (TSU) Message

The "Type" field is set to "5" for TSU messages.

N-MADP (or C-MADP) may send out a TSU message if downlink (or uplink) traffic splitting configuration has changed.

A TSU message consists of the following fields:

  o Connection ID (1 Byte): an unsigned integer to identify the anchor connection;
  o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class;
  o Sequence Number (2 Bytes): an unsigned integer to identify the TSU message.
  o Flags (1 Byte)
      + Bit #0: a Reverse Path bit flag to indicate if the traffic splitting configuration is for the reverse path (1) or not (0);
      + Bit #1: a Bit-Reversal bit flag to indicate if bit-reversal is used in traffic splitting
      + Others: reserved.
  o Traffic Splitting Configuration Parameters (5 + (N -1) Bytes):

+ StartSN (4 Bytes): the sequence number of the first MX SDU using the traffic splitting configuration provided by the TSU message
+ L (1 Byte): the traffic splitting burst size
+ K(i): the traffic splitting threshold of the i-th delivery connection, where connections are ordered according to their Connection ID.

Let’s use f(x) to denote the traffic splitting function, which maps a MX SDU Sequence Number "x" to the i-th delivery connection.

f(x)=i, if K[i-1]< or = mod(x - StartSN, L) < K[i]

Wherein, 1 < or = i < N, K[0]=0, and K[N]=L.

N is the total number of connections for delivering a data flow, identified by (anchor) Connection ID and Traffic Class ID.

When the bit-reversal bit is set to 1, the burst size L MUST be a power of 2, and the traffic splitting function is

f(x)=i, if K[i-1]< or = F(mod(x - StartSN, L)) < K[i]

Wherein F(.) is the bit reversal function [BITR] of the input variable.

5.7 Acknowledgement Message

The "Type" field is set to "6" for ACK messages.

C-MADP (or N-MADP) SHOULD send out the ACK message in response to the successful reception of a PLR, FSN, or TSU message.

C-MADP SHOULD send out the ACK message in response to a Probe message with the ACK flag set to "1".

The ACK message consists of the following fields:

- o Acknowledgment Number (2 Bytes): the sequence number of the received message.
- o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.

6 Security Considerations

User data in MAMS framework rely on the security of the underlying network transport paths. When this cannot be assumed, NCM configures use of appropriate protocols for security, e.g. IPsec [RFC4301] [RFC3948], DTLS [RFC6347].

7 IANA Considerations

This draft makes no requests of IANA.

8 Contributing Authors

The editors gratefully acknowledge the following additional contributors in alphabetical order: Salil Agarwal/Nokia, Hema Pentakota/Nokia.

9 References

9.1 Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC4301] Kent, S. and K. Seo, "Security Architecture for the Internet Protocol", RFC 4301, DOI10.17487/RFC4301, December 2005, <http://www.rfc-editor.org/info/rfc4301>.

9.2 Informative References

[RFC6347] Rescorla, E. and N. Modadugu, "Datagram Transport Layer Security Version 1.2", RFC 6347, January 2012, <http://www.rfc-editor.org/info/rfc6347>.

[RFC7296] Kaufman, C., Hoffman, P., Nir, Y., Eronen, P., and T. Kivinen, "Internet Key Exchange Protocol Version 2 (IKEv2)", STD 79, RFC 7296, DOI 10.17487/RFC7296, October 2014, <http://www.rfc-editor.org/info/rfc7296>.

[RFC3948] Huttunen, A., Swander, B., Volpe, V., DiBurro, L., and M. Stenberg, "UDP Encapsulation of IPsec ESP Packets", RFC 3948, DOI 10.17487/RFC3948, January 2005, <http://www.rfc-editor.org/info/rfc3948>.

[MPProxy] X. Wei, C. Xiong, and E. Lopez, "MPTCP proxy mechanisms", https://tools.ietf.org/html/draft-wei-mptcp-proxy-mechanism-02

[MPPlain] M. Boucadair et al, "An MPTCP Option for Network-Assisted MPTCP", https://www.ietf.org/id/draft-boucadair-mptcp-plain-mode-09.txt

[MAMS] S. Kanugovi, S. Vasudevan, F. Baboescu, and J. Zhu, "Multiple Access Management Protocol", https://tools.ietf.org/html/draft-kanugovi-intarea-mams-protocol-03

[GMA] J. Zhu, "Trailer-based Encapsulation Protocols for Generic Multi-Access Convergence", https://tools.ietf.org/html/draft-zhu-intarea-gma-01

[GRE2784] D. Farinacci, et al., "Generic Routing Encapsulation (GRE)", RFC 2784 March 2000, <http://www.rfc-editor.org/info/rfc2784>.

[GRE2890] G. Dommety, "Key and Sequence Number Extensions to GRE", RFC 2890 September 2000, <http://www.rfc-editor.org/info/rfc2890>.

[IANA] https://www.iana.org/assignments/protocol-numbers/protocol-numbers.xhtml

[LWIPEP] 3GPP TS 36.361, "Evolved Universal Terrestrial Radio Access (E-UTRA); LTE-WLAN Radio Level Integration Using Ipsec Tunnel (LWIP) encapsulation; Protocol specification"

[RFC791] Internet Protocol, September 1981

[CRLNC] S Wunderlich, F Gabriel, S Pandi, et al. Caterpillar RLNC (CRLNC): A Practical Finite Sliding Window RLNC Approach, IEEE Access, 2017

[CTCP] M. Kim, et al. Network Coded TCP (CTCP), eprint arXiv:1212.2291, 2012

[RLNC] J. Heide, et al. Random Linear Network Coding (RLNC)-Based Symbol Representation, https://www.ietf.org/id/draft-heide-nwcrg-rlnc-00.txt

[BITR] Alan H. Karp, "Bit reversal on uniprocessors", SIAM Review, 38 (1): 1-26, 1996.

Authors' Addresses

Jing Zhu

Intel

Email: jing.z.zhu@intel.com

SungHoon Seo

Korea Telecom

Email: sh.seo@kt.com

Satish Kanugovi

Nokia

Email: satish.k@nokia.com

Shuping Peng

Huawei

Email: pengshuping@huawei.com

INTAREA   J. Zhu
Internet Draft   Intel
Intended status: Standards Track   S. Seo
Expires: April 1,2020   Korea Telecom
                                                            S. Kanugovi
                                                                  Nokia
                                                                S. Peng
                                                                 Huawei
                                                        October 1, 2019

User-Plane Protocols for Multiple Access Management Service
draft-zhu-intarea-mams-user-protocol-07

Status of this Memo

Copyright Notice

Abstract

Today, a device can be simultaneously connected to multiple communication networks based on different technology implementations and network architectures like WiFi, LTE, and DSL. In such multi-connectivity scenario, it is desirable to combine multiple access networks or select the best one to improve quality of experience for a user and improve overall network utilization and efficiency. This document presents the u-plane protocols for a multi access management services (MAMS) framework that can be used to flexibly select the combination of uplink and downlink access and core network paths having the optimal performance, and user plane treatment for improving network utilization and efficiency and enhanced quality of experience for user applications.

Table of Contents

1. Introduction

Multi Access Management Service (MAMS) [MAMS] is a programmable framework to select and configure network paths, as well as adapt to dynamic network conditions, when multiple network connections can serve a client device. It is based on principles of user plane interworking that enables the solution to be deployed as an overlay without impacting the underlying networks.

This document presents the u-plane protocols for enabling the MAMS framework. It co-exists and complements the existing protocols by providing a way to negotiate and configure the protocols based on client and network capabilities. Further it allows exchange of network state information and leveraging network intelligence to optimize the performance of such protocols. An important goal for MAMS is to ensure that there is minimal or no dependency on the actual access technology of the participating links. This allows the scheme to be scalable for addition of newer access technologies and for independent evolution of the existing access technologies.

2. Terminologies

Anchor Connection: refers to the network path from the N-MADP to the Application Server that corresponds to a specific IP anchor that has assigned an IP address to the client.

Delivery Connection: refers to the network path from the N-MADP to the C-MADP.

"Network Connection Manager" (NCM), "Client Connection Manager" (CCM), "Network Multi Access Data Proxy" (N-MADP), and "Client Multi Access Data Proxy" (C-MADP) in this document are to be interpreted as described in [MAMS].

3. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

The terminologies "Network Connection Manager" (NCM), "Client Connection Manager" (CCM), "Network Multi Access Data Proxy" (N-MADP), and "Client Multi Access Data Proxy" (C-MADP) in this document are to be interpreted as described in [MAMS].

4 MAMS User-Plane Protocols

Figure 1 shows the MAMS u-plane protocol stack as specified in [MAMS].

```
             +-----------------------------------------------------+
             |      User Payload (e.g. IP PDU)                     |
             |-----------------------------------------------------|
          +--|-----------------------------------------------------|--+
          |  |-----------------------------------------------------|  |
          |  | Multi-Access (MX) Convergence Sublayer              |  |
          |  |-----------------------------------------------------|  |
          |  |-----------------------------------------------------|  |
          |  | MX Adaptation   | MX Adaptation | MX Adaptation     |  |
          |  | Sublayer        | Sublayer      | Sublayer          |  |
          |  | (optional)      | (optional)    | (optional)        |  |
          |  |-----------------------------------------------------|  |
          |  | Access #1 IP    | Access #2 IP  | Access #3 IP      |  |
          |  +-----------------------------------------------------+  |
          +-----------------------------------------------------------+
```

Figure 1: MAMS U-plane Protocol Stack

It consists of the following two Sublayers:

o Multi-Access (MX) Convergence Sublayer: This layer performs multi-access specific tasks, e.g., access (path) selection, multi-link (path) aggregation, splitting/reordering, lossless switching, fragmentation, concatenation, keep-alive, and probing etc.
o Multi-Access (MX) Adaptation Sublayer: This layer performs functions to handle tunneling, network layer security, and NAT.

The MX convergence sublayer operates on top of the MX adaptation sublayer in the protocol stack. From the Transmitter perspective, a User Payload (e.g. IP PDU) is processed by the convergence sublayer first, and then by the adaptation sublayer before being transported over a delivery access connection; from the Receiver perspective, an IP packet received over a delivery connection is processed by the MX adaptation sublayer first, and then by the MX convergence sublayer.

4.1 MX Adaptation Sublayer

The MX adaptation sublayer supports the following mechanisms and protocols while transmitting user plane packets on the network path:

o UDP Tunneling: The user plane packets of the anchor connection can be encapsulated in a UDP tunnel of a delivery connection between the N-MADP and C-MADP.

o IPsec Tunneling: The user plane packets of the anchor connection are sent through an IPsec tunnel of a delivery connection.
o Client Net Address Translation (NAT): The Client IP address of user plane packet of the anchor connection is changed, and sent over a delivery connection.
o Pass Through: The user plane packets are passing through without any change over the anchor connection.

The MX adaptation sublayer also supports the following mechanisms and protocols to ensure security of user plane packets over the network path.

o IPsec Tunneling: An IPsec [RFC7296] tunnel is established between the N-MADP and C-MADP on the network path that is considered untrusted.
o DTLS: If UDP tunneling is used on the network path that is considered "untrusted", DTLS (Datagram Transport Layer Security) [RFC6347] can be used.

The Client NAT method is the most efficient due to no tunneling overhead. It SHOULD be used if a delivery connection is "trusted" and without NAT function on the path.

The UDP or IPsec Tunnelling method SHOULD be used if a delivery connection has a NAT function placed on the path.

4.2 GMA-based MX Convergence Sublayer

Figure 2 shows the MAMS u-plane protocol stack based on trailer-based encapsulation [GMA]. Multiple access networks are combined into a single IP connection. If NCM determines that N-MADP is to be instantiated with GMA as the MX Convergence Protocol, it exchanges the support of GMA convergence capability in the discovery and capability exchange procedures [MAMS].

```
          +-----------------------------------------------------+
          |                        IP PDU                       |
          |-----------------------------------------------------|
          |                GMA  Convergence Sublayer            |
          |-----------------------------------------------------|
          | MX Adaptation    | MX Adaptation | MX Adaptation    |
          | Sublayer         | Sublayer      | Sublayer         |
          | (optional)       | (optional)    | (optional)       |
          |-----------------------------------------------------|
          | Access #1 IP     | Access #2 IP  | Access #3 IP     |
          +-----------------------------------------------------+
```
Figure 2: MAMS U-plane Protocol Stack with GMA as MX Convergence Layer

Figure 3 shows the trailer-based Multi-Access (MX) PDU (Protocol Data Unit) format [GMA]. If the MX adaptation method is UDP tunneling and "MX header optimization" in the "MX_UP_Setup_Configuration_Request" message [MAMS] is true, the "IP length" and "IP checksum" header fields of the MX PDU SHOULD remain unchanged. Otherwise, they should be updated after adding or removing the GMA trailer in the convergence sublayer.

```
          +-------------------------------------------------------+
          | IP hdr |         IP payload               | GMA Trailer |
          +-------------------------------------------------------+
                         Figure 3: GMA PDU Format
```

4.3 MPTCP-based MX Convergence Sublayer

Figure 4 shows the MAMS u-plane protocol stack based on MPTCP. Here, MPTCP is reused as the "MX Convergence Sublayer" protocol. Multiple access networks are combined into a single MPTCP connection. Hence, no new u-plane protocol or PDU format is needed in this case.

```
          |------------------------------------------------------|
          |                        MPTCP                         |
          |------------------------------------------------------|
          |  TCP             |   TCP           |      TCP        |
          |------------------------------------------------------|
          | MX Adaptation    | MX Adaptation | MX Adaptation     |
          | Sublayer         | Sublayer      | Sublayer          |
          | (optional)       | (optional)    | (optional)        |
          |------------------------------------------------------|
          | Access #1 IP     | Access #2 IP  | Access #3 IP      |
          +------------------------------------------------------+
```

Figure 4: MAMS U-plane Protocol Stack with MPTCP as MX Convergence Layer

If NCM determines that N-MADP is to be instantiated with MPTCP as the MX Convergence Protocol, it exchanges the support of MPTCP capability in the discovery and capability exchange procedures [MAMS]. MPTCP proxy protocols [MPProxy][MPPlain] SHOULD be used to manage traffic steering and aggregation over multiple delivery connections.

4.4 GRE as MX Convergence Sublayer

Figure 5 shows the MAMS u-plane protocol stack based on GRE (Generic Routing Encapsulation) [GRE2784]. Here, GRE is reused as the "MX Convergence sub-layer" protocol. Multiple access networks are combined

into a single GRE connection. Hence, no new u-plane protocol or PDU format is needed in this case.

```
+-----------------------------------------------------+
|      User Payload (e.g. IP PDU)                     |
|-----------------------------------------------------|
|              GRE as MX Convergence Sublayer         |
|-----------------------------------------------------|
|        GRE Delivery Protocol (e.g. IP)              |
|-----------------------------------------------------|
| MX Adaptation    | MX Adaptation | MX Adaptation    |
| Sublayer         | Sublayer      | Sublayer         |
| (optional)       | (optional)    | (optional)       |
|-----------------------------------------------------|
| Access #1 IP     | Access #2 IP  | Access #3 IP     |
+-----------------------------------------------------+
```

Figure 5: MAMS U-plane Protocol Stack with GRE as MX Convergence Layer

If NCM determines that N-MADP is to be instantiated with GRE as the MX Convergence Protocol, it exchanges the support of GRE capability in the discovery and capability exchange procedures [MAMS].

4.4.1 Transmitter Procedures

Transmitter is the N-MADP or C-MADP instance, instantiated with GRE as the convergence protocol that transmits the GRE packets. The Transmitter receives the User Payload (e.g. IP PDU), encapsulates it with a GRE header and Delivery Protocol (e.g. IP) header to generate the GRE Convergence PDU.

When IP is used as the GRE delivery protocol, the IP header information (e.g. IP address) can be created using the IP header of the user payload or a virtual IP address. The "Protocol Type" field of the delivery header is set to 47 (or 0X2F, i.e. GRE)[IANA].

The GRE header fields are set as specified below,

- If the transmitter is a C-MADP instance, then sets the LSB 16 bits to the value of Connection ID for the Anchor Connection associated with the user payload or sets to 0xFFFF if no Anchor Connection ID needs to be specified.
- All other fields in the GRE header including the remaining bits in the key fields are set as per [GRE_2784][GRE_2890].

4.4.2 Receiver Procedures

Receiver is the N-MADP or C-MADP instance, instantiated with GRE as the convergence protocol that receives the GRE packets. The receiver processes the received packets per the GRE procedures [GRE_2784, GRE_2890] and retrieves the GRE header.

- If the Receiver is an N-MADP instance,
  - o Unless the LSB 16 Bits of the Key field are 0xFFFF, they are interpreted as the Connection ID of Anchor Connection for the user payload. This is used to identify the network path over which the User Payload (GRE Payload) is to be transmitted.
- All other fields in the GRE header, including the remaining bits in the Key fields, are processed as per [GRE_2784][GRE_2890].

The GRE Convergence PDU is passed onto the MX Adaptation Layer (if present) before delivery over one of the network paths.

4.5 Co-existence of MX Adaptation and MX Convergence Sublayers

MAMS u-plane protocols support multiple combinations and instances of user plane protocols to be used in the MX Adaptation and the Convergence sublayers.

For example, one instance of the MX Convergence Layer can be MPTCP Proxy [MPProxy][MPPlain] and another instance can be Trailer-based. The MX Adaptation for each can be either UDP tunnel or IPsec. IPsec may be set up for network paths considered as untrusted by the operator, to protect the TCP subflow between client and MPTCP proxy traversing that network path.

Each of the instances of MAMS user plane, i.e. combination of MX Convergence and MX Adaptation layer protocols, can coexist simultaneously and independently handle different traffic types.

5. MX Convergence Control Message

A UDP connection may be configured between C-MADP and N-MADP to exchange control messages for keep-alive or path quality estimation. The N-MADP end-point IP address and UDP port number of the UDP connection is used to identify MX control PDU. Figure 6 shows the MX control PDU format with the following fields:

- o Type (1 Byte): the type of the MX control message
- o CID (1 Byte): an unsigned integer to identify the anchor and delivery connection of the MX control message
  - + Anchor Connection ID (MSB 4 Bits): an unsigned integer to identify the anchor connection

+ Delivery Connection ID (LSB 4 Bits): an unsigned integer to identify the delivery connection

o MX Control Message (variable): the payload of the MX control message

Figure 7 shows the MX convergence control protocol stack, and MX control PDU goes through the MX adaptation sublayer the same way as MX data PDU.

```
                        <----MX Control PDU Payload --------------->
+--------------------------------------------------------------------+
| IP header | UDP Header| Type | CID |       MX Control Message     |
+--------------------------------------------------------------------+
                      Figure 6: MX Control PDU Format

          |------------------------------------------------------|
          |          MX Convergence Control Messages              |
          |------------------------------------------------------|
          |                  UDP/IP                               |
          |------------------------------------------------------|
          | MX Adaptation   | MX Adaptation | MX Adaptation        |
          | Sublayer        | Sublayer      | Sublayer             |
          | (optional)      | (optional)    | (optional)           |
          |------------------------------------------------------|
          | Access #1 IP    | Access #2 IP  | Access #3 IP         |
          +------------------------------------------------------+
              Figure 7: MX Convergence Control Protocol Stack
```

5.1 Keep-Alive Message

The "Type" field is set to "0" for Keep-Alive messages. C-MADP may send out Keep-Alive message periodically over one or multiple delivery connections, especially if UDP tunneling is used as the adaptation method for the delivery connection with a NAT function on the path.

A Keep-Alive message is 6 Bytes long, and consists of the following fields:

o Keep-Alive Sequence Number (2 Bytes): the sequence number of the keep-alive message

o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.

5.2 Probe Message

The "Type" field is set to "1" for Probe messages.

N-MADP may send out the Probe message for path quality estimation. In response, C-MADP may send back the ACK message.

A Probe message consists of the following fields:

- o Probing Sequence Number (2 Bytes): the sequence number of the Probe REQ message
- o Probing Flag (1 Byte):
  - + Bit #0: a ACK flag to indicate if the ACK message is expected (1) or not (0);
  - + Bit #1: a Probe Type flag to indicate if the Probe message is sent during the initialization phase (0) when the network path is not included for transmission of user data or the active phase (1) when the network path is included for transmission of user data;
  - + Bit #2: a bit flag to indicate the presence of the Reverse Connection ID (R-CID) field.
  - + Bit #3~7: reserved
- o Reverse Connection ID (1 Byte): the connection ID of the delivery connection for sending out the ACK message on the reverse path
- o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.
- o Padding (variable)

The "R-CID" field is only present if both Bit #0 and Bit #2 of the "Probing Flag" field are set to "1". Moreover, Bit #2 of the "Probing Flag" field SHOULD be set to "0" if the Bit #0 is "0", indicating the ACK message is not expected.

If the "R-CID" field is not present but the Bit #0 of the "Probing Flag" field is set to "1", the ACK message SHOULD be sent over the same delivery connection as the Probe message.

The "Padding" field is used to control the length of Probe message.

5.3 Packet Loss Report (PLR) Message

The "Type" field is set to "2" for PLR messages.

C-MADP may send out the PLR messages to report lost MX SDU for example during handover. In response, C-MADP may retransmit the lost MX SDU accordingly.

A PLR message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection which the ACK message is for;

- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the anchor connection which the ACK message is for;
- o ACK number (4 Bytes): the next (in-order) sequence number (SN) that the sender of the PLR message is expecting
- o Number of Loss Bursts (1 Byte)
  For each loss burst, include the following
  - + Sequence Number of the first lost MX SDU in a burst (4 Bytes)
  - + Number of consecutive lost MX SDUs in the burst (1 Byte)

```
          C-MADP                                             N-MADP
              |                                                 |
              |<------------------ MX SDU (data packets)--------|
              |                                                 |
             +----------------------------------+               |
             |Packet Loss detected              |               |
             +----------------------------------+               |
              |                                                 |
              |----- PLR Message ------------------------------>|
              |<-------------retransmit(lost)MX SDUs -----------|
```

Figure 8: MAMS Retransmission Procedure

Figure 8 shows the MAMS retransmission procedure in an example where the lost packet is found and retransmitted.

5.4 First Sequence Number (FSN) Message

The "Type" field is set to "3" for FSN messages.

N-MADP may send out the FSN messages to indicate the oldest MX SDU in its buffer if a lost MX SDU is not found in the buffer after receiving the PLR message from C-MADP. In response, C-MADP SHALL only report packet loss with SN not smaller than FSN.

A FSN message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection which the FSN message is for;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the anchor connection which the FSN message is for;
- o First Sequence Number (4 Bytes): the sequence number (SN) of the oldest MX SDU in the (retransmission) buffer of the sender of the FSN message.

Figure 9 shows the MAMS retransmission procedure in an example where the lost packet is not found.

```
          C-MADP                                              N-MADP
              |                                                  |
              |<------------------ MX SDU (data packets)--------|
              |                                                  |
             +----------------------------------+                |
             |Packet Loss detected              |                |
             +----------------------------------+                |
              |                                                  |
              |----- PLR Message ------------------------------>|
              |                               +---------------------+
              |                               |Lost packet not found|
              |                               +---------------------+
              |<-------------FSN message -----------------------|
```

Figure 9: MAMS Retransmission Procedure with FSN

5.5 Coded MX SDU (CMS) Message

The "Type" field is set to "4" for CMS messages.

N-MADP (or C-MADP) may send out the CMS message to support downlink (or uplink) packet loss recovery through coding, e.g. [CRLNC], [CTCP], [RLNC]. A coded MX SDU is generated by applying a network coding algorithm to multiple consecutive (uncoded) MX SDUs, and it is used for fast recovery without retransmission if any of the MX SDUs is lost.

A Coded MX SDU message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection of the coded MX SDU;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the coded MX;
- o Sequence Number (4 Bytes): the sequence number of the first (uncoded) MX SDU used to generate the coded MX SDU.
- o Fragmentation Control (FC) (1 Byte): to provide necessary information for re-assembly, only needed if the coded MX SDU is too long to transport in a single MX control PDU.
- o N (1 Byte): the number of consecutive MX SDUs used to generate the coded MX SDU
- o K (1 Byte): the length (in terms of bits) of the coding coefficient field
- o Coding Coefficient (N x K / 8 Bytes)
  - + a(i): the coding coefficient of the i-th (uncoded) MX SDU

    + padding
 o Coded MX SDU (variable): the coded MX SDU

If K = 0, the simple XOR method is used to generate the Coded MX SDU from N consecutive uncoded MX SDUs, and the a(i) fields are not included in the message.

If the coded MX SDU is too long, it can be fragmented, and transported by multiple MX control PDUs. The N, K, and a(i) fields are only included in the MX PDU carrying the first fragment of the coded MX SDU.

```
          C-MADP                                             N-MADP
              |                                                 |
              |<------------------ MX SDU #1 -------------------|
              |      lost<-------- MX SDU #2 -------------------|
              |<---- CMS Message (MX SDU #1 XOR MX SDU #2)------|
             +----------------------+                           |
             | MX SDU #2 recovered  |                           |
             +----------------------+                           |
              |                                                 |
```

Figure 10: MAMS Packet Recovery Procedure with XOR Coding

5.6 Traffic Splitting Update (TSU) Message

The "Type" field is set to "5" for TSU messages.

N-MADP (or C-MADP) may send out a TSU message if downlink (or uplink) traffic splitting configuration has changed.

A TSU message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class;
- o Sequence Number (2 Bytes): an unsigned integer to identify the TSU message.
- o Flags (1 Byte)
    - + Bit #0: a Reverse Path bit flag to indicate if the traffic splitting configuration is for the reverse path (1) or not (0);
    - + Bit #1: a Bit-Reversal bit flag to indicate if bit-reversal is used in traffic splitting
    - + Others: reserved.
- o Traffic Splitting Configuration Parameters (5 + (N -1) Bytes):

+ StartSN (4 Bytes): the sequence number of the first MX SDU using the traffic splitting configuration provided by the TSU message
+ L (1 Byte): the traffic splitting burst size
+ K(i): the traffic splitting threshold of the i-th delivery connection, where connections are ordered according to their Connection ID.

Let's use f(x) to denote the traffic splitting function, which maps a MX SDU Sequence Number "x" to the i-th delivery connection.

f(x)=i, if K[i-1]< or = mod(x - StartSN, L) < K[i]

Wherein, 1 < or = i < N, K[0]=0, and K[N]=L.

N is the total number of connections for delivering a data flow, identified by (anchor) Connection ID and Traffic Class ID.

When the bit-reversal bit is set to 1, the burst size L MUST be a power of 2, and the traffic splitting function is

f(x)=i, if K[i-1]< or = F(mod(x - StartSN, L)) < K[i]

Wherein F(.) is the bit reversal function [BITR] of the input variable.

5.7 Acknowledgement Message

The "Type" field is set to "6" for ACK messages.

C-MADP (or N-MADP) SHOULD send out the ACK message in response to the successful reception of a PLR, FSN, or TSU message.

C-MADP SHOULD send out the ACK message in response to a Probe message with the ACK flag set to "1".

The ACK message consists of the following fields:

- o Acknowledgment Number (2 Bytes): the sequence number of the received message.
- o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.

6 Security Considerations

User data in MAMS framework rely on the security of the underlying network transport paths. When this cannot be assumed, NCM configures use of appropriate protocols for security, e.g. IPsec [RFC4301] [RFC3948], DTLS [RFC6347].

7 IANA Considerations

This draft makes no requests of IANA.

8 Contributing Authors

The editors gratefully acknowledge the following additional contributors in alphabetical order: Salil Agarwal/Nokia, Hema Pentakota/Nokia.

9 References

9.1 Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC4301] Kent, S. and K. Seo, "Security Architecture for the Internet Protocol", RFC 4301, DOI10.17487/RFC4301, December 2005, <http://www.rfc-editor.org/info/rfc4301>.

9.2 Informative References

[RFC6347] Rescorla, E. and N. Modadugu, "Datagram Transport Layer Security Version 1.2", RFC 6347, January 2012, <http://www.rfc-editor.org/info/rfc6347>.

[RFC7296] Kaufman, C., Hoffman, P., Nir, Y., Eronen, P., and T. Kivinen, "Internet Key Exchange Protocol Version 2 (IKEv2)", STD 79, RFC 7296, DOI 10.17487/RFC7296, October 2014, <http://www.rfc-editor.org/info/rfc7296>.

[RFC3948] Huttunen, A., Swander, B., Volpe, V., DiBurro, L., and M. Stenberg, "UDP Encapsulation of IPsec ESP Packets", RFC 3948, DOI 10.17487/RFC3948, January 2005, <http://www.rfc-editor.org/info/rfc3948>.

[MPProxy] X. Wei, C. Xiong, and E. Lopez, "MPTCP proxy mechanisms", https://tools.ietf.org/html/draft-wei-mptcp-proxy-mechanism-02

[MPPlain] M. Boucadair et al, "An MPTCP Option for Network-Assisted MPTCP", https://www.ietf.org/id/draft-boucadair-mptcp-plain-mode-09.txt

[MAMS] S. Kanugovi, S. Vasudevan, F. Baboescu, and J. Zhu, "Multiple Access Management Protocol", https://tools.ietf.org/html/draft-kanugovi-intarea-mams-protocol-03

[GMA] J. Zhu, "Trailer-based Encapsulation Protocols for Generic Multi-Access Convergence", https://tools.ietf.org/html/draft-zhu-intarea-gma-01

[GRE2784] D. Farinacci, et al., "Generic Routing Encapsulation (GRE)", RFC 2784 March 2000, <http://www.rfc-editor.org/info/rfc2784>.

[GRE2890] G. Dommety, "Key and Sequence Number Extensions to GRE", RFC 2890 September 2000, <http://www.rfc-editor.org/info/rfc2890>.

[IANA] https://www.iana.org/assignments/protocol-numbers/protocol-numbers.xhtml

[LWIPEP] 3GPP TS 36.361, "Evolved Universal Terrestrial Radio Access (E-UTRA); LTE-WLAN Radio Level Integration Using Ipsec Tunnel (LWIP) encapsulation; Protocol specification"

[RFC791] Internet Protocol, September 1981

[CRLNC] S Wunderlich, F Gabriel, S Pandi, et al. Caterpillar RLNC (CRLNC): A Practical Finite Sliding Window RLNC Approach, IEEE Access, 2017

[CTCP] M. Kim, et al. Network Coded TCP (CTCP), eprint arXiv:1212.2291, 2012

[RLNC] J. Heide, et al. Random Linear Network Coding (RLNC)-Based Symbol Representation, https://www.ietf.org/id/draft-heide-nwcrg-rlnc-00.txt

[BITR] Alan H. Karp, "Bit reversal on uniprocessors", SIAM Review, 38 (1): 1-26, 1996.

Authors' Addresses

Jing Zhu

Intel

Email: jing.z.zhu@intel.com

SungHoon Seo

Korea Telecom

Email: sh.seo@kt.com

Satish Kanugovi

Nokia

Email: satish.k@nokia.com

Shuping Peng

Huawei

Email: pengshuping@huawei.com

INTAREA   J. Zhu
Internet Draft   Intel
Intended status: Standards Track   S. Seo
Expires: April 1,2020   Korea Telecom
                                                            S. Kanugovi
                                                                  Nokia
                                                                S. Peng
                                                                 Huawei
                                                        October 1, 2019

User-Plane Protocols for Multiple Access Management Service
draft-zhu-intarea-mams-user-protocol-07

Abstract

Today, a device can be simultaneously connected to multiple communication networks based on different technology implementations and network architectures like WiFi, LTE, and DSL. In such multi-connectivity scenario, it is desirable to combine multiple access networks or select the best one to improve quality of experience for a user and improve overall network utilization and efficiency. This document presents the u-plane protocols for a multi access management services (MAMS) framework that can be used to flexibly select the combination of uplink and downlink access and core network paths having the optimal performance, and user plane treatment for improving network utilization and efficiency and enhanced quality of experience for user applications.

Table of Contents

1. Introduction

Multi Access Management Service (MAMS) [MAMS] is a programmable framework to select and configure network paths, as well as adapt to dynamic network conditions, when multiple network connections can serve a client device. It is based on principles of user plane interworking that enables the solution to be deployed as an overlay without impacting the underlying networks.

This document presents the u-plane protocols for enabling the MAMS framework. It co-exists and complements the existing protocols by providing a way to negotiate and configure the protocols based on client and network capabilities. Further it allows exchange of network state information and leveraging network intelligence to optimize the performance of such protocols. An important goal for MAMS is to ensure that there is minimal or no dependency on the actual access technology of the participating links. This allows the scheme to be scalable for addition of newer access technologies and for independent evolution of the existing access technologies.

2. Terminologies

Anchor Connection: refers to the network path from the N-MADP to the Application Server that corresponds to a specific IP anchor that has assigned an IP address to the client.

Delivery Connection: refers to the network path from the N-MADP to the C-MADP.

"Network Connection Manager" (NCM), "Client Connection Manager" (CCM), "Network Multi Access Data Proxy" (N-MADP), and "Client Multi Access Data Proxy" (C-MADP) in this document are to be interpreted as described in [MAMS].

3. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

The terminologies "Network Connection Manager" (NCM), "Client Connection Manager" (CCM), "Network Multi Access Data Proxy" (N-MADP), and "Client Multi Access Data Proxy" (C-MADP) in this document are to be interpreted as described in [MAMS].

4 MAMS User-Plane Protocols

Figure 1 shows the MAMS u-plane protocol stack as specified in [MAMS].

```
             +-----------------------------------------------------+
             |      User Payload (e.g. IP PDU)                     |
             |-----------------------------------------------------|
          +--|-----------------------------------------------------|--+
          |  |-----------------------------------------------------|  |
          |  | Multi-Access (MX) Convergence Sublayer              |  |
          |  |-----------------------------------------------------|  |
          |  |-----------------------------------------------------|  |
          |  | MX Adaptation   | MX Adaptation | MX Adaptation     |  |
          |  | Sublayer        | Sublayer      | Sublayer          |  |
          |  | (optional)      | (optional)    | (optional)        |  |
          |  |-----------------------------------------------------|  |
          |  | Access #1 IP    | Access #2 IP  | Access #3 IP      |  |
          |  +-----------------------------------------------------+  |
          +-----------------------------------------------------------+
```

Figure 1: MAMS U-plane Protocol Stack

It consists of the following two Sublayers:

o Multi-Access (MX) Convergence Sublayer: This layer performs multi-access specific tasks, e.g., access (path) selection, multi-link (path) aggregation, splitting/reordering, lossless switching, fragmentation, concatenation, keep-alive, and probing etc.
o Multi-Access (MX) Adaptation Sublayer: This layer performs functions to handle tunneling, network layer security, and NAT.

The MX convergence sublayer operates on top of the MX adaptation sublayer in the protocol stack. From the Transmitter perspective, a User Payload (e.g. IP PDU) is processed by the convergence sublayer first, and then by the adaptation sublayer before being transported over a delivery access connection; from the Receiver perspective, an IP packet received over a delivery connection is processed by the MX adaptation sublayer first, and then by the MX convergence sublayer.

4.1 MX Adaptation Sublayer

The MX adaptation sublayer supports the following mechanisms and protocols while transmitting user plane packets on the network path:

o UDP Tunneling: The user plane packets of the anchor connection can be encapsulated in a UDP tunnel of a delivery connection between the N-MADP and C-MADP.

o IPsec Tunneling: The user plane packets of the anchor connection are sent through an IPsec tunnel of a delivery connection.
o Client Net Address Translation (NAT): The Client IP address of user plane packet of the anchor connection is changed, and sent over a delivery connection.
o Pass Through: The user plane packets are passing through without any change over the anchor connection.

The MX adaptation sublayer also supports the following mechanisms and protocols to ensure security of user plane packets over the network path.

o IPsec Tunneling: An IPsec [RFC7296] tunnel is established between the N-MADP and C-MADP on the network path that is considered untrusted.
o DTLS: If UDP tunneling is used on the network path that is considered "untrusted", DTLS (Datagram Transport Layer Security) [RFC6347] can be used.

The Client NAT method is the most efficient due to no tunneling overhead. It SHOULD be used if a delivery connection is "trusted" and without NAT function on the path.

The UDP or IPsec Tunnelling method SHOULD be used if a delivery connection has a NAT function placed on the path.

4.2 GMA-based MX Convergence Sublayer

Figure 2 shows the MAMS u-plane protocol stack based on trailer-based encapsulation [GMA]. Multiple access networks are combined into a single IP connection. If NCM determines that N-MADP is to be instantiated with GMA as the MX Convergence Protocol, it exchanges the support of GMA convergence capability in the discovery and capability exchange procedures [MAMS].

```
          +-----------------------------------------------------+
          |                       IP PDU                        |
          |-----------------------------------------------------|
          |               GMA  Convergence Sublayer             |
          |-----------------------------------------------------|
          | MX Adaptation   | MX Adaptation | MX Adaptation     |
          | Sublayer        | Sublayer      | Sublayer          |
          | (optional)      | (optional)    | (optional)        |
          |-----------------------------------------------------|
          | Access #1 IP    | Access #2 IP  | Access #3 IP      |
          +-----------------------------------------------------+
```

Figure 2: MAMS U-plane Protocol Stack with GMA as MX Convergence Layer

Figure 3 shows the trailer-based Multi-Access (MX) PDU (Protocol Data Unit) format [GMA]. If the MX adaptation method is UDP tunneling and "MX header optimization" in the "MX_UP_Setup_Configuration_Request" message [MAMS] is true, the "IP length" and "IP checksum" header fields of the MX PDU SHOULD remain unchanged. Otherwise, they should be updated after adding or removing the GMA trailer in the convergence sublayer.

```
+-------------------------------------------------------+
| IP hdr |        IP payload              | GMA Trailer |
+-------------------------------------------------------+
```

Figure 3: GMA PDU Format

4.3 MPTCP-based MX Convergence Sublayer

Figure 4 shows the MAMS u-plane protocol stack based on MPTCP. Here, MPTCP is reused as the "MX Convergence Sublayer" protocol. Multiple access networks are combined into a single MPTCP connection. Hence, no new u-plane protocol or PDU format is needed in this case.

```
|-----------------------------------------------------|
|                        MPTCP                        |
|-----------------------------------------------------|
|  TCP             |   TCP           |      TCP       |
|-----------------------------------------------------|
| MX Adaptation    | MX Adaptation | MX Adaptation    |
| Sublayer         | Sublayer      | Sublayer         |
| (optional)       | (optional)    | (optional)       |
|-----------------------------------------------------|
| Access #1 IP     | Access #2 IP  | Access #3 IP     |
+-----------------------------------------------------+
```

Figure 4: MAMS U-plane Protocol Stack with MPTCP as MX Convergence Layer

If NCM determines that N-MADP is to be instantiated with MPTCP as the MX Convergence Protocol, it exchanges the support of MPTCP capability in the discovery and capability exchange procedures [MAMS]. MPTCP proxy protocols [MPProxy][MPPlain] SHOULD be used to manage traffic steering and aggregation over multiple delivery connections.

4.4 GRE as MX Convergence Sublayer

Figure 5 shows the MAMS u-plane protocol stack based on GRE (Generic Routing Encapsulation) [GRE2784]. Here, GRE is reused as the "MX Convergence sub-layer" protocol. Multiple access networks are combined

into a single GRE connection. Hence, no new u-plane protocol or PDU format is needed in this case.

```
+-------------------------------------------------------+
|      User Payload (e.g. IP PDU)                       |
|-------------------------------------------------------|
|              GRE as MX Convergence Sublayer           |
|-------------------------------------------------------|
|        GRE Delivery Protocol (e.g. IP)                |
|-------------------------------------------------------|
| MX Adaptation    | MX Adaptation | MX Adaptation      |
| Sublayer         | Sublayer      | Sublayer           |
| (optional)       | (optional)    | (optional)         |
|-------------------------------------------------------|
| Access #1 IP     | Access #2 IP  | Access #3 IP       |
+-------------------------------------------------------+
```

Figure 5: MAMS U-plane Protocol Stack with GRE as MX Convergence Layer

If NCM determines that N-MADP is to be instantiated with GRE as the MX Convergence Protocol, it exchanges the support of GRE capability in the discovery and capability exchange procedures [MAMS].

4.4.1 Transmitter Procedures

Transmitter is the N-MADP or C-MADP instance, instantiated with GRE as the convergence protocol that transmits the GRE packets. The Transmitter receives the User Payload (e.g. IP PDU), encapsulates it with a GRE header and Delivery Protocol (e.g. IP) header to generate the GRE Convergence PDU.

When IP is used as the GRE delivery protocol, the IP header information (e.g. IP address) can be created using the IP header of the user payload or a virtual IP address. The "Protocol Type" field of the delivery header is set to 47 (or 0X2F, i.e. GRE)[IANA].

The GRE header fields are set as specified below,

- If the transmitter is a C-MADP instance, then sets the LSB 16 bits to the value of Connection ID for the Anchor Connection associated with the user payload or sets to 0xFFFF if no Anchor Connection ID needs to be specified.
- All other fields in the GRE header including the remaining bits in the key fields are set as per [GRE_2784][GRE_2890].

4.4.2 Receiver Procedures

Receiver is the N-MADP or C-MADP instance, instantiated with GRE as the convergence protocol that receives the GRE packets. The receiver processes the received packets per the GRE procedures [GRE_2784, GRE_2890] and retrieves the GRE header.

- If the Receiver is an N-MADP instance,
  - o Unless the LSB 16 Bits of the Key field are 0xFFFF, they are interpreted as the Connection ID of Anchor Connection for the user payload. This is used to identify the network path over which the User Payload (GRE Payload) is to be transmitted.
- All other fields in the GRE header, including the remaining bits in the Key fields, are processed as per [GRE_2784][GRE_2890].

The GRE Convergence PDU is passed onto the MX Adaptation Layer (if present) before delivery over one of the network paths.

4.5 MX Adaptation and Convergence Co-existence

MAMS u-plane protocols support multiple combinations and instances of user plane protocols to be used in the MX Adaptation and the Convergence sublayers.

For example, one instance of the MX Convergence Layer can be MPTCP Proxy [MPProxy][MPPlain] and another instance can be Trailer-based. The MX Adaptation for each can be either UDP tunnel or IPsec. IPsec may be set up for network paths considered as untrusted by the operator, to protect the TCP subflow between client and MPTCP proxy traversing that network path.

Each of the instances of MAMS user plane, i.e. combination of MX Convergence and MX Adaptation layer protocols, can coexist simultaneously and independently handle different traffic types.

5. MX Convergence Control Message

A UDP connection may be configured between C-MADP and N-MADP to exchange control messages for keep-alive or path quality estimation. The N-MADP end-point IP address and UDP port number of the UDP connection is used to identify MX control PDU. Figure 6 shows the MX control PDU format with the following fields:

- o Type (1 Byte): the type of the MX control message
- o CID (1 Byte): an unsigned integer to identify the anchor and delivery connection of the MX control message
  - + Anchor Connection ID (MSB 4 Bits): an unsigned integer to identify the anchor connection

+ Delivery Connection ID (LSB 4 Bits): an unsigned integer to identify the delivery connection
- o MX Control Message (variable): the payload of the MX control message

Figure 7 shows the MX convergence control protocol stack, and MX control PDU goes through the MX adaptation sublayer the same way as MX data PDU.

```
                        <----MX Control PDU Payload --------------->
+------------------------------------------------------------------+
| IP header | UDP Header| Type | CID |       MX Control Message    |
+------------------------------------------------------------------+
```

Figure 6: MX Control PDU Format

```
|-----------------------------------------------------|
|          MX Convergence Control Messages            |
|-----------------------------------------------------|
|                  UDP/IP                             |
|-----------------------------------------------------|
| MX Adaptation   | MX Adaptation | MX Adaptation     |
| Sublayer        | Sublayer      | Sublayer          |
| (optional)      | (optional)    | (optional)        |
|-----------------------------------------------------|
| Access #1 IP    | Access #2 IP  | Access #3 IP      |
+-----------------------------------------------------+
```

Figure 7: MX Convergence Control Protocol Stack

5.1 Keep-Alive Message

The "Type" field is set to "0" for Keep-Alive messages. C-MADP may send out Keep-Alive message periodically over one or multiple delivery connections, especially if UDP tunneling is used as the adaptation method for the delivery connection with a NAT function on the path.

A Keep-Alive message is 6 Bytes long, and consists of the following fields:

- o Keep-Alive Sequence Number (2 Bytes): the sequence number of the keep-alive message
- o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.

5.2 Probe Message

The "Type" field is set to "1" for Probe messages.

N-MADP may send out the Probe message for path quality estimation. In response, C-MADP may send back the ACK message.

A Probe message consists of the following fields:

- o Probing Sequence Number (2 Bytes): the sequence number of the Probe REQ message
- o Probing Flag (1 Byte):
  - + Bit #0: a ACK flag to indicate if the ACK message is expected (1) or not (0);
  - + Bit #1: a Probe Type flag to indicate if the Probe message is sent during the initialization phase (0) when the network path is not included for transmission of user data or the active phase (1) when the network path is included for transmission of user data;
  - + Bit #2: a bit flag to indicate the presence of the Reverse Connection ID (R-CID) field.
  - + Bit #3~7: reserved
- o Reverse Connection ID (1 Byte): the connection ID of the delivery connection for sending out the ACK message on the reverse path
- o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.
- o Padding (variable)

The "R-CID" field is only present if both Bit #0 and Bit #2 of the "Probing Flag" field are set to "1". Moreover, Bit #2 of the "Probing Flag" field SHOULD be set to "0" if the Bit #0 is "0", indicating the ACK message is not expected.

If the "R-CID" field is not present but the Bit #0 of the "Probing Flag" field is set to "1", the ACK message SHOULD be sent over the same delivery connection as the Probe message.

The "Padding" field is used to control the length of Probe message.

5.3 Packet Loss Report (PLR) Message

The "Type" field is set to "2" for PLR messages.

C-MADP may send out the PLR messages to report lost MX SDU for example during handover. In response, C-MADP may retransmit the lost MX SDU accordingly.

A PLR message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection which the ACK message is for;

- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the anchor connection which the ACK message is for;
- o ACK number (4 Bytes): the next (in-order) sequence number (SN) that the sender of the PLR message is expecting
- o Number of Loss Bursts (1 Byte)
  For each loss burst, include the following
  - + Sequence Number of the first lost MX SDU in a burst (4 Bytes)
  - + Number of consecutive lost MX SDUs in the burst (1 Byte)

```
          C-MADP                                                N-MADP
              |                                                   |
              |<------------------- MX SDU (data packets)--------|
              |                                                   |
             +----------------------------------+                 |
             |Packet Loss detected              |                 |
             +----------------------------------+                 |
              |                                                   |
              |----- PLR Message ------------------------------>|
              |<-------------retransmit(lost)MX SDUs -----------|
```

Figure 8: MAMS Retransmission Procedure

Figure 8 shows the MAMS retransmission procedure in an example where the lost packet is found and retransmitted.

5.4 First Sequence Number (FSN) Message

The "Type" field is set to "3" for FSN messages.

N-MADP may send out the FSN messages to indicate the oldest MX SDU in its buffer if a lost MX SDU is not found in the buffer after receiving the PLR message from C-MADP. In response, C-MADP SHALL only report packet loss with SN not smaller than FSN.

A FSN message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection which the FSN message is for;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the anchor connection which the FSN message is for;
- o First Sequence Number (4 Bytes): the sequence number (SN) of the oldest MX SDU in the (retransmission) buffer of the sender of the FSN message.

Figure 9 shows the MAMS retransmission procedure in an example where the lost packet is not found.

```
          C-MADP                                              N-MADP
              |                                                  |
              |<------------------ MX SDU (data packets)--------|
              |                                                  |
            +---------------------------------+                  |
            |Packet Loss detected             |                  |
            +---------------------------------+                  |
              |                                                  |
              |----- PLR Message ------------------------------>|
              |                               +---------------------+
              |                               |Lost packet not found|
              |                               +---------------------+
              |<-------------FSN message -----------------------|
```

Figure 9: MAMS Retransmission Procedure with FSN

5.5 Coded MX SDU (CMS) Message

The "Type" field is set to "4" for CMS messages.

N-MADP (or C-MADP) may send out the CMS message to support downlink (or uplink) packet loss recovery through coding, e.g. [CRLNC], [CTCP], [RLNC]. A coded MX SDU is generated by applying a network coding algorithm to multiple consecutive (uncoded) MX SDUs, and it is used for fast recovery without retransmission if any of the MX SDUs is lost.

A Coded MX SDU message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection of the coded MX SDU;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class of the coded MX;
- o Sequence Number (4 Bytes): the sequence number of the first (uncoded) MX SDU used to generate the coded MX SDU.
- o Fragmentation Control (FC) (1 Byte): to provide necessary information for re-assembly, only needed if the coded MX SDU is too long to transport in a single MX control PDU.
- o N (1 Byte): the number of consecutive MX SDUs used to generate the coded MX SDU
- o K (1 Byte): the length (in terms of bits) of the coding coefficient field
- o Coding Coefficient (N x K / 8 Bytes)
  - + a(i): the coding coefficient of the i-th (uncoded) MX SDU

      + padding
- o Coded MX SDU (variable): the coded MX SDU

If K = 0, the simple XOR method is used to generate the Coded MX SDU from N consecutive uncoded MX SDUs, and the a(i) fields are not included in the message.

If the coded MX SDU is too long, it can be fragmented, and transported by multiple MX control PDUs. The N, K, and a(i) fields are only included in the MX PDU carrying the first fragment of the coded MX SDU.

```
          C-MADP                                             N-MADP
              |                                                  |
              |<------------------ MX SDU #1 -------------------|
              |      lost<-------- MX SDU #2 -------------------|
              |<---- CMS Message (MX SDU #1 XOR MX SDU #2)------|
             +-----------------------+                           |
             | MX SDU #2 recovered  |                           |
             +-----------------------+                           |
              |                                                  |
```

Figure 10: MAMS Packet Recovery Procedure with XOR Coding

5.6 Traffic Splitting Update (TSU) Message

The "Type" field is set to "5" for TSU messages.

N-MADP (or C-MADP) may send out a TSU message if downlink (or uplink) traffic splitting configuration has changed.

A TSU message consists of the following fields:

- o Connection ID (1 Byte): an unsigned integer to identify the anchor connection;
- o Traffic Class ID (1 Byte): an unsigned integer to identify the traffic class;
- o Sequence Number (2 Bytes): an unsigned integer to identify the TSU message.
- o Flags (1 Byte)
  - + Bit #0: a Reverse Path bit flag to indicate if the traffic splitting configuration is for the reverse path (1) or not (0);
  - + Bit #1: a Bit-Reversal bit flag to indicate if bit-reversal is used in traffic splitting
  - + Others: reserved.
- o Traffic Splitting Configuration Parameters (5 + (N -1) Bytes):

+ StartSN (4 Bytes): the sequence number of the first MX SDU using the traffic splitting configuration provided by the TSU message
+ L (1 Byte): the traffic splitting burst size
+ K(i): the traffic splitting threshold of the i-th delivery connection, where connections are ordered according to their Connection ID.

Let's use f(x) to denote the traffic splitting function, which maps a MX SDU Sequence Number "x" to the i-th delivery connection.

f(x)=i, if K[i-1]< or = mod(x - StartSN, L) < K[i]

Wherein, 1 < or = i < N, K[0]=0, and K[N]=L.

N is the total number of connections for delivering a data flow, identified by (anchor) Connection ID and Traffic Class ID.

When the bit-reversal bit is set to 1, the burst size L MUST be a power of 2, and the traffic splitting function is

f(x)=i, if K[i-1]< or = F(mod(x - StartSN, L)) < K[i]

Wherein F(.) is the bit reversal function [BITR] of the input variable.

5.7 Acknowledgement Message

The "Type" field is set to "6" for ACK messages.

C-MADP (or N-MADP) SHOULD send out the ACK message in response to the successful reception of a PLR, FSN, or TSU message.

C-MADP SHOULD send out the ACK message in response to a Probe message with the ACK flag set to "1".

The ACK message consists of the following fields:

o Acknowledgment Number (2 Bytes): the sequence number of the received message.
o Timestamp (4 Bytes): the current value of the timestamp clock of the sender in the unit of 100 microseconds.

6 Security Considerations

User data in MAMS framework rely on the security of the underlying network transport paths. When this cannot be assumed, NCM configures use of appropriate protocols for security, e.g. IPsec [RFC4301] [RFC3948], DTLS [RFC6347].

7 IANA Considerations

This draft makes no requests of IANA.

8 Contributing Authors

The editors gratefully acknowledge the following additional contributors in alphabetical order: Salil Agarwal/Nokia, Hema Pentakota/Nokia.

9 References

9.1 Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC4301] Kent, S. and K. Seo, "Security Architecture for the Internet Protocol", RFC 4301, DOI10.17487/RFC4301, December 2005, <http://www.rfc-editor.org/info/rfc4301>.

9.2 Informative References

[RFC6347] Rescorla, E. and N. Modadugu, "Datagram Transport Layer Security Version 1.2", RFC 6347, January 2012, <http://www.rfc-editor.org/info/rfc6347>.

[RFC7296] Kaufman, C., Hoffman, P., Nir, Y., Eronen, P., and T. Kivinen, "Internet Key Exchange Protocol Version 2 (IKEv2)", STD 79, RFC 7296, DOI 10.17487/RFC7296, October 2014, <http://www.rfc-editor.org/info/rfc7296>.

[RFC3948] Huttunen, A., Swander, B., Volpe, V., DiBurro, L., and M. Stenberg, "UDP Encapsulation of IPsec ESP Packets", RFC 3948, DOI 10.17487/RFC3948, January 2005, <http://www.rfc-editor.org/info/rfc3948>.

[MPProxy] X. Wei, C. Xiong, and E. Lopez, "MPTCP proxy mechanisms", https://tools.ietf.org/html/draft-wei-mptcp-proxy-mechanism-02

[MPPlain] M. Boucadair et al, "An MPTCP Option for Network-Assisted MPTCP", https://www.ietf.org/id/draft-boucadair-mptcp-plain-mode-09.txt

[MAMS] S. Kanugovi, S. Vasudevan, F. Baboescu, and J. Zhu, "Multiple Access Management Protocol", https://tools.ietf.org/html/draft-kanugovi-intarea-mams-protocol-03

[GMA] J. Zhu, "Trailer-based Encapsulation Protocols for Generic Multi-Access Convergence", https://tools.ietf.org/html/draft-zhu-intarea-gma-01

[GRE2784] D. Farinacci, et al., "Generic Routing Encapsulation (GRE)", RFC 2784 March 2000, <http://www.rfc-editor.org/info/rfc2784>.

[GRE2890] G. Dommety, "Key and Sequence Number Extensions to GRE", RFC 2890 September 2000, <http://www.rfc-editor.org/info/rfc2890>.

[IANA] https://www.iana.org/assignments/protocol-numbers/protocol-numbers.xhtml

[LWIPEP] 3GPP TS 36.361, "Evolved Universal Terrestrial Radio Access (E-UTRA); LTE-WLAN Radio Level Integration Using Ipsec Tunnel (LWIP) encapsulation; Protocol specification"

[RFC791] Internet Protocol, September 1981

[CRLNC] S Wunderlich, F Gabriel, S Pandi, et al. Caterpillar RLNC (CRLNC): A Practical Finite Sliding Window RLNC Approach, IEEE Access, 2017

[CTCP] M. Kim, et al. Network Coded TCP (CTCP), eprint arXiv:1212.2291, 2012

[RLNC] J. Heide, et al. Random Linear Network Coding (RLNC)-Based Symbol Representation, https://www.ietf.org/id/draft-heide-nwcrg-rlnc-00.txt

[BITR] Alan H. Karp, "Bit reversal on uniprocessors", SIAM Review, 38 (1): 1-26, 1996.

Authors' Addresses

Jing Zhu

Intel

Email: jing.z.zhu@intel.com

SungHoon Seo

Korea Telecom

Email: sh.seo@kt.com

Satish Kanugovi

Nokia

Email: satish.k@nokia.com

Shuping Peng

Huawei

Email: pengshuping@huawei.com