Protocol for interactive low-latency media transmission system
draft-liu-protocol-interactive-media-transmission-00

Abstract

This document introduces a protocol used for interactive low-latency
media transmission network.

Requirements Language

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 RFC 2119 [1].

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

Emerging real-time interactive video/audio communication applications
bring new challenges for existing protocols. This documents
introduces the use cases, requirements and protocol for interactive
low-latency multimedia transmission network over the Internet.

Interactive real-time media communication is getting popular with the
growth of short video, on-line education, on-line gaming and other
similar applications. Some application providers build their own
interactive real-time media communication network to support their
applications yet facing high cost and technical issues. For example,
interactive communication between users is unpredictable, which
results in high cost when dedicated entity for interaction is used
and the wastage of reserved resources for interaction.

To avoid the aforementioned issues and challenges, some other
application providers attempt to use third party's interactive real-
time media communication network provided by giant cloud operator.
However, there are several challenges of existing protocol to support
the above mentioned scenarios.

1. Interactive Online broadcasting service is flexible and much more
complicated compared with traditional media broadcasting service.
For interactive Online broadcasting applications, audiences may

occasionally request to setup bidirectional real-time communication
with the broadcaster and all the other audiences should be able to
receive the merged interactive media traffic containing the
broadcaster and connected audience. To meet this end, there is a
need for standardized signaling protocol which can support media
stream merging switching and pulling to support those complicated
scenarios.

2. Applications such as interactive online broadcasting, short
video, on-line education, on-line gaming are very delay sensitive.
Thus, the protocols for media stream merging switching and pulling
should be able to meet the latency requirement for those
applications.

3. There are many different media transmission protocols (e.g.
QUIC, WebRTC, etc) across different layers, which are widely used in
the ecosystem, the protocols for media stream merging switching and
pulling should be able to compatible with different transmission
protocols.

This document specifies a protocol for media stream merging, pulling
and switching that used for Interactive real-time media communication
system.

2. **System Architecture**

This section specifies the system architecture of the Interactive
real-time media communication system.

```
                                              Sever for media streaming
control
                                          +-----------+
                                          |           |
                                          |           |
                                          |           |
                                          |           |
                                          |           |
                                          +-----+-----
+                         +-----+
   +-----+                      |       |
|        |                      |       |
    |     |                     |       |
|        |                      |       |
    |     |                     |       |
|        |                      |       |
    |     |<------+             |       |
|              +------->|       |
    |     |     |              +----------------v-------------------
+    |         |     |                         |
    |     |     |                              |
|   |         +-----+                          |
   +-----+     |                               |
|   |       Audience                           |
  Broadcaster    |                             |
+----+        +-----+                          |
              +--------------------->|
|         |         |     |
                                      |Interactive real-time communication
|         |         |     |
                                      |            network
+------------>|      |
   +-----+        +--------------------->|
|        |         |     |
    |     |    |        |                      |
|        |    |        |                       |
    |     |    |        |                      |
+----+         +-----+
    |     |<------+                            |
|   |        Audience                          |
    |     |                                    |
|   |         +-----+                          |
    |     |                         +-------------^-----+---------------
+    |         |     |                            |
   +-----+                                        |
|                    |         |     |            |
   Audience                                       |
|                        +------->|      |
```

```
connected with                                     |
|                              |      |
 the broadcaster                                   +-+-----v---
+                             |      |
                                                   |
|                            +-----+
                                                   |
|                            Audience
                                                   |          |
                                                   |          |
                                                   |          |
                                                   +----------+
                                             Server for media stream merging
```
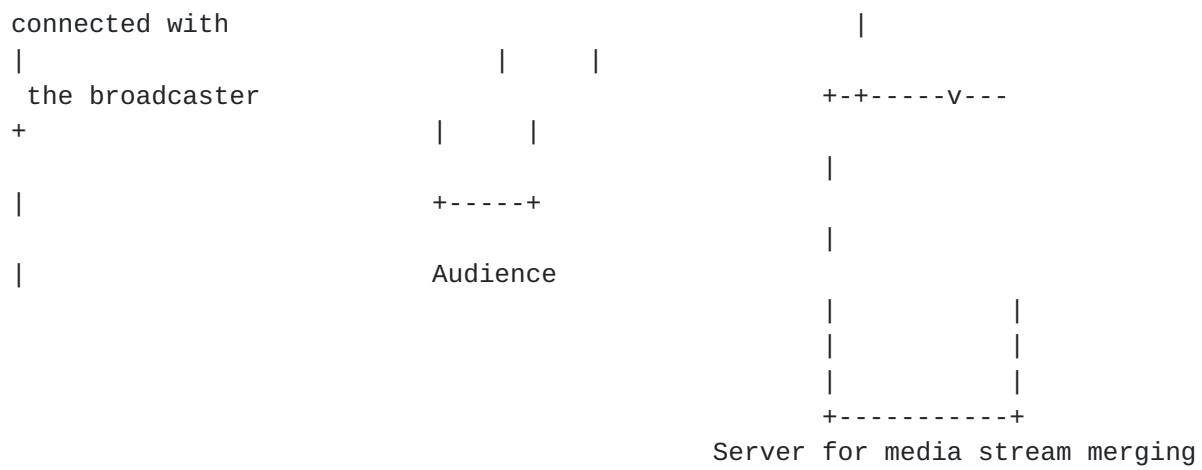
Figure 1: Architecture

The interactive real-time communication network can be provided by
giant cloud provider. The communication network can provide
fundamental capabilities of media stream, including media pulling,
media pushing. In addition, the network can also support
capabilities such as media merging and media switching. The
capabilities can be triggered by control server and server for media
streaming merging, which are provided by 3rd party. Based on those
capabilities, the audience can receive corresponding media from
broadcaster or merged media between broadcaster and requested
audience for interaction seamlessly.

[3]. Signaling procedure

This section defines the signaling procedure of Interactive real-time
media communication network.

Figure 2 shows the signaling procedure of Interactive real-time media
communication among broadcaster, requested audience for interaction
and other audience. The broadcaster and audience firstly push their
media streams to the interactive real-time media communications
network. A audience wishes to interact with the broadcaster and thus
sends a request to the control server for interaction. The control
server processes the request and sends comment for media merging to
the server for media stream merging. Upon the receipt of merging
request from control server, the server for media stream merging
pulls the corresponding streams from both the broadcaster and the
requested audience for interaction and processes with the media
merging.

After the completion of media merging, the server for media stream
merging pushes the merged media to the Interactive real-time media
communication network which then sends the merged media to
corresponding edge media distribution servers which connects the
audiences who watch the media. After the distribution, the control
server sends the command for media switching to the Interactive real-
time media communication network. The network then forwards the
switching signaling message to the edge node. Up the receipt of the
signaling message, the edge node performs the media switching by
pushing the merged media to the audiences.

Audience
Interactive real-time
                              connected
with   Server for media   media
communication
Broadcaster   the broadcaster   Control Server
stream merging   network   Audience

```
+--------------+             +--------------+             +--------------
+            +--------------+           +--------------+
+--------------+
|              |             |             |             |
|              |             |             |             |             |
|
|              |             |             |             |
|              |             |             |             |             |
|
|              |             |             |             |
|              |             |             |             |             |
|
|              |             |             |             |
|              |             |             |             |             |
|
```

```
+------+-------+           +-------+------+          +-------+------
+             +------+-------+         +-------+------+          +------
+-------+
      |                          |
|                          |
|                          |
      |                          |
|                          |
|                          |
      |                          |
|                          |
|                          |
      |                          |
|                          |
|                          |
      |                          |
|                          |
|                          |
      |                          |
|                          |
|                          |
      |                          |
|                          |
|                          |
      |                          |           Push media stream
|                          |
|                          |
      +---------------------------+----------------------------
+----------------------+----------------------------
>|                          |
      |                          |
|                          |
|                          |
      |                          |
|                          |
|                          |
      |                          |
|                          |
|                          |
      |                          |
|                          |            |   Pull media
stream     |
      |                          |
|                          |            |
<------------------------+
      |                          |
|                          |
|                          |
      |                          |
|                          |
|                          |
      |                          |            |      Push media
stream      |                          |            |
      |                          +----------------------------
```

```
+--------------------------+------------------------
>|                          |
     |                          |
 |                          |
 |                          |
     |                          |
 |                          |
 |                          |
     |                          |
 |                          |
 |                          |
     |                          |
 |                          |
 |                          |
     |    Pull media stream     |
 |                          |
 |                          |
     +--------------------------+--------------------------
+-------------------------+------------------------
>|                         |
     |                         |
 |                         |
 |                         |
     |                         |
 |                         |
 |                         |
     |                         |
 |                         |
 |                         |
     |                         |
 |                         |
 |                         |
     |                         |   Pull media stream
 |                         |
 |                         |
     |                         +--------------------------
+-------------------------+------------------------
>|                         |
     |                         |
 |                         |
 |                         |
     |                         |                        |Command for
stream merging  |                         |                        |
     |                         |
+------------------------->| Pull stream for merging
 |                         |
     |                         |
 |                         +--------------------------
 |                         |
>|                         |
     |                         |
 |                         |
```

```
|                            |
    |                            |
|                            | Push merged stream
|                            |
    |                            |
|                            +------------------------
>|                            |
    |                            |                        |Command for
stream switching|                            |                    |
    |                            |
+---------------------------+------------------------
>|                            |
    |                            |
|                            |                        +--
+                            |
    |                            |
|                            |                        |
|                            |
    |                            |
|                            |                        |<-
+                            |
    |                            |
|                            |                        |Perform stream
switch     |
    |                            |
|                            |
|                            |
```

Figure 2: Procedure
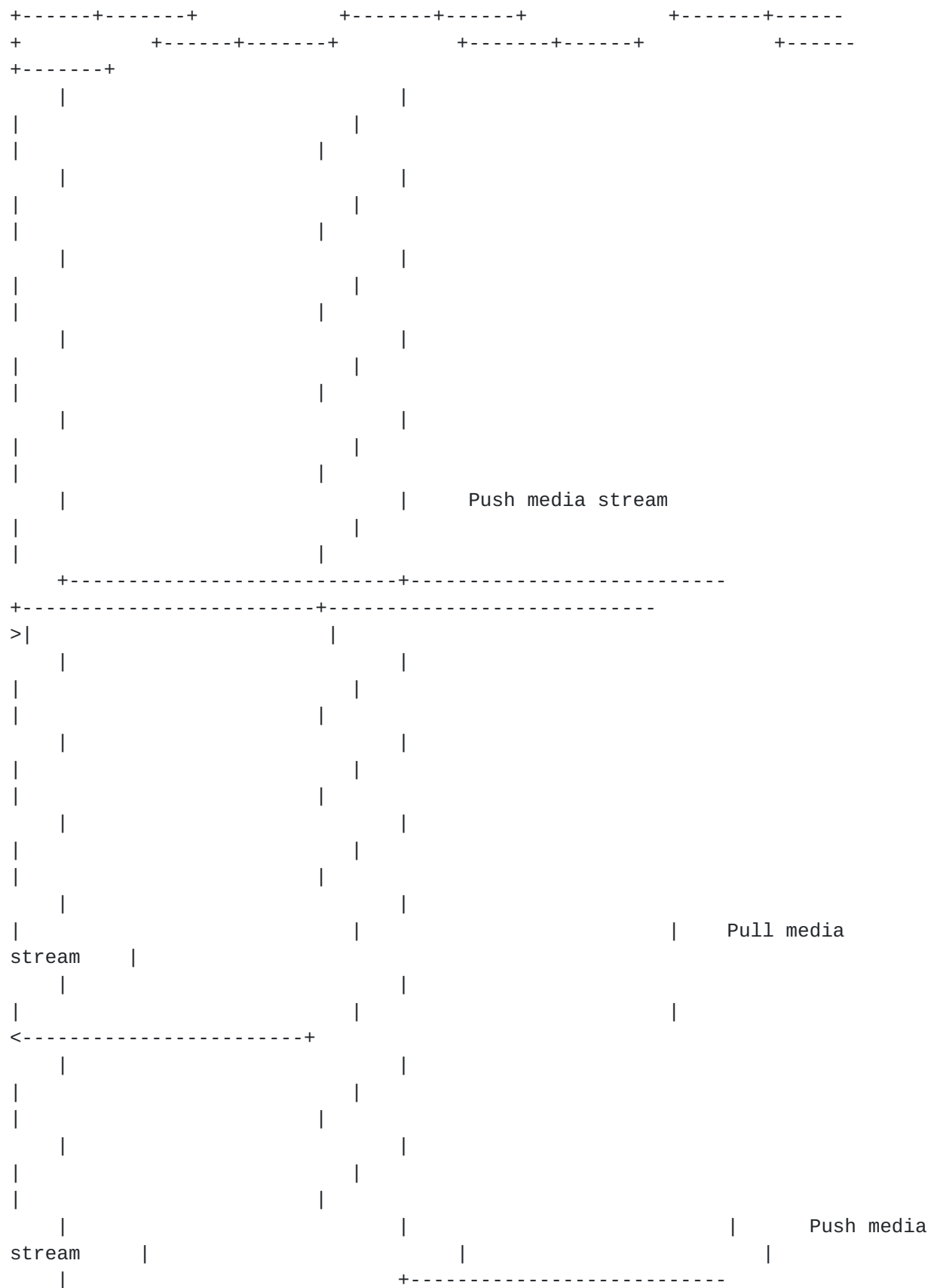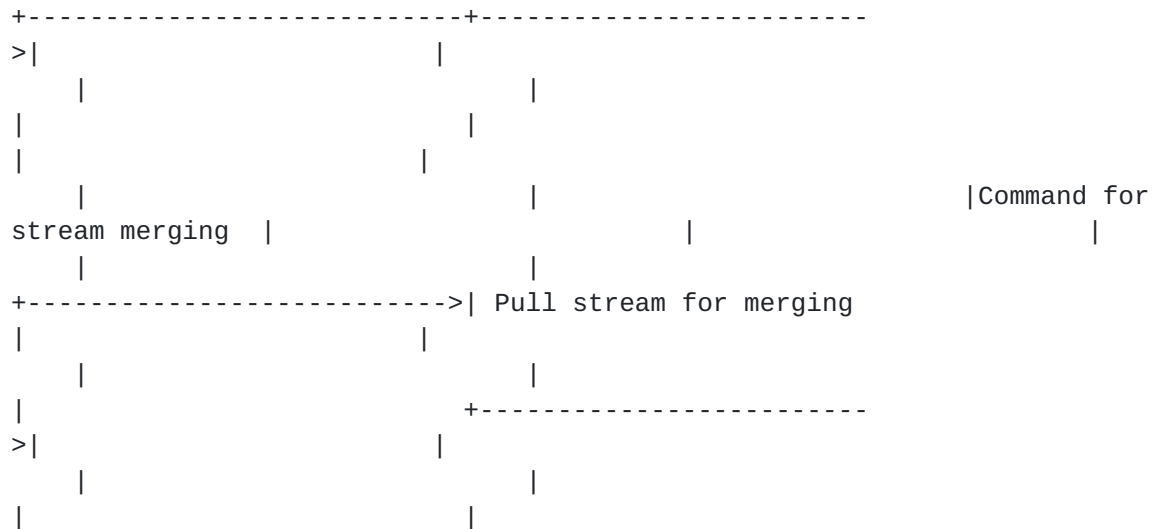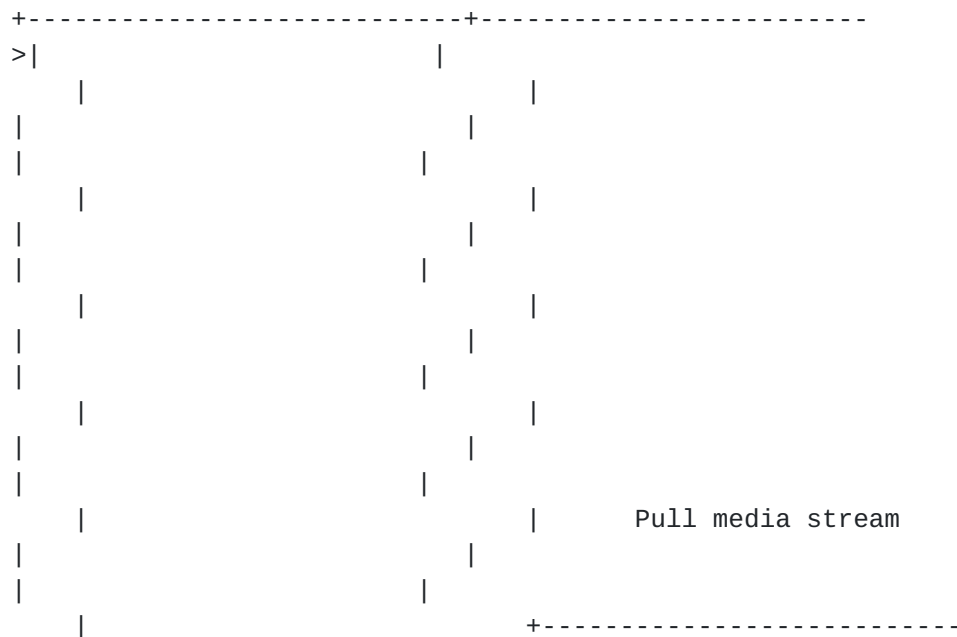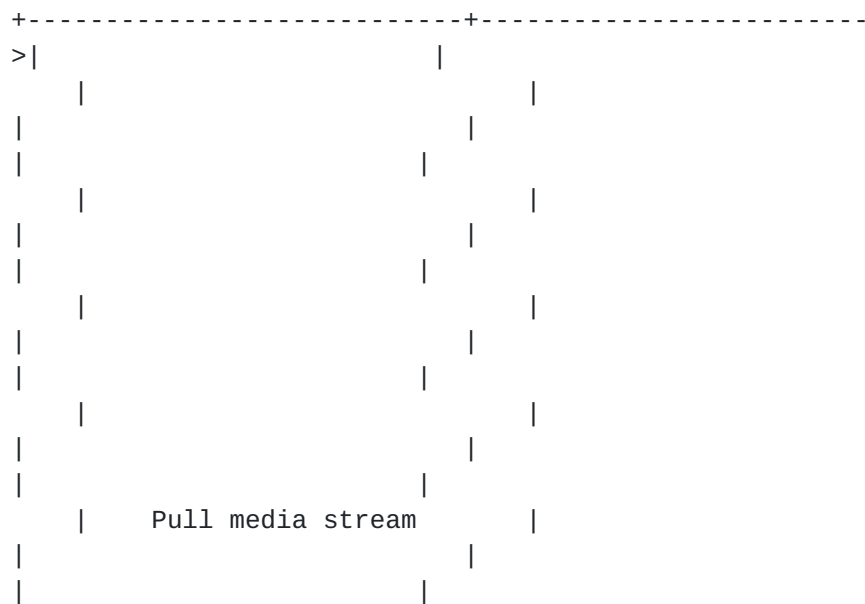
4. **Signaling Specification**

This section defines the signaling specification for the interactive
real time media communication. In order to achieve the merging and
switching functionalities for different media source, signaling
messages need to be delivered to the corresponding entities (e.g.
control server, edge node, etc) in order to perform the proper
operations. The signaling message of interactive media control
protocol is shown as follows:

```
Interactive Media Control Message {
  Message Type (i),
  Message Length (i),
  Message Payload (..),
}
```

Figure 3: Interactive media signaling message

To process with the signaling message, the corresponding entities
need to identify the type of signaling message. This can be achieved
via using message type which can be carried by the message header.
The message types of Interactive media control protocol can be
described as follows:

```
+=====+===========+
|  ID | Messages  |
+=====+===========+
| 0x0 | Merging   |
+-----+-----------+
| 0x1 | Switching |
+-----+-----------+
| 0x2 | Grabbing  |
+-----+-----------+
| 0x3 | Pulling   |
+-----+-----------+
| 0x4 | Pushing   |
+-----+-----------+
```

Table 1: Message types
of Interactive media
control protocol

The message length indicates the total length of the message payload
filed in bytes. Message payload contains the information for
controlling media merging and media switching. The subsequent sub-
section describes these two message types and related payload in
detail.

4.1. Merging signaling message

Merging signaling message is used to request the server for media
stream merging to perform media merging between a broadcaster and an
audience. The merging signaling message is shown as follows:

```
Merging Message {,
  Payload Type (i),
  first media info {
    1st media ID (i),
    1st media URL (b),
  }
  2nd media info {
    2nd media ID (i),
    2nd media URL (b),
  }
}
```

Figure 4: Merging signaling message

The payload type field in the header indicates the merging signaling
message. First media ID and second media ID represent the IDs of the
media from broadcaster and the requested audience for interaction,
respectively. The ID is comprised of a string which represents the
unique ID of an media source. Each media info contains the media ID,
media URL. The media URL represents the address of edge node which
interacts with the audience.

4.2. Switching signaling message

Switching signaling message is used to instruct the Interactive real-
time media communication system to perform media switching upon the
receipt of the request from the control server. The switching
signaling message is shown as follows:

```
Switching Message {
  Payload Type (i),
  Source media info {
    Src media ID (i),
    Src media URL (b),
  },
  Destination media info {
    Dst media ID (i),
    Dst media URL (b),
  }
}
```

Figure 5: Switching signaling message

The payload type field in the header indicates the switching
signaling message. Source media info contains the information
regarding source media from the broadcaster. Destination media info
contains the information regarding destination media which is the
merged media between the broadcaster and the requested audience for
interaction. Each media info contains the media ID, media URL.

The switch signaling message is sent to the edge node which manages
the media delivery for the audience. If the edge node acknowledges
the media switching, it re-directs the media content with the
destination media using media transmission protocols (e.g. QUIC,
WebRTC, etc). Upon the receipt of the switching signaling message,
the media transmission protocol decides time-stamp, information
regarding I-frame, and optionally the sequence number to achieve the
re-direction of the new merged media. This is to make sure that the
audience can smoothly switch to the merged media without the negative
impact on user experience.

4.3. Grabbing signaling message

Grabbing signaling message is used to instruct the Interactive real-
time media communication system to switch edge node for audience, for
example, in mobility scenario. In the mobility case, the Interactive
real-time media communication system may decide to switch a more
suitable edge node for media pushing for an audience according the
location information. The grabbing signaling message is shown as
follows:

```
Grabbing Message {
  Payload Type (i),
  new media info {
    new media ID (i),
    new media URL (b),
  },
  error_code,

  }
}
```

Figure 6: Grabbing signaling message

The grabbing signaling message is sent from Interactive real-time
media communication system to the edge node. A new edge node firstly
start pushing media to the audience. Meanwhile, it registers the
service to the Interactive real-time media communication system. The
system detects that the media pushing service already exists and thus
sends the grabbing signaling message to the old edge node. For the
old edge node, the grabbing signaling message is used to instruct the
node to drop the media pushing to the audience. The error code
indicates the reason for dropping. The reasons are shown below:

```
+========+====================+
| Reason | Code               |
+========+====================+
|    0x0 | Dropped by Mobility |
+--------+--------------------+
|    0x1 | Proactive dropping  |
+--------+--------------------+
|    0x2 | Passive dropping    |
+--------+--------------------+
```

Table 2: Reason code for
grabbing signaling message

Dropped by Mobility indicates the case where a new edge node has
taken place and pushes the media to the audience instead of the old
edge node. Proactive dropping indicates the case where an edge node
gets issues on the media pushing and the audience can request for re-
connection for the delivery of the media. Passive dropping indicates
the case where the corresponding media has been banned and thus can
not be pushed anymore.

4.4. Pulling signaling message

Pulling signaling message is sent from audience to the edge node.
Once the pulling signaling message is acknowledged, the edge node
sends the corresponding media to the audience. The pulling signaling
message is shown below:

```
Pulling Message {
  Payload Type (i),
  Media info {
    Media URL (b),
  }
}
```

Figure 7: Pulling signaling message

The payload type field in the header indicates the pulling signaling
message. The media URL indicates the address of the target media
which can be obtained from the edge node.

The edge node allocates a media ID for the broadcaster or the
requested audience for interaction so that the media can be uniquely
identified in the communication system. Upon the receipt of the
pulling signaling message, the edge node acknowledges the signaling
message with the media ID which uniquely identifies the target media.

4.5. Pushing signaling message

Pushing signaling message is sent from broadcaster or the requested
audience for interaction to the edge node in order to start pushing
media to the edge node. The pulling signaling message is shown
below:

```
Pushing Message {
  Payload Type (i),
  Media info {
    Media URL (b),
  }
}
```

Figure 8: Pushing signaling message

The payload type field in the header indicates the pushing signaling
message. The media ID indicates the media that is about to sent to
the edge node. The URL represents the address of the broadcaster or
the requested audience for interaction which push media to the edge
node. Upon the receipt of the pulling signaling message, the edge
node acknowledges the signaling message with a media ID which
uniquely identifies the pushed media.

5. Acknowledgements

6. IANA Considerations

TBD.

7. Security Considerations

The signaling messages defined in this document should be protected
by security mechanism.

8. Normative References

   [1]   Bradner, S., "Key words for use in RFCs to Indicate
              Requirement Levels", March 1997,
              <http://xml.resource.org/public/rfc/html/rfc2119.html>.

   [2]   Rose, M., "Writing I-Ds and RFCs using XML", RFC 2629,
              DOI 10.17487/RFC2629, June 1999,
              <https://www.rfc-editor.org/info/rfc2629>.

Authors' Addresses

   Dapeng(Max) Liu
   Alibaba Group
   Email: max.ldp@alibaba-inc.com

   Yaming He
   Alibaba Group
   Email: heyaming.hym@alibaba-inc.com

   Xiaobo Yu
   Alibaba Group
   Email: shibo.yxb@alibaba-inc.com

   Xiao Kai
   Alibaba Group
   Email: xiaokaikai.xk@alibaba-inc.com

   Songlin Li
   Alibaba Group
   Email: songlin.lsl@alibaba-inc.com