Media Handling Considerations for Wireless Networks

draft-kaippallimalil-tsvwg-media-hdr-wireless-04

Authors: John Kaippallimalil, Sri Gundavelli, Spencer Dawkins

IETF 119 Brisbane, March 2024

Requirements overview

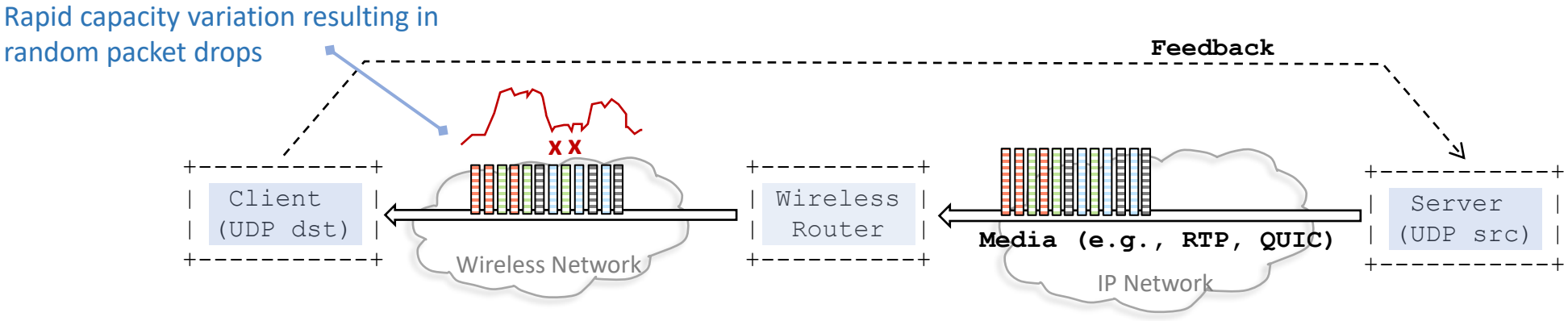
metadata for media frames

metadata for media streams

independent of frame/stream

transport requirements

Media Data Unit Handling



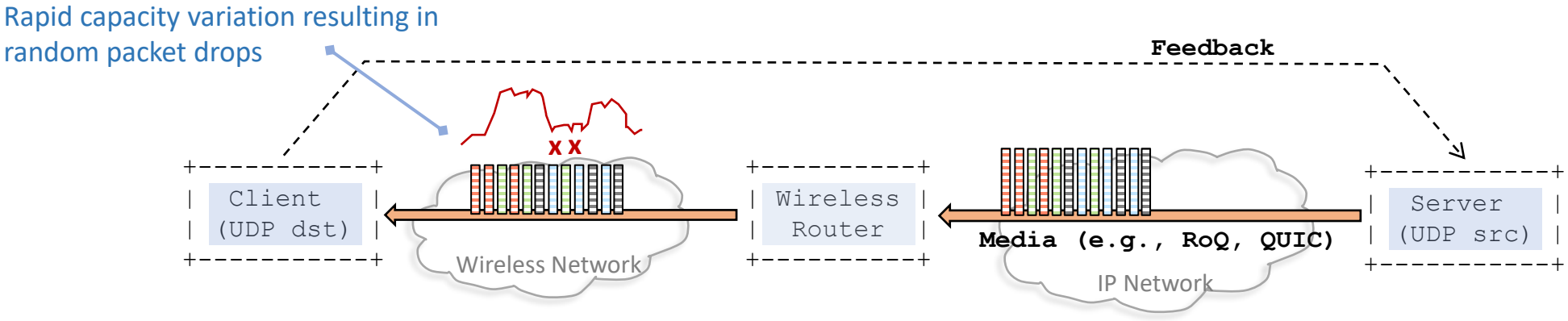
1. **Identify media frames**
   Treat packets of a media frame in the same manner.

2. **Relative Priority of media frames**
   Relative importance of a media frame over others in the flow.
   (used to set drop priority in case of congestion/high load)

3. **Tolerance to delay**
   Some media frames may be able to tolerate delay in the network. In this case, the network may delay packets rather than drop.

> 3.1. Requirement 1: Classification of downlink media frames
>
> Feedback provided by ECN/L4S to the server (UDP sender) is not fast enough to adjust the sending rate when available wireless capacity changes significantly in very short periods of time (~ 1 millisecond).
> Differentiating using multiple DSCP codes does not provide the resolution required to classify media frames and adapt to changes in coding due to dynamic content or resulting from network conditions.
>
> Relative priority of media frames, tolerance to delay, identification of media frame boundaries are provided by the application to optimize traffic shaping at the wireless router. Alternatively, an application may prefer to provide only the information about streams and their relative priority (see Section 3.2). In such cases it does not provide any information to classify media frames.

Media stream handling



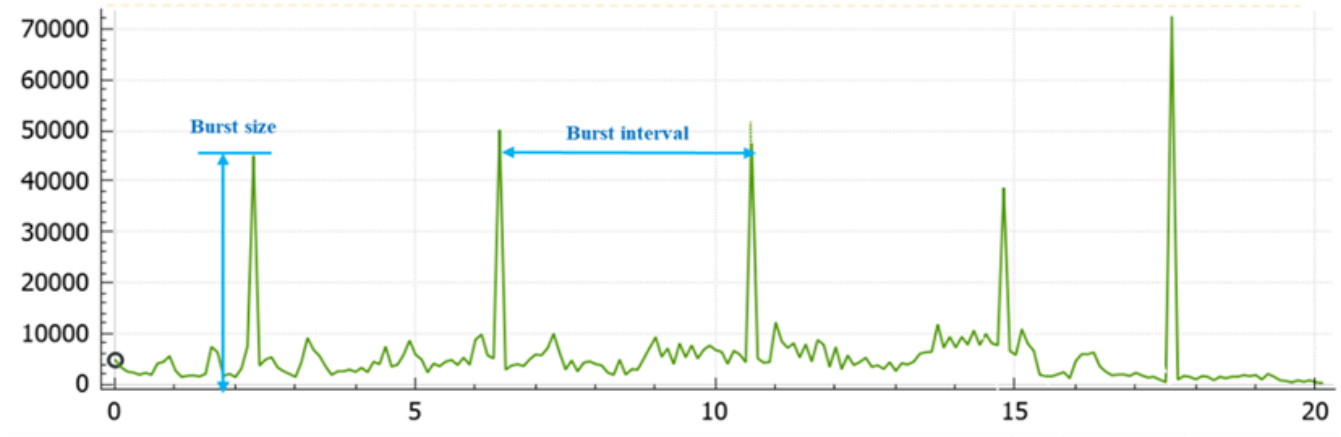
1. **Identify media streams**
   Treat packets of a media stream in the same manner.

2. **Relative Priority of media stream**
   Relative importance of a media stream over other streams in the flow.
   (used to set drop priority in case of congestion/high load)

3. **Tolerance to delay**
   Some media streams may be able to tolerate delay in the network. In this case, the network may delay packets rather than drop.

> 3.2. Requirement 2: Classification of downlink streams
>
> In some deployments, multiple media streams with different priorities are sent in a single transport connection (e.g., WebRTC [RFC8854], QUIC transports for multimodal media, audio, video, control, haptics). In such cases, an application may want to prioritize one stream over another in the event of extreme congestion (e.g., audio stream prioritized over video stream).
>
> The application may prefer to provide only the information about streams and their relative priority, and in such cases it does not provide any information to classify media frames (see Section 3.1). In this case, the application conveys to the wireless network the relative priority of media streams in a single transport connection.

Burst size, jitter and packet loss



Ref: Support of dynamic change in traffic characteristics, 3GPP S2-2403095.

Burst size is defined in the draft as a requirement.
The assumption is that the server knows the size of the burst at the beginning since it will use it to pace packets, etc.

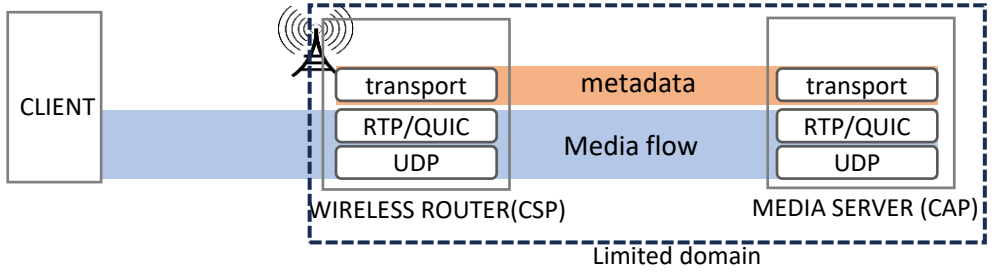
Burst interval is useful for periodic data (may need to be added)

Jitter is useful for the wireless router to convey packet delays (early/late arrivals) to the scheduler.

Detection of lost packets is needed to identify if media frame boundary packets are lost. (not used for

```
3.3.  Requirement 3: Size of a burst of packets

   Media flows can have large and unexpected variations in packet bursts
   due to dynamic changes in content, server estimation of network
   conditions and pacing behavior.  Encoding of live video, and
   multimodal media can only increase the burst size that a server has
   to contend with sending out in a relative smoothed out manner.  The
   burst size is observable on the wire, but can only be determined by
   the end of the burst of packets.  Wireless networks on the other hand
   cannot reserve resources for the maximum burst size as that will lead
   to poor utilization of the radio resources.

   The server should provide expected size of a burst of packets at the
   beginning of the burst to allow the scheduler to reserve sufficient
   resources (and avoid have too few resources that lead to a tail
   drop).  Application servers that are not able to reliably estimate
   the burst size at the beginning of a burst should not provide any
   information about the burst size.

3.4.  Requirement 4: Jitter

   Packets of a burst that experience a high level of jitter are likely
   experiencing higher than usual delay before arriving at the wireless
   network.  The wireless router can use packet timestamps to derive
   jitter and optimize downlink scheduling.

   A timestamp with resolution to the microsecond level (e.g., short
   format in [RFC5905]) is sufficient for these purposes.

3.5.  Requirement 5: Detection of loss of packets

   The wireless router should be aware of any loss of packets belonging
   to a media frame.  For example, the loss of a packet that is a start
   or end of a media frame can cause confusion in estimating media frame
   boundaries.  However, the wireless router does not re-order packets
   arriving out of order at the wireless router as this will increase
   latency experienced by the flow.

   The server should provide a sequence number that identifies the
   sequence of packets for a transport connection carrying the media
   flows.
```

Privacy Considerations



Media flows are between client and server.
Metadata is sent in a side-channel separate from the media flow between CAPs to CSPs in a limited domain only.

The privacy and security aspects of the e2e media flow itself is out of scope.

Burst size, sequence number, timestamp is either observable on-path or can be inferred.

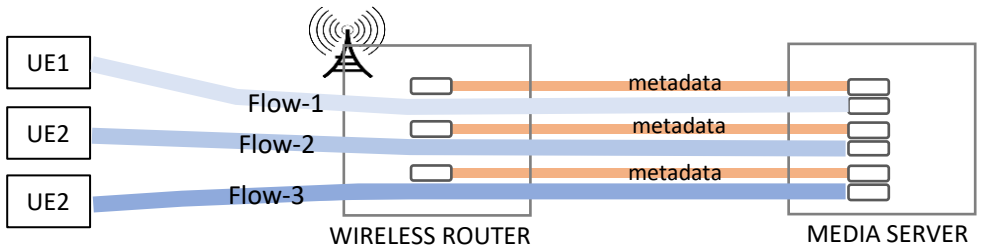
Relative importance, frame boundaries and tolerance to delay convey DSCP-like information to make better QoS decisions.

Note: Frequency analysis on RTP/SRTP flows can determine what content is sent. In QUIC with dynamic stream multiplexing, this can be obscured.

All metadata must be integrity protected to detect modification.

> 3.6. Requirement 6: Privacy Considerations
>
> Encrypted media payloads along with temporary IP addresses between a server and user (client) provide a measure of privacy for the content and the identity of the user. It should however be noted that media flows (e.g., encrypted video payloads in SRTP) exhibit a pattern of bursts and intervals that amounts to a signature and is vulnerable to frequency analysis. To avoid this kind of frequency analysis, media sent by the server would need to be scheduled or multiplexed differently to each user/recipient. This may be possible in transports like QUIC which allows flexibility in scheduling each stream. Transports like QUIC also fully encrypt the entire stream and therefore no media headers are observable on path either. The security aspects of the media payload/ transport are not in the scope of these requirements and is described here only to provide context for metadata privacy. Privacy considerations for the metadata itself should ensure that no additional information about either the content or the user of the content is revealed.
>
> Some of the metadata like the size of a burst of packets, sequence number and timestamp are information that can be plainly observed or inferred by an entity on path. These and all other metadata sent from server to the wireless router are vulnerable to modification on path. All metadata should therefore have secure integrity protection (e.g., a secure message digest) to detect any modification or tampering on path.

Scalability, Continuity after handover



Router should not have additional per flow (metadata) state for each new flow with the same types of media (streams, media frames).

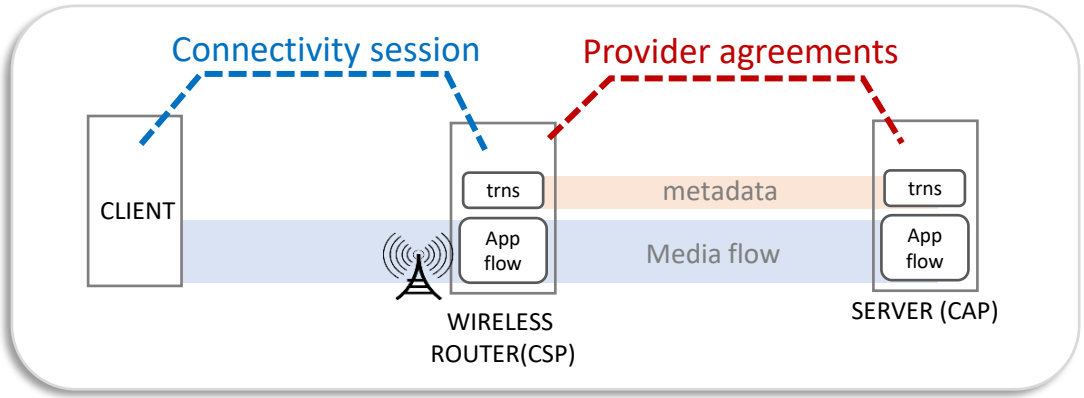
Per-flow metadata setup delay or complexity (e.g., key management) should be avoided.

Induced latency, i.e., the additional latency incurred per packet to process the metadata should be low.

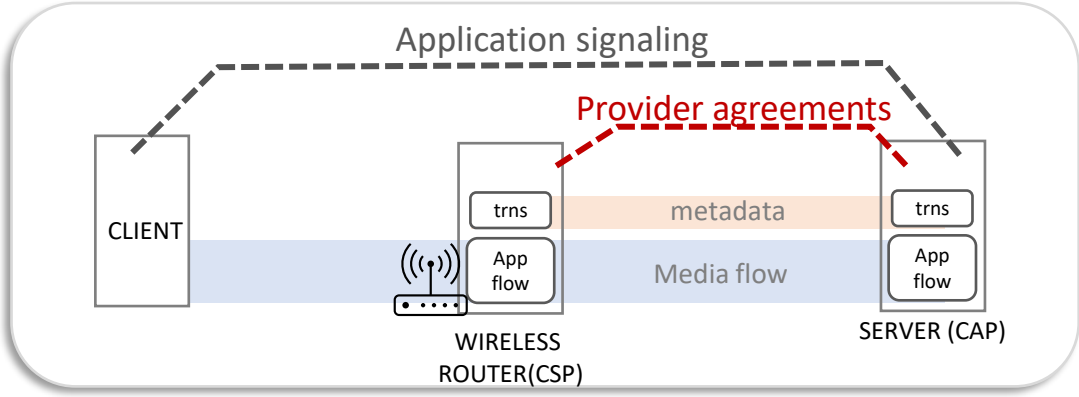
The server should not induce additional delay (e.g., to configure parameters) to provide metadata for initial packets or following handover.

> 3.7. Requirement 7: Scalable to large number of media flows
>
> There may be a large number of media flows handled by the server and wireless router, and more generally between CAP and CSP.
>
> Per flow information (state) at a wireless router for optimizing the flow can negate the advantages offered as the number of flows handled increase. The metadata other state information that a wireless router has to maintain for each additional media flow it handles should be very low or none.
>
> 3.8. Requirement 8: Continuity following user mobility
>
> The general trend in wireless networks is to distribute the wireless router closer to the user. This can help with low latency but it results in more handovers from one router to another as the user moves. The number of handovers also increase as a user moves faster or the media session lasts longer.
>
> There should no additional delay incurred during handover in configuring/setting up the metadata of a media session in progress.

Some assumptions



Scenario A: 3GPP Network



Scenario B: Wi-fi/ other Network

Is the end-user (client, host) aware?

Two scenarios by which the [host (client) – wireless router – server] are in sync.

a) 3GPP network: connectivity session requests capabilities + provider agreements (edge)
b) Other case: application signaling + provider agreements

Can the server decide what metadata is sent (incl. avoid misuse of metadata)?

The server can use multiple input points like congestion, delay and feedback information and determine how to mark the packets.

Next Steps

We request adoption by the working group as a basis for work on solutions:

- Problem has been presented at IETF 116, 117, 118
  (IETF 116: adapting for rapid resource change in wireless; IETF 117: metadata, transport details)
- Considerable discussion of solution at TSVWG mailing list
  (UDP transport/network options, IPv6 HBH, efficiency, fragmentation)
- There is interest in the group for solutions in this space.
  (26 participants indicated that this problem was worth addressing)