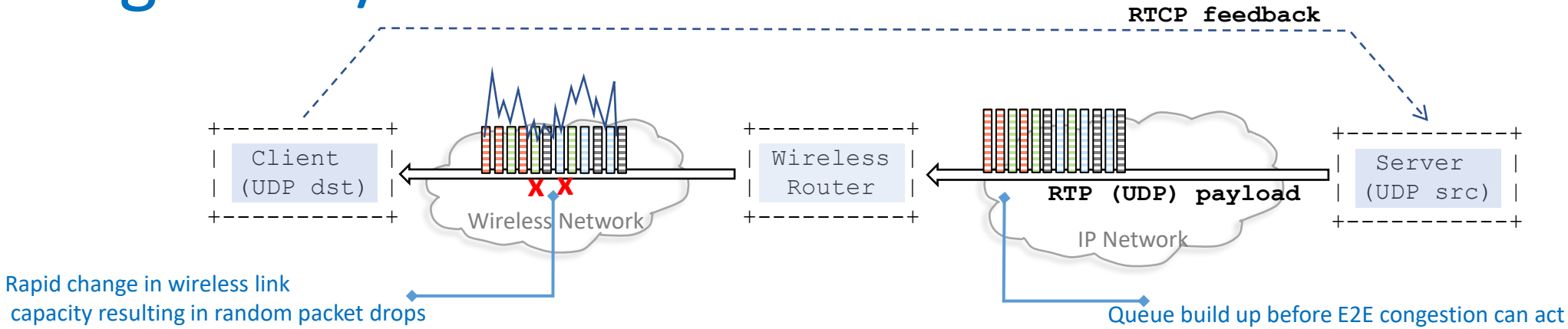
Alternative Optimizations for Low Latency Media Handling

John Kaippallimalil, Sri Gundavelli, Spencer Dawkins

Background/context - media metadata



Large transient variations in link capacity in wireless access, and variations will be even more with millimeter wave radio.

3GPP Rel 18 has specified L4S, selective packet drops for RTP. L4S/ECN feedback reacts in ~100 ms; selective drops ~1 ms.

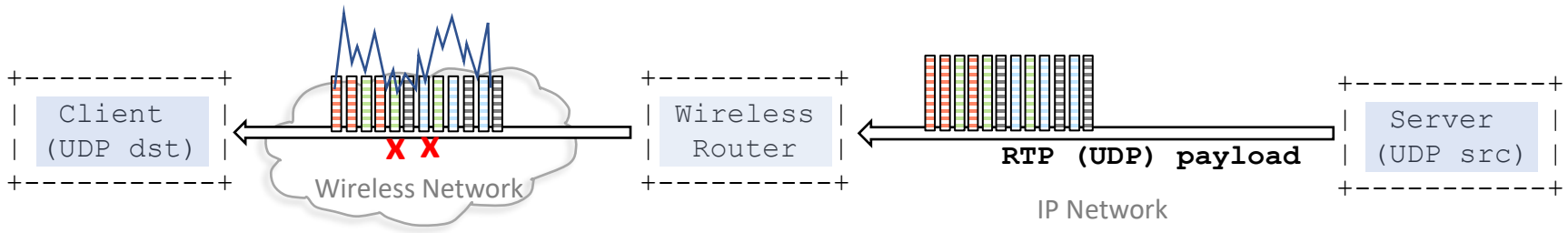
For encrypted media, media-hdr-wireless draft submitted to tsvwg proposes media metadata sent in UDP extensions (used by wireless router to classify, and wireless network optimizes shaping/scheduling and selective drops)

See tsvwg-media-hdr-wireless draft (tsvwg session on Tuesday 17:00 – 18:00 @ Continental 5)

Next 2 slides outline some related but new issues that need further discussion.

- Evolving media encoding
- Feedback to server

I. Evolving Media Encoding



Previous slide looks at metadata to help wireless networks.

However, media applications and encoding are evolving:

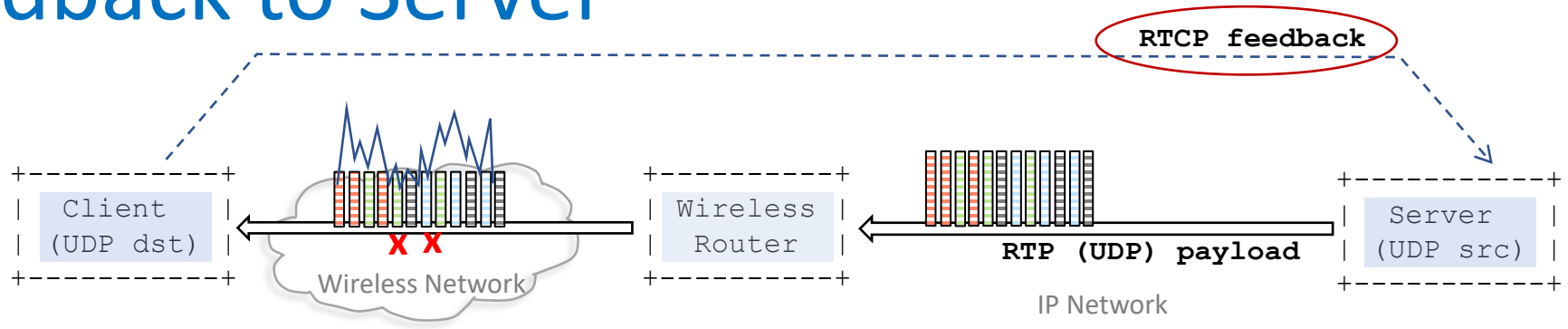
- How will AI-generated content (avatars ?), and video, audio, haptics be encoded?
  - wireless schedulers currently optimize assuming periodic handling (like video I-frames, P-frames, audio …)

- Will different encodings be sent as streams within a single transport?
  Or separate transport connections that need to be coordinated?

- How can applications (server side) provide additional information so that UDP packets on the wire have this additional metadata?

- And the UDP source maybe a streaming server, or it can be a wireless client (UE) that generates upstream content.

II. Feedback to Server



In tsvwg-media-hdr-wireless, the wireless network avoids random drops and instead drops a set of packets affect the media application less (i.e., lower priority, localize to one group of packets)

The feedback via RTCP, etc. is used by the server for pacing and adjusting the sending rate.
However, when a large number of packets are dropped, the server may reach an unexpected conclusion:

- Should the server reduce the sending rate?
  In some cases, perhaps not as capacity variation is transient and still need high throughput/max utilization
- What other impacts/behavior should be indicated in the feedback loop?