3GPP SA2 Meeting #147-e
Elbonia, October 18 – 22, 2021

S2-2107180

Time Critical Communication: Latency Critical High-Rate Adaptive services & Low Latency Low Loss Scalable Throughput (L4S)

Ericsson

© 3GPP 2020

Outline

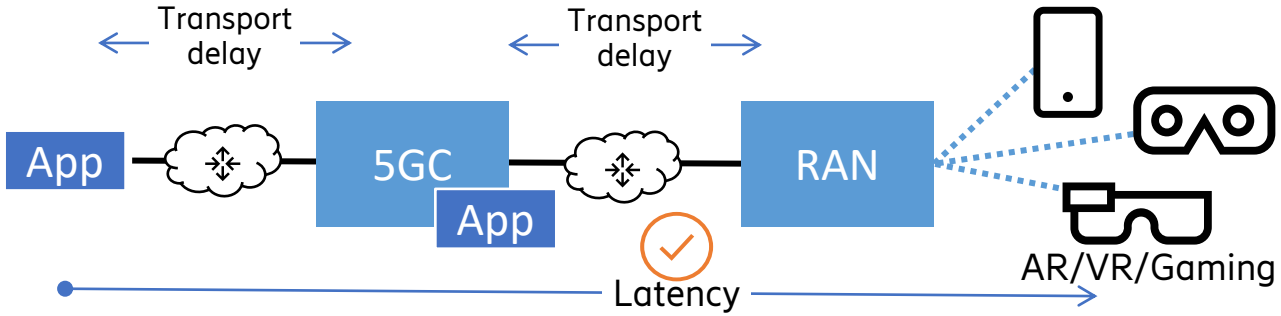
- L4S background & rate adaption principals
- L4S description
- Lab setup
- Lab results

© 3GPP 2020

Background

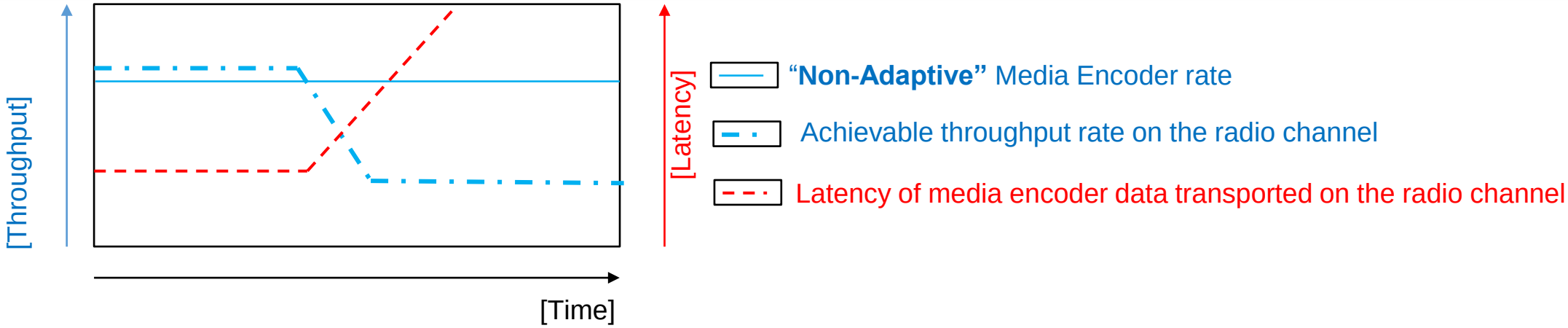
- Many 5G value-adding use cases are time-critical
- …and there's a market interest in realizing time-critical services in public networks
  - low-latency applications like AR/VR streaming and cloud gaming
  - Need tools and methods to remove sources for latency spikes, typically caused by network queue delays

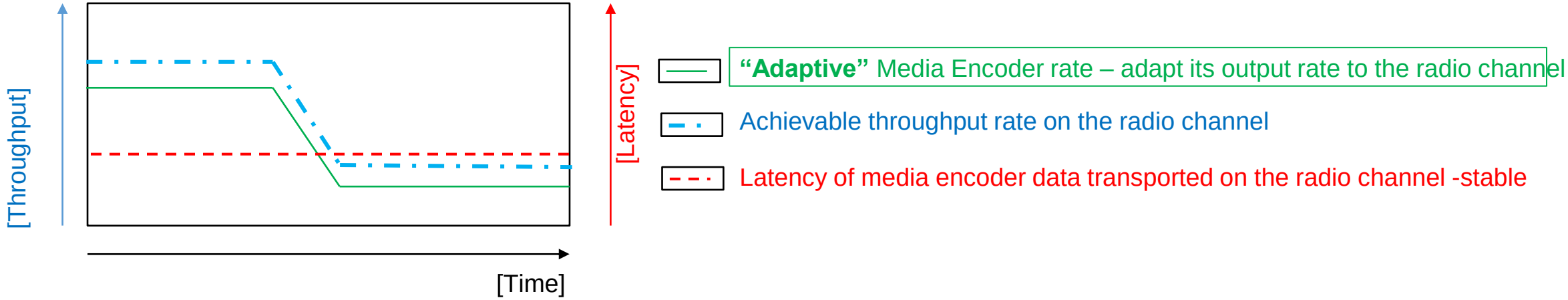
  ⇒ provide responsive rate adaptation to adapt bitrate to the available network throughput
- To secure satisfactory transport delays, assume app deployed sufficiently close to user



© 3GPP 2020

Rate Adaption and Latency Principles

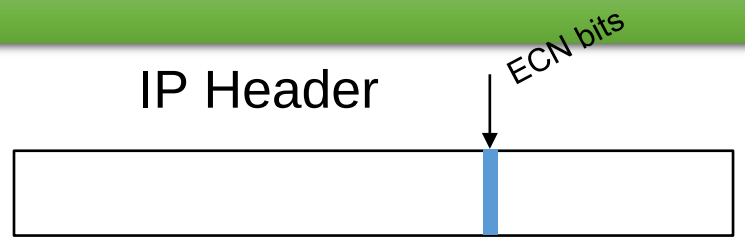




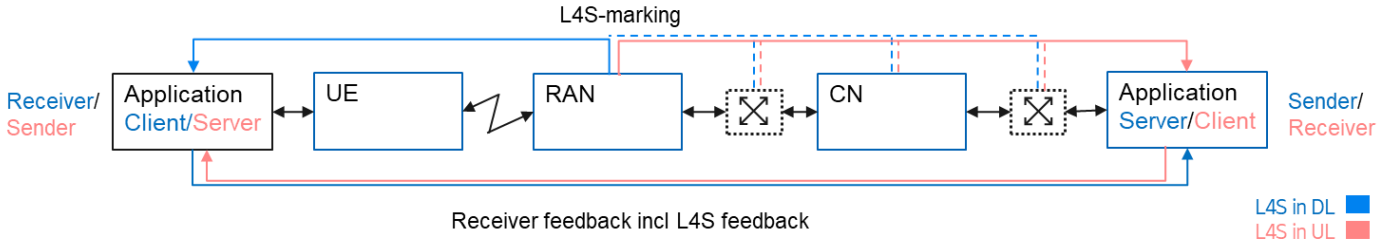
© 3GPP 2020

L4S in RAN, overview

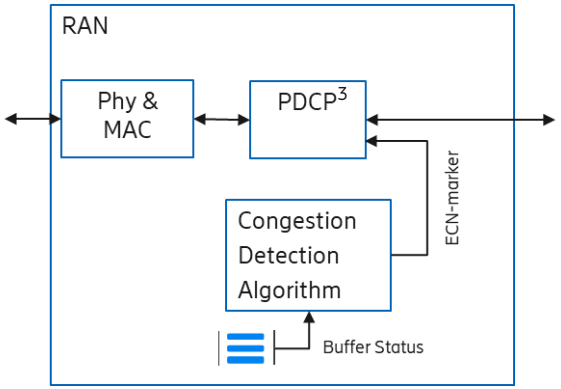
- L4S (IETF draft)
  - Congestion Marking in IP header (two ECN[1] bits)
  - Transparent through normal IP NW
  - Separate Queues for MBB and L4S traffic
  - Congestion management in any node possible

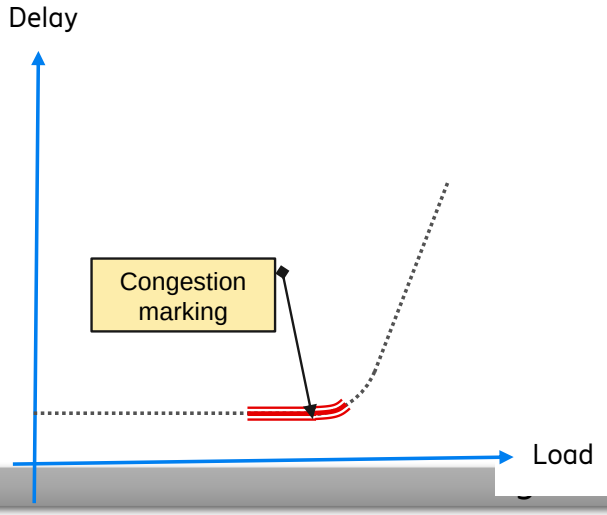


- Basic Principle
  - Congestion markings are seen on application level
  - Applications can act before congestion will result in poor latency
  - Proactive/continuous feedback to rate adaptation



- RAN implementation example
  - Dedicated bearer for latency-critical traffic
  - Congestion detector supporting CE[2] marking in IP header
  - The L4S dedicated bearer is treated according to existing 3GPP procedures (L1-L3) in the RAN such as Link Adaption, HARQ ,power, measurements etc.





1) ECN: Explicit Congestion Notification 2) CE: Congestion Experienced 3) PDCP: Packet Data Convergence Protocol

© 3GPP 2020

Dedicated bearer for Latency-critical traffic

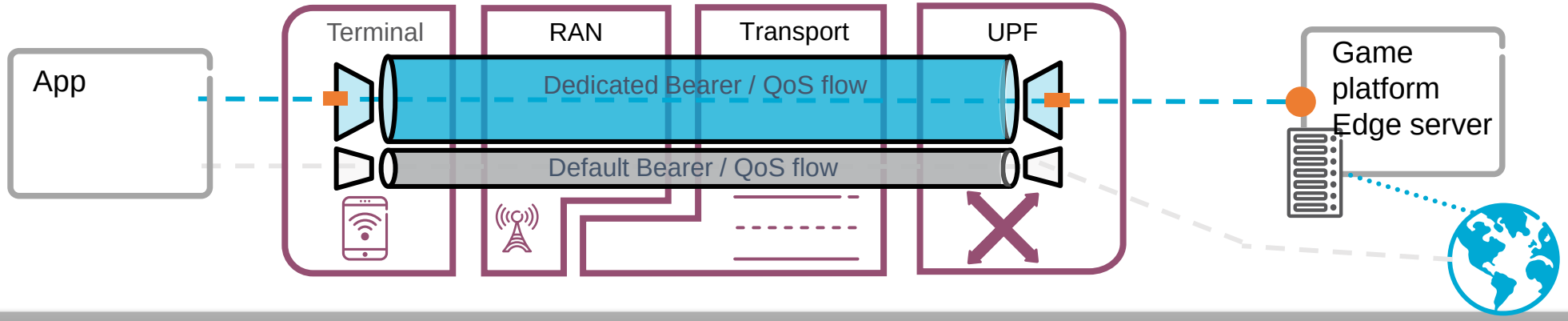
Reusing QoS/bearer features

Why a dedicated bearer/QoS Flow?

- Separate queue for latency-critical traffic (avoid disturbance from legacy traffic)
- Provide L4S feedback on this traffic
- Optional: provide QoS/priority
- Opportunity to tune other functions for latency

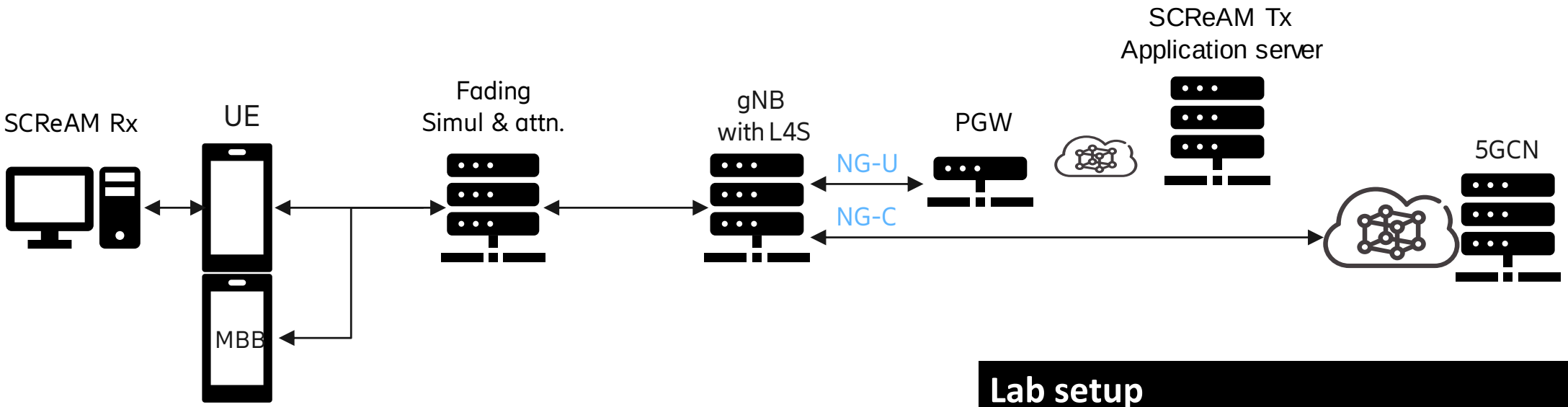
How to use dedicated bearer/QoS Flow?

- Traffic filters to map latency-critical traffic to bearer based on ECN/L4S markings
- Possibly combined with IP on Edge server



© 3GPP 2020

Proof of Concept (PoC) Lab setup



Lab setup

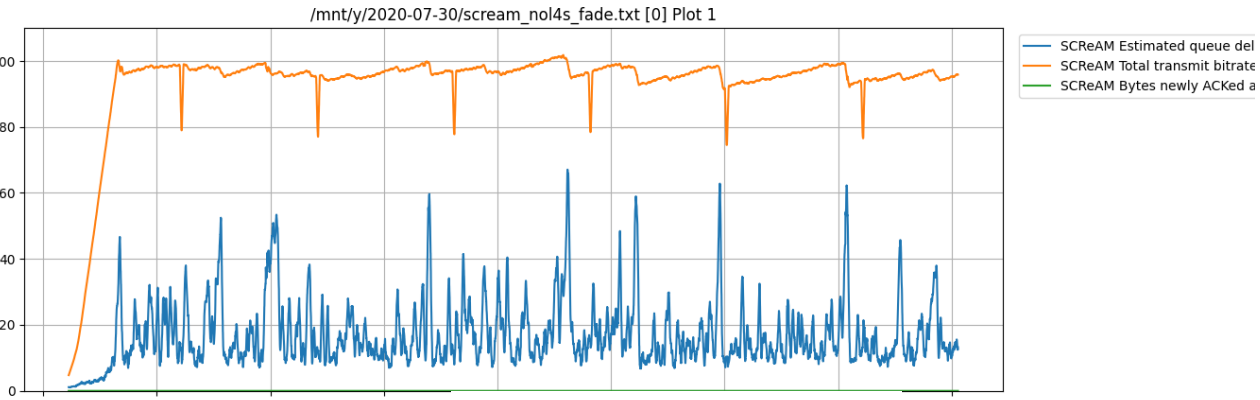
- Fading
- Programmable Attenuation
- Local GW and Application server
- Background traffic

Link to open source of SCReAM https://github.com/EricssonResearch/scream

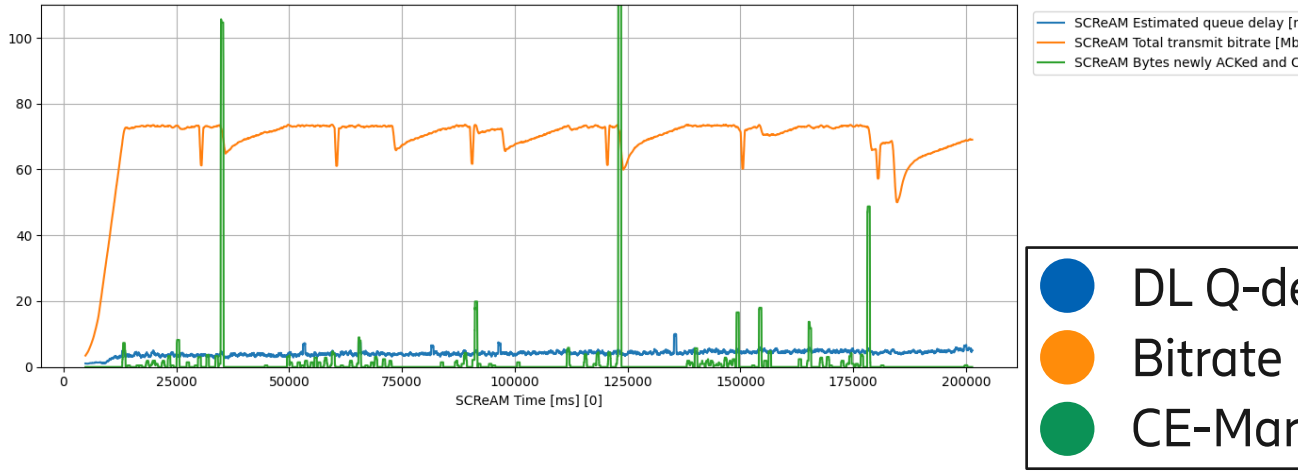
© 3GPP 2020

Results (70Hz fading and background traffic)

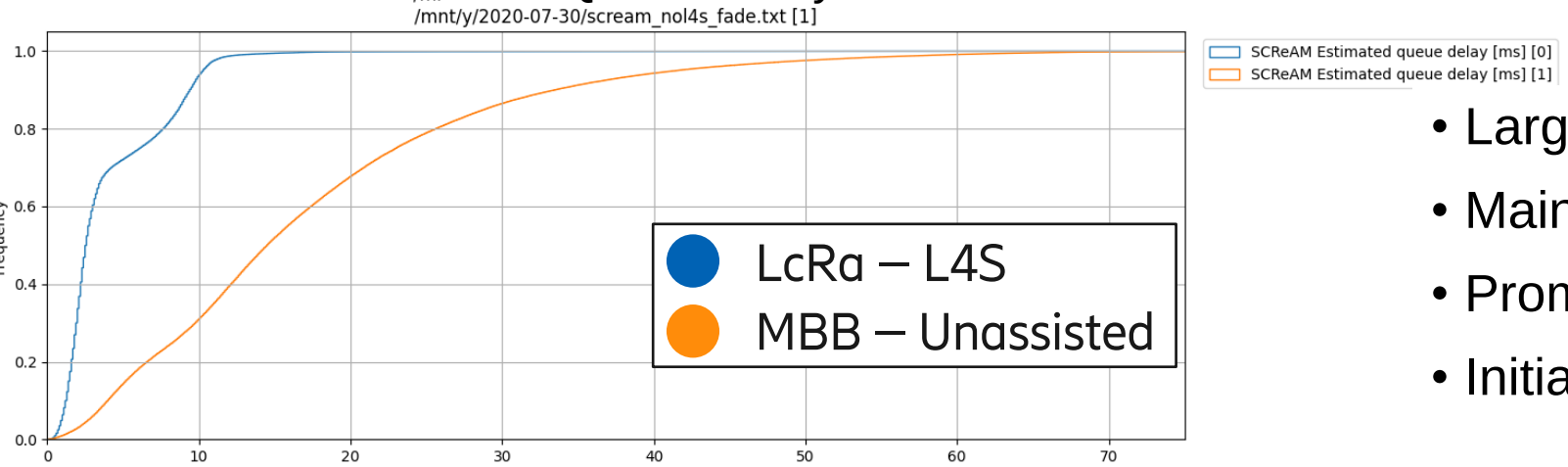
MBB (Unassisted rate adaptation)



LcRa[1] (L4S Assisted rate adaptation)



DL Queue Delay



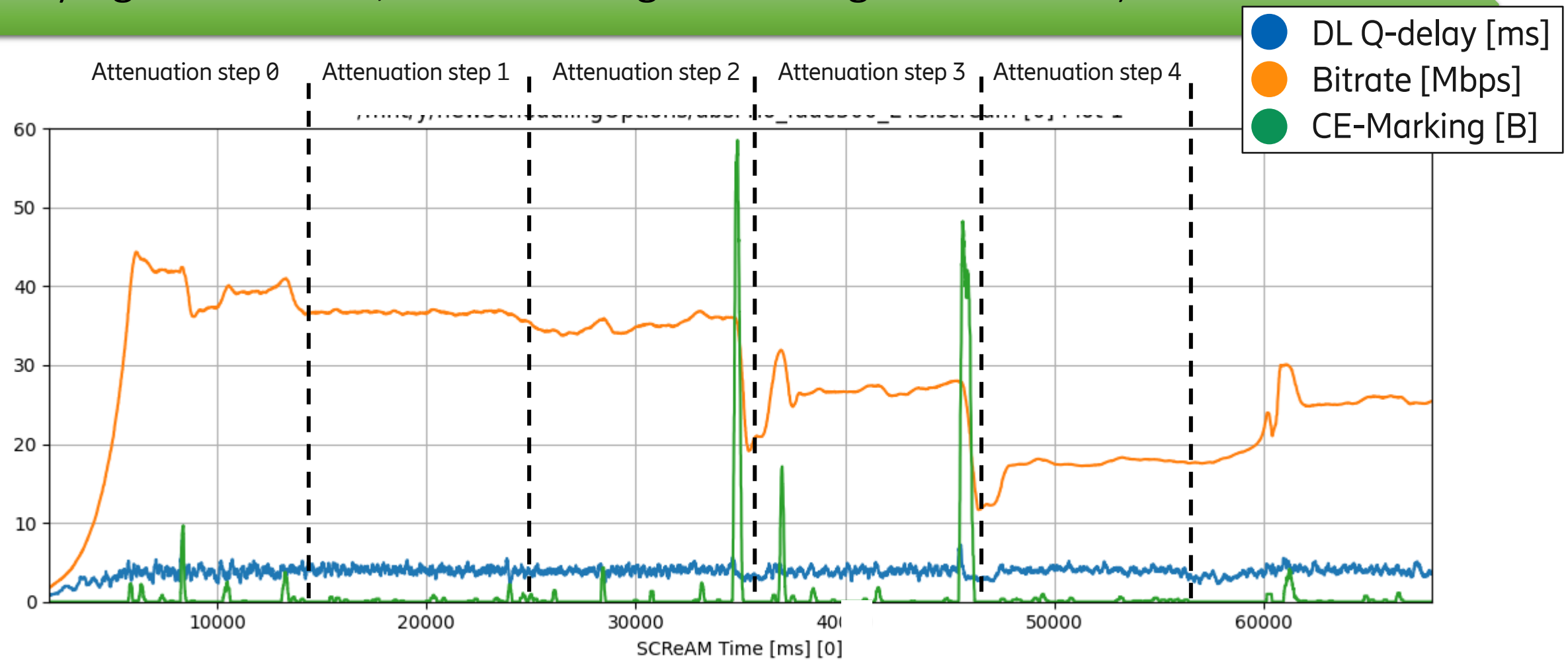
- Large improvements in DL latency
- Maintained high throughput
- Promising results from L4S lab tests
- Initial field tests confirm lab results

1) LcRa: Latency-Critical High-Rate Adaptive services

© 3GPP 2020

Some Additional L4S lab results

(Varying attenuation, 300Hz fading and background traffic)

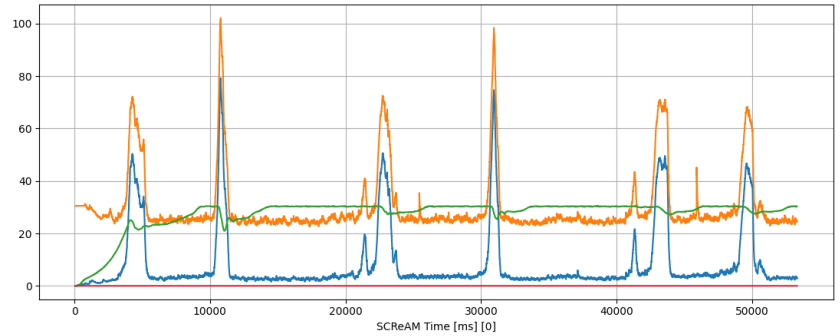


© 3GPP 2020

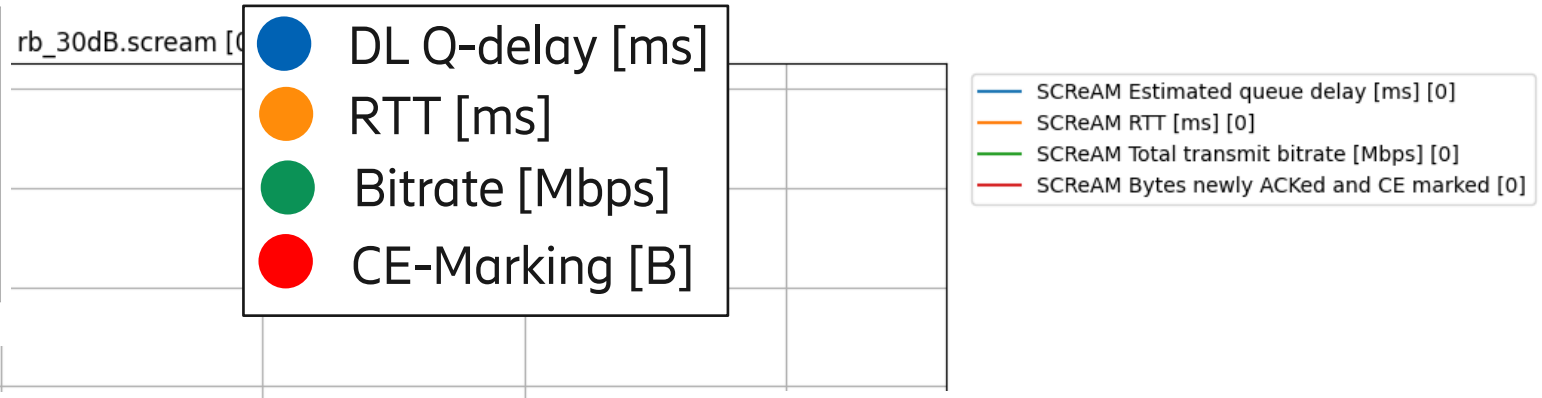
”City Busride” Testcase Results

Ramped drop (1s, 20dB), second 5dB attenuation, bursty background traffic

MBB (Unassisted rate adaptation)

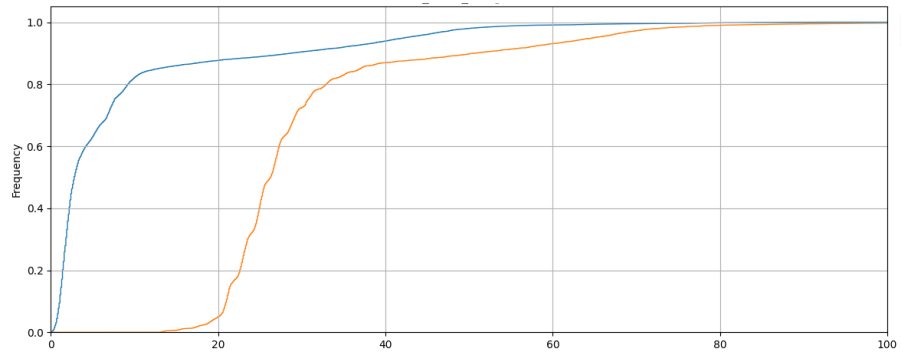


LcRa (L4S Assisted rate adaptation)



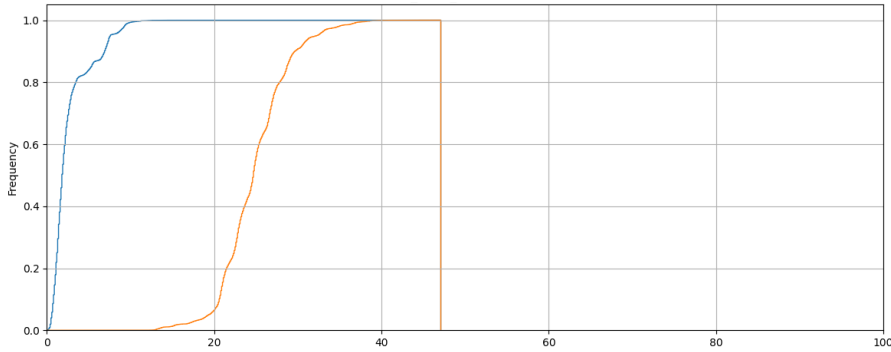
- DL Q-delay [ms]
- RTT [ms]
- Bitrate [Mbps]
- CE-Marking [B]

MBB (Unassisted rate adaptation) CDF



- DL Q-delay [ms]
- RTT [ms]

LcRa (L4S Assisted rate adaptation) CDF



Ericsson

© 3GPP 2020

Summary and proposal

- L4S has proven to be an enabler for high-rate adaptive services
- L4S can easily be introduced in 3GPP directly without changing architecture principles (see slide 6)
- Ericsson proposes to introduce L4S in Rel-18

© 3GPP 2020