Low-Latency Communications and the Internet Architecture

Mobile Networks, Analytics, and Edge Computing Workshop,
Technical University of Munich

Jari Arkko
September 2017

Based on draft-arkko-arch-low-latency (co-author: Jeff Tantsura)
Goals

• Trying to understand what the recent interest on low-latency / critical communications applications implies for the Internet

• New work for improving latency? Architectural evolution?

• Work in progress; please contribute!
Recent General Requirements for Low-Latency Communications

- Tactile Internet requires 1 ms reaction time
- Self-driving cars require 1ms latency
- Mission-critical 5G MTC requires low latency & high reliability and availability
Plenty of Wild Claims in This Space

Can we now use circuit switching?

Low-latency applications will pay for everything.

Need to build a new protocol!

Finally a reason to use QoS.

All your server businesses are belong to us.

Cloud is > IMS away.

Leads to ideas...

This comic strip was created at MakeBeliefsComix.com. Go there to make one yourself!
Does The World Care about Low-Latency?

- Data centers distributed around the globe
- Including content served from operator premises
- Advanced optimisation techniques for connecting to data centers (DNS etc)
- Industry working HTTP2, QUIC, TLS.1 (0-RTT), L4S, DETNET, 802.1 TSN, 5G radios, …
- SDN and SFC can optimise long chains of processing functions
- Industry working on ServiceWorker, AMP, …

Clearly caring, and in some cases paying a price for low latency
Let's Recap To Be Clear

- Latency in L2 is being improved
- Latency in routing/forwarding is being improved
- Latency in transport is being improved
- Latency in security is being improved
- Latency in application protocols is being improved
- Network deployments are changing to take into account latency
All Done? Or Work Ahead?

• Not necessarily the big revolution some might claim; a lot of the tools are there

• Obviously much of this is work in progress

• But, the Internet may be changing and this would cause strain for the architecture
Architectural Thoughts 1

• Need to consider the system whenever thinking about this topic

• Trend of service placement in different locations: from global datacenters to more regional ones, cooperative solutions, edge computing

• Impacts on architectures that employ tunnelling

• There are and will be demands on cross-layer optimisation, is that a good thing?
Architectural Thoughts 2

• Tension between local networks (e.g., cars braking and informing nearby cars) and Internet-networking

• Designing applications entirely in their own silo vs. applications that also talk to peers in the Internet

• Tension between application/edge and network control of forwarding decisions (MPTCP vs. routing)

• Deployment story for new QoS or low-latency tech

• Findings from claffy & Clark: https://www.caida.org/publications/papers/2015/
Opportunities

• There is willingness to change & deploy new tech

• Virtualization technology allows us to move elements more freely than before

• And there is willingness to deploy resources closer to users on a global scale (perhaps even to edge computing)

• We have a more realistic ability to use multi-path communications than perhaps before
Thank you