Quality of Outcome — From Complex Latency Measures to Application Outcome

DOMOS

Users consume applications, not networks (and certainly not latency)

Agenda

1. What's so difficult about network quality and application performance
    a. Measurements
    b. Aggregation
    c. Application design
    d. Users
2. Quality of Outcome
    a. Design Goals
        i. Useful for:
            1. End-Users
            2. Network Operators and Vendors
            3. Applications
    b. Proposed Solution
    c. The foundation
    d. The middleground between QoS and QoE
    e. The formula
3. Some results

DOMOS

What's so difficult about going from latency measurements to application outcome?

Measurements

How to measure network quality?

- Bandwidth
- Latency
  - Idle
  - Under Load/working latency
  - 50th/90th/99th... percentile latency
- Packet loss

- Active probing
  - Speedtests
  - Ping (ICMP)
  - TWAMP / STOPM
  - RRUL

- Passive
  - Monitoring load (bandwidth utilization)
  - Monitoring latency (SYN ACKs, DNS lookups etc.)

The fundamental difficulty

Network Quality changes frequently and radically

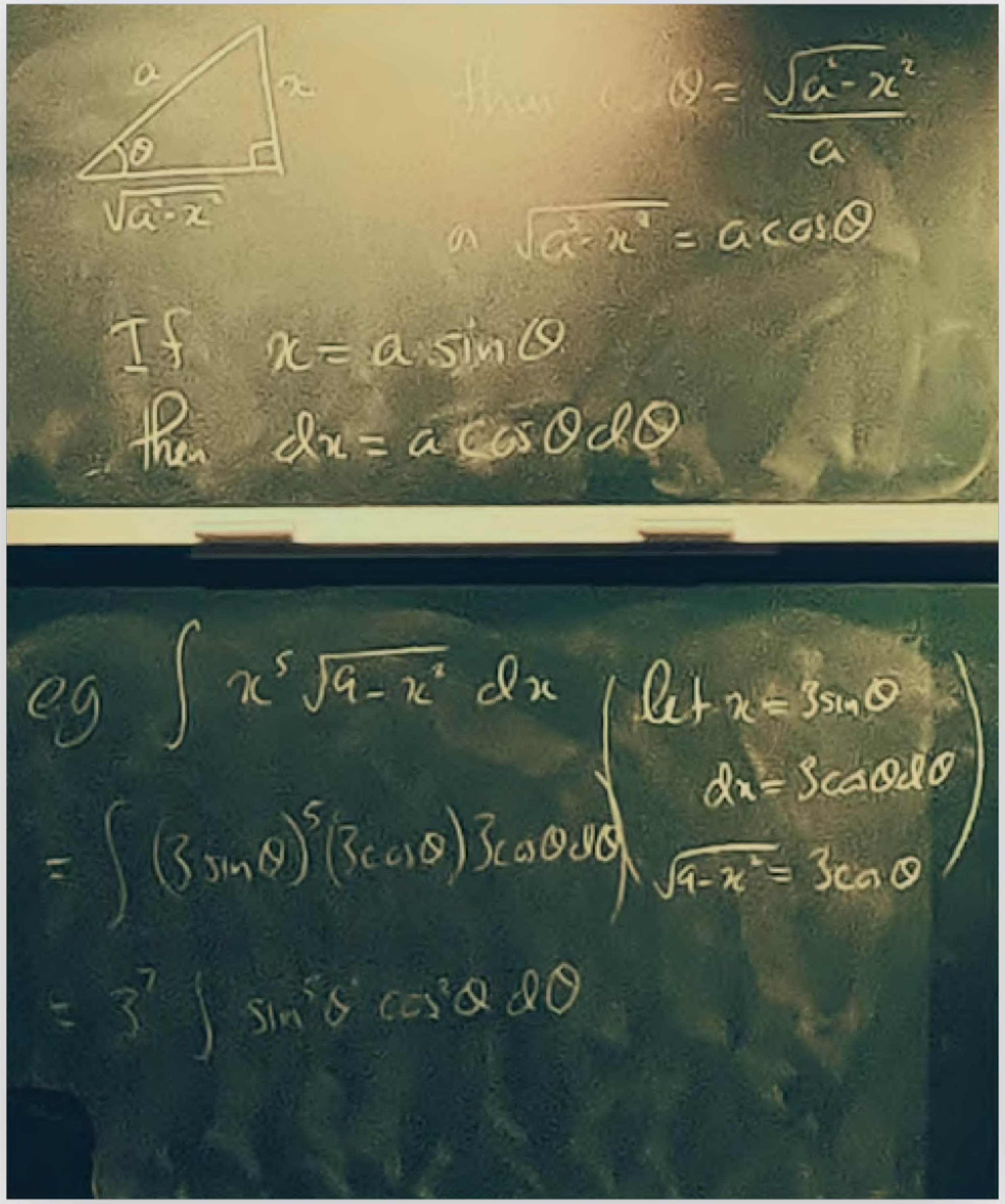
Aggregation

- Time is complex
  - There can be thousands of packets per second
  - Bits per millisecond, megabits per second, average megabits per second are all very different

- Statistics
  - Applications can affected by the peaks, the average, and by the type of variation
  - A network quality metric that needs all samples will not be feasible to use

- End-to-end vs for a segment ?
  - How can you meaningfully aggregate segments?

- Perspective matters
  - From within the network or from the outside
  - Networks are complex
  - Multi-technology, multi-operator
  - Not all technology measures bandwidth/latency the same way
  - Not all operators test bandwidth/latency the same way



Applications have built techniques to counteract the rapid changes of network quality

- Multiple layers of optimization and decades of techniques
  - TCP (QUIC)
  - UDP - WebRTC

- App layer adjustments
  - Resolution
  - Codecs
  - Packets per second
  - Jitter buffers
  - Bit rate adaption

- Application Outcome depends on the App and App type
  - For gaming lower latency is always better
  - For Video Conference, it always keeps a jitter buffer, so below a point it makes no difference
  - For downloads latency matters very little (in most cases)

User Experience — Let's not go there

- Two peoples experience of the same video conference may be different.
- Your mood may affect you experience
- Your expectations may affect your experience

It is just not measurable in a scalable way

While User Experience is not measurable (at scale)

Application Outcomes are (and will strongly correlate with user experience)

Quality of Outcome

Creating a standardized middle ground

Quality of Experience

MOS, NPS, gMOS, ...

Relatable, unreliable

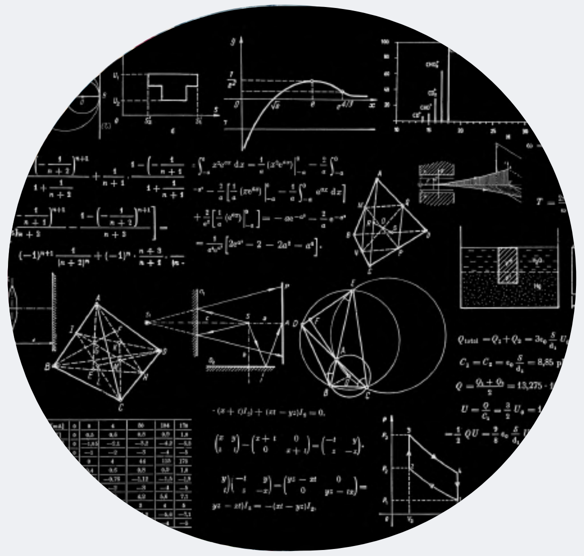
Quality of Outcome

What is the likelihood of perfect video conference?

Relatable, reliable

Quality of Service

Bandwidth, latency (in many forms), packet loss



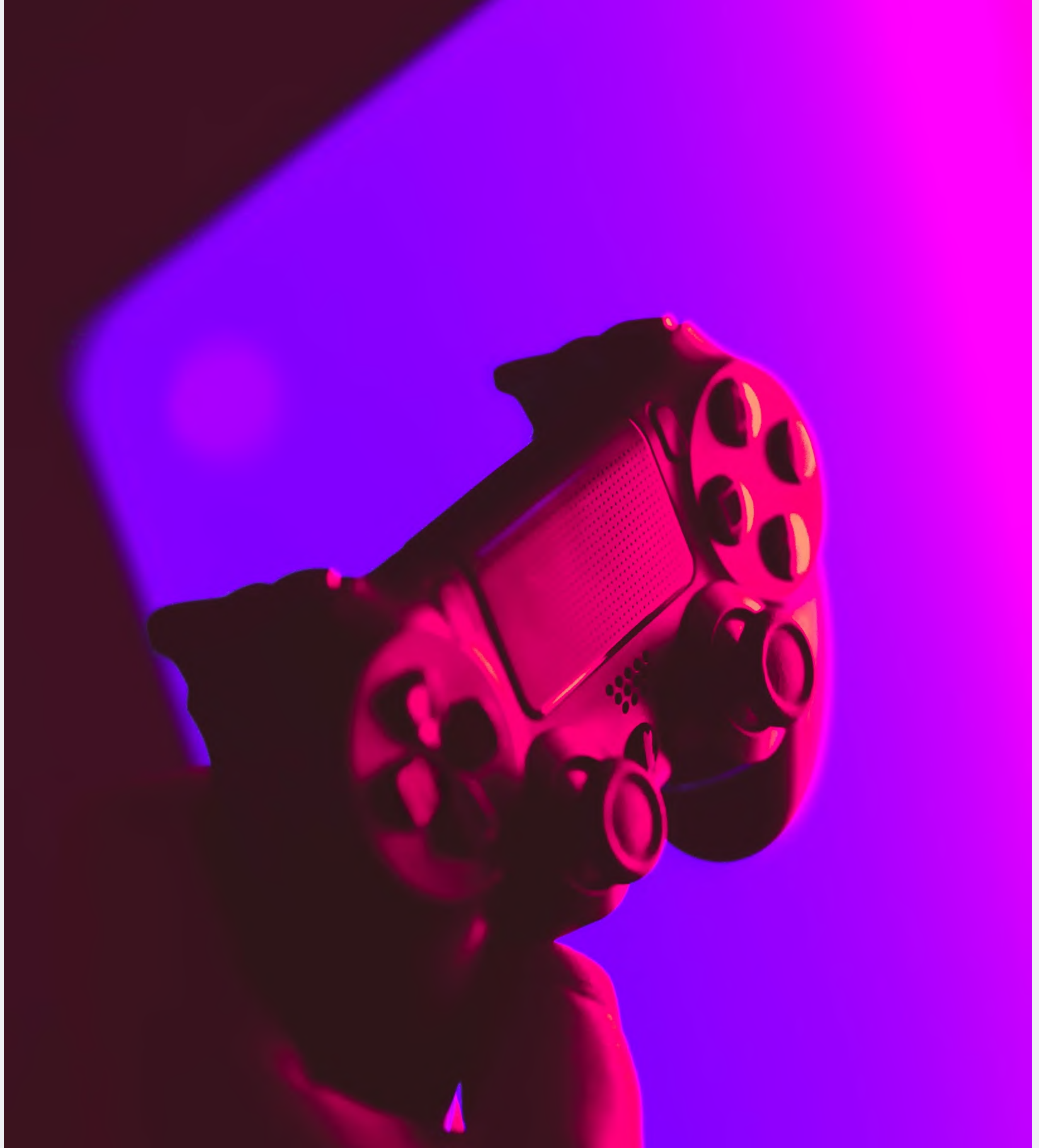
Unrelatable, reliable

Design Goals

How can we make it useful for applications, users and operators?

For the end-users

- Understandable
- Maps to application outcome
- Honest and objective



For the applications

- Objective
- Can be used to understand if the application will work as intended
- Can describe complex network conditions
- Can map both very simple and complex network requirements
- Probabilistic

For the network operators and vendors

- Objective
- Strong mathematical foundation
- Composability
  - e.g., Measure LAN and WAN separately. Add them together
- Useful throughout the life cycle
  - Lab Testing
  - Active Testing
  - Monitoring

Proposed Solution

Calculating the probability of meeting network requirements

With a defined way of measuring the network and aggregating and compose it

https://datatracker.ietf.org/doc/draft-teigen-ippm-app-quality-metric-reqs/

Working with the aggregation and measurement complexities

Building on top of TR-452 (QED)

- Mathematical framework for network quality
- Network Quality is how latency distributes at different loads
- Captures jitter, peaks, packet loss..
- Composable
- Useful throughout the life cycle
- Can describe complex networks and requirements


broadband forum

Technical Report

TR-452.1

Quality Attenuation Measurement Architecture and Requirements

Revision 1

Date: September 2020

Allowing complex app requirements

But take it back to a probability

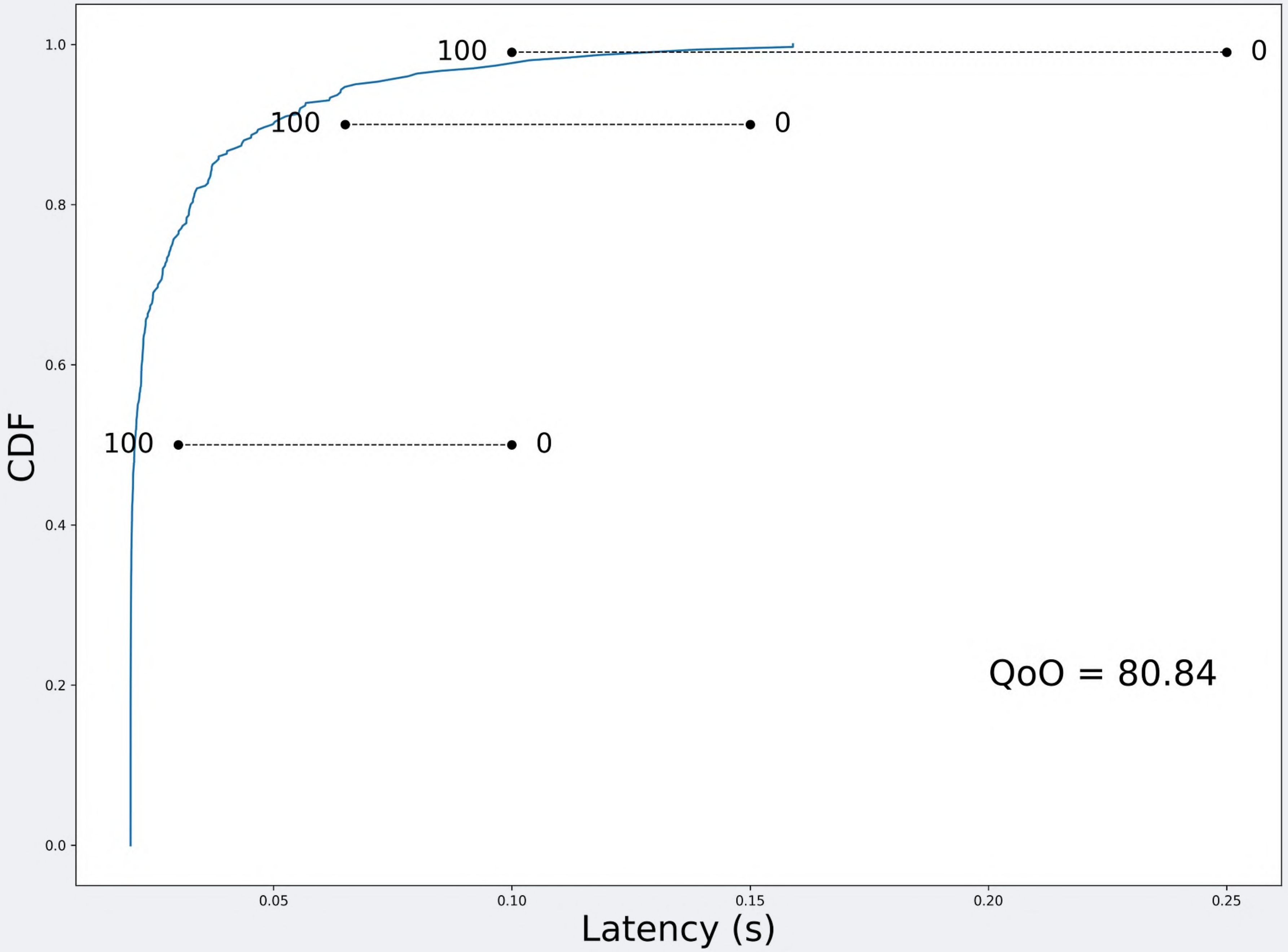
The formula visualized - How likely is the network to reach the network requirements?

- Blue line is measured latency(ML), shown as a cumulative distribution function (CDF) for some time-period. Note that the QoO framework is indifferent to the length of the time-period, a key advantage of using percentiles and distributions

- The dots marked with '100' are the Network Requirements for Perfection (NRP). Those marked with '0' are the Network Requirements points of uselessness (NRPoU)

- In this example, there is set up network requirements at 3 percentiles (50th, 90th and 99th). You can setup as many as you like

- The ML breaches the 99th percentile requirement. At a point that is roughly 80% closer to NRP than the NRPoU. Hence the QoO is 80 (the worst breach is always used.



5. Calculating Quality of Outcome (QoO)

At this point we have everything we need to calculate the quality of the application outcome. The QoO. There are 3 scenarios:

1. The network meets all the requirements for perfection. There is a 100% chance that the application is not lagging because of the network

2. The network does meet one of the criteria of uselessness, including bandwidth. There is a 0% chance that the application will work because of the network

3. The network does not meet NRP but is not beyond NRPoU.

1 and 2 require nothing more from the framework. For 3, we will now specify the calculation between to translate these distances to a 0 to 100 measure. We use the percentile pair where the measured latency is the closest to the NRPoU as the application is only as good as its weakest link.

Mathematically: QoO = min(ML, NRP, NRPoU) = (1-(ML-NRP)/(NRPoU-NRP)) * 100)

Essentially, where on the relative distance between Network Requirement for Perfection (NRP) and Network Requirement Point of Uselessness (NRPoU) the Measured Latency (ML) lands, normalized to a percentage.

The formula

- Measure latency distributions in any way you like (but note how, so others can evaluate statistical confidence)

- Applications requirements or SLAs:
  - Is described with *network requirement for perfection:*
    - e.g., *Video Conference needs 5Mbps, and 90% of packets to arrive within 100ms and 99% within 150ms*
  - And *network requirement points of uselessness:*
    - e.g., *If 10% of packets to arrive after 400ms or 1% after 500ms*

- The distance between is normalized, and we do assume linearity in degradation, and call it a probability

- We go from complex network statistics to probability of reaching outcomes

How do people like QoO?

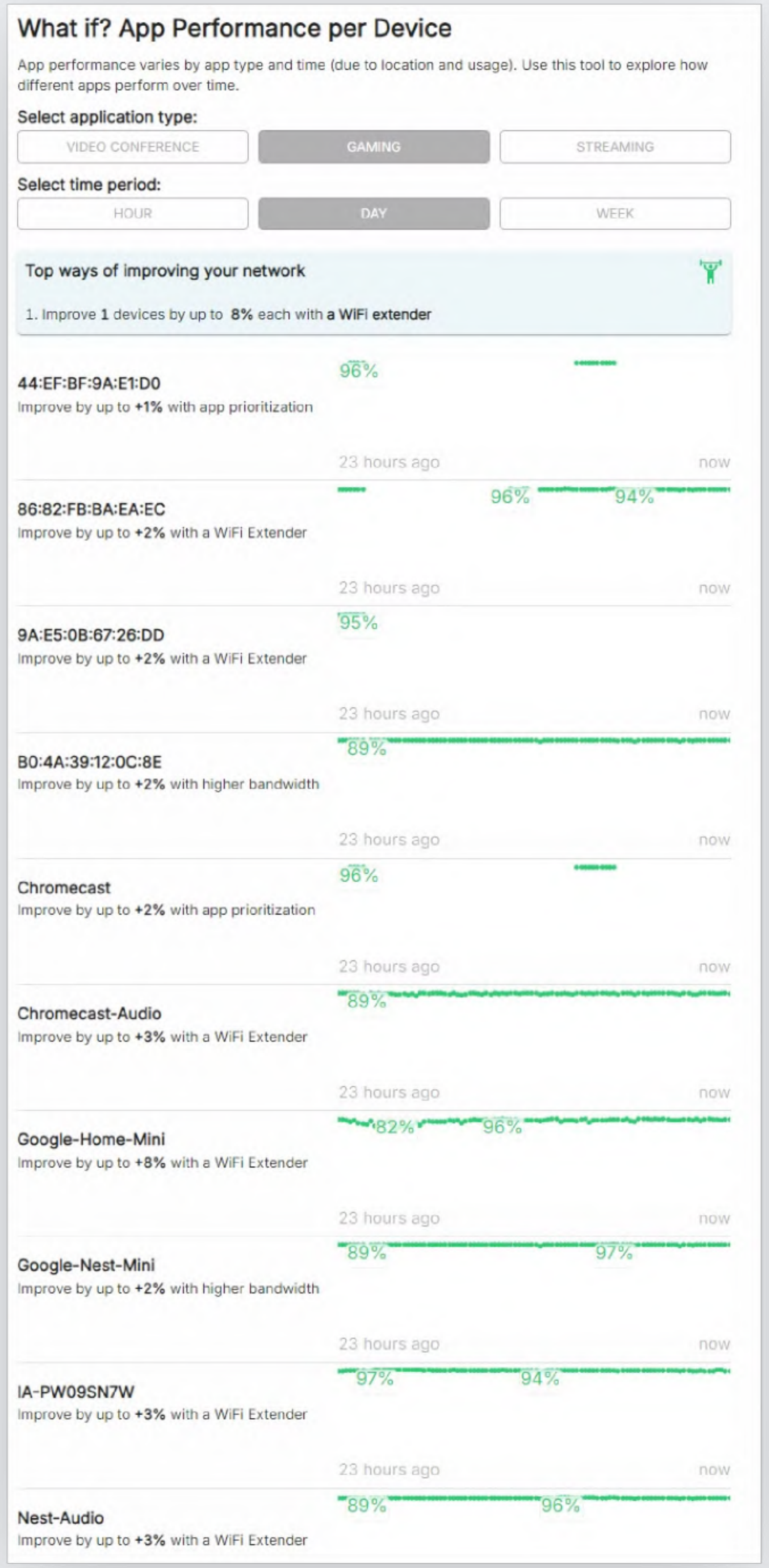
Comcast Innovation Fund User Trial

25 routers were sent to volunteer Comcast customers, with adjustments made to make it a representative population. The goals:

1) Measuring network quality as app performance in the QoO Framework

2) Giving participants access to a web tool showing them network quality and recommendations in the QoO framework (as app performance) and giving them a questionnaire on how they liked it

How did people like it?



People like QoO better

86% | Says QoO more **accurately** reflects experience than their bandwidth
14% neutral - 0% negative

100% | Says the tool that showing network quality in QoO is **better** than speedtest
and latency and jitter

86% | Says it is somewhat, very or extremely **important** that QoO is independent

86% | Would be likely or very likely to **recommend** the QoO tool
to a friend or colleague

QoO as a selling point in itself

53%

Says they would *rather* buy app guarantees than bandwidth - 40% unsure - only 7% no

80%

Says they would rather buy from an ISP with app performance guarantees in addition to bandwidth - 20% unsure, 0% no

Thank you for Listening!

Call to Action: Get in touch and contribute!

magnus@domos.ai
bjorn@domos.ai