

A Single Common Metric to Characterize Varying Packet Delay

Bob Briscoe, Independent <ietf@bobbriscoe.net>
Koen De Schepper, **NOKIA** Bell Labs <koen.de_schepper@nokia.com>
Greg White, **CableLabs**[®] <g.white@CableLabs.com>
Vidhi Goel, Apple <Vidhi_Goel@Apple.com>

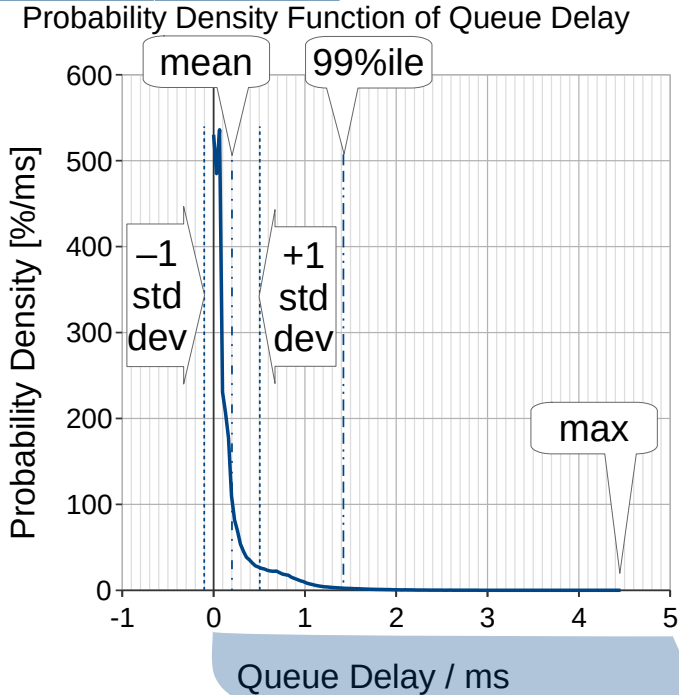


IAB Workshop on Measuring Network Quality for End-Users
Sep 2021

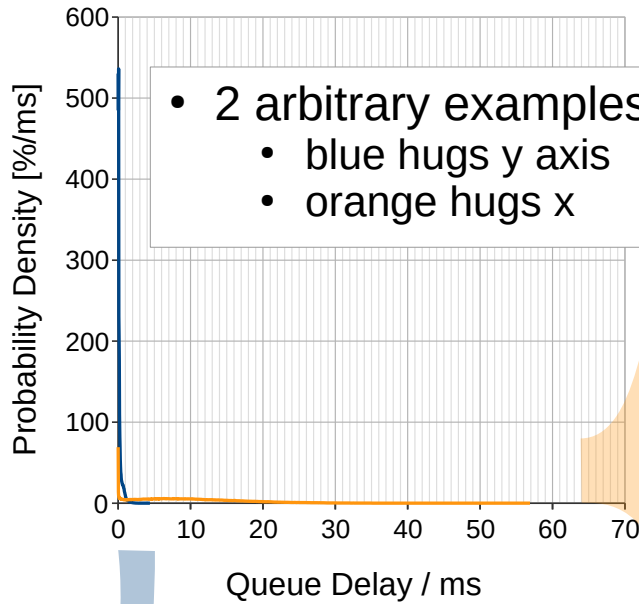
Which metric best characterizes the experience of varying packet delay?

mean	0.2 ms
std dev	0.3 ms
median	0.07 ms
99%ile	1.4 ms
max	4.5 ms

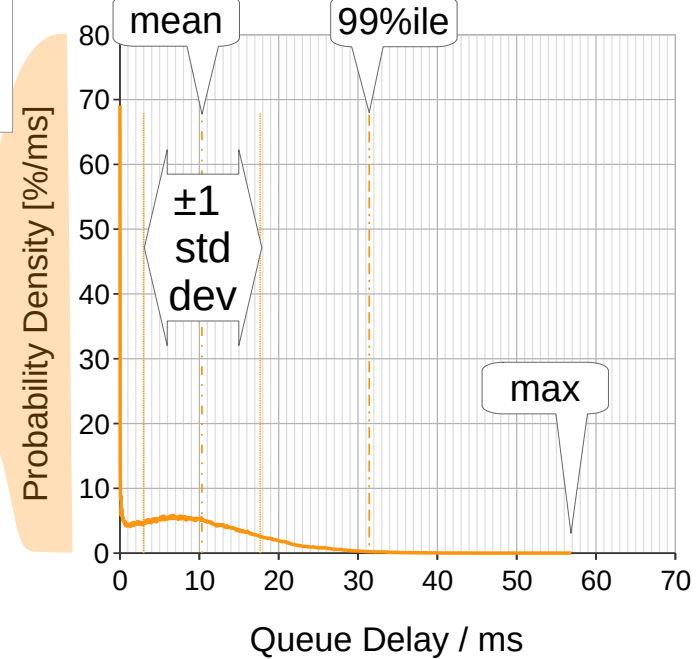
mean	10 ms
std dev	7 ms
median	9 ms
99%ile	31 ms
max	57 ms



Probability Density Function of Queue Delay

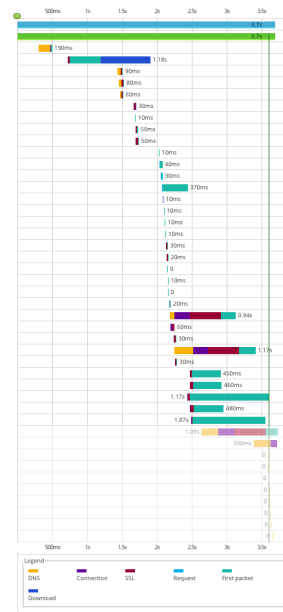


Probability Density Function of Queue Delay



mean or median are distractions

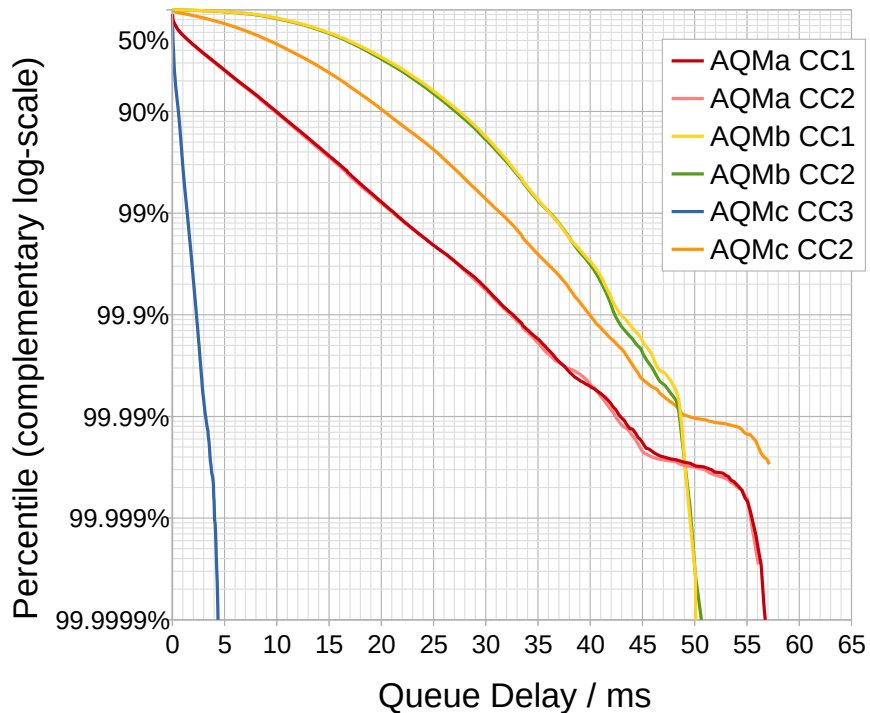
- Real time
 - play-out after median delay would discard 50% of packets
- TCP short flows (e.g. RPC, web)
 - wait for straggler packets to deliver to app in order
- Multiple objects / streams (e.g. http2, quic, webrtc)
 - even if no protocol sequence, typically inter-object dependencies in the app logic
- **Generalization (mostly true):**
 - the user experiences the delay delivering the evolving assembled product, not the pieces
- Nearly all packet delay distributions are asymmetric with a long-tail
 - mean, median, standard deviation, etc. all characterize the irrelevant body, not the tail
- **Proposal: Standardize at least one high percentile to enable comparisons...**



Which high percentile?

Queue Delay CCDFs

120 Mb/s link rate, 10 ms base RTT



- Not too high

- otherwise too slow to calculate accurately
- but high enough to reflect typical delay experience

- Strawman: **99%ile**

- one imperfect number better than many different perfect ones
- can most apps conceal 1% discard well?

- IETF (ippm): appropriate body to forge consensus

- anyone interested? arguments against?

Clarifications

- Not saying won't need to specify what, where and how as well
 - 1-way/2-way; layer7/4; at queue/e2e; capacity; RTT; load pattern; etc
 - that's for each scenario, whereas the present question is for *all* scenarios
- Not saying you shouldn't specify other percentiles as well or ideally a whole log-scale CCDF
 - as long as we have one common metric – for comparisons

A Single Common Metric to Characterize Varying Packet Delay

Thank you

Q&A