

# L4S Operational Guidance Draft

Greg White, Editor  
TSVWG Session @ IETF108  
July 30, 2020

# Coexistence of L4S & Classic Traffic

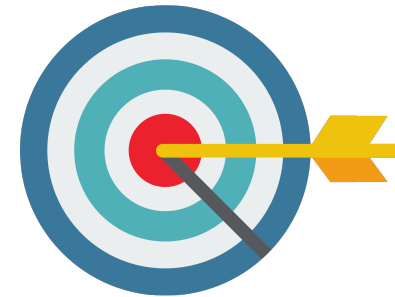
Network Bottleneck		Prevalence	Fairness Result	
AQM	Sched.			
L4S ECN	DualQ	<i>Future</i>	DualQ-Coupled-AQM provides reasonable fairness across a range of conditions	✓
L4S ECN	FQ	<i>Future</i>	FQ enforces fairness	✓
Classic ECN	FQ	Common	FQ enforces fairness	✓
Classic ECN	FIFO	?	<b>Problem: Classic flows get less throughput than L4S flows</b>	<b>X</b>
Non-ECN	any	Common	L4S flows become Reno-friendly in response to loss	✓
Taildrop	any	Very common	L4S flows become Reno-friendly in response to loss	✓

Classic flow/ECN = RFC3168

# Why this draft?



- Provide Operational Guidance for the L4S experiment
  - Motivation: Ensure reasonable fairness between L4S flows & Classic flows sharing a bottleneck
  - Informational Draft identifying opportunities to prevent and/or detect and resolve fairness problems
    - Minimize likelihood of introducing fairness problems
    - Facilitate resolution of any issues
  - Scope of IETF Guidance
    - Proactive: preventing potential issues
    - Real-time Reactive: detecting and resolving issues in real-time
    - Reactive: address any problems via management intervention
      - Assumes L4S is safe enough, given the unfairness does not result in starvation

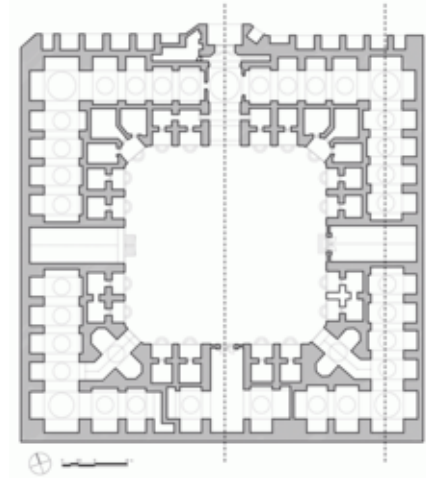


# Target Audience / Outline

- Operator of an L4S Host
  - CDN servers
  - Other hosts
- Operator of a network
  - Options for existing Classic ECN FIFO Bottlenecks
- Researchers
  - Detection of Classic ECN FIFO bottlenecks
  - End-to-end measurement of L4S vs. Classic performance

# Plan forward

- Draft-00 to be posted soon
  - Including contributions from Bob Briscoe, Jake Holland, Koen De Schepper, Olivier Tilmans, Tom Henderson, Asad Ahmed and others on the TSVWG mailing list.
  - Please review and provide comments/contributions via mailing list
- Interim meeting?



# Backup Material

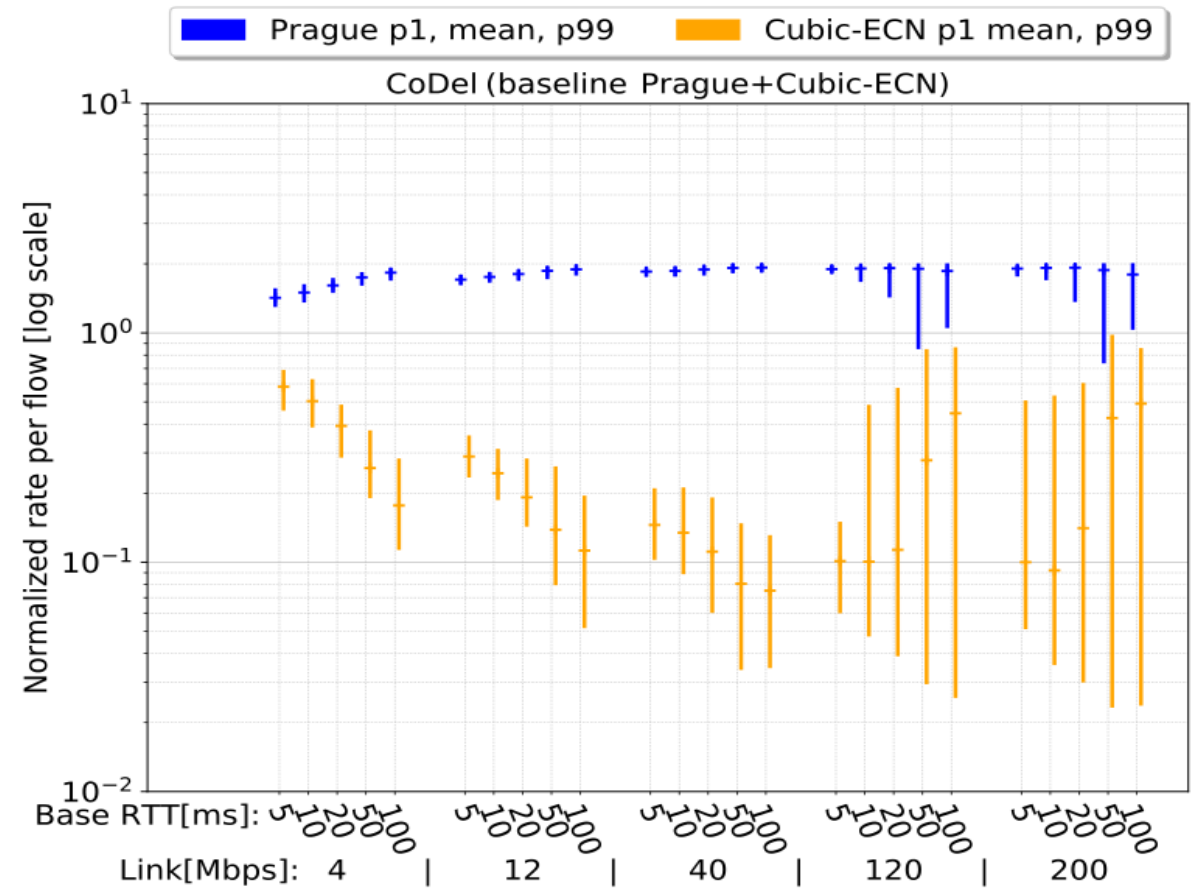
# Mailing List Summary from Jake Holland

1. A robust classic bottleneck detection mechanism
2. Changing L4S to use a 2-signal approach, using ECT(1)->ECT(0) for the  $1/p$  signal and ECT(1|0)->CE as a  $1/\sqrt{p}$  signal.
3. A flag day to deprecate ECT(1)->CE marking by classic queues (instead treating ECT(1) as NECT if no non-3168 meaning is implemented).
4. Operational considerations to recommend changing ECT(1) to NECT at ingress to networks that have marking classic queues deployed
5. Operational considerations to recommend policing strategies that can solve the general case of non-compliant traffic that does not respond with the expected backoff to AQM congestion signaling.
6. An experiment-linked public whitelist of participant-registered IP ranges that have a L4S compatible dualq in their reachability path at the likely bottleneck, which would be checked by endpoints before negotiating L4S.

[https://mailarchive.ietf.org/arch/msg/tsvwg/yPX9QCitf2g\\_Kf2vGijl1mLMzhM/](https://mailarchive.ietf.org/arch/msg/tsvwg/yPX9QCitf2g_Kf2vGijl1mLMzhM/)

# Classic ECN FIFO Fairness

- RFC3168 does not differentiate between ECT0 (Classic) and ECT1 (L4S)
  - Existing Classic ECN FIFOs presumably treat both identically
- L4S & Classic senders respond differently to CE marks
  - L4S senders expect frequent CE marking
    - E.g. 2 CE-marks per RTT
  - Classic senders expect very rare marking
    - E.g. 1 CE-mark per N RTTs (N=BDP, in segments)
- L4S flows thus induce frequent ECN marking, causing Classic flows to yield to “apparent” congestion
  - Cubic (classic) throughput shown to suffer in competition with Prague (L4S)
  - Not believed to be a starvation risk
  - Less unfairness at lower link rates





# Coexistence Caveats & Observations

- Prevalence of FIFO Classic ECN AQMs
  - Believed to be uncommon
  - None specifically identified yet
  - Some evidence of possible new deployments
- What is “reasonable fairness” today?
  - Flow rates are already unfair during convergence, by definition
    - Even Cubic typically takes 10-30 seconds to converge
    - How frequently do >1 simultaneous flows all last longer?
  - Flow sharding is commonplace
  - Existing deployed CC algorithms have differing performance<sup>1</sup>
  - RTT-unfairness is common

<sup>1</sup>. Kunze, R uth, Hohlfeld, “Congestion Control in the Wild – Investigating Content Provider Fairness”, IEEE Transactions on Network and Service Management, December 2019.

# Options for CDN\* Servers

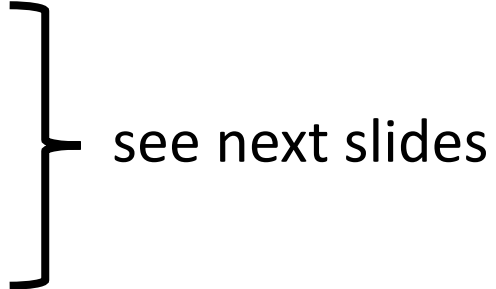
- Prior to deploying L4S on servers:
  - Consult with network operators on presence of RFC3168 bottlenecks
  - Perform downstream tests per access network
    - Tests (TBD) to detect absence of RFC 3168
    - Enable AccECN feedback, but enable/disable L4S per access network
  - Upstream tests?
- In-band RFC3168 detection and monitoring
  - Real-time response (fallback)
  - Non-real-time response (disable for future connections)

\*and other hosts that serve a similarly constrained set of networks/clients

# Options for non-CDN Hosts

- In-band RFC3168 detection (and possibly fallback)
- Per-dst path test:
  - For a connection capable of L4S feedback
  - If CE feedback, perform active test (TBD) for RFC3168 presence
  - Could cache result per-dst
- Query a TBD public whitelist of domains that are participating in L4S experiment

# Options for Operator of a Classic ECN FIFO

- Configure AQM to treat ECT1 as NotECT
  - Configure Non-Coupled Dual Queue
  - WRED with ECT1 Differentiation
  - ECT1 tunnel bypass
  - Disable RFC3168 ECN marking
  - Re-mark ECT1 to NotECT prior to AQM (e.g. ingress)
- 
- see next slides

# Configure Non-coupled DualQ AQM

## Option 1:

- Configure 2 queues, both with ECN
- 50:50 WRR scheduler
- Queue #1: ECT1 & CE packets
  - Shallow immediate AQM target
- Queue #2: ECT0 & NotECT packets
  - Classic AQM target
- Outcome
  - $n$  L4S flows and  $m$  long-running Classic flows
  - if  $m$  &  $n$  are non-zero, get  $1/2n$  &  $1/2m$  of the capacity, otherwise  $1/n$  or  $1/m$
  - never  $< 1/2$  each flow's rate if all had been Classic

## Option 2:

- Configure 2 queues, both with AQM
- 50:50 WRR scheduler
- Queue #1: ECT1 & NotECT packets
  - ECN disabled
- Queue #2: ECT0 & CE packets
  - ECN enabled
- Outcome
  - ECT1 treated as NotECT
  - Flow balance for the 2 queues the same as in option 1

# WRED with ECT1 Differentiation

- Similar to option 2 on prev. slide, but...
  - Single Queue, with two WRED classes, both with AQM
  - Class #1: ECT1 & NotECT packets
    - ECN disabled
  - Class #2: ECT0 & CE packets
    - ECN enabled

# ECT1 tunnel bypass

- Tunnel ECT1 traffic with outer header indicating Not-ECT
  - RFC6040 compatibility mode tunnel
- Two Variants
  1. per-domain: tunnel ECT1 pkts to domain edge towards dst
  2. per-dst: tunnel ECT1 pkts to dst

# Researchers

- Possible Research Goals:
  - Identify prevalence (and locations?) of Classic ECN FIFO bottlenecks
  - Additional studies of L4S/Classic interaction
- This draft should:
  - Provide guidance on metrics to monitor
    - e.g. use RFC3168 detection in Prague CC but without fall-back
  - Describe active testing methods
    - in-band or out-of-band
    - distinguish FIFO from FQ



# Standards Actions

- Could update RFC3168 and/or RFC8311 to clarify that ECT1->CE marking in classic AQMs is “deprecated”
  - Not absolutely necessary - any of the above approaches is already compliant with existing standards track RFCs
- Could update RFC3168 to add:
  - “It SHOULD be possible to configure a device that supports RFC3168 ECN marking to only apply marking to certain ECN codepoints, e.g. only ECT(0) or only ECT(1).”