Identifying Modified ECN Semantics for Ultra-Low Queuing Delay
draft-briscoe-tsvwg-ecn-l4s-id-02
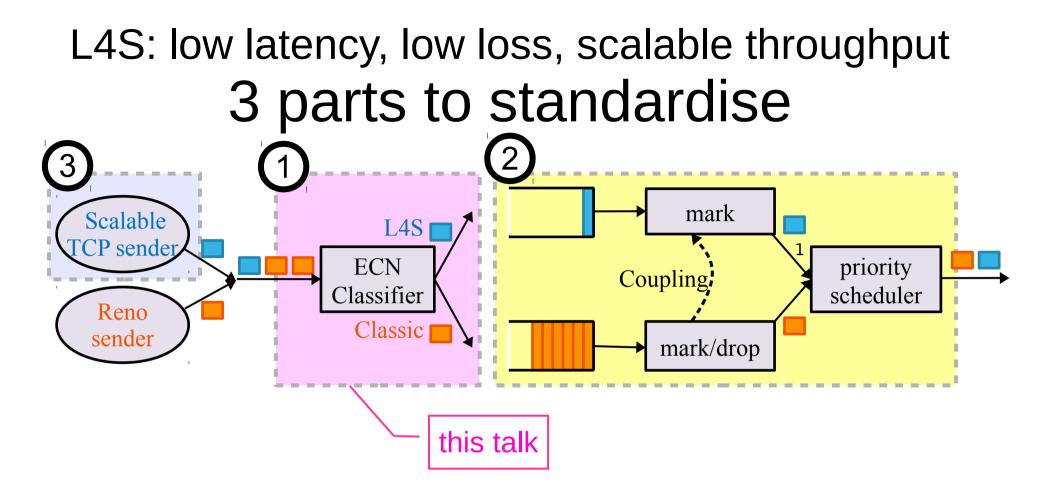
Bob Briscoe

[simula . research laboratory]

• Koen De Schepper, Inton Tsang

NOKIA Bell Labs

IETF-97 Jul 2017

L4S: low latency, low loss, scalable throughput
3 parts to standardise

this talk

1)	The identifier	draft-briscoe-tsvwg-ecn-l4s-id	tsvwg
2)	The DualQ AQM	draft-briscoe-tsvwg-aqm-dualq-coupled	aqm?
3)	Scalable transports	many	?

The goal

- Experimental RFC to assign an identifier for L4S

- Previously focused on choice of identifier: ECT(1)

  - Prerequisite:
    release ECT(1) from prior experimental use as ECN Nonce
    <draft-black-tsvwg-ecn-experimentation>

- This presentation:

  - how we've defined the meaning of ECT(1)

  - adoption call?

ECN Codepoints		
00	Not-ECT	Not-ECT
10	ECT(0)	ECN-Capable Transport
01	ECT(1)	
11	CE	Congestion Experienced

the MUSTs, SHOULDs, etc. pt1/2

- To use L4S, the sender:
  - **MUST** set ECT(1)
  - **SHOULD** ensure rate is inversely proportional to CE marking
  - details for each transport to be specified separately

- To support L4S, a network node:
  - **MUST** classify ECT(1) and **SHOULD** classify CE as L4S
  - **MUST** (?) also implement a Classic AQM treatment
  - **MUST** classify ECT(0) and Not-ECT as Classic

ECN Codepoints		
00	Not-ECT	Not-ECT
10	ECT(0)	ECN-Capable Transport
01	ECT(1)	
11	CE	Congestion Experienced

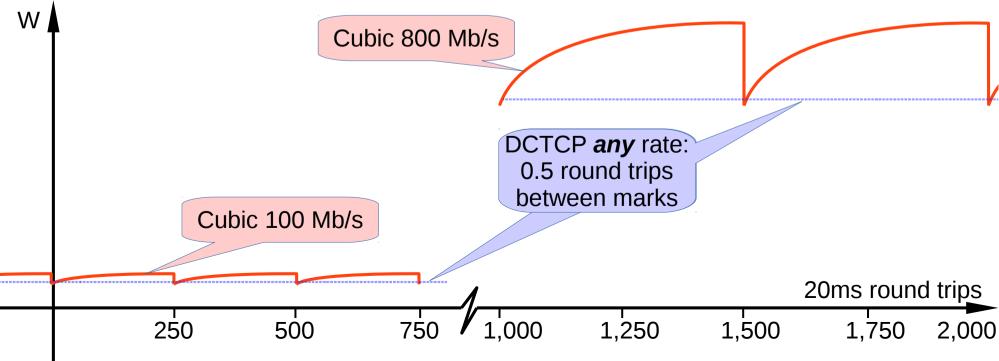# Why should rate be inversely proportional to marking?

- Rationale: Scalable
  - invariant number of control signals per RTT
- The rule is easy to derive:

  requirement: no. of marked segments per round trip = constant, C

  $\rightarrow$ segments per round trip (W) x probability each will be marked (p) = C

  $\rightarrow Wp = C$

  $\rightarrow W = C/p \rightarrow$ rate should be inversely proportional to marking

W

Cubic 800 Mb/s

DCTCP *any* rate:
0.5 round trips
between marks

Cubic 100 Mb/s

20ms round trips

250   500   750   1,000   1,250   1,500   1,750   2,000

the MUSTs, SHOULDs, etc. pt2/2

- Meaning of Classic ECN
    - AQM will mark ECT(0) packets as CE under the same conditions as it would drop Not-ECT packets [RFC3168]

- Meaning of L4S ECN
    - Likelihood that an AQM drops a Not-ECT Classic packet, p_C **MUST** be roughly proportional to the square of the likelihood that it would mark it, if it was an L4S packet, p_L

    $$p_C \approx (p_L / k)^2$$

    - no need to standardize k for interoperability,
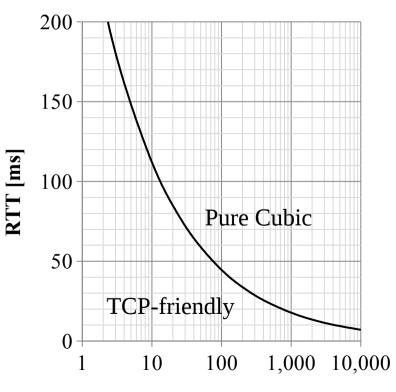   2 is **RECOMMENDED** experimentally

Why squared?

$$p_C \approx (p_L / k)^2$$

- To shift to the scalable regime of L4S,
  by counterbalancing the square root in std TCP congestion avoidance [RFC5681]

$$W = \kappa / \sqrt{p_C}$$

  which has become the gold-standard rate per flow
- Not all traffic behaves like this
  - not all traffic is standard TCP in congestion avoidance
  - short flows
  - not all TCPs are standard, e.g. Cubic, Compound
- Principle:
  - Avoid starvation of any long flows
  - CC of short flows only needs *any* congestion signal
  - Do no harm to the lamest TCP
- Pragmatic:
  - Cubic, Compound are often in their
    TCP-friendly mode over typical low RTT paths

Next Steps

- adoption call
- consider carefully before reassigning a scarce IP header codepoint for a new experiment
- Please review, comment, implement
  - brief draft (8pp without boilerplate & appendices)
- Plenty of discussion already
  - on aqm@ietf.org when issue first raised
  - on tcpprague@ietf.org
  - in L4S BoF
- pls discuss L4S ID on tsvwg@ietf.org for now
  - cc: tcpprague@ietf.org if you like

Q&A

large saw teeth can ruin the quality of your experience