Identifying Modified ECN Semantics for Ultra-Low Queuing Delay

draft-briscoe-tsvwg-ecn-l4s-id-01

Bob Briscoe

[simula . research laboratory]

- Koen De Schepper, Inton Tsang

NOKIA Bell Labs

IETF-96 Jul 2016

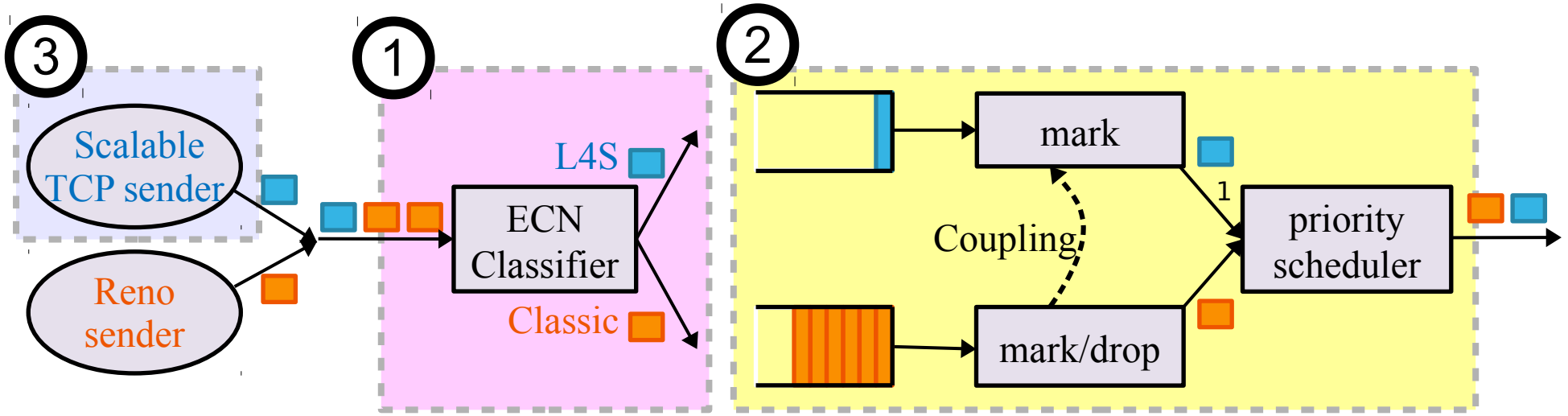
The authors were part-funded by the European Community under its Seventh Framework Programme through the Reducing Internet Transport Latency (RITE) project (ICT-317700). The views expressed here are solely those of the authors.

Low Latency Low Loss Scalable throughput (L4S) – recap

Latency so low it sticks to your finger

- Recall: demo at Prague IETF (aqm wg & bits-n-bites) and L4S BoF yesterday
  - see https://riteproject.eu/dctth/ for videos, papers, etc
- L4S could incrementally replace “best efforts”
  - ultra-low queuing delay
  - zero congestion loss
  - scalable throughput (beyond Reno, Compound, Cubic)
- Eventually for *all* Internet traffic
- Aim: to be worth the deployment hassle – so much better than today

3 parts to standardise



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

this talk

- #1, #2 are as general as possible
- #3 is more specific to each transport

identifier requirements

- Chosen ID likely to involve compromises
  - 'cos limited IP header space
- All requirements expressed as SHOULDs in draft:
  - **end-to-end?** traverses boundaries: host-network-network-middlebox-host
  - protocol agnostic: common to IPv4 & IPv6 and transport agnostic
  - incrementally deployable
  - works for AQMs in **tunnels** or **lower layers**
  - consume minimal **codepoints**
  - avoid **reordering**: whole flow handled by same queue
    - including **control packets**
- Words in bold turned out to be the distinguishing issues

choice of identifier

- Three possibilities; all involve compromises
  - two other possibilities quickly discounted
  - ECT(1) + CE chosen
  - reasoning recorded in Appendix A of draft
  - table highlights solely the distinguishing issues

Issue	DSCPx + ECN initial	DSCPx + ECN eventual	ECN* initial	ECT(1) + CE initial	ECT(1) + CE eventual
end-to-end?		?			
tunnels			?	?	
lower layers		?			?
codepoints		?			?
reordering					
control packets					?

Legend	
	Poor
	Ordinary
	Good
?	Optimistic

CAVEAT: The table is not meant to be understandable without referring to the text.

---

* only feasible to use ECN alone if Classic ECN becomes obsolete
  "ECN" shares the eventual scores of "ECT(1) + CE"

the chosen identifier: compromises

- end-to-end?
  - ECN pretty good traversal
  - Don't need to traverse L4+ middleboxes
- tunnels
  - RFCs OK, but non-compliant tunnels likely
  - if operators deploy ECN AQMs,
    they will at least fix their own tunnels
- lower layers
  - ECN protocol when AQM is at lower layer:
    some specified, but implementation very patchy
- codepoints
  - Any use of the ECN field will consume the last codepoint
- reordering
  - Corner case: Classic ECN AQM upstream AND multi-bottleneck
    Effect of a few Classic CE packets arriving early to be assessed
- TCP control packets
  - Unless ECT is allowed*, control packets will get Classic delay

Issue	. .	ECT(1) + CE initial	eventual
end-to-end?	.		
tunnels	.	?	
lower layers	.		?
codepoints	.		?
reordering	.		
control packets	.		?

* draft-bagnulo-tsvwg-generalized-ecn

meaning of this new identifier?

- Original goals of ECN included lower delay with modified TCP
  - but too many combinations to standardise a winner
  - so [RFC3168] defined 'Classic' ECN behaviour as *equivalent* to drop

    ...could allow new criteria to be developed for setting the CE codepoint, and new congestion control mechanisms for end-node reaction to CE packets. However, this is a research issue, and as such is not addressed in this document.
  - so ECN inherited the hexlemma of drop-based TCP...

- Proposed meaning of L4S identifier:

  The likelihood that an AQM drops a Not-ECT Classic packet MUST be roughly proportional to the square of the likelihood that it would have marked it if it had been an L4S packet.

  The constant of proportionality does not have to be standardised for interoperability, but a value of 1 is RECOMMENDED.

  – if we (IETF) define TCP Prague differently, this will have to be revisited to match

Process

- L4S-ID draft written as Experimental
  - cannot update pre-existing PSs that mention ECT(1) alongside ECT(0)

- Current idea is a Proposed Standard that
  - obsoletes ECN Nonce [RFC3540]
  - reserves ECT(1) for future experiments (e.g. L4S).
  - It would update:
    - ECN in IP [RFC3168]
    - ECN in TCP [RFC3168]
    - ECN in SCTP [RFC4960]
    - ECN in RTP [RFC6679]
    - ECN in DCCP [RFC4340]

Next Steps

- consider carefully before consuming the last ECN codepoint
- Please review and comment
  - 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

Q&A

large saw teeth can ruin the quality of your experience

If we choose ECT(1): what do we preclude?

- would obsolete:
  - ECN nonce [RFC3540]
    - other ways to do feedback integrity without a codepoint
  - various “less severe than CE” schemes from the research community, incl. for flow start-up
  - see Appendix B