More Accurate ECN Feedback in TCP

draft-ietf-tcpm-accurate-ecn-20





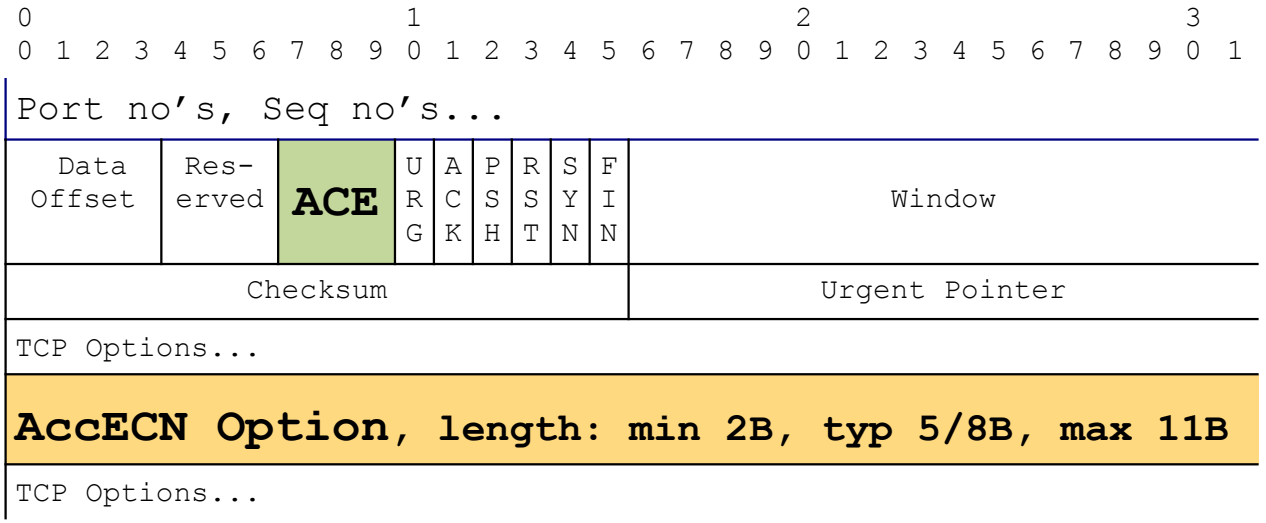


Bob Briscoe, Independent
Mirja Kühlewind, Ericsson
Richard Scheffenegger, NetApp

IETF-114 Jul 2022

Solution (recap)
Congestion extent, not just existence

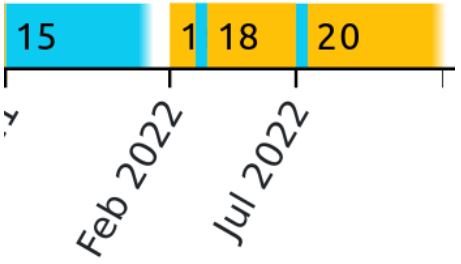
- AccECN: Change to TCP wire protocol
  - Repeated count of CE packets (ACE) - essential
  - and CE bytes (AccECN Option) – supplementary



- Key to congestion control for low queuing delay
  - 0.5 ms (vs. 5-15 ms) over public Internet

Recent draft history

draft-ietf-tcpm-accurate-ecn



- 22-Mar-22: -18
- 11-Jul-22: -19 [summary of diffs to list]
- 25-Jul-22: -20 [summary of diff to list]
- 18 → 19: 2 changes promised last meeting (slides 4&5):
  - one in response to Gorry on ACK filtering
  - one in response to a suggestion from Ilpo on optionality of implementation of the AccECN Option
  - plus one other minor change from Ilpo
    - reorder IP-ECN mangling tests to match implementation order
- 19 → 20: 2 recent changes (slides 5&6):
  - follow-on from change suggested by Ilpo
  - IANA exID assignments - NCardwell – RScheff uncovered in interop

Change #1: draft 18 → 19
Interaction with ACK Filtering (§3.3.3)

- Previously, draft-18 updated RFC3449*; saying that an ACK filtering node...
  `SHOULD determine if an ACK is part of a connection using AccECN and SHOULD then preserve the correct operation of AccECN feedback.`
- Gorry:
  - no need for AccECN to update RFC3449,
    which already says (in its section on ACK filtering):
    `need to preserve correct operation of ECN feedback [RFC3168].`
  - AccECN will update RFC3168,
    so this statement in RFC3449 will apply to AccECN feedback too
- Text changes in draft-19:
  - Emphasizes that the RFC3449 statement applies to AccECN too;
  - Also, more technical detail in the "how-to":
    - states worst-case wrap of ACE field (8 ACKed packets)
    - presence of AccECN TCP Option doesn't remove ACE wrap issue,
      because the Option might not be read

*TCP Performance Implications of Network Path Asymmetry

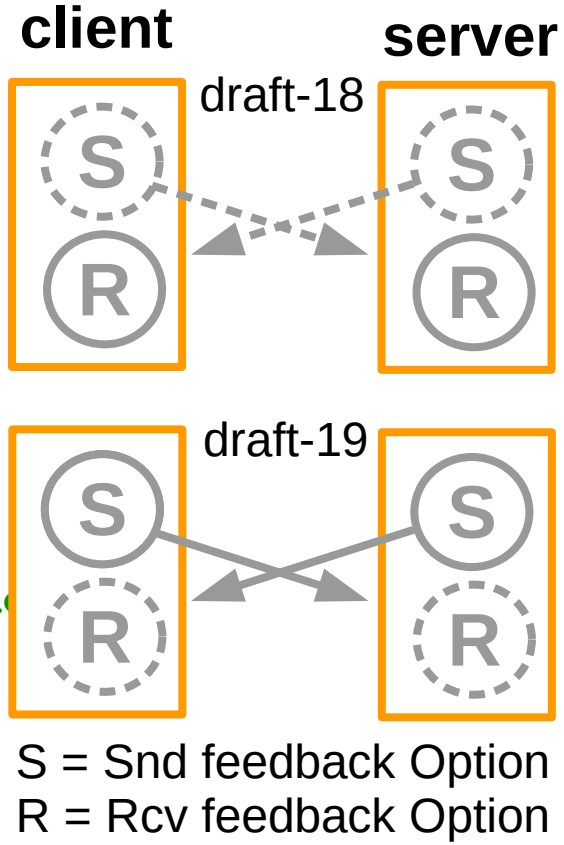
Change #2 draft 18 → 20
The AccECN Option (§3.2.3)

- Draft-19 and -20 say (paraphrasing):

  `Even if a developer does not implement logic to understand received AccECN Options it is RECOMMENDED that they still implement logic to send AccECN Options.`

- Draft-20: added additional text more strongly motivating the RECOMMENDATION:

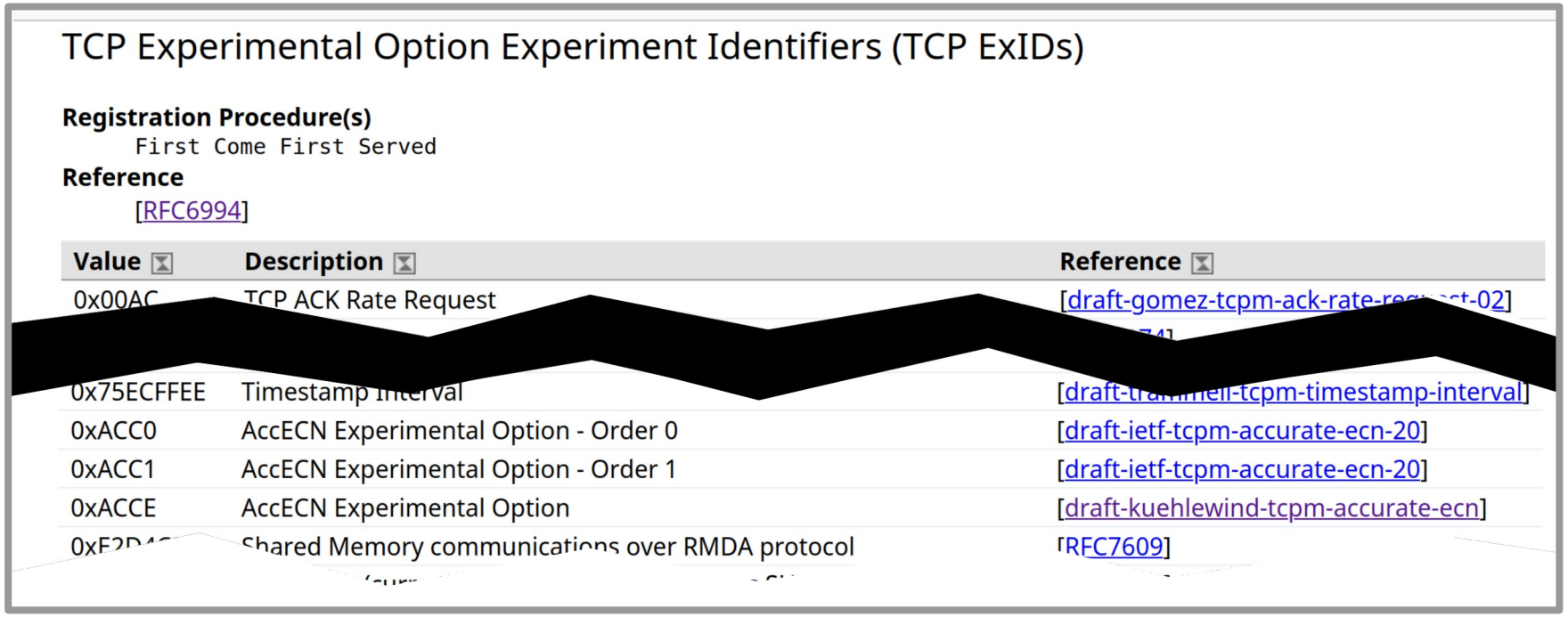
  `Otherwise, those remote peers that implement the receiving logic will still be excluded from congestion feedback that is robust against the increasingly aggressive ACK filtering in the Internet.` `The logic to send AccECN Options is the simpler to implement of the two sides.`



- BTW: Both snd & rcv logic already implemented for Linux.
  Assuming that covers majority of servers, implementing snd logic in other clients at least completes a robust feedback loop for most downstream flows

Change #3 draft 19 → 20
IANA Considerations (§7)

- Retrospectively registered Experimental IDs
  - used for pair of TCP Options in Linux implementation
  - interop testing discovered different from FBSD pair



TCP Experimental Option Experiment Identifiers (TCP ExIDs)

Registration Procedure(s)
First Come First Served

Reference
[RFC6994]

Value	Description	Reference
0x00AC	TCP ACK Rate Request	[draft-gomez-tcpm-ack-rate-req…st-02]
0x75ECFFEE	Timestamp Interval	[draft-…-tcpm-timestamp-interval]
0xACC0	AccECN Experimental Option - Order 0	[draft-ietf-tcpm-accurate-ecn-20]
0xACC1	AccECN Experimental Option - Order 1	[draft-ietf-tcpm-accurate-ecn-20]
0xACCE	AccECN Experimental Option	[draft-kuehlewind-tcpm-accurate-ecn]
0xE2D4…	Shared Memory communications over RMDA protocol	[RFC7609]

- Recorded these ExIDs in draft

Status & Next Steps

draft-ietf-tcpm-accurate-ecn-20

- SECDIR (Scott Kelly) review: tracker status
  - he intended 'Almost ready' after draft was fixed
  - but still showing as 'Has issues'
- Now ready for WGLC
- Early pair of TCP Option Kind assignments?
- draft-ietf-tcpm-generalized-ecn (EXP)
  also ready but dependent on AccECN

AccECN

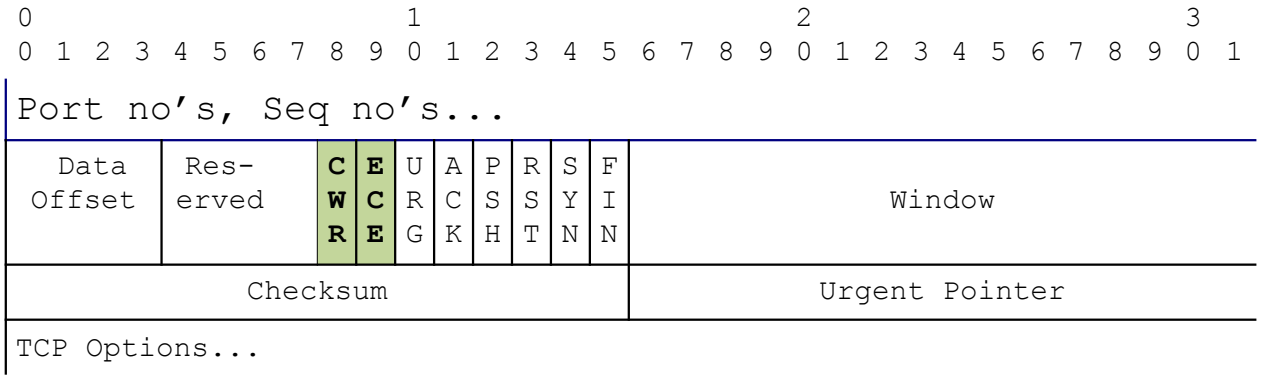
Q&A

spare slides

Problem (Recap)
Congestion Existence, not Extent

- Explicit Congestion Notification (ECN)
  - routers/switches mark more packets as load grows
  - RFC3168 added ECN to IP and TCP

IP-ECN	Codepoint	Meaning
00	not-ECT	No ECN
10	ECT(0)	ECN-Capable Transport
01	ECT(1)	
11	CE	Congestion Experienced



- Problem with RFC3168 ECN feedback:
  - only one TCP feedback per RTT
  - rcvr repeats ECE flag for reliability, until sender's CWR flag acks it
  - suited TCP at the time – one congestion response per RTT