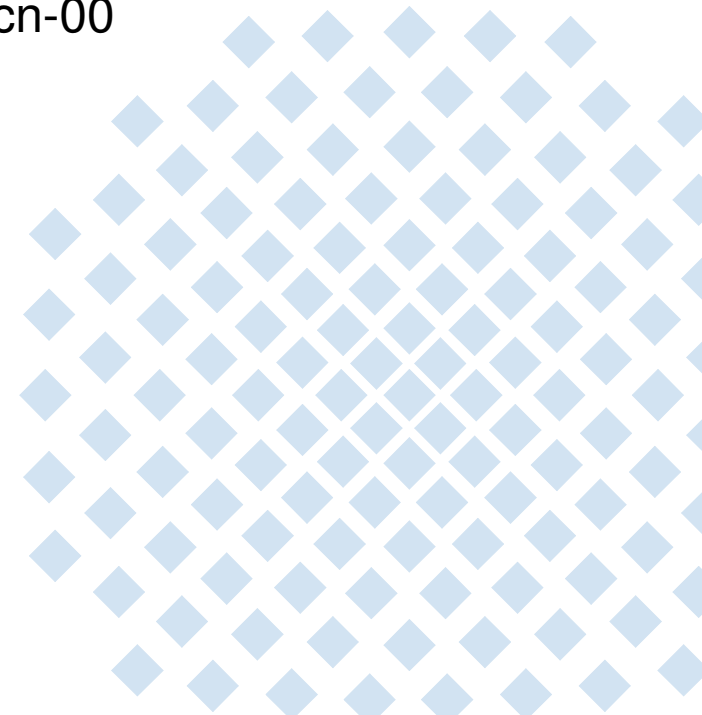


Accurate ECN Feedback in TCP

tsvarea – 81. IETF Quebec – July 28, 2011

draft-kuehlewind-conex-accurate-ecn-00

Mirja Kühlewind <mirja.kuehlewind@ikr.uni-stuttgart.de>
Richard Scheffenegger <rs@netapp.com>



Motivation

draft-kuehlewind-conex-accurate-ecn-00

Congestion Exposure (ConEx)

- Mechanism by which senders inform the network about the congestion encountered by previous packets on the same flow
- Mechanism to retrieve more accurate ECN feedback needed

Problem

ECN provides only one congestion feedback signal per RTT
as designed for current congestion control mechanisms
that will react only once per RTT on congestion

Scope

- Can also be used by other TCP mechanisms. e.g. DCTCP
- Congestion control might react differently on ECN in future (ICCRG and LEDBAT ML)
- Not ConEx specific; Feedback of a wider community needed

Accurate ECN Feedback in TCP

draft-kuehlewind-conex-accurate-ecn-00

- Requirements on resilience, timeliness, integrity, accuracy and complexity listed
- Re-use of the ECN/ECN-Nonce TCP bits (or TCP Option)
- Discussion (ACK loss, ECN Nonce) not exhaustive yet...
→ Please read draft

Three coding options proposed (and discussed)

1. One bit feedback flag
 - Set ECE only in one (or N subsequent) ACKs
 - Remark: CWR is unused; can be used for redundancy in subsequent ACK (not in draft...)
 2. Three bit field with counter feedback
 - Use ECE/CWR/NS signal a counter value (mod8) in every ACK (as with re-ECN)
 - Does not allow ECN Nonce
 3. Codepoints with dual counter feedback
 - Have 2 counter (CE, ECT(1)) encoded in 8 codepoints (send congestion value by default)
- Chose one of the schemes (remove the other option from the draft) + specify protocol

Question?

Backup

Accurate ECN Feedback in TCP

Design Choices

- Re-use of the ECN/ECN-Nonce TCP bits
 - Classic ECN should not be used in parallel anymore
- No additional bits from three reserved bits in TCP header
 - No additional benefit (only shift of problems in time)
- No extra TCP Option
 - Deployment issues because of middleboxes
 - Growth of header length (goal would be to have this mechanism activated by default)
 - Could provides more information e.g. explicit the number of ECT(0), ECT(1), CE, non- ECT marked and lost packets (as in ECN for RTP/UDP), but is this needed?