Guidelines for Adding Congestion Notification to Protocols that Encapsulate IP

draft-briscoe-tsvwg-ecn-encap-guidelines-01

Bob Briscoe

IETF-85 Nov 2012

explicit congestion notification (ECN)

- growing interest again
  - in recognition of the importance of low delay
  - particularly in L2 networks (backhaul, data centres)

- drop: both congestion signal and impairment
  - compromise: deliberately delay the signals (bufferbloat)
- ECN: a signal without impairment
  - can signal as early as needed

problem

- AQM* & ECN are for queues at any layer
  - not just IP
- ECN has to be explicitly propagated
  - up the layers
- in contrast drop is easy
  - it naturally propagates up the layers

* AQM = active queue management (e.g. RED)

aim of this draft

- guidelines for writing specs to propagate ECN up to IP from:
  - L2 protocols (e.g. IEEE802, TRILL)
  - tunnelling protocols (L2TP, GRE, PPTP, GTP,…)
- for authors who may not be ECN experts

draft status

- resurrected individual draft, -00 posted 18months ago
- fell down my priority list, but has become important again
- intended status: best current practice

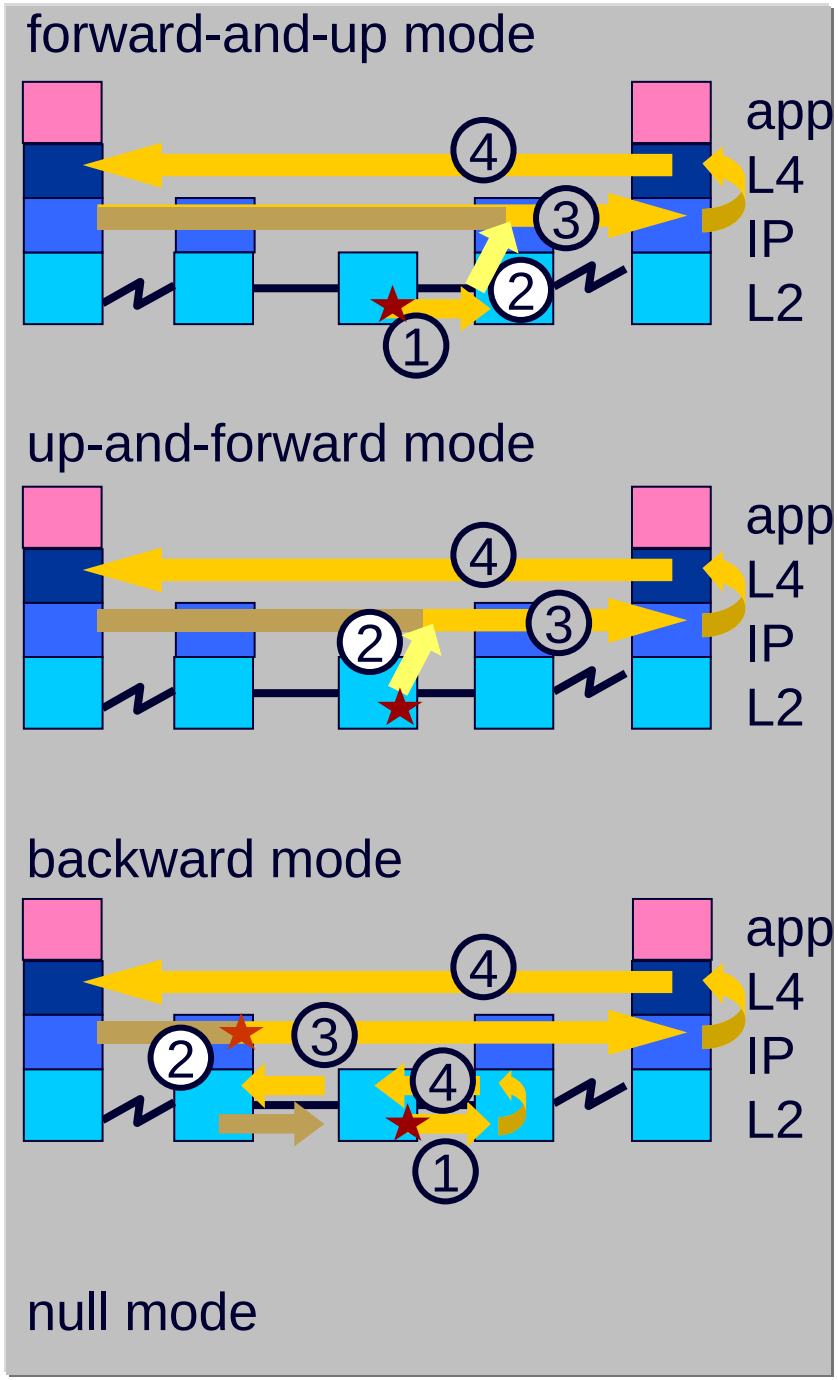
L2TP = layer 2 tunnelling protocol [RFC2661]
PPTP = Point-to-point Tunnelling Protocol [RFC2637]
GRE = generic routing encapsulation [RFC1701, RFC2784]
QCN = quantised congestion notification [IEEE 802.1Qau]
GTP = GPRS tunnelling protocol [3GPP TS 29.060]

a variety of arrangements

- avoid precluding L2 innovation
- must not be over-prescriptive

- guidelines for each mode
  - see draft (or spare slides)

- wide expertise needed for authoring & review



next steps

- process
  - aim to request adoption as wg item at next IETF
  - will require liaison with other standards bodies
    - setting requirements for interfacing IP with their protocols
- document
  - I’m conscripting expert help
    - L2 & tunnelling experts
  - text & structure largely complete
    - new co-authors & reviewers may disagree
  - reviews pls

Guidelines for Adding Congestion Notification to Protocols that Encapsulate IP

draft-briscoe-tsvwg-ecn-encap-guidelines-01

Q&A

& spare slides

status of congestion notification
in protocols that encapsulate IP

- IETF

  done: MPLS-in-MPLS, IP-in-MPLS [RFC5129], IP-in-IP [RFC6040]

  to do: trill-rbridge-options (in progress),
  & pass ECN thru tunnel protocols, eg. L2TP, GRE

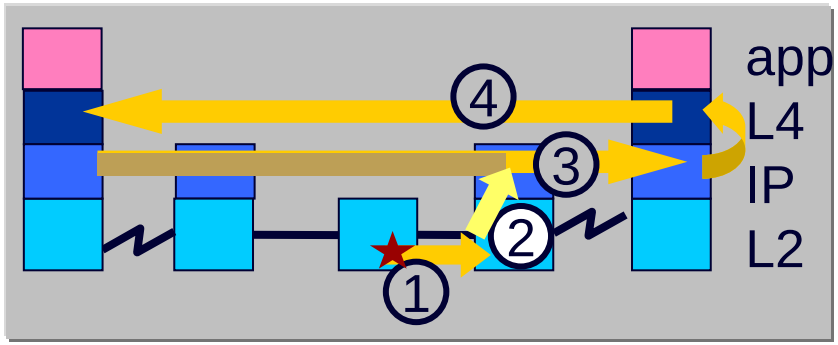
- Other standards bodies:

  done: QCN [802.1Qau], Frame Relay, ATM [I.371]
  (all subnet-local)

  todo: IEEE 802.1, (802.3, 802.11), …?
  & pass ECN thru tunnel protocols, eg. 3GPP GTP

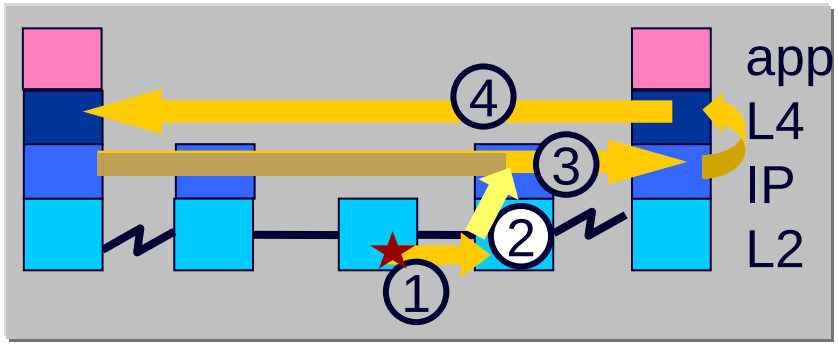
L2TP = layer 2 tunnelling protocol [RFC2661]
GRE = generic routing encapsulation [RFC1701, RFC2784]
QCN = quantised congestion notification
GTP = GPRS tunnelling protocol - user plane [3GPP TS 29.281]

forward and upward mode requirements



- identifying whether transport will understand ECN
- identifying whether egress will understand ECN
- propagating ECN on encapsulation
- propagating ECN on decapsulation
- reframing issues

forward and upward mode guidelines



- identifying whether transport will understand ECN
  - ‘ECN-capable transport’ codepoint or other approaches
- identifying whether egress will understand ECN
  - new problem
- propagating ECN on encapsulation
  - copying ECN down for monitoring purposes
- propagating ECN on decapsulation
  - combining inner & outer
- reframing issues
  - marked bytes in ≈ marked bytes out
  - timeliness – don’t hold back any remainder

the main problem: incremental deployment

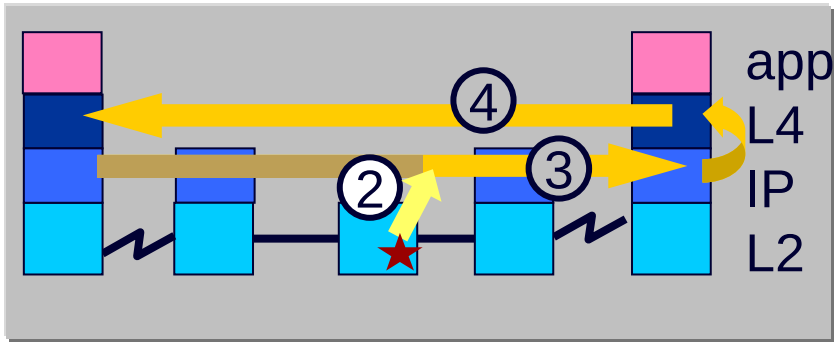
- IP-ECN designed for incremental deployment

		congested queue supports ECN?	
transport supports ECN?	IP header	N	Y
N	Not-ECT	drop	drop
Y	ECT	drop	CE

- if transport only understands drop
  - lower layer must not send it congestion indications
- need not mimic IP mechanism (grey)
  - but needs to achieve same outcome (white)
  - also, must check egress understands ECN too

ECT = ECN-capable transport
CE = Congestion Experienced

up and forward mode guidelines

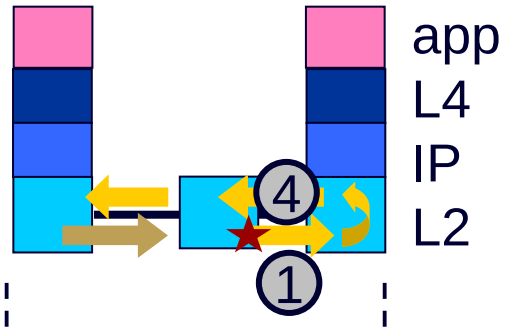


- identifying whether transport will understand ECN
  - use IP mechanism
- ~~identifying whether egress will understand ECN~~
- ~~propagating ECN on encapsulation~~
- ~~propagating ECN on decapsulation~~
- ~~reframing issues~~
- a layering violation
  - but safe if guidelines apply

backward mode

IEEE 802.1Qau (QCN)
ATM ITU-T-I.371
Frame Relay

- often designed for where the subnet is the whole network



not a good fit

- doesn't interwork efficiently with IP's forwards-only mode