Byte and Packet Congestion Notification

draft-ietf-tsvwg-byte-pkt-congest-00.txt

Bob Briscoe, BT & UCL
IETF-73 tsvwg Nov 2008

status

- Byte and Packet Congestion Notification
    - **new WG draft:** draft-ietf-tsvwg-byte-pkt-mark-00.txt as of 8-Aug-08
    - **previously:** draft-briscoe-tsvwg-byte-pkt-mark-02
    - **intended status:** informational (update to RFC2309 advice)
    - **RFC publication milestone:** Sep '09
    - **immediate intent:** encourage review
    - **w-gs & r-gs affected:** TSVWG, DCCP, PCN, ICCRG & PWE3

reminder (exec summary)

- scope
  - in any AQM (e.g. RED drop, RED ECN, PCN)
    should we allow for packet-size when network writes or when transport reads a loss or ECN mark?
- what little advice there is in the RFC series (RFC2309 on RED) is unclear
  - gives both options with 'more research needed'
  - slight bias to favouring small packets, which give perverse incentives to create small packets and seems to encourage a dangerous DoS vulnerability
- unequivocal UPDATE to advice in RFC2309
  - AQM SHOULD NOT give smaller packets preferential treatment
  - adjust for byte-size when transport reads NOT when network writes
- all known network layer implementations follow this advice anyway
  - retrospective tidy-up to RFC series
  - avoids complexity of catering for all possibilities, when no-one uses them
  - includes detailed advice on buffer design etc, gathered from experts & literature

---

- Terminology: RED's 'byte mode queue measurement' (often called just 'byte mode') is OK, only 'byte mode packet drop' deprecated
- NOTE: don't turn off RED completely: drop-tail is as bad or worse

why decide now?
between transport & network

- near-impossible to design transports to meet guidelines [RFC5033]
    - if we can't agree whether transport or network should handle packet size

- DCCP CCID standardisation
    - hard to assess TFRC small packet variant experiment [RFC4828]

- PCN marking algorithm standardisation [draft-ietf-pcn-marking-behaviour-01]
    - stds track draft depends on this decision

- part of answering ICCRG question
    - what's necessary & sufficient forwarding hardware for future cc?
    - [draft-irtf-iccrg-welzl-congestion-control-open-research-02] incorporates this I-D by ref

- given no-one seems to have implemented network layer bias
    - advise against it before we're stuck with an incompatible deployment fork

- encouraging larger PMTUs by not favouring smaller ones
    - may start to solve other scaling problems

text updates

[draft-briscoe-tsvwg-byte-pkt-mark-02]
  $\rightarrow$ [draft-ietf-tsvwg-byte-pkt-congest-00]

- few changes since previous (individual) draft

  - summarised at head of document

- added note for RFC Editor

  - "intended to update RFC2309" (RED manifesto)

- added question to outstanding issues section (for ICCRG)

  - will congestion of packet processing become more common?

- updated refs (some in various w-gs have become w-g items)

reviews & comments

- reviews of previous drafts
  - current draft is result of extensive previous reviews

- current version: few comments on list since Aug
  - off-list with Iljitsch van Bejnum, Rob Hancock, Phil Eardley
    - discussion continuing – I'm trying to bring it to the tsvwg list
  - Iljitsch: wanted positive discrimination for large packets by policing small
    - I resisted: congestion notification should reflect probability of congestion no less, no more – otherwise creates unintended consequences
    - also controversy over advice IETF gives to transports
  - Phil: suggestions to make draft clearer

- need reviews
  - signed up: Joe Touch, Wes Eddy, Jukka Manner

conclusion

- unequivocal UPDATE to RFC2309 ('RED manifesto')

  - adjust for byte-size when transport reads NOT when network writes

  - previously gave both options with 'more research needed'

- all known implementations follow this advice anyway

  - retrospective tidy-up to RFC series

- reviews pending

Byte and Packet Congestion Notification

draft-ietf-tsvwg-byte-pkt-congest-00.txt

Q&A