ECN and Congestion Feedback Using the NSH and IPFIX
draft-eastlake-sfc-nsh-ecn-support-01

Donald Eastlake – Huawei d3e3e3@gmail.com
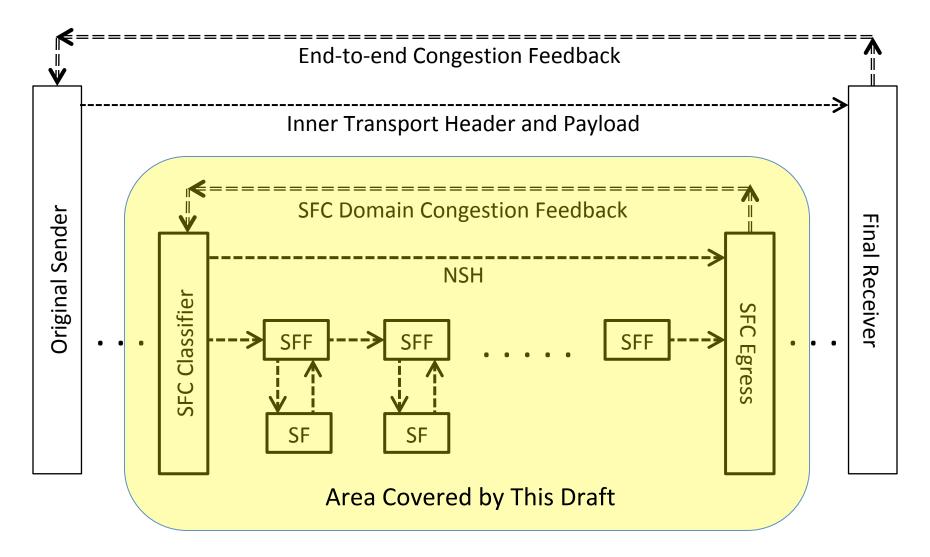
Bob Briscoe – Independent

Andrew Malis – Huawei

Goals of Proposal in this Draft

- Collect congestion information within a Service Function Chaining (SFC) domain with minimal packet drops.

- Communicate congestion information to the Classifier(s) so they might take action to reduce congestion.

High Level Overview

- Congestion encountered is communicated downstream towards the SFC domain egress by Explicit Congestion Notification (ECN, RFC 3168) bits in the Network Service Header (NSH, RFC 8300).

- Congestion information is communicated back upstream to the Classifier using IP Flow Information Export (IPFIX, RFC 7011).

High Level Overview

End-to-end Congestion Feedback

Inner Transport Header and Payload

Original Sender

Final Receiver

SFC Domain Congestion Feedback

NSH

SFC Classifier

SFF → SFF → ... ... → SFF → SFC Egress

SF   SF

Area Covered by This Draft

Downstream ECN

- Congestion information is conveyed downstream by two ECN bits in the NSH. Only in severe cases (or where the end-to-end transport does not support ECN) are packets dropped.

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Ver|O|U|      TTL      |     Length    |U|U|U|U|MD Type| Next Protocol |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                                        ^ ^
                                        | |
                                    +-------+
                                    |NSH ECN|
                                    | field |
                                    +-------+
```

Upstream IPFIX

- IPFIX, as extended by draft-ietf-tsvwg-tunnel-congestion-feedback, provides mechanisms for communicating statistics from which congestions can be determined from an egress to a classifier.

  – Such statistics are cumulative so occasional lost upstream packets are tolerable.
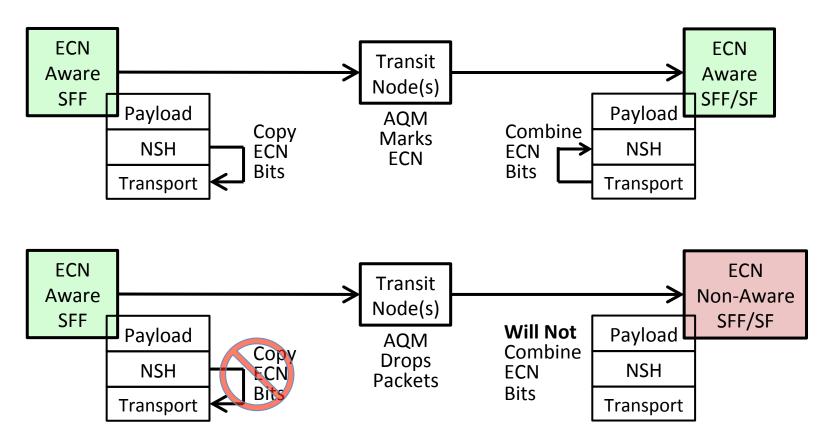
Possible Classifier Actions

- Possible classifier actions on learning of congestion include:
  - Traffic throttling.
  - Congestion feedback further upstream.
  - Traffic re-direction.
    - Yes, you have to be super careful to avoid oscillation. For example if you have long lived flows, the Classifier can choose less congested paths for newly appearing flows.

Some Details

- If the end of an NSH hop (SFF <-> SF, SFF -> SFF) can properly combine Information from the outer transport header into the NSH, then the ingress of such a hop copies the NSH ECN to that transport header. Otherwise, it leaves the outer transport header showing no ECN support and congestion is indicated by packet drop.
  - This requires adding one bit of configuration in each entry at an SFF under a SPI/Index. This bit indicating whether or not the end of the next hop supports ECN when it de-encapsulates.

Some Details

- More graphically

ECN Aware SFF → (Payload / NSH / Transport) — Copy ECN Bits → Transit Node(s) AQM Marks ECN → Combine ECN Bits → (Payload / NSH / Transport) ECN Aware SFF/SF

ECN Aware SFF → (Payload / NSH / Transport) — Copy ECN Bits (crossed out) → Transit Node(s) AQM Drops Packets → **Will Not** Combine ECN Bits → (Payload / NSH / Transport) ECN Non-Aware SFF/SF

Some Details

- This all works better if ECN is implemented throughout the SFC Domain. If an SF does not support NSH or ECN it will have a proxy which should support ECN but even then it is better if the non-NSH supporting SF supports ECN. Basically, any bottleneck where there might be congestion that does not support ECN means that congestion is unmanaged.

Next Step

- Call for WG Adoption?

END

ECN AND CONGESTION FEEDBACK USING THE NSH AND IPFIX