

Fragment Reassembly & ECN

Propagating ECN across IP tunnel Headers Separated by a Shim

draft-ietf-tsvwg-rfc6040update-shim-10

Guidelines for Adding Congestion Notification to Protocols that Encapsulate IP

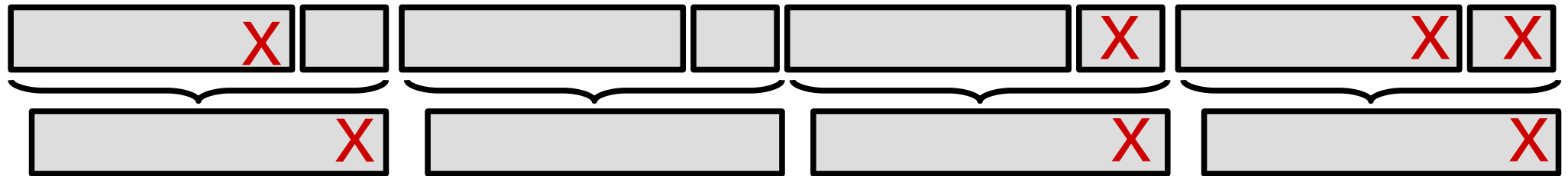
draft-ietf-tsvwg-ecn-encap-guidelines-11

Bob Briscoe
Independent

IETF-107+ tsvwg interim#1, Apr 2020

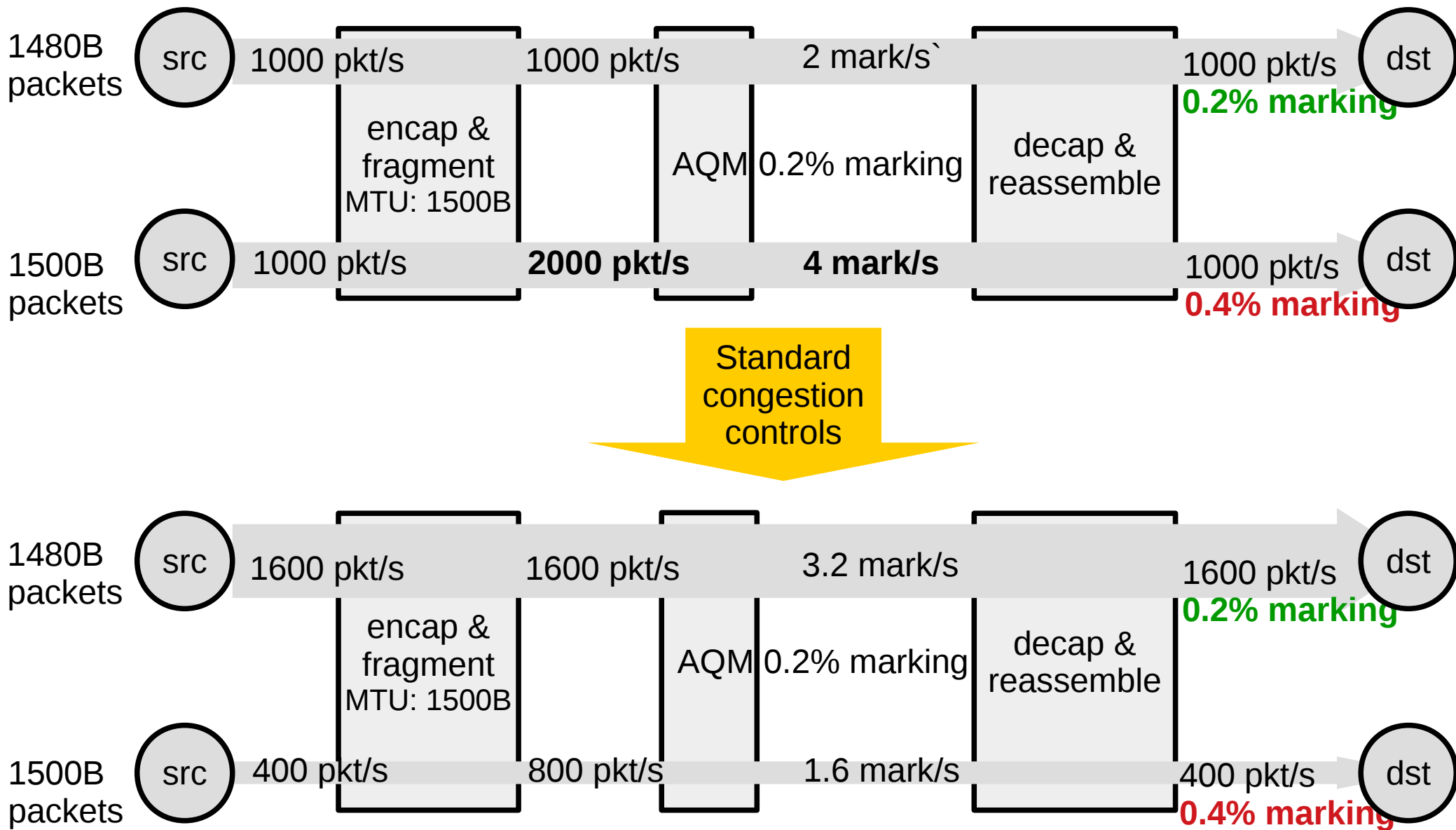
Message

- RFC3168 reassembly preserves the time of each ECN mark



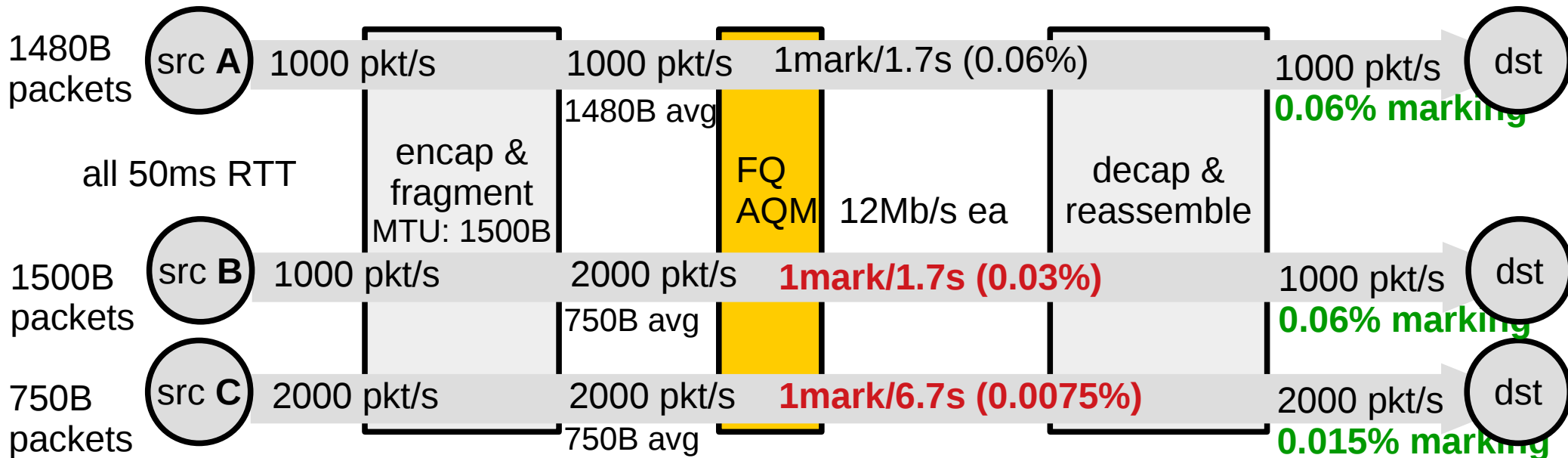
- Certainly, Codel controls the time between marks
 - but it doesn't aim for a specific time between marks
 - it merely adjusts the time until it balances with the CC
- No different to adjusting a probability
- Preserving the time of each mark achieves nothing
 - except 'unfairness'

Example#1/2: RFC3168 ECN & Reassembly



Note: Fragmentation/reassembly & encap/decap are orthogonal, but they often live together

Example#2/2: RFC3168 ECN & Reassembly



- TCP adjusts the marking prob's to the rate the FQ scheduler gives
 - and to the doubling of markings at reassembly
- AQM marks flows A and B every 1.7s
 - even tho packet rate of B is twice as fast
- AQM marks flow C every 6.7s
 - even tho packet rate and packet size of B&C are identical
 - (even tho AQM doesn't know B packets are fragments)
- Proves RFC3168 ECN reassembly is broken

Status and Next Steps

draft-ietf-tsvwg-rfc6040update-shim-06

draft-ietf-tsvwg-ecn-encap-guidelines-11

- Content Development & Socialization (incl. IEEE/3GPP Liaisons):
 - **Complete, except...**
- Decided both these drafts will state:
 - fragmentation / reassembly out of scope
- Mention RFC3168 discusses ECN frag & reassembly
 - but not a normative ref. for encap/decap
- Update RFC3168 frag/reassembly
 - in a separate short standards track RFC

Propagating ECN across IP tunnel Headers Separated by a Shim

draft-ietf-tsvwg-rfc6040update-shim-10

Guidelines for Adding Congestion Notification to
Protocols that Encapsulate IP

draft-ietf-tsvwg-ecn-encap-guidelines-11

Q&A

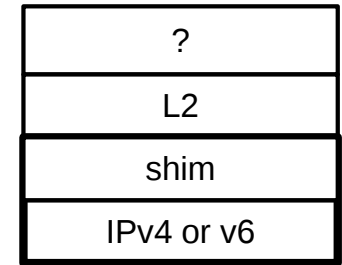
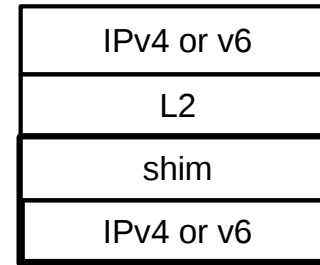
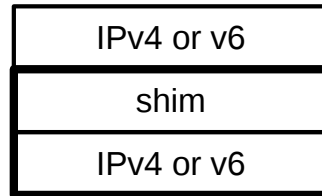
Spare Slides

draft-ietf-tsvwg-rfc6040update-shim-06

Addresses 2 Problems with “Tunnelling of ECN” [RFC6040]

1) Scope omitted shims

- IP-IP tunnels, but not IP-shim-(L2)-IP

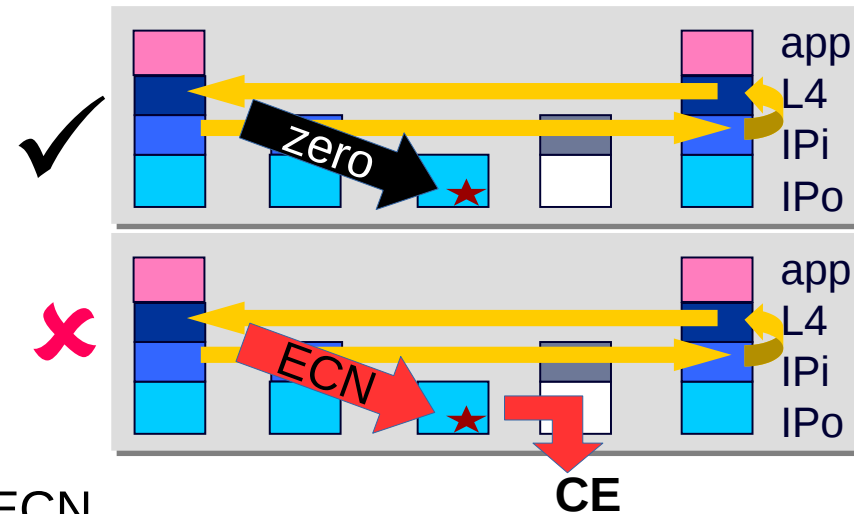


↓
outer

2) If decap non-ECN or unknown, encap MUST zero ECN outer

- How to require this, without making all pre-existing encaps non-compliant?
- RFC6040 did not take a position
- rfc6040update-shim says (paraphrasing):

“if decap does not, or might not, propagate ECN, if possible, **the operator** MUST configure the ingress to zero the outer ECN field”



Survey of IP-shim-(L2)-IP encaps

Protocol	RFC	STDs or widely deployed	AOK	NOK: 6040shim updates	NOK: non-IETF: update recommended
Geneve	nvo3-geneve	✓	✓		
GUE	intarea-gue	✓	✓		
SFC	7665	✓	N/A?		
VXLAN	7348	✓			✓
VXLAN-GPE	nvo3-vxlan-gpe	✗			
LISP	6830	✓	✓		
CAPWAP	5415	✓	✓		
Teredo	4380	✓		✓	
GTP	v1, v1U,v2C	✓			✓
GRE	2784	✓		✓	
NVGRE	7637	✓		↑	
(P)MIP{4 6}	5944,6275,5845	✓		✗	
L2TPv3	3931	✓		✓	
L2TPv2	2661	✓		✓	
PPTP	2637	✗			
NSH	8300	✓		✗ ⇨ new draft	