Tunnelling of Explicit Congestion Notification

draft-briscoe-tsvwg-ecn-tunnel-03.txt

Bob Briscoe, BT

IETF-75 saag Jul 2009

This work is partly funded by Trilogy, a research project supported by the European Community www.trilogy-project.org

status

- Tunnelling of Explicit Congestion Notification
  - **new WG draft:** draft-ietf-tsvwg-ecn-tunnel-03.txt 21 Jul '09
  - **intended status:** standards track
  - **updates (if approved):** 3168, 4301
  - **RFC pub target:** Dec ‘09
  - **immediate intent:** socialise in security area

  - only changes to 4301 are at decap
  - adds new behaviours for previously unused combinations of inner and outer header
    - operators who want the new behaviours can require compliance
    - backward compatible; can update remaining decapsulators lazily
  - as with ECN in 4301: no modes, no capability negotiation

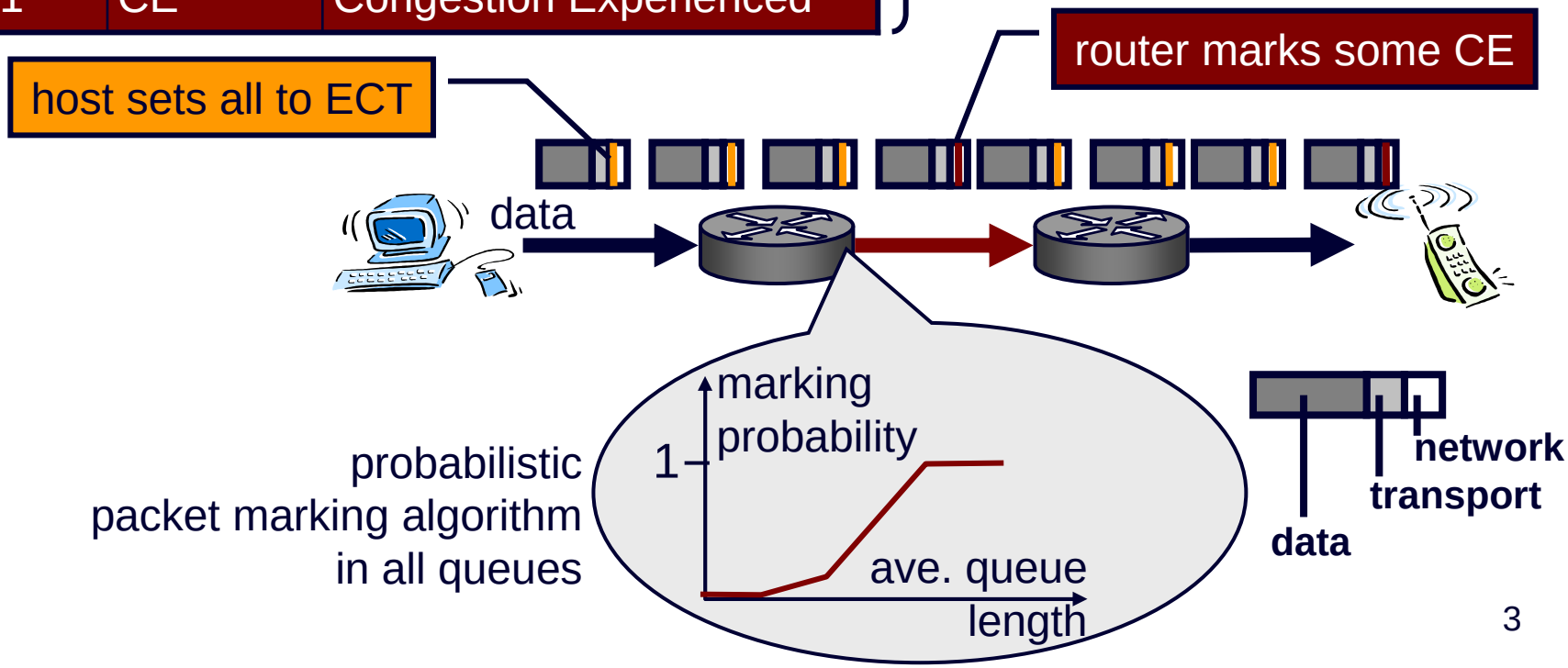
explicit congestion notification (ECN RFC3168) recap

bits 6 & 7 of IP Diffserv field

ECN field	codepoint	meaning
00	Not-ECT	Not ECN-capable transport
10	ECT(0)	ECN-capable transport
01	ECT(1)	ECN-capable transport
11	CE	Congestion Experienced

transport only understands drop

transport understands ECN



motivation for change

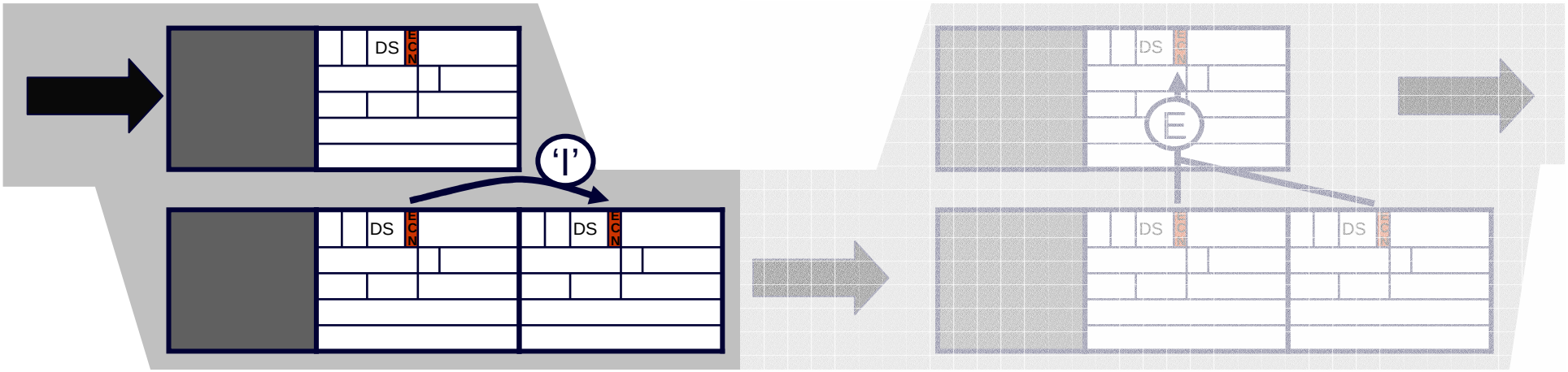
ECN field	codepoint	meaning
00	Not-ECT	Not ECN-capable transport
10	ECT(0)	ECN-capable transport
01	ECT(1)	ECT or low severity congestion
11	CE	Congestion Experienced

- introduce 2 severity levels of congestion (1 level still works too)
  - for pre-congestion notification (PCN – RFC5559)
  - or other alternate uses of the ECN field (RFC4774)

- in RFC4103 (and 3168) ECN propagation restricted to 1 level
  - vestige of earlier covert channel restriction
  - RFC4301 removed restriction from ingress, but not egress

- tunnels and ECN schemes get deployed independently
- should “just work”
  - whatever tunnels happen to intervene, consistent ECN behaviour
  - whatever ECN scheme is in use, tunnels need no config

proposed encap – RFC4301 unchanged



encapsulation at tunnel ingress

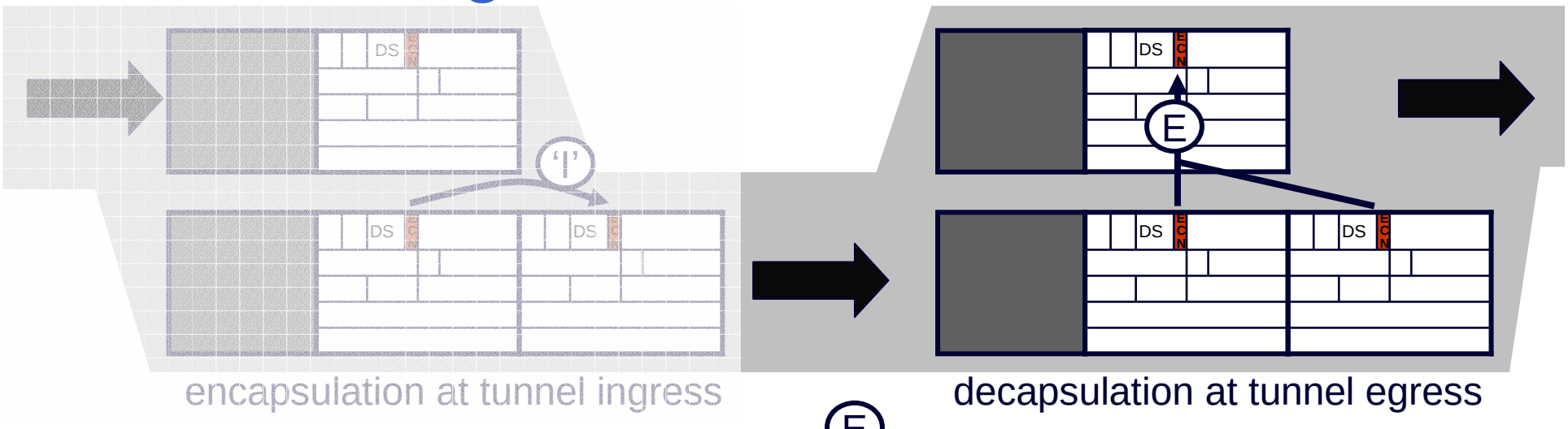
decapsulation at tunnel egress

'I'

incoming header (also = outgoing inner)	outgoing outer		
	RFC3168 ECN limited functionality	RFC3168 ECN full functionality	RFC4301 IPsec
Not-ECT	Not-ECT	Not-ECT	Not-ECT
ECT(0)	Not-ECT	ECT(0)	ECT(0)
ECT(1)	Not-ECT	ECT(1)	ECT(1)
CE	Not-ECT	**ECT(0)**	CE
proposal shown in red	unchanged **compatibility mode** for legacy	**'reset' CE no longer used**	'copy' CE becomes **normal mode** for all IP in IP

- non-IPsec ECN encap brought into line with RFC4301
- required for PCN
- tidies up perversity
  - 4301 decided 2-bit covert channel is manageable
  - IPsec tunnels don't block it
  - non-IPsec tunnels block it

current egress behaviour

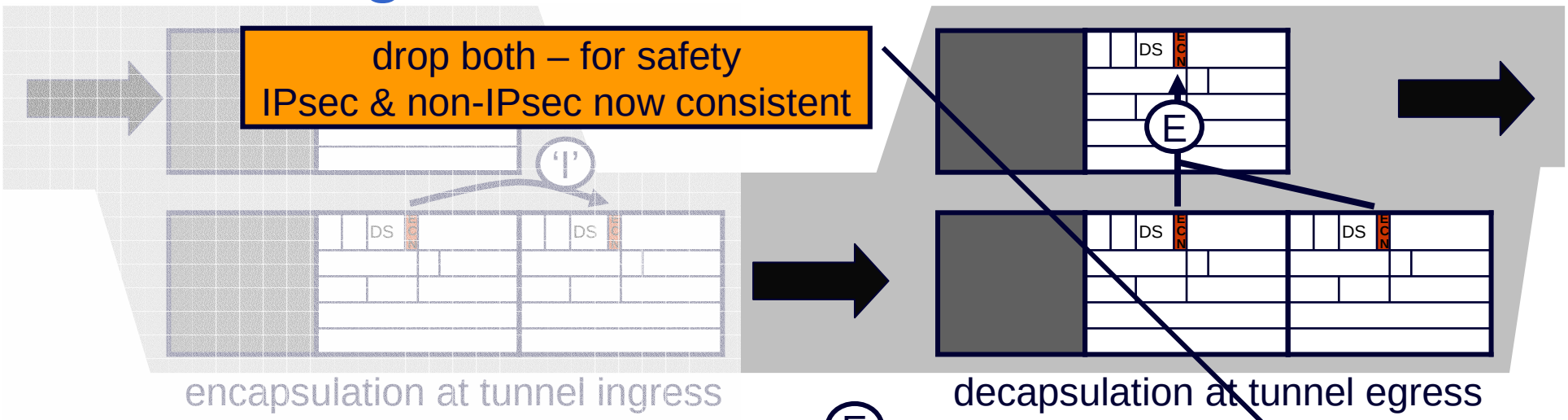


- OK for current ECN
  - 1 severity level of congestion
- any outer changes to ECT(0/1) lost
  - originally to restrict covert channel (but 2-bit now considered manageable)
  - effectively wastes ½ bit in IP header

E

incoming inner	incoming outer			
	Not-ECT	ECT(0)	ECT(1)	CE
Not-ECT	Not-ECT	Not-ECT	Not-ECT	drop / Not-ECT
ECT(0)	ECT(0)	ECT(0)	ECT(0)	CE
ECT(1)	ECT(1)	ECT(1)	ECT(1)	CE
CE	CE	CE	CE	CE
	Outgoing header (RFC4301 \ RFC3168)			

new egress rules



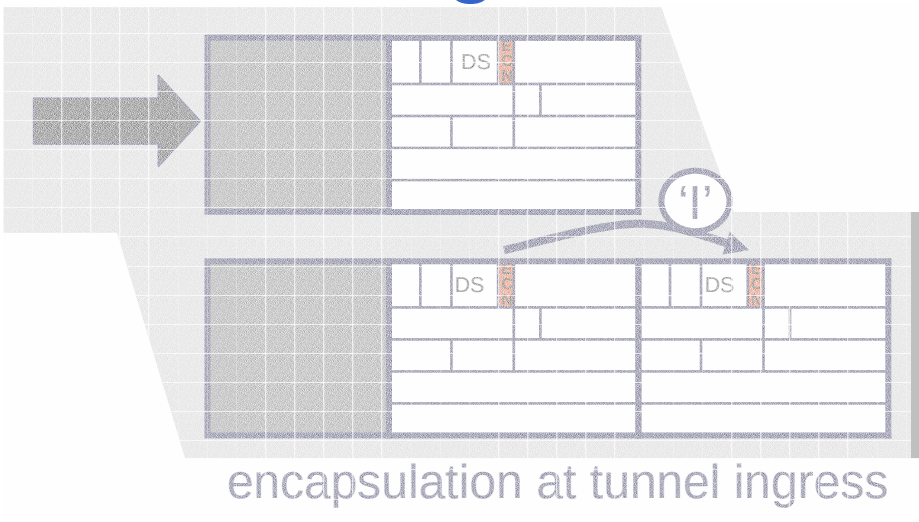
- cater for ECT(1) meaning either more severe or same severity as ECT(0)
  - for PCN or similar schemes that signal 2 severity levels
- drop potentially unsafe unused combinations
  - where congestion marked in outer but inner says transport won't understand

E

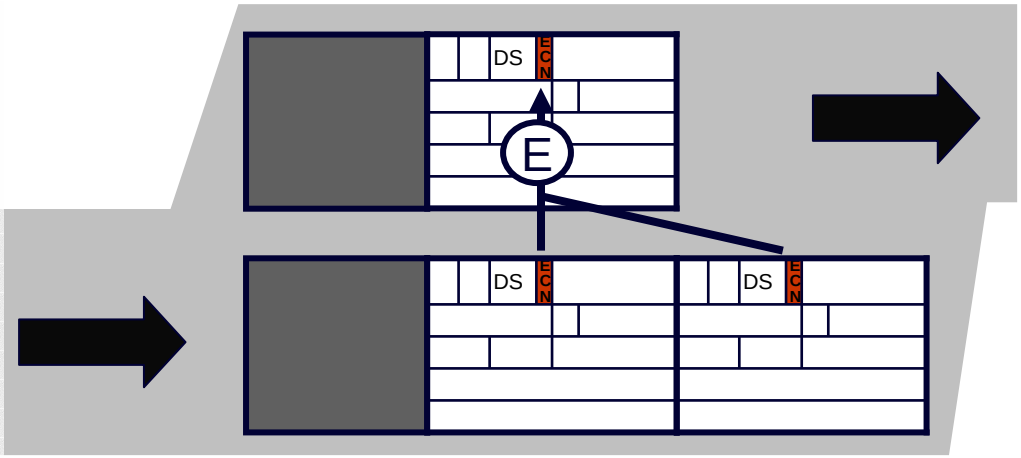
incoming inner	incoming outer			
	Not-ECT	ECT(0)	ECT(1)	CE
Not-ECT	Not-ECT	Not-ECT	**drop**	**drop**
ECT(0)	ECT(0)	ECT(0)	**ECT(1)**	CE
ECT(1)	ECT(1)	ECT(1)	ECT(1)	CE
CE	CE	CE	CE	CE
	Outgoing header (proposed update) **(bold = proposed change for all IP in IP)**			

a change into ECT(1) propagates from outer

new egress rules



encapsulation at tunnel ingress



decapsulation at tunnel egress

- only changing currently unused combinations
  - optional alarms added to all unused combinations
- only tunnels that need the new capability need to comply
  - an update, not a fork
  - no changes to combinations used by existing protocols (backward compatible)

E

incoming inner	incoming outer			
	Not-ECT	ECT(0)	ECT(1)	CE
Not-ECT	Not-ECT	Not-ECT (!!!)	**drop (!!!)**	**drop (!!!)**
ECT(0)	ECT(0)	ECT(0)	**ECT(1) (!!!)**	CE
ECT(1)	ECT(1)	ECT(1) **(!!!)**	ECT(1)	CE
CE	CE	CE	CE **(!!!)**	CE
	Outgoing header (proposed update) **(bold = proposed change for all IP in IP)**			

(**!!!**) = currently unused combination, egress MAY raise an alarm

next steps

- review from security area?
  - pref before tsvwg last call (Nov ’09?)
  - or during tsvwg last call / IESG review

Tunnelling of Explicit Congestion Notification

draft-briscoe-tsvwg-ecn-tunnel-03.txt

Q&A