Tunnelling of Explicit Congestion Notification

draft-briscoe-tsvwg-ecn-tunnel-02.txt

Bob Briscoe, BT

IETF-74 tsvwg Mar 2009

draft-ietf-tsvwg-ecn-tunnel-02.txt
exec summary

Tech changes:

- ingress (no change from -01 draft):
  - brings into line with RFC4301 IPsec
- egress:
  - save two wasted codepoint combinations
  - one proposed at IETF-73:
    generally agreed to go for it
    - needed by PCN but more general
  - one proposed by Anil Agarwal on list
  - both have no backward compatibility issues
    - because they use previously unused codepoint combinations

- Baked: ready for review
  - apologies for late posting
  - complete re-write
  - solely standards action text (17pp)
  - shifted motivation, impact analysis etc to appendices or trash
- Plan
  - list of 6 volunteer reviewers
    - question: all 3 changes ok?
  - socialise in PCN now
  - socialise with IPsec w-g
    once rough concensus in tsvwg (Jul)

Tunnelling of Explicit Congestion Notification

draft-briscoe-tsvwg-ecn-tunnel-02.txt

Bob Briscoe, BT

IETF-74 PCN Mar 2009

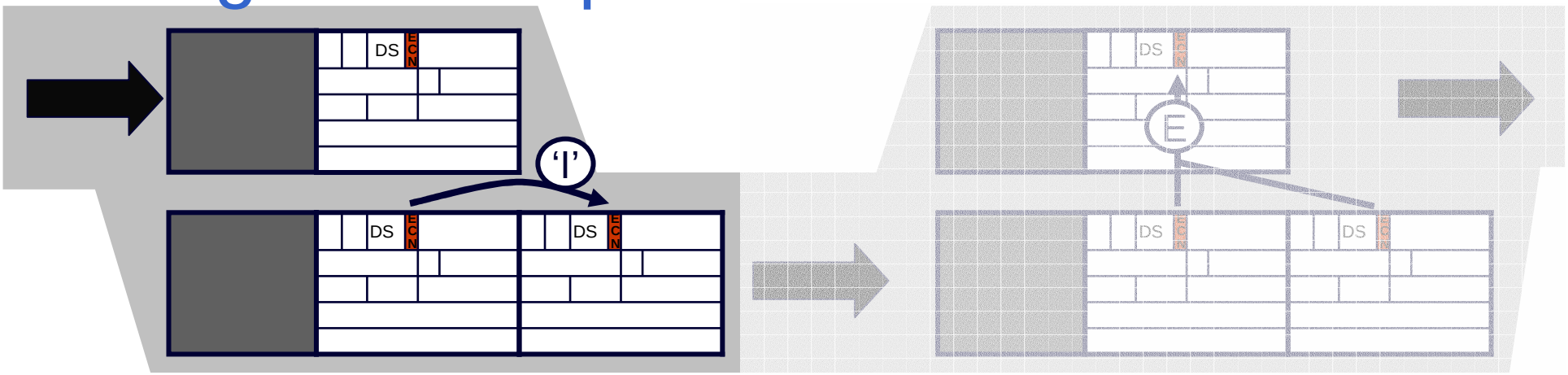
status

- Layered Encapsulation of Congestion Notification
  - **new WG draft:** draft-ietf-tsvwg-ecn-tunnel-02.txt 24 Mar '09
  - **intended status:** standards track
  - **RFC pub target:** ? TBA
  - **immediate intent:** review specifically: fix to decap as well as encap?
  - **w-gs & r-gs affected:** TSVWG, PCN, ICCRG, IPsec, Internet Area?

recap (exec summary)

- scope
  - all IP in IP (v4, v6) tunnels, all DSCPs
  - solely wire protocol processing of tunnelled ECN, not marking or response algorithms
- sequence of standards actions led to perverse position
  - non-IPsec ECN tunnels [RFC3168] have vestige of stronger security than even IPsec [RFC4301] decided was necessary**!**
  - limits usefulness of 3168 tunnels
    - ingress: PCN stds track "excess rate marking" works with 4301 but not 3168
    - egress: PCN 2-level marking lost
      requires complex work-rounds or reduced function
- ingress: bring ECN tunnelling [RFC3168] into line with IPsec [RFC4301]
- egress: use two wasted combinations of inner & outer codepoints
  - absolutely no backwards compatibility issues

ingress recap



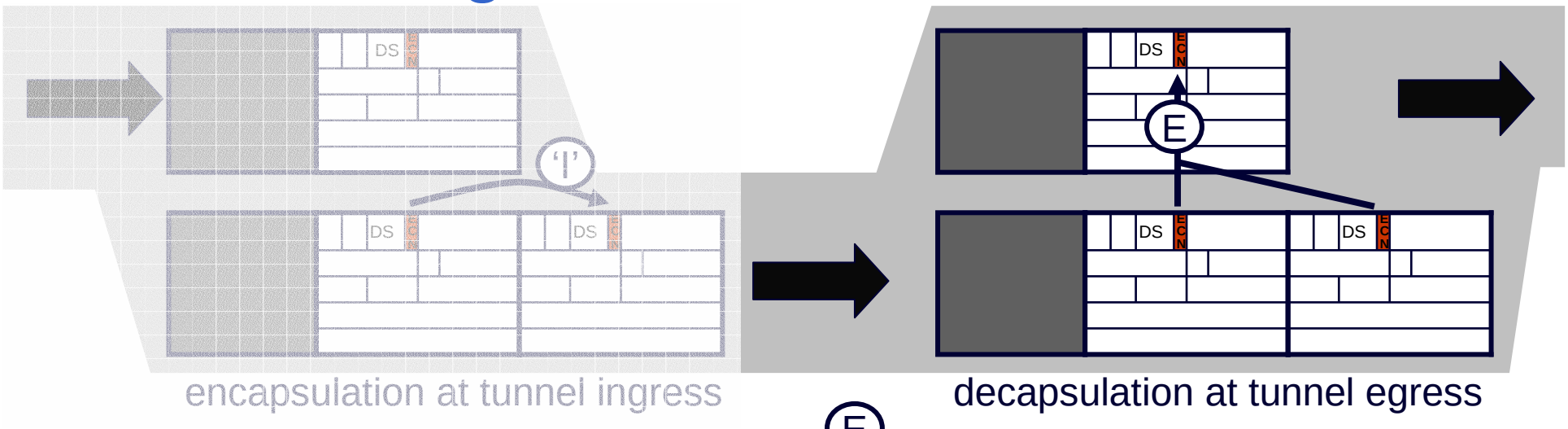
encapsulation at tunnel ingress

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	unchanged **compatibility state** for legacy	**'reset' CE no longer used**	'copy' CE becomes **normal state** for all IP in IP

current egress behaviour



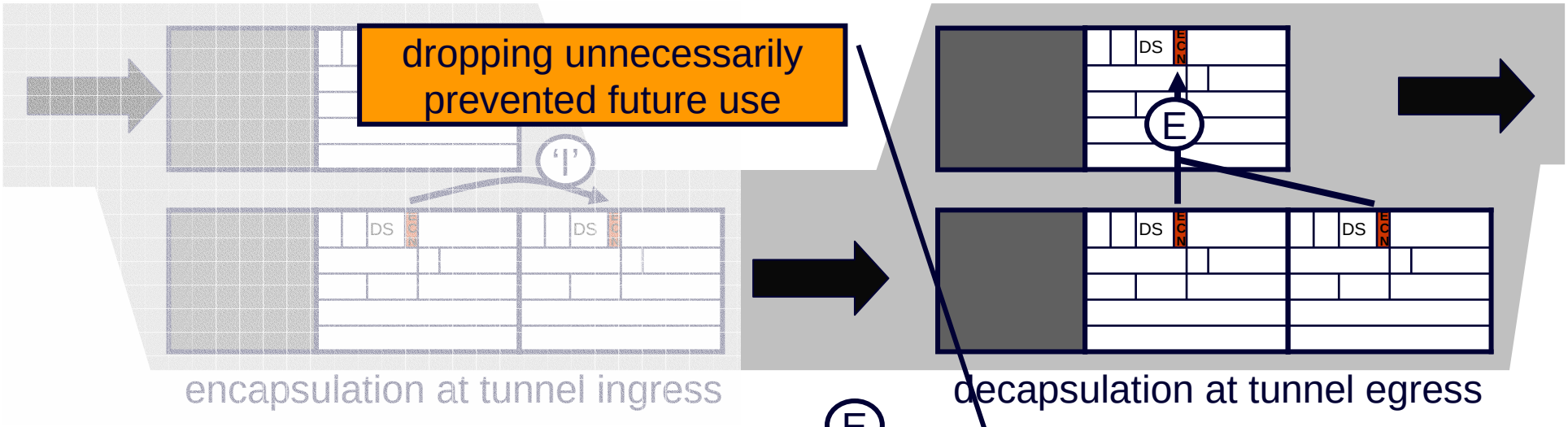
- OK for current ECN
- but any changes to ECT lost
  - effectively wastes ½ bit in IP header
  - again, for safety against marginal threat that IPsec decided was manageable
- PCN tried to use ECT(0/1)
  - but having to waste DSCPs instead
  - or other complex work-rounds
  - or hobbled function

E

incoming inner	incoming outer			
	Not-ECT	ECT(0)	ECT(1)	CE
Not-ECT	Not-ECT	drop (**!!!**)	drop (**!!!**)	drop (**!!!**)
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 (RFC3168 & RFC4301)			

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

new egress rules (appendix in -01, normative in -02)



- no effect on any legacy
  - adds new capability using previously illegal combinations of inner & outer
  - only tunnels that need the new capability need to comply
  - an update, not a fork

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)

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

propagates changed outer

text changes draft-01→ 02

- scope reduced solely to ECN in IP in IP tunnels
  - removed ECN design guidelines for any layered encapsulation (e.g. ethernet)
- changes to egress made normative
  - one was tentative in appendix (proposed last IETF)
  - other suggested by Anil Agarwal on list
- completely restructured and largely rewritten
  - solely standards action text
  - bloat (justification, analysis) removed or shifted to appendices

next steps

- ready for full review now
  - list of 6 volunteers
  - main question: all three changes ok?
  - remember, these are nuances to the behaviour of the neck of the hour-glass
- socialise in PCN
- once rough concensus in tsvwg, socialise in IPsec (Jul)
  - will need to assure IPsec folks that they don't have to change (again)

backward & forward compatibility

ingress		mode		egress: **I-D ecn-tunnel**	RFC 4301	RFC 3168		RFC 2481		RFC 2401/ 2003
				compreh ensive	4301	full	lim	2481	lim?	-
			action	**calc C**	calc B	calc B	inner	calc A	inner	inner
compre-hensive	**I-D.ecn-tunnel**	**normal**	**'copy'**	C	B	B	n/a	n/a	n/a	n/a
		compat	**'zero'**	C	n/a	n/a	inner	inner	inner	inner
'3g IPsec'	RFC4301	4301	'copy'	C	B	B	n/a	n/a	n/a	n/a
ECN	RFC3168	full	'reset CE'	C	n/a	B	n/a	n/a	n/a	n/a
		limited	'zero'	C	n/a	n/a	inner	inner	inner	inner
ECN expt	RFC2481	2481	'copy'?	C	n/a	B	n/a	A	n/a	n/a
		limited?	'zero'	C	n/a	n/a	inner	n/a	inner	inner
'2g IPsec' IP in IP	RFC2401 RFC2003	-	'copy'	C	n/a	n/a	inner	A	inner	broken: loses CE

C: calculation C (more severe multi-level markings prevail)
B: calculation B (preserves CE from outer)
A: calculation A (for when ECN field was 2 separate bits)
inner: forwards inner header, discarding outer
n/a: not allowed by configuration

Tunnelling of Explicit Congestion Notification

draft-briscoe-tsvwg-ecn-tunnel-02.txt

Q&A