# Layered Encapsulation of Congestion Notification

draft-briscoe-tsvwg-ecn-tunnel-00.txt

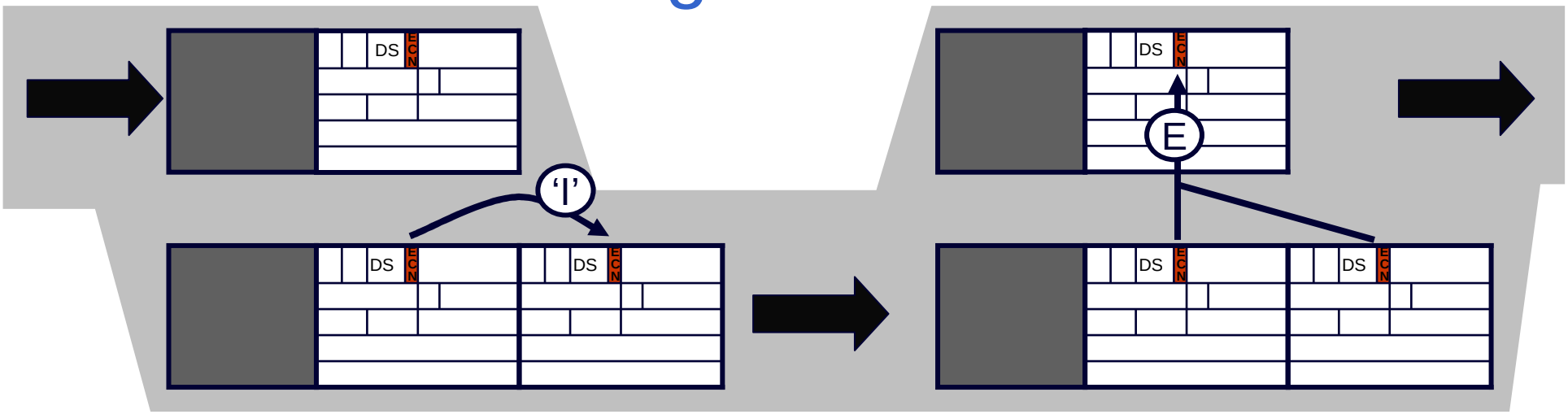
**Bob Briscoe**, BT

IETF-69 tsvwg Jul 2007

# initial draft

- Layered Encapsulation of Congestion Notification
  - **initial draft:** draft-briscoe-tsvwg-ecn-tunnel-00.txt
  - **intended status:** standards track
  - **immediate intent:** move to WG item<br>discuss widening scope
- exec summary
  - propose to update RFC3168 ECN tunnel behaviour for all IP in IP
    - only wire protocol processing, not marking or response algorithms
  - to bring into line with new RFC4301 IPsec ECN behaviour
  - defines default tunnel processing of ECN field for all Diffserv PHBs
    - but also gives guidance on alternatives for specific PHBs (e.g. PCN) and for specific link encapsulations (e.g. MPLS)

# one main change to RFC3168 ECN




encapsulation at tunnel ingress

decapsulation at tunnel egress

'I'

| incoming header | outgoing outer | | | | |
|---|---|---|---|---|---|
| | RFC3168 ECN limited functionality | RFC3168 ECN full functionality | RFC430 1 IPsec | **proposed** all IP in IP compatibility mode | **proposed all IP in IP** normal mode |
| Not-ECT | Not-ECT | Not-ECT | Not-ECT | Not-ECT | Not-ECT |
| ECT(0) | Not-ECT | ECT(0) | ECT(0) | Not-ECT | ECT(0) |
| ECT(1) | Not-ECT | ECT(1) | ECT(1) | Not-ECT | ECT(1) |
| CE | Not-ECT | **ECT(0)** | **CE** | Not-ECT | **CE** |

**'reset CE'** **'copy CE'**

# why update ECN RFC3168 now?

- despite everyone's best intentions
  - unfortunate sequence of standards actions led to a perverse position..
  - 2001: ECN RFC3168
    - IETF Security Area were concerned about covert channels
    - so RFC3168 didn't copy CE at ingress for IPsec
    - for consistency, also didn't copy CE for non-IPsec tunnels
  - 2005: RFC4301 IPsec
    - Security Area decided 2-bit ECN covert channels can be managed
    - RFC4301 IPsec now copies CE at ingress
- non-IPsec tunnels left not copying CE at ingress
  - lost consistency between IPsec & non-IPsec
  - vestige of security no longer used by IPsec now limits usefulness of non-IPsec tunnels
- copying of whole ECN field at tunnel ingress is more straightforward
- PCN & ECN in MPLS currently being defined; simply copying ECN
  - update RFC3168 now, so all consistent: IPsec, non-IPsec, PCN, MPLS

# widen scope of draft?

- PCN will probably do 2-level congestion marking
  - will require different rules at tunnel egress
  - should we try to make all tunnels consistent with that too?
- while we're updating guidance on ECN tunnelling
  - should we also update guidance on Diffserv tunnelling?

discuss (here or on tsvwg list)

- no time for (spare slides)...
  - exception to tunnel ingress copying CE
  - minor changes at egress (corner case & simplification: single mode)
    - tried really hard not to change IPsec behaviour (except corner cases)
  - guidance for alternative congestion control

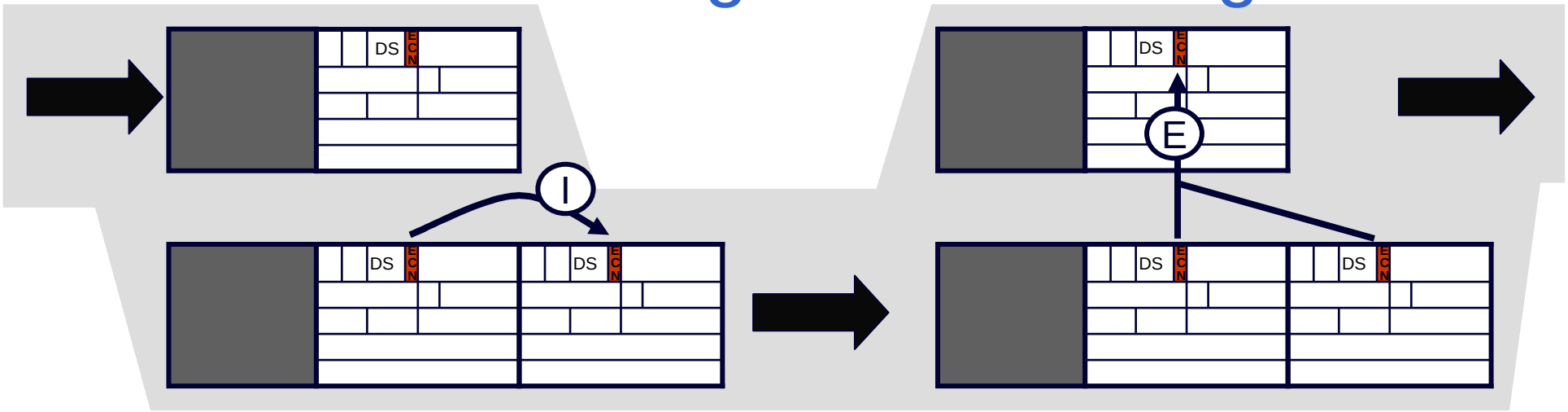
please read & review draft

# Layered Encapsulation of Congestion Notification

draft-briscoe-tsvwg-ecn-tunnel-00.txt

# Q&A

# also minor changes at tunnel egress




encapsulation at tunnel ingress

decapsulation at tunnel egress

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 full & RFC4301) **(bold red = proposed for all IP in IP)** | | | | |

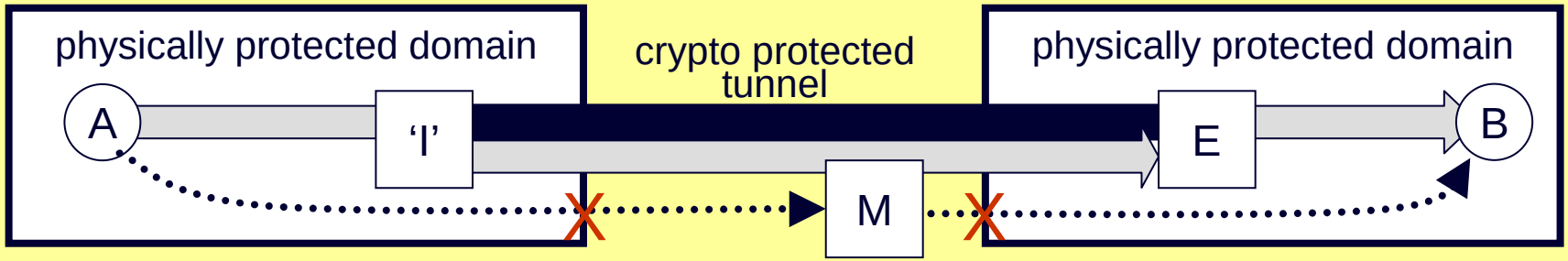
(!!!) = illegal transition, E MAY raise an alarm

- **propose only one mode at egress**
  - **limited functionality mode no longer necessary at E**

conflicting design constraints

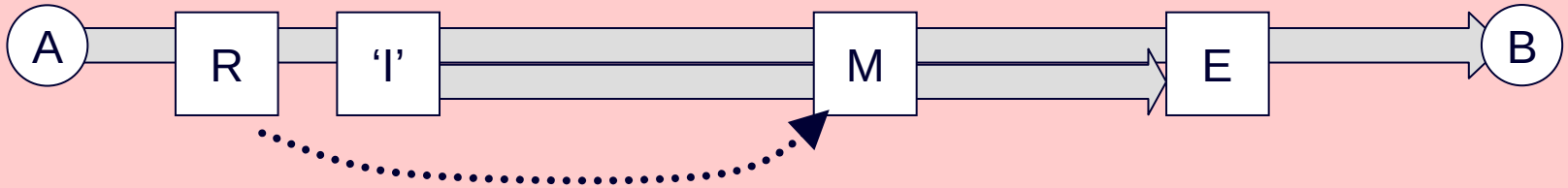
# security vs. management & control

- information security constraint (lesser known IPsec reqm't)




  - I can prevent covert channel A→M with encryption
  - E an prevent covert channel M→B with integrity checking

- tunnel ingress control / management constraints



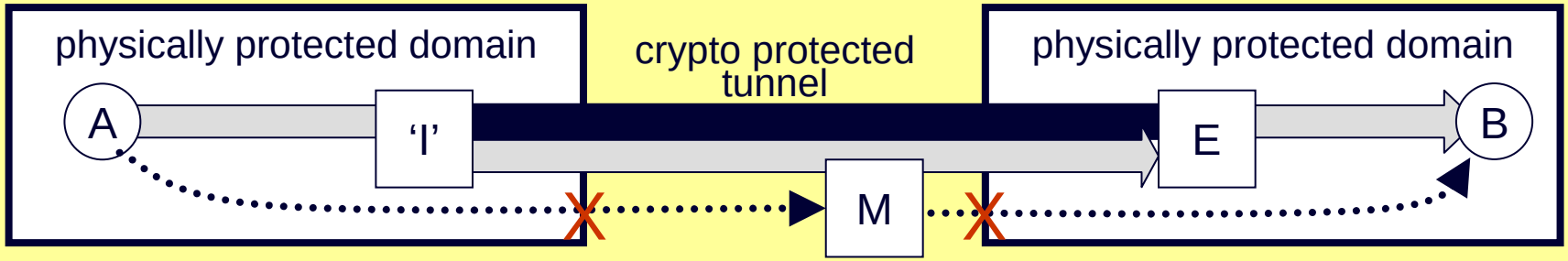

  - marking algorithm at M may depend on prior markings (since A)
    - e.g. a number of PCN marking proposals work this way
  - M may need to monitor congestion since A
    - e.g. if M is monitoring an SLA at a border

- IPsec crypto cannot cover mutable fields (ECN, DS & TTL)
  - if 'I' copies ECN CE, it opens up 2-bit covert channel A→M or R→M

conflicting design constraints

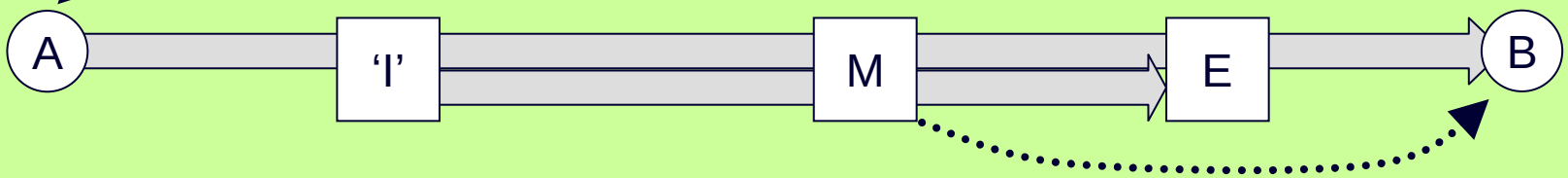
# security vs. congestion control

- information security constraint (lesser known IPsec reqm't)




  - I can prevent covert channel A→M with encryption
  - E an prevent covert channel M→B with integrity checking

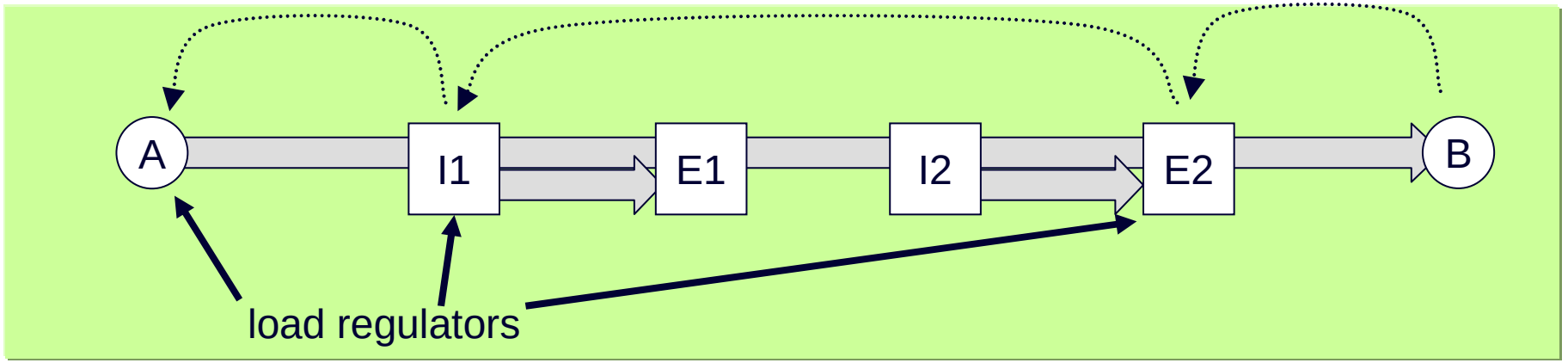
- tunnel egress control constraint
  - explicit congestion notification control channel M→B→A




- IPsec crypto cannot cover mutable fields (ECN, DS & TTL)
  - if E copies ECN CE, it opens up 2-bit covert channel M→B

exception
# in-path load regulators

- typically load regulation at source A (e2e principle)
- reasonable in-path load regulator proposals exist
  - e.g. PCN admission control (& PWE3?)




  - new normal rule for tunnel ingress (e.g. I2)
    - copy CE to outer header
  - exception if ingress also in-path load regulator (I1)
    - copy ECN to outer header but reset CE to ECT(0)