Google

ConEx Concepts and Abstract Mechanism

draft-ietf-conex-abstract-mech-07.txt

Matt Mathis, Google

Bob Briscoe, BT

IETF-87 ConEx Jul 2013

Bob Briscoe's contribution is partly funded by trilogy 2, a research project supported by the European Community www.trilogy2.org

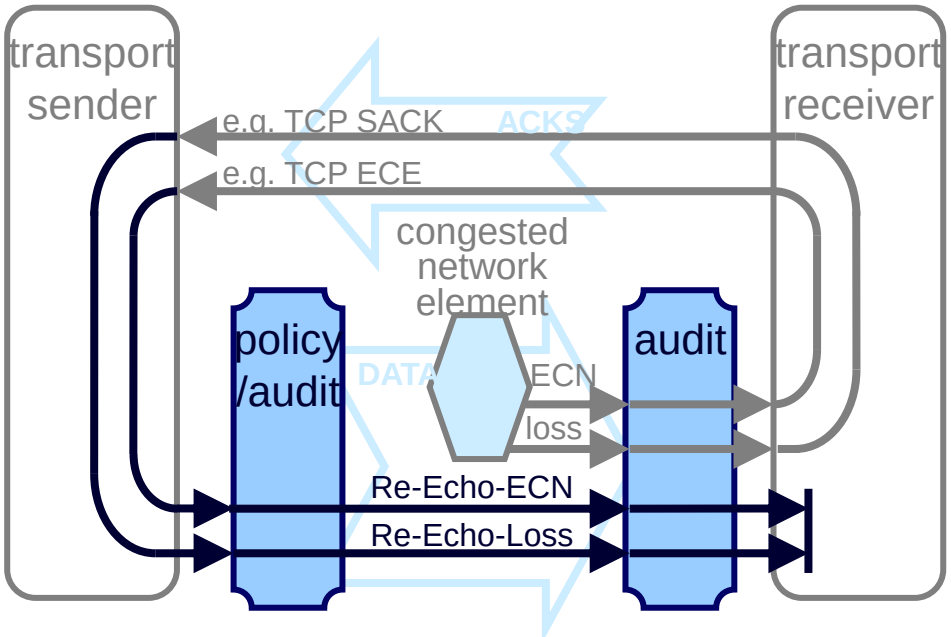
BT

ConEx Concepts and Abstract Mechanism

- **working group draft:** draft-ietf-conex-abstract-mech-07.txt
- **intended status:** informational
- **immediate intent:** minor rev to -08 this week, then WGLC
- **milestone target:** Jul 2011

recall

- abstract design of algorithms & protocol: TCP & IP encoding follows
- scope
  - loss-based and ECN
  - any transport
  - the structure of audit



normative improvements to draft (I)
deleted a ‘pious’ requirement on other protocols

3.1. Requirements for ConEx Signals

c. The ConEx signal SHOULD be timely. There will be a minimum delay of one RTT, and often longer if the transport protocol sends infrequent feedback (consider RTCP [RFC3550] for example). ~~This delay complicates auditing, and SHOULD be minimized.~~

normative improvements to draft (II)

consolidated network protocol requirements

3.3. Requirements for non-abstract ConEx specifications

An experimental ConEx specification SHOULD describe the following protocol details:

Network Layer:

A. The specific ConEx signal encodings with packet formats, bit fields and/or code points;

B. An inventory of invalid combinations of flags or invalid codepoints in the encoding. Whether security gateways should normalise, discard or ignore such invalid encodings, and what values they should be considered equivalent to by ConEx-aware elements;

C. An inventory of any conflated signals or any other effects that are known to compromise signal integrity;

D. Whether the source is responsible for allowing for the round trip delay in ConEx signals (e.g. using a Credit marking), and if so whether Credit is maintained for the duration of a flow or degrades over time, and what defines the end of the duration of a flow;

E. A specification for signal units (bytes vs packets, etc), any approximations allowed and algorithms to do any implied conversions or accounting;

F. If the units are bytes a definition of which headers are included in the size of the packet;

G. How tunnels should propagate the ConEx encoding;

H. Whether the encoding fields are mutable or not, to ensure that header authentication, checksum calculation, etc. process them correctly. A ConEx encoding field SHOULD be immutable end-to-end, then end points can detect if it has been tampered with in transit;

I. if a specific encoding allows mutability (e.g. at proxies), an inventory of invalid transitions between codepoints. In all encodings, transitions from any ConEx marking to Not-ConEx MUST be invalid;

J. A statement that the ConEx encoding is only applicable to unicast and anycast, and that forwarding elements should silently ignore any ConEx signalling on multicast packets (they should be forwarded unchanged)

K. Definition of any extensibility;

L. Backward and forward compatibility and potential migration strategies. In all cases, a ConEx encoding MUST be arranged so that legacy transport senders implicitly send Not-ConEx;

M. Any (optional) modification to data-plane forwarding dependent on the encoding (e.g. preferential discard, interaction with Diffserv, ECN etc.);

N. Any warning or error messages relevant to the encoding.

black: no change

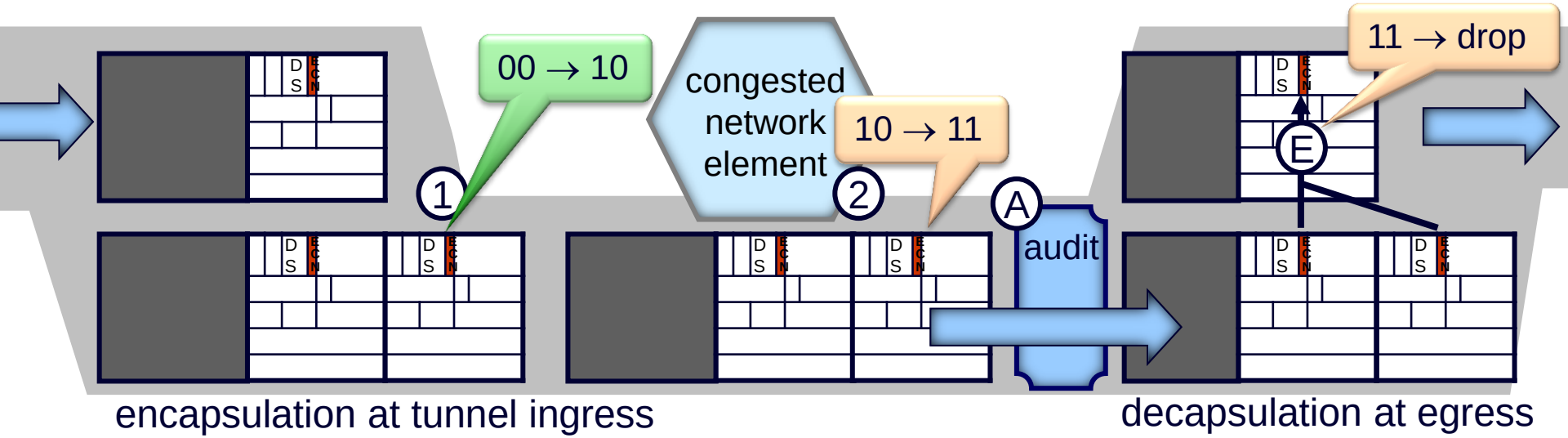
green: normative text elsewhere made lower case, and consolidated into this list by ref

amber: new

technical improvements to draft

added unilateral deployment technique for audit

- exploits a side-effect of standard tunnelling (IP-in-IP or any ECN link encap)



- even for e2e transports that don't support ECN, the operator can:
  - 1 at encap: alter 00 to 10 in outer
  - 2 at interior buffers: turn on ECN
- defers any drops until egress (E)
- audit (A) just before egress can see packets to be dropped:
  - CE outer + Not-ECT inner (D)

recap of standard ECN decap [RFC6040, RFC3168]

	incoming inner (E)	incoming outer			
		Not-ECT	ECT(0)	ECT(1)	CE
00	Not-ECT	Not-ECT	Not-ECT	Not-ECT	**drop** (D)
10	ECT(0)	ECT(0)	ECT(0)	ECT(1)	CE
01	ECT(1)	ECT(1)	ECT(1)	ECT(1)	CE
11	CE	CE	CE	CE	CE
		Outgoing header			

Editorial mods

2. Replaced detail in Overview with forward ref to body

Preserved the text on flow-state and byte-pkt, just moved it

4.4. Encoding ConEx: Independent Bits

Added “A packet with ConEx set combined with all the three other flags cleared implies ConEx-Not-Marked”

5.5. Audit

“Generic loss auditing ... not believed to be possible” moved from last bullet to first

5.5.1. Using Credit to Simplify Audit:

Added sentence on the need to specify whether credit expires etc in a specific encoding doc.

5.4.3. Congestion Policers

Referred to [I-D.briscoe-conex-policing] instead of an academic paper

6. Support for Incremental Deployment

Moved “A network operator can create incentives for senders...” from senders bullet to networks bullet (and referred to it from senders as well).

8. Security Considerations

It is planned to document all known attacks and their defences (including all the above) in the RFC series against a concrete ConEx protocol specification. In the interim, [Refb-dis] and its references should be referred to for details and ways to address these attacks in the case of re-ECN.

items for next -08 rev

5. Audit

New text (suggested by Mirja) on why its OK for audit to ignore Not-ConEx packets (because only policy devices can deal with Not-ConEx), and discuss implications in the case of loss.

9. Acknowledgements

Added Ingemar Johansson and David Wagner, but ooops!...

missed ack for an earlier review by Marcelo

status & plans

- Thanks for additional review (esp. Mirja)
- Feels very ready for second WGLC
  ... once -08 posted

ConEx Concepts and Abstract Mechanism

draft-ietf-conex-abstract-mech-07.txt

Q&A