



# congestion exposure BoF candidate protocol: re-ECN

Bob Briscoe  
Chief Researcher, BT  
Nov 2009

This work is partly funded by Trilogy, a research project supported by the European Community  
[www.trilogy-project.org](http://www.trilogy-project.org)

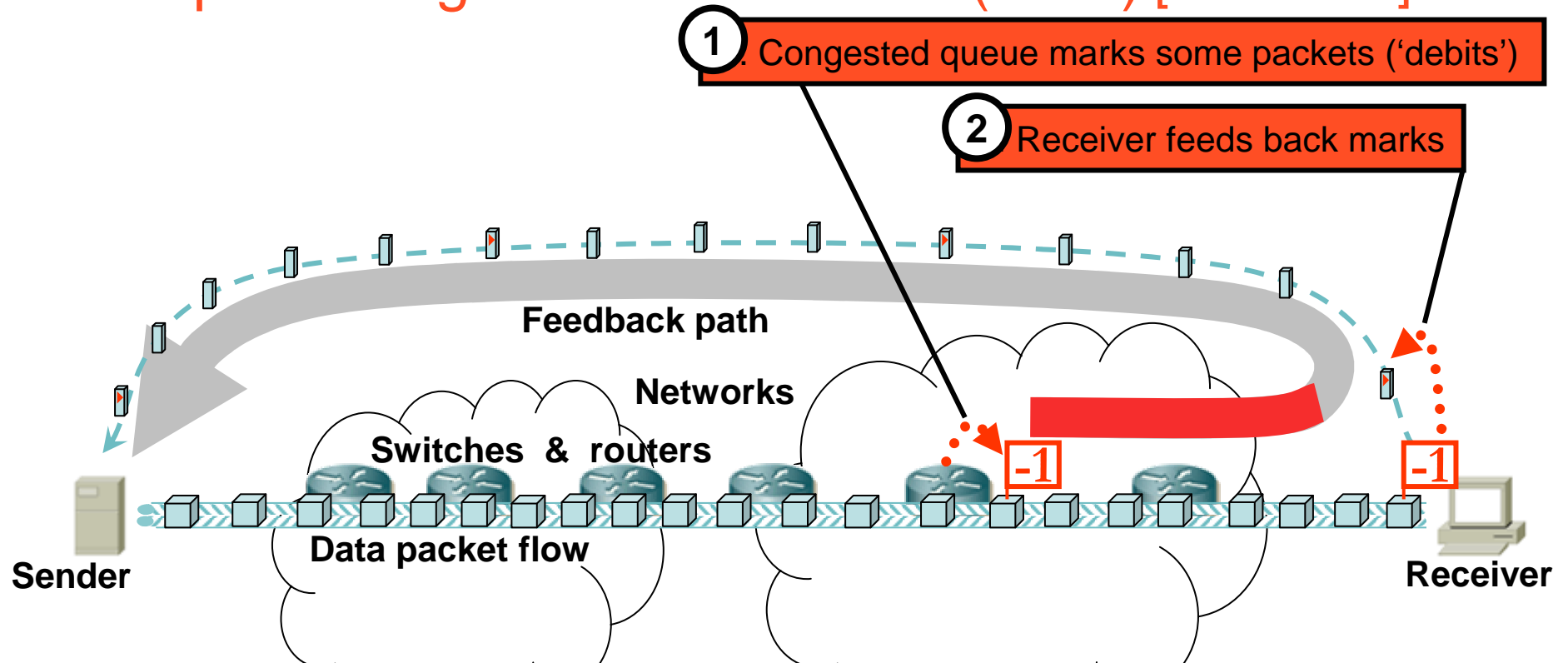


This work is investigative.  
It does not yet indicate the  
direction of BT's production architecture.

## goals

- network can measure contribution to congestion as easily as it measures volume today
- metric for neutral but sufficient capacity sharing
- Internet designed so endpoints deal with congestion
- endpoints expose congestion in packets to network
- purpose of this talk
  - one protocol exists & implemented (x2) – concrete
  - not asking BoF to bless this solution – a strong contender

# congestion exposure uses drop or explicit congestion notification (ECN) [RFC3168]

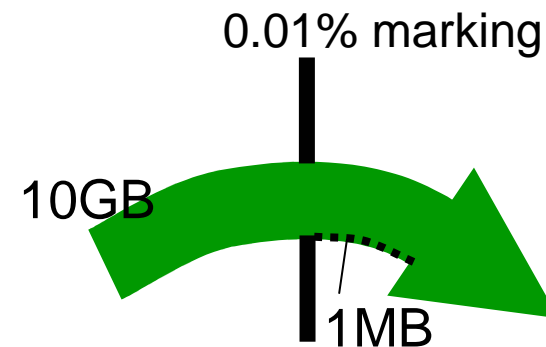
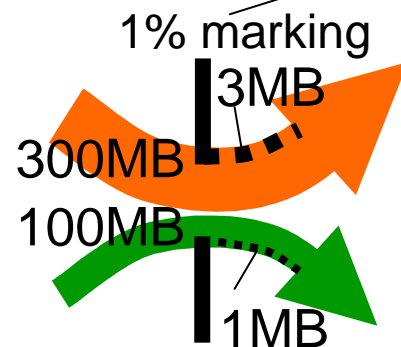
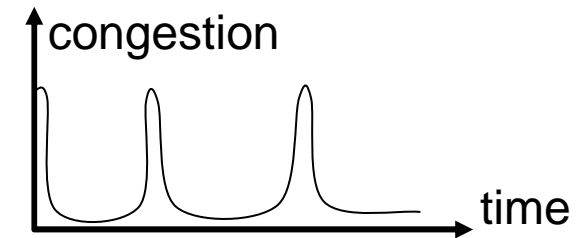
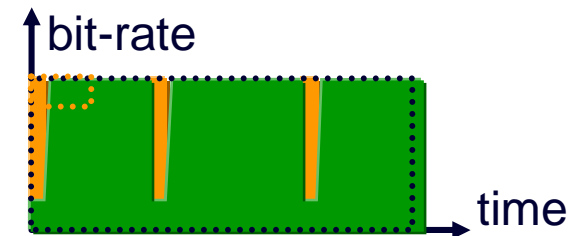


congestion signal *without* impairment

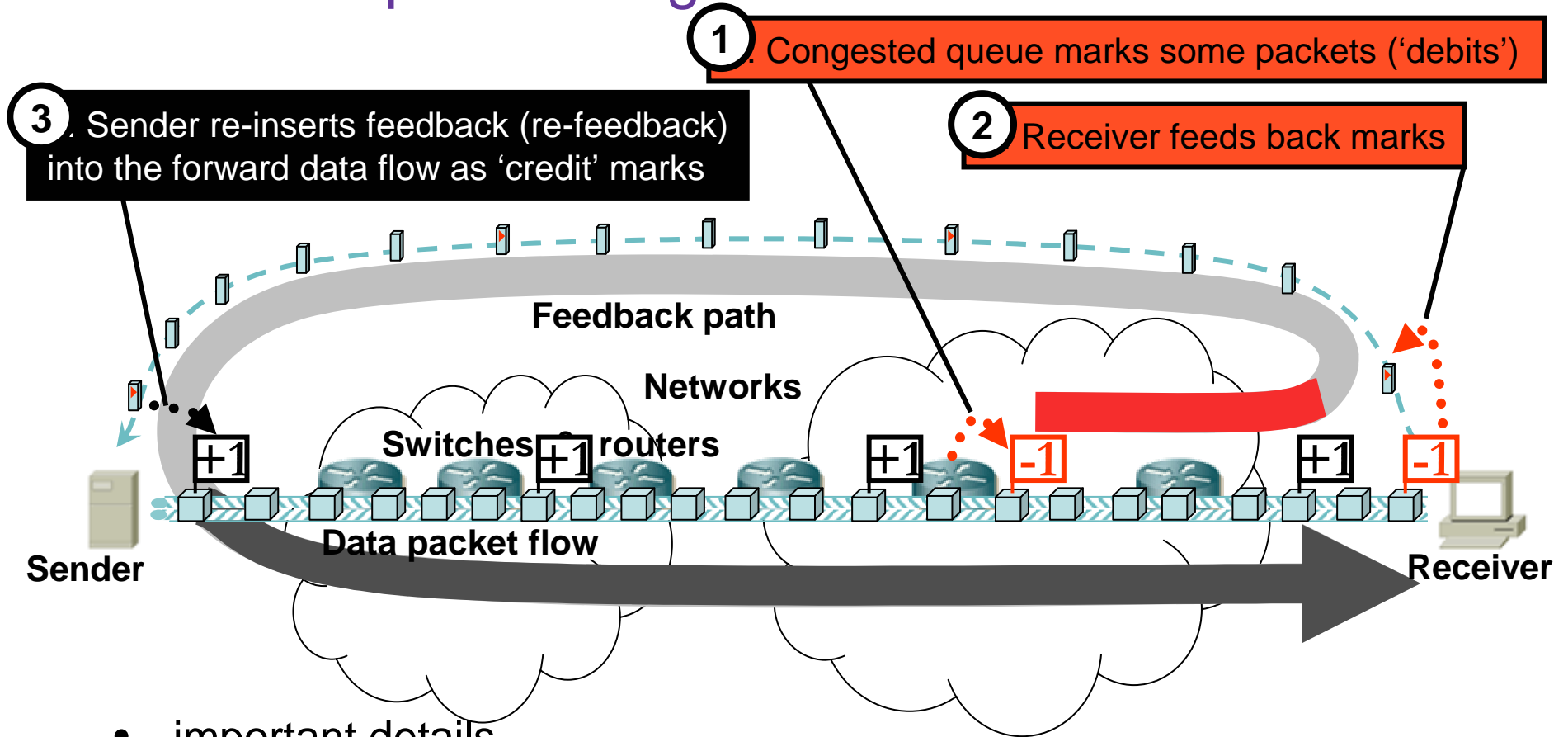
- then tiny queuing delay and tiny tiny loss for all traffic
- no need to avoid congestion to prevent impairment
- whether core, access or borders

# measuring contribution to congestion

- user's contribution to congestion  
= bytes marked
- can transfer very high volume
  - but keep congestion-volume very low
  - similar trick for video streaming

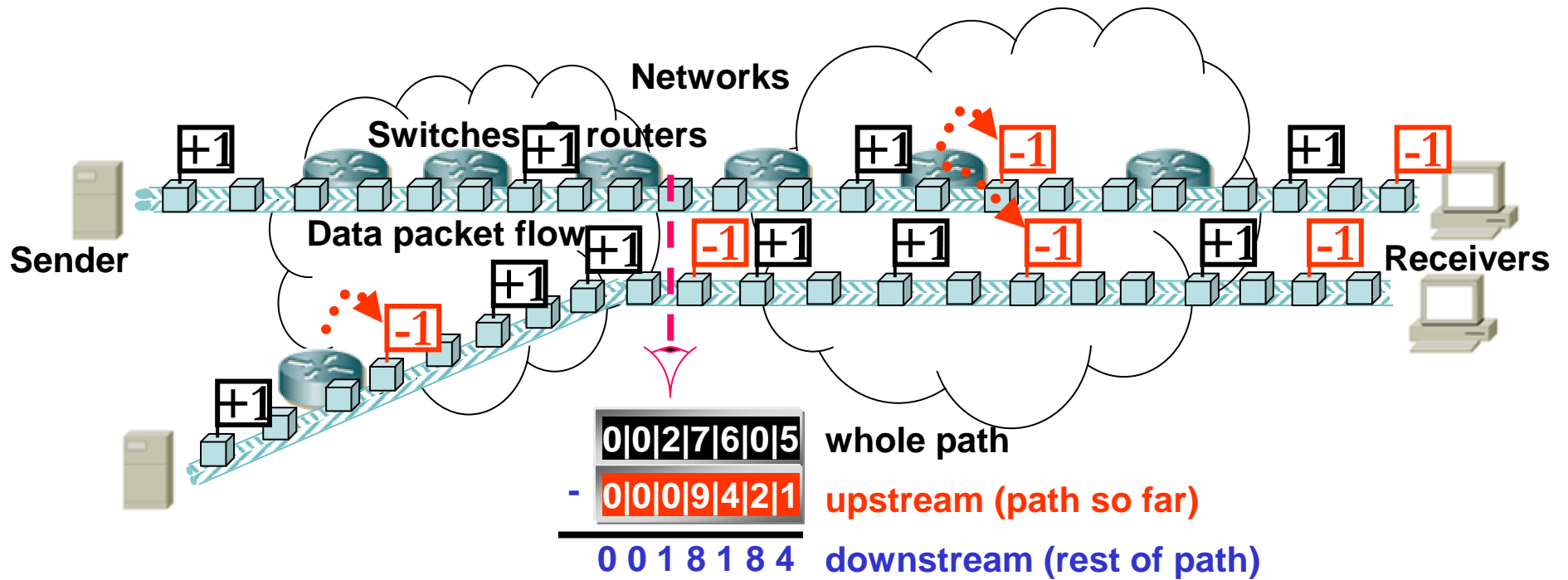


re-inserted feedback (re-feedback) = re-ECN  
sender exposes congestion to network



- important details
  - bootstrap: send no less credit than likely debit in 1 RTT
  - sender re-inserts feedback whether triggered by ECN or loss
- no changes required to IP or MPLS data forwarding

# packets expose congestion over rest of path from wherever you look at them

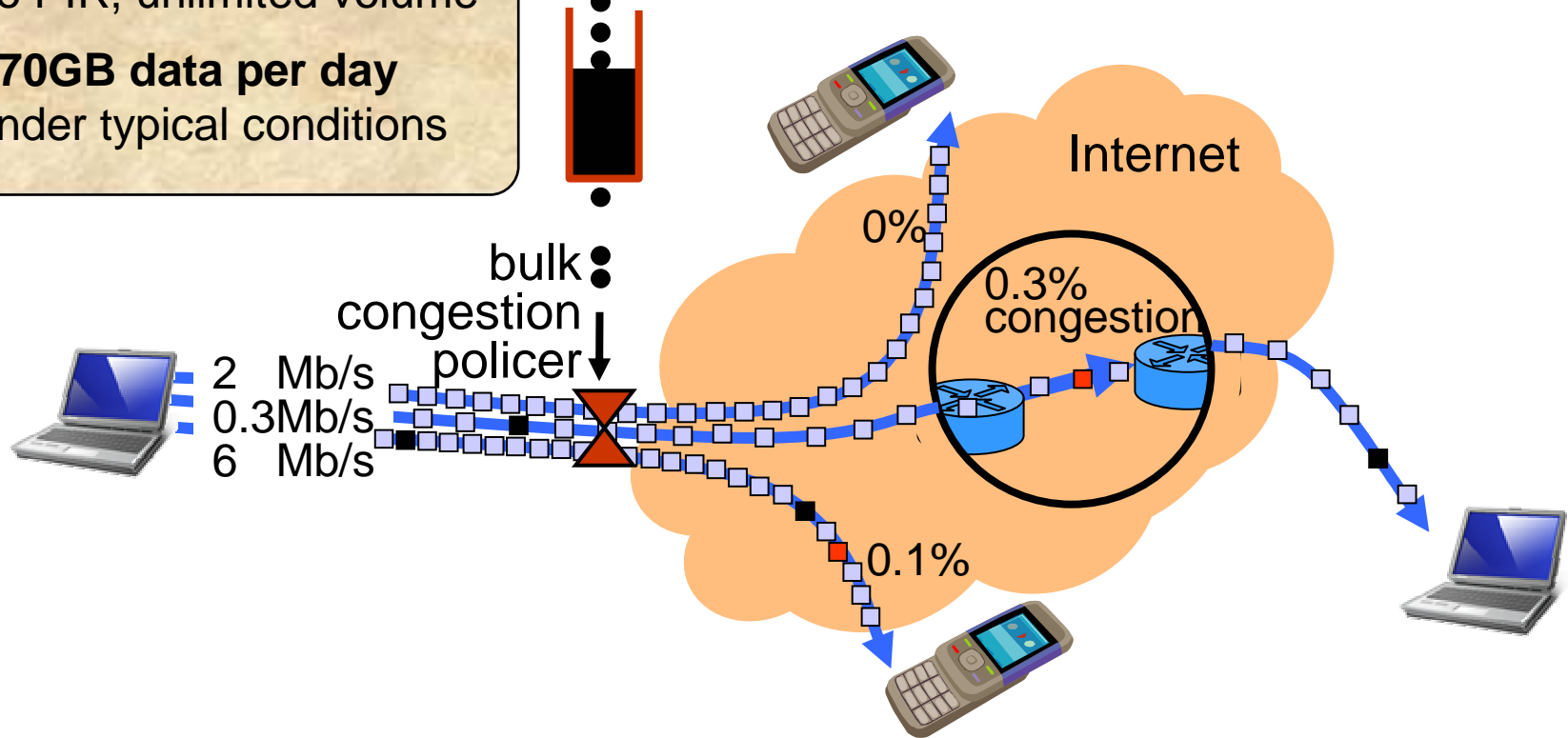


# bulk congestion policing

## ④ example use of ConEx

**Acceptable Use Policy**  
'congestion-volume'  
allowance: 35MB/day  
no other limits needed;  
no PIR, unlimited volume  
~70GB data per day  
under typical conditions

- *not* proposing this for standardisation
  - but need models like this to be possible

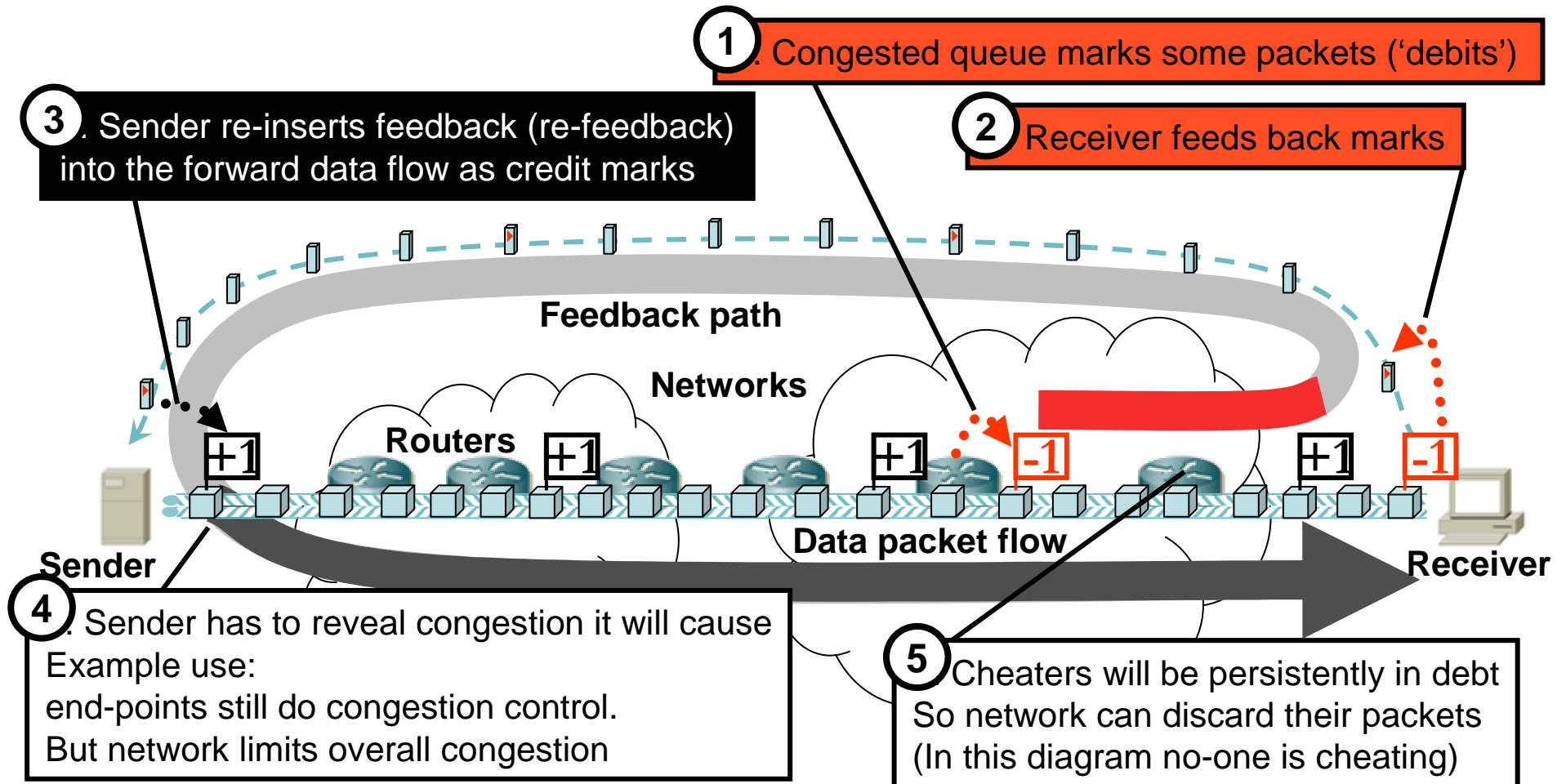


## no time for other potential uses...

- see motivation draft & papers for...
  - bulk congestion policing (or per flow)
  - DDoS mitigation
  - e2e QoS, all within best efforts, with no flow signalling
  - relaxes unnecessary constraints on transport design
  - self-admission control
  - server / middlebox flow state exhaustion control
  - wholesale & interconnect SLAs
- more speculative
  - inter-domain traffic engineering?
  - all-optical interconnects more feasible?
  - replaces multiple access in shared access networks?



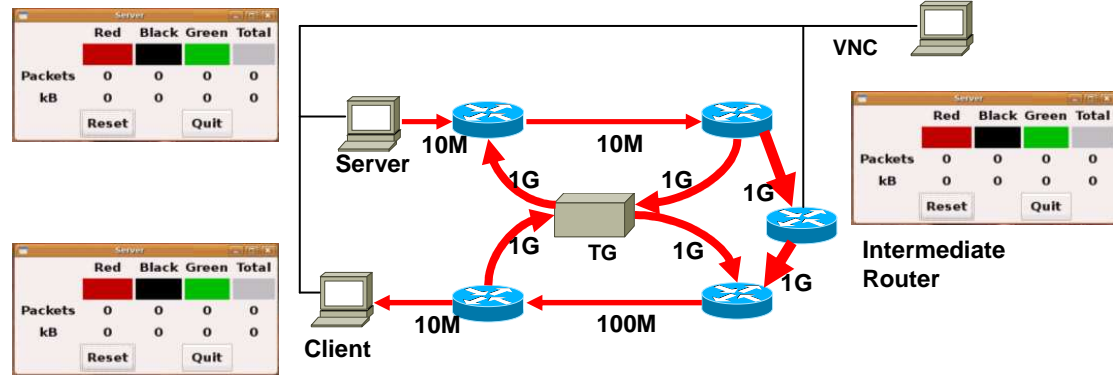
# why won't sender under-expose congestion?



(5) cheat detection: haven't been able to avoid per-flow state

- but designed so flow state does not break shared fate principle
- agnostic to flow *behaviour* – just checks diff between 2 numbers per flow

## re-ECN status



- relatively stable draft of spec in IPv4&6
  - with TCP as transport – exemplar & full spec
- two independent prototype implementations (Linux)
  - quick simple demo afterwards
- *ns-2* implementation
- full security analysis
  - resisted several perverse research community attacks
- Global Info Infrastructure Commission analysis
  - public policy
  - commercial
  - technical feasibility



## goals

- network can measure contribution to congestion as easily as it measures volume today
- metric for neutral but sufficient capacity sharing
- purpose of this talk
  - one protocol exists & implemented (x2) – concrete
  - not asking BoF to bless this solution – a strong contender

congestion exposure BoF

# candidate protocol: re-ECN



[<draft-briscoe-tsvwg-re-ecn-tcp>](#)  
[<draft-briscoe-tsvwg-re-ecn-tcp-motivation>](#)

re-ECN & re-feedback project page:  
[<http://bobbriscoe.net/projects/refb/>](http://bobbriscoe.net/projects/refb/)

## Q&A

