



Internet capacity sharing for packets not flows

Bob Briscoe
Chief Researcher, BT
Jul 2009

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

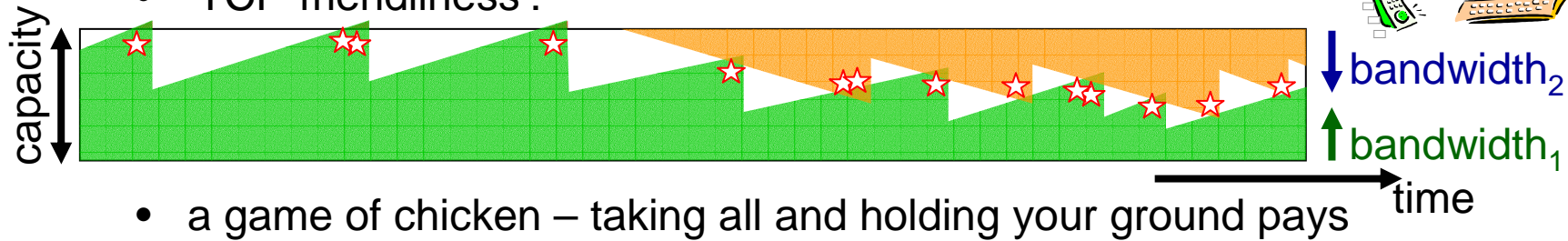
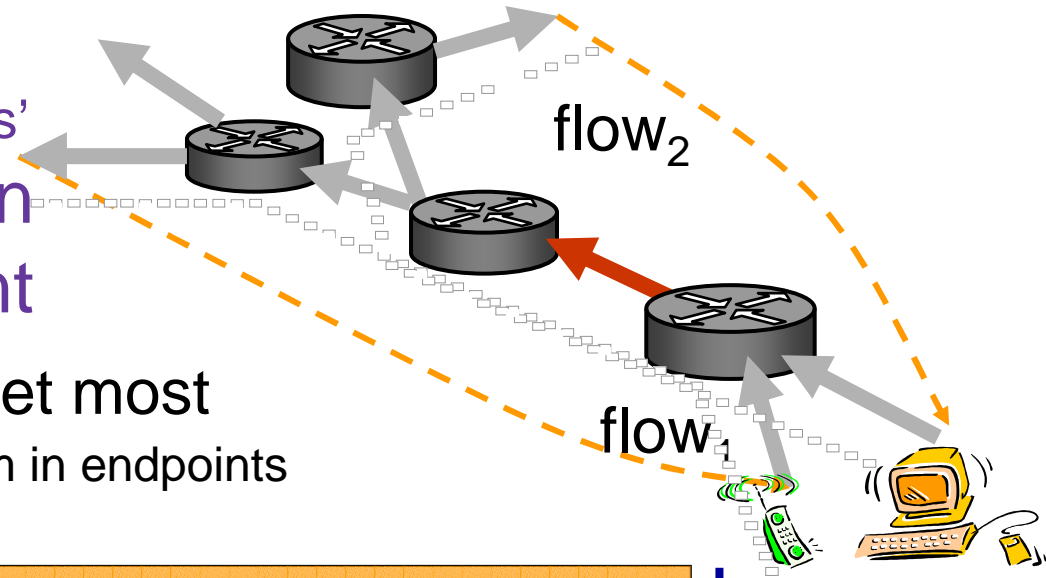


how to share a bandwidth cloud?

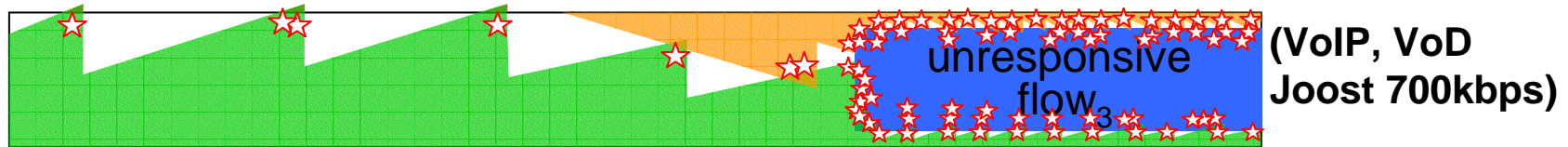
- transport area consensus reversed since 2006
 - ‘TCP-friendly’ was useful, but not a way forward
 - rewrite of IETF capacity sharing architecture in process
 - commercial/policy review in process driven by ‘captains of industry’
- TCP’s dynamic response to congestion is fine
 - but the way it shares capacity is very wrong
- ISP's homespun alternatives have silently overridden TCP
 - result: blocks, throttles & deep packet inspection
 - if it’s new, it won’t get through (if it’s big, it won’t either)
- “good fences make good neighbours;” IETF challenge:
 - protocols for good fences, before industry builds bad ones
- accept: transport protocols don’t do fairness (not on their own)
 - new challenge: liberal but effective capacity sharing function?
 - capacity sharing for packets not flows

how Internet sharing 'works' endemic congestion & voluntary restraint

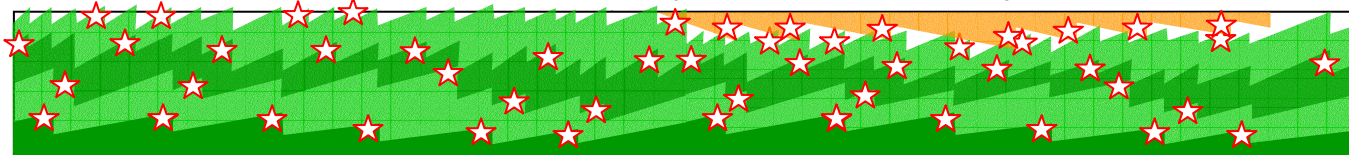
- those who take most, get most
 - voluntarily polite algorithm in endpoints
 - 'TCP-friendliness':



- a game of chicken – taking all and holding your ground pays

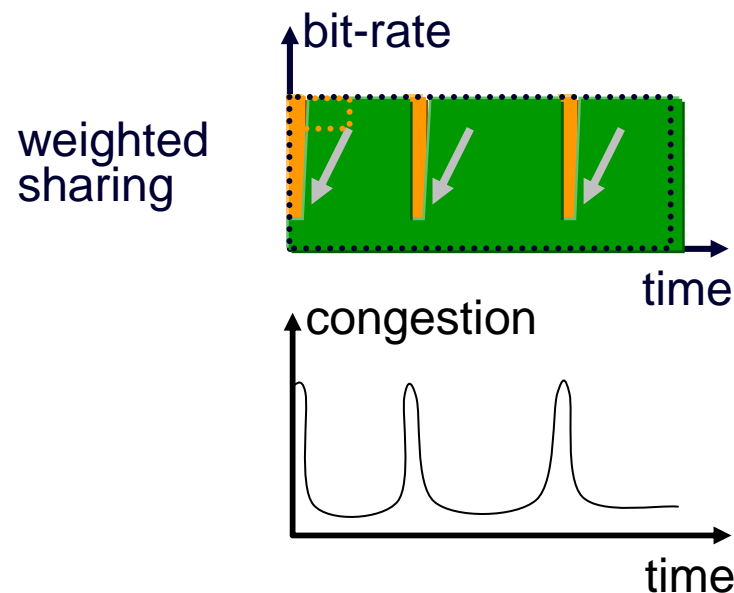
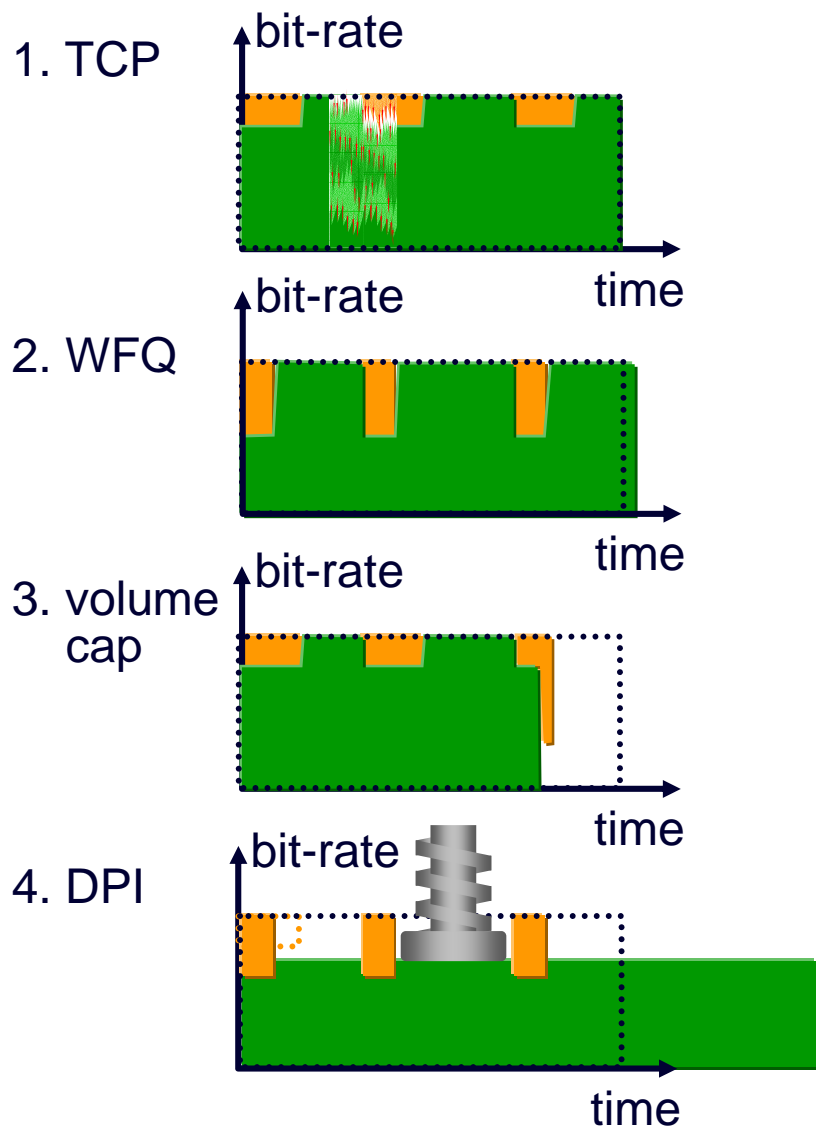


- or start more 'TCP-friendly' flows than anyone else (**Web: x2, p2p: x5-100**)



- or for much longer than anyone else (file transfer x200)
- net effect of both (p2p: x1,000-20,000 higher traffic intensity)

none of the above harness end-system flexibility



- light usage can go much faster
- hardly affects completion time of heavy usage

NOTE: weighted sharing doesn't imply differentiated network service

- just weighted aggressiveness of end-system's rate response to congestion cf. LEDBAT

powerful resource accountability metric

congestion-volume

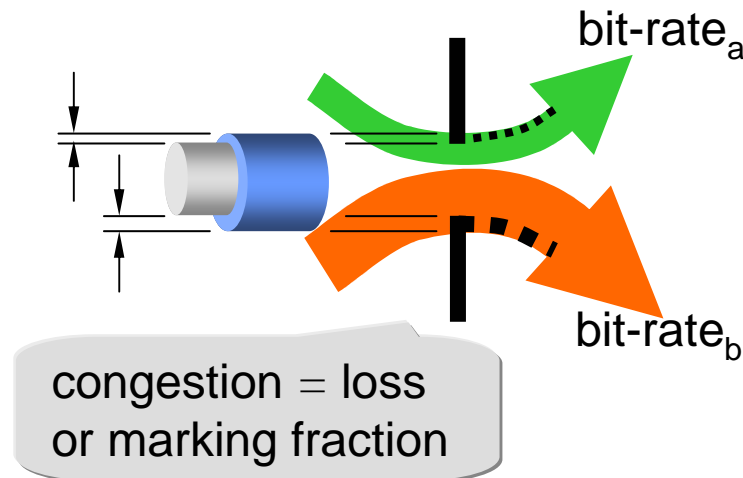
- volume weighted by congestion when it was sent

- takes into account all three factors

• bit-rate	✓	✓	✓	✓	✓
• weighted by congestion	✓	~	~	x	~
• activity over time	✓	x	x	✓	✓
congestion-volume	TCP	WFQ	Vol	DPI	

- a dual metric
 - of customers to ISPs (too much traffic)
 - and ISPs to customers (too little capacity)
- a) cost to other users of your traffic
- b) marginal cost of equipment upgrade
 - so it wouldn't have been congested
 - so traffic wouldn't have affected others
- competitive market matches a) & b)

- how to measure
 - volume that is marked with explicit congestion notification (ECN)
 - can't be gamed by strategising machines



if only...

ingress net could see congestion...
flat fee congestion policing

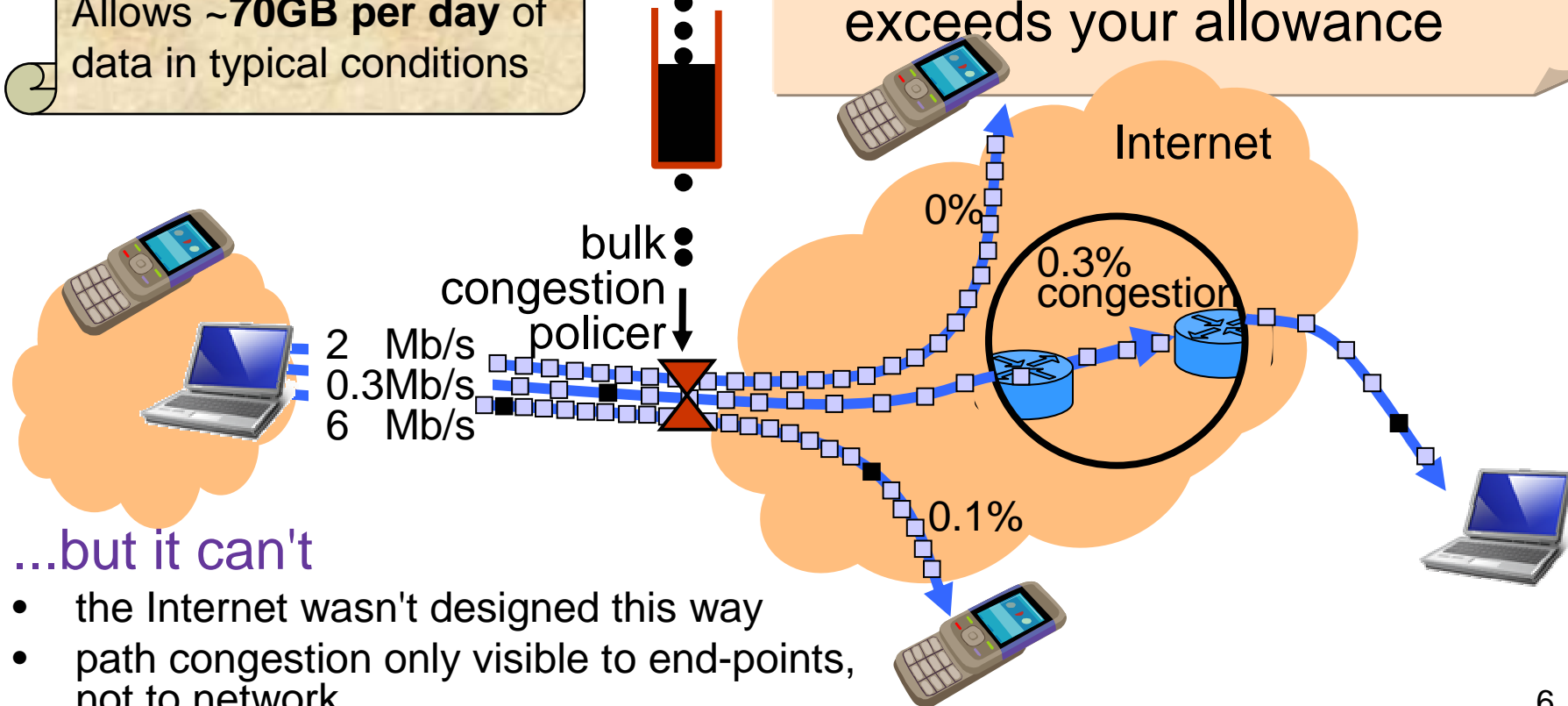
Acceptable Use Policy

'congestion-volume'
allowance: 1GB/month

@ €15/month

Allows ~70GB per day of
data in typical conditions

- incentive to avoid congestion
- only throttles traffic when your contribution to congestion in the cloud exceeds your allowance



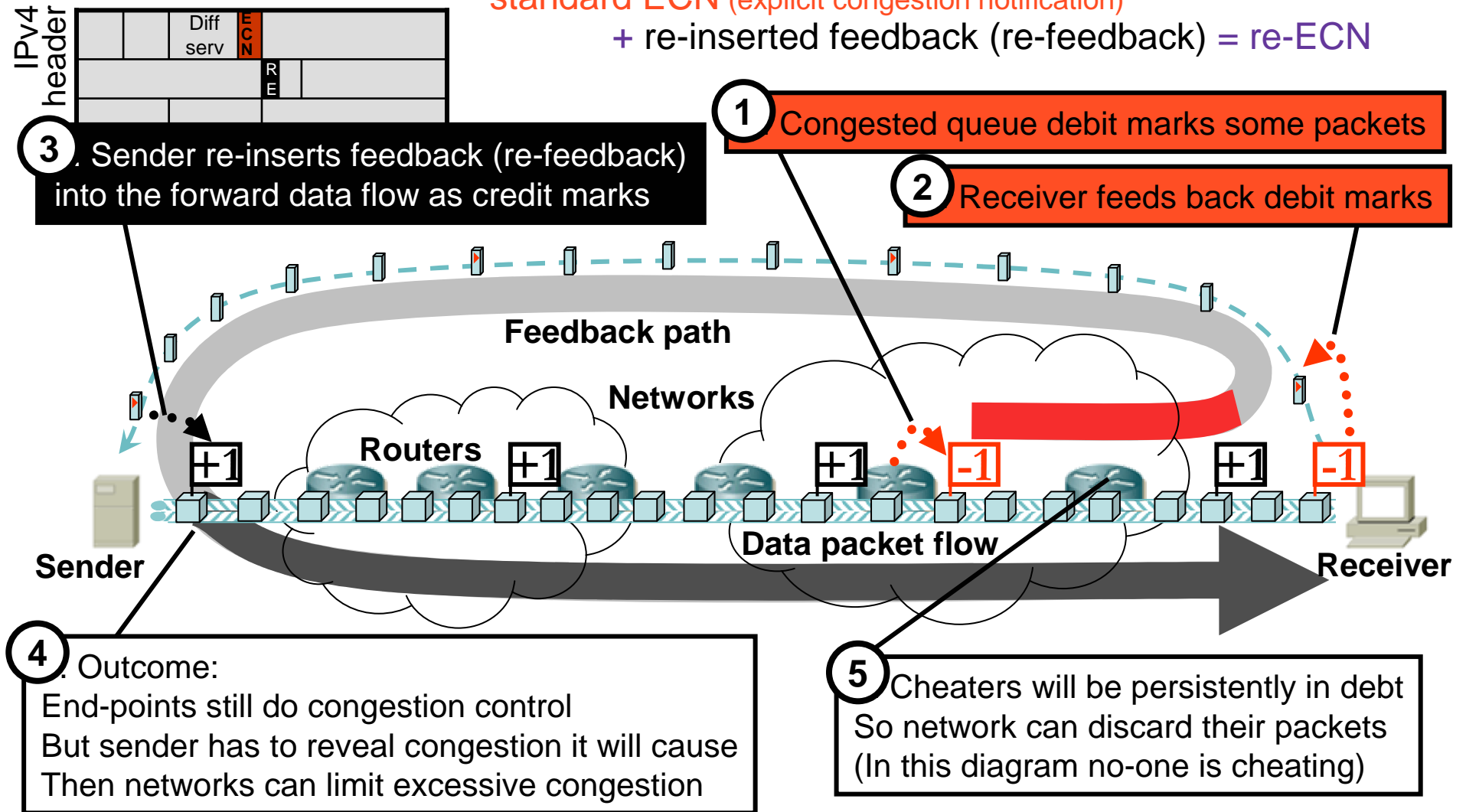
...but it can't

- the Internet wasn't designed this way
- path congestion only visible to end-points, not to network

congestion transparency in one bit

standard ECN (explicit congestion notification)

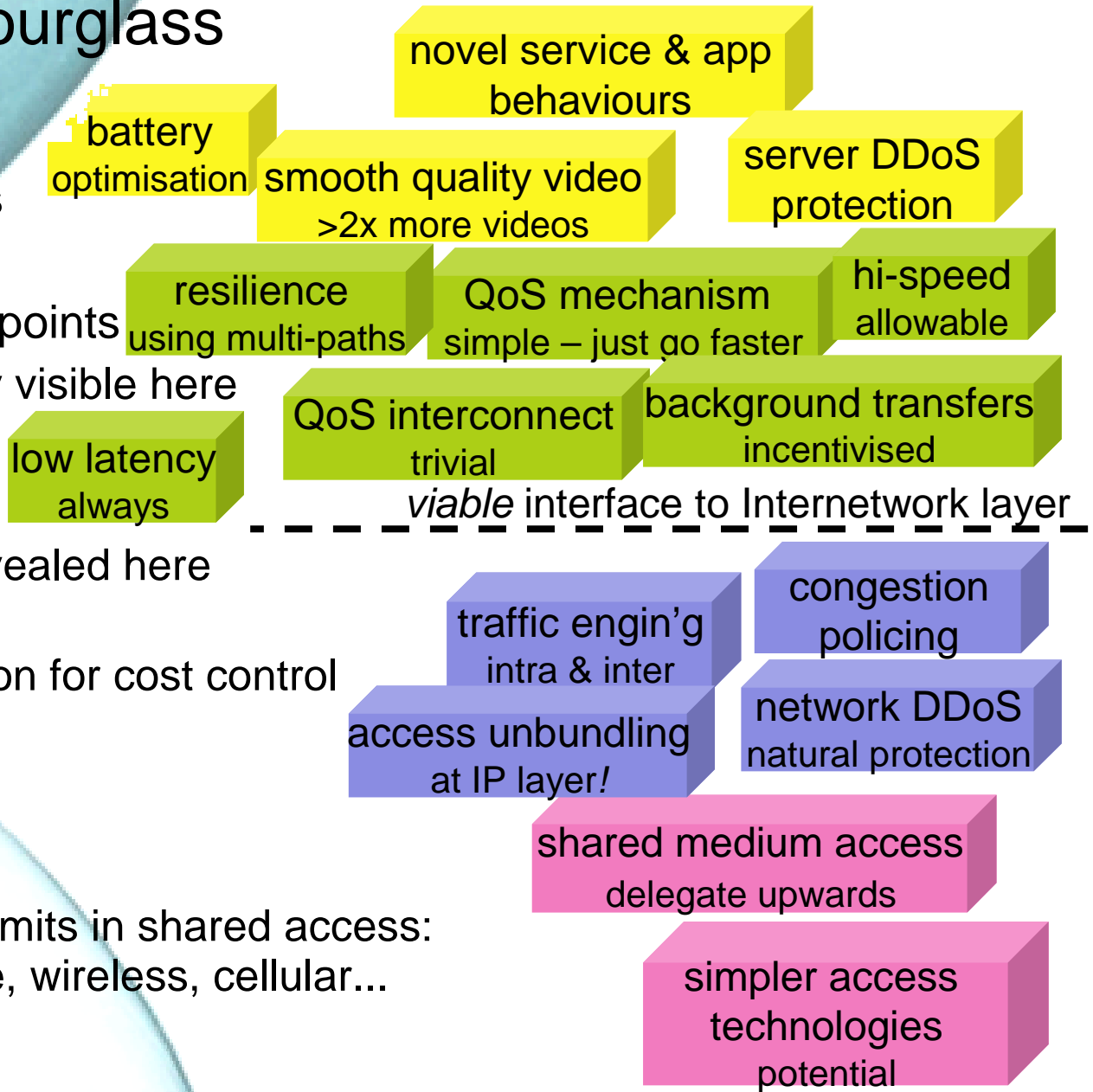
+ re-inserted feedback (re-feedback) = re-ECN



no changes required to IP data forwarding

the neck of the hourglass ...but for control

- applications & services
- transport layer on end-points
 - usage costs currently visible here
- internetwork layer
 - once usage costs revealed here
 - ISPs won't need deep packet inspection for cost control
- link layer
 - can remove bit-rate limits in shared access: passive optical, cable, wireless, cellular...



more info...

- The whole story in 7 pages
 - Bob Briscoe, "Internet Fairer is Faster", BT White Paper (Jun 2009) ...this formed the basis of:
 - Bob Briscoe, "[A Fairer, Faster Internet Protocol](#)", IEEE Spectrum (Dec 2008)
- Slaying myths about fair sharing of capacity
 - [Briscoe07] Bob Briscoe, "[Flow Rate Fairness: Dismantling a Religion](#)" ACM Computer Communications Review 37(2) 63-74 (Apr 2007)
- How wrong Internet capacity sharing is and why it's causing an arms race
 - Bob Briscoe et al, "[Problem Statement: Transport Protocols Don't Have To Do Fairness](#)", IETF Internet Draft (Jul 2008)
- re-ECN protocol spec
 - Bob Briscoe et al, "[Adding Accountability for Causing Congestion to TCP/IP](#)" IETF Internet Draft (Mar 2009)
- Re-architecting the Internet:
 - The [Trilogy](#) project <www.trilogy-project.org>

IRTF Internet Capacity Sharing Architecture design team

<<http://trac.tools.ietf.org/group/irtf/trac/wiki/CapacitySharingArch>>

re-ECN & re-feedback project page:

<<http://bobbriscoe.net/projects/refb/>>

BoF planning for following IETF: [subscribe](#), re-ecn@ietf.org

implementation (linux or ns2) bob.briscoe@bt.com

Internet capacity sharing for packets not flows

discuss...



main steps to deploy re-feedback / re-ECN

summary

rather than control sharing in the access links,
pass congestion info & control upwards

- network
 - turn on explicit congestion notification in data forwarding
 - already standardised in IP & MPLS
 - standards required for meshed network technologies at layer 2 (ECN in IP sufficient for point to point links)
 - deploy simple active policing functions at customer interfaces around participating networks
 - passive metering functions at inter-domain borders
- terminal devices
 - (minor) addition to TCP/IP stack of sending device
 - or sender proxy in network
- then new phase of Internet evolution can start
 - customer contracts & interconnect contracts
 - endpoint applications and transports
- requires update to the IP standard (v4 & v6)
 - started process in Autumn 2005
 - using last available bit in IPv4 header or IPv6 extension header

routing money

and simple internalisation of all externalities

