



design for tussle

Bob Briscoe
Chief Researcher, BT
Jun 2009



re-architecting the Internet



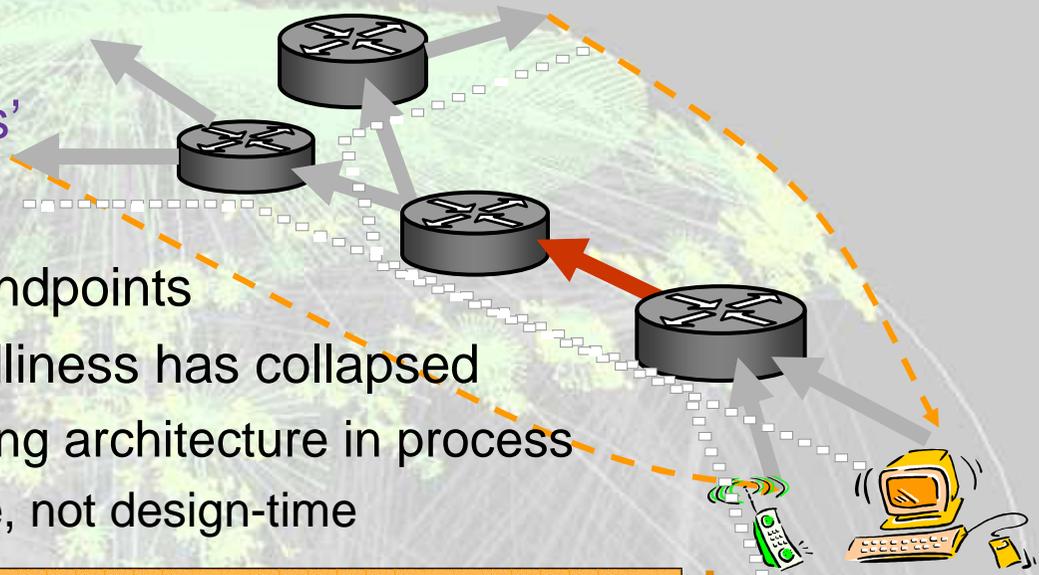
design for tussle

- enduring struggles over economic & social reward, power, business models, etc
- futile for architects to shape the outcome of these tussles
 - otherwise those in power violate the architecture to achieve their ends
 - result: unstructured heap
 - bizarre feature interactions, broken evolution potential
- role of designers: allow tussles to play out at run-time
 - technical excellence still necessary, but not enough
 - not to be confused with indecision over technical choices
- examples
 - extracting value vs. value neutral
 - self-supply vs. service provision
 - traceability vs. anonymity

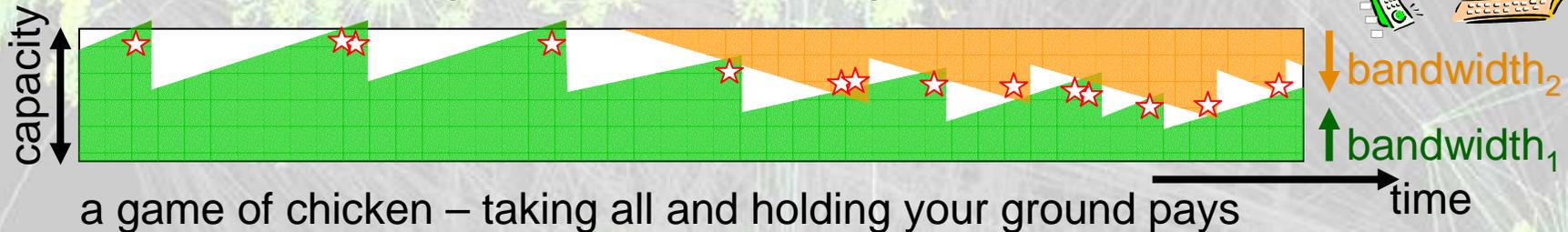


how Internet sharing 'works'

TCP-friendliness



- voluntarily polite algorithm in endpoints
- since 2006 belief in TCP-friendliness has collapsed
 - rewrite of IETF capacity sharing architecture in process
 - to control sharing at run-time, not design-time



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



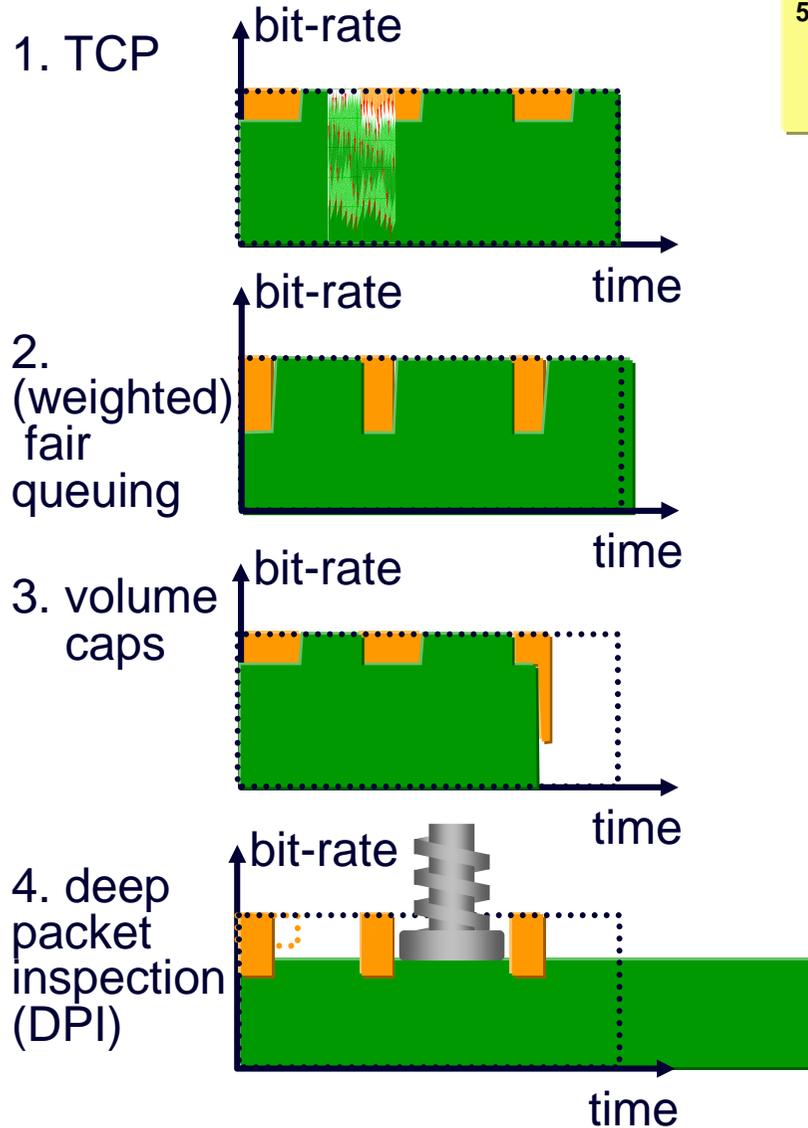
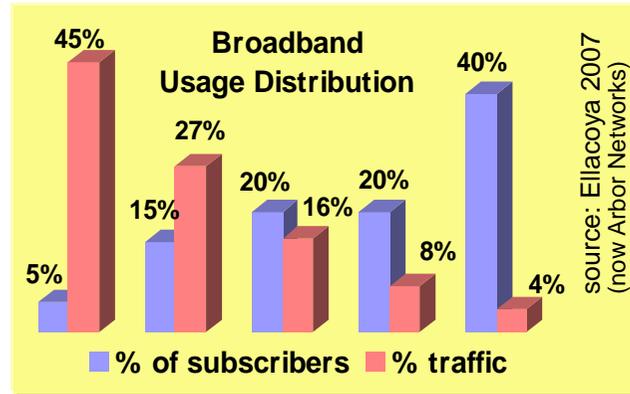
or start more 'TCP-friendly' flows than anyone else (Web: x2, p2p: x5-100)



or for much longer than anyone else (p2p file-sharing x200)

- net effect of both (p2p: x1,000-20,000 higher traffic intensity)

ISP's have quietly overridden TCP



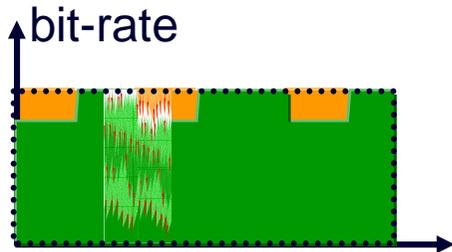


closing off the future

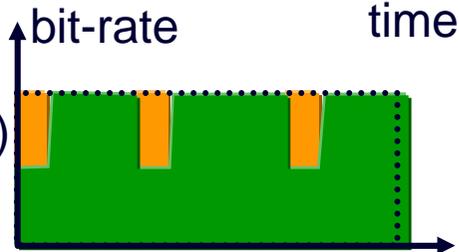
- without correct metric, ISPs resort to application analysis
 - getting impossible to deploy a new use of the Internet
 - must negotiate the arbitrary blocks and throttles en route
- two confusable motives
 - fairer cost sharing
 - competitive advantage to own services
- how to deconfuse: make cost of usage transparent
 - fixing Internet technology should avoid need for legislation

ISP's have quietly overridden TCP

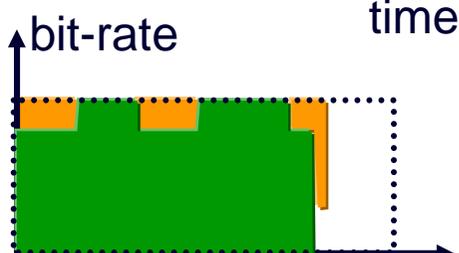
1. TCP



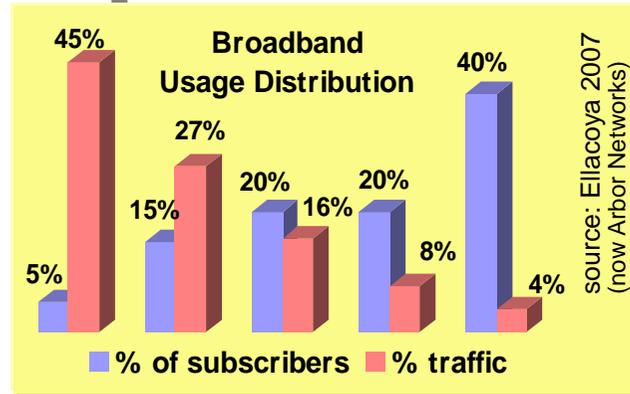
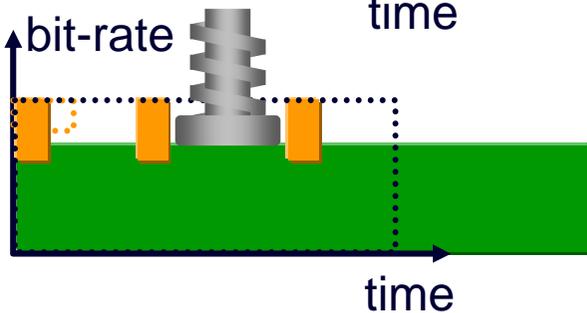
2. (weighted) fair queuing



3. volume caps

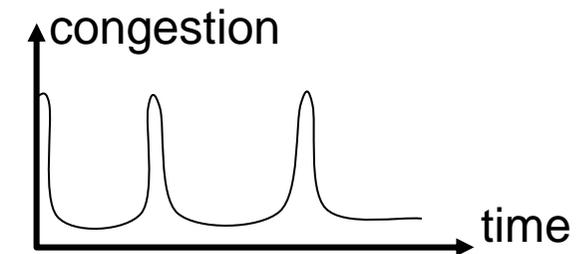
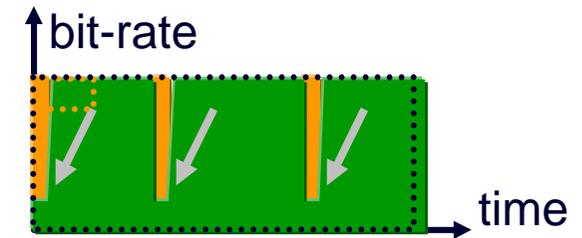


4. deep packet inspection (DPI)



simpler & better...

weighted TCP sharing



- 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

flat fee congestion policing



if ingress net could see congestion cost...

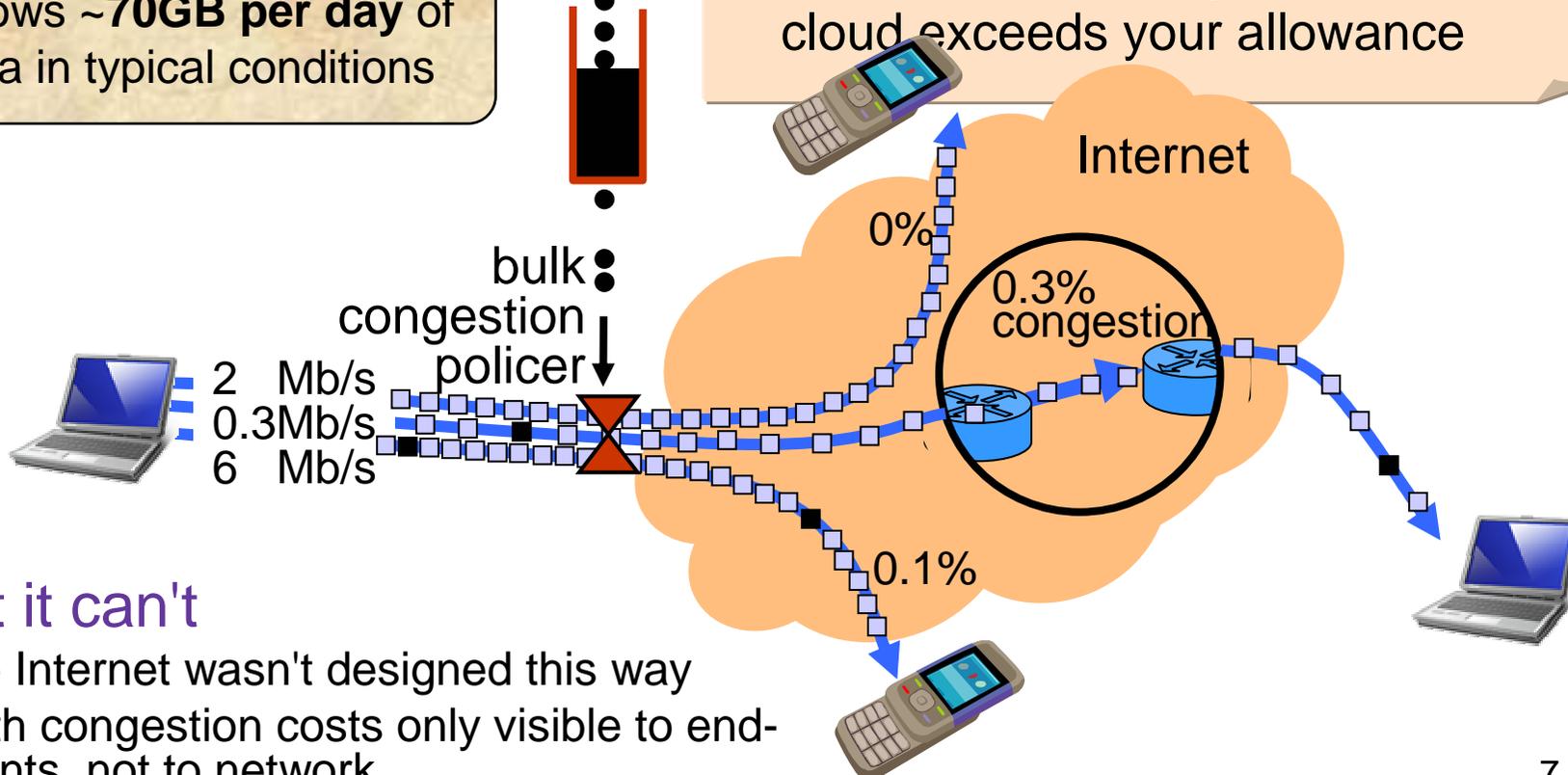
Acceptable Use Policy

'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

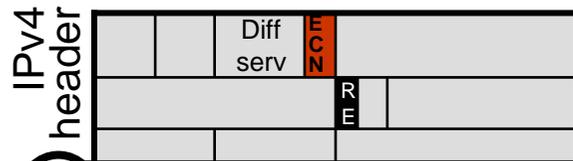
- the Internet wasn't designed this way
- path congestion costs only visible to end-points, not to network

cost transparency in one bit



standard ECN (explicit congestion notification)

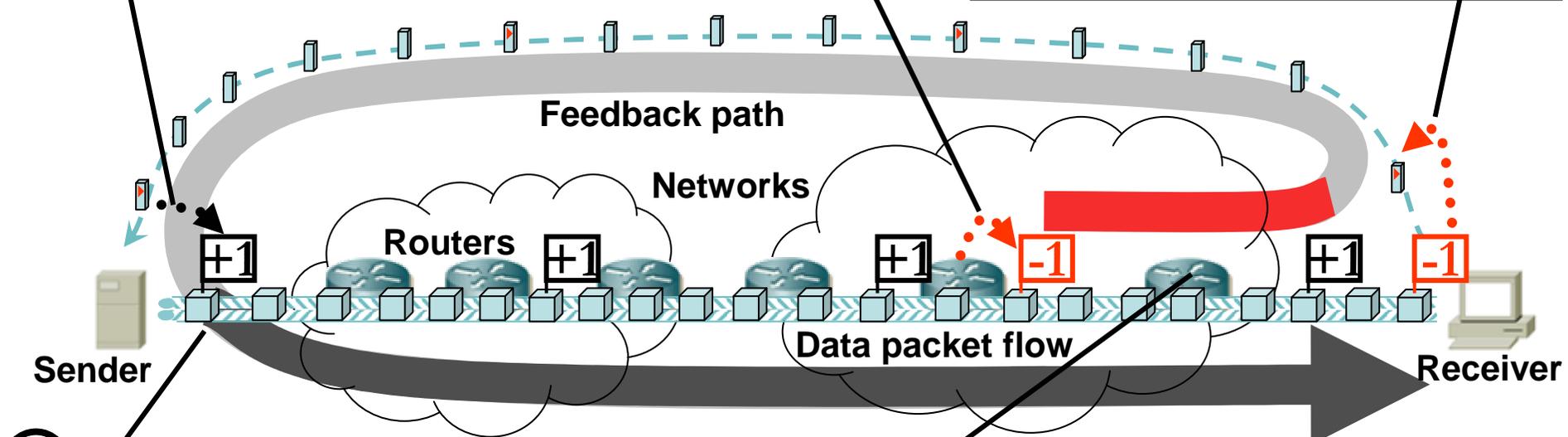
+ re-inserted feedback (re-feedback) = re-ECN



3. Sender re-inserts feedback (re-feedback) into the forward data flow as credit marks

1. Congested queue debit marks some packets

2. Receiver feeds back debit marks



4. Outcome:
End-points still do congestion control
But sender has to reveal congestion it will cause
Then networks can limit excessive congestion

5. Cheaters will be persistently in debt
So network can discard their packets
(In this diagram no-one is cheating)

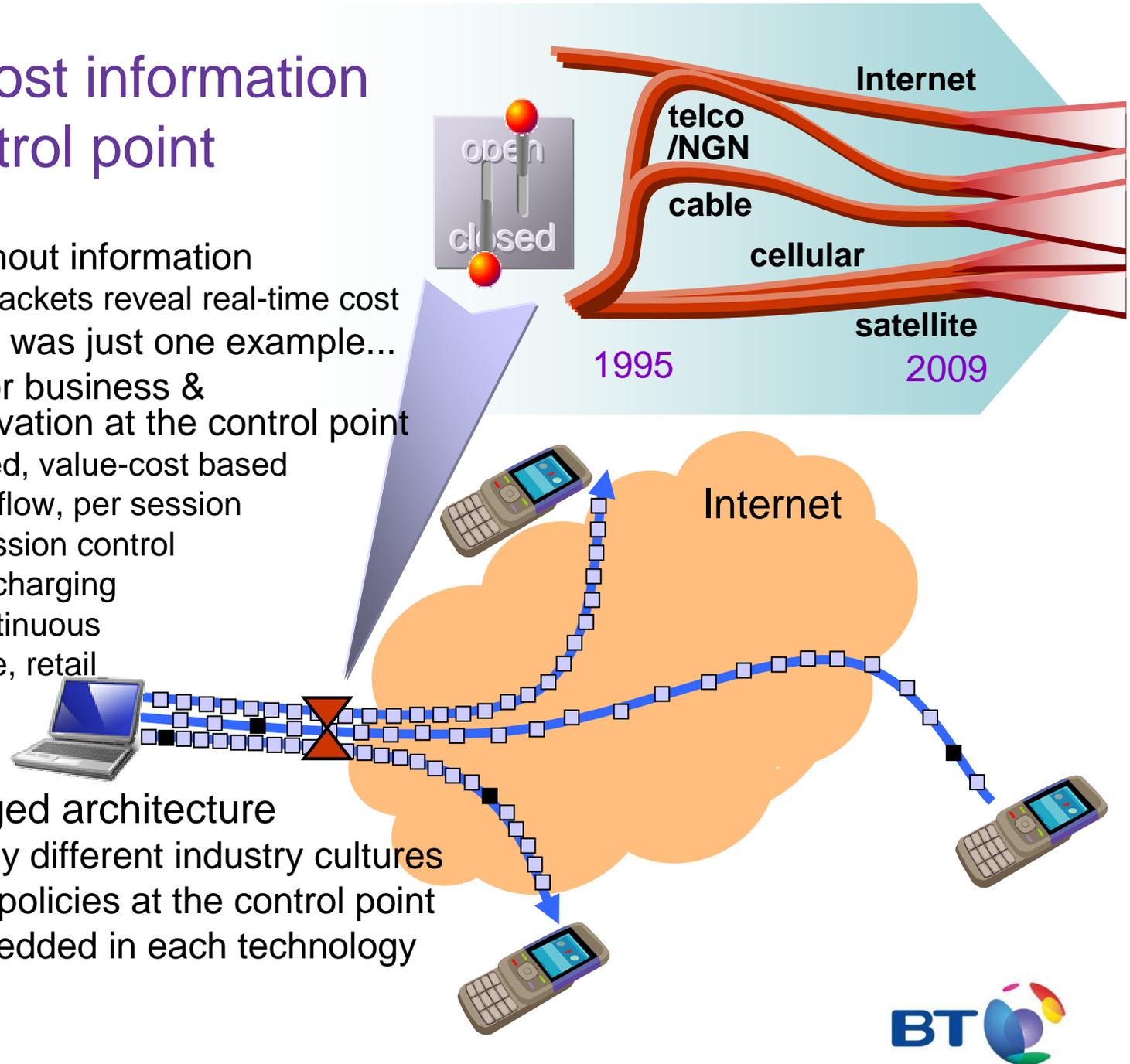
no changes required to IP data forwarding

bringing cost information to the control point

- no control without information
 - re-ECN packets reveal real-time cost
- flat fee policer was just one example...
- huge space for business & technical innovation at the control point

- cost based, value-cost based
- bulk, per flow, per session
- call admission control
- policing, charging
- tiers, continuous
- wholesale, retail

- truly converged architecture
 - can apply different industry cultures
 - through policies at the control point
 - not embedded in each technology



a new chapter of innovation

novel service & app behaviours



- applications & services

battery optimisation

smooth quality video
>2x more videos

server DDoS protection

- transport layer on end-points

resilience
using multi-paths

QoS mechanism
simple – just go faster

hi-speed transfers

- usage costs currently visible here

low latency
always

QoS interconnect
trivial

background transfers
incentivised

- internetwork layer

commercially viable interface to Internet layer

- once usage costs revealed here

- ISPs won't need deep packet inspection for cost control

traffic engin'g
intra & inter

congestion
policing

access unbundling
at IP layer!

network DDoS
natural protection

- link layer

- can remove bit-rate limits in shared access:
passive optical, cable, wireless, cellular...

shared medium access
delegate upwards

simpler access
technologies
potential

- all due to 'design for tussle'

trilogy

re-architecting the Internet

the neck of the hourglass, for control

www.trilogy-project.eu

This work is partly funded by Trilogy, a research project (ICT-216372) supported by the European Community under its Seventh Framework Programme. The views expressed here are those of the author(s) only. The European Commission is not liable for any use that may be made of the information in this document.





more info...

- Design for Tussle
 - David Clark, John Wroclawski, Karen Sollins and Robert Braden, "Tussle in Cyberspace: Defining Tomorrow's Internet," in IEEE/ACM Transactions on Networking **13**(3) 462-475 (2005)
 - Alan Ford, Philip Eardley, Barbara van Schewick, "New Design Principles for the Internet," in Proc *IEEE ICC Future networks* (2009)
- The whole capacity sharing story in 5 pages
 - Bob Briscoe, "[A Fairer, Faster Internet Protocol](#)", IEEE Spectrum (Dec 2008)
- Slaying myths about fair sharing of capacity
 - 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-architecting the Internet:

The [Trilogy](http://www.trilogy-project.org) project <www.trilogy-project.org>

congestion transparency, re-ECN & re-feedback project page:

<http://www.cs.ucl.ac.uk/staff/B.Briscoe/projects/refb/>

bob.briscoe@bt.com

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

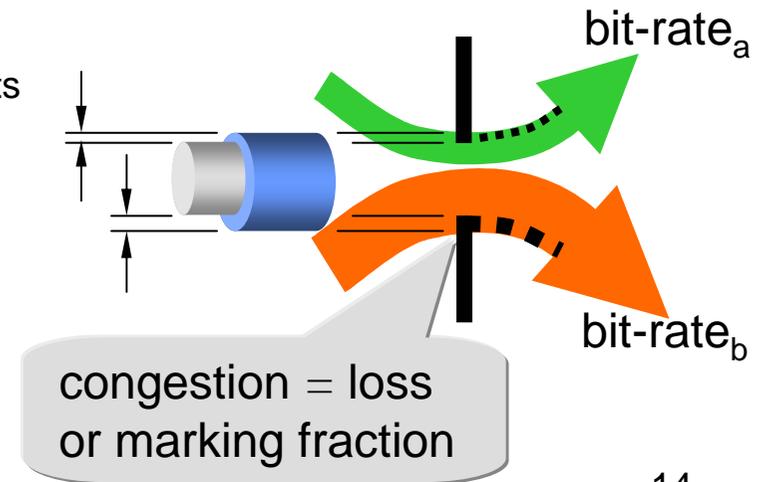
a new resource accountability metric – a bandwidth trading unit



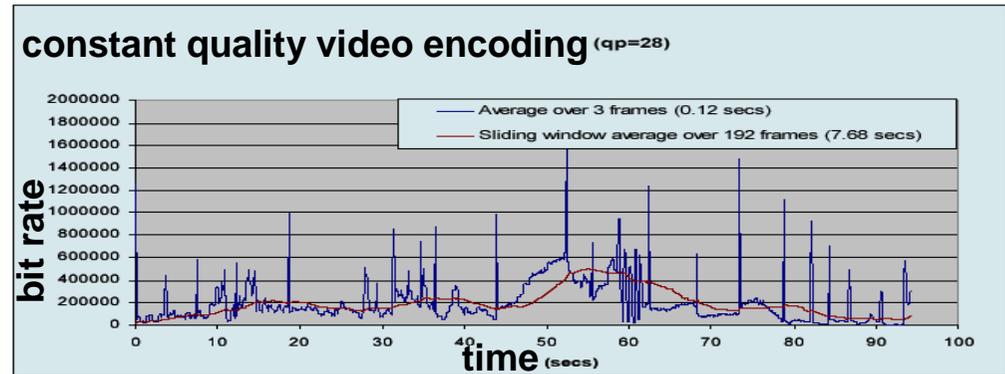
- how to measure
 - volume that is marked with explicit congestion notification (ECN)
 - can't be gamed by strategising machines
- a resource accountability metric
 - of customers to ISPs (too much traffic)
 - and ISPs to customers (too little capacity)
- a) cost to other users of your traffic
- b) marginal cost of equipment upgrade
 - so it wouldn't have been congested
 - so traffic wouldn't have affected others
- competitive market matches a) & b)

- cost of network usage
 - unforgivable for a business not to understand its costs
- answer: **congestion-volume**
 - volume weighted by congestion when it was sent
- takes into account all three factors

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



guaranteed bit-rate?
 or much faster 99.9% of the time?
 harnessing flexibility



- the idea that humans want to buy a known fixed bit-rate
 - comes from the needs of media delivery technology
 - hardly ever a human need or desire
- services want freedom & flexibility
 - access to a large shared pool, not a pipe
- when freedoms collide, congestion results
 - many services can adapt to congestion
 - shift around resource pool in time/space



% figures =
 no. of videos
 that fit into the
 same capacity

Constant Bit Rate **100%** Constant Quality **125%** Equitable Quality **216%**
 sequences encoded at same average of 500kb/s [Crabtree09]

routing money

and simple internalisation of all externalities

