fixing the Internet for sustainable business models

Bob Briscoe
Chief Researcher, BT Group

Dec 2008

BT future communications architecture programme

- instigated 2002
- to lead global moves to fix the Internet architecture
  - top-down (pressure for national funding, set research agenda etc)
  - bottom-up as peer researchers
- IP: the foundation of BT's 21C architecture
  - rather than BT-specific comms architecture fixes
  - make the off-the shelf architecture fit for the whole value chain
- scope: ICT infrastructure
  - multi-provider, high volume, low margin, generic with hooks

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.

BT UCL roke NEC Deutsche Telekom NOKIA UCL Université catholique de Louvain EURESCOM

how to share the resources of a cloud

known problem since early Internet

- tremendous idea
  - anyone can use any link anywhere on the Internet without asking, as much as they like
- when freedoms collide
  - what share does each party get?

- keeping one-way datagrams
- allowing for
  - self-interest & malice
    - of users and of providers
  - evolvability
    - of new rate dynamics from apps
    - of new business models
  - viability of supply chain
  - simplicity

- if we do nothing
  - the few are ruining it for the many
  - massive capacity needed to keep interactive apps viable
  - poor incentives to invest in capacity
  - operators are kludging it with deep packet inspection
  - solely today's apps frozen into net
  - complex, ugly feature interactions

Internet topology visualization produced by Walrus (Courtesy of Young Hyun, CAIDA)

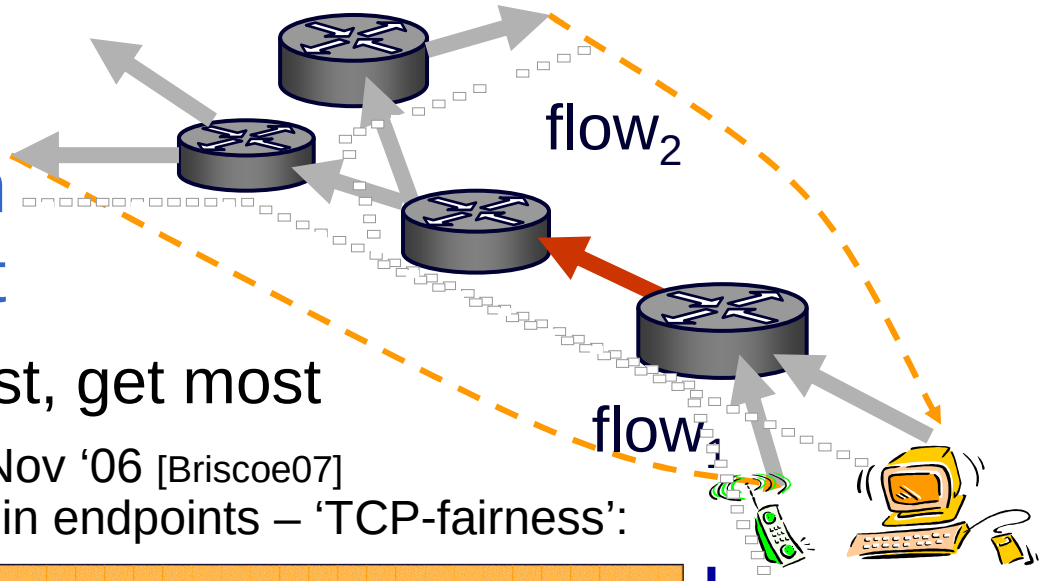
moving mountains

Internet Engineering Task Force

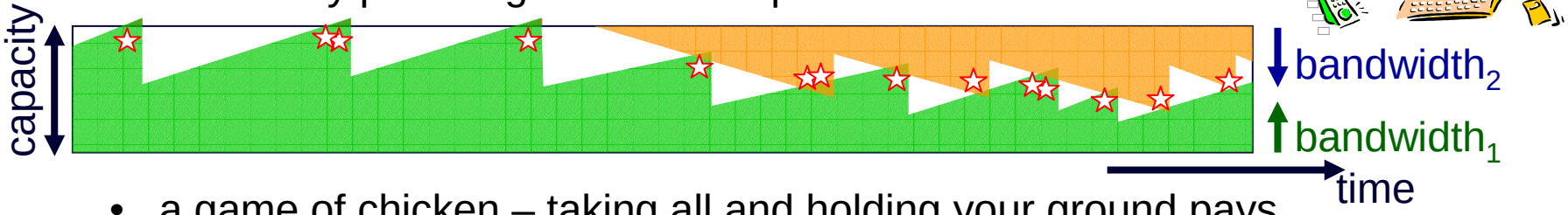
- Nov 2005
  - proposed replacement resource sharing architecture to IETF
  - general response: "What's the problem? TCP prevalent, so sharing OK"
- Nov 2006
  - Dismantled TCP-Friendliness religion at IETF transport plenary
- Nov 2008
  - agreed to draft a major change to the Internet architecture
    - initially in IRTF Internet Congestion Control Research Group
    - eventual intent: Internet Architecture Board RFC
- main points likely to feature in the new architecture
  - primary resource sharing function in network, not end-points
  - congestion control still primarily in end-points

how Internet sharing 'works'

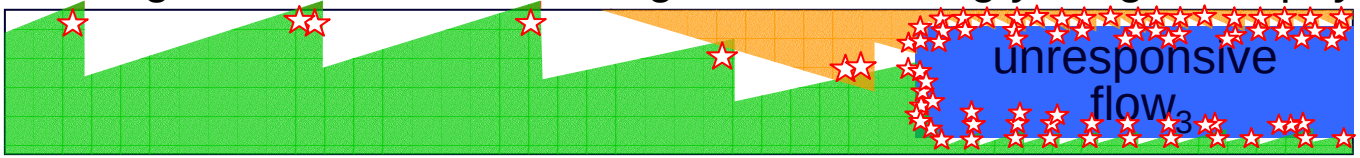
endemic congestion & voluntary restraint



- aka. those who take most, get most
  - technical consensus until Nov '06 [Briscoe07] voluntarily polite algorithm in endpoints – 'TCP-fairness':

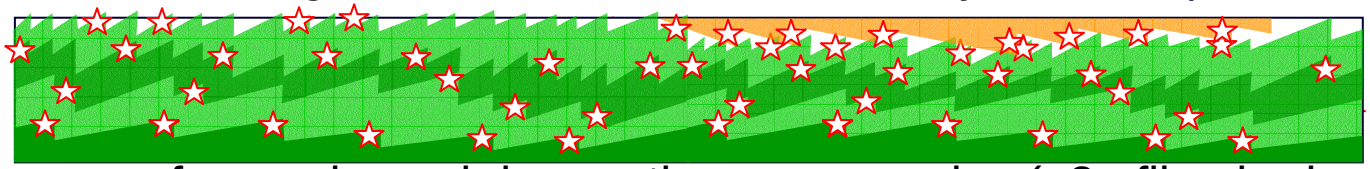


  - a game of chicken – taking all and holding your ground pays



(VoIP, VoD Joost 700kbps)

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



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

BT

ITU working definition of NGN

A Next Generation Network (NGN) is a packet-based network able to provide services including Telecommunication Services and able to make use of multiple broadband, QoS-enabled transport technologies and in which service-related functions are independent from underlying transport-related technologies. It offers unrestricted access by users to different service providers. It supports generalized mobility which will allow consistent and ubiquitous provision of services to users.

The NGN is characterized by the following fundamental aspects:

- ...
- Decoupling of service provision from network, and provision of open interfaces
- ...

<www.itu.int/ITU-T/studygroups/com13/ngn2004/working_definition.html>

BT

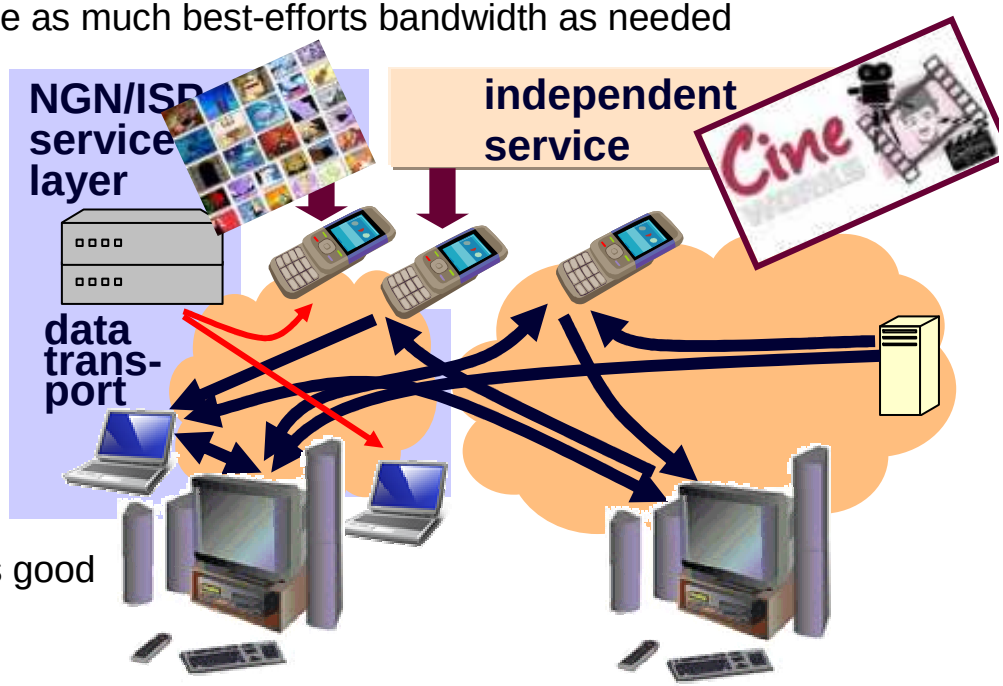
just saying it, doesn't make it true

- service-network independence: nice ideal
- but the economics makes it idealistic

- recovering network costs through services: nice ideal
- but IP technology makes it idealistic

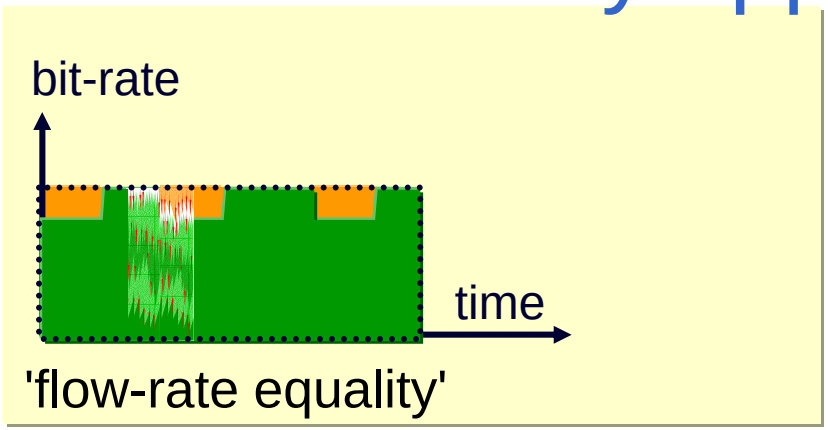
cost-shifting between services

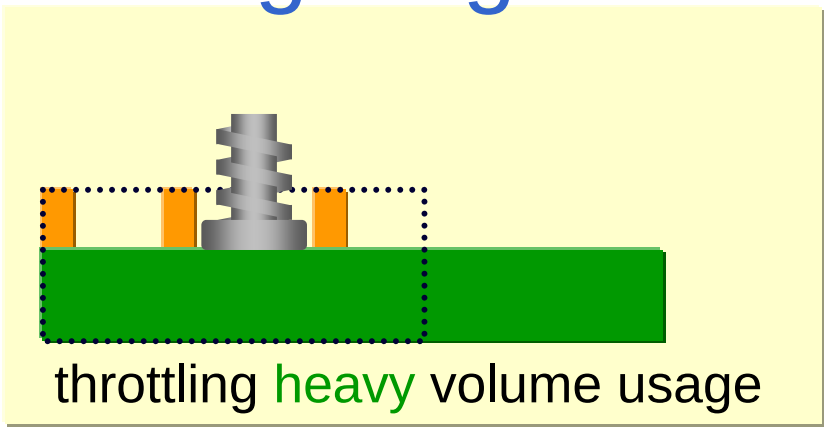
- scenario
  - ISP/NGN also a higher level service provider (TV, video phone, etc)
  - competing with independent service providers (Skype, YouTube, etc)
- capacity & QoS costs for high value services
  - ISP buys capacity & QoS internally
  - independent service & their customers use as much best-efforts bandwidth as needed
  - because of how Internet sharing 'works'
- cost of heavy usage service subsidised by ISP's lighter users
- knee-jerk reaction of ISP/NGN
  - block p2p or independent services
- No! don't blame your customers
- fix the cost accountability foundations
  - separation between network & services is good
  - but need to add cost accountability to IP



BT

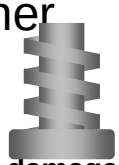
two arbitrary approaches fighting





degree of freedom	‘flow rate equality’ the Internet way (TCP)	‘volume accounting’ operators (& users)
multiple flows	✗	✓
activity factor	✗	✓
congestion variation	✓	✗
application control	✓	✗

- each cancels out the worst failings of the other
- Internet looks like 'it works OK'
- but the resulting arms race leaves collateral damage



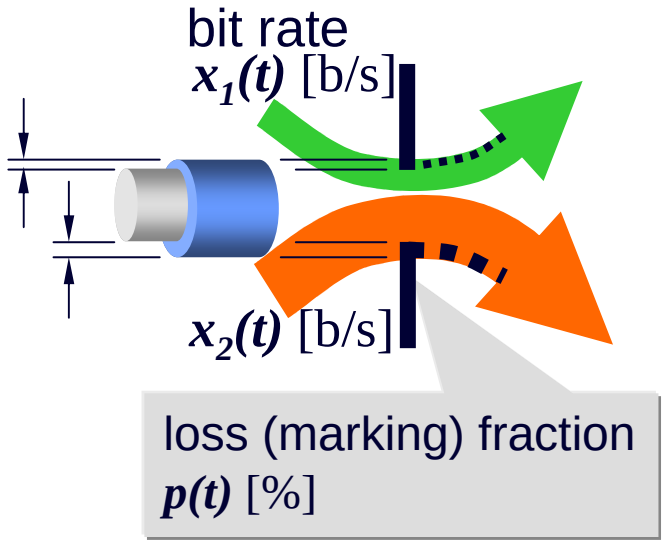
underlying problems
blame our choices, not p2p

- commercial
  - Q. what is cost of network usage?
  - A. volume? NO; rate? NO
  - A. 'congestion volume' (later slide)
  - *our own* unforgivable sloppiness over what *our* network costs are
- technical
  - lack of cost accountability in the Internet protocol (IP)
  - p2p file-sharers finding loopholes in technology *we* chose
- *we* haven't designed *our* contracts & technology for machine-powered customers

BT

core of solution

congestion-volume metric



- congestion-volume
  - your volume weighted by link congestion when each packet is served
- intuition
  - some ISPs count volume only during peak
  - like counting (100% x volume) during peak and (0% x volume) otherwise
  - congestion-volume counts $p \cdot x_i$ over time
- measurement
  - the amount of data discarded from your traffic
  - or marked with explicit congestion notification (ECN)
  - end-point function in current architecture

1. cost to other users of your traffic
2. the marginal cost of upgrading equipment
   - so it wouldn't have been congested
   - so traffic wouldn't have affected others

- competitive market matches 1 & 2

metric for customers to judge ISPs, and ISPs to judge customers

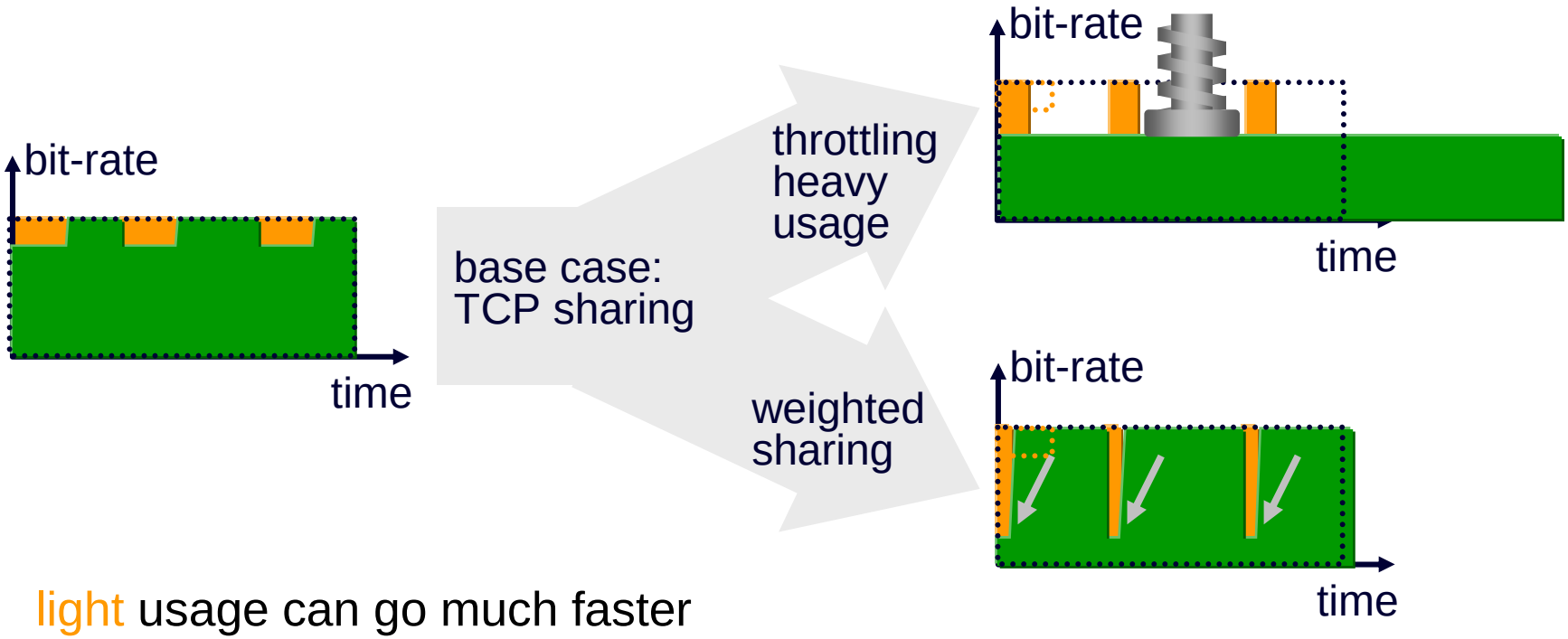
congestion = too much traffic meets too little capacity

most interesting when 'congestion' = marking, not loss

note: diagram is conceptual
congestion volume & capital cost of equipment would be accumulated over time

BT

there are better solutions than fighting



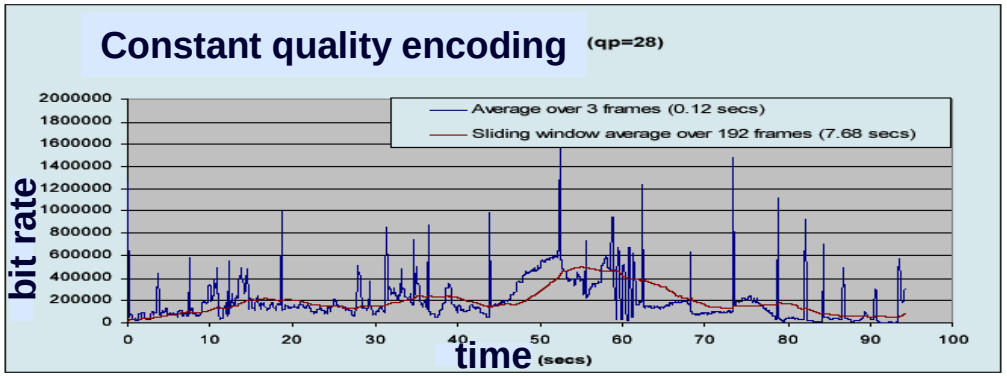
- light usage can go much faster
- hardly affecting completion times of heavy usage

NOTE: weighted sharing doesn't imply differentiated services

- can be weighted aggressiveness of end-point rate control

BT

there are better solutions than buying bit-rate



- 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
  - when freedoms collide, congestion results
  - many services can adapt to congestion
  - shift around the resource pool in time/space

Constant Bit Rate 100% Constant Oualitv 125% Equitable Quality 200%

sequences encoded at same average of 500kb/s

[Crabtree09]

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

BT

if ingress could see congestion

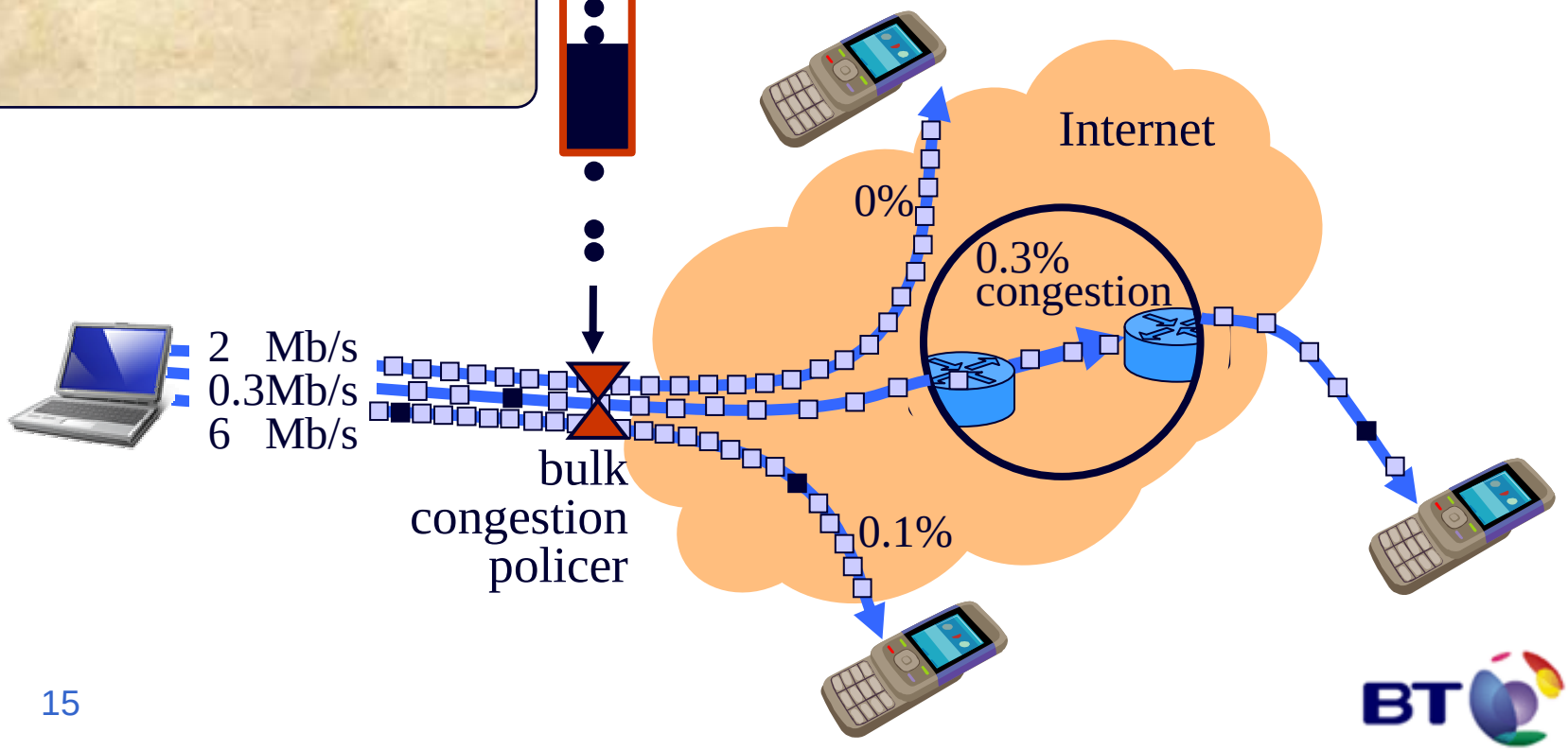
congestion policing

Acceptable Use Policy

Your 'congestion volume' allowance: 1GB/month (= 3kb/s continuous)
Only limits excess traffic above the Internet 'high-water-mark'

Under typical conditions this will allow you to transfer about **70GB per day**.

- only throttles traffic when your contribution to congestion in the cloud exceeds your allowance
- creates incentives for weighted sharing, equitable quality video, etc



problems using congestion in contracts

	1. loss	2. ECN	3. re-ECN
can't justify selling an impairment	☹	☺	☺
absence of packets is not a contractible metric	☹	☺	☺
congestion not visible to upstream network nodes	☹	☹	☺
congestion is outside a customer's control	☹	☹	☺
customers don't like variable charges	☹	☹	☺
congestion is not an intuitive contractual metric	☹	☹	☹

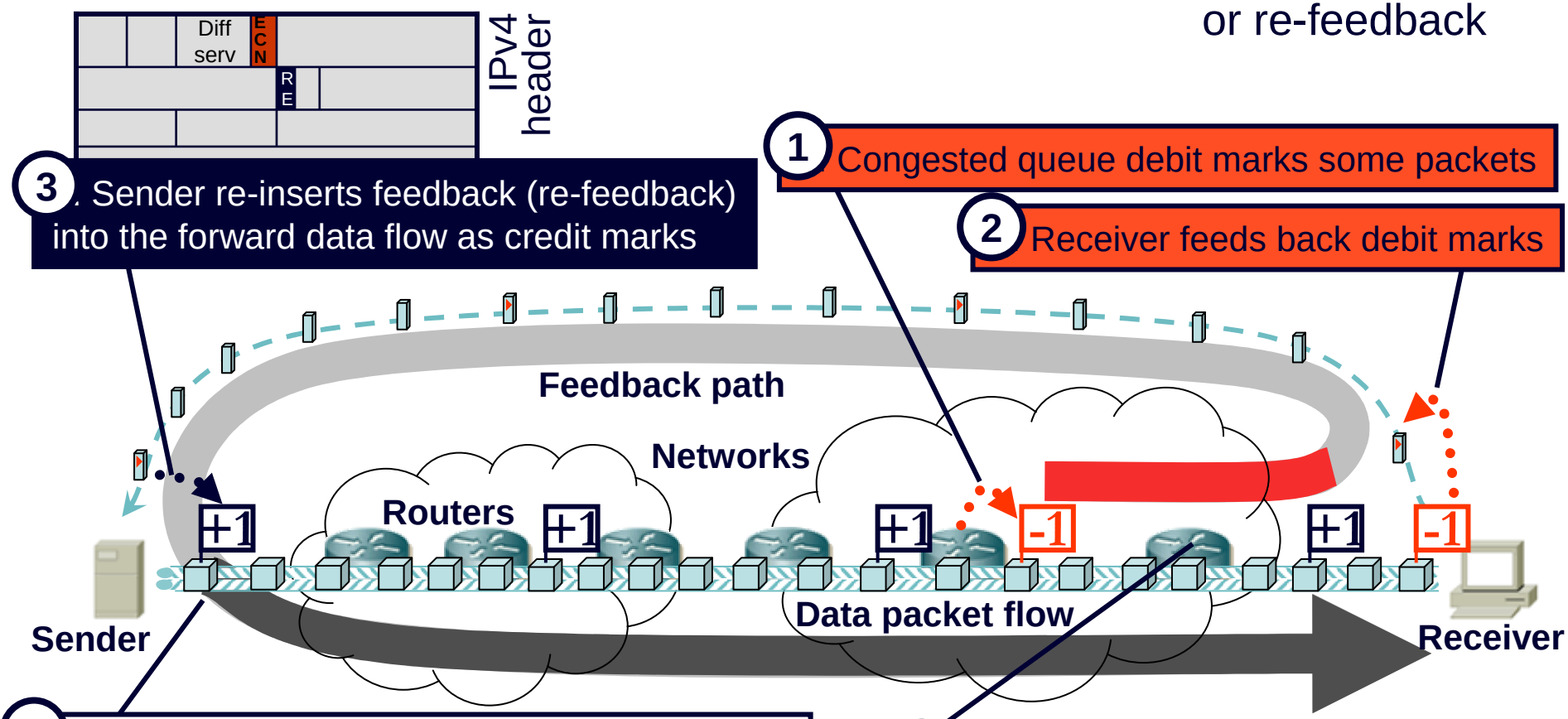
1. **loss:** used to signal congestion since the Internet's inception
   - computers detect congestion by detecting gaps in the sequence of packets
   - computers can hide these gaps from the network with encryption
2. **explicit congestion notification (ECN):** standardised into TCP/IP in 2001
   - approaching congestion, a link marks an increasing fraction of packets
   - implemented in Windows Vista (but off by default) and Linux, and IP routers (often off by default)

3. **re-inserted ECN (re-ECN):** standards proposal since 2005 (later slides)
   - packet delivery conditional on sender declaring expected congestion
   - uses ECN equipment in the network unchanged

BT

re-ECN = standard ECN + re-inserted feedback

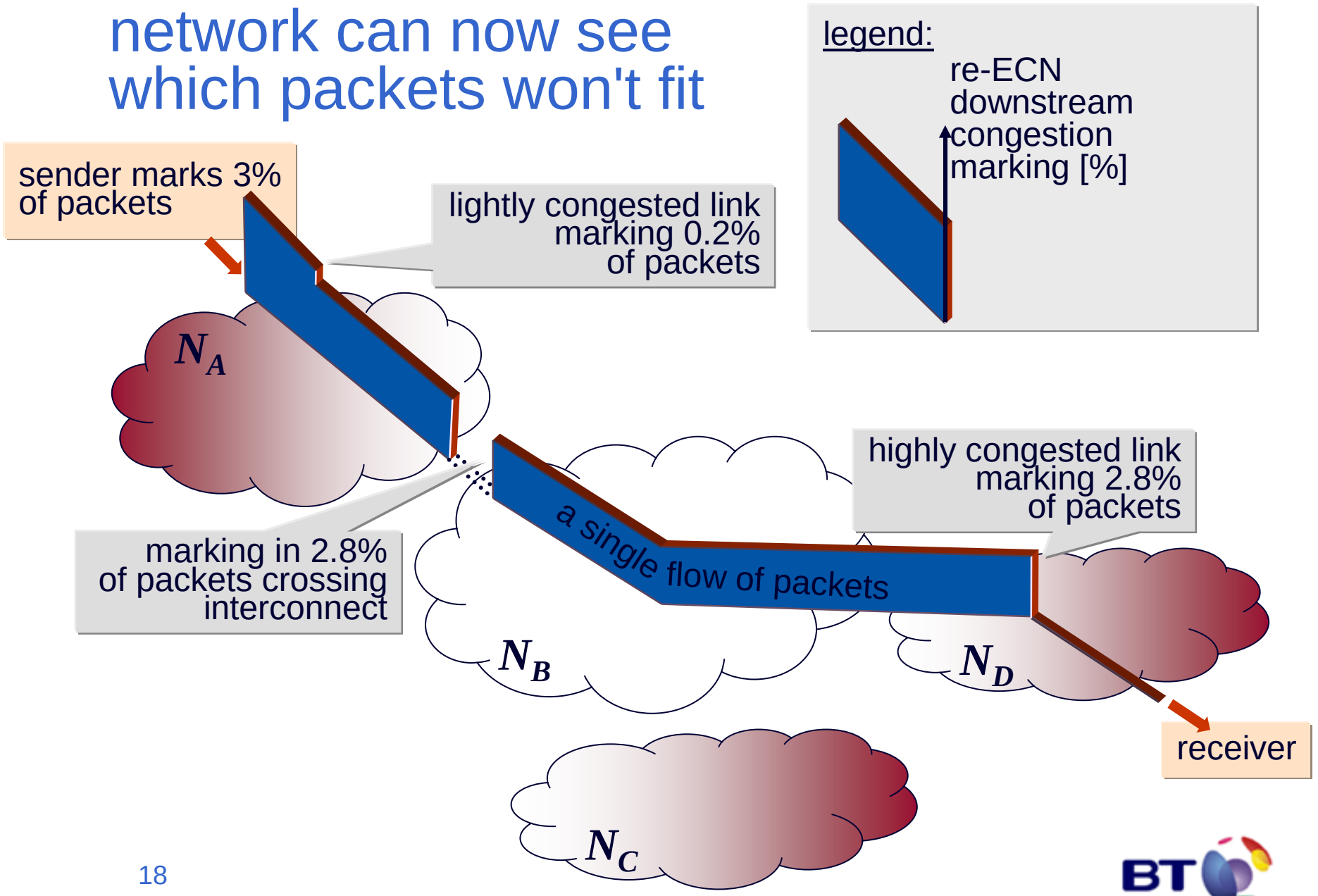
or re-feedback



- No changes required to data forwarding
- Realisation of network control & economics research stretching back to 1991 [Kelly05]

BT

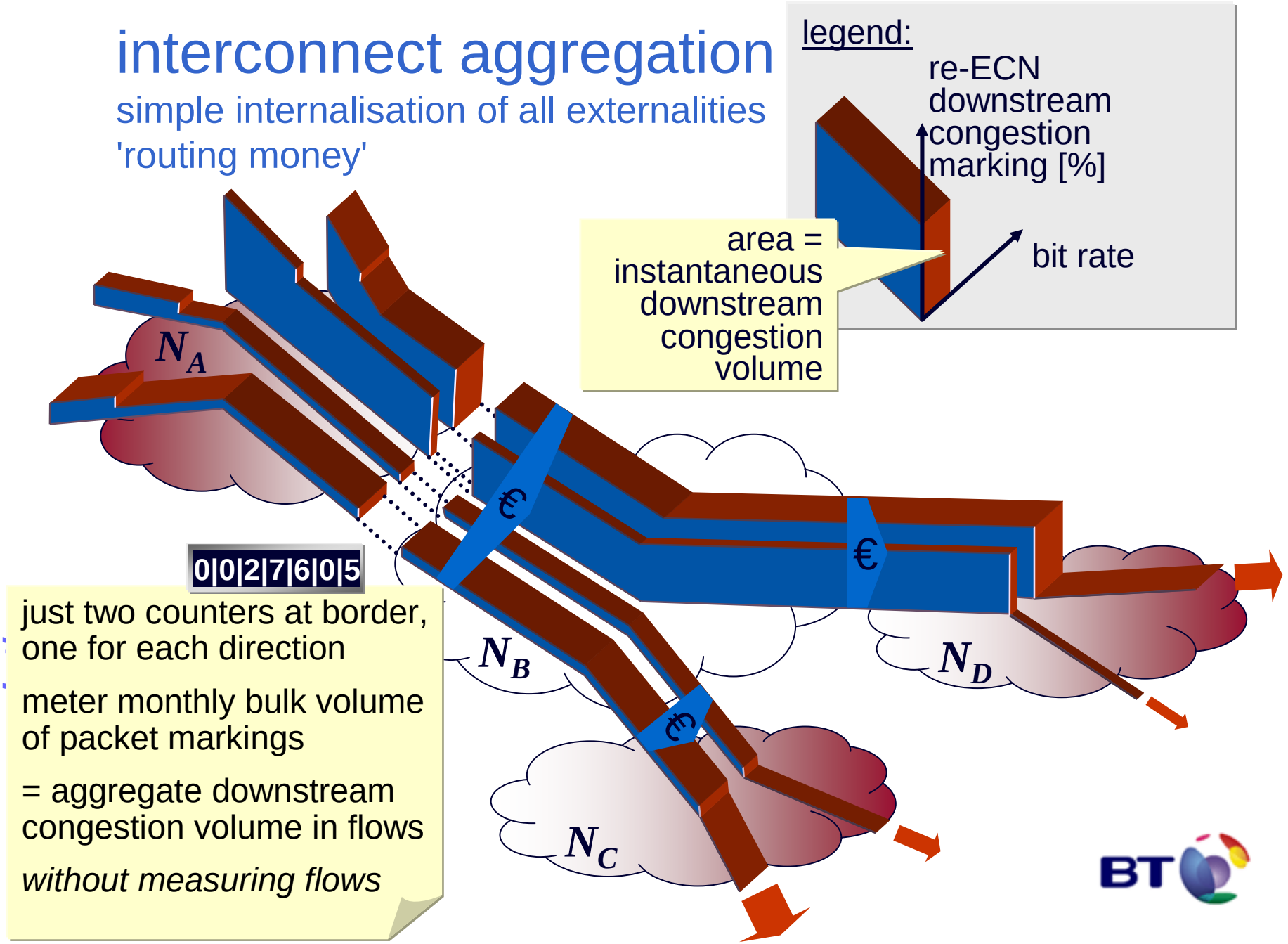
network can now see which packets won't fit



interconnect aggregation

simple internalisation of all externalities

'routing money'



just two counters at border, one for each direction

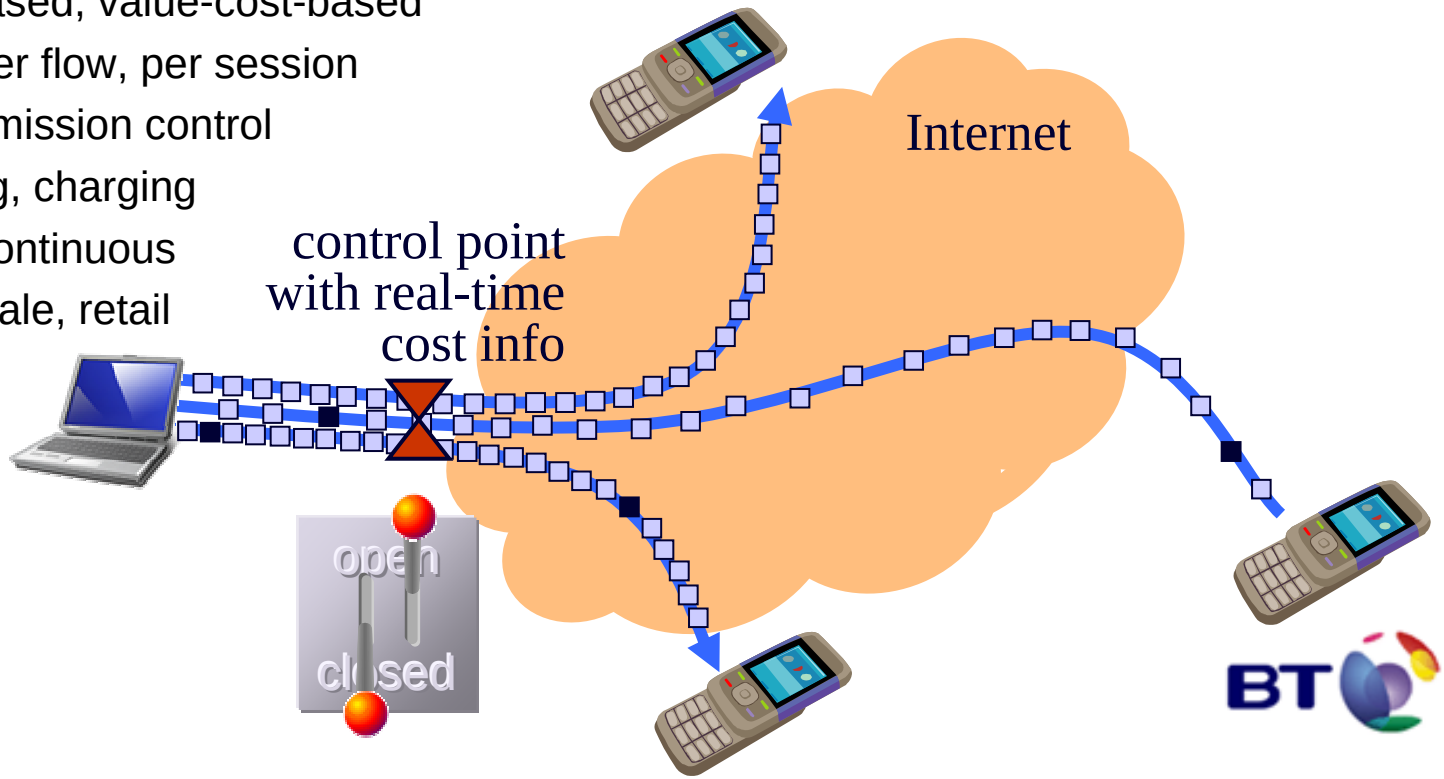
meter monthly bulk volume of packet markings

= aggregate downstream congestion volume in flows

without measuring flows

richer ingress control point

- no control without information
  - re-ECN packets carry info on their real-time cost implications
- control point is <u>designed for tussle</u>
  - bulk policer design given earlier was merely the most open possible example...
- huge space for business & technical innovation at policer
  - cost-based, value-cost-based
  - bulk, per flow, per session
  - call admission control
  - policing, charging
  - tiers, continuous
  - wholesale, retail



a new chapter of innovation

- hugely opens space for apps / services
- costs currently only visible at transport layer

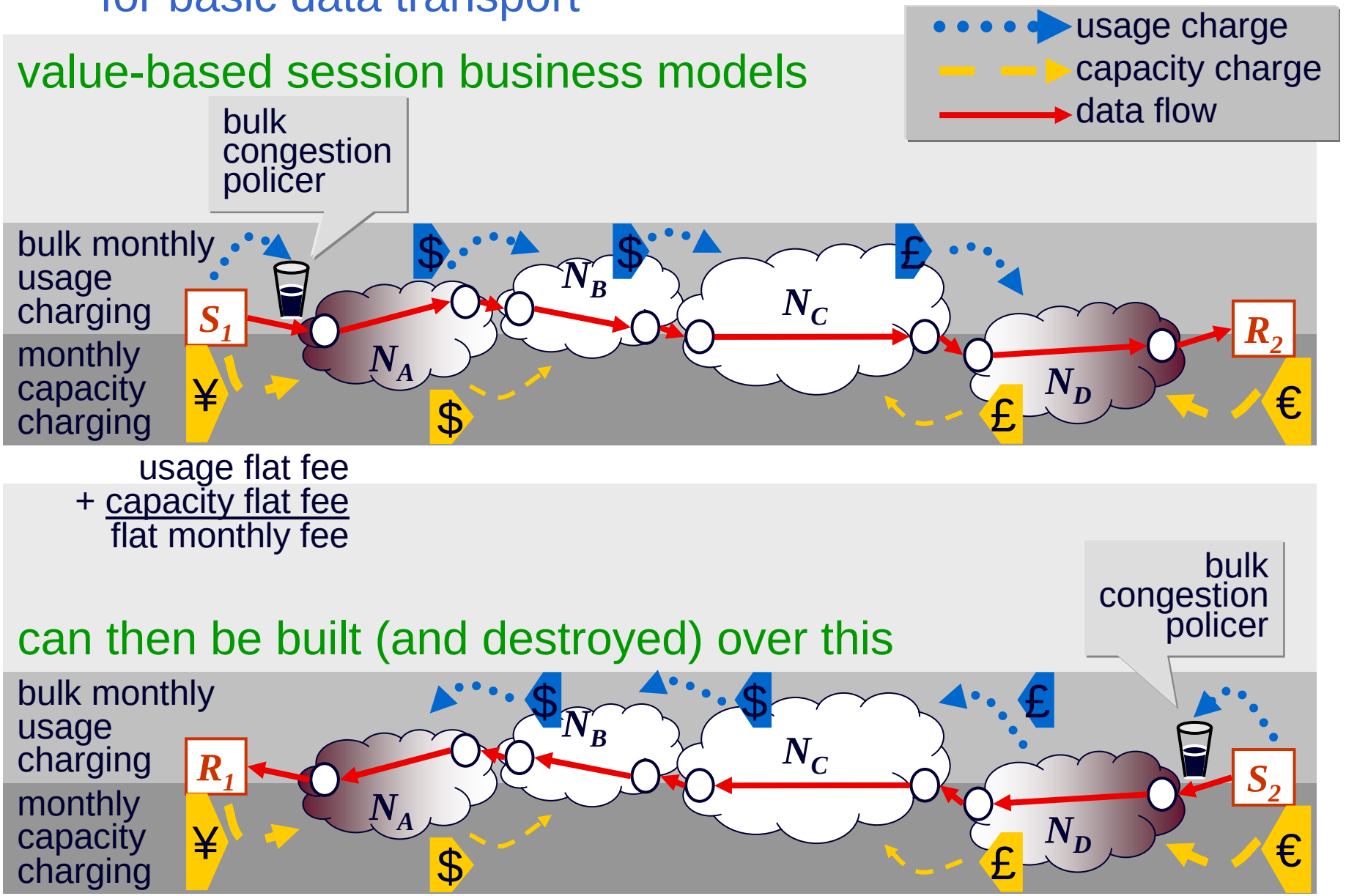
- once costs revealed at network layer
  - ISPs won't need deep packet inspection for cost control

- can remove restrictions in shared access networks
  - passive optical networks, cable, wireless, cellular
  - won't need bit-rate limits once network layer can limit congestion

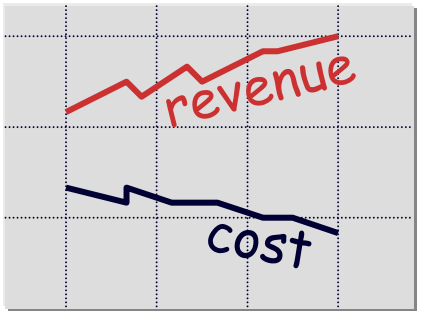
BT

example sustainable business model

for basic data transport



wrap up



- separation of service & network: fine industry goal
  - but idealistic if networks cannot even know their costs
- numerous deep preconceptions to discard
  - x flow rate equality / TCP friendliness badly shares the resource cloud
  - x volume represents cost
  - x humans want known bit-rate
- the elusive problem:
  - traffic cost designed to only be handled by end-points (transport layer)
- solution:
  - reinsert cost information into network layer = re-feedback
- IETF/IRTF drafting architectural shift on layering of resource sharing
  - next mountain to move: add cost accountability (re-ECN) to IP
- once resource sharing fixed properly at the neck of the hourglass
  - over-restrictive lower layer controls can be removed
  - opens new space for service innovation

BT

more info...

- The whole story in 5 pages
  - Bob Briscoe, "A Fairer, Faster Internet Protocol", IEEE Spectrum (Dec 2008)
- Inevitability of policing
  - The Broadband Incentives Problem, Broadband Working Group, MIT, BT, Cisco, Comcast, Deutsche Telekom / T-Mobile, France Telecom, Intel, Motorola, Nokia, Nortel (May '05 & follow-up Jul '06) <cfp.mit.edu>
- 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)
- Understanding why QoS interconnect is better understood as a congestion issue
  - Bob Briscoe and Steve Rudkin "Commercial Models for IP Quality of Service Interconnect" BT Technology Journal 23 (2) pp. 171--195 (April, 2005)
- Network utility optimisation & stability analysis
  - [Kelly05] Frank kelly and Thomas Voice, "Stability of End-to-End Algorithms for Joint Routing and Rate Control" ACM CCR 35(2) 5-12 (Jan 06)
- Equitable quality video streaming
  - [Crabtree09] B. Crabtree, M. Nilsson, P. Mulroy and S. Appleby "Equitable quality video streaming" Computer Communications and Networking Conference, Las Vegas, (January 2009)
- Re-architecting the Internet:
  - The Trilogy project
- **Re-ECN & re-feedback project page: <http://www.cs.ucl.ac.uk/staff/B.Briscoe/projects/refb/>**

BT

sustainable IP resource sharing

Q&A