

session QoS vs bulk QoS

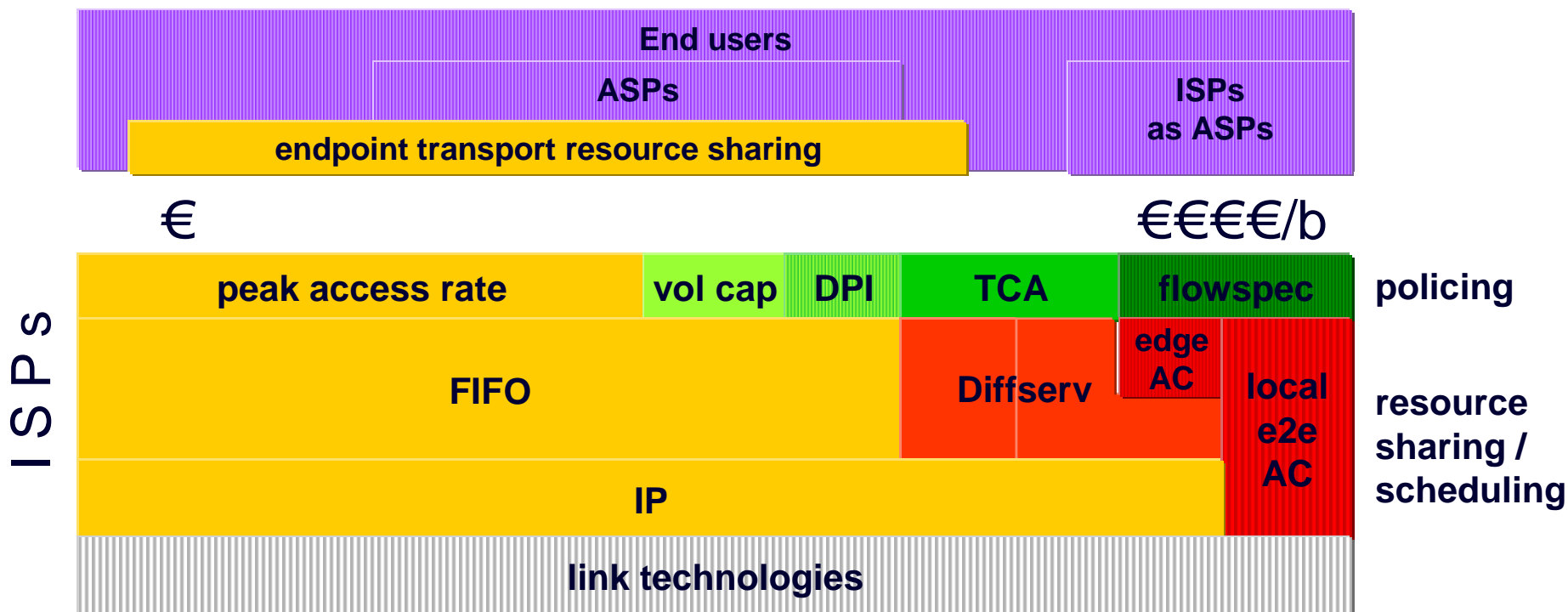
Bob Briscoe
Chief Researcher, BT Group
Oct 2008



QoS bypass

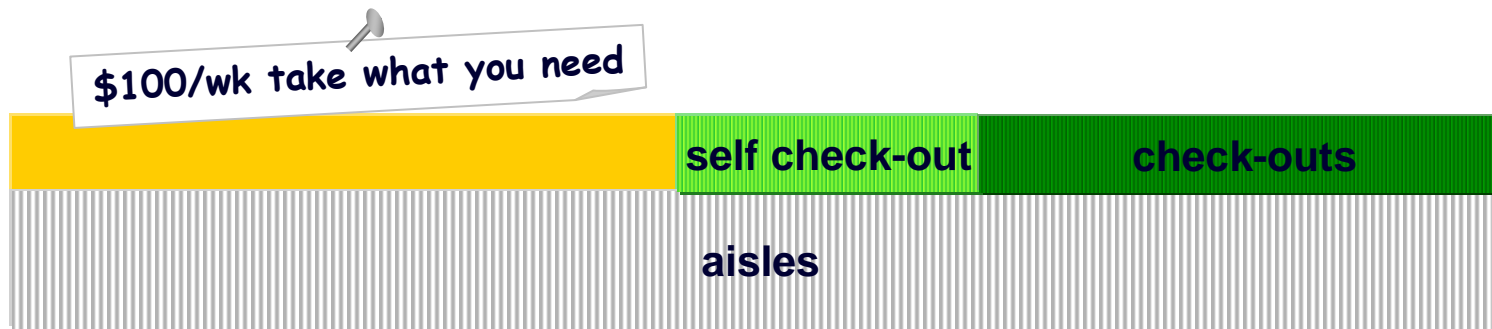
Q. **bulk** or **session** QoS?

A. bulk, but **BE bulk** QoS

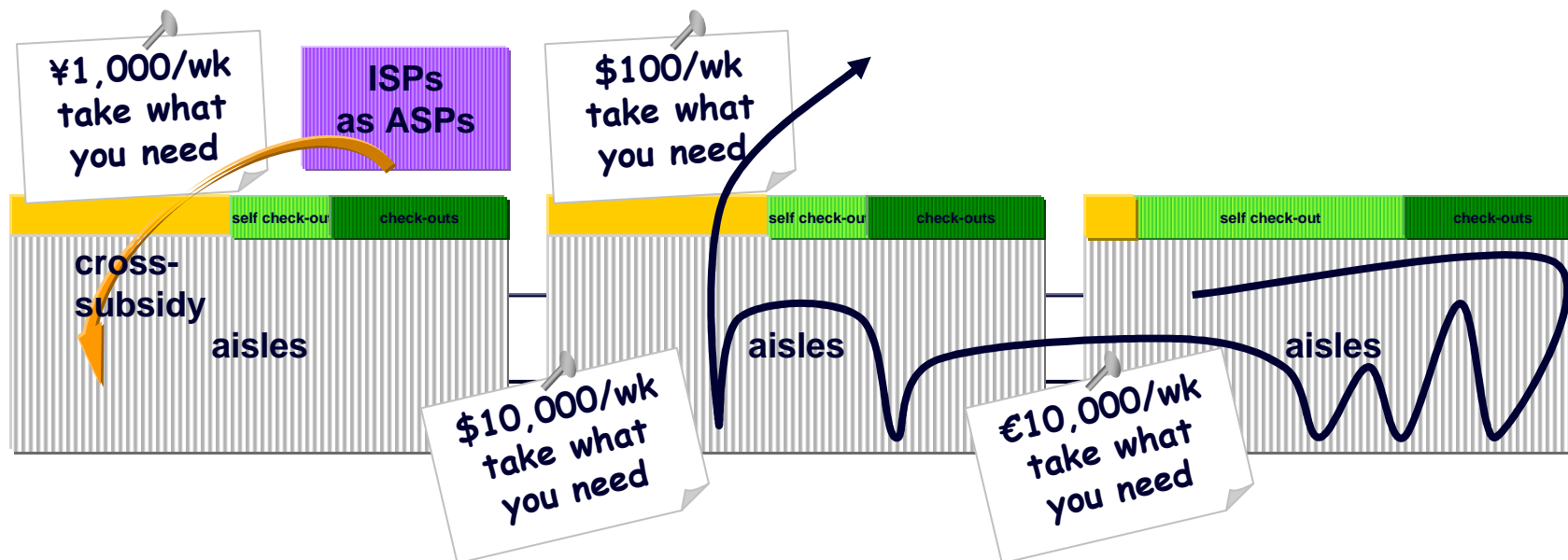


- QoS = differentiated congestion delay & bandwidth
- as link rates increase, congestion delay becoming a non-problem
- all the bandwidth-demanding applications are taking the QoS they need
 - just taking a larger than average cost share of the best efforts service

the information supermarket



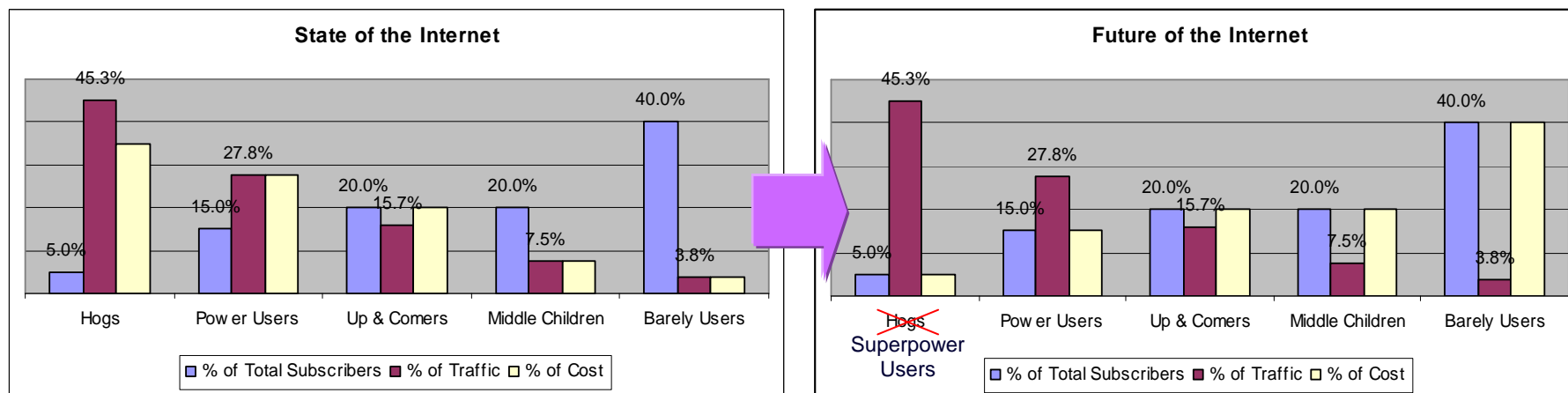
QoS interconnect



- evolution by company death is too slow
 - years
- need market evolution (by financial perf)
 - months or weeks

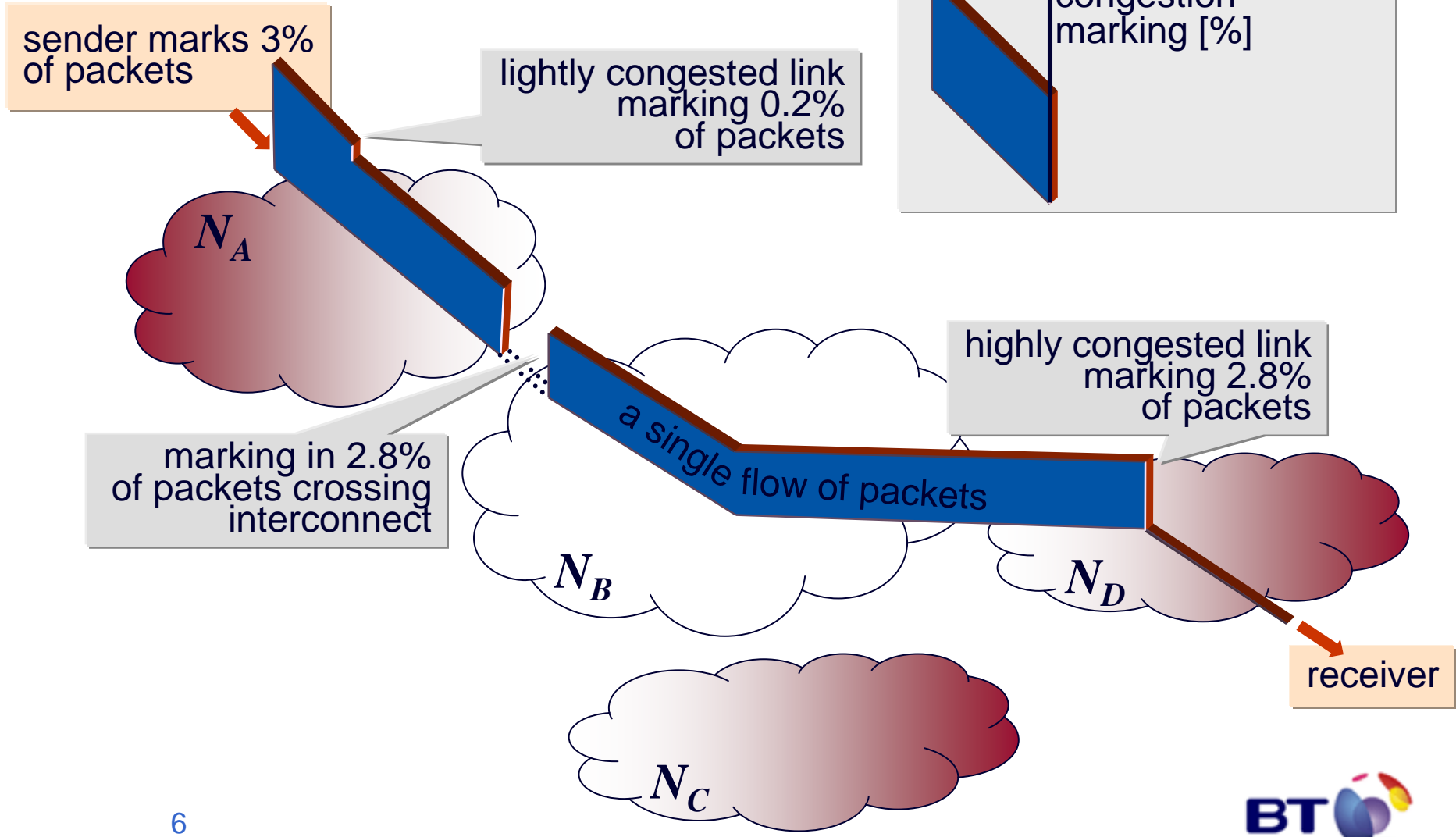
QoS interconnection includes BE QoS

- QoS interconnection is **not** just about explicit QoS mechanisms
- starts with visibility of BE costs
- including at interconnect [Laskowski06, Briscoe05]...



- this is how to get to this future
 - where apps minimise cost, even if they transfer large volumes
 - (limiting peak volume will wrongly cap BitTorrent DNA)

automatic interconnect usage cost allocation



interconnect aggregation

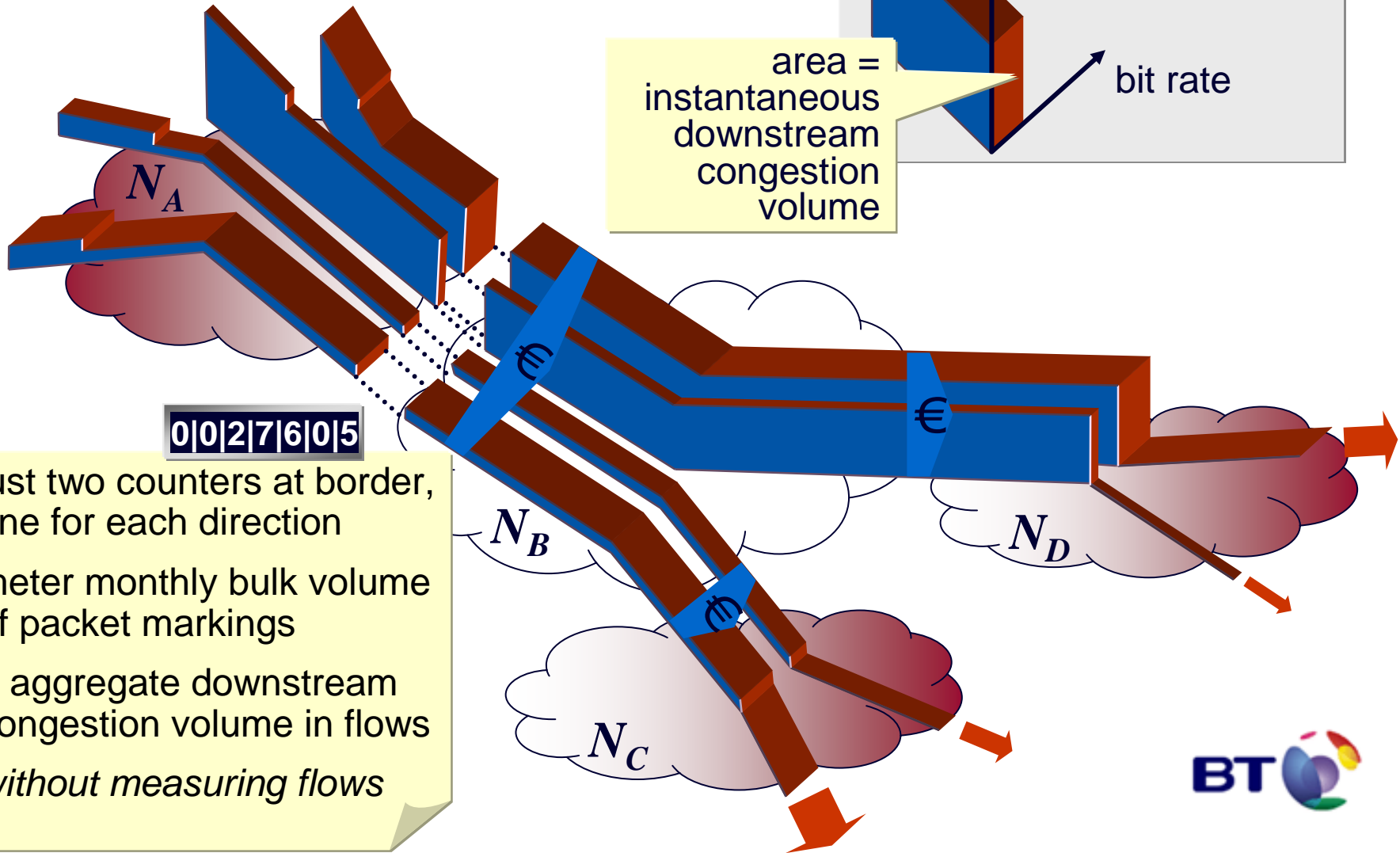
simple internalisation of all externalities
'routing money'

legend:

re-ECN
downstream
congestion
marking [%]

bit rate

area =
instantaneous
downstream
congestion
volume



0|0|2|7|6|0|5

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



Acceptable Use Policy

Your 'congestion volume' allowance:
1GB/month (= 3kb/s continuous)

This only limits the traffic you can try to transfer above the maximum the Internet can take when it is congested.

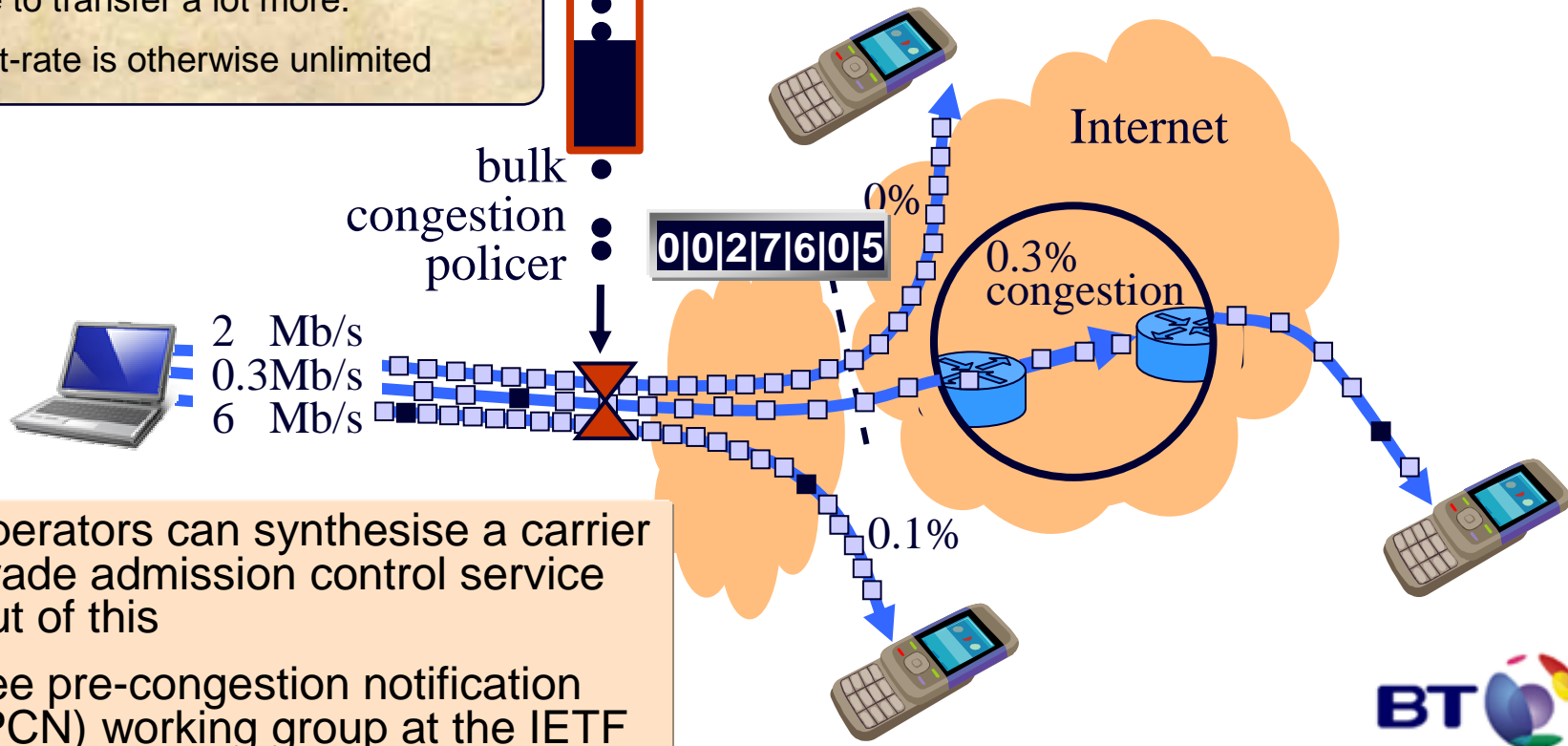
Under typical conditions this will allow you to transfer about **70GB per day**.

If you use software that seeks out uncongested times and routes, you will be able to transfer a lot more.

Your bit-rate is otherwise unlimited

differentiated services
& admission control
just happen

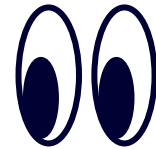
- as an attribute of the customer contract, not the network
- customer can roam without changing network



- operators can synthesise a carrier grade admission control service out of this
- see pre-congestion notification (PCN) working group at the IETF

summary

- everyone's got their eye on the wrong balls



- volume → cost (congestion)
- AF, EF & session QoS → BE QoS cost policing
 - intra-domain & inter-domain

session QoS vs bulk QoS

Q&A
refs
spare slides



more info

interconnected visibility of BE cost

- The Internet's missing link: rest of path metrics at interconnect
[Laskowski06] Paul Laskowski and John Chuang, "Network Monitors and Contracting Systems: Competition and Innovation" In: Proc. ACM SIGCOMM'06, Computer Communication Review 36 (4) pp. 183--194 (September, 2006)
- A way to do rest of path metrics
[Briscoe05] Bob Briscoe, Arnaud Jacquet, Carla Di-Cairano Gilfedder, Andrea Soppera and Martin Koyabe, "Policing Congestion Response in an Inter-Network Using Re-Feedback" In: Proc. ACM SIGCOMM'05, Computer Communication Review 35 (4) (September, 2005)

pre-congestion notification (PCN)

- Diffserv's scaling problem
[Reid05] Andy B. Reid, *Economics and scalability of QoS solutions*, BT Technology Journal, 23(2) 97–117 (Apr'05)
- PCN interconnection for commercial and technical audiences:
[Briscoe05] Bob Briscoe and Steve Rudkin, *Commercial Models for IP Quality of Service Interconnect*, in BTTJ Special Edition on IP Quality of Service, 23(2) 171–195 (Apr'05) <www.cs.ucl.ac.uk/staff/B.Briscoe/pubs.html#ixqos>
- IETF PCN working group documents
<tools.ietf.org/wg/pcn/> in particular:
[PCN] Phil Eardley (Ed), *Pre-Congestion Notification Architecture*, Internet Draft <www.ietf.org/internet-drafts/draft-ietf-pcn-architecture-06.txt> (Sep'08)
[re-PCN] Bob Briscoe, *Emulating Border Flow Policing using Re-PCN on Bulk Data*, Internet Draft <www.cs.ucl.ac.uk/staff/B.Briscoe/pubs.html#repcn> (Sep'08)

these slides

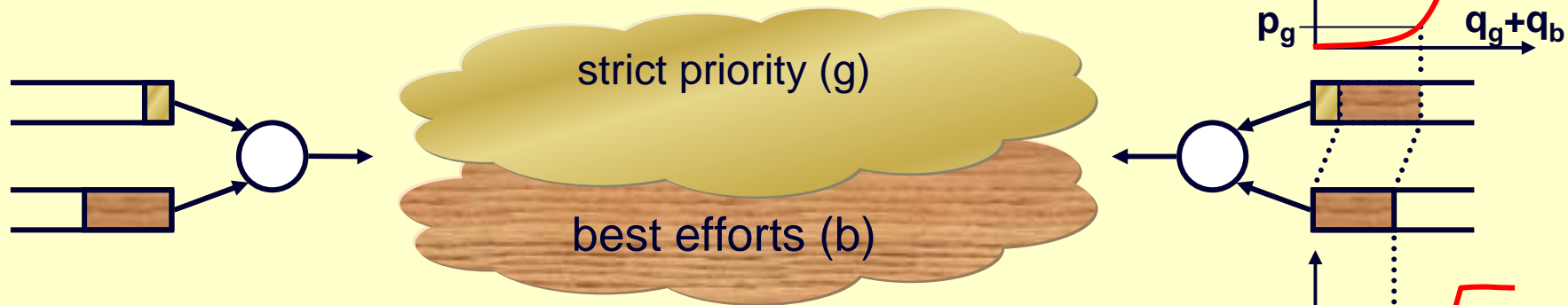
<www.cs.ucl.ac.uk/staff/B.Briscoe/present.html>



shouldn't network charge more for lower congestion?

- apologies for my sleight of hand
 - actually aiming to *avoid* congestion impairment (loss / delay)
 - congestion marking = congestion *avoidance* marking
 - alternatively, congestion marking = price marking
- clearly should charge more for higher 'price marking'

- Diffserv example may help [Gibbens02]






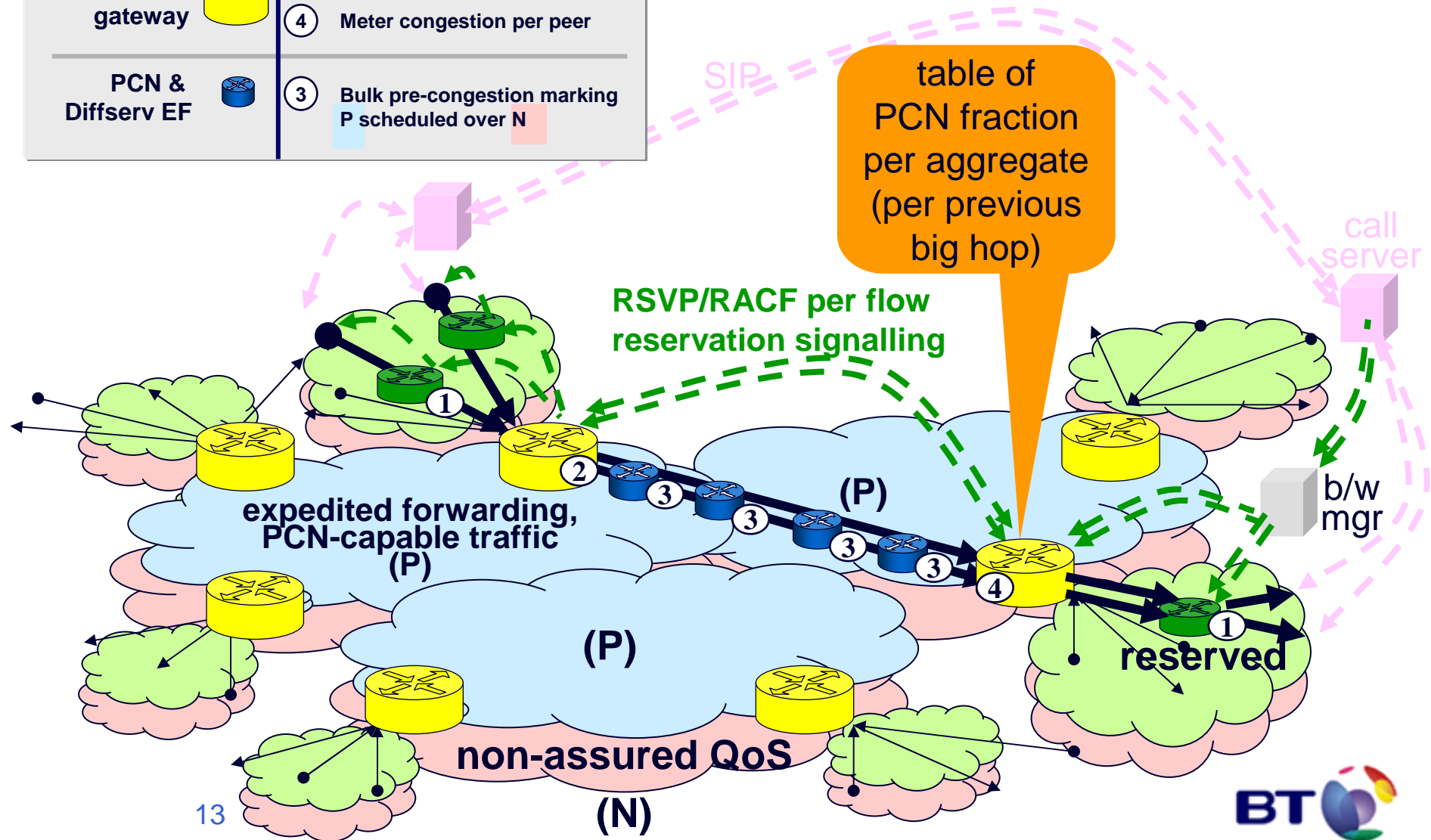
weak competition
 price of *expectation* of better service
 arbitrarily higher $p_g \gg p_b$

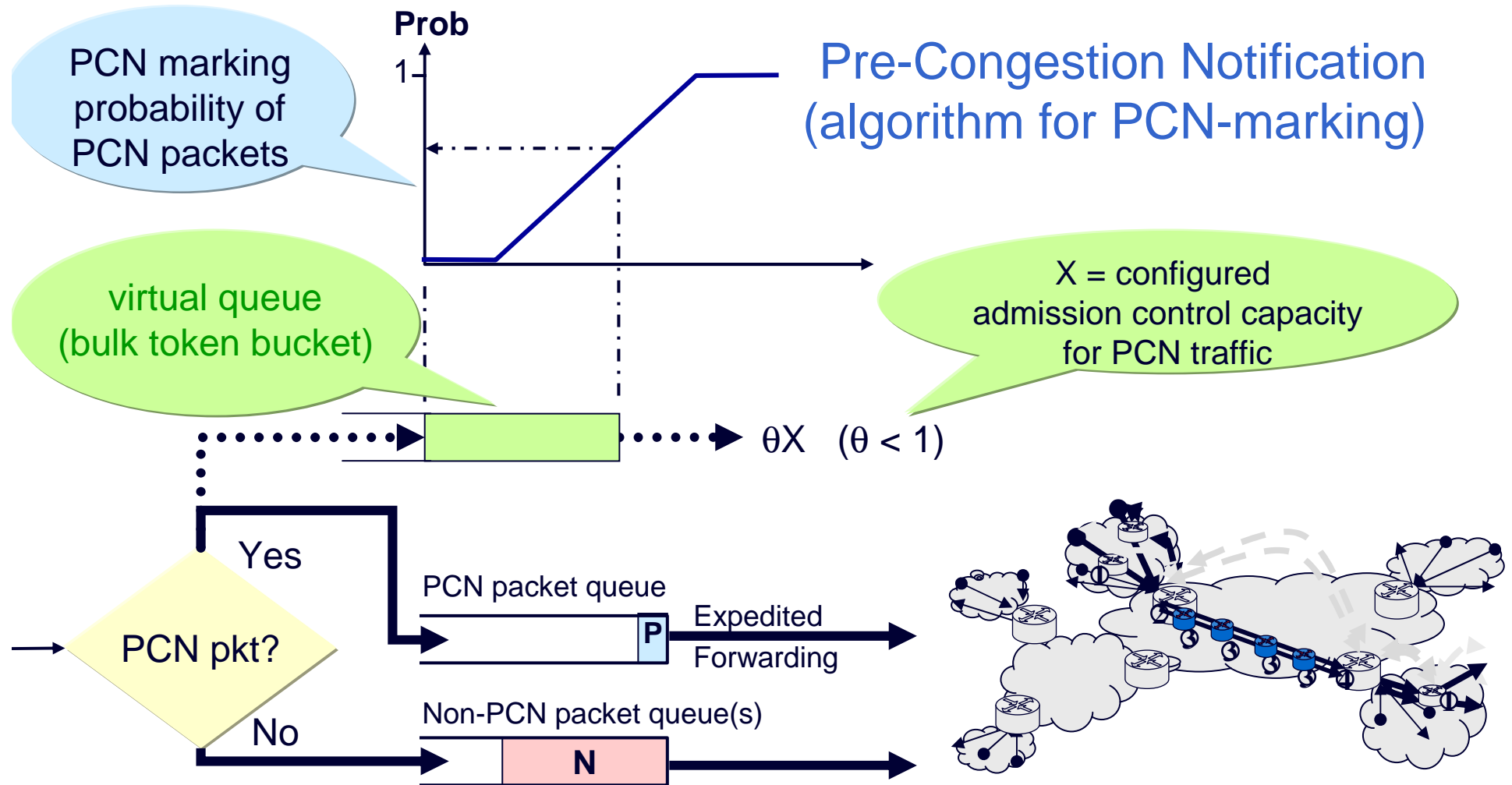
perfect competition
 price differential \propto cost differential
 $p_g \geq p_b$

PCN system arrangement

highlighting 2 flows

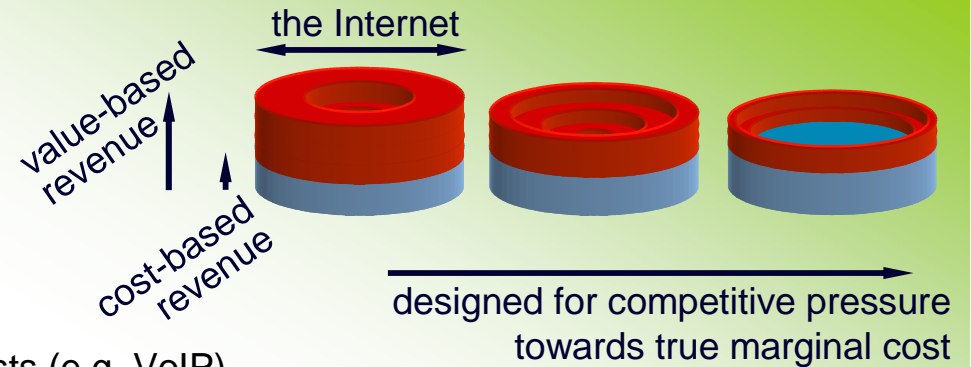
IP routers	Data path processing
Reservation enabled 	① Reserved flow processing
RSVP/PCN gateway 	② Policing flow entry to P ④ Meter congestion per peer
PCN & Diffserv EF 	③ Bulk pre-congestion marking P scheduled over N



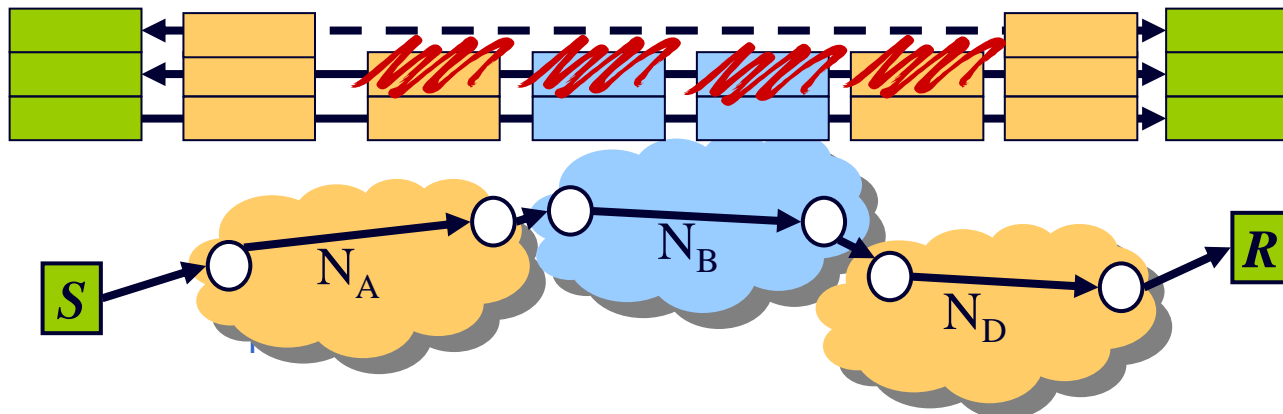


- virtual queue (a conceptual queue – actually a simple counter):
 - drained somewhat slower than the rate configured for adm ctrl of PCN traffic
 - therefore build up of virtual queue is 'early warning' that the amount of PCN traffic is getting close to the configured capacity
 - NB mean number of packets in real PCN queue is still very small

value-based charges over low cost floor



- over IP, currently choice between
 - A. “good enough” service with no QoS costs (e.g. VoIP)
 - but can brown-out during peak demand or anomalies
 - B. fairly costly QoS mechanisms – either admission control or generous sizing
- this talk: where the premium end of the market (B) is headed
 - a new IETF technology: pre-congestion notification (PCN)
 - service of ‘B’ but mechanism cost competes with ‘A’
 - assured bandwidth & latency + PSTN-equivalent call admission probability
 - fail-safe fast recovery from even multiple disasters
- core networks could soon fully guarantee sessions without touching sessions
 - some may forego falling session-value margins to compete on cost



app signal (SIP)	per session
QoS admission	
priority forwarding	bulk data
& PCN	

