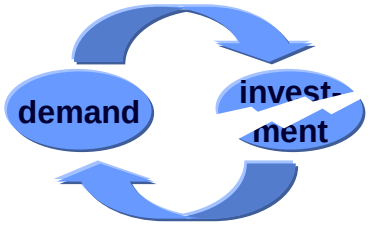
investment and load control incentives:

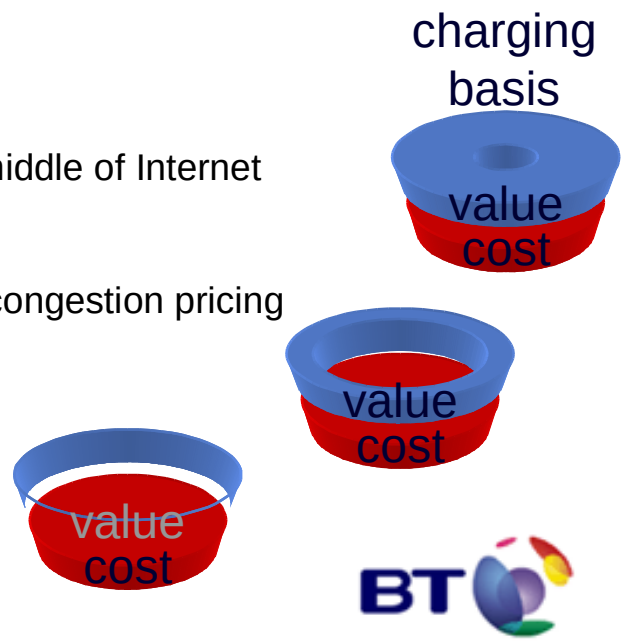
broadband evolution from value to cost

Bob Briscoe
BT Networks Research Centre
May 2005

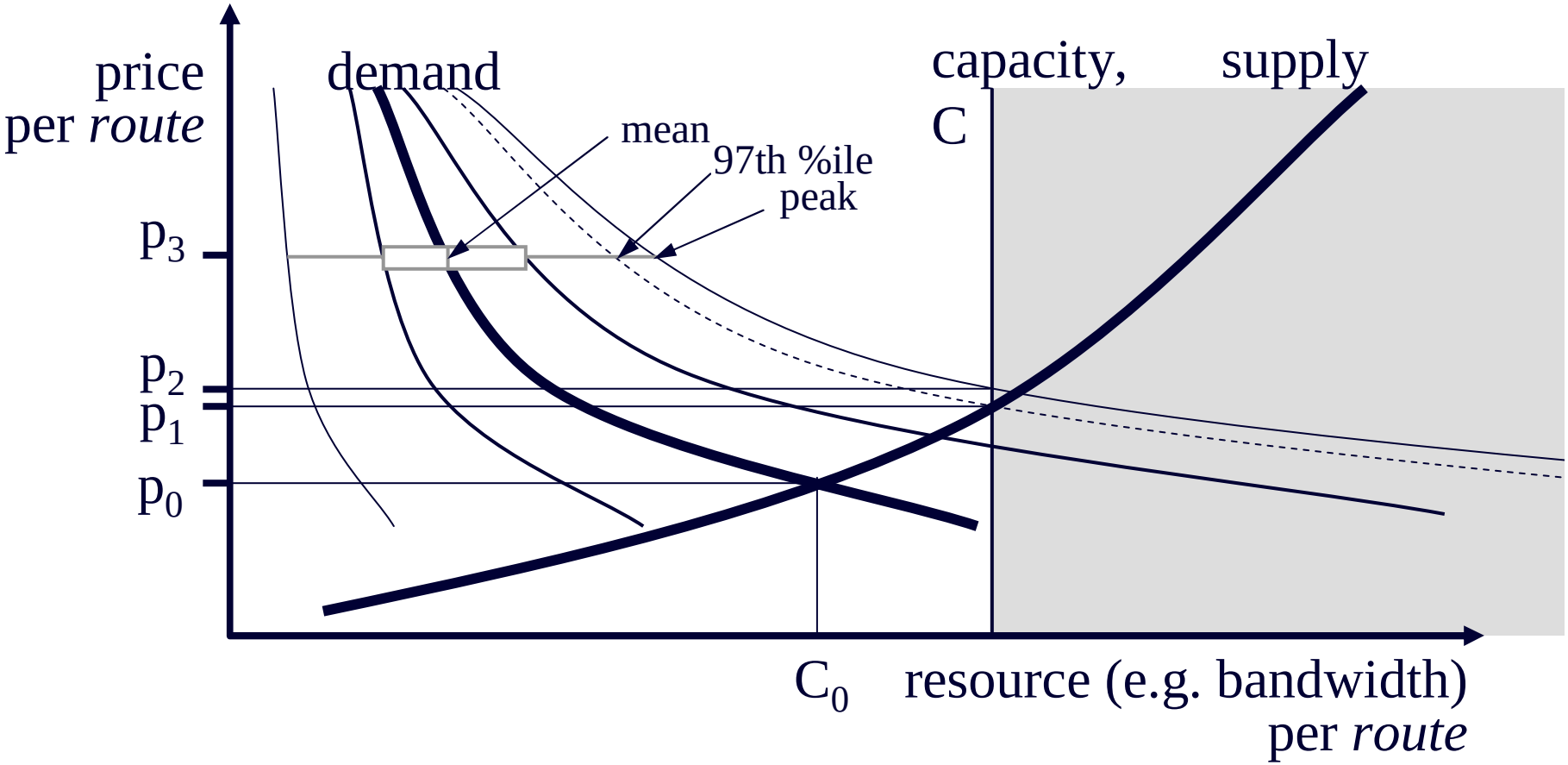
menu

- congestion pricing
  - tutorial: economics & engineering
  - ✓ computer-assisted user incentives: v. cheap, strategy-proof sol'n
  - × investment incentives: poor – commoditised, highly competitive
- salvation?
  - competition far from perfect
  - ⇒ value-based not cost-based charging
- evolution to end-game
  - competition: cost-based charging hole grows from middle of Internet
- end-game
  - internal markets (wholesale/interconnect) driven to congestion pricing
  - retail human-customer markets layered on top
- googly
  - fast or total commoditisation





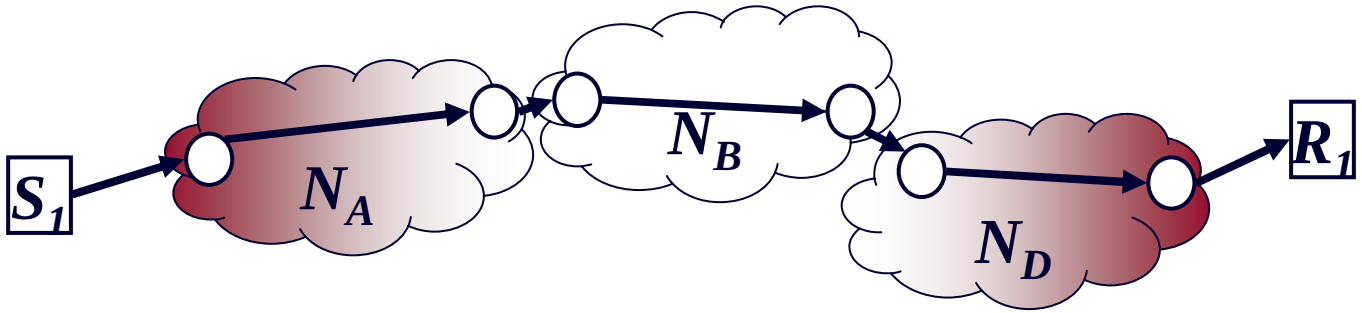
context: demand varies fast, supply slowly



- mix of pricing & throttling incentives – but how?
  - note: 'throttling' = caps, quotas, rate policing, shaping

BT

context: investment costs



- selling QoS = managing risk of congestion
  - if no risk of congestion, can't sell QoS
  - congestion risk highest in access nets (cost economics of fan-out)
  - but *small* risk in cores/backbones (failures, anomalous demand)

BT

congestion pricing
tutorial: economics & engineering

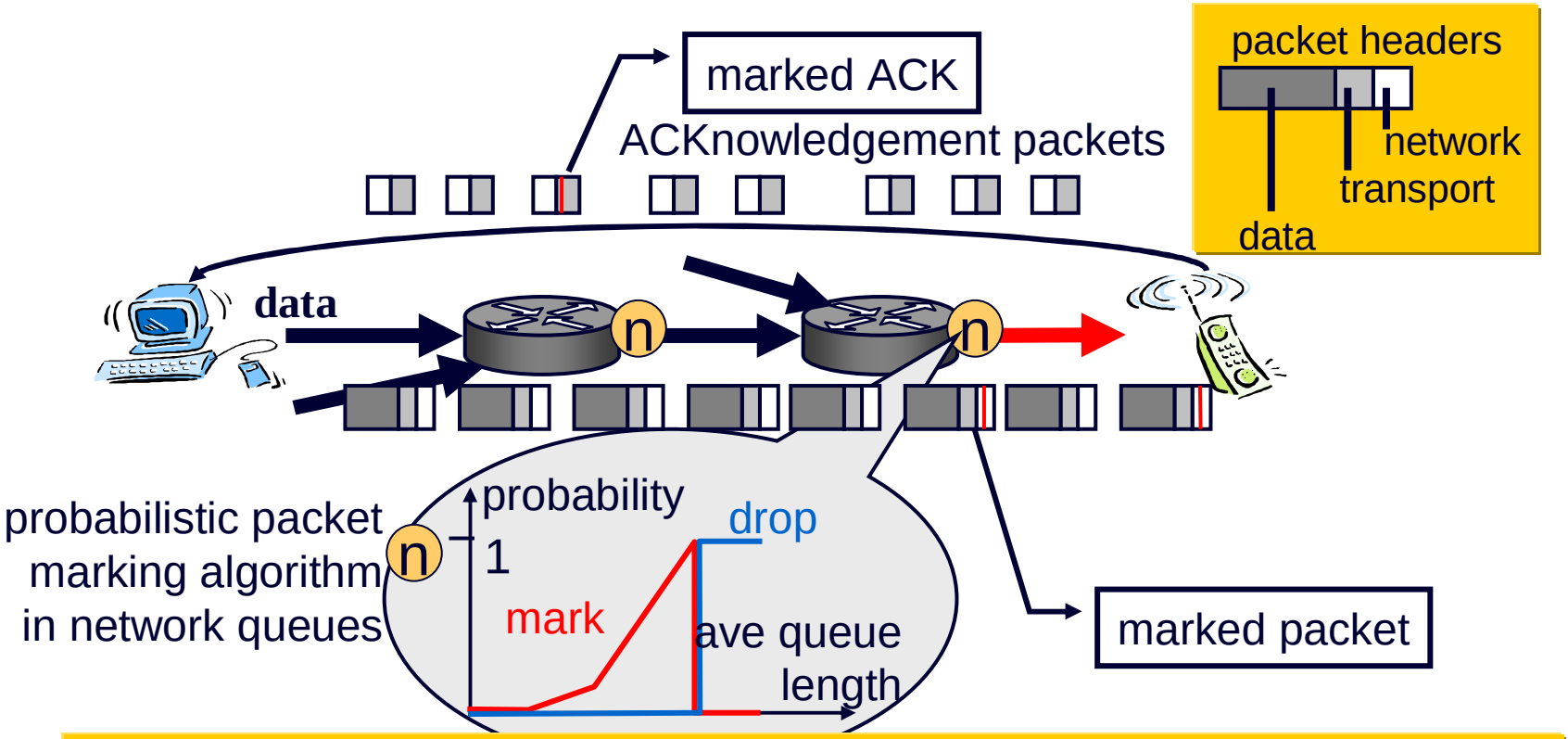
an Internet proof against
strategising machines

costs

- infrastructure costs: sunk
- operational costs: usage independent
- usage and congestion: cost operator nothing
- congestion: costs those sharing each resource
  - **congestion definition:**
    probability that serving one (packet) will cause another not to be served to its reqs

- approximations to congestion metrics (we'll come back to these)
  1. by time: time-of-day volume pricing
  2. by route: on/off-net, domain hops, distance
  3. by class of service: flat fee for each class, volume price for each class
- accurate congestion metrics (in all 3 dimensions)
  - loss rate
  - explicit congestion notification…

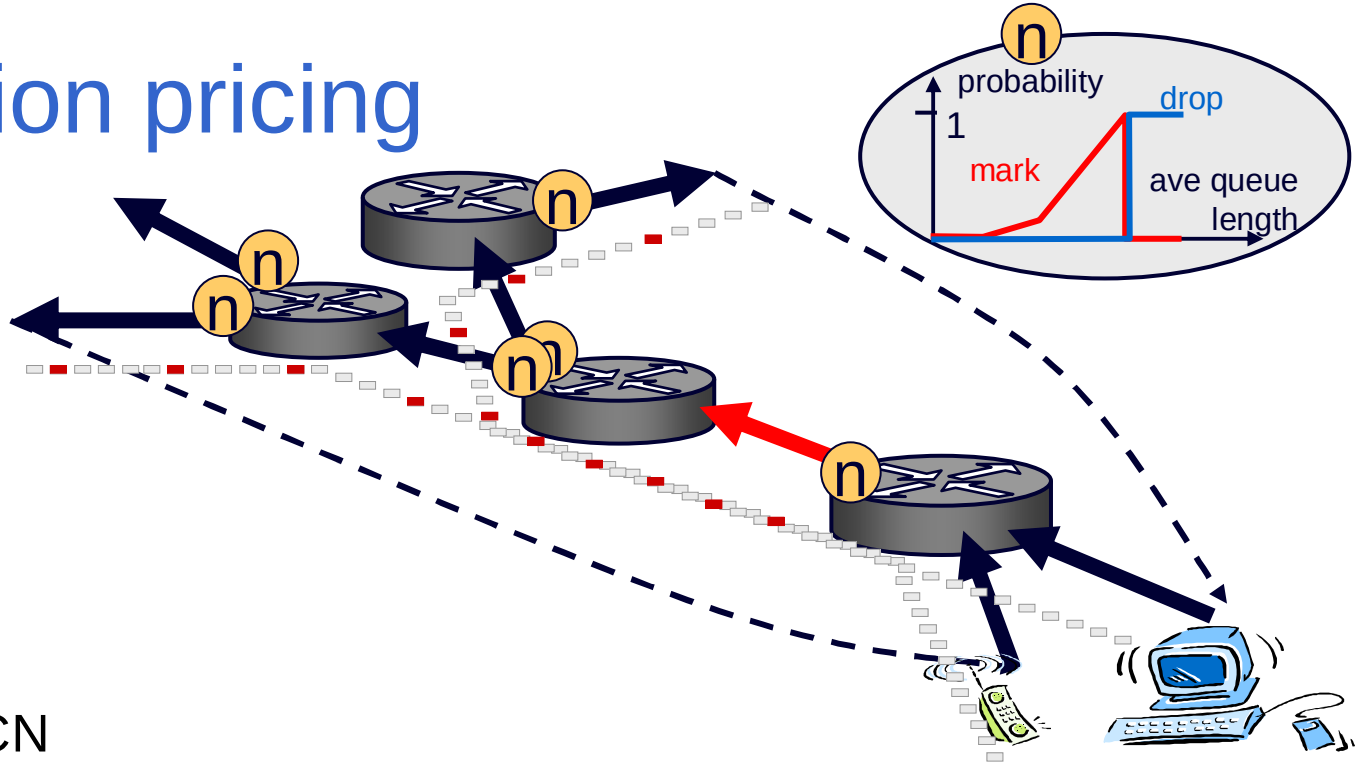
pre-requisite knowledge:
explicit congestion notification (ECN)



IETF proposed std: RFC3168; most recent change to IPv4&6 (Sep 2001)
implemented in commercial routers & Linux servers but not Windows

BT

congestion pricing

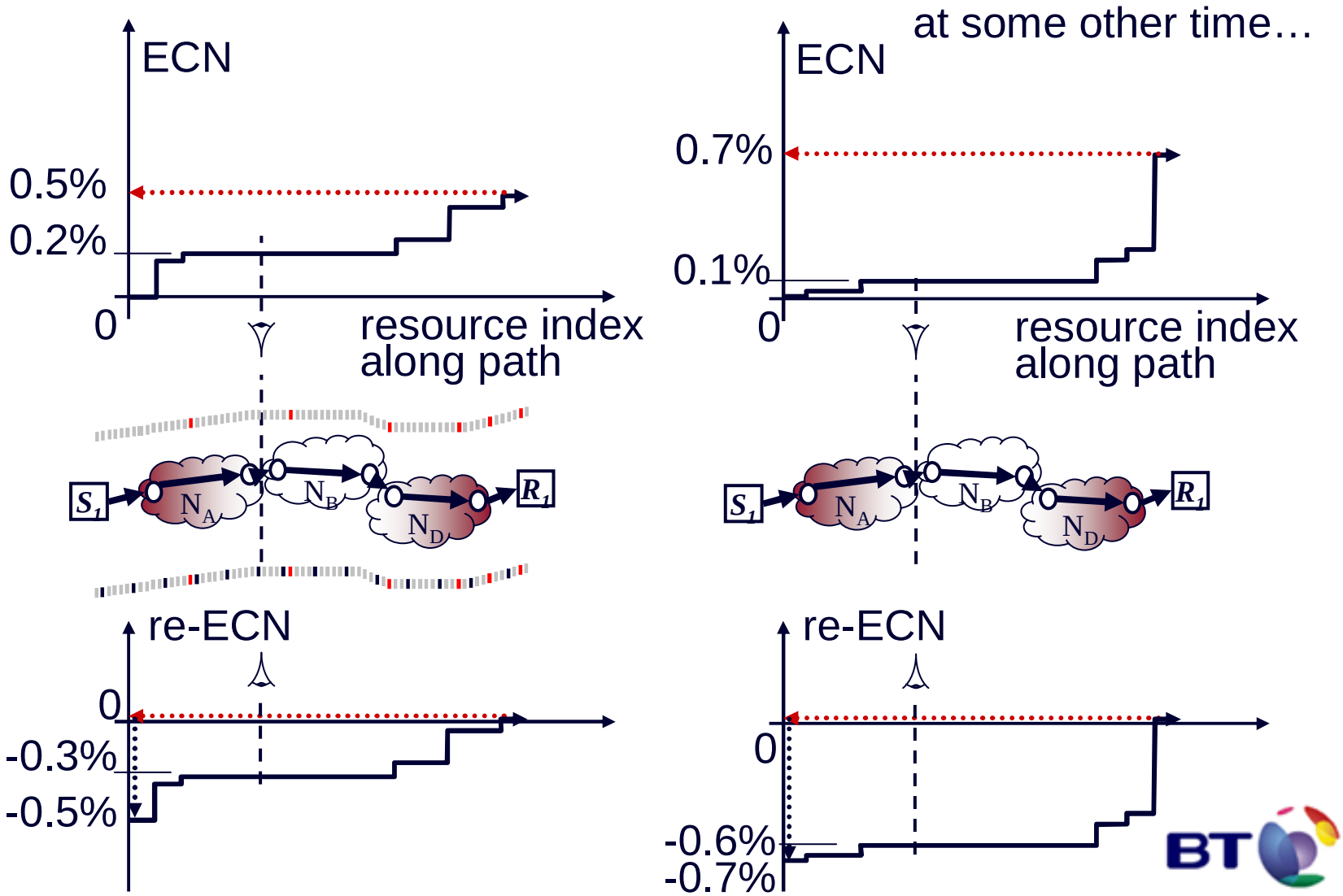


- without ECN
  - first sign of congestion is loss
  - loss is an impractical metric for charging (metering holes)
- with ECN
  - notifies incipient congestion before service degrades
  - volume charging but only of marked packets ⇒ congestion charging

BT

re-ECN: receiver-aligned ECN [Briscoe05]

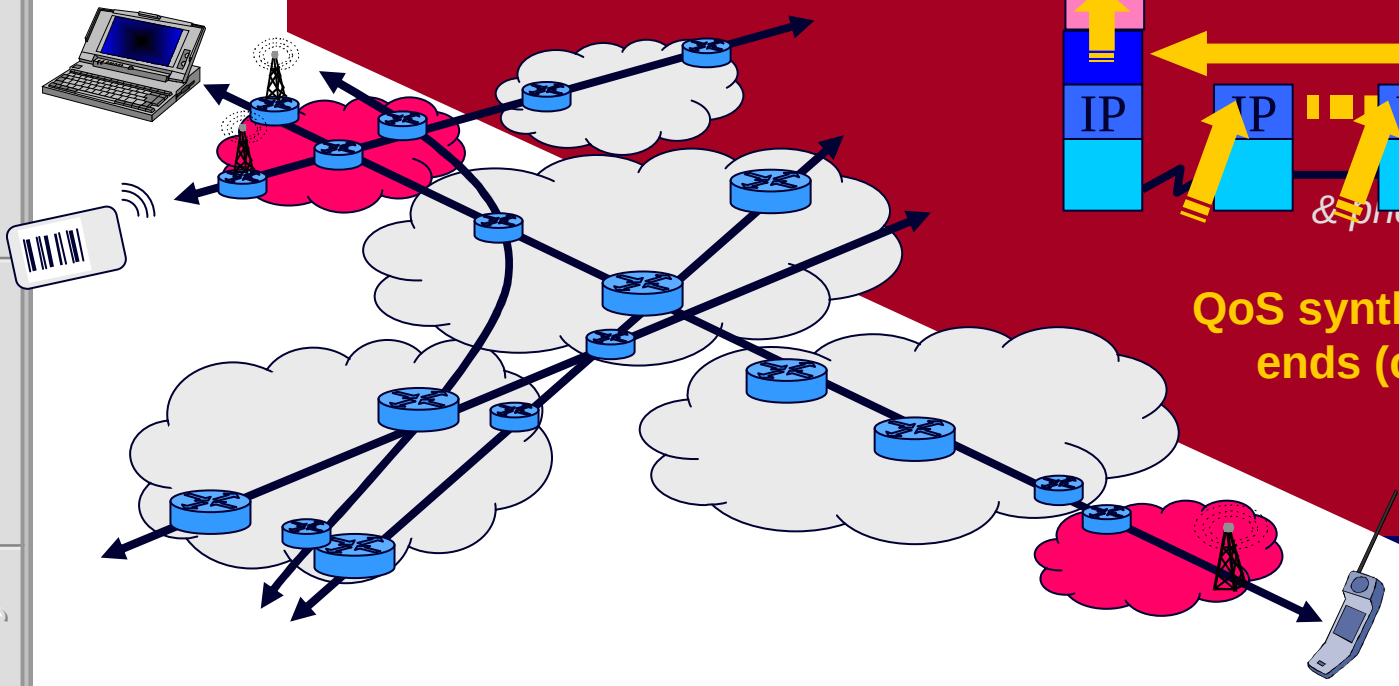
downstream path characterisation

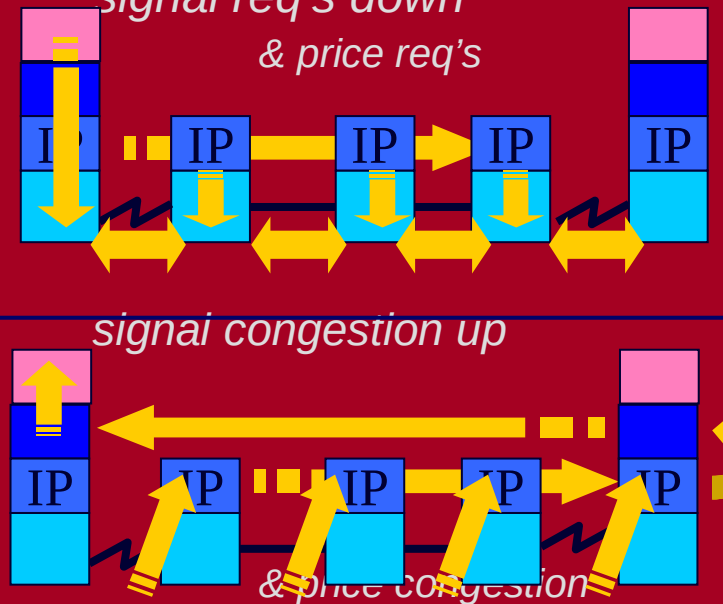


seamless resource control

😐 **traditional (optional):**
optimise ea subnet separately
e.g. Diffserv (open-loop)

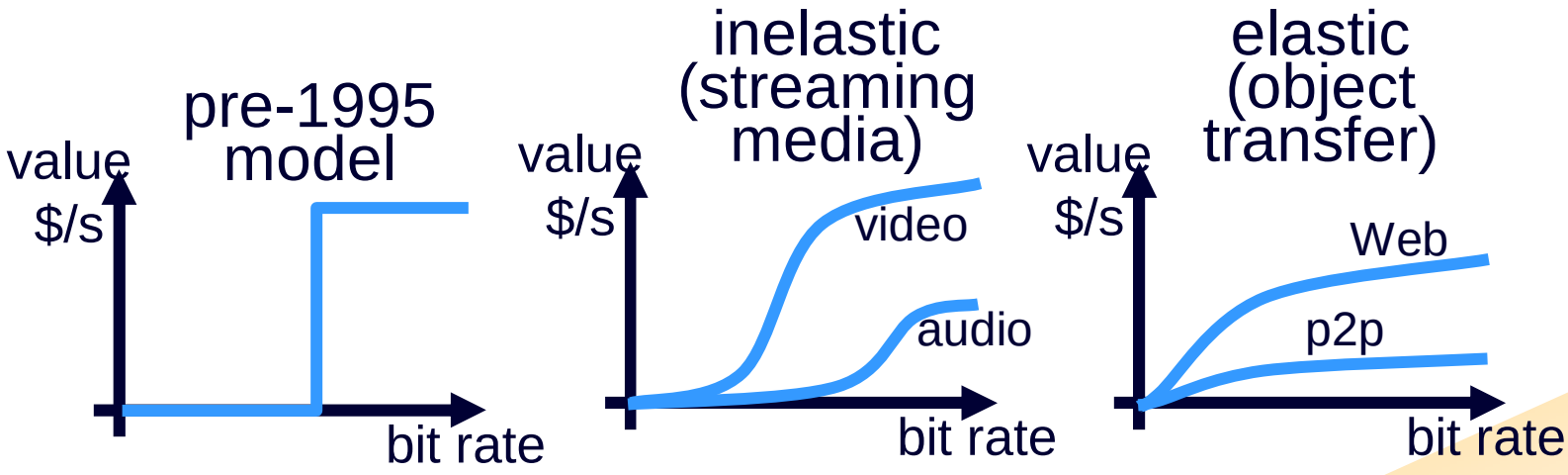
☺ **new (required):**
optimise all paths together





QoS synthesised by the ends (closed-loop)

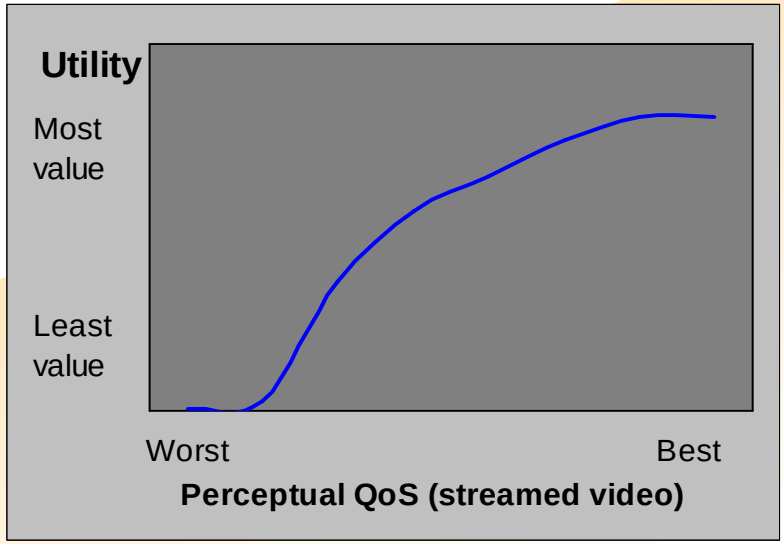
value: curve families



theoretical
[Shenker95]

&

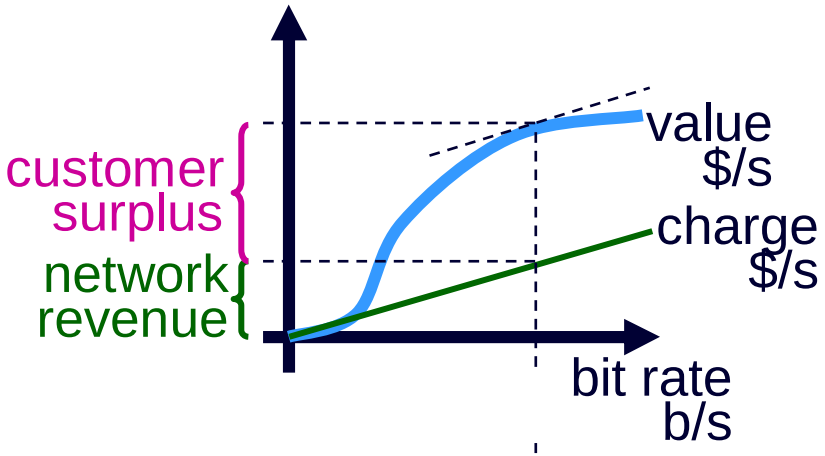
actual
value models

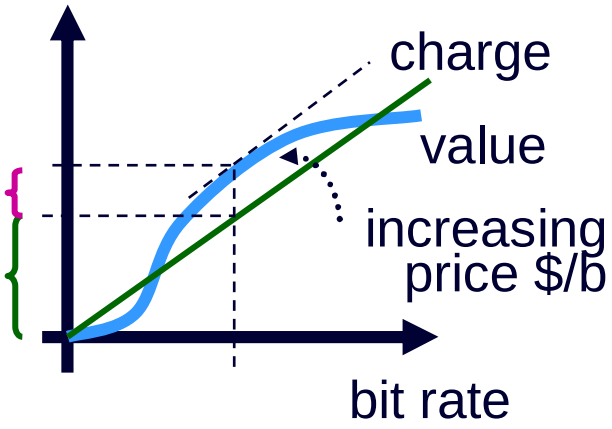


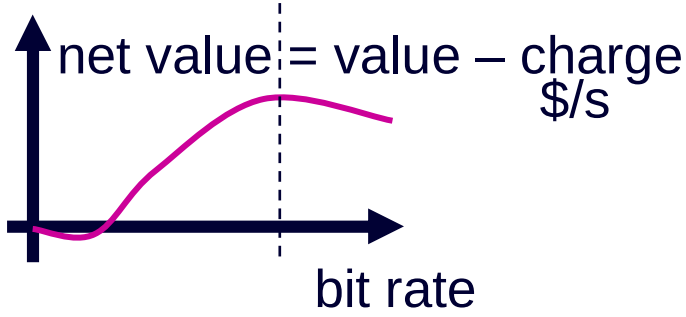
average of normalised curves from a set of experiments on paying customers [Hands02]

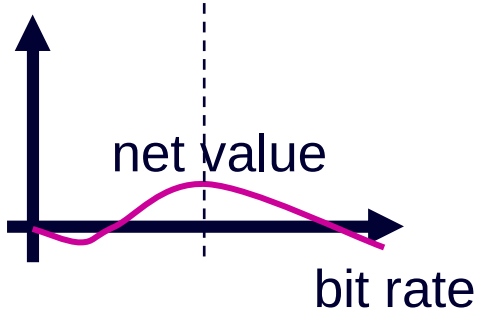
BT

value – cost: customer's optimisation

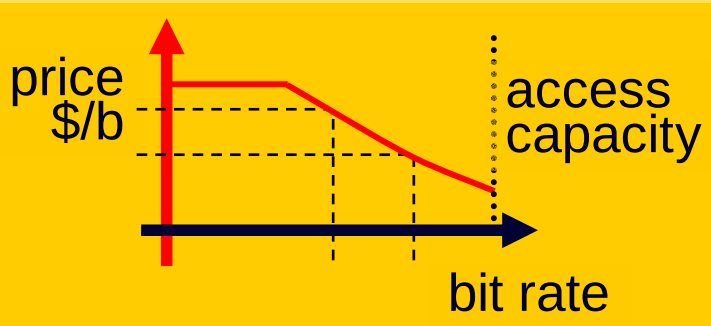




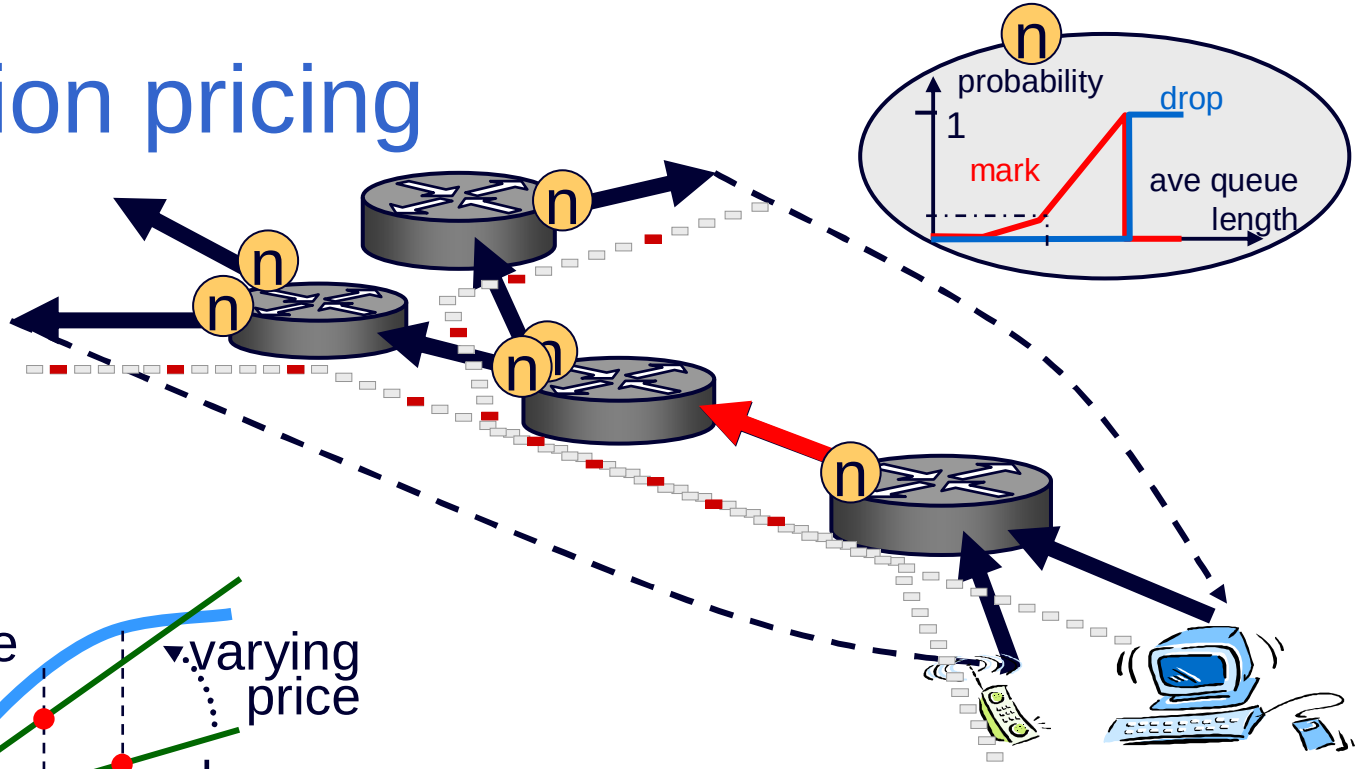


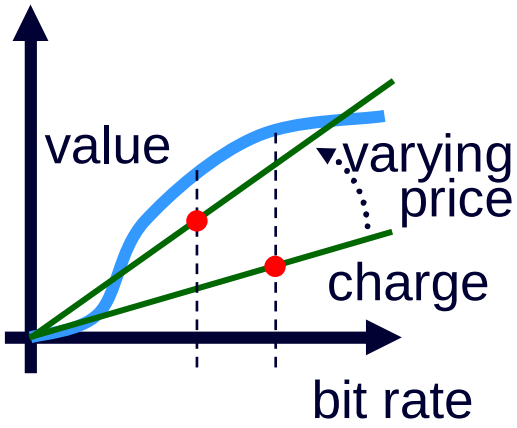


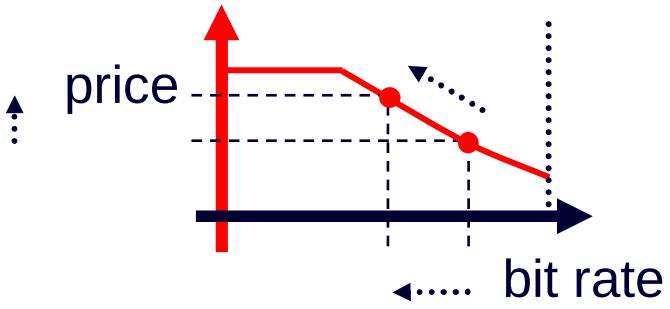
demand curve derivable from value curves



congestion pricing





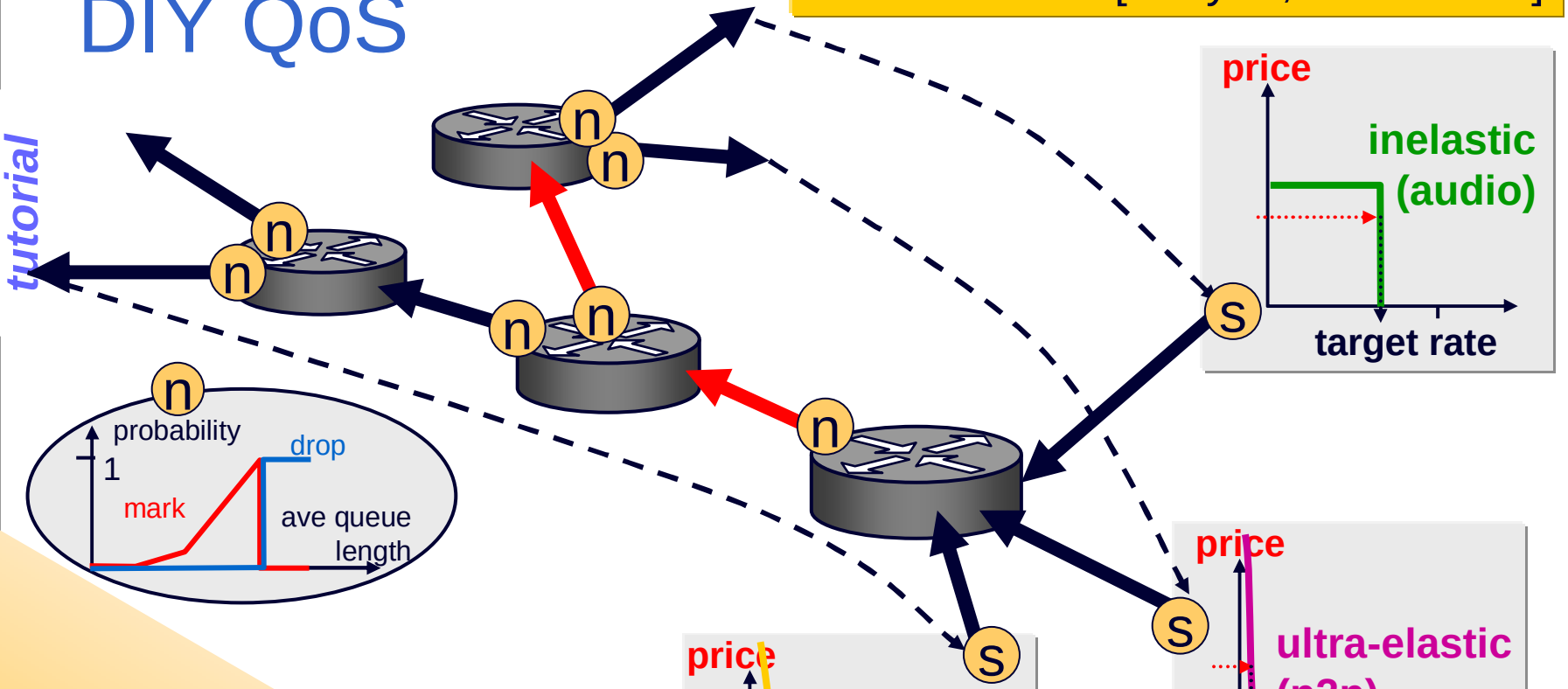


- volume charging
  - but only of marked packets

⇒ congestion charging

DIY QoS

maximises social welfare across whole Internet [Kelly98, Gibbens99]



(shadow) price = ECN

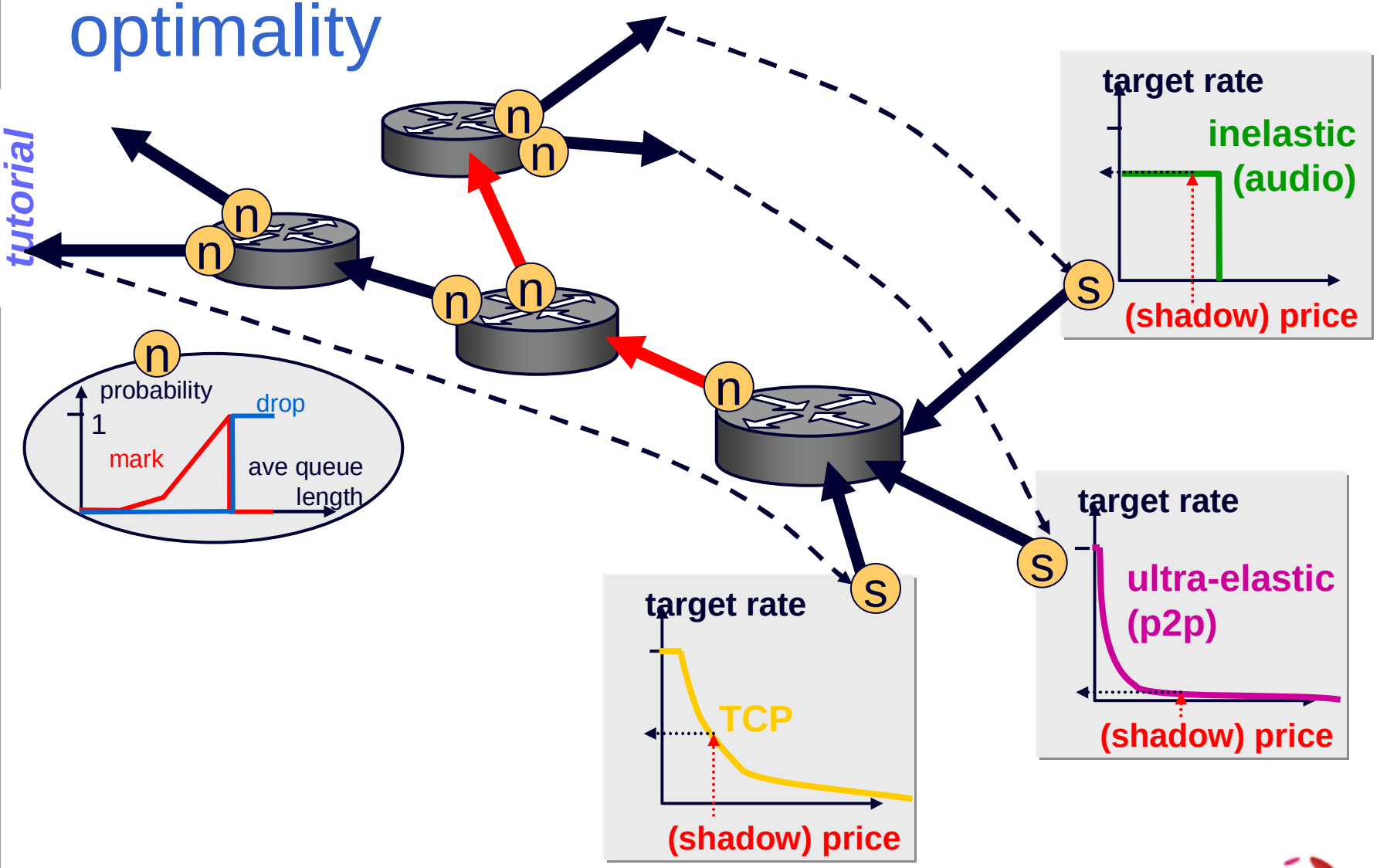
n network algorithm — supply

s sender algorithms — demand

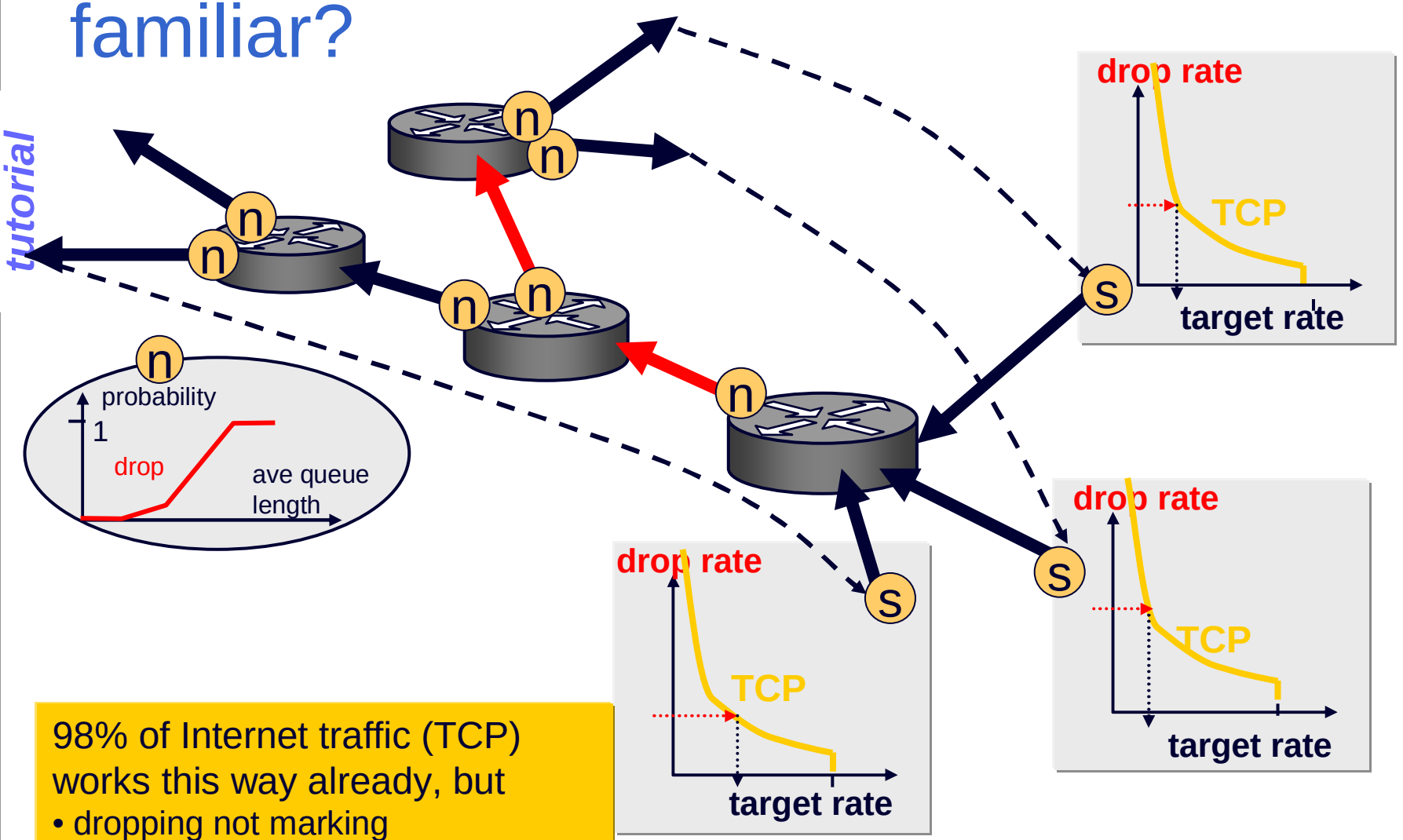
alternative version of previous slide

for those who prefer their graphs with the independent variable horizontal

optimality



familiar?



98% of Internet traffic (TCP) works this way already, but

- dropping not marking
- senders respond voluntarily
  ***as if* congestion charged**
- every sender responds identically

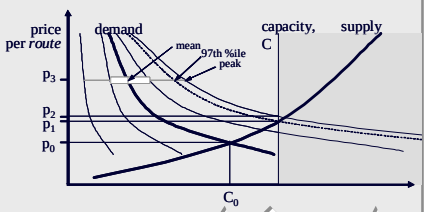
BT

shaping short-term demand with flat pricing

recall…

context: demand varies fast, supply slowly

- mix of pricing & throttling incentives – but how?
  - note: 'throttling' = caps, quotas, rate policing, shaping



- human customers highly averse to unpredictable pricing
- answer: congestion-based throttling – example [Briscoe05]:
  - customer pays monthly flat fee subscription (congestion credit limit)
  - congestion 'cost' metered by customer's access provider
  - if (variable) cost in danger of exceeding (flat) income, throttle traffic
  - can focus throttling proportionate to congestion on each route
- cf. volume caps (but better)

BT

supply side

recall…

costs

- congestion: costs those sharing each resource
- usage and congestion: **cost operator nothing**

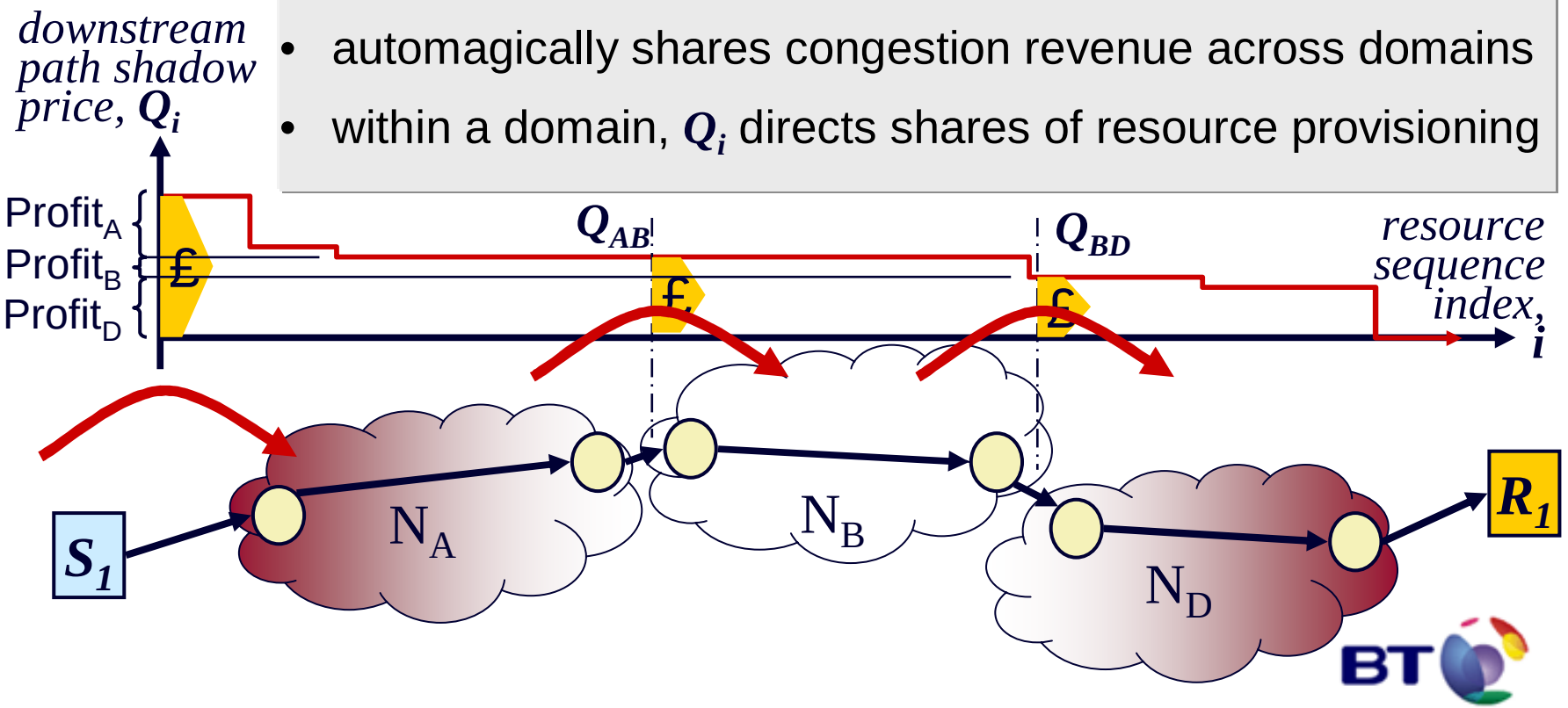
Q so who should collect the congestion charge?

A the operator – offsets the marginal cost of capacity…

BT

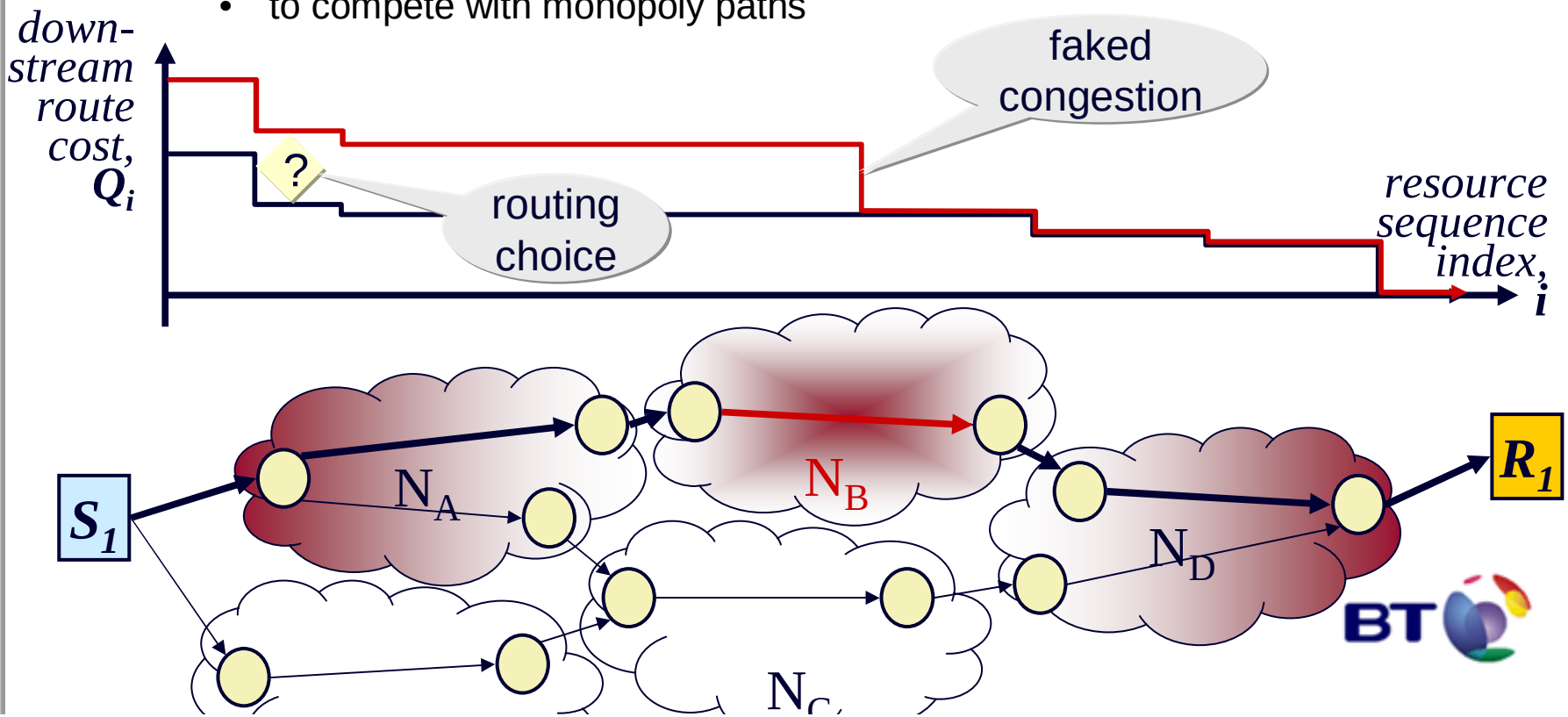
congestion pricing - inter-domain [Briscoe05]

- Q_i = fraction of volume marked with ECN (the shadow price)
- Q_i metered between domains by single bulk counter
- sending domain pays receiving domain congestion charge
  $C = \lambda Q_i$ @ relatively fixed price λ
- automagically shares congestion revenue across domains
- within a domain, Q_i directs shares of resource provisioning



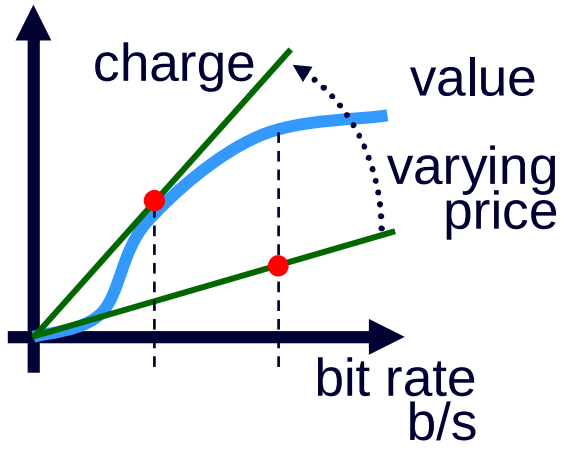
congestion competition – inter-domain routing

- why won't a network overstate congestion?
  - upstream networks will route round more highly congested paths
  - N_A can see relative costs of paths to R_1 thru N_B & N_C
- also incentivises new provision
  - to compete with monopoly paths

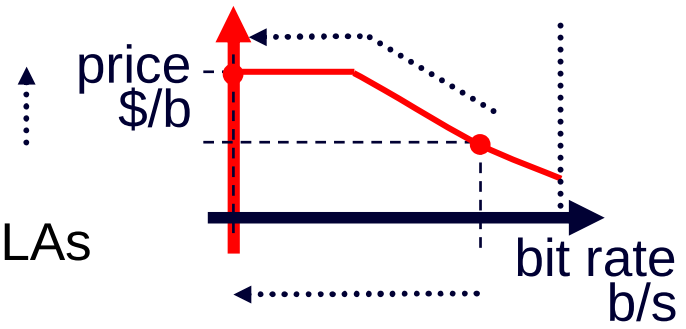


congestion notification also underlies…

- scalable flow admission control
  - for S-shaped value curves (inelastic streaming media)



- class of service pricing
- verifying impairment budgets in SLAs
- resource allocation for VPNs
- …



BT

what's wrong with what we've got?

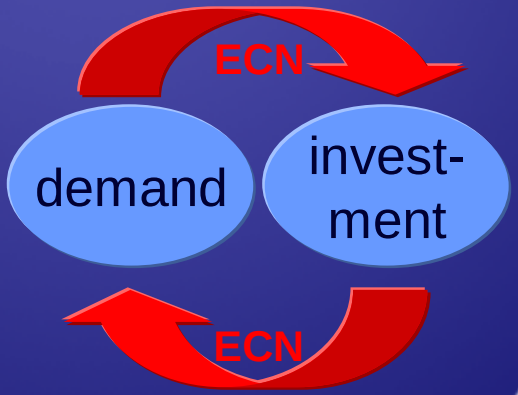
recall…

costs

- approximations to congestion metrics
  1. by time
     - time-of-day volume pricing
  2. by route
     - on/off-net
     - domain hops
     - distance
  3. by class of service
     - flat fee for each class
     - volume price for each class
- dilemma
  - nothing wrong with these… for humans
  - but computers will exploit every gap in every approximation

BT

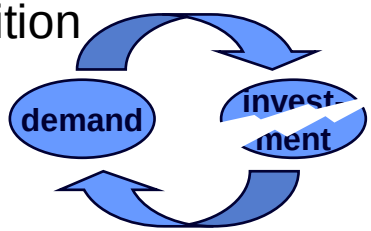
congestion pricing
the hammer for every nail



but…

the sting

- congestion price
  - the minimum price at any time to keep each route fully utilised
  - the price you would expect under perfect competition
  - investment incentives: poor – commoditised



- saving graces
  - competition far from perfect in access networks
  - perfect competition would have to be for every route
  - customers willing to pay premium for predictable price & service

price discrimination by value: feasibility?

inference vs. hiding

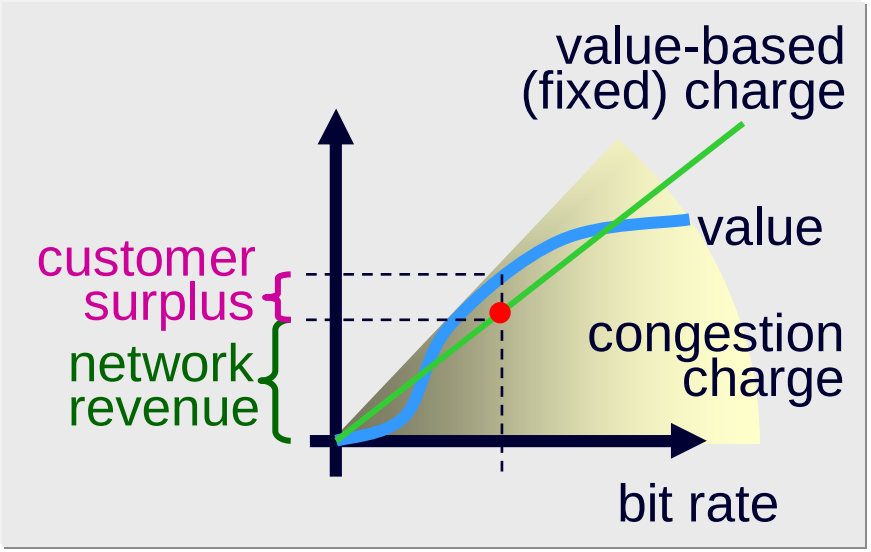
- value is task-specific not application-specific [Bouch00]
  - remote collaboration vs. talking head (both video apps)
  - directory-style look-ups vs. browsing pages (both Web apps)
  - network needs to infer customer intent…
- SMS 10p/100B
  - £1k /1MB audio track?
  - £1M /1GB video?

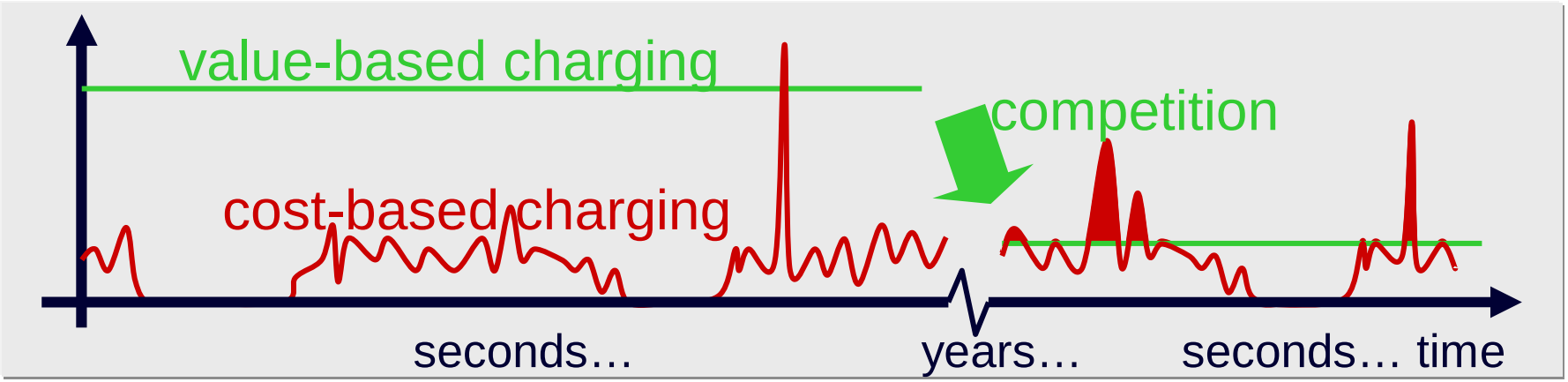
- deep packet inspection (DPI)
  - × expensive
  - × regulatory issues
    - × anti-trust
    - × anti-competitive behaviour
    - × common carrier immunity threatened
  - × routine encryption (VPNs, e-commerce) thwarts
  - × knowledgeable customers can thwart (encryption)
  - ✓ mass market likely to be naïve
  - × even naïve customers eventually notice cheaper identical service
  - × edge networks not naïve – will hide value from interior networks
- per session QoS
  - request to network for specific QoS reservation
  - ✓ network can infer broad task family (e.g. audio or video)
  - × edge networks will hide value from interior networks

BT

value-based charging
& competitive pressure

- instead of flapping around
  - why not just fix the price high?
- fine if you can get away with it





- if charge more than "cost plus normal profit"
  - competitors undercut
- demand exceeds supply
  - nearly half the time

BT

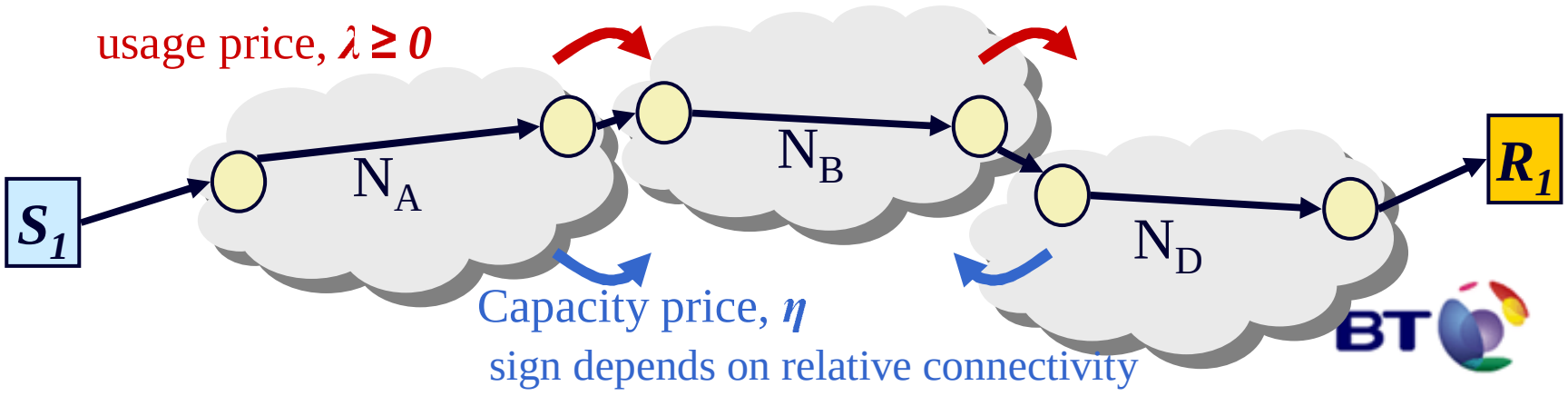
value-based capacity charge

- two-part tariff
  - capacity & usage (congestion or an approximation to it)
- capacity charge encourages stickiness
  - to switch providers based on usage price
    you must hold multiple subscriptions
  - the higher the capacity charges
    the less subscriptions you can afford
- competition reduces capacity subscription element
  - usage (congestion) charges offset marginal cost of capacity
  - if try to maintain high capacity charges, competitors will undercut
  - reduces relative contribution of capacity charge
  - increases multi-homing, reduces stickiness

BT

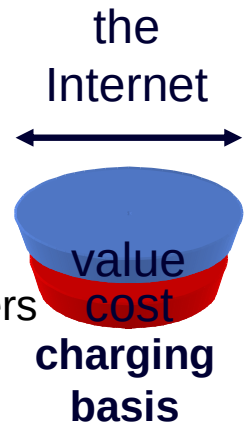
two part tariffs

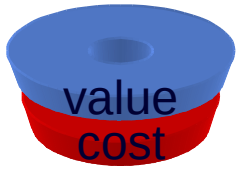
- sending domain pays $C = \eta X + \lambda Q$ to receiving domain per accounting period
- X is capacity @ price η
- Q is QoS/usage-related (volume, congestion) @ price λ
- both prices relatively fixed
- usage related price $\lambda \geq 0$ (safe against 'denial of funds')

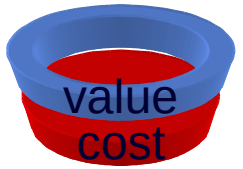


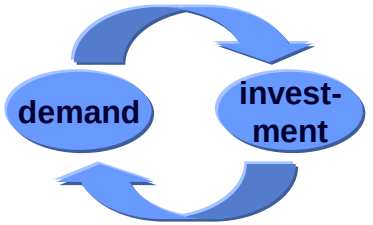
market structure evolution

- layered market
  - value-based charging over cost-based substrate
  - cost-based is most generic, proof against strategising machines
  - value-based charging layered over it, priced for human customers
- edge networks
  - will prevent backbones inferring value of traffic
- competition
  - most intense in middle – low cost to switch providers
  - will drive prices to floor of “cost plus normal profit”
  - hole devoid of value-based charging will grow from middle
- virtuous circle?
  - edge networks can still extract value
  - edge networks most need investment





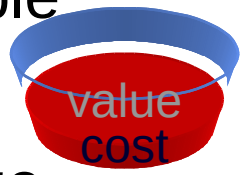




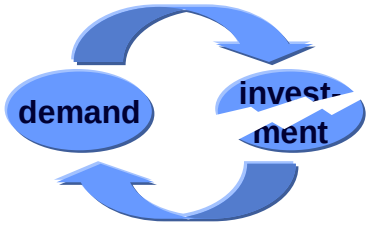
BT

googly: watch your backs

- commoditisation can move fast, once it's feasible
  - QoS commoditisation is now feasible
- the Web commoditised data transport for a huge number of applications
  - TCP just quietly gets on with allocating capacity between them all



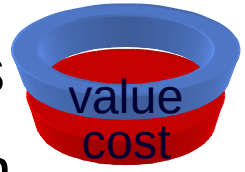
- we have the benefit of hindsight
  - but fierce competition could ruin your whole day



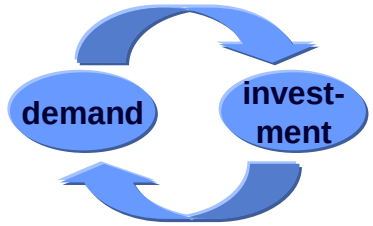
BT

summary

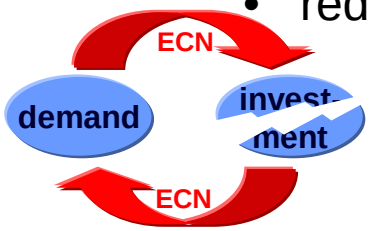
- congestion pricing is a hammer for every nail
- hole in value-based charging will grow outwards



- congestion (cost-based) pricing layered beneath
  - coordinates cost sharing between the networks
  - (spare slides: how broadband access operators share value over this hole)
- edge networks
  - need most investment and can capture most value



- googly: market might commoditise fast
  - feasible with latest congestion control advances
  - reducing role for subscription charging: more multi-homing



BT

references

- [Shenker95] Scott Shenker. Fundamental design issues for the future Internet. IEEE Journal on Selected Areas in Communications, 13(7):1176–1188, 1995
- [Hands02] David Hands (Ed.). M3I user experiment results. Deliverable 15 Pt2, M3I Eu Vth Framework Project IST-1999-11429, URL: http://www.m3i.org/private/, February 2002. (Partner access only)
- [Kelly98] Frank P. Kelly, Aman K. Maulloo, and David K. H. Tan. Rate control for communication networks: shadow prices, proportional fairness and stability. Journal of the Operational Research Society, 49(3):237–252, 1998
- [Gibbens99] Richard J. Gibbens and Frank P. Kelly, Resource pricing and the evolution of congestion control, Automatica 35 (12) pp. 1969—1985, December 1999 (lighter version of [Kelly98])
- [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) (to appear)

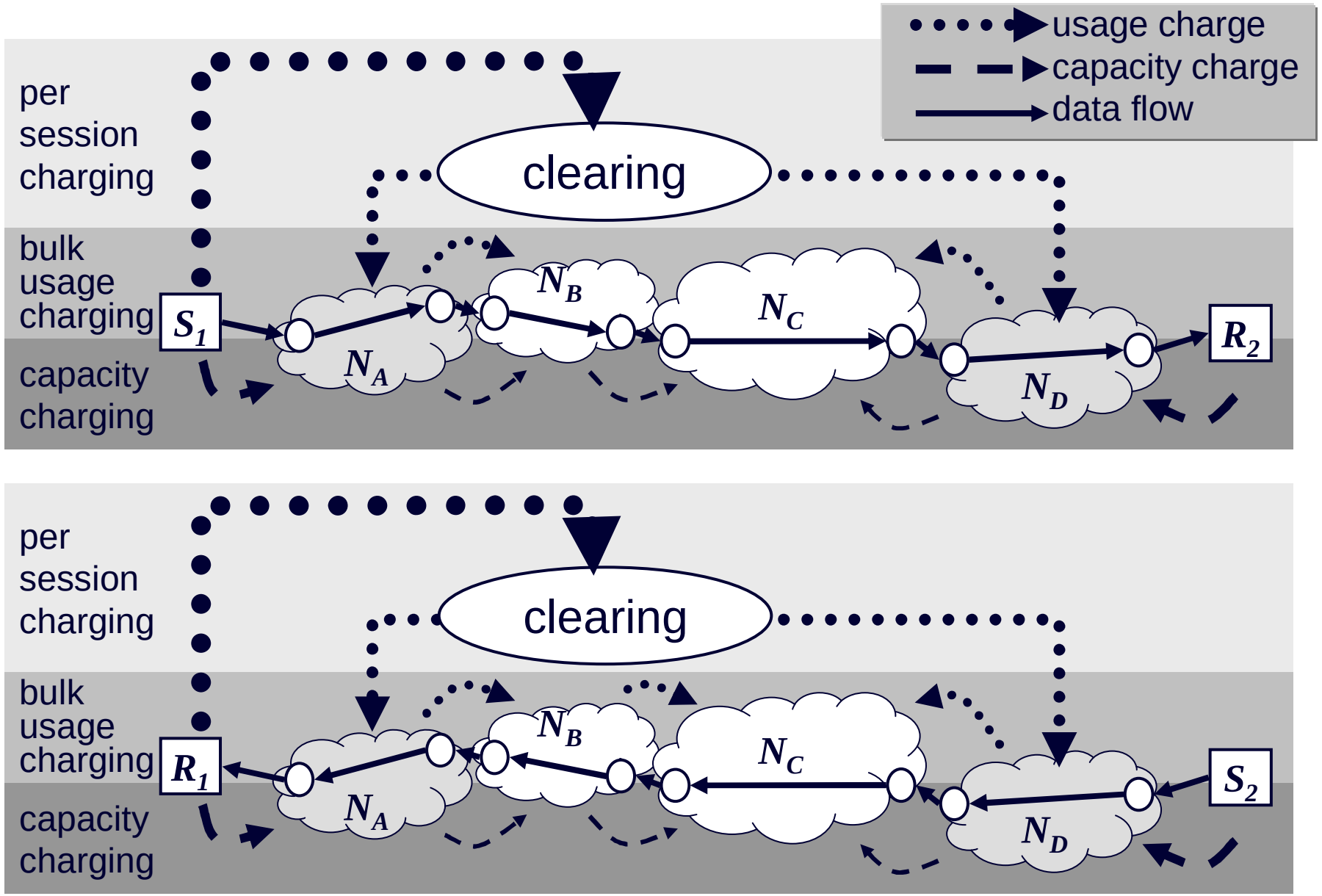
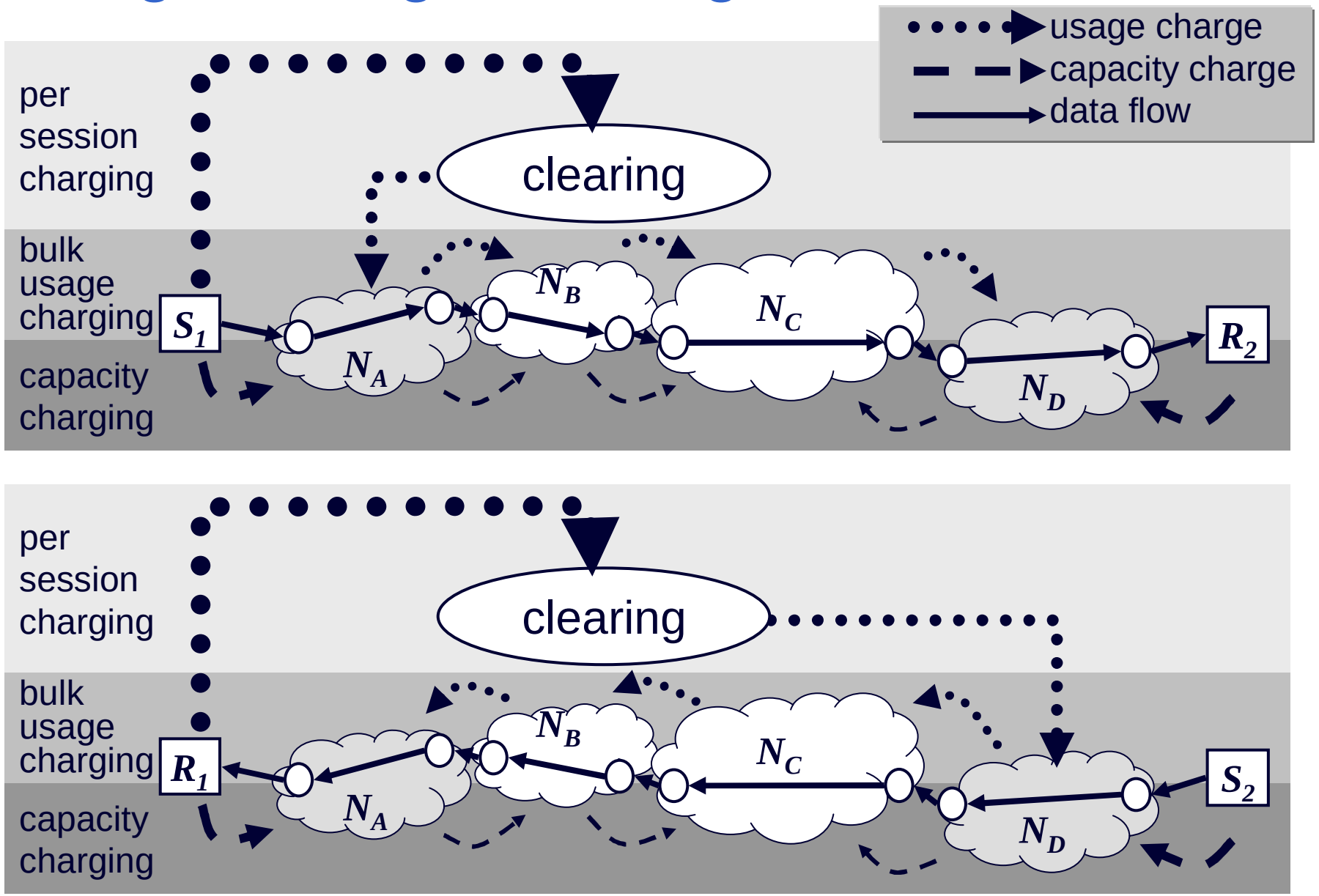
- Market Managed Multi-service Internet consortium <http://www.m3i.org/>

BT

bridging the value-hole

spare slides

edge-to-edge clearing - value-based



edge-to-edge clearing – cost-based