practical microeconomics and Internet resource sharing protocols

Bob Briscoe

Sep 2009

“The gap between theory and practice is greater in practice than in theory” Steve Crocker

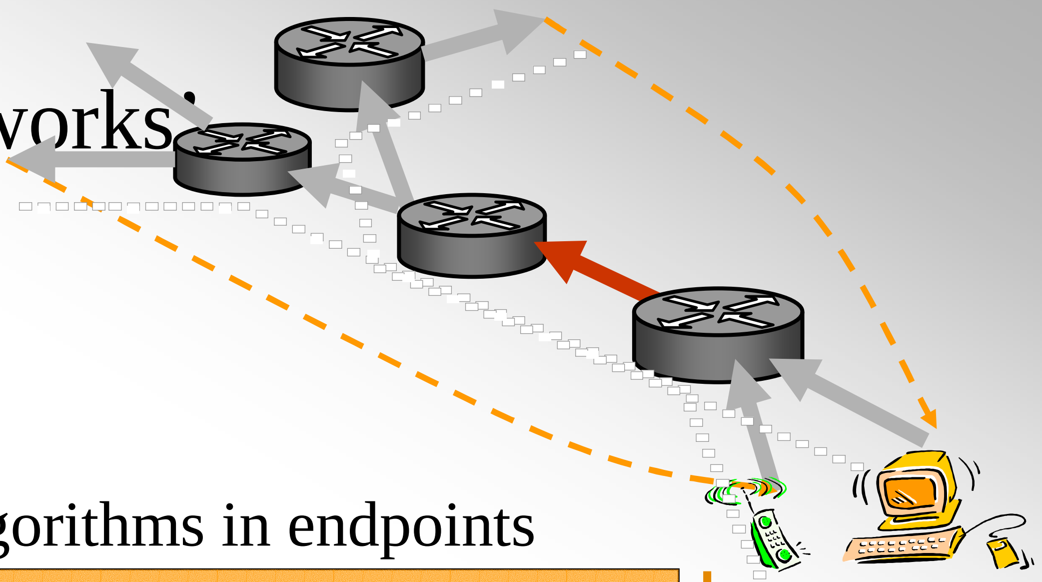
how to share a packet network?

- anyone can use any capacity anywhere on the Internet, as much as they like, without asking
  - fantastic ideal
  - but when freedoms collide, what share do you get?
- freedom with accountability
- decades of misunderstanding to undo

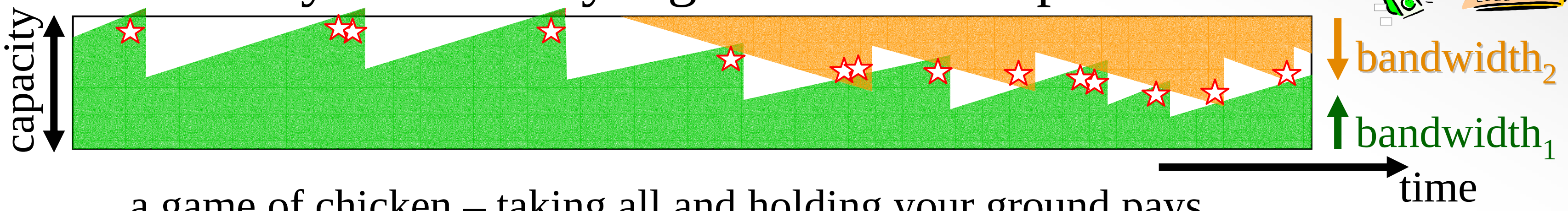
- need solutions that cater for
  - self-interest & malice
    - of users and of providers
    - without killing cooperation
  - evolvability
    - of new rate dynamics from apps
    - of new business models
  - viability of supply chain
  - simplicity (e.g. one-way datagrams)

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

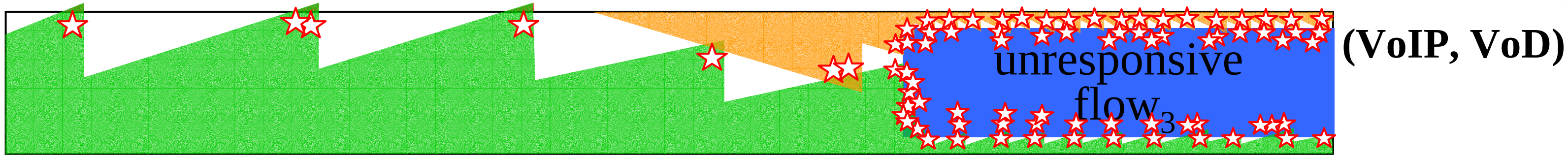
how Internet sharing 'works' TCP-friendliness



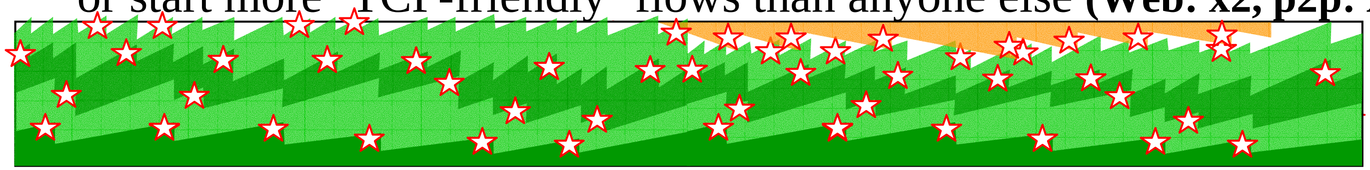
- endemic congestion
- voluntarily restraint by algorithms in endpoints



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



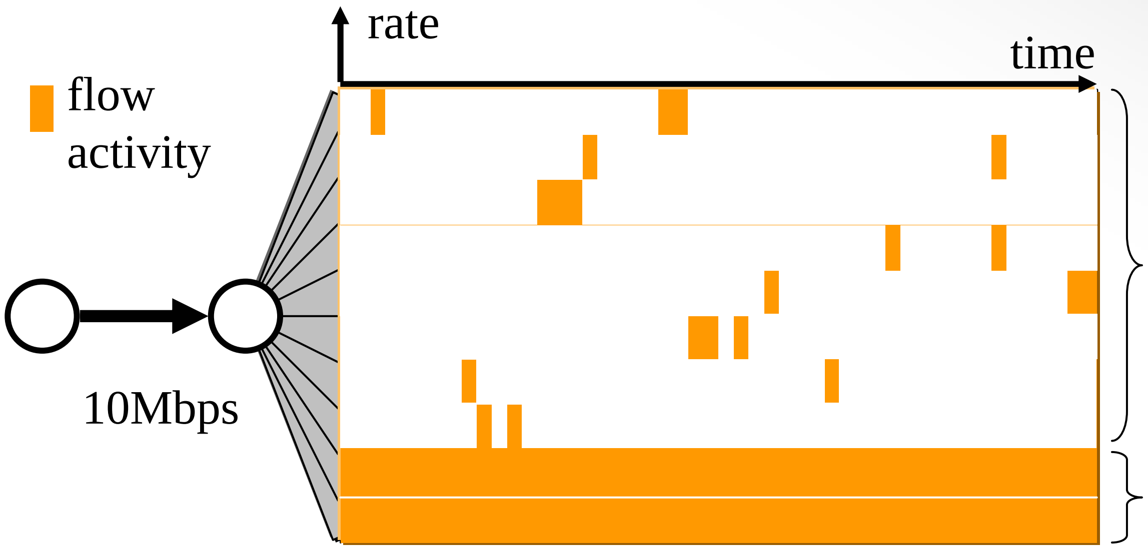
or start more 'TCP-friendly' flows than anyone else **(Web: x2, p2p: x5-100)**



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

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

TCP's broken resource sharing
base example: different activity factors



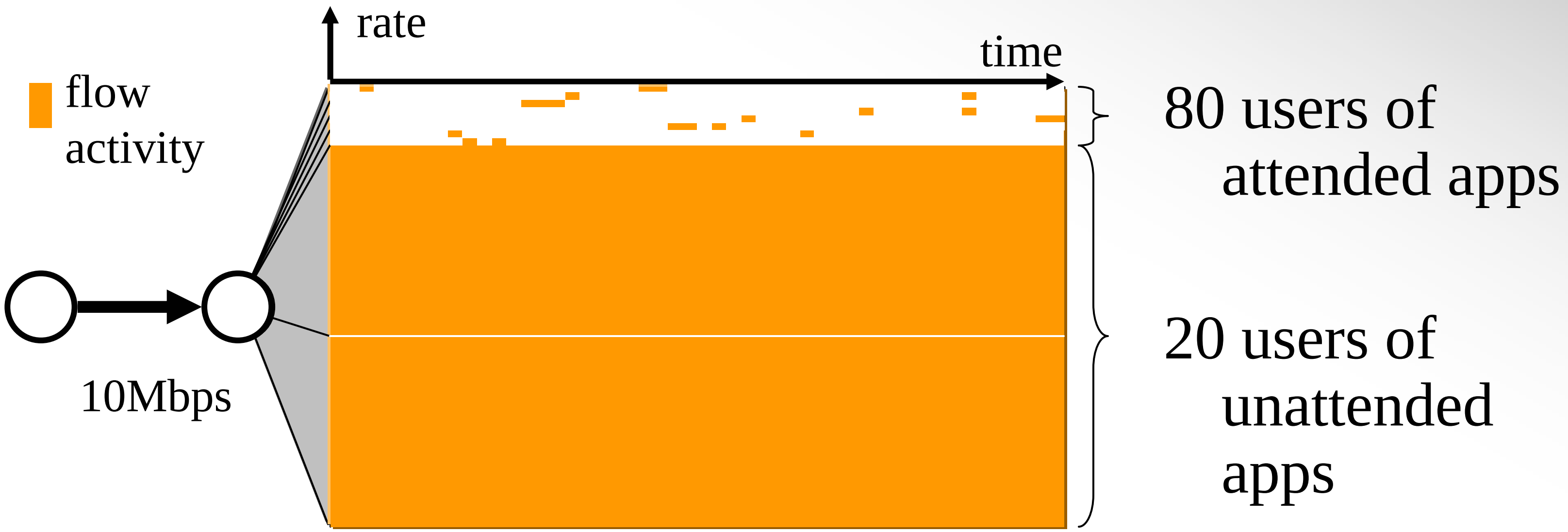
2Mbps access each

80 users of attended apps

20 users of unattended

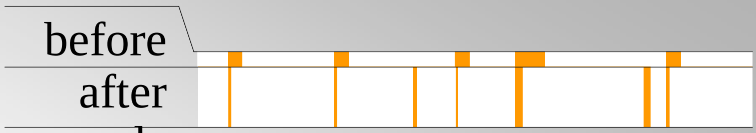
usage type	no. of users	activity factor	ave.simul flows /user	TCP bit rate /user	vol/day (16hr) /user	traffic intensity /user
attended	80	5%	=	417kbps	150MB	21kbps
unattended	20	100%	=	417kbps	3000MB	417kbps
				x1	x20	x20

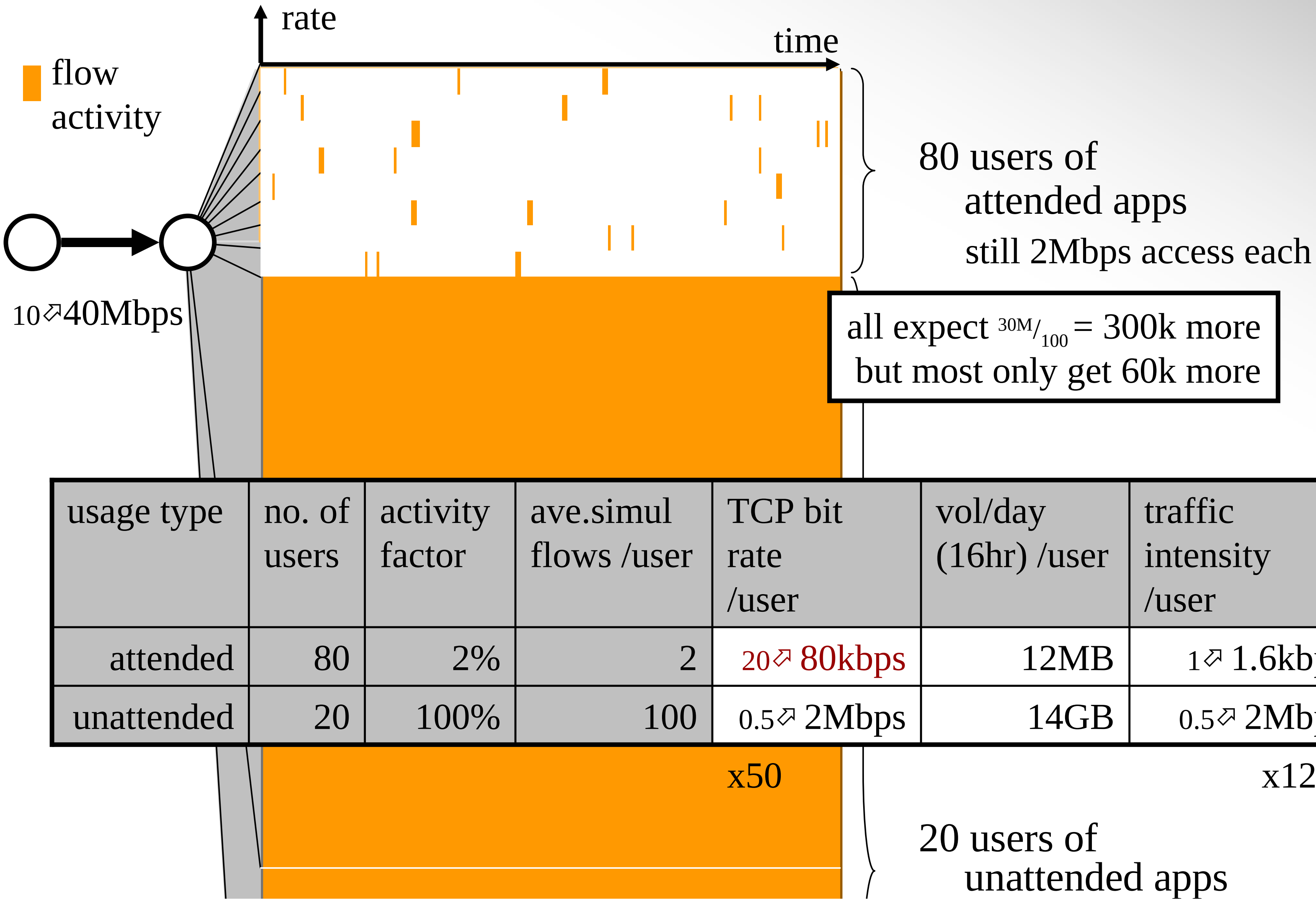
TCP's broken resource sharing
compounding activity factor & multiple flows



usage type	no. of users	activity factor	ave.simul flows /user	TCP bit rate /user	vol/day (16hr) /user	traffic intensity /user
attended	80	5%	2	20kbps	7.1MB	1kbps
unattended	20	100%	50	500kbps	3.6GB	500kbps
				x25	x500	x500

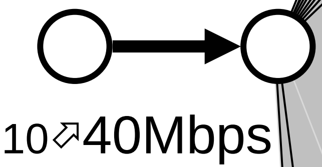
most users hardly benefit from bottleneck upgrade





usage type	no. of users	activity factor	ave.simul flows /user	TCP bit rate /user	vol/day (16hr) /user	traffic intensity /user
attended	80	2%	2	20⬀ 80kbps	12MB	1⬀ 1.6kbps
unattended	20	100%	100	0.5⬀ 2Mbps	14GB	0.5⬀ 2Mbps

consequence #1
higher investment risk



- recall

all expect $^{30M}/_{100}$ = 300k more
but most only get 60k more

- but ISP needs everyone to pay for 300k more
- if most users unhappy with ISP A's upgrade
- they will drift to ISP B who doesn't invest
- competitive ISPs will stop investing...

- if those willing to spend more can't get more, they won't spend more
- then we all share a smaller Internet

consequence #2
trend towards bulk enforcement

- as access rates increase
  - attended apps leave access unused more of the time
  - anyone might as well fill the rest of their own access capacity
- operator choices:
  a) either continue to provision sufficiently excessive shared capacity
  b) or enforce usage limits

---

see joint industry/academia (MIT) white paper “Broadband Incentives” [BBincent06]

consequence #3
networks making choices for users

- characterisation as two user communities over-simplistic
  - heavy *users* mix heavy and light *usage*
- ISP sees two prioritisation choices

a) bulk: network throttles all a heavy user's traffic indiscriminately
  - should encourage the user to self-throttle least valued traffic
  - ***but*** many users have neither the software nor the expertise

b) selective: network *infers* what the user would do
  - using deep packet inspection (DPI) and/or addresses to identify apps

- even if DPI intentions honourable
  - confusable with attempts to discriminate against certain apps
  - user's priorities are task-specific, not app-specific
  - customers understandably get upset when ISP guesses wrongly

ISP's homespun alternatives
have silently overridden TCP

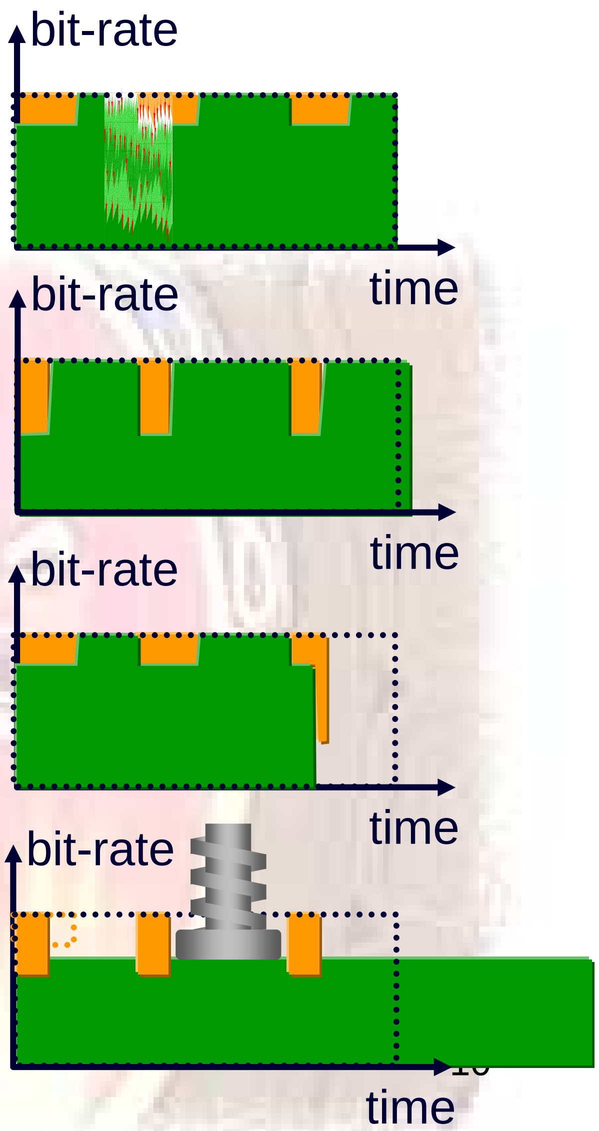
who is the fairest of them all?

1. equal bottleneck flow rates
   (TCP, XCP, RCP)?

2. access rate shared between active users,
   but weighted by fee (WFQ)?

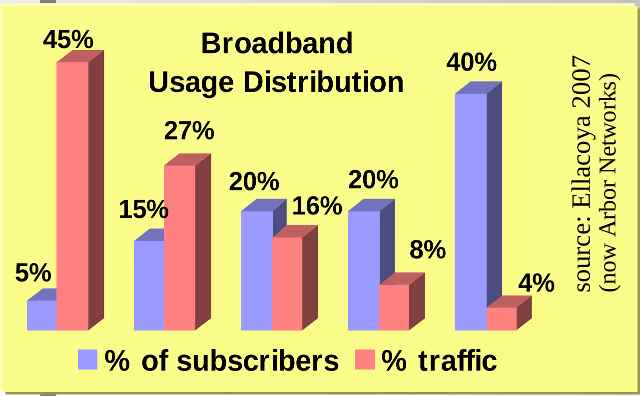
3. volume caps
   tiered by fee?

4. heaviest applications of heaviest users
   throttled at peak times by deep packet inspection (DPI)?

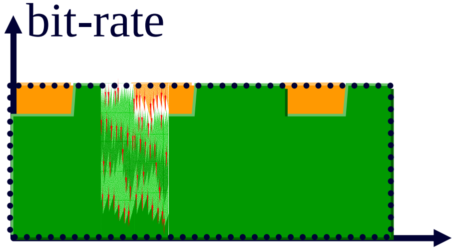


none of the above

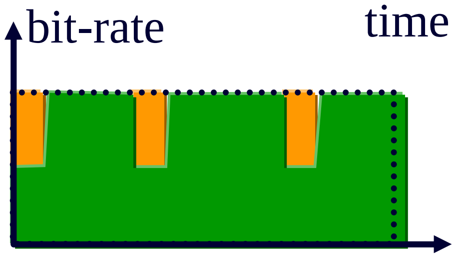
harness end-system flexibility



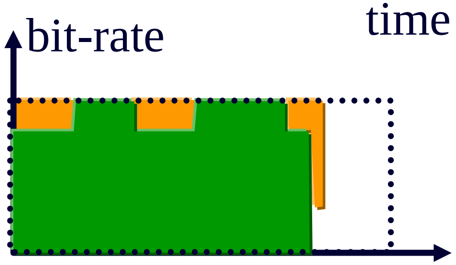
1. TCP
   XCP
   RCP



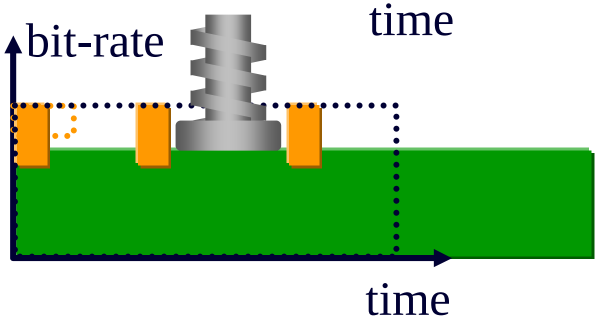
2. (weighted) fair queuing



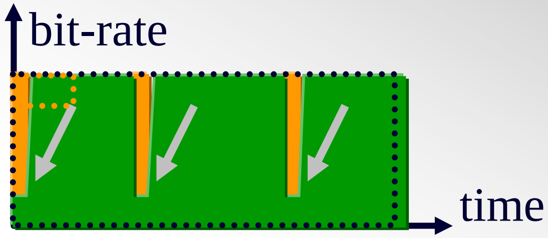
3. volume caps

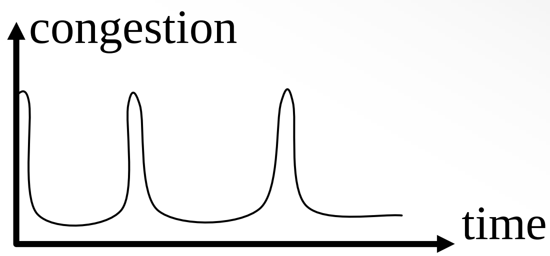


4. deep packet inspection (DPI)



weighted TCP sharing



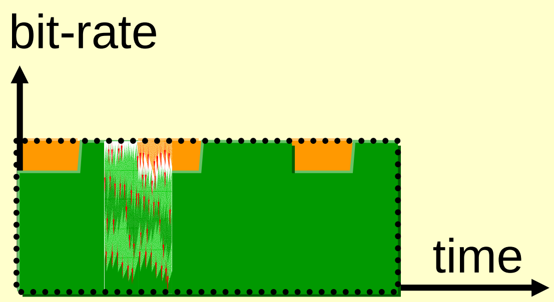


- light usage can go much faster
- hardly affects completion time of heavy usage

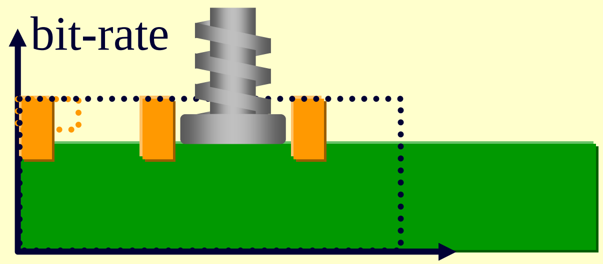
NOTE: weighted sharing doesn't imply differentiated network service

- just weighted aggressiveness of end-system's rate response to congestion, e.g. [LEDBAT]

two arbitrary approaches fighting

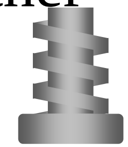


'flow-rate equality'

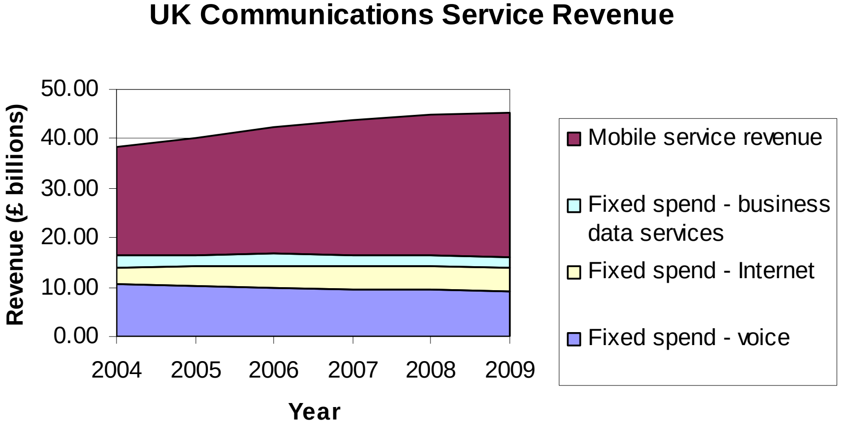


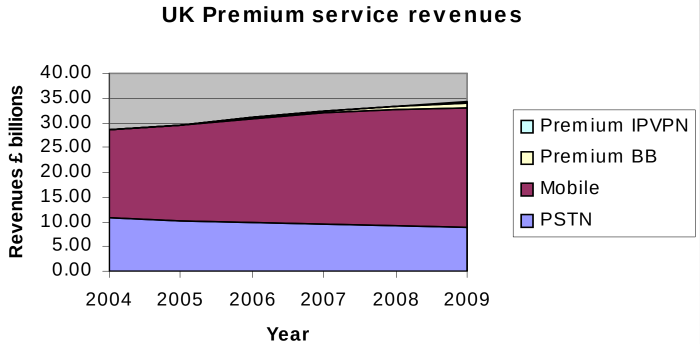
throttling heavy volume usage

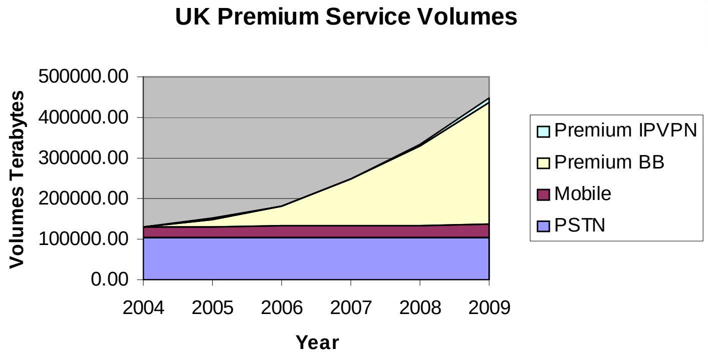
degree of freedom	the Internet way (TCP) ‘flow rate equality’	operators (& users) ‘volume accounting’
multiple flows	✗	✓
activity factor	✗	✓
congestion variation	✓	✗
application control	✓	✗

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

very large sums involved
very large distortions involved







- definition of 'premium'
  - services *requiring* better than normal QoS (latency or b/w)
  - not necessarily *using* network QoS mechanisms (e.g. VoIP)

Sources: Analysys Research (2005) and S Rudkin, BT internal report (Oct 2005)

in which fields of knowledge should we look for solutions?

- philosophy
- economics
  - microeconomics
  - political economy
  - industrial organisation
- engineering
  - data networking
  - control theory
- computer science
  - information theory
- mathematics

11	religion
10	politics
9	legal
8	commercial
7	application
4	transport
3	network
2	link
1	physical

11 layer OSI stack ☺

philosophy
fairness / justice

- 350 BCE Aristotle distinguished:
  - distributive justice
    - is the overall distribution of resources fair? (centralised)
  - commutative (rectifactory) justice
    - is each redistributive transaction fair? (distributed)
    - if voluntary, yes, by definition
- proposed approach
  - microeconomics for globally distributed resource sharing
    - in the process, we must sort out correct metrics, incentives, etc
    - invent technology to mitigate failings of market mechanisms
  - groups can override market allocations amongst themselves
    - e.g. country, university, multinational business, consortium, NATO, club, Internet café, ISP

organisation of lecture

- the problem: how to share a packet network?
- in theory – use a market mechanism
- in practice – failings of market mechanisms
- technical fixes for the failings of markets?
- fallacies
- specifics

terminological exactitude

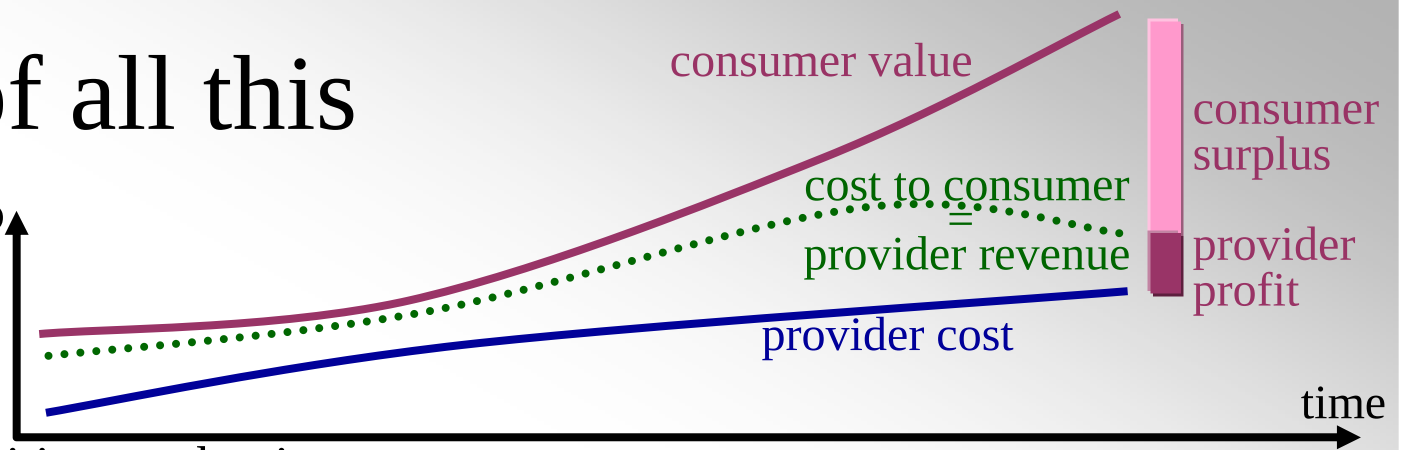
- tariff
  - e.g. where V is volume [B] & t is time [month]
  - charge, $G = aV + bt + c$
- price
  - undefined unless wrt to something

    price wrt V $$p_V = \frac{\partial G}{\partial V} = a$$
- cost
  - undefined unless state to whom

    cost to consumer = charge levied by producer

    $\neq$ cost to producer

the point of all this economics



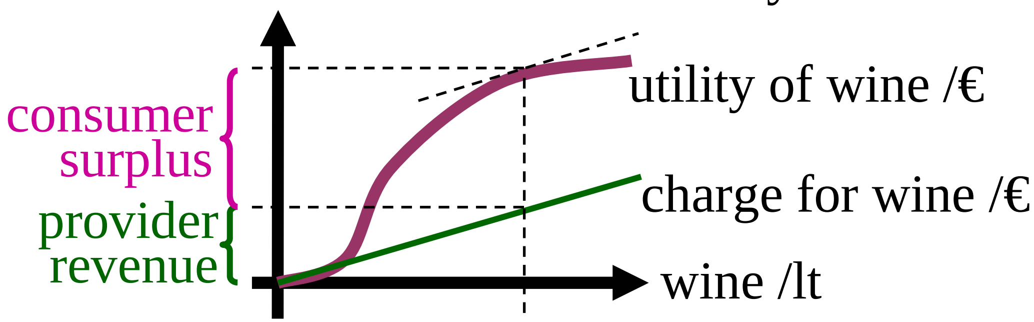
over time a competitive market is meant to

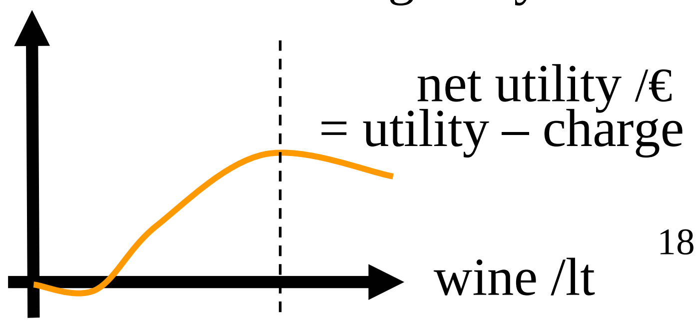
a) ensure resources get allocated most to those willing to pay most for them

b) provide the funds to invest where supply is short of demand

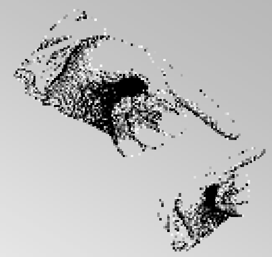
c) reduce the cost of what consumers buy to the cost of providing it

- a) & b) operate within a market (e.g. Internet usage) and between markets (e.g. Internet vs. travel vs. wine)
- c) squeezes profits and grows consumer surplus
- a) should ensure everyone optimises their utility (happiness) given their limited wealth and that they must cover the cost of the things they want





the invisible hand of the market often needs a helping hand



- if you don’t want the rich to pay more & get more (a), don’t use a market
  - but market is simplest distributed way to optimise utility (a) & match supply to demand (b)
  - so governments typically prefer to give pensioners €10/month to spend freely, rather than a €10 Internet voucher
- a poorly competitive market won’t squeeze profits (c) well
  - governments often prefer to regulate an uncompetitive market, e.g. by capping prices close to the cost to the provider (as if c)
  - then utility optimisation (a) & matching supply to demand (b) can still proceed automagically

cost vs value in Internet architecture

- user value per bit varies over ~10^{10} (video vs SMS)
- not role of network architecture to reveal user value
- revealing cost (to consumer) *is* role of architecture
  - lowest cost routes (without traffic)
  - traffic cost
- then net can make user accountable for cost of actions
- user decides if private value of act is worth the cost

- harder as cost to consumer approaches true cost
  - dynamic cost of traffic congestion
  - allocating traffic costs between networks

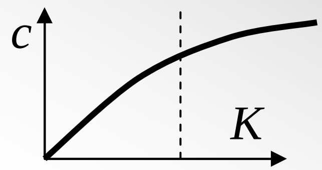
relaxing the economics

- don't confuse *being able* to hold users accountable for true costs with a desire that every ISP *should*
- as long as ISPs can put in constraints, they can also relax them
- as market gets more competitive, ISPs need to *be able* to tend towards true cost
- architecture must *be able* to allow tussle between profit & consumer surplus to play out
- reference: “Tussle in Cyberspace” [Clark05]

usage vs subscription prices

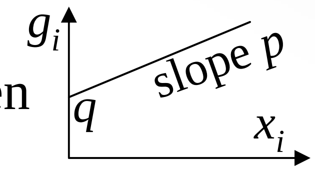
Pricing Congestible Network Resources [MacKieVarian95]

- assume competitive providers buy capacity K [b/s] at
  cost rate [€/s] of *c(K)*

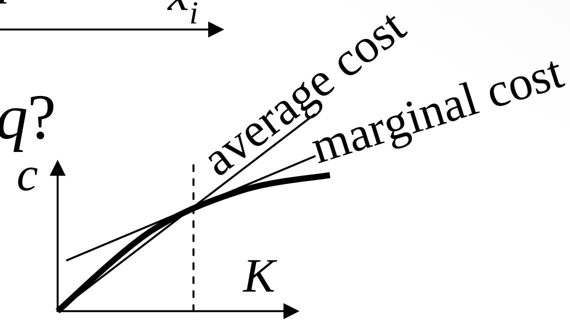


- assume they offer a dual tariff to customer *i*
  - subscription price q [€/s]
  - usage price p [€/b] for usage x_i [b/s], then

  charge rate [€/s], $g_i = q + px_i$



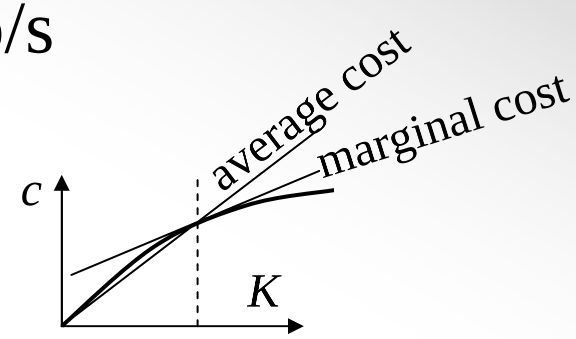
- what's the most competitive choice of p & q?
- $$\frac{\text{usage revenue}}{\text{capacity cost}} = \frac{1}{e}$$ where e is elasticity of scale
  - if charge less for usage and more for subscription, quality will be worse than competitors
  - if charge more for usage and less for subscription, utilisation will be poorer than competitors



$$e = \frac{\text{average cost}}{\text{marginal cost}} = \frac{c(K)}{Kc'(K)}$$

for example

- if a 10Gb/s link costs €1000
- and it costs €67 to upgrade to 11Gb/s
  - average cost = €100 per Gb/s
  - marginal cost ~ €67 per Gb/s



$$e = \frac{\text{average cost}}{\text{marginal cost}} = \frac{3}{2}$$

$$\therefore \frac{\text{usage revenue}}{\text{capacity cost}} = \frac{1}{e} = \frac{2}{3} \qquad \frac{\text{subscription revenue}}{\text{capacity cost}} = \frac{1}{3}$$

- ie usage revenue covers marginal cost
  subscription revenue covers the rest

typology of goods

free-riding problems

prevent non-contributors benefiting?

		excludable	non-excludable
use by A prevents simultaneous use by B?	rivalrous (‘use-up-able’)	private goods food, clothing, toys, furniture, cars	common goods fish, hunting game, water
	non-rivalrous (irreducible)	club goods satellite television	public goods national defense, free-to-air TV, air, published info

tragedy of the commons problems

- shared Internet bandwidth: a common good
  - ‘use-up-able’ and non-excludable (if ‘pure’ Internet)
  - also instantly perishable (the extreme of non-durable goods)
- free-riding typically reduces the incentive to supply
- common goods tend to be under-supplied *and* over-consumed
  - network congestion = too much traffic meets too little capacity
- public (e.g. Wikipedia) easier than common goods for creating a sharing economy

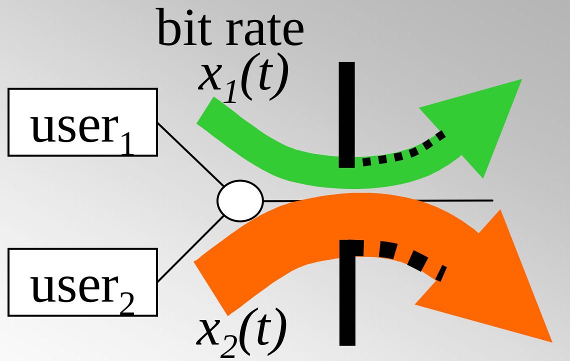
externalities

- an externality occurs where the actions of one agent directly affect the environment of another agent
- reference: Varian, *Microeconomic Analysis*
- positive externalities
  - others use software compatible with yours
  - others connect to your network (‘network effects’)
- negative externalities
  - pollution, road congestion, network congestion

aligning incentives
in the presence of externalities

- a market doesn't 'work' if externalities present
  - when deciding how much gas to use, *homo economicus* only takes account of the cost to him, not to others
- solution: internalise the externality
  - increase his charge by the cost to others of his actions
  - he will use less gas – the correct amount to optimise everyone's utility (a) and match supply to demand (b)

dual view of
congestion harm metric



loss (marking) fraction
$p_j(t)$

$$p(t) \equiv \frac{excess\_load(t)^+}{offered\_load(t)}$$

A. what each user i got,
weighted by congestion at the time
   - bit rate [bs^{-1}] weighted by congestion []

B. the bits each user contributed to excess load
   - congestion weighted by each user's bit-rate

$\boxed{p_j(t)x_r(t)}$

- a precise instantaneous measure of harm during dynamics
that easily integrates over time and sums over the resources j on the route r of a flow and over allthe flows of a user i, where $p_r = \Sigma_{j\in r} p_j$

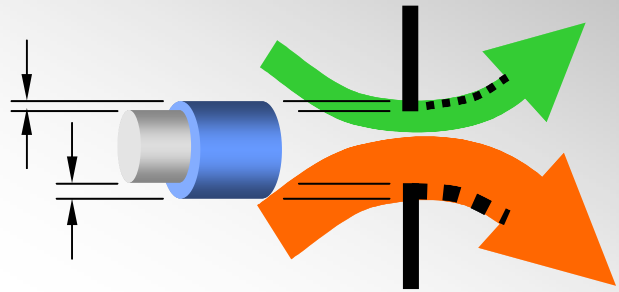
$\boxed{v_i \equiv \Sigma_{r\in i} \int p_r(t)x_r(t)\, dt}$

- termed **congestion-volume** [byte]
- result is easy to measure and compare per user
   - volume of bytes discarded or ECN marked

- intuition: compare with volume, $V_i \equiv \Sigma_{r\in i} \int x_r(t)\, dt$
which is bit rate over time summed over all a sender's flows
- network operators often count volume only over peak period
   - as if $p(t)=1$ during peak and $p(t)=0$ otherwise

dual demand & supply role of

congestion-volume metric



- a resource accountability metric
  1. of customers to ISPs (too much traffic)
  2. and ISPs to customers (too little capacity)

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

- competitive market matches 1 & 2

note: diagram is conceptual

congestion volume would be accumulated over time

capital cost of equipment would be depreciated over time

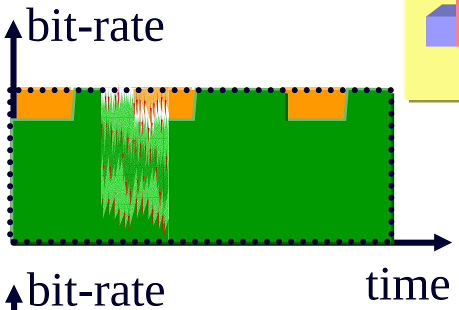
NOTE: congestion volume isn't an extra cost

- part of the flat charge we already pay
- we ***might*** see tiered pricing like this...

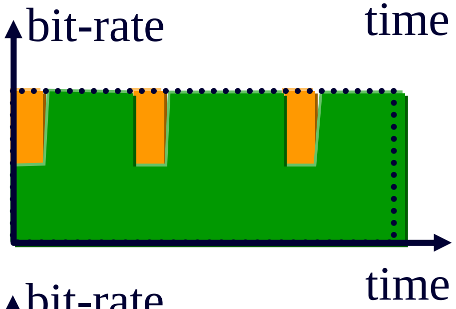
access link	**congestion volume allow'ce**	**charge**
100Mbps	50MB/month	€15/month
100Mbps	100MB/month	€20/month

congestion-volume

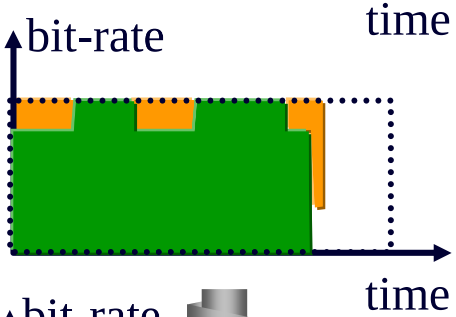
1. TCP
   XCP
   RCP



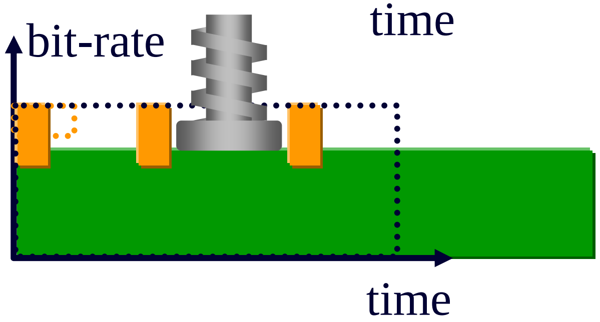
2. (weighted) fair queuing

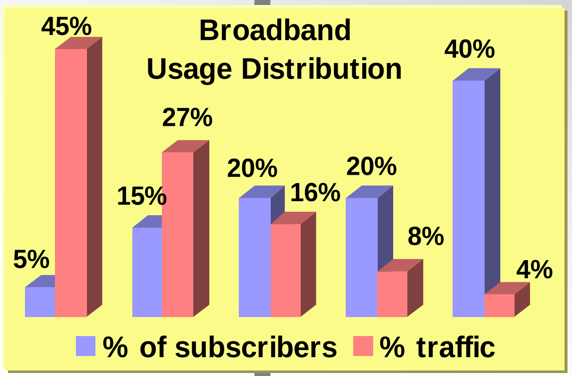


3. volume caps

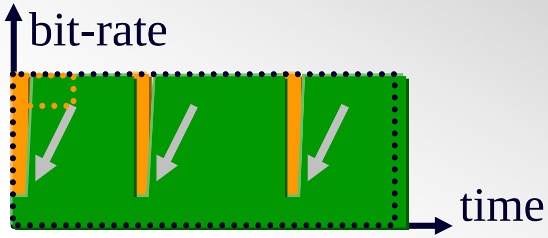


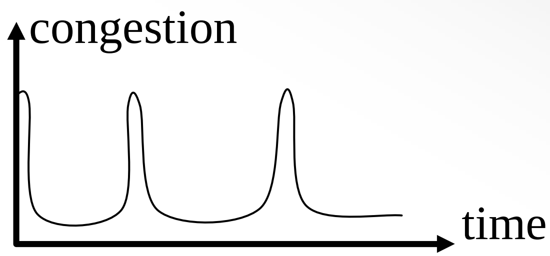
4. deep packet inspection (DPI)





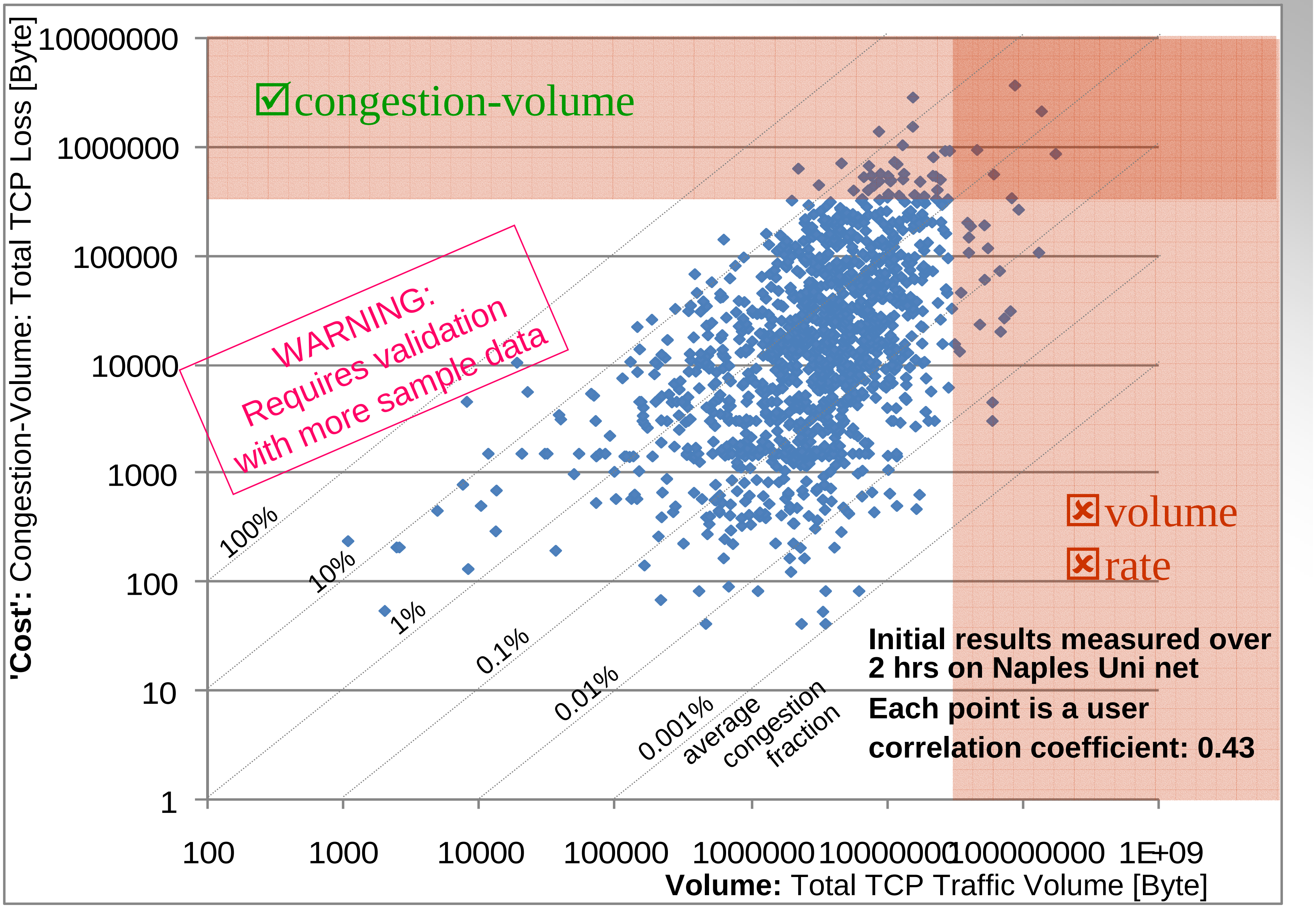
weighted
TCP
sharing





- takes into account all three factors

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



☑congestion-volume
☒volume
☒rate
WARNING:
Requires validation
with more sample data
Initial results measured over
2 hrs on Naples Uni net
Each point is a user
correlation coefficient: 0.43
100%
10%
1%
0.1%
0.01%
0.001%
average
congestion
fraction
'Cost': Congestion-Volume: Total TCP Loss [Byte]
Volume: Total TCP Traffic Volume [Byte]
10000000
1000000
100000
10000
1000
100
10
1
100
1000
10000
100000
1000000
10000000
100000000
1E+09

sneak preview: flat fee, best without effort QoS

if ingress net could see congestion...

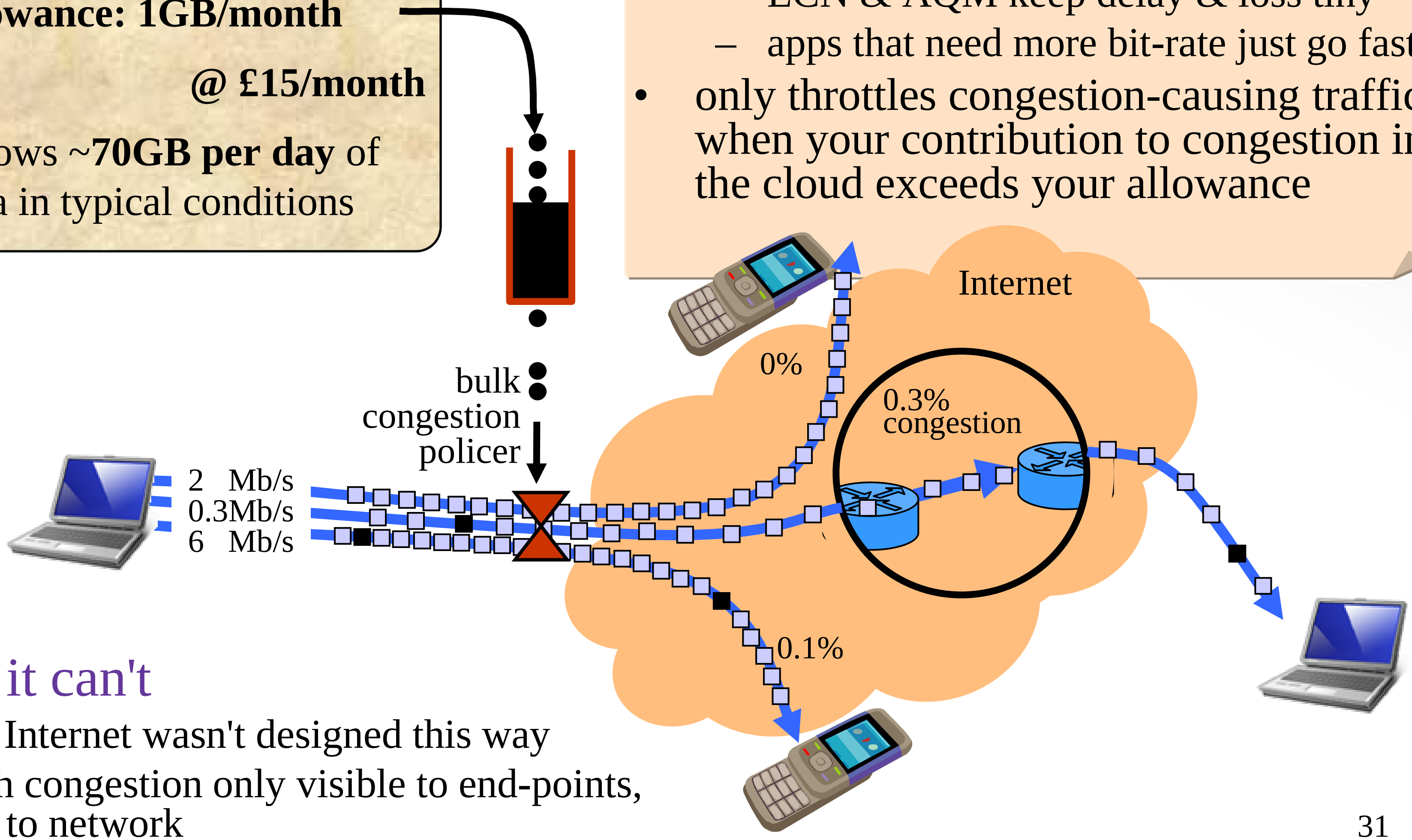
<u>Acceptable Use Policy</u>

'congestion-volume' allowance: 1GB/month

@ £15/month

Allows ~**70GB per day** of data in typical conditions

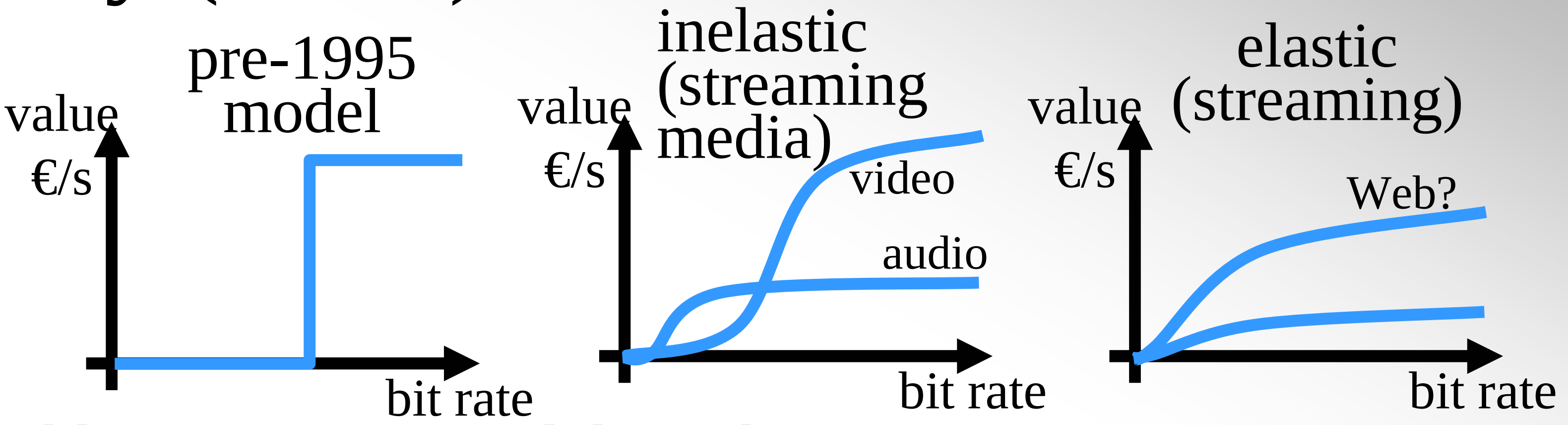
- simple invisible QoS mechanism
  - ECN & AQM keep delay & loss tiny
  - apps that need more bit-rate just go faster
- only throttles congestion-causing traffic when your contribution to congestion in the cloud exceeds your allowance



...but it can't

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

utility (value) wrt bit rate: curve families

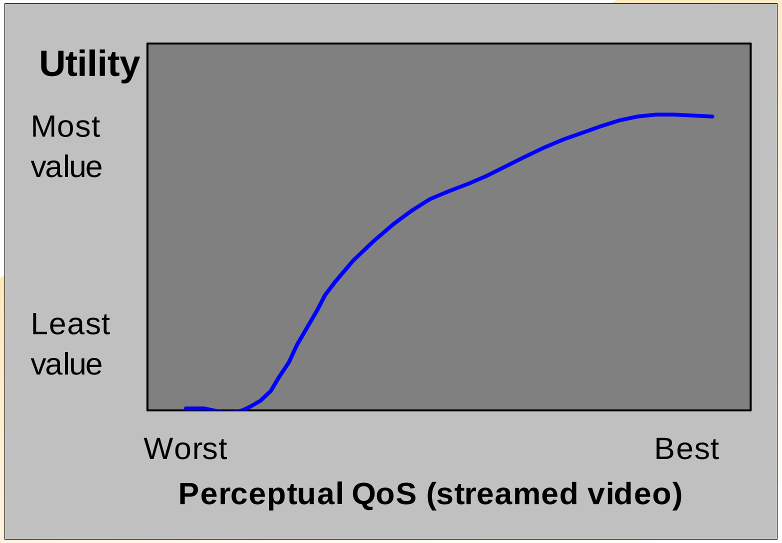


reasonable assumption used throughout economics:
utility satiates (concave):
• slope (marginal utility) monotonically decreases
• utility monotonically increases

theoretical
[Shenker95]

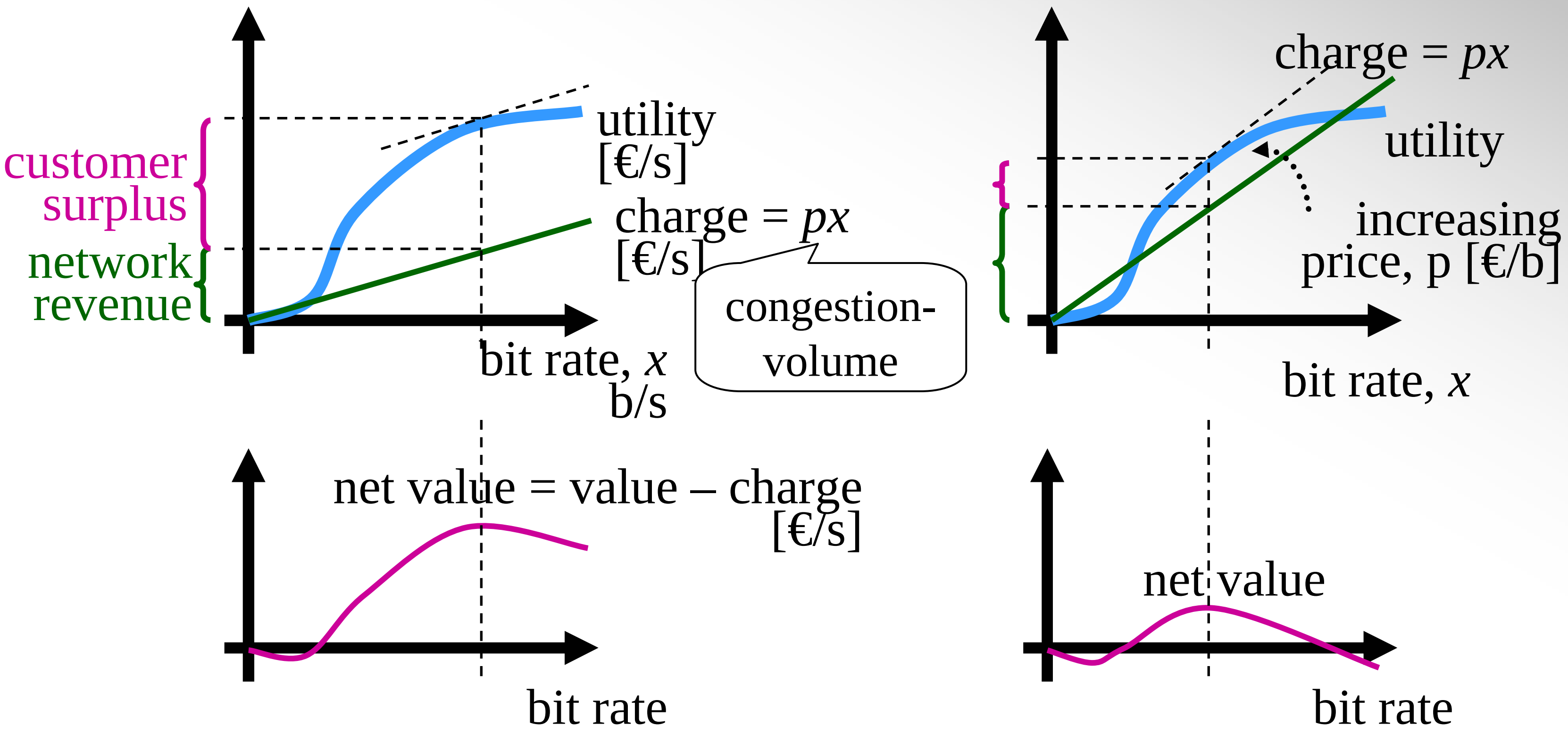
&

actual
value models



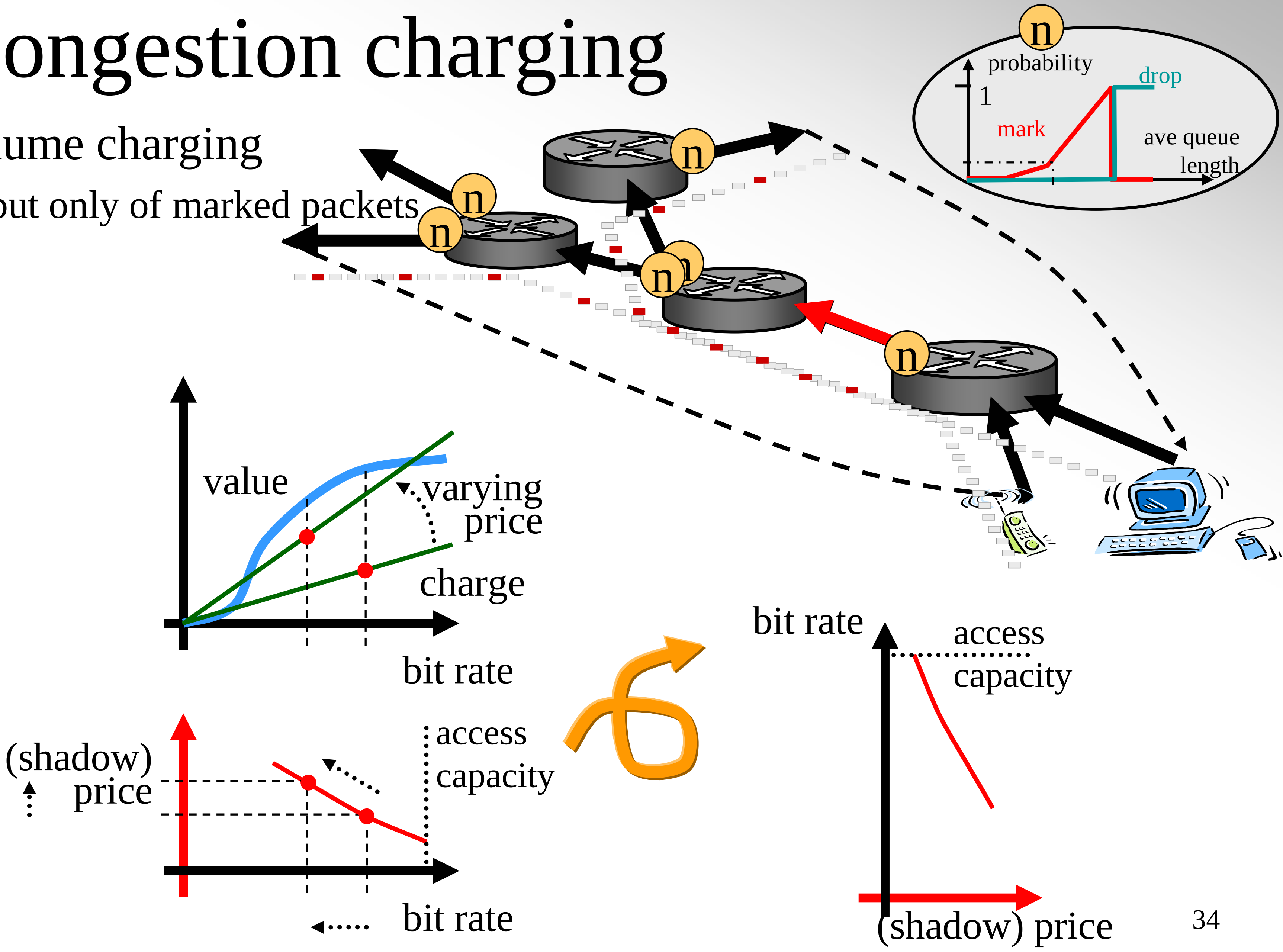
average of normalised curves from a set of experiments on paying customers [Hands02]

value – charge: consumer's optimisation

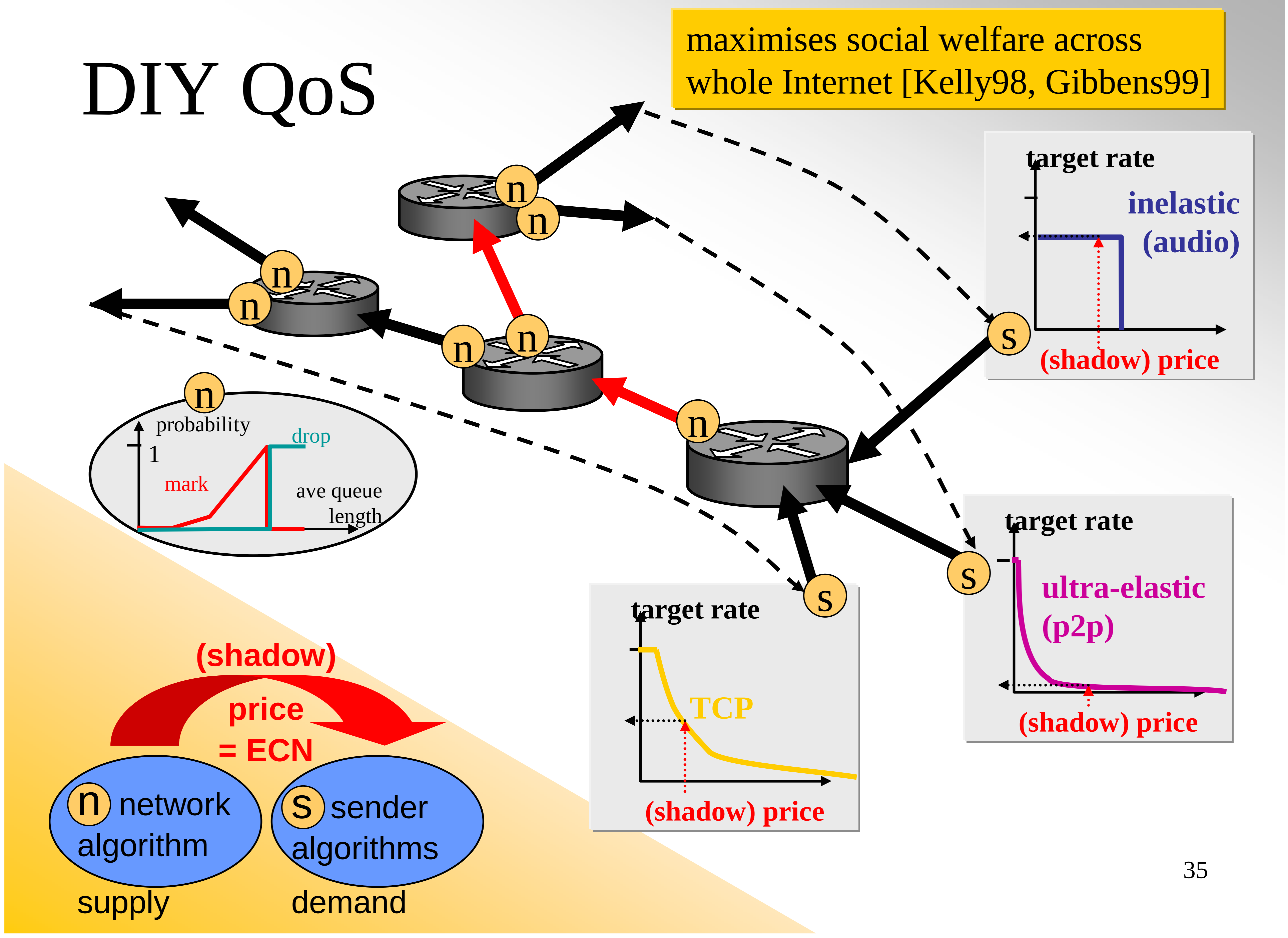


congestion charging

- volume charging
  - but only of marked packets



DIY QoS

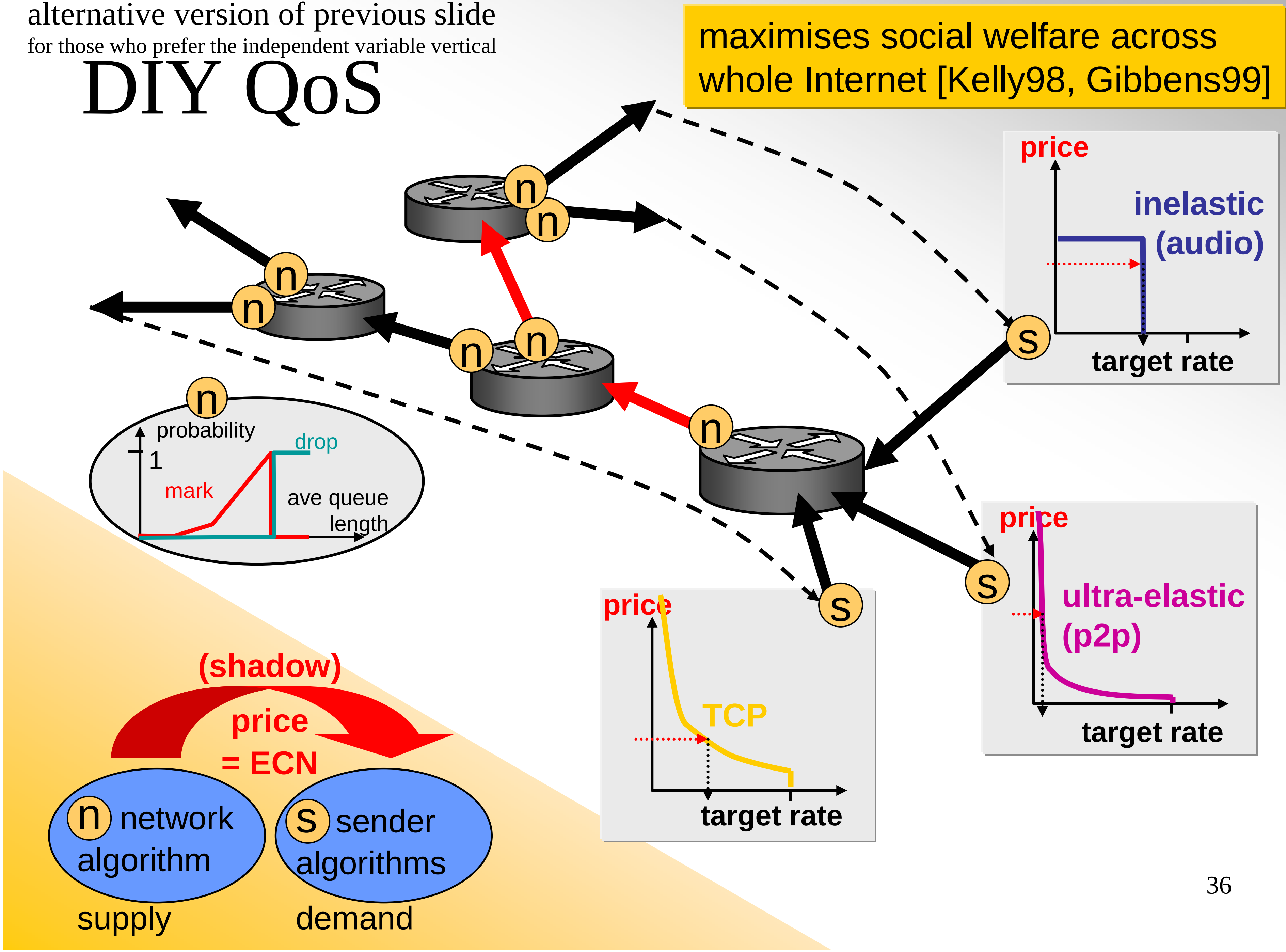


maximises social welfare across whole Internet [Kelly98, Gibbens99]
n
s
target rate
inelastic (audio)
(shadow) price
ultra-elastic (p2p)
TCP
probability
1
drop
mark
ave queue length
(shadow) price = ECN
network algorithm
sender algorithms
supply
demand

alternative version of previous slide

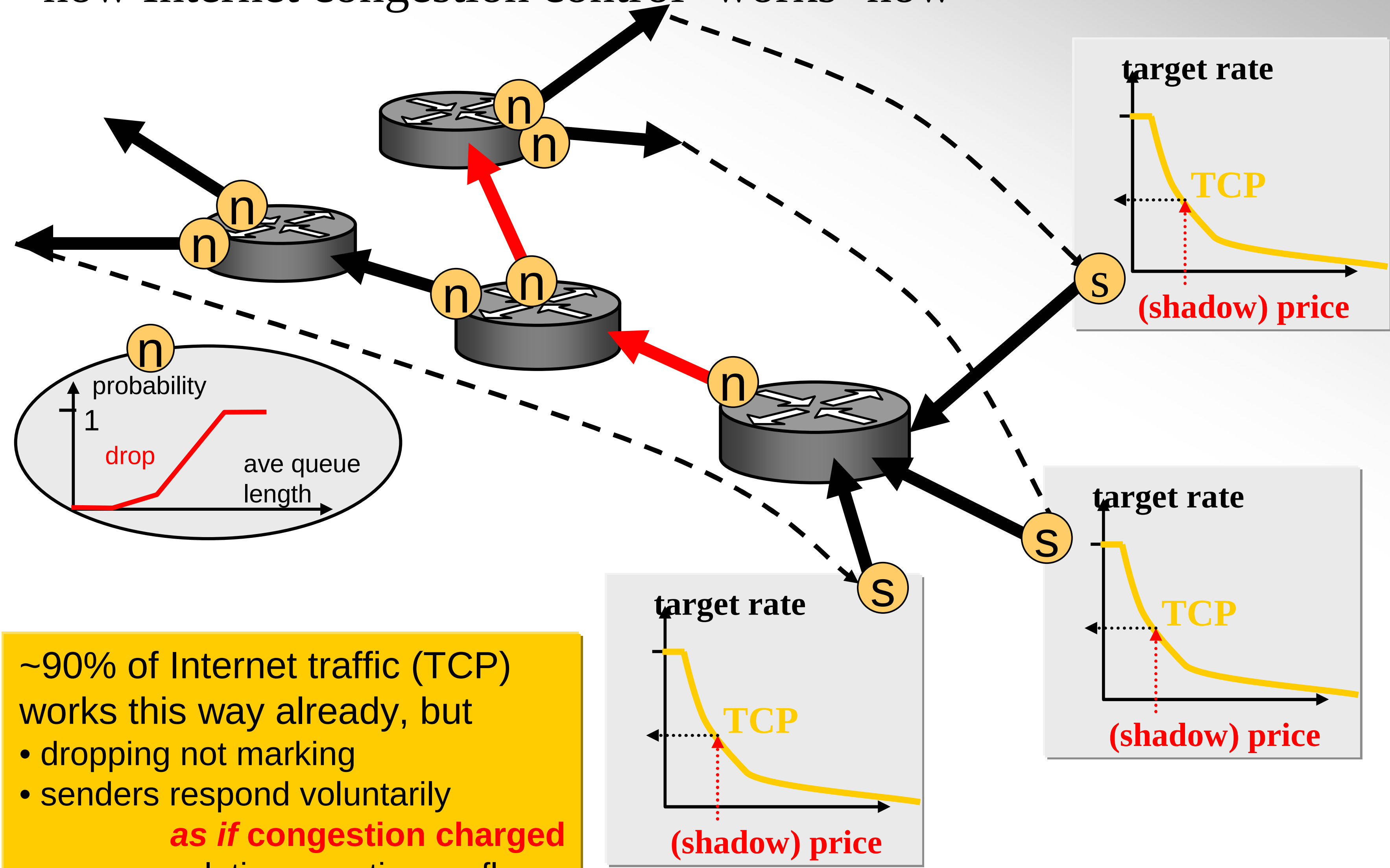
for those who prefer the independent variable vertical

DIY QoS



familiar?

– how Internet congestion control ‘works’ now



~90% of Internet traffic (TCP) works this way already, but

- dropping not marking
- senders respond voluntarily *as if* **congestion charged**
- no accumulation over time or flows
- every sender responds identically

microeconomics or ‘just’ optimisation?

- some use a ‘price’ purely as a name for a slack variable introduced in order to solve a distributed optimisation problem
- microeconomics solves a distributed optimisation problem
- some choose to connect a technical optimisation to the real economy through applying market prices
- others don’t
- for instance, today’s TCP uses losses as a ‘price’
    - although no-one applies market prices to TCP losses
    - there are numerous connections between TCP and the Internet market within which it exists
- an optimisation can choose to optimise anything
    - comparing an optimisation to real-world economics can hilite bad choices

reverse engineering TCP's economics
(rough model) as if derived from a utility curve

- window of W packets per round trip time T

$$\text{time}\ \frac{T}{W}\ \text{per ACK}$$

$$W \text{ increases by } \frac{1}{W} \text{ per ACK} \Rightarrow \frac{1}{T} \text{ per RTT}$$

$$W \text{ decreases by } \frac{W}{2} \text{ per NACK} \Rightarrow$$

$$\frac{Wp}{2} \text{ per ACK} \Rightarrow \frac{W^2 p}{2T} \text{ per RTT}$$

$$\text{hence } \frac{dW}{dt} = \frac{1}{T} - \frac{W^2 p}{2T}$$

$$\text{which gives steady state throughput } \frac{\overline{W}}{T} = \frac{1}{T}\sqrt{\frac{2}{p}}$$

reverse engineering TCP's economics
(rough model) as if derived from a utility curve

$$\text{steady state packet rate, } \bar{x} = \frac{\overline{W}}{T}$$

$$= \frac{1}{\mathrm{T}}\sqrt{\frac{2}{p}}$$

$$\frac{d\bar{x}}{dt} = \frac{1}{T}\frac{d\overline{W}}{dt}$$

$$= \frac{1}{\mathrm{T}^2} - \frac{\bar{x}^2 p}{2}$$

$$\text{TCP behaves as if user maximising net utility} = \text{utility} - \text{cost}$$

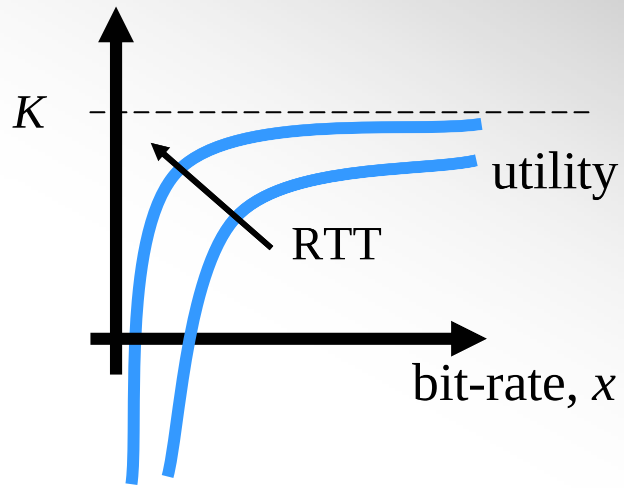
$$= U(\bar{x}) - p\bar{x}$$

$$\text{max of concave function is where derivative} = 0$$

$$U'(\bar{x}) = p$$

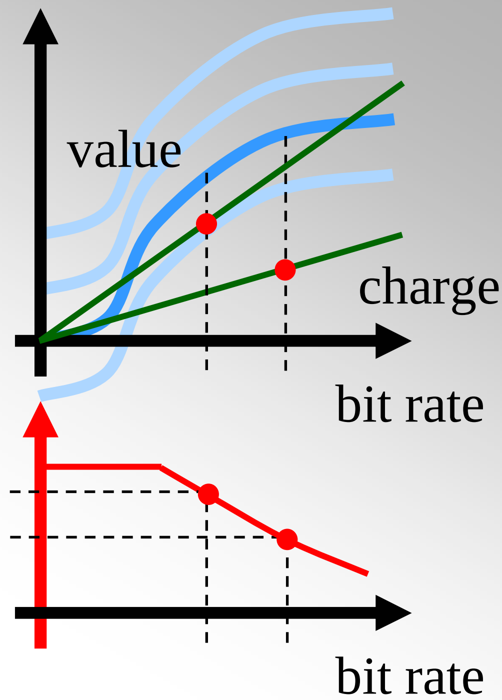
$$\text{in steady state} \qquad = \frac{2}{T^2\bar{x}^2}$$

$$\text{integrating, } U(\bar{x}) = K - \frac{2}{T^2\bar{x}}$$



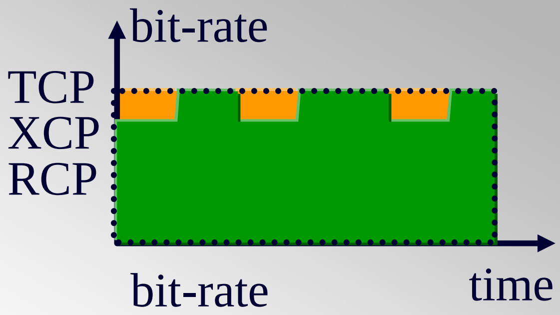
- TCP packet rate is more sensitive to RTT than bandwidth aside

utility: ordinal not cardinal

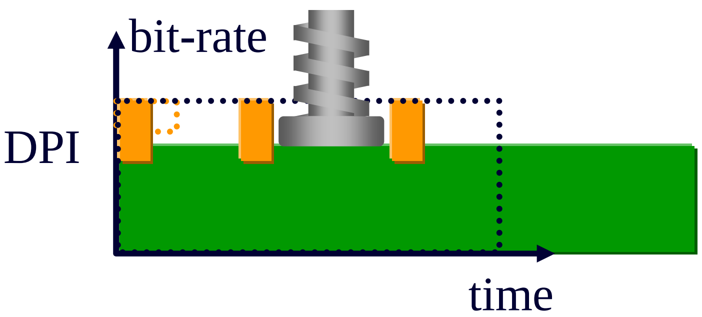


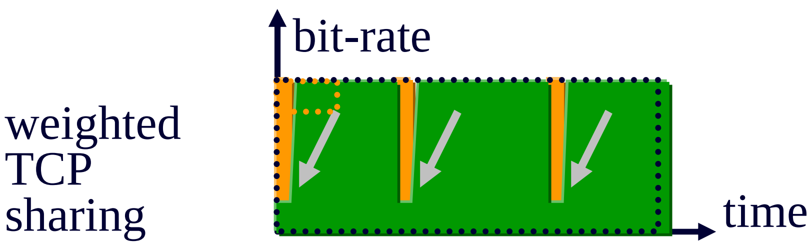
- utility itself never actually needed
- endpoint algo solely maps congestion to bit-rate
- no change if utility curve shifted up or down
- only slope (marginal utility) is ever used

‘good enough’ or optimal?



- optimisation can be oversold
  - in life ‘good enough’ is everywhere
  - history gets stuck down paths that end at good enough
  - to jolt onto better path higher effort than value gained
- but highly sub-optimal outcomes cause distortions
  - if architecture leads to extreme suboptimum (e.g. TCP)
  - economics will win another way (e.g. deep pkt inspection)
  - architecture that prevents tussle (optimisation) gets violated
  - result: a mess
- see “Tussle in Cyberspace” [Clark05]

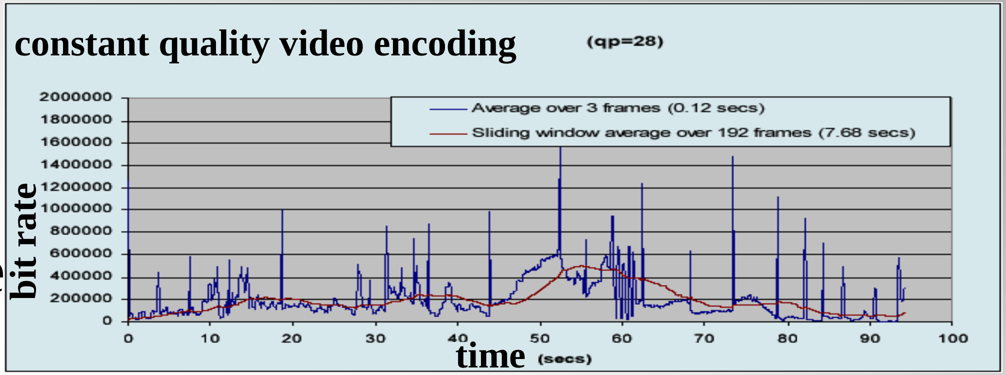




motivating congestion-volume

harnessing flexibility

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



- the idea that humans want to have 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

Constant Bit Rate **100%** Constant Quality **125%** Equitable Quality **216%**

sequences encoded at same average of 500kb/s

[Crabtree09]

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

market failures

the Internet suffers from them all!

market failures

the Internet suffers from them all!

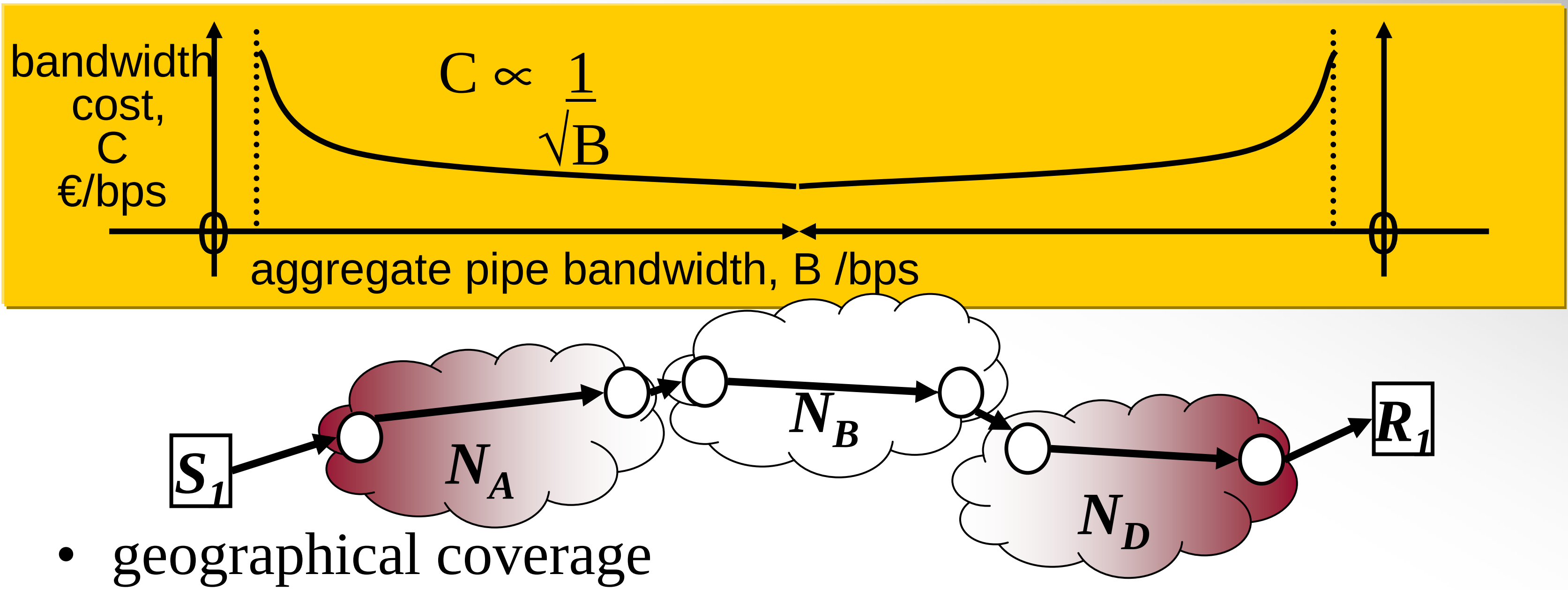
- externalities
  - (-) congestion
  - (+) network effects
- non-excludability
- market power
  - natural monopoly
- switching costs
- transaction costs
  - 2-sided market
  - termination monopoly
- information asymmetry

- the bit-rates people choose will be ‘wrong’
  a) global utility won’t be maximised
  b) supply won’t match demand
  c) profit won’t be squeezed
- technical fix(es)?
  - more helping hands for the invisible hand?

not too perfect, please!

- Internet can’t be isolated from the economy
- driving charges to cost and other benefits (a,b,c) can’t happen if market can’t function well for technical reasons, e.g.
    - true cost information only visible asymmetrically
    - high barriers to entry for new providers
    - high costs for customers to switch providers
- but, if Internet market is too ‘efficient’
    - investment will go to less ‘efficient’ markets i.e. with higher profitability

natural monopoly of access networks



- geographical coverage
  - two operators each build networks covering an area
  - if they each get half the customers spread evenly
  - costs nearly twice as much per customer
- solutions are primarily regulatory
  - a ‘layer 2 problem’ necessary to correct at L2
  - e.g. ‘local loop unbundling’
    - monopolist must lease out copper lines and equipment space in exchange
    - at regulated price and quality, incl. installation time, access to building, etc

switching costs
(switching in the economic sense)

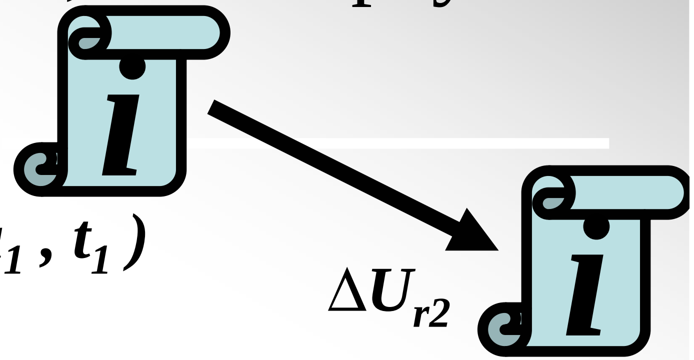
- consumer cost of switching between providers
  - identifier portability (e.g. email, IP address)
  - reconfiguration of stored profiles, data etc
  - contractual lock-in (e.g. 1yr min contract)
- regulatory remedies
- technical remedies:
  - simultaneous contracts
    - multihoming
    - multipath TCP

communications: a 2-sided market

the direction of value flow

who to make accountable for usage costs?
sending net (content)? rcving net (eyeballs)?

- if use principle of cost causation, sender pays
    - safe against denial of funds (DoF)



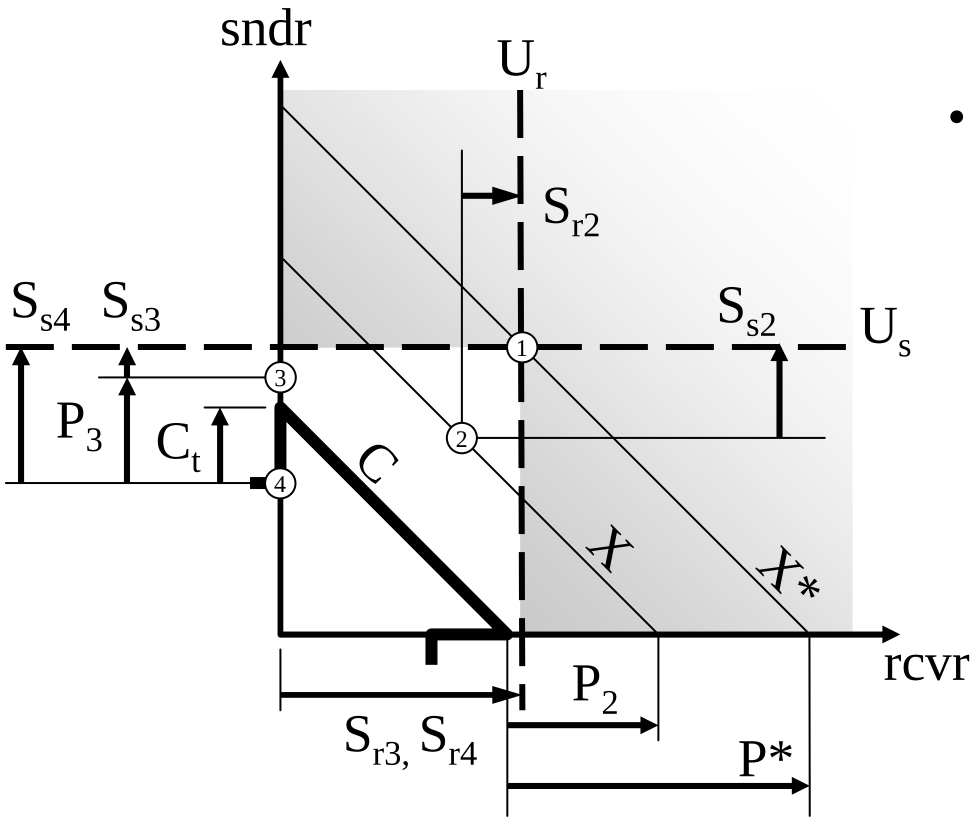
info value U = f(i, place , time)

xmt value $\Delta U_s = f(i, a_1 , t_2) - f(i, a_1 , t_1)$

- xmt value /leg = ΔU_j
- if sender pays and ΔU_s < cost, no transmission, even if $\Sigma\, \Delta U_j$ >> cost
- two-sided market (cf. credit card, night club, auction)

charge apportionment

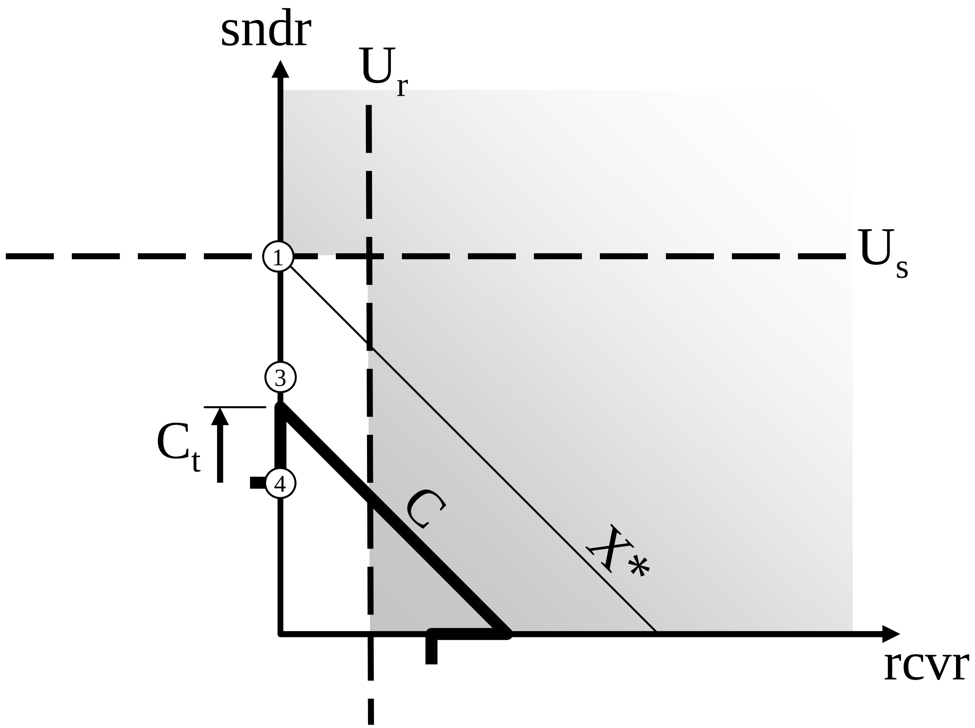
U : utility (to consumer)
$_{s/r}$: sender/receiver subscript
C : cost (to provider)
X : charge (paid by consumer)
S = U-X : consumer surplus
P = X-C : provider profit
C_t : apportionment transaction cost



- charge frontier represents apportionment choices
  - shaded region is provider's upper bound
- cost frontier is provider's lower bound
  - odd discontinuities due to apportionment transaction cost
- market evolution
  1) max provider profit, P*
  2) immature market
  3) commoditised market
  4) max consumer surplus, $S_{s4}+S_{r4}$
  - as market commoditises, need for retail apportionment reduces ('bill and keep' becomes predominant)

‘spam’ effect

U : utility (to consumer)
$_{s/r}$: sender/receiver subscript
C : cost (to provider)
X : charge (paid by consumer)



- rcvr’s utility is <u>expected</u> utility averaged over many messages
  - reduces considerably if some messages are low utility (irritatingly chatty friends or spam)
- if $U_r \leq C_t$, it’s never worth reapportioning some charge to the receiver

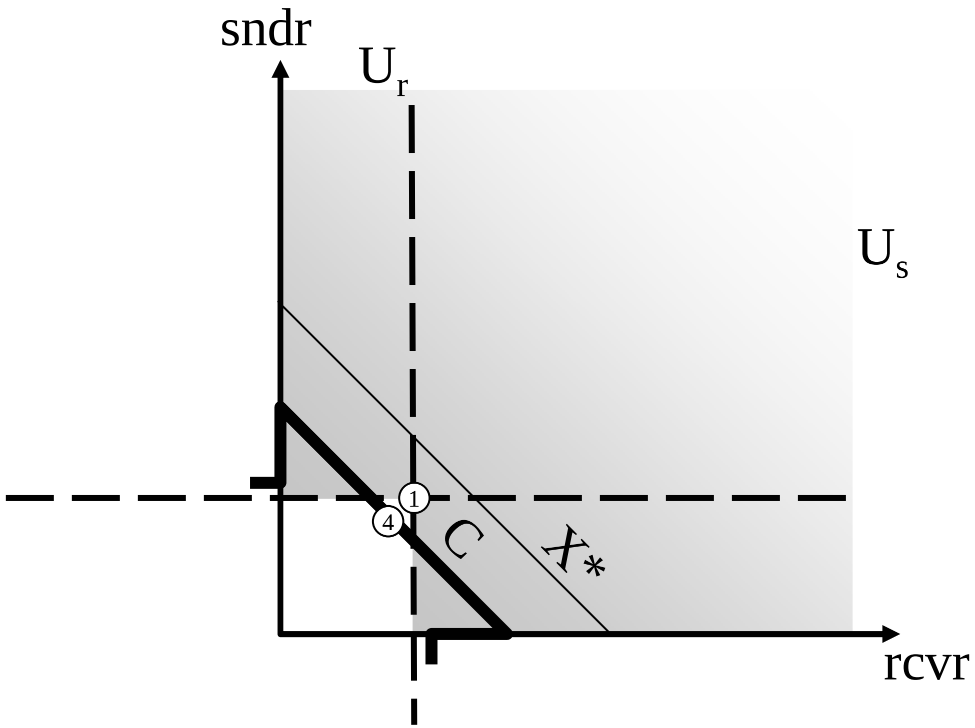
messages of marginal value

U : utility (to consumer)

$_{s/r}$: sender/receiver subscript

C : cost (to provider)

X : charge (paid by consumer)



- some messages only have sufficient value to leave profit after costs if charges are shared
- if these represent a large part of the market, charge reapportionment is the only way to grow market volume

termination monopoly
(the term originated in telephony)

- if sender-pays
- what if there is no alternative route?
  - e.g. the receiver is only attached to one ISP
- could be solved by regulation
- technical fix(es) possible
  - reciprocity?
  - receiver-pays at higher end-to-end layer (see later)

information asymmetry

competition & quality,
choice, routing & congestion

“The market for ‘lemons’:
Quality, uncertainty and market mechanisms” [Akerlof70]

- won Nobel Prize in Economics, 2001
- if seller not buyer knows which items are duds
  - buyer only willing to risk price of below average quality
  - seller makes sales for less than average quality
  - sellers unwilling to buy stock when will lose on average
  - market collapses
- Internet exhibits strange information asymmetry
  - buyer knows quality of goods but not seller
  - similar outcome [Briscoe08], see consequence #1 earlier

Internet congestion information asymmetry

- Internet architecture designed so that
  - transport layer detects congestion
  - hard for network to see congestion
    - gaps in transport sequence space
    - can be obfuscated by IPsec or multipath
    - if net intercepted feedback, transport could encrypt it
- ISP cannot limit costs it cannot see
  - can detect drop at its own equipment
  - perhaps collect to a control point using management messages
  - but not whole path congestion
- drop is a dodgy contractual metric
  - highly disputable
  - an absence – did it ever exist? Complex to prove [Argyraki07]

- ECN reveals congestion
- but only at receiver
  - problematic if net charges or limits by congestion received
  - receiver not in control of received packets
    - unwanted traffic, DoS, spam
    - wanted traffic, but unwanted high rate during congestion
  - receiving network not in control of received packets
    - cannot advertise or choose routes without rest-of-path congestion
    - networks cannot reward each for doing so [Constantiou01, Laskowski06]

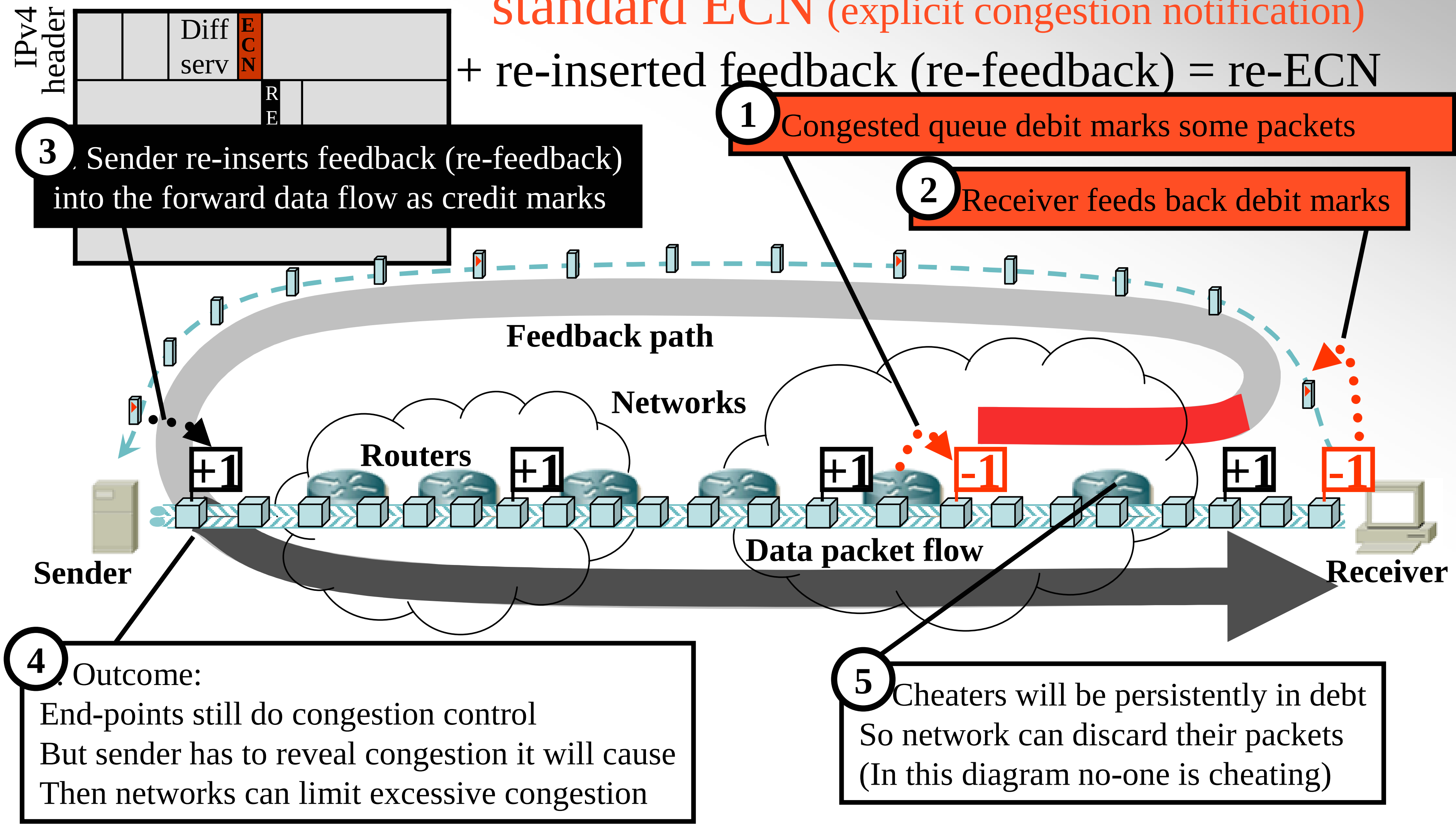
re-feedback (re-ECN)
re-inserted explicit congestion notification

a panacea?

one bit opens up the future

standard ECN (explicit congestion notification)

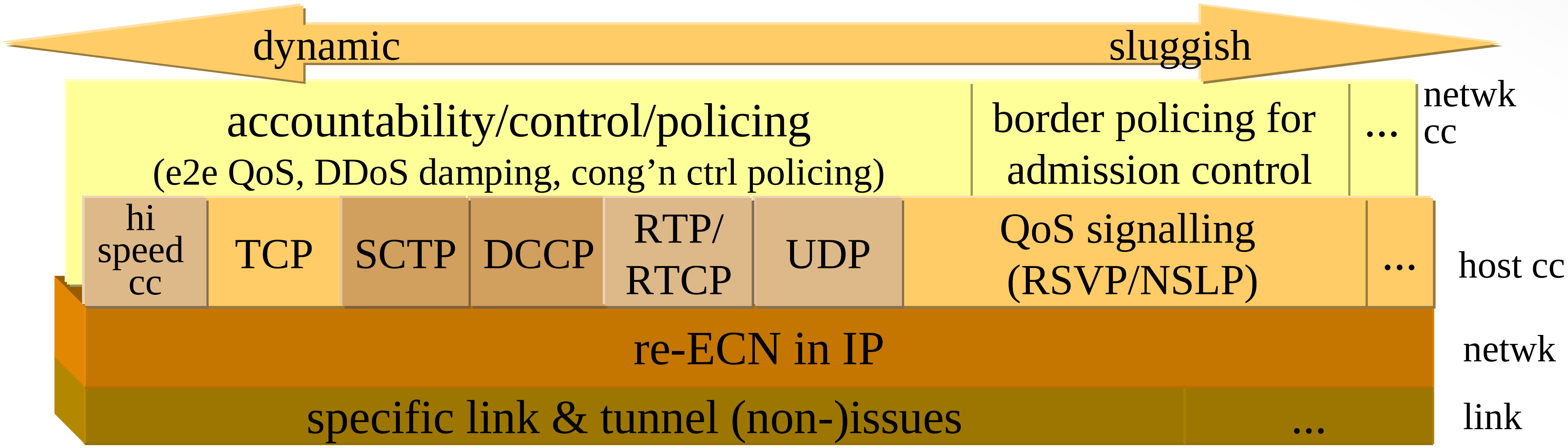
+ re-inserted feedback (re-feedback) = re-ECN



no changes required to IP data forwarding

standards agenda
re-ECN

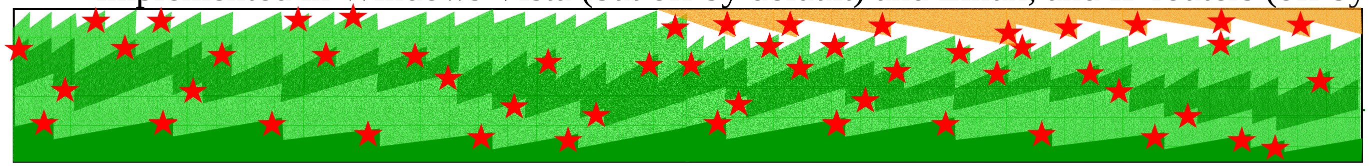
- layered beneath all transports
- for initial protocol specs see [re-ECN, re-PCN]
- implementations available (Linux & ns2) – just ask



problems using congestion in contracts

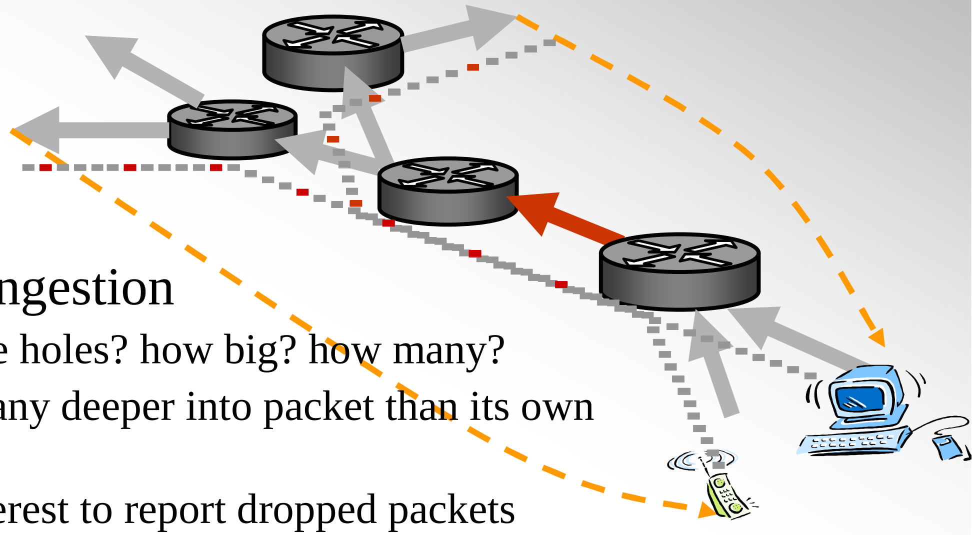
	1. loss	2. ECN	3. re-ECN
can't justify selling an impairment	☹	☺	☺
absence of packets is not a contractible metric	☹	☺	☺
congestion is outside a customer's control	☹	☹	☺
consumers don't like variable charges	☹	☹	☺
congestion is not an intuitive contractual metric	☹	☹	☹

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 (off by default)

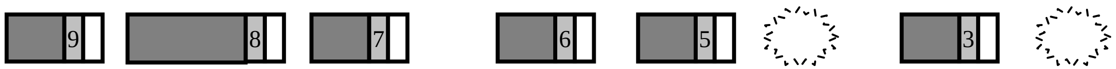


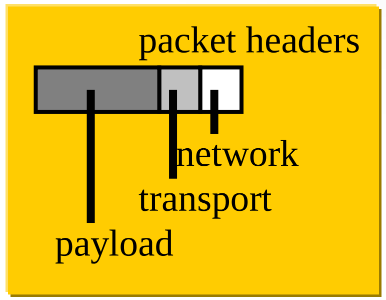
3. **re-inserted ECN [re-ECN]:** standards proposal since 2005
   - packet delivery conditional on sender declaring expected congestion
   - uses ECN equipment in the network unchanged

solution step #1: ECN
make congestion visible to network layer

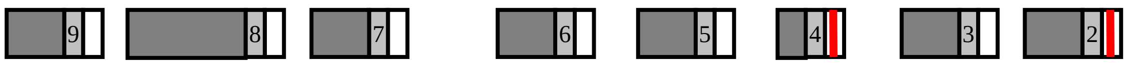


- packet drop fraction is a measure of congestion
  - but how does network at receiver measure holes? how big? how many?
  - can't presume network operator allowed any deeper into packet than its own header
  - not in other networks' (or endpoints') interest to report dropped packets





- solution: Explicit Congestion Notification (ECN)
  - mark packets as congestion ***approaches*** - to avoid drop
  - already standardised into IP (RFC3168 – 2001)
  - implemented by most router vendors – very lightweight mechanism
  - but rarely turned on by operators (yet) – mexican stand-off with OS vendors



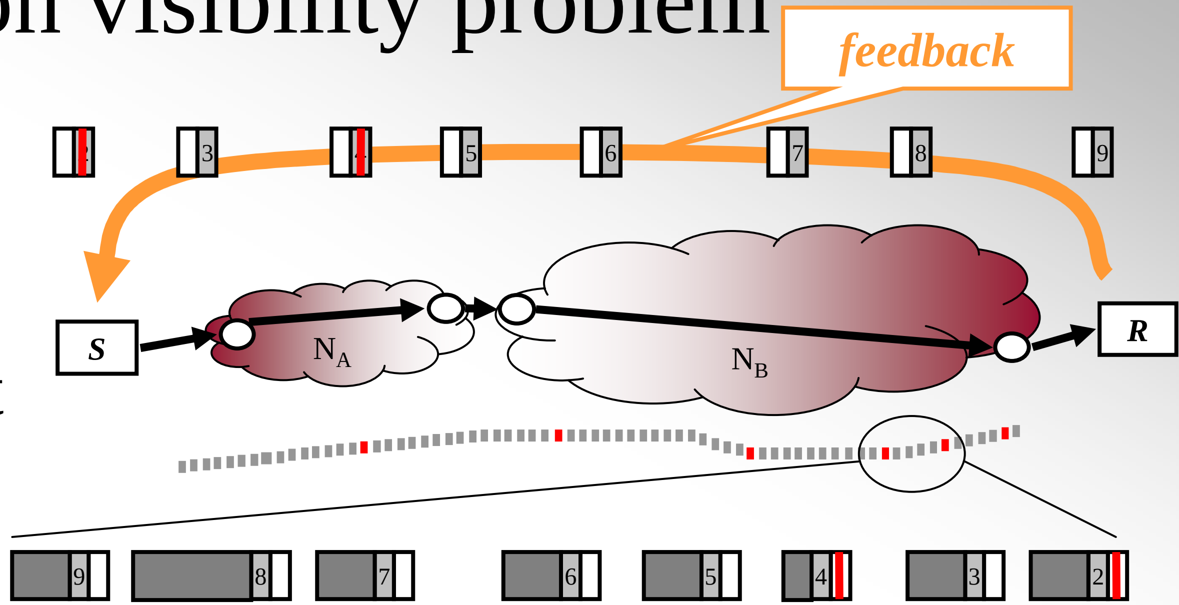
00:	**Not ECN Capable Transport (ECT)**
01 or 10:	**ECN Capable Transport - no Congestion Experienced (sender initialises)**
11:	**ECN Capable Transport - and Congestion Experienced (CE)**

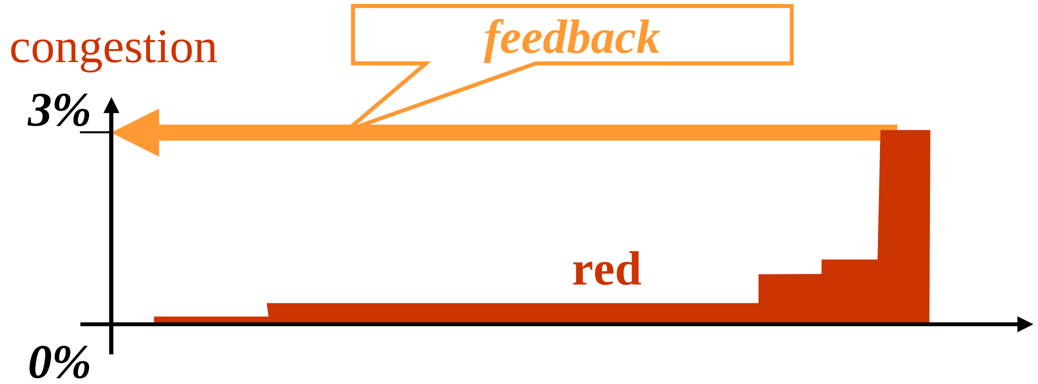
bits 6 & 7 of **IP** DS byte

new information visibility problem

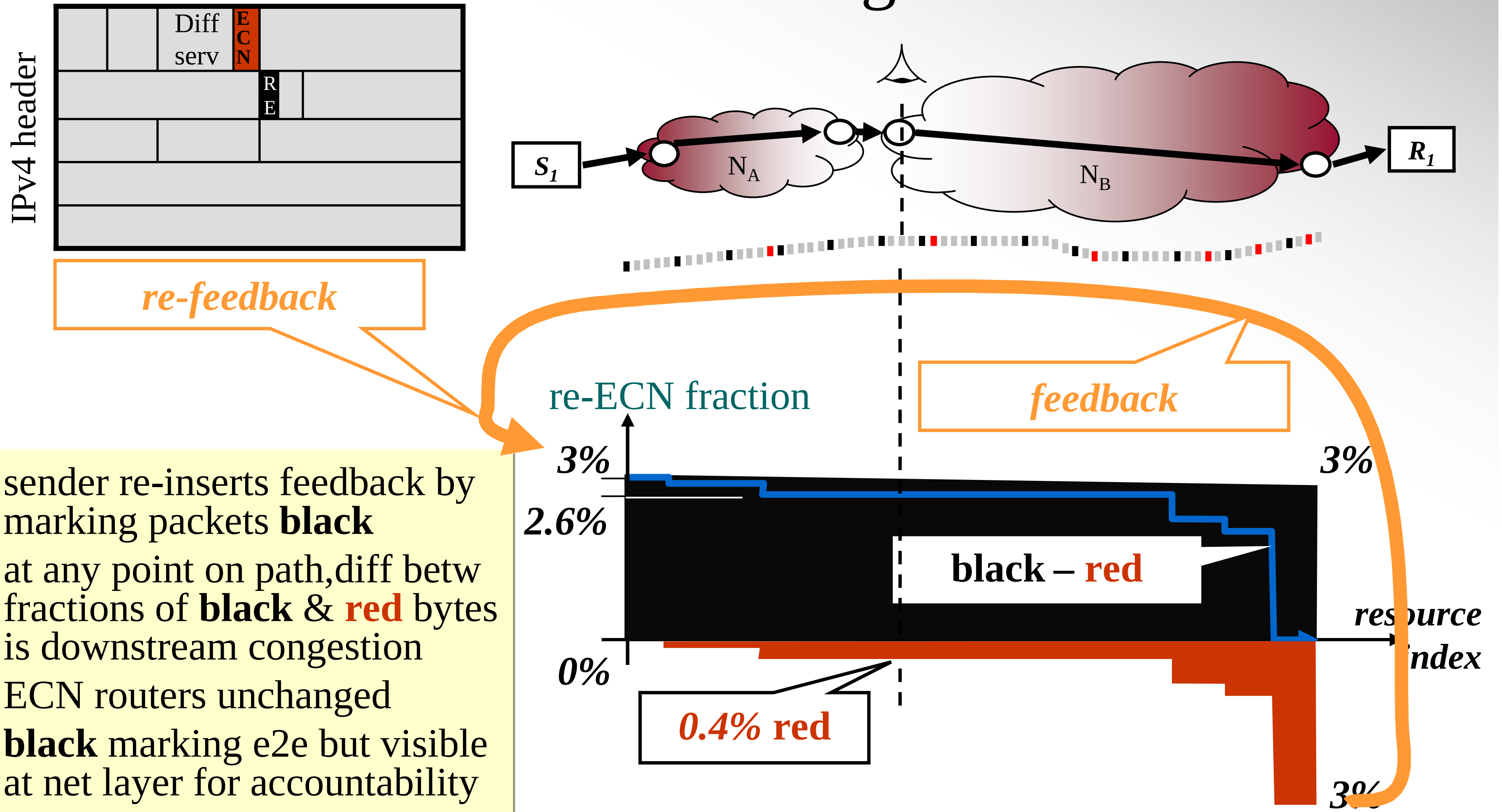
ECN is not enough

- path congestion only measurable at exit
- can’t measure path congestion at entry
  - can’t presume allowed deeper into feedback packets





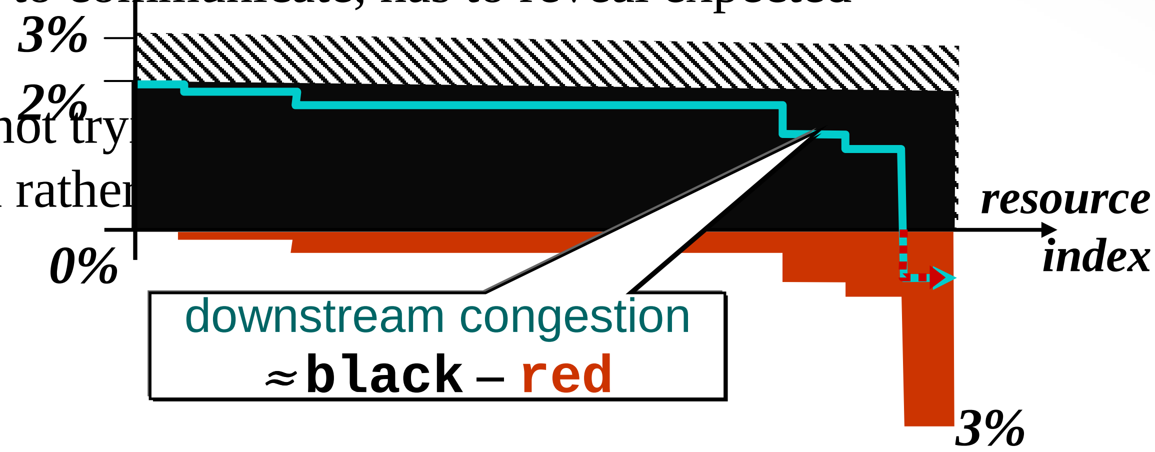
solution step #2: re-ECN
measurable downstream congestion



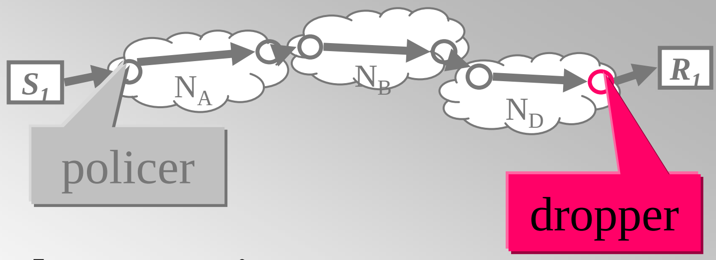
- sender re-inserts feedback by marking packets **black**
- at any point on path,diff betw fractions of **black** & **red** bytes is downstream congestion
- ECN routers unchanged
- **black** marking e2e but visible at net layer for accountability

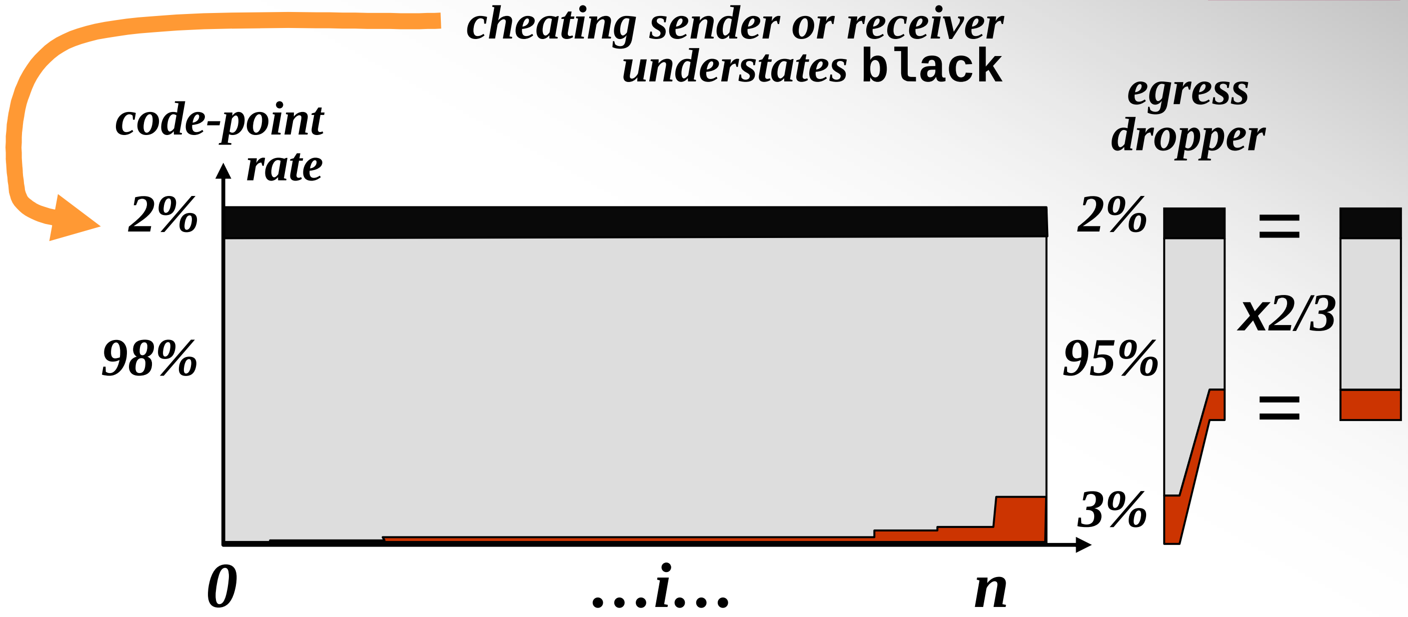
proposed re-ECN service model

- to encourage sender (or proxy) to indicate sufficient expected congestion...
- Internet won't try to deliver packet flows beyond the point where more congestion has been experienced than expected
  - if sender wants to communicate, has to reveal expected congestion
  - even if sender not try
    can be dropped rather
    congestion



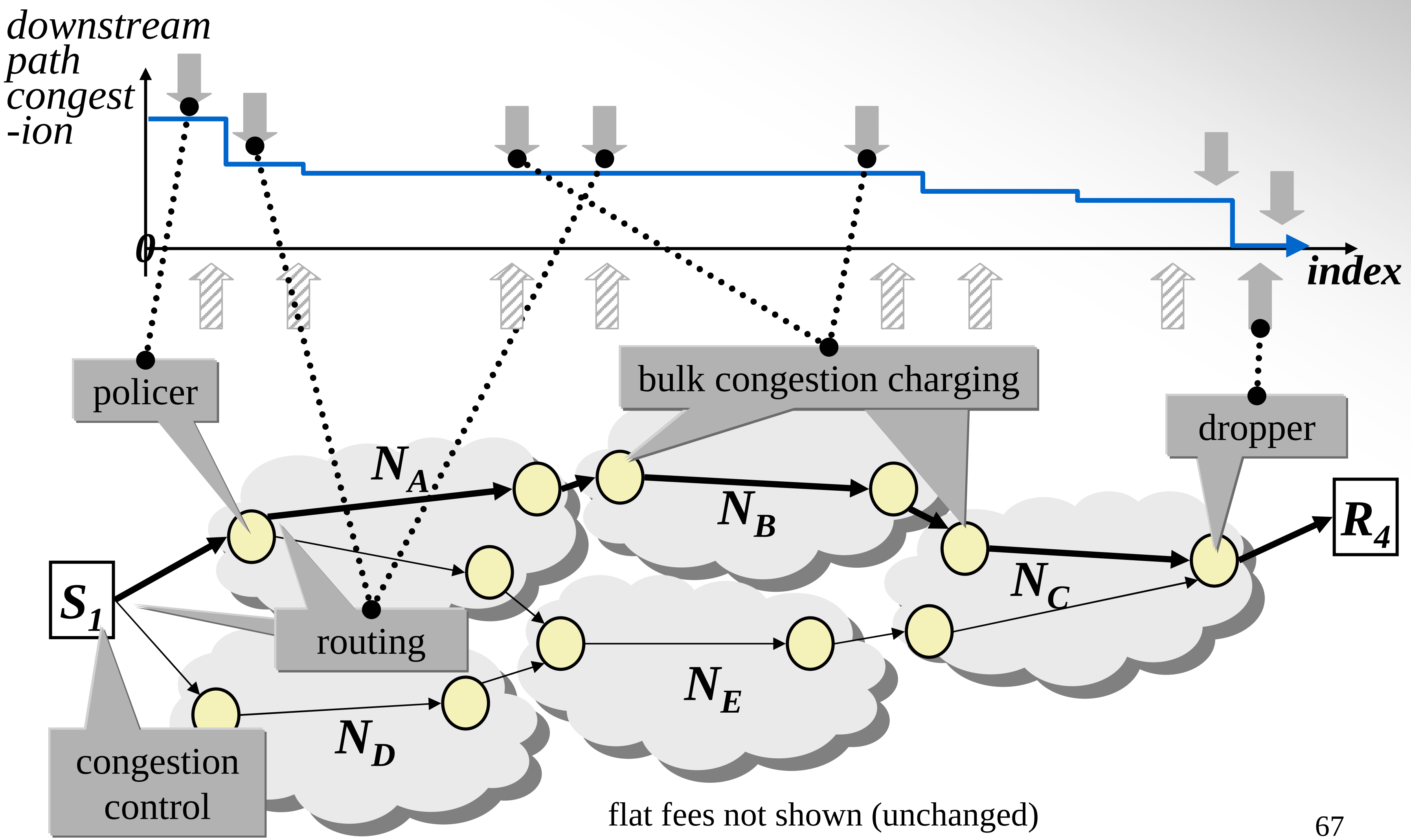
egress dropper (sketch)





- drop enough traffic to make fraction of **red** = **black**
- goodput best if rcvr & sender honest about feedback & re-feedback

incentive framework



flat fees not shown (unchanged)

how to limit congestion with flat fee pricing

Acceptable Use Policy

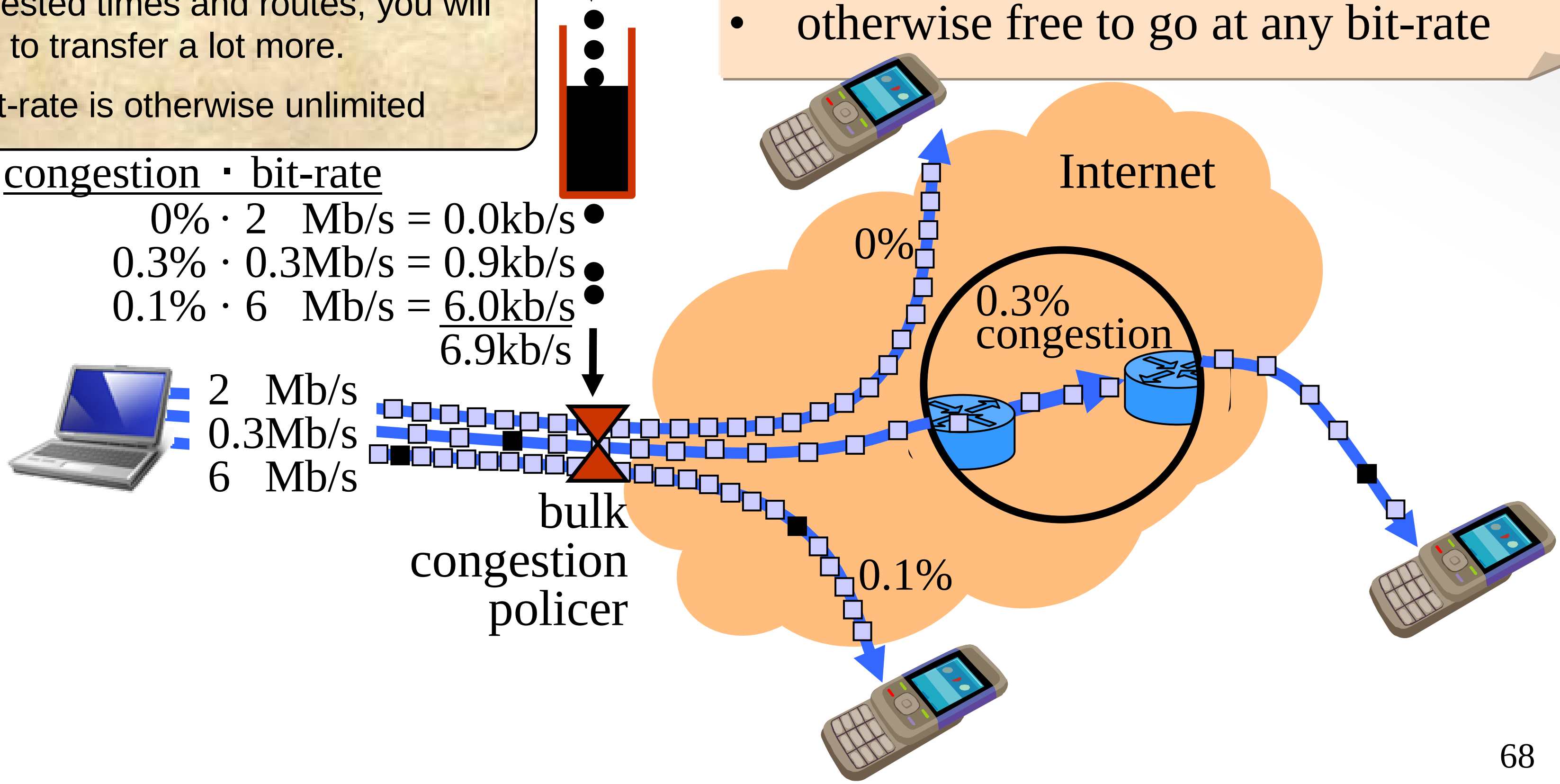
Your 'congestion volume' allowance: 1GB/month (= 3kb/s continuous)
This only limits the traffic you can try to transfer above the maximum the Internet can take when it is congested.

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

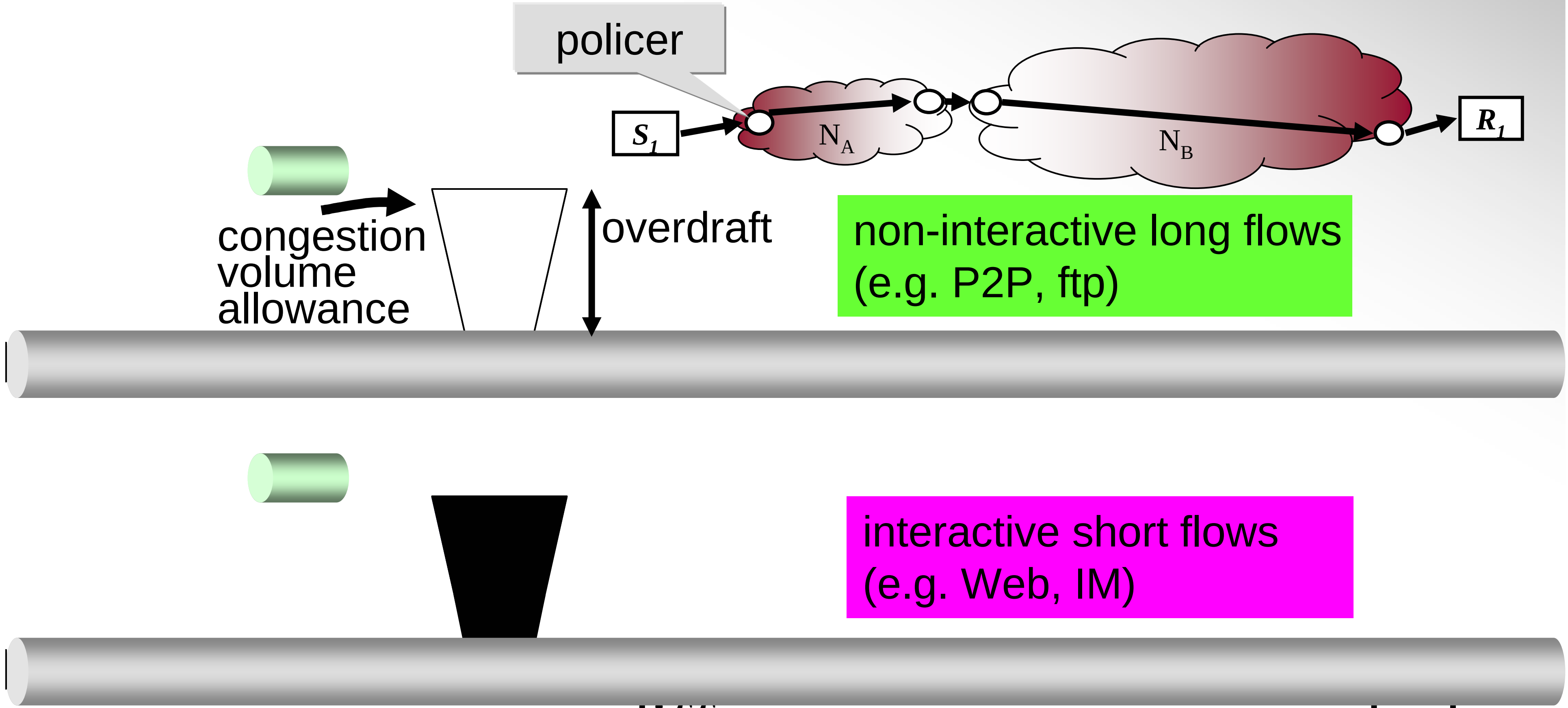
If you use software that seeks out uncongested times and routes, you will be able to transfer a lot more.

Your bit-rate is otherwise unlimited

- simple invisible QoS mechanism
  - apps that need more, just go faster
- only throttles traffic when your contribution to congestion in the cloud exceeds your allowance
- otherwise free to go at any bit-rate

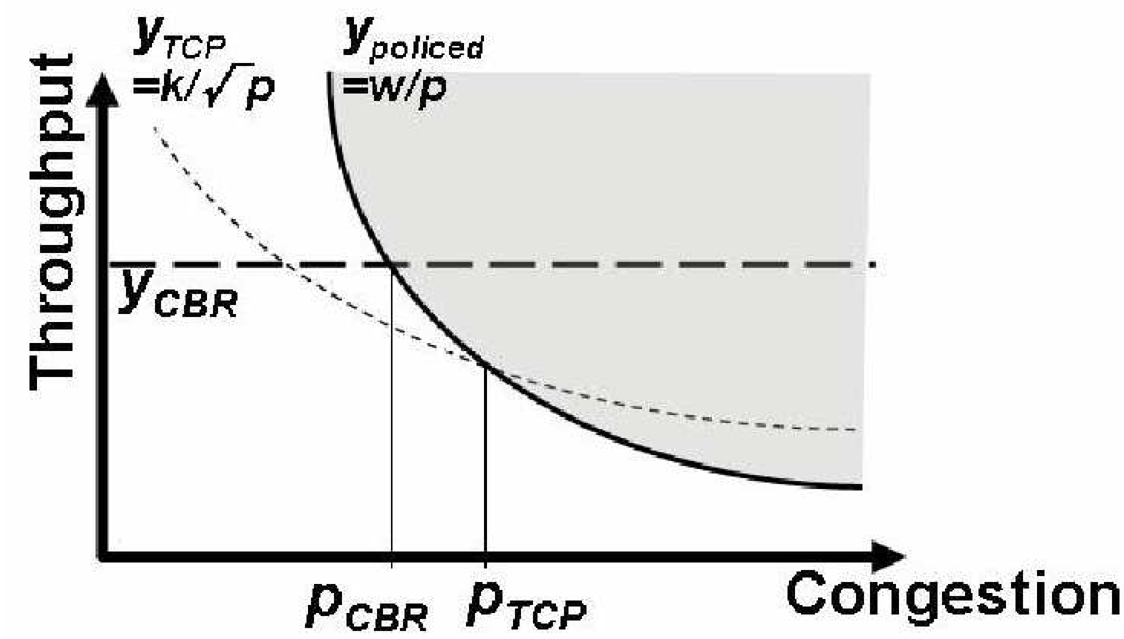


congestion policer – one example: per-user



two different customers, same deal

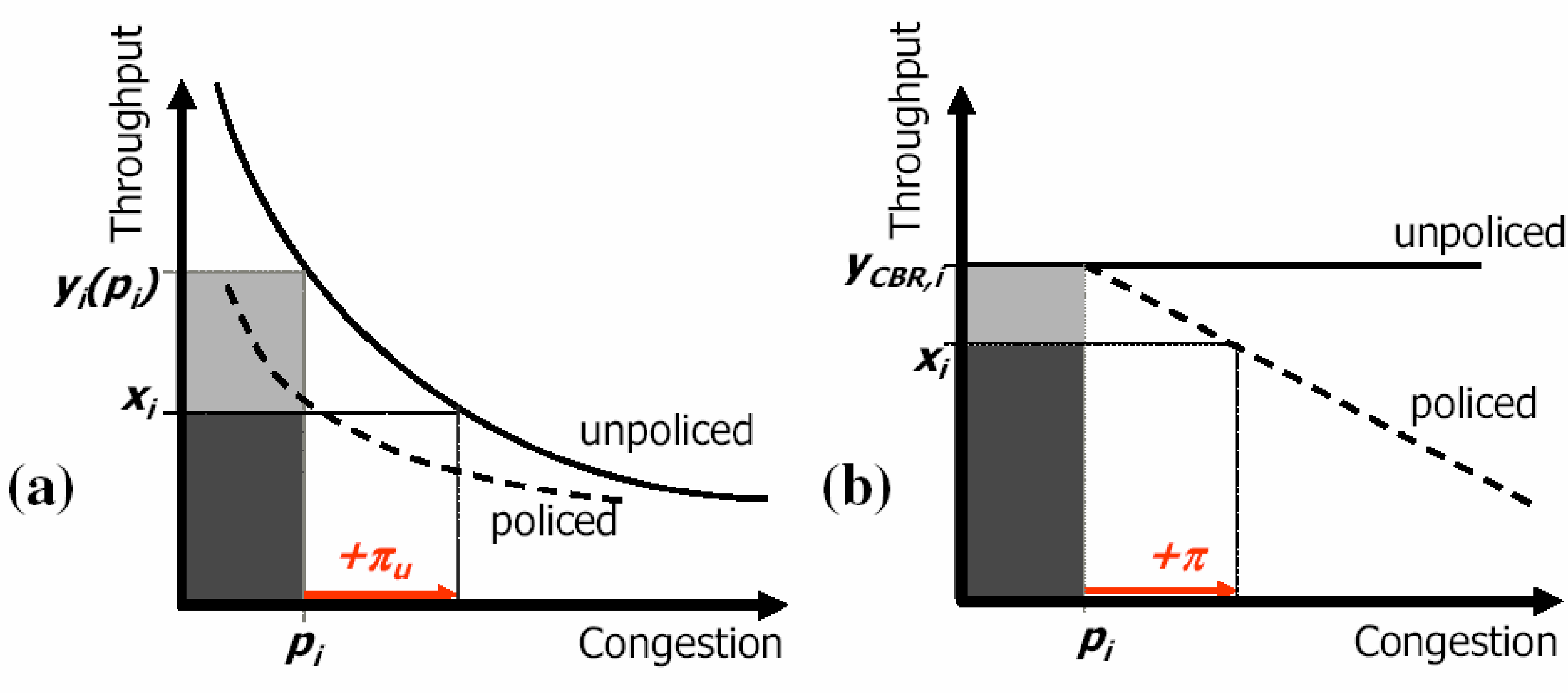
bulk congestion policer



- policer filled with congestion-volume at w [b/s]
- mix numerous flows
  - TCP
  - constant bit-rate (CBR)
- no policer intervention while in white region
- if congestion-volume consumed faster than w [b/s]
  - e.g. too many flows or passing through high congestion or both
  - if each flow r causes congestion p_r, policer limits that flow's bit-rate to

$$y_{policed} = w/p_r$$

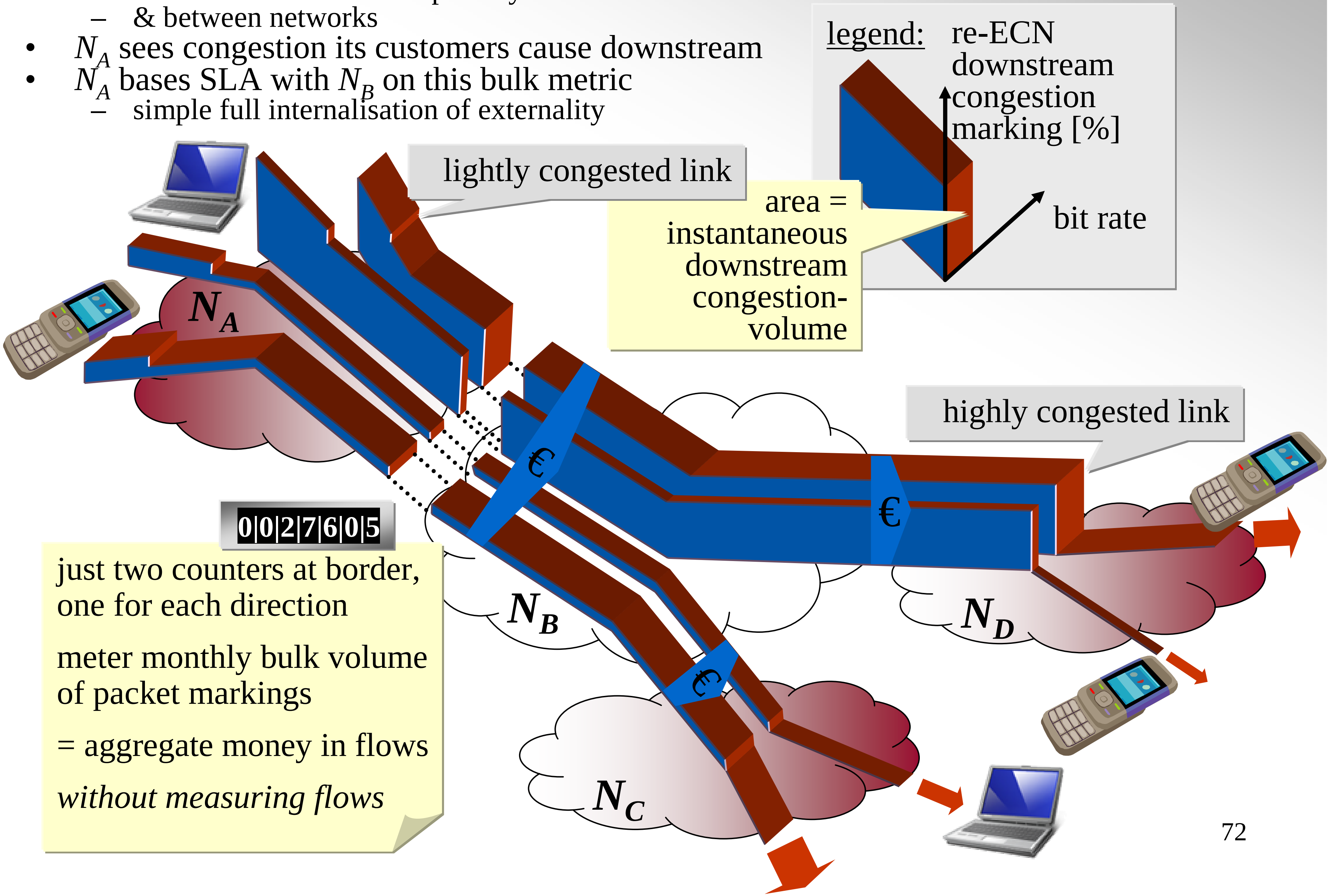
bulk congestion policer
incentive for self-management



- simplest bulk policer (ns2) smoothly takes over congestion control
- if mix of CBR & elastic flows
  - policer losses degrade CBR but it survives – elastic flows compensate
- additional policer losses (π) can be avoided by smart endpoint slowing itself down
- smarter to keep within congestion-volume allowance, but dumb endpoint works OK

routing money

- information symmetry
  - between network & transport layer
  - & between networks
- N_A sees congestion its customers cause downstream
- N_A bases SLA with N_B on this bulk metric
  - simple full internalisation of externality



just two counters at border, one for each direction

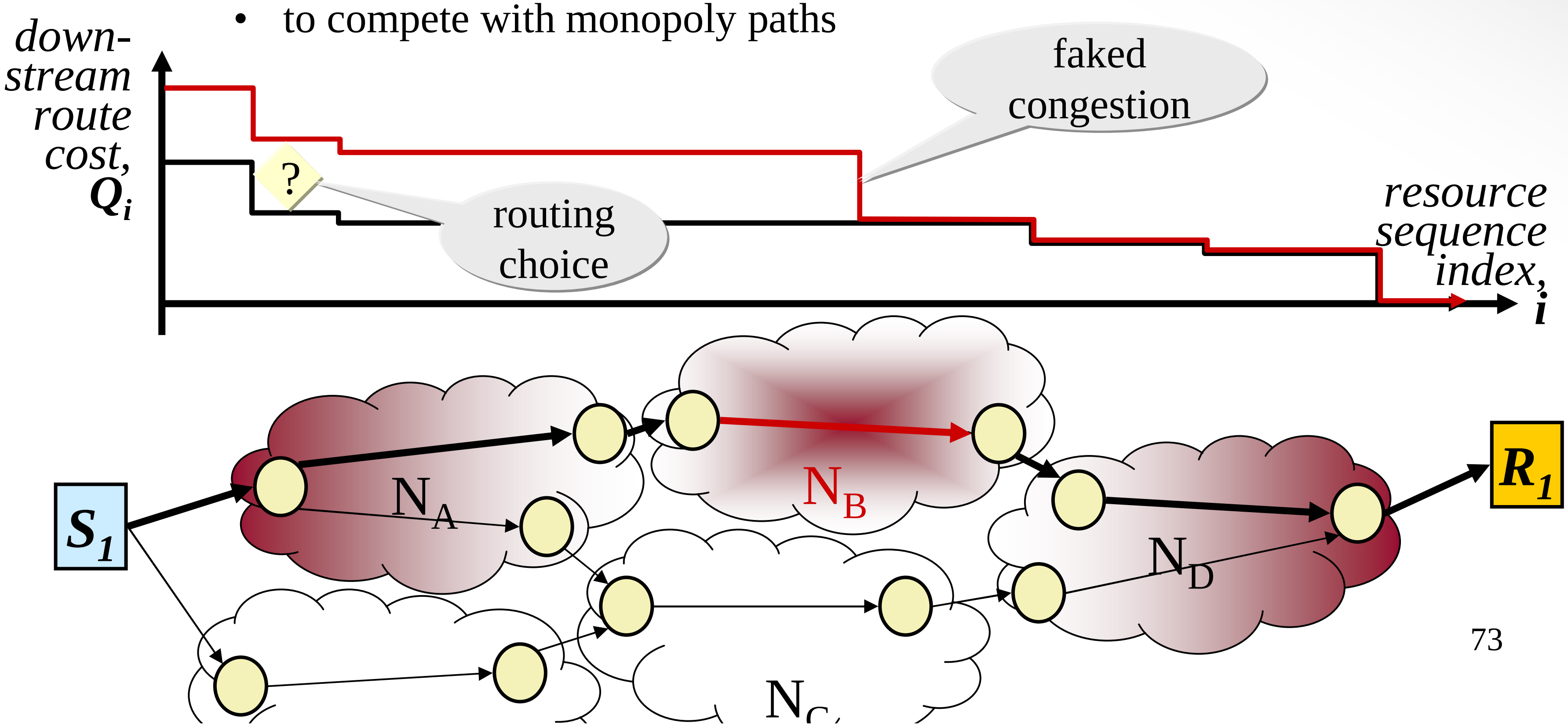
meter monthly bulk volume of packet markings

= aggregate money in flows

without measuring flows

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



fixing re-ECN termination monopoly

- an externality due to ‘sender-pays’
  - sender pays for congestion in the terminating network
  - but receiver chooses the terminating network
  - receiver’s choice causes hidden cost to senders
- solution is *not* ‘receiver-pays’ at network layer
  - no receiver control over packets sent at network layer
  - no control for receiving networks either
- solution
  - implement any receiver-pays sessions directly with sender (e2e)
  - sufficient in some sessions only
  - removes externality, and therefore termination monopoly
  - (assumes natural access monopoly already removed by regulation)

market failures

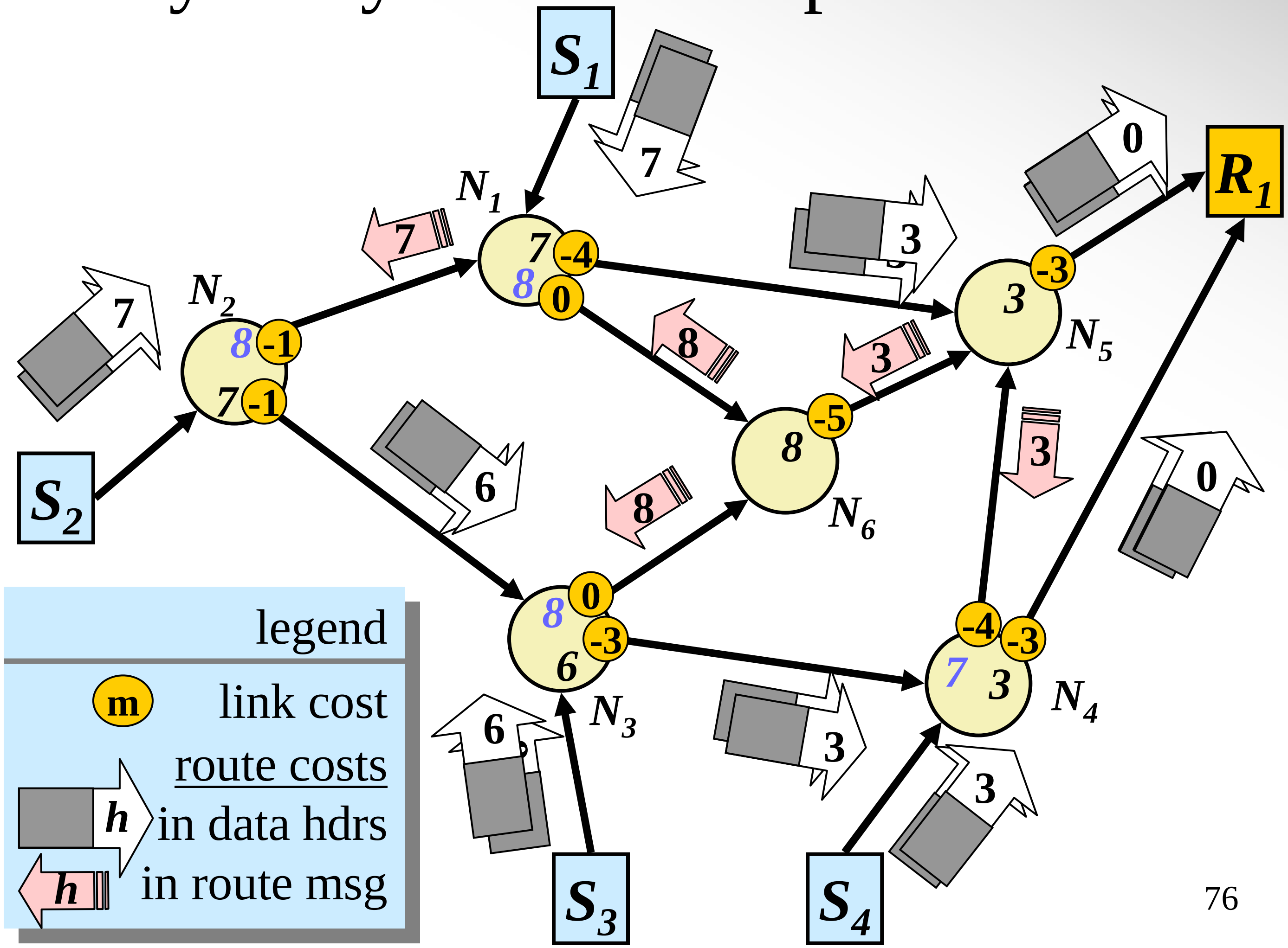
possibly all fixable

- ✓ externalities
  - ✓ (-) congestion
  - ✓ (+) network effects
- ✓ non-excludability
- ✓ market power
  - ✓ natural monopoly
- ✓ switching costs
- ✓ transaction costs
  - ✓ 2-sided market
  - ✓ termination monopoly
- ✓ information asymmetry

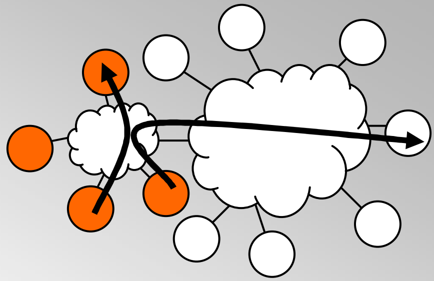
- generally the Internet has solved failures in other markets
  - market mechanisms require ubiquitous information
- the bit-rates people choose could be ‘right’
  a) global utility maximised
  b) supply matches demand
  c) profit squeezed

re-feedback & routing support

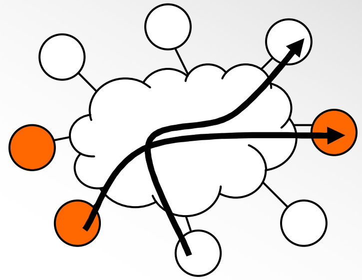
- not done any analysis on this aspect



fairness between fairnesses

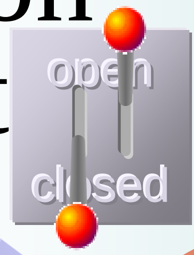


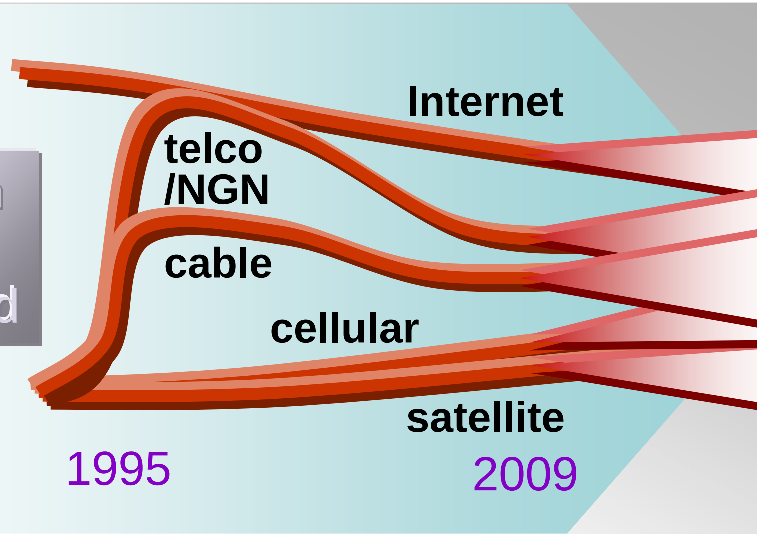
- to isolate a subgroup who want their own fairness regime between them
  - must accept that network between them also carries flows to & from other users
- in life, local fairnesses interact through global trade
  - e.g. University assigns equal shares to each student
    - but whole Universities buy network capacity from the market
  - further examples: governments with social objectives, NATO etc



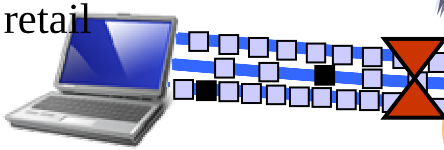
- cost fairness sufficient to support allocation on global market
  - then subgroups can reallocate tokens (the right to cause costs) amongst their subgroup
    - around the edges (higher layer)
  - naturally supports current regime as one (big) subgroup
    - incremental deployment
- different fairness regimes will grow, shrink or die
  - determined by market, governments, regulators, society – around the edges
  - all built over solely congestion marking at the IP layer – neck of the hourglass

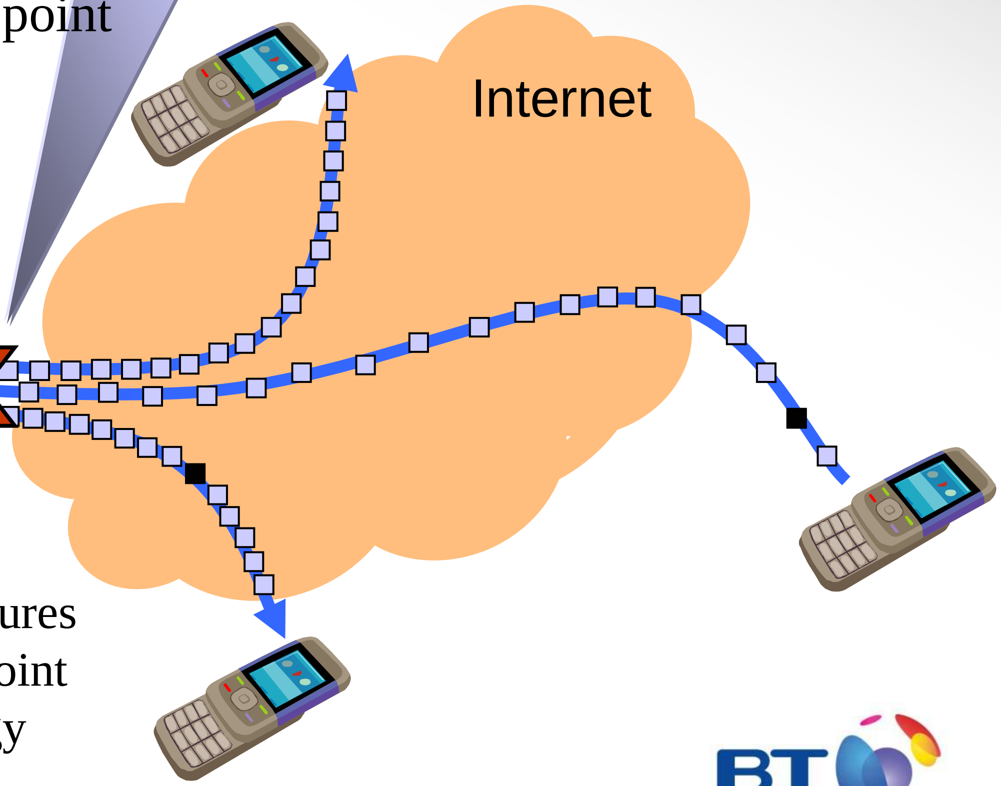
bringing information to the control point





- 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

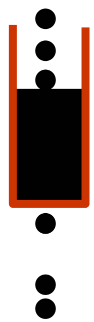
BT

different traffic types

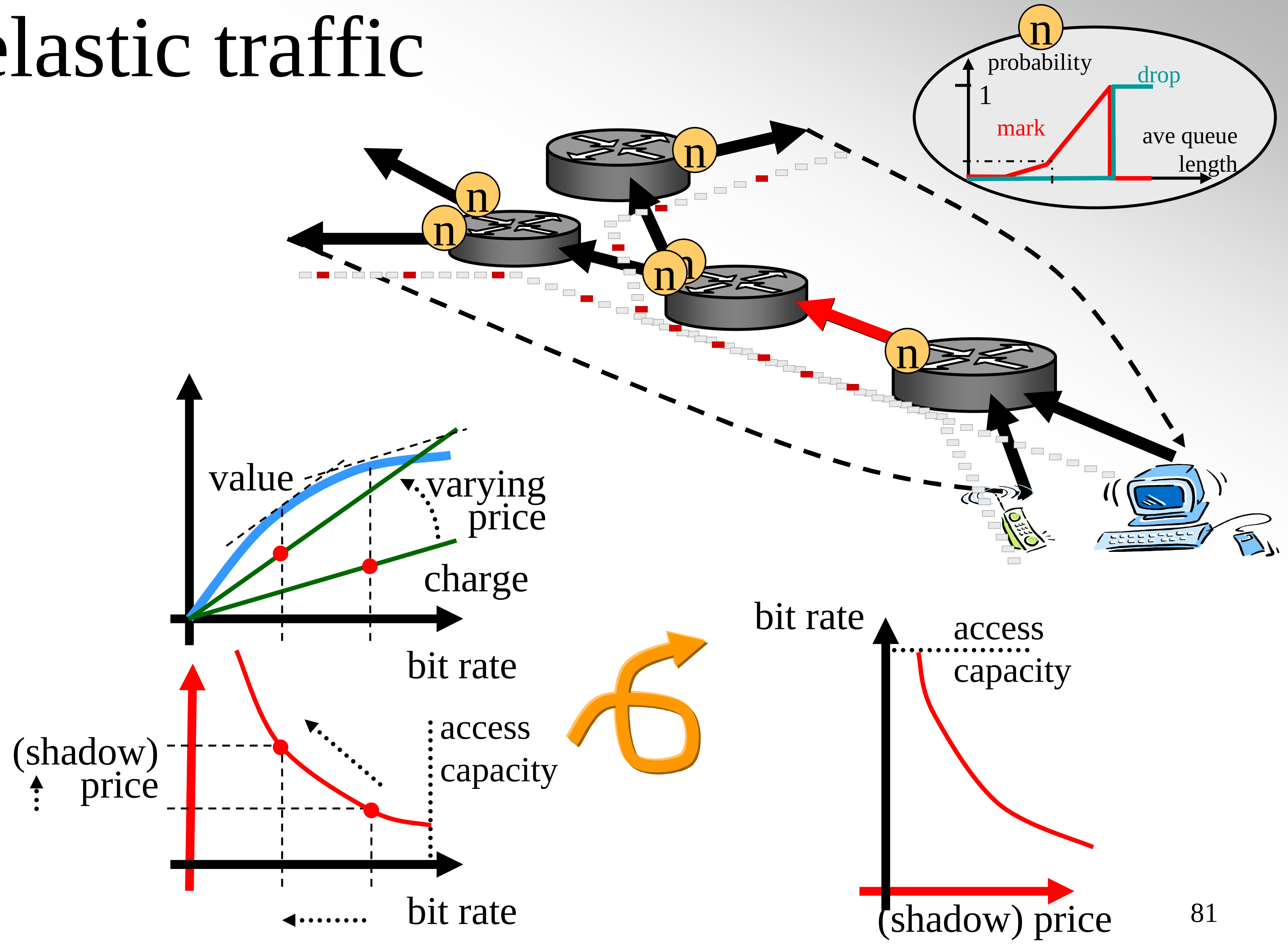
- different congestion controls
- always same accountability & incentive alignment using congestion-volume

delay-intolerant & loss-intolerant

- ECN requires active queue managem’t (AQM)
  - e.g. random early detection (RED)
- AQM keeps queues short (statistically)
  - low delay nearly always (whether ECN or drop)
- ECN keeps drop extremely low
- the remaining QoS dimension: bit-rate
  - re-ECN policing is sufficient control
  - via congestion-volume



elastic traffic

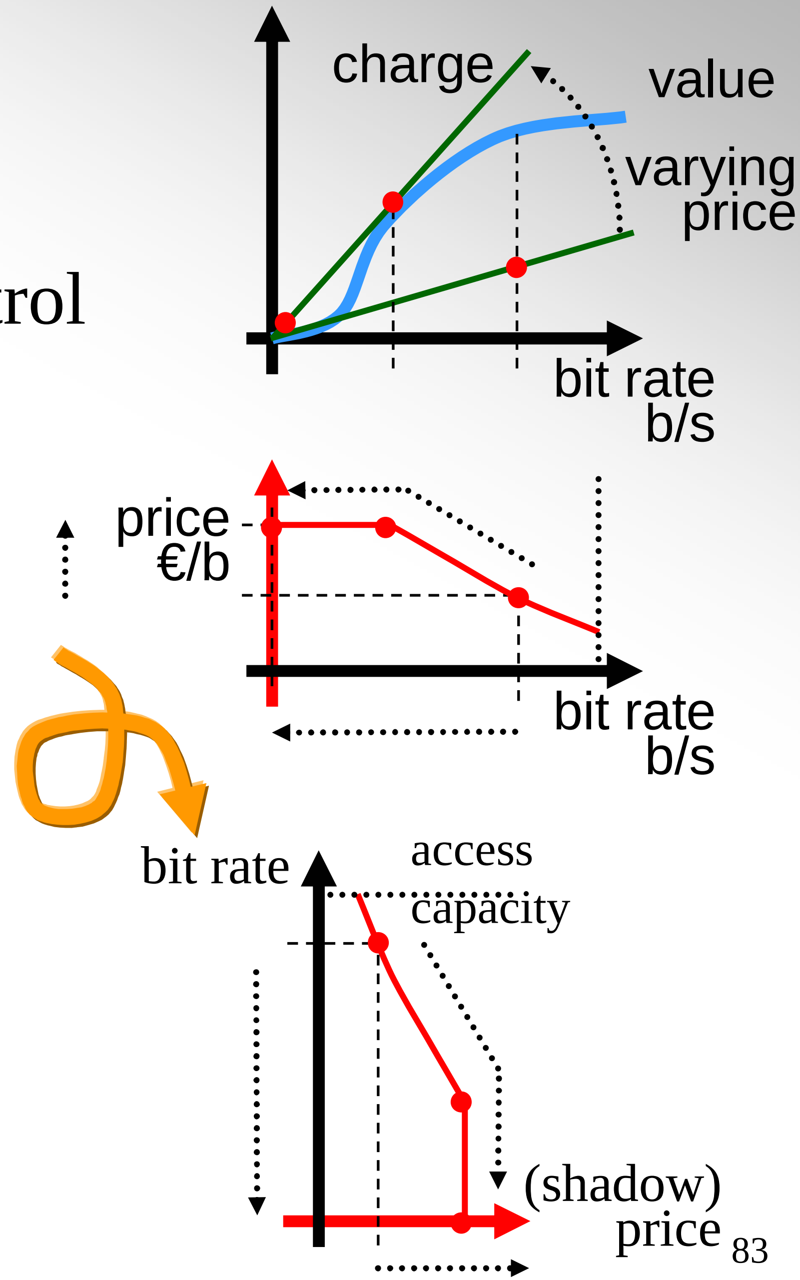


file transfer
fixed volume with utility for completion time

- [Key99] predicts people will flip
    - whenever congestion level drops below a threshold
    - from zero rate to their line rate back to zero otherwise
- [Key04] stabilised if mixed with streaming traffic
- [Gibbens99] adapting to congestion level still pays off
- still active area of research
    - analysis hasn't allowed for round trip delay
        - uncertainty could cause less extreme behaviour
    - TCP has survived well for this class of utility
        - reverse engineering TCP to economics would imply elastic utility
    - a series of files is not strictly a fixed object size
        - lower congestion leads to downloading more bits in total
        - some files more optional than others

inelastic traffic

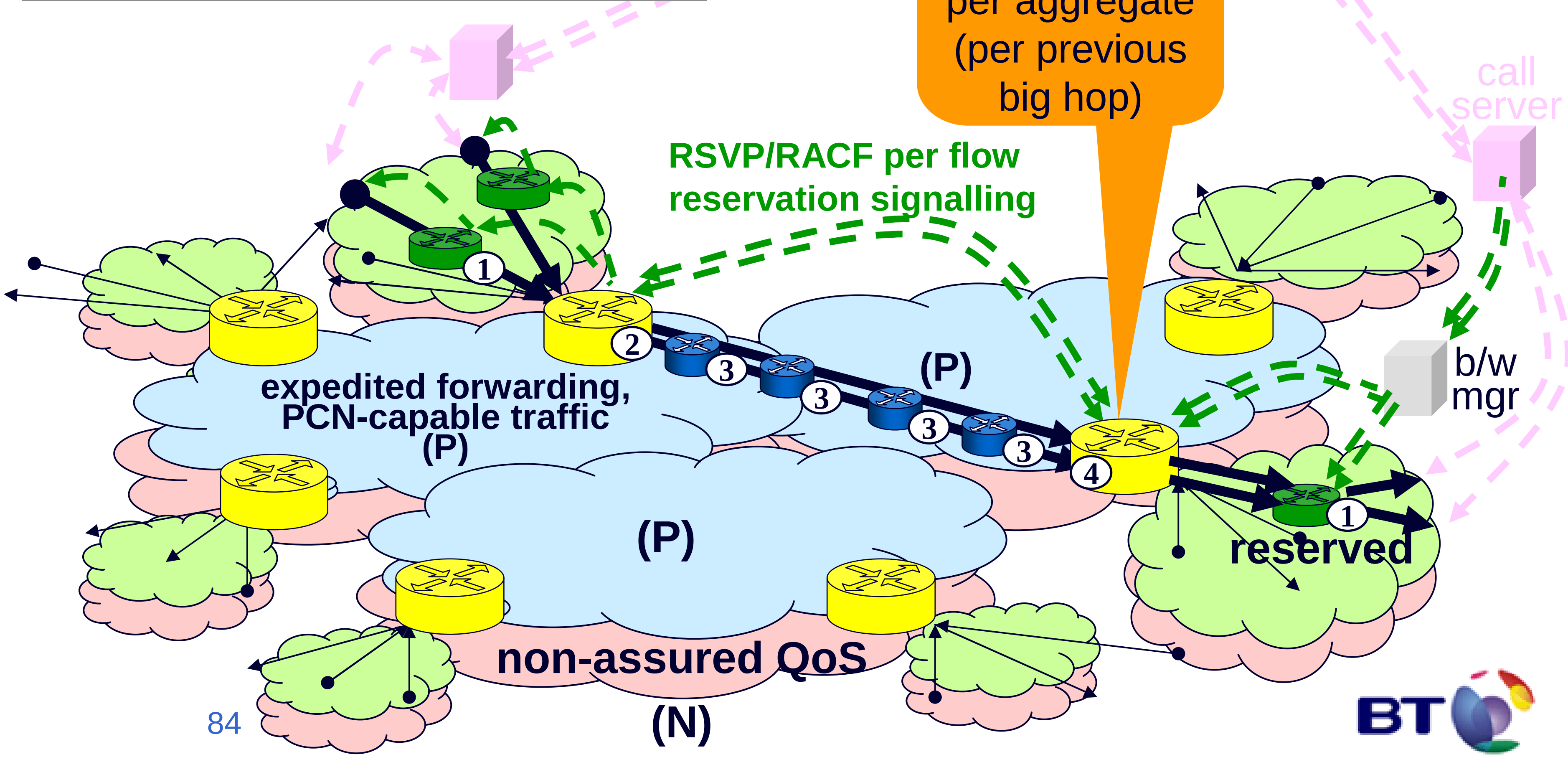
- scalable flow admission control
  - for sigmoid-shaped value curves (inelastic streaming media)
  - see [PCN] for single domain
  - see [re-PCN] for inter-domain



PCN system arrangement

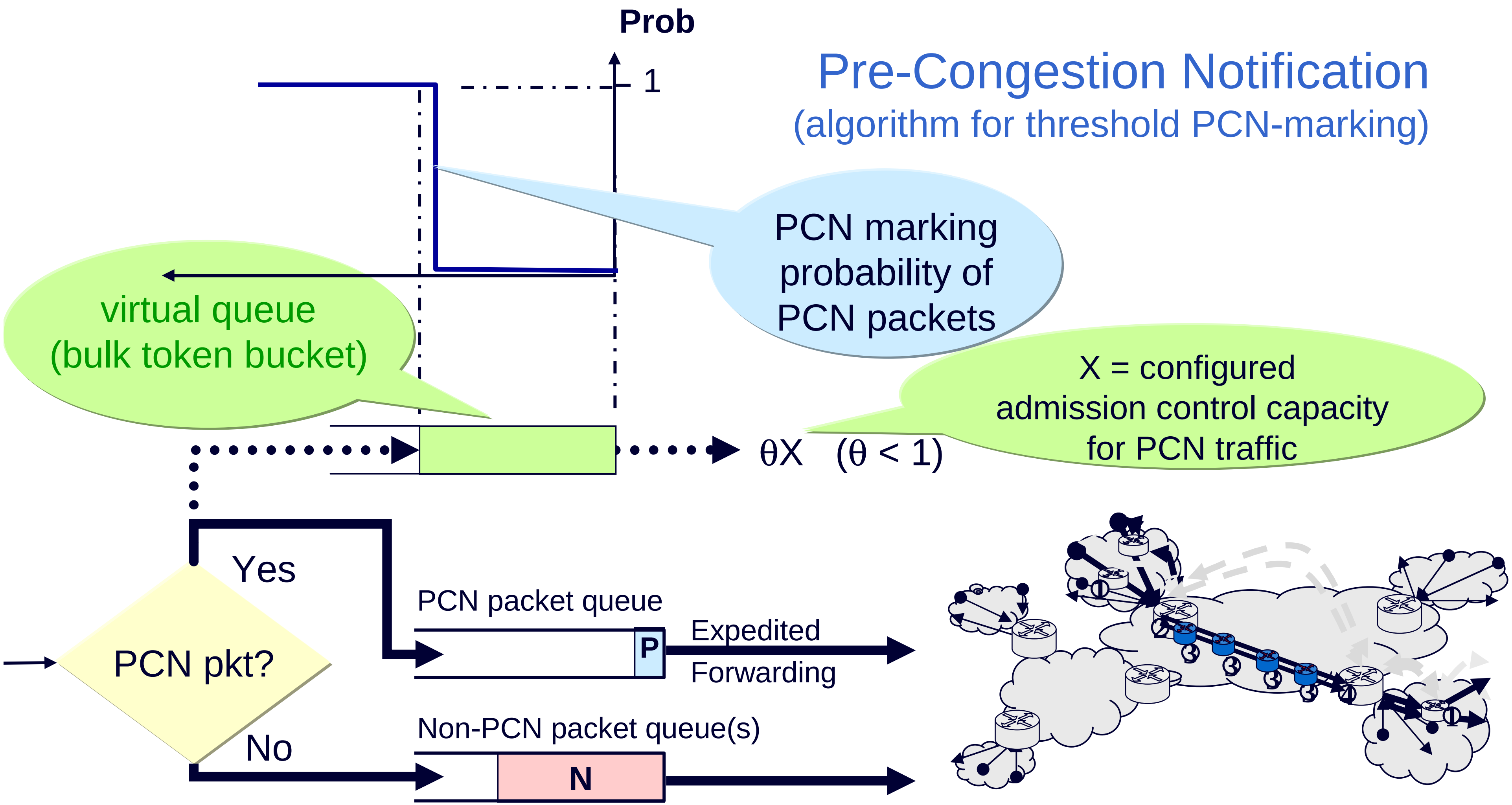
highlighting 2 flows

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



Pre-Congestion Notification

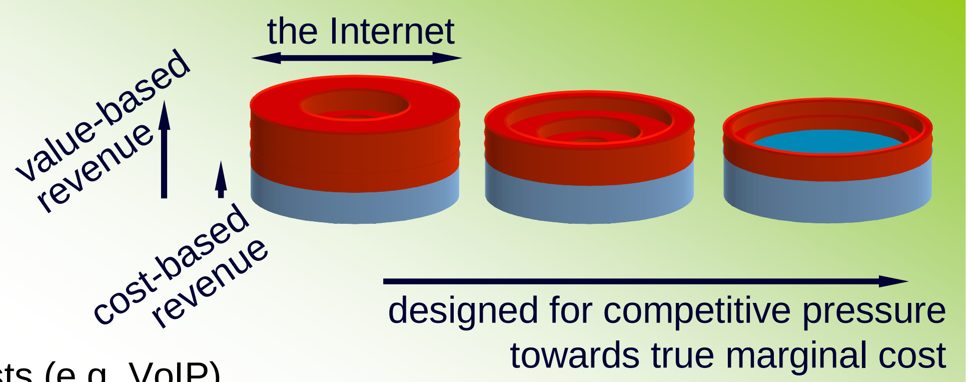
(algorithm for threshold PCN-marking)



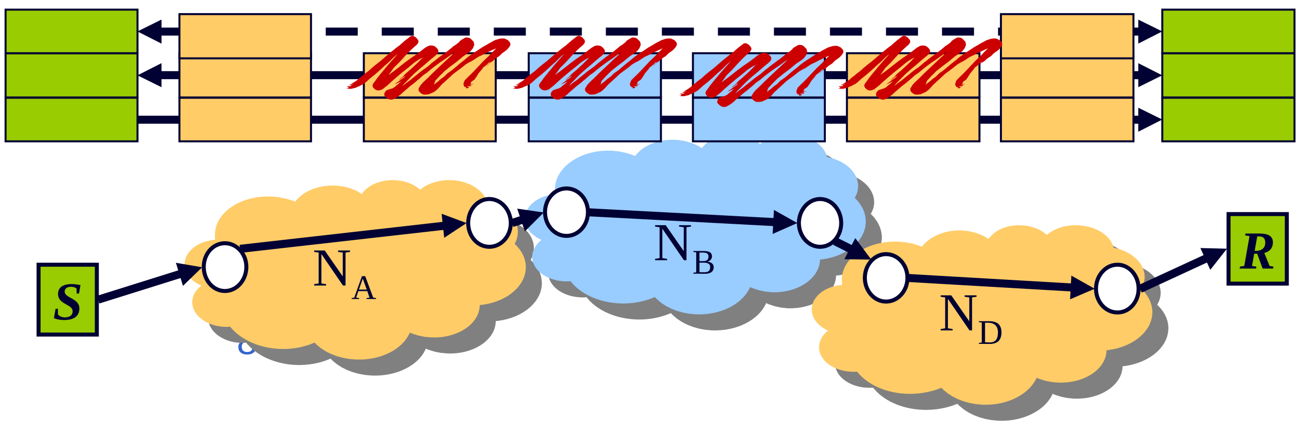
- virtual queue (a conceptual queue – actually a simple counter):
  - drained somewhat slower than the rate configured for adm ctrl of PCN traffic
  - therefore build up of virtual queue is ‘early warning’ that the amount of PCN traffic is getting close to the configured capacity
  - NB mean number of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

BT

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) QoS admission	**per session**
priority forwarding **& PCN**	**bulk data**

BT

legend

connection-oriented (CO) QoS

PCN QoS

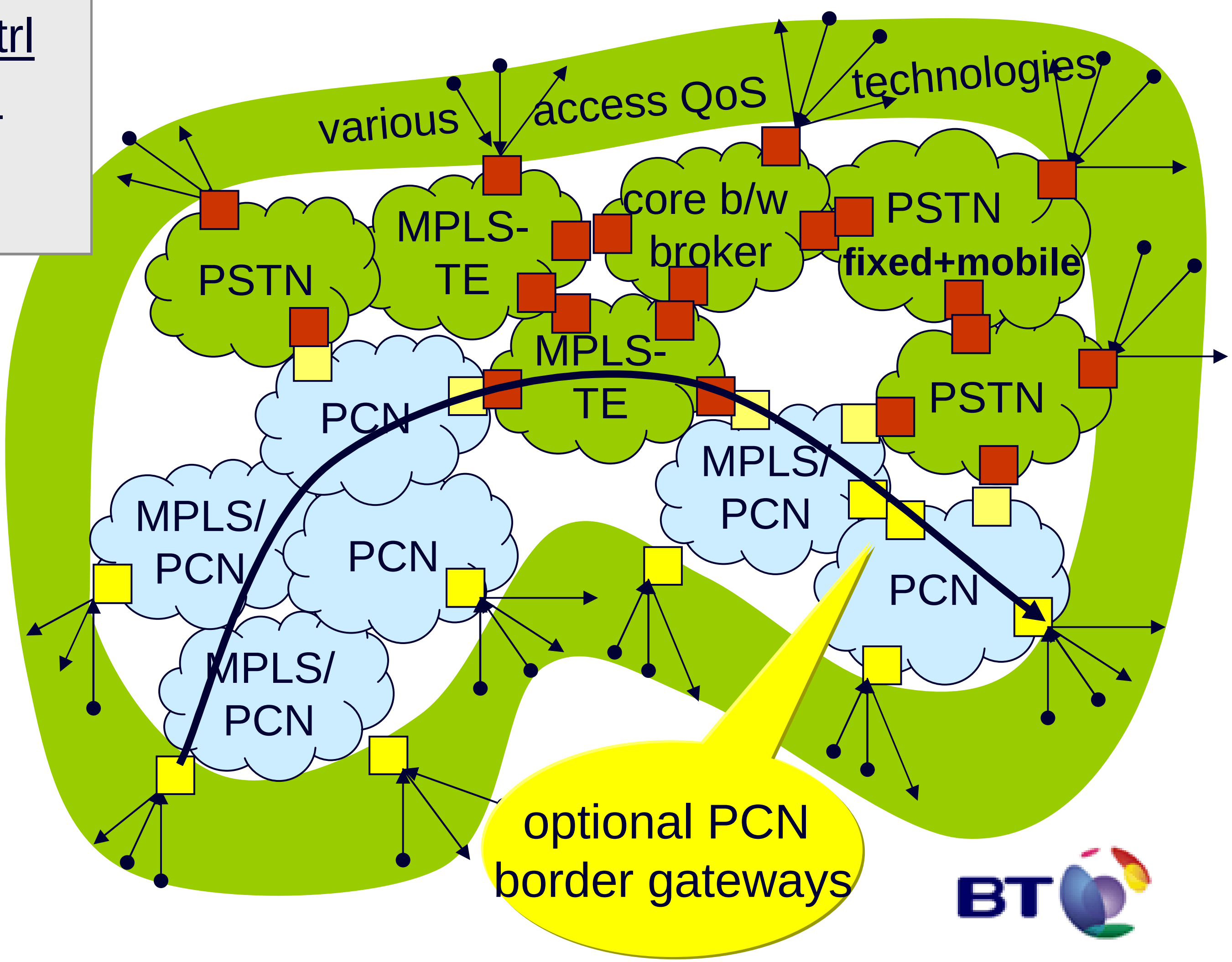
flow admission ctrl & border policing

PCN / CO

CO / CO

PCN

the wider it is deployed the more cost it saves



Still initiated by end to end app layer signalling (SIP)

Figure focuses on layers below

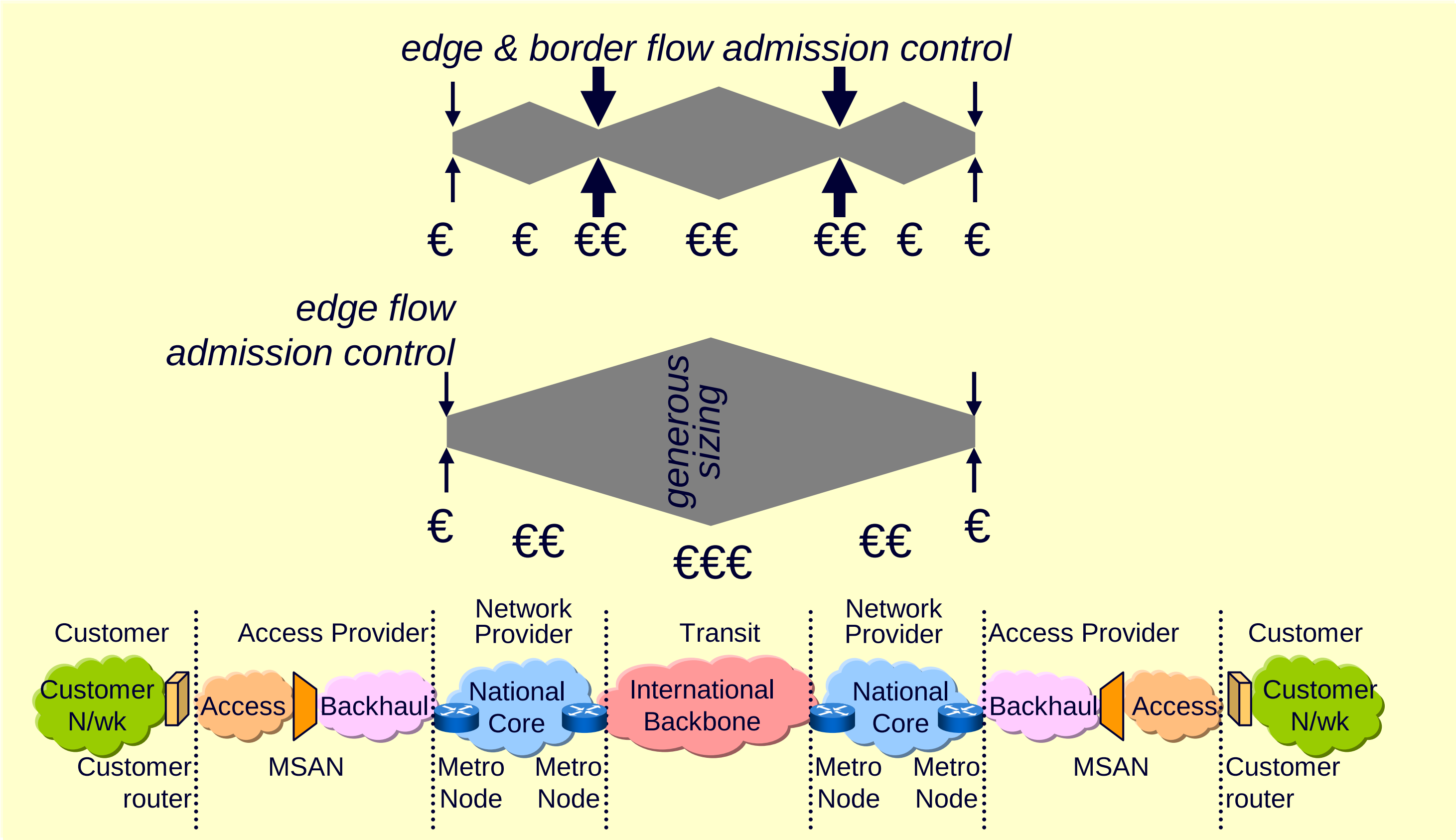
PCN status

- main IETF PCN standards appearing through 2009
  - main author team from companies on right (+Universities)
  - wide & active industry encouragement (no detractors)
- IETF initially focusing on *intra*-domain
  - but chartered to “keep *inter*-domain strongly in mind”
  - re-charter likely to shift focus to interconnect around Mar’09
- detailed extension for interconnect already tabled (BT)
  - holy grail of last 14yrs of IP QoS effort
  - fully guaranteed global internetwork QoS with economy of scale
- ITU integrating new IETF PCN standards
  - into NGN resource admission control framework (RACF)

classic trade-off with diseconomy of scale either way

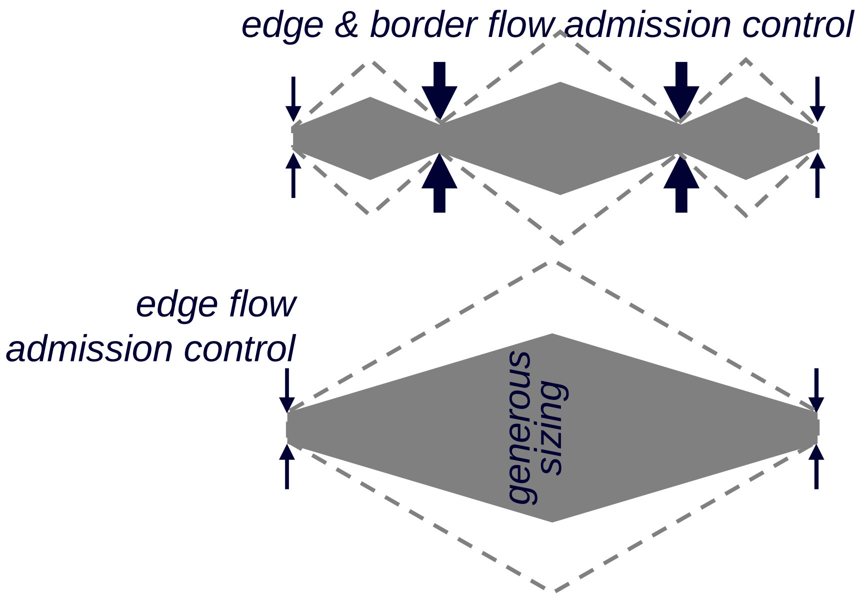
seen in all QoS schemes before PCN

- flow admission ctrl (smarts) vs. generous sizing (capacity)
    - the more hops away from admission control smarts
    - the more generous sizing is needed for the voice/video class



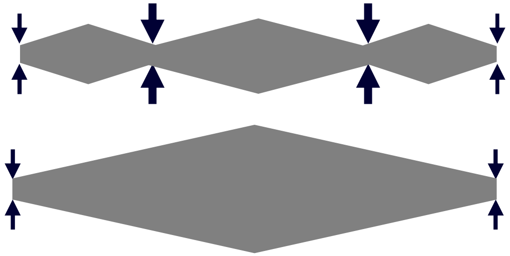
current Diffserv interior link provisioning
for voice/video expedited forwarding (EF) class

- admission control at network edge but not in interior
  - use typical calling patterns for base size of interior links, then...
  - add normal, PSTN-like over-provisioning to keep call blocking probability low
  - add extra Diffserv generous provisioning in case admitted calls are unusually focused



- residual risk of overload
  - reduces as oversizing increases
- stakes
  - brown out of *all* calls in progress

new IETF simplification
pre-congestion notification [PCN]



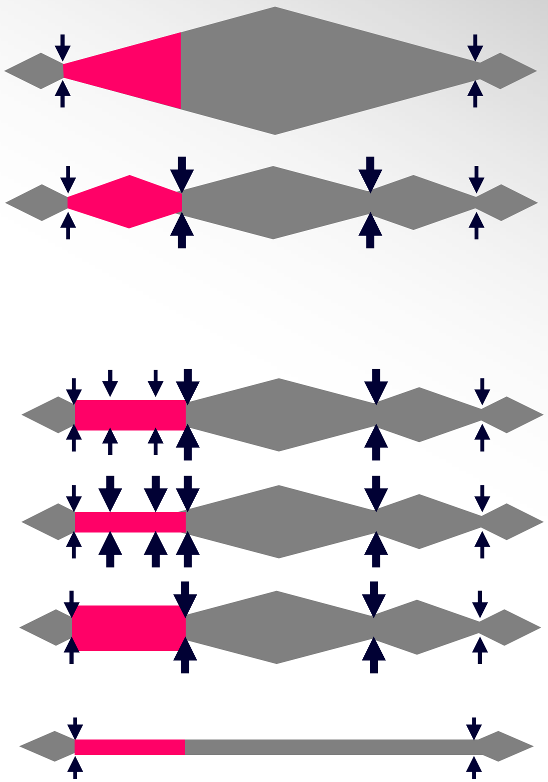
- PCN: radical cost reduction
  - compared here against simplest alternative – against 6 alternatives on spare slide
  - no need for any Diffserv generous provisioning between admission control points
    - 81% less b/w for BT's UK PSTN-replacement
    - ~89% less b/w for BT Global's premium IP QoS
    - still provisioned for low (PSTN-equivalent) call blocking ratios
      as well as carrying re-routed traffic after any dual failure
  - no need for interior flow admission control smarts, just one big hop between edges
- PCN involves a simple change to Diffserv

  - interior nodes randomly *mark* packets as the class nears its provisioned rate
  - pairs of edge nodes use level of marking between them to control flow admissions
  - much cheaper and more certain way to handle very unlikely possibilities
- interior nodes can be IP, MPLS or Ethernet
  - can use existing hardware, tho not all is ideal

core & interconnect QoS

comparative evaluation

	inter-connect	brown-out risk	opex	capex	
				capacity	flow smarts
Diffserv with edge AC but no border AC	bulk rate	finite	££	£££	£
Diffserv with edge and border AC	flow AC	finite	££	££	££
core bandwidth broker	vapour-ware?	finite?	££	£	£££
MPLS-TE hard LSPs and border AC	flow AC	~0	£	££	££
MPLS-TE soft LSPs and border AC	flow AC	~0	£	£	£££
non-blocking core and border AC	flow AC	~0	£	££	££
PCN	bulk congestion	~0	£	£	£

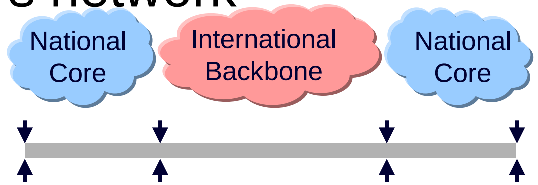


downside to PCN: not available quite yet!

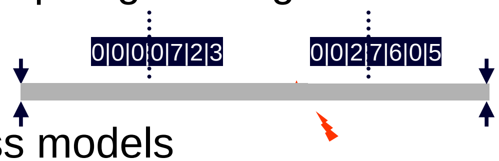
PCN best with new interconnect business model

bulk border QoS [re-PCN]

- can deploy independently within each operator's network
  - with session border controllers & flow rate policing
  - preserves traditional interconnect business model



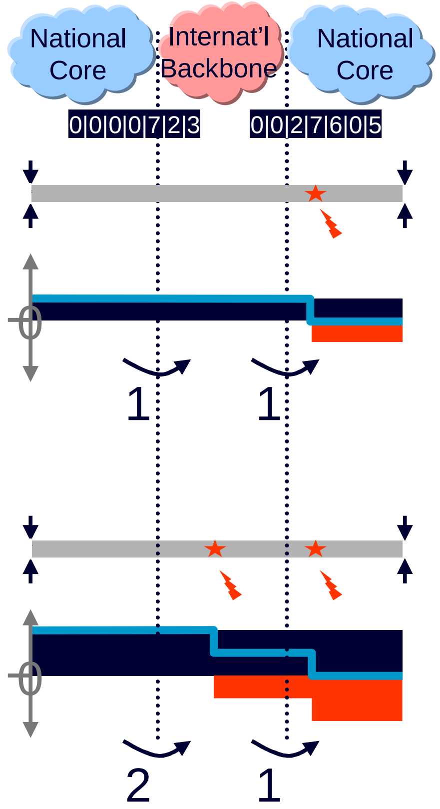
- but most benefit from removing all per-flow border controls
  - instead, simple bulk count of bytes in PCN marked packets crossing border
    - out of band (also helps future move to all-optical borders)
  - each flow needs just one per-flow admission control hop edge to edge
- new business model only at interconnect



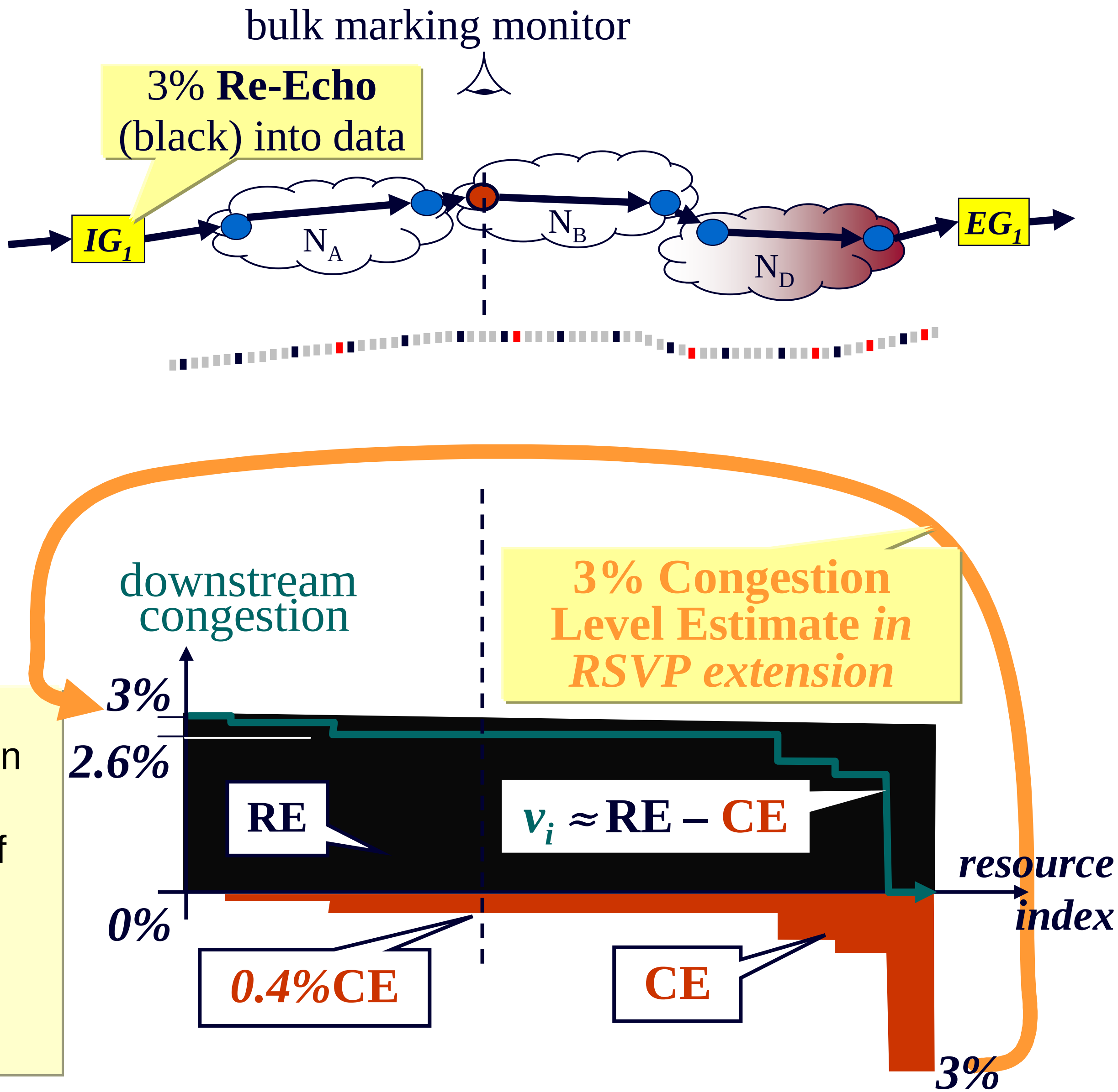
  - no change needed to edge / customer-facing business models
  - not selling same things across interconnects as is sold to end-customer
  - but bulk interconnect SLAs with penalties for causing pre-congestion can create the same guaranteed retail service

accountability of sending networks

- in connectionless layers (IP, MPLS, Ethernet)
  - marks only meterable downstream of network being congested
  - but sending network directly controls traffic
- trick: introduce another colour marking (black) [re-PCN]
  - contractual obligation for flows to carry as much black as red
    - sending net must insert enough black
  - black minus red = pre-congestion being caused downstream
  - still measured at borders in bulk, not within flows
- apportionment of penalties
  - for most metrics, hard to work out how to apportion them
  - as local border measurements decrement along the path they naturally apportion any penalties



re-PCN



- ingress gateway blanks **RE**, in same proportion as fraction of **CE** arriving at egress
- at any point on path, bulk diff betw fractions of **RE** & **CE** is downstream congestion
- routers unchanged

fallacies

rate fairness (esp. max-min)?

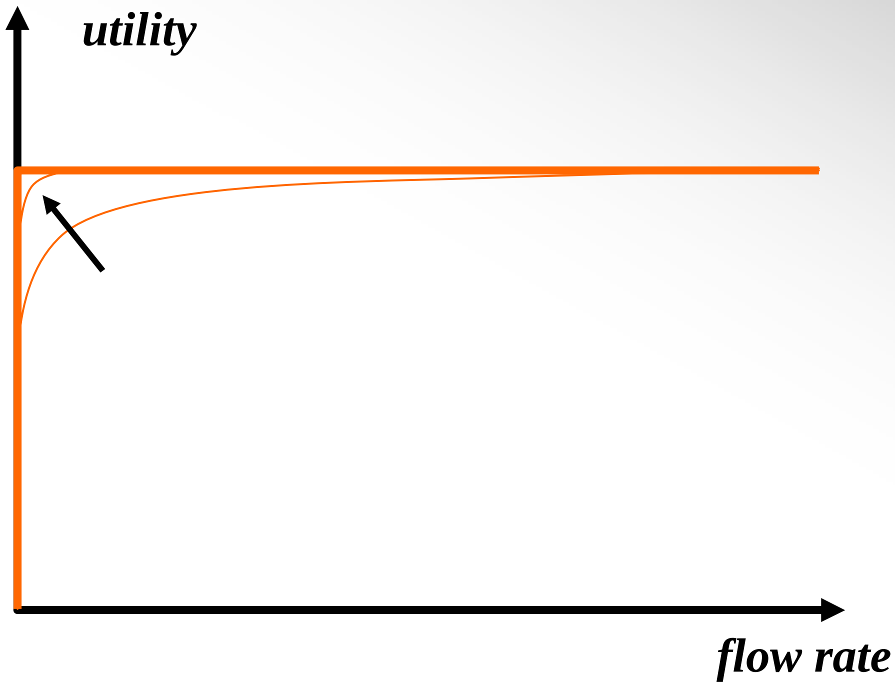
XCP: fairness / efficiency separation?

weighted fair queuing & flow isolation?

TCP-friendly rate-control (TFRC)?

problems with rate fairness
illustration: max-min rate fairness

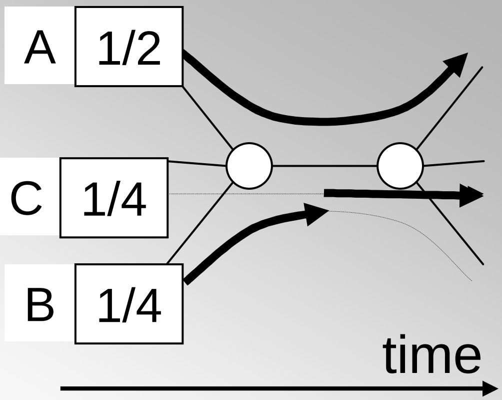
- max-min rate fairness
  - maximise the minimum share
  - then the next minimum & so on
- if users take account of the congestion they cause to others…
- max-min rate fairness would result if all users' valuation of rate were like the sharpest of the set of utility curves shown [Kelly97]
  - they all value high rate exactly the same as each other
  - they all value very low rate just a smidgen less
  - ie, they are virtually indifferent to rate



- users aren't that weird

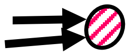
∴ max-min is seriously unrealistic

fair allocation
☒ not between flow rates as shown.
☑ but among users, over time

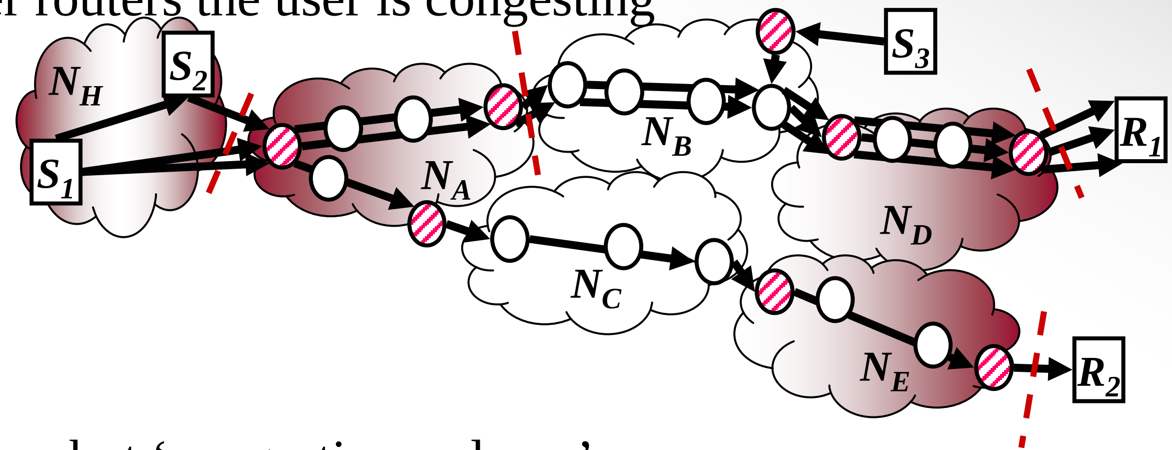


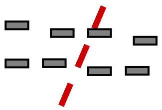
- users A & B congest each other
  - then A & C cause similar congestion, then A & D...
  - is it fair for A to get equal shares to each of B, C & D each time?
- in life fairness is not just instantaneous
  - even if Internet doesn't always work this way, it must *be able* to
  - efficiency and stability might be instantaneous problems, but not fairness
- need somewhere to integrate cost over time (and over flows)
  - the sender's transport and/or network edge are the natural place(s)
- places big question mark over router-based fairness (e.g. XCP)
  - at most routers, data from any user might appear
    - each router would need per-user state
    - and co-ordination with every other router
- XCP claims to be able to separate fairness from efficiency
  - only applies to flow rate fairness, not economic fairness (congestion-volume)
  - false information in XCP protocol hard / impossible to verify [Katabi04]

target structure: *network* fairness



- ✗ bottleneck policers: active research area since 1999 (cf. XCP)
  - detect flows causing unequal share of congestion
  - located at each potentially congested router
  - takes no account of how active a source is over time
  - nor how many other routers the user is congesting
  - based on cheap pseudonyms (flow IDs)





- ✓ re-ECN / ECN
  - like counting volume, but ‘congestion-volume’
  - reveals congestion caused in all Internet resources by all sources (or all sinks) behind a physical interface, irrespective of addressing
  - accumulates over time
  - no advantage to split IDs
- focus of fairness moves from flows to packets

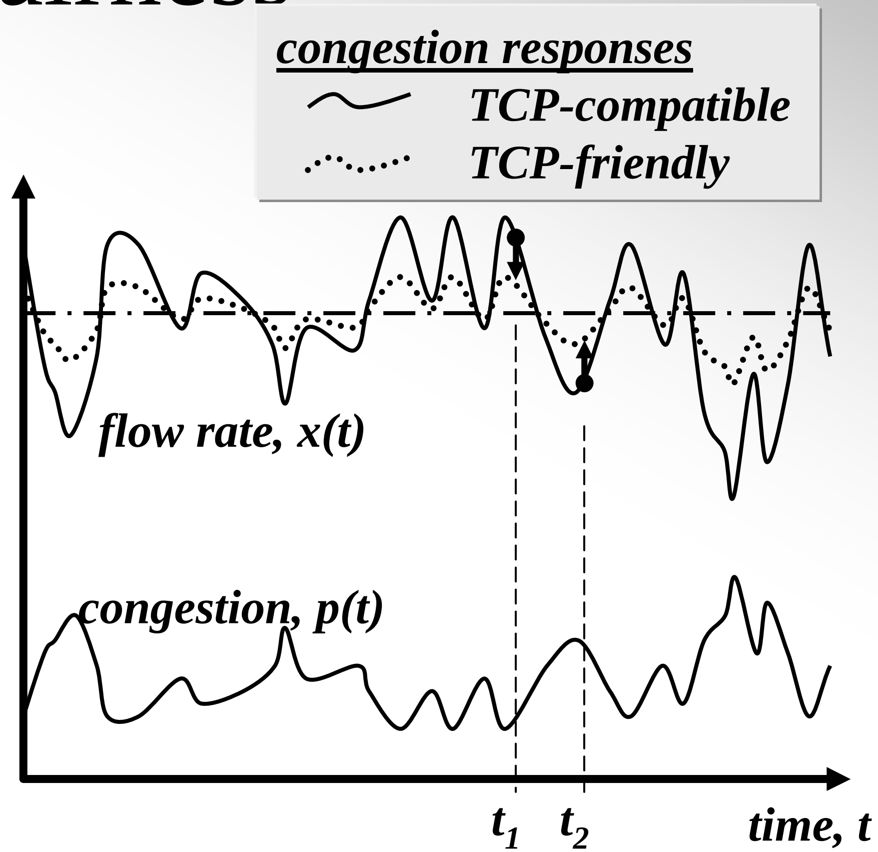
(W)FQ prevents me helping you

- isolation: goal of (weighted) fair queuing (W)FQ
  - separate queues for each user (or each flow)
  - scheduler divides time between active users (or active flows)
  - an excessive user grows own queue, but others unaffected
- user isolation
  - prevents me helping you (e.g. with LEDBAT)
    - I can only help myself
    - isolation between users also isolates me from other users' congestion signals
    - can't respond even though I would be willing to
- flow isolation
  - can't even help my own flows by shuffling others
- as interim, per-user rate policing doesn't close off much
    - just as if a shared link were multiple separate links
  - but per-flow rate policing closes off a lot of future flexibility
    - and it's unnecessary to satisfy anyone's interests

illustration: TCP-friendly rate control (TFRC)
problems with rate fairness

- TCP-friendly
  - same ave rate as TCP
  - congestion response can be more sluggish
- compared to TCP-compatible
  - higher b/w during high congestion
  - lower b/w during low congestion
- giving more during times of plenty doesn’t compensate for taking it back during times of scarcity

- TCP-friendly flow causes more congestion volume than TCP
- need lower rate if trying to cause same congestion cost

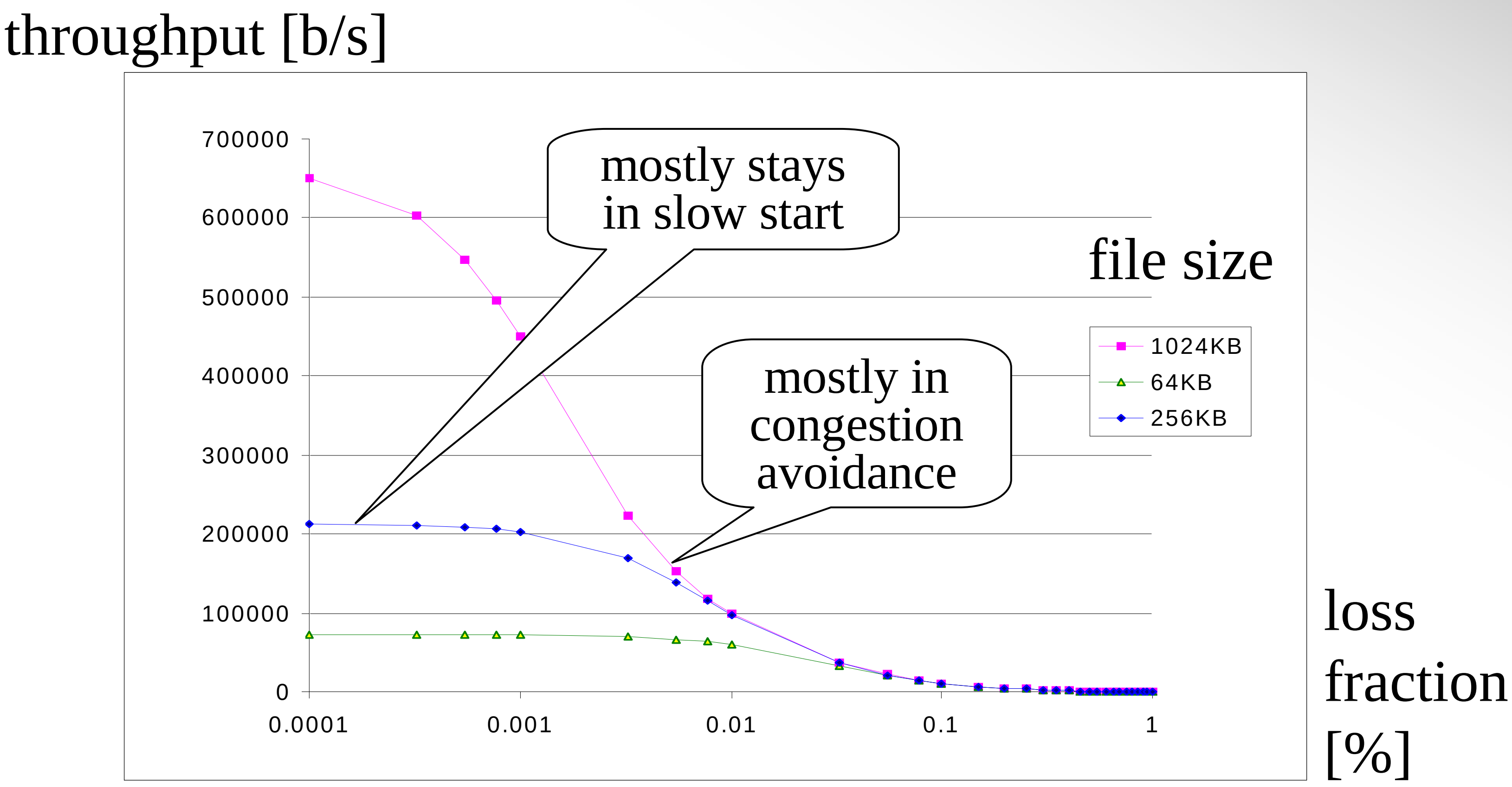


- TFRC vs TCP is a minor unfairness
  - compared to the broken per flow notion common to both

specifics

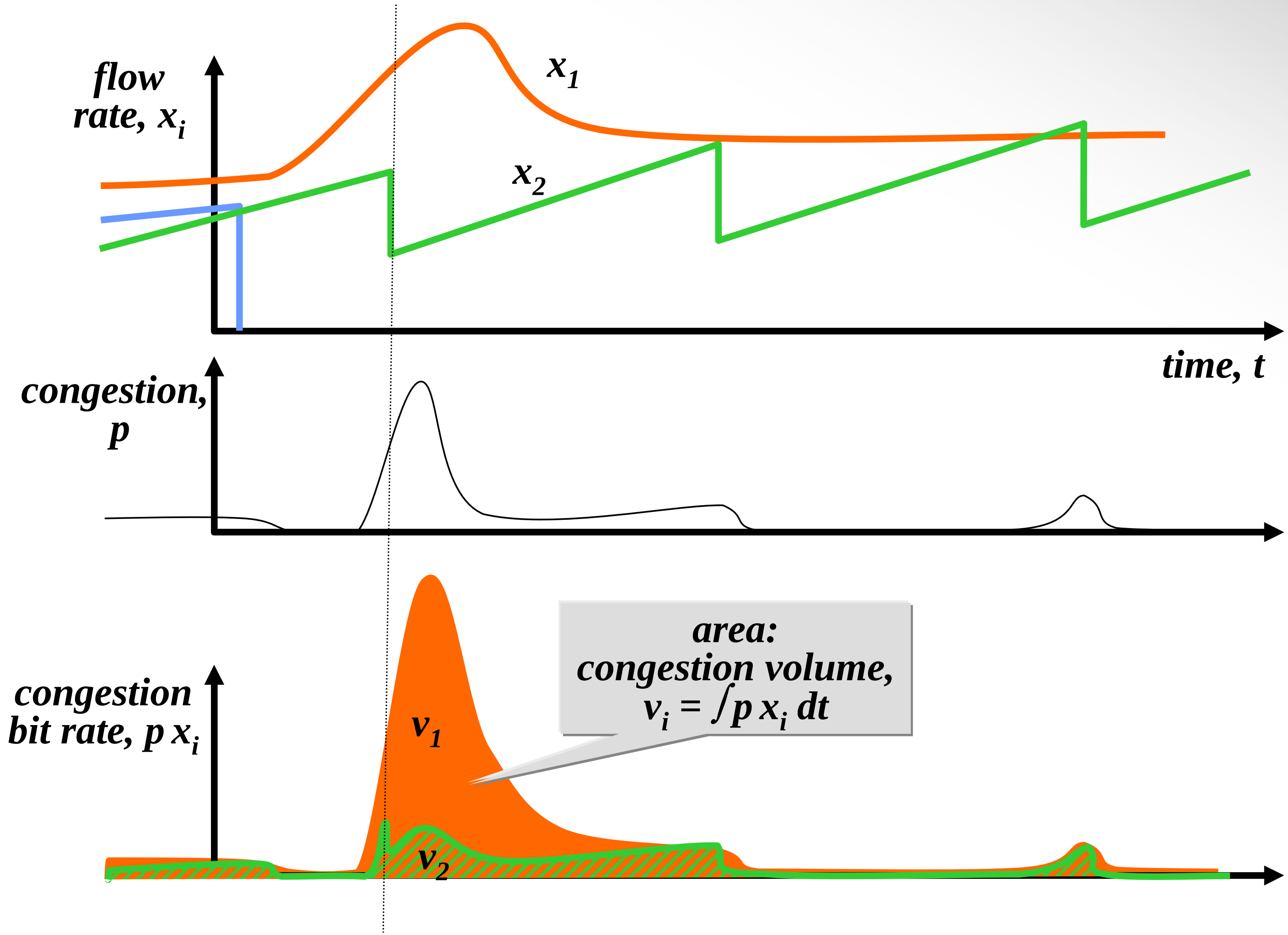
- flow start & transients
- weighted congestion controls
  - multipath transports
  - dependence of bit-rate on RTT
  - dependence of bit-rate on packet size
- marking algorithms
  - scaling congestion signals
  - combining congestion marks – multi-bottleneck paths
  - marking across Diffserv classes – independent vs interdependent
- multicast

short flows
TCP inadequate – can economics help?



- Above model from [Cardwell00]
- [Key99] derives flow start behaviour as strategy a sender would adopt if subject to congestion pricing – exponential – very similar to TCP slow-start

congestion volume
captures (un)fairness during dynamics



re-ECN flow bootstrap

- at least one **green** packet(s) at start of flow or after >1sec idle
  - means “feedback not established”
  - ‘credit’ for safety due to lack of feedback
  - a **green** byte is ‘worth’ same as a **black** byte
- a different colour from black
  - distinguishes expected congestion based on experience from based on conservatism
  - gives deterministic flow state mgmt (policers, droppers, firewalls, servers)
  - rate limiting of state set-up
  - congestion control of memory exhaustion

- **green** also serves as state setup bit [Clark, Handley & Greenhalgh]
  - protocol-independent identification of flow state set-up
  - for servers, firewalls, tag switching, etc
  - don’t create state if not set
  - may drop packet if not set but matching state not found
  - firewalls can permit protocol evolution without knowing semantics
  - some validation of encrypted traffic, independent of transport
  - can limit outgoing rate of state setup
- to be precise **green** is ‘idempotent soft-state set-up codepoint’

weighted congestion controls

- important to enable *w<1*, negates weight inflation
- new app parameter overloading socket API
  - will require app & policy integration
- existing cc's where TCP-friendliness doesn't apply:

  <http://trac.tools.ietf.org/group/irtf/trac/wiki/CapacitySharingArch#CongestionControlsforwhichTCP-FriendlinessDoesntApply>
  - IETF activities
    - Low extra delay background transport (LEDBAT)
    - Pre-congestion notification (PCN)
    - Pseudowire Congestion Control Framework
    - multipath TCP (MPTCP)
  - Research implementations & proposals
    - Relentless Congestion Control
    - Weighted Window-based Congestion Control [Siris02]
    - mulTFRC [Damjan09]
    - mulTCP [Crowcroft98]

multipath transports

- congestion accountability
    - naturally works for multipath
    - volume of congested bytes crossing trust boundary
    - irrespective of how many or which flows they are in
- whole MPTCP bundle currently TCP-friendly
    - to comply with current IETF process
        - – until consensus reached on new non-TCP-friendly principles
    - MPTCP could be weighted
    - as any cc could (see weighted congestion control)

dependence of bit-rate on RTT?

- dependence on RTT arises from packet conservation
    - basis of TCP design
    - ACK clocking very powerful for robust implementation
    - but fallacy to say packet conservation is a *principle*…
- control theorists [Vinnicombe, Low] have proved
    - acceleration needs to depend on 1/RTT
    - but steady-state rate does not
- implementations:
    - FAST TCP [Jin04]
    - Kelly's primal unipath algorithm [Siris02]

$$\frac{d}{dt} x_r(t) = \frac{\kappa}{T}\left(w_r - x_r(t)\, p_r(t)\right) \qquad \text{in steady state } \bar{x}_r = \frac{w_r}{p_r}$$

$x_r(t)$: bit-rate, κ: gain constant,
T: round trip time, $p_r(t)$: path congestion

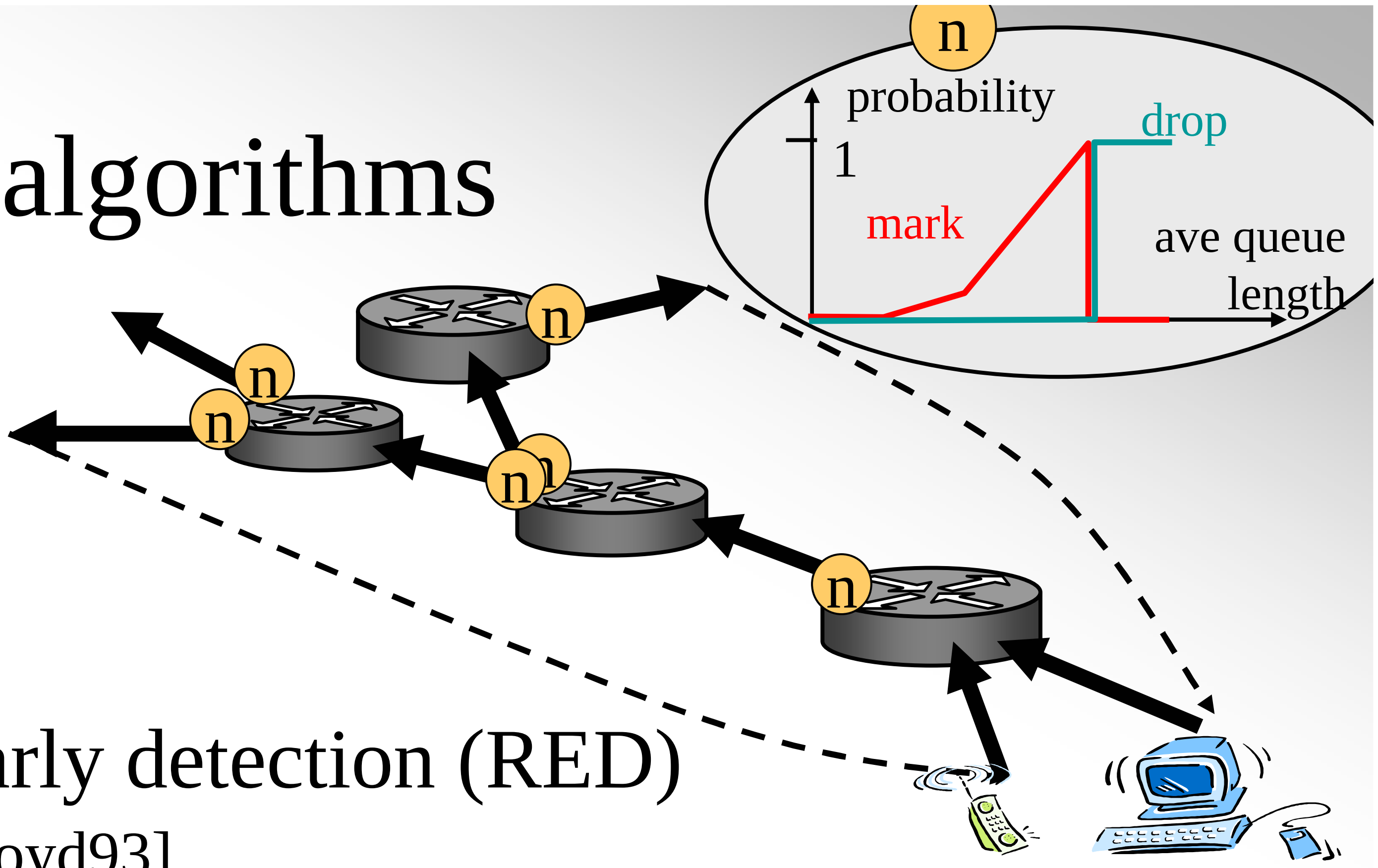
(independent of T)

dependence of bit-rate on packet size?

- TCP controls no. of packets in flight (window)
    - larger packets give faster bit-rate
    - ACK clocking makes for robust implementations
    - but another fallacy, not a principle…
- tempting to reduce drop for small packets
    - drops less control packets, which tend to be small
        - SYNs, ACKs, DNS, SIP, HTTP GET etc
    - but small != control
    - favouring smallness encourages smallness, not ‘controlness’
        - malice: small packet DoS
        - innocent experimentation: “Hey, smaller packets go faster” OS tweaks, application evolution
- AQM in network SHOULD NOT give smaller packets preferential treatment
    - opens DoS vulnerability
- adjust for byte-size when transport reads NOT when network writes congestion notifications

- lots of details, see [byte-pkt]

marking algorithms

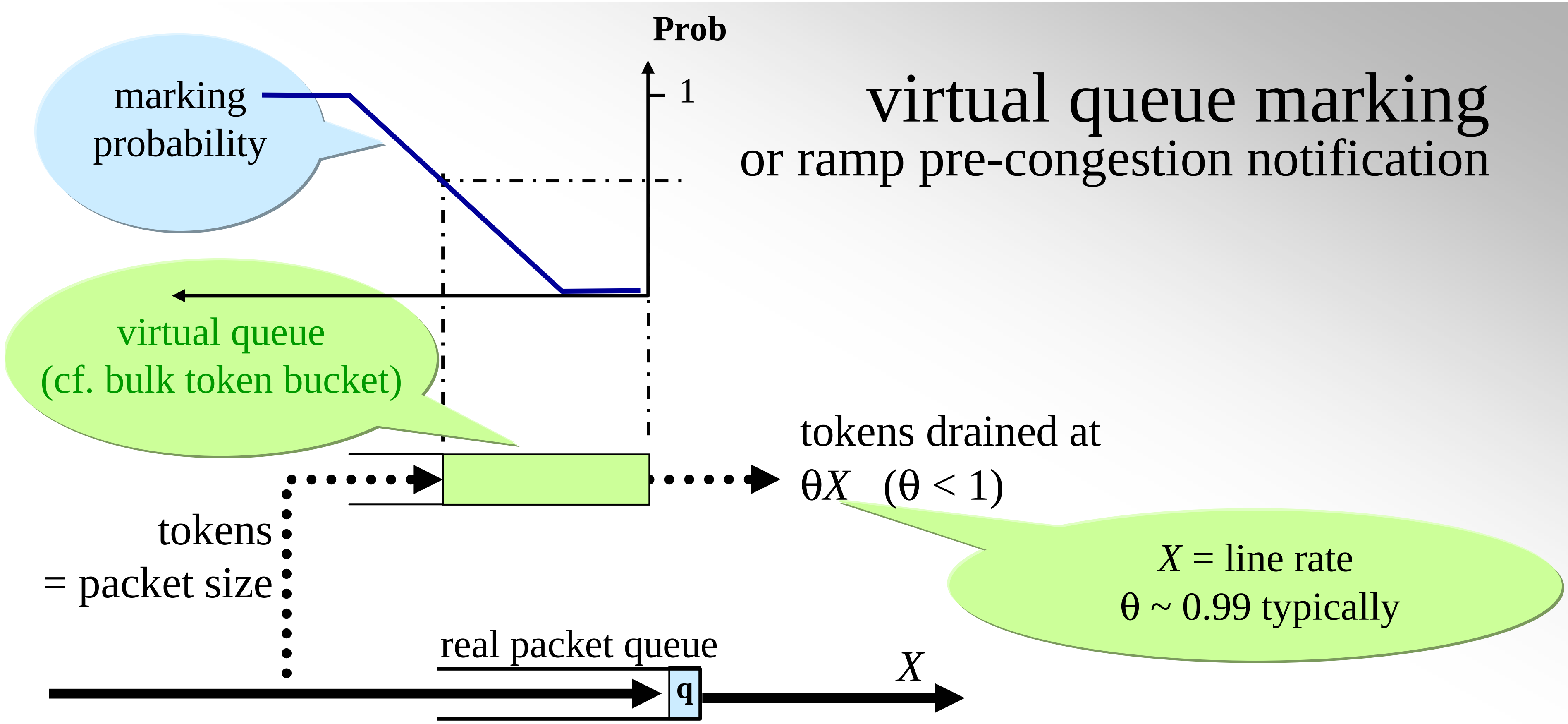


- random early detection (RED)
  - see [Floyd93]
- diagram shows gentle RED
  - queue length smoothed through EWMA
  - RED sensitive to parameter settings
  - still active area of research

scaling congestion signals
1/p congestion controls (e.g. Relentless CC)

- TCP's $W \propto {}^{1}/_{\sqrt{p}}$ window doesn't scale
  - congestion signals /window reduce as speed grows, $O({}^{1}/_{W})$
  - root cause of TCP taking hours / saw tooth at hi-speed
- $W \propto {}^{1}/_{p}$ scales congestion signals / window $O(1)$
  - Relentless, Kelly's primal algorithm
  - IOW, get same no of losses per window whatever the rate
- an alternative way of getting more precise congestion signals than more bits per packet

virtual queue marking
or ramp pre-congestion notification



- virtual queue (a conceptual queue – actually a simple counter):
  - drained somewhat slower than the line rate
  - build up of virtual queue is ‘early warning’ that traffic is getting close to capacity
  - mean number of packets in real queue, q, is kept very small by closed loop congestion control based on marks from virtual queue

combining congestion marks – costs

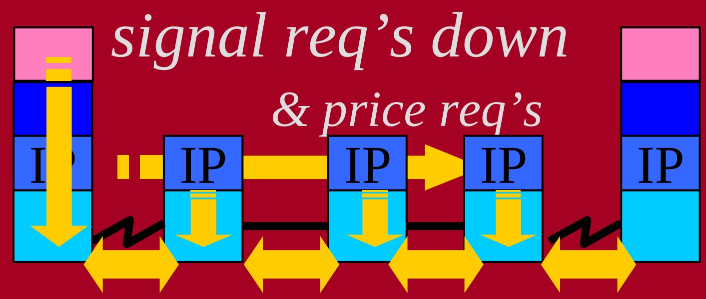
- up layers
  - congestion info must rise up the layers (even beyond transport)
    - unlike most header fields where requests pass down the layers
  - all congestion starts as a physical phenomenon
  - where higher layer takes over from lower
    - convert specific link congestion metric to forward it
- across layers
  - multiple congested bottlenecks on path
  - optimisation maths is based on linearly adding them
  - can use combinatorial probability, either approximately or directly

    $p = 1 - (1-p_1)(1-p_2)...$

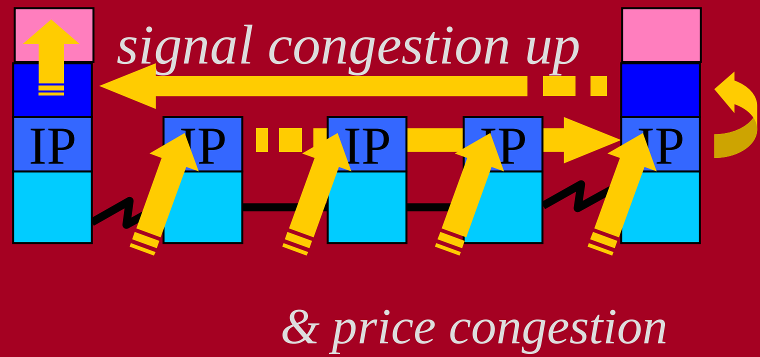
    $\approx p_1 + p_2 + ... \qquad\qquad if\ p << 1$
  - can define marking algo curve as exponential, so probabilistic addition becomes exact addition [REM]

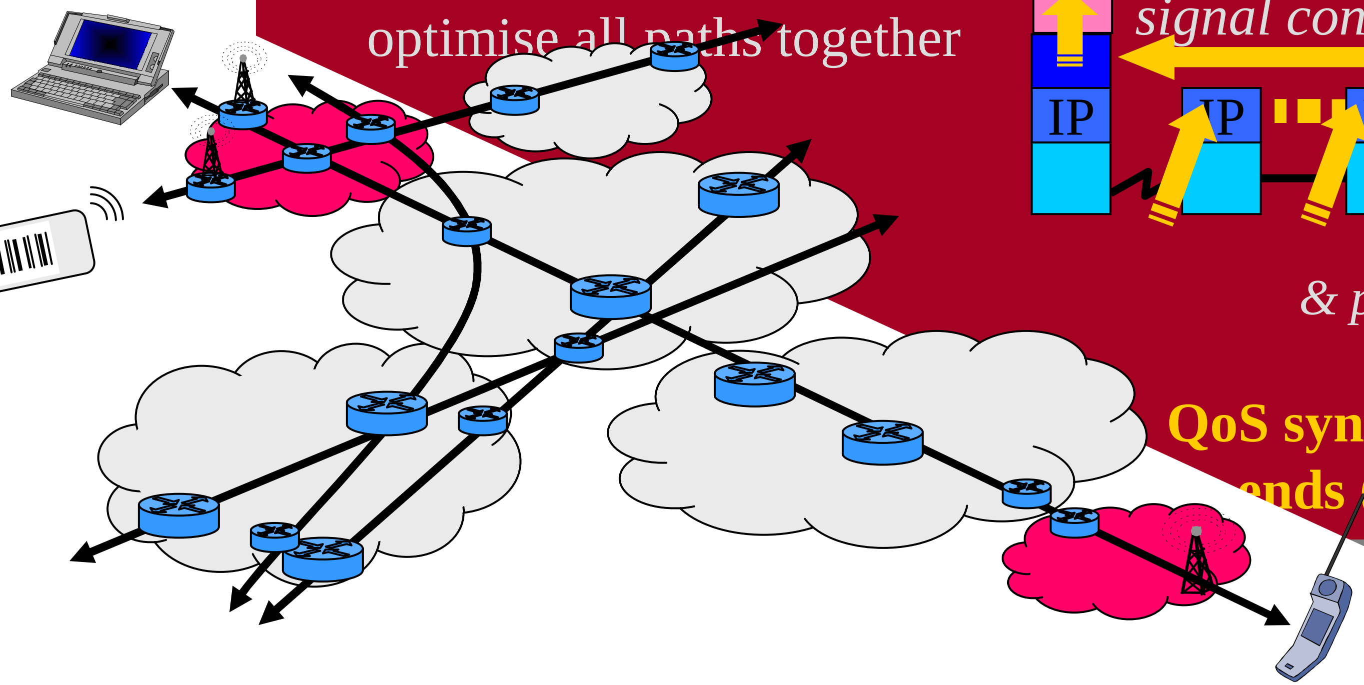
layered congestion notification (e2e principle)

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



☺ **new:**
optimise all paths together



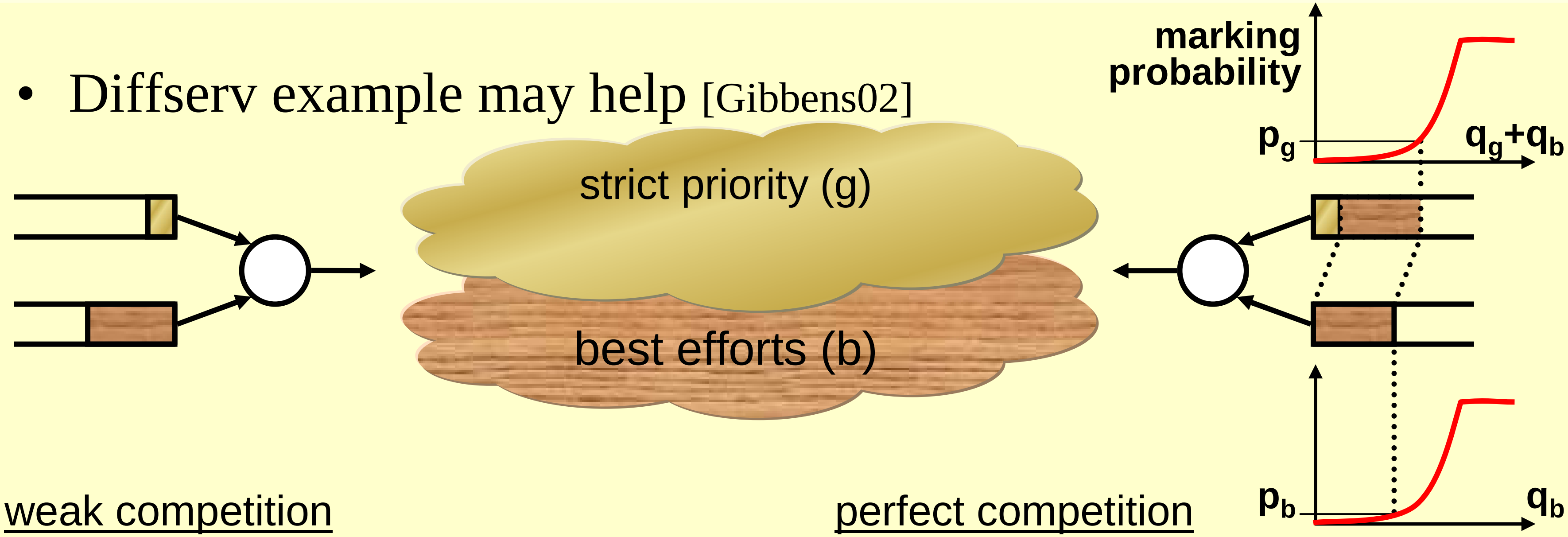


QoS synthesised by the ends (closed-loop)

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]

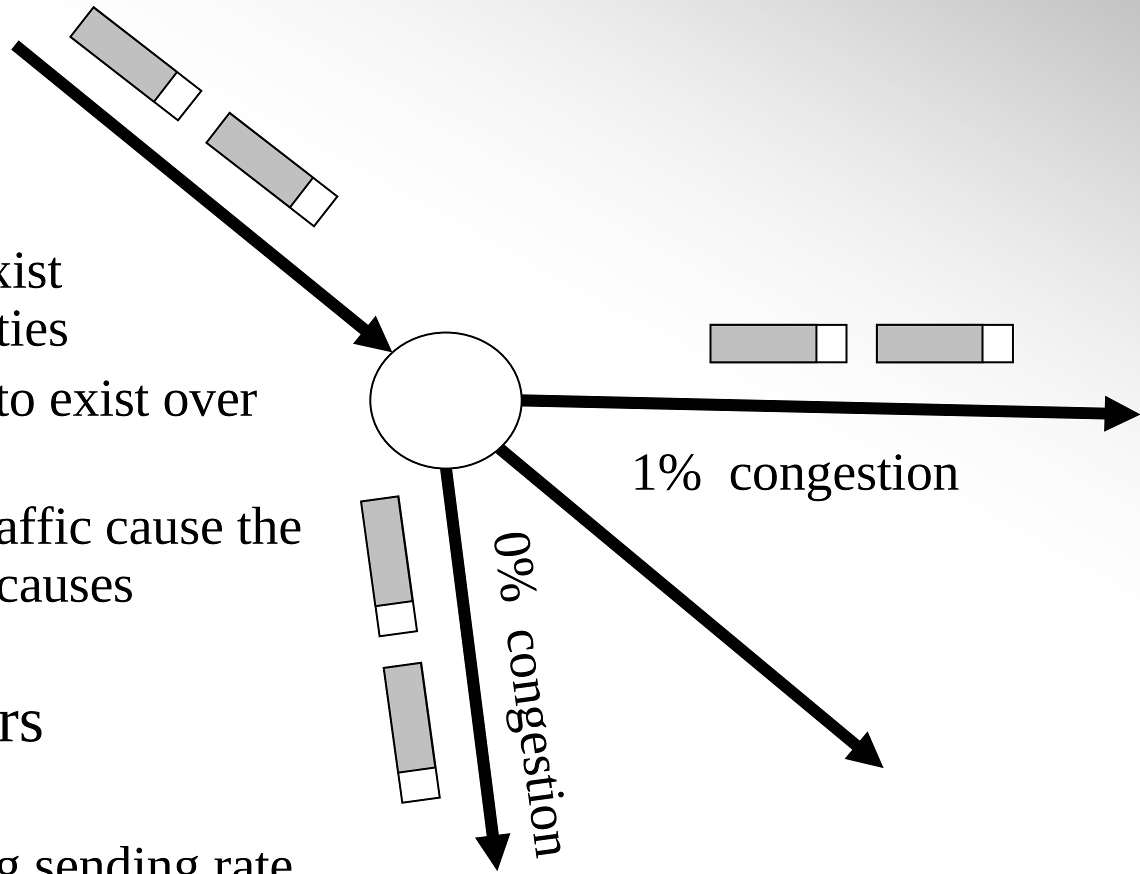


<u>weak competition</u>
price of *expectation* of better service
arbitrarily higher $p_g >> p_b$

<u>perfect competition</u>
price differential ∝ cost differential
$p_g \geq p_b$

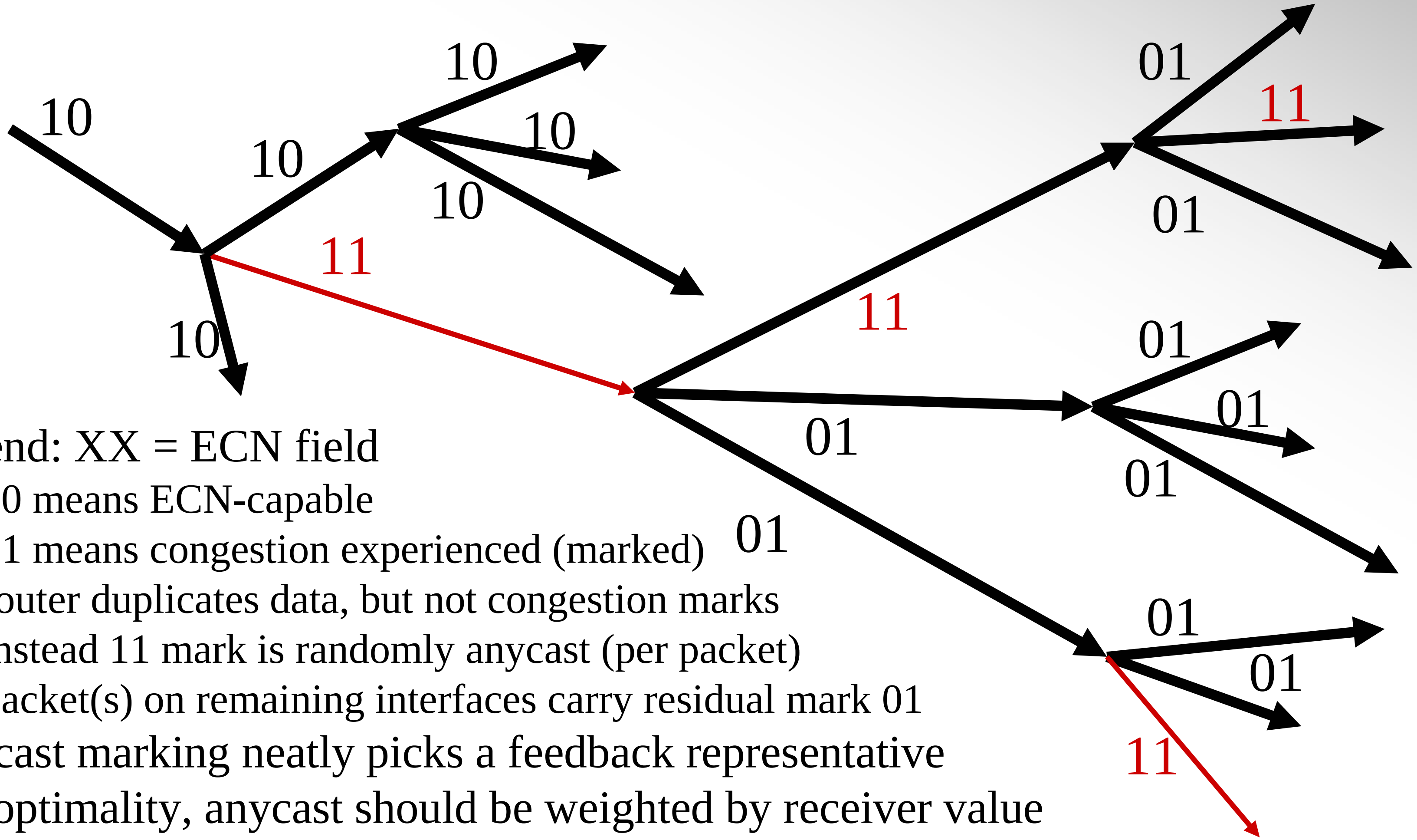
multicast congestion cost causation?

- strictly
  - operator causes packet duplication service to exist and chooses link capacities
  - receivers cause session to exist over link
  - sender & background traffic cause the traffic rate that directly causes congestion
- easier to make receivers responsible for costs
  - but receivers not causing sending rate, only existence of *some* traffic
  - to remove cost, need all downstream receivers to leave, but each has little incentive given cost should be shared



multicast & congestion notification

antidote to arbitrary 'research' on fairness between unicast & multicast

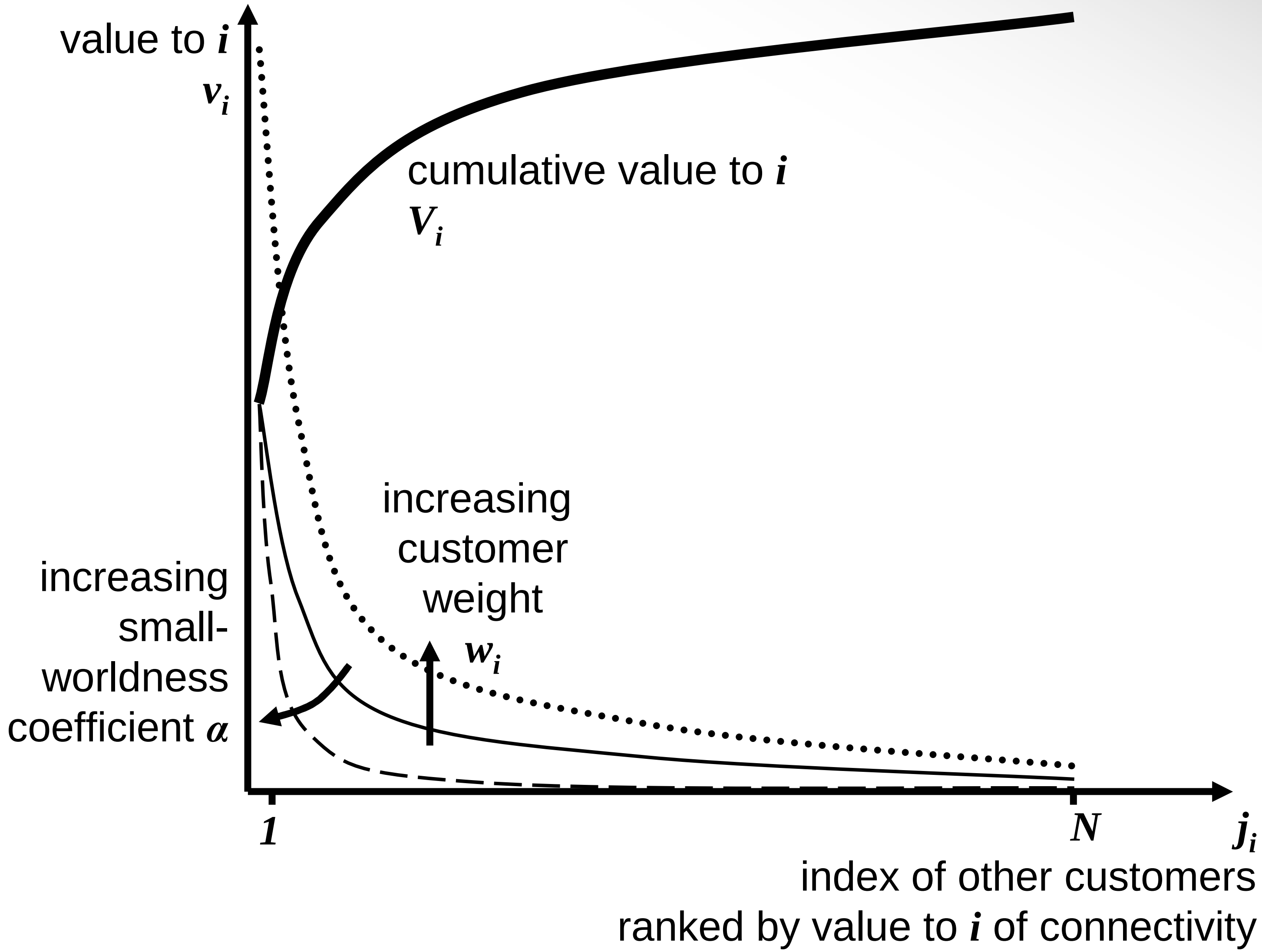


- legend: XX = ECN field
  - 10 means ECN-capable
  - 11 means congestion experienced (marked)
  - router duplicates data, but not congestion marks
  - instead 11 mark is randomly anycast (per packet)
  - packet(s) on remaining interfaces carry residual mark 01
- anycast marking neatly picks a feedback representative
- for optimality, anycast should be weighted by receiver value
- none of this is easy to implement
- can't do any of this with drop

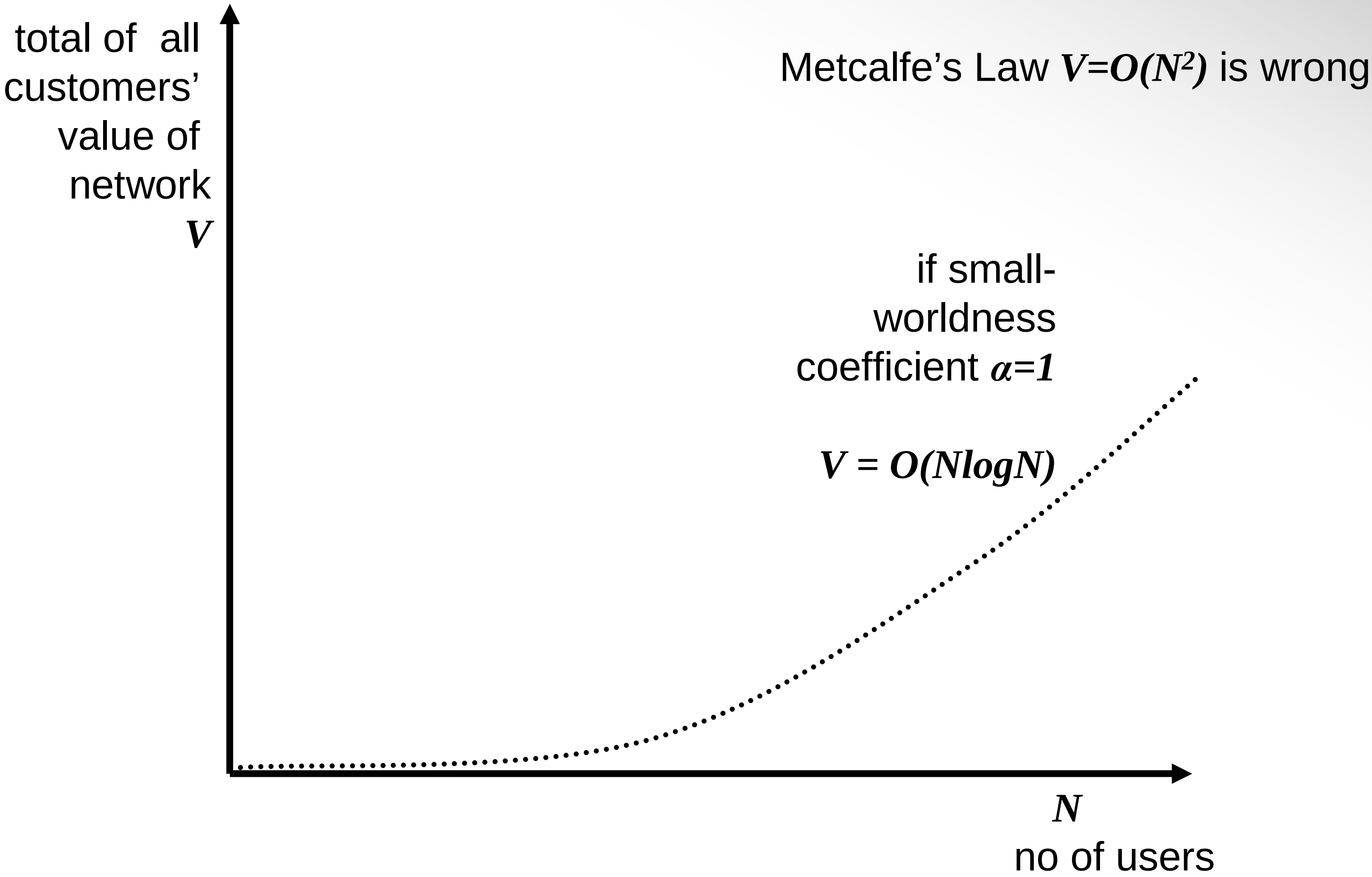
value of connectivity

(BGP tries to conflate this with *cost* of usage)

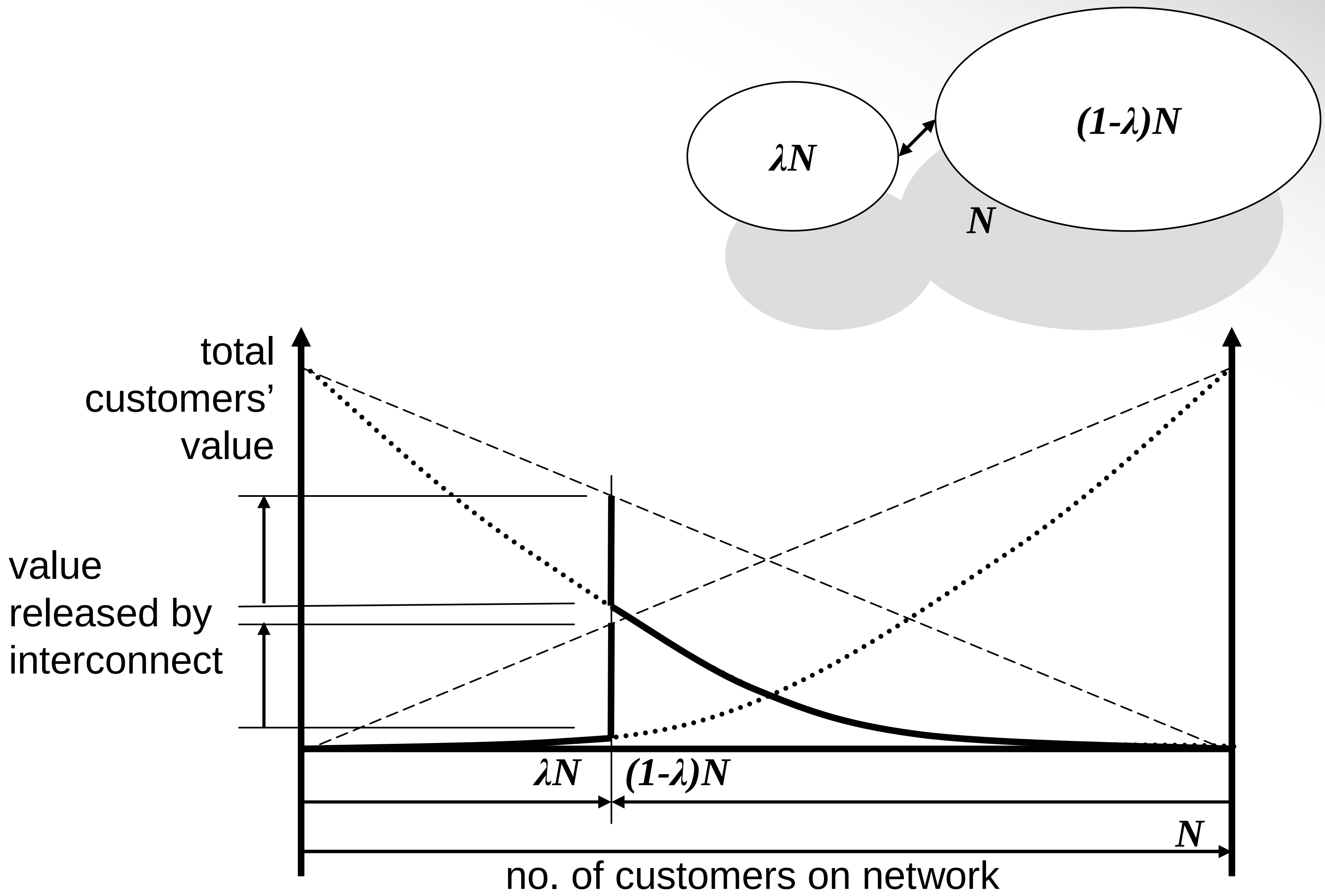
value to *i* of other customers



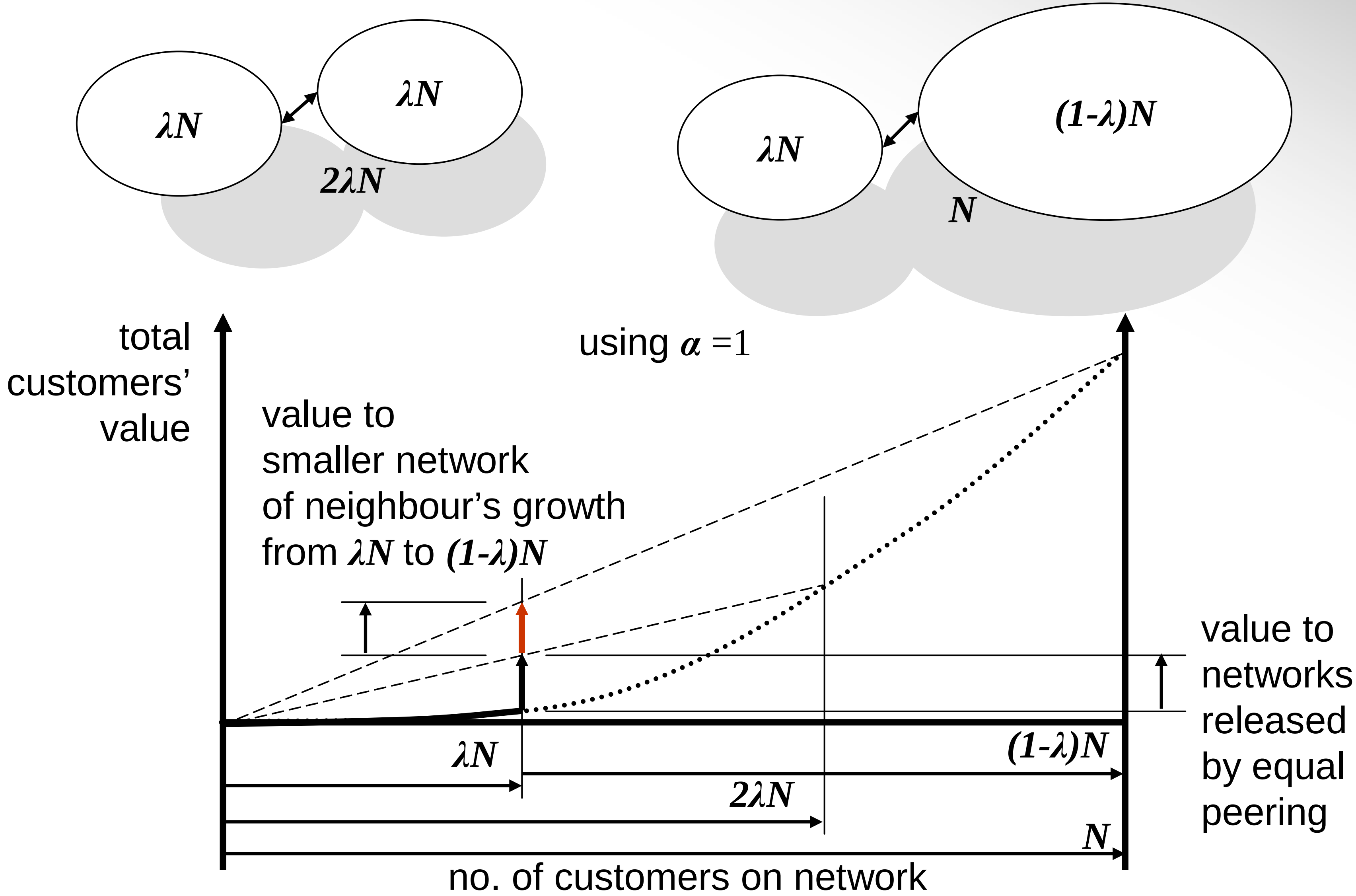
how the value of a network scales
with no. & weight of users



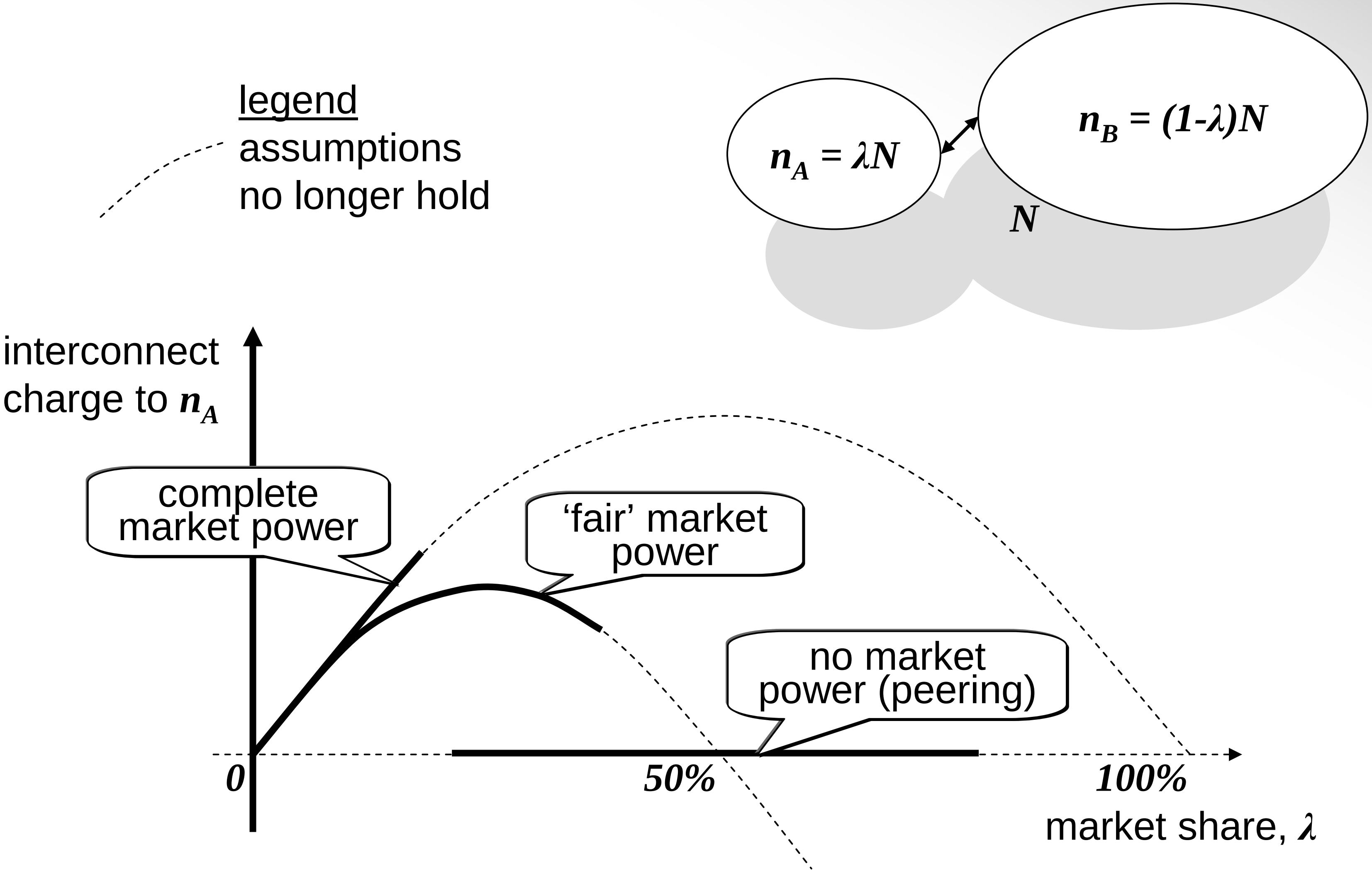
growth in network value
by scaling & interconnect



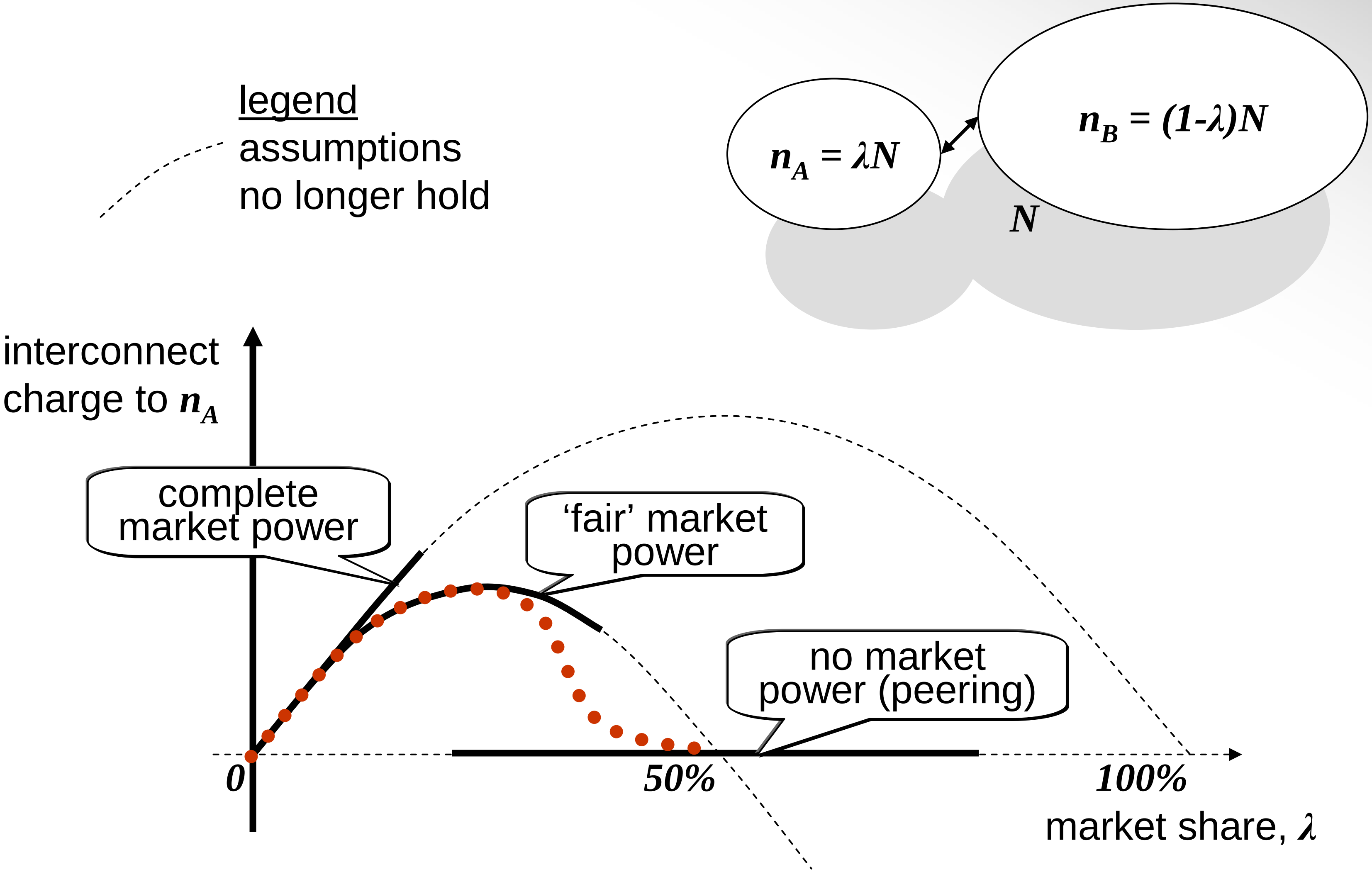
interconnect settlement



charging for interconnect
within the same market



charging for interconnect
within the same market



more info...

- The whole story in 7 pages
  - Bob Briscoe, "Internet: Fairer is Faster", BT White Paper TR-CXR9-2009-001 (May 2009)

  the following abridged article was based on the the above white paper
  - Bob Briscoe, "A Fairer, Faster Internet Protocol", IEEE Spectrum (Dec 2008)
- Inevitability of policing
  - [BBincent06] 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>
- Stats on p2p usage across 7 Japanese ISPs with high FTTH penetration
  - [Cho06] Kenjiro Cho et al, "The Impact and Implications of the Growth in Residential User-to-User Traffic", In Proc ACM SIGCOMM (Oct '06)
- 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
  - [Briscoe08] 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
  - [Briscoe05] Bob Briscoe and Steve Rudkin "Commercial Models for IP Quality of Service Interconnect" BT Technology Journal 23 (2) pp. 171--195 (April, 2005)
- Growth in value of a network with size
  - [Briscoe06] Bob Briscoe, Andrew Odlyzko & Ben Tilly, "Metcalfe's Law is Wrong", IEEE Spectrum, Jul 2006
- Re-architecting the Future Internet:
  - The Trilogy project
- Re-ECN & re-feedback project page:

  [re-ECN] http://bobbriscoe.net/projects/refb/
- These slides

  <bobbriscoe.net/present.html>

more info on 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) <bobbriscoe.net/pubs.html#ixqos>
- IETF PCN working group documents <tools.ietf.org/wg/pcn/> in particular:

  [PCN] Phil Eardley (Ed), *Pre-Congestion Notification Architecture*, RFC5559 (2009)

  [re-PCN] Bob Briscoe, *Emulating Border Flow Policing using Re-PCN on Bulk Data*, Internet Draft <bobbriscoe.net/pubs.html#repcn> (Sep'08)
- These slides

  <bobbriscoe.net/present.html>

references I – mostly economics

- [Clark05] David D Clark, John Wroclawski, Karen Sollins and Bob Braden, "Tussle in Cyberspace: Defining Tomorrow's Internet," IEEE/ACM Transactions on Networking (ToN) 13(3) 462–475 (June 2005) <portal.acm.org/citation.cfm?id=1074049>
- [MacKieVarian95] MacKie-Mason, J. and H. Varian, “Pricing Congestible Network Resources,” IEEE Journal on Selected Areas in Communications, `Advances in the Fundamentals of Networking' 13(7)1141--1149, 1995 http://www.sims.berkeley.edu/~hal/Papers/pricing-congestible.pdf
- [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])
- [Gibbens02] Richard J. Gibbens and Frank P. Kelly. On packet marking at priority queues. *IEEE Transactions on Automatic Control*, 47(6):1016–1020, June 2002.
- [Akerlof70] G.A. Akerlof. The market for ‘lemons’: Quality, uncertainty and market mechanisms. *Quarterly Journal of Economics*, 84:488–500, August 1970.
- [Laskowski06] Paul Laskowski and John Chuang. Network monitors and contracting systems: Competition and innovation. *Proc. ACM SIGCOMM’06, Computer Communication Review*, 36(4):183–194, September 2006.
- [Constantiou01] Ioanna D. Constantiou and Costas A. Courcoubetis. Information asymmetry models in the Internet connectivity market. In *Proc. 4th Internet Economics Workshop*, May 2001.
- [PCN] Phil Eardley (Ed), “Pre-Congestion Notification Architecture,” IETF Internet Draft draft-ietf-pcn-architecture-02.txt (Nov ’07)
- [Siris] Future Wireless Network Architecture <www.ics.forth.gr/netlab/wireless.html>
- Market Managed Multi-service Internet consortium <www.m3i_project.org/>

references II – mostly technical

- [Cardwell00] N. Cardwell, S. Savage and T. Anderson, "Modeling TCP latency,“ In Proceedings of IEEE INFOCOM, Tel Aviv, Israel, March 2000. http://citeseer.ist.psu.edu/cardwell00modeling.html
- [Key99] Peter Key and Laurent Massouli´e. User policies in a network implementing congestion pricing. In *Proc. Workshop on Internet Service Quality and Economics*. MIT, December 1999.
- [Key04] Key, P., Massoulié, L., Bain, A., and F. Kelly, “Fair Internet traffic integration: network flow models and analysis,” Annales des Télécommunications 59 pp1338--1352, 2004 http://citeseer.ist.psu.edu/641158.html
- [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)
- [REM] S. Athuraliya, V.H. Li, Steven H. Low and Qinghe Yin, “REM: Active Queue Management,” IEEE Network 15 (3) pp. 48--53 (May/June, 2001)
- [Jin04] Cheng Jin, David Wei, and Steven Low. FAST TCP: Motivation, architecture, algorithms, performance. In *Proc. IEEE Conference on Computer Communications (Infocomm’04)*. IEEE, March 2004.
- [Siris02] Vasilios A. Siris, Costas Courcoubetis, and George Margetis. Service differentiation and performance of weighted window-based congestion control and packet marking algorithms in ECN networks. *Computer Communications*, 26(4):314–326, 2002.
  - implementations in FreeBSD & ns2: <http://tracer.ucnet.uoc.gr/wtp/>
- [byte-pkt] Bob Briscoe, “Byte and Packet Congestion Notification” IETF Internet Draft draft-ietf-tsvwg-byte-pkt-mark-00.txt (Aug 2008)
- [Argyraki07] Katerina Argyraki, Petros Maniatis, Olga Irzak, Subramanian Ashish, and Scott Shenker. Loss and delay accountability for the internet. In *Proc. IEEE International Conference on Network Protocols*. IEEE, October 2007.
- [Katabi04] Dina Katabi, XCP Performance in the Presence of Malicious Flows In Proc PFLDnet’04 (Feb 2004)