

flow rate fairness dismantling a religion

[<draft-briscoe-tsvarea-fair-01.pdf>](#)

status:	individual draft
final intent:	informational
intent next:	tsvwg WG item after (or at) next draft

Bob Briscoe
Chief Researcher, BT Group
IETF-68 tsvwg Mar 2007



updated 00⇒01 draft

- diffs and alt formats (courtesy of rfcdiff & xml2rfc tools) at: <http://www.cs.ucl.ac.uk/staff/B.Briscoe/pubs.html#rateFairDis>
- lots of (on & off list) email comments from presenting at IETF-67 tsv-area, IRTF iccrg & e2erg
- changes from previous draft-00 (⇒ = “focused on in this talk”)
 - Toned down the polemic, but some still think it’s too hot for a WG item
 - ⇒ Added importance of the issue and implications (§1 Introduction)
 - Added §8 "Critiques of Specific Schemes":
 - pulls together critiques of: max-min, TCP, TFRC & router-based fairness (e.g. XCP)
 - ⇒ added material on fairness wrt RTT, packet size and WFQ
 - Clarified how to calibrate the cost of congestion from equipment costs (in §5.2)
 - Clarified §5.3.2 "Enforcing Cost Fairness"
 - unofficial BoF Wed 15:10-16:40 Karlin I
 - ⇒ Added substantial new §9 "Implications for the RFC Series"

importance and implications (§1)

- if we do nothing about fairness
 - the few are ruining it for the many
 - so, keeping interactive apps viable requires massive capacity
 - or poor incentives to invest in capacity
 - so, operators are kludging it with DPI (deep packet inspection)
 - so, today's apps are getting frozen into the Internet
 - and we're getting complex, ugly feature interactions

recap

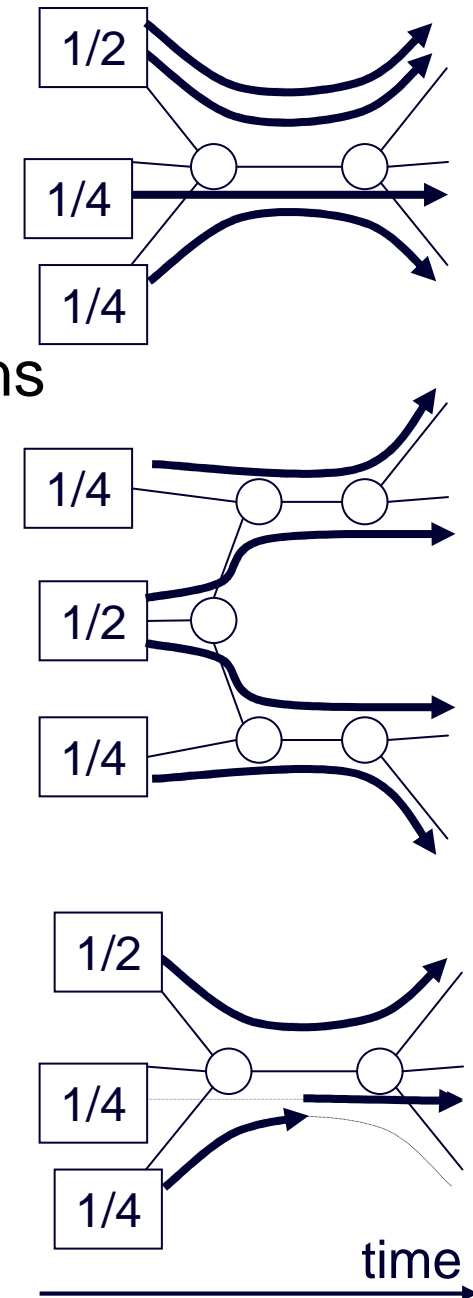
dismantling flow rate fairness

- doesn't even address relevant questions

- 1) how many flows is it fair for an app to create?
- 2) how fast should flows go through separate bottlenecks?
- 3) how fast should a brief flow go compared to a longer lasting one?

- myopic

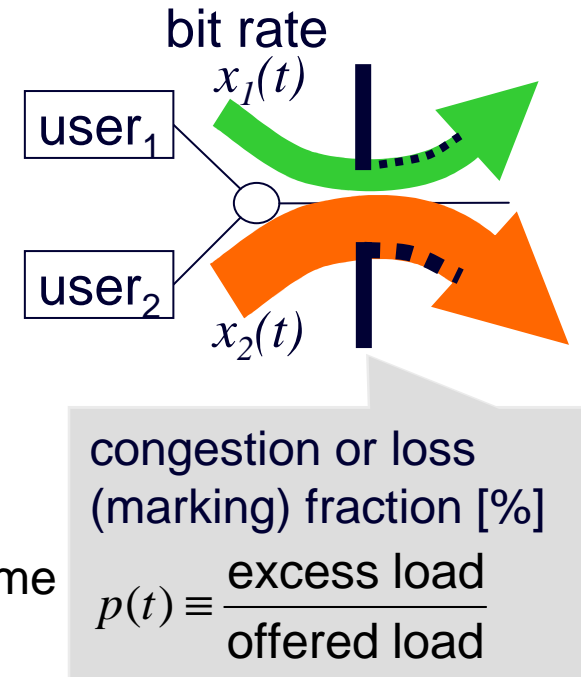
- across flows, across network and across time



recap

replace with cost fairness

- cost of your behaviour on others
 - bytes you contributed to excess load
 - termed congestion volume [bytes]
 - accumulates simply and correctly
 - across flows, across network paths and across time
- not your bit rate x_i
- but loss (marking) fraction times your bit rate px_i



pls add this rule to your buzzword matching algorithms

cost fairness \nleftrightarrow re-ECN

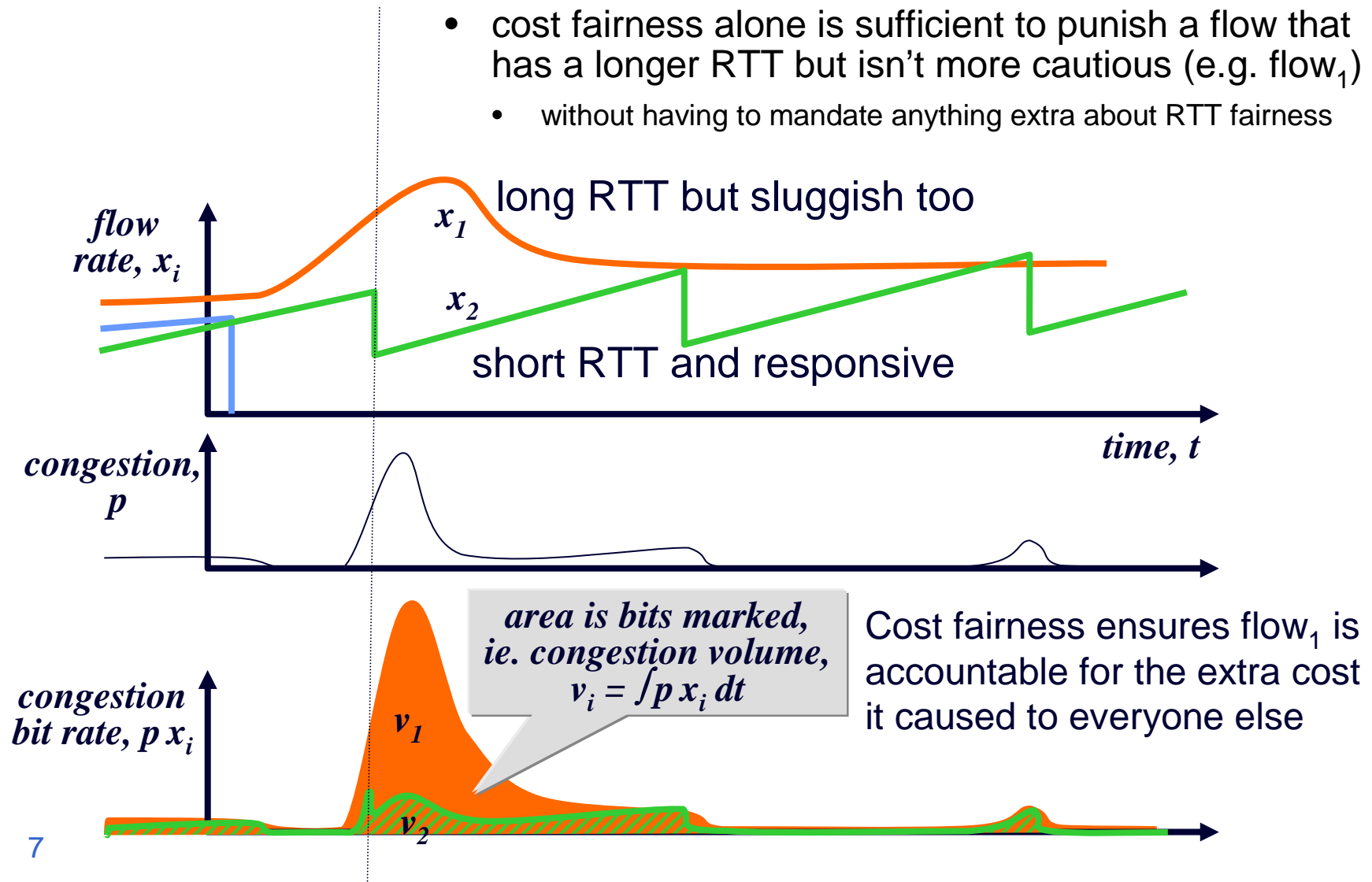
[draft-briscoe-tsvarea-fair-01.pdf](#)

[draft-briscoe-tsvwg-re-ecn-tcp-03.txt](#)

- re-ECN is not limited to enforcing cost fairness
 - re-ECN appendix shows how to police TCP (flow rate fairness)
 - fairness I-D shows how to do other forms of fairness with it
- cost fairness could be done with something else
 - but no other practical schemes (yet)

RTT fairness

- TCP's stable rate is proportional to $1/\text{RTT}$
 - doesn't need to be, but change of rate (dynamics) should be
 - FAST is an example of congestion control like this
- cost fairness alone is sufficient to punish a flow that has a longer RTT but isn't more cautious (e.g. flow₁)
 - without having to mandate anything extra about RTT fairness



implications of this draft for the RFC series

cc RFCs sorted by who must/should to do what

- cc algo as impl'n building block without saying where to use
 - 3448 TFRC
- spec of cc impl'n for a specific transport
 - 2581 TCPcc, 2960 SCTP, 4341 4342 DCCP CCIDs, 3551 RTP/AVP, 4585 RTP/AVPF
- hosts must impl't a specific transport
 - 1122 Host Reqs
- what hosts must do if they impl't a specific cc enhancement
 - 3124 Congestion Mgr
- spec semantics of cong'n notification impl'n
 - 2309 AQM, 3168 ECN
- apps must impl't safe cc behaviour
 - 2616 HTTP/1.1, 3550 RTP cc applicability
- best practice, guidelines & principles for cc design
 - 1254 cc survey, 2309 AQM, 2914 cc Principles, 3426 Arch considerations, 3714 Voice cc concerns
- recommend how new cc designs should interact with old
 - 2309 AQM, 2357 Criteria for RMT, 2914 cc Principles

acks: Michael Welzl & Wes Eddy draft-irtf-iccr-g-cc-rfcs-00, adding to Sally Floyd's RFC2914

implications of this draft for the RFC series if we add app/user policy-control over congestion control

- cc algo as impl'n building block without saying where to use
 - 3448 TFRC
 - spec of cc impl'n for a specific transport
 - 2581 TCPcc, 2960 SCTP, 4341 4342 DCCP CCIDs, 3551 RTP/AVP, 4585 RTP/AVPF
 - hosts must impl't a specific transport
 - 1122 Host Reqs
 - what hosts must do if they impl't a specific cc enhancement
 - 3124 Congestion Mgr
 - spec semantics of cong'n notification impl'n
 - 2309 AQM, 3168 ECN
 - apps must impl't safe cc behaviour
 - 2616 HTTP/1.1, 3550 RTP cc applicability
 - best practice, guidelines & principles for cc design
 - 1254 cc survey, 2309 AQM, 2914 cc Principles, 3426 Arch considerations, 3714 Voice cc concerns
 - recommend how new cc designs should interact with old
 - 2309 AQM, 2357 Criteria for RMT, 2914 cc Principles
- stand as they are, for apps that don't need or user policy ctrl
- OK. Must impl't means available for use, not must be used
- critical to cost fairness; OK, except tighten up open issues (byte-mode drop & ECN tunnels)
- All good sound general advice
- Generally sound advice, except definitions of fairness based on flow rate, and TCP-fair advice in 2357 needs qualifying

next steps

aim, fire, ready



1. make this fairness I-D suitable for WG item
3. need *to be able* to make senders accountable for congestion caused (e.g. with re-ECN)
4. need weighting parameter added to transport APIs (e.g. MulTCP)

?

2. transition from what we have now?

- we have no fairness, so there's nothing to transition from
- but there is a danger of getting it *more* unfair than we have already
- therefore should have step 3 largely in place before step 4
- hi-speed congestion ctrl in progress could be designed *as if* we have step 3
 - voluntary cost fairness (cf. voluntary TCP fairness)

flow rate fairness dismantling a religion

[<draft-briscoe-tsvarea-fair-01.pdf>](#)

spare slides:

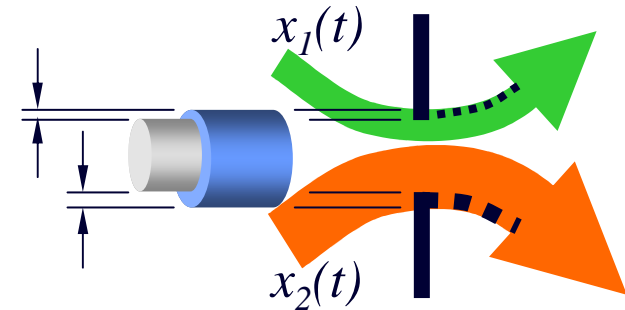
- calibrating congestion cost as equipment cost
- packet size fairness
- WFQ & cost fairness

Q&A



calibrating 'cost to other users'

- a monetary value can be put on 'what you unsuccessfully tried to get'
 - the marginal cost of upgrading network equipment
 - so it wouldn't have marked the volume it did
 - so your behaviour wouldn't have affected others
- competitive market matches...
 - the cost of congestion volume
 - with the cost of alleviating it



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

- congestion volume is not an extra cost
 - part of the flat charge we already pay
 - but we can't measure who to blame for what
 - if we could, we *might* see pricing like this...

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

- NOTE WELL
 - IETF provides the metric
 - industry does the business models

packet size fairness

- new I-D written but not posted
- intended as informational, through tsvwg WG?
- gives principles for handling different packet sizes
 - for any active queue mgmt (AQM) scheme, eg:
 - RED drop/marketing (open issue in RFC2309)
 - PCN (pre-congestion notification) marking (deliverable of newly chartered WG)
- in summary: answers two questions
 1. byte congestible or packet congestible resource?
 - RED should usually use byte-mode queue measurement
 2. if byte congestible, which layer should account for packet size?
 - transport not network
 - transport should respond to congestion volume in bytes, not packets
 - TFRC-SP (small packets) is the correct place to do this
 - RED byte mode drop considered harmful

weighted fair queuing (WFQ)

- WFQ typically allocates capacity per flow, not per user
 - vulnerable to flow splitting games described in draft
- controls fairness over flow lifetimes not over user history
 - but for high utilisation customer lines this approximates to the same thing
 - but not over all the congestion caused in the Internet – just one interface
- implications of WFQ not being cost fair
 - doesn't mean WFQ is 'incorrect'
 - just means WFQ can't ensure customers pay their rightful costs
 - a future competing solution that did might be preferred by operators

Bar BoF “re-ECN architectural intent”

Wed 21 March 1510-1640, Karlin I, Prague Hilton
background papers on re-ECN:

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

including particularly

<[draft-briscoe-tsvwg-re-ecn-tcp-03.txt](#)>