Flow Rate Fairness: Dismantling a Religion

Bob Briscoe
BT Research & UCL
bob.briscoe@bt.com

ABSTRACT

Resource allocation and accountability keep reappearing on every list of requirements for the Internet architecture. The reason we never resolve these issues is a broken idea of what the problem is. The applied research and standards communities are using completely unrealistic and impractical fairness criteria. The resulting mechanisms don't even allocate the right thing and they don't allocate it between the right entities. We explain as bluntly as we can that thinking about fairness mechanisms like TCP in terms of sharing out flow rates has no intellectual heritage from any concept of fairness in philosophy or social science, or indeed real life. Comparing flow rates should never again be used for claims of fairness in production networks. Instead, we should judge fairness mechanisms on how they share out the 'cost' of each user's actions on others.

Categories and Subject Descriptors

K.4.1 [**Computers and Society**]: Public Policy Issues; C.2.1 [**Computer-communication networks**]: Network Architecture and Design

General Terms

Economics, Security

Keywords

Resource allocation, congestion control, fairness, accountability, identity

1. INTRODUCTION

"But he has nothing on at all."

The Emperor's New Clothes, Hans Christian Andersen

This paper is deliberately destructive. It sets out to destroy an ideology that is blocking progress—the idea that fairness between multiplexed packet traffic can be achieved by controlling relative flow rates alone. Flow rate fairness was the goal behind fair resource allocation in widely deployed protocols like weighted fair queuing (WFQ), TCP congestion control and TCP-friendly rate control [8, 1, 11]. But it is actually just unsubstantiated dogma to say that equal flow rates are fair. This is why resource allocation and accountability keep reappearing on every list of requirements for the Internet architecture (e.g. [2]), but never get solved. Obscured by this broken idea, we wouldn't know a good solution from a bad one.

Controlling relative flow rates *alone* is a completely impractical way of going about the problem. To be realistic for large-scale Internet deployment, relative flow rates should be the *outcome* of another fairness mechanism, not the mechanism itself. That other mechanism should share out the 'cost' of one user's actions on others—how much each user's transfers restrict other transfers, given capacity constraints. Then flow rates will depend on a deeper level of fairness that has so far remained unnamed in the literature, but is best termed 'cost fairness'.

It really is only the idea of flow rate fairness that needs destroying—nearly everything we've engineered can remain. The Internet architecture needs some minor additions, but otherwise it is largely already suited to cost fairness.

The metric required to arbitrate cost fairness is simply volume of congestion, that is congestion times the bit rate of each user causing it, taken over time. In engineering terms, for each user it can be measured very easily as the amount of data the user sent that was dropped. Or with explicit congestion notification (ECN [30]) the amount of each user's data to have been congestion marked. Importantly, unlike flow rates, this metric integrates easily and correctly across different flows on different paths and across time.

What we call cost fairness has been in the literature for nearly a decade, but it hasn't been put so bluntly before. We were moved to spell it out unambiguously (and avoiding maths), because this isn't just some dry academic fairness debate that might change allocation percentages somewhere in the third decimal place. The outcomes due to flow rate fairness that we see on the Internet today are hopelessly unlike the outcomes that would result from cost fairness.

Not that the outcomes we see are the deliberate intent of flow rate fairness. They are the random result of an absence of fairness control, because flow rate fairness isn't even capable of reasoning about questions like, "How many flows is it fair to start between two endpoints?" or, "What rate is fair for a flow that has been running longer than another?".

While everyone prevaricates, novel p2p applications have started to thoroughly exploit this architectural vacuum with no guilt or shame, by just running more flows for longer (after all, they are using TCP, which is fair isn't it?). In response some ISPs are deploying kludges like volume caps or throttling specific applications using deep packet inspection. Innocent experimental probing has turned into an arms race. The p2p community's early concern for the good of the Internet is being set aside, aided and abetted by commercial concerns, in pursuit of a more pressing battle against the ISPs that are fighting back. Bystanders sharing the same capacity are suffering heavy collateral damage.

This trend has spread beyond the p2p community. There is now no shame in opening multiple TCP connections, or offering VoIP or video streaming software without any congestion control.

Whether the prevailing notion of flow rate fairness has been the root cause or not, there will certainly be no solution until the networking community gets its head out of the sand and understands how unrealistic its view is, and how important this issue is. Certainly fairness is not a question of technical function—any allocation 'works'. But getting it hopelessly wrong badly skews the outcome of conflicts between the vested interests of real businesses and real people.

But isn't it a basic article of faith that multiple views of fairness should be able to co-exist, the choice depending on policy? Absolutely correct—and we shall return to how this can be done later. But that doesn't mean we have to give the time of day to any random idea of fairness.

Fair allocation of rates between flows isn't based on any respected definition of fairness from philosophy or the social sciences. It has just gradually become the way things are done in networking. But it's actually self-referential dogma. Or put more bluntly, bonkers.

We expect to be fair to people, groups of people, institutions, companies—things the security community would call 'principals'. But a flow is merely an information transfer between two applications. Where does the argument come from that information transfers should have equal rights? It's equivalent to claiming food rations are fair because the boxes are all the same size, irrespective of how many boxes each person gets or how often they get them.

Because flows don't deserve rights in real life, it is not surprising that two loopholes the size of barn doors appear when trying to allocate rate fairly to flows in a non-co-operative environment. If at every instant a resource is shared among the flows competing for a share, any real-world entity can gain by i) creating more flows than anyone else, and ii) keeping them going longer than anyone else.

We appeal to the networking community to quietly set aside rate fairness between flows. It had its time, but now it has been shown to be unfounded, unrealistic and impractical. Papers and standards proposals using it should be rejected. No-one should ever have to cater for it in future Internet protocols—it places arbitrary requirements on the system that can't be met and wouldn't achieve any meaningful sort of fairness even if they could be met.

Alternatively, someone should write a defence of flow rate fairness. Continuing to use flow rate fairness as the dominant ideology, without rebutting Kelly's seminal 1997 paper that undermined it, just leaves the Internet community divided into religious sects, making a mockery of the scientific process towards consensus.

2. FAIR ALLOCATION OF WHAT AMONG WHAT?

The issue with flow rate fairness is far more basic than whether allocations should be max-min, proportional or whatever. Flow rate fairness doesn't even allocate the correct thing. And it doesn't allocate it among the correct entities either. At this most basic level we will contrast the two main contending views:

- Allocate rate among flows (flow rate fairness)
- Allocate congestion among the bits sent by economic entities (cost fairness)

When cost fairness was proposed, it stated its case in terms of the dominant belief system—flow rate fairness. Unfortunately, this meant that the dominant belief system didn't notice it had been struck an intellectual death blow. Its believers carried on regardless and it remains dominant today.

As a result, one sees talk of weighted proportional fairness in the same context as proportional, max-min or min-max fairness as if they are all members of the same set. They are not. Weighted proportional fairness has an extra weight parameter w that the others lack. With weighted proportional fairness, the interesting bit is what regulates users in their choice of w. Otherwise, it would hardly be a useful definition of fairness to say it is fair for flow A to go w times as fast as flow B, if the user behind A has free choice of w.

In fact, it turns out that the interesting bit is nothing to do with flows, or their weights. The *only* worthwhile definition of fairness depends on the allocation of cost among the bits sent by economic entities, regardless of which flows the bits are in. A user's choice of w then depends on that.

2.1 Structure of Paper

The body of this paper is structured around our question, "Fair allocation of what among what?":

- §3 answers the "...of what...?" question, explaining why fair allocation of costs is a sufficient and realistic form of fairness, and allocation of rate is not.
- §4 answers the "...among what?" question, explaining why fairness among flows can only be myopic whereas fairness among economic entities naturally spans all flows and spans time. Also we show fairness among flows is hard, if not impossible, to enforce while enforcing fairness among economic entities seems practical.

Having debunked the dominant ideology of flow rate fairness, and replaced it with cost fairness, in §5 we discuss how other forms of fairness can be asserted locally. Then, before we draw conclusions, §6 maps the progression of seminal ideas in the literature on which this paper is based and §7 outlines concrete criticisms of specific fairness schemes: max-min flow rate fairness, TCP, TFRC and WFQ.

3. COST, NOT BENEFIT

The issues of fair allocation of resources comes under the domain of political economy, with philosophy reasoning about our judgements.[1] In §5 we will discuss how different fairness policies can co-exist. But to answer our question, "Fair allocation of what?" we start from the premise used in economics (and life) that fairness concerns comparing benefits, costs or both.

The benefit of a data transfer can be assumed to increase with flow rate, but the shape and size of the function relating the two (the utility function) is unknown, subjective and

[1]For instance Nozick argues that Rawls's influential 'Justice as Fairness' [31] defines fairness only as an end, whereas it is actually possible to ensure each transfer is fair even though overall allocations start out unfair and end unfair.

private to each user. Flow rate itself is an extremely inadequate measure for comparing benefits: user benefit per bit rate might be ten orders of magnitude different for different types of flow (e.g. SMS and video). So different applications might derive completely different benefits from equal flow rates and equal benefits might be derived from very different flow rates.

Turning to the cost of a data transfer across a network, flow rate alone is not the measure of that either. Cost is also dependent on the level of congestion on the path. This is counter-intuitive for some people so we shall explain a little further. Once a network has been provisioned at a certain size, it doesn't cost a network operator any more whether a user sends more data or not. But if the network becomes congested, each user restricts every other user, which can be interpreted as a cost *to all*—an externality in economic terms. For any level of congestion, Kelly showed [20] that the system is optimal if the blame for congestion is attributed among all the users causing it, in proportion to their bit rates. That's exactly what routers are designed to do anyway. During congestion, a queue randomly distributes the losses so all flows see about the same loss (or ECN marking) rate; if a flow has twice the bit rate of another it should see twice the losses. In this respect random early detection (RED [12]) is slightly fairer than drop tail, but to a first order approximation they both meet this criterion.

So in networking, the cost of one flow's behaviour depends on the congestion volume it causes which is the product of its instantaneous flow rate and congestion on its path, integrated over time. For instance, if two users are sending at 200kbps and 300kbps into a 450kbps line for 0.5s, congestion is $(200 + 300 - 450)/(200 + 300) = 10\%$ so the congestion volume each causes is $200\text{k} \times 10\% \times 0.5 = 10\text{kb}$ and 15kb respectively.

So cost depends not only on flow rate, but on congestion as well. Typically congestion might be in the fractions of a percent but it varies from zero to tens of percent. So, flow rate can never alone serve as a measure of cost.

To summarise so far, flow rate is a hopelessly incorrect proxy both for benefit and for cost. Even if the intent was to equalise benefits, equalising flow-rates wouldn't achieve it. Even if the intent was to equalise costs, equalising flow-rates wouldn't achieve it.

But actually a realistic resource allocation mechanism only needs to concern itself with costs. If we set aside political economy for a moment and use pure microeconomics, we can use a competitive market to arbitrate fairness, which handles the benefits side, as we shall now explain. Then once we have a feasible, scalable system that at least implements one defined form of fairness, we will show how to build other forms of fairness within that.

In life, as long as people cover the cost of their actions, it is generally considered fair enough. If one person enjoys a hot shower more than their neighbour enjoys the toast they made with equal units of electricity, no-one expects the one who enjoyed the shower to have to pay more. If someone makes more of their lot in life than another, some complain it's not fair, but most call this envy, not unfairness.

Market economics works on the same premise (unsurprisingly given life and market economics are closely related). In fact, a market adjusts supply to meet demand so that benefits are as fairly distributed as is consistent with the pre-existing inequalities in life. But this fairness between benefits is an *outcome* and it is only met as long as a market mechanism makes people accountable for the costs of their actions (and various market failures are avoided).

We deliberately say 'make people accountable' to avoid the phrase 'make people pay', because Internet users prefer flat rate subscriptions. So, ISPs want to be able to limit the congestion costs their users can cause (§4.3.2), not just charge them for whatever unlimited costs they cause.

If we do make users truly accountable for the cost of the congestion they cause, a form of fairness between flow rates emerges automatically. As everyone increases the rate of each of their flows, congestion rises. As congestion rises, everyone pays due regard to the share of the cost attributed to them. So, each individual will want their congestion control algorithm to continuously adjust its rate to maximise their net utility—benefit minus cost. Kelly [20] shows that even if each user keeps their utility function private but we *model* all the different users by an arbitrary weight that scales their utility function relative to others, users will allocate themselves flow rates so that the cost they cause will equal the weight they choose—weighted proportional fairness.

But such a flow rate allocation is not the measure of fairness, it is merely a possible *outcome* caused by cost fairness, given some assumptions about how to model users' private utility functions. Enforcing underlying cost fairness is in itself a sufficient form of fairness. We repeat: *the resulting relative flow rates are not the measure of fairness*.

Most importantly, Kelly proved cost fairness would lead everyone to maximise their combined aggregate utility across the whole Internet. In other words, if anyone was allocated more and someone else less, the outcome would be worse as a whole. This is why cost fairness is so important, as other forms of fairness cannot be better, unless some major flaw is found in Kelly's assumptions. Kelly *et al* also proved that, even though relative flow rates would likely be very different from those seen today, the Internet would remain stable given reasonable constraints and assumptions [22].

While on the subject of assumptions, we should add that the benefit of a real-time application depends on jitter, not just transfer rate. But simple scaling arguments show that it will be possible for network operators to minimise congestion delay as networks increase in capacity [21, §2], an argument supported by recent research showing that router buffers are often significantly oversized [14].

On the other hand, no-one has even tried to claim that flow rate equality achieves any fairness objective. It has just been asserted as an arbitrary engineer's dogma. This is why flow rate fairness is so open to criticism as unrealistic—having no basis in any recognised form of fairness in real life, science or philosophy.

Proponents of flow-rate fairness might be forgiven for aiming for an 'unrealistic' form of fairness if a 'realistic' form was difficult to implement in practice. In fact, it is flow rate fairness that is completely impractical to enforce (§4.3). The reason we are resurrecting cost fairness is that we believe there are now much more practical ways to enforce it—ways that are built around existing Internet congestion control but, unlike Kelly's, they don't require all ISPs to change their retail model to congestion charging.

But how would users "allocate themselves flow rates so that the cost they cause will equal the weight they choose"? If they were made accountable for congestion, they would install a version of TCP with a weight parameter (for example MulTCP [7]), at least for TCP-based applications. Of course, most users wouldn't want the fuss of weighting each individual flow. But if they chose to set policies on average for large classes of flows (or to accept the defaults set by application developers), the resulting suboptimal outcome for themselves would be their own private choice to trade optimality against hassle. The underlying fairness criterion would still be met: that people should be accountable for the costs they cause to others.

In contrast, with flow-rate fairness, two flows may cause orders of magnitude different costs to others (for instance if one has been running orders of magnitude longer) by running at equal rates. Nowhere do we find any justification for the dogma that flow rates must be equal to be fair. Nowhere do we find any rebuttal of Kelly's destruction of flow rate fairness, even after nine years.

4. ECONOMIC ENTITIES NOT FLOWS

4.1 Something to Integrate the Allocations

Imagine loaves of bread are regularly delivered to a famine-struck refugee camp. Each time a loaf is brought out, a queue forms and the loaf is divided equally among those in the queue. If the individuals who appear in each queue are always different, except for one who always appears in every queue, would it still be fair to share each loaf equally among those in each queue? How do we tell whether the individual is needy or greedy?

This example shows that realistic fairness policies must depend on an individual's history. But if that isn't a convincing argument, it doesn't have to be. We don't have to show that fairness policies *must* depend on history, only that realistic ones *probably will*. So a fairness mechanism that claims to support commercially realistic fairness policies must be structured to hold individual history without destroying scalability. And here, 'individual' means some real-world entity with an economic existence, not a flow.

Router-based flow rate fairness mechanisms tend to have to be myopic. To be otherwise would seem to require holding the history of most Internet connected individuals on most routers, because a flow from nearly any individual in the world might appear at nearly any router. So instead, router-based schemes tend to share out flow rate at each instant without regard to individual history—and unfortunately without regard to commercial reality.

Instead of arbitrating fairness on routers, fairness can be and already is arbitrated where state can be held scalably—at the endpoints where the congestion costs of each individual are already collected together. One reason for our frustration with the networking community's focus on flow rate fairness is that the TCP/IP-based architecture of the Internet already has a structure very close to that required to arbitrate fairness based on the costs that individuals cause.

Congested routers generate cost signals (losses or ECN marks) that are carried to the transport causing the congestion, piggy-backed in the packet stream either as gaps in the transport stream or as ECN marks. These congestion signals are already fed back to the sending transport by nearly all transport protocols. And congestion control algorithms like TCP already adapt their flow rates in response to congestion. So all we would need to change would be to use a weighted TCP algorithm [7] (or equivalent for inelastic applications) that could weight itself under the control of a process overarching all the flows of one user, which would take into account the user's cost history across all flows.

Of course, there is no incentive for anyone to voluntarily subject themselves to such fairness (nonetheless, they already subject themselves to TCP which voluntarily halves its rate whenever it senses congestion). But as we shall see in §4.3.1, policing fairness between individuals (and between networks) at their point of attachment to the Internet now seems tractable, whereas getting every router to police fairness between every individual connected to the Internet is a pipedream, because it would be extremely complicated for routers to have to know about individuals globally.

4.2 Comparing Costs

So, how come one attachment point can arbitrate fairness between everyone on the Internet when it only knows about locally attached individuals? Do we have to add some fully connected mesh of co-ordination messages between every endpoint in the world? The answer is no, because, in a very subtle sense, we already have such a mesh. Fairness at one endpoint is kept in line with all the others by the commonly aligned discipline of *cost* throughout the globe. Cost in any part of the world has an exchange value with cost in any other part, because, wherever there's an Internet attachment, there's a connection with the global economy.

Different types of users (heavy users, light users, servers, server farms, companies) will want to be able to cause different volumes of congestion. As long as congestion can be equated to cost, it can be related to the amount each user has paid for their attachment to the Internet. And if the fairness of the market is unpalatable, any organisation (e.g. universities, corporations, governments) can shield its members from market fairness by asserting any other form of fairness between them, even if they are spread around the Internet. But the organisation as a whole can still be held accountable for each of its members' costs (see §5).

To be able to compare costs globally, we cannot merely talk of volume of congestion as a cost to other users without calibrating it—without specifying how it relates to monetary cost. In a competitive market, the monetary cost that should be assigned to congestion volume turns out to be the marginal cost of the capacity needed to alleviate the congestion [25] (the term 'marginal' cost is used in economics for the slope, $\frac{\partial C}{\partial X}$, of the curve of cost, C, against capacity, X).

Note that when we say that the cost of congestion equates to the marginal cost of capacity, we are not introducing any additional cost; we are merely categorising cost into subdivisions. So, an existing flat fee Internet charge should be considered to consist of parts to cover:

- operational (non-capacity) costs;
- capacity upgrade costs to alleviate congestion;
- the balance of the cost of capacity.

The distinction between the last two is important, because the cost of capacity is traditionally shared out in proportion to access link capacity. But different users with the same access link capacity can cause *hugely* different volumes of congestion across time and across all the Internet links they regularly use, so it is fair to share out the upgrade cost part in proportion to congestion caused, not access link capacity.

Operators might well invest in physical capacity ahead of time but set a logical limit above which congestion marking starts. Such operators continually receive information on how much real demand there is for capacity while collecting revenue to repay their investments. Such congestion marking controls demand without risk of actual congestion deteriorating service.

Once a cost is assigned to congestion that equates to the cost of alleviating it, users will only cause congestion if they want extra capacity enough to be willing to pay its cost. Of course, there will be no need to be too precise about that rule. Perhaps some people might be allowed to get more than they pay for and others less. Perhaps some people will be prepared to pay for what others get, and so on.

But, in a system the size of the Internet, there has to be be some handle to arbitrate how much cost some users cause to others. Flow rate fairness comes nowhere near being up to the job. It just isn't realistic to create a system the size of the Internet and define fairness within the system without reference to fairness outside the system—in the real world where everyone grudgingly accepts that fairness usually means "you get what you pay for".

Note that we we use the phrase "you get what you pay for" not just "you pay for what you get". In Kelly's original formulation, users had to pay for the congestion they caused, which was unlikely to be taken up commercially. But the reason we are revitalising Kelly's work is that recent advances (§4.3.2) should allow ISPs to keep their popular flat fee pricing packages by aiming to ensure that users cannot cause more congestion costs than their flat fee pays for.

The details of all this dirty commercial reality don't have to concern the research or the networking standards communities. It is sufficient to design protocols so that congestion costs *can* be integrated together at some higher layer across different flows and across time, so that senders *can* be made accountable for the congestion they cause. Systems and protocols intended for Internet deployment do not have to *always* realise the sort of fairness over time that we find around us in the real world, but they must *be able* to.

This subtle connection with the global economy at every Internet attachment point ensures that there is no need for some system to decide how far back the history of each individual's costs should still be taken into account. Once the cost that one entity causes to others (integrated over time and over all its flows) has been suffered by that entity itself, it can be forgotten. Just like the costs for all the other benefits everyone assimilates in their daily lives.

4.3 Enforcement of Fairness

This section drives the final nail into the coffin of flow rate fairness, exposing flaws that even those within the box have to turn a blind eye to, in order to convince themselves that the world within the box is perfectly consistent.

4.3.1 Cheating with Whitewashed or Split Flow IDs

In the real world of deployed networks, if it is easy to cheat the fairness mechanism to get an unfair allocation, it's hardly a useful fairness mechanism. All known flow rate fairness mechanisms are wide open to cheating.

For instance, if I am the customer of a system giving max-min flow rate allocations, it is in my interest to split the identities of my flows into lots of little flows until they are all less than the minimum allocation. Then the system will dance to my tune and reduce the allocations of everyone else in order to increase all the allocations of my little flows. The

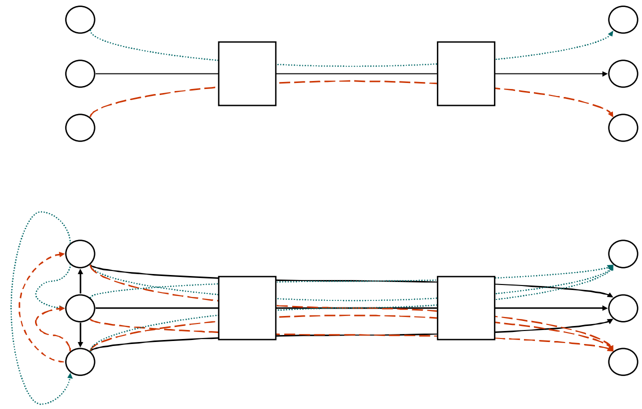

Figure 1: Splitting flow identifiers to cheat against flow rate fairness.

more I split my traffic down across more and more identifiers, the larger share of the resource all my flows taken together will get.

Further, if a history-based fairness mechanism (§4.1) believes it should allocate fewer resources to one flow identifier that it considers has already been given enough, it is trivially easy for the source behind that identifier to create a new identifier with a whitewashed reputation for its traffic.

And it's no good imagining that a router will be able to tell which flow IDs are actually all from the same entity (either in the security sense or the economic sense), because routers have to arbitrate between flows emanating from networks many domains away. They cannot be expected to know which sets of flow identifiers should be treated as a single entity. Flows between a pair of IP addresses may even be attributable to more than one entity, for instance, an IP address may be shared by many hundreds of accounts on a Web or e-mail hosting site or behind a NAT.

Bottleneck policers [10, 5, 29], suffer from the same inherent problem. They look for a flow ID at a bottleneck that is consuming much more bit rate than other flows in order to police use of TCP. But anyone can cheat by simply running multiple TCP flows. If the policer looks for cheating pairs of source-destination IP addresses, without regard to port numbers, a pair of corresponding nodes can still cheat by creating extra flows from spoofed source addresses after telling each other out of band where to send acknowledgements (or just using error correcting coding, not acks).

Alternatively, pairs of corresponding nodes can collude to share parts of each other's flows. For instance, if the three pairs of nodes in Fig 1 are trying to communicate, the senders can act as stepping stones for each other so that their three (n) flows appear as nine (n^2) across the bottleneck link in the middle. In effect, they have created a routing overlay, much like BitTorrent file-sharing software does. If one pair of naïve nodes competes for this bottleneck against n pairs of nodes adopting this strategy, it will get about n times smaller share than each other pair, assuming n is large.

Given identifiers can generally be freely created in cyberspace, it is well-known that they shouldn't be relied on for resource allocation (or more generally for negative reputation) [9, 13]. Kelly [20] chose cost-based fairness (his term was 'pricing per unit share') because it was immune to this problem—it allocates cost to bits not to flows and hence doesn't rely on any cyber-identifiers.

In summary, once one accepts that fairness should be based on concepts from social science, fairness can only be meaningful between entities with real-world identities—humans, institutions, businesses. Otherwise two entities can claim to have arbitrarily many flows between them, making fairness between flows completely meaningless.

4.3.2 Enforcing Cost Fairness

If enforcing flow rate fairness is impractical, is enforcing cost fairness any more achievable? Happily, the Internet's architecture is already suited to carrying the right cost information for cost fairness mechanisms to be enforced in a non-co-operative environment.

Kelly's stated motivation for his focus on pricing was so that the system would be applicable in a non-co-operative environment. In 1999, Gibbens and Kelly went further, pointing out [16] that ECN [30] provided an ideal basis on which to base cost fairness. The idea was simply for network operators to ECN mark traffic at congested routers without regard to flows, then to apply a price to the volume of traffic carrying ECN marks, which would make the transport endpoints accountable for the congestion they caused.

However, understandably, the idea of Internet retailers charging their end-customers directly for congestion met strong resistance. Customers are known to be highly averse to unpredictable charges for services [28, §5] so duration charging for each Internet flow was unlikely to replace flat monthly charging.

Many threw out the baby with the bath water, associating Kelly's theoretical work solely with its suggested pricing model. But over the ensuing years, an active research community has sought to keep the underlying theory but wrapped around with more realistic incarnations.

Indeed the recent proposal called re-ECN [4] claims to do just that. Here the discussion is confined to whether the economic structure and functional effect on the network service that re-ECN aspires to is valid. If it is, the research agenda should be focused on producing that outcome, even if re-ECN itself isn't the answer[2].

Re-ECN aims not to constrain retail pricing, requiring no change to typical flat rate Internet contracts. But it enables addition of a policer that can limit the volume of congestion a customer's sent traffic causes over, say, a moving month. Thus, if endpoint congestion control (e.g. MulTCP [7]) doesn't voluntarily act fairly the network ingress can force it to. It is expected that various styles of policing (including none) will evolve through market selection.

Although Gibbens & Kelly rightly identified that standard ECN reveals the necessary information for cost-based fairness, it doesn't reveal it in the right place for network layer policing—at the *sender's* network attachment. In the current TCP/IP architecture, congestion information emerges from the end of a forward data path, which is the last point in the feedback loop that any network operator can reliably intercept it—the wrong end for policing the sender.

Re-ECN preserves IP's datagram model, but makes delivery conditional on the sender 'pre-loading' packet streams with enough 'credit' to remain non-negative despite being decremented by congestion experienced along the path. It should then be in *both* the endpoints' interests for the sender to use a pattern of feedback where the sender re-inserts the feedback from each congestion event into the next sent packet as a 'credit' (re-feedback [3]). It should also be in the sender's interest to start every flow slowly and with some initial 'credit' while it establishes the path's congestion level.

Like Kelly's original proposal, re-ECN uses ECN routers (and receivers) unchanged to ensure the cost of congestion is communicated to each transport causing it, precisely in proportion to their bit rates, without any per-flow processing in the network. But, unlike Kelly, sources not receivers are held responsible and the network cannot raise unsolicited charges without the sender deliberately marking packets itself.

Re-ECN also aims to ensure cost-fairness between whole networks. Because the congestion level in every stream of packets decrements towards zero, at an inter-domain border both neighbouring networks can count the bulk volume of congestion that the passing packets are causing downstream of the border. If the downstream neighbour charges the upstream neighbour this cost (complementing fixed capacity charges), the upstream network should in turn want to ensure its upstream users (or networks) are accountable for their share of these costs arriving from their borders.

A provider could profit from faking higher congestion notification, but to remain competitive it will be upper bounded by the least congested competing path and lower bounded by actual congestion. Therefore, 'notified congestion' is not really congestion at all, but simply the market price. However, as competition intensifies, it tends towards actual congestion. In any market, if a provider overprices slightly, even occasionally, it risks losing customers and all the revenue it had hoped would pay back its investments. As in current communication markets, we expect price comparison robots will lubricate this competition. But in those jurisdictions where there is ineffective regulation of access networks, the tendency towards natural monopoly at the link layer will distort congestion pricing, but no more than it would distort *any* network layer pricing.

Each network could choose to share out its downstream costs between its upstream customers by some other fairness policy than cost (including absence of policy, which ensures incremental deployment). So, on the grander scale, re-ECN aims to ensure that networks have to be fair to each other, and that different fairness policies can co-exist, which is the subject of the next section.

5. FAIRNESS BETWEEN FAIRNESSES

A social anthropologist would be able to give numerous examples of tribes and societies holding differing opinions on fairness. But, we must also recognise that societal views of fairness are heavily influenced by the fairness that a market would produce [34]. Just as gravity pre-dated Newton, the invisible hand of the (maturing) market had been allocating resources in society long before Adam Smith noticed, particularly where the larger picture of trade between societies was concerned. Equality is sometimes considered fair for life's essentials, but in life few expect to get an equal share of every cake for nothing. As a society, we accept that a reasonably competitive market mechanism does produce a 'realistic' form of fairness; a form of fairness that people grudgingly accept they have to live with, where the buyer

[2]Readers tempted to game re-ECN should use the full spec. [4], which, as of early 2007, documents one outstanding vulnerability and defences against other known attacks.

gets no more than she pays for, at a competitive price that reflects the effort expended by the seller.

However, monarchs, governments, charities and so on have also been stamping their own view of fairness on this backdrop, sometimes less equal sometimes more. But even if different allocation schemes are chosen locally, perhaps taking account of social inequality, on a global scale arbitration between local views on fairness has largely been through market economics—we are not asking anyone to judge whether this is good or bad, it just is. The Internet should at least be able to cope with the world as it is (as well as how it might be). This doesn't imply we believe that economic forces are somehow above policy control. Rather, we observe that market forces (aside from wars) have been the default *global* resource allocation mechanism over many centuries. In the Greco-Roman civilisations, in the Buddhist, Confucian and later in the Islamic world, trade was a necessary but not central aspect of life. And over the last two decades, Western civilisations have been going through a phase of 'economics imperialism', where attempting to exert policy control over economics is even viewed as counter-productive.

However, we must not assume the current globalisation trend [32] heralds the end of history. The Internet should be able to reflect the shifting of societal forces as different local fairness regimes come and go—'design for tussle' [6]. On the whole, interworking of resource allocation between most parts of the Internet must *be able* to be based on market economics, but it should be possible to apply other fairness criteria locally. For instance, a University might choose to allocate resources to each student equally rather than by how much their parents can afford. But the resources one whole University gets relative to another institution depend on how much each pays their service provider.

With arbitration of fairness at the network edge, these enclaves where local fairness prevails can be virtual overlays; they need not align with physical network boundaries. A distance-learning University or company with a mobile sales-force could buy quotas from different networks and redistribute the aggregate among its members using its own view of fairness. Or whole countries might arrange to subsidise a minimum universal service obligation for Internet *usage*, but still, the country as a whole would be expected to pay its way in the world.

On the other hand, in market-led countries, commercial ISPs might solely allocate resources proportionate to customer subscriptions. Local pockets of heterogeneity will exist, from computer clubs to NATO, but the overall fabric of resource allocation gluing all these pockets together at the (inter)network layer is likely to be market-based.

This is what we mean by 'realistic'—fitting the commercial reality of a global market economy. We are fully aware that the power of market economics can be stretched too far; controlling aspects of society where economic assumptions break down. But we are not advocating that one religion should replace another—market economics replacing flow rate fairness. However, in the case of Internet resource allocation, it must at least *be possible* to use market economics, despite its known failings, given it is currently the most appropriate tool for managing conflicting demands on resources from any part of the globe.

A market is meant to optimise allocations in the face of conflicts of self-interest. If we want to assert other fairness regimes, we must recognise this acts against self-interest. If we don't understand how to overcome self-interest, its invisible hand will force its will on us some other way, distorting our attempts to work against it. This is why the loop holes in flow rate fairness are being so thoroughly exploited.

And this is our point. A market *mechanism* has to be *designed*. A weak design will be exploited mercilessly. The designs behind flow rate fairness are worse than weak. They are not even aware that, as resource allocation mechanisms, they *should* be able to meet the stringent requirements of a good market mechanism, such as forgery-resistant 'currency', information symmetry, internalisation of externalities and so forth.

If we did wish to promote the cause of equality, equalising flow rates would in no way achieve our ends. In fact, it would only promote the cause of selfishness and malice, because flows don't equate to people, so its broken logic can be thoroughly exploited. Only by providing a bullet-proof mechanism to arbitrate self-interest, can we then move on to allocate resources locally in other ways.

6. THE SEMINAL LITERATURE

Max-min flow rate fairness has a long history in networking, with research to find distributed (router-based) max-min algorithms starting in 1980 [18] and Nagle proposing a novel approach in 1985 (RFC970 [27]). All these early 'fair queuing' algorithms gave equal rights to each source. In 1989, to solve the problem of some sources deserving more rate than others, the authors of 'weighted fair queuing' (WFQ) proposed per-source destination pair as a better model of the size of different sources. They admitted that a source could increase its share by faking transfers with numerous destinations, but a reasonable tradeoff between efficiency and security was required [8]. All these 'fair queuing' approaches allocate bit rate as their measure of fairness.

TCP congestion control was also introduced in the late 1980s [17], inheriting the assumption that it would be fair if flow rates through a single bottleneck converged on equality.

In 1991, Mazumdar *et al* [26] pointed out that there was nothing special about max-min fair rate allocation, and that other *ad hoc* definitions of fairness perhaps based on ratios of individual demands would be no less valid. Mazumdar *et al* advocated that it would be precise to base a definition of fairness on game theory, specifically the Nash bargaining solution. This resulted in proportional fairness, but still using the rate allocated to flows as the measure of fairness.

Kelly considered that Mazumdar's use of co-operative game theory was unlikely to be relevant to public networks where fairness would have to be enforced. Instead, in 1997, he introduced *weighted* proportional fairness [20], which finally broke the link between fairness and flow rates. However, the break in tradition wasn't obvious because the new form of fairness could easily be expressed in terms of flow rates, essentially using the weight of a flow as a 'fiddle-factor'.

Kelly showed that all a network had to do to achieve fairness in its economic sense (cost fairness) was to share the cost of congestion among bits (not flows). Then, as long as the network made users experience the cost of their bits, users could choose any size flows they wished. But their choice would be regulated by their own trade off between their valuation of bit rate and the charge for congestion.

Kelly's fairness with respect to bit rate per unit charge could also be (and was) framed in terms of fairness between flows by allowing the user an arbitrary choice of weight per flow. But Kelly pointed out that a flow could be divided into sub-flows without changing the overall rate allocation to all the sub-flows taken together; the user merely had to imagine that the weight she assigned to one flow could be subdivided proportionately into its sub-flows.

Kelly's work built on MacKie-Mason & Varian's seminal paper on the economics of networks from 1995, "Pricing Congestible Network Resources" [25]. This work explained the dual role of congestion costs in controlling demand and regulating supply, in welfare maximising, competitive and monopoly markets.

In his 1997 paper, Kelly framed cost fairness in terms of weighted proportional fairness of flow rates in order to relate to an ATM technology context. With ATM's flow-based user-network interface, users had to declare the weight they chose for their flows to the network. But by 1998 Kelly *et al* applied this work to an Internet setting where flows were not part of the user's interface with the network, so flow weights could become a purely private device, internal to the user's rate control algorithm [22]. Nonetheless, the *outcome* at the flow level was still weighted proportional fairness, and the underlying fairness that produced this outcome was still based solely on sharing the cost of congestion among bits.

Back in 1995, Shenker had identified two main types of network traffic: elastic and inelastic, distinguished respectively by their concave and sigmoid utility functions [33]. Whatever the utility function, Kelly teaches us that covering congestion costs is sufficient to achieve fairness. But then the outcome (in terms of flow rates) depends on the type of utility function:

- Weighted proportionally fair flow rates will be the outcome for elastic traffic streaming;

- Inelastic traffic flows hit a discontinuity once congestion rises beyond a certain level, at which point each is better off with zero rate, leading to a need for some form of admission control, whether self-admission control or arbitrated by the network [15].

- Key & Massoulié identified a third major class of network traffic where utility is derived solely from the duration required to complete transfer of a fixed volume of data [23]. They showed that, if cost fairness applied, self-interested congestion control would toggle between full line rate and zero (with occasional probes). Such behaviour alone destabilises the network, but it can be stabilised by mixing with streaming traffic [24]. Research on the second order incentives necessary to encourage stability continues.

Since these seminal papers in the late 1990s, theoretical refinement has continued, but the main thrust of research has been to find more realistic and practical ways of applying the insights, a process which is now bearing fruit (see §4.3.2).

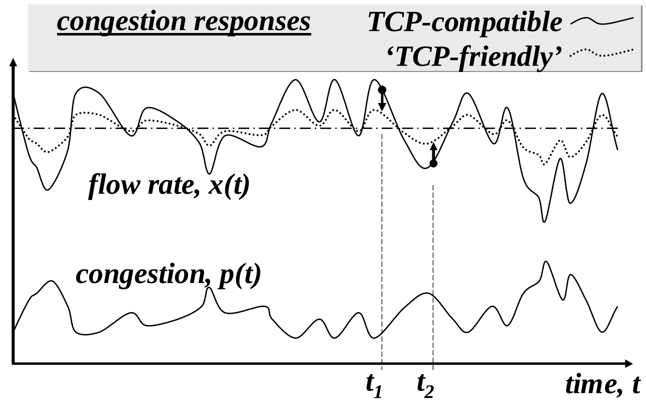

Figure 2: Schematic showing 'TCP-friendly' flows cause more congestion than TCP. A TCP-friendly flow is smoother than a TCP-compatible one but with the same mean rate if measured over long enough time. Therefore at times of high congestion (t_2) it uses more bandwidth than TCP while at times of low congestion (t_1) it uses less.

7. CRITIQUES OF SPECIFIC SCHEMES

7.1 Max-min flow rate fairness

In 1997, Kelly demonstrated [20] that realistic users would not choose max-min flow rate fairness if they were accountable for the congestion they caused to others. Users would only have chosen max-min if they valued bit rate with an unrealistically extreme set of utility functions that were all identical and that all valued low bit rate infinitesimally less than high bit rate. To spell Kelly's result out even more bluntly, max-min fair rate allocation would only be considered fair if *everyone* valued bit rate in a really weird way: that is, they they all valued very low bit rate hardly any less than very high bit rate and they all valued bit rate exactly the same as each other. (Note that max-min could be meaningful if allocating something like utility among users, but not rate among flows.)

7.2 TCP

TCP's congestion avoidance [1] leads to a form of fairness similar to cost fairness, except it is myopic, only being concerned with each instant in time and with each flow, as explained in §4. To be cost fair each user would have to take account of costs across time and across flows, and weight each class of TCP flow according to its importance to them, as can be done with MulTCP [7].

7.3 TFRC

An algorithm that converges on the same flow rate as TCP at equilibrium is called TCP-friendly. It can only claim to be TCP-compatible if it also exhibits the same dynamics as the TCP specification [1]. Certain streaming applications won't work unless they are allowed a more sluggish response to congestion than TCP's, so researchers invented TCP-friendly rate control (TFRC [11]) to define fair use of the network in competition with TCP-compatible flows.

Given TFRC aims to emulate TCP, by far its most significant fairness problems are those it shares with TCP as just mentioned. However, even if we set aside this myopia in time and within flows, TFRC exhibits an extra fairness problem because its design was based wholly on the broken idea that it is fair for a TCP-friendly flow to get the same rate as a TCP-compatible flow.

To explain, we need to remember that both congestion and flow rate vary over time. A more nimble congestion response like TCP's can mirror changing congestion fairly faithfully. It reduces its rate quickly during periods of higher congestion and increase again more quickly whenever congestion falls. In Fig 2 the resulting schematic plots of congestion and flow rate are shown as mirror images of each other. A more sluggish rate response is not as good at tracking the fast-changing congestion process. So the sluggish flow more often uses higher bandwidth when congestion is high, and more often uses lower bandwidth when congestion is low, causing more volume of congestion on average. Giving more during times of plenty doesn't compensate for taking it back during times of scarcity.

7.4 XCP and router-based fairness schemes

It is well-known that TCP congestion control is running out of dynamic range and many proposals for replacements that can take advantage of higher link capacities by accelerating faster have been put forward. XCP was the first of a family of router-based hi-speed congestion control mechanism, but it is particularly of interest because it claims to allow different fairness criteria to be configured.

However, XCP fairness is based on the myopic flow-rate-based view that we have so roundly criticised in this paper. For instance, XCP claims to be able to achieve a weighted proportional fair rate allocation [19, §6] by adding a weight field to each packet, but it glosses over how anyone could regulate each user's choice of the weight. If we compare weighted fair XCP with Kelly's original ATM-based weighted proportional fairness, the weight parameter advises network equipment on what allocation it should give each flow, but there is no direct congestion information in the XCP protocol that could be used at the ingress to make each source accountable for its choice of weight.

Further, we believe it will be necessary to be able to apply different fairness criteria to different subsets of users of a network and subsets across an internetwork as outlined in §5. We cannot immediately see how this would be feasible with router-based approaches like XCP, where routers would seem to have to share information on the history of each user with potentially every other router in the world (as explained in §4.1).

A combination of XCP's protocol fields could yield approximate congestion information to integrate each sender's congestion cost history at the access network close to the sender. This would allow the user's choice of weight to be regulated and enable different forms of fairness to be asserted locally. But one then has to question whether it would be simpler for the end system to do the rate control, given it has to give routers all the information they need to arbitrate fairness between flows anyway.

7.5 WFQ

Weighed fair queueing aims to isolate the capacity that a flow receives from excessive load applied by other flows, while at the same time ensuring the router's capacity is fully utilised. WFQ allocates capacity per-flow not per-user, so it is vulnerable to the flow ID splitting games described in §4.3.1 and it only controls fairness over flow lifetimes, not over user history. A comparison of cost fairness against WFQ (both as originally defined and as sold commercially) would be interesting given features of the two approaches overlap even though they don't have the same goals. But this subject would require a dedicated paper.

8. CONCLUSIONS

The outstanding barrier to a realistic resource allocation architecture for the Internet is purely religious. In much of the networking community you have to put fairness in terms of flow rates, otherwise your work is 'obviously' irrelevant. At minimum, you are an outcast, if not a heretic. But actually basing fairness on flow rates is a false god—it has no grounding in philosophy, science, or for that matter 'commercial reality'.

It is a classic case of a hegemony where those living within the box don't recognise the existence of the box, let alone the world outside the box. This paper was written from frustration that no-one inside the box believes that voices outside the box should be listened to. We expect complaints about the blunt style of this paper, but it seemed the only way forward was to force the issue, by making the box look ridiculous in its own terms.

Outside the box, cost fairness was derived from economic concepts of fairness back in 1997. When flow rate fairness is seen through the wider eyes of this economic analysis it is clearly broken, even on its own terms. The criticism is far more damning than merely whether allocations are fair. Both the thing being allocated (rate) and what it is allocated among (flows) are completely daft—both unrealistic and impractical. However, the Internet community continues to judge fairness using flow rates, apparently unaware that this approach has been shown to have no intellectual basis. These flow rate fairness algorithms are myopic in both space and time—they are completely unable to control fairness at all, because they don't adjust depending on how many flows users create and for how long.

To be clear, this accusation applies to the so-called 'fairness' that emerges from the TCP algorithm and the various fair queuing algorithms used in production networks. And, more worryingly, this broken idea of flow rate fairness has carried over into the community working on replacing the Internet architecture.

In real life, fairness generally concerns costs or benefits. Flow rate doesn't come anywhere near being a good model of either. User benefit per bit rate might be ten orders of magnitude different for different types of flow. And cost depends on the product of bit rate with congestion, which is very variable and nothing like bit rate alone.

Worse, there is no evidence whatsoever that fairness between flows relates in any way to fairness between any real-world entities that one would expect to treat fairly, such as people or organisations. If fairness is defined between flows, users can just create more flows to get a larger allocation. Worse still, fairness between flows is only defined instantaneously, which bears no relation to real-world fairness over time. Once the idea of fairness based on integrating costs over time is understood, we cannot see any reason to take any form of instantaneous per-flow rate fairness seriously, ever again—whether max-min or TCP.

Even if a system is being designed somehow isolated from the economy, where costs will never have to relate to real economic costs, we cannot see why anyone would adopt these forms of fairness that so badly relate to real-life fairness. For instance, how can people still be designing schemes to achieve max-min flow rate fairness years after Kelly's proof that users would have to value bit rate in a really weird way in order for max-min fairness to be desirable?

In contrast, cost fairness promises realistic solutions to all these issues. Further, it seems more tractable to enforce, unlike flow rate fairness, which seems inherently broken in this respect. We believe cost fairness is a coherent way forward with all the technical barriers overcome, or close to being overcome. This is where the research agenda should be focused.

If anyone with aspirations to scientific credentials still wants to cling to flow rate fairness, they must justify their preposterous position with reference to some previously respected fairness notions in philosophy or social science. In this paper, we have shown how the whole ideology is unlikely to be up to such rigour.

Acknowledgements

Thanks are due to Scott Shenker, Louise Burness, Ben Strulo, Marc Wennink, Arnaud Jacquet, Jon Crowcroft, Frank Kelly, Peter Key, Damon Wischik and the anonymous reviewers for their comments. However, the author alone shoulders the blame for any offence caused by the bluntness of style.

9. REFERENCES

[1] M. Allman, V. Paxson, and W. Stevens. TCP congestion control. Request for comments 2581, IETF, Apr. 1999.

[2] R. Braden, D. Clark, S. Shenker, and J. Wroclawski. Developing a next-generation Internet architecture. White paper, DARPA, July 2000.

[3] B. Briscoe, A. Jacquet, C. Di Cairano-Gilfedder, A. Salvatori, A. Soppera, and M. Koyabe. Policing congestion response in an internetwork using re-feedback. *Proc. ACM SIGCOMM'05, CCR*, 35(4):277–288, Aug. 2005.

[4] B. Briscoe, A. Jacquet, A. Salvatori, and M. Koyabe. Re-ECN: Adding accountability for causing congestion to TCP/IP. Internet Draft draft-briscoe-tsvwg-re-ecn-tcp-03.txt, IETF, Oct. 2006. (Work in progress).

[5] P. Chhabra, S. Chuig, A. Goel, A. John, A. Kumar, H. Saran, and R. Shorey. XCHOKe: Malicious source control for congestion avoidance at Internet gateways. In *Proc. IEEE International Conference on Network Protocols (ICNP'02)*. IEEE, Nov. 2002.

[6] D. Clark, K. Sollins, J. Wroclawski, and R. Braden. Tussle in cyberspace: Defining tomorrow's Internet. *Proc. ACM SIGCOMM'02, CCR*, 32(4):347–356, Oct. 2002.

[7] J. Crowcroft and P. Oechslin. Differentiated end to end Internet services using a weighted proportional fair sharing TCP. *CCR*, 28(3):53–69, July 1998.

[8] A. Demers, S. Keshav, and S. Shenker. Analysis and simulation of a fair-queueing algorithm. *CCR (SIGCOMM'89)*, 19(4):1–12, Sept. 1989.

[9] M. Feldman, C. Papadimitriou, J. Chuang, and I. Stoica. FreeRiding and whitewashing in peer-to-peer systems. In *Proc. Workshop on Practice and Theory of Incentives in Networked Systems (PINS'04)*, pages 228–236. ACM SIGCOMM, 2004.

[10] S. Floyd and K. Fall. Promoting the use of end-to-end congestion control in the Internet. *IEEE/ACM Transactions on Networking*, 7(4):458–472, Aug. 1999.

[11] S. Floyd, M. Handley, J. Padhye, and J. Widmer. TCP friendly rate control (TFRC): Protocol specification. Request for comments rfc3448.txt, IETF, Jan. 2003.

[12] S. Floyd and V. Jacobson. Random early detection gateways for congestion avoidance. *IEEE/ACM Transactions on Networking*, 1(4):397–413, Aug. 1993.

[13] E. Friedman and P. Resnick. The social cost of cheap pseudonyms. *Journal of Economics and Management Strategy*, 10(2):173–199, 1998.

[14] Y. Ganjali and N. McKeown. Update on buffer sizing in Internet routers. *CCR*, 36, Oct. 2006.

[15] R. J. Gibbens and F. P. Kelly. Distributed connection acceptance control for a connectionless network. In *Proc. International Teletraffic Congress (ITC16), Edinburgh*, pages 941–952, 1999.

[16] R. J. Gibbens and F. P. Kelly. Resource pricing and the evolution of congestion control. *Automatica*, 35(12):1969–1985, Dec. 1999.

[17] V. Jacobson. Congestion avoidance and control. *Proc. ACM SIGCOMM'88 Symposium, CCR*, 18(4):314–329, Aug. 1988.

[18] J. Jaffe. A decentralized, "optimal", multiple-user, flow control algorithm. In *Proc. Fifth Int'l. Conf. On Computer Communications*, pages 839–844, Oct. 1980.

[19] D. Katabi, M. Handley, and C. Rohrs. Congestion control for high bandwidth-delay product networks. *Proc. ACM SIGCOMM'02, CCR*, 32(4):89–102, Oct. 2002.

[20] F. P. Kelly. Charging and rate control for elastic traffic. *European Transactions on Telecommunications*, 8:33–37, 1997. (Correction by R. Johari & F. Kelly at URL: `http://www.statslab.cam.ac.uk/~frank/elastic.html`).

[21] F. P. Kelly. Models for a self-managed Internet. *Philosophical Transactions of the Royal Society*, 358(1773):2335–2348, Aug. 2000.

[22] F. P. Kelly, A. K. Maulloo, and D. 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.

[23] P. Key and L. Massoulié. User policies in a network implementing congestion pricing. In *Proc. Workshop on Internet Service Quality and Economics*. MIT, Dec. 1999.

[24] P. Key, L. Massoulie, A. Bain, and F. Kelly. Fair Internet traffic integration: network flow models and analysis. *Annales des Télécommunications*, 59:1338–1352, 2004.

[25] J. K. MacKie-Mason 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.

[26] R. Mazumdar, L. G. Mason, and C. Douligeris. Fairness in network optimal flow control: Optimality of product forms. *IEEE Trans. on Comms*, 39(5):775–782, May 1991.

[27] J. Nagle. On packet switches with infinite storage. Request for comments 970, IETF, Dec. 1985. (Status: unknown).

[28] A. Odlyzko. A modest proposal for preventing Internet congestion. Technical report TR 97.35.1, AT&T Research, Sept. 1997.

[29] R. Pan, L. Breslau, B. Prabhaker, and S. Shenker. Approximate fairness through differential dropping. *CCR*, 33(2):23–40, Apr. 2003.

[30] K. K. Ramakrishnan, S. Floyd, and D. Black. The addition of explicit congestion notification (ECN) to IP. Request for comments 3168, IETF, Sept. 2001.

[31] J. Rawls. *Justice as Fairness: A Restatement*. Harvard University Press, Cambridge, MA, 2001.

[32] J. R. Saul. *The Collapse of Globalism and the Reinvention of the World*. Viking, Canada, 2005.

[33] S. Shenker. Fundamental design issues for the future Internet. *IEEE Journal on Selected Areas in Communications*, 13(7):1176–1188, 1995.

[34] S. Siderenko. *Characteristics of Perceptions of Social Justice in the Contemporary USSR*, chapter 3, pages 41–45. Centre for Social Anthropology and Computing, University of Kent, 1991.