Principles for Internet Congestion Management

Lloyd Brown[1], Albert Gran Alcoz[2], Frank Cangialosi[3], Akshay Narayan[4], Mohammad Alizadeh[5], Hari Balakrishnan[5], Eric Friedman[1,9], Ethan Katz-Bassett[6], Arvind Krishnamurthy[7], Michael Schapira[8], Scott Shenker[1,9]

[1] UC Berkeley, [2] ETH Zürich, [3] BreezeML, [4] Brown University, [5] MIT, [6] Columbia University, [7] University of Washington, [8] Hebrew University of Jerusalem, [9] ICSI

Abstract

Given the technical flaws with—and the increasing non-observance of—the TCP-friendliness paradigm, we must rethink how the Internet should manage bandwidth allocation. We explore this question from first principles, but remain within the constraints of the Internet's current architecture and commercial arrangements. We propose a new framework, Recursive Congestion Shares (RCS), that provides bandwidth allocations independent of which congestion control algorithms flows use but consistent with the Internet's economics. We show that RCS achieves this goal using game-theoretic calculations and simulations as well as network emulation.

CCS Concepts

• **Networks** → **Network design principles**.

Keywords

Network Architecture

ACM Reference Format:

Lloyd Brown, Albert Gran Alcoz, Frank Cangialosi, Akshay Narayan, Mohammad Alizadeh, Hari Balakrishnan, Eric Friedman, Ethan Katz-Bassett, Arvind Krishnamurthy, Michael Schapira, Scott Shenker. 2024. Principles for Internet Congestion Management. In *ACM SIGCOMM 2024 Conference (ACM SIGCOMM '24), August 4–8, 2024, Sydney, NSW, Australia.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3651890.3672247

1 Introduction

In addition to being a technological marvel whose architecture has accommodated mind-boggling changes in size, speed, technologies, and uses, the Internet is also a massive experiment in decentralized resource sharing. Because computer communications are bursty, the Internet relies on packet-level statistical multiplexing to achieve reasonable efficiency. To deal with the inevitable overloads, the Internet relies on host-based congestion control algorithms (CCAs).

With this approach, the bandwidth a flow receives can depend heavily on the aggressiveness of its CCA. The Internet community quickly recognized that users would have an incentive to deploy ever more aggressive CCAs, thereby leading to overloads. To prevent this, the Internet community informally required all CCAs to be *TCP-friendly* (hereafter TCPF), as defined by [24]: "a flow is TCP-friendly if its arrival rate does not exceed the arrival of a conformant TCP connection in the same circumstances." [1] TCPF primarily applies to wide-area traffic on the public Internet, and we focus on that case in this paper. Specialized bandwidth allocation solutions are available in private deployments such as datacenters, enterprises, and private wide-area networks (WANs), in which there is a single administrative authority.

There are numerous practical and technical problems with TCPF. Prior work has shown that it is difficult to enforce [46] and that our understanding of the dynamics of CCAs breaks down at scale [38]. In addition, TCPF limits CCAs' ability to ramp up quickly and achieve full efficiency [49] and hinders the emergence of new delay-sensitive CCAs (as shown by Copa [5] and Nimbus [28]). It has also become clear that TCPF is no longer a strict requirement in deploying new CCAs, and that, in practice, non-TCPF CCAs will be deployed widely. For example, the TCP-unfriendly CCA BBR [17, 18, 47] has been widely adopted at Google, Amazon, Akamai, Dropbox, and Spotify for significant portions of their traffic.[2]

Given that TCPF is both deeply flawed and no longer adhered to by the major Internet actors, we should consider whether there are suitable alternatives to the TCPF paradigm. *This simple but central issue is the focus of this paper.* To that end, we explore from first principles what new conceptual framework might replace TCPF. However, while we reason from first principles, we do not start with a clean slate. We assume that, within our design/deployment timeframe, there will be no fundamental changes in the Internet architecture (*e.g.,* IP, BGP, and the best-effort service model) and its commercial arrangements (*e.g.,* how ISPs charge for service and peer with each other, and the widespread adherence to valley-free routing [26]). As such, we seek a conceptual foundation for how the Internet should share bandwidth that (i) can be implemented within the current architecture (though requiring additional protocols and mechanisms) and (ii) provides bandwidth allocations that are consistent with the current commercial arrangements between the parties involved.

This paper makes the following contributions:

- We articulate the goal of CCA independence (CCAI) (§2) as the foundational aim for sharing bandwidth.
- In contrast to the specific claims in prior work [12] and the general assumptions in literature (such as [20] and the

[1] At the time of [24], the term "TCP" prescribed a specific CCA: NewReno, as standardized in RFC2582. Also, even the staunchest of the early advocates recognized that TCPF was not tenable at high speeds, but the intent of proposals like High-Speed TCP [23] was to retain TCPF at lower speeds and create new standards for behavior at these higher speeds.

[2] While Google claims recent BBR versions are less unfair than the original [19], researchers dispute this claim [50] and BBRv3 remains TCP-unfriendly. However, our concern is not the degree of BBR's violation of TCPF but rather the lack of resistance to deploying CCAs that do not satisfy TCPF, and this applies to all BBR versions.

ACM SIGCOMM '24, August 4–8, 2024, Sydney, NSW, Australia

ACM ISBN 979-8-4007-0614-1/24/08
https://doi.org/10.1145/3651890.3672247