



Conference on Networked Systems 2021
(NetSys 2021)

Early Work: Path Selection in a Path-aware Network Architecture

Thorben Krüger and David Hausheer

5 pages

Early Work: Path Selection in a Path-aware Network Architecture

Thorben Krüger¹ and David Hausheer²

¹ thorben.krueger@ovgu.de

² david.hausheer@ovgu.de

Networks and Distributed Systems Lab
Otto-von-Guericke University Magdeburg, Germany

Abstract: Modern path-aware networking (PAN) architectures based on packet-carried forwarding state (PCFS) promise to support practical multipath communication with increased availability and redundancy, even for single-homed hosts. How exactly hosts can choose paths in a meaningful way while maintaining good network utilization as well as user satisfaction is the subject of this research. We demonstrate how overall network utilization is affected by the introduction of multipath communication and propose the term “Cost of Multipath” to reason about this observation. We also identify a collection of concrete practical techniques from networking research and engineering that can be adapted to allow hosts to make more informed path decisions. These tie into the practical implementation of an experimental path-selection engine with a high-level API that we aim to utilize to quantitatively investigate host-based approaches for traffic optimization in the existing path-aware SCION network architecture

Keywords: Path-awareness, SCION, traffic engineering, cost of multipath, multipath networking

1 Introduction

In contrast to the current BGP-based Internet, modern path-aware networking (PAN) architectures guarantee verifiable packet forwarding across the network over precise paths specified by the sending host. In a PAN, a host selects one or more desired paths to the destination from a set of alternatives, rather than merely delegating forwarding decisions to routers along the way. As a direct consequence, traffic distribution patterns across such a network are the product of the collective path-choices made on the hosts. Moreover, even a single-homed host can easily distribute its communication with a given destination over an almost arbitrary number of alternative paths.

We argue that these new capabilities enjoyed by the end hosts in the next-generation Internet require more sophisticated mechanisms for intelligent path choices, beyond the usual set of simple heuristics that merely compare path lengths without any consideration for, e.g., network utilization patterns, congestion or other factors.

The remainder of this paper is structured as follows: In Section 2, we give some necessary context for our early work and identify the techniques we aim to use in our implementation work. In Section 3, we discuss the impact that multipath communication can have on network utilization. We share a glimpse into our currently ongoing work before concluding with a summary in Section 4.

2 Background and Related Work

In PAN architectures, the ability to dictate guaranteed paths across the network resides exclusively with the end hosts [PSRC17]. In the SCION architecture, dedicated network-level services provide the necessary inter-domain topology information for a host to select among several possible paths to any destination. Currently, there are no mechanisms to prohibit the sub-optimal use of paths that are, e.g., overly long or congested. A related problem is known from research about peer-to-peer (P2P) networks: The resources shared in P2P systems often exist in multiple replicas [SB09], i.e., they are available from different peers. However, peer selection for resource retrieval in P2P overlay networks is often arbitrary and very rarely takes underlying network topology into account. This results in inefficient traffic patterns on the network, where resources that would also be available from nearby peers are instead retrieved from far away and usually with lower performance, while needlessly increasing network transit traffic and costs. As a remedy, [AFS07] proposes an “oracle” service for P2P overlays that could be deployed by ISPs to help their customers identify and contact peers that are nearby. In our work, we aim to generalize this approach to network paths instead of peers, where a network operator can offer comparable functionality to hosts that wish to adjust their multipath communication to current network conditions for mutual benefit.

In the absence of such centralized mechanisms, another way for a host to obtain some early path quality estimates is an adaption of the “connection racing” approach detailed in [WY12]. Connection handshakes to a destination are simultaneously initiated across a set of alternative paths. The different round-trip-times until handshake completion can then be used as estimates for “path latency” properties.

Efforts are underway at the IETF to develop and agree upon a suitable vocabulary to reason about path property concepts [EK20]. Our theoretical work will follow the current IETF consensus in the relevant terminology wherever possible. Our practical implementation of a suitable path-selection mechanism will also mirror the corresponding IETF proposal [TWE+21].

Lack of coordination between independent selfish agents in game theory and distributed mechanism design is often referred to as the “Price of Anarchy” [CK05]. We aim to develop our own notion of the “Price of Multipath” in deference to this metaphor.

3 Early and Ongoing Work

Intuitively, communication between two nodes across a network should happen over the shortest paths possible. An abstract example is depicted in Figure 1. Needlessly using longer paths, i.e., taking “detours” consumes more resources, resulting in worse network utilization, which is demonstrated in Figure 2. Where multipath communication combines the use of optimal shortest paths with the additional use of longer detours, network utilization is still worse than the optimum but improved when compared to the previous case. See Figure 3.

For the purposes of quick failover and redundancy, the capacity to engage in communication over multiple paths is assumed to be desirable, even in cases where global network utilization suffers. However, we need a qualitative understanding of this trade-off to further reason about traffic engineering under path-awareness.

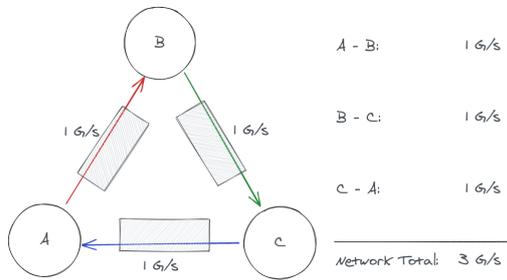


Figure 1: Singlepath direct communication: 3 Gbit/s total bandwidth

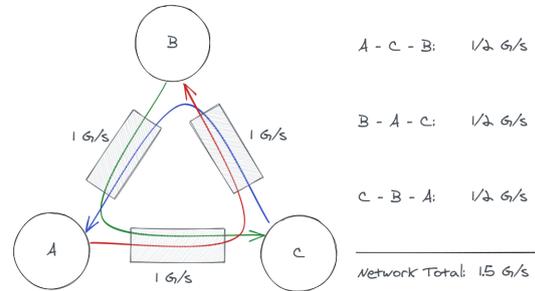


Figure 2: Singlepath indirect communication: 1.5 Gbit/s total bandwidth

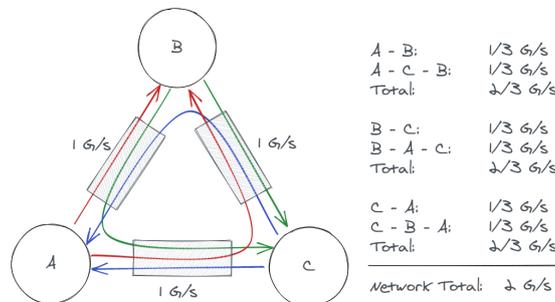


Figure 3: Multipath communication using both direct and indirect paths: 2 Gbit/s total bandwidth

3.1 Research Questions

How can a *Price of Multipath* be defined?

We aim to define a term for the *Price of Multipath (PoM)*, deriving the utilization penalties for a given network topology in terms of the node degree distribution and other parameters.

Is this *PoM* constant for a given network?

We will need to work confirm or disprove the suspicion that the *PoM* is generally a fixed property of a network, in the way that Figure 4 suggests.

What is *fair* path selection?

Can we come up with sane defaults for the path selection mechanism that avoids the pathological case depicted in Figure 4a)? How large should the fraction of bandwidth over a redundant, longer path be? Can there be special allowances and acceptable unfair behavior to ensure QoS?

What effects can path-selection have for network utilization?

Using a suitably sophisticated path-selection mechanism, could a host explore multiple alternative paths to a destination simultaneously, autonomously avoiding congested links and ultimately

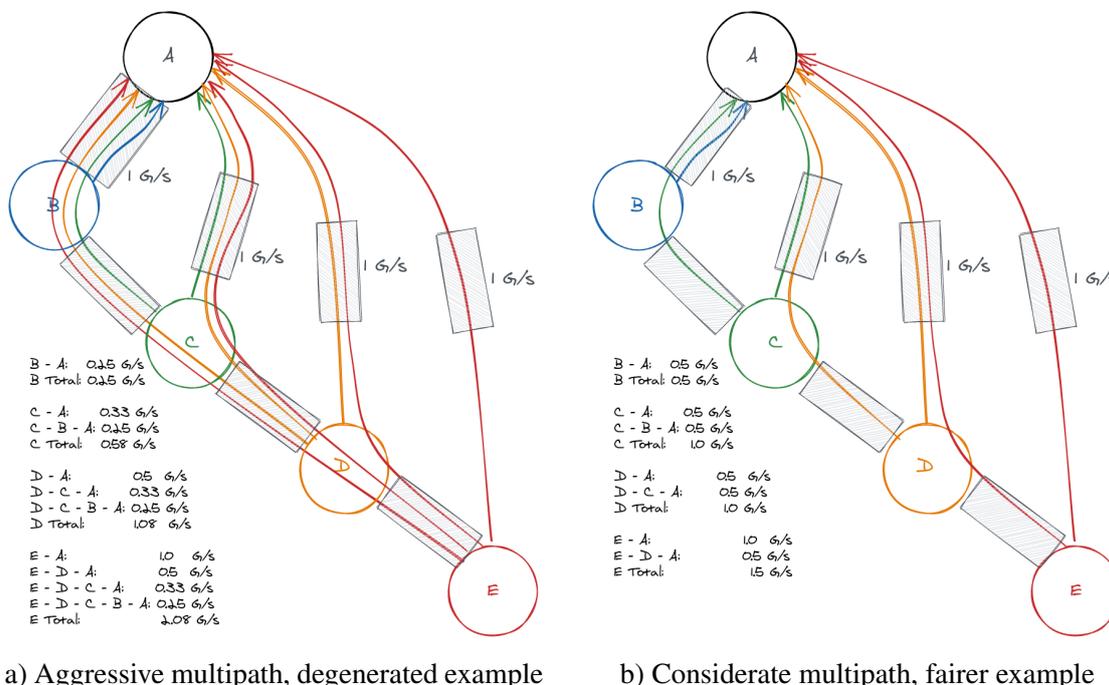


Figure 4: Total network utilization is a constant 4 Gbit/s in both cases

helping to improve global network utilization? On the other hand, does an overly naïve approach to path choice result in an under-utilization of viable alternative routes while compounding traffic congestion issues in common choke points?

3.2 Methodology

To also approach these questions from a practical side, we are currently in the process of designing and implementing a next-generation networking API that translates high-level user demands into actual path choices on the scriptable back-end, leveraging information about path performance and network conditions to automatically optimize for throughput, latency, etc.

4 Summary

Upcoming PAN architectures promise a world with easy and ubiquitous multipath. This requires the reconsideration of a lot of assumptions and thereby poses some interesting challenges. Our notion of the *Cost of Multipath* reflects the observation that network utilization appears to be slightly penalized when compared to traditional shortest-path approaches.

Once the theoretical underpinnings are more firmly in place, we seek to demonstrate the viability of our ideas through a practical implementation of a path-selection engine that incorporates these mechanisms behind a high-level next-generation networking API that echoes current proposals under development at the IETF.

Bibliography

- [AFS07] V. Aggarwal, A. Feldmann, C. Scheideler. Can ISPs and P2P users cooperate for improved performance? *ACM SIGCOMM Computer Communication Review* 37(3):29–40, 2007.
- [CK05] G. Christodoulou, E. Koutsoupias. The price of anarchy of finite congestion games. In *Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*. Pp. 67–73. 2005.
- [EK20] T. Enhardt, C. Krähenbühl. A Vocabulary of Path Properties. Internet-draft draft-irtf-panrg-path-properties-00, Internet Engineering Task Force, Mar 2020. Work in Progress.
<https://datatracker.ietf.org/doc/html/draft-irtf-panrg-path-properties-00>
- [PSRC17] A. Perrig, P. Szalachowski, R. M. Reischuk, L. Chuat. *SCION: a secure Internet architecture*. Springer, 2017.
- [SB09] J. Seedorf, E. Burger. Application-Layer Traffic Optimization (ALTO) Problem Statement. RFC 5693, Oct 2009.
[doi:10.17487/RFC5693](https://doi.org/10.17487/RFC5693)
<https://rfc-editor.org/rfc/rfc5693.txt>
- [TWE⁺21] B. Trammell, M. Welzl, T. Enhardt, G. Fairhurst, M. Kühlewind, C. Perkins, P. S. Tiesel, C. A. Wood, T. Pauly, K. Rose. An Abstract Application Layer Interface to Transport Services. Internet-draft draft-ietf-taps-interface-12, Internet Engineering Task Force, Apr 2021. Work in Progress.
<https://datatracker.ietf.org/doc/html/draft-ietf-taps-interface-12>
- [WY12] D. Wing, A. Yourtchenko. Happy Eyeballs: Success with Dual-Stack Hosts. RFC 6555, Apr 2012.
[doi:10.17487/RFC6555](https://doi.org/10.17487/RFC6555)
<https://rfc-editor.org/rfc/rfc6555.txt>