Path Selection for Multiple Paths In QUIC
draft-dawkins-quic-multipath-selection-01

Abstract

In QUIC working group discussions about proposals to use multiple paths, an obvious question came up - given multiple paths, how does QUIC select paths to send packets over?

The answer to that question may inform decisions in the QUIC working group about the scope of any multipath extensions considered for experimentation and adoption.

This document is intended to summarize aspects of path selection from those contributions and conversations.

It is recognized that path selection is not the only important open question about QUIC Multipath, but other open questions are out of scope for this document.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on December 4, 2021.

Copyright Notice

Copyright (c) 2021 IETF Trust and the persons identified as the document authors. All rights reserved.
1. Introduction

In QUIC working group [QUIC-charter] discussions about proposals to use multiple paths, an obvious question came up - given multiple paths, how does QUIC select paths to send packets over?
The answer to that question may inform decisions in the QUIC working group about the scope of any multipath extensions considered for experimentation and adoption.

This document is intended to summarize aspects of path selection from those contributions and conversations.

It is recognized that path selection is not the only important open question about QUIC Multipath, but other open questions are out of scope for this document.

1.1. Why We Should Look at Path Selection Strategies Now

One of the first questions that’s come up in discussions about multiple paths for QUIC has been about path selection. As soon as an implementation has multiple paths available, it must decide how to use these paths. The [RFC9000] answer, relying on connection migration, is "if you have multiple paths available, you can validate more than one connection at a time, but you only send on one connection at a time, and you migrate to another connection when you decide sending on the current connection is no longer appropriate. How you decide to migrate to another connection is up to you".

That has been a fine answer for many of the implementers who have worked on the first version of QUIC, and have deployed it in their networks. For other implementers, targeting other use cases and other networking environments, it may not be sufficient.

To take only one example, one of several presentations at [QUIC-interim-20-10] described aspects of 3GPP Access Traffic Steering, Switch and Splitting support (ATSSS), which contained four "Steering Modes" as part of Rel-16 in 2019 [TS23501], each of which corresponding roughly to a path selection strategy described in Section 3 of this document. A study on "ATSSS Phase 2" [TR23700-93] included four more Steering Modes for Rel-17, expected to be finalized in mid-2021, and none of these eight (so far) Steering Modes are based on QUIC - they are based on Multipath TCP ([RFC8684] or simple IP packet forwarding. And if that were not enough, a proposal for a study on "ATSSS Phase 3" [S2-2104599] was provided to the SA2 145-e meeting in May 2021. Some of the ATSSS strategies rely in 5G network internals and don’t translate to the broader Internet, but most do translate, and 3GPP participants certainly aren’t the only people thinking about path selection strategies.

Since the various proposals presented at [QUIC-interim-20-10] were developed in isolation from each other, the path selection strategies that they reflect may be similar between proposals, but not quite the
same, and none of the proposals presented had more than two strategies in common with any other proposal.

Given the number of path selection strategies being considered, implemented, and even deployed in any number of venues, and the potential for combinatorial explosion as described in Section 3.10, it seems that identifying common aspects of path selection strategies, sooner rather than later, is important.

1.2. Notes for Readers

This document is an informational Internet-Draft, not adopted by any IETF working group, and does not carry any special status within the IETF.

Please note well that this document reflects the author’s current understanding of past working group discussions and proposals. Contributions that add or improve considerations are welcomed, as described in Section 1.4.

1.3. Minimal Terminology

In this document, "QUIC multipath" is only used as shorthand for "QUIC using multiple paths". It does not refer to a specific proposal.

In this document, "path selection strategy" means the policy that a QUIC sender uses to guide its choice between multiple interfaces of a QUIC connection for "the next packet".

This document adopts three terms, stolen from [TS23501], that seemed helpful in previous discussions about multipath in the QUIC working group.

- Traffic Steering - selecting an initial path (in [RFC9000], this would be "validating a connection, and then using it". Although an [RFC9000] client can validate multiple connections, the client will only use one validated connection at a time.

- Traffic Switching - selecting a different validated path (in [RFC9000], this is something like "migrating to a new validated connection", although whether connection migration as defined in [RFC9000]) would be sufficient is discussed in Section 4).

- Traffic Splitting - using multiple validated paths simultaneously (this would almost certainly require an extension beyond connection migration as defined in [RFC9000]).
"Traffic Steering" does not require any extension to [RFC9000], and is not discussed further in this document. The focus will be on "Traffic Switching" and "Traffic Splitting".

1.4. Contribution and Discussion Venues for this draft.

(Note to RFC Editor - if this document ever reaches you, please remove this section)

This document is under development in the Github repository at https://github.com/SpencerDawkins/quic-multipath-selection.

Readers are invited to open issues and send pull requests with contributed text for this document, but since the document is intended to guide discussion for the QUIC working group, substantial discussion of this document should take place on the QUIC working group mailing list (quic@ietf.org). Mailing list subscription and archive details are at https://www.ietf.org/mailman/listinfo/quic.

2. Background for this document

A number of individual draft proposals for "QUIC over multiple paths" have been submitted to the IETF QUIC and INTAREA working groups, dating back as far as 2017. The author thinks that the complete list is as follows (and reminders for proposals he missed are welcomed):

- [I-D.an-multipath-quic]
- [I-D.an-multipath-quic-application-policy]
- [I-D.an-multipath-quic-traffic-distribution]
- [I-D.chan-quic-owl]
- [I-D.deconinck-quic-multipath]
- [I-D.deconinck-quic-multipath],
- [I-D.huitema-quic-mpath-req]
- [I-D.huitema-quic-mpath-option]
- [I-D.liu-multipath-quic]
- [I-D.piraux-intarea-quic-tunnel-session]

[I-D.bonaventure-iccrg-schedulers] has also been submitted to the Internet Congestion Control Research Group [ICCRG-charter] in the
Internet Research Task Force. It contains specific proposals for implementing some multipath schedulers, and includes some discussion of path selection relevant to this document.

One point of confusion in QUIC working group discussions was that the various proposals (also using Multipath TCP [RFC8684], so not all proposals were QUIC-specific) discussed in working group meetings and on the QUIC mailing list were from various proponents who weren’t solving the same problem. This meant that no two of the use cases presented at the QUIC working group virtual interim on QUIC Multipath [QUIC-interim-20-10] were relying on the same strategies.

It seemed useful to collect the path selection strategies described in those proposals, to look for common elements, and to write them down in one place, to allow more focused discussion. [I-D.dawkins-quic-what-to-do-with-multipath] was intended to summarize, at a high level, various proposals for the use of multipath capabilities in QUIC, both inside the IETF and outside the IETF, in order to identify elements that were common across proposals. This draft tries to describe the impact of these various strategies on potential QUIC Multipath extensions.

One element that is certainly worth considering is whether the path selection strategies that have been proposed can be satisfied using a small number of "building block" strategies.

3. Overview of Proposed Path Selection Strategies

The following strategies were discussed at [QUIC-interim-20-10], and afterwards on the QUIC mailing list. These are summarized in this section, are described in more detail in [I-D.dawkins-quic-what-to-do-with-multipath], and are attributed to various proposals in that document.

- Active-Standby - described in Section 3.1
- Latency Versus Bandwidth - described in Section 3.2
- Bandwidth Aggregation/Load Balancing - described in Section 3.3
- Minimum RTT Difference - described in Section 3.4
- Round-Trip-Time Thresholds - described in Section 3.5
- RTT Equivalence - described in Section 3.6
- Priority-based - described in Section 3.7
3.1. Active-Standby

The traffic associated with a specific flow will be sent via a specific path (the 'active path') and switched to another path (called 'standby path') when the active access is unavailable.

3.2. Latency Versus Bandwidth

Some traffic might be sent over a network path with lower latency and other traffic might be sent over a different network path with higher bandwidth.

3.3. Bandwidth Aggregation/Load Balancing

Traffic is sent using all available paths simultaneously, so that all available bandwidth is utilized, likely based on something like weighted round-robin path selection. This strategy is often used for bulk transfers.

3.4. Minimum RTT Difference

Traffic is sent over the path with the lowest smoothed RTT among all available paths, in order to minimize latency for latency-sensitive flows.

3.5. Round-Trip-Time Thresholds

Traffic is sent over the first path with a smoothed RTT that meets a certain threshold.

3.6. RTT Equivalence

When multiple paths each have sufficiently similar smoothed RTTs, traffic is sent over all of these paths. This is similar to "Bandwidth Aggregation/Load Balancing", with the additional qualification that not all available paths are used for this traffic.

3.7. Priority-based

Priorities are assigned to each path (often by association with network interfaces). Traffic is sent on a highest-priority path...
until it becomes congested, and then "overflows" onto a lower-priority path.

3.8. Redundant

Traffic is replicated over two or more paths. This strategy could be used continuously, but is more commonly used when measured network conditions indicate that redundant sending may be necessary to increase the likelihood that at least one copy of each packet will arrive at the receiver.

3.9. Control Plane Versus Data Plane

An application might stream media over one or more available paths (based on one of the other strategies named in this section), but might send ACK traffic or retransmission over a path specifically chosen for that purpose. This is more likely to be beneficial if the path characteristics differ significantly between available paths - for example, satellite uplink/downlink stations connected by both higher-bandwidth Low Earth Orbit satellite paths and lower-bandwidth cellular or landline paths.

3.10. Combinations of Strategies

In addition to the strategies described above, it is also possible to combine these strategies. For example, a scheduler might use load-balancing over three paths, but when one of the paths becomes unavailable, the scheduler might switch to the two paths that are still available, in a way similar to Active-Standby. This is very much an example chosen at random - potentially, there are many combinations that could be useful.

4. Implications for QUIC Multipath

This section summarizes potential implications for "Multipath QUIC" of path selection strategies described in Section 3, dividing them between "Traffic Switching" (Section 4.1) and "Traffic Splitting" (Section 4.2).

4.1. Selecting a Single Path Among Multiple Validated Paths ("Traffic Switching")

If a sender using Active-Standby (described in Section 3.1) does not perform frequent path switching, it can likely be supported using connection migration as defined in [RFC9000] without change.

- The caveat here is that connection migration can include the also-implicit assumption that an endpoint can free up resources
associated with the previously-active path. If connection migration happens often enough, the endpoint may spend considerable time "thrashing" between allocating resources and quickly freeing them. Of course, if a sender is frequently selecting a new path for connection migration, this probably degenerates into one of the other path selection strategies.

Some path selection strategies could be supported by a mechanism as simple as the one proposed in [I-D.huitema-quic-mpath-option], which replaces "the implicit signaling of path migration through data transmission, by means of a new PATH_OPTION frame" (this isn't intended to imply the proposal is simple, only the explicit signaling), if the receiver uses this option to notify the sender of the preferred path. For example, Minimum RTT Difference (described in Section 3.4) and Round-Trip-Time Thresholds (described in Section 3.5) likely fall into this category.

Some path selection strategies are exploiting a relatively long-lived difference between paths - for example, Latency Versus Bandwidth (described in Section 3.2), Priority-based (described in Section 3.7), and Control Plane Versus Data Plane (described in Section 3.9) may fall into this category. One might wonder why these senders would need to use a single "multipath connection", rather than multiple [RFC9000] connections, for these cases, but if there is a reason to use a single multipath connection, a mechanism similar to the one proposed in [I-D.huitema-quic-mpath-option] could also be used in these cases.

4.2. Selecting Multiple Active Paths ("Traffic Splitting")

Some path selection strategies are treating more than one path as a set of active paths, whether the sender is performing "Traffic Splitting" (as defined in Section 1.3), as is the case for Bandwidth Aggregation/Load Balancing (described in Section 3.3) and RTT Equivalence (described in Section 3.6), or simply transmitting the same packet across multiple paths, as is the case for Redundant (described in Section 3.8).

For these cases, a more complex mechanism is likely required.

4.3. Arbitrary Combinations

Because it is simple enough to imagine various combinations of strategies (as described in Section 3.10), it seems important to understand what basic building blocks are required in order to support the strategies that seem common across a variety of use cases, because interactions between strategies may have significant
implications for QUIC Multipath that might not arise when considering strategies in isolation.

This seems especially important because existing proposals for QUIC Multipath don’t use the same vocabulary to describe path selection strategies, so implementations may not behave in the same way, even if they are each using a strategy that seems to be common.

5. Next Steps

If this discussion is useful, it may also be useful to take the next step, and identify potential building blocks that can be used to construct the path selection strategies described in Section 4.1 and Section 4.2.

6. IANA Considerations

This document does not make any request to IANA.

7. Security Considerations

QUIC-specific security considerations are discussed in Section 21 of [RFC9000].

Section 6 of [I-D.ietf-quic-datagram] discusses security considerations specific to the use of the Unreliable Datagram Extension to QUIC.

Some "Multipath QUIC"-specific security considerations can be found in the corresponding section of [I-D.deconinck-quic-multipath].

Having said that, it may be best to repeat the security considerations section from [I-D.huitema-quic-mpath-option]: "TBD. There are probably ways to abuse this."

8. Acknowledgments

Your name could appear here. Please comment and contribute, as per Section 1.4.

9. Informative References

[I-D.an-multipath-quic]


Huitema, C., "QUIC Multipath Negotiation Option", draft-huitema-quic-mpath-option-00 (work in progress), October 2020.


Piraux, M., Bonaventure, O., and A. Masputra, "Session mode for multiple QUIC Tunnels", draft-piraux-intarea-quic-tunnel-session-00 (work in progress), November 2020.

ICCRG-charter

QUIC-charter

QUIC-interim-20-10


Author's Address

Spencer Dawkins
Tencent America LLC

Email: spencerdawkins.ietf@gmail.com
Gateway Based Trust Relationship Between the Endpoint and the Intermediate Node
draft-du-panrg-gateway-based-trust-relationship-01

Abstract

This document describes a mechanism about establishing trust relationship between the endpoint and the intermediate node along the path based on the gateway of the endpoint.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on July 4, 2022.

Copyright Notice

Copyright (c) 2021 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (https://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents...
1. Introduction

In future, many new services would emerge in the network, such as the 5G URLLC (Ultra Reliable Low Latency Communication) service, and the holographic type communications. Many of the new services need a higher QoS (Quality of Service) level than the current Internet services, and some of them have a critical SLA (Service-Level Agreement) requirement. The SLA differences between the new services and traditional services would become larger and lager. However, current networks can only provide the Best Effort bearing, in which all the traffic are treated as the same kind. In summary, current networks are short of negotiation abilities between the network and the applications. PANRG in the IRTF has proposed a research direction to enable the path aware networking. A lot of analyses have been done in the [RFC9049], which explains reasons why various Path Aware techniques have seen limited or no deployment.

One of the reasons is that it is hard to establish a trust relationship between the Endpoint and Intermediate Node. In the current network structure, the Endpoints only needs to be aware of the each other, and assume that the network can provide a good connection service for them. On the other hand, traditionally, Intermediate Nodes only need to support IP forwarding and do not need to be aware of up-layer information. In addition, the network nodes work in a per-packet model, not a per-flow model. Also in the [RFC9049], it is said that "per-connection state in intermediate nodes has been an impediment to adoption and deployment".

However, we can find that the gateway of the Endpoint is able to maintain a per-connection state and a trust-relationship for each
user. For example, the users in the fixed network need to be authorized by the BNG (Broadband Network Gateway), and the BNG also needs to do the accounting for each user. It is hard and unnecessary to make every intermediate node along the path has the same ability as the BNG; however, if they can have some communication with the BNG, perhaps they can make a better path choice for the user. Following this direction, this document proposes a mechanism about how to enable the communication between the BNG and the Head-End node in the network, because the Head-End node is the main node to select the path for a flow in the network. If any future work on the trust relationship between the Endpoint and the Intermediate Node is considered, the mechanism in this document can be a reference.

2. Proposed Mechanism for the Trust Problem

As shown in the Figure 1, in the fixed network, the BNG works as the gateway for the Client, and provides the Internet connection service for the Applications. The Client and Server are the EndPoints, and the BNG, Head-End, Mid-Node, End-Node are the nodes along the path from the Client to the Server. There are three paths, i.e., A, B, C, with different properties such as high bandwidth or low latency, between the Head-End and the End-Node in the network.

By default, all the traffic from the APPs are forwarded from the Head-End to the End-Node with the same treatment in the network. In the Head-end, perhaps a load balance mechanism can be enabled, but normally without any per-flow mechanism, because the Head-End does not know the requirements of each flow. If the Applications need different treatments in the network, and the Head-End can schedule the traffic to a proper path, the user can have a better experience and the network resource can be used more efficiently.

![Path-aware Mechanism in the Fixed Network](image)

The following paragraphs are about the trust problems and the potential solutions for them.
The first problem is the path information collection for the Endpoints. The Endpoints should be able to trust the path information that the Intermediate Nodes signal. As a first step, we only consider the situation that information is limited and does not need to be updated frequently. In this case, if the Head-End needs to inform the Endpoints something, it can send the information with its signature generated by using a private key. The Endpoints can check the information using the corresponding public key. For example, the public key can be obtained by the Endpoint in the authentication procedure.

The second problem is the Head-End should trust the Endpoints if it receives some path selection suggestions from the Endpoints. In this case, we think that the BNG has authenticated the Endpoints, so that the BNG can send some information to the Head-End indicating that the Endpoint is not a fake one. For example, the BNG and the Head-End can using an IPSec to transfer the traffic that needs specific treatment. Another option is that the BNG can forward the traffic that needs specific treatment with its signature generated by using a private key. The Head-End can check the information using the corresponding public key of the BNG.

The reason that we do not suggest that the Endpoints make the signature is because their number is much larger than the number of BNGs. We do not think the Head-End can handle a large number of keys. Meanwhile, in this mechanism, the Intermediate Node does not need to maintain per-connection state.

3. IANA Considerations

TBD.

4. Security Considerations

TBD.

5. Acknowledgements

TBD.

6. References

6.1. Normative References

6.2. Informative References


Authors’ Addresses

Zongpeng Du
China Mobile
No.32 XuanWuMen West Street
Beijing 100053
China

Email: duzongpeng@foxmail.com

Peng Liu
China Mobile
No.32 XuanWuMen West Street
Beijing 100053
China

Email: liupengyjy@chinamobile.com
Abstract

Path properties express information about paths across a network and the services provided via such paths. In a path-aware network, path properties may be fully or partially available to entities such as endpoints. This document defines and categorizes path properties. Furthermore, the document specifies several path properties which might be useful to endpoints or other entities, e.g., for selecting between paths or for invoking some of the provided services.

Discussion Venues

This note is to be removed before publishing as an RFC.

Discussion of this document takes place on the "Path-Aware Networking Research Group" mailing list (PANRG), which is archived at https://mailarchive.ietf.org/arch/browse/panrg/. Subscription information is at https://www.ietf.org/mailman/listinfo/panrg/.

Source for this draft and an issue tracker can be found at https://github.com/panrg/path-properties/.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 8 September 2022.
1. Introduction

The current Internet architecture does not explicitly support endpoint discovery of forwarding paths through the network as well as the discovery of properties and services associated with these paths. Path-aware networking, as defined in Section 1.1 of [I-D.irtf-panrg-questions], describes "endpoint discovery of the properties of paths they use for communication across an internetwork, and endpoint reaction to these properties that affects routing and/or data transfer". This document provides a generic definition of path properties, addressing the first of the questions in path-aware networking [I-D.irtf-panrg-questions].
As terms related to paths have been used with different meanings in different areas of networking, first, this document provides a common terminology to define paths, path elements, and flows. Based on these terms, the document defines path properties. Then, this document provides some examples of use cases for path properties. Finally, the document lists several path properties that may be useful for the mentioned use cases.

Note that this document does not assume that any of the listed path properties are actually available to any entity. The question of how entities can discover and distribute path properties in a trustworthy way is out of scope for this document.

2. Terminology

Entity: A physical or virtual device or function, or a collection of devices or functions, which plays a role related to path-aware networking for particular paths and flows. An entity can be on-path or off-path: On the path, an entity may participate in forwarding the flow, i.e., what may be called data plane functionality. On or off the path, an entity may influence aspects of how the flow is forwarded, i.e., what may be called control plane functionality, such as Path Selection or Service Invocation. An entity influencing forwarding aspects is usually aware of path properties, e.g., by observing or measuring them or by learning them from another entity.

Node: An on-path entity which processes packets, e.g., sends, receives, forwards, or modifies them. A node may be physical or virtual, e.g., a physical device, a service function provided as a virtual element, or even a single queue within a switch. A node may also be an entity which consists of a collection of devices or functions, e.g., an entire Autonomous System (AS).

Link: A medium or communication facility that connects two or more nodes with each other. A link enables a node to send packets to other nodes. Links can be physical, e.g., a Wi-Fi network which connects an Access Point to stations, or virtual, e.g., a virtual switch which connects two virtual machines hosted on the same physical machine. A link is unidirectional. As such, bidirectional communication can be modeled as two links between the same nodes in opposite directions.

Path element: Either a node or a link. For example, a path element can be an Abstract Network Element (ANE) as defined in [I-D.ietf-alto-path-vector].

Path: A sequence of adjacent path elements over which a packet can
be transmitted, starting and ending with a node. A path is unidirectional. Paths are time-dependent, i.e., the sequence of path elements over which packets are sent from one node to another may change. A path is defined between two nodes. For multicast or broadcast, a packet may be sent by one node and received by multiple nodes. In this case, the packet is sent over multiple paths at once, one path for each combination of sending and receiving node; these paths do not have to be disjoint. Note that an entity may have only partial visibility of the path elements that comprise a path and visibility may change over time. Different entities may have different visibility of a path and/or treat path elements at different levels of abstraction. For example, a path may be given as a sequence of physical nodes and the links connecting these nodes, or it may be given as a sequence of logical nodes such as a sequence of ASes or an Explicit Route Object (ERO). Similarly, the representation of a path and its properties, as it is known to a specific entity, may be more complex and include details about the physical layer technology, or it may be more abstract and only consist of a specific source and destination which is known to be reachable from that source.

Endpoint: The endpoints of a path are the first and the last node on the path. For example, an endpoint can be a host as defined in [RFC1122], which can be a client (e.g., a node running a web browser) or a server (e.g., a node running a web server).

Reverse Path: The path that is used by a remote node in the context of bidirectional communication.

Subpath: Given a path, a subpath is a sequence of adjacent path elements of this path.

Flow: One or multiple packets to which the traits of a path or set of subpaths may be applied in a functional sense. For example, a flow can consist of all packets sent within a TCP session with the same five-tuple between two hosts, or it can consist of all packets sent on the same physical link.

Property: A trait of one or a sequence of path elements, or a trait
of a flow with respect to one or a sequence of path elements. An example of a link property is the maximum data rate that can be sent over the link. An example of a node property is the administrative domain that the node belongs to. An example of a property of a flow with respect to a subpath is the aggregated one-way delay of the flow being sent from one node to another node over this subpath. A property is thus described by a tuple containing the path element(s), the flow or an empty set if no packets are relevant for the property, the name of the property (e.g., maximum data rate), and the value of the property (e.g., 1Gbps).

Aggregated property: A collection of multiple values of a property into a single value, according to a function. A property can be aggregated over multiple path elements (i.e., a subpath), e.g., the MTU of a path as the minimum MTU of all links on the path, or over multiple packets (i.e., a flow), e.g., the median one-way latency of all packets between two nodes, or over both, e.g., the mean of the queueing delays of a flow on all nodes along a path. The aggregation function can be numerical, e.g., median, sum, minimum, it can be logical, e.g., "true if all are true", "true if at least 50% of values are true", or an arbitrary function which maps multiple input values to an output value.

Observed property: A property that is observed for a specific path element, subpath, or path, e.g., using measurements. For example, the one-way delay of a specific packet transmitted from one node to another node can be measured.

Assessed property: An approximate calculation or assessment of the value of a property. An assessed property includes the reliability of the calculation or assessment. The notion of reliability depends on the property. For example, a path property based on an approximate calculation may describe the expected median one-way latency of packets sent on a path within the next second, including the confidence level and interval. A non-numerical assessment may instead include the likelihood that the property holds.

2.1. Terminology usage for specific technologies

The terminology defined in this document is intended to be general and applicable to existing and future path-aware technologies. Using this terminology, a path-aware technology can define and consider specific path elements and path properties on a specific level of abstraction. For instance, a technology may define path elements as IP routers, e.g., in source routing ([RFC1940]). Alternatively, it may consider path elements on a different layer of the Internet.
Architecture ([RFC1122]) or as a collection of entities not tied to a specific layer, such as an AS or an ERO. Even within a single path-aware technology, specific definitions might differ depending on the context in which they are used. For example, the endpoints might be the communicating hosts in the context of the transport layer, ASes that contain the hosts in the context of routing, or specific applications in the context of the application layer.

3. Use Cases for Path Properties

When a path-aware network exposes path properties to endpoints or other entities, these entities may use this information to achieve different goals. This section lists several use cases for path properties.

Note that this is not an exhaustive list, as with every new technology and protocol, novel use cases may emerge, and new path properties may become relevant. Moreover, for any particular technology, entities may have visibility of and control over different path elements and path properties, and consider them on different levels of abstraction. Therefore, a new technology may implement an existing use case related to different path elements or on a different level of abstraction.

3.1. Path Selection

Nodes may be able to send flows via one (or a subset) out of multiple possible paths, and an entity may be able to influence the decision which path(s) to use. Path Selection may be feasible if there are several paths to the same destination (e.g., in case of a mobile device with two wireless interfaces, both providing a path), or if there are several destinations, and thus several paths, providing the same service (e.g., Application-Layer Traffic Optimization (ALTO) [RFC5693], an application layer peer-to-peer protocol allowing endpoints a better-than-random peer selection). Care needs to be taken when selecting paths based on path properties, as path properties that were previously measured may not be helpful in predicting current or future path properties and such path selection may lead to unintended feedback loops.

Entities may select their paths to fulfill a specific goal, e.g., related to security or performance. As an example of security-related path selection, an entity may allow or disallow sending flows over paths involving specific networks or nodes to enforce traffic policies. In an enterprise network where all traffic has to go through a specific firewall, a path-aware entity can implement this policy using path selection. As an example of performance-related path selection, an entity may prefer paths with performance
properties that best match application requirements. For example, for sending a small delay sensitive query, the entity may select a path with a short One-Way Delay, while for retrieving a large file, it may select a path with high Link Capacities on all links. Note, there may be trade-offs between path properties (e.g., One-Way Delay and Link Capacity), and entities may influence these trade-offs with their choices. As a baseline, a path selection algorithm should aim to not perform worse than the default case most of the time.

Path selection can be done either by the communicating node(s) or by other entities within the network: A network (e.g., an AS) can adjust its path selection for internal or external routing based on path properties. In BGP, the Multi Exit Discriminator (MED) attribute is used in the decision-making process to select which path to choose among those having the same AS PATH length and origin [RFC4271]; in a path-aware network, instead of using this single MED value, other properties such as Link Capacity or Link Usage could additionally be used to improve load balancing or performance [I-D.ietf-idr-performance-routing].

3.2. Protocol Selection

Before sending data over a specific path, an entity may select an appropriate protocol or configure protocol parameters depending on path properties. For example, an endpoint may cache state on whether a path allows the use of QUIC [I-D.ietf-quic-transport] and if so, first attempt to connect using QUIC before falling back to another protocol when connecting over this path again. A video streaming application may choose an (initial) video quality based on the achievable data rate or the monetary cost of sending data (e.g., volume-base or flat-rate cost model).

3.3. Service Invocation

In addition to path or protocol selection, an entity may choose to invoke additional functions in the context of Service Function Chaining [RFC7665], which may influence what nodes are on the path. For example, a 0-RTT Transport Converter [I-D.ietf-tcpm-converters] will be involved in a path only when invoked by an endpoint; such invocation will lead to the use of MPTCP or TCPinc capabilities while such use is not supported via the default forwarding path. Another example is a connection which is composed of multiple streams where each stream has specific service requirements. An endpoint may decide to invoke a given service function (e.g., transcoding) only for some streams while others are not processed by that service function.
4. Examples of Path Properties

This Section gives some examples of path properties which may be useful, e.g., for the use cases described in Section 3.

Within the context of any particular technology, available path properties may differ as entities have insight into and are able to influence different path elements and path properties. For example, an endpoint may have some visibility into path elements that are on a low level of abstraction and close, e.g., individual nodes within the first few hops, or it may have visibility into path elements that are far away and/or on a higher level of abstraction, e.g., the list of ASes traversed. This visibility may depend on factors such as the physical or network distance or the existence of trust or contractual relationships between the endpoint and the path element(s). A path property can be defined relative to individual path elements, a sequence of path elements, or "end-to-end", i.e., relative to a path that comprises of two endpoints and a single virtual link connecting them.

Path properties may be relatively dynamic, e.g., the one-way delay of a packet sent over a specific path, or non-dynamic, e.g., the MTU of an Ethernet link which only changes infrequently. Usefulness over time differs depending on how dynamic a property is: The merit of a momentary measurement of a dynamic path property diminishes greatly as time goes on, e.g. the merit of an RTT measurement from a few seconds ago is quite small, while a non-dynamic path property might stay relevant for a longer period of time, e.g. a NAT typically stays on a specific path during the lifetime of a connection involving packets sent over this path.

Access Technology: The physical or link layer technology used for transmitting or receiving a flow on one or multiple path elements. If known, the Access Technology may be given as an abstract link type, e.g., as Wi-Fi, Wired Ethernet, or Cellular. It may also be given as a specific technology used on a link, e.g., 2G, 3G, 4G, or 5G cellular, or 802.11a, b, g, n, or ac Wi-Fi. Other path elements relevant to the access technology may include nodes related to processing packets on the physical or link layer, such as elements of a cellular backbone network. Note that there is no common registry of possible values for this property.

Monetary Cost: The price to be paid to transmit or receive a specific flow across a network to which one or multiple path elements belong.

Service function: A service function that a path element applies to
a flow, see [RFC7665]. Examples of abstract service functions include firewalls, Network Address Translation (NAT), and TCP optimizers. Some stateful service functions, such as NAT, need to observe the same flow in both directions, e.g., by being an element of both the path and the reverse path.

Transparency: When a node performs an action A on a flow F, the node is transparent to F with respect to some (meta-)information M if the node performs A independently of M. M can for example be the existence of a protocol (header) in a packet or the content of a protocol header, payload, or both. A can for example be blocking packets or reading and modifying (other protocol) headers or payloads. Transparency can be modeled using a function f, which takes as input F and M and outputs the action taken by the node. If a taint analysis shows that the output of f is not tainted (impacted) by M or if the output of f is constant for arbitrary values of M, then the node is considered to be transparent. An IP router could be transparent to transport protocol headers such as TCP/UDP but not transparent to IP headers since its forwarding behavior depends on the IP headers. A firewall that only allows outgoing TCP connections by blocking all incoming TCP SYN packets regardless of their IP address is transparent to IP but not to TCP headers. Finally, a NAT that actively modifies IP and TCP/UDP headers based on their content is not transparent to either IP or TCP/UDP headers. Note that according to this definition, a node that modifies packets in accordance with the endpoints, such as a transparent HTTP proxy, as defined in [RFC2616], and a node listening and reacting to implicit or explicit signals, see [RFC8558], are not considered transparent.

Administrative Domain: The identity of an individual or an organization that owns a path element (or several path elements). Examples of administrative domains are an IGP area, an AS, or a service provider network.

Routing Domain Identifier: An identifier indicating the routing domain of a path element. Path elements in the same routing domain are in the same administrative domain and use a common routing protocol to communicate with each other. An example of a routing domain identifier is the globally unique autonomous system number (ASN) as defined in [RFC1930].

Disjointness: For a set of two paths or subpaths, the number of
shared path elements can be a measure of intersection (e.g., Jaccard coefficient, which is the number of shared elements divided by the total number of elements). Conversely, the number of non-shared path elements can be a measure of disjointness (e.g., 1 - Jaccard coefficient). A multipath protocol might use disjointness as a metric to reduce the number of single points of failure.

Symmetric Path: Two paths are symmetric if the path and its reverse path consist of the same path elements on the same level of abstraction, but in reverse order. For example, a path which consists of layer 3 switches and links between them and a reverse path with the same path elements but in reverse order are considered "routing" symmetric, as the same path elements on the same level of abstraction (IP forwarding) are traversed in the opposite direction.

Path MTU: The maximum size, in octets, of an IP packet that can be transmitted without fragmentation.

Transport Protocols available: Whether a specific transport protocol can be used to establish a connection over a path or subpath, e.g., whether the path is QUIC-capable or MPTCP-capable, based on cached knowledge.

Protocol Features available: Whether a specific protocol feature is available over a path or subpath, e.g., Explicit Congestion Notification (ECN), or TCP Fast Open.

Some path properties express the performance of the transmission of a packet or flow over a link or subpath. Such transmission performance properties can be measured or approximated, e.g., by endpoints or by path elements on the path, or they may be available as cost metrics, see [I-D.ietf-alto-performance-metrics]. Transmission performance properties may be made available in an aggregated form, such as averages or minimums. Properties related to a path element which constitutes a single layer 2 domain are abstracted from the used physical and link layer technology, similar to [RFC8175].

Link Capacity: The link capacity is the maximum data rate at which data that was sent over a link can correctly be received at the node adjacent to the link. This property is analogous to the link capacity defined in [RFC5136] but not restricted to IP-layer traffic.

Link Usage: The link usage is the actual data rate at which data
that was sent over a link is correctly received at the node adjacent to the link. This property is analogous to the link usage defined in [RFC5136] but not restricted to IP-layer traffic.

One-Way Delay: The one-way delay is the delay between a node sending a packet and another node on the same path receiving the packet. This property is analogous to the one-way delay defined in [RFC7679] but not restricted to IP-layer traffic.

One-Way Delay Variation: The variation of the one-way delays within a flow. This property is similar to the one-way delay variation defined in [RFC3393] but not restricted to IP-layer traffic and defined for packets on the same flow instead of packets sent between a source and destination IP address.

One-Way Packet Loss: Packets sent by a node but not received by another node on the same path after a certain time interval are considered lost. This property is analogous to the one-way loss defined in [RFC7680] but not restricted to IP-layer traffic. Metrics such as loss patterns [RFC3357] and loss episodes [RFC6534] can be expressed as aggregated properties.

5. Security Considerations

If entities are basing policy or path selection decisions on path properties, they need to rely on the accuracy of path properties that other devices communicate to them. In order to be able to trust such path properties, entities may need to establish a trust relationship or be able to verify the authenticity, integrity, and correctness of path properties received from another entity.

Security related properties such as confidentiality and integrity protection of payloads are difficult to characterize since they are only meaningful with respect to a threat model which depends on the use case, application, environment, and other factors. Likewise, properties for trust relations between entities cannot be meaningfully defined without a concrete threat model, and defining a threat model is out of scope for this draft. Properties related to confidentiality, integrity, and trust are orthogonal to the path terminology and path properties defined in this document. Such properties are tied to the communicating nodes and the protocols they use (e.g., client and server using HTTPS, or client and remote network node using VPN) while the path is typically oblivious to them. Intuitively, the path describes what function the network applies to packets, while confidentiality, integrity, and trust describe what function the communicating parties apply to packets.
6. IANA Considerations

This document has no IANA actions.

7. Informative References

[I-D.ietf-alto-path-vector]

[I-D.ietf-alto-performance-metrics]

[I-D.ietf-idr-performance-routing]

[I-D.ietf-quic-transport]

[I-D.ietf-tcpm-converters]
[I-D.irtf-panrg-questions]
Trammell, B., "Current Open Questions in Path Aware Networking", Work in Progress, Internet-Draft, draft-irtf-panrg-questions-12, 25 January 2022,

DOI 10.17487/RFC1122, October 1989,

BCP 6, RFC 1930, DOI 10.17487/RFC1930, March 1996,

DOI 10.17487/RFC1940, May 1996,

DOI 10.17487/RFC2616, June 1999,


[RFC3393] Demichelis, C. and P. Chimento, "IP Packet Delay Variation Metric for IP Performance Metrics (IPPM)", RFC 3393,
DOI 10.17487/RFC3393, November 2002,

DOI 10.17487/RFC4271, January 2006,

RFC 5136, DOI 10.17487/RFC5136, February 2008,
Acknowledgments

Thanks to the Path-Aware Networking Research Group for the discussion and feedback. Specifically, thanks to Mohamed Boudacair for the detailed review and various text suggestions, thanks to Brian Trammell for suggesting the flow definition, thanks to Adrian Perrig and Matthias Rost for the detailed feedback, thanks to Paul Hoffman for the editorial changes, thanks to Luis M. Contreras and Jake Holland for the reviews, and thanks to Spencer Dawkins for the comments and suggestions.

Authors’ Addresses