PANRG
Internet-Draft
Intended status: Informational
Expires: May 7, 2020

T. Enghardt
TU Berlin
C. Kraehenbuehl
ETH Zuerich

November 04, 2019

A Vocabulary of Path Properties
draft-enghardt-panrg-path-properties-03

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 hosts. This document defines and categorizes a set of path properties.

Furthermore, the document specifies several path properties which might be useful to hosts or other entities (e.g., for placing communications or invoking some of the provided services).

Commentaire [Med1]: To make it clear the document is not about identifying an exhaustive list.

Table of Contents

1. Introduction

In the current Internet architecture, hosts generally do not have information about forwarding paths through the network and about services associated with these paths. A path-aware network, as introduced in [I-D.irtf-panrg-questions], exposes information about paths to hosts or to other entities. This document defines such information as path properties, addressing the first of the questions in path-aware networking [I-D.irtf-panrg-questions].

As terms related to paths have different meanings in different areas of networking, first, this document provides a common base terminology to define paths, path elements, and path properties. Then, this document provides some examples for use cases for path properties. Finally, the document lists several path properties that may be useful for the mentioned use cases.

2. Terminology

Node: An entity which processes packets, e.g., sends, receives, forwards, or modifies them. A node may be physical or virtual, e.g., a virtual machine or a service function. A node may also be the collection of multiple entities which, as a collection, processes packets, e.g., an entire Autonomous System (AS).

Host: A node that generally executes application programs on behalf of user(s), employing network and/or Internet communication services in support of this function, as defined in [RFC1122].

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, ~~and~~a bidirectional communication can be modeled as two links between the same nodes in opposite directions.

Path element: Either a node or a link.

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 ~~frequently~~. A path is defined between two nodes for unicast communications. 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 entities may treat path elements at different levels of abstraction. Hosts involved in a communication may not have the same path visibility on the path(s) to place communications between these two hosts. The paths visible to a host are volatile; that is, a host may start a communication with one single path available but multi-paths may be available during the lifetime of a connection.

Commentaire [Med2]: No need to mention this as this will depend on each path.

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

Symmetric path: Denotes the case where the reverse path can be inferred from the path. For example, if the path is composed of path elements (PE1, PE2, PE3), the reverse path is (PE3, PE2, PE1).

Subpath: Given a path, ~~a subpath~~a segment of a path. That is ~~is~~ a partial sequence of adjacent path elements of this path.

Flow: An entity made of 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, 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.

Discovered node: A node that is advertised by a path-aware network and is discovered by a host. A host can decide to invoke a discovered node in none, some, or all its communications according to local policies. The characterization of such policies is out of scope.

3. Use Cases for Path Properties

When a path-aware network exposes path properties to hosts 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.

3.1. Performance Monitoring and Enhancement

Network operators can observe path properties (e.g., measured by on-path devices), to monitor Quality of Service (QoS) characteristics of recent end-user traffic on a path or subpath through their network. Such properties may help identify potential performance problems or trigger countermeasures to enhance performance. Robots and probes may be located in strategic network points to assess these properties.

Note that performance measurement may also be undertaken by hosts when a communication is active. The outcome of such measurement will feed traffic scheduling and path selection (see Section 3.2).

3.2. Path Selection

Entities can choose what traffic to send over which path or subset of paths. 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 traffic 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 host can implement this policy using path selection, in which case the host needs to be aware of paths involving that firewall.

As an example of performance-related path selection, an entity may prefer paths with performance properties that best match its traffic, e.g., retrieving a small webpage as quickly as possible over a path with short One-Way Delays in both directions, or retrieving a large file over 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 both by hosts and by entities within the network: A network (e.g., an AS) can adjust its path selection for internal or external routing based on the path properties. In BGP, the Multi Exit Discriminator (MED) attribute is used to fed the decision-making process to select ~~decides~~ which path to choose ~~if other attributes~~ among those having the same AS PATH length and origin [RFC4271]~~are equal~~; in a path aware network, instead of using this single MED value, other properties such as maximum or available/expected data rate could additionally be used to improve load balancing or the one which optimizes the OWD (e.g., [I-D.ietf-idr-performance-routing]). A host might be able to select between a set of paths, either if there are several paths to the same destination (e.g., if the host is 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 hosts 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 have become outdated and, thus, useless to predict the path properties of packets sent now.

3.3. ~~Traffic Configuration~~Host-initiated Service Invocation

Commentaire [Med3]: I suggest to change the title to better reflect the content

When sending traffic over a specific path, entities can adjust this traffic based on the properties of the path. For example, an entity may select an appropriate protocol depending on the capabilities of the on-path devices, or adjust protocol parameters to an existing path. An example of traffic configuration is a video streaming application choosing an (initial) video quality based on the achievable data rate, or the monetary cost to send data across a network, eventually on a given path, using a volume-based or flat-rate cost model.

Conversely, the selection of a protocol may influence the ~~devices~~nodes that will be involved in a path. For example, a 0-RTT Transport Converter [I-D.ietf-tcpm-converters] will be involved in a path only

when invoked by a host; 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 of traffic policies is a connection which may be composed of multiple streams; each stream with specific service requirements. A host 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~~ path ~~Properties~~ properties which may be useful, e.g., for the use cases described in Section 3.

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~~ 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.

From the point of view of a host, path properties may relate to path elements close to the host, i.e., within the first few hops, or they may include path elements far from the host, e.g., list of ASes traversed. The visibility of path properties to a specific entity may depend on factors such as the physical or network distance or the existence of trust or contractual relationships between the entity and the path element(s).

Furthermore, entities may or may not be able to influence the path elements on their path and their path properties. For example, a user might select between multiple potential adjacent links by selecting between multiple available Wi-Fi Access Points. Or when connected to an Access Point, the user may move closer to enable their device to use a different access technology, potentially increasing the data rate available to the device. Another example is a user changing their data plan to reduce the Monetary Cost to transmit a given amount of data across a network.

Access Technology: The physical or link layer technology used for transmitting or receiving a flow on one or multiple path elements. The Access Technology may be given in an abstract way, e.g., as a Wi-Fi, Wired Ethernet, or Cellular link. It may also be given as a specific technology, e.g., as a 2G, 3G, 4G, or 5G cellular link, or an 802.11a, b, g, n, or ac Wi-Fi link. Other path elements

relevant to the access technology may include on-path devices, 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 to 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 (e.g., NAT) requires the same instance is involved in both directions (path and the reverse path).

Administrative Domain: The administrative domain, e.g., the ICP area, AS, or Service provider network to which a path element or subpath belongs.

Disjointness: For a set of two paths, 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 of paths as a metric to reduce the number of single points of failure.

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

Transport Protocols available: Whether a specific transport protocol can be used to establish a connection over a path or subpath. A host may cache its knowledge about recent successfully established connections using specific protocols, e.g., ~~a~~ cache locally visible QUIC-capable ~~connection,~~ or ~~an~~ MPTCP-capable paths ~~subflow~~. Such cache may also be fed by the outcome of probes that might be sent by a host, for example, upon bootstrapping to assess the compatibility of its available paths with packet transfer using specific transport protocols.

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 hosts or by path elements on the path. They might be made available in an aggregated form, such as averages or minimums. See [ANRW18-Metrics] for a discussion of how to measure some transmission performance properties at the host. 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].

Commentaire [Med4]: to cover the roaming case where you may also pay when you receive.

Mis en forme : Surlignage

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 nodes 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, nodes may need to establish a trust relationship or be able to verify the authenticity, integrity, and correctness of path properties received from another node.

6. IANA Considerations

This document has no IANA actions.

7. Informative References

[ANRW18-Metrics]
Enghardt, T., Tiesel, P., and A. Feldmann, "Metrics for access network selection", Proceedings of the Applied Networking Research Workshop on - ANRW '18, DOI 10.1145/3232755.3232764, 2018.

[I-D.ietf-tcpm-converters]
Bonaventure, O., Boucadair, M., Gundavelli, S., Seo, S., and B. Hesmans, "0-RTT TCP Convert Protocol", draft-ietf-tcpm-converters-13 (work in progress), October 2019.

[I-D.irtf-panrg-questions]
Trammell, B., "Current Open Questions in Path Aware Networking", draft-irtf-panrg-questions-03 (work in progress), October 2019.

[RFC1122] Braden, R., Ed., "Requirements for Internet Hosts - Communication Layers", STD 3, RFC 1122, DOI 10.17487/RFC1122, October 1989, <https://www.rfc-editor.org/info/rfc1122>.

[RFC3357] Koodli, R. and R. Ravikanth, "One-way Loss Pattern Sample Metrics", RFC 3357, DOI 10.17487/RFC3357, August 2002, <https://www.rfc-editor.org/info/rfc3357>.

[RFC3393] Demichelis, C. and P. Chimento, "IP Packet Delay Variation Metric for IP Performance Metrics (IPPM)", RFC 3393, DOI 10.17487/RFC3393, November 2002, <https://www.rfc-editor.org/info/rfc3393>.

[RFC5136] Chimento, P. and J. Ishac, "Defining Network Capacity", RFC 5136, DOI 10.17487/RFC5136, February 2008, <https://www.rfc-editor.org/info/rfc5136>.

[RFC5693] Seedorf, J. and E. Burger, "Application-Layer Traffic Optimization (ALTO) Problem Statement", RFC 5693, DOI 10.17487/RFC5693, October 2009, <https://www.rfc-editor.org/info/rfc5693>.

[RFC6534] Duffield, N., Morton, A., and J. Sommers, "Loss Episode Metrics for IP Performance Metrics (IPPM)", RFC 6534, DOI 10.17487/RFC6534, May 2012, <https://www.rfc-editor.org/info/rfc6534>.

[RFC7665] Halpern, J., Ed. and C. Pignataro, Ed., "Service Function Chaining (SFC) Architecture", RFC 7665, DOI 10.17487/RFC7665, October 2015, <https://www.rfc-editor.org/info/rfc7665>.

[RFC7679] Almes, G., Kalidindi, S., Zekauskas, M., and A. Morton, Ed., "A One-Way Delay Metric for IP Performance Metrics (IPPM)", STD 81, RFC 7679, DOI 10.17487/RFC7679, January 2016, <https://www.rfc-editor.org/info/rfc7679>.

[RFC7680] Almes, G., Kalidindi, S., Zekauskas, M., and A. Morton, Ed., "A One-Way Loss Metric for IP Performance Metrics (IPPM)", STD 82, RFC 7680, DOI 10.17487/RFC7680, January 2016, <https://www.rfc-editor.org/info/rfc7680>.

[RFC8175] Ratliff, S., Jury, S., Satterwhite, D., Taylor, R., and B. Berry, "Dynamic Link Exchange Protocol (DLEP)", RFC 8175, DOI 10.17487/RFC8175, June 2017, <https://www.rfc-editor.org/info/rfc8175>.

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, and thanks to Adrian Perrig and Matthias Rost for the detailed feedback. Thanks to Paul Hoffman for the editorial changes.

Authors' Addresses

Theresa Enghardt
TU Berlin

Email: theresa@inet.tu-berlin.de

Cyrill Kraehenbuehl
ETH Zuerich

Email: cyrill.kraehenbuehl@inf.ethz.ch