Network Working Group
Internet-Draft
Updates: 2330 (if approved)
Intended status: Standards Track
Expires: April 29, 2018

J. Alvarez-Hamelin
Universidad de Buenos Aires
A. Morton
AT&T Labs
J. Fabini
TU Wien

October 26, 2017

Advanced Unidirectional Route Assessment
draft-amf-ippm-route-01

Abstract

This memo introduces an advanced unidirectional route assessment metric and associated measurement methodology, based on the IP Performance Metrics (IPPM) Framework RFC 2330. This memo updates RFC 2330 in the areas of path-related terminology and path description, primarily to include the possibility of parallel subpaths between a given Source and Destination pair, owing to the presence of multi-path technologies.

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 April 29, 2018.

Table of Contents

1. Introduction

The IETF IP Performance Metrics (IPPM) working group first created a framework for metric development in [RFC2330]. This framework has stood the test of time and enabled development of many fundamental

metrics. It has been updated in the area of metric composition [RFC5835], and in several areas related to active stream measurement of modern networks with reactive properties [RFC7312].

The [RFC2330] framework motivated the development of "performance and reliability metrics for paths through the Internet," and Section 5 of [RFC2330] defines terms that support description of a path under test. However, metrics for assessment of path components and related performance aspects had not been attempted in IPPM when the [RFC2330] framework was written.

This memo takes-up the route measurement challenge and specifies a new route metric, two practical frameworks for methods of measurement (using either active or hybrid active-passive methods [RFC7799]), and round-trip delay and link information discovery using the results of measurements.

1.1. Issues with Earlier Work to define Route

Section 7 of [RFC2330] presented a simple example of a "route" metric along with several other examples. The example is reproduced below (where the reference is to Section 5 of [RFC2330]):

"route: The path, as defined in Section 5, from A to B at a given time."

This example provides a starting point to develop a more complete definition of route. Areas needing clarification include:

Time: In practice, the route will be assessed over a time interval, because active path detection methods like [PT] rely on TTL limits for their operation and cannot accomplish discovery of all hosts using a single packet.

Type-P: The legacy route definition lacks the option to cater for packet-dependent routing. In this memo, we assess the route for a specific packet of Type-P, and reflect this in the metric definition. The methods of measurement determine the specific Type-P used.

Parallel Paths: This a reality of Internet paths and a strength of advanced route assessment methods, so the metric must acknowledge this possibility. Use of Equal Cost Multi-Path (ECMP) and Unequal Cost Multi-Path (UCMP) technologies are common sources of parallel subpaths.

Cloud Subpath: May contain hosts that do not decrement TTL or Hop Limit, but may have two or more exchange links connecting

"discoverable" hosts or routers. Parallel subpaths contained within clouds cannot be discovered. The assessment methods only discover hosts or routers on the path that decrement TTL or Hop Count, or cooperate with interrogation protocols. The presence of tunnels and nested tunnels further complicate assessment by hiding hops.

Hop: Although the [RFC2330] definition was a link-host pair, only hosts are discoverable or have the capability to cooperate with interrogation protocols where link information may be exposed.

The refined definition of Route metrics begins in the sections that follow.

2. Scope

The purpose of this memo is to add new route metrics and methods of measurement to the existing set of IPPM metrics.

The scope is to define route metrics that can identify the path taken by a packet or a flow traversing the Internet between any two hosts.

<@@@@ or only hosts communicating at the IP layer? We would have to re-define the Src and Dst Parameters and Host Identity if we generalize beyond IP. Should we include MPLS and the capabilities of [RFC8029], with explicit multipath identification (section 6.2.6)? >

Also, to specify a framework for active methods of measurement which use the techniques described in [PT] at a minimum, and a framework for hybrid active-passive methods of measurement, such as the Hybrid Type I method [RFC7799] described in [I-D.ietf-ippm-ioam-data](intended only for single administrative domains), which do not rely on ICMP and provide a protocol for explicit interrogation of nodes on a path. Combinations of active methods and hybrid active-passive methods are also in-scope.

Further, this memo provides additional analysis of the round-trip delay measurements made possible by the methods, in an effort to discover more details about the path, such as the link technology in use.

This memo updates Section 5 of [RFC2330] in the areas of path-related terminology and path description, primarily to include the possibility of parallel subpaths between a given Source and Destination address pair (possibly resulting from Equal Cost Multi-Path (ECMP) and Unequal Cost Multi-Path (UCMP) technologies).

There are several simple non-goals of this memo. There is no attempt to assess the reverse path from any host on the path to the host attempting the path measurement. The reverse path contribution to delay will be that experienced by ICMP packets (in active methods), and may be different from UDP or TCP packets. Also, the round trip delay will include an unknown contribution of processing time at the host that generates the ICMP response. Therefore, the ICMP-based active methods are not supposed to yield accurate, reproducible estimations of the round-trip delay that UDP or TCP packets will experience.

3. Route Metric Terms and Definitions

This section sets requirements for the following components to support the Route Metric:

Note: the definitions concentrate on the IP-layer, but can be extended to other layers, and follow agreements on the scope.

Host Identity For hosts communicating at the IP-layer, the globally routable IP address(es) which the host uses when communicating with other hosts under normal or error conditions. The Host Identity revealed (and its connection to a Host Name through reverse DNS) determines whether interfaces to parallel links can be associated with a single host, or appear to be unique hosts.

Discoverable Host For hosts communicating at the IP-layer, compliance with Section 3.2.2.4 of [RFC1122] when discarding a packet due to TTL or Hop Limit Exceeded condition, MUST result in sending the corresponding Time Exceeded message (containing a form of host identity) to the source. This requirement is also consistent with section 5.3.1 of [RFC1812] for routers.

Cooperating Host Hosts MUST respond to direct queries for their host identity as part of a previously agreed and established interrogation protocol. Hosts SHOULD also provide information such as arrival/departure interface identification, arrival timestamp, and any relevant information about the host or specific link which delivered the query to the host.

Hop A Hop MUST contain a Host Identity, and MAY contain arrival and/or departure interface identification.

3.1. Formal Name

Type-P-Route-Ensemble-Method-Variant, abbreviated as Route Ensemble.

Note that Type-P depends heavily on the chosen method and variant.

3.2. Parameters

This section lists the REQUIRED input factors to specify a Route metric.

- Src, the IP address of a host
- Dst, the IP address of a host
- i, the TTL or Hop Limit of a packet sent from the host at Src to the host at Dst.
- MaxHops, the maximum value of i used, (i=1,2,3,...MaxHops).
- T0, a time (start of measurement interval)
- Tf, a time (end of measurement interval)
- T, the host time of a packet as measured at MP(Src), meaning Measurement Point at the Source.
- Ta, the host time of a reply packet’s *arrival* as measured at MP(Src), assigned to packets that arrive within a "reasonable" time (see parameter below).
- Tmax, a maximum waiting time for reply packets to return to the source, set sufficiently long to disambiguate packets with long delays from packets that are discarded (lost), thus the distribution of delay is not truncated.
- F, the number of different flows simulated by the method and variant.
- flow, the stream of packets with the same n-tuple of designated header fields that (when held constant) results in identical treatment in a multi-path decision (such as that taken in load balancing).
- Type-P, the complete description of the packets for which this assessment applies (including the flow-defining fields).

3.3. Metric Definitions

This section defines the REQUIRED measurement components of the Route metrics (unless otherwise indicated):

M, the total number of packets sent between T0 and Tf.

N, the smallest value of i needed for a packet to be received at Dst (sent between T0 and Tf).

Nmax, the largest value of i needed for a packet to be received at Dst (sent between T0 and Tf). Nmax may be equal to N.

Next, define a *singleton* definition for a Hop on the path, with sufficient indexes to identify all Hops identified in a measurement interval.

A Hop, designated h(i,j), the IP address and/or identity of one of j Discoverable Hosts (or Cooperating Hosts) that are i hops away from the host with IP address = Src during the measurement interval, T0 to Tf. As defined above, a Hop singleton measurement MUST contain a Host Identity, hid(i,j), and MAY contain one or more of the following attributes:

- a(i,j) Arrival Interface ID
- d(i,j) Departure Interface ID
- t(i,j) Arrival Timestamp (where t(i,j) is ideally supplied by the hop, or approximated from the sending time of the packet that revealed the hop)
- Measurements of Round Trip Delay (for each packet that reveals the same Host Identity and attributes, but not timestamp of course, see next section)

Now that Host Identities and related information can be positioned according to their distance from the host with address Src in hops, we introduce two forms of Routes:

A Route Ensemble is defined as the combination of all routes traversed by different flows from the host at Src address to the host at Dst address. The route traversed by each flow (with addresses Src and Dst, and other fields which constitute flow criteria) is a member of the ensemble and called a Member Route.

Using h(i,j) and components and parameters, further define:

A Member Route is an ordered graph {h(1,j), ... h(Nj, j)} in the context of a single flow, where h(i-1, j) and h(i, j) are by 1 hop away from each other and Nj=Dst is the minimum TTL value needed by the packet on Member Route j to reach Dst. Member Routes must be unique. This uniqueness requires that any two Member routes j and k that are part of the same Route Ensemble differ either in terms of minimum hop count Nj and Nk to reach the destination Dst, or, in the

case of identical hop count Nj=Nk, they have at least one distinct hop: h(i,j) != h(i, k) for at least one i (i=1..Nj).

The Route Ensemble from Src to Dst, during the measurement interval T0 to Tf, is the aggregate of all m distinct Member Routes discovered between the two hosts with Src and Dst addresses. More formally, with the host having address Src omitted:

Route Ensemble = {
{h(1,1), h(2,1), h(3,1), ... h(N1,1)=Dst},
{h(1,2), h(2,2), h(3,2),..., h(N2,2)=Dst},
...
{h(1,m), h(2,m), h(3,m), ...h(Nm,m)=Dst}
}

where the following conditions apply: i <= Nj <= Nmax (j=1..m)

Note that some h(i,j) may be empty (null) in the case that systems do not reply (not discoverable, or not cooperating).

h(i-1,j) and h(i,j) are the Hops on the same Member Route one hop away from each other.

Hop h(i,j) may be identical with h(k,l) for i!=k and j!=l ; which means there may be portions shared among different Member Routes (parts of various routes may overlap).

3.4. Related Round-Trip Delay and Loss Definitions

RTD(i,j,T) is defined as a singleton of the [RFC2681] Round-trip Delay between the host with IP address = Src and the host at Hop h(i,j) at time T.

RTL(i,j,T) is defined as a singleton of the [RFC6673] Round-trip Loss between the host with IP address = Src and the host at Hop h(i,j) at time T.

3.5. Discussion

Depending on the way that Host Identity is revealed, it may be difficult to determine parallel subpaths between the same pair of hosts (i.e. multiple parallel links). It is easier to detect parallel subpaths involving different hosts.

- If a pair of discovered hosts identify two different IP addresses, then they will appear to be different hosts.

- If a pair of discovered hosts identify two different IP addresses, and the IP addresses resolve to the same host name (in the DNS), then they will appear to be the same hosts.

- If a discovered host always replies using the same IP address, regardless of the interface a packet arrives on, then multiple parallel links cannot be detected at the IP layer.

- If parallel links between routers are aggregated below the IP layer, In other words, all links share the same pair of IP addresses, then the existence of these parallel links can’t be detected at IP layer.

Section 9.2 of [RFC2330] describes Temporal Composition of metrics, and introduces the possibility of a relationship between earlier measurement results and the results for measurement at the current time (for a given metric). If this topic is investigated further, there may be some value in establishing a Temporal Composition relationship for Route Metrics. However, this relationship does not represent a forecast of future route conditions in any way.

When a route assessment employs packets at the IP layer (for example), the reality of flow assignment to parallel subpaths involves layers above IP. Thus, the measured Route Ensemble is applicable to IP and higher layers (as described in the methodology’s packet of Type-P and flow parameters).

@@@@ Editor’s Note: There is an opportunity to investigate and discuss the RFC 2330 notion of equal treatment for a class of packets, "...very useful to know if a given Internet component treats equally a class C of different types of packets", as it applies to Route measurements. Knowledge of "class C" parameters on a path potentially reduces the number of flows required for a given method.

3.6. Reporting the Metric

@@@@ to be provided

4. Route Assessment Methodologies

There are two classes of methods described in this section, active methods relying on the reaction to TTL or Hop Limit Exceeded condition to discover hosts on a path, and Hybrid active-passive methods that involve direct interrogation of cooperating hosts (usually within a single domain). Description of these methods follow.

@@@@ Editor's Note: We need to incorporate description of Type-P packets (with the flow parameters) used in each method below.

4.1. Active Methodologies

We have chosen to describe the method based on that employed in current open source tools, thereby providing a practical framework for further advanced techniques to be included as method variants. This method is applicable to use across multiple administrative domains.

Paris-traceroute [PT] provides some measure of protection from path variation generated by ECMP load balancing, and it ensures traceroute packets will follow the same path in 98% of cases according to [SCAMPER]. If it is necessary to find every path possible between two hosts, Paris-traceroute provides "exhaustive" mode while scamper provides "tracelb" (stands for traceroute load balance).

The Type-P of packets used could be ICMP (as ones in the original traceroute), UDP and TCP. The later are used when a particular characteristic is needed to verify, such as filtering or traffic shaping on specific ports (i.e., services).

The advanced route assessment methods used in Paris-traceroute [PT] keep the critical fields constant for every packet to maintain the appearance of the same flow. Since route assessment can be conducted using TCP, UDP or ICMP packets, this method REQUIRES the Diffserv field, the protocol number, IP source and destination addresses, and the port settings for TCP or UDP kept constant. For ICMP probes, the method additionally REQUIRES the type, code, and ICMP checksum constant; which take the same position in the header of an IP packet, e.g., bytes 20 to 23 when the header IP has no options.

Maintaining a constant checksum in ICMP is most challenging because the ICMP Sequence Number is part of the calculation. The advanced traceroute method requires calculations using the IP Sequence Number Field and the Identifier Field, yielding a constant ICMP checksum in successive packets. For an example of calculations to maintain a constant checksum, see Appendix A of [RFC7820], where revision of a timestamp field is complemented by modifying the 2 octet checksum complement field (these fields take the roles of the ICMP Sequence Number Identifier Fields, respectively).

For TCP and UDP packets, the checksum must also be kept constant. Therefore, the first four bytes of UDP (or TCP) data field are modified to compensate for fields that change from packet to packet.

Note: other variants of advanced traceroute are planned be described.

Finally, the return path is also important to check. Taking into account that it is an ICMP time exceeded (during transit) packet, the source and destination IP are constant for every reply. Then, we should consider the fields in the first 32 bits of the protocol on the top of IP: the type and code of ICMP packet, and its checksum. Again, to maintain the ICMP checksum constant for the returning packets, we need to consider the whole ICMP message. It contains the IP header of the discarded packet plus the first 8 bytes of the IP payload; that is some of the fields of TCP header, the UDP header plus four data bytes, the ICMP header plus four bytes. Therefore, for UDP case the data field is used to maintain the ICMP checksum constant in the returning packet. For the ICMP case, the identifier and sequence fields of the sent ICMP probe are manipulated to be constant. The TCP case presents no problem because its first eight bytes will be the same for every packet probe.

Formally, to maintain the same flow in the measurements to a certain hop, the Type-P-Route-Ensemble-Method-Variant packets should be[PT]:

- TCP case: Fields Src, Dst, port-Src, port_Dst, and Diffserv Field should be the same.

- UDP case: Fields Src, Dst, port-Src, port-Dst, and Diffserv Field should be the same, the UDP-checksum should change to maintain constant the IP checksum of the ICMP time exceeded reply. Then, the data length should be fixed, and the data field is used to fixing it (consider that ICMP checksum uses its data field, which contains the original IP header plus 8 bytes of UDP, where TTL, IP identification, IP checksum, and UDP checksum changes).

- ICMP case: The Data field should compensate variations on TTL, IP identification, and IP checksum for every packet.

Then, the way to identify different hops and attempts of the same flow is:

- TCP case: The IP identification field.

- UDP case: The IP identification field.

- ICMP case: The IP identification field, and ICMP Sequence number.

4.2. Hybrid Methodologies

The Hybrid Type I methods provide an alternative method for Route Member assessment. As mentioned in the Scope section, [I-D.ietf-ippm-ioam-data] provides a possible set of data fields that would support route identification.

In general, nodes in the measured domain would be equipped with specific abilities:

1. The ingress node adds one or more fields to the measurement packets, and identifies to other nodes in the domain that a route assessment will be conducted using one or more specific packets. The packets typically originate from a host outside the domain, and constitute normal traffic on the domain.

2. Each node visited by the specific packet within in the domain identifies itself in a data field of the packet (the field has been added for this purpose).

3. When a measurement packet reaches the edge node of the domain, the edge node adds its identity to the list, removes all the identities from the packet, forwards the packet onward, and communicates the ordered list of node identities to the intended receiver.

In addition to node identity, nodes may also identify the ingress and egress interfaces utilized by the tracing packet, the time of day when the packet was processed, and other generic data (as described in section 4 of [I-D.ietf-ippm-ioam-data]).

4.3. Combining Different Methods

In principle, there are advantages if the entity conducting Route measurements can utilize both forms of advanced methods (active and hybrid), and combine the results. For example, if there are hosts involved in the path that qualify as Cooperating Hosts, but not as Discoverable Hosts, then a more complete view of hops on the path is possible when a hybrid method (or interrogation protocol) is applied and the results are combined with the active method results collected across all other domains.

In order to combine the results of active and hybrid/interrogation methods, the network hosts that are part of a domain supporting an interrogation protocol have the following attributes:

1. Hosts at the ingress to the domain SHOULD be both Discoverable and Cooperating, and SHOULD reveal the same Host Identity in response to both active and hybrid methods.

2. Any Hosts within the domain that are both Discoverable and Cooperating SHOULD reveal the same Host Identity in response to both active and hybrid methods.

3. Hosts at the egress to the domain SHOULD be both Discoverable and Cooperating, and SHOULD reveal the same Host Identity in response to both active and hybrid methods.

When Hosts follow these requirements, it becomes a simple matter to match single domain measurements with the overlapping results from a multidomain measurement.

In practice, Internet users do not typically have the ability to utilize the OAM capabilities of networks that their packets traverse, so the results from a remote domain supporting an interrogation protocol would not normally be accessible. However, a network operator could combine interrogation results from their access domain with other measurements revealing the path outside their domain.

5. Background on Round-Trip Delay Measurement Goals

The aim of this method is to use packet probes to unveil the paths between any two end-hosts of the network. Moreover, information derived from RTD measurements might be meaningful to identify:

1. Intercontinental submarine links

2. Satellite communications

3. Congestion

4. Inter-domain paths

This categorization is widely accepted in the literature and among operators alike, and it can be trusted with empirical data and several sources as ground of truth (e.g., [RTTSub] [bdrmap][IDCong]).

The first two categories correspond to the physical distance dependency on Round Trip Delay (RTD) while the last one binds RTD with queueing delay on routers. Due to the significant contribution of propagation delay in long distance hops, RTD will be at least 100ms on transatlantic hops, depending on the geolocation of the vantage points. Moreover, RTD is typically greater than 480ms when two hops are connected using geostationary satellite technology (i.e., their orbit is at 36000km). Detecting congestion with latency implies deeper mathematical understanding since network traffic load is not stationary. Nonetheless, as the first approach, a link seems to be congested if after sending several traceroute probes, it is possible to detect congestion observing different statistics parameters (e.g., see [IDCong]).

6. Tools to Measure Delays in the Internet

Internet routing is complex because it depends on the policies of thousands Autonomous Systems (AS). While most of the routers perform load balancing on flows using Equal Cost Multiple Path (ECMP), a few still divide the workload through packet-based techniques. The former scenario is defined according to [RFC2991] while the latter generates a round-robin scheme to deliver every new outgoing packet. ECMP keeps flow state in the router to ensure every packet of a flow is delivered by the same path, and this avoids increasing the packet delay variation and possibly producing overwhelming packet reordering in TCP flows.

Taking into account that Internet protocol was designed under the "end-to-end" principle, the IP payload and its header do not provide any information about the routes or path necessary to reach some destination. For this reason, the well-known tool traceroute was developed to gather the IP addresses of each hop along a path using the ICMP protocol [RFC0792]. Besides, traceroute adds the measured RTD from each hop. However, the growing complexity of the Internet makes it more challenging to develop accurate traceroute implementation. For instance, the early traceroute tools would be inaccurate in the current network, mainly because they were not designed to retain flow state. However, evolved traceroute tools, such as Paris-traceroute [PT] [MLB] and Scamper [SCAMPER], expect to encounter ECMP and achieve more accurate results when they do.

Paris-traceroute-like tools operate in the following way: every packet should follow the same path because the sensitive fields of the header are controlled to appear as the same flow. This means that source and destination IP addresses, source and destination port numbers are the same in every packet. Additionally, Differentiated Services Code Point (DSCP), checksum and ICMP code should remain constant since they may affect the path selection.

Today’s traceroute tools can send either UDP, TCP or ICMP packet probes. Since ICMP header does not include transport layer information, there are no fields for source and destination port numbers. For this reason, these tools keep constant ICMP type, code, and checksum fields to generate a kind of flow. However, the checksum may vary in every packet, therefore when probes use ICMP packets, ICMP Identifier and Sequence Number are manipulated to maintain constant checksum in every packet. On the other hand, when UDP probes are generated, the expected variation in the checksum of each packet is again compensated by manipulating the payload.

Paris-traceroute allows its users to measure RTD in every hop of the path for a particular flow. Furthermore, either Paris-traceroute or

Scamper is capable of unveiling the many available paths between a source and destination (which are visible to this method). This task is accomplished by repeating complete traceroute measurements with different flow parameters for each measurement. The Framework for IP Performance Metrics (IPPM) ([RFC2330] updated by[RFC7312]) has the flexibility to require that the round-trip delay measurement [RFC2681] uses packets with the constraints to assure that all packets in a single measurement appear as the same flow. This flexibility covers ICMP, UDP, and TCP. The accompanying methodology of [RFC2681] needs to be expanded to report the sequential hop identifiers along with RTD measurements, but no new metric definition is needed.

7. RTD Measurements Statistics

Several articles have shown that network traffic presents a self-similar nature [SSNT] [MLRM] which is accountable for filling the queues of the routers. Moreover, router queues are designed to handle traffic bursts, which is one of the most remarkable features of self-similarity. Naturally, while queue length increases, the delay to traverse the queue increases as well and leads to an increase on RTD. Due to traffic bursts generate short-term overflow on buffers (spiky patterns), every RTD only depicts the queueing status on the instant when that packet probe was in transit. For this reason, several RTD measurements during a time window could begin to describe the random behavior of latency. Loss must also be accounted for in the methodology.

To understand the ongoing process, examining the quartiles provides a non-parametric way of analysis. Quartiles are defined by five values: minimum RTD (m), RTD value of the 25% of the Empirical Cumulative Distribution Function (ECDF) (Q1), the median value (Q2), the RTD value of the 75% of the ECDF (Q3) and the maximum RTD (M). Congestion can be inferred when RTD measurements are spread apart, and consequently, the Inter-Quartile Range (IQR), the distance between Q3 and Q1, increases its value.

This procedure requires to compute quartile values "on the fly" using the algorithm presented in [P2].

This procedure allow us to update the quartiles value whenever a new measurement arrives, which is radically different from classic methods of computing quartiles because they need to use the whole dataset to compute the values. This way of calculus provides savings in memory and computing time.

To sum up, the proposed measurement procedure consists in performing traceroutes several times to obtain samples of the RTD in every hop

from a path, during a time window (W) and compute the quantiles for every hop. This could be done for a single path flow or for every detected path flow.

Even though a particular hop may be understood as the amount of hops away from the source, a more detailed classification could be used. For example, a possible classification may be identify ICMP Time Exceeded packets coming from the same routers to those who have the same hop distance, IP address of the router which is replying and TTL value of the received ICMP packet.

Thus, the proposed methodology is based on this algorithm:

```
 ================================================================
  1  input:   W (window time of the measurement)
  2           i_t (time between two measurements)
  3           E (True: exhaustive, False: a single path)
  4           Dst (destination IP address)
  5  output:  Qs (quartiles for every hop and alt in the path(s) to Dst)
 ----------------------------------------------------------------
  6  T <? start_timer(W)
  7  while T is not finished do:
  8  |       start_timer(i_t)
  9  |       RTD(hop,alt) = advanced-traceroute(Dst,E)
 10  |       for each hop and alt in RTD do:
 11  |       |     Qs[Dst,hop,alt] <? ComputeQs(RTD(hop,alt))
 12  |       done
 13  |       wait until i_t timer is expired
 14  done
 15  return (Qs)
 ================================================================
```

In line 9 the advance-traceroute could be either Paris-traceroute or Scamper, which will use "exhaustive" mode or "tracelb" option if E is set True, respectively. The procedure returns a list of tuples (m,Q1,Q2,Q3,M) for each intermediate hop in the path towards the Dst. Additionally, it could also return path variations using "alt" variable.

8. Conclusions

Combining the method proposed in Section 4 and statistics in Section 7, we can measure the performance of paths interconnecting two endpoints in Internet, and attempt the categorization of link types and congestion presence based on RTD.

9. Security Considerations

The security considerations that apply to any active measurement of live paths are relevant here as well. See [RFC4656] and [RFC5357].

The active measurement process of "changing several fields to keep the checksum of different packets identical" does not require special security considerations because it is part of synthetic traffic generation, and is designed to have minimal to zero impact on network processing (to process the packets for ECMP).

@@@@ add reference to security considerations from [I-D.ietf-ippm-ioam-data].

When considering privacy of those involved in measurement or those whose traffic is measured, the sensitive information available to potential observers is greatly reduced when using active techniques which are within this scope of work. Passive observations of user traffic for measurement purposes raise many privacy issues. We refer the reader to the privacy considerations described in the Large Scale Measurement of Broadband Performance (LMAP) Framework [RFC7594], which covers active and passive techniques.

10. IANA Considerations

This memo makes no requests of IANA.

11. Acknowledgements

The authors acknowledge Ruediger Geib, for his penetrating comments on the initial draft. Carlos Pignataro challenged the authors to consider a wider scope, and applied his substantial expertise with many technologies and their measurement features in his extensive comments. Frank Brockners also shared useful comments. We thank them all!

12. References

12.1. Normative References

[I-D.ietf-ippm-ioam-data]
Brockners, F., Bhandari, S., Pignataro, C., Gredler, H., Leddy, J., Youell, S., Mizrahi, T., Mozes, D., Lapukhov, P., Chang, R., and d. daniel.bernier@bell.ca, "Data Fields for In-situ OAM", draft-ietf-ippm-ioam-data-00 (work in progress), September 2017.

[RFC0792] Postel, J., "Internet Control Message Protocol", STD 5, RFC 792, DOI 10.17487/RFC0792, September 1981, <https://www.rfc-editor.org/info/rfc792>.

[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>.

[RFC1812] Baker, F., Ed., "Requirements for IP Version 4 Routers", RFC 1812, DOI 10.17487/RFC1812, June 1995, <https://www.rfc-editor.org/info/rfc1812>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC2330] Paxson, V., Almes, G., Mahdavi, J., and M. Mathis, "Framework for IP Performance Metrics", RFC 2330, DOI 10.17487/RFC2330, May 1998, <https://www.rfc-editor.org/info/rfc2330>.

[RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, DOI 10.17487/RFC2460, December 1998, <https://www.rfc-editor.org/info/rfc2460>.

[RFC2675] Borman, D., Deering, S., and R. Hinden, "IPv6 Jumbograms", RFC 2675, DOI 10.17487/RFC2675, August 1999, <https://www.rfc-editor.org/info/rfc2675>.

[RFC2681] Almes, G., Kalidindi, S., and M. Zekauskas, "A Round-trip Delay Metric for IPPM", RFC 2681, DOI 10.17487/RFC2681, September 1999, <https://www.rfc-editor.org/info/rfc2681>.

[RFC2991] Thaler, D. and C. Hopps, "Multipath Issues in Unicast and Multicast Next-Hop Selection", RFC 2991, DOI 10.17487/RFC2991, November 2000, <https://www.rfc-editor.org/info/rfc2991>.

[RFC4494] Song, JH., Poovendran, R., and J. Lee, "The AES-CMAC-96 Algorithm and Its Use with IPsec", RFC 4494, DOI 10.17487/RFC4494, June 2006, <https://www.rfc-editor.org/info/rfc4494>.

[RFC4656] Shalunov, S., Teitelbaum, B., Karp, A., Boote, J., and M. Zekauskas, "A One-way Active Measurement Protocol (OWAMP)", RFC 4656, DOI 10.17487/RFC4656, September 2006, <https://www.rfc-editor.org/info/rfc4656>.

[RFC5357] Hedayat, K., Krzanowski, R., Morton, A., Yum, K., and J. Babiarz, "A Two-Way Active Measurement Protocol (TWAMP)", RFC 5357, DOI 10.17487/RFC5357, October 2008, <https://www.rfc-editor.org/info/rfc5357>.

[RFC5644] Stephan, E., Liang, L., and A. Morton, "IP Performance Metrics (IPPM): Spatial and Multicast", RFC 5644, DOI 10.17487/RFC5644, October 2009, <https://www.rfc-editor.org/info/rfc5644>.

[RFC5835] Morton, A., Ed. and S. Van den Berghe, Ed., "Framework for Metric Composition", RFC 5835, DOI 10.17487/RFC5835, April 2010, <https://www.rfc-editor.org/info/rfc5835>.

[RFC6282] Hui, J., Ed. and P. Thubert, "Compression Format for IPv6 Datagrams over IEEE 802.15.4-Based Networks", RFC 6282, DOI 10.17487/RFC6282, September 2011, <https://www.rfc-editor.org/info/rfc6282>.

[RFC6437] Amante, S., Carpenter, B., Jiang, S., and J. Rajahalme, "IPv6 Flow Label Specification", RFC 6437, DOI 10.17487/RFC6437, November 2011, <https://www.rfc-editor.org/info/rfc6437>.

[RFC6564] Krishnan, S., Woodyatt, J., Kline, E., Hoagland, J., and M. Bhatia, "A Uniform Format for IPv6 Extension Headers", RFC 6564, DOI 10.17487/RFC6564, April 2012, <https://www.rfc-editor.org/info/rfc6564>.

[RFC6673] Morton, A., "Round-Trip Packet Loss Metrics", RFC 6673, DOI 10.17487/RFC6673, August 2012, <https://www.rfc-editor.org/info/rfc6673>.

[RFC7045] Carpenter, B. and S. Jiang, "Transmission and Processing of IPv6 Extension Headers", RFC 7045, DOI 10.17487/RFC7045, December 2013, <https://www.rfc-editor.org/info/rfc7045>.

[RFC7312] Fabini, J. and A. Morton, "Advanced Stream and Sampling Framework for IP Performance Metrics (IPPM)", RFC 7312, DOI 10.17487/RFC7312, August 2014, <https://www.rfc-editor.org/info/rfc7312>.

[RFC7799] Morton, A., "Active and Passive Metrics and Methods (with Hybrid Types In-Between)", RFC 7799, DOI 10.17487/RFC7799, May 2016, <https://www.rfc-editor.org/info/rfc7799>.

[RFC7820] Mizrahi, T., "UDP Checksum Complement in the One-Way Active Measurement Protocol (OWAMP) and Two-Way Active Measurement Protocol (TWAMP)", RFC 7820, DOI 10.17487/RFC7820, March 2016, <https://www.rfc-editor.org/info/rfc7820>.

[RFC8029] Kompella, K., Swallow, G., Pignataro, C., Ed., Kumar, N., Aldrin, S., and M. Chen, "Detecting Multiprotocol Label Switched (MPLS) Data-Plane Failures", RFC 8029, DOI 10.17487/RFC8029, March 2017, <https://www.rfc-editor.org/info/rfc8029>.

12.2. Informative References

[bdrmap] Luckie, M., Dhamdhere, A., Huffaker, B., Clark, D., and KC. Claffy, "bdrmap: Inference of Borders Between IP Networks", In Proceedings of the 2016 ACM on Internet Measurement Conference, pp. 381-396. ACM, 2016.

[I-D.brockners-inband-oam-data] Brockners, F., Bhandari, S., Pignataro, C., Gredler, H., Leddy, J., Youell, S., Mizrahi, T., Mozes, D., Lapukhov, P., Chang, R., and d. daniel.bernier@bell.ca, "Data Fields for In-situ OAM", draft-brockners-inband-oam-data-07 (work in progress), July 2017.

[IDCong] Luckie, M., Dhamdhere, A., Clark, D., and B. Huffaker, "Challenges in inferring Internet interdomain congestion", In Proceedings of the 2014 Conference on Internet Measurement Conference, pp. 15-22. ACM, 2014.

[MLB] Augustin, B., Friedman, T., and R. Teixeira, "Measuring load-balanced paths in the Internet", Proceedings of the 7th ACM SIGCOMM conference on Internet measurement, pp. 149-160. ACM, 2007., 2007.

[MLRM] Fontugne, R., Mazel, J., and K. Fukuda, "An empirical mixture model for large-scale RTT measurements", 2015 IEEE Conference on Computer Communications (INFOCOM), pp. 2470-2478. IEEE, 2015., 2015.

[P2] Jain, R. and I. Chlamtac, "The P 2 algorithm for dynamic calculation of quantiles and histograms without storing observations", Communications of the ACM 28.10 (1985): 1076-1085, 2015.

[PT] Augustin, B., Cuvellier, X., Orgogozo, B., Viger, F., Friedman, T., Latapy, M., Magnien, C., and R. Teixeira, "Avoiding traceroute anomalies with Paris traceroute", Proceedings of the 6th ACM SIGCOMM conference on Internet measurement, pp. 153-158. ACM, 2006., 2006.

[RFC7594] Eardley, P., Morton, A., Bagnulo, M., Burbridge, T., Aitken, P., and A. Akhter, "A Framework for Large-Scale Measurement of Broadband Performance (LMAP)", RFC 7594, DOI 10.17487/RFC7594, September 2015, <https://www.rfc-editor.org/info/rfc7594>.

[RTTSub] Bischof, Z., Rula, J., and F. Bustamante, "In and out of Cuba: Characterizing Cuba's connectivity", In Proceedings of the 2015 ACM Conference on Internet Measurement Conference, pp. 487-493. ACM, 2015.

[SCAMPER] Matthew Luckie, M., "Scamper: a scalable and extensible packet prober for active measurement of the Internet", Proceedings of the 10th ACM SIGCOMM conference on Internet measurement, pp. 239-245. ACM, 2010., 2010.

[SSNT] Park, K. and W. Willinger, "Self-Similar Network Traffic and Performance Evaluation (1st ed.)", John Wiley & Sons, Inc., New York, NY, USA, 2000.

Authors' Addresses

Jose Ignacio Alvarez-Hamelin
Universidad de Buenos Aires
Av. Paseo Colon 850
Buenos Aires C1063ACV
Argentine

Phone: +54 11 5285-0716
Email: ihameli@cnet.fi.uba.ar
URI: http://cnet.fi.uba.ar/ignacio.alvarez-hamelin/

Al Morton
AT&T Labs
200 Laurel Avenue South
Middletown, NJ 07748
USA

Phone: +1 732 420 1571
Fax: +1 732 368 1192
Email: acmorton@att.com
URI: http://home.comcast.net/~acmacm/

Joachim Fabini
TU Wien
Gusshausstrasse 25/E389
Vienna 1040
Austria

Phone: +43 1 58801 38813
Fax: +43 1 58801 38898
Email: Joachim.Fabini@tuwien.ac.at
URI: http://www.tc.tuwien.ac.at/about-us/staff/joachim-fabini/