ALTO Performance Metrics

draft-ietf-alto-performance-metrics-20->28

Qin Wu
Y. Richard Yang
Young Lee
D. Dhody
Sabine Randriamasy
Luis Miguel Contreras Murillo

March 23, 2022

IETF 113, Virtual Meeting

Status

- Updates between IETF 112 and IETF 113 (v20 -> v28)
  - Diff: https://tools.ietf.org/rfcdiff?url1=https://www.ietf.org/archive/id/draft-ietf-alto-performance-metrics-20.txt&url2=https://www.ietf.org/archive/id/draft-ietf-alto-performance-metrics-28.txt
  - Address IESG reviews
    - Provide structures and organizations
    - Specify clear grammar and requirements
    - Give as much info as possible on derivations, conditions
    - Update references (e.g., avoid downref)
    - Update examples, including check json grammar, ipv6 example
  - Have ballots but there are large updates which can benefit from one more round of reviews

Main Update: Structure

- Clarify that ALTO performance metrics are systematically derived from routing metrics (e.g., OSPF/ISIS/GBP-LS routing TE metrics are main source of ALTO performance metrics)

1		IPPM 7679 OW delay	IPPM 2681 RTT	IPPM 3393 Delay Var	RFC3630 OSPF TE	RFC7471 OSPF TE Metric Ext	RFC5305 ISIS TE	RFC8570 ISIS TE Metric Ext	RFC 8571 BGP-LS	ALTO PM
2	One-way delay	x				x		x	x	x (delay-ow)
3	min-max delay					x		x	x	x (using max-min op)
4	RTT		x							x (delay-rt)
5	Delay variation			x		x		x	x	x (delay-variation)
6	Loss rate					x		x	x	x (lossrate)
7	TE default value				x		x			x (routingcost)
8	Max (link) BW				x		x			x (max bw-residual)
9	Residual BW (max link bw - allocated to RSVP-TE LSPs tunels) aggred across levels					x		x	x	x (bw-residual)
10	Available BW (residual - measured actual forwarding of non-RSVP-TE packets)					x		x	x	x (bw-available)
11	Utilized BW (actual util of link)					x		x	x	
12	Max reservable BW (> max link if oversub)				x		x			
13	Unreserved (max reservable - tunnels, per priority, 8 levels)				x (8 prior levels)		x			
14	Hop count									x (hopcount)
15	Tput									x (tput)

Main Update: Structure

Metric	Definition	Origin
One-way Delay	Section 3.1	[RFC7679]
Round-trip Delay	Section 3.2	[RFC2681]
Delay Variation	Section 3.3	[RFC3393]
Hop Count	Section 3.4	[RFC7285]
Loss Rate	Section 3.5	[RFC7680]
TCP Throughput	Section 4.1	[RFC6349]
Residual Bandwidth	Section 4.2	[RFC8570]
Max Reservable Bandwidth	Section 4.3	[RFC5305]

Table 1. Cost Metrics Defined in this Document.

Metric	Definition in this doc	Semantics Based On
One-way Delay	Section 4.1	Base: [RFC7471,8570,8571] sum Unidirectional Delay
Round-trip Delay	Section 4.2	Base: Sum of two directions from above
Delay Variation	Section 4.3	Base: [RFC7471,8570,8571] sum of Unidirectional Delay Variation
Loss Rate	Section 4.4	Base: [RFC7471,8570,8571] aggr Unidirectional Link Loss
Residual Bandwidth	Section 5.2	Base: [RFC7471,8570,8571] min Unidirectional Residual BW
Available Bandwidth	Section 5.3	Base: [RFC7471,8570,8571] min Unidirectional Avail. BW
TCP Throughput	Section 5.1	[I-D.ietf-tcpm-rfc8312bis]
Hop Count	Section 4.5	[RFC7285]

Table 1. Cost Metrics Defined in this Document.

Main Update: Structure

The purpose of this document is to ensure proper usage of the performance metrics defined in Table 1; it does not claim novelty of the metrics. The "Origin" column of Table 1 gives an earlier RFC that has defined each metric.

The first 6 metrics listed in Table 1 (i.e., One-way Delay, Round-trip Delay, Delay Variation, Loss Rate, Residual Bandwidth, and Available Bandwidth) are derived from the set of traffic engineering performance metrics commonly defined in OSPF [RFC3630], [RFC7471]; IS-IS [RFC5305], [RFC8570]; and BGP-LS [RFC8571]. Deriving ALTO cost performance metrics from existing network-layer traffic engineering performance metrics, to expose to application-layer traffic optimization, can be a typical mechanism by network operators to deploy ALTO [RFC7971], [FlowDirector]. This document defines the base semantics of these metrics by extending them from link metrics to end-to-end metrics for ALTO. The "Semantics Based On" column specifies at a high level how the end-to-end metric is computed from link metrics; the details will be specified in the following sections.

The performance metrics can be classified into two categories: those derived from the performance of individual packets (i.e., one-way delay, round-trip delay, delay variation, hop count, and loss rate), and those related to bandwidth (TCP throughput, residual bandwidth, and maximum reservable bandwidth). These two categories are defined in Section 3 and Section 4 respectively. Note that all metrics except round trip delay in Table 1 are unidirectional; hence, a client will need to query both directions if needed.

The common metrics Min/Max Unidirectional Delay defined in [RFC8570][RFC8571] and Max Link Bandwidth defined in [RFC3630,RFC5305] are not listed in Table 1 because they can be handled by applying the statistical operators defined in this document. The metrics related with utilized bandwidth and reservable bandwidth (i.e., Max Reservable BW and Unreserved BW defined in [RFC3630,RFC5305]) are outside the scope of this document.

The 7th metric (the estimated TCP-flow throughput metric) provides an estimation of the bandwidth of a TCP flow, using TCP throughput modeling, to support use cases of adaptive applications [Prophet], [G2]. Note that other transport-specific metrics can be defined in the future. For example, QUIC-related metrics [RFC9000] can be considered when the methodology to measure such metrics is more mature (e.g., [I-D.corre-quic-throughput-testing]).

The 8th metric (the hop count metric) in Table 1 is mentioned in the ALTO base protocol [RFC7285], but not defined, and this document defines it to be complete.

These 8 performance metrics can be classified into two categories: those derived from the performance of individual packets (i.e., One-way Delay, Round-trip Delay, Delay Variation, Loss Rate, and Hop Count), and those related to bandwidth/throughput (Residual bandwidth, and Available Bandwidth, and TCP throughput). These two categories are defined in Sections 4 and 5 respectively. Note that all metrics except Round-trip Delay are unidirectional. An ALTO client will need to query both directions if needed.

Update: Grammar and Requirements

A cost metric string consists of a base metric identifier (or base identifier for short) string, followed by an optional statistical operator string, connected by the ASCII character colon (':', U+003A), if the statistical operator string exists.

Examples of cost metric strings then include "delay-ow", "delay-ow:min", "delay-ow:p99", where "delay-ow" is the base metric identifier string; "min" and "p99" are example statistical operator strings.

A cost metric string consists of a base metric identifier (or base identifier for short) string, followed by an optional statistical operator string, connected by the ASCII character colon (':', U+003A), if the statistical operator string exists. The total length of the cost metric string MUST NOT exceed 32, as required by [RFC7285].

If a cost metric string does not have the optional statistical operator string, the statistical operator SHOULD be interpreted as the default statistical operator in the definition of the base metric. If the definition of the base metric does not provide a definition for the default statistical operator, the metric MUST be considered as the median value.

Note that RFC 7258 limits the overall cost metric identifier to 32 characters. The cost metric variants with statistical operator suffixes defined by this document are also subject to the same overall 32-character limit, so certain combinations of (long) base metric identifier and statistical operator will not be representable. If such a situation arises, it could be addressed by defining a new base metric identifier that is an "alias" of the desired base metric, with identical semantics and just a shorter name.

Update: Conditions and Derivations

"estimation": The exact estimation method is out of the scope of this document. An example of estimating hopcounts is by importing from IGP routing protocols. It is RECOMMENDED that the "parameters" field of an "estimation" hop count metric provides a link ("link") to a description of the "estimation" method.

"estimation": See the "estimation" entry in Section 4.1.4. For estimation by aggregation of routing protocol link metrics, the default aggregation of the average of loss rate is the sum of the link link loss rates. But this default aggregation is valid only if two conditions are met: (1) it is valid only when link loss rates are low, and (2) it assumes that each link's loss events are uncorrelated with every other link's loss events. When loss rates at the links are high but independent, the general formula for aggregating loss assuming each link is independent is to compute end-to-end loss as one minus the product of the success rate for each link. Aggregation when losses at links are correlated can be more complex and the ALTO server should be cognizant of correlated loss rates. For estimation using IPPM, the IPPM metric MUST be packet loss (i.e., IPPM OWLoss* metrics). The statistical operator of the ALTO metric MUST be consistent with the IPPM statistical property (e.g., 95-th percentile).

Intended Semantics: To specify spatial and temporal aggregated packet, one-way loss rate from the specified source and the specified destination. The spatial aggregation level is specified in the query context (e.g., PID to PID, or endpoint to endpoint).

Intended Semantics: To specify the number of hops in the path from the specified source to the specified destination. The hop count is a basic measurement of distance in a network and can be exposed as the number of router hops computed from the routing protocols originating this information. A hop, however, may represent other units. The spatial aggregation level is specified in the query context (e.g., PID to PID, or endpoint to endpoint).

Update: References

"estimation": The exact estimation method is out of the scope of this document. See [Prophet] for a method to estimate TCP throughput. It is RECOMMENDED that the "parameters" field of an "estimation" TCP throughput metric provides two fields: (1) a congestion-control algorithm name (a field named "congestion-control-alg"); and (2) a link (a field named "link")to a description of the "estimation" method. Note that as TCP congestion control algorithms evolve (e.g., TCP Cubic Congestion Control [I-D.ietf-tcpm-rfc8312bis]), it helps to specify as many details as possible on the congestion control algorithm used. This description can be either free text for possible presentation to the user, or a formal specification. The semantics are out of the scope of this document.

9.1. Normative References

[RFC2818] Rescorla, E., "HTTP Over TLS", RFC 2818, DOI 10.17487/RFC2818, May 2000, <https://www.rfc-editor.org/info/rfc2818>.

"estimation": The exact estimation method is out of the scope of this document. It is RECOMMENDED that the "parameters" field of an "estimation" TCP throughput metric include the following information: (1) the congestion-control algorithm; and (2) the estimation methodology. To specify (1), it is RECOMMENDED that the "parameters" field (object) include a field named "congestion-control-algorithm", which provides a URI for the specification of the algorithm; for example, for an ALTO server to provide estimation to the throughput of a Cubic Congestion control flow, its "parameters" includes a field "congestion-control-algorithm", with value being set to [I-D.ietf-tcpm-rfc8312bis]; for an ongoing congestion control algorithm such as BBR, a a link to its specification. To specify (2), the "parameters" includes as many details as possible; for example, for TCP Cubic throughout estimation, the "parameters" field specifies that the throughput is estimated by setting _C_ to 0.4, and the Equation in Figure 8 of [I-D.ietf-tcpm-rfc8312bis] is applied; as an alternative, the methodology may be based on the NUM model [Prophet], or the G2 model [G2]. The exact specification of the parameters field is out of the scope of this document.

10.1. Normative References

[I-D.ietf-tcpm-rfc8312bis] Xu, L., Ha, S., Rhee, I., Goel, V., and L. Eggert, "CUBIC for Fast and Long-Distance Networks", Work in Progress, Internet-Draft, draft-ietf-tcpm-rfc8312bis-07, 4 March 2022, <https://www.ietf.org/archive/id/draft-ietf-tcpm-rfc8312bis-07.txt>.

[RFC7230] Fielding, R., Ed. and J. Reschke, Ed., "Hypertext Transfer Protocol (HTTP/1.1): Message Syntax and Routing", RFC 7230, DOI 10.17487/RFC7230, June 2014, <https://www.rfc-editor.org/info/rfc7230>.

Need WG Decision

- Take advantage of new round of review to add (back) metrics
  - bw-utilized (added in -21 (Sec. 4.4), removed in -23), bw-reservable, bw-reservable-level-[0-7]) => fully compatible w/ routing TE metrics set

1		IPPM 7679 OW delay	IPPM 2681 RTT	IPPM 3393 Delay Var	RFC3630 OSPF TE	RFC7471 OSPF TE Metric Ext	RFC5305 ISIS TE	RFC8570 ISIS TE Metric Ext	RFC 8571 BGP-LS	ALTO PM
2	One-way delay	x				x		x	x	x (delay-ow)
3	min-max delay					x		x	x	x (using max-min op)
4	RTT		x							x (delay-rt)
5	Delay variation			x		x		x	x	x (delay-variation)
6	Loss rate					x		x	x	x (lossrate)
7	TE default value				x		x			x (routingcost)
8	Max (link) BW				x		x			x (max bw-residual)
9	Residual BW (max link bw - allocated to RSVP-TE LSPs tunels) aggred across levels					x		x	x	x (bw-residual)
10	Available BW (residual - measured actual forwarding of non-RSVP-TE packets)					x		x	x	x (bw-available)
11	Utilized BW (actual util of link)					x		x	x	
12	Max reservable BW (> max link if oversub)				x		x			
13	Unreserved (max reservable - tunnels, per priority, 8 levels)				x (8 prior levels)		x			
14	Hop count									x (hopcount)
15	Tput									x (tput)