ALTO Working Group
Internet-Draft
Intended status: Standards Track
Expires: December 13, 2020

Q. Wu
Huawei
Y. Yang
Yale University
Y. Lee
Samsung
D. Dhody
Huawei
S. Randriamasy
Nokia Bell Labs
L. Contreras
Telefonica
June 11, 2020

ALTO Performance Cost Metrics
draft-ietf-alto-performance-metrics-11

Abstract

Cost metric is a basic concept in Application-Layer Traffic Optimization (ALTO), and is used in basic ALTO services including both the cost map service and the endpoint cost service.

Different applications may use different cost metrics, but the ALTO base protocol [RFC7285] defines only a single cost metric, i.e., the generic "routingcost" metric; see Sec. 14.2 of [RFC7285]. Hence, if the ALTO client of an application wants to issue a cost map or an endpoint cost request to determine the resource provider that offers better delay performance (i.e., low-delay) to a resource consumer, the base protocol does not define the cost metric to be used.

This document addresses the issue by introducing network performance metrics, including network delay, jitter, packet loss rate, hop count, and bandwidth. The ALTO server may derive and aggregate such performance metrics from routing protocols such as BGP-LS, OSPF-TE and ISIS-TE, or from end-to-end traffic management tools, and then expose the information to allow applications to determine "where" to connect based on network performance criteria.

There are multiple sources to derive the performance metrics. For example, whether the metric reported is an estimation based on measurements or it is a service-level agreement (SLA) can define the meaning of the performance metric. Hence, an application may need additional contextual information beyond the metric value. This document introduces an additional "cost-context" field to the ALTO "cost-type" field to convey such information.

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 [RFC2119].

Status of This Memo

This Internet-Draft will expire on December 13, 2020.

Copyright Notice

Table of Contents

1. Introduction 4
2. Performance Metric Attributes 5
  2.1. Performance Metric Context: cost-context 5
  2.2. Performance Metric Statistics 7
3. Packet Performance Metrics 8
  3.1. Cost Metric: One-Way Delay (delay-ow) 9
    3.1.1. Base Identifier 9
    3.1.2. Value Representation 9
    3.1.3. Intended Semantics and Use 9

3.1.4. Cost-Context Specification Considerations 11
3.2. Cost Metric: Round-trip Delay (delay-rt) 11
3.2.1. Base Identifier 11
3.2.2. Value Representation 11
3.2.3. Intended Semantics and Use 11
3.2.4. Cost-Context Specification Considerations 13
3.3. Cost Metric: Delay Variation (delay-variation) 13
3.3.1. Base Identifier 13
3.3.2. Value Representation 13
3.3.3. Intended Semantics and Use 13
3.3.4. Cost-Context Specification Considerations 15
3.4. Cost Metric: Hop Count (hopcount) 15
3.4.1. Base Identifier 15
3.4.2. Value Representation 15
3.4.3. Intended Semantics and Use 15
3.4.4. Cost-Context Specification Considerations 17
3.5. Cost Metric: Loss Rate (lossrate) 17
3.5.1. Base Identifier 17
3.5.2. Value Representation 17
3.5.3. Intended Semantics and Use 17
3.5.4. Cost-Context Specification Considerations 18
4. Bandwidth Performance Metrics 19
4.1. Cost Metric: TCP Throughput (tput) 19
4.1.1. Base Identifier 19
4.1.2. Value Representation 19
4.1.3. Intended Semantics and Use 19
4.1.4. Cost-Context Specification Considerations 20
4.2. Cost Metric: Residue Bandwidth 21
4.2.1. Base Identifier 21
4.2.2. Value Representation 21
4.2.3. Intended Semantics and Use 21
4.2.4. Cost-Context Specification Considerations 22
4.3. Cost Metric: Maximum Reservable Bandwidth 23
4.3.1. Base Identifier 23
4.3.2. Value Representation 23
4.3.3. Intended Semantics and Use 23
4.3.4. Cost-Context Specification Considerations 24
5. Operational Considerations 25
5.1. Source Considerations 25
5.2. Metric Timestamp Consideration 26
5.3. Backward Compatibility Considerations 26
5.4. Computation Considerations 26
5.4.1. Configuration Parameters Considerations 26
5.4.2. Availability Considerations 27
6. Security Considerations 27
7. IANA Considerations 27
8. Acknowledgments 28
9. References 28

9.1. Normative References 28
9.2. Informative References 29
Authors' Addresses 30

1. Introduction

Cost Metric is a basic concept in Application-Layer Traffic Optimization (ALTO). It is used in both the ALTO cost map service and the ALTO endpoint cost service in the ALTO base protocol [RFC7285].

Since different applications may use different cost metrics, the ALTO base protocol introduces an ALTO Cost Metric Registry (Section 14.2 of [RFC7285]), as a systematic mechanism to allow different metrics to be specified. For example, a delay-sensitive application may want to use latency related metrics, and a bandwidth-sensitive application may want to use bandwidth related metrics. The ALTO base protocol, however, has registered only one single cost metric, i.e., the generic "routingcost" metric; no latency or bandwidth related metrics are defined.

This document registers a set of new cost metrics specified in Table 1, to allow applications to better determine "where" to connect based on network performance criteria. This document follows the guideline defined in Section 14.2 of the ALTO base protocol [RFC7285]) on registering ALTO cost metrics. Hence it specifies the identifier, the intended semantics, and the security considerations of each one of the metrics defined in Table 1.

Metric	Definition	Origin Example
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]
Residue Bandwidth	Section 4.2	[RFC7810]
Max Reservable Bandwidth	Section 4.3	[RFC5305]

Table 1. Cost Metrics Defined in this Document.

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. For each performance metric, the Origin column of Table 1 gives an earlier RFC which has defined the metric. We can

rough classify the performance metrics 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 with bandwidth (TCP throughput, residue bandwidth and max reservable bandwidth). These two categories are defined in Section 3 and Section 4 respectively. Note that all metrics except round trip delay are unidirectional. Hence, a client will need to query both directions if needed.

An ALTO server may provide only a subset of the metrics described in this document. For example, those that are subject to privacy concerns should not be provided to unauthorized ALTO clients. Hence, all cost metrics defined in this document are optional and not all of them need to be exposed to a given application. When an ALTO server supports a cost metric defined in this document, it should announce this metric in its information resource directory (IRD).

An ALTO server introducing these metrics should consider security issues. As a generic security consideration on the reliability and trust in the exposed metric values, applications SHOULD rapidly give up using ALTO-based guidance if they detect that the exposed information does not preserve their performance level or even degrades it. This document discusses security considerations in more details in Section 6.

Following the ALTO base protocol, this document uses JSON to specify the value type of each defined metric. See [RFC8259] for JSON data type specification.

2. Performance Metric Attributes

2.1. Performance Metric Context: cost-context

The semantics of a performance metric depends on the source of the information. Specifically, this document defines four information sources when defining performance metrics: "nominal", and "sla" (service level agreement), "import", and "estimation".

Even given the source, precise interpretation of a performance metric value, if needed, depends on an additional set of measurement and computation parameters. For example, see Section 3.8 of [RFC7679] on items which a more complete measurement-based report should include.

To make it possible to specify both the source and the additional parameters, this document introduces an optional "cost-context" field to the "cost-type" field defined by the ALTO base protocol (Section 10.7 of [RFC7285]) as the following:

```
object {
  CostMetric   cost-metric;
  CostMode     cost-mode;
  [CostContext cost-context;]
  [JSONString  description;]
} CostType;

object {
  JSONString    cost-source;
  [JSONValue    parameters;]
} CostContext;
```

The "cost-source" field of the "cost-context" field MUST be one of four category values: "nominal", "sla", "import", and "estimation". "cost-context" will not be used as a key to distinguish among performance metrics. Hence, an ALTO information resource SHOULD NOT announce multiple CostType with the same "cost-metric" and "cost-mode". They can be placed into different information resources.

The "nominal" category indicates that the value of the metric is statically configured by the underlying devices. Not all metrics have reasonable "nominal" values. For example, throughput can have a nominal value, which indicates the configured transmission rate of the devices; latency typically do not have a nominal value.

The "sla" category indicates that the value of the metric is derived from some commitment which this document refers to as service-level agreement (SLA). Some operators also use terms such as "target" or "committed" values. For a "sla" metric, it is RECOMMENDED that the "parameters" field provides a link to the SLA definition.

The "import" category indicates that the value of the metric is derived from importing from a specific existing protocol or system. For an "import" metric, it is RECOMMENDED that the "parameters" field provides details to the system from which raw data is imported. In particular, one may notice that the set of end-to-end metrics defined in Table 1 has large overlap with the set defined in [RFC8571], in the setting of IGP traffic engineering performance metrics for each link (i.e., unidirectional link delay, min/max unidirectional link delay, unidirectional delay variation, unidirectional link loss, unidirectional residual bandwidth, unidirectional available bandwidth, unidirectional utilized bandwidth). Hence, an ALTO server may use "import" to indicate that its end-to-end metrics are computed from link metrics imported from [RFC8571].

The "estimation" category indicates that the value of the metric is computed through an estimation process. An ALTO server may compute "estimation" values by retrieving and/or aggregating information from routing protocols (e.g., [RFC8571]) and traffic measurement management tools (e.g., TWAMP), with corresponding operational issues. A potential architecture on estimating these metrics is shown in Figure 1 below. Section 5 will discuss in more detail the operational issues and how a network may address them.

```
  +--------+   +--------+  +--------+
  | Client |   | Client |  | Client |
  +----^---+   +---^----+  +---^----+
       |           |           |
       +-----------|-----------+
             NBI   |ALTO protocol
                   |
                   |
                +--+-----+  retrieval      +-----------+
                |  ALTO  |<----------------| Routing   |
                | Server |  and aggregation|           |
                |        |<-------------+  | Protocols |
                +--------+              |  +-----------+
                                        |
                                        |  +-----------+
                                        |  |Management |
                                        ---|           |
                                           |  Tool     |
                                           +-----------+
```

Figure 1. Potential framework to compute estimation to performance metrics

There can be overlap in deciding the cost-source category. It is the operator of an ALTO server who chooses the category. If a metric does not include a "cost-source" value, the application MUST assume that the value of "cost-source" is "estimation".

2.2. Performance Metric Statistics

Even with a specified cost context, a performance metric may be observed with values from an observation distribution. To address this issue, this document allows each performance metric's identifier to indicate a summary statistics of the distribution, to become <metric-base-identifier>-<stat>, where <stat> MUST be one of the following:

letter p followed by a number:

the value represents the percentile: less than or equal to number percent of observations are lower than the value (for example,

delay-ow-p75 gives the value that 75% of observed one-way delays will be less or equal to it). To avoid complex identifiers, the number MUST be a JSON number (Section 6 of [RFC8259]) without the minus or the exp component.

max:

the maximal value of the observation distribution.

min:

the minimal value of the observation distribution.

median:

the mid point of the observation distribution.

mean:

the arithmetic mean value of the observations.

stand-deviation:

the standard deviation of the observations.

If a metric has no <stat> (and hence no - as well), the metric is the 50 percentile (median). Since this scheme is common for all metrics defined in this document, below we only specify the base identifier.

3. Packet Performance Metrics

This section introduces ALTO network performance metrics including one way delay, round trip delay, delay variation, hop count, and packet loss rate. They measure the "quality of experience" of the stream of packets sent from a resource provider to a resource consumer. The measures of each individual packet (pkt) can include

the delay from the time that the packet enters the network to the time that the packet leaves the network (pkt.delay); the number of network hops that the packet traverses (pkt.hopcount); and whether the packet is dropped before reaching destination (pkt.dropped). The semantics of the performance metrics defined in this section is that they are statistics (percentiles) computed from these measures; for example, the x-percentile of the one-way delay is the x-percentile of the set of delays {pkt.delay} for the packets in the stream.

3.1. Cost Metric: One-Way Delay (delay-ow)

3.1.1. Base Identifier

The base identifier for this performance metric is "delay-ow".

3.1.2. Value Representation

The metric value type is a single 'JSONNumber' type value conforming to the number specification of [RFC8259] Section 6. The unit is expressed in milliseconds. Hence, the number can be a floating point number to express delay that is smaller than milliseconds. The number MUST be non-negative.

3.1.3. Intended Semantics and Use

Intended Semantics: To specify spatial and temporal aggregated delay of a stream of packets 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).

Use: This metric could be used as a cost metric constraint attribute or as a returned cost metric in the response.

Example 1: Delay value on source-destination endpoint pairs

```
POST /endpointcost/lookup HTTP/1.1
Host: alto.example.com
Content-Length: TBA
Content-Type: application/alto-endpointcostparams+json
Accept:
  application/alto-endpointcost+json,application/alto-error+json

{
  "cost-type": {"cost-mode" : "numerical",
                "cost-metric" : "delay-ow"},
  "endpoints" : {
    "srcs": [ "ipv4:192.0.2.2" ],
    "dsts": [
      "ipv4:192.0.2.89",
      "ipv4:198.51.100.34",
      "ipv6:2000::1:2345:6789:abcd"
    ]
  }
}

HTTP/1.1 200 OK
Content-Length: TBA
Content-Type: application/alto-endpointcost+json
{
  "meta" :{
    "cost-type": {"cost-mode" : "numerical",
                  "cost-metric" : "delay-ow"
     }
   },
    "endpoint-cost-map" : {
      "ipv4:192.0.2.2": {
        "ipv4:192.0.2.89"    : 10,
        "ipv4:198.51.100.34" : 20,
        "ipv6:2000::1:2345:6789:abcd"  : 30,
    }
  }
}
```

Comment: Since the "cost-type" does not include the "cost-source" field, the values are based on "estimation". Since the identifier does not include the -<percentile> component, the values will represent median values.

3.1.4. Cost-Context Specification Considerations

"nominal": Typically network one-way delay does not have a nominal value.

"sla": Many networks provide delay in their application-level service level agreements. It is RECOMMENDED that the "parameters" field of an "sla" one-way delay metric provides a link ("link") to the SLA definition.

"import": There can be multiple sources to import one-way delay. For example, if the import is from [RFC8571] (by using unidirectional link delay, min/max unidirectional link delay), it is RECOMMENDED that "parameters" provides "protocol" as a field and "RFC8571" as the value. During import, the server should be cognizant of potential issues when computing an end-to-end summary statistics from a link statistics. Another example import source is the IPPM framework. For IPPM, it is recommended that "parameters" provides "protocol" as a field and "ippm" as the value; see Section 4 of [I-D.ietf-ippm-initial-registry] for additional fields which can be specified for "ippm" in "parameters".

"estimation": The exact estimation method is out of the scope of this document. It is RECOMMENDED that the "parameters" field of an "estimation" one-way delay metric provides a link ("link") to a description of the "estimation" method.

3.2. Cost Metric: Round-trip Delay (delay-rt)

3.2.1. Base Identifier

The base identifier for this performance metric is "delay-rt".

3.2.2. Value Representation

The metric value type is a single 'JSONNumber' type value conforming to the number specification of [RFC8259] Section 6. The number MUST be non-negative. The unit is expressed in milliseconds.

3.2.3. Intended Semantics and Use

Intended Semantics: To specify spatial and temporal aggregated round-trip delay between the specified source and specified destination. The spatial aggregation level is specified in the query context (e.g., PID to PID, or endpoint to endpoint).

Note that it is possible for a client to query two one-way delays and then compute the round-trip delay. The server should be cognizant of the consistency of values.

Use: This metric could be used either as a cost metric constraint attribute or as a returned cost metric in the response.

Example 2: Round-trip Delay value on source-destination endpoint pairs

```
POST /endpointcost/lookup HTTP/1.1
Host: alto.example.com
Content-Length: TBA
Content-Type: application/alto-endpointcostparams+json
Accept:
  application/alto-endpointcost+json,application/alto-error+json

{
 "cost-type": {"cost-mode" : "numerical",
               "cost-metric" : "delay-rt"},
  "endpoints" : {
     "srcs": [ "ipv4:192.0.2.2" ],
     "dsts": [
       "ipv4:192.0.2.89",
       "ipv4:198.51.100.34",
       "ipv6:2000::1:2345:6789:abcd"
     ]
    }
}

  HTTP/1.1 200 OK
  Content-Length: TBA
  Content-Type: application/alto-endpointcost+json
  {
    "meta" :{
      "cost-type": {"cost-mode" : "numerical",
                    "cost-metric" : "delay-rt"
       }
     },
      "endpoint-cost-map" : {
        "ipv4:192.0.2.2": {
          "ipv4:192.0.2.89"    : 4,
          "ipv4:198.51.100.34" : 3,
          "ipv6:2000::1:2345:6789:abcd"  : 2,
      }
    }
  }
```

3.2.4. Cost-Context Specification Considerations

"nominal": Typically network round-trip delay does not have a nominal value.

"sla": It is RECOMMENDED that the "parameters" field of an "sla" round-trip delay metric provides a link ("link") to the SLA definition.

"import": There can be multiple sources to import round-trip delay. If the import is from [RFC8571] (by using unidirectional link delay, min/max unidirectional link delay), it is RECOMMENDED that "parameters" provides "protocol" as a field and "RFC8571" as the value; see Section 3.1.4 for discussions on summing up link metrics to obtain end-to-end metrics. If the import is from the IPPM framework, it is recommended that "parameters" provides "protocol" as a field and "ippm" as the value; see Section 4 of [I-D.ietf-ippm-initial-registry] for additional fields which can be specified for "ippm" in "parameters".

"estimation": The exact estimation method is out of the scope of this document. It is RECOMMENDED that the "parameters" field of an "estimation" round-trip delay metric provides a link ("link") to a description of the "estimation" method.

3.3. Cost Metric: Delay Variation (delay-variation)

3.3.1. Base Identifier

The base identifier for this performance metric is "delay-variation".

3.3.2. Value Representation

The metric value type is a single 'JSONNumber' type value conforming to the number specification of [RFC8259] Section 6. The number MUST be non-negative. The unit is expressed in milliseconds.

3.3.3. Intended Semantics and Use

Intended Semantics: To specify spatial and temporal aggregated delay variation (also called delay jitter)) with respect to the minimum delay observed on the stream over the specified source and destination. The spatial aggregation level is specified in the query context (e.g., PID to PID, or endpoint to endpoint).

Note that in statistics, variations are typically evaluated by the distance from samples relative to the mean. In networking context,

it is more commonly defined from samples relative to the min. This definition follows the networking convention.

Use: This metric could be used either as a cost metric constraint attribute or as a returned cost metric in the response.

Example 3: Delay variation value on source-destination endpoint pairs

```
POST /endpointcost/lookup HTTP/1.1
Host: alto.example.com
Content-Length: TBA
Content-Type: application/alto-endpointcostparams+json
Accept:
   application/alto-endpointcost+json,application/alto-error+json

{
  "cost-type": {"cost-mode" : "numerical",
   "cost-metric" : "delay-var"},
  "endpoints" : {
    "srcs": [ "ipv4:192.0.2.2" ],
    "dsts": [
      "ipv4:192.0.2.89",
      "ipv4:198.51.100.34",
      "ipv6:2000::1:2345:6789:abcd"
    ]
  }
}
HTTP/1.1 200 OK
 Content-Length: TBA
 Content-Type: application/alto-endpointcost+json
{
  "meta": {
           "cost type": {
           "cost-mode": "numerical",
           "cost-metric":"delay-var"
    }
   },
  "endpoint-cost-map": {
           "ipv4:192.0.2.2": {
           "ipv4:192.0.2.89"    : 0
           "ipv4:198.51.100.34" : 1
           "ipv6:2000::1:2345:6789:abcd"  : 5
         }
      }
   }
```

3.3.4. Cost-Context Specification Considerations

"nominal": Typically network delay variation does not have a nominal value.

"sla": It is RECOMMENDED that the "parameters" field of an "sla" delay variation metric provides a link ("link") to the SLA definition.

"import": There can be multiple sources to import delay variation. If the import is from [RFC8571] (by using unidirectional delay variation), it is RECOMMENDED that "parameters" provides "protocol" as a field and "RFC8571" as the value; see Section 3.1.4 for discussions on summing up link metrics to obtain end-to-end metrics. If the import is from the IPPM framework, it is recommended that "parameters" provides "protocol" as a field and "ippm" as the value; see Section 4 of [I-D.ietf-ippm-initial-registry] for additional fields which can be specified for "ippm" in "parameters".

"estimation": The exact estimation method is out of the scope of this document. It is RECOMMENDED that the "parameters" field of an "estimation" delay variation metric provides a link ("link") to a description of the "estimation" method.

3.4. Cost Metric: Hop Count (hopcount)

The metric hopcount is mentioned in [RFC7285] Section 9.2.3 as an example. This section further clarifies its properties.

3.4.1. Base Identifier

The base identifier for this performance metric is "hopcount".

3.4.2. Value Representation

The metric value type is a single 'JSONNumber' type value conforming to the number specification of [RFC8259] Section 6. The number MUST be a non-negative integer (greater than or equal to 0). The value represents the number of hops.

3.4.3. Intended Semantics and Use

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 router hops, in direct relation to the routing protocols originating this information. The spatial aggregation level is specified in the query context (e.g., PID to PID, or endpoint to endpoint).

Use: This metric could be used as a cost metric constraint attribute or as a returned cost metric in the response.

Example 4: hopcount value on source-destination endpoint pairs

```
POST /endpointcost/lookup HTTP/1.1
Host: alto.example.com
Content-Length: TBA
Content-Type: application/alto-endpointcostparams+json
Accept:
  application/alto-endpointcost+json,application/alto-error+json

  {
    "cost-type": {"cost-mode" : "numerical",
     "cost-metric" : "hopcount"},
    "endpoints" : {
      "srcs": [ "ipv4:192.0.2.2" ],
      "dsts": [
        "ipv4:192.0.2.89",
        "ipv4:198.51.100.34",
        "ipv6:2000::1:2345:6789:abcd"
      ]
    }
  }

HTTP/1.1 200 OK
Content-Length: TBA
Content-Type: application/alto-endpointcost+json
{
    "meta": {
               "cost type": {
             "cost-mode": "numerical",
             "cost-metric":"hopcount"}
       }
    },
   "endpoint-cost-map": {
           "ipv4:192.0.2.2": {
           "ipv4:192.0.2.89"   : 5,
           "ipv4:198.51.100.34": 3,
           "ipv6:2000::1:2345:6789:abcd" : 2,
                             }
             }
 }
```

3.4.4. Cost-Context Specification Considerations

"nominal": Typically hop count does not have a nominal value.

"sla": Typically hop count does not have an SLA value.

"import": There can be multiple sources to import hop count such as IGP routing protocols.

"estimation": The exact estimation method is out of the scope of this document. It is RECOMMENDED that the "parameters" field of an "estimation" hop count metric provides a link ("link") to a description of the "estimation" method.

3.5. Cost Metric: Loss Rate (lossrate)

3.5.1. Base Identifier

The base identifier for this performance metric is "lossrate".

3.5.2. Value Representation

The metric value type is a single 'JSONNumber' type value conforming to the number specification of [RFC8259] Section 6. The number MUST be non-negative. The value represents the percentage of packet losses.

3.5.3. Intended Semantics and Use

Intended Semantics: To specify spatial and temporal aggregated packet 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).

Use: This metric could be used as a cost metric constraint attribute or as a returned cost metric in the response.

Example 5: Loss rate value on source-destination endpoint pairs

```
POST /endpointcost/lookup HTTP/1.1
Host: alto.example.com
Content-Length: TBA
Content-Type: application/alto-endpointcostparams+json
Accept:
  application/alto-endpointcost+json,application/alto-error+json

  {
    "cost-type": {"cost-mode" : "numerical",
                  "cost-metric" : "lossrate"
    },
    "endpoints" : {
      "srcs": [ "ipv4:192.0.2.2" ],
      "dsts": [
        "ipv4:192.0.2.89",
        "ipv4:198.51.100.34",
        "ipv6:2000::1:2345:6789:abcd"
      ]
    }
  }

HTTP/1.1 200 OK
Content-Length: TBA
Content-Type: application/alto-endpointcost+json
{
    "meta": {
      "cost-type": {
        "cost-mode": "numerical",
        "cost-metric":"lossrate"
      }
    },
   "endpoint-cost-map": {
      "ipv4:192.0.2.2": {
        "ipv4:192.0.2.89"   : 0,
        "ipv4:198.51.100.34": 0,
        "ipv6:2000::1:2345:6789:abcd" : 0,
      }
    }
 }
```

3.5.4. Cost-Context Specification Considerations

"nominal": Typically packet loss rate does not have a nominal value, although some networks may specify zero losses.

"sla": It is RECOMMENDED that the "parameters" field of an "sla" packet loss rate provides a link ("link") to the SLA definition.

"import": There can be multiple sources to import packet loss rate. If the import is from [RFC8571] (by using unidirectional link loss), it is RECOMMENDED that "parameters" provides "protocol" as a field and "RFC8571" as the value; see Section 3.1.4 for discussions on summing up link metrics to obtain end-to-end metrics. If the import is from the IPPM framework, it is recommended that "parameters" provides "protocol" as a field and "ippm" as the value; see Section 4 of [I-D.ietf-ippm-initial-registry] for additional fields which can be specified for "ippm" in "parameters".

"estimation": The exact estimation method is out of the scope of this document. It is RECOMMENDED that the "parameters" field of an "estimation" packet loss rate metric provides a link ("link") to a description of the "estimation" method.

4. Bandwidth Performance Metrics

This section introduces three bandwidth related metrics. Given a specified source to a specified destination, these metrics reflect the volume of traffic that the network can carry from the source to the destination.

4.1. Cost Metric: TCP Throughput (tput)

4.1.1. Base Identifier

The base identifier for this performance metric is "tput".

4.1.2. Value Representation

The metric value type is a single 'JSONNumber' type value conforming to the number specification of [RFC8259] Section 6. The number MUST be non-negative. The unit is bytes per second.

4.1.3. Intended Semantics and Use

Intended Semantics: To give the throughput of a TCP flow from the specified source to the specified destination. The spatial aggregation level is specified in the query context (e.g., PID to PID, or endpoint to endpoint).

Use: This metric could be used as a cost metric constraint attribute or as a returned cost metric in the response.

Example 5: TCP throughput value on source-destination endpoint pairs

```
POST /endpointcost/lookup HTTP/1.1
Host: alto.example.com
Content-Length: TBA
Content-Type: application/alto-endpointcostparams+json
Accept:
  application/alto-endpointcost+json,application/alto-error+json

{
  "cost-type": {"cost-mode" : "numerical",
                "cost-metric" : "tput"},
  "endpoints" : {
    "srcs": [ "ipv4:192.0.2.2" ],
    "dsts": [
       "ipv4:192.0.2.89",
       "ipv4:198.51.100.34",
       "ipv6:2000::1:2345:6789:abcd"
    ]
  }
}

HTTP/1.1 200 OK
Content-Length: TBA
Content-Type: application/alto-endpointcost+json
{
  "meta": {
     "cost type": {
        "cost-mode": "numerical",
        "cost-metric":"tput"
    }
  }
  "endpoint-cost-map": {
    "ipv4:192.0.2.2": {
      "ipv4:192.0.2.89"   : 256000,
      "ipv4:198.51.100.34": 128000,
      "ipv6:2000::1:2345:6789:abcd" : 428000,
  }
}
```

4.1.4. Cost-Context Specification Considerations

"nominal": Typically TCP throughput does not have a nominal value.

"sla": Typically TCP throughput does not have an SLA value.

"import": Typically there is not a routing protocol through which one can import TCP throughput. If the import is from the IPPM framework,

it is recommended that "parameters" provides "protocol" as a field and "ippm" as the value; see Section 4 of [I-D.ietf-ippm-initial-registry] for additional fields which can be specified for "ippm" in "parameters".

"estimation": The exact estimation method is out of the scope of this document. See [ProphetINFOCOM18] for a method to estimate TCP throughput. It is RECOMMENDED that the "parameters" field of an "estimation" TCP throughput metric provides a link ("link") to a description of the "estimation" method.

4.2. Cost Metric: Residue Bandwidth

4.2.1. Base Identifier

The base identifier for this performance metric is "bw-residue".

4.2.2. Value Representation

The metric value type is a single 'JSONNumber' type value that is non-negative. The unit of measurement is bytes per second.

4.2.3. Intended Semantics and Use

Intended Semantics: To specify spatial and temporal residual bandwidth from the specified source and the specified destination. The value is calculated by subtracting tunnel reservations from Maximum Bandwidth (motivated from [RFC7810], Section 4.5.). The spatial aggregation unit is specified in the query context (e.g., PID to PID, or endpoint to endpoint).

Use: This metric could be used either as a cost metric constraint attribute or as a returned cost metric in the response.

Example 7: bw-residue value on source-destination endpoint pairs

```
POST/ endpointcost/lookup HTTP/1.1
Host: alto.example.com
Content-Length: TBA
Content-Type: application/alto-endpointcostparams+json
Accept:
  application/alto-endpointcost+json,application/alto-error+json

  {
   "cost-type": { "cost-mode":   "numerical",
                  "cost-metric": "bw-residue"},
   "endpoints":  {
     "srcs": [ "ipv4 : 192.0.2.2" ],
     "dsts": [
       "ipv4:192.0.2.89",
       "ipv4:198.51.100.34",
       "ipv6:2000::1:2345:6789:abcd"
     ]
   }
  }

HTTP/1.1 200 OK
Content-Length: TBA
Content-Type: application/alto-endpointcost+json
{
  "meta": {
    "cost-type" {
      "cost-mode": "numerical",
      "cost-metric": "bw-residue"
    }
  },
  "endpoint-cost-map" {
    "ipv4:192.0.2.2" {
      "ipv4:192.0.2.89" :    0,
      "ipv4:198.51.100.34": 2000,
      "ipv6:2000::1:2345:6789:abcd":  5000,
    }
  }
}
```

4.2.4. Cost-Context Specification Considerations

"nominal": Typically residue bandwidth does not have a nominal value.

"sla": Typically residue bandwidth does not have an "sla" value.

"import": There can be multiple sources to import residue bandwidth. If the import is from [RFC8571] (by using unidirectional residue bandwidth), it is RECOMMENDED that "parameters" provides "protocol" as a field and "RFC8571" as the value. The server should be cognizant of issues when computing end-to-end summary statistics from link statistics. For example, the min of the end-to-end path residue bandwidth is the min of all links on the path.

"estimation": The exact estimation method is out of the scope of this document. It is RECOMMENDED that the "parameters" field of an "estimation" residue bandwidth metric provides a link ("link") to a description of the "estimation" method.

4.3. Cost Metric: Maximum Reservable Bandwidth

4.3.1. Base Identifier

The base identifier for this performance metric is "bw-maxres".

4.3.2. Value Representation

The metric value type is a single 'JSONNumber' type value that is non-negative. The unit of measurement is bytes per second.

4.3.3. Intended Semantics and Use

Intended Semantics: To specify spatial and temporal maximum reservable bandwidth from the specified source to the specified destination. The value is corresponding to the maximum bandwidth that can be reserved (motivated from RFC 3630 Sec. 2.5.7.). The spatial aggregation unit is specified in the query context (e.g., PID to PID, or endpoint to endpoint).

Use: This metric could be used either as a cost metric constraint attribute or as a returned cost metric in the response.

Example 6: bw-maxres value on source-destination endpoint pairs

```
POST/ endpointcost/lookup HTTP/1.1
Host: alto.example.com
Content-Length: TBA
Content-Type: application/alto-endpointcostparams+json
Accept:
  application/alto-endpointcost+json,application/alto-error+json

  {
    "cost-type" { "cost-mode":   "numerical",
                  "cost-metric": "bw-maxres"},
    "endpoints":  {
      "srcs": [ "ipv4 : 192.0.2.2" ],
      "dsts": [
        "ipv4:192.0.2.89",
        "ipv4:198.51.100.34",
        "ipv6:2000::1:2345:6789:abcd"
      ]
    }
  }

HTTP/1.1 200 OK
Content-Length: TBA
Content-Type: application/alto-endpointcost+json
{
  "meta": {
    "cost-type": {
      "cost-mode":   "numerical",
      "cost-metric": "bw-maxres"
    }
  },
  "endpoint-cost-map": {
    "ipv4:192.0.2.2" {
      "ipv4:192.0.2.89" :    0,
      "ipv4:198.51.100.34": 2000,
      "ipv6:2000::1:2345:6789:abcd":  5000,
    }
  }
}
```

4.3.4. Cost-Context Specification Considerations

"nominal": Typically maximum reservable bandwidth does not have a nominal value.

"sla": Typically maximum reservable bandwidth does not have an "sla" value.

"import": There can be multiple sources to import maximum reservable bandwidth. For example, Maximum reservable bandwidth is defined by IS-IS/OSPF TE, and measures the reservable bandwidth between two directly connected IS-IS neighbors or OSPF neighbors; see Section 3.5 of [RFC5305]. If the import is from [RFC8571] (by using unidirectional maximum reservable bandwidth), it is RECOMMENDED that "parameters" provides "protocol" as a field and "RFC8571" as the value.

"estimation": The exact estimation method is out of the scope of this document. It is RECOMMENDED that the "parameters" field of an "estimation" maximum reservable bandwidth metric provides a link ("link") to a description of the "estimation" method.

5. Operational Considerations

The exact measurement infrastructure, measurement condition and computation algorithms can vary from different networks, and are outside the scope of this document. Both the ALTO server and the ALTO clients, however, need to be cognizant of the operational issues discussed below.

Also, the performance metrics specified in this document are similar, in that they may use similar data sources and have similar issues in their calculation. Hence, we specify common issues unless one metric has its unique challenges.

5.1. Source Considerations

The addition of the "cost-source" field is to solve a key issue: An ALTO server needs data sources to compute the cost metrics described in this document and an ALTO client needs to know the data sources to better interpret the values.

To avoid too fine-grained information, this document introduces "cost-source" to indicate only the high-level type of data sources: "estimation" or "sla", where "estimation" is a type of measurement data source and "sla" is a type that is more based on policy.

For estimation, for example, the ALTO server may use log servers or the OAM system as its data source [RFC7971]. In particular, the cost metrics defined in this document can be computed using routing systems as the data sources. Mechanisms defined in [RFC2681], [RFC3393], [RFC7679], [RFC7680], [RFC3630], [RFC3784], [RFC7471], [RFC7810], [RFC7752] and [I-D.ietf-idr-te-pm-bgp] that allow an ALTO Server to retrieve and derive the necessary information to compute the metrics that we describe in this document.

5.2. Metric Timestamp Consideration

Despite the introduction of the additional cost-context information, there is no built-in field to indicate the timestamps of the data used to compute a metric. To indicate this attribute, the ALTO server SHOULD return HTTP "Last-Modified", to indicate the freshness of the data used to compute the performance metrics.

If the ALTO client obtains updates through an incremental update mechanism (e.g., [ALTO SSE]), the client SHOULD assume that the metric is computed using a snapshot at the time that is approximated by the receiving time.

5.3. Backward Compatibility Considerations

One potential issue introduced by the optional "cost-source" field is backward compatibility. Consider that an IRD which defines two cost-types with the same "cost-mode" and "cost-metric", but one with "cost-source" being "estimation" and the other being "sla". Then an ALTO client that is not aware of the extension will not be able to distinguish between these two types. A similar issue can arise even with a single cost-type which has "cost-source" being "sla", but the backward client will ignore this field and consider the metric estimation.

To address this issue, the only defined "routingcost" metric can be ONLY "estimation".

5.4. Computation Considerations

The metric values exposed by an ALTO server may result from additional processing on measurements from data sources to compute exposed metrics. This may involve data processing tasks such as aggregating the results across multiple systems, removing outliers, and creating additional statistics. There are two challenges on the computation of ALTO performance metrics.

5.4.1. Configuration Parameters Considerations

Performance metrics often depend on configuration parameters. For example, the value of packet loss rate depends on the measurement interval and varies over time. To handle this issue, an ALTO server may collect data on time periods covering the previous and current time or only collect data on present time. The ALTO server may further aggregate these data to provide an abstract and unified view that can be more useful to applications. To make the ALTO client better understand how to use these performance data, the ALTO server

may provide the client with the validity period of the exposed metric values.

5.4.2. Availability Considerations

Applications value information relating to bandwidth availability whereas bandwidth related metrics can often be only measured at the link level. This document specifies a set of link-level bandwidth related values that may be exposed as such by an ALTO server. The server may also expose other metrics derived from their aggregation and having different levels of endpoint granularity, e.g., link endpoints or session endpoints. The metric specifications may also expose the utilized aggregation laws.

6. Security Considerations

The properties defined in this document present no security considerations beyond those in Section 15 of the base ALTO specification [RFC7285].

However concerns addressed in Sections "15.1 Authenticity and Integrity of ALTO Information", "15.2 Potential Undesirable Guidance from Authenticated ALTO Information" and "15.3 Confidentiality of ALTO Information" remain of utmost importance. Indeed, TE performance is a highly sensitive ISP information, therefore, sharing TE metric values in numerical mode requires full mutual confidence between the entities managing the ALTO Server and Client. Numerical TE performance information will most likely be distributed by ALTO Servers to Clients under strict and formal mutual trust agreements. On the other hand, ALTO Clients must be cognizant on the risks attached to such information that they would have acquired outside formal conditions of mutual trust.

7. IANA Considerations

IANA has created and now maintains the "ALTO Cost Metric Registry", listed in Section 14.2, Table 3 of [RFC7285]. This registry is located at <http://www.iana.org/assignments/alto-protocol/alto-protocol.xhtml#cost-metrics>. This document requests to add the following entries to "ALTO Cost Metric Registry".

Identifier	Intended Semantics
delay-ow	See Section 3.1
delay-rt	See Section 3.2
delay-var	See Section 3.3
hopcount	See Section 3.4
lossrate	See Section 3.5
tput	See Section 4.1
bw-residue	See Section 4.2
bw-maxres	See Section 4.3

This document requests the creation of the "ALTO Cost Source Registry" with the following currently defined values:

Identifier	Intended Semantics
nominal	Values in nominal cases
sla	Values reflecting service level agreement
import	Values from a given protocol
estimation	Values by estimation

8. Acknowledgments

The authors of this document would also like to thank Brian Trammell, Haizhou Du, Kai Gao, Lili Liu, Geng Li, Danny Alex Lachos Perez for the reviews and comments. Young Lee is an author of an earlier version of the document.

9. References

9.1. Normative References

[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>.

[RFC2679] Almes, G., Kalidindi, S., and M. Zekauskas, "A One-way Delay Metric for IPPM", RFC 2679, DOI 10.17487/RFC2679, September 1999, <https://www.rfc-editor.org/info/rfc2679>.

[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>.

[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>.

[RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, DOI 10.17487/RFC5305, October 2008, <https://www.rfc-editor.org/info/rfc5305>.

[RFC6349] Constantine, B., Forget, G., Geib, R., and R. Schrage, "Framework for TCP Throughput Testing", RFC 6349, DOI 10.17487/RFC6349, August 2011, <https://www.rfc-editor.org/info/rfc6349>.

[RFC7285] Alimi, R., Ed., Penno, R., Ed., Yang, Y., Ed., Kiesel, S., Previdi, S., Roome, W., Shalunov, S., and R. Woundy, "Application-Layer Traffic Optimization (ALTO) Protocol", RFC 7285, DOI 10.17487/RFC7285, September 2014, <https://www.rfc-editor.org/info/rfc7285>.

[RFC7810] Previdi, S., Ed., Giacalone, S., Ward, D., Drake, J., and Q. Wu, "IS-IS Traffic Engineering (TE) Metric Extensions", RFC 7810, DOI 10.17487/RFC7810, May 2016, <https://www.rfc-editor.org/info/rfc7810>.

[RFC8259] Bray, T., Ed., "The JavaScript Object Notation (JSON) Data Interchange Format", STD 90, RFC 8259, DOI 10.17487/RFC8259, December 2017, <https://www.rfc-editor.org/info/rfc8259>.

9.2. Informative References

[ProphetINFOCOM18] Gao, K., Zhang, J., and YR. Yang, "Prophet: Fast, Accurate Throughput Prediction with Reactive Flows", IEEE INFOCOM 2018 - IEEE Conference on Computer Communications 16-19 April 2018, 2018.

[RFC6390] Clark, A. and B. Claise, "Guidelines for Considering New Performance Metric Development", BCP 170, RFC 6390, DOI 10.17487/RFC6390, October 2011, <https://www.rfc-editor.org/info/rfc6390>.

[RFC7971] Stiemerling, M., Kiesel, S., Scharf, M., Seidel, H., and S. Previdi, "Application-Layer Traffic Optimization (ALTO) Deployment Considerations", RFC 7971, DOI 10.17487/RFC7971, October 2016, <https://www.rfc-editor.org/info/rfc7971>.

Authors' Addresses

Qin Wu
Huawei
101 Software Avenue, Yuhua District
Nanjing, Jiangsu 210012
China

Email: bill.wu@huawei.com

Y. Richard Yang
Yale University
51 Prospect St
New Haven, CT 06520
USA

Email: yry@cs.yale.edu

Young Lee
Samsung
1700 Alma Drive, Suite 500
Plano, TX 75075
USA

Email: leeyoung@huawei.com

Dhruv Dhody
Huawei
Leela Palace
Bangalore, Karnataka 560008
INDIA

Email: dhruv.ietf@gmail.com

Sabine Randriamasy
Nokia Bell Labs
Route de Villejust
Nozay 91460
FRANCE

Email: sabine.randriamasy@nokia-bell-labs.com

Luis Miguel Contreras Murillo
Telefonica

Email: luismiguel.contrerasmurillo@telefonica.com