ALTO Working Group Internet-Draft Intended status: Standards Track Expires: April 21, 2016

ALTO Extension: Endpoint Cost Service for Flows draft-wang-alto-ecs-flows-00

Abstract

The Endpoint Cost Service (ECS) has limitations to illustrate the network condition and to work with the OpenFlow protocol. This document extends ECS to improve the Application-Layer Traffic Optimization (ALTO) protocol by defining more types of endpoint address such as port number, protocol type, domain and etc.

Status of This Memo

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

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 <u>http://datatracker.ietf.org/drafts/current/</u>.

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 21, 2016.

Copyright Notice

Copyright (c) 2015 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to <u>BCP 78</u> and the IETF Trust's Legal Provisions Relating to IETF Documents (<u>http://trustee.ietf.org/license-info</u>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in <u>Section 4</u>.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

$\underline{1}$. Introduction								
<u>1.1</u> . Terminology								
1.2. Requirements Language								
<u>2</u> . Motivations								
<u>3</u> . Overview Of Approach								
3.1. Multi-Field Typed Endpoint Address Format								
3.2. Compatibility With Legacy Clients								
<u>3.3</u> . Endpoint Cost Resources								
4. Protocol Extension for Flow-Extended ECS								
<u>4.1</u> . Endpoints Extensions								
<u>4.1.1</u> . Multi-Field Typed Endpoint Addresses								
<u>4.1.2</u> . Address Type								
<u>4.1.3</u> . Endpoint Address								
<u>4.1.4</u> . Field Name								
<u>4.1.5</u> . Field Value								
<u>4.2</u> . Endpoint Cost Service Extensions								
<u>4.2.1</u> . Accept Input Parameters								
<u>4.2.2</u> . Capabilities								
<u>4.2.3</u> . Response								
<u>4.3</u> . ALTO Address Type Registry Extensions <u>11</u>								
<u>5</u> . Examples								
5.1. Information Resource Directory								
<u>5.2</u> . Endpoint Cost Service								
<u>6</u> . IANA Considerations								
7. Privacy And Security Considerations								
$\underline{8}$. Normative References								
Authors' Addresses								

<u>1</u>. Introduction

ECS is a basic service of the ALTO services defined in [RFC7285]. Based on the simple host model when defining endpoints, ECS defined in [RFC7285] may not work well in an emerging settings such as Software Defined Networking (SDN) based settings, where network routing decisions can be flow based. In this document, we present an extended ECS for such new settings.

<u>1.1</u>. Terminology

This document uses terms defined as follows:

o {1.2.3}: References of this form are to sections in the ALTO
protocol specification [<u>RFC7285</u>].

[Page 2]

o And other terms defined in {8.2} of [RFC7285].

<u>1.2</u>. 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 <u>RFC 2119</u> [<u>RFC2119</u>].

Motivations

Below is the acceptable input parameters of ECS in $\{11.5.1.3\}$ of [RFC7285].

oł	oject {					
	CostType	cost-type;				
	[JSONString	<pre>constraints<0*>;]</pre>				
	EndpointFilter	endpoints;				
}	ReqEndpointCostMap;					

object {
 [TypedEndpointAddr srcs<0..*>;]
 [TypedEndpointAddr dsts<0..*>;]
} EndpointFilter;

Hence, the granularity is TypedEndpointAddr, which is defined in {10.4.1} of [<u>RFC7285</u>]. In particular, [<u>RFC7285</u>] defines two address types: ipv4 and ipv6. This, however, may limit the usage of ECS in multiple settings. Below we give some use cases.

Use case 1:

ECS is not compatible with virtual host technology, a popular on the Internet, which allows different hosts to share the same IP address. For example, a reverse proxy p1 hosts three sites shown in Figure 1. These sites share the same public network address: 202.180.1.11. Suppose the link in the private network from p1 to server s1 is busy, but the link to server s2 is free. It will cost client c1 more to access www.a.com than www.b.com. Suppose c1 wants to know the cost to www.a.com. Because ECS only supports IP addresses, it will query the DNS server to transfer the domain name to IP address. Therefore, c1 can only obtain the cost to p1.

As a result, c1 will get the same result to three different domain names because ECS is only capable of measuring the cost between IP addresses.

[Page 3]

				+					
					Privat	е	+ -		+
					Networ	k		s1	
						-	+>	www.a.com	
				Í				192.168.1.10	Í
				I			i i		i
				i i			, , +-		+
				1			1 1		
				I					
+-		+	+	+	·	-+	+-		+
	c1			p1	-			s2	
	Web Browser	+	->	Reverse	e Proxy	+	+>	www.b.com	
	60.20.100.11			202.18	80.1.11			192.168.1.11	
T				192.16	8.1.20				
+-		+	+	+		-+	+-		+
				1			I		i
				1			' +-		+
				1			, . , ,	c2	
				I				55	
						-	+>	WWW.C.COM	
								192.168.1.12	
							+ -		+
				1					
				ا ب					ا ب

Figure 1: Using reverse proxy to operate virtual hosts.

Use case 2:

ECS is not compatible with port-based or protocol-based routing systems. For example, the OpenFlow protocol can forward packets to different destinations by the port in TCP/IP protocols. A simple topology is shown in Figure 2, the link between switch sw1 and switch sw2 has a low speed but a low latency, while sw3 is a high speed but high latency switch. And there are two services running on host h2, SSH and FTP, using port 22 and port 20, respectively. An efficient flow configuration supported by OpenFlow, is to use a low latency link to transfer SSH packets and a high speed route to transfer files. So sw1 and sw2 will exchange the SSH flows with each other to achieve a lower latency and forward FTP flows to sw3 to achieve a higher bandwidth.

In this case, the SDN network uses suitable links to transfer different packets, so the cost between IP addresses is protocol or port related. However, ECS cannot use this information to give a precise result.

[Page 4]



Figure 2: A simple example of protocol or port based routing.

Use case 3:

ECS is not compatible with other addresses such as MAC addresses or physical connectors. For example, some protocols such as ARP send packets by MAC addresses. ECS is unable to measure the cost between two NICs without IP addresses. The ALTO, as an information source, cannot compute the cost between two physic ports, either. These knowledge seems useless for the Internet users, but useful for ISPs.

3. Overview Of Approach

This section contains the non-normative overview of the ECS extension for flows defined in this document. It assumes that the readers are familiar with the ALTO Protocol ([RFC7285]).

3.1. Multi-Field Typed Endpoint Address Format

The typed endpoint address used by ECS is defined in {10.4} of [RFC7285]. That section only defines two address types: ipv4 and ipv6 to refer IPv4 addresses and IPv6 addresses, respectively. However, the flow-extended ECS may contain MAC addresses, domain names and port numbers to give a cost between hosts.

Therefore, this document extends the address type and the typed endpoint address to allow ECS to measure the cost more precisely.

[Page 5]

ECS Flow

For example, the following are some Multi-Field Type Endpoint Addresses. These fields in an address are interpreted as being related to each other with a logical AND.

> "ipv4:10.60.10.1;protocol:ssh;port:22" "ipv6:2001:da8::10;protocol:ftp;port:21" "domainname:www.a.com;protocol:http;port:80"

And a request to query the cost between hosts looks like this:

3.2. Compatibility With Legacy Clients

The extension defined in this document should be compatible with legacy implementations, which means clients and servers are not aware of this extension. A good way to achieve this goal is defining new media types for extended endpoint cost map. Based on the fact that the extended address looks alike the original typed endpoint address, it would be a simpler way to implement a parser which can handle both typed endpoint addresses.

Therefore, no new media type is defined in this document. Instead, this document extends the specifications of Information Resource Directory (IRD) in the ALTO protocol. Because the legacy ALTO clients MUST ignore unknown fields (see {8.3.7} of [<u>RFC7285</u>]), the legacy implementations will not use the extended typed endpoint address and are not aware of the existence of this extension.

<u>3.3</u>. Endpoint Cost Resources

This document extends the endpoint cost service in IRD to allow the same resource to receive either legacy typed endpoint addresses as defined in $\{10.4\}$ of [RFC7285], or extended typed endpoint addresses as defined in this document. The extended endpoint cost resource has a new capability called "flow-extension", which indicates supported

[Page 6]

address types and fields. The existence of this value means that the resource understands this extension.

```
For example, an extended endpoint cost resource in IRD is shown below:
```

The legacy implementations SHOULD ignore the unknown capability: "flow-extension", and will send requests with the typed endpoint address, as defined in [<u>RFC7285</u>]. So the ALTO server should return a non-extended legacy cost map.

However, an extended client will realize that this resource supports the extension to flows, and can POST the request with extended typed endpoint addresses. The server can understand that the client supports the extension in this document by parsing the extended endpoint addresses, and hence response with the extended cost map.

4. Protocol Extension for Flow-Extended ECS

<u>4.1</u>. Endpoints Extensions

This document extends Endpoint, as defined in {10.4} of [<u>RFC7285</u>], by adding ';' character to separate different fields in an address, and by adding more address types to indicate other fields.

4.1.1. Multi-Field Typed Endpoint Addresses

The Typed Endpoint Addresses, as defined in {10.4.1} of [<u>RFC7285</u>], are encoded as strings of the format: AddressType:EndpointAddr. This document extends it as the format: AddressType:EndpointAddr((;FieldNa me:FieldValue)|(;AddressType:EndpointAddr))*.

[Page 7]

ECS Flow

4.1.2. Address Type

This document extends Address Type, as defined in {10.4.2} of [<u>RFC7285</u>], by adding two values to AddressType: "mac" and "domainname" to refer to MAC addresses and domain names.

<u>4.1.3</u>. Endpoint Address

This document extends Endpoint Address, as defined in {10.4.3} of [<u>RFC7285</u>], by defining more EndpointAddr when AddressType is "mac" or "domainname".

4.1.3.1. MAC

MAC Endpoint Addresses are encoded as specified in <u>Section 3 of</u> [RFC7042].

4.1.3.2. Domain Name

Domain Name Addresses are encoded as specified in <u>Section 3.1 of</u> [RFC1034].

4.1.4. Field Name

The FieldName component of MultiFiledTypedEndPointAddr is defined as a string consisting of only US-ASCII alphanumeric characters (U+0030-U+0039, U+0041-U+005A, and U+0061-U+007A). The type FieldName is used in this document to indicate a string of this format. This document does not define any value of FieldName. Hence, ALTO clients and servers SHOULD interpret the meanings of FieldName by themselves.

4.1.5. Field Value

The FieldValue component of MultiFiledTypedEndPointAddr is defined as a string consisting of only US-ASCII alphanumeric characters (U+0030-U+0039, U+0041-U+005A, and U+0061-U+007A). The type FieldValue is used in this document to indicate a string of this format. This document does not define any value of FieldValue. Hence, ALTO clients and servers SHOULD interpret the meanings of FieldValue by themselves.

4.2. Endpoint Cost Service Extensions

This document extends the Endpoint Cost Service, as defined in {11.5.1} of [<u>RFC7285</u>], by adding new capabilities and extending TypedEndpointAddr.

[Page 8]

The media type ($\{11.5.1.1\}$ of [<u>RFC7285</u>]), HTTP method ($\{11.5.1.2\}$ of [<u>RFC7285</u>]) and "uses" specifications ($\{11.5.1.5\}$ of [<u>RFC7285</u>]) are unchanged.

<u>4.2.1</u>. Accept Input Parameters

The ReqEndpointCostMap object in {11.5.1.3} of [<u>RFC7285</u>] is extended as follows:

```
object {
 CostType
                             cost-type;
  [JSONString
                             constraints<0..*>;]
 EndpointFilter
                             endpoints;
  [MultiFieldEndpointFilter multi-field-endpoints;]
} RegEndpointCostMap;
object {
  [TypedEndpointAddr srcs<0..*>;]
  [TypedEndpointAddr dsts<0..*>;]
} EndpointFilter;
object {
  [MultiFieldTypedEndpointAddr srcs<0..*>;]
  [MultiFieldTypedEndpointAddr dsts<0..*>;]
```

```
} MultiFieldEndpointFilter;
```

With fields:

```
cost-type, constrains, endpoints: As defined in {11.5.1.3} of
[RFC7285].
```

multi-field-endpoints: If present, this field is the same as endpoints defined in {11.5.1.3} of [RFC7285], with the additional requirement that the ALTO server MUST ignore the endpoints field, and use srcs and dsts in the multi-field-endpoints to calculate the cost and return an extended response in this document. The AddressType and FieldName in each MultiFieldTypedEndpointAddr MUST be declared in the capability of this resource.

4.2.2. Capabilities

The EndpointCostCapabilities object in {11.5.1.4} of [<u>RFC7285</u>] is extended as follows:

[Page 9]

```
object {
   JSONString cost-type-names<1..*>;
   JSONBool cost-constraints;
   [FlowExtension flow-extension;]
} EndpointCostCapabilities;
object {
```

```
[JSONString address-types<1..*>;]
[JSONString field-names<1..*>;]
} FlowExtension;
```

With fields:

- cost-type-names, cost-constrains: As defined in {11.5.1.4} of
 [RFC7285].
- flow-extension: Defines lists of supported address types and field names of this resource. If present with one or two lists, this resource understands the flow-extension in this document and MUST support MultiFieldTypedEndpointAddr. If omitted, the default value is an empty object.
- address-types: Defines a list of supported address type. The values in this list MUST be defined in <u>Section 4.1.2</u> in this document. If present with one or more values, the ALTO server MUST support these values as extended endpoint address type. If omitted, the default value is an empty list, which means the server does not support any address type other than ipv4 and ipv6.
- field-names: Defines a list of supported field names. The values in this list are defined in Section 4.1.4 in this document. If present with one or more values, the ALTO server MUST support these values as field names. If omitted, the default value is an empty list.

4.2.3. Response

If the client does not provide "multi-field-endpoints" input parameter, the response is exactly as defined in {11.5.1.6} of [<u>RFC7285</u>]. If the client provides the "multi-field-endpoints", the response is changed as follows:

- o In the EndpointCostMapData object, replace the TypedEndpointAddr field with MultiFieldTypedEndpointAddr.
- o In the EndpointDstCosts object, replace the TypedEndpointAddr field with MultiFieldTypedEndpointAddr.

4.3. ALTO Address Type Registry Extensions

This document requests registration of the identifiers "mac" and "domainname" as shown in Table 1.

Identifier 	Address Encoding	Prefix Encoding	++ Mapping to/from IPv4/v6
mac domainname 	See Section 6.1.3 See Section 6.1.3	No compact encoding is available No compact encoding is available	<pre>Mapping from IPv4 by [RFC0826]. Mapping to IPv4 by [RFC0903]. Mapping from IPv6 by [RFC4861]. Mapping to IPv6 by [RFC3122]. Mapping to/from IPv4 by [RFC1034]. Mapping to/from IPv6 by [RFC3596]. </pre>

Table 1: New ALTO Address Types

5. Examples

<u>5.1</u>. Information Resource Directory

Here is an example of an ALTO server's Information Resource Directory with an Endpoint Cost Service which supports the flow-based ECS extensions.

```
GET /directory HTTP/1.1
Host: alto.example.com
Accept: application/alto-directory+json,application/alto-error+json
HTTP/1.1 200 OK
Content-Length: [TODO]
Content-Type: application/alto-directory+json
{
  "meta" : {
    "default-alto-network-map" : "my-default-network-map",
    "cost-types" : {
       "num-routing" : {
         "cost-mode" : "numerical",
         "cost-metric" : "routingcost"
       },
       "num-hopcount" : {
         "cost-mode" : "numerical",
```

}

```
"cost-metric" : "hopcount"
     },
  }
},
"resources" : {
    "my-default-network-map" : {
      "uri" : "http://alto.example.com/networkmap",
      "media-type" : "application/alto-networkmap+json"
    },
    "numerical-routing-cost-map" : {
      "uri" : "http://alto.example.com/costmap/num-routing",
      "media-types" : [ "application/alto-costmap+json" ],
      "uses" : [ "my-default-network-map" ],
      "capabilities" : {
        "cost-type-names" : [ "num-routing" ]
      }
    },
    "numerical-hopcount-cost-map" : {
      "uri" : "http://alto.example.com/costmap/num-hopcount",
      "media-types" : [ "application/alto-costmap+json" ],
      "uses" : [ "my-default-network-map" ],
      "capabilities" : {
        "cost-type-names" : [ "num-hopcount" ]
      }
    },
    . . . . . . . . .
    And other information resources described in RFC7285
    . . . . . . . . .
    "endpoint-multicost-map" : {
      "uri" : "http://alto.example.com/multi/endpointcost/lookup",
      "media-types" : [ "application/alto-endpointcost+json" ],
      "accepts" : [ "application/alto-endpointcostparams+json" ],
      "uses" : [ "my-default-network-map" ],
      "capabilities" : {
        "cost-constraints" : true,
        "flow-extension" : {
          "address-types" : [ "mac", "domainname" ],
          "field-names" : [ "port", "protocol" ]
        },
        "cost-type-names" : [ "num-routingcost",
                               "num-hopcount" ],
      }
  }
}
```

5.2. Endpoint Cost Service

```
This example uses multi-field typed endpoint addresses to query the
"routingcost" for selected endpoints.
       POST /endpointcost/lookup HTTP/1.1
       Host: alto.example.com
       Content-Length: 345
       Content-Type: application/alto-endpointcostparams+json
       Accept: application/alto-endpointcost+json,
               application/alto-error+json
       {
         "cost-type": {"cost-mode" : "ordinal",
                       "cost-metric" : "routingcost"},
         "multi-field-endpoints" : {
           "srcs": [ "ipv4:192.0.2.2" ],
           "dsts": [
             "domainname:www.a.com",
             "mac:01-23-45-67-89-AB",
             "ipv4:198.51.100.34;protocol:ssh",
             "ipv4:198.51.100.34;protocol:ftp",
             "ipv4:203.0.113.45;port:20"
           1
         }
       }
       HTTP/1.1 200 OK
       Content-Length: 402
       Content-Type: application/alto-endpointcost+json
       {
         "meta" : {
           "cost-type": {"cost-mode" : "ordinal",
                         "cost-metric" : "routingcost"
           }
         },
         "endpoint-cost-map" : {
           "ipv4:192.0.2.2": {
             "domainname:www.a.com"
                                                : 1,
             "mac:01-23-45-67-89-AB"
                                                : 2,
             "ipv4:198.51.100.34;protocol:ssh" : 3,
             "ipv4:198.51.100.34;protocol:ftp" : 4,
             "ipv4:203.0.113.45;port:20"
                                            : 5
           }
         }
       }
```

6. IANA Considerations

This document does not define any new media type or introduce any new IANA consideration.

7. Privacy And Security Considerations

This document does not introduce any privacy or security issue not already present in the ALTO protocol.

8. Normative References

- [RFC0826] Plummer, D., "Ethernet Address Resolution Protocol: Or Converting Network Protocol Addresses to 48.bit Ethernet Address for Transmission on Ethernet Hardware", STD 37, <u>RFC 826</u>, DOI 10.17487/RFC0826, November 1982, <<u>http://www.rfc-editor.org/info/rfc826</u>>.
- [RFC0903] Finlayson, R., Mann, T., Mogul, J., and M. Theimer, "A Reverse Address Resolution Protocol", STD 38, <u>RFC 903</u>, DOI 10.17487/RFC0903, June 1984, <<u>http://www.rfc-editor.org/info/rfc903</u>>.
- [RFC1034] Mockapetris, P., "Domain names concepts and facilities", STD 13, <u>RFC 1034</u>, DOI 10.17487/RFC1034, November 1987, <<u>http://www.rfc-editor.org/info/rfc1034</u>>.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", <u>BCP 14</u>, <u>RFC 2119</u>, DOI 10.17487/ <u>RFC2119</u>, March 1997, <<u>http://www.rfc-editor.org/info/rfc2119</u>>.
- [RFC3122] Conta, A., "Extensions to IPv6 Neighbor Discovery for Inverse Discovery Specification", <u>RFC 3122</u>, DOI 10.17487/ <u>RFC3122</u>, June 2001, <<u>http://www.rfc-editor.org/info/rfc3122</u>>.
- [RFC3596] Thomson, S., Huitema, C., Ksinant, V., and M. Souissi, "DNS Extensions to Support IP Version 6", <u>RFC 3596</u>, DOI 10.17487/RFC3596, October 2003, <<u>http://www.rfc-editor.org/info/rfc3596</u>>.
- [RFC4861] Narten, T., Nordmark, E., Simpson, W., and H. Soliman, "Neighbor Discovery for IP version 6 (IPv6)", <u>RFC 4861</u>, DOI 10.17487/RFC4861, September 2007, <<u>http://www.rfc-editor.org/info/rfc4861</u>>.

ECS Flow

- [RFC7042] Eastlake 3rd, D. and J. Abley, "IANA Considerations and IETF Protocol and Documentation Usage for IEEE 802 Parameters", <u>BCP 141</u>, <u>RFC 7042</u>, DOI 10.17487/RFC7042, October 2013, <<u>http://www.rfc-editor.org/info/rfc7042</u>>.
- [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", <u>RFC 7285</u>, DOI 10.17487/RFC7285, September 2014, <<u>http://www.rfc-editor.org/info/rfc7285</u>>.

Authors' Addresses

Junzhuo Austin Wang Tongji University 4800 Cao'an Road, Jiading District Shanghai China

Email: wangjunzhuo200@gmail.com

Qiao Xiang Tongji/Yale University 4800 Cao'an Road, Jiading District Shanghai China

Email: xiangq27@gmail.com