ALTO   L. M. Contreras, Ed.
Internet-Draft   Telefonica
Intended status: Informational   S. Randriamasy
Expires: 14 September 2023   Nokia Bell Labs
                                                                X. Liu
                                                        IBM Corporation
                                                         13 March 2023

ALTO extensions for handling Service Functions
draft-lcsr-alto-service-functions-02

Abstract

   This document proposes the usage of ALTO (and its extensions) to
   provide information about service functions to clients (e.g.,
   external systems) that could consume such information for decisions
   requiring network information (service composition, traffic steering
   to service chains, etc).

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 14 September 2023.

Copyright Notice

Table of Contents

[1](#). Introduction

   Network services are commonly formed by means of the concatenation of
   several atomic service functions (SF), resulting in a connected graph
   of functions. Those functions can be topologically spread across the
   network. In addition to that, there will be typically more than one
   instance of any particular atomic service function in the network for
   different purposes such as load balancing, redundancy, traffic
   optimization, etc.

During the definition phase of a network service there will be a
process for defining the type of service functions needed to
implement a given network service, as well as the way in which they
should be connected to steer the traffic flows through them. The
type of a SF can be for instance a User Plane Function (UPF) of the
mobile packet core, a cache of a Content Delivery Network (CDN), etc.
Thus when having multiple instances of a function (i.e., multiple
UPFs or multiple caches as in the example before), a decision process
should be in place to determine the particular instances for each
type of service function (i.e., what instance of UPF and CDN cache)
that will be part of the realization of the network service, that we
refer to as network service instance in this document.

Information on the SFs and their instances is available from entities
such as Service Orchestrators. Examples of such information are SF
type, service instance ID, SF locator, etc. Information about a
network service instance can be more comprehensively defined in terms
of performace and resources efficiency if this information is
complemented with network capabilities. This draft proposes to
complement information available on SF or their instances with
information on the paths that interconnect them.

At this point of network service realization, having timely
information of the characteristics of the interconnection paths among
SFs can be crucial. Aspects such as number of hops, associated
performance metrics, etc., can enrich (or even determine) the
decision of which instances of the service function consider as final
election.

This document proposes the usage of ALTO [RFC7285] and its extensions
to provide information about service functions or their
interconnection paths to clients (e.g., external systems) that could
consume such information for decisions requiring network information.

2. Service Function information

2.1. Cases relevant for exposing information related to Service Functions

Several initiatives in IETF deal with the interconnection of service
functions.

[RFC7665] defines the Service Function Chain (SFC) architecture.
There, the traffic is steered through SFC domains with the objective
of making the flow passing through a number of service functions to
run a service. When entering the domain, the traffic is classified
and assigned to an SFC Path. Specific information is added to the
packet flows within the domain, being this SFC encapsulation

containing metadata and contextual information useful for the
processing of the flows by the service functions and other components
in the architecture. In all this process, there is no explicit
identification of the service function to direct the traffic to, as
it is implicit in the definition of a specific SFC Path.

Similarly, in [RFC8986] the Segment Identifiers of SRv6 structures
the 128 bits of the IPv6 address in the form LOC:FUNCT:ARG. LOC is
the locator used to route a packet to the endpoint and encoded in the
L most significant bits of the Segment Identifier (SID). FUNCT
represents a Function ID and uses F bits. ARG represents optional
parameters to be interpreted by the function, and uses A bits.
Furthermore, [I-D.ietf-spring-sr-service-programming] defines data
plane functionalities required to implement service segments, in a
similar way as [RFC7665] for SFC.

Moreover, [I-D.ietf-teas-sf-aware-topo-model] defines a YANG data
model able to integrate both network topology and service location on
the same traffic engineering topology. In this model, the service
functions are represented by service-function-id and sf-connection-
point-id.

Finally, [I-D.ldbc-cats-framework] proposes a framework in which the
metrics associated to different service instances of the same service
function type are used for taking traffic steering decisions using
overlay connections. The metrics considered in such decision are
expected to be a combination of networking and computing metrics.

In all these previous cases, the information relative to the service
functions can be limited, if present. Richer information could be
needed for an integration between the control systems responsible for
the service operation and the control systems responsible for the
network actions that could optimize the delivery of services relying
on network information (that is, acting in an integrated fashion).

For instance, taking as example OpenStack [OpenStack], a network
service relies on descriptors providing information about Virtual
Deployment Units (VDUs), Connection Points (CPs) and Virtual Links
(VLs).

A VDU describes the properties of the virtual construct that hosts
the service function. Important information is the function
identifier and its type. The CPs contain the IP and MAC addresses
for such function, showing the binding as well between a VDU and a
VL. Finally, the VL identifies the connectivity between VDUs.

The level of information is not the same in all the solutions
overviewed, however a solution like ALTO could help to reconcile all
these different approaches by mapping and matching information on the
service and the network planes.

2.2. Retrieval of Service Function information

Different options can be considered, in some cases being
complementary and in some other cases being actual alternatives.

One option is the retrieval of information as provided by Service
Function Orchestration systems, as described by [OpenStack] or
similar solutions.

In addition to that, or as an alternative, different routing protocol
extensions can provide information about existing Service Functions
in the network. For instance [RFC9015] can advertise using BGP the
Service Function Type and Service Function Instances in a network.
Furthermore, [I-D.xu-lsr-isis-service-function-adv] and
[I-D.xu-lsr-ospf-service-function-adv] can advertise at IGP level the
Service Function Identifier as unique identifier representing a
service function within a domain.

3. Usage of ALTO for retrieving information relative to service functions

ALTO can expose combined information of the service overlay together
with the underlying network characteristics. This section details
the potential usage of ALTO in this respect.

3.1. Information of interest

There can be several kinds of ALTO exposure information in response
to client requests that can be taken into consideration. Some
examples are listed below:

* Path characteristics, from a PID, to any instance of a service
   function type.

* Path characteristics, from a PID, to a specific instance of a
   service function type.

* Path characteristics among any instance of a service function type
   X to any other instance of a service function type Y.

* Path characteristics among a specific instance of a service
   function type X to any other instance of a service function type
   Y.

    * Path characteristics, from a PID, to a chain of service functions.

    * Path characteristics, from a PID, to a chain of specific instances
       of service functions.

    Other type of requests could be further identified. The information
    to be exposed can be a combination of network and service function
    related characteristics.

    An ALTO server could be able to provide information for a limited set
    of requests. Thus, some indication of the possible requests to be
    served should be in place when interacting with the client.

3.2. ALTO mechanisms to support the requests about service functions

    ALTO can determine the path characteristics between two endpoints as
    described in [RFC7285]. ALTO also can provide the view of chain of
    functions by leveraging on the path vector concept developed in
    [RFC9275], where the endpoints considered represent service
    functions.

    [RFC9275] introduces the concept of Abstract Network Element (ANE) to
    specify a component or an aggregation of components sharing some
    characteristics in a network. Furthermore, [RFC9240] generalizes the
    concept of endpoint properties to entity properties, where entities
    may be defined in semantic domains such as as IPv4 or IPv6, or PIDs
    or ANEs.

    This draft makes use of these capabilities to support the retrieval
    of information relative to service functions.

4. ALTO architecture for service function information retrieval

    The following logical architecture defines the usage of ALTO for the
    retrieval of information about service functions or interconnection
    of service functions.

    Figure 1 represents the information exchange of an ALTO Server for
    providing information about service functions to ALTO clients (e.g.,
    external systems) that require combined information of network and
    service functions.

```
                        +--------+  Topological  +---------+
                        |        |  Information  |  e.g.,  |
                        |        |<------------->|  BGP,   |
              ALTO      |        |               | BGP-LS  |
    +--------+ protocol |        |               |         |
    | Client |<-------->|  ALTO  |               +---------+
    +--------+          | Server |
                        |        | Service Function+---------+
                        |        |  Information  |  e.g.,  |
                        |        |<------------->| Service |
                        |        |               | Function|
                        |        |               |Orchestr.|
                        +--------+               +---------+
```
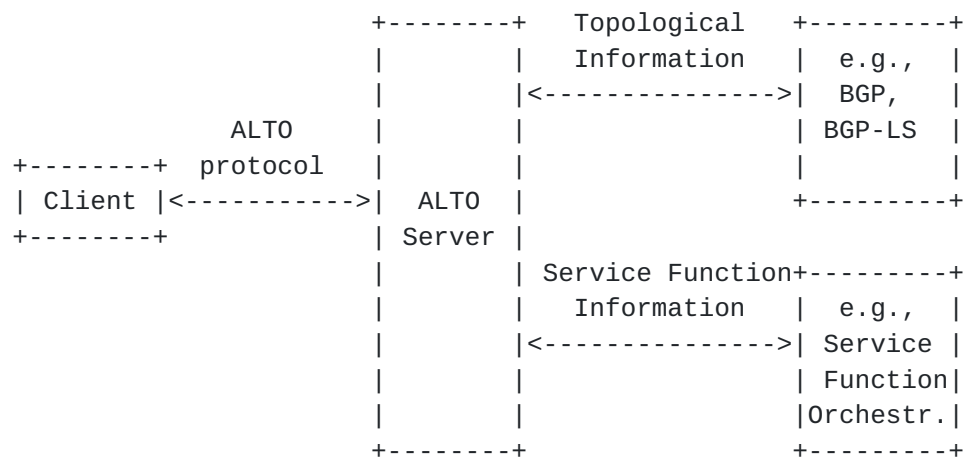
Figure 1. Information exchange

The network topological information will be complemented with
information relative to the service functions as provided by the
orchestration system managing and controlling that part.

The ALTO server will integrate the information of the service
functions based on some parameters, such as the IP address of the
service functions.

5. Proposed ALTO extensions

As proposed extension to existing ALTO specifications, the following
aspects are considered:

*   Extension to ALTO protocol to allow ALTO clients to express
    detailed requests in line with the information of interest
    described in Section 3.1.

*   Extensions to ALTO in order to combine and expose both service and
    network information, in line with the architecture depicted in
    Section 3.3. These extensions can involve particularizations of
    both [RFC9275] and [RFC9240].

*   Identification of mechanisms to be used by ALTO for collecting
    service functions' information.

Further extensions could be required.

Next iterations of this draft will further analyze the gap between
existing ALTO features and requirements to support the provisioning
of infrastructure information needed to perform efficient SF
management.

6. Security Considerations

To be provided.

7. Informative References

[I-D.ietf-spring-sr-service-programming]
          Clad, F., Xu, X., Filsfils, C., Bernier, D., Li, C.,
          Decraene, B., Ma, S., Yadlapalli, C., Henderickx, W., and
          S. Salsano, "Service Programming with Segment Routing",
          Work in Progress, Internet-Draft, draft-ietf-spring-sr-
          service-programming-07, 15 February 2023,
          <https://datatracker.ietf.org/doc/html/draft-ietf-spring-
          sr-service-programming-07>.

[I-D.ietf-teas-sf-aware-topo-model]
          Bryskin, I., Liu, X., Lee, Y., Guichard, J., Contreras, L.
          M., Ceccarelli, D., Tantsura, J., and D. Shytyi, "SF Aware
          TE Topology YANG Model", Work in Progress, Internet-Draft,
          draft-ietf-teas-sf-aware-topo-model-11, 12 March 2023,
          <https://datatracker.ietf.org/doc/html/draft-ietf-teas-sf-
          aware-topo-model-11>.

[I-D.ldbc-cats-framework]
          Li, C., Du, Z., Boucadair, M., Contreras, L. M., Drake,
          J., Huang, D., and G. S. Mishra, "A Framework for
          Computing-Aware Traffic Steering (CATS)", Work in
          Progress, Internet-Draft, draft-ldbc-cats-framework-01, 10
          March 2023, <https://datatracker.ietf.org/doc/html/draft-
          ldbc-cats-framework-01>.

[I-D.xu-lsr-isis-service-function-adv]
          Xu, X., Huang, H., Shah, H. C., and L. M. Contreras,
          "Advertising Service Functions Using IS-IS", Work in
          Progress, Internet-Draft, draft-xu-lsr-isis-service-
          function-adv-00, 9 March 2023,
          <https://datatracker.ietf.org/doc/html/draft-xu-lsr-isis-
          service-function-adv-00>.

[I-D.xu-lsr-ospf-service-function-adv]
          Xu, X., Huang, H., Shah, H. C., and L. M. Contreras,
          "Advertising Service Functions Using OSPF", Work in
          Progress, Internet-Draft, draft-xu-lsr-ospf-service-
          function-adv-00, 9 March 2023,
          <https://datatracker.ietf.org/doc/html/draft-xu-lsr-ospf-
          service-function-adv-00>.

[OpenStack]
          "VNF Descriptor (VNFD) Template Guide,
          https://docs.openstack.org/tacker/latest/contributor/
          vnfd_template_description.html", n.d..

[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>.

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

[RFC8986] Filsfils, C., Ed., Camarillo, P., Ed., Leddy, J., Voyer,
          D., Matsushima, S., and Z. Li, "Segment Routing over IPv6
          (SRv6) Network Programming", RFC 8986,
          DOI 10.17487/RFC8986, February 2021,
          <https://www.rfc-editor.org/info/rfc8986>.

[RFC9015] Farrel, A., Drake, J., Rosen, E., Uttaro, J., and L.
          Jalil, "BGP Control Plane for the Network Service Header
          in Service Function Chaining", RFC 9015,
          DOI 10.17487/RFC9015, June 2021,
          <https://www.rfc-editor.org/info/rfc9015>.

[RFC9240] Roome, W., Randriamasy, S., Yang, Y., Zhang, J., and K.
          Gao, "An Extension for Application-Layer Traffic
          Optimization (ALTO): Entity Property Maps", RFC 9240,
          DOI 10.17487/RFC9240, July 2022,
          <https://www.rfc-editor.org/info/rfc9240>.

[RFC9275] Gao, K., Lee, Y., Randriamasy, S., Yang, Y., and J. Zhang,
          "An Extension for Application-Layer Traffic Optimization
          (ALTO): Path Vector", RFC 9275, DOI 10.17487/RFC9275,
          September 2022, <https://www.rfc-editor.org/info/rfc9275>.

Authors' Addresses

   Luis M. Contreras (editor)
   Telefonica
   Ronda de la Comunicacion, s/n
   28050 Madrid
   Spain
   Email: luismiguel.contrerasmurillo@telefonica.com
   URI: http://lmcontreras.com

Sabine Randriamasy
Nokia Bell Labs
Email: sabine.randriamasy@nokia-bell-labs.com

Xufeng Liu
IBM Corporation
Email: xufeng.liu.ietf@gmail.com