     Operating the Network Service Header (NSH) with Next Protocol "None"

Abstract

   This document describes a network that supports Service Function
   Chaining (SFC) using the Network Service Header (NSH) with no payload
   data and carrying only metadata. This is achieved by defining a new
   NSH "Next Protocol" type value of "None".

   This document illustrates some of the functions that may be achieved
   or enhanced by this mechanism, but it does not provide an exhaustive
   list of use cases, nor is it intended to be definitive about the
   functions it describes. It is expected that other documents will
   describe specific use cases in more detail and will define the
   protocol mechanics for each use case.

Status of This Memo

   This is an Internet Standards Track document.

   This document is a product of the Internet Engineering Task Force
   (IETF). It represents the consensus of the IETF community. It has
   received public review and has been approved for publication by the
   Internet Engineering Steering Group (IESG). Further information on
   Internet Standards is available in Section 2 of RFC 7841.

   Information about the current status of this document, any errata,
   and how to provide feedback on it may be obtained at
   https://www.rfc-editor.org/info/rfc8393.

Copyright Notice

Table of Contents

1. Introduction

   An architecture for Service Function Chaining (SFC) is presented in
   [RFC7665]. That architecture enables packets to be forwarded along
   Service Function Paths (SFPs) to pass through various Service
   Functions (SFs) that act on the packets. Each packet is encapsulated
   with a Network Service Header (NSH) [RFC8300] that identifies the SFP
   that the packet travels along (by means of a Service Path Identifier
   -- SPI) and the hop (i.e., the next SF to be executed) along the SFP
   that the packet has reached (by means of a Service Index -- SI). The
   SPI and SI are fields encoded in the NSH.

   Packets are classified at the SFC network ingress boundaries by
   classifiers (Section 4.4 of [RFC7665]) and have an NSH applied to
   them. Such packets are forwarded between Service Function Forwarders
   (SFFs) using tunnels across the underlay network, and each SFF may
   hand the packet off to one or more Service Function Instances (SFIs)
   according to the definition of the SFP.

   The SFC classifier or any SFC-aware SFI may wish to share information
   (possibly state information) about the SFP, the traffic flow, or a
   specific packet, and they may do this by adding metadata to packets
   as part of the NSH. Metadata may be used to enhance or enable the
   function performed by SFC-aware SFs, may enable coordination and data
   exchange between SFIs, or may be used to assist a network operator in
   the diagnosis and monitoring of an SFP. The nature of metadata to be
   supplied and consumed is implementation- and deployment-specific.

   This document defines a mechanism for metadata to be carried on an
   SFP without the need for payload data. This mechanism enables
   diagnosis and monitoring of SFPs, and coordination between SFC-aware
   SFIs. The mechanism can be applied without the need for traffic to
   be flowing; if traffic is flowing, it can be applied without the need
   to insert what might be substantial amounts of metadata into data
   packets (an operation that may be costly in some hardware).

   This document describes how this function is achieved through the use
   of a new value for the NSH "Next Protocol" field to indicate "None".
   Like any NSH packets, such packets are contained within the
   SFC-enabled domain.

   This document illustrates some of the functions that may be achieved
   or enhanced by this mechanism, but it does not provide an exhaustive
   list of use cases, nor is it intended to be definitive about the
   functions it describes (see Section 6).

   This document uses the terms defined in [RFC7665] and [RFC8300].

2. Requirements Language

   The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
   "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and
   "OPTIONAL" in this document are to be interpreted as described in
   BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all
   capitals, as shown here.

3. The Network Service Header

   The NSH includes a field called "Next Protocol" that is used to
   indicate the nature of the payload data that follows the NSH. The
   field can be used by any component that processes the NSH (for
   example, to understand how to interpret and parse the payload) and by
   nodes at the end of the SFP that remove the NSH and forward the
   payload data.

3.1. Next Protocol "None"

   This document defines a new value (0x00) for the "Next Protocol"
   field to indicate that the next protocol is "None", which means that
   there is no user/payload data following the NSH.

   When the next protocol is "None", the rest of the NSH still has
   meaning; in particular, the metadata carried in the NSH may still be
   present. It is not intended that a packet with next protocol set to
   "None" be sent with no metadata (see Section 4). Thus, an SFC-aware
   node SHOULD NOT create a packet with "Next Protocol" set to "None",
   Metadata Type set to 0x2, and with an NSH Length of 0x2.

4. Processing Rules

   A packet with no payload data may be inserted at the head end of an
   SFP (such as at a classifier) and may be easily forwarded by an SFF
   or SFI on the SFP using the processing rules defined in [RFC8300].

   A packet with no payload may also be generated by an SFC-aware SFI as
   a result of processing an incoming packet (i.e., triggered by a
   condition arising from processing a normal NSH packet with a
   payload). In such cases, the SPI/SI can be inherited from the
   original packet or can be set according to information supplied in
   one of three ways:

   o through the control plane,

   o through the management plane, or

   o through information carried in the metadata of the data packet.

This document does not further specify the triggers to generate an
NSH packet with a "Next Protocol" set to "None".

An SFC-aware node wishing to send metadata without a data packet
(i.e., a node that conforms to this specification):

o MUST create a packet carrying an NSH and the desired metadata.

o MUST set the "Next Protocol" field to 0x00.

o SHOULD ensure that there are no bytes following the end of the NSH
   (i.e., that there is no payload data).

o MUST encapsulate and send the packet as normal for tunneling to
   the next hop on the SFP as would be done for any NSH packet (i.e.,
   for a data packet following the SFP).

A transit node (SFF, SFI, or classifier) that conforms to this
specification and that receives a packet with "Next Protocol"
indicating "None" MUST NOT attempt to parse or process beyond the end
of the NSH, but it SHOULD process the NSH and the metadata as normal.
Processing for nodes that do not support "Next Protocol" set to
"None" is described in Section 5. Note, however, that an
intermediate node that is instructed to strip all metadata from
packets will create a packet with an NSH but no metadata and no
payload. Such packets SHOULD NOT continue to be forwarded along the
SFP.

A node that is the egress of an SFP would normally strip the NSH and
forward the payload according to the setting of the "Next Protocol"
field. Such nodes MUST NOT forward packets with "Next Protocol"
indicating "None" even if there are some bytes after the NSH.

In deployments where use of next protocol "None" is not desired,
administrators SHOULD instruct SFC-aware nodes to not create such
packets and to discard packets with next protocol "None".

5. Backward Compatibility

SFC-aware nodes that do not understand the meaning of a value
contained in the "Next Protocol" field of the NSH are unable to parse
the payload. Such nodes silently drop packets with unknown "Next
Protocol" values unless explicitly configured to forward them
(Section 2.2 of [RFC8300]).

This means that legacy SFC-aware nodes that are unaware of the
meaning of the "Next Protocol" value "None" will act as follows:

o SFFs can be configured to forward the packets.

o SFC proxies will drop the packets.

o SFIs will most likely drop the packets.

o Classifiers (i.e., nodes performing reclassification) will most
   likely drop the packets.

SFC-aware nodes at the end of an SFP possibly forward packets with no
knowledge of the payload in a "pop and forward" form of processing
where the NSH is removed, the packet is simply put on an interface,
and the payload protocol is known a priori (or assumed). It is a
general processing rule for all packet forwarding engines that they
should not attempt to send packets with zero length. Packets with
the NSH "Next Protocol" field set to "None" are expected to have zero
payload length and so should not be forwarded once the NSH has been
stripped. In any case, as noted in Section 4, SFC-aware nodes at the
end of an SFP do not forward packets with "Next Protocol" set to
"None".

6. Overview of Use Cases

6.1. Per-SFP Metadata

Per-SFP metadata is metadata that applies to an SFP and any data
packets on that SFP. It does not need to be transmitted with every
packet, but it can be installed at the points of consumption (such as
at SFIs) and applied to all packets on the SFP as they pass through
this point. It could be installed by inclusion in the NSH of a data
packet sent on the SFP by out-of-band control- or management-plane
mechanisms, or by separate metadata-only packets using "Next
Protocol" set to "None" as described in this document.

Per-SFP metadata-only packets may be sent along the path of an SFP
simply by setting the correct SPI in the NSH, and setting the SI to
the correct value for the hop of the SFP at which the metadata is to
be introduced. SFC-aware nodes (e.g., classifiers) will know the
correct SI values to be used from information supplied by the control
or management plane as is the case for NSH packets with payload data.

6.2. Per-Flow Metadata

   Per-flow metadata is metadata that applies to a subset of the packets
   on an SFP, such as packets matching a particular 5-tuple of source
   address, destination address, source port, destination port, and
   payload protocol. Also, this metadata does not need to be
   transmitted with every packet, but it can be installed at the SFIs on
   the SFP and applied to the packets that match the flow description.

   If there is just one flow on an SFP, then there is no difference
   between per-flow metadata and per-SFP metadata as described in
   Section 6.1.

   In normal processing, the flow to which per-flow metadata applies can
   be deduced by looking at the payload data in the context of the value
   of the "Next Protocol" field. However, when "Next Protocol"
   indicates "None", this cannot be done. In this case, the identity of
   the flow is carried in the metadata itself.

6.3. Coordination between SFC-Aware SFIs

   A pair of SFC-aware SFIs (adjacent or not) on an SFP may desire to
   coordinate state and may do this by sending information encoded in
   metadata.

   To do this using the mechanisms defined in this document:

   o There must be an SFP that passes through the two SFIs in the
      direction of sender to receiver.

   o The sender must know the correct SPI to use.

   o The sender must know the correct SI to use for the point at which
      it resides on the SFP.

   o Ideally, the receiver will know to remove the packet from the SFP
      and not forward it further as this might share metadata wider than
      desirable and would cause unnecessary packets in the network.
      Note, however, that continued forwarding of such packets would not
      be substantially harmful in its own right.

   Note that technically (according to the SFC architecture) the process
   of inserting a packet into an SFP is performed by a classifier.
   However, a classifier may be co-resident with an SFI so that an
   implementation of an SF may also be able to generate NSH packets as
   described in this document.

Note also that a system with SFIs that need to coordinate between
each other may be configured so that there is a specific, dedicated
SFP between those service functions that is used solely for this
purpose. Thus, such an SFI does not need to insert NSH packets onto
SFPs used to carry payload data, but it can use (and know the SPI of)
this special, dedicated SFP.

6.4. Operations, Administration, and Maintenance (OAM)

Requirements for Operations, Administration, and Maintenance (OAM) in
SFC networks are discussed in [SFC-OAM-FRAME]. The NSH definition in
[RFC8300] includes an O bit that indicates that the packet contains
OAM information.

If OAM information is carried in packets that also include payload
data, that information might be carried between the NSH and the
payload. Therefore, the mechanism defined in this document can also
be used to carry OAM information independent of payload data.

Sending OAM separate from (but interleaved with) packets that carry
payload data may have several advantages including:

o Sending OAM when there is no other traffic flowing.

o Sending OAM at predictable intervals.

o Measuring path qualities distinct from behavior of SFIs.

o Sending OAM without needing to rewrite payload data buffers.

o Keeping OAM processing components separate from other processing
   components.

Mechanisms for providing active OAM [RFC7799] in an SFC network have
been proposed [OAM-SFC]. This use case is not intended to define
another mechanism for active OAM, but it does illustrate a further
option for discussion by the working group.

6.5. Control-Plane and Management-Plane Uses

As described in Section 6.3, SFPs can be established specifically to
carry metadata-only packets. And as described in Section 6.1,
metadata-only packets can be sent down existing SFPs. This means
that metadata-only packets can be used to carry control-plane and
management-plane messages used to control and manage the SFC network.

In effect, SFPs can be established to serve as a Data Control Network
(DCN) or a Management Control Network (MCN). Further details of this
process are out of scope for this document, but it should be
understood that, just as for OAM, an essential feature of using a
control channel is that the various speakers are assigned identifiers
(i.e., addresses). In this case, those identifiers could be SPI/SI
pairs or could be IP addresses as used in the normal control and
management plane of the SFC network.

6.6. Non-applicable Use Cases

Per-packet metadata is metadata that applies specifically to a single
payload packet. It informs an SFI how to handle the payload packet
and does not apply to any other packet.

The mechanisms described in this document are not applicable to per-
packet metadata because, by definition, if the "Next Protocol"
indicates "None", then there is no packet following the NSH for the
metadata to be associated with.

7. Management and Congestion Control Considerations

The mechanisms described in this document allow SFC-aware nodes in an
SFC network to generate additional packets. These are not intended
to be sent frequently for any flow, but there is still a risk that
they might flood the network. For example, if an attempt is made to
use this mechanism for "per-packet metadata" (see Section 6.6) then
this might double the number of packets in the network. Similarly,
if this mechanism is used for a form of aliveness detection OAM that
requires very frequent test messages, then the number of additional
messages may be very high. Such additional messages risk causing
congestion in the network.

The underlay network (that is, the tunnels across the underlay
between SFC nodes) will not distinguish between data-carrying packets
and those packets with "Next Protocol" set to "None". All packets
will be treated the same and will need to fall within the
capabilities of the underlay network to process and forward packets.

Nodes in the SFC overlay network will need to perform special
processing on the additional packets according to their roles and
according to the application for the metadata. For example, an SFF
will likely only have to forward per-SFP metadata, while an SF will
need to extract it and process it as it would if the metadata was
carried in a packet with user data. On the other hand, metadata
might also be used to cause actions at all nodes (see Sections 6.3,
6.4, and 6.5) and could increase the processing load.

   In view of these potential issues, all implementations SHOULD
   implement rate limits on the generation of per-SFP packets with "Next
   Protocol" set to "None". Furthermore, these rate limits SHOULD be
   configurable and applied per SFP and per application so that one
   application on one SFP does not encumber a different application on
   this or a different SFP. When an implementation finds that it is
   unable to generate or send a packet, it SHOULD increment a counter
   that is accessible by the operator and MAY raise an alert (although
   such alerts SHOULD, themselves, be rate limited).

   Additionally, an SFC node needs to protect itself against another
   node in the network not applying suitable rate limits. Therefore,
   implementations SHOULD apply incoming rate limits for SFC packets
   with "Next Protocol" set to "None". Such rate limits MAY be
   application aware, per SFC or interface, and SHOULD be configurable,
   but implementations MAY be more subtle if they are aware of internal
   processing loads and have access to queues/buffers. In any case,
   when an implementation drops a received packet because of these rate
   limits, it SHOULD increment a counter that is accessible by the
   operator and MAY raise an alert (although such alerts SHOULD,
   themselves, be rate limited).

   Suitable default rate limits will restrict an SFC node to not send
   more than one packet with "Next Protocol" set to "None" per ten data
   packets on any flow in a unit of time equal to the end-to-end
   delivery time on the flow.

8. Security Considerations

   Metadata-only packets as enabled by this document provide a covert
   channel. However, this is only different from the metadata feature
   in the normal NSH in that it can be sent without the presence of a
   data flow.

   Metadata may, of course, contain sensitive data and may also contain
   information used to control the behavior of SFIs in the network. As
   such, this data needs to be protected according to its value and
   according to the perceived vulnerabilities of the network.
   Protection of metadata may be achieved by using encrypted transport
   between SFC entities or by encrypting the metadata in its own right,
   and by authenticating the sender of the metadata. The need to
   protect the metadata is not modified by this document and forms part
   of the NSH definition found in [RFC8300].

   The mechanism described in this document might be used to introduce
   packets into the SFC overlay network and might be used to
   illegitimately introduce false metadata to the nodes on an SFC.

   Therefore, measures SHOULD be taken to ensure authorization of
   sources of such packets, and tunneling of such packets into the
   network SHOULD be prevented.

   The amount of packets with "Next Protocol" set to "None" on an SFP
   SHOULD be rate limited at each point on the SFP to provide additional
   network security.

   Further discussion of NSH security is presented in [RFC8300].

9. IANA Considerations

   IANA maintains a registry called "Network Service Header (NSH)
   Parameters" with a sub-registry called "NSH Next Protocol". IANA has
   allocated a new value to the sub-registry as follows:

   Next Protocol | Description | Reference
   ---------------+---------------+-------------
   0x00 | None | RFC 8393

10. References

10.1. 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>.

   [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC
              2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174,
              May 2017, <https://www.rfc-editor.org/info/rfc8174>.

   [RFC8300] Quinn, P., Ed., Elzur, U., Ed., and C. Pignataro, Ed.,
              "Network Service Header (NSH)", RFC 8300,
              DOI 10.17487/RFC8300, January 2018,
              <https://www.rfc-editor.org/info/rfc8300>.

10.2. Informative References

   [OAM-SFC] Mirsky, G., Meng, W., Khasnabish, B., and C. Wang, "Multi-
              Layer Active OAM for Service Function Chains in Networks",
              Work in Progress, draft-wang-sfc-multi-layer-oam-10,
              September 2017.

   [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>.

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

   [SFC-OAM-FRAME]
              Aldrin, S., Pignataro, C., Kumar, N., Akiya, N., Krishnan,
              R., and A. Ghanwani, "Service Function Chaining (SFC)
              Operation, Administration and Maintenance (OAM)
              Framework", Work in Progress, draft-ietf-sfc-oam-
              framework-04, March 2018.

Acknowledgements

Contributors

   Lucy Yong
   Retired

Authors' Addresses

   Adrian Farrel
   Juniper Networks

   Email: afarrel@juniper.net

   John Drake
   Juniper Networks

   Email: jdrake@juniper.net